Cours de mathématiques en MPSI

Eliott Perrière-Larraux

21 février 2025

Introduction

Hey ! Bienvenue à toi cher lecteur : si tu lis ceci, tu es probablement un élève de Ginette ou un petit chanceux qui a reçu ce document de mes propres mains. Je vais expliquer ici en quelques mots (voire un peu plus que "quelques") les motivations et spécificités de ce document, ainsi que les consignes d'utilisation. Si cela ne t'intéresse pas, passe directement à la suite, mais sache que vu le temps que j'ai passé à faire ce travail, j'apprécierais que tu écoutes ce que j'ai à dire.

Tout d'abord, j'ai eu l'idée de commencer ce document durant l'été entre ma sup et ma spé, face au volume colossal qu'occupe la totalité de mes cours de mathématiques. Comme je n'habite pas en région parisienne, cela m'arrangerait de pouvoir éviter de me trimbaler tous mes cours de maths de sup pendant la spé, c'est pourquoi j'ai compilé ici l'entièreté de mes cours de mathématiques de sup. Oui, tu as bien lu, l'entièreté. Le moindre résultat écrit au tableau durant les cours du matin avec M. Morlot est recensé ici.

Ce document représente au moins 200 heures de travail, réparties sur 5 semaines. J'ai tapé chaque chapitre à la main, en LaTeX (je crois que je suis bilingue LaTeX maintenant). Chaque chapitre contient toutes les définitions, tous les théorèmes (et variantes), ainsi que tous les exemples et toutes les remarques que j'ai jugés utiles et pertinents (la plupart en fait). J'y ai aussi mentionné les exercices donnés par M. Morlot en cours, avec des solutions ou des pistes de résolution pour la plupart. Les résultats hors-programme sont repérés par le sigle HP : ils sont alors toujours démontrés car tout résultat hors-programme doit être redémontré à l'écrit, et doit pouvoir l'être à la demande de l'examinateur à l'oral. Les parties entières du cours qui sont hors-programme sont aussi signalées par le sigle HP dans leur titre. Les autres démonstrations des résultats de cours ont été retirées car elles sont longues, et je souhaitais finir avant la fin des vacances d'été (ainsi que préserver le peu de santé mentale qu'il me reste).

Je vous demande par avance d'excuser toute faute de français (qui risque fort d'être une faute de frappe) ou toute faute relative aux mathématiques : je n'ai pas relu l'ensemble du document, j'ai d'autres bouquins de 400 pages à lire pour la rentrée. Si vous voyez une faute quelconque (mathématique ou linguistique), merci de m'envoyer un message avec la référence de la faute (chapitre, proposition par exemple) afin que je puisse la corriger.

Pour réaliser ce document, j'ai utilisé les cours que j'ai pris en note durant les cours dispensés par M. Morlot ainsi que les différents polycopiés qu'il nous a distribués (merci pour les PAPL). J'ai aussi utilisé les polycopiés de cours qu'il m'a donnés en raison de ma maladie, ce qui revient un petit peu au même que les cours pris en notes. Qu'il soit remercié ici pour tout cela. De plus,

le chapitre "Compléments sur les anneaux commutatifs" est issu d'un document éponyme distribué par M. Lafitte aux Taupins, et qui a un jour atterri sur mon bureau par miracle (c'est vrai). Un grand merci à lui aussi pour ce document très intéressant. Je sais que M. Morlot n'apprécie pas que ses documents tournent (en tout cas c'est le cas pour ses polycopiés de cours), donc je vous demande de ne pas faire tourner ce document en-dehors de Ginette, et si vous n'êtes pas de Ginette, de ne le partager à personne. Je sais que ce document n'est pas à proprement parler son cours, mais il a tellement fait pour moi que je lui dois bien ça.

Update : le poly de Lafitte a été tapé, j'y ai rajouté quelques résultats sur les anneaux noethériens, les radicaux d'idéaux, les idéaux maximaux et les idéaux premiers. Absolument tous les PAPL et tous les cours ont été tapés. J'ai aussi rajouté un paragraphe de compléments dans le chapitre "Polynômes" qui expose brièvement les notions de contenu et de polynôme primitif, qui aboutit sur une démonstration élégante du critère d'Eisenstein. A terme, je ne pense pas rajouter quoi que ce soit sur ce document, il contient déjà bien plus que le simple programme de sup, et est déjà assez long comme ça.

Si tu utilises, comme je compte le faire personnellement, ce document pour réviser et pour ne pas avoir à traîner tout ton cours de sup, profite bien de ce document, je serai ravi s'il peut t'être utile.

Avant de vous laisser, je voudrais saluer et remercier tous mes amis de Ginette qui ont rendu cette première année magique, malgré les aspérités. En particulier, un grand merci à mes cos Théa, Célia, Alexis et Nicolas, ainsi qu'à Angélique et Alexandre que j'adore et qui m'ont permis de décompresser un peu entre deux séquences d'écriture de ce document, je vous embrasse ! Et oui j'ai ptêtre un peu pété les plombs en écrivant toute cette introduction !

Bonne fin de vacances, ou bonne année de spé à tous. Des bisous.

EPL

Table des matières

Chapitre 1

Logique et raisonnements

La logique, bien que son développement formel soit récent dans l'histoire des mathématiques (un siècle environ), est aujourd'hui perçue comme la base sur laquelle se fondent celles-ci. Notre objectif, modeste, est de présenter ici les principaux outils dont nous aurons besoin pour cette année. Ils nous aideront à produire une rédaction claire et débarrassée de toute ambiguïté, selon les canons actuels de la rigueur.

1 De l'importance de bien rédiger

Avant de passer à des considérations plus théoriques, indiquons comment bien rédiger une démonstration mathématique. On n'hésitera pas à relire les exemples qui suivent de nombreuses fois : c'est par l'imitation qui s'acquiert la technique ! Pour ceux qui ne connaîtraient pas l'usage du mot "ssi" ou des symboles $\forall$,$\exists$, $\Longrightarrow$, $\Longleftrightarrow$, tout est expliqué dans le deuxième paragraphe.

Voici quelques grands principes qu'il ne faut jamais perdre de vue lorsque vous rédigez. Pourquoi ? Tout d'abord, parce que la rigueur du raisonnement est à ce prix ! Comme toute langue étrangère, les mathématiques requièrent une pratique régulière pour être correctement écrites. Et au-delà de cela, la rigueur ainsi acquise vous sera utile dans bien d'autres domaines. En second lieu, n'oubliez pas que le jour des concours, le correcteur aura à lire beaucoup de copies en un temps très court. il est donc essentiel que vous sachiez dire les choses clairement et simplement. Dans les épreuves de mathématiques, vous rencontrerez plusieurs types de questions, que nous listons maintenant.

1.1 Questions du type "pour tout"

Premier type : une question commençant par "montrer que pour tout réel x, etc.". Dans ce cas, il importe de commencer la rédaction par **soit** $x \in \mathbb{R}$ ou **soit** x **un réel**. On dit qu'on a **présenté** la variable x.

Exemple 1.1 (Montrer que $\forall x \in \mathbb{R},\ x^2 \geqslant$)**.** Soit $x \in \mathbb{R}$. Si $x \geqslant 0$, alors on a bien $x \times x \geqslant 0$. Et si $x \leqslant 0$, on a aussi $x \times x \geqslant 0$. Donc dans tous les cas, le résultat est prouvé.

Exemple 1.2 (Montrer que $\forall \theta \in \mathbb{R},\ e^{i\theta} \neq 0$)**.** Soit $\theta \in \mathbb{R}$. Alors $e^{i\theta} \times e^{-i\theta} = e^0 = 1$. En particulier, on a nécessairement $e^{i\theta} \neq 0$, sans quoi on aurait $e^{i\theta} \times e^{-i\theta} = 1 = 0$.

De manière générale, **il est interdit de calculer avec une variable avant de l'avoir présentée**. Par exemple, si on demande de trouver tous les réels dont le carré vaut 1, on ne peut pas écrire directement "supposons que $x^2 = 1$". En effet, dans ce cas, le correcteur serait en droit de demander "qui est x ?". Il faut d'abord écrire "soit x un réel".

Remarque 1.1 (Portée du quantificateur "pour tout")**.** Attention à la chose suivante : si une phrase commence par "pour tout x" ou "$\forall x$", **la portée de la variable x s'arrête à la fin de la phrase**. Autrement dit, une rédaction du type

$$\text{Pour tout } x \in \mathbb{R},\ x^2 \geqslant 0. \text{ Donc } x^2 + 1 \geqslant 0.$$

est incorrect en toute rigueur. a la place, il faudrait écrire

$$\text{Pour tout } x \in \mathbb{R},\ x^2 \geqslant 0. \text{ Donc pour tout } x \in \mathbb{R},\ x^2 + 1 \geqslant 0.$$

Comme c'est un peu pénible, si on prévoit d'utiliser une même variable x sur plusieurs phrases, il vaut mieux la présenter définitivement avec "soit x". Dans ce cas, la portée s'arrête à la fin de la question (ou au prochain "soit x", qui écrase en quelque sorte le précédent). Ainsi, on aurait aussi pu écrire

$$\text{Soit } x \in \mathbb{R}. \text{ On a } x^2 \geqslant 0. \text{ Donc } x^2 + 1 \geqslant 0.$$

1.2 Questions d'existence

Supposons qu'on nous demande de montrer une **existence**. Dans ce cas, il est demandé d'**exhiber un candidat**, puis de **démontrer qu'il convient**. Souvent, il vaut mieux commencer par chercher le candidat au brouillon.

1.3 Questions d'unicité

Supposons qu'on nous demande de montrer une **unicité**. Dans ce cas, on se donne deux candidats, et et on montre qu'ils sont égaux.

1.4 Équations, inéquations

Voici la méthode de résolution.

- Si ce n'est pas immédiate, on calcule le domaine de définition. Ne pas oublier de présenter l'inconnue ! Par exemple, supposons qu'elle s'appelle x. Si on nous demande de résoudre dans $\mathbb{R}$, on écrit : "Soit $x \in \mathbb{R}$. L'équation est définie en x ssi, etc.". Appelons $\mathcal{D}$ le domaine obtenu.
- Pour la résolution à proprement parler, on écrit d'abord "Soit $x \in \mathcal{D}$", puis on cherche l'ensemble exact des valeurs de x qui conviennent. Il faut être certain que les x donnés sont solution, mais aussi qu'on n'a oublié aucune solution ! Autrement dit, il s'agit de prouver une **équivalence** logique : les x donnés sont tous solution, et **réciproquement**, toute solution est une des valeurs données.
  - Soit on tente une résolution par équivalences successives en partant de l'équation (ou de l'inéquation). Dans ce cas, on écrit : "x est solution ssi ...".

- Soit on traite séparément chacune des deux implications : c'est très courant car il est rare de pouvoir mener une chaîne d'équivalences à son terme.
  * D'abord, on écrit "Supposons que x est solution", et on en déduit un ensemble de valeurs potentielles de x : c'est la condition **nécessaire**.
  * Puis, réciproquement, on n'oublie pas de vérifier si chacune de ces valeurs potentielles est solution ou non : c'est la condition **suffisante**. On obtient alors les solutions définitives.

2 Propositions logiques

Définition 1.1 (Proposition)**.** Une **proposition** est une phrase qui est soit vraie soit fausse.

Exemple 1.3. Voici des propositions (vraies ou fausses) :

- Paris est la capitale de la France.
- Lyon est la capitale de la France.
- Les Sigmas sont les meilleurs.

Définition 1.2 (Équivalence logique)**.** Deux propositions p et q sont dites **logiquement équivalentes** lorsqu'elles ont même table de vérité. On note alors $p \equiv q$.

Définition 1.3 (Conjonction, disjonction)**.** Soit p et q deux propositions. A partir de p et q, on définit les propositions "p et q" et "p ou q", appelées respectivement **conjonction** et **disjonction**, et notées respectivement $p \wedge q$ et $p \vee q$, par la table de vérité suivante :

p	q	$p \wedge q$	$p \vee q$
V	V	V	V
V	F	F	V
F	V	F	V
F	F	F	F

Définition 1.4 (Négation)**.** La **négation** d'une proposition p, notée "$\neg p$" ou "non p" est définie par la table de vérité ci-contre :

p	$\neg p$
V	F
F	V

On a donc $\neg(\neg p) \equiv p$.

Définition 1.5 (Quantificateurs)**.** Le symbole **"pour tout"**, dit **quantificateur universel**, s'écrit $\forall$. Le symbole **"il existe"**, dit **quantificateur existentiel**, s'écrit $\exists$. Une **phrase quantifiée** s'écrit de la façon suivante :

- D'abord l'un de ces deux symboles, appelés **quantificateurs** :
- Puis une relation d'appartenance d'un élément à un ensemble ;
- Une virgule (bien noter qu'après le symbole $\exists$, la virgule se lit "tel que") ;
- Une autre phrase quantifiée.

Ainsi, une phrase quantifiée peut se voir comme la succession de quantificateurs séparés par des virgules, et d'une affirmation finale.

Remarque 1.2. Il existe un troisième quantificateur, qui est $\exists!$ ("il existe un unique"), mais il est moins utilisé. De toute façon, on peut le reconstituer à partir des deux premiers. Par exemple, "$\exists! x \in X,\ p(x)$" pourra se réécrire :

$$[\exists x \in X,\ p(x)] \wedge \left[\forall (x,y) \in X^2,\ p(x) \wedge p(y) \implies x = y\right]$$

Théorème 1.1 (Négations de "et" et "ou")**.** *Soit p et q deux propositions. On a $\neg(p \wedge q) \equiv (\neg p) \vee (\neg q)$ et $\neg(p \vee q) \equiv (\neg p) \wedge (\neg q)$.*

Démonstration. Il suffit de dresser les tables de vérité. □

Proposition 1.1 (Associativité)**.** *Soit p, q et r trois propositions. On a :*

- *Associativité du "ou" : $(p \vee q) \vee r \equiv p \vee (q \vee r)$*
- *Associativité du "et" : $(p \wedge q) \wedge r \equiv p \wedge (q \wedge r)$*

Proposition 1.2 (Commutativité)**.** *Soit p et q deux propositions. On a :*

- *Commutativité du "ou" : $p \vee q \equiv q \vee p$*
- *Commutativité du "et" : $p \wedge q \equiv q \wedge p$*

Proposition 1.3 (Distributivité)**.** *Soit p, q et r trois propositions. On a :*

- *Distributivité du "ou" sur le "et" : $p \vee (q \wedge r) \equiv (p \vee q) \wedge (p \vee r)$*
- *Distributivité du "et" sur le "ou" : $p \wedge (q \vee r) \equiv (p \wedge q) \vee (p \wedge r)$*

Démonstration. Pour les trois propositions précédentes, il suffit de dresser les tables de vérité. □

Remarque 1.3 (Portée du quantificateur $\exists$)**.** La portée du quantificateur $\exists$ est aussi longue que celle de $\forall$. Si on veut réutiliser une variable énoncée après un quantificateur $\exists$, il faut écrire **"Fixons un tel ..."**. On peut alors réutiliser l'objet fixé, muni de ses propriétés supposées dans la phrase quantifiées.

Théorème 1.2. *Pour écrire la négation d'une phrase quantifiée :*

- *On inverse les quantificateurs : les $\forall$ deviennent des $\exists$, et les $\exists$ deviennent des $\forall$;*
- *Puis on nie l'affirmation finale.*

Démonstration. On le montre facilement par récurrence sur le nombre de quantificateurs. □

Définition 1.6 (Implication)**.** Soit p et q deux propositions. L'**implication** $p \implies q$ est définie par la table de vérité exposée ci-dessous :

p	q	$p \implies q$
V	V	V
V	F	F
F	V	V
F	F	V

Dans un énoncé, cela correspond à un raisonnement du type "**si** p**, alors** q".

Proposition 1.4. *Soit p et q deux propositions. On a $p \implies q \equiv (\neg p) \vee q$.*

Démonstration. Il suffit de comparer les tables de vérité. □

Remarque 1.4. Attention au fait que lorsque p est fausse, $p \implies q$ est vraie. On a coutume de dire que **le faux implique n'importe quoi**. Cela dit, ce problème est relativement marginal. Il peut cependant survenir dans des initialisations de récurrences.

Remarque 1.5 ("Donc" $\neq \implies$)**.** On ne peut pas utiliser $\implies$ comme une abréviation de "donc", ce sont deux choses différentes ! En effet, un "p donc q" correspond à $[p \wedge (p \implies q)]$, qui a même valeur de vérité que $p \wedge q$. Un "donc" par de quelque chose de vrai pour arriver à quelque chose de vrai, alors que $p \implies q$ ne suppose pas la véracité de p.

En somme, **le symbole $\implies$ est exclu dans une démonstration rédigée. Son usage n'est autorisé que dans une phrase quantifiée ou dans la résolution formelle d'une équation**.

Corollaire 1.1. *Soit p et q deux propositions logiques. On a :*

$$\neg(p \implies q) \equiv p \wedge (\neg q)$$

Démonstration. Passer à la négation la propriété précédente. □

Définition 1.7 (Contraposée)**.** Soit p et q deux propositions logiques. L'implication **contraposée** de $p \implies q$ est l'implication $(\neg q) \implies (\neg p)$.

Proposition 1.5 (Équivalence logique entre une implication et sa contraposée)**.** *Soit p et q deux propositions logiques. On a :*

$$p \implies q \equiv (\neg q) \implies (\neg p)$$

Démonstration. On a $(\neg q) \implies (\neg p) \equiv (\neg(\neg q)) \vee (\neg p) \equiv q \vee (\neg p) \equiv (\neg p) \vee q \equiv p \implies q$. □

Théorème 1.3. *Soit p et q deux propositions logiques. Pour prouver $p \implies q$, il y a trois manières.*

- *La **manière directe** : on part de p pour arriver à q. Dans ce cas, la bonne rédaction est "Supposons p, etc." puis plus loin "Donc q".*
- *La **contraposée** : on part de $\neg q$ pour arriver à $\neg p$.*
- *Le **raisonnement par l'absurde** : on suppose que p est vraie, puis on suppose que q est fausse pour arriver à une contradiction. Dans ce cas, après avoir supposé que p est vraie, la bonne rédaction est : "Supposons que q est fausse. Alors, comme on a p, cela impliquerait que, etc." puis plus loin "Contradiction. Donc q est vraie."*

Démonstration. Pour le premier point, il est inutile de considérer le cas où p est fausse, puisque le faux implique n'importe quoi. Il suffit donc de prouver que si p est vraie alors q est vraie. Pour la contraposée, on utilise la propriété précédente. Ppur le raisonnement par l'absurde, on montre que $p \wedge (\neg q)$ est fausse. C'est donc que sa négation, $(\neg p) \vee q$ est vraie, soit encore que $p \implies q$ est vraie. □

Définition 1.8 (Équivalence)**.** Soit p et q deux propositions logiques. Une **équivalence** entre p et q, notée $p \iff q$, est définie par la table ci-contre :

p	q	$p \iff q$
V	V	V
V	F	F
F	V	F
F	F	V

Cela signifie que p et q ont même valeur de vérité. Dans un énoncé, cela se traduira par une phrase du type "p si, et seulement si, q" ou en abrégé "p ssi q". On peut remarquer qu'on a :

$$p \iff q \equiv (p \implies q) \wedge (q \implies p)$$

Théorème 1.4. *Soit p et q deux propositions. Pour prouver $p \iff q$, il y a deux manières.*

- *Si on se sent suffisamment sûr de soi, la **manière directe** : on part de p, et par équivalences successives, on essaye d'arriver à q. C'est une méthode qui ne marche pas à tous les coups.*
- *Sinon, et c'est de loin le cas le plus fréquent, on démontre successivement $p \implies q$ et $q \implies p$. Dans la rédaction, au moment-clé où on "inverse le sens de la marche" et on passe à $q \implies p$, il est très important que le correcteur lise le mot **"réciproquement"**. A la rigueur, on pourra le remplacer par un "d'une part/d'autre part".*

Remarque 1.6. Pour montrer qu'une équivalence est fausse, il suffit de prouver qu'une des deux implications est fausse.

Proposition 1.6. *Soit p et q deux propositions logiques. On a :*

$$p \iff q \equiv (\neg p) \iff (\neg q)$$

Démonstration. Il suffit de passer chacune des deux implications à la contraposée. □

Définition 1.9 (Résolution d'une équation/inéquation)**. Résoudre** une équation dans un ensemble E consiste à trouver successivement :

- l'ensemble D des $x \in E$ en lesquels cette équation soit définie ;
- puis l'ensemble des $x \in D$ qui vérifient l'équation.

De même avec une inéquation. Autrement dit, à l'issue de la résolution, on doit être capable d'écrire

$$\forall x \in E,\ x \text{ solution ssi } x \in S$$

où S est un sous-ensemble de D (lui-même sous-ensemble de E). S est appelé **l'ensemble des solutions**.

Définition 1.10 (Raisonnement par analyse-synthèse)**.** A chaque fois qu'un énoncé commencera par "trouver tous les $x \in E$ tels que ...", il est recommandé de travailler par **analyse-synthèse**. Concrètement, cela consiste à réfléchir en deux temps.

- **Analyse** : si $x \in E$ convient, alors **nécessairement**, etc. Jusqu'à ce qu'on arrive à un ensemble de valeurs possibles pour x, qu'on appelle l'ensemble des "candidats". On travaille donc par **implication successives**. Ne pas chercher à écrire des équivalences, cela est **inutile et dangereux**.
- **Synthèse** : **réciproquement**, on examine un à un tous les candidats trouvés, on garde ceux qui sont effectivement solution du problème.

Autrement dit, l'analyse consiste à restreindre l'ensemble possible des solutions, de sorte qu'il ne reste plus qu'un nombre limité de cas à examiner durant la synthèse.

Démonstration. Formalisons un peu les choses.

- Pendant l'analyse, on détermine un ensemble $F \subset E$ tel que toutes les solutions éventuelles soient dans F. On a donc démontré la chose suivante :

$$\forall x \in E,\ x \text{ solution } \implies x \in F$$

- Puis, pendant la synthèse, on détermine un ensemble $S \subset F$ tel que

$$\forall x \in F,\ x \text{ solution } \iff x \in S$$

On vérifie alors que la conjonction des deux énoncés précédents a même valeur de vérité que

$$\forall x \in E,\ x \text{ solution } \iff x \in S$$

Autrement dit, S est bien l'ensemble des solutions. □

Remarque 1.7. Dans les cas extrêmes (mais pas si rares !), il ne reste déjà plus qu'un seul candidat à la fin de l'analyse. Dans ce cas, cette première phase montre l'**unicité** de la solution. Ensuite la synthèse permet de montrer

- soit l'**existence** d'une solution (si le candidat trouvé répond au problème) ;
- soit qu'il n'y a aucune solution (si le candidat trouvé ne répond pas au problème).

Remarque 1.8. Il arrive que la phase d'analyse produise des conditions nécessaires si restrictives que réciproquement, toutes les valeurs trouvées soient des solutions. Dans ce cas, la synthèse est très rapide, elle se borne à une simple vérification. De manière équivalente, on peut direz qu'on a procédé par **condition nécessaire / condition suffisante** : la phase d'analyse nous a amené à déterminer une condition nécessaire, qui s'est également révélée suffisante pendant la phase de synthèse.

Définition 1.11 (SPG)**.** Enfin, citons la technique du "sans perte de généralité" ou "SPG". il s'agit, dans une démonstration dont les notations sont un peu lourdes, de se débarrasser de cette complexité en se restreignant à un sous-cas qui couvre en fait tous les cas par analogie. Cette technique fonctionne bien à chaque fois que le problème possède une certaine symétrie. Attention, la notation "SPG" est à éviter sur une copie, elle n'est pas standard.

3 Raisonnement par récurrence

Cette partie est d'une importance capitale pour acquérir de bons réflexes à l'écrit. La validité du raisonnement par récurrence sera établie au cours de la construction de $\mathbb{N}$: ici, il s'agit seulement d'apprendre les différents types de récurrences et de savoir les rédiger proprement.

3.1 Récurrence simple

Théorème 1.5 (Principe de récurrence simple)**.** *Soit $(H_n)_{n\in\mathbb{N}}$ une suite d'assertions définies pour tout $n \in \mathbb{N}$. On suppose que les deux points suivants sont vrais :*

- *Initialisation : H_0 est vraie*
- *Hérédité : $\forall n \in \mathbb{N},\ H_n \implies H_{n+1}$*

Alors pour tout $n \in \mathbb{N}$, H_n est vraie.

Remarque 1.9. Souvent, la phase d'initialisation est courte et facile. Pour l'hérédité, il peut être utile de se faire la main sur de petites valeurs de n afin d'appréhender le fonctionnement général.

Théorème 1.6 (Rédaction d'une récurrence)**.** *Pour bien rédiger un raisonnement par récurrence :*

- *Si ce n'est pas clair, on précise "soit pour tout $n \in \mathbb{N}$, la proposition H_n : "..." ".*
- *On écrit "Initialisation :" et on prouve H_0.*
- *Ensuite on écrit "Hérédité : Soit $n \in \mathbb{N}$ tel que H_n soit vraie." puis on prouve H_{n+1}. On peut aussi écrire "Soit $n > 0$ et tel que H_{n-1} puis on prouve H_n.*
- *A l'endroit de la preuve où on utilise le fait que H_n est vraie, on écrit "par hypothèse de récurrence". A la fin de la récurrence, on écrit "ce qui achève la récurrence".*

3.2 Quelques variantes

Proposition 1.7 (Récurrence à partir d'un certain rang)**.** *Soit $n_0 \in \mathbb{Z}$ et soit $(P_n)_{n \geqslant n_0}$ une suite d'assertions. On suppose que les deux points suivants sont vrais :*

- *Initialisation : P_{n_0} est vraie*
- *Hérédité : $\forall n \geqslant n_0,\ P_n \implies P_{n+1}$*

Alors pour tout $n \geqslant n_0$, P_n est vraie.

Démonstration. Il suffit de poser pour tout $n \in Q_n = P_{n+n_0}$ et d'appliquer le principe de récurrence simple à $(Q_n)_{n\in\mathbb{N}}$. Ensuite, pour tout $n \geqslant n_0$, $n - n_0 \in \mathbb{N}$ donc $P_n = Q_{n-n_0}$ est vraie. □

Remarque 1.10. Cette technique peut se révéler utile lorsqu'on a une suite de propriétés à prouver pour $n \geqslant 0$ mais que l'hérédité ne fonctionne qu'à partir d'un certain indice $n_0 > 0$. On traite alors les premières valeurs de n (de 0 à $n_0 - 1$) séparément, puis on initialise la récurrence à n_0.

Proposition 1.8 (Récurrence à deux pas)**.** *Soit $n_0 \in \mathbb{Z}$ et soit $(P_n)_{n \geqslant n_0}$ une suite d'assertions. On suppose que les deux points suivants sont vrais :*

- *Initialisation : P_{n_0} et P_{n_0+1} sont vraies*
- *Hérédité : $\forall n \geqslant n_0,\ (P_n \wedge P_{n+1}) \implies P_{n+2}$*

Alors pour tout $n \geqslant n_0$, P_n est vraie.

Démonstration. Il suffit d'appliquer le principe de récurrence à partir d'un certain rang à la suite d'assertions définies par $Q_n = (P_n \wedge P_{n+1})$. Comme Q_n implique P_n, on obtient le résultat. □

Remarque 1.11. Sur le même principe, on peut effectuer des récurrences à trois pas, à quatre pas... L'important est de bien le préciser sur sa copie.

Proposition 1.9 (Récurrence forte)**.** *Soit $n_0 \in \mathbb{Z}$ et soit $(P_n)_{n \geqslant n_0}$ une suite d'assertions. On suppose que les deux points suivants sont vrais :*

- *Initialisation : P_{n_0} est vraie*
- *Hérédité : $\forall n \geqslant n_0,\ (\forall k \in [\![n_0, n]\!],\ P_k \text{ est vraie}) \implies P_{n+1}$*

Alors pour tout $n \geqslant n_0$, P_n est vraie. Ainsi, une récurrence forte consiste dans sa phase d'hérédité à supposer que tous les rangs $\leqslant n$ sont vrais (et non pas seulement n) pour en déduire P_{n+1}.

Démonstration. Cette fois, il suffit d'appliquer le principe de récurrence à partir d'un certain rang à la suite d'assertions définies par $Q_n = (\forall k \in [\![n_0, n]\!],\ P_k)$. Comme Q_n implique P_n, on obtient le résultat. □

Proposition 1.10 (Récurrence finie)**.** *Soit $(P_n)_{n_0 \leqslant n \leqslant n_1}$ une suite d'assertion avec $n_0 \leqslant n_1$. On suppose que les deux points suivants sont vrais :*

- *Initialisation : P_{n_0} est vraie*
- *Hérédité : $\forall n \in [\![n_0, n_1 - 1]\!],\ P_n \implies P_{n+1}$*

Alors pour tout $n \in [\![n_0, n_1]\!]$, P_n est vraie

Démonstration. Cette fois, il suffit d'appliquer le principe de récurrence à partir d'un certain rang à la suite d'assertions définies par

$$Q_n = \begin{cases} P_n & \text{si } n \in [\![n_0, n_1]\!] \\ \text{vrai} & \text{si } n > n_1 \end{cases}$$

Q_n est alors vraie quel que soit $n \in \mathbb{N}$, donc en particulier pour $n \in [\![n_0, n_1]\!]$. □

Proposition 1.11 (Récurrence descendante)**.** *Soit $(P_n)_{n_0 \leqslant n \leqslant n_1}$ une suite d'assertion avec $n_0 \leqslant n_1$. On suppose que les deux points suivants sont vrais :*

- *Initialisation : P_{n_1} est vraie*
- *Hérédité : $\forall n \in [\![n_0 + 1, n_1]\!],\ P_{n+1} \implies P_n$*

Alors pour tout $n \in [\![n_0, n_1]\!]$, P_n est vraie

Démonstration. Il suffit d'appliquer le principe de récurrence finie avec $Q_n = P_{n_1+n_0-n}$. □

3.3 Suites définies par récurrence

Une fois que l'on maîtrise le concept de définition par récurrence, on peut établir des **définitions** par récurrence.

Définition 1.12. Définir une suite $(u_n)_{n\in\mathbb{N}}$ par récurrence consiste à donner son premier terme u_0, puis à définir une relation de récurrence (assertion logique quelconque) entre u_n et u_{n+1}.

Remarque 1.12. On prouvera la validité d'une telle définition en construisant $\mathbb{N}$.

Remarque 1.13. On peut même imaginer définir une suite par une relation de récurrence à plusieurs pas, ou même par une relation de récurrence forte.

Chapitre 2

Ensembles et applications

1 Introduction à la théorie des ensembles

Une assertion du type $\forall x \in X,\ p(x)$ est en fait une abréviation de l'assertion $\forall x,\ x \in X \implies p(x)$. Ainsi, toute assertion du type $\forall x \in \emptyset,\ p(x)$ est toujours vraie puisque le faux implique n'importe quoi.

Une assertion du type $\exists x \in X,\ p(x)$ est en fait une abréviation de l'assertion $\forall x,\ x \in X \wedge p(x)$. Ainsi, toute assertion du type $\exists x \in \emptyset,\ p(x)$ est toujours fausse.

Mais pour effectuer ces raisonnements, on admet l'existence d'un ensemble, appelé ensemble vide, qui ne contient aucun élément. Le besoin de cet ensemble nous amène donc à construire une théorie des ensembles.

1.1 Généralités

Définition 2.1 (Ensemble)**.** Un **ensemble** E est une collection d'objet. Lorsqu'un objet x appartient à l'ensemble E, on dit que c'est un **élément** de E et on note $x \in E$. Les objets de E sont notés entre accolades pour décrire E.

Remarque 2.1. On ne se préoccupera pas des fondements théoriques de cette définition, bien trop abstraits et éloignés du programme de classe préparatoire.

Remarque 2.2. $E = \{1, 2, 3\}$ est un ensemble.

Définition 2.2 (Ensemble vide)**.** On admet l'existence d'un ensemble appelé **ensemble vide** et noté $\emptyset$ qui ne contient aucun élément. Quel que soit x, l'assertion $x \in \emptyset$ sera toujours fausse.

Remarque 2.3. De même, on ne se préoccupera pas des fondements de l'existence de l'ensemble vide.

Définition 2.3 (Inclusion)**.** Soit E et F deux ensembles. E est dit **inclus** dans F lorsque

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

On note alors $E \subset F$ ou encore $E \subseteq F$.

Exemple 2.1. On a toujours $\emptyset \subset F$.

Exemple 2.2. On a toujours $E \subset E$.

Définition 2.4 (Égalité ensembliste)**.** Soit E et F deux ensembles. E est dit **égal** à F lorsque

$$\forall x,\ x \in E \iff x \in F$$

On note alors $E = F$. On dit aussi que E est un **sous-ensemble** de F.

Proposition 2.1. *Soit E et F deux ensembles. On a :*

$$E = F \iff (E \subset F) \wedge (F \subset E)$$

Méthode 2.1 (Montrer une égalité ensembliste : double inclusion)**.** On en déduit la méthode générale pour montrer que deux ensembles E et F sont égaux, dite de "double inclusion" :

- Soit $x \in E$. Alors ... donc $x \in F$.
- Réciproquement, soit $x \in F$. Alors ... donc $x \in E$.

Remarque 2.4 (Insensibilité à la permutation et à la répétition)**.** En vertu du principe de double-inclusion, l'ordre et la répétition 'ont pas d'importance dans la définition d'un ensemble. Par exemple, on a $\{1, 2, 3\} = \{1, 3, 2\}$ et $\{1, 2, 3\} = \{1, 1, 2, 3\}$.

Proposition 2.2 (Unicité de l'ensemble vide, HP)**.** *L'ensemble vide est unique.*

Démonstration. Soit $\emptyset$ et $\emptyset'$ deux ensembles vides. On a $\emptyset \subset \emptyset'$ et $\emptyset' \subset \emptyset$ donc $\emptyset = \emptyset'$. □

Remarque 2.5 (Note historique)**.** A l'occasion de cette courte démonstration, le grand FM avait déclaré : "on passe en mode hors-programme de ta mère".

Définition 2.5 (Ensemble des parties d'un ensemble)**.** Soit E un ensemble. Les sous-ensembles de E forment un nouvel ensemble qu'on appelle **l'ensemble des parties de** E. On le note $\mathcal{P}(E)$.

Remarque 2.6. Là non plus, on ne se préoccupera pas de savoir pourquoi les sous-ensembles de E forment un nouvel ensemble.

Exemple 2.3. On a $\mathcal{P}(\emptyset) = \{\emptyset\}$ et $\mathcal{P}(\mathcal{P}(\emptyset)) = \{\emptyset, \{\emptyset\}\}$.

Définition 2.6 (Définition d'un ensemble par compréhension)**.** Soit E un ensemble et $p(x)$ une assertion qui dépend de $x \in E$. On peut définir l'ensemble des $x \in E$ qui satisfont $p(x)$. On le note $\{x \in E \mid p(x) \text{ vraie}\}$ ou $\{x \in E,\ p(x) \text{ vraie}\}$ voire $\{x \in E : p(x) \text{ vraie}\}$. On dit qu'il s'agit d'une définition par **compréhension**.

Remarque 2.7. Le choix de la notation pour un ensemble défini par compréhension dépend de ce qui est utilisé pour définir l'assertion $p(x)$. Par exemple, lorsqu'on travaille avec de la divisibilité, puisque le symbole "divise" s'écrit "|", on préfèrera utiliser " :" ou ",".

Exemple 2.4. $\mathbb{U} = \{z \in \mathbb{C}\ : |z| = 1\}$

Méthode 2.2 (Montrer une égalité ensembliste bis)**.** Voici une deuxième façon de montrer que $A = B$. Elle ne fonctionne que si A et B sont inclus dans un même sur-ensemble de référence E. Il suffit alors de montrer que $\forall x \in E,\ x \in A \iff x \in B$ Bien noter la différence avec la définition, on se restreint ici à un sur-ensemble et non pas à l'ensemble des éléments des mathématiques.

1.2 Opérations ensemblistes

Jusqu'à ce qu'on l'écrase on fixe E un ensemble de référence et on s'intéresse à ses parties.

Définition 2.7 (Union)**.** Soit A et B deux sous-ensembles de E. L'**union de A et B** ou **réunion**, notée $A \cup B$ est définie par compréhension comme :

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

Définition 2.8 (Intersection)**.** Soit A et B deux sous-ensembles de E. L'**intersection de A et B**, notée $A \cap B$ est définie par compréhension comme :

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

Définition 2.9 (Ensemble disjoints)**.** Soit A et B deux sous-ensembles de E. On dit que A et B sont **disjoints** lorsque $A \cap B = \emptyset$.

Exemple 2.5. Voici quelques cas particuliers :

- Si $A \subset B$, alors on a $A \cup B = B$ et $A \cap B = A$.
- Si $A = \emptyset$, alors on a $A \cup B = \emptyset \cup B = B$ et $A \cap B = \emptyset \cap B = \emptyset$.
- Si $B = E$, alors on a $A \cup B = A \cup E = E$ et $A \cap B = A \cap E = A$.

Remarque 2.8. Attention : l'assertion $\forall x \in A,\ x \in A \iff x \in B$ ne montre que $A \subset B$.

Proposition 2.3 (Compatibilité de l'inclusion avec l'union et l'intersection)**.** *Soit A, B et C des sous-ensembles de E. On a*

$$A \subset B \implies \begin{cases} (A \cup C) \subset (B \cup C) \\ (A \cap C) \subset (B \cap C) \end{cases}$$

Proposition 2.4 (Commutativité de l'union et de l'intersection)**.** *Soit A et B deux sous-ensembles de E. On a :*

$$A \cup B = B \cup A \quad et \quad A \cap B = B \cap A$$

Définition 2.10 (Complémentaire)**.** Soit A un sous-ensemble de E. Le **complémentaire** de A dans E est défini par compréhension comme :

$$\complement_E^A := \{x \in E \mid x \notin A\}$$

On le note plus souvent $\overline{A}$ lorsqu'il n'y a pas d'ambiguïté par rapport au sur-ensemble de référence.

Exemple 2.6. Attention, on n'a pas toujours $\complement_E^A \neq A$. Cela n'est vrai que dès que $E \neq \emptyset$.

Exemple 2.7. Si A est un sous ensemble de E, on a toujours $A \cup \complement_E^A = E$ et $A \cap \complement_E^A = \emptyset$.

Définition 2.11 (Différence ensembliste)**.** Soit A et B des sous-ensembles de E. Plus généralement, on définit la **différence ensembliste de A et B** par compréhension comme :

$$A \setminus B := \{x \in A \mid x \notin B\}$$

Exemple 2.8. Soit A un sous-ensemble de E. On a $\complement_E^A = E \setminus A$.

Proposition 2.5 (Expression de la différence ensembliste, utile pour les exercices)**.** *Soit A et B deux sous-ensembles de E. On a $A \setminus B = A \cap \overline{B}$.*

Proposition 2.6 (Décroissance du passage au complémentaire)**.** *Soit A et B deux sous-ensembles de E. Si $A \subset B$, alors $\overline{B} \subset \overline{A}$.*

Proposition 2.7 (Involutivité du passage au complémentaire)**.** *Soit A un sous-ensemble de E. On a :*

$$\overline{(\overline{A})} = A$$

Théorème 2.1 (Lois de De Morgan)**.** *Soit A et B deux sous-ensembles de E. On a :*

$$\begin{cases} \overline{A \cup B} = \overline{A} \cap \overline{B} \\ \overline{A \cap B} = \overline{A} \cup \overline{B} \end{cases}$$

Remarque 2.9 (Note historique)**.** A propos de ce théorème, le grand FM avait déclaré : "Je crois que ça s'appelle aussi les règles de Boole pour les ensembles. J'ai déjà entendu ça quelque part, et je crois pas que j'étais bourré".

Proposition 2.8 (Associativité de l'union et de l'intersection)**.** *Soit A, B et C des sous-ensembles de E. On a :*

$$\begin{cases} (A \cup B) \cup C = A \cup (B \cup C) \\ (A \cap B) \cap C = A \cap (B \cap C) \end{cases}$$

Cela permet d'écrire $A \cup B \cup C$ et $A \cap B \cap C$ sans ambiguïté.

Remarque 2.10. Attention : on n'écrit jamais $A \cup B \cap C$ ou $A \cap B \cup C$ sans parenthèse, car il y a alors une grosse ambiguïté !

Proposition 2.9 (Distributivité de $\cup$ sur $\cap$ et de $\cap$ sur $\cup$)**.** *Soit A, B et C des sous-ensembles de E. On a :*

$$\begin{cases} A \cup (B \cap C) = (A \cup B) \cap (A \cup C) \\ A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \end{cases}$$

Définition 2.12 (Couple)**.** Soit E et F des ensembles. Soit $x \in E$ et $y \in F$. A partir de x et y, on construit le **couple** (x, y).

Remarque 2.11. Ici, on ne se préoccupe pas de la construction d'un couple.

Théorème 2.2 (Propriété fondamentale des couples)**.** *Soit E et F des ensembles. On a :*

$$\forall x \in E, \forall x' \in E, \forall y \in F, \forall y' \in F,\ (x, y) = (x', y') \iff (x = x') \wedge (y = y')$$

Remarque 2.12. La négation de cette propriété fondamentale s'écrit donc :

$$(x, y) \neq (x', y') \iff (x \neq x') \vee (y \neq y')$$

Définition 2.13 (Produit cartésien)**.** Soit E et F deux ensembles. L'ensemble des couples (x, y) pour $x \in E$ et $y \in F$ est appelé **produit cartésien de E et F**. On le note $E \times F$.

Remarque 2.13. Le produit cartésien n'a aucune raison d'être commutatif !

Remarque 2.14. Soit E et F deux ensembles. On a toujours $\emptyset \times F = \emptyset$ et $E \times \emptyset = \emptyset$. *A contrario*, si E et F sont non vides, alors $E \times F$ est non vide.

On aimerait généraliser. Pour commencer, on voudrait, si l'on dispose de E, F et G des ensembles ainsi que de $x \in E$, $y \in F$ et $z \in G$, construire un objet (x, y, z), dit **triplet**, qui vérifie une propriété fondamentale du type :

$$(x, y, z) = (x', y', z') \iff (x = x') \wedge (y = y') \wedge (z = z')$$

On peut procéder de deux façons :

- On pose $(x, y, z) := ((x, y), z)$, une sorte de couple de couple, et on vérifie alors aisément qu'il vérifie la propriété fondamentale souhaitée.
- On peut procéder d'une autre manière que nous verrons plus tard dans ce chapitre.

Puis, dans le cas général, si $n \geqslant 2$ et $E_1, \ldots, E_n$ sont des ensembles, à partir de $x_1 \in E_1, \ldots, x_n \in E_n$, on peut construire un **n-uplet** (aussi dit **n-liste**) $(x_1, \ldots, x_n)$ qui vérifie une propriété fondamentale du type :

$$(x_1, \ldots, x_n) = (x'_1, \ldots, x'_n) \iff (x_1 = x'_1) \wedge \ldots \wedge (x_n = x'_n)$$

Définition 2.14. L'ensemble des n-uplets à valeurs dans $E_1, \ldots, E_n$ est noté $E_1 \times \ldots \times E_n$.

Définition 2.15. E^n est une notation standard pour le produit cartésien $E \times \ldots \times E$ où l'ensemble E est présent n fois.

Théorème 2.3 (Axiome du choix, HP)**.** *Lorsqu'on réalise le produit cartésien infini d'une infinité d'ensemble tous non vides, on décide que ce produit cartésien est non vide et il existe alors une* ***fonction de choix*** *qui nous permet d'en choisir un élément.*

Remarque 2.15. On n'est pas obligé de mettre des parenthèses intérieures dans n-uplet, même si les ensembles du produit cartésien sont distincts. Formellement, on identifie

$$E_1 \times \ldots \times E_l \times (E_{l+1} \times \ldots \times E_m) \times E_{m+1} \times \ldots \times E_n$$

à $E_1 \times \ldots \times E_n$. C'est ce qu'on peut appeler l'**associativité du produit cartésien**.

2 Applications

Le programme ne fait pas de distinction entre les mots "application" et "fonction" (même s'il y en a une légère en réalité) : nous considérerons donc que ces mots sont synonymes.

2.1 Définition formelle d'une application

Définition 2.16 (Application, fonction)**.** Soit E et F deux ensembles. Une **application**, ou **fonction**, de E dans F est un objet f défini par :

$$f = (E, F, \Gamma)$$

où $\Gamma \subset E \times F$ est appelée **graphe de** f et est telle que $\forall x \in E,\ \exists! y \in F,\ (x, y) \in \Gamma$. On la note plus souvent

$$\begin{array}{rccc} f & : & E & \to & F \\ & & x & \mapsto & f(x) \end{array}$$

f **associe** à chaque $x \in E$ une **unique image** $f(x)$. On dit que x est **un antécédent** de $f(x)$ par f.

Théorème 2.4. *Soit E, E', F et F' des ensembles. Soit $f : E \to F$ et $g : E' \to F'$ deux applications. Alors on a :*

$$f = g \iff \begin{cases} E = E' \\ F = F' \\ \forall x \in E,\ f(x) = g(x) \end{cases}$$

Remarque 2.16. Le troisième point traduit l'égalité des graphes.

Remarque 2.17. Pour représenter f, si E et F sont des parties de $\mathbb{R}$ on peut dessiner son graphe. Sinon, on faire un diagramme sagittal.

Définition 2.17 (Ensembles des fonctions)**.** Soit E et F des ensembles. L'ensemble des fonctions $f : E \to F$ est noté $\mathcal{F}(E, F)$ ou F^E.

Définition 2.18 (Fonction indicatrice)**.** Soit E un ensemble et A un sous-ensemble de E. L'**indicatrice de A dans E** est la fonction :

$$\begin{array}{rccc} \mathbb{1}_A & : & E & \to & \mathbb{R} \\ & & x & \mapsto & \begin{cases} 0 \text{ si } x \notin A \\ 1 \text{ si } x \in A \end{cases} \end{array}$$

Proposition 2.10 (Propriétés des indicatrices)**.** *Soit E un ensemble et A et B des sous-ensembles de E. On a :*

- $\mathbb{1}_{A \cap B} = \mathbb{1}_A \times \mathbb{1}_B$
- $\mathbb{1}_{\overline{A}} = 1 - \mathbb{1}_A$
- $\mathbb{1}_{A \cup B} = \mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_{A \cap B} = \mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_A \times \mathbb{1}_B$

Définition 2.19 (Restriction)**.** Soit E et F deux ensembles, ainsi que A un sous-ensemble de E. Soit $f : E \to F$ une fonction. La **restriction de f à A** est la fonction :

$$\begin{array}{rccc} f_{|A} & : & A & \to & F \\ & & x & \mapsto & f(x) \end{array}$$

Définition 2.20 (Corestriction)**.** Soit E et F deux ensembles, ainsi que B un sous-ensemble de F. Soit $f : E \to F$ une fonction. Si $\forall x \in E,\ f(x) \in B$, alors on peut **corestreindre** f à B et la **corestriction de f à B** est la fonction :

$$\begin{array}{rccc} f^{|B} & : & E & \to & B \\ & & x & \mapsto & f(x) \end{array}$$

Remarque 2.18. On peut même considérer $f_{|A}^{|B}$ à condition que $\forall x \in A,\ f(x) \in B$.

Définition 2.21 (Prolongement)**.** Soit E, E', F et F' des ensembles tels que $E \subset E'$ et $F \subset F'$. Soit $f : E \to F$. **Un prolongement** de f est une fonction $g : E' \to F'$ telle que

$$g_{|E}^{|F} = f$$

Remarque 2.19. Attention : il peut exister plusieurs prolongements d'une même fonction !

2.2 Injectivité, surjectivité, bijectivité

Dans toute cette partie, on fixe E et F deux ensembles ainsi que $f : E \to F$ une fonction.

Définition 2.22 (Injectivité)**.** f est dite **injective** lorsque

$$\forall (x, x' \in E^2),\ f(x) = f(x') \implies x = x'$$

ou de manière équivalente lorsque

$$\forall (x, x') \in E^2,\ x \neq x' \implies f(x) \neq f(x')$$

En rédaction automatique, pour montrer que f est injective, on écrira "Soit $(x, x') \in E^2$ tel que $f(x) = f(x')$" puis on montrera que $x = x'$ (ou alors on utilisera l'autre définition équivalente).

Remarque 2.20. La non injectivité s'écrit donc : $\exists (x, x') \in E^2,\ (x \neq x') \wedge (f(x) = f(x'))$.

Proposition 2.11. *Toute fonction réelle strictement monotone est injective.*

Remarque 2.21. Attention : la réciproque est fausse !

Remarque 2.22. La fonction $(\emptyset, \emptyset, \emptyset)$ est à la fois constante, croissante, décroissante, strictement croissante et strictement décroissante. Il en va de même pour les fonctions $(\emptyset, F, \emptyset)$ et $(\{a\}, \mathbb{R}, \{(a, b)\})$.

Définition 2.23 (Injection canonique)**.** Soit E un ensemble et A un sous-ensemble de E. **L'injection canonique de A dans E** est l'application :

$$\begin{array}{rccc} \iota & : & A & \to & E \\ & & x & \mapsto & x \end{array}$$

Elle est injective.

Proposition 2.12. *L'injectivité s'hérite par restriction et corestriction.*

Définition 2.24. f est dite **surjective** lorsque

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

En rédaction automatique, on écrira "Soit $y \in F$" puis on exhibera un candidat après avoir cherché par condition nécessaire. On pourra aussi utiliser des théorèmes d'existence comme le TVI dans $\mathbb{R}$.

Exemple 2.9. L'application $\begin{array}{ccc} \mathbb{R} & \to & \mathbb{U} \\ \theta & \mapsto & \exp(i\theta) \end{array}$ est surjective.

Injectivité et surjectivité ne sont pas incompatibles. C'est ce qui fait l'objet de la

Définition 2.25 (Bijectivité)**.** f est dite **bijective** lorsqu'elle est à la fois injective et surjective. De manière équivalente :

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

Méthode 2.3 (Montrer une bijectivité)**.** Pour montrer qu'une fonction $f : E \to F$ est bijective on peut :

- Se donner $y \in F$ et $x \in E$ et montrer que l'équation $y = f(x)$ admet une unique solution.

- Montrer séparément que f est injective et surjective.
- Si on est dans $\mathbb{R}$, réaliser un joli tableau de variations.

Définition 2.26 (Fonction identité)**.** La **fonction identité de** E est définie par :

$$\begin{array}{rccc} \mathrm{id}_E & : & E & \to & E \\ & & x & \mapsto & x \end{array}$$

Elle est bijective.

Définition 2.27 (Composée)**.** Soit E, F et G des ensembles ainsi que $f : E \to F$ et $g : F \to G$ des fonctions. Alors, la **composée de** g **par** f est la fonction définie par :

$$\begin{array}{rccc} g \circ f & : & E & \to & G \\ & & x & \mapsto & g(f(x)) \end{array}$$

Remarque 2.23. Soit E et F des ensembles ainsi que $f : E \to F$. On a :

$$\begin{cases} f \circ \mathrm{id}_E = f \\ \mathrm{id}_F \circ f = f \end{cases}$$

Remarque 2.24. Attention : la loi $\circ$ est non commutative en général ! Déjà, elle n'est souvent même pas définie dans les deux sens pour une question de domaines, et même si les fonctions étaient définies, la coïncidence n'a aucune raison d'avoir lieu. On trouvera un contre-exemple avec une fonction constante et une fonction linéaire dans $\mathbb{R}$.

Proposition 2.13 (Associativité de la composition)**.** *Soit E, F, G et H des ensembles. Soit $f : E \to F$, $g : F \to G$ et $h : G \to H$ des fonctions. Alors, on a :*

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

Théorème 2.5. *Soit E, F et G des ensembles. Soit $f ; E \to F$ et $g : F \to G$ des fonctions.*

- *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.*

Théorème 2.6 (Réciproque partielle)**.** *Soit E, F et G des ensembles. Soit $f : E \to F$ et $g : F \to G$ des fonctions. Réciproquement :*

- *Si $g \circ f$ est injective, alors f est injective.*
- *Si $g \circ f$ est surjective, alors g est surjective.*

Remarque 2.25. La construction de $g \circ f$ reste valable si $f : E \to F'$ et $g : F \to G$ avec $F' \subset F$. L'essentiel des résultats précédents reste valable, **sauf la composée de deux surjections (resp. bijections) qui n'a plus aucune raison d'être une surjection (resp. bijection)**.

2.3 Bijection réciproque

Définition 2.28 (Bijection réciproque)**.** Soit E et F deux ensembles et $f : E \to F$ une bijection. La **bijection réciproque de** f est la fonction définie de la manière suivante : soit $y \in F$. f est bijective donc $\exists! x \in E,\ y = f(x)$. Fixons-le. On pose alors

$$f^{-1}(y) := x$$

La fonction f^{-1} est alors elle aussi bijective, et elle vérifie

$$\begin{cases} f^{-1} \circ f = \mathrm{id}_E \\ f \circ f^{-1} = \mathrm{id}_F \end{cases}$$

Théorème 2.7. *Soit E et F des ensembles. Soit $f : E \to F$ une fonction quelconque. S'il existe $g : F \to E$ telle que $g \circ f = \mathrm{id}_E$ et $f \circ g = \mathrm{id}_F$ alors :*

1. *f est bijective*
2. *g est la bijection réciproque de f*

Remarque 2.26. Il s'agit donc là d'une nouvelle méthode pour prouver la bijectivité d'une fonction, qui par la même occasion nous fournit sa bijection réciproque.

Corollaire 2.1 (Involutivité du passage à la bijection réciproque)**.** *Soit E et F des ensembles. Soit $f : E \to F$ une fonction. Si f est bijective, de bijection réciproque f^{-1}, alors f^{-1} est bijective de bijection réciproque f.*

Corollaire 2.2 (Bijection réciproque d'une composée de bijections)**.** *Soit E, F et G des ensembles. Soit $f : E \to F$ et $g : F \to G$ des fonctions. Si f et g sont bijectives, alors :*

- *$g \circ f$ est bijective (on le savait déjà par un théorème précédent)*
- *La bijection réciproque de $g \circ f$ est donnée par $f^{-1} \circ g^{-1}$. Autrement dit, on a alors $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.*

Remarque 2.27. Attention à bien inverser f et g lorsqu'on passe à la bijection réciproque ! Il s'agit en fait là d'un résultat plus général que nous reverrons dans le chapitre "Introduction à l'algèbre".

Définition 2.29 (Involution)**.** Soit E un ensemble et $f : E \to E$. On dit que f est **involutive** lorsque $f \circ f = \mathrm{id}_E$, *ie* lorsque $\forall x \in E,\ f(f(x)) = x$.

Proposition 2.14. *Toute fonction involutive est bijection, de bijection réciproque elle-même.*

Définition 2.30 (Permutation)**.** Soit E un ensemble. Une permutation de E est une bijection $E \to E$. On note $\mathcal{S}_E$ ou $\mathfrak{S}_E$ l'ensemble des permutations de E.

Exemple 2.10. Soit E un ensemble. L'application $\begin{array}{ccc} \mathfrak{S}_E & \to & \mathfrak{S}_E \\ f & \mapsto & f^{-1} \end{array}$ est une involution.

Méthode 2.4 (Bonne définition d'une fonction)**.** Soit E et F des ensembles. Voici comment montrer que $\begin{array}{ccccc} f & : & E & \to & F \\ & & x & \mapsto & f(x) \end{array}$ est bien définie. Il y a trois points de vigilance, à vérifier dans l'ordre dans lequel ils sont énoncés :

- Est-ce que x est un objet primaire ? Vérifier qu'il ne s'agit pas 'une formule d'un autre objet plus simple. Dans le cas où x n'est pas un objet primaire et possède plusieurs représentants (rationnels, éléments d'ensembles quotients par exemple), il faut montrer que la valeur de $f(x)$ ne dépend pas du représentant choisi.
- Calculabilité ? Soit $x \in E$. Montrer que $f(x)$ existe.
- Bon ensemble d'arrivée ? Soit $x \in E$. Montrer que $f(x) \in F$.

Méthode 2.5 (Prouver une propriété sur f^{-1})**.** Pour montrer une propriété sur f^{-1}, on montre f de cette propriété puis on se débrouille pour simplifier par f.

Exercice 2.1. Pour s'exercer, on pourra montrer que si $f : \mathbb{R} \to \mathbb{R}$ est bijective et impaire, alors f^{-1} est impaire.

Exemple 2.11. $\begin{array}{ccc} \mathbb{R}_+ & \to & \mathbb{R}_+ \\ y & \mapsto & \sqrt{y} \end{array}$ est définie comme la bijection de réciproque de $\begin{array}{ccc} \mathbb{R}_+ & \to & \mathbb{R}_+ \\ x & \mapsto & x^2 \end{array}$. Or, la deuxième est strictement croissante, donc la première est aussi strictement croissante (appliquer la méthode et utiliser la contraposée de la croissance).

Définition 2.31. Soit E et F deux ensembles. Les deux assertions suivantes sont équivalentes :

1. $\exists f : E \to F$ bijective
2. $\exists g : F \to E$ bijective

Lorsque l'une de ces deux assertions est vérifiée, on dit que E et F sont **équipotents**.

Théorème 2.8 (Théorème de Cantor-Bernstein HP)**.** *Soit E et F deux ensembles. S'il existe $f : E \to F$ injective et $g : F \to E$ injective, alors E et F sont équipotents.*

2.4 Image directe, image réciproque

Définition 2.32 (Image directe)**.** Soit E et F des ensembles et $f : E \to F$. Soit A un sous-ensemble de E. L'**image directe de A par f** est l'ensemble défini par :

$$f(A) := \{y \in F \mid \exists x \in E,\ y = f(x)\}$$

Définition 2.33 (Image réciproque)**.** Soit E et F des ensembles et $f : E \to F$. Soit B un sous-ensemble de F. L'**image réciproque de B par f** est l'ensemble défini par :

$$f^{-1}(B) := \{x \in E \mid f(x) \in B\}$$

Remarque 2.28. Même si les fonctions précédemment définies vont d'un ensemble de parties dans un autre, on les note quand même f et f^{-1} à l'usage. **En cas de doute, vérifier le type d'objet d'entrée et de sortie**.

Remarque 2.29. Si f est bijective, $f^{-1}(B)$ peut être interprété comme l'image directe de B par f^{-1} ou comme l'image réciproque de B par f. Heureusement, les deux notions coïncident alors, il s'agit d'une simple double inclusion.

Exemple 2.12. Soit E et F deux ensembles et $f : E \to F$. On a :

- $f(\emptyset) = \{y \in F \mid \exists x \in \emptyset,\ f(x) = y\}\ \emptyset$

- $f \in (\emptyset) = \{x \in E \mid f(x) \in \emptyset\} = \emptyset$
- $f(E) = \{y \in F \mid \exists x \in E, f(x) = y\} =: \mathrm{Im}(f)$
- $f^{-1}(F) = \{x \in E \mid f(x) \in F\} = E$
- Soit $x_0 \in E$. $f(\{x_0\}) = \{y \in F \mid \exists \{x_0\},\ y = f(x)\} = \{f(x_0)\}$

Proposition 2.15. *Soit E et F deux ensembles. $f : E \to F$ est surjective si, et seulement si, $\mathrm{Im}(f) = F$.*

Proposition 2.16. *Soit E et F deux ensembles ainsi que $f : E \to F$. Alors $f^{|\mathrm{Im}(f)}$ existe toujours et est surjective par construction.*

Définition 2.34 (Ensemble défini par image directe)**.** L'ensemble $f(A)$ peut aussi être noté **symboliquement** :

$$\{f(x) \mid x \in A\}$$

C'est ce qu'on appelle la définition d'un ensemble par **image directe**.

Exercice 2.2. Soit E, F et G des ensembles. Soit $f : E \to F$ et $g : F \to G$ des fonctions.

1. Soit A un sous-ensemble de E. Montrer que $(g \circ f)(A) = g(f(A))$.
2. Soit B un sous-ensemble de G. montrer que $(g \circ f)^{-1}(B) = f^{-1}(g^{-1}(B))$.

(Procéder par double-inclusion à chaque fois)

Proposition 2.17 (Croissance des fonctions image directe et image réciproque)**.** *Soit E et F deux ensembles et $f : E \to F$. On a :*

- $\forall A_1, A_2 \subset E,\ A_1 \subset A_2 \implies f(A_1) \subset f(A_2)$
- $\forall B_1, B_2 \subset F,\ B_1 \subset B_2 \implies f^{-1}(B_1) \subset f^{-1}(B_2)$

Proposition 2.18. *Soit E et F deux ensembles et $f : E \to F$. On a :*

- $\forall B_1, B_2 \subset F,\ f^{-1}(B_1 \cup B_2) = f^{-1}(B_1) \cup f \in (B_2)$
- $\forall B_1, B_2 \subset F,\ f^{-1}(B_1 \cap B_2) = f^{-1}(B_1) \cap f \in (B_2)$
- $\forall B \subset F,\ f^{-1}(\overline{B}) = \overline{f^{-1}(B)}$
- $\forall A_1, A_2 \subset E,\ f(A_1 \cup A_2) = f(A_1) \cup f(A_2)$

Remarque 2.30. Attention, on a $f(A_1 \cap A_2) \neq f(A_1) \cap f(A_2)$ en général (sauf lorsque f est injective) et de même on a $f(\overline{A}) \neq \overline{f(A)}$ en général.

2.5 Familles

Définition 2.35 (Famille)**.** Soit I et E des ensembles. Formellement, une **famille $(x_i)_{i \in I}$ indexée par I et à valeurs dans** E est une application :

$$\begin{array}{rccc} (x_i)_{i \in I} & : & I & \to & E \\ & & i & \mapsto & x_i \end{array}$$

i s'appelle l'**indice**. On note donc E^I l'ensemble des familles indexées par I à valeurs dans E.

Exemple 2.13. Soit X un ensemble. Une **suite à valeurs dans** X est une application $(x_n)_{n \in \mathbb{N}} \in X^{\mathbb{N}}$.

Définition 2.36 (Sous-famille)**.** Soit I et E des ensembles. Soit $(x_i)_{i\in I} \in E^I$ et $J \subset I$. Une **sous-famille** de la famille $(x_i)_{i\in I}$ est une famille du type $(x_i)_{i\in J}$.

Remarque 2.31 (Retour sur la construction des n-uplets, HP)**.** Soit $n \in \mathbb{N}^\star$. Une façon de construire le n-uplet $(x_1, \dots, x_n)$ à valeurs dans E est de considérer la famille $(x_i)_{i\in[\![1,n]\!]} \in E^{[\![1,n]\!]}$ On dispos alors de la propriété fondamentale des n-uplets. En quelque sorte, on construit E^n sous la forme $E^{[\![1,n]\!]}$. Notons que cela nécessitait avant de construire les triplets comme couples de couple, puisque les fonctions qui nous servent ici à définir les n-uplets sont elles-mêmes des triplets.

Définition 2.37 (Union et intersection d'une famille de parties)**.** Soit I et E des ensembles. Soit $(A_i)_{i\in I} \in \mathcal{P}(E)^I$. On pose :

$$\bigcup_{i\in I} A_i := \{x \in E \mid \exists i \in I,\ x \in A_i\}$$

$$\bigcap_{i\in I} A_i := \{x \in E \mid \forall i \in I,\ x \in A_i\}$$

Définition 2.38 (Famille de parties deux à deux disjointes)**.** Soit I et E des ensembles. Soit $(A_i)_{i\in I} \in \mathcal{P}(E)^I$. On dit que **les A_i sont deux à deux disjoints** lorsque

$$\forall (i,j) \in I^2,\ i \neq j \implies A_i \cap A_j = \emptyset$$

Définition 2.39 (Union carrée/union disjointe)**.** Soit I et E des ensembles. Soit $(A_i)_{i\in I} \in \mathcal{P}(E)^I$. Si les A_i sont deux à deux disjoints, $\bigcup_{i\in I} A_i$ sera plutôt notée $\bigsqcup_{i\in I} A_i$. On parle d'**union disjointe** ou d'**union carrée**.

Définition 2.40 (Recouvrement)**.** Soit I et E des ensembles. Soit $(A_i)_{i\in I} \in \mathcal{P}(E)^I$. Si $\bigcup_{i\in I} A_i = E$, on dit que les A_i forment un **recouvrement** de E.

Définition 2.41 (Recouvrement disjoint)**.** Soit I et E des ensembles. Soit $(A_i)_{i\in I} \in \mathcal{P}(E)^I$. Si $\bigsqcup_{i\in I} A_i = E$ (on suppose donc implicitement que les A_i sont deux à deux disjoints), on dit que les A_i forment un **recouvrement disjoint** de E.

Exercice 2.3 (NE+FHP)**.** Soit A un ensemble appelé **alphabet**. Intuitivement, un **mot** de longueur $n \in \mathbb{N}$ sera un élément de A^n. On voudrait définir le lexique d'une langue comme une partie de $\bigcup_{n\in\mathbb{N}} A^n$. Mais cette union a-t-elle un sens ? Autrement dit, les A^n peuvent-ils être vus comme appartenant à un même $\mathcal{P}(E)$.

1. Construire un ensemble E tel que $\forall n \in \mathbb{N},\ A^n \subset E$ (voir A^n comme $A^{[\![1,n]\!]}$ puis on prendra $E = \mathcal{P}(\mathbb{N}) \times \{A\} \times \mathcal{P}(\mathbb{N} \times A)$)
2. On voudrait définir la longueur d'un mot comme l'unique $n \in \mathbb{N}$ tel que ce mot appartienne à A^n. Mais un tel n est-il bien unique ? Montrer que les A^n sont deux à deux disjoints.

Dans un ensemble, la plupart des résultat vus avec deux parties restent vraies avec une famille de parties, et nous en faisons la liste maintenant.

Proposition 2.19 (Lois de De Morgan)**.** *Soit I et E des ensembles. Soit $(A_i)_{i\in I} \in \mathcal{P}(E)^I$. On a*

$$\overline{\bigcup_{i\in I} A_i} = \bigcap_{i\in I} \overline{A_i}$$

$$\overline{\bigcap_{i\in I} A_i} = \bigcup_{i\in I} \overline{A_i}$$

Proposition 2.20 (Distributivités)**.** *Soit I et E des ensembles. Soit $A \subset E$ et $(B_i)_{i\in I} \in \mathcal{P}(E)^I$. On a*

$$A \cup \left(\bigcap_{i\in I} B_i\right) = \bigcap_{i\in I} (A \cup B_i)$$

$$A \cap \left(\bigcup_{i\in I} B_i\right) = \bigcup_{i\in I} (A \cap B_i)$$

Proposition 2.21. *Soit I, E et F des ensembles. Soit $f : E \to F$, $(A_i)_{i\in I} \in \mathcal{P}(E)^I$ et $(B_i)_{i\in I} \in \mathcal{P}(F)^I$. On a :*

- $f^{-1}\left(\bigcup_{i\in I} B_i\right) = \bigcup_{i\in I} f^{-1}(B_i)$
- $f^{-1}\left(\bigcap_{i\in I} B_i\right) = \bigcap_{i\in I} f^{-1}(B_i)$
- $f\left(\bigcup_{i\in I} A_i\right) = \bigcup_{i\in I} f^{-1}(A_i)$

Proposition 2.22 (Cas des unions disjointes)**.** *Soit I, E et F des ensembles. Soit $A \subset F$ et $(B_i)_{i\in I} \in \mathcal{P}(F)^I$. Supposons que les B_i sont deux à deux disjoints. Alors :*

- $A \cap \left(\bigsqcup_{i\in I} B_i\right) = \bigsqcup_{i\in I} A \cap B_i$
- $f^{-1}\left(\bigsqcup_{i\in I} B_i\right) = \bigsqcup_{i\in I} f^{-1}(B_i)$

Définition 2.42 (Généralisation du produit cartésien)**.** Soit I et E des ensembles. Soit $(A_i)_{i\in I} \in \mathcal{P}(E)^I$. On pose :

$$\prod_{i\in I} A_i := \left\{(x_i)_{i\in I} \in E^I \mid \forall i \in I,\ x_i \in A_i\right\}$$

Remarque 2.32. A une famille on peut associer un ensemble et vice-versa.

- A la famille $(x_i)_{i\in I} \in E^I$ on peut associer l'ensemble $\{x_i \mid i \in I\}$
- A l'ensemble A, on peut associer la famille $(a)_{a\in A}$

A chaque fois, famille et ensemble auront les mêmes éléments. Toutefois, on perde toute information sur l'ordre ou la répétition en passant aux ensembles.

Dans le même esprit, soit $\mathcal{A} \in \mathcal{P}(E)$ ($\mathcal{A}$ n'est plus une famille de parties de E mais un ensemble de parties de E). On posera alors

$$\bigcup_{A \in \mathcal{A}} A := \{x \in E \mid \exists A \in \mathcal{A},\ x \in A\}$$

$$\bigcap_{A \in \mathcal{A}} A := \{x \in E \mid \forall A \in \mathcal{A},\ x \in A\}$$

3 Ensembles finis

Lemme 2.1 (Principe des tiroirs, cas $[\![1, n]\!]$)**.** *Soit $m, n \in \mathbb{N}$ tels que $m > n$ et $f : [\![1, m]\!] \to [\![1, n]\!]$. Alors, f n'est pas injective.*

Définition 2.43 (Ensemble fini)**.** Soit E un ensemble. E est dit **fini** lorsqu'il existe $n \in \mathbb{N}$ tel que E soit équipotent à $[\![1, n]\!]$.

Définition 2.44 (Cardinal d'un ensemble fini)**.** Soit E un ensemble fini. Alors le $n \in \mathbb{N}$ tel que E soit équipotent à $[\![1, n]\!]$ est unique. On définit alors le cardinal de E par :

$$\mathrm{Card}(E) := n$$

Exemple 2.14. Soit $a, b \in \mathbb{Z}$ tels que $a \leqslant b$. Alors $[\![a, b]\!]$ est fini, de cardinal $b - a + 1$.

Définition 2.45 (Famille finie, taille d'une famille finie)**.** Une famille $(x_i)_{i \in I} \in E^I$ est dite **finie** lorsque I est fini. La **taille** de la famille est alors définie comme le cardinal de I.

Définition 2.46 (Dénombrer)**.** Dénombrer un ensemble fini consiste à calculer son cardinal.

Méthode 2.6 (Montrer une égalité d'ensembles finis)**.** Soit E un ensemble finie et A un sous-ensemble de E. Si $A \subset E$ et $\mathrm{Card}(A) = \mathrm{Card}(E)$, alors $A = E$.

Proposition 2.23. *Soit A et B deux ensembles finis. Alors $A \cup B$ est un ensemble fini, et on a*

$$\mathrm{Card}(A \cup B) \leqslant \mathrm{Card}(A) + \mathrm{Card}(B)$$

Proposition 2.24. *Soit $A_1, \ldots, A_n \subset E$ toutes finies. Alors $A_1 \cup \ldots \cup A_n$ est finie et on a*

$$\mathrm{Card}(A_1 \cup \ldots \cup A_n) \leqslant \mathrm{Card}(A_1) + \ldots + \mathrm{Card}(A_n)$$

Proposition 2.25. *L'image directe d'un ensemble fini est finie.*

Proposition 2.26 (Caractérisation des ensembles infinis)**.** *Soit E un ensemble. E est infini si, et seulement si, il existe une suite à valeurs dans E injective (ie ses termes sont deux à deux distincts).*

Exemple 2.15. L'ensemble $[0, 1]$ est infini : il suffit de considérer la suite $\left(\dfrac{1}{n+1}\right)_{n \in \mathbb{N}}$ qui est strictement décroissante donc injective.

Proposition 2.27. *Soit E un ensemble fini et F un ensemble quelconque. Les assertions suivantes sont équivalentes :*

1. *F est fini de même cardinal que E*
2. *Il existe une bijection de E dans F*
3. *Il existe une bijection de F dans E*

Théorème 2.9 (Principe des tiroirs, cas général)**.** *Soit E et F des ensembles finis tels que* $\mathrm{Card}(E) > \mathrm{Card}(F)$*. Alors, il n'existe pas d'injection de E dans F. Autrement dit, toute fonction $f : E \to F$ n'est pas injective.*

Remarque 2.33 (Extension au cas où E est infini)**. Le résultat s'étend au cas où E est infini**. En effet, s'il existait une injection de E dans F, alors celle-ci induirait par restriction une injection d'un sous-ensemble fini de E de cardinal strictement supérieur à celui de F (qui existe car E est infini) dans F, ce qui est absurde en vertu du théorème précédent.

Théorème 2.10. *Soit E et F des ensembles finis de même cardinal. Alors les trois assertion suivantes sont équivalentes :*

- *f est injective*
- *f est surjective*
- *f est bijective*

4 Compléments, HP

4.1 Théorème de Cantor-Bernstein

Nous proposons ici une démonstration du théorème de Cantor-Bernstein.

Théorème 2.11 (Théorème de Cantor-Bernstein)**.** *S'il existe une injection de E dans F et et une injection de F dans E, alors il existe une bijection de E dans F.*

Démonstration. Soit f une injection de E dans F et g une injection de F dans E. Pour chaque élément $x \in E$, on regarde la liste de ses "ancêtres" successifs : on cherch s'il possède un antécédent y par g. Si c'est le cas, on cherche si cet antécédent possède un antécédent par g dans F, et ainsi de suite. Formellement :

- On définit $E_0 := \{x \in E \mid \forall y \in F,\ g(y) \neq x\}$ l'ensemble des $x \in E$ sans ancêtre.
- De même, on définit $F_0 := \{y \in F \mid \forall x \in E,\ f(x) \neq y\}$ l'ensemble des $y \in F$ sans ancêtre.
- Par récurrence, on définit $E_{n+1} := g(F_n)$ et $F_{n+1} := f(E_n)$.
- On pose $E_p := \bigcup_{n\in\mathbb{N}} E_{2n}$, $E_i := \bigcup_{n\in\mathbb{N}} E_{2n+1}$ et $E_\infty := E \setminus (E_p \cup E_i)$.
- De même, on pose $F_p := \bigcup_{n\in\mathbb{N}} F_{2n}$, $F_i := \bigcup_{n\in\mathbb{N}} F_{2n+1}$ et $F_\infty := F \setminus (F_p \cup F_i)$.

Finalement, on construit notre bijection φ comme suit.

- Si $x \in E_p$, on pose $\varphi(x) := f(x)$. $\varphi(x)$ possède un ancêtre de plus que x (car c'est son "fils"), donc $\varphi(x) \in F_i$.

- Si $x \in E_i$, x possède un antécédent $y \in F$ puisqu'il a un nombre impair d'ancêtres (et que tout nombre naturel impair est supérieur ou égal à 1). De plus, cet antécédent y est unique car g est injective. On pose alors $\varphi(x) := y$. y possède un antécédent de moins que x (car c'est son "père"), donc $y \in F_p$.
- Si $x \in E_\infty$, on pose $\varphi(x) := f(x)$. x possède une infinité d'ancêtres, donc $\varphi(x)$ aussi car c'est son "fils".

On vient de montrer que φ envoyait E_p sur F_i, E_i sur F_p et E_∞ sur F_∞. Comme ces ensembles sont disjoints, il suffit de montrer que φ est bijectives *séparément* de E_p dans F_i, de E_i dans F_p et de E_∞ dans E_∞.

- Par construction, φ est injective sur E_p et E_∞.
- Sur E_i, si on a $\varphi(x) = \varphi(x')$ alors x et x' ont le même antécédent par g, donc comme une application ne peut avoir qu'une seule image, on a $x = x'$.
- Par construction, φ est surjective sur F_p (l'antécédent de $y \in F_p$ est donné par $g(y)$, c'est un élément qui possède un nombre pair d'ancêtres donc il est bien dans E_i).
- Sur F_i, tout élément y possède par définition au moins un antécédent x par f (puisqu'il possède un nombre impair d'ancêtres). L'élément x sera alors dans E_p et s'enverra bien sur y par φ.
- Sur F_∞, tout élément y possède par définition au moins un antécédent x par f (puisqu'il possède un nombre infini d'ancêtres). L'élément x possède alors une infinité d'ancêtres et est donc dans E_∞ ; par conséquent, il s'enverra bien sur y par φ.

Ainsi, la preuve est achevée. □

Plusieurs variantes du théorème de Cantor-Bernstein existent. Avant toute chose, démontrons le

Lemme 2.2. *Il existe une injection de E dans F si, et seulement si, il existe une surjection de F dans E.*

Démonstration. S'il existe une injection i de E dans F, alors on construit une surjection s de F dans E de la manière suivante :

- Si $y \in F$ possède un antécédent x par i, alors x est unique car i est injective et on pose $s(y) := x$.
- On choisit un "élément-poubelle" $x_0 \in E$ pour tous les y qui n'ont pas d'antécédent par i, et on pose $s(y) := x_0$.

Maintenant, s'il existe une surjection s de F dans E, on construit une injection i de E dans F comme suit : tout élément $x \in E$ possède au moins un antécédent $y \in F$ par s, on en choisit un et on pose $i(x) := y$. □

Remarque 2.34 (Axiome du choix)**.** Dans notre construction, nous utilisons **l'axiome du choix**, qui garantit l'existence d'une manière de choisir un élément dans un ensemble quelconque (ce qu'on appelle une **fonction de choix**). Malgré son apparente simplicité, cette question n'a rien de trivial, et les logiciens ont même démontré qu'on pouvait faire des mathématiques consistantes en niant cet axiome.

Corollaire 2.3 (Variante du théorème de Cantor-Bernstein)**.** *Soit E et F deux ensembles. Dans les quatre cas de figure suivants, il existe une bijection de E dans F :*

- *injection de E dans F / injection de F dans E*
- *injection de E dans F / surjection de E dans F*
- *surjection de F dans E / injection de F dans E*
- *surjection de F dans E / surjection de E dans F*

Démonstration. Il suffit d'utiliser à chaque fois le lemme précédent et le théorème de Cantor-Bernstein. □

4.2 Axiomes de Zermelo-Fraenkel

Cette partie est excessivement ardue et abstraite. Aussi est-il formellement déconseillé de s'y aventurer si l'on n'est pas parfaitement à l'aise avec le reste du programme.

Peu à peu, les mathématiciens se sont rendus compte qu'ils ne pouvaient plus se contenter des définitions intuitives que nous avons données dans ce chapitre. Il leur fallait une théorie beaucoup plus rigoureuse. Il s'agit d'un travail de longue haleine initié par Cantor à la fin su XIX^e siècle et continué par ses successeurs au XX^e. Parmi eux, Zermelo et Fraenkel ont laissé leur nom au système axiomatique le plus usuel aujourd'hui. Il est constitué de plusieurs axiomes, mais tous ne nous seront pas utiles. Nous présentons donc ici seulement ceux dont nous avons absolument besoin.

Cette théorie part d'un modèle appelé **univers**, dans lequel tous les objets sont des ensembles (on suppose qu'il en existe au moins un). On peut voir l'univers comme un nuage de points, chaque point désignant un ensemble. Entre certains ensembles de notre univers existent des flèches qui les relient. Elles symbolisent **l'appartenance**. L'écriture $a \in b$ signifie que "l'ensemble a appartient à l'ensemble b". On dit aussi que a est un **élément** de b. Aussi étrange que cela puisse paraître au premier abord, répétons qu'absolument *tous* les objets de notre univers sont des ensembles. Ainsi, les éléments d'un ensemble ne peuvent être que des ensemble à leur tour.

Pourtant, lorsqu'on considère un ensemble comme $\{1, 2, 3\}$ par exemple, on n'a guère envie de considérer ses trois éléments comme des ensembles. Mais ce serait oublier comment les entiers naturels sont obtenus. Nous verrons plus loin dans l'année (cf. compléments du chapitre "Réels et suites") qu'ils peuvent être vus eux-mêmes comme des ensembles. C'est d'ailleurs ce qui a fait naître le slogan : "Dieu créa l'ensemble vide, les hommes firent le reste".

Enfin, notons que l'usage de $\forall$ dans les phrases quantifiées sera légèrement différent : d'habitude, un tel quantificateur est toujours suivi du symbole $\in$ (par exemple : $\forall x \in \mathbb{R}$, $\forall n \in \mathbb{Z}$, etc.). Ici, ce n'est pas la peine, puisque tous les ensembles appartiennent à un même univers (inutile donc le préciser).

Définition 2.47 (Inclusion)**.** L'**inclusion** de deux ensembles est définie par :

$$\forall a,\ \forall b,\ a \subset b \iff (\forall c\ \ c \in a \implies c \in b)$$

Axiome 2.1 (Axiome d'extensionnalité)**.** Si deux ensembles ont les mêmes éléments, alors ils sont égaux :

$$\forall a,\ \forall b,\ (\forall x,\ x \in a \iff x \in b) \implies a = b$$

Axiome 2.2 (Axiome de la paire)**.** Étant donnés deux ensembles a et b, il existe un ensemble qui a pour éléments a et b, et eux seuls :

$$\forall a, \forall b,\ \exists c,\ (\forall x,\ x \in c \iff (x \in a \ \text{ ou } \ x \in b))$$

Notons que l'ensemble c est unique d'après l'axiome d'extensionnalité. On le note $\{a, b\}$. Si $a \neq b$, on dit que c'est une **paire**. A l'inverse, pour tout a, on pose $\{a\} := \{a, a\}$ et on dit que c'est un **singleton**.

Définition 2.48 (Couple)**.** Soit a et b des ensembles. Le **couple** (a, b) est défini par

$$(a, b) := \{\{a\}, \{a, b\}\}$$

.

Lemme 2.3. *Si* $\{x, y\} = \{x', y'\}$*, alors on a* $\begin{cases} x = x' \\ y = y' \end{cases}$ *ou* $\begin{cases} x = y' \\ y = x' \end{cases}$.

Démonstration. Dans un premier temps, supposons que $x = y$. Alors $x' \in \{x', y'\} = \{x\}$ donc $x' = x$. De même, $y' = x$. Finalement, on a $x = x' = y = y'$ d'où le résultat.

A partir de maintenant, supposons que $x \neq y$ et $x' \neq y'$. On a $x \in \{x, y\} = \{x', y'\}$ donc $x = x'$ ou $x = y'$. De même, $y = x'$ ou $y = y'$. Maintenant :

- Si $x = x'$, comme $x \neq y$, on ne peut pas avoir $y = x'$ donc $y = y'$.
- Si $x = y'$, le raisonnement est symétrique.

Ainsi, la preuve est achevée. □

Corollaire 2.4. *On a* $(a, b) = (a', b')$ *si, et seulement si,* $\begin{cases} a = a' \\ b = b' \end{cases}$.

Démonstration. Le sens réciproque est trivial. Quant au sens direct, si $(a, b) = (a', b')$, alors $\{\{a\}, \{a, b\}\} = \{\{a'\}, \{a', b'\}\}$. Appliquons le lemme :

- Si $\{a\} = \{a'\}$ et $\{a, b\} = \{a', b'\}$, la première égalité fournit $a = a'$. Puis la seconde aboutit à deux sous-cas :
  - Si $a = a'$ et $b = b'$ nous avons terminé.
  - Si $a = b'$ et $b = a'$, alors $a = a' = b = b'$ et nous avons terminé.
- Si $\{a\} = \{a', b'\}$ et $\{a'\} = \{a, b\}$, alors $a = a' = b = b'$ et nous avons terminé.

Ainsi, la preuve est achevée. □

Axiome 2.3 (Axiome de la réunion)**.** Si on se donne un ensemble a, il existe un ensemble qui est la réunion des élément de a (rappelons que les éléments de a sont aussi des ensembles) :

$$\forall a,\ \exists b,\ \forall x,\ (x \in b \iff \exists c \in a\, (a \in a \ \text{ et } \ x \in c))$$

On le notera $\bigcup a$. Il est unique d'après l'axiome d'extensionnalité.

Définition 2.49. Si a et b sont deux ensembles, on pose $a \cup b := \bigcup \{a, b\}$.

Définition 2.50. Si $n \geqslant 3$, la notation $\{a_1, \ldots, a_n\}$ désigne en fait $\{a_1, \ldots, a_{n-1}\} \cup \{a_n\}$ (il s'agit donc là d'une définition par récurrence).

Axiome 2.4 (Axiome de l'ensemble des parties)**.** Pour tout ensemble a, il existe un ensemble constitué des parties de a :

$$\forall a,\ \exists b,\ \forall x,\ (x \in b \iff x \subset c)$$

On le notera $\mathcal{P}(a)$. il est unique d'après l'axiome d'extensionnalité.

Axiome 2.5 (Axiome de compréhension)**.** Nous en donnons une version édulcorée, mais suffisante pour notre usage. Soit P une proposition logique ne faisant pas intervenir le symbole b. Si on se donne un ensemble a, alors il existe un sous-ensemble de a constitué des éléments qui vérifient la proposition P :

$$\forall a,\ \exists b,\ \forall x,\ (x \in b \iff (x \in a \ \text{ et } \ P))$$

Encore une fois, b est unique par l'axiome d'extensionnalité. On le notera $\{x \in a \mid P\}$. Pour être tout à fait exact, il n'y a pas *un* mais *des* axiomes de compréhension (un par proposition P, ce qui fait beaucoup !). On parlera donc de plutôt de **schéma d'axiomes de compréhension**.

Comme application, on peut construire le produit cartésien de deux ensembles. En effet si $a \in A$ et $b \in B$, alors $(a, b) = \{\{a\}, \{a, b\}\} \in \mathcal{P}(\mathcal{P}(A \cup B))$.

Définition 2.51 (Produit cartésien)**.** Si on se donne A et B deux ensembles, leur **produit cartésien** est défini par :

$$A \times B := \{z \in \mathcal{P}(\mathcal{A} \cup \mathcal{B}) \mid \exists a,\ \exists b,\ (a \in A \ \text{ et } \ b \in B \ \text{ et } \ z = (a, b))\}$$

Comme autre application (parfois introduite comme un axiome, mais ici cela s'en déduit de l'axiome de l'existence d'un ensemble), on a la :

Définition 2.52 (Ensemble vide)**.** On sait qu'il existe au moins un ensemble a. On pose :

$$\emptyset := \{x \in a \mid x \neq x\}$$

Remarquons que $\emptyset$ ne peut contenir aucun élément. Notons aussi que $\emptyset$ ne dépend pas du choix de a. En effet, si on se donne un autre ensemble b, et qu'on définit $\emptyset' := \{x \in b \mid x \neq x\}$, montrons que $\emptyset = \emptyset'$. Par extensionnalité, il suffit de montrer $\forall x,\ x \in \emptyset \iff x \in \emptyset'$. or, les deux propositions sont systématiquement fausses, donc elles sont bien équivalentes. On peut alors nommer $\emptyset$ **l'ensemble vide**.

Définition 2.53. Si x et y sont deux ensembles, il existe un unique ensemble z tel que

$$\forall t,\ \{t \in z \iff (t \in a \ \text{ et } \ t \in b)\}$$

On le note $x \cap y$.

Démonstration. L'unicité vient de l'axiome d'extensionnalité. Pour l'existence, on pose par exemple $x \cap y := \{t \in x \mid t \in y\}$. □

Citons un dernier axiome, qui ne fait pas à proprement parler partie des axiomes de Zermelo-Fraenkel :

Axiome 2.6 (Axiome de la fondation)**.** Dans tout ensemble x, il existe un élément y qui est disjoint de x :

$$\forall x,\ \exists y,\ (y \in x \ \text{ et } \ x \cap y = \emptyset)$$

Comme conséquence, on note qu'aucun ensemble ne peut appartenir à lui-même (si $x \in x$, appliquer l'axiome de fondation à $\{x\}$ pour obtenir une contradiction).

Il manque un axiome important, dit **axiome de l'infini**, que nous reverrons en détail dans les compléments du chapitre "Réels et suites" dans le retour sur la construction de $\mathbb{N}$.

Chapitre 3

Introduction à l'algèbre

1 Relations

Définition 3.1 (Relation binaire)**.** Soit E un ensemble. Une **relation** (binaire) $\mathcal{R}$ indique si deux éléments sont "liés". On dira que x et y **sont en relation**, et on écrira $x\mathcal{R}y$.

Exemple 3.1. La relation "x et y ont la même parité" sur $\mathbb{N}$, ou la relation "$x \leqslant y$" sur $\mathbb{R}$.

Remarque 3.1. Formellement, on peut définir $\mathcal{R}$ comme une partie de E^2 qui correspond aux couples qui sont en relations. mais il importe surtout de retenir deux choses :

- A part les relations usuelles, comme celles de l'exemple ci-dessus, on peut définir des relations abstraites. Dans ce cas, typiquement, on définira $\mathcal{R}$ sous la forme :

  $$\forall (x,y) \in E^2,\ x\mathcal{R}y \iff p(x,y)$$

  où $p(x,y)$ est une assertion qui dépend de x et y. On trouvera quelques exemples dans la feuille d'exercices.
- Pour montrer que deux relations $\mathcal{R}$ et $\mathcal{R}'$ sont égales, on montre que

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

Définition 3.2 (Relation d'ordre)**.** Une **relation d'ordre** vérifie les trois axiomes fondamentaux qui suivent :

1. Elle est **réflexive** : $\forall x \in E,\ x\mathcal{R}x$.
2. Elle est **antisymétrique** : si $x\mathcal{R}y$ et $y\mathcal{R}x$, alors $x = y$.
3. Elle est **transitive** : si $x\mathcal{R}y$ et $y\mathcal{R}z$, alors $x\mathcal{R}z$.

Si la relation permet de comparer deux éléments quelconques - autrement dit, si

$$\forall (x,y) \in E^2,\ x\mathcal{R}y \text{ ou } y\mathcal{R}x$$

elle est dite **totale**. Sinon, elle est dite **partielle**.

Exemple 3.2. La relation $\leqslant$ est une relation d'ordre totale sur $\mathbb{N}$. En revanche, la relation $\subset$ est une relation d'ordre partielle sur $\mathcal{P}(\mathbb{N})$, car par exemple, l'ensemble des entiers pairs et l'ensemble des entiers impairs ne peuvent être comparés : aucun n'est inclus dans l'autre. On parle aussi d'ensemble **totalement ordonnée** ou d'ensemble **partiellement ordonné**.

Remarque 3.2. Si E est muni de la relation d'ordre $\leqslant$, on définit les relations $<$, $\geqslant$ et $>$ à partir de $\leqslant$ de la manière suivante :

$$\forall (x,y) \in E^2,\ \begin{cases} x < y \iff x \leqslant y \text{ et } x \neq y \\ x \geqslant y \iff y \leqslant x \\ x > y \iff y \leqslant x \text{ et } x \neq y \end{cases}$$

$<$ et $>$ ne sont pas des relations d'ordre, puisqu'elles ne sont pas réflexives (sauf si $E = \emptyset$). En revanche, on vérifie sans peine que $\geqslant$ est une relation d'ordre. Elle est d'ailleurs total si, et seulement si, $\leqslant$ est totale.

Remarque 3.3. On a le résultat suivant : si $x_1 \leqslant x_2 \leqslant \ldots \leqslant x_n$, alors $x_1 \leqslant x_n$, avec égalité si, et seulement si, tous les x_i son égaux (récurrence sur $n \in \mathbb{N}^\star$). En particulier, il suffit qu'une seule des inégalités $x_i \leqslant x_{i+1}$ soit stricte pour que $x_1 < x_n$.

Définition 3.3 (Partie majorée, minorée, majorant, minorant)**.** Soit E un ensemble muni d'une relation d'ordre qu'on notera $\leqslant$. Alors une partie A de E est **majorée** quand elle possède un **majorant** M :

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

Elle est **minorée** quand elle possède un **minorant** m :

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

Remarque 3.4. Bien que notée $\leqslant$, cette relation d'ordre pourrait très bien être l'inclusion. En pratique, ce seront les deux seuls cas que nous considérerons.

Définition 3.4 (Maximum, minimum)**.** Il faut noter que m ou M n'appartient pas nécessairement à A. Si $m \in A$, on dit que c'est un **plus petit élément**, ou un **minimum**. Si $M \in A$, on dit que c'est un **plus grand élément**, ou **maximum**.

Exemple 3.3. $\mathbb{R}_+^\star$ possède plusieurs minorants (par exemple 0, ou encore -1), mais il ne possède pas de plus petit élément. En effet, s'il possédait un plus petit élément x, on aurait $x \in \mathbb{R}_+^\star$ donc $x > 0$, si bien que $\dfrac{x}{2}$ serait dans $\mathbb{R}_+^\star$, tout en étant strictement inférieur à x, ce qui constituerait une contradiction.

Proposition 3.1 (Unicité du maximum et du minimum)**.** *S'il existe, le maximum est unique. De même avec un minimum.*

Théorème 3.1. *Soit $(E, \leqslant)$ un ensemble totalement ordonné. Alors toute partie finie non vide de E admet un minimum et un maximum.*

Remarque 3.5. Si E totalement ordonné est de cardinal n, on montre par récurrence qu'il existe une unique bijection $\sigma : [\![1, n]\!] \to E$ strictement croissante. En posant $x_i = \sigma(i) \in E$, les éléments de E sont égaux à :

$$x_1 < \ldots < x_n$$

On dit qu'on a **ordonné** les éléments de E.

Définition 3.5. Si $x_1, \ldots, x_n$ sont des éléments de E totalement ordonné (avec $n \geqslant 1$), le maximum et le minimum de l'ensemble fini $\{x_1, \ldots, x_n\}$ sont usuellement notés $\max(x_1, \ldots, x_n)$ et $\min(x_1, \ldots, x_n)$. On peut aussi les noter $\max\limits_{1 \leqslant i \leqslant n} x_i$ et $\min\limits_{1 \leqslant i \leqslant n} x_i$

Proposition 3.2. *Le maximum et le minimum sont associatifs au sens suivant :*

$$\max\left(\max(x_1, \ldots, x_n), x_{n+1}\right) = \max(x_1, \ldots, x_{n+1})$$

$$\min\left(\min(x_1, \ldots, x_n), x_{n+1}\right) = \min(x_1, \ldots, x_{n+1})$$

Définition 3.6 (Élément maximal, élément minimal)**.** Supposons désormais que $(E, \leqslant)$ n'est pas nécessairement totalement ordonné. On dit qu'un élément M est **maximal** lorsque pour tout $x \in E$, si x et M sont comparables, alors $x \leqslant M$. De manière équivalente, on peut écrire :

$$\forall x \in E,\ x \geqslant M \implies x = M$$

ou même

$$\forall x \in E,\ x \geqslant M \implies x \leqslant M$$

Encore autrement dit, on n'a jamais $x > M$. Tout maximum est un élément maximal, mais la réciproque est fausse. De plus, un élément maximal n'est pas nécessairement unique. De même, on dit que m est **minimal** lorsque :

$$\forall x \in E,\ x \leqslant m \implies x = m$$

Définition 3.7 (Relation d'équivalence)**.** Une **relation d'équivalence** vérifie les trois axiomes fondamentaux qui suivent :

1. Elle est **réflexive** : $\forall x \in E,\ x \mathcal{R} x$.
2. Elle est **symétrique** : si $x \mathcal{R} y$ alors $y \mathcal{R} x$.
3. Elle est **transitive** : si $x \mathcal{R} y$ et $y \mathcal{R} z$, alors $x \mathcal{R} z$.

Pour chaque élément x, on peut considérer l'ensemble de tous les éléments en relation avec x (notons que c'est grâce à l'axiome de symétrie que la notion de "être en relation" a du sens). On l'appelle la **classe d'équivalence** de x et on la note $\overline{x}$. Une classe d'équivalence (sans préciser de quel x) est une partie A de E telle que $\exists x \in E,\ A = \overline{x}$.

Proposition 3.3. *Une classe d'équivalence peut être considérée comme la classe d'équivalence de n'importe lequel de ses éléments.*

Définition 3.8 (Ensemble-quotient, surjection canonique)**.** Si $\mathcal{R}$ est une relation d'équivalence sur E, l'ensemble des classes d'équivalences de E est appelé l'**ensemble-quotient** de E par $\mathcal{R}$. On le note souvent $E/\mathcal{R}$. L'application $x \mapsto \overline{x}$ est automatiquement surjective, on l'appelle la **surjection canonique** de E dans $E/\mathcal{R}$.

Définition 3.9 (Partition)**.** Une **partition** d'un ensemble E est un ensemble $\mathcal{A}$ de parties de E tel que :

- Aucune partie n'est vide : $\forall A \in \mathcal{A},\ A \neq \emptyset$.
- Deux parties distinctes sont toujours disjointes : $\forall (A, B) \in \mathcal{A}^2,\ A \neq B \implies A \cap B = \emptyset$

- La réunion des parties vaut E : $E = \bigcup_{A \in \mathcal{A}} A$

Les deux derniers points peuvent se réécrire

$$E = \bigsqcup_{A \in \mathcal{A}} A$$

Ils signifient que tout élément de E appartient à un et un seul A. Finalement, une partition est un recouvrement disjoint dont aucun élément n'est vide.

Exemple 3.4. $\mathbb{N}$ est partitionné entre entiers pairs et impairs. $\mathbb{R}$ est partitionné entre réels strictement positifs, réels strictement négatifs et $\{0\}$.

Théorème 3.2. *Soit $\mathcal{R}$ une relation d'équivalence sur un ensemble E. Alors, ses classes d'équivalence forment une partition de E.*

Remarque 3.6. On peut montrer par analyse-synthèse que, réciproquement, pour toute partition X de E, il existe une unique relation d'équivalence dont les classes sont exactement les éléments de X.

2 Lois de composition

Dans toute cette partie, nous nous donnons un ensemble de base E. Nous verrons dans les chapitres d'algèbre que E peut par exemple être un groupe, un anneau, un corps, un espace vectoriel...

Définition 3.10 (Loi de composition interne, stabilité, structure algébrique)**.** Une **loi de composition interne** sur E est une application

$$\begin{array}{rccc} * & : & E^2 & \longrightarrow & E \\ & & (x,y) & \longmapsto & x*y \end{array}$$

La loi est dite interne parce qu'on "reste dans E en mélangeant deux éléments de E". on dit aussi que E est **stable** par $*$. Le couple $(E,*)$ est appelé une **structure algébrique**.

Remarque 3.7. Nous verrons ultérieurement qu'une structure algébrique peut comporter plusieurs lois.

Exemple 3.5. L'addition est une loi de composition interne sur $\mathbb{R}_+$. La soustraction est aussi une loi de composition, mais elle n'est pas interne.

Définition 3.11 (Commutativité)**.** La loi de composition est **commutative** lorsqu'on peut "permuter les facteurs" :

$$\forall (x,y) \in E^2,\ x*y = y*x$$

Exemple 3.6. Sur $\mathbb{R}$, l'addition est commutative, mais pas la soustraction.

Définition 3.12 (Associativité)**.** La loi de composition est **associative** lorsqu'on peut "jouer avec les parenthèses" :

$$\forall (x,y,z) \in E^3,\ (x*y)*z = x*(y*z)$$

Exemple 3.7. Sur $\mathbb{N}$, l'addition et la multiplication sont des lois de composition internes associatives. Sur $\mathbb{Z}$, la soustraction est une loi de composition interne, mais elle n'est pas associative.

Définition 3.13 (Élément neutre)**.** Un **élément neutre** e pour la loi $*$ est un élément qui "ne change rien" :

$$\forall x \in E,\ x * e = e * x = x$$

Proposition 3.4. *S'il existe, l'élément neutre est unique. cela permet de parler de "l'élément neutre" sans ambiguïté.*

Exemple 3.8. Sur $\mathbb{Z}$, l'élément neutre de l'addition est 0, celui de la multiplication est 1. Sur E^E, l'élément neutre de la loi $\circ$ est id_E.

Définition 3.14. Soit $(E, *)$ une structure algébrique, A et B deux parties de E et $x \in E$. On pose respectivement :

$$A * x = \{a * x \mid a \in A\}$$

$$x * B = \{x * b \mid b \in B\}$$

$$A * B = \{a * b \mid (a, b) \in A \times B\}$$

Exemple 3.9. Dans $(\mathbb{Z}, +)$, on a $0 + \mathbb{Z} = \mathbb{Z} + 0 = \mathbb{Z}$. Dans $(\mathbb{Z}, \times)$, l'ensemble des multiples de n est noté $n\mathbb{Z}$.

Exemple 3.10. Si $*$ est associative, on a $(A * B) * C = A * (B * C)$ (chaîne d'équivalences). Si $*$ est commutative, on a $A * B = B * A$ (chaîne d'équivalences).

A partir de maintenant, on suppose que E est muni d'une loi interne associative $*$ et d'un élément neutre e.

Définition 3.15 (Inversibilité)**.** On dit qu'un élément $x \in E$ est inversible lorsqu'il "possède un inverse" :

$$\exists y \in E,\ x * y = y * x = e$$

Proposition 3.5. *S'il existe, l'inverse de x est unique. Cela permet de parler de "l'inverse de x" sans ambiguïté. On parle aussi du* ***symétrique*** *de x.*

Remarque 3.8. A propos de la notation :

- Si la loi est notée $+$, le neutre sera souvent noté 0 et l'inverse de x sera souvent noté $-x$: dans ce cas, $x - y$ désignera en fait $x + (-y)$.
- Dans le cas général, l'inverse de x sera souvent noté x^{-1}.

Exemple 3.11. e est toujours inversible, d'inverse lui-même.

Corollaire 3.1. *Si x est inversible, alors l'inverse de x^{-1} est x. Autrement dit, l'application de passage à l'inverse est involutive.*

Corollaire 3.2. *Soit x et y inversibles. Alors $x * y$ est inversible, d'inverse $y^{-1} * x^{-1}$.*

Remarque 3.9. Attention, il ne faut alors pas oublier d'inverse l'ordre de x et y, sinon cela devient faux dès que la loi n'est pas commutative.

3 Construction des entiers, HP

3.1 Construction de $\mathbb{N}$

Dans cette partie, nous construisons $\mathbb{N}$, ce qui constitue véritablement la pierre de fondation des mathématiques, au moins au niveau des classes préparatoires. Nous allons voir qu'à partir de trois axiomes, qui sont des "axiomes de bon sens" compte tenu de la perception pratique que nous avons de $\mathbb{N}$, on peut formaliser et démontrer quasiment la totalité du programme.

Définition 3.16 (Axiomes fondateurs de $\mathbb{N}$)**.** On pose l'existence d'un ensemble $\mathbb{N}$ non vide, munit d'une relation d'ordre $\leqslant$, tel que :

- Toute partie non vide possède un plus petit élément.
- Toute partie non vide et majorée possède un plus grand élément.
- $\mathbb{N}$ n'admet pas de plus grand élément.

Exemple 3.12. Soit $(u_n)_{n\in\mathbb{N}}$ une suite à valeurs dans $\mathbb{N}$. En considérant l'ensemble

$$u(\mathbb{N}) = \{u_n \mid n \in \mathbb{N}\}$$

on pourra montrer que

- si u est décroissante, alors elle est stationnaire ;
- si u est croissante et majorée, alors elle est stationnaire.

Proposition 3.6. *La relation d'ordre $\leqslant$ est totale.*

Démonstration. Pour comparer m et n, il suffit de considérer l'ensemble $\{m, n\}$, qui est un ensemble non vide. □

Définition 3.17. On pose $[\![n_1, n_2]\!] = \{n \in \mathbb{N} \mid n_1 \leqslant n \leqslant n_2\}$. En particulier, si $n_1 > n_2$, on a $[\![n_1, n_2]\!] = \emptyset$.

En un certain sens, le lemme suivant formalise ce que fait un enfant lorsqu'il compte.

Lemme 3.1. *Soit $n \in \mathbb{N}$. Alors n possède un unique successeur, c'est-à-dire un entier $n^+ > n$ tel que $]\!]n, n^+[\![= \emptyset$.*

Démonstration. Considérer l'ensemble $A = \{k \in \mathbb{N} \mid k > n\}$ qui est une partie non vide de $\mathbb{N}$ et prendre son minimum. Pour montrer que ce minimum convient, raisonner par l'absurde pour l'ensemble discret. Pour l'unicité, prendre un autre candidat et les ordonner SPG. Conclure par un mini-RPA lié à l'ensemble vide précédemment cité. □

Définition 3.18. On note 0 le plus petit élément de $\mathbb{N}$, et 1 son successeur. On définit l'ensemble $\mathbb{N}^\star = \mathbb{N} \setminus \{0\}$. On introduit également : $2 = 1^+$, $3 = 2^+$, $4 = 3^+$, $5 = 4^+$, $6 = 5^+$, $7 = 6^+$, $8 = 7^+$ et $9 = 8^+$.

Théorème 3.3 (Raisonnement par récurrence simple)**.** *Soit P_n une propriété indexée sur $\mathbb{N}$. On suppose que*

- *Initialisation : P_0 est vraie*
- *Hérédité : $\forall n \in \mathbb{N},\ P_n \implies P_{n+1}$*

Alors, pour tout $n \in \mathbb{N}$, P_n est vraie.

Démonstration. Considérer $A = \{n \in \mathbb{N} \mid P_n \text{ est fausse}\}$ et raisonner par l'absurde en supposant que A est non vide. Considérer alors son minimum, qui est supérieur ou égal à 1. Le minimum moins 1 n'est donc pas dans A, mais alors par hérédité le minimum n'est pas non plus dans A. Absurde. □

Le théorème qui suit est fondamental, il est la pierre de fondation d'un nombre incalculable de constructions mathématiques.

Théorème 3.4 (Définition d'une suite par récurrence, admis)**.** *Soit E un ensemble, $x \in E$ et $f : E \to E$. Alors il existe une unique suite $(u_n)_{n\in\mathbb{N}}$ telle que*

$$\begin{cases} u_0 = x \\ \forall n \in \mathbb{N},\ u_{n^+} = f(u_n) \end{cases}$$

Comme application, nous allons à présent définir une addition et une multiplication sur $\mathbb{N}$. Nous verrons par la suite comment les prolonger sur $\mathbb{Z}$, $\mathbb{Q}$ et enfin $\mathbb{R}$. On prolongera finalement à $\mathbb{C}$.

Définition 3.19 (Définition de l'addition)**.** Pour $m \in \mathbb{N}$, on définit $m + n$ par récurrence sur n :

$$\begin{cases} m + 0 = m \\ \forall n \in \mathbb{N},\ m + n^+ = (m + n)^+ \end{cases}$$

Remarque 3.10 (Changement de notation)**.** Notons que pour tout $n \in \mathbb{N}$, on a

$$n^+ = (n + 0)^+ = n + 0^+ = n + 1$$

A partir de maintenant, on abandonnera la notation n^+ au profit de la notation $n+1$. Typiquement, le principe de récurrence sera écrit sous la forme $P_n \implies P_{n+1}$, et la construction d'une suite par récurrence s'écrira $u_{n+1} = f(u_n)$.

La petite proposition qui suit est d'une importance capitale dans nombre d'inégalités, en particulier en analyse.

Proposition 3.7 (Théorème de Roger, ou théorème de la licorne, NS)**.** *Si $m < n+1$, alors $m \leqslant n$.*

Démonstration. Si on avait $m > n$, alors m viendrait s'intercaler entre n et $n^+ = n + 1$. □

On démontre sans trop de difficulté que l'addition est associative et commutative, et qu'elle est compatible avec $\leqslant$ au sens suivant.

Démonstration. Pour tout $(m, n, p) \in \mathbb{N}^3$, on a : $m \leqslant n \implies m + p \leqslant n + p$. □

Voici maintenant la définition de la multiplication par récurrence.

Définition 3.20 (Définition de la multiplication)**.** Pour $m \in \mathbb{N}$ fixé, on définit $m \times n$ par récurrence sur n :

$$\begin{cases} m \times 0 = 0 \\ \forall n \in \mathbb{N},\ m \times (n + 1) = (m \times n) + m \end{cases}$$

Là aussi, on démontre assez facilement que la multiplication est associative, commutative et distributive sur l'addition. On montre également sa compatibilité avec $\leqslant$ au sens suivant.

Proposition 3.8. *On a : $\forall (m, n, p) \in \mathbb{N}^3,\ m \leqslant n \implies m \times p \leqslant n \times p$*

3.2 Extension à $\mathbb{Z}$

Jusqu'ici, nous avons construit les entiers naturels, leur addition et leur multiplication. Mais nous savons bien que dans $\mathbb{N}$, aucun entier non nul n'est inversible pour l'addition. D'où l'idée de construire une structure plus riche, contenant $\mathbb{N}$, dans laquelle on puisse inverse ces entiers pour obtenir ce qu'on appelle des "entiers négatifs".

Une telle construction est proposée en annexe. Donnons sans démonstration les principaux résultats qui en découlent.

- Il existe un ensemble $\mathbb{Z}$ contenant $\mathbb{N}$, dans lequel on peut prolonger l'addition et la multiplication définies sur $\mathbb{N}$. Ces deux lois gardent leurs propriétés (associativité, commutativité, distributivité). De plus, tout élément de $\mathbb{Z}$ est inversible pour l'addition.
- La relation d'ordre $\leqslant$ définie peut être prolongée sur $\mathbb{Z}$. Elle reste totale, et elle reste compatible avec l'addition :
$$\forall (m,n,p) \in \mathbb{Z}^3,\ m \leqslant n \implies m+p \leqslant n+p$$
Cela montre que dans une addition, on peut "faire changer les termes de côté" de part et d'autres du signe $\leqslant$ ou du signe $\geqslant$ à condition d'inverser le signe. Elle est également compatible avec la multiplication au sens suivant :
$$\forall (m,n,p) \in \mathbb{Z}^3,\ \left\{ \begin{array}{l} m \leqslant n \\ p \geqslant 0 \end{array} \right. \implies m \times p \leqslant n \times p$$
- Tout élément de $\mathbb{Z}$ s'écrit d'une unique manière 0, n ou $-n$ avec $n \in \mathbb{N}^\star$. Cette écriture sous la forme $\pm n$ s'appelle usuellement un **entier relatif**. C'est celle que nous garderons systématiquement à partir de maintenant.
- Enfin, on peut généraliser le résultat sur les inégalités strictes d'entiers :
$$\forall (m,n) \in \mathbb{Z}^2,\ m < n+1 \implies m \leqslant n$$

4 Sommes et produits

4.1 Retour sur les formules sommatoires

Définition 3.21. Soit $(x_k)_{k\in\mathbb{N}}$ une famille de complexes, et $m \in \mathbb{N}$ fixé. On définit $\displaystyle\sum_{k=m}^{n} x_k$ et $\displaystyle\prod_{k=m}^{n} x_k$ par récurrence sur $n \geqslant m$.

- Si $n = m$, on les définit comme étant égal à x_k.
- Et pour tout $n \geqslant m$, on pose $\displaystyle\sum_{k=m}^{n+1} x_k = \left(\sum_{k=m}^{n} x_k\right) + x_{n+1}$ et $\displaystyle\prod_{k=m}^{n+1} x_k = \left(\prod_{k=m}^{n} x_k\right) \times x_{n+1}$

Si $n = m-1$, on pose $\displaystyle\sum_{k=m}^{n} x_k = 0$ et $\displaystyle\prod_{k=m}^{n} x_k = 1$ par convention. Et si $n < m-1$, on évite de définir ces grandeurs.

Remarque 3.11. Par récurrence, on peut démontrer les relations admises dans le chapitre sur les complexes : séparation et regroupement de termes, distributivité, relation de Chasles. A chaque fois, il s'agit d'une récurrence rapide.

Exemple 3.13. Soit $n \in \mathbb{N}$. Par une récurrence immédiate, on a $\sum_{k=1}^{n} 1 = n$ et $\prod_{k=1}^{n} 1 = 1$. Soit $\lambda \in \mathbb{C}$. Par distributivité, on a $\sum_{k=1}^{n} \lambda = \lambda n$ et $\prod_{k=1}^{n} \lambda = \lambda^n$.

Grâce à la commutativité, on se dit intuitivement que nous pouvons sommer les c_k dans l'ordre que nous le souhaitons. Cela permet en particulier d'indexer les éléments sur un autre ensemble que $[\![1, n]\!]$, et de donner un sens à des formules comme $\sum_{\substack{k=1 \\ k \text{ pair}}}^{n} x_k$ ou $\sum_{1 \leqslant i < j \leqslant n} x_{i,j}$.

Définition 3.22. Soit I un ensemble **fini** d'indices, et $(x_i)_{i \in I}$ une famille de complexes indexée sur I. Puisque I est fini, il existe une bijection σ d'un certain $[\![1, n]\!]$ sur I. On pose alors :

$$\sum_{i \in I} x_i = \sum_{k=1}^{n} x_{\sigma(k)} \quad \text{et} \quad \prod_{i \in I} x_i = \prod_{k=1}^{n} x_{\sigma(k)}$$

Remarque 3.12. On admet que le résultat final reste le même si on change de choix pour σ.

Exemple 3.14. Voici quelques cas très simples.

- Si $I = \emptyset$, on a $\sum_{i \in I} x_i = 0$ et $\prod_{i \in I} x_i = 1$.
- Si $I = \{a\}$, on a $\sum_{i \in I} x_i = x_a$ et $\prod_{i \in I} x_i = x_a$.
- Si $I = \{a, b\}$ avec $a \neq b$, on a $\sum_{i \in I} x_i = x_a + x_b$ et $\prod_{i \in I} x_i = x_a x_b$.
- Si $I = [\![1, n]\!]$, on a $\sum_{i \in I} x_i = \sum_{k=1}^{n} x_k$ et $\prod_{i \in I} x_i = \prod_{k=1}^{n} x_k$.
- Si la famille $(x_i)_{i \in I}$ est constante égale à λ, on a $\sum_{i \in I} x_i = \lambda \operatorname{Card}(I)$ et $\prod_{i \in I} x_i = \lambda^{\operatorname{Card}(I)}$.

Théorème 3.5. *Soit I est J deux ensembles finis, $\varphi : J \to I$ une bijection, et $(x_i)_{i \in I}$ une famille de complexes. Alors, on a :*

$$\sum_{i \in I} x_i = \sum_{j \in J} x_{\varphi(j)} \quad \textit{et} \quad \prod_{i \in I} x_i = \prod_{j \in J} x_{\varphi(j)}$$

Autrement dit, on "interprète i comme un $\varphi(j)$".

Les formules de séparation et regroupement de termes ainsi que de distributivité restent évidemment vraies. L'analogue de la relation de Chasles serait le

Théorème 3.6 (Associativité de la somme et du produit)**.** *Si $I_1, \ldots, I_n$ sont des parties deux à deux disjointes de I, alors on a :*

$$\sum_{i\in I_1\cup\ldots\cup I_n} = \sum_{i\in I_1} x_i + \ldots + \sum_{i\in I_n} x_i \quad et \quad \prod_{i\in I_1\cup\ldots\cup I_n} = \left(\prod_{i\in I_1} x_i\right) \times \ldots \times \left(\prod_{i\in I_n} x_i\right)$$

Démonstration. Une démonstration se trouve dans les compléments. □

Un particulier est celui des sommes doubles.

Corollaire 3.3 (Théorème de Fubini)**.** *Soit $(x_{i,j})_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}}$ une famille de complexes doublement indexée. Alors :*

$$\sum_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}} x_{i,j} = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} x_{i,j}\right) = \sum_{j=1}^{n}\left(\sum_{i=1}^{m} x_{i,j}\right)$$

et

$$\prod_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}} x_{i,j} = \prod_{i=1}^{m}\left(\prod_{j=1}^{n} x_{i,j}\right) = \prod_{j=1}^{n}\left(\prod_{i=1}^{m} x_{i,j}\right)$$

Remarque 3.13. On retiendra que dans une somme double de termes qui commutent (toujours le cas dans $\mathbb{C}$ donc), on peut sommer les termes dans l'ordre que l'on veut. Avec le même raisonnement, on généralise immédiatement à des sommes indexées sur des ensembles finis quelconques. Idem avec le produit.

Exemple 3.15. On veut calculer $\sum_{1\leqslant i,j\leqslant n} \min(i,j)$. On peut choisir de sommer d'abord selon i ou d'abord selon j. Commençons par exemple avec i. On obtient :

$$\begin{aligned}
\sum_{1\leqslant i,j\leqslant n} \min(i,j) &= \sum_{i=1}^{n}\left(\sum_{j=1}^{n} \min(i,j)\right)\\
&= \sum_{i=1}^{n}\left(\sum_{j=1}^{i} j + \sum_{j=i+1}^{n} i\right)\\
&= \sum_{i=1}^{n}\left(\frac{i(i+1)}{2} + (n-i)i\right)\\
&= -\frac{1}{2}\sum_{i=1}^{n} i^2 + \left(n+\frac{1}{2}\right)\sum_{i=1}^{n} i\\
&= -\frac{n(n+1)(2n+1)}{6} + \frac{(2n+1)n(n+1)}{4}\\
&= \frac{n(n+1)(2n+1)}{6}
\end{aligned}$$

De manière imagée, on vient de sommer les termes $x_{i,j}$ pour (i,j) variant dans un carré de taille $n \times n$, et le résultat est le même que l'on somme par les lignes ou par les colonnes. Le théorème de Fubini donne un résultat légèrement plus général, lorsque (i,j) parcourt un rectangle de taille $m \times n$. Mais l'associativité de la somme et du produit permet de traiter des cas plus complexes, comme celui où (i,j) varie dans un triangle.

Dans de tels cas, on peut commencer par fixer l'une des deux variables, mettons i, puis observer dans quelle plage (en fonction de i) varie j. Il est plus que recommandé de représenter graphiquement sur un plan le domaine dans lequel varie (i,j).

Exemple 3.16. On souhaite calculer $\displaystyle\sum_{1\leqslant i\leqslant j\leqslant n} ij$. On a :

$$\begin{aligned}
\sum_{1\leqslant i\leqslant j\leqslant n} ij &= \sum_{i=1}^{n}\left(i\sum_{j=i}^{n} j\right)\\
&= \sum_{i=1}^{n}\left[i\left(\frac{n(n+1)}{2}-\frac{(i-1)i}{2}\right)\right]\\
&= \frac{n(n+1)}{2}\sum_{i=1}^{n} i-\frac{1}{2}\sum_{i=1}^{n} i^3+\frac{1}{2}\sum_{i=1}^{n} i^2\\
&= \left(\frac{n(n+1)}{2}\right)^2-\frac{1}{2}\left(\frac{n(n+1)}{2}\right)^2+\frac{n(n+1)(2n+1)}{12}\\
&= \frac{n(n+1)}{2}\left[\frac{n(n+1)}{4}+\frac{2n+1}{6}\right]\\
&= \frac{n(n+1)}{2}\times\frac{3n^2+7n+2}{12}\\
&= \frac{n(n+1)(n+1)(3n+1)}{24}
\end{aligned}$$

4.2 Itération d'une loi de composition interne

Dans toute cette partie, on fixe une structure algébrique $(E,*)$ associative, admettant un élément neutre e.

Définition 3.23. Soit $x \in E$. On définit x^n par récurrence de la manière suivante :

$$\begin{cases} x^0 = e\\ \forall n\in\mathbb{N},\ x^{n+1} = x^n * x\end{cases}$$

Remarque 3.14. Plus généralement, on définira par récurrence toute expression du type

$$\prod_{k=1}^{n} x_k = x_1 * \ldots * x_n$$

- En théorie, il s'agit donc de

$$(((x_1 * x_2) * x_3) * \ldots * x_{n-1}) * x_n$$

Mais en fait, puisque $*$ est associative, on peut placer les parenthèses "où l'on veut !

$$x_1 * \ldots * x_n = x_1 * \ldots * x_k * (x_{k+1} * \ldots * x_m) * x_{m+1} * \ldots * x_n$$

Cela se démontre rigoureusement par récurrence.

- Autre exemple : donnons-nous $A_1, \ldots, A_n \subset E$. On peut alors considérer

$$A_1 * \ldots * A_n$$

au sens de la structure algébrique $(\mathcal{P}(E), *)$. Une récurrence immédiate montre que

$$A_1 * \ldots * A_n = \{x_1 * \ldots * x_n \mid (x_1, \ldots, x_n) \in A_1 \times \ldots \times A_n\}$$

Remarque 3.15. Si de plus la loi est commutative, on peut définir des expressions comme $\prod_{i \in I} x_i$ avec I fini, et l'ensemble des résultats précédents restent vrais. Par exemple, on pourra penser à l'union ou à l'intersection d'un sur les parties d'un ensemble.

Proposition 3.9. *Soit $x \in E$. On a :*

$$\forall (m,n) \in \mathbb{N}^2, \ \begin{cases} x^m * x^n = x^{m+n} \\ (x^m)^n = x^{mn} \end{cases}$$

Démonstration. Fixer m et raisonner par récurrence sur n. □

Remarque 3.16. En revanche, attention : si x et y ne commutent pas, on n'a pas nécessairement $(x * y)^n = x^n * y^n$.

Proposition 3.10. *Soit $(x, y) \in E^2$. Si $x * y = y * x$, alors :*

$$\begin{cases} \forall (m,n) \in \mathbb{N}^2, \ x^m * y^n = y^n * x^m \\ \forall n \in \mathbb{N}, \ (x * y)^n = x^n * y^n \end{cases}$$

Démonstration. Pour la première, fixer m et raisonner par récurrence sur n. Pour la seconde, raisonner par récurrence sur n. □

On généralise désormais l'inverse d'un produit d'éléments inversibles.

Proposition 3.11. *Soit $x_1, \ldots, x_n$ inversibles. Alors $x_1 * \ldots * x_n*$ est inversible, d'inverse*

$$x_n^{-1} * \ldots x_1^{-1}$$

Démonstration. C'est une récurrence immédiate. □

Exemple 3.17. Soit $x \in E$ inversible. Alors, pour tout $n \in \mathbb{N}^\star$, x^n est inversible, d'inverse $(x^{-1})^n$.

Définition 3.24. Soit $x \in E$ inversible et $n \in \mathbb{N}^\star$. On pose

$$x^{-n} = (x^n)^{-1} = (x^{-1})^n$$

Remarque 3.17. Les relations précédentes continuent à être vraies pour des exposants dans $\mathbb{Z}$. Ainsi, si x est inversible, on a :

$$\forall (m,n) \in \mathbb{Z}^2,\ \begin{cases} x^m * x^n = x^{m+n} \\ (x^m)^n = x^{mn} \end{cases}$$

Si y est inversible et $x * y = y * x$, on a de plus :

$$\begin{cases} \forall (m,n) \in \mathbb{Z}^2,\ x^m * y^n = y^n * x^m \\ \forall n \in \mathbb{Z},\ (x*y)^n = x^n * y^n \end{cases}$$

Les vérifications sont fastidieuses, mais faciles. Il suffit à chaque fois de revenir à la définition.

5 Deux théorèmes d'arithmétique

Théorème 3.7. *Il existe une* ***division euclidienne*** *sur* $\mathbb{Z}$ *:*

$$\forall (a,b) \in \mathbb{Z} \times \mathbb{N}^\star,\ \exists !(q,r) \in \mathbb{Z} \times [\![0, b-1]\!],\ a = bq + r$$

La division euclidienne sur $\mathbb{Z}$ *est d'une utilité considérable en arithmétique et en théorie des groupes.*

Démonstration. Pour l'existence, fixons a et b comme dans l'énoncé et considérons

$$A = \{a - bq \mid q \in \mathbb{Z}\} \cap \mathbb{N}$$

Puisque $b \geqslant 1$, on a $\forall q \leqslant 0,\ a - bq \geqslant a - q$. Donc pour $q = \min(0, a)$, on voit que A est non vide. Posons alors $r = \min(A)$ et soit q associé. On a bien $a = bq + r$ et $r \geqslant 0$. Par ailleurs, on a $r < b$ sans quoi on pourrait écrire

$$r - b = a - b(q+1) \in A$$

ce qui contredirait la minimalité de r. Donc $r \leqslant b - 1$.

Pour l'unicité, si deux couples (q, r) et (q', r') conviennent, supposons que $q > q'$. Alors on aurait $q \geqslant q' + 1$, puis $r' \geqslant r' - r = b(q - q') \geqslant b$, ce qui est absurde. Donc $q \leqslant q'$. Symétriquement, on prouve que $q' \leqslant q$, d'où $q = q'$ par antisymétrie, puis $r = r'$ en réinjectant dans l'égalité $a = a$ avec les deux divisions euclidiennes. □

Théorème 3.8. *Soit* b *un entier supérieur ou égal à* 2*. Alors tout entier* $n > 0$ *possède une unique* ***décomposition en base*** b *:*

$$n = \overline{a_k a_{k-1} \dots a_1 a_0}^b = \sum_{i=0}^{k} a_i b^i$$

où $k \geqslant 0$*,* $\forall i \in [\![0, k]\!],\ 0 \leqslant a_i < b$ *et* $a_k \neq 0$*.*

Démonstration. Il faut raisonner par récurrence et pour l'existence, et pour l'unicité (et j'ai la flemme de le taper parce que c'est pas beau). □

Exemple 3.18. 2014 s'écrit $\overline{11111011110}^2$ en base 2, aussi appelée **numération binaire**.

Exemple 3.19. Très provisoirement, notons à nouveau 9^+ le successeur de 9. En base 9^+, on a donc

$$\overline{10}^{9^+} = 1 \times 9^+ + 0 \times 1 = 9^+$$

Sauf mention expresse du contraire, tous les nombre seront écrits en base 9^+. on pourra donc omettre la barre horizontale au-dessus de $\overline{10}$, et écrire $9^+ = 10$ (ce qui ne serait pas nécessairement vrai dans une autre base : examiner le cas d'une base strictement supérieure à 10 par exemple). On parlera de **base 10** ou de **base décimale**.

6 Compléments, HP

6.1 Démonstration du théorème d'existence / unicité des suites définies par récurrence avec une fonction

Théorème 3.9. *Soit E un ensemble, $x \in E$ fixé et $f : E \to E$. Alors il existe une unique suite $(u_n)_{n\in\mathbb{N}}$ telle que*

$$\begin{cases} u_0 = x \\ \forall n \in \mathbb{N}, u_{n^+} = f(u_n) \end{cases}$$

Démonstration. Pour l'unicité, on suppose que deux suites $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ conviennent, et on démontre par récurrence sur n que $u_n = v_n$. Passons à l'existence, qui est loin d'être aussi évidente qu'il n'y paraît.

- Premier point : pour $n \in \mathbb{N}$, on commence par définir l'ensemble E_n (très grand !) des suites $(v_k)_{k\in\mathbb{N}}$ telles que

  $$\begin{cases} v_0 = x \\ \forall k < n, v_{k^+} = f(v_k) \end{cases}$$

  Autrement dit, E_n est l'ensemble des suites telles que "les termes de 0 à n conviennent". En particulier, E_0 est l'ensemble des suites telles que $v_0 = x$.
- Deuxième point : aucun E_n n'est vide. Prouvons le résultat par récurrence sur n.
  - Initialisation : $E_0 \neq \emptyset$. En effet, la suite constante égale à x convient.
  - Hérédité : si $E_n \neq \emptyset$, considérons une suite $(v_k)_{k\in\mathbb{N}}$ dans cet ensemble. En définissant la suite $(v'_k)_{k\in\mathbb{N}}$ par

    $$\begin{cases} \forall k \neq n^+, v'_k = v_k \\ v'_{n^+} = f(v_n) \end{cases}$$

    on obtient alors un élément de E_{n^+}.
- Troisième point : si $(v_k)_{k\in\mathbb{N}}$ et $(v'_k)_{k\in\mathbb{N}}$ appartiennent à E_n, alors $v_n = v'_n$. Il s'agit d'une récurrence facile sur n.
- Dernier point : on construit alors notre suite $(u_n)_{n\in\mathbb{N}}$ comme suit. Pour $n \in \mathbb{N}$ fixé, on choisit $(v_k)_{k\in\mathbb{N}}$ dans E_n (ce qui est possible puisque $E_n \neq \emptyset$) puis on pose $u_n = v_n$. Le point précédent montre que le résultat ne dépend pas du choix de (v_k). Ainsi, (u_n) est bien définie. Montrons maintenant qu'elle satisfait les conditions demandées.
  - Déjà, par construction on a bien $u_0 = x$.

– Puis si $n > 0$, montrons que $f(u_n) = u_{n^+}$. Soit $(v_k)_{k\in\mathbb{N}} \in E_{n^+}$ quelconque. Par construction on a que $E_{n^+} \subset E_n$. Ainsi, pour définir u_n et $u_{n^+} = v_{n+1}$, il est légitime de choisir $(v_k)_{k\in\mathbb{N}} \in E_n$, et on obtient $u_n = v_n$. Enfin, par définition de E_{n^+} on a $v_{n^+} = f(v_n)$ d'où $u_{n^+} = f(u_n)$.

Ainsi la preuve est achevée. □

À présent, précisons un peu la notion de suite définie par récurrence forte. Pour plus de simplicité dans les notations, nous indexerons nos suites sur $\mathbb{N}^*$. Intuitivement, il s'agit d'initialiser une suite $(u_n)_{n\in\mathbb{N}^*}$ à $u_1 = x \in E$, puis pour $n \in \mathbb{N}^*$ de définir u_{n+1} en fonction de $u_1, ..., u_n$. De manière formelle, il faudrait pouvoir écrire $u_{n+1} = f(u_1, ..., u_n)$ pour une application f bien choisie, dont le nombre d'arguments serait variable. Tout cela a un sens, car on rappelle que $F = \bigcup_{n\in\mathbb{N}^*} E^n$ est bien défini (cf. l'exemple sur les alphabets du chapitre "Ensembles et applications").

De plus on pourra parler de la taille d'un élément de F : il s'agit de l'unique entier $n \in \mathbb{N}^*$ tel que $\varphi \in E^n$. Dans le reste du paragraphe, on le notera $T(\varphi)$.

Revenons à notre problème : on définit une application $f : F \to E$, puis on cherche à prouver l'existence et l'unicité d'une suite $(u_n)_{n\in\mathbb{N}^*}$ vérifiant

$$\begin{cases} u_1 = x \\ \forall n \in \mathbb{N}^*, u_{n+1} = f[(u_1, ..., u_n)] \end{cases}$$

On pourra noter $f(u_1, ..., u_n)$ plutôt que $f[(u_1, ..., u_n)]$.

L'unicité se montre facilement par récurrence forte. Pour l'existence, commençons par définir une opération de concaténation : si $\varphi = (x_1, ..., x_n) \in F, x \in E$, on pose

$$\varphi \sqcup x = (x_1, ..., x_n, x)$$

Formellement, il s'agit de l'unique prolongement de φ à $[\![1, T(\varphi)+1]\!]$ qui vaut x en $T(\varphi)+1$. Ensuite, appliquons le théorème démontré précédemment en posant

$$\begin{cases} \varphi_1 = (x) \\ \forall n \in \mathbb{N}^*, \varphi_{n+1} = \varphi_n \sqcup f(\varphi_n) \end{cases}$$

Par une récurrence immédiate, on montre que $T(\varphi_n) = n$. Si on note $u_n = \varphi_n(n)$ le dernier terme du n-uplet φ_n, une autre récurrence immédiate montre que $\forall n \in \mathbb{N}^* \varphi_n = (u_1, ..., u_n)$. On en déduit que la suite $(u_n)_{n\in\mathbb{N}^*}$ convient.

6.2 Retour sur la construction de $\mathbb{N}$

Dans cette partie, on donne les démonstrations détaillées des résultats évoqués dans le cours au sujet de l'addition et de la multiplication dans $\mathbb{N}$.

Proposition 3.12. *L'addition est associative.*

Démonstration. On montre par récurrence sur $k \in \mathbb{N}$ que

$$\forall (m, n) \in \mathbb{N}^2, m + (n + k) = (m + n) + k$$

- Initialisation : $m + (n + 0) = m + n = (m + n) + 0$.
- Hérédité : supposons que $m + (n + k) = (m + n) + k$. Alors

$$m + (n + k^+) = m + (n + k)^+ = (m + (n + k))^+ = ((m + n) + k)^+ = (m + n) + k^+$$

Ainsi la récurrence est achevée. □

Lemme 3.2. *Pour tout $n \in \mathbb{N}$, on a $n = 0 + n$ et $n^+ = 1 + n$*

Démonstration. Montrons par récurrence sur n que $n = 0 + n$.

- $0 = 0 + 0$.
- $n = 0 + n$, alors $0 + n^+ = (0 + n)^+ = n^+$

Montrons par récurrence sur n que $n^+ = 1 + n$.

- $0^+ = 1 = 1 + 0$.
- Si $n^+ = 1 + n$, alors $(n^+)^+ = (1 + n)^+ = 1 + n^+$.

Ainsi la preuve est achevée. □

Corollaire 3.4. *L'addition est commutative.*

Démonstration. Montrons par récurrence sur k que $\forall n \geqslant 0, n + k = k + n$

- $\forall n \geqslant 0, n + 0 = 0 + n = n$ d'après le lemme précédent.
- Soit $n \geqslant 0$. Alors $n + k^+ = (n + k)^+ = (k + n)^+ = k + n^+$, par hypothèse de récurrence. Or d'après le lemme précédent, $k + n^+ = k + (1 + n) = (k + 1) + n$ par associativité. Enfin : $(k + 1) + n = (k + 0^+) + n = (k + 0)^+ + n = k^+ + n$.

Ainsi la récurrence est achevée. □

Lemme 3.3. *Si $m^+ = n^+$ alors $m = n$.*

Démonstration. Supposons que $m < n$. Alors n viendrait (strictement) s'intercaler entre m et $n^+ = m^+$, absurde. Donc $m \geqslant n$. Symétriquement on a $n \geqslant m$, puis on conclut par antisymétrie. □

Proposition 3.13. *$k + n = k' + n$ alors $k = k'$. Ainsi, à partir de maintenant, il sera permis de simplifier par un même terme de part et d'autre du signe dans une addition.*

Démonstration. On montre par récurrence sur n que $k + n = k' + n$ alors $k = k'$

- Initialisation : c'est trivial.
- Hérédité : supposons que $k + n^+ = k' + n^+$ alors $k + n^+ = k' + n^+ \Rightarrow (k + n)^+ = (k' + n)^+ \Rightarrow k + n = k' + n$, et on conclut en utilisant l'hypothèse de récurrence.

Ainsi la récurrence est achevée. □

Lemme 3.4. *On a $m \geqslant n \iff \exists k \in \mathbb{N}, m = n + k$, et k est alors unique.*

Démonstration. Commençons par le sens direct ; on fixe $n \in \mathbb{N}$ et on montre par récurrence sur $m \geqslant n$ l'existence de k tel que $m = n + k$.

- Initialisation : $m = m + 0$.

- Hérédité : supposons que la propriété est vraie au rang m. Alors on écrit $m^+ = (n+k)^+ = n + k^+$.

Quant à l'unicité, c'est une simple conséquence de la propriété de simplification. Venons-en maintenant à la réciproque, que nous montrons par récurrence sur k.

- Initialisation : $m + 0 = m \geqslant m$.
- Hérédité : $m + k^+ = (m+k)^+ \geqslant m + k \geqslant m$.

Ainsi la récurrence est achevée. □

Corollaire 3.5. *On en déduit $m \leqslant n \Rightarrow m + p \leqslant n + p$.*

Démonstration. On peut introduire k tel que $n = m + k$. En utilisant l'associativité et la commutativité on en déduit : $n + p = (m + k) + p = (m + p) + k \geqslant m$. □

Lemme 3.5. *La multiplication est distributive à droite sur l'addition : $(m + n)p = mp + np$.*

Démonstration. On montre par récurrence sur $p \in \mathbb{N}$ que

$$\forall (m, n) \in \mathbb{N}^2,\ (m + n)p = mp + np$$

- Initialisation : $(m + n) \cdot 0 = 0 = 0 + 0 = m \cdot 0 + n \cdot 0$.
- Hérédité : $(m + n)(p + 1) = (m + n)p + (m + n) = (mp + np) + (m + n)$ par hypothèse de récurrence. Puis $(mp+np)+(m+n) = m(p+1)+n(p+1)$ par associativité et commutativité de de l'addition.

Ainsi la récurrence est achevée. □

Lemme 3.6. *Pour tout $n \in \mathbb{N}$ on a $0 = 0 \cdot n$ et $n = 1 \cdot n$.*

Démonstration. Déjà, on a $0 \cdot n = (0 + 0)n = 0 \cdot n + 0 \cdot n$ d'après le lemme précédent, et on conclut par simplification. Puis on montre par récurrence sur $n \in \mathbb{N}$ que $1 \cdot n = n$.

- Initialisation : $1 \cdot 0 = 0$ par définition.
- Hérédité : $1 \cdot (n + 1) = 1 \cdot n + 1$ par définition, et $1 \cdot n = n$ par hypothèse de récurrence.

Ainsi la récurrence est achevée. □

Corollaire 3.6. *La multiplication est commutative.*

On montre par récurrence sur $n \in \mathbb{N}$ que $\forall m \in \mathbb{N}, mn = nm$.

- Initialisation : $0 \cdot m = 0$ d'après la proposition précédente. Par ailleurs, on a par définition $m \cdot 0 = 0$.
- Hérédité : $m(n + 1) = mn + m = nm + m$ par hypothèse de récurrence. Puis $nm + m = nm + 1 \cdot m = (n + 1) \cdot m = (n + 1)m$ par distributivité, ce qui achève la récurrence.

Corollaire 3.7. *On en déduit la distributivité à gauche de la multiplication sur l'addition.*

Démonstration. Il suffit d'écrire $m(n + p) = (n + p)m = nm + pm$. □

Corollaire 3.8. *La multiplication est associative.*

Démonstration. On montre par récurrence sur $p \in \mathbb{N}$ que $\forall (m, n) \in \mathbb{N}^2,\ (mn)p = m(np)$.

- Initialisation : $(mn) \cdot 0 = 0$ et $m(n \cdot 0) = m \cdot 0 = 0$.
- Hérédité : $(mn)(p+1) = (mn)p + mn = m(np) + mn$ par hypothèse de récurrence. Par distributivité à gauche, on a donc $m(np) + mn = m(np + n) = m(n(p+1))$, ce qui achève la récurrence.

□

Corollaire 3.9. *On a* $\forall (m,n,p) \in \mathbb{N}^3,\ m \leqslant n \Rightarrow mp \leqslant np$.

Démonstration. On introduit k tel que $n = m + k$ puis on a : $np = mp + kp \geqslant mp$. □

Corollaire 3.10. *On a* $n_1 ... n_k = 0 \iff \exists i \in [\![1,k]\!],\ n_i = 0$

Démonstration. On montre la partie directe par récurrence sur $k \in \mathbb{N}^\star$.

- Initialisation : pour $k = 1$, il n'y a rien à montrer.
- Hérédité : supposons $n_1 ... n_{k+1} = 0$ avec $k \geqslant 1$. Si $n_{k+1} = 0$ c'est terminé. Sinon on a $n_{k+1} \geqslant 1$. D'après le corollaire précédent, on a $0 = n_1 ... n_{k+1} \geq n_1 ... n_k$. Donc $n_1 \dots n_k = 0$, et on conclut par hypothèse de récurrence.

Pour la réciproque, il suffit d'utiliser la commutativité. □

Corollaire 3.11. *On a*

$$\forall (m,n,n') \in \mathbb{N}^* \times \mathbb{N}^2,\ mn = mn' \Rightarrow n = n'$$

Ainsi, à partir de maintenant, il sera permis de simplifier par un terme ***non nul*** *à gauche et à droite du signe dans une multiplication.*

Démonstration. Parmi n et n', l'un des deux est inférieur ou égal à l'autre, car $\leqslant$ est une relation d'ordre totale. Sans perte de généralité, supposons que $n \leqslant n'$ et écrivons $n' = n + k$. Alors $mn' = mn + mk$ donc $mk = 0$ par simplification. Comme $m \neq 0$, on en déduit $k = 0$ puis $n = n'$. □

Il nous reste à montrer les propriétés élémentaires sur les puissances.

Proposition 3.14. *On a* $m^k n^k = (mn)^k m^k m^l = m^{k+l}$

Démonstration. On montre par récurrence sur k que $m^k n^k = (mn)^k$.

- Initialisation : $m^0 n^0 = 1 \cdot 1 = 1 = (mn)^0$
- Hérédité : $m^{k+1} n^{k+1} = (m^k m)(n^k n) = (mn)^k (mn) = (mn)^{k+1}$

On montre maintenant par récurrence sur l que $m^k m^l = m^{k+l}$.

- Initialisation : $m^k m^0 = m^k \cdot 1 = m^k = m^{k+0}$
- Hérédité : $m^k m^{l+1} = (m^k m^l) m = m^{k+l} m = m^{k+l+1}$

Ainsi la preuve est achevée. □

6.3 Retour sur la construction de $\mathbb{Z}$

Comment construire un entier négatif à partir des entiers positifs ? Justement en le voyant comme la différence de deux entiers positifs. Par exemple, on pourrait voir le nombre -1 comme le couple $(2,3)$, parce que $2-3=-1$. Le problème, c'est qu'il n'y a pas unicité : on peut aussi écrire $-1=4-5$ ou $-1=79-80$. Il va donc falloir *quotienter* $\mathbb{N}^2$ par une relation d'équivalence.

On voudrait donc identifier les couples (a,b) et (a',b') tels que $a-b=a'-b'$. Malheureusement, on n'a pas encore défini la soustraction. Mais suffit de "faire passer b et b' de l'autre côté" pour se ramener à des additions. On pose donc la

Définition 3.25. On définit sur l'ensemble $\mathbb{N}^2$ la relation suivante.

$$(a,b)\sim(a',b')\Leftrightarrow a+b'=b+a'$$

Lemme 3.7. *C'est une relation d'équivalence.*

Démonstration. On vérifie les trois axiomes d'une relation d'équivalence.

- Réflexivité : comme l'addition sur $\mathbb{N}$ est commutative, on a $a+b=b+a$, soit $(a,b)\sim(a,b)$.
- Symétrie : si $a+b'=b+a'$, alors $a'+b=b'+a$ (par commutativité).
- Transitivité : supposons qu'on a $a+b'=b+a'$ et $a'+b''=b'+a''$. Si on avait le droit d'utiliser des soustractions, on écrirait tout simplement

  $$a-b=a'-b'=a''-b''$$

  Mais encore une fois, nous n'avons le droit d'utiliser que l'addition. L'astuce consiste à ajouter un terme "fantôme" qui sert d'intermédiaire et que l'on simplifie à la fin. Ici, en l'occurrence il s'agit de b. Rappelons qu'on veut montrer $a+b''=b+a''$. Si on ajoute b de part et d'autre, en utilisant l'associativité et la commutativité de l'addition on a

  $$(a+b'')+b'=(a+b')+b''=(b+a')+b''=b+(a'+b'')=b+(b'+a'')=(b+a'')+b'$$

  Il ne reste plus qu'à simplifier par b'.

Ainsi la preuve est achevée. □

Définition 3.26. On peut donc quotienter $\mathbb{N}^2$ par $\mathbb{N}$ et considérer $\overline{(a,b)}$, la classe d'équivalence de (a,b). On note $\mathbb{Z}$ cet ensemble.

Puisque formellement, on peut écrire $\overline{(a,b)}=a-b$, on voudrait additionner deux entiers relatifs de la manière suivante : $(a-b)+(a'-b')=(a+a')-(b+b')=\overline{(a+a',b+b')}$. Pour l'instant, il ne s'agit que d'un sommes donc amenés à considérer le

Lemme 3.8. *Si* $(a,b)\sim(a',b')$ *et* $(c,d)\sim(c',d')$ *alors* $(a+c,b+d)\sim(a'+c',b'+d')$.

Démonstration. Si $a+b'=b+a'$ et $c+d'=d+c'$, alors $a+c+b'+d'=b+d+a'+c'$ (on a utilisé l'associativité et la commutativité de l'addition). □

Corollaire 3.12. *La définition suivante de l'addition sur* $\mathbb{Z}$ *est consistante :*

$$\overline{(a,b)}+\overline{(c,d)}=\overline{(a+c,b+d)}$$

C'est-à-dire qu'elle ne dépend pas du représentant choisi pour chaque classe d'équivalence.

Ainsi, pour additionner deux éléments de $\mathbb{Z}$, on prend deux couples d'entiers qui représentent les deux classes d'équivalence, on les additionne terme à terme, et on prend la classe du tout. On résume cela par :

$$\overline{(a,b)} + \overline{(c,d)} = \overline{(a+c, b+d)}$$

Comme le procédé ne dépend pas du choix des représentants des classes, l'addition sur $\mathbb{Z}$ est bien définie.

Théorème 3.10. *L'addition sur $\mathbb{Z}$ est associative et commutative.*

Démonstration. Pour l'associativité, on a :

$$\begin{aligned}\left(\overline{(a,b)} + \overline{(c,d)}\right) + \overline{(e,f)} &= \overline{(a+c, b+d)} + \overline{(e,f)} \\ &= \overline{((a+c)+e, (b+d)+f)} \\ &= \overline{(a+(c+e), b+(d+f))} \\ &= \overline{(a,b)} + \left(\overline{(c,d)} + \overline{(e,f)}\right)\end{aligned}$$

On voit qu'elle hérite directement de l'associativité de $+$ dans $\mathbb{N}$. De même, la commutativité vient de celle de $+$ dans $\mathbb{N}$. □

Définition 3.27. On choisit d'inclure $\mathbb{N}$ dans $\mathbb{Z}$ en identifiant $n \in \mathbb{N}$ avec $\overline{(n,0)} \in \mathbb{Z}$. Cette inclusion est bien injective. L'addition sur $\mathbb{Z}$ est alors un prolongement de celle sur $\mathbb{N}$.

Démonstration. Si $n \neq n'$, alors $\overline{(n,0)} \neq \overline{(n',0)}$ puisque $n + 0 \neq 0 + n'$. De plus, pour tout couple $(n,n') \in \mathbb{N}^2$, on a $(n,0) + (n',0) = (n+n',0)$. □

Toujours par le même calcul formel on introduit la

Définition 3.28. On définit sur $\mathbb{Z}$ l'ordre $\leqslant$ par : $\overline{(a,b)} \leqslant \overline{(c,d)} \Leftrightarrow a + d \leqslant b + c$. Cette définition est bien consistante.

Démonstration. Supposons que $a + d \leqslant b + c$, $(a,b) \sim (a',b')$ et $(c,d) \sim (c',d')$. Écrivons $b + c = a + d + k$. Alors $a' + d' + k + b + c = a + d + k + b' + c' = b + c + b' + c'$. Donc $a' + d' + k = b' + c'$ par simplification. On a alors $a' + d' \leqslant b' + c'$. Réciproquement, si $a' + d' \leqslant b' + c'$ alors $a + d \leqslant b + c$ en échangeant les rôles des variables. □

Proposition 3.15. *$\leqslant$ est une relation d'ordre totale, et elle prolonge la définition de $\leqslant$ sur $\mathbb{N}$.*

Démonstration. On vérifie les trois axiomes d'une relation d'ordre.

- Réflexivité : $a + b = b + a$, donc $(a,b) \leqslant (a,b)$.
- Antisymétrie : si $(a,b) \leqslant (c,d)$ et $(c,d) \leqslant (a,b)$, alors $a + d \leqslant b + c$ et $c + b \leqslant d + a$, donc $a + d = b + c$, soit $\overline{(a,b)} = \overline{(c,d)}$.
- Transitivité : si $(a,b) \leqslant (c,d)$ et $(c,d) \leqslant (e,f)$, alors $a + d \leqslant b + c$ et $c + f \leqslant d + e$, donc $a + d + f \leqslant b + c + f \leqslant b + d + e$. Il existe donc k tel que $b + d + e = a + d + f + k$, puis par simplification $b + e = a + f + k$, ce qui implique $a + f \leqslant b + e$, soit $(a,b) \leqslant (e,f)$.

Il reste à vérifier que $\leqslant$ est totale : si on n'a pas $a + d \leqslant b + c$ alors on a $b + c \leqslant a + d$ car $\leqslant$ est totale sur $\mathbb{N}$. En utilisant la commutativité dans $\mathbb{N}$ on en déduit $c + b \leqslant d + a$, soit $\overline{(c,d)} \leqslant \overline{(a,b)}$. Enfin, pour $(m,n) \in \mathbb{N}^2$: $m \leqslant_{\mathbb{N}} n \Leftrightarrow m + 0 \leqslant_{\mathbb{N}} 0 + n \Leftrightarrow \overline{(m,0)} \leqslant_{\mathbb{Z}} (n,0)$. □

Proposition 3.16. *La relation $\leqslant$ reste compatible avec l'addition :*

$$m \leqslant n \Rightarrow m + p \leqslant n + p$$

Démonstration. Supposons qu'on a $\overline{(a,b)} \leqslant \overline{(c,d)}$, et donnons-nous $p = \overline{(e,f)}$. Alors on a $a + d \leqslant b + c$. En sommant $e + f$ et par commutativité, on en déduit $a + e + d + f \leqslant b + f + c + e$, soit encore $\overline{(a+e,b+f)} \leqslant \overline{(c+e,d+f)}$. □

La proposition capitale vue sur $\mathbb{N}$ reste vraie (en revanche, elle deviendra fausse dans $\mathbb{Q}$ et *a fortiori* dans $\mathbb{R}$).

Proposition 3.17. *Soit $m, n \in \mathbb{Z}, m < n + 1$ alors $m \leqslant n$.*

Démonstration. Si $\overline{(a,b)} < \overline{(c,d)} + 1 = \overline{(c+1,d)}$, alors $a + d \leqslant b + c + 1, a + d \neq b + c + 1$, donc $a + d < b + c + 1$. D'après ce qui a été vu sur $\mathbb{N}$ on peut donc écrire $a + d \leqslant b + c$. □

Proposition 3.18. *L'addition sur $\mathbb{Z}$ admet un élément neutre, et tout $x \in \mathbb{Z}$ admet un inverse.*

Démonstration. L'élément neutre est donné par $\overline{(0,0)}$ et l'inverse de $\overline{(a,b)}$ est $\overline{(b,a)}$. □

On a vu que $\mathbb{N}$ pouvait être considéré comme inclus dans $\mathbb{Z}$. Ainsi, si $n \in \mathbb{N}$ l'écriture $n \in \mathbb{Z}$ un sens. D'après la proposition précédente, l'écriture $-n$ a elle aussi un sens. Réciproquement, on a la

Proposition 3.19. *Tout élément de $\mathbb{Z}$ s'écrit d'une unique manière 0, n ou $-n$ avec $n \in \mathbb{N}^*$.*

Démonstration. Pour l'existence, on se donne $(a,b) \in \mathbb{Z}$. Si $a > b$ alors on écrit $a = b + n$ avec $n \in \mathbb{N}^*$, puis on a

$$\overline{(a,b)} = \overline{(n,0)} = n$$

De même, si $a < b$, alors on écrit $b = a + n$, puis on a

$$\overline{(a,b)} = \overline{(0,n)} = -\overline{(n,0)} = -n$$

Enfin, si $a = b$ on a $\overline{(a,b)} = \overline{(0,0)} = 0$. Pour ce qui est de l'unicité, soit $x \in \mathbb{Z}$

- Si $x = 0$, alors il ne peut pas s'écrire sous la forme $\pm n$ avec $n \neq 0$ donc son écriture est unique.
- Si $x > 0$, alors il ne peut pas s'écrire sous la forme $-n$ avec $n \neq 0$ donc il ne peut s'écrire que sous la forme n avec $n \neq 0$. n est alors unique (rappelons que l'inclusion de $\mathbb{N}$ dans $\mathbb{Z}$ est injective).
- Si $x < 0$, alors il ne peut pas s'écrire sous la forme n avec $n \neq 0$ donc il ne peut s'écrire que sous la forme $-n$ avec $n \neq 0$. Puis si $-n = -n'$, alors en ajoutant $n + n'$ de part et d'autre on obtient $n = n'$.

Ainsi la preuve est achevée. □

À présent, comment définir une multiplication sur $\mathbb{Z}$? Le même petit calcul formel donne

$$(a - b)(c - d) = (ac + bd) - (ad + bc)$$

Définition 3.29. La définition suivante de la multiplication sur $\mathbb{Z}$ est consistante :

$$\overline{(a,b)} \times \overline{(c,d)} = \overline{(ac + bd, ad + bc)}$$

De plus, elle prolonge la définition de la multiplication sur $\mathbb{N}$.

Démonstration. Il suffit de vérifier la même indépendance par rapport au choix des représentants. Si $a+b'=b+a'$ et $c+d'=d+c'$, alors on veut montrer que

$$(ac+bd, ad+bc) \sim (a'c'+b'd', a'd'+b'c')$$

On utilise le même genre d'astuce que dans $\mathbb{N}$, avec l'introduction d'un "terme fantôme". Le but est de montrer que

$$ac+bd+a'd'+b'c' = ad+bc+a'c'+b'd'$$

Avec un peu plus de mal, on trouve que qu'un terme fantôme possible est $ad'+bc'+bd'+ac'$:

$$\begin{aligned}(ac+bd+a'd'+b'c')+(ad'+bc'+bd'+ac') &= a(c+d')+b(c'+d)+(a'+b)d'+(a+b')c' \\ &= a(c'+d)+b(c+d')+(a+b')d'+(a'+b)c' \\ &= (ad+bc+a'c'+b'd')+(ad'+bc'+bd'+ac')\end{aligned}$$

Après simplification, on obtient le résultat. De plus, on a

$$\overline{(n,0)}\cdot\overline{(n',0)} = \overline{(nn'+0\cdot 0, n\cdot 0+0\cdot n')} = \overline{(nn',0)}$$

Donc la multiplication de $\mathbb{Z}$ prolonge bien celle de $\mathbb{N}$. □

Théorème 3.11. *La multiplication sur $\mathbb{Z}$ est associative, commutative, et distributive sur l'addition.*

Démonstration. Vérifions l'associativité.

$$\begin{aligned}\Big(\overline{(a,b)}\cdot\overline{(c,d)}\Big)\cdot\overline{(e,f)} &= \overline{(ac+bd, ad+bc)}\cdot\overline{(e,f)} \\ &= \overline{((ac+bd)e+(ad+bc)f, (ac+bd)f+(ad+bc)e)} \\ &= \overline{(a(ce+df)+b(cf+de), a(cf+de)+b(ce+df))} \\ &= (a,b)\cdot(ce+df, cf+de) \\ &= \overline{(a,b)}\cdot\Big(\overline{(c,d)}\cdot\overline{(e,f)}\Big)\end{aligned}$$

Vérifions maintenant la commutativité.

$$\begin{aligned}\overline{(a,b)}\cdot\overline{(c,d)} &= \overline{(ac+bd, ad+bc)} \\ &= \overline{(ca+db, cb+da)} \\ &= \overline{(c,d)}\cdot\overline{(a,b)}\end{aligned}$$

Quant à la distributivité, par commutativité il suffit de la vérifier d'un côté seulement.

$$\begin{aligned}\Big(\overline{(a,b)}+\overline{(c,d)}\Big)\cdot\overline{(e,f)} &= \overline{(a+c, b+d)}\cdot\overline{(e,f)} \\ &= \overline{((a+c)e+(b+d)f, (a+c)f+(b+d)e)} \\ &= \overline{(ae+bf+ce+df, af+be+cf+de)} \\ &= \overline{(ae+bf, af+be)}+\overline{(ce+df, cf+de)} \\ &= \overline{(a,b)}\cdot\overline{(e,f)}+\overline{(c,d)}\cdot\overline{(e,f)}\end{aligned}$$

Ainsi la preuve est achevée. □

Remarque 3.18. On vérifie immédiatement que 1 continue à être élément neutre.

Corollaire 3.13. *On a*

$$\begin{cases} \forall n \in \mathbb{Z},\ n \times 0 = 0 \times n = 0 \\ \forall (m,n) \in \mathbb{Z}^2,\ m(-n) = (-m)n = -(mn) \end{cases}$$

On dit que 0 est absorbant.

Démonstration. Pour la première ligne, il suffit d'écrire $n = n \times 1 = n(1+0) = n + n \times 0$, puis de simplifier par n. Le raisonnement est le même avec $0 \times n$.

Pour la seconde on écrit $m(-n) + mn = m(-n+n) = m \times 0 = 0$. Le raisonnement est le même avec $(-m)n$. □

Proposition 3.20. *On sait que dans $\mathbb{Z}$, la relation $\leqslant$ reste compatible avec l'addition. Pour la multiplication, il faut être plus prudent.*

$$\begin{cases} m \leqslant n \\ p \geqslant 0 \end{cases} \Rightarrow mp \leqslant np$$

Démonstration. En utilisant la distributivité et la compatibilité avec l'addition, tout revient à montrer que $n \geqslant 0$ et $p \geqslant 0$, alors $np \geqslant 0$ Pour cela on peut invoquer le fait que la multiplication sur $\mathbb{Z}$ prolonge celle sur $\mathbb{N}$. □

Remarque 3.19. On en déduit qu'un carré (d'entier) est toujours positif ou nul.

6.4 Formules sommatoires

On se donne un ensemble E muni d'une loi associative $*$ et d'un élément neutre e.

Proposition 3.21. *Quels que soient $0 \leqslant k \leqslant m \leqslant n$,*

$$x_1 * ... * x_n = x_1 * ... * x_k * (x_{k+1} * ... * x_m) * x_{m+1} * ... * x_n$$

(avec la convention habituelle si $k = 0$ ou $m = n$).

Démonstration. Dans un premier temps, fixons k et supposons que $n > m$. Il s'agit alors simplement de montrer $x_1 * ... * x_m = x_1 * ... * x_k * (x_{k+1} * ... * x_m)$. Si $k = 0$ il n'y a rien à prouver. Sinon, effectuons une récurrence sur m.

- Initialisation : si m=k, il n'y a rien à prouver.
- Hérédité : si la propriété est vraie au rang m, écrivons

$$\begin{aligned} x_1 * ... * x_{m+1} &= (x_1 * ... * x_m) * x_{m+1} \\ &= (x_1 * ..., * x_k * (x_{k+1} * ... * x_m)) * x_{m+1} \\ &= ((x_1 * ... * x_k) * (x_{k+1} * ... * x_m)) * x_{m+1} \\ &= (x_1 * ... * x_k) * ((x_{k+1} * ... * x_m) * x_{m+1}) \\ &= x_1 * ... * x_k * (x_{k+1} * ... * x_{m+1}) \end{aligned}$$

Montrons à présent le cas général par récurrence sur n.

- Initialisation : si $n = m$, c'est le cas précédent.
- Hérédité : si la propriété est vraie au rang n, on a

$$\begin{aligned} x_1 * ... * x_{n+1} &= (x_1 * ... * x_n) * x_{n+1} \\ &= (x_1 * ... * x_k * (x_{k+1} * ... * x_m) * x_{m+1} * ... * x_n) * x_{n+1} \\ &= x_1 * ... * x_k * (x_{k+1} * ... * x_m) * x_{m+1} * ... * x_{n+1} \end{aligned}$$

On vérifie que le calcul reste valable pour $k = 0$. Ainsi la preuve est achevée. □

À partir de maintenant, nous travaillons sur une loi associative **et** commutative. C'est pourquoi nous la noterons plutôt $+$, afin de mieux évoquer la commutativité, mais on pourrait aussi la noter $\times$ dans le cas de la multiplication. En particulier, l'élément neutre sera noté 0, et la composée $x_1 * ... * x_n$ sera notée $\sum_{k=1}^{n} x_k$.

Définition 3.30 (Permutations)**.** Soit $n \in \mathbb{N}$. Une **permutation** de $[\![1, n]\!]$ est une bijection de l'ensemble $[\![1, n]\!]$ sur lui-même. L'ensemble des permutations de $[\![1, n]\!]$ est noté $\mathcal{S}_n$ voire plus anciennement (et quand on arrive à le calligraphier) $\mathfrak{S}_n$.

Théorème 3.12. *Avec la loi $+$, on a le droit de composer les termes dans l'ordre que l'on veut. Autrement dit,*

$$\forall \sigma \in \mathcal{S}_n, x_1 + ... + x_n = x_{\sigma(1)} + ... + x_{\sigma(n)}$$

Démonstration. Prouvons le résultat par récurrence sur n.

- Initialisation : pour $n = 0$, le seul élément de $\mathcal{S}_0$ est "l'application vide".
- Hérédité : supposons la propriété vraie au rang n avec $n \geqslant 1$ et montrons-la au rang $n + 1$. Supposons dans un premier temps que $\sigma(n + 1) = n + 1$ Alors σ induit une permutation de $[\![1, n]\!]$ et il suffit d'invoquer l'hypothèse de récurrence. Dans un second temps, supposons que $\sigma(n+1) \neq n+1$. Grâce à l'hypothèse de récurrence, on peut supposer sans perte de généralité que $\sigma(n) = n + 1$. Alors, par un simple argument de commutativité, on permute $x_{\sigma(n)}$ et $x_{\sigma(n+1)}$ et on se ramène au cas précédent.

Ainsi la récurrence est achevée. □

Définition 3.31. Soit I un ensemble d'indices fini, et $(x_i)_{i \in I}$ une famille à valeurs dans $(E, +)$. Si on pose $n := |I|$, il existe une bijection σ de $[\![1, n]\!]$ sur I. On pose alors

$$\sum_{i \in I} x_i = \sum_{k=1}^{n} x_{\sigma(k)}$$

D'après le théorème précédent, la définition est bien indépendante du choix de σ. En effet, si on se donne une autre permutation σ', en notant $\tau := \sigma^{-1} \circ \sigma'$ on obtient

$$\sum_{k=1}^{n} x_{\sigma(k)} = \sum_{k=1}^{n} x_{\sigma(\tau(k))} = \sum_{k=1}^{n} x_{\sigma'(k)}$$

Remarque 3.20. On observera que la formule $\sum_{i\in I}(x_i+y_i)=\sum_{i\in I}x_i+\sum_{i\in I}y_i$ est une conséquence directe de l'associativité et de la commutativité de $+$.

Lemme 3.9. *Si I_1 et I_2 sont des parties disjointes de I, alors* $\sum_{i\in I_1\cup I_2}x_i=\sum_{i\in I_1}x_i+\sum_{i\in I_2}x_i$.

Démonstration. Raisonnons par récurrence sur $\mathrm{Card}(I_1\cup I_2)$.

- Initialisation : si $\mathrm{Card}(I_1\cup I_2)=0$, alors $I_1=I_2=\emptyset$ donc il n'y a rien à prouver.
- Hérédité : donnons-nous une bijection σ de $[\![1,\dots,n+1]\!]$ sur $I_1\cup I_2$ et supposons sans perte de généralité que $\sigma(n+1)\in I_2$.
  - Si $I_1=\emptyset$, il n'y a toujours rien à prouver.
  - Sinon, introduisons $I_2'=I_2\setminus\{\sigma(n+1)\}$. Comme $I_1\neq\emptyset$, I_2' est strictement inclus dans $I_1\cup I_2$ et par hypothèse de récurrence on a

$$\sum_{i\in I_2}x_i=\left(\sum_{i\in I_2'}x_i\right)+x_{\sigma(n+1)}$$

De même, $I_1\cup I_2'$ est strictement inclus dans $I_1\cup I_2$ donc :

$$\sum_{i\in I_1\cup I_2'}x_i=\sum_{i\in I_1}x_i+\sum_{i\in I_2'}x_i$$

Enfin, remarquons que σ induit une bijection de $[\![1,\dots,n]\!]$ sur $I_1\cup I_2'$. Par associativité, on en déduit

$$\begin{aligned}\sum_{i\in I_1}x_i+\sum_{i\in I_2}x_i&=\sum_{i\in I_1}x_i+\left(\sum_{i\in I_2'}x_i+x_{\sigma(n+1)}\right)\\&=\left(\sum_{i\in I_1}x_i+\sum_{i\in I_2'}x_i\right)+x_{\sigma(n+1)}\\&=\left(\sum_{i\in I_1\cup I_2'}x_i\right)+x_{\sigma(n+1)}\\&=\left(\sum_{k=1}^{n}x_{\sigma(k)}\right)+x_{\sigma(n+1)}\\&=\sum_{k=1}^{n+1}x_{\sigma(k)}\end{aligned}$$

ce qui achève la récurrence.

□

Théorème 3.13. *Si $I_1,\dots,I_n$ sont des parties deux à deux disjointes de I, alors*

$$\sum_{i\in I_1\cup\ldots\cup I_n}x_i=\sum_{i\in I_1}x_i+\ldots+\sum_{i\in I_n}x_i$$

Démonstration. C'est une récurrence immédiate sur n. □

Chapitre 4

Groupes, anneaux, corps

1 Groupes

Pour commencer, rappelons que la composition des applications est associative. En revanche, elle n'a aucune raison d'être commutative.

1.1 Premières définitions

De manière schématique, un groupe est un ensemble dans lequel on peut toujours composer deux éléments de manière associative, qui contient un élément neutre, et tel que tout élément possède un inverse.

Définition 4.1 (Groupe)**.** Soit un ensemble G muni d'une loi de composition interne $*$. La structure algébrique $(G, *)$ est dite un **groupe** lorsqu'elle vérifie les trois axiomes suivants.

- $*$ est associative.
- $(G, *)$ est muni d'un **élément neutre**.
- Tous les éléments de G sont **inversibles** pour $*$.

Par un léger abus de terminologie, on dira souvent que G lui-même est un groupe (pour la loi $*$). Généralement l'élément neutre de G est noté e, et l'inverse de $x \in G$ sera noté x^{-1}. Assez rapidement, s'il n'y a pas d'ambiguïté, on omettra le symbole $*$ et on notera simplement xy pour $x * y$. Si de plus, $*$ est commutative, on dit que G est **commutatif** (on rencontrera parfois le mot **abélien**). Sa loi pourra alors être notée $+$, et son neutre pourra être noté 0. Dans ce cas, on rappelle que l'inverse de x est usuellement noté $-x$, et que l'écriture $x - y$ désigne en fait $x + (-y)$.

Exemple 4.1. $(\mathbb{Z}, +)$ est un groupe commutatif.

Exemple 4.2. Jusqu'au baccalauréat, intentionnellement ou non, on rencontre plusieurs autres exemples de groupes.

- Tout d'abord, l'ensemble des nombres complexes $\mathbb{C}$ muni de l'addition : la somme de deux complexes est un complexe, et l'opposé d'un complexe est un complexe.
- On aurait aussi pu considérer $\mathbb{C}^\star$ muni de la multiplication : le produit de deux complexes non nuls est un complexe non nul, et l'inverse d'un complexe non nul est un complexe non nul.

- Ensuite, l'ensemble des nombres complexes de module 1 muni de la multiplication (autrement dit, le cercle unité) : si on multiplie z et z' de module 1, alors zz' est de module 1, et l'inverse d'un nombre complexe de module 1 est de module 1.
- Ensuite, le premier groupe fini rencontré : les racines n-èmes de l'unité. Le produit de deux racines de l'unité est une racine n-ème de l'unité, et l'inverse d'une racine n-ème de l'unité est une racine n-ème de l'unité.
- Après, la grande famille de la géométrie : les translations, les homothéties-translations, les rotations-translations, les similitudes... Tous ces ensembles forment des groupes pour la composition des applications. Attention, la composition est notre premier exemple de loi qui ne soit pas commutative. Par exemple, deux similitudes qui n'ont pas le même centre ne commutent pas en général.
- On peut citer le cas de vecteurs de $\mathbb{R}^2$ ou de $\mathbb{R}^3$, voire de $\mathbb{R}^n$ en général, qui sont des groupes pour l'addition.

On voit à quel point les groupes sont omniprésents en mathématiques, dans des branches très variées. L'intérêt d'identifier cette structure est de mettre au point une fois pour toutes des théorèmes généraux, valables dans n'importe quel groupe, et ainsi on n'aura pas besoin de redémontrer à chaque fois le même résultat "déguisé" sous une forme différente.

Proposition 4.1. *Dans un groupe, on peut "simplifier" une égalité à droite ou à gauche :*

$$\forall (x,y,z) \in G^3, \ \begin{cases} xy = xz \implies y = z \\ yx = zx \implies y = z \end{cases}$$

*On dit que x est **régulier**.*

Remarque 4.1. Rappelons que dans un groupe, l'inverse de xy est $y^{-1}x^{-1}$. Plus généralement, l'inverse de $x_1 \dots x_n$ est $x_n^{-1} \dots x_1^{-1}$.

Exemple 4.3. L'ensemble $\mathcal{S}_X$ des permutations d'un ensemble quelconque X forme un groupe pour la composition. En particulier, $\mathcal{S}_{[\![1,n]\!]}$ est un groupe non commutatif dès que $n \geqslant 3$. On l'appelle **groupe symétrique d'ordre** n et on le note plutôt $\mathcal{S}_n$.

Remarque 4.2. Soit G un groupe, pas nécessairement commutatif. Si $xy = e$, alors en composant par x^{-1}, on obtient $y = x^{-1}$, d'où $yx = e$.

Proposition 4.2 (Groupe-produit)**.** *Soit $(G_1, *_1)$ et $(G_2, *_2)$ deux groupes. Alors $G_1 \times G_2$ muni de la loi*

$$(x_1, x_2) * (y_1, y_2) := (x_1 *_1 y_1, \ x_2 *_2 y_2)$$

*est un groupe, dit **groupe-produit** de G_1 et G_2. On omettra souvent les signes $*$, $*_1$ et $*_2$.*

Remarque 4.3. Si G_1 et G_2 sont commutatifs, alors $G_1 \times G_2$ est commutatif.

Remarque 4.4. Bien sûr, on peut étendre les résultat précédents à un produit $G_1 \times \dots \times G_n$ pour $n \geqslant 2$.

Remarque 4.5. Comme $(\mathbb{R}, +)$ est un groupe, on en déduit immédiatement les structures de groupe de $(\mathbb{R}^2, +)$, $(\mathbb{R}^3, +)$ et $(\mathbb{R}^n, +)$.

1.2 Sous-groupes

Définition 4.2 (Sous-groupe)**.** Soit G un groupe. Un **sous-groupe** H de G est une partie $H \subset G$ telle que :

- $e \in H$
- $\forall (x,y) \in H^2,\ xy \in H$ (stabilité par la l.c.i)
- $\forall x \in H,\ x^{-1} \in H$ (stabilité par passage à l'inverse)

Proposition 4.3. *Soit G un groupe et H un sous-groupe. Muni de la loi induite par celle de G, H est un groupe.*

Remarque 4.6. Si G est commutatif, alors tous ses sous-groupes sont commutatifs.

Exemple 4.4. L'ensemble des nombre complexes de module 1 forme un sous-groupe multiplicatif de $\mathbb{C}^\star$. C'est donc un groupe commutatif.

Proposition 4.4 (Recette du sous-groupe)**.** *Soit G un groupe et $H \subset G$. Pour que H soit un sous-groupe de G, il faut et il suffit que*

- $H \neq \emptyset$ *(par exemple $e \in H$)*
- $\forall (x,y) \in H^2,\ xy^{-1} \in H$

Remarque 4.7. En pratique, on n'utilise que la condition suffisante. En effet, bien souvent, on pourra utiliser cette recette lorsqu'un énoncé nous demandera de démontrer que tel ou tel ensemble est un groupe : en l'interprétant comme un sous-ensemble d'un groupe connu, on montrera directement qu'il s'agit d'un sous-groupe.

Exemple 4.5. $\{e\}$ est le plus petit sous-groupe de G.

Exemple 4.6. Soit $(G,+)$ un groupe commutatif, et A et B deux sous-groupes. Alors $A+B$ est un sous-groupe, comme nous le vérifions maintenant.

- $0 = 0+0 \in A+B$
- Soit $a+b \in A+B$ et $a'+b' \in A+B$. Alors :

$$\begin{aligned}(a+b)-(a'+b') &= (a+b)+(-b'-a') \\ &= (a-a')+(b-b') && \text{par commutativité et associativité} \\ &\in A+B\end{aligned}$$

Exemple 4.7. Soit G un groupe non nécessairement commutatif. On appelle **centre** de G l'ensemble des éléments qui commutent avec tous les éléments de G :

$$Z(G) = \{x \in G \mid \forall y \in G,\ xy = yx\}$$

(le "Z" vient de l'allemand "*Zentrum*", les Allemands étant omniprésents dans l'histoire de l'algèbre). $Z(G)$ est un sous-groupe de G, comme nous le vérifions maintenant.

- Comme $\forall x \in G,\ ex = xe = x$, on a $e \in Z(G)$.

- Soit $(x, y) \in \mathbb{Z}(G)^2$. Soit $z \in G$. On a :

$$\begin{aligned}
(xy^{-1})z &= x(y^{-1}z) && \text{par associativité}\\
&= x(z^{-1}y)^{-1}\\
&= x(yz^{-1})^{-1} && \text{par } y \in Z(G)\\
&= x(zy^{-1})\\
&= (xz)y^{-1} && \text{par associativité}\\
&= (zx)y^{-1} && \text{car } x \in Z(G)\\
&= z(xy^{-1}) && \text{par associativité}
\end{aligned}$$

  si bien que $xy^{-1} \in Z(G)$

Théorème 4.1. *Soit $\mathcal{H}$ un ensemble de sous-groupes de G. Alors $\bigcap_{H\in\mathcal{H}} H$ est un sous-groupe de G.*

Ce théorème justifie la définition suivante.

Définition 4.3 (Sous-groupe engendré)**.** Soit X une partie de G. L'intersection de tous les sous-groupes de G qui contiennent X s'appelle le **sous-groupe engendré par** X. Puisqu'il est inclus dans tout sous-groupe contenant X, c'est le plus petit sous-groupe contenant X. On le note $< X >$, voire (X). Si x est un élément de G, on notera plus simplement $< x >$ plutôt que $< \{x\} >$.

Exemple 4.8. Dans $(\mathbb{R}, +)$, on a $< 1 >= \mathbb{Z}$.

Définition 4.4 (Itération de la l.c.i.)**.** Soit G un groupe dont la loi est notée multiplicativement, et $x \in G$. On rappelle qu'on pose

$$\begin{cases} x^0 = e \\ \forall n \in \mathbb{N},\ x^{n+1} = x^n x \end{cases}$$

Puis on rappelle qu'on étend la définition à $n \in \mathbb{Z}$ en posant

$$\forall n \in \mathbb{N}^\star,\ x^{-n} = (x^n)^{-1}$$

On rappelle qu'on a aussi qu'on a les relations suivantes :

$$\forall (m, n) \in \mathbb{Z}^2,\ \begin{cases} x^m x^n = x^{m+n} \\ (x^m)^n = x^{mn} \end{cases}$$

Et si x et y commutent, alors :

$$\begin{cases} \forall (m, n) \in \mathbb{Z}^2,\ x^m y^n = y^n x^m \\ \forall n \in \mathbb{Z},\ (xy)^n = x^n y^n \end{cases}$$

Exemple 4.9. On a $\forall n \in \mathbb{Z},\ e^n = e$.

Remarque 4.8. Dans le cas où le groupe est commutatif, il n'est pas rare que sa loi soit notée additivement. Dans ce cas, l'usage veut que x^n soit plutôt noté nx (pour correspondre à l'idée qu'on se fait de $x + \ldots + x$). La définition se réécrit ainsi :

$$\begin{cases} 0x = 0 \\ \forall n \in \mathbb{N}, (n + 1)x = nx + x \\ \forall n \in \mathbb{N}^\star,\ (-n)x = -(nx) \end{cases}$$

Par ailleurs, on obtient les propriétés suivantes.

$$\begin{cases} \forall (m,n) \in \mathbb{Z}^2,\ mx + nx = (m+n)x & \\ \forall (m,n) \in \mathbb{Z}^2,\ n(mx) = (mn)x & (= (nm)x = m(nx)) \\ \forall (m,n) \in \mathbb{Z}^2,\ mx + ny = ny + mx & (\text{par commutativité}) \\ \forall n \in \mathbb{Z},\ n(x+y) = nx + ny & (\text{par commutativité}) \end{cases}$$

Proposition 4.5. *Soit G un groupe et $x \in G$. Alors le sous-groupe engendré par x vaut*

$$< x >= \{x^n \mid n \in \mathbb{Z}\}$$

Remarque 4.9. Si la loi du groupe est notée additivement, on a $< x >= \{nx \mid n \in \mathbb{Z}\}$

Définition 4.5 (Groupe monogène, groupe cyclique)**.** Un groupe est **monogène** lorsqu'il est engendré par un seul élément. Il est **cyclique** lorsqu'il est monogène et fini.

Exemple 4.10. $(\mathbb{Z}, +)$ est monogène puisqu'il est engendré par 1. $(\mathbb{U}_n, \times)$ est cyclique, engendré par $\omega = \exp\left(i\dfrac{2\pi}{n}\right)$.

Remarque 4.10. Tout groupe monogène est nécessairement commutatif, puisque tous ses éléments s'écrivent en itérant un unique élément.

Pour finir, citons le célèbre

Théorème 4.2 (Théorème de Lagrange, HP)**.** *Soit G un groupe fini, et H un sous-groupe de G. Alors,* $\mathrm{Card}(H)$ *divise* $\mathrm{Card}(G)$*.*

Démonstration. On trouvera une preuve dans les compléments sur les groupes. □

1.3 Morphismes

Définition 4.6 (Morphisme de groupes)**.** Soit G et G' deux groupes. Un **morphisme** de G dans G' est une application $\varphi : G \to G'$ telle que

$$\forall (x,y) \in G^2,\ \varphi(xy) = \varphi(x)\varphi(y)$$

Autrement dit, φ "transporte la loi de G dans celle de G'". On note $Hom(G, G')$ l'ensemble des morphismes de G dans G'.

Remarque 4.11. La notation $Hom(G, G')$ vient de ce qu'anciennement, un morphisme s'appelait un homomorphisme.

Définition 4.7 (Isomorphisme, endomorphisme, automorphisme)**.** Un **isomorphisme** est un morphisme bijectif. Un **endomorphisme** est un morphisme de G dans lui-même. On note $Hom(G)$ l'ensemble des endomorphismes de G. Enfin, un **automorphisme** est un endomorphisme qui est aussi un isomorphisme.

Proposition 4.6 (Propriétés des morphismes)**.** *Pour tout morphisme $\varphi : G \to G'$, on a :*

$$\begin{cases} \varphi(e_G) = e_{G'} \\ \forall x \in G,\ \varphi(x^{-1}) = \varphi(x)^{-1} \end{cases}$$

Remarque 4.12 (Morphisme de groupes et itération)**.** Plus généralement, on montre que

$$\forall x \in G,\ \forall k \in \mathbb{Z},\ \varphi\left(x^k\right) = \varphi(x)^k$$

Définition 4.8 (Noyau)**.** Soit φ un morphisme de G dans G'. Le **noyau de** φ est défini par

$$\ker(\varphi) := \varphi^{-1}\left(\{e_{G'}\}\right) = \{x \in G \mid \varphi(x) = e_{G'}\}$$

la notation ker vient de l'allemand "*der Kern*" qui signifie "noyau".

Définition 4.9 (Image)**.** Soit φ un morphisme de G dans G'. L'**image de** φ est défini par

$$\mathrm{Im}(\varphi) := \varphi(G) = \{y \in G' \mid \exists x \in G,\ y = \varphi(x)\}$$

Théorème 4.3. *L'image d'un sous-groupe H de G par un morphisme de groupes est un sous-groupe de G'. L'image réciproque d'un sous-groupe H' de G' par un morphisme de groupes est un sous-groupe de G.*

Corollaire 4.1. $\ker(\varphi)$ *est un sous-groupe de* G. $\mathrm{Im}(\varphi)$ *est est un sous-groupe de* G'.

Théorème 4.4. *Soit φ un morphisme de G dans G'. Alors, φ est injectif si, et seulement si,* $\ker(\varphi) = \{e_G\}$.

Exemple 4.11. Un exemple typique d'utilisation est celui où G et G' sont finis de même cardinal. Si on montre que $\ker(\varphi) = \{e_G\}$, alors φ sera injectif, puis bijectif. On pourra ainsi démontrer très efficacement que φ est surjectif, ce qui autrement est toujours un peu délicat.

Proposition 4.7. *La composée de deux morphismes est encore un morphisme. La bijection réciproque d'un isomorphisme est un isomorphisme.*

Exemple 4.12. $(Aut(G), \circ)$ (avec $Aut(G)$ l'ensemble des automorphismes de G) est un groupe.

Corollaire 4.2. *Il existe un isomorphisme de G dans G' si, et seulement si, il existe un isomorphisme de G' dans G. De manière condensée, on dira alors que G et G' sont* ***isomorphes****. On pourra écrire :*

$$G \cong G'$$

2 Anneaux

2.1 Premières définitions

Définition 4.10 (Anneau)**.** Soit A un ensemble muni de deux lois de composition internes $+$ et $\times$. On dit que $(A, +, \times)$ est un anneau lorsque :

- $(A, +)$ est un groupe commutatif.
- $\times$ est associative et possède un élément neutre. En revanche, les éléments ne sont *a priori* pas inversibles pour $\times$, et $\times$ n'est pas nécessairement commutative.
- $\times$ est **distributive** à gauche et à droite sur $+$:

$$\forall (x, y, z) \in A^3,\ \begin{cases} x \times (y + z) = x \times y + x \times z \\ (x + y) \times z = x \times z + y \times z \end{cases}$$

Si $\times$ est commutative, on dit que l'anneau est **commutatif**.

Remarque 4.13. Si on veut montrer que A est un anneau commutatif, on a intérêt à montrer la commutativité avant la distributivité. Ainsi, il suffira de vérifier la distributivité d'un côté seulement.

Remarque 4.14 (Notations)**.** Traditionnellement, l'élément neutre pour $+$ est noté 0, et l'élément neutre pour $\times$ est noté 1. Rapidement, s'il n'y a pas d'ambiguïté, on ne notera même plus la loi $\times$: $x \times y$ sera noté xy. S'il existe, l'inverse de x pour la deuxième loi sera noté x^{-1}. Enfin, par "priorité d'opérateur", l'expression $xy + x'y'$ est définie comme étant égale à $(xy) + (x'y')$.

Exemple 4.13. $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ et $\mathbb{C}$ sont des anneaux commutatifs.

Exemple 4.14. Soit f et g deux fonctions d'un intervalle $I \subset \mathbb{R}$ dans $\mathbb{R}$. On définit de manière évidente :

$$\begin{cases} f+g : x \mapsto f(x) + g(x) \\ fg : x \mapsto f(x)g(x) \end{cases}$$

Avec cette notation, $(\mathbb{R}^I, +, \times)$ est un anneau (commutatif). On vérifie que chacune des propriétés s'hérite des mêmes propriétés dans l'anneau $\mathbb{R}$.

Remarque 4.15. Plus généralement, si X est un ensemble quelconque et $(A, +, \times)$ un anneau, alors $(A^X, +, \times)$ est un anneau au même titre que A.

Proposition 4.8. *On a :*

$$\begin{cases} \forall x \in A,\ 0 \times x = x \times 0 = 0 \\ \forall (x,y) \in A^2,\ x(-y) = (-x)y = -xy \end{cases}$$

On dit que 0 *est* ***absorbant****.*

Définition 4.11 (Anneau trivial, anneau non trivial)**.** Supposons que A est réduit à un singleton : $A = \{x\}$. Alors, on a nécessairement :

$$\begin{cases} x + x = x \\ x \times x = x \end{cases}$$

Réciproquement, tout singleton $\{x\}$ muni des deux lois exposées ci-dessus est bien un anneau. On l'appelle **anneau trivial**. Notons qu'on a nécessairement $0_A = 1_A$ et $A = \{0_A\}$. C'est pourquoi on parle aussi d'**anneau nul**. Si A n'est pas l'anneau trivial, on dira que A est un **anneau non trivial**.

Remarque 4.16. Dès que A est non trivial, on a $0 \neq 1$. En effet, on choisit alors un élément $x \neq 0$, on a $x \times 1 = x \neq x \times 0 = 0$.

Y a-t-il d'autres éléments dont le produit vaut 0 ? Cela dépend et fait l'objet de la définition suivante.

Définition 4.12 (Anneau intègre)**.** Un anneau est dit **intègre** lorsqu'il est non trivial, commutatif, et qu'il ne contient pas de "diviseurs de 0" :

$$\forall (x,y) \in A^2,\ xy = 0 \implies x = 0 \text{ ou } y = 0$$

Exemple 4.15. $\mathbb{Z}$ est un anneau intègre. En effet, supposons que $(\pm m)(\pm n) = 0$ avec $(m, n) \in \mathbb{N}^2$. Alors on a $mn = 0$ d'où $m = 0$ ou $n = 0$.

Proposition 4.9. *Dans un anneau intègre, tout élément non nul est régulier pour la deuxième loi.*

Démonstration. Par commutativité, il suffit de démontrer la régularité d'un seul côté. Supposons donc que $xy = xz$ avec x non nul. Alors $x(y - z) = 0$, et comme $x \neq 0$, on en déduit $y - z = 0$ puis $y = z$. □

Proposition 4.10 (Anneau-produit)**.** *De même qu'on a un groupe-produit, on peut définir un anneau-produit pour $n \geqslant 2$ anneaux.*

Définition 4.13 (Unités)**.** Les **unités** de A sont les éléments de A inversibles pour la deuxième loi. Leur ensemble est noté $\mathcal{U}(A)$, voire $A^\times$.

Proposition 4.11 (Structure de groupe des unités)**.** *La loi $\times$ de A induit une structure de groupe sur $\mathcal{U}(A)$. De manière légèrement abusive, on pourra dire que $(\mathcal{U}(A), \times)$ est un groupe.*

2.2 Calculs algébriques dans un anneau

Définition 4.14. On a rappelle que dans un groupe dont la loi est notée additivement, le n-ème itéré de x est noté nx. Dans $(A, +)$, on pose donc :

$$\begin{cases} 0_{\mathbb{Z}}.x = 0_A \\ \forall n \in \mathbb{N}, (n + 1_{\mathbb{Z}})x = nx + x \\ \forall n \in \mathbb{N}^\star,\ (-n)x = -(nx) \end{cases}$$

Au passage, on notera la priorité d'opérateur : $mx + ny$ désigne en fait $(mx) + (ny)$.

Exemple 4.16. Si $A = \mathbb{Z}$, on montre facilement que cette définition coïncide avec la multiplication (c'est-à-dire la deuxième l.c.i.). Si $n \in \mathbb{N}$, c'est une simple récurrence, et si $n < 0$, on passe à l'opposé. Plus généralement, ce raisonnement est valable dans tout sur-anneau de $\mathbb{Z}$.

Cela dit, dans le cas général, lorsqu'on écrit nx, il ne s'agit pas de la multiplication dans A, mais plutôt d'une loi **externe** : on compose un élément de $\mathbb{Z}$ et de A, qui n'ont pas la même nature, et obtient un élément de A. Cependant, on se rend vite compte qu'en termes d'associativité et de distributivité, tout fonctionne comme si on travaillait avec la multiplication dans A.

Proposition 4.12. *On a les relations suivantes :*

$$\begin{cases} \forall (m, n) \in \mathbb{Z}^2,\ \forall x \in A,\ (mn)x = m(nx) & (\textit{pseudo-associativité}) \\ \forall n \in \mathbb{Z},\ \forall (x, y) \in A^2,\ n(x + y) = nx + ny & (\textit{pseudo-distributivité à gauche}) \\ \forall (m, n) \in \mathbb{Z}^2,\ \forall x \in A, (m + n)x = mx + nx & (\textit{pseudo-distributivité à droite}) \\ \forall x \in A,\ 1_{\mathbb{Z}} \cdot x = x & (\textit{opérateur neutre}) \\ \forall n \in \mathbb{Z},\ \forall (x, y) \in A^2,\ (nx) \times y = x \times (ny) = n(x \times y) & (\textit{compatibilité avec le produit}) \end{cases}$$

Remarque 4.17. La dernière propriété permet d'écrire nxy sans ambiguïté.

Exemple 4.17. On en déduit que le produit est bilinéaire au sens suivant :

$$\begin{cases} \forall (m, n) \in \mathbb{Z}^2,\ \forall (x, y, z) \in A^3,\ x(my + nz) = m(xy) + n(xz) \\ \forall (m, n) \in \mathbb{Z}^2,\ \forall (x, y, z) \in A^3,\ (mx + ny)z = m(xz) + n(yz) \end{cases}$$

Proposition 4.13 (Pseudo-absorbance et pseudo-règle des signes)**.** *Soit $n \in \mathbb{Z}$ et $x \in A$. On a :*

$$\begin{cases} 0_{\mathbb{Z}}.x = n.0_A = 0_A \\ (-n).x = n.(-x) = -(n.x) \\ (-n).(-x) = n.x \end{cases}$$

Définition 4.15 (Itération de la loi $\times$)**.** On rappelle qu'on définit :

$$\begin{cases} x^0 = 1_A \\ \forall n \in \mathbb{N},\ x^{n+1} = x^n x \end{cases}$$

Mais attention, en général, on ne pourra pas définir x^n pour $n < 0$, sauf si x est inversible pour $\times$, auquel cas on retrouve la définition de x^n dans le groupe $\mathcal{U}(A)$.

Définition 4.16 (Élément nilpotent, indice de nilpotence)**.** Un élément x de l'anneau A est **nilpotent** s'il existe $n \in \mathbb{N}$ tel que $x^n = 0$. Dans ce cas, le plus petit n vérifiant cette égalité s'appelle l'**indice de nilpotence** de x.

Exemple 4.18. On rencontrera les exemples les plus classiques dans le chapitre "Matrices". En attendant, en anticipant un tout petit peu sur l'arithmétique, un premier exemple est donné par l'anneau $\mathbb{Z}/4\mathbb{Z}$ des entiers comptés modulo 4, où 2 est nilpotent d'indice 2.

Proposition 4.14 (Distributivité généralisée)**.** *Soit $(a_i)_{1 \leqslant i \leqslant m}$ et $(b_j)_{1 \leqslant j \leqslant n}$ deux familles finies à valeurs dans un anneau A. Alors on a*

$$\left(\sum_{i=1}^{m} a_i\right)\left(\sum_{j=1}^{n} b_j\right) = \sum_{(i,j) \in [\![1,m]\!] \times [\![1,n]\!]} a_i b_j$$

Proposition 4.15. *On peut généraliser la formule précédente à un nombre quelconque (fini) de sommes, et on obtient :*

$$\left(\sum_{i_1 \in I_1} a_{i_1,1}\right) \times \ldots \times \left(\sum_{i_n \in I_n} a_{i_n,n}\right) = \sum_{(i_1,\ldots,i_n) \in I_1 \times \ldots \times I_n} a_{i_1,1} \times \ldots \times a_{i_n,n}$$

Démonstration. Il s'agit d'un récurrence facile sur le nombre de sommes en utilisant la propriété précédente. □

Corollaire 4.3 (Développement d'un carré)**.** *Dans tout anneau, on a :*

$$\left(\sum_{i=1}^{n} a_i\right)^2 = \sum_{i=1}^{n} a_i^2 + \sum_{i \neq j} a_i a_j$$

Si de plus les a_i commutent deux à deux, on a :

$$\left(\sum_{i=1}^{n} a_i\right)^2 = \sum_{i=1}^{n} a_i^2 + 2\sum_{i<j} a_i a_j$$

Dans le cas où deux éléments a et b commutent, on a deux résultats importants.

Théorème 4.5. *Si a et b des éléments de A **commutent**, alors on a pour $n \in \mathbb{N}$:*

- *Formule du binôme de Newton :*

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

- *Formule de Bernoulli :*

$$a^n - b^n = (a-b) \sum_{k=0}^{n-1} a^{n-1-k} b^k = \left(\sum_{k=0}^{n-1} a^{n-1-k} b^k \right) (a-b)$$

2.3 Sous-anneaux

Définition 4.17 (Sous-anneau)**.** Un **sous-anneau** d'un anneau A est une partie $B \subset A$ telle que

- B est un sous-groupe de A
- $1 \in B$
- $\forall (x,y) \in B^2,\ xy \in B$

Dans ce cas, comme on le vérifie immédiatement, B muni des lois induites est lui-même un anneau.

Proposition 4.16 (Recette du sous-anneau)**.** *Soit A un anneau et $B \subset A$. Pour que B soit un sous-anneau de A, il faut et il suffit que*

- $1 \in B$
- $\forall (x,y) \in B^2,\ x - y \in B$
- $\forall (x,y) \in B^2,\ xy \in B$

Remarque 4.18. De même que pour les sous-groupes, on utilise principalement la condition suffisante, qui permet de montrer qu'un ensemble B est un anneau en l'interprétant comme un sous-anneau d'un anneau de référence.

Exemple 4.19. Tout sous-anneau d'un anneau commutatif est commutatif. Tout sous-anneau d'un anneau intègre est intègre.

Proposition 4.17. *Soit $\mathcal{B}$ un ensemble de sous-anneaux de A. Alors $\bigcap_{B \in \mathcal{B}} B$ est encore un sous-anneau de A.*

Définition 4.18 (Sous-anneau engendré)**.** Si X est une partie d'un anneau A, le **sous-anneau engendré par** X est l'intersection de tous les sous-anneaux contenant X. C'est le plus petit sous-anneau contenant X.

Exemple 4.20. Dans $\mathbb{Q}$, le sous-anneau engendré par $\dfrac{1}{2}$ est égal à $\left\{ \dfrac{k}{2^n} \mid (k,n) \in \mathbb{Z} \times \mathbb{N} \right\}$.

2.4 Morphismes

Définition 4.19 (Morphisme d'anneaux)**.** Un **morphisme** φ d'un anneau A dans un anneau B est une application qui transporte les deux lois, ainsi que l'élément neutre de la deuxième loi :

$$\begin{cases} \forall (x,y) \in A^2,\ \varphi(x+y) = \varphi(x) + \varphi(y) \\ \forall (x,y) \in A^2,\ \varphi(xy) = \varphi(x)\varphi(y) \\ \varphi(1_A) = 1_B \end{cases}$$

Proposition 4.18. *φ induit un morphisme du groupe $\mathcal{U}(A)$ dans le groupe $\mathcal{U}(B)$.*

Démonstration. En effet, si $x \in \mathcal{U}(A)$, alors $\varphi(x)$ est inversible, d'inverse $\varphi(x^{-1})$. □

Remarque 4.19. En particulier, on voit que φ est un morphisme du groupe $(A,+)$ dans le groupe $(B,+)$. Donc $\ker(\varphi)$ est toujours défini comme $\varphi^{-1}(\{0_B\})$, et il pourra toujours servir à montrer que φ est injectif (si, et seulement si, $\ker(\varphi) = \{0_A\}$). Attention à ne pas confondre les éléments neutres et à ne pas écrire $\ker(\varphi) = \varphi^{-1}(\{1_B\})$!

Proposition 4.19. *Soit $\varphi : A \to B$ un morphisme d'anneaux. On a les résultats suivants.*

- *Itération de la loi $+$:*
$$\forall x \in A,\ \forall n \in \mathbb{Z},\ \varphi(nx) = n\varphi(x)$$
- *Itération de la loi $\times$:*
$$\forall x \in A,\ \forall n \in \mathbb{N},\ \varphi(x^n) = \varphi(x)^n$$

Démonstration. Le premier point a déjà été prouvé. Le deuxième se démontre par une récurrence facile. □

Proposition 4.20. *L'image d'un sous-anneau par un morphisme d'anneaux est un sous-anneau. L'image réciproque d'un sous-anneau par un morphisme d'anneaux est un sous-anneau. On en déduit que le noyau et l'image d'un morphisme d'anneaux sont des sous-anneaux.*

Proposition 4.21. *La composée de deux morphismes d'anneaux est un morphisme d'anneaux. La bijection réciproque d'un isomorphisme d'anneaux est un isomorphisme d'anneaux.*

3 Corps

3.1 Premier exemple : construction de $\mathbb{Q}$, HP

Après avoir construit $\mathbb{N}$, on a éprouvé le besoin de construire $\mathbb{Z}$ parce que $\mathbb{N}$ "ne passait pas à l'inverse" pour l'addition. Pourrait-on effectuer le même travail sur la multiplication ? C'est ce qu'on va voir maintenant en construisant $\mathbb{Q}$. De même qu'on a écrit tout entier relatif comme différence de deux entiers naturels, on va écrire tout rationnel comme quotient de deux entiers relatifs. A un rationnel, on associe donc un couple d'entiers.

Proposition 4.22. *Sur l'ensemble $\mathbb{Z} \times \mathbb{Z}^\star$, on considère la relation définie par*

$$(p,q) \sim (p',q') \iff pq' = qp'$$

C'est une relation d'équivalence.

Définition 4.20 (Ensemble $\mathbb{Q}$)**.** On note $\mathbb{Q}$ l'ensemble $\mathbb{Z} \times \mathbb{Z}^\star$ quotienté par $\sim$. Ses éléments sont appelés les **rationnels**, il seront notés $\overline{(p,q)}$, ou encore plus simplement $\frac{p}{q}$. En particulier, si $k \in \mathbb{Z}^\star$, on peut "simplifier en haut et en bas par k" :

$$\frac{pk}{qk} = \frac{p}{q}$$

Proposition 4.23 (Addition sur $\mathbb{Q}$)**.** *On définit une addition sur $\mathbb{Q}$ par :*

$$\frac{p}{q} + \frac{r}{s} = \frac{ps + qr}{qs}$$

Cette opération est bien définie.

Remarque 4.20. On a $\frac{a}{n} + \frac{b}{n} = \frac{an + nb}{n^2} = \frac{a+b}{n}$.

Proposition 4.24 (Multiplication sur $\mathbb{Q}$)**.** *On définit une multiplication sur $\mathbb{Q}$ par :*

$$\frac{p}{q} \times \frac{r}{s} = \frac{pr}{qs}$$

Cette opération est bien définie.

Proposition 4.25 (Inclusion abusive $\mathbb{Z} \subset \mathbb{Q}$)**.** *L'application $n \mapsto \frac{n}{1}$ de $\mathbb{Z}$ dans $\mathbb{Q}$ est injective. Elle permet donc de voir $\mathbb{Z}$ comme un sous-ensemble de $\mathbb{Q}$. Avec cette inclusion, l'addition et la multiplication sur $\mathbb{Q}$ prolongent celles sur $\mathbb{Z}$.*

Théorème 4.6. *$(\mathbb{Q}, +, \times)$ est un anneau commutatif.*

La grande nouveauté de $\mathbb{Q}$ par rapport à $\mathbb{Z}$ réside dans l'inversibilité pour la deuxième loi.

Théorème 4.7. *Tout rationnel non nul est inversible pour la multiplication.*

Théorème 4.8. *On peut prolonger $\leqslant$ à $\mathbb{Q}$ tout entier. Le relation d'ordre obtenue reste total, et compatible avec l'addition et la multiplication au même sens que dans $\mathbb{Z}$.*

Démonstration. Ce théorème est démontré dans les compléments. □

3.2 Généralisation

Définition 4.21 (Corps)**.** Un **corps** $(K, +, \times)$ est un anneau commutatif, non trivial, tel que tout élément non nul est inversible pour la deuxième loi. On note $K^\star = K \setminus \{0\}$.

Remarque 4.21. Certaines sources un peu anciennes ne mentionnent pas la commutativité. Mais progressivement, cela tend à disparaître, afin de s'aligner sur la terminologie anglo-saxonne (selon laquelle tout corps est commutatif, la traduction exacte étant le mot *field*). Un "corps" dont la deuxième loi ne serait pas nécessairement commutative est maintenant appelé un **corps gauche** (*skew field* en anglais), et la terminologie semble bien stabilisée.

De toute manière, le théorème de Wedderburn montre que s'il est fini, tout corps gauche est commutatif. Il faut donc aller chercher des corps gauches infinis pour espérer rencontrer un cas de non-commutativité. Si on le souhaite, on pourra s'intéresser aux quaternions (cf. TD Matrices).

Remarque 4.22. Réciproquement, si x est inversible, alors $x \neq 0$, puisque $\forall y \in K,\ 0 \times y = 0 \neq 1$. Donc $(K^\star, \times)$ est égal au groupe des unités de K. En particulier, c'est un groupe.

Exemple 4.21. $\mathbb{Q}$, $\mathbb{R}$ et $\mathbb{C}$ sont des corps.

Définition 4.22 (Corps des fractions, HP)**.** Le travail que nous avons effectué pour passer de $\mathbb{Z}$ à $\mathbb{Q}$ peut s'effectuer avec n'importe quel anneau intègre A. En quotientant $A \times (A \setminus \{0\})$ par la relation définie par :

$$(a, b) \sim (c, d) \iff ad = bc$$

on obtient ce qu'on appelle le **corps des fractions de** A, et A se plonge naturellement dedans. Plus précisément, si $(a, b) \in A \times (A \setminus \{0\})$, la classe de (a, b) sera plutôt notée $\dfrac{a}{b}$, et tout élément $a \in A$ pourra être identifié à $\dfrac{a}{1}$.

Démonstration. Ce fait sera développé dans des compléments (de ce chapitre, ou du chapitre "Arithmétique", je ne sais pas encore) □

Remarque 4.23 (Localisation d'un anneau, HP)**.** Le corps des fractions est un cas particulier de ce que l'on appelle la **localisation d'un anneau** qui consiste à rendre inversible de force une partie stable par multiplication d'un anneau. Toutefois, le plongement n'est alors pas injectif, à moins que l'anneau de départ soit intègre. Si l'anneau de départ A est intègre et qu'on effectue ce travail sur la partie $A \setminus \{0\}$, on obtient le corps des fractions de A.

Remarque 4.24 (Calcul fractionnaire)**.** Si nous prenons un anneau intègre A et que nous construisons son corps des fractions K, rappelons que si $(a, b) \in A \times (A \setminus \{0\})$, on a

$$\begin{cases} \dfrac{1}{b} \times b = \dfrac{b}{b} = 1 \text{ donc } \dfrac{1}{b} = b^{-1} \\ \dfrac{a}{b} = \dfrac{a \times 1}{1 \times b} = \dfrac{a}{1} \times \dfrac{1}{b} = ab^{-1} = b^{-1}a \end{cases}$$

Plus généralement, dans un corps (et notamment dans $\mathbb{Q}$,$\mathbb{R}$ ou $\mathbb{C}$), l'inverse d'un élément $y \neq 0$ pourra être noté $\dfrac{1}{y}$, et l'écriture $\dfrac{x}{y}$ désignera indifféremment :

$$x \times \frac{1}{y} = \frac{1}{y} \times x$$

On vérifie immédiatement que les formules bien connus

$$\frac{x}{y} + \frac{x'}{y'} = \frac{xy' + x'y}{yy'} \quad , \quad \frac{x}{y} \times \frac{x'}{y'} = \frac{xx'}{yy'} \quad \text{et} \quad x \times \frac{x'}{y'} = \frac{xx'}{y}$$

sont justifiées. On en déduit que si $\dfrac{x}{y} \neq 0$, alors $\left(\dfrac{x}{y}\right)^{-1} = \dfrac{y}{x}$.

Proposition 4.26. *Tout corps est un anneau intègre.*

Définition 4.23 (Sous-corps)**.** Un **sous-corps** d'un corps K est une partie $L \in K$ telle que

- L est un sous-anneau de K

- $\forall x \in L,\ x \neq 0 \implies x^{-1} \in L$ (stabilité par passage à l'inverse d'un élément non nul)

Dans ce cas, comme on le vérifie immédiatement, L est lui-même un corps.

Proposition 4.27 (Recette du sous-corps)**.** *Soit K un anneau et $L \subset K$. Pour que L soit un sous-corps de K, il faut et il suffit que*

- $1 \in L$
- $\forall (x,y) \in L^2,\ x - y \in L$
- $\forall (x,y) \in L^2,\ y \neq 0 \implies xy^{-1} \in L$

Remarque 4.25. De même que pour les sous-groupes, on utilise principalement la condition suffisante, qui permet de montrer qu'un ensemble L est un corps en l'interprétant comme un sous-corps d'un corps de référence.

Proposition 4.28. *Soit $\mathcal{L}$ un ensemble de sous-corps de K. Alors $\bigcap_{L \in \mathcal{L}} L$ est encore un sous-corps de K.*

Définition 4.24 (Sous-corps engendré)**.** Si X est une partie de K, le **sous-corps engendré par** X est l'intersection de tous les sous-corps contenant X. C'est le plus petit sous-corps contenant X.

Exemple 4.22. Dans $\mathbb{R}$, le sous corps engendré par 1 est $\mathbb{Q}$. En effet, il contient 1 donc par stabilité, une récurrence immédiate montre qu'il contient $\mathbb{N}$, puis $\mathbb{Z}$. Par passage à l'inverse, il contient donc les nombres de la forme $\frac{1}{n}$, puis par stabilité il contient $\mathbb{Q}$. Comme $\mathbb{Q}$ est un corps, c'est le plus petit sous-corps de $\mathbb{R}$ qui contient 1.

Définition 4.25 (Morphisme de corps)**.** Un **morphisme** de corps est un morphisme pour les anneaux sous-jacents.

Exemple 4.23. Dans $\mathbb{C}$, la conjugaison est un automorphisme involutif de corps.

Remarque 4.26. On peut montrer sans difficulté que l'image d'un sous-corps par un morphisme de corps est un sous-corps. De même avec l'image réciproque.

Remarque 4.27. Attention ! Contrairement au cas des groupes et des anneaux, le produit cartésien de deux corps K_1 et K_2 n'est pas un corps ! En effet, $(1_{K_1}, 0_{K_2})$ et $(0_{K_1}, 1_{K_2})$ ne peuvent pas être inversés.

Proposition 4.29 (HP, mais à connaître sur le bout des doigts)**.** *Tout morphisme de corps est injectif.*

Démonstration. Soit φ un morphisme du corps K vers le corps L. Raisonnons par l'absurde : supposons que φ n'est pas injectif. On peut alors fixer $x \in \ker(\varphi) \setminus \{0_K\}$. Or, x est non nul, donc on a :

$$1_L = \varphi(1_K) = \varphi(xx^{-1}) = \varphi(x)\varphi(x^{-1}) = 0_L$$

par propriétés des morphismes de corps ainsi que par absorbance. Or, nos corps sont toujours supposés non triviaux, donc ceci est **absurde**. Ainsi, φ est bien injectif, ce qui conclut. □

Proposition 4.30 (HP, mais à connaître)**.** *Tout anneau intègre fini est un corps.*

Démonstration. Se donner un anneau intègre fini A et a un élément non nul de A. Considérer l'application

$$\begin{array}{rccc} \varphi_a : & A & \longrightarrow & A \\ & x & \longmapsto & ax \end{array}$$

Déjà, cette application est bien définie par stabilité de l'anneau par la deuxième loi. Ensuite, elle injective par intégrité de A. Enfin, elle est alors surjective par égalité de cardinaux. Donc on peut fixer un antécédent de 1_A par φ_a, ce qui montre que a est inversible. Ainsi, A est bien un corps car seule l'inversibilité des éléments non nuls manquait. □

4 Compléments sur les groupes et les corps, HP

4.1 Groupes finis

Pour commencer, voici la démonstration du théorème de Lagrange.

Théorème 4.9 (Théorème de Lagrange)**.** *Soit G un groupe fini, et H un sous-groupe de G. Alors,* $\mathrm{Card}(H)$ *divise* $\mathrm{Card}(G)$.

Démonstration. Introduisons la relation sur G :

$$\forall (x,y) \in G^2,\ x \sim y \iff x^{-1}y \in H$$

et vérifions que c'est une relation d'équivalence. Soit $(x,y,z) \in G^3$.

- Réflexivité : puisque $e \in H$, on a bien $x \sim x$.
- Symétrie : si $x^{-1}y \in H$, alors par passage à l'inverse $(x^{-1}y)^{-1} = y^{-1}x \in H$
- Transitivité : si $x^{-1}y \in$ et $y^{-1}z \in H$, alors par stabilité $(x^{-1}y)(y^{-1}z) = x^{-1}z \in H$

Ainsi, $\sim$ partitionne g en classes d'équivalence. Ces classes sont en nombre fini, car $G/\sim$ est l'image directe de G par la surjection canonique. Notons n le nombre de classes.

Par ailleurs, si $\overline{x}$ est une de ces classes d'équivalence, montrons qu'elle est égale à xH. Soit $y \in G$. On a

$$y \in \overline{x} \iff \exists h \in H,\ x^{-1}y = h \iff \exists h \in H,\ y = xh \iff y \in xH$$

Considérons alors l'application

$$\begin{array}{rccc} \varphi & : & H & \to & xH \\ & & h & \mapsto & xh \end{array}$$

- φ est injective par régularité de x
- φ est surjective par définition

Donc φ est bijective, si bien que $\mathrm{Card}(xH) = \mathrm{Card}(H)$. Ainsi, toutes les classes d'équivalence ont le même cardinal, à savoir $\mathrm{Card}(H)$. Finalement

$$\mathrm{Card}(G) = n \times \mathrm{Card}(H)$$

d'où le résultat. □

4.2 Retour sur la construction de $\mathbb{Q}$

Nous souhaitons maintenant prolonger $\leqslant_{\mathbb{Z}}$ à $\mathbb{Q}$. Pour cela, on sait bien intuitivement "quand une fraction est négative" : c'est que son numérateur et son dénominateur n'ont pas le même signe. Autrement dit, c'est quand le produit du numérateur est strictement négatif. De même, on sait "quand une fraction est plus petite qu'une autre : c'est que leur différence est négative ou nulle.

Proposition 4.31. *On choisit de définir $\frac{p}{q} \leqslant \frac{r}{s}$ par $(ps-qr)qs \leqslant_{\mathbb{Z}}$. cette définition est consistante, et $\leqslant$ est une relation d'ordre totale qui prolonge $\leqslant$ sur $\mathbb{Z}$. Enfin, on garde les mêmes compatibilités avec l'addition et la multiplication.*

Démonstration. Prouvons les résultats dans l'ordre de l'énoncé.

- Premier point : supposons qu'on ait plusieurs écritures du type $\frac{p}{q} = \frac{p'}{q'}$ et $\frac{r}{s} = \frac{r'}{s'}$, et montrons que
$$(ps - qr)qs \leqslant 0 \iff (p's' - q'r')q's' \leqslant 0$$
Par symétrie, il suffit de montrer le sens direct, le sens direct se démontrant en échangeant les variables. Supposons donc que $(ps - qr)qs \leqslant 0$, alors
$$\begin{aligned}(ps - qr)qsq'^2s'^2 &\text{donc } ps^2qq'^2s'^2 - q^2rsq'^2s'^2 \leqslant 0\\ &\text{donc } p's^2q^2q's'^2 - q^2r's^2q'^2s' \leqslant 0\\ &\text{donc } (p's'^2q' - r's'q'^2)q^2s^2 \leqslant 0\\ &\text{donc } (p's' - q'r')q's' \leqslant 0\end{aligned}$$
Voilà pour la consistance de la définition.
- Conséquence : si on veut comparer deux rationnels, en les "mettant au même dénominateur" on peut toujours se ramener à comparer $\frac{p}{q}$ et $\frac{p'}{q}$ avec $q > 0$, et l'inégalité $\frac{p}{q} \leqslant \frac{p'}{q}$ est alors équivalente à $(pq - p'q)q^2 \leqslant 0$, soit encore $p \leqslant p'$.
- Axiomes d'une relation d'ordre totale : ils viennent directement du point précédent.
- Le prolongement de $\leqslant_{\mathbb{Z}}$ est immédiat :
$$\frac{m}{1} \leqslant \frac{n}{1} \iff (m \times 1 - 1 \times n) \times 1 \times 1 \leqslant 0 \iff m \leqslant n$$
- Compatibilité avec l'addition : il s'agit de montrer l'assertion du type
$$\frac{p}{q} \leqslant \frac{p'}{q'} \implies \frac{p}{q} + \frac{r}{s} \leqslant \frac{p'}{q'} + \frac{r}{s}$$
Là aussi, on suppose sans perte de généralité de $q = s$ et le résultat est alors immédiat.
- Le raisonnement est le même avec la multiplication.

Ainsi, la preuve est achevée. □

4.3 Caractéristique d'un corps

Dans ce paragraphe, on se place dans un corps K, et on note $n \cdot x$ la n-ème itérée de x au sens de l'addition (avec $n \in \mathbb{Z}$ et $x \in K$).

Définition 4.26 (Caractéristique)**.** S'il existe, le plus petit entier $p \in \mathbb{N}^\star$ tel que $p \cdot 1_K = 0_K$ est appelé la **caractéristique de** K. S'il n'existe pas, on dit par convention que K est de caractéristique nulle.

Exemple 4.24. $\mathbb{R}$ et $\mathbb{C}$ sont de caractéristique nulle. $\mathbb{Z}/p\mathbb{Z}$, où p est premier, est un corps de caractéristique p (cf. le chapitre "Arithmétique").

Théorème 4.10. *Un corps fini est de caractéristique non nulle.*

Démonstration. Soit n le cardinal du corps. Alors dans l'ensemble $\{m \cdot 1_K \mid m \in [\![0, n]\!]\}$, au moins deux éléments sont égaux par le principe des tiroirs. On peut donc écrire $m \cdot 1_K = m' \cdot 1_K$ avec $m < m'$, d'où on en déduit $(m' - m) \cdot 1_K = 0_K$. L'ensemble $\{q \in \mathbb{N}^\star \mid q \cdot 1_K = 0_K\}$ est alors non vide, et il admet un plus petit élément. □

Théorème 4.11. *Si la caractéristique d'un corps est non nulle, alors c'est un nombre premier.*

Démonstration. Notons p la caractéristique du corps, et supposons que $p \neq 0$. Si p n'était pas premier, il suffirait d'écrire $p = ab$, $1 \leqslant a, b < p$. On aurait alors

$$(a \cdot 1_K) \times (b \cdot 1_K) = 0_K$$

Comme K est intègre, on aurait $a \cdot 1_K = 0_K$ ou $b \cdot 1_K = 0_K$, en contradiction avec la minimalité de p. Donc p est un nombre premier. □

Proposition 4.32. *Soit K de caractéristique $p > 0$, $x \in K^\star$, et $m \in \mathbb{Z}$. Alors $m \cdot x = 0_K$ si, et seulement si, p divise m.*

Démonstration. Il suffit de montrer que x est d'ordre p dans le groupe $(K, +)$.

- D'une part, par les opérations usuelles, on

$$\begin{aligned} p \cdot x &= p \cdot (1_K \times x) \\ &= (p \cdot 1_K) \times x \\ &= 0_K \times x \\ &= 0_K \end{aligned}$$

- D'autre part, soit $m \in [\![1, p-1]\!]$. Toujours par les mêmes règles, on a

$$m \cdot x = (m \cdot 1_K) \times x$$

qui est non nul par intégrité.

Ainsi, la preuve est achevée. □

Chapitre 5

Arithmétique

L'arithmétique, ou science des nombres entiers, trouve ses origines chez les Babyloniens côté occidental et chez les Indiens et les Chinois côté oriental. C'est donc une branche très ancienne des mathématiques, qui a fourni et fournit encore certains des plus beaux problèmes (le grand théorème de Fermat, résolu seulement en 1994 par Andrew Wiles et son ancien étudiant Richard Taylor). La plupart comportent un énoncé extrêmement simple, alors que leur résolution, quand elle est connue, est souvent excessivement ardue et fait appel à des domaines *a priori* très éloignés (analyse complexe, courbes elliptiques, etc.). A titre d'exemple, on citera la célèbre "conjecture de Goldbach" qui, comme son nom l'indique, n'est toujours pas résolue à l'heure où M. Morlot écrit ce cours (été 2013) ni à celle où je tape ce cours (été 2024) : est-ce que tout nombre pair plus grand que 2 peut s'écrire comme somme de deux nombres premiers ? Il semblerait bien que oui, en tout cas c'est ce que montrent des tests menés sur ordinateur jusqu'à des bornes impressionnantes. Mais quant à démontrer le théorème en toute généralité, il faut encore attendre... Plus généralement, les nombres premiers sont un thème central de la recherche en théorie des nombres, et leur réalité sous-jacente n'est encore que très partiellement comprise.

1 Divisibilité

Lemme 5.1. *Les unités de* $\mathbb{Z}$ *sont* -1 *et* 1.

Définition 5.1 (Divisibilité)**.** Soit $(a, d) \in \mathbb{Z}^2$. Lorsqu'il existe $k \in \mathbb{Z}$ tel que $a = kd$, on dit que a est un **multiple** de d et que d **divise** a. On écrit $d \mid a$.

Exemple 5.1. les diviseurs de 1 sont 1 et -1.

Exemple 5.2. Soit $a, d, \lambda \in \mathbb{Z}$. Si $d \mid a$, alors $\lambda d \mid \lambda a$. Si $\lambda \neq 0$, la réciproque est également vraie.

Remarque 5.1. Si d divise a et b, alors d divise toute $\mathbb{Z}$-combinaison linéaire $au + bv$ avec $(u, v) \in \mathbb{Z}^2$.

Remarque 5.2 (Notations)**.** L'ensemble des multiples de d est donc égal à $d\mathbb{Z}$. Quant à l'ensemble des diviseurs de a, il est souvent noté $\mathcal{D}(a)$, et l'ensemble de ses diviseurs strictement positifs est souvent noté $\mathcal{D}^+(a)$. Cependant, il peut être prudent de rappeler ces notations avant de les utiliser.

Remarque 5.3. Soit $d \in \mathbb{Z}$. $d\mathbb{Z}$ est égal à l'ensemble des itérés de d pour la loi $+$. Il s'agit donc du sous-groupes de $(\mathbb{Z}, +)$ engendré par d.

Proposition 5.1. *Sur $\mathbb{N}$, la divisibilité est une relation d'ordre partielle.*

Remarque 5.4. Sur $\mathbb{Z}$, elle reste réflexive et antisymétrique.

Remarque 5.5. 1 est le plus petit élément pour cette relation d'ordre (car il divise tous les entiers naturels), et 0 est le plus grand élément (car tout entier naturel divise 0).

Remarque 5.6 (Lien avec la relation d'ordre $\leqslant$)**.** Dans $\mathbb{N}^\star$, si $m \mid n$, alors $m \leqslant n$. En pratique, il suffit de vérifier que $n \in \mathbb{N}^\star$.

Nous en venons à la définition du pgcd et du ppcm de deux entiers. Dans un premier temps, restreignons-nous au cas où les entiers sont non nuls.

Définition 5.2 (Pgcd, ppcm)**.** Soit a et b deux entiers non nuls.

- L'ensemble des diviseurs communs strictement positifs de a et b est non vide, puisqu'il contient au moins 1. De plus, il est majoré (par exemple par $|a|$). Il admet donc un plus grand élément noté $a \wedge b$. On l'appelle le **plus grand commun diviseur** (ou **pgcd**) de a et b.
- L'ensemble des multiples communs strictement positifs de a et b est non vide, puisqu'il contient au moins $|ab|$. Il admet donc un plus petit élément noté $a \vee b$. On l'appelle **plus petit commun multiple** (ou **ppcm**) de a et b.

On rappelle que pour tout $n \in \mathbb{Z}$, $n\mathbb{Z}$ est un sous-groupe de $\mathbb{Z}$ (c'est le sous-groupe engendré par n).

Théorème 5.1 (Sous-groupes de $\mathbb{Z}$)**.** *Réciproquement, pour tout sous-groupe H de $\mathbb{Z}$, il existe un unique $n \in \mathbb{N}$ tel que $H = n\mathbb{Z}$.*

Corollaire 5.1. *Soit a et b des entiers non nuls. Alors*

$$\begin{cases} a\mathbb{Z} + b\mathbb{Z} = (a \wedge b)\mathbb{Z} \\ a\mathbb{Z} \cap b\mathbb{Z} = (a \vee b)\mathbb{Z} \end{cases}$$

Ce corollaire justifie la définition suivante.

Définition 5.3. On choisit de prolonger les notions de pgcd et de ppcm dans le cas où a et b sont deux entiers quelconques. Puisque $a\mathbb{Z} + b\mathbb{Z}$ et $a\mathbb{Z} \cap b\mathbb{Z}$ sont des sous-groupes additifs de $\mathbb{Z}$, on décide que le pgcd $a \wedge b$ est l'unique $\delta \in \mathbb{N}$ tel que $a\mathbb{Z} + b\mathbb{Z} = \delta\mathbb{Z}$, et que le ppcm $a \vee b$ est l'unique $\mu \in \mathbb{N}$ tel que $a\mathbb{Z} \cap b\mathbb{Z} = \mu\mathbb{Z}$.

Exemple 5.3. On pourra montrer que $\forall n \in \mathbb{Z},\ n \wedge 0 = |n|$ et $n \vee 0 = 0$. De même, $n \wedge 1 = 1$ et $n \vee 1 = |n|$.

Remarque 5.7. En reprenant le principe des démonstrations des théorèmes précédents, on constate les faits suivants.

- D'une part, $a \wedge b$ reste un diviseur commun de a et b, et $a \vee b$ reste un multiple commun de a et b. En revanche, ils n'ont plus de raison d'être maximal et minimal au sens de $\leqslant$.
- D'autre part, par définition, $\exists (u, v) \in \mathbb{Z}^2,\ au + bv = a \wedge b$

Remarque 5.8 (Utile pour les exercices calculatoires)**.** On ne change pas $a \wedge b$ et $a \vee b$:

- En remplaçant a par $-a$ ou b par $-b$. En effet, $a\mathbb{Z} = (-a)\mathbb{Z}$ et $b\mathbb{Z} = (-b)\mathbb{Z}$;
- En échangeant a et b. En effet, $a\mathbb{Z} + b\mathbb{Z} = b\mathbb{Z} + a\mathbb{Z}$ et $a\mathbb{Z} \cap b\mathbb{Z} = b\mathbb{Z} \cap a\mathbb{Z}$. Ainsi, le pgcd et le ppcm sont **commutatifs**.

De même, comme on le vérifie immédiatement, $\wedge$ et $\vee$ sont **associatifs**. C'est une conséquence immédiate de l'associativité de $+$ et $\cap$ dans $\mathcal{P}(\mathbb{Z})$.

Proposition 5.2. *Soit $(a, b) \in \mathbb{Z}^2$. Les diviseurs communs de a et b sont exactement les diviseurs de $a \wedge b$. Les multiples communs de a et b sont exactement les multiples de $a \vee b$.*

Définition 5.4 (Pgcd et ppcm de plus de deux entiers)**.** Soit $(a_1, \ldots, a_n) \in \mathbb{Z}^n$. Pour définir leur pgcd et leur ppcm, on peut le faire par récurrence en se plaçant dans les structures algébriques $(\mathbb{Z}, \wedge)$ et $(\mathbb{Z}, \vee)$. De même, on peut définir par récurrence $a_1\mathbb{Z} + \ldots + a_n\mathbb{Z}$ et $a_1\mathbb{Z} \cap \ldots \cap a_n\mathbb{Z}$ en se plaçant dans les structures algébriques $(\mathcal{P}(\mathbb{Z}), +)$ et $(\mathcal{P}(\mathbb{Z}), \cap)$. Et alors, par une récurrence immédiate, on s'aperçoit que tout revient à fonctionner de la manière suivante.

- Le pgcd $a_1 \wedge \ldots \wedge a_n$ est l'unique $\delta \in \mathbb{N}$ tel que tel que $a_1\mathbb{Z} + \ldots + a_n\mathbb{Z} = \delta\mathbb{Z}$. Il divise chacun des a_i, et l'ensemble des diviseurs communs des a_i est exactement égal à l'ensemble des diviseurs de δ.
- Le ppcm $a_1 \vee \ldots \vee a_n$ est l'unique $\mu \in \mathbb{N}$ tel que tel que $a_1\mathbb{Z} \cap \ldots \cap a_n\mathbb{Z} = \mu\mathbb{Z}$. Il est multiple de chacun des a_i, et l'ensemble des multiples communs des a_i est exactement égal à l'ensemble des multiples de μ.

Remarque 5.9. Comme précédemment, si on note δ le pgcd des a_i, alors

$$\exists (u_1, \ldots, u_n) \in \mathbb{Z}^n,\ \sum_{i=1}^{n} u_i a_i = \delta$$

En tant qu'éléments inversibles, 1 et -1 divisent trivialement tout entier. Réciproquement, on a la

Définition 5.5 (Entiers premiers entre eux)**.** Deux entiers sont **premiers entre eux** lorsque leurs diviseurs communs sont exactement 1 et -1. Plus généralement, soit $(a_1, \ldots, a_n) \in \mathbb{Z}^n$. Les a_i sont dits **premiers entre eux dans leur ensemble** lorsque leurs diviseurs communs sont exactement 1 et -1.

Remarque 5.10. Soit $(a, b) \in \mathbb{Z}^2$.

- Pour montrer que a et b sont premiers entre eux, il suffit donc de se donner $d \in \mathcal{D}(a) \cap \mathcal{D}(b)$ puis de montrer que $d \in \{-1, 1\}$. En effet, l'inclusion réciproque est toujours vraie.
- On constate facilement que a et b sont premiers entre eux si, et seulement si, $\mathcal{D}^+(a) \cap \mathcal{D}^+(b) = \{1\}$. Une variante consiste donc à se donner $d \in \mathcal{D}^+(a) \cap \mathcal{D}^+(b)$, et à montrer que $d = 1$. En effet, là aussi, l'inclusion réciproque est toujours vraie.

Exemple 5.4. Deux entiers consécutifs sont toujours premiers entre eux.

Remarque 5.11. Soit $(a, b) \in \mathbb{Z}^2$. il est facile de voir qu'on a $a \wedge b = 1$ si, et seulement si, a et b sont premiers entre eux.

- Si $a \wedge b = 1$, alors tout diviseur de a et b doit diviser 1, donc être une unité de $\mathbb{Z}$.

- Réciproquement, si a et b sont premiers entre eux, alors $a \wedge b$, en tant que diviseur commun, est égal à ± 1. Et comme il est dans $\mathbb{N}$, il est égal à 1.

Plus généralement, des entiers sont premiers entre eux dans leur ensemble si, et seulement si, leur pgcd vaut 1.

Remarque 5.12. Si les a_i sont deux à deux premiers entre eux, alors ils sont en particulier premiers entre eux dans leur ensemble (au moins si $n \geqslant 2$), mais la réciproque est fausse ! Considérer par exemple le tripler $(2, 3, 4)$.

Proposition 5.3. *Soit $(a, b) \in \mathbb{Z}^2$ et $k \in \mathbb{Z}$. Alors*

$$(ka) \wedge (kb) = k(a \wedge b)$$

Exercice 5.1. On peut prouver le symétrique pour le ppcm. Il est recommandé de commencer traiter à part le cas $k = 0$.

Proposition 5.4. *Soit $(a, b) \in \mathbb{Z}^2$ et $\delta = a \wedge b$. Alors*

$$\exists (a', b') \in \mathbb{Z}^2, \ \begin{cases} a = \delta a' \\ b = \delta b' \\ a' \wedge b' = 1 \end{cases}$$

Remarque 5.13. Ces deux dernières propositions restent vraies avec plus de deux entiers relatifs.

Théorème 5.2 (Théorème de Bézout)**.** *Soit $(a, b) \in \mathbb{Z}$. Alors a et b sont premiers entre eux si, et seulement si,*

$$\exists (u, v) \in \mathbb{Z}^2, \ au + bv = 1$$

Corollaire 5.2 (Corollaire du théorème de Bézout)**.** *Si a est premier avec chacun des $(b_i)_{1 \leqslant i \leqslant n}$, alors il est premier avec leur produit.*

Exemple 5.5. Soit $(a_1, \ldots, a_m) \in \mathbb{Z}^m$ et $(b_1, \ldots, b_n) \in \mathbb{Z}^n$ tels que

$$\forall (i, j) \in [\![1, m]\!] \times [\![1, n]\!], \ a_i \wedge b_j = 1$$

Alors, une double application de ce corollaire montre que

$$(a_1 \times \ldots \times a_m) \wedge (b_1 \times \ldots \times b_n) = 1$$

En particulier, si $a \wedge b = 1$, alors $\forall (p, q) \in (\mathbb{N}^\star)^2, \ a^p \wedge b^q = 1$.

Proposition 5.5 (Généralisation du théorème de Bézout)**.** *Soit $(a_1, \ldots, a_n)$ une famille finie d'entiers. Alors les a_i sont premiers entre eux dans leur ensemble si, et seulement si,*

$$\exists (u_1, \ldots, u_n) \in \mathbb{Z}^n, \ \sum_{i=1}^{n} a_i u_i = 1$$

Étant donnés a et b, il est bon de connaître une méthode pratique pour déterminer un couple de Bézout (u, v) tel que $au + bv = a \wedge b$ (notons qu'il n'y a pas unicité). Cette méthode est fournie par **l'algorithme d'Euclide**, qui permet à la fois d'obtenir le pgcd de a et b et un couple (u, v) correspondant. Avant tout, démontrons le

Lemme 5.2 (Invariance du pgcd par transvection)**.** *Soit $(a, b, \lambda) \in \mathbb{Z}^3$. Alors on a :*

$$a \wedge b = a \wedge (b + \lambda a)$$

Remarque 5.14. Si $a, b \in \mathbb{N}^\star$, on en déduit une méthode pour calculer $a \wedge b$. En notant r le reste de la division euclidienne de a par b, le lemme montre que $a \wedge b = b \wedge r$. Ainsi, on se ramène à des entiers plus petits.

Méthode 5.1 (Algorithme d'Euclide pour calculer le pgcd de $a > 0$ et $b > 0$)**.** Soit $a > 0$ et $b > 0$.

1. On pose $a_0 = a$ et $b_0 = b$.
2. r_0 est le reste de la division euclidienne de a par b.
3. Tant que $r_n \neq 0$:
   - On pose $a_{n+1} = b_n$ et $b_{n+1} = r_n$;
   - r_{n+1} est le reste de la division euclidienne de a_{n+1} par b_{n+1}.
4. En sortie de boucle, en notant n_f le rang pour lequel $r_{n_f} = 0$, le pgcd cherché est b_{n_f} (ou de manière équivalente r_{n_f-1}).
5. Version "étendue" de l'algorithme : si à chaque étape on exprime r_n comme combinaison de a et b, à la fin on obtient une combinaison égale à $a \wedge b$.

Théorème 5.3 (Lemme de Gauss)**.** *Soit $(a, b, c) \in \mathbb{Z}^3$. Si a divise bc et si a et b sont premiers entre eux, alors a divise c.*

Corollaire 5.3. *Si les $(a_i)_{1 \leqslant i \leqslant n}$ divisent tous b et sont deux à deux premiers entre eux, alors leur produit divise b.*

Exercice 5.2. Montrer que si les a_i sont dans $\mathbb{N}$ en étant deux à deux premiers, alors

$$a_1 \vee \ldots \vee a_n = a_1 \times \ldots \times a_n$$

Définition 5.6 (Forme irréductible d'un rationnel)**.** Tout rationnel non nul s'écrit d'une unique façon sous la forme $\dfrac{p}{q}$ avec $(p, q) \in \mathbb{Z} \times \mathbb{N}^\star$ et $p \wedge q = 1$. Dans ce cas, on dit que le rationnel est écrit sous **forme irréductible**.

2 Nombres premiers

Définition 5.7 (Nombres premiers)**.** Un entier naturel $p \geqslant 2$ est un **nombre premier** lorsque ses seuls diviseurs positifs sont 1 et lui-même. Dans la suite de ce chapitre, l'ensemble des nombres premiers sera noté $\mathcal{P}$, mais il est bon de rappeler cette notation sur une copie.

Remarque 5.15. Par convention, 1 n'est donc pas premier, on verra plus tard pourquoi ce choix. Ensuite, à part 2, tous les nombres premiers sont impairs puisque tous les nombres pairs sont divisibles par 2.

Exemple 5.6. Les premiers nombres premiers sont $2, 3, 5, 7, 11, 13, \ldots$

Proposition 5.6. *Soit $n \geqslant 2$. les assertions suivantes sont équivalentes :*

1. *n n'est pas premier (on dit parfois que n est **composé**)*

2. $\exists d_1, d_2 \in \mathbb{N}^\star,\ n = d_1 d_2 \quad et \quad d_1, d_2 > 1$
3. $\exists d_1, d_2 \in \mathbb{N}^\star,\ n = d_1 d_2 \quad et \quad d_1, d_2 < n$
4. $\exists d_1, d_2 \in \mathbb{N}^\star,\ n = d_1 d_2 \quad et \quad 1 < d_1, d_2 < n$

Remarque 5.16. Soit p un nombre premier et $n \in \mathbb{Z}$. Si p ne divise pas n, alors p est premiers avec n. En particulier, deux nombres premiers distincts sont toujours premiers entre eux.

Proposition 5.7 (Lemme d'Euclide)**.** *Soit $(a_1, \ldots, a_n) \in \mathbb{Z}^n$ et p premier. Si p divise le produit des a_i, alors p divise l'un des a_i.*

Exemple 5.7. Soit $a \in \mathbb{Z}$, $n \in \mathbb{N}^\star$ et p premier. Alors p divise a^n si, et seulement si, p divise a.

Lemme 5.3. *Tout entier $n \geqslant 2$ admet un diviseur premier.*

Corollaire 5.4. *Si n n'est divisible par aucun nombre premier de l'intervalle $[\![2, \lfloor\sqrt{n}\rfloor]\!]$, alors n est premier.*

Remarque 5.17. Soit $(a, b) \in \mathbb{Z}^2$. Si a et b ne sont pas premiers entre eux, alors ils admettent un diviseur premier commun. Par l'absurde, cela permet régulièrement de montrer que deux entiers sont premiers entre eux.

Exercice 5.3. Soit a et b deux entiers premiers entre eux. Montrer que $a^2 + b^2$ et ab sont premiers entre eux.

Théorème 5.4. *Il existe une infinité de nombres premiers.*

Définition 5.8 (Valuation p-adique)**.** Soit $n \in \mathbb{Z}^\star$ et p premier. La **valuation p-adique** de n est définie par :

$$v_p(n) = \max\left\{k \in \mathbb{N} : p^k \mid n\right\}$$

Remarque 5.18. Si $k \leqslant v_p(n)$, alors p^k divise toujours n, et si $k > v_p(n)$, p^k ne divise jamais n. En clair, p^k divise n si, et seulement si, $k \leqslant v_p(n)$. Autrement dit, on a :

$$\left\{k \in \mathbb{N} : p^k \mid n\right\} = [\![0, v_p(n)]\!]$$

En particulier, p divise n si, et seulement si, $v_p(n) \geqslant 1$, et donc p ne divise pas n si, et seulement si, $v_p(n) = 0$.

Exemple 5.8. On a toujours $v_p(-n) = v_p(n)$.

Exemple 5.9. Si p est premier et $\alpha \in \mathbb{N}$, alors $v_p(p^\alpha) = \alpha$ (utiliser la stricte croissance de $k \mapsto p^k$).

Lemme 5.4. *Soit $n \in \mathbb{Z}^\star$, p premier et $k \in \mathbb{N}$. Alors*

$$v_p(n) = k \iff \exists n' \in \mathbb{Z},\ \begin{cases} n = p^k n' \\ n' \wedge p = 1 \end{cases}$$

Proposition 5.8. *Soit $m, n \in \mathbb{Z}^\star$. pour tout nombre premier p, on a :*

$$v_p(mn) = v_p(m) + v_p(n)$$

Exemple 5.10. Soit $k \in \mathbb{N}^\star$. Par une récurrence immédiate, on montre que

$$\forall (n_1, \dots, n_k) \in (\mathbb{Z}^\star)^k,\ v_p\left(\prod_{i=1}^{k} n_i\right) = \sum_{i=1}^{k} v_p(n_i)$$

En particulier,

$$\forall n \in \mathbb{Z}^\star,\ v_p(n^k) = k v_p(n)$$

Définition 5.9 (Décomposition en facteurs premiers)**.** Soit $n \geqslant 1$. Une **décomposition en facteurs premiers** (ou **DFP**) est une écriture de n sous la forme

$$n = \prod_{p \in I} p^{\alpha_p}$$

où I est une partie finie de $\mathcal{P}$ et les α_p sont dans $\mathbb{N}^\star$. Pour $n = 1$, on peut poser $I = \emptyset$, et on obtient ce qu'on appelle la "DFP vide".

Remarque 5.19. On évitera l'acronyme "DFP" sur une copie, car il n'est pas standard.

Théorème 5.5. *Tout entier $n \geqslant 1$ admet une décomposition en facteurs premiers.*

Proposition 5.9. *Soit $n \geqslant 1$, et une DFP de n écrite sous la forme $n = \prod_{p \in I} p^{\alpha_p}$. Alors, pour tout nombre premier p_0 :*

- *Si $p_0 \in I$, on a $v_{p_0}(n) = \alpha_{p_0}$*
- *Sinon, on a $v_{p_0}(n) = 0$*

Corollaire 5.5. *Soit $n \geqslant 1$. Alors sa DFP est unique.*

Remarque 5.20. Pour cette raison (existence et unicité de la DFP), on dit que $\mathbb{Z}$ est un anneau **factoriel**.

Remarque 5.21. Au passage, on a montré que p divise n si, et seulement si, $p \in I$.

Proposition 5.10. *Soit $m, n \geqslant 1$. Alors m et n sont premiers entre eux si, et seulement si, ils n'ont aucun facteur commun dans leurs DFP respectives.*

Proposition 5.11. *Soit $m, n \in \mathbb{Z}^\star$. Alors $m \mid n$ si, et seulement si, pour tout nombre premier p, on a $v_p(m) \leqslant v_p(n)$.*

Exemple 5.11. Si $m, n \in \mathbb{N}^\star$, alors $m = n$ si, et seulement si, pour tout nombre premier p, on a $v_p(m) = v_p(n)$.

Théorème 5.6. *Les diviseurs strictement positifs de $n = \prod_{p \in I} p^{\alpha_p}$ sont exactement les entiers $m = \prod_{p \in I} p^{\beta_p}$ avec $\forall p \in I,\ \beta_p \leqslant \alpha_p$.*

Corollaire 5.6. *$n = \prod_{p \in I} p^{\alpha_p}$ possède $\prod_{p \in I} (\alpha_p + 1)$ diviseurs strictement positifs.*

Remarque 5.22. En numérotant les éléments de I, on peut écrire la DFP de n sous la forme bien connue

$$n = p_1^{\alpha_1} \dots p_k^{\alpha_k}$$

On obtient alors que "tout entier $n \geqslant 1$ admet une unique DFP, à l'ordre des facteurs près". En pratique, c'est cette écriture que nous garderons dans la suite, car elle est plus commode pour les calculs. On s'entraînera avec profit à reformuler les résultats précédents dans ce formalisme.

Proposition 5.12. *Soit a et b deux entier non nuls. Pour tout nombre premier p, on a :*

$$\begin{cases} v_p(a \wedge b) = \min(v_p(a), v_p(b)) \\ v_p(a \vee b) = \max(v_p(a), v_p(b)) \end{cases}$$

Remarque 5.23. Ce résultat se généralise à plus de deux entiers.

Corollaire 5.7. *On a*

$$\forall (a,b) \in \mathbb{Z}^2,\ (a \wedge b)(a \vee b) = |ab|$$

Exemple 5.12. Soit $a, b > 0$, $\delta = a \wedge b$ et $\mu = a \vee b$. En écrivant $a = \delta a'$ et $b = \delta b'$ avec $a' \wedge b' = 1$, on en tire $\mu = \delta ab$.

Exemple 5.13. Si $a, b \in \mathbb{N}$ sont premiers entre eux, alors on retrouve que $a \vee b = ab$.

3 Congruences

Dans toute cette partie, n est un entier naturel strictement positif.

Définition 5.10 (Congruence modulo n)**.** On rappelle que a et b sont **congrus modulo** n, et on écrit $a \equiv b\ [n]$ (voire $a = b\ [n]$) lorsque $n \mid b - a$ (soit encore lorsque $b - a \in n\mathbb{Z}$).

Exemple 5.14. Soit $a, b, k \in \mathbb{Z}$ avec $k \wedge n = 1$. Si $ka \equiv kb\ [n]$, alors $a \equiv b\ [n]$.

Théorème 5.7. *La congruence modulo n est une relation d'équivalence sur $\mathbb{Z}$.*

Définition 5.11 ($\mathbb{Z}/n\mathbb{Z}$)**.** L'ensemble des classes d'équivalence modulo n (autrement dit l'ensemble-quotient de $\mathbb{Z}$ par la relation de congruence modulo n) est noté $\mathbb{Z}/n\mathbb{Z}$. C'est un ensemble fini de cardinal n, et ses éléments sont donnés par :

$$\mathbb{Z}/n\mathbb{Z} = \left\{\overline{0}, \overline{1}, \dots, \overline{n-1}\right\}$$

Remarque 5.24. Ainsi, deux entiers sont congrus modulo n si, et seulement si, ils sont le même reste dans la division euclidienne par n.

Proposition 5.13. *L'addition et la multiplication de $\mathbb{Z}$ passent au quotient : si $a \equiv a'\ [n]$ et $b \equiv b'\ [n]$, alors $a + b \equiv a' + b'\ [n]$ et $ab \equiv a'b'\ [n]$.*

Remarque 5.25. La compatibilité avec l'addition est vrai modulo n'importe quoi. Toutefois, elle tombe souvent en défaut pour la multiplication et ne fonctionne ici que parce que le modulo, *ie* n, est entier.

Exemple 5.15. Par une récurrence immédiate, on montre que pour tout $k \in \mathbb{N}$, si $a \equiv b\ [n]$, alors $a^k \equiv b^k\ [n]$. Attention à ne pas écrire cela pour $k < 0$, cela n'a aucun sens !

Cela justifie la

Définition 5.12. On munit $\mathbb{Z}/n\mathbb{Z}$ d'une addition et d'une multiplication comme suit.

$$\begin{cases} \overline{a} + \overline{b} = \overline{a+b} \\ \overline{a} \times \overline{b} = \overline{ab} \end{cases}$$

Ces opérations son bien définies : elles ne dépendent pas des représentants choisis.

Théorème 5.8. *$(\mathbb{Z}/n\mathbb{Z}, +, \times)$ est un anneau commutatif, d'élément neutre $\overline{1}$ pour $\times$.*

Remarque 5.26. Autrement dit, la surjection canonique $\varphi : x \mapsto \overline{x}$ devient par définition un morphisme d'anneaux. En particulier,

- par itération de la loi $+$, on a $\forall x \in \mathbb{Z},\ \forall k \in \mathbb{Z},\ k \cdot \overline{x} = \overline{kx}$;
- par itération de la loi $\times$, on a $\forall x \in \mathbb{Z},\ \forall k \in \mathbb{N},\ \overline{x}^k = \overline{x^k}$.

Exemple 5.16. L'application de $(\mathbb{Z}/n\mathbb{Z}, +)$ dans $(U_n, \times)$ qui à $\overline{k}$ associe $\exp\left(i\frac{2k\pi}{n}\right)$ est bien définie et est un isomorphisme de groupes.

Théorème 5.9. *$\overline{a}$ est inversible dans $\mathbb{Z}/n\mathbb{Z}$ si, et seulement si, $a \wedge n = 1$.*

Démonstration. Si $\overline{a}$ est inversible, on peut écrire $\overline{a}\overline{b} = \overline{1}$ puis $n \mid ab - 1$, donc d'après le théorème de Bézout, a et n sont premiers entre eux. Réciproquement, utilisons à nouveau le théorème de Bézout. Si $a \wedge n = 1$, on écrit $au + nv = 1$, puis en passant dans $\mathbb{Z}/n\mathbb{Z}$, on obtient $\overline{au} = \overline{1}$. □

Théorème 5.10. *Soit $n \in \mathbb{N}^\star$. les trois assertion suivantes sont équivalentes :*

1. *n est premier*
2. *$\mathbb{Z}/n\mathbb{Z}$ est un corps*
3. *$\mathbb{Z}/n\mathbb{Z}$ est un anneau intègre*

Démonstration. Si n est premier, donnons-nous $\overline{a} \neq \overline{0}$. Alors $a \wedge n = 1$, donc $\overline{a}$ est inversible. De plus, tout corps est un anneau intègre. Enfin, si $\mathbb{Z}/n\mathbb{Z}$ est intègre, supposons que n n'est pas premier ; On ne peut pas avoir $n = 1$, sans quoi $\mathbb{Z}/n\mathbb{Z}$ serait l'anneau nul, donc $n \geqslant 2$. Écrivons donc $n = ab$ avec $a, b < n$. On obtient $\overline{a} \times \overline{b} = \overline{0}$, ce qui est absurde par intégrité, donc n est premier. □

Pour finir, citons le célèbre "petit théorème de Fermat". Avant toute chose, démontrons le

Lemme 5.5. *Soit p un nombre premier et $a, b \in \mathbb{Z}$. Alors*

$$(a+b)^p \equiv a^p + b^p \ [p]$$

Démonstration. Soit $k \in [\![1, p-1]\!]$. p ne divise pas k, donc $p \wedge k = 1$. D'après la formule du capitaine, $k\binom{p}{k} = p\binom{p-1}{k-1}$ puis par le lemme de Gauss, p divise $\binom{p}{k}$. Il suffit alors d'utiliser le binôme de Newton, de passer dans $\mathbb{Z}/p\mathbb{Z}$ et d'éliminer alors les termes non extrémaux. □

Théorème 5.11 (Petit théorème de Fermat)**.** *Soit p un nombre premier et $a \in \mathbb{Z}$. Alors :*

$$a^p \equiv a \ [p]$$

Démonstration. Sans perte de généralité, on se ramène au reste r de a modulo p. puis il suffit d'effectuer une récurrence finie sur r en appliquant le lemme qui précède pour l'hérédité. □

4 Compléments d'arithmétique, HP

Dans ce paragraphe, on démontre bon nombre de résultats arithmétiques hors programme majeurs, qui permettent souvent de faciliter la résolution d'un grand nombre d'exercice. Il est donc important de connaître ces résultats et les grandes lignes de leurs démonstrations.

Pour commencer, voici une application d'un théorème que nous avons vu dans le cas des groupes.

Théorème 5.12 (Petit théorème de Fermat - autre démonstration)**.** *Soit p un nombre premier, et a non divisible par p. Alors $a^{p-1} \equiv 1\ [p]$*

Démonstration. Puisque, $\overline{a}$ appartient au groupe $(\mathbb{Z}/p\mathbb{Z})^\star$ qui est de cardinal $p-1$. Il suffit alors d'appliquer le "théorème de Fermat" vu dans le chapitre de théorie des groupes : un élément du groupe élevé à la puissance du cardinal du groupe vaut l'élément neutre. □

Définition 5.13 (Indicatrice d'Euler)**.** L'**indicatrice d'Euler** est la fonction définie par

$$\varphi(n) = \text{Card}\left(\{k \in [\![1, n]\!] \mid k \wedge n = 1\}\right)$$

Plus généralement, on peut alors démontrer le :

Théorème 5.13 (Théorème de Fermat-Euler)**.** *Si $a \wedge n = 1$, alors $a^{\varphi(n)} \equiv 1\ [n]$.*

Démonstration. Si a est premier avec n, alors $\overline{a}$ appartient au groupes des unités $\mathcal{U}\left(\mathbb{Z}/n\mathbb{Z}\right)$, qui est de cardinal $\varphi(n)$. il suffit alors d'appliquer le "théorème de Fermat" vu sur les groupes. □

Théorème 5.14 (Théorème des restes chinois)**.** *Si $a \wedge b = 1$, alors l'application*

$$\begin{array}{rccc} \psi : & \mathbb{Z}/ab\mathbb{Z} & \longrightarrow & \mathbb{Z}/a\mathbb{Z} \times \mathbb{Z}/b\mathbb{Z} \\ & \overline{x}[ab] & \longmapsto & (\overline{x}[a], \overline{x}[b]) \end{array}$$

est un isomorphisme d'anneaux.

Démonstration. Pour commencer, ψ est bien définie, car si $x \equiv x'\ [ab]$, alors en particulier $x \equiv x'\ [a]$ et $x \equiv x'\ [b]$. Ensuite, c'est évidemment un morphisme d'anneaux d'après les définitions des opérations sur les ensembles $\mathbb{Z}/m\mathbb{Z}$. Il reste à montrer que c'est un isomorphisme.

Pour montrer que ψ est injectif, supposons que $x \equiv 0\ [a]$ et $x \equiv 0\ [b]$, alors $x \equiv 0\ [ab]$ d'après le corollaire du lemme de Gauss.

Pour montrer que ψ est surjective, il suffit alors de comparer les cardinaux. Si on souhaite une preuve plus constructive, donnons $\alpha \in [\![1, a]\!]$, $\beta \in [\![1, b]\!]$, et cherchons x tel que $x \equiv \alpha\ [a]$ et $x \equiv \beta\ [b]$. Il suffit alors de considérer un couple de Bézout (u, v) tel que $ua + bv = 1$ puis de poser $x = ua\beta + vb\alpha$. □

Remarque 5.27. Ce théorème s'appelle ainsi car il trouve son origine dans une récréation mathématique publiée dans un ancien manuscrit chinois.

Corollaire 5.8. *Si $a \wedge b = 1$, alors $\varphi(ab) = \varphi(a)\varphi(b)$.*

Démonstration. Un élément d'un anneau-produit est inversible si, et seulement si, ses deux composantes le sont. Or, $\varphi(ab)$ représente le nombre d'inversibles de $\mathbb{Z}/ab\mathbb{Z}$, et $\varphi(a)\varphi(b)$ celui des éléments inversibles de $\mathbb{Z}/a\mathbb{Z} \times \mathbb{Z}/b\mathbb{Z}$. On conclut par le théorème des restes chinois. □

Le théorème qui suit constitue un résultat important, utile dans de nombreux cas (et notamment dans le calcul du **déterminant de Smith**).

Théorème 5.15. *On a l'identité suivante :*

$$\forall n \in \mathbb{N}^\star,\ \sum_{d|n} \varphi(d) = n$$

Démonstration. Soit $n \in \mathbb{N}^*$. On définit $E_n = \left\{ \frac{k}{n} \mid k \in [\![1, n]\!] \right\}$. On définit alors pour tout diviseur positif d de n l'ensemble $D_d = \left\{ \frac{k}{d} \mid k \in [\![1, d]\!], k \wedge d = 1 \right\}$. Il suffit alors de montrer l'égalité suivante, ce qui suffira à conclure par passage aux cardinaux : $E_n = \bigsqcup_{d|n} D_d$. En effet, chaque D_d est immédiatement en bijection avec l'ensemble par lequel est défini $\varphi(d)$. Passons à la preuve de cette égalité.

- Montrons qu'il s'agit là d'une union disjointe. Soit d_1 et d_2 des diviseurs distincts de n. Raisonnons par l'absurde : supposons qu'il existe un $x \in D_{d_1} \cap D_{d_2}$ et fixons-le. Conformément aux définitions de D_{d_1} et D_{d_2}, on peut fixer $(k_1, k_2) \in [\![1, d_1]\!] \times [\![1, d_2]\!]$ tel que

  $$x = \frac{k_1}{d_1} = \frac{k_2}{d_2} \text{ avec } k_1 \wedge d_1 = 1 \text{ et } k_2 \wedge d_2 = 1$$

  On a alors $k_1 d_2 = k_2 d_1$ puis, d'après le lemme de Gauss, on a $k_1 \mid k_2$ et $k_2 \mid k_1$, puis il s'ensuit que $k_1 \leq k_2$ et $k_2 \leq k_1$ car k_1 et k_2 sont des entiers naturels non nuls. Par antisymétrie, on obtient que $k_1 = k_2$, ce qui est **absurde**. L'union considérée est donc disjointe.
- Montrons l'inclusion directe. Soit $x \in E_n$. Fixons $k \in [\![1, n]\!]$ tel que $x = \frac{k}{n}$. On peut fixer alors k' et d tels que $k = k' \times (k \wedge n)$ et $n = d \times (k \wedge n)$. Puisque $k \leq n$ alors $k' \leq d$ puis $x = \frac{k'}{d}$ avec $d \mid n$ donc $x \in D_d$. L'inclusion directe est donc bien vérifiée.
- Montrons l'inclusion réciproque. Soit $x \in \bigsqcup_{d|n} D_d$. Fixons $d \mid n$ tel que $x \in D_d$. Fixons $k \in [\![1, d]\!]$ tel que $x = \frac{k}{d}$. Or, $d \mid n$ donc on peut fixer l tel que $dl = n$. On sait que $1 \leq k \leq d$ donc $1 \leq l \leq kl \leq n$. Puis

  $$x = \frac{kl}{n} \in E_n$$

Ainsi, la preuve est achevée. □

Théorème 5.16. *On a*

$$\forall n \in \mathbb{N}^*,\ \varphi(n) = n \prod_{p|n} \left(1 - \frac{1}{p}\right)$$

Démonstration. On expose ici une démonstration probabiliste de ce résultat, moins lourde que la démonstration arithmétique pure. Soit $n \in \mathbb{N}^*$. On considère $\Omega = [\![1, n]\!]$ muni de la probabilité uniforme qu'on note $\mathbb{P}$. Si $d \in [\![1, n]\!]$ est un diviseur positif de n, on note

$$A_d = \{x \in [\![1, n]\!] \mid \exists k \in [\![1, n]\!],\ x = kd\}$$

Soit $d \in [\![1, n]\!]$ un diviseur de n. On montre aisément que $\text{Card}\,(A_d) = \left\lfloor \frac{n}{d} \right\rfloor = \frac{n}{d}$. Il s'ensuit que

$$\mathbb{P}\,(A_d) = \frac{1}{d}$$

Ensuite, on note $p_1 < ... < p_r$ les diviseurs premiers de n rangés dans l'ordre croissant.

On montre tout d'abord que les événements $(A_{p_i})_{1 \leq i \leq r}$ sont mutuellement indépendants. En effet, soit $I \subset [\![1, r]\!]$. Soit $\omega \in \Omega$. On a alors, en vertu du corollaire du lemme de Gauss

$$\omega \in \bigcap_{i \in I} A_{p_i} \iff \forall i \in I,\ p_i \mid \omega \iff \prod_{i \in I} p_i | \omega \iff \omega \in A_{\prod_{i \in I} p_i}$$

Il s'ensuit alors, d'après notre premier résultat et puisque $\prod_{i \in I} p_i | n$:

$$\mathbb{P}\left(\bigcap_{i \in I} A_{p_i}\right) = \mathbb{P}\left(A_{\prod_{i \in I} p_i}\right) = \frac{1}{\prod_{i \in I} p_i} = \prod_{i \in I} \frac{1}{p_i} = \prod_{i \in I} \mathbb{P}\,(A_i)$$

L'indépendance est démontrée.

Enfin, on considère $A = \{m \in [\![1, n]\!] \mid m \wedge n = 1\}$ et on veut montrer que $\varphi(n) = n \prod_{p|n} \left(1 - \frac{1}{p}\right)$.

Montrons que $\frac{\varphi(n)}{n} = \prod_{p|n} \left(1 - \frac{1}{p}\right)$, ce qui suffira à conclure. Or, on a d'une part, puisque la probabilité est uniforme :

$$\mathbb{P}(A) = \frac{\varphi(n)}{n}$$

Et d'autre part, puisque $A = \bigcap_{i=1}^{r} \overline{A_{p_i}}$ et puisque les événements $\left(\overline{A_{p_i}}\right)_{1 \leq i \leq r}$ sont mutuellement indépendants, on a :

$$\mathbb{P}(A) = \mathbb{P}\left(\bigcap_{i=1}^{r} A_{p_i}\right) = \prod_{i=1}^{r} \mathbb{P}\left(\overline{A_{p_i}}\right) = \prod_{i=1}^{r} \left(1 - \mathbb{P}\left(A_{p_i}\right)\right) = \prod_{i=1}^{r} \left(1 - \frac{1}{p_i}\right)$$

D'où :

$$\frac{\varphi(n)}{n} = \prod_{p|n} \left(1 - \frac{1}{p}\right)$$

Ce qui achève la preuve. □

Voici pour finir un théorème assez célèbre, un des premiers qu'on rencontre en arithmétique élémentaire.

Théorème 5.17 (Théorème de Wilson)**.** *p est premier si, et seulement si, $(p-1)! \equiv -1\ [p]$.*

Démonstration. Si $(p-1)! \equiv -1\ [p]$, alors toute classe non nulle modulo p est inversible. En effet, il suffit d'écrire

$$\overline{x} \times \left(- \prod_{\substack{1 \leqslant i \leqslant p-1 \\ \overline{i} \neq \overline{x}}} \overline{i} \right) = \overline{1}$$

Ainsi, $\mathbb{Z}/p\mathbb{Z}$ est un corps et p est premier.

Réciproquement, supposons que p est premier. Le cas $p = 2$ étant immédiat, passons à $p \geqslant 3$. Toute classe $\overline{x} \neq \overline{0}$ est inversible puisque $\forall x \in [\![1, p-1]\!],\ x \wedge p = 1$. Cherchons les classes qui sont leur propre inverse. On a :

$$\overline{x}^2 = \overline{1} \iff (\overline{x} + \overline{1})(\overline{x} - \overline{1}) = \overline{0} \iff \overline{x} = \pm\overline{1}$$

par intégrité dans le corps $\mathbb{Z}/p\mathbb{Z}$. A part ces deux classes, toute autre classe non nulle a un inverse qui n'est pas elle-même. Dans le produit $\overline{1} \times \ldots \times \overline{p-1}$, mettons à part $\overline{1}$ et $\overline{p-1} = \overline{-1}$, et regroupons les autres par paires d'inverses mutuels. On obtient alors le produit de $\overline{1}$ et de $\overline{-1}$, d'où le résultat. □

5 Petite généralisation : idéaux et anneaux principaux, HP

Dans toute cette partie, l'anneau A est supposé commutatif.

Définition 5.14 (Idéal)**.** Un **idéal** I de A est un sous-groupe additif de A absorbant pour la multiplication :

$$\forall (a, i) \in A \times I,\ ai \in I$$

Exemple 5.17 (Noyau d'un morphisme d'anneaux commutatifs)**.** Le noyau d'un morphisme d'anneaux commutatifs est toujours un idéal. il s'agit évidemment d'un sous-groupe de A, et il est absorbant par absorbance de 0 dans l'anneau et par propriété de morphisme d'anneau pour le produit.

Proposition 5.14 (Idéal engendré par un élément)**.** *Soit $x \in A$. Alors xA est un idéal de A, appelé* ***idéal engendré par*** *x. on le note souvent (x).*

Démonstration. Montrons déjà que xA est un sous-groupe additif. Déjà, $0 = x \times 0 \in xA$. Ensuite, soit $xa, xa' \in xA$. Alors $xa - xa' = x(a - a') \in xA$. Pour l'absorbance, la commutativité est essentielle. Soit $xa \in A$ et $b \in A$. Alors $b(xa) = x(ba) \in xA$. □

On peut se demander si la réciproque est vraie. C'est ce qui fait l'objet de la

Définition 5.15 (Anneau principal)**.** Si réciproquement

- les seuls idéaux de A sont de la forme xA
- A est intègre

alors A est dit **principal**.

Exemple 5.18 ($\mathbb{Z}$ est principal)**.** Les idéaux de $\mathbb{Z}$ sont de la forme $n\mathbb{Z}$, avec $n \in \mathbb{N}$ (ce sont mêmes les sous-groupes additifs de $\mathbb{Z}$). En particulier, $\mathbb{Z}$ est principal.

Définition 5.16 (Divisibilité)**.** Soit A un anneau intègre. On dit que a divise b et on note $a \mid b$ lorsque $b \in aA$.

Définition 5.17 (Éléments premiers entre eux)**.** Soit a et b deux éléments de A. Les unités de A divisent toujours a et b. Si réciproquement, les seuls diviseurs communs de a et b sont les unités de A, on dit que a et b sont **premiers entre eux**.

Exemple 5.19. Dans $\mathbb{Z}$, les seuls unités étant -1 et 1, on retrouve bien la définition usuelle selon laquelle deux entiers sont premiers entre eux lorsque leurs seuls diviseurs communs sont -1 et 1.

Théorème 5.18. *La somme de deux idéaux est un idéal.*

Démonstration. On sait déjà que dans un groupe commutatifs, la somme de deux sous-groupes est un sous-groupe. La propriété d'absorbance se vérifie immédiatement par distributivité dans l'anneau de départ. □

Corollaire 5.9. *Dans un anneau principal, le théorème de Bézout (y compris sa généralisation à n éléments) reste vrai. a et b sont premiers entre eux si, et seulement si,*

$$\exists (u, v) \in A^2,\ ua + vb = 1$$

Démonstration. La démonstration s'adapte sans difficulté. Si $ua + bv = 1$, tout diviseur commun de a et b doit diviser 1, donc doit être une unité. Réciproquement, soit a et b premiers entre eux. On sait que $aA + bA$ est un idéal de A, donc il est de la forme δA. En particulier, a et b sont dans δA, donc δ est une unité. On écrit ensuite $au + bv = \delta$ puis

$$(\delta^{-1})a + (\delta^{-1})b = 1$$

Ainsi, la preuve est achevée. □

Corollaire 5.10. *Dans un anneau principal, le lemme de Gauss reste vrai. De même, les corollaires du théorème de Bézout et du lemme de Gauss restent également vrais.*

Démonstration. Les preuves sont intégralement transposables. □

Remarque 5.28. En revanche, la définition d'un pgcd et d'un ppcm étant plus délicate, nous ne nous attarderons pas sur ce point ici. Nous reviendrons dessus plus en détail dans les compléments sur les anneaux commutatifs et en verrons un exemple dans le chapitre des polynômes.

Chapitre 6

Compléments sur les anneaux commutatifs, HP

Avant de commencer ce chapitre *Way too HP* (comme dirait l'autre...), je tiens à remercier le taupin inconnu qui a un jour déposé ce polycopié de compléments sur les anneaux commutatifs de M. Lafitte, duquel toute cette partie (ou presque) est tirée, sur mon bureau à ND3. Si tu lis ceci un jour, je te suis très reconnaissant (et je sais que ce n'est pas mon co Nicolas, que j'embrasse au passage).

Dans tous ces compléments, les anneaux considérés seront commutatifs. Pour A un anneau commutatif, on notera 0 pour 0_A et 1 pour 1_A.

1 Idéal d'un anneau commutatif

1.1 Diviseurs de zéro

Certains éléments d'un anneau ont des propriétés particulières par rapport à la multiplication, ce qui justifie quelques définitions.

Définition 6.1 (Élément simplifiable, diviseur de zéro)**.** Soit A un anneau commutatif. On dit qu'un élément $a \in A$ non nul est **simplifiable** si

$$\forall b \in A,\ ab = 0 \implies b = 0$$

Dans le cas contraire, on dit que a est un **diviseur de zéro**.

Définition 6.2 (Anneau intègre)**.** On dit qu'un anneau A est **intègre** si A est non trivial, commutatif (ce qui est supposé dans tout le chapitre) et si tout élément non nul de A est simplifiable.

Définition 6.3 (Ensemble des inversibles)**.** On note $\mathcal{U}(A)$, ou $A^\times$ ou encore $A^\star$ (mais cette dernière notation est parfois ambigüe), l'ensemble des éléments inversibles de l'anneau A. On l'appelle l'ensemble des unités de A.

Proposition 6.1 (Groupes des unités)**.** *L'ensemble des éléments inversibles d'un anneau A est un groupe pour la multiplication. On l'appelle le* ***groupe des unités de*** *A.*

Démonstration. Déjà fait dans le chapitre "Groupes, anneaux, corps". □

Définition 6.4 (Relation d'association)**.** Pour a et b deux éléments d'un anneau commutatif A, on dit que a est **associé** à b s'il existe un élément inversible $u \in A^\times$ tel que $a = bu$. La relation d'association est une relation d'équivalence sur A.

Exemple 6.1. L'ensemble des associés à 1 est exactement égal au groupe des unités de A.

Proposition 6.2. *Un morphisme d'anneaux $f : A \to B$ (non nécessairement commutatifs)induit par restriction et corestriction un morphisme de groupes de $A^\times$ dans $B^\times$.*

Démonstration. Soit $f : A \to B$ un morphisme d'anneaux. Soit $a \in A^\times$ et b l'inverse de a. Comme $ab = ba = 1$, on a $1 = f(1) = f(a)f(b) = f(b)f(a)$. par suite, $f(a)$ est inversible, d'inverse $f(b)$. On peut donc définir la restriction de f à $A^\times$ et corestreinte à $B^\times$. C'est clairement un morphisme de groupes par héritage du morphisme d'anneaux. □

Définition 6.5 (Élément nilpotent, indice de nilpotence)**.** Soit A un anneau (non nécessairement commutatif). On dit que $a \in A$ est **nilpotent** lorsqu'il existe $n \in \mathbb{N}^\star$ tel que $a^n = 0$. On appelle alors **indice de nilpotence de** a le plus petit entier $p \in \mathbb{N}^\star$ vérifiant cette propriété.

1.2 Idéaux

Dans tous ces compléments, on traite uniquement le cas des anneaux commutatifs. Cette hypothèse permet de simplifier de nombreuses notions et de nombreuses preuves, mais il existe une théorie des idéaux à droite et à gauche.

Définition 6.6 (Idéal)**.** Un **idéal** I d'un anneau A est une partie de A vérifiant les deux points suivants :

- L'ensemble I est un sous-groupe de $(A, +)$.
- Pour tout $a \in I$ et tout $b \in A$, on a $ba \in I$. On dit que I est **absorbant**.

Exemple 6.2. $\{0\}$ et A sont des idéaux (triviaux) de A.

Définition 6.7 (Idéal engendré par un élément)**.** Pour $a \in A$, $(a) := \{ax \mid x \in A\}$ est un idéal. On dit que c'est **l'idéal engendré par** a. On le note aussi aA.

Proposition 6.3 (Recette de l'idéal)**.** *Comme (-1_A) est un élément de A, pour prouver qu'une partie I de A est un idéal, il suffit de montrer les propriétés suivantes :*

1. $0 \in I$
2. $\forall (a, b) \in I^2,\ a + b \in I$
3. $\forall (a, b) \in A \times I,\ ab \in I$

Démonstration. Le troisième point donne l'absorbance. Le troisième point donne aussi la stabilité par passage à l'opposé car $(-1_A) \in A$. Les deux premiers points, couplés à la stabilité par passage à l'inverse, donnent le caractère de sous-groupe. □

Exemple 6.3. Si $\mathbb{K}$ est un corps, ses seuls idéaux sont les idéaux triviaux. En effet, soit I un idéel à gauche de $\mathbb{K}$ distinct de de $\{0\}$ et soit a un élément non nul de I. Soit $b \in \mathbb{K}$. Comme a est non nul, on peut considérer l'élément ba^{-1} de $\mathbb{K}$ et, par définition d'un idéal à gauche $(ba^{-1})a \in \mathbb{K}$. On a donc $b \in I$ puis $I = \mathbb{K}$.

Exemple 6.4. Plus généralement, si un idéal contient un élément inversible, alors il est égal à A tout entier.

Proposition 6.4 (Idéaux de $\mathbb{Z}$)**.** *Si I est un idéal de $\mathbb{Z}$, alors il existe un unique $n \geqslant 0$ tel que $I = (n)$.*

Démonstration. On a déjà prouvé ceci car tout sous-groupe de $(\mathbb{Z}, +)$ est en fait aussi un idéal. □

1.3 Opérations sur les idéaux

On dispose d'un certain nombre d'opérations intéressantes sur les idéaux.

Proposition 6.5 (Intersection d'idéaux)**.** *Si I et J deux idéaux de A, alors l'ensemble $I \cap J$ est encore un idéal de A. Plus généralement, l'intersection d'une famille d'idéaux de A est encore un idéal de A.*

Démonstration. Il suffit de revenir à la définition. C'est exactement la même démonstration que pour l'intersection d'un nombre quelconque de sous-groupes. □

Proposition 6.6 (Idéal engendré par une partie)**.** *Si S est une partie de A, il existe un plus petit idéal (pour l'inclusion) I de A contenant S, qu'on appelle* ***idéal engendré par*** *S. On le note $< S >$. Cela signifie que I est un idéal contenant S contenant S et que si J est un idéel contenant S, alors J contient I.*

Démonstration. Comme dans le cas des groupes, il suffit de définir I comme l'intersection des idéaux de A qui contiennent S. C'est un idéal d'après la proposition précédente et il contient clairement S. C'est aussi le plus petit pour l'inclusion. □

Définition 6.8 (Famille presque nulle)**.** Soit $(a_S)_{s \in S}$ une famille d'éléments de A. On dit qu'elle est **presque nulle** s'il existe une partie finie S' de S telle que

$$\forall s \in S \setminus S',\ a_s = 0$$

Remarque 6.1. Dans le cas où S est elle-même une partie finie, toutes la familles d'éléments de A indexées par S sont presque nulles.

Proposition 6.7. *Avec les mêmes notations que dans la proposition précédente, $I =< S >$ est l'ensemble des combinaisons linéaires à coefficients dans A presque nulles d'éléments de S, c'est-à-dire l'ensemble des éléments de la forme*

$$\sum_{s \in S} a_s s$$

avec $(a_s)_{s \in S}$ une famille d'éléments de A presque nulle.

Démonstration. Notons J l'ensemble des combinaisons linéaires presque nulles d'éléments de S à coefficients dans A. Si $(a_s)_{s\in S}$ est une famille presque nulle d'éléments de A, alors $\sum_{s\in S} a_s s$ est un élément de tout idéal de A contenant S puisqu'un tel idéal est stable par $+$ et absorbant. Il reste donc à prouver que J est un idéal de A. D'une part, il contient $0 = \sum_{s\in S} 0 \times s$. D'autre part, si $x = \sum_{s\in S} a_s s$ et $y = \sum_{s\in S} b_s s$ sont deux éléments de J, alors la famille $(a_s + b_s)_{s\in S}$ est presque nulle et $x + y \in J$. Enfin, pour $a \in A$ et $x = \sum_{s\in S} b_s s$ un élément de J, on a $ax = \sum_{s\in S} (ab_s)s$ ce qui prouve que $ax \in J$. Finalement, on a montré que J est bien un idéal de A, et il est clair qu'il contient chaque élément de S en prenant pour $(a_s)_{s\in S}$ la famille nulle sauf pour un élément a_{s_0} qui vaut 1_A. □

Proposition 6.8. *Pour $a \in A$, l'idéal engendré par $\{a\}$ est noté (a) et on a $(a) = aA$.*

Démonstration. Immédiat par la proposition précédente. □

Proposition 6.9. *Le noyau d'un morphisme d'anneaux $f : A \to B$ est un idéal.*

Démonstration. Un morphisme d'anneaux étant un morphisme de groupes abéliens, $\ker(f)$ est un sous-groupe de A. De plus, pour $x \in \ker(f)$ et $a \in A$, on a $f(ax) = f(a)f(x) = f(a) \times 0 = 0$ donc $ax \in \ker(f)$. □

Proposition 6.10 (Image réciproque d'un idéal)**.** *Soit $f : A \to B$ un morphisme d'anneaux et J un idéal de B. l'image réciproque de J par f est un idéal de A.*

Démonstration. Comme $f(0) = 0 \in J$, on a $0 \in f^{-1}(J)$. Soit $(a, b) \in f^{-1}(J)^2$. On a $f(a + b) = f(a) + f(b) \in J$ car $f(a)$ et $f(b)$ sont dans J puisque J est stable par $+$. Enfin, pour $a \in A$ et $b \in f^{-1}(J)$, $f(ab) = f(a)f(b) \in J$ puisque $f(b) \in J$ et J est absorbant. □

Remarque 6.2. En revanche, l'image d'un idéal par un morphisme d'anneaux n'est pas forcément un idéal.

Proposition 6.11 (Somme de deux idéaux)**.** *Soit I et J deux idéaux de A. L'ensemble des sommes $a + b$ avec $a \in I$ et $b \in J$ est un idéal de A, noté $I + J$. C'est aussi l'idéal engendré par la partie $I \cup J$.*

Proposition 6.12 (Somme quelconque d'idéaux)**.** *Plus généralement, si $(I_s)_{s\in S}$ est une famille d'idéaux de A, l'ensemble des sommes presque nulles $\sum_{s\in S} a_s$ où, pour tout $s \in S$, $a_s \in I_s$ est un idéal de A, noté $\sum_{s\in S} I_s$. C'est aussi l'idéal de A engendré par la partie $\bigcup_{s\in S} I_s$*

Démonstration. On note $I = \sum_{s\in S} I_s$.

- Comme $0 = \sum_{s\in S} 0$ et comme $0 \in I_s$ pour tout s, $0 \in I$.

- Soit $a = \sum_{s\in S} a_s$ et $b = \sum_{s\in S} b_s$ deux éléments de I. On a $a + b = \sum_{s\in S} a_s + b_s$, et cette somme est, bien entendu, presque nulle. On a donc $a + b \in I$ puisque $a_s + b_s$ pour tout $s \in S$ par stabilité de I_s par $+$.
- Soit $a = \sum_{s\in S} a_s$ un élément de I et $b \in A$. $ba = \sum_{s\in S} ba_s$ avec $s \in S$, $bas \in I_s$ puisque I_s est absorbant.
- On a montré que I est un idéal de A.
- Reste à prouver I est l'idéal de A engendré par la partie $\bigcup_{s\in S} I_s$. Nous devons établir deux inclusions. Comme I est un idéal de A contenant la réunion des I_s, on a déjà $\left\langle \bigcup_{s\in S} I_s \right\rangle \subset I$. D'autre part, $\left\langle \bigcup_{s\in S} I_s \right\rangle$ est un idéal de A contenant chaque I_s donc il est stable par $+$ et absorbant ce qui montre que tout somme presque nulle $\sum_{s\in S} a_s$ où, pour tout $s \in S$, $a_s \in I_s$, est aussi un élément de $\left\langle \bigcup_{s\in S} I_s \right\rangle$.

□

Définition 6.9 (Idéal produit)**.** Soit I et J deux idéaux d'un anneau commutatif A. On appelle **idéal produit de I et J**, et on note $I.J$, l'idéal engendré par l'ensemble des xy où $x \in I$ et $y \in J$. C'est donc l'ensemble des sommes $\sum_{k\in E} x_k y_k$ avec E fini, les x_k dans I et les y_k dans J.

Remarque 6.3. Attention ! La notation IJ peut désigner l'ensemble des produits xy pour $x \in I$ et $j \in J$, mais cette structure n'est pas intéressante car n'est même pas un groupe en général. On prendra donc garde à noter l'idéal produit $I.J$ avec le point entre les deux idéaux, c'est le plus petit idéal qui contient alors IJ.

Proposition 6.13. *Soit I et J deux idéaux d'un anneau commutatif A. On a $I.J \subset I \cap J$.*

Démonstration. Il suffit de remarquer que $IJ \subset I \cap J$ par absorbance (avec IJ l'ensemble des produits xy pour $x \in I$ et $y \in J$). Puis $I.J$ contieznt IJ par définition donc $I.J \subset I \cap J$. □

Proposition 6.14. *Soit I et J deux idéaux d'un anneau commutatif A. On a :*

$$(I + J).(I \cap J) \subset I.J$$

Démonstration. Soit $a = i + j \in I + J$ et $b \in I \cap J$. Alors $ab = ib + jb \in I.J$ car $b \in J$ et $b \in I$. Puisque $(I + J).(I \cap J)$ est l'ensemble des sommes finies d'éléments de ce type et $I.J$ est un idéal, on a bien $(I + J).(I \cap J) \subset I.J$. □

Proposition 6.15. *Soit I et J deux idéaux d'un anneau commutatif A. Si I et J sont comaximaux (ie $I + J = A$, cf. le pararagraphe 2), alors $I.J = I \cap J$.*

Démonstration. En vertu d'une proposition précédente, il suffit de prouver que $I \cap J \subset I.J$. Or, d'après la proposition précédente et l'hypothèse, on a $A.(I \cap J) \subset I.J$. Puis, si on se donne $x \in I \cap J$, alors $1x \in A.(I \cap J)$ donc $x \in I \cap J$. On a alors $I.J = I \cap J$. □

Définition 6.10 (Idéal quotient)**.** Soit I et J deux idéaux d'un anneau commutatif A. On appelle **idéal quotient de I par J**, et on note $I : J$ l'ensemble défini par :

$$I : J = \{x \in A \mid xJ \subset I\}$$

Démonstration. Montrons qu'il mérite bien son nom d'idéal.

- $0J = \{0\} \subset I$ donc $0 \in I : J$.
- Soit $(a, b) \in (I : J)^2$. Soit $x \in (a + b)J$. Fixons $y \in J$ tel que $x = (a + b)y = ay + by$. Par hypothèse, $ay \in I$ et $by \in I$, donc par stabilité de I par $+$, $x = (a+b)y \in I$. Donc $(a+b)J \subset I$. Donc $a + b \in I : J$.
- Soit $a \in I : J$ et $b \in A$. Soit $x \in abJ$. Fixons $y \in J$ tel que $x = aby$. Comme J est un idéal, $by \in J$ donc $x \in aJ$. Donc $x \in I$. Donc $abJ \subset I$. Donc $ab \in I : J$.

Ainsi, $I : J$ est bien un idéal. □

1.4 Radical d'un idéal

Définition 6.11 (Nilradical)**.** Le **nilradical** d'un anneau commutatif est l'ensemble de ses éléments nilpotents.

Définition 6.12 (Radical d'un idéal)**.** Plus généralement, on définit le **radical** d'un idéal I de A commutatif par

$$\sqrt{I} := \{a \in A \mid \exists n \in \mathbb{N}^\star,\ a^n \in I\}$$

Le nilradical de A est donc le radical de l'idéal nul.

Proposition 6.16. *Le radical $\sqrt{I}$ d'un idéal I de A est un idéal de A qui contient I.*

Démonstration. On voit immédiatement qu'on a $I \subset \sqrt{I}$ puisqu'il suffit de prendre $n = 1$ dans la définition de $\sqrt{I}$ pour tous les éléments de I. Montrons désormais que $\sqrt{I}$ est un idéal.

- Comme $0^1 = 0 \in I$, $0 \in \sqrt{I}$.
- Soit $a \in \sqrt{I}$ et $b \in \sqrt{I}$. Il existe n et m dans $\mathbb{N}^\star$ tels que $a^n \in I$ et $b^m \in I$, fixons-les. Comme l'anneau est commutatif, on peut utiliser la formule du binôme de Newton :

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

  Dans cette somme, tous les termes appartiennent à I. En effet, c'est vrai pour ceux d'indice $k \geqslant n$ puisque $a^k = a^n a^{n-k}$ et $a^n \in I$. De même, pour ceux d'indice $k \leqslant n$, on a $n+m-k \geqslant m$ et b$n + m - k = b^m b^{n-k}$ et $b^m \in I$. On a donc $(a + b)^{n+m} \in I$, d'où $a + b \in I$.
- Enfin, si $a \in \sqrt{I}$ et $b \in A$, choisissons $n \in \mathbb{N}^\star$ tel que $a^n \in I$. Alors, $(ba)^n = b^n a^n \in I$ et $ba \in \sqrt{I}$.
- $\sqrt{I}$ est donc bien un idéal de A.

□

Exemple 6.5. En particulier, le nilradical de d'un anneau commutatif est donc un idéal.

Définition 6.13 (Idéal radiciel)**.** Soit A un anneau commutatif. On dit qu'un idéal I de A est **radiciel** lorsqu'il est égal à son radical, *ie* I est radiciel lorsque $I = \sqrt{I}$.

Proposition 6.17 (Radicialité des radicaux)**.** *Soit A un anneau commutatif et I un idéal de A. On a $\sqrt{\sqrt{I}} = \sqrt{I}$. Autrement dit, le radical d'un idéal est radiciel.*

Démonstration. On a déjà l'inclusion $\sqrt{I} \subset \sqrt{\sqrt{I}}$. Montrons l'inclusion réciproque. Soit $x \in \sqrt{\sqrt{I}}$. Par hypothèse, il existe $n \in \mathbb{N}^\star$ tel que $x^n \in \sqrt{I}$. Fixons n. Or, puisque $x^n \in \sqrt{I}$, il existe $m \in \mathbb{N}^\star$ tel que $(x^n)^m \in I$. Alors, $nm \in \mathbb{N}^\star$ et $x^{nm} \in I$ donc $x \in \sqrt{I}$. On a bien $\sqrt{\sqrt{I}} \subset \sqrt{I}$, d'où lr résultat. □

Proposition 6.18 (Croissance des radicaux)**.** *Soit A un anneau commutatif et I et J deux idéaux de A. SI $I \subset J$, alors :*

$$\sqrt{I} \subset \sqrt{J}$$

Démonstration. Supposons que $I \subset J$. Soit $x \in \sqrt{I}$. Il existe $n \in \mathbb{N}^\star$ tel que $x^n \in I$. Fixons-le. Alors $x^n \in J$ par hypothèse, et comme $n \in \mathbb{N}^\star$, $x \in \sqrt{J}$. On a donc $\sqrt{I} \subset \sqrt{J}$. □

Proposition 6.19. *Soit I et J deux idéaux d'un anneau commutatif A. On a :*

$$\sqrt{I} + \sqrt{J} \subset \sqrt{I + J}$$

Démonstration. Soit $x \in \sqrt{I}$ et $y \in \sqrt{J}$. Fixons $m, n \in \mathbb{N}^\star$ tels que $x^m \in I$ et $y^n \in J$. Grâce au binôme de Newton puis en factorisant, on montre aisément que $(a + b)^{m+n} \in I + J$ donc $a + b \in \sqrt{I + J}$ ce qui conclut. □

Proposition 6.20 (Radical et intersection)**.** *Soit A un anneau commutatif et I et J deux idéaux de A. On a :*

$$\sqrt{I \cap J} = \sqrt{I} \cap \sqrt{J}$$

Démonstration. Déjà, $I \cap J$ est un idéal tel que $I \cap J \subset I$ et $I \cap J \subset J$. Par croissance, $\sqrt{I \cap J} \subset \sqrt{I}$ et $\sqrt{I \cap J} \subset \sqrt{J}$ donc $\sqrt{I \cap J} \subset \sqrt{I} \cap \sqrt{J}$. Montrons l'inclusion réciproque. Soit $x \in \sqrt{I} \cap \sqrt{J}$. Alors il existe $m \in \mathbb{N}^\star$ et $n \in \mathbb{N}^\star$ tels que $x^n \in I$ et $x^m \in J$. Alors, par absorbance de I et J, $x^{\max(m,n)} \in I \cap J$. Or, $\max(m, n) \in \mathbb{N}^\star$, donc $x \in \sqrt{I \cap J}$. On a bien $\sqrt{I \cap J} = \sqrt{I} \cap \sqrt{J}$. □

Proposition 6.21. *Soit I et J deux idéaux d'un anneau commutatif A. On a :*

$$\sqrt{I.J} = \sqrt{I \cap J} = \sqrt{I} \cap \sqrt{J}$$

Démonstration. Il suffit de montrer la première égalité. On sait que $I.J \subset I \cap J$ donc par croissance des radicaux, on a $\sqrt{I.J} \subset \sqrt{I \cap J} = \sqrt{I} \cap \sqrt{J}$. Réciproquement, soit $x \in \sqrt{I} \cap \sqrt{J}$. Fixons $m, n \in \mathbb{N}^\star$; tels que $x^m \in I$ et $x^n \in J$. Alors $x^{m+n} \in IJ$ donc $x^{m+n} \in I.J$. Donc $x \in \sqrt{I.J}$. On a bien

$$\sqrt{I.J} = \sqrt{I \cap J} = \sqrt{I} \cap \sqrt{J}$$

□

2 Anneaux quotients

Étant donné un anneau et une relation d'équivalence convenable sur cet anneau, l'objectif est de munir l'ensemble des classes d'équivalence d'une structure d'anneau. Cela revient en fait à "rendre nuls" les éléments d'un idéal de l'anneau sans modifier les autres règles de calcul.

2.1 Construction

Rappelons qu'une relation $\mathcal{R}$ sur un ensemble X est dite relation d'équivalence si elle est réflexive, symétrique et transitive. L'ensemble des classes d'équivalence de X pour la relation $\mathcal{R}$ est noté $X/\mathcal{R}$.

Soit maintenant un anneau A. On peut alors chercher les relations d'équivalence sur A qui sont **compatibles** avec la structure d'anneau. Autrement dit, on veut que soit vérifiée la propriété suivante :

$$\forall (x,y,x',y') \in A^4,\ (x\mathcal{R}y \wedge x'\mathcal{R}y') \implies ((x+x')\mathcal{R}(y+y') \wedge (xx')\mathcal{R}(yy'))$$

Proposition 6.22 (Analyse)**.** *Une relation d'équivalence compatible avec la structure d'anneau est nécessairement de la forme $\mathcal{R}_I$ avec I un idéal et*

$$\forall (x,y) \in A^2,\ x\mathcal{R}y \iff x-y \in I$$

Démonstration. Notons I la classe de 0. Si $x\mathcal{R}y$, comme $(-y)(-\dagger)$, alors $x-y\mathcal{R}0$, soit $x-y \in I$. Réciproquement, si $x-y \in I$, on a $(x-y)\mathcal{R}0$ et, puisque $y\mathcal{R}y$, il vient $x\mathcal{R}y$. Ce petit calcul montre que la relation $\mathcal{R}$ est définie par $x\mathcal{R}y$ si, et seulement si, $x-y \in I$.

Montrons par ailleurs que I est un idéal de A. On a déjà $0 \in I$. De plus, si $x \in I$ et $y \in I$, il vient $x\mathcal{R}0$ et $y\mathcal{R}0$ donc $(x+y)\mathcal{R}0$ ce qui prouve que $x+y \in I$. Enfin, si $x \in I$ et $a \in A$, on a $x\mathcal{R}0$ et $a\mathcal{R}a$ donc $(ax)\mathcal{R}(a\times 0)$ ce qui montre que $ax \in I$. Donc I est bien un idéal. □

Proposition 6.23 (Synthèse)**.** *Soit A un anneau commutatif et I un idéal de A. La relation $\mathcal{R}$ définie sur A par*

$$\forall (x,y) \in A^2,\ x\mathcal{R}y \iff x-y \in I$$

est une relation d'équivalence compatible avec la structure d'anneau. l'ensemble quotient $A/\mathcal{R}$ possède alors une structure canonique d'anneau.

De plus, la surjection canonique $\pi : A \to A/\mathcal{R}$ est un morphisme d'anneaux surjectif de noyau I.

L'anneau quotient $A/\mathcal{R}$ est noté A/I et, pour $a \in A$, on note $\overline{a}$ la classe d'équivalence de a modulo I s'il n'y a pas d'ambiguïté.

Démonstration. Il suffit d'effectuer les mêmes calculs dans l'autre sens. □

2.2 Idéaux comaximaux

Définition 6.14 (Idéaux comaximaux)**.** Deux idéaux I et J d'un anneau A sont dits **comaximaux** lorsque $I+J=A$.

Proposition 6.24 (Théorème chinois)**.** *On considère un anneau A ainsi que deux idéaux I et J de A comaximaux. Dans ce cas, le morphisme*

$$\begin{array}{rccc} f & : & A & \to & (A/I)\times(A/J) \\ & & a & \mapsto & (Cl_I(a), Cl_J(a)) \end{array}$$

est surjectif et son noyau est l'idéal $I\cap J$. Il en résulte, par passage au quotient, un isomorphisme d'anneaux :

$$\begin{array}{rccc} \tilde{f} & : & A/(I\cap J) & \to & (A/I)\times(A/J) \\ & & Cl_{I\cap J}(a) & \mapsto & (Cl_I(a), Cl_J(a)) \end{array}$$

Démonstration. Déjà, f est clairement un morphisme d'anneaux pour la structure d'anneau produit puisque la surjection canonique est morphisme d'anneaux. Ensuite, si on se donne $Cl_I(x)\in A/I$ et $Cl_J(y)\in A/J$, alors on peut fixer $i\in I$ et $j\in J$ tel que $x-y=i+j$ puisque I et J sont comaximaux. On pose alors $a:=x-i=y+j$. Puisque $a-x=-i\in I$ et $a-y=j\in J$, on a bien

$$f(a)=(Cl_I(x), Cl_J(y))$$

donc f est surjectif. Ensuite, soit $a\in A$. a appartient à $\ker(f)$ si, et seulement si, $Cl_I(a)=I$ et $Cl_J(a)=J$, donc a appartient à $\ker(f)$ si, et seulement si, $a\in I\cap J$. Donc on a bien $\ker(f)=I\cap J$.

Pour $\tilde{f}$, déjà cette fonction est bien définie car deux éléments de la même classe pour $I\cap J$ ont *a fortior* même classe pour I et pour J. Ensuite, es propriétés de morphisme et la surjectivité s'héritent de f. Enfin, $\tilde{f}$ est injective car son noyau est réduit à la classe de $I\cap J$, le neutre pour $+$ de l'anneau quotient $A/(I\cap J)$. □

Corollaire 6.1. *Soit I et J deux idéaux comaximaux d'un anneau A. Pour tout couple $(x,y)\in A^2$, il existe $a\in A$ tel que $a\in x+I$ et $a\in y+J$.*

Démonstration. Il suffit d'utiliser la surjectivité du morphisme f. □

2.3 Idéaux maximaux

Définition 6.15 (Idéaux propres)**.** Soit A un anneau commutatif. On appelle **idéal propre** de A tout idéal de A strictement inclus dans A.

Définition 6.16 (Idéaux maximaux)**.** Soit A un anneau commutatif. Un idéal I de A est dit **maximal** lorsqu'il est maximal parmi les idéaux propres (au sens de l'inclusion). En d'autres termes, I est un idéal propre et si J est un idéal propre tel que $I\subset J$, alors $J=I$.

Proposition 6.25. *Soit A un anneau commutatif. Un idéal I est maximal si, et seulement si, A/I est un corps.*

Démonstration. On prouve successivement les deux implications.

- $\Longrightarrow$: Supposons que I est maximal. On sait déjà que A/I est un anneau commutatif, il reste donc à prouver l'inversibilité des classes non nulles. Soit $\overline{x}\in A/I$ telle que $\overline{x}\neq\overline{0}$. Autrement dit, $x\notin I$. L'ensemble $I+xA$ est alors un idéal. Puisque $I\subset I+xA$ et $x\notin I$, alors $I\subsetneq I+xA$. Par maximalité de I, $I+xA$ ne peut pas être un idéal propre sans quoi on aurait $x\in I$. Donc les idéaux I et xA sont comaximaux, *ie* $I+xA=A$. On peut donc trouver $i\in I$ et $a\in A$ tels que $i+xa=1$. En passant au quotient, on obtient $\overline{xa}=\overline{1}$ si bien que $\overline{x}$ est inversible. Ainsi, A/I est un corps.

- $\Longleftarrow$: Réciproquement, supposons que A/I soit un corps et montrons que I est un idéal maximal. Soit J un idéal de A tel que $I \subsetneq J$. Montrons que $J = A$. Soit $x \in J \setminus I$. Alors $x \notin I$ donc $\overline{x} \neq \overline{0}$. Il s'ensuit que $\overline{x}$ est inversible dans A/I car A/I est un corps. On peut donc fixer $y \in A$ tel que $\overline{xy} = \overline{1}$. On peut donc fixer $z \in I$ tel que $xy - 1 = z$. On a donc $1 = xy - z$. Or, $x \in J$ donc $xy \in J$ et $z \in I$ donc $z \in J$. Donc $1 \in J$ puis $J = A$. Ainsi, nécessairement, I est un idéal maximal.

□

Exemple 6.6. Les idéaux maximaux de $\mathbb{Z}$ sont exactement les $n\mathbb{Z}$ où n est premier. En effet, soit I un idéal de $\mathbb{Z}$. Il est de la forme $I = n\mathbb{Z}$ avec $n \in \mathbb{N}$. Or, il est maximal si, et seulement si, $\mathbb{Z}/n\mathbb{Z}$ est un corps, c'est-à-dire si, et seulement si, n est un nombre premier.

2.4 Idéaux premiers

Définition 6.17 (Idéal premier)**.** Soit A un anneau commutatif. Un idéal I de A est dit premier lorsque tout produit d'éléments n'appartenant pas à I n'appartient pas à I, autrement dit lorsque $A \setminus I$ est stable par $\times$. Cela revient à dire que A/I est intègre.

Exemple 6.7. Dans $\mathbb{Z}$, l'idéal $\{0\}$ est premier car $\mathbb{Z}$ est un anneau intègre.

Proposition 6.26. *Soit A un anneau commutatif. Alors tout idéal maximal est premier.*

Démonstration. Soit I un idéal maximal de A. Soit $(x, y) \in (A \setminus I)^2$. Montrons que $xy \in A \setminus I$. On sait d'après le paragraphe précédent que A/I est un corps. Or, $\overline{x} \neq \overline{0}$ et $\overline{y} \neq \overline{0}$ donc, par intégrité du corps A/I, $\overline{xy} = \overline{x}\,\overline{y} \neq \overline{0}$. Donc $xy \notin I$, ce qui achève la preuve. Plus succinctement, si I est maximal alors A/I est un corps donc A/I est intègre donc I est premier. □

Remarque 6.4. Attention, la réciproque est fausse ! Par exemple, $\{0\}$ est un idéal premier de $\mathbb{Z}$ mais il n'est pas maximal car il est contenu strictement par l'idéal $2\mathbb{Z}$ qui est un idéal propre de $\mathbb{Z}$.

Proposition 6.27. *Tout idéal premier est radiciel.*

Démonstration. Soit A un anneau commutatif et I un idéal premier de A. On sait déjà que $I \subset \sqrt{I}$. Montrons l'inclusion réciproque. I est premier donc A/I est intègre. Soit $x \in \sqrt{I}$ et fixons $n \in \mathbb{N}^\star$ tel que $x^n \in I$. Alors $\overline{x}^n = \overline{0}$ donc par intégrité de A/I, $\overline{x} = \overline{0}$. Donc $x \in I$, si bien que $\sqrt{I} \subset I$. Un idéal premier est donc bien radiciel. □

Remarque 6.5. Attention : la réciproque est fausse !

3 L'anneau $\mathbb{Z}/n\mathbb{Z}$

Définition 6.18 ($\mathbb{Z}/n\mathbb{Z}$)**.** Pour $n \in \mathbb{N}^\star$, on note $\mathbb{Z}/n\mathbb{Z}$ l'anneau quotient de $\mathbb{Z}$ par l'idéal $n\mathbb{Z}$. Pour $x \in \mathbb{Z}$, on note $\overline{x}$ la classe de x. On a alors

$$\mathbb{Z}/n\mathbb{Z} = \left\{\overline{k} \mid k \in [\![0, n-1]\!]\right\}$$

De plus, la surjection canonique $\pi : \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ est un morphisme d'anneaux.

Remarque 6.6. De façon générale, il existe un unique morphisme d'anneaux f de $\mathbb{Z}$ dans un anneau A. En effet, $\mathbb{Z}$ est engendré par 1, comme $f(1) = 1$, cette égalité permet de définir sans ambiguïté f.

Proposition 6.28 (Inversible de $\mathbb{Z}/n\mathbb{Z}$)**.** *L'ensemble des éléments inversibles de $\mathbb{Z}/n\mathbb{Z}$ est*

$$(\mathbb{Z}/n\mathbb{Z})^\times = \left\{\overline{k} \mid k \in [\![0, n-1]\!],\ k \wedge n = 1\right\}$$

Démonstration. Il suffit d'utiliser le théorème de Bézout. □

Définition 6.19 (Indicatrice d'Euler)**.** On appelle **fonction indicatrice d'Euler** la fonction φ de $\mathbb{N}^\star$ dans $\mathbb{N}$ qui à n associe le nombre d'éléments inversibles de l'anneau $\mathbb{Z}/n\mathbb{Z}$, c'est-à-dire le nombre d'entiers $p \in [\![0, n-1]\!]$ premiers avec n.

Remarque 6.7. L'indicatrice d'Euler joue un grand rôle en arithmétique et en codage informatique comme par exemple dans le chiffrement RSA.

Exemple 6.8. $\varphi(2) = 1$, $\varphi(3) = 2$, $\varphi(4) = 2$. Pour p un nombre premier, on a $\varphi(p) = p - 1$. On a même la réciproque pour $p \geqslant 2$. Pour $\alpha \in \mathbb{N}^\star$ et p premier, on a $\varphi(p^\alpha) = p^{\alpha-1}(p-1)$.

Proposition 6.29. *Soit $n \in \mathbb{N}^\star$. Les trois assertions suivantes sont équivalentes :*

1. *n est premier*
2. *$\mathbb{Z}/n\mathbb{Z}$ est un corps*
3. *$\mathbb{Z}/n\mathbb{Z}$ est un anneau intègre*

Démonstration. Pour $1 \implies 2$, se donner $m \in [\![1, m-1]\!]$. n et m sont alors premiers entre eux et il suffit d'appliquer le théorème de Bézout avant de passer au quotient. $2 \implies 3$ est trivial, cela est toujours vrai. Pour $3 \implies 1$, supposons par l'absurde que n n'est pas premier et écrivons $n = ab$ avec $1 < a, b < n$. Alors $\overline{n} = \overline{a}\overline{b} = \overline{0}$ donc $\overline{a} = 0$ ou $\overline{b} = 0$ par intégrité, contradiction. Donc n est premier. □

Proposition 6.30 (Théorème de Fermat-Euler)**.** *Soit $n \in \mathbb{N}^\star$. pour tout $a \in \mathbb{Z}$ tel que $a \wedge n = 1$, on a :*

$$a^{\varphi(n)} \equiv 1\ [n]$$

Démonstration. Si a est premier avec n, alors $\overline{a}$ appartient au groupes des unités $\mathcal{U}(\mathbb{Z}/n\mathbb{Z})$, qui est de cardinal $\varphi(n)$. il suffit alors d'appliquer le "théorème de Fermat" vu sur les groupes (c'est le cas "facile" où le groupe est commutatif et où il suffit d'utiliser la bijection $x \mapsto ax$). □

Remarque 6.8. Pour $a \in \mathbb{Z}$ tel que $a \wedge n = 1$, on peut utiliser ce théorème pour déterminer un inverse modulaire de a qui n'est autre que $a^{\varphi(n)-1}$.

Proposition 6.31 (Théorème de Wilson)**.** *Soit $p \in \mathbb{N}^\star$. p est premier si, et seulement si, $(p-1)! \equiv -1\ [p]$.*

Démonstration. Si $(p-1)! \equiv -1\ [p]$, alors toute classe non nulle modulo p est inversible. En effet, il suffit d'écrire

$$\overline{x} \times \left(- \prod_{\substack{1 \leqslant i \leqslant p-1 \\ \overline{i} \neq \overline{x}}} \overline{i} \right) = \overline{1}$$

Ainsi, $\mathbb{Z}/p\mathbb{Z}$ est un corps et p est premier.

Réciproquement, supposons que p est premier. Le cas $p = 2$ étant immédiat, passons à $p \geqslant 3$. Toute classe $\overline{x} \neq \overline{0}$ est inversible puisque $\forall x \in [\![1, p-1]\!],\ x \wedge p = 1$. Cherchons les classes qui sont leur propre inverse. On a :

$$\overline{x}^2 = \overline{1} \iff (\overline{x} + \overline{1})(\overline{x} - \overline{1}) = \overline{0} \iff \overline{x} = \pm\overline{1}$$

par intégrité dans le corps $\mathbb{Z}/p\mathbb{Z}$. A part ces deux classes, toute autre classe non nulle a un inverse qui n'est pas elle-même. Dans le produit $\overline{1} \times \ldots \times \overline{p-1}$, mettons à part $\overline{1}$ et $\overline{p-1} = \overline{-1}$, et regroupons les autres par paires d'inverses mutuels. On obtient alors le produit de $\overline{1}$ et de $\overline{-1}$, d'où le résultat. □

Proposition 6.32 (Théorème chinois)**.** *Soit p et q deux nombres entiers premiers entre eux ainsi que $(y, z) \in \mathbb{Z}^2$. Alors, il existe un entier $x \in \mathbb{Z}$ tel que $y \equiv y\ [p]$ et $x \equiv z\ [q]$. De plus, toutes les solutions de ce système sont congrues modulo pq.*

Démonstration. On peut démontrer ce résultat "à la main", mais il s'agit d'un résultat général. En effet, puisque $p \wedge q = 1$, $p\mathbb{Z} + q\mathbb{Z} = \mathbb{Z}$ et ainsi les idéaux $p\mathbb{Z}$ et $q\mathbb{Z}$ sont comaximaux. Il suffit alors d'utiliser le théorème chinois vu dans la partie sur les idéaux comaximaux. On a en effet $p\mathbb{Z} \cap q\mathbb{Z} = pq\mathbb{Z}$ car p et q sont premiers entre eux. □

Corollaire 6.2. *Si p et q sont deux entiers premiers entre eux, alors il existe un isomorphisme d'anneaux de $Z/pq\mathbb{Z}$ sur $(\mathbb{Z}/p\mathbb{Z}) \times (\mathbb{Z}/q\mathbb{Z})$ (muni de la structure produit).*

Démonstration. Il s'agit encore là d'un résultat général, c'est la deuxième partie du théorème chinois sur les idéaux comaximaux. On peut évidemment le refaire à la main en comme dans la démontstration du théorème sur les idéaux comaximaux. □

Corollaire 6.3 (Multiplicativité de φ)**.** *Si p et q sont deux entiers premiers entre eux, alors $\varphi(pq) = \varphi(p)\varphi(q)$. On dit que φ est une fonction* ***arithmétique multiplicative****.*

Démonstration. Il suffit de comparer les cardinaux de $(\mathbb{Z}/pq\mathbb{Z})^\times$ et $((\mathbb{Z}/p\mathbb{Z}) \times (\mathbb{Z}/q\mathbb{Z}))^\times = (\mathbb{Z}/p\mathbb{Z})^\times \times (\mathbb{Z}/q\mathbb{Z})^\times$. En effet, on montre que la restriction et corestriction de l'isomorphisme fourni par le théorème précédent aux groupes des unités est un isomorphisme et on montre aisément l'égalité ensembliste énoncée précédemment. □

Corollaire 6.4 (Formule explicite de l'indicatrice d'Euler)**.** *Pour $n \in \mathbb{N}^\star$, on note S la partie finie de $\mathcal{P}$ constituée des nombres premiers p tel que $v_p(n) > 0$. On a alors :*

$$\varphi(n) = \prod_{p \in S} p^{v_p(n)-1}(p-1) = n \prod_{p \in S} \left(1 - \frac{1}{p}\right)$$

4 Anneaux principaux

On considère dans ce paragraphe uniquement des anneaux intègres (et toujours commutatifs).

4.1 Divisibilité dans un anneau intègre

Définition 6.20 (Divisibilité)**.** Soit A un anneau intègre et $(a,b) \in A^2$. On dit que a **divise** b (ou encore que b est un **multiple** de a) s'il existe $q \in A$ tel que $b = aq$. On note alors $a \mid b$.

La relation $\mid$ est réflexive et transitive (on dit que c'est une relation de **préordre** sur A) mais elle n'est pas antisymétrique en général. On a toutefois le résultat suivant :

Proposition 6.33. *Soit A un anneau intègre et $(a,b) \in A^2$. On a*

$$a \mid b \wedge b \mid a \iff \exists u \in A^\times,\ b = au$$

On a vu que cela revenait à dire que a et b sont associés.

Démonstration. Le sens réciproque est immédiat. Pour le sens direct, si $a = 0$ alors $b = 0$ et il suffit de prendre $u = 1$. Si $a \neq 0$, il existe $(k,k') \in A^2$ tel que $a = bk$ et $b = ak'$. On a donc $a(1-kk') = 0$ puis, par intégrité de A et par commutativité, k et k' sont inversibles et conviennent. □

Remarque 6.9. Rappelons que l'ensemble $(a) = aA$ des multiples de a est l'idéal engendré par a. On peut alors reformuler le résultat précédent de la manière suivante :

$$(a) = (b) \iff a \mid b \wedge b \mid a \iff \exists u \in A^\times,\ b = au$$

Proposition 6.34. *Soit $(a,b,c) \in A^3$ avec A un anneau commutatif et intègre. Si $c \mid a$ et $c \mid b$, alors*

$$\forall (u,v) \in A^2,\ c \mid au + bv$$

Démonstration. La vérification est immédiate. □

4.2 Anneaux principaux

Définition 6.21 (Idéal principal)**.** Un idéal I d'un anneau intègre A est dit **principal** s'il est engendré par un seul élément, c'est-à-dire s'il existe $a \in A$ tel que $I = (a)$.

Définition 6.22 (Anneau principal)**.** Un anneau intègre A est dit **principal** lorsque tous ses idéaux sont principaux.

Proposition 6.35. *$\mathbb{Z}$ est un anneau principal.*

Démonstration. Déjà prouvé car $\mathbb{Z}$ est intègre et tous es idéaux sont de la forme $n\mathbb{Z}$. □

Proposition 6.36. *$\mathbb{K}[X]$ est un anneau principal dès que le corps $\mathbb{K}$ est commutatif.*

Démonstration. La preuve est réalisée dans le chapitre "Polynômes". □

Ainsi, les deux objets que sont l'anneau des entiers relatifs $\mathbb{Z}$ et celui des polynômes à coefficients dans un corps commutatifs sont des anneaux principaux.

Dans les paragraphes suivants, on va définir le pgcd et le ppcm de deux éléments d'un anneau principal. On retrouvera alors les notions déjà vues dans $\mathbb{Z}$ et on utilisera la proposition précédente pour construire la même théorie pour les polynômes.

4.3 PGCD dans un anneau principal

Définition 6.23 (PGCD)**.** Soit A un anneau principal et $(a,b) \in A^2$. L'idéal $(a)+(b)$ est principal et, par conséquent, il existe $\delta \in A$ tel que $(a)+(b)=(\delta)$. L'élément δ n'est pas unique en général, mais il est unique à un coefficient multiplicatif inversible près. On dit que δ est **un pgcd** de a et b et on note $\delta = a \wedge b$.

Proposition 6.37. *Soit A un anneau principal et $(a,b) \in A^2$. On suppose que $(a,b) \neq (0,0)$. Dans ce cas, tout pgcd de a et b est non nul. Soit δ un tel pgcd. Toujours à un inversible près, δ est le "plus grand" (pour la relation $|$) diviseur commun à a et b.*

Démonstration. Soit d un diviseur commun à a et b. On a $(a)+(b) \subset (d)$ donc $(\delta) \subset (d)$ puis $d \mid \delta$. □

Remarque 6.10. Comme on l'a vu précédemment, la relation $|$ est seulement une relation de préordre sur A. Il n'y a donc pas véritablement de plus grand élément à l'ensemble des diviseurs. On peut avoir une véritable unicité si on travaille non plus dans A mais dans l'ensemble quotient $A/\mathcal{R}$ où $\mathcal{R}$ est la relation définie par $x\mathcal{R}y$ si, et seulement si, x est associé avec y. Dans ce cas, la relation $|$ induit une vraie relation d'ordre sur $A/\mathcal{R}$ mais la manipulation des objets de l'ensemble quotient étant plus compliquée, on préfère dire simplement "à un inversible près".

Exemple 6.9. Dans le cas où $A=\mathbb{Z}$, pour avoir une véritable unicité, on fait le choix de prendre le pgcd dans $\mathbb{N}$ donc positif.

Exemple 6.10. Dans le cas où $A=\mathbb{K}[X]$, pour avoir une véritable unicité, on fera le choix de prendre le polynôme unitaire (c'est-à-dire avec un coefficient dominant qui vaut 1).

Définition 6.24 (Ensemble des diviseurs)**.** Soit a un élément d'un anneau principal A. On note $\mathcal{D}(a)$ l'ensemble des diviseurs de a.

Proposition 6.38. *Soit a et b deux éléments d'un anneau principal A. On a :*

$$\mathcal{D}(a) \cap \mathcal{D}(b) = \mathcal{D}(a \wedge b)$$

Démonstration. Comme $\delta = a \wedge b \in (a)+(b)$, il existe $(u,v) \in A^2$ tel que $au+bv=\delta$. Fixons-les. Ainsi, tout diviseur de a et de b est aussi un diviseur de δ. De plus, $a \in (a)+(b)$ car $a=a+0$. Ainsi, $a \in (\delta)$ et $\delta \mid a$. De même, $\delta \mid b$. On a donc, par transitivité de la divisibilité, $D(\delta) \subset \mathcal{D}(a) \cap \mathcal{D}(b)$, ce qui achève la preuve. □

Proposition 6.39. *Soit A un anneau principal. A multiplication par un élément inversible près, on a :*

1. $\forall (a,b) \in A^2,\ a \wedge b = b \wedge a$ *(commutativité du PGCD)*
2. $\forall (a,b,c) \in A^3,\ (a \wedge b) \wedge c = a \wedge (b \wedge c)$ *(associativité du PGCD)*
3. $\forall (a,b,x) \in A^3,\ (xa) \wedge (xb) = x(a \wedge b)$

Démonstration.
1. Clair par la proposition précédente et par commutativité de l'intersection ensembliste.
2. Clair par la proposition précédente et par associativité de l'intersection ensembliste.

3. On a

$$\begin{aligned}((xa) \wedge (xb)) &= (xa) + (xb) \quad \text{(définition)} \\ &= x[(a) + (b)] \quad \text{(double inclusion)} \\ &= x(a \wedge b) \quad \text{(définition)} \\ &= (x(a \wedge b)) \quad \text{(double inclusion)}\end{aligned}$$

□

Proposition 6.40. *Soit a et b deux éléments d'un anneau principal A. On a :*

$$\forall q \in A,\ \mathcal{D}(a) \cap \mathcal{D}(b) = \mathcal{D}(b)\mathcal{D}(a - bq)$$

Ainsi, à multiplication par un élément inversible près :

$$a \wedge b = b \wedge (a - bq)$$

Démonstration. Exactement comme dans le cas de l'anneau $\mathbb{Z}$! □

Définition 6.25 (Éléments premiers entre eux)**.** Soit a et b deux éléments d'un anneau principal A. On dit que a et b sont **premiers entre eux** lorsque $a \wedge b = 1$ (c'est-à-dire lorsque les seuls diviseurs communs de a et b sont les unités de l'anneau).

Proposition 6.41 (Théorème de Bézout)**.** *Soit a et b deux éléments d'un anneau principal A. Les assertions suivantes sont équivalentes :*

1. *a et b sont premiers entre eux*
2. *Il existe $(u, v) \in A^2$ tel que $au + bv = 1$*

Démonstration. Si a et b sont premiers entre eux, alors $(a) + (b) = (1) = A$. Comme $1 \in (1)$, il existe $(u, v) \in A^2$ tel que $au + bv = 1$. Réciproquement, supposons qu'il existe $(u, v) \in A^2$ tel que $au+bv = 1$ et fixons-les. On a $au+bv \in (a)+b(a)$ donc $1 \in (a)+(b)$ donc $(1) \subset (a)+(b)$, c'est-à-dire $A \subset (a)+(b)$. Comme l'inclusion réciproque est toujours vraie, on a $(a)+(b) = (1)$ donc a et b sont premiers entre eux. □

Corollaire 6.5. *Soit A un anneau principal et $(a, b, c) \in A^3$. On suppose que $a \wedge b = 1$ et $a \wedge c = 1$. Dans ce cas, $a \wedge bc = 1$.*

Démonstration. D'après le théorème de Bézout, on peut fixer $(u, v, u', v') \in A^4$ tel que $au + bv = 1$ et $au' + cv' = 1$. On a donc $a(auu' + ucv' + u'bv) + bc(vv') = 1$. D'après le théorème de Bézout, cela signifie que $a \wedge bc = 1$. □

Remarque 6.11. On généralise sans problème par récurrence pour $n \geqslant 2$.

Proposition 6.42 (Lemme de Gauss)**.** *Soit A un anneau principal et $(a, b, c) \in A^3$. Si $a \mid bc$ et $a \wedge b = 1$ alors $a \mid c$.*

Proposition 6.43. *D'après le théorème de Bézout, on fixer $(u, v) \in A^2$ tel que $au + bv = 1$. On a donc $auc + bvc = c$ et a divise bc donc a divise c car divise le membre de gauche.*

Corollaire 6.6. *Soit A un anneau principal et $(a, b, c) \in A^3$. Si $a \wedge b = 1$, $a \mid c$ et $b \mid c$ alors $ab \mid c$.*

Démonstration. Comme $a \mid c$, il existe $q \in A$ tel que $c = aq$. Fixons-le. Par le lemme de Gauss, $b \mid q$ donc il existe $q' \in A$ tel que $q = bq'$. Donc $c = abq'$ puis $ab \mid c$. □

Dans un anneau principal quelconque, le calcul du PGCD n'est pas aisé. On s'intéresse ici à un cas particulier, celui où il existe une division euclidienne.

Définition 6.26 (Anneau euclidien, stathme)**.** On considère un anneau A commutatif et intègre. On dit que A est un **anneau euclidien** lorsqu'il existe une application

$$\nu \quad : \quad A \setminus \{0\} \quad \to \quad \mathbb{N}$$

vérifiant la propriété suivante :

$$\forall (a,b) \in A \times (A \setminus \{0\}),\ \exists (q,r) \in A^2,\ a = bq + r \text{ et } (r = 0 \text{ ou } \nu(r) < \nu(b))$$

On dit alors que ν est un **stathme** sur l'anneau euclidien A. On note (A, ν) un tel anneau euclidien, ou tout simplement A quand le stathme est fixé.

Remarque 6.12 (Étymologie)**.** Le nom "stathme" provient du grec "stathmos" signifiant arrêt.

Remarque 6.13. Attention, on ne demande pas l'unicité du couple (q, r) !

Exemple 6.11. L'anneau $\mathbb{Z}$ est euclidien avec le stathme $x \mapsto |x|$.

Remarque 6.14. Avec les mêmes notations, on dit que q est **un** quotient et r est **un** reste dans la division euclidienne de a par b.

Proposition 6.44 (Principalité des anneaux euclidiens)**.** *Tout anneau euclidien (A, ν) est principal.*

Démonstration. Soit (A, ν) un anneau euclidien. Soit I un idéal de A non réduit à $\{0\}$. Considérons l'ensemble

$$S := \{\nu(a) \mid a \in I \setminus \{0\}\}$$

Par hypothèse, puisque $I \setminus \{0\}$ est non vide, S est non vide. C'est aussi une partie de $\mathbb{N}$, donc elle admet un plus petit élément et il existe $a_0 \in I \setminus \{0\}$ tel que $\nu(a_0) = \min(S)$. Montrons alors que $I = (a_0)$. Par absorbance de I, on a immédiatement $(a_0) \subset I$. Réciproquement, soit $a \in I$. Effectuons la division euclidienne de a par a_0 : il existe $(q, r) \in A^2$ tel que $a = a_0 q + r$ et $r = 0$ ou $\nu(r) < \nu(a_0)$. Or, $r = a - a_0 q \in I$ car $a_0 q \in I$ d'après la première inclusion. Mais par minimalité de $\nu(a_0)$, on ne peut avoir $\nu(r) < \nu(a_0)$, donc nécessairement $r = 0$ puis $a = a_0 q \in (a_0)$. On a bien $I = (a_0)$ donc I est un idéal principal. Comme l'idéal nul est évidemment principal, l'anneau A est alors principal. □

Méthode 6.1 (Algorithme d'Euclide)**.** Soit (A, ν) un anneau euclidien. Puisque A est principal, pour $q \in A$, on a, à multiplication par un élément inversible près, $a \wedge b = b \wedge (a - bq)$. On utilise cette propriété avec q un quotient de a par b quand $b \neq 0$ et il vient avec les notations habituelles :

$$a \wedge b = b \wedge r$$

En itérant cette propriété, on construit une suite (r_k) d'éléments de A tels que $\nu(r_{k+1}) < \nu(r_k)$ tant que $r_{k+1} \neq 0$ en effectuant la division euclidienne de r_{k-1} par r_k après avoir posé $r_0 = a$ et $r_1 = b$. Comme il n'existe pas de suite strictement décroissante d'entiers naturels, cet algorithme se termine et le pgcd est un le dernier reste non nul.

4.4 PPCM dans un anneau principal

Définition 6.27 (PPCM)**.** Soit A un anneau principal et $(a,b) \in A^2$. L'idéal $(a) \cap (b)$ est principal et, par conséquent, il existe $\mu \in A$ tel que $(a) \cap (b) = (\mu)$. L'élément δ n'est pas unique en général, mais il est unique à un coefficient multiplicatif inversible près. On dit que μ est **un ppcm** de a et b et on note $\mu = a \vee b$.

Proposition 6.45. *Soit A un anneau principal et $(a,b) \in A^2$. On suppose que $(a,b) \neq (0,0)$. Dans ce cas, tout ppcm de a et b est non nul. Soit μ un tel ppcm. Toujours à un inversible près, μ est le "plus petit" (pour la relation $|$) multiple commun à a et b.*

Démonstration. Soit m un multiple commun à a et b. On a $(m) \subset (a) \cap (b)$ donc $(m) \subset (\mu)$ puis $\mu \mid m$. □

Proposition 6.46. *Soit A un anneau principal. A multiplication par un élément inversible près, on a :*

1. $\forall (a,b) \in A^2,\ a \vee b = b \vee a$ *(commutativité du ppcm)*
2. $\forall (a,b,c) \in A^3,\ (a \vee b) \vee c = a \vee (b \vee c)$ *(associativité du ppcm)*
3. $\forall (a,b,x) \in A^3,\ (xa) \vee (xb) = x(a \vee b)$

Démonstration. 1. Clair par commutativité de l'intersection ensembliste.

2. Clair par associativité de l'intersection ensembliste.
3. Le dernier point provient des égalités suivantes faciles à vérifier :

$$(xa) \cap (xb) = x\left[(a) \cap (b)\right] = x(a \vee b) = (x(a \vee b))$$

On peut aussi se ramener au cas où $(a,b) \neq (0,0)$, poser $\mu = a \vee b$ et montrer que $x\mu$ est un multiple commun à xa et xb (facile) et que c'est le "plus petit" pour $|$. Pour cela, on considère $m \in A$ tel que $xa \mid m$ et $xb \mid m$. On a alors $x \mid m$ par transitivité et il existe ù$' \in A$ tel ue $m = xm'$. Ainsi, $a \mid m'$ et $b \mid m'$ donc $\mu \mid m'$ et il existe $u \in A$ tel que $m' = \mu u$. Finalement, $m = x\mu u$.

□

Proposition 6.47. *Soit A un anneau principal et $(a,b) \in A^2$. On a, à multiplication par un élément inversible près :*

$$(a \wedge b)(a \vee b) = ab$$

Démonstration. Comme dans le cas de l'anneau $\mathbb{Z}$, on pose $\delta = a \wedge b$ et il existe alors $(a', b') \in A^2$ tel que $a = \delta a'$, $b = \delta b'$ et $a' \wedge b' = 1$. Il suffit alors de prouver que $a' \vee b' = a'b'$. On procède comme dans les démonstrations précédentes. On a clairement $a' \mid a'b'$ et $b' \mid a'b'$. On considère alors $m \in A$ tel que $a' \mid m$ et $b' \mid m$. Comme $a' \wedge b' = 1$, on a $a'b' \mid m$. On peut en conclure que $a'b'$ est bien le plus petit multiple commun de a' et b', ce qui achève la preuve. □

Remarque 6.15. En pratique, pour calculer un ppcm de deux éléments, on utilisera la relation $(a \wedge b)(a \vee b) = ab$ pour se ramener au calcul d'un pgcd de a et b.

5 Anneaux noethériens

Définition 6.28 (Anneau noethérien)**.** Soit A un anneau commutatif. Les trois propriétés suivantes sont équivalentes :

1. Tout idéal de A est engendré par un nombre fini d'élément.
2. Toute suite croissante d'idéaux de A $(I_n)_{n\in\mathbb{N}}$ est stationnaire à partir d'un certain rang.
3. Toute famille non vide d'idéaux admet un élément maximal pour l'inclusion.

Dans ce cas, on dit que l'anneau A est **noethérien**.

Démonstration. On démontre l'équivalence en faisant une circulante.

- $1 \implies 2$: Supposons que tout idéal de A est engendré par un nombre fini d'éléments. Soit $(I_n)_{n\in\mathbb{N}}$ une suite croissante d'idéaux. Alors $I := \bigcup_{n\in\mathbb{N}} I_n$ est encore un idéal de A. En effet, $0 \in I_0$ donc $0 \in I$ et soit $(a,b) \in I^2$. Alors il existe $n \in \mathbb{N}$ tel que $(a,b) \in I_n^2$ par croissance de la suite. Puis $a+b \in I_n$ donc $a+b \in I$. Enfin, soit $a \in I$ et $b \in A$. Il existe $n \in \mathbb{N}$ tel que $a \in I_n$ puis $ab \in I_n$ par absorbance de I_n donc $ab \in I$. Ensuite, I est engendré par un nombre fini d'éléments $a_1, \dots, a_m$. On peut alors trouver un $n_0 \in \mathbb{N}$ tel que $a_1, \dots, a_m$ soient des éléments de I_{n_0}. On a immédiatement $I_{n_0} \subset I$ par définition de I, et réciproquement, I_{n_0} étant un idéal contenant tous les a_i, il est contenu par I car I est engendré par les a_i. Donc $I = I_{n_0}$. Soit $n \geqslant n_0$. On a $I_{n_0} \subset I_n$ par croissance de la suite. Or, $I_{n_0} = I$ donc par définition de I, $I_n \subset I_{n_0}$. On a donc $I_n = I_{n_0}$ et la suite $(I_n)_{n\in\mathbb{N}}$ est stationnaire.
- $2 \implies 3$: On raisonne par contraposée. Supposons qu'il existe une famille $(I_s)_{s\in S}$ non vide d'idéaux de A qui n'admette pas d'élément maximal pour l'inclusion. On va construire par récurrence une suite $(J_n)_{n\in\mathbb{N}}$ d'idéaux de A strictement croissante. Pour l'initialisation, on choisit $s_0 \in S$ car S est non vide, et on pose $J_0 := I_{s_0}$. Désormais, supposons avoir construit J_n pour $n \in \mathbb{N}$, tel que $\forall k \in [\![0,n]\!]$, $\exists s \in S,\ J_k = I_s$ et telle que $\forall k \in [\![0, n-1]\!],\ J_k \subsetneq J_{k+1}$. Puisque $(I_s)_{s\in S}$ ne possède pas d'élément maximal $\exists s \in S,\ J_n \subsetneq I_s$. Fixons ce s et posons $J_{n+1} := I_s$. Ainsi, on a construit une suite $(J_n)_{n\in\mathbb{N}}$ d'idéaux de A strictement croissante, donc il existe une suite d'idéaux croissante et non stationnaire.
- $3 \implies 1$: Supposons que tout famille non vide d'idéaux admette un élément maximal pour l'inclusion. Soit I un idéal de A et considérons l'ensemble des idéaux de I engendrés par un nombre fini d'éléments et qui sont inclus dans I. Cet ensemble est non vide car il contient l'idéal nul. Par hypothèse, il possède alors un élément maximal J. On a alors $J \subset I$. Soit $x \in I$. Alors $J + xA$ est un idéal, comme somme de deux idéaux, et il engendré par un nombre fini d'éléments puisqu'il est engendré par le nombre fini d'éléments qui engendrent J et x. Or, $xA \subset I$ donc par stabilité de I par $+$, $J + xA \subset I$. Or, $J \subset J + xA$, donc, par maximalité de J, $J = J + xA$. Or, $x = 0 + x \times 1 \in J + xA$ donc $x \in J$. Ainsi, $J = I$, donc I est bien engendré par un nombre fini d'éléments.

Ainsi, la preuve est achevée. □

Proposition 6.48. *Tout anneau principal est noethérien.*

Démonstration. Tous les idéaux d'un anneau principal sont engendrés par un seul élément, donc en particulier par un nombre fini d'éléments. Tout anneau principal vérifie donc la condition 1, et est alors noethérien. □

6 Anneaux factoriels

6.1 Définition et première propriétés

Définition 6.29 (Élément irréductible)**.** Soit A un anneau intègre. Un élément $a \in A$ est dit **irréductible** s'il est non nul, non inversible et si ses seuls diviseurs sont les diviseurs triviaux u et ua pour $u \in A^\times$. Ce dernier point revient à dire qu'il vérifie la propriété suivante :

$$\forall (u,v) \in A^2,\ (a = uv) \implies (u \in A^\times \text{ ou } v \in A^\times)$$

Exemple 6.12. Les éléments irréductibles et positifs de $\mathbb{Z}$ sont les nombres premiers.

Exemple 6.13. Si $\mathbb{K}$ est un corps, les éléments inversibles de l'anneau $\mathbb{K}[X]$ sont les polynômes constants non nuls. Par suite, un polynôme est irréductible si et seulement si ses seuls diviseurs sont les constantes non nulles et les polynômes qui lui sont associés.

Définition 6.30 (Anneau factoriel)**.** Un anneau A est dit **factoriel** lorsqu'il est intègre et et que tout élément non nul admet une décomposition en produit de facteurs irréductibles et si cette décomposition est unique à l'ordre près et au produit par un inversible près. Formellement, cela signifie que :

1. Pour tout $a \in A \setminus \{0\}$, il existe $u \in A^\times$, des éléments irréductibles $p_1, \dots, p_k$ de A et des entiers naturels non nuls $\alpha_1, \dots, \alpha_k$ tels que :
$$a = u p_1^{\alpha_1} \dots p_k^{\alpha_k} = u \prod_{i=1}^{k} p_i^{\alpha_i}$$
2. Il y a unicité de la décomposition en ce sens : si
$$a = u p_1^{\alpha_1} \dots p_k^{\alpha_k} = v q_1^{\beta_1} \dots q_m^{\alpha_m}$$
avec $\alpha_1, \dots, \alpha_k, \beta_1, \dots, \beta_m \in \mathbb{N}^\star$ et où les $p_1, \dots, p_k$ d'une part, et les $q_1, \dots, q_m$ d'autre part, ne sont pas associés, alors $k = m$ et il existe une permutation σ de $[\![1, k$ et des éléments inversibles u_i tels que $q_{\sigma(i)} = u_i p_i$ et $\beta_{\sigma(i)} = \alpha_i$ pour tout $i \in [\![1, k]\!]$.

On notera (1) la condition d'existence et (2) la condition d'unicité.

Proposition 6.49. *Sous réserve que la condition* (1) *soit vérifiée, alors la condition* (2) *équivaut à cette nouvelle condition, notée* (2′) *: tout élément irréductible a de A vérifie*

$$\forall (b,c) \in A^2,\ a \mid bc \implies a \mid b \ \ ou \ \ a \mid c$$

Démonstration. Supposons que A est intègre et vérifie la condition (1).

- (2) $\implies$ (2′) : Supposons la condition (2) vérifiée par A. Si a est irréductible et divise bc, on écrit $bc = ad$, puis les décompositions de b, c et d en produits de facteurs irréductibles. L'unicité de la décomposition montre alors que a figure dans la décomposition de bc, donc obligatoirement dans celle de b ou dans celle de c.
- (2′) $\implies$ (2) : Supposons la condition (2′) vérifiée par A. On raisonne par récurrence sur $s \in \mathbb{N}^\star$ où s sera la somme des exposants figurant dans une des décompositions de a. L'hypothèse de récurrence $H(s)$ est donc la suivante :

"Si $a \in A\backslash\{0\}$ vérifie $a = up_1^{\alpha_1} \dots p_k^{\alpha_k} = vq_1^{\beta_1} \dots q_m^{\alpha_m}$ avec $\alpha_1, \dots, \alpha_k, \beta_1, \dots, \beta_m \in \mathbb{N}^\star$ et où les $p_1, \dots, p_k$ d'une part, et les $q_1, \dots, q_m$ d'autre part, ne sont pas associés, et $\alpha_1 + \dots + \alpha_k = s$, alors $k = m$ et il existe une permutation σ de $[\![1, k$ et des éléments inversibles u_i tels que $q_{\sigma(i)} = u_i p_i$ et $\beta_{\sigma(i)} = \alpha_i$ pour tout $i \in [\![1, k]\!]$."

Initialisation : si $a = up_1 = vq_1^{\beta_1} \dots q_m^{\alpha_m}$ alors q_i divise p_1 pour tout i en utilisant $(2')$ et le fait que q_i ne divise pas u puisque u est inversible (donc $u \mid q_i$) et que q_i n'est pas inversible. Cela impose à q_i d'âtre associé à p_1. Comme les q_i ne sont pas associés entre eux, on aura $m = 1$, q_1 associé à p_1 et $\beta_1 = 1$. Ainsi, $H(1)$ est vraie.

Hérédité : Supposons que $H(s)$ soit vraie pour un certain $s \in \mathbb{N}^\star$ et montrons la propriété pour $(s+1)$. On considérer alors les décompositions

$$a = up_1^{\alpha_1} \dots p_k^{\alpha_k} = vq_1^{\beta_1} \dots q_m^{\alpha_m}$$

avec $\alpha_1 + \dots + \alpha_k = s+1$. On a q_m irréductible et q_m divise le produit des p_i et de u mais q_m ne divise pas u car u est inversible (donc $u \mid q_m$) et q_m n'est pas inversible. Ainsi, q_m divise l'un des p_i d'après $(2')$. Quitte à permuter les p_i, on peut supposer que q_m divise p_k. Comme p_k est irréductible, il existe $w \in A^\times$ tel que $q_m = wp_k$ et l'on obtient, après simplification car l'anneau A est intègre :

$$a = up_1^{\alpha_1} \dots p_k^{\alpha_k - 1} = vq_1^{\beta_1} \dots q_m^{\alpha_m - 1}$$

Il suffit alors d'appliquer l'hypothèse de récurrence à cette décomposition en distinguant deux cas pour s'assurer que les exposants qui interviennent son vraiment tous strictement positifs : si $\alpha_k = 1$, alors $\beta_m = 1$ autrement q_m (et donc p_k) diviserait l'un des p_i avec $i \neq k$, ce qui est absurde. Si $\alpha_k > 1$, alors $\beta_k > 1$ autrement p_k diviserait l'un des q_i avec $i \neq m$. La propriété $H(s+1)$ est alors démontrée.

La récurrence est alors achevée.

Ainsi, la preuve est achevée. □

Proposition 6.50. *Soit A un anneau factoriel. $a \in A$ est irréductible si, et seulement si, il est non nul, non inversible et vérifie la propriété $a \mid bc \implies a \mid b$ ou $a \mid c$.*

Proposition 6.51. *Le sens direct a déjà été prouvé lors de la démonstration précédente. Réciproquement, si a vérifie cette propriété, et si $a = uv$, alors $a \mid uv$ donc $a \mid u$ ou $a \mid v$. Si par exemple $u = aa'$ alors par intégrité $1 = a'v$ et v sera inversible. Idem si $v = aa'$.*

6.2 Anneaux atomiques

On va donner ici la définition d'un type d'anneau intègre qui vérifie la condition (1) et montrer une propriété sur ce type d'anneau qui servira pour la partie suivante.

Définition 6.31 (Anneau atomique)**.** Un anneau A est dit **atomique** lorsqu'il est intègre et qu'il vérifie la condition (1), c'est-à-dire l'existence d'une décomposition en produit d'irréductibles pour tout élément non nul.

Exemple 6.14. Un anneau factoriel est nécessairement atomique.

Proposition 6.52 (Atomicité des anneaux noethériens intègres)**.** *Tout anneau intègre et noethérien est atomique.*

Démonstration. Soit A un anneau intègre et noethérien. On raisonne par l'absurde en supposant l'existence d'un élément non nul a de A sans décomposition en produit de facteurs irréductibles. Alors a n'est pas irréductibles donc il existe a_1 et b_1 deux éléments de A non inversibles tels que $a = a_1 b_1$. Au moins l'un des deux facteurs, disons a_1, ne se décompose pas en produit de facteurs irréductibles, sans quoi en réunissant les deux décompositions a se décomposerait en produit de facteurs irréductibles. Il existe alors a_2 et b_2 des éléments de A non inversibles tels que $a_1 = a_2 b_2$. Par exemple, a_2 ne se décompose pas en produit de facteurs irréductibles, et l'on continue de la même manière par récurrence. On obtient alors une suite infinie strictement croissante d'idéaux

$$(a) \subsetneq (a_1) \subsetneq (a_2) \subsetneq \ldots \subsetneq (a_n) \subsetneq \ldots$$

en contradiction avec le caractère noethérien de A. On notera que la suite d'idéaux est bien strictement croissante puisque $(a_i) = (a_{i+1})$ entraîne l'existence de q et q' des éléments de A tels que $a_i = qa_{i+1}$ et $a_{i+1} = q'a_i$ donc par intégrité $1 = qq'$ et q' est inversible, puis, en remplaçant dans $a_i = a_{i+1}b_{i+1}$, on trouve par intégrité $1 = ub_{i+1}$ ce qui est en contradiction avec le fait que $b_{i+1} \notin A^\times$. Ainsi, A est bien atomique. □

6.3 Factorialité des anneaux principaux

On arrive finalement à la proposition fondamentale :

Proposition 6.53 (Factorialité des anneaux principaux)**.** *Tout anneau principal est factoriel.*

Démonstration. Soit A un anneau principal et montrons que A vérifie les conditions (1) et (2′).

- (1) : A est principal donc A est noethérien. Or, A est principal donc A est intègre. Comme A est intègre et noethérien, il est atomique d'après le paragraphe qui précède et alors il vérifie la condition (1).
- (2′) : Soit a un élément irréductible de l'anneau A et $(b, c) \in A^2$ tel que $a \mid bc$. Si $a \mid b$,a lors c'est bon. Si a ne divise pas b, alors puisque a est irréductible, a est premier avec b et d'après le lemme de Gauss il divise c et la condition (2′) est vérifiée.
- Comme (1) est vérifiée, alors (2′) est équivalente à (2), et puisque (2′) est vérifiée, alors (2) est vérifiée. Par conséquent, A est bien factoriel.

Ainsi, la preuve est achevée. □

Remarque 6.16. Notons l'importance capitale de ce résultat. Il prouve d'un seul coup les théorèmes de décomposition en facteurs premiers dans $\mathbb{Z}$ et de décomposition en facteurs irréductibles dans $\mathbb{K}[X]$. Plus généralement, puisqu'on a vu qu'il suffisait d'être un anneau euclidien pour être un anneau principal, donc le simple fait de démontrer l'existence (et même pas l'unicité !) d'une division euclidienne sur un anneau garantit l'existence et l'unicité d'une décomposition en produit de facteurs irréductibles. Nous aurions donc pu nous arrêter aux théorèmes de division euclidienne dans $\mathbb{Z}$ et $\mathbb{K}[X]$ (lorsque $\mathbb{K}$ est un corps) plutôt que de démontrer tout un tas d'autres propriétés pour garantir l'existence et l'unicité de la décomposition en produit d'irréductibles.

6.4 PGCD et PPCM dans un anneau factoriel

Dans tout le paragraphe, A est un anneau factoriel fixé.

Soit a et b deux éléments de A. Il existe $p_1, \dots, p_k$ des éléments irréductibles de A deux à deux non associés et des éléments inversibles u et v tels que

$$a = up_1^{\alpha_1} \dots p_k^{\alpha_k} \quad \text{et} \quad b = vp_1^{\beta_1} \dots p_k^{\beta_k}$$

avec $\alpha_1, \dots, \alpha_k, \beta_1, \dots \beta_k \in \mathbb{N}$. En effet, on regroupe les irréductibles associés et on autorise les exposants nuls de façon à avoir les mêmes irréductibles pour a et b.

Dans tout le paragraphe, on conserve ces décompositions.

Proposition 6.54 (Critère de divisibilité)**.** *Avec les mêmes notations :*

$$a \mid b \iff \forall i \in [\![1, k]\!],\ \alpha_i \leqslant \beta_i$$

Démonstration. Pour le sens direct, il suffit de remarquer que la décomposition de b "contient" alors celle de a, ce qui force les inégalités. Réciproquement, si on a les inégalités, il suffit de "compléter" les exposants dans la décomposition de a pour obtenir un élément c tel que $ac = b$. □

Définition 6.32 (Expression du PGCD et du PPCM)**.** Avec les mêmes notation, et toujours à un inversible près, on a :

$$a \wedge b = \prod_{i=1}^{k} p_i^{\gamma_i}$$

et

$$a \vee b = \prod_{i=1}^{k} p_i^{\lambda_i}$$

avec pour tout $i \in [\![1, k]\!]$, $\gamma_i := \min(\alpha_i, \beta_i)$ et $\lambda_i := \max(\alpha_i, \beta_i)$.

Démonstration. Il suffit d'utiliser le critère de divisibilité et de montrer que les nombres donnés ici sont des multiples communs ou diviseurs communs de a et b et le critère de divisibilité montre qu'ils sont les "plus petit" et "plus grand" pour la divisibilité, toujours à un inversible près. □

Remarque 6.17. Là encore, cette proposition effectue d'un seul coup ce que nous avons longuement prouvé dans $\mathbb{Z}$ et dans $\mathbb{K}[X]$. Remarquons que dans ces anneaux, on impose toutefois une condition supplémentaire pour avoir l'unicité du PGCD et du PPCM : dans $\mathbb{Z}$ on le prend positif car les unités de $\mathbb{Z}$ sont -1 et 1, et dans $\mathbb{K}[X]$ on le prend unitaire car les unités de $\mathbb{K}[X]$ sont les polynômes constants non nuls.

7 Corps des fractions (cas commutatif)

Dans le deuxième paragraphe qui traite des anneaux quotients, nous avons d'une certaine façon "forcé" des éléments d'un anneau à être nuls. Nous voulons maintenant effectuer une opération "opposée" : rendre inversibles les éléments d'un anneau convenable.

Dans cette partie, l'anneau A est commutatif et intègre.

Proposition 6.55. *On considère la relation $\sim$ définie sur $A \times (A \setminus \{0\})$ par :*

$$\forall (a,b,a',b') \in (A \times (A \setminus \{0\}))^2,\ (a,b) \sim (a',b') \iff ab' = a'b$$

La relation binaire $\sim$ est une relation d'équivalence.

Démonstration. Les vérifications sont immédiates. □

Définition 6.33 (Notation $frac(A)$)**.** On note alors $frac(A) = (A \times (A \setminus \{0\}))/\sim$.

Définition 6.34 (Addition et multiplication sur $frac(A)$)**.** On définit une multiplication et une addition sur $frac(A)$ par :

$$\begin{array}{rccc} \tilde{+} & : & frac(A) & \to & frac(A) \\ & & \left(\overline{(a,b)}, \overline{(a',b')}\right) & \mapsto & \overline{(ab'+a'b, bb')} \end{array}$$

et

$$\begin{array}{rccc} \tilde{\times} & : & frac(A) & \to & frac(A) \\ & & \left(\overline{(a,b)}, \overline{(a',b')}\right) & \mapsto & \overline{(aa', bb')} \end{array}$$

Ainsi, ce que l'on veut imiter, c'est tout simplement le calcul de fractions que l'on apprend au collège.

Proposition 6.56. *Les opérations ci-dessus sont bien définies sur $frac(A)$.*

Démonstration. Soient (a_1,b_1), (a_1',b_1'), (a_2,b_2) et (a_2',b_2') des éléments de $A \times (A \setminus \{0\})$ tels que $(a_1,b_1) \sim (a_1',b_1')$ et $(a_2,b_2) \sim (a_2',b_2')$. Dans ce cas, on a

$$(a_1b_2 + a_2b_1, b_1b_2) \sim (a_1'b_2' + a_2'b_1', b_1'b_2')$$

car

$$\begin{aligned} (a_1b_2 + a_2b_1)b_1'b_2' &= a_1b_1'b_2b_2' + a_2b_2'b_1b_1' \\ &= a_1'b_1b_2b_2' + a_2'b_2b_1b_1' \\ &= (a_1'b_2' + a_2'b_1')b_1b_2 \end{aligned}$$

Ceci montre que l'application $\tilde{+}$ est bien définie car deux classes d'équivalences ont la même images, et aussi car par intégrité de A bb' est non nul lorsque b et b' sont non nuls. De plus, on a :

$$(a_1a_2, b_1b_2) \sim (a_1'a_2', b_1'b_2')$$

car

$$\begin{aligned} a_1a_2b_1'b_2' &= (a_1b_1')(a_2b_2') \\ &= (a_1'b_1)(a_2'b_2) \\ &= a_1'a_2'b_1b_2 \end{aligned}$$

ce qui prouve que $\tilde{\times}$ est bien définie. □

Proposition 6.57. *Muni des deux opérations définies ci-dessus, $frac(A)$ est un corps commutatif. L'élément neutre puir $\tilde{+}$ est $\overline{(0,1)}$ et l'élément neutre pour $\tilde{\times}$ est $\overline{(1,1)}$.*

Démonstration. Vérifions les différents axiomes.

- L'addition est associative. En effet, on a

$$\begin{aligned}\left(\overline{(a_1,b_1)}\tilde{+}\overline{(a_2,b_2)}\right)\tilde{+}\overline{(a_3,b_3)} &= \overline{(a_1b_2+a_2b_1,b_1b_1)}\tilde{+}\overline{(a_3,b_3)}\\ &= \overline{((a_1b_2+a_2b_1)b_3+a_3(b_1b_2),(b_1b_2)b_3))}\\ &= \overline{(a_1b_2b_3+a_2b_1b_3+a_3b_1b_2,b_1b_2b_3))}\\ &= \overline{(a_1,b_1)}\tilde{+}\overline{(a_2b_3+a_3b_2,b_2b_3)}\\ &= \overline{(a_1,b_1)}\tilde{+}\left(\overline{(a_2,b_2)}\tilde{+}\overline{(a_3,b_3)}\right)\end{aligned}$$

- L'addition est clairement commutative par commutativité de $+$ et $\times$ sur A.
- L'élément neutre pour $\tilde{+}$ est $\overline{(0,1)}$ et l'opposé de $\overline{(a,b)}$ est $\overline{(-a,b)}$ (vérifications immédiates).
- La multiplication est clairement associative et commutative par associativité et commutativité de $\times$ sur A.
- L'élément neutre pour $\tilde{\times}$ est $\overline{(1,1)}$, cela s'hérite immédiatement du fait que 1 est le neutre pour $\times$ dans A.
- Le produit est distributif sur la somme. En effet, on a

$$\begin{aligned}\left(\overline{(a_1,b_1)}\tilde{\times}\overline{(a_3,b_3)}\right)\tilde{+}\left(\overline{(a_2,b_2)}\tilde{\times}\overline{(a_3,b_3)}\right) &= \overline{(a_1a_3,b_1b_3)}\tilde{+}\overline{(a_2a_3,b_2b_3)}\\ &= \overline{((a_1b_2+a_2b_1)a_3b_3,b_1b_2b_3^2)}\\ &= \overline{((a_1b_2+a_2b_1)a_3,b_1b_2b_3)}\\ &= \overline{(a_1b_2+a_2b_1,b_1b_2)}\tilde{\times}\overline{(a_3,b_3)}\\ &= \left(\overline{(a_1,b_1)}\tilde{+}\overline{a_2,b_2}\right)\tilde{\times}\overline{(a_3,b_3)}\end{aligned}$$

- Enfin, pour $\overline{(a,b)}\neq 0$, on a $a\neq 0$. En effet, les éléments de classe de 0 sont tous les éléments de $frac(a)$ de la forme $(0,c)$, avec $c\neq 0$. On peut donc conclure que $\overline{(a,b)}$ est inversible dans $frac(A)$ puisque son inverse n'est autre que $\overline{(b,a)}$ qui a un sens puisque $a\neq 0$.

Ainsi, la preuve est achevée. □

Proposition 6.58 (Inclusion abusive $A\subset frac(A)$)**.** *L'application*

$$\begin{array}{rccc} i & : & A & \longrightarrow & frac(A)\\ & & a & \longmapsto & \overline{(a,1)}\end{array}$$

est un morphisme injectif d'anneaux. Ainsi, l'anneau A est isomorphe à l'anneau $i(A)$ qui est un sous-anneau de $frac(A)$. Grâce à cet isomorphisme, tout calcul dans A peut se voir comme un calcul dans A peut se voir comme un calcul dans $i(A)$ et c'est pourquoi on considère que A est inclus dans $frac(A)$ alors que, rigoureusement, ce n'est pas vraiment une inclusion mais seulement un isomorphisme entre A et un sous-anneau de $frac(A)$.

Démonstration. On vérifie facilement que l'application i est un morphisme d'anneaux puisque les applications $\tilde{+}$ et $\tilde{\times}$ sont bien définies et que le neutre pour $\tilde{\times}$ est $\overline{(1,1)}$. L'injectivité provient directement de l'étude du noyau et de l'intégrité de A. □

Remarque 6.18. Cela revient aussi à considérer qu'un anneau n'a d'importance ou n'existe vraiment qu'à un isomorphisme près.

Définition 6.35 (Notations finales)**.** Grâce à cette inclusion abusive, on note désormais $+$ et $\times$ pour $\tilde{+}$ et $\tilde{\times}$ puisque ces dernières opérations prolongent $+$ et $\times$. Souvent, quand il n'y a pas d'ambiguïté, on notera plutôt $\dfrac{a}{b}$ l'élément $\overline{(a,b)}$.

Remarque 6.19. Le corps des fractions d'un corps est lui-même.

Exemple 6.15. Le corps des fractions de l'anneau intègre $\mathbb{Z}$ est le corps $\mathbb{Q}$.

Exemple 6.16. Lorsque $\mathbb{K}$ est un corps, l'anneau des polynômes à coefficients dans $\mathbb{K}$ $\mathbb{K}[X]$ est intègre et son corps des fractions, noté $\mathbb{K}(X)$ est le corps des fractions rationnelles à coefficients dans $\mathbb{K}$.

Remarque 6.20 (Généralisation : localisation d'un anneau)**.** Il existe une construction encore plus générale à l'aide d'une relation d'équivalence similaire qui consiste à rendre toute partie S de A stable par multiplication et contenant 1 inversible de force. On parle de **localisation** de la partie S de l'anneau A. Tout fonction de même, les vérifications sont plus pénibles car la relation d'équivalence est alors plus complexes. En revanche, le morphisme i n'est plus nécessairement injectif si l'anneau A n'est pas intègre, donc l'inclusion abusive n'a plus lieu d'être. Lorsqu'on localise $A \setminus \{0\}$ pour un anneau intègre (cette partie étant alors stable par multiplication et contenant 1), on obtient le corps de fractions de A, ce qui prouve bien que la méthode de localisation est encore plus générale.

8 Exercices

Note : Je n'ai pas pris le temps de tenter de résoudre un seul de ces exercices, donc je ne connais pas les méthodes pour les résoudre. Si toutefois vous trouvez une solution, je serai ravi que vous me la partagiez.

Exercice 6.1. Soit $n \geqslant 2$ un entier. Déterminer les éléments nilpotents et les éléments inversibles de $\mathbb{Z}/n\mathbb{Z}$.

Exercice 6.2. Soit n et m des entiers non nuls. Montrer que l'application canonique de $\mathbb{Z}/nm\mathbb{Z}$ dans $\mathbb{Z}/n\mathbb{Z}$ est un morphisme d'anneaux. Montrer qu'il induit une surjection de $(\mathbb{Z}/nm\mathbb{Z})^\times$ sur $(Z/n\mathbb{Z})^\times$.

Exercice 6.3. Soit n et m des entiers non nuls. A quelle condition peut-on trouver un morphisme d'anneaux de $\mathbb{Z}/n\mathbb{Z}$ dans $\mathbb{Z}/m\mathbb{Z}$?

Exercice 6.4. Exhiber un morphisme d'anneaux $f : A \to B$ qui soit surjectif mais tel que le morphisme de groupes de $A^\times$ dans $B^\times$ déduit par restriction de f ne le soit pas.

Exercice 6.5. Soit $\mathbb{K}$ un corps et A un anneau non trivial. Montrer que tout morphisme d'anneaux de $\mathbb{K}$ dans A est injectif.

Exercice 6.6. Quel est le radical de l'idéal (12) dans $\mathbb{Z}$? (Oui, c'est un poly de Taupins au départ, je le rappelle...)

Exercice 6.7 (Généralisation du précédent)**.** Soit n un entier strictement positif qu'on décompose en facteurs premiers $n = \prod_{i=1}^{k} p_i^{\alpha_i}$ avec les α_i des entiers strictement positifs et les p_i des nombre premiers deux à deux distincts. On pose $m := \prod_{i=1}^{k} p_i$. Montrer que $\sqrt{n\mathbb{Z}} = m\mathbb{Z}$.

Exercice 6.8. Soit A un anneau commutatif et $(a, b) \in A^2$. Montrer que si a et b sont associés, alors les idéaux (a) et (b) sont égaux. Réciproquement, si A est intègre et si $(a) = (b)$, montrer que a et b sont associés. Et si A n'est pas intègre ?

Exercice 6.9. Montrer qu'un anneau intègre ne possédant qu'un nombre fini d'idéaux est un corps. On pourra montrer que tout élément non nul x est inversible en introduisant les idéaux $x^n A$ pour $n \in \mathbb{N}^\star$ et utiliser le principe des tiroirs.

Exercice 6.10. Soit I et J deux idéaux d'un anneau commutatif. Montrer que si I et J sont comaximaux (*ie* si $I + J = A$), alors pour tout $n \in \mathbb{N}^\star$, I^n et J^n (pas au sens du produit cartésien mais au sens de l'idéal produit) sont comaximaux (*ie* $I^n + J^n = A$).

Exercice 6.11. Soit A un anneau non nécessairement commutatif.

1. Montrer par un contre-exemple que l'ensemble des éléments nilpotents de A ne forme par un sous-groupe abélien en général. On pourra choisir $A = \mathcal{M}_2(\mathbb{C})$.
2. Soit N l'ensemble des éléments $a \in A$ tel que ax soit nilpotent pour tout $x \in A$. Montrer que N est un idéal de A dont tout élément est nilpotent.
3. Soit I un idéal de A dont tout élément est nilpotent. Montrer que $I \subset N$.

Exercice 6.12. Soit A un anneau non nécessairement commutatif et soit I l'idéal engendré par les $xy - yx$ pour $(x, y) \in A^2$.

1. Montrer que l'anneau A/I est commutatif.
2. Soit J un idéal de A tel que A/J soit commutatif. Montrer que $I \subset J$.

Exercice 6.13. On s'intéresse au polynôme $X^2 + 1$.

1. Soit $\mathbb{K}$ un corps commutatif et $P \in \mathbb{K}[X]$ un polynôme à coefficients dans $\mathbb{K}$. Montrer que l'anneau $\mathbb{K}[X]/(P)$ est un corps si, et seulement si, P est irréductible dans $\mathbb{K}[X]$.
2. Montrer que le polynôme X^2+1 est irréductible dans l'anneau $\mathbb{Z}[X]$. L'anneau $A := \mathbb{Z}[X]/(X^2+1)$ est-il un corps ? On pourra définir un isomorphisme de A sur l'anneau $\mathbb{Z}[i]$.
3. Soit p un nombre premier. Montrer que le polynôme $X^2 + 1$ est irréductible dans $(\mathbb{Z}/p\mathbb{Z})[X]$ si, et seulement si, $p \equiv 3$ [4].

Exercice 6.14. Soit A un anneau et I un idéal de A. On note $I[X]$ l'ensemble des polynômes $P \in A[X]$ dont tous les coefficients appartiennent à I.

1. Montrer que si I est un idéal bilatère de A, alors $I[X]$ est un idéal bilatère de $A[X]$.
2. Construir un isomorphisme de l'anneau $A[X]/I[X]$ sur l'anneau $(A/I)[X]$.

Exercice 6.15. Soit A un anneau et I un idéal de A. On note $\mathcal{M}_n(I)$ l'ensemble des matrices de $\mathcal{M}_n(A)$ dont tous les coefficients appartiennent à I.

1. Montrer que $\mathcal{M}_n(I)$ est un idéal de $\mathcal{M}_n(A)$ et construire un isomorphisme d'anneaux de $\mathcal{M}_n(A)/\mathcal{M}_n(I)$ sur $\mathcal{M}_n(A/I)$.
2. Inversement, montrer que tout idéal de $\mathcal{M}_n(A)$ est de la forme $\mathcal{M}_n(I)$, pour I un idéal bilatère de A.

Chapitre 7

Réels et suites

La secte des Pythagoriciens possédait un joyau dont elle ne faisait don à personne d'extérieur : la démonstration du théorème de Pythagore. Ce théorème impliquait notamment que la diagonale d'un carré de côté 1 valait $\sqrt{2}$. Lorsque les Pythagoriciens, pour qui "tout était nombre" (au sens de "tout est quotient de nombres entiers"), découvrirent que $\sqrt{2}$ était irrationnel, ils furent catastrophés : il était possible de faire coexister géométriquement des grandeurs "incommensurables" numériquement ! Finalement, cette crise des fondements ne fut résolue de manière satisfaisante que beaucoup plus tard, avec les premières constructions rigoureuses de $\mathbb{R}$ proposées au XIX^e siècle.

C'est une telle construction que nous proposons à la fin de ce chapitre, sans doute une des plus classiques. Comme elle n'est pas exigible au programme, on pourra admettre le résultat fondamental qui en découle, à savoir la propriété de la borne supérieure. C'est sur cette propriété fondatrice que s'appuie l'essentiel de l'analyse réelle.

1 Propriétés fondatrices de $\mathbb{R}$

Exposons sans démonstration les principales propriétés qui proviennent directement de la construction de $\mathbb{R}$.

Théorème 7.1 (Existence de $\mathbb{R}$ - admis)**.** *On peut construire $(\mathbb{R}, +, \times)$ un sur-corps de $\mathbb{Q}$, totalement ordonné par $\leqslant$. La relation $\leqslant$ est compatible avec l'addition et la multiplication au même sens que dans $\mathbb{Z}$ et $\mathbb{Q}$:*

$$\forall (x,y,z) \in \mathbb{R}^3,\ x \leqslant y \implies x+z \leqslant y+z$$

$$\forall (x,y,z) \in \mathbb{R}^3,\ \left\{ \begin{array}{l} x \leqslant y \\ z \geqslant 0 \end{array} \right. \implies xz \leqslant yz$$

Définition 7.1. Soit $(a,b) \in \mathbb{R}^2$. A partir de $\leqslant$, on définit de la manière habituelle les ensembles $[a,b]$, $[a,b[$, $]a,b]$ et $]a,b[$. Si $a > b$, ces ensembles sont tous vides.

$$\left\{ \begin{array}{l} [a,b] = \{t \in \mathbb{R} \mid a \leqslant t \leqslant b\} \\ [a,b[= \{t \in \mathbb{R} \mid a \leqslant t < b\} \\ ]a,b] = \{t \in \mathbb{R} \mid a < t \leqslant b\} \\ ]a,b[= \{t \in \mathbb{R} \mid a < t < b\} \end{array} \right.$$

Corollaire 7.1 (Variantes)**.** *Au niveau de l'addition, on a la variante suivante :*

$$\forall (x,y,z) \in \mathbb{R}^3,\ x \leqslant y \implies x+z \leqslant y+z$$

Au niveau de la multiplication, on a les trois variantes suivantes :

$$\forall (x,y,z) \in \mathbb{R}^3,\ \left\{ \begin{array}{l} x \leqslant y \\ z \leqslant 0 \end{array} \right. \implies xz \geqslant yz$$

$$\forall (x,y,z) \in \mathbb{R}^3,\ \left\{ \begin{array}{l} x < y \\ z > 0 \end{array} \right. \implies xz < yz$$

$$\forall (x,y,z) \in \mathbb{R}^3,\ \left\{ \begin{array}{l} x < y \\ z < 0 \end{array} \right. \implies xz > yz$$

Remarque 7.1. On retiendra que dans une inéquation, on peut faire "changer de côté" ce que l'on veut.

1. Dans une somme, à condition que les $+$ deviennent des $-$ et vice-versa.
2. Dans un produit, à condition que le terme soit non nul, que les $\times$ deviennent des $/$ et vice-versa. Si le facteur est strictement positif le sens de l'inéquation ne change pas, sinon il est inversé.

Remarque 7.2 (Passage à l'inverse dans une inéquation)**.** La fonction $t \mapsto \dfrac{1}{t}$ envoie $\mathbb{R}_+^\star$ dans lui-même et est strictement décroissante sur ce domaine. même remarque avec $\mathbb{R}_-^\star$.

Théorème 7.2. *Soit $x_1, \ldots, x_n \geqslant 0$. Alors $\displaystyle\sum_{i=1}^{n} x_i \geqslant 0$, avec égalité si, et seulement si, tous les x_i sont nuls.*

Corollaire 7.2. *Soit $a_1, \ldots, a_n$ et $b_1, \ldots, b_n$ tels que $\forall i \in [\![1,n]\!],\ a_i \leqslant b_i$. Alors*

$$\sum_{i=1}^{n} a_i \leqslant \sum_{i=1}^{n} b_i$$

avec égalité si, et seulement si, $\forall i \in [\![1,n]\!],\ a_i = b_i$.

Définition 7.2 (Borne supérieure)**.** Soit A une partie non vide de $\mathbb{R}$. Si elle existe, la **borne supérieure de** A est le plus petit des majorants de A. On la note $\sup(A)$.

Définition 7.3 (Borne inférieure)**.** Soit A une partie non vide de $\mathbb{R}$. Si elle existe, la **borne inférieure de** A est le plus grand des minorants de A. On la note $\inf(A)$.

Remarque 7.3. Notons que si elles existent, $\sup(A)$ et $\inf(A)$ sont nécessairement uniques, puisque ce sont respectivement un maximum et un minimum.

Remarque 7.4. Si A admet un maximum, alors A admet une borne supérieure égale à ce maximum. De même avec le minimum et la borne inférieure.

Théorème 7.3 (Propriété fondamentale de $\mathbb{R}$ - admise)**.** *Tel qu'il est construit, $\mathbb{R}$ vérifie la propriété fondamentale suivante : toute partie non vide et majorée de $\mathbb{R}$ admet une borne supérieure ; et toute partie non vide et minoré de $\mathbb{R}$ possède une borne inférieure.*

Remarque 7.5. La seconde propriété peut se voir comme une conséquence de la première, et réciproquement.

Exemple 7.1. La borne supérieure de $]-\infty, 1[$ vaut 1.

Remarque 7.6. Attention ! Selon les cas, la borne supérieure (ou inférieure) peut appartenir ou non à A. Par exemple, la borne supérieure de $]-\infty, 1[$ vaut 1 mais la borne supérieure de $]-\infty, 1]$ vaut 1 aussi.

Proposition 7.1 (Passage au sup/ à l'inf dans une inégalité large)**.** *Soit $A \subset \mathbb{R}$ non vide.*

- *Si $\exists M \in \mathbb{R}, \forall x \in A,\ x \leqslant M$, alors A admet une borne supérieure et on a $\sup(A) \leqslant M$.*
- *Si $\exists m \in \mathbb{R}, \forall x \in A,\ x \geqslant m$, alors A admet une borne inférieure et on a $\inf(A) \geqslant m$.*

Remarque 7.7. *A fortiori*, si $\forall x \in A,\ x < M$, alors on a encore $\sup(A) \leqslant M$. Mais **attention**, l'inégalité stricte pour le majorant est devenue large pour la borne supérieure ! Et c'est le meilleur résultat que l'on puisse obtenir, il suffit de considérer l'exemple de $]-\infty, 1[$ donné précédemment.

Cette proposition est souvent utilisée comme méthode alternative pour déterminer la borne supérieure (ou inférieure) d'une partie. On procède alors par double inégalité.

Exemple 7.2. Si A est une partie non vide et majorée de $\mathbb{R}$, et λ est un réel positif ou nul, on a $\sup(\lambda A) = \lambda \sup(A)$.

Théorème 7.4 (Partie entière)**.** *Soit x un réel. Il existe un unique $n \in \mathbb{Z}$ tel que $n \leqslant x < n+1$. On l'appelle la **partie entière de** x, et on le note $\lfloor x \rfloor$.*

Définition 7.4 (Partie fractionnaire)**.** Soit x un réel. La **partie fractionnaire de** x, généralement notée $\mathrm{Frac}(x)$ (voire $\{x\}$ s'il n'y a pas d'ambiguïté), est définie par $\mathrm{Frac}(x) = x - \lfloor x \rfloor$. Elle appartient donc à $[0, 1[$.

Remarque 7.8. Intuitivement, au moins si $x \geqslant 0$, $\lfloor x \rfloor$ représente "les chiffres de x avant la virgule" et $\mathrm{Frac}(x)$ représente "les chiffres de x après la virgule.

Exemple 7.3. Dans une division euclidienne $a = bq + r$, on a toujours $q = \left\lfloor \frac{a}{b} \right\rfloor$.

Proposition 7.2. *Si $x \in \mathbb{R}$ et $m \in \mathbb{Z}$, alors on a $\lfloor x + m \rfloor = \lfloor x \rfloor + m$.*

Remarque 7.9. Attention, si $m \in \mathbb{Z}$, ne pas écrire $\lfloor mx \rfloor = m \lfloor x \rfloor$, c'est faux en général ! Prendre par exemple $x = 1/2$ et $m = 2$. De même, on n'a pas $\lfloor x + y \rfloor = \lfloor x \rfloor + \lfloor y \rfloor$. Prendre par exemple $x = y = 1/2$.

Proposition 7.3. *La fonction $x \mapsto \lfloor x \rfloor$ est croissante de $\mathbb{R}$ dans $\mathbb{R}$. La fonction $x \mapsto \{x\}$ est 1-périodique de $\mathbb{R}$ dans $\mathbb{R}$.*

Exemple 7.4. Si $x \geqslant n$ avec $x \in \mathbb{R}$ et $n \in \mathbb{Z}$, alors $\lfloor x \rfloor \geqslant n$

Définition 7.5 (Valeur absolue)**.** La **valeur absolue** d'un réel x est définie par

$$|x| = \begin{cases} x & \text{si} x \geqslant 0 \\ -x & \text{si} x \leqslant 0 \end{cases}$$

Au passage, on vérifie que la définition est bien consistante lorsque $x = 0$. Voici à présent quelques résultats de bon sens, que l'on admet généralement "d'instinct" mais qu'il n'est pas forcément mauvais de démontrer une fois.

Proposition 7.4. *On a toujours $|x| \geqslant 0$, avec égalité si, et seulement si, $x = 0$.*

Proposition 7.5. *On a toujours $-|x| \leqslant x \leqslant |x|$.*

Lemme 7.1. *On a $x = \pm y$ si, et seulement si, $|x| = |y|$.*

Corollaire 7.3. *On a $\forall (x, y) \in \mathbb{R}^2,\ |xy| = |x| \times |y|$.*

Remarque 7.10. De même, on montre que

$$\forall y \in \mathbb{R}^\star,\ \left|\frac{1}{y}\right| \quad \text{et} \quad \forall (x, y) \in \mathbb{R} \times \mathbb{R}^\star,\ \left|\frac{x}{y}\right| = \frac{|x|}{|y|}$$

Remarque 7.11. On lit couramment que $\forall x \in \mathbb{R},\ \sqrt{x^2} = |x|$. cette assertion est vraie, mais encore faut-il avoir construit la fonction $\sqrt{\cdot}$. Cela dit, dès à présent, on peut montrer que l'application de $\mathbb{R}_+$ dans $\mathbb{R}_+$ $t \mapsto t^2$ est strictement croissante. Soit $0 \leqslant t < u$. Alors $t^2 \leqslant tu$ et $tu < u^2$. On conclut par transitivité.

Une fois ce point démontré, on en déduit que tout $y \in \mathbb{R}_+$ admet au plus une racine carrée dans $\mathbb{R}_+$. En particulier, soit $x \in \mathbb{R}$. Alors $|x|$ est l'unique racine carrée de x^2 dans $\mathbb{R}_+$. En effet, on a $|x|^2 = |x^2| = x^2$.

Proposition 7.6. *Soit $\delta \geqslant 0$ fixé. Alors $\forall x \in \mathbb{R},\ -\delta \leqslant x \leqslant \delta \iff |x| \leqslant \delta$.*

Remarque 7.12. Plus généralement, soit $a \in \mathbb{R}$ et $\delta \geqslant 0$. On a :

$$\begin{cases} \forall x \in \mathbb{R},\ x \in [a - \delta, a + \delta] \iff |x - a| \leqslant \delta \\ \forall x \in \mathbb{R},\ x \in]a - \delta, a + \delta[\iff |x - a| < \delta \end{cases}$$

Proposition 7.7 (Formules du min et du max)**.** *Soit $(x, y) \in \mathbb{R}^2$. On a*

$$\begin{cases} \max(x, y) = \dfrac{x + y}{2} + \dfrac{|x - y|}{2} \\ \min(x, y) = \dfrac{x + y}{2} - \dfrac{|x - y|}{2} \end{cases}$$

Proposition 7.8 (Inégalités triangulaires)**.** *Soit $(x, y) \in \mathbb{R}^2$. On a :*

$$\begin{cases} |x + y| \leqslant |x| + |y| \\ \big||x| - |y|\big| \leqslant |x - y| \end{cases}$$

Définition 7.6 (Partie positive, partie négative)**.** Soit $x \in \mathbb{R}$. La **partie positive** de x est définie par $x^+ = \max(x, 0)$. La **partie négative** de x est définie par $x^- = \max(-x, 0)$. On a

$$\begin{cases} 0 \leqslant x^+, x^- \leqslant x \\ x = x^+ - x^- \end{cases}$$

Définition 7.7 (Droite réelle achevée $\overline{\mathbb{R}}$)**.** La **droite réelle achevée** $\overline{\mathbb{R}}$ est égale à l'ensemble $\mathbb{R}$ auquel on adjoint deux nouveaux éléments $-\infty$ et $+\infty$. La relation $\leqslant$ se prolonge de manière évidente par

$$\forall x \in \overline{\mathbb{R}},\ -\infty \leqslant x \leqslant +\infty$$

On vérifie sans difficulté que $\leqslant$ reste une relation d'ordre totale sur $\overline{\mathbb{R}}$.

Remarque 7.13. Insistons sur le fait que $-\infty$ et $+\infty$ sont de **nouveaux** éléments. Autrement dit, ils ne sont pas réels et ils sont distincts. En particulier, on en déduit que

$$\forall x \in \mathbb{R},\ -\infty < x < +\infty$$

Définition 7.8. Soit $a, b \in \overline{\mathbb{R}}$. Pour généraliser le cas fini, on définit les ensembles suivants.

$$\begin{cases} [a,b] = \left\{t \in \overline{\mathbb{R}} \mid a \leqslant t \leqslant b\right\} \\ [a,b[= \left\{t \in \overline{\mathbb{R}} \mid a \leqslant t < b\right\} \\ ]a,b] = \left\{t \in \overline{\mathbb{R}} \mid a < t \leqslant b\right\} \\ ]a,b[= \left\{t \in \overline{\mathbb{R}} \mid a < t < b\right\} \end{cases}$$

Dans le cas où a et b sont réels, on vérifie sans peine que l'ancienne définition coïncide avec l'ancienne.

Exemple 7.5. On a $\overline{\mathbb{R}} = [-\infty, +\infty]$.

Exemple 7.6. En particulier, si $a > b$, on obtient que $[a,b] = \emptyset$.

Remarque 7.14. En revanche, l'addition et la multiplication sont plus délicates à prolonger. on ne peut le faire que partiellement en posant

$$\begin{cases} \forall a \in]-\infty, +\infty],\ a + (+\infty) = (+\infty) + a = +\infty \\ \forall a \in [-\infty, +\infty[,\ a + (-\infty) = (-\infty) + a = -\infty \end{cases}$$

$$\begin{cases} \forall a \in]0, +\infty],\ (+\infty) \times a = a \times (+\infty) = +\infty \\ \forall a \in]0, +\infty],\ (-\infty) \times a = a \times (-\infty) = -\infty \\ \forall a \in [-\infty, 0[,\ (+\infty) \times a = a \times (+\infty) = -\infty \\ \forall a \in [-\infty, 0[,\ (-\infty) \times a = a \times (-\infty) = +\infty \end{cases}$$

Autrement dit, il y a ce qu'on appelle deux **formes indéterminées**, qui sont du type $\infty - \infty$ et $0 \times \infty$. Nous verrons dans le chapitre "Sommabilité" un ultime prolongement, mais nous n'en faisons pas usage dans l'immédiat.

Exemple 7.7. Sur $]-\infty, +\infty]$, la loi $+$ reste associative et commutative. 0 reste neutre, et $+\infty$ absorbant. Remarque symétrique pour $[-\infty, +\infty[$.

Remarque 7.15. Soit $A \subset \mathbb{R}$ non vide. On montre facilement que A admet une borne supérieure au sens de $\overline{\mathbb{R}}$.

- Si A est majorée par un réel, on montre que $\sup_{\overline{\mathbb{R}}}(A) = \sup(A)$
- Sinon, on montre que $\sup_{\overline{\mathbb{R}}} = +\infty$

Le premier point permet d'écrire $\sup(A)$ sans risque de confusion. En résumé, on peut dire que $\sup(A) \in \mathbb{R} \cup \{+\infty\}$, et que A est majorée par un réel si, et seulement si, $\sup(A) < +\infty$. Bine sûr, on peut faire la remarque symétrique avec la borne inférieure.

Définition 7.9 (Intervalle)**.** Un **intervalle** de $\mathbb{R}$ est une **partie convexe** de $\mathbb{R}$. Autrement dit, c'est une partie $I \subset \mathbb{R}$ telle que

$$\forall (x,y) \in I^2,\ x \leqslant y \implies [x,y] \subset I$$

Remarque 7.16. En pratique, on peut se restreindre à $x < y$ puisque si $x = y$, alors $[x,y] = \{x\} \subset I$ car $x \in I$. C'est ce que l'on fera à chaque fois dans les exercices.

Théorème 7.5 (Caractérisation des intervalles non vides de $\mathbb{R}$)**.** *Les intervalles non vides de* $\mathbb{R}$ *sont :*

- *les* $[a,b]$ *avec* $-\infty < a \leqslant b < +\infty$
- *les* $]a,b]$ *avec* $-\infty \leqslant a < b < +\infty$
- *les* $[a,b[$ *avec* $-\infty < a < b \leqslant +\infty$
- *les* $]a,b[$ *avec* $-\infty \leqslant a < b \leqslant +\infty$

Remarque 7.17. Soit $I \subset \mathbb{R}$ un intervalle non vide. La démonstration précédente montre qu'il suffit de choisir $a = \inf(I)$ et $b = \sup(I)$. Mais on est en droit de se demander s'il s'agit d'une condition nécessaire. La réponse est oui.

- Si I s'écrit d'une des quatre manières ci-dessus, on montre que nécessairement $a = \inf(I)$ et $b = \sup(I)$. En particulier, a et b sont uniques. On les appelle les **bornes** de I. Elles sont dites **ouvertes** ou **fermées** selon que l'inégalité qui les concerne est stricte ou large.
- Au passage, on en déduit que les quatre cas ci-dessus sont deux à deux exclusifs : un intervalle ne peut pas être à la fois ouvert et fermé, etc.
- Si a et b l'intervalle est dit **borné**. Un intervalle fermé borné s'appelle aussi un **segment**.

Définition 7.10 (Intérieur)**.** Soit $I \subset \mathbb{R}$ un intervalle non vide, ainsi que $a = \inf(I)$ et $b = \sup(I)$ ses deux bornes, avec $a \leqslant b$ dans $\overline{\mathbb{R}}$. L'**intérieur** de I est l'intervalle $]a,b[$. On le note $\mathring{I}$.

Exemple 7.8. Un intervalle non vide I est d'intérieur non vide si, et seulement si, il contient au moins deux points.

Proposition 7.9. *L'intersection de deux intervalles est un intervalle.*

Remarque 7.18. En revanche, l'union de deux intervalles n'a aucun raison d'être un intervalle en général. Considérer par exemple $[0,1] \cup [2,3]$.

Définition 7.11 (Partie dense dans $\mathbb{R}$)**.** Soit A une partie de $\mathbb{R}$. Les deux propositions suivantes sont équivalentes :

- A rencontre tout intervalle ouvert borné non trivial :

$$\forall (a,b) \in \mathbb{R}^2,\ a < b \implies A \cap]a,b[\neq \emptyset$$

- A rencontre tout intervalle fermé borné non trivial :

$$\forall (a,b) \in \mathbb{R}^2,\ a < b \implies A \cap [a,b] \neq \emptyset$$

Lorsque A vérifie l'une de ces assertions, on dit que A **est dense dans** $\mathbb{R}$.

On sait qu'il n'existe pas de rationnel r tel que $r^2 = 2$. Sinon, il existerait deux entiers $p, q \in \mathbb{Z}^\star$ tels que $2q^2 = p^2$ donc on aurait $2v_2(p) = v_2(p^2) = v_2(2q^2) = 1 + 2v_2(q)$. Or, un entier ne pouvant pas à la fois être pair et impair, on aurait une contradiction. On en tire plusieurs conséquences.

Théorème 7.6. *$\mathbb{Q}$ et $\mathbb{R} \setminus \mathbb{Q}$ sont denses dans $\mathbb{R}$.*

Remarque 7.19. Dans $\mathbb{Q}$, la propriété de la borne supérieure est fausse. En effet, il suffit de considérer l'ensemble $]0, \sqrt{2}[\cap\mathbb{Q}$. Supposons qu'il possède une borne supérieure $\alpha \in \mathbb{Q}$. Si on a $\alpha > \sqrt{2}$, alors, comme il existe un rationnel dans $]\sqrt{2}, \alpha[$, ce rationnel est un majorant, ce qui est une contradiction car α doit être le plus petit majorant. De même, on ne peut pas avoir $\alpha < \sqrt{2}$. Donc on doit avoir $\alpha = \sqrt{2}$. Contradiction, car $\sqrt{2}$ n'est pas rationnel.

Nous verrons ultérieurement que le TVI provient essentiellement de la propriété de la borne supérieure, puis que beaucoup de théorèmes d'analyse réelle proviennent eux-mêmes du TVI. On voit donc que la propriété de la borne supérieure est véritablement fondatrice. Le fait qu'elle soit fausse dans $\mathbb{Q}$ limite considérablement les théorèmes que l'on peut y trouver.

Définition 7.12 (Ensemble des décimaux $\mathbb{D}$)**.** L'ensemble des **nombres décimaux** est défini par

$$\mathbb{D} = \left\{ \frac{p}{10^n} \mid (p, n) \in \mathbb{Z} \times \mathbb{N} \right\}$$

Exemple 7.9. On peut montrer que $\mathbb{D}$ est un anneau intègre.

Définition 7.13 (Valeur approchée par excès, par défaut)**.** Soit $x \in \mathbb{R}$ et $n \in \mathbb{N}$. Alors il existe un unique $p \in \mathbb{Z}$ tel que

$$\frac{p}{10^n} \leqslant x < \frac{p}{10^n} + \frac{1}{10^n}$$

Par définition, cet entier p vaut $\lfloor 10^n x \rfloor$. $\dfrac{p}{10^n}$ s'appelle la **valeur décimale approchée de x à 10^{-n} près**. On dit que c'est la valeur approchée **par défaut**, tandis que $\dfrac{p}{10^n} + \dfrac{1}{10^n}$ est la valeur approchée **par excès**.

Le théorème qui suit se révèle souvent bien utile, même s'il ne figure pas au programme.

Théorème 7.7 (Sous-groupes additifs de $\mathbb{R}$, HP à connaître)**.** *Les sous-groupes additifs de $\mathbb{R}$ sont soit de la forme $a\mathbb{Z}$ avec $a \in \mathbb{R}_+$, soit denses dans $\mathbb{R}$.*

Démonstration. Soit H un sous-groupe additif de $\mathbb{R}$. Si $H = \{0\}$, alors il est bien de la forme $a\mathbb{Z}$ avec $a = 0$. Sinon, il contient un réel non nul, et quitte à passer à l'opposé, il contient un réel strictement positif. Donc $H \cap \mathbb{R}_+^\star$ est non vide, et il admet une borne inférieure a.

- Supposons que $a \in H$, et montrons alors que $H = a\mathbb{Z}$.
  D'une part, on a $< a >= a\mathbb{Z} \subset H$ par définition du sous-groupe engendré. Réciproquement, soit $x \in H$. Comme $a \in \mathbb{R}_+^\star$, on peut considérer $q = \left\lfloor \dfrac{x}{a} \right\rfloor \in \mathbb{Z}$ et $r = x - aq$. Alors, d'après la première inclusion, $r \in H$, et comme $r \in [0, a[$, on a $r = 0$ par minimalité de a. Donc $x = aq \in a\mathbb{Z}$.
- Supposons que $a \notin H \cap \mathbb{R}_+^\star$., et montrons alors que H est dense dans $\mathbb{R}$.
  Donnons-nous un intervalle $]x, y[$ de largeur $\varepsilon = y - x > 0$.

- Montrons qu'on peut trouver un $h \in]0, \varepsilon[$ dans H. Prenons dans $H \cap \mathbb{R}_+^\star$ un h_1 tel que $h_1 \in [a, a+\varepsilon[$. Comme $a \notin H \cap \mathbb{R}_+^\star$, on a même $h_1 \in]a, a+\varepsilon[$. Par le même raisonnement, prenons dans H un $h_2 \in]a, h_1[$. il suffit alors de poser $h = h_1 - h_2$.
- Ensuite, posons $q = \left\lfloor \frac{x}{h} \right\rfloor + 1 \in \mathbb{Z}$. Alors qh appartient à H et il vérifie $x < qh < y$.

- Enfin, aucun $H = a\mathbb{Z}$ ne peut être dense dans $\mathbb{R}$ (il s'agit bien d'un "soit/soit"). En effet :
  - Si $a = 0$, alors l'intervalle $]0, 1[$ ne contient aucun élément de H.
  - Si $a > 0$, alors l'intervalle $]0, a[$ ne contient aucun élément de H.

Ainsi, la preuve est achevée. □

2 Suites réelles

Définition 7.14 (Suite réelle)**.** Une **suite réelle** $(u_n)_{n\in\mathbb{N}}$ est une application de $\mathbb{N}$ dans $\mathbb{R}$ qui à n associe u_n. n s'appelle l'**indice**. On peut plus prosaïquement la noter u s'il n'y a pas d'ambiguïté sur l'ensemble que parcourt l'indice.

Définition 7.15 (Relation $\leqslant$ sur $\mathbb{R}^\mathbb{N}$)**.** On définit une relation d'ordre sur $\mathbb{R}^\mathbb{N}$ par

$$u \leqslant v \iff \forall n \in \mathbb{N},\ u_n \leqslant v_n$$

C'est une relation d'ordre partielle.

Définition 7.16 (Suite majorée, minorée, bornée)**.** Une suite u est **majorée** lorsque l'ensemble $\{u_n \mid n \in \mathbb{N}\}$ est majoré. Elle est **minorée** si cet ensemble est minoré. Elle est **bornée** si elle est à la fois majorée et minorée.

Remarque 7.20. Pour montrer efficacement qu'une suite $(u_n)_{n\in\mathbb{N}}$ est bornée, on montrera de manière équivalente de $(|u_n|)_{n\in\mathbb{N}}$ est majorée (une seule opération au lieu de deux).

Exemple 7.10. Pour montrer que $(\cos(n))_{n\in\mathbb{N}}$ est bornée, il suffit de remarquer que $\forall n \in \mathbb{N},\ |\cos(n)| \leqslant 1$.

Exemple 7.11. Les suites bornées forment un sous-anneau de $\mathbb{R}^\mathbb{N}$.

Définition 7.17 (Suite croissante, décroissante, monotone, stationnaire)**.** Une suite u est **croissante** lorsque

$$\forall (m, n) \in \mathbb{N}^2,\ m \leqslant n \implies u_m \leqslant u_n$$

En pratique, il suffira de vérifier que $\forall n \in \mathbb{N},\ u_n \leqslant u_{n+1}$. Elle est **décroissante** lorsque

$$\forall (m, n) \in \mathbb{N}^2,\ m \leqslant n \implies u_m \geqslant u_n$$

En pratique, il suffira de vérifier que $\forall n \in \mathbb{N},\ u_n \geqslant u_{n+1}$. Elle est **monotone** si elle croissante ou décroissante. Elle est **stationnaire** si elle est constante à partir d'un certain rang.

Remarque 7.21. On définit de même les notions de suite strictement croissante, strictement décroissante et strictement monotone. En revanche, les notions de suite strictement majorée, strictement minorée et strictement bornée sont sans intérêt, car elles sont équivalentes aux mêmes notions sans le "strictement".

Définition 7.18 (Extension)**.** On peut définir de même des suites réelles indexées sur $\mathbb{Z}$: $(u_n)_{n\in\mathbb{Z}}$; sur $[\![n_0, +\infty[\![$ pour $n_0 \in \mathbb{Z}$: $(u_n)_{n\in[\![n_0,+\infty[\![}$; ou même sur deux indices à la fois, c'est ce qu'on appelle une **suite double** : $(u_{m,n})_{(m,n)\in\mathbb{N}^2}$ ou $(u_{m,n})_{(m,n)\in\mathbb{Z}^2}$.

Les relations d'ordre $\leqslant$ et $\geqslant$ sont toujours définies, ainsi que les notions de suites majorées, minorées et bornées. En revanche, si les notions de suite croissante, décroissante et monotone ont encore du sens pour $(u_n)_{n\in\mathbb{Z}}$ et $(u_n)_{n\in[\![n_0,+\infty[\![}$, elles n'ont plus de sens pour les suites doubles.

Définition 7.19 (Convergence, divergence)**.** Soit $l \in \mathbb{R}$. On dit que la suite réelle $(u_n)_{n\in\mathbb{N}}$ **converge vers** l ou que u_n **tend vers** l lorsque

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N},\ \forall n \in \mathbb{N},\ n \geqslant N \implies |u_n - l| \leqslant \varepsilon$$

Le nombre l s'appelle **une limite de** u et on écrit $u_n \xrightarrow[n\to+\infty]{} l$ ou encore $u_n \to l$, voire $\lim\limits_{n\to+\infty} u_n = l$. Une suite est dite **convergente** lorsqu'elle admet une limite finie : $\exists l \in \mathbb{R},\ u_n \to l$. Au contraire, elle est dite **divergente** lorsqu'elle n'admet pas de limite finie :

$$\forall l \in \mathbb{R},\ \exists \varepsilon > 0,\ \forall N \in \mathbb{N},\ \exists n \in \mathbb{N},\ (n \geqslant N) \wedge (|u_n - l| > \varepsilon)$$

Remarque 7.22. On montre facilement que la définition de la convergence est équivalente si on remplace les $\leqslant \varepsilon$ par des $< \varepsilon$. On utilisera régulièrement cette remarque pour des raisons de confort lors de certains démonstrations.

Remarque 7.23. De manière plus intuitive, on peut dire que $(u_n)_{n\in\mathbb{N}}$ converge vers l lorsque, aussi petit soit le couloir autour de l, u_n finit par se "laisser apprivoiser" et se retrouve alors coincé dans ce couloir au bout 'un certain temps.

Au contraire, si la suite diverge, cela signifie qu'elle est "sauvage" (grrr Fred) et que toute tentative pour la faire tendre vers une limite échouera. Plus précisément, quel que soit $l \in \mathbb{R}$, u_n ne se "laisse pas apprivoiser" et il existe un "couloir récalcitrant" d'épaisseur non nulle autour de l tel que u_n ressorte régulièrement de ce couloir, aussi loin qu'on aille chercher.

Exemple 7.12. Soit $l \in \mathbb{R}$. Si $(u_n)_{n\in\mathbb{N}}$ est constante égale à l, alors elle converge vers l. Plus généralement, si elle stationne à la valeur l, alors elle tend vers l.

Remarque 7.24. Soit $l \in \mathbb{R}$. Quel que soit $\varepsilon > 0$, on a $|u_n - l| \leqslant \varepsilon$ ssi $|(u_n - l) - 0| \leqslant \varepsilon$ ssi $||u_n - l| - 0| \leqslant \varepsilon$. Donc

$$u_n \to l \iff u_n - l \to 0 \iff |u_n - l| \to 0$$

en particulier, pour $l = 0$, on obtient

$$u_n \to l \iff |u_n| \to 0$$

Proposition 7.10 (Unicité de la limite)**.** *Si u possède une limite, alors celle-ci est unique. A partir de maintenant, on pourra donc parler de* ***la*** *limite d'une suite convergente.*

Remarque 7.25. On utilisera souvent ce résultat pour montrer l'égalité de deux nombres, en calculant par exemple une limite de deux façons différentes et en concluant par unicité de la limite.

Théorème 7.8. *Toute suite convergente est bornée.*

Remarque 7.26. On ne peut pas prendre le maximum d'un nombre infini de réels, or, il y a un nombre infini de $|u_n|$. On voit donc toute l'importance de "circonscrire" le domaine sur lequel évolue n pour prendre le maximum, ce qui nous avons pu faire grâce à l'hypothèse de convergence.

C'est une technique extrêmement générale, qui consiste à couper la suite en deux parties (voire plus). Une première partie finie, sur laquelle on peut faire à peu près ce que l'on veut (prendre le minimum, le maximum, calculer un produit, une somme...) ; et une deuxième partie infinie sur laquelle on a besoin d'hypothèses supplémentaires (majorant, minorant, convergence, etc.) pour pouvoir agir en une fois sur une infinité d'indices. Souvent, la première partie sera plus facile à traiter que la deuxième.

On trouvera un exemple de cette technique dans la partie des compléments sur la moyenne de Cesàro.

Définition 7.20. On dit que u_n tend vers $+\infty$ lorsque

$$\forall M \in \mathbb{R},\ \exists N \in \mathbb{N},\ \forall n \in \mathbb{N},\ n \geqslant N \implies u_n \geqslant M$$

On note $u_n \xrightarrow[n\to+\infty]{} +\infty$. Là aussi, la définition est équivalente si on remplace $\geqslant M$ par $> M$.

Remarque 7.27. La définition est également équivalente en remplaçant $M \in \mathbb{R}$ par $M \in \mathbb{R}_+$ ou $M \in \mathbb{R}_+^\star$. Elle est plus généralement équivalente avec $M \in]A, +\infty[$ ou $M \in [A, +\infty[$ avec $A \in \mathbb{R}$ fixé.

Remarque 7.28. De manière imagée, lorsque $u_n \to +\infty$, on peut dire que aussi haute soit l'altitude, u_n finit par la dépasser définitivement à partir d'un certain temps.

Définition 7.21. On dit que u_n tend vers $-\infty$ lorsque

$$\forall m \in \mathbb{R},\ \exists N \in \mathbb{N},\ \forall n \in \mathbb{N},\ n \geqslant N \implies u_n \leqslant m$$

On note $u_n \xrightarrow[n\to+\infty]{} -\infty$. Là aussi, la définition est équivalente si on remplace $\leqslant m$ par $< m$.

Remarque 7.29. La définition est également équivalente en remplaçant $m \in \mathbb{R}$ par $m \in \mathbb{R}_-$ ou $m \in \mathbb{R}_-^\star$. Elle est plus généralement équivalente avec $m \in]-\infty, A[$ ou $m \in]-\infty, A]$ avec $A \in \mathbb{R}$ fixé.

Remarque 7.30. De manière imagée, lorsque $u_n \to +\infty$, on peut dire que aussi basse soit la profondeur, u_n finit par la dépasser définitivement à partir d'un certain temps.

Remarque 7.31. Si $u_n \to +\infty$, alors u est divergente. De même avec $-\infty$. C'est la contraposée du théorème qui stipule que tout suite convergente est bornée.

Par ailleurs, elle ne peut pas tendre à la fois vers $-\infty$ et $+\infty$, ou vers $\pm\infty$ et un réel. Pour toutes ces raisons, si u admet une limite dans $\overline{\mathbb{R}}$, celle-ci continue à être unique.

Exemple 7.13. La suite définie par $u_n = n$ tend vers $+\infty$. Plus généralement, pour tout $k \in \mathbb{N}^\star$, la suite définie par $u_n = n^k$ tend vers $+\infty$.

Exemple 7.14. La suite définie sur $\mathbb{N}^\star$ par $u_n = \dfrac{1}{n}$ tend vers 0. Plus généralement, pour tout $k \in \mathbb{N}^\star$, la suite définie sur $\mathbb{N}^\star$ par $u_n = \dfrac{1}{n^k}$ tend vers 0.

Proposition 7.11 (Principe de troncature). *Soit $p \in \mathbb{N}^\star$ et $l \in \overline{\mathbb{R}}$. La suite $(u_n)_{n\in\mathbb{N}}$ tend vers l si, et seulement si, la suite $(u_n)_{n\geqslant p}$ tend vers l.*

Exemple 7.15. Si $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ coïncident à partir d'un certain rang, alors $u_n \to l$ si, et seulement si, $v_n \to l$. On dit que la limite est une notion **asymptotique**.

Exemple 7.16. Soit $u \in \mathbb{R}^{\mathbb{N}}$ une suite qui ne s'annule pas à partir d'un certain rang. Lorsqu'on écrire une assertion comme

$$\frac{1}{u_n} \to l$$

il sera inutile de se préoccuper du rang à partir duquel la suite $\left(\frac{1}{u_n}\right)$ a été définie (il suffit de savoir qu'il en existe au moins un).

Proposition 7.12 (Principe de décalage d'indice). *Soit $p \in \mathbb{N}^\star$ et $l \in \overline{\mathbb{R}}$. La suite $(u_n)_{n\geqslant p}$ tend vers l si, et seulement si, la suite $(u_{n+p})_{n\geqslant 0}$.*

Exemple 7.17. Si $u_n \to l$, alors par troncature et décalage d'indice, on obtient que $u_{n+1} \to l$.

Théorème 7.9 (Théorème d'encadrement). *Soit $l \in \mathbb{R}$, et u, v et w trois suites telles que*

$$\begin{cases} u_n \to l \\ w_n \to l \\ \forall n \in \mathbb{N},\ u_n \leqslant v_n \leqslant w_n \end{cases}$$

Alors, $v_n \to l$.

Remarque 7.32. Ce théorème est aussi appelé "théorème des gendarmes" à cause de sa forme imagée. Si on imagine v_n comme la trajectoire d'un voleur qui se fait poursuivre par deux gendarmes u_n et w_n, l'un à sa gauche et l'autre à sa droite, si les deux gendarmes convergent vers la même limite, alors ils "coincent le voleur en le prenant en sandwich".

Corollaire 7.4. *Si $|u_n| \leqslant \alpha_n$ avec $\alpha_n \to 0$, alors $u_n \to 0$.*

Exemple 7.18. La suite de terme général $\dfrac{\cos(n)}{n}$ définie pour $n \geqslant 1$ converge vers 0.

Théorème 7.10 (Théorème de minoration/majoration). *Supposons que $\forall n \in \mathbb{N},\ u_n \leqslant v_n$. Si $u_n \to +\infty$, alors $v_n \to +\infty$ (minoration). Si $v_n \to -\infty$, alors $u_n \to -\infty$ (majoration).*

Remarque 7.33. Le théorème d'encadrement, son corollaire et le théorème de comparaison restent vrais si on n'a les inégalités qu'à partir d'un certain rang.

Le théorème suivant est à mettre en regard avec ce que nous avons vu sur les bornes supérieures et les bornes inférieures.

Théorème 7.11 (Passage à la limite dans une inégalité dans une inégalité large). *Soit u une suite qui converge vers une limite $l \in \mathbb{R}$. Soit m et M deux réels.*

- *Si $\forall n \in \mathbb{N},\ u_n \leqslant M$, alors $l \leqslant M$.*
- *Si $\forall n \in \mathbb{N},\ u_n \geqslant m$, alors $l \geqslant m$.*

Remarque 7.34. Attention, quand bien même on aurait une inégalité stricte $u_n < M$, celle-ci deviendrait large en passant à la limite !

Remarque 7.35. Le théorème reste vrai si les inégalités n'ont lieu qu'à partir d'un certain rang. Le théorème reste également vrai si on autorise l à être infinie.

Exemple 7.19. Soit u et v deux suites telles que $u_n \to l \in \mathbb{R}$ et $v_n \to l' \in \mathbb{R}$. Si

$$\forall n \in \mathbb{N},\ u_n \leqslant v_n$$

alors $l \leqslant l'$ (considérer $u_n - v_n$ et lui appliquer le théorème qui précède).

Voici maintenant une sorte de réciproque partielle.

Théorème 7.12. *Soit $a < b$ des réels, et $(u_n)_{n\in\mathbb{N}}$ une suite de réels qui converge vers $l \in \mathbb{R}$.*

- *Si $l > a$, alors $u_n > a$ à partir d'un certain rang.*
- *Si $l < b$, alors $u_n < b$ à partir d'un certain rang.*
- *Si $l \in]a, b[$, alors $u_n \in]a, b[$ à partir d'un certain rang.*

Exemple 7.20. Soit $(u_n)_{n\in\mathbb{N}}$ qui tend vers $l \in \overline{\mathbb{R}}$. Si $l \neq 0$, alors $u_n \neq 0$ APCR.

Exercice 7.1. Soit $u \in \mathbb{Z}^{\mathbb{N}}$ qui converge vers une limite finie l. Montrer que $l \in \mathbb{Z}$ et que u est stationnaire.

Nous arrivons à une séquence importante. Depuis le lycée, on effectue régulièrement des "opérations algébriques sur les limites"., sans trop savoir si on la droit de le faire. Nous nous employons maintenant à le justifier.

Proposition 7.13 (Combinaison linéaire)**.** *Soit u et v des suites réelles qui convergent respectivement vers l et l'. Alors, pour tous $\lambda, \mu \in \mathbb{R}$, $\lambda u + \mu v$ converge vers $\lambda l + \mu l'$.*

Exemple 7.21. Si $u_n \to l$, alors $\lambda u_n \to \lambda l$.

Exemple 7.22. Si u converge, alors nécessairement $u_{n+1} - u_n \to 0$.

Corollaire 7.5. *Si u est bornée et $v_n \to 0$, alors $u_n v_n \to 0$.*

Proposition 7.14 (Produit)**.** *Soit u et v des suites réelles qui convergent respectivement vers l et l'. Alors uv converge vers ll'.*

Exemple 7.23. Les suites convergentes forment donc un sous-anneau des suites bornées.

En utilisant la deuxième inégalité triangulaire, on obtient la

Proposition 7.15 (Valeur absolue)**.** *Si $u_n \to l$, alors $|u_n| \to |l|$.*

Avant de traiter le passage à l'inverse, nous avons besoin du

Lemme 7.2. *Si $u_n \to l \in \mathbb{R}_+^\star$, alors u est minorée par un certain $\eta > 0$ à partir d'un certain rang.*

Remarque 7.36. Attention à ce résultat assez subtil : ce n'est pas parce qu'on a $u_n > 0$ pour tout n que u peut être minorée par un réel strictement positif. Prendre comme contre-exemple la suite $\left(\frac{1}{n}\right)_{n\in\mathbb{N}^\star}$.

Proposition 7.16 (Passage à l'inverse)**.** *Supposons que u ne s'annule par et qu'elle converge vers un réel $l \neq 0$. Alors*

$$\frac{1}{u_n} \to \frac{1}{l}$$

Remarque 7.37. Dans un exercice, même si on ne suppose pas *a priori* que u ne s'annule pas, le fait que $l \neq 0$ garanti ue la suite ne s'annule pas à partir d'un certain rang, ce qui permet de conserver le résultat.

Proposition 7.17. *Les limites passent encore aux opérations algébriques (combinaison linéaire, produit, valeur absolue) dans $\overline{\mathbb{R}}$, pourvu qu' on n'ait pas de forme indéterminée.*

Définition 7.22. On écrit $u_n \to 0^+$ lorsque $u_n \to 0$ et $u_n > 0$ à partir d'un certain rang. On écrit $u_n \to 0^-$ lorsque $u_n \to 0$ et $u_n < 0$ à partir d'un certain rang.

Proposition 7.18. *Supposons que u ne s'annule pas. Si $u_n \to 0^+$, alors $\dfrac{1}{u_n} \to +\infty$, et si $u_n \to +\infty$, alors $\dfrac{1}{u_n} \to 0^+$. De même, si $u_n \to 0^-$, alors $\dfrac{1}{u_n} \to -\infty$, et si $u_n \to -\infty$, alors $\dfrac{1}{u_n} \to 0^-$.*

Remarque 7.38. En revanche, une limite du type $\dfrac{1}{0}$ doit être considérée comme une forme indéterminée. La proposition ci-dessus montre qu'on peut obtenir $-\infty$ comme $+\infty$. On peut même imaginer que la suite $\dfrac{1}{u_n}$ n'ait pas de limite : on pourra considérer $u_n = \dfrac{(-1)^n}{n}$.

Théorème 7.13. *Soit α un réel.*

- *Si $\alpha > 1$, alors $\alpha^n \to +\infty$.*
- *Si $|\alpha| < 1$, alors $\alpha^n \to 0$.*

Intuitivement, une suite extraite de u consiste à ne garder que certains termes en en "sautant" d'autres, et en s'interdisant de "revenir en arrière". Formellement, on la

Définition 7.23 (Suite extraite, extractrice)**.** Une **suite extraite** de u est une suite de la forme $(u_{\varphi(n)})_{n\in\mathbb{N}}$, où φ est strictement croissante de $\mathbb{N}$ dans $\mathbb{N}$. C'est en fait l'application $u \circ \varphi$ de $\mathbb{N}$ dans $\mathbb{R}$. On dit que φ est une **extractrice**. Intuitivement, les $\varphi(n)$ son les indices que l'on garde dans u.

Proposition 7.19. *Si φ est strictement croissante de $\mathbb{N}$ dans $\mathbb{N}$, alors c'est une injection et on a*

$$\forall n \in \mathbb{N},\ \varphi(n) \geqslant n$$

Théorème 7.14. *Si u tend vers $l \in \overline{\mathbb{R}}$, alors toute suite extraite de u tend vers l.*

Corollaire 7.6. *Pour montrer qu'une suite n'admet pas de limite, il suffit d'exhiber deux suites extraites qui tendent vers des limites différentes.*

Exemple 7.24. Considérer les suites extraites de rang pair et impair pour montrer que la suite de terme général $u_n = (-1)^n$ n'admet pas de limite. Plus généralement, on peut montrer e cette façon qui si $\alpha \leqslant -1$, alors la suite de terme général α^n n'admet pas de limite.

Réciproquement, on rencontre souvent le cas particulier suivant dans les exercices.

Proposition 7.20. *Supposons que $u_{2n} \to m$ et $u_{2n+1} \to l$, avec $l \in \overline{\mathbb{R}}$. Alors, $u_n \to l$.*

Théorème 7.15 (Caractérisation séquentielle de la densité)**.** *Une partie $X \subset \mathbb{R}$ est dense dans $\mathbb{R}$ si, et seulement si, tout réel est limite d'une suite $(x_n)_{n\in\mathbb{N}}$ à valeurs dans X.*

Corollaire 7.7. *Tout réel est limite d'une suite de rationnels.*

Exemple 7.25. $\mathbb{D}$ est dense dans $\mathbb{R}$. En effet, si on se donne $x \in \mathbb{R}$, il suffit de considérer la suite de ses approximations décimales (ce que l'on avait déjà fait pour la densité de $\mathbb{Q}$ dans $\mathbb{R}$).

Les deux théorèmes qui suivent constituent ce qu'on appelle la "**caractérisation séquentielle de la borne supérieure/inférieure**".

Théorème 7.16. *Soit X une partie non vide de $\mathbb{R}$. Alors il existe une suite $(x_n)_{n\in\mathbb{N}}$ à valeurs dans X qui tend vers $\sup(X) \in \mathbb{R} \cup \{+\infty\}$. On a le résultat symétrique avec la borne inférieure.*

Théorème 7.17. *Réciproquement, soit $M \in \mathbb{R} \cup \{+\infty\}$ un majorant de X, et supposons qu'il existe une suite $(x_n)_{n\in\mathbb{N}}$ à valeurs dans X telle que $x_n \to M$. Alors $M = \sup(X)$. de même avec un minorant et la borne inférieure.*

Exemple 7.26. On peut alors retrouver le résultat $\sup(\lambda A) = \lambda \sup(A)$ avec cette caractérisation séquentielle.

3 Théorèmes d'existence de limites

Théorème 7.18 (Théorème de la limite monotone)**.** *Toute suite réelle monotone admet une limite dans $\overline{\mathbb{R}}$.*

Remarque 7.39. De manière précise, la preuve nous montre qu'on a les résultats suivants.

- Toute suite réelle croissante admet une limite dans $\mathbb{R} \cup \{+\infty\}$.
  - Si $(u_n)_{n\in\mathbb{N}}$ est majorée par M, alors elle converge et sa limite vérifie $l \leqslant M$. Attention ! Même si tous les u_n sont strictement inférieurs à M, on doit garder l'inégalité au sens large pour la limite.
  - Si $(u_n)_{n\in\mathbb{N}}$ n'est pas majorée, alors elle tend vers $+\infty$.
- Toute suite réelle décroissante admet une limite dans $\mathbb{R} \cup \{-\infty\}$.
  - Si $(u_n)_{n\in\mathbb{N}}$ est minoré par m, alors elle converge et sa limite vérifie $l \geqslant m$.
  - Si $(u_n)_{n\in\mathbb{N}}$ n'est pas minorée, alors elle tend vers $-\infty$.

Remarque 7.40. On abrègera souvent le nom de ce théorème en "TLM" (NS).

Définition 7.24 (Suites adjacentes)**.** Deux suites sont dites adjacentes lorsque l'une est décroissante, l'autre est croissante et leur différence tend vers 0.

Lemme 7.3. *Si u et v sont adjacentes (avec u décroissante et v croissante), on a :*

$$\forall n \in \mathbb{N},\ u_n \geqslant v_n$$

Théorème 7.19. *Deux suivantes adjacentes convergent vers une même limite finie.*

Exemple 7.27. On montre aisément que $\forall n \in \mathbb{N},\ u_n \geqslant l \geqslant v_n$.

Théorème 7.20 (Théorème de Bolzano-Weierstrass)**.** *Soit $(u_n)_{n\in\mathbb{N}}$ une suite de réels bornée. Alors on peut trouver une suite extraite qui converge.*

Remarque 7.41. On abrègera souvent le nom de ce théorème en "TBW" (NS).

Définition 7.25 (Valeur d'adhérence)**.** On dit que l est une valeur d'adhérence de la suite $(u_n)_{n\in\mathbb{N}}$ lorsqu'il existe une suite extraite de $(u_n)_{n\in\mathbb{N}}$ qui converge vers l.

Remarque 7.42. Le théorème de Bolzano-Weierstrass peut donc être reformulé de la manière suivante : toute suite bornée admet une unique valeur d'adhérence.

Proposition 7.21 (HP)**.** *Réciproquement, tout suite réelle bornée admettant une unique valeur d'adhérence est convergente, et converge alors vers cette valeur d'adhérence.*

Démonstration. Il faut raisonner par l'absurde. Notons l l'unique valeur d'adhérence On peut alors fixer un $\varepsilon > 0$ tel qu'une infinité de termes de la suite soient au moins ε-loin de l. On construit une extractrice avec ces points. Or, la suite obtenue avec cette extractrice reste bornée, donc admet une valeur d'adhérence l' par le théorème de Bolzano-Weierstrass. Or, par PLIL, $l \neq l'$. Donc la suite ne possède pas une unique valeur d'adhérence. Absurde. Donc la suite converge vers l'unique valeur d'adhérence. □

Remarque 7.43 (Compacts, HP, spé)**.** Ce théorème est plus généralement valable dans tout **espace métrique compact** (*ie* un espace muni d'une notion de convergence et tel que toute suite à valeurs dans cet espace admette au moins une valeur d'adhérence). La démonstration est rigoureusement la même que celle présentée ci-dessus.

4 Quelques compléments

4.1 Suites complexes

Définition 7.26. Une **suite à valeurs complexes** est une application $(u_n)_{n\mathbb{N}}$ de $\mathbb{N}$ dans $\mathbb{C}$. La **partie réelle de** u est la suite réelle $(\mathrm{Re}(u_n))_{n\in\mathbb{N}}$. La **partie imaginaire de** u est la suite réelle $(\mathrm{Im}(u_n))_{n\in\mathbb{N}}$. La **suite conjuguée de** u est la suite complexe $(\overline{u_n})_{n\in\mathbb{N}}$. La **suite des modules de** u est la suite réelle $(|u_n|)_{n\in\mathbb{N}}$.

Définition 7.27. Une suite complexe est dite **bornée** lorsque la suite de ses modules est majorée. Si la suite est réelle, on retrouve bien l'ancienne définition.

Exemple 7.28. La suite définie par $u_n = \exp(in\pi/6)$ est bornée.

Définition 7.28 (Convergence)**.** La suite $(u_n)_{n\in\mathbb{N}}$ est **convergente de limite** $l \in \mathbb{C}$ lorsque

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N},\ \forall n \in \mathbb{N},\ n \geqslant N \implies |u_n - l| \leqslant \varepsilon$$

Cette fois, il ne s'agit plus d'une valeur absolue mais d'un module. Cela dit, puisque le module est un prolongement de la valeur absolue à $\mathbb{C}$, la définition ci-dessus coïncide bien dans le cas des suites réelles.

Remarque 7.44. On montre comme dans $\mathbb{R}$ que si la limite existe, alors elle est unique. On raisonne par l'absurde, et à la place des valeurs absolues, on utilise des modules.

Remarque 7.45. Les principes de troncature et de décalage d'indice fonctionnent encore.

Remarque 7.46. On a $u_n \xrightarrow[n\to+\infty]{\mathbb{C}} l$ ssi $u_n - l \xrightarrow[n\to+\infty]{\mathbb{C}} 0$ ssi $|u_n - l| \xrightarrow[n\to+\infty]{\mathbb{R}} 0$. En particulier, $u_n \xrightarrow[n\to+\infty]{\mathbb{C}} 0$ ssi $|u_n| \xrightarrow[n\to+\infty]{\mathbb{C}} 0$.

Exemple 7.29. Soit $(u_n)_{n\in\mathbb{N}}$ une suite complexe et $(v_n)_{n\in\mathbb{N}}$ une suite réelle. Supposons que $\forall n \in \mathbb{N},\ |u_n| \leqslant v_n$. Si $v_n \xrightarrow[n\to+\infty]{\mathbb{R}} 0$, alors $u_n \xrightarrow[n\to+\infty]{\mathbb{C}} 0$.

Théorème 7.21 (Théorème-passerelle)**.** *Soit $(u_n)_{n\in\mathbb{N}}$ une suite complexe. On a :*

$$u_n \xrightarrow[n\to+\infty]{\mathbb{C}} l = a + ib \in \mathbb{C} \iff \begin{cases} Re(u_n) \xrightarrow[n\to+\infty]{\mathbb{R}} a \\ Im(u_n) \xrightarrow[n\to+\infty]{\mathbb{R}} b \end{cases}$$

On vérifie alors que de nombreux résultats vus dans le cas réels restent vrais. Les voici dans l'ordre.

Proposition 7.22. *Toute suite complexe convergente est bornée.*

Exemple 7.30. Si $|u_n| \xrightarrow[n\to+\infty]{\mathbb{R}} +\infty$, alors $(u_n)_{n\in\mathbb{N}}$ diverge dans $\mathbb{C}$.

Théorème 7.22 (Opérations algébriques)**.** *Les opérations algébriques usuelles sur les limites sont valables : combinaison linéaire, produit, passage à l'inverse (pour une limite non nulle), module, conjugaison.*

Proposition 7.23. *Si une suite complexe converge vers $l \in \mathbb{C}$, alors toute suite extraite converge vers l. Réciproquement, si $u_{2n} \xrightarrow[n\to+\infty]{} l$ et $u_{2n+1} \xrightarrow[n\to+\infty]{} l$, alors $u_n \xrightarrow[n\to+\infty]{} l$.*

Théorème 7.23 (Théorème de Bolzano-Weierstrass)**.** *Le théorème de Bolzano-Weierstrass reste vrai : toute suite complexe bornée admet une suite extraite convergente.*

4.2 Moyenne de Cesàro, HP mais à connaître par cœur

Cette partie n'est pas *stricto sensu* au programme ; mais elle est tellement classique que nous la donnons comme résultat de cours.

Soit une suite complexe $(u_n)_{n\geqslant 1}$. Au lieu de s'intéresser à la convergence de u, on va s'intéresser à la convergence des moyennes arithmétiques successives :

$$\forall n \in \mathbb{N}^\star,\ \mu_n = \frac{u_1 + \ldots + u_n}{n}$$

Intuitivement, on se dit que si u_n tend vers l, alors à la longue, les moyennes μ_n vont "gommer" les écarts initiaux de u_n autour de l, et "lisser les irrégularités", de sorte que μ_n tendra aussi vers l.

Formalisons maintenant cela. Tout d'abord, on commence par démontrer un cas particulier. C'est l'objet du

Lemme 7.4. *Si* $u_n \xrightarrow[n\to+\infty]{} 0$*, alors* $\mu_n \xrightarrow[n\to+\infty]{} 0$*.*

Démonstration. Donnons-nous $\varepsilon > 0$, et soit N un rang à partir duquel on ait $|u_n| \leqslant \dfrac{\varepsilon}{2}$. Posons ensuite $M = |u_1 + \ldots + u_N|$, puis donnons-nous un rang N' à partir duquel on ait $\dfrac{M}{n} \leqslant \dfrac{\varepsilon}{2}$. Alors, pour $n \geqslant \max(N, N')$, on a

$$\begin{aligned} |\mu_n| &\leqslant \frac{M + |u_{N+1} + \ldots + u_n}{n} \\ &\leqslant \frac{M}{n} + \frac{n-N}{n}\frac{\varepsilon}{2} \\ &\leqslant \frac{\varepsilon}{2} + 1 \times \frac{\varepsilon}{2} \\ &\leqslant \varepsilon \end{aligned}$$

La preuve est achevée. □

Théorème 7.24. *Si* $u_n \xrightarrow[n\to+\infty]{} l$*, alors* $\mu_n \xrightarrow[n\to+\infty]{} l$*.*

Démonstration. Il suffit d'appliquer le lemme précédent à $u_n - l$ qui tend alors vers 0. On obtient que

$$\frac{(u_1 - l) + \ldots + (u_n - l)}{n} = \mu_n - l$$

tend vers 0. On conclut par somme de limites. □

Remarque 7.47. En revanche, la réciproque est fausse. Comme contre-exemple, prendre par exemple la suite alternée définie par $u_n = (-1)^n$. On a $\mu_n \to 0$ comme on s'en convainc en extrayant les sous-suites de rangs pairs et impairs. Pourtant, on n'a pas $u_n \to 0$, puisque par exemple $u_{2n} \to 1$.

A partir de maintenant, on suppose que (u_n) est réelle. On peut se demander si le résultat reste vrai dans le cas où la limite vaut $\pm\infty$. La réponse est oui, comme le montre le théorème suivant.

Théorème 7.25. *Si* $u_n \xrightarrow[n\to+\infty]{} +\infty$*, alors* $\mu_n \xrightarrow[n\to+\infty]{} +\infty$*. De même avec* $-\infty$*.*

Démonstration. Quitte à changer u en $-u$, démontrons sans perte de généralité le résultat sur $+\infty$. Soit $M \in \mathbb{R}_+$. Au-delà d'un certain rang N on aura $u_n \geqslant 2M + 2$. Posons alors $S = u_1 + \ldots + u_N$, puis donnons-un rang N' à partir duquel on ait $\dfrac{S}{n} \geqslant -1$. Enfin, donnons-nous un rang N'' à partir duquel on ait $1 - \dfrac{N}{n} \geqslant \dfrac{1}{2}$. Pour $n \geqslant \max(N, N', N'')$, on peut alors écrire :

$$\begin{aligned} \mu_n &\geqslant \frac{S}{n} + \frac{1}{n}\sum_{k=N+1}^{n}(2M+2) \\ &\geqslant -1 + \frac{n-N}{n} \times (2M+2) \\ &\geqslant -1 + \frac{1}{2} \times (2M+2) \\ &\geqslant M \end{aligned}$$

puisque $2M + 2 \geqslant 0$. La preuve est achevée. □

Remarque 7.48. Là aussi, on pourra montrer que la réciproque est fausse.

4.3 Suites particulières

Suites arithmético-géométriques

On sait calculer le terme général d'une suite arithmétique, d'une suite géométrique, mais qu'en est-il d'un "mélange des deux" ?

On considère la suite définie par

$$\begin{cases} u_0 = x \\ \forall n \in \mathbb{N},\ u_{n+1} = au_n + b \end{cases}$$

où x, a et b sont des complexes. On impose $a \neq 1$, car sinon on a affaire à ue suite arithmétique que l'on connaît déjà. On cherche alors une solution constante q de la relation de récurrence. Or

$$\forall q \in \mathbb{C},\ q = aq + b \iff q = \frac{b}{1-a}$$

On parle de **point fixe**. Ensuite, on considère la suite auxiliaire v de terme général $v_n = u_n - q$. Elle vérifie $v_{n+1} = av_n$, si bien qu'elle est géométrique. On sait alors calculer son terme général, puis on se ramène à celui de u en ajoutant q.

Suites récurrentes linéaires d'ordre 2

Théorème 7.26 (Cas complexe)**.** *Soit $(a, b) \in \mathbb{C} \times \mathbb{C}^\star$, et cherchons l'ensemble des suites complexes qui vérifient la relation de récurrence double*

$$\forall n \in \mathbb{N},\ u_{n+2} + au_{n+1} + bu_n = 0$$

Tout d'abord, on cherche dans $\mathbb{C}$ les racines de ce qu'on appelle le "polynôme caractéristique" : $X^2 + aX + b$.

- *Si ce polynôme possède deux racines distinctes $r_1 \neq r_2$, les suites cherchées sont exactement celles de la forme*
$$u_n = \lambda r_1^n + \mu r_2^n$$
avec $(\lambda, \mu) \in \mathbb{C}^2$.
- *Si ce polynôme possède une racine double r, les suites cherchées sont exactement celles de la forme*
$$u_n = (\lambda + \mu n) r^n$$
avec $(\lambda, \mu) \in \mathbb{C}^2$.

Théorème 7.27 (Cas réel)**.** *Soit $(a, b) \in \mathbb{R} \times \mathbb{R}^\star$, et cherchons l'ensemble des suites complexes qui vérifient la relation de récurrence double*

$$\forall n \in \mathbb{N},\ u_{n+2} + au_{n+1} + bu_n = 0$$

Tout d'abord, on cherche dans $\mathbb{C}$ les racines de ce qu'on appelle le "polynôme caractéristique" : $X^2 + aX + b$.

- *Si ce polynôme possède deux racines réelles distinctes $r_1 \neq r_2$, les suites cherchées sont exactement celles de la forme*
$$u_n = \lambda r_1^n + \mu r_2^n$$
avec $(\lambda, \mu) \in \mathbb{R}^2$.
- *Si ce polynôme possède une racine double r, les suites cherchées sont exactement celles de la forme*
$$u_n = (\lambda + \mu n) r^n$$
avec $(\lambda, \mu) \in \mathbb{R}^2$.
- *Si ne possède pas de racine réelle, on écrit les deux racines complexes conjuguées sous la forme $\rho \exp(\pm i\theta)$, puis les suites cherchées sont exactement celles de la forme*
$$u_n = \rho^n \left(\lambda \cos(n\theta) + \mu \sin(n\theta)\right)$$
avec $(\lambda, \mu) \in \mathbb{R}^2$.

Remarque 7.49. A chaque fois, les constantes λ et μ pourront se déterminer à partir des premiers termes, sortes de "conditions initiales" (kssssss).

Exemple 7.31. En appliquant ceci à la suite de Fibonacci définie par
$$\begin{cases} F_0 = 0 \\ F_1 = 1 \\ \forall n \in \mathbb{N},\ F_{n+2} = F_{n+1} + F_n \end{cases}$$
on obtient
$$\forall n \in \mathbb{N},\ F_n = \frac{1}{\sqrt{5}} \left[\left(\frac{1+\sqrt{5}}{2} \right)^n - \left(\frac{1-\sqrt{5}}{2} \right)^n \right]$$

Suites définies par une fonction

Considérons une fonction f **continue** à valeurs réelles et un intervalle $I \subset \mathbb{R}$. On rappelle que I est dit **stable** par f lorsque $f(I) \subset I$. Dans ce cas, on peut définir une suite de la manière suivante :
$$\begin{cases} u_0 \in I \\ \forall n \in \mathbb{N}, u_{n+1} = f(u_n) \end{cases}$$

Proposition 7.24. *Voici la méthode générale pour étudier une suite du type $u_{n+1} = f(u_n)$, avec f continue.*

- *On cherche un intervalle I stable par f contenant u_0.*
- *On étudie les points fixes de f.*
- *On étudie la monotonie éventuelle de $(u_n)_{n_i n\mathbb{N}}$.*
- *On étudie la limite éventuelle de $(u_n)_{n\in\mathbb{N}}$*

Chacun des points mérite d'être commenté séparément.

Recherche d'un intervalle stable D'une manière générale, on aura toujours intérêt à choisir un intervalle le plus petit possible autour de u_0, ainsi l'étude de la suite sera plus rapide. Si f n'est pas continue sur l'ensemble de son domaine de définition, il faudra choisir I de telle sorte que $f_{|I}$ soit continue.
Dans la recherche de I, on aura bien souvent intérêt à étudier les variations de f et à tracer l'allure de sa courbe représentative.

Recherche des points fixes Pour commencer, on donne le

Théorème 7.28. *Si $u_n \to l$ et si f est continue en l, alors $f(l) = l$.*

Ce théorème est souvent utilisé dans le cadre d'un raisonnement par analyse-synthèse. Si u admet une limite l en laquelle f est définie, l sera nécessairement un point fixe de f par continuité. Notons qu'il est inutile de rechercher les points fixes "trop loin de I". En pratique, notre théorème de passage à la limite dans les inégalités nous montre que si u converge vers l, alors l appartient à I ou est une borne de I (en effet, les inégalités d'encadrement deviennent larges en passant à la limite). Donc, **la plupart du temps, il vaudra mieux avoir défini f sur un intervalle fermé**.
En tout cas, on comprend l'intérêt de choisir x petit : cela permet de limiter le nombre de points fixes éventuels. En pratique, on étudiera souvent $g = f - \mathrm{id}$.

Étude de la monotonie

- Notons que $\forall n \in \mathbb{N}, u_n n + 1 - u_n = g(u_n)$. Ainsi, on pourra se ramener à étudier le signe de g. bien souvent, l'étude précédente sur les zéros de g fournira en même temps le signe sur le reste du domaine à l'aide d'un tableau de variations.
- Parfois, on peut gagner du temps en remarquant que f est croissante ou décroissante. Dans ce cas, il suffit de comparer u_0 et u_1, puis une récurrence immédiate montre que u est monotone.

Étude de la limite de u Une fois qu'on a trouvé une limite éventuelle l, pour montrer que $u_n \to l$, il existe différente méthodes. Le plus souvent, il s'agira d'utiliser le théorème de la limite monotone. Pour cela, il faudra que l'étude de la monotonie de u ait été concluante. On pourra aussi utiliser l'inégalité des accroissements finis.

Exemple 7.32. En appliquant cette méthode à la suite définie par

$$\begin{cases} u_0 \in [0,1] \\ \forall n \in \mathbb{N},\ u_{n+1} = \sin(u_n) \end{cases}$$

on montre que $u_n \to 0$.

5 Compléments : construction de $\mathbb{R}$ et complétude, HP

5.1 Construction de $\mathbb{R}$

Nous nous proposons ici de construire $\mathbb{R}$ comme une extension de $\mathbb{Q}$, à l'aide des coupures de Dedekind, du nom de Richard Dedekind, un très grand mathématicien allemand du XIX^e siècle.

Construction de l'ensemble des réels

Définition 7.29. Une **section commençante ouverte** (ou SCO) de $\mathbb{Q}$ est une partie $A \subset \mathbb{Q}$ telle que

- $A \neq \emptyset$ et $A \neq \mathbb{Q}$
- $\forall a \in A, \forall a' \in \mathbb{Q}, a' < a \Rightarrow a' \in A$
- A ne possède pas de maximum.

L'ensemble des SCO est noté $\mathbb{R}$, et à partir de maintenant ses éléments seront plutôt appelés les **réels**.

Remarque 7.50. À la fin de cette partie, il s'avèrera en fait que les SCO sont les ensembles de la forme $]-\infty, A[\cap \mathbb{Q}$ avec $A \in \mathbb{R}$. Cela peut aider à les visualiser.

Remarque 7.51. Pour éviter au maximum les confusions, on réservera les majuscules aux éléments de $\mathbb{R}$ et les minuscules aux éléments de $\mathbb{Q}$.

Proposition 7.25. *L'application $r \mapsto A_r = \{r' \in \mathbb{Q} \mid r' < r\}$ va de $\mathbb{Q}$ dans $\mathbb{R}$ et elle est injective. Ainsi, elle permet d'identifier $\mathbb{Q}$ à un sous-ensemble de $\mathbb{R}$.*

Démonstration. Soit $r \in \mathbb{Q}$. Montrons déjà que $A_r \in \mathbb{R}$:

- $r - 1 \in A_r$ donc $A_r \neq \emptyset$ et $r \notin A_r$ donc $A_r \neq \mathbb{Q}$
- Soit $s \in A_r$ et $s' < s$. Alors $s' < r$ donc $s' \in A_r$
- Supposons que A_r admette un plus grand élément s. Alors on aurait $s < r$, puis $\frac{r+s}{2} \in A_r$ avec $\frac{r+s}{2} > s$, contradiction.

Soit maintenant $r' \in \mathbb{Q}$ tel que $A_r = A_{r'}$, et montrons $r = r'$. Supposons par l'absurde que $r < r'$ alors on aurait $r \in A_{r'} = A_r$, contradiction. Donc $r \geqslant r'$. En échangeant le rôle de r et r', on obtient $r' \geqslant r$. Donc $r = r'$. □

Définition 7.30. On définit une relation sur $\mathbb{R}$ par $A \leqslant B \Leftrightarrow A \subset B$.

Proposition 7.26. *$\leqslant$ est une relation d'ordre totale qui prolonge $\leqslant$ définie sur $\mathbb{Q}$.*

Démonstration. On sait déjà que $\leqslant$ est une relation d'ordre (partielle) sur les parties de n'importe quel ensemble. Ici, il faut vérifier qu'elle est totale. Soit $(A, B) \in \mathbb{R}^2$ et supposons qu'on n'a pas $A \subset B$. Montrons que $B \subset A$. Pour cela, prenons $a \in A \backslash B$. Soit $b \in B$ quelconque, le but est de montrer que $b \in A$. Or on ne peut pas avoir $b > a$ sans quoi on aurait $a \in B$. On en déduit $b \leqslant a$. Ensuite, si $b = a$ alors c'est terminé. Si $b < a$, alors on a aussi $b \in A$.

Pour finir, soit $(r, r') \in \mathbb{Q}^2$ et montrons que $r \leqslant r' \Leftrightarrow A_r \subset A_{r'}$. Le sens direct est trivial. Réciproquement, si $A_r \subset A_{r'}$, supposons $r > r'$, alors on a $r' \in A_r$ donc $r' \in A_{r'}$, absurde. Donc $r \leqslant r'$. □

Voici à présent un petit lemme dont nous ferons régulièrement usage.

Lemme 7.5. *Soit $A > 0$, alors A contient un rationnel strictement positif.*

Démonstration. On n'a pas $A \leqslant 0$, c'est-à-dire que A n'est pas inclus dans A_0. Autrement dit, A contient un rationnel $r \geqslant 0$. Si $r = 0$, il suffit de remarquer que par hypothèse, A ne possède pas de maximum. □

Construction de l'addition

Proposition 7.27. *L'opération* $A + B = \{a + b \mid (a, b) \in A \times B\}$ *est bien définie, et elle prolonge l'addition sur* $\mathbb{Q}$.

Démonstration. Soit $(A, B) \in \mathbb{R}^2$. Commençons par vérifier que $A + B$ est bien une SCO.

- $A \neq \emptyset$ donc il existe $a \in A$. De même il existe $b \in B$. Finalement, $A + B \neq \emptyset$. De plus, soit $x \in \mathbb{Q}\backslash A$ et $y \in \mathbb{Q}\backslash B$. Quel que soit $a \in A$, on ne peut pas avoir $a > x$ sans quoi on aurait $x \in A$. De même on ne peut pas avoir $a = x$, donc on a $a < x$. De même, tout $b \in B$ vérifie $b < y$. Finalement, tout $a + b \in A + B$ vérifie $a + b < x + y$, si bien que $A + B \neq \mathbb{Q}$
- Prenons $a + b \in A + B$ et $c < a + b$. En notant $\varepsilon = a + b - c > 0$, on a

$$c = \underbrace{\left(a - \frac{\varepsilon}{2}\right)}_{\in A} + \underbrace{\left(b - \frac{\varepsilon}{2}\right)}_{\in B} \in A + B$$

- Enfin, montrons que $A + B$ ne possède pas de maximum ; si on se donne $a + b \in A + B$ alors il suffit de choisir $a' > a$ dans A, $b' > b$ dans B, et de considérer $a' + b'$.

Maintenant, soit $(r, r') \in \mathbb{Q}^2$ et montrons que $A_r + A_{r'} = A_{r+r'}$.

- D'une part, si on se donne $r' + s' \in A_r + A_{s'}$ alors $r' < r$ et $s' < s$ donc $r' + s' < r + s$.
- Réciproquement, si on se donne $t \in A_{r+s}$, alors en notant $\varepsilon = r + s - t > 0$ on écrit de la même manière que tout à l'heure

$$t = \underbrace{\left(r - \frac{\varepsilon}{2}\right)}_{\in A_r} + \underbrace{\left(s - \frac{\varepsilon}{2}\right)}_{\in A_s} \in A_r + A_{s'}$$

Ainsi la preuve est achevée. □

Avant de continuer, nous avons besoin de démontrer le

Lemme 7.6. *Soit* $\frac{p}{q} \in \mathbb{Q}$. *Alors il existe* $n \in \mathbb{N}$ *tel que* $n > \frac{p}{q}$.

Démonstration. Supposons sans perte de généralité que $q \in \mathbb{N}^*$. Si $p \leqslant 0$ alors $n = 1$ convient, et sinon $n = 2p$ convient. En effet, on a $q \geqslant 1$ d'où on tire

$$2p > p \geqslant \frac{p}{q}$$

ce qui achève la preuve. □

Théorème 7.29. $(\mathbb{R}, +)$ *est un groupe commutatif, et l'addition est compatible avec* $\leqslant$ *au même sens que* $\mathbb{Q}$.

Démonstration. Vérifions les différents axiomes d'un groupe commutatif.

- Associativité et commutativité : elles s'héritent directement de $\mathbb{Q}$.
- Élément neutre : c'est $0 := A_0$
- Opposé : soit $A \in \mathbb{R}$
  - Si A est de la forme A_r alors on sait que $A_r + A_{-r} = A_0 = 0$.

– Si A ne peut pas s'écrire sous la forme A_r montrons que $B = \{-b' | b' \notin A\}$ convient. Pour commencer, montrons que $B \in \mathbb{R}$.

  * Comme $A \neq \emptyset$ on peut choisir $a \in A$ et on aura $-a \notin B$ Donc $B \neq \mathbb{Q}$.
    Comme $A \neq \mathbb{Q}$ on peut choisir $b' \in \mathbb{Q}\backslash A$ et on aura $-y \in B$. Donc $B \neq \emptyset$.
  * Soit $b \in Bc < b$ Ecrivons $b = -b'$ avec $b' \notin Ac = -c'$. Alors $c' > b'$ donc $c' \notin A$ sans quoi on aurait $b' \in A$ Donc $c \in B$.
  * Soit $b \in B$ et écrivons $b = -b'$. Par hypothèse on a $A \neq A_{b'}$. Supposons $A > A_{b'}$ c'est-à-dire $A \supseteq A_{b'}$ alors on pourrait trouver un $a \in A\backslash A_{b'}$ On aurait donc $b' \leq a$ d'où on déduirait $b' \in A$ dans tous les cas, contradiction. Ainsi on a $A < A_{b'}$ Autrement dit, il existe un $c' < b'$ avec $c' \notin A$. On a alors $-c' > -b'$ avec $-c' \in B$. Donc B ne possède pas de maximum.

  Ensuite, montrons que $A + B = A_0$.

  * Soit $c \in A + B$, écrivons $c = a - b'$ avec $a \in A$ et $b' \notin A$. On a donc $b' > a$ (sans quoi on aurait $b' \in A$ dans tous les cas), d'où on déduit $c \in A_0$.
  * Réciproquement soit $c < 0$, et supposons par l'absurde que $\forall a \in A,\ a - c \in A$. Si on choisit $a_0 \in A$ fixé, une récurrence immédiate montre qu'on aurait $a_0 - nc \in A$ pour tout $n \in \mathbb{N}$. Or si on prend $x \in \mathbb{Q}\backslash A$, on peut trouver $n \in \mathbb{N}$ tel que $n > \dfrac{a_0 - x}{c}$ par le lemme. On en déduirait $nc < a_0 - x$ d'où $x < a_0 - nc$ et finalement $x \in A$, contradiction. Donc on peut trouver $a \in A$ vérifiant $b' = a - c \notin A$. Il suffit alors de poser $b = -b' \in B$ et de remarquer que $c = a + b$.

Pour finir, soit $(A, B, C) \in \mathbb{R}^3$. Il faut vérifier la compatibilité de $\leq$ avec $+$ au sens suivant :

$$A \leq B \Rightarrow A + C \leq B + C$$

En fait, c'est à peu près immédiat. Si on se donne $a + c \in A + C$, alors *a fortiori* on $a \in B$, d'où le résultat. □

Construction de la multiplication

Nous en venons maintenant à la définition de la multiplication, qui est un peu plus compliquée que celle de l'addition.

Proposition 7.28. *L'opération* $A \cdot B = \{c \in \mathbb{Q} \mid \exists (a, b) \in A \times B, a \geq 0,\ b \geq 0, c \leq ab\}$ *est bien définie sur* $\mathbb{R}_+^*$ *et elle prolonge la multiplication sur* $\mathbb{Q}_+^*$.

Démonstration. Soit $A, B \in \mathbb{Q}_+^*$ et commençons par vérifier que $A \cdot B \in \mathbb{R}_+^*$

- Comme on n'a pas $A \subset A_0$, c'est qu'on peut trouver $a \in A$ avec $a \geqslant 0$. De même, on peut trouver $b \in B$ avec $b \geqslant 0$. Alors $ab \in A \cdot B$ donc $A \cdot B \neq \emptyset$. De plus, considérons $x \in \mathbb{Q}\backslash A$ et $y \in \mathbb{Q}\backslash B$. Rappelle que $a \in A$ vérifie $a < x$ et tout $b \in B$ vérifie $b < y$. Donc si $a, b \geqslant 0$ on a $xy > ay \geqslant ab$. Ainsi $xy \notin A \cdot B$ et $A \cdot B \neq \mathbb{Q}$.
- Soit $c \in A \cdot B$ et $c' < c$. Écrivons $c \leqslant ab$, avec $a \geqslant 0$ et $b \geqslant 0$. Alors *a fortiori* on a bien $c' \leqslant ab$ et $(a, b) \in A \times B$.
- Enfin, soit $c \in A \cdot B$ et écrivons $c \leqslant ab$ de la même manière. On peut trouver $a' > a$ dans A et $b' > b$ dans B, si bien que $a', b' > 0$ puis $a'b' > a'b \geqslant ab$. Donc $a'b' > c$ avec $a'b' \in A \cdot B$.
- Le point précédent montre que $A \cdot B$ n'est pas inclus dans A_0, d'où on tire $A \cdot B > 0$.

Maintenant, soit $r, s \in \mathbb{Q}_+^*$ et montrons que $A_r \cdot A_s = A_{rs}$.

- D'une part, soit $t \in A_r \cdot A_s$. Alors on peut écrire $t \leqslant r's'$, avec $(r', s') \in A_r \times A_s$ et $r', s' \geqslant 0$. Comme précédemment, on en déduit $r's' \leqslant r's < rs$ donc $t \in A_{rs}$.
- Réciproquement, soit $t \in A_{rs}$, alors $t < rs$. Introduisons alors $\varepsilon = \min(r, s, \frac{rs-t}{r+s})$ de telle sorte que $0 < \varepsilon \leqslant r, s, \frac{rs-t}{r+s}$. On a alors $(r-\varepsilon)(s-\varepsilon) = rs - \varepsilon(r+s) + \varepsilon^2 > rs - \varepsilon(r+s) \geqslant t$, d'où $t \in A_\varepsilon \cdot A_s$.

Ainsi la preuve est achevée. □

Définition 7.31. On prolonge $A \cdot B$ à tout entier de la manière suivante :

- Si $A = 0$ ou $B = 0$, on pose $A \cdot B = 0$.
- Si $A < 0$ et $B > 0$, on pose $A \cdot B = -[(-A) \cdot B]$.
- Si $A > 0$ et $B < 0$, on pose $A \cdot B = -[A \cdot (-B)]$.
- Si $A < 0$ et $B < 0$, on pose $A \cdot B = (-A) \cdot (-B)$.

Proposition 7.29. *Ce prolongement coïncide bien avec la multiplication sur tout entier.*

Démonstration. C'est immédiat en remarquant que $\forall r \in \mathbb{Q}, -A_r = A_{-r}$. À titre d'exemple, soit $r < 0$ et $s > 0$, On a $A_r \cdot A_s = -[(-A_r) \cdot A_s] = -[A_{-r} \cdot A_s] = -[A_{(-r)s}] = -A_{-rs} = A_{rs}$ et ainsi de suite. □

Ce petit lemme nous aidera pour les vérifications sur l'associativité et la distributivité.

Lemme 7.7. *On a* $\forall (A, B) \in \mathbb{R}^2, (-A) \cdot B = A \cdot (-B) = -(A \cdot B)$.

Démonstration. Soit $(A, B) \in \mathbb{R}^2$. Si A ou B est nul, le résultat est immédiat. Sinon, on a quatre cas.

- Si $A > 0$ et $B > 0$, alors $(-A) \cdot B = -(A \cdot B)$ par définition, et de même avec le deuxième terme.
- Si $A < 0$ et $B < 0$, alors $(-A) \cdot B = -[(-A) \cdot (-B)] = -(A \cdot B)$, et de même avec le deuxième terme.
- Si $A > 0$ et $B < 0$ alors $(-A) \cdot B = A \cdot (-B)$ par définition, et $A \cdot B = -[A \cdot (-B)]$ d'où $-(A \cdot B) = A \cdot (-B)$.
- Si $A < 0$ et $B > 0$ la preuve est symétrique.

Ainsi la preuve est achevée. □

Là aussi, nous avons besoin d'un autre lemme préliminaire, symétrique du lemme pour le groupe pour $+$.

Lemme 7.8. *Soit* $\alpha > 1$ *dans* $\mathbb{R}$ *et* $m \in \mathbb{Q}$. *Alors on peut trouver* $n \in \mathbb{N}$ *tel que* $\alpha^n > m$.

Démonstration. Pour $n \in \mathbb{N}^*$, on a

$$\alpha^n - 1 = (\alpha - 1)(\alpha^{n-1} + ... + 1) \geqslant n(\alpha - 1)$$

d'où $\alpha^n \geqslant 1 + n(\alpha - 1)$. Il suffit ensuite de prendre $n \in \mathbb{N}^*$ tel que $n > \dfrac{m-1}{\alpha - 1}$ à un lemme précédent □

Théorème 7.30. $(\mathbb{R}, +, \cdot)$ *est un corps.*

Démonstration. Nous avons déjà vérifié que c'était un groupe commutatif pour l'addition, et ce n'est pas l'anneau trivial car il contient $\mathbb{Q}$. Vérifions donc les axiomes sur la multiplication.

- **Associativité** : Soit $(A, B, C) \in \mathbb{R}^3$. Il faut montrer que $(A \cdot B) \cdot C = A \cdot (B \cdot C)$. Si l'un des trois réels est nul c'est immédiat. Sinon, sans perte de généralité on peut se ramener au cas où $A, B, C > 0$, les autres cas s'en déduisant immédiatement à partir du lemme.
  - Soit $r \in (A \cdot B) \cdot C$, et écrivons $r \leqslant sc$ et $s \leqslant ab$ avec les conditions habituelles. On a alors $r \leqslant a(bc)$ avec $bc \in B \cdot C$ et $a \in A$, d'où $r \in A \cdot (B \cdot C)$.
  - L'inclusion réciproque se démontre de la même manière.
- **Commutativité** : c'est une conséquence de la définition, qui est symétrique en A et B.
- **Élément neutre** : c'est $1 = A_1$. En effet, soit $A \in \mathbb{R}$ et montrons que $A \cdot A_1 = A$ (par commutativité, ce sera suffisant). Si $A = 0$ le résultat est évident, et si $A < 0$ on se ramène à $A > 0$ par le lemme 8.69. Supposons donc $A > 0$ et commençons par montrer $A \cdot A_1 \subset A$. Soit $r \in A \cdot A_1$ et écrivons $r \leqslant aa_1$ avec les conditions habituelles. Comme $a \geqslant 0$ et $a_1 < 1$ nous avons $r < a$ donc $r \in A$. Réciproquement, soit $a \in A$. Si on choisit $a' > a$ dans A (puisque A ne possède pas de maximum), alors $a' > 0$ puis $\dfrac{a}{a'} \in A_1$ avec $\dfrac{a}{a'} \geqslant 0$. Il ne reste plus qu'à écrire $a = a' \cdot \frac{a}{a'} \in A \cdot A_1$.
- Passage à l'inverse : soit $A \neq 0$. Grâce au lemme précédent, supposons même $A > 0$.
  - Si A est de la forme A_r alors $r > 0$ puis on sait que $A_r \cdot A_{1/r} = A_1 = 1$.
  - Si A ne peut pas s'écrire sous la forme A_{r_i} montrons que $B = \mathbb{R}_- \cup \{\dfrac{1}{b'} \mid b' \notin A\}$ convient. Pour commencer, montrons que B est bien définie et que $B \in \mathbb{R}$.
    * A n'est pas égal à tout entier, donc il existe des $b' \in \mathbb{Q}\backslash A$. Par ailleurs A contient un rationnel strictement positif par un lemme précédent, d'où on déduit que tous les $b' \in \mathbb{Q}\backslash A$ sont strictement positifs. Donc $\{\dfrac{1}{b'} | b' \notin A\}$ est bien défini, non vide et inclus dans $\mathbb{Q}_+^*$. Conclusion : B est bien défini et possède au moins un rationnel strictement positif.
    * Comme $\mathbb{R}_- \subset B$ on a $B \neq \emptyset$ par ailleurs on peut trouver $\alpha > 0$ dans A, d'où on déduit $\frac{1}{a} \notin B$ puis $B \neq \mathbb{Q}$.
    * Soit $b \in B, c < b, c \leqslant 0$ alors $c \in B$ sinon c'est qu'on a $0 < c < b$. Écrivons $c = \dfrac{1}{c'}$ et $b = \dfrac{1}{b'}$ avec $0 < b' < c'$. $b \notin A$. On en déduit $c' \notin A$, d'où $c \in B$.
    * Soit $b \in B$ et trouvons un $c > b$ dans B. Si $b \leqslant 0$ nous avons vu que B contenait un rationnel strictement positif. Sinon, écrivons $b = \dfrac{1}{b'}$ avec $b' \notin A$ et $b' > 0$. Par hypothèse on a $A \neq A_{b'}$. Supposons $A > A_{b'}$ -à-dire $A \supseteq A_{b'}$. Alors on pourrait trouver un $a \in A\backslash A_{b'}$ puis on aurait $a \geqslant b'$. On en déduirait donc $b' \in A$ dans tous les cas, contradiction. Ainsi on a $A < A_{b'}$. Autrement dit, il existe un $c' < b'$ avec $c' \notin A$. On en déduit $c' > 0$ puis $\dfrac{1}{c'} > \dfrac{1}{b'}$. Il suffit alors de poser $c := \dfrac{1}{c'}$.

    Montrons maintenant que $A \cdot B = A_1$

* Soit $c \in A \cdot B$ écrivons $c \leqslant ab$ avec les conditions habituelles et montrons que $ab < 1$. Si $b = 0$ le résultat est immédiat, sinon écrivons $b = \frac{1}{b'}$ avec $b' > 0$ et $b' \notin A$. On a nécessairement $a < b'$, d'où le résultat.
* Réciproquement soit $c \in A_1$, et montrons que $c \in A \cdot B$. Si $c \leqslant 0$ il suffit de remarquer que $0 \in A \cap B$, donc considérons à de maintenant que $c > 0$. Supposons que pour tout $a > 0$ dans A, on ait $\frac{a}{c} \in A$. Si nous choisissons $a_0 > 0$ fixé dans A (ce qui est possible par un lemme précédent), une récurrence immédiate montre qu'on aurait $\frac{a_0}{c^n} \in A$ pour tout $n \in \mathbb{N}$. Or si on prend $x \in \mathbb{Q}\backslash A$, on peut trouver $n \in \mathbb{N}$ tel que $(\frac{1}{c})^n > \frac{x}{a_0}$ un lemme. On en déduirait $x < \frac{a_0}{c^n}$ et finalement $x \in A$ contradiction. Donc il existe $a > 0$ dans A tel que $b' = \frac{a}{c} \notin A$. Si on pose $b = \frac{1}{b'} \in B$, on obtient alors $c = ab \in A \cdot B$.

- **Distributivité** : Par commutativité il suffit de la vérifier d'un côté seulement. Soit $(A, B, C) \in \mathbb{R}^3$ et montrons par exemple que $A \cdot (B + C) = A \cdot B + A \cdot C$. Là aussi, si A, B ou C est nul le résultat est immédiat ; sinon, quitte à changer A en $-A$ et (B, C) en $(-B, -C)$, un lemme montre qu'on peut supposer $A, B > 0$. Commençons par supposer aussi $C > 0$ et montrons le résultat.
  - Soit $r \in A \cdot (B + C)$, et écrivons $r \leqslant a(b + c)$ avec $a \geqslant 0$ et $b + c \geqslant 0$. On a $B > 0$ donc B contient un rationnel strictement positif. On en déduit $b_2 = max(0, b) \in B$ Or $ab \leqslant ab_2$ d'où $ab \in A \cdot B$. De même on montre $ac \in A \cdot C$. On a donc $a(b + c) \in A \cdot B + A \cdot C$ et de même pour r.
  - Réciproquement, soit $r \in A \cdot B + A \cdot C$ et écrivons $r = r_1 + r_2$ avec $r_1 \leqslant a_1 b$, $r_2 \leqslant a_2 c$ et les conditions habituelles Si on pose $a = max(a_1, a_2) \in A$ on a $r \leqslant a(b + c)$ (et bien sûr et $a \geqslant 0$ et $b + c \geqslant 0$) d'où $r \in A \cdot (B + C)$.

  Maintenant, si $C < 0$ montrons plutôt $A \cdot (B - C) = A \cdot B - A \cdot C$ avec $C > 0$.
  - Si $C < B$ on peut appliquer le point précédent, et on obtient

$$\begin{aligned} A \cdot (B - C) + A \cdot C &= A \cdot B + A \cdot (-C) + A \cdot C \\ &= A \cdot B - A \cdot C + A \cdot C \\ &= A \cdot B \end{aligned}$$

  - Si $C = B$ le résultat est immédiat.
  - Si $C > B$ on applique un lemme pour se ramener au premier cas.

$$\begin{aligned} A \cdot (B - C) &= -[A \cdot (C - B)] \\ &= -[A \cdot C - A \cdot B] \\ &= A \cdot B - A \cdot C \end{aligned}$$

Ainsi la preuve est achevée. □

Proposition 7.30. *On garde la compatibilité de $\leqslant$ avec la multiplication.*

Démonstration. De manière précise, montrons que

$$\forall (A, B, C) \in \mathbb{R}^3 [A \leq B \text{ et } C \geq 0] \Rightarrow AC \leq BC$$

Par distributivité et compatibilité avec l'addition, tout revient à se donner $A, B \geq 0$ et à montrer que $A \cdot B \geq 0$.

- Si A ou B est nul, le résultat est immédiat.
- Si $A, B > 0$, montrons que $A \cdot B \supset A_0$. Soit donc $x < 0$. Comme on peut trouver $a > 0$ dans A et $b > 0$ dans B, il vient immédiatement $x \leq ab$.

Ainsi la preuve est achevée. □

Théorème 7.31. *Soit X une partie non vide et majorée de $\mathbb{R}$. Alors elle admet une borne supérieure.*

Démonstration. Posons $M = \bigcup_{A \in X} A \subset \mathbb{Q}$, et montrons que M est la borne supérieure cherchée. Pour cette démonstration, il est particulièrement important de se rappeler que les minuscules désignent des rationnels, donc des éléments de Q, et les majuscules des réels, donc des parties de $\mathbb{R}$.

- Montrons déjà que $M \in \mathbb{R}$
  - X est non vide, donc on peut choisir $A \in X$. Ensuite on a $M \supset A \neq \emptyset$, donc $M \neq \emptyset$. Par ailleurs, si r est majorant de X, prenons $r \in \mathbb{Q} \backslash M_1$ et montrons que $r \notin M$. Ainsi aurons $M \neq \mathbb{Q}$. Soit donc $A \in X$ quelconque, et montrons que $r \notin A$. Comme $A \subset M_1$, si on avait $r \in A$ on aurait $r \in M_1$, contradiction.
  - Soit maintenant $m \in M$ et $m' < m$. Alors on peut trouver $A \in X$ tel que $m \in A$ puis on a $m' \in A$ d'où $m' \in M$.
  - Enfin, soit $m \in M$. Et trouvons $m' > m$ dans M. De même, on prend $A \in X$ tel que $m \in A$, puis comme A ne possède de maximum, on peut trouver $m' > m$ dans A. On a alors $m' \in M$.
- Montrons maintenant que $M = sup(X)$.
  - Par construction, M contient tous les A, donc c'est un majorant de X.
  - Donnons-nous $M_2 < M$, et trouvons $A > M_2$ dans X. Comme M contient strictement M_2 on peut trouver $m \in M \backslash M_2$. Soit $A \in X$ tel que $m \in A$. Puisque $m \notin M_2$, A n'est pas inclus dans M_2. Et comme sur $\mathbb{R}$, $\leqslant$ est une relation d'ordre totale, c'est donc que $A > M_2$.

Ainsi la preuve est achevée. □

Corollaire 7.8. *Soit X une partie non vide et minorée de $\mathbb{R}$. Alors elle admet une borne inférieure.*

Démonstration. Considérons $-X = \{-A | A \in X\}$. Cette partie est non vide et majorée, donc elle admet une borne supérieure qu'on note m. On en déduit que m est un minorant de X. Puis si on se donne $m_1 > m$, alors on a $-m_1 < -m$, donc $-m_1$ n'est pas un majorant de $-X$. Ainsi on peut trouver $-A > -m_1$ avec $A \in X$. Autrement dit on a $A < m_1$, ce qui montre que m_1 n'est pas un minorant de X. Donc $m = inf(X)$. □

5.2 Complétude de $\mathbb{R}$

Définition 7.32. La suite $(u_n)_{n \in \mathbb{N}}$ est une **suite de Cauchy** lorsque "ses termes sont de plus en plus proches les uns des autres" :

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall m, n \geqslant N, |u_m - u_n| \leqslant \varepsilon$$

Proposition 7.31. *Toute suite convergente est de Cauchy.*

Démonstration. Soit $(u_n)_{n\in\mathbb{N}}$ une suite de limite l, et soit $\varepsilon > 0$. Alors il existe $N \in \mathbb{N}$ tel que pour $n \geqslant N$ on ait $|u_n - l| \leqslant \frac{\varepsilon}{2}$. Pour $m, n \geqslant N$ on aura bien $|u_m - u_n| \leqslant \varepsilon$ par l'inégalité triangulaire. □

Plus globalement, ce dernier résultat est vrai dans beaucoup d'espaces, dès que l'on peut y définir une distance. En revanche, la question de la réciproque est éminemment délicate dans le cas général. C'est même elle qui motive la construction de $\mathbb{R}$, puisqu'elle est fausse dans $\mathbb{Q}$ et vraie dans $\mathbb{R}$.

Avant toute chose, démontrons le

Lemme 7.9. *Si une suite de Cauchy possède une suite extraite convergente, alors elle est convergente.*

Démonstration. Soit $(u_n)_{n\in\mathbb{N}}$ une suite de Cauchy telle que $(u_{\varphi(n)})_{n\in\mathbb{N}}$ converge vers une limite l, et donnons-nous $\varepsilon > 0$. Pour $m, n \geq N_1$ on a $|u_m - u_n| \leq \frac{\varepsilon}{2}$; pour $n \geq N_2$ on a$|u_{\varphi(n)} - l| \leq \frac{\varepsilon}{2}$. De plus, rappelons qu'on a $\varphi(n) \geq n$. Donc pour $n \geq max(N_1, N_2)$ on a

$$|u_n - l| \leq |u_n - u_{\varphi(n)}| + |u_{\varphi(n)} - l| \leq \varepsilon_i$$

ce qui achève la preuve. □

Théorème 7.32 (Complétude de $\mathbb{R}$)**.** *Toute suite de Cauchy converge dans* $\mathbb{R}$*. On dit que* $\mathbb{R}$ *est* ***complet****.*

Démonstration. Soit $(u_n)_{n\in\mathbb{N}}$ une suite de Cauchy, et $N \in \mathbb{N}$ tel que $\forall m, n \geq N |u_m - u_n| \leq 1$. En particulier, pour $n \geq N$ a

$$|u_n| \leq |u_n - u_N| + |u_N| \leq 1 + |u_N|$$

Donc l'ensemble des $|u_n|$ est majoré par $\max(|u_0|, ..., |u_{N-1}|, 1+|u_N|)$ et $(u_n)_{n\in\mathbb{N}}$ est bornée. D'après le théorème de Bolzano-Weierstrass extraite convergente, et on conclut par le lemme. □

Théorème 7.33. $\mathbb{Q}$ *n'est pas complet.*

Démonstration. Avant toute chose, précisons ce que nous entendons par là. Une suite de rationnels $(r_n)_{n\in\mathbb{N}}$ converge (dans $\mathbb{Q}$) vers $l \in \mathbb{Q}$ lorsque

$$\forall \varepsilon \in \mathbb{Q}_+^*,\ \exists N \in \mathbb{N},\ \forall n \geq N, |r_n - l| \leq \varepsilon$$

. On vérifie que les résultats suivants, déjà prouvés sur $\mathbb{R}$ dans le cours, restent valables sans modification.

- Lorsqu'elle existe, la limite est unique.
- On peut effectuer les mêmes opérations algébriques (somme, produit, etc.).
- La suite $(\frac{1}{2^n})_{n\in\mathbb{N}}$ tend vers 0 dans $\mathbb{Q}$ (conséquence d'un lemme).

De même, une suite de rationnels $(r_n)_{n\in\mathbb{N}}$ est une suite de Cauchy (dans $\mathbb{Q}$) lorsque

$$\forall \varepsilon \in \mathbb{Q}_+^*,\ \exists N \in \mathbb{N},\ \forall m, n \geq N,\ |r_m - r_n| \leq \varepsilon$$

Ensuite, si on veut passer par $\mathbb{R}$, il suffit d'écrire $\sqrt{2}$ comme limite d'une suite de rationnels. Mais si on cherche une démonstration plus intrinsèque qui passe uniquement par les rationnels, en voici une. La preuve fonctionne en plusieurs étapes.

- Première étape : on construit une suite (r_n) telle que $r_n > 0$ et $r_n^2 \to 2$ dans $\mathbb{Q}$. Soit la suite (p_n, q_n) à valeurs dans $\mathbb{N}^2$ définie par

$$\begin{cases} (p_0, q_0) = (1, 0) \\ \forall n \geq 0, (p_{n+1}, q_{n+1}) = (p_n + 2q_n, p_n + q_n) \end{cases}$$

On montre alors par récurrence la propriété suivante : $p_n \geq 1, q_n \geq np_n^2 - 2q_n^2 = (-1)^n$.

- Initialisation : $p_0 \geqslant 1$, $q_0 \geqslant 0$ et $p_0^2 - 2q_0^2 = 1 = (-1)^0$.
- Hérédité : soit $n \in \mathbb{N}$ tel que la propriété soit vraie.
  * On a $p_{n+1} = p_n + 2q_n \geqslant 1 + 2n \geqslant 1$.
  * Par ailleurs, $q_{n+1} = p_n + q_n \geqslant 1 + n$.
  * Enfin, on a

$$\begin{aligned} p_{n+1}^2 - 2q_{n+1}^2 = (p_n + 2q_n)^2 - 2(p_n + q_n)^2 &= -p_n^2 + 2q_n^2 \\ &= -(-1)^n \\ &= (-1)^{n+1} \end{aligned}$$

On en déduit $\forall n > 0$, $q_n \neq 0$ et $\dfrac{p_n^2}{q_n^2} = 2 + \dfrac{(-1)^n}{q_n^2}$ Or $\left|\dfrac{(-1)^n}{q_n^2}\right| = \dfrac{1}{q_n^2} \leqslant \dfrac{1}{n^2} \leqslant \dfrac{1}{2^n} \to 0$. Donc $\frac{p_n^2}{q_n^2} \to 2$

- Deuxième étape : la suite (r_n) est de Cauchy dans $\mathbb{Q}$. En effet, soit $\varepsilon \in \mathbb{Q}_+^*$. À partir d'un certain rang N_1, on a $r_n^2 \geqslant 1$ (puisque ce nombre tend vers 2), $r_n \geqslant 1$. À partir d'un certain rang N_2, $|r_n^2 - 2| \leqslant \varepsilon$. Pour $m, n \geqslant max(N_1, N_2)$, alors

$$|r_n - r_m| = \frac{|r_n^2 - r_m^2|}{r_n + r_m} \leqslant \frac{2\varepsilon}{2} = \varepsilon$$

- Troisième étape : supposons que la suite (r_n) converge vers $r \in \mathbb{Q}$. Alors (r_n^2) convergerait vers r^2 puis par unicité de la limite on obtiendrait $r^2 = 2$, contradiction.

Conclusion : $\mathbb{Q}$ n'est pas complet ! □

Remarque 7.52. Nous connaissons donc deux propriétés de $\mathbb{R}$ que ne possède pas $\mathbb{Q}$: la propriété de la borne supérieure et la complétude. En fait, ces deux propriétés ne sont que les deux versants d'une même montagne.

En construisant $\mathbb{R}$ comme nous l'avons fait, nous avons mis l'accent sur la propriété de la borne supérieure, et la complétude en a découlé. Mais il existe une autre construction de $\mathbb{R}$, qui consiste à quotienter l'anneau des suites rationnelles de Cauchy par l'idéal des suites rationnelles qui tendent vers 0 (cf. DM Avancé). Cette construction (qui peut être généralisée à d'autres ensembles) rend *de facto* $\mathbb{R}$ complet, et la propriété de la borne supérieure est alors une conséquence aisément démontrable (pas sûr du call là...) : on commence par démontrer le théorème des suites adjacentes, et le reste s'en déduit.

En définissant la notion de suite de Cauchy dans $\mathbb{C}$ comme dans $\mathbb{R}$ (à condition de remplacer les valeurs absolues par des modules), on obtient le

Théorème 7.34. $\mathbb{C}$ *est complet.*

Démonstration. Il suffit de remarquer que si la suite complexe $(a_n + ib_n)_{n\in\mathbb{N}}$ est de Cauchy, alors $(a_n)_{n\in\mathbb{N}}$ et $(b_n)_{n\in\mathbb{N}}$ le sont aussi en utilisant l'inégalité sur les modules des parties réelles et imaginaires. □

6 Compléments : retour sur la théorie des ensembles, HP

Nous arrivons maintenant au terme de notre voyage ensembliste. Après avoir posé axiomatiquement l'existence de $\mathbb{N}$, nous avons construit $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$... puis toute l'analyse s'en déduit, comme le montre la suite du cours. Il est assez merveilleux de contempler la façon dont quelques axiomes de base très simples sur $\mathbb{N}$ nous ont permis de construire un socle inamovible pour une grande partie des mathématiques... Mais on peut aller plus loin...

Rappelons qu'à partir de $\mathbb{N}$, les éléments de $\mathbb{Z}$ ont été définis comme des classes d'équivalence sur $\mathbb{N}^2$, les éléments de $\mathbb{Q}$ comme des classes d'équivalence sur $\mathbb{Z}\times\mathbb{Z}^*$, les éléments de $\mathbb{R}$ comme des parties de $\mathbb{Q}$, et enfin les éléments de $\mathbb{C}$ comme des couples dans $\mathbb{R}^2$. Bref, à chaque fois, les nouveaux nombres que nous introduisions étaient des ensembles. De même pour l'analyse, une fonction ou une application est un ensemble. D'où l'idée un peu folle qui a germé dans le cerveau des mathématiciens : et si tout était ensemble ? Il ne reste plus qu'à construire les éléments de $\mathbb{N}$ eux-mêmes comme des ensembles, et ce sera effectivement vrai. Nous allons voir qu'une telle construction est possible.

Il y a quelques chapitres, nous avions posé axiomatiquement l'existence d'un ensemble non vide $\mathbb{N}$, muni d'une relation d'ordre, et qui vérifiait trois propriétés fondamentales. Pour mémoire :

- toute partie non vide admet un minimum ;
- toute partie non vide et majorée admet un maximum ;
- $\mathbb{N}$ n'admet pas de maximum.

Nous en avions notamment déduit la validité du principe de récurrence, qui est à la base de l'arithmétique (définition par récurrence des lois de $\mathbb{N}$, division euclidienne). Mais le point de vue inverse est possible : il existe une construction de $\mathbb{N}$, due au mathématicien italien Giuseppe Peano, qui postule d'emblée la validité du principe de récurrence (c'est un axiome), pour en déduire les trois propriétés fondamentales sus-citées. Les deux constructions sont donc équivalentes.

Or quel que soit le point de vue adopté (à partir des trois propriétés fondamentales ou à partir des axiomes de Peano), on peut se demander si un tel ensemble $\mathbb{N}$ doit vraiment être posé axiomatiquement. En clair : peut-on démontrer l'existence et l'unicité de $\mathbb{N}$?

La question de l'unicité est assez simple. Par exemple, si deux ensembles ordonnés non vides vérifient les trois propriétés fondamentales, il est facile de construire une bijection de l'un dans l'autre qui respecte la relation d'ordre (on peut le faire par récurrence). On peut donc dire que, sous réserve d'existence, $\mathbb{N}$ est unique "à isomorphisme près".

Reste la question de l'existence. Il faudrait pouvoir construire $\mathbb{N}$ de manière effective à partir des axiomes de Zermelo-Fraenkel. Ainsi, les trois propriétés fondamentales (ou les axiomes de

Peano selon le point de vue) perdraient leur statut d'axiome, pour devenir des théorèmes. Sur le plan mathématique ce serait satisfaisant, car on est toujours un peu gêné d'introduire des axiomes supplémentaires : par souci d'économie et d'esthétique, on peut se demander si tout cela n'est pas redondant. Cette construction est presque possible, à ceci près que nous allons devoir introduire un dernier axiome, dit "axiome de l'infini".

Construisons donc $\mathbb{N}$ de manière effective. Ses éléments seront eux-mêmes des ensembles. Intuitivement, on choisit pour 0 le "plus petit" ensemble possible, c'est-à-dire qu'on pose $0 = \emptyset$. Une fois qu'on a construit n (comme un ensemble, rappelons-le), il est légitime de chercher un ensemble "juste un peu plus grand". Le mieux est de réunir n avec un singleton $\{m\}$, tel que $m \notin n$. L'axiome de fondation nous suggère alors de choisir $m = n$. Ainsi, pour tout n, on pose $n+1 = n \cup \{n\} = \{0, 1, ..., n\}$. Par exemple, on a

$$\begin{cases} 0 = \emptyset \\ 1 = \{\emptyset\} \\ 2 = \{\emptyset, \{\emptyset\}\} \\ 3 = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\} \\ \vdots \end{cases}$$

Il reste à formaliser cela.

Définition 7.33. Si X est un ensemble, on note $\mathcal{Q}(X)$ la propriété

$$\begin{cases} \alpha \in X \\ \forall x \in X, x \cup \{x\} \in X \end{cases}$$

X est donc un ensemble qui contient notre 0, et qui est stable par l'opération "successeur". On pose alors l'axiome suivant.

Axiome 7.1 (Axiome de l'infini)**.** Il existe au moins un ensemble X qui vérifie $\mathcal{Q}(X)$.

Maintenant qu'il existe au moins un tel ensemble X, on va définir $\mathbb{N}$ comme le plus petit de ces ensembles au sens de l'inclusion. Cela peut se faire par compréhension. Soit X un ensemble qui vérifie $\mathcal{Q}(X)$ et posons la

Définition 7.34 ($\mathbb{N}$)**.** On définit $\mathbb{N} = \{x \in X \mid \forall Y, \mathcal{Q}(Y) \Rightarrow x \in Y\}$.

On remarque tout de suite que $\mathbb{N}$ vérifie $\mathcal{Q}(\mathbb{N})$. Notons aussi que la définition de $\mathbb{N}$ est indépendante du choix de X. En effet, si nous choisissons un autre ensemble $\tilde{X}$, dont dérive un certain $\tilde{\mathbb{N}}$ alors quel que soit $x \in \mathbb{N}$ on a $\mathbb{Q}(\mathbb{N})$ donc $x \in \mathbb{N}$. Autrement dit, on a $\mathbb{N} \subset \tilde{N}$. Par symétrie on a l'inclusion réciproque, puis conclut par l'axiome d'extensionnalité.

Tel que nous l'avons contruit, $\mathbb{N}$ vérifie automatiquement un certain principe de récurrence, mais d'un genre bien particulier. Une récurrence simple, de type ensembliste, avec "initialisation à $\emptyset$". Nous en déduirons les trois propriétés fondatrices de $\mathbb{N}$. Pour commencer, on a le

Théorème 7.35 (Principe de récurrence ensembliste)**.** *Soit A une partie de $\mathbb{N}$ vérifiant les deux points suivants.*

- *Initialisation : $\emptyset \in A$*

- *Hérédité :* $\forall n,\ n \in A \Rightarrow n \cup \{n\} \in A$.

Alors $A = \mathbb{N}$.

Démonstration. Il suffit de remarquer que A vérifie $\mathcal{Q}(A)$. Ainsi, si $n \in \mathbb{N}$ en posant $Y = A$ on a $n \in A$. On en déduit que $\mathbb{N} \subset A$, et comme $A \subset \mathbb{N}$ a $A = \mathbb{N}$ par extensionnalité. □

Définition 7.35. Sur $\mathbb{N}$, on définit la relation par $m \leqslant n$ (m est inférieur ou égal à n).

Nous allons bien sûr démontrer que $\leqslant$ est une relation d'ordre, mais pour nous simplifier la tâche, nous allons utiliser un petit lemme à l'énoncé pour le moins surprenant.

Lemme 7.10. *Soit $m, n \in \mathbb{N}$. Si $m \in n$ alors $m \subset n$ (non, il n'y a pas de faute de frappe !).*

Démonstration. On démontre le résultat par récurrence sur n. Plus précisément, définissons par compréhension

$$A := \{n \in \mathbb{N} \mid \forall m, [m \in n \Rightarrow m \subset n]\}$$

- Initialisation : si $n = \emptyset$, alors on ne peut pas avoir $m \in n$ et comme le faux implique n'importe quoi, on en déduit que $n \in A$
- Hérédité : soit $n \in A$ et montrons que $n \cup \{n\} \in A$. Soit donc $m \in n \cup \{n\}/$
  - Si $m \in n$, alors $m \subset n$ car $n \in A$ (on pourrait dire qu'on applique l'hypothèse de récurrence) et *a fortiori* on a $m \subset n \cup \{n\}$.
  - Si $m \in \{n\}$, alors $m = n$ donc $m \subset n$ et on se ramène au cas précédent.

Ainsi la récurrence est achevée. □

Corollaire 7.9. *Soit $m, n \in \mathbb{N}$. $m \leqslant n$ alors $m \subset n$*

Démonstration. Si $m \in n$ c'est une conséquence du lemme précédent, et si $m = n$ le résultat est immédiat. □

Corollaire 7.10. *La relation $\leqslant$ est une relation d'ordre sur $\mathbb{N}$.*

Démonstration. Vérifions les axiomes :

- Réflexivité : soit $n \in \mathbb{N}$, on a $n = n$ donc $n \leqslant n$.
- Symétrie : soit $m, n \in \mathbb{N}$ tels que $m \leqslant n$ et $n \leqslant m$. Alors $m \subset n$ et $n \subset m$ donc $m = n$ par extensionnalité.
- Transitivité : soit $m, n, p \in \mathbb{N}$ tels que $m \leqslant n$ et $n \leqslant p$. Si $m = n$, on a bien $m \leqslant p$. Sinon, on a $m \in n$, et comme $b \subset p$, on a $m \in p$, donc $m \leqslant p$.

Ainsi, la preuve est achevée. □

Lemme 7.11. *Soit $m, n \in \mathbb{N}$. Si $m \in n$, alors $m \cup \{m\} \leqslant n$.*

Démonstration. On montre encore le résultat par récurrence sur n. Définissons par compréhension

$$A := \{n \in \mathbb{N} \mid \forall m,\ (m \in \mathbb{N} \text{ et } m \in n) \implies m \cup \{m\} \leqslant n\}$$

- Initialisation : quel que soit m, on n'a jamais $m \in \emptyset$, et comme le faux implique n'importe quoi, on en déduit que $\emptyset \in A$.
- Hérédité : supposons que $n \in A$, et soit $m \in \mathbb{N}$ tel que $m \in n \cup \{n\}$.

– Si $m \in n$, comme $n \in A$, on a $m \cup \{m\} \leqslant n$ et comme $n \leqslant n \cup \{n\}$ (car $n \in n \cup \{n\}$), on conclut par transitivité.
– Si $m \in \{n\}$, alors $m = n$ donc $m \cup \{m\} = n \cup \{n\}$ et on conclut par réflexivité.

Ainsi, la récurrence est achevée. □

Nous en arrivons au théorème final.

Théorème 7.36. $\mathbb{N}$ *ainsi construit est non vide, et il vérifie les trois propriétés suivantes.*

- *N n'admet pas de maximum.*
- *Toute partie non vide et majorée admet un maximum.*
- *Toute partie non vide admet un minimum.*

Démonstration. Montrons que $\mathbb{N}$ vérifie ce qu'on lui demande.

- $\mathbb{N}$ est non vide. En effet, il contient $\emptyset$.
- $\mathbb{N}$ n'admet pas de maximum. Supposons par l'absurde que $\mathbb{N}$ admette un maximum n. Alors comme $n \cup \{n\} \in \mathbb{N}$, on aurait $n \cup \{n\} \subseteq n$ et en particulier on aurait $n \in n$, contradiction d'après l'axiome de fondation.
- Toute partie non vide et majorée de $\mathbb{N}$ admet un maximum. Soit B non vide et majorée par $a \in \mathbb{N}$, et montrons par récurrence sur a que B admet un maximum. Plus précisément, nous introduisons par compréhension

$$A := \{a \in \mathbb{N} \mid \forall B, B \neq \emptyset \text{ et } B \text{ majorée par } a \Rightarrow B \text{ admet un maximum}\}$$

  – Initialisation : si B est majorée par $\emptyset$ alors tous les éléments de B sont $\leqslant \emptyset$ donc $\subseteq \emptyset$ et ils sont donc $= \emptyset$. B étant non vide, elle admet donc un maximum, qui est $\emptyset$. Ainsi, $\emptyset \in A$.
  – Hérédité : supposons que $a \in A$ et montrons que $a \cup \{a\} \in A$. Soit B non vide et majorée par $a \cup \{a\}$. Si $a \cup \{a\} \in B$, alors B admet un maximum. Sinon, tous les éléments de B appartiennent à $a \cup \{a\}$ donc ils appartiennent à a ou ils sont égaux à a. Dans tous les cas ils sont $\leqslant a$, on conclut par hypothèse de récurrence.
- Toute partie non vide de $\mathbb{N}$ admet un minimum. Soit A non vide. Alors A admet au moins un minorant, qui est $\emptyset$. Donc si on définit par compréhension l'ensemble de ses minorants

$$B := \{b \in \mathbb{N} \mid \forall a, a \in A \Rightarrow b \leqslant a\}$$

  alors B est non vide. De plus, A est non vide donc elle admet au moins un élément, et cet élément majore B. Ainsi, d'après le point précédent B admet un maximum b. Montrons que b est le minimum de A. Déjà il minore bien A, donc il ne reste plus qu'à montrer qu'il appartient à A.
  Supposons que $b \notin A$. Alors pour tout $a \in A$ on pourrait écrire $b \in a$. Mais alors on aurait $b \cup \{b\} \leqslant a$ le lemme, donc $b \cup \{b\} \in B$. En particulier, on aurait $b \cup \{b\} \leqslant b$, donc $b \cup \{b\} \subseteq b$; puis $b \in b$, contradiction l'axiome de fondation.

Ainsi la preuve est achevée. □

Remarque 7.53. On rappelle que $\leqslant$ est automatiquement totale, puisque si $(m, n) \in \mathbb{N}^2$, l'ensemble $\{m, n\}$ est non vide donc il admet un minimum.

Chapitre 8

Nombres complexes

1 Premières définitions

1.1 Construction de $\mathbb{C}$, HP

Définition 8.1 (Ensemble des nombres complexes et i)**.** On définit un nombre complexe comme un couple de réels $(a, b) \in \mathbb{R}^2$. On note $\mathbb{C}$ l'ensemble des nombres complexes et on définit le nombre complexe i par $i = (0, 1)$.

Remarque 8.1. L'application de $\mathbb{R}$ dans $\mathbb{C}$ qui à x associe $(x, 0)$ est injective, ce qui permet de considérer abusivement que $\mathbb{R} \subset \mathbb{C}$.

Définition 8.2 (Lois de $\mathbb{C}$)**.** On définit les lois $+$ et $\times$ sur $\mathbb{C}$ de la manière suivante :

1. $\forall((a, b), (c, d)) \in \mathbb{C}^2,\ (a, b) + (c, d) = (a + c, c + d)$
2. $\forall((a, b), (c, d)) \in \mathbb{C}^2,\ (a, b) \times (c, d) = (ac - bd, ad + bc)$

On vérifie immédiatement que $\mathbb{C}$ est stable par ces lois, et que ces lois sont associatives et commutatives. On vérifie de même que $\times$ est distributive sur $+$. Leurs éléments neutres respectifs sont $(0, 0)$ et $(1, 0)$. $(0, 0)$ est absorbant pour $\times$. Ces lois prolongent les lois usuelles de $\mathbb{R}$.

Théorème 8.1 (Unicité de la forme algébrique)**.**

$$\forall z \in \mathbb{C},\ \exists!(a, b) \in \mathbb{R}^2,\ z = a + ib$$

C'est ce que on appelle la forme ***algébrique*** *de z.*

Proposition 8.1. $(\mathbb{C}, +, \times)$ *est un corps.*

Remarque 8.2. Attention : il n'y a pas de relation d'ordre convenable sur $\mathbb{C}$.

1.2 Conjugaison

Définition 8.3 (Conjugué)**.** Soit $z \in \mathbb{C}$ qu'on écrit sous forme algébrique $z = a + ib$. Le **conjugué** de z, noté $\overline{z}$, est défini par :

$$\overline{z} := a - ib$$

Proposition 8.2 (Involutivité du passage au conjugué)**.** *On a* $\forall z \in \mathbb{C},\ \overline{(\overline{z})} = z$.

Proposition 8.3. *On a :*

$$\forall z \in \mathbb{C},\ \begin{cases} Re(f) = \dfrac{z+\overline{z}}{2} \\ Im(f) = \dfrac{z-\overline{z}}{2i} \end{cases}$$

Proposition 8.4 (Caractérisation de $\mathbb{R}$ et $i\mathbb{R}$ par la conjugaison)**.** *Soit* $z \in \mathbb{C}$.

$$z \in \mathbb{R} \iff \overline{z} = z$$

$$z \in i\mathbb{R} \iff \overline{z} = -z$$

Remarque 8.3. On pourra souvent utiliser cette proposition pour montrer qu'un complexe appartient à $\mathbb{R}$ ou $i\mathbb{R}$, en particulier en présence de nombres complexes de module 1, avec lesquels le conjugué fait bon ménage.

Proposition 8.5.

$$\forall (z, z') \in \mathbb{C}^2,\ \begin{cases} \overline{z+z'} = \overline{z} + \overline{z'} \\ \overline{zz'} = \overline{z}\overline{z'} \end{cases}$$

Proposition 8.6. *Soit* $z \in \mathbb{C}$. *On a* $\overline{-z} = -\overline{z}$. *De plus, supposons que* $z \neq 0$. *Alors* $\overline{z} \neq 0$, *et on a* $\overline{\left(\dfrac{1}{z}\right)} = \dfrac{1}{\overline{z}}$.

Exemple 8.1.

$$\forall (z, z') \in \mathbb{C}^2,\ \overline{z - z'} = \overline{z} - \overline{z'}$$

$$\forall (z, z') \in \mathbb{C} \times \mathbb{C}^\star,\ \overline{\left(\frac{z}{z'}\right)} = \frac{\overline{z}}{\overline{z'}}$$

Proposition 8.7 ($\mathbb{R}$-linéarité des parties réelle et imaginaire)**.** *Soit* $(z, z') \in \mathbb{C}^2$ *et* $(\lambda, \mu) \in \mathbb{R}^2$. *Alors :*

$$\begin{cases} Re(\lambda z + \mu z') = \lambda Re(z) + \mu Re(z') \\ Im(\lambda z + \mu z') = \lambda Im(z) + \mu Im(z') \end{cases}$$

1.3 Module

Définition 8.4. Soit $z \in \mathbb{C}$ qu'on écrit sous forme algébrique $z = a + ib$. Le **module** de z, noté $|z|$, est défini par :

$$|z| := \sqrt{a^2 + b^2}$$

Remarque 8.4. Le module prolonge la valeur absolue réelle.

Définition 8.5 (Cercle unité $\mathbb{U}$)**.** On pose

$$\mathbb{U} := \{z \in \mathbb{C} \mid |z| = 1\}$$

C'est l'ensemble des **unimodulaires**, aussi appelé **cercle unité** ou **cercle trigonométrique**.

Proposition 8.8. $\forall z \in \mathbb{C},\ z\overline{z} = |z|^2$

Méthode 8.1 (Calculs impliquant des modules)**.** En présence de calculs de modules, penser à élever au carré, ce qui fait apparaître des conjugués très pratiques par la proposition précédente.

Proposition 8.9. $\forall z \in \mathbb{U},\ \overline{z} = \dfrac{1}{z}$

Proposition 8.10 (Séparation)**.** $\forall z \in \mathbb{C},\ |z| = 0 \iff z = 0$

Proposition 8.11. $\forall z \in \mathbb{C},\ |z| = |\overline{z}| = |-z| = |-\overline{z}|$

Proposition 8.12 (Vers l'inégalité triangulaire)**.** *On a :*

$$\forall z \in \mathbb{C},\ \begin{cases} |Re(z)| \leqslant |z| \\ |Im(z)| \leqslant |z| \end{cases}$$

Méthode 8.2 (Montrer que deux complexes sont différents)**.** Pour montrer que deux complexes sont différents, il suffit de montrer que leurs parties réelles sont différentes, que leurs parties imaginaires sont différentes ou que leurs modules sont différents.

Méthode 8.3 (Factorisation d'une somme de carrés dans $\mathbb{C}$)**.** Ne pas oublier dans $\mathbb{C}$ la dernière identité remarquable $a^2 + b^2 = (a + ib)(a - ib)$.

Proposition 8.13. *On a :*

$$\begin{cases} \forall (z, z') \in \mathbb{C}^2,\ |zz'| = |z||z'| \\ \forall (z, z') \in \mathbb{C} \times \mathbb{C}^\star,\ \left|\dfrac{z}{z'}\right| = \dfrac{|z|}{|z'|} \end{cases}$$

Théorème 8.2 (Inégalité triangulaire)**.**

$$\forall (z, z') \in \mathbb{C},\ |z + z'| \leqslant |z| + |z'|$$

Corollaire 8.1 (Inégalité triangulaire généralisée)**.**

$$\forall n \geqslant 2,\ \forall (z_1, \ldots, z_n) \in \mathbb{C}^n,\ \left|\sum_{i=1}^{n} z_i\right| \leqslant \sum_{i=1}^{n} |z_i|$$

Corollaire 8.2 (Inégalité triangulaire renversée)**.**

$$\forall (z, z') \in \mathbb{C}^2,\ \big||z| - |z'|\big| \leqslant |z - z'|$$

Définition 8.6 (Colinéarité directe)**.** Soit $(z, z') \in \mathbb{C}^2$. Les trois assertions suivantes sont équivalentes :

1. $\begin{cases} \exists \lambda \in \mathbb{R}_+,\ z' = \lambda z \\ \text{ou} \\ \exists \mu \in \mathbb{R}_+,\ z = \mu z' \end{cases}$
2. $\begin{cases} \exists \lambda \in \mathbb{R}_+,\ z' = \lambda z \\ \text{ou} \\ z = 0 \end{cases}$

3. $\begin{cases} z' = 0 \\ \text{ou} \\ \exists \mu \in \mathbb{R}_+,\ z = \mu z' \end{cases}$

Lorsque ces assertions sont vérifiées, on dit que z et z' sont **colinéaires directs**.

Théorème 8.3 (Cas d'égalité de l'inégalité triangulaire)**.** *Soit $(z, z') \in \mathbb{C}^2$. On a $|z+z'| = |z|+|z'|$ si, et seulement si, z et z' sont colinéaires directs.*

2 Exponentielle complexe, trigonométrie

2.1 Exponentielle complexe

Définition 8.7 (Prolongement de l'exponentielle à $\mathbb{C}$)**.** La fonction $\exp : \mathbb{R} \to \mathbb{R}_+^\star$ admet un prolongement, encore noté exp et donné par :

$$\begin{array}{rccc} \exp & : & \mathbb{C} & \to & \mathbb{C}^\star \\ & & z & \mapsto & \displaystyle\sum_{n=0}^{+\infty} \frac{z^n}{n} \end{array}$$

$\exp(z)$ peut aussi être noté e^z. On admettra que :

$$\begin{cases} \forall (z, z') \in \mathbb{C}^2,\ e^{z+z'} = e^z e^{z'} \\ \forall z \in \mathbb{C},\ e^{\overline{z}} = \overline{e^z} \end{cases}$$

Proposition 8.14. *On a :*

1. $\forall \theta \in \mathbb{R},\ |e^{i\theta}| = 1$
2. $\forall (a, b) \in \mathbb{R}^2,\ |e^{a+ib}| = e^a$

Définition 8.8 (Cosinus et sinus)**.** Les fonctions **cosinus** et **sinus**, notées respectivement cos et sin sont définies par :

$$\begin{array}{rccc} \cos & : & \mathbb{R} & \to & \mathbb{R} \\ & & \theta & \mapsto & \mathrm{Re}(e^{i\theta}) \end{array}$$

et

$$\begin{array}{rccc} \sin & : & \mathbb{R} & \to & \mathbb{R} \\ & & \theta & \mapsto & \mathrm{Im}(e^{i\theta}) \end{array}$$

Proposition 8.15. $\forall \theta \in \mathbb{R},\ -1 \leqslant \cos(\theta), \sin(\theta) \leqslant 1$

Proposition 8.16. $\forall \theta \in \mathbb{R},\ \cos(\theta)^2 + \sin(\theta)^2 = 1$

Proposition 8.17 (Formules d'Euler)**.** *On a les formules suivantes, dites **formules d'Euler** :*

$$\forall \theta \in \mathbb{R},\ \begin{cases} \cos(\theta) = \dfrac{e^{i\theta} + e^{-i\theta}}{2} \\ \sin(\theta) = \dfrac{e^{i\theta} - e^{-i\theta}}{2i} \end{cases}$$

Théorème 8.4 (Admis)**.** sin *et* cos *sont dérivables et on a* $\sin' = \cos$ *et* $\cos' = -\sin$.

Théorème 8.5 (Technique de l'arc-moitié)**.** *Soit* $(a, b) \in \mathbb{R}^2$*. On a les formules dites d'****arc-moitié***

$$\begin{cases} e^{ia} + e^{ib} = 2\cos\left(\dfrac{a-b}{2}\right)\exp\left(i\dfrac{a+b}{2}\right) \\ e^{ia} - e^{ib} = 2i\sin\left(\dfrac{a-b}{2}\right)\exp\left(i\dfrac{a+b}{2}\right) \end{cases}$$

Remarque 8.5. Penser au cas où $b \equiv 0\ [2\pi]$, *ie* au cas où $e^{ib} = 1$.

2.2 Développements, linéarisations

Théorème 8.6 (Formule de Moivre)**.** *Soit* $\theta \in \mathbb{R}$*. On a la* ***formule de Moivre*** *:*

$$\forall n \in \mathbb{Z},\ (\cos(\theta) + i\sin(\theta))^n = \cos(n\theta) + i\sin(n\theta)$$

Autrement dit, $\forall n \in \mathbb{Z},\ (e^{i\theta})^n = e^{in\theta}$.

Méthode 8.4 (Développement de $\cos(n\theta)$ ou $\sin(m\theta)$)**.** On souhaite développer $\cos(n\theta)$ ou $\sin(m\theta)$, c'est-à-dire les exprimer comme des polynômes en $\cos(\theta)$ ou en $\sin(\theta)$. Pour cela :

- On écrit ce cosinus ou ce sinus comme une partie réelle ou imaginaire d'une exponentielle complexe.
- On utilise la formule de Moivre sur cette exponentielle complexe.
- On développe par le binôme de Newton après avoir écrit l'exponentielle sous la puissance sous forme algébrique.
- Enfin, on passe à la partie réelle ou imaginaire selon celle qu'on avait au départ pour obtenir le résultat.

Méthode 8.5 (Linéarisation de $\cos(\theta)^p \sin(\theta)^q$)**.** On souhaite linéariser $\cos(\theta)^p \sin(\theta)^q$, c'est-à-dire le transformer en combinaison linéaire de $\cos(n\theta)$ et de $\sin(n\theta)$. Pour cela :

- On exprime le cosinus et / ou le sinus à partir des formules d'Euler.
- Ensuite, on développe le tout avec le binôme de Newton, en mettant la puissance de 2 en facteur pour simplifier les calculs. Si on en en présence d'un produit d'un cosinus et d'un sinus, il faut développer normalement après avoir utilisé le binôme.
- On utilise la formule de Moivre après avoir développé.
- On regroupe les termes conjuguées grâce aux formules d'Euler pour former des $\cos(n\theta)$ et des $\sin(n\theta)<$;

2.3 Le nombre π

Définition 8.9 (π, HP)**.** L'ensemble $\{\theta > 0 \mid \cos(\theta) = 0\}$ admet un plus petit élément. On pose alors :

$$\pi := 2\min\{\theta > 0 \mid \cos(\theta) = 0\}$$

Théorème 8.7 (Admis)**.** *On a* $e^{i\frac{\pi}{2}} = i$ *et* $e^{i\pi} = -1$.

Proposition 8.18. *Soit $\theta \in \mathbb{R}$.*

$$\begin{cases} \cos(-\theta) = \cos(\theta) \\ \sin(-\theta) = -\sin(\theta) \end{cases}$$

$$\begin{cases} \cos(\pi - \theta) = -\cos(\theta) \\ \sin(\pi - \theta) = \sin(\theta) \end{cases}$$

$$\begin{cases} \cos(\pi + \theta) = -\cos(\theta) \\ \sin(\pi + \theta) = -\sin(\theta) \end{cases}$$

$$\begin{cases} \cos\left(\frac{\pi}{2} - \theta\right) = \sin(\theta) \\ \sin\left(\frac{\pi}{2} - \theta\right) = \cos(\theta) \end{cases}$$

$$\begin{cases} \cos\left(\frac{\pi}{2} + \theta\right) = -\sin(\theta) \\ \sin\left(\frac{\pi}{2} + \theta\right) = \cos(\theta) \end{cases}$$

On trouvera en **fin de chapitre** un **formulaire de trigonométrie** réunissant toutes les formules de trigonométries à connaître.

Proposition 8.19 (Parité et périodicité)**.** cos *est paire et 2π-périodique.* sin *est impaire et 2π-périodique.*

Théorème 8.8 (Valeurs particulières)**.** *Voici quelques valeurs particulières des fonctions cosinus et sinus :*

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

Définition 8.10 (Congruence modulo α)**.** Soit $\alpha > 0$ (en pratique, souvent π ou 2π dans ce chapitre). Soit $(x, y) \in \mathbb{R}^2$. On dira que x **et** y **sont congrus modulo** α, et on notera $x \equiv y\ [\alpha]$ lorsque :

$$\exists k \in \mathbb{Z},\ x = y + k\alpha$$

Proposition 8.20. *La congruence modulo $\alpha > 0$ est une relation d'équivalence sur $\mathbb{R}$.*

Proposition 8.21 (Quelques propriétés immédiates)**.** *Soit $\alpha > 0$.*

- $\forall (x, y) \in \mathbb{R}^2,\ x \equiv y\ [\alpha] \iff -x \equiv -y\ [\alpha]$

- $\forall (x,y,z) \in \mathbb{R}^3,\ x \equiv y\ [\alpha] \iff x+z \equiv y+z\ [\alpha]$
- $\forall \lambda > 0,\ \forall (x,y) \in \mathbb{R}^2,\ x \equiv y\ [\alpha] \iff \lambda x \equiv \lambda y\ [\lambda\alpha]$

Proposition 8.22. *Soit $\alpha > 0$. Les lois $+$ et $-$ de $\mathbb{R}$ passent au quotient de la relation de congruence modulo α. Précisément :*

$$\forall (x,x',y,y') \in \mathbb{R}^4,\ \begin{cases} x \equiv x'\ [\alpha] \\ y \equiv y'\ [\alpha] \end{cases} \implies \begin{cases} x+y \equiv x'+y'\ [\alpha] \\ x-y \equiv x'-y'\ [\alpha] \end{cases}$$

Remarque 8.6. Attention, cela ne fonctionne pas pour la multiplication en général, sauf pour les $\alpha \in \mathbb{Z}$.

Remarque 8.7. Grâce à une partie entière, il est toujours possible d'encadrer un $\theta \in \mathbb{R}$ par deux multiples de 2π consécutifs, avec un des encadrements strict.

Théorème 8.9 (Quelques équations trigonométriques)**.** *Soit $(\theta,\psi) \in \mathbb{R}^2$. On a :*

$$\cos(\theta) = 1 \iff \theta \equiv 0\ [2\pi]$$

$$\cos(\theta) = 0 \iff \theta \equiv \frac{\pi}{2}\ [\pi]$$

$$\cos(\theta) = -1 \iff \theta \equiv \pi\ [2\pi]$$

$$\cos(\theta) = \cos(\psi) \iff \theta \equiv \psi\ [2\pi] \quad ou \quad \theta \equiv -\psi\ [2\pi]$$

et

$$\sin(\theta) = 1 \iff \theta \equiv \frac{\pi}{2}\ [2\pi]$$

$$\sin(\theta) = 0 \iff \theta \equiv 0\ [\pi]$$

$$\sin(\theta) = -1 \iff \theta \equiv -\frac{\pi}{2}\ [2\pi]$$

$$\sin(\theta) = \sin(\psi) \iff \theta \equiv \psi\ [2\pi] \quad ou \quad \theta \equiv \pi - \psi\ [2\pi]$$

Remarque 8.8. Bien souvent, un dessin d'un cercle trigonométrique pourra servir de preuve pour des inéquations trigonométriques, en résolvant le cas d'égalité avec le théorème qui précède.

2.4 La fonction tangente

Définition 8.11 (Tangente)**.** On définit la fonction **tangente**, notée tan, par :

$$\begin{array}{rccc} \tan & : & \mathbb{R} \setminus \left\{ \frac{\pi}{2} + k\pi \mid k \in \mathbb{Z} \right\} & \to & \mathbb{R} \\ & & \theta & \mapsto & \dfrac{\sin(\theta)}{\cos(\theta)} \end{array}$$

Théorème 8.10 (Étude de la fonction tangente)**.** tan *est impaire et π-périodique. Elle est dérivable en tout point de son domaine de définition et on a :*

$$\tan' = 1 + \tan^2 = \frac{1}{\cos^2}$$

Elle est strictement croissante sur $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ *et vérifie*

$$\tan(x) \xrightarrow[x \to -\frac{\pi}{2}^+]{} -\infty$$

et

$$\tan(x) \xrightarrow[x \to \frac{\pi}{2}^-]{} +\infty$$

Proposition 8.23 (Formules impliquant la tangente de l'arc-moitié)**.** *Soit* $\theta \in \mathbb{R}\backslash\{\pi + 2k\pi \mid k \in \mathbb{Z}\}$ *et posons* $t := \frac{\theta}{2}$*. On a :*

$$\begin{cases} \cos(\theta) = \dfrac{1-t^2}{1+t^2} \\ \sin(\theta) = \dfrac{2t}{1+t^2} \\ \tan(\theta) = \dfrac{2t}{1-t^2} \end{cases}$$

Remarque 8.9. Ces formules se révèleront capitales en intégration.

Remarque 8.10. Pour retenir ces formules, penser aux parités et au fait que la tangente n'est pas toujours définie, alors que le cosinus et le sinus si.

Théorème 8.11 (Valeurs particulières)**.** *Voici quelques valeurs particulières de la fonction tangente :*

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

Définition 8.12 (Cotangente, HP)**.** On définit la fonction **cotangente**, notée cotan, par :

$$\begin{array}{rccc} \text{cotan} & : & \mathbb{R} \setminus \{k\pi \mid k \in \mathbb{Z}\} & \to & \mathbb{R} \\ & & \theta & \mapsto & \dfrac{\cos(\theta)}{\sin(\theta)} \end{array}$$

Remarque 8.11 (Lien cotangente/tangente)**.** Si $\theta \in \mathbb{R} \setminus \left\{k\frac{\pi}{2} \mid k \in \mathbb{Z}\right\}$, alors $\text{cotan}(\theta) = \dfrac{1}{\tan(\theta)}$

Théorème 8.12 (Équations trigonométriques)**.** *Soit* $\theta, \psi \in \mathbb{R} \setminus \left\{\frac{\pi}{2} + k\pi \mid k \in \mathbb{Z}\right\}$*. On a :*

$$\tan(\theta) = 0 \iff \theta \equiv 0 \ [\pi]$$

$$\tan(\theta) = \tan(\psi) \iff \theta \equiv \psi \ [\pi]$$

3 Première incursion dans les formules sommatoires

3.1 Préliminaires

Définition 8.13 (Sommes et produits)**.** Soit $(a_k)_{k\in\mathbb{N}}$ une suite de nombres complexes. Soit $n \in \mathbb{N}$.

- La **somme pour** k **allant de** 0 **à** n, notée $\sum_{k=0}^{n}$, désigne en fait $a_0 + a_1 + \ldots + a_n$.
- Le **produit pour** k **allant de** 0 **à** n, noté $\prod_{k=0}^{n}$, désigne en fait $a_0 \times a_1 \times \ldots \times a_n$.
- k (ou la **variable muette** utilisée à la place de k) s'appelle **l'indice courant**. Par défaut, l'indice courant augmente de 1 en 1.

On définit de même des sommes et des produits pour k allant de n_0 à $n_1 \geqslant n_0$. Par convention, on pose :

- $\sum_{k=n}^{n-1} a_k = 0$ ("somme vide")
- $\prod_{k=n}^{n-1} a_k = 1$ ("produit vide")

Si $n_1 \leqslant n_0 - 2$, on évite tout simplement de définir ces expressions.

Remarque 8.12 (Définition formelle)**.** En réalité, ces expressions sont définitions par récurrence et la convention sur les sommes et produits vides en est l'initialisation. Pour plus de détails, se référer à la quatrième partie "Sommes et produits" du chapitre "Introduction à l'algèbre".

Remarque 8.13. Il est fondamental de maîtriser la technique de changement de variable ou de changement d'indice : il faut savoir passer d'une somme indexée par k à une somme indexée par $k+1$, $2k$, etc.

Théorème 8.13 (Séparation / regroupement de termes)**.** *Soit* $(a_k)_{k\in\mathbb{N}}$ *et* $(b_k)_{k\in\mathbb{N}}$ *deux suites de nombres complexes. On a :*

$$\forall n \in \mathbb{N},\ \sum_{k=1}^{n}(a_k + b_k) = \left(\sum_{k=1}^{n} a_k\right) + \left(\sum_{k=1}^{n} b_k\right)$$

$$\forall n \in \mathbb{N},\ \prod_{k=1}^{n}(a_k \times b_k) = \left(\prod_{k=1}^{n} a_k\right) \times \left(\prod_{k=1}^{n} b_k\right)$$

Théorème 8.14 (Distributivité)**.** *Soit* $(a_k)_{k\in\mathbb{N}}$ *une suite de nombres complexes. On a :*

$$\forall n \in \mathbb{N},\ \forall \lambda \in \mathbb{C},\ \sum_{k=1}^{n}(\lambda a_k) = \lambda \left(\sum_{k=1}^{n} a_k\right)$$

$$\forall n \in \mathbb{N},\ \forall \lambda \in \mathbb{C},\ \prod_{k=1}^{n}(\lambda a_k) = \lambda^n \left(\prod_{k=1}^{n} a_k\right)$$

Théorème 8.15 (Relation de Chasles)**.** *Soit $(a_k)_{k\in\mathbb{N}}$ une suite de nombres complexes. On a :*

$$\forall (n_0,n_1,n_2)\in\mathbb{N}^3,\ \begin{cases} n_0\leqslant n_1\leqslant n_2 \\ ou \\ n_0\leqslant n_1+1\leqslant n_2 \end{cases} \Longrightarrow \left(\sum_{k=n_0}^{n_1} a_k\right)+\left(\sum_{k=n_1+1}^{n_2} a_k\right)=\sum_{k=n_0}^{n_2} a_k$$

$$\forall (n_0,n_1,n_2)\in\mathbb{N}^3,\ \begin{cases} n_0\leqslant n_1\leqslant n_2 \\ ou \\ n_0\leqslant n_1+1\leqslant n_2 \end{cases} \Longrightarrow \left(\prod_{k=n_0}^{n_1} a_k\right)\times\left(\prod_{k=n_1+1}^{n_2} a_k\right)=\prod_{k=n_0}^{n_2} a_k$$

Proposition 8.24 (Sommes usuelles)**.** *Soit $n\in\mathbb{N}$. On a :*

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

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

$$\sum_{k=0}^{n} k^3=\frac{n^2(n+1)^2}{4}=\left(\sum_{k=0}^{n} k\right)^2$$

Définition 8.14 (Suite arithmétique de raison r)**.** Soit $(u_n)_{n\in\mathbb{N}}\in\mathbb{C}^{\mathbb{N}}$ et $r\in\mathbb{C}$. $(u_n)_{n\in\mathbb{N}}$ est dite **arithmétique de raison** r lorsque $\forall n\in\mathbb{N},\ u_{n+1}=u_n+r$. Dans ce cas, on a :

$$\forall n\in\mathbb{N},\ u_n=u_0+nr$$

Définition 8.15 (Suite géométrique de raison q)**.** Soit $(v_n)_{n\in\mathbb{N}}\in\mathbb{C}^{\mathbb{N}}$ et $q\in\mathbb{C}$. $(v_n)_{n\in\mathbb{N}}$ est dite **géométrique de raison** q lorsque $\forall n\in\mathbb{N},\ v_{n+1}=v_n\times q$. Dans ce cas, on a :

$$\forall n\in\mathbb{N},\ v_n=v_0\times q^n$$

Théorème 8.16 (Somme des termes d'une suite arithmétique)**.** *Soit $(u_n)_{n\in\mathbb{N}}\in\mathbb{C}^{\mathbb{N}}$ une suite arithmétique raison $r\in\mathbb{C}$. On a :*

$$\forall n\in\mathbb{N},\ \sum_{k=0}^{n} u_k=(n+1)\frac{u_0+u_n}{2}$$

Théorème 8.17 (Somme des termes d'une suite géométrique de raison $q\neq 1$)**.** *Soit $(v_n)_{n\in\mathbb{N}}\in\mathbb{C}^{\mathbb{N}}$ une suite géométrique raison $q\in\mathbb{C}\setminus\{1\}$. On a :*

$$\forall n\in\mathbb{N},\ \sum_{k=0}^{n} v_k=v_0\frac{1-q^{n+1}}{1-q}$$

Théorème 8.18 (Formule de Bernoulli)**.** *Soit $(a,b)\in\mathbb{C}^2$ et $n\in\mathbb{N}$. Alors on a la* ***formule de Bernoulli*** *:*

$$a^n-b^n=(a-b)\sum_{k=0}^{n-1} a^{n-1-k}b^k$$

Remarque 8.14 (Cas n impair)**.** Lorsque n est impair, on peut obtenir une formule pour $a^n + b^n$ en effectuant $b \leftarrow -b$ dans la formule de Bernoulli :

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

Exemple 8.2. $a^2 - b^2 = (a-b)(a+b)$

Exemple 8.3. $a^3 - b^3 = (a-b)(a^2+ab+b^2)$

Exemple 8.4. $a^3 + b^3 = (a+b)(a^2-ab+b^2)$

Remarque 8.15. On a aussi $a^n - b^n = \prod_{\omega\in\mathbb{U}_n}(a-\omega b)$ (on verra dans la suite du chapitre que $\mathbb{U}_n$ désigne l'ensemble des racines n-èmes de l'unité).

3.2 Coefficients binomiaux

Définition 8.16 (Factorielle)**.** On pose $\forall n \in \mathbb{N}^\star,\ n! := 1 \times \ldots \times n$ et $0! := 1$ par convention. Plus rigoureusement, cette **factorielle** est définie par récurrence :

$$\begin{cases} 0! := 1 \\ \forall n \in \mathbb{N}, (n+1)! := (n!) \times (n+1) \end{cases}$$

Définition 8.17 (Coefficients binomiaux)**.** Soit $k \in \mathbb{Z}$ et $n \in \mathbb{N}$. On définit le **coefficient binomial** $\binom{n}{k}$ par :

$$\binom{n}{k} := \begin{cases} 0 & \text{si } k < 0 \\ \dfrac{n!}{k!(n-k)!} & \text{si } k \in [\![0,n]\!] \\ 0 & \text{si } k > n \end{cases}$$

Exemple 8.5 (Cas particuliers fréquents)**.** Voici quelques cas particuliers fréquents à connaitre par cœur :

- $\forall n \in \mathbb{N},\ \binom{n}{0} = \binom{n}{n} = 1$
- $\forall n \in \mathbb{N}^\star,\ \binom{n}{1} = \binom{n}{n-1} = n$
- $\forall n \geqslant 2,\ \binom{n}{2} = \binom{n}{n-2} = \dfrac{n(n-1)}{2}$

Proposition 8.25 (Relation de Pascal)**.** *On a la* ***relation de Pascal*** *:*

$$\forall n \in \mathbb{N},\ \forall k \in \mathbb{Z},\ \binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$$

Proposition 8.26 (Symétrie des coefficients binomiaux)**.** *Les coefficients binomiaux sont symétriques au sens suivant :*

$$\forall n \in \mathbb{N},\ \forall k \in \mathbb{Z},\ \binom{n}{k} = \binom{n}{n-k}$$

Proposition 8.27 (Formule du capitaine)**.** *On a la formule dite* ***du capitaine*** *:*

$$\forall (k,n) \in (\mathbb{N}^\star)^2,\ k\binom{n}{k} = n\binom{n-1}{k-1}$$

Théorème 8.19 (Binôme de Newton)**.** *On a la* ***formule du binôme de Newton***

$$\forall n \in \mathbb{N},\ \forall (a,b) \in \mathbb{C}^2,\ (a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$$

Théorème 8.20 (Généralisation : formule du multinôme de Newton, HP)**.** *Il existe une formule plus générale qui généralise le binôme, appelée* ***formule du multinôme de Newton*** *:*

$$\forall n \in \mathbb{N},\ \forall p \in \mathbb{N}^\star,\ \forall (a_i)_{1 \leqslant i \leqslant p} \in \mathbb{C}^p,\ \left(\sum_{i=1}^{p} a_i\right)^n = \sum_{k_1+\ldots+k_p=n} \frac{n!}{k_1! \ldots k_p!} a_1^{k_1} \ldots a_p^{k_p}$$

Démonstration. La preuve peut se faire par récurrence sur p avec une hypothèse en $\forall n \in \mathbb{N}$et en utilisant le binôme. Elle peut aussi se faire par dénombrement. □

Exemple 8.6. $\forall n \in \mathbb{N},\ \sum_{k=0}^{n} \binom{n}{k} = 2^n$

Exemple 8.7. $\forall n \in \mathbb{N},\ \sum_{k=0}^{n} (-1)^k \binom{n}{k} = 0^n = \delta_{n,0}$

Exemple 8.8. $\forall n \in \mathbb{N},\ \forall x \in \mathbb{C},\ \sum_{k=0}^{n} \binom{n}{k} x^k = (x+1)^n$

Exemple 8.9. $\forall n \in \mathbb{N},\ \forall x \in \mathbb{C},\ \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} = 1$

Méthode 8.6 (Calculs de sommes)**.** Voici différentes manières de calculer des sommes :

- On pourra utiliser les formules de Bernoulli, du binôme de Newton, du multinôme de Newton, des sommes de termes d'une suite arithmétique ou géométrique.
- On pourra faire apparaître un télescopage.
- On pourra utiliser les formules du capitaine ou la relation de Pascal pour se ramener à des cas usuels ou des télescopages.
- Si ce sont des sommes qui mettent en jeu des $\cos(k\theta)$ ou des $\sin(k\theta)$, on peut les interpréter comme des parties réelles ou imaginaires de sommes d'exponentielles complexes, puis appliquer le binôme de Newton, la formule de Moivre ou la formule des sommes géométriques à ces sommes exponentielles. On pourra finir par un arc-moitié et il ne faut pas oublier de repasser à la partie réelle ou imaginaire.

- Si la somme dépend d'un paramètre θ défini sur un intervalle, on peut interpréter la somme comme la dérivée d'une autre sommes. On travaille alors sur la somme "primitivée", puis le résultat d'obtient par dérivation. On peut utiliser les autres méthodes pour calculer cette somme "primitivée". On pourra considérer la sommes des $k\sin(k\theta)$ pour $\theta \in]0, 2\pi[$ en guise d'exemple.

4 Racines de l'unité

4.1 Écriture trigonométrique d'un nombre complexe

Lemme 8.1. *Soit $(a, b) \in \mathbb{R}^2$ tel que $a^2 + b^2 = 1$. Alors :*

$$\exists \theta \in \mathbb{R},\ \begin{cases} a = \cos(\theta) \\ b = \sin(\theta) \end{cases}$$

Corollaire 8.3. *Soit $z \in \mathbb{U}$. Alors $\exists \theta \in \mathbb{R},\ z = e^{i\theta}$. Autrement dit, la fonction*

$$\begin{array}{ccc} \mathbb{R} & \to & \mathbb{U} \\ \theta & \mapsto & e^{i\theta} \end{array}$$

est surjective.

Définition 8.18 (Forme trigonométrique)**.** Toute nombre complexe $z \in \mathbb{C}$ admet une **écriture trigonométrique**, ou **forme trigonométrique**, *ie* une écriture de la forme $z = \rho e^{i\theta}$ avec $\rho \in \mathbb{R}_+$ et $\theta \in \mathbb{R}$.

Exemple 8.10. $1 + i = \sqrt{2}e^{i\frac{\pi}{4}}$

Exemple 8.11. $-2024 = 2024e^{i\pi}$

Définition 8.19 (Arguments)**.** Soit $z \in \mathbb{C}^\star$. **Un argument** de z est un $\theta \in \mathbb{R}$ tel qu'il existe $\rho \in \mathbb{R}_+$ tel qu'on ait $z = \rho e^{i\theta}$. Remarquons-qu'on a alors nécessairement $\rho \in \mathbb{R}_+^\star$.

Remarque 8.16. Par convention, 0 n'a pas d'argument.

Lemme 8.2. *Soit $(\alpha, \beta) \in \mathbb{R}^2$. On a*

$$e^{i\alpha} = e^{i\beta} \iff \alpha \equiv \beta\ [2\pi]$$

Théorème 8.21 (Pseudo-unicité des écritures trigonométriques)**.** *L'ensemble des écritures trigonométriques de 0 est l'ensemble ds couples $(0, \theta)$ pour $\theta \in \mathbb{R}$. Soit $z \in \mathbb{C}^\star$. z possède un argument θ_0. L'ensemble des écritures trigonométriques de z est l'ensemble des couples $(|z|, \theta_0 + 2k\pi)$ pour $k \in \mathbb{Z}$*

Proposition 8.28. *Soit $(z, z') \in \mathbb{C}^2$. On a :*

$$e^z = e^{z'} \iff z - z' \in 2i\pi\mathbb{Z}$$

Exemple 8.12. Soit $a \in \mathbb{C}$. Si $a = 0$, l'équation $e^z = a$ n'a pas de solution. Si $a \neq 0$, fixons θ_a un argument de a. L'ensemble des solutions vaut alors $\{\ln|a| + i(\theta_a + 2k\pi) \mid k \in \mathbb{Z}\}$.

Définition 8.20 (Argument principal)**.** Soit $z \in \mathbb{C}^\star$. **L'argument principal de** z, noté $\arg(z)$ (NS), est l'unique argument θ de z tel que $\theta \in]-\pi, \pi]$.

Remarque 8.17. Toujours bien préciser ce que signifie cette notation quand on l'utilise (elle est NS), et aussi préciser qu'on choisit et argument dans $]-\pi, \pi]$ car d'autres le prennent dans $[0, 2\pi[$.

Remarque 8.18. 0 n'a pas d'argument, donc *a fortiori* par d'argument principal.

Proposition 8.29 (Propriété logarithmiques de l'argument principal)**.** *Soit* $(z, z') \in (\mathbb{C}^\star)^2$*. On a :*

- $\arg(zz') \equiv \arg(z) + \arg(z')\ [2\pi]$
- $\arg\left(\dfrac{z}{z'}\right) \equiv \arg(z) - \arg(z')\ [2\pi]$
- $\forall n \in \mathbb{Z},\ \arg(z^n) \equiv n\arg(z)\ [2\pi]$

Remarque 8.19. Attention, il faut bien mettre des $\equiv$ et non des $=$. En effet, par exemple, $n\arg(z)$ à très peu de chances de rester dans $]-\pi, \pi]$ lorsque n est grand.

Théorème 8.22 (Superposition de deux ondes)**.** *Soit* $(\lambda, \mu) \in \mathbb{R}^2 \setminus \{(0,0)\}$*. Alors :*

$$\exists (A, \varphi) \in \mathbb{R}_+^\star \times \mathbb{R},\ \forall t \in \mathbb{R},\ \lambda\cos(t) + \mu\sin(t) = A\cos(t - \varphi)$$

Remarque 8.20. On peut utiliser ce théorème pour résoudre des inéquations ou équations trigonométriques (mais il faut alors expliciter A et φ). Ce théorème sert aussi en physique (kssssss).

4.2 Racines n-èmes

Dans tout le paragraphe, on fixe $n \in \mathbb{N}^\star$.

Théorème 8.23 (Racines n-èmes de l'unité)**.** *L'équation* $z^n = 1$ *admet exactement* n *solutions dans* $\mathbb{C}$*. Ce sont les* $e^{i\frac{2k\pi}{n}}$ *pour* $k \in [\![0, n-1]\!]$*, qu'on appelle les* ***racines*** n***-èmes de l'unité****. On note alors :*

$$\mathbb{U}_n := \left\{ e^{i\frac{2k\pi}{n}} \mid k \in [\![0, n-1]\!] \right\}$$

Remarque 8.21 (Interprétation géométrique)**.** Géométriquement, dans le plan complexe, les racines n-èmes de l'unité sont exactement les sommets d'un n-gone régulier présent sur le cercle unité (dès que $n \geqslant 2$ pour que la notion de n-gone ait du sens).

Remarque 8.22. On a aussi : $\mathbb{U}_n := \left\{ e^{i\frac{2k\pi}{n}} \mid k \in [\![1, n]\!] \right\}$

Exemple 8.13. On a $\mathbb{U}_4 = \{1, i, -1, -i\}$.

Exemple 8.14. Soit $n \geqslant 2$. On a $\displaystyle\sum_{\omega \in \mathbb{U}_n} \omega = 0$.

Définition 8.21 (Nombre j)**.** On pose $j := e^{i\frac{2\pi}{3}}$.

Exemple 8.15. On a $\mathbb{U}_3 = \left\{1, j, j^2\right\}$.

Proposition 8.30 (Quelques propriétés de j)**.** *On a* $1 + j + j^2 = 0$ *et* $j^2 = \overline{j} = \dfrac{1}{j}$.

Méthode 8.7. Soit $u \in \mathbb{C}^\star$ et $n \in \mathbb{N}^\star$. Voici comment trouver les racines n-èmes de u dans $\mathbb{C}$.

- On cherche une **solution particulière**. Pour cela, on écrit u sous forme trigonométrique $u = \rho e^{i\theta}$ et alors $z_0 := \sqrt[n]{\rho} e^{i\frac{\theta}{n}}$ convient comme solution particulière.
- On en déduit toutes les solutions : $z^n = u$ ssi $z^n = z_0^n$ ssi $\left(\frac{z}{z}\right)^n = 1$. Donc z est solution si, et seulement si,
$$\exists k \in [\![0, n-1]\!],\ z = z_0 e^{i\frac{k\theta}{n}}$$
- Conclusion : 0 admet une racine n-ème dans $\mathbb{C}$ qui est 0, et u admet exactement n racines n-èmes dans $\mathbb{C}$.

Remarque 8.23 (Interprétation géométrique)**.** De même, les n racines n-èmes d'un complexe non nul sont exactement les sommets d'un n-gone régulier centré en $(0,0)$ (du moins dès que $n \geqslant 2$).

Méthode 8.8 (Recherche de racines carrées sous forme algébrique)**.** Voici comment chercher les racines carrées sous forme algébrique d'un complexe non nul. On raisonne par analyse-synthèse.

- Analyse : On écrit $z = a + ib$ une solution sous forme algébrique. En mettant au carré et en égalant avec le nombre complexe dont on cherche la racine, on peut obtenir une expression de $a^2 - b^2$ en passant aux parties réelles, une expression de $a^2 + b^2$ en passant aux modules et une expression de ab en passant aux parties imaginaires. On combine les deux premières par demi-somme et demi-différence pour obtenir a et b au signe près. On conclut sur le signe de a et b. On obtient alors deux solutions opposées.
- Synthèse : Réciproquement, puisque le nombre complexe dont on cherche les racines carrées est non nul, il admet deux racines carrées dans $\mathbb{C}$, qui sont nécessairement celles obtenues à la fin de l'analyse.

4.3 Équations du second degré

Théorème 8.24 (Résolution d'une équation du second degré à coefficients complexes)**.** *Soit $(a,b,c) \in \mathbb{C}^\star \times \mathbb{C}^2$. On considère l'équation $az^2 + bz + c = 0$ d'inconnue $z \in \mathbb{C}$. Notons $S_{a,b,c}$ l'ensemble des solutions. On pose $\Delta := b^2 - 4ac$, dit **discriminant** de cette équation. Soit δ une racine carrée de Δ dans $\mathbb{C}$ (si Δ en possède 2, on prend celle que l'on veut cela n'a pas d'importance). Alors,*

$$S_{a,b,c} = \left\{ \frac{-b-\delta}{2a}, \frac{-b+\delta}{2a} \right\}$$

Remarque 8.24. En particulier, si $\Delta = 0$ alors $\delta = 0$ et $S_{a,b,c} = \left\{ \frac{-b}{2a} \right\}$.

Remarque 8.25 (Notations et formulations pour δ)**.** Pour δ, bien dire que δ est **une** racine carrée de Δ dans $\mathbb{C}$ ou que $\delta^2 = \Delta$.

Remarque 8.26 (Notations et formulations pour l'ensemble des solutions)**.** Voici les trois **seules** bonnes manières de conclure la résolution :

- Soit $z \in \mathbb{C}$. z est solution ssi $z = \ldots$ ou $z = \ldots$.
- Les solutions sont données par : $z_1 := \ldots$ et $z_2 := \ldots$.
- L'ensemble des solutions vaut : $\{\ldots;\ldots\}$.

Remarque 8.27. On retiendra qu'il y a toujours au moins une solution. Précisément, si $\Delta = 0$, une seule solution, et si $\Delta \neq 0$, exactement deux solutions.

Corollaire 8.4. *Soit $(a,b,c) \in \mathbb{C}^\star \times \mathbb{C}^2$. On considère l'équation $az^2+bz+c=0$ d'inconnue $z \in \mathbb{C}$. Soit δ une racine carrée du discriminant de cette équation et posons $z_1 := \dfrac{-b-\delta}{2}$ et $z_2 := \dfrac{-b+\delta}{2}$. Alors on a :*

$$\begin{cases} z_1 + z_2 = \dfrac{-b}{a} \\ z_1 z_2 = \dfrac{c}{a} \end{cases}$$

Remarque 8.28. On a coutume de dire que "la somme des racines comptées avec multiplicité vaut $\dfrac{-b}{a}$" et que "le produit des racines comptées avec multiplicité vaut $\dfrac{c}{a}$".

Corollaire 8.5. *Soit $(S,P,u,v) \in \mathbb{C}^4$. Alors $u+v=S$ et $uv=P$ si, et seulement si, u et v sont les deux racines (éventuellement confondues) de l'équation $z^2 - Sz + P = 0$ d'inconnue $z \in \mathbb{C}$.*

Remarque 8.29. Rigoureusement, c'est si, et seulement si, l'ensemble $\{u,v\}$ est égal à l'ensemble des solutions de l'équation $z^2 - Sz + P = 0$ d'inconnue $z \in \mathbb{C}$.

Méthode 8.9 (Résoudre une équation polynomiale de degré > 2)**.** Pour résoudre une équation du type $P(z) = 0$ d'inconnue $z \in \mathbb{C}$ avec P un polynôme à coefficients complexes de degré strictement supérieure à deux, voici la méthode :

- On teste des valeurs simples à la recherche d'une racine $\lambda \in \mathbb{C}$. Dans ce cas, on rappelle que $P(z)$ est factorisable par $z - \lambda$.
- Cela nous permet de nous ramener à une équation de degré inférieur (strictement). On itère jusqu'à arriver à une équation de degré 2, puis on applique les théorèmes qui précèdent.

5 Application à la géométrie

Dans cette partie, on considère le plan affine usuel $\mathcal{P}$, dit "plan des points". On le munit d'un repère orthonormal direct $(O; \overrightarrow{e_1}, \overrightarrow{e_2})$. On note alors $\overrightarrow{\mathcal{P}}$ le "plan des vecteurs" correspondant.

5.1 Affixe

Définition 8.22 (Affixe d'un point)**.** Soit $M \in \mathcal{P}$ de coordonnées (x,y) dans notre repère. L'**affixe de** M est le complexe :

$$z_M := x + iy$$

Définition 8.23 (Affixe d'un vecteur)**.** Soit $\overrightarrow{u}$ un vecteur de coordonnées $\begin{pmatrix} a \\ b \end{pmatrix}$ dans notre repère. L'**affixe de** $\overrightarrow{u}$ est le complexe :

$$z_{\overrightarrow{u}} := a + ib$$

Exemple 8.16. Soit $(A,B) \in \mathcal{P}^2$. On a : $z_{\overrightarrow{AB}} = z_B - z_A$.

Proposition 8.31. *Soit $\overrightarrow{u}, \overrightarrow{v} \in \overrightarrow{\mathcal{P}}$ et $\lambda \in \mathbb{R}$. On a :*

$$\begin{cases} z_{\overrightarrow{u}+\overrightarrow{v}} = z_{\overrightarrow{u}} + z_{\overrightarrow{v}} \\ z_{\lambda\overrightarrow{u}} = \lambda z_{\overrightarrow{u}} \end{cases}$$

Remarque 8.30. Attention, λ **doit** être un **réel** !

Proposition 8.32 (Relation de Chasles)**.** *Soit $(A, B, C) \in \mathcal{P}^3$. On a $z_{\overrightarrow{AB}+\overrightarrow{BC}} = z_{\overrightarrow{AC}}$, soit la* ***relation de Chasles*** $\overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC}$.

5.2 Interprétation géométrique du module

Proposition 8.33. *Soit $\overrightarrow{u} \in \overrightarrow{\mathcal{P}}$. On a : $|z_{\overrightarrow{u}}| = \|\overrightarrow{u}\|$.*

Exemple 8.17. Pour tout $M \in \mathcal{P}$, $|z_M| = OM$.

Définition 8.24 (Cercle)**.** Soit $\Omega \in \mathcal{P}$ et $\mathbb{R} \geqslant 0$. Le **cercle de centre Ω et de rayon** R est l'ensemble :

$$\{M(z) \in \mathcal{P} : |z - z_\Omega| = R\}$$

On dit qu'il "a pour équation" $|z - z_\Omega| = R$.

Définition 8.25 (Disque fermé)**.** Soit $\Omega \in \mathcal{P}$ et $\mathbb{R} \geqslant 0$. Le **disque fermé de centre Ω et de rayon** R est l'ensemble :

$$\{M(z) \in \mathcal{P} : |z - z_\Omega| \leqslant R\}$$

On dit qu'il "a pour équation" $|z - z_\Omega| \leqslant R$.

Définition 8.26 (Disque ouvert)**.** Soit $\Omega \in \mathcal{P}$ et $\mathbb{R} \geqslant 0$. Le **disque ouvert de centre Ω et de rayon** R est l'ensemble :

$$\{M(z) \in \mathcal{P} : |z - z_\Omega| < R\}$$

On dit qu'il "a pour équation" $|z - z_\Omega| < R$.

5.3 Interprétation géométrique des arguments

Théorème 8.25. *Soit $\overrightarrow{u}, \overrightarrow{v}$ des vecteurs* ***non nuls****. Une mesure particulière de l'angle orienté $(\widehat{\overrightarrow{u}, \overrightarrow{v}})$ est donnée par* $\arg\left(\dfrac{z_{\overrightarrow{v}}}{z_{\overrightarrow{v}}}\right)$ *où* arg *désigne l'argument principal. Formellement :*

$$(\widehat{\overrightarrow{u}, \overrightarrow{v}}) \equiv \arg\left(\frac{z_{\overrightarrow{v}}}{z_{\overrightarrow{v}}}\right) \ [2\pi]$$

Autrement dit, les mesures de $(\widehat{\overrightarrow{u}, \overrightarrow{v}})$ sont exactement les arguments de $\dfrac{z_{\overrightarrow{v}}}{\overrightarrow{v}}$.

Corollaire 8.6 (Formule de l'emmerdement maximal)**.** *Soit $(A, B, C, D) \in \mathcal{P}^4$. Alors, une mesure de $(\widehat{\overrightarrow{AB}, \overrightarrow{CD}})$ est donnée par*

$$\arg\left(\frac{z_D - z_C}{z_B - z_A}\right)$$

Remarque 8.31 (Origine du nom du théorème)**.** Le nom quelque peu vulgaire de ce théorème provient de mon professeur de terminale et permet de retenir la formule : les lettres sont à l'envers par rapport à l'ordre de départ, et il y a des signes $-$.

Exemple 8.18. Soit $M \in \mathcal{P}$. Alors $\arg(z_M)$ est une mesure de $(\widehat{\overrightarrow{e_1}, \overrightarrow{OM}})$.

Proposition 8.34 (Relation de Chasles pour les mesures d'angles orientés)**.** *Soit $\overrightarrow{u}, \overrightarrow{v}$ et $\overrightarrow{w}$ des vecteurs non nuls. On peut écrire symboliquement :*

$$(\widehat{\overrightarrow{u}, \overrightarrow{v}}) + (\widehat{\overrightarrow{v}, \overrightarrow{w}}) \equiv (\widehat{\overrightarrow{u}, \overrightarrow{w}}) \ [2\pi]$$

Définition 8.27 (Vecteurs colinéaires)**.** Soit $\overrightarrow{u}, \overrightarrow{v} \in \overrightarrow{\mathcal{P}}$. On dira que $\overrightarrow{u}$ et $\overrightarrow{v}$ sont **colinéaires** lorsque :

- $\overrightarrow{u}$ ou $\overrightarrow{v}$ est le vecteur nul
- $\overrightarrow{u}$ et $\overrightarrow{v}$ sont non nuls et $(\widehat{\overrightarrow{u}, \overrightarrow{v}}) \equiv 0 \ [2\pi]$

Remarque 8.32. La définition donnée coïncide avec les définitions analogues à celles données pour les nombres complexes.

Définition 8.28 (Vecteurs orthogonaux)**.** Soit $\overrightarrow{u}, \overrightarrow{v} \in \overrightarrow{\mathcal{P}}$. On dira que $\overrightarrow{u}$ et $\overrightarrow{v}$ sont **orthogonaux** lorsque :

- $\overrightarrow{u}$ ou $\overrightarrow{v}$ est le vecteur nul
- $\overrightarrow{u}$ et $\overrightarrow{v}$ sont non nuls et $(\widehat{\overrightarrow{u}, \overrightarrow{v}}) \equiv -\dfrac{\pi}{2}$ ou $\dfrac{\pi}{2} \ [2\pi]$

Théorème 8.26 (CNS de colinéarité et d'orthogonalité)**.** *Soit $\overrightarrow{u}, \overrightarrow{v} \in \overrightarrow{\mathcal{P}}$. On a :*

- *$\overrightarrow{u}$ et $\overrightarrow{v}$ sont colinéaires si, et seulement si, $\overline{z_{\overrightarrow{u}}} \times z_{\overrightarrow{v}} \in \mathbb{R}$*
- *$\overrightarrow{u}$ et $\overrightarrow{v}$ sont orthogonaux si, et seulement si, $\overline{z_{\overrightarrow{u}}} \times z_{\overrightarrow{v}} \in i\mathbb{R}$*

Remarque 8.33. En travaillant dans le plan complexe, on peut prouver que les expressions du cosinus, du sinus et de la tangente données en 3^e pour le triangle rectangle sont bien valables au vu de nos définitions actuelles de ces fonctions. Il suffit de placer le sommet à l'angle droit au centre du plan complexe et d'utiliser les formules sur les angles, puis d'utiliser les définitions de cosinus et sinus.

5.4 Quelques transformations du plan

Dans toute cette partie, on exclut les homothéties de rapport 0.

Définition 8.29 (Translation)**.** Soit $\overrightarrow{u}\overrightarrow{\mathcal{P}}$. Soit $M \in \mathcal{P}$. Alors il existe un unique point $M' \in \mathcal{P}$ tel que $\overrightarrow{MM'} = \overrightarrow{u}$. La **translation de vecteur** $\overrightarrow{u}$ est alors la fonction :

$$\begin{array}{rccc} t_{\overrightarrow{u}} & : & \mathcal{P} & \to & \mathcal{P} \\ & & M & \mapsto & M' \end{array}$$

Soit $M(z) \in \mathcal{P}$ et $M'(z') \in \mathcal{P}$ son image. On a : $z' = z + z_{\overrightarrow{u}}$

Définition 8.30 (Homothétie)**.** Soit $\Omega \in \mathcal{P}$ et $\lambda \in \mathbb{R}^\star$. Soit $M \in \mathcal{P}$. Alors il existe un unique point $M' \in \mathcal{P}$ tel que $\overrightarrow{\Omega M'} = \lambda \overrightarrow{\Omega M}$. L'**homothétie de centre Ω et de rapport λ** est alors la fonction :

$$\begin{array}{rccc} h_{\Omega,\lambda} & : & \mathcal{P} & \to & \mathcal{P} \\ & & M & \mapsto & M' \end{array}$$

Soit $M(z) \in \mathcal{P}$ et $M'(z') \in \mathcal{P}$ son image. On a : $z' - z_\Omega = \lambda(z - z_\Omega)$

Définition 8.31 (Rotation)**.** Soit $\Omega \in \mathcal{P}$ et $\theta \in \mathbb{R}^\star$. Soit $M \in \mathcal{P}$. Alors il existe un unique point $M' \in \mathcal{P}$ tel que $\Omega M' = \Omega M$ et $(\widehat{\overrightarrow{\Omega M}, \overrightarrow{\Omega M'}}) \equiv \theta\ [2\pi]$. La **rotation de centre Ω et d'angle θ** est alors la fonction :

$$\begin{array}{rccc} r_{\Omega,\theta} & : & \mathcal{P} & \to & \mathcal{P} \\ & & M & \mapsto & M' \end{array}$$

Soit $M(z) \in \mathcal{P}$ et $M'(z') \in \mathcal{P}$ son image. On a : $z' - z_\Omega = e^{i\theta}(z - z_\Omega)$

Remarque 8.34 (Application affine, partie linéaire, HP)**.** Soit f une des trois transformations précédentes et $\overrightarrow{v} \in \overrightarrow{\mathcal{P}}$ Pour savoir comment f transforme $\overrightarrow{v}$, on voudrait écrire $f(\overrightarrow{v})$. Problème : $\overrightarrow{v} \notin \mathcal{P}$ A défaut, pourrait-on créer $\varphi : \overrightarrow{\mathcal{P}} :\to \overrightarrow{\mathcal{P}}$ qui serait le pendant de f côté $\overrightarrow{\mathcal{P}}$? Pour cela, il faudrait :

- Fixer $(A, B) \in \mathcal{P}^2$ tel que $\overrightarrow{AB} = \overrightarrow{v}$
- Calculer $A' := f(A)$ et $B' := f(B)$
- Poser $\varphi(\overrightarrow{v}) := \overrightarrow{A'B'}$ et vérifier que cette définition ne dépend pas du couple (A, B) choisi.

On appelle alors φ la **partie linéaire** de f et on dit que f est une **application affine**.

Définition 8.32 (Similitude directe)**.** Soit $(a, b) \in \mathbb{C}^\star \times \mathbb{C}$. Une **similitude directe** est une application de la forme :

$$\begin{array}{ccc} \mathcal{P} & \to & \mathcal{P} \\ M(z) & \mapsto & M'(az + b) \end{array}$$

Exemple 8.19. Les translations, homothéties et rotations sont des cas particuliers similitudes directes.

- Translation : $a = 1$ et $b = z_{\overrightarrow{u}}$
- Homothétie : $a = \lambda$ et $b = z_\Omega - \lambda z_\Omega$
- Rotation : $a = e^{i\theta}$ et $b = z_\Omega - e^{i\theta} z_\Omega$

Lemme 8.3. *Soit h une homothétie de rapport $\lambda \in \mathbb{R}^\star$ et r une rotation d'angle $\theta \in \mathbb{R}$ de **même centre** Ω. Alors, h et r commutent, c'est-à-dire que*

$$\forall M \in \mathcal{P},\ h(r(M)) = r(h(M))$$

Définition 8.33. $h \circ r = r \circ h$ s'appelle la **similitude directe de centre Ω, de rapport λ et d'angle θ.**

Méthode 8.10 (Étude d'une similitude directe)**.** Soit $(a, b) \in \mathbb{C}^\star \times \mathbb{C}$. Soit s la similitude directe d'équation $z' = az + b$. Voici comment étudier la similitude s.

- Premier cas : $a = 1$. On reconnaît une translation selon un vecteur d'affixe b.

- Second cas : $a \neq 1$.
  - On cherche un point fixe. Soit $M(z) \in \mathcal{P}$. On a $s(M) = M$ ssi $z = \dfrac{b}{1-a}$. Notons Ω le point d'affixe $\dfrac{b}{1-a}$.
  - Soit $M(z) \in \mathcal{P}$ et $M'(z') \in \mathcal{P}$ son image par s. On a
$$\begin{cases} z' = az + b \\ z_\Omega = az_\Omega + b \end{cases}$$
donc $z' - z_\Omega = a(z - z_\Omega)$
  - On écrit a sous forme trigonométrique $a = \lambda e^{i\theta}$ et on obtient $z' - z_\Omega = \lambda e^{i\theta}(z - z_\Omega)$. Ainsi, s est la similitude directe de centre Ω, de rapport λ et d'angle θ

Remarque 8.35. Dans le lemme qui précède, on avait vu un cas particulier de similitudes directes. Désormais, le point méthode montre que, mises à part les translations, on est toujours dans le cas $h \circ r = r \circ h$. On obtient le centre avec le point fixe et le rapport et la l'angle avec la forme trigonométrique.

Remarque 8.36 (Mise en œuvre pratique du point méthode)**.** En pratique, on pourra ne pas appliquer la deuxième étape du second cas du point méthode et directement passer à la troisième étape après la première, puisque l'on sait que la deuxième étape fournit toujours le même résultat.

Exemple 8.20 (L'ensemble des similitudes directes est stable par composition)**.** Donnons-nous deux similitudes directes s_1 et s_2 d'équations respectives $z' = a_1 z + b_1$ et $z' = a_2 z + b_2$. Étudions $s_1 \circ s_2$. Soit $M \in \mathcal{P}$. On a : $z_{(s_1 \circ s_2)(M)} = a_1 a_2 z_M + (a_1 b_2 + b_1)$ Or, $(a_1 a_2, a_1 b_2 + b_1) \in \mathbb{C}^\star \times \mathbb{C}$, donc $s_1 \circ s_2$ est une similitude directe. Attention, on a $z_{(s_2 \circ s_1)(M)} = a_2 a_1 z_M + (a_2 b_1 + b_2)$ donc les similitudes n'ont aucune raison de commuter : les deux composées ont même rapport et même angle, mais probablement pas même centre.

Définition 8.34 (Symétries d'axes $(O; \overrightarrow{e_1})$ et $(O; \overrightarrow{e_2})$)**.** La symétrie d'axe $(O; \overrightarrow{e_1})$ a pour équation $z' = \overline{z}$. La symétrie d'axe $(O; \overrightarrow{e_2})$ a pour équation $z' = -\overline{z}$.

Définition 8.35 (Similitude indirecte, HP)**.** La composée d'une similitude directe et d'une symétrie axiale s'appelle une **similitude indirecte**.

6 Formulaire de trigonométrie circulaire

Voici un formulaire de trigonométrie circulaire à connaître sur le bout de doigts. On rappelle que s'en déduit le formulaire de trigonométrie hyperbolique par la "règle d'Osborn" (cf. chapitre "Fonctions usuelles", paragraphe "3.2 Trigonométrie hyperbolique").

$$\begin{cases} \cos(-\theta) = \cos(\theta) \\ \sin(-\theta) = -\sin(\theta) \end{cases}$$

$$\begin{cases} \cos(\pi - \theta) = -\cos(\theta) \\ \sin(\pi - \theta) = \sin(\theta) \\ \cos(\pi + \theta) = -\cos(\theta) \\ \sin(\pi + \theta) = -\sin(\theta) \end{cases}$$

$$\begin{cases} \cos\left(\dfrac{\pi}{2}-\theta\right)=\sin(\theta) \\ \sin\left(\dfrac{\pi}{2}-\theta\right)=\cos(\theta) \\ \cos\left(\dfrac{\pi}{2}+\theta\right)=-\sin(\theta) \\ \sin\left(\dfrac{\pi}{2}+\theta\right)=\cos(\theta) \end{cases}$$

$$\begin{cases} \cos(a+b)=\cos(a)\cos(b)-\sin(a)\sin(b) \\ \cos(a-b)=\cos(a)\cos(b)+\sin(a)\sin(b) \\ \sin(a+b)=\sin(a)\cos(b)+\cos(a)\sin(b) \\ \sin(a-b)=\sin(a)\cos(b)-\cos(a)\sin(b) \end{cases}$$

$$\begin{cases} \cos(a)\cos(b)=\dfrac{1}{2}\left(\cos(a-b)+\cos(a+b)\right) \\ \sin(a)\sin(b)=\dfrac{1}{2}\left(\cos(a-b)-\cos(a+b)\right) \\ \sin(a)\cos(b)=\dfrac{1}{2}\left(\sin(a-b)+\sin(a+b)\right) \end{cases}$$

$$\begin{cases} \cos(p)+\cos(q)=2\cos\left(\dfrac{p+q}{2}\right)\cos\left(\dfrac{p-q}{2}\right) \\ \cos(p)-\cos(q)=-2\sin\left(\dfrac{p+q}{2}\right)\sin\left(\dfrac{p-q}{2}\right) \\ \sin(p)+\sin(q)=2\sin\left(\dfrac{p+q}{2}\right)\cos\left(\dfrac{p-q}{2}\right) \\ \cos(p)-\cos(q)=2\cos\left(\dfrac{p+q}{2}\right)\sin\left(\dfrac{p-q}{2}\right) \end{cases}$$

$$\begin{cases} \cos(2a)=2\cos(a)^2-1 \\ \cos(2a)=\cos(a)^2-\sin(a)^2 \\ \cos(2a)=1-2\sin(a)^2 \\ \sin(2a)=2\sin(a)\cos(a) \end{cases}$$

$$\begin{cases} \cos(a)^2=\dfrac{1+\cos(2a)}{2} \\ \sin(a)^2=\dfrac{1-\cos(2a)}{2} \end{cases}$$

$$\begin{cases} \tan(a+b)=\dfrac{\tan(a)+\tan(b)}{1-\tan(a)\tan(b)} \\ \tan(a-b)=\dfrac{\tan(a)-\tan(b)}{1+\tan(a)\tan(b)} \\ \tan(2a)=\dfrac{2\tan(a)}{1-\tan(a)^2} \end{cases}$$

$$t := \tan\left(\frac{\theta}{2}\right) \implies \begin{cases} \cos(\theta) = \dfrac{1-t^2}{1+t^2} \\ \sin(\theta) = \dfrac{2t}{1+t^2} \\ \tan(\theta) = \dfrac{2t}{1-t^2} \end{cases}$$

7 Compléments : retour sur la notion d'angle orienté, HP

La définition que le lycée donne de la mesure d'un angle orientée peut paraître un peu floue. De fait, elle est plus difficile à mettre en place qu'il n'y paraît. C'est ce à quoi nous nous attelons maintenant, avant de démontrer le théorème qui donne une mesure d'un angle en fonction de l'argument d'un certain complexe.

Considérons le cercle unité, c'est-à-dire le cercle de centre O et de rayon 1. Nous avons vu précédemment que tout point du cercle unité pouvait s'écrire $z = e^{i\theta}$. De plus, nous avons vu que la fonction $\theta \mapsto e^{i\theta}$ était continue, 2π-périodique, et nous avons étudié le sens de variation de $\theta \mapsto \mathrm{Re}(e^{i\theta}) = \cos(\theta)$ et $\theta \mapsto \mathrm{Im}(e^{i\theta}) = \sin(\theta)$ sur $[0, 2\pi[$. De tout cela, on déduit que lorsque θ croît de 0 à 2π, le point d'affixe $e^{i\theta}$ parcourt (continûment) le cercle trigonométrique en partant de 1 dans le "sens inverse des aiguilles d'une montre", aussi appelé sens direct.

Théorème 8.27. *Un cercle de rayon r a pour périmètre $2\pi r$.*

Démonstration. Posons un repère au centre du cercle. Par définition, le périmètre du cercle est la limite de la longueur d'une ligne brisée de plus en plus fine le long de sa circonférence. Considérons donc n sommets ($n \geqslant 2$) équirépartis (au sens de la distance) autour du cercle, avec les affixes $e^{i\frac{2k\pi}{n}}$ ($0 \leqslant k \leqslant n-1$). La longueur de la ligne brisée vaut

$$\sum_{k=0}^{n-1} \left| re^{i\frac{2(k+1)\pi}{n}} - re^{i\frac{2k\pi}{n}} \right| = \sum_{k=0}^{n-1} 2r \sin\left(\frac{\pi}{n}\right) = 2rn \sin\left(\frac{\pi}{n}\right)$$

On remarque ensuite que

$$2rn \sin\left(\frac{\pi}{n}\right) = 2\pi r \cdot \frac{\sin\left(\frac{\pi}{n}\right)}{\frac{\pi}{n}} \xrightarrow[n\to+\infty]{} 2\pi r \cdot \sin'(0) = 2\pi r$$

Ainsi la preuve est achevée. □

Maintenant, on va prouver le lien entre angle orienté et argument. Mais au juste, que signifie "mesurer" un angle orienté ? Tout commence avec la notion d'angle géométrique que l'on apprend au collège, voire en primaire : il s'agit d'un couple de deux demi-droites issues du même sommet. Ensuite, comment définit-on la mesure d'un angle géométrique lorsqu'on est écolier ? Intuitivement, c'est "quelque chose qui se mesure avec un rapporteur". Et un rapporteur, ce n'est ni plus ni moins qu'un cercle dont le pourtour est "régulièrement" gradué, au sens de la circonférence. On en arrive donc à la définition suivante :

Définition 8.36. Soit (AB) et (AC) deux demi-droites issues de A. Quel que soit le cercle de centre A, les demi-droites le coupent en deux arcs. La longueur du petit arc (au sens large) divisée par la

longueur du rayon est constante (comme on s'en rend compte en appliquant une homothétie). C'est une grandeur sans dimension comprise entre 0 et π, que l'on appelle la mesure de l'angle $\widehat{BAC}$ (en radians). On peut partir du cercle unité, ce qui donne directement la mesure de l'angle sans avoir à diviser par le rayon.

Remarque 8.37. Dans le cas d'un angle plat, les deux arcs de cercle ont la même longueur, donc peu importe le choix de l'arc, on obtiendra la même mesure.

Exemple 8.21. Un angle droit vaut $\frac{\pi}{2}$ radians.

Définition 8.37. On se donne deux vecteurs $\overrightarrow{u}$ et $\overrightarrow{v}$ non nuls. Pour mesurer l'angle orienté $(\overrightarrow{u}, \overrightarrow{v})$, on mesure l'angle géométrique qu'ils forment, c'est-à-dire qu'on calcule la longueur du petit arc de cercle trigonométrique qui joint $\overrightarrow{u}$ et $\overrightarrow{v}$. On obtient un réel $\theta \in [0, \pi]$. Puis, selon que l'on tourne dans le sens direct ou indirect pour aller de $\overrightarrow{u}$ à $\overrightarrow{v}$ le long de l'arc, on pose $\alpha = \pm\theta$. On obtient $\alpha \in [-\pi, \pi]$, et comme dans le cas de l'argument, on décide que la mesure d'un angle orienté est définie à 2π près. On dit donc que α est une mesure de l'angle.

Remarque 8.38. À vrai dire, pour aller de $\overrightarrow{u}$ à $\overrightarrow{v}$, il est possible de tourner dans les deux sens. Supposons que l'angle géométrique associé au petit arc vaut θ et que l'angle orienté mesure $\alpha = \pm\theta$. Sans perte de généralité, mettons $\alpha = \theta$ pour fixer les idées.

On peut généraliser la mesure d'un angle géométrique aux angles plus grands que π, en convenant que l'angle géométrique associé au grand arc mesure $\theta' = 2\pi - \theta$. En passant aux angles orientés, on remarque que le sens de parcours a été inversé. On obtient donc une nouvelle mesure $\alpha' = -(2\pi - \theta) = \theta - 2\pi$. Ainsi, on retrouve bien la même mesure à 2π près.

Et même possible de faire plus d'un tour pour aller de $\overrightarrow{u}$ à $\overrightarrow{v}$. Dans tous les cas, on trouvera un résultat de la forme $\theta + 2k\pi (k \in \mathbb{Z})$.

Rappelons que la mesure géométrique d'un angle n'est rien d'autre que la longueur d'arc parcourue sur le cercle trigonométrique. Si on convient de noter la longueur avec un signe moins lorsqu'elle est parcourue dans le sens indirect, on obtient la

Définition 8.38 (Définition équivalente)**.** Une mesure de l'angle orienté $(\widehat{\overrightarrow{u}, \overrightarrow{v}})$ est la distance (comptée en valeur algébrique) d'un arc de cercle trigonométrique qui joint $\overrightarrow{u}$ à $\overrightarrow{v}$, y compris en faisant plusieurs tours.

On peut alors prouver le

Théorème 8.28. *Donnons-nous un angle orienté du plan $(\widehat{AB, CD})$ (avec $A \neq B$ et $C \neq D$). Alors* $\arg\left(\dfrac{z_D - z_C}{z_B - z_A}\right)$ *est une mesure de $(\widehat{AB, CD})$*

Démonstration. Pour commencer, on effectue les deux opérations suivantes.

- On ramène $\overrightarrow{AB}$ et $\overrightarrow{CD}$ sur l'origine (c'est-à-dire qu'on suppose sans perte de généralité que $O = A = C$). Ainsi, on se ramène à calculer $\arg(z_D/z_B)$.
- Et on renormalise nos deux vecteurs (c'est-à-dire qu'on suppose $OB = OD = 1$). Pour cela, il suffit de remplacer z_B par $z_B/|z_B|$ et z_D par $z_D/|z_D|$. Cela ne change ni l'angle orienté $(\widehat{\overrightarrow{OB}, \overrightarrow{OD}})$ ni la valeur de $\arg(z_D/z_B)$.

On obtient ainsi deux complexes $z_B = e^{i\theta_1}$ $z_D = e^{i\theta_2}$ sur le cercle trigonométrique. Quitte à ajouter un multiple de 2π à θ_2, ce qui ne changera pas la position de D et laissera donc l'angle $(\widehat{\overrightarrow{OB}, \overrightarrow{OD}})$ inchangé, on suppose $\theta_1 \leq \theta_2 < \theta_1 + 2\pi$. Ainsi, on peut compter la longueur de l'arc de cercle positivement. Puis comme précédemment, il s'agit de calculer la limite d'une ligne brisée qui joint $e^{i\theta_1}$ à $e^{i\theta_2}$.

$$\begin{aligned}\sum_{k=0}^{n-1} \left| e^{i\left(\theta_1 + \frac{(k+1)(\theta_2-\theta_1)}{n}\right)} - e^{i\left(\theta_1 + \frac{k(\theta_2-\theta_1)}{n}\right)} \right| &= \sum_{k=0}^{n-1} 2 \left| \sin\left(\frac{\theta_2-\theta_1}{2n}\right) \right| \\ &= 2n \left| \sin\left(\frac{\theta_2-\theta_1}{2n}\right) \right| \\ &= 2n \sin\left(\frac{\theta_2-\theta_1}{2n}\right) \\ &\xrightarrow[n\to\infty]{} \theta_2 - \theta_1\end{aligned}$$

Donc une mesure de l'angle orienté est donnée par $\theta_2 - \theta_1$ et par ailleurs on a bien

$$\arg(z_D/z_B) = \arg\left(e^{i(\theta_2-\theta_1)}\right) \equiv \theta_2 - \theta_1 \ [2\pi]$$

Ainsi la preuve est achevée. □

Chapitre 9

Espaces vectoriels

Nous considérons dans tout le chapitre un corps $\mathbb{K}$. En pratique, ce corps sera souvent égal à $\mathbb{R}$ ou $\mathbb{C}$, mais la quasi-totalité des résultats restent vrais pour un corps quelconque.

1 Premières définitions

Au lycée, on apprend que $\mathbb{R}^2$ ou $\mathbb{R}^3$ peut être vu comme un ensemble de *vecteurs*, sur lesquels on peut effectuer des *combinaisons linéaires*. Plus généralement, nous avons la

Définition 9.1 ($\mathbb{K}$-espace vectoriel)**.** Un **$\mathbb{K}$-espace vectoriel** ou **$\mathbb{K}$-ev** est un ensemble E muni d'une loi de composition interne $+$ et d'une loi de composition externe :

$$\begin{array}{ccc} \mathbb{K} \times E & \to & E \\ (\lambda, x) & \mapsto & \lambda x \end{array}$$

vérifiant les axiomes suivants.

- $(E, +)$ est un groupe commutatif.
- On a une pseudo-distributivité :

$$\begin{cases} \forall (\lambda, x, y) \in \mathbb{K} \times E^2,\ \lambda(x+y) = \lambda x + \lambda y \\ \forall (\lambda, \mu, x) \in \mathbb{K}^2 \times E,\ (\lambda + \mu)x = \lambda x + \mu x \end{cases}$$

- On a une pseudo-associativité :

$$\forall (\lambda, \mu, x) \in \mathbb{K}^2 \times E,\ (\lambda\mu)x = \lambda(\mu x)$$

- On a une propriété d'opérateur neutre :

$$\forall x \in E,\ 1x = x$$

Les éléments de E sont appelés les **vecteurs**, et ceux de $\mathbb{K}$ les **scalaires**.

Remarque 9.1. À part dans certains cas lorsqu'il pourra y avoir ambiguïté, on cessera de noter les vecteurs avec des flèches.

Remarque 9.2. Si E est un $\mathbb{C}$-ev, alors *a fortiori* c'est un $\mathbb{R}$-ev, et même un $\mathbb{Q}$-ev.

Remarque 9.3 ($\mathbb{Z}$-module, HP)**.** Ces quatre axiomes ont déjà été vus lorsque nous avons introduit la loi externe $n \cdot x$ dans un anneau A (avec $n \in \mathbb{Z}$ et $x \in A$). Plus généralement, ils restent vrais sur $\mathbb{Z} \times E$ dès que $(E, +)$ est un groupe commutatif. La seule différence est que $\mathbb{Z}$ n'est pas un corps, mais un anneau. C'est pour cela qu'on ne parle pas de $\mathbb{Z}$-ev mais de $\mathbb{Z}$-module.

Définition 9.2 (Combinaison linéaire)**.** Soit $x_1, ..., x_n$ des vecteurs de E. Une **combinaison linéaire** des x_i est une expression de la forme $\sum_{i=1}^{n} \lambda_i x_i$, où les λ_i sont des scalaires.

Remarque 9.4 (Linéarité de la somme)**.** Par une récurrence immédiate, on montre que la pseudo-distributivité se généralise :

$$\lambda \sum_{i=1}^{n} x_i + \mu \sum_{i=1}^{n} y_i = \sum_{i=1}^{n} (\lambda x_i + \mu y_i)$$

On montre tout de suite quelques propriétés élémentaires.

Exemple 9.1. Si $x = \lambda y$ avec $\lambda \neq 0$, alors $y = \dfrac{1}{\lambda} x$

Proposition 9.1 (Pseudo-absorbance, pseudo-intégrité)**.** $\forall (\lambda, x) \in \mathbb{K} \times E,\ \lambda x = 0 \Leftrightarrow \lambda = 0_{\mathbb{K}}$ *ou* $x = 0_E$.

Corollaire 9.1. $\forall (\lambda, x) \in \mathbb{K} \times E,\ (-\lambda)x = \lambda(-x) = -(\lambda x)$

Exemple 9.2. Si $\lambda u = \mu u$ avec $u \neq 0_E$, alors $\lambda = \mu$.

Remarque 9.5 (Itération de la loi + et l.c.e.)**.** On pourra montrer le résultat suivant :

$$\forall (n, \lambda, x) \in \mathbb{Z} \times \mathbb{K} \times E,\ (n\lambda)x = \lambda(nx) = n(\lambda x)$$

En particulier, on pourra écrire $n\lambda x$ sans risque d'être ambigu. De même, si $\mathbb{K}$ est un sur-ensemble de $\mathbb{Z}$, on pourra vérifier que la notation λx désigne indifféremment l'itération du $+$ de E, ou la l.c.e.

Définition 9.3 (Colinéarité)**.** Soit x, y deux vecteurs. Les trois assertions suivantes sont équivalentes :

1. $\exists \lambda \in \mathbb{K}, y = \lambda x$ ou $\exists \mu \in \mathbb{K}, x = \mu y$
2. $\exists \lambda \in \mathbb{K}, y = \lambda x$ ou $x = 0$,
3. $\exists \mu \in \mathbb{K}, x = \mu y$ ou $y = 0$.

Dans ce cas, on dit que x et y sont **colinéaires**. Bien sûr, par symétrie de définition, x et y sont colinéaires ssi y et x sont colinéaires.

Exemple 9.3. Le vecteur nul est colinéaire à tout vecteur.

Remarque 9.6. Attention : à moins de travailler avec $\mathbb{K} = \mathbb{Q}$ ou $\mathbb{R}$, cela n'a aucun sens de parler de colinéarité directe.

Exemple 9.4. $\mathbb{K}$ lui-même est un $\mathbb{K}$-ev.

Proposition 9.2 (Espace vectoriel produit)**.** *Le produit cartésien $E_1 \times E_2$ de deux espaces vectoriels E_1 et E_2 muni des lois suivantes :*

$$\begin{cases} (x_1, x_2) + (x_1', x_2') = (x_1 + x_1', x_2 + x_2') \\ \lambda(x_1, x_2) = (\lambda x_1, \lambda x_2) \end{cases}$$

est un espace vectoriel.

Remarque 9.7. On peut généraliser ce résultat à un produit quelconque (même infini) d'espaces vectoriels.

Exemple 9.5 (Structure canonique de $\mathbb{K}$-ev de $\mathbb{K}^n$)**.** Pour $n \in \mathbb{N}^*$. $\mathbb{K}^n$ est un espace vectoriel pour les lois suivantes.

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{pmatrix} = \begin{pmatrix} x_1 + x_1' \\ x_2 + x_2' \\ \vdots \\ x_n + x_n' \end{pmatrix}$$

et

$$\lambda \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Il s'agit simplement d'une application de la remarque précédente et du fait que $\mathbb{K}$ est lui-même un $\mathbb{K}$-ev. On dit qu'on a muni $\mathbb{K}^n$ de sa **structure canonique** d'espace vectoriel.

Exemple 9.6. Soit X un ensemble quelconque et E un espace vectoriel. Un autre exemple très important est fourni par E^X muni des lois suivantes.

$$\begin{array}{rccccc} f+g & : & X & \mapsto & E \\ & & x & \to & f(x)+g(x) \end{array}$$

et

$$\begin{array}{rccccc} \lambda f & : & X & \mapsto & E \\ & & x & \to & \lambda f(x) \end{array}$$

Vérifications sur la loi $+$.

- L'associativité et la commutativité s'héritent de celles dans $(E, +)$.
- Elément neutre : c'est l'application $x \mapsto 0$,
- Opposé de : c'est l'application $x \mapsto -f(x)$,

Vérifications sur la loi externe.

- La pseudo-distributivité et la pseudo-associativité s'héritent de celles dans E.
- L'opérateur neutre est $1_{\mathbb{K}}$

Remarque 9.8. On pourra utiliser largement ce dernier exemple lorsqu'il s'agira de montrer que tel ensemble de fonctions est un espace vectoriel. Bien souvent, il suffira de montrer que c'est un sous-espace vectoriel d'un certain E^X ce qui sera toujours plus rapide.

Remarque 9.9. Un cas particulier est donné par $X = \mathbb{N}$. E^X est alors l'ensemble des suites à valeurs dans E, et c'est donc un espace vectoriel. Pour $X = [\![1, n]\!]$, on retrouve $\mathbb{K}^n$ muni de sa structure canonique.

2 Sous-espaces

Définition 9.4 (Sous-espace vectoriel)**.** Un **sous-espace vectoriel** ou **sev** de E est une partie $F \subset E$ telle que

- $(F, +)$ est un sous-groupe de $(E, +)$
- $\forall (\lambda, x) \in F, \lambda x \in F$ (stabilité par la l.c.e.).

Remarque 9.10. Soit F un sev de E et $n \in \mathbb{N}$, par une récurrence immédiate, F est stable par toute combinaison linéaire de n vecteurs.

Proposition 9.3 (Recette pratique pour montrer que F est un sev de E)**.** *Soit E un K-ev et $F \subset E$. Pour que F soit un sev de E, il faut et il suffit que*

- *$F \neq \emptyset$ (par exemple $0 \in F$)*
- *$\forall (\lambda, \mu, x, y) \in \mathbb{K}^2 \times F^2, \lambda x + \mu y \in F$ (stabilité par combinaison linéaire de deux vecteurs)*

Remarque 9.11. En pratique, on utilise la condition suffisante pour montrer que F est un sev de E avec E un espace vectoriel de référence, ce qui prouve que F un espace vectoriel.

Remarque 9.12. De manière équivalente, le deuxième point de la recette peut être remplacé par

$$\forall (\lambda, x, y) \in \mathbb{K} \times F^2,\ \lambda x + y \in F$$

Exemple 9.7. Dans le $\mathbb{R}$-ev des suites réelles, les suites bornées forment un sev, à l'intérieur duquel les suites convergentes forment un sev, à l'intérieur duquel les suites qui convergent vers 0 forment un autre sev.

Proposition 9.4. *Soit $\mathcal{F}$ un ensemble de sev de E. Alors $\bigcap_{F \in \mathcal{F}} F$ est un sev de E.*

Définition 9.5 (Sous-espace vectoriel engendré)**.** Soit A une partie de E. Alors le **sous-espace engendré par** A, noté $\text{Vect}(A)$, est l'intersection de tous les sev de E contenant A. C'est aussi le plus petit sev contenant A. Donc si un sev contient A, alors il contient $\text{Vect}(A)$.

Remarque 9.13. Si $A = \{x\}$, $\text{Vect}(\{x\})$ pourra plus simplement être noté $\text{Vect}(x)$.

Exemple 9.8. On a $\text{Vect}(0) = \{0\}$ et $\text{Vect}(E) = E$. Si F est un sev de E on a $\text{Vect}(F) = F$.

Proposition 9.5 (Vect d'une partie finie)**.** *Si $A = \{x_1, ..., x_n\}$, alors $\text{Vect}(A)$ est l'ensemble des combinaisons linéaires des éléments de A :*

$$\text{Vect}(A) = \{\lambda_1 x_1 + ... + \lambda_n x_n \mid \lambda_1, ..., \lambda_n \in \mathbb{K}\}$$

Exemple 9.9. Voici quelques cas très simples.

- Soit $x \in E$. On a $\text{Vect}(x) = \{\lambda x \mid \lambda \in \mathbb{K}\}$. On pourra le noter $\mathbb{K}x$. Si de plus $x \neq 0$, $\text{Vect}(x)$ est appelé une **droite vectorielle**. Géométriquement, on peut le représenter comme une droite qui passe par l'origine.
- Soit $(x, y) \in E^2$. On a $\text{Vect}(\{x, y\}) = \{\lambda x + \mu y \mid (\lambda, \mu) \in \mathbb{K}^2\}$. On pourra le noter $\mathbb{K}x + \mathbb{K}y$. Si x et y sont non colinéaires, $\text{Vect}(\{x, y\})$ est appelé un **plan vectoriel**. Géométriquement, on peut le représenter comme un plan qui passe par l'origine.

3 Applications linéaires

De même qu'un morphisme de groupe préserve la loi sous-jacente, ou qu'un morphisme d'anneaux ou de corps préserve les lois sous-jacentes, on a la

Définition 9.6 (Application linéaire, ensemble $\mathcal{L}(E,F)$)**.** Une **application linéaire** est un morphisme f d'un espace vectoriel E dans un espace vectoriel F, c'est-à-dire qu'elle vérifie

$$\forall(\lambda,\mu,x,y)\in\mathbb{K}^2\times E^2,\ f(\lambda x+\mu y)=\lambda f(x)+\mu f(y)$$

Si f est bijective, on rappelle qu'elle s'appelle un **isomorphisme**. Si $E=F$, on rappelle que f s'appelle un **endomorphisme**. Et si f est à la fois un isomorphisme et un endomorphisme, c'est un **automorphisme**.

L'ensemble des applications linéaires de E dans F est noté $\mathcal{L}(E,F)$. L'ensemble des endomorphismes de E est plus simplement noté $\mathcal{L}(E)$.

Remarque 9.14. De manière équivalente, on peut demander que f vérifie

$$\forall(\lambda,x,y)\in\mathbb{K}\times E^2,\ f(\lambda x+y)=\lambda f(x)+f(y)$$

Remarque 9.15. Soit $f\in\mathcal{L}(E,F)$. Par une récurrence immédiate, on vérifie que, plus généralement

$$\forall(\lambda_1,...,\lambda_n)\in\mathbb{K}^n,\ \forall(x_1,...,x_n)\in E^n,\ f\left(\sum_{i=1}^{n}\lambda_i x_i\right)=\sum_{i=1}^{n}\lambda_i f(x_i)$$

Exemple 9.10. Prenons $E=F=\mathbb{R}^2$. Chaque vecteur de $\mathbb{R}^2$ s'exprime d'une unique manière comme combinaison linéaire de $e_1=\begin{pmatrix}1\\0\end{pmatrix}$ et $e_2=\begin{pmatrix}0\\1\end{pmatrix}$:

$$x=\lambda e_1+\mu e_2$$

On en déduit $f(x)=\lambda f(e_1)+\mu f(e_2)$. En particulier, le maillage $\{me_1+ne_2\mid(m,n)\in\mathbb{Z}^2\}$ se transforme en $\{mf(e_1)+nf(e_2)\mid(m,n)\in\mathbb{Z}^2\}$. Plus généralement, il faut s'imaginer que f "déforme" continûment un maillage infiniment fin. f ne préserve pas forcément les longueurs ni les angles, mais elle préserve au moins les parallélismes.

Remarque 9.16. On a les deux résultats suivants.

- Si $f\in\mathcal{L}(E,F)$ et si on considère $0\in\mathcal{L}(E,E)$, alors $f\circ 0=0$.
- Et bien sûr, si $f\in\mathcal{L}(E,F)$ et si on considère $0\in\mathcal{L}(F,G)$, alors $0\circ f=0$ (mais ce dernier résultat n'utilise pas la linéarité).

Remarque 9.17. En particulier, f est un morphisme du groupe $(E,+)$ dans le groupe $(F,+)$. Cela permet de définir $\ker(f)$ et de caractériser l'injectivité de f. On en déduit aussi que

$$\begin{cases}f(0)=0\\\forall x\in E, f(-x)=-f(x)\end{cases}$$

(mais de toute façon, on le voit déjà par linéarité).

Proposition 9.6. *Soit $f \in \mathcal{L}(E,F)$. L'image d'un sous-espace vectoriel de E est un sous-espace vectoriel de F, et l'image réciproque d'un sous-espace vectoriel de F est un sous-espace vectoriel de E.*

Corollaire 9.2. *Soit $f \in \mathcal{L}(E,F)$. Alors* $\ker(f)$ *et* $\mathrm{Im}(f)$ *sont des sous-espaces vectoriels de E et F respectivement.*

Proposition 9.7. *$\mathcal{L}(E,F)$ est un $\mathbb{K}$-espace vectoriel. En particulier, $\mathcal{L}(E)$ est un $\mathbb{K}$-espace vectoriel.*

Lemme 9.1. *Si $f \in \mathcal{L}(F,G)$ et $g \in \mathcal{L}(E,F)$, alors $f \circ g \in \mathcal{L}(E,G)$. De plus, la composition est bilinéaire au sens suivant :*

$$\begin{cases} \forall(\lambda,\mu) \in \mathbb{K}^2,\ \forall(f,g,h) \in \mathcal{L}(F,G)^2 \times \mathcal{L}(E,F),\ (\lambda f + \mu g) \circ h = \lambda(f \circ h) + \mu(g \circ h) \\ \forall(\lambda,\mu) \in \mathbb{K}^2,\ \forall(f,g,h) \in \mathcal{L}(F,G) \times \mathcal{L}(E,F)^2,\ f \circ (\lambda g + \mu h) = \lambda(f \circ g) + \mu(f \circ h) \end{cases}$$

Exemple 9.11. En particulier, on a $(\lambda f) \circ g = f \circ (\lambda g) = \lambda(f \circ g)$.

Proposition 9.8. *On pourra montrer qu'on a toujours*

$$\begin{cases} Im(f \circ g) \subset Im(f) \\ Ker(f \circ g) \supset Ker(g) \end{cases}$$

De même, si $f \circ g = 0$ on montrera que

$$Im(g) \subset Ker(f)$$

Corollaire 9.3. *$(\mathcal{L}(E), +, \circ)$ est un anneau (non commutatif en général).*

Remarque 9.18. On s'entraînera avec profit à réaliser directement des calculs sur les endo- morphismes, à la manière des éléments d'un anneau, sans passer par leurs images élément par élément. Dans ce contexte, f^n désignera la n-ième itérée de f au sens de la loi $\circ$.

Si f et g commutent, on rappelle que la formule du binôme est toujours valable :

$$(f+g)^m = \sum_{k=0}^{m} \binom{m}{k} f^k g^{m-k}$$

ainsi que la formule de Bernoulli :

$$f^m - g^m = (f-g)(\sum_{k=0}^{m-1} f^k g^{m-k-1}) = (\sum_{k=0}^{m-1} f^k g^{m-k-1})(f-g)$$

En particulier, on utilisera régulièrement ces formules avec $g = \mathrm{id}$.

Exemple 9.12 (Suites récurrentes linéaires d'ordre 2)**.** Considérons les suites $(u_n)_{n\in\mathbb{N}}$ à valeurs complexes qui vérifient une relation de récurrence du type

$$u_{n+2} + a u_{n+1} + b u_n = 0$$

avec $b \neq 0$. Alors l'ensemble E de ces suites forme un sev de $\mathbb{C}^{\mathbb{N}}$ comme on peut le vérifier directement par le calcul. Mais une autre manière de voir les choses consiste à considérer l'application

$$\begin{array}{ccc} \varphi : \mathbb{C}^{\mathbb{N}} & \to & \mathbb{C}^{\mathbb{N}} \\ (u_n)_{n\in\mathbb{N}} & \mapsto & (u_{n+1})_{n\in\mathbb{N}} \end{array}$$

qui est linéaire. E est alors le noyau de l'application $\varphi^2 + a\varphi + b\,\mathrm{id}$, qui est linéaire.

Définition 9.7 ($\mathbb{K}$-algèbre)**.** On dit que $(A, +, \cdot, \times)$ est une **algèbre** sur le corps $\mathbb{K}$, ou $\mathbb{K}$-algèbre lorsque

- $(A, +, \cdot)$ est un K-ev ;
- $(A, +, \times)$ est un anneau :
- la deuxième loi de l'anneau est compatible avec la loi externe :

$$\lambda(x \times y) = (\lambda x) \times y = x \times (\lambda y)$$

Exemple 9.13. $(\mathcal{L}(E), +, \cdot, \times)$ est une $\mathbb{K}$-algèbre. $(\mathbb{K}, +, \cdot, \times)$ est une $\mathbb{K}$-algèbre. On verra d'autres exemples de $\mathbb{K}$-algèbres dans les chapitres sur les polynômes et les matrices.

Définition 9.8 (Sous-algèbre)**.** Si A est une algèbre, une **sous-algèbre** de A sera donc une partie de A qui est à la fois un sev et un sous-anneau de A. Il suffira de vérifier que c'est un sev (par la recette), qu'elle est stable par produit et qu'elle contient 1. Comme c'est un sev elle sera stable par différence, et ce sera donc aussi un sous-anneau. La compatibilité s'héritera automatiquement.

Remarque 9.19 ($\mathbb{Z}$-algèbre, HP)**.** Si $(A, +, \times)$ est un anneau, en considérant la loi externe $n \cdot x$ sur $\mathbb{Z} \times A$ nous avons vu qu'elle était également compatible (cf. chapitre "Groupes, anneaux, corps). Une fois de plus, la seule différence est que $\mathbb{Z}$ n'est pas un corps, mais un anneau. A cause de cela, on dira que A est une algèbre sur l'anneau $\mathbb{Z}$, ou $\mathbb{Z}$-algèbre. On peut résumer les choses dans le petit tableau ci-contre.

	pseudo-associativité pseudo-distributivité à gauche et à droite opérateur neutre	axiomes précédents + compatibilité
$\mathbb{Z}$	$\mathbb{Z}$-module exemple : groupe$(G,+)$commutatif	$\mathbb{Z}$-algèbre exemple : anneau$(A,+,\times)$
$\mathbb{K}$	$\mathbb{K}$-ev	$\mathbb{K}$-algèbre

Lemme 9.2. *Si f est un isomorphisme de E dans F, alors f^{-1} est un isomorphisme de F dans E.*

Définition 9.9 (Espaces vectoriels isomorphes)**.** Soit E et F deux $\mathbb{K}$-ev. Alors il existe un isomorphisme de E dans F ssi il existe un isomorphisme de F dans E. Dans ce cas, on dit que E et F sont **isomorphes**. On pourra écrire $E \simeq F$.

Définition 9.10 (Groupe linéaire)**.** L'ensemble des automorphismes de E est noté $GL(E)$. C'est un groupe pour la composition. On l'appelle le **groupe linéaire** de E.

4 Familles de vecteurs

4.1 Cas particulier des familles finies

Afin de donner une meilleure intuition à l'étudiant, nous commençons par traiter le cas de familles finies de vecteurs $(e_i)_{1\leqslant i\leqslant n}$. Nous verrons ensuite ce qu'il en est du cas général. On cherche maintenant à exprimer tout vecteur de E comme combinaison linéaire d'une famille donnée, afin de voir si celle-ci engendre tout l'espace ou non. Soit donc une famille finie $(e_i)_{1\leqslant i\leqslant n}$, et considérons

$$\begin{array}{ccc} \varphi : \mathbb{K}^n & \to & E \\ (\lambda_i)_{1\leqslant i\leqslant n} & \mapsto & \displaystyle\sum_{i=1}^{n} \lambda_i e_i \end{array}$$

On va s'intéresser de près à la surjectivité et/ou l'injectivité de φ. Pour cela, il est plus simple de vérifier d'abord le

Lemme 9.3. *φ est linéaire.*

Définition 9.11 (Famille génératrice)**.** La famille $(e_i)_{1\leqslant i\leqslant n}$ est **génératrice** si tout vecteur de E peut s'écrire comme une combinaison linéaire des e_i :

$$\forall x \in E,\ \exists(\lambda_1,\dots,\lambda_n) \in \mathbb{K}^n,\ x = \sum_{i=1}^{n} \lambda_i e_i$$

Autrement dit, φ est surjective.

Exemple 9.14. Si u est un vecteur non nul, (u) est une famille génératrice de la droite vectorielle Ku. Si u_1 et u_2 sont deux vecteurs non colinéaires de l'espace, (u_1, u_2) est une famille génératrice du plan vectoriel $\mathbb{K}u_1 + \mathbb{K}u_2$. La famille vide engendre $\{0\}$.

On s'intéresse maintenant à la question de savoir si l'écriture comme une combinaison linéaire des e_i est unique.

Définition 9.12 (Famille libre)**.** La famille $(e_i)_{1\leqslant i\leqslant n}$ est libre lorsqu'il y a unicité d'écriture dans $\displaystyle\sum_i \lambda_i e_i$:

$$\forall(\lambda_1,\dots,\lambda) \in \mathbb{K}^n,\ \forall(\mu_1,\dots,\mu_n) \in \mathbb{K}^n,\ \left(\sum_{i=1}^{n} \lambda_i e_i = \sum_{i=1}^{n} \mu_i e_i \implies \forall i \in [\![1,n]\!], \lambda_i = \mu_i\right)$$

Autrement dit, φ est injective. On dit aussi que les e_i sont **linéairement indépendants**. Une famille qui n'est pas libre est dite **liée**.

Remarque 9.20. Attention, lorsque la famille est libre, il y a unicité mais pas forcément existence !

Proposition 9.9 (Caractérisation des familles libres)**.** *La famille $(e_i)_{1\leqslant i\leqslant n}$ est libre si et seulement si*

$$\forall(\lambda_i) \in \mathbb{K}^n,\ \left(\sum_{i=1}^{n} \lambda_i e_i = 0 \Rightarrow \forall i \in [\![1,n]\!],\ \lambda_i = 0\right)$$

Au contraire, la famille est liée si et seulement si

$$\exists(\lambda_i) \in \mathbb{K}^n \backslash \{0\},\ \sum_{i=1}^{n} \lambda_i e_i = 0$$

Exemple 9.15. La famille (x, y) est liée si et seulement si x et y sont colinéaires.

Remarque 9.21. Dès qu'une famille contient le vecteur nul, elle est liée.

Proposition 9.10 (Caractérisation des familles liées)**.** *La famille $(e_i)_{1 \leqslant i \leqslant n}$ est liée si et seulement si l'un de ses vecteurs peut s'écrire combinaison linéaire des autres.*

Remarque 9.22. En un certain sens, une famille liée est une famille "redondante". En effet, supposons qu'on ait par exemple

$$e_{i_0} = \sum_{i \neq i_0}^{n} \frac{\lambda_i}{\lambda_{i_0}} e_i$$

Alors toute combinaison linéaire des $(e_i)_{1 \leqslant i \leqslant n}$ peut se récrire comme combinaison linéaire des $(e_i)_{i \neq i_0}$ en redistribuant le coefficient de e_{i_0} sur les autres e_i.

$$\sum_{i=1}^{n} \mu_i e_i = \sum_{i \neq i_0} \left(\mu_i - \frac{\mu_{i_0} \lambda_i}{\lambda_{i_0}} \right) e_i$$

Ainsi, e_{i_0} est "inutile".

Le caractère générateur est donc une affaire *d'existence*, le caractère libre une affaire *d'unicité* et le caractère lié une affaire de *redondance*.

Définition 9.13 (Base)**.** Si $(e_i)_{1 \leqslant i \leqslant n}$ est à la fois libre et génératrice, alors c'est une **base** de E. De manière équivalente, on peut dire que tout vecteur s'écrit d'une unique manière comme combinaison linéaire des e_i :

$$\forall x \in E, \exists!(\lambda_1, \ldots, \lambda_n) \in \mathbb{K}^n,\ x = \sum_{i=1}^{n} \lambda_i e_i$$

De manière encore équivalente, φ est bijective. Les λ_i s'appellent les **coordonnées de x dans la base** $(e_i)_{1 \leqslant i \leqslant n}$.

Exemple 9.16. Dans $\mathbb{R}^2$, (e_1, e_2) est une base.

Plu généralement, on a la

Définition 9.14 (Base canonique de $\mathbb{K}^n$)**.** La **base canonique** de $\mathbb{K}^n$ est définie par :

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots, \quad e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

Remarque 9.23. Puisque φ est un isomorphisme, on en déduit que φ^{-1} est aussi un isomorphisme, et en particulier c'est une application linéaire. Ainsi, si les coordonnées de x sont (λ_i) et celles de y sont (μ_i), alors celles de $\alpha x + \beta y$ sont $(\alpha\lambda_i + \beta\mu_i)$. Nous reviendrons plus longuement sur ce point au moment des formes linéaires.

4.2 Généralisation aux familles quelconques

Définition 9.15 (Famille presque nulle, support, somme presque nulle)**.** Une famille de vecteurs $(x_i)_{i\in I}$ est **presque nulle** lorsque tous les x_i sont nuls sauf un nombre fini d'entre eux. L'ensemble

$$S = \{i \in I \mid x_i \neq 0\}$$

s'appelle le **support** de la famille $(x_i)_{i\in I}$. Autrement dit, une famille est presque nulle lorsque son support est fini. Dans ce cas, la **somme presque nulle** $\sum_{i\in I} x_i$ désigne en fait $\sum_{i\in S} x_i$.

Remarque 9.24 (Ensemble intermédiaire fini)**.** Soit S le support de $(x_i)_{i\in I}$ et T un ensemble fini tel que $S \subset T \subset I$. Alors

$$\sum_{i\in I} x_i = \sum_{i\in T} x_i$$

En effet, $(x_i)_{i\in I}$ et $(x_i)_{i\in T}$ ont alors même support.

Définition 9.16 (Ensemble des familles presque nulle)**.** L'ensemble des familles presque nulles de vecteurs de E indexées sur I est noté $E^{(I)}$.

Proposition 9.11. *$E^{(I)}$ est un sev de E^I.*

Exemple 9.17. $\mathbb{K}^{(I)}$ est un sev de $\mathbb{K}^I$.

Théorème 9.1 (Opérations sur les sommes presque nulles)**.** *Toutes les opération usuelles sur les sommes finies restent licites avec les sommes presque nulles :*

- *Linéarité de la somme*
- *Changement de variable*
- *Associativité de la somme*
- *Théorème de Fubini*
- *Distributivité généralisée dans une $\mathbb{K}$-algèbre*

Remarque 9.25. En revanche, attention : pour ce qui est des sommes qui ne sont pas presque nulles, la théorie est beaucoup plus contraignante. On pourra se référer au chapitre "Sommabilité".

Remarque 9.26. Plus généralement, on peut définir des sommes presque nulles dans tout groupe commutatif $(G, +)$, voire même dans tout monoïde (structure algébrique associative munie d'un élément neutre) commutatif.

Définition 9.17 (Combinaison linéaire presque nulle)**.** Soit $(e_i)_{i\in I} \in E^I$ et $(\lambda_i)_{i\in I} \in \mathbb{K}^I$. Alors $(\lambda_i e_i)_{i\in I}$ est presque nulle, donc on peut considérer

$$\sum_{i\in I} \lambda_i e_i$$

qu'on appelle encore une **combinaison linéaire des** e_i. Si on note S le support de $(\lambda_i)_{i\in I}$, le support de $(\lambda_i e_i)_{i\in I}$ est inclus dans S, si bien que :

$$\sum_{i\in I} \lambda_i e_i = \sum_{i\in S} \lambda_i e_i$$

Remarque 9.27. Lorsque $I = [\![1, n]\!]$, il s'agit d'une combinaison linéaire au sens usuel, puisque toute famille indexée sur $[\![1, n]\!]$ est presque nulle. Autrement dit, on a alors $\mathbb{K}^{(I)} = \mathbb{K}^I$, et

$$\sum_{i\in I} \lambda_i e_i = \sum_{i=1}^{n} \lambda_i e_i$$

car $[\![1, n]\!]$ est un sur-ensemble fini du support. Il s'agit donc bien là d'une généralisation du paragraphe précédent.

Remarque 9.28. Un sev reste par combinaison linéaire quelconque, puisqu'il s'agit en fait d'une combinaison linéaire finie.

Proposition 9.12. *Si f est linéaire, alors on a toujours la relation*

$$f\left(\sum_{i\in I} \lambda_i e_i\right) = \sum_{i\in I} \lambda_i f(e_i)$$

Fixons une famille quelconque $(e_i)_{i\in I}$, et considérons la nouvelle application

$$\begin{array}{cccc} \varphi & \mathbb{K}^{(I)} & \to & E \\ & (\lambda_i)_{i\in I} & \mapsto & \sum_{i\in I} \lambda_i e_i \end{array}$$

Par linéarité de la somme, φ reste linéaire (le principe est exactement le même qu'avec les familles finies).

Définition 9.18. On garde les mêmes définitions pour une famille génératrice, libre et une base.

- En utilisant l'injectivité de φ, la caractérisation des familles libres reste valable.
- De même, les deux caractérisations des familles liées restent valables.
- Dès qu'une famille contient le vecteur nul, ou qu'elle contient une répétition, elle est liée.

Théorème 9.2. *Toute sur-famille d'une famille génératrice est génératrice. Toute sous-famille d'une famille libre est libre.*

Théorème 9.3 (HP)**.** *Une famille est libre si, et seulement si, toutes ses sous-familles finies sont libres.*

Démonstration. Le sens direct est immédiat d'après le théorème précédent. Pour le sens réciproque, se ramener au support de la famille lorsqu'on utilise la caractérisation. □

Remarque 9.29. Si F est un sev de E et $(e_i)_{i\in I}$ une famille de vecteurs de F. Alors $(e_i)_{i\in I}$ est libre en tant que famille de vecteurs de F si, et seulement si, elle est libre en tant que famille de vecteurs de E.

En revanche, attention : les choses changent si on modifie le corps de base. Par exemple, on pourra montre que dans l'espace vectoriel $\mathbb{R}$, la famille $(1, \sqrt{2})$ est $\mathbb{Q}$-libre mais $\mathbb{R}$-liée.

Définition 9.19. Si $(e_i)_{i\in I}$ est une famille presque nulle de vecteurs de E, l'ensemble $\mathrm{Vect}((e_i)_{i\in I})$ est défini comme l'ensemble des combinaisons linéaires de e_i. On le notera plus simplement $\mathrm{Vect}(e_i)_{i\in I}$. C'est un sev de E (puisqu'il est égal à $\mathrm{Im}(\varphi)$). C'est le plus petit sev contenant chacun des e_i.

Exemple 9.18. Par définition, $(e_i)_{i\in I}$ est une famille génératrice de $\mathrm{Vect}(e_i)_{i\in I}$.

- Si elle est libre, alors c'est une base de $\mathrm{Vect}(e_i)_{i\in I}$. La réciproque est trivialement vraie.
- Elle est génératrice si, et seulement si, $\mathrm{Vect}(e_i)_{i\in I} = E$.

Exemple 9.19. Posons

$$E = \left\{ x \mapsto \sum_{k=0}^{n} a_k x^k \mid n \in \mathbb{N},\ (a_0, \dots, a_n) \in \mathbb{R}^{n+1} \right\}$$

l'ensemble des fonctions polynomiales réelles, et considérons la famille $(f_k)_{k\in\mathbb{N}}$ définie par

$$\forall k \in \mathbb{N},\ f_k : x \mapsto x^k$$

On vérifie facilement que dans le $\mathbb{R}$-ev $\mathbb{R}^{\mathbb{R}}$, on a $E = \mathrm{Vect}(f_k)_{k\in\mathbb{N}}$. Donc E est un $\mathbb{R}$-ev et $(f_k)_{k\in\mathbb{N}}$ en est une famille génératrice. Puis on montre que $(f_k)_{k\in\mathbb{N}}$ est libre, si bien que c'est une base.

Exemple 9.20. Dans l'espace vectoriel $\mathbb{R}^{\mathbb{N}}$, définissons pour tout $k \in \mathbb{N}$ la suite $(e_n^k)_{n\in\mathbb{N}}$ par

$$\forall n \in \mathbb{N},\ e_n^k = \delta_{n,k}$$

Autrement dit, la suite e^k est la suite qui vaut 0 partout, sauf 1 en k-ème position. On vérifie facilement que dans $\mathbb{R}^{\mathbb{N}}$, on a $\mathbb{R}^{(\mathbb{N})} = \mathrm{Vect}(e_n^k)_{k\in\mathbb{N}}$. Donc (e^k) est une famille génératrice de $\mathbb{R}^{(\mathbb{N})}$, puis on montre qu'elle est libre, si bien qu'elle est une base de $\mathbb{R}^{(\mathbb{N})}$.

Exemple 9.21. On peut généraliser l'exemple précédent à $\mathbb{K}^{(I)}$.

Remarque 9.30. Supposons que l'un des e_i soit combinaison linéaire des autres. Comme dans le cas fini, on ne modifie pas le Vect de la famille en le retirant.

Remarque 9.31. On rappelle que les notions de famille et d'ensemble communiquent à travers les opérations suivantes.

- A une famille $(e_i)_{i\in I}$, on peut associer l'ensemble $\{e_i \mid i \in I\}$.
- Réciproquement, à un ensemble A on peut associer la famille $(a)_{a\in A}$.

A chaque fois, la famille et l'ensemble ont exactement les même éléments (il faut juste prendre gade aux répétitions). Donc un sev contient les éléments de l'ensemble si, et seulement si, il contient les éléments de la famille. On en déduit que l'ensemble et la famille ont le même Vect.

Proposition 9.13 (Vect d'une partie quelconque)**.** *Soit A une partie quelconque de E. Alors* $\mathrm{Vect}(A)$ *est égal à l'ensemble des combinaisons linéaires des vecteurs de A :*

$$\mathrm{Vect}(A) = \left\{ \sum_{a\in A} \lambda_a a \mid (\lambda_a)_{a\in A} \in \mathbb{K}^{(A)} \right\}$$

4.3 Applications linéaires et bases

Théorème 9.4 (Premier théorème de caractérisation : image d'une base)**.** *Soit $(e_i)_{i\in I}$ une base de E et F un espace vectoriel sur le même corps que E. Alors :*

$$\forall (f_i)_{i\in I} \in F^I,\ \exists! u \in \mathcal{L}(E,F),\ \forall i \in I,\ u(e_i) = f_i$$

Proposition 9.14. *Si $u \in \mathcal{L}(E,F)$ et $(e_i)_{i\in I}$ est une famille de vecteurs de E, alors on a :*

$$\mathrm{Vect}(u(e_i))_{i\in I} = u\left(\mathrm{Vect}(e_i)_{i\in I}\right)$$

Exemple 9.22. Si $(e_i)_{i\in I}$ est génératrice, on a $\mathrm{Im}(u) = \mathrm{Vect}(u(e_i))_{i\in I}$.

Théorème 9.5. *Soit $u \in \mathcal{L}(E,F)$, et $(e_i)_{i\in I}$ une base de E. Alors*

- *u est injective si, et seulement si, $(u(e_i))_{i\in I}$ est libre*
- *u est surjective si, et seulement si, $(u(e_i))_{i\in I}$ est génératrice*
- *u est un isomorphisme si, et seulement si, $(u(e_i))_{i\in I}$ une base*

5 Somme de sous-espaces

Proposition 9.15. *Soit E_1 et E_2 deux sev de E. Alors leur somme*

$$E_1 + E_2 = \{x \in E \mid \exists (x_1,x_2) \in E_1 \times E_2,\ x = x_1 + x_2\}$$

est encore un sev.

Exemple 9.23. On a $E_1 + \{0\} = E_1$ et $\{0\} + E_2 = E_2$.

Proposition 9.16. *Soit $(e_i)_{i\in I}$ une famille de vecteurs de E, et $I_1, I_2 \subset I$ tels que $I = I_1 \sqcup I_2$. Alors*

$$\mathrm{Vect}(e_i)_{i\in I} = \mathrm{Vect}(e_i)_{i\in I_1} + \mathrm{Vect}(e_i)_{i\in I_2}$$

Exemple 9.24. On a $\mathrm{Vect}(E_1 \cup E_2) = E_1 + E_2$.

On voit que tout élément de E_1+E_2 s'écrit comme la somme d'un élément de E_1 et d'un élément de E_2, mais y a-t-il unicité de l'écriture dans cette somme ? La petite définition suivante donne un moyen commode de le savoir.

Définition 9.20 (Somme directe)**.** Les deux assertions suivantes sont équivalentes.

1. $\forall x \in E_1 + E_2,\ \exists!(x_1,x_2) \in E_1 \times E_2,\ x = x_1 + x_2$
2. $\forall (x_1,x_2) \in E_1 + E_2,\ x_1 + x_2 = 0 \implies x_1 = x_2 = 0$

On dit alors que E_1 et E_2 sont en **somme directe**, et on note leur somme $E_1 \oplus E_2$.

Proposition 9.17 (Commutativité de la somme directe)**.** *Si E_1 et E_2 sont en somme directe, alors E_2 et E_1 sont en somme directe et on a :*

$$E_1 \oplus E_2 = E_2 \oplus E_1$$

Voici à présent une caractérisation extrêmement commode.

Proposition 9.18. *E_1 et E_2 sont en somme directe si, et seulement si, $E_1 \cap E_2 = \{0_E\}$.*

Exemple 9.25. Soit F un sev de E et $x \in E \setminus F$. Alors F et $\mathbb{K}x$ sont en somme directe. Il suffit de se donner $y \in F \cap \mathbb{K}x$ et de l'écrire sous la forme λx. On a nécessairement $\lambda = 0$ sans quoi x appartiendrait à F en dilatant de λ^{-1}, donc $x = 0_E$ et on conclut par la caractérisation précédente.

Proposition 9.19 (Associativité de la somme directe)**.** *Soit F, G et H des sev de E. Supposons que F et G soient en somme directe, et que $F \oplus G$ et H soient en somme directe. Alors, G et H sont en somme directe, et F et $G \oplus H$ sont en somme directe, et on a naturellement :*

$$(F \oplus G) \oplus H = F \oplus (G \oplus H)$$

Définition 9.21 (Sous-espaces supplémentaires)**.** Si de plus, $E_1 \oplus E_2 = E$, on dit que E_1 et E_2 sont **supplémentaires** (on pourra aussi dire que E_1 est un supplémentaire de E_2 et vice-versa).

Remarque 9.32. Montrer que E_1 et E_2 sont supplémentaires correspond donc à un problème d'existence-unicité.

- Existence pour pouvoir écrire $x = x_1 + x_2$ quel que soit $x \in E$
- Unicité pour que cette écriture soit unique

Comme toujours, l'unicité sera plus facile à montrer que l'existence (en l'occurrence, il suffit de montrer que $E_1 \cap E_2 = \{0_E\}$). On pourra souvent raisonner par analyse-synthèse pour montrer le caractère supplémentaire de deux sev.

Remarque 9.33. Si E_1 est un sev de E, on peut montrer qu'il admet un supplémentaire : la preuve en dimension finie sera donnée dans le chapitre "Dimension finie", et la preuve en dimension infinie, plus délicate, nécessite le lemme de Zorn. En revanche, ce supplémentaire n'a aucune raison d'être unique.

Exemple 9.26. Soit I un intervalle de $\mathbb{R}$ centré autour de 0. Alors les fonctions paires et les fonctions impaires de I dans $\mathbb{R}$ forment deux sev supplémentaires de $\mathbb{R}^I$. En effet, on montre aisément que ce sont des sev de $\mathbb{R}^I$ avec la recette, puis on montre la supplémentarité avec une analyse-synthèse qui fait apparaître ce qu'on appelle les parties paire et impaire de la fonction. Par exemple, cosh et sinh son respectivement les parties paire et impaire de exp.

Théorème 9.6 (Deuxième théorème de caractérisation : supplémentaires)**.** *Soit E_1 et E_2 deux sev de E supplémentaires. Soit $u_1 \in \mathcal{L}(E_1, F)$ et $u_2 \in \mathcal{L}(E_2, F)$. Alors*

$$\exists! u \in \mathcal{L}(E, F),\ \begin{cases} u_{|E_1} = u_1 \\ u_{|E_2} = u_2 \end{cases}$$

Théorème 9.7 (Isomorphisme induit d'un supplémentaire du noyau sur l'image)**.** *Soit $f \in \mathcal{L}(E, F)$. Alors f induit un isomorphisme de tout supplémentaire de* $\ker(f)$ *sur* $\mathrm{Im}(f)$.

Théorème 9.8 (Théorème de la base adaptée)**.** *Soit $(e_i)_{i\in I}$ une base de E, supposons qu'on ait $I = I_1 \sqcup I_2$. Alors $E_1 = \mathrm{Vect}(e_i)_{i\in I_1}$ et $E_2 = \mathrm{Vect}(e_i)_{i\in I_2}$ sont supplémentaires dans E.*

Réciproquement, si E_1 et E_2 sont supplémentaires, alors en concaténant une base de E_1 avec une base de E_2, on obtient une base de E. On dit que cette base de E est ***adaptée à la décomposition*** *$E_1 \oplus E_2$.*

Exemple 9.27. Soit $(e_i)_{i\in I}$ une famille libre et $I_1, I_2 \subset I$ tels que $I = I_1 \sqcup I_2$. Alors

$$\mathrm{Vect}(e_i)_{i\in I} = \mathrm{Vect}(e_i)_{i\in I_1} \oplus \mathrm{Vect}(e_i)_{i\in I_2}$$

Exemple 9.28. Dans le $\mathbb{R}$-ev E des fonctions polynomiales réelles, on rappelle que la famille $f_k : x \mapsto x^k$ forme une base. Si on pose $E_1 = \mathrm{Vect}(f_{2k})_{k\in\mathbb{N}}$ et $E_2 = \mathrm{Vect}(f_{2k+1})_{k\in\mathbb{N}}$, on obtient :

$$E = E_1 \oplus E_2$$

On peut vérifier que E_1 (resp. E_2) est exactement l'ensemble des fonctions polynomiales paires (resp impaires).

6 Applications linéaires importantes

Définition 9.22 (Homothétie de rapport λ)**.** Soit E un $\mathbb{K}$-ev et $\lambda \in \mathbb{K}$. L'**homothétie de rapport** λ est l'endomorphisme de E $x \mapsto \lambda x$.

Proposition 9.20. *L'application $\lambda \mapsto \lambda \,\mathrm{id}$ est un morphisme de $(\mathbb{K}^\star, \times)$ dans $(GL(E), \circ)$.*

Définition 9.23 (Projecteur)**.** Soit F et G deux espaces supplémentaires $E = F \oplus G$. Le **projecteur sur F parallèlement à G** est défini par

$$p : x = f + g \in F \oplus G \mapsto f$$

Comme il y a unicité de l'écriture $x = f + g$, p est bien défini.

Proposition 9.21. *p est un endomorphisme de E et il vérifie $p^2 = p$. On dit qu'il est **idempotent**.*

Remarque 9.34. Attention ! L'exposant 2 dans $p^2 = p$ est à interpréter au sens de la composition : $p^2 = p \circ p$.

Exemple 9.29. Le seul projecteur inversible est l'identité (raisonner par analyse-synthèse et composer par l'inverse de ce projecteur dans l'égalité $p^2 = p$).

Remarque 9.35. Si p est le projecteur sur F parallèlement à G, alors $p|_F^F = \mathrm{id}_F$ et $p|_G^G = 0$.

Proposition 9.22. *Soit p le projecteur sur F parallèlement à G. Alors on a :*

$$\begin{cases} F = \mathrm{Im}(p) \\ G = \ker(p) \end{cases}$$

Remarque 9.36. On en déduit que F et G sont déterminés de manière unique par p.

De manière intéressante, la réciproque est également vraie.

Proposition 9.23. *Si $p \in \mathcal{L}(E)$ vérifie $p^2 = p$, alors c'est le projecteur sur $\mathrm{Im}(p)$ parallèlement à $\ker(p)$.*

Définition 9.24 (Projecteur associé)**.** SI p est le projecteur sur F parallèlement à G, alors $q = \mathrm{id} - p$ est le projecteur sur G parallèlement à F. On a donc

$$\begin{cases} \ker(q) = \mathrm{Im}(p) \\ \mathrm{Im}(q) = \ker(p) \end{cases}$$

q est le **projecteur associé** de p. Bien sûr, réciproquement, p est le projecteur associé de q. De manière générique, on dit que p et q sont associés.

Remarque 9.37. Si p et q sont associés, on voit immédiatement que $pq = qp = 0$.

Définition 9.25 (Symétrie)**.** Si $E = F \oplus G$, la **symétrie par rapport à F parallèlement à G** est définie par

$$s : x = f + g \in F \oplus G \mapsto f - g$$

Proposition 9.24. *s est un endomorphisme de E, et il vérifie $s^2 = \mathrm{id}$.*

Remarque 9.38. s est une involution, et par conséquent une bijection.

Proposition 9.25. *Si p est le projecteur sur F parallèlement à G, alors on a*

$$s = 2p - \mathrm{id}_E$$

De manière équivalente, si q est le projecteur associé de p, on a

$$s = p - q$$

Remarque 9.39. On peut alors retrouver le fait que s est un endomorphisme et un calcul rapide montre que $s^2 = \mathrm{id}_E$.

Remarque 9.40. Si s est la symétrie par rapport F parallèlement à G, alors $p|_F^{|F} = \mathrm{id}_F$ et $p|_G^{|G} = -\mathrm{id}_G$.

Proposition 9.26. *Soit s la symétrie par rapport F parallèlement à G. Alors on a :*

$$\begin{cases} F = \ker(s - \mathrm{id}) \\ G = \ker(s + \mathrm{id}) \end{cases}$$

Remarque 9.41. On en déduit que F et G sont déterminés de manière unique par s.

Là aussi, la réciproque est également vraie.

Proposition 9.27. *Si $s \in \mathcal{L}(E)$ vérifie $s^2 = \mathrm{id}$, alors c'est la symétrie parallèlement à $\ker(s - \mathrm{id})$ parallèlement à $\ker(s + \mathrm{id})$.*

Remarque 9.42 (Symétries en caractéristique 2, HP)**.** Attention, les deux propositions précédentes deviennent fausses lorsque $\mathbb{K}$ est de caractéristique 2 ! En effet, dans ce cas, on ne peut plus effectuer de divisions par 2, puisque $2 = 0$.

Voici d'ailleurs un contre-exemple. Si $\mathbb{K} = \mathbb{Z}/2\mathbb{Z}$, on considère le $\mathbb{K}$-ev $\mathbb{K}^2$ muni de la base canonique (e_1, e_2). Alors s définie par

$$\begin{cases} s(e_1) = e_1 \\ s(e_2) = e_1 + e_2 \end{cases}$$

vérifie $s^2 = \mathrm{id}$ et pourtant $s \neq \mathrm{id}$. Or, si s était une symétrie, on aurait $s(f + g) = f - g = f + g$ d'où $s = \mathrm{id}$.

7 Hyperplans et formes linéaires

Définition 9.26 (Forme linéaire, espace dual)**.** On appelle **forme linéaire** une application $f \in \mathcal{L}(E, \mathbb{K})$. $\mathcal{L}(E, \mathbb{K})$ est noté $E^\star$, on l'appelle **espace dual** de E.

Proposition 9.28. *Toute forme linéaire non nulle est surjective.*

Définition 9.27 (Hyperplan)**.** Un supplémentaire d'une droite vectorielle s'appelle un **hyperplan** de E.

Remarque 9.43. Le vocabulaire provient manifestement de la dimension 3, où le supplémentaire d'une droite vectorielle est un plan vectoriel comme nous le verrons dans le chapitre "Dimension finie". En dimension supérieure, l'analogue d'un plan s'appelle un "hyperplan" pour suggérer qu'on est monté en dimension (de même, on peut parler d'hypercube, d'hypersphère...).

Théorème 9.9 (Caractérisation des hyperplans par les formes linéaires)**.** *H est un hyperplan de E si, et seulement si, c'est le noyau d'une forme linéaire non nulle.*

Corollaire 9.4. *Si H est un hyperplan de E et $u \notin H$, alors on a toujours $H \oplus \mathbb{K}u = E$.*

Remarque 9.44. En d'autres termes, un hyperplan a été défini comme un supplémentaire d'**une** droite vectorielle (qu'il ne contient nécessairement pas). Et *a posteriori*, il s'avère qu'il est supplémentaire de **toute** droite vectorielle qu'il ne contient pas.

Définition 9.28 (Formes coordonnées associées à une base)**.** Soit $(e_i)_{i\in I}$ une base de E. Les **formes coordonnées** associées à la base $(e_i)_{i\in I}$ sont les formes linéaires $(e_i^\star)_{i\in I}$ définies par

$$\forall (i,j) \in I^2,\ \delta_{i,j} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{sinon} \end{cases}$$

Remarque 9.45. Notons que les formes coordonnées sont bien définies, car une application linéaire est fixée de manière unique par l'image d'une base.

Remarque 9.46. Même si la notation ne le suggère pas (et cela est malheureux, mais rendrait la notation trop lourde), les formes coordonnées peuvent toutes changer si on changer un seul vecteur de la base pour obtenir une autre base. Cependant, les autres formes coordonnées ne changent pas de notation, bien qu'elles puissent changer puisque la base a changé.

Définition 9.29 (Symbole de Kronecker)**.** Le **symbole de Kronecker**, ou **delta de Kronecker**, est défini par

$$\forall (i,j) \in I^2,\ \delta_{i,j} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{sinon} \end{cases}$$

Remarque 9.47. Avec cette notation, la définition des formes coordonnées devient

$$\forall (i,j) \in I^2,\ e_i^\star(e_j) = \delta_{i,j}$$

Le théorème suivant justifie l'appellation de "forme coordonnée".

Théorème 9.10. *Soit $(e_i)_{i\in I}$ une base de E et $(e_i^\star)_{i\in I}$ les formes coordonnées associées. Alors, pour tout $x \in E$, la coordonnée selon e_i vaut $e_i^\star(x)$.*

Exemple 9.30. Dans le $\mathbb{R}$-ev $\mathbb{C}$, les fonctions $\mathrm{Re}(\cdot)$ et $\mathrm{Im}(\cdot)$ sont les formes coordonnées associées à la base $(1, i)$.

8 Translations, sous-espaces affines

8.1 Notions de base

Définition 9.30 (Translation)**.** Soit E un $\mathbb{K}$-ev et $a \in E$ fixé. La **translation de vecteur** a est l'application $x \mapsto x + a$. Attention, en général, ce n'est **pas** un endomorphisme (sauf si $a \neq 0$).

Jusqu'ici, nous avons vu que tout sev contenait au moins 0. Ainsi, les droites vectorielles, les plans vectoriels, et plus généralement tous les sev "passent par l'origine" au sens géométrique du terme. Or, en géométrie élémentaire, on sait qu'une droite ou un plan peut très bien ne pas passer par l'origine, tout en étant parallèle à une droite ou un plan qui passe par l'origine. Il nous faudrait donc une généralisation de la notion de sev. Ce ce qu'on appelle un sea, o sous-espace affine. Il consiste à se donner un point de l'espace comme origine, puis à définir un sous-espace qui passe par ce point et qui est parallèle à un sev donné.

Définition 9.31 (Sous-espace affine)**.** Soit F un sev de E et $a \in E$. Alors, le **sous-espace affine de E passant par a et parallèle à F** est défini par

$$a + F = \{a + f \mid f \in F\}$$

C'est l'image de F par la translation de vecteur a. Les éléments de $a + f$ sont appelés **points** de $a + F$. F s'appelle **la direction** du sea, et a s'appelle **une origine** du sea.

Proposition 9.29. *La direction d'un sea est unique. En revanche, n'importe lequel de ses points peut servir de direction. De manière précise, soit $\mathcal{F} = a + F$ un sea, et soit $a' \in E$ et F' un sev de E. Alors $\mathcal{F} = a' + F'$ si, et seulement si, $F = F'$ et $a' \in \mathcal{F}$.*

Exemple 9.31 (Équation affine)**.** Soit $f \in \mathcal{L}(E)$, et cherchons les solutions de l'équation $f(x) = b$, où $b \in E$ est fixé. Si l'équation possède au moins une solution x_0, alors l'ensemble des solutions est le sea passant par x_0 et de direction $\ker(f)$.

Exemple 9.32. Soit φ une forme linéaire non nulle, et soit $y \in \mathbb{K}$ fixé. Puisque φ est surjective, l'équation $\varphi(x) = y$ admet au moins une solution. Et donc, son ensemble de solutions est un hyperplan affine dirigé par $\ker(\varphi)$.

8.2 Notions plus avancées

Définition 9.32 (Parallélisme)**.** Deux sea $\mathcal{A}$ et $\mathcal{B}$ sont dits **parallèles** lorsqu'ils ont même direction. Le parallélisme est une relation d'équivalence sur l'ensemble des sea.

Proposition 9.30 (Intersection de sea)**.** *Soit $\mathcal{A}$ un sea de direction F et $\mathcal{B}$ un sea de direction G. Si $\mathcal{A} \cap \mathcal{B} \neq \emptyset$, alors $\mathcal{A} \cap \mathcal{B}$ est un sea de direction $F \cap G$.*

Remarque 9.48. Attention aux fausses intuitions ! Dans le plan, on sait que deux droites non parallèles se coupent toujours, mais dans l'espace (et au-delà), ce résultat devient faux. Par exemple, dans $\mathbb{R}^3$, la droite $y = z = 0$ n'a aucun point commun avec la droite $\begin{cases} y = 0 \\ z = 1 \end{cases}$.

Exemple 9.33. Dans $\mathbb{R}^3$, on peut montrer que l'intersection de deux hyperplans non parallèles est une droite. Pour l'instant, on se contentera de faire un dessin ; pour une démonstration élégante, on préfèrera attendre le chapitre sur la dimension finie.

Remarque 9.49. Plus généralement, si on se donne un ensemble de sea d'intersection non vide, alors cette intersection est un sea, de direction l'intersection des directions. La démonstration est rigoureusement la même.

Définition 9.33 (Sea engendré par une partie)**.** Soit X une partie non vide de E. Alors l'intersection de tous les sea contenant X est non vide (puisqu'elle contient X). C'est donc un sea, qu'on appelle **sea engendré par X**. C'est le plus petit sea contenant X.

Exemple 9.34. Le sea engendré par deux points non confondus est une droite affine. Le sea engendré par trois points non alignés est un plan affine.

Chapitre 10

Dimension finie

Dans tout le chapitre, on fixe $\mathbb{K}$ un corps et E un $\mathbb{K}$-espace vectoriel.

1 Dimension

Définition 10.1 (Espace de dimension finie)**.** E dit de **dimension finie** lorsqu'il admet une famille génératrice finie.

Exemple 10.1. $\{0\}$ est de dimension finie car engendré par la famille vide qui est finie.

Lemme 10.1 (Lemme d'adjonction, NS)**.** *Soit $(e_1, \dots, e_p)$ une famille libre de vecteurs de E et $x \in E \setminus \mathrm{Vect}(e_1, \dots, e_p)$. Alors, la famille $(e_1, \dots, e_p, x)$ est encore libre.*

Remarque 10.1. Le lemme reste vrai pour tout autre ensemble d'indexation que $[\![1, p]\!]$, même si cet ensemble est infini.

Théorème 10.1 (Théorème de la base incomplète, cas fini)**.** *Soit $\mathcal{L}$ une famille libre finie de vecteurs de E et $\mathcal{G}$ une famille génératrice finie de E, qui va nous servir de "réservoir". Alors $\mathcal{L}$ peut être complétée en une base à l'aide de certains vecteurs de $\mathcal{G}$.*

Théorème 10.2 (Théorème de la base extraite, cas fini)**.** *Soit $\mathcal{G}$ une famille génératrice finie de E. Alors, il existe une sous-famille de $\mathcal{G}$ qui est une base de E, et une base de ce type est appelée* ***base extraite****.*

Théorème 10.3. *Tout espace de dimension finie admet une base finie.*

Remarque 10.2. En fait, on remarque qu'il y a équivalence entre le fait de posséder une famille génératrice finie et une base finie.

Lemme 10.2 (Lemme du pivot)**.** *Supposons que E admette une famille génératrice finie de taille $n \in \mathbb{N}$. Alors, $n+1$ vecteurs de E sont toujours liés.*

Exemple 10.2. Trois vecteurs du plan $\mathbb{R}^2$ sont toujours liés. Quatre vecteurs de l'espace $\mathbb{R}^3$ sont toujours liés.

Remarque 10.3. Plus généralement, on montre, puisque toute sur-famille d'une famille liée est liée, que si E une famille génératrice de n vecteurs, alors toute famille de $m \geqslant n+1$ vecteurs est liée. Cela reste donc vrai si on a une famille infinie de vecteurs.

Définition 10.2 (Dimension d'un espace de dimension finie)**.** Supposons que E est de dimension finie. Alors toutes les bases de E ont la même taille, notons-la $n \in \mathbb{N}$ par exemple. On pose alors la **dimension de** E comme étant égale à :

$$\dim(E) := n$$

Exemple 10.3 ($\mathbb{K}^n$ pour $n \in \mathbb{N}^\star$)**.** Soit $n \in \mathbb{N}^\star$. Alors $\mathbb{K}^n$ est un $\mathbb{K}$-espace vectoriel de dimension finie, et $\dim(\mathbb{K}^n) = n$ (prendre la base canonique).

Exemple 10.4 (Droite vectorielle)**.** Soit une droite vectorielle $\mathbb{K}u$ avec $u \neq 0_E$. (u) est génératrice de $\mathbb{K}u$ et libre, donc c'est une base de $\mathbb{K}u$. Puis $\dim(\mathbb{K}u) = 1$.

Exemple 10.5 (Plan vectoriel)**.** Soit un plan vectoriel $\mathbb{K}u + \mathbb{K}v$ avec u et v non colinéaires. (u, v) en et une base, donc $\dim(\mathbb{K}u + \mathbb{K}v) = 2$.

Exemple 10.6 (Espace nul)**.** La famille vide est une base de l'espace nul, donc $\dim(\{0\}) = 0$.

Théorème 10.4 (Lien entre taille et caractères libre et générateur)**.** *Supposons que E est de dimension $n \in \mathbb{N}$. Alors*

1. *Pour toute famille libre $\mathcal{L}$ de vecteurs de E, on a Taille$(\mathcal{L}) \leqslant n$ avec égalité si, et seulement si Taille$(\mathcal{L}) = n$.*
2. *Pour toute famille génératrice $\mathcal{G}$ de vecteurs de E, on a Taille$(\mathcal{G}) \geqslant n$ (possiblement infinie) avec égalité si, et seulement si Taille$(\mathcal{G}) = n$.*

Lemme 10.3. *Supposons que E est de dimension finie. Alors toute famille génératrice $\mathcal{G}$ de E possède une sous-famille génératrice finie.*

Théorème 10.5 (Théorème de la base incomplète, cas général)**.** *Supposons que E est de dimension finie. Soit $\mathcal{L}$ une famille libre de vecteurs de E quelconque et $\mathcal{G}$ une famille génératrice de E quelconque. Alors, $\mathcal{L}$ peut être complétée en une base de E à partir de certains vecteurs de $\mathcal{G}$.*

Théorème 10.6 (Théorème de la base extraite, cas général)**.** *Supposons que E est de dimension finie. Alors on peut extraire une base de E de toute famille génératrice $\mathcal{G}$ de E.*

Théorème 10.7 (Dimension et isomorphisme)**.** *Supposons que E est de dimension finie et que F est un $\mathbb{K}$-espace vectoriel quelconque. Alors :*

- *Si E et F sont isomorphes, alors F est de dimension finie, et on a $\dim(F) = \dim(F)$.*
- *Réciproquement, si F est de dimension finie et si $\dim(F) = \dim(E)$, alors E et F sont isomorphes.*

Exemple 10.7. Tout $\mathbb{K}$-espace vectoriel de dimension $n \in \mathbb{N}$ est isomorphe à $\mathbb{K}^n$, avec la convention $\mathbb{K}^0 = \{0\}$.

Théorème 10.8 (Dimension d'un produit cartésien)**.** *Soit E un $\mathbb{K}$-ev de dimension finie égale à $p \in \mathbb{N}$ et F un $\mathbb{K}$-ev de dimension finie égale à n. Alors $E \times F$ est de dimension finie égale à $p+n$, ie on a alors*

$$\dim(E \times F) = \dim(E) + \dim(F)$$

Théorème 10.9 (Dimension de $\mathcal{L}(E,F)$)**.** *Soit E un $\mathbb{K}$-ev de dimension finie égale à $p \in \mathbb{N}$ et F un $\mathbb{K}$-ev de dimension finie égale à n. Alors $\mathcal{L}(E,F)$ est de dimension finie égale à $p \times n$, ie on a alors*

$$\dim(\mathcal{L}(E,F)) = \dim(E) \times \dim(F)$$

Exemple 10.8. Si E est de dimension finie, alors $\mathcal{L}(E)$ est de dimension finie et on a :

$$\dim(\mathcal{L}(E)) \dim(E)^2$$

Exemple 10.9. Si E est de dimension finie, alors $E^\star$ est de dimension finie et on a :

$$\dim(E^\star) = \dim(E)$$

2 Dimension d'un sous-espace vectoriel

Théorème 10.10 (Dimension et sev)**.** *Soit E de dimension finie et F un sev de E. Alors, F est de dimension finie et on a $\dim(F) \leqslant \dim(E)$, avec égalité si, et seulement si, $F = E$.*

Remarque 10.4. On pourra souvent se servir du cas d'égalité pour montrer une égalité d'ensembles. Retenir que inclusion et égalité de dimension donne égalité d'ensembles.

Exemple 10.10. On retrouve que les seuls sous-espaces vectoriels de $\mathbb{K}$ sont $\{0\}$ et $\mathbb{K}$.

Exemple 10.11. On retrouve le fait que toute forme linéaire non nulle est surjective (en utilisant l'exemple précédent sur l'image de cette forme linéaire).

Définition 10.3 (Rang d'une famille de vecteurs)**.** Soit $(x_1, \dots, x_p) \in E^p$ avec $p \in \mathbb{N}$. Le **rang de la famille** $(x_j)_{1 \leqslant j \leqslant p}$ est défini par :

$$rg(x_j)_{1 \leqslant j \leqslant p} := \dim\left(\mathrm{Vect}(x_j)_{1 \leqslant j \leqslant p}\right)$$

Proposition 10.1. *Soit $(x_1, \dots, x_p) \in E^p$ avec $p \in \mathbb{N}$. On a $rg(x_j)_{1 \leqslant j \leqslant p} \leqslant p$ avec égalité si, et seulement si, $(x_1, \dots, x_p)$ est libre.*

Théorème 10.11 (Existence du supplémentaire en dimension finie)**.** *Soit E un espace de dimension finie et F un sev de E. Alors F admet un supplémentaire dans E.*

Remarque 10.5. Attention : on rappelle que ce supplémentaire n'a **aucune raison** d'être unique !

Remarque 10.6 (Cas de la dimension infinie, HP)**.** On rappelle qu'en dimension, ce résultat reste vrai, mais il faut montrer que l'ensemble des sev en somme directe avec F est un ensemble inductif puis lui appliquer le lemme de Zorn. Il faut alors montrer la somme totale. ON peut par exemple raisonner par l'absurde et utiliser la maximalité offerte par le lemme de Zorn.

Lemme 10.4 (Lemme pré-Grassmann, NS)**.** *Soit E un espace vectoriel quelconque, et F et G des sous-espaces vectoriels de E de dimension finie. Si F et Gsont en somme directe, alors $F \oplus G$ est de dimension finie et on a :*

$$\dim(F \oplus G) = \dim(F) + \dim(G)$$

Exemple 10.12 (Cas fondamental)**.** Supposons que E est de dimension finie et que F et G sont en somme directe. Si $\dim(F) + \dim(G) = \dim(E)$ (comme on pourra souvent le montrer plus tard par le **théorème du rang**), alors $E = F \oplus G$, *ie* F et G sont supplémentaires dans E.

Théorème 10.12 (Formule de Grassmann)**.** *Soit E un espace vectoriel quelconque, et F et G des sous-espaces vectoriels de E de dimension finie. Alors $F+G$ est de dimension finie et on a :*

$$\dim(F+G) = \dim(F) + \dim(G) - \dim(F \cap G)$$

Exemple 10.13. Soit E un espace vectoriel quelconque, et F et G des sous-espaces vectoriels de E de dimension finie. Alors $F+G$ est de dimension finie et on a :

$$\dim(F+G) \leqslant \dim(F) + \dim(G)$$

Exemple 10.14. On peut généraliser par récurrence : si $F_1, \ldots, F_n$ sont des sev de dimension finie de E, alors $F_1 + \ldots + F_n$ est de dimension finie et on a :

$$\dim\left(\sum_{i=1}^{n} F_i\right) \leqslant \sum_{i=1}^{n} \dim(F_i)$$

3 Rang d'une application linéaire

Définition 10.4 (Rang d'une application linéaire)**.** Soit $u \in \mathcal{L}(E,F)$. u est dite de **rang fini** lorsque $\mathrm{Im}(u)$ est de dimension finie. Dans ce cas, le **rang de** u est défini par :

$$rg(u) := \dim\left(\mathrm{Im}(u)\right)$$

Remarque 10.7. Si E ou F est de dimension finie, il est inutile de prendre toutes ces précautions puisqu'alors $\mathrm{Im}(u)$ est de dimension finie. En effet, si F est de dimension finie alors $\mathrm{Im}(u)$ est un sev de F donc de dimension finie. Si E est de dimension finie, u envoie une base de E sur une famille génératrice de $\mathrm{Im}(u)$, donc $\mathrm{Im}(u)$ est de dimension finie.

Théorème 10.13 (Théorème du rang)**.** *Soit E un $\mathbb{K}$-ev de dimension finie, F un $\mathbb{K}$-ev quelconque et $u \in \mathcal{L}(E,F)$. Alors :*

$$\dim(E) = \dim(\ker(u)) + rg(u)$$

Corollaire 10.1. *Soit E et F des $\mathbb{K}$-ev de dimension finie tels que $\dim(E) = \dim(F)$. Soit $u \in \mathcal{L}(E,F)$. Alors les trois assertions suivantes sont équivalentes :*

1. *u est injective*
2. *u est surjective*
3. *u est bijective*

Remarque 10.8. L'exemple typique d'utilisation est lorsque E est de dimension finie et $u \in \mathcal{L}(E)$.

Corollaire 10.2. *Soit E de dimension finie et $u, v \in \mathcal{L}(E)$. Si $u \circ v = \mathrm{id}_E$, alors $v \circ u = \mathrm{id}_E$.*

Remarque 10.9. Plus généralement, soit E et F de même dimension (finie) ainsi que $u \in \mathcal{L}(E,F)$ et $v \in \mathcal{L}(F,E)$. Si $u \circ v = \mathrm{id}_E$, alors $v \circ u = \mathrm{id}_F$.

Proposition 10.2. *Soit E, F et G de dimension finie. Soit $u \in \mathcal{L}(E,F)$ et $v \in \mathcal{L}(F,G)$. Alors :*

$$rg(v \circ u) \leqslant \min(rg(u), rg(v))$$

Remarque 10.10. Par une récurrence immédiate, on a plus généralement des inégalités du type :

$$rg(u_1 \circ \ldots \circ u_p) \leqslant \min_{1 \leqslant i \leqslant p} (rg(u_i))$$

Proposition 10.3 (Invariance du rang par composition par un isomorphisme)**.** *Soit E et F des $\mathbb{K}$-espaces vectoriels quelconques et soit $u \in \mathcal{L}(E,F)$ de rang fini. Alors :*

- *Soit F' un espace vectoriel quelconque. Pour tout isomorphisme $\varphi : F \to F'$, $\varphi \circ u$ est de rang fini et on a :*
$$rg(\varphi \circ u) = rg(u)$$
- *Soit E' un espace vectoriel quelconque. Pour tout isomorphisme $\psi : E' \to E$, $u \circ \psi$ est de rang fini et on a :*
$$rg(u \circ \psi) = rg(u)$$

Exemple 10.15. Soit $u \in \mathcal{L}(E,F)$ de rang fini. Alors $-u$ est de rang fini et on a $rg(-u) = rg(u)$ (prendre $-\mathrm{id}_E$ comme isomorphisme).

4 Hyperplans et dualité

Dans tout le paragraphe, on fixe E un $\mathbb{K}$-espace vectoriel de dimension finie $n \in \mathbb{N}$ et $(e_1, \ldots, e_n)$ une base de E.

Définition 10.5 (Base duale)**.** $(e_1^\star, \ldots, e_n^\star)$ est une base de $E^\star$, appelée **base duale** de $(e_1, \ldots, e_n)$.

Remarque 10.11 (Résolution d'exercices en dualité)**.** Dans les exercices de dualité, deux techniques générales se distinguent.

- Si on dispose d'une égalité de vecteurs, il est judicieux d'appliquer une forme linéaire des deux côtés de cette égalité (le plus souvent, ce sera une $e_i^\star$).
- Si on dispose d'une égalité de formes linéaires, il est judicieux d'évaluer cette égalité en un vecteur précis (le plus souvent, ce sera en un des e_i).

Remarque 10.12 (Note historique)**.** A propos de cette remarque, le grand FM avait déclaré : "En prépa, beaucoup d'élèves ont peur de la dualité. C'est normal, parce qu'ils ont des mauvais profs. Mais moi, j'suis un bon prof !" (d'une voix enjouée pour la dernière phrase).

Proposition 10.4 (Expression d'une forme linéaire dans la base duale)**.** *Soit $\varphi \in E^\star$. Ses coordonnées dans la base $(e_i^\star)_{1 \leqslant i \leqslant n}$ sont $(\varphi(e_i))_{1 \leqslant i \leqslant n}$. Autrement dit, on a :*

$$\varphi = \sum_{i=1}^{n} \varphi(e_i) e_i^\star$$

Théorème 10.14 (Caractérisation des hyperplans par les équations)**.** *On a la caractérisation suivante des hyperplans par les équations en dimension finie.*

- *Soit $\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{K}^n \setminus \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ et considérons l'ensemble $H_{a_1,\ldots,a_n} : \sum_{i=1}^{n} a_i x_i = 0_{\mathbb{K}}$. Alors $H_{a_1,\ldots,a_n}$ est un hyperplan.*

- *Réciproquement, soit H un hyperplan de E. Alors* $\exists \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{K}^n \setminus \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$, $H = H_{a_1,\dots,a_n}$.
- *De plus, soit* $\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{K}^n \setminus \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ *et* $\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in \mathbb{K}^n \setminus \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ *tels que* $H = H_{a_1,\dots,a_n} = H_{b_1,\dots,b_n}$.

  Alors

$$\exists \lambda \in \mathbb{K}^\star, \ \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \lambda \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

Théorème 10.15 (Caractérisation des hyperplans par la dimension en dimension finie)**.** *En dimension finie, les hyperplans de E sont exactement les sev de dimension* $\dim(E) - 1$.

Théorème 10.16. *Soit $p \in [\![1, n]\!]$.*

- *Soit $H_1, \dots, H_p$ des hyperplans. Alors :*

$$\dim(H_1 \cap \dots \cap H_p) \geqslant n_p$$

- *Réciproquement, soit F un sev de dimension $n - p$. Alors il existe des hyperplans $H_1, \dots, H$ tels que*

$$F = H_1 \cap \dots \cap H_n$$

Chapitre 11

Matrices et systèmes linéaires

Dans tout le chapitre, on fixe $\mathbb{K}$ un corps.

1 Calcul matriciel

1.1 Structure de $\mathbb{K}$-espace vectoriel

Définition 11.1 (Matrices, ensemble $\mathcal{M}_{n,p}(\mathbb{K})$)**.** Soit $n,p \in \mathbb{N}^\star$. Intuitivement, une matrice à n lignes et p colonnes est un tableau de taille $n \times p$ à coefficients das $\mathbb{K}$.

Formellement, une matrice à n lignes et p colonnes à coefficients dans $\mathbb{K}$ est un triplet $\left(n,p,(a_{i,j})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant p}}\right)$.

Les $a_{i,j}$ sont des scalaires appelés **termes**, ou **coefficients** de la matrice.
L'ensemble des matrices à n lignes, p colonnes et à coefficients dans $\mathbb{K}$ est noté $\mathcal{M}_{n,p}(\mathbb{K})$.

Définition 11.2 (Matrices-colonne, matrices-ligne)**.** Soit $n,p \in \mathbb{N}^\star$. $\mathcal{M}_{n,1}(\mathbb{K})$ est l'ensemble des **matrices-colonne** de taille n à coefficients dans $\mathbb{K}$. $\mathcal{M}_{1,p}(\mathbb{K})$ est l'ensemble des **matrices-ligne** de taille p à coefficients dans $\mathbb{K}$.

Définition 11.3 (Lois de $\mathbb{K}$-ev de $\mathcal{M}_{n,p}(\mathbb{K})$)**.** On munit $\mathcal{M}_{n,p}(\mathbb{K})$ des lois suivantes :

- Un lci notée $+$:

$$\left(n,p,(a_{i,j})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant p}}\right)+\left(n,p,(b_{i,j})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant p}}\right):=\left(n,p,(a_{i,j}+b_{i,j})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant p}}\right)$$

- Une lce notée $\cdot$:

$$\lambda.\left(n,p,(a_{i,j})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant p}}\right)=\left(n,p,(\lambda a_{i,j})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant p}}\right)$$

Proposition 11.1 (Structure de $\mathbb{K}$-ev de $\mathcal{M}_{n,p}(\mathbb{K})$)**.** *Muni de ces deux lois, $\mathcal{M}_{n,p}(\mathbb{K})$ est un $\mathbb{K}$-ev.*

Définition 11.4 (Matrices élémentaires)**.** Soit $n,p \in \mathbb{N}^\star$. Les **matrices élémentaires** de $\mathcal{M}_{n,p}(\mathbb{K})$ sont définies par :

$$\forall (i_0,j_0) \in [\![1,n]\!]\times[\![1,p]\!],\ E_{i_0,j_0}:=\left(n,p,(\delta_{i,i_0}\delta_{j,j_0})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant p}}\right)$$

Si le contexte est ambigu, on pourra les noter $E_{i_0,j_0}^{(n,p)}$.

Théorème 11.1 (Base canonique de $\mathcal{M}_{n,p}(\mathbb{K})$, dimension de $\mathcal{M}_{n,p}(\mathbb{K})$)**.** *La famille* $\left(E_{i,j}^{(n,p)}\right)_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant p}}$ *est une base de* $\mathcal{M}_{n,p}(\mathbb{K})$*, appelée* ***base canonique*** *de* $\mathcal{M}_{n,p}(\mathbb{K})$*. En particulier,* $\mathcal{M}_{n,p}(\mathbb{K})$ *est un* $\mathbb{K}$*-espace vectoriel de dimension finie égale à* $n \times p$*.*

Remarque 11.1. Dès lors, on identifiera $\mathbb{K}^n$ à $\mathcal{M}_{n,1}(\mathbb{K})$ car ces espaces sont isomorphes.

1.2 Produit matriciel

Définition 11.5 (Produit matriciel)**.** On définit une loi $\times$ sur les matrices par :

$$\left(n, p, (a_{i,j})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant p}}\right) \times \left(p, q, (b_{i,j})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant q}}\right) := \left(n, q, \left(\sum_{k=1}^{p} a_{i,k}b_{k,j}\right)_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant q}}\right)$$

Remarque 11.2. Attention à bien prendre garde à la compatibilité des tailles ! Le nombre de lignes de la deuxième matrice doit être égal au nombre de colonnes de la première.

Exemple 11.1. Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B \in \mathcal{M}_{p,q}(\mathbb{K})$. On a :

$$\begin{cases} 0_{n,p} \times B = 0_{n,q} \\ A \times 0_{p,q} = 0_{n,q} \end{cases}$$

Exemple 11.2 (Produit d'une matrice-ligne et d'une matrice-colonne)**.** Soit $p \in \mathbb{N}^\star$. On a :

$$(\lambda_1, \ldots, \lambda_p) \times \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_p \end{pmatrix} = \left(\sum_{k=1}^{p} \lambda_k \mu_k\right)$$

et

$$\begin{pmatrix} \mu_1 \\ \vdots \\ \mu_p \end{pmatrix} \times (\lambda_1, \ldots, \lambda_p) = \left(p, p, (\mu_i\lambda_j)_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant p}}\right)$$

Exemple 11.3 ($\mathcal{M}_{1,1}(\mathbb{K})$)**.** On a $(\lambda) \times (\mu) = (\lambda \times \mu)$. Comme de plus $(\lambda) + (\mu) = (\lambda + \mu)$, on pourra identifier $\mathbb{K}$ et $\mathcal{M}_{1,1}(\mathbb{K})$.

Exemple 11.4 (Utile)**.** Soit $\lambda \in \mathbb{K}$ et $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n$. Alors on a :

$$\lambda . \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \times (\lambda)$$

Théorème 11.2 (Produit de matrices élémentaires)**.** *Soit $n, p, q \in \mathbb{N}^\star$. Soit $(i,j) \in [\![1,n]\!] \times [\![1,p]\!]$. et $(k,l) \in [\![1,p]\!] \times [\![1,q]\!]$. On a :*

$$E_{i,j}^{(n,p)} \times E_{k,l}^{(p,q)} = \delta_{j,k} E_{i,l}^{(n,q)}$$

Théorème 11.3. *La loi $\times$ est associative et bilinéaire.*

Exemple 11.5. On en déduit la compatibilité du produit matriciel avec la loi externe er la distributivité du produit matriciel sur $+$.

Théorème 11.4 (Produit par blocs)**.** *Soit $n_1, n_2, p_1, p_2, q_1, q_2 \in \mathbb{N}^\star$, $A \in \mathcal{M}_{n_1,p_1}(\mathbb{K})$, $B \in \mathcal{M}_{n_1,p_2}(\mathbb{K})$, $C \in \mathcal{M}_{n_2,p_1}(\mathbb{K})$, $D \in \mathcal{M}_{n_2,p_2}(\mathbb{K})$, $A' \in \mathcal{M}_{p_1,q_1}(\mathbb{K})$, $B' \in \mathcal{M}_{p_1,q_2}(\mathbb{K})$, $C' \in \mathcal{M}_{p_2,q_1}(\mathbb{K})$ et $D' \in \mathcal{M}_{p_2,q_2}(\mathbb{K})$ (en gros, il faut que les tailles de blocs soient compatibles). Alors on a :*

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \times \begin{pmatrix} A' & B' \\ C' & D' \end{pmatrix} = \begin{pmatrix} AA' + BC' & AB' + BD' \\ CA' + DC' & CB' + DD' \end{pmatrix}$$

On retiendra qu'on peut multiplier les blocs comme si c'étaient les coefficients d'une matrice normale (du moment que les blocs sont compatibles).

Exemple 11.6. Voici trois cas particuliers à connaître sur le bout des doigts :

- Colonne $\times$ ligne :

$$\begin{pmatrix} \\ C \\ \\ \end{pmatrix} \times \begin{pmatrix} \lambda_1 & | & \dots & | & \lambda_n \end{pmatrix} = \begin{pmatrix} C \times (\lambda_1) & | & \dots & | & C \times (\lambda_n) \\ \\ \end{pmatrix} = \begin{pmatrix} \lambda_1 C & | & \dots & | & \lambda_n C \\ \\ \end{pmatrix}$$

- Matrice $\times$ colonne :

$$\begin{pmatrix} C_1 & | & \dots & | & C_p \\ \\ \end{pmatrix} \times \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_p \end{pmatrix} = \begin{pmatrix} C_1 \times (\lambda_1) + \dots + C_p \times (\lambda_p) \\ \\ \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{p} \lambda_i C_i \\ \\ \end{pmatrix}$$

- Matrice $\times$ matrice :

$$\begin{pmatrix} \\ A \\ \\ \end{pmatrix} \times \begin{pmatrix} C_1 & | & \dots & | & C_q \\ \\ \end{pmatrix} = \begin{pmatrix} AC_1 & | & \dots & | & AC_q \\ \\ \end{pmatrix}$$

1.3 Matrices carrées

Définition 11.6 (Matrices carrées)**.** $\mathcal{M}_{n,n}(\mathbb{K})$ est noté plus simplement $\mathcal{M}_n(\mathbb{K})$. C'est l'ensemble des **matrices carrées d'ordre n à coefficients dans $\mathbb{K}$**. Il s'agit donc d'un $\mathbb{K}$-ev de dimension n^2.

Définition 11.7 (Matrice identité)**.** La **matrice identité d'ordre n** est la matrice carrée d'ordre n définie par :

$$I_n := \begin{pmatrix} 1 & & (0) \\ & \ddots & \\ (0) & & 1 \end{pmatrix}$$

Théorème 11.5. *$(\mathcal{M}_n(\mathbb{K}), +, \cdot, \times)$ est une $\mathbb{K}$-algèbre, d'élément neutre I_n pour $\times$.*

Remarque 11.3. Plus généralement, soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$. On a :

$$\begin{cases} A \times I_p = A \\ I_n \times A = A \end{cases}$$

Remarque 11.4. $\mathcal{M}_n(\mathbb{K})$ pour $n \in \mathbb{N}^\star$ est commutatif si, et seulement si $n = 1$. Dès que $n \geqslant 2$, il suffit de considérer les produits $E_{1,2} \times E_{2,1}$ et $E_{2,1} \times E_{1,2}$.

Remarque 11.5. Dès que $n \geqslant 2$, $\mathcal{M}_n(\mathbb{K})$ n'est pas intègre car $E_{1,2} \times E_{1,2} = 0_n$.

Définition 11.8 (Groupe linéaire d'ordre n)**.** On appelle **groupe linéaire d'ordre** n le groupe des unités de l'anneau $\mathcal{M}_n(\mathbb{K})$. On le note $GL_n(\mathbb{K})$ et on a donc :

$$GL_n(\mathbb{K}) := \mathcal{U}\left(\mathcal{M}_n(\mathbb{K})\right) = \{A \in \mathcal{M}_n(\mathbb{K}) \mid \exists B \in \mathcal{M}_n(K),\ AB = BA = I_n\}$$

Exemple 11.7. On rappelle que d'après le chapitre sur les groupes, on a alors :

$$\forall (A, B) \in GL_n(\mathbb{K})^2,\ (AB)^{-1} = B^{-1}A^{-1}$$

Remarque 11.6. Deux matrices de $GL_n(\mathbb{K})$ n'ont aucune raison de commuter dès que $n \geqslant 2$. Pour $n = 2$, il suffit de considérer $\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$ et $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$. Pour l'ordre $n \geqslant 3$, il suffit de considérer les mêmes matrices mais complétées par un bloc I_{n-2} et des zéros partout ailleurs.

Méthode 11.1. Pour trouver un contre-exemple à l'ordre $n \geqslant 2$, on peut souvent trouver un contre-exemple pur $n = 2$ puis le généraliser pour $n \geqslant 2$ avec des blocs.

Par exemple, si P et Q sont des matrices inversibles, alors la matrice $\begin{pmatrix} P & (0) \\ (0) & Q \end{pmatrix}$ est inversible d'inverse $\begin{pmatrix} P^{-1} & (0) \\ (0) & Q^{-1} \end{pmatrix}$, comme on le vérifie immédiatement par un calcul par blocs.

Proposition 11.2. *Soit $A, B \in \mathcal{M}_n(\mathbb{K})$. Si $AB = BA$. Alors, on a :*

- *La **formule du binôme de Newton** :*

$$\forall p \in \mathbb{N},\ (A + B)^p = \sum_{k=0}^{p} \binom{p}{k} A^k B^{p-k}$$

- *La **formule de Bernoulli** :*

$$\forall p \in \mathbb{N}^\star,\ A^p - B^p = (A - B)\left(\sum_{k=0}^{p_1} A^{p-1-k}B^k\right) = \left(\sum_{k=0}^{p_1} A^{p-1-k}B^k\right)(A - B)$$

Remarque 11.7. On pensera souvent au cas particulier $B = I_n$ qui fonctionne toujours, puisque I_n commute avec toute matrice en tant que neutre.

Définition 11.9 (Trace)**.** Soit $A = (a_{i,j})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant n}} \in \mathcal{M}_n(\mathbb{K})$. La **trace** de A est le **scalaire** défini par :

$$\mathrm{Tr}(A) := \sum_{i=1}^{n} a_{i,i}$$

Autrement dit, la trace d'une matrice est la somme de ses coefficients diagonaux.

Remarque 11.8. Attention : la trace n'est définie que pour les matrices carrées !

Théorème 11.6. *L'application* $\mathrm{Tr}(\cdot)$ *est une forme linéaire non-nulle sur* $\mathcal{M}_n(\mathbb{K})$.

Exemple 11.8. Ainsi, $\ker(\mathrm{Tr}(\cdot))$ est un hyperplan de $\mathcal{M}_n(\mathbb{K})$. C'est donc un espace de dimension $n^2 - 1$.

Théorème 11.7 (Exercice classique / Utile pour exos difficiles, HP)**.** *Soit* $\varphi \in \mathcal{M}_n(\mathbb{K})^\star$. *Alors,*

$$\exists! A \in \mathcal{M}_n(\mathbb{K}),\ \varphi = \mathrm{Tr}(A \times \cdot)$$

Démonstration. Pour tout $A \in \mathcal{M}_n(\mathbb{K})$, on considère l'application

$$\begin{array}{rccc} \varphi_A & : & \mathcal{M}_n(\mathbb{K}) & \to & \mathbb{K} \\ & & M & \mapsto & \mathrm{Tr}(AM) \end{array}$$

puis on considère

$$\begin{array}{rccc} \theta & : & \mathcal{M}_n(\mathbb{K}) & \to & \mathcal{M}_n(\mathbb{K})^\star \\ & & A & \mapsto & \varphi_A \end{array}$$

Déjà, θ est bien définie puisque toutes les application φ_A sont des formes linéaires par linéarité à droite du produit matriciel et par linéarité de la trace.

Ensuite, θ est linéaire par linéarité à gauche du produit matriciel et par linéarité de la trace.

De plus, θ est injective. En effet, puisque θ est linéaire, soit $A \in \mathcal{M}_n(\mathbb{K})$ telle que φ_A soit la forme linéaire nulle. Alors pour tous $(i,j) \in [\![1,n]\!]^2$, on a $\varphi_A(E_{j,i}) = 0_{\mathbb{K}}$, ce qui après calcul, puisque

$$\begin{aligned} \varphi_A(E_{j,i}) &= \mathrm{Tr}(AE_{j,i}) \\ &= \sum_{k=1}^{n} (AE_{j,i})_{k,k} \\ &= \sum_{k=1}^{n}\sum_{l=1}^{n} a_{k,l}\delta_{j,l}\delta_{i,k} \\ &= a_{i,j} \end{aligned}$$

donne $a_{i,j} = 0_{\mathbb{K}}$. Ainsi, A est la matrice nulle, donc θ est injective.

Enfin, par égalité de dimensions, θ est bijective, ce qui conclut. □

Exercice 11.1 (Oral X/ENS)**.** On peut utiliser ce théorème pour prouver que pour $n \geqslant 2$, tout hyperplan de $\mathcal{M}_n(\mathbb{K})$ rencontre $GL_n(\mathbb{K})$ en écrivant cet hyperplan comme le noyau d'une $\mathrm{Tr}(A \times \cdot)$

avec A non nulle. Il faut alors raisonner sur le rang de A en l'écrivant comme une matrice équivalente à une matrice J_r. On pourra utiliser la matrice

$$N := \begin{pmatrix} 0 & & & 0 & 1 \\ 1 & \ddots & & (0) & 0 \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ (0) & & & 1 & 0 \end{pmatrix}$$

Théorème 11.8. *Soit $A, B \in \mathcal{M}_n(\mathbb{K})$. On a :*

$$\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$$

Remarque 11.9. Si $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B \in \mathcal{M}_{p,n}(\mathbb{K})$, on montre de même que $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$.

1.4 Matrices carrées particulières

Dans tout le paragraphe, on fixe $n \in \mathbb{N}^\star$ et on s'intéresse à des matrices particulières de $\mathcal{M}_n(\mathbb{K})$

Définition 11.10 (Matrice diagonale)**.** Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est dite **diagonale** lorsqu'elle est de la forme

$$\begin{pmatrix} \lambda_1 & & (0) \\ & \ddots & \\ (0) & & \lambda_n \end{pmatrix}$$

On la note alors $diag(\lambda_1, \ldots, \lambda_n)$ Formellement, A est diagonale lorsque tous ses coefficients non diagonaux sont nuls, *ie* lorsque

$$\forall (i,j) \in [\![1,n]\!]^2,\ i \neq j \implies a_{i,j} = 0_{\mathbb{K}}$$

On notera $D_n(\mathbb{K})$ l'ensemble des matrices diagonales d'ordre n (NS).

Définition 11.11 (Matrice scalaire)**.** Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est dite **scalaire** lorsque

$$\exists \lambda \in \mathbb{K},\ A = \lambda I_n$$

Les matrices scalaires sont donc en particulier des matrices diagonales.

Proposition 11.3 (Produit de matrices diagonales)**.** *Le produit de deux matrices diagonales est une matrice diagonale dont le coefficient en position (i,i) est le produit des coefficients en position (i,i) des deux matrices diagonales de départ.*

Théorème 11.9 (Produit diagonal par blocs)**.** *Soit A_1, A_2, B_1 et B_2 des matrices compatibles en taille. On a :*

$$\begin{pmatrix} A_1 & (0) \\ (0) & A_2 \end{pmatrix} \times \begin{pmatrix} B_1 & (0) \\ (0) & B_2 \end{pmatrix} = \begin{pmatrix} A_1B_1 & (0) \\ (0) & A_2B_2 \end{pmatrix}$$

Théorème 11.10. *$D_n(\mathbb{K})$ est une sous-algèbre commutative de $\mathcal{M}_n(\mathbb{K})$ de dimension n. En effet, on a $D_n(\mathbb{K}) = \mathrm{Vect}(E_{i,i})_{1 \leqslant i \leqslant n}$.*

Proposition 11.4. *$D_n(\mathbb{K}) \setminus \{0_{n,n}\}$ est un sous-groupe commutatif de*

Définition 11.12 (Matrice triangulaire supérieure)**.** Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est dite **triangulaire supérieure** lorsqu'elle est de la forme

$$\begin{pmatrix} * & & (*) \\ & \ddots & \\ (0) & & * \end{pmatrix}$$

Formellement, A est triangulaire supérieure lorsque

$$\forall (i,j) \in [\![1,n]\!]^2,\ i > j \implies a_{i,j} = 0_{\mathbb{K}}$$

On notera $T_n^+(\mathbb{K})$ l'ensemble des matrices triangulaires supérieures d'ordre n (NS).

Définition 11.13 (Matrice triangulaire supérieure stricte)**.** Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est dite **triangulaire supérieure stricte** lorsqu'elle est de la forme

$$\begin{pmatrix} 0 & & (*) \\ & \ddots & \\ (0) & & 0 \end{pmatrix}$$

Formellement, A est triangulaire supérieure stricte lorsque

$$\forall (i,j) \in [\![1,n]\!]^2,\ i \geqslant j \implies a_{i,j} = 0_{\mathbb{K}}$$

On notera $T_n^{++}(\mathbb{K})$ l'ensemble des matrices triangulaires supérieures strictes d'ordre n (NS).

Exemple 11.9. On a $T_n^{++}(\mathbb{K}) \subset T_n^+(\mathbb{K})$.

Définition 11.14 (Matrice triangulaire inférieure)**.** Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est dite **triangulaire inférieure** lorsqu'elle est de la forme

$$\begin{pmatrix} * & & (0) \\ & \ddots & \\ (*) & & * \end{pmatrix}$$

Formellement, A est triangulaire inférieure lorsque

$$\forall (i,j) \in [\![1,n]\!]^2,\ i < j \implies a_{i,j} = 0_{\mathbb{K}}$$

On notera $T_n^-(\mathbb{K})$ l'ensemble des matrices triangulaires inférieures d'ordre n (NS).

Définition 11.15 (Matrice triangulaire inférieure stricte)**.** Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est dite **triangulaire inférieure stricte** lorsqu'elle est de la forme

$$\begin{pmatrix} 0 & & (0) \\ & \ddots & \\ (*) & & 0 \end{pmatrix}$$

Formellement, A est triangulaire inférieure stricte lorsque

$$\forall (i,j) \in [\![1,n]\!]^2,\ i \leqslant j \implies a_{i,j} = 0_{\mathbb{K}}$$

On notera $T_n^{--}(\mathbb{K})$ l'ensemble des matrices triangulaires inférieures strictes d'ordre n (NS).

Exemple 11.10. On a $T_n^{--}(\mathbb{K}) \subset T_n^{-}(\mathbb{K})$.

Proposition 11.5. *Le produit de deux matrices triangulaires supérieures (resp. inférieures) est une matrice triangulaire supérieure (resp. inférieure) dont les termes diagonaux sont le produit des termes diagonaux correspondants.*

Théorème 11.11. *$T_n^{+}(\mathbb{K})$ et $T_n^{-}(\mathbb{K})$ sont des sous-algèbres de $\mathcal{M}_n(\mathbb{K})$ de dimension $n \times \dfrac{n+1}{2}$.*

Théorème 11.12. *$T_n^{++}(\mathbb{K})$ et $T_n^{--}(\mathbb{K})$ sont des sous-espaces vectoriels de $\mathcal{M}_n(\mathbb{K})$ de dimension $n \times \dfrac{n-1}{2}$.*

Remarque 11.10. On a $T_n^{+}(\mathbb{K}) \oplus T_n^{--}(\mathbb{K}) = \mathcal{M}_n(\mathbb{K})$ et $T_n^{-}(\mathbb{K}) \oplus T_n^{++}(\mathbb{K}) = \mathcal{M}_n(\mathbb{K})$.

Proposition 11.6. *Soit $T \in T_n^{+}(\mathbb{K}) \cap GL_n(\mathbb{K})$. Alors $T^{-1} \in T_n^{+}(\mathbb{K})$. Idem avec $T_n^{-}(\mathbb{K})$.*

Théorème 11.13 (HP)**.** *Plus généralement, soit $\mathcal{A}$ une sous-algèbre de $\mathcal{M}_n(\mathbb{K})$. Alors*

$$\forall A \in \mathcal{A} \cap GL_n(\mathbb{K}),\ A^{-1} \in \mathcal{A}$$

Démonstration. La preuve du théorème précédent s'adapte. Soit $A \in \mathcal{A} \cap GL_n(\mathbb{K})$. Posons

$$\begin{array}{rccc} f_A & : & \mathcal{A} & \to & \mathcal{A} \\ & & B & \mapsto & AB \end{array}$$

f_A est bien définie par stabilité de $\mathcal{A}$ par $\times$. Elle est linéaire par linéarité à droite par linéarité du produit matriciel, et elle est injective car $A \in GL_n(\mathbb{K})$ (il suffit de considérer un élément B du noyau de f_A puis multiplier $AB = 0_n$ par A^{-1} pour obtenir $B = 0_n$). Par égalité de dimensions, f_A est surjective donc il existe $B \in \mathcal{A}$ telle que $AB = I_n$. Donc $A^{-1} = B \in \mathcal{A}$. □

Exercice 11.2. Soit $A \in GL_n(\mathbb{K})$. Montrer que A^{-1} peut s'écrire comme un polynôme en A. Indication : considérer $\mathrm{Vect}(A^k)_{k\in\mathbb{N}}$ et montrer que c'est une sous-algèbre de $\mathcal{M}_n(\mathbb{K})$.

1.5 Transposition

Dans tout le paragraphe, on fixe $n, p \in \mathbb{N}^\star$.

Définition 11.16 (Transposée)**.** Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$. La **transposée de** A, notée $A^\top$ ou anciennement ${}^tA$, est la matrice de $\mathcal{M}_{p,n}(\mathbb{K})$ définie par :

$$A^\top := \begin{pmatrix} a_{1,1} & \dots & a_{n,1} \\ \vdots & & \vdots \\ a_{1,p} & \dots & a_{n,p} \end{pmatrix} = (a_{j,i})_{\substack{1\leqslant j\leqslant p \\ 1\leqslant i\leqslant n}}$$

Proposition 11.7 (Propriétés de la transposée)**.** *La transposée vérifie les propriétés suivantes :*

- $\forall A \in \mathcal{M}_{n,p}(\mathbb{K}),\ (A^\top)^\top = A$ *(involutivité)*
- $\forall (\lambda,\mu) \in \mathbb{K}^2,\ \forall (A,B) \in \mathcal{M}_n(\mathbb{K})^2,\ (\lambda A + \mu B)^\top = \lambda A^\top + \mu B^\top$ *(linéarité)*
- $\forall (A,B) \in \mathcal{M}_{n,p}(\mathbb{K}) \times \mathcal{M}_{p,q}(\mathbb{K}),\ (AB)^\top = B^\top A^\top$ *(transposée d'un produit)*
- $\forall A \in\ GL_n(\mathbb{K}),\ A^\top \in GL_n(\mathbb{K}) \wedge (A^\top)^{-1} = (A^{-1})^\top$ *(inversibilité)*

- $\forall A \in \mathcal{M}_n(\mathbb{K}),\ \mathrm{Tr}(A^\top) = \mathrm{Tr}(A)$ *(invariance de la trace)*

Remarque 11.11. L'application de $\mathcal{M}_{n,p}(\mathbb{K})$ dans $\mathcal{M}_{p,n}(\mathbb{K})$ qui à une matrice associe sa transposée est bijective. En effet, on peut exhiber sa bijection réciproque, qui est la même fonction mais avec les domaines inversés.

Définition 11.17 (Matrice symétrique)**.** Soit $A \in \mathcal{M}_n(\mathbb{K})$. A est dite **symétrique** lorsque $A^\top = A$. On note $S_n(\mathbb{K})$ l'ensemble des matrices symétriques

Définition 11.18 (Matrice antisymétrique)**.** Soit $A \in \mathcal{M}_n(\mathbb{K})$. A est dite **antisymétrique** lorsque $A^\top = -A$. On note $A_n(\mathbb{K})$ l'ensemble des matrices antisymétriques

Proposition 11.8. *$S_n(\mathbb{K})$ et $A_n(\mathbb{K})$ sont des sev de $\mathcal{M}_n(\mathbb{K})$.*

Théorème 11.14. *On a*

$$S_n(\mathbb{K}) \oplus A_n(\mathbb{K}) = \mathcal{M}_n(\mathbb{K})$$

En effet, l'application "transposition" (qu'on notera ici pour quelques instants s)est une symétrie vectorielle de $\mathcal{M}_n(\mathbb{K})$ et $S_n(\mathbb{K})$ et $A_n(\mathbb{K})$ sont respectivement $\ker(s - \mathrm{id})$ *et* $\ker(s + \mathrm{id})$.

Théorème 11.15. *On a* $\dim(S_n(\mathbb{K})) = n \times \dfrac{n+1}{2}$ *et* $\dim(A_n(\mathbb{K})) = n \times \dfrac{n-1}{2}$.

Démonstration. Soit $(i,j) \in [\![1,n]\!]^2$. Si $i = j$, on pose $F_{i,j} := E_{i,j}$. Si $i < j$, on pose $F_{i,j} := E_{i,j} + E_{j,i}$. $(F_{i,j})_{1 \leqslant i \leqslant j \leqslant n}$ est alors une famille libre de $S_n(\mathbb{K})$ donc $\dim(S_n(\mathbb{K})) \geqslant n \times \dfrac{n+1}{2}$.

Aussi, la famille $(E_{i,j} - E_{j,i})_{1 \leqslant < j \leqslant n}$ est une famille libre de $A_n(\mathbb{K})$ donc $\dim(A_n(\mathbb{K})) \geqslant n \times \frac{n-1}{2}$.

En sommant, par supplémentarité, on remarque que ces inégalités doivent être des égalités. Donc on a bien les égalités de dimensions souhaitées, et les familles que nous avons considérées sont en réalités des bases de leurs espaces respectifs. □

2 Matrices et applications linéaires

2.1 Principe de correspondance

Définition 11.19 (Matrice d'un vecteur dans une base)**.** Soit E de dimension $n \in \mathbb{N}^\star$ et $(e_1, \ldots, e_n)$ une base de E. A tout vecteur $x \in E$, on peut faire correspondre une matrice de $\mathcal{M}_{n,1}(\mathbb{K})$ appelée **matrice du vecteur x dans la base** (e_i), notée $\mathrm{mat}_{(e_i)}(x)$, et définie par :

$$\mathrm{mat}_{(e_i)}(x) := \begin{pmatrix} e_1^\star(x) \\ \vdots \\ e_n^\star(x) \end{pmatrix}$$

Proposition 11.9. *L'application*

$$\begin{array}{rccc} \mathrm{mat}_{(e_i)} & : & E & \to & \mathcal{M}_{n,1}(\mathbb{K}) \\ & & x & \mapsto & \mathrm{mat}_{(e_i)}(x) \end{array}$$

est un isomorphisme de $\mathbb{K}$-espaces vectoriels.

Exemple 11.11. Si on prend $E = \mathbb{K}^n$ et $(e_1, \dots, e_n)$ la base canonique de $\mathbb{K}^n$, puisqu'on identifie $\mathbb{K}^n$ et $\mathcal{M}_{n,1}(\mathbb{K})$, on a alors $x = \mathrm{mat}_{(e_i)}(x)$.

Remarque 11.12. Désormais, les lettres X, Y et Z seront réservées pour des matrices-colonne, tandis que A, B et C seront plutôt utilisées pour des matrices moins spécifiques.

Définition 11.20 (Matrice d'une famille dans une base)**.** Soit E de dimension $n \in \mathbb{N}^\star$ et $(e_1, \dots, e_n)$ une base de E. Soit $p \in \mathbb{N}^\star$ et $(x_1, \dots, x_p) \in E^p$. La matrice de la famille $(x_j)_{1 \leqslant j \leqslant p}$ dans la base $(e_i)_{1 \leqslant i \leqslant n}$ est définie en juxtaposant les matrices-colonne de x_j dans la base (e_i) :

$$\mathrm{mat}_{(e_i)}\left[(x_1, \dots, x_p)\right] := \left(\begin{array}{c|c|c} \mathrm{mat}_{(e_i)}(x_1) & \dots & \mathrm{mat}_{(e_i)}(x_p) \end{array}\right)$$

Autrement dit, $\mathrm{mat}_{(e_i)}\left[(x_1, \dots, x_p)\right]$ est la matrice des $e_i^\star(x_j)$.

Exemple 11.12. En reprenant $E = \mathbb{K}^n$ et (e_i) la base canonique, on a :

$$\mathrm{mat}_{(e_i)}\left[(x_1, \dots, x_p)\right] = \left(\begin{array}{c|c|c} x_1 & \dots & x_p \end{array}\right)$$

Définition 11.21 (Matrice d'un application linéaire relativement à une base de départ et à une base d'arrivée)**.** Soit E un espace vectoriel de dimension finie $p \in \mathbb{N}^\star$ et $(e_1, \dots, e_p)$ une base de E ; et F un espace vectoriel de dimension finie $n \in \mathbb{N}^\star$ et $(f_1, \dots, f_n)$ une base de F. Soit $u \in \mathcal{L}(E, F)$. On rappelle que u est entièrement caractérisée par l'image de la base (e_j) par u. La **matrice de** u **relativement aux base** (e_j) **et** (f_i) est la matrice définie par :

$$\mathrm{mat}_{(e_j),(f_i)}(u) := \mathrm{mat}_{(f_i)}\left[(u(e_1), \dots, u(e_p))\right]$$

Autrement dit, $\mathrm{mat}_{(e_j),(f_i)}(u)$ est la matrice des $f_i^\star(u(e_j))$. La base (e_j) est dite **base de départ** et la base (f_i) est dite **base d'arrivée**.

Théorème 11.16. *L'application*

$$\begin{array}{rccc} \mathrm{mat}_{(e_j),(f_i)} & : & \mathcal{L}(E,F) & \to & \mathcal{M}_{n,p}(\mathbb{K}) \\ & & x & \mapsto & \mathrm{mat}_{(e_j),(f_i)}(x) \end{array}$$

est un isomorphisme de $\mathbb{K}$*-espaces vectoriels.*

Remarque 11.13 (Cas particulier des endomorphismes)**.** Dans le cas d'un endomorphisme (on prend $E = F$), et si on prend $(e_i) = (f_i)$, la matrice $\mathrm{mat}_{(e_i),(e_i)}(u)$ sera plutôt notée $\mathrm{mat}_{(e_i)}(u)$.

Exemple 11.13 (Homothéties)**.** Soit $\lambda \in \mathbb{K}$ et $(e_i)_{1 \leqslant i \leqslant n}$ une base de E. On a :

$$\mathrm{mat}_{(e_i)}(\lambda \,\mathrm{id}_E) = \lambda I_n$$

Exemple 11.14 (Formes linéaires)**.** Soit $(e_1, \dots, e_p)$ une base de E et choisissons $(1_\mathbb{K})$ comme base du $\mathbb{K}$-ev $\mathbb{K}$. Soit $\psi \in E^\star$. On a :

$$\mathrm{mat}_{(e_j),(1_\mathbb{K})}(\psi) = \begin{pmatrix} \psi(e_1) & \dots & \psi(e_p) \end{pmatrix} \in \mathcal{M}_{1,p}(\mathbb{K})$$

On réserve donc les matrices-ligne aux formes linéaires.

Exemple 11.15 (Similitude directe linéaire)**.** On voit $\mathbb{C}$ comme un $\mathbb{R}$-ev de dimension 2, dont on choisit la base $(1, i)$. Soit $a \in \mathbb{C}^\star$ et on considère la similitude directe $s : z \mapsto az$ qui est un endomorphisme de $\mathbb{C}$. On a :

$$\mathrm{mat}_{(1,i)}(s) = \begin{pmatrix} \mathrm{Re}(a) & -\mathrm{Im}(a) \\ \mathrm{Im}(a) & \mathrm{Re}(a) \end{pmatrix}$$

Corollaire 11.1 (Dimension de $\mathcal{L}(E,F)$)**.** *$\mathcal{L}(E,F)$ est de dimension finie égale à* $\dim(E) \times \dim(F)$.

Remarque 11.14. On avait décidé dans le chapitre "Dimension finie", d'admettre ce fait car la démonstration est beaucoup plus aisée lorsqu'on dispose du théorème qui précède.

Définition 11.22 (Application linéaire canoniquement associée à une matrice)**.** Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$. **L'application linéaire canoniquement associée à** A (abrégée en **ALCA**) est l'unique $u \in \mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$ telle que

$$\mathrm{mat}_{(e_j),(f_i)}(u) = A$$

où (e_j) désigne la base canonique de $\mathbb{K}^p$ et (f_i) désigne la base canonique de $\mathbb{K}^n$.

Définition 11.23 (Image d'une matrice)**.** Soit A une matrice. L'**image de** A, notée $\mathrm{Im}(A)$, est définie comme étant égale à $\mathrm{Im}(u)$ où u est l'application linéaire canoniquement associée à A.

Définition 11.24 (Noyau d'une matrice)**.** Soit A une matrice. Le **noyau de** A, notée $\ker(A)$, est défini comme étant égal à $\ker(u)$ où u est l'application linéaire canoniquement associée à A.

Théorème 11.17 (Traduction matricielle d'une équation linéaire)**.** *"L'équation $u(x) = y$ se traduit matriciellement $AX = Y$." Formellement, soit $x \in E$, $u \in \mathcal{L}(E,F)$ et posons $y := u(x)$. Soit $(e_1, \ldots, e_p)$ une base de E et $(f_1, \ldots, f_n)$ une base de f. Posons $X := \mathrm{mat}_{(e_j)}(x) \in \mathcal{M}_{p,1}(\mathbb{K})$, $A := \mathrm{mat}_{(e_j),(f_i)}(u) \in \mathcal{M}_{n,p}(\mathbb{K})$ et $Y := \mathrm{mat}_{(f_i)}(y) \in \mathcal{M}_{n,1}(\mathbb{K})$. On a alors :*

$$AX = Y$$

Corollaire 11.2 (Formule explicite de l'ALCA)**.** *L'application linéaire canoniquement associée à $A \in \mathcal{M}_{n,p}(\mathbb{K})$ est l'application :*

$$\begin{array}{ccc} \mathbb{K}^p & \to & \mathbb{K}^n \\ X & \mapsto & AX \end{array}$$

Remarque 11.15 (Détermination pratique de l'image de A)**.** L'image de A est le sous-espace engendré par ses colonnes vues comme des vecteurs de $\mathbb{K}^n$.

Remarque 11.16 (Détermination pratique du noyau de A)**.** Les lignes de A fournissent un système d'équations du noyau de A.

Méthode 11.2. Soit $A \in \mathcal{M}_n(\mathbb{K})$. On voudrait savoir si A est inversible, et si oui, on voudrait calculer son inverse.

1. On résout $AX = Y$ pour $X, Y \in \mathbb{K}^n$. Si on obtient une unique solution, alors A est inversible. Si on n'obtient pas une unique solution quel que soit Y, alors A n'est pas inversible.
2. On se place dans le cas où A est inversible. Voici comment obtenir A^{-1}. Par hypothèse,

$$\forall Y \in \mathbb{K}^n,\ \exists! X \in \mathbb{K}^n,\ AX = Y$$

On remplace alors successivement Y par e_1, ..., e_n la base canonique de $\mathbb{K}^n$. On obtient alors à chaque fois une unique solution X_1, ..., X_n. On a alors :

$$A^{-1} = \left(\begin{array}{c|c|c} X_1 & \dots & X_n \end{array}\right)$$

Théorème 11.18. *"Composer les applications linéaires revient à multiplier les matrices."*

Formellement, soit E, F et G des $\mathbb{K}$-ev de dimension finie, $u \in \mathcal{L}(E,F)$ et $v \in \mathcal{L}(F,G)$, $\mathcal{B}_E$, $\mathcal{B}_F$ et $\mathcal{B}_G$ des bases respectives de E, F et G. Alors on a :

$$\mathrm{mat}_{\mathcal{B}_E,\mathcal{B}_G}(v \circ u) = \mathrm{mat}_{\mathcal{B}_F,\mathcal{B}_G}(v) \times \mathrm{mat}_{\mathcal{B}_E,\mathcal{B}_F}(u)$$

Corollaire 11.3. *Soit $(e_1, \dots, e_n)$ une base de E. L'application :*

$$\begin{array}{rccc} \mathrm{mat}_{(e_i)} & : & \mathcal{L}(E) & \to & \mathcal{M}_n(\mathbb{K}) \\ & & u & \mapsto & \mathrm{mat}_{(e_i)}(u) \end{array}$$

est un isomorphisme de $\mathbb{K}$-algèbres.

Théorème 11.19. *Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est inversible si, et seulement si, son ALCA est bijective.*

Théorème 11.20 (Interprétation géométrique des matrices par blocs)**.** *Soit $u \in \mathcal{L}(E)$ et E_1 et E_2 des sev supplémentaires dans E. Notons $(e_1, \dots, e_k)$ une base de E_1 et $(e_{k+1}, \dots, e_n)$ une base de E_2. Alors $(e_1, \dots, e_n)$ est une base adaptée à la décomposition $E = E_1 \oplus E_2$. Écrivons $\mathrm{mat}_{(e_i)}(u)$ sous la forme :*

$$\mathrm{mat}_{(e_i)}(u) = \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right)$$

Alors $B = 0_{k,n-k}$ si et seulement si, E_2 est stable par u, et $C = 0_{n-k,k}$ si, et seulement si, E_1 est stable par u.

2.2 Cas des matrices carrées

Théorème 11.21 (Lien matrice-base)**.** *Soit $(e_1, \dots, e_n)$ une base de E et $(x_1, \dots, x_n) \in E^n$ une famille quelconque. Posons $A := \mathrm{mat}_{(e_i)}[(x_1, \dots, x_n)]$. Alors : A est inversible si, et seulement si, la famille $(x_1, \dots, x_n)$ est une base de E.*

Exemple 11.16 (Important)**.** Une matrice carrée de taille n est donc inversible si, et seulement si, ses colonnes forment une base de $\mathbb{K}^n$. Cela est aussi équivalent, par égalité entre la taille de la famille et la dimension, au fait que ses colonnes forment une famille libre ou génératrice de $\mathbb{K}^n$.

Théorème 11.22 (Lien matrice-isomorphisme)**.** *Soit E et F de même dimension $n \in \mathbb{N}^\star$, $(e_1, \dots, e_n)$ une base de E et $(f_1, \dots, f_n)$ une base de F. Soit $u \in \mathcal{L}(E,F)$, et posons $A := \mathrm{mat}_{(e_j),(f_i)}(u) \in \mathcal{M}_n(\mathbb{K})$. Alors A est inversible si, et seulement si, u est un isomorphisme.*

Remarque 11.17. Dans ce cas, on a $\mathrm{mat}_{(f_i),(e_j)}(u^{-1}) = A^{-1}$.

Exemple 11.17 (Cas des endomorphismes)**.** Si $E = F$ et $(e_1, \dots, e_n) = (f_1, \dots, f_n)$, on obtient que $u \in GL(E)$ si, et seulement si, $\mathrm{mat}_{(e_i)}(u) \in GL_n(\mathbb{K})$.

Théorème 11.23 (Inversibilité des matrices triangulaires)**.** *Soit T une matrice triangulaire (supérieure ou inférieure). Alors T est inversible si, et seulement si, tous ses coefficients diagonaux sont non nuls.*

Théorème 11.24 (Nilpotence des triangulaires strictes)**.** *Soit T une matrice triangulaire stricte (supérieure stricte ou inférieure stricte) de taille $n \times n$. Alors T est nilpotente, d'indice de nilpotence inférieur ou égal à n.*

Exemple 11.18 (Calcul de puissances de matrices par le binôme de Newton et les matrices nilpotentes)**.** Lorsqu'on peut écrire $A = \lambda I_n + T$ avec T triangulaire stricte, on peut utiliser la nilpotence de T et le binôme de Newton pour calculer les puissances de A. Cela fonctionne plus généralement si on remplace T par une matrice nilpotente N.

3 Équivalence et similitude

3.1 Cas général

Définition 11.25 (Matrice de passage)**.** Soit E de dimension $n \in \mathbb{N}$, $(e_1, \dots, e_n)$ une base de E, et $(e'_1, \dots, e'_n)$ une base de E. On appelle **matrice de passage de la base** (e_i) **vers la base** (e'_i) la matrice :

$$P := \mathrm{mat}_{(e_i)}\left[(e'_1, \dots, e'_n)\right]$$

P est inversible (car $(e_1, \dots_e n)$ est une base) et on a $P = \mathrm{mat}_{(e'_i),(e_i)}(\mathrm{id}_E)$.

Remarque 11.18. $P^{-1} = \mathrm{mat}_{(e_i),(e'_i)}(\mathrm{id}_E)$ est la matrice de passage de la base (e'_i) vers la base (e_i).

Théorème 11.25 (Changement de base pour un vecteur)**.** *Soit E de dimension $n \in \mathbb{N}$, $(e_1, \dots, e_n)$ une base de E, $(e'_1, \dots, e'_n)$ une base de E et P la matrice de passage de la base (e_i) vers la base (e'_i). Soit $x \in E$ et posons $X := \mathrm{mat}_{(e_i)}(x)$ et $X' := \mathrm{mat}_{(e'_i)}(x)$. Alors on a :*

$$X = PX'$$

Théorème 11.26 (Changement de base pour une application linéaire)**.** *Soit E de dimension $p \in \mathbb{N}$, $(e_1, \dots, e_p)$ une base de E, $(e'_1, \dots, e'_p)$ une base de E et P la matrice de passage de la base (e_j) vers la base (e'_j). Soit F de dimension $n \in \mathbb{N}$, $(f_1, \dots, f_n)$ une base de F, $(f'_1, \dots, f'_n)$ une base de F et Q matrice de passage de la base (f_i) vers la base (f'_i). Soit $u \in \mathcal{L}(E, F)$. Posons $A := \mathrm{mat}_{(e_j),(f_i)}(u)$ et $A' := \mathrm{mat}_{(e'_j),(f'_i)}(u)$. Alors on a :*

$$A' = Q^{-1}AP$$

Définition 11.26 (Matrices équivalentes)**.** Soit $A, A' \in \mathcal{M}_{n,p}(\mathbb{K})$. A et A' sont dites **matriciellement équivalentes** lorsque :

$$\exists (M, N) \in GL_n(\mathbb{K}) \times GL_p(\mathbb{K}),\ A' = MAN$$

Théorème 11.27. *L'équivalence matricielle est une relation d'équivalence sur $\mathcal{M}_{n,p}(\mathbb{K})$.*

Théorème 11.28 ("2 matrices sont équivalentes ssi elles représentent une même application linéaire)**.** *Formellement, soit E de dimension $p \in \mathbb{N}$, $(e_1, \dots, e_p)$ une base de E et F de dimension $n \in \mathbb{N}$, $(f_1, \dots, f_n)$ une base de F. Soit $u \in \mathcal{L}(E, F)$ et posons $A := \mathrm{mat}_{(e_j),(f_i)}(u) \in \mathcal{M}_{n,p}(\mathbb{K})$.*

Soit $A' \in \mathcal{M}_{n,p}(\mathbb{K})$. Alors A' est équivalente à A si, et seulement si, il existe $(e'_1, \dots, e'_p)$ une base de E et $(f'_1, \dots, f'_n)$ une base de F telles que

$$A' = \mathrm{mat}_{(e'_j),(f'_i)}(u)$$

3.2 Cas des matrices carrées et des endomorphismes

Théorème 11.29. *Soit E de dimension $p \in \mathbb{N}$, $(e_1, \dots, e_p)$ une base de E, $(e'_1, \dots, e'_p)$ une base de E et P la matrice de passage de la base (e_i) vers la base (e'_i). Soit $u \in \mathcal{L}(E)$. Posons $A := \mathrm{mat}_{(e_i)}(u)$ et $A' := \mathrm{mat}_{(e'_i)}(u)$. Alors on a :*

$$A' = P^{-1}AP$$

Définition 11.27 (Matrices semblables)**.** Deux matrices $A; B \in \mathcal{M}_n(\mathbb{K})$ sont dites **semblables** lorsque :

$$\exists P \in GL_n(\mathbb{K}),\ B = P^{-1}AP$$

Théorème 11.30 ("2 matrices sont semblables ssi elles représentent un même endomorphisme)**.** *Formellement, soit E de dimension $n \in \mathbb{N}^\star$ et $(e_1, \dots, e_n)$ une base de E. Soit $u \in \mathcal{L}(E)$, et posons $A := \mathrm{mat}_{(e_i)}(u) \in \mathcal{M}_n(\mathbb{K})$. Soit $A' \in \mathcal{M}_n(\mathbb{K})$. Alors A' est semblable à A si, et seulement si, il existe une base $(e'_1, \dots, e'_n)$ de E telle que :*

$$A' = \mathrm{mat}_{(e'_i)}(u)$$

Exemple 11.19. Soit $\lambda \in \mathbb{K}$. La seule matrice semblable à la matrice λI_n est λI_n (il s'agit d'une analyse-synthèse rapide qui utilise le fait que λI_n commute avec toute matrice). Il suffit donc de prendre des λI_n et μI_n avec $\lambda \neq \mu$ pour trouver deux matrices non semblables.

Proposition 11.10 (Semblables $\implies$ équivalentes)**.** *Si deux matrices sont semblables, alors elles sont équivalentes.*

Remarque 11.19. Attention : la réciproque est fausse ! Déjà, ne serait-ce que pour une question de taille, mais même lorsque les tailles sont compatibles, on a des contre-exemples. Par exemple, $2I_n$ et I_n sont équivalentes car $2I_n = (\sqrt{2}I_n)I_n(\sqrt{2}I_n)$, mais ne sont pas semblables d'après l'exemple précédent.

Proposition 11.11 (Puissances naturelles de matrices semblables)**.** *Soit $A, B \in \mathcal{M}_n(\mathbb{K})$ semblables. Fixons $P \in GL_n(\mathbb{K})$ telle que $B = P^{-1}AP$. Par une récurrence immédiate, on a :*

$$\forall k \in \mathbb{N},\ B^k = P^{-1}A^kP$$

Remarque 11.20 (Diagonalisation, HP, spé)**.** Ce résultat est à la base de ce que l'on appelle la **diagonalisation** : on cherche si A peut être une matrice diagonale. En effet, on sait alors calculer très aisément les puissances de A, et donc celles de B. La diagonalisation sera développée dans le cours de spé.

Exercice 11.3. Montrer que toute matrice nilpotente est semblable à une matrice triangulaire stricte (supérieure stricte par exemple). On pourra procéder par récurrence sur la taille de la matrice, et on pourra utiliser l'ALCA dans l'hérédité pour se ramener à une matrice dont la première colonne est nulle en utilisant l'hypothèse de récurrence.

Exercice 11.4. Notons $\mathcal{N}_n$ l'ensemble des matrices carrées de taille $n \in \mathbb{N}^\star$ nilpotentes. Montre que la famille $(A)_{A\in\mathcal{N}_n}$ n'est pas génératrice de $\mathcal{M}_n(\mathbb{K})$.

Proposition 11.12 (Invariance de la trace par similitude)**.** *Deux matrices semblables ont même trace.*

Remarque 11.21 (Important)**.** Par contraposée, il suffit de montrer que deux matrices n'ont pas même trace pour montrer qu'elles ne sont pas semblables.

Définition 11.28 (Trace d'un endomorphisme en dimension finie)**.** Soit E de dimension finie $n \in \mathbb{N}^\star$ et $u \in \mathcal{L}(E)$. Le scalaire $\mathrm{Tr}(\mathrm{mat}_{(e_i)}(u))$ est indépendant de la base (e_i) de E choisie. Soit $(e_1, \dots, e_n)$ une base quelconque de E. La **trace de** u est définie par :

$$\mathrm{Tr}(u) := \mathrm{Tr}(\mathrm{mat}_{(e_i)}(u))$$

Proposition 11.13. *Soit E un $\mathbb{K}$-ev de dimension finie. Alors $\mathrm{Tr}(\cdot) : \mathcal{L}(E) \to \mathbb{K}$ est une forme linéaire.*

Proposition 11.14. *On a :*

$$\forall (u, v) \in \mathcal{L}(E),\ \mathrm{Tr}(v \circ u) = \mathrm{Tr}(u \circ v)$$

Proposition 11.15. *Soit E de dimension finie et $p \in \mathcal{L}(E)$ un projecteur. Alors on a :*

$$\mathrm{Tr}(p) = rg(p)$$

Remarque 11.22 (Utile pour exos X/ENS)**.** Pour montrer que la trace d'une application linéaire est dans $\mathbb{N}$, on peut se demander si cette application linéaire n'est pas une combinaison linéaire de projecteurs ou une composée de projecteurs. Réciproquement, si on travaille sur le rang d'une somme de projecteurs, il peut être utile de passer par la trace, notamment si on a des égalités du type $rg(p + q) = rg(p) + rg(q)$.

Proposition 11.16 (Trace et rang d'une symétrie, HP)**.** *Soit E de dimension finie $n \in \mathbb{N}^\star$ et $s \in \mathcal{L}(E)$ une symétrie de rang r. On a :*

$$\mathrm{Tr}(s) = 2rg(s) - n$$

Démonstration. Soit $(e_1, \dots, e_r)$ une base de $\ker(s - \mathrm{id}_E)$ et $(e_{r+1}, \dots, e_n)$ une base de $\ker(s + \mathrm{id}_E)$. Alors $(e_1, \dots, e_n)$ est une base adaptée à $E = \ker(s - \mathrm{id}_E) \oplus \ker(s + \mathrm{id}_E)$, et on a dans cette base :

$$\mathrm{mat}_{(e_i)}(s) = \left(\begin{array}{c|c} I_r & (0) \\ \hline (0) & -I_{n-r} \end{array}\right)$$

Si bien que

$$\mathrm{Tr}(s) = 2r - n = 2rg(s) - n$$

□

Exemple 11.20. Toute matrice $A \in \mathcal{M}_n(\mathbb{K})$ telle que $A^2 = A$ est semblable à une matrice par blocs de la forme :

$$\left(\begin{array}{c|c} I_r & (0) \\ \hline (0) & (0) \end{array}\right)$$

avec $r \in [\![0, n]\!]$. En effet, son ALCA est alors un projecteur, et on utilise le théorème "2 matrices sont équivalentes ssi elles représentent un même endomorphisme".

4 Rang d'une matrice

4.1 Définition et premières propriétés

Définition 11.29 (Rang d'une matrice)**.** Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$. Le **rang de la matrice** A, noté $rg(A)$, est défini des deux façons équivalentes qui suivent :

1. Le rang de A est défini comme étant égal au rang de la famille de ses colonnes interprétées comme des vecteurs de $\mathbb{K}^n$.
2. Le rang de A est défini comme étant égal au rang de son ALCA.

Exemple 11.21. Soit $(i, j) \in [\![1, n]\!] \times [\![1, p]\!]$. On a $rg(E_{i,j}) = 1$.

Proposition 11.17 (Minoration du rang par les tailles)**.** *On a :*

$$rg(A) \leqslant \min(n, p)$$

Théorème 11.31 (CNS d'inversibilité sur le rang, rang plein)**.** *Soit $A \in \mathcal{M}_n(\mathbb{K})$. On a :*

$$A \in GL_n(\mathbb{K}) \iff rg(A) = n$$

*On dit alors que A est de **rang plein**.*

Théorème 11.32 ("Le rang d'une famille de vecteurs est égal au rang de sa matrice dans n'importe quelle base")**.** *Soit E de dimension de $n \in \mathbb{N}^\star$, $(e_1, \dots, e_n)$ une base de E et $(x, 1, \dots, x_p) \in E^p$. Posons $A := \mathrm{mat}_{(e_i)}[(x_1, \dots, x_p)] \in \mathcal{M}_{n,p}(\mathbb{K})$. Alors on a :*

$$rg(A) = rg(x_j)_{1 \leqslant j \leqslant p}$$

Théorème 11.33 ("Le rang d'une application linéaire est égal au rang de sa matrice dans n'importe quel couple de bases")**.** *Soit E de dimension de $p \in \mathbb{N}^\star$, $(e_1, \dots, e_p)$ une base de E et F de dimension $n \in \mathbb{N}^\star$, $(f_1, \dots, f_n)$ une base de F. Soit $u \in \mathcal{L}(E, F)$. Posons $A := \mathrm{mat}_{(e_j),(f_i)}(u)$. Alors on a :*

$$rg(A) = rg(u)$$

Dans la séquence qui suit, on traduit matriciellement des résultats du chapitre "Dimension finie".

Théorème 11.34. *Soit $A, B \in \mathcal{M}_n(\mathbb{K})$. Si $AB = I_n$, alors $BA = I_n$ puis $A \in GL_n(\mathbb{K})$ et $A^{-1} = B$.*

Remarque 11.23. En particulier, il suffit de vérifier l'inverse d'un seul côté.

Théorème 11.35. *Soit $(A, B) \in \mathcal{M}_{n,p}(\mathbb{K}) \times \mathcal{M}_{p,q}(\mathbb{K})$. Alors on a :*

$$rg(AB) \leqslant \min(rg(A), rg(B))$$

Théorème 11.36 (Invariance du rang par multiplication par une matrice inversible)**.** *Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$. On a :*

- $\forall P \in GL_p(\mathbb{K}),\ rg(AP) = rg(A)$
- $\forall Q \in GL_n(\mathbb{K}),\ rg(QA) = rg(A)$

Exemple 11.22. Deux matrices équivalentes (et aussi deux matrices semblables), ont donc même rang.

Exemple 11.23. On a : $\forall \lambda \in \mathbb{K}^\star,\ rg(\lambda A) = rg(A)$.

4.2 Rang et équivalence

Définition 11.30 (Matrice J_r)**.** Soit $n, p \in \mathbb{N}^\star$ et $r \in [\![0, \min(n,p)]\!]$. On note $(e_1, \ldots, e_n)$ la base canonique de $\mathbb{K}^n$. On pose :

$$J_r^{(n,p)} := \mathrm{mat}_{(e_i)}\left[(e_1, \ldots, e_r, 0, \ldots, 0)\right] \in \mathcal{M}_{n,p}(\mathbb{K})$$

Autrement dit,

$$J_r^{(n,p)} := \left(\begin{array}{c|c} I_r & 0_{r,p-r} \\ \hline 0_{n-r,r} & 0_{n-r,p-r} \end{array}\right)$$

Si le contexte sur les tailles est clair, on pourra la noter J_r.

Exemple 11.24. On a :

- $J_1^{(n,p)} = E_{1,1}^{(n,p)}$
- $J_n^{(n,n)} = I_n$
- $J_0^{(n,p)} = 0_{n,p}$

Remarque 11.24. On a toujours $rg(J_r) = r$.

Lemme 11.1. *Soit E et F de dimensions respectives $p \in \mathbb{N}^\star$ et $n \in \mathbb{N}^\star$. Soit $u \in \mathcal{L}(E,F)$ et posons $r := rg(u)$. Alors, il existe $(e_1, \ldots, e_p)$ une base de E et $(f_1, \ldots, f_n)$ une base de F telles que :*

$$\mathrm{mat}_{(e_j),(f_i)}(u) = J_r^{(n,p)}$$

Théorème 11.37. *Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et $r \in [\![0, \min(n,p)]\!]$. Alors $rg(A) = r$ si, et seulement si, A est **matriciellement équivalente** à $J_r^{(n,p)}$.*

Remarque 11.25. Attention : ce théorème utilise bien la notion d'équivalence matricielle et non pas de similitude ! Ne pas dire cela avec la similitude, cela est faux. Il y a beaucoup plus de classes d'équivalences pour la relation de similitude, c'est l'objet de la réduction de Jordan (spé).

Corollaire 11.4. *Soit $A, B \in \mathcal{M}_{n,p}(\mathbb{K})$. Alors A et B sont équivalentes si, et seulement si, elles ont même rang.*

Exemple 11.25. Il y a donc exactement $1 + \min(n,p)$ classes d'équivalence pour la relation d'équivalence matricielle sur $\mathcal{M}_{n,p}(\mathbb{K})$, dont un système de représentant est fourni par les $J_r^{(n,p)}$ pour $r \in [\![0, \min(n,p)]\!]$.

Corollaire 11.5. *On a :*

$$\forall A \in \mathcal{M}_{n,p}(\mathbb{K}),\ rg(A^\top) = rg(A)$$

Corollaire 11.6. *Le rang d'une matrice est donc aussi égal au rang de la famille constituée de ses lignes vues comme des éléments de $\mathcal{M}_{1,p}(\mathbb{K})$.*

Exemple 11.26. Soit $A \in \mathcal{M}_n(\mathbb{K})$. Alors $A \in Gl_n(\mathbb{K})$ si, et seulement si, ses lignes forment une base de $\mathcal{M}_{1,n}(\mathbb{K})$ (ce qui est aussi équivalent au fait qu'elles en forment une famille libre ou génératrice).

4.3 Matrices extraites

Définition 11.31 (Matrice extraite/sous-matrice)**.** Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$. Une **matrice extraite**, ou **sous-matrice**, de A est une matrice obtenue en sélectionnant certaines lignes et certaines colonnes de A (pas nécessairement adjacentes).

Théorème 11.38. *Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et $r \in [\![0, \min(n,p)]\!]$. Les deux assertions suivantes sont équivalentes :*

1. *A admet une sous-matrice (r,r) inversible*
2. *$rg(A) \geqslant r$*

En particulier, $rg(A)$ est le plus grand entier r tel que l'on puisse trouver une matrice extraite de A de taille (r,r) inversible.

Corollaire 11.7. *Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et B une sous-matrice de A. Alors on a :*

$$rg(B) \leqslant rg(A)$$

4.4 Opérations élémentaires

Définition 11.32 (Opérations élémentaires)**.** On a trois types d'**opérations élémentaires sur les lignes** (abrégées en **OEL**), et de même symétriquement sur les colonnes :

- Transvections : $L_i \leftarrow L_i + \lambda L_j$ avec $\lambda \in \mathbb{K}^\star$ et $i \neq j$
- Échanges de lignes : $L_i \leftrightarrow L_j$
- Dilatations : $L_i \leftarrow \lambda L_i$ avec $\lambda \in \mathbb{K}^\star$

Lemme 11.2 (Effet de la multiplication par une $E_{i,j}$)**.** *Soit $(i,j) \in [\![1,n]\!]^2$. Dans le produit $E_{i,j}^{(n,n)} \times A$:*

- *$L_i \leftarrow L_j$*
- *Toutes les autres lignes sont mises à 0*

Dans le produit $A \times E_{i,j}^{(n,n)}$:

- *$C_j \leftarrow C_i$*
- *Toutes les autres colonnes sont mises à 0*

Théorème 11.39 (Traduction matricielle des OE)**.** *Une OE sur les lignes revient à multiplier à gauche par une matrice inversible. Une OE sur les colonnes revient à multiplier à droite par une matrice inversible. Chacune des 6 OE laisse le rang invariant.*

Méthode 11.3 (Détermination du rang d'une matrice par opérations élémentaires)**.** Voici comment calculer le rang de $A \in \mathcal{M}_{n,p}(\mathbb{K})$.

- Si $A = 0_{n,p}$, alors le rang est nul.
- Sinon, on peut trouver un terme $a_{i,j} \neq 0_{\mathbb{K}}$.
- Si A est une matrice-ligne ou une matrice-colonne, on sait alors que son rang vaut 1.
- Sinon, on ramène $a_{i,j}$ en position $(1,1)$ par une permutation de lignes et / ou une permutation de colonnes.
- Ce terme nous sert alors de **pivot**. En effectuant les transvections $L_i \leftarrow L_i - \dfrac{a_{i,1}}{a_{1,1}} L_1$ (pour $2 \leqslant i \leqslant n$), on annule le reste de la première colonne.
- Par le lemme sur le calcul du rang, on se ramène alors à une matrice de taille $(n-1, p-1)$ et on itère jusqu'à arriver à une matrice nulle ou une matrice-ligne ou une matrice-colonne, dont on connaît toujours le rang.

Théorème 11.40 (Matrices d'opérations élémentaires)**.** *Voici les matrices qui correspondent aux OE :*

- *Transvection $L_i \leftarrow L_i + \lambda L_j$ ou $C_j \leftarrow C_j + \lambda C_i$ avec $i \neq j$ et $\lambda \in \mathbb{K}^\star$:*

$$T_{i,j}(\lambda) := I_n + \lambda E_{i,j}$$

- *Échange $L_i \leftrightarrow L_j$ ou $C_i \leftarrow C_j$:*

$$Ech_{i,j} := I_n - E_{i,i} - E_{j,j} + E_{i,j} + E_{j,i}$$

- *Dilatation $L_i \leftarrow \lambda L_i$ ou $C_i \leftarrow \lambda C_i$ avec $\lambda \in \mathbb{K}\star$:*

$$D_i(\lambda) := I_n + (\lambda - 1) E_{i,i}$$

Remarque 11.26. Toute matrice inversible reste donc inversible après multiplication par une matrice d'OE.

Lemme 11.3. *Soit $n, p \geqslant 2$ et $A \in \mathcal{M}_{n,p}(\mathbb{K})$ de la forme :*

$$A = \left(\begin{array}{c|c} \lambda & (*) \\ \hline (0) & B \end{array}\right)$$

ou

$$A = \left(\begin{array}{c|c} \lambda & (0) \\ \hline (*) & B \end{array}\right)$$

avec $\lambda \in \mathbb{K}$ et $B \in \mathcal{M}_{n-1,p-1}(\mathbb{K})$. Si $\lambda \neq 0_{\mathbb{K}}$, alors

$$rg(A) = 1 + rg(B)$$

Proposition 11.18. *Soit $A \in GL_n(\mathbb{K})$. Par OEL (ie en multipliant A par des matrices d'OEL), on peut se ramener à une matrice triangulaire supérieure.*

Méthode 11.4 (Algorithme de Gauss-Jordan)**.** Voici comment inverser une matrice, ou du moins apprendre qu'elle n'est pas inversible.

- On trace deux colonnes : en haut de la colonne gauche, on écrit A, en haut de celle de droite, I_n.
- A partir de là, toutes les fois que l'on fera une opération élémentaire sur les lignes de A, on effectuera la même opération sur celles de la matrice de la colonne de droite.
- On s'intéresse d'abord à la première colonne de A.
  - Si tous les $a_{i,1}$ sont nuls, alors A n'est pas inversible et on s'arrête là.
  - Sinon, on choisit un pivot $a_{i,1} \neq 0_{\mathbb{K}}$ puis on permute L_1 et L_i.
- Par transvections, on annule le reste de la première colonne.
- On recommence en travaillant sur la matrice extraite

$$\begin{pmatrix} a_{2,2} & \dots & a_{2,n} \\ \vdots & & \vdots \\ a_{n,2} & \dots & a_{n,n} \end{pmatrix}$$

  - Si à un moment on ne peut pas trouver de pivot non nul, alors la matrice n'est pas inversible et on s'arrête là.
  - Sinon, on arrive à la fin à une matrice triangulaire supérieure dont la diagonale ne s'annule pas.
- Puis on "repart dans l'autre sens". Comme $a_{n,n} \neq 0_{\mathbb{K}}$, on l'utilise comme pivot pour annuler tous les $a_{i,n}$ (pour $1 \leqslant i \leqslant n-1$) par transvection.
- On recommence alors sur la matrice extraite

$$\begin{pmatrix} a_{1,1} & \dots & a_{1,n-1} \\ & \ddots & \vdots \\ 0 & & a_{n-1,n-1} \end{pmatrix}$$

  et ainsi de suite. A chaque fois, on tombe directement sur un pivot non nul (c'est le terme diagonal), il n'y a plus besoin de permuter les lignes.
- A la fin, on obtient une matrice diagonale, dont la diagonale ne s'annule pas.
- On dilate chaque ligne pour ramener les termes diagonaux à 1 (on peut aussi effectuer les dilatations avant les transvections si cela simplifie les calculs).
- Alors, la matrice écrite dans la colonne de droite est égale à A^{-1}.

Proposition 11.19 (Les OEL engendrent $GL_n(\mathbb{K})$, HP, exo classique)**.** *Les matrices d'OEL engendrent $GL_n(\mathbb{K})$. Cela fonctionne encore sans les échanges de ligne.*

Démonstration. En effet, soit $A \in GL_n(\mathbb{K})$. Puisque $A^{-1} \in GL_n(\mathbb{K})$, on peut passer de A^{-1} à I_n par multiplication à gauche par un nombre fini de matrices d'OEL : $M_N \times \ldots \times M_1 \times A^{-1} = I_n$. Donc $A = M_N \times \ldots \times M_1$ puis les matrices d'OEL engendrent $GL_n(\mathbb{K})$.Ensuite, on peut émuler l'échange de lignes $L_i \leftrightarrow L_j$ à partir des autres OEL en effectuant successivement : $L_i \leftarrow L_i + L_j$, $L'_j \leftarrow L'_j - L'_i$, $L''_i \leftarrow L''_i + L''_j$ et enfin $L'''_j \leftarrow -L'''_j$. □

5 Systèmes linéaires

5.1 Premières définitions

Définition 11.33 (Système linéaire, inconnues)**.** Un **système linéaire de n équations à p inconnues** (dans $\mathbb{K}$) est un système de la forme :

$$E : \begin{cases} a_{1,1}x_1 & + & \dots & + & a_{1,p}x_p & = & b_1 \\ \vdots & & & & \vdots & \vdots & \vdots \\ a_{n,1}x_1 & + & \dots & + & a_{n,p}x_p & = & b_n \end{cases}$$

Les $a_{i,j}$ et les b_i sont des scalaires fixés, ce sont les données du problème. Formellement, l'**inconnue** est le p-uplet $\begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} \in \mathbb{K}^p$, mais on a l'habitude de dire que les x_j sont les inconnues.

Remarque 11.27 (Traduction matricielle)**.** Posons

$$A := \begin{pmatrix} a_{1,1} & \dots & a_{1,p} \\ \vdots & & \dots \\ a_{n,1} & \dots & a_{n,p} \end{pmatrix} \in \mathcal{M}_{n,p}(\mathbb{K})$$

et $B := \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in \mathbb{K}^n$. Soit $X := \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} \in \mathbb{K}^p$. Alors X est solution du système si, et seulement si, $AX = B$.

Définition 11.34 (Rang d'un système linéaire)**.** Le **rang du système** est défini par $rg((E)) := rg(A)$.

Définition 11.35 (Système homogène)**.** Le **système homogène** (E_0) associé au système (E) est le système :

$$E : \begin{cases} a_{1,1}x_1 & + & \dots & + & a_{1,p}x_p & = & 0_{\mathbb{K}} \\ \vdots & & & & \vdots & \vdots & \vdots \\ a_{n,1}x_1 & + & \dots & + & a_{n,p}x_p & = & 0_{\mathbb{K}} \end{cases}$$

Matriciellement, il est équivalent à $AX = 0_{\mathbb{K}^n}$.

Théorème 11.41 (Structure de l'ensemble des solutions d'un système linéaire)**.** *Notons u l'ALCA à A, $\mathcal{S}$ l'ensemble des solutions de (E) et $\mathcal{S}_0$ l'ensemble des solutions de (E_0) (donc $\mathcal{S}_0 = \ker(u)$). On rappelle que*

- *Soit $\mathcal{S} = \emptyset$.*
- *Soit $\mathcal{S}$ est un sea dirigé par $\mathcal{S}_0$, où $\mathcal{S}_0$ est de dimension $p - rg((E))$ par le théorème du rang.*

*En quelque sorte, $p - rg((E))$ mesure le nombre de **degrés de liberté** du système (E).*

5.2 Différentes interprétations de A

Une première interprétation serait de savoir si $B \in \mathrm{Im}(A)$.

Une deuxième interprétation peut se faire en écrivant A sous la forme

$$\left(\begin{array}{c|c|c} C_1 & \dots & C_p \end{array}\right)$$

Alors, le système est équivalent à $\sum_{j=1}^{p} x_j \cdot C_j = B$ (produit par blocs). On cherche donc à savoir si B peut s'écrire comme combinaison linéaire des C_j, et les x_j sont précisément les coefficients de cette combinaison linéaire.

Une troisième façon de voir les choses est d'écrire A sous la forme

$$\left(\begin{array}{c} L_1 \\ \hline \vdots \\ \hline L_n \end{array}\right)$$

Alors, le système est équivalent à

$$\begin{cases} L_1 X = & b_1 \\ \vdots & \vdots \\ L_n X = & b_n \end{cases}$$

Soit $i \in [\![1, n]\!]$ et notons φ_i l'ALCA à L_iLe système homogène se réécrit :

$$\begin{cases} \varphi_1(X) = & b_1 \\ \vdots & \vdots \\ \varphi_n(X) = & b_n \end{cases}$$

ie $X \in \bigcap_{i=1}^{n} \ker(\varphi_i)$. A nouveau, soit $i \in [\![1, n]\!]$:

- Si $L_i = (0, \dots, 0)$, soit $b_i \neq 0_{\mathbb{K}}$, et alors on a une **incompatibilité** et $\mathcal{S} = \emptyset$; soit $b_i = 0_{\mathbb{K}}$ et il est équivalent de supprimer la ligne.
- Si $L_i \neq (0, \dots, 0)$, alors φ_i est non nulle, et comme c'est une forme linéaire, elle est surjective. Puis l'équation $L_i X = b_i$ admet au moins une solution puis l'ensemble des solutions de cette équation est un hyperplan affine dirigé par $\ker(\varphi_i)$.

En somme, si $\mathcal{S} \neq \emptyset$, il s'agit d'une intersection d'hyperplan affine.

5.3 Résolution

Dans toute cette partie, posons $r := rg(A)$.

Définition 11.36 (Système de Cramer)**.** Le système (E) est dit **de Cramer** lorsque $n = p = r$ (dans ce cas, $A \in GL_n(\mathbb{K})$). On a alors $AX = B \iff X = A^{-1}B$ donc $\mathcal{S} = \left\{A^{-1}B\right\}$.

Méthode 11.5 (Résolution d'un système linéaire par le pivot de Gauss)**.**

- On travaille sur les lignes du système par opérations élémentaires.
- On part du coin en haut à gauche de A et tant que le scalaire $a_{j,j}$ sur la diagonale est non nul (c'est le **pivot**), on effectue des transvections $L_i \leftarrow L_i - \dfrac{a_{i,j}}{a_{i,i}} L_j$ pour $i > j$ afin d'annuler la j-ème colonne en-dessous de $a_{j,j}$.
- Si à l'étape j on ne peut pas trouver de pivot non nul, c'est qu'on a :

$$A = \begin{pmatrix} q_1 & & & & & \dots & \dots & \dots \\ 0 & \ddots & & & & \dots & \dots & \dots \\ & \ddots & \ddots & & & \dots & \dots & \dots \\ & & \ddots & q_{j-1} & & \dots & \dots & \dots \\ 0 & \dots & \dots & 0 & 0 & \dots & \dots & \dots \\ \vdots & & & \vdots & \vdots & \dots & \dots & \dots \\ 0 & \dots & \dots & 0 & 0 & \dots & \dots & \dots \end{pmatrix}$$

  où les $q_1, \dots, q_{j-1}$ sont des pivots non nuls. On essaye alors de permuter la j-ème colonne avec une autre colonne à sa droite (il suffit d'échanger deux inconnues), à la recherche d'un pivot non nul quelque par dans la sous matrice en bas à droite, qui commence juste une case en-dessous et à droite de q_{j-1}. Puis on continue.
- Si à un moment ce n'est pas possible, alors c'est que pour toutes les lignes pour $j \leqslant i \leqslant n$ dans le membre de gauche de l'équation du système sont nulles. Notamment, on arrivera, dans cette situation à chaque fois que $n > p$, c'est-à-dire s'il y a plus d'équations que d'inconnues. On est alors face à ce qu'on appelle un **risque d'incompatibilité** :
  - Si le second membre correspondant n'est pas nul, alors on a une contradiction, et il n'y a pas de solution.
  - Sinon, on est face à des lignes du type "$0 = 0$", et il est équivalent de les supprimer.
- On arrive finalement à une situation du type :

$$\begin{pmatrix} q_1 & \dots & \dots & \dots & \dots & \dots \\ & \ddots & & & & \\ & & \ddots & & & \\ (0) & & & q_r & \dots & \dots \end{pmatrix}$$

  avec $q_1, \dots, q_r \neq 0_{\mathbb{K}}$ et r le rang du système initial.
- Si $p = r$, on a un système de Cramer **échelonné** qu'on résout de bas en haut (on peut trouver la dernière inconnue, puis on réinjecte en remontant pour obtenir toutes les inconnues une par une). Si $p > r$, les r première variables sont alors appelées **variables principales**, et les $p - r$ restantes **variables auxiliaires**. On fait passer à droites les variables auxiliaires, qui deviennent des paramètres, et on se ramène au cas précédent.

5.4 Récapitulatif

Récapitulons les fois où nous nous sommes servis d'un pivot de Gauss ou d'une de ses variantes.

- Trouver une relation de liaison entre des vecteurs de $\mathbb{K}^n$, ou montrer qu'une famille de vecteurs de K^n est libre (ce qui est essentiellement la même chose)
- Inverser $A \in GL_n(\mathbb{K})$
- Calculer le rang de $A \in \mathcal{M}_{n,p}(\mathbb{K})$
- Passer de l'écriture d'un sev de $\mathbb{K}^n$ sous forme d'un système d'équations à une écriture sous la forme d'un Vect, et vice-versa
- Extraire une base du Vect d'une famille de vecteurs de $\mathbb{K}^n$. Profitons-en pour rappeler la méthode, qui n'est pas décrite dans ce document jusqu'à présent :

Méthode 11.6 (Extraire d'une famille de vecteurs une base de son Vect)**.** Soit $(C_1, \dots, C_p)$ une famille de vecteurs de $\mathbb{K}^n$. Posons $F := \text{Vect}(C_j)_{1 \leqslant j \leqslant p}$ et cherchons une base de F sous la forme $(C_{j_1}, C_{j_2}, \dots)$ (qui existe d'après le théorème de la base extraite). Notons $r := rg(C_j)_{1 \leqslant j \leqslant p}$ (*ie* $r := \dim(F)$).

- On résout le système $\sum_{j=1}^{p} x_j C_j = 0$ *ie* $A \times \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} = 0$, avec

$$A = \left(\begin{array}{c|c|c} C_1 & \dots & C_p \end{array}\right)$$

- On repère les "indices principaux".

On rappelle maintenant la justification de cette méthode, à reproduire au cas par cas une copie :

- SPG, supposons que les variables principales sont $x_1, \dots, x_r$ (quitte à les renommer). On a alors $rg(A) = r$. En effet, $rg(A)$ est alors égal à r en appliquant l'algorithme de détermination du rang. **Note** : dans un exercice calculatoire, ce point est inutile.
- Montrons que $(C_1, \dots, C_r)$ est une base de F.
  - Liberté : soit $x_1, \dots, x_r \in \mathbb{K}$ tels que $x_1 C_1 + \dots + x_r C_r = 0_{\mathbb{K}^n}$. On pose artificiellement $x_{r+1} = \dots = x_p = 0_{\mathbb{K}}$. Le vecteur $\begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$ est alors solution du système homogène résolut précédemment. Mais puisque $x_1, \dots, x_r$ s'expriment de façon linéaire en fonction des $x_{r+1}, \dots, x_p$, on a $x_1 = \dots = x_p = 0_{\mathbb{K}}$. La famille est libre.
  - Caractère générateur : jusqu'à présent, on montrait que $\dim(F) = r$, ce qui suffisait à conclure. Pour ce faire, on montrait que $\forall j \in [\![r+1, p]\!],\ C_j \in \text{Vect}(C_l)_{1 \leqslant l \leqslant r}$ en posant $x_j = -1$ et 0 pour les autres et on réinjectait dans le système. Mais désormais, cela est inutile en vertu de l'algorithme de calcul du rang qui nous assure immédiatement cette égalité puisque nous avons résolu le système.

Chapitre 12

Déterminants

1 Groupe symétrique

1.1 Généralités

Définition 12.1 (Groupe symétrique d'ordre n)**.** Les permutations de $[\![1,n]\!]$ forment un groupe pour la composition. On l'appelle le **groupe symétrique d'ordre** n et on le note $\mathcal{S}_n$. Assez rapidement, on omettra le symbole $\circ$.

Définition 12.2 (Notation d'une permutation de $[\![1,n]\!]$)**.** Souvent, une permutation σ de $\mathcal{S}_n$ sera notée sous la forme d'un tableau à deux lignes : en haut, les entiers de 1 à n ; en bas, leurs images par σ. Par exemple, dans $\mathcal{S}_5$, on peut avoir

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 1 & 5 & 2 \end{pmatrix}$$

Définition 12.3 (Support, points fixes)**.** Soit $\sigma \in \mathcal{S}_n$. L'ensemble des $x \in [\![1,n]\!]$ tels que $\sigma(x) \neq x$ s'appelle le **support** de σ. Les x tels que $\sigma(x) = x$ sont les **points fixes** de σ.

Remarque 12.1. Puisque σ est injective, elle stabilise son support.

Proposition 12.1. *Deux permutations à supports disjoints commutent.*

Définition 12.4 (Orbite)**.** Soit $\sigma \in \mathcal{S}_n$. La relation binaire

$$\forall (x,y) \in [\![1,n]\!],\ x \sim y \iff \exists k \in \mathbb{Z},\ y = \sigma^k(x)$$

est une relation d'équivalence. La classe de x est appelée **orbite** de x. De plus, il existe un plus petit $p > 0$ tel que $\sigma^p(x) = x$, et l'orbite de x est alors égale à $\left\{x, \sigma(x), \ldots, \sigma^{p-1}(x)\right\}$, avec les $\sigma^k(x)$ deux à deux distincts pour $0 \leqslant k \leqslant p-1$.

Exemple 12.1. Si x est un point fixe, son orbite est réduite à $\{x\}$.

Définition 12.5 (p-cycle)**.** Soit $p \geqslant 2$ et $a_1, \ldots, a_p$ des entiers deux à deux distincts dans $[\![1,n]\!]$. L'application σ définie sur $[\![1,n]\!]$ définie par :

1. $\forall x \notin \{a_1, \ldots, a_p\},\ \sigma(x) = x$
2. $\forall i \in [\![1, p-1]\!],\ \sigma(a_i) = a_{i+1}$
3. $\sigma(a_p) = a_1$

est une permutation de $[\![1, n]\!]$ notée $(a_1 \ldots a_p)$. Une telle permutation est appelée **p-cycle** ou **cycle d'ordre** p.

Remarque 12.2. De manière imagée, on peut imaginer que les a_i sont placés sur le tour d'un cadran. Appliquer la permutation revient alors à "faire tourner le cadran" d'un cran. C'est ce qui explique qu'un cycle s'appelle aussi une **permutation circulaire**.

Remarque 12.3. Le support du cycle $(a_1 \ldots\ a_p)$ est exactement égal à $\{a_1, \ldots, a_p\}$

Remarque 12.4. On a $(a_1 \ldots\ a_p) = (a_2 \ldots\ a_p a_1) = \ldots = (a_p a_1 \ldots\ a_{p-1})$

Remarque 12.5. On a $(a_1 \ldots\ a_p)^{-1} = (a_p a_{p-1} \ldots\ a_1)$

Théorème 12.1 (Décomposition en produit de cycles à supports disjoints)**.** *A l'ordre près des facteurs, toute permutation admet une unique décomposition en produit de cycles à supports disjoints.*

Remarque 12.6. Puisque les cycles sont à supports disjoints, ils commutent entre eux. Cela justifie l'expression "à l'ordre des facteurs près".

Exemple 12.2. Reprenons $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 1 & 5 & 2 \end{pmatrix}$. 1 s'envoie sur 3, 3 s'envoie sur 1. Voilà notre premier cycle, c'est (1 3) Puis 2 s'envoie sur 4, qui s'envoie sur 5, qui s'envoie sur 2. Finalement, $\sigma = (1\ 3)(2\ 4\ 5)$.

1.2 Signature

Définition 12.6 (Transposition)**.** Un 2-cycle s'appelle plus couramment une **transposition**. La transposition $(a\ b)$ consistant à "changer a et b, elle est involutive.

Remarque 12.7. Pour $n \geqslant 3$, $\mathcal{S}_n$ n'est pas commutatif. Il suffit de prendre a, b et c distincts et de considérer $(a\ b)$ et $(b\ c)$ puis de tester les deux compositions qui ne sont alors pas égales. En revanche, pour $n = 1$ ou $n = 2$, on observe immédiatement que $\mathcal{S}_n$ est bien commutatif.

Théorème 12.2 (Décomposition en produit de transpositions)**.** *Toute permutation s'écrit comme produit de transpositions : $\sigma = t_1 \ldots t_k$*

Théorème 12.3 (Signature, permutations paires et impaires)**.** *Il existe un unique morphisme de groupes $\varepsilon : \mathcal{S}_n \to \{-1, 1\}$ qui vaut -1 sur les transpositions. Le nombre $\varepsilon(\sigma)$ s'appelle la **signature** de σ. Si $\varepsilon(\sigma) = 1$ (resp. $\varepsilon(\sigma) = -1$), on dit que la permutation σ est **paire** (resp. **impaire**).*

Remarque 12.8. $\{-1, 1\}$ est bien un groupe en tant que groupe des unités de l'anneau $(\mathbb{Z}, +, \times)$.

Corollaire 12.1. *Dans la décomposition de $\sigma \in \mathcal{S}_n$ en produit de transpositions $\sigma = t_1 \ldots t_k$, la parité de k est indépendante de la décomposition (en effet, il suffit de passer à la signature).*

Remarque 12.9. Un p-cycle a une signature qui vaut $(-1)^{p-1}$. En pratique, on pourra l'utiliser pour déterminer la signature d'une permutation après l'avoir décomposée en produit de cycles.

Définition 12.7 (Groupe alterné d'ordre n, HP)**.** Les permutations paires forment un sous-groupe de $\mathcal{S}_n$, puisqu'il s'agit de $\ker(\varepsilon)$. Ce sous-groupe est appelé **groupe alterné d'ordre** n, on le note $\mathcal{A}_n$.

Proposition 12.2 (Cardinal de $\mathcal{A}_n$ dès que $n \geqslant 2$, HP)**.** *Supposons que $n \geqslant 2$. Si τ est une transposition fixée, l'application $\sigma \mapsto \sigma\tau$ envoie $\mathcal{A}_n$ sur $\mathcal{S}_n \setminus \mathcal{A}_n$ comme on le constate par un simple calcul de signature. De plus, elle est bijective car involutive. On en déduit alors que*

$$\mathrm{Card}(\mathcal{A}_n) = \frac{\mathrm{Card}(\mathcal{S}_n)}{2} = \frac{n!}{2}$$

2 Déterminant

Dans toute cette partie, $\mathbb{K}$ désigne un corps et E un $\mathbb{K}$-espace vectoriel de dimension n.

2.1 Déterminant d'une famille de vecteurs

Définition 12.8 (Formes p-linéaires)**.** Une application $f : E^p \to \mathbb{K}$ est une forme p-linéaire lorsqu'elle est linéaire en chacune de ses variables, ie lorsque l'on fige $p-1$ variables, l'application est linéaire par rapport à la p-ème variable.

Remarque 12.10. Dès que l'un des x_j est nul, on a $f(x_1, \ldots, x_p) = 0$

Remarque 12.11. Les formes p-linéaires forment un sous-espace vectoriel de $\mathbb{K}^{E^p}$.

Exemple 12.3. Pour $p = 1$, on retrouve la définition d'une forme linéaire.

Exemple 12.4. Dans $\mathcal{M}_n(\mathbb{K})$, l'application $(A_1, \ldots, A_p) \mapsto \mathrm{Tr}(A_1 \times \ldots \times A_p)$ est une forme p-linéaire.

Définition 12.9 (HP)**.** Si $f_1, \ldots, f_p$ sont des formes linéaires de E, alors l'application

$$\begin{array}{rccc} f_1 \otimes \ldots \otimes f_p : & E^p & \longrightarrow & \mathbb{K} \\ & (x_1, \ldots, x_p) & \longmapsto & f_1(x_1) \times \ldots \times f_p(x_p) \end{array}$$

est une forme p-linéaire.

Lemme 12.1. *Soit $(e_i)_{1 \leqslant i \leqslant n}$ une base de E, f une forme p-linéaire et $x_1, \ldots, x_p$ des vecteurs qu'on écrit sous la forme*

$$x_j = \sum_{i=1}^{n} \lambda_{i,j} e_i$$

On a alors

$$f(x_1, \ldots, x_p) = \sum_{1 \leqslant i_1, \ldots, i_p \leqslant n} \lambda_{i_1,1} \ldots \lambda_{i_p,p} f(e_{i_1}, \ldots, e_{i_p})$$

Remarque 12.12. On pourrait généraliser le lemme : si $(u_{i,1})_{i \in I_1}, \ldots, (u_{i,p})_{i \in I_p}$ sont des familles finies de E, alors on a

$$f\left(\sum_{i \in I_1} \lambda_{i,1} u_{i,1}, \ldots, \sum_{i \in I_1} \lambda_{i,1} u_{i,1}\right) = \sum_{(i_1, \ldots, i_p) \in I_1 \times \ldots \times I_p} \lambda_{i_1,1} \ldots \lambda_{i_p,p} f(u_{i_1,1}, \ldots, u_{i_p,p})$$

A partir de maintenant, on suppose que $\boxed{\text{p=n}}$

Définition 12.10 (Forme n-linéaire alternée)**.** Une forme n-linéaire est dite **alternée** lorsqu'elle s'annule dès que sont égaux :

$$\forall i \neq j,\ (x_i = x_j \implies f(x_1, \ldots, x_n) = 0)$$

Remarque 12.13. Les formes n linéaires alternées forment un sous-espace vectoriel des formes n-linéaires.

Théorème 12.4 (Invariance par transvection des formes n-linéaires alternées)**.** *Si f est une forme n-linéaire alternée, alors le scalaire $f(x_1, \ldots, x_n)$ est invariant si on ajoute à un x_j une combinaison linéaire des autres x_i.*

Théorème 12.5 (Échange de termes, forme antisymétrique)**.** *Si f est une forme n-linéaire alternée, alors elle vérifie*

$$\forall i \neq j,\ f(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_n) = -f(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_n)$$

*où on a échangé les positions de x_i et x_j. Une telle forme est dite **antisymétrique**.*

Remarque 12.14 (HP)**.** En caractéristique différente de 2, la réciproque est vraie : toute forme antisymétrique est alternée. Il suffit d'échanger les deux termes identiques puis de conclure par pseudo-intégrité, car dans ce cas, $2 \neq 0$.

Corollaire 12.2. *Si f est alternée, alors elle vérifie*

$$\forall \sigma \in \mathcal{S}_n,\ f(x_{\sigma(1),\ldots,x_{\sigma(n)}}) = \varepsilon(\sigma) f(x_1, \ldots, x_n)$$

Lemme 12.2. *Soit $(e_i)_{1 \leqslant i \leqslant n}$ une base de E, f une forme n-linéaire alternée et $x_1, \ldots, x_n$ des vecteurs qu'on écrit sous la forme*

$$x_j = \sum_{i=1}^{n} \lambda_{i,j} e_i$$

Alors on a :

$$f(x_1, \ldots, x_n) = \left(\sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} \lambda_{\sigma(i),i} \right) f(e_1, \ldots, e_n)$$

Définition 12.11 (Déterminant dans une base)**.** Soit $(e_i)_{1 \leqslant i \leqslant n}$ une base de E. Alors, il existe une unique forme n-linéaire alternée f telle que $f(e_1, \ldots, e_n) = 1$. On l'appelle **déterminant dans la base** $(e_i)_{1 \leqslant i \leqslant n}$, et on la note $\det_{(e_i)}$. Elle vérifie :

$$\det_{(e_i)}(x_1, \ldots, x_n) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} \lambda_{\sigma(i),i}$$

Corollaire 12.3. *Toute forme n-linéaire alternée est donc proportionnelle à $\det_{(e_i)}$. Le sous-espace vectoriel des formes n-linéaires alternées est donc de dimension 1, engendré par $\det_{(e_i)}$.*

Corollaire 12.4 (Relation de Chasles pour les déterminants)**.** *Soit A et B deux bases de E. Alors*

$$\det_A(\cdot) = \det_A(B) \times \det_B(\cdot)$$

Théorème 12.6 (Lien entre déterminants et bases)**.** *Soit $(e_i)_{1\leqslant i\leqslant n}$ une base de E et $(x_1,\dots,x_n)$ une famille de vecteurs de E. Alors $(x_1,\dots,x_n)$ est une base de E si, et seulement si,*

$$\det_{(e_i)}(x_1,\dots,x_n)\neq 0$$

Remarque 12.15. C'est là la principale utilité du déterminant : savoir si une famille est une base, ou savoir si une matrice est inversible (ce qui revient au même, le second point étant une traduction du premier). **Se rappeler de montrer que la matrice d'une famille dans une certaine base est inversible permet donc de montrer efficacement que cette famille est une base, sans passer par les caractères libre, générateur ou des questions de taille et de dimension. C'est une méthode très importante à retenir.**

2.2 Déterminant d'un endomorphisme

Lemme 12.3. *Soit $u\in\mathcal{L}(E)$ et $(e_i)_{1\leqslant i\leqslant n}$ une base de E. On a*

$$\forall(x_1,\dots,x_n)\in E^n,\ \det_{(e_i)}(u(x_j))=\det_{(e_i)}(u(e_i))\times\det_{(e_i)}(x_j)$$

Définition 12.12 (Déterminant d'un endomorphisme)**.** Soit $u\in\mathcal{L}(E)$ et $(e_i)_{1\leqslant i\leqslant n}$ une base de E. Le scalaire $\det_{(e_i)}(u(e_i))$ est indépendant du choix de la base $(e_i)_{1\leqslant i\leqslant n}$. On l'appelle le **déterminant de** u et on le note $\det(u)$.

Proposition 12.3. *Avec les mêmes hypothèses, on en tire l'identité :*

$$\forall(x_1,\dots,x_n)\in E^n,\ \det_{(e_i)}(u(x_1),\dots,u(x_n))=\det(u)\det_{(e_i)}(x_1,\dots,x_n)$$

Exemple 12.5. En particulier, on a $\det(\mathrm{id})=1$.

Théorème 12.7. *u est un automorphisme de E si, et seulement si,* $\det(u)\neq 0$.

Théorème 12.8 (Propriétés du déterminant des endomorphismes)**.** *On a les relations suivantes :*

1. $\forall\lambda\in\mathbb{K},\ \forall u\in\mathcal{L}(E),\ \det(\lambda u)=\lambda^n\det(u)$
2. $\forall(u,v)\in\mathcal{L}(E)^2,\ \det(u\circ v)=\det(u)\det(v)$

Corollaire 12.5. *Si u est inversible, alors on a* $\det(u^{-1})=\dfrac{1}{\det(u)}$.

Remarque 12.16. Attention, on a aucune raison d'avoir $\det(u+v)=\det(u)+\det(v)$!

2.3 Déterminant d'une matrice carrée

Définition 12.13 (Déterminant d'une matrice carré - Formule des signatures)**.** Soit $A\in\mathcal{M}_n(\mathbb{K})$. Le **déterminant** de A, noté $\det(A)$, est défini comme le déterminant des colonnes de A dans la base canonique de $\mathbb{K}^n$:

$$\det(A)=\sum_{\sigma\in\mathcal{S}_n}\varepsilon(\sigma)\prod_{i=1}^{n}a_{\sigma(i),i}$$

Remarque 12.17. Attention le déterminant d'une matrice non carrée n'a aucun sens !

Exemple 12.6. Si $A = (\lambda) \in \mathcal{M}_1(\mathbb{K})$, alors $\det(A) = \lambda$. Si $n = 0$, les différentes conventions sur l'ensemble vide montrent que $\det(A) = 1$.

Proposition 12.4. *Soit $(e_i)_{i\leqslant i\leqslant n}$ une base de E et $(x_i)_{1\leqslant i\leqslant n}$ une famille de vecteurs. Alors :*

$$\det_{(e_i)}(x_j) = \det(\mathrm{mat}_{(e_i)}(x_j))$$

De même, si $u \in \mathcal{L}(E)$, on a

$$\det(u) = \det(\mathrm{mat}_{(e_i)}(u))$$

Théorème 12.9 (Règle du gamma - Calcul d'un déterminant 2×2)**.** *On a la formule*

$$\begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} = x_1y_2 - y_1x_2$$

Théorème 12.10 (Règle de Sarrus - Calcul d'un déterminant 3×3)**.** *On a la formule*

$$\begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix} = x_1y_2z_3 + y_1z_2x_3 + z_1x_2y_3 - x_1z_2y_3 - y_1x_2z_3 - z_1y_2x_3$$

Pour la retenir, on fait toutes les diagonales. Celles dans le sens de la première bissectrice du repère prennent un moins, celles dans le sens de la seconde bissectrice du repère prennent un plus.

Théorème 12.11 (Propriétés du déterminant matriciel)**.** *On a les relations suivantes :*

1. $\forall \lambda \in \mathbb{K},\ \forall A \in \mathcal{M}_n(\mathbb{K}),\ \det(\lambda A) = \lambda^n \det(A)$
2. $\forall (A, B) \in \mathcal{M}_n(\mathbb{K})^2,\ \det(AB) = \det(A)\det(B)$

De plus, A est inversible si, et seulement si, $\det(A) \neq 0$*, et dans ce cas on a*

$$\det(A^{-1}) = \frac{1}{\det(A)}$$

Remarque 12.18. Attention, on a aucune raison d'avoir $\det(A + B) = \det(A) + \det(B)$!

Proposition 12.5 (Invariance par similitude du déterminant)**.** *Le déterminant est un invariant de similitude.*

Proposition 12.6 (Formules de Cramer, HP)**.** *Soit un système de Cramer $AX = B$. On dispose des formules de Cramer*

$$\forall i \in [\![1, n]\!],\ x_i = \frac{\det(A_1|\dots|A_{i-1}|B|A_{i+1}|\dots|A_n)}{\det(A)}$$

Démonstration. En effet, on sait qu'une des interprétations possibles du système de Cramer consiste à chercher B comme combinaison linéaire des colonnes de A. Ici, on cherche donc $x_1, \dots, x_n$ tels que $B = x_1A_1 + \dots + x_nA_n$. Il suffit alors d'utiliser la n-linéarité et le caractère alterné à la matrice dont on a remplacé A_i par B. □

Remarque 12.19. Cette technique n'est pratique qu'à l'ordre 2 ou 3 en général.

Théorème 12.12 (Invariance du déterminant par transposition)**.** *Le déterminant est invariant par transposition :*

$$\det(A) = \det(A^\top)$$

Corollaire 12.6. $\det(A)$ *n'est pas seulement n-linéaire alterné par rapport aux colonnes de A, mais aussi par rapport aux lignes de A.*

2.4 Calcul pratique

Théorème 12.13 (Effet des opérations élémentaires sur le déterminant d'une matrice)**.** *Qu'elles soient sur les lignes ou sur les colonnes, les opérations élémentaires ont l'effet suivant sur le déterminant d'une matrice.*

1. *Transvection : Aucun effet*
2. *Dilatation d'un facteur λ : Multiplication du déterminant par λ*
3. *Échange : Multiplication du déterminant par -1*

Lemme 12.4 (Calcul d'un déterminant par pivot)**.** *Soit $n \geqslant 2$ et $A \in \mathcal{M}_n(\mathbb{K})$ une matrice de la forme*

$$A = \left(\begin{array}{c|ccc} \lambda & * & \dots & * \\ \hline 0 & & & \\ \vdots & & \tilde{A} & \\ 0 & & & \end{array}\right)$$

Alors on a $\det(A) = \lambda \det(A)$

Remarque 12.20. On pourra utiliser ce lemme pour calculer très efficacement un déterminant.

Corollaire 12.7 (Déterminant d'une matrice triangulaire)**.** *Le déterminant d'une matrice triangulaire est égal au produit de ses termes diagonaux.*

Corollaire 12.8 (Déterminant d'une matrice triangulaire)**.** *Le déterminant d'une matrice diagonale est égal au produit de ses termes diagonaux.*

Définition 12.14 (Mineur, cofacteur)**.** Soit $n \geqslant 2$ et $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$. On appelle **mineur** du terme $a_{i,j}$ le déterminant noté $\Delta_{i,j}$ obtenu en supprimant la i-ème ligne et la j-ème colonne (concrètement, on supprime la ligne et la colonne du terme $a_{i,j}$ et on calcule le déterminant de "ce qu'il reste"). Le scalaire $A_{i,j} = (-1)^{i+j}\Delta_{i,j}$ s'appelle le **cofacteur** du terme $a_{i,j}$.

Théorème 12.14 (Développement selon une ligne ou une colonne)**.** *Soit $n \geqslant 2$ et $A \in \mathcal{M}_n(\mathbb{K})$. En notant $A_{i,j}$ les cofacteurs de A, on a :*

1. $\forall i \in [\![1, n]\!],\ \displaystyle\sum_{j=1}^{n} a_{i,j}A_{i,j} = \det(A)$ *(développement selon la i-ème ligne)*
2. $\forall j \in [\![1, n]\!],\ \displaystyle\sum_{i=1}^{n} a_{i,j}A_{i,j} = \det(A)$ *(développement selon la j-ème colonne)*

Théorème 12.15 (Déterminant d'une matrice triangulaire par blocs)**.** *Soit A une matrice triangulaire par blocs. Alors le déterminant de A est égal au produit des déterminants des blocs diagonaux.*

Corollaire 12.9 (Déterminant d'une matrice diagonale par blocs)**.** *Si A est diagonale par blocs, son déterminant vaut le produit des déterminants des blocs diagonaux.*

Définition 12.15 (Comatrice)**.** Si $n \geqslant 2$, la matrice des cofacteurs de A s'appelle la **comatrice** de A. On la note $\mathrm{Com}(A)$.

Théorème 12.16 (Formule de la comatrice)**.** *On a la relation*

$$A\,\mathrm{Com}(A)^\top = \mathrm{Com}(A)^\top A = \det(A) I_n$$

Remarque 12.21. Si $n = 1$, les différentes conventions sur l'ensemble vide montrent que $\mathrm{Com}(A) = 1$ et la formule reste valable.

Remarque 12.22. Dans le cas où A est inversible, cette dernière formule peut éventuellement servir à calculer

$$A^{-1} = \frac{1}{\det(A)} \mathrm{Com}(A)^\top$$

mais elle a un intérêt plus théorique que pratique, car elle devient rapidement inexploitable. On lui préfèrera l'utilisation du pivot de Gauss. Toutefois, cette formule est **très utile** dans les exercices faisant appel à la comatrice ou à la densité de de $\mathcal{GL}_n(\mathbb{K})$ dans $\mathcal{M}_n(\mathbb{K})$.

Théorème 12.17. *(Déterminant de Vandermonde) Soit $a_1, \ldots, a_n$ des scalaires de $\mathbb{K}$. Alors*

$$\begin{vmatrix} 1 & a_1 & \ldots & a_1^{n-1} \\ 1 & a_2 & \ldots & a_2^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & a_n & \ldots & a_n^{n-1} \end{vmatrix} = \prod_{i<j} (a_j - a_i)$$

En particulier, le déterminant est non nul si, et seulement, les a_i sont deux à deux distincts.

Remarque 12.23. On rencontrera parfois ce déterminant sous sa forme transposée, ce qui ne change rien car le déterminant est invariant par transposition.

Chapitre 13

Polynômes

Dans tout le chapitre, on fixe $\mathbb{K}$ un corps.

1 Définition formelle

Définition 13.1 (Polynôme à une indéterminée sur le corps $\mathbb{K}$)**.** Un **polynôme à une une indéterminée sur le corps** $\mathbb{K}$ est une suite presque nulle $(a_k)_{k\in\mathbb{N}} \in \mathbb{K}^{(\mathbb{N})}$.

En particulier, deux polynômes sont égaux si, et seulement si, ils ont les mêmes coefficients.

Proposition 13.1 (Rappels : structure de $\mathbb{K}$-espace vectoriel et base canonique)**.** *On rappelle que* $\mathbb{K}^{(\mathbb{N})}$ *est un* $\mathbb{K}$*-espace vectoriel pour les lois suivantes :*

1. $(a_k)_{k\in\mathbb{N}} + (b_k)_{k\in\mathbb{N}} = (a_k + b_k)_{k\in\mathbb{N}}$
2. $\lambda.(a_k)_{k\in\mathbb{N}} = (\lambda \times a_k)_{k\in\mathbb{N}}$

Aussi, on pose

$$\forall k \in \mathbb{N},\ e_k = (\delta_{k,l})_{l\in\mathbb{N}} \in \mathbb{K}^{(\mathbb{N})}$$

La famille $(e_k)_{k\in\mathbb{N}}$ *est une base de* $\mathbb{K}^{(\mathbb{N})}$*, appelée* ***base canonique*** *de* $\mathbb{K}^{(\mathbb{N})}$*.*

Remarque 13.1. Soit $(a_k)_{k\in\mathbb{N}} \in \mathbb{K}^{\mathbb{N}}$. Les deux assertions suivantes sont équivalentes :

1. $(a_k)_{k\in\mathbb{N}} \in \mathbb{K}^{(\mathbb{N})}$
2. $\exists n \in \mathbb{N},\ \forall k > n,\ a_k = 0_{\mathbb{K}}$

Définition 13.2 (Produit de polynômes)**.** On munit $\mathbb{K}^{(\mathbb{N})}$ d'une loi $\times$ de la manière suivante. Soit $P = (a_k)_{k\in\mathbb{N}} \in \mathbb{K}^{(\mathbb{N})}$ et $Q = (b_k)_{k\in\mathbb{N}} \in \mathbb{K}^{(\mathbb{N})}$. On définit alors le polynôme $P \times Q$ comme le polynôme $P \times Q = (c_k)_{k\in\mathbb{N}} \in \mathbb{K}^{(\mathbb{N})}$ où on a posé

$$\forall k \in \mathbb{N},\ c_k = \sum_{i+j=k} a_i b_j = \sum_{i=0}^{k} a_i b_{k-i} = \sum_{j=0}^{k} a_{k-j} b_j$$

Exemple 13.1. Le terme constant d'un produit de polynômes est égal au produit des termes constants.

Proposition 13.2 (Propriétés de $\times$)**.** *Sur $\mathbb{K}^{(\mathbb{N})}$, la loi $\times$ définie comme précédemment est :*

1. ***associative***
2. ***commutative***
3. ***bilinéaire** ; c'est-à-dire linéaire par rapport au premier facteur et linéaire par rapport au deuxième facteur*

Exemple 13.2. En particulier, $\times$ est ainsi compatible avec $\cdot$ et distributive sur $+$.

Corollaire 13.1 (Structure de $\mathbb{K}$-algèbre)**.** $(\mathbb{K}^{(\mathbb{N})}, +, \times, \cdot)$ *est une $\mathbb{K}$-algèbre commutative, d'élément neutre pour $\times$ le polynôme $e_0 = (\delta_{j,0})_{j\in\mathbb{N}}$*

Définition 13.3 (Indéterminée)**.** On pose

$$X = (\delta_{k,1})_{k\in\mathbb{N}}$$

On appelle X l'**indéterminée** (même si X est parfaitement déterminé et est une constante dans l'ensemble des polynômes, ie ce polynôme ne change pas, mais ce n'est pas non plus un polynôme constant).

Lemme 13.1. *On a*

$$\forall k \in \mathbb{N},\ X^k = (\delta_{j,k})_{j\in\mathbb{N}}$$

Définition 13.4 (Notation)**.** Dans ce contexte, on notera plutôt $\mathbb{K}^{(\mathbb{N})}$ de la façon suivante :

$$\mathbb{K}[X]$$

qu'on lira "$\mathbb{K}$ crochet X" ou plus exactement "l'ensemble des polynômes à une indéterminée sur le corps $\mathbb{K}$".

Corollaire 13.2. *Ainsi, $(X^k)_{k\in\mathbb{N}}$ est donc la **base canonique** de $\mathbb{K}[X]$ et tout polynôme $P = (a_k)_{k\in\mathbb{N}} \in \mathbb{K}[X]$ se décompose sur cette base sous la forme*

$$P = \sum_{k\in\mathbb{N}} a_k X^k$$

Définition 13.5 (Polynôme constant)**.** Soit $\lambda \in \mathbb{K}$. le **polynôme constant égal à** λ est le polynôme $\lambda.X^0$. C'est donc la suite $(\lambda, 0, 0, \ldots)$. En pratique, on pourra l'écrire λ car l'application

$$\begin{array}{ccc} \mathbb{K} & \longrightarrow & \mathbb{K}[X] \\ \lambda & \longmapsto & \lambda.X^0 \end{array}$$

est un morphisme injectif de $\mathbb{K}$-algèbres. On identifiera alors $\mathbb{K}$ à l'ensemble des polynômes constants.

Définition 13.6 (Degré)**.** Soit $P = \sum_{k\in\mathbb{N}} a_k X^k \in \mathbb{K}[X]$. On définit son **degré** de la manière suivante :

1. Si $P = 0_{\mathbb{K}[X]}$, on pose $\deg(P) = -\infty$
2. Si $P \neq 0_{\mathbb{K}[X]}$, alors le support de la famille $(a_k)_{k\in\mathbb{N}}$ est non vide et fini par hypothèse, donc on pose alors $\deg(P)$ comme étant le maximum de ce support. C'est donc alors le plus grand $n \in \mathbb{N}$ tel que $a_n = 0_{\mathbb{K}}$.

Définition 13.7 (Coefficient dominant, terme dominant)**.** Lorsque $P = \sum_{k\in\mathbb{N}} a_k X^k \in \mathbb{K}[X]$ est non nul de degré $\deg(P) = n \in \mathbb{N}$, on appelle **coefficient dominant de** P le scalaire a_n, et on appelle **terme dominant de** P le polynôme $a_n X^n$.

Définition 13.8 (Polynôme unitaire)**.** Si $P = \sum_{k\in\mathbb{N}} a_k X^k \in \mathbb{K}[X]$ est non nul de degré $\deg(P) = n \in \mathbb{N}$, on dit que P est **unitaire** lorsque son coefficient dominant vaut $1_k\mathbb{K}$, c'est-à-dire lorsque $a_n = 1_{\mathbb{K}}$.

Exemple 13.3. Pour tout $n \in \mathbb{N}$, X^n est unitaire de degré n.

Définition 13.9. Soit $n \in \mathbb{N}$. On pose :

$$\mathbb{K}_n[X] = \{P \in \mathbb{K}[X] \mid \deg(P) \leqslant n\}$$

Remarque 13.2. On a

$$\bigcup_{n\in\mathbb{N}} \mathbb{K}_n[X] = \mathbb{K}[X]$$

et

$$\forall n \in \mathbb{N},\ \mathbb{K}_n[X] \subset \mathbb{K}_{n+1}[X]$$

Proposition 13.3 (Propriété utile)**.** *Soit $n \in \mathbb{N}$ et $P \in \mathbb{K}_[X]$. On a :*

1. $\deg(P) \leqslant n \iff \exists (a_0, \dots, a_n) \in \mathbb{K}^{n+1},\ P = \sum_{k=0}^{n} a_k X^k$
2. $\deg(P) = n \iff \exists (a_0, \dots, a_n) \in \mathbb{K}^{n} \times \mathbb{K},\ P = \sum_{k=0}^{n} a_k X^k$

Proposition 13.4. *Soit $n \in \mathbb{N}$. On a les propriétés suivantes pour $\mathbb{K}_n[X]$:*

1. *$\mathbb{K}_n[X]$ est un sous-espace vectoriel de $\mathbb{K}[X]$*
2. *$\mathbb{K}_n[X]$ est de dimension finie $n+1$*
3. *$(X^0, \dots, X^n)$ est une base de $\mathbb{K}_n[X]$, appelée **base canonique** de $\mathbb{K}_n[X]$*

Théorème 13.1 (Degré d'un produit et d'une somme)**.** *Soit $P, Q \in \mathbb{K}[X]$. On a les propriétés suivantes sur le degré de $P+Q$:*

1. $\deg(P+Q) \leqslant \max(\deg(P), \deg(Q))$
2. *Si $\deg(P) \neq \deg(Q)$, alors $\deg(P+Q) = \max(\deg(P), \deg(Q))$ et le coefficient dominant de $P+Q$ est égal au coefficient du polynôme de plus haut degré.*

On a les propriétés suivantes sur le degré de PQ :

1. $\deg(PQ) = \deg(P) + \deg(Q)$
2. *Si P et Q sont tous deux non nuls, alors PQ est non nul et le coefficient dominant de PQ est égal au produit des coefficients dominant de P et de Q.*

Remarque 13.3. On peut généraliser ces propriétés à un nombre fini de termes ou de facteurs en effectuant une récurrence.

Corollaire 13.3 (Intégrité de l'anneau des polynômes)**.** $(\mathbb{K}[X],+,\times)$ *est un anneau intègre.*

Proposition 13.5 (HP)**.** *Soit* $P\in\mathbb{K}[X]$ *et* $n\in\mathbb{N}$*. On a*

$$\deg(P)=n \iff \exists(\lambda,Q)\in\mathbb{K}^\star\times\mathbb{K}[X],\ P=\lambda X^n+Q \quad et \quad \deg(Q)<n$$

Démonstration. Pour $\Rightarrow$, utiliser la "propriété utile" et distinguer les cas $n=0$ et $n\geqslant$. Pour $\Leftarrow$, fixer les objets dont on suppose l'existence et utiliser les propriétés du degré. □

Corollaire 13.4 (Renormalisé)**.** *Supposons que* P *est un polynôme non nul. En notant* $CD(P)$ *le coefficient dominant de* P*, on a le fait que*

$$\frac{1}{CD(P)}.P$$

est unitaire. On dit qu'on a ***renormalisé*** P*.*

Théorème 13.2 (Groupe des unités de l'anneau des polynômes)**.** *L'ensemble des unités de l'anneau* $(\mathbb{K}[X],+,\times)$ *vaut :*

$$\mathcal{U}(\mathbb{K}[X])=\{P\in\mathbb{K}[X]\mid \deg(P)=0\}=\mathbb{K}_0[X]\setminus\left\{0_{\mathbb{K}[X]}\right\}$$

Remarque 13.4 (Importante pour X/ENS)**.** Si $(A,+,\times)$ est un anneau **intègre**, tout le début du chapitre fonctionne de même, à l'exception du théorème précédent. Dans ce cas, les unités de $A[X]$ sont exactement les polynômes constants égaux aux unités de l'anneau A.

Définition 13.10 (Polynôme composé)**.** Soit $P,Q\in\mathbb{K}[X]$. Le **polynôme composé** $P\circ Q$ est défini de la manière suivante : on écrit P sous la forme $\sum_{k\in\mathbb{N}} a_kX^k$ puis on pose

$$P\circ Q=\sum_{k\in\mathbb{N}} a_kQ^k$$

Proposition 13.6 (Degré de $P\circ Q$ avec Q non constant, HP)**.** *Si* Q *est un polynôme non constant (c'est-à-dire que* $\deg(Q)\geqslant 1$*), alors on a :*

$$\deg(P\circ Q)=\deg(P)\times\deg(Q)$$

Démonstration. Si P est nul, on conclut immédiatement. Si P est non nul,on pose $n=\deg(P)$ et on écrit

$$P\circ Q=a_nQ^n+\sum_{k=0}^{n-1}a_kQ^k$$

Or, on a $\deg(a_nQ^n)=n\deg(Q)$ et si $n=0$, le deuxième polynôme est nul, et si n est non nul, le deuxième polynôme est de degré inférieur à $\max(\deg(Q^0),\dots,\deg(Q^{n-1}))$ donc de degré strictement inférieur à $n\deg(Q)$. On conclut alors par les propriétés sur le degré d'une somme. □

2 Arithmétique dans $\mathbb{K}[X]$

Définition 13.11 (Divisibilité (rappel))**.** Soit $A, B \in \mathbb{K}[X]$. On dit que B **divise** A, et on note $B \mid A$ lorsque

$$\exists P \in \mathbb{K}[X],\ A = BP$$

c'est-à-dire lorsque $A \in B\mathbb{K}[X]$ ou encore lorsque $A\mathbb{K}[X] \subset B\mathbb{K}[X]$.

Proposition 13.7 ($\mid$ est un préordre)**.** *La relation $\mid$ est réflexive et transitive, mais pas antisymétrique. On dit que c'est un* ***préordre****.*

Proposition 13.8. *(Propriétés de $\mid$) La relation $\mid$ vérifie les propriétés suivantes :*

1. *Soit D un diviseur de A et B. Alors*
$$\forall (U, V) \in \mathbb{K}[X]^2,\ D \mid AU + BV$$
2. *Soit $Q \in \mathbb{K}[X]$. Si $D \mid A$, alors $QD \mid QA$. Réciproquement, si $QD \mid QA$ **et** si $Q \neq 0$, alors $D \mid A$.*
3. *Si $D \mid A$ et $A \neq 0$, alors* $\deg(D) \leqslant \deg(A)$

Définition 13.12 (Association)**.** Soit $A, B \in \mathbb{K}[X]$. Les deux assertion suivantes sont équivalentes :

1. $A \mid B$ et $B \mid A$
2. $\exists \lambda \in \mathbb{K}^\star,\ B = \lambda A$

Dans ce cas, A et B sont dits **associés**. La relation d'**association** est une relation d'équivalence. Nous la noterons ici $\sim$ de façon non standard.

Théorème 13.3. *Chaque classe d'équivalence pour la relation d'association admet un unique représentant unitaire ou nul (qu'on abrègera en UON, non standard).*

Proposition 13.9. *Deux polynômes associés ont le même degré. En revanche la réciproque est fausse, on pourra considérer X et $X + 1$.*

Exemple 13.4. $A \sim B \iff A\mathbb{K}[X] = B\mathbb{K}[X]$ (utiliser une des définitions alternatives de la divisibilité)

Remarque 13.5 (Transformation d'un préordre en ordre, HP)**.** On rappelle que la divisibilité n'est pas antisymétrique. Toutefois, on peut la rendre antisymétrique en créant une divisibilité induite sur les classes d'association.

De façon générale, si on dispose d'un préordre $\mathcal{R}$ sur un ensemble E, on peut "transformer" ce préordre en ordre en quotientant E par le relation définie par $x \sim y \iff (x\mathcal{R}y \wedge y\mathcal{R}x)$ puis en faisant passer le préordre au quotient. Il s'agira alors toujours d'une relation d'ordre (cf. exercice étoilé d'introduction à l'algèbre).

Théorème 13.4 (Division euclidienne)**.** *Soit $(A, B) \in \mathbb{K}[X] \times (\mathbb{K}[X] \setminus \{0_{\mathbb{K}[X]}\})$. Alors :*

$$\exists !(Q, R) \in \mathbb{K}[X]^2,\ A = BQ + R \ \text{ et } \ \deg(R) < \deg(B)$$

Théorème 13.5 (Classification des idéaux de l'anneau des polynômes)**.** *Soit $\mathcal{I}$ un idéal de $\mathbb{K}[X]$. Alors, il existe un unique polynôme $P \in \mathbb{K}[X]$ unitaire ou nul tel que*

$$\mathcal{I} = P\mathbb{K}[X]$$

Corollaire 13.5 (Principalité de l'anneau des polynômes)**.** *L'anneau* $(\mathbb{K}[X],+,\times)$ *est principal.*

Remarque 13.6 (HP)**.** On aurait déjà pu obtenir la principalité de $\mathbb{K}[X]$ puisque $\mathbb{K}[X]$ dispose d'une division euclidienne. En effet, tout anneau euclidien est principal (voir compléments sur les anneaux).

Corollaire 13.6. *Les résultats arithmétiques usuels s'appliquent dans* $\mathbb{K}[X]$ *:*

1. *Soit* $A,B\in\mathbb{K}[X]$*. Les unités de* $\mathbb{K}[X]$ *(ie les* $\lambda\in\mathbb{K}^\star$*) divisent toujours* A *et* B*. Réciproquement, si ce sont les seuls diviseurs communs de* A *et* B*,* A *et* B *sont dits* ***premiers entre eux****.*
2. *On définit de même les polynômes* $A_1,\dots,A_n$ ***premiers entre eux dans leur ensemble****.*
3. ***Théorème de Bézout*** *: Soit* $(A,B)\in\mathbb{K}[X]^2$*.* A *et* B *sont premiers ente eux si, et seulement si,* $\exists(U,V)\in\mathbb{K}[X]^2,\ AU+BV=1$
4. ***Généralisation du théorème de Bézout*** *: Soit* $n\geqslant 2$ *et* $(A_1,\dots,A_n)\in\mathbb{K}[X]^n$*.* $A_1,\dots,A_n$ *sont premiers entre eux dans leur ensemble si, et seulement si,*
$$\exists(U_1,\dots,U_n)\in\mathbb{K}[X]^n,\ \sum_{i=1}^{n}A_iU_i=1$$
5. ***Corollaire du théorème de Bézout*** *: Soit* $n\geqslant 2$ *et* $(A,B_1,\dots,B_n)\in\mathbb{K}[X]^{n+1}$*. Si* A *est premier avec chacun des* B_i*, alors* A *est premier avec leur produit. On peut par exemple appliquer deux fois ce résultat si des* $A_1,\dots,A_m$ *sont tous premiers avec chacun des* B_i*, alors le produit des* A_i *est premier avec le produit des* B_i*.*
6. ***Lemme de Gauss*** *: Soit* $(A,B,C)\in\mathbb{K}[X]^3$*. Si* $A\mid BC$ *et* A *est premier avec* B*, alors* $A\mid C$*.*
7. ***Corollaire du lemme de Gauss*** *: Soit* $n\geqslant 2$ *et* $(A_1,\dots,A_n,B)\in\mathbb{K}[X]^{n+1}$*. Si les* A_i *sont deux à deux premiers entre eux et si* $\forall i\in[\![1,n]\!],\ A_i\mid B$*, alors*
$$A_1\times\ldots\times A_n\mid B$$

Définition 13.13 (**Un** PGCD, **un** PPCM)**.** Soit A et B des polynômes tous deux non nuls.

1. On appelle **un PGCD** de A et B tout diviseur commun non nul de A et B de degré maximal.
2. On appelle **un PPCM** de A et B tout multiple commun non nul de A et B de degré minimal.

Théorème 13.6. *Soit* A *et* B *des polynômes tous deux non nuls.*

1. *Les PGCD de* A *et* B *forment une classe d'association.*
2. *Les PPCM de* A *et* B *forment une classe d'association.*

Définition 13.14 (**Le** PGCD, **le** PPCM)**.** Soit A et B des polynômes tous deux non nuls.

1. La classe d'association des PGCD de A et B admet un unique représentant UON. On l'appelle **le PGCD** de A et B, et on le note $A\wedge B$. Il vérifie :
$$A\mathbb{K}[X]+B\mathbb{K}[X]=(A\wedge B)\mathbb{K}[X]$$
2. La classe d'association des PPCM de A et B admet un unique représentant UON. On l'appelle **le PPCM** de A et B, et on le note $A\vee B$. Il vérifie :
$$A\mathbb{K}[X]\cap B\mathbb{K}[X]=(A\vee B)\mathbb{K}[X]$$

Définition 13.15 (Extension du PGCD et du PPCM)**.** Soit A et B des polynômes quelconques (éventuellement nuls).

1. $A\mathbb{K}[X] + B\mathbb{K}[X]$ est un idéal de $\mathbb{K}[X]$ dont
$$\exists! P \text{ UON},\ A\mathbb{K}[X] + B\mathbb{K}[X] = P\mathbb{K}[X]$$
On pose alors $A \wedge B = P$.
2. $A\mathbb{K}[X] \cap B\mathbb{K}[X]$ est un idéal de $\mathbb{K}[X]$ dont
$$\exists! P \text{ UON},\ A\mathbb{K}[X] \cap B\mathbb{K}[X] = P\mathbb{K}[X]$$
On pose alors $A \vee B = P$.

Cette définition prolonge celle donnée précédemment, et on conserve les appellations "le PGCD" et "le PPCM".

Proposition 13.10 (Théorème de Bézout généralisé)**.** *Soit A et B deux polynômes quelconques. Alors*
$$\exists (U, V) \in \mathbb{K}[X]^2,\ AU + BV = A \wedge B$$

Remarque 13.7. L'ensemble des diviseurs communs de A et B est égal à l'ensemble des diviseurs de $A \wedge B$. L'ensemble des multiples communs de A et B est égal à l'ensemble des multiples de $A \vee B$.

Remarque 13.8. Dans l'ensemble des polynômes UON, la divisibilité induit une relation d'ordre : en effet, deux polynômes UON associés sont égaux. Alors, $A \wedge B$ est le maximum pour la divisibilité induite des polynômes UON divisant A et B, et $A \vee B$ est le minimum pour la divisibilité induite des polynômes UON multiples de A et de B.

Définition 13.16 (PGCD et PPCM de plusieurs polynômes)**.** Soit $n \geqslant 2$ et $A_1, \ldots, A_n$ des polynômes quelconques (éventuellement nuls).

1. $A_1\mathbb{K}[X] + \ldots + A_n\mathbb{K}[X]$ est un idéal de $\mathbb{K}[X]$ dont
$$\exists! P \text{ UON},\ A_1\mathbb{K}[X] + \ldots + A_n\mathbb{K}[X] = P\mathbb{K}[X]$$
On pose alors $A_1 \wedge \ldots \wedge A_n = P$. On dit que P est le PGCD de $A_1, \ldots, A_n$.
2. $A_1\mathbb{K}[X] \cap \ldots \cap A_n\mathbb{K}[X]$ est un idéal de $\mathbb{K}[X]$ dont
$$\exists! P \text{ UON},\ A_1\mathbb{K}[X] \cap \ldots \cap A_n\mathbb{K}[X] = P\mathbb{K}[X]$$
On pose alors $A_1 \vee \ldots \vee A_n = P$. On dit que P est le PPCM de $A_1, \ldots, A_n$.

Proposition 13.11 (Associativité et commutativité du PGCD et du PPCM)**.** *Le PGCD et le PPCM sont associatifs et commutatifs.*

Proposition 13.12 (invariance par association du PGCD et du PPCM)**.** *Le PGCD et le PPCM sont invariants si on remplace A ou B par un de leurs associés.*

Proposition 13.13 (Distributivité du PGCD)**.** *Le PGCD est distributif :*
$$\forall (A, B) \in \mathbb{K}[X]^2,\ \forall P \text{ UON},\ (PA) \wedge (PB) = P(A \wedge B)$$

Proposition 13.14. *A et B sont premiers entre eux si, et seulement si, $A \wedge B = 1$. $A_1, \ldots, A_n$ sont premiers entre eux dans leur ensemble si, et seulement si, $A_1 \wedge \ldots \wedge A_n = 1$*

Proposition 13.15 (Propriété utile)**.** *Soit $(A, B) \in \mathbb{K}[X]^2$. On note $\Delta = A \wedge B$? Alors, il existe $(\tilde{A}, \tilde{B}) \in \mathbb{K}[X]^2$ tel que $A = \Delta\tilde{A}$, $B = \Delta\tilde{B}$ et $\tilde{A} \wedge \tilde{B} = 1$.*

Théorème 13.7 (Invariance par transvection)**.** *Soit $(A, B) \in \mathbb{K}[X]^2$. On a*

$$\forall Q \in \mathbb{K}[X],\ A \wedge B = (A - BQ) \wedge B$$

En conséquence, l'algorithme d'Euclide est valable et peut être appliqué pour déterminer un PGCD.

Définition 13.17 (Congruence modulo P)**.** Soit $P \in \mathbb{K}[X]$ non nul. On dit que A est **congru à** B **modulo** P, et on note $A \equiv B\ [P]$ lorsque $P \mid A - B$. Les propriétés usuelles restent valables :

1. $\cdot \equiv \cdot\ [P]$ est une relation d'équivalence
2. Si $A_1 \equiv B_1\ [P]$ et $A_2 \equiv B_2\ [P]$ alors $A_1 + A_2 \equiv B_1 + B_2\ [P]$ et $A_1A_2 \equiv B_1B_2\ [P]$.
3. Si $A \equiv B\ [P]$, alors $\forall n \in \mathbb{N},\ A^n \equiv B^n\ [P]$

3 Lien avec les fonctions

3.1 Évaluation d'un polynôme

Définition 13.18 (Évaluation en un point de $\mathbb{K}$)**.** Soit $P \in \mathbb{K}[X]$ et $x \in \mathbb{K}$. Pour évaluer P en x, on écrit P sous la forme $P = \sum_{k\in\mathbb{N}} a_k X^k$ puis on pose :

$$P(x) = \sum_{k\in\mathbb{N}} a_k x^k$$

On dit aussi qu'on a **substitué** x **à** X **dans** P.

Théorème 13.8 (Évaluation et opérations algébriques)**.** *Soit $(\lambda, \mu) \in \mathbb{K}^2$ et $(P, Q) \in \mathbb{K}[X]^2$. Soit $x \in \mathbb{K}$. On a :*

1. $(\lambda P + \mu Q)(x) = \lambda P(x) + \mu Q(x)$
2. $(PQ)(x) = P(x)Q(x)$
3. $(P \circ Q)(x) = P(Q(x))$

Remarque 13.9. En fait, on a montré que $\varphi_x : \begin{array}{ccc} \mathbb{K}[X] & \longrightarrow & \mathbb{K} \\ A & \longmapsto & A(x) \end{array}$ est un morphisme de $\mathbb{K}$-algèbres. Et on peut généraliser, ce qui fait l'objet du théorème suivant.

Théorème 13.9 (Évaluation en un point d'une $\mathbb{K}$-algèbre, HP, spé)**.** *Soit $\mathcal{A}$ une $\mathbb{K}$-algèbre quelconque et $P \in \mathbb{K}[X]$ qu'on écrit sous la forme $P = \sum_{k\in\mathbb{N}} a_k X^k$. Alors on peut encore poser*

$$\forall x \in \mathcal{A},\ P(x) = \sum_{k\in\mathbb{N}} a_k X^k$$

Alors, pour tout $x \in \mathcal{A}$, l'application

$$\begin{array}{rccc} \varphi_x : & \mathbb{K}[X] & \longrightarrow & \mathcal{A} \\ & Q & \longmapsto & Q(x) \end{array}$$

est un morphisme de $\mathbb{K}$-algèbres et le théorème qui précède reste valable.

On pensera notamment aux $\mathbb{K}$-algèbres suivantes : $\mathcal{L}(E)$ où E est un $\mathbb{K}$-espace vectoriel, $\mathcal{M}_n(\mathbb{K})$ ou encore $\mathbb{K}[X]$.

Théorème 13.10 (Lemme des noyaux, HP, vu en spé)**.** *Pour tout $u \in \mathcal{L}(E)$, pour tous polynômes P et Q premiers entre eux, on a*

$$\ker((PQ)(u)) = \ker(P(u)) \oplus \ker(Q(u))$$

Démonstration. Utiliser le théorème de Bézout et évaluer en u. Si on note $PA + QB = 1$ cette relation de Bézout, montrer ensuite que $\mathrm{Im}(P(u) \circ A(u)) \subset \ker(Q(u))$ et $\mathrm{Im}(Q(u) \circ B(u)) \subset \ker(P(u))$ en se plaçant dans $\ker((PQ)(u))$. La somme est alors totale en évaluant en x le Bézout évalué en u. On montre que la somme est directe en remarquant que $P(u)$ et $A(u)$ commutent, de même que $Q(u)$ et $B(u)$, car ce sont des polynômes en u. □

3.2 Racines

Définition 13.19 (Racine)**.** Soit $P \in \mathbb{K}[X]$. une **racine** de P (ou un **zéro** de P) est un scalaire $x \in \mathbb{K}$ tel que $P(x) = 0_{\mathbb{K}}$. On notera ici de façon non standard $\mathcal{R}(P)$ l'ensemble des racines d'un polynôme P.

Remarque 13.10. Si on cherche les racines communes de A et B avec B non nul, alors elles sont aussi des racines du reste dans la division euclidienne de A par B (écrire cette division et l'évaluer en cette racine). Pn peut alors itérer pour se ramener à des polynômes de faible degré.

Théorème 13.11. *Soit $x \in \mathbb{K}$ et $P \in \mathbb{K}[X]$. On a*

$$x \in \mathcal{R}(P) \iff X - x \mid P$$

Remarque 13.11 (Très utile pour les exercices X/ENS)**.** En utilisant la formule de Bernoulli, on obtient le résultat suivant :

$$\forall P \in \mathbb{Z}[X],\ \forall (a,b) \in \mathbb{Z}^2,\ a - b \mid P(a) - P(b)$$

Remarque 13.12. On reprend l'exemple des racines communes de A et B avec B non nul.On a $X - x \mid A$ et $X - x \mid B$ si, et seulement si, $X - x \mid A \wedge B$, donc les racines communes de A et B sont exactement les racines de $A \wedge B$.

Proposition 13.16. *Soit $P \in \mathbb{K}[X]$ et $x_1, \ldots, x_n$ des scalaires de $\mathbb{K}$ **deux à deux distincts**. Les deux assertions suivantes sont équivalentes :*

1. $\forall i \in [\![1, n]\!],\ P(x_i) = 0_{\mathbb{K}}$
2. $(X - x_1) \times \ldots \times (X - x_n) \mid P$

Méthode 13.1 (Factorisation d'un polynôme aux racines distinctes)**.** Voici comment on peut factoriser un polynôme P dont les racines sont deux à deux distinctes.

1. On cherche les racines.
2. On montre que celles-ci sont deux à deux distinctes (si c'est le cas évidemment, sinon on ne peut pas appliquer la méthode) A ce moment, le produit des $X-$racines divise P. On fixe alors le polynôme Q correspondant à cette divisibilité.
3. On examine le degré de Q : s'il y a autant de racines distinctes que le degré de P, Q est une constante.
4. On examine le coefficient dominant de P : il permet de déterminer Q si Q est une constante.

Exemple 13.5. On a par exemple la factorisation suivante dans $\mathbb{C}[X]$:

$$\forall n \in \mathbb{N}^\star,\ X^n - 1 = \prod_{i=0}^{n-1} \left(X - \exp\left(i\frac{2k\pi}{n} \right) \right)$$

dont on en déduit l'identité :

$$\forall n \in \mathbb{N}^\star,\ \forall (a,b) \in \mathbb{C}^2,\ a^n - b^n = \prod_{i=0}^{n-1} \left(a - \exp\left(i\frac{2k\pi}{n} \right) b \right)$$

Corollaire 13.7. *Soit $n \in \mathbb{N}$ et $P \in \mathbb{K}_n[X]$. Si P admet au moins $n+1$ racines 2 à 2 distinctes, alors $P = 0$.*

Corollaire 13.8. *On en déduit les deux résultats qui suivent :*

1. *Soit $n \in \mathbb{N}$ et $(P,Q)_i n \mathbb{K}_n[X]^2$. Si P et Q coïncident (au moins) en $n+1$ points deux à deux distincts, alors $P = Q$.*
2. *Soit $(P,Q) \in \mathbb{K}[X]$. Si P et Q coïncident en une infinité de points, alors $P = Q$.*

Exemple 13.6. Soit $n \in \mathbb{N}^\star$ et $(A,B) \in \mathbb{C}[X]^2$. Alors on a

$$A^n - B^n = \prod_{i=0}^{n-1} \left(A - \exp\left(i\frac{2k\pi}{n} \right) B \right)$$

Exemple 13.7. Soit $P \in \mathbb{C}[X]$. Si $\forall x \in \mathbb{R},\ P(x) \in \mathbb{R}$, alors $P \in \mathbb{R}[X]$. En effet, P et son "polynôme conjugué" coïncident sur $\mathbb{R}$ qui est une partie infinie de $\mathbb{C}$. Donc tous les coefficients de P sont égaux à leur conjugué, donc sont tous réels. Donc $P \in \mathbb{R}[X]$.

Définition 13.20 (Ordre de multiplicité)**.** Soit $P \in \mathbb{K}[X]$ un polynôme **non nul** et $x \in \mathbb{K}$. On appelle **ordre de multiplicité de x en tant que racine de P** l'entier noté de façon non standard $m_P(x)$ défini par :

$$m_P(x) = \max\left\{ k \in \mathbb{N} : (X-x)^k \mid P \right\}$$

Remarque 13.13. On a

$$x \in \mathcal{R}(P) \iff m_P(x) \geqslant 1$$

De même, on a

$$x \notin \mathcal{R}(P) \iff m_P(x) = 0$$

On peut alors parler de "racine d'ordre 0".

Définition 13.21 (Racine simple, double, triple)**.** Si $m_P(x) = 1$, on parle de racine **simple**. Si $m_P(x) = 2$, on parle de racine **double**. Si $m_P(x) = 3$, on parle de racine **triple**.

Remarque 13.14. Comme en arithmétique, on montre que

$$\forall k \in \mathbb{N},\ (X-x)^k \mid P \iff k \leqslant m_P(x)$$

Corollaire 13.9. *Soit P de degré $n \in \mathbb{N}$. Alors P admet au plus n racines comptées avec leur multiplicité. Formellement, on a :*

$$\sum_{x \in \mathcal{R}(P)} m_P(x) \leqslant n$$

3.3 Fonctions polynomiales

Définition 13.22 (Fonction polynomiale associée à un polynôme)**.** Soit $P \in \mathbb{K}[X]$. La **fonction polynomiale associée à** P est la fonction :

$$\begin{array}{rccc} \tilde{P}: & \mathbb{K} & \longrightarrow & \mathbb{K} \\ & x & \longmapsto & P(x) \end{array}$$

Remarque 13.15. Attention, on ne peut identifier un polynôme et sa fonction polynômiale associée **que** si le corps de base $\mathbb{K}$ est infini. Dès qu'il est fini, il suffit de considérer le polynôme $\prod_{x \in \mathbb{K}} (X-x)$ qui est non nul en tant que polynôme, mais dont la fonction polynômiale associée est nulle.

3.4 Dérivation formelle

Définition 13.23 (Dérivée formelle)**.** Soit $p \in \mathbb{K}[X]$. Écrivons-le sous la forme $P = \sum_{k \in \mathbb{N}} a_k X^k$. On définit alors la **dérivée formelle de** P par :

$$P' = \sum_{k \in \mathbb{N}^\star} k a_k X^{k-1}$$

Remarque 13.16 (HP, caractéristique du corps et dérivation)**.** Supposons que $k \neq 0_{\mathbb{Z}}$ et $a_k \neq 0_{\mathbb{K}}$. On n'a pas nécessairement $k.a_k \neq 0_{\mathbb{K}}$. En effet, si $\mathbb{K}$ est de caractéristique k, rien ne va plus. Ainsi, **on suppose dans ce paragraphe que $\mathbb{K}$ est de caractéristique nulle**.

Remarque 13.17. Insistons bien sur le fait qu'il s'agit d'une dérivation **formelle**. Cela dit, si $\mathbb{K} = \mathbb{R}$, on retrouve la dérivation usuelle des fonctions polynomiales.

Exemple 13.8. $(X^0)' = 0$ et $\forall n \in \mathbb{N}^\star, (X^n)' = nX^{n-1}$

Proposition 13.17 (Combinaison linéaire)**.** *Soit $(\lambda, \mu) \in \mathbb{K}^2$ et $(P, Q) \in \mathbb{K}[X]^2$. On a :*

$$(\lambda P + \mu Q)' = \lambda P' + \mu Q'$$

Exemple 13.9. Ainsi, $\begin{array}{rccc} D: & \mathbb{K}[X] & \longrightarrow & \mathbb{K}[X] \\ & P & \longmapsto & P' \end{array}$ est un endomorphisme du $\mathbb{K}$-espace vectoriel $\mathbb{K}[X]$. A la réflexion, on aurait en fait pu le caractériser comme l'unique endomorphisme D de $\mathbb{K}[X]$ vérifiant $D(X^0) = 0_{\mathbb{K}[X]}$ et $\forall n \in \mathbb{N}^\star,\ D(X^n) = nX^{n-1}$

Exemple 13.10. Si on souhaite dériver $P = \sum_{k=0}^{n} a_k X^k$, la définition n'est en théorie pas applicable. Mais en utilisant la linéarité de D, on obtient tout de même :

$$P' = \sum_{k=1}^{n} a_k k X^{k-1}$$

Proposition 13.18 (Produit)**.** *Soit* $(P, Q) \in \mathbb{K}[X]^2$*. On a :*

$$(P \times Q)' = P' \times Q + P \times Q'$$

Proposition 13.19 (Composée)**.** *Soit* $(P, Q) \in \mathbb{K}[X]^2$*. On a :*

$$(P \circ Q)' = (P' \circ Q) \times Q'$$

Proposition 13.20. *Soit* $P \in \mathbb{K}[X]$ *non constant. Alors :*

$$\deg(P') = \deg(P) - 1$$

Remarque 13.18. Si P est constant, on a $\deg(P') = -\infty$. Ainsi, dans tous les cas, on a :

$$\deg(P') \leqslant \deg(P) - 1$$

On peut aller plus loin : P est constant si, et seulement si, $P' = 0$

Définition 13.24 (n-ème dérivée formelle)**.** Soit $P \in \mathbb{K}[X]$. Sa n-ème dérivée formelle est définie par récurrence : on pose $P^{(0)} = P$ et $\forall n \in \mathbb{N},\ P^{(n+1)} = (P^{(n)})'$

Exemple 13.11. Soit $d \in \mathbb{N}$. On a

$$\forall n \in [\![0, d]\!],\ (X^d)^{(n)} = \frac{d!}{(d-n)!} X^{d-n}$$

et

$$\forall n > d,\ (X^d)^{(n)} = 0$$

Exemple 13.12. Soit $P \in \mathbb{K}[X]$. On a

$$\forall (m, n) \in \mathbb{N}^2,\ (P^{(m)})^{(n)} = P^{(m+n)}$$

Exemple 13.13. La famille $\left(\frac{X^0}{0!}, \ldots, \frac{X^d}{d!}\right)$ est une base de $\mathbb{K}_d[X]$. Dans cette base, la matrice de l'endomorphisme dérivation est une matrice comportant des 0 partout sauf sur la surdiagonale qui ne comporte que des 1.

Proposition 13.21 (généralisation de la propriété précédente)**.** *Soit* $P \in \mathbb{K}[X]$ *et* $n \in \mathbb{N}$*. Si* $\deg(P) \geqslant n$*, alors* $\deg(P^{(n)}) = \deg(P) - n$

Remarque 13.19. Si $\deg(P) < n$ alors on vérifie que $P^{(n)} = 0$. Ainsi, dans tous les cas, on a toujours :

$$\deg(P^{(n)}) \leqslant \deg(P) - n$$

Remarque 13.20. On a

$$\forall n \in \mathbb{N},\ \forall P \in \mathbb{K}[X],\ \deg(P) \leqslant n \iff P^{(n+1)} = 0$$

On en déduit que $K_n[X]$ est un sev en tant que noyau de D^{n+1}.

Exercice 13.1. Dans le cas où $\deg(P) \geqslant n$, calculer $CD(P^{(n)})$ en fonction de $CD(P)$ et de $\deg(P)$

Proposition 13.22 (Combinaison linéaire)**.** *Soit $n \in \mathbb{N}$, $(\lambda, \mu) \in \mathbb{K}^2$ et $(P, Q) \in \mathbb{K}[X]^2$. On a :*

$$(\lambda P + \mu Q)^{(n)} = \lambda P(n) + \mu Q^{(n)}$$

Proposition 13.23 (Produit, formule de Leibniz)**.** *Soit $n \in \mathbb{N}$ et $(P, Q) \in \mathbb{K}[X]^2$. On a la formule de Leibniz :*

$$(PQ)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} P^{(k)} Q^{(n-k)}$$

Lemme 13.2. *Soit $P \in \mathbb{K}[X] \setminus \left\{0_{\mathbb{K}[X]}\right\}$, $x \in \mathbb{K}$ et $\alpha \in \mathbb{N}$. On a :*

$$m_P(x) = \alpha \iff \exists Q \in \mathbb{K}[X],\ P = (X - x)^\alpha Q \quad et \quad Q(x) \neq 0$$

Lemme 13.3. *Soit $P \in \mathbb{K}[X] \setminus \left\{0_{\mathbb{K}[X]}\right\}$, $x \in \mathbb{K}$ et $\alpha \in \mathbb{N}^\star$. Supposons que $m_P(x) = \alpha$. Alors :*

1. *$P' \neq 0_{\mathbb{K}[X]}$*
2. *$m_{P'}(x) = \alpha - 1$*

Remarque 13.21. Toute racine de multiplicité $\geqslant 2$ de P est donc racine de P'. Ainsi, pour rendre un polynôme P à racines simples de force, il suffit de lui substituer $\dfrac{P}{P \wedge P'}$.

Théorème 13.12 (CNS de multiplicité)**.** *Soit $P \in \mathbb{K}[X] \setminus \left\{0_{\mathbb{K}[X]}\right\}$, $x \in \mathbb{K}$ et $\alpha \in \mathbb{N}$. On a :*

$$m_P(x) = \alpha \iff (\forall k \in [\![0, \alpha - 1]\!],\ P^{(k)}(x) = 0) \quad et \quad P^{(\alpha)}(x) \neq 0$$

Remarque 13.22. Au-delà de α, on ne peut rien dire ! Même si APCR cela sera toujours nul, on ne sait rien de ce qu'il se passe entre $_a lpha$ et ce rang.

4 Bases de $\mathbb{K}_n[X]$

4.1 Base de Taylor

Théorème 13.13. *Soit $(P_i)_{i \in I}$ une famille de polynômes de $\mathbb{K}[X]$* ***tous non nuls*** *ayant des degrés deux à deux distincts. Alors, la famille $(P_i)_{i \in I}$ est* ***libre****.*

Corollaire 13.10 (Familles échelonnées en degré)**.** *On a les deux résultats suivants :*

1. *Soit $n \in \mathbb{N}$ et $(P_k)_{0 \leqslant k \leqslant n} \in \mathbb{K}[X]^{n+1}$ telle que $\forall k \in [\![0, n]\!],\ \deg(P_k) = k$. Alors, $n \in \mathbb{N}$ et $(P_k)_{0 \leqslant k \leqslant n}$ est une base de $\mathbb{K}_n[X]$.*
2. *Soit $(P_k)_{k \in \mathbb{N}} \in \mathbb{K}[X]^{\mathbb{N}}$ telle que $\forall k \in \mathbb{N},\ \deg(P_k) = k$. Alors $(P_k)_{k \in \mathbb{N}}$ est une base de $\mathbb{K}[X]$.*

Théorème 13.14. *Soit $a \in \mathbb{K}$ et $n \in \mathbb{N}$. D'après le corollaire précédent, $((X-a)^k)_{0\leqslant k\leqslant n}$ est une base de $\mathbb{K}_n[X]$. Soit $P \in \mathbb{K}_n[X]$. Sur cette base, P se décompose en :*

$$P(X) = \sum_{k=0}^{n} \frac{P^{(k)}(a)}{k!}(X-a)^k$$

Pa substitution, on a aussi :

$$P(X+a) = \sum_{k=0}^{n} \frac{P^{(k)}(a)}{k!}X^k$$

Remarque 13.23. Par $k!$, on entend $(1_{\mathbb{Z}}.1_{\mathbb{K}}) \times \ldots \times (k_{\mathbb{Z}}.1_{\mathbb{K}})$. On travaille donc en caractéristique nulle.

Remarque 13.24. Puisque toutes les dérivées suivantes sont nulles, on peut en fait écrire :

$$P(X) = \sum_{k\in\mathbb{N}} \frac{P^{(k)}(a)}{k!}(X-a)^k$$

4.2 Base de Lagrange

Dans tous le paragraphe, on fixe $n \in \mathbb{N}$ et $x_0, \ldots, x_n$ $n+1$ scalaires deux à deux distincts.

Définition 13.25 (Polynômes interpolateurs de Lagrange)**.** Les **polynômes interpolateurs de Lagrange associés aux** x_k sont les polynômes $L_0, \ldots, L_n$ définis par :

$$\forall k \in [\![0,n]\!],\ L_k = \prod_{\substack{0\leqslant i\leqslant n\\ i\neq k}} \frac{X-x_i}{x_k-x_i}$$

Proposition 13.24 (Propriété fondamentale des polynômes interpolateurs de Lagrange)**.** *Les polynômes interpolateurs de Lagrange sont des éléments de $\mathbb{K}_n[X]$ et vérifient la propriété fondamentale :*

$$\forall (k,j) \in [\![0,n]\!]^2,\ L_k(x_j) = \delta_{k,j}$$

Théorème 13.15. *La famille $(L_k)_{0\leqslant k\leqslant n}$ est une base de $K_n[X]$. Soit $P \in \mathbb{K}_n[X]$. Sur cette base, P se décompose en :*

$$P = \sum_{k=0}^{n} P(x_k)L_k$$

Remarque 13.25. Soit $k \in [\![0,n]\!]$. A quoi est égale la k-ème forme coordonnée $L_k^\star$ associée à la base $(L_k)_{0\leqslant k\leqslant n}$? Pour tout $P \in \mathbb{K}_n[X]$, on a

$$L_k^\star(P) = P(x_k)$$

donc en fait, $L_k^\star$ est le morphisme d'évaluation $P \mapsto P(x_k)$.

Corollaire 13.11. *Soit $(y_0, \ldots, y_n) \in \mathbb{K}^{n+1}$ quelconque. Alors :*

$$\exists! P \in \mathbb{N}_n[X],\ \forall k \in [\![0,n]\!],\ P(x_k) = y_k$$

Ce polynôme est donné par :

$$P = \sum_{k=0}^{n} y_k L_k$$

4.3 Application à l'interpolation

On veut approximer une fonction $f : I \to \mathbb{K}$. parfois, le calcul de f est tellement coûteux qu'on préfère remplacer f par une approximation polynomiale. Généralement, on demande à ce polynôme de passer par des points obligés appelés **points d'interpolation** qui sont deux à deux distincts. D'après le paragraphe précédent, cela est toujours possible.

Formellement, posons $\forall k \in [\![0, n]\!], y_k = f(x_k)$. Puis on pose

$$\begin{array}{rccc} \varphi : & \mathbb{K}[X] & \longrightarrow & \mathbb{K}^{n+1} \\ & P & \longmapsto & \begin{pmatrix} P(x_0) \\ \vdots \\ P(x_n) \end{pmatrix} \end{array}$$

Le but est donc formellement de résoudre l'équation

$$\varphi(P) = \begin{pmatrix} y_0 \\ \vdots \\ y_n \end{pmatrix}$$

Or φ est linéaire et on sait qu'il existe au moins une solution

$$P_{\text{fond}} = \sum_{k=0}^{n} y_k L_k$$

Ainsi, l'ensemble des solutions vaut $P_{\text{fond}} + \ker(\varphi)$: c'est un sous-espace affine dirigé par $\ker(\varphi)$. Or, en notant $Q = \prod_{i=0}^{n}(X - x_k)$, on obtient que $P \in \mathbb{K}_n[X]$ appartient à $\ker(\varphi)$ si, et seulement si, $Q \mid P$.

En somme, P interpole correctement f si, et seulement si,

$$P \equiv P_{\text{fond}} \ [Q]$$

5 Polynômes irréductibles

5.1 Cas général

Définition 13.26 (Polynôme irréductible)**.** Soit $P \in \mathbb{K}[X]$ **non constant**. On dit que P est **irréductible** lorsque ses seuls diviseurs sont $1_{\mathbb{K}[X]}$, P et leurs associés. Sinon, P est dit **réductible**.

Proposition 13.25. *Soit P irréductible et $Q \in \mathbb{K}[X]$.*

1. *Si $P \nmid Q$, alors $P \wedge Q = 1$.*
2. *Supposons que Q est irréductible. Si $P \nsim Q$, alors $P \wedge Q = 1$.*

Proposition 13.26 (Lemme d'Euclide)**.** *Soit P irréductible, $n \in \mathbb{N}^\star$ et $A_1, \ldots, A_n \in \mathbb{K}[X]$. Si $P \mid A_1 \times \ldots \times A_n$, alors*

$$\exists i \in [\![1, n]\!], \ P \mid A_i$$

Proposition 13.27. *Soit $P \in \mathbb{K}[X]$ non constant. Posons $n = \deg(P) \in \mathbb{N}^\star$. Les quatre assertion suivantes sont équivalentes :*

1. *P est réductible*
2. *$\exists (A,B) \in \mathbb{K}[X]^2,\ P = AB$ et $\deg(A), \deg(B) > 0$*
3. *$\exists (A,B) \in \mathbb{K}[X]^2,\ P = AB$ et $\deg(A), \deg(B) < n$*
4. *$\exists (A,B) \in \mathbb{K}[X]^2,\ P = AB$ et $0 < \deg(A), \deg(B) < 0$*

Théorème 13.16 (Lien entre racines et irréductibilité)**.** *Soit $P \in \mathbb{K}[X]$ non constant.*

Supposons que P est irréductible. Déjà, tout polynôme de degré 1 est irréductible. En revanche, dès que $\deg(P) \geqslant 2$, alors P ne possède pas de racine.

Réciproquement, supposons que P ne possède pas de racine. Déjà, P n'est pas de degré 1. Supposons que P est de degré 2 ou 3. En raisonnant par l'absurde et pour une question de degré, P ne peut pas être réductible (sans quoi un polynôme de degré 1 le diviserait et il possèderait alors une racine). Dès que $\deg(P) \geqslant 4$, tout peut arriver. Par exemple, dans $\mathbb{R}[X]$, $X^4+2X^2+1 = (X^2+1)^2$ est réductible mais ne possède pas de racine. En revanche, dans $\mathbb{Q}[X]$, X^4+1 n'a pas de racine et est irréductible.

Jusqu'à la fin du chapitre, on notera de façon **non standard** $\mathcal{P}$ l'ensemble des polynômes irréductibles unitaires.

Théorème 13.17 (Existence de la DFI)**.** *Soit $A \in \mathbb{K}[X] \setminus \{0\}$. Alors A admet une décomposition en facteurs irréductibles (DFI), c'est-à-dire que*

$$A = \lambda \prod_{P \in I} P^{\alpha_P}$$

où $\lambda \in \mathbb{K}^\star$, I est une partie finie de $\mathcal{P}$ et $\forall P \in I,\ \alpha_P \in \mathbb{N}^\star$.

Exemple 13.14. On a nécessairement $\lambda = CD(A)$ puisque tous les $P \in I$ sont unitaires.

Exemple 13.15. La forme suivante est beaucoup plus opérationnelle :

$$A = \lambda P_1^{\alpha_1} \times \ldots \times P_n^{\alpha_n}$$

avec $\lambda \in \mathbb{K}^\star$, les $P_i \in \mathcal{P}$ deux à deux distincts et les α_i dans $\mathbb{N}^\star$.

Définition 13.27 (Valuation P-adique)**.** Soit $P \in \mathcal{P}$ et $A \in \mathbb{K}[X] \setminus \{0\}$. La **valuation P-adique de A** est l'entier défini par :

$$v_P(A) = \max\left\{k \in \mathbb{N} : P^k \mid A\right\}$$

Exemple 13.16. Soit $x \in \mathbb{K}$. On a $X - x \in \mathcal{P}$. Puis $v_{X-x}(A) = m_A(x)$

Remarque 13.26. On a la même phénomène que dans $\mathbb{Z}$ avec la valuation p-adique et que dans $\mathbb{K}[X]$ avec la multiplicité. Soit $P \in \mathcal{P}$. On a

$$\forall k \in \mathbb{N},\ P^k \mid A \iff k \leqslant v_P(A)$$

d'où on tire

$$\left\{k \in \mathbb{N} : P^k \mid A\right\} = [\![0, v_P(A)]\!]$$

Exemple 13.17. Soit A et B des polynômes non nuls et P irréductible unitaire. Si $A \sim B$, alors $v_P(A) = v_P(B)$

Exemple 13.18. Soit $\alpha \in \mathbb{N}$ et P irréductible unitaire. On a $v_P(P^\alpha) = \alpha$.

Lemme 13.4. *Soit A un polynôme non nul, P irréductible unitaire et $\alpha \in \mathbb{N}$. On a*

$$v_P(A) = \alpha \iff \exists A_1 \in \mathbb{K}[X],\ A = P^\alpha A_1 \quad et \quad P \wedge A_1 = 1$$

Corollaire 13.12. *Soit A et B des polynômes non nuls et P un polynôme irréductible unitaire. On a :*

$$v_P(AB) = v_P(A) + v_P(B)$$

Remarque 13.27. Par une récurrence immédiate, on généralise à un produit fini de polynômes non nuls.

Corollaire 13.13. *Soit $A \in \mathbb{K}[X] \setminus \{0\}$ dont on considère une DFI $A = \lambda \prod_{P \in I} P^{\alpha_P}$. Soit $Q \in \mathcal{P}$. On a :*

1. *Si $Q \in I$, alors $v_Q(A) = \alpha_Q$*
2. *Si $Q \notin I$, alors $v_Q(A) = 0$*

Corollaire 13.14 (Unicité de la DFI)**.** *Dans $\mathbb{K}[X] \setminus \{0\}$, la DFI est unique.*

Remarque 13.28. Un polynôme irréductible unitaire apparaît dans la DFI de A si, et seulement si, il divise A.

Proposition 13.28. *A et B non nuls sont premiers entre eux si, et seulement si, leurs DFI sont disjointes.*

Proposition 13.29. *Soit A et B des polynômes non nuls. On a :*

1. *$A \mid B \iff \forall P \in \mathcal{P},\ v_P(A) \leqslant v_P(B)$*
2. *Supposons de plus que A et B sont unitaires. Alors $A = B \iff \forall P \in \mathcal{P},\ v_P(A) = v_P(B)$*

Proposition 13.30 (Valuation P-adique du PPCM et du PGCD)**.** *Soit A et B deux polynômes non nuls et P un polynôme irréductible unitaire. On a :*

$$v_P(A \wedge b) = \min(v_P(A), v_P(B)) \quad et \quad v_P(A \vee B) = \max(v_P(A), v_P(B))$$

Corollaire 13.15. *Soit A et B des polynômes **unitaires**. On a :*

$$AB = (A \wedge B)(A \vee B)$$

Définition 13.28. Soit $A \in \mathbb{K}[X] \setminus \{0\}$. A est dit **scindé** lorsque, dans sa DFI, tous les facteurs sont de degré 1. En pratique, on pourra donc écrire :

$$A = \lambda \prod_{i=1}^{r} (X - x_i)^{\alpha_i}$$

où $\lambda \in \mathbb{K}^\star$, $r \in \mathbb{N}$, les x_i sont deux à deux distincts et les α_i sont dans $\mathbb{N}^\star$.

Exemple 13.19. Dans ce cas, on a $\forall i \in [\![1,r]\!],\ \alpha_i = v_{X-x_i} = m_A(x_i)$

Remarque 13.29. Dans la suite, pour un polynôme scindé, on pourra écrire $A = \lambda(X-x_1)\dots(X-x_n)$ où les x_i ne seront plus nécessairement supposés deux à deux distincts.

Théorème 13.18 (Formules de Viète)**.** *Soit $A = \lambda(X-x_1)\times\ldots\times(X-x_n)$. Écrivons-le sous la forme $A = \sum_{k=0}^{n} a_k X^k$. Alors, on a :*

$$\forall k \in [\![0,n]\!],\ \sum_{1\leqslant i_1<\ldots<i_k\leqslant n} x_{i_1}\times\ldots\times x_{i_k} = (-1)^k \frac{a_{n-k}}{a_n}$$

Définition 13.29 (Fonctions symétriques élémentaires, HP)**.**

$$(x_1,\dots,x_n) \mapsto \sum_{1\leqslant i_1<\ldots<i_k\leqslant n} x_{i_1}\times\ldots\times x_{i_k}$$

est appelée n**-ème fonction symétrique élémentaire d'ordre** k. On la note plus souvent $\sigma_{n,k}$. Ces fonctions vérifient une propriété fondamentale de décomposition des polynômes à n indéterminées.

5.2 Cas de $\mathbb{R}[X]$ et de $\mathbb{C}[X]$

Théorème 13.19 (Théorème de D'Alembert-Gauss)**.** *Dans $\mathbb{C}[X]$, tout polynôme non constant admet au moins une racine. On dit que $\mathbb{C}$ est* ***algébriquement clos****.*

Proposition 13.31. *Soit $P \in \mathbb{R}[X]\setminus\{0\}$ et $z \in \mathbb{C}$. Alors :*

$$m_P(z) = m_P(\overline{z})$$

Exemple 13.20. En particulier, z est racine de P à coefficients réels non nul si, et seulement si, $\overline{z}$ est racine de P.

Lemme 13.5. *Soit $(A,B) \in \mathbb{R}[X]^2$ tel que B divise A dans $\mathbb{C}[X]$? Alors, B divise A dans $\mathbb{R}[X]$.*

Démonstration. Ce théorème consiste à effectuer la division euclidienne dans $\mathbb{R}[X]$ puis dans $\mathbb{C}[X]$ avant de plonger celle de $\mathbb{R}[X]$ dans $\mathbb{C}[X]$ puis d'invoquer l'unicité de la division euclidienne dans $\mathbb{C}[X]$. □

Remarque 13.30. Plus généralement, ce lemme fonctionne dans deux corps quelconques $\mathbb{L}$ et $\mathbb{K}$ dès que $\mathbb{L}$ est une extension de corps de $\mathbb{K}$.

Théorème 13.20 (Classification des polynômes irréductibles à coefficients complexes)**.** *Dans $\mathbb{C}[X]$, les polynômes irréductibles sont exactement les polynômes de degré* 1*.*

Remarque 13.31. Contrairement aux autres corps où on ne dispose que d'une inégalité, si on prend $P \in \mathbb{C}[X]\setminus\{0\}$ de degré $n \in \mathbb{N}$, alors on a :

$$\sum_{x\in\mathcal{R}(P)} m_P(x) = n$$

Théorème 13.21 (Classification des polynômes irréductibles à coefficients réels)**.** *Dans $\mathbb{R}[X]$, les polynômes irréductibles sont exactement les polynômes de degré 1 et les trinômes du second degré de discriminant strictement négatif.*

Remarque 13.32. Pour factoriser un polynôme dans $\mathbb{R}[X]$, on peut le factoriser dans $\mathbb{C}[X]$ puis regrouper les racines complexes conjuguées.

Corollaire 13.16. *Soit A et B des polynômes à coefficients complexes tous deux différents du polynôme nul. On a :*

$$B \mid A \iff \forall z \in \mathbb{C},\ m_B(z) \leqslant m_A(z)$$

6 Compléments, HP

6.1 Relations coefficients-racines

Voici la démonstration formelle du résultat évoqué en cours.

Théorème 13.22. *Soit $P = \lambda(X-\lambda_1)(X-\lambda_2)...(X-\lambda_n) = \sum_{k=0}^{n} a_k X^k$ un polynôme scindé. Alors*

$$\forall k \in [\![0,n]\!],\ \sum_{1\leq i_1<...<i_k\leq n} \lambda_{i_1}...\lambda_{i_k} = (-1)^k \frac{a_{n-k}}{a_n}$$

Démonstration. Pour $i \in [\![1,n]\!]$ posons $t_{i,0} = X$ et $t_{i,1} = -\lambda_i$..

- Les $t_{i,j}$ sont à valeurs dans l'anneau $\mathbb{K}[X]$ donc la formule de distributivité généralisée permet d'écrire

$$P = \lambda \prod_{i=1}^{n} \left(\sum_{j\in\{0,1\}} t_{i,j} \right) = \lambda \sum_{(j_1,...,j_n)\in\{0,1\}^n} t_{1,j_1}...t_{n,j_n}$$

  Pour chaque n-uplet $(j_1,...,j_n)$ considérons son support S. L'application

$$\begin{array}{ccc} \{0,1\}^n & \to & \mathcal{P}([\![1,n]\!]) \\ (j_1,...,j_n) & \mapsto & S \end{array}$$

  est bijective. Pour s'en convaincre, le mieux est d'exhiber sa bijection réciproque :

$$\begin{array}{ccc} \mathcal{P}([\![1,n]\!]) & \to & \{0,1\}^n \\ S & \mapsto & (\mathbb{1}_s(1),...,\mathbb{1}_s(n)) \end{array}$$

- Par changement de variable, puis par associativité, on en déduit

$$P = \lambda \sum_{S\subset\{1,n\}} \left(\prod_{i\in S}(-\lambda_i) \prod_{i\in S} X \right) = \lambda \sum_{k=0}^{n} (-1)^k \left[\sum_{S\subset\{1,n\}} \left(\prod_{i\in S} \lambda_i \right) \right] X^{n-k}$$

  Pour $k=0$, on obtient $a_n = \lambda$ Plus généralement, fixons $k \in [\![0,n]\!]$. Tout revient à montrer que

$$\sum_{1\leq i_1<...<i_k\leq n} \lambda_{i_1}...\lambda_{i_k} = \sum_{S\subset[\![1,n]\!]} \left(\prod_{i\in S} \lambda_i \right)$$

- Posons $A = \{(i_1, ..., i_k) \in [\![1, n]\!]^k \mid i_1 < ... < i_k\}$ et $B = \{S \subset [\![1, n]\!] : |S| = k\}$. L'application
$$\begin{array}{rccc} \varphi & : & A & \to & B \\ & & (i_1, ..., i_k) & \mapsto & \{i_1, ..., i_k\} \end{array}$$
est bien définie. En effet, les $\{i_j\}$ sont deux à deux disjoints donc le cardinal de leur union vaut bien la somme de leurs cardinaux, c'est-à-dire k. Ensuite elle est bijective. Là aussi, pour s'en convaincre, le mieux est encore d'exhiber sa bijection réciproque ; celle-ci consiste à se donner un ensemble S de cardinal k, à ordonner ses éléments $i_1 < ... < i_k$, puis à retourner le k-uplet $(i_1, ..., i_k)$. Par un dernier changement de variable, on en déduit le résultat final.

Ainsi la preuve est achevée. □

6.2 Théorème de d'Alembert-Gauss

Théorème 13.23. *Tout polynôme non constant de $\mathbb{C}[X]$ admet une racine.*

Démonstration. Il s'agit d'un des théorèmes les plus difficiles de l'année, et sa preuve est très clairement hors-programme. Nous la donnons pour les plus valeureux. Il s'agit peu ou prou de la démonstration originale de Cauchy. Sans perte de généralité, on suppose que P est unitaire.

1. Si on écrit $P(z) = z^n + a_{n-1}z^{n-1} + ... + a_0$ avec $n \geqslant 1$, alors
$$\forall z \neq 0,\ |a_{n-1}z^{n-1} + ... + a_0| = |z|^n \cdot \left|\frac{a_{n-1}}{z} + ... + \frac{a_0}{z^n}\right| \leq |z|^n(\frac{|a_{n-1}|}{|z|} + ... + \frac{|a_0|}{|z|^n})$$
Puisque $\frac{|a_{n-1}|}{R} + ... + \frac{|a_0|}{R^n} \underset{R\to+\infty}{\longrightarrow} 0$, on peut choisir $R > 0$ tel que
$$\forall z \in \mathbb{C}, |z| \geqslant R \Rightarrow \frac{|a_{n-1}|}{|z|} + ... + \frac{|a_0|}{|z|^n} \leq \frac{1}{2}$$
Soit $z \in \mathbb{C}$. Par la deuxième inégalité triangulaire, on a
$$|z| \geqslant R \Rightarrow |P(z)| \geq |z^n| - |a_{n-1}z^{n-1} + ... + a_0| \geq \frac{|z|^n}{2} \geq \frac{R^n}{2}$$
Or on a $\frac{|a_0|}{R^n} \leq \frac{1}{2}$, d'où finalement, $|z| \geqslant R \Rightarrow |P(z)| \geq |a_0| = |P(0)|$.
2. Considérons alors le disque fermé $D = \{z \in \mathbb{C} : |z| \leqslant R\}$. L'ensemble $\{|P(z)||z \in D\}$ étant non vide et minoré, il admet une borne inf m et il existe une suite à valeurs dans D telle que $|P(z_k)| \to m$. Puisque (z_k) est bornée, quitte à considérer une suite extraite, on peut supposer que $z_k \to \tilde{z}$. Par passage au module, on a alors $|z_k| \to |\overline{z}|$ puis par passage à la limite on a $|\tilde{z}| \leqslant R$.
Soit $k \in \mathbb{N}$. On a ensuite
$$\begin{aligned} \big||P(\tilde{z})| - |P(z_k)|\big| &\leq |P(\tilde{z}) - P(z_k)| \\ &\leq |\tilde{z}^n - z_k^n| + |a_{n-1}| \cdot |\tilde{z}^{n-1} - z_k^{n-1}| + ... + |a_1| \cdot |\tilde{z} - z_k| \end{aligned}$$
Si on peut prouver que chacun des $|\tilde{z}^j - z_k^j|$ tend vers 0 lorsque $k \to +\infty$, on en déduira par passage à la limite que $\big||P(\vec{z})| - m\big| = 0$. Or on a
$$|\tilde{z}^j - z_k^j| \leq |\tilde{z} - z_k| \cdot (|\tilde{z}|^{j-1} + |\tilde{z}|^{j-2}|z_k| + ... + |z_k|^{j-1}) \leq |\tilde{z} - z_k| \cdot jR^{j-1}$$
ce qui permet de conclure.

3. Par le point 2., $|P(z)|$ admet un minimum m sur D, qu'il atteint en $\tilde{z}$. Et d'après le point 1., ce minimum est en fait un minimum global sur $\mathbb{C}$ Supposons maintenant que $m > 0$ et aboutissons à une contradiction.
Considérons la fonction $z \mapsto P(\overline{z} + z)$. D'après la formule de Taylor on peut écrire $\forall z \in \mathbb{C},\ P(\overline{z}+z) = b_n z^n + ... + b_1 z + b_0$ avec $|b_0| = |P(\tilde{z})| = m \neq 0$. Puisque P n'est pas constant, le polynôme $P(\overline{z} + X)$ n'est pas constant non plus, et il existe donc $k \in [\![1, n]\!]$ tel que $b_k \neq 0$. Considérons k minimal pour cette propriété. On a

$$\forall z \in \mathbb{C},\ P(\tilde{z}+z) = b_n z^n + ... + b_k z^k + b_0$$

Soit alors c une racine k-ème de $-\overline{b_k} b_0$. On a

$$\begin{aligned}\forall t > 0, |P(\overline{z} + ct)| &\leqslant \left| \sum_{k<j\leqslant n} b_j c^j t^j \right| + \left| - |b_k|^2 b_0 t^k + b_0 \right| \\ &\leqslant t^{k+1} \sum_{0\leqslant j\leqslant n-k} |b_{j+k+1} c^{j+k+1}| \cdot t^j + m \left| 1 - |b_k|^2 t^k \right|.\end{aligned}$$

Soit maintenant $t \in [0, 1]$. Alors la somme est majorée par un certain $M > 0$ fixé (notons que si elle est vide, il n'y a rien à faire). Quitte à diminuer encore t, supposons que $1 - |b_k|^2 t^k \geqslant 0$, si bien que

$$|P(\tilde{z} + ct)| \leqslant t^{k+1} M + m \left(1 - |b_k|^2 t^k\right)$$

Quitte à diminuer une dernière fois t, supposons que $|t|M \leq \frac{1}{2} m |b_k|^2$, ce qui est possible car on a supposé $m > 0$ (c'est le point-clé !). On a alors

$$|P(\overline{z} + ct)| \leqslant m + t^k \left(tM - m |b_k|^2\right) \leqslant m - \frac{1}{2} m |b_k|^2 t^k < m$$

Contradiction.

Ainsi la preuve est achevée. □

6.3 Polynômes à coefficients dans $\mathbb{Z}$: contenu et critère d'Eisenstein

On présente ici la notion de contenu d'un polynôme à coefficients dans $\mathbb{Z}$ et de polynôme primitif. Cela nous sert à démontrer le célèbre **critère d'Eisenstein**, grand classique d'X/ENS. On étend enfin la définition du contenu aux polynômes à coefficients dans $\mathbb{Q}$. Ce travail peut être généralisé plus généralement avec un anneau factoriel et son corps des fractions.

Définition 13.30. On rappelle que l'ensemble des polynômes à coefficients dans $\mathbb{Z}$, noté $\mathbb{Z}[X]$, muni des lois usuelles, est un anneau intègre puisque $\mathbb{Z}$ est intègre.

Proposition 13.32. *Les éléments inversibles de $\mathbb{Z}[X]$ sont les polynômes constants égaux à -1 et 1.*

Démonstration. Comme pour $\mathbb{K}[X]$, on raisonne d'abord sur le degré, mais il faut ensuite raisonner sur le coefficient dominant qui doit être une unité de $\mathbb{Z}$. □

Définition 13.31 (Contenu d'un polynôme à coefficients dans $\mathbb{Z}$)**.** Soit P un polynôme de $\mathbb{Z}[X]\backslash\{0\}$ qu'on écrit sous la forme $P = \sum_{k=0}^{n} a_k X^k$ avec $a_n \neq 0$. On appelle **contenu de** P, et on note $c(P)$, le PGCD des coefficients de P :

$$c(P) := a_0 \wedge \ldots \wedge a_n$$

Définition 13.32 (Polynôme primitif)**.** On dit que $P \in \mathbb{Z}[X] \setminus \{0\}$ est **primitif** lorsque $c(P) = 1$.

Proposition 13.33. *On a*

$$\forall \lambda \in \mathbb{Z}^\star,\ \forall P \in \mathbb{Z}[X] \setminus \{0\},\ c(\lambda P) = |\lambda| c(P)$$

Démonstration. C'est une simple réécriture de la distributivité du PGCD. □

Lemme 13.6. *Soit $P \in \mathbb{Z}[X] \setminus \{0\}$. Alors le polynôme $\dfrac{P}{c(P)}$ est primitif.*

Démonstration. On a $c(P) = c\left(c(P) \cdot \dfrac{P}{c(P)}\right) = c(P)c\left(\dfrac{P}{c(P)}\right)$ donc $c\left(\dfrac{P}{c(P)}\right) = 1$. □

Proposition 13.34. *Le produit de deux polynômes primitifs est primitif.*

Démonstration. Soit $P \in \mathbb{Z}[X] \setminus \{0\}$ et $Q \in \mathbb{Z}[X] \setminus \{0\}$ tels que $c(P) = c(Q) = 1$. Déjà, par intégrité de $\mathbb{Z}[X]$, $PQ \neq 0$ donc $c(PQ)$ est bien défini.

Raisonnons par l'absurde et supposons que $c(PQ) \neq 1$. Alors $c(PQ) > 1$ donc $c(PQ)$ possède un diviseur premier p. En passant dans $\mathbb{Z}/p\mathbb{Z}[X]$, on a $\overline{PQ} = \overline{P} \times \overline{Q} = \overline{0}$, donc puisque $\mathbb{Z}/p\mathbb{Z}$ est un corps, $\mathbb{Z}/pZ[X]$ est un anneau intègre et donc $\overline{P} = \overline{0}$ ou $\overline{Q} = \overline{Q}$. Ainsi, les coefficients de P ou de Q sont tous divisibles par p. Contradiction avec le fait que P et Q sont primitifs. Donc PQ est primitif, *ie* $c(PQ) = 1$. □

Corollaire 13.17 (Contenu d'un produit de polynôme à coefficients dans $\mathbb{Z}$)**.** *Soit $P \in \mathbb{Z}[X] \setminus \{0\}$ et $Q \in \mathbb{Z}[X] \setminus \{0\}$. On a*

$$c(PQ) = c(P)c(Q)$$

Démonstration. On a

$$c(PQ) = c\left(c(P)c(Q) \cdot \frac{PQ}{c(P)c(Q))}\right) = c(P)c(Q)c\left(\frac{PQ}{c(P)c(Q)}\right) = c(P)c(Q)$$

puisque $\dfrac{PQ}{c(P)c(Q)}$ est primitif. □

On démontre un dernier lemme avant d'arriver au critère d'Eisenstein.

Lemme 13.7. *Soit $A \in \mathbb{Z}[X]$ non irréductible dans $\mathbb{Q}[X]$. Alors il existe B et C dans $\mathbb{Z}[X]$ tels que $\deg(B) < \deg(A)$, $\deg(C) < \deg(A)$ et $A = BC$.*

Démonstration. On pose $A_1 := \dfrac{A}{c(A)}$. Alors A_1 n'est pas irréductible dans $\mathbb{Q}[X]$ non plus et il existe B_1 et C_1 dans $\mathbb{Q}[X]$ tels que $\deg(B_1) < \deg(A_1)$, $\deg(C_1) < \deg(A_1)$ et $A_1 = B_1C_1$. Notons β le

produit des dénominateurs de B_1 et γ le produit des dénominateurs de C_1. Alors $B_2 := \beta B_1 \in \mathbb{Z}[X]$ et $C_2 := \gamma C_1 \in \mathbb{Z}[X]$. Puis :

$$\beta\gamma A_1 = \beta\gamma B_1 c_1 = B_2 C_2$$

En passant au contenus :

$$\beta\gamma = c(\beta\gamma A_1) = c(B_2C_2) = c(B_2)c(C_2)$$

Puis

$$A = c(A)A' = c(A)\frac{BC}{\beta\gamma} = \left(c(A)\frac{B_2}{c(B_2)}\right)\left(\frac{C_2}{C_2}\right)$$

Ainsi, $B := c(A)\dfrac{B_2}{c(B_2)}$ et $C := \dfrac{C_2}{c(C_2)}$ conviennent. □

Théorème 13.24 (Critère d'Eisenstein)**.** *Soit $A = \sum_{k=0}^{n} a_k X^k \in \mathbb{Z}[X]$ et p un nombre premier. Supposons que*

$$\begin{cases} p \nmid a_n \\ \forall k \in [\![0, n-1]\!],\ p \mid a_k \\ p^2 \nmid a_0 \end{cases}$$

alors A est irréductible dans $\mathbb{Q}[X]$.

Démonstration. Raisonnons par l'absurde et supposons que A ne soit pas irréductible dans $\mathbb{Q}[X]$. On peut alors fixer B et C dans $\mathbb{Z}[X]$ tels que $\deg(B), \deg(C) < \deg(A)$ et $A = BC$. Écrivons $B = \sum_{k=0}^{i} b_k X^k$ et $C = \sum_{k=0}^{j} c_k X^k$.

En passant dans $\mathbb{Z}/p\mathbb{Z}[X]$, on a $\overline{A} = \overline{a_n} X^n$. Or, $a_n = b_i c_j$ donc $\overline{a_n} = \overline{b_i} \times \overline{c_j}$. Par intégrité de $\mathbb{Z}/p\mathbb{Z}$, $\overline{b_i} \neq \overline{0}$ et $\overline{c_j} \neq \overline{0}$ donc $\overline{B}$ et $\overline{C}$ sont de degrés respectifs i et j dans $\mathbb{Z}/p\mathbb{Z}$. On a donc :

$$\overline{A} = \overline{b_i} X^i \times \overline{c_j} X^j = \overline{B} \times \overline{C}$$

Puisque X est irréductible dans $\mathbb{Z}/p\mathbb{Z}[X]$, en comparant les DFI et les degrés, on a $\overline{B} = \overline{b_i} X^i$ et $\overline{C} = \overline{c_j} X^j$. Or, par hypothèse, on a nécessairement $i > 0$ et $j > 0$. Donc p divise b_0 et c_0. Or, $a_0 = b_0 c_0$ donc p^2 divise a_0. Absurde d'après le troisième point du critère. Donc A est irréductible dans $\mathbb{Q}[X]$. □

Exemple 13.21 (Application)**.** Comme application, on montre qu'il existe des polynômes à coefficient dans $\mathbb{Z}$ d'un degré quelconque irréductibles dans $\mathbb{Q}[X]$. Il suffit de considérer $X^n - 2$ et de lui appliquer le critère d'Eisenstein avec $p = 2$. On peut utiliser ce fait et un autre résultat qui découle du critère pour montrer que $\mathbb{R}$ est un $\mathbb{Q}$-ev de dimension infinie.

On peut étendre la définition du contenu aux polynômes à coefficients dans $\mathbb{Q}$ comme suit.

Définition 13.33 (Contenu d'un polynôme à coefficients dans $\mathbb{Q}$)**.** Soit $P \in \mathbb{Q}[X] \setminus \{0\}$. Il existe alors au moins un $\lambda \in \mathbb{N}^\star$ tel que $\lambda P \in \mathbb{Z}[X]$. On pose alors

$$c(P) := \frac{c(\lambda P)}{\lambda}$$

Démonstration. Pour λ, il suffit de prendre le PPCM des dénominateurs des coefficients de P ou la valeur absolue du produit des dénominateurs des coefficients de P.

Ensuite, montrons que la définition de $c(P)$ ne dépend pas du choix de λ. Soit $\lambda \in \mathbb{N}^\star$ et $\mu \in \mathbb{N}^\star$ tels que $\lambda P \in \mathbb{Z}[X]$ et $\mu P \in \mathbb{Z}[X]$. Alors $\lambda\mu P \in \mathbb{Z}[X]$ et on a $c(\lambda\mu P) = \lambda c(\mu P) = \mu c(\lambda P)$ donc $\dfrac{c(\lambda P)}{\lambda} = \dfrac{c(\mu P)}{\mu}$.

Enfin vérifions que la définition prolongée celle donnée dans $\mathbb{Z}[X]$. Supposons que $P \in \mathbb{Z}[X]\setminus\{0\}$. Alors $c_{\mathbb{Q}}(P) = \frac{c_{\mathbb{Z}}(1P)}{1} = c_{\mathbb{Z}}(P)$. La définition prolonge bien celle de $\mathbb{Z}$. □

Corollaire 13.18 (Contenu d'un produit de polynôme à coefficients dans $\mathbb{Q}$)**.** *Soit $P \in \mathbb{Q}[X] \setminus \{0\}$ et $Q \in \mathbb{Q}[X] \setminus \{0\}$. On a*

$$c(PQ) = c(P)c(Q)$$

Démonstration. Fixons $\lambda \in \mathbb{N}^\star$ tel que $\lambda P \in \mathbb{Z}[X]$ et $\mu \in \mathbb{N}^\star$ tel que $\mu Q \in \mathbb{Z}[X]$. Alors $\lambda\mu PQ \in \mathbb{Z}[X]$ et on a :

$$c(PQ) = \frac{c(\lambda\mu PQ)}{\lambda\mu} = \frac{c(\lambda P)c(\mu Q)}{\lambda\mu} = c(P)c(Q)$$

d'après le cas de $\mathbb{Z}[X]$. □

Proposition 13.35. *Soit $P \in \mathbb{Q}[X] \setminus \{0\}$. Si $c(P) = 1$, alors $P \in \mathbb{Z}[X] \setminus \{0\}$.*

Démonstration. Écrivons $P = \sum_{k=0}^{n} a_k X^k$ avec $a_n \neq 0$. Soit $\lambda \in \mathbb{N}^\star$ tel que $\lambda P \in \mathbb{Z}[X]$. On a alors $c(\lambda P) = \lambda$. En particulier, λ divise chacun des entiers λa_k. Soit $k \in [\![0, n]\!]$. On peut donc écrire $\lambda a_k = \lambda q_k$ avec $q_k \in \mathbb{Z}$. Comme $\lambda \neq 0$, on en déduit que $a_k = q_k \in \mathbb{Z}$. □

Remarque 13.33 (Extension de la définition de primitif)**.** On peut alors étendre la notion de polynôme primitif à $\mathbb{Q}$ sans problème, puisqu'alors tout polynôme primitif est nécessairement à coefficients dans $\mathbb{Z}$.

Remarque 13.34 (Contenu et anneau factoriel)**.** Plus généralement, on peut réaliser le même travail de définition du contenu dans n'importe quel anneau factoriel ($\mathbb{Z}$ est factoriel car principal) et son corps des fractions ($\mathbb{Q}$ ici). En revanche, il sera alors défini "à un inversible près" (là où on impose la positivité du PGCD dans $\mathbb{Z}$). Pour plus de détail sur les anneaux factoriels, se référer au chapitre "Compléments sur les anneaux commutatifs".

Chapitre 14

Fractions rationnelles

Dans tout le chapitre, on fixe $\mathbb{K}$ un corps. En pratique, cela sera souvent $\mathbb{R}$ ou $\mathbb{C}$, voire $\mathbb{Q}$.

1 Premières définitions

Définition 14.1 (Corps des fractions rationnelles)**.** $(\mathbb{K}[X], +, \times)$ étant un anneau intègre, on peut considérer son corps des fractions. Il sera noté $K(X)$ et, conformément à l'usage, ses éléments seront notés $\dfrac{P}{Q}$ (où $(P, Q) \in \mathbb{K}[X] \times (\mathbb{K}[X] \setminus \{0\})$). Pour mémoire, $\dfrac{P}{Q}$ désigne la classe d'équivalence du couple (P, Q).

Remarque 14.1. On vérifie que la structure d'espace vectoriel de $\mathbb{K}[X]$ se prolonge à $\mathbb{K}(X)$.

Définition 14.2 (Degré)**.** Soit $F \in \mathbb{K}[X]$. On fixe $(P, Q) \in \mathbb{K}[X] \times (\mathbb{K}[X] \setminus \{0\})$ tel que $F = \dfrac{P}{Q}$ puis on définit le degré de F par

$$\deg(F) = \deg(P) - \deg(Q)$$

Cette définition est indépendante du choix du représentant de F.

Remarque 14.2. Attention, contrairement au cas de $\mathbb{K}[X]$, les fractions rationnelles de degré 1 ne sont pas nécessairement constantes ! Par exemple, on pourra considérer $\dfrac{X+1}{X}$.

Proposition 14.1 (Propriétés du degré)**.** *Soit $F_1, F_2 \in \mathbb{K}(X)$. On a :*

1. $\deg(F_1 + F_2) \leqslant \max(\deg(F_1), \deg(F_2))$
2. $\deg(F_1 F_2) = \deg(F_1) + \deg(F_2)$

La propriété sur la somme reste vraie pour une combinaison linéaire de F_1 et F_2.

Définition 14.3. Soit $F \in \mathbb{K}(X)$. Il existe un unique couple $(P, Q) \in \mathbb{K}[X] \times (\mathbb{K}[X] \setminus \{0\})$ tel que $F = \dfrac{P}{Q}$, $P \wedge Q = 1$ et Q est unitaire. $\dfrac{P}{Q}$ s'appelle alors **l'écriture irréductible** de F.

Définition 14.4 (Racines, pôles)**.** Soit $F \in \mathbb{K}(X)$ qu'on écrit sous forme irréductible $\dfrac{P}{Q}$. On appelle **racine** de F toute racine de P. On appelle **pôle** de F toute racine de Q.

Définition 14.5 (Évaluation)**.** Soit $F \in \mathbb{K}(X)$ qu'on écrit sous forme irréductible $\dfrac{P}{Q}$. On note $\mathcal{P}$ l'ensemble de ses pôles. Pour tout $x \in \mathbb{K} \setminus \mathcal{P}$, on pose :

$$F(x) = \frac{P(x)}{Q(x)}$$

Définition 14.6 (Fonction rationnelle)**.** La fonction $\begin{array}{rccc} \tilde{F} : & \mathbb{K} \setminus \mathcal{P} & \longrightarrow & \mathbb{K} \\ & x & \longmapsto & F(x) \end{array}$ s'appelle la **fonction rationnelle** associée à F.

Proposition 14.2. *Soit $F \in \mathbb{K}(X)$ et $\dfrac{A}{B}$ une écriture quelconque de F. Soit $x \in \mathbb{K}$. Si $B(x) \neq 0_{\mathbb{K}}$, alors $x \notin \mathcal{P}$ donc $F(x)$ existe, et on a*

$$F(x) = \frac{A(x)}{B(x)}$$

Proposition 14.3 (Évaluation et opérations algébriques)**.** *Du moment que x n'est ni un pôle de F_1, ni un pôle de F_2, alors x n'est pas un pôle de toute combinaison linéaire de F_1 et F_2 ni du produit F_1F_2 et l'évaluation passe alors à ces opérations :*

$$(\lambda F_1 + \mu F_2)(x) = \lambda F_1(x) + \mu F_2(x)$$

et

$$(F_1F_2)(x) = F_1(x)F_2(x)$$

Corollaire 14.1. *Si F_1 et F_2 coïncident en une infinité de points, alors $F_1 = F_2$.*

Définition 14.7. Soit $F \in \mathbb{K}(X)$ qu'on écrit sous forme quelconque $\dfrac{A}{B}$. On pose alors :

$$F(-X) = \frac{A(-X)}{B(-X)}$$

Définition 14.8 (Parité, imparité)**.** F est dite **paire** lorsque $F(-X) = F$. F est dite **paire** lorsque $F(-X) = -F$.

Proposition 14.4. *On a*

$$(\lambda F_1 + \mu F_2)(-X) = \lambda F_1(-X) + \mu F_2(-X)$$

et

$$(F_1F_2)(-X) = F_1(-X)F_2(-X)$$

Définition 14.9 (Généralisation)**.** Plus généralement, soit $U \in \mathbb{K}(X)$ non constant. $F(U) = \dfrac{A(U)}{B(U)}$ est bien défini et ne dépend pas du représentant $\dfrac{A}{B}$ et les résultats de la proposition précédente restent vrais

$$(\lambda F_1 + \mu F_2)(U) = \lambda F_1(U) + \mu F_2(U)$$

et

$$(F_1F_2)(U) = F_1(U)F_2(U)$$

Ainsi, on pourra passer de F à $F(X+1)$, $F(X^2)$ et même $F(X)$!

2 Décomposition en éléments simples

Définition 14.10 (Partie entière)**.** Soit $F \in \mathbb{K}(X)$. Alors il existe un unique couple $(E, G) \in \mathbb{K}[X] \times \mathbb{K}(X)$ tel que $F = E + G$ et $\deg(G) < 0$. E s'appelle la **partie entière** de F.

Définition 14.11. On pose

$$\mathbb{K}_{-1} = \{G \in \mathbb{K}(X) \mid \deg(G) \leqslant -1\}$$

C'est un sous-espace vectoriel de $\mathbb{K}(X)$ qui vérifie

$$\mathbb{K}(X) = \mathbb{K}[X] \oplus \mathbb{K}_{-1}(X)$$

Remarque 14.3. Après avoir effectué cette première décomposition, on peut se demander si G ne peut pas se décomposer en éléments plus simples. C'est précisément l'objet de la suite du chapitre.

Définition 14.12. Une **élément simple** de $\mathbb{K}(X)$ est une fraction rationnelle $\dfrac{P}{Q^j}$ avec $j \in \mathbb{N}^\star$, Q irréductible unitaire et $\deg(P) < \deg(Q)$

Théorème 14.1. *La famille* $\left(\dfrac{X^k}{Q^j}\right)$ *avec* Q *irréductible unitaire,* $j \in \mathbb{N}^\star$ *et* $0 \leqslant k < \deg(Q)$ *est une base de* $\mathbb{K}_{-1}(X)$*. De plus, si* $\dfrac{P}{Q}$ *est une écriture irréductible de* F *et si* $\lambda Q_1^{\alpha_1} \dots Q_n^{\alpha_n}$ *est la DFI de* Q*, alors la décomposition peut s'écrire*

$$\frac{P}{Q} = E + \sum_{i=1}^{n} \left(\sum_{j=1}^{\alpha_i} \frac{P_{i,j}}{Q_i^j} \right)$$

Autrement dit, on peut se restreindre aux éléments simples dont le dénominateur est de la forme Q_i^j *avec* $1 \leqslant j \leqslant \alpha_i$*.*

Remarque 14.4. Le premier point est utile pour les exercices théoriques. Pour les exercices calculatoires, utiliser la forme irréductible.

3 Méthodes pratiques

Méthode 14.1 (Méthode 1 - Identification)**.** On effectue la DES théorique, puis on met tout au même dénominateur que F avant de multiplier par ce dénominateur puis d'identifier les coefficients du polynôme obtenu. On obtient alors un système linéaire à résoudre.

Théorème 14.2 (DES de $\frac{P'}{P}$ pour P scindé)**.** *Soit $P \in \mathbb{K}[X]$ un polynôme scindé dont on écrit la DFI sous la forme $P = \lambda \prod_{i=1}^{r} (X - x_i)^{\alpha_i}$. Alors, la DES de $\frac{P'}{P}$ est donnée par*

$$\frac{P'}{P} = \sum_{i=1}^{r} \frac{\alpha_i}{X - x_i}$$

Théorème 14.3 (Coefficient pour un pôle simple)**.** *Soit $(P, Q) \in \mathbb{K}[X] \times (\mathbb{K}[X] \setminus \{0\})$ et λ une racine simple de Q. Fixons $Q_1 \in \mathbb{K}[X]$ tel que $Q = (X - \lambda)Q_1$ et $Q_1(\lambda) \neq 0$. Alors, dans la DES de $\frac{P}{Q}$, le terme en $\frac{a}{X - \lambda}$ est donné par :*

$$a = \frac{P(\lambda)}{Q_1(\lambda)} = \frac{P(\lambda)}{Q'(\lambda)}$$

Théorème 14.4 (Coefficient pour un pôle de multiplicité quelconque, HP)**.** *Soit $(P, Q) \in \mathbb{K}[X] \times (\mathbb{K}[X] \setminus \{0\})$ et λ une racine de multiplicité $p \in \mathbb{N}^\star$ de Q. Fixons $Q_p \in \mathbb{K}[X]$ tel que $Q = (X - \lambda)^p Q_p$ et $Q_p(\lambda) \neq 0$. Alors, dans la DES de $\frac{P}{Q}$, le terme en $\frac{a}{(X - \lambda)^p}$ est donné par :*

$$a = \frac{P(\lambda)}{Q_p(\lambda)} = \frac{p! P(\lambda)}{Q^{(p)}(\lambda)}$$

Méthode 14.2 (Méthode 2 - "Évaluations" par les théorèmes qui précèdent)**.** On utilise les deux théorèmes précédents et on itère. On peut utiliser les premiers résultats obtenus sur certains coefficients pour se ramener à des calculs plus simples pour les coefficients suivants.

Remarque 14.5. On peut toujours évaluer la DES en un point qui n'est pas un pôle de la fraction rationnelle de départ.

Méthode 14.3 (Méthode 3 - Évaluation en un point qui n'est pas un pôle)**.** On peut évaluer en un point qui n'est pas un pôle pour obtenir une équation linéaire reliant les différentes inconnues. En couplant cela avec les autres méthodes, cela peut accélérer les calculs finaux.

Théorème 14.5 (Passage dans $\mathbb{C}$)**.** *Soit $F \in \mathbb{C}(X)$ et supposons qu'il existe A et B des polynômes à coefficients réels tels que $F = \frac{A}{B}$. La DFI de B peut s'écrire sous la forme*

$$B = \lambda \prod_{i=1}^{r} (X - \lambda_i)^{\alpha_i} \prod_{j=1}^{s} (\textit{Trinôme irréductible unitaire})^{\beta_j}$$

avec les $\lambda_i \in \mathbb{R}$. Or, chaque trinôme irréductible unitaire peut se factoriser sous la forme $(X - \mu_j)(X - \overline{\mu_j})$ avec les $\mu_j \in \mathbb{C} \setminus \mathbb{R}$. Puis on a alors

$$B = \lambda \prod_{i=1}^{r} (X - \lambda_i)^{\alpha_i} \prod_{j=1}^{s} (X - \mu_j)^{\beta_j} (X - \overline{\mu_j})^{\beta_j}$$

Dans $\mathbb{C}(X)$, la DES de $F = \dfrac{A}{B}$ va donc être de la forme :

$$F = E + \sum_i \sum_k \frac{a_{i,k}}{(X-\lambda_i)^k} + \sum_j \sum_l \left(\frac{b_{j,l}}{(X-\mu_j)^l} + \frac{b'_{j,l}}{(X-\overline{\mu_j})^l} \right)$$

Alors, dans cette DES :

1. *Tous les $a_{i,k}$ sont réels*
2. *Chaque $b'_{j,l}$ vaut $\overline{b_{j,l}}$*

Méthode 14.4 (Méthode 4 - Écriture à coefficients réels et passage dans $\mathbb{C}(X)$)**.** Si F admet une écriture à coefficients réels, utiliser le théorème précédent.

Méthode 14.5 (Méthode 5 - Parité et imparité)**.** Supposons que F est paire ou impaire. On commence par écrire sa DES

$$F = E + \sum_{i,j} \frac{a_{i,j}}{(X-\lambda_i)^j}$$

Or, on a

$$F(-X) = E(-X) + \sum_{i,j} \frac{a_{i,j}}{(-X-\lambda_i)^j}$$

En utilisant l'unicité de la DES, on peut obtenir des relations sur les coefficients.

Méthode 14.6 (Méthode 6 - Passage à la limite dans $\mathbb{R}(X)$)**.** Soit $d \in \mathbb{N}$. on peut écrire $x \mapsto x^d F(x)$ de 2façons différentes avec la forme initiale de F ou avec sa DES. Lorsqu'on fait tendre x vers $+\infty$, on obtient des relations entre les coefficients par unicité de la limite puisque l'on connaît bien les limites de fractions rationnelles en $+\infty$. Pour que la méthode ait un intérêt, il faut que d ne soit pas trop grand. Régulièrement, on prend $d = -\deg(F)$, mais on peut le prendre différemment parfois. On peut aussi faire des limites en $-\infty$.

Méthode 14.7 (Méthode 7 - Plongement dans $\mathbb{C}(X)$)**.** Soit $F \in \mathbb{R}(X)$. on l'interprète temporairement comme une fraction rationnelle de $\mathbb{C}(X)$ puis on regroupe les termes conjugués pour revenir à $\mathbb{R}(X)$.

4 Primitives de fonctions rationnelles

On considère $f \in \mathbb{K}(X)$ avec $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$. On note $\mathcal{P}$ l'ensemble de ses pôles. On cherche des primitives $\mathbb{R} \setminus \mathcal{P} \to \mathbb{K}$

Pour cela, on considère la DES de F. La partie entière étant polynomiale, on ne s'y intéressera pas et on considèrera uniquement $\mathbb{K}_{-1}(X)$.

4.1 Cas de $\mathbb{C}(X)$

Les éléments simples sont, à un scalaire près et transformées en fonction rationnelles, les

$$x \mapsto \frac{1}{(x-\lambda)^p}$$

avec $\lambda \in \mathbb{C}$ et $p \in \mathbb{N}^\star$.

Théorème 14.6. *Il faut distinguer les cas :*

1. *Supposons que $p \geqslant 2$. Une primitive de $x \mapsto (x-\lambda)^{-p}$ est donnée par*

$$x \mapsto \frac{1}{1-p}(x-\lambda)^{1-p}$$

Elle est définie sur $]-\infty, \lambda[$ ou $]\lambda, +\infty[$ si $\lambda \in \mathbb{R}$ et sur $\mathbb{R}$ si $\lambda \in \mathbb{C} \setminus \mathbb{R}$.

2. *Supposons que $p = 1$. **Attention**, pas de $x \mapsto \ln(|x-\lambda|)$, cela **ne fonctionne pas** dès que $\lambda \in \mathbb{C} \setminus \mathbb{R}$. Il faut passer par la **quantité conjuguée** et on se ramène à du* ln *réel et une* arctan *qui peut être affectée d'un scalaire i.*

Proposition 14.5 (Cas général pour $p = 1$, HP)**.** *On cherche à primitive $x \mapsto \dfrac{1}{x-(\alpha+i\beta)}$ avec $\alpha \in \mathbb{R}$ et $\beta \in \mathbb{R}^\star$. La primitive est donnée par*

$$x \mapsto \frac{1}{2}\ln\left((x-\alpha)^2+\beta^2\right) + i\arctan\left(\frac{x-\alpha}{\beta}\right)$$

cette primitive est alors définie sur $\mathbb{R}$.

Démonstration. A savoir retrouver vite. On peut le retrouver rapidement en passant par la quantité conjuguée et en effectuant les manipulations usuelles de primitivation. □

4.2 Cas de $\mathbb{R}(X)$

On peut passer dans $\mathbb{R}(X)$ et adapter les résultats du paragraphe précédent. Pour les éléments simples de première espèce réels, on le résultat est toujours valables. Pour ceux complexes non réels, on les regroupe directement avec leur conjugué.

Proposition 14.6 (Cas général, HP)**.** *Supposons que $\lambda = \alpha + i\beta \notin \mathbb{R}$ et cherchons une primitive de*

$$x \mapsto \frac{a}{x-\lambda} + \frac{\overline{a}}{x-\overline{\lambda}}$$

Une primitive est donnée par

$$x \mapsto \frac{a+\overline{a}}{2}\ln\left((x-\alpha)^2+\beta^2\right) + i(a-\overline{a})\arctan\left(\frac{x-\alpha}{\beta}\right)$$

Démonstration. Idem, à savoir retrouver vite. Pour le retrouver, mettre tout au même dénominateur puis effectuer les manipulations usuelles. □

Exercice 14.1. Trouver une primitive de $x \mapsto \dfrac{1}{(1+x+x^2)^3}$

1) Par un changement de variable affine, se ramener à une fonction de la forme $t \mapsto \dfrac{1}{(1+t^2)^3}$

2) 1ère méthode : Écrire t sous la forme $\tan(u)$

3) 2ème méthode : Par une IPP, trouver par récurrence une primitive de $t \mapsto \dfrac{1}{(1+t^2)^p}$

5 Compléments, HP

5.1 Démonstration du théorème de décomposition en éléments simples

Théorème 14.7. *La famille* $(\frac{X^k}{Q^j})$ *avec* Q *irréductible unitaire,* $j \in \mathbb{N}^*$ *et* $0 \leqslant k < deg(Q)$ *est une base de* $\mathbb{K}_{-1}(X)$*. De plus, si* $\frac{P}{Q}$ *est une écriture de* F *et si* $\lambda Q_1^{\alpha_1} \dots Q_n^{\alpha_n}$ *est la DFI de* Q*, alors la décomposition peut s'écrire*

$$\frac{P}{Q} = E + \sum_{i=1}^{n} \left(\sum_{j=1}^{\alpha_i} \frac{P_{i,j}}{Q_i^j} \right)$$

Autrement dit, on peut restreindre la somme aux éléments simples dont le dénominateur est de la forme Q_i^j *avec* $1 \leqslant j \leqslant \alpha_i$*.*

Démonstration. Montrons d'abord le caractère générateur. Si Q est constant il n'y a rien à faire. Sinon, tout revient à montrer que si $\deg(P) < \deg(Q)$ alors on peut écrire

$$\frac{P}{Q} = \sum_{i=1}^{n} (\sum_{j=1}^{\alpha_i} \frac{P_{i,j}}{Q_i^j})$$

Pour chaque i, considérons le polynôme $R_i = \prod_{\substack{k=1 \\ k \neq i}}^{n} Q_k^{\alpha_k}$.

Si un polynôme Δ divise tous les R_i, alors il divise *a fortiori* Q donc sa DFI s'écrit $\mu Q_1^{\beta_1} ... Q_n^{\beta_n}$. Par ailleurs, pour i fixé, comme Δ divise R_i on doit avoir $\beta_i \leq 0$ Ainsi, Δ est constant. On en déduit que les R_i sont premiers entre eux dans leur ensemble, et le théorème de Bézout généralisé permet d'écrire

$$U_1 R_1 + \ldots + U_n R_n = 1 \text{ puis} \frac{P}{Q} = \sum_{i=1}^{n} \frac{P U_i}{Q_i^{\alpha_i}}$$

On est donc ramené au cas d'une fraction rationnelle $\frac{P}{Q^\alpha}$ avec Q irréductible. Quitte à prendre à nouveau la partie entière, on peut supposer que $\deg(P) < \deg(Q^\alpha)$. On effectue alors une récurrence sur α.

- Initialisation : pour $\alpha = 1$, il n'y a rien à faire
- Hérédité : soit $\alpha \geq 2$ supposons la propriété vraie au rang $\alpha - 1$ et montrons-la au rang α. Soit donc P tel que $\deg(P) < \deg(Q)$, et écrivons la division euclidienne de P par $Q : P = BQ + R$ avec $\deg(R) < \deg(Q)$. Alors $\frac{P}{Q^\alpha} = \frac{B}{Q^{\alpha-1}} + \frac{R}{Q^\alpha}$ Comme

  $$\deg(B) = \deg(P - R) - deg(Q) \leqslant \deg(P) - deg(Q) < deg(Q^{\alpha-1})$$

  il n'y a plus qu'à invoquer l'hypothèse de récurrence pour décomposer $\frac{B}{Q^{\alpha-1}}$.

Passons maintenant à la liberté, et montrons que si

$$\sum_{i=1}^{n} \left(\sum_{j=1}^{a_i} \frac{P_{ij}}{Q_i^j} \right) = 0$$

alors tous les $P_{i,j}$ sont nuls. Supposons que pour un certain i_0, les $P_{i_0,j}$ soient non tous nuls, et soit j_0 maximal tel que $P_{i_0,j_0} \neq 0$. On multiplie alors l'égalité par $Q_{i_0}^{j_0} \prod_{\substack{k=1 \\ k\neq i_0}}^{n} Q_k^{\alpha_k}$

- Chaque $\dfrac{P_{i,j}}{Q_i^j}$ pour $i \neq i_0$ devient un polynôme multiple de $Q_{i_0}^{j_0}$ donc de Q_{i_0}.
- Quant aux termes de la forme $\dfrac{P_{i_0,j}}{Q_{i_0}^j}$:
  - si $j > j_0$ ils sont nuls ;
  - si $j \leqslant j_0$ ils deviennent des polynômes multiples de $P_{i_0,j}Q_{i_0}^{j_0-j}$. En particulier, si $j < j_0$ ils sont encore multiples de Q_{i_0}.

Finalement, modulo Q_{i_0} il ne reste plus que $P_{i_0,j_0} \prod_{\substack{k=1 \\ k\neq i_0}}^{n} Q_k^{\alpha_k}$. Or Q_{i_0} est premier avec chacun des Q_k pour $k \neq i_0$ donc d'après le corollaire du théorème de Bézout il est premier avec $\prod_{\substack{k=1 \\ k\neq i_0}}^{n} Q_k^{\alpha_k}$. Par le théorème de Gauss il divise donc P_{i_0,j_0} et comme $\deg(P_{i_0,j_0}) < \deg(Q_{i_0})$ on a finalement $P_{i_0,j_0} = 0$. Contradiction. □

5.2 Plongement de $\mathbb{R}(X)$ dans $\mathbb{C}(X)$

Régulièrement, nous souhaitons interpréter une fraction rationnelle de $\mathbb{R}(X)$ comme une fraction rationnelle de $\mathbb{C}(X)$ à coefficients réels. Mais tout cela a-t-il un sens ? Afin de ne pas nous y perdre, pour noter la classe d'équivalence de $(P,Q) \in \mathbb{K}[X] \times (\mathbb{K}[X]\backslash\{0\})$ nous allons momentanément abandonner la notation $\dfrac{P}{Q}$ au profit de $\dfrac{P}{Q}_{|\mathbb{K}}$.

La question est alors de savoir si, lorsque P et Q sont des polynômes réels, on peut identifier $\dfrac{P}{Q}_{|\mathbb{R}}$ et $\dfrac{P}{Q}_{|\mathbb{C}}$. On considère donc

$$\begin{array}{rccc} \varphi & : & \mathbb{R}(X) & \to & \mathbb{C}(X) \\ & & \dfrac{P}{Q}_{|\mathbb{R}} & \mapsto & \dfrac{P}{Q}_{|\mathbb{C}} \end{array}$$

- Premier point : si $\dfrac{P_1}{Q_1}_{|\mathbb{R}} = \dfrac{P_2}{Q_2}_{|\mathbb{R}}$ alors $P_1Q_2 = P_2Q_1$ dans $\mathbb{R}[X]$. Mais comme les lois de $\mathbb{C}[X]$ prolongent celles de $\mathbb{R}[X]$, on a aussi $P_1Q_2 = P_2Q_1$ dans $\mathbb{C}[X]$. Ainsi, on a $\dfrac{P_1}{Q_1}_{|\mathbb{C}} = \dfrac{P_2}{Q_2}_{|\mathbb{C}}$ et φ est bien définie.
- Deuxième point : on a $\varphi\left(\dfrac{P}{Q}_{|\mathbb{R}} + \dfrac{R}{S}_{|\mathbb{R}}\right) = \varphi\left(\dfrac{PS+QR}{QS}_{|\mathbb{R}}\right) = \dfrac{PS+QR}{QS}_{|\mathbb{C}}$. De même, puisque les lois de $\mathbb{C}[X]$ prolongent celles de $\mathbb{R}[X]$, on $\dfrac{PS+QR}{QS}_{|\mathbb{C}} = \dfrac{P}{Q}_{|\mathbb{C}} + \dfrac{R}{S}_{|\mathbb{C}}$ donc φ est un morphisme pour l'addition.

- De la même manière, on montre que φ est un morphisme pour la multiplication. Par ailleurs, on a bien sûr $\varphi(1_{|\mathbb{R}}) = 1_{|\mathbb{C}}$.
- Dernier point : en tant que morphisme de corps, φ est injectif.

Le dernier point montre que $\mathbb{R}(X)$ peut légitimement être considéré comme une sous-partie de $\mathbb{C}(X)$ puis le fait que φ soit morphisme montre que dans ce cas, les définitions des lois (cas réel et cas complexe) coïncident bien.

5.3 Division selon les puissances croissantes

Dans le cas d'un pôle multiple, les calculs de décomposition en éléments simples peuvent être rapidement compliqués. Voici une technique très puissante qui permet d'obtenir d'un coup l'ensemble des coefficients associés à un pôle.

Théorème 14.8 (Division selon les puissances croissantes)**.** *Soit A et B deux polynômes de $\mathbb{K}[X]$. On suppose que le terme constant de B est non nul. Alors*

$$\forall p \in \mathbb{N},\ \exists!(Q,R) \in \mathbb{K}[X]^2,\ \begin{cases} A = BQ + X^{p+1}R \\ \deg(Q) \leq p \end{cases}$$

Démonstration. Commençons par montrer l'unicité. Si (Q,R) et (Q',R') conviennent, alors $B(Q-Q') = X^{p+1}(R'-R)$. Comme 0 n'est pas racine de B, celui-ci est premier avec X^{p+1}, qui divise donc $Q-Q'$ d'après le lemme de Gauss. Comme $\deg(Q-Q') \leqslant p$, on a $Q = Q'$ donc $R = R'$.

Pour l'existence à présent, on la prouve par récurrence sur $p \in \mathbb{N}$.

- Initialisation : pour $p = 0$, notons λ le coefficient constant de A et $\mu \neq 0$ celui de B. Il suffit alors de choisir $Q = \dfrac{\lambda}{\mu}$. Ainsi $A - BQ$ n'a pas de coefficient constant et on a bien $A - BQ = XR$.
- Hérédité : soit $p \geqslant 1$ supposons la propriété vraie au rang $p-1$ et écrivons par hypothèse de récurrence
$$\begin{cases} A = BQ + X^p R \\ \deg(Q) \leq p-1 \end{cases}$$
Notons toujours $\mu \neq 0$ coefficient constant de B : $B = \mu + XB_2$. Notons également ν le coefficient constant de R : $R = \nu + XR_2$. On a alors
$$\begin{aligned} A &= BQ + \nu X^p + X^{p+1}R_2 \\ &= B\left(Q + \frac{\nu}{\mu}X^p\right) + X^{p+1}\left(R_2 - \frac{\nu}{\mu}B_2\right) \end{aligned}$$
avec $\deg\left(Q + \dfrac{\nu}{\mu}X^p\right) \leqslant p$.

Ainsi, la récurrence est achevée. □

Exemple 14.1 (Pratique de la division selon les puissances croissantes)**.** Effectuons la division de X^3+X par X^2+X+1 à l'ordre $p=4$. On écrit les polynômes dans l'ordre des puissances croissantes. À part cela, la technique la même que pour la division euclidienne.

$$\begin{array}{cccccc|ccc}
X & & +X^3 & & & & 1 & +X & +X^2 \\
-X & -X^2 & -X^3 & & & & \multicolumn{3}{c}{\cline{1-3}} \\
\multicolumn{3}{c}{\cline{1-3}} & & & & X & -X^2 & +X^3 \\
 & -X^2 & & & & & & & \\
 & \multicolumn{3}{c}{\cline{2-4}} & & & & & \\
 & X^2 & +X^3 & +X^4 & & & & & \\
 & \multicolumn{3}{c}{\cline{2-4}} & & & & & \\
 & & X^3 & +X^4 & & & & & \\
 & & -X^3 & -X^4 & -X^5 & & & & \\
 & & \multicolumn{3}{c}{\cline{3-5}} & & & & \\
 & & & & -X^5 & & & &
\end{array}$$

On a donc $X^3+X=(X^2+X+1)(X^3-X^2+X)-X^5$.

On en déduit une technique très efficace de décomposition en éléments simples, exposée ci-contre.

Méthode 14.8 (Décomposition en éléments simples pour un pôle multiple)**.** Voici une méthode de décomposition en éléments simples pour un pôle multiple qui utilise la division selon les puissances croissantes.

- On note toujours $\dfrac{P}{Q}$ notre fraction rationnelle, avec $P\wedge Q=1$ et Q unitaire.
- Quitte à effectuer une division euclidienne, on suppose que $\deg(P)<\deg(Q)$ (cela permet d'alléger les calculs).
- Soit λ un pôle d'ordre p. On a donc $Q=(X-\lambda)^pQ_p$ avec $Q_p(\lambda)\neq 0$. On effectue le changement de variable $X=Y+\lambda$. On a alors
$$F=\frac{P(Y+\lambda)}{Y^pQ_p(Y+\lambda)}=\frac{\tilde{P}(Y)}{Y^p\tilde{Q}_p(Y)}$$
- La division selon les puissances croissantes de $\overline{P}$ par $\tilde{Q}_p$ à l'ordre $p-1$ donne
$$\tilde{P}(Y)=\tilde{Q}_p(Y)S(Y)+Y^pR(Y)$$
avec $\deg(S)<p$, ce qui permet d'écrire $S(Y)=\displaystyle\sum_{i=0}^{p-1}s_iY^i$.
- On écrit alors $F=\dfrac{S(Y)}{Y^p}+\dfrac{R(Y)}{\tilde{Q}_p(Y)}=\displaystyle\sum_{i=0}^{p-1}\frac{s_i}{Y^{p-i}}+\frac{R(Y)}{\tilde{Q}_p(Y)}$.
- On revient à $Y=X-\lambda$: on a décomposé le pôle λ.

Démonstration. Justifions cette méthode.

- Lorsqu'on parle de "changement de variable", tout se passe comme si Y devenait la nouvelle inconnue à la place de X. En fait, en toute rigueur il faut plutôt considérer qu'on introduit un polynôme $Y=X-\lambda\in\mathbb{K}[X]$.
- La formule du binôme montre que $P(Y+\lambda)$ peut s'écrire $\tilde{P}(Y)$ et que $Q_p(Y+\lambda)$ peut s'écrire $\tilde{Q}_p(Y)$. De plus, cette même formule du binôme montre que le terme constant de $\tilde{Q}_p$ est égal à $Q_p(\lambda)\neq 0$. On peut donc bien effectuer une division selon les puissances croissantes.

- La méthode donne l'impression qu'on effectue la division selon les puissances croissantes avec l'inconnue Y. En fait, en toute rigueur il faut d'abord l'effectuer avec X :
$$\tilde{P} = \tilde{Q_p}S + X^pR$$
puis dans un second temps, on évalue cette égalité polynomiale en Y. Cela dit, il est vrai que formellement tout se passe comme si on travaillait avec Y dès le départ.
- Enfin, il faut justifier que les termes $\displaystyle\sum_{i=0}^{p-1} \frac{s_i}{Y^{p-i}}$ correspondent bien à la décomposition du pôle λ. Notons qu'on a
$$F = \sum_{i=0}^{p-1} \frac{s_i}{(X-\lambda)^{p-i}} + \frac{R(X-\lambda)}{\tilde{Q_p}(Y+\lambda)} = \sum_{i=0}^{p-1} \frac{s_i}{(X-\lambda)^{p-i}} + \frac{R(X-\lambda)}{Q_p}$$
$R(X-\lambda)$ est un polynôme de $\mathbb{K}[X]$ et Q_p est un polynôme dont les facteurs irréductibles sont exactement ceux de Q, sauf $X-\lambda$. En effectuant la décomposition en éléments simples de la fraction rationnelle $\dfrac{R(X-\lambda)}{Q_p}$ et en lui ajoutant $\displaystyle\sum_{i=0}^{p-1} \frac{s_i}{(X-\lambda)^{p-i}}$, on conclut par unicité.

□

Exemple 14.2 (Décomposition en éléments simples de $\dfrac{X^2+1}{(X-1)^3X^2}$)). Ce n'est pas la peine d'effectuer une division euclidienne, le numérateur est déjà de degré strictement inférieur à celui du dénominateur. On commence par le pôle 0 d'ordre 0, pas besoin d'effectuer un changement de variable. La division selon les puissances croissantes de X^2+1 par $(X-1)^3 = X^3-3X^2+3X-1$ donne $X^2+1 = (X-1)^3(-3X-1) + X^2(3X^2-8X+7)$, d'où
$$F = -\frac{1}{X^2} - \frac{3}{X} + \frac{3X^2-8X+7}{(X-1)^3}$$
Le changement de variable $Y = X-1$ donne
$$\frac{3X^2-8X+7}{(X-1)^3} = \frac{3(Y+1)^2-8(Y+1)+7}{Y^3} = \frac{3Y^2-2Y+2}{Y^3} = \frac{2}{Y^3} - \frac{2}{Y^2} + \frac{3}{Y}$$
Finalement
$$\frac{X^2+1}{(X-1)^3X^2} = -\frac{1}{X^2} - \frac{3}{X} + \frac{2}{(X-1)^3} - \frac{2}{(X-1)^2} + \frac{3}{X-1}$$

Exemple 14.3 (Décomposition de $\dfrac{3X^4-4X^3-2X^2+4X+3}{(X^2+1)^2(X-1)}$)). Comme 1 est un pôle simple, son coefficient est donné par $\dfrac{P(1)}{Q_2(1)} = \dfrac{4}{2^2} = 1$. Pour $(X^2+1)^2 = (X-i)^2(X+i)^2$; il suffit de calculer les coefficients de $(X-i)(X+i)^2$, les autres seront conjugués. Posons donc $Y = X-i$, ce qui donne
$$\begin{aligned} F &= \frac{3(Y+i)^4-4(Y+i)^3-2(Y+i)^2+4(Y+i)+3}{Y^2(Y+2i)^2(Y+i-1)} \\ &= \frac{3Y^4+(-4+12i)Y^3-(20+12i)Y^2+(16-16i)Y+(8+8i)}{Y^2(Y^3+(-1+5i)Y^2+(-8-4i)Y+(4-4i))} \end{aligned}$$

et effectuons notre division selon les puissances croissantes.

$$\begin{aligned}3Y^4+(-4+12i)Y^3-(20+12i)Y^2+(16-16i)Y+(8+8i)&\\ =(Y^3+(-1+5i)Y^2+(-8-4i)Y+(4-4i))(2i+(1+i)Y)&\\ +Y^2((-6+2i)+(2+6i)Y+(2-i)Y^2)&\end{aligned}$$

On obtient

$$F=\frac{2i}{Y^2}+\frac{1+i}{Y}+\dots$$

et finalement

$$\begin{aligned}F&=\frac{1}{X-1}+\frac{1+i}{X-i}+\frac{2i}{(X-i)^2}+\frac{1-i}{X+i}+\frac{-2i}{(X+i)^2}\\ &=\frac{1}{X-1}+\frac{2X-2}{X^2+1}-\frac{8X}{(X^2+1)^2}\end{aligned}$$

Chapitre 15

Espaces euclidiens

Dans tout le chapitre, les corps de base des espaces vectoriels est expressément fixé à $\mathbb{K} = \mathbb{R}$.

1 Produit scalaire

Définition 15.1 (Produit scalaire)**.** Soit E un espace vectoriel réel. Un **produit scalaire** sur E est une application $\langle \cdot, \cdot \rangle : E^2 \to \mathbb{R}$ qui vérifie les 4 axiomes suivants :

1. Bilinéarité : Pour tout $x \in E$, $y \mapsto \langle x, y \rangle$ est linéaire et pour tout $y \in E$, $x \mapsto \langle x, y \rangle$ est linéaire
2. Symétrie : $\forall (x, y) \in E^2,\ \langle x, y \rangle = \langle y, x \rangle$
3. Caractère positif : $\forall x \in E,\ \langle x, x \rangle \geqslant 0$
4. Caractère défini : $\forall x \in E,\ \langle x, x \rangle = 0 \Longrightarrow x = 0_E$

On résume cela en disant qu'un produit scalaire est une **forme bilinéaire symétrique définie positive**. Le produit scalaire peut aussi être noté $< x|y >$, $(x|y)$ voire $x \cdot y$.

Remarque 15.1. Si on prouve la symétrie d'abord, il suffit de prouver la linéarité par rapport à l'une des variables pour prouver la bilinéarité.

Définition 15.2 (Espace préhilbertien réel, espace euclidien)**.** Un **espace préhilbertien réel** est un $\mathbb{R}$-espace vectoriel muni d'un produit scalaire. Si cet espace vectoriel est de dimension finie sur $\mathbb{R}$, on dit alors qu'il s'agit d'un **espace euclidien**.

Remarque 15.2. Si E est préhilbertien, tout sev de E est naturellement muni d'une structure préhilbertienne en restreignant le produit scalaire. De même, tout sev d'un espace euclidien est naturellement muni d'une structure euclidienne.

Définition 15.3 (Produit scalaire canonique de $\mathbb{R}^n$)**.** On définit le **produit scalaire canonique de** $\mathbb{R}^n$ par

$$\left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\rangle = \sum_{i=1}^{n} x_i y_i$$

Matriciellement, on peut écrire

$$\langle X, Y \rangle = X^\top Y$$

Définition 15.4 (Produit scalaire canonique de $\mathcal{M}_{n,p}(\mathbb{R})$)**.** $\mathcal{M}_{n,p}(\mathbb{R})$ est euclidien pour son produit scalaire canonique

$$\langle A, B\rangle = \mathrm{Tr}(A^\top B)$$

Proposition 15.1 (Un produit scalaire sur les polynômes)**.** *Dans* $\mathbb{R}[X]$*, l'application*

$$\left\langle \sum_{k\in\mathbb{N}} a_k X^k, \sum_{k\in\mathbb{N}} b_k X^k \right\rangle = \sum_{k\in\mathbb{N}} a_k b_k$$

est un produit scalaire qui munit $\mathbb{R}[X]$ *d'un structure préhilbertienne.*

Théorème 15.1 (Produit scalaire sur les fonctions continues)**.** *Soit* $a < b$ *des réels. Dans le* $\mathbb{R}$*-espace vectoriel* $\mathcal{C}^0([a,b],\mathbb{R})$*, l'application définie par*

$$\langle f, g\rangle = \int_{[a,b]} fg$$

est un produit scalaire.

Définition 15.5 (Norme euclidienne associée à un produit scalaire)**.** La **norme euclidienne** associée au produit scalaire $\langle\cdot,\cdot\rangle$ est définie par

$$\|x\| = \sqrt{\langle x, x\rangle}$$

Elle vérifie $\|\lambda.x\| = |\lambda| \times \|x\|$ et $\|x\| = 0 \iff x = 0_E$

Définition 15.6. (Distance euclidienne associée à un produit scalaire) La **distance euclidienne** associée au produit scalaire $\langle\cdot,\cdot\rangle$ est définie par

$$d(x,y) = \|x - y\|$$

où $\|\cdot\|$ est la norme euclidienne associée au produit scalaire $\langle\cdot,\cdot\rangle$. Elle vérifie la propriété de symétrie

$$d(x,y) = d(y,x)$$

Proposition 15.2 (Identités remarquables)**.** *Soit* $(E, \langle\cdot,\cdot\rangle)$ *un espace préhilbertien réel. On a les* ***identités remarquables suivantes*** *:*

1. $\forall (x,y) \in E^2,\ \|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x, y\rangle$
2. $\forall (x,y) \in E^2,\ \|x-y\|^2 = \|x\|^2 + \|y\|^2 - 2\langle x, y\rangle$
3. $\forall (x,y) \in E^2,\ \langle x+y, x-y\rangle = \|x\|^2 - \|y\|^2$

Proposition 15.3 (Identité du parallélogramme)**.** *Soit* $(E, \langle\cdot,\cdot\rangle)$ *un espace préhilbertien réel. On a*

$$\forall (x,y) \in E^2,\ \|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2$$

Corollaire 15.1 (Identités de polarisation)**.** *Soit* $(E, \langle\cdot,\cdot\rangle)$ *un espace préhilbertien réel et* $\|\cdot\|$ *la norme euclidienne associée. on a alors les identités de polarisation :*

1. $\langle x, y\rangle = \dfrac{1}{2}\left(\|x+y\|^2 - \|x\|^2 - \|y\|^2\right)$

2. $\langle x,y\rangle = \dfrac{1}{2}\left(\|x\|^2+\|y\|^2-\|x-y\|^2\right)$
3. $\langle x,y\rangle = \dfrac{1}{4}\left(\|x+y\|^2-\|x-y\|^2\right)$

Corollaire 15.2. *Une norme euclidienne est associée à un unique produit scalaire.*

Définition 15.7 (Vecteur unitaire)**.** Un vecteur $x \in E$ est dit **unitaire** lorsque $\|x\| = 1$.

Définition 15.8 (Normalisation)**.** Lorsque $x \neq 0_E$, le vecteur

$$\frac{x}{\|x\|} = \|x\|^{-1}.x$$

est unitaire. On l'appelle le **renormalisé** de x.

Remarque 15.3. Plus généralement, pour $x \in E$ et $\lambda \in \mathbb{R}^\star$, on s'autorisera à noter $\dfrac{x}{\lambda}$ à la place de $\lambda^{-1}.x$

Théorème 15.2 (Inégalité de Cauchy-Schwarz)**.** *Soit $(E,\langle\cdot,\cdot\rangle)$ un espace préhilbertien réel. On a*

$$\forall (x,y)\in E^2,\ |\langle x,y\rangle| \leqslant \|x\|\|y\|$$

avec égalité si, et seulement si, x et y sont colinéaires.

Remarque 15.4. L'inégalité reste vraie sans le caractère défini (mais pas le cas d'égalité). On pourra donc appliquer l'inégalité dans le cas de formes bilinéaires symétriques positives.

Corollaire 15.3 (Inégalité triangulaire)**.** *La norme euclidienne associée à $\langle\cdot,\cdot\rangle$ vérifie l'inégalité triangulaire :*

$$\forall (x,y)\in E^2,\ \|x+y\| \leqslant \|x\|+\|y\|$$

avec égalité si, et seulement si, x et y sont colinéaires directs.

Corollaire 15.4 (Seconde inégalité triangulaire)**.** *On en déduit la second inégalité triangulaire :*

$$\forall (x,y)\in E^2,\ \big|\|x\|-\|y\|\big| \leqslant \|x-y\|$$

Définition 15.9 (Norme)**.** Une **norme** sur un espace vectoriel E est une application $\|\cdot\| : E \to \mathbb{R}_+$ qui vérifie les axiomes suivants :

1. Séparation : $\forall x\in E,\ \|x\| = 0 \Rightarrow x = 0$
2. Homogénéité : $\forall x\in E,\ \forall \lambda\in\mathbb{R},\ \|\lambda.x\| = |\lambda|\times\|x\|$
3. Inégalité triangulaire : $\forall (x,y)\in E^2,\ \|x+y\| \leqslant \|x\|+\|y\|$

Remarque 15.5. Toutes les normes ne dérivent pas d'un produit scalaire. Pour ces normes, dites non euclidiennes et étudiées en deuxième année, on ne peut plus rien dire sur le cas d'égalité de l'inégalité triangulaire.

2 Orthogonalité

Dans toute cette partie, $(E,\langle\cdot,\cdot\rangle)$ est un espace préhilbertien réel.

2.1 Définitions de base

Définition 15.10 (Orthogonalité)**.** Deux vecteurs x et y sont dits **orthogonaux** lorsque $\langle x, y\rangle = 0$. Dans ce cas, on écrit $x \perp y$. Un vecteur x est dit **orthogonal** à une partie B lorsqu'il est orthogonal à tout vecteur de B :

$$\forall y \in B,\ \langle x, y\rangle = 0$$

Dans ce cas, on écrit $x \perp B$. Deux parties A et B sont dites **orthogonales** lorsque tout vecteur de l'une est orthogonal à tout vecteur de l'autre :

$$\forall (x, y) \in A \times B,\ \langle x, y\rangle = 0$$

Dans ce cas, on écrit $A \perp B$.

Définition 15.11 (Orthogonal d'une partie)**.** Soit A une partie de E. L'ensemble des vecteurs orthogonaux à tous les vecteurs de A est un sous-espace vectoriel de E, appelé **orthogonal** de A. On le note

$$A^\perp = \{x \in E \mid \forall a \in A,\ \langle x, a\rangle = 0\}$$

C'est aussi la plus grande partie orthogonale à A au sens de l'inclusion.

Exemple 15.1. On a $E^\perp = \{0\}$ et $\{0\}^\perp = E$.

Proposition 15.4 (Décroissance de l'othogonal)**.** *Si $A \subset B$ alors $B^\perp \subset A^\perp$.*

Corollaire 15.5. *Soit A une partie de E. Alors $A^\perp = \mathrm{Vect}(A)^\perp$.*

2.2 Familles orthogonales

Définition 15.12 (Famille orthogonale, famille orthonormale)**.** Une famille $(x_i)_{i\in I}$ de vecteurs de E est dite **orthogonale** si les x_i sont deux à deux orthogonaux. Si de plus les x_i sont tous unitaires, la famille est dite **orthonormale** (ou **orthonormée**).

Exemple 15.2. Dans $\mathbb{R}^n$ muni de son produit scalaire canonique, la base canonique est orthonormale.

Exemple 15.3. Dans $\mathbb{R}[X]$ muni de son produit scalaire usuel, la base canonique est orthonormale.

Proposition 15.5. *Toute sous-famille d'une famille orthogonale est orthogonale. Toute sous-famille d'une famille orthonormale est orthonormale.*

Proposition 15.6. *Toute famille orthogonale de vecteurs* ***non nuls*** *est libre.*

Corollaire 15.6. *Toute famille orthonormale est libre.*

Théorème 15.3 (Théorème de Pythagore)**.** *Deux vecteurs x et y sont orthogonaux si, et seulement si,*

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2$$

Théorème 15.4. *Soit $(x_i)_{1\leqslant i\leqslant n}$ une famille orthogonale finie. Alors*

$$\left\|\sum_{i=1}^{n} x_i\right\|^2 = \sum_{i=1}^{n} \|x_i\|^2$$

Remarque 15.6. Attention, la réciproque est fausse dès que $n \geqslant 3$. On pourra chercher un contre-exemple dans $\mathbb{R}^3$ muni de son produit scalaire canonique.

Théorème 15.5 (Existence d'une base orthonormale pour un espace euclidien)**.** *Tout espace euclidien admet une base orthonormale.*

Proposition 15.7 (Expression dans une base orthonormale)**.** *Soit $(e_i)_{1\leqslant i\leqslant n}$ une base orthonormale de E euclidien. Alors tout vecteur s'exprime dans cette base par :*

$$x = \sum_{i=1}^{n} \langle x, e_i \rangle e_i$$

De plus, si on écrit $x = \sum_{i=1}^{n} x_i e_i$ et $y = \sum_{i=1}^{n} y_i e_i$, on a :

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i$$

et

$$\|x\|^2 = \sum_{i=1}^{n} x_i^2$$

2.3 Supplémentaire orthogonal

Proposition 15.8. *L'orthogonalité est une condition suffisante (mais pas nécessaire !) pour que deux sous-espaces vectoriels F et G soient en somme directe.*

Définition 15.13 (Supplémentaires orthogonaux)**.** Soit F et G deux sous-espaces vectoriels à la fois supplémentaires et orthogonaux. On dit alors que F et G sont **supplémentaires orthogonaux** et on écrit

$$E = F \overset{\perp}{\oplus} G$$

On dit que G est un supplémentaire orthogonal F, et vice-versa.

Théorème 15.6 (Théorème de la base adaptée, version orthonormale)**.** *Soit $(e_i)_{i\in I}$ une base orthonormale de E et supposons que $I = I_1 \sqcup I_2$. Alors $E_1 = \mathrm{Vect}(e_i)_{i\in I_1}$ et $E_2 = \mathrm{Vect}(e_i)_{i\in I_2}$ sont supplémentaires orthogonaux. Réciproquement, si E_1 et E_2 sont supplémentaires orthogonaux, alors en concaténant des bases orthonormales de E_1 et E_2, on obtient une base orthonormale de E.*

Proposition 15.9. *Si G est un supplémentaire orthogonal de F, alors on a nécessairement $G = F^\perp$. S'il existe, le supplémentaire orthogonal est donc unique.*

Corollaire 15.7. *Si F admet un supplémentaire orthogonal, alors $F^\perp$ admet un supplémentaire orthogonal et on a $(F^\perp)^\perp = F$.*

Définition 15.14. Si F admet un supplémentaire orthogonal G, le **projecteur orthogonal** est le projecteur p sur F parallèlement à G. $p(x)$ s'appelle le **projeté orthogonal** de x sur F.

Remarque 15.7. Le projecteur associé $q = \mathrm{id} - p$ est donc le projecteur orthogonal sur $F^\perp$.

Définition 15.15 (Symétrie orthogonale)**.** Si F admet un supplémentaire orthogonal G, la **symétrie orthogonale d'axe** F est la symétrie par rapport à F parallèlement à G. Si on l'appelle s, on rappelle qu'on a

$$s = 2p - \mathrm{id}$$

Définition 15.16 (Réflexion)**.** Une symétrie orthogonale par rapport à un hyperplan s'appelle une **réflexion**.

Proposition 15.10 (Caractérisation de l'orthogonalité pour les projecteurs et les symétries)**.** *Un projecteur p est orthogonal si, et seulement si,* $\ker(p)$ *et* $\mathrm{Im}(p)$ *sont orthogonaux. Une symétrie s est orthogonale si, et seulement si,* $\ker(s - \mathrm{id})$ *et* $\ker(s + \mathrm{id})$ *sont orthogonaux.*

Définition 15.17. Soit F un sous-espace vectoriel de E et $x \in E$. La **distance** de x à F est définie par

$$d(x, F) = \inf_{y \in F} \|x - y\|$$

Elle et bien définie car l'ensemble est non vide car contient $\|x - 0\|$ et minoré par 0 en tant qu'ensemble de normes euclidiennes.

Théorème 15.7. *Supposons que F admette un supplémentaire orthogonal, et soit p le projecteur orthogonal sur F. Alors, pour tout $x \in E$, $d(x, F)$ est un minimum et $p(x)$ est l'unique élément de F qui le réalise.*

Théorème 15.8 (Synthèse partielle - cas fondamental)**.** *Si F est de dimension finie, alors on a $E = F \overset{\perp}{\oplus} F^\perp$. Dans ce cas, $F^\perp$ admet un supplémentaire orthogonal égal à F et on pourra toujours définir le projecteur orthogonal sur F et la symétrie orthogonale par rapport à F.*

Remarque 15.8. Attention on a des contre-exemples en dimension infinie. On pourra considérer $\mathrm{Vect}(X^n - 1)_{n \in \mathbb{N}^\star}$ dans $\mathbb{R}[X]$ muni de son produit scalaire usuel, dont l'orthogonal vaut $\{0\}$. Sinon, on pourra considérer l'ensemble

$$H = \left\{ f \in \mathcal{C}^0([0,1], \mathbb{R}) \mid f(0) = 0 \right\}$$

dans $\mathcal{C}^0([0,1], \mathbb{R})$ muni de son produit scalaire usuel, qui est un hyperplan car le noyau d'un morphisme d'évaluation non-nul. Son orthogonal vaut alors $\{0\}$ (le prouver en prenant $f \in H^\perp$ puis en effectuant le produit scalaire avec $t \mapsto tf(t) \in H$ et conclure par signe et intégrale nulle pour les points autre que 0, et par continuité en 0).

Corollaire 15.8. *Si E est euclidien et F est un sous-espace vectoriel de E, on a toujours*

$$\dim(F^\perp) = \dim(E) - \dim(F)$$

Corollaire 15.9 (Expression du projecteur orthogonal en dimension finie)**.** *Si F est de dimension finie et $(e_i)_{1 \leqslant i \leqslant k}$ est une base orthonormale de F, le projecteur orthogonal sur F est donné par*

$$\forall x \in E,\ p(x) = \sum_{i=1}^{k} \langle x, e_i \rangle e_i$$

Remarque 15.9. Si on ne dispose pas d'une base orthonormale, mais simplement d'une base $(e_i)_{1 \leqslant i \leqslant k}$, on peut soit appliquer l'algorithme de Gram-Schmidt expliqué après, soit raisonner de la manière suivante.

On sait qu'il existe des scalaires $(\lambda_1, \dots, \lambda_k) \in \mathbb{R}^k$ tels que $p(x) = \sum_{j=1}^{k} \lambda_j e_j$. Ensuite, on sait que $x - p(x) \in F^\perp$, donc on obtient

$$\forall i \in [\![1, k]\!],\ \langle x, e_i \rangle = \sum_{j=1}^{k} \lambda_j \langle e_j, e_i \rangle$$

On obtient ainsi un système linéaire de k équations à k inconnues (dont on peut démontrer qu'il est de Cramer) qu'on peut alors toujours résoudre pour obtenir les coefficients.

Proposition 15.11 (Théorème de la base incomplète, version orthonormale)**.** *Si E est euclidien, alors toute famille orthonormale de E peut être complétée en une base orthonormale.*

Théorème 15.9 (Orthonormalisation de Gram-Schmidt)**.** *Soit $(x_1, \dots, x_n)$ une famille libre de E. Alors, il existe une famille orthonormale $(\varepsilon_1, \dots, \varepsilon_n)$ telle que*

$$\forall k \in [\![1, n]\!],\ \mathrm{Vect}(x_1, \dots, x_k) = \mathrm{Vect}(\varepsilon_1, \dots, \varepsilon_k)$$

Remarque 15.10 (Mise en œuvre pratique du procédé d'orthonormalisation)**.** On construit la famille par récurrence finie sur k.
Initialisation : On pose $\varepsilon_1 = \dfrac{x_1}{\|x_1\|}$
Hérédité : Soit $k \in [\![2, n]\!]$ et supposons avoir construit $(\varepsilon_1, \dots, \varepsilon_{k-1})$ On pose alors

$$\tilde{\varepsilon}_k = x_k - \sum_{i=1}^{k-1} \langle x_k, \varepsilon_i \rangle \varepsilon_i$$

puis on pose alors $\varepsilon_k = \dfrac{\tilde{\varepsilon}_k}{\|\tilde{\varepsilon}_k\|}$

Remarque 15.11. Le théorème fonctionne encore avec une famille dénombrable, c'est-à-dire une famille de la forme $(x_k)_{k \in \mathbb{N}}$. On pensera notamment aux polynômes.

Remarque 15.12. On peut montrer par analyse-synthèse que la famille $(\varepsilon_1, \dots, \varepsilon_n)$ est la seule famille qui vérifie la condition supplémentaire :

$$\forall k \in [\![1, n]\!],\ \langle x_k, \varepsilon_k \rangle > 0$$

3 Compléments : introduction à la géométrie affine, HP

On se donne un espace euclidien E (mais on généraliserait sans problème à un $\mathbb{K}$-ev quelconque).

Jusqu'ici, nous avons considéré $\mathbb{R}^2$ ou $\mathbb{R}^3$ comme des espaces vectoriels. Leurs éléments sont des vecteurs, et intuitivement on les assimile à des déplacements. Mais en géométrie élémentaire, lorsqu'on "pose un repère", on sait que $\mathbb{R}^2$ et $\mathbb{R}^3$ peuvent être aussi vus comme des ensembles de points. Cette fois, leurs éléments sont plutôt assimilés à des positions fixes. Cela n'a pas tellement de sens d'additionner deux positions. En revanche, leur différence correspond à un déplacement, et

on peut lui associer un vecteur.

Pour bien faire la différence entre $\mathbb{R}^2$ espace de points et $\mathbb{R}^2$ espace de vecteurs, on peut noter les points en ligne : (x_1, x_2) et les vecteurs en colonne : $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ De même avec $\mathbb{R}^3$. Formalisons

Définition 15.18 (Espace affine)**.** Soit $\mathcal{E}$ un ensemble non vide. $\mathcal{E}$ est un **espace affine** de **direction** E lorsqu'il existe une application $(A, B) \in \mathcal{E}^2 \mapsto \overrightarrow{AB} \in E$ vérifiant les deux axiomes suivants.

- $\forall A \in \mathcal{E},\ \forall u \in E,\ \exists! B \in \mathcal{E},\ \overrightarrow{AB} = u$. On notera $B = A + u$.
- Relation de Chasles : $\forall (A, B, C) \in \mathcal{E}^3,\ \overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC}$

Remarque 15.13. Par souci de clarté, on essayera au maximum de garder les lettres majuscules pour les points de l'espace affine, et les lettres minuscules ou les flèches pour les vecteurs de l'espace vectoriel.

Exemple 15.4. $\mathbb{R}^n$ peut être considéré comme un espace affine : à deux n-uplets $(x_1, ..., x_n)$ et $(y_1, ..., y_n)$ on associe le vecteur $\begin{pmatrix} y_1 - x_1 \\ \vdots \\ y_n - x_n \end{pmatrix}$

Exemple 15.5. À deux points A et B de E, faisons correspondre le vecteur $B - A$. On vérifie immédiatement que E est un espace affine de direction lui-même. La notation $B = A + u$ reste parfaitement valable.

Définition 15.19 (Translation)**.** Si $u \in E$, la translation de vecteur u est l'application $t_u : A \mapsto A + u$.

Proposition 15.12. *On a les propriétés suivantes.*

- $\forall (A, B) \in \mathcal{E}^2,\ A = B \Leftrightarrow \overrightarrow{AB} = \overrightarrow{0}$
- $\forall (A, B) \in \mathcal{E}^2, \overrightarrow{BA} = -\overrightarrow{AB}$.

Démonstration. Soit A point de $\mathcal{E}$ fixé, et écrivons $\overrightarrow{AA} = \overrightarrow{AA} + \overrightarrow{AA}$, si bien que $\overrightarrow{AA} = \overrightarrow{0}$. Donc si $A = B$ alors $\overrightarrow{AB} = \overrightarrow{0}$. Réciproquement, si $\overrightarrow{AB} = \overrightarrow{0}$, alors comme $\overrightarrow{AA} = \overrightarrow{0}$, par ... on en déduit $A = B$ Pour le deuxième point, il suffit d'écrire $\overrightarrow{AB} + \overrightarrow{BA} = \overrightarrow{AA} = \overrightarrow{0}$. □

Définition 15.20 (Barycentre)**.** Soit $A_1, ..., A_n$ des points de E ($n \geqslant 1$), et $\alpha_1, ..., \alpha_n$ des scalaires de somme non nulle. Alors il existe un unique G tel que $\sum_i \alpha_i \overrightarrow{GA_i} = \overrightarrow{0}$. C'est le **barycentre** de $((A_1, \alpha_1), ..., (A_n, \alpha_n))$. De plus on a

$$\forall O \in \mathcal{E},\ \overrightarrow{OG} = \frac{\sum_i \alpha_i \overrightarrow{OA_i}}{\sum_i \alpha_i}.$$

Plus généralement, on définit de la même manière le barycentre de $(A_i, \alpha_i)_{i \in I}$ avec I fini.

Démonstration. Existence : utilisons la fin de l'énoncé avec $O = A_1$ et montrons que

$$G = A_1 + \frac{\sum_i \alpha_i \overrightarrow{A_1A_i}}{\sum_i \alpha_i}$$

convient. On a

$$\sum_i \alpha_i \overrightarrow{GA_i} = \sum_i \alpha_i(\overrightarrow{GA_1} + \overrightarrow{A_1A_i}) = \overrightarrow{GA_1}\sum_i \alpha_i + \sum_i \alpha_i \overrightarrow{A_1A_i} = (\sum_i \alpha_i)(\overrightarrow{GA_1} + \overrightarrow{A_1G}) = \overrightarrow{0}$$

Unicité : si G_1 et G_2 conviennent, alors en soustrayant $\sum_i \alpha_i \overrightarrow{G_1A_i} = \overrightarrow{0}$ et $\sum_i \alpha_i \overrightarrow{G_2A_i} = \overrightarrow{0}$ on obtient $\sum_i \alpha_i \overrightarrow{G_1G_2} = \overrightarrow{0}$ et on conclut en remarquant que $\sum_i \alpha_i \neq 0$. Enfin, si on se donne $O \in \mathcal{E}$ quelconque, alors par la relation de Chasles on a

$$\frac{\sum_i \alpha_i \overrightarrow{OA_i}}{\sum_i \alpha_i} = \frac{\sum_i \alpha_i(\overrightarrow{OG} + \overrightarrow{GA_i})}{\sum_i \alpha_i} = \overrightarrow{OG_i}$$

ce qui achève la preuve. □

Proposition 15.13 (Commutativité du barycentre)**.** *Pour toute permutation $\sigma \in \mathcal{S}_n$ le barycentre de $((A_1, \alpha_1), ..., (A_n, \alpha_n))$ est aussi le barycentre de $((A_{\sigma(1)}, \alpha_{\sigma(1)}), ..., (A_{\sigma(n)}, \alpha_{\sigma(n)}))$*

Démonstration. C'est une conséquence immédiate de la commutativité de la somme. □

Proposition 15.14 (Homogénéité du barycentre)**.** *Si λ est un scalaire non nul, alors le barycentre de $((A_1, \alpha_1), ..., (A_n, \alpha_n))$ est aussi le barycentre de $((A_1, \lambda\alpha_1), ..., (A_n, \lambda\alpha_n))$.*

Démonstration. Il suffit de multiplier l'égalité $\sum_i \alpha_i \overrightarrow{GA_i} = \overrightarrow{0}$ par λ. □

Proposition 15.15 (Associativité du barycentre)**.** *Soit $\{I_1, ..., I_p\}$ une partition de $[\![1, n]\!]$ telle que*

$$\forall r \in [\![1, p]\!],\ \sigma_r = \sum_{i \in I_r} \alpha_i \neq 0$$

Notons G_r le barycentre de $((A_1, \alpha_i))_{i \in I_r}$ (pour tout $r \in [\![1, p]\!]$). Alors le barycentre de $((A_1, \alpha_1), ..., (A_n, \alpha_n))$ est aussi barycentre de $((G_1, \sigma_1), ..., (G_p, \sigma_p))$.

Démonstration. Notons G le barycentre de $((A_1, \alpha_1), ..., (A_n, \alpha_n))$ et notons provisoirement G' le barycentre de $((G_1, \sigma_1), ..., (G_p, \sigma_p))$. Il suffit d'écrire

$$\begin{aligned}\sum_{i \in I} \alpha_i \overrightarrow{G'A_i} &= \sum_{r=1}^{p} \left(\sum_{i \in I_r} \alpha_i \overrightarrow{G'G_r} + \sum_{i \in I_r} \alpha_i \overrightarrow{G_rA_i} \right) \\ &= \sum_{r=1}^{p} \sigma_r \overrightarrow{G'G_r} \\ &= \overrightarrow{0}\end{aligned}$$

Donc par unicité on a $G = G'$ □

Définition 15.21 (Repère affine)**.** Un **repère affine** de E est la donnée d'un point O et d'une base $(e_i)_{1\leq i\leq n}$ de E. Le point O est l'**origine** du repère. n s'appelle la **dimension** de E. Le repère est **orthonormal** lorsque la base $(e_i)_{1\leq i\leq n}$ est orthonormale.

Proposition 15.16. *L'application*

$$\begin{array}{rccc} \varphi & : & \mathbb{R}^n & \longrightarrow & E \\ & & (x_1,...,x_n) & \longmapsto & M = O + \sum_i x_i e_i \end{array}$$

est une bijection.

Démonstration. Si $M = O + \sum_i x_i e_i = M' = O + \sum_i x'_i e_{i_i}$ alors $\sum_i x_i e_i = \sum_i x'_i e_i$, donc est injective. Pour la surjectivité, si $M \in E$ alors il existe $x_1, ..., x_n$ tels que $\sum_i x_i e_i = \overrightarrow{OM}$. □

Définition 15.22 (Coordonnées dans un repère affine)**.** Les scalaires $x_1, ..., x_n$ s'appellent les **coordonnées** de M dans le repère. En particulier, le point O a toutes ses coordonnées nulles.

Exemple 15.6. Si A a pour coordonnées $(x_1^A, ..., x_n^A)$ dans le repère et B a pour coordonnées $(x_1^B, ..., x_n^B)$ dans le repère, alors $\overrightarrow{AB}$ a pour coordonnées $\begin{pmatrix} x_1^B - x_1^A \\ \vdots \\ x_n^B - x_n^A \end{pmatrix}$ dans la base $(e_i)_{1\leq i\leq n}$.

Exemple 15.7. Si G est le barycentre de $((A_1, \alpha_1), ..., (A_p, \alpha_p))$ alors ses coordonnées sont les moyennes des coordonnées des A, pondérées par les α_i.

Exemple 15.8. Dans $\mathbb{R}^n$ vu comme un espace affine, on pose souvent $O = (0, ..., 0)$ et $(e_i)_{1\leq i\leq n}$ la base canonique (on l'appelle le repère canonique). Dans ce cas, si on se donne un élément $M - (x_1, ..., x_n) \in \mathbb{R}^n$, ses coordonnées dans le repère canonique sont exactement les x_i.

Définition 15.23. Un **sous-espace affine** (ou sea) de E est une partie de E de la forme $\mathcal{F} = A+F$ où A est un point de E et F un sev de E. F est lui-même un espace affine de direction F. A est appelé une origine de F.

Remarque 15.14. Lorsqu'on considère E comme un espace affine de direction lui-même, on retrouve bien la définition d'un sea vue précédemment. Il s'agit donc d'une généralisation.

Remarque 15.15. N'importe quel point de F peut servir d'origine. En revanche, la direction est bien fixée de manière unique. La démonstration est la même que dans le chapitre "Espaces vectoriels".

Définition 15.24 (Droite affine)**.** Si A et B sont deux points distincts, la **droite affine** (AB) est le sous-espace affine $A + \mathbb{R}\overrightarrow{AB}$. On vérifie que si C et D sont deux points distincts de (AB), alors $(CD) = (AB)$.

Définition 15.25. Si A et B sont deux points distincts, la **demi-droite** $[AB)$ est l'ensemble $A + \mathbb{R}_+\overrightarrow{AB}$ ce n'est plus un espace affine).

Soit $\mathcal{E}$ un espace affine de direction E. Une **application affine** de $\mathcal{E}$ est une application $f : \mathcal{E} \to \mathcal{E}$ telle qu'il existe une application linéaire $\varphi \in \mathcal{L}(E)$ vérifiant

$$\forall (A, B) \in \mathcal{E}^2, \overrightarrow{f(A)f(B)} = \varphi(\overrightarrow{AB})$$

On voit que si elle existe, φ est nécessairement unique. On l'appelle **partie linéaire** de f. À ce titre, on relira avec profit la remarque de la partie Géométrie du chapitre "Nombres complexes".

Proposition 15.17. *La composée de deux applications affines est une application affine.*

Démonstration. Soit f et g nos deux applications affines, avec φ et ψ associées. Alors

$$\forall (A, B) \in \mathcal{E}^2, \overrightarrow{f(g(A))f(g(B))} = \varphi(\overrightarrow{g(A)g(B)}) = \varphi(\psi(\overrightarrow{AB}))$$

Ainsi la preuve est achevée. □

Remarque 15.16. Au passage, on voit que lorsqu'on compose les applications affines, on compose les parties linéaires.

Chapitre 16

Fonctions usuelles

1 Révisions de lycée (et plus si affinités...)

1.1 Équations, inéquations

Pour les rédactions à adopter afin de résoudre des équations et des inéquations, on pourra se référer au chapitre "Logique et raisonnements". Pour les méthodes de résolution, nous n'en donnons ici que les noms : équivalences successives, analyse-synthèse, discussion selon un paramètre, tableau de signes, théorèmes de cours explicitant un ensemble de solutions (équations trigonométriques par exemple)...

A ce propos, on rappelle que dans un tableau de signe, le signe + signifie **strictement positif**, et le signe − signifie **strictement négatif** !

1.2 Tableaux de variations

On rappelle que dans un tableau de variations, le symbole $\nearrow$ signifie **strictement croissante et continue**, et le symbole $\searrow$ signifie **strictement décroissante et continue**.

Méthode 16.1 (Tableau de variations)**.** Voici comment réaliser le tableau de variations.

- On définit le domaine d'étude D et on étudie la une fonction définie sur D.
- On justifie la dérivabilité de la fonction : combinaison linéaire, produit, quotient dont le dénominateur ne s'annule pas, composée... de fonctions dérivables. Si ces théorèmes généraux ne suffisent pas, on peut être forcé de revenir à la définition de la dérivabilité sur des points précis du domaine.
- On calcule la dérivée. Pour le rédiger proprement, on écrit "$\forall x \in D,\ f(x) = \ldots$ donc $\forall x \in D,\ f'(x) = \ldots$.
- On étudie le signe de f'. En théorie, on a besoin de savoir le signe au sens strict de f', donc trois informations (stricte positivité, stricte négativité, nullité), mais en pratique seules deux suffisent car elles permettent de déterminer la troisième avec exactitude. Le plus souvent, le signe est facile (fonctions polynomiales de degré 1 ou 2), mais lorsque f' est plus complexe,

on pourra utiliser la rédaction suivante, plus efficace : "Soit $x \in D$. $f'(x) \geqslant 0$ ssi ..., avec égalité ssi ....

- De cette étude du signe de f', on en déduit les variations de f grâce à divers théorèmes. Pour l'énoncé de ces théorèmes et de leurs raffinements, on se réfèrera au chapitre "Dérivation".

1.3 Transformations de graphes

Définition 16.1 (Courbe représentative)**.** Si $\mathcal{P}$ désigne le plan usuel muni d'un repère orthonormal $(O; \overrightarrow{e_1}, \overrightarrow{e_2})$, on peut dessiner la **courbe représentative** d'une fonction $f : I \to \mathbb{R}$:

$$\mathcal{C}_f := \{M(x,y) \in \mathcal{P} \mid x \in I,\ y = f(x)\}$$

Proposition 16.1 (Translation horizontale)**.** *Soit $a \in \mathbb{R}$ et*

$$\begin{array}{rccc} t_{h,a} & : & \mathcal{P} & \to & \mathcal{P} \\ & & (x,y) & \mapsto & (x+a,y) \end{array}$$

Alors $t_{h,a}(\mathcal{C}_f)$ a pour équation $y = f(x-a)$, ie

$$t_{h,a}(\mathcal{C}_f) = \{M'(x,y) \mid x \in I + a,\ y = f(x-a)\}$$

Proposition 16.2 (Dilatation d'un facteur λ selon Ox)**.** *Soit $\lambda \in \mathbb{R}^\star$ et*

$$\begin{array}{rccc} d_{h,\lambda} & : & \mathcal{P} & \to & \mathcal{P} \\ & & (x,y) & \mapsto & (\lambda x,y) \end{array}$$

Alors $d_{h,\lambda}(\mathcal{C}_f)$ a pour équation $y = f\left(\frac{x}{\lambda}\right)$, , ie

$$d_{h,\lambda}(\mathcal{C}_f) = \left\{M'(x,y) \mid x \in \lambda I,\ y = f\left(\frac{x}{\lambda}\right)\right\}$$

Proposition 16.3 (Symétrie d'axe la droite d'équation $x = a$)**.** *Soit $a \in \mathbb{R}$ et*

$$\begin{array}{rccc} s_{h,a} & : & \mathcal{P} & \to & \mathcal{P} \\ & & (x,y) & \mapsto & (2a-x,y) \end{array}$$

Alors $s_{h,a}(\mathcal{C}_f)$ a pour équation $y = f(2a-x)$, ie

$$s_{h,a}(\mathcal{C}_f) = \{M'(x,y) \mid x \in 2a - I,\ y = f(2a-x)\}$$

Proposition 16.4 (Translation verticale)**.** *Soit $a \in \mathbb{R}$ et*

$$\begin{array}{rccc} t_{v,a} & : & \mathcal{P} & \to & \mathcal{P} \\ & & (x,y) & \mapsto & (x,y+a) \end{array}$$

Alors $t_{v,a}(\mathcal{C}_f)$ a pour équation $y = f(x) + a$, ie

$$t_{v,a}(\mathcal{C}_f) = \{M'(x,y) \mid x \in I,\ y = f(x) + a\}$$

Proposition 16.5 (Dilatation d'un facteur λ selon Oy)**.** *Soit $\lambda \in \mathbb{R}^\star$ et*

$$\begin{array}{rccc} d_{v,\lambda} & : & \mathcal{P} & \to & \mathcal{P} \\ & & (x,y) & \mapsto & (x,\lambda y) \end{array}$$

Alors $d_{v,\lambda}(\mathcal{C}_f)$ a pour équation $y = \lambda f(x)$, , ie

$$d_{h,\lambda}(\mathcal{C}_f) = \{M'(x,y) \mid x \in I,\ y = \lambda f(x)\}$$

Proposition 16.6 (Symétrie d'axe la droite d'équation $y = a$)**.** *Soit $a \in \mathbb{R}$ et*

$$\begin{array}{rccc} s_{h,a} & : & \mathcal{P} & \to & \mathcal{P} \\ & & (x,y) & \mapsto & (x,2a-y) \end{array}$$

Alors $s_{v,a}(\mathcal{C}_f)$ a pour équation $y = 2a - f(x)$, ie

$$s_{v,a}(\mathcal{C}_f) = \{M'(x,y) \mid x \in I,\ y = 2a - f(x)\}$$

Proposition 16.7 (Symétrie d'axe la droite d'équation $y = x$ lorsque f est bijective)**.** *Supposons que $f : \mathbb{R} \mapsto \mathbb{R}$ est bijective et soit*

$$\begin{array}{rccc} s & : & \mathcal{P} & \to & \mathcal{P} \\ & & (x,y) & \mapsto & (y,x) \end{array}$$

Alors $s(\mathcal{C}_f)$ a pour équation $y = f^{-1}(x)$, ie

$$s(\mathcal{C}_f) = \left\{M'(x,y) \mid x \in \mathbb{R},\ y = f^{-1}(x)\right\}$$

Définition 16.2 (Équations de droites du plan)**.** On dispose de deux types d'équations de droites du plan :

- Celles de la forme $y = ax + b$. Le **seul** inconvénient de ce type d'équations (et c'est le seul, car elles sont beaucoup plus manipulables que le type suivant) est qu'elles ne fournissent pas les droites verticales.
- Celles de la forme $ax+by+c = 0$, dites "**équations cartésienne**" avec $(a,b) \neq (0,0)$. Dans ce cas, on rappelle qu'un **vecteur directeur** de la droite est donné par $\begin{pmatrix} -b \\ a \end{pmatrix}$ et qu'un **vecteur normal** de la droite est donné par $\begin{pmatrix} a \\ b \end{pmatrix}$.

1.4 Systèmes linéaires

On se réfèrera à la dernière partie du chapitre "Matrices" qui traite en détail les systèmes linéaires. La méthode du pivot de Gauss doit être maîtrisée sur le bout des doigts.

1.5 Deux théorèmes admis

Dans ce paragraphe, on se donne I et J deux intervalles non triviaux de $\mathbb{R}$ et $f : I \to J$ une bijection. On pose $g := f^{-1} : J \to I$. On donne ici deux résultat admis qui seront démontrés dans la suite de chapitres d'analyse.

Théorème 16.1 (Théorème de la bijection réciproque, cas continu)**.** *Si f (avec les hypothèses du paragraphe) est strictement monotone, alors g est continue.*

Remarque 16.1. On montre aisément que g a mêmes variations que f, et on en déduit alors que f elle-même est continue en appliquant le théorème à g.

Théorème 16.2. *Soit $a \in I$ et $b := f(a)$. Si f est strictement monotone et dérivable en a, alors g est dérivable en b si, et seulement si, $f'(a) \neq 0$, et dans ce cas, on a $g'(b) = \dfrac{1}{f'(a)}$.*

Remarque 16.2. Lorsque f' ne s'annule pas, on en déduit une généralisation : g est dérivable, et on a $g' = \dfrac{1}{f' \circ f}$.

2 Logarithme, exponentielle

2.1 Logarithme népérien, logarithme en base a

Définition 16.3 (Logarithme népérien)**.** La fonction **logarithme népérien**, notée ln est l'unique primitive s'annulant en 1 de la fonction

$$\begin{array}{ccc} \mathbb{R}_+^\star & \to & \mathbb{R}_+^\star \\ x & \mapsto & \dfrac{1}{x} \end{array}$$

Autrement dit, ln est caractérisée par le fait que $\ln(1) = 0$, que ln est dérivable et que $\forall x > 0,\ \ln'(x) = \dfrac{1}{x}$.

Remarque 16.3. La justification de cette définition est admise pour l'instant et prendra son sens avec le théorème fondamental de l'analyse vu dans le chapitre "Intégration".

Exemple 16.1. En particulier, ln est continue.

Exemple 16.2. On déduit de la dérivabilité en 1 que $\dfrac{\ln(1+h)}{h} \xrightarrow[h\to 0]{} 1$

Théorème 16.3 (Primitive de $\dfrac{u'}{u}$)**.** *Soit $I \subset \mathbb{R}$ un intervalle non trivial et $u \in \mathcal{D}^1(I, \mathbb{R})$ telle que $\forall x \in I,\ u(x) \neq 0$. Alors, une primitive de $\dfrac{u'}{u}$ est donnée par $\ln|u|$.*

Remarque 16.4. Attention à ne pas oublier les valeurs absolues qui sont importantes ! Toujours les mettre, puis les enlever si l'intérieure est positif.

Théorème 16.4 (Propriété fonctionnelle du logarithme népérien)**.** *On a :*

$$\forall x, y > 0,\ \ln(xy) = \ln(x) + \ln(y)$$

Exemple 16.3. Ainsi, ln est un morphisme du groupe $(\mathbb{R}_+^\star, \times)$ dans le groupe $(\mathbb{R}, +)$.

Corollaire 16.1 (Autres propriétés fonctionnelles du logarithme népérien)**.** *On a les propriétés fonctionnelles suivantes pour le* ln *:*

1. $\forall x > 0,\ \ln\left(\frac{1}{x}\right) = -\ln(x)$
2. $\forall x, y > 0,\ \ln\left(\frac{x}{y}\right) = \ln(x) - \ln(y)$
3. $\forall x > 0,\ \forall n \in \mathbb{Z},\ \ln(x^n) = n\ln(x)$

Proposition 16.8 (Limites du logarithme népérien)**.** *On a :*

$$\begin{cases} \ln(x) \xrightarrow[x\to+\infty]{} +\infty \\ \ln(x) \xrightarrow[x\to 0^+]{} -\infty \end{cases}$$

Proposition 16.9 (Inégalité de concavité pour le logarithme népérien)**.** *On a :*

$$\forall x > -1,\ \ln(1+x) \leqslant x$$

avec égalité si, et seulement si $x = 0$.

Exemple 16.4. En appliquant l'inégalité de concavité à $\frac{-x}{1+x}$, on montre que

$$\forall x > -1,\ \ln(1+x) \geqslant \frac{x}{1+x}$$

Définition 16.4 (Logarithme en base a)**.** Soit $a \in \mathbb{R}_+^\star \setminus \{1\}$. Le **logarithme en base** a est la fonction définie par :

$$\begin{array}{rccc} \log_a & : & \mathbb{R}_+^\star & \to & \mathbb{R} \\ & & x & \mapsto & \dfrac{\ln(x)}{\ln(a)} \end{array}$$

2.2 Exponentielle

Définition 16.5 (Exponentielle)**.** On définit la **fonction exponentielle**, notée exp par :

$$\exp = \ln^{-1}$$

Exemple 16.5. $\exp(0) = 1$

Définition 16.6 (Constante e)**.** On pose $e := \exp(1)$. On a $e \approx 2,71828$.

Exemple 16.6. On a $\ln(e) = 1$ et $\log_e = \ln$.

Proposition 16.10 (Dérivabilité et dérivée de exp)**.** exp *est dérivable et vérifie* $\exp' = \exp$.

Corollaire 16.2. exp *est strictement croissante.*

Exemple 16.7. De la dérivabilité de exp en 0 on déduit $\dfrac{\exp(h)-1}{h} \xrightarrow[h\to 0]{} 1$.

Théorème 16.5 (Propriétés fonctionnelles de l'exponentielle)**.** *On a :*

1. $\forall (x,y) \in \mathbb{R}^2,\ \exp(x+y) = \exp(x)\exp(y)$
2. $\forall x \in \mathbb{R},\ \exp(-x) = \dfrac{1}{\exp(x)}$

3. $\forall (x,y) \in \mathbb{R}^2,\ \exp(x-y) = \dfrac{\exp(x)}{\exp(y)}$
4. $\forall x \in \mathbb{R},\ \forall n \in \mathbb{Z},\ \exp(nx) = \exp(x)^n$

Exemple 16.8. Ainsi, exp réalise un morphisme du groupe $(\mathbb{R}, +)$ dans le groupe $(\mathbb{R}_+^\star, \times)$.

Proposition 16.11 (Limites de l'exponentielle)**.** *On a :*

$$\begin{cases} \exp(x) \xrightarrow[x\to+\infty]{} +\infty \\ \exp(x) \xrightarrow[x\to-\infty]{} 0 \end{cases}$$

Proposition 16.12 (Inégalité de convexité pour l'exponentielle)**.** *On a :*

$$\forall x \in \mathbb{R},\ \exp(x) \geqslant 1 + x$$

avec égalité si, et seulement si $x = 0$.

2.3 Exponentielle quelconque

Définition 16.7 (Puissance quelconque d'un réel strictement positif)**.** Soit $a \in \mathbb{R}_+^\star$ et $x \in \mathbb{R}$. On **pose** :

$$a^x := \exp(x \ln(a))$$

Exemple 16.9. On a donc $\forall x \in \mathbb{R},\ \exp(x) = e^x$. On rencontrera souvent cette forme, plus courte que la notation $\exp(x)$.

Remarque 16.5. On peut prolonger cette définition à $a \in \mathbb{R}_+^\star$ et $x \in \mathbb{C}$.

Exemple 16.10. Soit $a \in \mathbb{R}_+^\star \setminus \{1\}$ et $b \in \mathbb{R}_+^\star$. On a :

$$\forall x \in \mathbb{R},\ a^x = b \iff x = \log_a(b)$$

Exemple 16.11. Soit $a \in \mathbb{R}_+^\star$ et $x \in \mathbb{R}$. On a $\ln(a^x) = x \ln(a)$.

Remarque 16.6. On vérifie que dans le cas où $x \in \mathbb{Z}$, la définition donnée ci-dessus (analytique) coïncide bien avec la définition algébrique (itérative) de a^x.

Proposition 16.13 (Propriétés fonctionnelles)**.** *Soit* $a, b > 0$ *et* $x, y \in \mathbb{R}$ *(voire* $\mathbb{C}$*). On a :*

- $a^x \times a^y = a^{x+y}$
- $a^{-x} = \dfrac{1}{a^x}$
- $a^{x-y} = \dfrac{a^x}{b^y}$
- $a^x \times b^x = (ab)^x$
- $(a^x)^y = a^{xy}$

Proposition 16.14 (Croissances comparées du logarithme népérien)**.** *On a*

$$\frac{\ln(x)}{x} \xrightarrow[x\to+\infty]{} 0$$

et

$$x \ln(x) \xrightarrow[x\to 0^+]{} 0$$

Corollaire 16.3 (Généralisation)**.** *Soit $\alpha, \beta > 0$. On a*

$$\frac{\ln(x)^\alpha}{x^\beta} \xrightarrow[x\to+\infty]{} 0$$

et

$$x^\beta |\ln(x)|^\alpha \xrightarrow[x\to 0^+]{} 0$$

Proposition 16.15 (Croissances comparées de l'exponentielle)**.** *On a*

$$\frac{x}{\exp(x)} \xrightarrow[x\to+\infty]{} 0$$

et

$$x\exp(x) \xrightarrow[x\to-\infty]{} 0$$

Corollaire 16.4 (Généralisation)**.** *Soit $\beta, \gamma > 0$. On a*

$$\frac{x^\beta}{\exp(x)^\gamma} \xrightarrow[x\to+\infty]{} 0$$

et

$$|x|^\beta \exp(x)^\gamma \xrightarrow[x\to-\infty]{} 0$$

Exemple 16.12. Soit $x > 0$ fixé. On a : $a^x \xrightarrow[a\to 0^+]{} 0$.

Définition 16.8 (0^x)**.** Par **convention**, on pose :

$$\forall x > 0,\ 0^x = 0$$

On rappelle que $0^0 = 1$ par **définition** et que pour $x < 0$, 0^x **n'existe pas**.

Remarque 16.7. Les identités restent vraies, du moment qu'elles sont définies. Pour être précis, il s'agit de :

- $\forall x, y > 0,\ 0^x \times 0^y = 0^{x+y}$
- $\forall x > 0,\ \forall b > 0,\ 0^x \times b^x = (0 \times b)^x$
- $\forall x, y > 0,\ (0^x)^y = 0^{xy}$

Exemple 16.13. On a

$$\forall a \in \mathbb{R}_+^\star,\ \forall n \in \mathbb{N}^\star,\ \sqrt[n]{a} = a^{\frac{1}{n}}$$

3 Trigonométrie

3.1 Trigonométrie circulaire (cosinus, sinus et tangente)

Proposition 16.16 (Une inégalité utile)**.** *On a :*

$$\forall x \in \mathbb{R},\ |\sin(x)| \leqslant |x|$$

Définition 16.9 (Fonction arc-cosinus)**.** La fonction $\cos_{|[0,\pi]}^{[-1,1]}$ est bijective. On appelle **arc-cosinus**, et on note $\arccos : [-1,1] \to [0,\pi]$ sa bijection réciproque.

Définition 16.10 (Fonction arc-sinus)**.** La fonction $\sin_{|[-\frac{\pi}{2},\frac{\pi}{2}]}^{[-1,1]}$ est bijective. On appelle **arc-sinus**, et on note $\arcsin : [-1,1] \to [-\frac{\pi}{2},\frac{\pi}{2}]$ sa bijection réciproque.

Définition 16.11 (Fonction arc-tangente)**.** La fonction $\tan_{|]-\frac{\pi}{2},\frac{\pi}{2}[}$ est bijective. On appelle **arc-tangente**, et on note $\arctan : \mathbb{R} \to]-\frac{\pi}{2},\frac{\pi}{2}[$ sa bijection réciproque.

Lemme 16.1. *On a :*

$$\forall x \in [-1,1],\ \cos(\arcsin(x)) = \sin(\arccos(x)) = \sqrt{1-x^2}$$

Théorème 16.6 (Étude de arccos)**.** arccos *est continue et strictement décroissante. Elle est dérivable en tout point de* $]-1,1[$ *et on a :*

$$\forall x \in]-1,1[,\ \arccos'(x) = \frac{-1}{\sqrt{1-x^2}}$$

Théorème 16.7 (Étude de arcsin)**.** arcsin *est impaire, continue et strictement croissante. Elle est dérivable en tout point de* $]-1,1[$ *et on a :*

$$\forall x \in]-1,1[,\ \arcsin'(x) = \frac{1}{\sqrt{1-x^2}}$$

Théorème 16.8 (Étude de arctan)**.** arctan *est impaire, continue et strictement croissante. Elle est dérivable en tout point de* $\mathbb{R}$ *et on a :*

$$\forall x \in \mathbb{R},\ \arctan'(x) = \frac{1}{1+x^2}$$

Proposition 16.17 (Trois relations utiles)**.** *On a les relations :*

- $\forall x \in]-1,1[,\ \arccos(x) + \arcsin(x) = \dfrac{\pi}{2}$
- $\forall x > 0,\ \arctan(x) + \arctan\left(\dfrac{1}{x}\right) = \dfrac{\pi}{2}$ *Cette relation s'avèrera très utile pour certains DL.*
- $\forall x < 0,\ \arctan(x) + \arctan\left(\dfrac{1}{x}\right) = -\dfrac{\pi}{2}$

3.2 Trigonométrie hyperbolique

Définition 16.12 (Partie paire, partie impaire)**.** Soit $f : \mathbb{R} \to \mathbb{R}$. Il existe un unique couple de fonctions $(f_1, f_2) \in \left(\mathbb{R}^{\mathbb{R}}\right)^2$ tel que f_1 soit paire, f_2 soit impaire et on ait $f = f_1 + f_2$. f_1 s'appelle la **partie paire** de f et f_2 s'appelle la **partie impaire** de f.

Remarque 16.8. Le résultat et la définition fonctionnent encore pour $f : I \to \mathbb{R}$ avec I centré en 0 (*ie* $\forall x \in I, -x \in I$)

Définition 16.13 (Cosinus hyperbolique)**.** La fonction **cosinus hyperbolique**, notée cosh ou ch est définie comme étant la partie paire de la fonction exp :

$$\begin{array}{rccc} \cosh & : & \mathbb{R} :\to & \mathbb{R} \\ & & x & \mapsto & \dfrac{e^x + e^{-x}}{2} \end{array}$$

Définition 16.14 (Sinus hyperbolique)**.** La fonction **sinus hyperbolique**, notée sinh ou sh est définie comme étant la partie impaire de la fonction exp :

$$\begin{array}{rccc} \sinh & : & \mathbb{R} :\to & \mathbb{R} \\ & & x & \mapsto & \dfrac{e^x - e^{-x}}{2} \end{array}$$

Définition 16.15 (Tangente hyperbolique)**.** La fonction **cosinus hyperbolique**, notée tanh ou th est définie par :

$$\begin{array}{rccc} \tanh & : & \mathbb{R} :\to & \mathbb{R} \\ & & x & \mapsto & \dfrac{\sinh(x)}{\cosh(x)} \end{array}$$

Théorème 16.9. *On a :*

$$\forall x \in \mathbb{R},\ \cosh(x)^2 - \sinh(x)^2 = 1$$

Remarque 16.9. C'est cette relation qui explique le nom de "trigonométrie hyperbolique" : elle définit le paramétrage d'une hyperbole.

Théorème 16.10 (Étude de cosh)**.** cosh *est paire, strictement décroissante sur* $]-\infty, 0]$ *et strictement croissante sur* $[0, +\infty[$*. Elle est dérivable et vérifie* $\cosh' = \sinh$*. On a* $\cosh(0) = 1$*,* $\cosh(x) \xrightarrow[x\to+\infty]{} +\infty$ *et* $\cosh(x) \xrightarrow[x\to-\infty]{} +\infty$*.*

Théorème 16.11 (Étude de sinh)**.** sinh *est impaire et strictement croissante. Elle est dérivable et vérifie* $\sinh' = \cosh$*. On a* $\sinh(0) = 0$*,* $\sinh(x) \xrightarrow[x\to+\infty]{} +\infty$ *et* $\sinh(x) \xrightarrow[x\to-\infty]{} -\infty$*.*

Théorème 16.12 (Étude de tanh)**.** tanh *est impaire et strictement croissante. Elle est dérivable et vérifie* $\tanh' = 1 - \tanh^2 = \dfrac{1}{\cosh^2}$*. On a* $\tanh(0) = 0$*,* $\tanh(x) \xrightarrow[x\to+\infty]{} 1$ *et* $\tanh(x) \xrightarrow[x\to-\infty]{} -1$*.*

Méthode 16.2 (Règle d'Osborn pour les formules de trigonométrie hyperbolique, HP)**.** Il existe des formules de trigonométrie hyperbolique, analogues aux formules de trigonométrie circulaire. Ces formules ne sont pas au programme et pas à connaître, mais elles peuvent parfois servir. Il existe une façon simple de retenir ces formules, qu'on appelle **règle d'Osborn** : une formule de trigonométrie hyperbolique est la même qu'une formule de trigonométrie circulaire en remplaçant les cos par des cosh, les sin par des sinh et les tan par des tanh, à ceci près qu'un $\sin(a)\sin(b)$ doit être remplacé par un $-\sinh(a)\sinh(b)$.

Exemple 16.14. La formule $\cos(a+b) = \cos(a)\cos(b) - \sin(a)\sin(b)$ devient

$$\cosh(a+b) = \cosh(a)\cosh(b) + \sinh(a)\sinh(b)$$

(on pourra le vérifier en revenant à la définition si on souhaite vérifier la validité de cette sorcellerie). La formule $\cos(x)^2 + \sin(x)^2 = 1$ devient $\cosh(x)^2 - \sinh(x)^2 = 1$, ce que nous avions prouvé par des calculs explicites.

Remarque 16.10 (Justification de la règle d'Osborn, HP)**.** Ceci est dû au fait qu'on peut prolonger les fonctions sin et cos à $\mathbb{C}$ grâce aux formules d'Euler, et alors sinh et cosh ne sont que des cas particuliers de sinus et de cosinus complexes. Par exemple, on a $\forall x \in \mathbb{R},\ \sinh(x) = -i\sin(ix)$ et $\forall x \in \mathbb{R},\ \cosh(x) = \cos(ix)$ ce qui explique l'apparition du signe $-$ devant ce qui était auparavant un produit de 2 sinus. On peut en fait généraliser les formules de trigonométrie circulaire aux sinus et cosinus prolongés sur $\mathbb{C}$.

4 Quelques rappels sur le calcul de primitives

On se réfèrera au chapitre "Intégration", et plus précisément à la partie "Intégration et dérivation", où tout est développé.

Voici un tableau des primitives usuelles à connaître par cœur :

$f(x)$	$\int f$
$x^\alpha \quad (\alpha \in \mathbb{R} \setminus \{-1\})$	$\dfrac{x^{\alpha+1}}{\alpha+1}$
$\dfrac{1}{x}$	$\ln\lvert x\rvert$
$e^{ax} \quad (a \in \mathbb{R}^\star)$	$\dfrac{e^{ax}}{a}$
$\cos(x)$	$\sin(x)$
$\sin(x)$	$-\cos(x)$
$\tan(x)$	$-\ln\lvert\cos(x)\rvert$
$\dfrac{1}{\cos(x)}$	$\ln\left\lvert\tan\left(\dfrac{x}{2}+\dfrac{\pi}{4}\right)\right\rvert$ (NE)
$\dfrac{1}{\sin(x)}$	$\ln\left\lvert\tan\left(\dfrac{x}{2}\right)\right\rvert$ (NE)
$\dfrac{1}{1+x^2}$	$\arctan(x)$
$\dfrac{1}{1-x^2}$	$\operatorname{argth}(x)$
$\dfrac{1}{\sqrt{1+x^2}}$	$\operatorname{argsh}(x)$
$\dfrac{1}{\sqrt{x^2-1}}$	$\operatorname{argch}(x)$
$\dfrac{1}{\sqrt{1-x^2}}$	$\arcsin(x)$
$\dfrac{-1}{\sqrt{1-x^2}}$	$\arccos(x)$

Chapitre 17

Limites et continuité

1 Préliminaires

Dans tout le paragraphe, on fixe $X \subset \mathbb{R}$.

1.1 Initiation à la topologie

On introduit ici quelques notions de topologie simples qui nous permettront d'unifier la suite du chapitre, et en particulier de ne pas à avoir à donner différentes définitions pour les différentes limites.

Définition 17.1 (Voisinages fondamentaux dans $\mathbb{R}$)**.** Soit $a \in \mathbb{R}$. Un **voisinage fondamental de** a est un intervalle de la forme $[a-\delta, a+\delta]$ où $\delta \in \mathbb{R}_+^\star$. On notera ici $\mathcal{V}(a)$ (NS) l'ensemble des voisinages fondamentaux de a.

Exemple 17.1. Soit $u \in \mathbb{R}^{\mathbb{N}}$ et $l \in \mathbb{R}$. On a $u_n \xrightarrow[n\to+\infty]{} l$ si, et seulement si, $\forall V \in \mathcal{V}(l),\ u_n \in V$ APCR.

Exemple 17.2. X est dense dans $\mathbb{R}$ si, et seulement si, $\forall a \in \mathbb{R},\ \forall V \in \mathcal{V}(a),\ X \cap V \neq \emptyset$.

Définition 17.2 (Adhérence dans $\mathbb{R}$)**.** Soit $a \in \mathbb{R}$. a est dit **adhérent à** X lorsque :

$$\forall V \in \mathcal{V}(a),\ X \cap V \neq \emptyset$$

L'**adhérence de** X **dans** $\mathbb{R}$ est l'ensemble $\{a \in \mathbb{R} \mid a \text{ est adhérent à } X\}$.

Exemple 17.3. Tout point de X est adhérent à X.

Exemple 17.4. Personne n'est adhérent à l'ensemble vide.

Exemple 17.5. X est dense dans $\mathbb{R}$ si, et seulement si, l'adhérence de X dans $\mathbb{R}$ vaut $\mathbb{R}$.

Théorème 17.1 (Caractérisation séquentielle de l'adhérence dans $\mathbb{R}$)**.** *Soit $a \in \mathbb{R}$. a est adhérent à X si, et seulement si,*

$$\exists (x_n)_{n\in\mathbb{N}} \in X^{\mathbb{N}},\ x_n \xrightarrow[n\to+\infty]{} a$$

Désormais, on étend nos définitions à $\overline{\mathbb{R}}$.

Définition 17.3 (Voisinages fondamentaux dans $\overline{\mathbb{R}}$)**.** Soit $a \in \overline{\mathbb{R}}$.

- Si $a \in \mathbb{R}$, la définition d'un voisinage fondamental reste la même.
- Si $a = +\infty$, on appelle voisinage fondamental de a tout intervalle de la forme $[M, +\infty]$ avec $M \in \mathbb{R}$.
- Si $a = -\infty$, on appelle voisinage fondamental de a tout intervalle de la forme $[-\infty, m]$ avec $m \in \mathbb{R}$.

Dans tous les cas, les voisinages fondamentaux sont des intervalles de $\overline{\mathbb{R}}$ qui contiennent a (mais ce n'est pas suffisant, $\{a\}$ n'est pas un voisinage fondamental de a mais contient a). On notera toujours $\mathcal{V}(a)$ l'ensemble des voisinages fondamentaux de a.

Exemple 17.6. Soit $u \in \mathbb{R}^{\mathbb{N}}$ et $l \in \overline{\mathbb{R}}$. On a $u_n \xrightarrow[n\to+\infty]{} l$ si, et seulement si, $\forall V \in \mathcal{V}(l),\ u_n \in V$ APCR.

Exemple 17.7. Pour tout $a \in \overline{\mathbb{R}}$, $\mathcal{V}(a)$ est totalement ordonné pour $\subset$.

Proposition 17.1. *Pour tout $a \in \overline{\mathbb{R}}$, $\mathcal{V}(a)$ est stable par intersection finie (conséquence de la relation d'ordre totale).*

Définition 17.4 (Adhérence dans $\overline{\mathbb{R}}$)**.** Soit $a \in \overline{\mathbb{R}}$. a est dit **adhérent à** X lorsque :

$$\forall V \in \mathcal{V}(a),\ X \cap V \neq \emptyset$$

L'**adhérence de X dans** $\overline{\mathbb{R}}$ est l'ensemble :

$$adh_{\overline{\mathbb{R}}}(X) := \left\{a \in \overline{\mathbb{R}} \mid a \text{ est adhérent à } X\right\}$$

Jusqu'à la fin de l'année, implicitement, l'adhérence de X sera toujours l'adhérence de X dans $\overline{\mathbb{R}}$. On la note plus souvent $\overline{X}$.

Théorème 17.2 (Caractérisation séquentielle de l'adhérence dans $\overline{\mathbb{R}}$)**.** *Soit $a \in \overline{\mathbb{R}}$. a est adhérent à X si, et seulement si,*

$$\exists (x_n)_{n\in\mathbb{N}} \in X^{\mathbb{N}},\ x_n \xrightarrow[n\to+\infty]{} a$$

Exemple 17.8. Soit $u, v \in \overline{\mathbb{R}}$ tels que $u < v$. Alors $\overline{]u, v[} = [u, v]$

Définition 17.5 (Propriété au voisinage de)**.** Soit $f : X \to \mathbb{R}$ et $a \in \overline{X}$. On dit que f **possède la propriété P au voisinage de** a lorsqu'il existe un voisinage fondamental V de a tel que $f_{|X\cap V}$ possède la propriété P.

Exemple 17.9. cos est à valeurs strictement positives au voisinage de 0. Il suffit de prendre $V = \left]-\frac{\pi}{4}, \frac{\pi}{4}\right[$.

Exemple 17.10. sin n'est pas à valeurs strictement positives au voisinage de 0.

1.2 Premières définitions

Définition 17.6 (Structure de $\mathbb{R}$-algèbre de $\mathbb{R}^X$, rappel)**.** On rappelle que si $f : X \to \mathbb{R}$, $g : X \to \mathbb{R}$ et $\lambda \in \mathbb{R}$, on définit :

$$\begin{array}{rccc} f+g & : & X & \to & \mathbb{R} \\ & & x & \mapsto & f(x)+g(x) \end{array}$$

$$\begin{array}{rccc} f\times g & : & X & \to & \mathbb{R} \\ & & x & \mapsto & f(x)\times g(x) \end{array}$$

$$\begin{array}{rccc} \lambda\cdot f & : & X & \to & \mathbb{R} \\ & & x & \mapsto & \lambda\cdot f(x) \end{array}$$

Ainsi, $(\mathbb{R}^X, +, \times)$ est un anneau et $(\mathbb{R}^X, +, \times, \cdot)$ est une $\mathbb{R}$-algèbre.

Définition 17.7 (Relation d'ordre partielle sur $\mathbb{R}^X$, rappel)**.** On définit une relation binaire sur $\mathbb{R}^X$ par :

$$\forall (f,g) \in (\mathbb{R}^X)^2,\ f \leqslant g \iff \forall x \in X,\ f(x) \leqslant g(x)$$

On vérifie que $\leqslant$ est une relation d'ordre (partielle en général) sur $\mathbb{R}^X$.

Définition 17.8 (Fonction majorée, minorée, bornée, rappel)**.** Soit $f \in \mathbb{R}^X$. On rappelle que :

- f est dite **majorée** lorsque : $\exists M \in \mathbb{R},\ \forall x \in X,\ f(x) \leqslant M$
- f est dite **minorée** lorsque : $\exists m \in \mathbb{R},\ \forall x \in X,\ f(x) \geqslant m$
- f est dite **bornée** lorsqu'elle est à la fois majorée et minorée

Exemple 17.11. L'ensemble des fonctions bornées de X dans $\mathbb{R}$ forme une sous-algèbre de $\mathbb{R}^X$.

Proposition 17.2. *f est bornée si, et seulement si, $|f|$ est majorée.*

Définition 17.9 (Maximum, minimum, maximum local, minimum local en un point)**.** Soit $f : X \to \mathbb{R}$ et $a \in X$.

- On dit que f admet un **maximum en** a lorsque $\forall x \in X,\ f(x) \leqslant f(a)$.
- On dit que f admet un **minimum en** a lorsque $\forall x \in X,\ f(x) \geqslant f(a)$.
- On dit que f admet un **maximum local en** a lorsque $\exists V \in \mathcal{V}(a),\ \forall x \in X \cap V,\ f(x) \leqslant f(a)$.
- On dit que f admet un **minimum local en** a lorsque $\exists V \in \mathcal{V}(a),\ \forall x \in X \cap V,\ f(x) \geqslant f(a)$.

Définition 17.10 (Maximum et minimum d'une fonction)**.** Soit $f : X \to \mathbb{R}$.

- S'il existe, le maximum de f peut être noté $\max\limits_{x\in X}(f(x))$.
- S'il existe, le minimum de f peut être noté $\min\limits_{x\in X}(f(x))$.

Définition 17.11 (Sup et inf d'une fonction)**.** Plus généralement, supposons que $X \neq \emptyset$ et soit $f : X \to \mathbb{R}$. Alors, $f(X) \neq \emptyset$ donc on peut définir :

- $\sup\limits_{x\in X}(f(x)) \in \mathbb{R} \cup \{+\infty\}$
- $\inf\limits_{x\in X}(f(x)) \in \mathbb{R} \cup \{-\infty\}$

Définition 17.12 (Monotonies)**.** Soit $f : X \to \mathbb{R}$.

- f est dite **croissante** lorsque $\forall (x,y) \in X^2,\ x \leqslant y \implies f(x) \leqslant f(y)$
- f est dite **décroissante** lorsque $\forall (x,y) \in X^2,\ x \leqslant y \implies f(x) \geqslant f(y)$
- f est dite **strictement croissante** lorsque $\forall (x,y) \in X^2,\ x < y \implies f(x) < f(y)$
- f est dite **strictement décroissante** lorsque $\forall (x,y) \in X^2,\ x < y \implies f(x) > f(y)$
- f est dite **monotone** lorsqu'elle est croissante ou décroissante
- f est dite **strictement monotone** lorsqu'elle est strictement croissante ou strictement décroissante

Exemple 17.12. Si $X = \emptyset$, f est la fois croissante, décroissante, strictement croissante et strictement décroissante par proposition vide. Si X est un singleton, f est croissante et décroissante car n'importe quoi implique le vrai, et f est strictement croissante et strictement décroissante car le faux implique n'importe quoi.

Définition 17.13 (Propriété sur)**.** Soit $f : X \to \mathbb{R}$ et $Y \subset X$. On dit que f **possède la propriété** P **sur** Y lorsque $f_{|Y}$ possède la propriété P.

Définition 17.14 (Fonctions paires et impaires)**.** Supposons que X soit centré en 0 (*ie* que $\forall x \in X,\ -x \in X$). Soit $f : X \to \mathbb{R}$. On dit que f est **paire** lorsque

$$\forall x \in X,\ f(-x) = f(x)$$

On dit que f est **impaire** lorsque

$$\forall x \in X,\ f(-x) = -f(x)$$

L'ensemble des fonctions paires et l'ensemble des fonctions impaires de X dans $\mathbb{R}$ forment des sev supplémentaires de $\mathbb{R}^X$.

Remarque 17.1. Soit $a \in \mathbb{R}$ et supposons que $\forall x \in X,\ 2a - x \in X$. On peut alors dire que f est paire par rapport à a lorsque $\forall x \in X,\ f(2a - x) = f(x)$ et que f est impaire par rapport à a lorsque $\forall x \in X,\ f(2a - x) = -f(x)$. Là encore, ces ensembles forment des sev supplémentaires de $\mathbb{R}^X$.

Définition 17.15 (Périodicité)**.** Soit $T > 0$ et supposons que $\forall x \in X,\ x + t \in X$. Soit $f : X \to \mathbb{R}$. f est dite T**-périodique** lorsque

$$\forall x \in X,\ f(x + T) = f(x)$$

Exemple 17.13. Les fonctions T-périodiques forment une sous-algèbre de $\mathbb{R}^X$.

Définition 17.16 (Fonctions lipschitziennes)**.** Soit $\lambda \geqslant 0$ et $f : X \to \mathbb{R}$. f est dite λ**-lipschitzienne** lorsque

$$\forall (x,y) \in X^2,\ |f(x) - f(y)| \leqslant \lambda |x - y|$$

Plus généralement, f est dite **lipschitzienne** lorsqu'il existe $\lambda \geqslant 0$ tel que f soit λ-lipschitzienne.

Exemple 17.14. Une fonction 0-lipschitzienne est constante.

Remarque 17.2. La lipschitziannité est une condition très forte.

Proposition 17.3 (Interprétation géométrique des fonctions lipschitziennes)**.** *Supposons que f soit λ-lipschitzienne. Imaginons un double cône de pentes λ et $-\lambda$ qui coulisse sur la courbe. Alors, la courbe de f se trouve toujours dans les parties Ouest et Est du double cône, peu importe la position de celui-ci.*

Exemple 17.15. La fonction racine carrée n'est pas lipschitzienne.

2 Étude locale d'une fonction

Dans tout le paragraphe, on fixe $X \subset \mathbb{R}$.

2.1 Limites

Définition 17.17 (Limite d'une fonction)**.** Soit $f : X \to \mathbb{R}$, $a \in \overline{X}$ et $b \in \overline{\mathbb{R}}$. On dit que $f(x)$ **tend vers** b **quand** x **tend vers** a, et on note $f(x) \xrightarrow[x\to a]{} b$, lorsque :

$$\forall V_b \in \mathcal{V}(b),\ \exists V_a \in \mathcal{V}(a),\ \forall x \in X,\ x \in V_a \implies f(x) \in V_b$$

Exemple 17.16 (Fonctions constantes)**.** Une fonction constante égale à C tend vers C en tout point de son adhérence.

Proposition 17.4. *Supposons que $(a, b) \in \mathbb{R}^2$. Les 4 assertions suivantes sont équivalentes :*

1. $f(x) \xrightarrow[x\to a]{} b$
2. $f(x) - b \xrightarrow[x\to a]{} 0$
3. $|f(x) - b| \xrightarrow[x\to a]{} 0$
4. $f(a + h) \xrightarrow[h\to 0]{} b$

Définition 17.18. La notation $f(x) \xrightarrow[x\to a]{} 0^+$ signifie que $f(x) \xrightarrow[x\to a]{} 0$ et $\exists V \in \mathcal{V}(a),\ \forall x \in X \cap V,\ f(x) > 0$. La notation $f(x) \xrightarrow[x\to a]{} 0^-$ signifie que $f(x) \xrightarrow[x\to a]{} 0$ et $\exists V \in \mathcal{V}(a),\ \forall x \in X \cap V,\ f(x) < 0$.

Lemme 17.1 ($\overline{\mathbb{R}}$ est séparé)**.** *On a :*

$$\forall (b, b') \in \overline{\mathbb{R}}^2,\ b \neq b' \implies (\exists (V, V') \in \mathcal{V}(b) \times \mathcal{V}(b'),\ V \cap V' = \emptyset)$$

Théorème 17.3 (Unicité de la limité)**.** *On a :*

$$\forall a \in \overline{X},\ \forall (b, b') \in \overline{\mathbb{R}}^2,\ (f(x) \xrightarrow[x\to a]{} b \wedge f(x) \xrightarrow[x\to a]{} b') \implies b = b'$$

La limite d'une fonction en un point, si elle existe, est donc unique.

Définition 17.19 (Limite d'une restriction)**.** Soit $f : X \to \mathbb{R}$, $Y \subset \mathbb{R}$ et $a \in \overline{X \cap Y}$. Posons $X^\star := X \cap Y$ et $f^\star := f_{|X^\star}$. Soit $b \in \overline{\mathbb{R}}$. Si $f^\star(x) \xrightarrow[x\to a]{} b$, alors on note symboliquement

$$f(x) \xrightarrow[\substack{x\to a \\ x\in Y}]{} b$$

On dit que b est la **limite de** f **en** a **relativement à** Y. Voici trois cas particuliers :

- $Y =]a, +\infty[$ On parle de **limite à droite** et on note $f(x) \xrightarrow[\substack{x\to a \\ x>a}]{} b$ ou $f(x) \xrightarrow[x\to a^+]{} b$, voire $f(a^+) = b$.
- $Y =]-\infty, a[$ On parle de **limite à gauche** et on note $f(x) \xrightarrow[\substack{x\to a \\ x<a}]{} b$ ou $f(x) \xrightarrow[x\to a^-]{} b$, voire $f(a^-) = b$.

- $Y = \mathbb{R} \setminus \{a\}$ On parle de **limite épointée** et on note $f(x) \xrightarrow[\substack{x\to a\\ x\neq a}]{} b$.

Remarque 17.3. Attention : dans le cas des limites à droite, à gauche et épointée, le point a n'appartient pas à $X^\star$, donc $f(a)$ peut exister et être différente de b. On pourra considérer la limite épointée en 0 de l'indicatrice de $\{0\}$.

Proposition 17.5 (Limite induite)**.** *Si* $f(x) \xrightarrow[x\to a]{} b$*, alors* $f(x) \xrightarrow[\substack{x\to a\\ x\in Y}]{} b$.

Remarque 17.4. Attention, la réciproque est fausse en générale : par exemple, si on a des limites à gauche et à droite égales, on ne peut en déduire que la limite épointée.

Proposition 17.6 (Réciproque partielle)**.** *Si* $Y \in \mathcal{V}(a)$ *et* $f(x) \xrightarrow[\substack{x\to a\\ x\in Y}]{} b$*, alors* $f(x) \xrightarrow[x\to a]{} b$.

Exemple 17.17. Si $f(x) \xrightarrow[\substack{x\to a\\ x\in Y_1}]{} b$, ... et $f(x) \xrightarrow[\substack{x\to a\\ x\in Y_n}]{} b$, alors $f(x) \xrightarrow[\substack{x\to a\\ x\in Y_1\cup\ldots\cup Y_n}]{} b$

Méthode 17.1 (Montrer que $f(x)$ ne tend pas vers b)**.** Pour montrer que $f(x)$ ne tend pas vers b lorsque x tend vers a, il suffit d'exhiber $Y \subset \mathbb{R}$ tel que $a \in \overline{X \cap Y}$, ainsi que $b' \neq b$ tels que $f(x) \xrightarrow[\substack{x\to a\\ x\in Y}]{} b'$.

Méthode 17.2 (Montrer que f n'admet pas de limite en a)**.** Pour montrer que f n'admet pas de limite en a, il suffit d'exhiber $Y_1, Y_2 \subset \mathbb{R}$ tels que $a \in \overline{X \cap Y_1}$ et $a \in \overline{X \cap Y_2}$ et $b_1 \neq b_2$ tels que $f(x) \xrightarrow[\substack{x\to a\\ x\in Y_1}]{} b_1$ et $f(x) \xrightarrow[\substack{x\to a\\ x\in Y_2}]{} b$.

Théorème 17.4 (Caractère local de la limite)**.** *Supposons que* $Y \in \mathcal{V}(a)$*. Alors :*

$$f(x) \xrightarrow[x\to a]{} b \iff f(x) \xrightarrow[\substack{x\to a\\ x\in Y}]{} b$$

Exemple 17.18. Si f et g coïncident au voisinage de a, alors elles ont le même comportement asymptotique en a.

Théorème 17.5 (Caractérisation séquentielle de la limite)**.** *Soit* $f : X \to \mathbb{R}$*,* $a \in \overline{X}$ *et* $b \in \overline{\mathbb{R}}$*. Alors*

$$f(x) \xrightarrow[x\to a]{} b \iff \forall (x_n) \in X^{\mathbb{N}},\ x_n \xrightarrow[n\to+\infty]{} a \implies f(x_n) \xrightarrow[n\to+\infty]{} b$$

Méthode 17.3 (Variante : montrer que $f(x)$ ne tend pas vers b)**.** Pour montrer que $f(x)$ ne tend pas vers b lorsque x tend vers a, il suffit d'exhiber $(x_n) \in X^{\mathbb{N}}$ telle que $x_n \xrightarrow[n\to+\infty]{} a$ et $b' \neq b$ tels que $f(x_n) \xrightarrow[n\to+\infty]{} b'$.

Méthode 17.4 (Variante : montrer que f n'admet pas de limite en a)**.** Pour montrer que f n'admet pas de limite en a, il suffit d'exhiber $(x_n) \in X^{\mathbb{N}}$ et $(y_n) \in X^{\mathbb{N}}$ telles que $x_n \xrightarrow[n\to+\infty]{} a$ et $y_n \xrightarrow[n\to+\infty]{} a$ ainsi que $b \neq b'$ tels que $f(x_n) \xrightarrow[n\to+\infty]{} b$ et $f(y_n) \xrightarrow[n\to+\infty]{} b'$.

Exemple 17.19. cos n'admet pas de limite en $+\infty$. En effet, il suffit de considérer les suites $(2\pi n)_{n\in\mathbb{N}}$ et $(\pi + 2\pi n)_{n\in\mathbb{N}}$

Proposition 17.7 (Combinaison linéaire de limites). *Soit $f, g : X \to \mathbb{R}$, $a \in \overline{X}$ et $b_1, b_2 \in \overline{\mathbb{R}}$ tels que $f(x) \xrightarrow[x\to a]{} b_1$ et $g(x) \xrightarrow[x\to a]{} b_2$. Soit $\lambda, \mu \in \mathbb{R}$. Pourvu qu'on n'ait pas de forme indéterminée, on a :*

$$\lambda f(x) + \mu g(x) \xrightarrow[x\to a]{} \lambda b_1 + \mu b_2$$

Proposition 17.8 (Produit de limites). *Soit $f, g : X \to \mathbb{R}$, $a \in \overline{X}$ et $b_1, b_2 \in \overline{\mathbb{R}}$ tels que $f(x) \xrightarrow[x\to a]{} b_1$ et $g(x) \xrightarrow[x\to a]{} b_2$. Pourvu qu'on n'ait pas de forme indéterminée, on a :*

$$f(x)g(x) \xrightarrow[x\to a]{} b_1 b_2$$

Proposition 17.9 (Quotient de limites). *Soit $f, g : X \to \mathbb{R}$, $a \in \overline{X}$ et $b_1, b_2 \in \overline{\mathbb{R}}$ tels que $f(x) \xrightarrow[x\to a]{} b_1$ et $g(x) \xrightarrow[x\to a]{} b_2$. Supposons que $\forall x \in X,\ g(x) \neq 0$ et $b_2 \neq 0$. Pourvu qu'on n'ait pas de forme indéterminée, on a :*

$$\frac{f(x)}{g(x)} \xrightarrow[x\to a]{} \frac{b_1}{b_2}$$

On a les raffinements habituels si g ne s'annule pas au voisinage de a, selon le signe de b_1 et si $g(x) \xrightarrow[x\to a]{} 0^+$ ou $g(x) \xrightarrow[x\to a]{} 0^-$.

Proposition 17.10 (Valeur absolue de limites). *Soit $f : X \to \mathbb{R}$, $a \in \overline{X}$ et $b \in \overline{\mathbb{R}}$ tels que $f(x) \xrightarrow[x\to a]{} b$. Alors on a :*

$$|f(x)| \xrightarrow[x\to a]{} |b|$$

Proposition 17.11 (Composition de limites). *Soit $f : X \to \mathbb{R}$, $a \in \overline{X}$, $Y \subset \mathbb{R}$ tel que $f(X) \subset Y$, $g : Y \to \mathbb{R}$, $b \in \overline{Y}$ et $c \in \overline{\mathbb{R}}$. Si $f(x) \xrightarrow[x\to a]{} b$ et $g(y) \xrightarrow[y\to b]{} c$, alors :*

$$(g \circ f)(x) \xrightarrow[x\to a]{} c$$

Exemple 17.20. Par caractérisation séquentielle, on obtient la limite de x quand x tend vers $+\infty$ (ou $-\infty$). On en déduit par ls propositions qui précèdent les limites des fonctions polynomiales et des fonctions rationnelles.

Théorème 17.6 (Théorème d'encadrement/sandwich/des gendarmes). *Soit $f, g, h : X \to \mathbb{R}$ telles que $f \leqslant g \leqslant h$. Soit $a \in \overline{X}$ et $b \in \overline{\mathbb{R}}$. Si $f(x) \xrightarrow[x\to a]{} b$ et $h(x) \xrightarrow[x\to a]{} b$, alors $g(x) \xrightarrow[x\to a]{} b$.*

Corollaire 17.1. *Soit $f, g : X \to \mathbb{R}$ telles que $\forall x \in X,\ |f(x)| \leqslant g(x)$ et $a \in \overline{X}$. Si $g(x) \xrightarrow[x\to a]{} 0$, alors $f(x) \xrightarrow[x\to a]{} 0$.*

Corollaire 17.2. *Soit $f, g : X \to \mathbb{R}$ avec g bornée. Si $f(x) \xrightarrow[x\to a]{} 0$, alors $f(x)g(x) \xrightarrow[x\to a]{} 0$.*

Théorème 17.7 (Théorème de comparaison). *Soit $f, g : X \to \mathbb{R}$ telles que $f \leqslant g$ et $a \in \overline{X}$.*

- *Principe de minoration : si $f(x) \xrightarrow[x\to a]{} +\infty$, alors $g(x) \xrightarrow[x\to a]{} +\infty$.*
- *Principe de majoration : si $g(x) \xrightarrow[x\to a]{} -\infty$, alors $f(x) \xrightarrow[x\to a]{} -\infty$.*

Théorème 17.8 (Passage à la limite dans une inégalité large / PLIL)**.** *Soit $f, g : X \to \mathbb{R}$, $a \in \overline{X}$ et $b, b' \in \overline{\mathbb{R}}$ tels que $f(x) \xrightarrow[x\to a]{} b$ et $g(x) \xrightarrow[x\to a]{} b'$. Si $\forall x \in X,\ f(x) \leqslant g(x)$, alors $b \leqslant b'$.*

Exemple 17.21 (Cas particuliers fréquents)**.** Si $\forall x \in X,\ f(x) \leqslant M$ avec M indépendant de x, alors $b \leqslant M$. Si $\forall x \in X,\ m \leqslant g(x)$ avec m indépendant de x, alors $m \leqslant b'$.

Remarque 17.5. Grâce au caractère local de la limite, tous les résultats précédents qui font appel à des inégalités sont encore vrais si les inégalités n'ont lieu qu'au voisinage de a.

Théorème 17.9 (Réciproque partielle du PLIL)**.** *Soit $f : X \to \mathbb{R}$, $a \in \overline{X}$ et $b \in \overline{\mathbb{R}}$ tels que $f(x) \xrightarrow[x\to a]{} b$. Soit $(\alpha, \beta) \in \mathbb{R}^2$.*

- *Si $b \in \mathbb{R}$, alors f est bornée au voisinage de a.*
- *Si $b > \alpha$, alors $f(x) > \alpha$ au voisinage de a.*
- *Si $b < \beta$, alors $f(x) < \beta$ au voisinage de a.*

Théorème 17.10 (Théorème de la limite monotone / TLM)**.** *Soit $(u, b) \in \overline{\mathbb{R}}^2$ tel que $u < v$ et $f :]u, v[\to \mathbb{R}$. Si f est monotone, alors f admet des limites en u et en v. De manière plus précise :*

- *Supposons que f est croissante. Alors*
  - *f admet une limite $b \in \mathbb{R} \cup \{+\infty\}$ (dans $\mathbb{R}$ si f est majorée, et valant $+\infty$ si f n'est pas majorée) en v*
  - *f admet une limite $b' \in \mathbb{R} \cup \{-\infty\}$ (dans $\mathbb{R}$ si f est minorée, et valant $-\infty$ si f n'est pas minorée) en u*
- *Supposons que f est décroissante. Alors*
  - *f admet une limite $b \in \mathbb{R} \cup \{-\infty\}$ (dans $\mathbb{R}$ si f est minorée, et valant $-\infty$ si f n'est pas minorée) en v*
  - *f admet une limite $b' \in \mathbb{R} \cup \{+\infty\}$ (dans $\mathbb{R}$ si f est majorée, et valant $+\infty$ si f n'est pas majorée) en u*

Exemple 17.22. Soit $f :]0, 1[\to \mathbb{R}$ croissante et majorée par un certain $M \in \mathbb{R}$. Alors elle admet une limite l en 1 d'après le TLM. Un PLIL montre que $l \leqslant M$. On montre que $\forall x \in]0, 1[,\ f(x) \leqslant l$ en fixant x, en utilisant la croissance et en faisant un PLIL sur $y \geqslant x$. Si on suppose que f est strictement croissante, on montre que $\forall x \in]0, 1[,\ f(x) < l$ en fixant $y \in]x, 1[$ et en utilisant le résultat qui précède sur $f(y)$ ainsi que la stricte croissance.

Corollaire 17.3. *Soit $(u, b) \in \overline{\mathbb{R}}^2$ tel que $u < v$ et $f :]u, v[\to \mathbb{R}$. Si f est monotone, alors en tout point a de $]u, v[$, f admet des limites à gauche et à droite qui encadrent $f(a)$. De manière précise, soit $a \in]u, v[$.*

- *Supposons que f est croissante. Alors $f(a^-)$ et $f(a^+)$ existent et sont finies, et on a $f(a^-) \leqslant f(a) \leqslant f(a^+)$.*
- *Supposons que f est décroissante. Alors $f(a^-)$ et $f(a^+)$ existent et sont finies, et on a $f(a^-) \geqslant f(a) \geqslant f(a^+)$.*

2.2 Continuité locale

Définition 17.20 (Continuité en un point)**.** Soit $f : X \to \mathbb{R}$ et $a \in X$. On dit que f est **continue en** a lorsque $f(x) \xrightarrow[x\to a]{} f(a)$, autrement dit lorsque :

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall x \in X,\ |x-a| \leqslant \delta \implies |f(x)-f(a)| \leqslant \varepsilon$$

Exemple 17.23. Une fonction constante est continue en chacun de ses points.

Exemple 17.24. La fonction $\mathrm{id}_{\mathbb{R}}$ est continue en tout point de $\mathbb{R}$.

Proposition 17.12. *Soit $f : X \to \mathbb{R}$ et $a \in X$. Soit $b \in \overline{\mathbb{R}}$. Si $f(x) \xrightarrow[x\to a]{} b$, alors $b = f(a)$ si bien que f est continue en a.*

Remarque 17.6. Cela provient du fait que désormais, a est dans X donc ne peut plus être en bord d'intervalle, là d'où proviennent la majorité de nos contre-exemples.

Définition 17.21 (Continuité à gauche et à droite)**.** Soit $f : X \to \mathbb{R}$ et $a \in X$. On dit que f est **continue à gauche en** a lorsque $f(a^-)$ existe et vaut $f(a)$. On dit que f est **continue à droite en** a lorsque $f(a^+)$ existe et vaut $f(a)$.

Proposition 17.13. *Si f est continue à gauche et à droite en a, alors f est continue en a.*

Exemple 17.25. Soit $f :]u, v[\to \mathbb{R}$ croissante et $a \in]u, v[$. On rappelle que f admet des limites finies à gauche et à droite en a qui vérifient $f(a^-) \leqslant f(a) \leqslant f(a^+)$. Pour montrer que f est continue en a, il suffit alors de montrer que $f(a^-) \geqslant f(a)$ et $f(a) \geqslant f(a^+)$.

Remarque 17.7. On peut adapter la proposition en bord de domaine, si une seule des deux limites existe. Elle reste alors vraie.

Théorème 17.11 (Caractérisation séquentielle de la continuité)**.** *Soit $f : X \to \mathbb{R}$ et $a \in X$. f est continue en a si, et seulement si,*

$$\forall (x_n)_{n\in\mathbb{N}} \in X^{\mathbb{N}},\ x_n \xrightarrow[n\to+\infty]{} a \implies f(x_n) \xrightarrow[n\to+\infty]{} f(a)$$

Exemple 17.26. Les suites récurrentes de la forme $u_{n+1} = f(u_n)$ avec f continue. Voir le dernier paragraphe du chapitre "Réels et suites".

Définition 17.22 (Prolongement par continuité)**.** Soit $f : X \to \mathbb{R}$, $a \in \mathbb{R}$ et $b \in \mathbb{R}$ tels que $a \in \overline{X} \setminus X$ et $f(x) \xrightarrow[x\to a]{} b$. Alors f admet un unique prolongement à $X \cup \{a\}$ qui soit continu en a. C'est la fonction :

$$\begin{array}{rccc} f^\star & : & X \cup \{a\} & \to & \mathbb{R} \\ & & x & \mapsto & \begin{cases} f(x) \text{ si } x \in X \\ b \text{ si } x = a \end{cases} \end{array}$$

On l'appelle **le prolongement par continuité de f en a.**

Proposition 17.14 (Combinaison linéaire)**.** *Soit $f, g : X \to \mathbb{R}$ et $a \in X$ tels que f et g soient continues en a. Soit $\lambda, \mu \in \mathbb{R}$. Alors $\lambda f + \mu g$ est continue en a.*

Proposition 17.15 (Produit)**.** *Soit $f, g : X \to \mathbb{R}$ et $a \in X$ tels que f et g soient continues en a. Alors fg est continue en a.*

Proposition 17.16 (Quotient)**.** *Soit $f, g : X \to \mathbb{R}$ et $a \in X$ tels que f et g soient continues en a. Si $\forall x \in X,\ g(x) \neq 0$, alors $\dfrac{f}{g}$ est continue en a.*

Proposition 17.17 (Valeur absolue)**.** *Soit $f : X \to \mathbb{R}$ et $a \in X$ tels que f soient continue en a. Alors $|f|$ est continue en a.*

Proposition 17.18 (Minimum et maximum)**.** *Soit $f, g : X \to \mathbb{R}$ et $a \in X$ tels que f et g soient continues en a. Alors $x \mapsto \min(f(x), g(x))$ et $x \mapsto \max(f(x), g(x))$ sont continues en a.*

Théorème 17.12 (Composée)**.** *Soit $f : X \to \mathbb{R}$, $a \in X$, $Y \subset \mathbb{R}$ tel que $f(X) \subset Y$ et $g : Y \to \mathbb{R}$. Si f est continue en a et g est continue en $f(a)$, alors $g \circ f$ est continue en a.*

3 Extension globale

Dans tout le paragraphe, on fixe $X \subset \mathbb{R}$.

3.1 Continuité

Définition 17.23 (Continuité globale)**.** Soit $f : X \to \mathbb{R}$. f est dite **continue** lorsque f est continue en tout point de X.

Définition 17.24 (Continuité sur)**.** Soit $f : X \to \mathbb{R}$ et $A \subset X$. f est dite **continue sur** A lorsque $f_{|A}$ est continue.

Proposition 17.19 (Continuité induite)**.** *Soit $f : X \to \mathbb{R}$ et $A \subset X$. Si f est continue, alors f est continue sur A.*

Remarque 17.8. Attention : la réciproque est fausse ! Par exemple, la partie entière est continue sur $[0, 1[$ mais pas en tout point de $[0, 1[$. Elle est discontinue en 0. Toutefois, il suffit A soit un intervalle ouvert pour que la réciproque soit vraie.

Exemple 17.27. Si f est continue et admet un prolongement par continuité, alors ce prolongement est continu.

Théorème 17.13. *La continuité globale passe aux opérations algébriques suivantes :*

- *Combinaison linéaire*
- *Produit*
- *Quotient dont le dénominateur ne s'annule pas*
- *Valeur absolue*
- *Minimum et maximum*
- *Composition*

Exemple 17.28. $\mathcal{C}^0(X, \mathbb{R})$ (l'ensemble des fonctions de X dans $\mathbb{R}$ continues) est une sous-algèbre de $\mathbb{R}^X$.

Exemple 17.29. Toute fonction polynomiale est continue. Toute fonction rationnelle est continue.

Définition 17.25 (Uniforme continuité)**.** Soit $f : X \to \mathbb{R}$. f est dite **uniformément continue** lorsqu'on "peut choisir δ indépendamment de a" :

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall (x,y) ß X^2,\ |x-y| \leqslant \delta \implies |f(x) - f(y) \leqslant \varepsilon$$

Théorème 17.14 (Lipschitzienne $\implies$ Uniformément continue $\implies$ Continue)**.** *Si f est lipschitzienne, alors est elle uniformément continue. Si f est uniformément continue, alors elle est continue.*

Exemple 17.30. Toute fonction constante est 0-lipschitzienne, donc uniformément continue ?

Exemple 17.31. D'après le seconde inégalité triangulaire, la valeur absolue est 1-lipschitzienne donc elle est uniformément continue puis continue.

Remarque 17.9. Attention, les réciproques sont fausses. Par exemple, la racine carrée restreinte à $[0,1]$ est uniformément continue par le théorème de Heine (cf. la suite du chapitre), mais n'est pas lipschitzienne. Ensuite, la fonction carré est continue mais n'est pas uniformément continue.

Théorème 17.15 (Théorème de Heine)**.** *"Toute fonction continue sur un segment est uniformément continue." Soit $(a,b) \in \mathbb{R}^2$ tel que $a \leqslant b$ et $f : [a,b] \to \mathbb{R}$. Si f est continue, alors f est uniformément continue.*

3.2 Théorèmes généraux

Théorème 17.16 (Théorème des valeurs intermédiaires / TVI, version segment)**.** *Soit $(a,b) \in \mathbb{R}^2$ tel que $a \leqslant b$. Soit $f : [a,b] \to \mathbb{R}$ continue. Alors, f "prend toutes les valeurs entre $f(a)$ et $f(b)$". Précisément :*

- *Supposons que $f(a) \leqslant f(b)$. Alors*

$$\forall y \in [f(a), f(b)]\,,\ \exists x \in [a,b],\ f(x) = y$$

- *Supposons que $f(a) \geqslant f(b)$. Alors*

$$\forall y \in [f(b), f(a)]\,,\ \exists x \in [a,b],\ f(x) = y$$

Remarque 17.10. Bien que non constructif, ce théorème permet de prouver une surjectivité.

Corollaire 17.4 (Variantes du TVI)**.** *On a les variantes suivantes du TVI :*

- *Soit $(a,b) \in \overline{\mathbb{R}}^2$ tel que $-\infty < a < b \leqslant +\infty$ et $f : [a,b[\to \mathbb{R}$ continue. Soit $l_b \in \overline{\mathbb{R}}$ tel que $f(x) \xrightarrow[x \to b]{} l_b$.*
  - *Supposons que $f(a) < l_b$. Alors*

$$\forall y \in [f(a), l_b[,\ \exists x \in [a,b[,\ y = f(x)$$

  - *Supposons que $f(a) > l_b$. Alors*

$$\forall y \in]l_b, f(a)],\ \exists x \in [a,b[,\ y = f(x)$$

- *Soit* $(a,b) \in \overline{\mathbb{R}}^2$ *tel que* $-\infty \leqslant a < b < +\infty$ *et* $f :]a,b] \to \mathbb{R}$ *continue. Soit* $l_a \in \overline{\mathbb{R}}$ *tel que* $f(x) \xrightarrow[x\to a]{} l_a$.
  - *Supposons que* $l_a < f(b)$. *Alors*
  $$\forall y \in]l_a, f(b)],\ \exists x \in]a,b],\ y = f(x)$$
  - *Supposons que* $l_a > f(b)$. *Alors*
  $$\forall y \in [f(b), l_a[,\ \exists x \in]a,b],\ y = f(x)$$
- *Soit* $(a,b) \in \overline{\mathbb{R}}^2$ *tel que* $-\infty \leqslant a < b \leqslant +\infty$ *et* $f :]a,b[\to \mathbb{R}$ *continue. Soit* $l_b \in \overline{\mathbb{R}}$ *tel que* $f(x) \xrightarrow[x\to b]{} l_b$ *et* $l_a \in \overline{\mathbb{R}}$ *tel que* $f(x) \xrightarrow[x\to a]{} l_a$.
  - *Supposons que* $l_a < l_b$. *Alors*
  $$\forall y \in]l_a, l_b[,\ \exists x \in]a,b[,\ y = f(x)$$
  - *Supposons que* $l_a > l_b$. *Alors*
  $$\forall y \in]l_b, l_a[,\ \exists x \in]a,b[,\ y = f(x)$$

Remarque 17.11. Le TVI version segment et l'ensemble des ses variantes peuvent être appelés TVI.

Corollaire 17.5 (Existence d'une racine réel pour les polynômes de degré impair)**.** *La troisième variante montre que toute fonction polynomiale de degré impair admet au moins une racine réelle.*

Corollaire 17.6 (L'image d'un intervalle par une fonction continue est un intervalle)**.** *Soit* $f : X \to \mathbb{R}$ *continue et* $I \subset X$ *un intervalle. Alors* $f(I)$ *est un intervalle.*

Remarque 17.12. Dans le cas où f est continue et monotone, elle envoie les bornes de I sur les bornes de $f(I)$, et I et $f(I)$ ont alors même nature (ouvert, fermé, semi-ouvert). En pratique, un tableau de variations suffira pour justifier ce fait, mais on peut le montrer en faisant une double inclusion et en appliquant le TVI.

Théorème 17.17 (Théorème des bornes atteintes)**.** *Soit* $(a,b) \in \mathbb{R}^2$ *tel que* $a \leqslant b$ *et* $f : [a,b] \to \mathbb{R}$. *Si* f *est continue, alors elle admet un minimum et un maximum. Autrement dit, toute fonction continue sur un segment est bornée et atteint ses bornes.*

Remarque 17.13. On utilise plus souvent le fait que f soit bornée que le fait que ses bornes soient atteintes, mais ce dernier fait peut toujours servir.

Exemple 17.32. En notant m le minimum de f et M son maximum, on a par le TVI $\mathrm{Im}(f) = [m, M]$.

Théorème 17.18. *Soit* I *un* ***intervalle****,* $f : I \to \mathbb{R}$ *une fonction* ***continue*** *et* ***injective****. Alors* f *est strictement monotone.*

Remarque 17.14. N'oublier aucune des trois hypothèses : un intervalle, continue et injective ! Dès que l'on enlève une des trois hypothèses, le théorème tombe en défaut.

Théorème 17.19 (Théorème de la bijection réciproque / TBR, cas continu)**.** *Soit I et J des intervalles de $\mathbb{R}$ non triviaux et $f : I \to J$ une bijection. Si f est strictement monotone, alors f^{-1} est continue.*

Exemple 17.33. La fonction racine carrée est donc continue en appliquant le TBR à la fonction carré.

Exemple 17.34. En appliquant le théorème à f^{-1} qui a les mêmes variations que f, si la bijection f est strictement monotone, alors elle est continue.

4 Fonctions à valeurs complexes

On garde $X \subset \mathbb{R}$.

Définition 17.26 (Fonctions de base)**.** Soit $f : X \to \mathbb{C}$. On peut considérer :

$$\begin{array}{rcccc} \mathrm{Re}(f) & : & X & \to & \mathbb{R} \\ & & x & \mapsto & \mathrm{Re}(f(x)) \end{array}$$

$$\begin{array}{rcccc} \mathrm{Im}(f) & : & X & \to & \mathbb{R} \\ & & x & \mapsto & \mathrm{Im}(f(x)) \end{array}$$

$$\begin{array}{rcccc} |f| & : & X & \to & \mathbb{R} \\ & & x & \mapsto & |f(x)| \end{array}$$

$$\begin{array}{rcccc} \overline{f} & : & X & \to & \mathbb{R} \\ & & x & \mapsto & \overline{f(x)} \end{array}$$

Définition 17.27 (Fonctions à valeurs complexes bornée)**.** $f : X \to \mathbb{C}$ est dite **bornée** lorsque $|f|$ est majorée (au sens de $\mathbb{R}$).

Remarque 17.15. Lorsque f est à valeurs réelles, cela coïncide bien puisqu'alors f est bornée si, et seulement si $|f|$ est majorée, et car le module prolonge la valeur absolue.

Définition 17.28 (Limite d'une fonction à valeurs complexes)**.** Soit $a \in \overline{X}$, $b \in \mathbb{C}$ et $f : X \to \mathbb{C}$. On dit que $f(x)$ tend vers b quand x tend vers a, et on note $f(x) \xrightarrow[x \to a]{} b$ lorsque

$$\forall \varepsilon > 0,\ \exists V_a \in \mathcal{V}(a),\ \forall x \in X,\ x \in V_a \implies |f(x) - b| \leqslant \varepsilon$$

Remarque 17.16. Attention : pour les fonctions à valeurs complexes, pas de limite infinie, $+\infty$ n'a pas de sens dans $\mathbb{C}$ (sauf si la fonction est à valeurs réelles bien sûr).

Proposition 17.20. *Toute fonction à valeurs complexes qui tend vers $b \in \mathbb{C}$ lorsque x tend vers $a \in \overline{X}$ est bornée au voisinage de a.*

Théorème 17.20 (Théorème passerelle)**.** *Soit $a \in \overline{X}$, $(u, v) \in \mathbb{R}^2$ et $f : C \to \mathbb{C}$. On a :*

$$f(x) \xrightarrow[x \to a]{} u + iv \iff \begin{cases} Re(f)(x) \xrightarrow[x \to a]{} u \\ Im(f)(x) \xrightarrow[x \to a]{} v \end{cases}$$

Remarque 17.17. On déduit de ce théorème :

- **L'unicité de la limite**, quand elle existe : il suffit de passer par le cas réel et l'unicité de la forme algébrique.
- La **caractérisation séquentielle complexe de la limite** : il suffit de passer par la caractérisation séquentielle réelle et le théorème passerelle pour les suites complexes.

Définition 17.29 (Continuité locale dans $\mathbb{C}$)**.** Soit $a \in X$ et $f : X \to \mathbb{C}$. f est dite **continue en** a lorsque $f(x) \xrightarrow[x \to a]{} f(a)$.

Remarque 17.18. Cette définition prolonge celle donnée dans $\mathbb{R}$.

Remarque 17.19 (Utile)**.** S'il existe $b \in \mathbb{C}$ tel que $f(x) \xrightarrow[x \to a]{} b$, alors on a nécessairement $b = f(a)$, si bien que f est continue en a. Il suffit d'appliquer la proposition analogue du cas réel et le théorème passerelle.

Remarque 17.20. On prouve aisément les deux résultats suivants :

- **Caractérisation séquentielle complexe de la continuité**
- **Théorème passerelle pour la continuité locale** : f est continue en a si, et seulement si, $\mathrm{Re}(f)$ et $\mathrm{Im}(f)$ sont continues en a.

Définition 17.30 (Continuité globale dans $\mathbb{C}$)**.** $f : X \to \mathbb{C}$ est dite **continue** lorsqu'elle est continue en tout point de X.

Remarque 17.21. Cette définition prolonge celle donnée dans $\mathbb{R}$.

Remarque 17.22. On prouve aisément le **théorème passerelle pour la continuité globale** : f est continue si, et seulement si, $\mathrm{Re}(f)$ et $\mathrm{Im}(f)$ sont continues.

Théorème 17.21 (Opérations algébriques sur les fonctions à valeurs complexes)**.** *Les opérations algébriques suivantes sont valables pour les limites, la continuité locale et la continuité globale :*

- *Combinaison linéaire*
- *Produit*
- *Quotient dont le dénominateur ne s'annule pas*
- *Module*
- *Conjugué*
- *Composition*

Remarque 17.23. Attention : pour la composition, seule la "dernière" fonction par laquelle on compose par la gauche peut être à valeurs complexes !

Chapitre 18

Dérivation

I est toujours un intervalle non trivial de $\mathbb{R}$ et sauf précision du contraire, f est une fonction de I dans $\mathbb{R}$

1 Dérivation locale

Définition 18.1 (Taux d'accroissement en un point)**.** Soit $a \in I$. Le taux d'accroissement de f en a est la fonction

$$\begin{array}{rccc} \tau_{a,f} : & I \setminus \{a\} & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & \dfrac{f(x)-f(a)}{x-a} \end{array}$$

Définition 18.2 (Dérivabilité en un point et nombre dérivé)**.** Lorsque $\tau_{a,f}$ admet une limite finie l en a, f est dite dérivable en a. l est appelé le nombre dérivé de f en a et on le note $f'(a)$. On dit que f est dérivable à droite (resp. à gauche) en a lorsque c$\tau_{a,f}$ admet une limite finie à droite (resp. à gauche). Cette limite est appelée nombre dérivé de f en a à droite (resp. à gauche) et est notée $f'_d(a)$ (resp. $f'_g(a)$)

Proposition 18.1 (Caractère local de la dérivabilité)**.** *Soit $f : I \to \mathbb{R}$ et $g : J \to \mathbb{R}$ et $a \in I \cap J$. Supposons que f et g coïncident au voisinage de a. Alors f est dérivable en a si, et seulement si, g est dérivable en a. Dans ce cas, $g'(a) = f'(a)$.*

Définition 18.3 (Équation de la tangente)**.**

$$y - f(a) = f'(a)(x - a)$$

Proposition 18.2 (Dérivable $\Rightarrow$ continue)**.** *Si f est dérivable en a, alors f est continue en a.*

Remarque 18.1. La réciproque est fausse. Considérer la valeur absolue en 0.

Théorème 18.1 (Condition nécessaire d'extremum local)**.** *Soit $a \in \mathring{I}$ et f dérivable en a.* ***Si*** *f admet un extremum local en a,* ***alors*** *$f'(a) = 0$.*

Remarque 18.2. La réciproque est fausse. Considérer la fonction cube en 0.

Théorème 18.2 (Opérations algébriques sur la dérivabilité en un point)**.** *La dérivabilité en un point passe aux combinaisons linéaires, au produit, à l'inverse et au quotient si le dénominateur ne s'annule pas, et à la composition.*

Théorème 18.3 (Théorème de la bijection réciproque, cas dérivable en un point)**.** *Soit I et J des intervalles non triviaux de $\mathbb{R}$ et $f : I \to J$ une bijection strictement monotone. Soit a dans I. Supposons que f est dérivable en a. Alors f^{-1} est dérivable en $f(a)$ si, et seulement si, $f'(a) \neq 0$. Dans ce cas, on a $(f^{-1})'(f(a)) = \dfrac{1}{f'(a)}$*

2 Dérivation globale

Définition 18.4 (Dérivabilité globale)**.** $f : I \to \mathbb{R}$ est dite dérivable lorsqu'elle est dérivable en tout point de I. On définit alors f' la **fonction dérivée de** f qui à tout x de I associe $f'(x)$. On note parfois cette fonction $\dfrac{df}{dx}$, voire Df.

Remarque 18.3. Plus généralement, on s'autorise à considérer des fonctions définies sur des unions d'INT. Les définitions restent les mêmes. Exemple : la fonction tangente.

Corollaire 18.1 (Dérivable $\Rightarrow$ continue)**.** *Toute fonction dérivable est continue*

Définition 18.5 (Dérivabilité sur)**.** Soit $f : I \to \mathbb{R}$ et $J \subset I$ un autre INT. f est dite dérivable sur J lorsque la restriction de f à J est dérivable.

Théorème 18.4 (Opérations algébriques sur la dérivabilité globale)**.** *La dérivabilité globale passe aux opérations algébriques suivantes (sous réserve que les fonctions considérées soient dérivables et compatibles) :*

1. *Combinaison linéaire :* $(\lambda f + \mu g)' = \lambda f' + \mu g'$
2. *Produit :* $(fg)' = f'g + fg'$
3. *Inverse (si g ne s'annule pas) :* $\left(\dfrac{1}{f}\right)' = -\dfrac{f'}{f^2}$
4. *Quotient (si g ne s'annule pas) :* $\left(\dfrac{f}{g}\right)' = \dfrac{f'g - fg'}{g^2}$
5. *Composée (si les domaines sont compatibles) :* $(g \circ f)' = (g' \circ f) \times f'$

Corollaire 18.2. *Soit $f : I \to \mathbb{R}$ dérivable.*

1. *Soit $n \in \mathbb{N}^\star$. Alors f^n est dérivable et $(f^n)' = nf^{n-1}f'$*
2. *Soit $n \in \mathbb{Z}$. Supposons que f est à valeurs dans $\mathbb{R}^\star$. Alors f^n est dérivable et $(f^n)' = nf^{n-1}f'$*
3. *Soit $\alpha \in \mathbb{R}$. Supposons que f est à valeurs dans $\mathbb{R}_+^\star$. Alors f^α est dérivable et $(f^n)' = \alpha f^{\alpha-1}f'$*

Théorème 18.5 (Théorème de la bijection réciproque, cas dérivable)**.** *Soit I et J des intervalles non triviaux de $\mathbb{R}$ et $f : I \to J$ une bijection strictement monotone. Supposons que f est dérivable et que f' ne s'annule pas. Alors f^{-1} est dérivable et*

$$(f^{-1})' = \frac{1}{f' \circ f^{-1}}$$

3 Théorèmes généraux

Théorème 18.6 (Théorème de Rolle)**.** *Soit $(a,b) \in \mathbb{R}^2$ tel que $a < b$ et $f : [a,b] \to \mathbb{R}$ telle que*

1. *f est continue en tout point de $[a,b]$*
2. *f est dérivable en tout point de $]a,b[$*
3. *$f(a) = f(b)$*

Alors, il existe $c \in]a,b[$ tel que $f'(c) = 0$.

Remarque 18.4. Ce théorème est donc *a fortiori* vrai lorsque la fonction est dérivable. Ces conditions sont typiques d'une **fonction non injective**.

Théorème 18.7 (Théorème des accroissements finis / Égalité des accroissements finis)**.** *Soit $(a,b) \in \mathbb{R}^2$ tel que $a < b$ et $f : [a,b] \to \mathbb{R}$ telle que*

1. *f est continue en tout point de $[a,b]$*
2. *f est dérivable en tout point de $]a,b[$*

Alors, il existe $c \in]a,b[$ tel que $f'(c) = \dfrac{f(b) - f(a)}{b - a}$.

Remarque 18.5. Là encore, ce résultat est *a fortiori* vrai si la fonction est dérivable en tout point de l'intervalle.

Corollaire 18.3 (Inégalité des accroissements finis)**.** *Soit I un INT de $\mathbb{R}$, $\lambda \geqslant 0$ et $f : I \to \mathbb{R}$ telle que*

1. *f est continue en tout point de I*
2. *f est dérivable en tout point de $\mathring{I}$*
3. *$\forall x \in \mathring{I},\ |f'(x)| \leqslant \lambda$*

Alors, f est λ-lipschitzienne.

Exemple 18.1. Permet de montrer que $\forall x \in \mathbb{R},\ |\sin(x)| \leqslant |x|$. Utilisé dans la méthode de Héron sur l'approximation des racines carrées avec $u_{n+1} = \frac{1}{2}(u_n + \frac{k}{u_n})$

Théorème 18.8 (Équivalences entre signe de la dérivée et variations)**.** *Soit I un INT de $\mathbb{R}$ et $f : I \to \mathbb{R}$. Supposons que f est continue en tout point de I et dérivable en tout point de $\mathring{I}$. Alors :*

1. *f est croissante si, et seulement si, $\forall x \in \mathring{I},\ f'(x) \geqslant 0$*
2. *f est décroissante si, et seulement si, $\forall x \in \mathring{I},\ f'(x) \leqslant 0$*
3. *f est constante si, et seulement si, $\forall x \in \mathring{I},\ f'(x) = 0$*

Corollaire 18.4 (Condition suffisante de stricte montonie)**.** *Soit I un INT de $\mathbb{R}$ et $f : I \to \mathbb{R}$. Supposons que f est continue en tout point de I et dérivable en tout point de $\mathring{I}$. Alors :*

1. ***Si*** *$\forall x \in \mathring{I},\ f'(x) \geqslant 0$,* ***alors*** *$f$ est strictement croissante.*
2. ***Si*** *$\forall x \in \mathring{I},\ f'(x) \leqslant 0$,* ***alors*** *$f$ est strictement décroissante.*

Corollaire 18.5 (Raffinement de la CS)**.** *Soit I un INT de $\mathbb{R}$ et $f : I \to \mathbb{R}$. Supposons que f est continue en tout point de I et dérivable en tout point de $\mathring{I}$. Alors :*

1. ***Si*** $\forall x \in \mathring{I},\ f'(x) \geqslant 0$ *et le nombre de points d'annulation de* f' *est* ***fini, alors*** f *est strictement croissante.*
2. ***Si*** $\forall x \in \mathring{I},\ f'(x) \leqslant 0$ *et le nombre de points d'annulation de* f' *est* ***fini, alors*** f *est strictement décroissante.*

Théorème 18.9 (Théorème de la limite de la dérivée)**.** *Soit* I *un INT de* $\mathbb{R}$, $f : I \to \mathbb{R}$ *et* $a \in I$. *Supposons que* f *est continue en tout point de* I *et que* f *est dérivable en tout point de* $I \setminus \{a\}$. ***Si*** *il existe* $\lambda \in \mathbb{R}$ *tel que* $f'(x) \xrightarrow[x \neq a]{} \lambda$, ***alors*** *:*

1. f *est dérivable en* a
2. $f'(a) = \lambda$

Remarque 18.6. Réciproquement, si f est dérivable en a de dérivée λ, **on ne peut pas** en déduire que $f'(x) \xrightarrow[x \neq a]{} \lambda$. Comme contre-exemple, on pourra considérer la fonction qui à x associe $x^2 \sin(\frac{1}{x})$ prolongée par 0 en 0.

4 Fonctions de classe $\mathcal{D}^n$ et de classe $\mathcal{C}^n$

Définition 18.6 (Classe $\mathcal{D}^n$)**.** Soit $f : I \to \mathbb{R}$ et $n \in \mathbb{N}$. f est dite de classe $\mathcal{D}^n$ lorsqu'il existe une suite finie de fonctions de I dans $\mathbb{R}$ $(\varphi_i)_{0 \leqslant i \leqslant n}$ telle que $\varphi_0 = f$, chacune des φ_k soit dérivable pour $0 \leqslant k \leqslant n-1$ et que pour tout $0 \leqslant k \leqslant n-1$, on ait $\varphi_k' = \varphi_{k+1}$. On note $\mathcal{D}^n(I, \mathbb{R})$ l'ensemble des fonctions de I dans $\mathbb{R}$ de classe $\mathcal{D}^n$.

Définition 18.7 (Classe $\mathcal{D}^\infty$)**.** Soit $f : I \to \mathbb{R}$ et $n \in \mathbb{N}$. f est dite de classe $\mathcal{D}^\infty$ lorsque f est de classe $\mathcal{D}^n$ quel que soit $n \in \mathbb{N}$. On note $\mathcal{D}^\infty(I, \mathbb{R})$ l'ensemble des fonctions de I dans $\mathbb{R}$ de classe $\mathcal{D}^\infty$. On a alors :

$$\mathcal{D}^\infty = \bigcap_{n \in \mathbb{N}} \mathcal{D}^n(I, \mathbb{R})$$

Remarque 18.7. On a la chaîne infinie d'inclusion suivante :

$$\mathcal{D}^\infty(I, \mathbb{R}) \subset ... \subset \mathcal{D}^1(I, \mathbb{R}) \subset \mathcal{D}^0(I, \mathbb{R})$$

Définition 18.8 (Dérivée n-ème)**.** Soit $f \in \mathcal{D}^n(I, \mathbb{R})$. On définit la dérivée n-ème de f, notée $f^{(n)}$, ou bien encore $\dfrac{d^n f}{dx^n}$, voire plus rarement $D^n f$ en fixant une chaîne admissible quelconque (en fait, cette chaîne est unique) $(\varphi_0,\ ...,\ \varphi_n)$ puis en posant $f^{(n)} = \varphi_n$

Proposition 18.3 (Relation de Chasles sur la classe $\mathcal{D}^n$)**.** *Soit* $(p, q) \in \mathbb{N}^2$ *et* $f : I \to \mathbb{R}$. *Alors,* f *est de classe* $\mathcal{D}^{p+q}$ *si, et seulement si,* f *est de classe* $\mathcal{D}^p$ *et* $f^{(p)}$ *est de classe* $\mathcal{D}^q$. *Dans ce cas, on a* $f^{(p+q)} = (f^{(p)})^{(q)}$

Corollaire 18.6. *Soit* $f : I \to \mathbb{R}$. f *est de classe* $\mathcal{D}^\infty$ *si, et seulement si, elle admet une chaîne admissible infinie.*

Proposition 18.4. *Soit* $f : I \to \mathbb{R}$. *Si* f *est de classe* $\mathcal{D}^1$ *et si* f' *est de classe* $\mathcal{D}^\infty$, *alors* f *est de classe* $\mathcal{D}^\infty$.

Définition 18.9 (Classe $\mathcal{D}^n$ sur)**.** Soit $f : I \to \mathbb{R}$ et $J \subset I$ un autre intervalle non trivial. f est dite de classe $\mathcal{D}^n$ sur J lorsque la restriction de f à J est de classe $\mathcal{D}^n$.

Proposition 18.5 (Classe $\mathcal{D}^n$ induite)**.** *Soit $f \in \mathcal{D}^n(I,\mathbb{R})$ et $J \subset I$ un autre INT. Alors f est de classe $\mathcal{D}^n$ sur J et toutes les dérivées k-èmes pour $0 \leqslant k \leqslant n$ de la restriction de f à J coïncident avec celles de f.*

Proposition 18.6 (Caractère local de la classe $\mathcal{D}^n$)**.** *Soit $(I_\omega)_{\omega\in\Omega}$ une famille d'intervalles ouverts de $\mathbb{R}$ et $f : \bigcup_{\omega\in\Omega} \to \mathbb{R}$. Si pour tout $\omega \in \Omega$, f est de classe $\mathcal{D}^n$, alors f est de classe $\mathcal{D}^n$.*

Remarque 18.8. Les deux résultats qui précèdent sont aussi vrai pour la classe $\mathcal{D}^\infty$, cela s'en déduit immédiatement.

Définition 18.10 (Classe $\mathcal{C}^n$)**.** Soit $f : I \to \mathbb{R}$ et $n \in \mathbb{N}$. f est dite de classe $\mathcal{C}^n$ lorsque f est de classe $\mathcal{C}^n$ et $f^{(n)}$ est continue. On note $\mathcal{C}^n(I,\mathbb{R})$ l'ensemble des fonctions de I dans $\mathbb{R}$ de classe $\mathcal{C}^n$.

Définition 18.11 (Classe $\mathcal{C}^\infty$)**.** On pose

$$\mathcal{C}^\infty = \bigcap_{n\in\mathbb{N}} \mathcal{C}^n(I,\mathbb{R})$$

Les fonctions qui appartiennent à cet ensemble sont dites de classe $\mathcal{C}^\infty$.

Remarque 18.9. On a la chaîne infinie d'inclusion suivante :

$$\mathcal{C}^\infty(I,\mathbb{R}) \subset ... \subset \mathcal{C}^1(I,\mathbb{R}) \subset \mathcal{C}^0(I,\mathbb{R})$$

Remarque 18.10. On a $\mathcal{C}^\infty(I,\mathbb{R}) = \mathcal{D}^\infty(I,\mathbb{R})$

Proposition 18.7 (Relation de Chasles sur la classe $\mathcal{C}^n$)**.** *Soit $(p,q) \in \mathbb{N}^2$ et $f : I \to \mathbb{R}$. Alors, f est de classe $\mathcal{C}^{p+q}$ si, et seulement si, f est de classe $\mathcal{D}^p$ et $f^{(p)}$ est de classe $\mathcal{C}^q$. Dans ce cas, on a $f^{(p+q)} = (f^{(p)})^{(q)}$*

Théorème 18.10 (Combinaison linéaire)**.** *Soit f et g des fonctions de $\mathcal{D}^n(I,\mathbb{R})$ (resp. $\mathcal{C}^n(I,\mathbb{R})$). Soit λ et μ des réels. Alors*

1. *$\lambda f + \mu g \in \mathcal{D}^n(I,\mathbb{R})$ (resp. $\mathcal{C}^n(I,\mathbb{R})$)*
2. *$(\lambda f + \mu g)^{(n)} = \lambda f^{(n)} + \mu g^{(n)}$*

Le résultat est donc aussi valable pour la classe $\mathcal{C}^\infty$.

Théorème 18.11 (Produit)**.** *Soit f et g des fonctions de $\mathcal{D}^n(I,\mathbb{R})$ (resp. $\mathcal{C}^n(I,\mathbb{R})$). Alors*

1. *$fg \in \mathcal{D}^n(I,\mathbb{R})$ (resp. $\mathcal{C}^n(I,\mathbb{R})$)*
2. *$(fg)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} f^{(k)} g^{(n-k)}$ C'est ce qu'on appelle la **formule de Leibniz**.*

Le résultat est donc aussi valable pour la classe $\mathcal{C}^\infty$.

Théorème 18.12 (Quotient)**.** *Soit f et g des fonctions de $\mathcal{D}^n(I,\mathbb{R})$ (resp. $\mathcal{C}^n(I,\mathbb{R})$) telles que $\forall x \in I,\ g(x) \neq 0$. Alors $\dfrac{f}{g} \in \mathcal{D}^n(I,\mathbb{R})$ (resp. $\mathcal{C}^n(I,\mathbb{R})$). Le résultat est donc aussi valable pour la classe $\mathcal{C}^\infty$.*

Théorème 18.13 (Composition). *Soit $f \in \mathcal{D}^n(I,\mathbb{R})$ (resp. $\mathcal{C}^n(I,\mathbb{R})$) et $g \in \mathcal{D}^n(J,\mathbb{R})$ (resp. $\mathcal{C}^n(I,\mathbb{R})$) telles que $f(I) \subset J$. Alors $g \circ f \in \mathcal{D}^n(I,\mathbb{R})$ (resp. $\mathcal{C}^n(I,\mathbb{R})$). Le résultat est donc aussi valable pour la classe $\mathcal{C}^\infty$.*

Théorème 18.14 (Théorème de la bijection réciproque pour les fonctions de classe $\mathcal{C}^n$). *Soit I et J des intervalles non triviaux de $\mathbb{R}$ et $f : I \to J$ une bijection strictement monotone. Soit $n \in \mathbb{N}^\star$. Si f est de classe $\mathcal{C}^n$ et f' ne s'annule pas, alors f^{-1} est de classe $\mathcal{C}^n$. Le résultat est donc aussi valable pour la classe $\mathcal{C}^\infty$.*

5 Extension aux fonctions à valeurs complexes

Définition 18.12 (Dérivabilité des fonctions à valeurs complexes). Soit $f : I \to \mathbb{C}$. On définit de même la dérivabilité en un point, la dérivabilité à gauche et à droite en un point, et la dérivabilité globale. On définit de même les classes $\mathcal{C}^n$ et $\mathcal{D}^n$.

Proposition 18.8 (Dérivable $\Rightarrow$ continue). *Soit $f : I \to \mathbb{C}$ $a \in I$. Si f est dérivable en a, alors f est continue en a.*

Théorème 18.15 (Théorème passerelle). *On a les équivalences suivantes :*

1. *f est dérivable en a si, et seulement si, $Re(f)$ et $Im(f)$ sont dérivables en a. Dans ce cas, $f'(a) = [Re(f)]'(a) + i[Im(f)]'(a)$*
2. *f est dérivable si, et seulement si, $Re(f)$ et $Im(f)$ sont dérivables. Dans ce cas, $f' = [Re(f)]' + i[Im(f)]'$*
3. *Soit $n \in \mathbb{N}$. f est de classe $\mathcal{D}^n$ si, et seulement si, $Re(f)$ et $Im(f)$ sont de classe $\mathcal{D}^n$. Dans ce cas, $f^{(n)} = [Re(f)]^{(n)} + i[Im(f)]^{(n)}$*
4. *Soit $n \in \mathbb{N}$. f est de classe $\mathcal{C}^n$ si, et seulement si, $Re(f)$ et $Im(f)$ sont de classe $\mathcal{C}^n$. Dans ce cas, $f^{(n)} = [Re(f)]^{(n)} + i[Im(f)]^{(n)}$*
5. *f est de classe $\mathcal{C}^\infty$ si, et seulement si, $Re(f)$ et $Im(f)$ sont de classe $\mathcal{C}^\infty$.*

Théorème 18.16 (Opérations algébriques sur la dérivabilité des fonctions à valeurs complexes). *Les opérations algébriques suivantes sur la dérivabilité des fonctions à valeurs complexes sont valables : combinaison linéaire, produit, inverse et quotient lorsque le dénominateur ne s'annule pas, passage au conjugué (via théorème passerelle) et composée $g \circ f$ lorsque f est **à valeurs réelles**. Il en va de même pour les classes $\mathcal{D}^n$ et $\mathcal{C}^n$. La formule de Leibniz reste valable. **Attention**, le passage au module ne fonctionne pas, tout comme la valeur absolue sur $\mathbb{R}$.*

Théorème 18.17 (Dérivabilité de l'exponentielle complexe). *Soit $f : I \to \mathbb{C}$ dérivable. Alors, $\exp \circ f : I \to \mathbb{C}$ est dérivable et on a*

$$(\exp \circ f)' = (\exp \circ f) \times f'$$

Corollaire 18.7. *Soit $f \in \mathcal{D}^1(I,\mathbb{C})$.*

1. *Soit $n \in \mathbb{N}^\star$. Alors f^n est dérivable et $(f^n)' = nf^{n-1}f'$*
2. *Soit $n \in \mathbb{Z}$. Supposons que f est à valeurs dans $\mathbb{C}^\star$. Alors f^n est dérivable et $(f^n)' = nf^{n-1}f'$*
3. *Soit $\alpha \in \mathbb{C}$. Supposons que f est à valeurs dans $\mathbb{R}_+^\star$. Alors f^α est dérivable et $(f^n)' = \alpha f^{\alpha-1}f'$*

Théorème 18.18 (Inégalité des accroissements finis, cas complexe)**.** *Soit $f : I \to \mathbb{C}$ continue et dérivable en tout point intérieur de I. Supposons qu'il existe $\lambda \geqslant 0$ tel que pour tout x de l'intérieur de I, $|f'(x)| \leqslant \lambda$. Alors, f est λ-lipschitzienne.*

Théorème 18.19. *Soit $f : I \to \mathbb{C}$ continue et dérivable en tout point intérieur de I. Alors f est constante si, et seulement si, pour tout x de l'intérieur de I, $f'(x) = 0$.*

6 Fonctions convexes

Définition 18.13 (Fonction convexe)**.** $f : I \to \mathbb{R}$ est dite convexe lorsque

$$\forall (a,b) \in I^2,\ \forall t \in [0,1]\,,\ f((1-t)a+tb) \leqslant (1-t)f(a)+tf(b)$$

Théorème 18.20 (Inégalité de Jensen)**.** *Soit $f : I \to \mathbb{R}$ convexe, $n \in \mathbb{N}^\star$ et $\lambda_1,\ ...,\ \lambda_n$ des réels* ***positifs*** *tels que $\sum_{k=1}^{n} \lambda_k = 1$. Alors*

$$\forall (a_1,\ ...,\ a_n) \in I^n,\ f\left(\sum_{k=1}^{n} \lambda_k a_k\right) \leqslant \sum_{k=1}^{n} \lambda_k f(a_k)$$

Lemme 18.1 (Inégalité des pentes)**.** *Soit $f : I \to \mathbb{R}$ convexe et $(a,b,c) \in I^3$ tel que $a < b < c$. Alors*

$$\frac{f(b)-f(a)}{b-a} \leqslant \frac{f(c)-f(a)}{c-a} \leqslant \frac{f(c)-f(b)}{c-b}$$

Corollaire 18.8 (Caractérisation de la convexité par la croissance des taux d'accroissements)**.** *Soit $f : I \to \mathbb{R}$. f est convexe si, et seulement, pour tout $a \in I$, la fonction $\tau_{a,f}$ (taux de variation de f en a) est croissante.*

Corollaire 18.9 (Position relative des cordes pour les fonctions convexes)**.** *Une fonction convexe est située sous ses cordes entre les points d'intersection de la courbe avec les cordes et au-dessus en dehors.*

Théorème 18.21 (Dérivabilité à gauche et à droite des fonctions convexes)**.** *Soit $f : I \to \mathbb{R}$ convexe. Alors f est dérivable à gauche et à droite en tout point de l'intérieur de I. En conséquence, f est nécessairement continue en tout point de l'intérieur de I. De plus, l'ensemble des points de non dérivabilité est au plus dénombrable (en utilisant le lemme de l'inégalité des pentes et la densité de $\mathbb{Q}$ dans $\mathbb{R}$).*

Remarque 18.11. On ne peut rien dire en bord d'intervalle ! Par exemple, la fonction indicatrice du singleton nul restreinte à $\mathbb{R}_+$ est convexe, mais n'est pas continue en 0.

Théorème 18.22 (Caractérisation de la convexité pour les fonctions dérivables)**.** *Soit $f \in \mathcal{D}^1(I,\mathbb{R})$. f est convexe si, et seulement si, f' est croissante. Dans ce cas, la courbe de f est située au-dessus de toutes ses tangentes.*

Corollaire 18.10 (Caractérisation de la convexité pour les fonctions deux fois dérivables)**.** *Soit $f \in \mathcal{D}^2(I,\mathbb{R})$. f est convexe si, et seulement si, $f'' \geqslant 0$.*

Chapitre 19

Intégration

1 Fonctions en escalier et continues par morceaux

Dans tout le paragraphe, on fixe a et b des réels tels que $a < b$.

Définition 19.1 (Subdivision)**.** Une **subdivision** de $[a,b]$ est une famille $\sigma = (c_i)_{0\leqslant i\leqslant n}$ avec $n \in \mathbb{N}^\star$ telle que $a = c_0 < ... < c_n = b$. Le **pas de la subdivision** est défini par $\max\limits_{1\leqslant i\leqslant n}(c_i - c_{i-1})$.

Définition 19.2 (Subdivision régulière)**.** Pour $n \in \mathbb{N}^\star$, la subdivision $\left(a + i\dfrac{b-a}{n}\right)_{0\leqslant i\leqslant n}$ est appelée **subdivision régulière**.

Définition 19.3 (Subdivision plus fine)**.** Soit σ et σ' des subdivisions de $[a,b]$. σ est dite **plus fine** que σ' lorsque tous les points de σ' sont des points de σ.

Proposition 19.1 (Réunion de subdivisions)**.** *Soit σ_1 et σ_2 des subdivisions de $[a,b]$. En réunissant les points de σ_1 et σ_2, on obtient une subdivision de $[a,b]$ qu'on peut noter de façon NS $\sigma_1 \vee \sigma_2$. Cette subdivision est à la fois plus fine que σ_1 et que σ_2.*

Définition 19.4 (Fonction en escalier)**.** Une fonction $\varphi : [a,b] \to \mathbb{R}$ est dite **en escalier** lorsqu'il existe une subdivision $\sigma = (c_i)_{0\leqslant i\leqslant n}$ telle que pour tout $i \in [\![1,n]\!]$, la restriction de φ à $]c_{i-1}, c_i[$ soit constante. On dit alors que σ est une subdivision **adaptée** à φ. On notera $\mathcal{E}([a,b],\mathbb{R})$ l'ensemble des fonctions de $[a,b]$ dans $\mathbb{R}$ en escalier.

Remarque 19.1. Si on besoin d'expliciter la valeur constante de φ sur $]c_{i-1}, c_i[$, on peut choisir $\varphi\left(\dfrac{c_{i-1}+c_i}{2}\right)$

Remarque 19.2. Toute subdivision plus fine d'une subdivision adaptée à une fonction en escalier reste adaptée à cette fonction en escalier.

Proposition 19.2. *$\mathcal{E}([a,b],\mathbb{R})$ est une sous-algèbre de $\mathbb{R}^{[a,b]}$*

Définition 19.5 (Fonction continue par morceaux)**.** Une fonction $f : [a,b] \to \mathbb{R}$ est dite **continue par morceaux** lorsqu'il existe une subdivision $\sigma = (c_i)_{0\leqslant i\leqslant n}$ telle que pour tout $i \in [\![1,n]\!]$, la restriction de f à $]c_{i-1}, c_i[$ soit continue et admette des limites finies à droite en c_{i-1} et à gauche en c_i. On notera $CPM([a,b],\mathbb{R})$ l'ensemble des fonctions de $[a,b]$ dans $\mathbb{R}$ continues par morceaux.

Proposition 19.3. *(En escalier ⇒ CPM) Toute fonction en escalier est continue par morceaux. Formellement,* $\mathcal{E}([a,b],\mathbb{R}) \subset CPM([a,b],\mathbb{R})$

Proposition 19.4 (CPM ⇒ bornée)**.** *Toute fonction continue par morceaux est bornée.*

Définition 19.6 (Fonctions continues par morceaux sur un INT quelconque)**.** Soit I un INT de $\mathbb{R}$. On dit qu'une fonction $f : I \to \mathbb{R}$ est continue par morceaux lorsque sa restriction à tout segment non trivial de I est continue par morceaux.

Théorème 19.1 (Théorème d'approximation)**.** *Soit* $f \in CPM([a,b],\mathbb{R})$ *et* $\varepsilon > 0$*. Alors, il existe une fonction* $\varphi \in \mathcal{E}([a,b],\mathbb{R})$ *telle que*

$$\forall x \in [a,b],\ |f(x) - \varphi(x)| \leqslant \varepsilon$$

Remarque 19.3. Utile pour démontrer le lemme de Riemann-Lebesgue dans le cas des fonctions CPM

Définition 19.7 (Norme infinie)**.** Soit $f : [a,b] \to \mathbb{R}$ bornée. Sa **norme infinie** est le réel défini par

$$||f||_\infty = \sup_{x\in[a,b]} |f(x)|$$

Définition 19.8 (Convergence uniforme)**.** Soit $f : [a,b] \to \mathbb{R}$ bornée et $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions de $[a,b]$ dans $\mathbb{R}$ bornées. Les deux assertions suivantes sont équivalentes :

1. $||f - f_n||_\infty \xrightarrow[n\to+\infty]{} 0$
2. $\forall \varepsilon > 0,\ \exists N \in \mathbb{N},\ \forall n \geqslant N,\ \forall x \in [a,b],\ |f(x) - f_n(x)| \leqslant \varepsilon$

Lorsqu'une de ces assertion est vérifiée, on dit que la suite de fonctions (f_n) **converge uniformément** vers f. On le note

$$f_n \xrightarrow[n\to+\infty]{CVU} f$$

Théorème 19.2 (Approximation des fonctions continues par morceaux par des fonctions en escalier)**.** *Soit* $f \in CPM([a,b],\mathbb{R})$*. Alors, il existe une suite de fonctions en escalier* $(\varphi_n)_{n\in\mathbb{N}}$ *telle que*

$$\varphi_n \xrightarrow[n\to+\infty]{CVU} f$$

2 Définition de l'intégrale

Définition 19.9 (Intégrale d'une fonction en escalier)**.** Soit $\varphi \in \mathcal{E}([a,b],\mathbb{R})$. Fixons $\sigma = (c_i)_{0\leqslant i\leqslant n}$ une subdivision adaptée à φ. Pour tout $i \in [\![1,n]\!]$, on note λ_i la valeur constante de φ sur $]c_{i_1}, c_i[$. On définit alors l'intégrale de φ par

$$I(\varphi,\sigma) = \sum_{i=1}^{n} \lambda_i (c_i - c_{i_1})$$

qui est indépendant du choix de la subdivision. On note plus souvent ce scalaire

$$\int_{[a,b]} \varphi$$

Proposition 19.5 (Linéarité de l'intégrale pour les fonctions en escalier)**.** *Soit λ et μ des réels ainsi que φ et ψ des fonctions en escalier. Alors*

$$\int_{[a,b]} (\lambda\varphi + \mu\psi) = \lambda \int_{[a,b]} \varphi + \mu \int_{[a,b]} \psi$$

Définition 19.10 (Intégrale "sur" pour les fonctions en escalier)**.** Soit $\varphi \in \mathcal{E}([a,b],\mathbb{R})$ et $(c,d) \in [a,b]^2$ tel que $c < d$. On définit l'intégrale de φ sur $[c,d]$ par

$$\int_{[c,d]} \varphi = \int_{[c,d]} \varphi_{|[c,d]}$$

Proposition 19.6 (Relation de Chasles pour les fonctions en escalier)**.** *Soit $\varphi \in \mathcal{E}([a,b],\mathbb{R})$ et $c \in]a,b[$. On a*

$$\int_{[a,b]} \varphi = \int_{[a,c]} \varphi + \int_{[c,b]} \varphi$$

Proposition 19.7 (Inégalité triangulaire pour les fonctions en escalier)**.** *Soit $\varphi \in \mathcal{E}([a,b],\mathbb{R})$. On a*

1. $|\varphi| \in \mathcal{E}([a,b],\mathbb{R})$
2. $$\left|\int_{[a,b]} \varphi\right| \leqslant \int_{[a,b]} |\varphi|$$

Proposition 19.8 (Positivité et croissance de l'intégrale pour les fonctions en escalier)**.** *On a les deux propriétés suivantes*

1. *Positivité :*
$$\forall \varphi \in \mathcal{E}([a,b],\mathbb{R}),\ \varphi \geqslant 0 \Rightarrow \int_{[a,b]} \varphi \geqslant 0$$
2. *Croissance :*
$$\forall (\varphi,\psi) \in (\mathcal{E}([a,b],\mathbb{R}))^2,\ \varphi \geqslant \psi \Rightarrow \int_{[a,b]} \varphi \geqslant \int_{[a,b]} \psi$$

Proposition 19.9 (Insensibilité de l'intégrale pour les fonctions en escalier)**.** *Soit $(\varphi,\psi) \in (\mathcal{E}([a,b],\mathbb{R}))^2$. Si φ et ψ ne diffèrent qu'en un nombre fini de points, alors*

$$\int_{[a,b]} \varphi = \int_{[a,b]} \psi$$

Définition 19.11 (Intégrale d'une fonction continue par morceaux)**.** Soit $f \in CPM([a,b],\mathbb{R})$. On fixe une suite $(\varphi_n)_{n\in\mathbb{N}}$ de fonctions en escalier telle que $\varphi_n \xrightarrow[n\to+\infty]{CVU} f$ puis on définit l'intégrale de f par

$$\int_{[a,b]} f = \lim_{n\to+\infty} \int_{[a,b]} \varphi_n$$

Proposition 19.10 (Linéarité de l'intégrale pour les fonctions CPM)**.** *Soit λ et μ des réels ainsi que f et g des fonctions continues par morceaux. Alors*

$$\int_{[a,b]} (\lambda f + \mu g) = \lambda \int_{[a,b]} f + \mu \int_{[a,b]} g$$

Définition 19.12 (Intégrale "sur" pour les fonctions CPM)**.** Soit $f \in CPM([a,b],\mathbb{R})$ et $(c,d) \in [a,b]^2$ tel que $c < d$. On définit l'intégrale de f sur $[c,d]$ par

$$\int_{[c,d]} f = \int_{[c,d]} f_{|[c,d]}$$

Proposition 19.11 (Relation de Chasles pour les fonctions CPM)**.** *Soit $f \in CPM([a,b],\mathbb{R})$ et $c \in]a,b[$. On a*

$$\int_{[a,b]} f = \int_{[a,c]} f + \int_{[c,b]} f$$

Proposition 19.12 (Inégalité triangulaire pour les fonctions CPM)**.** *Soit $f \in CPM([a,b],\mathbb{R})$. On a*

1. $|f| \in CPM([a,b],\mathbb{R})$
2. $$\left|\int_{[a,b]} f\right| \leqslant \int_{[a,b]} |f|$$

Proposition 19.13 (Positivité et croissance de l'intégrale pour les fonctions CPM)**.** *On a les deux propriétés suivantes*

1. *Positivité :*
$$\forall f \in CPM([a,b],\mathbb{R}),\ f \geqslant 0 \Rightarrow \int_{[a,b]} f \geqslant 0$$
2. *Croissance :*
$$\forall (f,g) \in (CPM([a,b],\mathbb{R}))^2,\ f \geqslant g \Rightarrow \int_{[a,b]} f \geqslant \int_{[a,b]} g$$

Proposition 19.14 (Insensibilité de l'intégrale pour les fonctions CPM)**.** *Soit $(f,g) \in (CPM([a,b],\mathbb{R}))^2$. Si φ et ψ ne diffèrent qu'en un nombre fini de points, alors*

$$\int_{[a,b]} f = \int_{[a,b]} g$$

Proposition 19.15 (Stricte positivité et stricte croissance de l'intégrale pour des fonctions continues.)**.** *On a les deux résultats suivant :*

1. *Stricte positivité : Soit $f \in \mathcal{C}^0([a,b],\mathbb{R})$. Si $f > 0$, alors*
$$\int_{[a,b]} f > 0$$
2. *Stricte croissance : Soit $(f,g) \in (\mathcal{C}^0([a,b],\mathbb{R}))^2$. Si $f > g$, alors*
$$\int_{[a,b]} f > \int_{[a,b]} g$$

Proposition 19.16. *Soit $f \in \mathcal{C}^0([a,b],\mathbb{R})$ de signe constant. Si $\int_{[a,b]} f = 0$, alors $f = 0$*

Théorème 19.3 (Sommes de Riemann)**.** *Soit $f \in CPM([a,b],\mathbb{R})$. Posons*

$$\forall n \in \mathbb{N}^\star,\ S_n = \frac{b-a}{n}\sum_{k=0}^{n-1} f\left(a + k\frac{b-a}{n}\right)$$

et

$$\forall n \in \mathbb{N}^\star,\ S'_n = \frac{b-a}{n}\sum_{k=1}^{n} f\left(a + k\frac{b-a}{n}\right)$$

Alors

$$S_n \xrightarrow[n\to+\infty]{} \int_{[a,b]} f$$

et

$$S'_n \xrightarrow[n\to+\infty]{} \int_{[a,b]} f$$

Définition 19.13 (Valeur moyenne)**.** Soit $f \in CPM([a,b],\mathbb{R})$. On appelle **valeur moyenne** de f le réel

$$\mu(f) = \frac{1}{b-a}\int_{[a,b]} f$$

Définition 19.14 (Intégrale orientée)**.** Soit $f \in CPM(I,\mathbb{R})$ avec I un INT de $\mathbb{R}$ et $(a,b) \in I^2$. On définit

1. Si $a < b$,
$$\int_a^b f(t)dt = \int_{[a,b]} f$$
2. Si $a = b$,
$$\int_a^b f(t)dt = 0$$
3. Si $a > b$,
$$\int_a^b f(t)dt = -\int_{[a,b]} f$$

Proposition 19.17 (Validité des propriétés de l'intégrale pour l'intégrale orientée)**.** *La linéarité, la relation de Chasles étendue à $(a,b,c) \in I^3$ et le théorème sur les sommes de Riemann restent valables pour l'intégrale orientée. En revanche, la positivité, la croissance, l'inégalité triangulaire ne sont valables que si $a \leqslant b$. La stricte positivité et la stricte croissance ne restent vraies que si $a < b$. Ces hypothèses devront **toujours** être mentionnées avant l'utilisation de ces théorèmes.*

Proposition 19.18 (Extension aux fonctions à valeurs complexes)**.** *On étend de manière naturelle les définitions aux fonctions à valeurs complexes. Tous les théorèmes restent vrais, sauf les inégalités, la seule inégalité qui reste vraie est l'inégalité triangulaire (car absence de relation d'ordre prolongeant celle de $\mathbb{R}$ sur $\mathbb{C}$)*

Théorème 19.4 (Théorème passerelle)**.** *Soit $f \in CPM([a,b],\mathbb{C})$. Alors $Re(f)$ et $Im(f)$ sont continues par morceaux et à valeurs dans $\mathbb{R}$, et on a*

$$\int_{[a,b]} f = \int_{[a,b]} Re(f) + i \int_{[a,b]} Im(f)$$

Autrement dit, on a $Re\left(\int_{[a,b]} f\right) = \int_{[a,b]} Re(f)$ *et* $Im\left(\int_{[a,b]} f\right) = \int_{[a,b]} Im(f)$

3 Intégration et dérivation

Dans tout le paragraphe, on fixe I un intervalle non trivial de $\mathbb{R}$ et $f : I \to \mathbb{K}$ avec $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

Définition 19.15 (Primitive)**.** Une primitive de f est une fonction $F \in \mathcal{D}^1(I,\mathbb{K})$ telle que $F' = f$.

Proposition 19.19 (Unicité à constante près des primitives)**.** *Si elles existent, les primitives de f diffèrent toutes d'une constante additive.*

Théorème 19.5 (Théorème fondamental de l'analyse)**.** *Supposons que f est continue. Soit $a \in I$ quelconque. Alors la fonction*

$$\begin{array}{rccc} Fv : & I & \longrightarrow & \mathbb{K} \\ & x & \longmapsto & \displaystyle\int_a^x f(t)dt \end{array}$$

est l'unique primitive de f qui s'annule en a.

Méthode 19.1 (Calcul d'une primitive)**.** Pour calculer une primitive d'une fonction, il suffit de s'assurer que celle-ci est continue, puis de choisir le a le plus simple possible, de faire monter x dans la borne puis de calculer l'intégrale grâce aux méthodes présentées plus loin.

Remarque 19.4. Attention, toute primitive ne s'obtient pas par le TFA. Par exemple, en considérant la fonction nulle, changer les bornes ne fera qu'ajouter 0, et ne nous donnera jamais les fonctions constantes.

Remarque 19.5. On peut trouver dommage de demander la forte hypothèse de continuité, mais on ne peut pas vraiment faire plus simple. Il existe même des fonctions discontinues qui admettent des primitives (considérer la dérivée de $x \mapsto x^2 \sin(\frac{1}{x})$) prolongée par 0 en 0, puisque la fonction précédente est de classe $\mathcal{D}^1$ mais pas de classe $\mathcal{C}^1$.

Remarque 19.6 (HP)**.** Si f est CPM et admet une primitive, alors elle est continue. Pour montrer cela, il faut appliquer le TLD aux bords de la subdivision.

Théorème 19.6 (Calcul direct à l'aide d'une primitive)**.** *Soit $(a,b) \in I^2$. Supposons que f soit continue et fixons F une primitive de f. On a*

$$\int_a^b f(t)dt = F(b) - F(a)$$

On note, pour alléger l'écriture $[F(t)]_a^b$: on parle de ***terme tout intégré****.*

Corollaire 19.1. *Supposons que f est de classe $\mathcal{C}^1$. Alors*

$$\forall (a,b) \in I^2,\ \int_a^b f'(t)dt = f(b) - f(a)$$

Remarque 19.7. Pour montrer l'invariance par T-translation d'une fonction continue T-périodique, il faut montrer que la fonction $x \mapsto \int_x^{x+T} f(t)dt$ est constante. Pour cela, l'exprimer à l'aide d'une primitive puis dériver pour obtenir la constance et évaluer en 0 pour conclure. Dans la cas CPM, il faut fixer une subdivision de l'intervalle de longueur T puis coincer un multiple de T dans cette subdivision. Ensuite, on travaille sur chacun des segments où la restriction est continue puis on recolle par la relation de Chasles.

Théorème 19.7 (Intégration par parties)**.** *Soit g ayant les mêmes hypothèses que f. Supposons que f et g soient de classe $\mathcal{C}^1$. Alors*

$$\forall (a,b) \in I^2,\ \int_a^b f'(t)g(t)dt = [f(t)g(t)]_a^b - \int_a^b f(t)g'(t)dt$$

Théorème 19.8 (Changement de variable)**.** *Soit I,J des INT de $\mathbb{R}$, $f \in \mathcal{C}^0(I,\mathbb{R})$ et $\varphi \in \mathcal{C}^1(J,I)$. Alors*

$$\int_{\varphi(a)}^{\varphi(b)} f(t)dt = \int_a^b f(\varphi(t))\varphi'(t)dt$$

Remarque 19.8. On a les changements de variables usuels dans les cas suivants :

1. $\sqrt{1-t^2}$: on utilise $t = \cos(u)$ ou $t = \sin(u)$
2. $\sqrt{t^2-1}$: on utilise $t = \cosh(u)$
3. $\sqrt{t^2+1}$: on utilise $t = \sinh(u)$

Proposition 19.20 (Règles de Bioche, HP)**.** *On souhaite calculer $\int F(\sin(\theta),\cos(\theta),\tan(theta))d\theta$ où F est une fraction rationnelle. On pose $\omega(\theta) = F(\theta)d\theta$. Les règles de Bioche énoncent des changements de variable à effectuer pour accélérer le calcul :*

1. *Supposons que $\omega(-\theta) = \omega(\theta)$: on pose alors $u = \cos(\theta)$*
2. *Supposons que $\omega(\pi-\theta) = \omega(\theta)$: on pose alors $u = \sin(\theta)$*
3. *Supposons que $\omega(\pi+\theta) = \omega(\theta)$: on pose alors $u = \tan(\theta)$*
4. *Supposons que deux des propriétés précédentes soient vérifiées. Alors, la troisième l'est aussi et on pose $u = \cos(2\theta)$*
5. *Si aucune des trois propriétés n'est vérifiée, on pose $u = \tan\left(\dfrac{\theta}{2}\right)$ et on utilise les formules $\cos(\theta) = \dfrac{1-u^2}{1+u^2}$, $\sin(\theta) = \dfrac{2u}{1+u^2}$ et $\tan(\theta) = \dfrac{2u}{1-u^2}$*

Dans tous les cas, on se ramène à un calcul d'intégrale fraction rationnelle qu'on sait de toute manière toujours réaliser. Les 4 premiers changements de variables sont en général beaucoup plus rapide que le dernier.

Remarque 19.9. Bien que HP, on peut toujours utiliser ces règles sans justification car ce n'est pas réellement un théorème mais une astuce. Pour retenir les trois premiers changements de variable, remarquer qu'il correspondent aux mêmes invariances trigonométriques que celles supposées sur la fonction ω. Pour le dernier, il faut absolument connaître ses formules de trigonométrie. Le 4ème est en revanche beaucoup plus rare.

Corollaire 19.2 (Parité et imparité, HP)**.** *Supposons que I soit centré en 0 et que f soit continue.*

1. *Supposons que f soit paire. Alors*

$$\forall a \in I,\ \int_{-a}^{a} f(t)dt = 2\int_{0}^{a} f(t)dt$$

2. *Supposons que f soit impaire. Alors*

$$\forall a \in I,\ \int_{-a}^{a} f(t)dt = 0$$

Démonstration. Couper en deux par la relation de Chasles et utiliser un changement de variable $u = -t$. □

Remarque 19.10. On a des résultats similaires pour les fonctions continues qui présentent des symétries en un certain réel a_0, qui se démontrent de la même manière.

Théorème 19.9 (Formule de Taylor avec reste intégral)**.** *Soit $a \in I$. Soit $n \in \mathbb{N}$ et supposons que $f \in \mathcal{C}^{n+1}(I, \mathbb{K})$. Alors*

$$\forall x \in I,\ f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + \int_a^x \frac{(x-t)^n}{n!} f^{(n+1)}(t)dt$$

Théorème 19.10 (Inégalité de Taylor-Lagrange)**.** *Soit $a \in I$. Soit $n \in \mathbb{N}$ et supposons que $f \in \mathcal{C}^{n+1}(I, \mathbb{K})$. Soit $x \in I$ et M_{n+1} un majorant de $|f^{(n+1)}|$ sur $[min(a,x), max(a,x)]$. Alors*

$$\left| f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k \right| \leqslant M_{n+1} \frac{|x-a|^{n+1}}{(n+1)!}$$

Remarque 19.11. Lorsque M_{n+1} ne dépend pas de n, on peut obtenir l'expression de $f(x)$ comme somme d'une série par croissances comparées. C'est par exemple le cas avec la fonction exp ou les fonctions sin et cos.

4 Compléments, HP

4.1 Retour sur les sommes de Riemann

Ici, nous nous proposons de démontrer le théorème sur les sommes de Riemann dans le cas le plus général, c'est-à-dire avec une fonction continue par morceaux et des subdivisions potentiellement irrégulières.

Théorème 19.11. *Soit f continue par morceaux sur $[a,b]$, et $\sigma = (a_k)_{0\leqslant k\leqslant n}$ une subdivision de $[a,b]$. Si le pas de σ tend vers 0 alors*

$$\sum_{k=0}^{n-1}(a_{k+1}-a_k)f(a_k) \to \int_{[a,b]} f$$

Démonstration. Dans un premier temps, prouvons le théorème pour une fonction en escalier. On se donne $(c_i)_{0\leqslant i\leqslant N}$ une subdivision adaptée à φ. De la même manière que dans le cours, on obtient

$$\left|\int_{[a,b]}\varphi - \sum_{k=0}^{n-1}(a_{k+1}-a_k)\varphi(a_k)\right| \leqslant \sum_{k=0}^{n-1}\int_{[a_k,a_{k+1}]}|\varphi-\varphi(a_k)|$$

Pour $i\in[\![0,N]\!]$ fixé, si on considère $K(i) := \{k\in[\![0,n-1]\!] \mid a_k\leqslant c_i < a_{k+1}\}$, on a $|K(i)|\leqslant 1$ car les $[a_k,a_{k+1}[$ sont deux à deux disjoints. Donc si on pose :

$$K := \bigcup_{0\leqslant i\leqslant N} K(i)$$

on a $|K|\leqslant N+1$. Maintenant, soit $k\in[\![0,n-1]\!]\backslash K$ fixé, et posons

$$i_0 = \max\{i\in[\![0,N]\!] \mid c_i\leqslant a_k\}$$

Comme $[a_k,a_{k+1}[$ ne contient aucun c_i, on a $c_{i_0} < a_k < a_{k+1}\leqslant c_{i_0+1}$. En particulier, $[a_k,a_{k+1}[\subset]C_{i_0},c_{i_0+1}[$ donc φ est constante sur $[a_k,a_{k+1}]$ égale à $\varphi(a_k)$. Puisque (a_k,a_{k+1}) est adaptée à $|\varphi-\varphi(a_k)|$ on en déduit

$$\int_{[a_k,a_{k+1}]}|\varphi-\varphi(a_k)| = 0$$

De plus, φ ne prend qu'un nombre fini de valeurs, donc $|\varphi|$ admet un majorant M. Finalement, si on note $p(\sigma)$ le pas de σ, alors

$$\left|\int_{[a,b]}\varphi - \sum_{k=0}^{n-1}(a_{k+1}-a_k)\varphi(a_k)\right| \leqslant \sum_{k\in K}\int_{[a_k,a_{k+1}]}2M \leqslant 2M(N+1)p(\sigma)\to 0$$

Dans un second temps, passons à f continue par morceaux. Soit $\varepsilon>0$, et donnons-nous φ en escalier telle que $|f-\varphi|\leqslant \dfrac{\varepsilon}{3(b-a)}$. On a

$$\begin{aligned}\left|\int_{[a,b]} f - \sum_{k=0}^{n-1}(a_{k+1}-a_k)f(a_k)\right| &\leqslant \int_{[a,b]}|f-\varphi| + \left|\left(\int_{[a,b]}\varphi\right) - \sum_{k=0}^{n-1}(a_{k+1}-a_k)\varphi(a_k)\right| \\ &\quad + \sum_{k=0}^{n-1}(a_{k+1}-a_k)|\varphi(a_k)-f(a_k)| \\ &\leqslant \frac{2}{3}\varepsilon + \left|\left(\int_{[a,b]}\varphi\right) - \sum_{k=0}^{n-1}(a_{k+1}-a_k)\varphi(a_k)\right|\end{aligned}$$

Enfin, pour $p(\sigma)$ assez petit, la dernière valeur absolue est $\leqslant \dfrac{\varepsilon}{3}$. □

4.2 Retour sur la construction de l'intégrale de Riemann

La construction de l'intégrale de Riemann s'inscrit dans un cadre beaucoup plus vaste (et très fécond !) que nous proposons maintenant. Il s'agit d'une technique de construction maintes fois utilisée par les mathématiciens.

Avant tout nous avons besoin de quelques définitions. Dans toute cette partie, E est un espace vectoriel réel ou complexe.

Définition 19.16 (Norme)**.** Une **norme** sur E est une application $||\cdot|| : E \to \mathbb{R}$, qui vérifie les axiomes suivants :

- Séparation : $\forall x \in E, ||x|| = 0 \Rightarrow x = 0$.
- Homogénéité : $\forall (\lambda, x) \in \mathbb{K} \times E, ||\lambda x|| = |\lambda| \cdot ||x||$.
- Inégalité triangulaire : $\forall (x, y) \in E^2, ||x + y|| \leqslant ||x|| + ||y||$.

Exemple 19.1. La valeur absolue et le module sont respectivement des normes sur $E = \mathbb{R}$ et sur $E = \mathbb{C}$.

Dans le cadre d'un espace vectoriel normé (evn en abrégé), on peut ensuite définir les notions de suite convergente et de suite de Cauchy au sens usuel. Par exemple, on a le très important

Définition 19.17. Soit $[a, b]$ un segment de $\mathbb{R}$ et posons $\mathbb{K} = \mathbb{R} ou \mathbb{C}$. Alors l'ensemble $\mathcal{B}([a, b], \mathbb{K})$ des fonctions bornées de $[a, b]$ dans $\mathbb{K}$ est un $\mathbb{K}$-ev et peut le munir de la norme uniforme, définie de la manière suivante :

$$||f||_\infty = \sup_{x \in [a,b]} |f(x)|$$

Démonstration. Déjà le nombre $||f||_\infty$ est bien défini, car $[a, b]$ est non vide et $|f|$ est majorée. Ensuite on a bien $||f||_\infty \geqslant 0$, avec égalité ssi f est la fonction nulle. La semi-linéarité est facile à prouver. On commence par écrire

$$(\forall x \in [a, b],\ |\lambda f(x)| \leqslant |\lambda| \cdot ||f||_\infty) \text{ donc } ||\lambda f||_\infty \leqslant |\lambda| \cdot ||f||_\infty$$

Puis si $\lambda = 0$ on obtient $||0 \cdot f||_\infty = 0$, et sinon on applique le résultat à $\frac{1}{\lambda}$ et λf pour obtenir l'inégalité inverse. Enfin pour l'inégalité triangulaire on écrit

$$(\forall x \in [a, b],\ |f(x) + g(x)| \leqslant ||f||_\infty + ||g||_\infty) \text{ donc } ||f + g||_\infty \leqslant ||f||_\infty + ||g||_\infty$$

Par ailleurs, on sait que $\mathcal{B}([a, b], \mathbb{K})$ est un sev de $\mathbb{K}^{[a,b]}$. □

Définition 19.18 (Adhérence d'une partie)**.** Conformément au cas réel, nous dirons que l'adhérence $\overline{A}$ d'une partie $A \subset E$ est l'ensemble des limites de points de A.

Remarque 19.12. On pourrait aussi donner une définition en termes de voisinages et montrer qu'elle est équivalente, mais ce n'est pas notre objet ici. Tout cela sera traité en détail en seconde année.

On arrive alors au fondamental

Théorème 19.12. *Soit E et F des $\mathbb{K}$-ev, avec F complet. Soit $A \subset E$. et $f : A \to F$ une application uniformément continue. Alors il existe un unique prolongement continu de f à $\overline{A}$.*

Démonstration. On travaille en deux temps.

- Existence
  Soit $x \in \overline{A}$, alors on peut écrire $x = \lim x_n$ avec $x_n \in A$. Comme la suite (x_n) converge elle est de Cauchy, puis comme f est uniformément continue sur A, la suite $(f(x_n))$ est de Cauchy dans F (considérer $\varepsilon > 0$ et $\delta > 0$ associé dans l'uniforme continuité de f). Puisque F est complet, $(f(x_n))$ admet une limite qu'on note $\overline{f}(x)$.
  Montrons que cette définition de $\overline{f}$ est consistante. Si (x'_n) est une autre suite qui converge vers x, alors $\|x_n - x'_n\| \leqslant \|x_n - x\| + \|x - x'_n\| \to 0$. Par uniforme continuité de f on a $\|f(x_n) - f(x'_n)\| \to 0$. Notant $\overline{f}_2(x)$ la limite de $f(x'_n)$, on a

  $$|\tilde{f}(x) - \tilde{f}_2(x)| \leqslant \|\tilde{f}(x) - f(x_n)\| + \|f(x_n) - f(x'_n)\| + \|f(x'_n) - \tilde{f}_2(x)\| \to 0$$

  donc $\overline{f}(x) = \overline{f}_2(x)$. Par ailleurs, $\overline{f}$ prolonge bien f, puisque si on se donne $x \in A$, il suffit de considérer la suite constante égale à x, la limite de la suite image étant égale à $f(x)$.
  Montrons enfin que $\overline{f}$ ainsi définie est continue. En fait, nous allons même montrer qu'elle est uniformément continue. Soit $\varepsilon > 0$ et $\delta > 0$ associé dans l'uniforme continuité de f. Soit ensuite $x, y \in \overline{A}$ tels que $\|x - y\| \leqslant \delta/2$. On choisit $x_n \to x, y_n \to y$, puis on écrit

  $$\|x_n - y_n\| \leqslant \|x_n - x\| + \|x - y\| + \|y - y_n\| \to \|x - y\|$$

  donc à partir d'un certain rang on a $\|x_n - y_n\| \leqslant \delta$ puis $\|f(x_n) - f(y_n)\| \leqslant \varepsilon$. Or de même on a

  $$\|\tilde{f}(x) - \tilde{f}(y)\| \leqslant \|\tilde{f}(x) - f(x_n)\| + \varepsilon + \|f(y_n) - \tilde{f}(y)\|$$

  puis par passage à la limite on obtient $\|\overline{f}(x) - \overline{f}(y)\| \leqslant \varepsilon$.
- Unicité.
  Supposons que $\overline{f}$ et $\overline{g}$ conviennent, et soit $x \in \overline{A}$. Écrivons $x = \lim x_n$ après quoi on a

  $$\begin{aligned} |\tilde{f}(x) - \tilde{g}(x)| &\leqslant \|\tilde{f}(x) - f(x_n)\| + \|f(x_n) - g(x_n)\| + \|g(x_n) - \tilde{g}(x)\| \\ &= \|\overline{f}(x) - \overline{f}(x_n)\| + \|\tilde{g}(x_n) - \tilde{g}(x)\| \\ &\to 0 \end{aligned}$$

  par continuité de $\overline{f}$ et $\overline{g}$. Donc $|\tilde{f}(x) - \tilde{g}(x)| = 0$, puis comme cette égalité est valable pour tout x, on a $\tilde{f} = \tilde{g}$.

Ainsi la preuve est achevée. □

Voici maintenant une application spectaculaire de ce procédé. On commence par considérer l'ensemble $A \subset \mathcal{B}([a,b], \mathbb{K})$ des fonctions en escalier, et dessus on construit une application

$$I : \varphi \mapsto \int_{[a,b]} \varphi$$

dont on montre qu'elle est linéaire. L'inégalité triangulaire, quant à elle, montre que I est $(b-a)$-lipschitzienne, et donc uniformément continue. Or, $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ est complet. Ainsi, il existe un unique prolongement continu de I à l'ensemble $\overline{A}$ des fonctions réglées (qui contient en particulier les fonctions continues par morceaux), on l'appelle **intégrale de Riemann**.

La linéarité de I sur A passe naturellement à la limite sur $\overline{A}$ (c'est un simple argument de continuité de I), et on obtient de la même manière l'inégalité triangulaire.

Chapitre 20

Analyse asymptotique

1 Relations de comparaison

1.1 Suites réelles

Dans tout le paragraphe, on fixe u et v deux suites réelles indexées sur $\mathbb{N}$.

Définition 20.1 (Négligeabilité)**.** u est dite **négligeable devant** v, et on note $u_n \underset{n\to+\infty}{=} o(v_n)$ (voire $u_n = o(v_n)$) lorsque :

$$\exists N \in \mathbb{N},\ \exists (w_n)_{n\geqslant N} \in \mathbb{R}^{[\![N,+\infty[\![},\ (\forall n \geqslant N,\ u_n = v_n w_n) \quad \text{et} \quad w_n \xrightarrow[n\to+\infty]{} 0$$

Définition 20.2 (Domination)**.** u est dite **dominée par** v, et on note $u_n \underset{n\to+\infty}{=} O(v_n)$ (voire $u_n = O(v_n)$) lorsque :

$$\exists N \in \mathbb{N},\ \exists (w_n)_{n\geqslant N} \in \mathbb{R}^{[\![N,+\infty[\![},\ (\forall n \geqslant N,\ u_n = v_n w_n) \quad \text{et} \quad (w_n)_{n\geqslant N} \text{ est bornée}$$

Définition 20.3 (Équivalence)**.** u est dite **équivalente à** v, et on note $u_n \underset{n\to+\infty}{\sim} v_n$ (voire $u_n \sim v_n$) lorsque :

$$\exists N \in \mathbb{N},\ \exists (w_n)_{n\geqslant N} \in \mathbb{R}^{[\![N,+\infty[\![},\ (\forall n \geqslant N,\ u_n = v_n w_n) \quad \text{et} \quad w_n \xrightarrow[n\to+\infty]{} 1$$

Remarque 20.1. Attention, les relations $\cdot = o(\cdot)$ et $\cdot = O(\cdot)$ ne sont pas des relations d'équivalence. Elles n'ont aucune raison d'être symétriques notamment. Et ce n'est pas parce que deux suites sont un petit o d'une même suite qu'elles sont égales !

Exemple 20.1 ($u_n \sim 0$ ssi (u_n) nulle APCR)**.** On a toujours $0 = o(v_n)$ mais la réciproque est beaucoup plus rare. En effet, on peut montrer que $u_n = o(0)$ si, et seulement si, (u_n) est nulle APCR. Ce dernier fait est aussi vrai avec les grands O et l'équivalence.

Théorème 20.1. *On a :*

$$u_n \sim v_n \iff u_n = v_n + o(v_n)$$

Théorème 20.2 (Caractérisation par le quotient)**.** *Supposons que v est non nulle APCR Nßℕ. Alors :*

1. $u_n = o(v_n) \iff \dfrac{u_n}{v_n} \xrightarrow[n\to+\infty]{} 0$
2. $u_n = O(v_n) \iff \left(\dfrac{u_n}{v_n}\right)_{n\geqslant N}$ *est bornée*
3. $u_n \sim v_n \iff \dfrac{u_n}{v_n} \xrightarrow[n\to+\infty]{} 1$

Exemple 20.2 ($u_n = o(1)$)**.** On a $u_n = o(1) \iff u_n \xrightarrow[n\to+\infty]{} 0$. Plus généralement, on a

$$\forall \lambda \in \mathbb{R}^\star, u_n = o(\lambda) \iff u_n \xrightarrow[n\to+\infty]{} 0$$

Exemple 20.3 ($u_n = O(1)$)**.** On a $u_n = O(1) \iff (u_n)$ est bornée. Plus généralement, on a

$$\forall \lambda \in \mathbb{R}^\star, u_n = O(\lambda) \iff (u_n) \text{ est bornée}$$

Exemple 20.4. Soit $\alpha > \beta \geqslant 0$ des réels. On a $n^\beta = o(n^\alpha)$ et $\dfrac{1}{n^\alpha} = o\left(\dfrac{1}{n^\beta}\right)$

Théorème 20.3 (Obtention d'un équivalent par encadrement)**.** *Soit u, v et w des suites réelles et α une suite à valeurs dans $\mathbb{R}_+^\star$ telles que $u_n \sim \alpha_n$, $w_n \sim \alpha_n$ et $u \leqslant v \leqslant w$ APCR. Alors, on a :*

$$v_n \sim \alpha_n$$

Remarque 20.2. On peut écrire les choses de façon plus condensée. Par exemple, si on veut écrire "Si $u_n = o(w_n)$ et $v_n = o(w_n)$, alors $u_n + v_n = o(w_n)$, on pourra écrire $o(w_n) + o(w_n) = o(w_n)$. Attention, cette notation est **symbolique** et **doit être lue de gauche à droite**. En particulier, le symbole égal n'est alors plus symétrique. Cependant, cela permet de condenser les écritures.

Dans ce contexte, on donne le théorème qui suit :

Théorème 20.4 (Composition des symboles)**.** *On a les relations **symboliques** (mais vraies si correctement reformulées) :*

1. $o(u_n) = O(u_n)$
2. $u_n \sim v_n \implies u_n = O(v_n)$
3. $\alpha_n o(v_n) = o(\alpha_n v_n)$
4. $o(\alpha_n v_n) = \alpha_n o(v_n)$
5. $\alpha_n O(v_n) = O(\alpha_n v_n)$
6. $O(\alpha_n v_n) = \alpha_n O(v_n)$
7. $o(v_n) + o(v_n) = o(v_n)$
8. $(v_n) + O(v_n) = O(v_n)$
9. $O(v_n) + O(v_n) = O(v_n)$
10. $o(o(v_n)) = o(v_n)$
11. $o(O(v_n)) = o(v_n)$

12. $O(o(v_n)) = o(v_n)$
13. $O(O(v_n)) = O(v_n)$
14. $o(u_n)o(v_n) = o(u_n v_n)$
15. $O(u_n)O(v_n) = O(u_n v_n)$

Remarque 20.3. Attention ! $o(u_n) + o(v_n) \neq o(u_n + v_n)$ en général ! De même, $O(u_n) + O(v_n) \neq O(u_n + v_n)$ en général !

Théorème 20.5. *$\sim$ est une relation d'équivalence sur $\mathbb{R}^{\mathbb{N}}$.*

Remarque 20.4. Il ne sert à rien d'écrire un équivalent à plus d'un terme ! Par exemple, écrire $u_n \sim 1 + \dfrac{1}{n}$ revient à dire que $u_n \sim 1$ voire $u_n \sim 1 - \dfrac{1}{n}$ par transitivité de $\sim$.

Remarque 20.5. $\forall \lambda \in \mathbb{R}^\star,\ u_n \sim \lambda \iff u_n \xrightarrow[n\to+\infty]{} \lambda$

Théorème 20.6 (Équivalence et héritage)**.** *Supposons que $u_n \sim v_n$. Alors la limite s'hérite par équivalence :*

$$\forall l \in \overline{\mathbb{R}},\ u_n \xrightarrow[n\to+\infty]{} l \iff v_n \xrightarrow[n\to+\infty]{} l$$

Aussi, le signe s'hérite par équivalence :

1. $u_n \neq 0$ *APCR* $\iff v_n \neq 0$ *APCR*
2. $u_n \geqslant 0$ *APCR* $\iff v_n \geqslant 0$ *APCR*
3. $u_n \leqslant 0$ *APCR* $\iff v_n \leqslant 0$ *APCR*
4. $u_n > 0$ *APCR* $\iff v_n > 0$ *APCR*
5. $u_n < 0$ *APCR* $\iff v_n < 0$ *APCR*

Enfin, soit $w \in \mathbb{R}^{\mathbb{N}}$. Alors, les relations de comparaison s'héritent :

1. $w_n = o(v_n) \iff w_n = o(v_n)$
2. $w_n = O(v_n) \iff w_n = O(v_n)$
3. $u_n = o(w_n) \iff v_n = o(w_n)$
4. $u_n = O(w_n) \iff v_n = O(w_n)$

Théorème 20.7 (Produit, quotient et puissances d'équivalents)**.** *Soit $a, b, c, d \in \mathbb{R}^{\mathbb{N}}$. Soit $\alpha \in \mathbb{R}$. Alors :*

1. *Si $a_n \sim b_n$ et $c_n \sim d_n$, alors $a_n c_n \sim b_n d_n$.*
2. *Supposons que c et d soient non nulles APCR. Si $a_n \sim b_n$ et $c_n \sim d_n$, alors $\dfrac{a_n}{c_n} \sim \dfrac{b_n}{d_n}$.*
3. *Supposons que a et b sont à valeurs dans $\mathbb{R}_+^\star$ APCR. Si $a_n \sim b_n$, alors $a_n^\alpha \sim b_n^\alpha$.*

Remarque 20.6 (Billet gratuit pour la 5/2)**. Attention ! On ne somme pas les équivalents ! ! !**

Théorème 20.8 (HP)**.** *Si $u_n \sim v_n$ et $u_n \xrightarrow[n\to+\infty]{} +\infty$ (et donc de même pour v_n), alors :*

$$\ln(u_n) \sim \ln(v_n)$$

Démonstration. A refaire à chaque fois. On a $u_n = v_n + o(v_n)$. Donc en factorisant dans le ln et en utilisant la propriété fonctionnelle du ln, on obtient :

$$\ln(u_n) = \ln(v_n) + \ln(1 + o(1))$$

Or, $1 + o(1)$ tend vers 1, donc par continuité du ln en 1, $\ln(1 + o(1)) = o(1)$. Puisque v_n tend vers $+\infty$, $\ln(v_n)$ aussi. A fortiori, on a donc $\ln(1 + o(1)) = o(\ln(v_n))$ (par une caractérisation par le quotient par exemple). D'où :

$$\ln(u_n) = \ln(v_n) + o(\ln(v_n))$$

Autrement dit,

$$\ln(u_n) \sim \ln(v_n)$$

□

Théorème 20.9 (Formule de Stirling)**.** *On a :*

$$n! \sim \left(\frac{n}{e}\right)^n \sqrt{2\pi n}$$

Exemple 20.5. Avec le théorème HP qui précède, on obtient facilement que $\ln(n!) \sim n \ln(n)$.

1.2 Brève extension aux suites complexes

Définition 20.4 (Extension des définitions à $\mathbb{C}$)**.** Soit $u, v \in \mathbb{C}^{\mathbb{N}}$. On définit de même $u_n = o(v_n)$, $u_n = O(v_n)$ et $u_n \sim v_n$ à ceci près que la suite (w_n) est désormais à valeurs dans $\mathbb{C}$.

Théorème 20.10 (Extension des résultats à $\mathbb{C}$)**.** *Dans $\mathbb{C}$, les résultats suivants restent vrais :*

1. $u_n \sim v_n \iff u_n = v_n + o(v_n)$
2. *$\sim$ reste une relation d'équivalence*
3. *Caractérisation par le quotient*
4. *Composition des symboles*
5. *Produit et quotient d'équivalents (les puissances restent vraies mais nous ramènent par hypothèse dans $\mathbb{R}_+^\star$)*
6. *Les héritages, exceptés ceux qui font intervenir des symboles d'inégalités et perdent alors leur sens.*

1.3 Généralisation aux fonctions

Dans cette section, on fixe $X \subset \mathbb{R}$ tel que $X \neq \emptyset$ ainsi que $a \in \overline{X}$. Les fonctions seront considérées de X dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$.

Définition 20.5. Soit $f, g : X \to \mathbb{K}$.

1. On écrira $f(x) \underset{x \to a}{=} o\left(g(x)\right)$ lorsque

$$\exists V \in \mathcal{V}(a),\ \exists h : X \cap V \to \mathbb{K},\ (\forall x \in X \cap V,\ f(x) = g(x)h(x)) \quad \text{et} \quad h(x) \xrightarrow[x \to a]{} 0$$

2. On écrira $f(x) \underset{x\to a}{=} O(g(x))$ lorsque

$$\exists V \in \mathcal{V}(a),\ \exists h : X \cap V \to \mathbb{K},\ (\forall x \in X \cap V,\ f(x) = g(x)h(x)) \quad \text{et} \quad h \text{ est bornée}$$

3. On écrira $f(x) \underset{x\to a}{\sim} o(g(x))$ lorsque

$$\exists V \in \mathcal{V}(a),\ \exists h : X \cap V \to \mathbb{K},\ (\forall x \in X \cap V,\ f(x) = g(x)h(x)) \quad \text{et} \quad h(x) \xrightarrow[x\to a]{} 1$$

Proposition 20.1 (Caractère local des relations de comparaison)**.** *On garde f et g comme précédemment. Soit $\tilde{V} \in \mathcal{V}(a)$. Posons $\tilde{f} = f_{|X\cap\tilde{V}}$ et $\tilde{g} = g_{|X\cap\tilde{V}}$. Alors :*

1. $f(x) \underset{x\to a}{=} o(g(x)) \iff \tilde{f}(x) \underset{x\to a}{=} o(\tilde{g}(x))$
2. $f(x) \underset{x\to a}{=} O(g(x)) \iff \tilde{f}(x) \underset{x\to a}{=} O(\tilde{g}(x))$
3. $f(x) \underset{x\to a}{\sim} g(x) \iff \tilde{f}(x) \underset{x\to a}{\sim} \tilde{g}(x)$

Remarque 20.7. $f(x) \underset{x\to a}{\sim} 0$ si, et seulement si, f est nulle au voisinage de a

Remarque 20.8. Soit $\lambda \in \mathbb{K}^\star$. On a $f(x) \underset{x\to a}{\sim} \lambda$ si, et seulement si, $f(x) \xrightarrow[x\to a]{} \lambda$

Proposition 20.2. *On a*

$$f(x) \underset{x\to a}{\sim} g(x) \iff f(x) \underset{x\to a}{=} g(x) + o(g(x))$$

Remarque 20.9. $\sim$ est une relation d'équivalence.

Remarque 20.10 (Caractérisation par le quotient)**.** La caractérisation par le quotient reste valable tant que g ne s'annule pas au voisinage de a.

Remarque 20.11. Soit $\alpha \geqslant \beta > 0$. On a $x^\beta \underset{x\to+\infty}{=} o(x^\alpha)$

Exemple 20.6. Soit $f : x \to a_n x^n + \ldots + a_0$ une fonction polynomiale (avec $n \in \mathbb{N}$ et $a_n \neq 0$). On a

$$f(x) \underset{x\to+\infty}{\sim} a_n x^n$$

et

$$f(x) \underset{x\to-\infty}{\sim} a_n x^n$$

Théorème 20.11. *Les résultats généraux suivants vus sur les suites restent valables :*

1. *Composition des symboles*
2. *Équivalents qui passent au produit, au quotient et à une puissance* ***indépendante de*** *x*
3. *Héritages par équivalence*
4. *Obtention d'un équivalent par encadrement*

Exemple 20.7. Soit $(p, q) \in \mathbb{N}^2$ tel que $p < q$. En revenant à la définition, on montre que

$$x^q \underset{x\to0}{=} o(x^p)$$

Exemple 20.8. Considérons la fonction rationnelle $f : x \to \dfrac{a_n x^n + \ldots + a_0}{b_p x^p + \ldots + b_0}$ (avec $n, p \in \mathbb{N}$ et $a_n \neq 0$ et $b_p \neq 0$). On a

$$f(x) \underset{x\to+\infty}{\sim} \frac{a_n}{b_p} x^{n-p}$$

et

$$f(x) \underset{x\to-\infty}{\sim} \frac{a_n}{b_p} x^{n-p}$$

1.4 Petit retour sur les croissances comparées

Théorème 20.12 (Croissances comparées pour les fonctions)**.** *Soit $\alpha, \beta, \gamma > 0$ Au voisinage de $+\infty$, on a :*

1. $\ln^\alpha(x) = o(x^\beta)$
2. $x^\beta = o(\exp(\gamma x))$

Et de plus, ces trois expressions tendent vers $+\infty$ quand x tend vers $+\infty$.

Corollaire 20.1. *On en déduit :*

$$|\ln^\alpha(x)| x^\beta \xrightarrow[x\to 0^+]{} 0$$

et

$$|x^\beta| \exp(\gamma x) \xrightarrow[x\to-\infty]{} 0$$

Théorème 20.13 (Croissances comparées pour les suites)**.** *Soit $\alpha, \beta, \gamma > 0$ Au voisinage de $+\infty$, on a :*

1. $\ln^\alpha(n) = o(n^\beta)$
2. $n^\beta = o(\exp(\gamma n))$
3. $\exp(\gamma n) = o(n!)$
4. $n! = o(n^n)$

Et de plus, ces cinq expressions tendent vers $+\infty$ quand n tend vers $+\infty$.

2 Développements limités

Dans tout le paragraphe, on fixe : I un intervalle non trivial de $\mathbb{R}$, $a \in I$, $f : I \to \mathbb{K}$ avec $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$ ainsi que $n \in \mathbb{N}$.

2.1 Premières définitions

Définition 20.6 (Développement limité à l'ordre n en a)**.** On dit que f admet un **développement limité à l'ordre n en a/au point a** (noté $DL_n(a)$) lorsque :

$$\exists (c_0, \ldots, c_n) \in \mathbb{K}^{n+1},\ f(x) \underset{x\to a}{=} \sum_{k=0}^{n} c_k (x-a)^k + o\left((x-a)^n\right)$$

Remarque 20.12. Sans perte de généralité, on pourra toujours supposer que la fonction ε que contient le o de la définition est définie sur I tout entier, et non sur un voisinage de a, puisque de toute façon la définition du o est **locale**.

Théorème 20.14 (Cas $n = 0$, équivalence avec la continuité)**.** *f admet un $DL_0(a)$ si, et seulement si, f est continue en a. Dans ce cas, ce DL est nécessairement donnée par :*

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

ie on a $c_0 = f(a)$ dans la définition.

Théorème 20.15 (Cas $n = 1$, équivalence avec la dérivabilité)**.** *f admet un $DL_1(a)$ si, et seulement si, f est dérivable en a. Dans ce cas, ce DL est nécessairement donnée par :*

$$f(x) \underset{x\to a}{=} f(a) + f'(a)(x-a) + o(x-a)$$

ie on a $c_0 = f(a)$ et $c_1 = f'(a)$ dans la définition.

Remarque 20.13. Attention ! Pour $n \geqslant 2$, on ne peut plus rien dire ! Par exemple, la fonction de $\mathbb{R}$ dans $\mathbb{R}$ qui à x associe $x^3 \sin\left(\dfrac{1}{x}\right)$ prolongée par 0 en 0 admet un $DL_2(0)$ (poser $c_0 = c_1 = c_2 = 0$) mais n'est pourtant pas deux fois dérivable en 0.

Théorème 20.16 (Lemme de troncature)**.** *Supposons que f admette un $DL_n(a)$ qu'on écrit sous la forme :*

$$f(x) \underset{x\to a}{=} \sum_{k=0}^{n} c_k(x-a)^k + o\left((x-a)^n\right)$$

Alors, pour tout $m \in [\![0, n]\!]$, f admet un $DL_m(a)$ donné par :

$$f(x) \underset{x\to a}{=} \sum_{k=0}^{m} c_k(x-a)^k + o\left((x-a)^m\right)$$

Théorème 20.17 (Unicité)**.** *Supposons que f admette un $DL_n(a)$ qu'on écrit sous la forme :*

$$f(x) \underset{x\to a}{=} \sum_{k=0}^{n} c_k(x-a)^k + o\left((x-a)^n\right)$$

Alors, les scalaires $c_0, \ldots, c_n$ sont uniques.

Définition 20.7 (Partie régulière)**.** La fonction polynomiale

$$\begin{array}{ccc} I & \longrightarrow & \mathbb{K} \\ x & \longmapsto & \displaystyle\sum_{k=0}^{n} c_k(x-a)^k \end{array}$$

est donc définie de façon unique. On l'appelle **la partie régulière du** $DL_n(a)$.

Corollaire 20.2 (Fonctions paires)**.** *Supposons que f est paire et que f admet un $DL_n(0)$ qu'on écrit sous la forme :*

$$f(x) \underset{x\to 0}{=} \sum_{k=0}^{n} c_k x^k + o\left(x^n\right)$$

*Alors, tous les c_k pour k **impair** sont nuls.*

Corollaire 20.3 (Fonctions impaires)**.** *Supposons que f est impaire et que f admet un $DL_n(0)$ qu'on écrit sous la forme :*

$$f(x) \underset{x\to 0}{=} \sum_{k=0}^{n} c_k x^k + o\left(x^n\right)$$

*Alors, tous les c_k pour k **pair** sont nuls.*

2.2 Calculs pratiques

Théorème 20.18 (Combinaison linéaire)**.** *Soit f et g des fonctions de I dans $\mathbb{K}$ admettant des $DL_n(a)$ qu'on écrit sous la forme :*

$$f(x) \underset{x\to a}{=} \sum_{k=0}^{n} c_k (x-a)^k + o\left((x-a)^n\right)$$

et

$$g(x) \underset{x\to a}{=} \sum_{k=0}^{n} d_k (x-a)^k + o\left((x-a)^n\right)$$

Soit $\lambda, \mu \in \mathbb{K}$. Alors, $\lambda f + \mu g$ admet un $DL_n(a)$ donné par :

$$(\lambda f + \mu g)(x) \underset{x\to a}{=} \sum_{k=0}^{n} (\lambda c_k + \mu d_k)(x-a)^k + o\left((x-a)^n\right)$$

Théorème 20.19 (Produit)**.** *Soit f et g des fonctions de I dans $\mathbb{K}$ admettant des $DL_n(a)$ qu'on écrit sous la forme :*

$$f(x) \underset{x\to a}{=} \sum_{k=0}^{n} c_k (x-a)^k + o\left((x-a)^n\right)$$

et

$$g(x) \underset{x\to a}{=} \sum_{k=0}^{n} d_k (x-a)^k + o\left((x-a)^n\right)$$

Alors fg admet un $DL_n(a)$ donné par : Soit f et g des fonctions de I dans $\mathbb{K}$ admettant des $DL_n(a)$ qu'on écrit sous la forme :

$$(fg)(x) \underset{x\to a}{=} \sum_{k=0}^{n} \left(\sum_{i+j=k} c_i d_j \right) (x-a)^k + o\left((x-a)^n\right)$$

Pour la composition, aucun résultat théorique n'est exigible, on se contente de faire de la substitution.

Proposition 20.3 (Substitution dans un DL)**.** *Supposons que f possède un $DL(0)$ qu'on écrit sous la forme :*

$$f(u) \underset{u\to 0}{=} \sum_{k=0}^{n} c_k u^k + o\left(u^n\right)$$

*et supposons qu'on possède une expression $u(x)$ telle que $u(x) \xrightarrow[x\to a]{} 0$. Alors on a par **substitution** :*

$$f(u(x)) \underset{x\to a}{=} \sum_{k=0}^{n} c_k (u(x))^k + o\left((u(x))^n\right)$$

*Sur une copie ou en khôlle, les mentions "$u(x) \xrightarrow[x\to a]{} 0$" et "substitution" devront **toujours** apparaître.*

Proposition 20.4 (Substitution dans les relations de comparaison)**.** *Supposons que $f(u) \underset{u\to 0}{\sim} g(u)$ et supposons qu'on possède une expression $u(x)$ telle que $u(x) \xrightarrow[x\to a]{} 0$. Alors, par substitution, on a :*

$$f(u(x)) \underset{x\to a}{\sim} g(u(x))$$

On a les résultats symétriques avec les petits o et les grands O.

Proposition 20.5 (Substitution de suites dans des DL ou relations de comparaison)**.** *Avec les mêmes hypothèses, mais cette fois-ci des suites qui tendent vers 0, on a les mêmes résultats de substitution.*

Proposition 20.6 (Quotient)**.** *Soit $u : I \to \mathbb{K}$ telle que $u(x) \xrightarrow[x\to a]{} 0$ et u admette un $DL_n(a)$. Alors, $\dfrac{1}{1-u}$ admet un $DL_n(a)$. En fait, on a :*

$$\frac{1}{1-u(x)} \underset{x\to a}{=} 1 + \ldots + (u(x))^n + o\left((x-a)^n\right)$$

et chacun des u^k admet un $DL_n(a)$ par produit. On conclut par combinaison linéaire.

Remarque 20.14. Attention ! Le résultat précédent n'est d'aucune utilité pratique, car si on anticipe mal les ordres, on se retrouve avec des ordres en trop. Moralité : il faut **anticiper les ordres** pour se faciliter le travail.

Remarque 20.15 (Recentrage en 0)**.** Soit $(c_0, \ldots, c_n) \in \mathbb{K}^{n+1}$. L'assertion :

$$f(x) \underset{x\to a}{=} \sum_{k=0}^{n} c_k (x-a)^k + o\left((x-a)^n\right)$$

est équivalente à l'assertion :

$$f(a+h) \underset{h\to 0}{=} \sum_{k=0}^{n} c_k h^k + o\left(h^n\right)$$

On peut donc toujours se recentrer en 0 pour se ramener à des DL usuels.

Théorème 20.20 (Forme normalisée, "règle du $c_m h^m$ (NS pour cette dernière))**.** *On suppose que*

$$f(a+h) \underset{h\to 0}{=} \sum_{k=0}^{n} c_k h^k + o\left(h^n\right)$$

Supposons que $\exists k \in [\![0,n]\!],\ c_k \neq 0$*. Considérons alors* m *le plus petit entier tel que* c_m *soit non nul. Alors on a :*

$$f(a+h) \underset{h\to 0}{=} h^m \left(c_m + c_{m+1}h + \ldots + c_n h^{n-m} + o\left(h^{n-m}\right)\right)$$

C'est ce qu'on appelle la ***forme normalisée du DL****. De plus, par troncature, on a :*

$$f(a+h) \underset{h\to 0}{=} c_m h^m + o(h^m)$$

donc on en déduit que

$$f(a+h) \underset{h\to 0}{\sim} c_m h^m$$

Exemple 20.9. Soit x au voisinage de 0. On a :

$$\forall \alpha \in \mathbb{R},\ (1+x)^\alpha - 1 \sim \alpha x$$

$$\exp(x) - 1 \sim x$$

$$\ln(1+x) \sim x$$

$$\cos(x) - 1 \sim \frac{x^2}{2}$$

$$\sin(x) \sim x$$

$$\tan(x) \sim x$$

$$\arctan(x) \sim x$$

2.3 Résultats théoriques

Théorème 20.21. *Supposons que* f *admette un* $DL_n(a)$ *qu'on écrit sous la forme :*

$$f(x) \underset{x\to a}{=} \sum_{k=0}^{n} c_k (x-a)^k + o\left((x-a)^n\right)$$

Supposons que f *admette une primitive* F*. Alors,* F *admet un* $DL_{n+1}(a)$ *donné par :*

$$F(x) \underset{x\to a}{=} F(a) + \sum_{k=0}^{n} \frac{c_k}{k+1}(x-a)^{k+1} + o\left((x-a)^{n+1}\right)$$

Corollaire 20.4 (Formule de Taylor-Young)**.** *Soit* $f \in \mathcal{C}^n(I,\mathbb{K})$*. Alors, quel que soit* $a \in I$*,* f *admet un* $DL_n(a)$ *donné par :*

$$f(x) \underset{x\to a}{=} \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + o\left((x-a)^n\right)$$

soit encore :

$$f(a+h) \underset{h\to 0}{=} \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}h^k + o\left(h^n\right)$$

La partie régulière de la première forme de ce DL s'appelle le ***polynôme de Taylor de*** f ***à l'ordre*** n ***en*** a.

Remarque 20.16. On rappelle qu'admettre un DL_n ne rend pas $\mathcal{C}^n$!!!

Remarque 20.17. Une fonction de classe $\mathcal{C}^\infty$ admet donc des DL à tous les ordres.

Proposition 20.7 (Dérivée d'un DL_n pour une fonction $\mathcal{C}^n$ avec $n \in \mathbb{N}^\star$)**.** *En général, on ne dérive pas un DL. Cependant, supposons que* $n \in \mathbb{N}^\star$ *et* f *soit de classe* $\mathcal{C}^n$*. Alors, grâce à la formule de Taylor-Young, si on part d'un DL de* f *sous la forme*

$$f(x) \underset{x\to a}{=} \sum_{k=0}^{n} c_k(x-a)^k + o\left((x-a)^n\right)$$

avec

$$\forall k \in [\![0,n]\!],\ c_k = \frac{f^{(k)}(a)}{k!}$$

, on obtient alors

$$f'(x) \underset{x\to a}{=} \sum_{k=1}^{n} kc_k(x-a)^{k-1} + o\left((x-a)^{n-1}\right)$$

sachant que pour mémoire, on avait

$$f(x) \underset{x\to a}{=} f(a) + \sum_{k=1}^{n} c_k(x-a)^k + o\left((x-a)^n\right)$$

Virtuellement, tout se passe comme si on avait dérivé le $DL_n(a)$ *de* f *pour obtenir de le* $DL_{n-1}(a)$ *de* f'*.* ***Sur une copie, on pourra utiliser ce résultat à la condition d'écrire :*** $n \in \mathbb{N}^\star$***,*** f ***est de classe*** $\mathcal{C}^n$ ***et "formule de Taylor-Young".***

Théorème 20.22 (DL usuels au voisinage de 0, à savoir par cœur)**.** *On a les DL usuels suivants au voisinage de* 0*,* ***à connaître par cœur :***

$$\frac{1}{1-x} \underset{x\to 0}{=} 1 + x + ... + x^n + o(x^n)$$

$$\frac{1}{1+x} \underset{x\to 0}{=} 1 - x + ... + (-1)^n x^n + o(x^n)$$

$$(1+x)^\alpha \underset{x\to 0}{=} 1 + \alpha x + \frac{\alpha(\alpha-1)}{2}x^2 + ... + \frac{\alpha(\alpha-1)...(\alpha-n+1)}{n!}x^n + o(x^n)$$

$$e^x \underset{x\to 0}{=} 1 + x + \frac{x^2}{2!} + ... + \frac{x^n}{n!} + o(x^n)$$

$$\ln(1+x) \underset{x\to 0}{=} x - \frac{x^2}{2} + \frac{x^3}{3} + ... + (-1)^{n-1}\frac{x^n}{n} + o(x^n)$$

$$\cos(x) \underset{x\to 0}{=} 1 - \frac{x^2}{2!} + \frac{x^4}{4!} + \ldots + (-1)^n \frac{x^{2n}}{(2n)!} + o(x^{2n+1})$$

$$\sin(x) \underset{x\to 0}{=} x - \frac{x^3}{3!} + \frac{x^5}{5!} + \ldots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + o(x^{2n+2})$$

$$\cosh(x) \underset{x\to 0}{=} 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \ldots + \frac{x^{2n}}{(2n)!} + o(x^{2n+1})$$

$$\sinh(x) \underset{x\to 0}{=} x + \frac{x^3}{3!} + \frac{x^5}{5!} + \ldots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + o(x^{2n+2})$$

$$\arctan(x) \underset{x\to 0}{=} x - \frac{x^3}{3} + \frac{x^5}{5} + \ldots + (-1)^n \frac{x^{2n+1}}{2n+1} + o(x^{2n+2})$$

$$\tan(x) = x + \frac{1}{3}x^3 + \frac{2}{15}x^5 + \frac{17}{315}x^7 + o(x^8)$$

$$\tanh(x) = x - \frac{1}{3}x^3 + \frac{2}{15}x^5 - \frac{17}{315}x^7 + o(x^8)$$

Définition 20.8 (DL au sens fort)**.** Soit $n \in \mathbb{N}^\star$ et x au voisinage de 0. On a :

$$\ln(1+x) \underset{x\to 0}{=} x - \frac{x^2}{2} + \frac{x^3}{3} + ... + (-1)^{n-1}\frac{x^n}{n} + (-1)^n \frac{x^{n+1}}{n+1} + o(x^{n+1})$$

ce qu'on peut aussi réécrire :

$$\ln(1+x) \underset{x\to 0}{=} x - \frac{x^2}{2} + \frac{x^3}{3} + ... + (-1)^{n-1}\frac{x^n}{n} + O(x^{n+1})$$

C'est ce qu'on appelle un **DL au sens fort**. En effet, $O(x^{n+1}) \Longrightarrow o(x^n)$ mais pas l'inverse.

Définition 20.9 (Développement asymptotique)**.** Il s'agit de la même idée que le DL, mais en bord de domaine avec $a \in \overline{I} \setminus I$. On n'a plus nécessairement des puissances de x, mais cela peut être des fonctions de forces croissantes avec par exemple les différents éléments des croissances comparées.

Méthode 20.1 (DA de suites implicites)**.** Voici la méthode à appliquer pour étudier le DA d'une suite définie de façon implicite (en vogue aux oraux des Mines) :

1. Prouver la bonne définition de la suite. Traditionnellement, un tableau de variation suffit pour prouver l'existence et l'unicité.
2. On détermine la limite de la suite (u_n) définie de façon implicite On peut par exemple l'interpréter comme la bijection réciproque de la fonction qui permet de la définir de façon implicite. Souvent, cette limite vaut $+\infty$.
3. On cherche un moteur, c'est-à-dire une équation qui mette en jeu u_n et un $o(u_n)$. Cette équation nous permettra de gagner peu à peu des ordres dans le DA. Souvent, ce moteur s'obtient soit avec la définition implicite de la suite, soit en passant cette définition au ln ou à d'autres fonctions du même genre.
4. Il faut initialiser le moteur, donc chercher un premier terme de Dl ou de DA. Souvent, un équivalent obtenu directement ou indirectement à partir de la limite fait l'affaire.

Exemple 20.10. Pour mettre en œuvre cette méthode, se référer aux exercices 22, 23 et 24 du TD16.

2.4 Application aux études de courbes

Soit $f : I \to \mathbb{R}$ et $a \in I$. Supposons que f admette un $DL_1(a) : f(x) \underset{x\to a}{=} c_0+c_1(x-a)+o(x-a)$. Alors f est dérivable en a et on a $f(a) = c_0$ et $f'(a) = c_1$. On rappelle que l'équation de la tangente à la courbe représentative de f en a est alors donnée par $y - f(a) = f'(a)(x - a)$ soit encore

$$y = c_0 + c_1(x - a)$$

Ainsi, on "lit" l'équation de la tangente à la courbe de f en a dans le $DL_1(a)$. Pousser le DL plus loin nous donne un équivalent de $f - y_{tangente}$, ce qui nous donne localement les positions relatives de la courbe et de la tangente.

Théorème 20.23 (CN et CS sur la nature des extrema locaux)**.** *Supposons que f soit $\mathcal{C}^2$ au voisinage de a et que $f'(a) = 0$. On a une condition nécessaire sur la nature de l'extremum local en a :*

1. *Si f admet un minimum local en a, alors $f''(a) \geqslant 0$.*
2. *Si f admet un maximum local en a, alors $f''(a) \leqslant 0$.*

On a ensuite une condition suffisante sur la nature d'un potentiel extremum local en a :

1. *Si $f''(a) > 0$, alors f admet un minimum local en a.*
2. *Si $f''(a) < 0$, alors f admet un maximum local en a.*

Remarque 20.18. Si $f''(a) = 0$, on ne peut rien dire. En guise de contre-exemple, considérer $x \mapsto x^3$, $x \mapsto x^4$ et $x \mapsto -x^4$ en 0.

Définition 20.10 (Asymptote au voisinage de $+\infty$ et $-\infty$)**.** Supposons que f admette un développement asymptotique au voisinage de $+\infty$ de la forme :

$$f(x) \underset{x\to+\infty}{=} ax + b + o(1)$$

Alors on dit que la droite $y = ax + b$ est **asymptote à la courbe au voisinage de** $+\infty$. On définit de même les asymptotes au voisinage de $-\infty$.

Chapitre 21

Introduction aux équations différentielles

Dans tout le chapitre, I désigne un intervalle non trivial de $\mathbb{R}$ et toutes les fonctions vont de I dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

1 Généralités

Définition 21.1 (Équation différentielle linéaire d'ordre n)**.** Une **équation différentielle linéaire d'ordre** $n \in \mathbb{N}^\star$ est une équation dont l'inconnue est une fonction $y \in \mathcal{D}^n(I, \mathbb{K})$. Elle s'écrit sous la forme

$$\sum_{k=a}^{n} a_k(t) y^{(k)} = b(t)$$

où $a_1, ..., a_n, b$ sont des fonctions continues de I dans $\mathbb{K}$ et où a_n n'est pas la fonction nulle. Plus précisément, cela signifie que y vérifie

$$\forall t \in I,\ \sum_{k=a}^{n} a_k(t) y^{(k)}(t) = b(t)$$

Définition 21.2 (EDL à coefficients constants)**.** L'EDL est dite **"à coefficients constants"** lorsque les a_k sont des fonctions constantes.

Définition 21.3 (Forme résolue)**.** On dit que l'équation

$$\sum_{k=a}^{n} a_k(t) y^{(k)} = b(t)$$

est sous **forme résolue** lorsque $a_n = 1$. En pratique, on peut s'y ramener dès lors que a_n ne s'annule pas : il suffit alors de diviser par a_n.

Définition 21.4 (Équation homogène)**.** L'équation **homogène** associée à l'équation différentielle linéaire d'ordre n

$$\sum_{k=a}^{n} a_k(t)y^{(k)} = b(t)$$

est l'équation

$$\sum_{k=a}^{n} a_k(t)y^{(k)} = 0$$

Proposition 21.1 (Structure d'espace affin des solutions, si elles existent)**.** *La solution générale d'une équation différentielle linéaire, si elle existe, est la somme d'une solution particulière de l'équation avec second membre et de la solution générale de l'équation homogène associée.*

Proposition 21.2. *On cherche une solution particulière de*

$$\sum_{k=a}^{n} a_k(t)y^{(k)} = \lambda b_1(t) + \mu b_2(t)$$

il suffit alors de trouver une solution particulière y_1 *de* $\sum_{k=a}^{n} a_k(t)y^{(k)} = b_1(t)$ *et une solution particulière* y_2 *de* $\sum_{k=a}^{n} a_k(t)y^{(k)} = b_2(t)$*. Alors, puisque l'équation est linéaire, la fonction* $y_p = \lambda y_1 + \mu y_2$ *convient pour le problème initial.*

Remarque 21.1. Pour le cos et le sin, il peut être plus rapide de les interpréter comme des parties réelle et imaginaire d'une exponentielle complexe. On résout alors l'équation de $\mathbb{C}$ puis on passe à la partie réelle ou à la partie imaginaire.

2 Équations différentielles linéaires d'ordre 1

Théorème 21.1 (Résolution de $y'+a(t)y = 0$)**.** *Soit A une primitive quelconque de a. Les solutions sont données par*

$$\forall t \in I,\ y(t) = \lambda \exp(-A(t))$$

avec $\lambda \in \mathbb{K}$ une constante. Les solutions forment un $\mathbb{K}$-espace vectoriel de dimension 1.

Théorème 21.2 (Résolution de $y' + a(t)y = b(t)$)**.** *Déjà, cette équation admet au moins une solution. On commence par résoudre l'équation homogène associée : $y' + a(t)y = 0$. Ensuite, on fait varier la constante en écrivant y sous la forme $y(t) = \lambda(t)\exp(-A(t))$. On détermine alors $\lambda(t)$ en réinjectant dans l'équation générale car on obtient $\lambda'(t)\exp(-A(t)) = b(t)$ et il suffit alors d'isoler $\lambda'(t)$ et de primitiver. On remonte alors à y et on obtient la solution générale.*

Corollaire 21.1 (Existence et unicité d'une solution à un problème de Cauchy)**.** *Soit $t_0 \in I$ et $C_0 \in \mathbb{K}$. Alors il existe une unique fonction de I dan $\mathbb{K}$ vérifiant $y' + a(t)y = b(t)$ **et** la condition de Cauchy $y(t_0) = C_0$.*

Définition 21.5 (Cas général $a(t)y' + b(t)y = c(t)$, recollement)**.** Dans le cas général, a peut s'annuler. Toutefois, par continuité, on peut trouver des INT maximaux sur lesquels a ne s'annule pas. Sur ces intervalles, on divise par a pour se ramener sous forme résolue et appliquer notre méthode sur ces sous-intervalles. Ensuite, lorsqu'on a résolu sur tous les sous-intervalles, on essaye de recoller nos solutions aux points d'annulation de a en utilisant les hypothèses de continuité ou de dérivabilité de y en ces points-là : c'est ce qu'on appelle le **recollement**.

3 Équations différentielles linéaires d'ordre 2 à coefficients constants

Le seul cas au programme pour les EDL d'ordre 2 est celui avec les coefficients constants. En particulier, le premier coefficient est donc non nul, et on peut se ramener à la forme générale :

$$y'' + ay' + by = f(t)$$

avec f continue. On remarquera que toute solution est nécessairement de classe $\mathcal{C}^2$.

Théorème 21.3 (Équation homogène, cas complexe)**.** *Soit (E_0) : $y'' + ay' + b = 0$ avec $(a, b) \in \mathbb{C}^2$. Voici comment trouver S_0. On considère le polynôme caractéristique $P = X^2 + aX + b$ puis on distingue les cas suivants :*

1. *Supposons que P possède deux racines complexes **distinctes** r_1 et r_2. Alors*

$$S_0 = \left\{t \mapsto \lambda \exp(r_1 t) + \mu \exp(r_2 t) \mid (\lambda, \mu) \in \mathbb{C}^2\right\}$$

2. *Supposons que P possède une racine **double** r. Alors*

$$S_0 = \left\{t \mapsto (\lambda + \mu t) \exp(rt) \mid (\lambda, \mu) \in \mathbb{C}^2\right\}$$

Théorème 21.4 (Équation homogène, cas réel)**.** *Soit (E_0) : $y'' + ay' + b = 0$ avec $(a, b) \in \mathbb{R}^2$. Voici comment trouver S_0. On considère le polynôme caractéristique $P = X^2 + aX + b$ puis on distingue les cas suivants :*

1. *Supposons que P possède deux racines réelles **distinctes** r_1 et r_2. Alors*

$$S_0 = \left\{t \mapsto \lambda \exp(r_1 t) + \mu \exp(r_2 t) \mid (\lambda, \mu) \in \mathbb{R}^2\right\}$$

2. *Supposons que P possède une racine réelle **double** r. Alors*

$$S_0 = \left\{t \mapsto (\lambda + \mu t) \exp(rt) \mid (\lambda, \mu) \in \mathbb{R}^2\right\}$$

3. *Supposons que P possède 2 racines complexes non réelles conjuguées $s \pm i\omega$. Alors*

$$S_0 = \left\{t \mapsto \exp(st)(\lambda \cos(\omega t) + \mu \sin(\omega t) \mid (\lambda, \mu) \in \mathbb{R}^2)\right\}$$

Corollaire 21.2. *Que ce soit dans $\mathbb{R}$ ou dans $\mathbb{C}$, S_0 est un $\mathbb{K}$-espace vectoriel de dimension 2.*

Théorème 21.5 (Existence et unicité d'une solution à un problème de Cauchy)**.** *Soit $(C_0, C_1) \in \mathbb{K}^2$ et $t_0 \in I$. Alors il existe une unique solution de $y'' + ay' + by = f(t)$ qui vérifie $y(t_0) = C_0$ et $y'(t_0) = C_1$*

Concrètement, les seuls cas au programme pour f sont les fonctions polynomiales et les exponentielles, ainsi que tout ce qui s'en déduit par superposition ou par passage aux parties réelle et imaginaire.

Théorème 21.6 (Cas polynomial)**.** *On garde l'équation $y'' + ay' + b = f(t)$ et on suppose que $f(t)$ est de la forme $Q(t)$ où $Q \in \mathbb{K}[X]$ est de degré $n_i n\mathbb{N}$. On cherche une solution polynomiale. On distingue alors différent cas :*

1. *Supposons que $b \neq 0$. Alors on cherche la solution sous la forme d'un polynôme de $\mathbb{K}_n[X]$ et on réinjecte dans l'équation pour obtenir un système linéaire.*
2. *Supposons que $b = 0$ et $a \neq 0$ On cherche la solution sous la forme d'un polynôme de $\mathbb{K}_{n+1}[X]$ et on réinjecte dans l'équation pour obtenir un système linéaire.*
3. *Supposons que $b = 0$ et $a = 0$ On cherche la solution sous la forme d'un polynôme de $\mathbb{K}_{n+2}[X]$ et on réinjecte dans l'équation pour obtenir un système linéaire.*

Démonstration. Tout cela est lié aux fonctions

$$\begin{array}{rccc} \varphi : & \mathbb{K}_p[X] & \longrightarrow & \mathbb{K}_n[X] \\ & R & \longmapsto & R'' + aR' + bR \end{array}$$

pour $p \in \{n, n+1, n+2\}$ dans nos différents cas (dans le même ordre). A chaque fois, on détermine le noyau par des considérations de degré. Puis par le théorème du rang ces application sont surjectives. On a alors l'existence d'une solution, puis on la cherche en réinjectant. Cela fonctionne de même avec des applications légèrement plus complexes pour des seconds membres de la forme $Q(t)\exp(rt)$. □

Théorème 21.7. *On garde l'équation $y'' + ay' + b = f(t)$ et on suppose que $f(t)$ est de la forme $\exp(rt)$ pour $r \in \mathbb{K}$. On note toujours $P = X^2 + aX + b$ le polynôme caractéristique. Voici comment trouver une solution particulière. On distingue différents cas selon la multiplicité de r en tant que racine de P :*

1. *Supposons que r n'est **pas racine** de P. Alors*

$$t \mapsto \frac{1}{P(r)}\exp(rt)$$

   est une solution particulière.

2. *Supposons que r est racine **simple** de P. Alors*

$$t \mapsto \frac{t}{P'(r)}\exp(rt)$$

   est une solution particulière.

3. *Supposons que r est racine **double** de P. Alors*

$$t \mapsto \frac{t^2}{P''(r)}\exp(rt)$$

   est une solution particulière.

Pour le retenir, on a de manière générique, en posant $m = m_P(r)$ la multiplicité de r en tant que racine de P, la solution particulière

$$t \mapsto \frac{t^m}{P^{(m)}(r)}\exp(rt)$$

4 Compléments, HP

4.1 Lien avec la science physique

Il est très important pour l'étudiant de faire le lien entre ce chapitre et certains chapitres du programme de physique. En mécanique par exemple, l'équation du mouvement d'un mobile soumis à des forces aboutit à la résolution d'une équation différentielle. Le principe fondamental de la dynamique donne l'accélération, ce qui aboutit à une équation différentielle d'ordre 2 (avec pourquoi pas un terme à l'ordre 1 s'il y a des forces de frottement proportionnelles à la vitesse).

Une autre application, et des plus notables, se rencontre en électricité dans l'étude des circuits RLC, RL ou RC. Prenons le circuit RLC par exemple, qui consiste à placer en série

- un conducteur ohmique de résistance R ;
- une bobine d'inductance L ;
- un condensateur de capacité C.

Comme inconnue, nous choisissons u, la tension aux bornes du condensateur. La tension aux bornes du conducteur ohmique vaut alors $Ri = RCu'$, et celle aux bornes de la bobine vaut $Li' = LCu''$. On a ensuite

$$u + RCu' + LCu'' = 0 \iff u'' + \frac{R}{L}u' + \frac{1}{LC}u = 0$$

Le polynôme caractéristique est $X^2 + \frac{R}{L}X + \frac{1}{LC}$, de discriminant $\frac{R^2}{L^2} - \frac{4}{LC}$

- Un cas intéressant est celui où $R = 0$. On a alors deux racines complexes conjuguées $\pm\frac{i}{\sqrt{LC}}$.
  On posera $\omega = \frac{1}{\sqrt{LC}}$ si bien que les solutions de l'équation sont de la forme
  $$y(t) = \lambda\cos(\omega t) + \mu\sin(\omega t)$$
  On a ce qu'on appelle un **oscillateur harmonique** à la **pulsation** ω.
- Plus généralement, si $R < 2\sqrt{\frac{L}{C}}$, alors en posant
  $$\omega := \frac{\sqrt{4L/C - R^2}}{2L}$$
  on aura toujours un oscillateur harmonique à la pulsation ω, mais il sera **amorti** au cours du temps :
  $$y(t) = (\lambda\cos(\omega t) + \mu\sin(\omega t))e^{-\frac{R}{2L}t}$$
- Si $R > 2\sqrt{\frac{L}{C}}$, notons r_1 et r_2 les deux racines du polynôme caractéristique ; les solutions sont de la forme
  $$y(t) = \lambda e^{r_1 t} + \mu e^{r_2 t}$$
  Or, comme le produit des racines vaut $\frac{1}{LC} > 0$ et leur somme vaut $-\frac{R}{L} < 0$, on en déduit qu'elles sont toutes les deux strictement négatives. Donc λ et μ ne peuvent être nuls sans risque d'obtenir des tensions infinies quand t grandit. Quoi qu'il arrive, il n'y a plus d'oscillations du tout, seulement un amortissement qui tend vers une tension nulle.

- Enfin, si $R = 2\sqrt{\dfrac{L}{C}}$, on est en "**régime critique**", à la limite des oscillations.

Si maintenant un générateur dans le circuit à la pulsation ω_2, la solution générale de l'équation différentielle est la somme de la solution précédente et d'une solution particulière à la pulsation ω_2. En régime permanent avec $R > 0$ lorsque la composante homogène a été amortie, il ne reste plus que la partie à la pulsation ω_2. On est sorti du "**régime transitoire**" pour entrer dans le "**régime forcé**". En passant en notations complexes (voir cours de physique), on démontre que l'amplitude du signal est maximale lorsque $\omega_2 = \omega$; le circuit entre alors en **résonance**.

4.2 Équations d'ordre 2

On se propose ici de prouver un théorème admis dont la démonstration est instructive.

Théorème 21.8. *Soit l'équation $y'' + ay' + by = f(t)$ avec a et b constants et f continue sur I. Si on se donne des conditions $y(t_0) = C_0$ et $y'(t_0) = C_1$, alors l'équation admet une unique solution.*

Démonstration. On va adapter la technique de variation de la constante à l'ordre 2. Il y aura maintenant deux "constantes" à faire varier. La démonstration est assez longue, elle fonctionne en plusieurs points.

- On rappelle que les solutions de l'équation sans second membre forment un $\mathbb{K}$-ev de dimension 2. Pour en trouver une base, on considère les racines du polynôme caractéristique X^2+aX+b.
  - Deux racines distinctes : une base est donnée par le couple $(t \mapsto e^{r_1 t}, t \mapsto e^{r_2 t})$ ou par le couple $(t \mapsto \cos(\omega t)e^{st}, t \mapsto \sin(\omega t)e^{st})$, selon qu'on est dans le cas complexe ou réel.
  - Une racine double : une base est donnée par $\langle t \mapsto e^{rt}, t \mapsto te^{rt}\rangle$.

  Dans tous les cas, nous noterons (y_1, y_2) notre base. La solution générale de l'équation homogène s'écrit donc $\lambda_1 y_1 + \lambda_2 y_2$.
- Ensuite, l'idée est d'associer à toute fonction y deux fois dérivable des fonctions λ_1 et λ_2 telles que
$$\forall t \in I,\ \begin{cases} y(t) = \lambda_1(t)y_1(t) + \lambda_2(t)y_2(t) \\ y'(t) = \lambda_1(t)y_1'(t) + \lambda_2(t)y_2'(t) \end{cases}$$
  Mais en avons-nous le droit ? Pour bien mettre en évidence les inconnues, écrivons le système sous la forme
$$\begin{cases} y_1\lambda_1 + y_2\lambda_2 = y \\ y_1'\lambda_1 + y_2'\lambda_2 = y' \end{cases}$$
  soit encore $A(t)\begin{pmatrix}\lambda_1 \\ \lambda_2\end{pmatrix} = \begin{pmatrix} y \\ y'\end{pmatrix}$ avec $A(t) := \begin{pmatrix} y_1(t) & y_2(t) \\ y_1'(t) & y_2'(t)\end{pmatrix}$.
  Nous allons commencer par introduire la fonction $W := y_1y_2' - y_1'y_2$ (c'est-à-dire le déterminant de notre système). La fonction W est dérivable, et on a :
$$\begin{aligned} W' + aW &= y_1'y_2' + y_1y_2'' - y_1''y_2 - y_1'y_2' + ay_1y_2' - ay_1'y_2 \\ &= y_1(y_2'' + ay_2') - y_2(y_1'' + ay_1') \\ &= y_1(-by_2) - y_2(-by_1) \\ &= 0 \end{aligned}$$

La fonction W joue un rôle très important dans l'étude des équations différentielles, on l'appelle le **wronskien**. Ici on a $\forall t \in I,\ W(t) = \mu e^{-at}$ avec $\mu \in \mathbb{K}$ fixé. Montrons à présent que $\mu = W(0) \neq 0$.

- Si $y_1(t) = e^{r_1 t}$ et $y_2(t) = e^{r_2 t}$ on a $W(0) = r_2 - r_1 \neq 0$.
- Si $y_1(t) = e^{rt}$ et $y_2(t) = te^{rt}$ on a $W(0) = 1 \neq 0$.
- Si $y_1(t) = \cos(\omega t)e^{st}$ et $y_2(t) = \sin(\omega t)e^{st}$ on a $W(0) = \omega \neq 0$ (car les racines du polynôme caractéristique sont non réelles).

On en déduit que

$$\forall t \in I, W(t) \neq 0$$

On vérifie ensuite que

$$\forall t \in I, A(t) \begin{pmatrix} y_1 & -y_1' \\ -y_2 & y_2' \end{pmatrix} = W(t) I_2,$$

Donc pour tout $t \in I$, la matrice $A(t)$ est inversible et

$$A(t)^{-1} = \frac{1}{W(t)} \begin{pmatrix} y_1(t) & -y_1'(t) \\ -y_2(t) & y_2'(t) \end{pmatrix}$$

Si on note $\alpha(t), \beta(t), \gamma(t), \delta(t)$ les coefficients de $A(t)^{-1}$, on voit que $\alpha, \beta, \gamma, \delta$ sont des fonctions dérivables (comme quotients dont le dénominateur $W(t)$ ne s'annule pas). On peut alors poser, pour toute fonction y deux fois dérivable,

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = A^{-1} \begin{pmatrix} y \\ y' \end{pmatrix}$$

- λ_1 et λ_2 s'expriment comme sommes et produits à partir de $\alpha, \beta, \gamma, \delta, y, y'$, donc elles sont dérivables. Puis on a
$$y' = (y_1\lambda_1 + y_2\lambda_2)' = y' + y_1\lambda_1' + y_2\lambda_2'$$
d'où $y_1\lambda_1' + y_2\lambda_2' = 0$ (notons que cette égalité est toujours satisfaite, que y soit solution ou non). On a aussi :
$$\begin{aligned} y'' + ay' + by &= (y_1''\lambda_1 + y_1'\lambda_1' + y_2''\lambda_2 + y_2'\lambda_2') + a(y_1'\lambda_1 + y_2'\lambda_2) + b(y_1\lambda_1 + y_2\lambda_2) \\ &= y_1'\lambda_1' + y_2'\lambda_2' \end{aligned}$$
Finalement, y est solution ssi $A \begin{pmatrix} \lambda_1' \\ \lambda_2' \end{pmatrix} = \begin{pmatrix} 0 \\ f \end{pmatrix}$ ssi $\begin{pmatrix} \lambda_1' \\ \lambda_2' \end{pmatrix} = A^{-1} \begin{pmatrix} 0 \\ f \end{pmatrix}$
- Or $A(t)^{-1} \begin{pmatrix} 0 \\ f(t) \end{pmatrix}$ est un vecteur-colonne dont les coordonnées s'expriment comme sommes et produits à partir de $\alpha, \beta, \gamma, \delta$ et f. On peut donc écrire
$$\forall t \in I,\ A(t)^{-1} \begin{pmatrix} 0 \\ f(t) \end{pmatrix} = \begin{pmatrix} \mu_1(t) \\ \mu_2(t) \end{pmatrix}$$
où μ_1 et μ_2 sont continues. Finalement, y est solution ssi
$$\forall t \in I,\ \begin{cases} \lambda_1'(t) = \mu_1(t) \\ \lambda_2'(t) = \mu_2(t) \end{cases}$$
Puisque toute fonction continue admet une primitive, on en déduit que l'équation possède des solutions.

- Si de plus on fixe une condition de Cauchy, on a

$$\begin{cases} y(t_0) = C_0 \\ y'(t_0) = C_1 \end{cases} \iff A(t_0) \begin{pmatrix} \lambda_1(t_0) \\ \lambda_2(t_0) \end{pmatrix} = \begin{pmatrix} C_0 \\ C_1 \end{pmatrix} \iff \begin{pmatrix} \lambda_1(t_0) \\ \lambda_2(t_0) \end{pmatrix} = A(t_0)^{-1} \begin{pmatrix} C_0 \\ C_1 \end{pmatrix}$$

  donc λ_1 et λ_2 sont fixées de manière unique.

Ainsi la preuve est achevée. □

Chapitre 22

Fonctions de deux variables

Dans tout le chapitre, on munit $\mathbb{R}^2$ de sa norme euclidienne canonique, notée $||\cdot||$.

1 Fonctions continues sur un ouvert de $\mathbb{R}^2$

Définition 22.1 (Disque ouvert)**.** Soit $p \in \mathbb{R}^2$ et $r > 0$ On appelle **disque ouvert** de centre p et de rayon r l'ensemble

$$D(p,r) = \left\{ z \in \mathbb{R}^2 \middle| ||z-p|| < r \right\}$$

Définition 22.2 (Ouvert de $\mathbb{R}^2$)**.** Soit U une partie de $\mathbb{R}^2$. On dit que U est un **ouvert** de $\mathbb{R}^2$ lorsque

$$\forall p \in U,\ \exists r > 0,\ D(p,r) \subset U$$

Dans la suite, U désigne toujours un ouvert non vide de $\mathbb{R}^2$.

Définition 22.3 (Limite, continuité locale et continuité globale)**.** Soit $f : U \to \mathbb{R}$ et $p \in U$. On donne les définitions suivantes :

1. Soit $l \in \mathbb{R}$. On dit que f **tend vers** l **en** p, et on écrit $f(z) \xrightarrow[z \to p]{} l$, lorsque
$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall z \in U,\ ||z-p|| \leqslant \delta \Longrightarrow |f(z) - l| \leqslant \varepsilon$$
2. On dit que la fonction f **est continue en** p lorsque $f(z) \xrightarrow[z \to p]{} f(p)$, c'est-à-dire lorsque
$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall z \in U,\ ||z-p|| \leqslant \delta \Longrightarrow |f(z) - f(p)| \leqslant \varepsilon$$
3. On dit que f est **continue** lorsque f est continue en tout point de U.

Proposition 22.1 (Caractérisation séquentielle de la limite, HP)**.** *On a*

$$f(x,y) \xrightarrow[(x,y) \to p=(a,b)]{} l$$

si, et seulement si,

$$\forall (x_n, y_n)_{n \in \mathbb{N}} \in U^{\mathbb{N}},\ \left[x_n \xrightarrow[n \to +\infty]{} a \text{ et } y_n \xrightarrow[n \to +\infty]{} b \right] \Longrightarrow f(x_n, y_n) \xrightarrow[n \to +\infty]{} l$$

Corollaire 22.1 (Caractérisation séquentielle de la continuité, HP)**.** *f est continue en $p=(a,b)$ si, et seulement si,*

$$\forall (x_n,y_n)_{n\in\mathbb{N}}\in U^{\mathbb{N}},\ \left[x_n \xrightarrow[n\to+\infty]{} a \text{ et } y_n \xrightarrow[n\to+\infty]{} b\right] \Longrightarrow f(x_n,y_n)\xrightarrow[n\to+\infty]{} f(a,b)$$

Corollaire 22.2 (Opérations algébriques sur la continuité des fonctions de deux variables)**.** *La continuité des fonctions de deux variables passe aux combinaisons linéaires, au produit et au quotient si le dénominateur ne s'annule pas.*

Remarque 22.1. Soit $\theta\in\mathcal{C}^0(\mathbb{R},\mathbb{R})$. Les fonctions de $\mathbb{R}^2$ dans $\mathbb{R}$ $(x,y)\mapsto\theta(x)$ et $(x,y)\mapsto\theta(y)$ sont continues. En particulier, cela est vrai avec l'identité, donc les projections $(x,y)\mapsto x$ et $(x,y)\mapsto y$ sont continues, ce qui permet de montrer la continuité de nombreuses fonctions par opérations algébriques.

Proposition 22.2. *Soit $f:U\to\mathbb{R}$, I un intervalle contenant $f(U)$ et $\theta:I\to\mathbb{R}$. Si f est continue en $p\in U$, et si θ est continue en $f(p)$, alors la composée $\theta\circ f$ est continue en p.*

Proposition 22.3. *Soit $f:U\to\mathbb{R}$, I un intervalle ainsi que γ_1 et γ_2 deux fonctions de I dans $\mathbb{R}$ telles que*

$$\forall t\in I,\ (\gamma_1(t),\gamma_2(t))\in U$$

Si γ_1 et γ_2 sont continues en $a\in I$, et si f est continue en $(\gamma_1(a),\gamma_2(a))$, alors la composée $t\mapsto f(\gamma_1(t),\gamma_2(t))$ est continue en a.

Proposition 22.4. *Soit $f:U\to\mathbb{R}$, V un ouvert de $\mathbb{R}^2$ et $\varphi,\psi:V\to\mathbb{R}$ deux fonctions telles que*

$$\forall z\in V,\ (\varphi(z),\psi(z))\in U$$

Si φ et ψ sont continues en $p\in V$, et si f est continue en $(\varphi(p),\psi(p))$, alors la composée $z\mapsto f(\varphi(z),\psi(z))$ est continue en p.

2 Dérivation

Définition 22.4 (Applications partielles)**.** Étant donné un point $p=(a,b)\in U$, on considère les ensembles

$$D_1=\{x\in\mathbb{R}|(x,b)\in U\} \text{ et } D_2=\{y\in\mathbb{R}|(a,y)\in U\}$$

ainsi que les applications partielles

$$\begin{array}{rccc}\varphi_1: & D_1 & \longrightarrow & \mathbb{R}\\ & x & \longmapsto & f(x,b)\end{array}$$

et

$$\begin{array}{rccc}\varphi_2: & D_2 & \longrightarrow & \mathbb{R}\\ & y & \longmapsto & f(a,y)\end{array}$$

Définition 22.5 (Dérivées partielles)**.** Soit $p=(a,b)\in U$.

1. Si l'application partielle φ_1 est dérivable en a, on dit que f admet une **première dérivée partielle** en (a,b) et l'on pose
$$\partial_1 f(a,b)=\varphi_1'(a)$$

2. Si l'application partielle φ_2 est dérivable en a, on dit que f admet une **deuxième dérivée partielle** en (a,b) et l'on pose
$$\partial_2 f(a,b) = \varphi_2'(b)$$

Remarque 22.2. En pratique, pour une fonction $f : (x,y) \mapsto f(x,y)$, on utilise les notations plus parlantes
$$\frac{\partial f}{\partial x}(a,b) = \partial_1 f(a,b)$$
et
$$\frac{\partial f}{\partial y}(a,b) = \partial_2 f(a,b)$$
en s'adaptant au nom des variables utilisées dans la définition de f.

Proposition 22.5 (Combinaison linéaire et produit)**.** *Soit f et g deux fonctions de U dans $\mathbb{R}$, λ et μ des réels et $p \in U$. Supposons que f et g admettent des dérivées partielles en p.*

1. *Alors $\lambda f + \mu g$ également, avec*
$$\frac{\partial(\lambda f + \mu g)}{\partial x}(p) = \lambda \frac{\partial f}{\partial x}(p) + \mu \frac{\partial g}{\partial x}(p)$$
et
$$\frac{\partial(\lambda f + \mu g)}{\partial y}(p) = \lambda \frac{\partial f}{\partial y}(p) + \mu \frac{\partial g}{\partial y}(p)$$
2. *Et fg également, avec*
$$\frac{\partial(fg)}{\partial x}(p) = \frac{\partial f}{\partial x}(p) g(p) + f(p) \frac{\partial g}{\partial x}(p)$$
et
$$\frac{\partial(fg)}{\partial y}(p) = \frac{\partial f}{\partial y}(p) g(p) + f(p) \frac{\partial g}{\partial y}(p)$$

Remarque 22.3. Attention, l'existence de dérivées partielles, même en tout point de U, n'entraîne pas la continuité de f. Comme contre-exemple, on peut citer $(x,y) \mapsto \dfrac{xy}{x^2+y^2}$ prolongée par 0 en $(0,0)$ qui admet des dérivées partielles en tout point mais qui n'est pourtant pas continue en $(0,0)$.

Définition 22.6 (Fonctions de classe $\mathcal{C}^1$)**.** La fonction f est dite **de classe $\mathcal{C}^1$** lorsqu'elle admet admet des dérivées partielles en tout point de U et que les fonctions $\dfrac{\partial f}{\partial x}$ et $\dfrac{\partial f}{\partial y}$ définies alors sur U sont continues.
On note $\mathcal{C}^1(U,\mathbb{R})$ l'ensemble des fonctions de classe $\mathcal{C}^1$ définies sur U.

Théorème 22.1 (Développement limité à l'ordre 1)**.** *Soit $f \in \mathcal{C}^1(U,\mathbb{R})$ et $p = (a,b) \in U$. Il existe une fonction $\varepsilon : U \to \mathbb{R}$ telle que $\varepsilon(z) \xrightarrow[z \to p]{} 0$ et*
$$\forall (x,y) \in U,\ f(x,y) = f(a,b) + (x-a)\frac{\partial f}{\partial x}(a,b) + (y-b)\frac{\partial f}{\partial y}(a,b) + \varepsilon(x,y)||(x,y)-(a,b)||$$

Définition 22.7 (Plan tangent)**.** Le théorème précédent affirme que la fonction affine

$$(x,y) \mapsto f(a,b) + (x-a)\frac{\partial f}{\partial x}(a,b) + (y-b)\frac{\partial f}{\partial y}(a,b)$$

approche au premier ordre f au voisinage de (a,b). On appelle alors **plan tangent du graphe de la fonction f en** (a,b) le plan d'équation

$$z = f(a,b) + (x-a)\frac{\partial f}{\partial x}(a,b) + (y-b)\frac{\partial f}{\partial y}(a,b)$$

Corollaire 22.3 ($\mathcal{C}^1 \Longrightarrow \mathcal{C}^0$)**.** *Si $f \in \mathcal{C}^1(U,\mathbb{R})$, alors f est continue.*

Corollaire 22.4. *Soit $(f,g) \in (\mathcal{C}^1(U,\mathbb{R}))^2$ et $\lambda,\mu \in \mathbb{R}$. Alors $\lambda f + \mu g$ et fg sont des fonctions de classe $\mathcal{C}^1$.*

Définition 22.8 (Gradient, champ de vecteurs)**.** Soit $f \in \mathcal{C}^1(U,\mathbb{R})$. Le **gradient** de f est l'application (lire ∇ "nabla")

$$\begin{array}{rccc} \nabla f : & U & \longrightarrow & \mathbb{R}^2 \\ & (a,b) & \longmapsto & \nabla f(a,b) = \begin{pmatrix} \dfrac{\partial f}{\partial x}(a,b) \\ \\ \dfrac{\partial f}{\partial y}(a,b) \end{pmatrix} \end{array}$$

Le gradient est un **champ de vecteurs**, c'est-à-dire une application dont les valeurs sont des vecteurs.

Remarque 22.4. On peut réécrire le DL à l'ordre 1 à l'aide du gradient et du produit scalaire usuel de $\mathbb{R}^2$ (avec les mêmes hypothèses que dans le théorème)

$$\forall z \in U,\ f(z) = f(p) + \langle \nabla f(p), z-p\rangle + \varepsilon(z)||z-p||$$

3 Dérivation des fonctions composées

Proposition 22.6. *Soit $f \in \mathcal{C}^1(U,\mathbb{R})$, I un intervalle non trivial contenant $f(U)$ et $\theta \in \mathcal{C}^1(I,\mathbb{R})$. Alors $\theta \circ f$ est une fonction de classe $\mathcal{C}^1$ et on a*

$$\forall p \in U,\ \frac{\partial(\theta\circ f)}{\partial x}(p) = \theta'(f(p))\frac{\partial f}{\partial x}(p)$$

et

$$\forall p \in U,\ \frac{\partial(\theta\circ f)}{\partial y}(p) = \theta'(f(p))\frac{\partial f}{\partial y}(p)$$

Théorème 22.2 (Première règle de la chaîne)**.** *Soit $f \in \mathcal{C}^1(U,\mathbb{R})$ et $(\gamma_1,\gamma_2) \in (\mathcal{C}^1(I,\mathbb{R}))^2$. On suppose que*

$$\forall t \in I,\ (\gamma_1(t),\gamma_2(t)) \in U$$

Alors

$$g : t \mapsto f(\gamma_1(t), \gamma_2(t))$$

est de classe $\mathcal{C}^1$ *et on a*

$$\forall a \in I,\ g'(a) = \frac{\partial f}{\partial x}(\gamma_1(a), \gamma_2(a))\gamma_1'(a) + \frac{\partial f}{\partial y}(\gamma_1(a), \gamma_2(a))\gamma_2'(a)$$

Définition 22.9 (Ligne de niveau, HP)**.** On appelle **ligne de niveau** d'une fonction $f : \mathbb{R}^2 \mapsto \mathbb{R}$ une partie de $\mathbb{R}^2$ d'équation $f(x, y) = c$ pour un certain $c \in \mathbb{R}$.
Si la fonction γ paramètre une ligne de niveau, c'est-à-dire si la fonction $f \circ \gamma$ est constante, on obtient l'égalité

$$\langle \nabla f(\gamma(a)), \gamma'(a) \rangle = 0$$

c'est-à-dire que le gradient $\nabla f(\gamma(a))$ est orthogonal au vecteur dérivé $\gamma'(a)$ qui dirige la tangente au point $\gamma(a)$ de la ligne de niveau. On dit que le gradient ∇f est orthogonal aux lignes de niveau.

Définition 22.10 (Dérivée selon un vecteur)**.** Soit $f \in \mathcal{C}^1(U, \mathbb{R})$, $p \in U$ et $v = \begin{pmatrix} h \\ k \end{pmatrix} \in \mathbb{R}^2$. L'application

$$\varphi_v : t \mapsto f(p + tv)$$

est définie sur une union d'intervalle ouverts dont au moins un contient 0. De plus, elle est dérivable en 0, de dérivée :

$$\varphi_v'(0) = h\frac{\partial f}{\partial x}(p) + k\frac{\partial f}{\partial y}(p) = \langle \nabla f(p), v \rangle$$

On note $D_v f(p) = \varphi_v'(0)$ cette dérivée, et on l'appelle **dérivée de** f **en** p **selon le vecteur** v.

Remarque 22.5. La notion de dérivée selon un vecteur généralise celle des dérivées partielles en ne faisant plus jouer un rôle particulier aux droites parallèles aux axes de coordonnées.

Remarque 22.6. Le gradient de f en p "pointe" dans la direction dans laquelle f croît le plus vite. Cela correspond à la direction de plus grande pente sur le graphe Γ_f.

Théorème 22.3 (Deuxième règle de la chaîne)**.** *Soit* $f \in \mathcal{C}^1(U, \mathbb{R})$*. Soit* V *un ouvert non vide de* $\mathbb{R}^2$ *et* $(\varphi, \psi) \in (\mathcal{C}^1(V, \mathbb{R}))^2$ *un couple de fonctions tel que*

$$\forall z \in V,\ (\varphi(z), \psi(z)) \in U$$

Alors l'application composée

$$\begin{array}{rccc} g : & V & \longrightarrow & \mathbb{R} \\ & z & \longmapsto & f(\varphi(z), \psi(z)) \end{array}$$

est de classe $\mathcal{C}^1$ *et vérifie*

$$\forall p \in V,\ \partial_1 g(p) = \partial_1 f(\varphi(p), \psi(p))\partial_1\varphi(p) + \partial_2 f(\varphi(p), \psi(p))\partial_1\psi(p)$$

et

$$\forall p \in V,\ \partial_2 g(p) = \partial_1 f(\varphi(p), \psi(p))\partial_2\varphi(p) + \partial_2 f(\varphi(p), \psi(p))\partial_2\psi(p)$$

4 Extrema

Définition 22.11 (Extrema)**.** Soit X une partie non vide de $\mathbb{R}^2$, $f : X \to \mathbb{R}$ et $p \in X$.

1. On dit que f admet un **maximum** en p lorsque $\forall z \in X,\ f(z) \leqslant f(p)$
2. On dit que f admet un **maximum local** lorsque
$$\exists \delta > 0,\ \forall z \in X \cap D(p, \delta),\ f(z) \leqslant f(p)$$
3. On dit que f admet un **minimum** en p lorsque $\forall z \in X,\ f(z) \geqslant f(p)$
4. On dit que f admet un **minimum local** lorsque
$$\exists \delta > 0,\ \forall z \in X \cap D(p, \delta),\ f(z) \geqslant f(p)$$
5. On dit que f un **extremum** en p en si f admet un maximum ou un minimum en p.
6. On dit que f un **extremum local** en p en si f admet un maximum local ou un minimum local en p.

Remarque 22.7. Pour insister sur la distinction avec les extrema locaux, les extrema sont parfois appelés **extrema globaux**.

Définition 22.12 (Point critique)**.** Soit $f \in \mathcal{C}^1(U, \mathbb{R})$ et $p \in U$. On dit que p est un **point critique** pour f lorsque
$$\frac{\partial f}{\partial x}(p) = \frac{\partial f}{\partial y}(p) = 0$$
c'est-à-dire lorsque $\nabla f(p) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

Théorème 22.4 (Condition nécessaire d'extremum local)**.** *Soit $f \in \mathcal{C}^1(U, \mathbb{R})$ et $p \in U$. Si p admet un extremum local en p, alors p est un point critique de f.*

Chapitre 23

Séries

Dans tout le chapitre, les suites sont considérées à valeurs dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

1 Généralités

Définition 23.1 (Série, somme partielle)**.** Soit $(u_n)_{n\in\mathbb{N}}$ une suite à termes dans $\mathbb{K}$. La **série de terme général** u_n désigne en fait la suite $\left(\sum_{k=0}^{n} u_k\right)_{n\in\mathbb{N}}$. On la note $\sum u_n$. Le nombre $\sum_{k=0}^{n} u_k$ s'appelle la **somme partielle de rang** n de la série $\sum u_n$. Dans ce chapitre, nous utiliserons la notation très répandue S_n pour la somme partielle de rang n.

Définition 23.2 (Série convergente, somme de la série)**.** La série $\sum u_n$ est dite **convergente** lorsque la suite $(S_n)_{n\in\mathbb{N}}$, et la limite des sommes partielles s'appelle alors la **somme de la série**. On la note

$$\sum_{n=0}^{+\infty} u_n$$

Définition 23.3. Plus généralement, si $n_0 \in \mathbb{N}$ et si $(u_n)_{n\geqslant n_0}$ est une suite à termes réels ou complexes, la série $\sum u_n$ désigne la suite des sommes partielles

$$\left(\sum_{k=0}^{n} u_k\right)_{n\geqslant n_0}$$

S'il y a ambiguïté, on pourra la noter $\sum_{n\geqslant n_0} u_n$. Si elle converge, sa somme est notée

$$\sum_{n=n_0}^{+\infty} u_n$$

Les résultats qui suivent sont énoncés dans le cas où $n_0 = 0$, mais leurs énoncés symétriques dans le cas général sont également vrais.

Proposition 23.1 (Troncature)**.** *Soit $\sum u_n$ une série et $p \in \mathbb{N}^\star$. Alors la série $\sum u_n$ converge si, et seulement si, la série tronquée $\sum_{n\geqslant p} u_n$ converge, et dans ce cas, on a :*

$$\sum_{n=0}^{+\infty} u_n = \sum_{n=0}^{p-1} u_n + \sum_{n=p}^{+\infty} u_n$$

Proposition 23.2 (Décalage d'indice)**.** *Soit $p \in \mathbb{N}^\star$ et $\sum_{n\geqslant p} u_n$ une série. Alors $\sum_{n\geqslant p} u_n$ converge si, et seulement si, $\sum_{n\geqslant 0} u_{n+p}$ converge, et dans ce cas on a*

$$\sum_{n=0}^{+\infty} u_{n+p} = \sum_{n=p}^{+\infty} u_n$$

Définition 23.4 (Reste d'ordre n d'une série convergente)**.** Soit $\sum u_n$ une série **convergente** et $n \in \mathbb{N}$. Le **reste d'ordre** n de la série $\sum u_n$ est défini par

$$R_n = \sum_{k=n+1}^{+\infty} u_k = \sum_{k>n}^{+\infty} u_k$$

Il vérifie

$$S_n + R_n = \sum_{k=0}^{+\infty} u_k$$

Corollaire 23.1 (Condition nécessaire de convergence sur les restes)**.** ***Si*** *$\sum u_n$ converge,* ***alors*** $R_n \xrightarrow[n\to+\infty]{} 0$.

Corollaire 23.2 (Condition nécessaire de convergence sur le terme général)**.** ***Si*** *$\sum u_n$ converge,* ***alors*** $u_n \xrightarrow[n\to+\infty]{} 0$.

Remarque 23.1. Attention, la réciproque est fausse ! Comme contre exemple, on pourra considérer la série harmonique.

Définition 23.5 (Divergence)**.** Une série est dite **divergente** lorsque la suite de ses sommes partielles diverge.

Définition 23.6 (Divergence grossière)**.** La série $\sum u_n$ est dite **grossièrement divergente** lorsque u_n ne tend pas vers 0 lorsque n tend vers $+\infty$.

Théorème 23.1 (Séries géométriques)**.** *Soit $z \in \mathbb{C}$. La série $\sum z^n$ converge si, et seulement si, $|z| < 1$, et dans ce cas se somme vaut $\dfrac{1}{1-z}$*

Théorème 23.2 (Linéarité). *Soit $\sum u_n$ et $\sum v_n$ des séries convergentes et λ et μ des scalaires. Alors la série $\sum(\lambda u_n + \mu v_n)$ est convergente et sa somme vaut*

$$\sum_{n=0}^{+\infty}(\lambda u_n + \mu v_n) = \lambda \sum_{n=0}^{+\infty} u_n + \mu \sum_{n=0}^{+\infty} v_n$$

Théorème 23.3 (Théorème passerelle pour les séries complexes). *On suppose que $\sum u_n$ est à termes complexes. Alors $\sum u_n$ converge si, et seulement si, les séries réelles $\sum Re(u_n)$ et $\sum Im(u_n)$ convergent, et dans ce cas on a :*

$$\sum_{n=0}^{+\infty} u_n = \sum_{n=0}^{+\infty} Re(u_n) + i\sum_{n=0}^{+\infty} Im(u_n)$$

Proposition 23.3 (Lien suite-série). *La suite $(u_n)_{n\in\mathbb{N}}$ converge si, et seulement si, la série $(u_{n+1} - u_n)$ converge.*

Remarque 23.2. Cette proposition évidente peut paraître inutile, mais elle permet d'étudier la nature d'une suite grâce à des outils sur les séries et peut donc se révéler très précieuse dans certains cas.

2 Séries réelles positives

Proposition 23.4. *Soit $\sum u_n$ une série à termes réels positifs. Alors elle est convergente si, et seulement si, la suite $(S_n)_{n\in\mathbb{N}}$ est majorée. Dans ce cas, on a :*

$$\forall n \in \mathbb{N},\ S_n \leqslant S$$

où S désigne la somme de la série $\sum u_n$.

Théorème 23.4. *Soit $\sum u_n$ et $\sum v_n$ des séries à termes réels positifs. Si $u_n \leqslant v_n$ pour tout $n \in \mathbb{N}$ et si $\sum v_n$, alors $\sum u_n$ converge, et dans ce cas on a :*

$$\sum_{n=0}^{+\infty} u_n \leqslant \sum_{n=0}^{+\infty} v_n$$

Remarque 23.3. Si l'inégalité $u_n \leqslant v_n$ n'a lieu qu'à partir d'un certain rang, alors $\sum u_n$ continue à converger, mais l'inégalité sur les sommes n'a plus aucune raison d'être valable.

Théorème 23.5. *Soit $\sum u_n$ et $\sum v_n$ des séries à termes réels positifs. Si $u_n \underset{n\to+\infty}{\sim} v_n$, alors $\sum u_n$ et $\sum v_n$ ont même nature.*

Théorème 23.6 (Critère spécial des séries alternées). *Soit $(u_n)_{n\in\mathbb{N}}$ une suite réelle positive décroissante qui tend vers 0. Alors*

1. *D'une part, la série* $\sum(-1)^n u_n$ *converge*
2. *(Contrôle de la somme et des restes) D'autre part, pour tout* $n \in \mathbb{N}$, $\sum_{k=n}^{+\infty}(-1)^k u_k$ *est du signe de* $(-1)^n$ *au sens large, et on a*
$$\left|\sum_{k=n}^{+\infty}(-1)^k u_k\right| \leqslant u_n$$

Remarque 23.4. On a le résultat symétrique avec la série $\sum(-1)^{n-1}u_n$.

3 Comparaison série-intégrale

Donnons-nous $f : \mathbb{R}_+ \to \mathbb{R}$ continue par morceaux. Lorsque l'on considère la série $\sum_{n\geqslant 1} f(n)$, la somme partielle de rang n, $\sum_{k=1}^{n} f(n)$, peut se réécrire un peu artificiellement $\sum_{k=1}^{n}\int_k^{k+1} f(k)dt$. Ici, on s'intéresse au cas où f est monotone. A chaque fois, il est recommandé de faire un dessin avec les rectangles.

Par exemple, supposons que f soit décroissante. On peut justifier à l'oral avec un dessin bien fait, mais formellement il faudrait écrire

$$\forall k \in \mathbb{N}^\star,\ \forall t \in [k-1,k],\ f(t) \geqslant f(k) \ \text{ donc } \ f(k) \leqslant \int_{k-1}^{k} f(t)dt$$

et

$$\forall k \in \mathbb{N}^\star,\ \forall t \in [k,k+1],\ f(k) \geqslant f(t) \ \text{ donc } \ f(k) \geqslant \int_{k}^{k+1} f(t)dt$$

d'où, en sommant et en utilisant la relation de Chasles, on obtient l'encadrement

$$\int_1^{n+1} f(t)dt \leqslant S_n \leqslant \int_0^n f(t)dt$$

On obtient évidemment un résultat analogue dans le cas où f est croissante. On peut adapter la méthode selon l'intervalle de définition de f.

Théorème 23.7 (Séries de Riemann)**.** *Soit* $\alpha \in \mathbb{R}$. *La série* $\sum \frac{1}{n^\alpha}$ *converge si, et seulement si* $\alpha > 1$.

Exemple 23.1. La série harmonique $\sum_{n\geqslant 1}\frac{1}{n}$ diverge car elle correspond au cas où $\alpha = 1$.

4 Convergence absolue

Définition 23.7 (Convergence absolue)**.** Soit $\sum u_n$ une série à termes complexes. On dit que la série $\sum u_n$ est absolument convergente lorsque la série à termes positifs $\sum|u_n|$ converge.

Théorème 23.8 (Convergence absolue $\Longrightarrow$ convergence)**.** *Soit $\sum u_n$ une série à termes complexes.* ***Si*** *$\sum |u_n|$ converge (ie si $\sum u_n$ converge absolument),* ***alors*** *$\sum u_n$ converge.*

Définition 23.8 (Prolongement de l'exponentielle à $\mathbb{C}$)**.** Pour tout $z \in \mathbb{C}$, la série $\sum \dfrac{z^n}{n!}$ est absolument convergente. On **définit** alors

$$\exp(z) = \sum_{n=0}^{+\infty} \frac{z^n}{n!}$$

Corollaire 23.3 (Règle du grand O)**.** *Soit $\sum u_n$ une série à termes complexes et $\sum v_n$ une série à termes réels positifs convergente. Si $u_n \underset{n\to+\infty}{=} O(v_n)$, alors la série $\sum u_n$ converge absolument, donc converge.*

Théorème 23.9 (Règle de D'Alembert, HP)**.** *Soit $\sum u_n$ une série à termes complexes non nuls. Supposons que*

$$\left|\frac{u_{n+1}}{u_n}\right| \xrightarrow[n\to+\infty]{} \lambda \in [0, +\infty]$$

Si $\lambda < 1$, alors la série converge absolument. Si $\lambda > 1$, alors elle diverge grossièrement. Si $\lambda = 1$, on ne peut rien dire.

Démonstration. Dans le premier cas, on prend $\alpha \in]\lambda, 1[$ et le module du quotient est plus petit que α APCR. Alors, par une récurrence immédiate ou un télescopage, on montre que $u_n = O(\alpha^n)$ et alors la série converge absolument. Dans l'autre cas, $|u_n| \geqslant |u_{n_0}|$ APCR, donc u_n ne tend pas vers 0 et la série diverge grossièrement. Pour le cas $\lambda = 1$, tous les cas de figure sont possibles. □

5 Représentation décimale des réels, HP

Définition 23.9 (Développement décimal)**.** Soit x un réel positif ou nul. Un **développement décimal** de x est une écriture de la forme

$$x = \sum_{n=0}^{+\infty} \frac{d_n}{10^n}$$

où $d_0 \in \mathbb{N}$ et $\forall n \in \mathbb{N}^\star,\ d_n \in [\![0, 9]\!]$. Les d_n (pour $n \in \mathbb{N}^\star$) sont appelés les **décimales** de x, et on écrit :

$$x = d_0, d_1 d_2 \dots$$

Remarque 23.5. Par convention, si $x < 0$, l'écriture $x = -d_0, d_1 d_2 \dots$ signifiera que x est l'opposé de $d_0, d_1 d_1 \dots$.

Définition 23.10 (Développement décimal propre, impropre)**.** Si tous les d_n valent 9 à partir d'un certain rang, le développement décimal est dit **impropre**. Sinon, il est dit **propre**.

Théorème 23.10. *Tout réel $x \geqslant 0$ admet un unique développement décimal propre.*

Remarque 23.6. On a le même résultat dans une base b entière supérieure ou égale à 2 quelconque.

Corollaire 23.4 (Indénombrabilité de $\mathbb{R}$)**.** $\mathbb{R}$ *est indénombrable.*

Démonstration. On utilise l'argument diagonal de Cantor qui utilise l'existence-unicité du développement décimal propre des réels positifs. L'argument est développé dans la partie "Dénombrabilité" du chapitre "Dénombrements". □

Chapitre 24

Sommabilité

Dans tout ce qui suit, $\mathcal{P}_f(I)$ désigne l'ensemble des parties finies de I.

1 Familles réelles positives

Dans tout le paragraphe, les familles sont à termes dans $[0, +\infty]$.

Définition 24.1 (Ajout de $+\infty$ à $\mathbb{R}_+$)**.** Considérons l'ensemble $[0, +\infty] = \mathbb{R}_+ \cup \{+\infty\}$. On prolonge $\leqslant$, $+$ et $\times$ de la manière suivante :

1. $\forall x \in [0, +\infty],\ x \leqslant +\infty$
2. $\forall x \in [0, +\infty],\ x + (+\infty) = (+\infty) + x = +\infty$
3. $\forall x \in]0, +\infty],\ x \times (+\times) = (+\infty) \times x = +\infty$
4. $0 \times (+\infty) = (+\infty) \times 0 = 0$

Proposition 24.1. *On vérifie immédiatement les points suivants :*

1. $\leqslant$ *reste une relation d'ordre totale*
2. $+$ *reste associative et commutative*
3. $\times$ *reste associative, commutative et distributive sur* $+$
4. $\leqslant$ *reste compatible avec* $+$ *et* $\times$

Proposition 24.2. *Dans* $[0, +\infty]$*, les résultats usuels sur les sommes et les produits finis restent valables : relation de Chasles, associativité, distributivité, etc.*

Théorème 24.1 (Propriété de la borne supérieure prolongée)**.** *Toute partie non vide de* $[0, +\infty]$ *admet une borne supérieure dans* $[0, +\infty]$*.*

Définition 24.2 (Somme d'une famille, sommabilité)**.** Soit $(a_i)_{i \in I}$ une famille à termes dans $[0, +\infty]$. l'ensemble des sommes finies

$$\left\{ \sum_{i \in J} a_i \middle| J \in \mathcal{P}_f(I) \right\}$$

est non vide car il contient la somme vide qui vaut 0. La **somme de la famille** $(a_i)_{i\in I}$ est définie par

$$\sum_{i\in I} a_i = \sup_{J\in\mathcal{P}_f(I)} \left(\sum_{i\in J} a_i\right)$$

La famille est dite **sommable** lorsque

$$\sum_{i\in I} a_i < +\infty$$

Remarque 24.1. Dès que l'un des a_i vaut $+\infty$, la somme de la famille vaut $+\infty$.

Remarque 24.2. La définition coïncide avec celle donnée pour les sommes finies et les sommes presque nulles.

Proposition 24.3. *Soit $(a_i)_{i\in I}$ une famille de somme S. Alors*

1. *Toute sous-famille est de somme $\leqslant S$*
2. *Toute famille majorée par la famille $(a_i)_{i\in I}$ est de somme $\leqslant S$*

Définition 24.3. Soit $(a_n)_{n\in\mathbb{N}}$ une famille à termes dans $\mathbb{R}_+$. Puisque la suite des sommes partielles de la série $\sum a_n$ admet de toute façon une limite, on s'autorisera à écrire $\sum_{n=0}^{+\infty} a_n$ dans tous les cas. Ainsi, si la série diverge, on écrira symboliquement

$$\sum_{n=0}^{+\infty} a_n = +\infty$$

Proposition 24.4 (Cas des séries, cas positif)**.** *Soit $(a_n)_{n\in\mathbb{N}}$ une famille à termes dans $\mathbb{R}_+$. Alors "la somme de la famille est égale à la somme de la série". Pour essayer d'éviter les confusions, on prendra garde à utiliser autant que possible deux notations différentes. Par exemple :*

$$\sum_{n\in\mathbb{N}} a_n = \sum_{n=0}^{+\infty} a_n$$

Corollaire 24.1. *La famille $(a_n)_{n\in\mathbb{N}}$ une famille à termes dans $\mathbb{R}_+$ est sommable si, et seulement si, la série $\sum a_n$ converge, et dans ce cas la somme de la famille est égale à la somme de la série.*

Remarque 24.3. On a le même résultat en considérant une famille $(a_n)_{n\geqslant n_0}$ une famille à termes dans $\mathbb{R}_+$ avec $n_0 \in \mathbb{Z}$.

Théorème 24.2 (Dénombrabilité du support d'une famille sommable, HP)**.** *Une famille sommable possède un support au plus dénombrable.*

Démonstration. Il suffit de remarquer que pour tout $n \in \mathbb{N}^\star$, la famille ne peut prendre ses valeurs qu'un nombre fini de fois dans $\left]\dfrac{1}{n+1}, \dfrac{1}{n}\right]$ et de même avec $]1, +\infty[$, sans quoi la famille ne serait pas sommable. Or ces ensembles partitionnent $\mathbb{R}_+^\star$ et leur ensemble est dénombrable. Le support de la famille est donc l'union au plus dénombrable (dénombrable en fait) d'ensemble au plus dénombrables (finis même) : c'est donc un ensemble au plus dénombrable. □

Exemple 24.1 (Séries de Riemann)**.** Si $s > 1$, la famille $\left(\frac{1}{n^s}\right)_{n\in\mathbb{N}^\star}$ est sommable.

Théorème 24.3 (Changement d'indice, cas positif)**.** *Soit $(a_i)_{i\in I}$ une famille à termes dans $[0,+\infty]$ et $\sigma : J \to I$ une bijection. Alors*

$$\sum_{i\in I} a_i = \sum_{j\in J} a_{\sigma(j)}$$

Remarque 24.4. En particulier, la sommabilité est une propriété invariante par permutation de termes.

Théorème 24.4 (Linéarité, cas positif)**.** *Soit $(a_i)_{i\in I}$ et $(b_i)_{i\in I}$ des familles à termes dans $[0,+\infty]$. Soit $\lambda,\mu \in [0,+\infty]$. On a*

$$\sum_{i\in I}(\lambda a_i + \mu b_i) = \lambda\sum_{i\in I} a_i + \mu\sum_{i\in I} b_i$$

Théorème 24.5 (Relation de Chasles, cas positif)**.** *Soit $(a_i)_{i\in I}$ une famille à termes dans $[0,+\infty]$. Soient I_1 et I_2 deux ensembles disjoints tels que $I = I_1 \sqcup I_2$. On a*

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

Remarque 24.5. Par une récurrence immédiate, on généralise à un nombre fini d'ensemble deux à deux disjoints.

Théorème 24.6 (Sommation par paquets, cas positif)**.** *Soit $(a_i)_{i\in I}$ une famille à termes dans $[0,+\infty]$. Soit $(I_t)_{t\in T}$ une famille de parties de I deux à deux disjointes telle que $I = \bigsqcup_{t\in T} I_t$. On a*

$$\sum_{i\in I} a_i = \sum_{t\in T}\left(\sum_{i\in I_t} a_i\right)$$

Corollaire 24.2 (Théorème de Fubini, cas positif)**.** *Soit $(a_{i,j})_{(i,j)\in I\times J}$ une famille à termes dans $[0,+\infty]$. On a :*

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

Remarque 24.6. Représenter l'ensemble de sommation, notamment pour les sommations triangulaires peut aider : on peut passer d'une sommation ligne par ligne à une sommation colonne par colonne. Il vaut mieux retenir cela qu'apprendre par cœur des formules complexes.

2 Famille réelles quelconques

On procèdera toujours en deux temps : d'abord, on vérifie l'hypothèse de sommabilité en effectuant tous les calculs que l'on souhaite après être passées à la valeur absolue grâce au paragraphe précédent ; ensuite on applique les théorèmes du paragraphe. Il existe une variante qui consiste à faire du calcul formel, c'est-à-dire à calculer comme si on avait l'hypothèse de sommabilité, puis à constater que cette hypothèse est vérifiée en expliquant pourquoi avec un passage aux valeurs absolues.

Définition 24.4 (Sommabilité d'une famille réelle quelconque)**.** Une famille réelle $(a_i)_{i\in I}$ est dite **sommable** lorsque

$$\sum_{i\in I}|a_i| < +\infty$$

autrement dit lorsque la famille réelle positive $(|a_i|)_{i\in I}$ est sommable. Dans ce cas, les familles à termes positifs $(a_i^+)_{i\in I}$ et $(a_i^-)_{i\in I}$ sont sommables (par comparaison) et on **pose** :

$$\sum_{i\in I} a_i = \sum_{i\in I} a_i^+ - \sum_{i\in I} a_i^-$$

Remarque 24.7. Dans le cas où la famille est positive, on retrouve bien la définition initiale. Même remarque dans le cas des familles finies ou presque nulles.

Théorème 24.7 (Inégalité triangulaire, cas réel)**.** *Si la famille réelle $(a_i)_{i\in I}$ est sommable, alors*

$$\left|\sum_{i\in I} a_i\right| \leqslant \sum_{i\in I}|a_i|$$

On observe désormais si les résultats du paragraphe précédent restent vrais.

Proposition 24.5. *Toute sous-famille d'une famille sommable reste sommable, mais on ne peut rien dire sur sa somme. Attention, on ne peut rien dire d'une famille majorée, à moins de majorer la famille $(|a_i|)_{i\in I}$.*

Théorème 24.8 (Cas des séries, cas réel)**.** *Soit $(a_i)_{i\in\mathbb{N}}$ une famille réelle quelconque. La famille $(a_i)_{i\in\mathbb{N}}$ est sommable si, et seulement si, la série $\sum a_n$ est absolument convergente, et dans ce cas on a :*

$$\sum_{i\in\mathbb{N}} a_i = \sum_{i=0}^{+\infty} a_i$$

Théorème 24.9 (Changement d'indice, cas réel)**.** *Soit $(a_i)_{i\in I}$ une famille réelle quelconque et $\sigma : J \to I$ une bijection. Alors la famille $(a_i)_{i\in I}$ est sommable si, et seulement si, la famille $(a_{\sigma(j)})_{j\in J}$, et dans ce cas on a :*

$$\sum_{i\in I} a_i = \sum_{j\in J} a_{\sigma(j)}$$

Théorème 24.10 (Linéarité, cas réel)**.** *Soit $\lambda, \mu \in \mathbb{R}$. Si $(a_i)_{i\in I}$ et $(b_i)_{i\in I}$ réelles sont sommables, alors $(\lambda a_i + \mu b_i)_{i\in I}$ est sommable et on a :*

$$\sum_{i\in I}(\lambda a_i + \mu b_i) = \lambda \sum_{i\in I} a_i + \mu \sum_{i\in I} b_i$$

Théorème 24.11 (Relation de Chasles, cas réel)**.** *Si I_1 et I_2 sont deux ensembles disjoints tels que $I = I_1 \sqcup I_2$, alors $(a_i)_{i\in I}$ est sommable si, et seulement si, $(a_i)_{i\in I_1}$ et $(a_i)_{i\in I_2}$ son sommables, et dans ce cas on a :*

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

Par une récurrence immédiate, on généralise à un nombre fini d'ensembles deux à deux disjoints.

Théorème 24.12 (Sommation par paquets, cas réel)**.** *Soit $(a_i)_{i\in I}$ une famille réelle quelconque. Soit $(I_t)_{t\in T}$ une famille de parties de I deux à deux disjointes telle que $I = \bigsqcup_{t\in T} I_t$.* ***Si*** *$(a_i)_{i\in I}$ est sommable,* ***alors*** *chaque $(a_i)_{i\in I_t}$ est sommable et la famille $\left(\sum_{i\in I_t} a_i\right)_{t\in T}$ est sommable et on a :*

$$\sum_{i\in I} a_i = \sum_{t\in T}\left(\sum_{i\in I_t} a_i\right)$$

Corollaire 24.3 (Théorème de Fubini, cas réel)**.** *Soit $(a_{i,j})_{(i,j)\in I\times J}$ une famille réelle sommable. On a :*

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

3 Familles complexes

Définition 24.5 (Sommabilité d'une famille complexe)**.** Une famille complexe $(a_i)_{i\in I} = (x_i + \mathrm{i}y_i)_{i\in I}$ est dite **sommable** lorsque

$$\sum_{i\in I} |a_i| < +\infty$$

Dans ce cas, les familles réelles $(x_i)_{i\in I}$ et $(y_i)_{i\in I}$ sont sommables et on **pose** :

$$\sum_{i\in I} a_i = \sum_{i\in I} x_i + \mathrm{i}\sum_{i\in I} y_i$$

L'ensemble des familles sommables est noté $l^1(I)$.

Remarque 24.8. Dans le cas où la famille est réelle, on retrouve bien la définition initiale. Même remarque dans le cas des familles finies ou presque nulles. On observe désormais si les résultats du paragraphe précédent restent vrais.

Proposition 24.6. *Toute sous-famille d'une famille sommable reste sommable, mais on ne peut rien dire sur sa somme. Attention, on ne peut rien dire d'une famille majorée, à moins de majorer la famille $(|a_i|)_{i\in I}$.*

Théorème 24.13 (Inégalité triangulaire, cas complexe)**.** *Si la famille complexe $(a_i)_{i\in I}$ est sommable, alors*

$$\left|\sum_{i\in I} a_i\right| \leqslant \sum_{i\in I} |a_i|$$

Théorème 24.14 (Cas des séries, cas complexe)**.** *Soit $(a_i)_{i\in\mathbb{N}}$ une famille complexe. La famille $(a_i)_{i\in\mathbb{N}}$ est sommable si, et seulement si, la série $\sum a_n$ est absolument convergente, et dans ce cas on a :*

$$\sum_{i\in\mathbb{N}} a_i = \sum_{i=0}^{+\infty} a_i$$

Théorème 24.15 (Changement d'indice, cas complexe)**.** *Soit $(a_i)_{i\in I}$ une famille complexe et $\sigma : J \to I$ une bijection. Alors la famille $(a_i)_{i\in I}$ est sommable si, et seulement si, la famille $(a_{\sigma(j)})_{j\in J}$, et dans ce cas on a :*

$$\sum_{i\in I} a_i = \sum_{j\in J} a_{\sigma(j)}$$

Théorème 24.16 (Linéarité, cas complexe)**.** *Soit $\lambda, \mu \in \mathbb{C}$. Si $(a_i)_{i\in I}$ et $(b_i)_{i\in I}$ complexes sont sommables, alors $(\lambda a_i + \mu b_i)_{i\in I}$ est sommable et on a :*

$$\sum_{i\in I}(\lambda a_i + \mu b_i) = \lambda \sum_{i\in I} a_i + \mu \sum_{i\in I} b_i$$

Théorème 24.17 (Relation de Chasles, cas complexe)**.** *Si I_1 et I_2 sont deux ensembles disjoints tels que $I = I_1 \sqcup I_2$, alors $(a_i)_{i\in I}$ complexe est sommable si, et seulement si, $(a_i)_{i\in I_1}$ et $(a_i)_{i\in I_2}$ sont sommables, et dans ce cas on a :*

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

Par une récurrence immédiate, on généralise à un nombre fini d'ensembles deux à deux disjoints.

Théorème 24.18 (Sommation par paquets, cas complexe)**.** *Soit $(a_i)_{i\in I}$ une famille complexe. Soit $(I_t)_{t\in T}$ une famille de parties de I deux à deux disjointes telle que $I = \bigsqcup_{t\in T} I_t$. **Si** $(a_i)_{i\in I}$ est sommable, **alors** chaque $(a_i)_{i\in I_t}$ est sommable et la famille $\left(\sum_{i\in I_t} a_i\right)_{t\in T}$ est sommable et on a :*

$$\sum_{i\in I} a_i = \sum_{t\in T}\left(\sum_{i\in I_t} a_i\right)$$

Corollaire 24.4 (Théorème de Fubini, cas complexe)**.** *Soit $(a_{i,j})_{(i,j)\in I\times J}$ une famille complexe sommable. On a :*

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

4 Produit de familles

Théorème 24.19 (Produit de familles, cas positif)**.** *Soit $(a_i)_{i\in I}$ et $(b_j)_{j\in J}$ deux familles à termes dans $[0, +\infty]$. Alors*

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

Théorème 24.20 (Produit de familles, cas complexe)**.** *Soit $(a_i)_{i\in I}$ et $(b_j)_{j\in J}$ deux familles complexes sommables. Alors la famille $(a_i b_j)_{(i,j)\in I\times J}$ est sommable et on a :*

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

Remarque 24.9. Par une récurrence immédiate, on généralise le résultat à un produit fini de sommes.

Corollaire 24.5 (Produit de Cauchy de deux séries absolument convergentes)**.** *Soit* $\sum u_n$ *et* $\sum v_n$ *deux séries à termes complexes absolument convergentes. On pose*

$$\forall n \in \mathbb{N},\ w_n = \sum_{k=0}^{n} u_k v_{n-k}$$

Alors, la série $\sum w_n$ *est absolument convergente et on a :*

$$\sum_{n=0}^{+\infty} w_n = \left(\sum_{n=0}^{+\infty} u_n\right)\left(\sum_{n=0}^{+\infty} v_n\right)$$

Corollaire 24.6.

$$\forall (z, z') \in \mathbb{C}^2,\ \exp(z + z') = \exp(z)\exp(z')$$

Démonstration. Utiliser le produit de Cauchy et faire apparaître un binôme de Newton en multipliant les w_n par $\dfrac{n!}{n!}$. □

Chapitre 25

Dénombrements

1 Généralités

Définition 25.1 (Cardinal d'un ensemble fini)**.** Soit E un ensemble. S'il existe une bijection de E sur un certain $[\![1, n]\!]$, alors ce n est unique. On dit que alors que E est **fini** et son **cardinal** est alors défini par :

$$\mathrm{Card}(E) = n$$

On rencontre aussi souvent la notation $|E|$, voire plus rarement la notation $\#E$.

Exemple 25.1. Pour tout ensemble fini E et pour tout $\lambda \in \mathbb{C}$ (marche plus généralement dans une structure algébrique comportant un + commutatif), on a

$$\sum_{x \in E} \lambda = \mathrm{Card}(E) \times \lambda$$

Théorème 25.1 (Principe des tiroirs, cas général)**.** *Soit E et F des ensembles finis tels que $\mathrm{Card}(E) > \mathrm{Card}(F)$. Alors il n'existe pas d'injection de E dans F.*

Proposition 25.1. *Si E est fini et si E et F peuvent être mis en bijection, alors F est fini de même cardinal que E. Réciproquement, si E et F sont de même cardinal, alors ils peuvent être mis en bijection.*

Théorème 25.2. *Soit E un ensemble fini. Toute partie A de E est finie, et son cardinal vérifie :*

$$\mathrm{Card}(A) \leqslant \mathrm{Card}(E)$$

De plus, on égalité si, et seulement si, $A = E$.

Corollaire 25.1. *Soit E et F finis de même cardinal ainsi que $f : E \to F$. Alors f est bijective si, et seulement si, f est injective si, et seulement si, f est surjective.*

Proposition 25.2 (Une formule utile)**.** *Soit E un ensemble fini et $A \subset E$. Alors*

$$\mathrm{Card}(A) = \sum_{x \in E} \mathbb{1}_A(x)$$

Remarque 25.1. Avec le formalisme des sommes presque nulles, la formule reste vraie si A est une partie fini d'un ensemble quelconque.

Lemme 25.1 (Simple additivité du cardinal)**.** *Soit A et B deux parties finies disjointes d'un ensemble E. Alors $A \cup B$ est finie de cardinal*

$$\mathrm{Card}(A \cup B) = \mathrm{Card}(A) + \mathrm{Card}(B)$$

Exemple 25.2. Si E est fini, $A \subset$ et $\overline{A}$ désigne le complémentaire de A dans E, alors A et $\overline{A}$ sont finis et on a

$$\mathrm{Card}(\overline{A}) = \mathrm{Card}(E) - \mathrm{Card}(A)$$

En pratique, on utilisera souvent cette formule car il est parfois beaucoup plus simple de dénombrer le complémentaire d'un ensemble.

Corollaire 25.2 (Généralisation)**.** *Soit $\mathcal{A} \subset \mathcal{P}(E)$ un ensemble fini de parties finies de deux à deux disjointes. Alors $\bigsqcup_{A\in\mathcal{A}} A$ est fini et on a*

$$\mathrm{Card}\left(\bigsqcup_{A\in\mathcal{A}} A\right) = \sum_{A\in\mathcal{A}} \mathrm{Card}(A)$$

Corollaire 25.3 (Lemme des bergers)**.** *Soit $f : E \to F$ avec F fini. Si chacun des n éléments de F admet exactement k antécédents, alors E est fini de cardinal*

$$\mathrm{Card}(E) = n \times k$$

Théorème 25.3 (Cardinal d'une union)**.** *Soit A et B deux parties finies d'un ensemble E. Alors $A \cup B$ est finie et on a*

$$\mathrm{Card}(A \cup B) = \mathrm{Card}(A) + \mathrm{Card}(B) - \mathrm{Card}(A \cap B)$$

Proposition 25.3. *L'image directe d'un ensemble fini est finie.*

Proposition 25.4 (Inégalité de Boole pour les cardinaux)**.** *Soit $\mathcal{A} \subset \mathcal{P}(E)$ un ensemble fini de parties finies. Alors $\bigcup_{A\in\mathcal{A}} A$ est finie et on a :*

$$\mathrm{Card}\left(\bigcup_{A\in\mathcal{A}} A\right) \leqslant \sum_{A\in\mathcal{A}} \mathrm{Card}(A)$$

Théorème 25.4 (Formule du crible de Poincaré / principe d'inclusion-exclusion, HP)**.** *Soit $A_1, \ldots, A_n$ des parties finies d'un ensemble E. Alors $\bigcup_{i=1}^{n} A_i$ est finie et on a la formule du **crible de Poincaré** :*

$$\mathrm{Card}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{k=1}^{n} (-1)^{k-1} \sum_{1\leqslant i_1<\ldots<i_k\leqslant n} \mathrm{Card}\left(\bigcap_{j=1}^{k} A_{i_j}\right)$$

Démonstration. Dénombrer le complémentaire en utilisant la distributivité généralisée sur les fonctions indicatrices des A_i □

Théorème 25.5 (Cardinal d'un produit cartésien)**.** *Si E et F sont des ensembles finis, alors l'ensemble $E \times F$ est fini de cardinal*

$$\mathrm{Card}(E \times F) = \mathrm{Card}(E) \times \mathrm{Card}(F)$$

Corollaire 25.4. *Si $E_1, \ldots, E_n$ sont des ensembles finis, alors $E_1 \times \ldots \times E_n$ est fini de cardinal*

$$\mathrm{Card}(E_1 \times \ldots \times E_n) = \mathrm{Card}(E_1) \times \ldots \times \mathrm{Card}(E_n)$$

Théorème 25.6 (Cardinal de l'ensemble des applications d'un ensemble fini vers un autre)**.** *Si E et F sont finis, alors F^E est fini de cardinal*

$$\mathrm{Card}\left(F^E\right) = \mathrm{Card}(F)^{\mathrm{Card}(E)}$$

Remarque 25.2. C'est ce résultat qui justifie a posteriori la notation F^E.

Théorème 25.7. *Si E est un ensemble fini, alors $\mathcal{P}(E)$ est fini de cardinal*

$$\mathrm{Card}\left(\mathcal{P}(E)\right) = 2^{\mathrm{Card}(E)}$$

2 Dénombrabilité, HP, au programme de spé

Définition 25.2 (Dénombrabilité)**.** Un ensemble E est dit **dénombrable** s'il peut être mis en bijection avec $\mathbb{N}$. Il est dit **au plus dénombrable** s'il est fini ou dénombrable. il est dit **indénombrable** s'il n'est pas au plus dénombrable.

Proposition 25.5. *E est au plus dénombrable si, et seulement si, il existe une injection de E dans $\mathbb{N}$.*

Corollaire 25.5. *Toute partie d'un ensemble au plus dénombrable est au plus dénombrable.*

Lemme 25.2. *Soit $\varphi : E \to F$. Si φ est injective et F est au plus dénombrable, alors E est au plus dénombrable. Si φ est surjective et E est au plus dénombrable, alors F est au plus dénombrable.*

Théorème 25.8. *Si $E_1, \ldots, E_n$ sont au plus dénombrables, alors $E_1 \times \ldots \times E_n$ est au plus dénombrable.*

Corollaire 25.6. *$\mathbb{Z}$ et $\mathbb{Q}$ sont dénombrables.*

Théorème 25.9. *Une union au plus dénombrable d'ensemble au plus dénombrables est au plus dénombrable.*

Proposition 25.6. *$\mathcal{P}(\mathbb{N})$ est indénombrable.*

Démonstration. Il s'agit d'un cas particulier du théorème de Cantor qui stipule que pour tout ensemble E, il n'existe pas de surjection de E dans $\mathcal{P}(E)$ (et donc *a fortiori* pas de bijection). Pour le démontrer, raisonner par l'absurde en considérant une éventuelle surjection puis en s'intéressant à l'ensemble des éléments de E qui n'appartiennent pas à leur image par f. Par hypothèse, cette partie de E admet un antécédent qui ne peut ni appartenir à son image, ni ne pas lui appartenir, ce qui est absurde. □

Proposition 25.7. $\mathbb{R}$ *est indénombrable.*

Démonstration. C'est l'argument diagonal de Cantor. Supposer que l'on puisse énumérer tous les réels, puis conclure à une absurdité en prenant le réel formé par des chiffres différents de 9 et différents des chiffres qui apparaissent sur la diagonale. Ce réel n'apparaît alors pas dans l'énumération, ce qui est absurde par existence et unicité du développement propre. □

3 Choisir p objets parmi n

3.1 Avec ordre et avec répétition

On dénombre des p-uplets.

Théorème 25.10. *Soit F de cardinal n. Le nombre de p-uplets à valeurs dans F est n^p.*

3.2 Avec ordre et sans répétition

On dénombre des p-uplets injectifs.

Théorème 25.11 (Nombre d'arrangements)**.** *Soit F de cardinal $n \geqslant p$. Alors le nombre de p-uplets à valeurs distinctes dans F vaut*

$$A_n^p = n(n-1)\dots(n-p+1) = \frac{n!}{(n-p)!}$$

*A_n^p s'appelle un nombre d'**arrangements**.*

Corollaire 25.7 (Nombre d'injections)**.** *Soit E de cardinal p et F de cardinal $n \geqslant p$. Alors le nombre d'injections de E dans F vaut A_n^p.*

Corollaire 25.8 (Nombre de permutations)**.** *Si E est fini de cardinal n alors il existe $n!$ permutations de E.*

3.3 Sans ordre et sans répétition

On dénombre des parties à p éléments.

Théorème 25.12 (Nombre de combinaisons)**.** *Soit F de cardinal n. Le nombre de parties de F de cardinal p vaut*

$$C_n^p = \binom{n}{p} = \frac{n(n-1)\dots(n-p+1)}{p!} = \frac{n!}{p!(n-p)!} = \frac{A_n^p}{p!}$$

*C'est pour cela que les coefficients binomiaux s'appellent aussi des nombres de **combinaisons**.*

Corollaire 25.9 (Nombre d'applications strictement croissantes)**.** *Soit E de cardinal p et F de cardinal $n \geqslant p$. Alors le nombre d'applications strictement croissantes de E dans F vaut $\binom{n}{k}$.*

3.4 Sans ordre et avec répétition, HP, exo classique

Théorème 25.13. *Le nombre de façons de choisir p objets parmi n sans ordre et avec répétition est égal à* $\binom{n+p-1}{p}$

Démonstration. On se ramène à $F = [\![1, n]\!]$ pour plus de clarté. Si on choisit les éléments deux à deux distincts, on se ramène au cas précédent. Si les répétitions sont permises, tout revient à choisir un p-uplet croissant au sens large. On prend alors un p-uplet croissant $(x_1 \leqslant \ldots, x_p)$ et on l'envoie sur le p-uplet $(x_1, x_2 + 1, \ldots, x_p + (p-1))$ qui est alors strictement croissant. On vérifie que cet "envoi" est bijectif vers les p-uplets à valeurs dans $[\![1, n+p-1]\!]$ puis on conclut alors par le premier sous-cas. □

Corollaire 25.10 (Nombre d'applications croissantes)**.** *Le nombre d'applications croissantes d'un ensemble à p éléments dans un ensemble à $n \leqslant p$ éléments vaut* $\binom{n+p-1}{p}$.

Remarque 25.3. On peut retenir la méthode du "stars and bars" : lorsqu'on choisit p objets parmi des objets de n types sans ordre avec répétition, cela revient à choisir les $n-1$ cloisons entre les multiples choix des différents éléments, donc on a bien $\binom{n+p-1}{p}$ choix possibles.

Chapitre 26

Probabilités

1 Probabilités sur univers fini

Définition 26.1 (Univers, issue, événement, événement élémentaire)**.** Soit Ω un ensemble non vide, appelé **univers**. Chaque élément $\omega \in \Omega$ correspond à une **issue** possible de l'expérience. Formellement, un **événement** est une partie $A \subset \Omega$. C'est donc un ensemble d'issues possibles. On dit que A est **réalisé** lorsque l'issue de l'expérience est un $\omega \in A$. Dans toute la suite, nous nous restreindrons au cas où Ω est **fini** (et toujours non vide). Si $A = \{\omega\}$ est un singleton, on parle d'**événement élémentaire**.

Remarque 26.1. Même si formellement un événement A est défini comme une partie de Ω, la plupart du temps on décrira A par une proposition logique $p(\omega)$ qui dépend de $\omega \in \Omega$. Dans ce cas, A est à comprendre comme l'ensemble $\{\omega \in \Omega | p(\omega)\}$

Définition 26.2 (Événements "ou", "et" et contraire)**.** Soit A et B deux événements, décrits respectivement par des propositions logiques p et q. L'événement A **ou** B désigne l'événement $A \cup B$. Il est décrit par $p \vee q$. L'événement A **et** B désigne l'événement $A \cap B$. Il est décrit par $p \wedge q$. L'événement **contraire** de A est l'événement $\overline{A} = \Omega \setminus A$. Il est décrit par $\neg p$.

Définition 26.3 (Événement impossible, événements incompatibles, système complet d'événements)**.** L'événement **impossible** est l'événement $\emptyset$. Il doit son nom au fait qu'il ne peut jamais être réalisé. Si $A \cap B = \emptyset$, on dit que A et B sont **incompatibles**. Enfin, si $\Omega = \bigsqcup_{1 \leqslant i \leqslant n} A_i$, on dit que les A_i forment un **système complet d'événements**.

Exemple 26.1. Les événements élémentaires forment un système complet d'événements.

Exemple 26.2. Si A est un événement, alors $(A, \overline{A})$ est un système complet d'événements.

Définition 26.4 (Probabilité, espace probabilisé)**.** Une **probabilité** sur Ω est une application

$$\begin{array}{rccc} \mathbb{P} : & \mathcal{P}(\Omega) & \longrightarrow & [0,1] \\ & A & \longmapsto & \mathbb{P}(A) \end{array}$$

telle que $\mathbb{P}(\Omega) = 1$ et que pour toutes parties disjointes A et B de Ω, on ait

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$$

On dit que $(\Omega, \mathbb{P})$ est un **espace probabilisé**.

Proposition 26.1. *Puisque Ω et $\emptyset$ sont disjoints, on a immédiatement $\mathbb{P}(\emptyset) = 0$.*

Proposition 26.2. *Si $A_1, \dots, A_n$ sont deux à deux disjoints, alors on a*

$$\mathbb{P}\left(\bigsqcup_{1 \leqslant i \leqslant n} A_i\right) = \sum_{1 \leqslant i \leqslant n} \mathbb{P}(A_i)$$

Proposition 26.3 (Formule des probabilités totales, V1)**.** *Si les A_i forment un système complet d'événements et si $B \subset \Omega$, on a*

$$\mathbb{P}(B) = \sum_{i=1}^{n} \mathbb{P}(B \cap A_i)$$

Théorème 26.1 (Caractérisation par les probabilités élémentaires)**.** *Soit un univers fini $\Omega = \{\omega_1, \dots, \omega_n\}$ et $p_1, \dots, p_n$ des réels positifs dont la somme vaut 1. Alors il existe une unique probabilité sur Ω telle que*

$$\forall i \in [\![1, n]\!],\ \mathbb{P}(\{\omega_i\}) = p_i$$

Corollaire 26.1. *Si deux probabilités coïncident sur les événements élémentaires, alors coïncident partout.*

Définition 26.5 (Probabilité uniforme)**.** La **probabilité uniforme** sur Ω est l'unique probabilité qui vaut $\dfrac{1}{\mathrm{Card}(\Omega)}$ sur chaque événement élémentaire. Elle vérifie donc

$$\forall A \subset \Omega,\ \mathbb{P}(A) = \frac{\mathrm{Card}(A)}{\mathrm{Card}(\Omega)}$$

Théorème 26.2. *Soit $(\omega, \mathbb{P})$ un espace probabilisé. On a les propriétés suivantes :*

1. $\forall (A, B) \in \mathcal{P}(\Omega)^2,\ A \subset B \Longrightarrow \mathbb{P}(A) \leqslant \mathbb{P}(B)$ *(croissance)*
2. $\forall A \in \mathcal{P}(\Omega),\ \mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$
3. $\forall (A, B) \in \mathcal{P}(\Omega)^2,\ \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

Proposition 26.4 (Inégalité de Boole)**.** *Soit $A_1, \dots, A_n$ des événements. On a*

$$\mathbb{P}\left(\bigcup_{1 \leqslant i \leqslant n} A_i\right) \leqslant \sum_{1 \leqslant i \leqslant n} \mathbb{P}(A_i)$$

Définition 26.6 (Événement presque sûr, HP)**.** Un événement est dit **presque sûr** lorsqu'il est de probabilité 1. L'abréviation **p.s.** est standard.

Proposition 26.5 (HP)**.** *Une intersection finie d'événements presque sûrs est presque sûre.*

Démonstration. Procéder par récurrence. On a simplement besoin du fait que si A et B sont presque sûrs, alors $A \cap B$ est presque sûr. Pour montrer cela, il suffit d'utiliser le fait que $\mathbb{P}(A \cap B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cup B)$. Or, par croissance de la probabilité, on a nécessairement $\mathbb{P}(A \cup B) = 1$, ce qui achève la preuve. □

Définition 26.7 (Probabilité conditionnelle)**.** Soit B un événement tel que $\mathbb{P}(B) > 0$. pour tout événement A, la **probabilité conditionnelle de A sachant B** est définie par

$$\mathbb{P}(A|B) = \mathbb{P}_B(A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

Proposition 26.6. *$\mathbb{P}(\cdot|B) = \mathbb{P}_B$ est une probabilité sur Ω.*

Remarque 26.2. L'événement B (et tout sur-événement de B) est $\mathbb{P}_B$-presque sûr.

Théorème 26.3 (Formule des probabilités composées)**.** *Soit $A_1, \ldots, A_n$ des événements tels que $\mathbb{P}(A_1 \cap \ldots \cap A_{n-1}) > 0$. Alors*

$$\mathbb{P}(A_1 \cap \ldots \cap A_n) = \mathbb{P}(A_1) \times \mathbb{P}(A_2|A_1) \times \ldots \times \mathbb{P}(A_n|A_1 \cap \ldots \cap A_{n-1})$$

Théorème 26.4 (Formule des probabilités totales, V2)**.** *Soit $(A_i)_{1 \leqslant i \leqslant n}$ un système complet d'événements tel que $\forall i \in [\![1, n]\!],\ \mathbb{P}(A_i) > 0$. Soit B un événement. On a*

$$\mathbb{P}(B) = \sum_{i=1}^{n} \mathbb{P}_{A_i}(B)\mathbb{P}(A_i)$$

Remarque 26.3. Il existe une convention qui stipule que $\mathbb{P}_A(B)\mathbb{P}(A) = 0$ lorsque $\mathbb{P}(A) = 0$. Dans ce cas, la formule des probabilités totales V2 reste vraie en toute généralité.

Théorème 26.5 (Première formule de Bayes)**.** *Soit A et B deux événements de probabilités non nulles. Alors*

$$\mathbb{P}_B(A) = \frac{\mathbb{P}_A(B)\mathbb{P}(A)}{\mathbb{P}(B)}$$

Théorème 26.6 (Seconde formule de Bayes)**.** *Soit $(A_i)_{1 \leqslant i \leqslant n}$ un système complet d'événements tel que $\forall i \in [\![1, n]\!],\ \mathbb{P}(A_i) > 0$. Soit B un événement de probabilité non nulle. On a*

$$\forall i \in [\![1, n]\!],\ \mathbb{P}_B(A_i) = \frac{\mathbb{P}_{A_i}(B)\mathbb{P}(A_i)}{\displaystyle\sum_{j=1}^{n} \mathbb{P}_{A_j}(B)\mathbb{P}(A_j)}$$

Définition 26.8 (Événements indépendants)**.** On dit que les événements A et B sont **indépendants** lorsque

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$$

Définition 26.9 (Événements mutuellement indépendants)**.** Soit $(A_i)_{1 \leqslant i \leqslant n}$ une famille finie d'événements. Les A_i sont dits **mutuellement indépendants** (ou plus simplement, **indépendants**) lorsque

$$\forall I \subset [\![1, n]\!],\ \mathbb{P}\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} \mathbb{P}(A_i)$$

On rappelle qu'on a par convention

$$\bigcap_{i \in \emptyset} A_i = \Omega$$

Remarque 26.4. Si $I = \emptyset$ ou I est un singleton, il n'y a rien à vérifier. En pratique, il y a donc $2^n - n - 1$ vérifications à effectuer.

Proposition 26.7. *Soit $A_1, \dots, A_n$ des événements mutuellement indépendants. Alors pour tout $\sigma \in \mathcal{S}_n$, les événements $A_{\sigma(i)}$ sont mutuellement indépendants. De plus, pour toute sous-famille $(A_{i_1}, \dots, A_{i_r})$, les événements restent indépendants.*

Théorème 26.7. *Soit $A_1, \dots, A_n$ des événements mutuellement indépendants. Soit $B_1, \dots, B_n$ des événements que*

$$\forall i \in [\![1, n]\!],\ B_i \in \left\{A_i, \overline{A_i}\right\}$$

Alors, les événements $B_1, \dots, B_n$ sont mutuellement indépendants.

Théorème 26.8 (Théorème de regroupement)**.** *Soit $A_1, \dots, A_n$ des événements mutuellement indépendants. et soit $p \in [\![1, n-1]\!]$. Alors*

1. $\bigcap_{1 \leqslant i \leqslant p} A_i$ *et* $\bigcap_{p < i \leqslant n} A_i$ *sont indépendants.*
2. $\bigcap_{1 \leqslant i \leqslant p} A_i$ *et* $\bigcup_{p < i \leqslant n} A_i$ *sont indépendants.*
3. $\bigcup_{1 \leqslant i \leqslant p} A_i$ *et* $\bigcap_{p < i \leqslant n} A_i$ *sont indépendants.*
4. $\bigcup_{1 \leqslant i \leqslant p} A_i$ *et* $\bigcup_{p < i \leqslant n} A_i$ *sont indépendants.*

2 Variables aléatoires

Définition 26.10 (Variable aléatoire)**.** Soit $(\Omega, \mathbb{P})$ un espace probabilisé fini. Une **variable aléatoire** définie sur Ω est une application

$$\begin{array}{rccc} X : & \Omega & \longrightarrow & E \\ & \omega & \longmapsto & X(\omega) \end{array}$$

En particulier, si $E \subset \mathbb{R}$, la variable aléatoire est dite **réelle**.

Définition 26.11. Soit A un événement. On définit :

$$\{X \in A\} = \{\omega \in \Omega \mid X(\omega) \in A\} = X^{-1}(A)$$

On pourra par exemple considérer des événements du type $\{X = x\}$ où $x \in E$ est fixé. Si X est réelle, on pourra considérer des événements du type $\{X \leqslant x\}$. La notation $\mathbb{P}(\{X \in A\})$ étant lourde, on lui préfèrera la notation suivante, légèrement abusive :

$$\mathbb{P}(X \in A)$$

Exemple 26.3. Lorsque x parcourt $X(\Omega)$, les événements $\{X = x\}$ forment un SCE.

Définition 26.12 (Loi d'une variable aléatoire)**.** La **loi de** X est définie sur $\mathcal{P}(E)$ par

$$\forall A \subset E,\ \mathbb{P}_X(A) = \mathbb{P}(X \in A) = \mathbb{P}(X^{-1}(A))$$

Elle induit une loi de probabilité sur $X(\Omega)$ (définie sur $\mathcal{P}(X(\Omega))$ donc).

Définition 26.13 (Même loi)**.** Soit $(\Omega, \mathbb{P}')$ un espace probabilisé et $Y : \Omega' \to E$ une variable aléatoire. On dit que X et Y **ont même loi** lorsque

$$\forall A \subset E,\ \mathbb{P}(X \in A) = \mathbb{P}'(Y \in A)$$

On note alors

$$X \sim Y$$

Notons que le seul pré-requis est que X et Y pointent vers un même ensemble E, les espaces probabilisés peuvent être les mêmes comme différents.

Proposition 26.8. *Si $A \subset E$, on a*

$$\mathbb{P}(X \in A) = \sum_{x \in A} \mathbb{P}(X = x)$$

Proposition 26.9 (La loi de X est caractérisée par les probabilités élémentaires $\mathbb{P}(X = x)$)**.** *X et Y ont même loi si, et seulement si, elles ont les mêmes probabilités élémentaires sur E. On dit que "la loi de X est caractérisée par les probabilités élémentaires $\mathbb{P}(X = x)$".*

Remarque 26.5. Attention, ce n'est pas parce que X et Y ont même loi qu'elles sont égales, il suffit de considérer les variables aléatoires qui correspondent au résultat pile et au résultat face d'un pièce équilibrée.

Définition 26.14 (Variable aléatoire image)**.** Soit X à valeurs dans E et f une application de E dans F. Alors $f \circ X$ est une variable aléatoire à valeurs dans F, appelée **image** de X par f. On la note plus simplement $f(X)$ et elle vérifie

$$\forall B \subset F,\ \mathbb{P}(f(X) \in B) = \mathbb{P}(X \in f^{-1}(B))$$

Autrement dit, on a

$$\mathbb{P}_{f(X)} = \mathbb{P}_X \circ f^{-1}$$

Proposition 26.10 (Transfert de loi)**.** *Si $X \sim Y$, alors $f(X) \sim f(Y)$.*

Définition 26.15 (Couple de variables aléatoires)**.** Soit $(\omega, \mathbb{P}$ un espace probabilisé et X et Y des variables aléatoires définies sur Ω à valeurs respectivement dans E et F. le **couple de variables aléatoires** (X, Y) est la variable aléatoire à valeurs dans $E \times F$ définie par

$$\begin{array}{rccc} (X,Y): & \Omega & \longrightarrow & E \times F \\ & \omega & \longmapsto & (X(\omega), Y(\omega)) \end{array}$$

La notation $\{X = x, Y = y\}$ désignera en fait l'événement $\{X = x\} \cap \{Y = y\}$. De même, on peut imaginer des événements du type $\{X \leqslant x, Y \geqslant y\}$, etc.

Définition 26.16 (Loi conjointe, lois marginales)**.** La **loi conjointe** de X et Y est la loi de (X, Y). Réciproquement, les **lois marginales** de (X, Y) sont les lois de X et Y.

Exemple 26.4. Lorsque x parcourt $X(\Omega)$ et y parcourt $Y(\Omega)$, les événements $\{X = x, Y = y\}$ forment un SCE.

Remarque 26.6. En général, certains des événements $\{X = x, Y = y\}$ sont de probabilité nulle. En effet, si l'on note $Z = (X, Y)$, on a $Z(\Omega) \subset X(\Omega) \times Y(\Omega)$, mais l'inclusion réciproque n'a aucune raison d'être vraie.

Proposition 26.11. *On a*

$$\forall x \in E,\ \mathbb{P}(X = x) = \sum_{y \in F} \mathbb{P}(X = x, Y = y)$$

et

$$\forall y \in F,\ \mathbb{P}(Y = y) = \sum_{x \in E} \mathbb{P}(X = x, Y = y)$$

Corollaire 26.2 (La loi conjointe détermine les lois marginales)**.** *Si (X, Y) et (X', Y') ont les mêmes probabilités élémentaires, alors $X \sim X'$ et $Y \sim Y'$.*

Remarque 26.7. La réciproque est fausse. Il faut considérer un lancer de deux pièces équilibrées dont le résultat est recensé par (X, Y), et d'un autre côté un lancer truqué où la deuxième pièce suit toujours le résultat de la première (équilibrée), dont le résultat est recensé par (X', Y'). Les lois marginales sont les mêmes, et pourtant les lois conjointes ne sont pas les mêmes.

Définition 26.17 (Variables aléatoires indépendantes)**.** Deux variables aléatoires X et Y définies sur le même espace probabilisé sont dites **indépendantes** lorsque

$$\forall (x, y) \in X(\Omega) \times Y(\Omega),\ \mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y)$$

Cela est équivalent au fait de vérifier

$$\forall (x, y) \in E \times F,\ \mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y)$$

Dans ce cas, on note alors $X \perp\!\!\!\perp Y$

Proposition 26.12. *Soit X et Y deux variables aléatoires indépendantes. Alors pour tout $A \subset E$ et tout $B \subset F$, les événements $\{X \in A\}$ et $\{Y \in B\}$ sont indépendants.*

Définition 26.18 (Loi conditionnelle)**.** Pour tout $x \in E$ tel que $\mathbb{P}(X = x) > 0$, on définit la **loi conditionnelle de Y sachant** $\{X = x\}$ par

$$\forall B \subset F,\ \mathbb{P}_{\{X=x\}}(Y \in B) = \mathbb{P}(Y \in B \mid X = x)$$

On définirait de même la loi de X sachant $\{Y = y\}$.

Proposition 26.13 (HP)**.** *Les trois propositions suivantes sont équivalentes :*

1. *X et Y sont indépendantes*
2. *Pour tout $x \in E$ tel que $\mathbb{P}(X = x) > 0$, la loi de Y sachant $\{X = x\}$ vaut la loi de Y*
3. *Pour tout $y \in F$ tel que $\mathbb{P}(Y = y) > 0$, la loi de X sachant $\{Y = y\}$ vaut la loi de X*

Démonstration. Il suffit de montrer l'équivalence entre 1. et 2. car celle entre 1. et 3. est symétrique. Pour le sens direct, utiliser la définition de le fait que les événements $\{X \in A\}$ et $\{Y \in B\}$ sont indépendants puis revenir à la définition de la loi conditionnelle. Pour le sens réciproque, utiliser la définition de la loi conditionnelle. Pour les événements de probabilité nulle, la probabilité d'un événement "et" comportant l'un de ces événements est aussi nulle par croissance de la probabilité. □

Proposition 26.14 (Transfert d'indépendance)**.** *Soit X et Y deux variables aléatoires indépendantes à valeurs respectivement dans E et F. Alors pour toutes fonctions $f : E \to E'$ et $g : F \to F'$, les variables aléatoires $f(X)$ et $g(Y)$ sont indépendantes.*

Définition 26.19 (Généralisation : n-uplet de variables aléatoires)**.** On suppose que $n \geqslant 2$. Si $X_1, \ldots, X_n$ sont des variables aléatoires définies sur un même espace probabilisé à valeurs respectivement dans $E, \ldots, E_n$, alors le n-uplet de variables aléatoires $(X_1, \ldots, X_n)$ est une variable aléatoire à valeurs dans $E_1 \times \ldots \times E_n$.
On définit de même la loi conjointe des X_i et les lois marginales de $(X_1, \ldots, X_n)$. Les lois marginales continuent à être déterminées par la la loi conjointe. On définit aussi de même la loi de X_i sachant $\{X_j = x_j\}$.

Définition 26.20 (Variables aléatoires mutuellement indépendantes)**.** Les variables aléatoires $X_1, \ldots, X_n$ sont dites **mutuellement indépendantes** (ou plus simplement **indépendantes**) lorsque

$$\forall (x_1, \ldots, x_n) \in X_1(\Omega) \times \ldots \times X_n(\Omega),\ \mathbb{P}(X_1 = x_1, \ldots, X_n = x_n) = \mathbb{P}(X_1 = x_1) \times \ldots \times \mathbb{P}(X_n = x_n)$$

Cela est équivalent au fait de vérifier

$$\forall (x_1, \ldots, x_n) \in E_1 \times \ldots \times E_n,\ \mathbb{P}(X_1 = x_1, \ldots, X_n = x_n) = \mathbb{P}(X_1 = x_1) \times \ldots \times \mathbb{P}(X_n = x_n)$$

Proposition 26.15. *Soit $X_1, \ldots, X_n$ des variables aléatoires mutuellement indépendantes. Alors pour tous $A_1 \subset E_1, \ldots, A_n \subset E_n$, les événements $\{X_1 \in A_1\}, \ldots, \{X_n \in A_n\}$ sont mutuellement indépendants.*

Proposition 26.16. *Soit $X_1, \ldots, X_n$ des variables aléatoires mutuellement indépendantes. Alors pour toute permutation $\sigma \in \mathcal{S}_n$, les variables aléatoires $X_{\sigma(1)}, \ldots, X_{\sigma(n)}$ sont mutuellement indépendantes. Pour toute sous-famille $(X_{i_1}, \ldots, X_{i_r})$, les variables aléatoires restent mutuellement indépendantes. Les variables aléatoires sont deux à deux indépendantes, mais la réciproque est fausse dès que $n \geqslant 3$.*

Proposition 26.17. *Si $X_1, \ldots, X_n$ sont mutuellement indépendantes, alors $f_1(X_1), \ldots, f_n(X_n)$ sont mutuellement indépendantes.*

Théorème 26.9 (Lemme des coalitions)**.** *Soit $X_1, \ldots, X_n$ des variables aléatoires mutuellement indépendantes et $p \in [\![1, n-1]\!]$ Soit $f : E_1 \times \ldots \times E_p \to F$ et $g : E_{P+1} \times \ldots \times E_n \to G$. Alors les variables aléatoires $f(X_1, \ldots, X_p)$ et $g(X_{P+1}, \ldots, X_n)$ sont indépendantes. Le résultat s'étend de manière naturelle à plus de deux coalitions.*

2.1 Lois usuelles

Définition 26.21. Soit $(\Omega, \mathbb{P})$ et $(E, \mathbb{Q})$ des espaces probabilisés et $X : \Omega \to E$ une variable aléatoire. De manière logique, on dit que X **suit** la loi $\mathbb{Q}$ lorsque $\mathbb{P}_X = \mathbb{Q}$:

$$\forall A \in E,\ \mathbb{P}_X(A) = \mathbb{Q}(A)$$

Dans ce cas, on note $X \sim \mathbb{Q}$.

Remarque 26.8. Soit $X, Y : \Omega \to E$. Supposons que $X \sim \mathbb{Q}$. Alors $Y \sim \mathbb{Q}$ si, et seulement si, $X \sim Y$. Supposons que $X \sim Y$. Alors $X \sim \mathbb{Q}$ si, et seulement si, $Y \sim \mathbb{Q}$.

Proposition 26.18. *$X \sim \mathbb{Q}$ si, et seulement si, $\forall x \in E,\ \mathbb{P}(X = x) = \mathbb{Q}(\{x\})$*

Définition 26.22 (Loi uniforme)**.** La **loi uniforme** est l'unique probabilité définie sur $E = [\![1, n]\!]$ caractérisée par

$$\forall i \in [\![1, n]\!],\ \mathbb{Q}(\{i\}) = \frac{1}{n}$$

Ainsi, on aura $X \sim \mathcal{U}([\![1, n]\!])$ lorsque $X : \Omega \to [\![1, n]\!]$ et

$$\forall i \in [\![1, n]\!],\ \mathbb{P}(X = i) = \frac{1}{n}$$

Définition 26.23 (Loi de Bernoulli de paramètre p)**.** Soit $p \in [0, 1]$. La **loi de Bernoulli de paramètre** p est l'unique probabilité définie sur $E = \{0, 1\}$ caractérisée par

$$\mathbb{Q}(\{1\}) = p \ \text{ et } \ \mathbb{Q}(\{0\}) = 1 - p$$

Ainsi, on aura $X \sim \mathcal{B}(p)$ lorsque $X : \Omega \to \{0, 1\}$, $\mathbb{P}(X = 1) = p$ et $\mathbb{P}(X = 0) = 1 - p$.

Exemple 26.5. Soit $(\Omega, \mathbb{P})$ un espace probabilisé et A un événement. Alors, l'indicatrice $\mathbb{1}_A$ suit la loi $\mathcal{B}(\mathbb{P}(A))$.

Définition 26.24 (Loi binomiale de paramètres n et p)**.** Soit $n \in \mathbb{N}^\star$ et $p \in [0, 1]$. La **loi binomiale de paramètres** n **et** p est l'unique probabilité sur $E = [\![0, n]\!]$ caractérisée par

$$\forall k \in [\![0, n]\!],\ \mathbb{Q}(\{k\}) = \binom{n}{k} p^k (1 - p)^{n-k}$$

Ainsi, on aura $X \sim \mathcal{B}(n, p)$ lorsque $X : \Omega \to [\![0, n]\!]$ et

$$\forall k \in [\![0, n]\!],\ \mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

Théorème 26.10 (Somme de n Bernoulli indépendantes de même paramètre)**.** *Soit $X_1, \ldots, X_n$ des variables aléatoires indépendantes de loi $\mathcal{B}(p)$. Alors*

$$X_1 + \ldots + X_n \sim \mathcal{B}(n, p)$$

2.2 Simulation de variables aléatoires

Théorème 26.11 (Simulation de variables aléatoires)**.** *Soit $(E_i, \mathbb{Q}_i)_{1 \leqslant i \leqslant n}$ des espaces probabilisés. Alors il existe un espace probabilisé $(\Omega, \mathbb{P})$ et des variables aléatoires $X_i : \Omega \to E_i$ tels que*

1. *les X_i soient mutuellement indépendantes*
2. *$\forall i \in [\![1, n]\!],\ X_i \sim \mathbb{Q}_i$*

Définition 26.25 (Variables aléatoires indépendantes et identiquement distribuées)**.** Lorsqu'on a n copies identiques d'une même variable aléatoire qui sont mutuellement indépendantes, on dit que ces variables aléatoires sont **indépendantes et identiquement distribuées**. L'abréviation **iid** est standard. Le théorème précédent nous permet d'avoir autant de variables iid que l'on souhaite qui suivent les lois de notre choix.

3 Espérance et variance

Dans cette partie, sauf mention expresse du contraire, on considère des variables aléatoires réelles.

3.1 Espérance

Définition 26.26 (Espérance)**.** Soit X une variable aléatoire réelle. L'**espérance de** X est définie par

$$\mathbb{E}[X] = \sum_{x \in X(\Omega)} \mathbb{P}(X = x) \times x$$

Remarque 26.9. En vertu du formalisme des sommes presque nulles, pour tout B tel que $X(\Omega) \subset B \subset \mathbb{R}$, on a

$$\mathbb{E}[X] = \sum_{x \in B} \mathbb{P}(X = x) \times x$$

Définition 26.27 (Variable aléatoire centrée)**.** On dit que la variable aléatoire X est **centrée** lorsque $\mathbb{E}[X] = 0$.

Proposition 26.19 (Espérance d'une indicatrice)**.** *Pour tout événement A, on a*

$$\mathbb{E}[\mathbb{1}_A] = \mathbb{P}(A)$$

Cette formule peut s'avérer très utile lorsqu'on on interprète une variable aléatoire comme une fonction d'indicatrices, car l'espérance est linéaire (et multiplicative pour les variables aléatoires indépendantes).

Proposition 26.20. *On a*

$$\mathbb{E}[X] = \sum_{\omega \in \Omega} \mathbb{P}(\{\omega\}) \times X(\Omega)$$

Proposition 26.21 (Formule de transfert)**.** *Soit $X : \Omega \to E$ une variable aléatoire quelconque et $f : E \to \mathbb{R}$. On a*

$$\mathbb{E}[f(X)] = \sum_{x \in X(\Omega)} \mathbb{P}(X = x) f(x)$$

Remarque 26.10. De même, si $X(\Omega) \subset B \subset \mathbb{R}$, on a

$$\mathbb{E}[f(X)] = \sum_{x \in B} \mathbb{P}(X = x) f(x)$$

Corollaire 26.3. *Soit $X : \Omega \to E$ et $Y : \Omega \to E$ deux variables aléatoires quelconques. Si $X \sim Y$, alors pour toute fonction $f : E \to \mathbb{R}$, on a*

$$\mathbb{E}[f(X)] = \mathbb{E}[f(Y)]$$

Remarque 26.11. Si Y est définie sur un autre espace probabilisé $(\Omega', \mathbb{P}')$, le résultat reste vrai en écrivant $\mathbb{E}[f(X)] = \mathbb{E}'[f(Y)]$

Corollaire 26.4. *Deux variables aléatoires réelles de même loi ont même espérance.*

Théorème 26.12 (Propriétés intégrales de l'espérance)**.** *L'espérance vérifie les propriétés suivantes*

1. *Linéarité :* $\mathbb{E}[\lambda X + \mu Y] = \lambda\mathbb{E}[X] + \mu\mathbb{E}[Y]$
2. *Inégalité triangulaire* $|\mathbb{E}[X]| \leqslant \mathbb{E}[|X|]$
3. *Positivité : si* $X \geqslant 0$*, alors* $\mathbb{E}[X] \geqslant 0$
4. *Croissance : si* $X \leqslant Y$*, alors* $\mathbb{E}[X] \leqslant \mathbb{E}[Y]$

Remarque 26.12. Les deux derniers points restent vrais si l'inégalité n'a lieu que presque sûrement.

Proposition 26.22 (Quelques espérances à connaître)**.** *On a les résultat suivants :*

1. *Si* $X = x$*, alors* $\mathbb{E}[X] = x$
2. *Si* $X \sim \mathcal{U}(\llbracket 1, n \rrbracket)$*, alors* $\mathbb{E}[X] = \dfrac{n+1}{2}$
3. *Si* $X \sim \mathcal{B}(p)$*, alors* $\mathbb{E}[X] = p$
4. *Si* $X \sim \mathcal{B}(n,p)$ *alors* $\mathbb{E}[X] = np$

Remarque 26.13. Par conséquent, la variable aléatoire $X - \mathbb{E}[X]$ est toujours centrée.

Proposition 26.23 (Formule du crible de Poincaré pour les probabilités, HP)**.** *Soit* $(\Omega, \mathbb{P})$ *et* $A_1, \dots, A_n$ *des événements. On a*

$$\mathbb{P}\left(\bigcup_{1\leqslant i\leqslant n}\right) = \sum_{1\leqslant k\leqslant n}(-1)^{k-1}\sum_{1\leqslant i_1<\ldots<i_k\leqslant n}\mathbb{P}\left(\bigcap_{1\leqslant j\leqslant k} A_{i_j}\right)$$

Démonstration. On passe par le complémentaire. Développer $(1-\mathbb{1}_{A_1})\dots(1-\mathbb{1}_{A_n})$ par distributivité généralisée puis passer aux espérances. □

Proposition 26.24 (Formules sur des égalités ayant lieu presque sûrement, HP)**.** *Si* $X = x$ *p.s., alors* $\mathbb{E}[X] = x$*. Si* $X = Y$ *P.S., alors* $\mathbb{E}[X] = \mathbb{E}[Y]$*.*

Démonstration. On calcule

$$\mathbb{E}[X] = \mathbb{P}(X = x) \times x + \sum_{t\in X(\Omega)\setminus\{x\}} \mathbb{P}(X = t) \times t$$

Or, pour tout $t \in X(\Omega) \setminus \{x\}$, $\{X = t\} \subset \{X \neq x\}$ donc de probabilité nulle par croissance de la probabilité. D'où on tire $\mathbb{E}[X] = x$. La deuxième s'en déduit : $X - Y = 0$ p.s., donc on en déduit le résultat par linéarité et par le premier cas. □

Proposition 26.25 (HP)**.** *Supposons que* $X \geqslant 0$*. On a* $\mathbb{E}[X] = 0$ *si, et seulement si,* $X = 0$ *p.s.*

Démonstration. Dans le sens direct, on a pour tout $x \in X(\Omega)$, $\mathbb{P}(X = x)x$. Donc, pour tout $x \in \mathbb{R}_+^\star$, $\mathbb{P}(X = x) = 0$. Or $\Omega = X^{-1}(X(\Omega)) = \bigsqcup_{x\in X(\Omega)} \{X = x\}$. Nécessairement, $0 \in X(\Omega)$. Puis, par union disjointe, $\mathbb{P}(\Omega) = \mathbb{P}(X = 0) = 1$. Donc $X = 0$ p.s. Réciproquement, on utilise la proposition précédente. □

Théorème 26.13. *(Inégalité de Cauchy-Schwarz pour l'espérance) Soit X et Y des variables aléatoires réelles. On a*

$$|\mathbb{E}[XY]| \leqslant \sqrt{\mathbb{E}[X^2]}\sqrt{\mathbb{E}[Y^2]}$$

car la fonction $(X,Y) \mapsto \mathbb{E}[XY]$ est bilinéaire symétrique et positive, et ce sont les seuls hypothèses nécessaires pour l'inégalité de Cauchy-Schwarz. En revanche, le cas d'égalité devient faux. En effet, si $\mathbb{E}[X^2] = 0$, on a seulement $X = 0$ p.s.

Remarque 26.14 (HP)**.** La solution consiste à quotienter le $\mathbb{R}$-ev $\mathbb{R}^\Omega$ par la relation d'équivalence "$X = Y$ p.s.". On vérifie que l'espérance passe au quotient et fait de l'espace quotient un espace préhilbertien réel.

Théorème 26.14 (Espérance du produit de variables aléatoires indépendantes)**.** *Si X et Y sont indépendantes, alors*

$$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$$

Remarque 26.15. La réciproque est fausse en général !

Remarque 26.16. Le théorème se généralise au cas où n variables aléatoires sont mutuellement indépendantes grâce au lemme des coalitions.

Théorème 26.15 (Inégalité de Markov)**.** *Soit X à valeurs dans $\mathbb{R}_+$. Alors*

$$\forall x > 0,\ \mathbb{P}(X \geqslant x) \leqslant \frac{\mathbb{E}[X]}{x}$$

3.2 Autres moments

Définition 26.28 (Moment d'ordre k)**.** Si X est une variable aléatoire réelle, son **moment d'ordre** k vaut $\mathbb{E}[X^k]$ pour $k \in \mathbb{N}$.

Définition 26.29. La **variance** de X est définie par

$$\mathbb{V}[X] = \mathbb{E}\left((X - \mathbb{E}[X])^2\right)$$

C'est un nombre positif ou nul, qui indique la "dispersion" de X.

Proposition 26.26 (HP)**.** *Soit X réelle d'espérance μ. On a $\mathbb{V}[X] = 0$ si, et seulement si, $X = \mu$ p.s.*

Démonstration. $(X - \mathbb{E}[X])^2$ est positive. En vertu d'une propriété précédente, $\mathbb{V}[X] = 0$ ssi $(X - \mathbb{E}[X])^2 = 0$ p.s. ssi $(X - \mathbb{E}[X]) = 0$ p.s. ssi $X = \mu$ p.s. □

Proposition 26.27 (Formule de König-Huygens)**.** *On a aussi*

$$\mathbb{V}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

Proposition 26.28. *Si deux variables aléatoires ont même loi, alors elles ont même moment et donc elles ont même variance.*

Proposition 26.29. *La variance vérifie*

$$\forall (a,b) \in \mathbb{R}^2,\ \mathbb{V}[aX + b] = a\mathbb{V}[X]$$

Définition 26.30 (Variable aléatoire réduite)**.** Si $\mathbb{V}[X] = 1$, la variable aléatoire X est dite **réduite**.

Définition 26.31 (Écart-type)**.** L'**écart-type** de X est défini par

$$\sigma(X) = \sqrt{\mathbb{V}[X]}$$

Il vérifie la relation

$$\sigma(aX + b) = |a|\sigma(X)$$

Proposition 26.30. *Si X est une variable aléatoire d'écart-type non nul, alors la variable aléatoire $\dfrac{X - \mathbb{E}[X]}{\sigma(X)}$ est centrée réduite.*

Théorème 26.16 (Inégalité de Bienaymé-Techebychev)**.** *Soit X une variable aléatoire réelle. Alors*

$$\forall a > 0,\ \mathbb{P}(|X - \mathbb{E}[X]| \geqslant a) \leqslant \frac{\mathbb{V}[X]}{a^2}$$

Définition 26.32 (Covariance)**.** La **covariance** de deux variables aléatoires X et Y est définie par

$$\mathrm{cov}(X, Y) = \mathbb{E}\left[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])\right]$$

Proposition 26.31 (Formule de König-Huygens pour la covariance)**.** *On a aussi*

$$\mathrm{cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

Proposition 26.32 (Inégalité de Cauchy-Schwarz pour la covariance)**.** *La covariance est bilinéaire, symétrique et positive, donc l'inégalité de Cauchy-Schwarz s'applique (mais pas la cas d'égalité !) :*

$$|\mathrm{cov}(X, Y)| \leqslant \sqrt{\mathbb{V}[X]}\sqrt{\mathbb{V}[Y]}$$

Corollaire 26.5. *Si X et Y sont indépendantes, alors leur covariance est nulle. La réciproque est fausse !*

Définition 26.33 (Variables aléatoires décorrélées)**.** Si $\mathrm{cov}(X, Y) = 0$, on dit que X et y sont décorrélées.

Théorème 26.17 (Variance d'une somme)**.** *Soit $X_1, \ldots, X_n$ sont des variables aléatoires. Alors*

$$\mathbb{V}[X_1 + \ldots + X_n] = \mathbb{V}[X_1] + \ldots + \mathbb{V}[X_n] + \sum_{i \neq j} \mathrm{cov}(X_i, X_j)$$

Par symétrie de la covariance, cela se réécrit

$$\mathbb{V}[X_1 + \ldots + X_n] = \mathbb{V}[X_1] + \ldots + \mathbb{V}[X_n] + 2\sum_{i < j} \mathrm{cov}(X_i, X_j)$$

Corollaire 26.6. *Si $X_1, \ldots, X_n$ sont deux à deux indépendantes, on a*

$$\mathbb{V}[X_1 + \ldots + X_n] = \mathbb{V}[X_1] + \ldots + \mathbb{V}[X_n]$$

Remarque 26.17. A fortiori, le résultat reste vrai lorsque les variables aléatoires sont mutuellement indépendantes.

Corollaire 26.7 (Quelques espérances à connaître)**.** *On a les résultats suivants :*

1. *Si* $X \sim \mathcal{B}(p)$*, alors* $\mathbb{V}[X] = p(1-p)$
2. *Si* $X \sim \mathcal{B}(n,p)$*, alors* $\mathbb{V}[X] = np(1-p)$

Théorème 26.18 (Loi faible des grands nombres, HP)**.** *Soit* $(X_n)_{n\in\mathbb{N}^\star}$ *une suite de variables aléatoires réelles toutes définies sur le même univers* Ω*, indépendantes et identiquement distribuées. Notons* μ *leur espérance (commune à toutes les variables aléatoires donc). On pose*

$$\forall n \in \mathbb{N}^\star,\ Y_n = \frac{X_1 + \ldots + X_n}{n}$$

Alors, "Y_n *a fortement tendance à se concentrer autour de* μ*". Plus précisément :*

$$\forall \varepsilon > 0,\ \mathbb{P}(|Y_n - \mu| > \varepsilon) \xrightarrow[n\to+\infty]{} 0$$

Ce théorème justifie l'approche fréquentiste des probabilités.

Démonstration. L'espérance de Y_n vaut toujours μ. Utiliser ensuite l'inégalité de Bienaymé-Tchebychev et les propriétés de la variance pour conclure. □

Bibliographie

[1] F. Morlot, Lycée Privé Sainte-Geneviève, Cours dispensés en MPSI1 durant l'année 2023-2024
[2] C. Lafitte, Lycée Privé Sainte-Geneviève, *Compléments sur les anneaux commutatifs*