
Cours de lycée

Cassis
Avril 2022

Avant-propos

Ce document reprend le programme du lycée en introduisant constructions formelles et preuves rigoureuses. Il est destiné principalement à un élève de lycée (ou entrant au lycée) qui souhaite se familiariser avec les mathématiques du supérieur. L'objectif étant de suivre le programme, nous limiterons autant que possible les écarts au programme, mis à part pour la première partie (pré requis) qui est nécessaire à formaliser rigoureusement les mathématiques : les autres chapitres aborderont chacun un point essentiel des mathématiques du lycée (géométrie / analyse / algèbre) en détails. Notons tout de même l'ajout de deux chapitre d'approfondissement en algèbre, pour poursuivre plus loin que le programme dans cette branche peu développée au lycée.

Il y aura plusieurs exercices, leur objectif premier est de familiariser le lecteur avec les éléments détaillés dans les parties qui précèdent l'exercice, mais certains sont plus compliqués ou visent à approfondir une notion. Dans ce cas, ces exercices porteront une étoile (*). Nous conseillons de plus de ne faire les exercices étoilés qu'en deuxième lecture d'un chapitre, pour s'être d'abord assuré d'avoir maîtrisé ce dont il est question. Certains exercices sont aussi utilisés dans des preuves. Ce sont en général des exercices techniques, pour ne pas poser problème dans l'acceptation de la validité des preuves qui les utilisent.

Table des matières

I	Prérequis	1
1	Logique de base	1
1.1	La proposition, phrase mathématique	1
1.1.1	Les relations	1
1.1.2	La conjonction	2
1.1.3	La disjonction	2
1.1.4	La négation	3
1.1.5	L'implication	3
1.1.6	L'équivalence	4
1.1.7	La quantification universelle	4
1.1.8	La quantification existentielle	4
1.1.9	A propos de la quantification	5
1.2	Raisonnement et preuve	5
1.2.1	L'hypothèse	5
1.2.2	Introduction d'hypothèse	6
1.2.3	L'introduction de la conjonction	6
1.2.4	L'élimination de la conjonction	6
1.2.5	L'introduction de la disjonction	6
1.2.6	L'élimination de la disjonction, ou disjonction de cas	6
1.2.7	L'introduction de la négation	7
1.2.8	Le raisonnement par l'absurde	7
1.2.9	L'introduction de l'implication	7
1.2.10	L'élimination de l'implication, le modus ponens	7
1.2.11	Le raisonnement par contraposée	8
1.2.12	L'équivalence	8
1.2.13	L'introduction du quantificateur universel	8
1.2.14	L'élimination du quantificateur universel	8
1.2.15	L'introduction du quantificateur existentiel	8
1.2.16	L'élimination du quantificateur existentiel	8
1.2.17	Le raisonnement par récurrence	9
2	Théorie des ensembles	10
2.1	Que faire d'un ensemble ?	10
2.1.1	Appartenance et égalité	10
2.1.2	Description par extension.	10
2.1.3	Description par compréhension	11
2.1.4	Inclusion d'ensembles	11
2.1.5	Union d'ensembles	12
2.1.6	Intersection d'ensembles	13
2.1.7	Le complémentaire relatif	13
2.1.8	Différence symétrique	14
2.1.9	Produit cartésien d'ensembles	14
2.1.10	Ensemble des parties	15
2.1.11	Intervalles de nombres réels	16
2.1.12	Raisonnement sur des ensembles	16
2.2	Relations binaires et fonctionnelles	16
2.2.1	Relation binaire	17
2.2.2	Relation d'équivalence	18
2.2.3	Partition d'un ensemble	20
2.2.4	Relation d'ordre	22
2.2.5	Fonctions	24

2.2.6	Fonctions et ensembles	26
3	Combinatoire et dénombrement	27
3.1	Cardinal	27
3.1.1	Cardinal d'une union disjointe	27
3.1.2	Cardinal d'un produit cartésien	28
3.1.3	Ensemble de fonctions	28
3.1.4	Permutations et arrangements	29
3.2	Sur les coefficients binomiaux	30
3.2.1	Propriétés	30
II	Géométrie	33
4	Construction des vecteurs	33
4.1	Propriétés d'un vecteur	34
4.2	Addition et multiplication de vecteurs	34
4.3	Système de coordonnées	36
4.4	Déterminant de deux vecteurs	38
5	Angles et trigonométrie	40
5.1	Cercle trigonométrique	40
5.2	Propriétés des fonctions trigonométriques	41
6	Produit scalaire	43
6.1	Définitions	43
6.1.1	Identités de polarisation	43
6.2	Propriétés	45
6.3	Application à un cercle	46
III	Algèbre	47
7	Équations et inégalités	47
7.1	Équation du premier degré	47
7.2	Système d'équations du premier degré	48
7.3	Inéquations	49
7.4	Résolution d'équation du second degré	49
8	Nombres complexes	52
8.1	Construction des nombres complexes	52
8.2	Le plan complexe	53
8.2.1	La conjugaison complexe	54
8.2.2	Partie réelle et imaginaire	55
8.2.3	Notation exponentielle	56
8.3	Les équations de degré 2	57
9	Suites : des éléments algébriques	58
9.1	Définition d'une suite	58
9.2	Sommes et produits	59
9.3	Suites particulières	61
9.3.1	Suite arithmétique	61
9.3.2	Suite géométrique	61
9.3.3	Somme géométrique	62
9.3.4	Suite arithmético-géométrique	62

10 Approfondissement : Polynômes à une indéterminée	64
10.1 Construction de l'ensemble des polynômes	64
10.2 Les polynômes forment un anneau euclidien	66
10.3 Polynôme scindé	66
11 Approfondissement : Un peu d'algèbre générale	68
11.1 Loi de composition interne, monoïde, groupe	68
11.2 Sous-groupe, morphisme de groupe, noyau et image	69
11.3 Exemple : des petits groupes	71
IV Analyse	73
12 Topologie des réels	73
12.1 Axiomes des l'ensemble des réels	73
12.2 Ensembles ouverts et voisinages	74
12.3 Valeur absolue et distance	75
12.4 Limite et caractérisation séquentielle	77
12.4.1 Calcul de limite de suites	79
12.5 Fermés et segments	82
12.5.1 Sur les suites extraites	84
13 Fonctions continues et limites	86
13.1 Limite d'une fonction	86
13.2 Limite à gauche, à droite	88
13.3 Continuité d'une fonction	88
13.4 Continuité uniforme	89
13.5 Théorèmes fondamentaux sur les fonctions continues	90
14 Dérivabilité	92
14.1 Nombre dérivé	92
14.2 Fonction dérivée	95
14.2.1 Opérations sur les dérivées	96
14.2.2 Fonction réciproque	99
14.3 Théorèmes sur la dérivation	100
15 Intégration de fonctions continues par morceaux sur un segment	102
15.1 Fonctions en escalier et continues par morceaux	102
15.2 Intégrale de Riemann	103
15.3 Propriétés des intégrales	104
15.4 Calcul d'intégrale	106
16 Généralités sur les fonctions	108
16.1 Logarithme	108
16.2 Exponentielle	110
16.3 Cosinus et sinus	111

Première partie

Prérequis

1 Logique de base

Nous allons ici voir quelques éléments de logique élémentaire pour permettre au lecteur de comprendre les exigences de rigueur des mathématiques. Cela permet aussi, bien sûr, de comprendre comment rédiger soi-même une preuve, par exemple. Nous commencerons par définir les termes que nous utilisons (proposition, implication...) puis nous détaillerons différents types de raisonnement.

1.1 La proposition, phrase mathématique

Pour commencer, tout langage se compose d'une grammaire, c'est-à-dire une façon de structurer ses phrases. Nous allons donc explorer la façon dont on construit une phrase mathématique.

Avant d'aller plus loin, une précision s'impose : il faudra distinguer deux styles dans la rédaction mathématique que sont le style formel et le style informel. Les deux doivent traduire une rigueur, mais le style formel est composé de symboles purement mathématiques, tandis que le style informel est une phrase écrite en toutes lettres. Un exemple serait

« pour tout réel x , il existe un réel y tel que $x + y = 0$ »

pour le style informel, dont l'équivalent formel est

$$\forall x \in \mathbb{R}, \exists y \in \mathbb{R}, x + y = 0$$

qui traduit le même sens mais n'est qu'une succession de symboles.

Nous verrons en détail ce que signifie chaque symbole, mais un point important est que l'on considère en permanence que tout ce qui peut se dire en mathématique peut s'exprimer à travers ces symboles de façon plus ou moins alambiquée. Cela ne veut pas dire que l'on souhaite recourir en permanence au style formel, loin s'en faut. Simplement, pour une étude des raisonnements et de la logique en général, il est préférable de réfléchir en utilisant le style formel, puisqu'il est bien plus facile à décortiquer.

Une phrase mathématique s'appelle une proposition et peut être soit vraie, soit fausse (c'est le principe qu'on appelle le tiers exclu). C'est, de plus, une phrase portant sur des objets mathématiques (des nombres, des fonctions, des ensembles...) Les mathématiciens s'intéressent principalement aux phrases vraies, et il convient donc d'avoir un moyen de vérifier qu'une phrase donnée est vraie : c'est ce que l'on appelle une preuve en mathématique. Nous allons donc découvrir quelles sont les briques de base pour écrire une proposition, et quel sens ont ces briques.

1.1.1 Les relations

Le premier élément constitutif d'une proposition est une relation. Une relation exprime une valeur de vérité sur le lien entre des objets mathématiques. La relation la plus utilisée et la plus évidente est l'égalité, notée $=$, qui exprime que deux objets sont en fait le même objet.

Exemple 1.1. La proposition

$$1 + 1 = 2$$

exprime que l'objet $1 + 1$ est en fait l'objet 2.

Il existe bien d'autres relations, comme par exemple le parallélisme, portant sur les droites du plan, ou la similitude portant sur les triangles (deux triangles sont semblables lorsqu'on peut tourner et agrandir / rétrécir le premier triangle pour obtenir le deuxième). Enfin, précisons qu'une relation peut ne comporter qu'un objet, dont elle exprime une propriété : on appelle souvent ce genre de relation paramétrée par un seul objet des prédicats, comme le prédicat « être égal à 3 » qui peut porter sur 2 (auquel cas la proposition correspondante est fausse) ou sur $1 + 2$ (auquel cas la proposition correspondante est vraie).

Les relations sont les briques élémentaires des propositions, car les autres éléments constitutifs des propositions ne peuvent que relier d'autres briques plus élémentaires. Le point essentiel d'une proposition est donc toujours une relation entre objets mathématiques.

1.1.2 La conjonction

Si l'on possède deux propositions P_1 et P_2 , alors on peut construire une nouvelle proposition appelée la conjonction de P_1 et P_2 , notée $P_1 \wedge P_2$. Cette proposition traduit que P_1 et P_2 sont toutes les deux vraies. On traduit en général la conjonction par « et » comme dans « il fait beau et je sors » pour dire à la fois que je sors, et qu'il fait beau.

Exemple 1.2. Si l'on veut exprimer que $1 + 1$ vaut 2 et que $2 + 2$ vaut 4, on peut écrire

$$(1 + 1 = 2) \wedge (2 + 2 = 4)$$

La conjonction est ce qu'on appelle un connecteur logique : elle permet de combiner plusieurs propositions en une seule nouvelle proposition.

Nous mettons des parenthèses pour rendre plus lisibles les séparations entre les relations et les connecteurs logiques.

Nous allons présenter aussi ce qu'on appelle une table de vérité, qui est un tableau montrant quand une formule est vraie en fonction de la vérité de ses parties. On note 1 pour dire « Vrai » et 0 pour dire « Faux ».

A	B	$A \wedge B$
0	0	0
0	1	0
1	0	0
1	1	1

TABLE 1 – Table de vérité de \wedge

1.1.3 La disjonction

En reprenant nos propositions P_1 et P_2 , on peut aussi construire la proposition $P_1 \vee P_2$, appelée la disjonction de P_1 et de P_2 , qui signifie qu'au moins l'une des deux propositions entre P_1 et P_2 est vraie. On traduit la disjonction par « ou », mais il convient de faire attention : le « ou » en mathématiques est dit inclusif, par défaut, ce qui signifie que l'on peut avoir à la fois P_1 et P_2 , là où le langage populaire utilise plutôt un « ou » exclusif (par exemple dans « tu veux un fromage ou un dessert ? »), la traduction n'est donc pas exactement identique entre les mathématiques et le langage populaire.

Exemple 1.3. Imaginons que x est un nombre entier. On note $\text{pair}(x)$ la relation « x est pair » et $\text{impair}(x)$ la relation « x est impair ». Alors la proposition suivante est vraie :

$$(\text{pair}(x)) \vee (\text{impair}(x))$$

La disjonction est aussi un connecteur logique, très semblable à la conjonction. Nous allons définir, par convention (qui est donc une règle arbitraire mais utile pour mieux se comprendre), que la conjonction est prioritaire sur la disjonction, ce qui signifie que

$$P \vee Q \wedge R$$

doit se lire

$$P \vee (Q \wedge R)$$

ce qui nous permet d'écrire moins de parenthèses.

A	B	$A \vee B$
0	0	0
0	1	1
1	0	1
1	1	1

TABLE 2 – Table de vérité de \vee

1.1.4 La négation

Nous ajoutons aussi un connecteur logique unaire, ce qui signifie qu'il permet de construire une nouvelle proposition à partir d'une seule proposition. La négation exprime à partir d'une proposition P la proposition « P est fausse », et se note $\neg(P)$. Elle est vraie quand P est fausse et inversement.

Exemple 1.4. Plutôt que d'écrire que $0 = 1$ est faux, un mathématicien préférera écrire la proposition vraie suivante :

$$\neg(0 = 1)$$

Dans le cas de l'égalité, $\neg(a = b)$ a aussi une notation standard, qui est $a \neq b$ (plus simple et rapide à écrire). Il existe beaucoup de symboles dénotant des relations pour lesquels barrer le symbole traduit la négation de la relation, comme $\not\subseteq$ pour la relation \subseteq .

A	$\neg A$
0	1
1	0

TABLE 3 – Table de vérité de \neg

1.1.5 L'implication

Il nous arrive souvent, quand on réfléchit, de déterminer des conséquences logiques entre les événements. En mathématiques aussi, nous pouvons écrire des formes de conséquences logiques à l'aide de l'implication. Si P_1 et P_2 sont des proposition, on écrit l'implication de P_1 sur P_2 par $P_1 \implies P_2$, que l'on lit « P_1 implique P_2 » ou encore « si P_1 alors P_2 ». Ce connecteur logique traduit le fait que lorsque la première partie (appelée prémisses) P_1 est vraie, la deuxième partie (appelée conclusion) P_2 est forcément vraie. Remarquons que par cette définition, lorsque P_1 est fausse, $P_1 \implies P_2$ est vraie peu importe la vérité de P_2 .

Exemple 1.5. Soit un triangle (ABC) , on note $\text{rectangle}(ABC)$ la relation « (ABC) est rectangle en B ». On peut alors écrire la proposition (qui est le théorème de Pythagore) :

$$\text{rectangle}(ABC) \implies AC^2 = AB^2 + BC^2$$

A	B	$A \implies B$
0	0	1
0	1	1
1	0	0
1	1	1

TABLE 4 – Table de vérité de \implies

1.1.6 L'équivalence

L'équivalence logique entre deux proposition peut s'écrire directement à partir de l'implication, mais elle traduit un lien plus fort. Si l'implication $P_1 \implies P_2$ traduit que P_2 découle toujours de P_1 , leur équivalence logique, notée $P_1 \iff P_2$ et que l'on lit « P_1 si et seulement si P_2 » traduit que les deux découlent l'un de l'autre, et dont que chaque fois qu'une des deux proposition est vraie, l'autre aussi (et chaque fois qu'une des deux propositions est fausse, l'autre aussi). Si l'on s'intéresse dans un premier temps à une proposition de la forme $A \implies B$, on appelle en générale la réciproque de cette proposition la proposition $B \implies A$.

Exemple 1.6. Il se trouve que le théorème de Pythagore est une équivalence : sa réciproque est vraie aussi, d'où :

$$\text{rectangle}(ABC) \iff AC^2 = AB^2 + BC^2$$

A	B	$A \iff B$
0	0	1
0	1	0
1	0	0
1	1	1

TABLE 5 – Table de vérité de \iff

Nous allons maintenant voir des briques dont l'utilisation est plus complexe mais qui sont très importantes : ce sont les quantificateurs.

Supposons qu'on ait une proposition P dans laquelle est présente une variable x , par exemple la proposition $x = 2$. Suivant la valeur de x , la proposition peut être vraie ou fausse, et on ne peut pas le déterminer *a priori*, cela va à l'encontre de ce qu'on a dit d'une proposition. C'est le cas car la variable dans P est dite libre : elle n'est pas fixée dans l'énoncé. Il existe alors deux choix relativement naturels pour attribuer un sens à P : est-ce que l'on veut dire que P est toujours vraie, peu importe la valeur de x ? Ou bien est-ce que l'on veut dire qu'il y a au moins une valeur de x telle que P est vraie ? Les quantificateurs servent à ajouter une précision de cet ordre, et permettent de ne pas avoir une variable libre dans une proposition (on dit alors que la variable est liée).

1.1.7 La quantification universelle

Le premier quantificateur se lit « pour tout » ou encore « quel que soit ». Soit $P(x)$ une proposition dépendant d'un paramètre x , la quantification universelle sur x dans $P(x)$ est la proposition écrite $\forall x, P(x)$, se lit « pour tout x , $P(x)$ est vraie » et traduit exactement ce sens : toute valeur de x rend la proposition $P(x)$ vraie.

Exemple 1.7. Supposons que nos paramètres soient des entiers. Alors on sait que tout entier est soit pair soit impair, ce qu'on peut donc noter :

$$\forall x, (\text{pair}(x) \vee \text{impair}(x))$$

Nous allons convenir d'un ordre de priorité pour les quantificateurs : ils sont moins prioritaires que les connecteurs logiques, donc

$$\forall x, \text{pair}(x) \vee \text{impair}(x)$$

se lit de la même façon.

1.1.8 La quantification existentielle

Le deuxième quantificateur se lit « il existe ». Soit $P(x)$ une proposition dépendant d'un paramètre x , on note cette quantification par $\exists x, P(x)$, qui se lit « il existe x tel que $P(x)$ », ce qui signifie que parmi tous les x possibles, il y en a au moins un (il peut aussi y en avoir plusieurs) tel que $P(x)$ est vraie.

Exemple 1.8. Supposons encore que nos paramètres sont des entiers. La proposition suivante est vraie :

$$\exists x, x = 2$$

de façon évidente puisque $2 = 2$.

1.1.9 A propos de la quantification

En logique formelle, il n'est pas choquant d'écrire par exemple $\forall x, \exists y, x = y$ sans plus de précisions, mais ça n'est pas le cas dans l'ensemble des mathématiques (et nous voulons avant tout prendre des habitudes compatibles avec l'ensemble des mathématiques). En effet, l'usage lorsque l'on quantifie une variable est de dire dans quel univers elle se trouve, c'est-à-dire préciser si ce que l'on quantifie est un ensemble, un entier, un réel, un nombre complexe, un triangle... Nous verrons plus en détail l'appartenance ensembliste dans le prochain chapitre, mais nous préciserons dès maintenant quand on introduit une variable à l'aide d'un « pour tout » ou d'un « il existe » de quelle espèce est cette variable.

Pour ce faire, plutôt que $\forall x$ nous écrirons par exemple $\forall x \in \mathbb{N}$, signifiant « pour tout x appartenant à l'ensemble des entiers naturels » ce qui est évidemment équivalent à « pour tout entier naturel x ». Les principaux ensembles sur lesquels nous quantifierons sont \mathbb{R} et \mathbb{N} (nous aurons l'occasion de voir d'autres ensembles mais cela suffira pour l'instant).

Passons désormais à l'étape suivante pour construire des raisonnements mathématiques : une fois que l'on sait faire des phrases, il faut apprendre à enchaîner les phrases de manière à faire du sens.

1.2 Raisonnement et preuve

Les logiciens, depuis le siècle dernier, ont réussi définir formellement les preuves et les hypothèses mathématiques. Cependant, comprendre ces formalismes et ces idées demande un certain recul préalable sur les mathématiques : nous nous contenterons donc de comprendre informellement ce qui fait une bonne preuve et ce qui est autorisé (ou non) dans un raisonnement mathématique. Une preuve mathématique est une suite de raisonnements élémentaires qui sont irréfutables de par leur évidence et, en s'enchaînant, forment un raisonnement plus large qui est tout aussi irréfutable. Nous appellerons les raisonnements élémentaires des règles de déduction, et allons en voir plusieurs. Chaque fois, nous allons définir une règle de déduction par trois éléments : les hypothèses que l'on utilise, les prémisses de la règle et ses conclusions. Les règles de déduction sont intimement liées à la forme des propositions, puisque pour chaque connecteur il existe une règle qui permet de construire une proposition avec ce connecteur, et une règle qui permet d'utiliser une proposition avec ce connecteur. Appliquer une règle se fait sans préciser de quelle règle l'on parle, en général, tant le choix de la règle à appliquer est évident (on utilisera seulement le mot-clé « donc », signifiant que l'on applique une règle dont on a déjà prouvé les prémisses en amont). La plupart des mathématiciens, par effet de style et pour rendre la lecture un peu plus agréable, varient les mots de liaisons et les tournures, évitant ainsi d'utiliser « donc » en permanence, mais il faut comprendre que les mots de liaisons servent à marquer l'utilisation d'une règle et que n'utiliser que « donc » ne saurait être vu comme une erreur.

Enfin, une dernière précision à propos des hypothèses : en mathématiques, toute preuve se fait sous des hypothèses, et il existe très peu (voire pas) de résultats réellement absolus, ce qui signifie qu'une preuve ne peut que montrer « ce résultat est vrai en supposant les hypothèses suivantes ». Nous mettons donc un point d'honneur dans cette partie à expliciter les hypothèses sous lesquelles tient un résultat. Si certains résultats apparaissent comme universellement vrais (par exemple que $1 + 1 = 2$), la formalisation des objets manipulés demandent en réalité un ensemble de résultats, appelés axiomes, qui sont des hypothèses tellement générales qu'on les tient toujours pour vraies.

1.2.1 L'hypothèse

Tout d'abord, la règle la plus basique d'un raisonnement est la suivante : si l'on a supposé une proposition P , alors dans ce contexte P est vraie. Cela est nécessaire pour pouvoir utiliser des hypothèses : celles-ci apparaissent au moment d'utiliser la règle d'hypothèse (comme son nom l'indique).

1.2.2 Introduction d'hypothèse

Si l'on veut ajouter une hypothèse P au contexte d'hypothèses \mathcal{H} dans lequel on écrit une preuve, il suffit d'écrire « supposons P ».

1.2.3 L'introduction de la conjonction

Pour justifier qu'une proposition de la forme $P \wedge Q$ est vraie, on justifie que P est vraie, dans un premier temps, puis que Q est vraie, dans un deuxième temps, et sous les mêmes hypothèses. La proposition $P \wedge Q$ est alors justifiée sous les hypothèses ayant permis de justifier P et Q .

Exemple 1.9. Supposons P , puis supposons aussi Q , nous allons prouver $P \wedge Q$:

Par hypothèse, P est vraie, et par hypothèse Q est vraie, donc $P \wedge Q$ est vraie.

1.2.4 L'élimination de la conjonction

Une proposition de la forme $P \wedge Q$ permet de justifier deux propositions : P et Q . En effet, avoir prouvé $P \wedge Q$ signifie qu'à la fois P et Q sont vrais, donc si, sous des hypothèses \mathcal{H} on a prouvé $P \wedge Q$, alors on peut au choix prouver P comme prouver Q sous les hypothèses \mathcal{H} . Cette règle affaiblit en quelque sorte la proposition que l'on a initialement, puisqu'elle permet de récupérer une partie seulement de la proposition initiale.

Exemple 1.10. Prenons un certain nombre x , et supposons la proposition $(x = 2) \wedge \text{pair}(x)$, alors par hypothèse, cette proposition est vraie, donc on en déduit que $x = 2$.

1.2.5 L'introduction de la disjonction

Pour prouver une proposition de la forme $P \vee Q$, il suffit au choix de prouver P ou de prouver Q .

Exemple 1.11. Prouvons que P définie comme $(0 = 0) \vee \text{pair}(2)$ est vraie par deux preuves distinctes :

- $0 = 0$ est vrai par définition de l'égalité, donc P est vraie.
- $\text{pair}(2)$ est vrai par définition de ce qu'est un entier pair, donc P est vraie.

Attention, deux preuves ont été données pour montrer qu'il y avait deux moyens de prouver $P \vee Q$ si l'on peut prouver P comme Q , mais cela ne veut pas dire qu'une preuve de $P \vee Q$ nécessite deux preuves, bien au contraire. Un autre exemple, prouvons P valant $(0 = 1) \vee \text{pair}(2)$:

$\text{pair}(2)$ est vrai par définition de ce qu'est un entier pair, donc P est vraie.

1.2.6 L'élimination de la disjonction, ou disjonction de cas

Cette règle étant un peu plus technique que les autres à utiliser, on précise en général lorsqu'on l'utilise. Supposons qu'on veuille prouver une proposition A par ce processus, la disjonction de cas est un raisonnement en trois temps :

1. Tout d'abord, on prouve une proposition de la forme $P \vee Q$.
2. Ensuite, on distingue les cas menant à $P \vee Q$, qui sont lorsque P est vraie et lorsque Q est vraie. On suppose donc P en plus, et on prouve alors A .
3. On suppose maintenant Q en plus, et on prouve A .

Si l'on a respecté ces étapes, alors on a prouvé A .

Exemple 1.12. Montrons que pour tout $x \in \mathbb{N}$, il existe $y \in \mathbb{N}$ tel que $x = 2 \times y$ ou $x = 2 \times y + 1$. Pour cela, on considère d'abord que tout entier naturel est soit pair soit impair (on le montrera dans un exemple ultérieur). On fait alors une distinction de cas :

1. x est pair ou impair.

2. Si x est pair, alors on peut écrire $x = 2 \times y$ avec $y \in \mathbb{N}$, donc on a bien prouvé qu'il existait le y que l'on cherchait.
3. Si x est impair, alors $x - 1$ est pair, donc on trouve y tel que $x - 1 = 2 \times y$, ce qui revient à dire que $x = 2 \times y + 1$: on a donc trouvé un y respectant ce que l'on voulait prouver.

Donc pour tout entier, il existe un entier y valant sa moitié ou sa moitié augmentée de 1.

1.2.7 L'introduction de la négation

L'introduction de la négation est ce qu'on appelle le raisonnement par contradiction. Ce raisonnement cherche à prouver une proposition de la forme $\neg A$ et se déroule en deux temps :

1. D'abord, on suppose A .
2. Ensuite, on prouve une absurdité, c'est-à-dire une phrase manifestement fausse (par exemple $0 = 1$).

Si l'on a fait ça, alors on a prouvé $\neg A$.

Exemple 1.13. Montrons que la proposition P donnée par « tous les nombres premiers sont impairs » est fausse en prouvant $\neg P$:

Supposons que tous les nombres premiers sont impairs, comme 2 est un nombre premier, cela signifie que 2 est impair, or il est pair : c'est une contradiction. Donc tous les nombres premiers ne sont pas impairs.

1.2.8 Le raisonnement par l'absurde

Le raisonnement par l'absurde est en quelque sorte la version inverse du raisonnement par contradiction. Si le raisonnement par contradiction suppose A pour prouver $\neg A$, le raisonnement par l'absurde suppose $\neg A$ pour prouver A . En effet, le reste est identique : on prouve en supposant $\neg A$ qu'on aboutit à une contradiction, et de cette contradiction on déduit directement A .

Certains auteurs confondent le raisonnement par contradiction et le raisonnement par l'absurde.

Exemple 1.14. Soit une proposition P , supposons $\neg\neg P$ et prouvons alors P :

Procédons par l'absurde en supposant $\neg P$, alors on a à la fois $\neg P$ et son contraire $\neg\neg P$, qui sont deux propositions incompatibles : c'est absurde, donc P est vraie.

1.2.9 L'introduction de l'implication

Si l'on veut prouver une proposition de la forme $A \implies B$, on procède en deux temps :

1. On suppose A .
2. On prouve B dans le contexte où l'on a supposé A .

Dans ce cas, on a prouvé $A \implies B$.

Exemple 1.15. L'exemple précédent nous permet de déduire que $\neg\neg P \implies P$.

1.2.10 L'élimination de l'implication, le modus ponens

Cette règle permet d'utiliser une proposition de la forme $A \implies B$. Son utilisation est simple : si l'on sait vraies les proposition $A \implies B$ et A alors on peut en déduire B . C'est l'un des principes logiques les plus basiques et son nom de *modus ponens* date de l'Antiquité.

Exemple 1.16. Les exemples ne manquent pas, mais supposons par exemple que l'on sache les choses suivantes :

- Socrate est un homme.
- Si Socrate est un homme, alors Socrate est mortel.

On en déduit que Socrate est mortel.

1.2.11 Le raisonnement par contraposée

Cette règle permet, en un sens, d'inverser le sens d'une implication. Si une proposition de la forme $A \implies B$ est vraie, alors la proposition $\neg B \implies \neg A$ est vraie aussi (et ceci marche dans l'autre sens : prouver $\neg B \implies \neg A$ suffit à prouver $A \implies B$).

1.2.12 L'équivalence

L'équivalence $A \iff B$ est exactement $(A \implies B) \wedge (B \implies A)$, ce qui signifie que prouver que deux propositions sont équivalentes se fait en prouvant que la première implique la deuxième puis que la deuxième implique la première (on appelle ça procéder par double implication) et une proposition de la forme $A \iff B$ s'utilise en faisant un *modus ponens* depuis le sens que l'on veut utiliser.

1.2.13 L'introduction du quantificateur universel

Si l'on veut prouver une proposition de la forme $\forall x \in X, P(x)$, on procède en deux temps :

1. On définit une nouvelle variable, disons α , dont on sait juste que $\alpha \in X$.
2. On prouve $P(\alpha)$

Alors on a prouvé $\forall x \in X, P(x)$.

Exemple 1.17. Montrons que pour tout $x \in \mathbb{R}$, $x + 0 = x$:

Soit $\alpha \in \mathbb{R}$, par définition de l'addition $\alpha + 0 = \alpha$, donc $\forall x \in \mathbb{R}, x + 0 = x$.

1.2.14 L'élimination du quantificateur universel

Si l'on a une proposition de la forme $\forall x \in X, P(x)$ on peut l'utiliser en remplaçant x par une valeur précise. Supposons que $\alpha \in X$ est une valeur particulière, $\forall x \in X, P(x)$ nous permet de déduire $P(\alpha)$.

Exemple 1.18. Reprenons l'exemple donné pour le *modus ponens*, mais en changeant légèrement l'énoncé :

- Tout homme est mortel.
- Socrate est un homme.

Alors, comme Socrate est un homme, ce que l'on va écrire $\text{Socrate} \in \text{Hommes}$, on peut appliquer $\forall x \in \text{Hommes}, \text{mortel}(x)$, ce qui nous fait déduire que Socrate est mortel.

1.2.15 L'introduction du quantificateur existentiel

Pour montrer $\exists x \in X, P(x)$, la règle que l'on utilise est simple : on trouve un certain $\alpha \in X$ tel que $P(\alpha)$ est vrai, ce qui nous permet de conclure qu'il existe bien $x \in X$ tel que $P(x)$.

Exemple 1.19. Montrons qu'il existe un nombre entier pair :

2 est pair, donc il existe un nombre entier pair.

1.2.16 L'élimination du quantificateur existentiel

Si l'on tient vraie une proposition de la forme $\exists x \in X, P(x)$, alors on peut introduire une variable, disons $\alpha \in X$, telle que $P(\alpha)$ est vraie.

Exemple 1.20. Montrons que s'il existe un nombre pair, alors il existe un nombre impair :

On suppose qu'il existe un nombre pair, soit x un tel nombre. Alors $x + 1$ est un nombre impair, donc il existe un nombre impair.

1.2.17 Le raisonnement par récurrence

Ce raisonnement est bien plus particulier. Il est introduit ici mais servira principalement dans des notions traitées dans des chapitres plus tardifs. Le principe du raisonnement par récurrence est, pour un prédicat P dépendant d'une variable $n \in \mathbb{N}$, de prouver $\forall n \in \mathbb{N}, P(n)$.

Pour ce faire, nous allons prouver deux choses :

- Tout d'abord, on prouve $P(0)$.
- Ensuite, on prouve $\forall n \in \mathbb{N}, P(n) \implies P(n+1)$.

En ayant prouvé ces deux éléments, on prouve alors que P est vraie pour n'importe quel entier naturel n , car $n = 0 + 1 + 1 + 1 + \dots + 1$, où l'on additionne n fois 1 à 0. Or $P(0) \implies P(1) \dots \implies P(n)$ en utilisant ce que l'on a prouvé. Ainsi, pour tout entier naturel, on a prouvé $P(n)$.

Remarque. Ajoutons un détail sur la façon dont il faut rédiger l'étape que l'on appelle l'hérédité, celle de prouver $\forall n \in \mathbb{N}, P(n) \implies P(n+1)$. Il faut tout d'abord fixer n quelconque pour pouvoir prouver l'implication à l'intérieur, puis supposer $P(n)$ vraie et prouver alors $P(n+1)$.

Exemple 1.21. Montrons que tout entier est soit pair soit impair, c'est-à-dire soit de la forme $2k$ soit de la forme $2k+1$.

- Tout d'abord, $0 = 2 \times 0$ donc 0 est pair.
- Soit $n \in \mathbb{N}$. Alors deux cas sont possibles :
 - Soit $n = 2k$, $k \in \mathbb{N}$. Dans ce cas, $n+1 = 2k+1$ donc $n+1$ est impair.
 - Soit $n = 2k+1$. Dans ce cas, $n+1 = 2k+2 = 2(k+1)$, donc $n+1$ est pair.

Dans tous les cas, $n+1$ est soit pair soit impair.

Par récurrence, on en déduit que pour tout entier naturel n , n est ou pair ou impair, i.e. qu'il existe k tel que $n = 2k$ ou $n = 2k+1$.

2 Théorie des ensembles

Cette section s'intéresse aux éléments basiques de la théorie des ensembles. Cette expression désigne à la fois la manipulation élémentaire des ensembles utilisée dans toutes les branches des maths, une forme de formalisme employé pour formaliser les mathématiques, et une théorie à part entière s'occupant d'ensembles particuliers. Cette section étant avant tout une introduction, nous nous contenterons évidemment de nous intéresser aux manipulations élémentaires autorisées en mathématiques. Nous ne rentrerons pas dans le formalisme de la théorie axiomatique de ZF à cause de sa complexité théorique, et admettrons simplement que ce qu'il est évident que l'on peut faire, peut être fait.

2.1 Que faire d'un ensemble ?

Tout d'abord, nous devons nous interroger rapidement sur ce qu'est un ensemble exactement. Les mathématiques utilisant cette notion comme primitive, un ensemble ne peut pas être décrit en termes mathématiques autres que « un ensemble est un objet qui se comporte comme un ensemble ». Ceci étant, il est important d'avoir une idée précise de ce que désigne un ensemble au niveau de l'intuition, dans ce qu'on pourrait appeler des méta-mathématiques. Un ensemble est un regroupement (appelé aussi une collection) d'objets, pouvant comporter virtuellement n'importe quoi. On appelle « élément » un objet appartenant à un ensemble donné. Un facteur important dans cette définition intuitive est que les collections que nous constituons sont insensibles à l'ordre et à la multiplicité des éléments. Par exemple un ensemble contenant x et x est exactement identique à un ensemble ne contenant x qu'une fois (ainsi chaque élément n'est listé qu'une fois, par convention). De même, un ensemble contenant x puis y est le même qu'un ensemble contenant y puis x : on dira donc seulement que l'ensemble contient x et y . Nous allons maintenant donner les différents éléments de base pour pouvoir désigner efficacement des ensembles : les définition d'ensembles par extension, compréhension, l'union, l'intersection, le complémentaire relatif, la différence symétrique, le produit cartésien et l'ensemble des parties.

2.1.1 Appartenance et égalité

Comme nous l'avons dit, un objet peut être élément d'un ensemble. Pour cela, nous utilisons le symbole \in , qui s'utilise ainsi :

$x \in X$ signifie que x appartient à l'ensemble X

De plus, la notion d'égalité, qui signifie que deux objets sont les mêmes, existe évidemment aussi pour les ensembles. Un ensemble X est égal à un autre ensemble Y lorsqu'il a les mêmes éléments, ce qui signifie que pour tout objet x , on a l'équivalence

$$x \in X \iff x \in Y$$

Exemple 2.1. En reprenant les définitions d'ensembles usuels données plus tôt, il est évident que

$$1 \in \mathbb{N}$$

et de plus, en imaginons que l'on nomme \mathbb{M} l'ensemble \mathbb{N} auquel on a ajouté 0, l'égalité

$$\mathbb{M} = \mathbb{N}$$

est vérifiée, comme le nombre d'apparitions d'un élément dans un ensemble n'a pas d'importance et que $0 \in \mathbb{N}$.

2.1.2 Description par extension.

La façon la plus simple de décrire un ensemble est encore de lister explicitement ses éléments : on appelle cela une description de l'ensemble par extension. La notation est d'écrire chaque élément de l'ensemble, séparé par une virgule, le tout entre accolades. Il n'y aura en général pas de confusion entre la virgule séparant deux éléments et celle séparant un nombre de sa partie décimale, mais dans un objectif de clarté nous utiliserons la notation anglophone où la partie décimale d'un nombre est séparée de sa partie entière par un point (par exemple « un demi » s'écrira 0.5).

Exemple 2.2. L'ensemble des entiers naturels inférieurs à 5 est $\{0, 1, 2, 3, 4, 5\}$.

Remarque (L'ensemble vide). Si un ensemble contient des éléments, il existe aussi un ensemble ne contenant pas d'éléments. Cet ensemble s'appelle l'ensemble vide et se note \emptyset : il a la caractéristique que pour tout objet x , la propriété suivante est vérifiée :

$$x \notin \emptyset$$

Remarquons d'ores et déjà que l'ensemble $\{\emptyset\}$ est parfaitement différent que \emptyset puisqu'il contient un élément. Il sera plus tard utile d'être à l'aise sur la différence entre un ensemble et ses éléments.

2.1.3 Description par compréhension

Nous connaissons plusieurs ensembles dont nous acceptons l'existence, comme \mathbb{N} ou \mathbb{R} , mais ces ensembles ont la particularité d'être infinis. Malheureusement, décrire par extension un ensemble infini est parfaitement impossible pour nous, et il nous faut donc recourir à une notation permettant d'écrire des ensembles infinis. La description par compréhension existe dans ce but : si l'on considère un ensemble X et un prédicat (propriété dépendant d'un paramètre) $P(x)$, on peut considérer l'ensemble des éléments appartenant à X qui vérifient ce prédicat. La description par compréhension permet donc d'écrire des parties de plus gros ensembles, la plupart du temps de façon concise. La grammaire pour écrire un ensemble par compréhension dépend parfois des auteurs, mais une écriture classique et admise de tout est la suivante, pour écrire « l'ensemble des éléments de X qui vérifient le prédicat P » :

$$\{x \mid x \in X, P(x)\}$$

Remarque. Selon l'auteur, la barre \mid est remplacée par deux points $:$. De plus, il arrive de vouloir plutôt que simplement sélectionner des éléments x , considérer l'image de chaque x par une certaine fonction f (par exemple au lieu de regarder les entiers impaires, on peut vouloir considérer toutes les moitiés de nombres entiers impaires). Dans ce cas, l'écriture devient dans le cas général :

$$\{f(x) \mid x \in X, P(x)\}$$

Enfin, dans le cas d'une propriété qui se déduit de façon évidente en donnant les premiers termes, on peut écrire seulement les premiers termes en utilisant de points de suspension (cette utilisation est moins rigoureuse mais beaucoup plus lisible, nous conseillons de l'utiliser seulement lorsque l'on est sûrs d'être compris sans ambiguïté).

Exemple 2.3. Explicitons l'exemple précédent : nous allons noter l'ensemble A constitué des moitiés de tous les nombres entiers impaires :

$$A = \left\{ \frac{1}{2}x \mid x \in \mathbb{N}, \text{impair}(x) \right\}$$

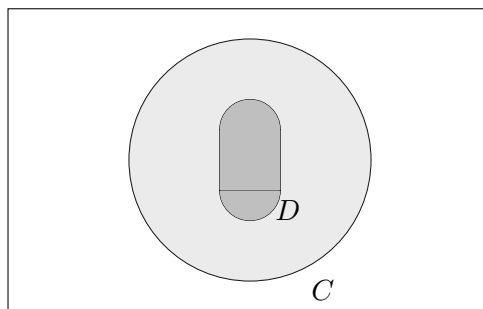
Avec des points de suspension, cet ensemble est

$$A = \{0.5, 1.5, 2.5, \dots\}$$

2.1.4 Inclusion d'ensembles

Nous allons nous intéresser à l'inclusion, permettant de comparer des ensembles. L'idée est naturelle : l'inclusion permet de déterminer si un ensemble est plus grand qu'un autre (attention, on ne parle pas ici du nombre d'éléments, qui est une notion plus fine appelée cardinale). Un ensemble E est inclus dans un ensemble F lorsque tout élément x de E est aussi un élément de F . On dit alors que E est un sous-ensemble de F , et on le note $E \subseteq F$ (ou parfois $E \subset F$, qui induit une ambiguïté sur si $E \neq F$ ou si les deux ensembles peuvent être égaux).

Nous donnerons, lorsque cela est possible, un dessin explicatif des différentes notions ensemblistes. Nous appelons diagrammes de Venn les diagrammes que nous dessineront, illustrant les ensembles. Ceux-ci sont intuitifs : une surface fermée représente un ensemble, un point dans une surface représente un élément de l'ensemble donné et une surface entièrement contenue dans une autre surface représente un ensemble inclus dans un autre.

FIGURE 1 – En gris clair C , foncé D , avec $D \subseteq C$

2.1.5 Union d'ensembles

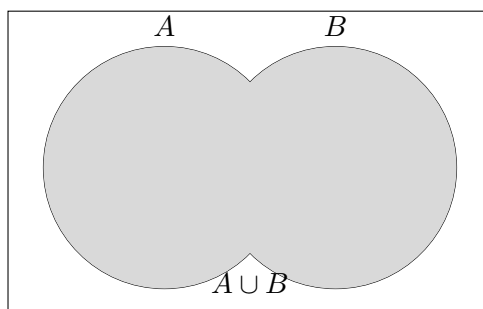
Maintenant que nous pouvons définir des ensembles et les comparer, nous allons définir des opérations sur nos ensembles, qui ressembleront fortement aux définitions que nous avons vu en logique. La première opération ensembliste est l'union : c'est une opération qui à partir de deux ensembles construit un nouvel ensemble regroupant les éléments des deux premiers ensembles. On note l'union grâce au symbole \cup , et appartenir à l'ensemble $A \cup B$ signifie exactement que l'on appartient à A ou à B . On peut donc vouloir dire

$$A \cup B = \{x \mid x \in A \vee x \in B\}$$

mais on remarque que cela ne respecte pas exactement la syntaxe que nous avons fixée pour un ensemble en compréhension car l'on ne connaît pas forcément d'ensemble plus grand dans se trouvent l'ensemble des x désignés. L'intuition reste cependant importante car elle montre que \cup est l'équivalent ensembliste du connecteur \vee .

Exemple 2.4. De façon évidente :

$$\mathbb{N} = \{x \mid x \in \mathbb{N}, \text{pair}(x)\} \cup \{x \mid x \in \mathbb{N}, \text{impair}(x)\}$$

FIGURE 2 – En gris l'union $A \cup B$

Remarque (Union généralisée). Si nous avons une famille d'ensembles (qui est une liste d'ensembles), que nous noterons X_1, X_2, \dots, X_n , alors nous pouvons définir l'union de cette famille comme les unions successives $((X_1 \cup X_2) \cup X_3) \cup \dots \cup X_n = \bigcup_{i=1}^n X_i$. La notation donnée à droite du $=$ permet d'écrire des unions avec beaucoup d'ensembles d'un coup. Si, au lieu d'une famille d'ensembles X_1, \dots, X_n , nous avons un ensemble $\mathbf{X} = \{X_1, \dots, X_n\}$, l'union $\bigcup_{i=1}^n X_i$ s'écrit directement $\bigcup \mathbf{X}$.

Exercice 2.1. Soient A et B deux ensembles, montrer d'abord que $A \subseteq A \cup B$ et $B \subseteq A \cup B$, puis montrer que $B \subseteq A$ si et seulement si $A = A \cup B$.

Correction. Soit $x \in A$, alors $x \in A \cup B$, donc $A \subseteq A \cup B$. De même, si $x \in B$ alors $x \in A \cup B$ donc $B \subseteq A \cup B$.

Si $A = A \cup B$, alors soit $x \in B$. Par inclusion, $x \in A \cup B = A$ donc $x \in A$. Donc $B \subseteq A$.

Si $B \subseteq A$, alors soit $x \in A \cup B$. Par disjonction de cas :

- si $x \in A$, alors $x \in A$
- si $x \in B$, alors par inclusion $x \in A$.

Donc dans tous les cas, $x \in A$. On en déduit donc que $A \cup B \subseteq A$, et par ce que nous avons prouvé avant $A \subseteq A \cup B$, donc $A = A \cup B$.

2.1.6 Intersection d'ensembles

Si l'union permet de construire, à partir de deux ensembles, un ensemble plus gros regroupant les éléments des deux ensembles, l'intersection est l'opération inverse : à partir de deux ensembles A et B , l'intersection $A \cap B$ est l'ensemble qui contient seulement les éléments à la fois dans A et dans B . Là encore, on pourrait écrire d'une façon proche d'un ensemble par compréhension cette opération :

$$A \cap B = \{x \mid x \in A \wedge x \in B\}$$

ce qui nous donne donc le pendant ensembliste de la conjonction logique.

Exemple 2.5. Sans le prouver, nous savons que 2 est le seul nombre à la fois pair et premier, d'où :

$$\{n \mid n \in \mathbb{N}, \text{premier}(n)\} \cap \{n \mid n \in \mathbb{N}, \text{pair}(n)\} = \{2\}$$

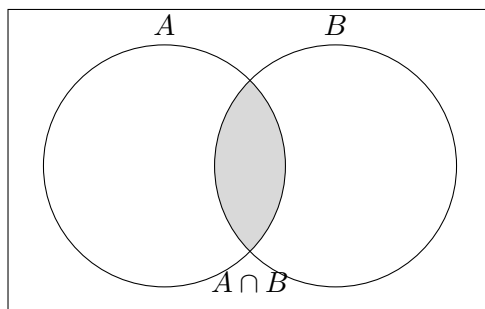


FIGURE 3 – En gris l'intersection $A \cap B$

1

Remarque. De la même façon que pour l'union, on peut définir pour X_1, \dots, X_n l'intersection $\bigcap_{i=1}^n X_i$ pour écrire avec concision l'intersection d'un grand nombre d'ensembles, et l'on pourra aussi écrire $\bigcap \mathbf{X}$ pour un ensemble regroupant tous les ensembles dont on prend l'intersection.

Exercice 2.2. Soient A et B des ensembles, montrer que $A \cap B \subseteq A$ et $A \cap B \subseteq B$, puis que $A \subseteq B$ si et seulement si $A \cap B = A$.

2.1.7 Le complémentaire relatif

Pour continuer notre analogie avec la logique, nous pouvons chercher un équivalent ensembliste à la négation. Cependant, dans le cadre des ensembles, cela signifierait chercher un ensemble de la forme

$$\{x \mid x \notin A\}$$

à partir d'un certain ensemble A . Or un ensemble pareil contiendrait presque tous les ensembles (tous sauf A), et en l'appliquant directement à \emptyset on obtiendrait ainsi l'ensemble de tous les ensembles, qui est une absurdité. En effet, il suffit alors de considérer l'ensemble des ensembles qui ne se contiennent pas eux-mêmes.

L'équivalent de la négation doit donc être relatif : nous n'allons pas écrire $\{x \mid x \notin A\}$ mais, pour deux ensembles A et B , nous allons construire $\{x \mid x \in A, x \notin B\}$. Nous listons donc les éléments qui ne sont pas dans un ensemble, mais en se fixant d'abord comme référentiel un premier ensemble. Cet ensemble, que l'on lit « A privé de B », se note $A \setminus B$ ou $\complement_A B$, voire, si l'on connaît déjà implicitement le référentiel (ce qui arrive souvent), simplement \overline{B} .

Exemple 2.6. Les nombres impairs sont exactement les nombres qui ne sont pas pairs, d'où :

$$\{n \mid n \in \mathbb{N}, \text{impair}(n)\} = \mathbb{N} \setminus \{n \mid n \in \mathbb{N}, \text{pair}(n)\}$$

De plus, le contexte de travailler avec des entiers peut parfois être évident, et dans ce cas nous pouvons directement écrire :

$$\{n \mid \text{impair}(n)\} = \overline{\{n \mid \text{pair}(n)\}}$$

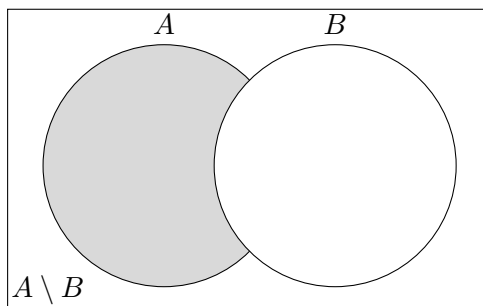


FIGURE 4 – En gris le complémentaire relatif $A \setminus B$

2.1.8 Différence symétrique

Le complémentaire permet donc de supprimer dans un ensemble les éléments qui sont en commun avec un autre ensemble. La différence symétrique, elle, applique cette idée aux deux ensembles en réunissant à la fois $A \setminus B$ et $B \setminus A$. On peut donc l'écrire $(A \setminus B) \cup (B \setminus A)$: ce sont les éléments appartenant à uniquement l'un des deux ensembles, mais pas les deux à la fois. On note cet ensemble $A \Delta B$. Il existe une opération qui lui correspond en logique mais elle est moins utile dans le strict cadre de la logique (elle peut se rencontrer dans l'informatique, par exemple en architecture d'ordinateur), on l'appelle le « ou exclusif » ou le *XOR*.

Exemple 2.7.

$$\{1, 2, 3, 4, 5\} \Delta \{5, 6, 7, 8, 9\} = \{1, 2, 3, 4, 6, 7, 8, 9\}$$

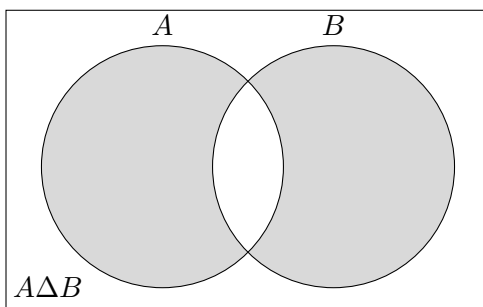


FIGURE 5 – En gris la différence symétrique $A \Delta B$

2.1.9 Produit cartésien d'ensembles

Nous allons désormais voir une forme plus complexe d'ensemble à travers le produit cartésien. Le produit cartésien de deux ensembles A et B se note $A \times B$ et correspond à l'ensemble des couples constitués d'un élément de A et d'un élément de B .

Plus précisément, un couple ressemble à un ensemble, mais avec deux éléments pour lesquels l'ordre et le nombre d'occurrences compte. On note un couple constitué de a comme premier élément et b comme deuxième élément par (a, b) , et on appelle respectivement a la première composante et b la deuxième composante du couple (a, b) . Ainsi deux couples (a, b) et (b, a) sont distincts (en supposant que $a \neq b$) et le couple (a, a) n'est pas (a) (qui n'est d'ailleurs même pas un couple). On peut étendre

cette notion à des groupements par trois, appelés triplets, ou des groupements par n avec $n \in \mathbb{N}$, appelés n -uplets. Ici, nous adopterons la convention qu'un n -uplet est simplement un couple avec un couple pour l'une de ses composantes (par exemple (a, b, c) est juste une façon commode d'écrire $((a, b), c)$).

Exemple 2.8. Commençons par un exemple simple en considérant $A = \{2, 4, 6\}$ et $B = \{3, 6, 9\}$, alors :

$$A \times B = \{(2, 3), (2, 6), (2, 9), (4, 3), (4, 6), (4, 9), (6, 3), (6, 6), (6, 9)\}$$

Remarque (Carré d'un ensemble). Pour un ensemble A , on écrit A^2 pour l'ensemble $A \times A$, et même A^3 pour $A \times A \times A$, jusqu'à A^n pour $A \times \dots \times A$ où l'on multiplie n fois A à lui-même.

Nous allons illustrer l'idée de produit cartésien en utilisant le plan usuel. Pour cela, nous allons considérer acquise la représentation d'un système de coordonnées et la correspondance entre \mathbb{R}^2 et le plan usuel (nous reviendront plus tard sur les fondements de ces affirmations, mais les utiliserons ici de façon intuitive). C'est-à-dire que l'on va considérer fixés deux axes gradués (x et y), orthogonaux, et nous assimilerons un point A à ses coordonnées dans ce repère, qui seront donc un couple (x_A, y_A) . Nous noterons I l'ensemble des nombres réels entre 1 et 2 : $I = \{x \mid x \in \mathbb{R}, 1 \leq x \leq 2\}$. L'ensemble $I^2 = I \times I$ est alors l'ensemble des points dont l'abscisse est entre 1 et 2 et l'ordonnée entre 1 et 2, puisque ce sont les couples formés de réels tous les deux entre 1 et 2.

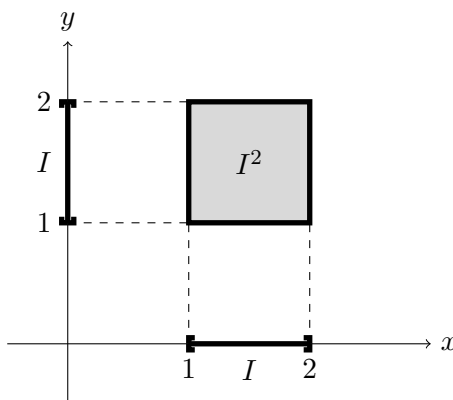


FIGURE 6 – En gris le produit cartésien $I \times I$

2.1.10 Ensemble des parties

Pour finir cette partie, nous allons nous intéresser à une construction permettant de lister toutes les parties d'un ensemble. En effet, si A est un ensemble, alors tous les ensembles X tels que $X \subseteq A$ existent, et l'on peut tous les regrouper dans un ensemble (dont les éléments sont donc des ensembles), que l'on appelle ensemble des parties de A , et que l'on note $\mathcal{P}(A)$. Remarquons que par définition \emptyset et A sont des éléments de $\mathcal{P}(A)$ puisque $\emptyset \subseteq A$ et $A \subseteq A$. Pour des ensembles quelconques A et E , on a donc l'équivalence suivante :

$$A \subseteq E \iff A \in \mathcal{P}(E)$$

Exemple 2.9. Nous allons construire plusieurs ensembles de parties, successivement, à partir de l'ensemble vide :

$$\mathcal{P}(\emptyset) = \{\emptyset\} \quad \mathcal{P}(\mathcal{P}(\emptyset)) = \{\{\emptyset\}, \emptyset\} \quad \mathcal{P}(\mathcal{P}(\mathcal{P}(\emptyset))) = \{\{\{\emptyset\}, \emptyset\}, \{\{\emptyset\}\}, \emptyset, \{\emptyset\}\}$$

On conviendra aisément que le dernier ensemble est plutôt difficile à lire, mais il permet de remarquer qu'à partir d'un ensemble à 0 éléments, nous obtenons successivement des ensembles à 1, puis 2, puis 4 éléments. Nous pourrions donc conjecturer sur le nombre de parties d'un ensemble à n éléments donnés, mais nous traiterons ce problème plus tard.

2.1.11 Intervalles de nombres réels

On appelle intervalle une partie E de \mathbb{R} convexe, c'est-à-dire que si $x \in E$ et $y \in E$ alors pour tout $x \leq t \leq y, t \in E$. Un intervalle s'écrit sous l'une des formes suivantes :

- $[a; b] = \{x \mid a \leq x \leq b\}$
- $[a; b[= \{x \mid a \leq x < b\}$
- $]a; b] = \{x \mid a < x \leq b\}$
- $]a; b[= \{x \mid a < x < b\}$
- $] - \infty; b] = \{x \mid x \leq b\}$
- $] - \infty; b[= \{x \mid x < b\}$
- $[a; +\infty[= \{x \mid a \leq x\}$
- $]a; +\infty[= \{x \mid a < x\}$

2.1.12 Raisonner sur des ensembles

Maintenant que nous avons vu les principales opérations et la grammaire élémentaire pour parler des ensembles, nous allons parler de la façon dont nous pouvons raisonner avec les ensembles. Le but ici ne sera pas de décrire des raisonnements alambiqués et virtuoses mais seulement de donner au lecteur une idée de ce qu'il devra écrire dans des exercices basiques de théorie des ensembles.

Introduire un élément L'un des procédés les plus souvent utilisés pour raisonner sur les ensembles est d'introduire un élément quelconque de cet ensemble pour raisonner dessus. Pour revenir à la partie sur les raisonnements, l'idée derrière ce procédé est simplement que $\forall x \in X \implies P(x)$ doit se prouver en introduisant un élément x de X quelconque.

Exemple 2.10. Montrons que $A \cap B \subseteq A$:

Soit $x \in A \cap B$, alors comme $x \in A \cap B$, on en déduit que $x \in A$, donc $x \in A \cap B \implies x \in A$, ce qui signifie par définition que $A \cap B \subseteq A$.

Raisonner par double inclusion La façon la plus simple de prouver qu'un ensemble est égal à un autre est de procéder par double inclusion. Ce principe se fait en deux temps. Supposons qu'on possède un ensemble X et un ensemble Y , on prouve :

- que $X \subseteq Y$, puis
- que $Y \subseteq X$

et alors $X = Y$.

Exemple 2.11. Soient les ensembles $A = \{n \mid n \in \mathbb{N}, \text{pair}(n)\}$ et $B = \{2 \times n \mid n \in \mathbb{N}\}$, et montrons que $A = B$:

Raisonnons par double inclusion :

- Soit $n \in A$, alors par définition de la parité, il existe un nombre $k \in \mathbb{N}$ tel que $n = 2 \times k$. On en déduit que n est le double d'un nombre entier, ce qui signifie que $n \in B$.
- Soit $n \in B$, alors $n = 2 \times k$ pour un certain $k \in \mathbb{N}$ par définition des éléments de B . n est alors pair par définition, donc $n \in A$.

Donc $A = B$.

2.2 Relations binaires et fonctionnelles

Cette partie va parler des relations binaires, qui sont un outil essentiel pour la construction et la compréhension des mathématiques. Nous verrons dans l'ordre les généralités sur les relations binaires (leur définition, principalement), les relations d'ordre, les relations d'équivalence et les fonctions.

2.2.1 Relation binaire

Commençons par donner la définition d'une relation binaire sur un ensemble donné. Faisons un *aparte* pour distinguer une définition d'une caractérisation. Une définition correspond à la première rencontre avec un objet mathématique, alors qu'une caractérisation est une propriété d'un objet qui *pourrait* remplacer la définition, au sens où tout objet suivant la définition est exactement un objet suivant la caractérisation. Nous verrons plus tard des caractérisations, mais l'important dans cette opposition, à l'heure actuelle, est que la définition d'un objet ne demande en général pas d'argument ou de démonstration : on se contente d'affirmer à quoi renvoie un certain terme mathématique.

Définition 2.1 (Relation binaire). Une relation binaire \mathcal{R} sur un ensemble E est une partie de E^2 (c'est-à-dire un ensemble $\mathcal{R} \subseteq E^2$). On dit que deux éléments $x \in E$ et $y \in E$ sont en relation pour \mathcal{R} lorsque le couple (x, y) est un élément de \mathcal{R} . Plutôt que d'écrire cela $(x, y) \in \mathcal{R}$, on écrit en général $x\mathcal{R}y$.

En général, une relation binaire est donnée par une description des éléments qui sont en relation plutôt que par la liste explicite des couples (cette liste explicite est en général impossible à faire puisque la plupart des relations binaires que nous utiliserons seront sur des ensembles infinis).

Listons quelques relations binaires usuelles pour mieux se faire à cette notion nouvelle.

Exemple 2.12.

- Sur \mathbb{N} , \leq est une relation binaire : on écrit $x \leq y$ pour dire que x est inférieur à y .
- Pour un ensemble E quelconque, l'égalité $=$ peut se considérer comme une relation binaire sur E , c'est alors l'ensemble $\{(x, x) \mid x \in E\}$ qu'on appelle aussi diagonale de E . En effet, chaque élément x n'est que dans le couple (x, x) , donc n'est en relation qu'avec lui-même, ce qui correspond bien à $x = x$.
- Si l'on définit \mathcal{D} comme l'ensemble des droites du plan, alors le parallélisme est une relation binaire sur \mathcal{D} .
- Sur \mathbb{N} , la relation \mid définie par $x \mid y$ s'il existe k tel que $y = k \times x$ (appelée divisibilité) est une relation binaire.

Nous allons maintenant présenter des notions liées aux relations permettant de mieux voir quelles en sont les manipulations possibles. Comprendre en profondeur le sens et l'utilité de ces définitions n'est pas nécessaire pour la suite.

Définition 2.2 (Complémentaire, relation réciproque). Soit E un ensemble et \mathcal{R} une relation binaire sur E . On appelle complémentaire de \mathcal{R} la relation $\overline{\mathcal{R}}$ décrite par l'ensemble $\overline{\mathcal{R}} = \{(x, y) \mid x \in E, y \in E, (y, x) \notin \mathcal{R}\}$. Cela signifie que deux éléments sont en relation pour $\overline{\mathcal{R}}$ si et seulement s'ils ne sont pas en relation pour \mathcal{R} .

On appelle la relation réciproque de \mathcal{R} , que l'on note \mathcal{R}^{-1} , la relation définie par $x\mathcal{R}^{-1}y \iff y\mathcal{R}x$.

Exemple 2.13. Dans \mathbb{N} , le complémentaire de \leq est $>$, tandis que la réciproque de \leq est \geq .

Donnons les propriétés qui sont habituellement utilisées pour les relations binaires.

Définition 2.3 (Réflexivité, symétrie, transitivité, antiréflexivité, antisymétrie). Soit un ensemble E . On définit les propriétés suivantes pour une relation binaire \mathcal{R} sur E :

Réflexivité : $\forall x \in E, x\mathcal{R}x$

Symétrie : $\forall x \in E, \forall y \in E, x\mathcal{R}y \implies y\mathcal{R}x$

Transitivité : $\forall x \in E, \forall y \in E, \forall z \in E, (x\mathcal{R}y \wedge y\mathcal{R}z) \implies x\mathcal{R}z$

Antiréflexivité : $\forall x \in E, \neg(x\mathcal{R}x)$

Antisymétrie : $\forall x \in E, \forall y \in E, (x\mathcal{R}y \wedge y\mathcal{R}x) \implies x = y$

Définition 2.4 (Finesse). Soit E un ensemble et \mathcal{R}_1 et \mathcal{R}_2 deux relations binaires sur E . On dit que \mathcal{R}_1 est plus fine que \mathcal{R}_2 lorsque $\mathcal{R}_1 \subseteq \mathcal{R}_2$. Cela revient à la proposition suivante :

$$\forall x \in E, \forall y \in E, x\mathcal{R}_1y \implies x\mathcal{R}_2y$$

Une relation plus fine qu'une autre possède donc moins de liaisons entre les éléments.

Exercice 2.3. Soit E un ensemble et \mathcal{R} une relation binaire réflexive sur E . Montrer que l'égalité est une relation binaire plus fine que \mathcal{R} .

Exercice 2.4. Soit E un ensemble et \mathcal{R} une relation binaire transitive et symétrique. Montrer que pour tout $x \in E, y \in E$, si $x\mathcal{R}y$ alors $x\mathcal{R}x$ et $y\mathcal{R}y$.

En déduire qu'une relation transitive et symétrique où tout élément est en relation avec au moins un élément est réflexive.

Exercice 2.5. Soit E un ensemble et \mathcal{R} une relation binaire symétrique et antisymétrique. Montrer que \mathcal{R} est l'égalité.

Exercice 2.6 (Construire une relation réflexive et transitive $*$). Soit E un ensemble et \mathcal{R} une relation binaire sur E . On définit \mathcal{R}^* , appelée la clôture réflexive et transitive de \mathcal{R} de la façon suivante : $x\mathcal{R}^*y$ si $x = y$ ou s'il existe une suite finie x_0, \dots, x_n où $x_0 = x, x_n = y$ et pour chaque $i \in \{1, \dots, n-1\}$, $x_i\mathcal{R}x_{i+1}$.

Montrer que \mathcal{R}^* est bien une relation réflexive et transitive, que de plus $\mathcal{R} \subseteq \mathcal{R}^*$ (ce qui est équivalent à dire que $x\mathcal{R}y \implies x\mathcal{R}^*y$), et enfin, que \mathcal{R}^* est la plus fine relation réflexive et transitive qui contient \mathcal{R} .

Exercice 2.7 (Construire une relation symétrique $*$). Soit E un ensemble et \mathcal{R} une relation binaire sur E . Montrer que $\mathcal{R} \cup \mathcal{R}^{-1}$ est une relation symétrique contenant \mathcal{R} . Montrer, de plus, que $\mathcal{R} \cup \mathcal{R}^{-1}$ est plus fine que toute les relations symétriques contenant \mathcal{R} .

Exercice 2.8 (Intersection et union de relations $**$). Soit E un ensemble, et \mathcal{R} un ensemble de relations binaires sur E (donc $\mathcal{R} \subseteq \mathcal{P}(E^2)$). Montrer que $\mathcal{R}' = \bigcap \mathcal{R}$ est aussi une relation binaire sur E , et que \mathcal{R}' est plus fine que toute relation $\mathcal{R} \in \mathcal{R}$. De même, montrer que $\mathcal{R}'' = \bigcup \mathcal{R}$ est une relation binaire sur E et que toute relation $\mathcal{R} \in \mathcal{R}$ est plus fine que \mathcal{R}'' .

Nous verrons plus tard que regrouper plusieurs relations ou objets divers pour en prendre l'intersection sera souvent fructueux, car beaucoup de propriétés restent stables par intersection (ce qui signifie que si l'on prend un ensemble d'objets possédant une propriété, alors l'intersection de tous ces objets possédera aussi cette propriété). La prochaine partie nous donnera un tel exemple.

2.2.2 Relation d'équivalence

Une relation d'équivalence est une relation binaire particulière. Celle-ci prend une importance particulière car elle permet de relever des propriétés partagées par des éléments, et devient un élément de construction mathématique essentiel lorsque l'on veut traiter ces propriétés en tant que telles. Par exemple, la propriété pour des droites d'être parallèles entre elles est intéressante, mais on peut en définir la notion de direction, qui est « l'endroit vers lequel pointent toutes les droites parallèles à une certaine droite ». Nous allons donc voir comment définir rigoureusement ce procédé.

Commençons par présenter des exemples de relations d'équivalences usuelles pour donner une idée de l'intuition derrière cette notion :

Exemple 2.14 (Relations d'équivalence classiques). Les relations suivantes sont des relations d'équivalences :

- La relation de parallélisme sur l'ensemble \mathcal{D} des droites du plan affine est une relation d'équivalence.
- La similitude sur l'ensemble \mathcal{T} des triangles du plan est une relation d'équivalence.
Pour rappel, deux triangles sont semblables si leurs angles sont les mêmes, ou de façon équivalente si l'on peut les superposer en en tournant un et en l'agrandissant ou en le réduisant.
- Soit $n \in \mathbb{N}$ fixé, la relation dite de congruence modulo n (qui sera détaillée plus tard) est une relation d'équivalence. On la définit par le fait que si x est congru à y modulo n alors n divise $y - x$, ou encore que x et y ont le même reste par la division euclidienne par n .

Nous pouvons maintenant donner la définition d'une relation d'équivalence :

Définition 2.5 (Relation d'équivalence). Une relation d'équivalence \sim sur un ensemble E est une relation binaire possédant les propriétés suivantes :

Réflexivité : $\forall x \in E, x \sim x$.

Symétrie : $\forall x \in E, \forall y \in E, x \sim y \implies y \sim x$.

Transitivité : $\forall x \in E, \forall y \in E, \forall z \in E, (x \sim y \wedge y \sim z) \implies x \sim z$.

L'intérêt d'une relation d'équivalence est de pouvoir regrouper les éléments par groupes possédant une propriété commune, c'est-à-dire étant tous en relation.

Définition 2.6 (Classe d'équivalence). Soit E un ensemble et \sim une relation d'équivalence sur E . Soit $x \in E$. On appelle classe d'équivalence de x , et l'on notera C_x ou \bar{x} l'ensemble :

$$C_x = \{y \mid y \in E, x \sim y\}$$

qui sont des ensembles non vides car $x \in C_x$ (par réflexivité de \sim).

L'utilisation de la notation \bar{x} est standard mais peut porter à confusion à cause de l'utiliser de pour de nombreuses notions mathématiques, aussi nous garderons la notation C_x pour qualifier la classe d'un élément x .

L'intérêt de cette notion est que la classe d'équivalence ne varie pas suivant l'élément choisi dans la classe. Le résultat suivant le montre :

Proposition 2.2.1. Soit E un ensemble et \sim une relation d'équivalence sur E , soient $x \in E$ et $y \in E$. Il y a alors deux possibilités :

- Si $x \sim y$, alors $C_x = C_y$.
- Si $\neg(x \sim y)$, alors $C_x \cap C_y = \emptyset$.

Démonstration. Prouvons le premier cas par double inclusion :

Soit $z \in C_x$, prouvons que $z \in C_y$. Par définition, comme $z \in C_x$, $x \sim z$. Or par hypothèse, $x \sim y$ donc, puisque \sim est symétrique, $y \sim x$, ce qui nous donne par transitivité que $y \sim z$, soit $z \in C_y$. Donc $z \in C_x \implies z \in C_y$, donc $C_x \subseteq C_y$.

Réciproquement, si $z \in C_y$, alors par définition $y \sim z$ et l'hypothèse $x \sim y$ nous permet par transitivité de déduire que $x \sim z$, donc que $z \in C_x$, donc $C_y \subseteq C_x$.

D'où par double inclusion que $C_x = C_y$.

Prouvons le deuxième cas par contradiction :

Supposons qu'il existe $z \in C_x \cap C_y$. Alors on en déduit que $z \in C_x$ et que $z \in C_y$. De ces deux hypothèses, respectivement, on déduit que $x \sim z$ et $y \sim z$, donc par symétrie et transitivité de \sim , on en déduit que $x \sim y$, ce qui contredit notre hypothèse de départ.

Donc $C_x \cap C_y = \emptyset$. □

Nous pouvons maintenant définir les ensembles quotients, qui seront essentiels pour construire des objets mathématiques de plus en plus complexes.

Définition 2.7 (Ensemble quotient). Soit E un ensemble et \sim une relation d'équivalence sur E . On appelle le quotient de E par \sim , et on note E/\sim l'ensemble

$$E/\sim = \{C_x \mid x \in E\}$$

qui est donc l'ensemble des regroupement de familles de même propriété pour cette relation d'équivalence.

Remarquons qu'il existe naturellement la fonction $\pi : E \longrightarrow E/\sim$. Nous détaillerons la notion $x \longmapsto C_x$

de fonction plus tard, mais nous donnons dès maintenant l'existence d'une telle fonction pour l'étudier plus tard.

Exercice 2.9. Montrer que les deux relations suivantes sont des relations d'équivalence :

- La relation de parallélisme sur l'ensemble \mathcal{D} des droites du plan (nous pouvons utiliser les résultats appris en 6e).
- (*) Pour $n \in \mathbb{N}$ fixé, la congruence modulo n .

Exercice 2.10 (Construire une relation d'équivalence *). Soit E un ensemble et \rightarrow une relation binaire sur E . On note \longleftrightarrow la relation symétrique décrite dans l'exercice 2.7 et \longleftrightarrow^* sa clôture réflexive et symétrique.

Montrer que \longleftrightarrow^* est une relation d'équivalence contenant \rightarrow , et qu'elle est la plus fine relation d'équivalence contenant \rightarrow .

Exercice 2.11 (Construction alternative d'une relation d'équivalence **). Soit E un ensemble et \mathcal{R} une relation binaire sur E . On note \mathcal{R} l'ensemble des relations d'équivalence contenant \mathcal{R} .

1. Montrer que \mathcal{R} n'est pas vide (car on veut en considérer l'intersection).
2. Montrer que si \mathcal{R} est constitué uniquement de relations réflexives, alors $\bigcap \mathcal{R}$ est réflexive, et contient \mathcal{R} .
3. Montrer que si \mathcal{R} est constitué uniquement de relations symétriques, alors $\bigcap \mathcal{R}$ est réflexive, et contient \mathcal{R} .
4. Montrer que si \mathcal{R} est constitué uniquement de relations transitives, alors $\bigcap \mathcal{R}$ est réflexive, et contient \mathcal{R} .
5. En déduire que $\bigcap \mathcal{R}$ est la relation d'équivalence la plus fine qui contient \mathcal{R} .

Remarque. D'après les deux exercices précédents, on remarque qu'il existe deux façon de définir une structure contenant un ensemble : la première est d'étendre cet ensemble jusqu'à obtenir la structure souhaitée, et la deuxième est de prendre une surestimation des structure qui contiennent l'ensemble en question pour en prendre l'intersection. L'avantage de la première méthode est d'en général donner une description explicite de la structure attendue, mais celle-ci peut-être difficilement manipulable, alors que la deuxième façon permet de traiter de façon systématique ce genre de problèmes (c'est ce que nous utiliserons dans la suite du document).

2.2.3 Partition d'un ensemble

Cette partie permet de mieux se familiariser avec les quotients par une relation d'équivalence. En effet, on appelle partition d'un ensemble l'ensemble de parties produit par un quotient de la forme E/\sim .

Définition 2.8 (Partition d'un ensemble). Soit E un ensemble et $\mathcal{P} \subseteq \mathcal{P}(E)$ un ensemble de parties de E . On dit que \mathcal{P} est une partition de E si :

- $\forall P \in \mathcal{P}, P \neq \emptyset$.
- $\bigcup \mathcal{P} = E$.
- $\forall P \in \mathcal{P}, \forall P' \in \mathcal{P}, P \neq P' \implies P \cap P' = \emptyset$.

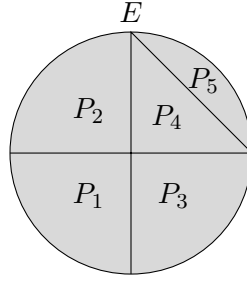
Une partition d'un ensemble est donc exactement une façon de découper un ensemble en paquets pour que chaque élément de l'ensemble aille dans exactement un ensemble.

Le dessin ci-dessous présente un ensemble E avec une partition $\mathcal{P} = \{P_1, P_2, P_3, P_4, P_5\}$.

Donnons une définition équivalente à une partition qui sera moins centrée sur les parties que sur les éléments de l'ensemble :

Proposition 2.2.2 (Caractérisation d'une partition). Soit E un ensemble et $\mathcal{P} \subseteq \mathcal{P}(E)$, les deux propositions sont équivalentes :

- \mathcal{P} est une partition de E .
- Les conditions suivantes sont respectées :
 - $\forall P \in \mathcal{P}, P \neq \emptyset$.

FIGURE 7 – Un ensemble E partitionné

- $\forall x \in E, \exists P \in \mathbf{P}, x \in P$.
- $\forall x \in E, \forall P \in \mathbf{P}, \forall P' \in \mathbf{P}, (x \in P \wedge x \in P') \implies P = P'$

Remarque. Les points 2 et 3 de la caractérisation précédente peuvent se réécrire

$$\forall x \in E, \exists! P \in \mathbf{P}, x \in P$$

signifiant « il existe un unique élément tel que » plutôt que simplement « il existe ». Nous éviterons d'utiliser ce symbole car il évite de voir l'importance que revêt l'unicité dans la phrase.

Démonstration. Montrons qu'une partition vérifie bien les conditions décrites :

- La première phrase est commune à la définition et à la caractérisation.
- Soit $x \in E$, alors comme $E = \bigcup \mathbf{P}$, $x \in \bigcup \mathbf{P}$, ce qui signifie par définition de l'union, qu'il existe $P \in \mathbf{P}$ tel que $x \in P$, ce qui est le résultat que nous voulons démontrer.
- Soit $x \in E$ et soient $P \in \mathbf{P}$, $P' \in \mathbf{P}$. Par contraposée de la dernière propriété d'une partition, comme $\neg(P \cap P' = \emptyset)$ (en effet, cette intersection contient au moins x), on en déduit $\neg(P \neq P')$, c'est-à-dire $P = P'$.

Montrons le sens réciproque, c'est-à-dire que les conditions suffisent à construire une partition :

- Le fait que toutes les parties sont non vides est vrai par hypothèse.
- Pour hypothèse, tout élément $x \in E$ appartient à une partie $P \in \mathbf{P}$, donc $E \subseteq \bigcup \mathbf{P}$, et l'autre inclusion est évidente puisque tous les éléments des parties de E sont des éléments de E . D'où $\bigcup \mathbf{P} = E$.
- Soient P et P' deux éléments de \mathbf{P} . Supposons que $P \neq P'$, alors il n'existe pas d'élément $x \in E$ tel que $x \in P \wedge x \in P'$, ce qui signifie qu'il n'existe pas d'élément $x \in E$ tel que $x \in P \cap P'$, donc $P \cap P' = \emptyset$.

Les deux définitions sont donc équivalentes.

□

Nous allons maintenant voir un théorème exprimant l'équivalence entre une relation d'équivalence et une partition.

Théorème 2.1 (Equivalence d'une partition et d'un ensemble quotient). *Soit E un ensemble. Alors pour toute partition \mathbf{P} il existe une unique relation d'équivalence $\sim_{\mathbf{P}}$ telle que $E/\sim_{\mathbf{P}} = \mathbf{P}$.*

Réciproquement, si \sim est une relation d'équivalence, alors E/\sim est une partition de E .

Démonstration. Comme pour chaque $x \in E$, il existe un unique $P \in \mathbf{P}$ tel que $x \in P$, on définit $\sim_{\mathbf{P}}$ par le fait que deux éléments x et y sont en relation quand ils appartiennent à la même partie $P \in \mathbf{P}$. Montrons alors que cela forme une relation d'équivalence ;

- Par définition, pour tout $x \in E$, en nommant P la partie de la partition à laquelle appartient x , $x \in P$, donc $x \sim_{\mathbf{P}} x$.
- Soient $x \in E$ et $y \in E$ tels que $x \sim_{\mathbf{P}} y$, nommons P la partie à laquelle appartiennent x et y . Il est évident que $y \sim_{\mathbf{P}} x$ comme P , un ensemble, n'est pas ordonné.

- Soient $x \in E$, $y \in E$ et $z \in E$ tels que $x \sim_P y$ et $y \sim_P z$, alors les trois éléments appartiennent à une même partie $P \in \mathbf{P}$, donc $x \sim_P z$.

Donc \sim_P est une relation d'équivalence.

Montrons maintenant que $E/\sim_P = \mathbf{P}$:

Soit $x \in E$, alors $C_x = P$ pour un certain $P \in \mathbf{P}$. En effet, puisque tous les éléments en relation avec x sont dans la même partie $P \in \mathbf{P}$, on fixe ce P et on en déduit que $C_x \subseteq P$. Réciproquement, tout élément de P est en relation avec x par définition puisque ceux-ci appartiennent à la même partie. Donc $C_x = P$.

Ainsi, pour toute partie $P \in \mathbf{P}$, on trouve un élément $x \in P$ (car P est non vide par hypothèse) et on en déduit que $P = C_x$, dont on déduit que $P \subseteq E/\sim_P$. Réciproquement, si l'on prend $P \in E/\sim_P$, alors pour tout élément $x \in P$ appartient à un certain $P' \in \mathbf{P}$, mais $P' = C_x = P$ d'après les argument précédents, donc $E/\sim_P \subseteq \mathbf{P}$.

Donc par double inclusion, $E/\sim_P = \mathbf{P}$

Montrons que E/\sim est une partition de E :

- Tout d'abord, soit $P \in E/\sim_P$, par définition on trouve $x \in P$ tel que $P = C_x$, or \sim est réflexive donc $x \in P$, donc $P \neq \emptyset$.
- Soit $x \in E$, par définition $x \in C_x$ et $C_x \in E/\sim_P$, donc on trouve $P \in E/\sim_P$ tel que $x \in P$.
- Soit $x \in E$, $P, P' \in E/\sim_P$ tels que $x \in P \wedge x \in P'$. Alors on trouve y et z tels que $P = C_x$ et $P' = C_z$. Alors par 2.2.1, comme $P \cap P' \neq \emptyset$, on en déduit que $P = P'$.

Donc E/\sim est bien une partition de E . □

Exercice 2.12 (Direction d'une droite). Soit \mathcal{D} l'ensemble des droites du plan. On décide de partitionner cet ensemble en regroupant les droites qui ont la même direction. Quelle est la relation d'équivalence correspondant à cette partition ?

2.2.4 Relation d'ordre

Outre les relations d'équivalence, un autre type de relations important pour toutes les maths usuelles est celui des relations d'ordre. Celles-ci servent, comme leur nom l'indique, à ordonner un ensemble. Leur exemple le plus évident est la relation \leq qui permet d'ordonner les entiers, et qui est la base du fait de pouvoir compter. En effet, l'intérêt de compter $1, 2, 3, 4, \dots$ dans cet ordre est juste que ces entiers viennent avec un ordre naturel. Voyons donc la définition d'une relation d'ordre.

Définition 2.9 (Relation d'ordre). Soit E un ensemble et \preceq une relation binaire sur E . On dit que \preceq est une relation d'ordre lorsqu'elle possède les propriétés suivantes :

Réflexivité : $\forall x \in E, x \preceq x$

Antisymétrie : $\forall x \in E, \forall y \in E, (x \preceq y \wedge y \preceq x) \implies x = y$

Transitivité : $\forall x \in E, \forall y \in E, \forall z \in E, (x \preceq y \wedge y \preceq z) \implies x \preceq z$

Nous ne nous attarderons pas beaucoup sur les relations d'ordre car, si celles-ci ouvrent le champ à des théories très larges, elles sont largement hors de notre étude et beaucoup trop abstraites. Nous allons donner dans un premier temps des exemples de relations d'ordre, puis nous verrons quelques éléments importants dans l'étude des relations d'ordre.

Exemple 2.15. Voici quelques relations d'ordre classiques :

- Dans \mathbb{N} , la relation \leq est une relation d'ordre.
- Dans \mathbb{N} , la relation $|$ définie précédemment est aussi une relation d'ordre.
- Si l'on fixe un ensemble E , alors \subseteq est une relation d'ordre sur $\mathcal{P}(E)$.

Une différence notable entre \leq et, par exemple, \subseteq , est que l'on peut toujours comparer deux entiers avec \leq , alors qu'il existe des parties de E que l'on ne peut pas comparer.

Exemple 2.16. En prenant $E = \{1, 2, 3\}$, on remarque que $\{1\} \subseteq \{1, 2\}$ et $\{1\} \subseteq \{1, 3\}$ mais que $\neg(\{1, 2\} \subseteq \{1, 3\})$, de même que $\neg(\{1, 3\} \subseteq \{1, 2\})$. Donc ces deux ensembles sont dits incomparables.

On peut donc définir une nouvelle propriété pour certains ordres.

Définition 2.10 (Ordre total). Soit E un ensemble et \preceq une relation d'ordre sur E . On dit que \preceq est total si :

$$\forall x \in E, \forall y \in E, x \preceq y \vee y \preceq x$$

signifiant que deux éléments peuvent toujours être comparés.

Nous pouvons aussi définir la notion de minorant et de majorant.

Définition 2.11 (Minorant, majorant). Soit (E, \preceq) un ensemble ordonné, et soit $A \subseteq E$ une partie de E . On dit que m est un minorant de A lorsque

$$\forall x \in A, m \preceq x$$

c'est donc un élément plus petit que l'ensemble des éléments de A . On dit que M est un majorant de A lorsque

$$\forall x \in A, x \preceq M$$

c'est donc un élément plus grand que l'ensemble des éléments de A .

Remarque. Un minorant, comme un majorant, n'est en général pas unique. En prenant par exemple $E = \mathbb{R}$ et $A = \{0, 1, 5, 6\}$, on remarque que -1 comme 0 sont des minorants de A . De même, 8 et 9 sont des majorants de A .

Ajoutons une notion plus fine que celle de majorant et de minorant que sont les bornes supérieures et inférieures.

Définition 2.12 (Borne supérieure, borne inférieure). Soit (E, \preceq) un ensemble ordonné, et soit $A \subseteq E$ une partie de E . On dit que m est la borne inférieure de A si m est un minorant de A tel que pour tout x minorant de A , on a

$$x \preceq m$$

et de même, on dit que M est la borne supérieure de A si M est un majorant de A tel que pour tout majorant x de A , on a

$$M \preceq x$$

Ce sont donc le plus grand minorant et le plus petit majorant. Attention cependant, ceux-ci n'existent pas forcément.

On remarquera l'emploi du déterminant « un » pour parler de minorant, et « le » pour parler de borne inférieure. Cette pratique est essentielle, car « le » et « la » insinuent qu'il y a unicité de l'objet considéré. Prouvons, sous réserve d'existence, cette unicité.

Proposition 2.2.3. Si A a une borne supérieure, alors celle-ci est unique. De même, si A a une borne inférieure, celle-ci est unique.

Démonstration. Nous ne traiterons que le cas de la borne supérieure : l'autre cas est laissé en exercice au lecteur.

Soit M borne supérieure de A et M' borne supérieure de A . Alors comme M est plus petit que tous les majorants de A et que M' est un majorant de A , on en déduit que $M \preceq M'$. De même, M est un majorant de A et M' est plus petit que tous les majorants de A , donc $M' \preceq M$.

On en déduit que $M = M'$.

□

2.2.5 Fonctions

Nous allons maintenant introduire ce qui est sûrement l'élément le plus important des mathématiques : les fonctions. En effet, que cela soit en algèbre, en analyse ou même en géométrie, la notion de fonction est omniprésente. Nous allons donc préciser ce qu'est une fonction.

Remarque. Dans ce document, nous parlerons indifféremment de fonctions ou d'applications. Les deux termes ont la même signification.

Définition 2.13 (Fonction). Soient E et F des ensembles. On appelle une fonction un triplet (E, F, Γ) où $\Gamma \subseteq E \times F$ est une relation entre E et F telle que pour tout $x \in E$, il existe un unique couple $\alpha \in \Gamma$ tel que x est le premier élément de α .

Autrement dit, Γ est un ensemble de couples de la forme (x, y) où $x \in E$, $y \in F$ et tel qu'à x fixé, il y a un seul couple de la forme (x, y) .

Pour noter E , appelé l'ensemble de départ (ou domaine) de f , et F appelé l'ensemble d'arrivée (ou codomaine) de f , on écrit $f : E \rightarrow F$. Comme nous connaissons déjà les éléments $x \in E$, décrire l'ensemble des couples (x, y) composant Γ (appelé le graphe de f) se fait en donnant la forme de y en fonction de x . Pour cela, on écrit $x \mapsto f(x)$. On appelle antécédent l'élément x et image l'élément $f(x)$.

Donnons quelques exemples de fonctions.

Exemple 2.17.

- $f : \mathbb{R} \rightarrow \mathbb{R}$
 $x \mapsto x^2$
- $g : \mathbb{N} \rightarrow \mathbb{N}$
 $n \mapsto n + 2$
- $0 : \mathbb{R} \rightarrow \mathbb{R}$
 $x \mapsto 0$
- $i : \mathbb{N} \rightarrow \mathbb{R}$
 $x \mapsto x$
- $\text{id}_{\mathbb{N}} : \mathbb{N} \rightarrow \mathbb{N}$
 $n \mapsto n$

Les fonctions sont essentielles car elles nous permettent de définir des liens entre des objets mathématiques. De plus, elles sont l'objet d'étude privilégié de la discipline appelée l'analyse.

Remarque. La fonction $\text{id}_{\mathbb{N}}$, appelée la fonction identité, peut se généraliser à n'importe quel ensemble E :

$$\begin{aligned} \text{id}_E : E &\rightarrow E \\ x &\mapsto x \end{aligned}$$

De plus, le graphe de cette fonction est la diagonale de l'ensemble : $\{(x, x) \mid x \in E\}$.

Enfin, nous allons définir une opération sur les fonctions, qui permet de s'abstraire de l'évaluation d'une fonction en lui donnant un argument :

Définition 2.14 (Composition de fonctions). Soient $f : E \rightarrow F$ et $g : F \rightarrow G$ deux fonctions. On définit $g \circ f$ de la façon suivante :

$$\begin{aligned} g \circ f : E &\rightarrow G \\ x &\mapsto g(f(x)) \end{aligned}$$

Cette relation définit bien une fonction.

Démonstration. Montrons que la relation donnée par le graphe de $g \circ f$ est fonctionnelle. En effet, ce qu'il est essentiel de montrer est que pour $x \in E$, il existe un unique élément $y \in G$ tel que $(g \circ f)(x) = y$.

Soit $x \in E$, alors $(g \circ f)(x) = g(f(x))$, or il existe un unique $y' \in F$ tel que $y' = f(x)$ et il existe un unique $y \in G$ tel que $y = g(y')$ (car f et g sont des fonctions). Donc il existe un unique $y \in G$ tel que $(g \circ f)(x) = y$.

On en déduit que $g \circ f$ est bien définie en tant que fonction. □

La composition de fonctions est, de plus, associative.

Proposition 2.2.4. *Si $f : E \rightarrow F$, $g : F \rightarrow G$ et $h : G \rightarrow H$ sont des fonctions, alors*

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

ce que l'on écrira directement $h \circ g \circ f$.

Démonstration. Le résultat est évident en regardant à un x fixé : les deux fonction s'évaluent en $h(g(f(x)))$. □

Nous allons maintenant étudier des propriétés liées aux fonctions.

Définition 2.15 (Injection, surjection, bijection). Soit $f : E \rightarrow F$ une fonction. On dit que f est injective lorsqu'elle envoie chaque élément sur une image différente. Cela s'écrit formellement

$$\forall x \in E, \forall y \in E, f(x) = f(y) \implies x = y$$

ce qui signifie que seul x a comme image $f(x)$.

On dit que f est surjective lorsque tout élément de F est l'image d'un élément de E par f . Cela se formalise par la phrase

$$\forall y \in F, \exists x \in E, f(x) = y$$

On dit que f est bijective lorsque f est à la fois injective et surjective. Une formulation équivalente à cela est

$$\forall y \in F, \exists! x \in E, f(x) = y$$

ce qui revient aussi à dire que pour chaque élément $y \in F$, il existe une unique solution à l'équation $f(x) = y$.

Donnons maintenant une caractérisation d'une fonction bijective.

Théorème 2.2 (Bijection réciproque). *Soit $f : E \rightarrow F$, f est bijective si et seulement s'il existe $g : F \rightarrow E$ telle que $f \circ g = \text{id}_F$ et $g \circ f = \text{id}_E$.*

Remarque. Ces équations se traduisent de la façon suivante :

$$(\forall x \in F, f(g(x)) = x) \wedge (\forall x \in E, g(f(x)) = x)$$

Démonstration. Montrons l'équivalence par double implication. On va donc montrer dans un premier temps que si une fonction est bijective alors il existe une fonction g telle que décrite plus haut, puis que s'il existe une telle fonction g alors f est bijective.

⇒ Supposons que f est bijective. Alors soit la fonction $g : F \rightarrow E$ définie par le graphe réciproque de $f : \forall x \in E, \forall y \in F, f(x) = y \iff x = g(y)$. On remarque que par construction, $x = g(y) = g(f(x))$ et $y = f(x) = f(g(y))$. Il suffit donc de vérifier que g est bien une fonction. Pour cela, on remarque que f étant une bijection, pour tout $y \in F$, il existe un unique $x \in E$ tel que $f(x) = y$, donc tel que $x = g(y)$. On en déduit donc que pour tout $y \in F$, il existe une unique image $g(y)$, donc g est bien définie en tant que fonction.

⇐ Supposons qu'il existe g telle que $f \circ g = \text{id}_F$ et $g \circ f = \text{id}_E$. Montrons que f est injective. Soient $x, y \in E$ tels que $f(x) = f(y)$. Alors $g(f(x)) = g(f(y))$, ce qui signifie que $x = y$ puisque $\forall a, g(f(a)) = a$. On en déduit que f est injective. Montrons maintenant qu'elle est surjective. Soit $y \in F$, alors $g(y) \in E$ et $f(g(y)) = y$ par hypothèse, donc $g(y)$ est un antécédent de y : tout élément $y \in F$ possède un antécédent, donc f est surjective.

L'équivalence est donc prouvée par double implication.

□

On peut prouver une version plus forte de ce résultat : en effet, on peut remarquer que pour l'injectivité, on utilise l'un des deux sens, et l'autre sens pour la surjectivité. En réalité, on peut caractériser ainsi l'injectivité et la surjectivité.

Exercice 2.13. Soit $f : E \rightarrow F$ une fonction. Montrer que :

- f est injective si et seulement s'il existe une fonction $g : F \rightarrow E$ telle que $g \circ f = \text{id}_E$.
- f est surjective si et seulement s'il existe une fonction $g : F \rightarrow E$ telle que $f \circ g = \text{id}_F$.

Exercice 2.14. Soient $f : E' \rightarrow F$ et $g : F \rightarrow F$. Montrer que :

- si f et g sont injectives, alors $g \circ f$ est injective.
- si f et g sont surjectives, alors $g \circ f$ est surjective.
- si f et g sont bijectives, alors $g \circ f$ est bijective.

Nous donnons en exercice d'approfondissement une autre caractérisation de l'injectivité et de la surjectivité.

Exercice 2.15 (*). Soit $f : E \rightarrow F$ une fonction. Montrer que :

- f est injective si et seulement si pour tout ensemble A et toutes fonctions $g : A \rightarrow E, h : A \rightarrow E$, si $f \circ g = f \circ h$ alors $g = h$.
- f est surjective si et seulement si pour tout ensemble B et toutes fonctions $g : F \rightarrow B, h : F \rightarrow B$, si $g \circ f = h \circ f$ alors $g = h$.

2.2.6 Fonctions et ensembles

Nous allons maintenant définir des opérations sur les fonctions par rapport à des ensembles, et sur des ensembles par rapport à des fonctions.

Définition 2.16 (Restriction, co-restriction). Soit $f : E \rightarrow F$ une fonction, $E' \subseteq E$ et $F' \subseteq F$. On appelle restriction de f à E' , et on note $f|_{E'}$, la fonction définie par le triplet (E', F, Γ'_f) où Γ'_f coïncide avec Γ_f sur E' , ce qui signifie que si $x \in E'$, alors $f(x) = f|_{E'}(x)$. On appelle co-restriction de f à F' , et on note $f|^{F'}$, la fonction définie par le triplet (E, F', Γ) .

Une restriction est toujours définie, mais une co-restriction nécessite de prouver que $\forall x \in E, f(x) \in F'$ pour être définie.

Définition 2.17 (Image directe, image réciproque). Soit $f : E \rightarrow F$ une fonction, soient $E' \subseteq E$ et $F' \subseteq F$. On définit l'image directe de E' par f , notée $f(E')$, par

$$f(E') = \{f(x) \mid x \in E'\}$$

et on définit l'image réciproque de F' par f , notée $f^{-1}(F')$, par

$$f^{-1}(F') = \{x \mid x \in E, f(x) \in F'\}$$

3 Combinatoire et dénombrement

Cette section est centrée sur l'étude du dénombrement, que l'on pourrait résumer en « l'art de compter efficacement ». Nous verrons donc dans un premier temps la notion de cardinal, puis quelques cardinaux classiques, puis nous étudierons les coefficients binomiaux, pierre angulaire du dénombrement.

3.1 Cardinal

Le cardinal d'un ensemble est le nombre d'éléments qui sont dans cet ensemble. Par exemple, le cardinal de $\{0\}$ est 1, puisqu'il y a 1 élément.

Définition 3.1 (Cardinal). Soit E un ensemble. S'il existe un élément $n \in \mathbb{N}$ tel que E est en bijection avec $\{1, \dots, n\}$, alors on dit que E est fini de cardinal n , ce que l'on écrit

$$\text{card}(E) = n$$

Le cardinal est donc une information préservée par des bijections. Ce qui s'exprime par le théorème qui suit.

Théorème 3.1 (Conservation du cardinal). Soient E et F deux ensembles et E de cardinal n , s'il existe une bijection $f : E \rightarrow F$ alors F est de cardinal n .

Démonstration. En effet, soit $g : \{1, \dots, n\} \rightarrow E$ bijective (qui existe par définition du fait que E est de cardinal n). Puisque f est aussi bijective, $f \circ g : \{1, \dots, n\} \rightarrow F$ est bijective, donc par définition F est de cardinal n . \square

Les injections aussi permettent de conserver une partie de l'information.

Théorème 3.2 (Injection et cardinal). Soit E et F deux ensembles, F de cardinal n et $f : E \rightarrow F$ une fonction injective. Alors $\text{card}(E) \leq n$.

Démonstration. Soit $A = f(E) \subseteq F$. On considère la fonction $g = f|_{f(E)}$. Par construction, g est surjective, puisque tout élément de $f(E)$ a un antécédent par f , donc par g . De plus comme f est injective, g l'est aussi. Ainsi $\text{card}(E) = \text{card}(A)$. En considérant la bijection entre F et $\{1, \dots, n\}$, et quitte à effectuer une permutation, on déduit que $\text{card}(A) \leq \text{card}(F)$, d'où $\text{card}(E) \leq n$. \square

Dénombrons maintenant des ensembles d'objets classiques.

3.1.1 Cardinal d'une union disjointe

Proposition 3.1.1. Soient E et F deux ensembles finis tels que $E \cap F = \emptyset$, on a l'égalité

$$\text{card}(E \cup F) = \text{card}(E) + \text{card}(F)$$

Ce résultat peut s'illustrer par le schéma ci-dessous.

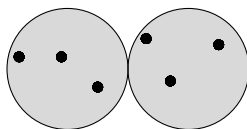


FIGURE 8 – Illustration d'une union disjointe

Démonstration. On va prouver cette propriété par récurrence sur le cardinal de F :

- Si $F = \emptyset$, alors $E \cup F = E$ et $\text{card}(E) + \text{card}(F) = \text{card}(E) + 0 = \text{card}(E)$, donc l'égalité est vérifiée pour $\text{card}(F) = 0$.

- Si $F = F' \cup \{x\}$, donc si $\text{card}(F) = \text{card}(F') + 1$, où par hypothèse de récurrence $\text{card}(E \cup F') = \text{card}(E) + \text{card}(F')$, alors soit $f : \{1, \dots, m\} \rightarrow E \cup F'$ la bijection correspondant au cardinal de $E \cup F'$, nous allons construire une bijection dans $E \cup F$:

On construit $f' : \{1, \dots, m+1\} \rightarrow E \cup F$ en ajoutant que $f'(m+1) = x$. Par définition, comme f est surjective et que $E \cup F = (E \cup F') \cup \{x\}$, f' est surjective. Si $f'(a) = f'(b)$, alors soit $f'(a) \in E \cup F'$, auquel cas l'injectivité de f nous fait déduire que $a = b$, soit $f'(a) = x$, auquel cas seul $m+1$ est un antécédent de x . On en déduit que f' est injective. Donc f est bijective.

Il en découle que $\text{card}(E \cup F) = m+1 = \text{card}(E) + \text{card}(F)$.

Donc, par récurrence, $\boxed{\text{card}(E \cup F) = \text{card}(E) + \text{card}(F)}$. □

L'hypothèse $E \cap F = \emptyset$ est importante pour définir f' ici : l'injectivité provient directement de cette hypothèse (sinon, x aurait pu appartenir à E et l'on n'aurait pas un unique antécédent). L'exercice qui suit est la généralisation au cas où $E \cap F$ est quelconque.

Exercice 3.1 (Cardinal d'une union quelconque). Soient E et F deux ensembles. Montrer que :

$$\text{card}(E \cup F) = \text{card}(E) + \text{card}(F) - \text{card}(E \cap F)$$

3.1.2 Cardinal d'un produit cartésien

Proposition 3.1.2. Soient E et F deux ensembles finis de cardinal respectif n et m . Alors

$$\text{card}(E \times F) = n \times m = \text{card}(E) \times \text{card}(F)$$

Démonstration. Prouvons ce résultat par récurrence sur E :

- Tout d'abord, si $E = \emptyset$ alors $E \times F = \emptyset$ donc l'égalité est vérifiée.
- Supposons que $E = E' \cup \{a\}$. Par définition,

$$E \times F = (E' \times F) \cup \{(a, x) \mid x \in F\}$$

or $\{(a, x) \mid x \in F\}$ est de cardinal F (la bijection est évidente), donc

$$\text{card}(E \times F) = \text{card}(E') \times \text{card}(F) + \text{card}(F) = (\text{card}(E') + 1) \times \text{card}(F) = \text{card}(E) \times \text{card}(F)$$

(le passage de \cup à $+$ se fait car aucun élément de $E' \times F$ ne contient a en première coordonnée)

Donc $\boxed{\text{l'égalité est démontrée par récurrence.}}$ □

Remarque. On comprend alors la notation \times pour désigner le produit cartésien.

Exercice 3.2 (Le lemme des bergers). Montrer que si $f : E \rightarrow F$ est tel qu'il existe k tel que pour tout $y \in F$, $f^{-1}(y) = k$, alors on a la relation

$$E = F \times k$$

3.1.3 Ensemble de fonctions

Nous allons ici travailler avec des ensembles de fonctions. Pour cela, nous allons commencer par définir des notations :

$$F^E = \{f : E \rightarrow F\} \quad \mathcal{F}(E, F) = F^E$$

Ces notations permettent de se demander combien il existe de fonctions $f : E \rightarrow F$.

Proposition 3.1.3. Soient E et F deux ensembles de cardinal respectif n et m . Alors

$$\text{card}(F^E) = \text{card}(F)^{\text{card}(E)}$$

Démonstration. Nous allons dénombrer les triplets (E, F, Γ) où Γ est un graphe fonctionnel. Par définition, E et F étant fixé, il y a autant de tels triplets que de graphes fonctionnels Γ . Nous allons alors raisonner par récurrence sur le cardinal de E :

- Si $E = \emptyset$, alors le seul graphe possible est le graphe vide, auquel cas $F^E = \{(\emptyset, F, \emptyset)\}$, ce qui vérifie l'égalité.
- Si $E = E' \cup \{a\}$ et $\text{card}(F^{E'}) = \text{card}(F)^{\text{card}(E')}$, alors on peut définir

$$\begin{aligned} \pi : \mathcal{F}(E, F) &\longrightarrow \mathcal{F}(E', F) \\ f &\longmapsto f|_{E'} \end{aligned}$$

Dans ce cas, $\pi(f) = \pi(f')$ si et seulement si f et f' ne diffèrent que sur leur image par a . Or il y a $\text{card}(F)$ images possibles pour a , donc par le lemme des bergers :

$$\text{card}(F^E) = \text{card}(F)^{\text{card}(E')} \times \text{card}(F) = \text{card}(F)^{\text{card}(E')+1} = \text{card}(F)^{\text{card}(E)}$$

□

Remarque. Là encore, la notation F^E prend son sens en regardant le cardinal.

Exercice 3.3 (Dénombrer les parties d'un ensemble). Soit E un ensemble.

- Montrer qu'il y a une bijection entre $\mathcal{P}(E)$ et $\mathcal{F}(E, \{0, 1\})$. *Indication :* Pour un sous-ensemble $E' \subseteq E$, la fonction associée, notée $\chi_{E'}$, envoie les éléments de E' sur 1 et les autres sur 0.
- En déduire que $\text{card}(\mathcal{P}(E)) = 2^{\text{card}(E)}$.

Remarque. On note parfois aussi l'ensemble $\mathcal{P}(E)$ par 2^E .

Exercice 3.4 (Sur les fonctions caractéristiques *). Soit E un ensemble. On définit

$$\begin{aligned} \chi : \mathcal{P}(E) &\longrightarrow \mathcal{F}(E, \{0, 1\}) \\ E' &\longmapsto \chi_{E'} \end{aligned}$$

Où $\chi_{E'}$ est telle qu'indiquée dans l'exercice précédent.

- Montrer que $\forall x \in E, \chi_{E' \cap E''}(x) = \chi_{E'}(x) \times \chi_{E''}(x)$.
- Montrer que $\forall x \in E, \chi_{E' \cup E''}(x) = \chi_{E'}(x) + \chi_{E''}(x) - \chi_{E' \cap E''}(x)$.
- Montrer que $\forall x \in E, \chi_{E \setminus E'}(x) = 1 - \chi_{E'}(x)$.

3.1.4 Permutations et arrangements

Nous aurons besoin pour les prochains calculs de dénombrement d'une nouvelle notation : celle des factorielles.

Définition 3.2 (Factorielle). Soit n un entier. On appelle factorielle de n , que l'on note $n!$, l'entier

$$n! = \prod_{i=1}^n i = 1 \times 2 \times \dots \times n$$

avec la convention que $0! = 1$.

Définition 3.3 (Arrangement). Soit E et F des ensembles finis. On note A_F^E le nombre d'injections (ou fonctions injectives) de E vers F . Ce nombre vaut

$$A_F^E = \frac{(\text{card}(F))!}{(\text{card}(F) - \text{card}(E))!}$$

si $\text{card}(E) \leq \text{card}(F)$ et 0 sinon.

On note A_k^n , appelé l'arrangement de k parmi n , le nombre d'injections d'un ensemble à k éléments dans un ensemble à n éléments.

Démonstration. Nous ne donnerons qu'une idée de la preuve, le lecteur assidu rédigera les morceaux manquants. L'ensemble des injections peut se définir par récurrence : on a 1 fonction depuis l'ensemble vide, puis $\text{card}(F)$ choix pour l'image du premier élément, puis $\text{card}(F) - 1$ choix pour l'image du deuxième éléments, etc. jusqu'à avoir $\text{card}(E)$ images choisies dans F . □

Nous allons définir maintenant le nombre de permutations.

Définition 3.4 (Permutation). Soit E un ensemble fini. On appelle l'ensemble des permutations de E l'ensemble $\mathcal{Bij}(E, E)$ des bijections de E dans E , et l'on le note $\mathfrak{S}(E)$, ou encore \mathfrak{S}_n lorsque $E = \{1, \dots, n\}$ (une permutation de \mathfrak{S}_n peut s'interpréter comme un mélange des entiers de 1 à n). De plus,

$$\text{card}(\mathfrak{S}_n) = n!$$

Démonstration. L'idée de la preuve est d'utiliser le fait qu'une bijection est injective pour obtenir $n!$ permutations. \square

3.2 Sur les coefficients binomiaux

Cette partie se concentrera sur les coefficients binomiaux, qui sont une part importante du dénombrement. Nous allons d'abord définir les coefficients binomiaux puis nous en donnerons des propriétés importantes, avant de donner des exemples d'utilisations classiques de ces nombres (principalement le binôme de Newton).

Définition 3.5. Soient k et n deux entiers naturels tels que $k \leq n$. On note $\binom{n}{k}$, que l'on lit k parmi n , le nombre

$$\binom{n}{k} = \{E \mid E \subseteq \{1, \dots, n\}, \text{card}(E) = k\}$$

C'est donc l'ensemble des parties à k éléments d'un ensemble à n éléments.

3.2.1 Propriétés

Nous allons maintenant donner différentes propriétés des coefficients binomiaux.

Proposition 3.2.1. Soient $k \leq n$ deux entiers naturels. On a

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

Démonstration. La bijection entre les parties à k éléments et les parties à $n-k$ éléments est la fonction définie par $E \mapsto \{1, \dots, n\} \setminus E$. \square

Cette propriété revient à dire que choisir $n-k$ éléments revient à choisir les k éléments à ne pas inclure dans la partie que l'on construit.

Proposition 3.2.2. Soient $k < n$. Alors

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

et

$$\binom{n}{0} = 1$$

Démonstration. Il n'y a, d'abord, qu'une partie à 0 éléments : c'est \emptyset . On en déduit que $\binom{n}{0} = 1$.

Soit une partie E à $k+1$ éléments. Alors soit $n+1 \in E$, soit $n+1 \notin E$, et ces deux cas sont disjoint. Traitons alors chaque cas :

- Si $n+1 \in E$, alors on a une bijection entre les telles parties E et les parties à k éléments : à E on associe $E \setminus \{n+1\}$ et la bijection réciproque est donnée par l'application qui à E associe $E \cup \{n+1\}$. Donc il y a $\binom{n}{k}$ telles parties.

- Si $n \notin E$, alors on a une bijection entre les telles parties E et les parties à $k+1$ éléments, puisque ces parties sont aussi, alors, des parties de $\{1, \dots, n\}$ à $k+1$ éléments. Réciproquement, toute partie de $k+1$ éléments de $\{1, \dots, n\}$ est une partie à $k+1$ éléments de $\{1, \dots, n+1\}$. Donc il y a $\binom{n}{k+1}$ telles parties.

Puisque l'union est disjointe, on en déduit que $\boxed{\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1} \text{ et } \binom{n}{0} = 1.}$ \square

Remarque. A partir des deux propriétés précédentes, on peut en déduire que $\binom{n}{n} = 1$. On peut ainsi pour chaque entier trouver directement la valeur de $\binom{n}{n}$ et de $\binom{n}{0}$, et calculer récursivement la valeur d'un $\binom{n}{k}$ quelconque à partir de là. C'est ce qu'on appelle le triangle de Pascal, du nom de son inventeur et à cause de la forme qu'a l'ensemble des coefficients binomiaux lorsque l'on les donne par liste (c.f. le dessin ci-dessous).

$$\begin{array}{ccccccc}
 & & & & 1 & & & & \\
 & & & & 1 & & 1 & & \\
 & & & 1 & & 2 & & 1 & \\
 & & 1 & & 3 & & 3 & & 1 \\
 & 1 & & 4 & & 6 & & 4 & & 1
 \end{array}$$

FIGURE 9 – Triangle de Pascal pour $n \leq 4$.

Nous allons maintenant voir une formule explicite pour le calcul d'un coefficient binomial.

Proposition 3.2.3 (Formule explicite). *Soient $k \leq n$ deux entiers. Alors*

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

Démonstration. Soit une partie $E \subseteq \{1, \dots, n\}$ de cardinal k . On considère l'ensemble des injections de E dans $\{1, \dots, n\}$: il y en a $\frac{n!}{(n-k)!}$. Or une injection est exactement une liste d'éléments de $\{1, \dots, n\}$ (en comptant l'ordre), et deux listes donnent le même ensemble d'arrivée si et seulement s'il existe une permutation faisant passer de la première liste à la deuxième. Il y a $k!$ permutations au sein d'une liste, et on a donc $k!$ listes possibles de taille k pour un ensemble fixé. On a donc $k \times \binom{n}{k} = \frac{n!}{(n-k)!}$, ce qui revient à dire $\boxed{\binom{n}{k} = \frac{n!}{k!(n-k)!}.}$ \square

Les exercices qui suivent sont importants pour comprendre l'intérêt d'avoir défini les coefficients binomiaux.

Exercice 3.5. Soient a, b deux nombres réels et n un entier naturel. Montrer que

$$(a+b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k} = b^n + nab^{n-1} + \dots + na^{n-1}b + a^n$$

Indication : Raisonner par récurrence sur n .

Enfin, donnons un résultat important, appelé lemme des tiroirs, dans sa version finie et sa version infinie.

Lemme 3.2.1 (Tiroirs finis). *Soient E et F deux ensembles finis tels que $\text{card}(E) > \text{card}(F)$, et $f : E \rightarrow F$. Alors il existe deux éléments $x, y \in E$ tels que $f(x) = f(y)$.*

Démonstration. Ce résultat est la contraposée du fait que s'il existe une fonction injective $f : E \rightarrow F$ alors $\text{card}(E) \leq \text{card}(F)$. Puisque $\text{card}(E) > \text{card}(F)$, toute fonction $f : E \rightarrow F$ est non injective, d'où le résultat. \square

Lemme 3.2.2 (Tiroirs infinis). *Soient E et F respectivement un ensemble infini et un ensemble fini, et $f : E \rightarrow F$. Il existe un élément $y \in F$ tel que $f^{-1}(\{y\})$ est infini.*

Démonstration. Supposons que pour tout élément $y \in F$, on ait $f^{-1}(\{y\})$ fini. Alors comme F est fini, on peut considérer $k = \max(f^{-1}(\{y\}))$, et en appliquant le lemme des bergers, on en déduit que le cardinal de E est fini, et inférieur à $k \times \text{card}(F)$, ce qui est une contradiction. \square

Deuxième partie

Géométrie

4 Construction des vecteurs

Nous allons dans cette section définir formellement les vecteurs et en donner les propriétés immédiates. Donnons d'abord les notions que l'on considère comme étant axiomatique, c'est-à-dire acceptées comme connues sans devoir être définie, dans le cadre de la géométrie affine (c'est la géométrie étudiant les points du plan et leurs transformations).

Nous admettrons l'existence de la relation de parallélisme ainsi que ses propriétés, la perpendicularité et ses propriétés, ainsi que la notion d'angle et de distance entre deux points, permettant donc de définir par exemple la médiatrice d'un segment ou le milieu de deux points.

De plus, on notera \mathcal{P} le plan, qui est considéré comme un ensemble de points, et \mathcal{D} l'ensemble des droites du plan : $\mathcal{D} \subseteq 2^{\mathcal{P}}$.

Un vecteur est un déplacement rectiligne du plan. On peut le voir comme un élément de $\mathcal{F}(\mathcal{P}, \mathcal{P})$ ou comme un ensemble de couples du plan, ce que l'on va considérer ici. Un vecteur est défini par sa norme, sa direction et son sens. On pourrait alors définir une relation d'équivalence sur les couples de points en cherchant à définir norme, direction et sens, de sorte que deux couples appartiennent au même vecteur s'ils définissent la même norme, direction et sens. Cependant, nous allons plutôt utiliser une propriété annexe des vecteurs, qui est que si $\overrightarrow{AB} = \overrightarrow{CD}$ alors $(ABDC)$ est un parallélogramme, et réciproquement. Définissons donc un vecteur à partir de cette propriété.

Définition 4.1 (Bipoint). On appelle « bipoint » un élément de \mathcal{P}^2 , qui est donc l'ensemble des bipoints. Ce bipoint (A, B) représente le segment orienté $[\overrightarrow{AB}]$.

Nous allons donc définir la relation d'équivalence signifiant moralement « ces deux bipoints définissent le même vecteur » :

Définition 4.2 (Equipollence). On définit la relation \sim d'équivalence sur les bipoints par :

$$(A, B) \sim (C, D) \iff I_{AD} = I_{BC}$$

où I_{AD} désigne le milieu du segment $[AD]$.

Cette définition est équivalente à dire que $(ABDC)$ est un parallélogramme.

Démonstration. Nous devons donc montrer que la relation \sim est une relation d'équivalence :

- $\boxed{(A, B) \sim (A, B)}$ puisque $I_{AB} = I_{AB}$.
- si $(A, B) \sim (C, D)$ alors par symétrie de l'égalité, $\boxed{(C, D) \sim (A, B)}$.
- si $(A, B) \sim (C, D)$ et $(C, D) \sim (E, F)$, alors on sait que (AB) et (EF) sont parallèles à (CD) , ce qui signifie que les deux droites sont parallèles entre elles. Nous allons maintenant utiliser la réciproque du théorème de Thalès généralisé, en utilisant le fait que $\frac{AC}{BD} = \frac{CE}{DF} = \frac{EA}{FB} = 1$ (de par le fait que $(ABDC)$ et $(EFDC)$ sont des parallélogrammes) et que (AC) et (BD) sont parallèles ainsi que (CE) et (DF) . On en déduit donc que (AE) et (BF) sont parallèles. Ainsi $(ABFE)$ est un parallélogramme, d'où $\boxed{(A, B) \sim (E, F)}$.

□

On définit alors \mathcal{V} l'ensemble des vecteurs du plan.

Définition 4.3 (Vecteur). L'ensemble \mathcal{V} est défini par

$$\mathcal{V} = \mathcal{P}^2 / \sim$$

On met donc dans la même classe deux bipoints qui désignent un mouvement dans la même direction et le même sens d'une longueur égale.

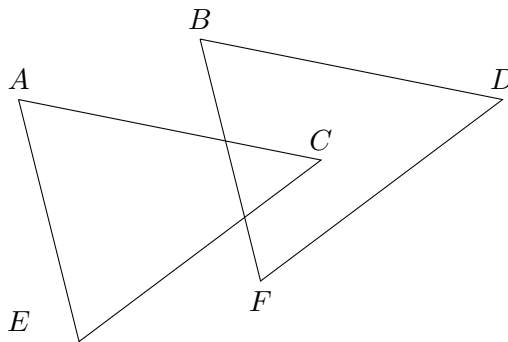


FIGURE 10 – Illustration du cas de la transitivité

4.1 Propriétés d'un vecteur

Nous allons maintenant montrer qu'un vecteur définit le graphe d'une fonction du plan.

Définition 4.4. Soit \vec{u} un vecteur. $(\mathcal{P}, \mathcal{P}, \vec{u})$ est une fonction qui à un point A associe un point B qu'on nommera l'image de A par le vecteur \vec{u} . On notera cela

$$B = \vec{u} A$$

Démonstration. Il faut donc prouver que si $(A, B) \sim (A, B')$ alors $B = B'$. Par hypothèse, $(ABB'A)$ est un parallélogramme, donc $BB' = AA = 0$, donc $B' = B$. Il n'y a donc qu'un couple possible dans une classe d'équivalence donnée, donc $(\mathcal{P}, \mathcal{P}, \vec{u})$ est bien une fonction du plan. \square

Nous allons montrer qu'un vecteur est caractérisé par sa norme, sa direction et son sens. Pour cela, nous allons montrer un lemme (c'est-à-dire un résultat intermédiaire).

Lemme 4.1.1. Soit \vec{u} un vecteur. Alors l'ensemble des bipoints appartenant à \vec{u} forment des segments de même longueur et des droites toutes parallèles entre elles.

Démonstration. Cela découle directement des propriétés d'un parallélogramme : si $(A, B) \sim (C, D)$ alors le parallélogramme $(ABDC)$ possède des côtés opposés de même longueur, donc $AB = CD$. De plus, les côtés opposés étant parallèles, (AB) et (CD) sont parallèles. \square

Il en découle qu'on peut définir la norme d'un vecteur, notée $\|\vec{u}\|$ comme étant la distance entre les points d'un bipoint du vecteur, et la direction comme l'ensemble des droites parallèles qui prolongent les segments des bipoints. Avoir le même sens signifie pour deux demi-droites parallèles que l'une des deux possède l'ensemble de ses projetés orthogonaux sur l'autre droite inclus dans la deuxième demi-droite.

Proposition 4.1.1. Soit \vec{u} et \vec{u}' de même norme, même direction et même sens. Alors $\vec{u} = \vec{u}'$.

Démonstration. Soit un point A , montrons que $\vec{u} A = \vec{u}' A$. Soit B et C respectivement l'image par \vec{u} et l'image par \vec{u}' de A . Soit alors le quadrilatère $(ABCA)$. Puisque \vec{u} et \vec{u}' ont la même direction, on en déduit que B et C sont alignés avec A . De plus, comme les vecteurs ont la même norme, soit ils sont de part et d'autre de A et A est le milieu de $[BC]$, soit ils sont superposés. Or puisque les vecteurs ont même sens, ils ne peuvent pas être de part et d'autre de A . Donc $B = C$, donc $\vec{u} = \vec{u}'$. \square

4.2 Addition et multiplication de vecteurs

Nous allons définir l'addition de deux vecteurs et la multiplication d'un vecteur par un nombre réel (appelé aussi scalaire).

Définition 4.5. Soient \vec{u} et \vec{v} deux vecteurs. On définit $\vec{u} + \vec{v}$ par

$$\vec{u} + \vec{v} = \vec{u} \circ \vec{v}$$

en considérant les vecteurs comme des fonctions.

Nous allons montrer plusieurs propriétés de cette addition de vecteurs.

Proposition 4.2.1. L'addition de vecteurs est commutative, c'est-à-dire que pour tous vecteurs \vec{u} , \vec{v} , on a $\vec{u} + \vec{v} = \vec{v} + \vec{u}$.

Proposition 4.2.2. Soient B et C les images d'un point A fixé par \vec{u} et \vec{v} . De plus, soit D l'image par \vec{u} de C . Montrons que $D = \vec{v}B$.

Par définition, $(ABDC)$ est un parallélogramme (puisque constitué de bipoints de \vec{u}). Ceci signifie que AC et BD sont parallèles, de même longueur, et de même sens. On en déduit donc que $\vec{v} = \vec{AC} = \vec{BD}$, donc que $D = \vec{v}B$. Ainsi $\boxed{\vec{u} + \vec{v} = \vec{v} + \vec{u}}$.

De plus, comme la composition est associative, l'addition l'est.

Nous allons montrer que le vecteur $\{(x, x) \mid x \in \mathcal{P}\}$, appelé vecteur nul et noté $\vec{0}$, est neutre pour l'addition.

Proposition 4.2.3. Pour tout $\vec{u} \in \mathcal{V}$, $\vec{0} + \vec{u} = \vec{u}$.

Démonstration. Notons B l'image d'un point A fixé. Alors $\vec{0}B = B$ par définition, ce qui signifie donc que $\boxed{\vec{0} + \vec{u} = \vec{u}}$. \square

Cette addition de vecteur possède une propriété appelée la relation de Chasles, elle dit qu'un mouvement d'un point A à un point B puis du point B à un point C équivaut à un mouvement du point A au point C directement.

Proposition 4.2.4 (Relation de Chasles). Soient A, B, C trois points du plan. L'identité suivante est vérifiée :

$$\vec{AB} + \vec{BC} = \vec{AC}$$

Démonstration. Par définition, la fonction donnée par $\vec{AB} + \vec{BC}$ envoie A sur C , comme \vec{AC} , donc ces deux vecteurs ont même norme, même direction et même sens : $\boxed{\text{ils sont égaux.}}$ \square

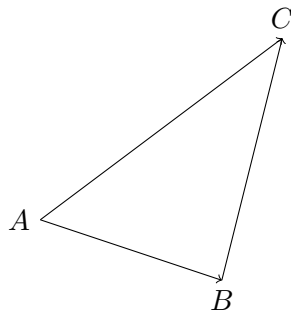


FIGURE 11 – Illustration de la relation de Chasles

De plus, un vecteur \vec{u} possède ce qu'on appelle un opposé.

Proposition 4.2.5. Pour tout \vec{u} , il existe un unique \vec{v} tel que $\vec{u} + \vec{v} = \vec{0}$. On note ce vecteur $-\vec{u}$.

De plus, $-\vec{AB} = \vec{BA}$.

Démonstration. Soit \vec{u} un vecteur. On pose $-\vec{u}$ comme ayant la même direction, la même norme mais le sens opposé : on en déduit que $A = -\vec{u}(\vec{u}A)$ ce qui signifie que la composée des fonctions représentant \vec{u} et $-\vec{u}$ vaut le vecteur $\vec{AA} = \vec{0}$. Donc on a trouvé $-\vec{u}$ tel que $\vec{u} + (-\vec{u}) = \vec{0}$.

Ce vecteur est unique car si on avait \vec{v} un autre opposé, alors

$$\vec{v} = -\vec{u} + \vec{u} + \vec{v} = -\vec{u} + \vec{0} = -\vec{u}$$

donc l'opposé d'un vecteur est unique.

Enfin, on remarque que $\vec{AB} + \vec{BA} = \vec{0}$ par la relation de Chasles. □

Nous allons désormais définir la multiplication d'un vecteur par un scalaire.

Définition 4.6. Soit \vec{u} un vecteur et $k \in \mathbb{R}$. On définit $k\vec{u}$ comme l'unique vecteur de même direction et même sens et de norme $k\|\vec{u}\|$.

Démonstration. Ce vecteur est bien défini par la caractérisation d'un vecteur par sa norme, sa direction et son sens. □

Exercice 4.1. Soit \vec{u} et \vec{v} deux vecteurs, k et k' deux réels. Montrer :

- $(k \times k')\vec{u} = k(k'\vec{u})$
- $k(\vec{u} + \vec{v}) = k\vec{u} + k\vec{v}$
- $(k + k')\vec{u} = k\vec{u} + k'\vec{u}$
- $1\vec{u} = \vec{u}$

4.3 Système de coordonnées

Nous allons maintenant étudier les repères et les bases du plan. Ceux-ci permettent de décrire de façon numérique les points du plan. Nous verrons d'abord ce qu'est une base puis ce qu'est un repère. Pour ce faire, nous allons commencer par définir la notion de colinéarité pour des vecteurs.

Définition 4.7. On dit que deux vecteurs \vec{u} et \vec{v} sont colinéaires quand il existe un réel k tel que $\vec{u} = k\vec{v}$. Cela est équivalent à dire que \vec{u} et \vec{v} ont même direction.

Démonstration. Raisonnons par double implication.

- Si $\vec{u} = k\vec{v}$, alors par définition \vec{u} et \vec{v} ont la même direction, un sens opposé si $k < 0$ et même sens si $k \geq 0$. Donc les deux vecteurs ont la même direction.
- Si les deux vecteurs ont la même direction, alors soit $\vec{u} = \vec{v}$ auquel cas $k = 1$, soient ils sont différents, auquel cas on raisonne par disjonction de cas :

- Si les deux vecteurs ont le même sens, alors on pose $k = \frac{\|\vec{u}\|}{\|\vec{v}\|}$, donc $k\vec{v}$ a la même direction, le même sens et la même norme que \vec{u} : ils sont égaux.
- De la même façon, si les deux vecteurs ont des sens opposés, alors on prend l'opposé de la valeur donnée plus tôt.

Dans tous les cas, on a trouvé $k \in \mathbb{R}$, $\vec{u} = k\vec{v}$.

Les deux propositions sont donc équivalentes. □

Remarque. Le vecteur nul est donc colinéaire à tous les vecteurs, en prenant $k = 0$.

Définition 4.8. On appelle base du plan \mathcal{P} un couple (\vec{u}, \vec{v}) où \vec{u} et \vec{v} ne sont pas colinéaires.

Une base est importante car elle permet d'écrire n'importe quel vecteur.

Proposition 4.3.1. Soit $\mathcal{B} = (\vec{u}, \vec{v})$ une base du plan. Alors pour tout vecteur \vec{w} il existe un unique couple (x, y) tel que $\vec{w} = x\vec{u} + y\vec{v}$.

Démonstration. Soit A un point quelconque, $B = \overrightarrow{w}A$, $C = \overrightarrow{v}A$, $D = \overrightarrow{u}A$. On trace les droites (AD) et (AC) puis les parallèles à (AD) passant par B et à (AC) passant par B . On note respectivement J et I les points d'intersections. Cette construction forme un parallélogramme, donc $\overrightarrow{w} = \overrightarrow{AI} + \overrightarrow{AJ}$ et comme ADI sont alignés et ACJ , on en déduit qu'il existe k_1, k_2 tels que $\overrightarrow{AI} = k_1 \overrightarrow{u}$ et $\overrightarrow{AJ} = k_2 \overrightarrow{v}$. D'où $\boxed{\overrightarrow{w} = k_1 \overrightarrow{u} + k_2 \overrightarrow{v}}$.

De plus, k_1 et k_2 sont uniques. En effet, s'il existe une autre décomposition alors on trouve un parallélogramme en A et B passant par C et D : il est unique et c'est déjà $(AIBJ)$. Donc les facteurs d'une autre décomposition sont k_1 et k_2 , donc ces facteurs sont uniques. \square

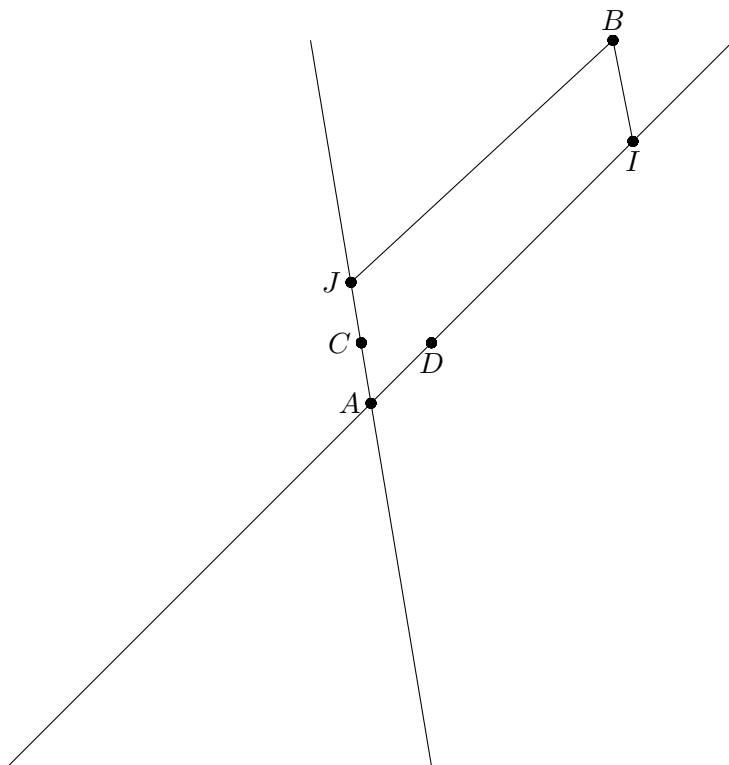


FIGURE 12 – Illustration de la décomposition dans une base

Une base permet donc de décomposer n'importe quel vecteur comme somme de deux vecteurs particuliers, qui sont colinéaires aux vecteurs de la base. On notera $(x_{\overrightarrow{u}}, y_{\overrightarrow{u}})$ les coefficients tels que $\overrightarrow{u} = x_{\overrightarrow{u}} \overrightarrow{e}_1 + y_{\overrightarrow{u}} \overrightarrow{e}_2$ pour une base $(\overrightarrow{e}_1, \overrightarrow{e}_2)$.

Exercice 4.2. Soient \overrightarrow{u} et \overrightarrow{v} deux vecteurs. Montrer que

$$x_{\overrightarrow{u} + \overrightarrow{v}} = x_{\overrightarrow{u}} + x_{\overrightarrow{v}} \quad y_{\overrightarrow{u} + \overrightarrow{v}} = y_{\overrightarrow{u}} + y_{\overrightarrow{v}}$$

Montrer de plus que

$$x_{k\overrightarrow{u}} = kx_{\overrightarrow{u}} \quad y_{k\overrightarrow{u}} = ky_{\overrightarrow{u}}$$

Nous pouvons alors définir un repère : celui-ci sert à repérer des points du plan, et à donner des coordonnées non pas aux vecteurs mais aux points directement.

Définition 4.9. On appelle repère un triplet (O, e_1, e_2) où O est un point fixé, nommé l'origine du repère, et $(\overrightarrow{e}_1, \overrightarrow{e}_2)$ forme une base.

Proposition 4.3.2. Pour tout point A , il existe un unique couple de réels (k_1, k_2) tel que $\overrightarrow{OA} = k_1 \overrightarrow{e}_1 + k_2 \overrightarrow{e}_2$.

Démonstration. La démonstration est évidente en utilisant la propriété précédente. \square

Nous allons maintenant voir les notions de base orthogonale et normée, puis étudier certaines propriétés des vecteurs liées aux coordonnées.

Définition 4.10. Soit \mathcal{B} une base. On dit que \mathcal{B} est orthogonale si les deux vecteurs qui la constituent ont des directions perpendiculaires. On dit que \mathcal{B} est normée si les vecteurs qui la constituent sont de norme 1. Une base à la fois orthogonale et normée est appelée une base orthonormée.

Proposition 4.3.3. Soit \vec{u} un vecteur, de coordonnées (x, y) et \vec{v} de coordonnées (x', y') . Alors $\vec{u} + \vec{v}$ a pour coordonnées $(x + x', y + y')$. De plus, $-\vec{u}$ a pour coordonnées $(-x, -y)$.

Démonstration. On a les égalités $\vec{u} = x \vec{e}_1 + y \vec{e}_2$ et $\vec{v} = x' \vec{e}_1 + y' \vec{e}_2$, donc

$$\vec{u} + \vec{v} = (x + x') \vec{e}_1 + (y + y') \vec{e}_2$$

grâce à la troisième égalité de l'exercice 4.1.

On peut aussi montrer que le vecteur nul a comme coordonnées $(0, 0)$: il vaut $0 \vec{e}_1 + 0 \vec{e}_2$.

Alors en notant (x', y') les coordonnées de $-\vec{u}$, on sait en passant par les coordonnées que $x + x' = 0$ et $y + y' = 0$, donc $x' = -x$ et $y' = -y$. D'où le fait que les coordonnées de $-\vec{u}$ sont $(-x, -y)$. \square

Proposition 4.3.4. Soit \vec{u} un vecteur, de coordonnées (x, y) . Alors $k \vec{u}$ a comme coordonnées (kx, ky) .

Démonstration. En exercice pour le lecteur assidu. \square

Proposition 4.3.5. Si \vec{AB} est un vecteur, que (x_A, y_A) sont les coordonnées de A et (x_B, y_B) sont celles de B , alors les coordonnées de \vec{AB} sont $(x_B - x_A, y_B - y_A)$.

Démonstration. On remarque que $\vec{AB} = \vec{AO} + \vec{OB}$ par la relation de Chasles, or \vec{OA} a pour coordonnées celles de A , et $\vec{OA} = -\vec{AO}$, donc

$$\vec{AB} = \vec{OB} - \vec{OA}$$

ce qui nous donne bien que les coordonnées de \vec{AB} sont $(x_B - x_A, y_B - y_A)$. \square

Proposition 4.3.6. Soit \vec{AB} un vecteur. Si la base \mathcal{B} est orthonormée et que les coordonnées du vecteur sont (x, y) dans cette base, alors la norme du vecteur s'exprime :

$$\|\vec{AB}\| = \sqrt{x^2 + y^2}$$

Démonstration. Tout d'abord, on note $C = (x \vec{e}_1)A$. Puisque la base est orthonormée, le triangle (ACB) est rectangle en C . Le théorème de Pythagore nous permet alors de déduire que $\|\vec{AB}\| = \sqrt{\|\vec{AC}\|^2 + \|\vec{CB}\|^2}$ or par colinéarité avec les vecteurs de la base, normés, et en connaissant les coordonnées de \vec{AB} , on en déduit que $\|\vec{AB}\| = \sqrt{x^2 + y^2}$. \square

4.4 Déterminant de deux vecteurs

Nous allons maintenant voir ce que l'on appelle le déterminant de deux vecteurs.

Définition 4.11. Soient \vec{u} et \vec{v} , deux vecteurs, \mathcal{B} une base dans laquelle les vecteurs ont respectivement comme coordonnées (x, y) et (x', y') . On appelle déterminant, que l'on note $\det(\vec{u}, \vec{v})$ ou encore $[\vec{u}, \vec{v}]$, la quantité

$$\det(\vec{u}, \vec{v}) = xy' - x'y$$

Cet outil nous donne un critère de colinéarité puissant.

Proposition 4.4.1. En reprenant les notations précédentes, \vec{u} et \vec{v} sont colinéaires si et seulement si $\det(\vec{u}, \vec{v}) = 0$.

Démonstration. En effet, si l'un des vecteurs est multiple de l'autre, on vérifie directement que le déterminant des deux vecteurs est nul. Réciproquement, si le déterminant est nul, alors on en déduit que $xy' = x'y$, ce qui signifie que soit $(y, y') = (0, 0)$, auquel cas les deux vecteurs sont colinéaires, soit en divisant par y et y' , $\frac{x}{y} = \frac{x'}{y'}$. Ce qui montre que les deux vecteurs sont proportionnels, donc qu'ils sont colinéaires. \square

Exercice 4.3. Soient deux droites passant respectivement par A et B , et par C et D . Montrer que les deux droites sont parallèles si et seulement si $\det(\vec{AB}, \vec{CD}) = 0$.

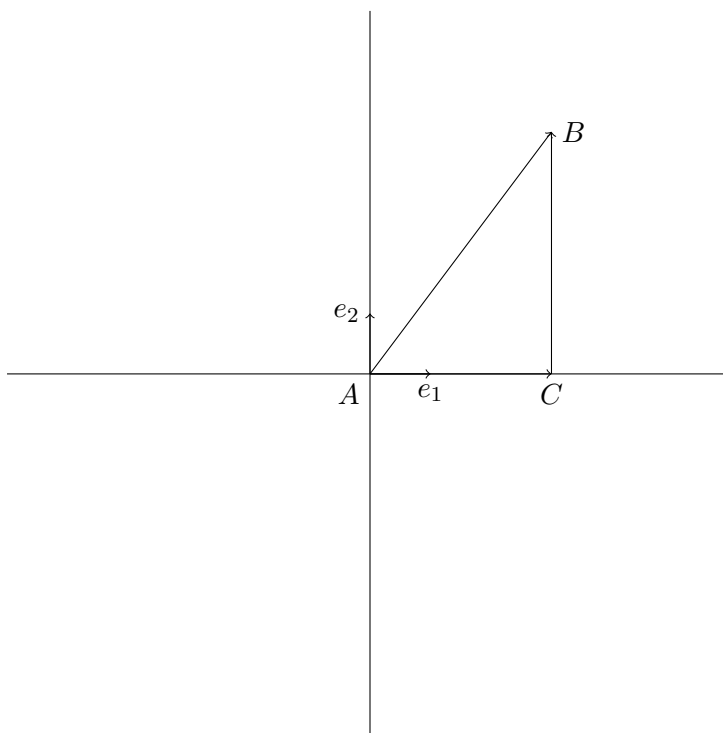


FIGURE 13 – Schéma d'une base orthonormée et de la décomposition d'un vecteur dans cette base

5 Angles et trigonométrie

Nous allons ici traiter des angles et de leur rapport aux longueurs. Nous commencerons par définir ce que nous appellerons un angle, bien qu'une part importante de ces définitions ne peut pas être explicitée (une définition parfaitement rigoureuse des angles demande d'étudier en profondeur les transformations du plan, et une définition rigoureuse de la trigonométrie demande des outils très poussés d'analyse). Cependant, nous mettrons ici l'accent sur l'utilisation de radians plutôt que de degrés pour mesurer les angles. En effet, cette mesure est bien plus efficace en mathématiques, puisqu'elle n'est qu'une mesure de distance.

Définition 5.1 (Angle). Soient deux vecteurs \vec{u} et \vec{v} . On appelle angle entre \vec{u} et \vec{v} , et l'on note (\vec{u}, \vec{v}) , l'angle formé en appliquant les deux vecteurs à un point A quelconque et en prolongeant les demi-droites ainsi créées. On appelle mesure de l'angle la longueur de l'arc de cercle de rayon 1 compris entre ces deux demi-droites, et l'unité de ces angles s'appelle les radians (cf figure 14).

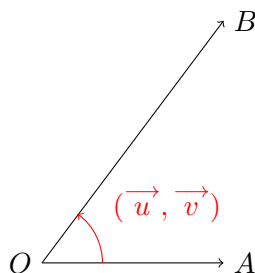


FIGURE 14 – La figure décrite par la définition

5.1 Cercle trigonométrique

Cette partie se concentrera sur l'étude du cercle trigonométrique, qui est une construction permettant de comprendre visuellement ce que mesure la trigonométrie. Ce cercle est un cercle centré en une origine O fixée (donc de coordonnées $(0,0)$) de rayon 1 sur lequel s'enroule la droite des réels : à un réel on associe un point du cercle, que l'on assimile avec la mesure de l'angle formé entre l'axe des abscisses et le segment reliant l'origine au point. On oriente l'angle : un angle dans un sens inverse de celui partant de A et allant vers le haut est négatif, et un angle dans le même sens est positif. Dans le dessin, on indique un angle θ quelconque, le segment en rouge est celui qui relie l'origine au point de la droite réelle (et les deux points rouges sont, de gauche à droite, le point θ sur le cercle et le point θ sur la droite des réels).

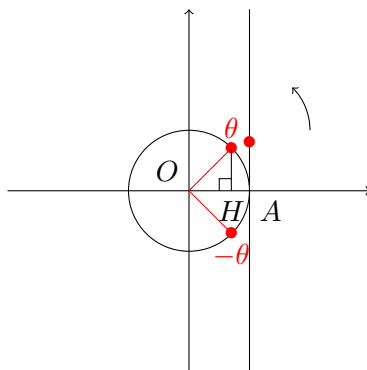


FIGURE 15 – Cercle trigonométrique

Définition 5.2 (Sinus, cosinus, tangente). Soit θ un point du cercle trigonométrique (d'angle θ). On définit les coordonnées de θ par $(\cos(\theta), \sin(\theta))$.

De façon équivalente, on définit dans le triangle rectangle en H , $(OH\theta)$, les deux valeurs par $\sin(\theta) = \frac{H\theta}{O\theta}$, $\cos(\theta) = \frac{OH}{O\theta}$.

Démonstration. L'idée de la preuve est simplement que puisque $O\theta = 1$, les quotients se simplifient. \square

Remarque. Puisque l'angle correspond à la partie parcourue sur un cercle de rayon 1, un tour entier correspond à un angle de 2π , un demi-tour à un angle de π , un quart de tour à un angle de $\frac{\pi}{2}$, etc.

Voici un tableau listant des valeurs importantes de \cos et \sin pour $\theta \in \left[0; \frac{\pi}{2}\right]$:

θ	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$
$\cos(\theta)$	1	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{1}{2}$	0
$\sin(\theta)$	0	$\frac{1}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{3}}{2}$	1

TABLE 6 – Table trigonométrique

Remarque. Soit un triangle (ABC) rectangle en B . Notons $\widehat{ABC} = \theta$. Alors $AB = AC \cos(\theta)$ et $BC = AC \sin(\theta)$, cela vient directement de la définition du sinus et du cosinus comme rapports de longueur. Ainsi connaître un angle non droit d'un triangle rectangle nous permet de déduire les longueurs de tous les autres côtés si l'on connaît la longueur de l'hypothénuse (le plus long côté).

5.2 Propriétés des fonctions trigonométriques

Les fonctions \sin et \cos sont très importantes à étudier, notamment de par leur omniprésence en physique. Nous les étudierons à nouveau quand nous aurons l'outil de la dérivation, mais nous pouvons déjà explorer des éléments importants. Nous allons commencer par l'égalité fondamentale de la trigonométrie.

Proposition 5.2.1. *Soit θ un angle. Alors*

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

Démonstration. Par définition, on a un triangle rectangle d'hypothénuse 1 et de longueur des côtés $\sin(\theta)$ et $\cos(\theta)$ (c.f. le cercle trigonométrique). En utilisant le théorème de Pythagore, on en déduit directement que $\cos^2(\theta) + \sin^2(\theta) = 1$. \square

En regardant le cercle trigonométrique, on remarque les identités suivantes (nous les admettrons ici) :

- $\cos(-\theta) = \cos(\theta)$
- $\sin(-\theta) = -\sin(\theta)$
- $\cos(\pi - \theta) = -\cos(\theta)$
- $\sin(\pi - \theta) = \sin(\theta)$
- $\cos(\pi + \theta) = -\cos(\theta)$
- $\sin(\pi + \theta) = -\sin(\theta)$

De plus, comme un angle est le même en ajoutant un multiple de 2π (faire un tour revient à ne rien faire), on en déduit que $\cos(x + 2\pi) = \cos(x)$ et que $\sin(x + 2\pi) = \sin(x)$: on dit que les deux fonctions sont 2π -périodiques.

Enfin, nous allons démontrer une identité essentielle pour étudier les fonctions trigonométriques.

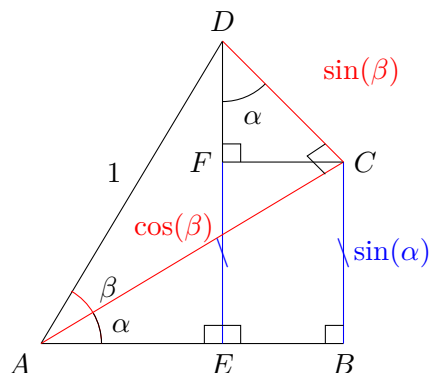


FIGURE 16 – Figure pour raisonner sur l'addition

Proposition 5.2.2 (Formule d'addition du sinus). *Soit α et β deux angles. Alors*

$$\cos(\alpha + \beta) = \cos(\alpha) \cos(\beta) - \sin(\alpha) \sin(\beta)$$

et

$$\sin(\alpha + \beta) = \sin(\alpha) \cos(\beta) + \sin(\beta) \cos(\alpha)$$

Démonstration. Nous allons utiliser la figure 16. En effet, avec les identités de symétrie données plus haut, on peut vérifier qu'il suffit de traiter le cas d'un angle aigu. Calculons en premier lieu $\cos(\alpha + \beta)$ puis $\sin(\alpha + \beta)$.

Par définition, $\cos(\alpha + \beta) = AE = AB - BE$. Calculons donc AB et BE . Comme (DFC) est un triangle rectangle en F et que l'angle en D est de mesure α , on en déduit que $FC = \sin(\alpha)DC$, or $DC = \sin(\beta)$, donc $FC = \sin(\alpha) \sin(\beta)$. De plus, comme $(FCBE)$ est un parallélogramme, ses côtés opposés sont de longueurs égales, donc $FC = EB$, d'où $EB = \sin(\alpha) \sin(\beta)$. Pour calculer AB , on remarque que le triangle (ABC) est rectangle en B , et d'hypoténuse $\cos(\beta)$, donc $AB = \cos(\alpha) \cos(\beta)$. On en déduit donc que $\boxed{\cos(\alpha + \beta) = \cos(\alpha) \cos(\beta) - \sin(\alpha) \sin(\beta)}$.

Par définition, $\sin(\alpha + \beta) = ED$, or $ED = EF + FD$ et par construction $FE = BC$. Par trigonométrie dans le triangle (ABC) , $BC = \cos(\beta) \sin(\alpha)$ et par trigonométrie dans (DFC) , $FD = \sin(\beta) \cos(\alpha)$, d'où $\boxed{\sin(\alpha + \beta) = \cos(\alpha) \sin(\beta) + \sin(\alpha) \cos(\beta)}$. \square

A partir des identités précédentes et des formules liées à la symétries dans le cercle trigonométrique, on peut déduire de nouvelles identités :

- $\cos(\alpha - \beta) = \cos(\alpha) \cos(-\beta) - \sin(\alpha) \sin(\beta) = \cos(\alpha) \cos(\beta) + \sin(\alpha) \sin(\beta)$
- $\sin(\alpha - \beta) = \cos(\alpha) \sin(-\beta) + \sin(\alpha) \cos(-\beta) = \sin(\alpha) \cos(\beta) - \cos(\alpha) \sin(\beta)$

Enfin, en considérant $\alpha = \beta$, on trouve les formules

$$\cos(2\alpha) = \cos^2(\alpha) - \sin^2(\alpha)$$

et

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

Exercice 5.1 (D'autres égalités). Cet exercice vise à démontrer d'autres égalités sur les fonctions trigonométriques. Certaines égalités sont plus faciles à obtenir en utilisant celles déjà démontrées dans l'exercice.

- En utilisant l'égalité fondamentale, réécrire $\cos(2\alpha)$ en une formule n'utilisant pas $\sin^2(\alpha)$, puis en une formule n'utilisant pas $\cos^2(\alpha)$.
- En additionnant $\cos(\alpha + \beta)$ et $\cos(\alpha - \beta)$, trouver une formule exprimant $\cos(\alpha) \cos(\beta)$.
- De même, exprimer $\sin(\alpha) \sin(\beta)$.
- De même, exprimer $\sin(\alpha) \cos(\beta)$ et $\cos(\alpha) \sin(\beta)$.
- En additionnant ou soustrayant les identités précédentes, déduire une formule exprimant $\cos(a) + \cos(b)$, $\cos(a) - \cos(b)$, $\sin(a) + \sin(b)$ et $\sin(a) - \sin(b)$. Indice : On posera le changement de variable $a = \alpha + \beta$, $b = \alpha - \beta$, ce qui équivaut à $\alpha = \frac{a+b}{2}$ et $\beta = \frac{a-b}{2}$.

6 Produit scalaire

Cette partie s'intéresse à un outil important à la fois en mathématiques et en physique : le produit scalaire. Nous le définirons puis en donnerons des définitions équivalentes, dans un premier temps. Ensuite, nous donnerons plusieurs propriétés du produit scalaire et nous verrons enfin une application du produit scalaire pour déterminer une équation de cercle. Pour donner des définitions équivalentes, il nous faut démontrer un premier théorème : le théorème d'Al Kashi.

6.1 Définitions

Le produit scalaire est une opérations prenant deux vecteurs et renvoyant un réel. On peut l'assimiler à une opération étudiant la projection d'un vecteur sur un autre (nous verrons une propriété rendant cette idée plus claire).

Définition 6.1. Soient \vec{u} et \vec{v} deux vecteurs du plan. On appelle produit scalaire, et on note $\vec{u} \cdot \vec{v}$, le réel

$$\vec{u} \cdot \vec{v} = \|\vec{u}\| \times \|\vec{v}\| \times \cos(\widehat{\vec{u}, \vec{v}})$$

6.1.1 Identités de polarisation

Théorème 6.1 (Al-Kashi). Soit (ABC) un triangle. En notant a, b, c les longueur des côtés opposés aux sommets, respectivement, A, B et C , alors

$$c^2 = a^2 + b^2 - 2ab \cos(\widehat{ACB})$$

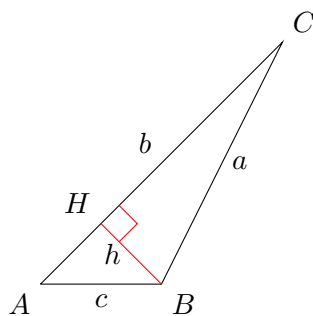


FIGURE 17 – Notations utilisées

Démonstration. On veut donc évaluer c^2 . Pour cela, on utilisera le projeté orthogonal de B sur (AC) (noté H , c.f. figure 17). De plus, nous voulons exprimer les longueurs avec $\hat{C} = \widehat{ACB}$.

Tout d'abord, dans le triangle (ABH) , par trigonométrie, on trouve que $CH = \cos(\hat{C})a$, d'où $AH = b - a \cos(\hat{C})$. De plus, toujours par trigonométrie, $h = a \sin(\hat{C})$.

En utilisant le théorème de Pythagore dans (ABH) , rectangle en H , on en déduit que $c^2 = h^2 + AH^2$, d'où en remplaçant AH et h :

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

ce qui, en développant, donne $c^2 = b^2 + a^2(\cos^2(\hat{C}) + \sin^2(\hat{C})) - 2ab \cos(\hat{C})$ or $\cos^2(\hat{C}) + \sin^2(\hat{C}) = 1$, d'où $\boxed{c^2 = b^2 + a^2 - 2ab \cos(\hat{C})}$. □

Grâce au théorème d'Al Kashi, on peut désormais trouver de nouvelles formules pour calculer le produit scalaire.

Proposition 6.1.1 (Identité de polarisation 1). Soient \vec{u} et \vec{v} deux vecteurs du plan, on a l'égalité suivante :

$$\vec{u} \cdot \vec{v} = \frac{1}{2} \left(\|\vec{u} + \vec{v}\|^2 - \|\vec{u}\|^2 - \|\vec{v}\|^2 \right)$$

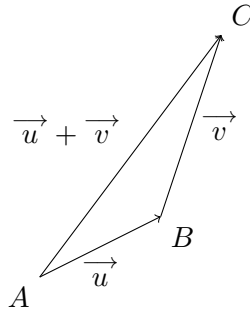


FIGURE 18 – Notations utilisées

Démonstration. Nous utiliserons les notations de la figure 18. On remarque d'abord que $\hat{B} = \pi - (\vec{u}, \vec{v})$ (en effet, en prolongeant le vecteur \vec{u} , on obtient un angle total plat et l'angle (\vec{u}, \vec{v}) de l'autre côté de \hat{B}).

On en déduit donc, en utilisant Al Kashi, que $\|\vec{u} + \vec{v}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2 + 2\|\vec{u}\|\|\vec{v}\|\cos(\vec{u}, \vec{v})$ or $\|\vec{u}\|\|\vec{v}\|\cos(\vec{u}, \vec{v}) = \vec{u} \cdot \vec{v}$. En isolant $\vec{u} \cdot \vec{v}$, on en déduit l'équation souhaitée, c'est-à-dire

$$\vec{u} \cdot \vec{v} = \frac{1}{2} \left(\|\vec{u} + \vec{v}\|^2 - \|\vec{u}\|^2 - \|\vec{v}\|^2 \right). \quad \square$$

De plus, le même théorème nous permet de déduire une autre formule.

Proposition 6.1.2 (Identité de polarisation 2). *Soient \vec{u} et \vec{v} deux vecteurs du plan. L'égalité suivante est vérifiée :*

$$\vec{u} \cdot \vec{v} = \frac{1}{2} \left(\|\vec{u}\|^2 + \|\vec{v}\|^2 - \|\vec{u} - \vec{v}\|^2 \right)$$

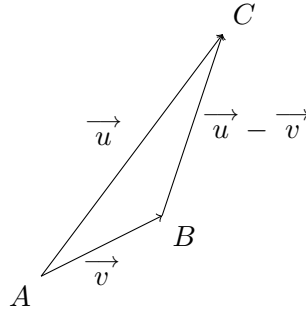


FIGURE 19 – Notations utilisées

Démonstration. Nous utiliserons les notations de la figure 19. La formule d'Al Kashi donne directement

$$\|\vec{u} - \vec{v}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2 - 2\vec{u} \cdot \vec{v}$$

d'où en isolant le produit scalaire $\vec{u} \cdot \vec{v} = \frac{1}{2} \left(\|\vec{u}\|^2 + \|\vec{v}\|^2 - \|\vec{u} - \vec{v}\|^2 \right).$ □

Les deux équations précédentes nous fournissent une dernière expression du produit scalaire.

Proposition 6.1.3 (Identité de polarisation 3). *Pour deux vecteurs \vec{u} et \vec{v} , le produit scalaire s'exprime*

$$\vec{u} \cdot \vec{v} = \frac{1}{4} \left(\|\vec{u} + \vec{v}\|^2 - \|\vec{u} - \vec{v}\|^2 \right)$$

Démonstration. Il suffit d'additionner les deux identités précédente puis de diviser ces identités par 2. □

Enfin, si l'on dispose d'une base orthonormée (\vec{e}_1, \vec{e}_2) , on obtient une nouvelle identité.

Proposition 6.1.4. Soit (\vec{e}_1, \vec{e}_2) une base orthonormée, \vec{u} et \vec{v} deux vecteurs de coordonnées respectives (x, y) et (x', y') , alors

$$\vec{u} \cdot \vec{v} = xx' + yy'$$

Démonstration. Il suffit de développer la troisième identité de polarisation, en utilisant le fait que $\|\vec{u} + \vec{v}\|^2 = (x+x')^2 + (y+y')^2$ et $\|\vec{u} - \vec{v}\|^2 = (x-x')^2 + (y-y')^2$ car la base est orthonormée. \square

6.2 Propriétés

Nous allons montrer les propriétés essentielles du produit scalaire.

Tout d'abord, le produit scalaire est dit symétrique.

Proposition 6.2.1 (Symétrie). Soient \vec{u} et \vec{v} deux vecteurs, alors

$$\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$$

Démonstration. En repassant par la définition, on remarque qu'il suffit de prouver que $\cos(\vec{u}, \vec{v}) = \cos(\vec{v}, \vec{u})$, or $\cos(\vec{v}, \vec{u}) = \cos(-(\vec{u}, \vec{v})) = \cos(\vec{u}, \vec{v})$. \square

De plus, le produit scalaire est dit bilinéaire.

Proposition 6.2.2 (Bilinéarité). Soient \vec{u} , \vec{v} et \vec{w} des vecteurs, et $k \in \mathbb{R}$. Alors

$$\vec{u} \cdot (\vec{v} + k\vec{w}) = (\vec{u} \cdot \vec{v}) + k(\vec{u} \cdot \vec{w})$$

Démonstration. Soit (\vec{e}_1, \vec{e}_2) une base orthonormée et (x, y) , (x', y') , (x'', y'') les coordonnées respectives de \vec{u} , \vec{v} et \vec{w} . Alors

$$\vec{u} \cdot (\vec{v} + k\vec{w}) = x(x' + kx'') + y(y' + ky'') = (xx' + yy') + k(xx'' + yy'') = (\vec{u} \cdot \vec{v}) + k(\vec{u} \cdot \vec{w})$$

\square

Remarque. La symétrie permet de déduire que

$$(\vec{u} + k'\vec{u}') \cdot (\vec{v} + k\vec{v}') = \vec{u} \cdot \vec{v} + k'(\vec{u}' \cdot \vec{v}) + k(\vec{u} \cdot \vec{v}') + kk'(\vec{u}' \cdot \vec{v}')$$

Nous avons dans l'introduction de cette partie mentionné un lien avec la projection d'un vecteur sur un autre. Voici la propriété qui illustre ce lien.

Proposition 6.2.3. Soient \vec{u} et \vec{v} deux vecteurs, ainsi que \vec{w} le projeté orthogonal de \vec{v} sur la droite prolongeant \vec{u} . Alors

$$\vec{u} \cdot \vec{v} = \vec{u} \cdot \vec{w}$$

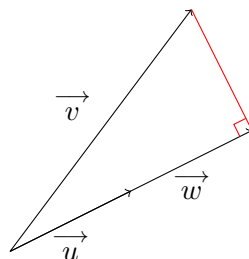


FIGURE 20 – Illustration de la proposition 6.2.3

Démonstration. La trigonométrie nous donne directement $\|\vec{w}\| = \|\vec{v}\| \cos(\vec{u}, \vec{v})$ or on sait de plus, par propriété du projeté orthogonal, que $(\vec{u}, \vec{w}) = 1$, d'où le résultat. \square

On peut de plus calculer la norme à partir du produit scalaire.

Proposition 6.2.4. Soit \vec{u} un vecteur du plan. Alors

$$\|\vec{u}\|^2 = \vec{u} \cdot \vec{u}$$

Démonstration. Le résultat est évident en repassant par la définition. \square

Enfin, le produit scalaire sert de critère d'orthogonalité (i.e. de critère pour savoir si deux vecteurs sont de directions perpendiculaires).

Proposition 6.2.5. Soient \vec{u} et \vec{v} deux vecteurs. Alors \vec{u} et \vec{v} sont orthogonaux si et seulement si $\vec{u} \cdot \vec{v} = 0$.

Démonstration. Si l'un des vecteurs est nul, il est par convention orthogonal à tout autre vecteur. Si les deux vecteurs sont non nuls, alors :

Si les deux vecteurs ont un angle absolu différent de $\frac{\pi}{2}$, par définition, $\cos(\vec{u}, \vec{v})$ est non nul, donc $\vec{u} \cdot \vec{v} \neq 0$.

Si les deux vecteurs ont un angle absolu de $\frac{\pi}{2}$, alors le cosinus de leur angle est nul donc les deux vecteurs sont nuls.

Donc $\boxed{\vec{u} \perp \vec{v} \iff \vec{u} \cdot \vec{v} = 0.}$ \square

6.3 Application à un cercle

Nous donnons dans cette partie un problème sous forme d'exercice.

Exercice 6.1. Soient A et B deux points. Le but de ce problème est de décrire l'ensemble des points formant le cercle \mathcal{C} de diamètre $[AB]$.

- Soit M . Montrer que $M \in \mathcal{C} \iff \vec{MA} \cdot \vec{MB} = 0$.
- En posant une base orthonormée, en déduire une équation du cercle en fonction des coordonnées de A et B . *Indication* : on posera les coordonnées du point M comme valant (x, y) , qui seront les variables de l'équation de cercle.

Troisième partie

Algèbre

7 Équations et inégalités

Cette section s'intéressera à la manipulation d'équations et d'inéquations. Une équation est une proposition composée uniquement de la relation « = » utilisée une fois avec une variable libre, appelée inconnue, souvent notée x , dont l'objectif est de trouver la valeur. On appelle membre gauche et membre droit, respectivement, les expressions se trouvant à la gauche et à la droite du « = ». C'est donc un prédicat égalitaire dont on cherche à établir le domaine de vérité : trouver ce domaine s'appelle résoudre l'équation.

Exemple 7.1. Voici plusieurs équations :

$$x + 1 = 0 \quad x^2 + 7x + 8 = 12 \quad 28 = 0$$

on remarque que la troisième équation est simplement fausse : il n'y a aucune solution à cette équation.

On utilisera ici un fait important et axiomatique (c'est-à-dire qu'il est accepté comme vrai car à la base même de nos raisonnements) :

Axiome 7.1. Si $a = b$ alors dans toute proposition de la forme $P(a)$, $P(b)$ a les mêmes valeurs de vérité.

Cela nous renseigne sur une condition suffisante pour résoudre une équation : une équation de la forme $x = a$ où $a \in \mathbb{R}$ admet comme solutions exactement $\{a\}$.

7.1 Équation du premier degré

Étudions tout d'abord la forme la plus simple d'une équation, l'équation du premier degré, c'est-à-dire de la forme suivante :

$$ax + b = c$$

où $a, b, c \in \mathbb{R}$ et $a \neq 0$ (car sinon, il n'y a pas de x , donc l'équation est directement résolue).

Nous allons pour cela montrer une propriété importante :

Proposition 7.1.1. Soient a et b deux objets, et f une fonction prenant ces objets en argument. Alors $a = b \implies f(a) = f(b)$.

Démonstration. On suppose que $a = b$. On sait qu'il existe un unique couple de la forme $(a, \alpha) \in \Gamma_f$, or $a = b$ donc le couple contenant l'image de b est aussi (a, α) , donc $f(b) = f(a) = \alpha$. \square

Cette proposition est évidente, mais elle est nécessaire pour plusieurs raisonnements. En effet, on peut maintenant trouver des solutions en appliquant des fonctions aux deux membres de l'équation. Remarquons que nous n'avons qu'une implication, or nous voulons l'ensemble exact des valeurs de x pour lesquelles l'équation est vérifiée : il faut donc si l'on recourt à cette méthode vérifier que chaque solution trouvée est bel et bien solution de l'équation en réinjectant la valeur dans l'équation (ceci signifie remplacer x par la valeur trouvée).

Exemple 7.2. Si l'on cherche x tel que $x = 5$ alors on peut appliquer la fonction carré, donnant $x^2 = 25$, or -5 aussi vérifie cette équation. Il faut donc vérifier pour chaque valeur de x trouvée si elle est bien appropriée ou si elle est en réalité un « faux positif ».

Grâce à cette proposition, nous allons prouver deux corollaires tout aussi importants.

Corollaire 7.1.1. Soient a et b deux réels et k un réel non nul. Alors

$$a = b \iff k \times a = k \times b$$

Démonstration. Soit $f : x \mapsto k \times x$ de réciproque $g : x \mapsto \frac{1}{k} \times x$, en composant par f puis par g , on obtient $a = b \implies k \times a = k \times b$ puis $k \times a = k \times b \implies a = b$, d'où $a = b \iff k \times a = k \times b$. \square

Corollaire 7.1.2. Soient a et b deux réels, et k un réel. Alors

$$a = b \iff a + k = b + k$$

Démonstration. La preuve se fait de la même façon que pour le corollaire précédent, avec les fonctions $f : x \mapsto x + k$ et $g : x \mapsto x - k$. \square

Remarque. En utilisant les deux corollaires, on trouve que $a = b \iff k \times a + k' = k \times b + k'$ pour $k \neq 0$.

Nous pouvons désormais résoudre une équation du premier degré :

Théorème 7.1 (Équations du premier degré). Soient a et b deux réels. Alors

$$ax + b = c \iff x = \frac{c - b}{a}$$

Démonstration. On additionne $-b$ dans chaque membre puis on multiplie chaque membre par $\frac{1}{a}$ pour obtenir l'équation sous la forme voulue. \square

Nous avons dès lors la solution directe à n'importe quelle équation du 1^{er} degré.

7.2 Système d'équations du premier degré

Nous allons maintenant nous intéresser à un système de deux équations. Un système de deux équations est un prédicat à deux variables libres contenant deux égalités séparées par un « et ». Plutôt que

$$(ax + by + c = d) \wedge (a'x + b'y + c' = d')$$

nous noterons le système sous la forme suivante :

$$\begin{cases} ax + by + c = d \\ a'x + b'y + c' = d' \end{cases}$$

Pour résoudre ce type d'équations, nous allons recourir à une nouvelle proposition. Nous allons utiliser ce que l'on appelle des combinaisons linéaires d'équations.

Proposition 7.2.1. Soit un système S de deux équations tel que décrit plus haut. Alors le système est équivalent, pour tout k non nul, au système

$$\begin{cases} ax + by + c = d \\ a'x + b'y + c' + k(ax + by + c) = d' + kd \end{cases}$$

Démonstration. En effet, on additionne des deux côtés kd à la deuxième équation, et on remplace ensuite d par $ax + by + c$. \square

Pouvoir faire des combinaisons linéaires d'équations va nous permettre d'éliminer directement une des variable dans la deuxième équation.

Exercice 7.1. Soit le système S suivant :

$$\begin{cases} 2x + 3y = 5 \\ 6x + y = 12 \end{cases}$$

trouver k tel qu'en additionnant k fois la première ligne à la deuxième ligne, l'équation restante est une équation du premier degré. La résoudre et l'injecter dans la première ligne pour trouver une deuxième équation du premier degré. La résoudre.

Exercice 7.2 (Formule générale *). Trouver à partir de la méthode précédente, pour le système

$$\begin{cases} ax + by + c = d \\ a'x + b'y + c' = d' \end{cases}$$

la formule générale du couple (x, y) solution. On supposera de plus que $ab' - b'a \neq 0$.

7.3 Inéquations

Cette section se concentrera plus précisément sur les inéquations. Comme les équations, nous parlons de prédicat à une inconnue, mais au lieu d'une relation égalitaire, nous utilisons comme relation une inégalité, c'est-à-dire l'une des relations suivantes : $\leq, <, >, \geq$.

Là encore, il est évident qu'une inéquation de la forme $x\mathcal{R}a$ où \mathcal{R} est l'une des relations citées plus tôt, se résout directement (par exemple $x \leq 2$ a comme solution évidente $] -\infty; 2]$).

Nous allons démontrer un résultat proche du premier résultat démontré dans la section 7.1, mais pour la conservation d'équations.

Pour cela, nous devons ajouter une définition.

Définition 7.1. Une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ est dite croissante si $\forall x \in \mathbb{R}, \forall y \in \mathbb{R}, x \leq y \implies f(x) \leq f(y)$.
 f est dite strictement croissante si $\forall x \in \mathbb{R}, \forall y \in \mathbb{R}, x < y \implies f(x) < f(y)$.
 f est dite décroissante si $\forall x \in \mathbb{R}, \forall y \in \mathbb{R}, x \leq y \implies f(x) \geq f(y)$.
 f est dite strictement décroissante si $\forall x \in \mathbb{R}, \forall y \in \mathbb{R}, x < y \implies f(x) > f(y)$.

Remarque. Une fonction strictement croissante (respectivement décroissante) est donc croissante (respectivement décroissante).

Proposition 7.3.1. Soit f une fonction croissante, alors $ax \leq b \implies f(ax) \leq f(b)$.

Si f est décroissante, alors $ax \leq b \implies f(ax) \geq f(b)$.

De même si f est strictement croissante ou décroissante en remplaçant l'inégalité large par une inégalité stricte.

Démonstration. Cela tient directement à la définition d'une fonction croissante. □

Cette propriété des fonctions croissantes est utile pour additionner et soustraire, car l'addition par un nombre fixé et la multiplication par un réel strictement positif sont des fonctions croissantes.

Proposition 7.3.2. Soit k un nombre réel, alors $x \mapsto x + k$ est strictement croissante.

Démonstration. On remarque directement que si $x < y$, alors $x - y < 0$. Or $x - y = x + k - (y + k)$ donc $x + k - (y + k) < 0$ ce qui est équivalent à $x + k < y + k$. D'où la stricte croissance de la fonction. □

Proposition 7.3.3. Soit k un nombre réel strictement positif, alors $x \mapsto k \times x$ est strictement croissante. Si $k < 0$ alors $x \mapsto k \times x$ est strictement décroissante.

Démonstration. On remarque que si $x < y$ alors $k(x - y) < 0$ puisque k est strictement positif et $x - y$ strictement négatif. Donc $kx < ky$, d'où le résultat.

Dans le cas où $k < 0$, $k(x - y) > 0$ donc $kx > ky$. □

De là, on en déduit comment résoudre une inéquation du premier degré en multipliant et additionnant des termes. En effet, ces deux fonctions donnent des inéquations équivalentes quand on compose, puisque ce sont des bijections et que leur réciproque est aussi croissante strictement.

Théorème 7.2 (Inéquation du premier degré). Soit une inéquation de la forme $ax + b < c$, alors la solution est $] -\infty; \frac{c-b}{a}[$ si $a > 0$ et $] \frac{c-b}{a}; +\infty[$ si $a < 0$

Démonstration. Nous montrerons uniquement le premier cas. On compose par $f : x \mapsto x - b$ puis par $g : x \mapsto \frac{x}{a}$ pour en déduire directement $x < \frac{c-b}{a}$. □

7.4 Résolution d'équation du second degré

Nous allons désormais nous concentrer sur les équations du second degré, c'est-à-dire où il y a un terme en x^2 , mais avant ça nous devons traiter un cas particulier d'équation, appelé équations à produit nul.

Théorème 7.3. \mathbb{R} est intègre, ce qui signifie que $ab = 0 \iff (a = 0) \vee (b = 0)$.

Démonstration. Supposons que $a = 0$ ou $b = 0$, alors de façon évidente $ab = 0$. Maintenant, si l'on suppose que $a \neq 0$ et $b \neq 0$, alors on peut diviser ab par b pour obtenir a , qui est par hypothèse différent de 0. Donc $ab \neq 0$. \square

Ce théorème nous permet de déduire une façon de résoudre une équation particulière.

Proposition 7.4.1. *L'équation $(ax + b)(a'x + b') = 0$ est équivalente à $ax + b = 0 \vee a'x + b' = 0$.*

Remarque. De plus, de par la définition même de l'union ensembliste, si l'on a un système de la forme $E_1 \vee E_2$ où E_1 et E_2 sont des équations, alors en notant S_1 et S_2 les ensembles de solution respectivement de E_1 et E_2 , la solution du système est $S_1 \cup S_2$.

Nous souhaitons désormais résoudre une équation de la forme

$$ax^2 + bx + c = 0$$

où $a \neq 0$. On peut se convaincre que résoudre cette équation nous permet de résoudre n'importe quelle équation avec un x^2 , quitte à devoir faire passer d'un seul côté du « = » tous les termes. Grâce au résultat précédent, on sait qu'il suffit pour résoudre cette équation de trouver une forme factorisée de l'équation, c'est-à-dire mettre l'équation sous la forme $(ax + b)(a'x + b') = 0$, auquel cas nous avons une équation à produit nul et deux équations du premier degré.

Pour commencer, nous allons nous intéresser à une forme restreinte de ce problème.

Proposition 7.4.2. *La solution à une équation de la forme*

$$x^2 = a$$

pour $a \geq 0$ est $\{\sqrt{a}; -\sqrt{a}\}$ (et il n'y a pas de solution pour $a < 0$).

Démonstration. Nous faisons une simple série de calculs :

$$\begin{aligned} x^2 = a &\iff x^2 - a = 0 \\ &\iff (x - \sqrt{a})(x + \sqrt{a}) = 0 \\ &\iff x - \sqrt{a} = 0 \quad \vee \quad x + \sqrt{a} = 0 \\ x^2 = a &\iff x = \sqrt{a} \quad \vee \quad x = -\sqrt{a} \end{aligned}$$

\square

Ce cas particulier nous renseigne sur la forme que nous pouvons vouloir donner à notre équation du second degré : si l'on peut mettre l'équation sous la forme $(ax + b)^2 = c$ alors on peut réutiliser cette méthode pour trouver l'ensemble (potentiellement vide) des solutions de l'équation. Nous voulons donc trouver α , β et γ tels que $ax^2 + bx + c = 0 \iff (\alpha x + \beta)^2 = \gamma$. Nous allons pour cela définir ce que l'on appelle la forme canonique.

Définition 7.2 (Forme canonique). Soit une équation du second degré de la forme

$$ax^2 + bx + c = 0$$

alors cette équation est équivalente à

$$a \left(x + \frac{b}{2a} \right)^2 - \frac{b^2 - 4ac}{4a} = 0$$

appelée forme canonique de l'équation.

Démonstration. Nous allons simplement développer la forme canonique :

$$\begin{aligned} a \left(x + \frac{b}{2a} \right)^2 - \frac{b^2 - 4ac}{4a} &= ax^2 + 2 \frac{abx}{2a} + \frac{b^2}{4a^2} - \frac{b^2 - 4ac}{4a} \\ a \left(x + \frac{b}{2a} \right)^2 - \frac{b^2 - 4ac}{4a} &= ax^2 + bx + c \end{aligned}$$

D'où l'équivalence des deux équations. \square

L'équation donnée ci-dessus est donc équivalence à

$$\left(x + \frac{b}{2a}\right)^2 = \frac{b^2 - 4ac}{4a^2}$$

or $a^2 > 0$ donc il y a (au moins) une solution si et seulement si $b^2 - 4ac \geq 0$. On en déduit le critère pour résoudre une équation du second degré.

Théorème 7.4 (Équation du second degré). *Soit une équation du second degré de la forme*

$$ax^2 + bx + c = 0$$

où $a \neq 0$. Soit $\Delta = b^2 - 4ac$, alors suivant le signe de Δ :

- Si $\Delta > 0$ alors il y a deux solutions : $x = \frac{-b + \sqrt{\Delta}}{2a} \vee x = \frac{-b - \sqrt{\Delta}}{2a}$.
- Si $\Delta = 0$ alors il y a une seule solution : $x = \frac{-b}{2a}$.
- Si $\Delta < 0$ alors il n'y a pas de solution.

Démonstration. Il suffit d'utiliser la proposition 7.4.2 pour l'équation sous la forme canonique. \square

Dans le cas où $\Delta < 0$, en réalité, il existe des solutions : simplement, celles-ci ne sont pas réelles mais complexes. Les nombres complexes sont le thème du prochain chapitre ; ils sont des nombres tels qu'il existe i , appelé nombre imaginaire, qui a la particularité que $i^2 = -1$.

8 Nombres complexes

Cette section se concentrera sur l'étude des nombres complexes. Ceux-ci sont définis comme des nombres de la forme $x + iy$ où $x, y \in \mathbb{R}$ et

$$i^2 = -1$$

Si l'introduction d'un nombre de carré négatif peut sembler illicite ou étrange (en effet, un carré est toujours positif), nous allons commencer par construire cet ensemble, noté \mathbb{C} , pour montrer qu'il existe bien un ensemble se comportant comme décrit. Nous verrons ensuite le plan complexe et les interprétations géométriques qui vont avec cette notion. Enfin, nous étendront notre résolution d'équations du second degré à la résolution dans \mathbb{C} de ces équations.

8.1 Construction des nombres complexes

Nous allons munir \mathbb{R}^2 d'une loi d'addition et d'une loi de multiplication.

Définition 8.1. Soit $(x, y), (x', y') \in \mathbb{R}^2$, on définit

$$(x, y) + (x', y') = (x + x', y + y')$$

et

$$(x, y) \times (x', y') = (x \times x' - y \times y', x \times y' + x' \times y)$$

On appelle nombre complexe un élément de \mathbb{R}^2 en utilisant les opérations ainsi définies. Ainsi, $\mathbb{R}^2 = \mathbb{C}$.

Nous allons maintenant montrer que les opérations ainsi construites conservent les mêmes propriétés que les opérations de \mathbb{R} .

Proposition 8.1.1 (Commutativité). Soient $(x, y), (x', y') \in \mathbb{C}$, alors $(x, y) + (x', y') = (x', y') + (x, y)$ et $(x, y)(x', y') = (x', y')(x, y)$.

Démonstration. Ce sont de simples vérifications de calculs :

$$\begin{aligned} (x, y) + (x', y') &= (x + x', y + y') \\ &= (x' + x, y' + y) \quad (x, y, x', y' \in \mathbb{R}) \\ (x, y) + (x', y') &= (x', y') + (x, y) \\ (x, y) \times (x', y') &= (xx' - yy', xy' + x'y) \\ &= (x'x - y'y, y'x + yx') \quad (x, y, x', y' \in \mathbb{R}) \\ (x, y) \times (x', y') &= (x', y') \times (x, y) \end{aligned}$$

□

Exercice 8.1. Vérifier que pour tous $x, y, z \in \mathbb{C}$, $x + (y + z) = (x + y) + z$ et que $x \times (y \times z) = (x \times y) \times z$. Cette propriété s'appelle l'associativité de l'addition.

Proposition 8.1.2 (Neutre pour l'addition). Pour tout $(x, y) \in \mathbb{C}$, $(x, y) + (0, 0) = (x, y)$

Démonstration. Cela est direct en revenant à la définition de l'addition dans \mathbb{C} .

□

Proposition 8.1.3 (Neutre pour le produit). Pour tout $(x, y) \in \mathbb{C}$, $(x, y) \times (1, 0) = (x, y)$.

Démonstration. En effet,

$$\begin{aligned} (x, y) \times (1, 0) &= (1x + 0y, 1y + 0x) \\ (x, y) \times (1, 0) &= (x, y) \end{aligned}$$

□

Nous allons maintenant définir i .

Définition 8.2. Le nombre imaginaire i est défini comme $i = (0, 1)$.

Démonstration. On vérifie que $i^2 = -1 : i \times i = (0 - 1, 0 + 0) = (-1, 0)$. \square

Enfin, soit la fonction $f : (x, 0) \mapsto x$, cette fonction est une bijection, de réciproque évidente $g : x \mapsto (x, 0)$ et on remarque que $f(x + y) = f(x) + f(y)$ et $f(x \times y) = f(x) \times f(y)$. C'est ce que l'on appelle un morphisme d'anneau, mais cette notion est hors du programme. Ceci étant, nous voyons que l'on peut identifier \mathbb{R} à un sous-ensemble de \mathbb{C} .

Théorème 8.1. Soit $(x, y) \in \mathbb{C}$. On peut écrire (x, y) de façon unique comme la somme d'un nombre réel (identifié à un nombre complexe) et du produit d'un nombre réel par i , et cette décomposition est

$$(x, y) = x + iy$$

Démonstration. L'existence est une simple vérification que $(x, y) = x + y \times i$.

Pour l'unicité, supposons deux décompositions $(x, y) = a + b \times i = a' + b' \times i$. Alors on vérifie que $a + b \times i = (a, b)$ et $a' + b' \times i = (a', b')$. Donc par égalité avec (x, y) on en déduit que $a = a'$ et $b = b'$. \square

Nous retrouvons donc la forme que nous souhaitions au début.

Si cet exercice de construction peut sembler inutile, il est essentiel pour s'assurer que l'ensemble \mathbb{C} existe bel et bien, ce qui ne semble pas chose aisée quand l'on voit qu'il contient un élément dont le carré est négatif.

8.2 Le plan complexe

En assimilant les nombres complexes à \mathbb{R}^2 , on remarque que \mathbb{C} est naturellement similaire au plan dans lequel, si l'on fixe une origine, on obtient un système de coordonnées qui est un ensemble de couple de réels. De plus, l'addition de vecteurs est, au niveau des coordonnées, une addition coordonnée par coordonnée, exactement comme l'addition complexe. On décide donc de représenter les nombres complexe comme des vecteurs du plan en fixant une origine O . On dit qu'un point est d'affixe $x + iy$ lorsqu'il correspond au nombre $x + iy$.

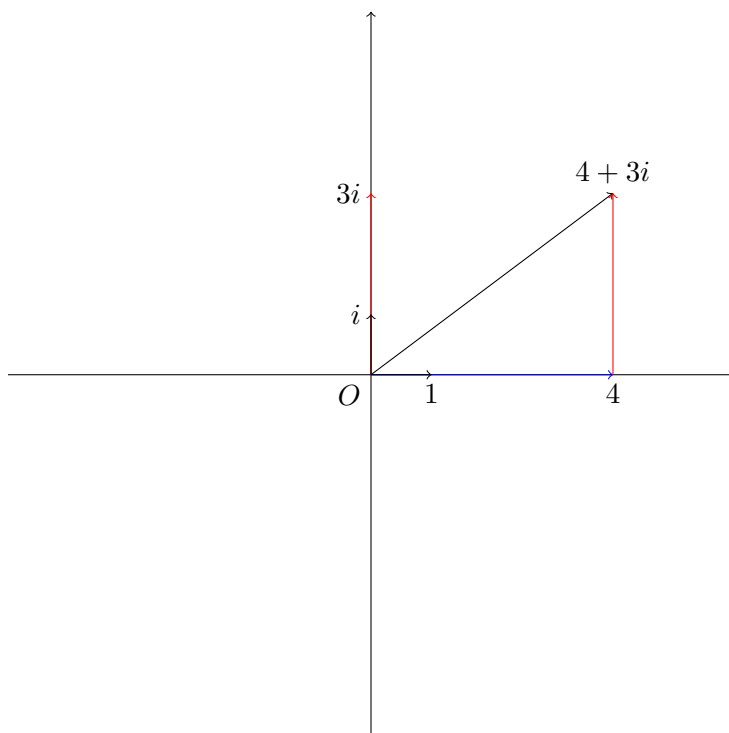


FIGURE 21 – Plan complexe avec un point d'affixe $4 + 3i$

Nous allons maintenant voir une nouvelle façon de déterminer la position d'un vecteur, qui est un système de coordonnées appelé coordonnées polaires. Pour cela, nous allons avoir besoin d'un théorème préalable.

Théorème 8.2. Soit \vec{u} un vecteur non nul et \vec{e} un vecteur du plan fixé. Alors \vec{u} est déterminé de façon unique par sa norme et son angle principal (c'est-à-dire l'angle qui fait moins d'un tour) avec \vec{e} .

Démonstration. Nous admettrons ce théorème dans ce document. □

Nous allons donc relier cette notion au plan complexe en considérant l'opération donnant la norme du vecteur associé à un complexe.

Définition 8.3 (Module). Soit $z \in \mathbb{C}$, on appelle module de z , et l'on note $|z|$, le nombre

$$|z| = \|\vec{u}_z\|$$

où \vec{u}_z est le vecteur d'affixe z .

De par la formule du calcul de la norme d'un vecteur dans un système de coordonnées fixé, on en déduit que si $z = x + iy$ alors $|z| = \sqrt{x^2 + y^2}$.

De plus, nous nommons argument d'un nombre complexe l'angle que fait son vecteur représentatif avec l'axe des origines.

Proposition 8.2.1. Soit un nombre complexe z de module r et d'argument θ . Alors $z = r \cos(\theta) + ir \sin(\theta)$.

Démonstration. On remarque que $\frac{z}{r}$ est de module 1 : il appartient donc au cercle trigonométrique et correspond au même angle que z . On en déduit donc que les coordonnées de $\frac{z}{r}$ sont $(\cos(\theta), \sin(\theta))$. Il en découle que les coordonnées de z sont $(r \cos(\theta), r \sin(\theta))$, donc par unicité de $z = x + iy$ on en déduit que $z = r \cos(\theta) + ir \sin(\theta)$. □

8.2.1 La conjugaison complexe

Cette partie va s'intéresser à une fonction appelée conjugaison, que l'on note \bar{z} pour un nombre complexe z , et qui est définie par $\bar{\cdot} : \mathbb{C} \rightarrow \mathbb{C}$.

$$x + iy \mapsto x - iy$$

Nous allons montrer que cette fonction est un morphisme d'anneau.

Proposition 8.2.2. Soient z et z' deux complexes. Alors $\overline{z + z'} = \bar{z} + \bar{z}'$ et de plus $\overline{-z} = -\bar{z}$.

Démonstration. On pose $z = x + iy$ et $z' = x' + iy'$. Alors $\overline{z + z'} = x + x' - iy - iy' = (x - iy) + (x' - iy') = \bar{z} + \bar{z}'$.

Pour la deuxième partie, on remarque que $\overline{z + (-z)} = 0 = \bar{z} + \overline{-z}$ donc, en passant \bar{z} de l'autre côté, on trouve bien que $\overline{-z} = -\bar{z}$. □

Exercice 8.2. Montrer que $\overline{z \times z'} = \bar{z} \times \bar{z}'$. En déduire que $\frac{\bar{1}}{z} = \frac{1}{\bar{z}}$ pour $z \neq 0$.

Cette fonction nous permet d'exprimer autrement le module d'un nombre complexe.

Proposition 8.2.3. Soit $z \in \mathbb{C}$. Alors

$$|z| = \sqrt{z\bar{z}}$$

Démonstration. Il suffit de vérifier que pour $z = x + iy$, $z\bar{z} = x^2 + y^2$. Un calcul nous donne cette égalité :

$$\begin{aligned} z\bar{z} &= (x + iy)(x - iy) \\ &= x^2 - ixy + ixy - i^2y^2 \\ z\bar{z} &= x^2 + y^2 \end{aligned}$$

□

Enfin, il existe un lien entre l'argument d'un nombre complexe et son conjugué.

Proposition 8.2.4. *Soit z un nombre complexe de module r et d'argument θ . Alors \bar{z} a le même module et son argument est $-\theta$.*

Démonstration. Il suffit de vérifier que $\cos(-\theta) + i \sin(-\theta) = \cos(\theta) - i \sin(\theta)$. Ce calcul est laissé au lecteur. \square

8.2.2 Partie réelle et imaginaire

Nous avons jusque là décomposé les nombres complexes sous la forme $z = x + iy$. Nous pouvons systématiser cette décomposition en formant les deux fonctions

$$\Re : \begin{array}{ccc} \mathbb{C} & \longrightarrow & \mathbb{R} \\ x + iy & \longmapsto & x \end{array}$$

et

$$\Im : \begin{array}{ccc} \mathbb{C} & \longrightarrow & \mathbb{R} \\ x + iy & \longmapsto & y \end{array}$$

Pour $z \in \mathbb{C}$, on appelle $\Re(z)$ sa partie réelle et $\Im(z)$ sa partie imaginaire. On remarque que $z \in \mathbb{R}$ si et seulement si $z = \Re(z)$ et que $z \in i\mathbb{R}$ si et seulement si $z = \Im(z)$ (l'ensemble $i\mathbb{R}$ est l'ensemble des multiples réels que i , qu'on appelle l'ensemble des imaginaires purs).

Proposition 8.2.5. *Soit $z \in \mathbb{C}$, alors $\Re(z) = \frac{1}{2}(z + \bar{z})$ et $\Im(z) = \frac{1}{2i}(z - \bar{z})$.*

Démonstration. Là encore, un calcul suffit :

$$\begin{aligned} \frac{1}{2}(z + \bar{z}) &= \frac{1}{2}(x + iy + x - iy) \\ &= \frac{1}{2}(2x) \end{aligned}$$

$$\frac{1}{2}(z + \bar{z}) = x$$

$$\begin{aligned} \frac{1}{2i}(z - \bar{z}) &= \frac{1}{2i}(x + iz - x - (-iy)) \\ &= \frac{1}{2i}2iy \end{aligned}$$

$$\frac{1}{2i}(z - \bar{z}) = y$$

\square

On peut exprimer, par trigonométrie, un lien entre l'argument d'un nombre complexe, ses parties réelle et imaginaire et son module.

Proposition 8.2.6. *Soit $z \in \mathbb{C}$, de module r et d'argument θ , alors*

$$\cos(\theta) = \frac{\Re(z)}{r}$$

et

$$\sin(\theta) = \frac{\Im(z)}{r}$$

Démonstration. La démonstration est laissée au lecteur. \square

8.2.3 Notation exponentielle

Nous allons désormais utiliser une nouvelle notation pour désigner $\cos(\theta) + i \sin(\theta)$, qu'on appelle notation exponentielle.

Définition 8.4. On appelle exponentielle de $i\theta$, et l'on note $e^{i\theta}$, le nombre

$$e^{i\theta} = \cos(\theta) + i \sin(\theta)$$

Remarque. Un nombre complexe z d'argument r et de module θ s'exprime alors directement $z = re^{i\theta}$.

Une propriété importante de la notation exponentielle est qu'elle permet d'écrire plus facilement les multiplications complexes.

Proposition 8.2.7. Soient z et z' deux nombres complexes d'arguments respectifs θ et θ' et de modules respectifs r et r' . Alors on a l'égalité suivante :

$$z \times z' = rr' e^{i(\theta+\theta')}$$

Démonstration. Nous décomposons ce produit en un module et un argument. Il est évident que le module du produit est le produit des modules (de par la formule du produit elle-même). Nous allons donc prouver que $e^{i\theta} e^{i\theta'} = e^{i(\theta+\theta')}$:

$$\begin{aligned} e^{i\theta} e^{i\theta'} &= (\cos(\theta) + i \sin(\theta))(\cos(\theta') + i \sin(\theta')) \\ &= [\cos(\theta) \cos(\theta') - \sin(\theta) \sin(\theta')] + i [\cos(\theta) \sin(\theta') + \sin(\theta) \cos(\theta')] \\ &= \cos(\theta + \theta') + i \sin(\theta + \theta') \\ e^{i\theta} e^{i\theta'} &= e^{i(\theta+\theta')} \end{aligned}$$

□

En reprenant les éléments précédents, on remarque que l'inverse d'un complexe de norme 1 est son conjugué.

Proposition 8.2.8. Soit $z \in \mathbb{C}$ tel que $|z| = 1$, et soit θ son argument. Alors $z^{-1} = \bar{z}$.

Démonstration. Tout d'abord, $z = e^{i\theta}$ d'après la remarque 8.2.3. Or on sait que $z^{-1}z = 1$, donc en écrivant $z^{-1} = e^{i\theta'}$, on en déduit que $e^{i(\theta+\theta')} = e^{0 \times i}$ donc $\theta + \theta' = 0$ (nous considérons les angles principaux, d'où l'absence de multiple de 2π). De cette égalité on en déduit que $\theta' = -\theta$, donc par la proposition 8.2.4 on en déduit que $z^{-1} = \bar{z}$. □

Cette propriété nous permet de déduire l'identité d'Euler.

Théorème 8.3 (Identité d'Euler). Soit $\theta \in]-\pi; \pi]$, alors

$$\begin{aligned} \cos(\theta) &= \frac{e^{i\theta} + e^{-i\theta}}{2} \\ \text{et} \\ \sin(\theta) &= \frac{e^{i\theta} - e^{-i\theta}}{2i} \end{aligned}$$

Démonstration. Cette identité vient directement de la définition de parties réelle et imaginaire d'un complexe de module 1, en utilisant la propriété précédente et la formule donnant les parties réelle et imaginaire d'un complexe en fonction de son conjugué. □

Nous allons enfin voir l'identité de Moivre.

Proposition 8.2.9. Soit $z \in \mathbb{C}$, d'argument θ . Alors z^n a comme argument $n\theta$.

Démonstration. La preuve est simplement une induction utilisant le fait que l'argument d'un produit est la somme des arguments. □

Théorème 8.4 (Formule de Moivre). Soit $\theta \in]-\pi; \pi]$, alors

$$(\cos(\theta) + i \sin(\theta))^n = \cos(n\theta) + i \sin(n\theta)$$

Démonstration. Il suffit de développer l'identité d'Euler et d'utiliser la propriété précédente. □

8.3 Les équations de degré 2

Si nous revenons sur les équations de degré 2 traitées plus tôt, nous remarquons que la forme factorisée fonctionne encore parfaitement dans le cadre des complexes. Simplement, nous pouvons maintenant trouver une racine carrée pour tout nombre complexe.

Définition 8.5 (Racine carrée). Soit $z \in \mathbb{C}$, on appelle racine carrée principale de z le complexe de module $\sqrt{|z|}$ et d'argument $\frac{1}{2}\theta$ où θ est l'argument de z . La racine carrée, mise au carré, de z , vaut z .

Démonstration. On remarque que le module de la racine mise au carré vaut $\sqrt{|z|}^2 = |z|$ et l'argument vaut $2 \times \frac{1}{2}\theta = \theta$, donc la racine mise au carré vaut bien z . \square

La résolution d'une équation de degré 2 peut alors se généraliser de la façon suivante.

Théorème 8.5. *Soit une équation de la forme*

$$az^2 + bz + c = 0$$

où $a \neq 0$ et de coefficients complexes. Soit $\Delta = b^2 - 4ac$ et soit δ la racine carrée principale de Δ . Alors

- *Si $\delta = 0$ alors il n'y a qu'une solution : $x = \frac{-b}{2a}$.*
- *Si $\delta \neq 0$ alors il y a deux solutions : $x = \frac{-b + \delta}{2a} \vee x = \frac{-b - \delta}{2a}$.*

Démonstration. Il suffit de reprendre la preuve du cas réel et de l'adapter en posant une racine carrée dans tous les cas à Δ . \square

9 Suites : des éléments algébriques

Dans cette section, nous définirons la notion de suite, où nous utiliserons pleinement le raisonnement par récurrence. Nous étudierons ensuite les sommes et leurs propriétés, puis nous verrons plus en précision les suites arithmétiques, géométriques et arithmético-géométriques.

9.1 Définition d'une suite

Nous allons voir la définition d'une suite, avec les notations associées.

Définition 9.1. On appelle suite réelle, ou simplement suite, une fonction $u : \mathbb{N} \rightarrow \mathbb{R}$. C'est donc une liste infinie de nombres réels ordonnés (d'abord l'élément d'indice 0, puis l'élément d'indice 1, etc.) On note généralement une suite u de l'une des façons suivantes : (u) , $(u_n)_{n \in \mathbb{N}}$ ou (u_n) s'il n'y a pas de confusion possible.

Pour une suite (u) , on notera u_i son terme d'indice i , ou encore $[u]_i$ (cette notation sera surtout utile pour une suite avec un nom plus long, pour écrire par exemple $[u + v]_i$).

Une suite peut être vue comme une succession de nombres réels allant à l'infini.

Exemple 9.1. Voici plusieurs suites classiques :

- la suite (n) qui est la suite $0, 1, 2, 3, 4, 5, \dots$
- la suite (n^2) qui est la suite $0, 1, 4, 9, 16, 25, \dots$
- la suite $(2n)$ qui est la suite $0, 2, 4, 6, 8, 10, \dots$
- la suite $((-1)^n)$ qui est la suite $1, -1, 1, -1, 1, -1, \dots$

Remarque. De par la définition même d'une suite, le raisonnement par récurrence est particulièrement adapté. En effet, un raisonnement par récurrence permet de prouver une propriété pour tout entier naturel. Ici, justement, une suite est définie par l'image qu'elle envoie à chaque entier naturel.

Définition 9.2 (Suite définie par une relation de récurrence). On peut définir une suite (u) par une relation de récurrence, c'est-à-dire par deux relations :

$$\begin{cases} u_0 \in \mathbb{R} \\ \forall n \in \mathbb{N}, u_{n+1} = f(u_n) \end{cases}$$

Où $f : \mathbb{R} \rightarrow \mathbb{R}$

Démonstration. Montrons que la suite est bien définie, par récurrence :

- (u) est bien définie en 0 puisqu'on a fixé u_0 .
- Supposons que (u) soit bien définie pour tout $k \leq n$, où $n \geq 0$. Alors $u_{n+1} = f(u_n)$ est bien définie, puisqu'image par une fonction de u_n , bien définie. Donc (u) est bien définie en $n + 1$.

Par récurrence, on en déduit que (u) est bien définie en n pour tout $n \in \mathbb{N}$. □

Exercice 9.1. Soit la suite (u) définie par la relation de récurrence suivante :

$$\begin{cases} u_0 = 0 \\ \forall n \in \mathbb{N}, u_{n+1} = u_n + 1 \end{cases}$$

Montrer que $(u) = (n)$, c'est-à-dire que $\forall n \in \mathbb{N}, u_n = n$.

Exercice 9.2. Soit la suite (u) définie par la relation de récurrence suivante :

$$\begin{cases} u_0 = 0 \\ \forall n \in \mathbb{N}, u_{n+1} = u_n^2 + 2u_n + 1 \end{cases}$$

Montrer que $(u) = (n^2)$, c'est-à-dire que $\forall n \in \mathbb{N}, u_n = n^2$.

9.2 Sommes et produits

Nous allons maintenant étudier le symbole \sum et le symbole \prod , qui permettent d'écrire rigoureusement ce que l'on peut écrire par une somme avec des points de suspension.

Définition 9.3. Soit (u_n) une suite. On note $\sum_{k=i}^j u_k$ et $\prod_{k=i}^j u_k$ (pour $i \leq j$ fixés) les réels suivants :

$$\sum_{k=i}^j u_k = u_i + u_{i+1} + u_{i+2} + \dots + u_{j-1} + u_j$$

et

$$\prod_{k=i}^j u_k = u_i \times u_{i+1} \times u_{i+2} \times \dots \times u_{j-1} \times u_j$$

La variable k dans ces expressions est muettes : elle est seulement propre à l'intérieur de l'expression décrite dans le symbole et peut être remplacée par n'importe quelle autre variable.

Nous allons définir, plus rigoureusement, les sommes et les produits par récurrence :

$$\left[\sum_{k=i}^i u_k = u_i \right] \wedge \left[\sum_{k=i}^{j+1} u_k = \sum_{k=i}^j u_k + u_{j+1} \right]$$

$$\left[\prod_{k=i}^i u_k = u_i \right] \wedge \left[\prod_{k=i}^{j+1} u_k = \prod_{k=i}^j u_k \times u_{j+1} \right]$$

et l'on prend la convention que si $j < i$, alors $\sum_{k=i}^j u_k = 0$ et $\prod_{k=i}^j u_k = 1$.

Nous allons voir plusieurs propriétés de ces notations. Nous ne prouverons que le cas de \sum car le cas de \prod sera en général similaire, et fait un bon exercice pour le lecteur.

Proposition 9.2.1 (Recollement). Soit (u) une suite, $i \leq j < n$ des entiers. Alors

$$\sum_{k=i}^j u_k + \sum_{k=j+1}^n u_k = \sum_{k=i}^n u_k$$

et de même

$$\prod_{k=i}^j u_k \times \prod_{k=j+1}^n u_k = \prod_{k=i}^n u_k$$

Démonstration. Nous allons prouver ce résultat par deux récurrences successives.

Tout d'abord, prouvons que

$$\sum_{k=i}^j u_k = u_i + \sum_{k=i+1}^j u_k$$

par récurrence sur j :

- Si $i = j$, alors

$$\sum_{k=i}^j u_k = u_i + 0$$

- Si la propriété est vraie pour un $i < j$, alors

$$\sum_{k=i}^{j+1} u_k = \sum_{k=i}^j u_k + u_{j+1} = u_i + \sum_{k=i+1}^{j+1} u_k$$

D'où la propriété par récurrence :

$$\sum_{k=i}^j u_k = u_i + \sum_{k=i+1}^j u_k$$

Prouvons maintenant la proposition par récurrence sur j , où i et n sont fixés.

- Si $j = i$, alors l'équation est directe avec la propriété précédente.
- Si la propriété est vraie pour j , montrons qu'elle est vraie pour $j + 1$:

$$\sum_{k=i}^j u_k + \sum_{k=j+1}^n u_k = \sum_{k=i}^j u_k + u_{j+1} + \sum_{k=j+2}^n u_k = \sum_{k=i}^{j+1} u_k + \sum_{k=j+2}^n u_k$$

Ainsi la propriété est vraie pour tout j , y compris pour $j = n$, donnant alors comme résultat

$$\boxed{\sum_{k=i}^j u_k + \sum_{k=j+1}^n u_k = \sum_{k=i}^n u_k + 0.}$$

□

Exercice 9.3. Montrer que

$$\prod_{k=i}^j k = \frac{j!}{(i-1)!}$$

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

Proposition 9.2.2. Soient $(u), (v)$ deux suites et $\lambda \in \mathbb{R}$. Alors

$$\sum_{k=i}^j (\lambda u_k + v_k) = \lambda \sum_{k=i}^j u_k + \sum_{k=i}^j v_k$$

et

$$\prod_{k=i}^j (\lambda u_k) = \lambda^{j-i+1} \prod_{k=i}^j u_k$$

Démonstration. Nous allons prouver ce résultat par récurrence sur j :

- Si $j = i$ alors l'égalité devient $\lambda u_i + v_i = \lambda u_i + v_i$, qui est vraie.
- Supposons la propriété vraie jusqu'au rang j . Alors

$$\sum_{k=i}^{j+1} (\lambda u_k + v_k) = \lambda \sum_{k=i}^j u_k + \sum_{k=i}^j v_k + \lambda u_{j+1} + v_{j+1}$$

d'où le résultat en regroupant les termes par factorisation et recollement.

Donc la propriété est vraie. □

Proposition 9.2.3. Soit (u) une suite. Alors

$$\sum_{k=i}^j u_k = \sum_{k=0}^{j-i} u_{j-k}$$

et

$$\prod_{k=i}^j u_k = \prod_{k=0}^{j-i} u_{j-k}$$

Démonstration. Démonstration laissée au lecteur. Indication : utiliser la propriété utilisée dans la preuve de la proposition 9.2.1. □

Proposition 9.2.4.

$$\sum_{k=i}^j 1 = j - i + 1$$

Démonstration. Nous allons prouver ce résultat par récurrence.

- Si $j = i$, alors la somme vaut $1 = j - i + 1$
- Si l'égalité vaut pour $j > i$, alors

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

d'où l'égalité dans le cas $j + 1$.

L'égalité est donc vraie pour tout $j > i$. □

Exercice 9.4 (Calcul d'une somme classique *). Nous allons étudier la somme $\sum_{k=0}^n k$.

1. Montrer que $\sum_{k=0}^n k + \sum_{k=0}^n (n - k) = \sum_{k=0}^n n$.
2. En déduire la valeur de $\sum_{k=0}^n k + \sum_{k=0}^n (n - k)$. *Indication : on peut écrire $n = n \times 1$ puis factoriser n .*
3. En réécrivant la somme précédente, en déduire la valeur de $\sum_{k=0}^n k$.

9.3 Suites particulières

Nous étudierons ici trois cas particuliers de suites définies par récurrence : les suites arithmétiques, géométriques et arithmético-géométriques.

9.3.1 Suite arithmétique

Définition 9.4. Une suite arithmétique, dite de raison r , est une suite définie par

$$\begin{cases} u_0 \in \mathbb{R} \\ \forall n \in \mathbb{N} \quad u_{n+1} = u_n + r \end{cases}$$

où $r \in \mathbb{R}$.

Nous allons déterminer le terme général d'une telle suite.

Proposition 9.3.1. *Le terme général d'une suite arithmétique de raison r est $u_n = u_0 + r \times n$.*

Démonstration. Nous allons prouver ce résultat par récurrence :

- Le résultat est évident pour u_0
- Si $u_n = u_0 + r \times n$ alors $u_{n+1} = r + u_n = u_0 + r \times (n + 1)$. D'où le résultat. □

9.3.2 Suite géométrique

Définition 9.5. Une suite géométrique, dite de raison q , est une suite définie par

$$\begin{cases} u_0 \in \mathbb{R} \\ \forall n \in \mathbb{N}, \quad u_{n+1} = q \times u_n \end{cases}$$

Donnons le terme général d'une telle suite.

Proposition 9.3.2. *Une suite de raison q a comme terme général*

$$u_n = u_0 \times q^n$$

Démonstration. Exercice laissé au lecteur. *Indication : refaire le modèle de la preuve précédente.* □

9.3.3 Somme géométrique

Nous allons désormais étudier un cas particulier de somme, appelé somme géométrique.

Proposition 9.3.3. *Soit q un réel différent de 1. Alors*

$$\sum_{k=0}^n q^k = \frac{q^{n+1} - 1}{q - 1}$$

Démonstration. Calculons d'abord $q \sum_{k=0}^n q^k$ puis $q \sum_{k=0}^n q^k - \sum_{k=0}^n q^k$:

$$\begin{aligned} q \sum_{k=0}^n q^k &= \sum_{k=0}^n q^{k+1} \\ q \sum_{k=0}^n q^k &= \sum_{k=1}^{n+1} q^k \\ q \sum_{k=0}^n q^k - \sum_{k=0}^n q^k &= q^{n+1} + \sum_{k=0}^n q^k - \sum_{k=0}^n q^k - q^0 \\ (q - 1) \sum_{k=0}^n q^k &= q^{n+1} - 1 \\ \sum_{k=0}^n q^k &= \frac{q^{n+1} - 1}{q - 1} \end{aligned}$$

D'où le résultat. □

9.3.4 Suite arithmético-géométrique

Nous allons définir ici une dernière forme de suite, qui est une combinaison des deux précédentes.

Définition 9.6. Une suite arithmético-géométrique est une suite de la forme

$$\begin{cases} u_0 \in \mathbb{R} \\ \forall n \in \mathbb{N}, \quad u_{n+1} = q \times u_n + r \end{cases}$$

Nous allons d'abord définir le point fixe d'une telle suite.

Définition 9.7. On appelle point fixe de la relation $u_{n+1} = q \times u_n + r$ un réel λ tel que $\lambda = q \times \lambda + r$. Trouver un point fixe équivaut à résoudre une équation du premier degré, donc l'existence d'un tel point fixe (pour $r \neq 1$) est évidente.

On peut alors trouver le terme général d'une suite arithmético-géométrique.

Théorème 9.1. *Soit (u) une suite arithmético-géométrique de point fixe λ . Alors $(u - \lambda)$ est une suite géométrique.*

Démonstration. On prouve par récurrence ce résultat :

- Au premier rang, il n'y a rien à vérifier.
- Puisque $\lambda = q \times \lambda + r$, on en déduit que $u_{n+1} - \lambda = q \times u_n + r - q \times \lambda - r = q(u_n - \lambda)$, ce qui est une relation de récurrence géométrique.

Donc la suite est bien géométrique, on en déduit la forme générale de la suite :

$$u_n = q^n(u_0 - \lambda) + \lambda$$

□

Exercice 9.5. Trouver le terme général de la suite

$$\begin{cases} u_0 = 2 \\ \forall n \in \mathbb{N}, \quad u_{n+1} = 3u_n + 2 \end{cases}$$

Exercice 9.6. Trouver le terme général de la suite

$$\begin{cases} u_0 = 5 \\ \forall n \in \mathbb{N}, \quad u_{n+1} = u_n + 2 \end{cases}$$

10 Approfondissement : Polynômes à une indéterminée

Cette section s'intéresse aux polynômes à coefficients dans \mathbb{R} et dans \mathbb{C} . C'est un approfondissement, et à ce titre peut être évité en première lecture : nous traiterons ici de problématiques hors du cadre du lycée. Dans la suite, \mathbb{K} désignera \mathbb{R} ou \mathbb{C} (au sens où ce que nous dirons pourra fonctionner pour les deux).

Un polynôme correspondra moralement à une fonction de la forme

$$x \mapsto a_0 + a_1x + a_2x^2 + \dots + a_nx^n$$

dont les zéros forment donc une équation de degré n .

La première partie définira les polynômes formels sur \mathbb{K} , la deuxième partie s'occupera de la structure de $\mathbb{K}[X]$, l'ensemble des polynômes, comme un anneau euclidien (ces termes seront expliqués). Enfin, la troisième partie utilisera cette structure pour prouver que la notion de polynôme que nous avons définie est la bonne, pour ensuite parler du théorème fondamental de l'algèbre.

10.1 Construction de l'ensemble des polynômes

Nous allons définir les polynômes non pas comme des fonctions, comme on pourrait s'y attendre, mais comme une suite. Cette suite sera la suite de coefficients a_0, a_1, \dots et contiendra un nombre fini de coefficients non nuls.

Définition 10.1. On appelle polynôme à coefficients dans \mathbb{K} un élément de $(\mathbb{K})^{\mathbb{N}}$ avec un nombre fini de coefficients non nuls. Puisque pour $(a_n) \in \mathbb{K}[X]$, $\{a_n \mid a_n \neq 0\}$ est fini, on peut trouver un i maximal tel que $a_i \neq 0$: on appelle ce nombre le degré du polynôme (a_n) et on le note $\deg(a)$. On note $\mathbb{K}[X]$ l'ensemble des polynômes à coefficients dans \mathbb{K} .

Nous pouvons maintenant définir les opérations sur les polynômes.

Définition 10.2. Soit (a) et (b) deux polynômes à coefficients dans \mathbb{K} . Alors on définit les deux suites $(a+b)$ et $(a \times b)$ par :

$$[a+b]_n = a_n + b_n$$

et

$$[a \times b]_n = \sum_{i=0}^n a_i b_{n-i}$$

Démonstration. Nous devons vérifier que ces opérations donnent bien des polynômes en retour.

Pour $(a+b)$, on remarque directement qu'à partir du terme $\max(\deg(a), \deg(b))$, les termes de l'addition sont tous nuls, donc $(a+b)$ possède seulement des coefficients nuls à partir de ce terme. On en déduit de plus que $\deg(a+b) \leq \max(\deg(a), \deg(b))$.

Pour $(a \times b)$, à partir du terme $\deg(a) + \deg(b)$, tous les termes sont nuls. En effet, $a_i b_{n-i} = 0$ car tous les termes non nuls de a_i sont inférieurs à $\deg(a)$, mais alors b_{n-i} est nul car d'indice supérieur strictement à $\deg(b)$. De plus, comme $a_{\deg(a)} b_{\deg(b)} \neq 0$, on en déduit que $\deg(a \times b) = \deg(a) + \deg(b)$. \square

Les lois ainsi définies sont commutatives et associatives.

Exercice 10.1. Montrer que $(a+b) = (b+a)$, que $(a+(b+c)) = ((a+b)+c)$ et que ces identités sont aussi valables pour la multiplication.

Les lois de distributivité habituelles fonctionnent encore ici.

Proposition 10.1.1. Soient $(a), (b), (c) \in \mathbb{K}[X]$, alors $(a+b) \times c = a \times c + b \times c$ et $a \times (b+c) = a \times b + a \times c$.

Démonstration. Nous regardons à un indice n fixé ces suites :

$$\begin{aligned}
 [(a+b) \times c]_n &= \sum_{i=0}^n [a+b]_i c_{n-i} \\
 &= \sum_{i=0}^n (a_i + b_i) c_{n-i} \\
 &= \sum_{i=0}^n (a_i c_{n-i} + b_i c_{n-i}) \\
 &= \sum_{i=0}^n a_i c_{n-i} + \sum_{i=0}^n b_i c_{n-i} \\
 &= [a \times c]_n + [b \times c]_n \\
 [(a+b) \times c]_n &= [a \times c + b \times c]_n
 \end{aligned}$$

$$\begin{aligned}
 [a \times (b+c)]_n &= \sum_{i=0}^n a_i [b+c]_{n-i} \\
 &= \sum_{i=0}^n a_i (b_{n-i} + c_{n-i}) \\
 &= \sum_{i=0}^n (a_i b_{n-i} + a_i c_{n-i}) \\
 &= \sum_{i=0}^n a_i b_{n-i} + \sum_{i=0}^n a_i c_{n-i} \\
 [a \times (b+c)]_n &= [a \times b + a \times c]_n
 \end{aligned}$$

Donc, puisque les suites coïncident sur chaque $n \in \mathbb{N}$, on en déduit les égalités. \square

Nous avons aussi des éléments neutres.

Proposition 10.1.2. 0 est neutre pour l'addition et 1 pour la multiplication, où $0 = (0, 0, \dots)$ et $1 = (1, 0, 0, \dots)$.

Démonstration. Pour tout $n \in \mathbb{N}$, $a_n = a_n + 0$ et $a_n = a_n \times 1 + 0$, d'où le résultat. \square

Les éléments ont, de plus, un opposé.

Proposition 10.1.3. Soit $(a) \in \mathbb{K}[X]$. Alors on définit $(-a)$ par $[-a]_n = a_n$. Cet élément a la propriété que $(a + (-a)) = (0)$.

Démonstration. Le résultat est direct. \square

Enfin, on définit la multiplication d'un polynôme par un élément de \mathbb{K} de la façon suivante :

Définition 10.3. Soit $(a) \in \mathbb{K}[X]$ et $b \in \mathbb{K}$, on définit $b \cdot a$ par $[b \cdot a]_n = b \times a_n$.

On va désormais noter $X = (0, 1, 0, 0, 0, 0, \dots)$ la suite valant 1 seulement à l'indice 1 et 0 ailleurs.

Proposition 10.1.4. Soit (a) un polynôme de degré n . Alors (a) s'écrit comme

$$a = \sum_{k=0}^n a_k X^k$$

Démonstration. En effet, on prouve par récurrence que X^k est la suite valant 1 à l'indice k et 0 ailleurs (récurrence laissée au lecteur). Alors on remarque de plus que $a_k X^k$ vaut exactement a_k à l'indice k et 0 ailleurs, d'où en calculant toute la somme, le résultat. \square

Nous noterons désormais $P, Q, P(X)$ ou $Q(X)$ les polynômes et respectivement (p_n) et (q_n) leurs coefficients. De plus, nous allons définir la composition et l'évaluation.

Définition 10.4. Soient P et Q deux polynômes, où P est de degré n . La composition de Q par P , notée $P \circ Q$ ou $P(Q)$, est définie par

$$P(Q(X)) = \sum_{k=0}^n p_k(Q(X))^k$$

Proposition 10.1.5. *Le polynôme X est neutre pour la composition.*

Démonstration. Ce résultat se déduit directement de la décomposition de la proposition 10.1.4. \square

Définition 10.5. Soit $\alpha \in \mathbb{K}$ et $P \in \mathbb{K}[X]$, on définit l'évaluation $P(\alpha) \in \mathbb{K}$ de P en α par

$$P(\alpha) = \sum_{k=0}^n p_k \alpha^k$$

10.2 Les polynômes forment un anneau euclidien

Nous avons vu que $\mathbb{K}[X]$ est un ensemble muni d'une addition, d'une multiplication, d'un neutre 0 pour l'addition, d'un neutre 1 pour la multiplication et pour tout élément, d'un opposé pour l'addition. On appelle cette structure un anneau. Remarquons que \mathbb{R}, \mathbb{C} et même \mathbb{Z} sont des anneaux.

Cette structure est large et la structure des polynômes est plus particulière. En effet, c'est ce qu'on appelle un anneau euclidien, car on peut y définir une division euclidienne.

Théorème 10.1. *Soient P et S deux polynômes, P non nul. Alors il existe un unique couple de deux polynômes (Q, R) tels que*

$$P = S \times Q + R \wedge \deg(R) < \deg(S)$$

On appelle cette décomposition la division euclidienne de P par S .

Démonstration. Ce théorème sera admis dans ce document. \square

Ce résultat nous permet directement d'en déduire un résultat fondamental dans l'étude des polynômes : le lien entre une racine et la factorisation d'un polynôme.

Proposition 10.2.1. *Soit P un polynôme non nul et $\alpha \in \mathbb{K}$ tel que $P(\alpha) = 0$. Alors il existe $Q \in \mathbb{K}[X]$ tel que $P = (X - \alpha) \times Q$.*

Démonstration. Nous allons faire la division euclidienne de P par $(X - \alpha)$. On trouve donc (Q, R) tels que $P = (X - \alpha)Q + R$. Or en évaluant cette égalité en α , on en déduit, comme $P(\alpha) = 0$ et que $(\alpha - \alpha) \times (Q(\alpha)) = 0$, que $R(\alpha) = 0$. Or on sait que le degré de R est strictement inférieur à 1 : il est donc constant et vaut 0. Donc $P = (X - \alpha)Q$. \square

10.3 Polynôme scindé

Nous pouvons maintenant déduire une propriété importante pour exprimer un polynôme de façon unique.

Proposition 10.3.1. *Soit P un polynôme de degré inférieur n . Si P s'annule au moins $n - 1$ fois, alors celui-ci est nul.*

Démonstration. En effet, D'après le théorème précédent, si P est non nul, on trouve $n + 1$ monômes de la forme $(X - \alpha_i)$, donnant un polynôme de degré strictement supérieur à n . On en déduit par l'absurde que P est nul. \square

Proposition 10.3.2. *Soient P et Q des polynômes de degré inférieur à n coïncidant sur au moins $n + 1$ points. Alors $P = Q$.*

Démonstration. On remarque que $P - Q$ a plus de n points d'annulation, donc est nul par la propriété précédente. D'où $P = Q$. \square

On en déduit alors qu'il y a une bijection entre les fonctions polynômiales et les polynômes ainsi construits.

Théorème 10.2. *L'ensemble $\mathbb{K}[X]$ est en bijection avec l'ensemble $\mathcal{F}_{\mathbb{K}[X]}$ des fonctions polynômiales.*

Démonstration. La surjectivité est évidente puisqu'une fonction polynômiale est exactement une fonction de la forme $x \mapsto P(x)$.

L'injectivité provient du fait que \mathbb{K} est infini, et donc que deux fonctions identiques ont une infinité de points communs. Les polynômes associés coïncident donc sur un nombre de points supérieur à leur degré : ils sont donc égaux.

Il y a donc bijectivité entre les deux ensembles. \square

Nous avons donc prouvé que notre définition de polynôme est cohérente pour travailler sur les fonctions polynômiales.

Nous allons enfin développer la définition de polynôme scindé.

Définition 10.6. Soit $P \in \mathbb{K}[X]$. On dit que P est scindé s'il existe un nombre fini de α_i tels que $P(X) = a \prod_{i=1}^n (X - \alpha_i)$ où a est le coefficient de X^n dans P .

Remarque. Un polynôme est scindé si et seulement s'il a autant de racines que son degré.

Tout polynôme n'est pas scindé, par exemple dans $\mathbb{R}[X]$, le polynôme $X^2 + 1$ est toujours strictement positif. Pourtant, il existe un théorème important, appelé théorème fondamental de l'algèbre, et démontré par d'Alembert et Gauss, disant que tout polynôme dans \mathbb{C} est scindé.

Théorème 10.3 (D'Alembert-Gauss). *Soit $P \in \mathbb{C}[X]$ de degré supérieur à 1. Alors il existe une racine à P .*

Exercice 10.2. Montrer l'équivalence entre cet énoncé et celui expliqué plus haut : « tout polynôme dans $\mathbb{C}[X]$ est scindé. »

11 Approfondissement : Un peu d'algèbre générale

Dans cette section, nous allons introduire des notions d'algèbre générale telle qu'on peut en fait dans le supérieur. Nous allons introduire les notions de loi de composition interne puis de groupe, avec les définitions usuelles accompagnant les groupes.

Le but de l'algèbre générale est de s'abstraire des cas particuliers et d'étudier des structures dans leur généralité. Historiquement cette façon d'aborder les mathématiques est née assez tard, puisqu'il a fallu étudier d'abord beaucoup de structures pour trouver quelles abstractions sont pertinentes. Même si nous chercherons à donner des exemples, nous allons procéder de façon plus dogmatique : nous allons introduire des notions dont l'utilité peut sembler nulle. L'élève qui ne sera pas intéressé par ce chapitre pourra donc aisément l'ignorer, mais l'auteur étant un amateur d'algèbre, il considère que la beauté de cette théorie de la généralité peut toucher même des lycéens.

11.1 Loi de composition interne, monoïde, groupe

L'étude de l'algèbre est surtout centrée sur l'étude d'ensembles munis d'opérations. Définissons donc ce qu'est une opération, aussi appelée loi de composition interne (abrégée l.c.i.) :

Définition 11.1 (Loi de composition interne, magma). Soit E un ensemble. On appelle une loi de composition interne une fonction $\cdot : E \times E \rightarrow E$. On notera $x \cdot y$ à la place de $\cdot(x, y)$ (on appelle cette notation la notation infixe).

Un couple (E, \cdot) où E est un ensemble et \cdot est une loi de composition interne sur E est appelé magma.

Exercice 11.1. Soit E de cardinal n , dénombrer le nombre de lci qu'il peut exister sur E .

Exemple 11.1. Voici plusieurs magmas :

- Pour un ensemble E , $(\mathcal{P}(E), \cup)$ est un magma, il en est de même pour les opérations ensemblistes binaires $(\cup, \Delta, \setminus \text{ etc.})$
- $(\mathbb{N}, +)$
- $(\mathbb{Z}, +)$
- $(\mathbb{R}, +)$
- (\mathbb{N}, \times)
- $(\mathbb{Z}, -)$
- Soit le plan \mathcal{P} usuel. Si A et B sont deux points, on note $f(A, B)$ le symétrique de B par rapport à A . Alors (\mathcal{P}, f) est un magma.

La notion de magma est cependant très faible : aucune hypothèse n'apparaît dans cette définition, ce qui rend cette structure presque impossible à utiliser. Nous allons donc chercher des propriétés que peuvent avoir les lci.

Définition 11.2 (Propriétés d'une lci). Soit (E, \cdot) un magma.

Nous allons donner ici une liste de propriétés classiques que peut respecter une lci :

- Commutativité : $\forall x \in E, \forall y \in E, x \cdot y = y \cdot x$.
- Associativité : $\forall x \in E, \forall y \in E, \forall z \in E, x \cdot (y \cdot z) = (x \cdot y) \cdot z$
- Existence d'un élément neutre : $\exists e \in E, \forall x \in E, e \cdot x = x \wedge x \cdot e = x$
- Si \cdot possède un élément neutre, que l'on notera e , alors pour $x \in E$ on dit que x' est le symétrique de x si $x \cdot x' = e$ et $x' \cdot x = e$.

Exercice 11.2. Montrer que si $x \in E$ admet un symétrique, alors ce symétrique est unique.

La première structure qui a un peu d'intérêt que nous pouvons introduire est celle de monoïde :

Définition 11.3 (Monoïde). Un monoïde (M, \cdot) est un magma (M, \cdot) tel que \cdot est associative et possède un élément neutre, qu'on notera généralement e_M voire e si le contexte n'est pas ambigu.

On dit que (M, \cdot) est un monoïde commutatif lorsque c'est un monoïde dont la loi est commutative.

Exercice 11.3. Parmi les exemples de 11.1, y a-t-il des monoïdes ? Si oui, lesquels ? Même question pour les monoïdes commutatifs.

Exercice 11.4. Montrer que dans un monoïde, le neutre est unique.

La notion encore plus forte, et qui est celle qui nous intéresse vraiment, est celle de groupe :

Définition 11.4 (Groupe). Un groupe (G, \cdot) est un monoïde pour lequel tout élément possède un inverse. Si $x \in G$, on notera l'inverse de x par x^{-1} .

Exercice 11.5. Soit (G, \cdot) un groupe et $x \in G$ un élément idempotent, c'est-à-dire tel que $x \cdot x = x$. Montrer que x est le neutre de G .

Exercice 11.6. Soit (G, \cdot) un groupe. Montrer que pour tout $x \in G$, $(x^{-1})^{-1} = x$. Montrer que pour tous $x, y \in G$, $(x \cdot y)^{-1} = y^{-1} \cdot x^{-1}$.

11.2 Sous-groupe, morphisme de groupe, noyau et image

Une notion fondamentale en algèbre est celle de sous-structure. L'idée est que si l'on prend un ensemble muni d'une structure (un magma, un monoïde, un groupe, ou d'autres structures plus exotiques), on peut considérer des parties de cet ensemble qui restent stables pour la structure, définissant alors une partie structurée à part entière. Nous allons donner une définition rigoureuse dans le cas d'un monoïde et d'un groupe.

Définition 11.5 (Sous-monoïde, sous-groupe). Soit (M, \cdot) un monoïde, de neutre e . On dit que M' est un sous-monoïde de M si :

- $M' \subseteq M$
- $e \in M'$
- si $m \in M'$ et $m' \in M'$ alors $m \cdot m' \in M'$.

Si M' est un sous-monoïde de M , alors (M', \cdot) est un monoïde (où \cdot est la restriction de la loi \cdot de M à M'). De plus, si M est un monoïde commutatif alors M' l'est aussi.

Soit (G, \cdot) un groupe de neutre e . On dit que G' est un sous-groupe de G si :

- $G' \subseteq G$
- $e \in G'$
- si $x \in G'$ et $y \in G'$ alors $x \cdot y \in G'$
- si $x \in G'$ alors $x^{-1} \in G'$.

Si G' est un sous-groupe de G , alors (G', \cdot) est un groupe (où \cdot est la restriction de la loi \cdot de G à G'). De plus, si G est un groupe commutatif alors G' l'est aussi.

Démonstration. Montrons que (M', \cdot) est un monoïde. Tout d'abord, l'opération \cdot restreinte à M' est bien définie (et totale), car si $x \in M'$ et $y \in M'$ alors $x \cdot y \in M'$. De plus, par définition, pour $x, y, z \in M'$,

$$(x \cdot_{M'} y) \cdot_{M'} z = (x \cdot_M y) \cdot_M z = x \cdot_M (y \cdot_M z) = x \cdot_{M'} (y \cdot_{M'} z)$$

puisque \cdot sur M est associative. De même, on peut vérifier que e est un élément neutre dans M' à partir du fait qu'il l'est dans M . Dans le cas où M est commutatif, on remarque que, pour $x, y \in M'$, $x \cdot y$ a la même valeur dans M et M' , et donc vaut $y \cdot x$.

Pour prouver que (G', \cdot) est un groupe, il nous suffit de montrer que chacun de ses éléments est inversible. En effet, on a déjà à partir du résultat précédent le fait que (G', \cdot) est un monoïde. Soit $x \in G'$, on trouve $x^{-1} \in G'$ d'après nos hypothèses. On vérifie alors que $x \cdot x^{-1} = e$ puisque c'est le cas dans G . \square

Exercice 11.7. Montrer que $2\mathbb{N}$, l'ensemble des entiers naturels pairs, est un sous-monoïde de $(\mathbb{N}, +)$. Montrer que $2\mathbb{Z}$, l'ensemble des entiers relatifs pairs, est un sous-groupe de $(\mathbb{Z}, +)$.

Exercice 11.8 (Caractérisation d'un sous-groupe). Soit (G, \cdot) un groupe, et $G' \subseteq G$ non vide. Montrer que G' est un sous-groupe de G si et seulement si pour tous $x, y \in G'$, $x \cdot y^{-1} \in G'$.

Exercice 11.9. Soit (G, \cdot) un groupe de neutre e . Montrer que G et $\{e\}$ sont des sous-groupes de G (on les appelle parfois les sous-groupes triviaux de G).

La prochaine définition que nous introduisons est un autre point essentiel de l'algèbre générale : celui de morphisme. L'idée d'un morphisme est une application qui va être compatible avec la structure. Dans le cas des groupes, cela se traduit par le fait que l'image du produit est le produit des images :

Définition 11.6 (Morphisme de groupes). Soient (G, \cdot_G) et (H, \cdot_H) deux groupes. $f : G \rightarrow H$ est un morphisme de groupes lorsque la proposition suivante est vérifiée :

$$\forall x \in G, \forall y \in G, f(x \cdot_G y) = f(x) \cdot_H f(y)$$

Exercice 11.10. En reprenant les notations de la définition précédente, et en notant e_G et e_H les neutres des groupes respectifs, montrer que $f(e_G) = e_H$. Montrer que pour tout $x \in G$, $f(x^{-1}) = (f(x))^{-1}$.

On notera en général \cdot pour toutes les lois de groupes, en considérant qu'il n'y a pas de confusion, mais un lecteur peu à l'aise avec la théorie des groupes devra faire attention à comprendre dans quel ensemble chaque objet écrit habite.

La propriété suivante sera particulièrement utile dans ses implications :

Proposition 11.2.1. Soient (G, \cdot) et (H, \cdot) deux groupes, et $f : G \rightarrow H$ un morphisme de groupes. Soient G' (respectivement H') un sous-groupe de G (respectivement de H). Alors $f(G')$ est un sous-groupe de H et $f^{-1}(H')$ est un sous-groupe de G .

Démonstration. Montrons que $f(G')$ est un sous-groupe de H :

- Comme $G' \subseteq G$ et par définition de l'image directe, on a $f(G') \subseteq H$.
- D'après l'exercice précédent, $f(e_G) = e_H$, et $e_G \in G'$ car G' est un sous-groupe de G , donc $e_H \in f(G')$.
- Soient $a, b \in f(G')$, alors par définition de $f(G')$ on trouve $x, y \in G'$ tels que $f(x) = a$ et $f(y) = b$. Alors

$$a \cdot b = f(x) \cdot f(y) = f(x \cdot y)$$

donc $a \cdot b \in f(G')$ car G' est stable par produit.

- Soit $a \in f(G')$, par définition on trouve $x \in G'$ tel que $f(x) = a$, or $f(x)^{-1} = f(x^{-1})$ donc $a^{-1} \in f(G')$ car $x^{-1} \in G'$.

Donc $f(G')$ est bien un sous-groupe de H .

Montrons que $f^{-1}(H')$ est un sous-groupe de G :

- Là encore, la définition d'image réciproque d'une partie $H' \subseteq H$ nous donne directement que $f^{-1}(H') \subseteq G$.
- Puisque $f(e_G) = e_H$ et que $e_H \in H'$ (car H' est un sous-groupe), on en déduit que $e_G \in f^{-1}(H')$.
- Soient $x, y \in f^{-1}(H')$. On peut alors écrire $f(x \cdot y) = f(x) \cdot f(y)$ or $f(x) \in H'$ et $f(y) \in H'$ par définition, et comme H' est un sous-groupe de H , il est stable par produit, donc $x \cdot y \in f^{-1}(H')$.
- Soit $x \in f^{-1}(H')$. Comme $f(x^{-1}) = (f(x))^{-1}$ et que H' est stable par passage au symétrique, on en déduit que $f(x^{-1}) \in H'$, donc que $x^{-1} \in f^{-1}(H')$.

Donc $f^{-1}(H')$ est bien un sous-groupe de G . □

Ce résultat signifie que les morphismes induisent des correspondances entre les sous-structures. Cependant, la considération inverse est très fructueuse aussi : étudier les sous-structures liées à un morphisme permet de mieux étudier ses propriétés.

Définition 11.7 (Noyau, Image). Soient (G, \cdot) et (H, \cdot) deux groupes, et $f : G \rightarrow H$ un morphisme de groupes. On définit l'image de f , notée $\text{Im}(f)$, par

$$\text{Im}(f) = f(G)$$

On définit le noyau de f , noté $\ker(f)$, par

$$\ker(f) = f^{-1}(\{e_H\})$$

L'intuition derrière ces deux définitions est la suivante : $\text{Im}(f)$ représente l'étendue des valeurs atteintes par f , et $\ker(f)$ représente à quel point les points s'accumulent en le neutre de l'image. Ces objets sont étroitement liés respectivement à la surjectivité et à l'injectivité de f , comme nous allons le prouver.

Proposition 11.2.2. *En reprenant les notations précédentes, on a les équivalences suivantes :*

- f est surjective si et seulement si $\text{Im}(f) = H$
- f est injective si et seulement si $\ker(f) = \{e_G\}$

Démonstration. Supposons que f est surjective. Soit $y \in H$, alors par surjectivité de f il existe $x \in G$ tel que $f(x) = y$, donc $y \in f(G) = \text{Im}(f)$. Comme par définition $\text{Im}(f) \subseteq H$, on en déduit que $H = \text{Im}(f)$. Réciproquement, si $\text{Im}(f) = H$, alors soit $y \in H$: puisque $H = \text{Im}(f)$, on en déduit que $y \in \text{Im}(f)$ donc on trouve $x \in G$ tel que $f(x) = y$, ce qui signifie que f est surjective.

Supposons que f est injective. Puisque $\ker(f)$ est un sous-groupe, $\{e_G\} \subseteq \ker(f)$. Comme f est injective, si l'on prend $x \in \ker(f)$, comme on sait que $f(x) = f(e_G)$, il en découle que $x = e_G$, donc $\{e_G\} = \ker(f)$. Réciproquement, si $\ker(f) = \{e_G\}$, alors soient $x, y \in G$, montrons que si $f(x) = f(y)$ alors $x = y$:

$$\begin{aligned} f(x) = f(y) &\implies f(x) \cdot f(y)^{-1} = e_H \\ &\implies f(x \cdot y^{-1}) = e_H \\ &\implies x \cdot y^{-1} \in \ker(f) \\ &\implies x \cdot y^{-1} = e_G \quad \text{car } \ker(f) = \{e_G\} \\ &\implies x = y \end{aligned}$$

donc f est injective. □

Cela nous donne donc un nouvel outil pour montrer qu'une fonction est bijective : si elle est un morphisme de groupes, il nous suffit de montrer que son image est l'ensemble d'arrivée, et que son noyau est réduit au neutre de l'ensemble de départ. Il se trouve que les morphismes bijectifs revêtent une importance particulière, car là où deux ensembles en bijections peuvent avoir des données différentes (par exemple \mathbb{N} et \mathbb{Z} ne sont pas équipés par défaut des mêmes structures), deux groupes tels qu'il existe un morphisme bijectif entre eux seront en général assimilés. En effet, du point de vue de la théorie des groupes, ces deux groupes auront strictement le même comportement et les résultats vrais sur l'un seront vrais sur l'autre. Nous allons donner quelques définitions pour pouvoir parler de ces aspects :

Définition 11.8. Soient (G, \cdot) et (H, \cdot) deux groupes. On dit que :

- Un morphisme de groupe $f : G \rightarrow G$ est un endomorphisme (ce terme s'applique à un morphisme dont le domaine est le codomaine).
- Un morphisme de groupe $f : G \rightarrow H$ bijectif est un isomorphisme.
- Un morphisme de groupe $f : G \rightarrow G$ bijectif est un automorphisme (ce terme s'applique à un isomorphisme d'un groupe dans lui-même).
- G et H sont isomorphes s'il existe un isomorphisme $f : G \rightarrow H$.

11.3 Exemple : des petits groupes

Nous allons tout d'abord étudier précisément un groupe comportant deux éléments. Comme nous l'avons dit, en théorie des groupes, nous nous importons peu du nom des éléments et l'important est de considérer les groupes à isomorphisme près. Cela signifie que nous parlerons *du* groupe à deux éléments quand nous aurons établi que tous deux groupes de cardinal 2 sont isomorphes.

Puisque nous voulons deux éléments, nous allons prendre comme support l'ensemble $2 = \{0, 1\}$, que nous noterons G_2 pour une meilleure lisibilité. Nous allons décrire la loi \cdot sur G_2 à partir de ce qu'on appelle une table de Cayley, qui est un tableau à deux entrées, présentant le résultat de l'opération $x \cdot y$ pour x et y des éléments de notre groupe. Tout d'abord, on sait qu'il existe un élément

neutre, nous déciderons que 0 est ce tel élément. Cela fixe donc que $0 \cdot 1 = 1 \cdot 0 = 1$, et que $0 \cdot 0 = 1$: il ne reste qu'à déterminer $1 \cdot 1$. Cependant, comme tout élément a un inverse, on sait qu'il existe x tel que $1 \cdot x = 0$, et $1 \cdot 0 \neq 0$, donc la seule possibilité est $1 \cdot 1 = 0$.

\cdot	0	1
0	0	1
1	1	0

TABLE 7 – Table de Cayley du groupe à deux éléments

On montre alors le résultat suivant pour justifier que ce groupe est le seul à isomorphisme près :

Proposition 11.3.1. *Soient (G, \cdot) et (H, \cdot) deux groupes à deux éléments, alors il existe un isomorphisme $\varphi : G \rightarrow H$.*

Démonstration. On note x_G et x_H respectivement l'élément de G différent du neutre e_G et l'élément de H différent du neutre e_H . On définit alors φ par $\varphi(e_G) = e_H$ et $\varphi(x_G) = x_H$.

φ est bien un morphisme de groupe par disjonction de cas : $\varphi(e_G \cdot e_G) = \varphi(e_G) \cdot \varphi(e_G)$ dans le premier cas, $\varphi(e_G \cdot x_G) = \varphi(e_G) \cdot \varphi(x_G)$ dans le deuxième, $\varphi(x_G \cdot e_G) = \varphi(x_G) \cdot \varphi(e_G)$ dans le troisième et enfin $\varphi(x_G \cdot x_G) = \varphi(x_G) \cdot \varphi(x_G)$.

Montrons que φ est surjective :

$$\text{Im}(\varphi) = \{\varphi(e_G), \varphi(x_G)\} = \{e_H, x_H\} = H$$

Montrons que φ est injective : $\ker(\varphi) = \{e_G\}$ puisque $\varphi(x_G) \neq e_H$.

Donc φ est un isomorphisme entre G et H . □

Ainsi il existe un unique groupe à deux éléments, à isomorphisme près.

Exercice 11.11 (*). Existe-t-il un unique groupe à trois éléments, à isomorphisme près ?

Quatrième partie

Analyse

12 Topologie des réels

Pour pouvoir étudier efficacement l'analyse, nous allons commencer par introduire la topologie de \mathbb{R} . Nous ne détaillerons cependant pas ce qu'est la topologie : nous utilisons ce terme uniquement comme indicateur de ce dans quoi se place notre étude. Nous allons principalement décrire \mathbb{R} , puis donner les notions de voisinage, de limite et de continuité, avant de donner le résultat essentiel de cette considération topologique : le théorème de Bolzano-Weierstrass.

12.1 Axiomes des l'ensemble des réels

D'abord, nous allons introduire ce qu'est l'ensemble \mathbb{R} . Nous allons présenter un ensemble d'axiomes paraissant évidents à propos de \mathbb{R} , mais qu'il sera trop laborieux de démontrer (car cela signifierait donner une construction explicite de \mathbb{R} , ce qui est hors du cadre de ce document).

Axiome 12.1 (Les réels forment un corps ordonné). *L'ensemble $(\mathbb{R}, +, \times, \leq)$ est un corps ordonné, ceci signifie que :*

- L'opération $+$: $\mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ est associative, càd

$$\forall x \in \mathbb{R}, \forall y \in \mathbb{R}, \forall z \in \mathbb{R}, (x + y) + z = x + (y + z)$$

ce qui signifie qu'on notera indifféremment $x + y + z$.

- L'opération $+$ est commutative, càd

$$\forall x \in \mathbb{R}, \forall y \in \mathbb{R}, x + y = y + x$$

- L'opération $+$ a l'élément neutre 0, càd

$$\forall x \in \mathbb{R}, x + 0 = x \wedge 0 + x = x$$

- Tout élément $x \in \mathbb{R}$ a un symétrique pour $+$, noté $-x$, tel que $x + (-x) = -x + x = 0$. On notera directement $a - b$ à la place de $a + (-b)$.
- L'opération \times est elle aussi associative, commutative et possède un élément neutre noté 1.
- Tout élément $x \in \mathbb{R} \setminus \{0\}$ possède un symétrique pour \times , noté $\frac{1}{x}$ ou x^{-1} , tel que $x \times x^{-1} = x^{-1} \times x = 1$.
- La relation \leq est une relation d'ordre totale sur \mathbb{R} .
- Pour tous $a, b, c, d \in \mathbb{R}$, si $a \leq b$ et $c \leq d$ alors $a + c \leq b + d$.
- Pour tous $a, b \in \mathbb{R}$, si $a \geq 0$ et $b \geq 0$ alors $a \times b \geq 0$.

Nous utiliserons en général des conséquences de ces axiomes, mais nous ne montreront pas en quoi ils découlent de ces définitions. Cette décision est motivée par l'évidence des manipulations en question (par exemple la règle des signes qui dit que si $a \leq 0$ et $b \geq 0$ alors $ab \leq 0$). Nous encourageons cependant un lecteur motivé à redémontrer ce genre de petits résultats à mesure qu'ils sont utilisés. De plus nous identifions évidemment \mathbb{N} comme une partie de \mathbb{R} , et de même pour \mathbb{Z} . Ainsi écrire $-7 \in \mathbb{R}$ est considéré comme valide (ce qui peut sembler évident, mais qui n'est pas forcément rigoureusement valide).

Enfin, remarquons un point important : ces axiomes sont aussi respectés par \mathbb{Q} , l'ensemble des nombres rationnels, qui n'est pourtant pas l'ensemble \mathbb{R} (puisque'il existe des nombres irrationnels). Que nous manque-t-il alors pour avoir un ensemble qui se comporte comme \mathbb{R} ? Nous allons justement introduire une propriété qui distingue \mathbb{Q} et \mathbb{R} , qui se trouve caractériser exactement \mathbb{R} .

Axiome 12.2 (Propriété de la borne supérieure). Soit $S \subseteq \mathbb{R}$ une partie majorée, c'est-à-dire telle qu'il existe $M \in \mathbb{R}$ tel que $\forall s \in S, s \leq M$. Alors il existe une borne supérieure à S , i.e. un élément \sup_S majorant S , donc tel que $\forall s \in S, s \leq \sup_S$, et tel que pour tout majorant M de S , on ait $\sup_S \leq M$.

Exercice 12.1 (*). Soit la partie $S = \{x \mid x \in \mathbb{Q}, x^2 \leq 2\}$. Montrer que $\sup_S = \sqrt{2}$ dans \mathbb{R} . En déduire que S n'a pas de borne supérieure dans \mathbb{Q} .

Remarque. Cet axiome signifie aussi que toute partie S minorée possède une borne inférieure, car la borne supérieure de l'ensemble $S' = \{-s \mid s \in S\}$ respecte la relation $-\sup_{S'} = \inf_S$. De plus, les bornes supérieure et inférieure sont uniques en procédant par double inégalité.

Nous allons donner caractérisation importante de la borne supérieure d'une partie majorée.

Proposition 12.1.1. Soit $S \subseteq \mathbb{R}$ une partie majorée et $s \in \mathbb{R}$. Alors s est la borne supérieure de S si et seulement si s est un majorant de S et pour tout $\varepsilon > 0$, il existe $x \in S$ tel que $s - \varepsilon < x$.

Démonstration. Si s est une borne supérieure de S , alors s est majorant de S . Soit $\varepsilon > 0$, comme $s - \varepsilon < s$ on en déduit que $s - \varepsilon$ n'est pas un majorant de S , ce qui signifie qu'il est faux que pour tous les éléments $x \in S$, $x \leq s - \varepsilon$, i.e. qu'il existe $x \in S$ tel que $s - \varepsilon < x$.

Réciproquement, si l'autre proposition est vérifiée, alors s est majorant de S et, par l'absurde, si M est un majorant de S tel que $M < s$, alors $s - M > 0$ donc en considérant $\varepsilon = s - M$, on trouve $x \in S$ tel que $s - (s - M) = M < x$, donc M n'est pas majorant de S . Donc $s \leq M$, donc s est la borne supérieure de S . \square

12.2 Ensembles ouverts et voisinages

Nous pouvons alors construire la topologie de \mathbb{R} : ceci signifie que l'on va donner une famille de parties de \mathbb{R} qui permettront de définir si un élément est plus proche d'un autre. Par exemple, si l'on prends trois points $x = 0$, $y = 1$ et $z = 2$, alors on peut considérer que y est plus proche de x que z parce que tout intervalle qui contient x et z contient aussi y . Plutôt que de considérer des intervalles, nous allons préciser notre description par la notion d'ouvert. Une façon classique d'introduire la topologie de \mathbb{R} est de donner la notion de distance sur \mathbb{R} , mais nous voulons commencer par développer le raisonnement plus ensembliste qui lui est lié, car cela permettra de donner plus de sens à des notions telles que les voisinages.

Pour que le lien avec la distance puisse s'établir facilement, et car la pure définition d'ouvert peut sembler absconse, nous allons d'abord définir les boules ouvertes.

Définition 12.1 (Boule ouverte). Soit $x \in \mathbb{R}$ et $r > 0$, on appelle boule ouverte de centre x et de rayon r , et l'on note $B(x, r)$, l'ensemble

$$B(x, r) =]x - r; x + r[= \{a \mid a \in \mathbb{R}, x - r < a < x + r\}$$

Définition 12.2 (Ouvert). On appelle ouvert de \mathbb{R} (ou simplement ouvert) une partie $U \subseteq \mathbb{R}$ qui vérifie la propriété suivante :

$$\forall x \in U, \exists \varepsilon > 0, B(x, \varepsilon) \subseteq U$$

Autrement dit, un ouvert U est un ensemble pour lequel, pour tout point x dans U , on peut trouver une boule ouverte assez petite centrée en x qui est contenue dans U .

Proposition 12.2.1. Soit $x \in \mathbb{R}$ et $r > 0$. Alors $B(x, r)$ est un ouvert.

Démonstration. Soit $y \in B(x, r)$. Par définition, $x - r < y < x + r$. Soit $m = \min(y - (x - r), x + r - y)$ la distance entre y et le bord de $B(x, r)$, montrons que $B(y, m) \subseteq B(x, r)$. Soit $z \in B(y, m)$, on en déduit que $y - m < z < y + m$ mais comme $m \leq y - (x - r)$ (i.e. $-m \geq x - r - y$) et $m \leq x + r - y$, par transitivité des inégalités on obtient $y + x - r - y < z < y + x + r - y$ soit $x - r < z < x + r$ ce qui signifie que $z \in B(x, r)$, d'où le résultat. \square

La notion d'ouvert étant assez peu intuitive pour la topologie de \mathbb{R} , nous allons directement définir les voisinages.

Définition 12.3 (Voisinage). Soit $x \in \mathbb{R}$, on appelle voisinage de x une partie $V \subseteq \mathbb{R}$ telle qu'il existe un ouvert $U \subseteq V$ contenant x .

On note \mathcal{V}_x l'ensemble des voisinages de x (c'est donc un ensemble d'ensembles de réels).

Proposition 12.2.2. Soit $x \in \mathbb{R}$, alors $V \in \mathcal{V}_x$ si et seulement s'il existe $\varepsilon > 0$ tel que $B(x, \varepsilon) \subseteq V$.

Démonstration. Supposons que V est un voisinage de x . Alors on trouve un ouvert U inclus dans V et contenant x . Par définition du fait que U est un ouvert, on trouve $\varepsilon > 0$ tel que $B(x, \varepsilon) \subseteq U$, donc par transitivité de l'inclusion on en déduit que $B(x, \varepsilon) \subseteq V$.

Réciproquement, supposons qu'on trouve $\varepsilon > 0$ tel que $B(x, \varepsilon) \subseteq V$. Alors comme $B(x, \varepsilon)$ est un ouvert, on a trouvé un ouvert contenant x inclus dans V : V est un voisinage de x . \square

Un voisinage est donc un ensemble « autour du point x » en ce sens que pour un voisinage donné de x , il existe une distance en-dessous de laquelle tous les points assez près de x sont dans ce voisinage. On peut donc par exemple avoir l'intuition de ce que signifie « x est plus près de y que z » par le fait qu'il existe plus de voisinages de y contenant x que de voisinages de z contenant x . Attention cependant, ce qui vient d'être écrit est seulement une intuition du résultat, car le terme de « plus de voisinages » n'est pas rigoureux.

Un ensemble de voisinages d'un point est ce que l'on appelle un filtre des parties de \mathbb{R} , mais nous n'allons pas détailler la théorie des filtres. Simplement nous allons donner des propriétés basiques des ensembles \mathcal{V}_a .

Proposition 12.2.3 (Les voisinages forment un filtre). Soit $a \in \mathbb{R}$, alors les propriétés suivantes sont vérifiées :

- Si $V \in \mathcal{V}_a$ et $V \subseteq W$ alors $W \in \mathcal{V}_a$ (en particulier l'union d'un voisinage de a avec toute partie est un voisinage de a).
- Si $V \in \mathcal{V}_a$ et $V' \in \mathcal{V}_a$ alors $V \cap V' \in \mathcal{V}_a$.
- $\mathbb{R} \in \mathcal{V}_a$.

Démonstration. Prouvons chaque propriété :

- Soit $V \in \mathcal{V}_a$, alors on trouve $\varepsilon > 0$ tel que $B(a, \varepsilon) \subseteq V$, donc par transitivité de l'inclusion $B(a, \varepsilon) \subseteq W$, donc $W \in \mathcal{V}_a$.
- Soit $V, V' \in \mathcal{V}_a$, alors on trouve $\varepsilon > 0$ et $\varepsilon' > 0$ tels que $B(a, \varepsilon) \subseteq V$ et $B(a, \varepsilon') \subseteq V'$. Alors en prenant $\varepsilon'' = \min(\varepsilon, \varepsilon')$ on déduit que $B(a, \varepsilon'') \subseteq V$ et $B(a, \varepsilon'') \subseteq V'$, d'où $B(a, \varepsilon'') \subseteq V \cap V'$, donc $V \cap V' \in \mathcal{V}_a$.
- Comme $B(a, 1) \subseteq \mathbb{R}$, on en déduit que $\mathbb{R} \in \mathcal{V}_a$.

\square

Introduisons la notion de base séquentielle de voisinage, qui sera utile plus tard.

Définition 12.4 (Base séquentielle de voisinages). Soit $a \in \mathbb{R}$. On dit que $(V_n)_{n \in \mathbb{N}}$ est une base séquentielle de voisinages si pour tout voisinage $V \in \mathcal{V}_a$, il existe un $n \in \mathbb{N}$ tel que $V_n \subseteq V$ et si $V_{n+1} \subseteq V_n$, i.e. (par récurrence) si $n > m \implies V_n \subseteq V_m$.

Remarque. Le fait de considérer que la base est décroissante pour l'inclusion est une convention prise dans ce document pour faciliter les démonstrations.

La plupart des propositions « vraies sur tous les voisinages » sont en fait équivalentes à être vraies sur une base séquentielle de voisinages, d'où l'intérêt de cette notion.

12.3 Valeur absolue et distance

Une autre façon de considérer si des points sont loin ou non les uns des autres est de définir une distance. Il existe dans \mathbb{R} une distance canonique, mais commençons par définir ce qu'est une distance en toute généralité :

Définition 12.5 (Distance). Soit E un ensemble, on dit que $d : E \times E \rightarrow \mathbb{R}_+$ est une distance quand les conditions suivantes sont vérifiées :

- Symétrie : $\forall x \in E, \forall y \in E, d(x, y) = d(y, x)$
- Séparation : $\forall x \in E, \forall y \in E, d(x, y) = 0 \iff x = y$
- Inégalité triangulaire : $\forall x \in E, \forall y \in E, \forall z \in E, d(x, z) \leq d(x, y) + d(y, z)$

Dans le cas qui nous intéresse (celui où $E = \mathbb{R}$), nous allons commencer par définir la valeur absolue.

Définition 12.6 (Valeur absolue). On définit la valeur absolue de x , notée $|x|$, par

$$|x| = \max(x, -x)$$

La valeur absolue d'un réel est donc le réel auquel on a supprimé l'information du signe. Par exemple $|2| = 2$, et $|-2| = 2$.

Exercice 12.2. Montrer que $|x| = |y|$ si et seulement si $x = y$ ou $x = -y$.

Exercice 12.3. Montrer que $|x| = \sqrt{x^2}$.

Exercice 12.4. Montrer que pour tous réels x, y le produit des valeurs absolues est la valeur absolue du produit : $|x| \times |y| = |x \times y|$

Remarque. On peut voir la valeur absolue comme le module complexe restreint à \mathbb{R} .

La valeur absolue peut se voir comme la distance à 0, et plus généralement la valeur absolue de $x - y$ correspond à la distance usuelle entre x et y sur la droite réelle. Montrons donc que c'est une distance.

Proposition 12.3.1. La fonction

$$(x, y) \mapsto |x - y|$$

est une distance sur \mathbb{R} .

Démonstration. Vérifions les axiomes d'une distance :

- Symétrie : soient x et y deux réels. Comme $x - y = -(y - x)$ on en déduit que $|x - y| = |y - x|$.
- Séparation : $|0| = 0$ donc $|x - x| = 0$ pour tout $x \in \mathbb{R}$. Réciproquement, pour $x, y \in \mathbb{R}$, si $|x - y| = 0$ alors $x - y = 0$ ou $x - y = 0$, ce qui dans les deux cas est équivalent à $x = y$. Donc $|x - y| = 0 \iff x = y$.
- Inégalité triangulaire : on veut montrer que $|x - z| \leq |x - y| + |y - z|$, il nous suffit pour cela de montrer que pour tout $a, b \in \mathbb{R}, |a + b| \leq |a| + |b|$ (puisque $x - z = (x - y) + (y - z)$). D'abord $2 \times |ab| \geq 2ab$, et $(|a| + |b|)^2 = a^2 + b^2 + 2|ab|$ car $a^2 = |a|^2$ et de même pour b . On en déduit que $(|a| + |b|)^2 \geq (a + b)^2$ donc, en en prenant la racine carrée, que $|a| + |b| \geq |a + b|$.

La valeur absolue de la différence est donc une distance sur \mathbb{R} . □

Corollaire 12.3.1 (Inégalité triangulaire). Nous avons montré un résultat intermédiaire qui est utile en tant que tel :

$$\forall x \in \mathbb{R}, \forall y \in \mathbb{R}, |x + y| \leq |x| + |y|$$

Corollaire 12.3.2 (Deuxième inégalité triangulaire). Un résultat plus fort mais qui découle de l'inégalité triangulaire est le suivant :

$$\forall x \in \mathbb{R}, \forall y \in \mathbb{R}, ||x| - |y|| \leq |x + y| \leq |x| + |y|$$

Démonstration. Comme $x = (x + y) - y$ on en déduit en passant à la valeur absolue et à l'inégalité triangulaire que $|x| \leq |x + y| + |y|$ et de même avec $y = (y + x) - x$ on déduit que $|y| \leq |x + y| + |x|$. Ces deux équations sont équivalentes respectivement à $|x| - |y| \leq |x + y|$ et $|y| - |x| \leq |x + y|$ mais cela signifie que $||x| - |y|| \leq |x + y|$ (car $|x + y|$ majore une valeur et son opposé, donc sa valeur absolue). □

Exercice 12.5. Soient a et r deux réels. Montrer que l'ensemble des solutions de l'inéquation

$$|x - a| < r$$

d'inconnue x est exactement $B(a, r)$.

Nous pouvons désormais donner de nouvelles caractérisations des boules ouvertes et des voisinages.

Proposition 12.3.2. La boule ouverte $B(x, r)$ est exactement l'ensemble $\{y \mid y \in \mathbb{R}, |x - y| < r\}$.

Démonstration. Le résultat vient directement de l'exercice 12.5. □

Proposition 12.3.3. Une partie $V \subseteq \mathbb{R}$ est un voisinage de x si et seulement s'il existe $\varepsilon > 0$ tel que la propriété suivante est vérifiée :

$$\forall y \in \mathbb{R}, |x - y| < \varepsilon \implies y \in V \quad (1)$$

Démonstration. V est un voisinage de x si et seulement s'il existe une boule $B(x, \varepsilon)$ contenue dans V , avec $\varepsilon > 0$. En utilisant la nouvelle caractérisation de boule ouverte et la définition d'inclusion, on en déduit que $|x - y| < \varepsilon \implies y \in V$ est équivalent à $B(x, \varepsilon) \subseteq V$.

Donc V est un voisinage de x si et seulement s'il existe $\varepsilon > 0$ tel que (1) est vérifiée. □

12.4 Limite et caractérisation séquentielle

Les notions de voisinage, d'ouvert et de distance servent avant tout à pouvoir définir la notion de limite. Nous allons définir la limite d'une suite en premier lieu, puis celle d'une fonction. L'idée derrière une limite est de définir ce que signifie « s'approcher aussi près que possible d'un point » : la notion de voisinage permet justement de traduire ça. Pour pouvoir unifier un peu les définitions sur les limites de suites, nous allons introduire une nouvelle notion de voisinage, appelé voisinage de $\pm\infty$.

Définition 12.7 (Voisinage aux infinis). On dit que V est un voisinage de $+\infty$ (respectivement $-\infty$) si V contient un intervalle de la forme $]a; +\infty[$ (respectivement $] - \infty; a[$). On note respectivement $\mathcal{V}_{+\infty}$ et $\mathcal{V}_{-\infty}$ les ensembles des voisinage de $+\infty$ et de $-\infty$.

Le lecteur voulant s'exercer peut vérifier que le comportement de $\mathcal{V}_{+\infty}$ et $\mathcal{V}_{-\infty}$ est celui attendu d'un ensemble de voisinages.

Définition 12.8 (Limite d'une suite réelle). Soit $(u_n)_{n \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ une suite réelle et $l \in \mathbb{R}$. On dit que (u_n) converge vers l si pour tout voisinage V de l , il existe un indice $n \in \mathbb{N}$ tel que pour tout $p > n$, $u_p \in V$. Décrit plus formellement, la définition est donc

$$\forall V \in \mathcal{V}_l, \exists n \in \mathbb{N}, \forall p > n, u_p \in V$$

on note alors $\lim_{n \rightarrow \infty} u_n = l$

On dit que (u_n) diverge vers $+\infty$ (respectivement $-\infty$) si pour tout voisinage V de $+\infty$ (respectivement $-\infty$) il existe un indice $n \in \mathbb{N}$ à partir duquel toutes les valeurs de u_n sont dans V . Plus formellement la définition est

$$\forall V \in \mathcal{V}_{\pm\infty}, \exists n \in \mathbb{N}, \forall p > n, u_p \in V$$

on note alors $\lim_{n \rightarrow \infty} u_n = +\infty$ (respectivement $\lim_{n \rightarrow \infty} u_n = -\infty$).

Remarque. S'il existe un réel l tel que (u_n) converge vers l , on dit que (u_n) converge. S'il n'existe pas de telle limite, on dit que (u_n) diverge (remarquons que diverger vers un infini est une condition plus forte que diverger, par exemple $((-1)^n)$ diverge mais ne diverge vers aucun infini).

Remarque. On note souvent $\lim u_n$ directement, auquel cas il est sous-entendu $\lim_{n \rightarrow \infty} u_n$.

Donnons une caractérisation plus pratique des limites.

Proposition 12.4.1 (Caractérisation des limites). *Soit (u_n) une suite réelle et l un réel. Alors*

$$\lim_{n \rightarrow \infty} u_n = l \iff \forall \varepsilon > 0, \exists n \in \mathbb{N}, \forall p > n, |u_p - l| < \varepsilon$$

De même dans le cas d'une divergence vers l'infini :

$$\lim_{n \rightarrow \infty} u_n = +\infty \iff \forall M \in \mathbb{R}, \exists n \in \mathbb{N}, \forall p > n, u_p > M$$

$$\lim_{n \rightarrow \infty} u_n = -\infty \iff \forall m \in \mathbb{R}, \exists n \in \mathbb{N}, \forall p > n, u_p < m$$

Démonstration. Supposons que $\lim u_n = l$. Soit $\varepsilon > 0$, alors $B(l, \varepsilon)$ est un voisinage de x , donc on trouve $n \in \mathbb{N}$ tel que $\forall p > n, u_p \in B(l, \varepsilon)$ ce qui est équivalent à $\forall p > n, |u_p - l| < \varepsilon$, donc $\boxed{\forall \varepsilon > 0, \exists n \in \mathbb{N}, \forall p > n, |u_p - l| < \varepsilon}$. Réciproquement, soit $V \in \mathcal{V}_l$, par la proposition 12.3.3, on trouve $\varepsilon > 0$ tel que $\forall y \in \mathbb{R}, |l - y| < \varepsilon \implies y \in V$. En utilisant notre hypothèse avec le ε ainsi introduit, on en déduit qu'il existe $n \in \mathbb{N}$ tel que $\forall p > n, |u_p - l| < \varepsilon$, ce qui signifie que $u_p \in V$. On en déduit donc que $\boxed{\lim_{n \rightarrow \infty} u_n = l}$.

Supposons que $\lim u_n = +\infty$. Soit $M \in \mathbb{R}$. Par définition, $]M; +\infty[\in \mathcal{V}_{+\infty}$ donc on trouve un rang à partir duquel les termes de u_n sont dans $]M; +\infty[$, ce qui signifie par définition qu'on trouve un rang à partir duquel les termes de u_n sont supérieurs strictement à M . D'où $\boxed{\forall M \in \mathbb{R}, \exists n \in \mathbb{N}, \forall p > n, u_p > M}$. Réciproquement, soit un voisinage $V \in \mathcal{V}_{+\infty}$, on trouve une partie de la forme $]M; +\infty[\subseteq V$ et d'après notre hypothèse, on trouve un rang à partir duquel tous les termes sont supérieurs strictement à M , c'est-à-dire tous les termes appartiennent à V . Ce qui signifie que $\boxed{\lim u_n = +\infty}$. \square

Cette caractérisation nous sera utile dans un premier temps pour démontrer l'unicité de la limite.

Exercice 12.6 (Une autre caractérisation). Soit $a \in \mathbb{R} \cup \{-\infty, +\infty\}$, (V_n) une base séquentielle de voisinages de a et (u_n) une suite réelle. Montrer que $\lim u_n = a$ si et seulement si pour tout $n \in \mathbb{N}$ il existe $n_0 \in \mathbb{N}$ tel que pour tout $p > n_0$, on a $u_p \in V_n$.

On peut donc désormais noter $\lim u_n$ l'unique (nous le montrons juste après) $l \in \mathbb{R}$ tel que $\lim u_n = l$, si (u_n) est une suite convergente.

Exercice 12.7. Montrer que $\lim u_n = +\infty \iff \lim -u_n = -\infty$. En déduire la troisième équivalence de la proposition précédente.

Proposition 12.4.2. *Soit (u_n) une suite réelle. Si (u_n) converge vers l et vers l' alors $l = l'$.*

Démonstration. Nous allons démontrer ce résultat par l'absurde. Supposons que $l \neq l'$. Soit alors $\delta = \frac{|l - l'|}{3} > 0$. Par définition de la limite en l , on trouve n_1 tel que $\forall p > n_1, |u_p - l| < \delta$ et de même en l' , on trouve n_2 tel que $\forall p > n_2, |u_p - l'| < \delta$. Considérons $n = \max(n_1, n_2)$: pour tout $p > n$, on a à la fois $|u_p - l| < \delta$ et $|u_p - l'| < \delta$. En additionnant nos deux inéquations, on trouve

$$|u_p - l| + |u_p - l'| < 2\delta < |l - l'|$$

ce qui contredit l'inégalité triangulaire, puisque $|l - l'| \leq |l - u_p| + |u_p - l'|$. Le résultat est donc absurde : on en déduit que $\boxed{l = l'}$. \square

Exercice 12.8. Montrer que si (u_n) diverge vers un infini, alors (u_n) ne converge pas vers un réel, et ne peut converger que vers un seul infini.

Exercice 12.9. Soit $a \in \mathbb{R}$, montrer que $(]a - \frac{1}{n}; a + \frac{1}{n}[)_{n \in \mathbb{N}}$ est une base séquentielle de voisinages de a . Montrer que $(]n; +\infty[)_{n \in \mathbb{N}}$ est une base séquentielle de voisinages de $+\infty$. Montrer que $(]-\infty, -n])_{n \in \mathbb{N}}$ est une base séquentielle de voisinages de $-\infty$.

Exercice 12.10. Soit (u_n) une suite convergente et $k \in \mathbb{R}$, montrer que $\lim ku_n = k \lim u_n$.

Pour clore cette sous-partie, nous allons revenir sur la définition de borne supérieure : nous allons maintenant en donner une caractérisation séquentielle, c'est-à-dire une caractérisation utilisant des suites.

Proposition 12.4.3. Soit S une partie majorée de \mathbb{R} . Alors s est la borne supérieure de S si et seulement si s est un majorant de S et il existe une suite $(x_n)_{n \in \mathbb{N}} \in S^{\mathbb{N}}$ d'éléments de S telle que $\lim_{n \rightarrow \infty} x_n = s$.

Démonstration. Supposons que s est la borne supérieure de S . s est donc un majorant de S . D'après la proposition 12.1.1, pour tout $n \in \mathbb{N}$, on peut trouver $x_n \in S$ tel que $s - \frac{1}{n} < x_n$. Montrons que $\lim x_n = s$. Soit $\varepsilon > 0$, comme $\lim \frac{1}{n}$ on trouve un rang n_0 pour lequel pour tout $p > n_0$, $\frac{1}{p} < \varepsilon$ (il n'y a pas de valeur absolue car toutes les grandeurs sont positives, et on regarde la différence à 0). Alors, pour $p > n_0$, on a d'un côté $s - \frac{1}{p} < x_p$ et de l'autre $\frac{1}{p} < \varepsilon$, d'où $s - x_p < \varepsilon$. Donc $\lim x_n = s$.

Réciproquement, supposons que s est un majorant de S et qu'il existe x_n une suite d'éléments de S telle que $\lim x_n = s$. Soit $\varepsilon > 0$, on trouve $n \in \mathbb{N}$ tel que $\forall p > n, |x_p - s| < \varepsilon$, mais comme s est un majorant, l'inégalité est exactement $s - x_p < \varepsilon$ soit $s - \varepsilon < x_p$, donc il existe bien $x \in S$ tel que $s - \varepsilon < x$, donc s est la borne supérieure de S . \square

Remarque. Par dualité, on remarque que la borne inférieure d'un ensemble minoré est exactement un minorant tel qu'il existe une suite d'éléments de la partie qui converge vers cette valeur.

12.4.1 Calcul de limite de suites

Nous allons nous intéresser à plusieurs théorèmes essentiels pour calculer des limites de suites : le théorème d'encadrement, le théorème de la limite monotone et le théorème des suites adjacentes. Avant cela, nous allons donner plusieurs exercices permettant d'avoir des propriétés basiques pour calculer des limites.

Exercice 12.11. Montrer que $\lim \frac{1}{n} = 0$. Montrer que $\lim n = +\infty$. Montrer que $\lim x = x$ pour tout $x \in \mathbb{R}$. Montrer que $\lim \frac{1}{2^n} = 0$.

Exercice 12.12. Montrer que pour deux suites (u_n) et (v_n) convergentes, on a $\lim(u_n + v_n) = (\lim u_n) + (\lim v_n)$, $\lim(u_n - v_n) = (\lim u_n) - (\lim v_n)$ ainsi que $\lim(u_n \times v_n) = (\lim u_n) \times (\lim v_n)$.

Exercice 12.13. Montrer que si $\lim u_n = +\infty$ alors il existe un rang n_0 tel que pour tout $n > n_0$ on ait $u_n > 0$, montrer de plus que $\lim \frac{1}{u_n} = 0$.

Remarque. La suite $\left(\frac{1}{u_n}\right)_{n \in \mathbb{N}}$ n'est pas forcément définie pour tout n , mais comme on s'est assurés qu'elle l'était à partir d'un certain rang, et que la limite reste la même en supprimant les premiers termes, il n'y a pas de problème. Cependant cela signifie que l'on fait un abus de notation en confondant la suite avec la suite définie à partir d'un certain rang, il est bon de garder cela à l'esprit.

On peut trouver d'autres propriétés pour calculer les limites de suites, mais nous allons écrire les principales règles de calcul de limites dans le tableau Figure 22, pour commencer. Il faut lire ce tableau comme on lit une table de vérité : chaque ligne correspond à une certaine limite attribuée à (u_n) et (v_n) et donne alors les valeurs des autres expression en fonctions de ces valeurs. Les cases où il est indiqué FI (pour *Forme Indéterminée*) sont celle où l'on ne peut pas calculer la limite directement à partir de ses composantes. Par exemple on peut montrer que $\lim n^2 = +\infty$, donc le quotient $\frac{n^2}{n} = n$ tend vers $+\infty$ là où $\frac{n}{n^2}$ tend vers 0.

Exercice 12.14. Compléter le tableau avec les cas où u_n et v_n ont des limites finies non nulles, suivant le signe de leurs limites.

Proposition 12.4.4. Soient $P = (p_i)_{i \in \mathbb{N}}$ et $Q = (q_i)_{i \in \mathbb{N}}$ deux polynômes réels, de degrés respectifs m et r . Alors :

- si $m > r$,

$$\lim \frac{P(n)}{Q(n)} = +\infty$$

- si $m = r$,

$$\lim \frac{P(n)}{Q(n)} = \frac{p_m}{q_r}$$

$\lim u_n$	$\lim v_n$	$\lim(u_n + v_n)$	$\lim(u_n - v_n)$	$\lim(u_n \times v_n)$	$\lim(u_n/v_n)$
$-\infty$	$-\infty$	$-\infty$	FI	$+\infty$	FI
0	$-\infty$	$-\infty$	$+\infty$	FI	0
$+\infty$	$-\infty$	FI	$+\infty$	$+\infty$	FI
$-\infty$	0	$-\infty$	$-\infty$	FI	$-\infty$
0	0	0	0	0	FI
$+\infty$	0	$+\infty$	$+\infty$	FI	$+\infty$
$-\infty$	$+\infty$	FI	$-\infty$	$-\infty$	FI
0	$+\infty$	$+\infty$	$-\infty$	FI	0
$+\infty$	$+\infty$	$+\infty$	FI	$+\infty$	FI

FIGURE 22 – Tableau des limites usuelles

- si $m > r$,

$$\lim \frac{P(n)}{Q(n)} = 0$$

Démonstration. Tout d'abord, on montre par récurrence que $\forall k \in \mathbb{N}^*$, $\lim \frac{1}{n^k} = 0$:

- On a montré dans un exercice précédent que $\lim \frac{1}{n} = 0$.
- Soit $k \in \mathbb{N}$, supposons que $\lim \frac{1}{n^k} = 0$, alors par produit avec la suite des inverse de limite nulle, on en déduit que $\lim \frac{1}{n^{k+1}} = 0$

D'où le résultat.

Maintenant, on réécrit $P(n)$ et $Q(n)$ de la façon suivante :

$$\begin{aligned}
 P(n) &= \sum_{i=0}^m p_i n^i \\
 &= n^m \left(\sum_{i=0}^m p_i n^{i-m} \right) \\
 &= n^m \left(\sum_{i=0}^m p_{m-i} n^{-i} \right) \\
 Q(n) &= n^r \left(\sum_{i=0}^r q_{r-i} n^{-i} \right)
 \end{aligned}$$

Mais pour $i > 0$, $n^{-i} = \frac{1}{n^i}$ donc chaque terme tend vers 0. On en déduit donc que

$$\lim_{n \rightarrow \infty} \sum_{i=0}^m p_{m-i} n^{-i} = p_m \quad \text{et} \quad \lim_{n \rightarrow \infty} \sum_{i=0}^r q_{r-i} n^{-i} = q_r$$

d'où, en écrivant le quotient, l'égalité suivante :

$$\frac{P(n)}{Q(n)} = \frac{n^m \left(\sum_{i=0}^m p_{m-i} n^{-i} \right)}{n^r \left(\sum_{i=0}^r q_{r-i} n^{-i} \right)} = n^{m-r} R(n)$$

où R est une suite telle que $\lim R(n) = \frac{p_m}{q_r}$.

Le résultat découle alors immédiatement :

- si $m > r$, alors $m - r > 0$ donc par produit de limite, sachant que $\lim n^{m-r} = +\infty$ et $\lim R(n)$ est finie,

$$\boxed{\lim \frac{P(n)}{Q(n)} = +\infty}$$

- si $m = r$ alors $m - r = 0$ donc $\frac{P(n)}{Q(n)} = R(n)$, d'où directement

$$\lim \frac{P(n)}{Q(n)} = \frac{p_m}{q_m}$$

- si $m < r$ alors $m - r < 0$ donc par produit d'une limite finie par une limite nulle,

$$\lim \frac{P(n)}{Q(n)} = 0$$

□

Théorème 12.1 (Encadrement). Soient $(u_n), (v_n)$ et (w_n) des suites réelles, telles que $\forall n \in \mathbb{N}, u_n \leq v_n \leq w_n$ et $\lim u_n = \lim w_n$ (sous-entendu que u_n et w_n convergent). Alors $\lim v_n = \lim u_n = \lim w_n$.

Démonstration. Soit $l = \lim u_n = \lim w_n$. Soit $\varepsilon > 0$, montrons qu'à partir d'un certain rang, $|v_p - l| < \varepsilon$. Pour cela, on nomme n_0 et n_1 les indices respectifs à partir desquels $\forall p > n_0, |u_p - l| < \varepsilon$ et $\forall p > n_1, |w_p - l| < \varepsilon$. Soit $n_2 = \max(n_0, n_1)$, on a alors pour $p > n_0$ les deux équation suivante :

$$\begin{cases} l - \varepsilon < u_p < l + \varepsilon \\ l - \varepsilon < w_p < l + \varepsilon \end{cases} \text{ ce qui signifie, par l'encadrement de } v_p \text{ par } u_p \text{ et } w_p, \text{ que}$$

$$l - \varepsilon < v_p < l + \varepsilon$$

soit $|v_p - l| < \varepsilon$, d'où le résultat.

□

Exercice 12.15. En utilisant le théorème d'encadrement, déterminer la nature de la suite $\left(\frac{(-1)^n}{n}\right)$, ce qui signifie qu'il faut déterminer si la suite converge, et donner sa limite si elle converge.

Théorème 12.2 (Majoration, minoration). Soient $(u_n)_{n \in \mathbb{N}}, (v_n)_{n \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ deux suites réelles telles que $\forall n \in \mathbb{N}, u_n \leq v_n$. Alors

- si $\lim u_n = +\infty$ alors $\lim v_n = +\infty$.
- si $\lim v_n = -\infty$ alors $\lim u_n = -\infty$.

Démonstration. Nous n'allons prouver que le premier point, le deuxième sera un exercice. Supposons que $\lim u_n = +\infty$. Soit $V \in \mathcal{V}_{+\infty}$, par définition on trouve $M \in \mathbb{R}$ tel que $]M; +\infty[\subseteq V$. Par définition de $\lim u_n = +\infty$ et car $]M; +\infty[\in \mathcal{V}_{+\infty}$, on trouve n_0 tel que $\forall p > n_0, u_p > M$, mais comme pour tout $n \in \mathbb{N}, v_n \geq u_n$, on en déduit que $\forall p > n_0, v_p > M$, donc $\forall p > n_0, v_p \in V$. D'où $\lim v_n = +\infty$. □

Exercice 12.16. Montrer le deuxième point du théorème précédent.

Le second théorème est celui de la limite monotone, qui permet d'obtenir un résultat de convergence avec des hypothèses relativement faibles.

Théorème 12.3 (Limite monotone). Soit (u_n) une suite croissante (i.e. telle que pour tout $n \in \mathbb{N}, m \in \mathbb{N}, n \leq m \implies u_n \leq u_m$). Soit (u_n) est majorée, auquel cas elle converge vers la borne supérieure de ses valeurs atteinte, soit elle n'est pas majorée, auquel cas elle diverge vers $+\infty$.

Démonstration. Soit (u_n) une suite croissante. Supposons qu'il existe M majorant $\{u_n \mid n \in \mathbb{N}\}$. Soit s la borne supérieure de cet ensemble. Par définition, il existe une suite (v_n) qui converge vers s . Soit $\varepsilon > 0$, on trouve par définition de la convergence de (v_n) vers s un élément v_p tel que $|v_p - s| < \varepsilon$. Comme s est majorant, on sait que cette inégalité est exactement $s - v_p < \varepsilon$, et puisque v_p appartient aux valeurs prises par (u_n) , on trouve k tel que $v_p = u_k$. Pour $p > k$, on a $0 \leq s - u_p \leq s - u_k < \varepsilon$ par croissance de (u_n) et par le fait que s majore les valeurs prises par (u_n) . On en déduit que $|s - u_p| < \varepsilon$, donc $\lim u_n = s$.

Supposons que (u_n) ne soit pas majorée. Cela signifie que pour tout $M \in \mathbb{R}$, on peut trouver $n \in \mathbb{N}$ tel que $u_n > M$, mais par croissance de (u_n) cela signifie que pour tout $p > n, u_p > M$. Donc $\lim u_n = +\infty$. □

Exercice 12.17. Montrer, de façon duale, que si (u_n) est décroissante alors au choix :

- (u_n) est minorée et converge vers la borne inférieure de ses valeurs.
- (u_n) est non minorée et diverge vers $-\infty$.

Le dernier théorème peut être vu comme un cas particulier du théorème de la limite monotone.

Théorème 12.4 (Suites adjacentes). Soient (u_n) et (v_n) deux suites réelles telles que :

- (u_n) est croissante
- (v_n) est décroissante
- $\lim(u_n - v_n) = 0$

Alors (u_n) et (v_n) sont convergentes et de même limite, et en notant l cette limite on a $\forall n \in \mathbb{N}, u_n \leq l \leq v_n$.

Démonstration. Nous allons d'abord prouver que $\forall n \in \mathbb{N}, u_n \leq v_n$. Par l'absurde, supposons qu'il existe n_0 tel que $u_{n_0} > v_{n_0}$. Alors par croissance de (u_n) et décroissance de (v_n) , on en déduit que pour tout $p > n_0$, $u_p - v_p > u_{n_0} - v_{n_0}$ et $u_{n_0} - v_{n_0} > 0$, ce qui contredit $\lim(u_n - v_n) = 0$. Ainsi $\forall n \in \mathbb{N}, u_n \leq v_n$. Par limite monotone, puisque (u_n) est croissante et majorée par v_0 (car $u_n \leq v_n \leq v_0$) et puisque (v_n) est décroissante et minorée par u_0 , on en déduit que (u_n) et (v_n) sont convergentes. Si l'on nomme l et l' les limites respectivement de (u_n) et (v_n) alors on sait que $l - l' = 0$ soit $l = l'$. Donc (u_n) et (v_n) sont convergentes de même limite.

L'inégalité sur la limite découle directement du fait que l est la borne supérieure des valeurs de (u_n) et la borne inférieure des valeurs de (v_n) . \square

12.5 Fermés et segments

Dans cette sous-partie, nous étudierons la notion de fermé, ainsi que celle de segments. Pour que la notion de fermé fasse sens, voyons déjà ce qu'est un point adhérent à une partie.

Définition 12.9 (Adhérence). Soit $F \subseteq \mathbb{R}$. On dit que x est adhérent à F si pour tout voisinage $V \in \mathcal{V}_x$, $V \cap F \neq \emptyset$. Autrement dit, un point x est adhérent à une partie F s'il est infiniment proche de F . On appelle adhérence de F , et l'on note \overline{F} , l'ensemble de ses points adhérents :

$$\overline{F} = \{x \mid x \in \mathbb{R}, \forall V \in \mathcal{V}_x, V \cap F \neq \emptyset\}$$

Donnons une caractérisation séquentielle d'un point adhérent :

Proposition 12.5.1 (Caractérisation séquentielle de l'adhérence). Soit $F \subseteq \mathbb{R}$ et $x \in \mathbb{R}$. x est adhérent à F si et seulement s'il existe une suite (x_n) d'éléments de F telle que $\lim x_n = x$.

Démonstration. Supposons que x est adhérent à F . On définit le voisinage $V_n = B\left(x, \frac{1}{n}\right) \in \mathcal{V}_x$ et puisque x est adhérent à F , on trouve $x_n \in V_n \cap F$. Montrons que $\lim x_n = x$. Soit $\varepsilon > 0$, comme $\lim \frac{1}{n} = 0$ on trouve n_0 tel que pour tout $p > n_0$, $\frac{1}{n} < \varepsilon$. Alors comme $x_n \in V_n$, par définition, $|x_n - x| < \frac{1}{n}$ donc pour $p > n_0$, on en déduit que $|x_p - x| < \varepsilon$, donc $\lim x_n = x$ et chaque x_n est un élément de F .

Réciproquement, s'il existe une suite (x_n) d'éléments de F telle que $\lim x_n = x$, alors soit un voisinage $V \in \mathcal{V}_x$. Par caractérisation des voisinages, on peut trouver un réel $\varepsilon > 0$ tel que pour tout $y \in \mathbb{R}$, $|y - x| < \varepsilon \implies y \in V$. Or puisque $\lim x_n = x$, on peut trouver un certain rang n_0 à partir duquel $|x_p - x| < \varepsilon$, donc en prenant ce rang n_0 , on en déduit que $x_{n_0} \in V$, or $x_{n_0} \in F$ donc $F \cap V \neq \emptyset$. On en déduit que x est adhérent à F . \square

Remarque. Comme $\lim x = x$, on en déduit que $F \subseteq \overline{F}$.

Définissons alors les fermés de \mathbb{R} :

Définition 12.10 (Fermé). On dit qu'une partie $F \subseteq \mathbb{R}$ est fermée si $\overline{F} = F$, c'est-à-dire si tous les points de F lui sont adhérents.

Remarque. La définition se réduit donc simplement à $\overline{F} \subseteq F$.

Une autre façon de voir un fermé est un ensemble stable par limite.

Proposition 12.5.2. *Une partie $F \subseteq \mathbb{R}$ est fermée si et seulement si pour toute suite $(x_n) \in F^{\mathbb{N}}$ convergente, $\lim x_n \in F$.*

Démonstration. Supposons que F est fermée. Soit (x_n) une suite à valeurs dans F , convergente. Par définition, il existe une suite à valeurs dans F dont $\lim x_n$ est la limite, donc $x \in \overline{F} = F$, donc $\lim x_n \in F$.

Réciproquement, si toute suite à valeurs dans F et convergente converge dans F , alors soit $x \in \overline{F}$. Par caractérisation séquentielle, on trouve x_n à valeurs dans F telle que $\lim x_n = x$, et par hypothèse $\lim x_n \in F$, donc $x \in F$. Donc $\overline{F} \subseteq F$, donc F est fermée. \square

Enfin, nous allons donner une dernière caractérisation, reliant les fermés et les ouverts.

Proposition 12.5.3. *Soit $U \subseteq \mathbb{R}$. Alors U est ouvert si et seulement si $\mathbb{R} \setminus U$ est fermé.*

Démonstration. Supposons que U est ouvert et que $\mathbb{R} \setminus U$ n'est pas fermé. Ceci signifie qu'il existe une suite (x_n) donc aucun x_n n'est dans U telle que $\lim x_n \in U$. Cependant, comme U est ouvert, on trouve $\varepsilon > 0$ tel que $B(\lim x_n, \varepsilon) \subseteq U$, qui est donc un voisinage de $\lim x_n$, et de plus une partie de U donc il n'existe aucune $x_n \in B(\lim x_n, \varepsilon)$, ce qui est contradictoire : on a trouvé un voisinage de la limite dans lequel aucun terme n'est, donc ça n'est pas la limite de la suite. Par l'absurde, on en déduit donc que $\mathbb{R} \setminus U$ est fermé.

Réciproquement, si $F := \mathbb{R} \setminus U$ est fermé, supposons par l'absurde que U n'est pas ouvert. Cela signifie qu'il existe un élément $x \in U$ tel que pour tout $\varepsilon > 0$, il existe un élément $x_\varepsilon \notin U$ tel que $x_\varepsilon \in B(x, \varepsilon)$. Mais si $x_\varepsilon \notin U$, cela signifie que $x_\varepsilon \in F$, donc pour tout voisinage V de x , $V \cap F \neq \emptyset$, donc $x \in \overline{F} = F$, donc $x \in F \cap U = \emptyset$, ce qui est une contradiction. Par l'absurde, on en déduit que U est ouvert. \square

Une famille de fermés est particulièrement intéressante pour son bon comportement : ce sont les segments.

Définition 12.11 (Segment). Un segment $I \subseteq \mathbb{R}$ est une partie de la forme $[a, b]$ avec $a, b \in \mathbb{R}$.

Exercice 12.18. Montrer qu'un segment est fermé.

Les segments sont un cas particulier de la notion de compact, très importante en topologie mais que nous n'aborderons pas dans ce document. Cependant, une propriété essentielle des compacts fonctionne très bien dans le cas des segments, et constitue le théorème dit des segments emboîtés.

Théorème 12.5. *Soit $([a_n, b_n])_{n \in \mathbb{N}}$ une suite de segments emboîtés non vides, c'est-à-dire telle que $[a_{n+1}, b_{n+1}] \subseteq [a_n, b_n]$, et telle que la suite des longueurs des segments tend vers 0. Alors*

$$\bigcap_{n \in \mathbb{N}} [a_n, b_n] = \{x\}$$

pour un certain $x \in [a_0, b_0]$.

Démonstration. Le fait que les segments sont emboîtés signifie que (a_n) est une suite croissante et (b_n) une suite décroissante, et que les segments soient non vides signifie que $\forall n \in \mathbb{N}, a_n < b_n$. De plus, comme la longueur des segments tend vers 0, on a $\lim a_n - b_n = 0$. Ainsi (a_n) et (b_n) forment des suites adjacentes : notons x leur limite commune. Comme pour tout n , $a_n \leq x \leq b_n$, on en déduit que $x \in \bigcap_{n \in \mathbb{N}} [a_n, b_n]$. S'il existe y dans cette intersection, alors y est la limite de (a_n) et (b_n) , et par unicité de la limite on en déduit que $x = y$. Donc $\bigcap_{n \in \mathbb{N}} [a_n, b_n] = \{x\}$. \square

Pour clore cette partie, donnons un dernier résultat, essentiel pour comprendre la topologie de \mathbb{R} , appelé théorème de Bolzano-Weierstrass.

Théorème 12.6 (Bolzano-Weierstrass). *Soit (u_n) une suite de réels bornées, c'est-à-dire qu'il existe un segment $[a, b]$ tel que $\{u_n \mid n \in \mathbb{N}\} \subseteq [a, b]$. Alors il existe une sous-suite convergente, i.e. il existe une fonction $\varphi : \mathbb{N} \rightarrow \mathbb{N}$ strictement croissante telle que $(u_{\varphi(n)})_{n \in \mathbb{N}}$ converge.*

Démonstration. Nous allons construire une suite de segment emboîtés et φ par récurrence, de telle sorte que pour chaque segment de la suite, il existe une infinité de valeurs de (u_n) comprises dans ce segment :

- Posons $[a_0, b_0] = [a, b]$ et $\varphi(0) = 0$
- Si l'on a défini $[a_i, b_i]$ et $(\varphi(n))_{n \leq i}$ pour un rang i donné, alors on considère $c = \frac{a_i + b_i}{2}$. D'après le principe des tiroirs infini, il existe une infinité de valeurs de (u_n) dans $[a_i, c]$ ou dans $[c, b_i]$. Dans le premier cas, on définit $[a_{i+1}, b_{i+1}] = [a_i, c]$ et dans le deuxième cas on définit $[a_{i+1}, b_{i+1}] = [c, b_i]$. De plus, on définit $\varphi(i+1)$ comme le plus petit indice supérieur à $\varphi(i)$ qui appartient à $[a_{i+1}, b_{i+1}]$.

Avec notre construction, on sait que $u_{\varphi(n)} \in [a_n, b_n]$, mais avec le théorème des segments emboîtés, on déduit que a_n et b_n convergent vers l'unique point d'intersection de la suite de segments emboîtés (la longueur des segments diminue de moitié à chaque fois, donc tend vers 0), et par encadrement on en déduit que $(u_{\varphi(n)})$ converge vers la même limite.

Donc on a extrait de (u_n) une suite $(u_{\varphi(n)})$ convergente. □

12.5.1 Sur les suites extraites

Nous allons ici détailler deux propriétés importantes des suites extraites. Tout d'abord, donnons-en une définition.

Définition 12.12 (Suite extraite). *Soit (u_n) une suite extraite. On appelle suite extraite de (u_n) une suite de la forme $(u_{\varphi(n)})$ où $\varphi : \mathbb{N} \rightarrow \mathbb{N}$ est une fonction strictement croissante.*

Proposition 12.5.4. *Une fonction $\varphi : \mathbb{N} \rightarrow \mathbb{N}$ strictement croissante est telle que $\forall n \in \mathbb{N}, \varphi(n) \geq n$.*

Démonstration. Prouvons ce résultat par récurrence :

- Par définition, $\varphi(0) \geq 0$.
- Si $\varphi(n) \geq n$, alors par croissance stricte de φ , on a $\varphi(n+1) > \varphi(n)$, ce qui signifie $\varphi(n+1) \geq \varphi(n) + 1$ par définition de $>$ sur \mathbb{N} . En utilisant l'hypothèse de récurrence et par transitivité de \geq , on en déduit que $\varphi(n+1) \geq n+1$.

D'où par récurrence que $\forall n \in \mathbb{N}, \varphi(n) \geq n$. □

De plus, le point essentiel dans l'utilité des suites extraites sera le suivant :

Proposition 12.5.5 (Stabilité de la limite). *Soit (u_n) une suite convergente et $(u_{\varphi(n)})$ une suite extraite de (u_n) . Alors $\lim u_n = \lim u_{\varphi(n)}$.*

Démonstration. Notons $b = \lim u_n$. Soit $V \in \mathcal{V}_b$. Par définitions de la limite de (u_n) on trouve n_0 tel que $\forall p > n_0, u_p \in V$. Alors pour $p > \varphi(n_0)$, comme $\varphi(p) \geq p$ et $\varphi(n_0) \geq n_0$, on en déduit que $u_{\varphi(p)} \in V$. Donc on a trouvé $n_1 = \varphi(n_0)$ tel que pour tout $p > n_1, u_{\varphi(p)} \in V$. D'où $\lim u_{\varphi(n)} = \lim u_n$. □

Un point de vue topologique des suites extraites est de considérer l'ensemble $\{u_n \mid n \in \mathbb{N}, p\}$: il existe une suite extraite φ de (u_n) qui converge vers $l \in \mathbb{R}$ si et seulement si l est dans l'adhérence de chaque ensemble précédemment décrit, pour $p \in \mathbb{N}$. Nous n'allons pas montrer ce résultat mais nous intéresser seulement à un cas plus restreint :

Proposition 12.5.6. *Soit (u_n) une suite bornée. Alors (u_n) converge si et seulement s'il existe $l \in \mathbb{R}$ tel que toute suite extraite convergente de (u_n) converge vers l .*

Démonstration. Si (u_n) converge, alors sa limite est la limite de ses suites extraites par le résultat précédent.

Si (u_n) diverge, alors par définition :

$$\forall l \in \mathbb{R}, \exists \varepsilon_l > 0, \forall n \in \mathbb{N}, \exists p > n, |u_p - l| \geq \varepsilon_l$$

Supposons que toute suite extraite converge vers une même limite l . Grâce au théorème de Bolzano-Weierstrass, il existe au moins une suite extraite convergente, que nous nommerons $(u_{\varphi(n)})$. Par définition de la divergence de (u_n) , en prenant l , on trouve $\varepsilon_l > 0$ et une fonction $\psi : \mathbb{N} \rightarrow \mathbb{N}$ définie par récurrence par $\psi(0)$ le plus petit $p > 0$ tel que $|u_p - l| \geq \varepsilon_l$ et par $\psi(n+1)$ le plus petit $p > \psi(n)$ tel que $|u_p - l| \geq \varepsilon_l$. Cette fonction est strictement croissante par construction, donc $\psi \circ \varphi : \mathbb{N} \rightarrow \mathbb{N}$ est une fonction strictement croissante : on en déduit que la suite $(u_{\psi \circ \varphi(n)})$ est une suite extraite de (u_n) , donc elle converge vers l . On trouve donc n_0 tel que $\forall p > n_0, |u_{\psi \circ \varphi(p)} - l| < \varepsilon_l$ or par définition de ψ , on sait que $|u_{\psi \circ \varphi(p)} - l| \geq \varepsilon_l$ pour tout p : c'est donc une contradiction. Ainsi toute suite extraite convergente ne peut avoir la même limite. Donc par contraposée, si toute suite extraite convergente a la même limite, (u_n) converge. □

13 Fonctions continues et limites

Cette section s'intéresse à la généralisation de la notion de limite pour une fonction, et la notion de continuité qui en est le prolongement direct. Enfin, nous allons donner les résultats fondamentaux sur les fonctions continues : le théorème des valeurs intermédiaires et le théorème des bornes atteintes.

13.1 Limite d'une fonction

Nous connaissons les limites de suites, mais la notion de limite s'étend à celle de fonctions. Cependant, un fait important pour les fonctions est que l'on peut étudier des limites en différents points (pas qu'en ∞). D'où la définition suivante de limite :

Définition 13.1 (Limite d'une fonction en un point). Soit $f : F \rightarrow \mathbb{R}, F \subseteq \mathbb{R}, a \in \overline{F}$ et $b \in \mathbb{R} \cup \{-\infty, +\infty\}$. On dit que f tend vers b en a , ce que l'on écrit $\lim_{x \rightarrow a} f(x) = b$, lorsque la propriété suivante est vérifiée :

$$\forall V \in \mathcal{V}_b, \exists V' \in \mathcal{V}_a, f(V') \subseteq V$$

Remarque. On note $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$, pour étendre notre notation d'adhérence aux limites infinies.

Dire que f tend vers b en a signifie donc qu'aussi près que l'on se place de b , il existe un voisinage assez petit de a qui est envoyé assez près de b . Nous allons reformuler cette définition avec une caractérisation par des ε dans le cas d'une limite finie.

Proposition 13.1.1. Soit $f : F \rightarrow \mathbb{R}, F \subseteq \mathbb{R}, a \in \overline{F}$ et $b \in \mathbb{R}$. Alors :

- si $a \in \mathbb{R}$, alors $\lim_{x \rightarrow a} f(x) = b$ est équivalent à la proposition suivante :

$$\forall \varepsilon > 0, \exists \delta > 0, \forall x \in \mathbb{R}, |x - a| < \delta \implies |f(x) - b| < \varepsilon$$

- si $a = +\infty$, alors $\lim_{x \rightarrow a} f(x) = b$ est équivalent à la proposition suivante :

$$\forall \varepsilon > 0, \exists M \in \mathbb{R}, \forall x \in \mathbb{R}, x > M \implies |f(x) - b| < \varepsilon$$

- si $a = -\infty$, alors $\lim_{x \rightarrow a} f(x) = b$ est équivalent à la proposition suivante :

$$\forall \varepsilon > 0, \exists m \in \mathbb{R}, \forall x \in \mathbb{R}, x < m \implies |f(x) - b| < \varepsilon$$

Démonstration. Traitons chaque cas :

- Soit $a \in \mathbb{R}$, et supposons que $\lim_{x \rightarrow a} f(x) = b$. Soit $\varepsilon > 0$, comme $B(b, \varepsilon) \in \mathcal{V}_b$ on trouve par définition de la limite un voisinage $V \in \mathcal{V}_a$ tel que $f(V) \subseteq B(b, \varepsilon)$. Or l'inclusion signifie $\forall x \in \mathbb{R}, x \in V \implies f(x) \in B(b, \varepsilon)$ soit $\forall x \in \mathbb{R}, x \in V \implies |f(x) - b| < \varepsilon$. Par propriété du voisinage V , on trouve $\delta > 0$ tel que $\forall x \in \mathbb{R}, |x - a| < \delta \implies x \in V$ d'où, par transitivité de l'implication, que $\boxed{\forall x \in \mathbb{R}, |x - a| < \delta \implies |f(x) - b| < \varepsilon.}$

Réciproquement, soit $V \in \mathcal{V}_b$ un voisinage de b . Par propriété de voisinage, on trouve $\varepsilon > 0$ tel que $\forall x \in \mathbb{R}, |x - b| < \varepsilon \implies x \in V$. Soit alors $\delta > 0$ déduit de l'hypothèse appliquée à ε . On pose alors $V' = B(a, \delta)$, montrons que $f(V') \subseteq V$: soit $x \in V'$, par construction de δ on sait que $|x - a| < \delta \implies |f(x) - b| < \varepsilon$, or la prémisse de cette implication est vraie car $x \in B(a, \delta)$, donc $|f(x) - b| < \varepsilon$, ce qui d'après la définition de ε appliquée à $f(x)$ signifie que $f(x) \in V$. Donc $\boxed{f(V') \subseteq V.}$

- Supposons que $\lim_{x \rightarrow +\infty} f(x) = b$. Soit $\varepsilon > 0$, $B(b, \varepsilon)$ est un voisinage de b donc on trouve un voisinage V' de $+\infty$ tel que $f(V') \subseteq B(b, \varepsilon)$, i.e. tel que $\forall x \in V', |f(x) - b| < \varepsilon$, et par propriété d'un voisinage de $+\infty$ on trouve M tel que $\forall x \in \mathbb{R}, x > M \implies x \in V'$, d'où $\boxed{\forall x \in \mathbb{R}, x > M \implies |f(x) - b| < \varepsilon.}$

Réciproquement, soit $V \in \mathcal{V}_b$, on trouve ε tel que $B(b, \varepsilon) \subseteq V$. Par hypothèse appliquée à ε , on trouve $M \in \mathbb{R}$ tel que $x > M \implies f(x) \in B(b, \varepsilon)$, ce qui équivaut à $f([M; +\infty[) \subseteq B(b, \varepsilon)$, or $[M; +\infty[\in \mathcal{V}_{+\infty}$, et par transitivité de l'inclusion, $\boxed{\text{on a trouvé } V' \text{ tel que } f(V') \subseteq V.}$

- Le cas $a = -\infty$ se traite de la même façon que pour $a = +\infty$ en inversant le sens des inégalités. \square

Exercice 13.1. Montrer que si $b = +\infty$ les énoncés équivalents à $\lim_{x \rightarrow a} f(x) = b$ sont les mêmes en remplaçant $\forall \varepsilon > 0$ par $\forall K \in \mathbb{R}$ et $|f(x) - b| < \varepsilon$ par $f(x) > K$. De même pour $b = -\infty$ en remplaçant par $\forall k \in \mathbb{R}$ et $f(x) < k$.

Une caractérisation importante est la caractérisation séquentielle : l'idée est que « tendre vers a » peut s'exprimer plus simplement par le fait de considérer toutes les suites de limite a .

Proposition 13.1.2 (Caractérisation séquentielle). *Soit $f : F \rightarrow \mathbb{R}, F \subseteq \mathbb{R}, a \in \overline{F}$ et $b \in \overline{\mathbb{R}}$. Alors $\lim_{x \rightarrow a} f(x) = b$ si et seulement si pour toute suite $(u_n)_{n \in \mathbb{N}} \in F^{\mathbb{N}}$ telle que $\lim u_n = a$ on a $\lim f(u_n) = b$.*

Remarque. Nous n'écrivons pas $n \rightarrow \infty$ pour accentuer la différence entre une limite de fonction et une limite de suite.

Démonstration. Supposons que $\lim_{x \rightarrow a} f(x) = b$, et soit (u_n) une suite réelle de limite a . Montrons que $\lim f(u_n) = b$. Pour cela, soit $V \in \mathcal{V}_b$, par définition de la limite de f en a on trouve $V' \in \mathcal{V}_a$ tel que $f(V') \subseteq V$. Par définition de $\lim u_n = a$, on trouve un rang n_0 tel que pour tout $p > n_0, u_p \in V'$. En utilisant l'inclusion plus tôt, on en déduit $\forall p > n_0, f(u_p) \in V$, donc $\lim f(u_n) = b$.

Supposons maintenant que pour toute suite (u_n) telle que $\lim u_n = a$, on ait $\lim f(u_n) = b$, et montrons que $\lim_{x \rightarrow a} f(x) = b$. Soit $V \in \mathcal{V}_b$, procédons par l'absurde : supposons que pour tout voisinage $V' \in \mathcal{V}_a$, on ait $f(V') \not\subseteq V$ et cherchons une contradiction. L'énoncé précédent est équivalent à dire que pour tout voisinage $V' \in \mathcal{V}_a$, il existe un élément $x_{V'}$ tel que $f(x_{V'}) \notin V$. Soit $(V_n)_{n \in \mathbb{N}}$ une base séquentielle de voisinages de a , qui existe car nous en avons exhibé une pour chaque élément $a \in \mathbb{R} \cup \{-\infty, +\infty\}$ dans un exercice précédent. Alors on pose $x_n = x_{V_n}$, c'est-à-dire l'élément de V_n tel que $f(x_n) \notin V$. Si l'on prend un voisinage W de a , alors on trouve $n \in \mathbb{N}$ tel que pour tout $p > n, V_p \subseteq W$, donc pour tout $p > n, x_p \in W$. Ainsi $\lim x_n = a$. On en déduit que $\lim f(x_n) = b$ par notre hypothèse, ce qui signifie qu'il existe $n_0 \in \mathbb{N}$ tel que pour tout $p > n_0, f(x_p) \in V$, pourtant par construction de $x_n, f(x_n) \notin V$: on a bien une contradiction. Ainsi $\lim_{x \rightarrow a} f(x) = b$. \square

Proposition 13.1.3 (Unicité de la limite). *Soit $f : F \rightarrow \mathbb{R}, F \subseteq \mathbb{R}, a \in \overline{F}$ et $b, c \in \overline{\mathbb{R}}$, si $\lim_{x \rightarrow a} f(x) = b$ et $\lim_{x \rightarrow a} f(x) = c$ alors $b = c$.*

Démonstration. On va utiliser la caractérisation séquentielle : nos deux hypothèses disent que pour toute suite (u_n) telle que $\lim u_n = a$, alors $\lim f(u_n) = b$ et $\lim f(u_n) = c$, donc par unicité de la limite d'une suite on en déduit que $b = c$. \square

On pourra donc parler, quand elle existe, de **la** limite de f en a .

Exercice 13.2. En utilisant la caractérisation séquentielle, démontrer le théorème d'encadrement pour les fonctions : soient $f, g, h : F \rightarrow \mathbb{R}, F \subseteq \mathbb{R}, a \in \overline{F}, b \in \mathbb{R}$ tels que $\forall x \in F, f(x) \leq g(x) \leq h(x)$ et $\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} h(x) = b$, alors $\lim_{x \rightarrow a} g(x) = b$.

Montrer que ce théorème est encore vrai si l'on considère que l'encadrement n'est vrai qu'au voisinage de a , c'est-à-dire en affaiblissant l'hypothèse d'encadrement à

$$\exists V \in \mathcal{V}_a, \forall x \in V, f(x) \leq g(x) \leq h(x)$$

Exercice 13.3. Grâce à la caractérisation séquentielle, montrer que le tableau des limites (Figure 22) est encore valide pour des limites de fonction en un point a donné. Plus généralement, montrer que les théorèmes de limites de suites (encadrement, majoration, valeurs usuelles) sont encore vrais pour les fonctions.

13.2 Limite à gauche, à droite

Une notion pouvant avoir son importance est celle de limite à gauche (respectivement à droite), qui correspond à étudier uniquement d'un côté du point de limite.

Définition 13.2 (Limite à gauche, à droite). Soit $f : F \rightarrow \mathbb{R}$, $F \subseteq \mathbb{R}$ et $a \in \overline{F}$. On dit que $f(x)$ tend vers b à gauche (respectivement à droite) en a lorsque

$$\forall V \in \mathcal{V}_b, \exists V' \in \mathcal{V}_a, f(V' \cap [a; +\infty[) \subseteq V$$

(respectivement $V' \cap]-\infty, a]$)

On note alors $\lim_{x \rightarrow a^+} f(x) = b$ (respectivement $\lim_{x \rightarrow a^-} f(x) = b$).

Remarque. La notion de limite à gauche ou à droite n'a de sens que pour $a \in \mathbb{R}$.

On peut relier la limite et les limites à gauche et à droite.

Proposition 13.2.1. Soit $f : F \rightarrow \mathbb{R}$, $a \in F$, alors $\lim_{x \rightarrow a} f(x) = b$ si et seulement si $\lim_{x \rightarrow a^+} f(x) = b$ et $\lim_{x \rightarrow a^-} f(x) = b$. Autrement dit la limite de f en a existe si et seulement si ses limites à gauche et à droite existent et qu'elles coïncident.

Démonstration. Si $\lim_{x \rightarrow a} f(x) = b$ alors soit $V \in \mathcal{V}_b$, on trouve par hypothèse $V' \in \mathcal{V}_a$ tel que $f(V') \subseteq V$, donc $f(V' \cap [a; +\infty[) \subseteq V$ et $f(V' \cap]-\infty, a]) \subseteq V$, donc $\lim_{x \rightarrow a^+} f(x) = \lim_{x \rightarrow a^-} f(x) = b$.

Réciproquement, supposons que les limites à gauche et à droite existent et coïncident. Soit $V \in \mathcal{V}_a$, on trouve par la limite à gauche $V' \in \mathcal{V}_a$ tel que $f(V' \cap]-\infty, a]) \subseteq V$ et par la limite à droite $V'' \in \mathcal{V}_a$ tel que $f(V'' \cap [a; +\infty[) \subseteq V$. Soit alors $W = V' \cap V'' \in \mathcal{V}_a$, montrons que $f(W) \subseteq V$. Soit $x \in W$, si $x \leq a$ alors $x \in W \cap]-\infty, a]$, et comme $W \subseteq V'$, $x \in V' \cap]-\infty, a]$, donc $f(x) \in V$. Si $x \geq a$ alors $x \in W \cap [a; +\infty[$ donc $x \in V'' \cap [a; +\infty[$, d'où $f(x) \in V$. Dans tous les cas, $f(x) \in V$, donc $f(W) \subseteq V$. Ainsi $\lim_{x \rightarrow a} f(x) = b$. \square

13.3 Continuité d'une fonction

Les fonctions continues sont fondamentales en analyse, et nous allons maintenant pouvoir les étudier. L'idée intuitive d'une fonction continue est souvent présentée par « une fonction dont on peut tracer le graphe sans lever le crayon », mais son formalisme mathématique nécessite de parler de limites. En fait, on peut considérer la continuité comme la préservation de voisinages, mais nous allons ici nous restreindre à un point de vue plus terre à terre : une fonction continue est une fonction dont les valeurs à des points proches sont proches. Nous allons d'abord présenter la notion de continuité locale, puis de continuité sur un ensemble.

Définition 13.3 (Continuité en un point). Soit $f : F \rightarrow \mathbb{R}$, $F \subseteq \mathbb{R}$ et $a \in F$. On dit que f est continue en a lorsque

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

Proposition 13.3.1. Une fonction f est continue en a si et seulement si la fonction f admet une limite (réelle) en a .

Démonstration. L'équivalence signifie que si la limite de f en a existe, alors elle vaut $f(a)$. En effet, supposons que $\lim_{x \rightarrow a} f(x) = b$ où $b \in \mathbb{R}$. Par caractérisation séquentielle de la limite appliquée à la suite constante (a) , on en déduit que $\lim f(a) = b$, mais comme $(f(a))$ est une suite constante, on peut prouver que $\lim f(a) = f(a)$. Par unicité de la limite, on en déduit que $b = f(a)$. \square

Définition 13.4 (Continuité sur un ensemble). Soit $f : F \rightarrow \mathbb{R}$, $F \subseteq \mathbb{R}$. On dit que f est continue si f est continue en x pour tout $x \in F$. On dit que f est continue sur $I \subseteq F$ si $f|_I$ est continue.

Proposition 13.3.2 (Caractérisation de la continuité). Une fonction $f : F \rightarrow \mathbb{R}$ est continue si et seulement si

$$\forall \varepsilon > 0, \forall x \in F, \exists \delta > 0, \forall y \in F, |y - x| < \delta \implies |f(y) - f(x)| < \varepsilon$$

Démonstration. La proposition est équivalente à si l'on avait écrit au début $\forall x \in F, \forall \varepsilon > 0$, et f est continue si et seulement si pour tout $x \in F$, $\lim_{y \rightarrow x} f(y) = f(x)$, ce qui se réécrit (comme y et $f(x)$ sont finis) $\forall \varepsilon > 0, \exists \delta > 0, \forall y \in F, |x - y| < \delta \implies |f(x) - f(y)| < \varepsilon$, d'où l'équivalence. \square

Donnons plusieurs résultats de stabilité de la classe des fonctions continues, c'est-à-dire plusieurs façons de justifier qu'une fonction est continue à partir de la continuité de ses composantes.

Proposition 13.3.3 (Stabilité par combinaisons linéaires). *Soient $f : F \rightarrow \mathbb{R}$ et $g : F \rightarrow \mathbb{R}$ deux fonctions continues, et $k, k' \in \mathbb{R}$. Alors*

$$\begin{aligned} kf + k'g : F &\longrightarrow \mathbb{R} \\ x &\longmapsto k \times f(x) + k' \times g(x) \end{aligned}$$

est continue.

Démonstration. Soit $a \in F$, alors par les propriétés précédentes sur les limites finies

$$\lim_{x \rightarrow a} (k \times f(x) + k' \times g(x)) = k(\lim_{x \rightarrow a} f(x)) + k'(\lim_{x \rightarrow a} g(x))$$

donc $kf + k'g$ a une limite finie en a . \square

Proposition 13.3.4 (Stabilité par quotient, par produit). *Soient $f, g : F \rightarrow \mathbb{R}$ deux fonctions continues, et g une fonction ne s'annulant pas. Alors les fonctions*

$$\begin{aligned} \frac{f}{g} : F &\longrightarrow \mathbb{R} \\ x &\longmapsto \frac{f(x)}{g(x)} \end{aligned}$$

et

$$\begin{aligned} f \times g : F &\longrightarrow \mathbb{R} \\ x &\longmapsto f(x) \times g(x) \end{aligned}$$

est continue (le produit est continu même lorsque g s'annule).

Démonstration. Là encore, le résultat découle directement des propriétés sur les limites finies. \square

Exercice 13.4. Soient $f : F \rightarrow \mathbb{R}$ et $g : G \rightarrow \mathbb{R}$ deux fonctions avec $F, G \subseteq \mathbb{R}$ et $f(F) \subseteq G$. Montrer que $g \circ f : F \rightarrow \mathbb{R}$ est continue.

Remarque. Cet exercice nous permet de déduire de la stabilité par produit celle par quotient puisqu'on peut composer par la fonction inverse.

13.4 Continuité uniforme

Donnons tout de suite la définition de continuité uniforme, qui ne servira pas maintenant mais qui est un renforcement du fait d'être continu, consistant en une inversion de quantificateurs.

Définition 13.5 (Continuité uniforme). Soit $f : F \rightarrow \mathbb{R}, F \subseteq \mathbb{R}$. On dit que f est uniformément continue si la propriété suivante est vérifiée :

$$\forall \varepsilon > 0, \exists \delta > 0, \forall x \in F, \forall y \in F, |x - y| < \delta \implies |f(x) - f(y)| < \varepsilon$$

Proposition 13.4.1. *Si f est uniformément continue, alors f est continue.*

Démonstration. Supposons f uniformément continue. Soit $\varepsilon > 0$, alors on trouve $\delta > 0$ vérifiant la propriété de la continuité uniforme. Soit alors $x \in F$: en prenant ce δ , on a bien

$$\forall y \in F, |x - y| < \delta \implies |f(x) - f(y)| < \varepsilon$$

donc f est continue. \square

Profitons-en pour donner le théorème de Heine.

Théorème 13.1 (Heine). *Soit $f : [a, b] \rightarrow \mathbb{R}$, $a, b \in \mathbb{R}$ une fonction continue. Alors f est uniformément continue.*

Démonstration. Procédons par contraposée : supposons f non uniformément continue et déduisons-en que f n'est pas continue. Si f n'est pas uniformément continue, alors on trouve ε tel que pour tout $\delta > 0$, on trouve x_δ, y_δ tels que $|x_\delta - y_\delta| < \delta \wedge |f(x_\delta) - f(y_\delta)| \geq \varepsilon$. En prenant la suite de valeurs $(1/n)$ pour δ , on en déduit deux suites (x_n) et (y_n) telles que $|x_n - y_n| < \frac{1}{n}$ et $|f(x_n) - f(y_n)| \geq \varepsilon$. De plus, comme $x_n \in [a, b]$, la suite (x_n) est bornée : par le théorème de Bolzano-Weierstrass, on en déduit qu'il existe une fonction φ strictement croissante telle que $(x_{\varphi(n)})$ converge. Dans ce cas, on remarque que comme $|x_{\varphi(n)} - y_{\varphi(n)}| < \frac{1}{\varphi(n)}$ et que $\lim \varphi(n) = +\infty$, on a $\lim |x_{\varphi(n)} - y_{\varphi(n)}| = 0$, dont on déduit directement que $\lim x_{\varphi(n)} = \lim y_{\varphi(n)}$ (et donc que $(y_{\varphi(n)})$ converge). Si f était continue, alors on aurait $\lim f(x_{\varphi(n)}) = \lim f(y_{\varphi(n)})$ grâce à l'égalité précédente, ce qui signifie que $\lim |f(x_{\varphi(n)}) - f(y_{\varphi(n)})| = 0$, ce qui est contredit par le fait que pour tout n , $|f(x_{\varphi(n)}) - f(y_{\varphi(n)})| \geq \varepsilon > 0$. Donc f n'est pas continue.

Donc, par contraposée, si f est continue alors elle est uniformément continue. □

13.5 Théorèmes fondamentaux sur les fonctions continues

Parmi les théorèmes essentiels sur les fonctions continues d'une variable réelle, le plus connu est celui des valeurs intermédiaires, que nous allons énoncer et démontrer maintenant.

Théorème 13.2 (Valeurs intermédiaires). *Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction continue et $y \in \mathbb{R}$ compris entre $f(a)$ et $f(b)$. Alors il existe $c \in [a, b]$ tel que $f(c) = y$.*

Démonstration. Pour démontrer ce résultat, nous allons construire une suite de fermés emboîtés $([a_n, b_n])$ qui convergera vers un singleton c tel que $f(c) = y$. Dans un premier temps, on pose $[a_0, b_0] = [a, b]$, et l'on construit ensuite les segments par récurrence, de telle sorte que y est compris entre $f(a_0)$ et $f(b_0)$.

Supposons construit le segment $[a_n, b_n]$. Soit alors $c_n = \frac{a_n + b_n}{2}$ le milieu du segment. Il y a alors deux possibilités : soit $y \in [f(a_n), f(c_n)]$ soit $y \in [f(c_n), f(b_n)]$ (ici on considère que $[f(a_n), f(c_n)] = [f(c_n), f(a_n)]$, c'est-à-dire qu'on ne considère pas l'ordre dans lequel les bornes sont données). Dans le premier cas on pose $[a_{n+1}, b_{n+1}] = [a_n, c_n]$ et dans le deuxième cas on pose $[a_{n+1}, b_{n+1}] = [c_n, b_n]$.

Comme $c_n \in [a_n, b_n]$, on en déduit que $([a_n, b_n])$ est bien une suite de fermés emboîtés. De plus, la longueur de $[a_{n+1}, b_{n+1}]$ est la moitié de celle de $[a_n, b_n]$, donc la suite des longueurs est une suite géométrique de raison $\frac{1}{2}$ et de premier terme $b_n - a_n$. On en déduit que la longueur des segments tend vers 0. Enfin, par récurrence (le résultat est vrai à l'initialisation par hypothèse et héréditaire par construction de $[a_{n+1}, b_{n+1}]$) on peut montrer que y est compris entre $f(a_n)$ et $f(b_n)$ pour tout $n \in \mathbb{N}$. Par le théorème des segments emboîtés, on en déduit qu'il existe un unique c tel que $\bigcap_{n \in \mathbb{N}} [a_n, b_n] = \{c\}$,

et $c = \lim a_n = \lim b_n$.

Comme y est encadré par $f(a_n)$ et $f(b_n)$, on en déduit que $\lim f(a_n) = \lim f(b_n) = y$ par théorème d'encadrement, d'où par continuité de f , $f(\lim a_n) = y$, d'où $f(c) = y$. □

Donnons une autre formulation équivalente du TVI (théorème des valeurs intermédiaires).

Proposition 13.5.1 (Énoncé équivalent du TVI). *Soit $f : F \rightarrow \mathbb{R}$, $F \subseteq \mathbb{R}$ une fonction continue et I un intervalle inclus dans F . Alors $f(I)$ est aussi un intervalle.*

Démonstration. Pour montrer que $f(I)$ est un intervalle, il suffit de montrer que pour $x, z \in f(I)$ et $y \in \mathbb{R}$ tel que $x < y < z$, on a $y \in f(I)$. Par définition de $f(I)$, on trouve $a, b \in I$ tels que $f(a) = x$ et $f(b) = z$. Supposons sans perte de généralité que $a < b$ (sinon on inverse simplement les deux). Alors $f|_{[a, b]}$ est une fonction continue et y est compris entre $f(a)$ et $f(b)$, donc on trouve $c \in [a, b]$ tel que $f(c) = y$. Or I est un intervalle, donc $c \in I$, donc $y \in f(I)$. Ainsi $f(I)$ est un intervalle. □

Un autre résultat important lié au TVI est appelé au choix « corollaire du TVI » ou « théorème de la bijection ». Nous emploierons plutôt la deuxième terminologie, pour mettre l'accent sur l'intérêt du théorème.

Théorème 13.3 (Bijection). *Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction continue et strictement monotone. Alors f établit une bijection entre $[a, b]$ et $[f(a), f(b)]$ si f est croissante, entre $[f(b), f(a)]$ si f est décroissante.*

Démonstration. Nous ne traiterons que le cas où f est strictement croissante, l'autre cas étant analogue. Montrons successivement que f est surjective sur $[f(a), f(b)]$ puis que f est injective. Soit $y \in [f(a), f(b)]$, alors par le théorème des valeurs on trouve $x \in [a, b]$ tel que $f(x) = y$, donc f est surjective. Si l'on prend deux éléments x, y tels que $x \neq y$, alors soit $x < y$ soit $x > y$, dans les deux cas, la croissance stricte de f nous permet de déduire que $f(x) \neq f(y)$. Par contraposée cela signifie que $\forall x, y \in [a, b], f(x) = f(y) \implies x = y$, donc f est injective. Donc f est bijective sur $[f(a), f(b)]$. \square

Exercice 13.5. Montrer que l'on peut étendre le théorème de la bijection à $f : I \rightarrow \mathbb{R}$ strictement monotone continue, où I est un intervalle : f établit une bijection entre I et $f(I)$. En déduire, pour les différents cas :

- Si f est strictement croissante, que $I = \mathbb{R}$, alors f établit une bijection entre \mathbb{R} et l'intervalle $] \lim_{x \rightarrow -\infty} f(x), \lim_{x \rightarrow +\infty} f(x) [$
- Si f est strictement décroissante, que $I = \mathbb{R}$, alors f établit une bijection entre \mathbb{R} et l'intervalle $] \lim_{x \rightarrow +\infty} f(x), \lim_{x \rightarrow -\infty} f(x) [$
- Si f est strictement croissante, que $I =] - \infty, a]$, alors f établit une bijection entre $] - \infty, a]$ et l'intervalle $] \lim_{x \rightarrow -\infty} f(x), f(a) [$

Donner les autres cas pour chaque forme de I intervalle et selon la monotonie de f , et le démontrer.

Enfin, concluons sur un dernier théorème essentiel.

Théorème 13.4 (Bornes atteintes). *Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction continue. Alors $f([a, b])$ est borné et f atteint ses bornes, c'est-à-dire qu'il existe $c \in [a, b]$ et $d \in [a, b]$ tels que $f([a, b]) = [f(c), f(d)]$.*

Démonstration. Soit une telle fonction f . Tout d'abord, montrons que f est bornée. Supposons que f n'est pas bornée. Alors elle n'est pas majorée ou elle n'est pas minorée. Supposons sans perte de généralité que f n'est pas majorée. On trouve alors pour chaque $M \in \mathbb{R}$ un élément $x_M \in [a, b]$ tel que $f(x_M) > M$. En prenant $M = n$ pour chaque $n \in \mathbb{N}$, on obtient donc une suite (x_n) telle que $\lim f(x_n) = +\infty$, mais comme (x_n) est bornée, on peut en extraire une sous-suite convergente $(x_{\varphi(n)})$. On a alors $\lim f(x_{\varphi(n)})$ finie puisque f est continue sur $[a, b]$ et que $\lim x_n \in [a, b]$, ce qui est une contradiction avec le fait que $\lim f(x_n) = +\infty$. Donc f est majorée. Le même raisonnement permet de montrer que f est minorée, donc f est bornée. \square

Soit $M = \sup(f([a, b]))$. Par caractérisation séquentielle de la borne supérieure, on trouve une suite (d_n) d'éléments de $[a, b]$ telle que $\lim f(d_n) = M$. Comme (d_n) est bornée, on peut en extraire une sous-suite convergente $(d_{\varphi(n)})$, nommons $d = \lim d_{\varphi(n)}$. Par unicité de la limite, cela signifie que $M = f(d)$. Avec un même raisonnement, on trouve $c \in [a, b]$ tel que $f(c) = \inf([a, b])$. Ainsi f atteint ses bornes. \square

De plus, par la formulation équivalente du TVI, on sait que $f([a, b])$ est un intervalle, et que sa borne supérieure (respectivement inférieure) est en fait un maximum (respectivement un minimum), donc cela signifie que cet ensemble est le segment entre le minimum et le maximum, i.e. que $f([a, b]) = [f(c), f(d)]$. \square

14 Dérivabilité

Cette section va développer la théorie de la dérivation, essentielle en analyse. L'idée a été amenée par Newton et Leibnitz, mais si l'histoire des mathématiques est passionnante, elle n'est pas l'objet de ce document : nous allons donc faire une présentation résolument moderniste de la dérivation (comme nous l'avons fait pour toutes les notions jusqu'ici, d'ailleurs). Nous allons introduire les nombres dérivés, puis nous parlerons de fonction dérivée. Nous donnerons ensuite plusieurs méthodes de calcul de dérivées, puis des théorèmes essentiels sur les dérivées : théorème de Rolle, théorème et inégalité des accroissements finis, relation entre la monotonie d'une fonction et le signe de sa dérivée.

14.1 Nombre dérivé

La première étape pour définir la dérivation est de s'intéresser à la notion de taux d'accroissement. Graphiquement, le taux d'accroissement d'une courbe \mathcal{C}_f (représentative de la fonction f) entre deux points x et y est la pente de la corde entre les points d'abscisse x et d'abscisse y .

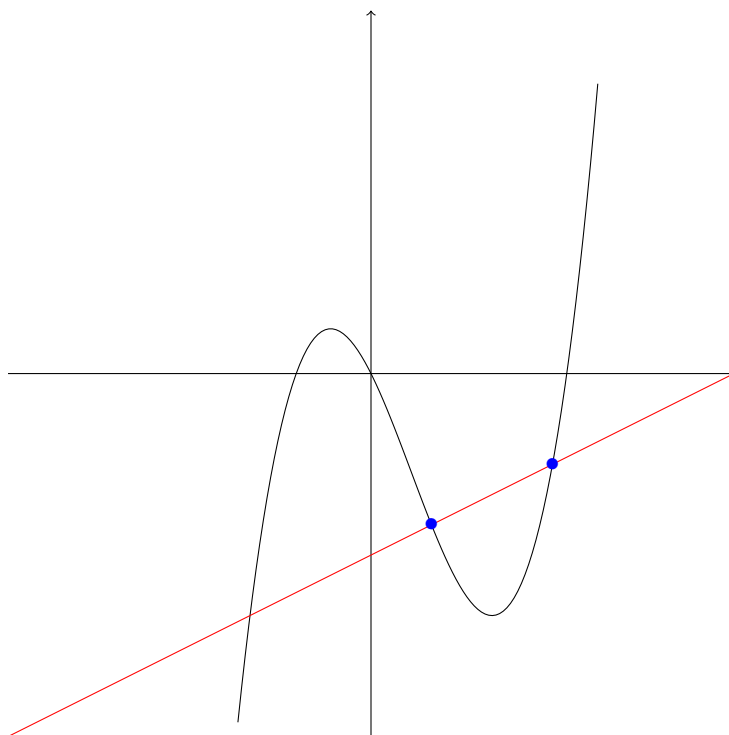


FIGURE 23 – Corde entre deux points d'une courbe

En prenant comme exemple la Figure 23, où $x = 1$ et $y = 3$, la corde est tracée en rouge (plus précisément la droite contenant la corde, la corde n'étant que le segment entre les deux points bleus). La pente correspond au coefficient directeur de la droite, et on peut donc le calculer par la formule

$$\frac{f(y) - f(x)}{y - x}$$

(remarquons que choisir de mettre x avant y dans la formule n'a pas d'importance, puisqu'échanger l'un avec l'autre donnerait deux changements de signes). Le taux d'accroissement mesure donc la position relative de deux points, en normalisant leur écart horizontal. Cependant, on remarque que la corde dessinée précédemment est totalement différente de la courbe, et nous allons essayer au contraire de rendre notre corde la plus proche possible de la courbe. Plus précisément, nous allons fixer un point x et essayer de rapprocher y de x . Ce faisant, notre corde sera très proche de la courbe autour de x , et l'on se rapprochera de ce qu'on appelle la tangente à la courbe au point x .

La Figure 24 donne une idée d'une meilleure approximation. Maintenant, on peut se demander ce qu'il se passe à *la limite* : la corde devient la tangente à la courbe au point x . C'est exactement cette donnée qui nous intéressera, car elle a l'avantage premier de ne dépendre que du point x considéré, et de

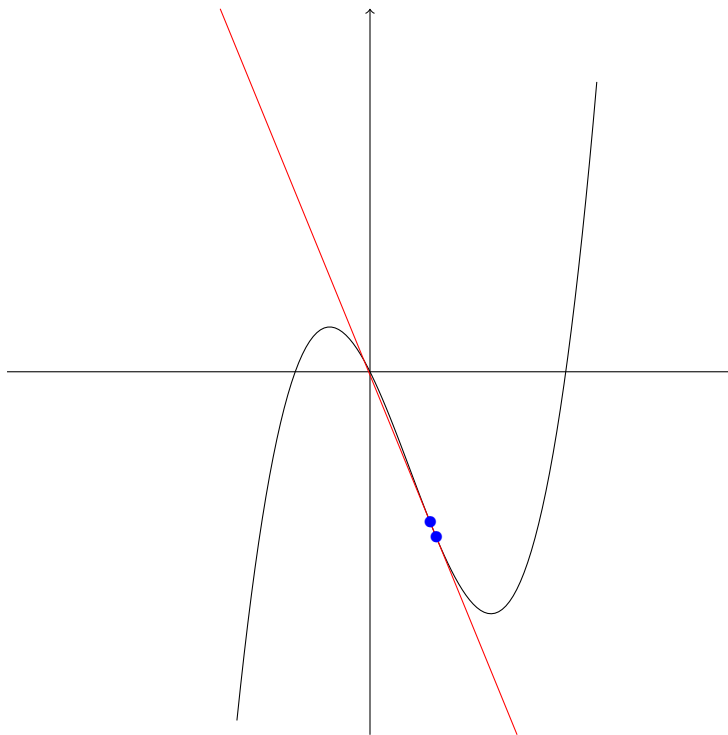


FIGURE 24 – Corde très proche de la tangente

donner une donnée purement locale. Cependant, la tangente en elle-même a assez peu d'importance, et seule sa pente compte vraiment. Celle-ci donne en effet l'information essentielle de comment évolue la courbe localement. On va donc définir le nombre dérivé d'une fonction comme la pente de sa tangente.

Définition 14.1 (Nombre dérivé). Soit $f : F \rightarrow \mathbb{R}$ une fonction, $a \in F$. On appelle nombre dérivé de f au point a , et l'on note df_a , la valeur

$$df_a = \lim_{y \rightarrow a} \frac{f(y) - f(a)}{y - a}$$

(qui n'existe pas forcément). Si df_a existe, on dit que f est dérivable en a . Une formulation équivalente du nombre dérivé est

$$df_a = \lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h}$$

par un simple changement de variable.

Remarque. La fonction $y \mapsto \frac{f(y) - f(a)}{y - a}$ est définie seulement pour $y \neq a$. Cela n'empêche pas de pouvoir considérer la limite de la fonction au point a .

Une propriété importante : cette limite indique est plus forte que la continuité au point considéré.

Proposition 14.1.1. Si df_a existe alors f est continue en a .

Proposition 14.1.2. Par définition, $(y - a)df_a = f(y) - f(a)$ or $\lim_{y \rightarrow a} y - a = 0$ donc par produit de limite et égalité, $\lim_{y \rightarrow a} f(y) - f(a) = 0$ ce qui revient à $\lim_{y \rightarrow a} f(y) = f(a)$ d'où le fait que f est continue en a .

Exercice 14.1. Soit $f : x \mapsto x^2$, calculer df_1 .

Ceci étant, donnons l'équation de la tangente à une courbe donnée par une fonction. Commençons par définir ce qu'est une tangente :

Définition 14.2 (Tangente à une courbe). Soit $f : F \rightarrow \mathbb{R}$ et $a \in F$. On dit que la droite Δ (vue comme une fonction, qui à x associe y tel que $(x, y) \in \Delta$) est tangente à f en a si $f(a) = \Delta(a)$ et si pour toute droite Δ' telle que $f(a) = \Delta'(a)$ alors il existe un voisinage V de a tel que

$$\forall x \in V, |f(x) - \Delta(x)| \leq |f(x) - \Delta'(x)|$$

Un premier résultat assez direct est qu'une tangente, pour une fonction et un point fixés, est unique.

Proposition 14.1.3. *En reprenant les notations précédente, si Δ est tangente à f en a et Δ' aussi, alors $\Delta = \Delta'$.*

Démonstration. On remarque qu'il existe un voisinage V tel que $|f(x) - \Delta(x)| \leq |f(x) - \Delta'(x)|$ et un voisinage W tel que $|f(x) - \Delta(x)| \geq |f(x) - \Delta'(x)|$, donc sur le voisinage $V \cap W$ on a $|f(x) - \Delta(x)| = |f(x) - \Delta'(x)|$ ce qui revient à $\Delta(x) = \Delta'(x)$ ou $\Delta(x) = 2f(x) - \Delta'(x)$. Dans le premier cas, on peut prendre deux points de ce voisinage et en déduire que Δ et Δ' coïncident sur deux points donc sont égales. Dans le deuxième cas, on peut construire $\Delta''(x) = \frac{\Delta(x) + \Delta'(x)}{2}$ mais alors $\Delta''(x) = f(x)$ donc, comme $\Delta'(x)$ est tangente, on en déduit que $|f(x) - \Delta'| \leq |f(x) - \Delta''(x)| = 0$ donc Δ' et f coïncident sur $V \cap W$, et on en déduit que $\Delta = \Delta'$ sur $V \cap W$, donc que $\Delta = \Delta'$. \square

La tangente est donc la droite qui approche le mieux f localement. Nous pouvons alors déduire l'équation de la tangente à la courbe au point considéré (et en déduire alors que notre définition de nombre dérivé est la bonne).

Proposition 14.1.4. *En reprenant les notations précédentes, f admet une tangente en a si et seulement si df_a existe, et dans ce cas son équation est*

$$y = df_a \times (x - a) + f(a)$$

Démonstration. Supposons que f est dérivable en a . Montrons alors que $\Delta : y = df_a \times (x - a) + f(a)$ est tangente à la courbe au point a (cela suffit étant donné que la tangente est unique). Soit Δ' une autre droite telle que $\Delta'(a) = f(a)$, en notant le coefficient directeur m on remarque qu'on peut mettre Δ' sous la forme $\Delta'(x) = m \times (x - a) + f(a)$ car l'expression de droite a le même coefficient directeur et vaut la même valeur en un point (donc a la même ordonnée à l'origine). On veut maintenant trouver $\delta > 0$ tel que pour $|x - a| < \delta$ on ait

$$|f(x) - \Delta(x)| \leq |f(x) - \Delta'(x)| \quad (2)$$

qui est vraie pour $x = a$ (de façon évidente). On cherche donc une condition pour avoir cette inégalité dans le cas $x \neq a$: réécrivons cette inégalité et divisons-la par $|x - a|$ pour faire apparaître un taux d'accroissement :

$$\begin{aligned} |f(x) - \Delta(x)| \leq |f(x) - \Delta'(x)| &\iff |f(x) - f(a) - df_a \times (x - a)| \leq |f(x) - f(a) - m \times (x - a)| \\ &\iff \left| \frac{f(x) - f(a)}{x - a} - df_a \right| |x - a| \leq \left| \frac{f(x) - f(a)}{x - a} - m \right| |x - a| \\ &\iff \left| \frac{f(x) - f(a)}{x - a} - df_a \right| \leq \left| \frac{f(x) - f(a)}{x - a} - m \right| \end{aligned}$$

Or comme le taux d'accroissement entre x et a tend vers df_a quand x tend vers a , on en déduit l'existence d'un $\delta > 0$ tel que l'inégalité est vraie : si $m = df_a$ l'inégalité est vraie par égalité (prenons par exemple $\delta = 1$), sinon on a une valeur strictement positive à droite donc on peut utiliser la définition de limite en prenant le terme de droite pour ε et on en déduit l'existence de δ tel que $|x - a| < \delta \implies (2)$. On en déduit donc que Δ est la tangente à f en a .

Pour le sens réciproque, nous allons raisonner par contraposition. On suppose donc que df_a n'a pas de limite en a , ce qui signifie qu'il existe une suite x_n telle que $\lim x_n = a$, la suite $y_n = \lim \frac{f(x_n) - f(a)}{x_n - a}$ diverge (et $x_n \neq a$ à partir d'un certain rang, car sinon la suite y_n n'est pas définie, donc on pourra diviser par $x_n - a$), et on veut montrer qu'alors f n'admet pas de tangente en a . Raisonnons alors par cas, suivant si la suite (y_n) est bornée.

- Si (y_n) n'est pas bornée, alors on peut trouver une suite extraite $y_{\varphi(n)}$ telle que $\lim y_{\varphi(n)} = +\infty$ ou $\lim y_{\varphi(n)} = -\infty$. Supposons qu'il existe une tangente Δ à f en a et notons m son coefficient

directeur. Dans le cas où $\lim y_{\varphi(n)} = +\infty$, posons $m' = m + 1$ et $\Delta' : y = (m + 1)(x - a) + f(a)$, par propriété de la tangente de Δ à f en a on trouve V voisinage de a tel que pour tout $x \in V$ l'inégalité (2) soit vérifiée. Comme on $\lim x_n = a$, on trouve un rang n_0 tel que pour tout $p > n_0, u_p \in V$. Ainsi pour tout $p > n_0$ on a l'inégalité

$$|y_n - m| \leq |y_n - (m + 1)|$$

mais comme $\lim y_{\varphi(n)} = +\infty$, on trouve n_1 tel que $\forall p > n_1, y_{\varphi(p)} > m + 1$. Ainsi pour $p > \max(n_0, n_1)$ on trouve $|y_{\varphi(p)} - (m + 1)| = y_{\varphi(p)} - m - 1$ et $|y_{\varphi(n)} - m| = y_{\varphi(n)} - m$ donc l'inégalité (2) devient $0 \leq -1$: c'est absurde. De la même façon, si $\lim y_{\varphi(n+1)} = -\infty$ on refait la même construction mais avec $m' = m - 1$, et on aboutit à $0 \geq 1$, qui est absurde aussi. Ainsi si (y_n) est non bornée alors f n'admet pas de tangente en a .

- Si (y_n) est bornée, alors par un résultat précédent sur les suites extraites, on sait qu'il existe au moins deux suites extraites qui convergent vers des valeurs différentes, notons φ et ψ les fonctions d'extraction et m et m' les limite associées. On construit alors $\Delta : y = m \times (x - a) + f(a)$ et $\Delta' : y = m' \times (x - a) + f(a)$. Soit $\Delta'' : y = m'' \times (x - a) + f(a)$ une droite quelconque différente de Δ qui passe par le point $(a, f(a))$, montrons que pour tout voisinage V de a il existe un $x \in V$ tel que $|f(x) - \Delta(x)| < |f(x) - \Delta''(x)|$.

Soit V un voisinage de A . Comme $\lim x_n = a$, on trouve un rang n_0 tel que pour tout $p > n_0, x_p \in V$. On a alors, pour $p > n_0$:

$$|f(x_{\varphi(p)}) - \Delta(x_{\varphi(p)})| < |f(x_{\varphi(p)}) - \Delta''(x_{\varphi(p)})| \iff |y_{\varphi(p)} - m| < |y_{\varphi(p)} - m'|$$

or on sait que $\lim y_{\varphi(n)} = m$ donc en fixant $\varepsilon_n = 0.5 \times |m - m'| > 0$ dans la définition de limite de $(y_{\varphi(n)})$ on trouve $n_1 > n_0$ tel que pour tout $p > n_1, |y_{\varphi(p)} - m| < 0.5 \times |m - m'|$ d'où

$$\begin{aligned} |y_{\varphi(p)} - m'| &= |y_{\varphi(p)} - m + m - m'| \\ &\geq ||y_{\varphi(p)} - m| - |m - m'|| \\ &> ||y_{\varphi(p)} - m| - 2 \times |y_{\varphi(p)} - m|| \\ &> |y_{\varphi(p)} - m| \end{aligned}$$

Donc en prenant $x = x_{\varphi(n_1+1)}$ on trouve $x \in V$ tel que $|f(x) - \Delta(x)| < |f(x) - \Delta''(x)|$.

Ainsi pour $\Delta'' \neq \Delta$, Δ'' ne peut être une tangente à f en a . Mais notre raisonnement peut être mené de l'exacte même manière pour prouver que Δ'' ne peut être une tangente à f en a si $\Delta'' \neq \Delta'$, et comme $\Delta \neq \Delta'$ par hypothèse, on en déduit qu'il n'existe pas de tangente à f en a .

Ainsi, par contraposée, si f admet une tangente en a , alors elle est dérivable en a . \square

14.2 Fonction dérivée

De façon analogue à l'extension de fonction continue en un point à fonction continue sur une partie, on peut considérer naturellement pour une fonction le fait d'admettre une dérivée sur chaque point de son ensemble.

Définition 14.3 (Dérivabilité). Soit $f : F \rightarrow \mathbb{R}, F \subseteq \mathbb{R}$. On dit que f est dérivable si pour tout $x \in F$, f est dérivable en x . On dit que f est dérivable sur $I \subseteq F$ si $f|_I$ est dérivable.

Remarque. Comme la dérivabilité en un point implique la continuité en ce point, une fonction dérivable est continue.

De plus, on remarque un point important qui différencie la continuité et la dérivabilité : être continu est une condition (la limite de $f(x)$ en a ne peut qu'être $f(a)$) mais être dérivable est l'information d'une valeur (la limite du taux d'accroissement). On peut donc associer à une fonction dérivable et un point, le nombre dérivé en ce point. Cela nous pousse à la définition suivante :

Définition 14.4 (Fonction dérivée). Soit $f : F \rightarrow \mathbb{R}$ une fonction dérivable. On appelle fonction dérivée, et l'on note df , ou f' , ou $\frac{d}{dx}f$, la fonction

$$\begin{aligned} f' : F &\longrightarrow \mathbb{R} \\ x &\longmapsto df_x \end{aligned}$$

On utilisera principalement la notation f' pour définir la fonction dérivée, pour sa concision.

14.2.1 Opérations sur les dérivées

Nous allons maintenant donner des outils pour calculer efficacement des fonctions dérivées. Le premier point est la stabilité par combinaison linéaire, qui permet donc de sommer des fonctions et de les multiplier par un réel.

Proposition 14.2.1 (Stabilité par combinaison linéaire). Soient $f, g : F \rightarrow \mathbb{R}$ deux fonctions dérivables, et $k, l \in \mathbb{R}$. Alors $kf + lg$ est dérivable et de plus

$$d(kf + lg)_x = kf'(x) + lg'(x)$$

Démonstration. Calculons le taux d'accroissement de la fonction entre un point $a \in F$ et un point $a + h$:

$$\begin{aligned} \frac{(kf + lg)(a + h) - (kf + lg)(a)}{h} &= \frac{kf(a + h) + lg(a + h) - kf(a) - lg(a)}{h} \\ &= \frac{k(f(a + h) - f(a)) + l(g(a + h) - g(a))}{h} \\ &= k \frac{f(a + h) - f(a)}{h} + l \frac{g(a + h) - g(a)}{h} \end{aligned}$$

Or le terme de gauche tend vers $kf'(a)$ et celui de droite vers $lg'(a)$ par opérations sur les limites finies. On en déduit que

$$\boxed{\lim_{h \rightarrow 0} \frac{(kf + lg)(a + h) - (kf + lg)(a)}{h} = kf'(a) + lg'(a)}$$

□

Calculons maintenant notre première classe de fonctions dont la dérivée est facile à calculer :

Proposition 14.2.2 (Dérivée de polynôme). Soit $P \in \mathbb{R}[X]$, $P(X) = \sum_{k=0}^n p_k X^k$. Alors la fonction f_P associée à P est dérivable, et sa dérivée est

$$\begin{aligned} f'_P : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto \sum_{k=0}^{n-1} ((k+1) \times p_{k+1}) x^k \end{aligned}$$

et correspond donc à la fonction polynomiale associée à

$$P'(X) = \sum_{k=0}^{n-1} ((k+1) \times p_{k+1}) X^k$$

Démonstration. La preuve se fera en deux temps : nous allons prouver notre résultat sur les monômes, puis l'étendre aux polynômes.

Montrons qu'une fonction $f : x \mapsto x^n$ avec $n \in \mathbb{N}$ est dérivable et que sa dérivée est $f' : x \mapsto nx^{n-1}$ en calculant le taux d'accroissement entre $a \in \mathbb{R}$ et $a + h$:

$$\begin{aligned} \frac{f(a+h) - f(a)}{h} &= \frac{1}{h}((a+h)^n - a^n) \\ &= \frac{1}{h} \left(\sum_{i=0}^n \binom{n}{i} a^i h^{n-i} - a^n \right) \\ &= \frac{1}{h} \left(\sum_{i=0}^{n-1} \binom{n}{i} a^i h^{n-i} \right) \\ &= \sum_{i=0}^{n-1} \binom{n}{i} a^i h^{n-i-1} \\ &= \binom{n}{1} a^{n-1} + \sum_{i=0}^{n-2} a^i h^{n-i-1} \end{aligned}$$

mais pour $i \in \{0, \dots, n-2\}$, $n-i-1 > 0$ donc $\lim_{h \rightarrow 0} h^{n-i-1} = 0$ d'où par somme, sachant que $\binom{n}{1} = n$,

$$\lim_{h \rightarrow 0} na^{n-1} \sum_{i=0}^{n-2} a^i h^{n-i-1} = na^{n-1}$$

donc f est dérivable et sa dérivée est $x \mapsto nx^{n-1}$.

Comme chaque monôme est dérivable, la combinaison linéaire $x \mapsto \sum_{k=0}^n p_k x^k$ est aussi dérivable, donc f_P est dérivable. De plus, d'après la propriété précédente, on sait que

$$d(x \mapsto \sum_{k=0}^n p_k x^k) = \sum_{k=0}^n p_k d(x \mapsto x^k)$$

d'où, comme la dérivée de la fonction $x \mapsto 1$ est nulle, et en changeant l'indice avec $i \leftarrow i + 1$, le résultat :

$$d(f_P)_x = \sum_{k=0}^{n-1} ((k+1)p_{k+1})x^k$$

□

Intéressons-nous maintenant à la stabilité par produit.

Proposition 14.2.3 (Dérivée d'un produit). *Soient $f, g : F \rightarrow \mathbb{R}$ deux fonctions dérivables. Alors la fonction $f \times g$ est dérivable et de plus on a*

$$d(f \times g)_x = f(x)g'(x) + f'(x)g(x)$$

Démonstration. Calculons le taux d'accroissement de $f \times g$ entre un point $a \in F$ et $a + h$:

$$\begin{aligned} \frac{(f \times g)(a+h) - (f \times g)(a)}{h} &= \frac{(f(a+h) \times g(a+h)) - (f(a) \times g(a))}{h} \\ &= \frac{1}{h} (f(a+h) \times g(a+h) - f(a+h) \times g(a) + (f(a+h) - f(a)) \times g(a)) \\ &= \frac{1}{h} (f(a+h)(g(a+h) - g(a)) + g(a)(f(a+h) - f(a))) \\ &= f(a+h) \frac{g(a+h) - g(a)}{h} + g(a) \frac{f(a+h) - f(a)}{h} \end{aligned}$$

Analysons maintenant les deux termes. Pour celui de gauche, par continuité de f (car f est dérivable) on déduit que $\lim_{h \rightarrow 0} f(a+h) = f(a)$, pour les deux quotients, on reconnaît directement $g'(a)$ et $f'(a)$ respectivement, ce qui nous donne la dérivabilité de $f \times g$ et la formule :

$$\boxed{d(f \times g)_x = f'(x)g(x) + f(x)g'(x)}$$

□

Nous allons montrer que la composition de fonctions dérivables reste dérivable.

Proposition 14.2.4. Soit $f : F \rightarrow \mathbb{R}$, $g : G \rightarrow \mathbb{R}$ avec $F, G \subseteq \mathbb{R}$ et $f(F) \subseteq G$. Alors la fonction $g \circ f$ est dérivable et on a, pour $x \in F$:

$$d(g \circ f)_x = f'(x) \times g' \circ f(x)$$

Démonstration. Montrons ce résultat en calculant le taux d'accroissement entre un point $a \in F$ et $a+h$. Si f s'annule au voisinage de a , alors $g \circ f$ aussi et la formule est vérifiée. Sinon alors on peut (en prenant h assez petit) considérer que $f(a+h) - f(a)$ ne s'annule pas, d'où :

$$\frac{(g \circ f)(a+h) - (g \circ f)(a)}{h} = \frac{g(f(a+h)) - g(f(a))}{f(a+h) - f(a)} \frac{f(a+h) - f(a)}{h}$$

Or comme f est continue, $\lim_{h \rightarrow 0} f(a+h) = f(a)$ donc le terme de gauche tend vers $f'(a)g(f(a))$. Le terme de droite tend vers $f'(a)$. D'où la dérivabilité de $g \circ f$ et la formule :

$$\boxed{d(g \circ f)_x = g' \circ f(x) \times f'(x)}$$

□

Une fonction dont la dérivée est importante est la fonction inverse :

Proposition 14.2.5. La fonction

$$\begin{aligned} i : \mathbb{R}^* &\longrightarrow \mathbb{R} \\ x &\longmapsto \frac{1}{x} \end{aligned}$$

est dérivable.

Démonstration. Calculons la limite de son taux d'accroissement entre a et $a+h$:

$$\begin{aligned} \frac{\frac{1}{a+h} - \frac{1}{a}}{h} &= \frac{\frac{a - (a+h)}{a(a+h)}}{h} \\ &= \frac{-h}{ha(a+h)} \\ &= \frac{-1}{a(a+h)} \end{aligned}$$

Or $\lim_{h \rightarrow 0} a+h = a$ d'où

$$\boxed{\lim_{h \rightarrow 0} \frac{\frac{1}{a+h} - \frac{1}{a}}{h} = \frac{-1}{a^2}}$$

donc la fonction inverse est dérivable et sa dérivée est

$$\begin{aligned} i' : \mathbb{R}^* &\longrightarrow \mathbb{R} \\ x &\longmapsto \frac{-1}{x^2} \end{aligned}$$

□

Exercice 14.2. On rappelle que pour $x \neq 0$, $x^{-k} = \frac{1}{x^k}$. En utilisant les résultats précédents, montrer que pour $k \in \mathbb{Z}$, la fonction $x \mapsto x^k$ est dérivable, sur \mathbb{R} si $k \geq 0$ et sur \mathbb{R}^* sinon, et que la dérivée de cette fonction est la fonction $x \mapsto kx^{k-1}$.

Exercice 14.3. Dédurre des résultats précédents que si une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ dérivable ne s'annule pas, alors son inverse $\frac{1}{f} : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto \frac{1}{f(x)}$ est aussi dérivable, et que la dérivée de cette fonction en x est

$$\frac{-f'(x)}{f(x)^2}$$

Exercice 14.4. A partir de l'exercice précédent, en déduire que si u et v sont deux fonctions dérivables et v ne s'annule pas, alors la fonction $x \mapsto \frac{u(x)}{v(x)}$ est dérivable et sa fonction dérivée est

$$x \mapsto \frac{u'(x)v(x) - u(x)v'(x)}{(v(x))^2}$$

14.2.2 Fonction réciproque

Nous allons donner un autre théorème important de dérivation, qui permet de dériver une fonction bijection par rapport à sa réciproque. Nous l'utiliserons pour montrer que les fonctions racines n èmes sont dérivables et calculer leur dérivée.

Théorème 14.1 (Dérivée d'une fonction réciproque). *Soit $f : F \rightarrow G$, où $F, G \subseteq \mathbb{R}$ une fonction dérivable bijective dont la dérivée ne s'annule pas. Alors la fonction $f^{-1} : G \rightarrow F$ est aussi dérivable, et sa dérivée est donnée par :*

$$f^{-1'}(x) = \frac{1}{f'(f^{-1}(x))}$$

Démonstration. Nous allons calculer le taux d'accroissement de f^{-1} entre a et b deux éléments de G , en remarquant que par bijectivité de f , on peut réécrire $a = f(f^{-1}(a))$:

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

Comme f' ne s'annule pas, on en déduit que

$$\lim_{b \rightarrow a} \frac{f^{-1}(b) - f^{-1}(a)}{b - a} = \lim_{b \rightarrow a} \frac{1}{\frac{f^{-1}(b) - f^{-1}(a)}{f^{-1}(f(b)) - f^{-1}(a)}} = \frac{1}{f'(f^{-1}(a))}$$

donc f^{-1} est dérivable et sa dérivée est

$$\boxed{\begin{array}{ccc} f^{-1'} : & G & \longrightarrow \mathbb{R} \\ & x & \longmapsto \frac{1}{f'(f^{-1}(x))} \end{array}}$$

□

Déduisons-en que la fonction $x \mapsto \sqrt[n]{x}$ est dérivable sur \mathbb{R}_+^* .

Proposition 14.2.6. *Pour tout $n \in \mathbb{N}^*$, la fonction $x \mapsto x^n$ est bijective de \mathbb{R}_+ dans \mathbb{R}_+ , et en notant $\sqrt[n]{x}$ sa réciproque alors la fonction $x \mapsto \sqrt[n]{x}$ est dérivable sur \mathbb{R}_+^* , de dérivée*

$$x \mapsto \frac{\sqrt[n]{x}}{nx}$$

Démonstration. Soit $x \in \mathbb{N}^*$. Montrons d'abord que $f := x \mapsto x^n$ est une bijection de \mathbb{R}_+ dans \mathbb{R}_+ :

Elle est d'abord strictement croissante : si $x > y$ alors $x^n > y^n$ pour x et y positifs. De plus, on a montré qu'elle était dérivable de dérivée $x \mapsto nx^{n-1}$, donc elle est continue. Enfin, elle vaut 0 en 0 et diverge vers $+\infty$ en $+\infty$ car pour tout $M > 0$, $M^n > 0$. La fonction $x \mapsto x^n$ est donc une fonction dérivable bijective.

Enfin, la dérivée de cette fonction est non nulle pour $x \neq 0$: si $x \neq 0$ alors $df_x = nx^{n-1} \neq 0$. La fonction induite de \mathbb{R}_+^* vers \mathbb{R}_+^* est évidemment toujours bijective.

En utilisant le théorème précédent, on en déduit que la fonction réciproque, que nous noterons $x \mapsto \sqrt[n]{x}$ (ou g ici pour des facilités de notation) et appelons racine n ième, est dérivable sur \mathbb{R}_+^* et on a $g'(x) = \frac{1}{f'(\sqrt[n]{x})}$, or $f'(\sqrt[n]{x}) = n(\sqrt[n]{x})^{n-1} = n \frac{x}{\sqrt[n]{x}}$, donc

$$g'(x) = \frac{\sqrt[n]{x}}{nx}$$

□

Exercice 14.5. On considère maintenant que pour $n > 0$ et $x \in \mathbb{R}_+^*$, $x^{1/n} = \sqrt[n]{x}$. Montrer tout d'abord que pour $p \in \mathbb{Z}, q \in \mathbb{N}^*$ et $x \in \mathbb{R}_+^*$, $x^{p/q} = \sqrt[q]{x^p} = (\sqrt[q]{x})^p$. Montrer alors que la formule de dérivation des puissances fonctionne aussi dans le cas rationnel, c'est-à-dire montrer que pour $a \in \mathbb{R}_+^*$ et $r \in \mathbb{Q}$:

$$d(x \mapsto x^r)_a = ra^{r-1}$$

14.3 Théorèmes sur la dérivation

Dans cette partie, nous allons démontrer les théorèmes essentiels à propos des fonctions dérivées, permettant par exemple de lier l'étude du signe d'une fonction dérivée et l'étude des variations de la fonction initiale.

Commençons par un lemme important :

Lemme 14.3.1. Soit $f : I \rightarrow \mathbb{R}$ une fonction dérivable et I un intervalle ouvert. Soit $x \in I$ tel que $\forall y \in I, f(y) \leq f(x)$, alors

$$f'(x) = 0$$

Démonstration. Le taux d'accroissement de f entre x et $x+h$ est du signe opposé à celui de h car $f(x+h) - f(x) \leq 0$ par hypothèse. Cela signifie donc que lorsque h tend vers 0 par valeurs supérieures, le taux d'accroissement tend vers une valeur négative. De même lorsque h tend vers 0 par valeurs inférieures, le taux d'accroissement limite est positif. Comme la limite existe, elle est égale à la limite à gauche et à droite, donc $f'(x) \leq 0$ et $f'(x) \geq 0$, donc $f'(x) = 0$. □

Exercice 14.6. Montrer que la version du lemme où x est un minimum est vraie aussi. Montrer de même que le lemme reste vrai si l'on suppose que x est extrema local, c'est-à-dire qu'il existe un voisinage $V \in \mathcal{V}_x$ tel que pour tout $y \in V, f(y) \leq f(x)$ (respectivement $f(y) \geq f(x)$ pour un minimum local).

Nous allons en déduire le théorème de Rolle :

Théorème 14.2 (Rolle). Soit $f : I \rightarrow \mathbb{R}$ une fonction dérivable, où I est un intervalle. S'il existe deux éléments $a, b \in I$ tels que $a < b$ et $f(a) = f(b)$ alors il existe $c \in]a, b[$ tel que $f'(c) = 0$.

Démonstration. Puisque f est dérivable, elle est continue, donc elle est continue sur $[a, b]$. Par le théorème des bornes atteintes, il existe donc un maximum de f sur $[a, b]$, dont nous noterons c l'antécédent et d l'image et un minimum dont nous noterons γ l'antécédent et δ l'image. Si $\delta = d = f(a)$ alors tout élément de $[a, b]$ vaut $f(a)$ et on trouve aisément un élément de dérivée nulle. Sinon, alors au choix $c \neq a \wedge c \neq b$ ou $\gamma \neq a \wedge \gamma \neq b$, et comme la fonction f est dérivable sur l'intervalle ouvert $]a, b[$ on en déduit qu'en c (respectivement γ) la dérivée est nulle. Ainsi dans tous les cas il existe c tel que $f'(c) = 0$. □

Remarque. Les hypothèses minimales pour appliquer ce théorème sont simplement que f est continue sur $[a, b]$ et dérivable sur $]a, b[$.

Ce théorème permet lui-même de démontrer un autre théorème :

Théorème 14.3 (Accroissements finis). *Soit $f : I \rightarrow \mathbb{R}$, I un intervalle. Soient $a, b \in I$ deux éléments distincts, $a < b$, alors il existe $c \in]a, b[$ tel que $f'(c) = \frac{f(b) - f(a)}{b - a}$.*

Démonstration. On pose la fonction $g : x \mapsto f(x) - \frac{f(b) - f(a)}{b - a}x$, qui est dérivable comme combinaison linéaire de fonctions dérivables. Sa fonction dérivée est $g' : x \mapsto f'(x) - \frac{f(b) - f(a)}{b - a}$ et de plus :

$$\begin{aligned} g(a) &= f(a) - \frac{f(b) - f(a)}{b - a}a \\ &= \frac{1}{b - a}(bf(a) - af(a) - af(b) + af(a)) \\ &= \frac{bf(a) - af(b)}{b - a} \\ g(b) &= f(b) - \frac{f(b) - f(a)}{b - a}b \\ &= \frac{1}{b - a}(bf(b) - af(b) - bf(b) + bf(a)) \\ &= \frac{bf(a) - af(b)}{b - a} \end{aligned}$$

Donc $g(a) = g(b)$, donc en appliquant le théorème de Rolle on trouve $c \in]a, b[$ tel que $g'(c) = 0$, c'est-à-dire tel que $f'(c) - \frac{f(b) - f(a)}{b - a} = 0$, donc

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

□

Remarque. Là encore, il nous suffit que f soit continue sur $[a, b]$ et dérivable sur $]a, b[$ puisque nous utilisons directement le théorème de Rolle.

Donnons ensuite le cas des inégalités :

Proposition 14.3.1 (Inégalité des accroissements finis). *Soit $f : I \rightarrow \mathbb{R}$ une fonction dérivable, où I est un intervalle. Soient $a < b$ deux réels et $M \in \mathbb{R}$ tels que $\forall x \in]a, b[, |f'(x)| \leq M$, alors $\left| \frac{f(b) - f(a)}{b - a} \right| \leq M$.*

Démontrons alors le lien croissance / signe de la dérivée :

Proposition 14.3.2 (Lien croissance / signe de la dérivée). *Soit $f : I \rightarrow \mathbb{R}$ une fonction dérivable, I un intervalle. Alors :*

- si $\forall x \in I, f'(x) \geq 0$ alors f est croissante.
- si $\forall x \in I, f'(x) \leq 0$ alors f est décroissante.
- si $\forall x \in I, f'(x) > 0$ alors f est strictement croissante.
- si $\forall x \in I, f'(x) < 0$ alors f est strictement décroissante.

Démonstration. Nous allons traiter le premier cas uniquement, tous les autres étant identiques en changeant le signe intervenant dans la démonstration. Supposons que $\forall x \in I, f'(x) \geq 0$. Soient $x, y \in I, x \geq y$. Le théorème des accroissements finis nous indique qu'il existe $z \in]x, y[$ tel que $f'(z) = \frac{f(y) - f(x)}{y - x}$ or $f'(x) \geq 0$ donc $\frac{f(y) - f(x)}{y - x} \geq 0$. Comme par hypothèse, $y - x \geq 0$, on en déduit que $f(y) - f(x) \geq 0$. Donc $\boxed{f \text{ est croissante.}}$ □

Remarque. A partir de ces résultats, on peut établir les variations d'une fonction en étudiant simplement le signe de sa dérivée : sur chaque intervalle où le signe de f' est constant, f est monotone.

15 Intégration de fonctions continues par morceaux sur un segment

Dans ce chapitre, nous allons construire les intégrales de Riemann. Les intégrales sont souvent vues comme l'opération inversant (dans la mesure du possible) la dérivation. Nous allons construire les intégrales sur un segment, et nous allons considérer pour cela les fonctions continues par morceaux.

L'intégrale s'interprète comme l'aire algébrique sous la courbe d'une fonction. Pour calculer cette aire, nous allons approcher la courbe par des fonctions dont l'aire sous la courbe est bien connue, puis calculer, en quelques sortes, la limite de ces aires de fonctions approchantes.

15.1 Fonctions en escalier et continues par morceaux

La première étape sera donc de définir les fonctions dont le calcul de l'intégral est simple, puis de montrer que les fonctions que nous voulons intégrer peuvent s'approcher par de telles fonctions. Commençons par définir ce que sont les fonctions dont nous voulons calculer l'intégrale.

Définition 15.1 (Fonction continue par morceaux). Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction, $a, b \in \mathbb{R}$. On dit que f est continue par morceaux s'il existe une famille finie strictement croissante $(x_i)_{1 \leq i \leq n}$ telle que $x_1 = a$, $x_n = b$ et f est continue sur chaque intervalle $]x_i, x_{i+1}[$.

Remarque. Une fonction continue par morceaux est donc simplement une fonction comprenant un nombre fini de discontinuités. Nous donnons cette définition sur un segment car nous ne travaillerons que sur des segments dans ce chapitre.

Ensuite, définissons les fonctions en escalier, qui sont les fonctions qui nous serviront à approcher les fonctions continues par morceaux.

Définition 15.2 (Fonction en escalier). Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction, $a, b \in \mathbb{R}$. On dit que f est une fonction en escalier s'il existe une famille finie strictement croissante $(x_i)_{1 \leq i \leq n}$ telle que $x_1 = a$, $x_n = b$ et f est constante sur chaque intervalle $]x_i, x_{i+1}[$. On note $\Psi[a, b]$ l'ensemble des fonction en escalier de domaine $[a, b]$.

Notre premier résultat va être l'approximation de fonctions continues par morceaux par des fonctions en escalier :

Proposition 15.1.1. Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction continue par morceaux, $a, b \in \mathbb{R}$. Soit $\varepsilon > 0$, alors il existe une fonction en escalier e_+ supérieure à f telle que $\forall x \in [a, b], |f(x) - e_+(x)| < \varepsilon$ et de même il existe une fonction en escalier e_- inférieure à f telle que $\forall x \in [a, b], |f(x) - e_-(x)| < \varepsilon$.

Démonstration. Nous allons d'abord prouver ce résultat dans le cas où f est continue. Si f est continue, alors d'après le théorème de Heine f est uniformément continue (puisque continue sur un segment). Étant donné le $\varepsilon > 0$ de l'énoncé, nous trouvons par continuité uniforme de f un $\delta > 0$ tel que $\forall x, y \in [a, b], |x - y| < \delta \implies |f(x) - f(y)| < \varepsilon$. On peut alors découper $[a, b]$ en morceaux, en se donnant la famille

$$x_i = a + \delta i$$

pour $i < n$ et $x_n = b$, où i varie entre 1 et $n = \left\lceil \frac{b-a}{\delta} \right\rceil$. Autrement dit, notre famille (x_i) a la forme $a, a + \delta, a + 2\delta, \dots, a + n\delta, b$. Sur chaque $]x_i, x_{i+1}[$, comme f est continue, f a un minimum et un maximum. On définit $e_+ = \max_{x \in]x_i, x_{i+1}[} f(x)$ et $e_- = \min_{x \in]x_i, x_{i+1}[} f(x)$ ainsi que $e_-(x_i) = e_+(x_i) = f(x_i)$.

Soit $x \in [a, b]$, soit $x = x_i$ pour un certain i auquel cas l'inégalité est vérifiée pour e_+ et e_- , soit on trouve i tel que $x \in]x_i, x_{i+1}[$. Par définition, le minimum et le maximum de f sur ce segment sont atteints à des points à une distance à x inférieure à δ , donc $|f(x) - e_+(x)| < \varepsilon$ et $|f(x) - e_-(x)| < \varepsilon$. De plus, par comme nous les avons construites pour être respectivement un maximum et un minimum, e_+ et e_- sont bien respectivement supérieure et inférieure à f .

Nous pouvons maintenant démontrer le cas général par récurrence sur la taille de la famille (x_i) découpant l'intervalle $[a, b]$ en des intervalles sur lesquels f est continue.

- Pour le cas où $i = 2$, cela signifie que f est continue, et nous venons d'en faire la démonstration.

- Si la propriété est vraie au rang k , alors pour une fonction découpée en fonctions continues par $k+1$ points, on peut considérer f sur l'intervalle $[a, x_k]$ et sur l'intervalle $[x_k, x_{k+1}]$. Par hypothèse de récurrence, on peut donc construire e_+ et e_- sur chacun des deux intervalles, mais la valeur en x_k n'est pas forcément la même. Cependant, en changeant le point $e_+(x_k) = e_-(x_k) = f(x_k)$ on obtient bien une fonction qui vérifie ce que l'on veut montrer.

Toute fonction continue par morceaux peut être approchée par des fonctions en escalier.

□

15.2 Intégrale de Riemann

Nous pouvons désormais définir l'intégrale de Riemann en tant que telle. La première étape est de la définir sur les fonctions en escalier. Une fonction en escalier peut se voir comme une succession de rectangle, donc l'aire sous la courbe d'une fonction en escalier est simplement une somme d'aires de rectangles.

Définition 15.3 (Intégrale de Riemann d'une fonction en escalier). Soit $f \in \Psi_{[a,b]}$ une fonction en escalier, avec un découpage $(x_i)_{1 \leq i \leq n}$ de l'intervalle $[a, b]$ sur lequel f est constante. On définit

$$\int_{[a,b]} f = \sum_{i=1}^{n-1} (x_{i+1} - x_i) f\left(\frac{x_{i+1} + x_i}{2}\right)$$

l'intégrale entre a et b de f .

Remarque. Si $f \leq g$, au sens où $\forall x \in [a, b], f(x) \leq g(x)$, alors $\int_{[a,b]} f \leq \int_{[a,b]} g$.

Exercice 15.1. Démontrer la remarque précédente.

Remarquons alors que pour deux fonctions en escalier, l'une étant plus fine que l'autre (au sens où le découpage associé à la première contient tous les points du découpage associé à la deuxième), si les deux sont des approximations par valeurs inférieures d'une fonction f alors l'intégrale de la fonction plus fine sera supérieure à l'intégrale de l'autre fonction. De même pour des approximations par valeurs supérieures, on peut remarquer que les intégrales vont en décroissant. Comme on voudrait moralement que l'intégrale de f soit la limite des intégrales des approximations par fonctions en escalier, on va utiliser la notion de borne supérieure et de borne inférieure.

Définition 15.4 (Intégrale de Riemann). Soit $f : [a, b] \rightarrow \mathbb{R}$, on définit l'intégrale de Riemann de f (ou juste « intégrale de f ») par

$$\int_{[a,b]} f = \sup_{\substack{g \in \Psi_{[a,b]} \\ g \leq f}} \left(\int_{[a,b]} g \right) = \inf_{\substack{g \in \Psi_{[a,b]} \\ g \geq f}} \left(\int_{[a,b]} g \right)$$

Démonstration. Montrons que les deux bornes sont égales. D'abord, par transitivité de l'inégalité, une fonction inférieure à f est inférieure à une fonction supérieure à f . Ainsi les intégrales de fonctions en escaliers inférieures à f sont toutes des minorants de l'ensemble des intégrales de fonctions en escalier supérieures. De même les intégrales de fonctions en escalier supérieures sont des majorants de l'ensemble des intégrales de fonctions en escalier inférieures. Enfin, si l'on prend $\varepsilon > 0$, on montre qu'il existe $g \leq f$ et $h \geq f$ telles que

$$\left(\int_{[a,b]} g \right) + \varepsilon \geq \int_{[a,b]} h$$

et de même dans l'autre sens (mais nous ne montrerons que le premier cas, l'autre étant identique). On sait qu'on peut approcher f par valeurs inférieures à tout $\varepsilon > 0$ près, on l'approche alors par g qui est à $\frac{\varepsilon}{b-a}$ près de f . En prenant h approchant f par valeurs supérieures à $\frac{\varepsilon}{2(b-a)}$ près, on en déduit que

$$\left(\int_{[a,b]} g \right) + \varepsilon \geq \int_{[a,b]} h$$

d'où le fait que l'intégrale de f est la valeur commune des deux bornes.

□

15.3 Propriétés des intégrales

Nous allons donner et démontrer les propriétés essentielles des intégrales. La première est la valeur d'une intégrale constante.

Proposition 15.3.1 (Intégrale d'une constante). *Soit $k \in \mathbb{R}$, l'intégrale de la fonction $f : x \mapsto k$ est*

$$\int_{[a,b]} f = k(b-a)$$

Démonstration. La fonction étant une fonction en escalier, son intégrale se calcule directement par la formule de base, et on peut considérer le découpage (a, b) . Si une fonction en escalier est inférieure à f , alors son intégrale sera inférieure, et de même pour une fonction en escalier supérieure. On en déduit que

$$\int_{[a,b]} f = k(b-a)$$

□

La propriété suivante est la linéarité, qui permet de simplifier les intégrales de sommes de fonctions.

Proposition 15.3.2 (Linéarité). *Soient $f, g : [a, b] \rightarrow \mathbb{R}$ deux fonctions continues par morceaux, et $k, l \in \mathbb{R}$. Alors*

$$\int_{[a,b]} (kf + lg) = k \int_{[a,b]} f + l \int_{[a,b]} g$$

Démonstration. Montrons que l'on peut approcher l'intégrale de $kf + lg$ par des approximations de f et g . Soit $\varepsilon > 0$, on trouve une approximation h de f par valeurs inférieures à $\frac{\varepsilon}{2k}$ près et une approximation j de g par valeurs inférieures à $\frac{\varepsilon}{2l}$ près. Alors $kh + lj$ est une approximation de $kf + lg$ à ε près. On peut donc approcher à n'importe quelle précision l'intégrale de $kf + lg$ par l'intégrale d'une fonction de la forme $kh + lj$ avec h et j approximations inférieures respectivement de f et de g . On en déduit que

$$\int_{[a,b]} (kf + lg) \leq k \int_{[a,b]} f + l \int_{[a,b]} g$$

Réciproquement, si l'on a une approximation h de f à ε près et une approximation j de g à ε' près, alors une approximation de $kf + lg$ à $\min(\varepsilon, \varepsilon')$ fournit une meilleure approximation que $kh + lj$, d'où l'inégalité inverse. On en déduit que

$$\int_{[a,b]} (kf + lg) = k \int_{[a,b]} f + l \int_{[a,b]} g$$

□

La positivité de l'intégrale sera utile pour en déduire la croissance, entre autre.

Proposition 15.3.3 (Positivité). *Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction continue par morceaux. Si $\forall x \in [a, b], f(x) \geq 0$ alors $\int_{[a,b]} f \geq 0$*

Démonstration. Supposons que $\forall x \in [a, b], f(x) \geq 0$. Alors toute fonction en escalier supérieure à f est positive, donc les intégrales de fonctions en escalier supérieures à f sont toutes positives, ce qui signifie que leur borne inférieure est positive :

$$\int_{[a,b]} f \geq 0$$

□

Corollaire 15.3.1. Si pour deux fonctions f, g continues par morceaux on a $\forall x \in [a, b], f(x) \leq g(x)$ alors $\int_{[a,b]} f \leq \int_{[a,b]} g$.

Démonstration. Il suffit de considérer la positivité de $g - f$. □

Enfin, voyons la relation de Chasles.

Proposition 15.3.4 (Relation de Chasles). Soit $f : [a, c] \rightarrow \mathbb{R}$ une fonction continue par morceaux, alors pour tout $b \in [a, b]$,

$$\int_{[a,c]} f = \int_{[a,b]} f + \int_{[b,c]} f$$

Démonstration. Pour une fonction en escalier, l'égalité est évidente en découpant la somme associée. Pour une fonction quelconque, on remarque que f est une fonction en escalier inférieure à f sur $[a, c]$ si et seulement si elle est une fonction en escalier inférieure à f sur $[a, b]$ et sur $[a, c]$, le résultat en découle donc. □

Remarque. On définit une nouvelle notation pour les intégrales :

$$\int_a^b f(t)dt := \int_{[a,b]} f$$

dans le cas où $a \leq b$, et si $a > b$ alors

$$\int_a^b f(t)dt = \int_{[b,a]} f$$

Présentons les sommes de Riemann, permettant un calcul plus commode des intégrales.

Proposition 15.3.5 (Somme de Riemann). Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction continue. Alors

$$\int_{[a,b]} f = \lim_{n \rightarrow \infty} \frac{b-a}{n} \sum_{k=1}^n f\left(a + \frac{k(b-a)}{n}\right)$$

Démonstration. Soit $n \in \mathbb{N}$, notons $x_k = a + k \frac{b-a}{n}$. Tout d'abord, on remarque que, à $k \leq n$ fixés, l'égalité suivante est vraie :

$$\frac{b-a}{n} f\left(a + k \frac{b-a}{n}\right) = \int_{[x_k, x_{k+1}]} f\left(a + k \frac{b-a}{n}\right)$$

donc

$$\frac{b-a}{n} f\left(a + k \frac{b-a}{n}\right) - \int_{[x_k, x_{k+1}]} f = \int_{[x_k, x_{k+1}]} \left(f\left(a + k \frac{b-a}{n}\right) - f\right)$$

et, en sommant sur chaque k cela nous donne

$$\frac{b-a}{n} \sum_{k=1}^n f(x_k) - \int_{[a,b]} f = \sum_{k=1}^n \int_{[x_k, x_{k+1}]} (f(x_k) - f)$$

Nous voulons majorer la différence de droite, donc nous allons introduire la valeur $\omega_n = \sup\{|f(x) - f(y)|, x \in [a, b], y \in [a, b], |x - y| \leq \frac{b-a}{n}\}$. Comme f est continue, d'après le théorème de Heine, elle est uniformément continue. Remarquons que cela signifie que $\lim \omega_n = 0$. Or

$$\left| \frac{b-a}{n} \sum_{k=1}^n f(x_k) - \int_{[a,b]} f \right| \leq \sum_{k=1}^n \int_{[x_k, x_{k+1}]} \omega_n = \sum_{k=1}^n \frac{b-a}{n} \omega_n$$

donc

$$\lim \left| \frac{b-a}{n} \sum_{k=1}^n f\left(a + \frac{k(b-a)}{n}\right) - \int_{[a,b]} f \right| = 0$$

d'où le résultat. □

Enfin, montrons le théorème fondamental de l'analyse :

Théorème 15.1 (Fondamental de l'analyse). *Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction continue. Alors la fonction*

$$\begin{aligned} F : [a, b] &\longrightarrow \mathbb{R} \\ x &\longmapsto \int_{[a, x]} f \end{aligned}$$

est dérivable et sa dérivée est f .

Démonstration. Soit $\varepsilon > 0$ et $x \in [a, b]$, par continuité de f , on trouve $\delta > 0$ tel que $\forall y \in [a, b], |x - y| < \delta \implies |f(x) - f(y)| < \varepsilon$. Alors pour $h \in B(x, \delta)$, on en déduit que

$$\begin{aligned} |F(x) - F(x+h) - hf(x)| &= \left| \int_x^{x+h} f(t) dt - hf(x) \right| \\ &= \left| \int_x^{x+h} f(t) - f(x) dt \right| \\ &\leq h\varepsilon \\ \left| \frac{F(x) - F(x+h)}{h} - f(x) \right| &\leq \varepsilon \end{aligned}$$

On a donc montré que pour tout voisinage de $f(x)$, il existait un voisinage de x tel que $\frac{F(x) - F(x+h)}{h}$ appartenait à ce voisinage. Autrement dit :

$$\lim_{h \rightarrow 0} \frac{F(x) - F(x+h)}{h} = f(x)$$

et c'est la limite d'un taux d'accroissement, donc on en déduit que F est dérivable et que $F' = f$. \square

Une formulation équivalente est la suivante :

Corollaire 15.3.2. *La fonction F définie précédemment est l'unique primitive de f (i.e. fonction dont la dérivée est f) qui s'annule en a , et toute primitive de f est égale à F à une constante près.*

Démonstration. La fonction s'annule bien en a , et si G est une primitive de f , alors $G - F$ est une primitive de 0 puisque la dérivée est linéaire. On en déduit que $G - F$ est constante (puisque nous travaillons sur un intervalle) donc que $F = G + k$ avec $k \in \mathbb{R}$. Si $k \neq 0$ alors $F \neq G$ donc F est bien l'unique primitive de f s'annulant en a . \square

Remarque. Une autre définition équivalente est que si f est une fonction continue de $[a, b]$ dans \mathbb{R} et que F est une primitive de f , alors $\int_a^b f(t) dt = F(b) - F(a)$. On marque souvent $[F(x)]_a^b$ pour désigner $F(b) - F(a)$.

Exercice 15.2. Montrer cette équivalence.

15.4 Calcul d'intégrale

Nous présenterons ici deux méthodes importantes de calcul d'intégrales : l'intégration par parties et le changement de variable.

Proposition 15.4.1 (Intégration par parties). *Soient $u, v : [a, b] \rightarrow \mathbb{R}$ deux fonctions dérivables. Alors*

$$\int_a^b u'(t)v(t) dt = [uv]_a^b - \int_a^b u(t)v'(t) dt$$

Démonstration. On remarque d'abord que $(uv)' = u'v + uv'$ donc $u'v = (uv)' - uv'$. En passant aux intégrales, et comme l'intégrale d'une dérivée vaut la fonction elle-même, on en déduit

$$\int_a^b u'(t)v(t)dt = [uv]_a^b - \int_a^b u(t)v'(t)dt$$

□

Proposition 15.4.2 (Changement de variable). *Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction continue, $a, b \in \mathbb{R}$ et $\varphi : [\alpha, \beta] \rightarrow [a, b]$ dérivable et de dérivée continue, $\alpha, \beta \in \mathbb{R}$. Alors*

$$\int_{\alpha}^{\beta} f(\varphi(t))\varphi'(t)dt = \int_{\varphi(\alpha)}^{\varphi(\beta)} f(t)dt$$

Démonstration. Soit la fonction $g : x \mapsto \int_{\varphi(\alpha)}^x f(t)dt$. On va calculer la dérivée de $g \circ \varphi$:

$$(g \circ \varphi)' = \varphi' \times g' \circ \varphi$$

mais la dérivée de g est f , donc

$$(g \circ \varphi)' = \varphi' \times f \circ \varphi$$

ce qui signifie que $h : x \mapsto \int_{\alpha}^x (f \circ \varphi)(t)\varphi'(t)dt$ et g ont la même dérivée. De plus, les deux fonctions sont égales en α et valent 0, donc elles sont égales en tant que primitives d'une même fonction qui ont une valeur commune. Ainsi, en évaluant en β , on en déduit

$$\int_{\alpha}^{\beta} f(\varphi(t))\varphi'(t)dt = \int_{\varphi(\alpha)}^{\varphi(\beta)} f(t)dt$$

□

16 Généralités sur les fonctions

Ce dernier chapitre aura pour but de réinvestir les différents outils des chapitres précédents pour donner des propriétés basiques sur les fonctions usuelles. Nous introduirons le logarithme, l'exponentielle ainsi que les fonctions trigonométriques réciproques.

16.1 Logarithme

Le logarithme népérien (ou naturel) est une fonction dont l'intérêt premier est de transformer les produits en somme. Cela servait avant à faciliter des calculs, et sert aussi pour la notion d'échelle logarithmique, dont le principe est de représenter plusieurs ordres de grandeur en une seule échelle.

Définition 16.1 (Logarithme népérien). On appelle logarithme népérien, et on note \ln , la fonction

$$\begin{aligned} \ln : \mathbb{R}_+^* &\longrightarrow \mathbb{R} \\ x &\longmapsto \int_1^x \frac{1}{t} dt \end{aligned}$$

On déduit de cette définition que la fonction \ln est dérivable, et donc aussi continue. De plus sa dérivée étant strictement positive, cette fonction est strictement croissante. Enfin, $\ln(1) = 0$ puisque \ln est la primitive de la fonction inverse qui s'annule en 1.

Proposition 16.1.1. Soient $a, b \in \mathbb{R}_+^*$, alors

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

Démonstration. Développons le terme de gauche :

$$\begin{aligned} \ln(a \times b) &= \int_1^{a \times b} \frac{1}{t} dt \\ &= \int_{1/a}^b \frac{1}{t} dt \\ &= \int_{1/a}^1 \frac{1}{t} dt + \int_1^b \frac{1}{t} dt \\ &= \int_a^1 \frac{-1}{t^2} \times t dt + \int_1^b \frac{1}{t} dt \\ &= \int_1^a \frac{1}{t} dt + \int_1^b \frac{1}{t} dt \\ \ln(a \times b) &= \ln(a) + \ln(b) \end{aligned}$$

En effectuant le changement de variable $\varphi(t) = a \times t$ puis le changement de variable $\varphi(t) = \frac{1}{t}$ dans l'intégrale de gauche. \square

Proposition 16.1.2. Soit $a \in \mathbb{R}_+^*$, alors

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

Démonstration. Puisque $\ln\left(a \times \frac{1}{a}\right) = \ln(1) = 0$ et que $\ln\left(a \times \frac{1}{a}\right) = \ln(a) + \ln\left(\frac{1}{a}\right)$, on en déduit que

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

\square

Exercice 16.1. Soit $n \in \mathbb{Z}$ et $a \in \mathbb{R}_+^*$, montrer que $\ln(a^n) = n \ln(a)$. *Indication : on raisonnera par récurrence pour \mathbb{N} puis on utilisera le passage à l'inverse pour calculer $\ln(a^{-n})$.*

On définit de plus le logarithme en base quelconque :

Définition 16.2 (Logarithme en base quelconque). Soit $a \in \mathbb{R}_+^*$, on définit la fonction

$$\begin{aligned} \log_a : \mathbb{R}_+^* &\longrightarrow \mathbb{R} \\ x &\longmapsto \frac{\ln(x)}{\ln(a)} \end{aligned}$$

Proposition 16.1.3. On a

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

Démonstration. Pour prouver cela, nous allons procéder par minoration. D'abord, par décroissance de la fonction inverse, on sait que

$$\int_n^{n+1} \frac{1}{t} dt \geq \frac{1}{n+1}$$

donc cela signifie que

$$\int_1^n \frac{1}{t} dt \geq \sum_{k=1}^{n-1} \frac{1}{k}$$

Si l'on montre que la somme de droite, que l'on va noter H_n , diverge vers ∞ pour $n \rightarrow \infty$, on en déduira que l'intégrale de gauche diverge, et cela permettra de déduire que $\lim_{x \rightarrow +\infty} \ln(x) = +\infty$ car cela signifie que la fonction est croissante non majorée.

Montrons que la somme H_n diverge. Pour cela, on remarque que pour tout $p \in \mathbb{N}$, $i \in \{2^p, \dots, 2^{p+1}\}$, $\frac{1}{i} \geq 2^{p+1}$, donc

$$\sum_{i=2^p}^{2^{p+1}-1} \frac{1}{i} \geq \frac{1}{2}$$

ce qui signifie que $H_{2^n} \geq \frac{n}{2}$ mais la suite de droite diverge, donc H_{2^n} diverge. De plus, H_n est croissante, donc comme elle est non majorée, elle diverge. Donc $\lim H_n = +\infty$, donc

$$\boxed{\lim_{x \rightarrow +\infty} \ln(x) = +\infty}$$

□

Corollaire 16.1.1. On a aussi

$$\lim_{x \rightarrow 0} \ln(x) = -\infty$$

Démonstration. Cela se déduit directement d'un changement de variable en $x = \frac{1}{h}$ avec $h \rightarrow \infty$ et du fait que $\ln\left(\frac{1}{h}\right) = -\ln(h)$. □

Proposition 16.1.4 (Croissance comparée). On a

$$\lim_{x \rightarrow +\infty} \frac{\ln(x)}{x} = 0$$

Démonstration. Étudions la fonction $f : x \mapsto \frac{\ln(x)}{x}$. Elle est dérivable sur \mathbb{R}_+^* et sa dérivée est $f' : x \mapsto \frac{1 - \ln(x)}{x^2}$ qui est strictement négative pour x assez grand. On en déduit qu'au voisinage de $+\infty$, f est strictement décroissante. De plus, f est positive pour $x > 1$, donc on en déduit que f a une limite en $+\infty$. On pose l la limite de f . Comme $x \mapsto f(2x)$ a la même limite que f , on en déduit que $\lim_{x \rightarrow +\infty} \frac{\ln(2x)}{2x} = l$ or l'expression de gauche est égale à $\frac{1}{2} \frac{\ln(x)}{x} + \frac{1}{x}$ qui tend vers $\frac{l}{2}$, donc par unicité de la limite on en déduit que $\frac{l}{2} = l$ soit

$$\boxed{l = 0}$$

□

Corollaire 16.1.2. *Il s'en déduit*

$$\lim_{x \rightarrow 0} x \ln(x) = 0$$

Démonstration. Cette limite est équivalente à $\lim_{x \rightarrow \infty} \frac{1}{x} \ln\left(\frac{1}{x}\right) = \lim_{x \rightarrow \infty} -\frac{\ln(x)}{x}$ par changement de variable, d'où le résultat. \square

16.2 Exponentielle

La fonction exponentielle est définie en générale comme solution d'une équation différentielle. Ici, nous la construisons comme réciproque de la fonction logarithme népérien, et nous montrons qu'elle vérifie l'équation différentielle caractéristique de l'exponentielle.

Définition 16.3 (Exponentielle). On définit la fonction exponentielle

$$\begin{aligned} \exp : \mathbb{R} &\longrightarrow \mathbb{R}_+^* \\ x &\longmapsto \exp(x) \end{aligned}$$

comme la réciproque de la fonction \ln .

Démonstration. Par stricte croissance et étude des limites de \ln , \ln est bien bijective de \mathbb{R}_+^* dans \mathbb{R} , donc cette fonction admet une réciproque $\exp : \mathbb{R} \rightarrow \mathbb{R}_+^*$. Comme $x \mapsto \frac{1}{x}$ ne s'annule pas, la fonction \exp est dérivable sur \mathbb{R} (donc continue). \square

Proposition 16.2.1 (Équation différentielle). *La fonction \exp vérifie les propriétés suivantes :*

$$\begin{cases} \exp' &= \exp \\ \exp(0) &= 1 \end{cases}$$

De plus, elle est l'unique fonction qui vérifie ces propriétés.

Démonstration. En utilisant le théorème de dérivée d'une fonction réciproque, on sait que

$$\exp'(x) = \frac{1}{\ln'(\exp(x))}$$

mais $\ln'(t) = \frac{1}{t}$ donc l'expression est équivalente à

$$\boxed{\exp'(x) = \exp(x)}$$

Comme $\ln(1) = 0$ et \exp est la réciproque de \ln , il en vient que $\exp(\ln(1)) = 1$ donc

$$\boxed{\exp(0) = 1}$$

Si une autre fonction vérifiait ces propriétés, disons g , alors $\exp - g$ serait de dérivée nulle sur \mathbb{R} : elle serait donc constante. Comme g et \exp coïncident en 0, cela signifie que $g = \exp + 0$ donc que

$$\boxed{g = \exp}$$

\square

Proposition 16.2.2. *Soient a et b deux réels, alors*

$$\exp(a + b) = \exp(a) \times \exp(b)$$

Démonstration. Puisque \exp est la réciproque de \ln , on va écrire a et b sous forme d'images de logarithmes :

$$\begin{aligned} \exp(a) \times \exp(b) &= \exp(\ln(\exp(a) \times \exp(b))) \\ &= \exp(\ln(\exp(a)) + \ln(\exp(b))) \\ &= \exp(a + b) \end{aligned}$$

D'où le résultat. \square

Exercice 16.2. Montrer que pour tout $n \in \mathbb{N}$, $a \in \mathbb{R}$, $\exp(na) = (\exp(a))^n$.

Proposition 16.2.3. Pour tout $a \in \mathbb{R}$, $\exp(-a) = \frac{1}{\exp(a)}$

Démonstration. On utilise simplement le fait que $\exp(a + (-a)) = 1$ et $\exp(a + (-a)) = \exp(a) \times \exp(-a)$. \square

Définition 16.4 (Notation puissance). On note $e = \exp(1)$. Comme pour tout $n \in \mathbb{N}$, $\exp(n) = e^n$, on étend cette notation pour tout $x \in \mathbb{R}$: $\exp(x) = e^x$.

Définition 16.5 (Exponentielle en base quelconque). Soit $a \in \mathbb{R}_+^*$, on définit l'exponentielle en base a , notée a^x en un $x \in \mathbb{R}$ donné, par

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

Proposition 16.2.4 (Limites de l'exponentielle). L'exponentielle a comme limites

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

Démonstration. Comme l'exponentielle est la réciproque de \ln et que $\lim_{x \rightarrow 0} \ln(x) = -\infty$ et $\lim_{x \rightarrow +\infty} \ln(x) = +\infty$, on déduit les limites voulues en remplaçant x tendant vers $-\infty$ (respectivement $+\infty$) par $\ln(x)$ pour x tendant vers 0 (respectivement $+\infty$) et la limite à calculer est alors simplement la limite de la fonction identité. \square

Proposition 16.2.5 (Croissance comparée). On a

$$\lim_{x \rightarrow \infty} \frac{x}{\exp(x)} = 0$$

Démonstration. Là encore, on effectue un changement de variable pour changer x en $\ln(x)$, où $x \rightarrow +\infty$. On a alors la limite de $\frac{\ln(x)}{x}$ que nous avons déjà calculé au préalable. \square

Corollaire 16.2.1. On a aussi

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

Démonstration. Cela se déduit encore par un changement de variable. \square

16.3 Cosinus et sinus

Les fonctions \cos et \sin sont importantes en analyse aussi, nous allons donner leurs dérivées en admettant qu'elles sont continues.

Proposition 16.3.1. La fonction \sin est dérivable et sa dérivée est \cos .

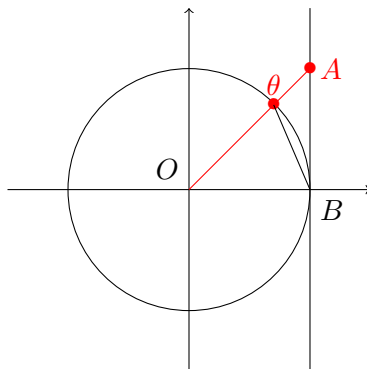


FIGURE 25 – Figure pour encadrer le sinus

Démonstration. En utilisant la figure précédente, on peut calculer l'aire de $(OB\theta)$, l'aire du secteur angulaire S entre $[OB]$ et $[O\theta]$ et celle de (OBA) , qui par inclusions successives sont l'une inférieure à la suivante. La hauteur $[H\theta]$ mesure $\sin(\theta)$ et la base $[OB]$ mesure 1, donc l'aire de $(OB\theta)$ vaut $\mathcal{A}_{OB\theta} = \frac{1}{2} \sin(\theta)$. La longueur du secteur angulaire est directement $\mathcal{A}_S = \frac{1}{2} \theta$. En utilisant le théorème de Thalès, la longueur de $[AB]$ vaut $\tan(\theta)$ donc l'aire de (OBA) vaut $\mathcal{A}_{OBA} = \frac{1}{2} \tan(\theta)$. On en déduit l'encadrement suivante :

$$\frac{1}{2} \sin(\theta) \leq \frac{1}{2} \theta \leq \frac{1}{2} \tan(\theta)$$

qui, si l'on divise par $\frac{1}{2} \sin(\theta)$ et qu'on applique la fonction inverse (strictement décroissante), donne

$$\cos(\theta) \leq \frac{\sin(\theta)}{\theta} \leq 1$$

et on sait que $\lim_{\theta \rightarrow 0} \cos(\theta) = 1$ par continuité. Par encadrement, on en déduit donc que

$$\boxed{\lim_{\theta \rightarrow 0^+} \frac{\sin(\theta)}{\theta} = 1}$$

La limite n'a été prouvée que pour les valeurs supérieures, comme θ était positif, mais par imparité de \sin on en déduit que la limite à gauche est la même, donc \sin est dérivable en 0 et sa dérivée vaut 1.

Nous allons maintenant en déduire la dérivabilité sur \mathbb{R} de \sin en calculant son taux d'accroissement entre un $a \in \mathbb{R}$ et $a + h$:

$$\begin{aligned} \frac{\sin(a+h) - \sin(a)}{h} &= \frac{\sin(a) \cos(h) + \sin(h) \cos(a) - \sin(a)}{h} \\ &= \frac{\sin(a)(\cos(h) - 1)}{h} + \frac{\sin(h)}{h} \cos(a) \\ &= \sin(a) \frac{-2 \sin^2\left(\frac{h}{2}\right)}{h} + \frac{\sin(h)}{h} \cos(a) \\ &= -\sin(a) \sin\left(\frac{h}{2}\right) \frac{\sin\left(\frac{h}{2}\right)}{\frac{h}{2}} + \frac{\sin(h)}{h} \cos(a) \end{aligned}$$

Le terme de gauche tend vers 0 car le sinus tend vers 0 et le quotient avec le sinus tend vers 1 d'après la dérivée de \sin en 0. Ainsi il ne reste que le terme de droite, qui tend vers $\cos(a)$. Donc

$$\boxed{\forall a \in \mathbb{R}, \sin'(a) = \cos(a)}$$

□

Proposition 16.3.2. *La fonction \cos est dérivable, de dérivée $-\sin$.*

Démonstration. On sait que $\cos(x) = \sin\left(\frac{\pi}{2} - x\right)$ donc en dérivant, cela nous donne $\cos'(x) = -\cos\left(\frac{\pi}{2} - x\right) = -\sin(x)$, donc $\boxed{\cos' = -\sin.}$ □