COURS DE MATHÉMATIQUES PREMIÈRE ANNÉE (L1)

UNIVERSITÉ DENIS DIDEROT PARIS 7

Marc HINDRY

Tous les chapitres sont importants. Le premier chapitre est volontairement bref mais fondamental : il y aura intérêt à revenir sur les notions de langage mathématique et de raisonnement tout au long du cours, à l'occasion de démonstrations. Les chapitre 19 et 20 reposent sur une synthèse de l'algèbre (linéaire) et de l'analyse (calcul différentiel et intégral) tout en étant assez géométriques. Le chapitre 21 (fonctions de plusieurs variables) appartient en pratique plutôt à un cours de deuxième année; il a été ajouté pour les étudiants désirant anticiper un peu ou ayant besoin, par exemple en physique, d'utiliser les fonctions de plusieurs variables et dérivées partielles, dès la première année.

L'ordre des chapitres. L'ordre choisi n'est que l'un des possibles. En particulier on pourra vouloir traiter l'"analyse" (chapitres 12-20) en premier : pour cela on traitera d'abord le chapitre sur les nombres réels et complexes (ou la notion de limite est introduite très tôt), le principe de récurrence et on grapillera quelques notions sur les polynômes et l'algèbre linéaire. La séquence d'algèbre linéaire (chapitres 7-11) est très inspirée de la présentation par Mike Artin (Algebra, Prentice-Hall 1991) mais on peut choisir bien d'autres présentations. On pourra aussi par exemple préférer étudier $\mathbf{Z}$ avant $\mathbf{R}$ et $\mathbf{C}$ (du point de vue des constructions, c'est même préférable!). Le chapitre 16 sur les fonctions usuelles peut être abordé à peu près à n'importe quel moment, quitte à s'appuyer sur les notions vues en terminale.

Nous refusons le point de vue : "... cet ouvrage part de zéro, nous ne supposons rien connu...". Au contraire nous pensons qu'il faut s'appuyer sur les connaissances de terminale et sur l'intuition (notamment géométrique). Il semble parfaitement valable (et utile pédagogiquement) de parler de droites, courbes, plans, fonction exponentielle, logarithme, sinus, etc ... avant de les avoir formellement introduit dans le cours. Il semble aussi dommage de se passer complètement de la notion très intuitive d'angle sous prétexte qu'il s'agit d'une notion délicate à définir rigoureusement (ce qui est vrai).

Illustrations : Nous avons essayé d'agrémenter le cours d'applications et de motivations provenant de la physique, de la chimie, de l'économie, de l'informatique, des sciences humaines et même de la vie pratique ou récréative. En effet nos pensons que même si on peut trouver les mathématiques intéressantes et belles en soi, il est utile de savoir que beaucoup des problèmes posés ont leur origine ailleurs, que la séparation avec la physique est en grande partie arbitraire et qu'il est passionnant de chercher à savoir à quoi sont appliquées les mathématiques.

Indications historiques Il y a hélas peu d'indications historiques faute de temps, de place et de compétence mais nous pensons qu'il est souhaitable qu'un cours contienne des allusions : 1) au développement historique, par exemple du calcul différentiel 2) aux problèmes ouverts (ne serait-ce que pour mentionner leur existence) et aux problème résolus disons dans les dernières années. Les petites images (mathématiques et philathéliques) incluses à la fin de certains chapitres sont donc une invitation à une recherche historique.

Importance des démonstrations Les mathématiques ne se réduisent pas à l'exactitude et la rigueur mais quelque soit le point de vue avec lequel ont les aborde la notion de démonstration y est fondamentale. Nous nous efforçons de donner presque toutes les démonstrations. L'exception la plus notable est la construction des fonctions cosinus et sinus, pour laquelle nous utiliserons l'intuition géométrique provenant de la représentation du cercle trigonométrique ; l'intégrabilité des fonctions continues sera aussi en partie admise.

Il y a là une difficulté qui sera levée avec l'étude des fonctions analytiques (faite en seconde année).

Difficulté des chapitres Elle est inégale et bien sûr difficile à évaluer. Certains chapitres développent essentiellement des techniques de calculs (chapitres 6, 7, 10, 16, 17, 18, 19, 20), le chapitre 11 reprend du point de vue de l'algèbre linéaire des notions vues en terminales, d'autres développent des concepts (chapitres 2, 3, 4, 5, 8, 9, 12, 13, 15) et sont donc en ce sens plus difficiles ; le chapitre 14 est intermédiaire dans cette classification un peu arbitraire. Enfin le chapitre 21 n'est destiné à être appronfondi qu'en deuxième année.

Résumés En principe les énoncés importants sont donnés sous l'entête "thèorème" suivis par ordre décroissant d'importance des "propositions" et des "lemmes". Un "résumé" de chaque chapitre peut donc être obtenu en rassemblant les énoncés des théorèmes (et les définitions indispensables à la compréhension des énoncés). Nous avons seulement inclus un chapitre résumant et synthétisant les différents points de vue développés en algèbre linéaire (après le chapitre 11).

Archimède [$A\rho\chi\iota\mu\acute{\eta}\delta\eta\varsigma$] ($\sim$ 287–$\sim$ 212)

Al Khwārizmī (fin VIII^e, début IX^e)

CHAPITRE 1 LE LANGAGE MATHÉMATIQUE

Ce chapitre, volontairement court, précise les modalités du raisonnement mathématique. En effet on n'écrit pas un texte mathématique comme un texte de langage courant : ce serait théoriquement possible mais totalement impraticable pour de multiples raisons (le raccourci des "formules" est notamment une aide précieuse pour l'esprit).

Une *définition* précise le sens mathématique d'un mot ; par exemple :

Définition: Un ensemble E est fini si il n'est pas en bijection avec lui-même privé d'un élement. Un ensemble est infini si il n'est pas fini.

On voit tout de suite deux difficultés avec cet exemple : d'abord il faut avoir défini "ensemble" (ce que nous ne ferons pas) et "être en bijection" (ce qu'on fera au chapitre suivant) pour que la définition ait un sens ; ensuite il n'est pas immédiat que la définition donnée coïncide avec l'idée intuitive que l'on a d'un ensemble fini (c'est en fait vrai).

Un *énoncé mathématique* (nous dirons simplement *énoncé*) est une phrase ayant un sens mathématique précis (mais qui peut être vrai ou faux) ; par exemple :

(A) 1=0

(B) Pour tout nombre réel x on a $x^2 \geq 0$

(C) $x^3 + x = 1$

sont des énoncés ; le premier est faux, le second est vrai, la véracité du troisième dépend de la valeur de la variable x. Par contre, des phrases comme "les fraises sont des fruits délicieux", "j'aime les mathématiques" sont clairement subjectives. L'affirmation : "l'amiante est un cancérogène provoquant environ trois mille décès par an en France et le campus de Jussieu est floqué à l'amiante" n'est pas un énoncé mathématique, même si l'affirmation est exacte. Nous ne chercherons pas à définir précisément la différence entre énoncé mathématique et énoncé non mathématique.

Un *théorème* est un énoncé vrai en mathématique ; il peut toujours être paraphrasé de la manière suivante : "Sous les hypothèses suivantes : .... , la chose suivante est toujours vraie :... ". Dans la pratique certaines des hypothèses sont omises car considérés comme vraies a priori : ce sont les *axiomes*. La plupart des mathématiciens sont d'accord sur un certain nombre d'axiomes (ceux qui fondent la théorie des ensembles, voir chapitre suivant) qui sont donc la plupart du temps sous-entendus.

Par exemple nous verrons au chapitre 5 que :

THÉORÈME: *Soit n un nombre entier qui n'est pas le carré d'un entier alors il n'existe pas de nombre rationnel x tel que $x^2 = n$ (en d'autres termes $\sqrt{n}$ n'est pas un nombre rationnel).*

Pour appliquer un théorème à une situation donnée, on doit d'abord vérifier que les hypothèses sont satisfaites dans la situation donnée, traduire la conclusion du théorème dans le contexte et conclure.

Par exemple : prenons $n = 2$ (puis $n = 4$) alors 2 n'est pas le carré d'un entier donc le théorème nous permet d'affirmer que $\sqrt{2}$ n'est pas un nombre rationnel. Par contre l'hypothèse n'est pas vérifiée pour $n = 4$ et le théorème ne permet pas d'affirmer que $\sqrt{4}$ n'est pas un nombre rationnel (ce qui serait d'ailleurs bien sûr faux!).

Les *connecteurs logiques* permettent de fabriquer de nouveaux énoncés à partir d'autres ; nous utiliserons exclusivement les connecteurs suivants :

non : non(A) est vrai si et seulement si (A) est faux

ou : (A) *ou* (B) est vrai si et seulement si (A) est vrai ou (B) est vrai.

et : (A) *et* (B) est vrai si et seulement si (A) est vrai et (B) est vrai.

implique (en symbole $\Rightarrow$) : (A) *implique* (B) est vrai si et seulement si chaque fois que (A) est vrai alors (B) est aussi vrai.

équivaut (en symbole $\Leftrightarrow$) : (A) équivaut (B) est vrai si (A) est vrai chaque fois que (B) est vrai et réciproquement.

Une *démonstration logique* (nous dirons ensuite simplement une démonstration) est un énoncé, comportant éventuellement comme variable d'autres énoncés de sorte qu'il soit vrai quel que soit les énoncés variables. Voici des exemples de démonstration :

Si $(A) \Rightarrow (B)$ et $(B) \Rightarrow (C)$ alors $(A) \Rightarrow (C)$

$non(non(A))$ équivaut à (A)

Si $(A) \Rightarrow (B)$ et $non(B)$ alors $non(A)$.

Si (A) *ou* (B) et $non(B)$ alors (A).

Bien entendu, les démonstrations "intéressantes" en mathématiques sont plus longues et sont composées de chaînes d'implications élémentaires comme celles qui précèdent. Une manière simple (mais fastidieuse) de vérifier ce type d'énoncé est faire un tableau avec les diverses possibilités : chaque énoncé est vrai ou faux (V ou F). Par exemple, pour le premier énoncé il y a huit possibilités :

A	B	C	$A \Rightarrow B$	$B \Rightarrow C$	$A \Rightarrow C$
V	V	V	V	V	V
V	V	F	V	F	F
V	F	V	F	V	V
V	F	F	F	V	F
F	V	V	V	V	V
F	V	F	V	F	V
F	F	V	V	V	V
F	F	F	V	V	V

On constate bien que chaque fois que $A \Rightarrow B$ et $B \Rightarrow C$ sont simultanément vrais alors $A \Rightarrow C$ est vrai aussi.

Exemples de raisonnements parmi les plus utilisés :

Raisonnement cas par cas :
Schéma : si (A) *ou* (B), $(A) \Rightarrow (C)$ et $(B) \Rightarrow (C)$, alors C

Raisonnement par contraposée :
Schéma : si $(A) \Rightarrow (B)$, alors $non(B) \Rightarrow non(A)$

Raisonnement par l'absurde :
Schéma : si $(B) \Rightarrow (A)$ *et* $non(A)$, alors $non(B)$.

On voit qu'il n'y a aucune difficulté fondamentale avec les raisonnements logiques, la seule difficulté est parfois d'arriver à enchaîner les déductions. A titre d'exercice on vérifiera les déductions suivantes :

$non((A)\ ou\ (B)) \Leftrightarrow (non(A)\ et\ non(B))$

$non((A)\ et\ (B)) \Leftrightarrow (non(A)\ ou\ non(B))$

$non(A)\ ou\ (B) \Leftrightarrow (A \Rightarrow B)$

$(A\ et\ B)\ ou\ (C) \Leftrightarrow (A\ ou\ C)\ et\ (B\ ou\ C)$

Les *quantificateurs* permettent de transformer un énoncé contenant une variable en un énoncé "absolu" : nous utiliserons exclusivement deux quantificateurs :

il existe (en symbole $\exists$)

pour tout (en symbole $\forall$)

Exemple : considérons les énoncés suivants contenant la variable $x \in \mathbf{R}$.

$$A(x)\ : x^2 - 1 = 0$$
$$B(x)\ : x^2 + x = x(x+1)$$
$$C(x)\ : x + 1 = x$$

L'affirmation $(\forall x \in \mathbf{R}\ non(C(x)))$ tout comme $(\exists x \in \mathbf{R}\ A(x))$ est vraie. Par contre il est faux que : $\forall x \in \mathbf{R}\ A(x)$

La négation de $\forall x\ A(x)$ est $\exists x\ non(A(x))$. La négation de $\exists x\ A(x)$ est $\forall x\ non(A(x))$.

Par exemple la négation de :

$$(A)\ :\ \forall x \in \mathbf{R},\ \forall \epsilon \in \mathbf{R}_+^*,\ \exists \delta \in \mathbf{R}_+^*,\ \forall y \in \mathbf{R},\ |x-y| \leq \delta \Rightarrow |f(x) - f(y)| \leq \epsilon$$

est :

$$non(A)\ :\ \exists x \in \mathbf{R},\ \exists \epsilon \in \mathbf{R}_+^*,\ \forall \delta \in \mathbf{R}_+^*,\ \exists y \in \mathbf{R},\ |x-y| \leq \delta \text{ et } |f(x) - f(y)| > \epsilon$$

Remarque : l'énoncé (A) écrit que la fonction f est continue en tout point alors que $non(A)$ écrit qu'il existe un point où f n'est pas continue (voir chapitre 13).

Commentaires : la nécessité de la formalisation du raisonnement mathématique et de la notion d'ensemble a accompagné historiquement l'apparition de *paradoxes* au tournant de ce siècle. Ceux-ci sont essentiellement de deux types : paradoxes sémantiques et paradoxes logiques.

Un exemple de paradoxe sémantique est le suivant : on choisit un dictionnaire de langue française et on considère l'ensemble S des nombres entiers que l'on peut définir à l'aide de moins de vingt mots de ce dictionnaire. Comme le nombre de mots est fini et le nombre de phrase de moins de vingt mots est fini, l'ensemble S est fini ; il existe donc "Le plus petit nombre entier que l'on ne peut pas définir en moins de vingt mots". Mais nous venons de le définir en moins de vingt mots!

Un exemple de paradoxe logique (dû à Russel) est le suivant : considérons l'ensemble S formé de tous les éléments qui ne s'appartiennent pas à eux-mêmes ; en symboles :

$$S := \{x \mid x \notin x\}$$

Cet ensemble à l'air inoffensif mais si on pense que $S \in S$ alors on en déduit $S \notin S$ et inversement!

La méthode pour éliminer les paradoxes du premier type est de se restreindre au langage purement mathématique (ou plus précisément de séparer langage et métalangage, nous ne précisons pas cette notion) : on se borne à travailler avec des notions qui peuvent s'écrire en langage symbolique (idéalement on pourrait penser à écrire tout en langage symbolique, mais on s'aperçoit vite que pour des raisons de longueur, c'est impraticable).

La méthode pour éliminer les paradoxes du type "Russel" est de restreindre la notion d'ensemble ; en particulier on déclare qu'on ne peut pas former un ensemble seulement à partir d'un énoncé avec variables. Ainsi $S := \{x \mid A(x)\}$ ne définit pas nécessairement un ensemble ; par contre, si T est un ensemble alors $S := \{x \in T \mid A(x)\}$ définit encore un (sous-)ensemble.

Terminons ce premier chapitre par une description lapidaire de l'usage et de la place des mathématiques au sein des autres sciences.

Un des paradigmes des sciences peut être succintement décrit par le diagramme suivant :

$$\begin{array}{ccc} \text{observation} & \longrightarrow & \text{modélisation} \\ \downarrow & & \downarrow \quad \text{Math.} \\ \text{expérience} & \longrightarrow & \text{prédiction} \end{array}$$

Concernant les *applications* des notions de ce cours en sciences indiquons par une flèche quelques unes des plus marquantes :

- Algèbre et Arithmétique → informatique;
- Théorie des groupes → chimie;
- Calcul différentiel et intégral → physique;
- Equations différentielles → physique, biologie, économie;

Exercice : (logique, inégalités, ...)

Sachant que les statistiques disponibles (code 163 de l'INSERM) indiquent 902 décès pour l'année 1994 par mésothéliome de la plèvre (cancer mortel, causé par l'inhalation de fibres d'amiante), discuter la compatibilité des déclarations suivantes du professeur Brochard, chercheur à l'INSERM, membre du Comité Permanent Amiante (C.P.A) :

(a) "Le mésothéliome est un cancer rare, moins de 200 cas par an [en France]" (C.P.A, l'amiante et la santé, page 13, 1994). (b) "Au moins 150 mésothéliomes dus à l'amiante [par an en France]" (déclaration sur TF1, fin 1994). (c) "On aurait en fait 440 mésothéliomes par an en France" (rapport destiné au ministère du travail, novembre 1994)

"Environ 600 mésothéliomes pleuraux en 1992, en France" (conférence internationale sur le mésothéliome à Créteil, 1995)(*)

Indications : on pourra utiliser les tables de vérité et aussi le fait que le C.P.A a été créé et financé par les industriels de l'amiante et géré par l'agence de communnication et lobbying "Communications Economiques et Sociales" (C.E.S. 10 Avenue de Messine, 75008 Paris).

(*) Post-Scriptum (1996) Le rapport INSERM sur "*les effets sur la santé de l'amiante* " conclut qu'il y a *au minimum* 750 décès par an en France dus aux mésothéliomes causés par l'amiante.

CHAPITRE 2 ENSEMBLES ET APPLICATIONS.

Georg Cantor, le fondateur de la théorie des ensembles définissait un ensemble comme "un groupement d'objets déterminés et bien distincts, de notre perception ou de notre entendement, et que l'on appelle les éléments de l'ensemble". Nous considèrerons la notion d'ensemble comme intuitive en gardant néanmoins en mémoire le fait qu'on ne peut pas considérer "n'importe quoi" comme un ensemble si l'on veut éviter les contradictions. Nous allons donc juste définir les opérations usuelles sur les ensembles (sous-ensembles, complémentaires, intersections, unions, produits, ensemble des parties) puis nous abordons les deux points cruciaux : la notion de fonction (ou application) qui est fondamentale dans toutes les mathématiques et le concept d'infini avec l'exemple fondamental : l'ensemble des entiers naturels, noté $\mathbf{N}$*, est infini.*

2.1 ENSEMBLES

Dans la pratique il y a deux façons de construire ou décrire des ensembles : en donnant la liste de ses éléments, par exemple $E := \{0, 1, 2, 3, 5, 7, 8\}$ est un ensemble, ou bien en décrivant une caractérisation des éléments, par exemple nous admettrons que $\mathbf{N} := \{n \mid n \text{ est un entier naturel}\}$ est un ensemble. Parmi les ensembles les plus importants nous étudierons outre $\mathbf{N}$ déjà cité, l'ensemble des nombres entiers relatifs, noté $\mathbf{Z}$, l'ensemble des nombres rationnels, noté $\mathbf{Q}$, l'ensemble des nombres réels, noté $\mathbf{R}$ et l'ensemble des nombres complexes, noté $\mathbf{C}$.

Ensemble vide : il s'agit de l'ensemble ne contenant aucun élément ; on le note $\emptyset$; on peut aussi le définir comme $\emptyset := \{x \mid x \neq x\}$

Relations entre éléments et ensembles :

Un ensemble E est donc une collection d'objets qu'on appelle éléments ; pour chaque élément x on écrit $x \in E$ (lire "x appartient à E"). Si l'élément x n'est pas dans l'ensemble E on écrira $x \notin E$ (lire "x n'appartient pas à E").

Par exemple il est clair que $4 \in \mathbf{N}$ et $4 \notin \emptyset$. Quelque soit l'élément x on a toujours $x \notin \emptyset$.

On dit qu'un ensemble E est *inclus* dans un autre ensemble F (ce qu'on note $E \subset F$), si tous les éléments de E sont aussi dans F ; en d'autres termes si $x \in E \Rightarrow x \in F$. Deux ensembles sont égaux si ils ont les mêmes éléments ; en particulier :

$$E \subset F \text{ et } F \subset E \Leftrightarrow E = F$$

Par exemple $\emptyset \subset \mathbf{N}$ mais les ensembles ne sont pas égaux (donc $non(\mathbf{N} \subset \emptyset)$ ou encore $\mathbf{N} \not\subset \emptyset$).

Opérations sur les ensembles :

Sous-ensemble : si E est un ensemble et $A(x)$ un énoncé avec une variable x dans E, on peut fabriquer l'ensemble :

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

Par exemple l'ensemble des nombres entiers pairs est décrit par :

$$\mathcal{P} := \{x \in \mathbf{N} \mid \exists y \in \mathbf{N},\ x = 2y\}$$

Complémentaire : Soit F un sous-ensemble de E ; on définit le complémentaire de F dans E que l'on note $\mathrm{C}_E F$ (ou simplement $\mathrm{C}\, F$ si E est sous-entendu) comme l'ensemble des éléments de E qui n'appartiennent pas à F :

$$\mathrm{C}_E F := \{x \in E \mid x \notin F\}$$

Si F n'est plus nécessairement un sous-ensemble de E on emploiera la notation : $E \setminus F$ pour désigner $\{x \in E \mid x \notin F\}$.

Par exemple le complémentaire de $\mathcal{P}$ dans $\mathbf{N}$ est l'ensemble des nombres impairs :

$$\mathrm{C}_{\mathbf{N}} \mathcal{P} = \mathcal{I} := \{x \in \mathbf{N} \mid \exists y \in \mathbf{N},\ x = 2y + 1\}$$

Intersection : si E et F sont deux ensembles on peut former un ensemble appelé leur intersection notée $E \cap F$ et définie par :

$$E \cap F := \{x \in E \mid x \in F\} = \{x \in F \mid x \in E\} = \{x \mid x \in E \text{ et } x \in F\}$$

Par exemple, si $E = \{0, 1, 2, 3, 5, 7, 8\}$ et $\mathcal{P}$ désigne l'ensemble des entiers pairs, alors $E \cap \mathcal{P} = \{0, 2, 8\}$.

Union : si E et F sont deux ensembles on peut former un ensemble appelé leur union et notée $E \cup F$ et définie par :

$$E \cup F := \{x \mid x \in E \text{ ou } x \in F\}$$

Par exemple si $E := \{0, 1, 2, 3, 5, 7, 8\}$ et $F := \{0, 1, 2, 4, 8, 16, 32\}$ alors $E \cup F = \{0, 1, 2, 3, 4, 5, 7, 8, 16, 32\}$

Produit : Si $x \in E$ et $y \in F$ on peut fabriquer un nouvel élément appelé *couple* et noté (x, y), caractérisé par le fait que $(x, y) = (z, t)$ si et seulement si $x = z$ et $y = t$. L'ensemble de ces couples s'appelle le produit (cartésien) de E et F et se note :

$$E \times F := \{(x, y) \mid x \in E \text{ et } y \in F\}$$

Pour se représenter un produit cartésien on aura avantage à avoir en tête l'exemple suivant : soit $E := [0, 3]$ (l'intervalle des nombres réels compris entre 0 et 3) et $F := [0, 1]$ alors $E \times F$ est le rectangle de la figure suivante

Un autre exemple familier est celui du plan que l'on peut représenter comme le produit $\mathbf{R} \times \mathbf{R}$.

Ensemble des parties : Soit E un ensemble, on peut former un nouvel ensemble dont les éléments sont les sous-ensembles de E et que l'on note $\mathcal{P}(E)$:

$$\mathcal{P}(E) := \{F \mid F \subset E\}$$

Par exemple $\mathcal{P}(\emptyset) = \{\emptyset\}$ (ensemble avec un élément) mais on a aussi $\mathcal{P}(\{0,1\}) = \{\emptyset, \{0\}, \{1\}, \{0,1\}\}$ (ensemble avec quatre éléments)

Remarque : on notera que l'on n'a pas donné de démonstration pour l'existence de l'union, du produit etc. En fait il faut comprendre ces énoncés comme des *axiomes* i.e. des énoncés élémentaires que l'on admet être vrais et à partir desquels on va démontrer toutes les autres affirmations. Le caractère extrêmement intuitif (on a envie de dire "évident" de ces axiomes fait qu'ils sont admis par presque tout le monde).

Calculs sur les ensembles : il est très important de savoir calculer et raisonner sur les ensembles ; il faut aussi remarquer que le calcul sur les ensembles est entièrement analogue au calcul sur les propositions ; en effet l'union correspond au connecteur *ou*, l'intersection correspond au connecteur *et* et la relation d'inclusion correspond à l'implication, prendre le complémentaire correspond au connecteur *non* : si les éléments x de A sont caractérisés par la propriété $P(x)$ et ceux de B par la propriété $Q(x)$ alors :

Les éléments x de $A \cup B$ sont caractérisés par la propriété $P(x)$ *ou* $Q(x)$.

Les éléments x de $A \cap B$ sont caractérisés par la propriété $P(x)$ *et* $Q(x)$.

La relation $A \subset B$ équivaut à l'implication $\forall x,\ P(x) \Rightarrow Q(x)$.

Les éléments x de $\mathrm{C}_E\, A$ sont caractérisés, parmi les éléments de E par la propriété $non(A(x))$.

Ainsi le calcul sur les ensembles peut toujours se ramener au calcul propositionnel ; voici une liste (non exhaustive) de formules où $A, B, C, \ldots$ sont des ensembles :

Formulaire

$A \cap B = B \cap A$ et $A \cup B = B \cup A$ (commutativité)
$A \cap (B \cap C) = (A \cap B) \cap C$ et $A \cup (B \cup C) = (A \cup B) \cup C$ (associativité)
$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ et $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ (distributivité)
$\mathrm{C}_E(\mathrm{C}_E\, A) = A$
$(A \subset B) \Rightarrow (\mathrm{C}_E\, B \subset \mathrm{C}_E\, A)$
$\mathrm{C}_E(A \cup B) = \mathrm{C}_E\, A \cap \mathrm{C}_E\, B$ et $\mathrm{C}_E(A \cap B) = \mathrm{C}_E\, A \cup \mathrm{C}_E\, B$ (loi de Morgan)
$A \times (B \cap C) = (A \times B) \cap (A \times C)$ et $A \times (B \cup C) = (A \times B) \cup (A \times C)$
$(A \subset B)$ *et* $(C \subset D) \Rightarrow A \times C \subset B \times D$

Démonstration: Démontrons la première formule de distributivité :

$x \in A \cap (B \cup C) \Leftrightarrow x \in A$ *et* $(x \in B$ *ou* $x \in C) \Leftrightarrow (x \in A$ *et* $x \in B)$*ou*$(x \in A$ *et* $x \in C) \Leftrightarrow x \in (A \cap B) \cup (A \cap C)$.

La loi de Morgan se démontre de manière similaire :

$$x \in \mathrm{C}(A \cup B) \Leftrightarrow non(x \in A \text{ ou } x \in B) \Leftrightarrow non(x \in A) \text{ et } non(x \in B) \Leftrightarrow x \in \mathrm{C}\, A \cap \mathrm{C}\, B$$

Les autre démonstrations sont similaires et laissées en exercice. □

2.2 APPLICATIONS

Définition: Une *application* (ou *fonction*) définie sur X et à valeurs dans Y est une loi qui, à tout élément de X fait correspondre un unique élément de Y. Si on note f cette application, l'élément associé à x par f est noté $f(x)$. L'ensemble X s'appelle l'*ensemble de départ*, l'ensemble Y s'appelle l'*ensemble d'arrivée* de f. On note souvent une fonction $f : X \to Y$ ou, si les ensembles X et Y sont sous-entendus $x \mapsto f(x)$. L'élément $f(x) = y$ s'appelle l'*image* de x par f et x s'appelle un *antécédent* de y par f.

Remarque : une fonction peut être définie par son *graphe*, un sous-ensemble $\Gamma \subset X \times Y$ qui possède la propriété suivante : $\forall x \in X,\ \exists y \in Y,\ (x, y) \in \Gamma$ et de plus $(x, y) \in \Gamma$ et $(x, y') \in \Gamma \Rightarrow y = y'$. Le graphe d'une fonction f est l'ensemble des couples $(x, f(x))$ pour $x \in X$.

Remarque : une phrase usuelle comme "la fonction $\cos(x)$" comporte une ambigüité qui devient transparente si on augmente la phrase en "la fonction $\cos(x)$ est une bijection" qui est manifestement fausse si on parle d'une fonction de $\mathbf{R}$ dans $\mathbf{R}$ et néanmoins vraie si l'on parle d'une fonction de $[0, \pi]$ vers $[-1, +1]$ (voir le chapitre 16).

Remarque : on ne fait pas de distinction entre fonction et application.

Exemples :

L'association $x \mapsto x^2 + 1$ définit une application de $\mathbf{R}$ dans $\mathbf{R}$.

L'association $x \mapsto \sqrt{x}$ définit une application de $\mathbf{N}$ dans $\mathbf{R}$ (mais pas de $\mathbf{N}$ dans $\mathbf{N}$).

L'association $x \mapsto \frac{1}{x^2-1}$ définit une application de $\mathbf{R} \setminus \{+1, -1\}$ dans $\mathbf{R}$.

La loi qui associe à un point du plan Π son symétrique par rapport à un point donné O, définit une application de Π dans Π.

L'association $F \mapsto \complement_E F$ définit une application de $\mathcal{P}(E)$ dans $\mathcal{P}(E)$.

L'application qui à tout élément $x \in X$ associe x s'appelle l'*application identique* et se note id_X.

Si f est une application de X dans Y et si X' est un sous-ensemble de X, on peut définir f' la *restriction* de f à X' par : $\forall x \in X',\ f'(x) := f(x)$.

Composition : Si $f : X \to Y$ et $g : Y \to Z$ sont deux applications, on peut définir la *composée* de f et g par $(g \circ f)(x) = g(f(x))$. Une propriété importante de la composition des applications est l'associativité :

PROPOSITION: *La composition des applications est associative. C'est-à-dire que si $h : X \to Y$, $g : Y \to Z$ et $f : Z \to W$ sont trois applications, alors $(f \circ g) \circ h = f \circ (g \circ h)$ (que l'on note donc simplement $f \circ g \circ h$).*

Démonstration: En effet $\forall x \in X,\ (f \circ (g \circ h)(x) = f((g \circ h)(x)) = f(g(h(x)))$ et $((f \circ g) \circ h)(x) = (f \circ g)(h(x)) = f(g(h(x)))$ sont bien égaux. □

Exemples : Si f est donnée par $x \mapsto \frac{1}{x^2-1}$ de $\mathbf{R} \setminus \{+1, -1\}$ dans $\mathbf{R}$ et g est donnée par $x \mapsto x^2 + 1$ de $\mathbf{R}$ dans $\mathbf{R}$; alors $g \circ f$ est une application de $\mathbf{R} \setminus \{+1, -1\}$ dans $\mathbf{R}$ décrite par $g \circ f(x) = (\frac{1}{x^2-1})^2 + 1$.

Si f est la symétrie du plan Π par rapport au point O, alors $f \circ f = id_\Pi$.

Il est souvent intéressant de décomposer une application (par exemple pour calculer sa dérivée) ; par exemple l'application définie par $f(x) := \left(e^{\cos(x)} + 1\right)^3$ se décompose en $f = g \circ h \circ k$ où $k(x) = \cos(x)$, $h(x) = e^x$ et $g(x) = (x + 1)^3$.

Il est naturel, disposant d'une fonction f d'étudier les équations du type : $f(x) = f(y)$ ou encore $y = f(x)$. Cela conduit à la notion d'application injective ou surjective.

Définition: Une application $f : X \to Y$ est *injective* si (pour tout $x, y \in X$) l'égalité $f(x) = f(y)$ entraîne $x = y$. En d'autres termes tout élément de Y a au plus un antécédent ou encore est l'image d'au plus un élément de X.

Exemple : les fonctions $x \mapsto x+2$ (de $\mathbf{R}$ dans $\mathbf{R}$) et $x \mapsto \log(x)$ (de $\mathbf{R}_+^*$ dans $\mathbf{R}$) sont injectives mais les fonctions $x \mapsto x^2$ et $x \mapsto \sin(x)$ de $\mathbf{R}$ dans $\mathbf{R}$ ne sont pas injectives.

Définition: Une application $f : X \to Y$ est *surjective* si, pour tout $y \in Y$ il existe $x \in X$ tel que $y = f(x)$. En d'autres termes tout élément de Y a au moins un antécédent.

Exemple : La fonction f définie par $f(x) = x + 2$ de $\mathbf{R}$ dans $\mathbf{R}$ est surjective. La fonction définie par $g(x) = x^2$ de $\mathbf{R}$ dans $\mathbf{R}$ n'est pas surjective. Par contre la "même" fonction considérée de $\mathbf{R}$ dans $\mathbf{R}_+$ est surjective. On voit donc qu'il faut bien préciser ensemble de départ et d'arrivée pour parler de surjectivité et d'injectivité.

Remarque : considérons les "mêmes" fonctions mais sur des ensembles différents. Les fonctions $x \mapsto x^2$ restreinte à $\mathbf{R}_+$ et $x \mapsto \sin(x)$ à l'intervalle $[-\frac{\pi}{2}, \frac{\pi}{2}]$ sont injectives. La fonction $x \mapsto x^2$ considérée de $\mathbf{R}$ dans $\mathbf{R}_+$ est surjective. On voit donc qu'il faut bien préciser ensemble de départ et d'arrivée pour parler de surjectivité et d'injectivité.

Définition: Une application $f : X \to Y$ est *bijective* si elle est à la fois injective et surjective. En d'autres termes tout élément de Y a exactement un antécédent.

Exemple : La fonction f de $\mathbf{R}$ dans $\mathbf{R}$ donnée par $x \mapsto x + 2$ est une bijection ; de même la fonction $x \mapsto \log(x)$ est une bijection de $\mathbf{R}_+^*$ dans $\mathbf{R}$.

Lorsque $f : X \to Y$ est une bijection, on peut définir une application de Y dans X par la loi qui à y associe l'unique élément x tel que $y = f(x)$ (le fait que f soit bijective garantit exactement l'existence et l'unicité d'un tel x).

Définition: On appelle *bijection réciproque* d'une bijection f et on note f^{-1} l'application caractérisée par : $x = f^{-1}(y) \Leftrightarrow y = f(x)$. Il est clair que f^{-1} est aussi une bijection.

Exemple : la bijection réciproque de $x \mapsto x+2$ est donnée par $x \mapsto x-2$. La bijection réciproque de $x \mapsto \log(x)$ de $\mathbf{R}_+^*$ dans $\mathbf{R}$ est la fonction $x \mapsto \exp(x)$ de $\mathbf{R}$ dans $\mathbf{R}_+^*$. La symétrie par rapport à un point du plan est sa propre bijection réciproque.

Définition: Soit $f : E \to F$ une application.

i) Si A est une partie de E on appelle *image directe* de A par f et on note $f(A)$ l'ensemble :

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

ii) Si B est une partie de F on appelle *image réciproque* de B par f et on note $f^{-1}(B)$ l'ensemble :

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

Remarques : On prendra bien garde à ne pas confondre l'application $f^{-1} : \mathcal{P}(F) \to \mathcal{P}(E)$ ainsi définie (qui existe pour toute fonction f) avec la bijection réciproque $f^{-1} : F \to E$ (qui n'existe que si f est bijective).

On pourra vérifier en exercice que :
(i) f est surjective si et seulement si $F = f(E)$
(ii) f est injective si et seulement si $f : E \to f(E)$ est une injection.

PROPOSITION: *Soit $f : E \to F$ une application, on a les formules suivantes*
(i) Pour toutes parties A, B de E

$$f(A \cup B) = f(A) \cup f(B) \ , \ A \subset B \Rightarrow f(A) \subset f(B) \text{ et } f(A \cap B) \subset f(A) \cap f(B)$$

(les deux derniers ensembles sont, en général, distincts).
(ii) Pour toutes parties A, B de F, on a $f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B)$, $f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B)$, $A \subset B \Rightarrow f^{-1}(A) \subset f^{-1}(B)$ et également $f^{-1}(\mathrm{C}_F\, A) = \mathrm{C}_E\, f^{-1}(A)$

Démonstration: Supposons $y \in f(A \cup B)$ c'est-à-dire $y = f(x)$ avec $x \in A \cup B$ soit encore $x \in A$ ou $x \in B$; alors $y = f(x) \in f(A)$ ou $y = f(x) \in f(B)$ donc $y = f(x) \in f(A) \cup f(B)$; ainsi $f(A \cup B) \subset f(A) \cup f(B)$. Si maintenant $y \in f(A) \cup f(B)$ alors $y = f(x')$ avec $x' \in A$ ou bien $y = f(x'')$ avec $x'' \in B$ donc il existe $x \in A \cup B$ (égal à x' ou x'') tel que $y = f(x)$ donc $y \in f(A \cup B)$ et $f(A) \cup f(B) \subset f(A) \cup f(B)$ et finalement l'égalité des deux ensembles.

Supposons $y \in f(A \cap B)$, alors $y = f(x)$ avec $x \in A \cap B$ donc $x \in A$ et $y = f(x) \in f(A)$ mais aussi $x \in B$ donc $y = f(x) \in f(B)$; on peut conclure $y \in f(A) \cap f(B)$ et $f(A \cap B) \subset f(A) \cap f(B)$. L'exemple suivant montre qu'on n'a pas en général égalité : prenons $E := \{a, b\}$, $F := \{c\}$, $f(a) = f(b) = c$, $A := \{a\}$ et $B := \{b\}$. Alors $\emptyset = f(A \cap B) \neq f(A) \cap f(B) = \{c\}$.

Pour changer un peu raisonnons par équivalence : $x \in f^{-1}(A \cup B)$ équivaut à $f(x) \in A \cup B$, qui équivaut à $f(x) \in A$ ou $f(x) \in B$, qui équivaut à $x \in f^{-1}(A)$ ou $x \in f^{-1}(B)$, qui équivaut à $x \in f^{-1}(A) \cup f^{-1}(B)$. Ainsi on a bien $f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B)$.

On fait de même avec l'intersection ; enfin $x \in f^{-1}(\mathrm{C}_F\, A)$ équivaut à $f(x) \in \mathrm{C}_F\, A$, qui équivaut à $f(x) \notin A$ ou encore $non(f(x) \in A)$, qui équivaut à $non(x \in f^{-1}(A))$ équivaut à $x \in \mathrm{C}_E\, f^{-1}(A)$; d'où l'égalité $f^{-1}(\mathrm{C}_F\, A) = \mathrm{C}_E\, f^{-1}(A)$. □

2.3 RELATION D'ORDRE ET D'EQUIVALENCE

Une relation sur un ensemble E est un énoncé $\mathcal{R}(x, y)$ (ou $x\mathcal{R}y$) à deux variables : si $\mathcal{R}(x, y)$ est vrai on dira que x est relié à y par la relation $\mathcal{R}$. Les deux exemples les plus importants sont les relations d'ordre et d'équivalence. Une relation d'ordre établit une règle de comparaison entre tous ou certains des éléments : par exemple dans un dictionnaire les mots sont classés suivant une certaine loi, on peut classer les habitants d'un pays par ordre croissant d'âge. Une relation d'équivalence regroupe les éléments d'un ensemble par des propriétés mutuellement exclusives. Par exemple on peut regrouper ensemble les mots commençant par la même lettre, on peut séparer les habitants d'un pays d'après leur sexe, leur année de naissance etc...

2.3.1 Relation d'ordre.

Définition: Une *relation d'ordre* sur un ensemble E est une relation $\mathcal{R}$ telle que :

(i) (Réflexivité) Pour tout $x \in E$ on a $x \mathcal{R} x$.
(ii) (Transitivité) Si $x \mathcal{R} y$ et $y \mathcal{R} z$ alors $x \mathcal{R} z$
(iii) (Antisymétrie) Si $x \mathcal{R} y$ et $y \mathcal{R} x$ alors $x = y$.

Remarques : ces propriétés correspondent aux propriétés de la relation "être plus petit que, ou égal". En fait en mathématique la phrase "être plus petit que" doit presque toujours s'interpréter comme "être plus petit ou égal à". Si l'on veut ajouter que les éléments sont distincts on dira "être *strictement* plus petit".

Exemples : Les ensembles $\mathbf{N}$, $\mathbf{Z}$, $\mathbf{Q}$, $\mathbf{R}$ sont tous munis d'une relation d'ordre $\leq$ que l'on peut décrire par : $x \leq y$ si et seulement si $y - x$ est positif (ou nul). Si $x \leq y$ et $x \neq y$ on écrit $x < y$. La notion d'ordre permet de caractériser les intervalles :

Définition: Soit E l'un des ensembles $\mathbf{N}$, $\mathbf{Z}$, $\mathbf{Q}$ ou $\mathbf{R}$. Un *intervalle* est un sous-ensemble I tel que, si $x, y \in I$ et $x \leq z \leq y$ alors $z \in I$.

Par contre l'ensemble $\mathbf{C}$ n'a pas de relation d'ordre naturelle et la notion d'intervalle n'y a pas de sens ; on peut néanmoins définir par exemple un ordre lexicographique ainsi (rappelons que tout nombre complexe s'écrit de manière unique $x + iy$ avec x, y réels) :

décrètons que $x + iy \mathcal{R} x' + iy'$ si et seulement si $x < x'$ ou $x = x'$ et $y \leq y'$ (comme pour classer les mots dans un dictionnaire, on compare d'abord les premières lettres et, si elles sont égales on compare les secondes lettres et ainsi de suite).

La relation d'inclusion est aussi une relation d'ordre ; en effet on a bien $F \subset F$ pour tout ensemble F ; si $E \subset F$ et $F \subset G$ alors $E \subset G$ et enfin si $E \subset F$ et $F \subset E$ alors $E = F$.

Il y a une différence importante entre les premiers exemples et ce dernier : dans les premiers cas deux éléments sont toujours comparables ; parmi deux éléments l'un est plus petit que l'autre. Par contre deux ensembles ne sont pas nécessairement comparables : si $E := \{0, 1, 2, 3\}$ et $F := \{0, 1, 4\}$ alors on a $E \not\subset F$ et $F \not\subset E$. Ceci suggère la définition suivante: un ordre $\mathcal{R}$ sur un ensemble E est dit *total* (ou encore l'ensemble E *totalement ordonné*) si deux éléments sont toujours comparables i.e. si :

$$\forall x, y \in E,\ x \mathcal{R} y \text{ ou } y \mathcal{R} x$$

Par exemple : la relation sur l'ensemble $\mathbf{N}$ définie par $x \mid y$ si et seulement si x *divise* y (ou encore y est un multiple entier de x) est une relation d'ordre (vérification laissée en exercice) mais ce n'est pas un ordre total ; en effet 5 ne divise pas 6 et 6 ne divise pas 5.

Si $f : X \to Y$ est une application entre deux ensembles ordonnés par les relations $\leq$ il est naturel de se demander si f préserve l'ordre :

Définition: Une application f est *croissante* si pour tout x, y dans l'ensemble de départ de f, la relation $x \leq y$ entraîne $f(x) \leq f(y)$. Si $x \leq y$ entraîne $f(y) \leq f(x)$ on dit que f est *décroissante* (ou renverse l'ordre). On dit que f est *monotone* si elle est croissante ou décroissante.

Exemples : Les fonctions de $\mathbf{R}$ dans $\mathbf{R}$ donnée par $x \mapsto 2x - 3$ ou $x \mapsto \exp(x)$ sont croissantes, l'application donnée par $x \mapsto -4x + 1$ est décroissante, l'application donnée par $x \mapsto \sin(x)$ n'est pas monotone (sur $\mathbf{R}$).

Soit $f : E \to F$ une fonction, alors les applications $A \mapsto f(A)$ (de $\mathcal{P}(E)$ dans $\mathcal{P}(F)$) et $B \mapsto f^{-1}(B)$ (de $\mathcal{P}(F)$ dans $\mathcal{P}(E)$) sont toutes deux croissantes (si $\mathcal{P}(E)$ et $\mathcal{P}(F)$ sont ordonnés par l'inclusion).

L'importance pratique des fonctions croissantes (ou décroissantes) est qu'elles permettent de transformer des inégalités ; par exemple :

$$\frac{x-y}{3} \leq z^2 \Rightarrow \exp(\frac{x-y}{3}) \leq \exp(z^2) \Rightarrow -4\exp(\frac{x-y}{3}) + 1 \geq -4\exp x(z^2) + 1$$

Nous verrons que la méthode la plus puissante pour voir si une fonction (de $\mathbf{R}$ dans $\mathbf{R}$) est monotone est le calcul différentiel. En effet nous *démontrerons* au chapitre 14 qu'une fonction dérivable sur un intervalle est monotone si et seulement si sa dérivée est de signe constant (résultat admis en terminale).

2.3.2 Plus grand élément, borne supérieure.

Une des traductions les plus fréquentes d'un problème est la recherche d'un minimum ou d'un maximum : si l'on veut placer son argent on cherchera naturellement à le placer de manière à obtenir un rendement maximum ; pour se déplacer d'un point à un autre on cherche le chemin le plus court ; ayant construit (ou dessiné) un pont il est important de connaître le poids maximal qu'il peut supporter ; de nombreux problèmes en physique (ou chimie) peuvent se formuler ainsi : par exemple un rayon lumineux se réfléchit ou se réfracte en suivant un chemin minimal (principe de Fermat) ; un solide posé sur un plan horizontal restera en équilibre seulement si son centre de gravité est situé dans une position minimale.

Définition: Soit $(E, \leq)$ un ensemble ordonné, un élément y de E est le *plus grand élément* de E si tous les autres éléments sont plus petits, c'est-à-dire si $\forall x \in E, x \leq y$.

Remarque : il y a bien sûr une définition analogue du *plus petit élément*.

Exemples : Le plus petit élément de $\mathbf{N}$ est 0 mais $\mathbf{N}$ n'a pas de plus grand élément. Considérons la relation d'inclusion sur l'ensemble $\mathcal{P}(E)$; ce n'est pas un ensemble totalement ordonné mais il a un plus petit élément : l'ensemble vide $\emptyset$ et un plus grand élément : l'ensemble E.

Considérons maintenant un sous-ensemble F d'un ensemble ordonné E ; il est souvent intéressant de connaître un élément de E qui est plus grand que tous les éléments de F ; on peut aussi chercher un tel élément le plus petit possible. C'est le but des définitions suivantes :

Définition: Soit $F \subset E$ un sous-ensemble d'un ensemble ordonné, un élément M de E est un *majorant* de F si pour tout x dans F on a $x \leq M$. Le plus petit des majorants de F (s'il existe) s'appelle la *borne supérieure* de F (dans E).

On peut bien sûr définir de la même façon un *minorant* et la *borne inférieure* comme le plus grand des minorants.

Exemples : Soit $\mathbf{N} \subset \mathbf{R}$, tout nombre réel négatif est un minorant de $\mathbf{N}$ et sa borne inférieure est donc 0 (qui est aussi le plus petit élément de $\mathbf{N}$).

Soit $E := [0,1[\subset \mathbf{R}$ l'intervalle des nombres réels positifs et strictement plus petits que 1. Il est clair que 0 est la borne inférieure de E (et son plus petit élément) et que 1 est sa borne supérieure bien que E n'ait pas de plus grand élément.

Soit $F := \{x \in \mathbf{Q} \mid x < \sqrt{2}\}$ considéré comme sous-ensemble de $\mathbf{Q}$, alors F admet des majorants (par exemple 2 ou $\frac{3}{2}$) mais pas de borne supérieure. En effet, si elle existait, la borne supérieure m de F vérifierait $m^2 \leq 2$ et $2 \leq m^2$ donc $m^2 = 2$, mais ceci est impossible (voir par exemple le chapitre 5). Bien sûr le même ensemble F, considéré comme sous-ensemble de $\mathbf{R}$ admet $\sqrt{2}$ comme borne supérieure.

Une caractérisation commode de la borne supérieure d'un ensemble de réels est la suivante :

PROPOSITION: *Un réel M est la borne supérieure d'un ensemble $E \subset \mathbf{R}$ si et seulement si :*

(i) $\forall x \in E,\ x \leq M$

(ii) $\forall \varepsilon > 0,\ \exists x \in E,\ M - \varepsilon \leq x$

Démonstration: En effet la première propriété dit que M est un majorant et la seconde que c'est le plus petit des majorants : si m est un majorant de E on voit que $\forall \varepsilon > 0,\ M - \varepsilon \leq m$ et donc $M \leq m$. □

En fait nous verrons qu'une propriété très importante de $\mathbf{R}$ est que tout sous-ensemble (non vide) majoré admet une borne supérieure ; cette dernière propriété est fausse dans l'ensemble $\mathbf{Q}$..

2.3.3 Relation d'équivalence.

Définition: Une *relation d'équivalence* sur un ensemble E est une relation $\mathcal{R}$ telle que :

(i) (Réflexivité) Pour tout $x \in E$ on a $x \mathcal{R} x$.

(ii) (Transitivité) Si $x \mathcal{R} y$ et $y \mathcal{R} z$ alors $x \mathcal{R} z$

(iii) (Symétrie) Si $x \mathcal{R} y$ entraîne $y \mathcal{R} x$.

Exemples : Sur l'ensemble $\mathbf{N}$ la relation $x \mathcal{R} y$ si $x - y$ est pair définit une relation d'équivalence.

Définition: La *classe d'équivalence* d'un élément x est l'ensemble des éléments qui lui sont reliés par $\mathcal{R}$:

$$C(x) := \{y \in E \mid x \mathcal{R} y\}$$

Remarque : Les classes d'équivalence forment une partition de E, i.e. on peut écrire E comme union disjointe de classes d'équivalence. En effet si $x, x' \in E$ ou bien $C(x) \cap C(x') = \emptyset$ ou bien $C(x) = C(x')$ (si $y \in C(x) \cap C(x')$ alors $y\mathcal{R}x$ et $y\mathcal{R}x'$ entraîne $x\mathcal{R}x'$).

Définition: L'ensemble des classes d'équivalence de E pour la relation $\mathcal{R}$ s'appelle l'*ensemble quotient* de E par $\mathcal{R}$ et se note $E/\mathcal{R}$.

Commentaire : il s'agit d'une notion délicate qui permet de nombreuses constructions : l'ensemble $\mathbf{Z}$ est construit à partir de $\mathbf{N}$ comme le quotient de $\mathbf{N} \times \mathbf{N}$ par la relation

d'équivalence $(x,y)\mathcal{R}(x',y') \Leftrightarrow x+y'=x'+y$ et l'ensemble $\mathbf{Q}$ est construit à partir de $\mathbf{Z}$ comme le quotient de $\mathbf{Z}\times(\mathbf{Z}\setminus\{0\})$ par la relation d'équivalence $(x,y)\mathcal{R}(x',y') \Leftrightarrow xy'=x'y$.

Le seul exemple d'ensemble quotient que nous approfondirons est le suivant : Soit n un entier ≥ 1, considérons la relation d'équivalence suivante sur $\mathbf{Z}$:

$$x\mathcal{R}_n y \Leftrightarrow \ n \text{ divise } x-y$$

Cette relation s'appelle *relation de congruence modulo n* et se note souvent (Cf chapitre 5) $x \equiv y \mod n$. L'ensemble quotient est un ensemble à n éléments : les classes de $0,1,\ldots,n-1$ et se note d'habitude $\mathbf{Z}/n\mathbf{Z}$.

2.4 CARDINAUX ET ENTIERS NATURELS

La notion de cardinal est probablement la première notion mathématique abstraite : il y a quelque chose de commun à trois carottes et trois étoiles, c'est le nombre de ces objets. L'idée de nombre entier est en fait issue de cette intuition. Avec les notions introduites précédemment on voit que deux ensembles ont même cardinal si on peut les mettre en bijection. Ainsi un nombre entier –un cardinal– apparaît comme une classe d'équivalence d'ensembles. Ces définitions qui peuvent paraître pédantes (et le sont) quand on parle de cardinaux finis deviennent indispensables pour aborder les cardinaux infinis, c'est-à-dire pour parler du "nombre" d'éléments d'un ensemble infini.

2.4.1 Ensembles et cardinaux finis

Définition: Deux ensembles ont même *cardinal* si il existe une bijection entre les deux ensembles. On dit aussi qu'ils sont *équipotents*.

Il s'agit bien de la traduction mathématique de "avoir le même nombre d'éléments" ; mais cette intuition correspond en fait au cas des ensembles finis et pour les ensembles infinis, la définition mathématique est la seule qui permette de raisonner.

Cardinaux finis : un entier naturel est le cardinal d'un ensemble fini. Par exemple $\mathrm{card}(\emptyset)=0$, $\mathrm{card}(\{\emptyset\})=1$, $\mathrm{card}(\{\emptyset,\{\emptyset\}\})=2$ etc... De manière plus parlante, si a,b,c,d sont des éléments distincts $\mathrm{card}\{a\}=1$, $\mathrm{card}\{a,b\}=2$ et $\mathrm{card}\{a,b,c,d\}=4$.

Un ensemble est donc fini et de cardinal n si et seulement si il est en bijection avec l'ensemble $\{0,1,\ldots,n-1\}$.

Une propriété très importante des ensembles finis (qui en fait les caractérise) est la suivante :

THÉORÈME: *Soient E et F des ensembles finis de même cardinal ; soit f une application de E dans F alors les propriétés suivantes sont équivalentes :*

(i) L'application f est bijective.

(ii) L'application f est injective.

(iii) L'application f est surjective.

Démonstration: Appelons n le cardinal de E. Pour que f soit surjective il faut et il suffit que $f(E)$ ait n éléments, mais $\mathrm{card}(f(E)) \leq \mathrm{card}(E)$ avec égalité si et seulement si

f est injective. On conclut que (ii) équivaut à (iii) ; par ailleurs (i) équivaut par définition à (ii) *et* (iii) d'où le théorème. □

Remarque : On voit en particulier que si $x \in E$ et E est fini alors $E \setminus \{x\}$ n'est pas en bijection avec E. Ceci est une caractérisation des ensembles finis.

Une autre application simple des ensembles finis est le *principe des tiroirs* ; "si on range $n+1$ chaussettes dans n tiroirs, l'un (au moins) des tiroirs contiendra deux chaussettes" (on laisse la preuve en exercice).

Il est naturel, sachant qu'une ensemble est fini de chercher à déterminer son cardinal (un entier naturel). On appelle combinatoire cette partie des mathématiques. Voici quelques résultats utiles.

THÉORÈME: *Soient E et F des ensembles finis de cardinaux m et n respectivement, alors :*

(i) $\mathrm{card}(E) + \mathrm{card}(F) = \mathrm{card}(E \cap F) + \mathrm{card}(E \cup F)$

(ii) $\mathrm{card}(E \times F) = mn$

(iii) Soit $\mathcal{F}(E,F)$ l'ensemble des applications de E vers F alors $card(\mathcal{F}(E,F)) = n^m$. En particulier $\mathrm{card}(\mathcal{P}(E)) = 2^m$.

(iv) Le nombre d'injection de E dans F est 0 si $m > n$ et $n(n-1)(n-2)\dots(n-m+1)$ si $m \leq n$.

(v) L'ensemble des bijections de F vers F a pour cardinal $n! = n(n-1)(n-2)\dots 2.1$

Démonstration: (i) Commençons par observer que dans le cas plus facile où $E \cap F = \emptyset$, la formule est évidente ; en effet si $X = A \cup B$ avec $A \cap B = \emptyset$ alors $\mathrm{card}(X) = \mathrm{card}(A) + \mathrm{card}(B)$. Revenons au cas général et posons $E' := E \setminus (E \cap F)$, alors $E \cup F$ est union disjointe de F et E' donc $\mathrm{card}(E \cup F) = \mathrm{card}(F) + \mathrm{card}(E')$. Mais E est union disjointe de E' et $E \cap F$ donc on a aussi : $\mathrm{card}(E) = \mathrm{card}(E') + \mathrm{card}(E \cap F)$ et on tire de ces deux égalités la formule : $\mathrm{card}(E) + \mathrm{card}(F) = \mathrm{card}(E \cap F) + \mathrm{card}(E \cup F)$

(ii) On peut écrire $E \times F = \cup_{x \in E} \{x\} \times F$; or ces ensembles sont disjoints donc on a $\mathrm{card}(E \times F) = \sum_{x \in E} \mathrm{card}(\{x\} \times F)$. Mais F est en bijection avec chacun des ensembles $\{x\} \times F$ par l'application $y \mapsto (x,y)$ donc $\mathrm{card}(\{x\} \times F) = n$ et $\mathrm{card}(E \times F) = \sum_{x \in E} n = mn$.

(iii) Pour construire une fonction de $E = \{a_1, a_2, \dots, a_m\}$ vers F il faut choisir $f(a_1)$ (il y a n choix possibles), $f(a_2)$ (il y a n choix possibles),... etc. Il y a donc $n \times n \dots n = n^m$ fonctions de E vers F.

Soit A un sous-ensemble de E, on lui associe la fonction $f_A : E \to \{0,1\}$ définie par $f_A(x) = 1$ si $x \in A$ et $f_A(x) = 0$ si $x \notin A$ (la fonction f_A s'appelle la *fonction caractéristique* de A). On obtient ainsi une bijection entre $\mathcal{P}(E)$ et $\mathcal{F}(E, \{0,1\})$ (la bijection réciproque est donnée par $f \mapsto \{x \in E \mid f(x) = 1\}$). On conclut que $\mathrm{card}(\mathcal{P}(E)) = \mathrm{card}(\{0,1\})^m = 2^m$.

(iv) Tout d'abord, il est clair que si $\mathrm{card}(E) > \mathrm{card}(F)$ il n'existe aucune injection de E dans F. Si maintenant $E = \{a_1, a_2, \dots, a_m\}$ et $m \leq n$, pour construire une application injective de E dans F on doit choisir $f(a_1) \in F$ (il y a n choix possibles) puis $f(a_2) \in F \setminus \{f(a_1)\}$ (il y a $n-1$ choix possibles) puis $f(a_3) \in F \setminus \{f(a_1), f(a_2)\}$ (il y a $n-2$ choix possibles) et ainsi de suite. On obtient donc bien en tout $n(n-1)(n-2)\dots(n-m+1)$ injections.

(v) Si $E = F$ est fini on sait qu'une fonction de E dans F est bijective si et seulement si elle est injective donc d'après le résultat précédent il y a $n(n-1)(n-2)\dots(n-n+1) = n!$ bijections. □

Introduisons maintenant une notation très utile en combinatoire :

Définition: Soit F un ensemble de cardinal n et soit $0 \leq p \leq n$, le nombre de parties de F ayant p éléments se note C_n^p ou $\binom{n}{p}$.

THÉORÈME: *On a les formules suivantes :*

(i) $C_n^p = \frac{n(n-1)(n-2)\dots(n-p+1)}{p!} = \frac{n!}{p!(n-p)!}$

(ii) $C_n^p = C_n^{n-p}$

(iii) $C_n^p = C_{n-1}^p + C_{n-1}^{p-1}$

Démonstration: (i) Un sous-ensemble à p éléments de F est donné à permutation près de ses éléments (il y a $p!$ permutations d'après le théorème précédent) par une injection de $\{0, 1, 2, \dots, p\}$ dans F ; il y a $n(n-1)(n-2)\dots(n-p+1)$ injections et donc $\frac{n(n-1)(n-2)\dots(n-p+1)}{p!}$ parties à p éléments.

Démontrons maintenant les deux propriétés (ii) et (iii). On peut bien sûr démontrer ces formules en utilisant la formule $C_n^p = \frac{n!}{p!(n-p)!}$ (vérifiez-le à titre d'exercice) mais nous trouvons plus instructive une démonstration en termes d'ensembles à partir de la définition des C_n^p.

(ii) Soit E un ensemble de cardinal n. L'application $A \mapsto \complement_E A$ définit une bijection entre l'ensemble des parties de E à p éléments et l'ensemble des parties de E à $n-p$ éléments, d'où la formule (ii).

(iii) Soit E un ensemble de cardinal n et soit $x \in E$. L'ensemble des parties de E à p éléments se répartit en deux sous-ensembles disjoints : l'ensemble des parties à p éléments de E contenant l'élément x et l'ensemble des parties à p éléments de E ne contenant pas l'élément x. Le premier est en bijection avec l'ensemble des parties à $p-1$ éléments de $E \setminus \{x\}$ qui a pour cardinal C_{n-1}^{p-1}, et le second est en bijection avec l'ensemble des parties à p éléments de $E \setminus \{x\}$ qui a pour cardinal C_{n-1}^p, d'où le résultat cherché. □

Remarque : Si l'on écrit dans un tableau les coefficients C_n^p (où n sera le numéro de la ligne et p le numéro de la colonne), les propriétés (i) et (ii) se traduisent par la symétrie de chaque ligne et en observant que chaque coefficient est la somme de deux coefficients de la ligne précédente : celui situé juste au-dessus et son prédécesseur. Ces remarques permettent d'ailleurs de calculer très facilement les premiers coefficients. Ce tableau s'appelle le *triangle de Pascal* (bien qu'il ait été connu par exemple des mathématiciens arabes avant sa redécouverte par Pascal).

Les coefficients C_n^p pour $0 \leq p \leq n \leq 7$:

$$\begin{array}{cccccccc}
1 & & & & & & & \\
1 & 1 & & & & & & \\
1 & 2 & 1 & & & & & \\
1 & 3 & 3 & 1 & & & & \\
1 & 4 & 6 & 4 & 1 & & & \\
1 & 5 & 10 & 10 & 5 & 1 & & \\
1 & 6 & 15 & 20 & 15 & 6 & 1 & \\
1 & 7 & 21 & 35 & 35 & 21 & 7 & 1
\end{array}$$

2.4.2 Ensembles infinis, $\mathbf{N}$ et principe de récurrence.

Nous ne donnerons pas de construction de l'ensemble $\mathbf{N}$ *bien que celle-ci puisse se faire dans le cadre de la théorie des ensembles. Il faut pour cela introduire l'axiome d'existence d'un ensemble infini. Quelque soit la présentation, l'ensemble des entiers naturels est le premier ensemble infini qu'on rencontre. Il peut être caractérisé par l'existence d'un élément initial (zéro) et pour chaque élément* n *d'un successeur* $n+1$ *(distinct de* $0, 1, \ldots, n$*) et pour chaque élément différent de zéro d'un prédécesseur ainsi que par le principe de récurrence.*

Nous supposons connu donc l'ensemble :

$$\mathbf{N} := \{0, 1, 2, 3, 4, 5, \ldots\}$$

Il est muni d'une loi d'addition et de multiplication et d'un ordre ; une loi moins évidente qui le caractérise essentiellement est la suivante :

THÉORÈME: *(principe de récurrence) Soit* S *un sous-ensemble de* $\mathbf{N}$ *contenant* 0 *et tel que :*

$$\forall n \in \mathbf{N},\ n \in S \Rightarrow (n+1) \in S$$

alors $S = \mathbf{N}$.

L'utilité du théorème est de permettre de vérifier une propriété $\mathcal{P}(n)$ pour tout entier naturel n en montrant que $\mathcal{P}(n) \Rightarrow \mathcal{P}(n+1)$ et en vérifiant $\mathcal{P}(0)$.

Exemple : Démontrons que

$$\sum_{i=0}^{n} i = \frac{n(n+1)}{2}$$

Pour cela appelons $\mathcal{P}(n)$ cette formule et $S := \{n \in \mathbf{N} \mid \mathcal{P}(n)\}$. On voit tout de suite que $\mathcal{P}(0)$ est vrai car $0 = 0$; supposons donc $\mathcal{P}(n)$ vrai et démontrons donc $\mathcal{P}(n+1)$ *à partir de* $\mathcal{P}(n)$: $\sum_{i=0}^{n+1} i = \sum_{i=0}^{n} i + (n+1)$ qui d'après $\mathcal{P}(n)$ vaut $\frac{n(n+1)}{2} + (n+1) = \frac{(n+1)(n+2)}{2}$ soit donc : $\sum_{i=0}^{n+1} i = \frac{(n+1)(n+2)}{2}$ ce qui est bien $\mathcal{P}(n+1)$. Le théorème permet de conclure que $S = \mathbf{N}$ ce qui signifie bien que pour tout entier n la formule $\mathcal{P}(n)$ est vraie.

Exercice : démontrer de la même manière les formules suivantes :

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

$$\sum_{i=0}^{n} i^3 = \left(\frac{n(n+1)}{2}\right)^2$$

Pouvez-vous trouver (et prouver) une formule semblable pour $\sum_{i=0}^{n} i^4$?

THÉORÈME: *L'ensemble* $\mathbf{N}$ *est infini.*

Démonstration: Le contraire serait surprenant, mais donnons néanmoins la démonstration complète. Considérons l'ensemble $\mathbf{N}^* := \mathbf{N} \setminus \{0\}$ et l'application de $\mathbf{N}$ vers $\mathbf{N}^*$ définie par $n \mapsto n+1$. C'est une bijection (vérification facile) mais nous avons vu qu'un ensemble fini ne peut pas être en bijection avec "lui-même moins un élément" donc $\mathbf{N}$ est bien infini. □

La théorie des ensembles permet de construire à partir de $\mathbf{N}$ les ensembles $\mathbf{Q}$, $\mathbf{R}$ et $\mathbf{C}$. Nous ne développerons pas ces constructions mais signalons qu'il y a beaucoup plus de nombres réels que de nombres entiers ou rationnels et en particulier qu'il existe "plusieurs infinis".

Définition: Un ensemble X est *dénombrable* s'il existe une injection de X dans $\mathbf{N}$. Il revient au même de dire que X est en bijection avec un sous-ensemble de $\mathbf{N}$.

THÉORÈME: (Cantor) *L'ensemble* $\mathbf{Q}$ *est dénombrable. L'ensemble* $\mathbf{R}$ *n'est pas dénombrable.*

Démonstration: Si $\mathbf{R}$ était dénombrable, l'intervalle $[0,1]$ le serait également. On pourrait donc écrire $[0,1] = \{x_1, x_2, \ldots, x_n, \ldots\}$. Notons $x_n = 0, a_1^{(n)} a_2^{(n)} \ldots a_m^{(n)} \ldots$ le développement décimal de x_n. Pour chaque $n \geq 1$, on peut choisir un chiffre b_n tel que $b_n \neq a_n^{(n)}$ et fabriquer le nombre réel $x := 0, b_1 b_2 \ldots b_m \ldots$. On voit alors immédiatement que, pour tout n, on a $x \neq x_n$, ce qui contredit l'hypothèse initiale. □

Un peu d'histoire :

Le théorème de Cantor affirme donc qu'il y a "beaucoup plus" de nombres réels que de nombres rationnels, en d'autre termes il n'existe pas de bijection entre $\mathbf{Q}$ *et* $\mathbf{R}$*. Introduisons une définition : un nombre réel est dit algébrique s'il est racine d'un polynôme à coefficients dans* $\mathbf{Q}$ *(ainsi* $1+\sqrt{5}$*,* $\sqrt[3]{4+\sqrt[5]{2}}$ *sont des nombres algébriques) ; il est dit transcendant s'il n'est pas algébrique. L'existence de nombres transcendants n'est pas évidente et historiquement ils ont été découverts dans l'ordre suivant :*

Liouville montre en 1844 qu'il existe des nombres transcendants ; par exemple les nombres du type $0,10\ldots010\ldots010\ldots$ *où, à chaque fois, la suite de zéros est beaucoup plus longue que la précédente, sont transcendants.*

Hermite prouve en 1873 que le nombre e (base du logarithme népérien) est transcendant. Il est très difficile de démontrer qu'un nombre donné est transcendant et c'est le premier nombre "naturel" pour lequel cela a été démontré.

Cantor établit en 1874 que "presque tous" les nombres sont transcendants. En effet l'ensemble des nombres algébriques a le même cardinal que $\mathbf{Q}$ *(ou* $\mathbf{N}$*).*

Lindemann montre en 1882, en adaptant la méthode de Hermite, que π *est transcendant. Ce résultat achève de démontrer l'impossibilité de la quadrature du cercle.*

Pascal Blaise (1623–1662)

CHAPITRE 3 GROUPES, STRUCTURES ALGÉBRIQUES

La formalisation des structures algébriques –groupes, anneaux, corps, espaces vectoriels– est relativement récente mais l'idée est présente partout dans les sciences et en particulier en mathématique. Il s'agit grosso modo d'extraire des règles opératoires, valables indépendemment de la nature des objets considérés. Par exemple les règles pour faire la somme de deux nombres, la somme de deux vecteurs du plan ou la composition de deux rotations sont les mêmes. L'idée sous-jacente à la notion de groupe est celle de la symétrie ; c'est pourquoi nous choisissons d'étudier dans une première partie les symétries de quelques figures simples avant d'introduire formellement la définition de groupe.

3.1 SYMÉTRIES ET GROUPES.

Considérons une figure simple comme un rectangle (avec sa largeur différente de sa longueur) :

On distingue deux axes de symétrie : l'axe horizontal L_1 et l'axe vertical L_2 ; on voit qu'on peut aussi appliquer le rectangle sur lui-même en le faisant pivoter d'un demi-tour autour du point O (on peut aussi interpréter cela par une symétrie par rapport au point O). On admettra que ce sont les seules transformations (avec l'identité!) qui appliquent le rectangle sur lui-même en respectant les formes.

On vérifie sans peine les faits suivants :

1) Appliquer deux fois la même transformation revient à appliquer l'identité

2) Appliquer la symétrie s_1 par rapport à L_1 puis la symétrie s_2 par rapport à L_2 revient à appliquer la symétrie s_O par rapport à O ; en fait appliquer deux de ces trois symétries revient à appliquer la troisième (l'ordre étant indifférent).

On peut regrouper cela dans un tableau où l'on inscrit dans la ligne de l'élément s et la colonne de l'élément t la composée $s \circ t$:

$\circ$	id	s_O	s_1	s_2
id	id	s_O	s_1	s_2
s_O	s_O	id	s_2	s_1
s_1	s_1	s_2	id	s_O
s_2	s_2	s_1	s_O	id

Considérons maintenant un carré :

Les transformations qui appliquent le carré sur lui-même, en respectant les formes, sont maintenant :

Les symétries par rapport à l'axe horizontal L_1 et à l'axe vertical L_2 (que nous noterons s_1 et s_2), les symétries par rapport à la diagonale D_1 et à la diagonale D_2 (que nous noterons s_3 et s_4), les rotations autour du point O faisant un quart de tour (que nous noterons r_1), un demi-tour (que nous noterons r_2), trois quarts de tour (que nous noterons r_3), et enfin bien sûr l'identité.

On vérifiera que : 1) Appliquer deux fois la même symétrie ou la rotation d'un demi-tour revient à appliquer l'identité ; mais appliquer deux fois la même rotation d'un quart ou trois quarts de tour revient à appliquer la rotation d'un demi-tour. Toutefois appliquer quatre fois la même rotation d'un quart ou trois quarts de tour revient à appliquer l'identité.

2) Appliquer la symétrie par rapport à L_1 puis la symétrie par rapport à L_2 revient à appliquer la rotation d'un demi-tour ; en fait appliquer deux des trois symétries revient à appliquer une des rotations, appliquer une des rotations et une des symétries revient à appliquer une des symétries. Toutefois, l'ordre n'est pas cette fois indifférent : par exemple $s_1 s_3 = r_3 \neq r_1 = s_3 s_1$.

On peut regrouper cela dans un tableau où l'on inscrit dans la ligne de l'élément s et la colonne de l'élément t la composée $s \circ t$:

$\circ$	id	r_1	r_2	r_3	s_1	s_2	s_3	s_4
id	id	r_1	r_2	r_3	s_1	s_2	s_3	s_4
r_1	r_1	r_2	r_3	id	s_3	s_4	s_2	s_1
r_2	r_2	r_3	id	r_1	s_2	s_1	s_4	s_3
r_3	r_3	id	r_1	r_2	s_4	s_3	s_1	s_2
s_1	s_1	s_4	s_2	s_3	id	r_2	r_3	r_4
s_2	s_2	s_3	s_1	s_4	r_2	id	r_1	r_2
s_3	s_3	s_1	s_4	s_2	r_1	r_3	id	r_2
s_4	s_4	s_2	s_3	s_1	r_3	r_1	r_2	id

Observons expérimentalement quelques faits : tous les éléments apparaissent une et une seule fois dans chaque ligne et colonne ; dans le premier tableau, l'ordre dans lequel on compose des éléments n'importe pas ; dans le second tableau, l'ordre est important, mais une chose est préservée : si on veut faire le produit : $s \circ t \circ u$ alors on sait qu'il n'est pas nécessaire de "mettre les parenthèses", c'est-à-dire que $(s \circ t) \circ u = s \circ (t \circ u)$.

Nous venons de décortiquer l'archétype d'un groupe ; de manière générale :

L'ensemble des transformations préservant une figure forme un groupe.

Pour voir l'intérêt de définitions plus abstraites, essayez de donner une description des 48 transformations préservant un cube.

3.2 GROUPES, EXEMPLES

Définition: Une *loi de composition* sur un ensemble E est une application de $E \times E$ vers E.

Exemples : La plupart des opérations usuelles sont des lois de composition : l'addition ou la multiplication sont des lois de composition sur $\mathbf{N}$,$\mathbf{Z}$,$\mathbf{Q}$,$\mathbf{R}$ ou $\mathbf{C}$; la soustraction définit une loi de composition sur $\mathbf{Z}$,$\mathbf{Q}$,$\mathbf{R}$ ou $\mathbf{C}$ (mais pas sur $\mathbf{N}$) ; l'application de $\mathcal{F}(E,E) \times \mathcal{F}(E,E)$ vers $\mathcal{F}(E,E)$ définie par $(f,g) \mapsto f \circ g$ est aussi une loi de composition.

Définition: Un *groupe* est la donnée d'un ensemble G et d'une loi de composition $(x, y) \mapsto x * y$ telle que :

(i) (élément neutre) Il existe e dans G tel que pour tout x dans G on a $e*x = x*e = x$.

(ii) (associativité) Pour tout x, y, z dans G on a : $(x * y) * z = x * (y * z)$.

(iii) (élément inverse) Pour tout x dans G il existe x' dans G tel que : $x*x' = x'*x = e$.

Si de plus pour tout x, y dans G on a : $x*y = y*x$, on dit que la loi $*$ est *commutative* et que le groupe $(G, *)$ est *commutatif*.

Convention : pour calculer dans un groupe, on omettra souvent le signe $*$ et on écrira gh au lieu de $g * h$.

Exemples : 1) L'ensemble des transformations du rectangle (respectivement du carré) avec la loi de composition naturelle forme un groupe de cardinal 4 (respectivement 8). Le premier groupe est commutatif, le second ne l'est pas.

2) Les ensembles $\mathbf{Z}$,$\mathbf{Q}$,$\mathbf{R}$ et $\mathbf{C}$, munis de l'addition sont des groupes (noter que $(\mathbf{N}, +)$ ne vérifie pas (iii)). Les ensembles $\mathbf{Q}^*$,$\mathbf{R}^*$ ou $\mathbf{C}^*$ munis de la mutiplication sont des groupes (noter que $(\mathbf{Z} \setminus \{0\}, \times)$ ne vérifie pas (iii)). Tous ces groupes sont commutatifs.

3) Soit E un ensemble et soit $\mathcal{S}(E)$ l'ensemble des bijections de E vers E ; soit $\circ$ la loi de composition naturelle de deux bijections, alors $(\mathcal{S}(E), \circ)$ est un groupe. En particulier l'ensemble des bijections de $\{1, 2, 3, \ldots, n\}$ vers lui-même, muni de la composition des applications, forme un groupe qu'on note $\mathcal{S}_n$. C'est un groupe avec $n!$ éléments, on l'appelle le *groupe des permutations* sur n éléments.

Définition: Un *sous-groupe* H d'un groupe $(G, *)$ est un sous-ensemble de G tel que la loi $*$ restreinte à $H \times H$ définisse une loi interne qui donne une loi de groupe sur H.

Ainsi un sous-groupe est stable pour la loi $*$ (c'est-à-dire que si $x, y \in H$ alors $x * y \in H$), l'élément neutre e appartient à H et si $x \in H$ alors $x^{-1} \in H$. Remarquons qu'il est inutile de vérifier l'associativité : puisque $\forall x, y, z \in G,\ (xy)z = x(yz)$, il est clair qu'on a $\forall x, y, z \in H,\ (xy)z = x(yz)$. En fait on peut même raccourcir ces vérifications :

PROPOSITION: *Soit H un sous-ensemble d'un groupe G, c'est un sous-groupe si et seulement si il satisfait :*

(i) $e \in H$

(ii) $x, y \in H$ entraîne $xy^{-1} \in H$.

Démonstration: Ces conditions sont nécessaires. Réciproquement, supposons les propriétés (i) et (ii) vérifiées et montrons qu'alors H est un sous-groupe. Si $y \in H$ alors $ey^{-1} = y^{-1} \in H$; si x est également dans H alors $xy = x(y^{-1})^{-1} \in H$ donc H est bien un sous-groupe. □

Exemples :

1) L'ensemble μ_n des racines complexes de l'équation $X^n = 1$, muni de la multiplication des nombres complexes forme un sous-groupe de $\mathbf{C}^*$: en effet si $z, z' \in \mu_n$ alors $(z/z')^n = z^n/z'^n = 1$ donc $z/z' \in \mu_n$.

2) L'ensemble $n\mathbf{Z} := \{nx \mid x \in \mathbf{Z}\}$ muni de l'addition est un sous-groupe de $\mathbf{Z}$. Nous verrons au chapitre 5 que ce sont les seuls sous-groupes de $\mathbf{Z}$.

3) L'ensemble des rotations préservant le carré s'écrit en reprenant les notations du premier paragraphe $\{id, r_1, r_2, r_3\}$ et est un sous-groupe du groupe des transformations préservant le carré.

4) les inclusions suivantes sont des inclusions de sous-groupes : $\mathbf{Z} \subset \mathbf{Q} \subset \mathbf{R} \subset \mathbf{C}$ (pour la loi d'addition) ; $\{+1, -1\} \subset \mathbf{Q}^* \subset \mathbf{R}^* \subset \mathbf{C}^*$ (pour la loi de multiplication). L'ensemble $\mathbf{R}_+^*$ (mais pas $\mathbf{R}_-^*$) est un sous-groupe de $\mathbf{R}$; le cercle $\{z \in \mathbf{C} \mid |z| = 1\}$ est un sous-groupe de $\mathbf{C}^*$.

Définition: Un *homomorphisme de groupe* est une application $f : (G, *) \to (H, \circ)$ telle que :

$$\forall x, y \in G,\ f(x * y) = f(x) \circ f(y)$$

Si de plus f est une bijection, on dit que f est un *isomorphisme* de groupe et que G et H sont *isomorphes*.

Exemples :

1) Considérons l'application $x \mapsto x^n$. C'est un homomorphisme de $\mathbf{Q}^*$ dans $\mathbf{Q}^*$ (resp. $\mathbf{R}^*$, resp. $\mathbf{C}^*$). Cette application donne un isomorphisme de groupe de $\mathbf{R}_+^*$ dans $\mathbf{R}_+^*$ (en effet tout réel positif possède une unique racine n-ème positive, voir chapitre 4).

2) Soit G le groupe des transformations du carré ; soit $E := \{A, B, C, D\}$ l'ensemble des sommets du carré et H l'ensemble des bijections de E dans E. Toute transformation du carré, préservant les formes, doit envoyer un sommet sur un sommet et donne donc une bijection de E sur E. L'application qui à un élément $s \in G$ associe sa restriction à E est un homomorphisme de groupes de G vers H.

3) Soit G un groupe et g un élément de ce groupe, définissons par récurrence $g^0 := e$ et $g^{n+1} := gg^n$ (pour $n \in \mathbf{N}$) et enfin $g^{-n} := (g^n)^{-1}$. L'application $n \mapsto g^n$ de $\mathbf{Z}$ vers G est un homomorphisme de groupes, c'est-à-dire que $g^{m+n} = g^m g^n$. Remarquons que si G est fini alors cette application n'est pas injective et il existe donc un plus petit entier positif et non nul d tel que $g^d = e$.

Définition: Le plus petit entier $d \geq 1$ tel que $g^d = e$, s'il existe, s'appelle l'*ordre* de g, s'il n'existe pas on dit que g est d'ordre infini.

Par exemple, l'élément 2 est d'ordre infini dans $\mathbf{Q}^*$ alors que -1 est d'ordre 2 dans le même groupe.

Nous avons vu qu'il est important de savoir si une application est injective ou surjective. Dans le cas d'homomorphismes de groupes il existe un critère simple qui nécessite les définitions suivantes :

Définition: Le *noyau* d'un homomorphisme de groupe $f : G \to H$ est l'ensemble $f^{-1}(\{e_H\}) = \{g \in G \mid f(g) = e_H\}$. On le note $Ker(f)$ (à cause de l'allemand "Kern").

L'importance du noyau vient du théorème suivant :

THÉORÈME: Un homomorphisme de groupe $f : G \to H$ est injectif si et seulement si $Ker(f) = \{e_G\}$. Le noyau de f est toujours un sous-groupe de G.

Démonstration: En effet $f(x) = f(y)$ équivaut à $f(x)f(y)^{-1} = e_H$ ou encore $f(xy^{-1}) = e_H$, ce qui signifie $xy^{-1} \in Ker(f)$. Si $Ker(f) = \{e_G\}$ on voit que $f(x) = f(y)$

entraîne $xy^{-1} = e$ ou encore $x = y$ donc f est injective. Si $Ker(f)$ contient un élément $g \neq e_G$ alors $f(g) = f(e_G) = e_H$ et f n'est pas injective.

La deuxième affirmation est facile : si $x, y \in Ker(f)$ alors $f(xy^{-1}) = f(x)f(y^{-1}) = f(x)f(y)^{-1} = ee^{-1} = e$ donc $xy^{-1} \in Ker(f)$. $\square$

Dans le paragraphe suivant nous étudions toutes les notions définies ici, sur l'exemple du groupe des permutations sur n éléments.

3.3 LE GROUPE $\mathcal{S}_n$.

Un élément s de $\mathcal{S}_n$ est une permutation de l'ensemble $\{1, 2, 3, \ldots, n\}$ et est donc défini par la suite $s(1), s(2), s(3), \ldots, s(n)$. On doit aussi se souvenir que si $i \neq j$ alors $s(i) \neq s(j)$. L'élément neutre sera noté id. On notera en général une permutation par un tableau :

$$s = \begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ s(1) & s(2) & s(3) & \ldots & s(n) \end{pmatrix}$$

Par exemple le groupe $\mathcal{S}_2$ possède 2 éléments : id et $t = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$; le groupe $\mathcal{S}_3$ possède 6 éléments : l'identité et les cinq permutations : $\tau_{23} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \tau_{12} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \tau_{13} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \rho_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$ et $\rho_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$ Le tableau de la loi de groupe de $\mathcal{S}_3$ est :

$\circ$	id	τ_{12}	τ_{23}	τ_{13}	ρ_1	ρ_2
id	id	τ_{12}	τ_{23}	τ_{13}	ρ_1	ρ_2
τ_{12}	τ_{12}	id	ρ_1	ρ_2	τ_{23}	τ_{13}
τ_{23}	τ_{23}	ρ_2	id	ρ_1	τ_{13}	τ_{12}
τ_{13}	τ_{13}	ρ_1	ρ_2	id	τ_{12}	τ_{23}
ρ_1	ρ_1	τ_{13}	τ_{12}	τ_{23}	ρ_2	id
ρ_2	ρ_2	τ_{23}	τ_{13}	τ_{12}	id	ρ_1

On voit en particulier que $\mathcal{S}_3$ n'est pas commutatif.

Sur ces deux exemples on peut facilement définir le *signe* d'une permutation : $\varepsilon(id) = +1$ et $\varepsilon(t) = -1$ pour $\mathcal{S}_2$ et ensuite $\varepsilon(id) = \varepsilon(\rho_1) = \varepsilon(\rho_2) = +1$ et $\varepsilon(\tau_{12}) = \varepsilon(\tau_{23}) = \varepsilon(\tau_{13}) = -1$ pour $\mathcal{S}_3$. On vérifie facilement que ε est un homomorphisme de groupes (à valeurs dans le groupe à deux éléments $\{+1, -1\}$).

Pour étudier les groupes $\mathcal{S}_n$, commençons par y définir des éléments particulièrement simples.

Définition: Un *m-cycle* ou *cycle de longueur m* dans $\mathcal{S}_n$ est une permutation s de l'ensemble $E := \{1, \ldots, n\}$ qui laisse fixes $n - m$ éléments et permute circulairement les autres. Plus précisément, il existe un sous-ensemble à m éléments $I = \{i_1, \ldots, i_m\}$ de E tel que : si $i \notin I$ alors $s(i) = i$ mais $s(i_k) = i_{k+1}$ (pour $k = 1, \ldots, m-1$) et $s(i_m) = i_1$. L'ensemble I s'appelle le *support* du cycle.

Une *transposition* est un cycle de longueur 2.

Nous noterons $s = (i_1, i_2, \ldots, i_m)$ le cycle décrit dans la définition. Une transposition ayant pour support $\{i, j\}$ sera aussi notée τ_{ij} (ce qui est cohérent avec la notation déjà utilisée pour les éléments de $\mathcal{S}_2$ et $\mathcal{S}_3$).

Exemple : l'élément $t \in \mathcal{S}_2$ est une transposition, tout comme $\tau_{12}, \tau_{13}, \tau_{23} \in \mathcal{S}_3$. Les éléments $\rho_1, \rho_2 \in \mathcal{S}_3$ sont des 3-cycles. Par contre la permutation $s = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}$ n'est pas un cycle. On peut vérifier que la permutation $s' = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 2 & 7 & 5 & 1 & 6 & 4 \end{pmatrix}$ est un cycle de longueur 5 et de support $\{1, 3, 7, 4, 5\}$, c'est-à-dire que $s' = (1, 3, 7, 4, 5)$.

THÉORÈME: *Toute permutation se décompose de manière unique (à l'ordre près) en produit de cycles dont les supports sont deux à deux disjoints.*

Démonstration: On utilise une récurrence sur l'entier n, l'affirmation étant claire pour $n \leq 3$ (puisque toutes les permutations sont alors des cycles). Supposons donc l'énoncé démontré pour les permutations de k éléments avec $k < n$ et considérons $s \in \mathcal{S}_n$. En regardant la suite $1, s(1), s^2(1) \ldots$ on voit qu'il existe un plus petit entier $m \geq 1$ tel que $s^m(1) = 1$ (on n'exclut pas que $m = 1$). Définissons l'ensemble $I := \{1, s(1), s^2(1) \ldots, s^{m-1}(1)\}$ et le m-cycle $r := (1, s(1), s^2(1) \ldots, s^{m-1}(1))$; alors la permutation $t := sr^{-1}$ laisse fixe les éléments de I et pour $i \notin I$ on a $t(i) = s(i)$. La restriction de t à $J := \{1, \ldots, n\} \setminus I$ est donc une permutation des éléments de J que nous notons s'. Comme $card(J) < n$ on sait (par l'hypothèse de récurrence) que $s' = s'_1 \ldots s'_r$ avec s'_i des cycles de J à supports disjoints. Définissons $s_i \in \mathcal{S}_n$ par $s_i(j) = s'_i(j)$ si $j \in J$ et $s_i(j) = j$ si $j \notin I$; on voit qu'alors on a $t = s_1 \ldots s_r$ et par conséquent $s = s_1 \ldots s_r r$. Ceci prouve l'existence de la décomposition en cycles ; pour l'unicité on observe que le cycle r est uniquement déterminé par s et que par hypothèse de récurrence $s'_1, \ldots, s'_r$ (et par conséquent $s_1, \ldots, s_r$) sont uniques. □

Voyons comment on obtient en pratique cette décomposition sur un exemple : Prenons la permutation $\rho = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 6 & 7 & 5 & 1 & 2 & 4 \end{pmatrix}$. On choisit un premier élément disons 1 et on calcule ses images successives par ρ : on a $\rho(1) = 3,\ \rho^2(1) = \rho(3) = 7,\ \rho^3(1) = \rho(7) = 4,\ \rho^4(1) = \rho(4) = 5$ et $\rho^5(1) = \rho(5) = 1$ et on obtient ainsi un premier cycle s' qui est le 5-cycle dans l'exemple précédant le théorème. On prend alors un autre élément qui n'est pas dans le suppport de s', par exemple 2 et on recommence : $\rho(2) = 6,\ \rho^2(2) = \rho(6) = 2$. on obtient ainsi la décomposition $\rho = s'\tau_{26}$.

Cette décomposition est très utile pour calculer l'ordre d'une permutation (si vous n'avez jamais vu la notion de PPCM – plus petit commun multiple– consultez le chapitre 5) :

PROPOSITION: *Soit s une permutation qui se décompose en le produit de r cycles à supports disjoints de longueurs $m_1, \ldots, m_r$, alors l'ordre de la permutation s est égal au* PPCM$(m_1, \ldots, m_r)$.

Démonstration: Démontrons d'abord que si la permutation s est un m-cycle, elle a pour ordre m : il suffit de le faire pour le cycle $s = (1, 2, \ldots, m)$. Or, si $i > m$ on a $s(i) = i$ et

donc $s^m(i) = i$; si maintenant $1 \leq i \leq m$ on a $s^m(i) = s^i(s^{m-i}(i)) = s^i(m) = s^{i-1}(1) = i$ donc au total $s^m = id$. Par ailleurs si $1 \leq k \leq m-1$ alors $s^k(1) = k+1 \neq 1$ donc $s^k \neq id$; ainsi l'ordre de s est bien m.

Dans le cas général où $s = s_1 \dots s_r$ avec s_i cycles de longueurs m_i à supports disjoints, notons $N := \text{PPCM}(m_1, \dots, m_r)$. Observons que, comme les s_i commutent, on a $s^k = s_1^k \dots s_r^k$ et que, d'après l'unicité de la décomposition en cycles on a $s^k = id$ si et seulement si $s_1^k = \dots = s_r^k = id$ donc si et seulement si l'ordre de s_i (c'est-à-dire m_i) divise k donc si et seulement si N divise k. □

Exemples : considérons les deux permutations suivantes dans $\mathcal{S}_{10}$:

$$s = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 4 & 10 & 6 & 8 & 9 & 3 & 1 & 7 & 2 & 5 \end{pmatrix}, \; t = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 2 & 3 & 4 & 5 & 1 & 7 & 8 & 6 & 10 & 9 \end{pmatrix}$$

alors les décompositions en cycles de s et t s'écrivent $s = (1,4,8,7)(2,10,5,9)(3,6)$ et $t = (1,2,3,4,5)(6,7,8)(9,10)$ et donc $ordre(s) = PPCM(4,4,2) = 4$ et $ordre(t) = PPCM(5,3,2) = 30$.

PROPOSITION: *Tout cycle peut s'écrire comme produit de transpositions et donc toute permutation peut s'écrire comme produit de transpositions.*

Démonstration: Quitte à changer de notation il suffit de montrer que le cycle $s = (1, 2, \dots, m)$ s'écrit comme produit de transpositions. Or considérons le produit $s' = \tau_{12}\tau_{23} \dots \tau_{i,i+1} \dots \tau_{m-1,m}$ on vérifie que $s'(m) = \tau_{12}\tau_{23} \dots \tau_{m-2,m-1}(m-1) = \dots = \tau_{12}\tau_{23} \dots \tau_{i,i+1}(i+1) = \dots = \tau_{12}(2) = 1$ et que si $i \leq m-1$ alors $s'(i) = \tau_{12}\tau_{23} \dots \tau_{i,i+1}(i) = \tau_{12}\tau_{23} \dots \tau_{i-1,i}(i+1) = i+1$ et finalement on a bien $s = s'$, ce qui achève la preuve. □

Remarque : la décomposition en produit de transpositions n'est pas du tout unique mais la parité du nombre de transposition ne change pas comme on pourra le vérifier à l'aide de la notion suivante.

Définition: Le *signe* d'une permutation $s \in \mathcal{S}_n$ est défini par le produit :

$$\varepsilon(s) = \prod_{1 \leq i < j \leq n} \frac{s(j) - s(i)}{j - i}$$

Il est aisé de vérifier que $\varepsilon(s) \in \{+1, -1\}$ et que le signe d'une transposition est -1 ; la principale propriété est la suivante :

PROPOSITION: *Le signe est un homomorphisme de $\mathcal{S}_n$ vers $\{+1, -1\}$. Son noyau (l'ensemble des permutations paires que l'on notera $\mathcal{A}_n$) est un sous-groupe de cardinal $\frac{n!}{2}$.*

Démonstration: Pour montrer la première propriété, on calcule le signe du produit de deux permutations s, t :

$$\varepsilon(st) = \prod_{1 \leq i < j \leq n} \frac{st(j) - st(i)}{j - i} = \prod_{1 \leq i < j \leq n} \left(\frac{st(j) - st(i)}{t(j) - t(i)} \right) \left(\frac{t(j) - t(i)}{j - i} \right) =$$

$$= \prod_{1 \leq i < j \leq n} \frac{s(j) - s(i)}{j - i} \prod_{1 \leq i < j \leq n} \frac{t(j) - t(i)}{j - i} = \varepsilon(s)\varepsilon(t)$$

Le signe d'une transposition τ est -1 ; considérons l'application $s \mapsto s\tau$. C'est une application de $\mathcal{A}_n$ vers $\mathcal{S}_n \setminus \mathcal{A}_n$ qui est injective (car $s\tau = s'\tau$ entraîne $s = s'$) et surjective (car $(s\tau)\tau = s$) donc bijective. Ainsi $n! = card(\mathcal{S}_n) = card(\mathcal{A}_n) + card(\mathcal{S}_n \setminus \mathcal{A}_n) = 2card(\mathcal{A}_n)$. □

Remarque : on voit donc $\varepsilon(s) = +1$ si s est le produit d'un nombre pair de transpositions et $\varepsilon(s) = -1$ si s est le produit d'un nombre impair de transpositions. Plus généralement un cycle de longueur m aura donc un signe $(-1)^{m+1}$, ce qui donne une méthode de calcul du signe d'une permutation connaissant sa décomposition en cycles.

3.4 STRUCTURE D'ANNEAU ET STRUCTURE DE CORPS.

Définition: Un *anneau* est la donnée d'un ensemble A et de deux lois de composition $+$ (addition) et $*$ (Multiplication) telles que :

(i) $(A, +)$ est un groupe commutatif (dont on note l'élément neutre $0 = 0_A$).

(ii) La loi $*$ est associative.

(iii) La loi $*$ possède un élément neutre (qu'on notera $1 = 1_A$)

(iv) La loi $*$ est distributive par rapport à l'addition :

$$\forall x, y, z \in A,\ x * (y + z) = (x * y) + (x * z) \ \text{ et } \ (y + z) * x = (y * x) + (z * x)$$

Si de plus la loi $*$ est commutative on dit que l'anneau A est commutatif.

Remarquons que l'on a toujours $x * 0 = 0 * x = 0$ dans un anneau ; en effet $x * 0 = x * (0 + 0) = x * 0 + x * 0$ et donc (la loi $+$ est une loi de groupe) $x * 0 = 0$.

Définition: Un *corps* est un anneau tel que :

(v) Tout élément $x \in A \setminus \{0_A\}$ possède un inverse.

Convention : Un anneau (ou un corps) est donc un triplet $(A, +, *)$, l'ensemble A s'appelle l'ensemble *sous-jacent* à l'anneau ; toutefois on parle souvent de l'anneau A en sous-entendant les lois $+$ et $*$ quand il est clair dans le contexte de quelles lois il s'agit.

Exemples : Nous étudierons tout spécialement l'anneau des entiers relatifs $(\mathbf{Z}, +, \times)$; ce n'est pas un corps car les seuls éléments de $\mathbf{Z}$ possédant un inverse pour la multiplication sont $+1$ et -1. Les corps les plus importants que nous étudierons sont le corps des nombres rationnels $\mathbf{Q}$, le corps des nombres réels $\mathbf{R}$ et le corps des nombres complexes $\mathbf{C}$. Un nombre rationnel peut bien sûr s'écrire comme une fraction $\frac{a}{b}$ avec $a \in \mathbf{Z}$ et $b \in \mathbf{Z} \setminus \{0\}$ avec la règle $\frac{a}{b} = \frac{a'}{b'}$ si $ab' = a'b$; l'addition et la multiplication sont définis par $\frac{a}{b} + \frac{c}{d} = \frac{ad+bc}{bd}$. Nous verrons aussi que, si K désigne $\mathbf{Q}$, $\mathbf{R}$ ou $\mathbf{C}$, l'ensemble des polynômes à coefficients dans K, que l'on note $K[X]$, muni de l'addition et de la multiplication naturelles, forme un anneau qui possède beaucoup de propriétés communes avec $\mathbf{Z}$. Tous ces anneaux sont commutatifs.

L'ensemble des matrices 2×2 à coefficients réels (voir chapitre 7) muni des lois :

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} = \begin{pmatrix} a + a' & b + b' \\ c + c' & d + d' \end{pmatrix}$$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} = \begin{pmatrix} aa' + bc' & ab' + bd' \\ ca' + dc' & cb' + dd' \end{pmatrix}$$

forme un anneau qui n'est pas commutatif ; par exemple :

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \neq \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

Règles de calcul dans un anneau :
(distibutivité généralisée) $x \sum_{i=1}^{n} y_i = \sum_{i=1}^{n} xy_i$
Attention : dans un anneau, il n'est pas vrai en général que lorsque $x \in A \setminus \{0\}$ on ait $xy = xz \Rightarrow y = z$; par exemple $\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 & 3 \\ 0 & 1 \end{pmatrix}$ mais $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \neq \begin{pmatrix} 2 & 3 \\ 0 & 1 \end{pmatrix}$
Si l'anneau est commutatif : $(xy)^n = x^n y^n$
L'expression de la puissance n-ème d'une somme est souvent utile :

THÉORÈME: *(Formule du binôme de Newton) Soient a, b deux éléments d'un anneau commutatif et soit n un entier ≥ 1, on a la formule :*

$$(a+b)^n = \sum_{p=0}^{n} C_n^p a^p b^{n-p}$$

où $C_n^p = \frac{n!}{p!(n-p)!}$ est le nombre de parties à p éléments dans un ensemble à n éléments.

A cause de cette formule, les coefficients C_n^p sont aussi appelés *coefficients binômiaux*. Les premiers exemples de cette formule s'écrivent :

$$(a+b)^1 = a+b,\ (a+b)^2 = a^2+2ab+b^2,\ (a+b)^3 = a^3+3a^2b+3ab^2+b^3$$

$$(a+b)^4 = a^4+4a^3b+6a^2b^2+4ab^3+b^4,\ (a+b)^5 = a^5+5a^4b+10a^3b^2+10a^2b^3+5ab^4+b^5$$

Démonstration: La démonstration se fait par récurrence sur le nombre n : la formule est évidente pour $n = 0$ ou $n = 1$, on la suppose donc vraie pour l'entier n, pour tout a, b et on cherche à en déduire la formule pour l'entier $n+1$.
On a : $(a+b)^{n+1} = (a+b)(a+b)^n$ qui d'après l'hypothèse de récurrence vaut :

$$(a+b)\sum_{p=0}^{n} C_n^p a^p b^{n-p} = \sum_{p=0}^{n} C_n^p a^{p+1} b^{n-p} + \sum_{p=0}^{n} C_n^p a^p b^{n-p+1},$$

cette dernière expression est égale à :

$$a^{n+1} + \sum_{h=1}^{n} (C_n^h + C_n^{h-1}) a^h b^{n+1-h} + b^{n+1}$$

et, si on se rappelle que $C_n^h + C_n^{h-1} = C_{n+1}^h$ celle-ci vaut :

$$\sum_{h=0}^{n+1} C_{n+1}^h a^h b^{n+1-h}$$

ce qui est bien la formule de Newton pour l'entier $n+1$. □

Remarque : L'hypothèse que l'anneau est commutatif ne peut pas être enlevée (dans un anneau non commutatif, en général $ba^p b^{n-p}$ n'est égal à $a^p b^{n+1-p}$ comme le montre l'exemple des matrices $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ et $B = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ puisque $(A+B)^2 = \begin{pmatrix} 3 & 6 \\ 3 & 6 \end{pmatrix}$ mais $A^2 + 2AB + B^2 = \begin{pmatrix} 4 & 7 \\ 1 & 5 \end{pmatrix}$.

Exercice : Le dessin suivant fournit une illustration de la formule $(a+b)^2 = a^2+2ab+b^2$ en décomposant un carré de côté $a+b$ en deux carrés de côtés a et b et deux rectangles de longueur b et largeur a.

Donner une illustration de la formule $(a+b)^3 = a^3+3a^2b+3ab^2+b^3$ en décomposant un cube de coté $a+b$ en deux cubes de côtés a et b, trois parallélipipèdes d'arêtes a, a et b et trois parallélipipèdes d'arêtes a, b et b.

Un peu d'histoire :

Les notions de groupes et corps ont tiré leur première illustration spectaculaire du problème de la "résolution des équations polynomiales". On connait depuis le lycée la résolution de $a+bx+cx^2=0$ à l'aide de la fonction racine carrée $\sqrt{}$; au XVIème siècle, Cardan (également inventeur du système d'articulation mécanique portant son nom) a donné des formules pour résoudre $a+bx+cx^2+dx^3=0$ à l'aide des fonctions $\sqrt{}$ et $\sqrt[3]{}$; son élève Ferrari (aucun rapport connu avec Enzo) a ensuite donné des formules pour résoudre $a+bx+cx^2+dx^3+ex^4=0$ à l'aide des fonctions $\sqrt{}$, $\sqrt[3]{}$ et $\sqrt[4]{}$. Les mathématiciens ont longtemps cherché à résoudre ainsi les équations de degré ≥ 5 avant que Abel (1802-29) et Galois (1811-32) ne montrent que cela est impossible. Par exemple les solutions de $x^5-x+1=0$ ne peuvent pas s'exprimer à l'aide de $\sqrt{}$, $\sqrt[3]{}$, $\sqrt[4]{}$ et $\sqrt[5]{}$. Ces propriétés des équations de degré 3,4,5, etc sont liées aux propriétés des groupes $\mathcal{S}_3, \mathcal{S}_4, \mathcal{S}_5$ etc. La théorie de Galois (à l'université Paris 7) s'étudie en maîtrise (M1) de mathématiques.

Galois Evariste (1811–1832)

CHAPITRE 4 LE CORPS DES RÉELS R ET DES COMPLEXES C

Les nombres "réels" ont été ainsi baptisés car on pensait que ce sont ceux qui permettaient de décrire les phénomènes physiques. Il est vrai que tout le calcul différentiel, et donc toute la mécanique classique repose sur la notion de nombre réel (même si cela n'est pas explicite chez Newton et Leibniz). Les nombres réels ont donc été utilisés très tôt bien que la démonstration de leurs propriétés et surtout de leur existence (du point de vue mathématique!) date du siècle dernier. Nous n'aborderons donc pas cet aspect et renvoyons aux traités classiques pour une description de $\mathbf{R}$ *par les coupures de Dedekind ou les classes d'équivalence de suites de Cauchy (Voir par exemple l'ouvrage de Dixmier cité en bibliographie). Quant aux nombres complexes, même les mathématiciens ont mis longtemps à accepter leur emploi (ils se sont longtemps appelés nombres imaginaires tant leur existence était sujette à doute). Néanmoins ils sont assez faciles à construire à partir des nombres réels et s'avèrent aussi utiles que les réels, y compris dans les autres sciences comme la physique.*

4.1 NOMBRES RÉELS.

La nécessité de considérer des nombres plus généraux que les nombres rationnels apparaît déjà avec l'absence de solution à l'équation $x^2 = 2$*, plus généralement l'existence de suite de nombres rationnels (ou de points d'une droite) "ayant l'air de converger" vers un point mais ne convergeant pas vers un nombre rationnel (ou un point commensurable) conduit à l'introduction des nombres réels que nous définirons ici de manière axiomatique, i.e. sans démontrer leur existence. Nous introduisons aussi la notion de limite –déjà abordée en terminale– qui est fondamentale dans toute l'analyse : les nombres réels permettent de nombreux procédés "infinitésimaux" ou de "passage à la limite". Ceci nous permet aussi de traiter précisément et rigoureusement le développement décimal des nombres réels : il est classique de représenter un nombre réel sous forme de développement décimal* $x = \pm a_0, a_1 a_2 a_3 \dots a_n \dots$ *avec* $a_0 \in \mathbf{N}$ et $a_i \in \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$*. Par exemple :*

$$\pi = 3,14159265358979323846264338327950288419716939937510582097494459230 78\dots$$

Mais si on cherche à définir un nombre réel comme une telle suite on trouve quelques difficultés ; considérons par exemple le "nombre" $x := 0,99999\dots9\dots$*, il est raisonnable de penser que* $10x = 9,99999\dots9\dots$ *et aussi que* $10x - x = 9$ *et donc* $x = 1$ *; la multiplication est assez difficile à définir sur les développements décimaux.*

Une notion fondamentale sur les réels est celle d'ordre ; l'ensemble des réels est muni d'une addition et d'une multiplication qui en font un corps ; la relation d'ordre pour être utile doit être compatible avec ces opérations, plus précisément elle doit vérifier les règles suivantes :

(i) Pour tous x, y, z réels, $x \leq y \Rightarrow x + z \leq y + z$

(ii) Pour tous x, y réels, pour tout a réel positif $x \leq y \Rightarrow ax \leq ay$

On peut aussi en déduire :

(iii) $0 < x \leq y \Rightarrow 0 < \frac{1}{y} \leq \frac{1}{x}$

(iv) $x \leq y \Rightarrow -x \geq -y$ et $\forall x,\ x^2 \geq 0$

Un corps satisfaisant ces règles est appelé un *corps ordonné*.

A ces règles il faut rajouter une propriété qui formalise une intuition :

Définition: Un corps ordonné est dit *archimédien* si pour tout $x > 0$ et $y > 0$ il existe un entier $n \geq 1$ tel que $nx = x + \ldots + x > y$ (ici 0 désigne l'élément neutre).

Autrement dit, une quantité, aussi petite soit-elle, ajoutée suffisamment de fois à elle-même dépasse n'importe quelle quantité donnée. Par exemple le groupe $(\mathbf{Z}, +)$ est bien sûr archimédien, de même que $(\mathbf{Q}, +)$; les réels forment aussi un corps archimédien :

CARACTÉRISATION : *Le corps $(\mathbf{R}, +, \times, \leq)$ contient le corps des rationnels, est un corps totalement ordonné archimédien et vérifie la propriété dite des intervalles emboités :*

Soit $I_n = [a_n, b_n]$ une suite décroissante d'intervalles fermés bornés non vides alors $\bigcap_{n\in\mathbf{N}} I_n$ est non vide (c'est-à-dire : il existe $x \in \mathbf{R}$ tel que pour tout n on ait $x \in I_n$).

Un élément de ce corps s'appelle un nombre réel.

Ainsi, $\mathbf{R}$ est caractérisé par le fait d'être un corps (il y a une addition et une multiplication avec les "bonnes" propriétés) d'être totalement ordonné (ce qui le différencie de $\mathbf{C}$), archimédien et enfin la dernière propriété le différencie de $\mathbf{Q}$.

La représentation la plus usuelle des réels est celle des points d'une droite, nous la supposons connue.

La relation d'ordre permet aussi de définir la distance entre deux réels et donc de dire si deux réels sont proches :

Définition: La *valeur absolue* d'un nombre réel x est $\max\{x, -x\}$ et se note $|x|$. La distance entre deux réels x et y est $|x - y|$.

La valeur absolue d'un nombre est donc toujours positive. Rappelons les deux propriétés bien connues et fondamentales de la valeur absolue :

THÉORÈME: *(i)* $|xy| = |x||y|$

(ii) (inégalité triangulaire) $|x + y| \leq |x| + |y|$

Démonstration: Laissée en exercice (ou voir les cours au lycée). □

La deuxième inégalité s'appelle triangulaire (bien qu'il n'y ait pas ici de vrai triangle : les points sont situés sur une droite) ; en effet, si l'on désigne par $d(x, y)$ la distance entre deux nombres réels x et y, on peut aussi exprimer l'inégalité (ii) sous la forme $d(a, c) \leq d(a, b) + d(b, c)$

Remarque : l'inégalité $|x - a| \leq b$ équivaut à $a - b \leq x \leq a + b$. Ainsi les ensembles du type $\{x \in \mathbf{R} \mid |x - a| \leq b\}$ (respectivement $\{x \in \mathbf{R} \mid |x - a| < b\}$) sont des intervalles fermés (respectivement ouvert) aux deux extrémités. Inversement un intervalle $[a, b]$ peut aussi s'écrire $[a, b] = \{x \in \mathbf{R} \mid |x - \frac{a+b}{2}| \leq \frac{b-a}{2}\}$.

La notion de distance permet de formaliser l'idée de "tendre vers un point". Intuitivement une suite u_n tend vers $\ell \in \mathbf{R}$ si u_n est de plus en plus proche de ℓ quand n augmente

ou encore si u_n se retrouve dans n'importe quel intervalle autour de ℓ, aussi petit soit-il, dès que n est assez grand.

Définition: Une suite u_n de nombres réels (ou rationnels) *tend vers* 0 si elle vérifie :

$$\forall \varepsilon > 0, \exists n_0 \in \mathbf{N}, n \geq n_0 \Rightarrow |u_n| \leq \varepsilon$$

Une suite u_n de nombres réels (ou rationnels) *tend vers* ℓ si $u_n - \ell$ tend vers 0. On dit aussi que u_n *converge* vers ℓ ou que ℓ est la *limite* de la suite u_n, ce que l'on note $\lim u_n = \ell$.

Autrement dit : soit n'importe quel (petit) intervalle centré en ℓ, alors tous les termes de la suite, sauf un nombre fini sont situés dans l'intervalle.

Exemples : La suite $u_n = \frac{1}{n+1}$ a pour limite $\ell = 0$; montrons cela directement à partir de la définition. Soit $\varepsilon > 0$, le corps $\mathbf{R}$ étant archimédien, il existe un entier n_0 plus grand que $1/\varepsilon$; soit alors $n \geq n_0$ alors $0 < 1/n \leq 1/n_0 \leq \varepsilon$ donc $|u_n| \leq \varepsilon$. Par contre, la suite $u_n = (-1)^n$ ne converge pas (si ℓ est différent de ± 1 un intervalle suffisamment petit centré en ℓ ne contient aucun terme u_n et si $\ell = \pm 1$, un nombre infini de termes éviteront un petit intervalle centré en ℓ).

THÉORÈME: *Soit u_n une suite convergente vers une limite ℓ, supposons que pour tout n on ait $u_n > a$ (respectivement $u_n \geq a$) alors $\ell \geq a$.*

Démonstration: Raisonnons par l'absurde et supposons que $\ell < a$. Choisissons un intervalle I contenant ℓ mais pas a (par exemple $I = \{x \in \mathbf{R} \mid |x - \ell| \leq \frac{a-\ell}{2}\}$) alors les éléments de la suite u_n sont dans I (sauf un nombre fini d'entre eux) mais pour tous les éléments x de I on a $x < a$ d'où une contradiction. □

Remarque : Si $u_n := \frac{1}{n+1}$ on a $u_n > 0$ mais $\lim u_n = 0$; on ne peut donc pas garder les inégalités strictes en passant à la limite.

Exploitons maintenant la propriété des intervalles emboités :

THÉORÈME: *(i) Tout sous-ensemble de $\mathbf{R}$ non vide et majoré admet une borne supérieure. Tout sous-ensemble de $\mathbf{R}$ non vide et minoré admet une borne inférieure.*

(ii) Toute suite croissante et majorée (respectivement décroissante et minorée) est convergente.

(Ce résultat est très important mais on peut omettre la démonstration assez technique)

Démonstration: (i) Soit E un ensemble non vide majoré de réels on va construire des intervalles emboités $I_n = [a_n, b_n]$ tels que l'intersection contienne au plus un point (et donc exactement un point) qui sera la borne supérieure. Soit $e \in E$ et M un majorant de E, on pose $I_0 := [e, M]$. Pour construire I_1 on distingue deux cas : si $\frac{M+e}{2}$ est un majorant de E on choisit $a_1 = e$ et $b_1 = \frac{M+e}{2}$; sinon il existe dans E un élément qui est plus grand que $\frac{M+e}{2}$ et on choisit a_1 égal à cet élément et $a_2 = M$. En itérant ce procédé on obtient une suite décroissante d'intervalles $I_n = [a_n, b_n]$ tels que b_n soit un majorant de E, tel que a_n

soit un élément de E et tel que $|b_{n+1} - a_{n+1}| \leq \frac{|a_n - b_n|}{2}$ donc $|a_n - b_n| \leq \frac{(M-e)}{2^n}$. Montrons maintenant qu'il ne peut y avoir qu'un seul point dans l'ensemble $S := \bigcap_{n\in\mathbf{N}} I_n$ et que c'est la borne supérieure. Tout d'abord soit $s, t \in S$ alors ces deux nombres appartiennent aussi I_n donc, pour tout n on a $|s-t| \leq \frac{(M-e)}{2^n}$ donc $|s-t| = 0$ et $s = t$. Par construction la suite des a_n comme celle des b_n converge vers s. Comme tous les b_n sont des majorants de E, s est aussi un majorant de E ; comme tous les a_n sont des éléments de E, on a que s est le plus petit majorant.

(ii) Considérons $E = \{u_n \mid n \in \mathbf{N}\}$, c'est un ensemble majoré par hypothèse donc il admet une borne supérieure ℓ. Montrons que u_n converge vers ℓ. Soit $\varepsilon > 0$, la définition de la borne supérieure entraîne qu'il existe un élément de E, disons u_{n_0} tel que $\ell - \varepsilon \leq u_{n_0} \leq \ell$; mais alors comme u_n est croissante on a pour tout $n \geq n_0$ les inégalités $\ell - \varepsilon \leq u_{n_0} \leq u_n \leq \ell$ et donc $|u_n - \ell| \leq \varepsilon$, ce qui prouve bien que u_n tend vers ℓ.
□

Notation : on sait que si $x \in \mathbf{R}$ alors il existe un unique entier relatif m tel que $m \leq x < m+1$ on l'appelle la *partie entière* de x et on le note $[x]$. Par exemple $[\pi] = 3$ et $[-3/2] = -2$.

APPLICATION: DÉVELOPPEMENT DÉCIMAL D'UN NOMBRE RÉEL.

On appelle bien sûr *chiffre* un élément de $C := \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ (on pourrait d'ailleurs faire les mêmes raisonnements dans une autre base que 10). Considérons une suite $a_1, a_2, a_3, \ldots$ de chiffres et associons lui la suite de nombres rationnels

$$s_n := \frac{a_1}{10} + \frac{a_2}{10^2} + \ldots + \frac{a_n}{10^n}$$

qu'on notera aussi $s_n := 0, a_1 a_2 \ldots a_n$.

1ère étape : la suite s_n converge vers un réel x appartenant à l'intervalle $[0, 1]$.

Démonstration: En effet la suite s_n est croissante (car $s_{n+1} = s_n + a_{n+1} 10^{-n-1} \geq s_n$) et majorée par 1 :

$$s_n \leq 9\left(\frac{1}{10} + \frac{1}{10^2} + \ldots + \frac{1}{10^n}\right) = 1 - \frac{1}{10^n} \leq 1$$

enfin comme $0 \leq s_n \leq 1$ on a bien $0 \leq x = \lim s_n \leq 1$. □

On introduit naturellement la notation : $x = 0, a_1 a_2 a_3 \ldots a_n \ldots$ et on appelle cette écriture *un développement décimal* de x. Deux questions se posent naturellement :

1) Est-ce-que tout nombre réel admet un développement décimal? Autrement dit tout nombre $x \in [0, 1]$ est-il limite d'une suite s_n?

2) Un tel développement est-il unique?

(en remarquant que $x - [x] \in [0, 1[$, on peut se borner à considérer les réels dans l'intervalle $[0, 1[$).

2ème étape : Tout nombre réel $x \in [0, 1[$ admet un développement décimal $x = 0, a_1 a_2 a_3 \ldots a_n \ldots$.

Démonstration: Fabriquons la suite $a_1 := [10x]$, $a_2 := [10^2x - 10a_1] \ldots a_n := [10^n x - 10^{n-1}a_1 - \ldots - 10a_{n-1}]$ et ensuite $s_n := \frac{a_1}{10} + \frac{a_2}{10^2} + \ldots + \frac{a_n}{10^n}$. Comme $0 \leq x < 1$ on a $0 \leq 10x < 10$ donc $a_1 \leq 10x < a_1 + 1$ donc $a_1 \leq 9$ et l'entier a_1 est bien un chiffre. Par ailleurs $s_1 = \frac{a_1}{10} \leq x < \frac{a_1}{10} + \frac{1}{10}$ donc $0 \leq x - s_1 < \frac{1}{10}$. Montrons par récurrence que $a_n \leq 9$ (i.e. l'entier a_n est un chiffre) et $0 \leq x - s_n < \frac{1}{10^n}$; ce qui prouvera que $x = 0, a_1a_2a_3 \ldots a_n \ldots$. Si $0 \leq x - s_n < \frac{1}{10^n}$ alors $0 \leq 10^{n+1}x - 10^{n+1}s_n = 10^{n+1}x - 10^n a_1 - \ldots - 10a_n < 10$ donc $0 \leq a_{n+1} < 10$ et donc a_{n+1} est bien un chiffre. Ensuite $a_{n+1} \leq 10^{n+1}x - 10^n a_1 - \ldots - 10a_n < a_{n+1} + 1$ donc $0 \leq x - \frac{a_1}{10} - \ldots \frac{a_n}{10^n} - \frac{a_{n+1}}{10^{n+1}} < 10^{-n-1}$; ce qu'il fallait démontrer. □

3ème étape : Le développement décimal $x = 0, a_1a_2a_3 \ldots a_n \ldots$ existe et est unique si l'on impose la condition :

$$\forall N \in \mathbf{N},\ \exists n > N,\ a_n \neq 9$$

Autrement dit on exclut les développements du type $0, a_1a_2 \ldots a_n 9999 \ldots 9 \ldots$ (avec $a_n \neq 9$) que l'on remplace par $0, a_1a_2 \ldots (a_n + 1)0 \ldots$

Par exemple 0,1234567899999. . . =0,12345679

Démonstration: Supposons $x = 0, a_1a_2a_3 \ldots a_n \ldots = 0, b_1b_2b_3 \ldots b_n \ldots$ et disons $a_1 = b_1, \ldots, a_{r-1} = b_{r-1}$ mais $a_r < b_r$. On obtient facilement $(b_r - a_r)10^{-r} = 0, 0 \ldots 0a_{r+1} \ldots - 0, 0 \ldots 0 \ldots b_{r+1} \ldots$. Le membre de gauche vaut au moins 10^{-r} car $b_r - a_r \geq 1$ mais

$$0, 0 \ldots 0a_{r+1} \ldots = \frac{a_{r+1}}{10^{r+1}} + \ldots + \frac{a_n}{10^n} \ldots < \frac{9}{10^{r+1}} + \ldots + \frac{9}{10^n} \ldots = 10^{-r}$$

L'inégalité est stricte car il existe des $a_n < 9$ par hypothèse ; on obtient donc une contradiction du type $10^{-r} < 10^{-r}$.
□

Remarque : on peut observer que les seuls nombres réels qui admettent "deux" développements sont exactement les nombres rationnels "décimaux" $x = \frac{m}{10^n}$

APPLICATION: RACINE n-IÈME D'UN RÉEL POSITIF

Soit $a \in \mathbf{R}_+$ et n un entier ≥ 1 alors il existe un unique $x \in \mathbf{R}_+$ tel que $x^n = a$. On l'appelle la *racine n-ième* de a et on le note $x = \sqrt[n]{a}$.

Démonstration: L'unicité est facile car si $0 < x < x'$ alors $0 < x^n < x'^n$. Pour montrer l'existence, considérons $S := \{y \in \mathbf{R}_+ \mid y^n \geq a\}$ alors S est non vide et minoré (par exemple par 0) donc possède une borne inférieure que nous baptisons x. Comme pour tout $y \in S$ on a $y^n \geq a$ on en déduit $x^n \geq a$. Si on avait $x^n < a$ alors pour $e > 0$ (mais très petit) on en déduirait $(x + e)^n < a$ (on donne une démonstration de ce fait ci-dessous) et donc $x + e \notin S$. Mais alors S ne contient aucun point de l'intervalle $[x, x + e]$ ce qui contredit le fait que x est la borne inférieure de S.

Il nous reste à montrer la "continuité" de la fonction $y \mapsto y^n$, c'est-à-dire à montrer que si y est très proche de x alors y^n est très proche de x^n. Nous verrons au chapitre 13 une méthode générale pour démontrer cela ; donnons néanmoins une démonstration directe (où l'on pourra supposer que $x > 0$).

Vérifions par récurrence que pour $0 \leq h \leq \frac{x}{2}$ on a $(x+h)^n \leq x^n + (2^n-1)hx^{n-1}$ en effet $(x+h)^{n+1} = (x+h)(x+h)^n \leq (x+h)(x^n+(2^n-1)hx^{n-1}) = x^{n+1} + hx^n(2^n + (2^n-1)\frac{h}{x}) \leq x^{n+1} + (2^{n+1}-1)hx^n$ d'où la propriété annoncée. On en déduit que si $x \leq y \leq x + \frac{\varepsilon}{2^n x^{n-1}}$ alors $x^n \leq y^n \leq x^n + \varepsilon$; ce qu'il fallait démontrer. □

4.2 NOMBRES COMPLEXES.

La nécessité d'étendre le corps des réels se fait sentir si on cherche à résoudre des équations comme $x^2 + 1 = 0$. Si on ajoute formellement un "nombre" i tel que $i^2 + 1 = 0$ alors on peut déjà résoudre les équations de degré 2 ; en effet pour étudier $ax^2 + bx + c = 0$ on introduit $\Delta := b^2 - 4ac$ et si $\Delta \geq 0$ les racines sont $\frac{-b \pm \sqrt{\Delta}}{2}$ alors que si $\Delta < 0$ il n'y a pas de racines réelles mais on peut "fabriquer" des racines par la formule $\frac{-b \pm i\sqrt{-\Delta}}{2}$. Ceci suggère d'étudier les "nombres" de la forme $x + iy$; il est clair ce que doivent être la somme et le produit de tels expressions ; nous prendrons ce guide pour définir les nombres complexes.

Définition: Un *nombre complexe* s'écrit $z = x + iy$ avec $x, y \in \mathbf{R}$; l'ensemble des nombres complexes se note $\mathbf{C}$ et est en bijection avec $\mathbf{R}^2 = \mathbf{R} \times \mathbf{R}$.

Définition: Soient $z = x + iy$ et $z' = x' + iy'$ deux nombres complexes.

On appelle partie *réelle* (respectivement *imaginaire*) de $z = x + iy$ le nombre réel x (respectivement le nombre y).

On définit la somme de deux nombres complexes par :

$$z + z' := (x + x') + i(y + y')$$

On définit le produit de deux nombres complexes par :

$$zz' := (xx' - yy') + i(yx' + xy')$$

Le *conjugué* de z est $\bar{z} := x - iy$. Le *module* de z est $|z| := \sqrt{x^2 + y^2} = \sqrt{z\bar{z}}$.

Remarque : on peut considérer un nombre réel x comme un nombre complexe en l'écrivant $x = x + i0$; un nombre réel est égal à son conjugué, la somme et le produit de deux nombres réels coïncident avec leur somme et produit comme nombres complexes, le module d'un nombre réel est sa valeur absolue.

THÉORÈME: *(i) L'ensemble $\mathbf{C}$ muni de la somme et de le multiplication est un corps commutatif. L'inverse d'un nombre complexe non nul $z = x + iy$ est donné par*

$$z^{-1} = \frac{\bar{z}}{|z|^2} = \frac{x}{x^2 + y^2} - i\frac{y}{x^2 + y^2}$$

(ii) La conjugaison complexe $z \mapsto \bar{z}$ est un isomorphisme de corps, c'est-à-dire que $\bar{1} = 1$, $\overline{x + y} = \bar{x} + \bar{y}$ et $\overline{xy} = \bar{x}\bar{y}$. La conjugaison est *involutive*, c'est-à-dire que $\bar{\bar{x}} = x$.

Démonstration: La vérification des axiomes d'un anneau ne pose aucune dificulté et est laissée au lecteur. Vérifions l'existence d'un inverse pour tout nombre complexe non

nul. Soit $z = x + iy \in \mathbf{C}^*$. Comme $|z|^2 = z\bar{z} = x^2 + y^2 \in \mathbf{R}^*$ on peut définir $z' := z/|z|^2$ et clairement $zz' = 1$. La deuxième partie de l'énoncé se vérifie par un calcul direct. □

Exemples : on vérifiera (en appliquant directement la définition) que :

$$(1+i)^2 = 2i,\ (1+i\sqrt{3})^3 = -8,\ \frac{(1+2i)}{(2+3i)} = \frac{8+i}{13},\ \left(\frac{-1+i\sqrt{3}}{2}\right)^2 + \left(\frac{-1+i\sqrt{3}}{2}\right) + 1 = 0$$

Donnons maintenant des représentations géométriques des nombres complexes :
Représentation dans le plan

On utilise la bijection $\mathbf{C} \to \mathbf{R}^2$ donnée par $z \mapsto (Re(z), Im(z))$ et on représente le nombre complexe z par le point $M = M(z)$ d'abscisse $Re(z)$ et d'ordonnée $Im(z)$. Le module $|z|$ est la distance entre O et M.

On peut définir la distance comme pour les nombres réels par $d(z, z') := |z - z'|$, on a alors :

THÉORÈME: *(i)* $|zz'| = |z||z'|$.
(ii) (inégalité triangulaire) Pour tous nombres complexes z, z' *on a* $|z + z'| \leq |z| + |z'|$.

Démonstration: (i) est immédiat car $|zz'|^2 = zz'\overline{zz'} = z\bar{z}z'\bar{z}' = |z|^2|z'|^2$. La preuve de (ii) est plus subtile : considérons la fonction de variable réelle $P(t) := |z + tz'|^2 = |z|^2 + t(z'\bar{z} + z\bar{z}') + t^2|z'|^2$; c'est un polynôme du second degré avec au plus une racine (aucune racine si z'/z n'est pas réel) donc $\Delta := (z'\bar{z} + z\bar{z}')^2 - 4|z|^2|z'|^2 \leq 0$ ou encore $|(z'\bar{z} + z\bar{z}')| \leq 2|z||z'|$. Nantis de cette inégalité, développons :

$$|z + z'|^2 = |z|^2 + z'\bar{z} + z\bar{z}' + |z'|^2 \leq |z|^2 + 2|z||z'| + |z'|^2 = (|z| + |z'|)^2$$

Ce qui donne bien l'inégalité cherchée. □

Remarque : Cette fois, l'inégalité (ii) peut se traduire par $d(M, M') \leq d(M, M'') + d(M'', M')$ qui est l'inégalité sur un triangle : la somme des longueurs de deux des côtés est plus grande que la longueur du troisième côté.

Ayant la notion de distance, on peut définir quand un point est proche d'un autre en particulier la notion de limite (on répète ici la définition par commodité) :

Définition: Une suite z_n de nombres complexes *tend vers* 0 si elle vérifie :

$$\forall \varepsilon > 0, \exists n_0 \in \mathbf{N}, n \geq n_0 \Rightarrow |z_n| \leq \varepsilon$$

Une suite z_n de nombres complexes *tend vers* $\ell \in \mathbf{C}$ si $z_n - \ell$ tend vers 0. On dit aussi que z_n *converge* vers ℓ ou que ℓ est la *limite* de la suite z_n, ce que l'on note $\lim u_n = \ell$.

Autrement dit : soit n'importe quel (petit) disque de centre ℓ, alors tous les termes de la suite, sauf un nombre fini sont situés dans le disque.

THÉORÈME: *Soit z_n une suite de nombres complexes ;* $\lim z_n = z$ *équivaut à* $\lim Re(z_n) = Re(z)$ *et* $\lim Im(z_n) = Im(z)$.

Démonstration: En remplaçant z_n par $z_n - z$ il suffit de prouver que $\lim z_n = 0$ si et seulement si $\lim Re(z_n) = 0$ et $\lim Im(z_n) = 0$. Mais comme $|Re(z_n)| \leq |z_n|$, $|Im(z_n)| \leq |z_n|$ et $|z_n| \leq |Re(z_n)| + |Im(z_n)|$ ceci est clair. □

Exemple : Si $|\alpha| < 1$ alors la suite $z_n := \alpha^n$ converge vers 0 car $|z_n|$ converge vers 0. Cependant si $\alpha = e^{i\pi\sqrt{2}}$ la suite α^n ne converge pas, bien que $\lim |\alpha^n| = 1$.

On admet ici l'existence des fonctions sinus et cosinus telles que si l'angle θ sur la figure ci-dessous est donné en radians (un tour complet vaut 2π, un demi-tour π, un quart de tour $\frac{\pi}{2}$) alors $OA = \cos(\theta)$ et $OB = \sin(\theta)$. Il y a là une difficulté qui sera levée en deuxième année après l'étude de fonctions analytiques.

On voit donc que tout nombre complexe peut s'exprimer comme :

$$z = r(\cos(\theta) + i\sin(\theta))$$

avec $r = |z| \in \mathbf{R}_+$ et $\theta \in \mathbf{R}$. Ou encore : si $z = a + ib \neq 0$ avec a, b réels, alors il existe un "angle" (i.e. un réel) θ tel que $\cos(\theta) = a/\sqrt{a^2+b^2}$ et $\sin(\theta) = b/\sqrt{a^2+b^2}$. Le nombre θ n'est déterminé qu'à un multiple entier de 2π près, il s'appelle l'argument de z et se note $Arg(z)$ (si on veut être tout-à-fait rigoureux, on doit dire *un* argument). Plus précisément :

THÉORÈME: *(i) Supposons $r(\cos(\theta)+i\sin(\theta)) = r'(\cos(\theta')+i\sin(\theta'))$ avec $r, r' \in \mathbf{R}_+^*$ alors $r = r'$ et il existe $n \in \mathbf{Z}$ tel que $\theta = \theta' + 2\pi n$.*
(ii) $|\bar{z}| = |z|$, $Arg(zz') = Arg(z) + Arg(z') + 2\pi n$ et $Arg(\bar{z}) = -Arg(z) + 2\pi n$.

Démonstration: (i) En prenant les modules on arrive à $|r| = |r'|$ et comme r et r' sont positifs on a bien $r = r'$. On en tire $\cos(\theta) = \cos(\theta')$ et $\sin(\theta) = \sin(\theta')$ ce qui entraîne $\theta = \theta' + 2\pi n$. (ii) La formule donnant le module du conjugué est claire, celle donnant son argument découle de celle donnant l'argument d'un produit : $Arg(z\bar{z}) = Arg(z)+Arg(\bar{z})+$

$2k\pi$ doit être un multiple de 2π car $z\bar{z}$ est réel et positif. La formule donnant l'argument d'un produit se déduit des formules classiques $\cos(y+y') = \cos(y)\cos(y') - \sin(y)\sin(y')$ et $\sin(y+y') = \cos(y)\sin(y') + \sin(y)\cos(y')$; en effet si $z = r(\cos(y) + i\sin(y))$ et $z' = r'(\cos(y') + i\sin(y'))$ alors le produit zz' vaut :

$$\begin{aligned} zz' =& rr'\{(\cos(y)\cos(y') - \sin(y)\sin(y')) + i(\cos(y)\sin(y') + \sin(y)\cos(y'))\} \\ =& rr'\{\cos(y+y') + i\sin(y+y')\} \end{aligned}$$

d'où $Arg(zz') = y + y' + 2n\pi$. □

La meilleure façon de décrire les cordonnées polaires à travers les nombres complexes est d'introduire la fonction exponentielle d'une variable complexe :

Définition: On pose $e^{i\theta} := \cos(\theta) + i\sin(\theta)$ et plus généralement si $z = x + iy$:

$$e^z = e^{x+iy} := e^x\cos(y) + ie^x\sin(y)$$

Exemples :

$$e^{2\pi i} = 1,\ e^{\pi i} = -1,\ e^{\frac{2\pi i}{3}} = \frac{-1+i\sqrt{3}}{2},\ e^{\log 2 + \frac{\pi i}{2}} = 2i.$$

THÉORÈME: *(i) Tout nombre complexe z non nul peut s'écrire $z = e^{z'}$ pour un certain nombre complexe z'.*
(ii) $e^z = e^{z'}$ équivaut à $Re(z) = Re(z')$ et $Im(z) = Im(z') + 2\pi n$.
(iii) $e^{z+z'} = e^z e^{z'}$
(iv) $\overline{e^z} = e^{\bar{z}}$

Démonstration: (i) provient du fait que tout point du cercle $|z| = 1$ peut s'écrire $z = \cos(\theta) + i\sin(\theta)$ pour un $\theta \in \mathbf{R}$ et du fait que l'exponentielle réelle est surjective de $\mathbf{R}$ sur $\mathbf{R}_+^*$.

(ii) est une redite du théorème précédent.

(iii) On sait que $e^{x+x'} = e^x e^{x'}$ pour $x, x' \in \mathbf{R}$; il suffit donc de vérifier que $e^{i(y+y')} = e^{iy}e^{iy'}$ pour $y, y' \in \mathbf{R}$. Mais cette dernière égalité équivaut à la formule donnée pour l'argument d'un produit de deux nombres complexes.

(iv) $e^{x-iy} = e^x(\cos(-y) + i\sin(-y)) = e^x(\cos(x) - i\sin(x)) = \overline{e^{x+iy}}$. □

On peut utiliser cette représentation pour déterminer les racines n-ième d'un nombre complexe :

THÉORÈME: *Soit $z_0 \in \mathbf{C}^*$ et n un entier ≥ 1, alors il existe n nombres complexes tels que $z^n = z_0$.*

Plus explicitement si $z_0 = r_0 e^{i\theta}$ alors les n racines n-ième sont :

$$z = \sqrt[n]{r_0}e^{i\left(\frac{\theta}{n} + \frac{2k\pi}{n}\right)} \text{ avec } k \in \{0, 1, \ldots, n-1\}$$

Démonstration: Cherchons z sous la forme $re^{i\alpha}$; l'équation $z^n = z_0$ équivaut alors à $r^n e^{in\alpha} = r_0 e^{i\theta}$ ou encore à $r^n = r_0$ et $n\alpha = \theta + 2k\pi$ (avec $k \in \mathbf{Z}$), d'où l'énoncé. □

En particulier les racines n-ièmes de 1 s'appellent *racine de l'unité* ; elles forment les sommets d'un polygone régulier à n côtés :

Racines 5-ième de l'unité :

Ce dernier théorème est en fait un cas un peu particulier du célèbre théorème de D'Alembert-Gauss (il fut énoncé pour la première fois par D'Alembert mais démontré rigoureusement plus tard par Gauss) :

THÉORÈME: *(D'Alembert-Gauss). Soit $P(X)$ un polynôme à coefficients complexes, si P est non constant, alors il possède une racine, i.e. $\exists \alpha \in \mathbf{C},\ P(\alpha) = 0$.*

Tout polynôme P de degré $d \geq 1$ s'écrit :

$$P(X) = a_0(X - \alpha_1)(X - \alpha_2)\dots(X - \alpha_d)$$

avec $a_0 \in \mathbf{C}^*$ et $\alpha_1, \alpha_2, \dots, \alpha_d \in \mathbf{C}$ (non nécessairement distincts).

Démonstration: Nous admettrons la première partie. Le fait que la première partie de l'énoncé entraîne la seconde est un résultat assez simple d'algèbre que nous démontrerons dans le chapitre 6 sur les polynômes. □

Une autre application classique de la représentation exponentielle est la formule de Moivre :

$$\cos(nx) + i\sin(nx) = (\cos(x) + i\sin(x))^n$$

où $x \in \mathbf{R}$ et $n \in \mathbf{Z}$.

Démonstration: On sait que $e^{inx} = (e^{ix})^n$ d'où la formule.

C'est un exercice classique, en utilisant la formule du binôme de Newton et la formule $\cos^2(x) + \sin^2(x) = 1$ d'en tirer une expression de $\cos(nx)$ et $\sin(nx)/\sin(x)$ comme polynôme en $\cos(x)$. Faisons-le pour $\cos(nx)$:

$(\cos(x) + i\sin(x))^n = \sum_{k=0}^{n} C_n^k (i\sin(x))^k (\cos(x))^{n-k}$ donc $\cos(nx)$ vaut :

$$Re\{(\cos(x) + i\sin(x))^n\} = \sum_{h=0}^{[\frac{n}{2}]} C_n^{2h} (-1)^h (\sin(x))^{2h} (\cos(x))^{n-2h} =$$

$$\sum_{h=0}^{[\frac{n}{2}]} C_n^{2h} (-1)^h (1 - \cos^2(x))^h (\cos(x))^{n-2h}$$

Ainsi $\cos(nx) = P_n(\cos(x))$ avec $P_n(X) = \sum_{h=0}^{[\frac{n}{2}]} C_n^{2h} (-1)^h (1 - X^2)^h X^{n-2h}$. Par exemple $P_2(X) = 2X^2 - 1$, $P_3(X) = 4X^3 - 3X$ et $P_4(X) = 8X^4 - 8X^2 + 1$.

Ces formules permettent aussi d'exprimer $\cos^n(x)$ comme combinaison linéaire de $\cos(nx)$, $\cos((n-2)x)$,... Par exemple :

$$\cos^2(x) = \frac{\cos(2x)+1}{2},\ \cos^3(x) = \frac{\cos(3x)+3\cos(x)}{4} \text{ et } \cos^4(x) = \frac{\cos(4x)+4\cos(2x)+3}{8}$$

4.3 GÉOMÉTRIE ET NOMBRES COMPLEXES

Nous avons vu que les nombres complexes peuvent être représentés par des points du plan ; inversement les nombres complexes permettent une formulation élégante de nombreux problèmes de géométrie du plan. Nous donnons dans cette partie deux exemples de ce phénomène.

4.3.1 Similitudes du plan.

En géométrie comme en physique, on étudie toujours les transformations préservant les distances ou les formes (ou des quantités bien adaptées au problème que l'on veut traiter) ; on s'intéresse ici aux transformations du plan préservant les formes au sens suivant :

Définition: Une *similitude* est une application f du plan vers lui-même telle que, pour tout x, y dans le plan, $d(f(x), f(y)) = \lambda d(x, y)$, où $\lambda \in \mathbf{R}_+^*$ est une constante qui s'appelle le *rapport* de la similitude f. Une similitude de rapport 1 s'appelle une *isométrie*.

Exemples : une translation, une rotation (autour d'un point selon un angle donné), une symétrie (orthogonale par rapport à une droite), une homothétie sont des similitudes ; ce sont des isométries sauf les dernières.

Les figures suivantes sont semblables deux à deux (il existe une similitude du plan qui transforme l'un en l'autre).

Remarque : l'ensemble des similitudes forme un groupe ; l'application qui à une similitude associe son rapport est un homomorphisme de groupe à valeurs dans $\mathbf{R}_+^*$ et dont le noyau est constitué par le sous-groupe des isométries.

Les nombres complexes permettent une description simple des similitudes :

THÉORÈME: *L'ensemble des similitudes est décrit par les transformations de* $\mathbf{C}$ *dans* $\mathbf{C}$ *données par :*

$$z \mapsto az + b \text{ ou } z \mapsto a\bar{z} + b$$

où $a \in \mathbf{C}^*$ *et* $b \in \mathbf{C}$. *Ces transformations sont des isométries si et seulement si* $|a| = 1$.

Démonstration: Il est immédiat de vérifier que les transformations décrites sont des similitudes (de rapport $|a|$) ; pour la réciproque, quitte à remplacer f par $g(z) = (f(z) - f(0))/(f(1) - f(0))$, on peut supposer que $f(0) = 0$ et $f(1) = 1$. Ecrivons alors les deux conditions $|f(z) - f(0)| = |z - 0|$ et $|f(z) - f(1)| = |z - 1|$, on obtient : $|f(z)| = |z|$

et $|f(z)|^2 - 2Re\, f(z) + 1 = |z|^2 - 2Re\, z + 1$ d'où $Re\, f(z) = Re\, z$ et $|f(z)| = |z|$. On obtient ainsi que $\forall z \in \mathbf{C}$, $f(z) = z$ ou $\bar{z}$. Reste à voir que si, disons $f(z_0) = z_0$, pour un nombre complexe non réel, alors pour tout $z \in \mathbf{C}$ on a $f(z) = z$. On écrit bien sûr que $|f(z) - f(z_0)| = |z - z_0|$ donc $Re\,(f(z)\bar{z}_0) = Re\,(z\bar{z}_0)$. Or l'équation $Re\,(\bar{z}\bar{z}_0) = Re\,(zz_0)$ entraîne (puisque $\mathrm{Im}(z_0) \neq 0$) que $\mathrm{Im}(z) = 0$ donc dans tous les cas $f(z) = z$. □

Exemples. La rotation de centre l'origine et d'angle θ correspond à la multiplication par $a = e^{i\theta}$; l'application $z \mapsto \bar{z}$ correspond à la symétrie orthogonale par rapport à l'axe des abscisses.

4.3.2 Droites, cercles et transformations homographiques.

Commençons par exprimer dans le plan complexe l'équation d'une droite et d'un cercle. Si $z = x + iy$ (avec $x, y \in \mathbf{R}$) on sait que $x = \frac{z+\bar{z}}{2}$ et $y = \frac{z-\bar{z}}{2i}$; comme l'équation cartésienne d'une droite est de la forme $ax + by + c = 0$ (avec $a, b, c \in \mathbf{R}$ et a ou b non nul) on en tire, en terme de z, l'équation de la droite : $\frac{a-ib}{2}z + \frac{a+ib}{2}\bar{z} + c$ ou encore, en posant $\alpha = \frac{a+ib}{2}$, on obtient l'équation $\alpha\bar{z} + \overline{\alpha}z + c$. L'équation d'un cercle de centre β et de rayon r peut s'écrire $|z - \beta| = r$ ou encore en élevant au carré : $z\bar{z} + \overline{\beta}z + \beta\bar{z} + |\beta|^2 = r^2$. Réciproquement considérons l'équation $az\bar{z} + \overline{\beta}z + \beta\bar{z} + c = 0$ (où $\beta \in \mathbf{C}$ et $a, c \in \mathbf{R}$) ; si $a = 0$ on retrouve l'équation d'une droite (sauf si β est aussi nul, cas trivial qu'on écarte) ; si $a \neq 0$, on peut diviser par a l'équation et en tirer : $z\bar{z} + \frac{\overline{\beta}}{a}z + \frac{\beta}{a}\bar{z} + \frac{|\beta|^2}{a^2} = -\frac{c}{a} + \frac{|\beta|^2}{a^2} = \frac{-ca+|\beta|^2}{a^2}$. On a ainsi montré :

THÉORÈME: *L'ensemble des cercles et droites du plan complexe est décrit par des équations du type :*

$$az\bar{z} + \overline{\beta}z + \beta\bar{z} + c = 0$$

(où $\beta \in \mathbf{C}$ et $a, c \in \mathbf{R}$ non tous nuls). On obtient ainsi une droite si $a = 0$ et $\beta \neq 0$, un cercle si $a \neq 0$ et $ac < |\beta|^2$ (resp. un point et l'ensemble vide si $ac = |\beta|^2$ ou $ac > |\beta|^2$).

Il est clair que les similitudes préservent l'ensemble des droites et des cercles mais il y a des transformations beaucoup plus générales qui font cela :

Définition: On appelle *fonction homographique* toute transformation du type $z \mapsto \frac{az+b}{cz+d}$ où $ad - bc \neq 0$. On appelle *fonction anti-homographique* toute transformation du type $z \mapsto \frac{a\bar{z}+b}{c\bar{z}+d}$ où $ad - bc \neq 0$.

Il faut tout de suite observer que, si $c \neq 0$, la fonction $f(z) = \frac{az+b}{cz+d}$ n'est pas définie en $z = -\frac{d}{c}$; on dit que $-\frac{d}{c}$ est le *pôle* de f ; de même la fonction f n'atteint pas la valeur $\frac{a}{c}$ puisque $\frac{az+b}{cz+d} = \frac{a}{c}$ entraînerait $bd = ac$. Pour ne pas alourdir les énoncés, on sous-entend souvent ce fait. Par ailleurs la condition $ad - bc \neq 0$ est mise pour éviter les fonctions constantes ; en effet si $ad - bc = 0$ avec disons $c \neq 0$ alors $b = \frac{ad}{c}$ et donc $f(z) = \frac{az+\frac{ad}{c}}{cz+d} = \frac{a}{c}$. D'un autre coté considérons $g(z) = \frac{dz-b}{-cz+a}$, alors $f \circ g(z) = \frac{ad-bc}{ad-bc}z = z$ (si $ad - bc \neq 0$).

Exemples :
Les translations $f(z) = z + a$ sont des homographies.
Les homothéties $f(z) = \lambda z$ (avec $\lambda \in \mathbf{R}^*$) sont des homographies.

Les rotations $f(z) = \alpha z$ (avec $\alpha = e^{i\theta}$) sont des homographies.
La symétrie $f(z) = \bar{z}$ est une anti-homographie.
L'inversion $f(z) = 1/\bar{z}$ est une anti-homographie.

Les quatre premiers exemples sont des similitudes et préservent donc toutes les formes ; ce n'est pas le cas de l'inversion, mais elle a tout de même la propriété remarquable de transformer une droite D en une droite (si $0 \in D$) ou un cercle (si $0 \notin D$) et de transformer un cercle C en une droite (si $0 \in C$) ou un cercle (si $0 \notin C$). Nous allons voir que c'est une propriété générale des (anti-)homographies.

THÉORÈME: *Les fonctions homographiques (resp. anti-homographiques) préservent l'ensemble des droites et des cercles. Une droite D est transformée en cercle par f, si le pôle de f n'est pas situé sur la droite D, ou en droite, si le pôle de f est situé sur la droite D. Un cercle C est transformé en cercle par f, si le pôle de f n'est pas situé sur le cercle C, ou en droite, si le pôle de f est situé sur le cercle C.*

Démonstration: Nous vérifions seulement que les (anti-)homographies transforment cercles et droites en cercles et droites et admettrons que ce sont les seules transformations ayant cette propriété. Une première démonstration est fournit par un calcul formel : par exemple l'équation $uz\bar{z} + \overline{\beta}z + \beta\bar{z} + v = 0$ se transforme par $z \mapsto \frac{az+b}{cz+d}$ en

$$(u|a|^2 + \overline{\beta}a\bar{c} + \beta\bar{a}c + v|c|^2)z\bar{z} + \overline{(u\bar{a}b + \overline{\beta}b\bar{c} + \beta\bar{a}d + v\bar{c}d)}z$$

$$+(u\bar{a}b + \overline{\beta}b\bar{c} + \beta\bar{a}d + v\bar{c}d)\bar{z} + (u|b|^2 + \overline{\beta}bd + \beta\bar{b}d + v|d|^2) = 0.$$

Une deuxième démonstration consiste à écrire $f(z) = \frac{az+b}{cz+d}$ comme composée de transformations simples ; il reste alors à vérifier la propriété pour transformations simples. Or, si $c \neq 0$ on a $f(z) = \frac{bc-ad}{c(cz+d)} + \frac{a}{c}$, donc si l'on pose $g(z) = cz + d, i(z) = 1/z$ et $h(z) = \frac{bc-ad}{c}z + \frac{a}{c}$ alors $f = h \circ i \circ g$. Il suffit donc de vérifier la propriété pour i ou encore pour l'inversion $z \mapsto /\bar{z}$, ce que nous avons déjà fait. □

Remarque : on a omis de préciser dans l'énoncé que si $\frac{a}{c}$ appartient à une droite ou un cercle, ce point n'est jamais dans l'image.

Exemple d'application : on veut transformer le demi-plan $\mathcal{H} := \{z \in \mathbf{C} \mid Im(z) > 0\}$ en le disque $\mathcal{D} := \{z \in \mathbf{C} \mid |z| < 1\}$. La transformation $f(z) = \frac{z-i}{z+i}$ transforme l'axe des imaginaires en l'axe réel et l'axe réel en le cercle de centre O et de rayon 1 et $\mathcal{H}$ en $\mathcal{D}$.

D'Alembert Jean (1717-1783)

CHAPITRE 5 L'ANNEAU DES ENTIERS Z.

La théorie des nombres est une des plus belles branches des mathématiques (Zut! L'auteur s'est dévoilé comme spécialiste de la théorie des nombres). Traditionnellement l'étude des propriétés de divisibilité fait apparaître la notion de nombres premiers : les entiers naturels divisibles uniquement par 1 et par eux-mêmes (on exclut 1 par convention) dont le début de la liste peut être obtenu par le crible d'Eratosthène : on raye les multiples de 2, puis les multiples de 3,5,7 et on obtient :

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97, 101, \ldots$$

ainsi que l'étude d'équations "diophantiennes comme $x^2 + y^2 = z^2$ (triangle pythagoriciens). Longtemps considérée comme une des branches les plus "pures", la théorie des nombres a trouvé des applications en informatique, cryptographie (code de cartes bancaires par exemple). Un des problèmes fondamentaux est de trouver (ou de prouver qu'il n'existe pas) un algorithme "rapide" de factorisation en produit de nombres premiers. L'arithmétique dans $\mathbf{N}$ *est souvent simplifiée par l'introduction des nombres négatifs, i.e. par l'introduction de l'anneau* $\mathbf{Z}$*.*

5.1 ARITHMÉTIQUE

Nous supposons connu l'ensemble

$$\mathbf{Z} := \{\ldots, -n, -n+1, \ldots, -2, -1, 0, 1, 2, \ldots, n-1, n, \ldots\}$$

qui est muni d'une loi d'addition et d'une loi de multiplication qui en font un anneau commutatif. Il est aussi muni d'une relation d'ordre qui permet de définir la *valeur absolue* d'un entier par la formule $|n| := max\{n, -n\}$.

Divisibilité : on dira que a divise b si b est un multiple de a ou encore si il existe $c \in \mathbf{Z}$ tel que $b = ac$. Un entier a est *inversible* si il existe $b \in \mathbf{Z}$ tel que $ab = 1$; on voit facilement que ceci équivaut à $a = \pm 1$. Ainsi, si a divise b et b divise a alors $a = \pm b$. Un nombre est *premier* si ses seuls diviseurs positifs sont 1 et lui-même (on exclut +1 et -1 par convention) ; on se restreint parfois aux nombres premiers positifs.

La propriété la plus fondamentale de l'anneau $\mathbf{Z}$ est l'existence de la *division euclidienne* qui est utilisée par l'étudiant depuis l'école primaire (au moins pour les nombres positifs) :

THÉORÈME: *Soit $a, b \in \mathbf{Z}$ avec $b \neq 0$ alors il existe $q, r \in \mathbf{Z}$, uniques, tels que :*
$a = bq + r$ $0 \leq r < |b|$

L'entier r s'appelle le *reste* de la division de a par b, et l'entier q s'appelle le *quotient* de la division de a par b.

Démonstration: Supposons d'abord pour simplifier que b est positif. On regarde la suite des multiples (positifs et négatifs) de b. On constate qu'il existe $q \in \mathbf{Z}$ tel que $qb \leq a < (q+1)b$ (il suffit de prendre pour q le plus grand entier tel que $qb \leq a$) ; posons $r := a - bq$ alors il vient $0 \leq r < b$ d'où le résultat. Si b est négatif, on procède de même avec $-a$ et $-b$: on obtient $-a = q_1(-b) + r_1$ avec $0 \leq r_1 < -b = |b|$ d'où $a = q_1 b - r_1$. Si

$r_1 = 0$ on a déjà la division, sinon on écrit $a = (q_1 + 1)b + (-b - r_1)$ et on note qu'on a bien $0 \leq -b - r_1 < |b|$.

Pour prouver l'unicité, on suppose que $a = bq + r = bq' + r'$ avec $0 \leq r, r' < |b|$. On en tire $|b||q - q'| = |r - r'| < |b|$ ce qui entraîne $|q - q'| = 0$ et donc $q = q'$ puis $r = r'$. □

Remarque : la démonstration donnée est proche mais un peu différente de l'algorithme appris à l'école primaire et qui peut être décrit ainsi : on cherche $c_0 \in \{0, 1, \ldots, 9\}$ et $n \geq 0$ tel que $(c_0 10^n)b \leq a < (c_0+1)10^n b$ et on remplace a par $a_1 = a - c_0 10^n b$ et on calcule c_1 tel que $(c_1 10^{n-1})b \leq a_1 < (c_1 + 1)10^{n-1} b$ et à la fin on obtient $a = b(c_0 10^n + \ldots + c_n) + a_{n+1}$.

THÉORÈME: *Les sous-groupes de* $\mathbf{Z}$ *sont tous de la forme* $n\mathbf{Z}$ *avec* $n \in \mathbf{N}$.

Démonstration: Soit G un sous groupe de $\mathbf{Z}$. On sait que $0 \in G$, si $G = \{0\}$ alors $G = 0.\mathbf{Z}$, sinon il existe n le plus petit élément strictement positif de G. L'ensemble des multiples de n est contenu dans G ; inversement, soit $g \in G$, effectuons la division euclidienne de g par n, on obtient $g = nq + r$ avec $0 \leq r < n$. On a donc l'égalité $r = g - nq$ et donc (comme g et n sont dans G) l'ensemble G contient r mais par choix de n ceci entraîne que $r = 0$ et que g est un multiple de n. On a donc bien $G = n\mathbf{Z}$. □

Remarque : les sous-groupes de $\mathbf{Z}$ sont aussi ses *idéaux* i.e. les sous-ensembles $I \subset \mathbf{Z}$ tels que I soit un sous-groupe et tels que $a \in \mathbf{Z}$ et $b \in I$ entraîne $ab \in I$.

Définition: Le plus grand commun diviseur (PGCD) de deux nombres a et b est un nombre d qui divise a et b et tel que tout diviseur commun de a et b divise d.

Définition: Le plus petit commun multiple (PPCM) de deux nombres a et b est un nombre m qui est un multiple a et b et tel que tout multiple commun de a et b est multiple de m.

Remarque : il n'est pas évident que le PGCD ou le PPCM existe mais ceci est garanti par la proposition suivante. Quand à l'unicité, la remarque faite sur les éléments inversibles montre que le PGCD (ou le PPCM) est unique au signe près. On choisit bien sûr le signe plus.

Remarque : si $a, b \in \mathbf{Z}$ on peut définir le sous-ensemble suivant de $\mathbf{Z}$:

$$a\mathbf{Z} + b\mathbf{Z} := \{au + bv \mid u, v \in \mathbf{Z}\}$$

dont on vérifie aisément que c'est un sous-groupe. De même $a\mathbf{Z} \cap b\mathbf{Z}$ est un sous-groupe.

PROPOSITION: *Soit* a *et* b *deux entiers non nuls alors il existe deux entiers* d *et* m *tels que* $a\mathbf{Z} + b\mathbf{Z} = d\mathbf{Z}$ *et* $a\mathbf{Z} \cap b\mathbf{Z} = m\mathbf{Z}$. *De plus l'entier* d *est un PGCD de* a *et* b, *et* m *est un PPCM de* a *et* b. *Enfin on a l'égalité* $ab = \pm md$.

Démonstration: Soit $d \in \mathbf{Z}$ tel que $a\mathbf{Z} + b\mathbf{Z} = d\mathbf{Z}$, montrons que d est un PGCD de a et b. Tout d'abord $a = a.1 + b.0$ est un multiple de d donc d divise a (et aussi b par le même raisonnement ; on peut aussi écrire $d = au + bv$ pour certains entiers u, v, par conséquent tout entier e diviseur commun de a et b divise au, bv et donc leur somme c'est-à-dire d.

Soit $m \in \mathbf{Z}$ tel que $a\mathbf{Z} \cap b\mathbf{Z} = m\mathbf{Z}$, montrons que d est un PPCM de a et b. Tout d'abord $m \in a\mathbf{Z}$ donc m est un multiple de a (et aussi de b par le même raisonnement) ;

si m' est un multiple de a et b alors $m' \in a\mathbf{Z}$ et $m' \in b\mathbf{Z}$ et donc $m' \in m\mathbf{Z}$ c'est-à-dire que m' est un multiple de m.

On sait donc que $a = a'd$ et $b = b'd$ donc $r := a'b'd$ est un multiple de a et b et est donc divisible par m ; donc md divise $rd = ab$. Par ailleurs, d'après la première partie du théorème, il existe $u, v \in \mathbf{Z}$ tels que $d = au + bv$ donc $md = aum + bvm$; mais ab divise am et bm donc md et on peut conclure que $md = \pm ab$. □

Si $\mathrm{PGCD}(a, b) = 1$ on dit que a et b sont *premiers entre eux*. Le résultat précédent nous permet de caractériser ces nombres :

THÉORÈME: *(Bézout) Soit* $d := \mathrm{PGCD}(a, b)$ *alors il existe deux entiers* u *et* v *tels que*

$$au + bv = d$$

En particulier deux entiers a *et* b *sont premiers entre eux si et seulement si il existe* u, v *entiers tels que* $au + bv = 1$.

Démonstration: La première partie de l'énoncé est une conséquence directe de la proposition précédente. Pour la deuxième partie, notons que si $\mathrm{PGCD}(a, b) = 1$ alors il existe $u, v \in \mathbf{Z}$ tels que $au + bv = 1$; inversement si de tels u, v existent, alors un diviseur d de a et b diviserait $au + bv$ et donc 1 ce qui donne bien que a et b sont premiers entre eux. □

Une des méthodes les plus rapides pour calculer le PGCD (et par conséquent le PPCM) est la suivante :

THÉORÈME: *(Algorithme d'Euclide)*
L'algorithme suivant fournit un calcul du PGCD de a *et* b *:*
$a = bq_1 + r_1$ (division de a par b)
$b = r_1q_2 + r_2$ (division de b par r_1)
$r_1 = r_2q_3 + r_3$ (division de r_1 par r_2)
. . . . . .
$r_{n-1} = r_nq_{n+1} + r_{n+1}$ (division de r_{n-1} par r_n)
Jusqu'à ce que $r_{n+1} = 0$ *et alors* $\mathrm{PGCD}(a, b) = r_n$

Démonstration: Elle consiste à vérifier que

$$\mathrm{PGCD}(a, b) = \mathrm{PGCD}(b, r_1) = \mathrm{PGCD}(r_1, r_2) = \ldots = \mathrm{PGCD}(r_{n-1}, r_n) = \mathrm{PGCD}(r_n, r_{n+1})$$

car clairement $\mathrm{PGCD}(r_n, r_{n+1}) = r_n$. Il suffit donc de montrer que pour a,b et q entier on a $\mathrm{PGCD}(a, b) = \mathrm{PGCD}(a - bq, b)$; ceci résulte du fait que d divise a et b équivaut à d divise b et $a - bq$. □

Remarque : si on le souhaite, une variante de cet algorithme permet de trouver u, v entiers tels que $au + bv = \mathrm{PGCD}(a, b)$. En effet il suffit d'écrire $\mathrm{PGCD}(a, b) = r_n = r_{n-2} - q_nr_{n-1}$, $r_{n-1} = r_{n-3} - q_{n-1}r_{n-2}$ etc pour en tirer r_n comme combinaison de r_{n-2}, r_{n-3} et ainsi de suite jusqu'à l'exprimer comme combinaison de a et b (ceci fournit d'ailleurs une autre démonstration du théorème de Bézout).

Faisons ce calcul sur un exemple ;
1932=6.301+126
301=2.126+49
126=2.49+28
49=1.28+21
28=1.21+7
21=3.7+0 (FIN du calcul du PGCD)
et
7=28–21=2.28–49=2.126–5.49=–5.301+12.126=12.1932–77.301 (FIN du calcul)
Résultat : $\mathrm{PGCD}(1932, 301) = 7 = 12.1932 - 77.301$

THÉORÈME: *(i) (Euclide) Soit p un nombre premier, si p divise ab alors p divise a ou b.*

(ii) (Gauss) Si $\mathrm{PGCD}(a, b) = 1$ *et a divise bc alors a divise c.*

Démonstration: (i) Supposons que p ne divise pas a alors $1 = \mathrm{PGCD}(a, p) = pu + av$ donc $b = pbu + abv$ donc p divisant pbu et abv divise b.

(ii) On a de même $1 = \mathrm{PGCD}(a, b) = au + bv$ donc $c = acu + bcv$ donc a divise c. □

THÉORÈME: *(Unicité de la décomposition en facteurs premiers) Soit n un entier distinct de $0, 1, -1$ alors il existe $\varepsilon = \pm 1$, il existe des nombres premiers $p_1, \ldots, p_r$ et des entiers $m_1, \ldots, m_r \geq 1$ tels que*

$$n = \varepsilon p_1^{m_1} \ldots p_r^{m_r}$$

de plus cette décomposition est unique à l'ordre près.

Exemples : $6440 = 2^3.5.7.23$, $1932 = 2^2.3.7.23$, $301 = 7.43$ Question : Avez-vous déjà factorisé votre numéro de téléphone? Celui de la police est un nombre premier alors que celui des pompiers se décompose en 2.3^2.

Démonstration: L'existence se prouve par récurrence sur n : si n est premier alors, on est content, sinon on a $n = ab$ avec $a < n$ et $b < n$ donc a et b, d'après l'hypothèse de récurrence se décomposent en produit de nombres premiers et donc n aussi.

L'unicité découle de l'application répétée du théorème d'Euclide (ou de Gauss). □

Remarques : si l'on connait la décomposition en facteurs premiers de deux nombres on peut facilement en déduire leur PGCD, mais ce n'est pas en général une méthode efficace de calcul. Par exemple on retrouve $\mathrm{PGCD}(1932, 301) = 7$ et on peut calculer $\mathrm{PGCD}(6440, 1932) = 23$ et $\mathrm{PGCD}(6440, 301) = 1$.

APPLICATION: Soit $n \in \mathbf{N}$ qui ne soit pas le carré d'un entier naturel, alors n n'est pas non plus le carré d'un nombre rationnel ; en d'autres termes le nombre $\sqrt{n}$ n'est pas un nombre rationnel.

Démonstration: Ecrivons $n = p_1^{m_1} \ldots p_r^{m_r}$; comme n n'est pas un carré, l'un des nombres premiers p_i apparaît dans la décomposition avec un exposant m_i impair. Si l'on pouvait écrire $\sqrt{n} = a/b$ avec $a, b \in \mathbf{N}$ on aurait $b^2 = na^2$; appelons m (resp. n) l'exposant de p_i dans la décomposition de a (resp. de b), alors l'unicité de la décomposition entraîne que $2n = 2m + m_i$ ce qui est absurde puisque m_i est impair. □

APPLICATION: Il existe une infinité de nombres premiers :

Démonstration: Soient $p_1, \ldots, p_r$ un ensemble fini de nombres premiers, montrons qu'il existe un nombre premier distinct de ceux-ci, ce qui achèvera la démonstration. Pour cela considérons $N := (p_1 \ldots p_r) + 1$ et q un facteur premier de N (il en existe) ; comme les p_i ne divisent pas N on doit avoir q distinct des p_i. □

Cet énoncé peut être considérablement affiné en quantifiant "combien" il y a de nombres premiers. Appelons $\pi(x)$ le nombre de nombres premiers $\leq x$; par exemple $\pi(2) = 1$, $\pi(3) = 2$, $\pi(10) = 4$ et $\pi(100) = 25$. Il y a un siècle Hadamard et De La Vallée-Poussin ont réussi à montrer que $\pi(x)$ valait à peu près $x/\log(x)$ (présisément $\pi(x)\log(x)/x$ tend vers 1 quand x tend vers l'infini). On peut interpréter ceci en disant que la probabilité pour qu'un nombre $\leq x$ soit premier est environ $1/\log(x)$; par exemple $100/\log 100 = 21,71\ldots$ et $\pi(100)\log(100)/100 = 1,15\ldots$ est déjà proche de 1 ; en poussant un peu plus loin le calcul on obtient $\pi(1000000)\log(1000000)/1000000 = 1,084\ldots\ldots$

5.2 CONGRUENCES

L'étudiant connait depuis l'école primaire les raisonnements de parité, la "preuve" par neuf et (peut-être) la "preuve" par onze du résultat d'une multiplication. La théorie des congruences est une généralisation de ce type de raisonnement.

5.2.1 Propriétés des congruences

Soit n un entier (strictement) positif, rappelons la définition de la relation de congruence modulo n

Définition: Deux nombres entiers a et b sont *congruents modulo* n si leur différence est divisible par n. On note cela $a \equiv b \bmod n$.

C'est une relation d'équivalence : elle est réflexive, symétrique, transitive (voir chapitre 2). Enonçons quelques unes de ses propriétés :

PROPOSITION: *1) Supposons que $a \equiv b \bmod n$ et $c \equiv d \bmod n$, alors $a + c \equiv b + d \bmod n$ et $ac \equiv bd \bmod n$ et si $r \geq 0$, on a $a^r \equiv b^r \bmod n$.*

2) Si $\mathrm{PGCD}(c, n) = 1$ alors il existe $c' \in \mathbf{Z}$ tel que $cc' \equiv 1 \bmod n$ et donc la congruence $ac \equiv bc \bmod n$ entraîne $a \equiv b \bmod n$.

3) $a \equiv b \bmod mn$ entraîne $a \equiv b \bmod n$.

Démonstration: 1) L'hypothèse se traduit par $a = b + kn$ et $c = d + jn$ donc $a + c = b + d + (j + k)n$ et $ac = bd + (kd + bj + kjn)n$. L'égalité $a^r = b^r + \ell n$ s'obtient par récurrence sur r.

2) Si c et n sont premiers entre eux il existe $u, v \in \mathbf{Z}$ tels que $cu + nv = 1$ et par conséquent $cu \equiv 1 \bmod n$. Si maintenant $ac \equiv bc \bmod n$, en multipliant par u on obtient bien $a \equiv b \bmod n$.

3) C'est immédiat. □

Remarque : sans l'hypothèse $\mathrm{PGCD}(c, n) = 1$ la conclusion de l'énoncé 2) peut être fausse car par exemple $2.4 \equiv 2.1 \bmod 6$ mais $4 \not\equiv 1 \bmod 6$.

Exemples de calculs : $1995 \equiv 5 \bmod 10$ donc $1995^4 \equiv 5^4 \bmod 10$ mais $5^2 \equiv 5 \bmod 10$ donc $1995^4 \equiv 5 \bmod 10$. De même on peut calculer $1991^{1991} \equiv 1 \bmod 10$.

APPLICATION: Preuves par 9 et 11.

L'écriture d'un nombre M en base décimale $M = c_n c_{n-1} \dots c_1 c_0$ signifie :

$$M = c_n 10^n + c_{n-1} 10^{n-1} + \dots + c_1 10 + c_0$$

et par conséquent $M \equiv c_n + c_{n-1} \dots + c_1 + c_0 \bmod 9$ (puisque $10 \equiv 1 \bmod 9$) et également $M \equiv c_0 - c_1 + c_2 \dots + (-1)^{n-1} c_{n-1} + (-1)^n c_n \bmod 11$ (puisque $10 \equiv -1 \bmod 11$). Supposons qu'on veuille vérifier le résultat d'une multiplication $M \times N = L?$, on calcule les restes ℓ, m, n de la division par 9 (idem avec 11) de L, M, N et on vérifie si $mn \equiv \ell \bmod 9$. Cela ne donne qu'une vérification et non une preuve, mais on ne fait pas souvent 9 erreurs de retenue... Le choix de 9 ou 11 provient du fait qu'on a une astuce simple permettant de calculer le reste modulo 9 ou 11 sans faire de division euclidienne. Exemple : $M = 1111114444$ et $N = 1234567$ alors $M \equiv 22 \equiv 4 \bmod 9$ et $M \equiv 0 \bmod 11$, $N \equiv 28 \equiv 1 \bmod 9$ et $N \equiv 4 \bmod 11$ donc $MN \equiv 4 \bmod 9$ et $MN \equiv 0 \bmod 11$.

THÉORÈME: *(théorème des restes chinois)*

Supposons que $\mathrm{PGCD}(m, n) = 1$ *alors le système de congruence :*

$$\begin{cases} x \equiv a \bmod m \\ x \equiv b \bmod n \end{cases}$$

possède une solution $x_0 \in \mathbf{Z}$ *et toutes les autres sont de la forme* $x_0 + mnk$ *avec* $k \in \mathbf{Z}$.

Démonstration: D'après le théorème de Bézout, il existe deux entiers u, v (dont on possède un algorithme de calcul) tels que $um + vn = 1$, posons donc $x_0 := a + (b - a)um$ alors $x_0 = b - (b - a)vn$ et c'est donc bien une solution du système de congruence. Soit x une autre solution, on a donc $x - x_0 \equiv 0 \bmod m$ et $x - x_0 \equiv 0 \bmod n$ donc (comme m et n sont premiers entre eux) $x - x_0 \equiv 0 \bmod mn$. □

Exemple : soit à résoudre

$$\begin{cases} x \equiv 5 \bmod 276 \\ x \equiv 2 \bmod 43 \end{cases}$$

On a vu que 276(=1932/7) et 43(=301/7) sont premiers entre eux et que $1 = 12.276 - 77.43$; on obtient donc une solution $x_0 = 5 + (2 - 5)12.276 = -9936$ et les autres solution sont donnés par $x = -9936 + 11868k$; la plus petite solution positive est 1932.

THÉORÈME: *("Petit" théorème de Fermat)*

Soit p *un nombre premier, alors pour tout entier* $a \in \mathbf{Z}$ *on a la congruence* $a^p \equiv a \bmod p$ *; si de plus* p *ne divise pas* a *alors* $a^{p-1} \equiv 1 \bmod p$.

Démonstration: Commençons par établir que $(x + y)^p \equiv x^p + y^p \bmod p$. En effet, d'après la formule du binôme de Newton, cette formule est surement vraie si tous les C_p^k pour $1 \leq k \leq p - 1$ sont divisibles par p. Mais $C_p^k k!(p - k)! = p!$ et p ne divise ni $k!$ ni $p - k$ (c'est là qu'on utilise que p est premier) car sinon, d'après le théorème d'Euclide, il diviserait un des facteurs c'est à dire un nombre $< p$; par conséquent en appliquant encore une fois Euclide on obtient que p divise C_p^k. On peut maintenant en déduire :

$$a^p - a \equiv (a - 1)^p - (a - 1) \equiv \dots \equiv 1^p - 1 \equiv 0 \bmod p$$

Si de plus p ne divise pas a alors a possède un inverse modulo p et donc $a^{p-1} \equiv 1 \bmod p$. □

COROLLAIRE: *Si $a \equiv b \not\equiv 0 \bmod p$ et si $c \equiv d \bmod (p-1)$ alors $a^c \equiv b^d \bmod p$.*

Démonstration: On sait déjà que $a^c \equiv b^c \bmod p$ et par ailleurs $c = d + k(p-1)$ donc $b^c \equiv b^d \left(b^{p-1}\right)^d \equiv b^d \bmod p$. □

Exemple : $112345^{149785} \equiv 2^5 \equiv 32 \equiv -1 \bmod 11$.

5.2.2 L'anneau $\mathbf{Z}/n\mathbf{Z}$

Notation : on désignera par $\bar{a}$ le reste de la division de a par n (c'est donc par convention un entier dans $\{0, 1, 2, \ldots, n-1\}$.

Définition: On appelle $\mathbf{Z}/n\mathbf{Z}$ l'ensemble $\{0, 1, 2, \ldots, n-1\}$ muni des lois d'addition et de multiplication suivantes : $(a, b) \mapsto \overline{(a+b)}$ et $(a, b) \mapsto \overline{ab}$.

PROPOSITION: L'ensemble $\mathbf{Z}/n\mathbf{Z}$, muni de ses lois d'addition et de multiplication est un anneau commutatif.

Démonstration: Découle immédiatement des propriétés de $\mathbf{Z}$ et des propriétés des congruences. □

On peut se demander s'il s'agit d'un corps, la réponse est la suivante :

THÉORÈME: *L'anneau $\mathbf{Z}/n\mathbf{Z}$ est un corps si et seulement si n est premier.*

Démonstration: Si n n'est pas premier, alors $n = ab$ avec $a, b \neq \pm 1$ et donc ni a, ni b n'est un multiple de n. Mais dans $\mathbf{Z}/n\mathbf{Z}$ on a donc $\bar{a} \neq \bar{0}$ et $\bar{b} \neq \bar{0}$ mais $\bar{a}\bar{b} = \overline{ab} = \bar{n} = 0$ donc $\mathbf{Z}/n\mathbf{Z}$ ne peut pas être un corps. Inversement si n est premier, montrons que tout élément $\bar{a} \in \mathbf{Z}/n\mathbf{Z} \setminus \{0\}$ possède un inverse : a et n sont premier entre eux donc il existe u, v tels que $au + vn = 1$ et donc $au \equiv 1 \bmod n$ et donc (dans $\mathbf{Z}/n\mathbf{Z}$) on a $\bar{a}\bar{u} = 1$. □

On peut traduire le théorème des restes chinois ainsi :

Considérons l'application $\rho : \mathbf{Z}/mn\mathbf{Z} \to \mathbf{Z}/m\mathbf{Z} \times \mathbf{Z}/n\mathbf{Z}$ qui a un élément $a \in \mathbf{Z}/nm\mathbf{Z}$ associe son reste modulo m et son reste modulo n ; les propriétés des congruences permettent de vérifier que c'est un morphisme d'anneaux : $\rho(a+b) = \rho(a) + \rho(b)$ et $\rho(ab) = \rho(a)\rho(b)$, le théorème des restes chinois dit que si $\mathrm{PGCD}(m, n) = 1$ alors ρ est une bijection (un isomorphisme).

Le calcul sur ordinateur s'effectue en fait modulo 2^N (avec N dépendant de la machine, du logiciel, etc). Par exemple, Turbo-Pascal recquiert, pour le type entier, des nombres compris entre -32768 et 32767 ; cela signifie qu'il travaille modulo 2^{16} (note : $2^{16} - 1 = 65535 = 32768 + 32767$.

Enfin indiquons que la cryptographie (cartes bancaires, transaction par internet, etc.) fait un usage massif, à travers le système RSA notamment, des groupes finis des éléments inversibles de $\mathbf{Z}/n\mathbf{Z}$ et des (grands) nombres premiers. les espaces vectoriels sur le corps $\mathbf{Z}/p\mathbf{Z}$ sont eux à la base des codes correcteurs d'erreurs (compact disque, transmission d'image, etc).

CHAPITRE 6 L'ANNEAU DES POLYNÔMES

Vous avez étudié depuis la seconde jusqu'à la terminale les fonctions de variable réelle de la forme $x \mapsto a_n x^n + \ldots + a_1 x + a_0$ *et appris à résoudre les équations du premier et du second degré. Il est commode pour approfondir cette étude de considérer les expressions formelles du type* $a_n x^n + \ldots + a_1 x + a_0$ *et de travailler directement sur elles. C'est ce point de vue qu'on adopte ici : un polynôme est défini comme la suite de ses coefficients ; cela permet notamment de développer l'analogie entre les propriétés de divisibilité dans l'anneau des polynômes et dans l'anneau* $\mathbf{Z}$*. Bien entendu la notation* $P = (a_0, \ldots, a_n)$*, même si elle présente l'avantage d'insister sur le rôle des coefficients, est impraticable et on utilisera la notation usuelle* $P = a_0 + \ldots + a_n X^n$*, celle de tout le monde, même des mathématiciens.*

6.1 POLYNÔMES

K désignera ici un sous-corps de $\mathbf{C}$ que l'on pourra prendre égal à $\mathbf{R}$ ou $\mathbf{C}$ pour simplifier.

Définition: Un *polynôme* à coefficient dans K est une suite d'éléments de K, disons $P = (a_0, a_1, \ldots, a_n, \ldots)$ telle qu'il existe n_0 avec $\forall n \geq n_0,\ a_n = 0$. Les a_i s'appellent les *coefficients* du polynôme P.

Un polynôme du type $(a_0, 0, 0, \ldots)$ s'appelle un *polynôme constant*. Le polynôme $(0, 0, \ldots)$ s'appelle le *polynôme nul*.

Le *degré* d'un polynôme non nul $P = (a_0, a_1, \ldots, a_n, \ldots)$ est l'entier

$$\deg(P) := \max\{n \in \mathbf{N} \mid a_n \neq 0\}$$

Il nous faut bien sûr définir l'addition et la multiplication :

Définition: Soit $P = (a_0, a_1, \ldots, a_n, \ldots)$ et $Q = (b_0, b_1, \ldots, b_n, \ldots)$ deux polynômes, alors leur somme et leur produit sont définis par : $P+Q := (a_0+b_0, a_1+b_1, \ldots, a_n+b_n, \ldots)$ et $PQ := (c_0, c_1, \ldots, c_n, \ldots)$ avec $c_n := \sum_{i=0}^{n} a_i b_{n-i}$.

THÉORÈME: *L'ensemble des polynômes, muni de l'addition et de la multiplication est un anneau commutatif ; l'élément neutre pour l'addition est le polynôme nul, l'élément neutre pour la multiplication est le polynôme constant* $\mathbf{1} := (1, 0, 0, \ldots)$.

On a les relations (lorsque ni P, ni Q ni $P+Q$ ne sont nuls) :

$$\deg(P+Q) \leq \max\{\deg(P), \deg(Q)\} \quad \text{et} \quad \deg(PQ) = \deg(P) + \deg(Q)$$

Démonstration: Il est immédiat de vérifier que l'addition définit une loi de groupe. Le polynôme constant dont le premier coefficient est 1 est bien l'élément neutre car si $b_0 = 1$ et $b_i = 0$ pour $i \geq 1$ on a bien $\sum_{i=0}^{n} a_i b_{n-i} = a_n$. Vérifier l'associativité est un exercice sur la notation "Sigma" que l'on laisse au lecteur.

Démontrons maintenant les formules sur les degrés : si $\deg(P) = d$ (resp. $\deg(Q) = e$) et p_d (resp. q_e) est le dernier coefficient non nul de P (resp. de Q) on voit facilement que

$p_i + q_i = 0$ dès que $i > \max(d, e)$ d'où la première inégalité. Remarquons que si $d > e$ le dernier coefficient non nul de $P + Q$ est p_d et donc dans ce cas on a $\deg(P + Q) = \max\{\deg(P), \deg(Q)\}$ (idem si $d < e$) alors que si $d = e$ tous les coefficients de $P + Q$ d'indice strictement supérieur à d sont nuls et le coefficient d'indice d vaut $p_d + q_e$ et donc peut fort bien être nul. Si $n > d + e$ alors $\sum_{i=0}^{n} p_i q_{n-i} = 0$ car chacun des termes est nul (ou bien $i > d$ ou bien $n - i > e$) ; par ailleurs le $(d+e)$-ème coefficient de PQ est $p_d q_e \neq 0$. De ces deux remarques on tire que $\deg(PQ) = d + e$. □

Voyons maintenant comment justifier et revenir à une notation plus usuelle : introduisons le polynôme $X = (0, 1, 0, 0 \ldots)$; on voit facilement que $X^2 = X.X = (0, 0, 1, 0, \ldots)$ et plus généralement que $X^n = (0, 0, \ldots, 0, 1, 0, \ldots)$ (où le 1 est le coefficient d'indice n). On en déduit une écriture plus commode (qui est celle que l'on utilisera dans toute la suite!) :

$$(a_0, a_1, \ldots, a_n, \ldots) = a_0\mathbf{1} + a_1 X + a_2 X^2 + \ldots + a_n X^n$$

Ceci justifie la
Notation : l'ensemble des polynômes à coefficients dans K se note $K[X]$. On dit que X est une *indéterminée*.

Remarque : nous distinguons donc le polynôme $a_0 + a_1 X + a_2 X^2 + \ldots + a_n X^n$ de la fonction $x \mapsto a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n$.

La propriété fondamentale de l'anneau des polynômes est, tout comme pour $\mathbf{Z}$, l'existence d'une division euclidienne :

THÉORÈME: *Soit $A \in K[X]$ et $B \in K[X] \setminus \{0\}$, il existe $Q, R \in K[X]$, uniques tels que :*

$$A = BQ + R \quad \text{et} \quad R = 0 \text{ ou } deg(R) < deg(B)$$

Démonstration: (unicité) Supposons $A = BQ + R = BQ' + R'$; alors $B(Q - Q') = R' - R$. Si R' était distinct de R alors $\deg(B) \leq \deg(B(Q - Q')) = \deg(R' - R) < \deg(B)$ amènerait une contradiction donc $R = R'$ et par conséquent $Q = Q'$.

(existence) La preuve se fait par récurrence sur $n := deg(A)$. Observons que si $\deg(A) < \deg(B)$ alors $A = 0.B + A$ fournit une division euclidienne. Supposons donc démontrée l'existence de la division euclidienne pour les polynômes de degré $\leq n - 1$ et établissons son existence pour A de degré n. On peut supposer $n \geq \deg(B) = m$ sinon on est dans un cas déjà traité ; écrivons $A = a_n X^n + \ldots$ et $B = b_m X^m + \ldots$ et considérons $A_1 := A - \frac{a_n}{b_m} X^{n-m} B$; si $A_1 = 0$ la démonstration est terminée et sinon, on voit aisément que $\deg(A_1) \leq n - 1$ car le coefficient de degré n s'annule (c'est fait pour!) donc d'après l'hypothèse de récurrence on sait que $A_1 = BQ_1 + R_1$ avec $\deg(R_1) < \deg(B)$ d'où l'on tire $A = B\left(Q_1 + \frac{a_n}{b_m} X^{n-m}\right) + R_1$ ce qui achève la démonstration de l'existence. □

Exemple : La démonstration fournit d'ailleurs un algorithme pour calculer Q et R ; illustrons cela avec $A = 2X^5 + 3X^3 + X^2 - X + 5$ et $B = X^2 + X - 1$: on peut présenter les calculs comme ceux de la division euclidienne usuelle (dans $\mathbf{Z}$) :

$$
\begin{array}{rcc}
 & 2X^5+3X^3+X^2-X+5 & X^2+X-1 \\
\ominus & 2X^5+2X^4-2X^3 & 2X^3-2X^2+7X-8 \\
 & -2X^4+5X^3+X^2-X+5 & \\
\ominus & -2X^4-2X^3+2X^2 & \\
 & 7X^3-X^2-X+5 & \\
\ominus & 7X^3+7X^2-7X & \\
 & -8X^2+6X+5 & \\
\ominus & -8X^2-8X+8 & \\
 & 14X-3 &
\end{array}
$$

ainsi $2X^5+3X3+X^2-X+5=(X^2+X-1)(2X^3-2X^2+7X-8)+(14X-3)$.

L'existence de la division euclidienne permet de développer les propriétés de divisibilité : PGCD,PPCM, théorème de Bézout, algorithme d'Euclide, théorème de Gauss, décomposition en produit de facteurs, de manière entièrement analogue à $\mathbf{Z}$. Nous donnons donc seulement les énoncés et renvoyons au chapitre précédent pour les démonstrations en signalant seulement les endroits où le vocabulaire introduit des différences. Les polynômes inversibles sont les constantes non nulles : en effet il est clair que ces polynômes sont inversibles et réciproquement si $PQ=1$ on a $\deg(P)+\deg(Q)=0$ et on conclut que P est constant. L'analogue des nombres premiers est donné par les polynômes *irréductibles*, i.e. par les polynômes P, non constants, qui ne peuvent s'écrire $P=QR$ avec Q,R deux polynômes non constants. Les polynômes inversibles sont les constantes non nulles.

Définition: Le *plus grand diviseur commun* ou PGCD de deux polynômes A et $B\in K[X]$ est un polynôme D qui divise A et B et tel que tout polynôme divisant A et B divise nécessairement D. Le *plus petit commun multiple* est un polynôme M multiple de A et B et tel que tout polynôme multiple de A et B soit divisible par M.

THÉORÈME: *Soient A,B deux polynômes, l'un d'entre eux non nul (au moins), le PGCD de A et B existe et, si l'on impose qu'il soit unitaire, il est unique. De même le PPCM existe et l'on a* $\mathrm{PPCM}(A,B)\ \mathrm{PGCD}(A,B)=AB$.

L'algorithme (d'Euclide) suivant fournit un calcul du PGCD :
$A=BQ_1+R_1$ (division de A par B)
$B=R_1Q_2+R_2$ (division de B par R_1)
$R_1=R_2Q_3+R_3$ (division de R_1 par R_2)
.....
$R_{n-1}=R_nQ_{n+1}+R_{n+1}$ (division de R_{n-1} par R_n)
Jusqu'à ce que $R_{n+1}=0$ *et alors* $\mathrm{PGCD}(A,B)=R_n$

Démonstration: La démonstration est identique au cas arithmétique : on doit seulement observer que $\deg(R_{i+1})<\deg(R_i)$ pour s'assurer que l'algorithme converge. □

Exemple de calcul : prenons $A:=X^6+X^5+X^4+X^2+X+1$ et $B=X^5+X^4+X^3+X^2+X+1$ alors
$A=BQ_1+R_1$ (avec $Q_1=X$ et $R_1=1-X^3$)
$B=R_1Q_2+R_2$ (avec $Q_2=-X^2-X-1$ et $R_2=2X^2+2X+2$)
$R_1=R_2Q_3+R_3$ (avec $Q_3=\frac{1}{2}$ et $R_3=0$)

donc $R_2 = 2(X^2+X+1)$ est le PGCD. Si l'on impose qu'ils soient unitaires $\mathrm{PGCD}(A,B) = X^2+X+1$ et $\mathrm{PPCM}(A,B) = (X^4+1)B = X^9+X^8+X^7+X^6+2X^5+2X^4+X^4+X^3+X^2+X+1$.

THÉORÈME: *(Bézout) Soit $A, B \in K[X]$ alors il existe $U, V \in K[X]$ tels que $AU + BV = \mathrm{PGCD}(A,B)$. De plus l'algorithme d'Euclide fournit également un calcul de U et V.*

Remarque : Les polynômes U et V ne sont pas uniques (en effet $U' = U + QB$ et $V' = V - QA$ font aussi l'affaire) mais on peut imposer (si A et B non constants) que $\deg(U) \leq \deg(B) - 1$ et $\deg(V) \leq \deg(A) - 1$.

Démonstration: On "copie" la démonstration faite pour $\mathbf{Z}$:
Considérons l'ensemble $I := \{AP + BQ \mid P, Q \in K[X]\}$; c'est un idéal de $K[X]$: la somme de deux éléments de I est dans I et le produit par un polynôme quelconque d'un élément de i est encore dans I ; l'existence de la division euclidienne entraîne, comme dans $\mathbf{Z}$, que tout idéal est engendré par un élément, c'est-à-dire que $I = DK[X] = \{DP \mid P \in K[X]\}$. Par définition de I il existe $U, V \in K[X]$ tels que $D = AU + BV$. Voyons que $D = \mathrm{PGCD}(A,B)$: tout d'abord $A \in I$ donc A est un multiple de D, idem pour B donc D divise A et B ; si maintenant C divise A et B alors C divise $D = AU + BV$ donc D est bien le PGCD. □

Exemple : reprenons le cas précédent $A := X^6+X^5+X^4+X^2+X+1$ et $B = X^5+X^4+X^3+X^2+X+1$ alors en remontant les étapes de l'algorithme d'Eclide on obtient : $R_2 = B - R_1Q_2 = B - (A - BQ_1)Q_2$ d'où $PGCD(A,B) = X^2+X+1 = (-\frac{1}{2}Q_2)A + \frac{1}{2}(1+Q_1Q_2)B$.

Définition: Un polynôme $P \in K[X]$ est dit *irréductible* s'il n'est pas constant et si les seules factorisations $P = QR$ (avec $Q, R \in K[X]$) s'obtiennent avec P ou Q constant.

Remarques : i) La notion de polynôme irréductible correspond à celle de nombre premier dans $\mathbf{Z}$.

ii) Les polynômes de degré 1 sont irréductibles car $X - a = QR$ entraîne Q ou R constant pour des raisons de degré. Néanmoins il y a beaucoup d'autres polynômes irréductibles en général ; par exemple X^2+1 est irréductible dans $\mathbf{R}[X]$, $X^3 - X + 1$ est irréductible dans $\mathbf{Q}[X]$

iii) Il est indispensable de préciser le corps K car par exemple X^2+1 n'est pas irréductible dans $\mathbf{C}[X]$ et $X^3 - X + 1$ n'est pas irréductible dans $\mathbf{R}[X]$ (ils ont chacun au moins une racine).

THÉORÈME:
(i) (Euclide) Soit P irréductible dans $K[X]$ et divisant QR alors P divise Q ou R.
(ii) (Gauss) Si $\mathrm{PGCD}(P,Q) = 1$ et P divise QR alors P divise R.

Démonstration: La démonstration est entièrement analogue à celle faite dans $\mathbf{Z}$. □

THÉORÈME: Soit $P \in K[X]$ un polynôme non constant, alors il existe $a \in K^*$ et des polynômes unitaires distincts $P_1, \ldots, P_r$ et des entiers $m_1, \ldots, m_r$ tous ≥ 1 tels que :

$$P = aP_1^{m_1} \ldots P_r^{m_r}$$

De plus les P_i, les m_i et a sont uniques.

Démonstration: La démonstration est entièrement analogue à celle faite dans **Z**. Il faut seulement observer que les polynômes inversibles (i.e. les P tels qu'il existe Q avec $PQ = 1$) sont les polynômes constants non nuls. □

Exemple : reprenons les polynômes A et B dont nous avons calculé le PGCD. Dans $\mathbf{Q}[X]$ on a $A = (X^2+X+1)(X^4+1)$ et $B = (X^2+X+1)(X^2-X+1)(X+1)$ alors que sur $\mathbf{R}[X]$ on a $A = (X^2+X+1)(X^2+\sqrt{2}X+1)(X^2-\sqrt{2}+1)$ et $B = (X^2+X+1)(X^2-X+1)(X+1)$ et sur $\mathbf{C}[X]$ on a $A = (X-j)(X-\bar{j})$ et $B = (X-j)(X-\bar{j})(X+j)(X-\bar{j})(X+1)$ (où l'on note $j := -\frac{1}{2} + i\frac{\sqrt{3}}{2}$).

Remarque : on peut montrer que $K[X]$ possède une infinité de polynômes irréductibles unitaires en "copiant" la démonstration faite pour les nombres premiers.

6.2 RACINES D'UN POLYNÔME

On étudie dans ce paragraphe les premières propriétés de la fonction associée à un polynôme : si $P := a_0 + a_1X + \ldots + a_nX^n \in K[X]$ alors on peut lui associer la fonction de K dans K définie par $x \mapsto a_0 + a_1x + \ldots + a_nx^n$. En particulier on s'intéresse aux valeurs de cette fonction ; en fait il nous suffira de regarder quand la fonction s'annule, ce qui nous amène à la notion de racine d'un polynôme.

PROPOSITION: *Soit $P \in K[X]$ et soit $\alpha \in K$ alors $P(\alpha) = 0$ si et seulement si $(X - \alpha)$ divise P.*

Démonstration: *Si $P = (X-\alpha)Q$ alors visiblement $P(\alpha) = 0$. Supposons inversement que $P(\alpha) = 0$ et effectuons la division de P par $X - \alpha$. On a $P = (X - \alpha)Q + R$ avec $R = 0$ ou $\deg(R) < \deg(X - \alpha) = 1$; donc R est constant et en calculant $P(\alpha)$ on trouve que $R = P(\alpha)$ donc $R = 0$ et $X - \alpha$ divise P.* □

Définition: On dit que α est une *racine* de P si $P(\alpha) = 0$ ou si $(X - \alpha)$ divise P. On dit que α est une *racine d'ordre r* de P si $(X - \alpha)^r$ divise P mais $(X - \alpha)^{r+1}$ ne divise pas P.

THÉORÈME: *Un polynôme de degré n possède au plus n racines (comptée avec multiplicités).*

Démonstration: Supposons que $\alpha_1, \ldots, \alpha_s$ soient des éléments distincts et racines d'ordre $m_1, \ldots, m_s$ de P alors les polynômes $(X - \alpha_i)^{m_i}$ sont premiers entre eux (deux à deux) et divisent P donc leur produit divise P. Or le produit $\prod_{i=1}^{s}(X - \alpha_i)^{m_i}$ a pour degré $\sum_{i=1}^{s} m_i$ donc $\sum_{i=1}^{s} m_i \leq \deg(P) = n$. □

Par analogie avec le calcul différentiel, on peut définir la dérivée d'un polynôme et il est raisonnable de penser que l'annulation des dérivées correspond à une racine multiple. Pour démontrer cela on établit la "formule de Taylor pour les polynômes" qui servira de prototype pour la formule de Taylor générale (chapitre 14).

Définition: Soit $P = a_nX^n + a_{n-1}X^{n-1} \ldots + a_1X + a_0$ un polynôme, le *polynôme dérivé* est $P' := na_nX^{n-1} + (n-1)a_{n-1}X^{n-2} \ldots + a_1$. On note $P^{(r)}$ la *dérivée n-ème* définit par $P^{(r+1)} = (P^{(r)})'$.

Cette opération de dérivation est donc définie sans passage à la limite mais jouit des mêmes propriétés que la dérivation des fonctions :

PROPOSITION: $(P+Q)' = P' + Q'$
$(PQ)' = P'Q + PQ'$
Plus généralement on a la formule Leibniz

$$(PQ)^{(n)} = \sum_{i=0}^{n} C_n^i P^{(i)} Q^{(n-i)}$$

Les propriétés suivantes sont équivalentes :
(i) P possède un racine d'ordre r en $X = \alpha$
(ii) $P(\alpha) = P'(\alpha) = \ldots = P^{(r-1)}(\alpha) = 0$ et $P^{(r)}(\alpha) \neq 0$

Démonstration: La démonstration de la première formule est laissée en exercice. Pour la deuxième formule on se ramène facilement au cas où $P = X^m$ et $Q = X^n$; alors $PQ' + QP' = X^n(mX^{m-1}) + X^m(nX^{n-1}) = (m+n)X^{m+n-1} = (PQ)'$. Un calcul par récurrence, à partir de la formule précédente donne la formule de Leibniz (ce calcul est fait au chapitre 14 pour la dérivation usuelle).

Si P possède une racine d'ordre r en α alors $P = (X-\alpha)^r Q$ avec $X - \alpha$ ne divisant pas Q donc $Q(\alpha) \neq 0$. En appliquant la formule de Leibniz on voit que

$$P(\alpha) = P'(\alpha) = \ldots = P^{(r-1)}(\alpha) = 0$$

et $P^{(r)}(\alpha) = r!Q(\alpha) \neq 0$. Pour établir la réciproque on va se servir de la formule suivante :

PROPOSITION: *(Formule de Taylor pour les polynômes) Soit P un polynôme de degré n et $\alpha \in K$ alors*

$$P = P(\alpha) + P^{(1)}(\alpha)(X-\alpha) + \frac{P^{(2)}(\alpha)}{2!}(X-\alpha)^2 + \ldots + \frac{P^{(n)}(\alpha)}{n!}(X-\alpha)^n$$

Démonstration: (de la formule de Taylor) Tout polynôme P de degré n peut s'écrire $P = \sum_{i=0}^{n} a_i (X-\alpha)^i$ (en effet il suffit de le vérifier pour $P = X^k$ et $X^k = (X - \alpha + \alpha)^k = \sum_{i=0}^{k} C_k^i \alpha^{k-i}(X-\alpha)^i$). La dérivation étant additive il suffit de vérifier la formule de Taylor pour le polynôme $P = (X-\alpha)^k$. Mais dans ce cas $P(\alpha) = P'(\alpha) = \ldots = P^{(k-1)}(\alpha) = 0$ et $P^{(k)}(\alpha) = k!$ donc la formule est vraie.

Terminons maintenant la preuve de la proposition :
Si $P(\alpha) = P'(\alpha) = \ldots = P^{(r-1)}(\alpha) = 0$ et $P^{(r)}(\alpha) \neq 0$ alors

$$P = \sum_{i=0}^{n} \frac{P^{(i)}(\alpha)}{i!}(X-\alpha)^i = (X-\alpha)^r \left(\frac{P^{(r)}(\alpha)}{r!} + \sum_{i=r+1}^{n} \frac{P^{(i)}(\alpha)}{i!}(X-\alpha)^{i-r} \right)$$

et on a bien $P = (X-\alpha)^r Q$ avec $Q(\alpha) = \frac{P^{(r)}(\alpha)}{r!} \neq 0$. □

Remarque : Jusqu'à présent nous aurions pu supposer le corps K commutatif quelconque, par exemple $K = \mathbf{Z}/p\mathbf{Z}$ mais la formule de Taylor n'est pas valable (telle quelle) sur $\mathbf{Z}/p\mathbf{Z}$ et de "drôles de choses" peuvent arriver en dérivant les polynômes : considérons le polynôme $P = X^{3p} + X^p + 1$ alors P n'est pas constant et pourtant $P' = 0$; par ailleurs, d'après le "petit" théorème de Fermat, le polynôme $P = X^p - X$ a pour racine tous les éléments du corps $\mathbf{Z}/p\mathbf{Z}$.

Explicitons maintenant la factorisation des polynômes à coefficients dans $\mathbf{R}$ et $\mathbf{C}$

THÉORÈME: *(Factorisation dans $\mathbf{R}[X]$ et $\mathbf{C}[X]$)*

(i) Les polynômes irréductibles dans $\mathbf{C}[X]$ sont les polynômes du premier degré ; tout polynôme de degré n se factorise sous la forme :

$$P = a_n X^n + \ldots + a_0 = a_n (X - \alpha_1)^{m_1} \ldots (X - \alpha_r)^{m_r}$$

avec les α_i distincts et $m_1 + \ldots + m_r = n$.

(ii) Les polynômes irréductibles dans $\mathbf{R}[X]$ sont les polynômes du premier degré et les polynômes du second degré de la forme $P = aX^2 + bX + c$ avec $b^2 - 4ac < 0$;

Remarque : on voit ainsi que, sur $\mathbf{R}$, tout polynôme de degré n se factorise sous la forme :

$$P = a_n X^n + \ldots + a_0 = a_n (X - \alpha_1)^{m_1} \ldots (X - \alpha_r)^{m_r} (X^2 + b_1 X + c_1)^{n_1} \ldots (X^2 + b_s X + c_s)^{n_s}$$

avec les α_i réels distincts, les couples (b_i, c_i) distincts vérifiant $b_i^2 - 4c_i < 0$ et $m_1 + \ldots + m_r + 2(n_1 + \ldots + n_s) = n$.

Démonstration: (i) Il faut démontrer que les seuls polynômes unitaires irréductibles sur $\mathbf{C}$ sont les $X - \alpha$ mais ceci est clair car tout polynôme non constant possède un facteur de ce type d'après le théorème de d'Alembert-Gauss.

(ii) Il faut démontrer que les seuls polynômes unitaires irréductibles sur $\mathbf{R}$ sont les $X - \alpha$ et les $X^2 + bX + c$ (avec $b^2 - 4c < 0$). Pour cela soit P un polynôme unitaire irréductible ; il possède une racine complexe α. Si $\alpha \in \mathbf{R}$ alors $X - \alpha$ divise P et donc $P = X - \alpha$. Sinon, observons que, comme P est à coefficient réel :

$$P(\overline{\alpha}) = \bar{P}(\overline{\alpha}) = \overline{P(\alpha)} = 0$$

d'autre part $(X - \alpha)(X - \overline{\alpha}) = X^2 - 2Re(\alpha)X + |\alpha|^2$ est à coefficient réel et divise P donc $P = X^2 - 2Re(\alpha)X + |\alpha|^2$. $\square$

Terminons ce paragraphe en étudiant, sur $\mathbf{R}$, le graphe des fonctions polynômes de degré ≤ 3 :

Si $P = aX + b$ on obtient une droite :

Si $P = aX^2 + bX + c$ on obtient une parabole qui a pour axe de symétrie la droite verticale $x = -\frac{b}{2a}$

Si $P = aX^3 + bX^2 + cX + d$ on obtient une cubique qui a toujours un point de symétrie ; pour étudier le signe de la dérivée $P = 3ax^2 + 2bX + c$, on a besoin du signe de $\Delta := b^2 - 3ac$

6.3 FRACTIONS RATIONNELLES

Une fraction rationnelle F à une indéterminée est une expression du type $F = \frac{P}{Q}$ avec P et Q polynômes (Q est supposé non nul et bien sûr $\frac{PR}{QR} = \frac{P}{Q}$). L'ensemble des fractions rationnelles forme un corps qu'on peut construire formellement à partir de l'anneau des polynômes de la même façon que $\mathbf{Q}$ *est construit à partir de* $\mathbf{Z}$.

Définition: Un *élément simple* de $\mathbf{C}(X)$ est une fraction rationnelle de la forme :

$$F = \frac{b}{(X-a)^n}$$

avec $a, b \in \mathbf{C}$ et $n \in \mathbf{N}^*$.

Un *élément simple* de $\mathbf{R}(X)$ est une fraction rationnelle de la forme :

$$F = \frac{b}{(X-a)^n} \quad ou \quad F = \frac{cX+d}{(X^2+aX+b)^n}$$

avec $a, b, c, d \in \mathbf{R}$ et $n \in \mathbf{N}^*$ et (dans le deuxième cas $a^2 - 4b < 0$).

L'intérêt de cette notion est illustré par le théorème suivant :

THÉORÈME: *Soit $K = \mathbf{R}$ ou $\mathbf{C}$, une fraction rationnelle de $K(X)$ peut s'écrire de manière unique comme somme d'un polynôme et d'éléments simples, cette écriture s'appelle la décomposition en éléments simples de la fraction rationnelle.*

Exemples : La décomposition en éléments simples de $F = \frac{X^3+X^2+X-1}{X^3-X}$ est $F = 1 + \frac{1}{(X-1)} - \frac{1}{(X+1)} + \frac{1}{X}$

Remarque : l'unicité de cette décomposition est très utile pour la calculer comme on le verra sur les exemples. Nous allons faire la démonstration sur **C** et laissons le lecteur adapter l'argument au corps des réels ; en fait si F est une fraction à coefficient réels, on peut la décomposer en éléments simples sur **R** et sur **C**.

Démonstration: (Unicité) Il suffit de voir que si :

$$F = P + \sum_{i=1}^{r} \sum_{j=1}^{m_i} \frac{b_{ij}}{(X-a_i)^j} = 0$$

alors les coefficients b_{ij} et le polynôme P sont nuls. Pour cela multiplions F par $(X-a_i)^{m_i}$; on obtient une égalité de la forme $0 = (X-a_i)^{m_i}F = b_{im_i} + (X-a_i)G$ avec G une fraction rationnelle sans pôle en a_i. En calculant les valeurs en a_i, on obtient donc $b_{im_i} = 0$. En répétant l'opération pour chaque coefficient on obtient $b_{ij} = 0$ et donc $P = 0$.

(existence) Commençons par observer que si P et Q sont des polynômes premiers entre eux alors, d'après le théorème de Bézout, il existe deux polynômes A, B tels que $AP+BQ = 1$; donc toute fraction rationnelle de la forme $F = \frac{R}{PQ}$ peut s'écrire $F = \frac{R(AP+BQ)}{PQ} = \frac{AR}{Q} + \frac{BR}{P}$. Par ailleurs tout polynôme s'écrit à un coefficient près $D = \prod_{i=1}^{r}(X-a_i)^{m_i}$ donc toute fraction rationnelle $F = \frac{C}{D}$ va se décomposer en $\sum_{i=1}^{r} \frac{P_i}{(X-a_i)^{m_i}}$ mais si on utilise maintenant la "formule de Taylor" pour P_i au point a_i on obtient $P_i = \sum_{j=0}^{\deg(P_i)} p_{ij}(X-a_i)^j$ d'où en reportant une expression de F comme somme d'éléments simples et de polynômes. □

L'utilisation la plus fréquente de la décomposition en éléments simples est le calcul de primitives (voir chapitre 17) mais elle peut être utilisée aussi pour calculer la dérivée n-ème ; par exemple si $F = \frac{X^3+X^2+X-1}{X^3-X} = 1 + \frac{1}{X-1} - \frac{1}{X+1} + \frac{1}{X}$ alors, comme on sait que $\int \frac{dt}{(t-a)} = Log|t-a| + C'$ on en tire

$$\int F(t)dt = t + \log|\frac{t^2-t}{t+1}| + C$$

En observant que $(\frac{d}{dt})^m(\frac{1}{t-a}) = (-1)^m \frac{m!}{(t-a)^{m+1}}$ on en tire :

$$F^{(m)}(t) = (-1)^m m! \left(\frac{1}{(t-1)^{m+1}} - \frac{1}{(t+1)^{m+1}} + \frac{1}{t^{m+1}} \right)$$

Pratique de la décomposition en éléments simples : On peut appliquer la méthode suivante : on factorise le dénominateur, puis on écrit formellement le type de la décomposition en éléments simples avec des coefficients inconnus, on calcule ensuite ces coefficients à l'aide des lemmes qui suivent.

La partie polynômiale de la décomposition simple s'appelle la *partie entière* de la fraction rationnelle ; elle se calcule ainsi :

LEMME: *Soit $F = P/Q$ une fraction rationnelle et soit E le quotient de la division de P par Q, i.e. $P = EQ + R$ avec $\deg(R) < \deg Q$ alors E est la partie entière de la fraction F.*

Démonstration: En effet, en réduisant au même dénominateur la somme des éléments simples, on obtient une égalité de la forme $F = E + R/Q$ avec P, R polynômes et $\deg(R) \leq \deg(Q) - 1$. Après multiplication par Q cette égalité devient $P = EQ + R$ qui indique que E est le quotient de P par Q et R le reste puisque $\deg(R) < \deg(Q)$. □

Il est également aisé de trouver le coefficient correspondant à un pôle d'ordre maximal :

LEMME: *Soit $Q = (X-a)^m Q_1$ avec $Q_1(a) \neq 0$ et $F = P/Q$ dont la décomposition s'écrit : $F = \frac{u_m}{(X-a)^m} + \ldots$ alors $u_m = P(a)/Q_1(a) = m!P(a)/Q^{(m)}(a)$.*

Démonstration: On a $(X-a)^m F = P/Q_1 = u_m + (X-a)G$ avec G fraction rationnelle sans pôle en a ; en calculant les valeurs pour $X = a$, on en déduit $u_m = P(a)/Q_1(a)$. Par ailleur la formule de Leibniz nous donne $Q^{(m)}(a) = m!Q_1(a)$ d'où la deuxième expression. □

Ces deux lemmes sont déjà suffisants pour calculer la décomposition d'une fraction rationnelle sans pôle double.

Exemple : soit $F = \frac{X^{2n}}{X^n-1}$ on effectue la division $X^{2n} = (X^n+1)(X^n-1)+1$; on sait que les racines de $X^n - 1$ sont les racines n-èmes de l'unité et on peut factoriser $X^n - 1 = \prod_{h=0}^{n-1}(X - \alpha_h)$ avec $\alpha_h := exp(\frac{2\pi i h}{n})$ d'où une expression a priori :

$$F = E + \sum_{h=0}^{n-1} \frac{u_h}{X - \alpha_h}$$

D'après le premier lemme on a $E = X^n + 1$ et si on applique le deuxième lemme avec $P = X^{2n}$ et $Q = X^n - 1$ on obtient $u_h = \frac{(\alpha_h)^{2n}}{n(\alpha_h)^{n-1}} = \alpha_h/n$ et donc

$$F = \frac{X^{2n}}{X^n - 1} = X^n + 1 + \frac{1}{n}\sum_{h=0}^{n-1} \frac{\alpha_h}{(X - \alpha_h)}$$

Pour traiter les calculs avec des pôles multiples le plus économique est d'utiliser le lemme suivant :

LEMME: *(division aux puissances croissantes) Soit $\alpha \in K$, $P, Q \in K[X]$ avec $Q(\alpha) \neq 0$ alors pour tout $k \in \mathbf{N}$ il existe $a_i \in K$ et $R \in K[X]$ tels que :*

$$P = (a_0 + a_1(X-\alpha) + \ldots + a_{k-1}(X-\alpha)^{k-1})Q + (X-\alpha)^k R$$

En particulier si $F := P/(X-\alpha)^k Q$ alors la partie de la décomposition en éléments simples de F correspondant au pôle α s'écrit :

$$\frac{a_0}{(X-\alpha)^k} + \ldots + \frac{a_{k-1}}{(X-\alpha)}$$

Démonstration: Observons que le polynôme $P-(P(\alpha)/Q(\alpha))Q$ s'annule en α donc $P-(P(\alpha)/Q(\alpha))Q=(X-\alpha)R$; raisonnons maintenant par récurrence sur k et supposons $P=(a_0+a_1(X-\alpha)+\ldots+a_{k-1}(X-\alpha)^{k-1})Q+(X-\alpha)^kR_k$; on sait que $R_k=rQ+(X-\alpha)R_{k+1}$ (avec $r:=R_k(\alpha)/Q(\alpha)$) d'où $P=(a_0+a_1(X-\alpha)+\ldots+a_{k-1}(X-\alpha)^{k-1}+r(X-\alpha)^k)Q+(X-\alpha)^{k+1}R_{k+1}$. La deuxième affirmation est immédiate. □

Exemple : $f=\dfrac{X^{10}+X^2+1}{X^9-2X^5+X}$ Nous allons calculer la décomposition en éléments simples sur $\mathbf{C}$ et en déduire celle sur $\mathbf{R}$. Cette fraction est aussi l'occasion de faire des remarques sur l'utilisation des symétries (coefficients réels, parité).

Factorisation : le polynôme $Q:=X^9-2X^5+X$ se décompose en :
$X(X^4-1)^2=X(X-1)^2(X+1)^2(X^2+1)^2=X(X-1)^2(X+1)^2(X+i)^2(X-i)^2$
Décomposition a priori de f avec $A,B,C,D,E,F,G,H,I\in\mathbf{C}$ et $R\in\mathbf{C}[X]$:

$$R+\frac{A}{X}+\frac{B}{X+1}+\frac{C}{(X+1)^2}+\frac{D}{(X-1)}+\frac{E}{(X-1)^2}+\frac{F}{(X-i)}+\frac{G}{(X-i)^2}+\frac{H}{(X+i)}+\frac{I}{(X+i)^2}$$

Utilisation des symétries : on sait que $\bar{f}=f$ et on observe que $f(X)=-f(-X)$; en reportant cela dans la décomposition et en utilisant l'unicité de la décomposition on obtient $R(X)=-R(-X)$, $B=D$, $E=-C$, $F=H$ et $I=-G$ d'une part et d'autre part $\bar{E}=E$, $\bar{A}=A$, $\bar{B}=B$, $\bar{C}=C$, $\bar{D}=D$, $\bar{E}=E$, $\bar{F}=H$, $\bar{G}=I$ d'où l'on tire que A, B, C, F sont réels (et $D=B$, $E=-C$) et G est imaginaire pur (et $I=-G$).

Le premier lemme permet de calculer R ; on trouve $R=X$ (qui est bien impair et à coefficients réels).

Le deuxième lemme permet de calculer A (et si on voulait C et G) : $f_0:=Xf=\frac{X^{10}+X^2+1}{(X^4-1)^2}$ donc $f_0(0)=A=1$

Le troisième lemme permet de calculer simultanément C et B :

$Q=(X+1)^2Q_1$ avec $Q_1=X^7-2X^6+3X^5-4X^4+3X^3-2X^2+X$ d'où l'on tire (par la formule de Taylor par exemple) $Q_1=-16+64(X+1)+(X+1)^2Q_2$ et de même $P=X^{10}+X^2+1=3-12(X+1)+(X+1)^2P_2$ d'où l'écriture de la division aux puissances croissantes de P par Q_1 par rapport à $(X+1)$:

$$3-12(X+1)=(a_0+a_1(X+1))(-16+64(X+1))+(X+1)^2R$$

d'où l'on tire $a_0=-3/16$ et $a_1=0$ soit $C=-3/16$ et $B=0$

On procède de même pour les coefficients G et F : On a $Q=(X-i)^2Q_i$ avec $Q_i=X(X^2-1)^2(X+i)^2$. Ceci permet de calculer $Q_i(i)=-16i$ et $Q_i'(i)=-64$ d'où $Q_i=-16i-64(X-i)+(X-i)Q_2$. De la même façon $P(i)=-1$ et $P'(i)=12i$ donc $P=-1+12(X-i)+(X-i)^2P_2$ d'où l'écriture de la division aux puissances croissantes de P par Q_i par rapport à $(X-i)$:

$$-1+12(X-i)=(b_0+b_1(X-i))(-16i-64(X-i))+(X-i)^2R$$

d'où l'on tire aisément $G = b_0 = -i/16$ et $F = b_1 = -1/2$ (on vérifie bien que G est purement imaginaire et F réel).

Conclusion : f s'écrit :

$$X + \frac{1}{X} + \frac{-3/16}{(X+1)^2} + \frac{3/16}{(X-1)^2} + \frac{-1/2}{X-i} + \frac{-i/16}{(X-i)^2} + \frac{-1/2}{X+i} + \frac{i/16}{(X+i)^2}$$

Pour obtenir la décomposition sur $\mathbf{R}$ il suffit ici de regrouper les termes complexes conjugués : $\frac{1}{(X+i)} + \frac{1}{(X-i)} = \frac{2X}{X^2+1}$ et $\frac{i}{(X+i)^2} - \frac{i}{(X-i)^2} = \frac{4X}{(X^2+1)^2}$ donc

$$f = X + \frac{1}{X} + \frac{-3}{16(X+1)^2} + \frac{3}{16(X-1)^2} - \frac{X}{X^2+1} + \frac{X}{4(X^2+1)^2}$$

Remarque. On peut souvent utiliser avantageusement un calcul de valeur particulière ou de limite pour calculer certains coefficients ou des relations entre eux. Par exemple, en gardant les notations de l'exemple précédent, une fois calculé R, on a facilement $F - R = (2X^6 + 1)/(X^9 - 2X^5 + X)$ donc

$$\lim_{x\to\infty} x(F(x) - R(x)) = 0 = A + B + D + F + H.$$

On voit qu'il n'y a pas de difficulté fondamentale pour calculer la décomposition en éléments simples sinon celle de bien ordonner les calculs.

Noether Emmy (1882–1935)

CHAPITRE 7 MATRICES

Ce chapitre introduit un outil de calcul très commode : les matrices. Celles-ci sont définies comme un tableau de nombres mais on en verra une interprétation plus abstraite au chapitre 9. La technique centrale de ce chapitre est celle dite du "pivot de Gauss" qui est à la fois simple, efficace et versatile.

7.1 OPÉRATIONS SUR LES MATRICES

Définition: Soient $m, n \geq 1$ des entiers ; une *matrice* $m \times n$ est un tableau de nombres avec m lignes et n colonnes :

$$A := \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ . & . & \dots & . \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

Les éléments a_{ij} s'appellent les *coefficients* de la matrice A. Par convention, le premier indice est le numéro de la ligne, le deuxième indice est le numéro de la colonne. On désignera par $Mat(m \times n, \mathbf{R})$ l'ensemble des matrices $m \times n$ à coefficients dans $\mathbf{R}$.

Remarque : on travaillera pour le moment avec des nombres réels mais il n'y a pas de difficulté à étendre les notions de ce chapitre au cas où les coefficients sont dans un corps commutatif K.

Exemple : $A = \begin{pmatrix} 0 & -1 & 3 & 4 \\ 1 & 2 & -3 & 4 \end{pmatrix}$ est une matrice 2×4 et son coefficient a_{22} est égal à 2, alors que $a_{13} = 3$.

Une matrice $1 \times n$ est un *vecteur ligne*. Une matrice $m \times 1$ est un *vecteur colonne*.

Définissons maintenant les opérations sur les matrices :

Addition :

Si $A := (a_{ij})$ et $B := (b_{ij})$ sont des matrices $m \times n$ alors $A + B$ est une matrice $m \times n$ dont les coefficients sont donnés par $c_{ij} := a_{ij} + b_{ij}$.

$$\text{Exemple : } \begin{pmatrix} 0 & -1 & 3 & 4 \\ 1 & 2 & -3 & 4 \end{pmatrix} + \begin{pmatrix} 1 & 2 & 0 & 0 \\ -1 & 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 3 & 4 \\ 0 & 5 & -1 & 5 \end{pmatrix}$$

Multiplication par un scalaire : Si $A := (a_{ij})$ est une matrice $m \times n$ et $\alpha \in \mathbf{R}$ alors αA est une matrice $m \times n$ dont les coefficients sont donnés par $c_{ij} := \alpha a_{ij}$.

$$\text{Exemple : } 3 \begin{pmatrix} 0 & 1 & -1 & 3 & 4 \\ 1 & 2 & -3 & 4 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 3 & -3 & 9 & 12 \\ 3 & 6 & -9 & 12 & 3 \end{pmatrix}$$

Enfin l'opération la plus intéressante (et la plus compliquée) :

Multiplication de deux matrices :

Si $A := (a_{ij})$ et $B := (b_{ij})$ sont des matrices $m \times n$ et $n \times p$ alors AB est une matrice $m \times p$ dont les coefficients sont donnés par $c_{ij} := \sum_{k=1}^{n} a_{ik} b_{kj}$.

Remarque : il faut bien noter que la multiplication de deux matrices n'est définie que si le nombre de colonnes de la première matrice est égal au nombre de lignes de la seconde matrice.

Cas particulier : Le produit d'un vecteur ligne par un vecteur colonne (de même longueur) est un nombre :

$$\begin{pmatrix} x_1 & x_2 & \dots & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ . \\ . \\ y_n \end{pmatrix} = x_1y_1 + x_2y_2 + \dots + x_ny_n$$

Exemple : $\begin{pmatrix} 3 & -1 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} = 0$

(ce produit est en fait le produit scalaire et la dernière égalité signifie que les vecteurs $(3, -1, 4)$ et $(1, -1, -1)$ sont orthogonaux).

Cas général : Le produit d'une matrice s'obtient en faisant le produit de chaque ligne de la première matrice par les colonnes de la seconde :

$$\begin{pmatrix} a_{11} & \dots & \dots & a_{1n} \\ . & \dots & \dots & . \\ a_{i1} & \dots & \dots & a_{in} \\ . & \dots & \dots & . \\ a_{m1} & \dots & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11} & \dots & b_{1j} & \dots & b_{1p} \\ & & . & & \\ & & . & & \\ & & . & & \\ & & . & & \\ & & . & & \\ b_{n1} & \dots & b_{nj} & \dots & b_{np} \end{pmatrix} = \begin{pmatrix} c_{11} & & & & c_{1p} \\ & & & & \\ & & c_{ij} & & \\ & & & & \\ c_{m1} & & & & c_{mp} \end{pmatrix}$$

Exemple :

$$\begin{pmatrix} 0 & -1 & 3 & 4 \\ 1 & 2 & -3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & -1 & -1 \\ -1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} -5 & -5 & -4 \\ -1 & 8 & 5 \end{pmatrix}$$

Plus généralement on peut faire le produit de matrices par blocs (de tailles compatibles) :

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} A' & B' \\ C' & D' \end{pmatrix} = \begin{pmatrix} AA' + BC' & AB' + BD' \\ CA' + DC' & CB' + DD' \end{pmatrix}$$

où $A, B, \dots, D'$ sont des matrices, en les manipulant comme des scalaires, mais en faisant attention à ne pas les faire commuter car :

La multiplication des matrices n'est pas commutative, on a en général $AB \neq BA$, ainsi par exemple : $\begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix} \neq \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}$.

On peut aussi rajouter une opération qui est plutôt une notation : on peut juxtaposer une matrice $m \times n$ et une matrice $m \times r$ pour obtenir une matrice $m \times (n + r)$. Si A et B sont les deux matrices initiales, on note la nouvelle matrice $(A|B)$

Les matrices permettent de compacter des formules et donc de les manipuler de façon plus efficace ; nous allons appliquer cela au problème de la résolution d'un *système linéaire*, c'est-à-dire d'un système d'équation du type :

$$\begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n = b_1 \\ \qquad\qquad \ldots \\ a_{m1}x_1 + \ldots + a_{mn}x_n = b_m \end{cases}$$

On peut réécrire un tel système en termes de matrices en posant

$$A := (a_{ij}), \quad X := \begin{pmatrix} x_1 \\ . \\ . \\ . \\ x_n \end{pmatrix} \quad \text{et} \quad b := \begin{pmatrix} b_1 \\ . \\ . \\ . \\ b_n \end{pmatrix}$$

Alors le système équivaut à :

$$AX = b$$

Avant de manipuler plus avant les matrices, il convient de connaître les règles de calcul :

PROPOSITION: *Les opérations sur les matrices vérifient les règles suivantes :*
1) (distributivité) $A(B+C) = AB + AC$ *et* $(A+B)C = AC + BC$
2) (associativité) $(AB)C = A(BC)$
3) (compatibilité) $\alpha(AB) = (\alpha A)B = A(\alpha B)$

Démonstration: La vérification ne présente pas de difficulté. □

On peut aussi introduire des matrices particulières :

La *matrice nulle* $m \times n$ est la matrice dont tous les coefficients sont nuls ; on la note 0_{mn} ou simplement 0 si le contexte rend clair sa taille. Observons que pour toute matrice A (de taille $n \times p$) on a $0_{mn}A = 0_{mp}$.

La *matrice identité* $m \times m$ est la matrice dont tous les coefficients sont nuls sauf ceux situés sur la diagonale qui valent 1 ; on la note I_m ou simplement I si le contexte rend clair sa taille. Ainsi

$$I = \begin{pmatrix} 1 & & & 0 \\ & 1 & & \\ & & . & \\ & & & . \\ 0 & & & 1 \end{pmatrix}$$

(où les coefficients laissés en blanc sont nuls). Observons que pour toute matrice A de taille $m \times n$ on a $I_m A = A = AI_n$.

Une *matrice diagonale* $m \times m$ est une matrice dont tous les coefficients sont nuls sauf ceux situés sur la diagonale ; ainsi

$$D = \begin{pmatrix} a_{11} & & & 0 \\ & a_{22} & & \\ & & . & \\ & & & . \\ 0 & & & a_{mm} \end{pmatrix}$$

(où les coefficients laissés en blanc sont nuls).

Une *matrice triangulaire supérieure* $m \times m$ est une matrice dont tous les coefficients sont nuls sauf ceux sur la diagonale et au dessus de la diagonale ; ainsi

$$T = \begin{pmatrix} a_{11} & * & * & * & * \\ & a_{22} & * & * & * \\ & & . & * & * \\ & & & . & * \\ 0 & & & & a_{mm} \end{pmatrix}$$

(où les coefficients laissés en blanc sont nuls et les étoiles désignent des coefficients quelconques).

COROLLAIRE: *L'ensemble des matrices carrées $Mat(n \times n, \mathbf{R})$, muni de l'addition et de la multiplication est un anneau ; si $n = 1$ c'est un corps "égal" à $\mathbf{R}$, si $n \geq 2$ c'est un anneau non commutatif qui n'est pas un corps.*

Démonstration: Les propriétés d'anneau sont contenues dans les propriétés générales des matrices, l'élément neutre de la loi d'addition étant 0_{nn} et l'élément neutre de la multiplication étant I_n. On a déjà vu que l'anneau n'est pas commutatif si $n = 2$ et de même ce n'est pas un corps car $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = 0_{22}$ est nul sans qu'aucun des facteurs ne soit nul. □

Il est donc intéressant de rechercher si une matrice a un inverse (à gauche ou à droite si elle n'est pas carrée). Faisons le explicitement sur les matrices 2×2 : observons que

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = (ad - bc) I_2$$

On voit donc que si $ad - bc \neq 0$ alors la matrice $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ est inversible et

$$A^{-1} = \begin{pmatrix} \frac{d}{ad-bc} & \frac{-b}{ad-bc} \\ \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{pmatrix}$$

D'autre part si on a $ad - bc = 0$ la matrice A n'est pas inversible car sinon on obtiendrait $\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = 0.I_2 = 0_{22}$ et donc $A = 0_{22}$ ce qui serait contradictoire.

Un intérêt clair du calcul de l'inverse d'une matrice (quand il existe) est la résolution des systèmes linéaires associés : en effet si A est inversible alors $AX = b$ équivaut à $X = A^{-1}b$. Il est néanmoins rare que l'on ait recours a cette méthode : tout d'abord elle ne s'adapte qu'à des cas particuliers et ensuite il existe une méthode permettant de traiter tous les cas et qui de plus est beaucoup plus pratique et performante du point de vue algorithmique (et qui fournit l'inverse quand il existe).

7.2 LA MÉTHODE DU PIVOT

On décrit une procédure de résolution des systèmes linéaires : l'idée est de se ramener à un système triangulaire que l'on peut ensuite résoudre simplement.

Définition: Une *opération élémentaire* sur les lignes d'une matrice consiste à :

– Remplacer une ligne L_i par la ligne $L_i + L_j$ (avec $i \neq j$).
– Echanger deux lignes.
– Multiplier une ligne par un scalaire non nul.

Remarque : en combinant ces opérations, on voit qu'on peut remplacer L_i par $L_i + \alpha L_j$.

Montrons sur un exemple que ces opérations permettent de simplifier considérablement une matrice ; indiquons par une flèche le fait de passer à une autre matrice (par une opération élémentaire).

$$A = \begin{pmatrix} 1 & 0 & 2 & 3 & 4 \\ 1 & 1 & -1 & 1 & 0 \\ 2 & 1 & -3 & 7 & 2 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 2 & 3 & 4 \\ 0 & 1 & -3 & -2 & -4 \\ 2 & 1 & -3 & 7 & 2 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 2 & 3 & 4 \\ 0 & 1 & -3 & -2 & -4 \\ 0 & 1 & -7 & 1 & -6 \end{pmatrix} \rightarrow$$
$$\begin{pmatrix} 1 & 0 & 2 & 3 & 4 \\ 0 & 1 & -3 & -2 & -4 \\ 0 & 0 & -4 & 3 & -2 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 2 & 3 & 4 \\ 0 & 1 & -3 & -2 & -4 \\ 0 & 0 & 1 & -3/4 & 1/2 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 2 & 3 & 4 \\ 0 & 1 & 0 & -17/4 & -5/2 \\ 0 & 0 & 1 & -3/4 & 1/2 \end{pmatrix}$$
$$\rightarrow \begin{pmatrix} 1 & 0 & 0 & 9/2 & 3 \\ 0 & 1 & 0 & -17/4 & -5/2 \\ 0 & 0 & 1 & -3/4 & 1/2 \end{pmatrix}$$

Le grand intérêt, du point de vue de la résolution des systèmes linéaires, est le suivant :

THÉORÈME: *Considérons le système linéaire*

$$(*) \quad AX = b$$

Supposons que l'on passe de la matrice $A_1 := (\, A \mid b \,)$ à la matrice $A_2 := (\, A' \mid b' \,)$ par une succession d'opérations élémentaires, alors les solutions du système linéaire :

$$(**) \quad A'X = b'$$

sont les mêmes que celles du système (*).

Démonstration: Echanger l'ordre de deux équations ou en multiplier une, par un scalaire non nul, ne change pas les solutions ; passer de deux équations $L_1 = L_2 = 0$ à deux autres équations $L_1 + \alpha L_2 = L_2 = 0$ n'en change pas les solutions, d'où l'énoncé. □

Exemple : Le système d'équations

$$\begin{cases} x_1 & +2x_3 & +3x_4 & = 4 \\ x_1 & +x_2 & -x_3 & +x_4 & = 0 \\ 2x_1 & +x_2 & -3x_3 & +7x_4 & = 2 \end{cases}$$

possède les mêmes solutions que le système :

$$\begin{cases} x_1 & & & -\frac{9}{2}x_4 & = 3 \\ & x_2 & & -\frac{17}{4}x_4 & = -\frac{5}{2} \\ & & x_3 & -\frac{3}{4}x_4 & = \frac{1}{2} \end{cases}$$

Or résoudre ce dernier système est immédiat ; on donne une valeur arbitraire à x_4 et on en tire :

$$\begin{cases} x_1 & = \frac{9}{2}x_4 + 3 \\ x_2 & = \frac{17}{4}x_4 - \frac{5}{2} \\ x_3 & = \frac{3}{4}x_4 + \frac{1}{2} \end{cases}$$

En vue de donner une autre preuve du théorème, interprétons les opérations élémentaires en termes de matrices : soit

$$E_{k,\ell} := \begin{pmatrix} 1 & & & & \\ & . & & & \\ & & \alpha & & \\ & & . & & \\ & & & . & \\ & & & & 1 \end{pmatrix}$$

la matrice dont tous les coefficients sont nuls sauf ceux situés sur la diagonale qui valent 1 et le coefficient ℓk (ℓ-ème ligne et k-ème colonne) qui vaut α. Calculons EA en appelant a_{ij} les coefficients de A et c_{ij} ceux de EA. Un calcul direct donne que si $i \neq \ell$ alors $c_{ij} = a_{ij}$ alors que $c_{\ell j} = a_{\ell j} + \alpha a_{kj}$, donc :

Multiplier à gauche par E revient à transformer A en ajoutant à sa ℓ-ème ligne α fois sa k-ème ligne.

On vérifiera que :
Multiplier à gauche par

$$E_{i,j} := \begin{pmatrix} 1 & & & & & \\ & . & & & & \\ & & 0 & & 1 & \\ & & & 1 & & \\ & & 1 & & 0 & \\ & & & & & 1 \end{pmatrix}$$

la matrice A revient à échanger les i-ème et j-ème lignes de A.

Exemple :

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} = \begin{pmatrix} d & e & f \\ a & b & c \end{pmatrix}$$

Multiplier à gauche par

$$E_i(\alpha) := \begin{pmatrix} 1 & & & & \\ & . & & & \\ & & \alpha & & \\ & & & . & \\ & & & & 1 \end{pmatrix}$$

la matrice A revient à multiplier la i-ème ligne de A par α.

En résumé on voit que l'on passe d'une matrice A_1 à une matrice A_2 par une suite d'opérations élémentaires si il existe une suite de matrices E_1,E_2, ..., E_r chacune de l'un des trois types précédents telles que $A_2 = E_1E_2 \dots E_rA_1$. En particulier, si l'on revient aux systèmes linéaires, l'égalité $(A' \mid b') = E_1E_2 \dots E_r(A \mid b)$ équivaut à $A' = E_1E_2 \dots E_rA$ et $b' = E_1E_2 \dots E_rb$ donc les systèmes $AX = b$ et $A'X = b'$ sont équivalents.

Voyons maintenant quelle est la forme la plus simple que l'on puisse obtenir pour une matrice (ou un système linéaire).

Définition: Une matrice est *échelonnée* si

1) Le premier coefficient non nul d'une ligne est 1 (on dit que c'est un *pivot*).

2) Le premier coefficient non nul de la $(i+1)$-ème ligne est à droite de celui de la i-ème ligne.

3) Les coefficients au dessus d'un pivot sont nuls.

Exemple :

$$\begin{pmatrix} 1 & 2 & 0 & 0 & 4 \\ 0 & 0 & 1 & -2 & 9 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} 1 & 0 & 0 & 5 & 1 \\ 0 & 1 & 0 & -2 & 1 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

sont échelonnées.

Remarques : Une ligne est soit nulle soit de la forme $(0, \dots, 0, 1, *, \dots, *)$; si une ligne est nulle, toutes les lignes situées en dessous sont nulles.

Indiquons comment, en pratique, on peut appliquer des transformations élémentaires à n'importe quelle matrice pour la rendre échelonnée : on choisit un coefficient non nul situé le plus à gauche possible ; par multiplication par un scalaire et échange des lignes on se ramène au cas ou ce coefficient est 1 et est situé sur la première ligne :

$$A' = \begin{pmatrix} 0 & 0 & 1 & * & * \\ 0 & 0 & * & * & * \\ 0 & 0 & * & * & * \end{pmatrix}$$

En ajoutant à chacune des lignes un multiple adéquat de la première ligne on fait apparaître des zéros en dessous du pivot 1 de la première ligne. On répète l'opération sans toucher la première ligne et au bout d'un certain temps on arrive à une matrice du type :

$$A'' = \begin{pmatrix} 0 & 0 & 1 & * & * & * & * & * \\ 0 & 0 & 0 & 1 & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 1 & * & * \end{pmatrix}$$

Et il ne reste qu'à faire apparaître des zéros au dessus de chaque pivot, ce qui peut se faire en retranchant un multiple adéquat de la ligne du pivot.

Il nous reste à discuter de la résolution des systèmes échelonnés.

THÉORÈME: *Soit $AX = b$ un système d'équation tel que la matrice $(A|b)$ soit échelonnée.*

1) Le système possède une solution si et seulement si il n'y a pas de pivot sur la dernière colonne.

2) Si il n'y a pas de pivot sur la dernière colonne, la solution générale du système s'obtient en fixant arbitrairement chacun des x_i tels que la i-ème colonne ne contienne pas de pivot et en calculant chacun des autres x_i en fonction de ceux-là grâce à l'équation correspondant à la i-ème ligne.

Exemples : La matrice

$$(A|b) = \begin{pmatrix} 1 & 2 & 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & 3 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

possède un pivot sur la dernière colonne et le système associé :

$$\begin{cases} x_1 & +2x_2 & & +x_4 & +2x_5 & = 0 \\ & & x_3 & +3x_4 & +4x_5 & = 0 \\ & & & & 0 & = 1 \end{cases}$$

n'a évidemment pas de solution.

La matrice

$$(A|b) = \begin{pmatrix} 1 & 0 & 0 & 1 & 2 & 2 \\ 0 & 0 & 1 & 3 & 4 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

ne possède pas de pivot sur la dernière colonne et le système associé a pour solutions $x_1 = 2 - x_4 - 2x_5$, $x_3 = 1 - 3x_4 - 4x_5$, les coordonnée x_2,x_4 et x_5 des solutions étant arbitraires.

COROLLAIRE: *Un système linéaire homogène (c'est-à-dire $b = 0$) avec m équations, n inconnues et $n > m$ possède au moins une solution non nulle.*

Démonstration: En effet un système homogène possède toujours une solution : le vecteur nul ; ensuite le décompte des équations et variables entraîne que, une fois le système rendu échelonné, une variable au moins va pouvoir prendre n'importe quelle valeur. □

APPLICATION: Calcul de l'inverse d'une matrice par opération élémentaires :

Pour calculer l'inverse d'une matrice A carrée $n \times n$ on réduit par opération élémentaires la matrice $(A \mid I_n)$ à une matrice échelonnée ; si la matrice A est inversible la matrice échelonnée obtenue sera $(I_n \mid A^{-1})$.

Démonstration: La seule matrice carrée inversible et échelonnée est l'identité, donc la réduction de la matrice A est de la forme suivante : $E_r \dots E_1 A = I$ avec E_i des matrices correspondant à des opérations élémentaires. La réduction à une matrice échelonnée de $(A \mid I)$ est donc $E_r \dots E_1(A \mid I) = (I \mid E_r \dots E_1)$ mais l'identité $E_r \dots E_1 A = I$ signifie précisément que $E_r \dots E_1 = A^{-1}$. □

Exemple : soit $A := \begin{pmatrix} 5 & 2 & 1 \\ 1 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}$ on peut calculer son inverse a l'aide des opérations :

$$\begin{pmatrix} 5 & 2 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 2 & 0 & 1 & 0 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 0 & -3 & 1 & 1 & -5 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & -2 & 1 & 0 & -2 & 1 \end{pmatrix} \to$$

$$\begin{pmatrix} 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & -1/2 & 0 & 1 & -1/2 \\ 0 & -3 & 1 & 1 & -5 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & -1/2 & 0 & 1 & -1/2 \\ 0 & 0 & -1/2 & 1 & -2 & -3/2 \end{pmatrix} \to$$

$$\begin{pmatrix} 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & -1/2 & 0 & 1 & -1/2 \\ 0 & 0 & 1 & -2 & 4 & 3 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 & 1 & -2 & -1 \\ 0 & 1 & 0 & -1 & 3 & 1 \\ 0 & 0 & 1 & -2 & 4 & 3 \end{pmatrix}$$

Ainsi l'on a :

$$A^{-1} = \begin{pmatrix} 1 & -2 & -1 \\ -1 & 3 & 1 \\ -2 & 4 & 3 \end{pmatrix}$$

ce que l'on peut d'ailleurs vérifier directement.

Remarque : si en échelonnant la matrice $(A|I)$ on trouve une matrice $(B|C)$ avec $B \neq I$ alors on peut conclure que A n'est pas inversible. En effet il suffit d'observer que les matrices élémentaires sont toutes inversibles (permuter deux fois deux lignes revient à ne rien changer, de même ajouter puis retrancher un multiple d'une ligne ou multiplier puis diviser par un scalaire non nul).

Exemple : cherchons par cette méthode l'inverse, s'il existe de $A = \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix}$

$$\begin{pmatrix} 1 & 3 & 1 & 0 \\ 2 & 6 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 3 & 1 & 0 \\ 0 & 0 & -2 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 3 & 1 & 0 \\ 0 & 0 & 1 & -\frac{1}{2} \end{pmatrix} \to \begin{pmatrix} 1 & 3 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & -\frac{1}{2} \end{pmatrix}$$

donc A n'est pas inversible.

Hamilton William Rowan (1805–1865)

CHAPITRE 8 ESPACES VECTORIELS

Le petit prince demanda "Dessine-moi un espace vectoriel". Le pilote commença par le dessin suivant :

et lui commenta : "Voila, les vecteurs c'est comme ça, on les note avec des flèches au-dessus, on peut les ajouter, les multiplier par un nombre".

Le petit prince, un peu étonné, lui répondit, non sans quelques raisons : "Mais pourquoi parler de vecteurs pour ajouter des nombres? Et mettre des flèches, c'est un peu enfantin, non? Tu crois que je ne saurai pas reconnaître un vecteur sans sa flèche?"

Un peu vexé, Antoine essaya un deuxième dessin en se justifiant ainsi : "C'était un espace vectoriel de dimension 1, en voila un de dimension 2 ; et je t'ai enlevé les flèches!"

L'enfant contempla le nouveau dessin et ajouta un peu perplexe : "Ecoute, ta planète n'est déjà pas vraiment plate, mais la mienne ne l'est vraiment pas !". L'aviateur grommela "Bon d'accord" et dessina ceci :

Le petit prince contempla avec plaisir ce troisième dessin mais au bout d'un moment il s'inquiéta : "Mais ... on ne peut pas voir les mouvements ... le temps ... est-ce-qu'on ne pourrait pas rajouter une dimension?". Carrément agacé, l'homme fit le dessin suivant :

et ajouta : "Tiens! J'ai mis dans ce coffre un cours d'algèbre linéaire où tu trouveras des espaces de dimension 4 et aussi de dimension n et peut-être même (je ne me souviens plus) des exemples d'espaces vectoriels de dimension infinie".

Le petit prince s'en fut, très content.

8.1 INTRODUCTION À L'ALGÈBRE LINÉAIRE

L'exemple type de l'espace vectoriel, au niveau de ce cours, est l'espace $\mathbf{R}^n$. Nous noterons le plus souvent $e = (x_1, \ldots, x_n)$ un élément (un *vecteur*) de $\mathbf{R}^n$ et nous appellerons

coordonnées du vecteur e les nombres réels x_i. On utilisera toutefois aussi, de temps en temps, la notation en vecteur colonne $e = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$.

On dispose de deux opérations fontamentales sur $\mathbf{R}^n$:
l'addition :
si $e = (x_1, \ldots, x_n)$ et $f = (y_1, \ldots, y_n)$ alors

$$e + f := (x_1 + y_1, \ldots, x_n + y_n)$$

la multiplication par un scalaire :
si $e = (x_1, \ldots, x_n)$ et $\alpha \in \mathbf{R}$ alors

$$\alpha.e := (\alpha x_1, \ldots, \alpha x_n)$$

Bien entendu on ne peut "dessiner" cet espace que lorsque $n = 1, 2$ ou 3. Nous ne formalisons les propriétés d'un espace vectoriel que dans le paragraphe suivant, mais sur cet exemple les propriétés sont bien connues.

Pourquoi considérer une dimension quelconque? Une première (et mauvaise?) réponse est que les mathématiciens aiment bien travailler dans la plus grande généralité possible et que ce sont eux qui décident du contenu des cours de première année. Une deuxième (et meilleure?) réponse est que presque tous les problèmes de la vie courante moderne conduisent à des espaces de dimension plus grande que 4. Nous ne pensons pas à l'espace-temps de dimension 4 dans la théorie d'Einstein, mais à n'importe quelle gestion de banque ou d'entreprise. Une entreprise qui évalue ses stocks va devoir noter la quantité de liquidités, de bien immobiliers, de machines, des produits stockés ou fabriqués. Prenons un petit exemple : un agriculteur produit des pommes de terre, du tournesol, du fourrage, du blé et des olives. Pour noter sa production annuelle (disons en tonnes) il a besoin d'un vecteur avec cinq coordonnées $p = (x_1, \ldots x_5)$. Si la production des mêmes produits par un autre agriculteur est $p' = (x'_1, \ldots x'_5)$ alors leur production totale sera représentée par la somme des deux vecteurs p et p'. Si l'on veut la production en kilo du premier, elle sera donnée par le vecteur $1000p = (1000x_1, \ldots 1000x_5)$.

Le bureau des douanes (ministère du commerce extérieur) comptabilise les importations/exportations de plusieurs milliers de produits (ayant chacun plusieurs paramètres : prix, code du pays, etc). Il est clair que la manipulation de telles données se fait sur ordinateur et qu'il y a donc besoin de procédures mathématiques (et de personnel sachant les utiliser!).

Quels sont les problèmes posés qu'on veut résoudre? L'algèbre linéaire est un outil précieux pour l'étude de la géométrie. Nous aborderons ceci à partir d'exemples. L'archétype du problème d'algèbre linéaire est un système linéaire :

Exemple (simple) :

On dispose d'un budget de 200F pour préparer une sangria en mélangeant deux proportions de jus de fruit pour une proportion de vin ; sachant que le vin coûte 20F le litre et le jus de fruit 10F le litre combien de litres de sangria pourra-t-on préparer?

Appelons x le nombre de litres de vin et y le nombre de litres de jus de fruits, le problème se traduit par le système d'équations :

$$\begin{cases} y = 2x \\ 20x + 10y = 200 \end{cases}$$

d'où l'on tire aisément $20x + 10(2x) = 200$ et donc $x = 5$ et $y = 10$. On pourra donc préparer 15 litres.

Un système linéaire général à n inconnues et m équations s'écrit :

$$\begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n = b_1 \\ \ldots \ \ldots \ \ldots \\ a_{m1}x_1 + \ldots + a_{mn}x_n = b_m \end{cases}$$

(où les a_{ij} et les b_i sont des constantes et les x_j sont les inconnues). Si n et m sont grands (de l'ordre de quelques milliers par exemple) on ne peut résoudre "à la main" ces systèmes et on a développé au chapitre précédent un algorithme pour les traiter. Les systèmes linéaires ont tous une structure similaire : si l'on considère le système obtenu en remplaçant les b_i par 0, qu'on appelle le *système linéaire homogène associé* :

$$\begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n = 0 \\ \ldots \ \ldots \ \ldots \\ a_{m1}x_1 + \ldots + a_{mn}x_n = 0 \end{cases}$$

on voit que les solutions forment un espace vectoriel contenu dans $\mathbf{R}^n$ (la définition formelle n'est donnée qu'au paragraphe suivant) : la somme de deux solutions est encore une solution, le produit par un scalaire d'une solution est encore une solution. De plus si l'on connaît une solution particulière du système de départ, toutes les autres sont sommes de la solution particulière et d'une solution du système homogène.

Exemple : une solution du système

$$\begin{cases} x + y + z = 1 \\ x - y + 2z = 0 \\ 3x - y + 5z = 1 \end{cases}$$

est donnée par $(2, 0, -1)$ (vérification directe) alors que le système homogène

$$\begin{cases} x + y + z = 0 \\ x - y + 2z = 0 \\ 3x - y + 5z = 0 \end{cases}$$

a pour solutions les vecteurs de la forme $(-3t, t, 2t)$ avec $t \in \mathbf{R}$. Ainsi les solutions du système de départ sont toutes données par $(-3t + 2, t, 2t - 1)$ quand t varie.

Ces équations linéaires définissent aussi les objets géométriques simples comme les droites, les plans :

Dans le plan $\mathbf{R}^2$, l'équation $ax + by = c$ (où a, b, c sont des constantes) définit une droite (sauf si $a = b = 0$). Dans l'espace $\mathbf{R}^3$ l'équation $ax + by + cz = d$ (où a, b, c, d sont

des constantes) définit un plan (sauf si $a = b = c = 0$) mais il faut deux équations pour définir une droite. Toutefois deux équations linéaires dans $\mathbf{R}^3$ ne définissent pas toujours une droite : par exemple les équations $2x - y + 3z = 2$, $-6x + 3y - 9z = -6$ définissent un plan.

8.2 ESPACES VECTORIELS

Ce paragraphe contient la formalisation de la notion d'espace vectoriel. On ne gagne rien à supposer que le corps de base est toujours $\mathbf{R}$*, on travaillera donc sur un corps* K*, mais dans les exemples, on pourra supposer* $K = \mathbf{R}$*. Si le goût du lecteur l'incline vers les situations concrètes ou le souci du "À quoi ça sert?", il pourra méditer le fait que la théorie des codes de télécommunication utilise largement les espaces vectoriels sur les corps finis* $\mathbf{Z}/p\mathbf{Z}$*.*

Définition: Un *espace vectoriel* sur un corps K est un ensemble E, muni de deux lois $(+, .)$, la loi $+$ étant une application de $E \times E$ vers E, la loi . ("multiplication par un scalaire") étant une application $K \times E$ vers E, telles que :

1) $(E, +)$ est un groupe commutatif (avec élément neutre noté 0 ou 0_E).

2) $\forall a \in K,\ x, y \in E\ \ a.(x + y) = a.x + a.y$

3) $\forall a, b \in K,\ x \in E\ \ (a + b).x = a.x + b.x$

3) $\forall a, b \in K,\ x \in E\ \ (ab).x = a.(b.x)$

4) $\forall x \in E\ \ 1.x = x$

Convention : On omettra très souvent le point notant la multiplication par un scalaire. On dira souvent un K-espace vectoriel au lieu d'un espace vectoriel sur K

Exemples : on vérifie (immédiatement) que $(K^n, +, .)$ (où l'addition et la multiplication par un scalaire sont définies comme au paragraphe précédent pour $\mathbf{R}^n$) est un espace vectoriel.

- L'ensemble des solutions d'un système linéaire homogène est un espace vectoriel (comme c'est un sous-ensemble de K^n l'addition et la multiplication par un scalaire sont déjà définies).

- L'ensemble des matrices $m \times n$ est un espace vectoriel.

- Le corps des complexes $\mathbf{C}$ est un espace vectoriel de deux façons (au moins) : c'est d'abord un espace vectoriel sur lui-même, mais c'est aussi un espace vectoriel sur $\mathbf{R}$.

- L'espace des fonctions de $\mathbf{R}$ dans $\mathbf{R}$ est un $\mathbf{R}$-espace vectoriel : si f, g sont des fonctions et α un réel, on pose $(f + g)(x) := f(x) + g(x)$ et $(\alpha f)(x) := \alpha f(x)$.

- L'espace des fonctions continues de $\mathbf{R}$ dans $\mathbf{R}$ est un $\mathbf{R}$-espace vectoriel. En effet la somme de deux fonctions continues est continue de même que le produit par une constante.

- L'espace des fonctions $f(t)$ deux fois dérivables de $\mathbf{R}$ dans $\mathbf{R}$ et vérifiant l'équation différentielle :

$$f''(t) + cos(t)f'(t) + 3tf(t) = 0$$

est un $\mathbf{R}$-espace vectoriel. En effet la somme de deux fonctions deux fois dérivables et vérifiant l'équation différentielle est encore deux fois dérivable et vérifie l'équation différentielle (voir le chapitre 20 pour l'utilisation de cet exemple).

- L'ensemble $K[X]$ des polynômes forme un K-espace vectoriel.

- L'ensemble des suites $u := (u_n)_{n\in\mathbf{N}}$ à valeurs réelles forme un espace vectoriel : si $u := (u_n)_{n\in\mathbf{N}}$ et $v := (v_n)_{n\in\mathbf{N}}$ sont des suites et α un réel, on pose $u + v := (u_n + v_n)_{n\in\mathbf{N}}$ et $\alpha u := (\alpha u_n)_{n\in\mathbf{N}}$

- L'ensemble des suites réelles $u := (u_n)_{n\in\mathbf{N}}$ vérifiant la relation de récurrence :

$$u_n + (2n+1)u_{n-1} + n^2 u_{n-2} = 0$$

forme un espace vectoriel. En effet la somme de deux telles suites et le produit par une constante est encore une suite du même type (voir le chapitre 12 pour l'utilisation de cet exemple).

On peut aussi à partir d'espaces vectoriels donnés en fabriquer d'autres.

Définition: Un sous-ensemble F d'un espace vectoriel E est un *sous-espaces vectoriel* si muni des deux lois de E (restreintes à F) il devient un espace vectoriel.

Cela signifie donc que la loi d'addition est une loi interne de $F \times F$ dans F, que pour tout $\lambda \in \mathbf{R}$ et $x \in F$ on a $\lambda x \in F$ et que F muni de ces deux opérations vérifie les axiomes d'un espace vectoriel.

PROPOSITION: *Soit $F \subset E$, alors F est un sous-espace vectoriel de E si et seulement si on a :*
(i) $0_E \in F$
(ii) $\forall \alpha, \beta \in K,\ \forall x, y \in F,\ \alpha x + \beta y \in F$

Démonstration: En effet cette condition équivaut au fait que l'addition et la multiplication par un scalaire (dans E) définissent bien deux lois $F \times F \to F$ et $K \times F \to F$ et les axiomes d'espace vectoriel sont alors vérifiés puisqu'ils le sont dans E. □

Exemples : Une droite passant par l'origine, un plan passant par l'origine sont des sous-espaces vectoriels de $\mathbf{R}^3$. Si $x \in E$ (et $x \neq 0$) alors l'ensemble des vecteurs αx (où α parcourt K) est un sous-espace vectoriel : c'est "la droite engendrée par x".

Comme application, on peut observer que l'intersection de deux sous-espaces vectoriels est encore un sous-espace vectoriel mais que l'union n'est pas un sous-espace vectoriel (sauf si l'un des deux espaces contient l'autre).

Si F, F' sont des sous-espaces vectoriels de E on peut définir leur *somme* comme :

$$F + F' := \{f + f' \mid f \in F \text{ et } f' \in F'\}$$

On dira que les deux sous-espaces vectoriels sont en *somme directe* si $F \cap F' = \{0\}$ et on notera, dans ce cas la somme $F \oplus F'$.

Si S est un sous-ensemble de E, on peut définir l'*espace vectoriel engendré par S* comme la somme des droites engendrées par les éléments de S. Par exemple si $e := (1, 0, 0)$,

$f := (0,1,0)$, $g := (1,1,0)$ et $h := (0,0,1) \in \mathbf{R}^3$ alors l'espace vectoriel engendré par $\{e,f,g\}$ est un plan alors que l'espace vectoriel engendré par $\{e,f,h\}$ est l'espace $\mathbf{R}^3$ tout entier. Les sous espaces engendrés par $\{e,f\}$ et par $\{f,g\}$ ne sont pas en somme directe mais les sous espaces engendrés par $\{e,f\}$ et par $\{h\}$ sont en somme directe.

Notation : Soit $e_1,\ldots,e_r$ des vecteurs d'un espace vectoriel E. L'ensemble des combinaisons linéaires des $e_1,\ldots,e_r$ est un sous-espace vectoriel de E, on le note :

$$\mathrm{Vect}(e_1,\ldots,e_r) = \{x = \lambda_1 e_1 + \ldots + \lambda_r e_r \mid \lambda_1,\ldots,\lambda_r \subset \mathbf{R}\}$$

Si E,F sont des espaces vectoriels, on peut définir de manière naturelle une structure d'espaces vectoriel sur le produit :

$$\forall x,x' \in E,\ y,y' \in F :\ \ (x,y)+(x',y') := (x+x',y+y') \quad \text{et} \quad \alpha(x,y) := (\alpha x, \alpha y)$$

Par exemple $\mathbf{R}^m \times \mathbf{R}^n$ s'identifie à $\mathbf{R}^{m+n}$.

8.3 BASES ET DIMENSION

Dans tout ce paragraphe on travaille dans un espace vectoriel E sur un corps K que l'on pourra prendre égal à $\mathbf{R}$. Ce paragraphe donne une définition précise de la notion de dimension d'un espace vectoriel et est fondamental pour la suite. Le point clef est qu'un espace vectoriel E (de dimension finie) est toujours isomorphe à K^n et plus précisément qu'il existe n vecteurs $e_1,\ldots,e_n$ tels que tout vecteur de E s'écrivent de façon unique comme combinaison linéaire des e_i. Ceci généralise la notion de repère du plan (deux vecteurs) et de l'espace (trois vecteurs).

Commençons par un peu de vocabulaire :

Une *famille* de vecteurs est une suite (finie dans la pratique) de vecteurs indexés par un ensemble I (dans la pratique on prend souvent $I := \{1,\ldots,n\}$) ; on note $(e_i)_{i\in I}$ une telle famille.

Une *combinaison linéaire* des vecteurs e_i est un vecteur x de la forme $x := \sum_{i\in I} \alpha_i e_i$ avec $\alpha_i \in K$; si l'ensemble I est infini on impose que pour tout $i \in I$, sauf un nombre fini, on ait $\alpha_i = 0$.

Définition: 1) Une famille $(e_i)_{i\in I}$ de vecteurs est *libre* si la seule combinaison linéaire nulle des e_i est obtenue en prenant tous les α_i nuls ; elle est *liée* si elle n'est pas libre.

2) Une famille $(e_i)_{i\in I}$ de vecteurs est *génératrice* de l'espace vectoriel E si tout vecteur de E est une combinaison linéaire des e_i.

3) Une famille $(e_i)_{i\in I}$ est une *base* si elle est libre et génératrice.

Exemples : 1) Une famille contenant le vecteur nul 0 ou bien contenant deux vecteurs égaux est liée. Plus généralement une famille est liée si et seulement l'un de ses vecteurs est combinaison linéaire des autres vecteurs.

2) Prenons $E := \mathbf{R}^3$ et

$$e_1 := \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix},\ e_2 := \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix},\ e_3 := \begin{pmatrix} 2 \\ -1 \\ 5 \end{pmatrix},\ e_4 := \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$$

On vérifie les relations : $e_3 = e_1 - 3e_2$ et $2e_4 = e_2 + e_3$ ainsi la famille $\{e_1, e_2, e_3, e_4\}$ n'est pas libre, la famille $\{e_1, e_2, e_3\}$ non plus. Par contre la famille $\{e_1, e_2\}$ est libre ; en effet si $\alpha_1 e_1 + \alpha_2 e_2 = 0$ alors $\alpha_1 = \alpha_1 + \alpha_2 = 2\alpha_1 - \alpha_2 = 0$ ce qui entraîne $\alpha_1 = \alpha_2 = 0$. Géométriquement cela traduit que les deux vecteurs e_1, e_2 ne sont pas situés sur une même droite : ils ne sont pas colinéaires ; par contre les quatre vecteurs e_1, e_2, e_3, e_4 sont tous situés dans un même plan P. On peut d'ailleurs vérifier que ce plan est donné par :

$$P = \left\{ \begin{pmatrix} x_1 \\ x_1 \\ x_2 \end{pmatrix} \in \mathbf{R}^3 \mid -2x_1 + x_2 + x_3 = 0 \right\}$$

Considérons maintenant le vecteur $e_5 := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ et vérifions que $\{e_1, e_2, e_5\}$ forme une base de $E = \mathbf{R}^3$. Tout d'abord prouvons que cette partie est libre : pour cela on considère une combinaison linéaire nulle : $\alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_5 = 0$; ceci équivaut à : $2\alpha_1 + \alpha_3 = 0$,$2\alpha_1 + \alpha_2 = 0$ et $4\alpha_1 - \alpha_2 = 0$; en sommant les deux dernières égalités on obtient $3\alpha_1 = 0$ donc $\alpha_1 = 0$ puis $\alpha_2 = \alpha_3 = 0$. Pour montrer (directement) que $\{e_1, e_2, e_5\}$ est génératrice on doit montrer que pour tout vecteur $e := \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ il existe une combinaison linéaire telle que $e = \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_5$. Ceci équivaut à $\alpha_1 + \alpha_3 = a$,$\alpha_1 + \alpha_2 = b$, $2\alpha_1 - \alpha_2 = c$ dont on détermine l'unique solution $\alpha_3 = (3a - b - c)/3$, $\alpha_2 = (2b - c)/3$ et $\alpha_1 = (b + c)/3$.

3) (Base canonique) Il est facile de vérifier que dans K^n, les vecteurs $e_1 := (1, 0, \ldots, 0)$, $e := (0, 1, 0, \ldots, 0)$, $\ldots$,$e_n := (0, \ldots, 0, 1)$ forment une base : en effet tout vecteur $x = (x_1, \ldots, x_n)$ s'écrit de manière unique comme combinaison linéaire des e_i de la façon suivante : $x = x_1 e_1 + \ldots + x_n e_n$. Cette base s'appelle la *base canonique* de K^n.

Commentaires : On voit que si l'on considère une famille finie $\{e_1, \ldots, e_n\}$ et si l'on examine les combinaisons linéaires de ces vecteurs :

1) Si la famille est libre, un vecteur, qui est combinaison linéaire, l'est de manière unique.

2) Si la famille est génératrice, tout vecteur est combinaison linéaire des vecteurs de cette famille, autrement dit $\mathrm{Vect}(e_1, \ldots, e_n)$ est l'espace tout entier.

3) Si la famille est une base, tout vecteur s'écrit de manière unique comme combinaison linéaire. En d'autres termes l'application de K^n vers E donnée par $(\alpha_1, \ldots, \alpha_n) \mapsto \alpha_1 e_1 + \ldots + \alpha_n e_n$ est une bijection.

Nous dirons (provisoirement) qu'un espace vectoriel E est de *dimension finie* si il admet une famille génératrice finie.

Exemple : L'espace vectoriel $K[X]$ n'est pas de dimension finie sur K. En effet si $P_1, \ldots, P_n$ est une famille finie de polynômes, soit $d := \max(deg(P_i))$, alors toute combinaison linéaire des P_i est un polynôme de degré inférieur ou égal à d.

L'espace vectoriel K^n est de dimension finie (heureusement ...) puisque les vecteurs $e_i := (0, \ldots, 1, \ldots, 0)$ (avec un 1 sur la i-ème coordonnée) forment une base. Nous allons maintenant voir que tout espace vectoriel de dimension fini est en fait de ce type.

PROPOSITION: *Soient $e_1, \ldots, e_n$ des vecteurs de E, soient $f_1, \ldots, f_{n+1}$ des combinaisons linéaires des $e_1, \ldots, e_n$, alors la famille des f_j est liée.*

Démonstration: Ecrivons $f_i = \sum_j a_{ij} e_j$, alors $\sum_{i=1}^{n+1} x_i f_i = \sum_j \left(\sum_i a_{ij} x_i\right) e_j$ mais on sait que le système $\sum_i a_{ij} x_i = 0$ pour $j = 1, \ldots, n$ possède une solution non nulle car c'est un système homogène avec plus d'inconnues que d'équations (la démonstration de ce fait a été donnée en étudiant la méthode du pivot au chapitre 7) ; donc celle-ci fournit une combinaison linéaire des f_i qui est nulle. □

COROLLAIRE: *Deux bases ont toujours le même cardinal.*

Démonstration: Soient $e_1, \ldots, e_n$ et $f_1, \ldots, f_m$ deux bases. Si on avait $m > n$, comme les f_j sont des combinaisons linéaires des e_i, on en tirerait que les f_j sont liés, ce qui n'est pas. On a donc établi que $m \leq n$ et par symétrie $n \leq m$ et donc $m = n$. □

Remarque : le corollaire reste vrai si l'espaces est de dimension infinie, mais la preuve est plus difficile.

THÉORÈME: *(de la base incomplète) Soit $\mathcal{L}$ une partie libre de E contenue dans $\mathcal{G}$ une partie génératrice de E. Alors il existe une base $\mathcal{B}$ de E contenant $\mathcal{L}$ et contenue dans $\mathcal{G}$.*

Le nom du théorème provient du fait qu'il indique que l'on peut compléter la partie libre $\mathcal{L}$ en une base (à l'aide d'éléments d'une partie génératrice). On pourrait aussi l'appeler "théorème de la base extraite" puisqu'il dit aussi que l'on peut extraire une base de toute partie génératrice.

Démonstration: Choisissons un sous-ensemble $\{b_1, b_2, \ldots\} = \mathcal{B}$ de $\mathcal{G}$ donnant une partie libre, contenant $\mathcal{L}$ et de cardinal maximal pour ces propriétés ; montrons que c'est bien une base. Il suffit de voir que c'est une partie génératrice ; mais comme tout vecteur est combinaison linéaire des éléments de $\mathcal{G}$, il suffit de voir qu'un vecteur de $\mathcal{G}$ est une combinaison linéaire des vecteurs de $\mathcal{B}$. Soit donc e un vecteur de $\mathcal{G}$ alors la famille $\mathcal{B} \cup \{e\}$ est liée et la relation de dépendance linéaire $\beta e + \sum \alpha_i b_i = 0$ ne peut pas correspondre à $\beta = 0$ (sinon les $b_i \in \mathcal{B}$ seraient liés) et permet donc d'exprimer e comme combinaison linéaire des vecteurs de $\mathcal{B}$. □

On en déduit immédiatement :

COROLLAIRE: *Tout espace vectoriel de dimension finie admet une base (finie).*

Démonstration: On prend une partie génératrice finie et on peut en extraire une base. □

Toutes les bases ayant même cardinal, on peut définir :

Définition: Le cardinal d'une base de E s'appelle la *dimension* de E. On le note $\dim(E)$.

Définition: La dimension de l'espace vectoriel engendré par un système de vecteurs $\{u_i \mid i \in I\}$ s'appelle le *rang* du système. En d'autres termes le rang de $\{u_i \mid i \in I\}$ est $\dim \mathrm{Vect}(u_i \mid i \in I)$.

Ainsi un espace est de dimension n si l'une de ses bases (et donc toutes) est de cardinal n. La dimension est l'invariant le plus important d'un espace vectoriel. L'espace K^n à une base de cardinal n et est donc bien de dimension n (ouf!) et en fait si E est de dimension n sur K et a pour base $e_1, \ldots, e_n$, l'application $(\alpha_1, \ldots, \alpha_n) \mapsto \alpha_1 e_1 + \ldots + \alpha_n e_n$ est une bijection (et même un isomorphisme dès que nous aurons vu la notion d'application linéaire).

COROLLAIRE: *Soit F un sous-espace vectoriel de E espace vectoriel de dimension finie, alors il existe un autre sous-espace vectoriel F' tel que $E = F \oplus F'$. On dit que F' est un supplémentaire de F dans E.*

Démonstration: On choisit une base de F, disons $e_1, \ldots, e_m$ que l'on complète en une base de E, disons $e_1, \ldots, e_n$; alors le sous-espace vectoriel F' engendré par $e_{m+1}, \ldots, e_n$ répond à la question. □

Un supplémentaire n'est pas unique ; un décompte des cardinaux des diverses bases montre que la dimension du supplémentaire est indépendante du choix du supplémentaire et vaut $\dim(E) - \dim(F)$.

COROLLAIRE: *Supposons que E est de dimension n alors :*

1) Toute famille libre est de cardinal au plus n (avec égalité si et seulement si c'est une base).

2) Toute famille génératrice est de cardinal au moins n (avec égalité si et seulement si c'est une base).

Démonstration: 1) Une partie libre peut être complétée en une base de cardinal n et a donc un cardinal $\leq n$ avec égalité si il n'y a pas besoin de compléter, i.e. si c'est une base.

2) Si une partie est génératrice, on peut trouver un sous-ensemble qui est une base de cardinal n, donc le cardinal est au moins n avec égalité si la partie génératrice est une base. □

En particulier, dans un espace de dimension n, le rang d'un système $u_1, \ldots, u_m$ est toujours inférieur ou égal à $\min(m, n)$, il est égal à m si et seulement si le système est libre et il est égal à n si et seulement si le système est générateur.

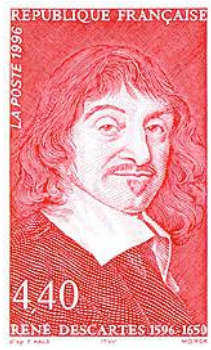
COROLLAIRE: *1) Si $F \subset E$ alors $\dim F \leq \dim E$, avec égalité si et seulement si $E = F$.*

2) $\dim(E \oplus F) = \dim E + \dim F$

3) $\dim(E \times F) = \dim E + \dim F$

Démonstration: Ces trois énoncés peuvent se démontrer en dénombrant des bases de chacun des espaces vectoriels. Ainsi pour 1) on complète une base $e_1, \ldots, e_m$ de F en une base $e_1, \ldots, e_n$ de E et bien sûr $m \leq n$, l'égalité ayant lieu si $E = F$. Pour 2) on constate que l'union d'une base de E et d'une base de F est disjointe et donne une base de $E \oplus F$.

Pour 3) on observe que si $e_1, \ldots, e_m$ est une base de E et $f_1, \ldots, f_n$ est une base de F alors $(e_1, 0), \ldots, (e_m, 0), (0, f_1), \ldots, (0, f_n)$ fournit une base de $E \times F$. $\square$



Descartes René (1596–1650)

CHAPITRE 9 APPLICATIONS LINÉAIRES

On formalise dans ce chapitre le deuxième concept abstrait fondamental en algèbre linéaire : celui d'application linéaire ; le lecteur connait déjà un bon nombre d'exemples : hormis les applications définies par des matrices, on étudie au lycée des rotations, symétries, projections, homothéties que l'on retrouve ici.

9.1 THÉORÈME DE LA DIMENSION

Définition: Soient E, F des K-espaces vectoriels, une application $u : E \to F$ est *K-linéaire* (ou *linéaire* si le contexte est clair) si elle vérifie :

$$\forall e_1, e_2 \in E,\ \forall \alpha_1, \alpha_2 \in K, \quad u(\alpha_1 e_1 + \alpha_2 e_2) = \alpha_1 u(e_1) + \alpha_2 u(e_2)$$

Exemple : si $E = K^n$ et $F = K^m$ alors toute matrice A $m \times n$ définit une application $X \mapsto AX$ (où X est un vecteur colonne) de E vers F qui est linéaire puisque $A(\alpha_1 X_1 + \alpha_2 X_2) = \alpha_1 AX_1 + \alpha_2 AX_2$.

On utilise fréquemment la linéarité de la dérivation, de l'intégrale ; traduit en termes d'espaces vectoriels cela signifie que, par exemple l'application $f \mapsto f'$ qui a une fonction associe sa dérivée est linéaire (de l'espace des fonctions dérivables dans l'espace des fonctions), que l'application $f \mapsto \int_a^b f(t)dt$ est linéaire (de l'espace des fonctions continues dans $\mathbf{R}$ disons). Voyons comment u agit sur les sous-espaces vectoriels :

PROPOSITION: *Soient E, F des K-espaces vectoriels et u une application linéaire de E vers F, l'image d'un sous-espace vectoriel de E par u est un sous-espace vectoriel de F, l'image réciproque d'un sous-espace vectoriel de F par u est un sous-espace vectoriel de E.*

Démonstration: Soit E' un sous-espace vectoriel de E, soient $y_1, y_2 \in u(E')$ alors il existe $x_1, x_2 \in E'$ tels que $y_i = u(x_i)$ donc $\alpha_1 y_1 + \alpha_2 y_2 = u(\alpha_1 x_1 + \alpha_2 x_2) \in u(E')$; donc $u(E')$ est un sous-espace vectoriel. Par ailleurs si F' est un sous-espace vectoriel de F et si $x_1, x_2 \in u^{-1}(F')$ alors $u(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 u(x_1) + \alpha_2 u(x_2) \in F'$ donc $\alpha_1 x_1 + \alpha_2 x_2 \in u^{-1}(F')$ et $u^{-1}(F')$ est bien un sous-espace vectoriel. □

En particulier le *noyau* de u, l'ensemble $\mathrm{Ker}(u) := \{x \in E \mid u(x) = 0\}$ est un sous-espace vectoriel de E et l'*image* de E par u, l'ensemble $\mathrm{Im}(u) := \{y \in F \mid \exists x \in E,\ u(x) = y\}$ est un sous-espace vectoriel de F. On vérifie aisément que :

- Une application linéaire u est injective si et seulement si son noyau est nul : $\mathrm{Ker}(u) = \{0\}$

- Une application linéaire u est surjective si et seulement si l'image d'une partie génératrice est génératrice.

Définition: Un *isomorphisme d'espaces vectoriels* est une application linéaire $u : E \to F$ qui est bijective.

Remarque : il est immédiat que si u est bijective et linéaire, la bijection réciproque est aussi linéaire et que u transforme une base de E en une base de F.

Exemples : Soit $u(x,y) := 2x + y$ de $\mathbf{R}^2$ vers $\mathbf{R}$ alors le noyau est la droite donnée par l'équation $2x + y = 0$ et l'image est $\mathbf{R}$ entier.

Soit $v(x,y,z) = (2x + y, 3y + z, z + y + x, x + y)$ de $\mathbf{R}^3$ dans $\mathbf{R}^4$, alors un petit calcul fournit que $\mathrm{Ker}(v) = \{0\}$ et que l'image est l'hyperplan d'équation $-3X - Y + Z + 5T = 0$.

Une rotation, une symétrie, une homothétie de rapport non nul est bijective (noyau nul et image égale à l'espace entier). Une projection se fait parallèlement à son noyau et sur son image.

Rotation dans le plan

Rotation dans l'espace

Symétrie dans le plan

Symétrie dans l'espace

Projection sur une droite
parallèlement à un plan

Projection sur un plan
parallèlement à une droite

Les rotations et les symétries sont des isomorphismes alors que les projections ne sont pas surjectives et ont un noyau non nul. On peut observer sur ces exemples la loi générale sur les dimensions que l'on énonce maintenant :

THÉORÈME: *Soient E, F des K-espaces vectoriels et u une application linéaire de$E$ vers F, alors*

$$\dim E = \dim \mathrm{Im}(u) + \dim \mathrm{Ker}(u)$$

Démonstration: Soit E' un supplémentaire de $\mathrm{Ker}(u)$ (i.e. $E' \oplus \mathrm{Ker}(u) = E$) alors la restriction de u à E' avec pour but $F' := \mathrm{Im}(u)$ est un isomorphisme $u' : E' \to F'$: en effet elle est surjective car si $y \in F'$ alors il existe $x \in E$ tel que $u(x) = y$ mais $x = x' + x''$ avec $x' \in E'$ et $x'' \in \mathrm{Ker}(u)$ donc $y = u(x) = u'(x')$; par ailleurs $\mathrm{Ker}(u') = \mathrm{Ker}(u) \cap E' = \{0\}$ donc u' est injective. On a donc $\dim E' = \dim \mathrm{Im}(u)$ et $\dim \mathrm{Ker}(u) + \dim E' = \dim E$ d'où le théorème. □

COROLLAIRE: $\dim u(E) \leq \dim E$ *(avec égalité si et seulement si u est injective).*

Démonstration: En effet $\dim u(E) = \dim E - \dim \mathrm{Ker}(u) \leq \dim E$ avec égalité si et seulement si $\mathrm{Ker}(u)$ est nul c'est-à-dire si u est injective. □

COROLLAIRE: *Soient E, F des K-espaces vectoriels et u une application linéaire de E vers F, supposons $\dim E = \dim F = n$, alors u est bijective si et seulement si elle est injective ou bien si et seulement si elle est surjective.*

Démonstration: L'application u est surjective si et seulement si $\dim \mathrm{Im}(u) = n$ donc si et seulement si $\dim \mathrm{Ker}(u) = 0$ c'est-à-dire si u est injective. □

COROLLAIRE: *Soient E et F des sous-espaces vectoriels d'un espace vectoriel G.*

$$\dim E + \dim F = \dim(E + F) + \dim(E \cap F)$$

Démonstration: Considérons l'application linéaire $u : E \times F \to E + F$ donnée par $(x, y) \mapsto x + y$. Elle est surjective par construction et son noyau est le sous-espace vectoriel $\{(x, y) \in E \times F \mid x + y = 0\} = \{(x, -x) \mid x \in E \cap F\}$ qui est isomorphe à $E \cap F$. On en déduit que $\dim E + \dim F = \dim E \times F = \dim(E + F) + \dim(E \cap F)$. □

9.2 MATRICE D'UNE APPLICATION LINÉAIRE

Un moyen simple de construire des applications linéaires est de fixer l'image des éléments d'une base de l'espace de départ $u(e_1), \ldots, u(e_n)$. Si l'on dispose d'une base de l'espace d'arrivée on peut exprimer les vecteur $u(e_i)$ comme combinaisons linéaires des vecteurs de la base. On arrive à une description de u par un ensemble de nombres qui se range naturellement dans un tableau, i.e. une matrice. Cette construction est l'inverse de celle qui a une matrice A associe l'applicaton linéaire $X \mapsto AX$ et est extrêmement utile pour l'étude des applications linéaires et des matrices.

Définition: Soient $e_1, \ldots, e_n$ une base de E et $f_1, \ldots, f_m$ une base de F et $u : E \to F$ une application linéaire, alors $u(e_j) := a_{1j} f_1 + \ldots + a_{mj} f_m$ et on définit la *matrice de u dans les bases $e = e_1, \ldots, e_n$, $f = f_1, \ldots, f_m$* comme :

$$Mat(u; (e), (f)) := (a_{ij})$$

Cette correspondance entre matrices et application linéaire est si naturelle que la composition d'application linéaires correspond à la multiplication des matrices :

THÉORÈME: *Soient $e_1, \ldots, e_n$ une base de E, $f_1, \ldots, f_m$ une base de F et $g_1, \ldots g_k$ une base de G, soient $u : E \to F$ et $v : F \to G$ des applications linéaires, alors*

$$Mat(v \circ u; (e), (g)) = Mat(v; (f), (g)) Mat(u; (e), (f))$$

Démonstration: Notons $B := Mat(v;(f),(g)) = (b_{ij})$ c'est-à-dire $v(f_j) = \sum_{i=1}^{k} b_{ij} g_i$ et $A := Mat(u;(e),(f)) = a_{ij}$ c'est-à-dire $u(e_j) = \sum_{i=1}^{m} a_{ij} f_i$ et enfin $BA = (c_{ij})$ donc $c_{ij} = \sum_{h=1}^{m} b_{ih} a_{hj}$. On peut calculer :

$$v \circ u(e_j) = v\left(\sum_{h=1}^{m} a_{hj} f_h\right) = \sum_{h=1}^{m} a_{hj} \left(\sum_{i=1}^{k} b_{ih} g_i\right) = \sum_{i=1}^{k} \left(\sum_{h=1}^{m} b_{ih} a_{hj}\right) g_i$$

□

Remarque : soit $e_1, \ldots, e_n$ la base canonique de $\mathbf{R}^n$ et $f_1, \ldots, f_m$ la base canonique de $\mathbf{R}^m$ alors les vecteurs colonnes de la matrice de $u : \mathbf{R}^n \to \mathbf{R}^m$ sont les vecteurs $u(e_1), \ldots, u(e_n)$.

Exemples : (on choisit la base canonique de $\mathbf{R}^2$ dans les exemples qui suivent)

La matrice d'une rotation u d'angle θ est donnée par $R_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$. En effet $u(e_1) = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}$ et $u(e_2) = \begin{pmatrix} -\sin(\theta) \\ \cos(\theta) \end{pmatrix}$.

La matrice de la symétrie par rapport à la droite faisant un angle $\theta/2$ avec l'axe des x est donnée par $\begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}$. En effet $u(e_1) = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}$ et $u(e_2) = \begin{pmatrix} \sin(\theta) \\ -\cos(\theta) \end{pmatrix}$.

La matrice de la projection sur l'axe des x parallèlement à l'axe des y est $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$.

La matrice d'une homothétie de rapport α est $\begin{pmatrix} \alpha & 0 \\ 0 & \alpha \end{pmatrix}$.

APPLICATION: Formules d'addition de cosinus et sinus :

En appliquant le théorème précédent à la composition de deux rotations d'angles θ et θ' on obtient $R_{\theta+\theta'} = R_\theta R_{\theta'}$ et donc :

$$\begin{pmatrix} \cos(\theta+\theta') & -\sin(\theta+\theta') \\ \sin(\theta+\theta') & \cos(\theta+\theta') \end{pmatrix} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} \cos(\theta') & -\sin(\theta') \\ \sin(\theta') & \cos(\theta') \end{pmatrix} =$$

$$= \begin{pmatrix} \cos(\theta)\cos(\theta') - \sin(\theta)\sin(\theta') & -\cos(\theta)\sin(\theta') - \cos(\theta')\sin(\theta) \\ \cos(\theta)\sin(\theta') + \cos(\theta')\sin(\theta) & \cos(\theta)\cos(\theta') - \sin(\theta)\sin(\theta') \end{pmatrix}$$

ce qui donne bien les formules d'addition de cosinus et sinus :

$$\cos(\theta+\theta') = \cos(\theta)\cos(\theta') - \sin(\theta)\sin(\theta') \text{ et } \sin(\theta+\theta') = \cos(\theta)\sin(\theta') + \cos(\theta')\sin(\theta)$$

9.3 CHANGEMENTS DE BASES

Reprenons l'exemple de la symétrie par rapport à une droite dont la matrice dans la base canonique est $\begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}$; si l'on choisit comme base les vecteurs f_1, f_2 comme indiqué ci-dessous on obtient comme matrice pour la symétrie $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

On se propose dans ce paragraphe de systématiser les calculs de ces changements de base. Le but est le plus souvent de trouver des bases dans lesquelles la matrice de l'application étudiée est la plus simple possible.

Définition: Soient $e_1, \ldots, e_n$ et $e'_1, \ldots, e'_n$ deux bases de E, la *matrice de passage* de e à e' est la matrice $n \times n$ dont la j-ème colonne est formée des composantes de e'_j dans la base $e_1, \ldots, e_n$.

En d'autres termes $P(e, e') = (a_{ji})$ avec $e'_j = \sum_{i=1}^{n} a_{ij} e_i$.

Commentaires : 1) La matrice $P = P(e, e')$ est aussi $Mat(id, (e'), (e))$ la matrice de l'application identité de E (muni de la base $e'_1, \ldots, e'_n$) vers E (muni de la base $e_1, \ldots, e_n$).

2) Si on pose $P' = P(e', e)$ alors $PP' = Id$ ou encore $P^{-1} = P(e', e)$.

3) Le sens d'écriture est bien sûr une convention ; l'important est d'être cohérent (on roule à gauche en Angleterre).

THÉORÈME: *Soit $u : E \to F$ une application linéaire et soient $e_1, \ldots, e_n$, $e'_1, \ldots, e'_n$ deux bases de E et $f_1, \ldots, f_m$, $f'_1, \ldots, f'_m$ deux bases de F ; notons A la matrice de u dans les bases e et f et A' la matrice de u dans les bases e' et f' ; notons encore $P = P(e, e')$ et $Q = P(f, f')$ les matrices de passage de e vers e' (resp. de f vers f'), alors :*

$$A' = Q^{-1}AP$$

Démonstration: Considérons la suite d'applications :

$$(E, e') \overset{id_E}{\to} (E, e) \overset{u}{\to} (F, f) \overset{id_F}{\to} (F, f')$$

On peut en déduire l'égalité de matrices :

$$Mat(u, (e'), (f')) = Mat(id_E, (f), (f'))Mat(u, (e), (f))Mat(id_E, (e'), (e))$$

ce qui correspond bien à l'égalité $A' = Q^{-1}AP$. $\square$

Exemples :

1) Considérons l'application linéaire u de $\mathbf{R}^2$ dans $\mathbf{R}^2$ dont la matrice dans la base canonique e_1, e_2 est $\begin{pmatrix} 4 & -3 \\ 2 & -1 \end{pmatrix}$ et choisissons $e'_1 := \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ et $e'_2 := \begin{pmatrix} 3 \\ 2 \end{pmatrix}$. La matrice de passage de e vers e' est $P = \begin{pmatrix} 1 & 3 \\ 1 & 2 \end{pmatrix}$ et son inverse se calcule facilement $P^{-1} = \begin{pmatrix} -2 & 3 \\ 1 & -1 \end{pmatrix}$. Par ailleurs $u(e'_1) = e'_1$ et $u(e'_2) = 2e'_2$ donc la matrice de u dans la base e' est $\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$; on vérifiera qu'on a bien

$$P^{-1}\begin{pmatrix} 4 & -3 \\ 2 & -1 \end{pmatrix} P = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

2) Soit $u : E \to F$, on peut choisir des bases "bien adaptées" à u ainsi : pour E on choisit $e_1, \ldots, e_n$ de sorte que $e_{r+1}, \ldots, e_n$ forment une base du noyau de u et $e_1, \ldots, e_r$ forment une base d'un supplémentaire du noyau. Les vecteurs $f_1 = u(e_1), \ldots, f_r = u(e_r)$ sont linéairement indépendants et on peut les compléter en une base $f_1, \ldots, f_m$ de F ; la matrice de u s'écrit alors :

$$Mat(u; (e), (f)) = \begin{pmatrix} 1 & 0 & & & \\ 0 & . & & & \\ 0 & & 1 & & \\ 0 & & & 0 & \\ 0 & & & & 0 \end{pmatrix} = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

En particulier on voit que pour toute matrice A il existe P, Q inversibles telles que $Q^{-1}AP$ soit de la forme précédente. L'entier r est le *rang* de la matrice ou de l'application linéaire, c'est-à-dire la dimension de l'image. 3) Soit e_1, e_2, e_3 la base canonique de $\mathbf{R}^3$ et soit u la rotation d'angle $2\pi/3$ d'axe $e_1 + e_2 + e_3$. Choisissons $f_3 = \frac{1}{\sqrt{3}}(e_1 + e_2 + e_3)$ et ensuite f_1, f_2 formant une base orthonormé du plan orthogonal à f_3 de sorte que f_1, f_2, f_3 forment une base directe de l'espace. Pour fixer les idées on peut prendre $f_1 = \frac{1}{\sqrt{2}}(e_1 - e_2)$ et $f_2 = \frac{1}{\sqrt{6}}(e_1 + e_2 - 2e_3)$. Alors

$$Mat(u; (f)) = \begin{pmatrix} \cos(2\pi/3) & -\sin(2\pi/3) & 0 \\ \sin(2\pi/3) & \cos(2\pi/3) & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{-1}{2} & \frac{-\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & \frac{-1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

et, en posant $P := P(e, f) = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{3} \\ \frac{-\sqrt{2}}{2} & \frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{3} \\ 0 & \frac{-\sqrt{6}}{3} & \frac{\sqrt{3}}{3} \end{pmatrix}$, on obtient

$$Mat(u; (e)) = P\begin{pmatrix} \cos(2\pi/3) & -\sin(2\pi/3) & 0 \\ \sin(2\pi/3) & \cos(2\pi/3) & 0 \\ 0 & 0 & 1 \end{pmatrix} P^{-1} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

(exercice : vérifier que u est une isométrie du cube $\mathcal{C} := [-1, 1]^3$).

CHAPITRE 10 INTRODUCTION AUX DÉTERMINANTS

On définit et étudie dans ce chapitre le déterminant de n vecteurs $e_1, \ldots, e_n$ dans $\mathbf{R}^n$; c'est un nombre réel (ou un élément de K si on travaille dans K^n). L'interprétation géométrique de la valeur absolue de ce déterminant est simple et importante : c'est le "volume" du parallélipipède engendré par les vecteurs $e_1, \ldots, e_n$:

Dans $\mathbf{R}^2$, on a $|\det(e_1, e_2)|$ = aire hachurée

Dans $\mathbf{R}^3$, on a $|\det(e_1, e_2, e_3)|$ = volume du parallélipipède

Avec cette interprétation, il est clair que le déterminant s'annule si et seulement si le parallélipipède est "plat", c'est-à-dire si les vecteurs $e_1, \ldots, e_n$ sont liés. Le signe du déterminant est plus subtil et correspond à l'orientation de l'espace, concept que nous n'élaborerons pas (voir néanmoins le chapitre suivant) ; voici le signe sur des exemples :

Les déterminants fournissent également un critère théorique (et parfois pratique) pour calculer le rang d'une matrice et résoudre certains systèmes linéaires ; c'est l'application qui est exposée ici. La construction du déterminant est assez abstraite et pourra être survolée en première lecture.

10.1 FORMES n-LINÉAIRES ALTERNÉES

Définition: Soit E un K-espace vectoriel et $n \geq 1$, une application $f : E \times \ldots \times E \to K$ est *n-linéaire* si elle est linéaire par rapport à chaque variable, c'est-à-dire si :
$f(x_1, \ldots, x_{i-1}, \alpha y_i + \beta z_i, x_{i+1}, \ldots, x_n) = \alpha f(x_1, \ldots, x_{i-1}, y_i, x_{i+1}, \ldots, x_n) +$
$\beta f(x_1, \ldots, x_{i-1}, z_i, x_{i+1}, \ldots, x_n)$

Elle est *symétrique* si elle est invariante par permutation des facteurs et *antisymétrique* ou *alternée* si l'interversion de deux facteurs change le signe, c'est-à-dire si

$$\forall x_1, \ldots, x_n \in E, \forall i < j,\ f(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_n) = -f(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_n)$$

Remarque : on déduit immédiatement que si f est alternée, alors $f(\ldots, x, \ldots, x, \ldots) = 0$ et puis plus généralement que si x_i est une combinaison linéaire des autres vecteurs $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$ alors $f(x_1, \ldots, x_n) = 0$.

Exemples : Le produit scalaire $\mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ est bilinéaire et symétrique. L'application de $\mathbf{R}^2 \times \mathbf{R}^2$ vers $\mathbf{R}$ définie par $((a,b),(c,d)) \mapsto ad - bc$ est bilinéaire alternée.

Il y a assez peu de formes alternées comme le prouve l'énoncé suivant :

PROPOSITION: *Si* $\dim(E) = n$*, l'ensemble des formes* n*-linéaires alternées est un espace vectoriel de dimension 1. Si* $e_1, \ldots, e_n$ *forment une base de* E *et si* f *est une forme* n*-linéaire alternée non nulle, alors* $f(e_1, \ldots, e_n) \neq 0$*.*

Démonstration: (la démonstration n'est compliquée que par les notations) Soient $x_1, \ldots, x_n$ des vecteurs de E ; écrivons leurs coordonnées dans la base $e_1, \ldots, e_n$ ainsi : $x_i = \sum_{j=1}^{n} a_{ij} e_j$ et donc

$$f(x_1, \ldots, x_n) = \sum_{j_1, \ldots, j_n} \left(\prod_{i=1}^{n} a_{ij_i} \right) f(e_{j_1}, \ldots, e_{j_n})$$

mais $f(e_{j_1}, \ldots, e_{j_n}) = 0$ si l'un des j_i apparaît deux fois et si les j_i sont une permutation de $1, 2, \ldots n$ alors $f(e_{j_1}, \ldots, e_{j_n}) = \varepsilon f(e_1, \ldots, e_n)$ où ε est le signe de la permutation en question. En définitive on obtient, en changeant un peu les notations :

$$f(x_1, \ldots, x_n) = \left(\sum_{s \in \mathcal{S}_n} \varepsilon(s) \left(\prod_{i=1}^{n} a_{is(i)} \right) \right) f(e_1, \ldots, e_n)$$

On voit donc que toutes les formes n-linéaires alternées sont proportionnelles et que f est non nulle si et seulement si $f(e_1, \ldots, e_n) \neq 0$. □

Exemple : Si f est n-linéaire alternée sur $\mathbf{R}^n$, on peut calculer explicitement :

si $n = 2$:

$f((a,b),(c,d) = acf((1,0),(1,0)) + adf((1,0),(0,1)) + cbf((01,),(1,0)) + bdf((0,1),(0,1)) = (ad - bc)f((1,0),(0,1))$

Si $n = 3$, on vérifiera par un calcul similaire la "règle de Sarrus" :

$$f((a,b,c),(d,e,f),(g,h,i)) = (aei + bfg + cdh - afh - bdi - ceg) f((1,0,0),(0,1,0),(0,0,1))$$

On définit maintenant le déterminant de n vecteurs dans un espace de dimension n.

Définition: Soit $e_1, \ldots, e_n$ une base de E, le *déterminant* par rapport à la base $e_1, \ldots, e_n$ est l'unique application n-linéaire alternée $\det : E^n \to K$ telle que $\det(e_1, \ldots, e_n) = 1$.

Exemple : Si $E = K^n$ on sous-entend que l'on a choisi comme base la base canonique. On obtient alors (en écrivant les vecteurs en colonnes) :

$$\det\left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix} \right) = ad - bc$$

$$\det\left(\begin{pmatrix} a \\ b \\ c \end{pmatrix}, \begin{pmatrix} d \\ e \\ f \end{pmatrix}, \begin{pmatrix} g \\ h \\ i \end{pmatrix} \right) = aei + bfg + cdh - afh - bdi - ceg$$

La propriété fondamentale du déterminant est la suivante :

THÉORÈME: *Le déterminant de n vecteurs $x_1, \ldots, x_n$ dans K^n est nul si et seulement si ces vecteurs sont liés.*

Démonstration: On a déjà vu que si les vecteurs sont liés le déterminant est nul ; inversement si les vecteurs $x_1, \ldots, x_n$ sont libres, ils forment une base (puisque K^n est de dimension n) et donc, d'après la proposition précédente, $\det(x_1, \ldots, x_n) \neq 0$, . □

Ayant défini le déterminant de n vecteurs de K^n, on peut facilement définir le déterminant d'une matrice carrée :

Définition: Le déterminant d'une matrice carrée est le déterminant de ses vecteurs colonnes par rapport à la base canonique de K^n.

Le théorème précédent se traduit alors ainsi pour les matrices :

THÉORÈME: *Soit A une matrice carrée, alors A est inversible si et seulement si* $\det(A) \neq 0$.

Démonstration: D'après le théorème précédent, le déterminant de A est non nul si et seulement si les vecteurs colonnes de la matrice A sont linéairement indépendants. Comme les vecteurs colonnes de la matrice A engendre l'image de l'application associée à A, cela équivaut à dire que cette application est surjective ou encore qu'elle bijective, ce qui signifie bien que A est inversible. □

THÉORÈME: *Le déterminant des matrices est multiplicatif, c'est-à-dire : si A et B sont deux matrices carrées de même ordre n, on a :*

$$\det(AB) = \det(A)\det(B)$$

Démonstration: Soient $x_1, \ldots, x_n$ des vecteurs de K^n, les applications n-linéaires alternées de $K^n \times \ldots \times K^n$ vers K définies par $(x_1, \ldots, x_n) \mapsto \det(Ax_1, \ldots, Ax_n)$ et $(x_1, \ldots, x_n) \mapsto \det(x_1, \ldots, x_n)$ sont proportionnelles d'après la première proposition de ce chapitre, donc $\det(Ax_1, \ldots, Ax_n) = \alpha \det(x_1, \ldots, x_n)$. En appliquant ceci aux vecteurs de la base canonique on obtient que $\alpha = \det(A)$ et ainsi :

$$\det(Ax_1, \ldots, Ax_n) = \det(A)\det(x_1, \ldots, x_n)$$

Maintenant on peut calculer :

$$\begin{aligned}\det(AB)\det(x_1, \ldots, x_n) &= \det(ABx_1, \ldots, ABx_n) \\ &= \det(A)\det(Bx_1, \ldots, Bx_n) \\ &= \det(A)\det(B)\det(x_1, \ldots, x_n)\end{aligned}$$

d'où la formule $\det(AB) = \det(A)\det(B)$, puisque $det(x_1, \ldots, x_n) \neq 0$. □

Les déterminants peuvent servir à calculer le rang d'une matrice de taille quelconque : le point est que l'on peut extraire d'une matrice des matrices carrées en ne gardant que les coefficients correspondant à certaines lignes et colonnes.

Définition: Soit $r \leq \min(m,n)$, on appelle *mineurs* d'ordre r d'une matrice $m \times n$ les déterminants d'une matrice extraite de taille $r \times r$.

Exemples : Les mineurs d'ordre 1 sont simplement les coefficients de la matrice ; les mineurs d'ordre 2 d'une matrice $A = (a_{ij})$ sont les expressions $M_{i,j;k,l} = a_{ik}a_{jl} - a_{jk}a_{il}$; quelques mineurs d'ordre trois de la matrice $A := \begin{pmatrix} 1 & 2 & 3 & 0 \\ 0 & 1 & 2 & 3 \\ 1 & 1 & 1 & 4 \\ 5 & 4 & 3 & 1 \end{pmatrix}$ sont donnés par

$$\det \begin{pmatrix} 1 & 3 & 0 \\ 1 & 1 & 4 \\ 5 & 3 & 1 \end{pmatrix} = 46, \quad \det \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix} = 0, \quad \det \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 3 \\ 5 & 4 & 1 \end{pmatrix} = 10$$

THÉORÈME: *Soit A une matrice $m \times n$ alors $Rang(A) = r$ si et seulement si :*
i) On peut extraire un mineur non nul d'ordre r de la matrice A.
ii) Tous les mineurs de A d'ordre $r+1$ sont nuls.

Démonstration: Si un mineur d'ordre r est non nul, alors il y a r vecteurs colonnes indépendants donc le rang est au moins r. Inversement pour prouver qu'une matrice de rang r contient un mineur d'ordre r, il est plus simple d'utiliser que le rang ainsi que l'existence d'un mineur de taille donnée non nul ne change pas par opérations élémentaires sur les lignes où les colonnes. On est amené alors à prouver l'énoncé pour les matrices de la forme $\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}$, ce qui est immédiat. □

Exemple : Le déterminant de la matrice A de taille 4×4 donnée ci-devant est nul et l'un (au moins) de ses mineurs d'ordre 3 est non nul, par conséquent cette matrice est de rang 3.

Donnons maintenant une des applications classiques des déterminants aux systèmes linéaires.

THÉORÈME: *(Systèmes de Cramer) Considérons un système linéaire carré :*

$$\begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n = b_1 \\ \qquad \ldots \quad \ldots \quad \ldots \\ a_{n1}x_1 + \ldots + a_{nn}x_n = b_n \end{cases}$$

dont la matrice associée est inversible ; alors il a pour unique solution

$$x_i = \frac{\det \begin{pmatrix} a_{11} & \ldots & b_1 & \ldots & a_{1n} \\ \ldots & \ldots & & \ldots & \ldots \\ a_{n1} & \ldots & b_n & \ldots & a_{1n} \end{pmatrix}}{\det \begin{pmatrix} a_{11} & \ldots & a_{1i} & \ldots & a_{1n} \\ \ldots & \ldots & & \ldots & \ldots \\ a_{n1} & \ldots & a_{ni} & \ldots & a_{1n} \end{pmatrix}}$$

Démonstration: Appelons $C_1, \ldots, C_n$ les vecteurs colonnes de la matrice associée au système linéaire et b le vecteur colonne de coordonnées b_i ; alors le système se traduit par $\sum_{i=1}^{n} x_i C_i = b$ donc $\det(C_1, \ldots, b, \ldots, C_n) = x_i \det(C_1, \ldots, C_i, \ldots, C_n)$. On obtient l'énoncé annoncé en divisant par $\det(C_1, \ldots, C_i, \ldots, C_n)$. □

10.2 CALCULS DE DÉTERMINANTS

Les théorèmes précédents ont motivé, nous l'espérons, l'apprentissage de quelques techniques de calcul de déterminant.

Remarque : D'après la propriété de multilinéarité du déterminant, les opérations élémentaires modifient de la façon suivante le déterminant :

– Remplacer une colonne C_i par $C_i + \alpha C_j$ ne modifie pas le déterminant.

– Inverser deux colonnes ne change pas la valeur absolue, mais change le signe.

– Multiplier une colonne par α multiplie le déterminant par α.

On est ramené alors au cas d'une matrice triangulaire qui est assez facile :

THÉORÈME: *Le déterminant d'une matrice triangulaire est égal au produit de ses coefficients diagonaux :*

$$\det \begin{pmatrix} a_{11} & * & * & * \\ & a_{22} & * & * \\ & & . & * \\ 0 & & & a_{nn} \end{pmatrix} = a_{11} \ldots a_{nn}$$

Démonstration: Notons $e_1, \ldots, e_n$ les vecteurs de la base canonique de $\mathbf{R}^n$. Ce déterminant est le déterminant des vecteurs $a_{11}e_1 + \ldots$, $a_{22}e_2 + \ldots$, $\ldots$, $a_{nn}e_n$ et est donc aussi le déterminant de $a_{11}e_1$, $a_{22}e_2$, $\ldots$, $a_{nn}e_n$. Ce dernier déterminant est égal à $a_{11} \ldots a_{nn} \det(e_1, \ldots, e_n) = a_{11} \ldots a_{nn}$. □

Une autre technique souvent utile et également basée sur la multilinéarité du déterminant est la suivante :

THÉORÈME: *(développement par rapport à une colonne) On a la formule suivante :*

$$\det \begin{pmatrix} & a_{1j} & \\ & . & \\ & . & \\ & . & \\ & a_{nj} & \end{pmatrix} = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det \begin{pmatrix} & a_{1j} & \\ & . & \\ a_{i1} & a_{ij} & a_{in} \\ & . & \\ & a_{nj} & \end{pmatrix}$$

où la "croix" signifie que l'on a retiré la i-ème ligne et la j-ème colonne.

Démonstration: La j-ème colonne s'écrit $\sum_i a_{ij} e_i$ donc par linéarité on peut se ramener au cas où la colonne est le vecteur e_i. Si l'on ramène la i-ème ligne sur la première ligne,

on multiplie le déterminant par $(-1)^{i+j}$ (compter le nombre de transpositions) et on doit juste vérifier que

$$\det \begin{pmatrix} 1 & ** \\ 0 & A \\ 0 & \end{pmatrix} = \det(A)$$

Ce qui est laissé au lecteur. □

Exemple (calcul de déterminant 3×3 à partir de déterminants 2×2 :

$$\det \begin{pmatrix} 3 & 4 & 6 \\ 8 & 1 & -1 \\ -2 & 0 & 1 \end{pmatrix} = (+3) \det \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} - (8) \det \begin{pmatrix} 4 & 6 \\ 0 & 1 \end{pmatrix} + (-2) \det \begin{pmatrix} 4 & 6 \\ 1 & -1 \end{pmatrix} = -9$$

Lagrange Joseph (1736–1813)

CHAPITRE 11 GÉOMÉTRIE DANS LE PLAN ET L'ESPACE

Ce chapitre reprend les notions d'algèbre linéaire en intégrant les concepts de produit scalaire et d'orthogonalité étudiés au lycée. Outre la description d'exemples d'isométries (rotations, symétries orthogonales, translations) l'illustration principale est le calcul de la distance d'un point à une droite ou un plan. D'un point de vue algébrique, on identifie le "plan" avec $\mathbf{R}^2$ *et l'espace à trois dimensions avec* $\mathbf{R}^3$*. On notera qu'avec ces identifications une "droite" dans le plan possède une équation du type* $ax + by + c$ *et n'est pas un espace vectoriel (sauf si* $c = 0$*); on appellera "droite vectorielle" une droite d'équation* $ax + by = 0$*. De même on parlera de plan d'équation* $ax + by + cz + d = 0$ *dans l'espace* $\mathbf{R}^3$*; un plan vectoriel est un plan d'équation* $ax + by + cz = 0$*.*

11.1 PRODUIT SCALAIRE ET ISOMÉTRIES.

Commençons par rappeler les définitions suivantes dans l'espace à trois dimensions $\mathbf{R}^3$ (nous vous laissons formuler les mêmes définitions dans le plan $\mathbf{R}^2$).

Définition: On appelle *produit scalaire* de deux vecteurs $x, y \in \mathbf{R}^3$ de coordonnées x_1, x_2, x_3, resp. y_1, y_2, y_3 le nombre :

$$(x \cdot y) := x_1y_1 + \ldots + x_ny_n.$$

Les vecteurs x et y sont *orthogonaux* si $(x \cdot y) = 0$.

Définition: On appelle *norme euclidienne* du vecteur x de $\mathbf{R}^3$ le nombre :

$$||x|| := \sqrt{(x \cdot x)} = \sqrt{x_1^2 + x_2^2 + x_3^2}.$$

La *distance euclidienne* entre deux vecteurs x et y de $\mathbf{R}^3$ est définie par la formule :

$$\text{distance}\,(x, y) := d(x, y) = ||x - y|| = \sqrt{(x_1 - y_1)^2 + (x_2 - x_3)^2 + (x_3 - y_3)^2}.$$

Enfin un vecteur x est dit *unitaire* si $||x|| = 1$.

Remarques. a) (Théorème de Pythagore!) Soient $x, y, z \in \mathbf{R}^n$ tels que $x - z$ soit orthogonal à $y - z$ (i.e. les trois points forment un triangle rectangle en z) alors :

$$\begin{aligned} d(x, y)^2 = ||x - y||^2 &= ||(x - z) - (y - z)||^2 = ||x - z||^2 - 2(x - z) \cdot (y - z) + ||y - z||^2 \\ &= ||x - z||^2 + ||y - z||^2 = d(x, z)^2 + d(y, z)^2. \end{aligned}$$

b) Soit x un vecteur non nul, on peut toujours écrire $x = ||x|| \, \frac{x}{||x||}$ donc comme un vecteur unitaire (la direction donnée par x) multiplié par un scalaire positif (sa longueur ou norme).
c) Pour écrire l'équation d'une droite vectorielle D dans le plan, disons $ax + by = 0$, on peut aussi simplement prendre le vecteur u de coordonnée a et b et observer que, si X est le vecteur de coordonnées x, y, on peut écrire l'équation du plan

$$(u \cdot X) = 0.$$

Ainsi la donnée de la droite (vectorielle) équivaut à la donnée d'un vecteur orthogonal à cette droite. Cette remarque permet de passer facilement d'une base de D à une équation.

Les principales propriétés de la norme euclidienne sont résumées dans la proposition suivante (les trois premières propriétés caractérisant les normes).

PROPOSITION: *La norme euclidienne sur* $\mathbf{R}^3$ *vérifie pour tout* $x, y \in \mathbf{R}^3$

(i) *On a* $||x|| \geq 0$ *avec égalité si et seulement si* $x = 0$*;*

(ii) *Pour* $a \in \mathbf{R}$*, on a* $||ax|| = |a|||x||$*;*

(iii) *(inégalité triangulaire)* $||x + y|| \leq ||x|| + ||y||$*. De plus on a l'inégalité de Cauchy-Schwarz :*

$$|(x \cdot y)| := |x_1y_1 + x_2y_2 + x_3y_3| \leq ||x||\,||y||.$$

(iv) distance $(x, y) \geq 0$ *(avec égalité seulement si* $x = y$*)*

(v) distance $(x, y) \leq$ distance (x, z) + distance (z, y).

Remarque. Dans le plan, on sait (et l'on va revoir cela) que, si θ est l'angle entre les deux vecteurs x et y, alors $(x \cdot y) = \cos\theta\,||x||\,||y||$; l'inégalité de Cauchy-Schwarz équivaut donc dans ce cas à $|\cos\theta| \leq 1$.

Démonstration: Les deux premières propriétés sont immédiates, les points (iv) et (v) se déduisent facilement de (i), (ii) et (iii) ; montrons la troisième à l'aide de l'inégalité de Cauchy-Schwarz. On écrit pour cela :

$$\begin{aligned} ||x + y||^2 &= (x_1 + y_1)^2 + (x_2 + y_2)^2 + (x_3 + y_3)^2 = ||x||^2 + 2\,(x \cdot y) + ||y||^2 \\ &\leq ||x||^2 + 2||x||\,||y|| + ||y||^2 = (||x|| + ||y||)^2\,. \end{aligned}$$

Pour démontrer l'inégalité de Cauchy-Schwarz on utilise l'astuce suivante. On calcule, pour $t \in \mathbf{R}$:

$$0 \leq ||x + ty||^2 = ||x||^2 + 2t(x \cdot y) + t^2||y||^2.$$

On remarque alors que, si un polynôme $a + 2bt + ct^2$ est constamment positif, alors $\Delta = b^2 - ac \leq 0$, ce qui s'écrit ici $(x \cdot y)^2 - ||x||^2||y||^2 \leq 0$, d'où l'inégalité cherchée se déduit. □

PROPOSITION: *Soit* D *une droite vectorielle de* $\mathbf{R}^3$*, l'orthogonal de* D*, noté* $D^\perp := \{x \in \mathbf{R}^3 \mid \forall y \in D,\ (x \cdot y) = 0\}$ *est un plan vectoriel supplémentaire de* D *(c'est-à-dire que* $\mathbf{R}^3 = D \oplus D^\perp$*). Soit* Π *une plan vectoriel de* $\mathbf{R}^3$*, l'orthogonal de* Π*, noté* $\Pi^\perp := \{x \in \mathbf{R}^3 \mid \forall y \in \Pi,\ (x \cdot y) = 0\}$ *est une droite vectorielle supplémentaire de* Π *(c'est-à-dire que* $\mathbf{R}^3 = \Pi \oplus \Pi^\perp$*).*

Démonstration: Du point de vue de l'algèbre linéaire, il est immédiat que $D^\perp$ est un sous-espace vectoriel et que $D \cap D^\perp = \{0\}$, il faut donc prouver que $\dim D^\perp = 2$. Pour cela introduisons une base f_1 de D et l'application linéaire $\Phi : \mathbf{R}^3 \to \mathbf{R}$ définie par $\Phi(x) = (f_1 \cdot x)$. Par construction $\mathrm{Ker}\,\phi = D^\perp$; montrons que Φ est surjective et donc que $\dim D^\perp = \dim \mathrm{Ker}\,\Phi = 2$ comme annoncé. Si Φ n'était pas surjective, l'image serait nulle et on aurait, pour tout $x \in \mathbf{R}^3$: $(f_1 \cdot x) = 0$, ce qui entraînerait $f_1 = 0$ et une contradiction. La deuxième affirmation se prouve de façon analogue. □

Remarque. L'orthogonal d'une droite dans le plan $\mathbf{R}^2$ est une droite, on peut résumer cela en disant que si E est un sous-espace, $E \oplus E^\perp$ est l'espace total.

Définition: Une base e_1, e_2, e_3 de $\mathbf{R}^3$ est *orthogonale* si les vecteurs e_i sont deux à deux orthogonaux; la base est *orthonormée* si de plus $||e_i|| = 1$.

Exemple : on voit immédiatement que la base canonique de $\mathbf{R}^3$ est orthonormée.

THÉORÈME: (Procédé d'orthogonalisation de Gram-Schmidt) *Soit u_1, u_2, u_3 une base (quelconque) de $\mathbf{R}^3$, on peut obtenir une base orthogonale v_1, v_2, v_3 de la forme suivante : $v_1 := u_1$, $v_2 = u_2 + au_1$, $v_3 = u_3 + bu_2 + cu_1$ pour certains réels a, b, c. Pour obtenir une base orthonormée, on remplace v_i par $v_i' := v_i/||v_i||$.*

Démonstration: La construction se réalise étape par étape. On pose donc $v_1 = u_1$ ensuite on cherche $a \in \mathbf{R}$ tel que $(u_2 + au_1) \cdot u_1 = 0$ et on trouve donc $a = -(u_2 \cdot u_1)/||u_1||^2$, ainsi on choisit $v_2 := u_2 - \frac{(u_2 \cdot u_1)}{||u_1||^2} u_1$. Ensuite on cherche v_3 de la forme $v_3 = u_3 + \lambda v_2 + \mu v_1$ (il sera bien de la forme $u_3 + bu_2 + cu_1$) tel que $(v_3 \cdot v_2) = (v_3 \cdot v_1) = 0$; on trouve comme conditions $(u_3 \cdot v_2) + \lambda||v_2||^2 = (u_3 \cdot v_1) + \mu||v_1||^2 = 0$ d'où le choix $v_3 = u_3 - \frac{(u_3 \cdot v_2)}{||v_2||^2} v_2 - \frac{(u_3 \cdot v_1)}{||v_1||^2} v_1$. La matrice faisant passer de u_1, u_2, u_3 à v_1, v_2, v_3 est inversible (elle est triangulaire avec des 1 sur la diagonale) donc v_1, v_2, v_3 est bien une base orthogonale. □

Définition: On appelle *isométrie* de $\mathbf{R}^3$ une application $f : \mathbf{R}^3 \to \mathbf{R}^3$ qui préserve la distance, c'est-à-dire telle que pour tout $x, y \in \mathbf{R}^3$ on a $||f(x) - f(y)|| = ||x - y||$.

PROPOSITION: *Soit $u : \mathbf{R}^3 \to \mathbf{R}^3$ une application linéaire, alors les propriétés suivantes sont équivalentes :*
(i) u préserve la distance (i.e. est une isométrie).
(ii) u préserve la norme (i.e.pour tout $x \in \mathbf{R}^3$ on a $||f(x)|| = ||x||$).
(iii) u préserve le produit scalaire (i.e.pour tout $x, y \in \mathbf{R}^3$ on a $(f(x) \cdot f(y)) = (x \cdot y)$).

Démonstration: Si une application linéaire f préserve la distance, elle préserve la norme car $||f(x)|| = d(f(x), 0) = d(f(x), f(0)) = d(x, 0) = ||x||$. Si f préserve la norme elle préserve le produit scalaire car

$$(x \cdot y) = \frac{1}{2} \left\{ ||x + y||^2 - ||x||^2 - ||y||^2 \right\}.$$

Enfin si f préserve le produit scalaire elle préserve la distance puisque $d(f(x), f(y))$ vaut

$$\sqrt{(f(x) - f(y)) \cdot (f(x) - f(y))} = \sqrt{f(x - y) \cdot f(x - y)} = \sqrt{(x - y) \cdot (x - y)} = d(x, y).$$

□

Remarque. On peut démontrer qu'une isométrie s'écrit nécessairement $f(x) = u(x) + a$ avec u isométrie linéaire (en particulier telle que $u(0) = 0$) et $a \in \mathbf{R}^3$. En effet quitte à remplacer $f(x)$ par $f(x) - f(0)$ on peut supposer que $f(0) = 0$. Soit e_1, e_2, e_3 une base orthonormée (par exemple la base canonique) alors les $e_i' := f(e_i)$ forment aussi une base

orthonormée. Soit $x \in \mathbf{R}^3$, on peut l'écrire $x = x_1e_1+x_2e_2+x_3e_3$ et on note que $(x\cdot e_i) = x_i$. Si on écrit de même $f(x) = y_1e'_1 + y_2e'_2 + y_3e'_3$ et on observe que $(f(x) \cdot e'_i) = y_i$ et ainsi :

$$x_i = (x \cdot e_i) = (f(x) \cdot f(e)_i) = (f(x) \cdot e'_i) = y_i,$$

ce qui montre bien que $f(x_1e_1 + x_2e_2 + x_3e_3) = x_1f(e_1) + x_2f(e_2) + x_3f(e_3)$ et que f est linéaire.

Caractérisation matricielle des isométries.

PROPOSITION: *Soit A la matrice d'une application linéaire $\mathbf{R}^3 \to \mathbf{R}^3$, l'application est une isométrie si et seulement si ${}^tAA = I$ ou encore si ses colonnes fournissent une base orthonormée de $\mathbf{R}^3$.*

Démonstration: Le produit scalaire de deux vecteurs (écrits en colonne) X et Y s'écrit en termes de matrices ${}^tXY$ donc si A est la matrice d'une application linéaire de $\mathbf{R}^3$ dans $\mathbf{R}^3$, celle-ci est une isométrie si et seulement si ${}^t(AX)(AY) = {}^tXY$ ou encore ${}^tX({}^tAA)Y = {}^tXY$ pour tout $X, Y \in \mathbf{R}^3$ ce qui entraîne facilement ${}^tAA = I$. Ensuite une autre interprétation d'une isométrie linéaire est une transformation linéaire qui transforme la base canonique en une base orthonormée, ou encore que ses vecteurs colonnes forment une base orthonormée de $\mathbf{R}^3$. □

Remarque : Si A est la matrice d'une isométrie, on observe que $1 = \det({}^tAA) = \det({}^tA)\det(A) = \det(A)^2$ et donc $\det(A) = \pm 1$. Ceci justifie la définition suivante :

Définition: Une base e_1, e_2, e_3 est une base *directe* (resp. *indirecte*) si $\det(e_1, e_2, e_3) > 0$ (resp. si $\det(e_1, e_2, e_3) < 0$). Une isométrie linéaire u est *directe* (resp. *indirecte*) si $\det(u) = +1$ (resp. si $\det(u) = -1$). Une isométrie est directe (resp. indirecte) si sa partie linéaire est directe (resp. indirecte).

Exemples d'isométrie de $\mathbf{R}^2$ ou $\mathbf{R}^3$.

(a) Translation dans $\mathbf{R}^3$. L'application $f(x) = x + a$ est clairement une isométrie, remarquons que, sauf le cas trivial $a = 0$, la transformation n'a aucun point fixe. C'est une isométrie directe.

(b) Rotation dans le plan (c'est une isométrie directe). On peut décrire en coordonnées cartésiennes la rotation f d'angle θ et centre $C = (x_0, y_0)$ en notant $\begin{pmatrix} x' \\ y' \end{pmatrix} = f\begin{pmatrix} x \\ y \end{pmatrix}$:

$$\begin{pmatrix} x' - x_0 \\ y' - y_0 \end{pmatrix} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix}.$$

(c) Symétrie par rapport à une droite dans le plan (resp. par rapport à un plan dans

l'espace). Soit Π un hyperplan de $\mathbf{R}^3$ et y un vecteur unitaire orthogonal à Π; tout vecteur x de $\mathbf{R}^3$ se décompose de façon unique en $x = \lambda y + x_2$ avec $\lambda \in \mathbf{R}$ et $x_2 \in \Pi$. On pose alors

$$f(x) := -\lambda y + x_2.$$

Par exemple la symétrie orthogonale par rapport au plan $(y \cdot x) = 0$ s'écrit :

$$f(x) = x - 2\frac{(y \cdot x)}{||y||^2} x_1.$$

On vérifie que f est bien une isométrie qui est indirecte.

Enfin on peut composer deux isométries ou prendre la bijection inverse (vérifier que le résultat est toujours une isométrie). Une notion un peu plus générale est celle de *similitude* i.e. de transformation de $\mathbf{R}^3$ qui dilate d'un facteur constant les distances, c'est-à-dire telle qu'il existe $\mu > 0$ tel que :

$$\forall x, y \in \mathbf{R}^3, \qquad d(f(x), f(y)) = \mu d(x, y).$$

On montre facilement qu'une similitude est la composée d'une isométrie et d'une homothétie.

11.2 LE PLAN EUCLIDIEN (DIMENSION 2)

Rappel. Nous avons vu que $e^{i\theta} = e^{i\theta'}$ si et seulement si $\theta' = \theta + 2k\pi$ (avec $k \in \mathbf{Z}$), ce qui se traduit par :

$$\begin{cases} \cos(\theta) = \cos(\theta') \\ \sin(\theta) = \sin(\theta') \end{cases} \iff \qquad \theta' = \theta + 2k\pi \qquad (\text{avec } k \in \mathbf{Z})$$

Nous dirons simplement que cosinus et sinus déterminent l'angle à 2π près. Inversement si $a, b \in \mathbf{R}$ vérifient $a^2 + b^2 = 1$ alors il existe un réel θ (unique à 2π près) tel que $\cos(\theta) = a$ et $\sin(\theta) = b$. Introduisons formellement la définition de l'angle.

Définition: Soit u, v deux vecteurs unitaires non nuls du plan euclidien orienté (disons muni de la base canonique), on appelle *angle de u vers v* le réel θ (unique modulo 2π) tel que v s'obtienne en appliquant la rotation $\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$. L'angle entre deux vecteurs non nuls u et v est l'angle entre $u/||u||$ et $v/||v||$.

Nous avons déjà vu deux exemples d'isométries linéaires : les rotations d'angle θ et les symétries orthogonales par rapport à une droite vectorielle, on peut montrer que ce sont les seules (on donne l'énoncé sous forme matricielle).

THÉORÈME: *Soit A une matrice orthogonale (la matrice d'une isométrie) et soit $\epsilon = \pm 1$ son déterminant, alors il existe $\theta \in \mathbf{R}$ (unique modulo 2π) tel que*

$$A = \begin{pmatrix} \cos(\theta) & -\epsilon \sin(\theta) \\ \sin(\theta) & \epsilon \cos(\theta) \end{pmatrix}.$$

Ainsi A correspond soit à la rotation d'angle θ, soit à la symétrie orthogonale par rapport à la droite vectorielle faisant un angle $\theta/2$ avec l'axe des x.

Démonstration: Posons $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Les colonnes sont des vecteurs de normes 1 donc on peut écrire $a = \cos(\theta)$, $c = sin(\theta)$ et $b = \cos(\phi)$, $d =: sin(\phi)$ et le produit scalaire des deux vecteurs colonne est nul, soit $\cos(\theta)\cos(\phi) + \sin(\theta)\sin(\phi) = 0$ ou encore $\cos(\phi - \theta) = 0$. On conclut donc que $\phi = \epsilon\frac{\pi}{2} + \theta$ modulo 2π (avec $\epsilon = \pm 1$) et ainsi $b = \cos(\epsilon\frac{\pi}{2} + \theta) = -\epsilon \sin(\theta)$ et $d = \sin(\epsilon\frac{\pi}{2} + \theta) = \epsilon \cos(\theta)$, d'où le résultat annoncé. □

rotation d'angle θ : symétrie orthogonale par rapport à D :

A titre d'exercice on pourra en déduire une liste des isométries affines; ce sont : les rotations d'angle θ et centre C, les symétries orthogonales par rapport à une droite affine, les translations et les *glissements* (symétrie orthogonale par rapport à la droite D composée avec une translation parallèle à la droite D).

Définition: On appelle *distance d'un point P à une droite D* (dans le plan) la borne inférieure des distances du point avec les points de la droite. En symbole :

$$d(P, D) := \inf_{Q \in D} d(P, Q).$$

En particulier on voit que $P \in D$ entraîne que $d(P, D) = 0$. Voici le calcul en général :

THÉORÈME: *Soit D une droite du plan ayant pour équation $ax + by + c = 0$ et $P = (x_0, y_0)$ un point du plan; la distance de P à D est donnée par*

$$d(P, D) = \frac{|ax_0 + by_0 + c|}{\sqrt{a^2 + b^2}}.$$

De plus il existe un unique point $Q \in D$ tel que $d(P, D) = d(P, Q)$ et il peut être caractérisé par le fait que la droite PQ est orthogonale à D.

Démonstration: Soit u un vecteur orthogonal à la droite D, la droite passant par P et de direction u rencontre D en un unique point Q. Commençons par montrer que Q

est l'unique point qui réalise la distance $d(P, D)$. Soit en effet un autre point $Q' \in D$, par construction le vecteur PQ (qui est proportionnel à u) est orthogonal à QQ' et donc d'après le "théorème de Pythagore" :

$$d(P, Q')^2 = d(P, Q)^2 + d(Q, Q')^2 \geq d(P, Q)^2$$

avec égalité si et seulement si $Q = Q'$. Calculons maintenant $d(P, Q) = d(P, D)$. Ecrivons pour cela l'équation de D ainsi : $(x_1 \cdot z) + c = 0$ et notons z_0 le vecteur correspondant à P. On peut aussi écrire $z = z_0 + \lambda x_1$ et on aura $||z - z_0|| = |\lambda| \cdot ||x_1||$. En écrivant que $z \in D$, on obtient $0 = (x_1 \cdot z) + c = (x_1 \cdot z_0) + c + \lambda ||x_1||^2$ d'où $|\lambda| = \frac{|(x_1 \cdot z_0) + c|}{||x_1||^2}$. On en tire $d(P, Q) = ||z - z_0|| = |(x_1 \cdot z_0) + c|/||x_1||$, ce qui est le résultat cherché. □

11.3 L'ESPACE EUCLIDIEN (DIMENSION 3)

L'observation de la géométrie en dimension trois devrait nous apparaître naturelle. Il faut toutefois faire attention à certains phénomènes liés aux deux orientations possibles de l'espace, ainsi la question naïve suivante est en fait dépourvue de sens : "dans quel sens tourne la terre autour du soleil?". Nous commencerons par y définir une notion d'angle entre deux vecteurs dont on verra qu'elle est différente de celle du plan (orienté).

Le problème de définir et calculer l'angle de deux vecteurs dans l'espace peut sembler simple : on se dit qu'il suffit de considérer le plan engendré par les deux vecteurs et de calculer l'angle dans le plan... Observons cependant sur le dessin ci-dessous ce qui se passe si l'on se déplace dans l'espace.

On s'aperçoit que l'on peut donner à l'angle plusieurs valeurs : θ, $-\theta$ ou $2\pi - \theta$... En particulier on peut toujours faire un choix tel que l'angle soit situé dans l'intervalle $[0, \pi]$.

Définition: L'*angle entre u et v dans l'espace* (N.B. ne pas séparer les mots) est le réel $\theta \in [0, \pi]$ tel qu'il existe une orientation du plan $\langle u, v \rangle$ tel que la rotation d'angle θ envoie u sur v.

Le procédé suivant permet de fabriquer des bases directes ou orthonormées directes. Observons que, si $u, v \in \mathbf{R}^3$ et $x \in \mathbf{R}^3$ (de coordonnées x_1, x_2 et x_3) alors :

$$\det(u, v, x) = Ax + By + Cz = \begin{pmatrix} A \\ B \\ C \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

ce qui permet de définir

Définition: Le *produit vectoriel* de u et v, noté $u \wedge v$ est le vecteur de $\mathbf{R}^3$ tel que, pour tout $x \in \mathbf{R}^3$ on ait :

$$(u \wedge v) \cdot x = \det(u, v, x).$$

On en tire l'expression analytique du produit vectoriel :

$$\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} u_2v_3 - u_3v_2 \\ u_3v_1 - u_1v_3 \\ u_1v_2 - u_2v_1 \end{pmatrix}$$

ainsi que ses principales propriétés :

PROPOSITION: *Le produit vectoriel est bilinéaire (i.e. linéaire en u et en v) et vérifie les formules suivantes.*
(i) $v \wedge u = -u \wedge v$; *en particulier* $u \wedge u = 0$.
(ii) $u \wedge v$ *est orthogonal à* u *et* v.
(iii) Si u, v *sont indépendants et si l'angle de* u *à* v *est* $\theta \in [0, \pi]$ *alors*

$$||u \wedge v|| = \sin(\theta)\, ||u|| \cdot ||v||.$$

(On observera que, comme $\theta \in [0, \pi]$, *on a* $\sin(\theta) = |\sin(\theta)| \geq 0$*).*

Démonstration: La première formule découle du fait que $\det(u, v, x) = -\det(v, u, x)$ et la seconde du fait que $(u \wedge v) \cdot u = \det(u, v, u) = 0$. Pour la démonstration de (iii)on va introduire le *déterminants de Gram* de deux (ou trois ou etc) vecteurs :

$$G(x_1, x_2) := \det \begin{pmatrix} ||x_1||^2 & (x_1 \cdot x_2) \\ (x_1 \cdot x_2) & ||x_2||^2 \end{pmatrix}.$$

Dans le cas de deux vecteurs u, v comme dans l'énoncé, on observera que l'on peut calculer ce dernier ainsi :

$$G(u, v) = ||u||^2||v||^2 - (u \cdot v)^2 = (1 - \cos^2(\theta))||u||^2||v||^2 = \sin^2(\theta)||u||^2||v||^2.$$

Ensuite, posons $w := u \wedge v$ et calculons de deux manières $G(u, v, w)$. D'abord $G(u, v, w) = ||w||^2 G(u, v)$ car w est orthogonal à u et v. Ensuite montrons que

$$G(u, v, w) = \det(u, v, w)^2 \qquad (*)$$

ce qui montrera que

$$||u \wedge v||^2 = (u \wedge v) \cdot w = \det(u, v, w) = \sqrt{G(u, v, w)} = ||w||\sqrt{G(u, v)} = ||w|| \sin(\theta)||u||||v||$$

ce qui est la formule voulue après avoir divisé par $||w||$. Pour prouver la formule $(*)$ on montre la règle de transformation suivante: $G(f(u), f(v), f(w)) = \det(f)^2 G(u, v, w)$; comme $\det(f(u), f(v), f(w)) = \det(f)\det(u, v, w)$ et que la formule $(*)$ est clairement vraie lorsque u, v, w est la base canonique, on en tire qu'elle est toujours vraie. □

COROLLAIRE: *Soit* u, v *deux vecteurs unitaires orthogonaux, posons* $w := u \wedge v$ *alors* u, v, w *forment une base orthonormée directe.*

Démonstration: Les propriétés (i) et (iii) garantissent que u, v, w est une base orthonormée et on a par construction $\det(u, v, w) = (u \wedge v) \cdot w = ||u \wedge v||^2 > 0$. □

Remarque. La donnée d'un plan vectoriel Π dans l'espace, d'équation disons $ax + by + cz = 0$, équivaut à la donnée du vecteur u de coordonnées a, b, c et on peut écrire en abrégé cette équation, en notant X le vecteur de coordonnée x, y, z :

$$(u \cdot X) = 0.$$

Connaissant une base u, v du plan Π on peut en déduire une équation du plan en calculant $w := u \wedge v$; en effet une équation de Π sera donnée par $(w \cdot X) = 0$.

Définition: On appelle *distance d'un point à P un plan* Π (dans l'espace) la borne inférieure des distances du point avec les points du plan. En symbole :

$$d(P, \Pi) := \inf_{Q \in \Pi} d(P, Q).$$

En particulier on voit que $P \in \Pi$ entraîne que $d(P, \Pi) = 0$. Voici le calcul en général :

THÉORÈME: *Soit Π un plan dans l'espace ayant pour équation $ax + by + cz + d = 0$ et $P = (x_0, y_0, z_0)$ un point de l'espace; la distance de P à Π est donnée par*

$$d(P, \Pi) = \frac{|ax_0 + by_0 + cz_0 + d|}{\sqrt{a^2 + b^2 + c^2}}.$$

De plus il existe un unique point $Q \in \Pi$ tel que $d(P, \Pi) = d(P, Q)$ et il peut être caractérisé par le fait que la droite PQ est orthogonale à Π.

Démonstration: Le lecteur observera la similarité de cet énoncé avec la formule donnant la distance d'un point à une droite dans le plan et, en regardant sa démonstration, verra qu'elle s'applique presque mot pour mot. □

Euclide [$Ευκλειδης$] (environ 300 ans avant JC)

APPENDICE : RÉSUMÉ D'ALGÈBRE LINÉAIRE

Cet appendice contient un résumé des résultats d'algèbre linéaire développés dans ce cours et une synthèse des différentes méthodes. On donne des applications du cours mais aucune démonstration (celles-ci sont contenues ou peuvent être extraites des chapitres précédents), en cela ce chapitre empiète un peu sur le travail de TD.

A. DÉFINITIONS ET PROBLÈMES

Doivent être connues les définitions d'un (sous-)espace vectoriel, d'une base, d'une partie libre ou génératrice.

Etant donné une famille de vecteurs $f_1, f_2, \ldots, f_s$, on veut savoir si elle est libre, génératrice ou forme une base. On veut savoir aussi quel est son rang, c'est-à-dire quelle est la dimension de l'espace vectoriel qu'elle engendre. Si elle est libre, on veut savoir comment la compléter en une base (ce qui est possible d'après le théorème de la base incomplète).

Un sous-espace vectoriel de $\mathbf{R}^n$ peut être donné par une base, une partie génératrice, par des équations ou par des équations en nombre minimal. Dans le premier et le dernier cas la dimension du sous-espace vectoriel est facile à calculer : c'est le cardinal d'une base et c'est aussi n moins le nombre minimal d'équations.

Exemple: le sous-espace vectoriel $E = \{(x, y, z) \in \mathbf{R}^3 \mid x+y+z = 0,\ x+y = 0,\ z = 0\}$ est aussi égal à $\{(x, y, z) \in \mathbf{R}^3 \mid x+y = 0,\ z = 0\}$; il est engendré par les vecteurs $(1, -1, 0)$, $(-2, 2, 0)$ et $(18, 18, 0)$ et admet pour base le vecteur $(1, -1, 0)$. Il a pour dimension 1.

Remarque: en géométrie analytique, on obtient souvent un sous-espace vectoriel sous forme paramétrique; en fait cela revient à se donner une partie génératrice: le sous-espace vectoriel engendré par $f_1, f_2, \ldots, f_s$ est l'ensemble $\{t_1 f_1 + t_2 f_2 + \ldots + t_s f_s \mid (t_1, \ldots, t_s) \in \mathbf{R}^s\}$.

Un problème se pose souvent : ayant des équations pour un sous-espace vectoriel, trouver une base et la dimension de ce sous-espace vectoriel. Inversement, ayant une base ou une partie génératrice, on veut trouver des équations.

On sait que l'intersection ou la somme de deux sous-espaces vectoriels est encore un sous-espace vectoriel; un problème naturel est de calculer une base ou des équations de cette intersection ou (somme). En particulier deux sous-espaces vectoriels E et F sont en somme directe si $E \cap F = \{0\}$.

Doivent être connues les définitions d'une application linéaire, du noyau et de l'image, de la matrice de l'application linéaire par rapport à une base de l'espace de départ et une base de l'espace d'arrivée.

Une application linéaire u, disons de $\mathbf{R}^n$ vers $\mathbf{R}^m$ peut être donnée explicitement (exemple : $u(x, y, z) = (x+y, z, x+y+z, x+y)$) ou par sa matrice dans des bases données de $\mathbf{R}^n$ et $\mathbf{R}^m$ ou encore par une description géomètrique (exemples : la projection de $\mathbf{R}^3$ vers $\mathbf{R}^3$ sur un plan parallèlement à une droite, la rotation d'angle $\pi/5$ autour d'une droite de $\mathbf{R}^3$). Dans le dernier cas, on veut trouver la matrice de u dans des base données, dans les premiers cas on veut trouver une interprétation géomètrique : ceci se fait le plus

souvent en choisissant de "bonnes" bases et en calculant la matrice de u dans ces bases. Les méthodes pour trouver de "bonnes" bases ne font pas partie du programme de première année, elles seront donc toujours données dans les problèmes et exemples.

Dans tous les cas on veut calculer le noyau et l'image de l'application linéaire.

On doit connaître les règles de calcul des matrices ainsi que la formule de changement de bases à l'aide des matrices de passages.

Faire le produit de matrices permet de calculer la composée de deux applications linéaires (et vice versa). La formule de changement de bases ("$A' = Q^{-1}AP$") permet de calculer la matrice d'une application linéaire dans de nouvelles bases, connaissant la matrice de l'application linéaire dans certaines bases.

B. CALCULS PRATIQUES

On donne une liste (non exhaustive) de problèmes d'algèbre linéaire et de méthodes pour les résoudre. Un autre titre à ce paragraphe aurait pu être : application de la méthode du pivot.

a) Résolution de système linéaire.

Soit à résoudre le système :

$$\begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n = b_1 \\ \ldots \ \ldots \ \ldots \\ a_{m1}x_1 + \ldots + a_{mn}x_n = b_m \end{cases}$$

- On forme la matrice $(A|b)$ avec A matrice $m \times n$ formée des coefficients a_{ij} et b vecteur colonne formé des coefficients b_i. - On échelonne la matrice $(A|b) \to (A'|b')$ et on regarde si il y a un pivot sur la dernière colonne. S'il y a un pivot, le système n'a pas de solution. Sinon, on repère les colonnes sans pivot de la matrice A', elles correspondent aux variables $x_{i_1}, \ldots, x_{i_s}$. Les solutions s'obtiennent toutes en fixant des valeurs arbitraires aux inconnues $x_{i_1}, \ldots, x_{i_s}$ et en exprimant les autres en fonction de celles-ci.

Commentaires : l'ensemble des solutions est de la forme "solution particulière + solution de l'équation homogène" et la dimension des solutions est le nombre de colonne de A' sans pivot. Si $b = 0$, il suffit d'échelonner $A \to A'$ et de raisonner sur A'.

b) calcul sur les vecteurs.

Soit $f_1, \ldots, f_s$ une famille de vecteurs d'un espace vectoriel (disons $\mathbf{R}^n$ pour fixer les idéees).

Pour vérifier si $f_1, \ldots, f_s$ est libre, on résoud le système linéaire $x_1 f_1 + \ldots + x_r f_r = 0$. Les vecteurs $f_1, \ldots, f_s$ sont libres si la seule solution est $x_1 = \ldots = x_r = 0$ (si la matrice échelonnée correspondante n'a pas de colonne sans pivot).

Pour vérifier si $f_1, \ldots, f_s$ est génératrice de $\mathbf{R}^n$, on résoud le système $x_1 f_1 + \ldots + x_r f_r = b$. Les vecteurs $f_1, \ldots, f_s$ sont générateurs si le système a une solution pour tout $b \in \mathbf{R}^n$.

Note: Les vecteurs $f_1, \ldots, f_s$ ne peuvent être indépendants dans $\mathbf{R}^n$ que si $s \leq n$ et générateurs que si $s \geq n$. Dans le cas particulier $s = n$, rappelons que libres entraîne générateurs (et base), ce qui économise souvent des calculs.

Pour extraire une partie libre maximale, on échelonne et on ne garde que les f_i correspondant à une colonne avec pivot.

Pour compléter une partie libre $f_1, \ldots, f_s$ en une base, on peut ajouter à ces vecteurs des générateurs de l'espace (par exemple la base canonique) et extraire de la partie génératrice $f_1, \ldots, f_s, e_1, \ldots e_n$ une partie libre maximale par la méthode précédente.

Pour exprimer un vecteur b dans une base $f_1, \ldots, f_s$, on résoud le système $x_1 f_1 + \ldots + x_r f_r = b$ (dont on sait qu'il possède une solution unique).

c) calcul sur les $sous - espaces$ vectoriels.

Soit un sous-espace vectoriel E de $\mathbf{R}^n$ défini par les équations:

$$\begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n = 0 \\ \cdots \cdots \cdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n = 0 \end{cases}$$

La résolution (par la méthode du pivot) du système linéaire décrit par ces équations donne la dimension de E et permet d'exprimer les variables "avec pivot" en fonction des autres, d'où le calcul d'une base.

Exemple : si le résultat des calculs donne comme solution que x_1 et x_2 sont des variables libres et que $x_3 = 2x_1 - 7x_2$, $x_4 = -2x_1 + 4x_2$ alors $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ 2 \\ -2 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ -7 \\ 4 \end{pmatrix}$ et donc une base est donnée par $\begin{pmatrix} 1 \\ 0 \\ 2 \\ -2 \end{pmatrix}$ et $\begin{pmatrix} 0 \\ 1 \\ -7 \\ 4 \end{pmatrix}$.

Soit maintenant un sous-espace vectoriel E de $\mathbf{R}^n$ donné par une partie génératrice $f_1, \ldots, f_s$, pour calculer des équations définissant E on cherche à résoudre le système $x_1 f_1 + \ldots + x_s f_s = \begin{pmatrix} b_1 \\ \cdot \\ \cdot \\ b_n \end{pmatrix}$. La méthode du pivot donne des conditions linéaires sur les b_i pour que le système ait une solution ; ces conditions donnent les équations pour que $\begin{pmatrix} b_1 \\ \cdot \\ \cdot \\ b_n \end{pmatrix}$ soit dans E. Une base de E s'obtient en extrayant une partie libre maximale de $f_1, \ldots, f_s$.

d) calcul d'intersection, somme de $sous - espaces$.

Si E et F sont donnés par des équations, alors $E \cap F$ est donné par la réunion des deux paquets d'équations. Si E et F sont donnés chacun par une partie génératrice, alors une partie génératrice de l'espace $E + F$ est donné par la réunion des deux parties génératrices de E et F. Grâce à c) ont peut donc calculer intersection et somme de sous-espace vectoriel (leurs dimensions, une base etc).

e) calcul du noyau et image d'une application linéaire.

Soit $u : \mathbf{R}^n \to \mathbf{R}^m$ donnée par

$$u \begin{pmatrix} x_1 \\ . \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \ldots + a_{1n}x_n \\ \ldots \\ a_{m1}x_1 + \ldots + a_{mn}x_n \end{pmatrix}$$

Pour calculer $\mathrm{Ker}(u)$ et $\mathrm{Im}(u)$, ainsi que leurs dimensions, on forme la matrice $(A|b)$ avec $A = (a_{ij})$ et $b \in \mathbf{R}^m$; on l'échelonne : $(A|b) \to (A|b')$. Le noyau est obtenu en calculant les solutions du système avec $b = 0$, sa dimension est le nombre de colonnes sans pivot de A'; l'image est définie par les équations linéaires en b_i écrivant que le système associé à $(A|b)$ a une solution, sa dimension est le nombre de colonnes avec pivot de A' (on peut vérifier que $\dim \mathrm{Ker}(u) + \dim \mathrm{Im}(u) = n$). De plus une base de $\mathrm{Im}(u)$ peut être obtenue à l'aide des vecteurs e_i de la base canonique en ne retenant que les indices correspondant à une colonne avec pivot.

f) calcul sur les matrices.

Le chapitre sur les matrices étant essentiellement calculatoire on ne révise que quelques points; en particulier il faut aller voir dans le chapitre 7 la méthode donnée pour calculer l'inverse A^{-1} d'une matrice (ou conclure qu'il n'existe pas) et réviser dans le chapitre 9 le calcul des matrices de passage et de changements de base.

Soit $u : \mathbf{R}^n \to \mathbf{R}^m$ donnée par

$$u \begin{pmatrix} x_1 \\ . \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \ldots + a_{1n}x_n \\ \ldots \\ a_{m1}x_1 + \ldots + a_{mn}x_n \end{pmatrix}$$

Soit $e_1, \ldots, e_n$ une base de $\mathbf{R}^n$ et $f_1, \ldots, f_m$ une base de $\mathbf{R}^m$. Pour calculer B la matrice de u dans ces bases, on exprime chacun des vecteurs $u(e_i)$ comme combinaison linéaire des f_j avec la méthode donnée en b):

$$u(e_i) = b_{1i}f_1 + \ldots + b_{mi}f_m$$

et les coefficients de B sont donnés par les b_{ij}.

CHAPITRE 12 SUITES DE NOMBRES RÉELS OU COMPLEXES

La notion de suite est connue et étudiée depuis la terminale ; elle permet de décrire des processus itératifs numériques et contient en germe la notion de limite. La raison fondamentale de son importance est la possibilité d'approcher des phénomènes continus par des quantités discrètes : on approche un nombre réel par son développement décimal à l'ordre n, on trace un graphe en calculant des valeurs discrètes – tout calcul numérique effectué sur un ordinateur repose sur ce principe. Inversement c'est un sujet qui se prête bien aux expérimentations sur machine ou calculatrice. Ceci explique aussi que souvent on ne se contente pas de savoir qu'une suite approche un nombre ; on veut aussi savoir avec quelle précision le n-ième terme approche le nombre.

12.1 PROPRIÉTÉS GÉNÉRALES DES SUITES

Répétons les définitions données au chapitre 4 :

Définition: Une *suite* réelle ou complexe est une application $u : \mathbf{N} \to \mathbf{R}$ ou $\mathbf{C}$. On note $u_n := u(n)$ et on l'appelle le n-ième *terme* de la suite. La suite elle-même est notée $(u_n)_{n\in\mathbf{N}}$ ou plus simplement (u_n) (cette dernière notation étant un peu abusive mais usuelle).

Définition: Une suite (réelle ou complexe) $(u_n)_{n\in\mathbf{N}}$ est *convergente* vers ℓ si

$$\forall \varepsilon > 0,\ \exists n_0 \in \mathbf{N},\ n \geq n_0 \Rightarrow |u_n - \ell| \leq \varepsilon$$

On dit aussi que ℓ est la *limite* de la suite. Une suite est *divergente* si elle ne converge vers aucun nombre.

Remarques : i) les premiers termes d'une suite n'ont pas d'importance ; en particulier leurs valeurs n'influent en rien sur la convergence de la suite ; c'est pourquoi on les néglige souvent et on commettra même l'abus de langage consistant par exemple à parler de la suite $u_n = 1/n$ alors que le terme u_0 n'est pas défini.

ii) Par ailleurs une suite n'est pas forcément convergente, par exemple les suites $u_n = (-1)^n$ ou $v_n = \sin(n)$ ne sont pas convergentes (essayez de le démontrer!).

iii) On peut généraliser la définition de limite (pour une suite réelle) en décrétant que $\lim u_n = +\infty$ si $\forall M > 0,\ \exists n_0 \in \mathbf{N},\ n \geq n_0 \Rightarrow u_n > M$ (idem pour $-\infty$). Nous réserverons néanmoins l'appellation "convergente" aux suites ayant une limite finie.

PROPOSITION: *Une suite convergente est bornée.*

Démonstration: Si (u_n) est convergente, il existe n_0 et ℓ tel que pour $n \geq n_0$ on ait $|u_n - \ell| \leq 1$ donc $|u_n| \leq |\ell| + 1$. Posons donc $M := \max\{|u_0|, \ldots, |u_{n_0}|, |\ell| + 1\}$ alors on a pour tout $n \in \mathbf{N}$ l'inégalité $|u_n| \leq M$. □

La réciproque est fausse puisque les deux suites déjà mentionnées $u_n = (-1)^n$ ou $v_n = \sin(n)$ ne sont pas convergentes mais pourtant bornées. Néanmoins une propriété saute aux yeux, au moins pour la première suite : la "sous-suite" u_{2n} converge (elle est en fait constante). L'existence d'une sous-suite convergente est une propriété générale des suites bornées (théorème de Bolzano-Weierstrass) que nous citons pour mémoire. Voyons maintenant quelles opérations on peut effectuer sur les limites de suites.

THÉORÈME: *Si* $\lim u_n = \ell$ *et* $\lim v_n = \ell'$ *alors*

i) $\lim(au_n + bv_n) = a\ell + b\ell'$

ii) $\lim(u_n v_n) = \ell\ell'$

iii) Si $\ell' \neq 0$ *alors* $\lim(u_n/v_n) = \ell/\ell'$

Dans le cas ou l'une de ces limites est infinie ces formules restent valables à condition d'appliquer les règles suivantes :

$+\infty + \infty = +\infty$

$+\infty + a = +\infty$

$-\infty - \infty = -\infty$

$-\infty + a = -\infty$

Si $a > 0$ *alors* $a(+\infty) = +\infty$

si $a < 0$ *alors* $a(+\infty) = -\infty$

$(+\infty).(+\infty) = +\infty$

$(-\infty).(-\infty) = +\infty$

Si $\lim u_n = 0$ *mais (au moins à partir d'un certain rang)* $u_n \neq 0$ *alors* $\lim \frac{1}{|u_n|} = +\infty$. *De plus si en fait* $u_n > 0$ *(resp.* $u_n < 0$*) alors* $\lim \frac{1}{u_n} = +\infty$ *(resp.* $= -\infty$*).*

Démonstration: (i) Soit $\varepsilon > 0$ donné, choisissons $\varepsilon' > 0$ tel que $(|a| + |b|)\varepsilon' < \varepsilon$; par définition des limites, il existe n_0 et n_1 tels que si $n \geq n_0$ alors $|u_n - \ell| \leq \varepsilon'$ et si $n \geq n_1$ alors $|v_n - \ell'| \leq \varepsilon'$. Si maintenant $n \geq \max(n_0, n_1)$ alors $|(au_n + bv_n) - (a\ell + b\ell')| \leq |a||u_n - \ell| + |b||v_n - \ell'| \leq (|a| + |b|)\varepsilon' < \varepsilon$; ce qui prouve bien que $\lim(au_n + bv_n) = a\ell + b\ell'$.

(ii) Observons que d'après les hypothèses et la proposition précédente u_n est bornée par, disons, M. Soit $\varepsilon > 0$ donné, choisissons $\varepsilon' > 0$ tel que $(M + |\ell'|)\varepsilon' < \varepsilon$; par définition des limites, il existe n_0 et n_1 tels que si $n \geq n_0$ alors $|u_n - \ell| \leq \varepsilon'$ et si $n \geq n_1$ alors $|v_n - \ell'| \leq \varepsilon'$. Si maintenant $n \geq \max(n_0, n_1)$ alors $|u_n v_n - \ell\ell'| \leq |u_n||v_n - \ell'| + |\ell'||u_n - \ell| \leq (M + |\ell'|)\varepsilon' < \varepsilon$; ce qui prouve bien que $\lim u_n v_n = \ell\ell'$.

(iii) On utilise cette fois-ci que si $n \geq n_0$ alors $\frac{1}{|v_n|} \leq \frac{2}{|\ell'|}$ et l'inégalité $\left|\frac{u_n}{v_n} - \frac{\ell}{\ell'}\right| \leq \frac{|u_n - \ell|}{|v_n|} + \frac{|v_n - \ell'|}{|v_n \ell'|}$.

La démonstration des autres énoncés est assez immédiate en appliquant le même raisonnement (il est recommandé toutefois d'essayer de faire la démonstration des cas laissés indéterminés et de voir "qu'est-ce-qui ne marche pas"). Démontrons par exemple le dernier énoncé lorsque $u_n > 0$ et $\lim u_n = 0$: soit donc $M > 0$ alors il existe $n_0 \in \mathbf{N}$ tel que si $n \geq n_0$ alors $0 < u_n < \frac{1}{M}$ et donc $\frac{1}{u_n} > M$; ce qui prouve bien que $\lim \frac{1}{u_n} = +\infty$. □

Remarque : par contre $0.\infty$, $\frac{\infty}{\infty}$ et $+\infty - \infty$ ne sont pas définis (c'est-à-dire que ces seules informations ne sont pas suffisantes pour conclure). En effet choisissons par exemple $u_n := 1/n$ et $v_n := n$, $w_n := \sqrt{n}$, $t_n := n^2$ alors la première tend vers zéro alors que les autres tendent vers plus l'infini ; pourtant $\lim u_n v_n = 1$, $\lim u_n w_n = 0$ et $\lim u_n t_n = +\infty$.

COROLLAIRE: *Soit* u_n *une suite convergente et* v_n *une suite divergente alors* $u_n + v_n$ *est divergente.*

Démonstration: En effet si $u_n + v_n$ était convergente cela entraînerait que la suite $v_n = (u_n + v_n) - (u_n)$ serait également convergente. □

Exemple : considérons la suite $u_n := \sin(n) + \frac{1}{n+1} + (-1)^n e^{-n}$. Alors les deux suites $\frac{1}{n+1}$ et $(-1)^n e^{-n}$ sont convergentes (vers 0) mais la suite $\sin(n)$ est divergente donc la suite u_n est divergente.

Rappelons qu'une suite (réelle) monotone et bornée est convergente (chapitre 4). La variante suivante est souvent utile :

THÉORÈME: *(suites adjacentes) Soient u_n et v_n deux suites réelles telles que :*
(i) u_n est croissante et v_n est décroissante
(ii) La suite $v_n - u_n$ est positive et converge vers 0
alors u_n et v_n convergent toutes deux vers la même limite.

Démonstration: La suite u_n est croissante et majorée par v_0 donc est convergente (disons vers ℓ) ; la suite v_n est décroissante et minorée par u_0 donc convergente (disons vers ℓ'). La dernière condition entraîne que $\ell = \ell'$. □

Exemples : 1) Un des exemples les plus classiques est le suivant : on choisit $0 < a < b$ et on pose $u_0 := a$ et $v_0 = b$ et on définit par récurrence $u_{n+1} := \sqrt{u_n v_n}$ et $v_{n+1} := (u_n + v_n)/2$. La limite commune s'appelle la "moyenne arithmético-géométrique" de a et b.

Vérifions que u_n et v_n sont adjacentes : on a $u_0 < v_0$ par hypothèse et si $n \geq 1$ alors $v_n - u_n = \frac{1}{2}(u_{n-1} + v_{n-1} - 2\sqrt{u_{n-1}v_{n-1}}) = \frac{1}{2}(\sqrt{v_{n-1}} - \sqrt{u_{n-1}})^2 \geq 0$ donc $u_n \leq v_n$. Ensuite $v_{n+1} = (u_n + v_n)/2 \leq v_n$ et $u_{n+1} := \sqrt{u_n v_n} \geq u_n$ donc u_n est croissante et majorée donc convergente, disons vers ℓ et v_n est décroissante et minorée donc convergente vers ℓ' mais $\ell' = (\ell + \ell')/2$ donc $\ell = \ell'$.

De $v_n - u_n = \frac{1}{2}(\sqrt{v_{n-1}} - \sqrt{u_{n-1}}) = \frac{1}{2(\sqrt{v_{n-1}} + \sqrt{u_{n-1}})^2}(v_{n-1} - u_{n-1})^2$ on tire $|v_n - u_n| \leq C(v_{n-1} - u_{n-1})^2$ (avec $C := \frac{1}{8u_0}$) donc la convergence de la suite est très rapide (chaque pas double le nombre de chiffres de précision). Par exemple si $a := 2$ et $b := 8$ on trouve $u_1 = 4$, $u_2 = 4,47213\ldots$, $u_3 = 4,48604\ldots$ et $v_1 = 5$, $v_2 = 4,5$, $v_3 = 4,48606\ldots$.

2) Nous avons déjà étudié (chapitre 4) un autre exemple de suites adjacentes : soit x un réel et soit u_n son développement décimal par défaut à l'ordre n et v_n son développement décimal par excès à l'ordre n alors u_n et v_n sont adjacentes et convergent vers x (bien sûr, chaque pas ajoute un chiffre de précision).

3) La suite $u_n := 1 + \frac{1}{1!} + \frac{1}{2!} + \ldots + \frac{1}{n!}$ est convergente. En effet introduisons $v_n := u_n + \frac{1}{nn!}$; visiblement u_n est croissante, $u_n \leq v_n$ et $\lim(v_n - u_n) = 0$ mais de plus $v_{n+1} - v_n = \frac{-1}{n(n+1)(n+1)!} < 0$ donc v_n est décroissante et les suites sont adjacentes. Si l'on appelle e la limite commune (nous verrons au chapitre 14 que c'est bien la base du logarithme népérien) on a : $0 < e - u_n < \frac{1}{nn!}$. La suite u_n fournit donc de très bonnes approximations de e. Par exemple $u_1 = 2$, $u_2 = 2,5$, $u_3 = 2,66666\ldots$, $u_4 = 2,708333\ldots$, $u_5 = 2,71666\ldots$ et $v_1 = 3$, $v_2 = 2,75$, $v_3 = 2,72222\ldots$, $v_4 = 2,71874\ldots$, $v_5 = 2,718333\ldots$.

12.2 SUITES RÉCURRENTES

Le cas le plus général est celui d'une suite où u_n s'exprime en fonction de $u_0, \ldots, u_{n-1}$; nous étudierons seulement deux cas.

12.2.1 Suites du type $u_{n+1} = f(u_n)$

Les exemples les plus simples de telles suites sont :

Les suites arithmétiques définies par $u_{n+1} = u_n + a$: on déduit immédiatement que $u_n = u_0 + na$ et que $u_0 + \ldots + u_n = (n+1)u_0 + \frac{n(n+1)}{2}a$. En particulier, si u_n est réelle $\lim u_n = +\infty$ si $a > 0$, $\lim u_n = -\infty$ si $a < 0$ (et $u_n = u_0$ si $a = 0$).

Les suites géométriques définies par $u_{n+1} = au_n$: on déduit immédiatement que $u_n = u_0 a^n$ et que $u_0 + \ldots + u_n = (n+1)u_0$ si $a = 1$ et $\frac{a^{n+1}-1}{a-1}u_0$ si $a \neq 1$. En particulier $\lim u_n = 0$ si $|a| < 1$ et $\lim |u_n| = +\infty$ si $|a| > 1$. Si $|a| = 1$ donc $a = e^{i\theta}$, la situation est plus compliquée : la suite reste sur le cercle $|z| = |u_0|$ et u_{n+1} s'obtient à partir de u_n par une rotation d'angle θ ; si $\theta/\pi \in \mathbf{Q}$ alors la suite est périodique, sinon elle tourne de façon "dense" sur le cercle (c'est nettement plus difficile à montrer).

On peut donner un exemple plus compliqué où l'on arrive néanmoins à exprimer u_n par une formule fermée : soit la suite définie par $u_{n+1} = \frac{1}{2}\left(u_n + \frac{1}{u_n}\right)$; posons $w_n = \frac{u_n+1}{u_n-1}$ alors un calcul direct fournit la relation $w_{n+1} = w_n^2$ d'où $w_n = w_0^{2^n}$ et

$$u_n = \frac{w_0^{2^n} + 1}{w_0^{2^n} - 1} = \frac{\left(\frac{u_0+1}{u_0-1}\right)^{2^n} + 1}{\left(\frac{u_0+1}{u_0-1}\right)^{2^n} - 1}$$

On en déduit par exemple que si $\left|\frac{u_0+1}{u_0-1}\right| < 1$ alors $\lim u_n = -1$ alors que si $\left|\frac{u_0+1}{u_0-1}\right| > 1$ alors $\lim u_n = 1$. On a vu au chapitre 4 que l'équation $\frac{|z+1|}{|z-1|} = 1$ définit, dans le plan complexe, un cercle ou une droite ; ici c'est la droite des imaginaires. Si $Re(u_0) > 0$ (resp. si $Re(u_0) < 0$) alors $\lim u_n = 1$ (resp. $\lim u_n = -1$) ; si $u_0 \in i\mathbf{R}$ la suite ne peut pas converger (ses valeurs sont purement imaginaires) et peut même ne pas être définie à partir d'un certain rang (si $\frac{u_0+1}{u_0-1} = e^{k\pi i/2^m}$).

En général toutefois il est impossible d'exprimer u_n par une formule simple et on doit recourir au raisonnement et à des considérations qualitatives. Une des propriétés les plus utiles de ces suites est la suivante (voir le cours de terminale ou le chapitre suivant pour la définition de la continuité) :

THÉORÈME: *Soit $f : \mathbf{R} \to \mathbf{R}$ une fonction continue et soit u_n une suite vérifiant la relation $u_{n+1} = f(u_n)$; si u_n converge vers ℓ alors $\ell = f(\ell)$*

Remarque : le théorème ne dit pas que cette suite converge mais seulement que l'ensemble des "limites potentielles" est restreint.

Démonstration: D'après une propriété des fonctions continues : $\lim f(u_n) = f(\lim u_n)$ donc vues les hypothèses

$$\ell = \lim u_{n+1} = \lim f(u_n) = f(\lim u_n) = f(\ell).$$

□

Pour nous aider à déterminer quand une telle suite converge effectivement vers un point fixe ℓ de f, il est très éclairant de tracer le diagramme suivant qui décrit visuellement l'itération de la fonction f : on représente le graphe $y = f(x)$ de la fonction f et la droite

$y = x$; connaissant u_0, l'intersection de la droite verticale $x = u_0$ avec le graphe donne le point $(u_0, f(u_0)) = (u_0, u_1)$, l'intersection de la droite horizontale passant par ce point avec la droite $y = x$ donne le point (u_1, u_1), l'intersection de la droite verticale passant par ce dernier point avec le graphe de f fournit le point $(u_1, f(u_1) = (u_1, u_2)$ et ainsi de suite

Représentation graphique de quelques suites récurrentes suivant la valeur de $\alpha := f'(\ell)$

Ces graphiques suggèrent que si $|\alpha| = |f'(\ell)| < 1$ alors la suite semble converger vers ℓ (disons que le point fixe ℓ est *attracteur* ou *stable*) et si $|\alpha| = |f'(\ell)| > 1$ alors la suite ne semble pas converger vers ℓ (disons que le point fixe ℓ est *répulsif* ou *instable*). Dans le premier cas, si $0 < \alpha < 1$ la suite semble converger de façon monotone vers ℓ, par contre si $-1 < \alpha < 0$ la suite semble converger de façon alternée vers ℓ. On peut démontrer (si f est deux fois continument dérivable) grâce à la formule de Taylor que c'est bien ce qui se passe pourvu que u_0 soit suffisamment proche de ℓ. Illustrons cela par quelques exemples :

$u_{n+1} = \sqrt{u_n + 5}$ (avec $u_0 \geq -5$) :

On constate que $u_1 \geq 0$ et ensuite $u_n \geq 0$ donc la suite est bien définie. Si elle converge vers ℓ alors ℓ doit être positive et vérifier $\ell = \sqrt{\ell + 5}$ donc $\ell = \frac{1+\sqrt{21}}{2}$ est la seule limite possible. Le tracé du graphe $y = \sqrt{x + 5}$ et de l'itération suggère que si $u_0 \leq \ell$ (resp. $u_0 \geq \ell$) alors u_n est croissante (resp. décroissante) et converge vers ℓ, démontrons cela. Tout d'abord la fonction $f(x) := \sqrt{x + 5}$ est croissante donc $u_n \leq \ell$ (resp. $u_n \geq \ell$)

entraîne $u_{n+1} = f(u_n) \leq f(\ell) = \ell$ (resp. $u_{n+1} \geq \ell$). Ensuite

$$u_{n+1} - u_n = \frac{u_n + 5 - u_n^2}{\sqrt{u_n + 5} + u_n} = -\frac{(u_n - \frac{1+\sqrt{21}}{2})(u_n - \frac{1-\sqrt{21}}{2})}{\sqrt{u_n + 5} + u_n}$$

est positif si $u_n \leq \ell$ et négatif si $u_n \geq \ell$. En conclusion si $u_0 \leq \ell$ alors u_n est croissante et majorée par ℓ donc convergente et la limite doit être ℓ ; si $u_0 \geq \ell$ alors u_n est décroissante et minorée par ℓ donc convergente et la limite doit être ℓ. On remarquera que la pente du graphe en $x = \ell$ est positive et inférieure à 1 (exactement : $0 < \alpha = \frac{1}{2\ell}$).

Si l'on reprend la suite $u_{n+1} = \frac{1}{2}(u_n + 1/u_n)$ on observe sur le graphe que les deux limites possibles, $+1$ et -1, sont attracteurs.

On "voit" que si $u_0 > 0$ alors la suite va converger vers 1 et si $u_0 < 0$ alors elle va converger vers -1 (ce que l'on a déjà démontré directement).

$u_{n+1} = \frac{au_n+b}{cu_n+d}$: on suppose $ad - bc$ et c différent de zéro (sinon la suite est constante ou linéaire) et également $u_0 \neq -d/c$. Si la suite converge vers une limite ℓ alors celle-ci vérifie $c\ell^2 + (d-a)\ell - b = 0$. On doit distinguer deux cas : ou bien $(d-a)^2 + 4bc \neq 0$ et il y a deux limites possibles (réelles ou complexes), disons α et β, ou bien $(d-a)^2 + 4bc = 0$ et il y a une seule limite possible, disons α.

Si a, b, c, d sont réels, pour tracer le graphe des fonctions $f(x) = \frac{ax+b}{cx+d}$ observons que $y = f(x)$ équivaut à $y - \frac{a}{c} = -\frac{ad-bc}{c}/(x + \frac{d}{c})$ et donc, à translation près le graphe est celui de la fonction C/x et dépend donc du signe de C.

Cas où $ad - bc < 0$:

Cas où $ad - bc > 0$

On constate que dans le premier cas il y a deux limites possibles et que la suite semble converger vers l'une d'elles (toujours la même) : l'un des points est attracteur, l'autre est répulsif. Dans le deuxième cas il peut y avoir deux limites possibles, une seule ou aucune (i.e. α et β sont complexes) ; dans ce dernier cas il est clair que la suite ne peut converger.

On peut étudier un peu plus profondément ces suites ainsi : dans le cas ou α et β sont distinctes, introduisons la suite $v_n = \frac{u_n-\beta}{u_n-\alpha}$. Un calcul direct utilisant que α et β sont les racines de $cX^2 + (d-a)X - b$ donne $v_{n+1} = \lambda v_n$ avec $\lambda := \frac{a-c\beta}{a-c\alpha}$. Alors si $|\lambda| < 1$ on en déduit $\lim v_n = 0$ et $\lim u_n = \beta$; si $|\lambda| > 1$ on en déduit $\lim |v_n| = +\infty$ et $\lim u_n = \alpha$. Le cas où $|\lambda| = 1$ est plus compliqué ; c'est notamment ce qui se produit si les coefficients sont réels et α, β complexes conjugués.

Si α est la seule limite possible, on introduit $v_n := \frac{1}{u_n-\alpha}$ et de nouveau un calcul direct fournit $v_{n+1} = v_n + \lambda$ avec $\lambda := \frac{c}{a-\alpha c}$. On en tire $\lim |v_n| = +\infty$ et donc $\lim u_n = \alpha$.

On pourra vérifier cette théorie sur les exemples suivants : $u_{n+1} = \frac{u_n+1}{u_n-2}$ (deux limites possibles $\alpha = \frac{3-\sqrt{13}}{2}$ et $\beta = \frac{3+\sqrt{13}}{2}$ et $\lambda > 1$ donc $\lim u_n = \frac{3-\sqrt{13}}{2}$ sauf si $u_0 = \beta$) ; $u_{n+1} = \frac{2u_n-1}{u_n}$ (une seule limite possible $\alpha = 1$) ; $u_{n+1} = \frac{u_n+1}{u_n-1}$ (les deux limites possibles sont $\alpha = 1+\sqrt{2}$ et $\beta = 1-\sqrt{2}$ mais $\lambda = -1$ et la suite est périodique non convergente sauf si $u_0 = \alpha$ ou β).

Terminons par deux exemples où la limite possible est un point "faiblement" attracteur ou répulsif (i.e. la pente vaut 1) :

$u_{n+1} = \sin(u_n)$. En traçant le graphe on constate que la suite semble converger vers 0 ;

Prouvons cela. Le point clef est l'inégalité (classique mais démontré au chapitre 15) : $0 < x \leq \pi/2 \Rightarrow 0 < \sin(x) < x$. Elle permet d'établir que la seule limite possible est 0 ; d'autre part $|u_1| \leq 1$ et si $u_1 = 0$ la suite est constamment nulle ensuite donc on peut supposer $0 < u_1 \leq 1$ (le cas $-1 \leq u_1 < 0$ se traite parallèlement en utilisant l'imparité de $\sin(x)$) et en démontre alors que $0 < u_{n+1} = \sin(u_n) < u_n \leq 1$. La suite est donc décroissante et minorée donc convergente et la limite ne peut être que 0.

$u_{n+1} = u_n + u_n^3$. La seule limite possible est 0 et le graphe semble indiquer que c'est un point répulsif.

En fait si $u_0 > 0$ (on peut faire un raisonnement parallèle si $u_0 < 0$) alors $u_1 > u_0$ et la suite u_n est strictement croissante. Si u_n était bornée, elle convergerait vers une limite finie et strictement positive, ce qui est absurde donc elle tend vers l'infini.

12.2.2 Suites linéaires récurrentes.

Il s'agit des suites vérifiant une relation du type :

$$(*) \quad u_n = au_{n-1} + bu_{n-2} \quad \text{pour } n \geq 2$$

On suppose que $b \neq 0$ sinon on est ramené à une suite géométrique. On vérifie immédiatement que l'ensemble des suites vérifiant une telle relation donnée est un espace vectoriel (un sous-espace vectoriel de l'espace des suites) : en effet soit u_n et v_n deux telles suites, considérons $w_n := cu_n + dv_n$ alors $w_n = c(au_{n-1} + bu_{n-2}) + d(av_{n-1} + bv_{n-2}) = a(cu_{n-1} + dv_{n-1}) + b(cu_{n-2} + dv_{n-2}) = aw_{n-1} + bw_{n-2}$. D'autre part les valeurs u_0 et u_1 déterminent entièrement le reste de la suite ; donc si v_n est la suite obtenue avec $v_0 := 1$ et $v_1 := 0$ et w_n est la suite obtenue avec $w_0 := 0$ et $w_1 := 1$ on voit que ces deux suites forment une base de l'espace des suites vérifiant (*) (qui est donc de dimension 2) . On peut trouver une base plus explicitement de la façon suivante : considérons l'équation $X^2 - aX - b = 0$ (qui s'appelle l'*équation caractéristique* de la suite linéaire) et choisissons une racine α alors la suite $u_n := \alpha^n$ vérifie :

$$u_n - au_{n-1} - bu_{n-2} = \alpha^{n-2}(\alpha^2 - a\alpha - b) = 0$$

Distinguons donc deux cas :

Cas n^o 1 : Les deux racines α et β de $X^2 - aX - b = 0$ sont distinctes et alors une solution de (*) s'écrit $u_n = \lambda\alpha^n + \mu\beta^n$; si l'on veut tout exprimer en terme de u_0 et u_1 on écrit le système :

$$u_0 = \lambda + \mu \quad \text{et} \quad u_1 = \lambda\alpha + \mu\beta$$

d'où l'on tire $\lambda = \frac{u_1 - \beta u_0}{\alpha - \beta}$ et $\mu = \frac{\alpha u_0 - u_1}{\alpha - \beta}$ d'où :

$$u_n = \frac{u_1 - \beta u_0}{\alpha - \beta}\alpha^n + \frac{\alpha u_0 - u_1}{\alpha - \beta}\beta^n$$

Cas n^o 2 : Les deux racines de $X^2 - aX - b = 0$ sont confondues (disons égales à α) et donc $\alpha^2 - a\alpha - b = 0$ mais aussi $a\alpha + 2b = 0$. La suite α^n est encore une solution mais il faut en trouver une autre : on remarque alors que $v_n := n\alpha^n$ vérifie la relation de récurrence linéaire (*) :

$$v_n - av_{n-1} - bv_{n-2} = \alpha^{n-2}(n\alpha^2 - a(n-1)\alpha - b(n-2)) = \alpha^{n-2}(a\alpha + 2b) = 0$$

Donc les deux suites α^n et $n\alpha^n$ forment une base des suites vérifiant (*) ; toute suite u_n vérifiant (*) s'écrit :

$$u_n = \lambda\alpha^n + \mu n\alpha^n$$

ou encore en terme de u_0 et u_1 on obtient :

$$u_n = u_0\alpha^n + \left(\frac{u_1}{\alpha} - u_0\right) n\alpha^n$$

Bien entendu, dans chaque cas, il est facile, en général, d'étudier la convergence de la suite : elle dépend des modules des racines α, β ; le seul cas délicat est celui où $|\alpha| = |\beta|$. Donnons quelques exemples :

1) (Suite de Fibonacci) $u_n = u_{n-1} + u_{n-2}$

L'équation caractéristique s'écrit $X^2 - X - 1 = 0$ et a pour racines $\alpha = \frac{1+\sqrt{5}}{2}$ (connu comme le "nombre d'or") et $\beta = \frac{1-\sqrt{5}}{2}$ et donc $u_n = \lambda\alpha^n + \mu\beta^n$ se comporte comme $\lambda\alpha^n$ (si $\lambda \neq 0$) où $\sqrt{5}\lambda = u_1 - \beta u_0$. On en tire $\lim u_n = +\infty$ si $u_1 > \beta u_0$ et $\lim u_n = -\infty$ si $u_1 < \beta u_0$. On observe aussi que (si $u_1 \neq \beta u_0$) alors $\lim \frac{u_{n+1}}{u_n} = \alpha$.

Par contre si $u_1 = \beta u_0$ alors $u_n = \mu\beta^n$ et $\lim u_n = 0$, $\lim \frac{u_{n+1}}{u_n} = \beta$.

2) Considérons les suites vérifiant $u_n = u_{n-1} - \frac{1}{4}u_{n-2}$.

L'équation caractéristique s'écrit $X^2 - X + \frac{1}{4} = 0$ et a pour racine double $\alpha = \frac{1}{2}$. On en déduit $u_n = u_0 2^{-n} + (2u_1 - u_0)n2^{-n}$ et donc $\lim u_n = 0$ et $\lim \frac{u_{n+1}}{u_n} = \frac{1}{2}$.

3) Considérons les suites vérifiant $u_n = u_{n-1} - u_{n-2}$.

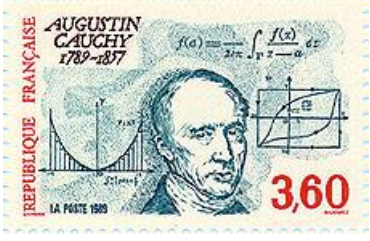
L'équation caractéristique s'écrit $X^2 - X + 1 = 0$ et a pour racines $\alpha = \frac{1+i\sqrt{3}}{2}$ et $\beta = \frac{1-i\sqrt{3}}{2}$. Observons que $\alpha^6 = \beta^6 = 1$ donc la suite $u_n = \lambda\alpha^n + \mu\beta^n$ vérifie $u_{n+6} = u_n$. La suite est donc périodique et elle n'est convergente que si elle est constante et nulle.

4) Considérons les suites vérifiant $u_n = 2\cos(\omega)u_{n-1} - u_{n-2}$.

L'équation caractéristique s'écrit $X^2 - 2\cos(\omega)X + 1 = 0$ et a pour racines $\alpha = e^{i\omega}$ et $\beta = e^{-i\omega}$. Si $\omega/\pi \in \mathbf{Z}$ alors $\cos(\omega) = \pm 1$ et l'équation caractéristique a une racine double $\alpha = \pm 1$ donc $u_n = (u_0 + (\pm u_1 - u_0)n)(\pm 1)^n$. Si $\omega/\pi \notin \mathbf{Z}$ alors on écrit

$$u_n = \frac{u_1 - e^{-i\omega}u_0}{2i\sin(\omega)}e^{in\omega} + \frac{e^{i\omega}u_0 - u_1}{2i\sin(\omega)}e^{-in\omega} = \frac{-u_0\sin((n-1)\omega) + u_1\sin(n\omega)}{\sin(\omega)}$$

Si $\omega/\pi \in \mathbf{Q}$ on voit que la suite est périodique et non constante (puisqu'on a supposé $\omega/\pi \notin \mathbf{Z}$) donc elle est bornée et non convergente. Si $\omega/\pi \notin \mathbf{Q}$ la suite est encore bornée (facile) et non convergente (un peu plus délicat à prouver).



Cauchy Augustin (1789–1857)

CHAPITRE 13 LIMITES ET CONTINUITÉ

La notion de continuité est intimement liée aux propriétés des nombres réels (Cantor appelait le cardinal de $\mathbf{R}$ *la puissance du continu). L'idée de la continuité est simple et peut se décrire de diverses façons imagées : "une fonction est continue si on peut tracer son graphe sans soulever le crayon du papier", "le graphe de la fonction n'a pas de saut" ou encore "si* x *se rapproche de* x_0 *alors* $f(x)$ *se rapproche de* $f(x_0)$*. En particulier on doit avoir la propriété suivante :*

$$\lim x_n = a \;\Rightarrow\; \lim f(x_n) = f(a)$$

On remarquera que cette propriété est fausse pour des fonctions simples comme les fonctions "en escalier" :

Tout comme la notion intuitive de limite, la notion de continuité est délicate à définir rigoureusement : la définition "epsilon-delta" donnée ci-dessous est due à Weierstrass et date donc du XIXe siècle. Voici une dernière approche avant la définition formelle : si l'on se donne un petit intervalle I *centré en* $f(a)$ *alors il existe un petit intervalle autour de* a *dont l'image par* f *est contenue dans* I*. Les définitions sont un peu arides et on aura intérêt à passer rapidement dessus jusqu'à arriver aux premières applications et revenir ensuite pour les approfondir.*

13.1 FONCTIONS CONTINUES

Commençons par définir la notion de limite ou de "tendre vers" pour une fonction.

Convention : dans tout ce chapitre on appellera constamment "intervalle" un "vrai" intervalle, c'est-à-dire un intervalle non vide et non réduit à un point (la plupart des énoncés seraient vides et sans intérêt pour des intervalles singleton).

Définition: Soit $f : I \to \mathbf{R}$ et x_0 un point de l'intervalle I, on dit que f *tend vers* ℓ quand x *tend vers* x_0 si

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ 0 < |x - x_0| \leq \delta \ \text{ et } \ x \in I \Rightarrow |f(x) - \ell| \leq \varepsilon$$

On écrit alors

$$\lim_{\substack{x \to x_0 \\ x \neq x_0}} f(x) = \ell$$

Remarques : 1) l'importance et la signification de prendre $x \neq x_0$ sont éclairées par l'exemple suivant : on regarde $f(x) := 1$ si $x \neq 0$ et $f(0) := 2$ alors

$$\lim_{\substack{x \to 0 \\ x \neq 0}} f(x) = 1 \neq f(0)$$

2) On peut définir de la même façon la limite d'une fonction à valeurs complexes et on montre alors facilement que comme pour les suites, $\lim f(x) = a + ib$ si et seulement si $\lim Re(f(x)) = a$ et $\lim Im(f(x)) = b$.

3) On peut définir des variantes de la notion de limite en faisant tendre la variable vers l'infini ou la fonction elle-même : par exemple une fonction $f(x)$ tends vers ℓ quand x tend vers l'infini si pour tout $\varepsilon > 0$ il existe $M > 0$ tel que $x \geq M$ entraîne $|f(x) - \ell| \leq \varepsilon$.

4) Il faut remarquer que la limite d'une fonction, si elle existe, est unique.

Définition: Une fonction $f : I \to \mathbf{R}$ est *continue* au point $x_0 \in I$ si

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ x \in I \text{ et } |x - x_0| \leq \delta \text{ entraîne } |f(x) - f(x_0)| \leq \varepsilon.$$

Il revient au même d'imposer :

$$\lim_{\substack{x \to x_0 \\ x \neq x_0}} f(x) = f(x_0)$$

Une fonction est *continue* sur I si elle est continue en tout point de I.

Remarque : à cause des quantificateurs, la définition ne change pas si on remplace les inégalités par des inégalités strictes, i.e. si on écrit $|x - x_0| < \delta$ et/ou $|f(x) - f(x_0)| < \varepsilon$.

Exemple : montrons directement à partir de la définition que la fonction $f(x) = x^2$ est continue sur $\mathbf{R}$.

Soit $x_0 \in \mathbf{R}$ et soit $\varepsilon > 0$; si $|x - x_0| \leq 1$ alors $|x + x_0| \leq 2|x_0| + 1$ donc si l'on choisit $\delta < \min\{1, \varepsilon/(2|x_0| + 1)\}$ et si $|x - x_0| \leq \delta$ on obtient $|x^2 - x_0^2| = |x + x_0||x - x_0| \leq \varepsilon$ ce qui prouve bien que la fonction $x \mapsto x^2$ est continue en x_0.

THÉORÈME: *Soit $f : I \to \mathbf{R}$ continue et soit $x_n \in I$; si $\lim x_n = \ell$ et si $\ell \in I$ alors $\lim f(x_n) = f(\ell)$.*

Démonstration: En effet soit $\varepsilon > 0$ alors il existe $\delta > 0$ tel que $|x - \ell| \leq \delta$ entraîne $|f(x) - f(\ell)| \leq \varepsilon$ (car f est continue en ℓ) mais comme $\lim x_n = \ell$, il existe $n_0 \in \mathbf{N}$ tel que si $n \geq n_0$ alors $|x_n - \ell| \leq \delta$ donc $|f(x_n) - f(\ell)| \leq \varepsilon$, ce qui est bien l'énoncé cherché. □

Remarque : on peut montrer que si pour toute suite l'implication du théorème est vraie alors la fonction f est continue.

On peut effectuer sur les limites de fonctions les mêmes opérations que sur les limites de suites (en fait si on prenait une définition plus abstraite on verrait qu'il s'agit du même concept).

THÉORÈME: *1) Si I est un intervalle contenant x_0 et si deux fonctions $f, g : I \to \mathbf{R}$ vérifient :*

$$\lim_{\substack{x \to x_0 \\ x \neq x_0}} f(x) = \ell \quad \text{et} \quad \lim_{\substack{x \to x_0 \\ x \neq x_0}} g(x) = \ell'$$

alors on a les formules suivantes :

i) $\lim\limits_{\substack{x \to x_0 \\ x \neq x_0}} af(x) + bg(x) = a\ell + b\ell'$

ii) $\lim\limits_{\substack{x \to x_0 \\ x \neq x_0}} f(x)g(x) = \ell\ell'$

iii) Si $\ell' \neq 0$ alors $\lim\limits_{\substack{x \to x_0 \\ x \neq x_0}} f(x)/g(x) = \ell/\ell'$

Dans le cas où l'une de ces limites est infinie ces formules restent valables à condition d'appliquer les règles suivantes :

$+\infty + \infty = +\infty$

$+\infty + a = +\infty$

$-\infty - \infty = -\infty$

$-\infty + a = -\infty$

Si $a > 0$ *alors* $a(+\infty) = +\infty$

si $a < 0$ *alors* $a(+\infty) = -\infty$

$(+\infty).(+\infty) = +\infty$

$(-\infty).(-\infty) = +\infty$

Si $\lim\limits_{\substack{x \to x_0 \\ x \neq x_0}} f(x) = 0$ *mais* $f(x) \neq 0$ *alors* $\lim \frac{1}{|f(x)|} = +\infty$. *De plus si en fait* $f(x) > 0$ *(resp.* $f(x) < 0$*) alors* $\lim \frac{1}{f(x)} = +\infty$ *(resp.* $= -\infty$*).*

2) Si I, J *sont deux intervalles contenant respectivement* a, x_0*, si* $f : I \to \mathbf{R}$ *et* $g : J \to \mathbf{R}$ *vérifient* $g(J) \subset I$ *(de sorte que* $f \circ g$ *est bien définie) et :*

$$\lim_{\substack{x \to x_0 \\ x \neq x_0}} g(x) = a$$

avec f *est continue en* a *avec* $f(a) = \ell$ *alors*

$$\lim_{\substack{x \to x_0 \\ x \neq x_0}} f \circ g(x) = \ell \quad .$$

Démonstration: 1) (i) On utilise l'inégalité $|af(x) + bg(x) - (a\ell + b\ell')| \leq |a||f(x) - \ell| + |b||g(x) - \ell'|$.

(ii) Commençons par observer que si f tend vers ℓ quand x tend vers x_0, alors il existe un intervalle autour de x_0 sur lequel f est bornée. On conclut alors en utilisant l'inégalité $|f(x)g(x) - \ell\ell'| \leq |f(x)||g(x) - \ell'| + |\ell'||f(x) - \ell|$.

(iii) On observe dans ce cas que $|1/g(x)| \leq 2/|\ell'|$ si x est suffisamment proche de x_0 et on utilise la majoration $|\frac{f(x)}{g(x)} - \frac{\ell}{\ell'}| \leq \frac{|f(x)|}{|g(x)||\ell'|}|g(x) - \ell'| + \frac{1}{|\ell'|}|f(x) - \ell|$.

La démonstration des cas particuliers se fait par les mêmes méthodes et est laissée au lecteur.

2) Soit $\varepsilon > 0$ alors il existe $\delta > 0$ tel que $|y - a| \leq \delta$ entraîne $|f(y) - \ell| \leq \varepsilon$ mais par ailleurs il existe $\eta > 0$ tel que $|x - x_0| \leq \eta$ entraîne $|g(x) - a| \leq \delta$ et donc $|f \circ g(x) - \ell| \leq \varepsilon$; ainsi $f \circ g$ tend bien vers ℓ quand x tend vers x_0. □

COROLLAIRE: *Soient $f, g : I \to$ deux fonctions continues alors :*
(i) $f + g$ et fg sont continues.
(ii) f/g et $f \circ g$ sont continues sur chaque intervalle où elles sont définies.

Démonstration: Immédiat à partir du théorème précédent en traduisant la continuité en terme de limites. Par exemple $\lim f \circ g(x) = f(\lim g(x)) = f \circ g(x_0)$ (la première égalité à cause de la continuité de f, la seconde à cause de la continuité de g) ; ainsi $f \circ g$ est bien continue. □

Comme les fonctions constantes et la fonction identité $x \mapsto x$ sont clairement continues, on obtient immédiatement

COROLLAIRE: Les fonctions polynômes sont continues, les fonctions rationnelles sont continues sur leurs intervalles de définition.

Si l'on sait (voir chapitre 16) que les fonctions "valeur absolue", $e^x, \log x, \cos(x), \sin(x)$ sont continues, on en tire que toutes les fonctions que l'on fabrique à partir de celles-ci (par opérations et composition) sont continues là où elles sont définies.

13.2 PROPRIÉTÉS DES FONCTIONS CONTINUES

Les énoncés de ce paragraphe sont assez parlants mais les démonstrations sont relativement délicates ; on pourra d'abord se concentrer sur l'application des résultats.

THÉORÈME: *(théorème des valeurs intermédiaires) Soit $I = [a, b]$ un intervalle fermé et $f : I \to \mathbf{R}$ une fonction continue ; si λ est un réel compris entre $f(a)$ et $f(b)$ alors il existe $c \in I$ tel que $f(c) = \lambda$. En d'autres termes l'image d'un intervalle par une fonction continue est un intervalle.*

Démonstration: Supposons par exemple $f(a) < f(b)$ et donc $f(a) \leq \lambda \leq f(b)$. Considérons l'ensemble $E := \{x \in I \mid f(x) \leq \lambda\}$; c'est un ensemble borné non vide (il contient a) donc il admet une borne supérieure $c := \sup(E)$. Montrons que $f(c) = \lambda$. Il existe une suite d'éléments x_n de E tels que $\lim x_n = c$ d'après les propriétés de la borne supérieure, par ailleurs $f(x_n) \leq \lambda$ par définition de E. Comme f est continue en c on en tire : $f(c) = \lim f(x_n) \leq \lambda$; Supposons $f(c) < \lambda$ et voyons que l'on aboutit à une contradiction. En effet il existerait alors $\varepsilon > 0$ tel que $f(c) + \varepsilon < \lambda$ et en utilisant de nouveau la continuité de f en c on obtient l'existence de $\delta > 0$ tel que $|x - c| < \delta$ entraîne $|f(x) - f(c)| < \varepsilon$ et donc $f(x) < f(c) + \varepsilon < \lambda$. On conclurait que $[c, c + \delta[\subset E$ ce qui contredirait que c est la borne supérieure de E. □

Graphiquement cela signifie qu'un graphe de fonction continue prenant des valeurs en dessous et au dessus d'une droite doit couper celle-ci en au moins un point.

APPLICATION: Un polynôme P à coefficients réels de degré impair possède une racine réelle. En effet si a_n est le coefficient du terme de plus haut degré n :

$$\lim_{x \to +\infty} P(x) = sgn(a_n)\infty \quad \text{et} \quad \lim_{x \to -\infty} P(x) = (-1)^n sgn(a_n)\infty$$

donc il existe a avec $P(a) < 0$ et il existe b avec $P(b) > 0$ donc il existe c avec $P(c) = 0$ d'après le théorème des valeurs intermédiaires.

THÉORÈME: *Soit $I = [a, b]$ un intervalle fermé et $f : I \to \mathbf{R}$ une fonction continue, alors :*

(i) f est bornée

(ii) f atteint son minimum et son maximum, c'est-à-dire que :

$$\exists c \in [a, b],\ f(c) = \sup\{f(x) \mid x \in [a, b]\} \quad \text{et} \quad \exists d \in [a, b],\ f(c) = \inf\{f(x) \mid x \in [a, b]\}$$

En particulier l'image d'un intervalle fermé par une application continue est un intervalle fermé.

Démonstration: (i) Considérons l'ensemble $E := \{x \in [a, b] \mid f \text{ est borné sur } [a, x]\}$. Visiblement E est un intervalle et si $c := \sup(E)$ on a $E = [a, c[$ ou $[a, c]$; on veut donc prouver que $c = b$ et que $c \in E$. Montrons que $c \in E$: comme f est continue en c il existe $\delta > 0$ tel que f soit bornée sur $[c - \delta, c]$, mais $c - \delta \in E$ donc f est bornée sur $[a, c - \delta]$ également et donc sur $[a, c]$.

Supposons maintenant $c < b$, la continuité de f en c entraîne qu'il existe un intervalle $[c-\delta, c+\delta] \subset [a, b]$ sur lequel f est bornée et donc comme précédemment on peut conclure que f est bornée sur $[a, c+\delta]$, donc $c+\delta \in E$ ce qui contredit que c est la borne supérieure de E.

(ii) Posons $M := \sup_{x\in[a,b]} f(x)$. On définit une suite d'intervalles emboités $I_0 := [a, b], \ldots, I_n := [a_n, b_n]$ de sorte que à chaque pas I_{n+1} soit égal soit à $[a_n, (a_n + b_n)/2]$ soit à $[(a_n + b_n)/2, b_n]$ et tel que $\sup_{x\in I_n} f(x) = M$. On choisit $c \in \cap_n I_n$ (cette intersection est non vide d'après la propriété des segments emboités) et on utilise la continuité de f en c pour conclure que pour tout $\varepsilon > 0$, il existe n tel que $x \in I_n$ entraîne $|f(x) - f(c)| \leq \varepsilon/2$. Par ailleurs, comme $\sup_{x\in I_n} f(x) = M$, il existe $x \in I_n$ tel que $f(x) \geq M - \varepsilon/2$ donc finalement $M \geq f(c) \geq M - \varepsilon$. Ceci étant valable pour tout $\varepsilon > 0$ on conclut bien que $f(c) = M$. $\square$

Remarquons que la conclusion peut être fausse si f n'est pas continue ou si l'intervalle n'est pas borné. Par exemple $f(x) = 1/x$ n'est pas bornée sur $]0, 1]$ et n'atteint pas son minimum sur $[1, +\infty)$.

THÉORÈME: *Soit f une fonction continue strictement croissante sur un intervalle d'extrémités a et b. Supposons que*

$$\lim_{\substack{x \to a \\ x > a}} f(x) = \alpha \quad \text{et} \quad \lim_{\substack{x \to b \\ x < b}} f(x) = \beta$$

Alors f détermine une bijection de l'intervalle d'extrémités a et b vers l'intervalle d'extrémités α et β et la bijection réciproque est également continue et croissante.

Remarque : en particulier si f est continue et croissante sur I contenant $[a, b]$ alors $f([a, b]) = [f(a), f(b)]$.

Démonstration: Comme $x < y$ entraîne $f(x) < f(y)$, la fonction f est injective. Soit $y \in]\alpha, \beta[$ alors il existe $x_1 \in]a, b[$ tel que $f(x_1) < y$ et il existe $x_2 \in]a, b[$ tel que $f(x_2) > y$ donc, d'après le théorème des valeurs intermédiaires, il existe $c \in]x_1, x_2[$ tel que $f(c) = y$. Ainsi f est surjective et définit une bijection de $]a, b[$ sur $]\alpha, \beta[$. Voyons que f^{-1} est croissante et continue. Si $x < y$ et $f^{-1}(x) \geq f^{-1}(y)$, on aurait alors $x = f(f^{-1}(x)) \geq f(f^{-1}(y)) = y$ ce qui est contradictoire. Choisissons maintenant $y_0 \in]\alpha, \beta[$ et montrons que f est continue en y_0. Soit donc $\varepsilon > 0$, on sait que $y_0 = f(x_0)$ avec $x_0 \in]a, b[$ et on peut choisir $a', b' \in]a, b[$ tels que $x_0 - \varepsilon \leq a' < x_0 < b' \leq x_0 + \varepsilon$. On sait que si $\alpha' := f(a')$ et $\beta' := f(b')$ alors $f(]a', b'[) =]\alpha', \beta'[$ et donc que si $y \in]\alpha', \beta'[$ alors $f^{-1}(y) \in]a', b'[$ et donc $|f(y) - f(y_0)| \leq \varepsilon$; ce qui montre que f^{-1} est continue en y_0. □

Remarque : on obtient ainsi des bijections de $[a, b]$ vers $[\alpha, \beta]$ ou de $]a, b[$ vers $]\alpha, \beta[$ etc.

On exploitera au chapitre 16 ce théorème pour définir les fonctions exponentielle (comme fonction réciproque du logarithme), Arcsin, Arcos, Arctg, Argsh, Argch et Argth. Donnons ici un exemple de construction de fonction réciproque : Soit la fonction $f(x) := x + x\log(x)$; on vérifie aisément qu'elle est strictement croissante sur $[1, +\infty)$ (en effet log est croissante et positive sur $[1, +\infty)$ donc $x > y$ entraîne $1 + \log(x) > 1 + \log(y) \geq 1$ donc $x(1 + \log(x)) > y(1 + \log(y))$). Ainsi f définit une bijection de $[1, +\infty)$ sur $[1, +\infty)$; pour tracer le graphe de la fonction $g = f^{-1}$ on prend simplement le symétrique du graphe de f par rapport à la droite $y = x$.

13.3 LIMITES DE FONCTIONS

On entame ici l'étude de quelques techniques de calculs de limites qui seront beaucoup plus développées au chapitre 19 (développements limités).

Définition: Soient deux fonctions f et g définies sur un intervalle I ; elles sont *équivalentes* au voisinage de $x_0 \in I$ (resp. de $\pm\infty$) si

$$\lim_{x\to x_0} (f(x)/g(x)) = \lim_{x\to x_0} (g(x)/f(x)) = 1$$

On écrit alors $f(x) \sim g(x)$ en sous-entendant le point où f et g sont équivalentes.

Commentaire : l'idée d'un équivalent est simple : dire que $f(x) \sim 1$ quand x tend vers 0, équivaut à dire que $\lim f(x) = 1$ mais $f(x) \sim x$ entraîne et est beaucoup plus précis que $\lim f(x) = 0$: cela dit avec quelle vitesse $f(x)$ tend vers zéro (par exemple moins vite que x^2 mais plus vite que $\sqrt[3]{x}$).

Exemples : On verra au chapitre 16 une démonstration géométrique du fait suivant : $\sin(x) \sim x$ lorsque x tend vers 0 ; de même $\cos(x) - 1 \sim x^2/2$.

Soit $F(x) := \frac{a_m x^m + \ldots + a_0}{b_n x^n + \ldots + b_0}$ une fraction rationnelle (avec $a_m b_n \neq 0$), alors lorsque x tend vers $\pm\infty$ on a $F(x) \sim \frac{a_m}{b_n} x^{m-n}$. En effet
$F(x)x^{n-m} = \left(a_m + \ldots + \frac{a_0}{x^m}\right) / \left(b_n + \ldots + \frac{b_0}{x^n}\right)$ tend bien vers a_m/b_n.

Remarque : Il ne faut pas confondre $f(x) \sim g(x) + h(x)$ et $f(x) - g(x) \sim h(x)$; par exemple si $f(x) = 2x^2 + x$, $g(x) = 2x^2$ et $h(x) = 1$ alors $f(x) \sim g(x) + h(x)$ mais $f(x) - g(x) \sim x \not\sim h(x)$ quand x tend vers l'infini. Parmi les autres pièges à éviter, signalons aussi que $f(x) \sim g(x)$ n'entraîne pas $h \circ f \sim h \circ f(x)$, même si h est continue ; en effet en choissant f et g comme dans l'exemple précédent et $h = \exp$ on a $h \circ f(x)/h \circ g(x) = e^x$ qui ne tend pas vers 1 (quand x tend vers l'infini).

On ne peut pas donc en général ajouter des équivalents mais on peut les multiplier :

PROPOSITION: *Supposons que $f(x) \sim g(x)$ et $h(x) \sim k(x)$ quand x tend vers a alors $f(x)h(x) \sim g(x)k(x)$ quand x tend vers a.*

Démonstration: En effet, si $\lim \left(\frac{f(x)}{g(x)}\right) = 1$ et $\lim \left(\frac{h(x)}{k(x)}\right) = 1$ alors $\lim \left(\frac{f(x)h(x)}{g(x)k(x)}\right) = 1$. □

Exemple : lorsque x tend vers l'infini $exp(x - 1/x) \sim exp(x)$ et $x^m + \ldots + a_0 \sim x^m$ donc $(x^m + \ldots + a_0)exp(x - 1/x) \sim x^m exp(x)$.

PROPOSITION: (prolongement par continuité) *Soit $c \in [a, b]$ et $f : [a, b] \setminus \{c\} \to \mathbf{R}$ une fonction continue et supposons que $\lim\limits_{\substack{x \to c \\ x \neq c}} f(x) = \ell$. Alors la fonction $\bar{f} : [a, b] \to \mathbf{R}$ définie par*

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

est continue.

Démonstration: Le seul problème est bien sûr la continuité en c mais on a par hypothèse : $\lim\limits_{\substack{x \to c \\ x \neq c}} \bar{f}(x) = \ell = \bar{f}(c)$ donc $\bar{f}$ est continue en c. □

Exemples : La fonction $f(x) = \sin(x)/x$ (si $x \neq 0$) et $f(0) = 1$ est continue.
La fonction $g(x) = (1 - \cos(x))/x^2$ (si $x \neq 0$) et $g(0) = 1/2$ est continue.
La fonction $h(x) = \sin(x)/x$ (si $x \neq 0$) et $h(0) = 1$ est continue.
La fonction $k(x) = \exp(-1/x^2)$ (si $x \neq 0$) et $k(0) = 0$ est continue.

Euler Leonhard (1707-1783)

CHAPITRE 14 DÉRIVÉES ET FORMULE DE TAYLOR

Vous connaissez déjà la notion de dérivée d'une fonction (en un point) dont nous rappelons néanmoins ci-dessous la définition ainsi que l'interprétation géométrique : la tangente à une courbe en un point (quand elle existe) est la droite qui "épouse" le mieux possible la forme de la courbe en ce point ; la dérivée d'une fonction f est la pente de la tangente à la courbe d'équation $y = f(x)$. Il est assez facile de voir que la dérivée d'une fonction croissante est positive, mais la réciproque qui est délicate et généralement admise en terminale est démontrée ici. Pour étudier plus finement une courbe, après avoir déterminé sa tangente en un point on peut chercher quel est le cercle ou la parabole qui l'approche le mieux ; on peut aussi vouloir savoir si le point est un point d'inflexion – la courbe traverse-t-elle sa tangente en ce point? – ou si la fonction atteint en un point un maximum ou un minimum. La réponse à toutes ces questions est donnée par la formule de Taylor, qui grosso modo décrit la courbe $y = f(x)$ près du point x_0 en fonction de ses dérivées successives en x_0.

14.1 RAPPELS SUR LE CALCUL DES DÉRIVÉES

Comme au chapitre précédent nous prenons comme convention que les intervalles sont des intervalles non vides et non réduits à un point.

Définition: Soit I un intervalle réel ouvert et $x_0 \in I$, soit $f : I \to \mathbf{R}$ une fonction à valeurs réelles ; on dit que f est *dérivable* en x_0 si la limite

$$\lim_{\substack{x \to x_0 \\ x \neq x_0}} \frac{f(x)-f(x_0)}{x-x_0}$$

existe et dans ce cas cette limite s'appelle la *dérivée* de f en x_0 et se note $f'(x_0)$ ou $\frac{df}{dx}(x_0)$.

Si f est dérivable en tout point, la fonction $x \mapsto f'(x)$ ainsi définie s'appelle la *fonction dérivée*. Si f' est elle même dérivable, sa dérivée s'appelle la *dérivée seconde* de f et se note f'' ou $f^{(2)}$. On définit de même la *dérivée n-ième* (si elle existe, on la note $f^{(n)}$ ou $\frac{d^n f}{dx^n}$) par la relation de récurrence

$$f^{(0)}(x) = f(x),\ f^{(1)}(x) = f'(x),\ f^{(n+1)}(x) = \left(f^{(n)}\right)'(x)$$

Exemples : Une fonction qui n'est pas continue en un point x_0 n'est pas dérivable ; la fonction $f(x) := |x|$ est continue au point $x_0 := 0$ mais n'y est pas dérivable (en effet la limite de $|x|/x$ en restreignant x à être positif est $+1$ alors qu'en restreignant x à être négatif la limite est -1).

Calculons directement à partir de la définition la dérivée de $f(x) := 3x^2 - 5x + 4$ en $x_0 = 1$: on a $\frac{f(x)-f(1)}{x-1} = \frac{3x^2-5x+2}{x-1} = 3x - 2$ et donc

$$f'(1) = \lim_{\substack{x \to 1 \\ x \neq 1}} \frac{f(x)-f(1)}{x-1} = \lim_{x \to 1} 3x - 2 = 1$$

Pour calculer les dérivées successives on calcule plus généralement $f'(x_0) = 6x_0 - 5$ puis $f''(x_0) = 6$ puis $f^{(n)}(x_0) = 0$ pour $n \geq 3$.

Interprétation : Si l'on trace la courbe $y = f(x)$ la dérivée $f'(x_0)$ s'interprète comme la pente de la droite tangente au point de coordonnées $(x_0, f(x_0))$ à la courbe.

L'équation de la tangente est :

$$y - f(x_0) = f'(x_0)(x - x_0)$$

On peut aussi écrire la définition de la dérivée sous la forme d'un "développement limité à l'ordre 1" (voir le chapitre 19) :

$$f(x_0 + h) = f(x_0) + f'(x_0)h + \varepsilon(h)h$$

avec $\lim_{h \to 0} \varepsilon(h) = 0$.

Sous cette forme il est clair qu'une fonction dérivable en x_0 est continue en x_0 puisqu'alors $\lim_{h \to 0} f(x_0 + h) = f(x_0)$. Une autre façon de dire que f est dérivable en x_0 est de dire que la fonction $g(x) := \frac{f(x) - f(x_0)}{x - x_0}$ peut être prolongée par continuité en x_0 en lui donnant la valeur $g(x_0) = f'(x_0)$.

Variantes :

1) Il est parfois utile de définir la notion de *dérivée à droite* notée $f'_d(x_0)$ (resp. *à gauche* notée $f'_g(x_0)$) d'une fonction en un point x_0 :

$$f'_d(x_0) = \lim_{\substack{x \to x_0 \\ x > x_0}} \frac{f(x) - f(x_0)}{x - x_0} \quad \text{et} \quad f'_g(x_0) = \lim_{\substack{x \to x_0 \\ x < x_0}} \frac{f(x) - f(x_0)}{x - x_0}$$

Par exemple si $f(x) = |x|$ alors $f'_d(0) = 1$ et $f'_g(0) = -1$. Une fonction est dérivable en x_0 si et seulement si elle est dérivable à droite et à gauche et ces deux dérivées sont égales.

2) Dans le cas où la limite de $\frac{f(x) - f(x_0)}{x - x_0}$ vaut $\pm\infty$, on ne considère pas la fonction dérivable mais on peut parler de tangente verticale : c'est le cas par exemple de la fonction $f(x) := \sqrt[3]{x}$ au point $x_0 = 0$:

3) Parfois une autre notation pour le calcul des dérivées est utile ; on aurait pu définir la dérivée en x_0 de f comme

$$f'(x_0) = \lim_{\substack{h \to 0 \\ h \neq 0}} \frac{f(x_0 + h) - f(x_0)}{h}$$

Révisons rapidement les règles principales de calcul des dérivées :

THÉORÈME: *Supposons f et g dérivables, alors $f+g$ et fg sont dérivables, f/g et $g\circ f$ sont dérivables là où elles sont définies ; si f est une bijection et $f'(x)\neq 0$ alors f^{-1} est dérivable en x. De plus on a les règles de calcul suivantes*

(i) (dérivée d'une somme) $(f+g)'=f'+g'$

(ii) (dérivée d'un produit) $(fg)'=f'g+fg'$

(iii) (dérivée d'un quotient) $\left(\frac{f}{g}\right)'=\frac{f'g-fg'}{g^2}$

(iv) (dérivée d'une fonction composée) $(g\circ f)'=f'.g'\circ f$

(v) (dérivée d'une fonction réciproque) Si f admet une bijection réciproque, si $y_0=f(x_0)$ (et donc $x_0=f^{-1}(y_0)$) et si $f'(x_0)\neq 0$ alors $\left(f^{-1}\right)'(y_0)=\frac{1}{f'(x_0)}$.

Démonstration: Laissons au lecteur le soin de réviser les trois premières formules et démontrons les deux dernières qui sont plus délicates.

(iv) Posons $f(x_0)=y_0$ et $g(y_0)=z_0$, d'après les hypothèses $f(x)-y_0=(x-x_0)(f'(x_0)+\varepsilon_1(x-x_0))$ et $g(y)-z_0=(y-y_0)(g'(y_0)+\varepsilon_2(y-y_0))$ avec $\lim_{h\to 0}\varepsilon_1(h)=\lim_{h\to 0}\varepsilon_2(h)=0$. En remplaçant y par $f(x)$, on en tire

$$\begin{aligned} g\circ f(x)-z_0 =&(f(x)-y_0)(g'(y_0)+\varepsilon_2(f(x)-y_0)\\ =&(x-x_0)(f'(x_0)+\varepsilon_1(x-x_0))(g'(y_0)+\varepsilon_2\left((x-x_0)(f'(x_0)+\varepsilon_1(x-x_0))\right)\\ =&(x-x_0)f'(x_0)g'(y_0)+(x-x_0)\eta(x-x_0)\end{aligned}$$

Il reste à vérifier que $\lim_{h\to 0}\eta(h)=0$ et on aura donc bien montré que $g\circ f$ est dérivable et que $(g\circ f)(x_0)=f'(x_0)g'(y_0)$. Or quand x tend vers x_0, la quantité $f'(x_0)+\varepsilon_1(x-x_0)$ tend vers $f'(x_0)$ et donc $(x-x_0)(f'(x_0)+\varepsilon_1(x-x_0))$ tend vers 0 et donc ε_2 de cette dernière quantité tend bien vers 0.

(v) Par hypothèse $f(x)-y_0=(x-x_0)(f'(x_0)+\varepsilon(x-x_0))$ avec $\lim_{h\to 0}\varepsilon(h)=0$; comme $f'(x_0)\neq 0$ on sait que si $|x-x_0|$ est très petit alors $|\varepsilon(x-x_0)|<|f'(x_0)|$ et $f'(x_0)+\varepsilon(x-x_0)\neq 0$ et on peut alors écrire :

$$\frac{x-x_0}{f(x)-f(x_0)}=\frac{1}{f'(x_0)+\varepsilon(x-x_0)}$$

Changeons de notation et appelons $g(y)=f^{-1}(y)=x$; l'égalité précédente peut se réécrire :

$$\frac{g(y)-g(y_0)}{y-y_0}=\frac{1}{f'(x_0)+\varepsilon(g(y)-g(y_0))}$$

or $\lim_{y\to y_0}\varepsilon(g(y)-g(y_0))=0$ puisque g est continue et $\lim_{h\to 0}\varepsilon(h)=0$ donc on obtient bien que $g=f^{-1}$ est dérivable et que $g'(y_0)=\frac{1}{f'(x_0)}$. □

Ces règles, jointes à la connaissance de quelques dérivées permettent de calculer toutes les dérivées "usuelles" (voir le chapitre 16 pour des révisions et compléments concernant les fonctions usuelles).

THÉORÈME: (calcul de dérivée de fonctions usuelles) *Les formules suivantes sont valables*

$(x^n)' = nx^{n-1}$ *pour* $n \in \mathbf{N}$ *et plus généralement (si* $x \in \mathbf{R}_+^*$ *et* $a \in \mathbf{R}$*) on* $(x^a)' = ax^{a-1}$
$(log|x|)' = \frac{1}{x}$
$(e^x)' = e^x$
$(\sin(x))' = \cos(x)$ *et* $(\cos(x))' = -\sin(x)$

Démonstration: La première formule s'obtient, en faisant une récurrence sur n et en utilisant la dérivée d'un produit : si $(x^n)' = nx^{n-1}$ alors $(x^{n+1})' = (xx^n)' = x^n + xnx^{n-1} = (n+1)x^n$. Par définition on a $x^a := e^{a\log(x)}$, en appliquant les deuxième et troisième formules et la règle de dérivation de fonctions composées on obtient : $(x^a)' = (a\log(x))'e^{a\log(x)} = ax^{a-1}$.

Nous réviserons dans le chapitre 16 (sur les "fonctions usuelles") le fait que, par construction, $\forall x \in \mathbf{R}_+^*, (\log x)' = \frac{1}{x}$. Si maintenant $x \in \mathbf{R}_-^*$ alors $\log|x| = \log(-x)$ et en utilisant la règle de calcul des dérivées de fonctions composées on obtient bien $(log|x|)' = -\left(\frac{-1}{x}\right) = \frac{1}{x}$.

Comme l'exponentielle est la fonction réciproque du logarithme, la règle (v) du théorème précédent permet de calculer en posant $y_0 = \log x_0$, $f(x) = \log x$ et $g(y) = e^y$:

$$g'(y_0) = \frac{1}{f'(x_0)} = x_0 = e^{y_0}$$

Pour calculer la dérivée de la fonction sinus admettons provisoirement la formule suivante, qui sera démontrée par un argument géométrique au chapitre 16 :

$$\lim_{\substack{h \to 0 \\ h \neq 0}} \frac{\sin(h)}{h} = 1$$

De la relation $\cos^2(h) - 1 = -\sin^2(h)$ on tire

$$\frac{\cos(h)-1}{h} = \frac{\cos^2(h)-1}{h}\frac{1}{\cos(h)+1} = -\left(\frac{\sin(h)}{h}\right)^2 \frac{h}{\cos(h)+1}$$

et donc $\frac{\cos(h)-1}{h}$ tend vers 0 quand h tend vers 0. On peut utiliser cela pour exprimer ainsi :

$$\begin{aligned}\frac{\sin(x+h)-\sin(x)}{h} &= \frac{\sin(x)\cos(h)+\cos(x)\sin(h)-\sin(x)}{h} \\ &= \cos(x)\left(\frac{\sin(h)}{h}\right) + \sin x\left(\frac{\cos(h)-1}{h}\right),\end{aligned}$$

d'où l'on tire

$$(\sin(x))' = \lim_{h\to 0}\frac{\sin(x+h)-\sin(x)}{h} = \cos(x).$$

Le calcul pour la fonction cosinus est similaire et laissé au lecteur. □

Exemples de calculs :

Soit $f(x) = \left(e^{\cos(x)} + 1\right)^3$ et posons comme au deuxième chapitre $k(x) := \cos(x)$, $h(x) := e^x$ et $g(x) := (x+1)^3$ de sorte que $f = g \circ h \circ k$. En appliquant deux fois la règle de dérivation de fonctions composées on obtient : $f'(x) = (h \circ k)'(x) g'(h \circ k(x)) = k'(x) h'(k(x)) g'(h \circ k(x))$; par ailleurs $k'(x) = -\sin(x)$, $h'(x) = e^x$ et $g'(x) = 3(x+1)^2$ d'où $f'(x) = -3\sin(x) e^{\cos(x)} \left(e^{\cos(x)} + 1\right)^2$.

Prenons maintenant $f(x) = (\cos(2\sqrt{x}+1) + e^{3x})/(\mathrm{tg}(x^2) + 4x + 1)^3$. Pour calculer la dérivée de f on la décompose : $f = uv^{-3}$ avec $u := \cos(2\sqrt{x}+1) + e^{3x}$ et $v := \mathrm{tg}(x^2) + 4x + 1$; la règle de dérivation d'un quotient jointe à celle d'une puissance donne $f' = (u'v - 3uv')v^{-4}$. Pour calculer u' posons $g(x) := 2\sqrt{x} + 1$ alors $g'(x) = \frac{1}{\sqrt{x}}$ et $(\cos(g(x))' = -g'(x)\sin(g(x)) = -\frac{1}{\sqrt{x}} \sin(2\sqrt{x}+1)$ et enfin $u' = -\frac{1}{\sqrt{x}} \sin(2\sqrt{x}+1) + 3e^{3x}$. D'autre part $(\mathrm{tg}(x))' = (\sin(x)/\cos(x))' = 1/\cos^2(x)$ donc $(\mathrm{tg}(x^2))' = 2x/\cos^2(x)$ et $v' = 4 + 2x/\cos^2(x)$.

APPLICATION: on peut déduire du théorème les règles suivantes :

(i) (dérivée d'une puissance) $(u^n)' = nu'u^{n-1}$

(ii) (dérivée "logarithmique" d'un produit) si $v = u_1^{m_1} \dots u_r^{m_r}$ alors $\frac{v'}{v} = m_1 \frac{u_1'}{u_1} + \dots + m_r \frac{u_r'}{u_r}$

(iii) la dérivée d'une fonction de période T (i.e. telle que $f(x+T) = f(x)$) est encore périodique de période T.

(iv) la dérivée d'une fonction paire (resp. impaire) est impaire (resp. paire).

Démonstration: (i) on applique la formule de dérivée d'une fonction composée à $f \circ u$ où $f(x) = x^n$.

(ii) Faisons le calcul lorsque $v = u_1 u_2$ et laissons au lecteur le soin d'écrire une démonstration par récurrence pour le cas général. $\frac{v'}{v} = \frac{u_1' u_2 + u_1 u_2'}{u_1 u_2} = \frac{u_1'}{u_1} + \frac{u_2'}{u_2}$.

(iii) Posons $g(x) = x + T$ alors $f = f \circ g$ et donc $f'(x) = (f \circ g)'(x) = f'(g(x))g'(x) = f'(x+T)$.

(iv) Supposons $f(-x) = \varepsilon f(x)$ (avec $\varepsilon = \pm 1$) et posons $i(x) = -x$ alors $(f \circ i)'(x) = \varepsilon f'(x) = f'(i(x))i'(x) = -f(-x)$ d'où l'énoncé. □

Le calcul de la dérivée n-ième d'une fonction ne pose pas plus de problème théorique mais peut être difficile à mener dans la pratique. En fait il est souvent intéressant seulement de savoir qu'une fonction possède une dérivée d'un certain ordre.

Définition: Un fonction $f :]a,b[\to \mathbf{R}$ est n fois continûment dérivable ou *de classe* $\mathcal{C}^n$ si elle possède des dérivées continues jusqu'à l'ordre n. Un fonction $f :]a,b[\to \mathbf{R}$ est *indéfiniment dérivable* ou *de classe* $\mathcal{C}^\infty$ si elle possède des dérivées à tout ordre.

Exemples : Les fonctions polynômes, exponentielles, logarithmes etc sont $\mathcal{C}^\infty$ de même que les fonctions "compliquées" dont on a calculé la première dérivée (pour la preuve, il suffit de formuler convenablement une récurrence). La fonction $|x|$ est seulement de classe $\mathcal{C}^0$, c'est-à-dire continue. La fonction définie par :

$$f(0) := 0 \quad \text{et} \quad f(x) = x^2 \sin\left(\frac{1}{x}\right) \quad \text{si } x \neq 0$$

est continue et dérivable ; en effet $f'(0) = \lim_{x\to 0} x\sin\left(\frac{1}{x}\right) = 0$ et si $x \neq 0$ on a $f'(x) = 2x\sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right)$. Toutefois $f'(x)$ n'est pas continue en 0 puisque $\cos\left(\frac{1}{x}\right)$ n'a pas de limite quand x tend vers 0. Cette fonction n'est donc pas de classe $\mathcal{C}^1$ sur $\mathbf{R}$.

Voici une formule bien utile donnant la dérivée n-ième d'un produit :

THÉORÈME: *(formule de Leibniz) Soient f et g deux fonctions n fois dérivables, alors le produit fg est n fois dérivable et la dérivée n-ième est donnée par la formule :*

$$(fg)^{(n)} = \sum_{i=0}^{n} C_n^i f^{(i)} g^{(n-i)}$$

Démonstration: La preuve se fait bien sûr par récurrence sur l'entier n. Pour $n = 0$ la formule dit que $fg = fg$, pour $n = 1$ la formule dit que $(fg)' = fg' + f'g$, ce que l'on sait déjà. Supposons donc la formule vrai pour $n-1$ et déduisons-en la formule pour n. Donc on suppose

$$(fg)^{(n-1)} = \sum_{i=0}^{n-1} C_{n-1}^i f^{(i)} g^{(n-1-i)}$$

on en déduit que $(fg)^{(n)} = \left((fg)^{(n-1)}\right)'$ vaut

$$\sum_{i=0}^{n-1} C_{n-1}^i \left(f^{(i+1)} g^{(n-1-i)} + f^{(i)} g^{(n-i)}\right) = fg^{(n)} + \sum_{j=1}^{n-1} f^{(j)} g^{(n-j)} \left(C_{n-1}^j + C_{n-1}^{j-1}\right) + f^{(n)} g$$

or d'après la règle du triangle de Pascal $C_{n-1}^j + C_{n-1}^{j-1} = C_n^j$ ce qui donne la formule de Leibniz pour n. □

14.2 THÉORÈME DES ACCROISSEMENTS FINIS, FORMULE DE TAYLOR

Commençons par une propriété simple des points où une fonction atteint un maximum (ou un minimum) :

THÉORÈME: *Soit $f :]a, b[\to \mathbf{R}$ une fonction qui atteint un maximum (ou minimum) en x_0 ; si f est dérivable en x_0, alors $f'(x_0) = 0$.*

Démonstration: Supposons que f atteigne un maximum en c (le raisonnement est analogue si c'est un minimum) ; si $x > c$ alors $\frac{f(x)-f(c)}{x-c} \leq 0$ et donc

$$f'(c) = \lim_{\substack{x \to c \\ x > c}} \frac{f(x)-f(c)}{x-c} \leq 0$$

de même si $x < c$ alors $\frac{f(x)-f(c)}{x-c} \geq 0$ et donc

$$f'(c) = \lim_{\substack{x \to c \\ x < c}} \frac{f(x)-f(c)}{x-c} \geq 0$$

et on peut conclure que $f'(c) = 0$. □

Le lemme qui suit est géométriquement assez intuitif :

LEMME: *(Lemme de Rolle) Soit $f : [a, b] \to \mathbf{R}$ une fonction continue sur $[a, b]$ et dérivable sur $]a, b[$ telle que $f(a) = f(b) = 0$, alors il existe $c \in]a, b[$ tel que $f'(c) = 0$.*

Démonstration: Si f est constamment nulle, la conclusion est évidente, sinon on a l'inégalité $\sup_{x\in[a,b]} |f(x)| > 0$ et de plus d'après un théorème du chapitre précédent cette borne supérieure est atteinte en un point c (nécessairement distinct de a et b). Il existe donc $c \in]a, b[$ tel que ou bien $\forall x \in [a, b],\ f(x) \leq f(c)$ ou bien $\forall x \in [a, b],\ f(x) \geq f(c)$. D'après le théorème précédent on doit avoir $f'(c) = 0$. □

Nous pouvons maintenant démontrer le théorème des accroissements finis

THÉORÈME: (théorème des accroissements finis) *Soit $f : [a, b] \to \mathbf{R}$ une fonction continue sur $[a, b]$ et dérivable sur $]a, b[$ alors :*

(1ère forme) Il existe $c \in]a, b[$ tel que $f(b) - f(a) = f'(c)(b - a)$.

(2ème forme) Supposons de plus la fonction dérivée bornée sur $]a, b[$ et posons $M := sup_{x\in]a,b[}|f'(x)|$ alors $|f(b) - f(a)| \leq M|b - a|$.

Démonstration: L'idée est de modifier la fonction f par une fonction linéaire de sorte que l'on puisse lui appliquer le lemme de Rolle. Considérons la fonction $g(x) := (b - a)f(x) + (f(a) - f(b))x + af(b) - bf(a)$ alors g est bien continue et dérivable avec $g'(x) = (b-a)f'(x)+f(a)-f(b)$. Par ailleurs par construction $g(b) = (b-a)f(b)+(f(a)-f(b))b+af(b)-bf(a) = 0$ et $g(a) = (b-a)f(a)+(f(a)-f(b))a+af(b)-bf(a) = 0$ donc il existe $c \in]a, b[$ tel que $g'(c) = (b - a)f'(c) + f(a) - f(b) = 0$ ce qui donne la première partie de l'énoncé. La deuxième est claire en notant que $|f'(c)| \leq M$. □

Commentaire. Bien que la deuxième forme du théorème soit moins précise, c'est la plus utile : c'est celle qui se généralise aux fonctions de plusieurs variables. C'est cet énoncé qui

permet de calculer sur ordinateur une valeur approchée du minimum (ou maximum) d'une fonction sur un intervalle, à condition que la fonction dérivable et qu'on sache controler sa dérivée.

COROLLAIRE: *Soit $f : [a, b] \to \mathbf{R}$ une fonction continue sur $[a, b]$ et dérivable sur $]a, b[$;*

La fonction f est constante si et seulement si $f' = 0$

La fonction f est croissante (resp. décroissante) si et seulement si $f' \geq 0$ (resp. $f' \leq 0$)

De plus si $f' > 0$ (resp. $f' < 0$) alors f est strictement croissante (resp. strictement décroissante).

Démonstration: Il est clair que si f est constante alors $f' = 0$; réciproquement supposons $f' = 0$ alors pour tout $x \in]a, b]$, le théorème des accroissements finis donne l'existence d'un $c \in]a, x[$ tel que $f(x) - f(a) = (x - a)f'(c) = 0$ donc $f(x) = f(a)$ est bien constante.

Si f est croissante alors le rapport $\frac{f(x)-f(x_0)}{x-x_0}$ est toujours positif, donc la limite quand x tend vers x_0 (si elle existe) est positive. Réciproquement supposons que pour $c \in]a, b[$ on ait $f'(c) \geq 0$; si $a \leq x < y \leq b$ alors le théorème des accroissements finis appliqué à l'intervalle $[x, y]$ donne $f(y) - f(x) = (y - x)f'(c) \geq 0$ (et si $f'(c) > 0$ on a même $f(y) - f(x) > 0$) d'où le résultat. Le raisonnement est bien sûr identique pour les fonctions décroissantes. □

On a vu (Cf Chapitre 6) que si $f(x)$ est une fonction polynôme de degré $\leq n$ alors :

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

Bien sûr une telle formule ne peut être exacte si f n'est pas un polynôme mais il est raisonnable de penser qu'elle donne une bonne approximation de la fonction f au voisinage du point a ; c'est le contenu de l'énoncé suivant :

THÉORÈME: *(Formule de Taylor-Lagrange) Soit $f : [a, b] \to \mathbf{R}$ une fonction continue sur $[a, b]$ et $n + 1$ fois dérivable sur $]a, b[$ alors il existe $c \in]a, b[$ tel que*

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

Si l'on suppose de plus que $|f^{(n+1)}|$ est bornée par disons M alors on en déduit

$$|f(b) - f(a) - f'(a)(b - a) - \frac{f''(a)}{2!}(b - a)^2 - \ldots - \frac{f^{(n)}(a)}{n!}(b - a)^n| \leq \frac{M}{(n + 1)!}|b - a|^{n+1}$$

Démonstration: Considérons la fonction $g(x) := f(x) - f(b) + \sum_{i=1}^{n} \frac{(b-x)^i}{i!} f^{(i)}(x)$ alors un calcul simple montre que $g'(x) = f^{(n+1)}(x)\frac{(b-x)^n}{n!}$. Choisissons maintenant λ tel que la fonction $h(x) := g(x) + \lambda\frac{(b-x)^{n+1}}{(n+1)!}$ s'annule en $x = a$. Comme on a $h(a) = h(b) = 0$, le

lemme de Rolle garantit l'existence de $c \in]a,b[$ tel que $0 = h'(c) = \frac{(b-c)^n}{n!}\left(f^{(n+1)}(c) - \lambda\right) = 0$. En reportant $\lambda = f^{(n+1)}(c)$ dans l'équation $h(a) = 0$ on obtient la formule de Taylor.

La deuxième affirmation est claire à partir de la formule. □

Par exemple cette formule appliquée à $b = 1$, $a = 0$ et $f(x) := e^x$ donne :

$$e = e^1 = 1 + \frac{1}{1!} + \frac{1}{2!} + \ldots + \frac{1}{n!} + \frac{e^c}{(n+1)!}$$

et donc en particulier $\lim_{n\to\infty}\left(1 + \frac{1}{1!} + \frac{1}{2!} + \ldots + \frac{1}{n!}\right) = e$.

14.3 EXTREMA D'UNE FONCTION

Calculer le tableau de variation d'une fonction et ses minima et maxima est une des utilisations les plus fréquentes du calcul différentiel ; on en donne ici quelques exemples.

Nous avons vu qu'une condition nécessaire pour qu'une fonction dérivable ait un maximum sur un intervalle ouvert est que sa dérivée s'annule en un point. Ce n'est clairement pas suffisant comme le montre l'exemple de $f(x) = x^3$ dont la dérivée s'annule en zéro mais qui est croissante. La formule de Taylor permet d'établir des conditions suffisantes lorsque la fonction est suffisamment de fois dérivable.

THÉORÈME: *Soit une fonction $f :]a,b[\to \mathbf{R}$ qui est $n+1$ fois continûment dérivable et soit $x_0 \in]a,b[$; supposons que $f'(x_0) = \ldots = f^{(n-1)}(x_0) = 0$ et $f^{(n)}(x_0) \neq 0$ alors*

(i) Si n est impair, f ne possède ni maximum ni minimum en c.

(ii) si n est pair alors il existe un intervalle $]x_0 - \varepsilon, x_0 + \varepsilon[\subset]a,b[$ sur lequel f admet $f(x_0)$ comme maximum si $f^{(n)}(x_0) < 0$ (resp. minimum si $f^{(n)}(x_0) > 0$).

Démonstration: C'est une application directe de la formule de Taylor : d'après les hypothèses on peut écrire $f(x) - f(x_0) = (x - x_0)^n\left(f^{(n)}(x_0) + (x - x_0)f^{(n+1)}(c)\right)$ (où c dépend de x). Comme $f^{(n+1)}(x)$ est continue, elle est bornée sur un intervalle autour de x_0 et donc si $|x - x_0|$ est assez petit $f^{(n)}(x_0) + (x - x_0)f^{(n+1)}(c)$ est du même signe que $f^{(n)}(x_0)$ et bien sûr $(x - x_0)^n$ est positif si n est pair et du signe de $(x - x_0)$ si n est impair ; ainsi le signe de $f(x) - f(x_0)$, quand x est dans un petit intervalle autour de x_0, est comme annoncé dans le théorème. □

Remarquons que la conclusion du théorème n'est pas que le maximum (ou minimum) sur tout l'intervalle $]a,b[$ de la fonction f est atteint mais seulement que f atteint en x_0 son maximum (ou minimum) sur un (petit) intervalle autour de x_0. On dit que f admet un maximum (ou minimum) *local*.

On peut illustrer ce théorème par le diagramme suivant :

CAS n impair :

$f^{(n)}(x_0) > 0$ $\qquad\qquad$ $f^{(n)}(x_0) > 0$

CAS n pair et $f^{(n)}(x_0) < 0$:

CAS n pair et $f^{(n)}(x_0) > 0$

COROLLAIRE: *Soit f une fonction continue de $[a,b]$ vers $\mathbf{R}$, dérivable sur $]a,b[$, alors f atteint son maximum (resp. son minimum) soit en a ou b, soit en un point $c \in]a,b[$ où $f'(c) = 0$*

Démonstration: On sait que f atteint ses extrema, si ce n'est pas en a ou b, d'après un théorème du paragraphe précédent, cela doit être en un point où f' s'annule. □

Terminons par l'étude de la position d'une courbe $y = f(x)$ par rapport à sa tangente.

Définition: Une courbe $y = f(x)$ possède un *point d'inflexion* en x_0 si la courbe est situé de part et d'autre de sa tangente au voisinage de x_0.

Exemple (en $x_0 = 0$) :

Point d'inflexion $f(x) = x^3 + x$ Pas d'inflexion $f(x) = x^4 + x$

Avec les mêmes méthodes ou en appliquant le théorème précédent à $g(x) := f(x) - f(x_0) - f'(x_0)(x - x_0)$ on démontre facilement la caractérisation suivante :

APPLICATION: (i) Supposons que f soit deux fois dérivable, alors si f a un point d'inflexion en x_0, on a $f''(x_0) = 0$.

(ii) Supposons que f soit n fois dérivable, $f''(x_0) = \ldots = f^{(n-1)}(x_0) = 0$ et $f^{(n)}(x_0) \neq 0$ alors x_0 est un point d'inflexion si et seulement si n est impair.

Leibniz Wilhelm Gotfried (1646–1716)

CHAPITRE 15 INTÉGRATION

L'objet de ce chapitre est le calcul des aires. Plus précisément on voit assez facilement que, en découpant les aires, on se ramène à savoir calculer les aires délimitées par deux axes verticaux, un axe horizontal et le graphe d'une fonction :

Il existe deux cas où le résultat est très simple à trouver : celui où la fonction est constante (un rectangle), celui où la fonction est linéaire (un trapèze) :

$$f(x) = c \qquad\qquad f(x) = ux + v$$
$$A = (b-a)c \qquad\qquad A = (b-a)\left(\tfrac{(ua+v)+(ub+v)}{2}\right) = \tfrac{1}{2}ub^2 + bv - \tfrac{1}{2}ua^2 - av$$

On peut dès à présent observer que l'aire, vue comme fonction de b est une fonction dérivable dont la dérivée dans le premier cas est $A' = c = f(b)$ et dans le second cas $A' = ub + v = f(b)$. Ce phénomène est général et constitue le théorème fondamental du calcul intégral, découvert par Newton. Avant de démontrer ce théorème il faut construire la théorie de l'intégration – du calcul d'aires– disons pour les fonctions continues. L'idée est assez simple : on approche toute fonction par des fonctions "en escalier", c'est-à-dire qu'on approche une surface par une somme de rectangle ; on prend un découpage de plus en plus fin et le résultat à la limite est l'aire cherchée. Ce programme comporte des difficultés techniques mais est essentiellement correct. En particulier Archimède est le premier à l'avoir utilisé et ses calculs d'aires délimitées par des arcs de cercles, des arcs de paraboles doivent être vus comme les premiers éléments du calcul intégral.

15.1 INTÉGRATION DES FONCTIONS CONTINUES.

Commençons par définir précisément les fonctions en escalier :

Définition: Une fonction $f : [a,b] \to \mathbf{R}$ est une fonction *en escalier* si il existe une subdivision de l'intervalle $[a,b]$ par $a = a_0 < a_1 < a_2 < \ldots < a_n = b$ telle que f soit constante sur chaque intervalle $]a_i, a_{i+1}[$.

Exemple : la fonction $E(x)$ "partie entière" est une fonction en escalier sur tout intervalle de $\mathbf{R}$.

Remarque : L'avantage des fonctions en escalier est qu'il est facile de calculer l'aire qu'elles délimitent. En effet si $f(x) = c_i$ pour $x \in]a_i, a_{i+1}[$ alors l'aire est donnée par le nombre

$$S(f) := \sum_{i=0}^{n-1} c_i(a_{i+1} - a_i)$$

à condition de compter positivement l'aire située au dessus de l'axe des x et négativement celle qui est située au dessous.

Comme nous voulons approcher les fonctions continues par des fonctions en escalier, il est commode d'introduire une classe de fonctions qui contient les deux types de fonctions :

Définition: Une fonction $f : [a, b] \to \mathbf{R}$ est une fonction *continue par morceaux* si il existe une subdivision de l'intervalle $[a, b]$ par $a = a_0 < a_1 < a_2 < \ldots < a_n = b$ telle que f soit continue chaque intervalle $]a_i, a_{i+1}[$ et admette une limite à droite et à gauche en chaque a_i.

Exemple : une fonction continue ou une fonction en escalier est continue par morceaux. Une fonction continue par morceaux a l'allure suivante :

Nous allons maintenant définir l'intégrale d'une fonction continue par morceaux, c'est-à-dire l'aire qu'elle délimite (en comptant positivement l'aire située au dessus de l'axe des x et négativement celle qui est située au dessous).

Soit donc f une fonction continue par morceaux sur $[a, b]$.

Comme nous allons considérer des subdivisions "de plus en plus fines" de l'intervalle $[a, b]$, il faut introduire une notion pour mesurer cela :

Définition: Soit $a = a_0 < a_1 < a_2 < \ldots < a_n = b$ une subdivision de l'intervalle $[a, b]$, le *pas* de la subdivision est la quantité $\max(a_{i+1} - a_i)$. On notera S_δ l'ensemble des subdivisions de pas inférieur à δ.

Soit σ une subdivision $a = a_0 < a_1 < a_2 < \ldots < a_n = b$ de l'intervalle $[a, b]$, on notera $I(f, \sigma)$ la somme (souvent appelée *somme de Riemann*) :

$$I(f, \sigma) := \sum_{i=0}^{n-1} f(a_i)(a_{i+1} - a_i)$$

Il est clair intuitivement que cette somme approche la valeur de l'aire délimité par le

graphe de la fonction f et qu'elle l'approche d'autant mieux que le pas de la subdivision est petit.

Il est donc raisonnable de définir :

Définition: Une fonction $f : [a, b] \to \mathbf{R}$ est *intégrable* si la limite de $I(f, \sigma)$ quand le pas de la subdivision σ tend vers 0 existe ; la valeur de cette limite s'appelle l'*intégrale* de f entre a et b et se note $\int_a^b f(t)dt$.

Une justification de cette notation (due à Leibniz) est qu'elle s'avère commode pour effectuer de façon mécanique des calculs (intégration par parties, changement de variables, etc, voir chapitre 17) ; une justification plus heuristique est que dt réprésente un accroissement infinitésimal de la variable et que $\int_a^b f(t)dt$ est la limite de $\sum_{i=1}^n f(a+\frac{i}{n}(b-a))(1/n)$.

THÉORÈME: *(i) Une fonction continue par morceaux est intégrable.*
(ii) Si f est une fonction en escalier et $f(]a_i, a_{i+1}[) = c_i$ alors

$$\int_a^b f(t)dt = \sum_{i=0}^{n-1} c_i(a_{i+1} - a_i)$$

.

Démonstration: Il n'y a pas de difficulté à prouver la seconde partie (sauf peut-être la difficulté des notations ...). Nous ne donnerons hélas pas la preuve complète de l'intégrabilité d'une fonction continue par morceaux. Indiquons néanmoins quelques unes des étapes de la démonstration. Le point clef est de montrer que si f est continue par morceaux sur $[a, b]$, alors on peut l'encadrer par deux suites monotones de fonctions en escalier l'approchant de mieux en mieux. Plus précisément on peut montrer :

LEMME: *Soit f une fonction continue par morceaux sur $[a, b]$. Alors il existe deux suites de fonctions en escalier g_n et h_n telles que sur $[a, b]$ telles que, pour tout $n \geq 1$ et $t \in [a, b]$ on ait*
(i) $g_n(t) \leq g_{n+1}(t) \leq f(t) \leq h_{n+1}(t) \leq h_n(t)$
(i) $0 \leq h_n(t) - g_n(t) \leq 1/n$.

Démonstration: (en partie admis) En découpant $[a, b]$ en un nombre fini d'intervalles et en prolongeant par continuité f sur chaque intervalle ouvert où elle est continue on se ramène au cas où f est continue sur $[a, b]$. On utilise alors le fait (admis ici) qu'elle est *uniformément continue*, c'est à dire qu'il existe $\delta = \delta(n)$ tel que $|x - x'| \leq \delta$ entraîne $|f(x) - f(x')| \leq 1/2n$. On choisit une subdivision de pas $\leq \delta(n)$ et on définit h_n comme le maximum (resp. g_n comme le minimum) de f sur chaque intervalle de la subdivision. Les suites aisni construites ont toutes les propriétés voulues sauf peut-être la monotonie. Il suffit alors de remplacer h_n et g_n par $h'_n = \min_{1 \leq m \leq n} h_m$ et $g'_n = \max_{1 \leq m \leq n} g_m$. □

Si l'on choisit g_n et h_n comme dans le lemme, alors les suites $v_n = \int_a^b h_n(t)dt$ et $u_n = \int_a^b g_n(t)$ sont adjacentes et leur limite définit $\int_a^b f(t)dt$. □

Observons aussi qu'il découle de ces constructions que si par exemple f est continue par morceaux alors

$$\lim \sum_{i=0}^{n-1} f\left(a + \frac{i}{n}(b-a)\right) \frac{(b-a)}{n} = \int_a^b f(t)dt$$

Par exemple si $f(t) = t^2$ et $[a,b] = [0,1]$ on en tire que

$$\int_0^1 t^2 dt = \lim_{n\to\infty} \sum_{i=0}^{n-1} \left(\frac{i}{n}\right)^2 \frac{1}{n} = \lim \frac{n(n+1)(2n+1)}{6n^3} = \frac{1}{3}$$

Ce que l'on vérifiera par la suite bien sûr par une autre méthode.

15.2 PROPRIÉTÉS DE L'INTÉGRALE

Rassemblons en un théorème les principales propriétés de l'intégrale :

THÉORÈME: *Soient f, g des fonctions continues par morceaux, l'intégrale vérifie les propriétés suivantes :*
i) (Linéarité)

$$\int_a^b (\lambda f(t) + \mu g(t))dt = \lambda \int_a^b f(t)dt + \mu \int_a^b g(t)dt,$$

ii) (Positivité)

$$f \geq g \Rightarrow \int_a^b f(t)dt \geq \int_a^b g(t)dt$$

de plus si les fonctions sont continues et on a $f(t_0) > g(t_0)$ en un point t_0 appartenant à l'intevalle $[a,b]$, alors $\int_a^b f(t)dt > \int_a^b g(t)dt$,
iii)

$$\left|\int_a^b f(t)dt\right| \leq \int_a^b |f(t)|dt$$

iv) ("formule de 'Chasles") Si $a < c < b$ alors

$$\int_a^b f(t)dt = \int_a^c f(t)dt + \int_c^b f(t)dt$$

Démonstration: Le principe est d'établir ces formules pour les fonctions en escalier et d'en déduire la formule dans le cas général par passage à la limite. (i) La linéarité de l'intégrale $\int_a^b f(t)dt = \sum_{i=0}^{n-1} c_i(a_{i+1} - a_i)$ est claire pour les fonctions en escaliers. En effet la somme ne dépend pas de la subdivision choisie (c'est clair sur le dessin, pouvez-vous en écrire la preuve formelle?) et en choisissant une subdivision adaptée aux deux fonctions en escalier f et g la linéarité provient de la distributivité de la multiplication par rapport à l'addition. Soit donc f_n et g_n des fonctions en escalier telles que $|f_n - f| \leq 1/n$ et

$|g_n - g| \leq 1/n$, alors $|(\lambda f_n - \mu g_n) - (\lambda f + \mu g)| \leq (|\lambda| + |\mu|)/n$ et donc, en passant à la limite

$$\begin{aligned}\int_a^b (\lambda f + \mu g)(t)dt &= \lim \int_a^b (\lambda f_n + \mu g_n)(t)dt \\ &= \lim \left(\lambda \int_a^b f_n(t)dt + \mu \int_a^b g_n(t)dt \right) \\ &= \lambda \int_a^b f(t)dt + \mu \int_a^b g(t)dt\end{aligned}$$

(ii) Vu la linéarité, il suffit de vérifier que si f est positive alors $\int_a^b f(t)dt \geq 0$. Mais il existe une suite de fonction en escalier $f_n \geq f$ telle que $\lim \int_a^b f_n(t)dt = \int_a^b f(t)dt$; or bien sûr $\int_a^b f_n(t)dt$ est somme de termes positifs donc est positif. La deuxième affirmation provient de ce que, si f est continue et, disons, $f(c) \neq 0$ avec $c \in [a, b]$ alors il existe un intervalle $[d, e] \subset [a, b]$ (avec $d < e$) sur lequel $f(x) \geq f(c)/2$ et donc $\int_a^b f(t)dt \geq (e - d)c/2 > 0$.

(iii) Si f est une fonction en escalier égale à c_i sur $]a_i, a_{i+1}[$ alors

$$\left| \int_a^b f(t)dt \right| = \left| \sum_{i=0}^{n-1} c_i(a_{i+1} - a_i) \right| \leq \sum_{i=0}^{n-1} |c_i|(a_{i+1} - a_i) = \int_a^b |f(t)|dt$$

L'inégalité pour une fonction intégrable s'en déduit par passage à la limite.

(iv) De nouveau il suffit de démontrer la formule pour une fonction en escalier, ce qui est immédiat. □

Remarques : 1) à condition de définir, lorsque $b < a$, l'intégrale par $\int_a^b f(t)dt := -\int_b^a f(t)dt$ la formule (iv) est valable quelque soit a, b, c.

2) Si $m \leq f(t) \leq M$ sur l'intervalle $[a, b]$, on déduit aisément de (ii) et du calcul $\int_a^b dt = (b - a)$ l'inégalité

$$m(b - a) \leq \int_a^b f(t)dt \leq M(b - a).$$

15.3 THÉORÈME FONDAMENTAL DE L'INTÉGRATION

Définition: Une *primitive* d'une fonction f est une fonction F dérivable telle que $F' = f$

Remarque : Sur un intervalle I une primitive (si elle existe) est unique à une constante près. En effet si F_1 et F_2 sont deux primitives, on voit que $F_1 - F_2$ a une dérivée nulle donc est constante.

Il est difficile de surestimer l'importance du théorème suivant :

THÉORÈME: (théorème fondamental de l'intégration) *Soit f une fonction continue sur l'intervalle $[a,b]$ alors une primitive F est donnée par la fonction $x \mapsto \int_a^x f(t)dt$ et si F est une primitive de f on a la formule :*

$$\int_a^b f(t)dt = F(b) - F(a).$$

Démonstration: Il suffit de vérifier que la fonction $G(x) := \int_a^x f(t)dt$ est dérivable et vérifie $G'(x) = f(x)$ car alors on aura bien prouvé l'existence d'une primitive et comme $G(a) = 0$ on a bien $\int_a^b f(t)dt = G(b) - G(a)$ et si F est une autre primitive de f on a $F(x) = G(x) + C$ donc $F(b) - F(a) = G(b) - G(a)$. Soit $\varepsilon > 0$ et soit $a \leq x < x + h \leq b$, les propriétés de l'intégrale permettent d'écrire :

$$\frac{G(x+h) - G(x)}{h} - f(x) = \frac{1}{h}\left(\int_a^{x+h} f(t)dt - \int_a^x f(t)dt\right) - f(x) = \frac{1}{h}\int_x^{x+h} (f(t) - f(x))dt$$

mais f est continue en x donc il existe δ tel que si $x \leq t \leq x+h \leq x+\delta$ alors $|f(x)-f(t)| \leq \varepsilon$ donc on obtient (si $|h| \leq \delta$) :

$$\left|\frac{G(x+h) - G(x)}{h} - f(x)\right| \leq \varepsilon\left(\frac{1}{h}\int_x^{x+h} dt\right) = \varepsilon$$

Ce qui exprime bien que G est dérivable et que $G'(x) = f(x)$ $\square$

L'importance de ce théorème vient du fait qu'il permet de calculer la plupart des aires (ou intégrales) de manière assez simple. Nous étudierons plus systématiquement le calcul des primitives au chapitre 17 et donnons l'exemple très simple suivant :

$$\int_a^b t^m dt = \frac{b^{m+1}}{m+1} - \frac{a^{m+1}}{m+1}$$

En particulier $\int_0^1 t^2 dt = 1/3$ comme nous l'avons déjà vérifié.

15.4 CALCULS APPROCHÉS D'INTÉGRALES

Nous développerons au chapitre 17 des techniques de calcul de primitives et donc d'intégrales mais ces techniques ne s'appliquent pas à toutes les fonctions et il est souhaitable de pouvoir évaluer ces intégrales. Par exemple la fonction logarithme est définie comme $\log(x) := \int_1^x dt/t$ *donc si l'on veut établir une table de logarithmes, il faut savoir calculer de manière approchée une intégrale. De nos jours, bien sûr, ces calculs sont programmés sur ordinateurs mais, demandez à vos parents (ou grands-parents) de vous montrer une vieille table de logarithmes Bouvard et Ratinet (dont le nom semble avoir inspiré Flaubert pour son oeuvre inachevée "Bouvard et Pécuchet").*

Commençons par le cas légèrement plus facile des fonctions monotones. Si $f(x)$ est croissante et $a < b$ on observe que $f(a)(b-a) \leq \int_a^b f(t)dt \leq f(b)(b-a)$ donc plus généralement si $a = a_0 < \ldots < a_n = b$ est une subdivision de l'intervalle $[a, b]$ on obtient :

$$\sum_{i=0}^{n-1} f(a_i)(a_{i+1} - a_i) \leq \sum_{i=0}^{n-1} \int_{a_i}^{a_{i+1}} f(t)dt = \int_a^b f(t)dt \leq \sum_{i=0}^{n-1} f(a_{i+1})(a_{i+1} - a_i)$$

En particulier si l'on choisit $a_i := a+(i/n)(b-a)$ et si l'on pose $S_n(f) := \left(\frac{b-a}{n}\right) \sum_{i=0}^{n-1} f(a_i)$ on en tire $S_n(f) \leq \int_a^b f(t)dt \leq S_n(f) + \frac{(b-a)}{n}(f(b) - f(a))$ d'où un calcul approché de l'intégrale. Ce type de calcul est illustré par la figure suivante :

On peut aussi utiliser le même calcul pour estimer des sommes. Par exemple en prenant la fonction décroissante $f(x) = 1/x^a$ (avec $a > 0$) on obtient

$$\sum_{n=2}^{N} 1/n^a \leq \int_1^N dt/t^a \leq \sum_{n=1}^{N-1} 1/n^a$$

En utilisant le théorème fondamental de l'intégration et, pour $a = 1$, la définition du logarithme, on sait que $\int_1^N dt/t^a = \left(\frac{1}{1-a}\right)\left(\frac{1}{N^{a-1}} - 1\right)$ si $a \neq 1$ et $\int_1^N dt/t^a = \log N$ si $a = 1$. Ainsi, si $a \neq 1$ on a :

$$\left(\frac{1}{1-a}\right)\left(\frac{1}{N^{a-1}} - 1\right) \leq \sum_{n=1}^{N} \frac{1}{n^a} \leq 1 + \left(\frac{1}{1-a}\right)\left(\frac{1}{N^{a-1}} - 1\right)$$

et si $a = 1$ on obtient :

$$\log N \leq \sum_{n=1}^{N} \frac{1}{n} \leq 1 + \log N$$

On peut conclure par exemple que $\lim_{N\to\infty} \sum_{n=1}^{N} \frac{1}{n} = +\infty$.

Remarque : il n'est peut-être pas inutile de remarquer, pour ceux qui aiment expérimenter sur machine, que, si l'on programme le calcul de la suite $u_N = \sum_{n=1}^{N} \frac{1}{n}$, celle-ci semble converger!

Donnons maintenant une manière un peu plus générale et plus fine d'approcher la valeur d'une intégrale. L'idée est d'approcher la fonction par une fonction linéaire par morceau comme dans la figure ci-dessous :

THÉORÈME: *(méthode des trapèzes) Soit* $f : [a,b] \to \mathbf{R}$ *une fonction deux fois continûment dérivable et soit* M *un majorant de* f'' *sur l'intervalle et* $a_i := a + \frac{i(b-a)}{n}$ *alors*

$$\left| \int_a^b f(t)dt - \frac{(b-a)}{n} \left(\frac{1}{2}f(a) + \frac{1}{2}f(b) + f(a_1) + \ldots + f(a_{n-1}) \right) \right| \leq \frac{M}{12n^2}(b-a)^3$$

Démonstration: On procède en intégrant par parties puis en choisissant judicieusement les constantes d'intégration.

$$\int_{a_i}^{a_{i+1}} f(t)dt = [(t-\lambda)f(t)]_{a_i}^{a_{i+1}} - \int_{a_i}^{a_{i+1}} (t-\lambda)f'(t)dt$$

$$= [(t-\lambda)f(t)]_{a_i}^{a_{i+1}} - \left[\left(\frac{t^2}{2} - \lambda t - \mu \right) f'(t) \right]_{a_i}^{a_{i+1}} + \int_{a_i}^{a_{i+1}} \left(\frac{t^2}{2} - \lambda t - \mu \right) f''(t)dt$$

Pour éviter d'avoir à calculer les dérivées $f'(a_i)$, on s'arrange pour que le deuxième crochet s'annule, c'est-à-dire que l'on choisit $t^2 - 2\lambda t - 2\mu = (t-a_i)(t-a_{i+1})$ donc $\lambda = (a_i + a_{i+1})/2$ et on obtient :

$$\int_{a_i}^{a_{i+1}} f(t)dt = \frac{f(a_{i+1}) + f(a_i)}{2n} + \frac{1}{2}\int_{a_i}^{a_{i+1}} (t-a_i)(t-a_{i+1})f''(t)dt$$

or cette dernière intégrale est bornée en module par $\frac{M}{2}\int_{a_i}^{a_{i+1}}(t-a_i)(a_{i+1}-t)dt = \frac{M(b-a)^3}{12n^3}$. En sommant les inégalités on obtient l'énoncé. □

Commentaires : la majoration énoncé est intéressante dès les premiers pas ; elle s'écrit pour $n = 1$:

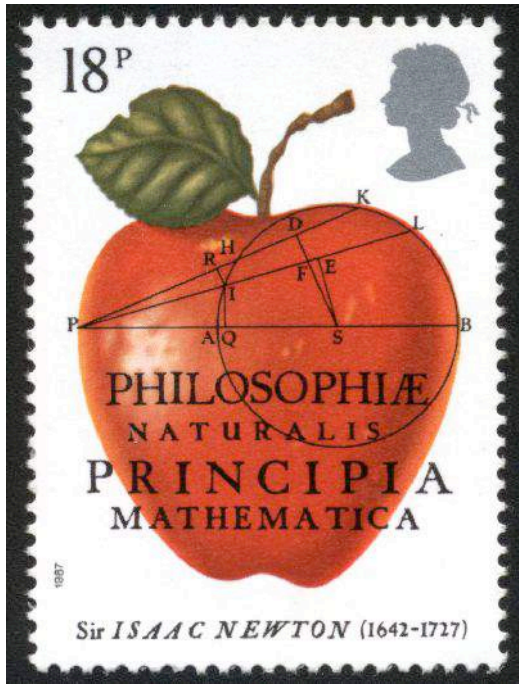
$$\left| \int_{a_i}^{a_{i+1}} f(t)dt - \frac{(b-a)}{2} \left(f(a) + f(b) \right) \right| \leq \frac{M}{12}(b-a)^3$$

Pour $n = 2$ on obtient :

$$\left| \int_{a_i}^{a_{i+1}} f(t)dt - \frac{(b-a)}{2} \left(\frac{1}{2}f(a) + \frac{1}{2}f(b) + f\left(\frac{a+b}{2} \right) \right) \right| \leq \frac{M}{48}(b-a)^3$$

La majoration est meilleure si M est petit (c'est-à-dire si la fonction n'oscille pas trop) et si n est grand (c'est-à-dire si la subdivision est fine) et moins bonne si $(b-a)$ est grand (c'est-à-dire si l'intervalle d'intégration est grand).

Exemple : (calcul approché de $\log 2$) Prenons $a=1$, $b=2$, $n=4$ et $f(t)=1/t$. On a $\max_{t\in[1,2]} |f''(t)| = 2$ et $S_4 = \frac{1}{4}\left(\frac{1}{2}+\frac{1}{4}+\frac{4}{5}+\frac{2}{3}+\frac{4}{7}\right) = 0,69702\ldots$ et donc la valeur approchée : $|\log 2 - S_4| \leq 0,010166\ldots$.



Newton Isaac (1642–1727)

CHAPITRE 16 QUELQUES FONCTIONS USUELLES

Ce chapitre est destiné à fournir une bibliothèque de fonctions utiles : les fonctions polynômes (déjà étudiées au chapitre 6), la fonction logarithme, l'exponentielle (et plus généralement les fonctions a^x ou x^a), les fonctions circulaires $\sin(x)$, $\cos(x)$ *et* $\mathrm{tg}(x)$ *ainsi que leurs "réciproques"* $\mathrm{Arcsin}(x)$, $\mathrm{Arcos}(x)$ *et* $\mathrm{Arctg}(x)$. *Les fonctions hyperboliques* $\mathrm{sh}(x)$, $\mathrm{ch}(x)$ *et* $\mathrm{th}(x)$ *et leurs "réciproques"* $\mathrm{Argsh}(x)$, $\mathrm{Argch}(x)$ *et* $\mathrm{Argth}(x)$ – *ces dernières étant moins indispensables puisque l'on peut les exprimer à partir du logarithme et de l'exponentielle.*

16.1 FONCTIONS LOGARITHME ET EXPONENTIELLE

Définition: On appelle *logarithme* la fonction définie sur $\mathbf{R}_+^*$ à valeurs réelles :

$$\log(x) := \int_1^x \frac{dt}{t}$$

On voit immédiatement que $\log(1) = 0$ et que $\log(x) > 0$ pour $x > 1$ (resp. $\log(x) < 0$ pour $x \in]0, 1[$).

THÉORÈME: *La fonction* $\log : \mathbf{R}_+^* \to \mathbf{R}$ *est croissante, continue, indéfiniment dérivable ; en fait*

$$\log'(x) = \frac{1}{x}, \quad \lim_{\substack{x \to 0 \\ x > 0}} \log x = -\infty \quad \text{et} \quad \lim_{x \to +\infty} \log x = +\infty$$

De plus la fonction log *est une bijection et un homomorphisme de groupe (elle transforme multiplication en addition)*

$$\log(xy) = \log(x) + \log(y)$$

Démonstration: D'après le théorème fondamental de l'intégration, la fonction log a pour dérivée $1/x$ et est donc croissante. Ensuite $\frac{d}{dx}\log(xy) = \frac{y}{xy} = \frac{1}{x}$ donc $\log(xy) - \log x$ est une fonction constante en x. En prenant $x = 1$, on voit que cette constante vaut $\log y$, d'où la formule $\log(xy) = \log(x) + \log(y)$. On en déduit que $\log(2^n) = n\log(2)$ tend vers l'infini quand n tend vers l'infini et comme log est croissante cela entraîne que $\log(x)$ tend vers l'infini quand x tend vers l'infini. Enfin comme $\log(x) = -\log(1/x)$ on en tire que lorsque x tend vers zéro (par valeurs positives) $\log(x)$ tend vers $-\infty$. □

Remarque : Nous ne considérerons que cette fonction logarithme appelée parfois logarithme néperien. Historiquement, on utilisait pour les calculs "à la main" le logarithme décimal c'est-à-dire $\log_{10} x = \log x / \log 10$. La généralisation des calculatrices et ordinateurs a rendu désuètes les tables de logarithmes et enlevé son intérêt au logarithme décimal.

Définition: La bijection réciproque de la fonction logarithme s'appelle la fonction *exponentielle* et se note $\exp : \mathbf{R} \to \mathbf{R}_+^*$.

THÉORÈME: La fonction exponentielle est indéfiniment dérivable et croissante ; en fait $\frac{d}{dx}\exp(x) = \exp(x)$. De plus

$$\lim_{x\to-\infty} \exp(x) = 0 \quad \text{et} \quad \lim_{x\to+\infty} \exp(x) = +\infty$$

et on a

$$\exp(x+y) = \exp(x)\exp(y)$$

Démonstration: On a déjà vu que $\exp(x)$ est égale à sa dérivée ; ensuite les limites découlent de celles calculées pour les logarithmes au théorème précédent ainsi que la dernière formule. □

Il est indispensable de mémoriser l'allure des graphes des fonctions usuelles

Remarque : Soit $n \in \mathbf{N}$ (ou même $n \in \mathbf{Z}$) alors $x^n = \exp(n\log(x))$; en effet d'après le théorème précédent on a $\exp(\log(x) + \ldots + \log(x)) = (\exp(\log(x)))^n = x^n$. Ceci suggère la possibilité de définir a^b ainsi :

Définition: (Puissances généralisées) Soit $a \in \mathbf{R}_+^*$ et $b \in \mathbf{R}$ on pose

$$a^b := \exp(b\log a)$$

En particulier, on retrouve pour $n \in \mathbf{N}^*$ les fonctions "racine n-ième" : $a^{1/n} = \exp(\frac{1}{n}\log a)$.

THÉORÈME: *On a les formules suivantes, pour $a, a' > 0$ et $b, c \in \mathbf{R}$:*
$a^{b+b'} = a^b a^{b'}$
$(aa')^b = a^b a'^b$
$\left(a^b\right)^c = a^{bc}$
$a^0 = 1 = 1^b$
De plus $\frac{d}{dx}\left(x^b\right) = bx^{b-1}$ *et* $\frac{d}{dx}\left(a^x\right) = \log(a)a^x$.

Démonstration: Ces formules découlent immédiatement de la définition et des propriétés de l'exponentielle et du logarithme : par exemple $a^{b+b'} = \exp((b+b')\log a) = \exp(b\log a)\exp(b'\log a)$ et $\frac{d}{dx}\left(x^b\right) = \frac{d}{dx}\left(\exp(b\log x)\right) = b\frac{d}{dx}(\log x)\exp(b\log x) = bx^{b-1}$. □

L'allure des graphes des fonctions $f(x) = x^a$ est la suivante :

Définition: On appelle *base du logarithme* le nombre e tel que $\log(e) = 1$ ou $1 = \int_1^e \frac{dt}{t}$ ou encore $e = \exp(1)$.

Remarque : on a vu que $e = \lim_{n\to\infty} \left(1 + \frac{1}{1!} + \ldots + \frac{1}{n!})\right) = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n$.

Notation : On a donc $\exp(x) = e^x$ et on utilisera désormais cette notation.

THÉORÈME:

i) $\lim_{\substack{x \to 0 \\ x > 0}} x^a (\log x)^n = 0$ *si* $a > 0$ *et* $n \in \mathbf{Z}$

$\lim_{x\to+\infty} x^a (\log x)^n = 0$ *si* $a < 0$

C'est-à-dire "les puissances l'emportent sur le logarithme"

ii) $\lim_{x\to+\infty} x^a e^x = +\infty$ *et*

$\lim_{x\to-\infty} x^a e^x = 0$

C'est-à-dire "l'exponentielle l'emporte sur les puissances"

Démonstration: i) Remarquons que pour x supérieur à un on a :

$$\log(x) = \int_1^x dt/t \leq \int_1^x dt/\sqrt{t} = (1/2)(\sqrt{x} - 1),$$

ainsi donc $\log(x)/x$ tend vers zéro quand x tend vers l'infini. De façon plus générale $(\log x)^n/x^a = ((n/a)^n \left(\log x^{a/n}/x^{a/n}\right)^n$ tend donc aussi vers zéro quand x tend vers l'infini. Les autres limites s'obtiennent par des calculs similaires. □

16.2 FONCTIONS CIRCULAIRES ET RÉCIPROQUES

Commençons par quelques rappels sur les fonctions $\cos(x)$ et $\sin(x)$. Tout d'abord prouvons l'énoncé admis aux chapitres précédents.

PROPOSITION:

$$\lim_{\substack{x \to 0 \\ x \neq 0}} \frac{\sin(x)}{x} = 1$$

Démonstration: Considérons un cercle de rayon r (avec des angles mesurés en radian) :

Nous allons utiliser deux faits géométriques assez simples : l'aire d'un secteur angulaire (d'angle x) est donné par $xr^2/2$ et la longueur de l'arc de cercle est donnée par xr (en particulier la surface du cercle est πr^2 et sa circonférence $2\pi r$), pour démontrer l'inégalité suivante :

LEMME: *Soit* $0 < x \leq \pi/2$ *alors* $\sin(x) < x < \mathrm{tg}(x)$.

Démonstration: Appelons x l'angle sur la figure ci-dessus en prenant $r = 1$; alors $\sin(x) = OA = CB \leq CD \leq arc(CD) = x$. Par ailleurs l'aire du secteur d'angle x est égale à $x/2$ et est majorée par l'aire du triangle OED qui vaut $ED/2 = \mathrm{tg}(x)/2$ d'où le lemme. □

COROLLAIRE: *Pour* $0 < |x| \leq \pi/2$ *on a* $\cos(x) < \frac{\sin(x)}{x} < 1$ *; en particulier* $\lim_{x\to 0} \dfrac{\sin(x)}{x} = 1$.

Démonstration: En effet on obtient pour $0 < x < \pi/2$ l'inégalité $\cos(x) < \sin(x)/x < 1$ qui est donc valable pour $0 < |x| < \pi/2$ puisque $\cos(x)$ et $\sin(x)/x$ sont paires ; enfin comme $\cos(x)$ est continue on a $\lim_{x\to 0} \cos(x) = \cos(0) = 1$. □

Traçons le graphe de la fonction sinus, celui de cosinus s'en déduit par translation puisque $\cos(t) = \sin(\frac{\pi}{2} + t)$:

FORMULAIRE :
$\sin(-t) = -\sin(t)$, $\cos(-t) = \cos(t)$ et $\mathrm{tg}(-t) = -\mathrm{tg}(t)$
$\sin(t+\pi) = -\sin(t)$, $\cos(t+\pi) = -\cos(t)$ et $\mathrm{tg}(t+\pi) = \mathrm{tg}(t)$
$\cos^2(t) + \sin^2(t) = 1$
$\sin(x+t) = \cos(t)\sin(x) + \cos(x)\sin(t)$ et $\cos(x+t) = \cos(x)\cos(t) - \sin(x)\sin(t)$
$\mathrm{tg}(x+t) = (\mathrm{tg}(t) + \mathrm{tg}(x))/(1 - \mathrm{tg}(x)\,\mathrm{tg}(t))$
$\sin(t) = 2\,\mathrm{tg}(t/2)/(1 + \mathrm{tg}^2(t/2))$ et $\cos(t) = (1 - \mathrm{tg}^2(t/2))/(1 + \mathrm{tg}^2(t/2))$

On peut aussi signaler que $\cos(t) = \cos(x)$ équivaut à $t = \pm x + 2k\pi$ (avec $k \in \mathbf{Z}$) alors que $\sin(t) = \sin(x)$ équivaut à $t = x + 2k\pi$ ou $t = \pi - x + 2k\pi$ (avec $k \in \mathbf{Z}$). On peut aussi observer que $a\cos(x) + b\sin(x) = \sqrt{a^2+b^2}\sin(x+\phi)$ où ϕ est tel que $\sin(\phi) = a/\sqrt{a^2+b^2}$ et $\cos(\phi) = b/\sqrt{a^2+b^2}$.

La fonction sin : $[-\frac{\pi}{2}, +\frac{\pi}{2}] \to [-1,1]$ est une bijection (strictement) croissante ; en effet il suffit d'utiliser le théorème du chapitre 13.

Définition: ("Arc sinus") La bijection réciproque de la fonction sin : $[-\frac{\pi}{2}, +\frac{\pi}{2}] \to [-1,+1]$ se note Arcsin : $[-1,+1] \to [-\frac{\pi}{2}, +\frac{\pi}{2}]$.
On peut en tracer le graphe par symétrie à partir de celui de $\sin(x)$:

THÉORÈME: *La fonction* Arcsin : $[-1,+1] \to [-\frac{\pi}{2}, +\frac{\pi}{2}]$ *est continue, croissante, impaire et vérifie :*
i) $y = \mathrm{Arcsin}(x) \Leftrightarrow \{x = \sin(y)$ et $y \in [-\frac{\pi}{2}, +\frac{\pi}{2}]\}$
ii) $\mathrm{Arcsin}(x)$ *est indéfiniment dérivable sur* $]-1,+1[$ *et*

$$\mathrm{Arcsin}'(x) = \frac{1}{\sqrt{1-x^2}}.$$

Démonstration: i) Découle de la définition de Arcsin comme bijection réciproque.
ii) Provient du théorème donnant la dérivée d'une fonction réciproque : si $x = \sin(y)$ et $y \in]-\frac{\pi}{2}, +\frac{\pi}{2}[$ alors $dx/dy = \cos(y) > 0$ et donc vaut $\sqrt{1-\sin^2(y)} = \sqrt{1-x^2}$ d'où $dy/dx = 1/\sqrt{1-x^2}$. □

On peut faire une construction analogue en observant que cos : $[0,\pi] \to [-1,+1]$ est une bijection décroissante.

Définition: ("Arc cosinus") La bijection réciproque de la fonction cos : $[0, \pi] \to [-1, +1]$ se note Arcos : $[-1, +1] \to [0, \pi]$.

On peut en tracer le graphe par symétrie à partir de celui de $\cos(x)$:

Toutes les propriétés de la fonction $\mathrm{Arcos}(x)$ se déduisent de celle de $\mathrm{Arcsin}(x)$ et de la formule suivante (visible sur le graphe) :

PROPOSITION: *Si* $x \in [-1, +1]$ *on a* $\mathrm{Arcsin}(x) + \mathrm{Arcos}(x) = \pi/2$

Démonstration: On calcule (comme pour Arcsinus) que $\mathrm{Arcos}'(x) = -1/\sqrt{1-x^2}$ donc la dérivée de la fonction $\mathrm{Arcsin}(x) + \mathrm{Arcos}(x)$ est nulle, donc la fonction est constante et en calculant la valeur en $x = 0$ on conclut. □

La fonction tg : $]-\frac{\pi}{2}, +\frac{\pi}{2}[\to \mathbf{R}$ est une bijection (strictement) croissante ; en effet il suffit d'utiliser le théorème du chapitre 13.

Définition: ("Arc tangente") La bijection réciproque de la fonction tg : $]-\frac{\pi}{2}, +\frac{\pi}{2}[\to \mathbf{R}$ se note Arctg : $\mathbf{R} \to]-\frac{\pi}{2}, +\frac{\pi}{2}[$.

On peut en tracer le graphe par symétrie à partir de celui de $\mathrm{tg}(x)$:

THÉORÈME: *La fonction* $\mathrm{Arctg} : \mathbf{R} \to]-\frac{\pi}{2}, +\frac{\pi}{2}[$ *est continue, croissante, impaire et vérifie :*

i) $y = \mathrm{Arctg}(x) \Leftrightarrow x = \mathrm{tg}(y)$ et $y \in]-\frac{\pi}{2}, +\frac{\pi}{2}[$

ii) $\mathrm{Arctg}(x)$ *est indéfiniment dérivable sur* $\mathbf{R}$ *et*

$$\mathrm{Arctg}'(x) = \frac{1}{x^2+1}$$

iii) $\lim_{x\to+\infty} \mathrm{Arctg}(x) = \frac{\pi}{2}$ et $\lim_{x\to-\infty} \mathrm{Arctg}(x) = -\frac{\pi}{2}$

Démonstration: i) et iii) découlent de la défintion de Arctg comme bijection réciproque.

ii) Le calcul est similaire à ceux déjà effectués : si $x = \mathrm{tg}(y)$ alors $dx/dy = (1 + \mathrm{tg}^2(y)) = 1 + x^2$ donc $dy/dx = 1/(1 + x^2)$. □

Remarque : La fonction Arctg vérifie une intéressante équation fonctionnelle :

PROPOSITION: *Soit* $x \in \mathbf{R}^*$ *alors* $\mathrm{Arctg}(x) + \mathrm{Arctg}(1/x) = sgn(x)\frac{\pi}{2}$ *(où le signe* $sgn(x)$ *vaut* $+1$ *si* $x > 0$ *et* -1 *si* $x < 0$*).*

Démonstration: Le calcul de la dérivée du membre de gauche donne : $(\mathrm{Arctg}(x) + \mathrm{Arctg}(1/x))' = \frac{1}{x^2+1} + \frac{-1}{x^2}\frac{1}{1+1/x^2} = 0$ donc cette fonction est constante sur $\mathbf{R}_+^*$ et $\mathbf{R}_-^*$ (notez qu'elle n'est pas constante sur $\mathbf{R}^*$ tout entier!). Par ailleurs $\mathrm{Arctg}(1) = \pi/4$ et $\mathrm{Arctg}(-1) = -\pi/4$ permettent de calculer ces constantes. □

16.3 FONCTIONS HYPERBOLIQUES ET RÉCIPROQUES

On peut définir des fonctions sinus, cosinus et tangente hyperboliques et travailler avec elles de manière tout-à-fait analogue. Nous serons brefs car ces fonctions sont moins importantes que les "vraies" fonctions sinus, cosinus et tangente.

Définition: La fonction *sinus hyperbolique* est définie par

$$\mathrm{sh}(x) := \frac{e^x - e^{-x}}{2}$$

La fonction *cosinus hyperbolique* est définie par

$$\mathrm{ch}(x) := \frac{e^x + e^{-x}}{2}$$

La fonction *tangente hyperbolique* est définie par

$$\mathrm{th}(x) := \frac{\mathrm{sh}(x)}{\mathrm{ch}(x)} = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

Le nom de ces fonctions vient du fait que $t \mapsto (\cos(t), \sin(t))$ paramétrise le cercle d'équation $x^2+y^2 = 1$ alors que $t \mapsto (\mathrm{ch}(t), \mathrm{sh}(t))$ paramétrise (une branche de) l'hyperbole d'équation $x^2 - y^2 = 1$.

Une autre explication (que nous ne formaliserons pas) aux noms et analogies entre fonctions circulaires et hyperboliques est la suivante : on se souvient que $\cos(x) = (e^{ix} + e^{-ix})/2$ *et* $\sin(x) = (e^{ix} - e^{-ix})/2i$ *et on en tire (formellement) que* $\mathrm{ch}(x) = \cos(ix)$ *et* $\mathrm{sh}(x) = i\sin(ix)$

Donnons les représentations graphiques des trois principales fonctions hyperboliques, les fonctions $\mathrm{sh}(x)$, $\mathrm{ch}(x)$ et $\mathrm{th}(x)$:

Passons à la construction des fonctions réciproques
La fonction $\mathrm{sh} : \mathbf{R} \to \mathbf{R}$ est une bijection croissante, donc :

Définition: (Argument sinus hyperbolique) La bijection réciproque de la fonction $\mathrm{sh} : \mathbf{R} \to \mathbf{R}$ se note $\mathrm{Argsh} : \mathbf{R} \to \mathbf{R}$.

La fonction $\mathrm{ch} : \mathbf{R}_+ \to [1, +\infty)$ est une bijection croissante, donc :

Définition: (Argument cosinus hyperbolique) La bijection réciproque de la fonction $\mathrm{ch} : \mathbf{R}_+ \to [1, +\infty)$ se note $\mathrm{Argch} : [1, +\infty) \to \mathbf{R}_+$.

La fonction $\mathrm{th} : \mathbf{R} \to]-1, +1[$ est une bijection croissante, donc :

Définition: (Argument tangente hyperbolique) La bijection réciproque de la fonction $\mathrm{th} : \mathbf{R} \to]-1, +1[$ se note $\mathrm{Argth} :]-1, +1[\to \mathbf{R}$.

THÉORÈME: *Les fonctions* Argsh, Argch *et* Argth *son continues sur leurs domaine de définition et dérivable à l'intérieur de celui-ci, de plus :*

i) $\mathrm{Argsh}'(t) = \dfrac{1}{\sqrt{1+t^2}}$ *et* $\mathrm{Argsh}(t) = \log(t + \sqrt{t^2+1})$

ii) $\mathrm{Argch}'(t) = \dfrac{1}{\sqrt{t^2-1}}$ *et* $\mathrm{Argch}(t) = \log(t + \sqrt{t^2-1})$

iii) $\mathrm{Argth}'(t) = \dfrac{1}{1-t^2}$ *et* $\mathrm{Argth}(t) = \dfrac{1}{2}\log\left(\dfrac{1+t}{1-t}\right)$

Démonstration: Les trois énoncés se démontrent de manière similaire ; prouvons le premier : si $x = \mathrm{sh}(y)$, alors $dx/dy = \mathrm{ch}(y) = \sqrt{1+\mathrm{sh}^2(y)} = \sqrt{1+x^2}$. Un calcul direct montre que la dérivée de $\log(x+\sqrt{x^2+1})$ vaut $\dfrac{1}{\sqrt{1+x^2}}$; on en tire donc $\mathrm{Argsh}(x) = \log(x+\sqrt{x^2+1}) + C$ et l'évaluation des deux fontion en $x = 0$ donne $C = 0$. □

FORMULAIRE :
$\mathrm{sh}(-t) = -\mathrm{sh}(t)$, $\mathrm{ch}(-t) = \mathrm{ch}(t)$
$\mathrm{ch}^2(t) - \mathrm{sh}^2(t) = 1$
$\mathrm{sh}(x+t) = \mathrm{sh}(x)\,\mathrm{ch}(t) + \mathrm{ch}(x)\,\mathrm{sh}(t)$
$\mathrm{ch}(x+t) = \mathrm{ch}(x)\,\mathrm{ch}(t) + \mathrm{sh}(t)\,\mathrm{sh}(x)$
$\mathrm{th}(x+t) = \frac{\mathrm{th}(x)+\mathrm{th}(t)}{1+\mathrm{th}(x)\,\mathrm{th}(t)}$
$\mathrm{sh}(t) = \frac{2\,\mathrm{th}(t/2)}{1-\mathrm{th}^2(t/2)}$ et $\mathrm{ch}(t) = \frac{1+\mathrm{th}^2(t/2)}{1-\mathrm{th}^2(t/2)}$
$\mathrm{sh}(x) + \mathrm{sh}(t) = 2\,\mathrm{sh}(\frac{x+t}{2})\,\mathrm{ch}(\frac{x-t}{2})$ et $\mathrm{ch}(x) + \mathrm{ch}(t) = 2\,\mathrm{ch}(\frac{x+t}{2})\,\mathrm{ch}(\frac{x-t}{2})$.

Napier (ou Neper) John (1550-1617)

CHAPITRE 17 CALCUL DE PRIMITIVES

Nous avons vu au chapitre 15 que le calcul d'une intégrale peut se ramener au calcul d'une primitive de la fonction à intégrer. Nous avons constitué au chapitre 16 une bibliothèque de fonctions usuelles. Un espoir un peu naïf consisterait à penser que les primitives de ces fonctions usuelles s'exprime de nouveau à partir de ces fonctions usuelles. Mais des fonctions comme $\exp(-t^2)$ *ou* $\frac{1}{\sqrt{t^3+1}}$ *détruisent cet espoir : les fonctions* $N(x) := \int_0^x \exp(-t^2)dt$ *et* $M(x) := \int_0^x \frac{dt}{\sqrt{t^3+1}}$ *ne peuvent pas s'exprimer à partir des fonctions usuelles ; on peut néanmoins étudier et utiliser ces fonctions (la fonction normale* N *est utilisée en statistiques et probabilités et la fonction* M *– liée aux fonctions elliptiques – est utilisée en physique et dans diverses branches des mathématiques). Après tout la fonction logarithme a été introduite pour donner une primitive à la fonction* $\frac{1}{x}$*. Il existe ainsi une théorie du calcul formel de primitives qui est assez complexe ; nous nous contenterons donc d'étudier un certain nombre de techniques ou "recettes" permettant de calculer (d'exprimer en termes de fonctions usuelles) quelques primitives.*

17.1 QUELQUES PRIMITIVES CLASSIQUES

Rappelons que sur un intervalle, une primitive d'une fonction continue existe et est unique à une constante près. Introduisons la notation fort commode suivante

Notation L'écriture $\int f(t)dt$ désignera une primitive de la fonction f.

Il ne faut pas confondre $\int_a^b f(t)dt$, qui est un nombre et $\int f(t)dt$ qui désigne une fonction ; par exemple on écrira $\int \frac{dt}{t^2+1} = \mathrm{Arctg}(t) + C$.

Il faut faire la remarque qu'il y a parfois une façon assez "bête" de calculer des primitives : si on trouve (par quelque procédé que ce soit) une fonction $F(t)$ dont on sait calculer la dérivée et qui vérifie $F'(t) = f(t)$, alors on a bien sûr trouvé une primitive de f. Commençons ici par une liste de primitives assez faciles à établir :

fonction	primitive	
t^a	$\frac{t^{a+1}}{a+1} + C$	si $a \neq -1$
t^{-1}	$\log\lvert t\rvert + C$	
e^{at}	$\frac{e^{at}}{a} + C$	
a^t	$\frac{a^t}{\log(a)} + C$	si $a > 0$
$\cos(t)$	$\sin(t) + C$	
$\sin(t)$	$-\cos(t) + C$	

fonction	primitive	
$\dfrac{1}{\cos(t)}$	$\log\left\lvert\operatorname{tg}\left(\dfrac{t}{2}+\dfrac{\pi}{4}\right)\right\rvert + C$	si $t \notin \dfrac{\pi}{2}+\pi\mathbf{Z}$
$\dfrac{1}{\sin(t)}$	$\log\left\lvert\operatorname{tg}\left(\dfrac{t}{2}\right)\right\rvert + C$	si $t \notin \pi\mathbf{Z}$
$\operatorname{tg}(t)$	$-\log\lvert\cos(t)\rvert + C$	si $t \notin \dfrac{\pi}{2}+\pi\mathbf{Z}$
$\operatorname{cotg}(t)$	$\log\lvert\sin(t)\rvert + C$	si $t \notin \pi\mathbf{Z}$
$\operatorname{ch}(t)$	$\operatorname{sh}(t) + C$	
$\operatorname{sh}(t)$	$\operatorname{ch}(t) + C$	
$\dfrac{1}{\operatorname{ch}(t)} + C$	$2\operatorname{Arctg}(e^t) + C$	
$\dfrac{1}{\operatorname{sh}^2(t)} + C$	$\log\left\lvert\operatorname{th}\left(\dfrac{t}{2}\right)\right\rvert + C$	si $t \neq 0$
$\operatorname{th}(t)$	$\log\operatorname{ch}(t) + C$	
$\operatorname{coth}(t)$	$\log\lvert\operatorname{sh}(t)\rvert + C$	si $t \neq 0$
$\dfrac{1}{t^2+a^2}$	$\dfrac{1}{a}\operatorname{Arctg}\left(\dfrac{t}{a}\right) + C$	
$\dfrac{1}{t^2-a^2}$	$\dfrac{1}{2a}\log\left\lvert\dfrac{t-a}{t+a}\right\rvert + C$	si $t \neq \pm a$
$\dfrac{1}{\sqrt{t^2+a^2}}$	$\log\lvert t+\sqrt{t^2+a^2}\rvert + C$	
$\dfrac{1}{\sqrt{t^2-a^2}}$	$\log\lvert t+\sqrt{t^2-a^2}\rvert + C$	si $\lvert t\rvert > \lvert a\rvert$
$\dfrac{1}{\sqrt{a^2-t^2}}$	$\operatorname{Arcsin}\left(\dfrac{t}{a}\right) + C$	si $\lvert t\rvert < \lvert a\rvert$

Remarque : on peut souvent donner plusieurs expressions d'une primitive et, dans ce cas, ces expressions sont égales à une constante près. Par exemple une primitive de $\frac{1}{\sqrt{t^2+a^2}}$ est aussi $\operatorname{Argsh}\left(\frac{t}{|a|}\right)$; une primitive de $\frac{1}{\sqrt{t^2-a^2}}$ est aussi $\operatorname{Argch}\left(\frac{t}{|a|}\right)$ (si $t > 0$) ou

$\text{Argch}\left(\frac{|t|}{|a|}\right)$ (si $t<0$).

Démonstration: La plupart de ces formules ont déjà été établies et de toutes façons peuvent se vérifier directement en dérivant ; dans la partie suivante on donne des techniques générales permettant d'ailleurs de les retrouver. Par exemple calculons la dérivée de la fonction $f(t) := \log|\operatorname{tg}(\frac{t}{2}+\frac{\pi}{4})|$:

$$f'(t) = \frac{1}{2}\frac{\operatorname{tg}'(\frac{t}{2}+\frac{\pi}{4})}{\operatorname{tg}(\frac{t}{2}+\frac{\pi}{4})} = \frac{1}{2\operatorname{tg}(\frac{t}{2}+\frac{\pi}{4})\cos^2(\frac{t}{2}+\frac{\pi}{4}))} = \frac{1}{\sin(2(\frac{t}{2}+\frac{\pi}{4}))} = \frac{1}{\cos(t)}$$

.

17.2 TECHNIQUES GÉNÉRALES

On considère dans ce paragraphe uniquement des fonctions f, g, etc continues réelles.
Linéarité : soient $\lambda,\mu\in\mathbf{R}$ et f,g continues :

$$\int_a^b(\lambda f(t)+\mu g(t))dt = \lambda\int_a^b f(t)dt + \mu\int_a^b f(t)dt$$

Exemple : nous savons (voir les formules étudiées au chapitre 4 donnant $\cos(nt)$ comme polynôme en $\cos(t)$) que :

$$\cos^4(t) = \frac{1}{8}(\cos(4t)+4\cos(2t)+3)$$

donc on en déduit :

$$\int\cos^4(t)dt = \frac{1}{32}(\sin(4t)+8\sin(2t)+12t)+C$$

Ainsi, par exemple, $\int_0^{\pi/2}\cos^4(t)dt = 3\pi/16$. Le même style de calcul permet d'intégrer sans difficultés $\cos^m(t)$ ou $\sin^m(t)$ (au moins si m n'est pas trop grand).
Intégration par parties :

THÉORÈME: *Soient f,g deux fonctions continûment dérivables, alors :*

$$\int_a^b f(t)g'(t)dt = f(b)g(b)-f(a)g(a)-\int_a^b f'(t)g(t)dt$$

Démonstration: Il suffit d'intégrer la relation entre dérivées $fg' = (fg)'-f'g$. □

Sous forme mnémotechnique on retiendra : $\int udv = uv - \int vdu$.
Exemples :
1) en prenant $u := \log(t)$, $v := t$ on obtient :

$$\int\log(t)dt = t\log(t)-\int dt = t\log(t)-t+C$$

Plus généralement si $a \in \mathbf{R} \setminus \{-1\}$ et $m \in \mathbf{N}$, posons $u = \log^m(t)$ et $v = \frac{t^{a+1}}{a+1}$; on obtient alors

$$\int t^a \log^m(t)dt = \frac{1}{a+1}t^{a+1}\log^m(t)dt - \frac{m}{a+1}\int t^a \log^{m-1}(t)dt$$

Cette formule, appliquée m fois permet d'exprimer la primitive cherchée sous la forme $t^{a+1}P(\log(t))$ où P est un polynôme de degré m.

2) en prenant $u := t^m$ et $v := e^{at}$ on obtient :

$$\int t^m e^{at}dt = \frac{1}{a}t^m e^{at} - \frac{m}{a}\int t^{m-1}e^{at}dt$$

d'où l'on tire, en itérant m fois, une expression d'une primitive de la forme $e^{at}P(t)$ avec P polynôme de degré m.

3) deux intégrations par parties permettent de calculer $I := \int \sin(t)e^{at}dt$:

$$I = \frac{1}{a}\sin(t)e^{at} - \frac{1}{a}\int \cos(t)e^{at}dt = \frac{1}{a}\sin(t)e^{at} - \frac{1}{a^2}\cos(t)e^{at} - \frac{1}{a^2}\int \sin(t)e^{at}dt$$

d'où

$$I = \frac{1}{1+a^2}(a\sin(t) - \cos(t))e^{at}.$$

4) l'intégration par parties permet aussi de démontrer la formule de Taylor "avec reste intégral" : si f est une fonction $n+1$ fois continument dérivable alors :

$$f(b) = f(a) + \frac{(b-a)}{1!}f'(a) + \ldots + \frac{(b-a)^n}{n!}f^{(n)}(a) + \int_a^b \frac{(b-t)^n}{n!}f^{(n+1)}(t)dt$$

Démonstration: La preuve s'effectue par récurrence sur l'entier n : pour $n = 0$, la formule équivaut au théorème fondamental de l'intégration $f(b) = f(a) + \int_a^b f'(t)dt$. Si l'on pose $u := f^{(n+1)}(t)$ et $v := -(b-t)^{n+1}/(n+1)!$ alors la formule d'intégration par parties fournit : $\int_a^b \frac{b-t)^n}{n!}f^{(n+1)}(t)dt = \frac{b-a)^{n+1}}{(n+1)!}f^{(n+1)}(a) + \int_a^b \frac{b-t)^{n+1}}{(n+1)!}f^{(n+2)}(t)dt$. Cette dernière égalité est exactement celle dont on a besoin pour la récurrence. □

Par exemple, pour la fonction e^x entre 0 et 1 on trouve :

$$e = 1 + \frac{1}{1!} + \frac{1}{2!} + \ldots + \frac{1}{n!} + \int_0^1 \frac{(1-t)^n}{n!}e^t dt$$

Changement de variables :

THÉORÈME: *(première forme) soit g continûment dérivable, f continue, alors :*

$$\int_a^b f \circ g(t)g'(t)dt = \int_{g(a)}^{g(b)} f(u)du$$

Démonstration: Considérons les deux fonctions $F_1(x) := \int_a^x f \circ g(t)g'(t)dt$ et $F_2(x) := \int_{g(a)}^{g(x)} f(u)du$ et soit F une primitive de f alors le théorème fondamental donne que $F_1'(x) = f \circ g(x)g'(x)$ et $F_2(x) = F \circ g(x) - F \circ g(a)$ d'où $F_2'(x) = F'(g(x)g'(x) = f \circ g(x)g'(x)$. On en tire que $F_1 - F_2$ est constante, mais $F_1(a) = F_2(a) = 0$ donc $F_1 = F_2$. □

En fait on emploie rarement le théorème de changement de variables sans que la fonction g ne détermine une bijection de $[a, b]$ sur l'intervalle d'extrémités $g(a), g(b)$. Dans ce cas la formule s'écrit (en notant $h := g^{-1}$) sous sa "deuxième forme" :

$$\int_{h(c)}^{h(d)} f \circ g(t)g'(t)dt = \int_c^d f(u)du$$

Après avoir vérifié que g détermine bien une bijection on peut présenter les calculs de la façon suivante :
(1) On pose $t = g(u)$ ($\Leftrightarrow u = g^{-1}(t)$).
(2) On calcule $dt = g'(u)du$.
(3) On obtient alors $\int f(t)dt = \int f(g(u))g'(u)du$.

Il n'est pas du tout évident de trouver un "bon" changement de variables mais voici quelques exemples (on donne quelques applications supplémentaires dans le paragraphe suivant) :

1) Dans le domaine $]0, a[$ on cherche à intégrer la fonction $f(t) := (t\sqrt{a^2 - t^2})^{-1}$. Pour cela on choisit le changement de variables $u := a/t$ (ou encore $t := a/u$) qui détermine bien une bijection de tout intervalle $[c, d]$ sur $[a/d, a/c]$ (pourvu que $0 < c < d$) on obtient alors :

$$\int \frac{dt}{t\sqrt{a^2 - t^2}} = -\frac{1}{a}\int \frac{du}{\sqrt{u^2 - 1}}.$$

On choisit alors le changement de variables $u = \mathrm{ch}(y)$ (possible car $u > 1$) en observant que $\sqrt{u^2 - 1} = \sqrt{\mathrm{ch}^2(y) - 1} = \sqrt{\mathrm{sh}^2(y)} = \mathrm{sh}(y)$ donc

$$\int \frac{du}{\sqrt{u^2 - 1}} = \int \frac{\mathrm{sh}(y)dy}{\mathrm{sh}(y)} = y + C = \mathrm{Argch}(u)$$

et enfin

$$\int \frac{dt}{t\sqrt{a^2 - t^2}} = -\frac{1}{a}\,\mathrm{Argch}(u) + C = -\frac{1}{a}\,\mathrm{Argch}\left(\frac{a}{t}\right) + C$$

2) En introduisant le changement de variables $u = -t$ on obtient :

$$\int_a^b f(t)dt = \int_{-a}^{-b} f(-u)(-du) = \int_{-b}^{-a} f(-u)du$$

En particulier si f est paire on voit que $\int_0^b f(t)dt = \int_{-b}^0 f(t)dt$ alors que si f est impaire $\int_0^b f(t)dt = -\int_{-b}^0 f(t)dt$; graphiquement on obtient :

f paire $\qquad\qquad$ f impaire

3) Le changement de variable $u := t + c$ (où c est une constante donne la formule $\int_a^b f(t)dt = \int_{a+c}^{b+c} f(u-c)du$ donc par exemple si la fonction f admet pour période T on obtient $\int_a^b f(t)dt = \int_{a+T}^{b+T} f(t)dt$ et ensuite $\int_a^{a+T} f(t)dt = \int_0^T f(u)du$. En effet

$$\int_a^{a+T} f(t)dt = -\int_0^a f(t)dt + \int_0^T f(t)dt + \int_T^{a+T} f(t)dt = \int_0^T f(t)dt.$$

17.3 PRIMITIVES DE FRACTIONS RATIONNELLES

La décomposition en éléments simples d'une fraction rationnelle (chapitre 6) permet de ramener le calcul des primitives de fractions rationelles au calcul d'une primitive de $\frac{1}{(t-a)^n}$ et de $\frac{ct+d}{(t^2+at+b)^n}$. La première est très simple : on obtient $\log|t-a|$ si $n=1$ et $\frac{-1}{(n-1)(t-a)^{n-1}}$ sinon ; la seconde se ramène au calcul d'une primitive de $\frac{1}{(t^2+1)^n}$ et est un peu plus complexe.

FORMULE 1 :

$$\int \frac{dt}{(t-a)^n} = \begin{cases} \dfrac{-1}{(n-1)(t-a)^{n+1}} + C & \text{si } n \neq 1 \\ \log|t-a| + C & \text{si } n = 1 \end{cases}$$

Remarquons maintenant que

$$\frac{ct+d}{(t^2+at+b)^n} = \frac{c}{2}\frac{2t+a}{(t^2+at+b)^n} + \left(d - \frac{ac}{2}\right)\frac{1}{(t^2+at+b)^n}$$

et donc on est ramené par linéarité à intégrer chacun des deux morceaux.

FORMULE 2 :

$$\int \frac{2t+a}{(t^2+at+b)^n}dt = \begin{cases} \frac{-1}{(n-1)(t^2+at+b)^{n-1}} + C & \text{si } n \neq 1 \\ \log|t^2+at+b| + C & \text{si } n = 1 \end{cases}$$

Pour traiter le deuxième membre rappelons que $4b - a^2 > 0$ et notons $\delta := \sqrt{4b - a^2}$ et utilisons la réduction standard d'un polynôme du second degré :

$$t^2 + at + b = \left(t + \frac{a}{2}\right)^2 + \frac{\delta^2}{4} = \frac{\delta^2}{4}\left(\left(\frac{2t+a}{\delta}\right)^2 + 1\right)$$

Ceci suggère d'effectuer le changement de variables $u := \frac{2t+a}{\delta}$; on obtient alors :

$$\int \frac{dt}{(t^2+at+b)^n} = \left(\frac{\delta}{2}\right)^{1-2n} \int \frac{du}{(u^2+1)^n}$$

On est donc bien ramené à calculer une primitive $I_n = \int \frac{du}{(u^2+1)^n}$. Cela peut se faire ainsi :

$$\left(\frac{u}{(u^2+1)^n}\right)' = \frac{1}{(u^2+1)^n} - \frac{2nu^2}{(u^2+1)^{n+1}} = \frac{1-2n}{(u^2+1)^n} + \frac{2n}{(u^2+1)^{n+1}}$$

En intégrant on obtient la formule de récurrence :

FORMULE 3 : (rappelons qu'on a posé $I_n = \int \frac{du}{(u^2+1)^n}$)

$$2nI_{n+1} = (2n-1)I_n + \frac{u}{(u^2+1)^n} + C$$

En particulier

$$I_1 = \int \frac{du}{(u^2+1)} = \mathrm{Arctg}(u) + C$$

$$I_2 = \int \frac{du}{(u^2+1)^2} = \frac{1}{2}\,\mathrm{Arctg}(u) + \frac{u}{2(u^2+1)} + C$$

$$I_3 = \int \frac{du}{(u^2+1)^3} = \frac{3}{8}\,\mathrm{Arctg}(u) + \frac{3u}{8(u^2+1)} + \frac{u}{4(u^2+1)^2} + C$$

(Comme il se doit, la Formule 1 est la plus rapide)

Exemples de calculs explicites :

1) reprenons la fraction rationnelle du chapitre 6, i.e. $f(t) := (t^{10}+t^2+1)/(t^9-2t^5+t)$ dont la décomposition en éléments simples est

$$f(t) = t + \frac{1}{t} + \frac{-3}{16(t+1)^2} + \frac{3}{16(t-1)^2} - \frac{t}{t^2+1} + \frac{t}{4(t^2+1)^2}$$

et ainsi

$$\int f(t)dt = \frac{t^2}{2} + \log|t| + \frac{3}{32}\left(\frac{1}{t+1} - \frac{1}{t-1}\right) - \frac{1}{2}\log|t^2+1| - \frac{1}{8(t^2+1)} + C$$

2) soit n impair, on a calculé durant le chapitre 6 la décomposition suivante :

$$t^{2n}/(t^n-1) = t^n + 1 + \frac{1}{n}\left\{\frac{1}{t-1} + \sum_{h=1}^{\frac{n-1}{2}} \frac{2t\cos(2\pi h/n) - 2}{t^2 - 2t\cos(2\pi h/n) + 1}\right\}$$

Pour calculer une primitive de chacun des derniers éléments simples on observe que $t^2 - 2\cos(a) + 1 = (t-\cos(a))^2 + \sin^2(a) = \left[(t-\cos(a)/\sin(a))^2 + 1\right]\sin^2(a)$ donc si l'on pose $u := (t-\cos(a))/\sin(a)$ on obtient

$$\begin{aligned}\int \frac{dt}{t^2 - 2\cos(a)t + 1} =& \frac{1}{\sin(a)}\int \frac{du}{u^2+1}\\ =& \frac{1}{\sin(a)}\operatorname{Arctg}(u) + C\\ =& \frac{1}{\sin(a)}\operatorname{Arctg}\left((t-\cos(a))/\sin(a)\right) + C\end{aligned}$$

et donc par linéarité on en tire

$$\begin{aligned}\int \frac{t^{2n}dt}{(t^n-1)} =& \frac{t^{n+1}}{n+1} + t + \frac{1}{n}\sum_{h=1}^{\frac{n-1}{2}} \left\{\cos(2\pi h/n)\log|t^2 - 2\cos(2\pi h/n)t + 1|\right.\\ & \left. - 2\sin(2\pi h/n)\operatorname{Arctg}\left(\frac{t-\cos(2\pi h/n)}{\sin(2\pi h/n)}\right)\right\} + C.\end{aligned}$$

Un grand nombre de calculs de primitives se ramène, grâce à un changement de variables adéquat, au calcul d'une primitive de fraction rationnelle. Nous citons ici quelques applications.

i) Si f est une fraction rationnelle en posant le changement de variables $u = e^x$ on obtient la formule $\int f(e^x)dx = \int \frac{f(u)du}{u}$ et l'on sait donc calculer une primitive des fractions rationnelles en e^x.

ii) Soit f une fraction rationnelle en deux indéterminées et cherchons une primitive de $f(\cos(x), \sin(x))$; pour cela on effectue le changement de variables $t := \operatorname{tg}(x/2)$ ou encore $x = 2\operatorname{Arctg}(t)$ et $dx = 2dt/(1+t^2)$ en utilisant $\cos(x) = (1-t^2)/(1+t^2)$ et $\sin(x) = 2t/(1+t^2)$ on en déduit :

$$\int f(\cos(x), \sin(x))dx = \int f\left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}\right)\frac{dt}{t^2+1}$$

donc on saura calculer une primitive.

Par exemple retrouvons par ce principe une primitive de $1/\sin(x)$ et $1/\cos(x)$. On peut écrire

$$\int \frac{dx}{\sin(x)} = \int \frac{dt}{t} = \log|t| + C = \log|\operatorname{tg}(x/2)| + C$$

et en faisant le changement $x := \frac{\pi}{2} - u$

$$\int \frac{dx}{\cos(x)} = -\int \frac{du}{\sin(u)} = -\log|\operatorname{tg}(u/2)| + C = -\log|\operatorname{tg}(\pi/4 - x/2)| + C$$

iii) un autre exemple intéressant est celui des intégrales (dites *abéliennes*) suivantes : il s'agit des primitives de $f(x, \sqrt[m]{(ax+b)/(cx+d)})$ où f est de nouveau une fraction rationnelle en deux indéterminées. On supposera $ad - bc \neq 0$ (sinon le calcul est déjà fait) et on posera

$$u := \sqrt[m]{\frac{ax+b}{cx+d}} \Rightarrow u^m = \frac{ax+b}{cx+d} \Leftrightarrow x = \frac{du^m - b}{-cu^m + a} \text{ donc } dx = (ad-bc)\frac{mu^{m-1}du}{(-cu^m+a)^2}$$

d'où la formule de changement de variables :

$$\int f(x, \sqrt[m]{(ax+b)/(cx+d)})dx = (ad-bc)\int f\left(\frac{du^m - b}{-cu^m + a}, u\right)\frac{mu^{m-1}du}{(-cu^m+a)^2}$$

Prenons un exemple concret : le changement $u := \sqrt[3]{\frac{x-1}{x+1}}$ donne

$$\int \sqrt[3]{\frac{x-1}{x+1}}dx = 6\int \frac{u^3 du}{(u^3-1)^2}$$

(il resterait bien sûr à calculer une primitive de $u^3/(u^3-1)^2$ en la décomposant en éléments simples).

iv) considérons les intégrales du type $\int f(x, \sqrt{ax^2+bx+c})dx$. On factorise le polynôme du second degré en posant $\Delta = b^2 - 4ac = \varepsilon\delta^2$ (où $\varepsilon = \pm 1$) et alors

$$ax^2 + bx + c = \frac{a\delta^2}{4a^2}\left[\left(\frac{2ax}{\delta} - \frac{b}{\delta}\right)^2 - \varepsilon\right]$$

on effectue naturellement le changement de variable $u := \frac{2ax}{\delta} - \frac{b}{\delta}$ et, si η désigne le signe de a on obtient une intégrale du type $\int g(u, \sqrt{\eta u^2 - \eta\varepsilon})du$ où g est une fraction rationnelle. On distingue alors les cas suivants :

CAS $\eta = \varepsilon = 1$: on pose $u = \operatorname{ch}(t)$ (avec disons $t > 0$) donc $\sqrt{u^2-1} = \operatorname{sh}(t)$ et $\int g(u, \sqrt{u^2-1})du$ devient $\int g(\operatorname{ch}(t), \operatorname{sh}(t))\operatorname{sh}(t)dt$ que l'on sait intégrer.

CAS $\eta = 1$,$\varepsilon = -1$: on pose $u = \operatorname{sh}(t)$ et donc on a $\sqrt{u^2+1} = \operatorname{ch}(t)$ et l'intégrale $\int g(u, \sqrt{u^2+1})du$ devient $\int g(\operatorname{sh}(t), \operatorname{ch}(t))\operatorname{ch}(t)dt$ que l'on sait intégrer.

CAS $\eta = -1$, $\varepsilon = 1$: on pose $u := \sin(t)$ (sur un intervalle convenable) donc $\sqrt{1-u^2} = |\cos(t)|$ et $\int g(u, \sqrt{1-u^2})du$ devient $\int g(\sin(t), |\cos(t)|)\cos(t)dt$ que l'on sait intégrer (le cas $\eta = \varepsilon = -1$ est, bien sûr, exclus).

CHAPITRE 18 INTÉGRALES IMPROPRES

Nous avons vu que l'intégrale d'une fonction continue par morceau sur un intervalle fermé $[a,b]$ est toujours définie. Une intégrale "indéfinie" ou "posant problème" est une intégrale soit d'une fonction non continue sur un intervalle $[a,b]$ (par exemple $\int_{-1}^{1} dx/x$) ou d'une fonction continue sur un intervalle non borné comme $[a,+\infty)$ (par exemple $\int_1^{+\infty} dx/x$). Ces intégrales n'existent pas toujours et on étudie dans ce chapitre des conditions de convergence et même dans certains cas des méthodes de calcul. Il faut remarquer que de nombreuses intégrales de ce type surgissent naturellement dans les sciences : les calculs de potentiel créé par une masse conduisent à de tels intégrales ; on peut signaler une formule importante (entre autres) en probabilité et statistique :

$$\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}$$

Rappelons que l'on ne peut pas exprimer par des fonctions élémentaires une primitive de la fonction e^{-x^2}.

18.1 INTÉGRALES IMPROPRES

Définition: Soit f une fonction continue $[a,b[\to \mathbf{R}$; si $\lim_{c\to b}\int_a^c f(t)dt$ existe, alors on appelle cette limite l'*intégrale impropre* de f entre a et b et on écrit :

$$\int_a^b f(t)dt := \lim_{\substack{c\to b\\ c<b}} \int_a^c f(t)dt$$

Convention : on devrait normalement se garder d'écrire $\int_a^b f(t)dt$ avant de savoir que cette limite existe ; il est néanmoins d'usage d'écrire "l'intégrale $\int_a^b f(t)dt$ est divergente" si cette limite n'existe pas.

Donnons quelques exemples :

1) Soit $f(t)=\log(t)$ alors $\int_x^1 \log(t)dt = [t\log(t)-t]_x^1 = x - x\log(x) - 1$ donc

$$\int_0^1 \log(t)dt = \lim_{x\to 0}(x - x\log(x) - 1) = -1$$

L'aire hachurée vaut 1

2) Soit $f(t)=t^a$ alors $\int_1^c t^a dt = [t^{a+1}/(a+1)]_1^c = c^{a+1}/(a+1) - 1/a + 1$ (resp. $[\log(t)]_1^c = \log(c)$) si $a\neq -1$ (resp. si $a=-1$). La limite lorsque c tend vers $+\infty$ existe

seulement quand $a < -1$, la limite quand c tend vers zéro existe seulement quand $a > -1$ et on obtient

$$\int_1^{+\infty} t^{-a}dt = \begin{cases} \frac{1}{a-1} & \text{si } a > 1 \\ \text{diverge} & \text{si } a \leq 1 \end{cases}$$

$$\int_0^1 t^{-a}dt = \begin{cases} \frac{1}{1-a} & \text{si } a < 1 \\ \text{diverge} & \text{si } a \geq 1 \end{cases}$$

Par exemple $\int_0^1 dt/\sqrt{t} = 2$ et $\int_1^{+\infty} dt/t^2\sqrt{t} = 2/3$.

Remarque : Si une intégrale est impropre en plusieurs points, on dira qu'elle est convergente si elle convergente au voisinage de chaque point. Par exemple l'intégrale $\int_{-\infty}^{+\infty} e^{-|t|}dt/\sqrt{|t|}$ est définie comme :

$$\lim_{Y\to-\infty} \int_Y^{-1} f(t)dt + \lim_{d\to 0^-} \int_{-1}^{d} f(t)dt + \lim_{c\to 0^+} \int_c^1 f(t)dt + \lim_{X\to+\infty} \int_1^X f(t)dt$$

ou $f(t) := e^{-|t|}dt/\sqrt{|t|}$. En termes concrets : pour étudier une intégrale impropre on "découpe les difficultés".

Pour démontrer le théorème on utilisera le critère assez intuitif suivant (dit *critère de Cauchy*) : soit $F : [a, b[\to \mathbf{R}$ telle que $F(c) - F(c')$ tende vers zéro lorsque $a \leq c \leq c' < b$ et c tend vers b, alors $F(x)$ tend vers une limite finie lorsque x tend vers b.

THÉORÈME: *Si l'intégrale $\int_a^b |f(t)|dt$, impropre en b, est convergente alors l'intégrale $\int_a^b f(t)dt$ est également convergente et l'on a $|\int_a^b f(t)dt| \leq \int_a^b |f(t)|dt$.*

Démonstration: D'après le critère évoqué précédemment (appliqué à $F(x) = \int_a^x f(t)dt$) l'intégrale $\int_a^b f(t)dt$ converge si et seulement si $\int_c^{c'} f(t)dt$ tend vers 0 lorsque $a \leq c \leq c' \leq b$ et c tend vers b. Le théorème découle alors de l'inégalité $|\int_c^{c'} f(t)dt| \leq \int_c^{c'} |f(t)|dt$. □

La réciproque est fausse ; par exemple nous verrons que $\int_0^{+\infty} \frac{\sin(x)}{x}dx$ est convergente alors que $\int_0^{+\infty} \left|\frac{\sin(x)}{x}\right| dx$ est divergente.

THÉORÈME: *Supposons que $|f(t)| \leq g(t)$ alors la convergence de l'intégrale $\int_a^b g(t)dt$ entraîne la convergence de $\int_a^b f(t)dt$; la divergence de $\int_a^b f(t)dt$ entraîne la divergence $\int_a^b g(t)dt$.*

Démonstration: D'après le théorème précédent, il suffit de voir que $\int_a^b |f(t)|dt$ est convergente. Or la fonction $F(x) := \int_a^x |f(t)|dt$ est croissante et bornée par $\int_a^b g(t)dt$ donc converge vers une valeur finie lorsque x tend vers b. □

THÉORÈME: *Soient $f(x), g(x)$ deux fonctions positives équivalentes quand x tend vers b alors l'intégrale $\int_a^b f(t)dt$ est convergente si et seulement si l'intégrale $\int_a^b g(t)dt$ est convergente.*

Démonstration: Comme les fonctions sont équivalentes, on a sur un voisinage de b des inégalités de la forme : $Cf(x) \leq g(x) \leq C'f(x)$; en appliquant le théorème précédent on en déduit que la convergence d'une des deux intégrales entraîne la convergence de l'autre. □

Exemples de calculs (les détails sont laissés en exercice) :

$$\int_a^b \frac{dt}{\sqrt{(t-a)(b-t)}} = \int_{-1}^1 \frac{dt}{\sqrt{1-t^2}} = \pi, \quad \int_{-\infty}^{+\infty} \frac{dt}{1+t^2} = \pi.$$

18.2 CALCULS D'INTÉGRALES IMPROPRES

Les techniques de calcul que nous abordons peuvent se résumer ainsi : on applique les techniques de calcul des intégrales finies (linéarité, intégration par parties, changement de variable) et on passe à la limite. Ce paragraphe est donc un paragraphe d'exercices et exemples, on détermine notamment quand $\int_0^{+\infty} (a_1 t^{s_1} + \ldots + a_r t^{s_r})\, e^{-t} dt/t$ *est convergente.*

PROPOSITION: *(Linéarité) Soient* $\alpha, \beta \in \mathbf{R}$ *et* $f, g : [a, b[\to \mathbf{R}$ *deux fonctions continues dont les intégrales sur* $[a, b]$ *convergent, alors l'intégrale de* $\alpha f + \beta g$ *converge et :*

$$\int_a^b (\alpha f(t) + \beta g(t))dt = \alpha \int_a^b f(t)dt + \beta \int_a^b g(t)dt.$$

Démonstration: On utilise la linéarité de l'intégrale usuelle et on passe à la limite. □

Exemple : Posons $\Gamma(s) := \int_0^{+\infty} e^{-t} t^s dt/t$, les critères du paragraphe précédent permettent d'établir que l'intégrale converge si et seulement si $s > 0$. Si donc $s_1, \ldots, s_r > 0$ alors

$$\int_0^{+\infty} (a_1 t^{s_1} + \ldots + a_r t^{s_r})\, e^{-t} dt/t = a_1 \Gamma(s_1) + \ldots + a_r \Gamma(s_r).$$

PROPOSITION: *(Intégration par parties) Soient* $f, g : [a, b[\to \mathbf{R}$ *continûment dérivables alors :*

$$\int_a^b f(t)g'(t)dt = \lim_{c \to b} f(c)g(c) - f(a)g(a) - \int_a^b f'(t)g(t)dt$$

où la convergence de deux termes entraîne la convergence du troisième.

Démonstration: On utilise l'intégration par parties de l'intégrale usuelle et on passe à la limite. □

Exemple :

$$\Gamma(s) = \int_0^{+\infty} e^{-t} t^{s-1} dt = \lim_{c \to +\infty} e^{-c} \frac{c'^s}{s} - \lim_{c' \to 0} e^{-c'} \frac{c'^s}{s} + \frac{1}{s} \int_0^{+\infty} e^{-t} t^s dt = \frac{\Gamma(s+1)}{s}$$

En calculant $\Gamma(1) = 1$, on en tire $\Gamma(m) = (m-1)!$ pour $m \in \mathbf{N}^*$.

PROPOSITION: *(Changement de variable) Soit $f : [c, d[\to \mathbf{R}$ continue dont l'intégrale sur $[c, d]$ converge, soit $g : [a, b[\to [c, d[$ une bijection continûment dérivable alors :*

$$\int_a^b f \circ g(t) g'(t) dt = \int_c^d f(u) du$$

Démonstration: On utilise le changement de variable de l'intégrale usuelle et on passe à la limite. □

Exemple : 1) Soit $I(b) := \int_{-\infty}^{+\infty} e^{-(x-a)^2/b} dt$. Faisons le changement de variable $\sqrt{b}u := x - a$, on obtient $I(b) = \sqrt{b} I(1)$. Si l'on sait que $I(1) = \sqrt{\pi}$ on obtient la formule :

$$\int_{-\infty}^{+\infty} e^{-(x-a)^2/b} dx = \sqrt{\pi b}$$

2) On peut aussi en déduire la valeur de $\Gamma(1/2)$ (et donc de $\Gamma(n+1/2)$) ; on effectue cette fois le changement de variable $u = \sqrt{t}$:

$$\Gamma(1/2) = \int_0^{+\infty} e^{-t} \frac{dt}{\sqrt{t}} = 2 \int_0^{+\infty} e^{-u^2} du = \int_{-\infty}^{+\infty} e^{-u^2} du = \sqrt{\pi}$$

Terminons ce chapitre par l'étude d'une intégrale "semi-convergente" :

L'intégrale $\int_0^{+\infty} \left| \frac{\sin(x)}{x} \right| dx$ est divergente mais $\int_0^{+\infty} \frac{\sin(x)}{x} dx$ est convergente (sa valeur est en fait $\pi/2$ mais nous ne prouverons pas cela).

Démonstration: Lorsque x tend vers 0 alors $\sin(x)/x$ tend vers 1 donc l'intégrale n'est qu'apparemment impropre en 0. Utilisons les inégalités suivantes :

$$\int_{N\pi}^{(N+1)\pi} \left| \frac{\sin(x)}{x} \right| dx \geq \int_{N\pi}^{(N+1)\pi} \frac{|\sin(x)|}{(N+1)\pi} dx = \frac{1}{(N+1)\pi} \int_0^{\pi} |\sin(x)| dx = \frac{2}{(N+1)\pi}$$

donc $\int_0^{(M+1)\pi} \left| \frac{\sin(x)}{x} \right| dx \geq \frac{2}{\pi} \sum_{N=0}^{M} \frac{1}{N+1}$, or cette dernière somme tend vers l'infini donc l'intégrale $\int_0^{+\infty} \left| \frac{\sin(x)}{x} \right| dx$ est divergente.

Pour prouver la convergence de $\int_0^{+\infty} \frac{\sin(x)}{x} dx$, on peut utiliser une intégration par parties :

$$\int_1^X \frac{\sin(x)}{x} dx = \frac{-\cos(X)}{X} + \cos(1) - \int_1^X \frac{\cos(x)}{x^2} dx$$

or $\int_1^{+\infty} \frac{\cos(x)}{x^2} dx$ est convergente car $|\frac{\cos(x)}{x^2}| \leq \frac{1}{x^2}$ et $\int_1^{+\infty} \frac{1}{x^2} dx = 1$ est convergente. Ainsi $\int_0^{+\infty} \frac{\sin(x)}{x} dx$ est convergente. □

Remarque : Pour calculer une valeur approchée d'une intégrale impropre comme la précédente on procède en deux temps. On majore $\int_X^{+\infty} \frac{\sin(x)}{x} dx$, ce qui peut se faire ainsi :

$$\int_X^{+\infty} \frac{\sin(x)}{x} dx = \frac{\cos(X)}{X} - \int_X^{+\infty} \frac{\cos(x)}{x^2} dx$$

La dernière intégrale est majorée,en valeur absolue, par $\int_X^{+\infty} dx/x^2 = 1/X$ donc on a $|\int_X^{+\infty} \frac{\sin(x)}{x} dx| \leq 2/X$. Ensuite on calcule une valeur approchée I de $\int_0^X \frac{\sin(x)}{x} dx$ à ε près, par exemple par la méthode des trapèzes et on conclut que $|\int_0^{+\infty} \frac{\sin(x)}{x} dx - I| \leq \varepsilon + 2/X$.

Abel Niels (1802–1829)

CHAPITRE 19 COURBES PARAMÉTRÉES ET DÉVELOPPEMENTS LIMITÉS

L'objet de ce chapitre est l'étude du comportement local (i.e. au voisinage d'un point) des courbes. Le premier paragraphe est consacré à des techniques de calculs de limites ou d'équivalents. Par exemple les deux limites suivantes

$$\lim_{x\to 0} \frac{\cos(x)-1}{x\sin(x)} \quad \text{et} \quad \lim_{x\to 0} \frac{e^x - 2\sqrt{1+x}+1}{x^2}$$

ne sont pas accessibles avec les méthodes développées avant ce chapitre ; le calcul des développements limités permet d'y remédier. La deuxième partie est consacrée à l'étude des courbes proprement dites ; plus précisément on étudie des courbes un peu plus générales que les graphes de fonctions et qui interviennent souvent par exemple en cinématique ou en mécanique (où les coordonnées d'un mobile sont connues en fonction du temps). L'exemple le plus basique en physique est le mouvement d'un solide soumis à la gravitation : avec les notations usuelles en physique $\vec{F} = m\vec{\gamma}$ *donc* $-mg = my''$ *et* $0 = mx''$ *d'où* $y(t) = -gt^2/2 + y'(0)t + y(0)$ *et* $x(t) = x'(0)t + x(0)$*, ce qui détermine une parabole.*

19.1 DÉVELOPPEMENTS LIMITÉS

Notation : on désignera dans tout ce chapitre par $\varepsilon(x)$ une fonction définie au voisinage de 0 et tendant vers 0 lorsque x tend vers 0.

Définition: Une fonction $f : I \to \mathbf{R}$ admet un *développement limité* (DL) à l'ordre n au voisinage de $x_0 \in I$ si il existe une fonction polynôme $P(x) = a_0 + a_1(x-x_0) + \ldots + a_n(x-x_0)^n$ tel que

$$f(x) = P(x) + (x-x_0)^n \varepsilon(x-x_0)$$

Exemple : une fonction admet un DL en un point x_0 d'ordre 1 si et seulement si elle est dérivable en ce point.

Remarque sur les notations : nous nous contenterons de la notation $\varepsilon(x)$ qui désignera dans tout ce chapitre une fonction tendant vers zéro quand x tend vers zéro (qui ne sera pas toujours la même!). Plusieurs auteurs utilisent (au moins) deux autres notations que nous signalons donc : la notation $O(x^n)$ ("grand O") désigne une fonction bornée en valeur absolue par $C|x|^n$ lorsque x tend vers zéro ; la notation $o(x^n)$ ("petit o") désigne une fonction dont le quotient par x^n tend vers zéro quand x tend vers zéro.

PROPOSITION: *Un développement limité d'ordre n, si il existe, est unique.*

Démonstration: Supposons que $f(x) = P(x) + x^n\varepsilon(x) = Q(x) + x^n\eta(x)$ où P, Q sont des polynômes de degré $\leq n$ et ε, η tendent vers zéro quand x tend vers zéro. Alors, si le polynôme $P - Q$ n'était pas nul, la fonction $(P-Q)(x)$ serait équivalente à une fonction ax^r avec $a \neq 0$ et $0 \leq r \leq n$ et ne pourrait être égale à une fonction du type $\varepsilon(x)x^n$ car $\lim \frac{\varepsilon(x)x^n}{ax^r} = 0$. □

Remarque : l'unicité permet de simplifier certains calculs car elle entraîne qu'une fonction paire (resp. impaire) n'aura que des termes de degré pair du type x^{2m} (resp. impair du type x^{2m+1}) dans un développement limité.

La méthode la plus courante pour montrer l'existence des DL est d'utiliser la formule de Taylor qui entraîne que :

THÉORÈME: (Formule de Taylor-Young)*Soit f une fonction n fois continument dérivable sur un intervalle I contenant x_0, alors*

$$f(x) = f(x_0)+f'(x_0)(x-x_0)+\frac{f''(x_0)}{2!}(x-x_0)^2+\ldots+\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n+(x-x_0)^n\varepsilon(x-x_0)$$

Démonstration: Quitte à faire une translation, on peut supposer que $x_0 = 0$. La formule de Taylor-Lagrange s'écrit :

$$f(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \ldots + \frac{f^{(n)}(c)}{n!}x^n$$

avec c compris entre 0 et x. Ainsi lorsque x tend vers zéro, c tend aussi vers zéro et comme $f^{(n)}$ est continue $f^{(n)}(c) - f^{(n)}(0)$ tend vers zéro quand x tend vers zéro donc $f^{(n)}(c)x^n = f^{(n)}(0)x^n + \varepsilon(x)x^n$; en reportant dans la formule de Taylor-Lagrange, on obtient le résultat escompté. □

Par application directe on trouve :

COROLLAIRE: *Les fonctions e^x, $\cos(x)$, $\sin(x)$ et $(1+x)^a$ admettent les DL suivants :*

$$\begin{aligned}
e^x =& 1 + x + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!} + x^n\varepsilon(x)\\
\cos(x) =& 1 - \frac{x^2}{2!} + \frac{x^4}{4!} + \ldots + (-1)^n\frac{x^{2n}}{(2n)!} + x^{2n}\varepsilon(x)\\
\sin(x) =& x - \frac{x^3}{3_!} + \frac{x^5}{5!} + \ldots + (-1)^n\frac{x^{2n+1}}{(2n+1)!} + x^{2n+1}\varepsilon(x)\\
(1+x)^a =& 1 + ax + \frac{a(a-1)}{2}x^2 + \ldots + \frac{a(a-1)\ldots(a-n+1)}{n!}x^n + x^n\varepsilon(x)
\end{aligned}$$

Exemple d'utilisation : $\lim\limits_{x\to 0}\dfrac{\cos(x)-1}{x\sin(x)} = -\dfrac{1}{2}$ et $\lim\limits_{x\to 0}\dfrac{e^x - 2\sqrt{1+x}+1}{x^2} = \dfrac{3}{4}$.

En effet $\cos(x)-1 = -x^2/2+x^2\varepsilon(x)$ et $x\sin(x) = x^2+x^2\varepsilon(x)$ donc $\frac{\cos(x)-1}{x\sin(x)} = -\frac{1}{2}+\varepsilon(x)$ tend vers $-\frac{1}{2}$. Par ailleurs $\sqrt{1+x} = 1 + x/2 - x^2/8 + x^2\varepsilon(x)$ donc $e^x - 2\sqrt{1+x}+1 = (1+x+x^2/2) - 2(1+x/2-x^2/8)+1+x^2\varepsilon(x) = \frac{3}{4}x^2 + x^2\varepsilon(x)$ d'où la deuxième limite.

THÉORÈME: *Soient f, g deux fonctions continues de variables réelles dont on suppose qu'elles admettent un DL à l'ordre n de la forme : $f(x) = a_0+a_1x+\ldots+a_nx^n+\varepsilon(x)x^n = P(x)+\varepsilon(x)x^n$ et $g(x) = b_0 + b_1x + \ldots + b_nx^n + \varepsilon(x)x^n = Q(x) + \varepsilon(x)x^n$*

(i) (Somme) La fonction $f + g$ admet un DL à l'ordre n en 0 qui s'écrit :

$$(f+g)(x) = (a_0+b_0) + (a_1+b_1)x + \ldots + (a_n+b_n)x^n + \varepsilon(x)x^n$$

(ii) (Produit) La fonction fg admet un DL à l'ordre n en 0 qui s'écrit : $(fg)(x) = c_0 + \ldots + c_n x^n + \varepsilon(x)x^n$ *avec* $c_i = \sum_{k=0}^{i} a_k b_{i-k}$

(iii) (Quotient) Supposons $b_0 \neq 0$ *(c'est-à-dire* $g(0) \neq 0$*) alors la fonction* f/g *admet un DL à l'ordre* n *en 0 qui s'écrit :* $(f/g)(x) = c_0 + \ldots + c_n x^n + \varepsilon(x)x^n$ *avec des* c_i *que l'on peut déterminer par la relation* $a_i = \sum_{k=0}^{i} c_k b_{i-k}$.

(iv) (Composée) Supposons $b_0 = 0$ *alors la fonction* $f \circ g$ *admet un DL à l'ordre* n *en 0 qui s'écrit :* $(f \circ g)(x) = R(x) + \varepsilon(x)x^n$ *où* $R(x)$ *est le polynôme* $P \circ Q(x)$ *auquel on a retiré les termes de degré* $> n$.

(v) (Intégration) Soit $F(x)$ *une primitive de* f *alors elle admet un DL d'ordre* $n+1$ *qui s'écrit* $F(x) = F(0) + a_0 x + a_1 x^2/2 + \ldots + a_n x^{n+1}/(n+1) + \varepsilon(x)x^{n+1}$.

Démonstration: Les deux premiers énoncés s'obtiennent immédiatement en additionnant ou multipliant les DL de f et g. Les points (iii) et (iv) sont également un simple calcul si l'on admet l'existence d'un DL, qui est un peu plus délicate à prouver et que nous admettrons. Enfin le dernier point s'obtient par intégration du DL de $f(x)$; en effet $\int_0^x t^m dt = x^{m+1}/(m+1)$ et, si l'on note $\eta(x) = sup_{t\in[0,x]}|\varepsilon(t)|$, on a :

$$\left|\int_0^x \varepsilon(t)t^n dt\right| \leq \eta(x)\int_0^x t^n dt = \eta(x)\frac{x^{m+1}}{(m+1)}$$

d'où le résultat. □

APPLICATION: Par application directe on obtient les DL suivants :

$$\begin{aligned}
\log(1+x) =& x - \frac{x^2}{2} + \frac{x^3}{3} + \ldots + (-1)^{n+1}\frac{x^n}{n} + x^n\varepsilon(x)\\
\mathrm{Arctg}(x) =& x - \frac{x^3}{3} + \ldots + (-1)^n\frac{x^{2n+1}}{2n+1} + \varepsilon(x)x^{2n+1}\\
\mathrm{Argsh}(x) =& x - \frac{x^3}{6} + \ldots + (-1)^n C_{2n}^n\frac{x^{2n+1}}{4^n(2n+1)} + \varepsilon(x)x^{2n+1}\\
\mathrm{tg}(x) =& x + \frac{x^3}{3} + \frac{2x^5}{15} + \varepsilon(x)x^6
\end{aligned}$$

Par la formule de Taylor on a calculé le DL de $(1+x)^{-1}$ et de $(1+x)^{-1/2}$ Le premier s'obtient en intégrant le DL de $1/(1+x)$; en utilisant le DL d'une fonction composée on obtient le DL de $1/(1+x^2)$ et $(1+x^2)^{-1/2}$ et en intégrant on obtient le DL de $\mathrm{Arctg}(x)$ et $\mathrm{Argsh}(x)$. En utilisant la formule de Taylor et le fait que la fonction tangente est impaire, on sait que $\mathrm{tg}(x) = c_1 x + c_3 x^3 + c_5 x^5 + \varepsilon(x)x^6$; on peut déterminer les coefficients c_i soit en calculant les dérivées successives de tg soit en utilisant les DL de $\sin(x)$ et $\cos(x)$ et la règle pour le quotient de deux DL.

19.2 COURBES PARAMÉTRÉES

Définition: On appelle *courbe paramétrée* une application de I dans $\mathbf{R}^2$ où I est un intervalle ou une union d'intervalles.

Notation : on écrira en général $M(t) = (x(t), y(t))$ le point du plan donné par la valeur t du paramètre.

Remarque : d'après la définition il faudrait distinguer l'ensemble $C = \{(x(t), y(t)) \in \mathbf{R}^2 \mid t \in I\}$ – qui correspond à l'appellation usuelle de courbe– et la courbe paramétrée elle-même qui est l'application $t \mapsto (x(t), y(t))$; mais, conformément à l'usage, on dira aussi que la courbe C est paramétrée par l'application. En particulier une courbe (au sens usuel) peut être paramétrée de diverses façons. De nombreux problèmes conduisent à des courbes paramétrées, commençons par donner des exemples de paramétrisations de courbes simples.

Exemples :

1) La courbe paramétrée définie par $x(t) = at + b$ et $y(t) = ct + d$ est une droite dont la pente est c/a.

2) L'ellipse d'équation $x^2/a^2 + y^2/b^2 = 1$ peut être paramétrée de plusieurs façons : d'abord à l'aide de fonctions circulaires : $M(t) = (a\cos(t), b\sin(t))$ mais aussi à l'aide de fractions rationnelles : si l'on coupe l'ellipse par la droite de pente t passant par le point $(a, 0)$ on obtient (après un petit calcul) un point

$$M(t) = \left(a\frac{a^2t^2 - b^2}{a^2t^2 + b^2}, t\left(a\frac{a^2t^2 - b^2}{a^2t^2 + b^2} - a\right)\right) = \left(a\frac{a^2t^2 - b^2}{a^2t^2 + b^2}, \frac{-2ab^2t}{a^2t^2 + b^2}\right)$$

Ceci fournit une paramétrisation de l'ellipse moins le point $(a, 0)$ qui n'est pas atteint.

Remarque : ces deux paramétrisations peuvent bien sûr être reliées. Explicitons cela dans le cas du cercle avec $a = b = 1$. Le point $\left(\frac{t^2-1}{t^2+1}, \frac{-2t}{t^2+1}\right)$ correspond à $(\cos(\theta), \sin(\theta))$ lorsque $t = -\operatorname{tg}(\theta/2)$.

Nous nous concentrons maintenant sur l'étude et le tracé des courbes paramétrées. Déterminons, dans certains cas, la pente de la tangente d'une courbe paramétrée :

THÉORÈME: *Soit $t \mapsto (f(t), g(t))$ une courbe paramétrée par un intervalle I (avec $t_0 \in I$) ; supposons les deux fonctions f et g continûment dérivables en t_0 et telle que $f'(t_0) \neq 0$ alors la pente de la tangente, au point $M(t_0) = (f(t_0), g(t_0))$ de la courbe paramétrée, est donnée par*

$$p = \frac{g'(t_0)}{f'(t_0)}.$$

Démonstration: Pour fixer les idées, disons que $f'(t_0) > 0$ alors $f'(t)$, étant continue, reste strictement positive sur un (petit) intervalle $]a, b[$ contenant t_0 et donc f détermine une bijection de $]a, b[$ sur un autre intervalle ; notons $h := f^{-1}$ la bijection réciproque,

on sait que si $t = h(x)$ alors $h'(x) = 1/f'(t)$, de plus on aura alors $y = g \circ h(x)$ et $dy/dx = g'(h(x))h'(x) = g'(t)/f'(t)$. $\square$

Remarques :

i) Mnémotechniquement on pourra utiliser : $\dfrac{dy}{dx} = \dfrac{dy}{dt} / \dfrac{dx}{dt}$.

ii) Si $f'(t_0) = 0$ mais $g'(t_0) \neq 0$ alors, en intervertissant les rôles des coordonnés on voit que la pente est verticale. Ceci nous laisse à traiter le cas où $f'(t_0) = g'(t_0) = 0$.

Définition: Le point $t_0 \in I$ (ou I est un intervalle) est dit *singulier* pour la courbe paramétrée $M : I \to \mathbf{R}^2$ donnée par $t \mapsto (f(t), g(t))$ si $f'(t_0) = g'(t_0) = 0$.

Pour étudier un point singulier, nous allons utiliser les DL des fonctions f, g au voisinage de t_0 (quand ils existent).

Notation : si $f(t) = a_0 + a_1 t + \ldots a_n t^n + \varepsilon(t)t^n$ et $g(t) = b_0 + b_1 t + \ldots b_n t^n + \varepsilon(t)t^n$, notons $M(t) = \begin{pmatrix} f(t) \\ g(t) \end{pmatrix}$ et $e_i = \begin{pmatrix} a_i \\ b_i \end{pmatrix}$, nous écrirons :

$$M(t) = e_0 + te_1 + \ldots + t^n e_n + t^n \varepsilon(t)$$

où $\varepsilon(t)$ désignera un vecteur dont les coordonnées tendent vers 0 quand t tend vers 0. On introduit aussi la notation $M'(t) = \begin{pmatrix} f'(t) \\ g'(t) \end{pmatrix}$ et plus généralement $M^{(n)}(t) = \begin{pmatrix} f^{(n)}(t) \\ g^{(n)}(t) \end{pmatrix}$. On obtient alors, sous l'hypothèse que f et g sont suffisamment dérivable une formule de Taylor-Young vectorielle :

$$M(t) = M(t_0) + \frac{(t-t_0)}{1!}M'(t_0) + \frac{(t-t_0)^2}{2!}M''(t_0) + \ldots + \frac{(t-t_0)^n}{n!}M^{(n)}(t_0) + (t-t_0)^n \varepsilon(t-t_0)$$

THÉORÈME: *Soient $m < n$ les plus petits entiers tels que les vecteurs $M^{(m)}(t_0)$ et $M^{(n)}(t_0)$ soient linéairement indépendants, notons $e_1 = M^{(m)}(t_0)$ et $e_2 = M^{(n)}(t_0)$ et supposons que $\det(e_1, e_2) > 0$, alors l'allure de la courbe paramétrée au voisinage de $M(t_0)$ est la suivante :*

$m,\ n$ impairs $\qquad$ m impair, n pair

m pair n impair $\qquad$ m pair, n pair

Si $\det(e_1, e_2) < 0$*, il faut opérer une symétrie (par rapport à la droite engendrée par* e_1*) pour obtenir l'allure de la courbe au voisinage de* $M(t_0)$*.*

Démonstration: (esquisse) Après éventuellement une symétrie et à un changement de repère près (qui ne modifie pas l'allure du graphe) on peut se ramener à étudier les courbes $M(t) := (t^m, t^n)$ ce qui est relativement aisé : si m est impair on aura $t = x^{1/m}$ et la courbe est le graphe de la fonction $f(x) = \left(x^{1/m}\right)^n$, c'est-à-dire :

Si m est pair alors $x \geq 0$ et, pour $t \geq 0$ on a $t = x^{1/m}$ et $y = \left(x^{1/m}\right)^n$ alors que pour $t \leq 0$, on a $y = (-1)^n \left(x^{1/m}\right)^n$ c'est-à-dire :

□

Comme illustration, incluons l'exemple suivant dont les détails sont laissés en exercice. On veut étudier le comportement au voisinage de $t = 0$ de la courbe $M(t) = (x(t), y(t)) = (t^2 + t^4, t^2 + t^5)$. On a $M'(0) = M'''(0) = 0$ et $M''(0) = (2, 2)$ indépendant de $M^{(4)}(0) = (24, 40)$. Pour déterminer la position de la demi-branche "$t > 0$" par rapport à celle "$t < 0$" on observe qu'ici, $x(t) = x(-t)$ et que, pour $t > 0$ on a $y(t) = y(-t) + 10t^5 > y(-t)$ d'où l'esquisse suivante :

Passons maintenant à l'étude pratique d'une courbe paramétrée ; un plan général d'étude possible est le suivant

– Détermination du domaine, des périodes et des symétries éventuelles ;
– Calcul de $x'(t)$ et de $y'(t)$, tableau de variations, asymptotes ;
– Etude des points singuliers, calcul de quelques tangentes ;
– Détermination des points doubles ;
– Représentation graphique.

Plutôt que de faire une étude théorique, illustrons cela sur un exemple :
Etudions la courbe :

$$M(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} t + \frac{4}{t-1} \\ t + 1 + \frac{4}{(t-1)^2} \end{pmatrix}$$

– Les fonctions $x(t)$ et $y(t)$ sont définies pour $t \in \mathbf{R} \setminus \{1\}$; il n'y a ni période ni symétrie apparente.

– On calcule $x'(t) = (t-3)(t+1)/(t-1)^2$ et $y'(t) = (t-3)(t^2+3)/(t-1)^3$ donc le point correspondant à $t = 3$ c'est-à-dire $M(3) = \begin{pmatrix} 5 \\ 5 \end{pmatrix}$ est singulier, et le point $M(-1) = \begin{pmatrix} -3 \\ 1 \end{pmatrix}$ a une tangente verticale. Les limites de $x(t)$ et $y(t)$ lorsque t tend vers 1 ou $\pm\infty$ sont faciles à calculer ainsi que les sens de variation ; on résume ceci dans un tableau :

t	$-\infty$		-1		1^-	1^+		3		$+\infty$
$x'(t)$		$+$	0	$-$			$-$	0	$+$	
$x(t)$	$-\infty$	$\nearrow$	-3	$\searrow$	$-\infty$	$+\infty$	$\searrow$	5	$\nearrow$	$+\infty$
$y(t)$	$-\infty$		$\nearrow$		$+\infty$	$+\infty$	$\searrow$	5	$\nearrow$	$+\infty$
$y'(t)$			$+$				$-$	0	$+$	

Pour déterminer si la courbe admet une asymptote lorsque t tend vers $1^{\pm}$ ou vers $\pm\infty$ on calcule le rapport $y(t)/x(t) = t^3 - t^2 - t + 5/(t-1)(t^2 - t + 4)$. Ce rapport ne tend pas vers une limite finie lorsque t tend vers 1 donc il n'y a pas d'asymptote dans ce cas. Par contre $\lim_{t\to\infty} y(t)/x(t) = 1$ et donc s'il existe une asymptote elle sera de la forme $y = x + b$; donc il nous faut maintenant voir si $y(t) - x(t)$ tend vers une limite finie : $y(t) - x(t) = 1 + 4/(t-1)^2 - 4/(t-1)$ tend vers 1 donc la droite $y = x + 1$ est asymptote ; de plus si t tend vers $+\infty$ on a $y(t) - x(t) - 1 < 0$ et si t tend vers $-\infty$ on a $y(t) - x(t) - 1 > 0$ donc dans le premier cas la branche de la courbe est au dessous de l'asymptote et dans le deuxième cas elle est au dessus de l'asymptote.

– On étudie le comportement de la courbe au voisinage du point singulier $M(3) = \begin{pmatrix} 5 \\ 5 \end{pmatrix}$: on calcule tout d'abord $M'(3) = 0$ (bien sûr) puis $M''(3) = \begin{pmatrix} 1 \\ 3/2 \end{pmatrix}$ et ensuite $M'''(3) = \begin{pmatrix} -3/2 \\ -3 \end{pmatrix}$. Ces deux derniers vecteurs sont linéairement indépendants et un calcul immédiat montre que $\det(M''(3), M'''(3)) < 0$. Ceci permet donc de déterminer l'allure de la courbe au voisinage de $M(3)$:

– Il n'y a pas de point double, c'est-à-dire que les équations $x(t) = x(s)$ et $y(t) = y(s)$ entraînent $t = s$. Vérifions cela : $t + 4/(t-1) = s + 4/(s-1)$ équivaut à $(t-s) = 4(t-s)/(s-1)(t-1)$ donc à $t = s$ ou $(s-1)(t-1) = 4$. Par ailleurs $t + 1 + 4/(t -$

$1)^2 = s + 1 + 4/(s-1)^2$ équivaut à $(t-s) = 4(s-t)(s+t-2)/(s-1)^2(t-1)^2$ donc à $t = s$ ou $s + t = 2 + (s-1)^2(t-1)^2/4$. Si l'on exclut $t = s$ on aboutit à ce que $(s-1)(t-1) = 4$ et $(t-1) + (s-1) = 4$ ce qui signifie que $s-1$ et $t-1$ sont les racines de $X^2 - 4X + 4 = (X-2)^2 = 0$ ce qui entraîne $s = t = 3$.

– Avant de tracer la courbe, notons que l'on peut préciser un peu le comportement de la courbe quand t tend vers $1^{\pm}$. En effet on observe que

$$y(t) - \left(\frac{x-1}{2}\right)^2 = (t-1)(1-(t-1)/4)$$

donc lorsque t tend vers 1, la courbe se rapproche de la parabole d'équation $y = (x-1)^2/4$. On peut même préciser que lorsque $t \to 1^+$ la courbe est *au dessus* de la parabole et lorsque $t \to 1^-$ la courbe est *au dessous* de la parabole. On généralise, sur cet exemple, la notion de courbe asymptote : on se contente souvent de rechercher les droites asymptotes, mais on peut rechercher plus généralement des courbes asymptotes.

– On peut aussi calculer les points d'intersection de la courbe avec l'asymptote $y = x + 1$, on trouve $t = 2$ et $M(2) = \begin{pmatrix} 6 \\ 7 \end{pmatrix}$ ainsi que les points d'intersection avec la parabole $y - \left(\frac{x-1}{2}\right)^2$, on trouve $t = 5$ et $M(5) = \begin{pmatrix} 6 \\ 25/4 \end{pmatrix}$.

On peut maintenant tracer le graphe de la courbe que l'on vient d'étudier (voir page suivante).

Comme autre exemple on décrit très succintement l'étude de la courbe (il s'agit d'un exemple de courbes dites de Lissajoux).

$$M(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} \sin(2t) \\ \cos(3t) \end{pmatrix}$$

La "caractéristique" de cette courbe est de posséder beaucoup de symétries; mettons en évidence les principales symétries.

– On a bien sûr $M(t+2\pi) = M(t)$ donc la courbe est périodique et il suffit de l'étudier pour t dans un intervalle de longueur 2π.

– On remarque aussi que $\begin{pmatrix} x(t+\pi) \\ y(t+\pi) \end{pmatrix} = \begin{pmatrix} x(t) \\ -y(t) \end{pmatrix}$ donc on peut prendre un intervalle de longueur π, on obtiendra la courbe complète par symétrie par rapport à l'axe des x.

– On a aussi $\begin{pmatrix} x(-t) \\ y(-t) \end{pmatrix} = \begin{pmatrix} -x(t) \\ y(t) \end{pmatrix}$ donc on peut prendre un intervalle centré en 0 et le réduire de moitié, on obtiendra la courbe complète par symétrie par rapport à l'axe des y.

Tracé de la courbe $M(t) = (x(t), y(t)) = \left(t + \frac{4}{t-1}, t + 1 + \frac{4}{(t-1)^2}\right)$:

– Après avoir fait un tableau de variations (ce qui est laissé en exercice), on peut tracer la courbe (en trait continu la partie correspondant à $t \in [0, \pi/2]$, en pointillé le reste de la courbe obtenu par symétrie) :

Tracé de la courbe $M(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} \sin(2t) \\ \cos(3t) \end{pmatrix}$.

Nous laissons en exercice le soin de déterminer les sept points doubles (si l'on restreint le paramètre t à l'intervalle $[-\pi, +\pi[$).

Remarque : on peut observer cette courbe sur l'écran d'un oscilloscope à condition d'y entrer deux tensions dont les fréquences sont dans un rapport $\frac{3}{2}$ (ou $\frac{2}{3}$).

Gauss Carl Friedrich (1777–1855)

CHAPITRE 20 EQUATIONS DIFFÉRENTIELLES

Il y a peu de concept mathématique aussi universellement utilisé dans les autres sciences que celui d'équation différentielle. On les retrouve aussi bien en biologie, en chimie qu'en économie. Newton disait que traiter mathématiquement un problème physique revient à trouver l'équation différentielle qui le décrit. Il n'y a donc que l'embarras du choix pour donner des exemples de modélisation de phénomènes par des équations différentielles. Observons que dans ce contexte l'idée d'unicité – qui apparaît si importante au mathématicien et parfois si inutile au profane– révèle son importance : si la solution est unique, on peut espérer que les équations que l'on a écrites décrivent complètement le phénomène. Remarquons qu'il serait aussi illusoire de penser que l'on peut "résoudre" toutes les équations différentielles explicitement. En effet nous avons déjà mentionné que des équations comme $y' = e^{x^2}$ n'avait pas de solution "élémentaire" ; on peut raffiner la question en s'autorisant les quadratures (i.e. on s'autorise à prendre des primitives), la réponse est de nouveau négative : cela ne suffit pas pour exprimer les solutions des équations différentielles en général. Il est donc important de développer une étude qualitative des équations différentielles ; nous abordons très succintement ce thème dans le premier paragraphe avant de nous concentrer sur des types très importants d'équations que l'on sait en fait résoudre explicitement.

20.1 INTRODUCTION, EXEMPLES

On donne quelques exemples d'équations différentielles et quelques considérations géométriques sur ces équations.

Commençons par donner une définition (un peu restrictive) d'une équation différentielle :

Définition: Une *équation différentielle d'ordre n* est une équation de la forme :

$$y^{(n)} = f(y^{(n-1)}, \ldots, y', y, x)$$

Une *solution* de l'équation différentielle est une fonction $g : I \to \mathbf{R}$ qui est n fois dérivable et vérifie $g^{(n)}(x) = f(g^{(n-1)}(x), \ldots, g'(x), g(x), x)$

Remarques :

i) Si $g : I \to \mathbf{R}$ est une solution et si $J \subset I$ alors visiblement la restriction de g à J est encore une solution, mais cela ne nous apprend rien sur l'équation ; on conviendra donc de ne s'intéresser qu'aux *solutions maximales*, i.e. les solutions qui ne sont pas des restrictions de solutions définies sur des intervalles plus grands.

ii) Le cas "général" serait plutôt une équation du type $f(y^{(n)}, y^{(n-1)}, \ldots, y', y, x) = 0$.

iii) Il est clair qu'une tel équation n'a presque jamais une solution unique (penser au calcul de primitives i.e. aux équations du type $y' = f(x)$) ; l'intuition physique peut nous guider : pour espérer avoir unicité des solutions il faut rajouter des *conditions initiales*, c'est-à-dire connaître $y^{(n-1)}(x_0)$,…,$y'(x_0)$, $y(x_0)$; par exemple pour une équation de la mécanique de Newton qui fait intervenir dérivée première (vitesse) et dérivée seconde (accélération), il faut connaître la position et la vitesse initiales d'un solide pour connaître son mouvement.

EXEMPLES : Etudions les solutions de l'équation : $y' = \sqrt[3]{y}$.

Soit I un intervalle où une solution y ne s'annule pas, alors on peut réécrire l'équation $y'/\sqrt[3]{y} = 1$ ce qui équivaut à $\frac{2}{3}y^{2/3} = x - C$ soit encore $y = \left(\frac{2}{3}(x-C)\right)^{3/2}$. On obtient ainsi des solutions sur tout intervalle $[C, +\infty)$. Par ailleurs $y = 0$ est une solution évidente ; ceci permet de décrire un certain nombre de solutions ; nous nous contenterons de les dessiner et d'observer que l'équation n'est pas "déterministe" : si disons $y(0) = 0$, rien ne permet de dire quand y "décollera" pour prendre des valeurs positives.

On n'a tracé que des solutions positives ; on peut en obtenir d'autres en utilisant la symétrie : si $y(x)$ est une solution, alors $y_1(x) := -y(-x)$ est également solution.

L'équation précédente est un cas particulier des *équations à variables séparées* i.e. celles que l'on peut écrire sous la forme

$$f(y)y' = g(x) \qquad (E)$$

Pour chercher les solutions de (E), on introduit F une primitive de f et G une primitive de g. Si $y(x)$ est une solution de (E) alors $F(y(x)) = G(x) + K$ (où K désigne une constante). Si on se place sur un intervalle où F définit une bijection, on poura alors écrire $y(x) = F^{-1}(G(x) + K)$. A titre d'exemple étudions l'équation

$$\sin^2(x)y' - \cos(x)y^2 = 0$$

En se plaçant hors de $x = 0$ et la où $y(x) \neq 0$, on peut réécrire l'équation $y'/y^2 = \cos(x)/\sin^2(x)$ qui donne $-1/y(x) = -1/\sin(x) + K$ d'où $y(x) = \sin(x)/(1 - K\sin(x))$.

Considérations géométriques

Définition: Un *champ de vecteurs* dans le plan est la donnée en chaque point du plan d'un vecteur du plan ; en d'autre termes c'est une application de $\mathbf{R}^2$ dans $\mathbf{R}^2$.

Le champ de vecteur $v(x, y) = (x, x + y^2)$

Le lien avec les équations différentielles est donné par l'observation suivante : si on souhaite étudier une équation différentielle du type $y' = f(x, y)$, on définit un champ de vecteur en choisissant en chaque point (x, y) un vecteur de pente $f(x, y)$. Alors il est clair qu'une solution de l'équation différentielle est une courbe tangente en chaque point au champ de vecteurs.

On peut représenter cela ainsi

Ceci permet déjà une certaine analyse qualitative des solutions. Par exemple le tracé des courbes $f(x,y) = a$ découpe le plan en région dans lesquelles la pente des solutions est $\leq a$ ou $\geq a$; en particulier la courbe $f(x,y) = 0$ partage le plan en plusieurs parties : celles où les solutions sont croissantes et celles où les solutions sont décroissantes.

Donnons des exemples de propriétés géométriques vérifiables par le calcul et par le dessin du champ de vecteurs.

Champ associé à $y' = ay + b$

L'allure des solutions sera confirmé au prochain paragraphe où l'on verra qu'elles sont de la forme $y(x) = \lambda e^{ax} - b/a$

Champ associé à $y' = f(x)$

Vu le parallélisme du champ de vecteur on voit que si $y(x)$ est solution $y(x) + C$ également ; ceci est clair car y doit être une primitive de f

Champ associé à $y' = f(y)$

Vu le parallélisme du champ de vecteur on voit que si $y(x)$ est solution $y(x + C)$ également ; ceci est clair car x doit être une primitive de $y'/f(y)$.

Exemples de modélisation par des équations différentielles :

1) Une piscine de 40000 litres a son eau renouvelée au rythme de 200 litres par heures. Par accident l'eau de regénération est polluée avec une concentration de 5% par un produit toxique dont on sait qu'une concentration supérieure à 3% peut être dangereuse. Comment varie la concentration du produit toxique en fonction du temps, la concentration dangereuse est-elle atteinte? Au bout de combien de temps?

Modélisation : appelons $y(t)$ la concentration et $Y(t) = 40000y(t)$ la quantité (en litres) du produit toxique dans la piscine. Considérons un "petit" accroissement du temps h, alors $Y(t+h)$ est égal $Y(t)$ augmenté de $200 \times 0,05 \times h$ (la quantité de produit déversé durant l'intervalle de longueur h) et diminué de $200 \times y(t) \times h$; ici on néglige des termes du second ordre (par rapport à h), i.e. on fait l'approximation que la concentration est constante, cette approximation est justifié (au moins intuitivement) car on va faire tendre h vers zéro. On obtient donc $Y(t + h) = Y(t) + 10h - 200y(t)h$ d'où en faisant tendre h vers zéro

$$Y'(t) + 0,0005Y(t) = 10$$

On verra (il s'agit d'une équation linéaire du premier ordre avec coefficients constants) que les solutions sont de la forme $Y(t) = Ce^{-0,005t} + 2000$ et donc si $Y(0) = 0$ (eau non polluée avant l'accident) on obtient $Y(t) = 2000(1 - e^{-0,005t})$ et $y(t) = 0,05(1 - e^{-0,005t})$. L'étude du champ de vecteurs associé donna d'ailleurs l'allure de ces courbes. On voit ainsi que la concentration tendra vers 5% avec le temps et atteindra 3% au bout de $200 \log(2,25) \sim 183$ heures.

VARIANTE : L'amiante est un poison dont l'inhalation provoque diverses maladies pulmonaires dont des cancers mortels (cancers broncho-pulmonaires, et mésothéliomes de la plèvre et du péritoine) dont le temps de latence est particulièrement long (respectivement 30 ans et 40 ans en moyenne) ; heureusement on n'est pas pressé! Sachant que l'air qu'on respire à Jussieu contient de l'ordre de quelques fibres par litre, que l'être humain respire plusieurs milliers de litres d'air par jour, que la concentration de fibres d'amiante dans l'air peut être multipliée par 50 par l'activité dans une pièce, par 500 lors de travaux dans les faux-plafonds (passage de câbles, réparation de circuit électrique, de canalisation), sachant que l'on ignore s'il existe un seuil au dessous duquel l'amiante cesse d'induire des cancers, pouvez-vous évaluer, à l'aide d'une équation différentielle, le risque encouru à Jussieu 1) par un étudiant (qui passe en moyenne 3 à 4 ans à l'université), 2) par un enseignant ou administratif (qui y travaille de façon permanente) 3) par un agent de maintenance ou de nettoyage (qui sera amené à opérer dans les faux-plafonds ou à nettoyer après)?

2) Si on considère qu'un ressort exerce une force proportionnelle à l'éloignement du solide qui lui est attaché, l'équation de la mécanique $\vec{F} = m\vec{\gamma}$ se traduit ainsi, en notant y la position du solide, m sa masse, 0 le point de fixation du ressort et k le rapport de proportionalité donnant la force exercée : $-ky = my''$.

Posons $\omega := \sqrt{k/m}$ alors l'équation s'écrit :

$$y'' + \omega^2 y = 0$$

Nous verrons que les solutions de cette équation sont du type $y(t) = A\sin(\omega t) + B\cos(\omega t)$ ou si l'on préfère $y(t) = C\cos(\omega t + \phi)$. Ceci permet de voir que les oscillations restent bornées (d'amplitude C) et périodiques (de période $2\pi/\omega$) ; on peut aussi constater que la connaissance de la position et vitesse initiale (disons $y(0)$ et $y'(0)$) permet de déterminer entièrement le mouvement du solide ($B = y(0)$ et $A = y'(0)/\omega$).

L'étude d'un circuit électrique "RLC" (avec résistance, inductance et capacité) conduit à des équations du même type.

3) On considère un pendule de longueur ℓ, de masse m accroché en un point O et on note y l'angle que la tige du pendule fait avec la verticale ; comme dans cette figure :

L'équation de la mécanique $\vec{F} = m\vec{\gamma}$ se traduit par $-mg\sin(y) = m\ell y''$ (où g est la constante universelle de gravitation) et en posant $\omega := \sqrt{g/\ell}$ on obtient :

$$y'' + \omega^2 \sin(y) = 0$$

Cette équation est nettement plus compliquée que les deux précédentes car elle n'est pas linéaire en y et l'on ne sait pas exprimer ses solutions par une formule. On peut néanmoins l'étudier qualitativement, par exemple essayer de savoir si il y a des solutions périodiques ou si une solution doit être périodique. Ces problèmes sont difficiles et nous nous contenterons d'observer que si les oscillations du pendule sont suffisamment petites, on aura $\sin(y(t)) \sim y(t)$ et "donc" que les solutions de l'équation $y'' + \omega^2 y = 0$ représenterons une bonne approximation des solutions de l'équation différentielle originale.

20.2 ÉQUATIONS DIFFÉRENTIELLES LINÉAIRES DU PREMIER ORDRE

Il s'agit des équations du type $y' = a(x)y + b(x)$. Une observation simple a montré que le champ de vecteurs associé à une forme simple au dessus de chaque intervalle où le signe de $a(x)$ est constant. On va voir que l'on peut en fait résoudre par quadrature ces équations.

Commençons par observer que l'ensemble des solutions forme un espace affine ; c'est-à-dire que si y_0 est une solution et y en est une autre alors $y - y_0$ est solution de l'équation $z' = a(x)z$. Or l'ensemble des solutions de cette dernière équation différentielle est visiblement un espace vectoriel.

Définition: On appelle équation différentielle *homogène* associée à l'équation $y' = a(x)y + b(x)$, l'équation $y' = a(x)y$.

Ainsi toute solution d'une équation différentielle linéaire du premier ordre est somme d'une solution particulière et d'une solution de l'équation homogène associée.

Commençons par la résolution de l'équation homogène (remarquez qu'il s'agit d'une équation à variables séparées) :

THÉORÈME: *Soit $(x_0, y_0) \in \mathbf{R}^2$ et soit $a : I \to \mathbf{R}$ une fonction continue avec $x_0 \in I$; il existe une unique solution de l'équation différentielle $y' = a(x)y$ telle que $y(x_0) = y_0$; celle-ci est donné explicitement par*

$$y(x) := y_0 \exp\left(\int_{x_0}^{x} a(t)dt\right)$$

La solution générale de l'équation différentielle est $y(x) = C \exp(A(x))$ où $A(x)$ désigne une primitive de $a(x)$.

Démonstration: On vérifie facilement que la fonction annoncée est une solution ; inversement, si y est une solution, posons $z(x) := y(x) \exp\left(-\int_{x_0}^{x} a(t)dt\right)$ alors z est dérivable et un calcul direct donne

$$z'(x) = (y'(x) - a(x)y(x)) \exp\left(-\int_{x_0}^{x} a(t)dt\right) = 0$$

d'où $z(x) = C$ et on trouve donc $y(x) = C \exp\left(\int_{x_0}^{x} a(t)dt\right)$. Imposer la condition $y(x_0) = y_0$ revient alors à imposer $C = y_0$. □

Exemples : soient λ, b des réels non nuls, on se restreint à $x \in \mathbf{R}_+^*$, alors l'équation $y' = \lambda x^b y$ a pour solution $y(x) = C \exp\left(\frac{\lambda x^{b+1}}{b+1}\right)$ si $b \neq -1$ et $y(x) = Cx^\lambda$ si $b = -1$.

Passons à la recherche d'une solution particulière pour une équation du type

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

Dans certains cas il est facile de trouver une solution simple ; par exemple l'équation différentielle $y' = \lambda y + \mu$ admet visiblement comme solution particulière $y = -\mu/\lambda$ par

conséquent la solution générale de l'équation est $y(x) = Ce^{\lambda x} - \mu/\lambda$. Toutefois il se peut qu'aucune solution ne soit "visible" et alors il est utile de recourir à une méthode générale dite de "variation de la constante" (sic).

L'idée est la suivante : l'équation homogène a pour solution $y_1(x) = C\exp\left(\int_{x_0}^x a(t)dt\right)$ et on va chercher une solution y sous la forme $y(x) = C(x)\exp\left(\int_{x_0}^x a(t)dt\right)$.

THÉORÈME: *Soit $(x_0, y_0) \in I \times \mathbf{R}$ et soit $a, b : I \to \mathbf{R}$ deux fonctions continues ; il existe une unique solution de l'équation différentielle $y' = a(x)y + b(x)$ telle que $y(x_0) = y_0$; celle-ci est donné explicitement par*

$$y(x) := y_0 \exp\left(\int_{x_0}^x a(t)dt\right) + \int_{x_0}^x b(t)\exp\left(-\int_x^t a(u)du\right)dt$$

La solution générale de l'équation différentielle est obtenue en faisant varier y_0.

Démonstration: Soit y une solution de l'équation différentielle étudiée, posons $C(x) := y(x)\exp\left(-\int_{x_0}^x a(t)dt\right)$; alors

$$C'(x) = (y'(x) - a(x)y(x))\exp\left(-\int_{x_0}^x a(t)dt\right) = b(x)\exp\left(-\int_{x_0}^x a(t)dt\right)$$

donc si on désigne par $y_1(x)$ la fonction $\exp\left(\int_{x_0}^x a(t)dt\right)$ on obtient $C(x) = C_0 + \int_{x_0}^x \frac{b(t)dt}{y_1(t)}$ et $y(x) = C_0 y_1(x) + y_1(x)\int_{x_0}^x \frac{b(t)dt}{y_1(t)}$. Inversement il est facile de vérifier qu'une telle fonction est bien solution de l'équation différentielle. La condition $y(x_0) = y_0$ équivaut à $C_0 = y_0$. □

Exemples

Considérons l'équation $y' = 2xy + 1$ alors le calcul donne (en prenant $x_0 = 0$) une solution homogène $y_1(x) = C\exp(x^2)$ et une solution générale $y(x) = C\exp(x^2) + \exp(x^2)\int_0^x \exp(-t^2)dt$.

L'équation $y' = -2y/x + e^x$ a pour solution homogène $y_1(x) = C|x|^{-2}$ et, sur chacun des intervalles $(-\infty, 0[$ et $]0, +\infty)$, la solution générale s'écrit $y(x) = C|x|^{-2} + |x|^{-2}\int_{x_0}^x e^t t^2 dt$ ou encore

$$y(x) = Cx^{-2} + \left(1 - 2x^{-1} + 2x^{-2}\right)e^x$$

20.3 ÉQUATIONS DIFFÉRENTIELLES LINÉAIRES DU SECOND ORDRE

On s'intéresse maintenant aux équations différentielles linéaires du second ordre. On traitera plus particulièrement le cas des coefficients constants mais le début de la discussion est générale ne nécessite pas cette hypothèse. Il s'agit d'équations du type :

$$y'' + a(x)y' + b(x)y = c(x)$$

La première observation est que de nouveau l'ensemble des solutions forme un espace affine et que toute solution sera somme d'une solution particulière et d'une solution de l'équation homogène correspondante

$$y'' + a(x)y' + b(x)y = 0$$

Commençons par traiter l'équation homogène en montrant le théorème général suivant :

THÉORÈME: *Soit $a, b : I \to \mathbf{R}$ deux fonctions continues, l'espace vectoriel des solutions de l'équation $y'' + a(x)y' + b(x)y = 0$ a une dimension égale au plus à deux. En fait il existe au plus une solution telle que $y'(x_0) = y'_0$ et $y(x_0) = y_0$.*

Démonstration: Pour chaque paire de solutions y_1, y_2 fabriquons la fonction

$$W(x) := \det \begin{pmatrix} y_1(x) & y'(x) \\ y_2(x) & y'_2(x) \end{pmatrix} = (y_1 y'_2 - y'_1 y_2)(x)$$

Un calcul direct donne

$$W'(x) = y_1 y''_2 - y''_1 y_2 = y_1(-a(x)y'_2 - b(x)y_2) - y_2(-a(x)y'_1 - b(x)y_1) = -a(x)W(x)$$

donc $W(x)$ est elle-même solution d'une équation différentielle linéaire du premier ordre et donc $W(x) = W(x_0) \exp\left(-\int_{x_0}^{x} a(t)dt\right)$; en particulier on a soit $\forall x \in I, W(x) \neq 0$ soit $\forall x \in I, W(x) = 0$.

Supposons qu'il existe deux solutions y_1, y_2 linéairement indépendantes (s'il n'en existe pas alors l'espace des solutions est de dimension ≤ 1) ; on sait que $(y_1 y'_2 - y'_1 y_2)(x_0) \neq 0$ donc pour toute autre solution y il existe λ, μ tels que $y(x_0) = \lambda y_1(x_0) + \mu y_2(x_0)$ et $y'(x_0) = \lambda y'_1(x_0) + \mu y'_2(x_0)$ (le système linéaire correspondant est de Cramer). Par conséquent la fonction W associée à y et $\lambda y_1 + \mu y_2$ est nulle et donc y est une combinaison linéaire de y_1 et y_2. □

Le déterminant W considéré dans la preuve s'appelle le *wronskien* en l'honneur du mathématicien H. Wronsky.

Remarquons que la démonstration contient des renseignements plus riches que ceux donnés dans l'énoncé du théorème ; on a en particulier montré :

COROLLAIRE: Soit y_1, y_2 deux solutions de $y'' + a(x)y' + b(x)y = 0$ telles que $(y_1 y'_2 - y'_1 y_2)(x_0) \neq 0$ pour un point x_0 alors y_1 et y_2 forment une base de l'espace vectoriel des solutions de l'équation différentielle.

Exemple : on peut vérifier que sur $\mathbf{R}_+^*$ les deux fonctions $y_1(x) = \cos(\sqrt{x})$ et $y_2(x) = \sin(\sqrt{x})$ forment une base des solutions de $4xy'' + 2y' + y = 0$. Remarquons que pour appliquer la théorie précédente il faut d'abord diviser par $4x$ et donc travailler sur un intervalle de $\mathbf{R}^*$. De même une base des solutions sur $\mathbf{R}_-^*$ est fournie par les deux fonctions $y_1(x) = \exp\left(\sqrt{-x}\right)$ et $y_1(x) = \exp\left(-\sqrt{-x}\right)$.

Nous allons maintenant considérer des équations avec a, b constants ; ce cas est très important d'abord parce que le principe général des développements limités nous a appris que toute fonction assez régulière peut être approchée par une fonction linéaire donc

les équations linéaires, à coefficients constants, décrivent au moins localement le comportement des solutions de beaucoup d'équation différentielles ; de plus un bon nombre de modélisations conduisent naturellement à des équations de ce type et elles présentent l'avantage qu'on peut décrire explicitement leurs solutions.

Définition: L'*équation caractéristique* associée à une équation différentielle $y'' + ay' + by = 0$ est l'équation polynomiale $X^2 + aX + b = 0$

Bien sûr cette notion est analogue à celle introduite pour les suites linéaires récurrentes (chapitre 12).

THÉORÈME: *Considérons l'équation différentielle à coefficients constants $y'' + ay' + by = 0$ et appelons α_1, α_2 les racines de l'équation caractéristique associée.*

i) Si $\alpha_1 \neq \alpha_2$ alors une base des solutions de l'équation différentielle est donnée par les fonctions $x \mapsto e^{\alpha_1 x}$ et $x \mapsto e^{\alpha_2 x}$.

ii) Si $\alpha_1 = \alpha_2 = \alpha$ alors une base des solutions de l'équation différentielle est donnée par les fonctions $x \mapsto e^{\alpha x}$ et $x \mapsto xe^{\alpha x}$.

De plus dans le premier cas, si les coefficients sont réels et les racines complexes $\alpha_1 = u + i\omega$ et $\alpha_2 = u - i\omega$, une base de solutions réelles est donnée par les fonctions $x \mapsto e^{ux}\cos(\omega x)$ et $x \mapsto e^{ux}\sin(\omega x)$.

Démonstration: D'après le théorème précédent, il suffit de vérifier que les deux fonctions sont solutions de l'équation différentielle (ce qui est immédiat à vérifier) et indépendantes en un point x_0, mais dans le premier cas $(y_1y_2' - y_1'y_2)(x_0) = (\alpha_2 - \alpha_1)e^{(\alpha_1+\alpha_2)x_0} \neq 0$ et dans le second cas $(y_1y_2' - y_1'y_2)(x_0) = e^{\alpha x_0} \neq 0$.Pour la dernière affirmation il suffit de se rappeler que $e^{(u+i\omega)x} + e^{(u-i\omega)x} = 2e^{ux}\cos(\omega x)$ et $e^{(u+i\omega)x} - e^{(u-i\omega)x} = 2ie^{ux}\sin(\omega x)$. □

Remarque : Le comportement général des solutions quand x tend vers l'infini est gouverné par les racines de l'équation caractéristique : si les racines vérifient $Re(\alpha_i) < 0$ alors toutes les solutions tendent vers zéro quand x tend vers $+\infty$; si les racines vérifient seulement $Re(\alpha_i) \leq 0$ et sont distinctes, toutes les solutions sont bornées, lorsque x tend vers $+\infty$.

La recherche d'une solution particulière de l'équation $y'' + ay' + by = c(x)$ est possible par une variante de la méthode de la variation des constantes mais nous ne traiterons que le cas où $c(x)$ est le produit d'un polynôme et d'une exponentielle (ou d'une fonction circulaire).

THÉORÈME: *Soit $P(x)$ un polynôme de degré m alors il existe une solution particulière de l'équation différentielle $y'' + ay' + by = P(x)e^{\beta x}$ de la forme suivante :*

i) Si $\beta^2 + a\beta + b \neq 0$ on peut prendre $y(x) = Q(x)e^{\beta x}$ avec $deg(Q) = m$

ii) Si $\beta^2 + a\beta + b = 0$ mais $2\beta + a \neq 0$ on peut prendre $y(x) = Q(x)e^{\beta x}$ avec $deg(Q) = m + 1$

iii) Si $\beta^2 + a\beta + b = 0$ et $2\beta + a = 0$ on peut prendre $y(x) = Q(x)e^{\beta x}$ avec $deg(Q) = m + 2$

Démonstration: Appelons E_m l'ensemble des fonctions f de la forme $f(x) = P(x)e^{\beta x}$ avec P polynôme de degré $\leq m$; c'est un espace vectoriel de dimension $m + 1$ (il est

clairement isomorphe à l'espace vectoriel des polynômes de degré $\leq m$. et notons L l'application qui à f associe $f'' + af' + f$. On calcule

$$L(f) = \left(P'' + (2\beta + a)P' + (\beta^2 + a\beta + b)P\right) e^{\beta x}$$

On voit tout de suite que L est donc une application linéaire de E_m vers E_m.

i) Comme $\beta^2 + a\beta + b \neq 0$, le degré de $P'' + (2\beta + a)P' + (\beta^2 + a\beta + b)P$ est égal à celui de P et donc $L(f) = 0$ équivaut à $f = 0$. L'application L étant injective doit être surjective, ce qui est la première affirmation du théorème.

ii) Comme $\beta^2 + a\beta + b = 0$ mais $2\beta + a \neq 0$ on a

$$deg\left(P'' + (2\beta + a)P' + (\beta^2 + a\beta + b)P\right) = deg(P) - 1$$

ainsi l'image de E_m par L se trouve dans E_{m-1} et le noyau de L est constitué des fonctions $Ce^{\beta x}$ donc est de dimension 1 ; l'image $\mathrm{Im}(L)$ a donc même dimension que E_{m-1} et lui est donc égale ; c'est-à-dire que pour tout polynôme P de degré $m-1$ il existe un polynôme Q de degré m tel que $L(Q(x)e^{\beta x} = P(x)e^{\beta x}$.

iii) Comme $\beta^2 + a\beta + b = 0$ et $2\beta + a = 0$, on a

$$deg\left(P'' + (2\beta + a)P' + (\beta^2 + a\beta + b)P\right) = deg(P) - 2$$

ainsi le noyau de L est constitué des fonctions $(C_1 + C_2 x)e^{\beta x}$ donc est de dimension deux ; l'image $\mathrm{Im}(L)$ est de dimension $m-1$ et contenue dans E_{m-2} donc lui est égale ; c'est-à-dire que pour tout polynôme P de degré $m-2$ il existe un polynôme Q de degré m tel que $L(Q(x)e^{\beta x} = P(x)e^{\beta x}$. $\square$

APPLICATION: (Résonance) Etudions le problème d'un ressort recevant une agitation périodique, ou d'un circuit électrique recevant un courant sinusoïdal. Un tel phénomène est décrit par une équation du type

$$y'' + \omega^2 y = a\cos(\rho x) + b\sin(\rho x)$$

Si $\rho \neq \pm\omega$ alors y est la superposition d'une solution générale de l'équation différentielle homogène : $\lambda\cos(\omega x) + \mu\sin(\omega x)$ et d'une solution particulière du type $\gamma\cos(\rho x) + \delta\sin(\rho x)$; en particulier les solutions sont bornées.

Si $\rho = \pm\omega$ alors y est la superposition d'une solution générale de l'équation différentielle homogène : $\lambda\cos(\omega x) + \mu\sin(\omega x)$ et d'une solution particulière du type $(ax + c)\cos(\rho x) + (bx + d)\sin(\rho x)$; en particulier les solutions ne sont pas bornées – ce qui veut dire que le ressort va casser ou que le circuit électrique va griller. Ce phénomène (le fait qu'un système possède une fréquence critique) s'appelle le phénomène de *résonance* ; c'est la raison pour laquelle les défilés militaires cessent de se faire au pas cadencé quand ils passent sur un pont.

Remarquons aussi que pour chercher une solution particulière de $y'' + \omega^2 y = g_1(x) + \ldots + g_s(x)$ il suffit de chercher une solution particulière y_i de $y'' + \omega^2 y = g_i(x)$ et ensuite d'additionner les y_i. Par exemple si $g_i(x) = A_i\cos(\omega_i x + \phi_i)$ avec $\omega_i^2 \neq \omega^2$ alors il y aura

une solution particulière de la forme $y_0(x) = C_1\cos(\omega_1 x + \psi_1) + \ldots + C_s\cos(\omega_s x + \psi_s)$. C'est le "principe de superposition".

Terminons ce chapitre en donnant un exemple d'équation différentielle (linéaire, à coefficients non constants) se ramenant à une équation différentielle linéaire à coefficients constants.

Equation d'Euler :

$$x^2y'' + axy' + by = g(x)$$

Plaçons nous sur $\mathbf{R}_+^*$ (ou $\mathbf{R}_-^*$) avec $\varepsilon := sgn(x) = \pm 1$ et posons $t := \log|x|$ (donc $x = \varepsilon e^t$) et $z(t) := y(\varepsilon e^t)$; on obtient aisément : $z'(t) = \varepsilon e^t y'(\varepsilon e^t)$ et $z''(t) = e^{2t}y''(\varepsilon e^t) + \varepsilon e^t y'(\varepsilon e^t)$ d'où $z'' + (a-1)z' + bz = g(\varepsilon e^t)$ qui est une équation linéaire à coefficients constants. Les solutions sont du type $z(t) = \lambda e^{\alpha_1 t} + \mu e^{\alpha_2 t} + z_0(t)$ (ou $(\lambda + \mu t)e^{\alpha t} + z_0(t)$) où $z_0(t)$ désigne une solution particulière. On en tire les solutions de l'équation originale sur $\mathbf{R}_+^*$ (ou $\mathbf{R}_-^*$) :

$y(x) = \lambda|x|^{\alpha_1} + \mu|x|^{\alpha_2} + z_0(\log|x|)$ (ou $(\lambda + \mu\log|x|)|x|^{\alpha} + z_0(\log|x|)$

Exemples : L'équation $x^2y'' + xy' + y = x + 1$ conduit à l'équation $z'' + z = \varepsilon e^t + 1$ et donc aux solutions de la forme : $z(t) = \lambda\cos(t) + \mu\sin(t) + \varepsilon e^t/2 + 1$ qui correspond aux solutions $y(x) = \lambda\cos(\log|x|) + \mu\sin(\log|x|) + x/2 + 1$.

L'équation $x^2y'' - 2xy' + 2y = x + 1$ conduit à l'équation $z'' - 3z' + 2z = \varepsilon e^t + 1$ et donc aux solutions de la forme : $z(t) = \lambda e^t + \mu e^{2t}) - te^t + 1/2$ qui correspond aux solutions $y(x) = \lambda|x| + \mu x^2 - (\log|x|)|x| + 1/2$.

L'étude de la possibilité de raccorder deux solutions sur $\mathbf{R}_+^*$ et $\mathbf{R}_-^*$ est laissée en exercice.

Exercice. (Méthode de *variation des constantes* pour les équations d'ordre deux). On souhaite calculer les solutions de

$$y''(x) + a(x)y'(x) + b(x)y(x) = g(x) \qquad (E)$$

lorsqu'on connait une base $y_1(x)$, $y_2(x)$ des solutions de l'équation homogène. En particulier $W(x) := y_1(x)y_2'(x) - y_2(x)y_1'(x)$ est une fonction ne s'annulant pas.

a) Montrer que si C_1 et C_2 sont deux fonctions dérivables vérifiant :

$$\begin{cases} C_1'y_1 + C_2'y_2 &= 0 \\ C_1'y_1' + C_2'y_2' &= g \end{cases}$$

alors la fonction $y(x) = C_1(x)y_1(x) + C_2(x)y_2(x)$ est solution de l'équation (E).

b) En déduire une expression de la solution générale de (E) :

$$y(x) = \left(A_1 - \int_{x_0}^{x} \frac{y_2(t)g(t)}{W(t)}dt\right)y_1(x) + \left(A_2 + \int_{x_0}^{x} \frac{y_1(t)g(t)}{W(t)}dt\right)y_2(x)$$

où A_1 et A_2 sont des constantes

CHAPITRE 21 FONCTIONS DE PLUSIEURS VARIABLES

Nous avons jusqu'à présent étudié les fonctions d'une variable réelle. Si ce thème est très riche et utile, il est cependant très loin de permettre d'aborder la totalité des situations. En effet on peut même dire que la plupart des phénomènes font intervenir plusieurs variables : un problème de thermodynamique (par exemple l'étude d'un gaz) fait intervenir pression, température, volume, un problème d'économie fait souvent intervenir de très nombreux paramètres, etc. L'étude des fonctions de plusieurs variables s'impose donc naturellement, les lois de la nature s'écrivant le plus souvent comme des relations entre plusieurs quantités. Comme on va le voir certains aspects des fonctions à plusieurs variables sont déjà présents dans l'étude des fonctions d'une variable, cependant elles font également apparaître des phénomènes nouveaux. D'un point de vue mathématique une première explication est que la topologie du plan, de l'espace (à trois dimensions) ou plus généralement de $\mathbf{R}^n$ est nettement plus compliquée que celle de $\mathbf{R}$. Ce point est succintement développé, on va surtout s'attacher à traiter l'analogue du calcul différentiel (dérivées partielles, accroissements finis et formule de Taylor, extrema locaux et globaux) pour une fonction de plusieurs variables. La généralisation des équations différentielles "ordinaires" (EDO), c'est-à-dire pour les fonctions d'une variable, s'appelle naturellement "équations aux dérivées partielles" (EDP). Dans cette première approche les EDP sont à peine évoquées, mais on s'est efforcé de mentionner, lors d'exemples et d'exercices, au moins trois des EDP les plus importantes en physique :

$$\Delta f := \frac{\partial^2 f}{\partial x_1^2} + \ldots + \frac{\partial^2 f}{\partial x_n^2} = 0 \qquad \textit{(équation de Laplace)}$$

$$\Delta f - c^2 \frac{\partial^2 f}{\partial t^2} = 0 \qquad \textit{(équation des ondes, D'Alembert)}$$

$$\Delta f - \frac{\partial f}{\partial t} = 0 \qquad \textit{(équation de la chaleur ou de la diffusion, Fourier)}$$

Cependant l'étude générale des solutions de ces équations dépasse largement le niveau de ce cours. Cette partie du cours se termine avec une discussion des méthodes de calcul différentiel visant à déterminer les extrema d'une fonction de plusieurs variables.

1. INTRODUCTION ET REPRÉSENTATION

Soit U une partie de $\mathbf{R}^n$, une fonction $f : U \to \mathbf{R}$ s'appelle une fonction de n variables. Par exemple les fonctions $f(x, y, z) = x^2 + y^2 + z^2$ et $g(x, y, z) = x^2 - y^2 - z^2$ sont deux fonctions (assez simples) de trois variables définies sur $U = \mathbf{R}^3$. La fonction $h(x, y) = \log(x^2 + y + 1)/(x - y)$ est une fonction de deux variables définie sur l'ensemble plus compliqué :

$$U := \{(x, y) \in \mathbf{R}^2 \mid x^2 + y + 1 > 0, \quad \text{et} \quad x \neq y\}.$$

L'analogue du *graphe* d'une fonction d'une variable est facile à définir. Si $f : U \to \mathbf{R}$ est une fonction de n variables, on définit le graphe de f comme l'ensemble :

$$\text{Graphe}_f := \{(x, y) \in U \times \mathbf{R} \mid y = f(x)\}.$$

Cependant on perçoit tout de suite une difficulté pour le dessiner : c'est un sous-ensemble de $\mathbf{R}^{n+1}$; or le tableau ou notre feuille de papier n'a que deux dimensions ! En utilisant les conventions visuelles de la perspective, on peut encore en partie représenter le graphe d'une fonction de deux variables (voir les exemples ci-dessous) mais cela devient impossible pour une fonction de 3 ou plus variables. On devra donc se contenter de représentations partielles. C'est-à-dire que, pour une fonction de 3 variables par exemple $f(x, y, z)$, on peut tracer soit un certain nombre de surfaces (qu'on pourra appeler "tranches")

$$u = f(x, y, c), \qquad \text{avec } c \text{ fixé},$$

soit tracer un certain nombre de *surfaces de niveau* i.e. les surfaces d'équation :

$$f(x, y, z) = a, \qquad \text{avec } a \text{ fixé}.$$

Exemples. Montrons comment tracer le graphe de la fonction $h(x, y) = x^2 - y^2$, c'est-à-dire la surface d'équation $z = x^2 - y^2$. On trace la courbe intersection de la surface avec les plans $x = 0$ et $y = 0$. Ce sont des paraboles dont l'une est renversée. En particulier on remarquera que le point $(0, 0)$ correspond à un minimum le long de la courbe obtenue en fixant $y = 0$ mais correspond à un maximum le long de la courbe obtenue en fixant $x = 0$.

Remarque : cette surface a la forme d'une "selle", ce qui justifiera d'appeler "point-selle" un point d'une surface ayant la même allure (on l'appelle aussi "col").

Considérons la fonction $f(x, y, z) = x^2 + y^2 + xyz$. Le tracé de quelques "tranches" $u = x^2 + y^2 + cxy$ (pour c fixé) donne :

$c = -2$ $\qquad$ $c = 0$ $\qquad$ $c = 2$

On obtient le tracé approximatif pour $c = 0$ en faisant tourner autour de l'axe vertical la parabole contenue dans le plan $y = 0$ d'équation $u = x^2$. Le tracé des courbes de niveau $x^2 + y^2 + xyz = a$ est laissé en exercice (choisir par exemple $c = -3$, $c = 0$ et $c = 3$).

Si l'on choisit maintenant $g(x, y, z) = x^2 - y^2 - z^2$. Le tracé de quelques "tranches" $u = x^2 - y^2 - c^2$ (pour c fixé) donne :

$c = 0$ $\qquad\qquad$ $c = 1$

Le tracé des courbes de niveau est également assez simple dans ce cas : il s'agit d'hyperboloïdes de révolution $x^2 - y^2 - z^2 = a$:

$a = 1$ $\qquad$ $a = 0$ $\qquad$ $a = -1$

Remarquons au passage qu'on a utilisé le mot "surface" en un sens que nous espérons être clair mais que nous n'avons pas défini.

Nous allons maintenant étudier les questions de continuité et de dérivabilité, mais pour cela il faut faire un détour et étudier un peu plus le plan et plus généralement l'espace $\mathbf{R}^n$.

2. UN PEU DE TOPOLOGIE DE $\mathbf{R}^n$

La notion de continuité sur $\mathbf{R}$ a été étudiée à partir de la distance entre deux réels, qui permet de définir la notion de "tendre vers" ou de "limite"; il existe bien sûr une notion analogue de distance dans $\mathbf{R}^n$ que nous rappelons. Notons $(x \cdot y) = x_1y_1 + \ldots + x_ny_n$ le produit scalaire de deux vecteurs de $\mathbf{R}^n$. Il est utile de se souvenir que $(x \cdot y) = 0$ équivaut à dire que les vecteurs x et y sont orthogonaux.

Définition: On appelle *norme euclidienne* d'un vecteur $x = (x_1, \ldots, x_n)$ de $\mathbf{R}^n$ le nombre :

$$||x|| := \sqrt{(x \cdot x)} = \sqrt{x_1^2 + \ldots + x_n^2}.$$

On définit la *distance euclidienne* entre deux vecteurs $x = (x_1, \ldots, x_n)$ et $y = (y_1, \ldots, y_n)$ de $\mathbf{R}^n$ par la formule :

$$\text{distance}\,(x, y) := ||x - y|| = \sqrt{(x_1 - y_1)^2 + \ldots + (x_n - y_n)^2}.$$

L'analogue d'un intervalle (ouvert) de $\mathbf{R}$ qui s'écrit $]a - r, a + r[= \{x \in \mathbf{R} \mid |x - a| < r\}$, est une *boule* (ouverte) de $\mathbf{R}^n$: on appelle *boule (ouverte) de centre* $a \in \mathbf{R}^n$ *et de rayon* r l'ensemble

$$B(a, r) := \{x \in \mathbf{R}^n \mid ||x - a|| < r\}.$$

Les principales propriétés de la norme euclidienne sont résumées dans la proposition suivante (les trois premières propriétés caractérisant les normes) vue au chapitre 11 en dimension 2 et 3.

PROPOSITION: *La norme euclidienne sur* $\mathbf{R}^n$ *vérifie pour tout* $x = (x_1, \ldots, x_n) \in \mathbf{R}^n$

(i) *On a* $||x|| \geq 0$ *avec égalité si et seulement si* $x = 0$;

(ii) *Pour* $a \in \mathbf{R}$, *on a* $||ax|| = |a|||x||$;

(iii) *(inégalité triangulaire)* $||x + y|| \leq ||x|| + ||y||$. *De plus on a l'inégalité de Cauchy-Schwarz :*

$$|(x \cdot y)| := |x_1 y_1 + \ldots + x_n y_n| \leq ||x||\,||y||.$$

(iv) $\text{distance}\,(x, y) \geq 0$ *(avec égalité seulement si* $x = y$*)*

(v) $\text{distance}\,(x, y) \leq \text{distance}\,(x, z) + \text{distance}\,(z, y)$.

Remarque. Dans le plan, on sait que, si θ est l'angle entre les deux vecteurs x et y, alors $(x \cdot y) = \cos\theta\,||x||\,||y||$; l'inégalité de Cauchy-Schwarz équivaut donc dans ce cas à $|\cos\theta| \leq 1$.

Remarque. Il est quelquefois commode d'utiliser d'autres normes comme par exemple :

$$||x||_1 = |x_1| + \ldots + |x_n|, \qquad \text{ou encore} \qquad ||x||' = \max_{i=1 \text{ à } n} |x_i|.$$

Les inégalités suivantes sont assez faciles à établir et sont laissées en exercice :

$$||x|| \leq ||x||_1 \leq n||x||, \qquad \text{et} \qquad \frac{||x||}{\sqrt{n}} \leq ||x||' \leq ||x||.$$

Elles permettent de voir que $||x||$ tend vers zéro si et seulement si $||x||_1$ (ou $||x||'$) tend vers zéro. Par contre les boules n'ont pas la même forme :

Boule $||x|| \leq 1$ dans $\mathbf{R}^3$, Boule $||x||_1 \leq 1$ dans $\mathbf{R}^3$, Boule $||x||' \leq 1$ dans $\mathbf{R}^3$.

On peut maintenant définir la notion de limite et continuité.

Définition: Soit U contenant une boule centrée en x_0 (dans $\mathbf{R}^n$) et soit $f : U \to \mathbf{R}$ une fonction de n variables. On dit que $f(x)$ tend vers ℓ quand x tend vers x_0 si on a

$$\forall \epsilon > 0, \exists \delta > 0, 0 < ||x - x_0|| \leq \delta \Rightarrow |f(x) - \ell| \leq \epsilon.$$

Si de plus $\ell = f(x_0)$, on dit que f est continue au point x_0. Si f est continue en tout point $x \in U$, on dit que f est continue sur U.

Les propriétés "usuelles" suivantes restent vraies (et se démontrent comme pour les fonctions d'une variables) :

(i) La somme et le produit de fonctions continues sont encore continus. Les polynômes (à plusieurs variables) sont des fonctions continues.

(ii) La composée de fonctions continues à plusieurs variables est encore continue. Plus précisément si $f(u_1, \ldots, u_n)$ est une fonction continue à n variables et si $g_i(x_1, \ldots, x_m)$ est une fonction continue à m variables pour $1 \leq i \leq n$, alors la fonction

$$F(x_1, \ldots, x_m) := f(g_1(x_1, \ldots, x_m), \ldots, g_n(x_1, \ldots, x_m))$$

est une fonction continue à m variables.

Exemple. La fonction $h(x, y) = \log(x^2+y+1)/(x-y)$ est continue sur son domaine de définition. En effet on peut l'écrire en composant polynômes, fonction inverse (c'est-à-dire $f(x) = 1/x$) et fonction logarithme.

Etudions deux exemples où la continuité est plus délicate et où il nous faut revenir à la définition. Considérons les deux fonctions

$$f(x,y) = \begin{cases} \frac{xy}{x^2+y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{cases} \qquad g(x,y) = \begin{cases} \frac{x^4+y^4}{x^2+y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{cases}$$

Par les mêmes arguments que précédemment, on voit que f et g sont continues en tout point de $\mathbf{R}^2 \setminus \{(0,0)\}$. La question de la continuité au point $(0,0)$ est plus délicate. Montrons que g est continue au point $(0,0)$ alors que f n'est pas continue au point $(0,0)$.

Pour montrer la continuité de g en $(0,0)$ on doit montrer que $g(x,y)$ devient arbitrairement petit quand la norme de (x,y) devient petite. Pour voir cela on peut majorer $x^4 + y^4 \leq (x^2 + y^2)^2$ et donc écrire

$$|g(x,y)| \leq x^2 + y^2$$

qui montre que $g(x,y)$ tend vers 0 quand (x,y) tend vers 0. En effet dès que $||(x,y)|| = \sqrt{x^2+y^2} \leq \sqrt{\epsilon}$ on aura $|g(x,y)| \leq \epsilon$.

Pour montrer que f n'est pas continue, on peut observer que, pour a fixé,

$$\lim_{x \to 0} f(x, ax) = \lim_{x \to 0} \frac{ax^2}{(1+a^2)x^2} = \frac{a}{1+a^2}$$

existe mais varie avec a, et en particulier n'est pas toujours égale à $f(0,0)$.

Un peu de vocabulaire. Nous utiliserons le minimum de topologie de $\mathbf{R}^n$ et nous concentrerons sur le calcul différentiel mais nous définissons néanmoins les notions suivantes abordées dans de nombreux ouvrages. Si $a \in A$ et $A \subset \mathbf{R}^n$, on dit que a est un *point intérieur* de A s'il existe une boule $B(a,r)$ contenue dans A; l'ensemble A est un *ouvert* si tout point de A est intérieur (il revient au même de dire que A est réunion de boules ouvertes). Un élément a est dans la *frontière* (ou *bord*) de A si toute boule $B(a,r)$ contient un élément de A et un élément hors de A. Un ensemble A est *fermé* si c'est le complémentaire d'un ouvert (il revient au même de dire que A contient sa frontière). L'ensemble A est *borné* s'il est contenu dans une boule; il est dit *compact* s'il est fermé et borné. Définissons enfin la notion d'ensemble n'ayant "qu'un seul morceau" : un ensemble ouvert A est *connexe* (ou "en un seul morceau") si on ne peut pas l'écrire comme réunion disjointe de deux ouverts; cette propriété équivaut au fait que deux points de A peuvent être joints, à l'intérieur de A par une ligne polygonale (une suite finie de segments).
Exemples et commentaires. Un ouvert de $\mathbf{R}$ est une réunion d'intervalles ouverts ; l'analogue d'un intervalle ouvert est un ouvert connexe. Les propriétés suivantes des fonctions continues se transposent :

(i) Une fonction continue sur un ensemble compact (fermé, borné) est bornée et atteint son maximum et son minimum.

(ii) Une fonction continue sur un ensemble connexe satisfait la propriété des valeurs intermédiaires.

3. DÉRIVÉES PARTIELLES

La définition des dérivées partielles est assez simple : on fixe toute les variables sauf une et on calcule, si elle existe, la dérivée de la fonction d'une variable ainsi obtenue. Plus précisément :

Définition: Soit $f : U \to \mathbf{R}$ une fonction de n variables $x = (x_1, \ldots, x_n)$ et soit $a = (a_1, \ldots, a_n)$ un point de U tel que U contienne une boule centré en a, on définit la *dérivée partielle de f* par rapport à la variable x_i en a comme la limite suivante (si elle existe) :

$$\frac{\partial f}{\partial x_i}(a) = \lim_{h \to 0} \frac{f(a_1, \ldots, a_{i-1}, a_i + h, a_{i+1}, \ldots, a_n) - f(a_1, \ldots, a_n)}{h}.$$

Si la fonction admet des dérivées partielles en tout point de U, on peut considérer la fonction $\frac{\partial f}{\partial x_i}(x)$ et étudier sa dérivabilité ; sa dérivée partielle par rapport à x_j au point a ne note

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(a) = \lim_{h \to 0} \frac{\frac{\partial f}{\partial x_i}(a_1, \ldots, a_{j-1}, a_j + h, a_{j+1}, \ldots, a_n) - \frac{\partial f}{\partial x_i}(a_1, \ldots, a_n)}{h}.$$

On peut bien sûr définir des dérivées partielles de n'importe quelle ordre mais nous n'utiliserons pas les dérivées d'ordre supérieur à 3. Nous utiliserons toujours la notation $\frac{\partial f}{\partial x_i}$ mais signalons que les notations f'_{x_i} ou encore $D_i f$ sont également utilisées.

On a vu que, pour une fonction d'une variable, l'existence de la dérivée en un point équivaut à l'existence d'un DL à l'ordre 1 en ce point. La situation est un peu plus compliqué en plusieurs variables. Ainsi la fonction $f(x,y) = xy/(x^2+y^2)$ vérifie $\frac{\partial f}{\partial x}(0,0) = \frac{\partial f}{\partial x}(0,0) = 0$ sans que la fonction soit même continue!

Définition: Une fonction f est *différentiable* en $a \in \mathbf{R}^n$ si elle admet un DL à l'ordre 1 au voisinage de ce point, c'est-à-dire si il existe $\ell_i \in \mathbf{R}$ tels que :

$$f(a+h) = f(a) + \ell_1 h_1 + \ldots + \ell_n h_n + \epsilon(h)||h||,$$

avec $\lim_{h\to 0} \epsilon(h) = 0$. La forme linéaire $g(h) := \ell_1 h_1 + \ldots + \ell_n h_n$ s'appelle la *différentielle* de f en a.

On voit facilement que, si f est différentiable, elle admet des dérivées partielles et on a $\ell_i = \frac{\partial f}{\partial x_i}(a)$. On verra plus loin que, si f admet des dérivées partielles continues, alors elle est différentiable.

Remarquons qu'il est dénué de sens de demander si une fonction de deux ou plusieurs variables est croissante ou non (il n'y a pas d'ordre naturel sur les couples de réels). On peut néanmoins se donner une direction (i.e. restreindre la fonction à une droite) et regarder si la fonction est croissante ou non dans cette direction. Ceci suggère d'introduire la notation suivante.

Définition: Soit f une fonction admettant des dérivées partielles en a, on appelle *gradient* de f en a le vecteur de $\mathbf{R}^n$ donné par :

$$\vec{\text{grad}} f(a) := \left(\frac{\partial f}{\partial x_1}(a), \ldots, \frac{\partial f}{\partial x_n}(a) \right).$$

Si f est différentiable au point a on peut donc écrire :

$$f(a+h) = f(a) + \left(\vec{\text{grad}} f(a) \cdot h \right) + \epsilon(h)||h||,$$

On constate alors que l'hyperplan affine d'équation $z = f(a) + \left(\vec{\text{grad}} f(a) \cdot (x-a) \right)$ est l'hyperplan approchant le mieux l'hypersurface $z = f(x)$.

Si par exemple, $f(x,y,z)$ est une fonction différentiable de trois variables, le plan tangent à la surface de niveau $f(x,y,z) = \lambda$ en (x_0,y_0,z_0) a pour équation :

$$\frac{\partial f}{\partial x}(x_0,y_0,z_0)(x-x_0) + \frac{\partial f}{\partial y}(x_0,y_0,z_0)(y-y_0) + \frac{\partial f}{\partial z}(x_0,y_0,z_0)(z-z_0) = 0.$$

En particulier, on constate que le vecteur gradient $\vec{\text{grad}} f(x_0, y_0, z_0)$ est orthogonal au plan tangent de la surface de niveau.

Exemple. Considérons la fonction $f(x, y) = x\cos(x^2 + yx)$. Ses dérivées partielles se calculent avec les règles de calcul de dérivées ordinaires (à une variables) en prenant soin de garder fixe une des deux variables; on obtient

$$\begin{aligned}
\frac{\partial f}{\partial x}(x, y) &= \cos(x^2 + yx) - x(2x + y)\sin(x^2 + yx) \\
\frac{\partial f}{\partial y}(x, y) &= -x^2 \sin(x^2 + yx) \\
\frac{\partial^2 f}{\partial x^2}(x, y) &= -(6x + 2y)\sin(x^2 + yx) - x(2x + y)^2 \cos(x^2 + yx) \\
\frac{\partial^2 f}{\partial x \partial y}(x, y) &= -2x \sin(x^2 + yx) - x^2(2x + y)\cos(x^2 + yx) \\
\frac{\partial^2 f}{\partial y^2}(x, y) &= -x^3 \cos(x^2 + yx) \\
\frac{\partial^2 f}{\partial y \partial x}(x, y) &= -2x \sin(x^2 + yx) - (2x^2 + xy)x\cos(x^2 + yx)
\end{aligned}$$

On remarque que, sur cet exemple on a $\frac{\partial^2 f}{\partial y \partial x}(x, y) = \frac{\partial^2 f}{\partial x \partial y}(x, y)$; on va voir que ce n'est pas une coïncidence.

THÉORÈME: (Théorème de Schwarz) *soit f une fonction de deux variables, deux fois continûment dérivable, alors on a :*

$$\frac{\partial^2 f}{\partial y \partial x}(a, b) = \frac{\partial^2 f}{\partial x \partial y}(a, b).$$

Démonstration: Posons $F(x) := f(x, b + k) - f(x, b)$ et $G(y) := f(a + h) - f(a, y)$, alors on a

$$A := f(a+h, b+k) - f(a+h, b) - f(a, b+k) + f(a, b) = F(a+h) - F(a) = G(b+k) - G(b).$$

On applique une première fois le théorème des accroissements finis (en une variable) en remarquant que $F'(x) = \frac{\partial f}{\partial x}(x, b + k) - \frac{\partial f}{\partial x}(x, b)$ donc

$$A = hF'(a + \theta_1 h) = h\left\{\frac{\partial f}{\partial x}(a + \theta_1 h, b + k) - \frac{\partial f}{\partial x}(a + \theta_1 h, b)\right\}.$$

Une deuxième application du théorème des accroissements finis (en une variable) à la fonction $H(y) = \frac{\partial f}{\partial x}(a + \theta_1 h, y)$ qui a pour dérivée $H'(y) = \frac{\partial^2 f}{\partial y \partial x}(a + \theta_1 h, y)$ donne

$$A = hkH'(b + \theta_2 k) = hk \frac{\partial^2 f}{\partial y \partial x}(a + \theta_1 h, b + \theta_2 k).$$

Un raisonnement symétrique à partir de la fonction G donnerait

$$A = kh \frac{\partial^2 f}{\partial x \partial y}(a + \theta_3 h, b + \theta_4 k).$$

On obtient ainsi l'existence de $\theta_1, \theta_2, \theta_3, \theta_4 \in]0, 1[$ (dépendant de h et k) tels que

$$\frac{\partial^2 f}{\partial y \partial x}(a + \theta_1 h, b + \theta_2 k) = \frac{\partial^2 f}{\partial x \partial y}(a + \theta_3 h, b + \theta_4 k).$$

En faisant tendre h et k vers zéro et en utilisant la continuité des deux dérivées secondes on obtient alors le résultat escompté :

$$\frac{\partial^2 f}{\partial y \partial x}(a, b) = \frac{\partial^2 f}{\partial x \partial y}(a, b).$$

□

Exercices. 1) On définit une fonction de $n + 1$ variables par la formule

$$F(x_1, \ldots, x_n, t) := t^{-n/2} \exp(-||x||^2/4t).$$

Déterminer le domaine de définition de F, calculer les dérivées partielles jusqu'au second ordre et montrer que F vérifie l'équation de la chaleur (ou équation de diffusion ou équation de Fourier) :

$$\Delta F - \frac{\partial F}{\partial t} = \frac{\partial^2 F}{\partial x_1^2} + \ldots + \frac{\partial^2 F}{\partial x_n^2} - \frac{\partial F}{\partial t} = 0.$$

2) Soit f, g deux fonctions d'une variable deux fois dérivables. Montrer que la fonction $F(x, t) = f(cx + t) + g(cx - t)$ vérifie l'équation aux ondes

$$\frac{\partial^2 F}{\partial x^2} - c^2 \frac{\partial^2 F}{\partial t^2} = 0.$$

Quelle interprétation physique peut-on donner à ce fait ?
3) Soient f, g deux fonctions dérivables de n variables, montrer les formules suivantes :

$$\begin{aligned}
\vec{\text{grad}}(f + g)(a) &= \vec{\text{grad}} f(a) + \vec{\text{grad}} g(a) \\
\vec{\text{grad}}(fg)(a) &= g(a)\vec{\text{grad}} f(a) + f(a)\vec{\text{grad}} g(a) \\
\vec{\text{grad}}\left(\frac{f}{g}\right)(a) &= \frac{1}{g^2(a)}\left(g(a)\vec{\text{grad}} f(a) - f(a)\vec{\text{grad}} g(a)\right).
\end{aligned}$$

Donnons maintenant en trois variables une des principale règles de calcul des dérivées partielles de fonctions composées (il n'y a pas de difficultés à faire ces calculs avec un autre nombre de variables).

THÉORÈME: *Soit $f(u,v,w)$ une fonction de trois variables ; soient $u(t), v(t), w(t)$ trois fonctions d'une variable et posons $F(t) = f(u(t), v(t), w(t))$. Si u, v, w sont continûment dérivables et f également, alors F est continûment dérivable et :*

$$F'(t) = u'(t)\frac{\partial f}{\partial x}(u(t),v(t),w(t)) + v'(t)\frac{\partial f}{\partial y}(u(t),v(t),w(t)) + w'(t)\frac{\partial f}{\partial z}(u(t),v(t),w(t)).$$

Démonstration: Posont $a(t) = (u(t), v(t), w(t)$. Comme u, v, w sont continûment dérivables on a $u(t+h) - u(t) = u'(t+\theta_1 h)h = u'(t)h + \epsilon(h)h$ (et de même pour v et w). On peut écrire

$$f(u(t+h),b,c) - f(u(t),b,c) = \frac{\partial f}{\partial x}(u(t)+\epsilon(h),b,c)u'(t+\theta h)h.$$

D'où l'on tire

$$\begin{aligned}
F(t+h) - F(t) =& f(u(t+h),v(t+h),w(t+h)) - f(u(t),v(t),w(t)) \\
=& f(u(t+h),v(t+h),w(t+h)) - f(u(t),v(t+h),w(t+h)) \\
&+ f(u(t),v(t+h),w(t+h)) - f(u(t),v(t),w(t+h)) \\
&+ f(u(t),v(t),w(t+h)) - -f(u(t),v(t),w(t)) \\
&= \frac{\partial f}{\partial x}(u(t)+\epsilon_1(h),v(t+h),w(t+h))u'(t+\theta_1 h)h \\
&+ \frac{\partial f}{\partial y}(u(t),v(t)+\epsilon_2(h),w(t+h))v'(t+\theta_2 h)h \\
&+ \frac{\partial f}{\partial z}(u(t),v(t),w(t)+\epsilon_3(h))w'(t+\theta_3 h)h \\
&= \left(u'(t)\frac{\partial f}{\partial x}(a(t)) + v'(t)\frac{\partial f}{\partial y}(a(t)) + w'(t)\frac{\partial f}{\partial z}(a(t))\right)h + \epsilon(h)h.
\end{aligned}$$

d'où le résultat. □

Dans le cas où $F(x_1,\ldots,x_n) = f(u_1(x_1,\ldots,x_n),\ldots,u_m(x_1,\ldots,x_n))$, la "formule générale" s'écrirait, :

$$\frac{\partial F}{\partial x_i}(x_1,\ldots,x_n) = \sum_{j=1}^{m}\frac{\partial u_j}{\partial x_i}(x_1,\ldots,x_n)\frac{\partial f}{\partial y_j}(u_1(x_1,\ldots,x_n),\ldots,u_m(x_1,\ldots,x_n)).$$

APPLICATION : Calcul du Laplacien en coordonnées polaires et sphériques.

Les coordonnées polaires dans le plan peuvent être définies par les formules $(x, y) = (r\cos(\theta), r\sin(\theta)$. Si $f(x, y)$ est une fonction de deux variables et si l'on pose $g(r, \theta) := f(r\cos(\theta), r\sin(\theta))$, alors le théorème précédent permet de calculer

$$\frac{\partial g}{\partial r} = \cos(\theta)\frac{\partial f}{\partial x} + \sin(\theta)\frac{\partial f}{\partial y} \qquad \text{et} \qquad \frac{\partial g}{\partial \theta} = -r\sin(\theta)\frac{\partial f}{\partial x} + r\cos(\theta)\frac{\partial f}{\partial y}$$
$$\frac{\partial^2 g}{\partial r^2} = \cos^2(\theta)\frac{\partial^2 f}{\partial x^2} + 2\cos(\theta)\sin(\theta)\frac{\partial^2 f}{\partial x \partial y} + \sin^2(\theta)\frac{\partial^2 f}{\partial y^2}$$
$$\frac{\partial^2 g}{\partial \theta^2} = r^2\sin^2(\theta)\frac{\partial^2 f}{\partial x^2} - 2r^2\cos(\theta)\sin(\theta)\frac{\partial^2 f}{\partial x \partial y} + r^2\cos^2(\theta)\frac{\partial^2 f}{\partial y^2} - r\cos(\theta)\frac{\partial f}{\partial x} - r\sin(\theta)\frac{\partial f}{\partial y}$$

et de vérifier :

$$\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = \frac{\partial^2 g}{\partial r^2} + \frac{1}{r}\frac{\partial g}{\partial r} + \frac{1}{r^2}\frac{\partial^2 g}{\partial \theta^2}.$$

En particulier, une fonction radiale (c'est-à-dire indépendante de θ) et harmonique (c'est-à-dire avec laplacien nul) vérifie donc l'équation différentielle ordinaire

$$g''(r) + r^{-1}g'(r) = 0,$$

ce qui entraîne $g'(r) = C_1/r$ puis $g(r) = C_1 \log r + C_2$.

Exercice. On définit également les coordonnées sphériques dans $\mathbf{R}^3$ par les formules

$$(x, y, z) = (r\cos(\phi)\cos(\theta), r\cos(\phi)\sin(\theta), r\sin(\phi))$$

correspondant géométriquement à la figure ci-dessous :

Coordonnées sphériques (r, θ, ϕ).

Si $f(x, y, z)$ est une fonction de trois variables et si l'on pose

$$g(r, \theta, \phi) = f(r\cos(\phi)\cos(\theta), r\cos(\phi)\sin(\theta), r\sin(\phi)),$$

vérifier la formule

$$\begin{aligned}\Delta f &= \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} \\ &= \frac{\partial^2 g}{\partial r^2} + \frac{2}{r}\frac{\partial g}{\partial r} + \frac{1}{r^2\cos^2(\phi)}\frac{\partial^2 g}{\partial \theta^2} + \frac{1}{r^2}\frac{\partial^2 g}{\partial \phi^2} - \frac{\mathrm{tg}(\phi)}{r^2}\frac{\partial g}{\partial \phi}.\end{aligned}$$

En déduire l'expression des fonctions radiales de laplacien nul sur $\mathbf{R}^3 \setminus \{0\}$ (on trouve $g(r) = C_1 + C_2/r$).

Plus généralement, posons $r := \sqrt{x_1^2 + \ldots + x_n^2}$ et $f(x_1, \ldots, x_n) := g(r)$. Montrer que

$$\Delta f(x_1, \ldots, x_n) = g''(r) + \frac{n-1}{r} g'(r)$$

et en déduire que les fonctions radiales harmoniques en $n \geq 3$ variables ont pour expression :

$$f(x_1, \ldots, x_n) = C_1 + C_2 r^{-n+2} = C_1 + \frac{C_2}{(x_1^2 + \ldots + x_n^2)^{(n-2)/2}}.$$

THÉORÈME: (Théorème des accroissement finis) *Soit $f : U \to \mathbf{R}$ une fonction continûment dérivable ; soient $x = (x_1, \ldots, x_n)$ et $y = (y_1, \ldots, y_n)$ deux points de U tels que le segment joignant x à y soit inclus dans U alors il existe un point c de ce segment tel que :*

$$f(x) - f(y) = \vec{\text{grad}} f(c) \cdot (x - y) = \frac{\partial f}{\partial x_1}(c)(x_1 - y_1) + \ldots + \frac{\partial f}{\partial x_n}(c)(x_n - y_n).$$

En particulier on l'inegalité suivante

$$|f(x) - f(y)| \leq M||x - y||$$

où M est un majorant de $||\vec{\text{grad}} f||$ sur le segment joignant x à y.

Démonstration: Posons $g(t) = f(y + t(x - y))$, alors g est une fonction d'une variable réelle t dont la dérivée peut se calculer comme fonction composée :

$$g'(t) = \frac{\partial f}{\partial x_1}(y + t(x - y))(x_1 - y_1) + \ldots + \frac{\partial f}{\partial x_n}(y + t(x - y))(x_n - y_n).$$

La dérivée de g est continue et on peut lui appliquer le théorème des accroissements finis (en une variable) et conclure qu'il existe $t_0 \in [0, 1]$ tel que $g(1) = g(0) + g'(t_0)$, ce qui donne la première formule avec $c = x + t_0(x - y)$. Pour obtenir la majoration, on applique l'inégalité de Cauchy-Schwarz : $\left|\left(\vec{\text{grad}} f(c) \cdot (x - y)\right)\right| \leq ||\vec{\text{grad}} f(c)||\, ||x - y||$. □

Remarques. 1) L'hypothèse que le segment $[x, y]$ soit contenue dans U ne peut pas être enlevée (voir par exemple l'exercice à la fin de ce chapitre). 2) Cet énoncé permet de voir qu'une fonction admettant des dérivées partielles continues est différentiable.

COROLLAIRE: *Soit $f : U \to \mathbf{R}$ une fonction continûment dérivable. Soit $a \in U$ tel que U contienne une boule centrée en a. Soit $\vec{v} \in \mathbf{R}^n$, alors la fonction $g(t) = f(x + t\vec{v})$ est croissante en $t = 0$ si et seulement si*

$$\left(\vec{\text{grad}} f(x) \cdot \vec{v}\right) \geq 0.$$

Démonstration: Soit h un vecteur de norme plus petit que le rayon de la boule de l'hypothèse, le théorème nous garantit l'existence d'un réel $\theta \in [0,1]$ tel que

$$f(a+h) - f(a) = \vec{\text{grad}} f(a+\theta h) \cdot h = \frac{\partial f}{\partial x_1}(a+\theta h)h_1 + \ldots + \frac{\partial f}{\partial x_n}(a+\theta h)h_n.$$

En remplaçant h par $t\vec{v}$ (pour t assez petit) on obtient

$$f(a+t\vec{v}) - f(a) = \vec{\text{grad}} f(a+\theta t\vec{v}) \cdot t\vec{v} = t\left((\vec{\text{grad}} f(a) \cdot \vec{v}) + \epsilon(t)\right)$$

d'où le résultat. □

COROLLAIRE: *Soit $f : U \to \mathbf{R}$ une fonction continûment dérivable. Soit $a \in U$ tel que U contienne une boule centrée en a et sur laquelle f admet un maximum ou un minimum en a (c'est-à-dire que a est un extremum local pour la fonction f), alors* $\vec{\text{grad}} f(a) = 0$.

Démonstration: En appliquant le corollaire précédent à $\vec{v}$ et $-\vec{v}$ on voit que le produit scalaire $\vec{\text{grad}} f(a) \cdot \vec{v}$ est nul. Le seul vecteur orthogonal à tous les vecteurs de $\mathbf{R}^n$ est le vecteur nul. □

Définition: Un point a où $\vec{\text{grad}} f(a) = 0$ sera appelé *point critique* de f.

Remarque. Une fonction différentiable sur un fermé borné (par exemple une boule fermée $||x-a|| \leq r$) atteint donc son maximum soit en un point du bord, soit en un point critique de l'intérieur (par exemple de la boule ouverte $||x-a|| < r$).

COROLLAIRE: *Soit $f : U \to \mathbf{R}$ une fonction continûment dérivable telle que pour tout $x \in U$ on ait* $\vec{\text{grad}} f(x) = 0$. *Supposons de plus que deux points de U puissent être joint par une ligne polygonale dans U, alors la fonction f est constante sur U.*

Démonstration: Il suffit de démontrer que f est constante sur un segment contenu dans U et ceci est clair à partir du théorème précédent. □

Remarque. L'hypothèse sur U est nécessaire comme le montre l'exemple suivant. Considérons la fonction

$$f(x,y) = \text{Arctg}(x/y) + \text{Arctg}(y/x).$$

On calcule aisément et on vérifie que $\frac{\partial f}{\partial x}(x,y) = \frac{\partial f}{\partial y}(x,y) = 0$; pourtant $f(1,1) = \pi$ et $f(-1,1) = -\pi$. Le problème vient de ce que l'ouvert de définition $U = \{(x,y) \in \mathbf{R}^2 \mid x \neq 0 \quad \text{et} \quad y \neq 0\}$ est constitué de 4 morceaux. Sur le quadrant supérieur droit $U_1 = \{(x,y) \in \mathbf{R}^2 \mid x > 0 \quad \text{et} \quad y > 0\}$ on a $f(x,y) = \pi$; sur le quadrant supérieur gauche $U_2 = \{(x,y) \in \mathbf{R}^2 \mid x < 0 \quad \text{et} \quad y > 0\}$ on a $f(x,y) = -\pi$; sur le quadrant inférieur gauche $U_3 = \{(x,y) \in \mathbf{R}^2 \mid x < 0 \quad \text{et} \quad y < 0\}$ on a $f(x,y) = \pi$ et sur le quadrant inférieur droit $U_4 = \{(x,y) \in \mathbf{R}^2 \mid x > 0 \quad \text{et} \quad y < 0\}$ on a $f(x,y) = -\pi$.

Nous avons vu qu'une fonction dérivable d'une variable ne peut avoir un extremum local qu'en un point ou sa dérivée s'annule ou en une extrêmité de l'intervalle ; de plus la considération de la valeur de la dérivée seconde (et d'ordre supérieur) permet en général de préciser si le point où la dérivée s'annule est bien un extremum ou non. Le point clef de la preuve était le développement de Taylor

$$f(x_0+h) = f(x_0) + f'(x_0)h + f''(x_0)\frac{h^2}{2} + \epsilon(h)h^2$$

qui permet de dire que, si $f''(x_0) \neq 0$, la fonction f se comporte au voisinage de x_0 comme le polynôme $f(x_0)+f'(x_0)h+f''(x_0)h^2/2$. Nous allons voir que l'analogue de ces résultats persiste en plusieurs variables. Tout d'abord donnons l'analogue de la formule de Taylor à l'ordre 2. Pour simplifier les notations et les calculs, nous écrirons la plupart des énoncés de ce paragraphe en 2 variables mais l'essentiel s'adapte sans difficulté au cas de n variables.

THÉORÈME: (formule de Taylor à l'ordre deux) *Soit $f : U \to \mathbf{R}$ une fonction de deux variables deux fois continûment dérivable au voisinage du point $x = (x_0, y_0) \in U$.*

$$\begin{aligned} f(x_0+h, y_0+k) =& f(x_0,y_0) + \frac{\partial f}{\partial x}(x)h + \frac{\partial f}{\partial y}(x)k \\ &+ \frac{1}{2!}\left(\frac{\partial^2 f}{\partial x^2}(x)h^2 + 2\frac{\partial^2 f}{\partial x \partial y}(x)hk + \frac{\partial^2 f}{\partial y^2}(x)k^2\right) + \epsilon(h,k)(h^2+k^2) \end{aligned}$$

où $\epsilon(h,k)$ est une fonction qui tend vers zéro quand (h,k) tend vers $(0,0)$.

Bien sûr 2! est égal à 2 et nous l'avons écrit ainsi uniquement pour suggérer que la formule à l'ordre 3 ferait intervenir un coefficient 1/3!, etc.

Démonstration: Comme pour la preuve du théorème des accroissement finis, on pose $g(t) = f(x_0+th, y_0+tk)$. Au vu des hypothèses, la fonction g est deux fois dérivable et :

$$\begin{aligned} g'(t) &= \frac{\partial f}{\partial x}(x_0+th, y_0+tk)h + \frac{\partial f}{\partial y}(x_0+th, y_0+tk)k \\ g''(t) &= \frac{\partial^2 f}{\partial x^2}(x_0+th, y_0+tk)h^2 + 2\frac{\partial^2 f}{\partial x \partial y}(x_0+th, y_0+tk)hk + \frac{\partial^2 f}{\partial y^2}(x_0+th, y_0+tk)k^2 \end{aligned}$$

et, en appliquant la formule de Taylor à l'ordre 2 à g on obtient

$$g(1) = g(0) + g'(0) + \frac{1}{2!}g''(t_0), \qquad \text{pour un } t_0 \in [0,1].$$

Il suffit maintenant d'observer que, comme les dérivées partielles secondes sont continues, on a par exemple

$$\frac{\partial^2 f}{\partial x \partial y}(x_0+t_0h, y_0+t_0k) = \frac{\partial^2 f}{\partial x \partial y}(x_0,y_0) + \epsilon(h,k)$$

d'où le théorème. □

Ce théorème permet essentiellement de ramener l'étude locale en (x_0, y_0) à celle de la fonction quadratique

$$g(h,k) := \frac{\partial^2 f}{\partial x^2}(x)h^2 + 2\frac{\partial^2 f}{\partial x \partial y}(x)hk + \frac{\partial^2 f}{\partial y^2}(x)k^2.$$

On peut étudier ce polynôme homogène de deux variables essentiellement comme un polynôme de degré 2 en 1 variable (ici le cas de deux variables est un peu plus simple) de la façon suivante. Posons

$$a := \frac{\partial^2 f}{\partial x^2}(x), \quad b := \frac{\partial^2 f}{\partial x \partial y}(x), \quad \text{et} \quad c := \frac{\partial^2 f}{\partial y^2}(x)$$

de sorte qu'on veut étudier $g(h,k) = ah^2 + 2bhk + ck^2$. Si $a \neq 0$ on peut remarquer que

$$g(h,k) = ah^2 + 2bhk + ck^2 = a\left\{(h + \frac{b}{a}k)^2 + \frac{ac - b^2}{a^2}k^2\right\}.$$

On voit immédiatement qu'il faut distinguer suivant le signe de $\Delta = b^2 - ac$; on obtient :

(i) Si $\Delta < 0$ alors $g(h,k)$ ne s'annule que pour $(h,k) = (0,0)$ et si $(h,k) \neq (0,0)$ alors $g(h,k)$ est du même signe que a. Ainsi on obtient un maximum (resp. un minimum) en $(h,k) = (0,0)$ si $a < 0$ (resp. si $a > 0$).

(ii) Si $\Delta > 0$ alors $g(h,k)$ s'annule le long de deux droites et prend des valeurs positives et négatives de part et d'autre de ces deux droites. Il n'y a pas d'extremum.

(iii) Si $\Delta = 0$ alors $g(h,k)$ s'annule le long d'une droite et chaque point de cette droite est un extremum pour la fonction $g(h,k)$. On dit que la forme quadratique g est dégénérée.

Traçons le graphe $z = g(h,k)$ dans quelques exemples

(i), $g(h,k) = h^2 + hk + k^2$ (ii), $g(h,k) = h^2 + 3hk + 2k^2$ (i), $g(h,k) = (h-k)^2$

Nous allons démontrer "à la main" une propriété générale des formes quadratiques qui nous sera utile pour le théorème suivant; le lemme dit qu'une forme quadratique positive se comporte comme $x^2 + y^2$.

LEMME: *Soit $Q(x,y) = ax^2 + 2bxy + cy^2$ une forme quadratique avec $\Delta := b^2 - ac \neq 0$.*

(i) Si $\Delta < 0$, alors il existe une constante $C > 0$ (ne dépendant que de a, b, c) telle que, si $a > 0$ alors $Q(x, y) \geq C(x^2 + y^2)$ et si $a < 0$ alors $Q(x, y) \leq -C(x^2 + y^2)$.
(ii) Si $\Delta > 0$, alors il existe $C > 0$ et λ, μ tels que $Q(x, \lambda x) \geq Cx^2$ et $Q(x, \mu x) \leq -Cx^2$.

Démonstration: Dans le cas (i), si $b = 0$ on peut prendre $C = \max\{|a|, |c|\}$; si $b \neq 0$ on peut reprendre l'expression $Q(x, y) = a\left\{(x + by/a)^2 - \frac{\Delta}{a^2}y^2\right\}$. Distinguons le cas où $|by/a| \leq |x/2|$ alors $x^2 + y^2 \leq (1 + a^2/4b^2)x^2$ et $(x + by/a)^2 - \frac{\Delta}{a^2}y^2 \geq |x/2|^2$ d'où l'inégalité cherchée. Dans le cas où $|by/a| \geq |x/2|$ alors $x^2 + y^2 \leq (1 + 2b^2/a^2)x^2$ et $(x + by/a)^2 - \frac{\Delta}{a^2}y^2 \geq -\frac{\Delta}{a^2}y^2$ d'où l'inégalité cherchée. Dans le cas (ii) il suffit de choisir λ et μ tels que $Q(1, \lambda) > 0$ et $Q(1, \mu) < 0$. □

Ces propriétés permettent de prouver le théorème suivant.

THÉORÈME: *Soit $f : U \to \mathbf{R}$ une fonction de deux variables deux fois continûment dérivable au voisinage du point $x = (x_0, y_0) \in U$. Supposons $\vec{\text{grad}} f(x_0, y_0) = 0$ et posons*

$$a := \frac{\partial^2 f}{\partial x^2}(x), \quad b := \frac{\partial^2 f}{\partial x \partial y}(x), \quad \text{et} \quad c := \frac{\partial^2 f}{\partial y^2}(x)$$

Alors on a :
(i) Supposons $\Delta := b^2 - ac < 0$. Si $a < 0$, alors $f(x, y)$ admet un maximum local en (x_0, y_0) ; si $a > 0$, alors $f(x, y)$ admet un minimum local en (x_0, y_0) ;
(ii) Si $\Delta := b^2 - ac > 0$ alors $f(x, y)$ ne possède pas d'extremum local en (x_0, y_0). On dit que f admet un "point-selle" en (x_0, y_0).
(iii) Si $\Delta = 0$ on ne peut pas conclure sans autres considérations.

Démonstration: Sous les hypothèses, la formule de Taylor s'écrit :

$$f(x_0 + h, y_0 + k) - f(x_0, y_0) = \frac{1}{2}\left(ah^2 + 2bhk + ck^2\right) + \epsilon(h, k)(h^2 + k^2).$$

Dans le cas (i) avec disons $a > 0$ on peut écrire, en utilisant le lemme précédent :

$$f(x_0 + h, y_0 + k) - f(x_0, y_0) \geq \left(\frac{C}{2} + \epsilon(h, k)\right)(h^2 + k^2),$$

qui montre bien que la fonction admet un minimum local en (x_0, y_0). On procède de même si $a < 0$. Dans le cas (ii) on observe que

$$f(x_0 + h, y_0 + \lambda h) - f(x_0, y_0) \geq \left(\frac{C}{2} + \epsilon(h, \lambda h)\right)h^2,$$

alors que

$$f(x_0 + h, y_0 + \mu h) - f(x_0, y_0) \leq \left(\frac{-C}{2} + \epsilon(h, \mu h)\right)h^2,$$

ce qui montre bien que le point (x_0, y_0) ne peut être un extremum. □

Exemple. Recherchons les extrema de la fonction $f(x,y) = (x^2 - y^2)e^{-x^2-y^2}$. La fonction admet des dérivées continues à tout ordre donc on sait a priori que $\frac{\partial^2 f}{\partial x \partial y}(x,y) = \frac{\partial^2 f}{\partial y \partial x}(x,y)$. Le calcul des dérivées ne présente pas de difficulté, on trouve:

$$\begin{aligned}
\frac{\partial f}{\partial x}(x,y) &= (2x - 2x(x^2 - y^2))e^{-x^2-y^2} = 2x(1 - x^2 + y^2)e^{-x^2-y^2} \\
\frac{\partial f}{\partial y}(x,y) &= (-2y - 2y(x^2 - y^2))e^{-x^2-y^2} = -2y(1 + x^2 - y^2)e^{-x^2-y^2} \\
\frac{\partial^2 f}{\partial x^2}(x,y) &= (2 - 6x^2 + 2y^2 - 2x(2x - 2x^3 + 2xy^2))e^{-x^2-y^2} \\
\frac{\partial^2 f}{\partial x \partial y}(x,y) &= \frac{\partial^2 f}{\partial y \partial x}(x,y) = (4xy - 2y(2x - 2x^3 + 2xy^2))e^{-x^2-y^2} \\
\frac{\partial^2 f}{\partial y^2}(x,y) &= (-2 - 2x^2 + 6y^2 - 2y(-2y - 2yx^2 + 2y^3))e^{-x^2-y^2}.
\end{aligned}$$

Comme la fonction exponentielle ne s'annule jamais, l'équation d'un point critique s'écrit $x(1 - x^2 + y^2) = y(1 + x^2 - y^2) = 0$. Comme $1 + x^2 - y^2 = 1 - x^2 + y^2 = 0$ est impossible, on obtient $x = y = 0$ ou $x = 1 + x^2 - y^2 = 0$ ou $y = 1 - x^2 + y^2 = 0$, soit encore $(x,y) = (0,0)$ ou $(0, \pm 1)$ ou $(\pm 1, 0)$. Il y a donc 5 points critiques.

Pour conclure, on évalue les dérivées secondes en chacun des points critiques; notons $a := \frac{\partial^2 f}{\partial x^2}(x,y)$, $b := \frac{\partial^2 f}{\partial x \partial y}(x,y)$, $c := \frac{\partial^2 f}{\partial y^2}(x,y)$ et $\Delta := b^2 - ac$.

En le point $(x,y) = (0,0)$, on obtient $a = 2$, $b = -2$, $c = 0$ et $\Delta = 4$ donc on est en présence d'un point-selle. En le point $(x,y) = (\pm 1, 0)$, on obtient $a = -4/e$, $b = 0$, $c = -4/e$ et $\Delta = -16/e^2$ donc on est en présence d'un maximum. En le point $(x,y) = (0, \pm 1)$, on obtient $a = 4/e$, $b = 0$, $c = 4/e$ et $\Delta = -16/e^2$ donc on est en présence d'un minimum. [On peut remarquer que, comme f tend vers zéro quand (x,y) tend vers l'infini, le maximum de f est donc $f(\pm 1, 0) = e^{-1}$ et le minimum de f est donc $f(0, \pm 1) = -e^{-1}$].

Nous terminons ce chapitre en considérant le problème un peu plus général suivant dit des "*extrema liés*. Le problème consiste à déterminer les extrema d'une fonction $f(x_1, \ldots, x_n)$ où les variables $x_1, \ldots, x_n$ sont restreintes en leur imposant une condition $g(x_1, \ldots, x_n) = 0$. Le théorème de Lagrange suivant est le résultat central.

THÉORÈME: *Soit f, g deux fonctions continûment dérivables $U \to \mathbf{R}$. Si la fonction f, avec les variables contraintes par $g(x_1, \ldots, x_n) = 0$, admet un extremum local en a alors il existe un réel λ (appelé multiplicateur de Lagrange) tel que*

$$\vec{\text{grad}} f(a) = \lambda \, \vec{\text{grad}} g(a).$$

Nous donnons ci-dessous une intuition géométrique et également une démonstration sous l'hypothèse (à vrai dire peu restrictive) que l'on peut remplacer l'équation $g(x_1, \ldots, x_n) = 0$ par une équation du type $x_n = h(x_1, \ldots, x_{n-1})$.

Interprétation géométrique. Nous allons pouvoir faire un dessin si le nombre de variables est 2 ou 3. Soit $a = (a_1, a_2)$ où la fonction $f(x, y)$ atteint un extremum sous la contrainte $g(x, y) = 0$. Traçons les deux courbes $\mathcal{C}_1 : f(x, y) = f(a_1, a_2)$ et $\mathcal{C}_2 : g(x, y) = 0$; si les tangentes à $\mathcal{C}_1$ et $\mathcal{C}_2$ en le point (a_1, a_2) ne sont pas confondues, c'est-à-dire si $\vec{\text{grad}} f(a)$ n'est pas colinéaire avec $\vec{\text{grad}} g(a)$, un petit déplacement le long de la courbe $\mathcal{C}_2$ fera traverser la courbe de niveau $\mathcal{C}_1$, ce qui contredit (au moins visuellement) l'hypothèse d'un extremum local.

Le même raisonnement peut se visualiser avec deux fonctions de trois variables et les surfaces $\mathcal{S}_1 : f(x, y, z) = f(a_1, a_2, a_3)$ et $\mathcal{S}_2 : g(x, y, z) = 0$.

Démonstration: Si l'on suppose $g(x_1, \ldots, x_n) = x_n - h(x_1, \ldots, x_{n-1})$ et on pose $x' = (x_1, \ldots, x_{n-1})$ on peut écrire

$$\vec{\text{grad}} f(x) = \left(-\frac{\partial h}{\partial x_1}(x'), \ldots, -\frac{\partial h}{\partial x_{n-1}}(x'), 1 \right).$$

En remarquant que, sous la contrainte $g(x_1, \ldots, x_n) = 0$ on a

$$f(x_1, \ldots, x_n) = f(x_1, \ldots, x_{n-1}, h(x_1, \ldots, x_{n-1})),$$

on voit qu'un extremum local ne peut exister que si les dérivées de la fonction de droite par rapport à $x_1, \ldots, x_{n-1}$ s'annulent, c'est-à-dire si :

$$\frac{\partial f}{\partial x_1} + \frac{\partial h}{\partial x_1}\frac{\partial f}{\partial x_n} = \ldots = \frac{\partial f}{\partial x_{n-1}} + \frac{\partial h}{\partial x_{n-1}}\frac{\partial f}{\partial x_n} = 0$$

ce qui donne bien, en posant $\partial f \partial x_n(x) = \lambda$:

$$\vec{\text{grad}} f(x) = \frac{\partial f}{\partial x_n}(x)\, \vec{\text{grad}} g(x) = \lambda\, \vec{\text{grad}} g(x).$$

□

Exemples. Soit trois réels $0 < c < b < a$, on cherche le vecteur le plus long situé sur l'ellipsoïde d'équation :

$$g(x, y, z) = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1.$$

Solution. Le problème se ramène à chercher les extremums de $f(x, y, z) := x^2 + y^2 + z^2$ sous la contrainte $g(x, y, z) - 1 = 0$. Le théorème nous garantit qu'un tel extremum ne

peut survenir qu'en un point où il existe $\lambda \in \mathbf{R}$ tel que $\vec{\text{grad}} f(x,y,z) = \lambda \vec{\text{grad}} g(x,y,z)$ ou encore

$$(2x, 2y, 2z) = \lambda \left(\frac{2x}{a^2} + \frac{2y}{b^2} + \frac{2z}{c^2} \right).$$

Si x est non nul on en tire $\lambda = a^2$ et ensuite $y = z = 0$ donc en fin de compte $(x,y,z) = (\pm a, 0, 0)$; de même si y est non nul on obtient $(x,y,z) = (0, \pm b, 0)$ et si z est non nul on obtient $(x,y,z) = (0,0,\pm c)$. On vérifie alors aisément que $f(\pm a, 0, 0) = a^2$ est bien le maximum et $f(0,0,\pm c) = c^2$ le minimum. Ce que l'on peut "voir" sur le dessin suivant.

L'ellipsoïde $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$.

Exercice/exemple. Considérons la région du plan $U := \{(x,y) \in \mathbf{R}^2 \mid y \neq 0 \quad \text{ou} \quad x > 0\}$ que l'on peut recouvrir par trois ouverts $U_1 = \{(x,y) \in \mathbf{R}^2 \mid y > 0\}$, $U_2 = \{(x,y) \in \mathbf{R}^2 \mid x > 0\}$ et $U_3 = \{(x,y) \in \mathbf{R}^2 \mid y < 0\}$. Définissons les trois fonctions

$$\begin{aligned} f_1(x,y) &:= \text{Arctg}(x/y) && \text{pour} \quad (x,y) \in U_1 \\ f_2(x,y) &:= \frac{\pi}{2} - \text{Arctg}(y/x) && \text{pour} \quad (x,y) \in U_2 \\ f_3(x,y) &:= \pi + \text{Arctg}(x/y) && \text{pour} \quad (x,y) \in U_3 \end{aligned}$$

Vérifier que ces trois fonctions se recollent pour donner une fonction $f : U \to \mathbf{R}$, c'est-à-dire que $f_1(x,y) = f_2(x,y)$ pour $(x,y) \in U_1 \cup U_2$ et ainsi de suite et que f est harmonique, i.e. :

$$\Delta f(x,y) = \frac{\partial^2 f}{\partial x^2}(x,y) + \frac{\partial^2 f}{\partial y^2}(x,y) = 0$$

Montrer de plus que, si $x < 0$, alors

$$\begin{aligned} \lim_{\epsilon \to 0^+} f(x,\epsilon) &= -\pi/2 \\ \lim_{\epsilon \to 0^-} f(x,\epsilon) &= +3\pi/2 \end{aligned}$$

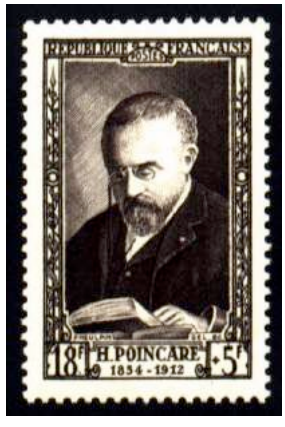
Remarques.

(i) On observera en particulier que l'inégalité du théorème des accroissements finis (dont les hypothèses ne sont pas vérifiées dans ce cas) est en défaut; en effet l'accroissement $|f(x,\epsilon) - f(x,-\epsilon)|$ ne tend pas vers zéro quand ϵ tend vers zéro.

(ii) Cette fonction permet de construire une solution de l'équation de Laplace $\Delta g(x,y) = 0$ sur U avec condition au bord $g(x,0^+) = +1$ et $g(x,0^-) = -1$ pour $x < 0$: il suffit de prendre $g(x,y) = \pi^{-1}(\pi/2 - f(x,y))$. Ceci permet de modéliser par exemple un potentiel électrique correspondant à une barre (l'axe des réels négatifs) avec une charge $+1$ au dessus et une charge -1 en dessous.

Kovalevskaïa Sofia (1850–1891)

Laplace Pierre Simon (1749-1827)



Poincaré Henri (1854-1912)