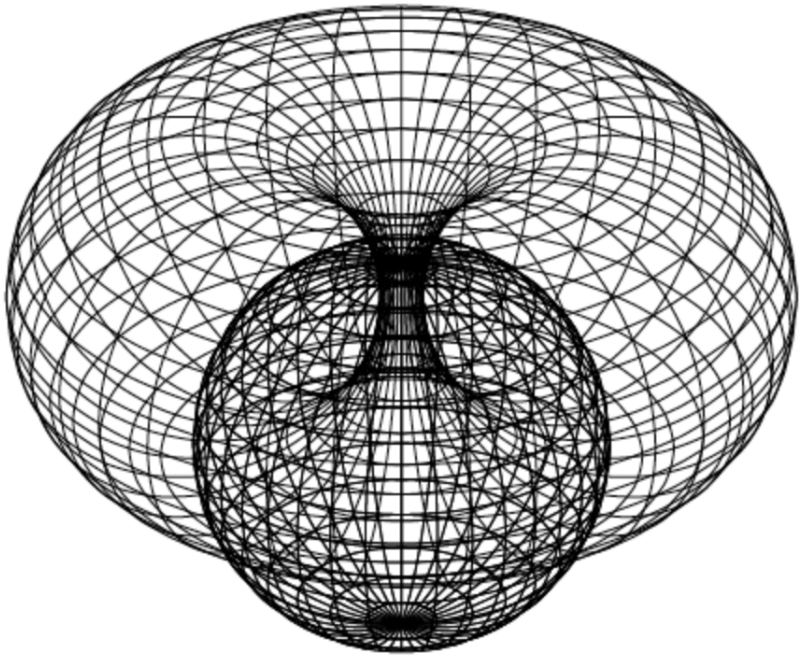
Fiches de Cours
Licence en Mathématiques

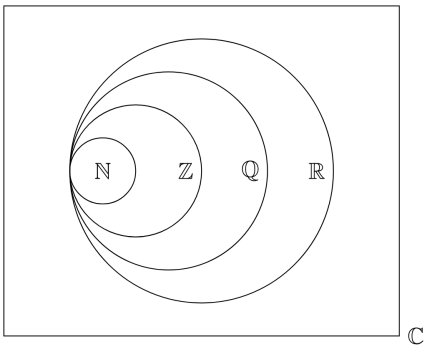


Université Jean François Champollion, Albi
Années universitaires 2022-2025

20 septembre 2025

Table des matières

Raisonnement et Ensembles



Chapitre 1

Raisonnements

Contents

Dans ce chapitre, nous allons définir les principaux quantificateurs et connecteurs logiques utilisés en mathématiques. Nous aborderons différents raisonnements très utilisés pour démontrer des propriétés. Nous présenterons aussi la structure la plus primitive des mathématiques qui, à elle seule, peut être développée en toute une théorie.

1.1 Assertions

En mathématiques, syntaxiquement, on utilise ce que l'on appelle des assertions composées de quantificateurs, de noms et de connecteurs logiques permettant d'affirmer des choses. Une assertion (phrase) peut être vraie ou fausse. On note généralement une assertion $\mathcal{A}$.

Exemple *"26 est plus petit que 50"* est un assertions vraie

1.1.1 Quantificateurs

Un quantificateur est un symbole mathématique permettant de donner une quantité d'un objet.

Définition (**Lettre Muette**) . Comme son num l'indique, une lettre muette est un lettre de l'alphabet (généralement x) muette. Elle peut être remplacée par n'importe quel objet en fonction de la définition de x, des propriétés que l'on décide qu'elle doit respecter.

Exemple Une lettre muette a peut représenter la valeur 2 ou 1.

On utilise les lettres muettes en mathématiques pour démontrer des résultats généraux. En manipulant des lettres muettes et en montrant des propriétés sur ces dernières, on n'a pas besoin de regarder chaque cas particulier.

Définition (**Quantificateur Universel**) . Le quantificateur universel $\forall$ nommé "pour tout" permet d'évoquer cette notion de généralité. Dans l'assertion $\forall x\ \mathcal{A}$, on veut dire "pour toute substitution de x par un objet donné, l'assertion $\mathcal{A}$ est vraie".

Exemple *"Pour toute voiture, celle-ci possède des roues".*

Définition (**Quantificateur Existenciel**) . Le quantificateur existenciel $\exists$ nommé "il existe" permet d'énoncer une propriété valable pour **au moins** un éléments. Ainsi, dans l'assertion $\exists x\ \mathcal{A}$ on veut dire que "il existe au moins un élément x qui vérifie la condition $\mathcal{A}$".

Exemple *L'assertion "il existe une voiture bleue est assurément vraie" tandis que l'assertion "toutes les voitures sont bleu ou rouge" n'était valable qu'en URSS.*

Chapitre 2

Ensembles

Contents

Dans ce chapitre, nous allons présenter plus en détail le concept d'ensemble. Les ensembles sont les structures les plus primitives des mathématiques, elles permettent par exemple d'effectuer des opérations à l'intérieur pour ensuite définir de nouveaux objets plus complexes.

2.1 Généralités

2.1.1 Définition

Définition (**Ensemble**) . Un ensemble est une collection non ordonnée d'objets appelés éléments.

L'ensemble vide, noté $\emptyset$, est l'unique ensemble ne contenant aucun élément. Un ensemble peut être vu comme un sac contenant divers éléments.

Exemple Les entiers positifs constituent un ensemble, de même que les entiers négatifs. On peut aussi parler au sens plus large de l'ensemble des jours de la semaine, ou de l'ensemble des voitures bleues.

2.1.2 Appartenance, Inclusion, Égalité

La théorie des ensembles est régie par une simple relation : l'appartenance.

Définition (**Appartenance**) . Soit E un ensemble et x un élément quelconque. On dit que x *appartient à* E si x est un élément de E. On peut aussi dire que E *contient* x. On note alors $x \in E$.
Dans le cas contraire, si "x n'est pas dans E", on dit que x *n'appartient pas à* E. On note alors $x \notin E$.

Par extension, on peut définir les notions d'inclusion et d'égalité entre ensembles.

Définition (**Inclusion**) . Soient E, F deux ensembles. On dit que F est inclus dans E ou que E contient F si tous les éléments de F appartiennent aussi à E. On note alors $F \subset E$. Plus formellement,

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

Définition (**Égalité**) . Soient E et F deux ensembles. On dira qu'ils sont égaux s'ils possèdent exactement les mêmes éléments. On notera alors $E = F$. Plus formellement,

$$E = F \iff (\forall x \in E,\ x \in F) \text{ et } (\forall x \in F,\ x \in E).$$

Proposition On peut aussi caractériser l'égalité d'ensembles en termes d'inclusions. Ainsi, deux ensembles égaux peuvent aussi être vus comme deux ensembles inclus l'un dans l'autre. D'où :

$$E = F \iff E \subset F \text{ et } F \subset E.$$

Pour montrer une égalité d'ensembles, on utilisera donc préférentiellement le principe de la double inclusion.

2.2 Opérations

Définissons maintenant quelques opérations sur les ensembles.

Définition (**Opérations**) . Soit E un ensemble et A et B deux parties de E. On définit les opérations suivantes :

- On appelle **complémentaire** de A dans E, noté $\overline{A}$, le sous-ensemble de E qui contient tous les éléments de E qui ne sont pas dans A.

$$\overline{A} := \{x \in E \mid x \notin A\}.$$

- On appelle la **différence** de A par B dans E, notée $A \setminus B$, le sous-ensemble de E contenant tous les éléments de A qui ne sont pas dans B.

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

- On appelle **intersection** de A et B, notée $A \cap B$, le sous-ensemble de E contenant tous les éléments à la fois dans A **et** dans B.

$$A \cap B := \{x \in E \mid x \in A \text{ et } x \in B\}.$$

- On appelle **union** de A et B, notée $A \cup B$, le sous-ensemble de E contenant tous les éléments appartenant à A **ou** à B.

$$A \cup B := \{x \in E \mid x \in A \text{ ou } x \in B\}.$$

Proposition
- On parlera d'union disjointe de deux ensembles A et B lorsque leur intersection est vide.
- On aura évidemment tout le temps $A \cap B \subset A$ et $A \cap B \subset B$.

Propriété (**Opérations**) . Soient E un ensemble et A, B et C des parties de E. On a les propriétés suivantes :

1. Commutativité de $\cap$ et $\cup$:
$$A \cap B = B \cap A \quad \text{et} \quad A \cup B = B \cup A.$$
2. Associativité de $\cap$ et $\cup$:
$$A \cap (B \cap C) = (A \cap B) \cap C \quad \text{et} \quad A \cup (B \cup C) = (A \cup B) \cup C.$$
3. Éléments neutres :
$$A \cup \emptyset = A, \quad A \cap E = A, \quad A \cap \emptyset = \emptyset, \quad A \cup E = E.$$
4. Distributivité :
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C) \quad \text{et} \quad A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$$
5. Complémentaire involutif : $\overline{\overline{A}} = A$.
6. Lois de De Morgan :
$$\overline{A \cap B} = \overline{A} \cup \overline{B} \quad \text{et} \quad \overline{A \cup B} = \overline{A} \cap \overline{B}.$$

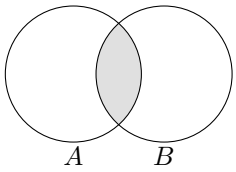


FIGURE 2.1 – Union $A \cup B$

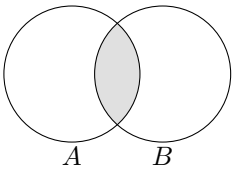


FIGURE 2.2 – Intersection $A \cap B$

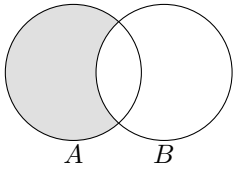


FIGURE 2.3 – Différence $A \setminus B$

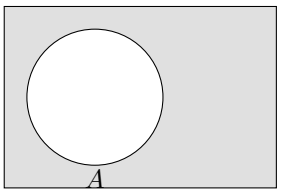


FIGURE 2.4 – Complémentaire $\overline{A}$

Définition (**Produit Cartésien**) . Soient A et B deux ensembles. On définit le produit cartésien de A et B, noté $A \times B$, comme l'ensemble composé de tous les couples possibles d'éléments de A et de B.
$$A \times B = \{(a, b) \mid a \in A, b \in B\}$$

Par définition, on dira que $(a, b) \in A \times B$ et $(c, d) \in A \times B$ sont égaux ssi $a = c$ et $b = d$.

Par extension, on peut définir le produit cartésien de plusieurs ensembles $A_1, \dots A_n, n \in \mathbb{N}$ comme l'ensemble des n-uplets de la forme $(a_1, \dots, a_n)$ où tous les $a_i \in A_i, \forall i \in [\![1, n]\!]$.

2.3 Ensembles de nombres

En mathématiques, nous travaillons avec différents ensembles de nombres possédant chacun différentes propriétés. Il est essentiel de bien les connaître ainsi que leurs relations d'inclusion.

2.3.1 Entiers Naturels

Définition (**Entiers Naturels**) . L'ensemble des entiers naturels, noté $\mathbb{N}$ est définit par :

$$0 \in \mathbb{N} \text{ et } \forall n \in \mathbb{N}, n + 1 \in \mathbb{N}$$

Autrement dit, $\mathbb{N}$ est le plus petit ensemble contenant 0 et fermé par l'opération successeur : $n \mapsto n + 1$.

On peut aussi définir $\mathbb{N}$ comme l'ensemble $\mathbb{N} := \{1, 2, 3, \dots\}$. Nous choisissons ici d'y inclure 0.

2.3.2 Entiers Relatifs

Définition (**Entiers Relatifs**) . On définit l'ensemble des entiers relatifs, noté $\mathbb{Z}$, comme :

$$\mathbb{Z} := \{\dots, -2, -1, 0, 1, 2, \dots\}$$

Plus formellement, $\mathbb{Z}$ peut être définit comme tous les nombres pouvant être formé de la soustraction de deux entiers naturels.

$$\mathbb{Z} = \{n - m \mid n, m \in \mathbb{N}\}$$

Proposition Cette définition permet d'en déduire les propriétés suivantes :

- $\mathbb{Z}$ contient $\mathbb{N}$
- $\mathbb{Z}$ est fermé par addition et soustraction.

2.3.3 Les Rationnels

Définition (**L'ensemble des Rationnels**) . L'ensemble des rationnels est définit comme l'ensemble des fractions d'entiers relatifs à dénominateur non nul. On le note $\mathbb{Q}$.

$$\mathbb{Q} := \left\{ \frac{a}{b} \mid a \in \mathbb{Z}, b \in \mathbb{Z}^* \right\}$$

Proposition $\mathbb{Q}$ est fermé par addition, soustraction, multiplication et division (sauf par zéro).

2.3.4 Ensemble des nombres réels

Définition (**Ensemble des nombres réels**) . L'ensemble des nombres réels, noté $\mathbb{R}$, est l'ensemble des limites de suites de rationnels convergentes. On peut aussi le définir comme

l'ensemble des nombres pouvant être représentés par une infinité décimale :

$$\mathbb{R} := \text{complétion de } \mathbb{Q} \text{ pour la norme } |\cdot|.$$

Remarque
- $\mathbb{R}$ contient tous les rationnels ($\mathbb{Q} \subset \mathbb{R}$) ainsi que tous les nombres irrationnels (comme $\sqrt{2}$, π).
- L'ensemble $\mathbb{R}$ est fermé par addition, soustraction, multiplication et division (sauf par zéro).
- Il est ordonné et complet : toute suite croissante et bornée converge dans $\mathbb{R}$.

Propriété (**Inclusions**) . On a toujours les inclusions :

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$$

2.4 Cardinalité d'un ensemble

Définition (**Ensemble fini et Cardinalité**) . Un ensemble E est dit *fini* s'il est vide ou s'il existe un entier naturel $n \in \mathbb{N}$ tel qu'il existe une suite finie d'éléments de E, $(x_1, \ldots, x_n)$, où chacun de ses éléments apparaît une et une seule fois.
Dans ce cas, on dira que n est le *cardinal* de E. On le note $|E| = n$.

Proposition Le cardinal $n \in \mathbb{N}$ d'un ensemble est unique. Le cardinal de l'ensemble vide est 0.

Exemple

$$|\{1, 2, 3\}| = 3, \quad |\emptyset| = 0$$

Remarque Pour les ensembles infinis (comme $\mathbb{N}$ ou $\mathbb{R}$), on parle de cardinal infini. Deux ensembles E et F ont le même cardinal si une bijection (voir chap. fonctions) existe entre eux.

Propriété (**Cardinal et Opérations**) . Soit E un ensemble fini de cardinal $n \in \mathbb{N}$.
- Tout sous-ensemble de E a un cardinal inférieur ou égal à n.
- Si A est un sous-ensemble de E alors $|E/A| = |E| - |A|$.
- Soient A et B deux sous-ensembles de E. On a alors :

$$\boxed{|A \cup B = |A| + |B| - |A \cap B|}$$

Chapitre 3

Fonctions et Applications

Contents

Les fonctions sont des objets essentiels en mathématiques. Outre leur utilité dans des domaines appliqués, elles constituent des objets mathématiques passionnants. Leur étude permet non seulement de comprendre la transformation d'éléments d'un ensemble à un autre, mais aussi de manipuler des ensembles et des relations abstraites. Dans ce cours, nous utiliserons les termes *fonctions* et *applications* de manière interchangeable.

3.1 Définition et notations

Une fonction peut être vue de deux façons différentes. La première la considère comme une sorte d'entité qui permet de « transformer » une valeur ou un élément en un autre. La seconde, que nous adopterons ici, la représente comme une correspondance entre deux ensembles.

Définition (**Application**) . Soient E et F deux ensembles. Une fonction f de E vers F est une correspondance qui à tout élément x de E associe un **unique** élément de F, noté $f(x)$. On écrit alors :

$$f : \begin{cases} E \longrightarrow F \\ x \longmapsto f(x) \end{cases}$$

Remarque Soit $f : E \longrightarrow F$ une fonction et $a \in E$, $b \in F$.

- L'ensemble E est appelé *ensemble de départ* ou *domaine* de f.
- L'ensemble F est appelé *ensemble d'arrivée* ou *codomaine* de f.
- Si $f(a) = b$, on dira que a *est un antécédent de* b *par* f et que b *est l'image de* a *par* f.

Attention : pour une fonction $f : E \longrightarrow F$ et $x \in E$, il ne faut pas confondre f, l'application en tant que correspondance, et $f(x)$ qui est un élément de F.

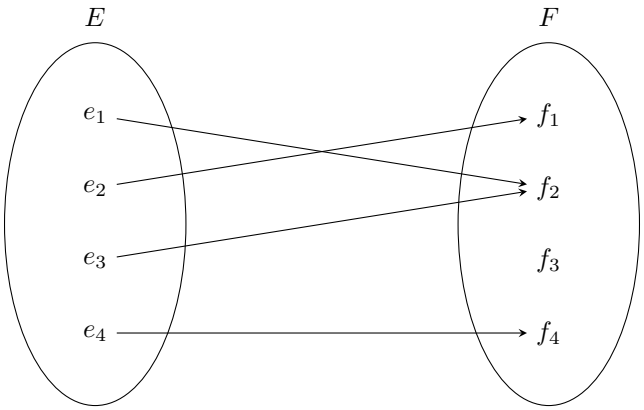


FIGURE 3.1 – Représentation sagittale d'une fonction $f : E \to F$

Exemple (Fonction carré) Soit la fonction

$$f : \mathbb{R} \longrightarrow \mathbb{R}, \quad x \longmapsto x^2.$$

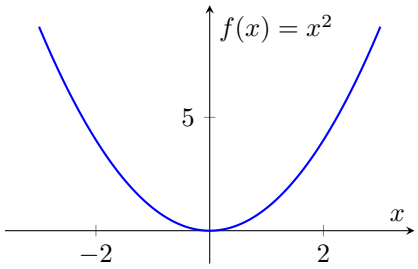


FIGURE 3.2 – Graphe de la fonction carré

C'est une application : à chaque réel x correspond un unique réel $f(x) = x^2$.
Domaine : $\mathbb{R}$, Codomaine : $\mathbb{R}$.
Image : l'ensemble des réels positifs ou nuls :

$$f(\mathbb{R}) = \{y \in \mathbb{R} \mid y \geq 0\}.$$

Quelques valeurs :

x	-2	-1	0	1	2
$f(x)$	4	1	0	1	4

On voit que 4 est l'image de 2 et de -2, et que $\sqrt{2}$ est un antécédent de 2.

Définition (Image et préimage) . Soit $f : A \longrightarrow B$ une fonction.

- **L'image directe** d'un sous-ensemble $X \subset A$ par f est :

$$f(X) := \{f(x) \mid x \in X\}.$$

- **La préimage** d'un sous-ensemble $Y \subset B$ par f est :

$$f^{-1}(Y) := \{x \in A \mid f(x) \in Y\}.$$

Exemple Soit $E = \{1, 2, 3\}$, $F = \{a, b\}$ et f définie par $f(1) = a$, $f(2) = a$, $f(3) = b$.

$$f(E) = \{a, b\}, \quad f^{-1}(\{a\}) = \{1, 2\}, \quad f^{-1}(\{b\}) = \{3\}.$$

Définition (Composition de fonctions) . Soient $f : E \longrightarrow F$ et $g : F \longrightarrow H$. La

composition $g \circ f : E \longrightarrow H$ est définie par :

$$\forall x \in E, \quad g \circ f(x) := g(f(x)).$$

Exemple Soient $f(x) = x^2$ et $g(x) = x + 1$. Alors :

$$g \circ f(x) = g(f(x)) = x^2 + 1.$$

Remarque Avant d'introduire l'injectivité et la surjectivité, on peut observer que :

- Certains éléments du codomaine peuvent avoir plusieurs antécédents (ex. 4 pour $f(x) = x^2$).
- Certains éléments du codomaine peuvent ne pas être atteints par la fonction (ex. -2 pour $f(x) = x^2$).

Ces observations motivent la définition des fonctions injectives et surjectives.

3.2 Injection et Surjection

Comme nous l'avons observé avec la fonction carré, certains éléments du codomaine peuvent avoir plusieurs antécédents et certains éléments peuvent ne pas être atteints. Cela conduit naturellement aux notions d'*injection*, *surjection* et *bijection*.

3.2.1 Fonction Injective

Définition (**Fonction injective**) . Une fonction $f : E \longrightarrow F$ est dite *injective* si des images égales impliquent des antécédents égaux. Formulée autrement :

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

Remarque L'injectivité garantit que chaque élément de l'image provient d'un seul élément du domaine. Intuitivement, la fonction ne "fusionne" pas d'éléments distincts du domaine sur le même élément du codomaine.

Exemple (Fonction injective) La fonction $f : \mathbb{R}_+ \longrightarrow \mathbb{R}$ définie par $f(x) = x^2$ est injective :

$$f(2) = 4, \quad f(3) = 9, \quad f(4) = 16$$

et chaque valeur de l'image correspond à un unique antécédent.

3.2.2 Fonction Surjective

Définition (**Fonction surjective**) . Une fonction $f : E \longrightarrow F$ est dite *surjective* si chaque élément du codomaine est atteint par au moins un antécédent :

$$\forall y \in F, \exists x \in E \text{ tel que } f(x) = y.$$

Remarque La surjectivité garantit que l'image couvre tout le codomaine. Intuitivement, aucun élément du codomaine n'est "laissé de côté".

Exemple (Fonction surjective) La fonction $f : \mathbb{R} \longrightarrow \mathbb{R}_+$ définie par $f(x) = x^2$ est surjective, car tout réel positif possède au moins un antécédent.

3.2.3 Fonction Bijective

Définition (**Fonction bijective**) . Une fonction $f : E \longrightarrow F$ est dite *bijective* si elle associe à chaque élément de F un unique antécédant de E. On peut alors construire sa *fonction inverse*, notée f^{-1} telle que :

$$f^{-1} : \begin{cases} F \longrightarrow E \\ f(x) \longmapsto x \end{cases}$$

Proposition Pour toute fonction bijective $f : E \longrightarrow F$, on a donc :

$$\forall x \in E, f^{-1}(f(x)) = x, \quad \forall y \in F, f(f^{-1}(y)) = y$$

Exemple (Fonction bijective) La fonction $f : \mathbb{R} \longrightarrow \mathbb{R}$ définie par $f(x) = x + 1$ est bijective d'inverse.

$$f^{-1}(y) = y - 1, \quad \forall y \in \mathbb{R}.$$

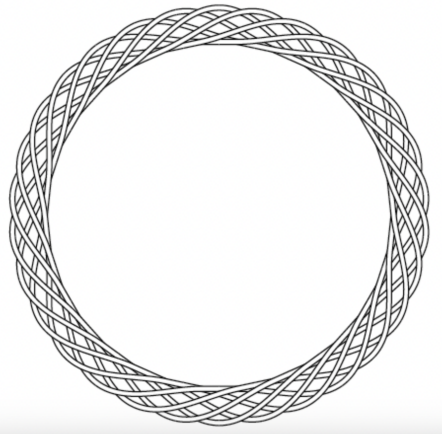
Propriété (**Caractérisation**) . Une fonction $f : E \longrightarrow F$ est bijective si et seulement si elle est *injective* et *surjective*.

Proposition Les bijections nous permettent de faire le lien avec la théorie des ensembles. En effet, deux ensembles fini A et B sont de même cardinal s'ils peuvent être mis en bijection.

On peut maintenant définir la notion d'ensemble dénombrable.

Définition (**Ensemble dénombrable**) . Un ensemble E est dit *dénombrable* s'il peut être mis en bijection avec l'ensemble des entiers naturels $\mathbb{N}$.

Algèbre des Structures



Chapitre 1

Groupes, Sous-Groupes

Contents

Grâce au cours de raisonnement et ensemble, nous connaissons bien la notion d'ensemble et les propriétés qu'ils ont. Cependant cette structure reste primitive et on commence à en faire le tour. Essayons maintenant de doter notre ensemble de quelques propriétés supplémentaires et d'une loi entre ses éléments.

1.1 Groupes

Définition (**Groupe**) . Un groupe est un couple $(G, *)$ où G est un ensemble et $*$ une application telle que $* : G \times G \to G$, que l'on appelle loi de composition interne. Une groupe satisfait les conditions suivantes :

1. $*$ **est associative** : $\forall x, y, z \in G,\ x * (y * z) = (x * y) * z$
2. **G est muni du neutre pour** $*$: $\exists e \in G, \forall x \in G,\ x * e = e * x = x$
3. **G est inversible** : $\forall x \in G, \exists y \in G,\ x * y = y * x = e$

Remarque

- Le groupe $(G, *)$ est dit **abélien** ou commutatif ssi $\forall g, h \in G, g * h = h * g$.
- **L'ordre** de G, noté $|G|$, est le cardinal de l'ensmeble si celui-ci est fini. Sinon on dit que G est d'ordre infini.

Définition (**Groupe Produit**) . Soient $(G_1, *)$ et $(G_2, *)$ deux groupes. On appelle groupe produit le groupe $(G_1 \times G_2, *)$ de G_1 et G_2 en posant :

$$(x_1, x_2) * (y_1, y_2) = (x_1 * y_1, x_2 * y_2)$$

1.2 Sous-groupes

Lorsque l'on a un groupe, on peut définir d'autre groupes à l'intérieur de celui-ci appelé sous-groupe. En général, pour montrer qu'une structure sur un ensemble connu est un groupe on essaye d'abord de montrer que c'est un sous-groupe d'un groupe connu.

Définition (**Sous-groupes**) . Un sous-groupe d'un groupe G est un sous-ensemble H de G sur lequel la multiplication de G induit une structure groupe. Il vérifie les propriétés suivantes :

- $e_G \in H$
- $\forall x, y \in H,\ x * y \in H$
- $\forall x \in H,\ x^{-1} \in H$

Théorème (**Neutre et inverse dans une sous-groupe**) . Si H est un sous groupe de G, le neutre de H est le même que le neutre de G et l'opposé dans H est le même que l'opposé dans G.

Proposition (CNS Sous-groupe) Soit $(G, *)$ un groupe. Une partie H de G est un sous-groupe de G ssi :

- $H \neq \emptyset$
- $\forall h_1, h_2 \in H, h_1 * h_2^{-1} \in H$

Définition (**Sous-groupe distingué**) . Un sous-groupe H de $(G, *)$ est dit **distingué** ssi

$$\forall g \in G, \forall h \in H, \quad ghg^{-1} \in H$$

On le notera $H \triangleleft G$

Remarque (Cas particulier des sous-groupes distingués)

- Pour tout groupe G, $\{e\} \triangleleft G$ et $G \triangleleft G$
- Une sous-groupe d'un groupe abélien est toujours distingué.

Définition (**Groupe Simple**) . Un groupe $G \neq \{e\}$ est appelé groupe simple si les seuls groupes distingués de G sont les sous-groupes triviaux $\{e\}$ et G.

Proposition L'intersection de sous-groupes distingués (ou non) de G est une sous-groupe distingué (ou non) de G.

1.3 Sous-groupe engendré

Dans certains espaces, on peut définir un groupe à partir d'un unique élément (ou de plusieurs) ce qui donne des groupes très intéressant à étudier.

Définition (**Sous-groupe engendré**) . Soit $(G, *)$ un groupe et X une partie de G. Il existe un plus petit sous-groupe de G contenant X appelé sous-groupe engendré par X et noté $\langle X \rangle$.
Un groupe fini est appelé **groupe cyclique** s'il existe $g \in G$ tel que $\langle g \rangle = G$.
Nous appelons l'ordre d'un élément $g \in G$ l'ordre du sous-groupe $\langle g \rangle$, engendré par g.

Corollaire () . Soient $(G, *)$ un groupe et $g \in G$ un élément d'ordre fini $n \in \mathbb{N}$. Alors n ets le plus petit entier strictement positif ayant la propriété $g^n = e$ et de plus :

$$\langle g \rangle = \{g, g^2, g^3, \ldots, g^n = e\}$$

Proposition (Sous-groupe des entiers) Soit H un sous-groupe de $(\mathbb{Z}, +)$, alors il existe un unique $n \in \mathbb{Z}$ tel que $H = n\mathbb{Z}$.

Chapitre 2

Groupe Symétrique

Contents

Maintenant que nous avons vu en détail la notion de groupe, sous-groupe et groupe engendré, attardons nous sur un des groupes les plus connus, le groupe symétrique. Ce groupe est un exemple presque canonique en algèbre des structures, il faut donc bien le connaître.

2.1 Le groupe symétrique

2.1.1 Définitions

Définition (**Groupe Symétrique**) . On appelle **groupe symétrique** le groupe des bijections de $\{1, \dots, n\}$ dans lui-même doté de la loi de composition, noté $(\mathcal{S}_n, \circ)$. Un élément de $\mathcal{S}_n$ quelconque est appelé une **permutation**.

Propriété (**Ordre et Bijection**) . Soit $n \in \mathbb{N}^*$ alors :

1. $\forall n \in \mathbb{N}^*, |\mathcal{S}_n| = n!$
2. si $|X| = n$ alors $\mathcal{S}(X) \cong \mathcal{S}_n$

Démonstration Soit $\mathcal{S}_n$ l'ensemble des bijections de $[\![1, n]\!], n \geq 3$ dans lui-même. Initilisation triviale.
Supposons que HR : $\forall n \geq 3, |\sigma_n| = n!$
Soient A_{n+1} et B_{n+1} deux ensembles à $n+1$ éléments deux à deux distincts tels que :

$$A_{n+1} = \{a_1, \dots, a_{n+1}\}$$

Soit σ une bijection telle que $\sigma : A_{n+1} \Longrightarrow B_{n+1}$. Combinatoirement, on a $n+1$ choix pour l'image de a_{n+1} par σ.

$$i.e \quad \sigma(a_{n+1}) \in B = \{b_1, \dots, b_{n+1}\}$$

Posons $\sigma(a_{n+1}) = b_{n+1}$. Notons que a_{n+1} n'est pas forcément égal à b_{n+1}. Pour respecter l'injectivité de σ, on n'a que $n!$ possibilités pour le choix de $\sigma(a_n)$ par hypothèse de récurrence. D'où, ici :

$$|\mathcal{S}_{n+1}| = |\mathcal{S}_n| \times (n+1) = n! \times (n+1) = (n+1)!$$

Démonstration Soit X un ensemble de cardinal $n \in \mathbb{N}^*$ et $\mathcal{S}(X)$ l'ensemble des bijections de X sur lui-même où :

$$X_n = \{x_1, \dots, x_n\}$$

Soit $\mathcal{S}_n$ l'ensemble des bijections de $\{1, \dots, n\}$ dans lui-même, alors :

$$i : \begin{cases} \{1, \dots, n\} \longrightarrow X \\ i \longmapsto x_i \end{cases} \quad \text{est une bijection}$$

Soit σ une bijection quelconque de $\mathcal{S}_n$ et τ une bijection quelconque de $X \longrightarrow X$

$$\begin{array}{ccc} X & \xrightarrow{\tau} & X \\ {\scriptstyle i}\uparrow & & \downarrow{\scriptstyle i^{-1}} \\ \{1, \dots, n\} & \xrightarrow{\sigma} & \{1, \dots, n\} \end{array}$$

Posons :

$$\Phi : \begin{cases} \mathcal{S}_n \longrightarrow \mathcal{S}(X) \\ \tau \longmapsto i^{-1} \circ \sigma \circ i \end{cases}$$

Alors Φ est une bijection de $\mathcal{S}(X)$ dans $\mathcal{S}_n$ et c'est aussi un morphisme de groupe car :

$$\Phi(\tau_1 \circ \tau_2) = i^{-1} \circ (\tau_1 \circ \tau_2) \circ i = i^{-1}\tau_1 i i^{-1}\tau_2 i = \Phi(\tau_1) \circ \Phi(\tau_2)$$

2.1.2 Structure du groupe symétrique

Théorème (**Centre d'un groupe**) . On désigne Z comme étant le centre d'un groupe. Ici, $\forall n \geq 3, Z(\mathcal{S}_n) = \{e\}$.

Démonstration (Par l'absurde) Soit $n \geq 3$ et $\sigma \in \mathcal{S}_n$ telle que $\sigma \neq e$.
Soit $x \in \{1, \dots, n\}, \sigma(x) \neq x$. Posons $y = \sigma(x)$ et $z = \sigma(y)$ et soit $\tau = (x\ y) \in \mathcal{S}_n$. Alors :

$$\begin{cases} \sigma \circ \tau(x) = \tau(y) = x \\ \tau \circ \sigma(x) = \sigma(y) = z \end{cases}$$

Donc $\exists x \in \{1, \dots, n\}$ tq $\sigma \circ \tau \neq \tau \circ \sigma$ donc $\sigma \notin Z(\mathcal{S}_n)$. Donc nécessairement, $Z(\mathcal{S}_n) = \{e\}$.

Théorème (**Cayley**) . Tout groupe fini G d'ordre n est isomorphe à un sous-groupe de $\mathcal{S}_n$.

2.1.3 Propriétés d'une permutation

Définition (**Points fixes, Support**) . Soient $n \in \mathbb{N}^*$ et $\sigma \in \mathcal{S}_n$.

- Les éléments $i \leq n$ tels que $\sigma(i) = i$ sont appelés **points fixes** de σ
- Une partie A de $\{1, \dots, n\}$ est dite stable par σ si $\sigma(A) \subseteq A$
- On définit le **support** de σ l'ensemble :

$$\boxed{\text{Supp}(\sigma) := \{k \in \{1, \dots, n\}, \sigma(k) \neq k\}}$$

Propriété (**Supports, Inclusion et Produit**) . Soient $n \in \mathbb{N}$ et $\sigma, \tau \in \mathcal{S}_n$ alors :

1. $\text{Supp}(\tau \circ \sigma) \subset \text{Supp}(\sigma) \cup \text{Supp}(\tau)$
2. si $\text{Supp}(\sigma) \cap \text{Supp}(\tau) = \emptyset$ alors :
   (a) $\text{Supp}(\pi \circ \tau) = \text{Supp}(\sigma) \cup \text{Supp}(\tau)$
   (b) $\begin{cases} \sigma \circ \tau = \tau \circ \sigma \\ \sigma \circ \tau = e \Longrightarrow \sigma = \tau = e \end{cases}$

Démonstration Soient $n \in \mathbb{N}$ et $\sigma, \tau \in \mathcal{S}_n$.

1. Soit $k \in \{1, \dots, n\}$ tel que $k \in \text{Supp}(\sigma \circ \tau)$. Par définition, $\sigma \circ \tau(k) \neq k$. Distinguons deux cas :
   (a) $\tau(k) = k$ alors $\sigma(\tau(k)) = \sigma(k) \neq k$ donc $k \in \text{Supp}(\sigma)$
   (b) $\sigma(k) = k$ alors puisque $\sigma \circ \tau(k) \neq k$, nécessairement, $\tau(k) \neq k$ donc $k \in \text{Supp}(\tau)$
   D'où $k \in \text{Supp}(\tau) \cup \text{Supp}(\sigma)$
2. Supposons que $\text{Supp}(\tau) \cap \text{Supp}(\sigma) = \emptyset$
   L'union des supports est égale au support du produit :
   (a) $\boxed{\subseteq}$ soit $k \in \text{Supp}(\sigma\tau)$ alors $\sigma \circ \tau(k) \neq k$ d'où $k \in \text{Supp}(\sigma) \cup \text{Supp}(\tau)$
   (b) $\boxed{\supseteq}$ soit $k \in \text{Supp}(\sigma) \cup \text{Supp}(\tau)$ alors $\sigma(k) \neq k$ ou $\tau(k) \neq k$ si $\tau(k) \neq k$ alors $\sigma(\tau(k)) \neq k$
   D'où $k \in \text{Supp}(\sigma\tau)$
3. Soient deux permutations σ et τ à supports disjoints et $k \in \{1, \dots, n\}$. Montrons que σ et τ commutent.
   Si $\sigma(k) = k$ alors $\tau(\sigma(k)) = \tau(k)$ et $\sigma(\tau(k)) = \tau(k)$ D'ou $\sigma \circ \tau = \tau \circ \sigma$. (On a le même raisonnement pour le second cas)
   Si $\sigma \circ \tau = e$, en raisonnant par l'absurde sur les supports, on a que $\sigma = \tau = e$.

Définition (**Cycle**) . Soit $p \in \mathbb{N}$ tel que $2 \leq p \leq n$ et soit $s \in \mathcal{S}_n$. On dit que s est un p-cycle de **support** $\{a_1, a_2, \dots, a_p\}$ et de **longueur** p ssi $\exists a_1, a_2, \dots, a_p \in \{1, 2, \dots, n\}$ deux à deux différents tels que :

$$\begin{cases} s(a_i) = a_{i+1} & \forall i < p \\ s(a_p) = a_1 & \\ s(k) = k & \forall k \notin \{a_1, \dots, a_p\} \end{cases}$$

Si s est un p-cycle alors on a : $s^p = Id$. Un cycle de longueur 2 est appelé **transposition**.

Théorème (**Décomposition de permutations**) . Toute permutation $\sigma \in \mathcal{S}_n$ peut s'écrire comme produit (i.e composition) de cycles $c_1, \dots, c_l$ à supports deux à deux disjoints. Cette décomposition est unique à l'ordre près.

Démonstration Soit $\sigma \in \mathcal{S}_n$. Soient $a, b \in \text{Supp}(\sigma)$ tels que :

$$\mathcal{O}(a) = \{\sigma^k(a); k \in \mathbb{N}\} \quad \mathcal{O}(b) = \{\sigma^k(b); k \in \mathbb{N}\}$$

Lemme Montrons que $\mathcal{O}(a) = \mathcal{O}(b)$ ou $\mathcal{O}(a) \cap \mathcal{O}(b) = \emptyset$
Supposons que $\mathcal{O}(a) \cap \mathcal{O}(b) \neq \emptyset$. Soit $x \in \mathcal{O}(a) \cap \mathcal{O}(b)$ alors $\exists k, l \in \mathbb{N}$ tels que :

$$\sigma^k(a) = x \text{ et } \sigma^l(b) = x \Longrightarrow \sigma^k(a) = \sigma^l(b)$$

or $\mathcal{O}(a)$ est un ensemble fini donc il existe $n \in \mathbb{N}$ tel que $\sigma^{k-l} = \sigma^n$ d'où $b \in \mathcal{O}(a)$ donc $\mathcal{O}(b) \subseteq \mathcal{O}(a)$
Symétriquement, on obtient facilement que $\mathcal{O}(a) \subseteq \mathcal{O}(b)$

Donc il existe $\mathcal{F} = (a_i, \dots, a_l), l \leq n$ tel que :

$$\text{Supp}(\sigma) = \bigcup_{i \leq l} \mathcal{O}(a_i)$$

Soit $c_i \in \mathcal{S}_n, i \leq l$ un cycle de longueur $p_i = |\mathcal{O}(a_i)|$ tel que :

$$c_i = (a_i \dots)$$

D'après le lemme précédent $\forall i, j \leq l$ et $i \neq j$:

$$\boxed{\text{Supp}(c_i) \cap \text{Supp}(c_j) = \emptyset}$$

Posons : $\boxed{\tau = c_1 \circ \cdots \circ c_i \circ \cdots \circ c_l}$

Montrons que $\sigma = \tau$: Soit $k \in [\![1, n]\!]$, distinguons deux cas :

- si $k \notin \text{Supp}(\sigma)$ alors $\sigma(k) = k$ et par construction $\mathcal{O}(k) = \{k\}$ d'où $\tau(k) = k$
- si $k \in \text{Supp}(\sigma)$ alors $\exists! i \leq l$ tel que $\tau(k) = c_i(k) = \sigma(k)$

On a donc montré que $\forall k \leq n, \tau(k) = \sigma(k)$. Donc $\sigma = \tau$

Propriété (Ordre d'une permutation) . Soit $\sigma \in \mathcal{S}_n$ et $c_1 \circ \cdots \circ c_m, m \leq n$ sa décomposition en produit de cycles à supports deux à deux disjoints. L'ordre de σ est égale au PPCM des longueurs des c_i.

Démonstration (Ordre d'une permutation) Soit σ une permutation de $\{1, \ldots, n\}$, d'après la propriété précédente :

$$\sigma = c_1 \circ \cdots \circ c_l$$

Où les c_i sont des cycles à supports deux à deux disjoints. Soient $k \, in\mathbb{N}, x \in [\![1, n]\!]$, alors :

$$\sigma^k(x) = x \Leftrightarrow c_1^k \circ \cdots \circ c_l^k(x) = x$$

Donc $\exists! i \leq l$ tel que :

$$\begin{aligned} \sigma^k(x) = x &\iff c_i^k(x) = x \\ &\iff k \in \text{ord}(c_i) \times \mathbb{N} \\ &\iff k \in \bigcap_{i \in [\![1,n]\!]} \text{ord}(c_i) \times \mathbb{N} \end{aligned}$$

or $\min\left\{\bigcap_{i \leq l} \text{ord}(c_i) \times \mathbb{N}\right\} = PPCM(\text{ord}_{i \leq l}(c_i))$. D'où le résultat.

2.1.4 Générateurs du groupe symétrique

Définition (Conjugaison) . Soient σ, τ deux permutations. On dit qu'elles sont conjuguées si :

$$\boxed{\exists \omega \in \mathcal{S}_n, \quad \sigma = \omega \circ \tau \circ \omega^{-1}}$$

Propriété (Conjugaison) . Deux permutations sont conjuguées ssi elles sont de même type.
En particulier, $\forall \sigma \in \mathcal{S}_n, \forall (i_1 \ldots i_l)$ un l-cycle, on a :

$$\sigma \circ (i_1 \ldots i_l)\sigma^{-1} = (\sigma(i_1) \ldots \sigma(i_l))$$

Théorème (Générateurs) . Le groupe symétrique est engendré par les transpositions.

Démonstration Soit $\sigma \in \mathcal{S}(n)$ alors $\exists c_1, \ldots, c_l \in \mathcal{S}_n$ des cycles à supports disjoints tels que :

$$\sigma = c_1 \circ \cdots \circ c_i \circ \cdots \circ c_l$$

où $c_i = (a_1 \ldots a_p)$. C'est un cycle donc on peut le décomposer en produit de transpositions.

$$\text{i.e } c_i = (a_1 \; a_2) \circ \cdots \circ (a_{p-1} \; a_p)$$

Soit $\tau \in \mathcal{S}_n$ le produit de la décomposition de chaque c_i en produit de transpositions. Soit $x \in [\![1, n]\!]$. Distinguons deux cas :

- si $x \notin \text{Supp}(\sigma)$ alors $\sigma(x) = x$ et par construction, $\tau(x) = x$ aussi.
- si $x \in \text{Supp}(\sigma)$ alors $\exists! i \leq l, x \in \text{Supp}(c_i)$ où $c_i = (a_1 \; a_2) \circ \cdots \circ \underset{\omega \in \mathcal{S}_n}{(x \; \sigma(x))} \circ (\sigma(x) \ldots) \circ \cdots \circ (a_{p-1} \; a_p)$ or ω est la seule permutation cd c_i telle que $x \in \text{Supp}(\omega)$ donc $c_i(x) = \sigma(x) = \tau(x)$

Donc $\sigma = \tau$. D'où $\mathcal{S}_n$ est engendré par les transpositions.

2.2 Signature et groupe alterné

2.2.1 Signature d'une permutation

Définition (**Signature**) . Soit $\sigma \in \mathcal{S}_n$, on appelle signature de σ l'entier $\varepsilon(\sigma)$ telle que :

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

Propriété (**La signature est un morphisme de groupes**) . $\varepsilon : \mathcal{S}_n \longrightarrow (\mathbb{Q}^*, \times)$ est un morphisme de groupes.

Propriété (**Calcul de la signature**) . Soit $z \in \mathcal{S}_n$ un cycle de longueur $p \geq 2$ alors :

$$\varepsilon(z) = (-1)^{p-1}$$

(Ce résultat se montre en décomposant z en produit de transpositions, puis en appliquant ε en tant que morphisme de groupes.)

2.2.2 Groupe Alterné

Définition (**Noyau, Groupe alterné**) . On appelle groupe alterné le noyau de ε. C'est un sous-groupe distingué de $\mathcal{S}_n$. On le note $\mathcal{A}_n$.

Propriété (**Groupe Alterné**) . Soit $n \geq 3$

- $\mathcal{S}_n$ est engendré par les permutations $(1, i), 2 \geq i \geq n$
- $\mathcal{A}_n$ est engendré par les cycles $(1, i, j), 2 \geq i \neq i \geq n$

En partiulier, $\mathcal{A}_n$ est engendré par les 3-cycles de $\mathcal{S}_n$.

Chapitre 3

Morphismes de Groupes

Contents

En mathématiques, dès que l'on découvre une nouvelle structure particulière, on essaye de définir des applications dessus pour essayer de voir si certaines structures se ressemblent ou pas. C'est ce que l'on va faire avec les groupes.

3.1 Morphismes, Image et Noyaux

3.1.1 Définitions Générales et premières propriétés

Définition (**Morphisme de groupes**) . Soient $(G, *)$ et $(K, \bullet)$ deux groupes. Un morphisme de groupes ϕ entre G et K est une application

$$\phi : \begin{cases} G \longrightarrow K \\ g \longmapsto \phi(g) \end{cases}$$

telle que :

$$\forall (g_1, g_2) \in G \times G, \quad \phi(g_1 * g_2) = \phi(g_1) \bullet \phi(g_2)$$

"Un morphisme de groupes est une application qui respecte la structure des groupes."

Définition (**Noyau et Image**) . Soit ϕ un morphime entre deux groupes $(G, *)$ et $(K, \bullet)$.

- Le **noyau** de ϕ est l'ensemble des éléments de G envoyés par ϕ sur e_K

$$\text{i.e } \ker(\phi) = \{g \in G, \quad \phi(g) = e_K\}$$

- **L'image** de ϕ est l'ensemble des $\phi(g), g \in G$ dans K.

$$\text{i.e } \phi(G) = \{h \in K, \quad \exists g \in G \text{ tq } \phi(g) = h\}$$

Définition (**Morphismes et bijection**) .

- Un **isomorphime** de G dans K est un morphisme de groupes **bijectif**.
- Un **endomorphisme** de G est un morphisme de G dans lui-même.
- Un **automorphisme** de G est un morphisme bijectif de G dans lui-même.

Propriété (**Morphismes**) . Soit $\phi : (G, *) \longrightarrow (K, +)$ un morphisme de groupes.

- $\phi(e_g) = e_K$
- $\phi(g^{-1}) = \phi(g)^{-1}$
- $\phi(g^n) = \phi(g)^n$

3.1.2 Image et Noyaux

Lemme (Image et Image réciproque d'un morphisme) Soit $\phi : (G, *) \longrightarrow (K, \bullet)$ un morphisme de groupes.

- L'image d'un sous-groupe H de G est un sous-groupe de G.
- Si $H \lhd G$ alors $\phi(H) \lhd \phi(G)$
  En particulier, si ϕ est surjectif, alors $\phi(H) \lhd K$
- L'image réciproque d'un sous-groupe de K est un sous-groupe de G.
- Si $H' \lhd G'$ alors $\phi^{-1}(H') \lhd G$

De plus, l'image de ϕ est un sous-groupe de K et $\ker(\phi)$ est un sous-groupe distingué de G.

Corollaire (**Morphisme et sous-groupe engendré**) . Soit $\phi : (G, *) \longrightarrow (K, +)$ un morphisme de groupes et $X \subset G$, alors

$$\boxed{\phi(\langle X \rangle) = \langle \phi(X) \rangle}$$

"L'image d'un générateur est le générateur de l'image."

Propriété (**Stabilité**) . Les morphismes de groupes sont stables par composition et réciproque (dans le cas d'un isomorphisme).

Lemme Un morphisme de groupes $\phi : G \longrightarrow K$ est injectif si et seulement si $\ker(\phi) = \{e_G\}$.

3.1.3 Le cas des entiers

Proposition Il existe un unique groupe cyclique d'ordre $n \in \mathbb{N}$ (à isomorphisme près).

Théorème (**Ordre Fini**) . Soit G un groupe et $g \in G$.

1. g est d'ordre **infini** ssi $\langle g \rangle \sim (\mathbb{Z}, +)$, dans ce cas :
$$\langle g \rangle = \{\dots, g^{-3}, g^{-2}, g^{-1}, e, g, g^2, g^3, \dots\}$$
2. g est d'ordre **fini** $n \in \mathbb{N}$ si et seulement si l'ensemble $\langle g \rangle = \{e, g, g^2, \dots, g^{n-1}\}$ est bien défini et si $g^n = e$.

3.2 Automorphismes

Pour rappel, un automorphisme d'un groupe G est un morphisme de groupe bijectif de G dans lui-même.

Définition (**Centre d'une groupe**) . Le centre d'un groupe G est l'ensemble $Z(G)$ des éléments de G qui commutent avec tous les éléments de G.

Théorème (**Automorphismes**) . L'ensemble $(Aut(G), \circ)$ des automorphismes d'un groupe G muni de la composition, est un groupe.

- C'est un sous-groupe de l'ensemble des permutations de G.
- Si G est d'ordre fini n, alors $Aut(G)$ est un groupe d'ordre au plus $(n-1)!$.

Définition (**Automorphisme intérieur**) . Soit $(G, *)$ un groupe noté multiplicativement, on définit l'automorphisme intérieur associé à $g \in G$ le morphime de groupes tel que :
$$Int_g \begin{cases} G \longrightarrow G \\ x \longmapsto gxg^{-1} \end{cases}$$
On définit $Int(G)$ comme l'ensemble des automorphismes intérieurs de G.

Remarque

- Tout automorphisme intérieur est un automorphisme de groupes.
- Deux éléments d'un groupes dont l'un est image de l'autre par automorphisme intérieur sont dits conjugués.
- On remarquera qu'un sous-groupe H de G est distingué dans G ssi H est stable par tous les automorphismes intérieurs de G.

Chapitre 4

Théorème de Lagrange

Contents

Joseph Louis de Lagrange était un mathématicien, astronome et mécanicien italien de la fin du XVIII° siècle. Né en 1736 à Turin en Sardaigne, il mourut en 1813 à Paris. Initiateur du Calcul des Variations, il étudia la mécanique de l'artillerie, l'astronomie, au travers des variations de l'orbite lunaire, l'analyse et se trouva très fort en arithmétique. On lui doit le Théorème des Quatre Carrés et le Théorème de Lagrange sur les groupes d'ordre fini.
La théorie des groupes n'existant pas à son époque, puisqu'elle émergeat qu'au XIX°s sous l'impulsion de Cauchy, Gauss et Galois, Lagrange n'énonça pas son théorème tel que nous allons le voir, mais il lui est quand même dû.

4.1 Relation d'équivalence

De légers rappels sur les relations et les classes d'équivalences ne sont jamais de trop...

Définition (**Relation d'équivalence**) . Une relation d'équivalence $\sim$ entre deux éléments x et y d'un ensemble X est une relation (RST) :

- **Réflexive :** $x \sim x$
- **Symétrique :** $x \sim y \iff y \sim x$
- **Transitive :** $\forall z \in X, x \sim y$ et $y \sim z \implies x \sim z$

Remarque Etre isomorphe est une relation d'équivalence sur les groupes.

Définition (**Classes d'équivalence**) . Soit $\sim$ une relation d'équivalence sur un ensemble X. Une classe d'équivalence de $x \in X$, souvent noté $\overline{x}$ est l'ensemble :

$$\overline{x} = \{y \in X, x \sim y\}$$

Remarque On remarquera que les classes d'équivalences sur un ensemble forment une partition de cet ensemble.

Définition (**Projection Canonique**) . Soit $\sim$ une relation d'équivalence sur un ensemble X. On appelle **ensemble quotient de X par $\sim$** l'ensemble :

$$X/\sim = \{\overline{x}, x \in X\}$$

On définit alors la **projection canonique** comme l'application :

$$\pi : \begin{cases} X \longrightarrow X/\sim \\ x \longmapsto \overline{x} \end{cases}$$

"C'est l'application, qui, à x lui associe sa classe d'équivalence pour la relation $\sim$."

Remarque La projection canonique est surjective puisqu'une classe d'équivalence n'est jamais vide.

4.2 Classes à gauche

Définition (**Classe à gauche**) . Soit G un groupe noté multiplicativement et H un sous-groupe de G. La relation $\sim_H$ est une relation d'équivalence sur G telle que :

$$\forall g_1, g_2 \in G, \quad g_1 \sim_H g_2 \Longleftrightarrow \exists h \in H, g_1 = h.g_2$$

Ses classes d'équivalences sont les ensembles gH notés $gH = \{gh, h \in H\}$ de G. On les appelles **classes de g à gauche** de G modulo H.

Remarque On peut définir de façon similaire les classes à droite de G modulo H, mais dans la pratique, nous n'utiliserons que les classes à gauche.

Propriété (**Classes à gauche et cardinal**) . Soit G un groupe et H un sous-groupe de G. Toutes les classes à gauche de G modulo H ont le même cardinal que H.

Démonstration Montrons que toutes les classes à gauche de G modulo H sont isomorphes à H. Soit $g \in G$. Considérons ϕ_g l'application telle que :

$$\phi_g : \begin{cases} H \longrightarrow gH \\ h \longmapsto gh \end{cases}$$

1. **Injectivité :** Soient $h_1, h_2 \in H$ tels que :

$$\begin{aligned} & \phi(h_1) = \phi(h_2) \\ \Longleftrightarrow\ & gh_1 = gh_2 \\ \Longleftrightarrow\ & h_1 = h_2 \end{aligned}$$

2. **Surjectivité :** Soit $x \in gH$, $\exists h \in H$ tel que $\phi(h) = gh = x$

Donc ϕ est bijective. D'où le résultat.

4.2.1 Ensemble Quotient et Indice

Définition (**Ensemble Quotient**) . Soit $(G, .)$ un groupe noté multiplicativement et $H \leq G$. L'ensemble quotient de G par la relation d'équivalence $\sim_H$, noté G/H est l'ensemble $\{gH,\ g \in G\}$ des classes à gauche de G modulo H.

L'indice de H dans G, noté $(G : H)$ est le cardinal de l'ensemble quotient G/H. Il correspond au nobre de classes d'équivalences différentes pour la relation $\sim_H$ dans G.

Théorème (**Cardinal du groupe et ensemble quotient**) . Soit $(G, .)$ un groupe et $H \leq G$ alors :

$$|G| = |H| \times (G : H)$$

4.2.2 Théorème de Lagrange appliqué aux groupes finis

Théorème (**Théorème de Lagrange**) . Soit G un groupe fini et $H \leq G$, alors l'ordre de H divise l'ordre de G.

Corollaire (**Ordre d'un élément**) . Soit G un groupe d'ordre fini et $g \in G$ alors $\text{ord}(g) \mid \text{Card}(G)$.

Proposition (Sous-groupe d'indice 2) Soit G un groupe et $H \leq G$, si H est d'indice 2 alors il est distingué dans G.

Démonstration Soit $g \in G$. Distinguons deux cas :

- si $g \in H$ alors comme H est un sous-groupe $\forall h \in H, ghg^{-1} \in H$
- si $g \in G \backslash H$. H est d'indice 2 donc G n'a que deux classes modulo H, $H = eH = He$ et puisque les classes forment une partition de G, $G \backslash H$ forme la deuxième classe à gauche et à droite.
  Donc $G \backslash H = gH = Hg$ et $\forall h \in H,\ \exists h' \in H$ tel que $gh = h'g$.
  C'est à dire que $ghg^{-1} = h' \in H$.

Donc $H \triangleleft G$.

Chapitre 5

Actions de Groupes

Contents

Depuis le début de l'étude des groupes nous voyons ces objets simplement comme des ensembles auxquels ont a donné une structure via une application et quelques autres propriétés. Ces objets, en apparence très simples, sont pourtant très riches et peuvent être très variés.
Une autre approche de l'étude de la théorie des groupes est de les faire agir sur un ensemble. L'un des exemples le plus intuitif pour comprendre cette notion est le groupe symétrique. En effet, il est définit comme l'ensemble des bijections sur un ensemble quelconque. Mais lorsqu'on choisit comme ensemble les entiers de 1 à $n \in \mathbb{N}$, on s'apperçoit que notre groupe permute tout simplement des suites d'entier.

5.1 Actions de groupes, premières définitions

Définition (**Actions de groupe**) . Soit G un groupe et X un ensemble quelconque. On définit l'action à gauche de G sur X, souvent noté $G \curvearrowright X$, l'application :

$$\bullet : \begin{cases} G \times X \longrightarrow X \\ (g, x) \longmapsto g.x \end{cases}$$

qui vérifie :

i) **Associativité Mixte :** $\forall g, h \in G,\ \forall x \in X,\ (gh).x = g.(h.x)$

ii) **Invariance par le neutre :** $\forall x \in X,\ e_G.x = x$

On appelle ainsi X un G-ensemble.

Définition (**Point fixe et partie stable**) . Soient G un groupe et X un G-ensemble :

- $x \in X$ est un point fixe de G si, $\forall g \in G,\ g.x = x$
- $Y \subset X$ est une partie stable de X si, $\forall y \in Y, \forall g \in G, g(Y) \subset Y$.

5.2 Morphisme Structurel

Théorème (**Morphisme structurel**) . Soient G un groupe et X un G-ensemble. Il existe une bijection entre les opérations au gauche de G sur X et l'ensemble des bijections de X sur lui-même ($S(X)$).

$$\text{i.e} \quad \begin{cases} G \times X \to X \\ (g, x) \mapsto g.x \end{cases} \quad \overset{\sim}{\longmapsto} \quad \phi : \begin{cases} G \to S(X) \\ g \mapsto \sigma_g \end{cases}$$

où $\forall g \in G, \forall x \in X,\ \sigma_g(x) = g.x$

Proposition Soient G un groupe et ϕ l'action de G sur le G-ensemble X.

- Soit ψ un morphisme de groupes tq $\psi : K \to G$, alors X est aussi un $\psi \circ \phi$-ensemble (i.e on peut composer une action de groupe par la droite par un morphisme de groupe dont l'image est le groupe opérant).
- Soit $Y \subset X$, alors Y est aussi un G-ensemble.
- Soit $H \leq G$, alors X est aussi un H-ensemble.

Remarque Notons qu'une action droite d'un groupe G sur un ensemble X correspond à un anti-morphisme de groupe (i.e du point de vue du morphisme structurel $\forall g, h \in G, \phi(gh) = \phi(h)\phi(g)$).

5.3 Orbites et Stabilisateurs

5.3.1 Définitions

Définition (**Orbite et Stabilisateur**) . Soit G un groupe et X un ensemble tels que $G \curvearrowright X$, et soit $x \in X$, on peut alors définir les deux ensembles suivants :

- **L'orbite** de x est le sous-ensemble $\text{Orb}(x) \subseteq X$ tel que :

$$\boxed{\text{Orb}_G(x) = \{g.x \mid g \in G\}}$$

- **Le stabilisateur** de x est le sous-groupe de G tel que :

$$\boxed{\text{Stab}(x) = \{g \in G \mid g.x = x\}}$$

L'orbite d'un élément x de X correspond à tous les éléments de X que l'on peut atteindre sous l'action de G sur x. Le stabilisateur de x de X correspond à tous les éléments de G qui laissent x invariant.

Définition (**Propriétés d'une action**) . Soit G un groupe et X un ensemble tels que $G \curvearrowright X$ on dit alors que G est :

- **Transitive** sur X s'il existe exactement une seule orbite dans X.
- **Libre** sur X si tous les statbilisateurs sont triviaux (i.e réduits à l'élément neutre de G).

Propriété (**Stabilité des orbites**) . Les orbites de X sont stables par G et toute union d'orbites de X est aussi stable par G.

5.3.2 Orbites, stabilisateurs et relation d'équivalence

Proposition Soit G un groupe et X un G-ensemble, alors :

- La relation $\sim$ définie sur X par

$$x \sim y \text{ si } x \in \mathrm{Orb}(y)$$

  est une relation d'équivalence sur X.
- Les classes d'équivalences pour cette relation sont exactement les orbites de X pour l'action de G.

Propriété () . Les orbites de X par l'action de G forment une partition de X.

Propriété (**Noyaux**) . Soient G un groupe et ϕ une action de groupe de G sur un ensemble X. Le noyaux de ϕ, $\ker(\phi)$, est l'intersection de tous les stabilisateurs de X.

$$\boxed{\text{i.e} \quad \ker(\phi) = \bigcap_{x \in X} \mathrm{Stab}_G(x)}$$

Définition (**Action fidèle**) . Soient G un groupe et X un G-ensemble pour une action ϕ, on dit que ϕ est fidèle si $\ker(\phi) = \{e_G\}$.

La fidélité d'une action de groupe peut être assimilé à l'injectivité.

5.4 Actions Particulières

5.4.1 Opération par translation

Ici, nous allons étudier les propriétés des actions par translations. Jusqu'à maintenant, nous avons étudié des actions de certains groupes sur des ensembles. Or on peut aussi considérer un groupe comme un ensemble et donc faire agir un groupe **sur lui-même**...

Définition (**Action par translation**) . Soit $(G, .)$ un groupe noté multiplicativement. On définit l'action à gauche de G sur lui-même :

$$\begin{aligned} G \times G &\longrightarrow G \\ (g, h) &\longmapsto g.h = gh \end{aligned}$$

Théorème de Cayley

Théorème (**Cayley, 1878**) . Tout groupe fini d'ordre $n \in \mathbb{N}$ est isomorphe à un sous-groupe de $\mathcal{S}_n$.

Ce théorème, bien que puissant n'est en réalité pas très utile. En effet pour un groupe d'ordre 6, on peut ainsi montrer qu'il est isomorhe à $\mathcal{S}_6$. On connaît bien le groupe symétrique mais $\mathcal{S}_6$ est d'ordre $6! = 720$, au final, on ne récolte pas spécialement plus d'informations.

5.4.2 Opération par conjugaison

Lorsque l'on a étudié le groupe symétrique, on a définit la conjugaison entre deux permuations. Ici, nous allons définir une action de groupe permettant de dire que deux permutations conjuguées sont dans la même orbite..intéressant non ?

Définition (**Action par conjugaison**) . Soit $(G, .)$ un groupe noté multiplicativement. On définit l'action par conjugaison de G sur lui-même :

$$\begin{aligned} G \times G &\longrightarrow G \\ (g, h) &\longmapsto g.h = ghg^{-1} \end{aligned}$$

On peut remarquer qu'elle est bien définie puisque g^{-1} appartient bien à G. Et ghg^{-1} est bien dans G par stabilité.

Remarque Permettons nous de faire quelques remarques...

- Si G est abélien, alors l'action par conjugaison de G sur lui-même en triviale.
- Ici, **l'orbite** d'un élément $h \in G$ est appelé **classe de conjugaison** de h.

On peut ainsi définir de nouveaux objects...

Définition (**Centralisateur**) . Soit $(G, .)$ un groupe noté multiplicativement. On définit le centralisateur de $h \in G$ comme l'ensemble :

$$Z_G(h) = \{g \in G \mid ghg^{-1} = h\}$$

Le centralisateur d'un élément $h \in G$ peut se voir comme l'ensemble des éléments $g \in G$ qui **commutent** avec h.

Autrement dit, le centralisateur d'un élément est le stabilisateur du même élément pour l'action par conjugaison.

Chapitre 6

Formule des Classes

Contents

Maintenant que nous avons vu en détail les actions de groupes, nous allons d'abord essayer de trouver des relations entre les différents ensembles qu'elles nous permettent de définir. Dans un second temps nous définirons proprement les groupes quotient du point de vue des actions de groupes, puis nous finirons par présenter les théorèmes d'isomorphismes.

6.1 Formule des classes

6.1.1 Relation Orbite/Stabilisateur

Quand on manipule des actions de groupes, on s'apperçoit vite que l'orbite d'un élément est étroitement lié à son stabilisateur...

Proposition Soient $(G, .)$ un groupe, X un ensemble et G_x le stabilisateur de $x \in X$ pour l'action de G sur X. Alors l'application

$$\begin{aligned} G/G_x &\longrightarrow \mathrm{Orb}_G(x) \\ gG_x &\longmapsto g.x \end{aligned}$$

Est une bijection entre l'ensemble des classes à gauche du stabilisateur de x et l'orbite de x sous G.

A partir de cette application et à l'aide du formidable outil qu'est le théorème de Lagrange, on peut en déduire le corollaire suivant...

Corollaire (**Relation Orbite/Stabilisateur**) . Soit X un G-ensemble, $x \in X$. On utilise les mêmes notations que précédement. On a alors :

- $|\mathrm{Orb}_G(x)| = (G : G_x)$
- $|G| = |G_x| \times \mathrm{Orb}_G(x)$

6.1.2 Formules des classes

Corollaire (**Formule des classes**) . Soient X G un groupe et X un G-ensemble. Alors on peut **partitionner X en orbites disjointes** sous l'action de G, d'où :

$$X = \bigsqcup_{i=1}^{r} \mathrm{Orb}_G(x_i) \quad \text{et} \quad |X| = \sum_{i=1}^{r} |\mathrm{Orb}_G(x_i)| = \sum_{i=1}^{r} (G : G_x) = \sum_{i=1}^{r} \frac{|G|}{|G_{x_i}|}$$

où r est le nombre d'orbites distinctes de X sous l'action de G.

6.2 Groupe Quotient

Nous savons déjà que le noyau d'un morphisme de groupes est un sous-groupe distingué. Nous cherchons ici, à partir d'un sous-groupe distingué, à construire un morphisme dont il serait le noyau. Pour cela, nous devons définir les groupes quotients.

6.2.1 "Quotientons !"

Soit G un groupe et H un sous-groupe de G. On définit l'ensemble $G/H = \{gH \mid g \in G\}$ comme l'ensemble quotient de G par la relation d'équivalence "appartenir à la même orbite" pour l'action par translation de H sur G.

Définition (**Projection Canonique**) . Soient G un groupe et H un sous-groupe distingué de G. La projection canonique :

$$\pi : \begin{cases} G \longrightarrow G/H \\ g \longmapsto gH \end{cases}$$

est **morphisme de groupes** qui envoie chaque élément de G sur sa classe d'équivalence modulo la relation "appartenir à la même orbite" évoquée précédement.

Théorème (**Sous-groupe distingué et projection canonique**) . Soit G un groupe. Sous H un sous-groupe de G.

H **est distingué** dans G ssi la formule $g_1H * g_2H = (g_1g_2)H,\ \forall g_1, g_2 \in G$ définit une loi de composition interne pour l'ensemble G/H telle que la projection canonique sur G est un morphisme de groupes.

$\iff$ G/H est un groupe pour la loi $*$

Corollaire (**Ordre du quotient**) . Si H est distingué dans G, alors l'ordre du quotient de G par H est égal au quotient de l'ordre de G par l'ordre de H.
D'où la formule suivante :

$$|G/H| = \frac{|G|}{|H|}$$

Corollaire (**Condition nécessaire et suffisante**) . Soit G un groupe et H un sous-groupe de G. H est distigué dans G ssi H est le noyau d'un morphisme de source G.

Remarque Pour montrer qu'un sous-groupe est distigué, on peut donc simplement vérifier que c'est le noyau de la projection canonique π sur G.

6.2.2 Propriétés du quotient

Théorème () . Soient G et K deux groupes et $H \triangleleft G$. Soient $\phi : G \longrightarrow K$ un morphisme de groupes et π la projection canonique sur G. Alors :

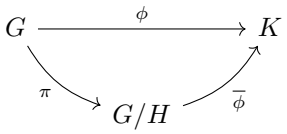
$$H \subset \ker(\phi)$$
$$\iff \phi(H) = \{e_K\}$$

Autrement dit, on peut factoriser le morphisme ϕ "en passant" par G/H. Il existe donc un morphisme de groupes $\overline{\phi} : G/H \longrightarrow K$ tel que $\phi = \overline{\phi} \circ \pi$. Ainsi, $\overline{\phi}$ est unique et définit par :

$$\overline{\phi} : \begin{cases} G/H \longrightarrow K \\ gH \longmapsto \overline{\phi}(gH) = \phi(g) \end{cases}$$

L'image de $\overline{\phi}$ est celle de ϕ et son noyau est $\ker(\phi)$/H.

On peut représenter cette "factorisation" par le diagramme commutatif suivant :



Où $H = \ker(\pi)$.

Dire que le diagramme est commutatif revient à dire que $\phi = \overline{\phi} \circ \pi$

6.3 Théorèmes d'isomorphismes

Admettons que l'on ait un morphisme $\phi : G \longrightarrow K$ entre deux groupes G et K. Mais ϕ n'est pas injectif, notons alors son noyau $H = \ker(\phi)$. Comme vu précédement, H est un sous-groupe de G. On va alors pouvoir construire un isomorphisme de groupe de G/H dans $\mathrm{Im}(\phi)$.

Cela revient à assimiler tous les éléments de G comme des classes d'équivalences et ainsi, ne plus avoir le problème du noyau qui "envoie" plusieurs éléments sur e_K. Pour obtenir un isomorphisme il nous faut donc aussi restreindre le domaine d'arrivée aux éléments atteignables par ϕ.

Théorème (**Premier théorème d'isomorphisme**) . Soit $\phi : G \longrightarrow K$ un morphisme de groupes. Alors il existe un isomorphisme $\overline{\psi}$ tel que :

$$\psi : \begin{cases} G/\ker(\phi) \longrightarrow \mathrm{Im}(\phi) \\ g\ker(\phi) \longmapsto \phi(g) \end{cases}$$

De plus, **si ϕ est surjectif**, ψ est un isomorphisme entre les groupes $G/\ker(\phi)$ et K. (si ϕ est surjectif, on peut étendre l'image de ψ à l'image de ϕ, donc à K tout entier)

Remarque Soit G un groupe monogène engendré par $g \in G$. D'après la propriété universelle de $\mathbb{Z}$, il existe un morphisme $\phi : \mathbb{Z} \twoheadrightarrow \langle g \rangle = G$ tel que $\ker(\phi) = n\mathbb{Z}$.
D'après le théorème 2.2, on peut quotienter l'espace de départ par le noyau du morphisme et ainsi obtenir le morphisme suivant :

$$\overline{\phi} : \begin{cases} \mathbb{Z}/n\mathbb{Z} \longrightarrow \langle g \rangle \\ \overline{k} \longmapsto g^k \in G \end{cases}$$

Enfin, d'après le premier théorème d'isomorphisme, puisque ϕ est surjectif, $\overline{\phi}$ est isomorphisme. D'où :

$$\langle g \rangle \simeq \mathbb{Z}/n\mathbb{Z}$$

Théorème (**Troisième théorème d'isomorphisme**) . Soient $K \subset H \subset G$ trois groupes. Tels que $H \lhd G$ et $K \lhd G$. On a alors :

$$(G/K)/(H/K) \simeq G/H$$

Autrement dit, le quotient de deux groupes quotientés par le même sous-groupe normal est isomorphe au quotient des autres groupes.

Ce théorème mérite quelques petites explications. Notons σ, π, μ les projections canoniques associés aux groupes quotient $G/H, G/K$ et $(G/K)/(H/K)$.
On a donc $K \subseteq H \Longrightarrow \sigma(H) = H/K$, un sous groupe de G/K. Or, d'après le théorème 2.2, il existe un unique morphisme $\overline{\sigma} : G/H \longrightarrow (G/K)/(H/K)$ tel que $\overline{\sigma} \circ \pi = \mu \circ \sigma$.

On a donc le diagramme commutatif suivant :

$$\begin{array}{ccc} G & \xrightarrow{\quad\sigma\quad} & G/K \\ {\scriptstyle\pi}\downarrow & & \downarrow{\scriptstyle\mu} \\ G/H & \underset{\exists!\overline{\sigma}}{\dashrightarrow} & (G/K)/(H/K) \end{array}$$

μ et σ sont surjectifs donc $\mu \circ \sigma$ est surjectif donc $\overline{\sigma}$ est surjectif. De plus, $\ker(\mu \circ \sigma) = H$ donc $\overline{\sigma}$ est injectif. On en conclut que $\overline{\sigma}$ est

Exemple Quelques exemples pour mieux apprécier le théorème précédent :

- A-t-on $(\mathbb{Z}/10\mathbb{Z})/(2\mathbb{Z}/10\mathbb{Z}) \simeq \mathbb{Z}/2\mathbb{Z}$?
- A-t-on $(\mathcal{S}_4/\mathcal{A}_4) \simeq (\mathcal{S}_4/\mathcal{V}_4)(\mathcal{A}_4/\mathcal{V}_4)$?

Chapitre 7

Anneaux

Contents

Résumons, on a pris un ensemble que l'on a muni d'une opération. On a trouvé pleins de propriétés à cette nouvelle structure. Serait-il intéressant de munir notre groupe d'une autre opération ?

7.1 Anneaux, Définitions et Exemples

Définition (**Anneaux**) . Un anneau est un ensemble A muni de deux lois de composition interne généralement notées $+$ et $\times$ tels que :

- $(A, +)$ est un groupe abélien de neutre 0_A.
- $\times$ est associative.
- $\times$ est distributive à gauche et à droite par rapport à $+$.

On notera alors notre anneau $(A, +, \times)$.

Définition (**Anneau Unitaire**) . Soit $(A, +, \times)$ un anneau. On dit que A est unitaire si il possède un neutre 1_A pour la $\times$. Autrement dit, si $\exists 1_A \in A$ tel que :

$$\forall a \in A, \quad 1_A \times a = a \times 1_A = a$$

Définition (**Anneau Commutatif**) . Soit $(A, +, \times)$ un anneau. On dit que A est commutatif si la loi $\times$ est commutative. Autrement dit si :

$$\forall a, b \in A, \quad a \times b = b \times a$$

Exemple Voyons quelques exemples d'anneaux :

1. $(GL_n(\mathbb{R}), +, \times)$ est un anneau unitaire non commutatif.
2. $(\mathbb{Q}, +, \times)$ est un anneau unitaire commutatif.
3. $(2\mathbb{Z}, +, \times)$ est un anneau commutatif non unitaire.

7.1.1 Sous-anneau

De même que pour les groupes, on peut définir la notion de sous-anneau. Très pratique pour montrer qu'un ensemble est un anneau.

Définition (**Sous-anneau**) . Soient $(A, +, \times)$ un anneau et $B \subseteq A$. On dit que B est un sous-anneau de A si

- $(B, +)$ est un sous-groupe de $(A, +)$
- $(B, +)$ est stable pour la loi $\times$ de A.

Tout comme pour les groupes, on peut restreindre ces conditions :

Proposition Soient $(A, +, \times)$ un anneau et $B \subseteq A$. B est un sous-anneau de A ssi

- $0_A \in B$
- $\forall x, y \in B, x + y \in B,\ -x \in B$ et $x \times y \in B$

Remarque Un sous-anneau d'un anneau commutatif est commutatif mais un sous-anneau d'un anneau unitaire n'est as forcément unitaire.

Exemple $(2\mathbb{Z}, +, \times)$ est un sous-anneau de $(Z, +, \times)$. Puisque $\mathbb{Z}$ est commutatif, $2\mathbb{Z}$ l'est aussi, en revanche, puisque $1 \in \mathbb{Z}$ n'est pas pair, $2\mathbb{Z}$ n'est pas unitaire.

7.2 Calculs dans un Anneau

Maintenant que nous avons muni notre groupe d'une nouvelle loi, il va falloir énoncer quelques règles de calculs. Puisque dans un anneau, il n'y a pas forcément d'inverse pour la loi $\times$, on ne pourra pas faire toutes les simplifications que l'on souhaite.

Propriété (**Règles de Calculs dans un Anneau**) . Soient $a, b \in (A, +, \times)$ et $n, m \in \mathbb{N}$. On a les règles suivantes :

i) $0_A \times a = a \times 0_A = 0_A$
ii) $(na)b = a(nb) = n(ab) = nab$
iii) $(na)(mb) = (nm)(ab)$

Supposons maintenant A unitaire :

i) $na = (n1_A)a = a(n1_A)$
ii) $(-1_A)^{2n} = 1_A$ et $(-1_A)^{2n+1} = -1_A$

Démonstration Soient $(A, +, \times)$ un anneau, $a, b \in A$ et $n, m \in \mathbb{N}$.

i) $0_A \times a = a \times 0_A = 0_A$

$$\begin{cases} (b + (-b)) \times a = ba + (-ba) = 0_A \\ a \times (b + (-b)) = ab + (-ab) = 0_A \end{cases}$$

ii) $(na)b = a(nb) = n(ab) = nab$

$$\begin{aligned} (na)b &= (a + \cdots + a)b \\ &= ab + \cdots + ab \\ &= n(ab) \end{aligned}$$

De plus, $ab + \cdots + ab = n(ab)$. Par associativité on a : $(na)b = nab$

iii) $(na)(mb) = (nm)(ab)$

$$\begin{aligned} (na)(nb) &= (\underbrace{a + \cdots + a}_{n \text{ fois}})(\underbrace{b + \cdots + b}_{m \text{ fois}}) \\ &= (\underbrace{ab + \cdots + ab}_{m \text{ fois}}) + \cdots + (\underbrace{ab + \cdots + ab}_{m \text{ fois}}) \\ &= n(\underbrace{ab + \cdots + ab}_{m \text{ fois}}) \\ &= (nm)(ab) \end{aligned}$$

iv) $na = (n1_A)a = a(n1_A)$: Supposons que $1_A \in A$.

$$\begin{cases} (n1_A)a = (\underbrace{1_A + \cdots + 1_A}_{n \text{ fois}})a = a + \cdots + a = na \\ a(n1_A) = a(\underbrace{1_A + \cdots + 1_A}_{n \text{ fois}}) = a + \cdots + a = na \end{cases}$$

Remarque Soit $(A, +, \times)$ un anneau. Si A est unitaire et que $1_A = 0_A$ alors $A = \{0_A\}$.

Démonstration Soit A un anneau unitaire tel que $0_A = 1_A$. Soit $x \in A$. Alors :

$$1_A x = 0_A x = x = 0_A$$

Donc $x = 0_A$ d'où $A = \{0_A\}$.

Proposition (Binôme de Newton) Soient $(A, +, \times)$ un anneau et $a, b \in A$ tels que $a \times b = b \times a$ on a alors :

$$\forall n \in \mathbb{N}^*, \quad (a+b)^n = \sum_{k=0}^{n} \binom{k}{n} a^k b^{n-k} = \sum_{k=0}^{n} \binom{k}{n} b^{n-k} a^k$$

7.3 Inverses dans un Anneau

Comme nous l'avons déjà vu, un anneau ne contient pas forcément d'inverses pour la loi $\times$. Cela peut être le cas dans certaines situations. Soit $(A, +, \times)$ un anneau unitaire non nul.

Définition (**Elements Inversibles**) . Un élément $x \in A$ est dit inversible dans A s'il existe $y \in A$ tel que $xy = yx = 1_A$. Dans ce cas, son inverse est unique et noté x^{-1}. On notera $\mathcal{U}(A)$ l'ensemble des éléments inversibles de A. Plus formellement :

$$\boxed{\mathcal{U}(A) = \{x \in A \mid \exists y \in A, xy = yx = 1_A\}}$$

Remarque $\mathcal{U}(A)$ est non vide puiqu'il contient 1_A d'inverse lui-même.

Théorème (**Structure des inverssibles**) . Soit $(A, +, \times)$ une anneau unitaire non vide, l'ensemble $\mathcal{U}(A)$ est un groupe pour la loi $\times$ de A de neutre 1_A.

Exemple Dans l'anneau $A = (\mathcal{M}_n(\mathbb{R}), +, \times)$ on a $\mathcal{U}(A) = GL_n(\mathbb{R})$.

7.4 Anneau Intègre et Diviseurs de zéro

Définition (**Diviseur de zéro**) . Soit $(A, +, \times)$ un anneau. On dit qu $x \in A$ est un diviseur de 0_A si :

- $x \neq 0_A$
- $\exists y \in A/\{0_A\}$, tel que $xy = 0_A$ ou $yx = 0_A$

On appelle **anneau intègre** tout anneau unitaire commutatif qui n'admet pas de diviseurs de zéro. Plus formellement, un anneau $(A, +, \times)$ est intègre si :

$$\forall x, y \in A, \quad xy = 0_A \Longrightarrow x = 0 \text{ ou } y = 0$$

Exemple $(\mathbb{Z}, +, \times)$ est intègre. En revanche $(\mathcal{M}_2(\mathbb{R}), +, \times)$ n'est pas intègre (utiliser les matrices nilpotentes).

Proposition Tout sous-anneau unitaire et non nul d'un anneau intègre est intègre.

Proposition Soit $(A, +, \times)$ un anneau unitaire. Si $x \in \mathcal{U}(A)$ alors x n'est pas un diviseur de zéro.

Démonstration Soit A un anneau unitaire non trivial. Raisonnons par l'absurde. Soit $x \in \mathcal{U}(A)$ un diviseur de zéro. Alors $\exists y \in A$ non nul tel que :

$$xy = 0_A \Longrightarrow x^{-1}xz = z = 0$$

Par hypothèse $z \neq 0_A$ donc x n'est pas un diviseur de zéro.

L'intégrité d'un anneau nous permet de démontrer une nouvelle règle de simplification :

Propriété (**Simplification dans un anneau intègre**) . Soit A un anneau intègre alors $\forall a \in A/\{0_A\}, \forall c, d \in A$, on a :

$$ab = ac \Longrightarrow b = c$$

Démonstration Soient A un anneau intègre, $a \in A^*$ et $b, c \in A$. Tels que $ab = ac$. Alors :

$$ab = ac \iff ab - ac = 0_A \iff a(b - c) = 0_A \iff a = 0 \text{ ou } b - c = 0_A$$

Par hypothèse $a \neq 0_A$ donc $b - c = 0_A$
D'où : $b = c$

7.5 Idéaux

Faisons la blague tout de suite, c'est pas idéal comme cours.

Soit $(A, +, \times)$ un anneau unitaire.

7.5.1 Définitions et premiers théorèmes

Définition (**Idéal**) . Une partie $I \subseteq A$ est un idéal à gauche de A si

i) $(I, +)$ est un groupe abélien

ii) I est absorbant à gauche dans A

$$\text{i.e} \quad \forall a \in A, \forall x \in I, \quad a.x \in I$$

Si un idéal est absorbant à gauche et à droite, on parle d'idéal bilatère. Dans un anneau commutatif, tous les idéaux sont bilatères.

Remarque Les ensembles $\{0_A\}$ et A sont des idéaux triviaux de A.

Proposition Une partie $I \subseteq A$ est un idéal de A ssi

i) $0_A \in I$

ii) $\forall x, y \in I, \quad x + y \in I$

iii) $\forall a \in A, \forall x \in I, \quad a.x \in I$

Théorème (**Idéaux de** $\mathbb{Z}$) . Les idéaux de $\mathbb{Z}$ sont exactement les $n\mathbb{Z}$ avec $n \in \mathbb{N}$.

Démonstration Soit $n \in \mathbb{N}$. On a déjà montré que les $(n\mathbb{Z}, +)$ sont des sous-groupes de $\mathbb{Z}$. Il suffit de montrer que ce $n\mathbb{Z}$ est bien un idéal de $\mathbb{Z}$.
Soit $k \in \mathbb{Z}$ alors $\forall i \in n\mathbb{Z}$ on a :

$$k \times i \times n\mathbb{Z} \in n\mathbb{Z}$$

Donc c'est bien un idéal.

Théorème (**C.N.S d'égalité anneau/idéal**) . Soit $I \subseteq A$ un idéal de A. On a alors :

$$I = A \iff I \cap \mathcal{U}(A) \neq \emptyset$$

Démonstration Soit $(A, +, \times)$ un anneau unitaire et I un idéal de A.

$\boxed{\Rightarrow}$ $1_A \in A = I$ or $1_A \times 1_A = 1_A$ donc $(1_A)^{-1} = 1_A$. Donc $I \cap \mathcal{U}(A) \neq \emptyset$

$\boxed{\Rightarrow}$ Soit $a \in I \cap \mathcal{U}(A)$ alors $\exists a^{-1} \in A$ tel que $a^{-1}a = 1_A$. Par absorption, $a^{-1}a = 1_A \in I$. Or $\forall a \in A, a \times 1_A = a \in A$. Donc $I = A$.

7.5.2 Caractéristiques des Idéaux

Proposition L'intersection finie d'idéaux est d'un anneau A est un idéal de A.

A partir de cette proposition, pour toute partie $X \subseteq A$ on sait que l'intersection de tous les idéaux de A contenant X est un idéal de A.

Définition (**Idéal engendré par une partie**) . Soit $(A, +, \times)$ un anneau. L'idéal engendré par X est l'intersection de tous les idéaux contenant X. On le note (X). On pourra noter $(X)_A$ pour bien identifier de quel anneaux (X) est un idéal.

Remarque Soit $X \subseteq A$ alors (X) est le plus petit idéal de A contenant X. Plus formellement, pout tout idéal J de A contenant X alors $(X) \subseteq J$.

Soit $X = \{x_1, \ldots, x_p\} \subseteq A$ avec $p \in \mathbb{N}$. L'idéal (X) est noté $x_1, \ldots, x_p$. On dit que $\forall i \in [\![1, p]\!]$, les x_i sont des générateurs de X.

7.5.3 Idéal Principal

Définition (**Idéal Principal**) . Soit A un anneau et $I \subseteq A$ un idéal de A. On dit que I est principal s'il est engendré par un set élément. Autrement dit s'il existe $x \in A$ tel que $X = (x)$.
On dit que A est principal s'il est intègre et que tous ses idéaux sont principaux.

Proposition Soit A un anneau commutatif, unitaire et $x \in A$ alors l'idéal de A engendré par (x) est de la forme :

$$(x) := \{ax \mid a \in A\}$$

Démonstration Montrons que $(x) := \{ax \mid a \in A\}$ par double inclusion.

$\boxed{\subseteq}$ L'ensemble $\{ax \mid a \in A\}$ est un idéal de A et contient (x) on a donc $(x) \subseteq \{ax \mid a \in A\}$ par minimalité.

$\boxed{\supseteq}$ Soit $ax \in \{ax \mid a \in A\}$ alors $ax \in (x)$ par absorption et car (x) est un idéal qui contient x.

Théorème (**Anneau Principal et** $\mathbb{Z}$) . $\mathbb{Z}$ est un anneau principal.

La preuve s'appuit sur le fait que $n\mathbb{Z} = (n)$.

Définition (**Somme d'idéaux**) . La somme d'idéaux est un idéal. Autrement dit, pour tout $I, J \subseteq A$ des idéaux de A, alors l'ensemble

$$I + J = \{x + y \mid x \in I, y \in J\} \subseteq A$$

Est un idéal de A.

7.6 Morphismes d'Anneaux

Définition (**Morphisme d'Anneaux**) . Soient A et B deux anneaux unitaires. Une application $f : A \longrightarrow B$ est un morphisme d'anneaux si

- f est un morphisme du groupe $(A, +)$ dans $(B, +)$.
- $\forall x, y \in A, \quad f(x \times y) = f(x) \times f(y)$
- $f(1_A) = 1_B$

Si f est, de plus bijective, on dira que c'est un **isomorphisme d'anneaux**.

Proposition Une fois que l'on a un mosphisme d'anneaux $f : (A, +, \times) \longrightarrow (B, +, \times)$, alors f induit un morphisme de groupes entre $(\mathcal{U}(A), \times)$ et $(\mathcal{U}(B), \times)$. Remarquons bien que nous ne parlons pas du morphisme de groupe de la définition de morphisme d'anneaux.

Définition (**Image et Noyau**) . Soit $f : (A, +, \times) \longrightarrow (B, +, \times)$ un mosphisme d'anneaux. On appelle image et noyau, l'image et le noyau du morphisme de groupes $f : (\mathcal{U}(A), \times) \longrightarrow (\mathcal{U}(B), \times)$.

Remarque On peut donc conclure de cette dernière définition que f est injective ssi $\ker f = \{0_A\}$.

Propriété (**Structure de l'image et du noyau**) . Soit $f : (A, +, \times) \longrightarrow (B, +, \times)$ un morphisme d'anneaux. Alors $\ker f$ est un idéal de A et $Im(f)$ est un sous-anneau de B.

Démonstration

Proposition La composée de deux morphismes d'anneaux est un morphisme d'anneaux. La réciproque d'un morphisme d'anneaux bijectif est aussi un morphisme d'anneaux.

7.7 Anneau Quotient

Théorème (**Quotient d'un anneau par un idéal**) . Soit A un anneau unitaire non trivial et I un idéal bilatère de A distinct de A. Soit A/I l'ensemble des classes à gauche du groupe $(A,+)$ modulo le sous-groupe I. On a les propriétés suivantes :

i) Il existe une loi de compisition interne sur A/I, notée $\times$ telle que :

$$\forall a,b \in A,\ \overline{a \times b} = \overline{a} \times \overline{b}$$

ii) Si $+$ désigne l'addition quotient du groupe quotient A/I alors $(A/I,+,\times)$ est un anneau unitaire non trivial. On l'appelle anneau-quotient de A par I.

Proposition L'anneau quotient d'un anneau commutatif est commutatif (pour la loi $\times$).

Proposition (Projection Canonique) Soit $(A/I,+,\times)$ un anneau quotient, alors l'application :

$$\pi : \begin{cases} A \longrightarrow A/I \\ a \longmapsto \overline{a} \end{cases}$$

est un morphisme d'anneaux surjectif de noyau I.

Théorème (**Isomorphisme**) . Soit $\phi : A \longrightarrow B$ un morphisme d'anneaux. On a alors :

$$\boxed{A/\ker\phi \simeq Im(\phi)}$$

7.8 L'anneau $\mathbb{Z}/n\mathbb{Z}$

Définition (**Anneau** $\mathbb{Z}/n\mathbb{Z}$) . Soit $n \in \mathbb{N}, n \geqslant 2$. L'anneau des entiers modulo n est l'anneau quotient de $\mathbb{Z}$ par on idéal $n\mathbb{Z}$, noté $\mathbb{Z}/n\mathbb{Z}$.

Théorème (**Inversibles dans** $\mathbb{Z}/n\mathbb{Z}$) . Soit l'anneau quotient $(\mathbb{Z}/n\mathbb{Z},+,\times)$. Les élements inversibles de cet anneau sont exactement les classes dont les représentant sont premiers avec n. Autrement dit :

$$\mathcal{U}(\mathbb{Z}/n\mathbb{Z}) = \{\overline{x} \mid x \in [\![1,n]\!] \text{ et } pgcd(1,n) = 1\}$$

Théorème (**Fondamental**) . Soit $n \geqslant 2$. On a :

$$\boxed{n \text{ premier} \iff \mathbb{Z}/n\mathbb{Z} \text{ est un corps} \iff \mathbb{Z}/n\mathbb{Z} \text{ est intègre}}$$

Ainsi, lorsque $n \geqslant 2$ est premier, on note $\mathbb{Z}/n\mathbb{Z}$, $\mathcal{F}_p$.

Proposition Les diviseurs de zéro dans l'anneau $\mathbb{Z}/n\mathbb{Z}$ sont exactement les classes dont les représentants divisent n et qui sont strictement compris entre 1 et n. Autrement dit,

$$\overline{d} \in \mathbb{Z}/n\mathbb{Z} \text{ est un diviseur de zéro} \iff 1 < d < n \text{ et } d|n$$

Théorème (**Chinois**) . Soient $n, m \in \mathbb{N}$ supérieurs à 2. Alors :

$$pgcd(n,m) = 1 \iff \mathbb{Z}/nm\mathbb{Z} \simeq (\mathbb{Z}/n\mathbb{Z}) \times (\mathbb{Z}/m\mathbb{Z})$$

Théorème (**Petit théorème de Fermat**) . Soit $p \in \mathbb{N}$ premier. Pour tout $x \in \mathbb{Z}$, on a $x^p \equiv x \mod p$.

Chapitre 8

Corps

Contents

Résumons, nous avons un groupe avec une loi supplémentaire qui peut être commutative et posséder un neutre. Grâce à cela, nous avons pu définir les idéeaux, un nouveau type de morphisme, etc...
Seulement, on souhaiterai pouvoir définir des structures plus complexes comme avec les groupes tels que les anneaux quotient, avoir une notion de divisibilité, bref, de l'arithmétique.
Pour cela, nous avons besoin d'étendre la notion d'anneaux aux corps...

8.1 Corps, définition et propriétés

Définition (**Corps**) . On appelle $(K, +, \times)$ un corps si $(K, +, \times)$ est un anneau et $(K^*, \times)$ un groupe. Si la loi $\times$ est commutative, on parlera alors de corps commutatif.

On peut bien évidement trouver une condition nécessaire et suffisante pour qu'un anneau soit un corps...

Proposition Soit $(K, +, \times)$ un anneau unitaire. K est un corps si tout élément non nul est inversible pour $\times$. Plus formellement ssi :

$$\forall x \in K^*,\ \exists y \in K,\ xy = yx = 1_K \quad \text{ssi}\, \mathcal{U}(K) = K^*$$

Proposition Un corps commutatif est aussi un anneau intègre.

Tout comme pour les groupes et les anneaux, on peut définir un sous-corps d'un corps.

Définition (**Sous-corps**) . Soit K un corps. Un partie $L \subseteq K$ est un sous-corps de K si :

i) L est un sous-anneau de K

ii) $1_K \in L$

iii) $\forall x \in L, x^{-1} \in L$

En pratique pour montrer qu'un anneau est un corps, on préfèrera montrer que c'est un sous-corps quand c'est possible.

Exemple $(\mathbb{C}, +, \times)$ est un corps. $(\mathbb{Q}, +, \times)$ et $(\mathbb{R}, +, \times)$ en sont des sous-corps.

8.2 Anneau Quotient

Dans cette section, on considère un anneau A unitaire, non trivial. On va quotienter A par un de ses idéaux I bilatère différent de A. On a donc le théorème suivant :

Théorème (**Structure de l'anneau quotient**) . Soit A un anneau unitaire, non trivial et I un idéal bilatère de A différent de A. L'ensemble A/I est définit comme l'ensemble des classes du groupe $(A, +)$ modulo le sous-groupe I de A. On définit des opérations dessus :

- **Loi multiplicative :** $\forall a, b \in A, \overline{a} \times \overline{b} = \overline{a \times b}$.
- **Loi additive :** $\forall a, b \in A, \overline{a} + \overline{b} \in A/I$.

Alors le groupe $(A, +)$ peut être étendu en un anneau $(A, +, \times)$ unitaire et non nul appelé **anneau quotient de A par I**.

Proposition Si A est commutatif, alors A/I est aussi commutatif.

Comme pour les groupes, on peut définir la projection canonique, porte d'entrée pour les théorèmes d'isomorphismes...

Définition (**Projection Canonique**) . Soit A un anneau et A/I son anneau quotient par I. On définit la projection canonique par :

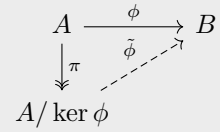
$$\pi : \begin{cases} A \longrightarrow A/I \\ a \longrightarrow \overline{a} \end{cases}$$

qui à chaque élément de A lui associe sa classe dans A/I. Elle "projette" les éléments de A sur leur classe. Cette application est :

- un morphisme d'anneaux
- surjective
- de noyau I

Théorème (**Isomorphisme**) . Soient A, B deux anneaux et $\phi : A \longrightarrow B$ une morphisme d'anneaux. On peut alors construire un isomorphisme d'anneaux $\overline{\phi}$ tel que :

$$A/\ker\phi \simeq Im\phi$$



8.3 L'anneau $\mathbb{Z}/n\mathbb{Z}$

Définissons maintenant proprement l'anneau $\mathbb{Z}/n\mathbb{Z}$ en nous servant du fait que $\mathbb{Z}$ est un anneau et $n\mathbb{Z}, \forall n \in \mathbb{N}$ est un de ses idéaux.

Définition (**Anneau** $\mathbb{Z}/n\mathbb{Z}$) . Soit $n \in \mathbb{N}, n \geqslant 2$. On appelle anneau des entiers modulo n l'anneua quotient de $\mathbb{Z}$ par son idéal $n\mathbb{Z}$ noté $\mathbb{Z}/n\mathbb{Z}$.

Théorème (**Inversibles dans** $\mathbb{Z}/n\mathbb{Z}$) . Le groupe des inversibles de $\mathbb{Z}/n\mathbb{Z}$ est :

$$\boxed{\mathcal{U}(\mathbb{Z}/n\mathbb{Z}) = \{\overline{x} \mid x \in [\![1, n]\!] \text{ tq } pgcd(x, n) = 1\}}$$

$(\mathcal{U}(\mathbb{Z}/n\mathbb{Z}), \times)$ est un groupe abélien d'ordre $\phi(n)$ où ϕ est l'indicatrice d'Euler.

Théorème (**Fondamental**) . On a les équivalences suivantes :

$$n \text{ premier} \iff \mathbb{Z}/n\mathbb{Z} \text{ est un corps} \iff \mathbb{Z}/n\mathbb{Z} \text{ est intègre}$$

Ce théorème nous permet de rapidement pouvoir affirmer qu'un $\mathbb{Z}/n\mathbb{Z}$ est un corps. En effet, l'utilisation d'un corps permet d'obtenir des inverses pour la loi $\times$.

Remarque (Notation) Lors que n est premier, le corps $\mathbb{Z}/n\mathbb{Z}$ est noté $\mathcal{F}_p$.

Enonçons maintenant quelques théorèmes en vrac très utiles dans l'étude de la théorie des corps.

Théorème (**Restes Chinois**) . Soient $n, m \in \mathbb{Z}$. Si $pgcd(n, m) = 1$ alors l'application :

$$\varphi : \begin{cases} \mathbb{Z}/nm\mathbb{Z} \longrightarrow \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z} \\ \overline{x} \longmapsto (x \mod n, x \mod m) \end{cases}$$

est un isomorphisme d'anneaux. En particulier :

$$\boxed{pgcd(n, m) = 1 \iff \mathbb{Z}/nm\mathbb{Z} \simeq \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}}$$

Comme vu dans le cadre de la théorie des groupes, nous pouvons aussi généraliser le petit théorème de Fermat.

Théorème (**Petit théorème de Fermat**) . Soit $p \in \mathbb{N}$. Si p est premier alors

$$\forall x \in \mathbb{Z}, \quad x^p \equiv x \mod p$$

8.4 Caractéristique d'un Anneau

Introduisons maintenant la caractéristique d'un anneau, concept fondamental en théorie des anneaux et en algèbre commutative.

Définition (**Caractéristique d'un anneau**) . Soit A un anneau unitaire non trivial. La **caractérisitique** de A est le plus petit entier naturel $n \in \mathbb{N}$ tel que $n \times 1_A = 0_A$. On la note $car(A)$. Si cet entier n'existe pas, alors $car(A) = 0_{\mathbb{N}}$.

Remarque La caractéristique d'un anneau permet de distinguer les anneaux divisibles comme $\mathbb{Q}, \mathbb{R}$ ou $\mathbb{C}$ des anneaux cycliques tels que les $\mathbb{Z}/n\mathbb{Z}$. Intuitivement elle "mesure" combien de fois on doit additionner 1 pour obtenir 0.

Proposition On peut définir la caractéristique d'un anneau plus formellement. Soit A un anneau initaire non trivial. Soit f le morphisme d'anneau suivant :

$$f : \begin{cases} \mathbb{Z} \longrightarrow A \\ k \longmapsto k \times 1_A \end{cases}$$

La caractéristique de A est donc l'unique entier naturel $n \in \mathbb{N}$ tel que :

$$\ker f = n\mathbb{Z}$$

On a alors $\forall x \in A$ que $nx = 0_A$.

Théorème (**Caractéristique des anneaux usuels**) .

- $car(\mathbb{Z}) = 0$
- $car(\mathbb{Z}/n\mathbb{Z}) = n$
- Un anneau de caractéristique nulle est infini.

Théorème (**Caractérisation des anneaux intègres**) . Soit A un anneau intègre. Alors $car(A) = 0$ ou $car(A) = p$ avec p premier.

8.5 Corps des fractions

Soit A un anneau commutatif, on cherche à construire une extension de A pour laquelle tout élément de A contient un inverse pour la loi $\times$ usuelle de A. Pour cela, nous allons être amenée à construire le corps des fractions de A.

En guise d'exemple introductif, on peut considérer l'anneau $(\mathbb{Z}, +, \times)$. Dans cet anneau, 2 ne possède pas d'inverse pour la loi $\times$. Pour cela, on peut "construire" un nouvel anneau (ici $\mathbb{Q}$), dans lequel $\frac{1}{2} \in \mathbb{Q}$ sera l'inverse de 2 pour la loi $\times$.

Proposition Soit A un anneau commutatif. On cherche à construire une structure pour dans laquelle tout élément de A possède inverse pour la loi $\times$. On définit ainsi le **corps des fractions** de A comme l'ensemble :

$$\boxed{Fr(A) = \left\{ \frac{a}{b} \mid a, b \in A \text{ et } b \neq 0_A \right\}}$$

Dans cette structure, ces objets sont définis comme des classes d'équivalences sous la forme :

$$\frac{a}{b} = \frac{c}{d} \iff ad = bc$$

Cette forme d'égalité garantit la compatibilité des opérations/propriétés arithmétiques déjà utilisées. Ainsi tout élément $b \in A$ qui ne possédait pas d'inverse dans A en possède maintenant un dans $Fr(A)$ tel que :

$$\forall b \in A, \exists \frac{1}{b} \in Fr(A), \quad b \times \frac{1}{b} = \frac{1}{b} \times b = 1_A$$

Chaque élément non nul de A possède maintenant un inverse pour $\times$ ce qui fait de $Fr(A)$ un corps.

Propriété (**Calculs**) . Abordons quelques propriétés de calculs dans $Fr(A)$. Soient $a, c \in A$ et $k, b, d, u, v \in A$ non nuls. On a les propriétés suivantes dans $Fr(A)$:

- $\frac{ka}{kb} = \frac{a}{b}$
- $\frac{a}{b} \times \frac{c}{d} = \frac{ac}{bd}$
- $\frac{a}{b} + \frac{c}{d} = \frac{ad+cb}{bd}$
- $\left\{\frac{u}{v}\right\}^{-1} = \frac{v}{u}$

Proposition Depuis le début, nous supposons que $Fr(A)$ est bien définit que tout se passe bien. Or, pour que $Fr(A)$ soit bien définit, il faut que l'on puisse "plonger" A dans son corps des fractions. On définit ainsi le morphisme :

$$\varphi : \begin{cases} A \longrightarrow Fr(A) \\ a \longmapsto \frac{a}{1_A} \end{cases}$$

C'est un morphisme d'anneaux par propriétés sur les fractions.
Pour "plonger" A dans $Fr(A)$, il faut que φ soit injectif.

$$\text{i.e } \forall a, b \in Fr(A), \quad \varphi(a) = \varphi(b) \Longrightarrow a = b$$

Autrement dit dans la cas des fractions, si $\frac{a}{1} = \frac{b}{1}$ alors $a = b$. Or si A n'est pas **intègre**, il peut contenir des diviseurs de zéro a et b tels que :

$$a \neq 0 \text{ et } b \neq 0 \ \text{ mais } \ ab = 0_A$$

En terme de fractions, cela se traduit par :

$$\exists a, b \in A, \ \text{tq} \ \ \frac{a}{b} = \frac{0}{b} = 0$$

Il se pourrait donc que φ ne soit pas injectif. Ainsi $Fr(A)$ ne serait pas définit.

Propriété (**Intégrité et corps des fractions**) . Soit A un anneau unitaire non nul et $Fr(A)$ son corps des fractions. Si il existe une morphisme d'anneau injectif $\varphi : A \longrightarrow Fr(A)$ injectif, alors nécessairement, A est intègre.

Théorème (**Existence corps des fractions**) . Tout anneau commutatif intègre A admet un corps de fractions :

$$Fr(A) = \left\{ \frac{a}{b} \mid a, b \in A, b \neq 0_A \right\}$$

Chapitre 9

Arithmétique dans les Anneaux

Contents

Une fois les corps définis (voir chap précédent), on peut parler d'arithmétique dans les anneaux. On va essayer dans ce chapitre de généraliser aux anneaux toutes les propriétés que l'on connaît sur les nombres relatifs usuels. L'objectif sera donc ensuite de pouvoir appliquer cette théorie à n'importe quel anneau.
L'étude des anneaux remonte aux travaux de Richard Dedekind (1831-1916) et David Hilbert (1862-1943), qui ont exploré les idéaux pour résoudre des problèmes en théorie des nombres. Ces concepts ont été développés pour comprendre la factorisation dans des extensions des entiers, notamment dans le cadre du dernier théorème de Fermat. L'approche moderne a été formalisée par Emmy Noether au début du XX° siècle.

Dans tout ce chapitre, on se place dans une anneau A unitaire, intègre et commutatif.

9.1 Divisibilité

Revenons aux fondamentaux...

Définition (**Divisibilité**) . Soient $a, b \in A$ on dit que a divise b (noté $a|b$) si il existe $c \in A$ tel que $ac = b$.
On note $\mathcal{D}(a)$ l'ensemble des diviseurs de $a \in A$ dans A :

$$\mathcal{D}(a) = \{d \in A \ \text{ tq } d|a\}$$

Proposition (Unicité) Soient $a, b \in A$. Si $b|a$ et $b \neq 0_A$ alors il existe un unique $q \in A$ appelé **quotient** tel que $a = bq$.

Propriété (**Divisibilité et Idéaux**) . Soient $a, b \in A$. On a alors l'implication suivante :

$$a|b \Longrightarrow (b) \subseteq (a)$$

Définition (**Éléménts associés**) . Soient $a, b \in A$ on dit que a et b sont **associés** si :

$$a|b \quad \text{et} \quad b|a$$

L'association pour la divisibilité est une relation d'équivalence noté $a\mathcal{R}b$.

Proposition Soient $a, b \in A$, on a alors :

$$a\mathcal{R}b \iff \exists u \in \mathcal{U}(A), a = ub \iff (a) = (b)$$

9.2 PGCD, PPCM et éléments irréductibles

Définition (**PGCD**) . Soit $(a_i)_{i\in I} \in A$ une famille dénombrable d'éléments de A. On définit le PGCD de (a_i) dans A comme le plus grand diviseur commun de tous les $a_i, i \in I$. Plus formellement, s'il existe :

$$\boxed{d = pgcd((a_i)_{i\in I}) \iff \begin{cases} \forall i \in I, d|a_i \\ \forall \delta \in A, \forall i \in I, \delta|a_i \implies d|\delta \end{cases}}$$

Attention : dans notre cadre le PGCD d'un famille d'éléments de A n'existe pas forcément.

Proposition Soit $(a_i)_{i\in I} \in A$ une famille dénombrable d'éléments de A. Le PGCD de la famille $(a_i)_{i\in I}$ est unique à un inversible près. Autrement dit,

$$\forall d, d' \in A, \begin{cases} d = pgcd((a_i)_{i\in I}) \\ d' = pgcd((a_i)_{i\in I}) \end{cases} \implies d\mathcal{R}d'$$

Proposition Soient $a, b, d \in A$. On a l'implication suivante :

$$(a) + (b) = (d) \implies d = pgcd(a, b)$$

Ainsi, si on trouve un générateur de la somme de deux idéaux, alors on trouve un PGCD à un inversible près.

Définition (**Éléments premiers entre eux**) . Soit $(a_i)_{i\in I} \in A$ une famille dénombrable d'éléments de A. On dit que tous les a_i sont premiers entre eux si

$$\forall i \in I, \forall j \in I, i \neq j, \quad pgcd(a_i, a_j) \text{ existe et } pgcd(a_i, a_j) \in \mathcal{U}(A)$$

Définition (**PPCM**) . Soit $(a_i)_{i\in I} \in A$ une famille dénombrable d'éléments de A. On dit que $(a_i)_{i\in I}$ admet un plus petit commun multiple dans A s'il existe un élément dans A multiple de tous les a_i et étant le plus petit multiple. Plus formellement :

$$\boxed{m = ppcm((a_i)_{i\in I}) \in A \iff \begin{cases} \forall i \in I, a_i|m \\ \forall \mu \in A, \forall i \in I, a_i|\mu \implies m|\mu \end{cases}}$$

Tout comme le PGCD, le PPCM est définit à un inversible près.

Proposition (Caractérisation des PPCM) On peut caractériser le PPCM de deux éléments à partir des idéaux de ces éléments. Ainsi, soient $a, b, m \in A$. On a alors :

$$\boxed{m = ppcm(a, b) \iff (m) = (a) \cap (b)}$$

Définition (**Élémént irréductible**) . Soit $a \in A$. On dit que p est irréductible dans A si il n'existe aucune décomposition de p dans A en éléments non inversibles. Plus formellement :

$$p \in A \text{ est irréductible } \iff \begin{cases} p \neq 0_A \text{ et } p \notin \mathcal{U}(A) \\ \forall d_1, d_2 \in A, p = d_1 \times d_2 \Longrightarrow d_1 \in \mathcal{U}(A) \textbf{ ou } d_2 \in \mathcal{U}(A) \end{cases}$$

Exemple Les irréductibles de $(\mathbb{Z}, +, \times)$ sont exactement les nombres premiers de $\mathbb{Z}$.

9.3 Divisibilité dans les anneaux principaux

Depuis le début de ce chapitre nous nous plaçons dans un cadre relativement général. Ce cadre, quoique très intéressant, ne nous permet pas d'avoir de "bonnes" propriétés pour la divisibilité. Ainsi, nous allons ici nous placer dans des anneaux principaux.

Remarque (Rappel) Un anneau A est dit principal si tout idéal peut être engendré par un seul élément de A.

Théorème (**Existence PGCD et PPCM**) . Soit A un anneau **principal**. Soit $(a_i)_{i\in I}$ une famille d'éléments de A. Alors le PGCD et le PPCM de $(a_i)_{i\in I}$ existe et de plus :

- $pgcd((a_i)_{i\in I}) = d$ tel que $d \in A$ et $(d) = (\{a_i \mid i \in I\})$
- $ppcm((a_i)_{i\in I}) = m$ tel que $m \in A$ et $(m) = \bigcap_{i\in I}(a_i)$

Théorème (**Bézout**) . Soit A un anneau principal et $a, b \in A$. On a alors :

$$\boxed{pgcd(a,b) = d \in A \iff \exists u, v \in A,\ au + bv = d}$$

En particulier :

$$pgcd(a,b) = 1 \iff \exists u, v \in A,\ au + bv = 1$$

Corollaire (**Bézout**) . Soit A un anneau principal et $a, b, \alpha, \beta \in A$. Alors :

$$a = d\alpha \text{ et } b = d\beta \text{ et } pgcd(\alpha, \beta) = 1 \Longrightarrow pgcd(a,b) = d$$

Théorème (**Gauss**) . Soit A un anneau principal et $a, b, c \in A$. On a :

- $pgcd(a,b) = 1$ et $a|bc \Longrightarrow a|c$
- p irréductible dans A et $p|ab \Longrightarrow p|b$

Définition (**Anneau Euclidien**) . On dit que A est un anneau euclidien s'il est possible d'y définir une division euclidienne.

Théorème (**Anneau euclidien, conséquences**) . Tout anneau euclidien est principal.

9.4 Anneaux Factoriels

Définition (**Anneau Factoriel**) . Un anneau A est dit factoriel s'il est intègre et si tout élément se factorise en éléments inversibles.

Théorème (**Factorialité des anneaux principaux**) . Tout anneau principal est factoriel.

Chapitre 10

Introduction à la théorie des corps

Contents

Dans ce chapitre, nous allons plonger plus profondément dans la théorie des corps en abordant les extensions de corps.
Dans tout ce chapitre, sauf mention contraire, nous considérons $\mathbb{K}$ comme un corps commutatif.

10.1 Extension de corps

Définition (**Extension de corps**) . Soit $\mathbb{K}$ un corps. On dit que $\mathbb{K}$ est une extension du corps $\mathbb{F}$ si $\mathbb{K}$ est un sous corps de $\mathbb{F}$. On la note $\mathbb{K}/\mathbb{F}$.

Exemple $\mathbb{C}$ est donc une extension de corps de $\mathbb{R}$ de même que $\mathbb{R}$ est une extension de corps de $\mathbb{Q}$.

10.2 Éléments algébriques et transcendants

Définition (**Éléments algébrique**) . Soit $\mathbb{K}/\mathbb{F}$ deux corps. Soit $\alpha \in \mathbb{K}$, on dit que α est algébrique dans $\mathbb{F}$ s'il est racine d'un polynôme P à coefficients dans $\mathbb{F}$ (i.e $P(\alpha) = 0$).

Exemple Ainsi, i est algébrique dans $\mathbb{C}$ puisqu'il est racine de $X + 1 \in \mathbb{R}[X]$ où $\mathbb{C}/\mathbb{R}$.

On peut remarquer que tout éléments d'un corps $\mathbb{K}$ est algébrique sur $\mathbb{K}$ car il est racine du polynôme $X - \alpha$.

Définition (**Élément Transcendant**) . Soit $\mathbb{K}/\mathbb{F}$ deux corps. Soit $\alpha \in \mathbb{K}$, on dit que α est transcendant sur $\mathbb{F}$ s'il n'est pas algébrique sur $\mathbb{F}$. Autrement dit, α est transcendant sur $\mathbb{F}$ ssi

$$\forall P \in \mathbb{F}[X], P(\alpha) \neq 0$$

Il est important de toujours spécifier dans quel corps on se trouve. En effet, un corps peut avoir une infinité de sous-corps, dont chacun a des propriétés différentes. Ainsi un élément transcendant/algébrique dans l'un peut ne pas l'être dans l'autre.

10.3 Polynômes et isomorphismes

10.3.1 Polynôme Minimal

Soient deux corps $\mathbb{K}/\mathbb{F}$. Soit $\alpha \in \mathbb{K}$. Soit le morphisme de substitution suivant :

$$\phi_\alpha : \begin{cases} \mathbb{F}[X] \longrightarrow \mathbb{K} \\ P \longmapsto P(\alpha) \end{cases}$$

Distinguons deux cas :

- Supposons que α est transcendant sur $\mathbb{F}$, alors $\ker \phi_\alpha = \{O_{\mathbb{F}[X]}\}$ et ϕ_α est injective.
- Supposons que α est algébrique sur $\mathbb{F}$. $\mathbb{F}[X]$ est principal et $\ker \phi_\alpha$ est un de ses idéaux donc il existe $P \in \mathbb{F}[X]$ tel que $P(\alpha) = 0$.

Définition (**Polynôme Minimal**) . Soit $\mathbb{K}$ une extension d'un corps $\mathbb{F}$. Soit $\alpha \in \mathbb{K}$. Les conditions suivantes sont équivalentes. $P_\alpha \in \mathbb{F}$ est appelé **polynôme minimal** de α sur $\mathbb{L}$ si :

P_α est l'unique polynôme unitaire de $\mathbb{F}[X]$ admettant α comme racine.

$\iff$ P_α est irréductible dans $\mathbb{F}[X]$ et α est une racine de P_α.

$\iff$ L'idéal de $\mathbb{F}[X]$ engendré par P_α est maximal.

$\iff$ Si $Q \in \mathbb{L}[X]$ admet α comme racine, alors $P_\alpha | Q$.

Le degré de P_α est appelé **degré de** α.

Proposition (Polynôme Minimal et noyau) Soit $\mathbb{K}$ une extension d'un corps $\mathbb{F}$. Soit $\alpha \in \mathbb{K}$. Soit $P_\alpha \in \mathbb{F}[X]$ le polynôme minimal de α. On peut montrer que :

$$\ker \phi_\alpha = (P_\alpha)$$

Démonstration Soit $\mathbb{K}$ une extension d'un corps $\mathbb{F}$. Soit $\alpha \in \mathbb{K}$. Soit le morphisme d'évaluation suivant :

$$\phi_\alpha : \begin{cases} \mathbb{F}[X] \longrightarrow \mathbb{K} \\ P \longmapsto P(\alpha) \end{cases}$$

Supposons que α est algébrique. Soit P le polynôme minimal de α. Montrons que $\ker \phi_\alpha = (P)$.

$\boxed{\subseteq}$ Soit $(P) = \{QP \mid Q \in \mathbb{F}[X]\}$. Soit $R \in (P)$, on a donc :

$$R(\alpha) = Q(\alpha) \times P(\alpha) = 0$$

donc $R \in \ker \phi_\alpha$. On a donc $(P) \subseteq \ker \phi_\alpha$.

$\boxed{\supseteq}$ Soit $Q \in \ker \phi_\alpha$, alors $Q(\alpha) = 0$.

si $Q = 0_{\mathbb{F}[X]}$ alors $Q \in (P)$

sinon on a :

$$(Q) := \{RQ \mid R \in \mathbb{F}[X]\}$$

or $\forall G \in (Q), G(\alpha) = 0$ donc $(Q) \subseteq (P) \Longrightarrow P|Q$.

d'où $\ker \phi_\alpha \subseteq (P)$.

Par double inclusion, on a donc $\ker \phi_\alpha = (P)$.

10.3.2 Sous-corps engendré

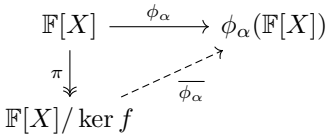
Définition (**Sous-corps engendré**) . Soient $\mathbb{K}/\mathbb{F}$ une extension de corps et $\alpha \in \mathbb{K}$. On définit le sous-corps engendré par $\mathbb{F}$ et α, noté $\mathbb{F}(\alpha)$, comme le plus petit sous-corps de $\mathbb{K}$ contenant α et $\mathbb{F}$.

On peut définir les sous-corps engendré de manière plus globale de la même façon en prenant une famille $(\alpha_1, \ldots, \alpha_k) \in \mathbb{K}$. De plus, comme pour les anneaux, $\mathbb{F}(\alpha)$ est exactement l'intersection de tous les sous-corps de $\mathbb{K}$ contenant α et $\mathbb{F}$.

Proposition Reprenons le morphisme d'évalusation précédant :

$$\phi_\alpha : \begin{cases} \mathbb{F}[X] \longrightarrow \mathbb{K} \\ P \longmapsto P(\alpha) \end{cases}$$

Supposons que $\alpha \in \mathbb{K}$ est algébrique. On a alors $\ker \phi_\alpha = (P_\alpha)$. Or, d'après le théorème d'isomorphisme, on a :



Donc :

$$\mathbb{F}[X]/(P_\alpha) \simeq \phi_\alpha(\mathbb{F}[X]) = \mathbb{F}[\alpha]$$

Or P_α est irréductible sur $\mathbb{F}[X]$ donc $\mathbb{F}[X]/(P_\alpha)$ est un corps. Par isomorphisme, $\mathbb{F}[\alpha]$ est aussi un corps.

Or $\mathbb{F}(\alpha)$ est le plus petit sous-corps de $\mathbb{K}$ contenant α et $\mathbb{F}$ et $\mathbb{F}[\alpha]$ est un sous-corps de $\mathbb{K}$ contenant α et $\mathbb{F}$. Donc c'est le plus petit, d'où $\mathbb{F}[\alpha] = \mathbb{F}(\alpha)$.
D'autre part, si α est transcendant, le quotient $\mathbb{F}[X]/(P_\alpha)$ n'est pas un corps. Donc $\mathbb{F}[\alpha]$ n'en est pas un. D'où $\mathbb{F}[\alpha] \not\subseteq \mathbb{F}(\alpha)$.

Théorème (**Cas transcendant**) . Soit $\alpha \in \mathbb{K}$ et $\mathbb{F} \subset \mathbb{K}$ un sous-corps de $\mathbb{K}$. On a :

$$\begin{aligned} & \alpha \text{ est transcendant sur } \mathbb{F} \\ \iff & \mathbb{F}[X] \simeq \mathbb{F}[\alpha] \\ \iff & \mathbb{F}(\alpha) \neq \mathbb{F}[\alpha] \end{aligned}$$

Théorème (**Cas algébrique**) . Soit $\alpha \in \mathbb{K}$ et $\mathbb{F} \subset \mathbb{K}$ un sous-corps de $\mathbb{K}$. On a :

$$\boxed{\alpha \text{ est algébrique sur } \mathbb{F} \iff \mathbb{F}[X] \simeq \mathbb{F}[\alpha]}$$

10.4 Degré d'une extension

Définition (**Degré**) . Soit $\mathbb{K}/\mathbb{F}$ une extension de corps. On considère $\mathbb{F}$ comme un $\mathbb{K}$ espace vectoriel et on définit le degré de l'extension noté $[\mathbb{K} : \mathbb{F}]$ tel que :

$$[\mathbb{K} : \mathbb{F}] = \dim_{\mathbb{K}}(\mathbb{F})$$

On dira qu'une extension de corps est **finie** si son degré est fini.

Le degré de l'extension de corps sera donc la dimension de $\mathbb{L}$ en tant que $\mathbb{K}$ espace vectoriel.

Exemple Si on prend $\mathbb{F} = \mathbb{Q}(\sqrt{2})$ et $\mathbb{K} = \mathbb{R}$ on a donc $[\mathbb{Q}(\sqrt{2}) : \mathbb{R}] = \dim_{\mathbb{R}}(\mathbb{Q}(\sqrt{2})) = 2$ car une base de $\mathbb{Q}(\sqrt{2})$ sur $\mathbb{R}$ est $(1, \sqrt{2})$.

Lemme Soit $\mathbb{K}/\mathbb{F}$ une extension de corps. On a les propriétés suivantes :

- $[\mathbb{K} : \mathbb{F}] = 1$ si et seulement si $\mathbb{K} = \mathbb{F}$
- $\alpha \in \mathbb{K}$ est de degré 1 sur $\mathbb{F}$ ssi $\alpha \in \mathbb{F}$

Propriété (**Degré et éléments algébriques**) . Soit $\mathbb{K}/\mathbb{F}$ une extension de corps. Alors :

- $\alpha \in \mathbb{K}$ est algébrique sur $\mathbb{F}$ ssi $[\mathbb{F}(\alpha) : \mathbb{F}]$ est le degré de α sur $\mathbb{F}$.
- $\alpha \in \mathbb{K}$ est algébrique est algébrique sur $\mathbb{F}$ ssi $[\mathbb{F}(\alpha) : \mathbb{F}] < \infty$.

Théorème (**Multiplicativité du degré**) . Soient $\mathbb{F} \subset \mathbb{K} \subset \mathbb{L}$ des corps, alors :

$$\boxed{[\mathbb{L} : \mathbb{F}] = [\mathbb{L} : \mathbb{K}] \times [\mathbb{K} : \mathbb{F}]}$$

En particulier : $[\mathbb{L} : \mathbb{K}] \mid [\mathbb{L} : \mathbb{F}]$ et $[\mathbb{K} : \mathbb{F}] \mid [\mathbb{L} : \mathbb{F}]$.

Chapitre 11

Nombres Complexes

L'idée sous-jascente des nombres complexes et d'étendre $\mathbb{R}$ en un corps $\mathbb{C}$ pour être capable de factoriser les polynômes irréductibles dans $\mathbb{R}$. On va donc introduire des "nombres" dont le carré sera négatif. Pour cela, nous supposerons l'existence d'un $i \in \mathbb{C}$ tel que $i^2 = -1$.

Contents

11.1 Définition et propriétés

En utilisant le chapitre sur la théorie des corps précédent, on peut définir $\mathbb{C}$ comme une extension de corps de $\mathbb{R}$.

Définition (**Corps des complexes**) . On définit l'ensemble des nombres complexes, noté $\mathbb{C}$, comme l'extention de corps du corps $\mathbb{R}$ par $i \in \mathbb{C}$ tel que i est racine de $X^2 + 1 \in \mathbb{R}[X]$. i est donc algébrique sur $\mathbb{R}$ et $i^2 = -1$. On a donc :

$$\mathbb{C} = \mathbb{R}(i) = \mathbb{R}[i] \quad \text{et} \quad [\mathbb{C} : \mathbb{R}] = 2$$

11.1.1 Construction et Opérations

Commençons tout d'abord par définir formellement l'ensemble des nombres complexes.

Définition (**Nombres complexes (en tant que quotient)**) . Soit $\mathbb{R}[X]$ l'anneau des polynômes à coefficients réel. On définit l'anneau $\mathbb{C}$ comme le quotient de $\mathbb{R}[X]$ par l'idéal

engendré par le polynôme $X^2 + 1 \in \mathbb{R}[X]$. On a ainsi :

$$\mathbb{C} := \mathbb{R}[X]/(X^2 + 1)$$

- $(X^1 + 1)$ est irréductible dans $\mathbb{R}[X]$ donc $\mathbb{C}$ est un corps. On parlera du **corps** des complexes.
- Tout élément $z \in \mathbb{C}$ s'écrit de la forme :

$$z = a + Xb, \quad a, b \in \mathbb{R}$$

et on identifie X à un élément $i \in \mathbb{C}$ tel que :

$$i^2 = -1$$

On définit les opérations sur $\mathbb{C}$ de la façon suivante $\forall a + ib, c + id \in \mathbb{C}$,

- **Addition** :

$$(a + ib) + (c + id) = (a + c) + i(b + d)$$

- **Multiplication** :

$$(a + ib) \times (c + id) = ac + iad + ibc - bd = (ac - bd) + i(ad + bc)$$

On peut donc écrire $\mathbb{C}$ de façon ensembliste comme une extension de $\mathbb{R}$:

$$\mathbb{C} := \{a + ib \mid a, b \in \mathbb{R}\}$$

11.1.2 Structure de Corps

Proposition Les propriétés de $\mathbb{R}$ et la construction de $\mathbb{C}$ nous permettent de lui donner une structure de corps. On a donc les propriétés suivantes :

- L'addition et la multiplication sont associatives et commutatives dans $\mathbb{C}$.
- La multiplication est distributive sur l'addition.
- Il existe des éléments neutres :

$$1_{\mathbb{C}} = 1_{\mathbb{R}} \quad \text{et} \quad 0_{\mathbb{C}} = 0 + i \times 0$$

Remarque Comme dit en introduction, $\mathbb{C}$ est construit comme extension de $\mathbb{R}$. Ainsi, l'application :

$$\phi : \begin{cases} \mathbb{C} \longrightarrow \mathbb{R} \\ a + 0 \times i \longmapsto a \end{cases}$$

est évidement injective. On peut donc en conlcure que $\mathbb{R} \subset \mathbb{C}$.

Proposition $\mathbb{C}$ est un corps de caractéristique nulle. Autrement dit il n'existe pas d'entier $n \in \mathbb{N}^*$ tel que :

$$1 + \cdots + 1 = n \times 1 = 0$$

Théorème (**Théorème Fondamental de l'Algèbre**) . Tout polynômes à coefficients complexes admet au moins une racine dans $\mathbb{C}$. Donc $\mathbb{C}$ est un corps alégriquement clos.

11.2 Forme algébrique

Le corps des complexes est assez facile à manipuler étant donné que chacun de ses éléments admets plusieurs formes d'écritures (algébrique, exponentielle, trigonométrique) qui permettent d'appréhender des manipulations de différentes façons.

11.2.1 Partie Réelle et Imaginaire

La forme algébrique d'un nombre complexe est sûrement la "plus simple" à comprendre, elle découle directement de la définition des nombres complexes comme extension de $\mathbb{R}$.

Définition (**Forme Algébrique d'un nombre complexe**) . Soit $z \in \mathbb{C}$. Alors il existe un unique couple $(a, b) \in \mathbb{R}^2$ tel que :

$$z = a + ib$$

On appelle cette forme la **forme algébrique** de z. Ainsi, on définit les parties réelles ($\Re(z)$) et imaginaires ($\Im(z)$) de z comme les applications :

$$\Re : \begin{cases} \mathbb{C} \longrightarrow \mathbb{R} \\ a + ib \longmapsto a \end{cases} \qquad \Im : \begin{cases} \mathbb{C} \longrightarrow \mathbb{R} \\ a + ib \longmapsto c \end{cases}$$

Ce sont deux applications linéaires.

Proposition Les parties imaginaires et réelles des nombres complexes nous permettent de les caractériser plus facilement. Ainsi, un nombre complexe est réell ssi sa partie imaginaire est nulle.

$$\text{i.e} \quad z \in \mathbb{R} \iff \Im(z) = 0_{\mathbb{R}}$$

Remarque Grâce à l'unicité des parties réelles et imaginaires, on a les règles d'identification usuelles suivantes :

$$a + ib = c + id \iff a = c \text{ et } b = d$$

$$a + ib = 0_{\mathbb{C}} \iff a = b = 0_{\mathbb{R}}$$

11.2.2 Conjugaison

Dans $\mathbb{C}$ il existe une opération appelée "conjugaison" qui, algébriquement n'a pas trop de signification, mais prend tout son sens sous forme trigonométrique.

Définition (**Conjugaison**) . On définit l'application conjugaison dans les complexes comme :

$$\overline{\bullet} : \begin{cases} \mathbb{C} \longrightarrow \mathbb{C} \\ a + ib \longmapsto \overline{a + ib} = a - ib \end{cases}$$

On appelera $\overline{z}$ le conjugué de z dans $\mathbb{C}$.

Propriété (**Conjugaison**) .

- La conjugaison est **involutive**.

$$\text{i.e} \quad \forall z \in \mathbb{C}, \overline{\overline{z}} = z$$

- La conjugaison est **linéaire**.
- La conjugaison respecte la multiplication.

$$\text{i.e} \quad \forall z, z' \in \mathbb{C}, \overline{z.z'} = \overline{z} \times \overline{z'}$$

- Pour tout $z \in \mathbb{C}$ on a les propriétés suivantes :

$$\Re(z) = \frac{z + \overline{z}}{2} \quad \text{et} \quad \Im(z) = \frac{z - \overline{z}}{2i}$$

11.2.3 Module d'un nombre complexe

Tout comme la conjugaison, le module d'un nombre complexe n'a pas beaucoup de signification sous forme algébrique mais il peut s'interpréter comme une norme dans $\mathbb{C}$.

Définition (**Module d'un nombre complexe**) . On appelle module dans $\mathbb{C}$ l'application :

$$|\cdot| : \begin{cases} \mathbb{C} \longrightarrow \mathbb{R}_+ \\ z = a + ib \longmapsto \sqrt{z\overline{z}} = \sqrt{a^2 + b^2} \end{cases}$$

On appelera ainsi module de z dans $\mathbb{C}$ le réel positif $\sqrt{z\overline{z}}$.

Tout comme la conjugaison, le module possède quelque propriétés bien utiles.

Propriété (**Module**) .

- Le module est défini.

$$\text{i.e} \quad \forall z \in \mathbb{C}, \quad |z| = 0_{\mathbb{R}} \iff z = 0_{\mathbb{C}}$$

- Le module est positif ($\forall z \in \mathbb{C}, |z| \geqslant 0$).
- Le module respecte la multiplication.

$$\text{i.e} \quad \forall z, z' \in \mathbb{C}, \quad |z \times z'| = |z| \times |z'|$$

- Le module vérifie l'inégalité triangulaire :

$$\forall z, z' \in \mathbb{C}, \quad |z + z'| \leqslant |z| + |z'|$$

Remarque On peut remarque que le module vérifie toutes les conditions pour être une norme. En effet, **le module est une norme sur** $\mathbb{C}$. De plus, il est bien compatible avec la norme réelle (valeur absolue) puisque :

$$\forall a \in \mathbb{R}, \quad \sqrt{a \times \overline{a}} = \sqrt{a^2} = |a|$$

11.3 Forme Trigonométrique d'un nombre complexe

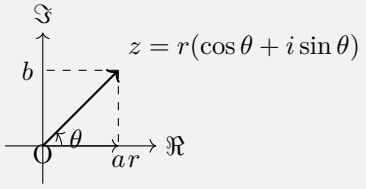
La forme trigonométrique d'un nombre complexe est moins facilement manipulable algébriquement mais reste très utile. En effet, elle permet de représenter facilement n'importe quel nombre complexe dans le plan.

Définition (**Forme Trigonométrique**) . Soit $z = a + ib \in \mathbb{C}$. On définit la forme trigonométrique de z comme :

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

où :

- $r = |z| = \sqrt{a^2 + b^2}$
- θ est appelé **argument de** z. Il représente l'angle que fait z avec l'axe des réels dans le plan.

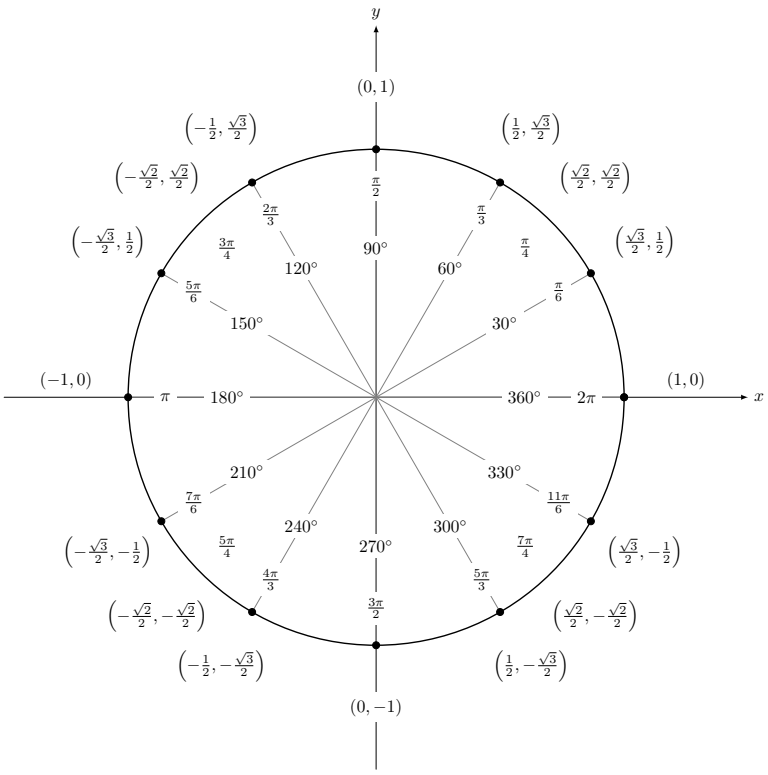


11.3.1 Nombres complexes de module 1

On peur remarquer, parmi les nombres complexes que ceux de moldule 1 sont assez caractéristiques. On note l'ensemble des nombres complexes de module 1 $\mathbb{U}$ défini par :

$$\mathbb{U} := \{z \in \mathbb{C},\ |z| = 1\}$$

Les nombres complexes forment ainsi ce que l'on appelle le **cercle trigonométrique** dans le plan :



Remarque On peut aussi remarquer que l'inverse d'un nombre complexe de module 1 est exactement son conjugué.

11.4 Forme Exponentielle d'un nombre complexe

Enfin, nous pouvons définir la forme exponentielle d'un nombre complexe.

Définition (**Forme Exponentielle**) . Soit $z = r(\cos\theta + i\sin\theta) \in \mathbb{C}^*$. On définit sa forme exponentielle ou polaire comme :

$$z = re^{i\theta}$$

Remarque La forme exponentielle d'un nombre complexe est unique modulo 2π pour θ.

$$\text{i.e} \quad \forall r_1 e^{i\theta_1}, r_2 e^{i\theta_2} \in \mathbb{C}, \quad r_1 e^{i\theta_1} = r_2 e^{i\theta_2} \iff \begin{cases} r_1 = r_2 \\ \theta_1 \underset{2\pi}{\equiv} \theta_2 \end{cases}$$

Propriété (**Formule de Moivre**) . Soit $\theta \in \mathbb{R}$ et $n \in \mathbb{Z}$ on a alors la formule suivante, dite de Moivre :

$$\left(e^{i\theta}\right)^n = e^{ni\theta}$$

en notation troginométrique, on a donc :

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

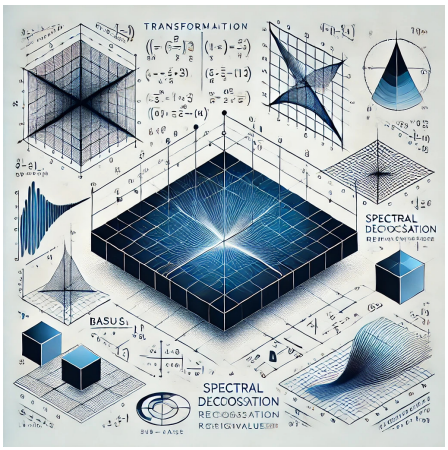
Proposition (Exponentielle complexe de module 1) Pour les exponentielles complexes de module 1, on a différentes propriétés notamment les formules d'Euler :

$$\forall \theta \in \mathbb{R}, \quad \cos\theta = \frac{e^{i\theta} + e^{-i\theta}}{2} \text{ et } \sin\theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}$$

Mais aussi :

- $\forall z \in \mathbb{C}, \quad |z| = 1 \iff \exists \theta \in \mathbb{R}, z = e^{i\theta}$
- $\forall \theta \in \mathbb{R}, \quad \overline{e^{i\theta}} = e^{i-\theta}$
- $\forall p, q \in \mathbb{R}, \quad e^{i(p+q)} = e^{ip} \times e^{iq}$

Algèbre Linéaire



Chapitre 1

Espaces Vectoriels

Contents

Dans ce chapitre nous allon définir les structures utulisées dans toute l'algèbre linéaire, les expaces vectoriels. Ce concept est fondamental pour tout le reste de l'algèbre linéaire puisque nous définissons les fondements même de la discipline.

1.1 Structure d'espace vectoriel

Les espaces vectoriels sont définis de façon axiomatique de la façon suivante.

Définition (**Axiomes d'un espace vectoriel**) . Soit F un corps et V un ensemble. On dit que V est un espace vectoriel sur F si les éléments de V et les opérations suivantes satisfont les propriétés suivantes :

1. **Commutativité de l'addition** : Pour tous $u, v \in V$, on a
$$u + v = v + u$$
2. **Associativité de l'addition** : Pour tous $u, v, w \in V$, on a
$$(u + v) + w = u + (v + w)$$
3. **Existence d'un élément neutre pour l'addition** : Il existe un élément $0 \in V$ tel que pour tout $v \in V$,
$$v + 0 = v$$

4. **Existence d'un opposé pour l'addition** : Pour chaque $v \in V$, il existe un élément $-v \in V$ tel que
$$v + (-v) = 0$$
5. **Compatibilité de la multiplication scalaire avec l'addition vectorielle** : Pour tous $\lambda \in F$ et $u, v \in V$, on a
$$\lambda \cdot (u + v) = \lambda \cdot u + \lambda \cdot v$$
6. **Compatibilité de la multiplication scalaire avec l'addition scalaire** : Pour tous $\lambda, \mu \in F$ et $v \in V$, on a
$$(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v$$
7. **Distributivité de la multiplication scalaire** : Pour tous $\lambda, \mu \in F$ et $v \in V$, on a
$$\lambda \cdot (\mu \cdot v) = (\lambda\mu) \cdot v$$
8. **Existence d'un élément neutre pour la multiplication scalaire** : Il existe un élément $1 \in F$ tel que pour tout $v \in V$,
$$1 \cdot v = v$$

En pratique, on utilise les propriétés des espaces vues dans les chapitres précédents et les structures déjà connnues.

Définition (**Espace Vectoriel**) . En résumé, un espace vectoriel E est un groupe abélien pour l'addition vectorielle muni d'une multiplication scalaire qui interagit de manière compatible par distributivité avec cette addition via les opérations du corps de base $\mathbb{K}$.

Dans toute la suite du cours, on notera E un espace vectoriel sur un corps de base $\mathbb{K}$ (en général $\mathbb{R}$ ou $\mathbb{C}$).

Remarque On parlera de **vecteur** pour les éléménts de E et de **scalaires** pour les éléments de $\mathbb{K}$.

Propriété (**Élémentaires**) .
- $\forall \lambda \in \mathbb{K}, \forall u \in E, \quad \lambda u = 0_E \Longrightarrow \lambda = 0_{\mathbb{K}}$ ou $u = 0_E$
- $\forall u \in E, \quad -u = -1 \times u$
- $\forall \lambda, \mu \in \mathbb{K}, \forall u \in E, \quad (\lambda - \mu)u = \lambda u - \mu u$

Définition (**Combinaison Linéaire**) . Soient $(u_1, \ldots, u_n)$ une famille de vecteur de E et $(\lambda_1, \ldots, \lambda_n)$ une famille de scalaires de $\mathbb{K}$. On définit une combinaison linéaire de ces vecteurs par ces scalaire comme le vecteur :
$$u' = \lambda_1 u_1 + \cdots + \lambda_n u_n \in E$$

La notion de combinaison linéaire est essentielle dans ce chapitre, notamment pour la compéhension des vect.

Exemple (Espace Vectoriels Usuels) Voyons quelques exemples d'espaces vectoriels.

- **Les complexes** : on peut voir l'ensemble des complexes $\mathbb{C}$ comme un $\mathbb{R}$-espace vectoriel. Ainsi, on muni $\mathbb{C}$ de son addition usuelle et de la multiplication externe suivante :

$$\forall \lambda \in \mathbb{R}, \forall z \in \mathbb{C}, \quad \lambda z \in \mathbb{C}$$

comme le produit complexe. Muni de ces deux lois, $\mathbb{C}$ est un espace vectoriel sur $\mathbb{R}$.

- **Les fonctions de X vers E** : Soient E un espace vectoriel et X un ensemble non vide. L'ensemble $\mathcal{F}(X, E)$ des applications de X dans E est un espace vectoriel. On définit l'égalité dans $\mathcal{F}(X, E)$ comme l'inégalité fonctionnelle usuelle. On le muni de l'addition suivante :

$$\forall f, g \in \mathcal{F}(X, E), \forall x \in X, \quad f(x) + g(x) = (f + g)(x)$$

et de la multiplication scalaire suivante :

$$\forall \alpha \in \mathbb{K}, \forall f \in \mathcal{F}(X, E), \forall x \in X, \quad \alpha f(x) = f(\alpha x)$$

- **Polynômes :** L'ensemble des polynômes à coefficients dans $\mathbb{K}$ noté $\mathbb{K}[X]$ est un $\mathbb{K}$-espace vectoriel pour l'addition et la multiplication usuelles.

Théorème (**Applications**) . Soient E un espace vectoriel et X un ensemble non vide. Alors l'ensemble $(\mathcal{F}(X, E), +, .)$ est un espace vectoriel sur $\mathbb{K}$.

Il peut être utile de définir les espaces vectoriels produits.

Définition (**Espaces Vectoriels Produits**) . Soient E, F deux $\mathbb{K}$-espaces vetcoriels. On note $G = E \times F$ l'espace vectoriel produit de E et F tel que :

$$\forall a, b \in E, \forall c, d \in F, \quad (a, b) + (c, d) = (a + c, b + d)$$

$$\forall \alpha \in \mathbb{K}, \forall (a, b) \in G, \quad \alpha(a, b) = (\alpha a, \alpha b) \in G$$

1.2 Sous-espaces vectoriels

Tout comme les sous-groupes, les sous-anneaux, etc, avant eux, les espaces vectoriels possèdent des sous-structures, appelées sous-espaces vectoriel qui ont les mêmes propriétés.

Définition (**Sous-espace Vectoriel**) . Soient E un $\mathbb{K}$-espace vectoriel et $F \subseteq E$. On dit que F est un sous-espace vectoriel de E si :

1. $0_E \in F$
2. $\forall x, y \in F, x + y \in F$
3. $\forall \lambda \in \mathbb{K}, \forall x \in F, \lambda.x \in F$

Proposition Ainsi, un sous-espace vectoriel F d'un $\mathbb{K}$-espace vectoriel E est aussi un $\mathbb{K}$-espace vectoriel pour les lois induites de E mais restreintes à F.

En pratique, on cherchera donc à montrer qu'un ensemble est un sous-espace vectoriel plutôt qu'à prouver tous les axiomes de la définition d'un espace vectoriel.

Exemple Soit $E = \mathbb{R}^3$ un $\mathbb{R}$-espace vectoriel. Par exemple, les droites de E sont des sous-espaces vectoriels de E. De même, pour $F = \mathbb{C}[X]$, l'ensemble $\{P \in \mathbb{C}[X] \mid P(0) = 0\}$ est un sous-espace vectoriel de F.

1.3 Sous-espaces vectoriels engendrés

Les sous-espaces vectoriels engendrés sont très utiles en algèbre linéaire. En effet, ils permettent de décrire très simplement un sous-espace vectoriel (et donc un espace vectoriel) grâce à une famille finie d'éléments.

Définition (**Sous-espace vectoriel engendré**) . Soient E un $\mathbb{K}$-espace vectoriel, X une partie de E et $n \in \mathbb{N}$. On définit le **sous-espace vectoriel engendré par** X, noté $\text{vect}(X)$, comme le plus petit sous-espace vectoriel de E contenant X, c'est-à-dire l'ensemble des combinaisons linéaires de tous les éléments de X. Plus formellement, on a donc :

$$\boxed{\text{vect}(X) := \{\alpha_1 u_1 + \cdots + \alpha_n u_n \mid \forall i \in [\![1, n]\!], \alpha_i \in \mathbb{K}, u_i \in X\}}$$

C'est un sous-espace vectoriel de E.

Proposition Dans le cas où X est un ensemble fini de la forme $X = \{u_1, \ldots, u_p\} \subset E$. Alors on a :

$$\text{vect}(X) = \{\alpha_1 u_1 + \cdots + \alpha_n u_n \mid \forall \alpha_1, \ldots, \alpha_n \in \mathbb{K}\}$$

On dit alors que X **engendre vect**(X) ou que X **est un générateur de vect**(X).

Lemme (Lemme d'expansion) Soient E un $\mathbb{K}$-espace vectoriel, F un sous-espace vectoriel de E

$$\boxed{X \subseteq F \Longrightarrow \text{vect}(X) \subseteq F}$$

Le lemme d'expansion est très utiile pour montrer des égalités ou des inclusions de sous-espaces vectoriels. Intuitivement, si X est une partie quelconque d'un espace vectoriel F, alors le plus petit sous-espace vectoriel qui contient X est lui-aussi inclus dans F. Un sous-espace vectoriel est donc "stable" par le vect.

Exemple Si $E = \mathbb{R}^3$ et $X = \{(1, 0, 1), (0, 1, 0)\}$ alors :

$$\begin{aligned}\text{vect}(X) &= \{\alpha(1, 0, 1) + \beta(0, 1, 0) \mid \alpha, \beta \in \mathbb{R}\} \\ &= \{(\alpha, \beta, \alpha) \mid \alpha, \beta \in \mathbb{R}\}\end{aligned}$$

Dans $\mathbb{R}^3$ ce sous-espace vectoriel peut se représenter comme un plan. De même, si $X = \{(1, 1, 0)\}$ alors :

$$\text{vect}(X) = \{(\alpha, \alpha, 0) \mid \alpha \in \mathbb{R}\}$$

Ici, $\text{vect}(X)$ sera donc une droite.

Propriété (**Sous-espaces vectoriels et inclusion**) . Soient $X, Y \subset E$ deux parties quelconques, alors :

$$X \subseteq \text{vect}(X) \quad \text{et} \quad X \subseteq Y \Longrightarrow \text{vect}(X) \subseteq \text{vect}(Y)$$

1.4 Opérations sur les sous-espaces vectoriels

Définissons maintetant quelques opérations ensemblistes sur les sous-espaces vectoriels et les propriétés qui en découlent.

Définition (**Intersection**) . Soient F et G deux sous-espaces vectoriels. On définit l'intersection de F et G notée $F \cap G$ comme l'intersection ensembliste :

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

et $F \cap G$ est aussi un sous-espace vectoriel de E.

Définition (**Somme**) . Soient F et G deux sous-espaces vectoriels de E. On définit la somme de F et G comme :

$$F + G := \{x + y \mid \forall x \in F, \forall y \in G\}$$

c'est aussi un sous-espace vectoriel de E.

Remarque On peut remarquer que la somme de sous-espaces vectoriels nous permet de décomposer chaque vecteur x de $F + G$ en somme de deux vecteurs $x_F \in \mathcal{F}$ et $x_G \in G$. Mais attention, cette décomposition n'est pas unique !

Définition (**Somme directe**) . Soit $F_1, \ldots, F_k$ une famille de sous-espaces vectoriels de E. On dit que les $F_i, \forall i \in [\![1,]\!]n$ sont **supplémentaires** dans E si pour tout vecteur $u \in E$, il admet une unique décomposition dans les F_i. Plus formellement, $F_1, \ldots, F_k$ sont supplémentaires dans E ssi

$$\forall u \in E, \exists!(u_1, \ldots, u_k) \in (F_1, \ldots, F_k) \text{ tels que } u = u_1 + \cdots + u_k$$

On note alors $E = F_1 \oplus \cdots \oplus F_k$.

La somme directe permet donc de décomposer une "gros" sous-espace vectoriel en "plus petits" sous espaces vectoriels. Attardons nous sur le cas particulier d'une somme directe de deux sous-espaces vectoriels.

Remarque (Cas particulier de deux sev) Soient F et G deux sous-espaces vectoriels de E. On dit que F et G sont supplémentaires dans E si leur intersection est réduite à 0_E et leur somme est égale à E. Plus formellement :

$$F \oplus G \iff \begin{cases} F \cap G = \{0_E\} \\ F + G = E \end{cases}$$

Définition (**Supplémentaire**) . Soit F un sous-espace vectoriel de E. On définit le supplémentaire de F dans E comme l'unique sous-espace vectoriel G de E tel que $F \oplus G = E$.

Propriété (**Existence**) . Sous E un espace vectoriel quelconque. Alors tout sous-espace vectoriel de E possède un supplémentaire dans E.

1.5 Familles Génératrices, Libres et Liées

Définition (**Famille**) . Soit F un sous-espace vectoriel de E. Une famille de vecteurs de F, notée $\mathcal{F}$ est un ensemble de vecteurs de F ordonnés ou non $\mathcal{F} = (u_1, \dots, u_p)$. Dans le cas d'un famille ordonnée, on parle de suite de vecteurs.

Définition (**Famille Génératrice**) . Soit F un sous-espace vectoriel et $\mathcal{F}$ une famille de F. On dit que $\mathcal{F}$ est génératrice de F ou engendre F si $F = \text{vect}(\mathcal{F})$.

Définition (**Famille Libre**) . Soit F un sous-espace vectoriel et $\mathcal{F} = (u_1, \dots, u_p)$ une famille de F. On dit que la famille $\mathcal{F}$ est libre si aucun de ses vecteurs n'est combinaison linéaire des autres.

Plus formellement, $\mathcal{F}$ est dite libre ssi pour toute combinaison linéaire nulle de ses vecteurs, alors tous les coefficients sont nuls.

$$\text{i.e} \quad \forall \alpha_1, \dots, \alpha_p \in \mathbb{K}, \alpha_1 u_1 + \cdots + \alpha_p u_p = 0_F \Longrightarrow \alpha_1 = \cdots = \alpha_p = 0_{\mathbb{K}}$$

Dans le cas d'une famille qui n'est pas libre, on parle de **famille liée**.

Remarque
- Par définition, deux vecteurs x et y sont dits colinéaires si la famille (x, y) est liée.
- Une famille qui contient 0_E est liée.
- Si la famille $(u_1, \dots, u_p)$ est libre et que $x \notin \text{vect}(u_1, \dots, u_p)$ alors la famille $(u_1, \dots, u_p, x)$ est aussi libre.

Exemple Si $E = \mathbb{R}^2$, alors la famille $((1,0),(0,2))$ est libre, tandis que la famille $((2,0),(3,0))$ est liée car le second vecteur est combinaison linéaire du premier.

Proposition Soient $F = \text{vect}(u_1, \dots, u_p)$ et $G = \text{vect}(a_1, \dots, a_n)$ deux sous espaces vectoriels. On a alors :

$$F + G = \text{vect}(u_1, \dots, u_p, a_1, \dots, a_n)$$

Définition (**Droite vectorielle**) . On appelle droite vectorielle un sous-espace vectoriel engendré par une famille d'un seul vecteur non nul.

Définition (**Plan vectoriel**) . On appelle plan vectoriel un sous-espace vectoriel engendré par une famille de deux vecteurs non nuls.

Remarque Soit $F = \text{vect}(u_1, \dots, u_p, u)$. On peut remarquer que si $u \in \text{vect}(u_1, \dots, u_p)$ alors

$$F = \text{vect}(u_1, \dots, u_p, u) = \text{vect}(u_1, \dots, u_p)$$

Autrement dit, il peut exister des générateurs redondants. Ce sont des vecteurs "inutiles" car ils sont combinaison linéaire des autres vecteurs $u_1, \dots, u_p$ de la famile. La famille est liée. On peut les enlever pour obtenir une famille libre.

1.6 Bases et coordonées

1.6.1 Bases et vecteurs colonnes

Dans cette section, nous allons aborder les notions de base et de coordonées. Ces notions permettent de donner une sorte de "repère" dans lequel nous nous trouvons. Les bases peuvent être vues comme des "façon de regarder" nos espaces.

Définition (**Base**) . Soit F un sous-espace vectoriel de E. Une base de F est une famille **libre et génératrice** de F dans E.

Remarque Ainsi, pour montrer qu'une famille $\mathcal{B} = (e_1, \dots, e_n)$ est une base de F, il faut procéder en deux temps :

i) Montrer que $\mathcal{B}$ est génératrice (i.e que $F = \text{vect}(\mathcal{B})$).

ii) Montrer que $\mathcal{B}$ est libre.

Définition (**Vecteur colonne**) . Un vecteur colonne d'un $\mathbb{K}$-sev F est une application :

$$X : \{1, \dots, n\} \longrightarrow \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n$$

où $x_1, \dots, x_n \in \mathbb{K}$.

Remarque Quelques remarques concernant les vecteurs colonnes :

- Un vecteur colonne est un **objet ordonné**, i.e si on change des éléménts de position, on n'a plus le même objet.
- L'écriture en colonne n'est qu'une écriture représentative de cette fonction.
- Deux vecteurs colonnes sont égaux ssi leur composantes sont égales deux à deux.

Proposition Définissons quelques opérations sur les vecteurs colonnes. Soient $X, Y \in \mathbb{K}^n$ deux vecteurs colonnes et $\alpha \in \mathbb{K}^n$, on a :

$$X + Y = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix} \quad \text{et} \quad \alpha X = \alpha . \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \alpha . x_1 \\ \vdots \\ \alpha . x_n \end{pmatrix}$$

Nous verrons dans le chapitre suivant que ces vecteurs colonnes sont en fait des cas particulier d'objects appelés matrices qui nous permettront de représenter analytiquement toute application linéaire.

1.6.2 Coordonnées et équations cartésiennes

Ces vecteurs colonnes nous permettent donc de définir la notion de coordonnées d'un vecteur dans un sev de la façon suivante.

Définition (**Coordonnées**) . Soit F un $\mathbb{K}$-sev et $\mathcal{B} = (e_1, \dots, e_n)$ une base de F. Pour tout $u \in F$, il existe un unique n-uplet $(\alpha_1, \dots, \alpha_n) \in \mathbb{K}^n$ tels que :

$$u = \alpha_1 e_1 + \cdots + \alpha_n e_n$$

On dit alors que le n-uplet $(\alpha_1, \dots, \alpha_n)$ sont les **coordonnées de u dans la base $\mathcal{B}$**.

On peut alors définir le **vecteur coordonnées** de u dans $\mathcal{B}$ noté $[u]^{\mathcal{B}}$ comme :

$$[u]^{\mathcal{B}} := \begin{cases} \{1, \dots, n\} \longrightarrow \mathbb{K}^n \\ i \longmapsto \alpha_i \end{cases}$$

Il est important de noter la différence entre un vecteur et ses coordonnées. En effet, un vecteur d'un sev F est un objet intrinsèque, qui ne dépend pas d'une base alors que ses coordonnées dépendent de la base dans laquelle on se place. Ainsi, lorsque l'on parlera des coordonnées d'un vecteur, il sera essentiel de toujours spécifier dans quelle base on se trouve.

Proposition Soient deux vecteurs $x, y \in F$ et $\mathcal{B}$ une base de F. On a :

$$x = y \iff [x]^{\mathcal{B}} = [y]^{\mathcal{B}}$$

Ainsi deux vecteurs sont égaux ssi ils ont les mêmes coordonnées dans une même base.

Remarque (Base canonique) On appelle base canonique d'un espace vectoriel sa base de référénce. Par exemple, la base canonique de $\mathbb{R}^3$ est $\mathcal{C}^3 := ((1,0,0),(0,1,0),(0,0,1))$.
De même, la base canonique de $\mathbb{R}_3[X]$ en tant que $\mathbb{R}$-espace vectoriel est $\mathcal{B} = (1, X, X^2, X^3)$. Intuivivement, il faut comprendre qu'à partir de cette famille, on peut générer tout polynôme de $\mathbb{R}_3[X]$ par une unique combinaison linéaire.

Théorème (**Équation Cartésienne**) . Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E. Soient $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ alors l'ensemble :

$$H := \{x_1 e_1 + \dots + x_n e_n \in E \mid \alpha_1 x_1 + \dots \alpha_n x_n = 0\}$$

est un sous-espace vectoriel de E. On dit que c'est un sous-espace vectoriel d'équation cartésienne $\alpha_1 x_1 + \dots \alpha_n x_n = 0$ dans la base $\mathcal{B}$.

Une équation cartésienne permet donc de caractériser un sous-espace vectoriel au même titre qu'une base. A partir d'une équation cartésienne, on peut facilement se ramener à une base du sev.

Théorème (**Bases et somme directe**) . Soient $F_1, \dots, F_n$ n sev de E de bases respectives $\mathcal{B}_1, \dots, \mathcal{B}_n$. Soit la famille $\mathcal{B}$ composée de l'union des bases précedentes. On a alors :

$$E = \bigoplus_{i \in [\![1,n]\!]} F_i \iff \mathcal{B} \text{ est une base de } E$$

Théorème (**Base incomplète**) . Soit E un espace vectoriel.

- Toute famille libre de E peut être complétée en une famille libre et génératrice (une base) de E.
- De toute famille génératrice de E, on peut en extraire une famille libre et génératrice de E.

Fondamental, ce théorème permet de construire des bases d'espaces vectoriels mais aussi de simplifier des familles génératrices en bases.

1.7 Dimension

Dans cette section, nous nous placeront dans des espaces vectoriels admettant une base composé d'un nombre fini de vecteurs. On dit que ces espaces vectoriel sont de *dimension finie*.

Définition (**Dimension**) . Soit E un espace vectoriel de dimension finie. On définit la dimension de E en tant que $\mathbb{K}$-espace vectoriel, notée $\dim_{\mathbb{K}} E = n$ comme le nombre de vecteur dans une base de E.

Remarque Toutes les bases d'un même espace vectoriel de dimension finie ont le même nombre de vecteurs. D'où l'unicité de la dimension pour un espace vectoriel.

Propriété (**Dimension**) . Soit E un espace vectoriel de dimension finie. On a les propriétés suivantes :

- Toute famille libre de E contient **au plus** $\dim E$ vecteurs.
- Toute famille génératrice de E contient **au moins** $\dim E$ vecteurs.
- Une famille de $\dim E$ vecteurs est une base de E si elle est *libre* ou *génératrice*.

Proposition On peut donc établir une relation entre un espace vectoriel E et n'importe lequel de ses sous-espaces vectoriels F :

$$\dim F \leqslant \dim E$$

Avec la relation : $\dim F = \dim E \iff F = E$

Exemple

$$\dim R^n = n, \quad \dim \mathbb{K}_n[X] = n + 1, \quad \dim \mathcal{M}_{n,p}(\mathbb{K}) = n \times p$$

Théorème (**Formule de Grassman**) . Soient F et G deux sous-espaces vectoriels d'un même espace vectoriel E. On a alors :

$$\boxed{\dim F + \dim G = \dim(F + G) + \dim(F \cap G)}$$

Si E est de dimension finie, alors F et G aussi, on a donc :

$$\dim(F + G) = \dim F + \dim G - \dim(F \cap G)$$

La supplémentarité des sev nous donne aussi une identité sur leur dimension :

Théorème (**Supplémentarité et dimension**) . Soit E un espace vectoriel en dimension finie et F, G deux sous-espaces vectoriels de E. On a alors :

$$F \oplus G = E \implies \dim F + \dim G = \dim E$$

Chapitre 2

Applications Linéaires et Matrices

Contents

Dans le chapitre 1 nous avons détaillé la notion d'espace vectoriel ainsi que leur dimension. Attardons nous maintenant sur un notion fondamentale de l'algèbre linéaire : les applications. Elles permettent de lier ces espaces entre eux et possèdent beaucoup de propriétés intéressantes. Ce chapitre est à lire et à étudier en même temps que le suivant concernant les matrices. Nous verrons à la fin de celui-ci qu'elles jouent un rôle fondamental dans la compréhension et la manipulation des applications linéaires.

Dans tout ce chapitre, sauf mention contraire, nous nous plaçons dans un $\mathbb{K}$-espace vectoriel E quelconque.

2.1 Définition et Propriétés

2.1.1 Généralités et structure

Définition (**Application Linéaire**) . Soit E et F deux $\mathbb{K}$-espaces vectoriels. Une application $f : E \longrightarrow F$ est dite *linéaire* si elle respecte la structure d'espaces vectoriels de E et F. Autrement dit si :

- f est **additive** : $\forall x, y \in E, f(x + y) = f(x) + f(y)$
- f est **homogène** : $\forall x \in E, \forall \lambda \in \mathbb{K}, f(\lambda.x) = \lambda.f(x)$

Que l'on peut résumer en :

$$\forall x, y \in E, \forall \lambda \in \mathbb{K}, \quad f(x+y) = f(x) + f(y) \quad f(\lambda.x) = \lambda.f(x)$$

Ou plus simplement :

$$\forall x, y \in E, \forall \lambda \in \mathbb{K}, \quad f(\lambda.x + y) = \lambda.f(x) + f(y)$$

L'ensemble des applications linéaires de E vers F est noté $\mathcal{L}(E, F)$ ou $\mathcal{L}_{\mathbb{K}}(E, F)$. Lorsque $E = F$, on note $\mathcal{L}(E)$ et on parle *d'endomorphisme* de E. Enfin, si $F = \mathbb{K}$, on parle alors de *forme linéaire* de E.

Exemple Quelques exemples d'applications linéaires :

- L'application $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^3, (x, y) \longmapsto (2x + 3y, -1/3y)$ est une *application linéaire* de $\mathbb{R}^2$ dans $\mathbb{R}^3$.
- L'application $P \longmapsto P'$ est un *endomorphisme* de $\mathbb{K}[X]$.
- L'application $P \longmapsto P(x)$ où $x \in \mathbb{K}$ est une *forme linéaire* de $\mathbb{K}[X]$.

Remarque À noter que la restriction d'une application linéaire est elle aussi linéaire.

Propriété (**Applications Linéaires**) . La linéarité de telles applications $f : E \longrightarrow F$ leur procure des propriétés intéressantes :

- $f(0_E) = 0_F$
- f conserve les combinaisons linéaires : soit $(x_1, \ldots, x_n)$ une suite à valeurs dans E, on a :

$$f\left(\sum_{i=0}^{n} x_i\right) = \sum_{i=0}^{n} f(x_i)$$

Définition (**Homothétie**) . Soit E un $\mathbb{K}$-espace vectoriel et $\lambda \in \mathbb{K}$. L'application $\lambda \mathrm{Id}_E$ est un endomorphisme de E appelé *homothétie de E par rapport à λ*.

Proposition Soient E et F deux $\mathbb{K}$-espaces vectoriels. Soient $f, g \in \mathcal{L}_{\mathbb{K}}(E, F)$ et $\alpha \in \mathbb{K}$. On définit $f + g$ et αf par :

$$\forall x \in \mathbb{E}, \quad (f+g)(x) = f(x) + g(x) \quad \text{et} \quad (\alpha f)(x) = \alpha f(x)$$

On en déduis donc que $\mathcal{L}_{\mathbb{K}}(E, F)$ est un $\mathbb{K}$ espace vectoriel.

Théorème (**Dimension de $\mathcal{L}_{\mathbb{K}}(E, F)$**) . Soient E, F deux $\mathbb{K}$-espaces vectoriels de dimension finie. Alors $\mathcal{L}_{\mathbb{K}}(E, F)$ est aussi de dimension finie et on a :

$$\dim \mathcal{L}_{\mathbb{K}}(E, F) = \dim E \times \dim F$$

2.1.2 Noyau et image d'une application linéaire

Proposition (Noyau et Image) Soient E et F deux $\mathbb{K}$-espaces vectoriels et $f : E \longrightarrow F$ une application linéaire.

- L'image par f d'un sous-espace vectoriel de E est un sous-espace vectoriel de F. En particulier, le sous-espace vectoriel $f(E)$ est appelé *image de f* et noté $\mathrm{Im}\, f$.

- La pré-image d'un sous-espace vectoriel de F par f est un sous-espace vectoriel de E. On appelle *noyau de* f le sous espace vectoriel de E définit par $f^{-1}(\{0_F\})$ et on le note $\ker f$.

Remarque Soient E et F deux $\mathbb{K}$-espaces vectoriels et $f : E \longrightarrow F$ une application linéaire. Alors :

$$\ker f = \{x \in E \mid f(x) = 0_F\}$$
$$x \in \ker f \iff f(x) = 0_F$$

De même :

$$\text{Im } f = \{f(x) \mid x \in F\}$$
$$y \in \text{Im } f \iff \exists y \in F, f(x) = y$$

Théorème (**Injectivité/Surjectivité et Noyau/Image**) . Soient E et F deux $\mathbb{K}$-espaces vectoriels et $f : E \longrightarrow F$ une application linéaire.

1. f est *injective* ssi $ker f = \{0_E\}$
2. f est *surjective* ssi Im $f = F$.

L'image et le noyau d'une application linéaire sont donc très importants pour caractériser une application linéaire. Attention à bien faire attention à l'espace de vide de ces deux ensembles. Le noyau est un sous-espace vectoriel de E alors que l'image est un sous-espace vectoriel de F.

Théorème (**Théorème du Rang**) . Soient E un $\mathbb{K}$-espace vectoriel de dimension finie et F un $\mathbb{K}$-espace vectoriel quelconque. Soit $f : E \longrightarrow F$ une application linéaire.

$$\boxed{\dim E = \dim \ker f + \dim \text{Im } f}$$

Proposition (Rappel) Soient E et F deux $\mathbb{K}$-espaces vectoriels et $f : E \longrightarrow F$ une application linéaire. Pour rappel, f est *bijective* ssi f est *injective* et *surjective*.

2.2 Matrices

La grande force des applications linéaires, outre leur propriétés pratiques, est le fait que l'on puisse les représenter très facilement par des matrices.

2.2.1 Définition et généralités

Définition (**Matrice**) . On appelle matrice $n \in \mathbb{N}$ lignes et $p \in \mathbb{N}$ colonnes toute application

$$M : \{1, \dots, n\} \times \{1, \dots, p\} \longrightarrow \mathbb{K}$$

On dit alors que M est de taille $n \times p$. On note $\mathcal{M}_{n,p}(\mathbb{K})$ l'ensemble des matrices de taille $n \times p$ à coefficients dans un corps $\mathbb{K}$.

Exemple Les matrices sont représentées de façon tabulaire. Par exemple une matrice M à 2 lignes et 3 colonnes à coefficients dans $\mathbb{R}$ peut être :

$$M = \begin{pmatrix} 1 & 26 & \frac{3}{2} \\ \pi & 6 & \sqrt{2} \end{pmatrix}$$

On notera $M(i, j)$ ou $M_{i,j}$ le coefficient de M de la i-ème ligne et j-ème colonne. Ici, on a donc $M(1, 1) = 1$ et $M(2, 3) = \sqrt{2}$.

Proposition (Matrices Remarquables) Le nombre de lignes et de colonnes d'une matrice pouvant varier dans $\mathbb{N} \times \mathbb{N}$, il existe des matrices dites remarquables :

- **Matrice ligne** : une matrice de taille $1 \times p$ de la forme : $(x_1, \ldots, x_p)$
- **Matrice colonne** : de la forme $\begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$ (souvent utilisée pour représenter les coordonnées d'un vecteur dans une base comme vu précédement).

Proposition (Matrice Égales) On dit que deux matrices M et N $\in \mathcal{M}_{n,m}(\mathbb{K})$ sont égales ssi elles sont de même taille et tout leurs coefficients sont égaux.

$$\text{i.e} \quad \forall i, j \in [\![1, n]\!] \times [\![1, m]\!], M(i, j) = N(i, j)$$

Définition (**Transposition**) . Soit $\mathcal{M}_{n,m}(\mathbb{K})$ l'ensemble des matrices à coefficients dans un corps $\mathbb{K}$. On défini l'application transposition comme :

$$T : \begin{cases} \mathcal{M}_{n,m}(\mathbb{K}) \longrightarrow \mathcal{M}_{m,n}(\mathbb{K}) \\ M = (m_{i,j}) \longrightarrow^t M = (m_{j,i}) \end{cases}$$

Cette application prend une matrice $n \times m$ et la retourne en une matrice $m \times n$ en "transformant" ses lignes en colonnes. Elle est **involutive**, autrement dit $\forall M \in \mathcal{M}_{n,m}(\mathbb{K})$ ${}^t({}^tM) = M$ et **linéaire**.

Définition (**Matrice symétrique**) . On dit qu'une matrice $M \in \mathcal{M}_n(\mathbb{K})$ est symétrique ssi ${}^tM = M$.

2.2.2 Opérations sur les matrices

Dans cette section, nous considérons l'ensemble des matrices $\mathcal{M}_{n,m}(\mathbb{K})$ de taille $n \times m$ à coefficients dans un corps $\mathbb{K}$.

Définition (**Somme Matricielle**) . Soient $A, B \in \mathcal{M}_{n,p}(\mathbb{K})$ deux matrices de même taille. On définit la matrice $A + B \in \mathcal{M}_{n,p}(\mathbb{K})$ comme la matrice dont chaque élément est l'exacte somme des éléments de A et B situés à la même place. Plus formellement, si $\forall (i, j) \in [\![1, n]\!] \times [\![1, p]\!]$, $A = (a_{i,j})$ et $B = (b_{i,j})$ alors :

$$A + B = (a_{i,j} + b_{i,j})$$

Exemple (Somme Matricielle) Soit $A, B \in \mathcal{M}_{3,2}(\mathbb{R})$ de la forme :

$$A = \begin{pmatrix} 1 & 3 \\ 6 & 9 \\ 2 & 0 \end{pmatrix} \quad B = \begin{pmatrix} 1/2 & 0 \\ 6 & 3 \\ 11 & 2 \end{pmatrix} \quad \text{alors} \quad A + B = \begin{pmatrix} 3/2 & 3 \\ 12 & 12 \\ 13 & 2 \end{pmatrix}$$

L'ensemble $\mathcal{M}_{n,p}(\mathbb{K})$ muni de l'addition comme nous venons de la définir forme un groupe abélien $(\mathcal{M}_{n,p}(\mathbb{K}), +)$.

Définition (**Produit externe**) . Soient $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et $\lambda \in \mathbb{K}$ un scalaire. On définit la matrice $\lambda.M$ comme le *produit externe* de M par le scalaire λ. C'est une matrice de

$\mathcal{M}_{n,p}(\mathbb{K})$. Elle est obtenue en multipliant chaque valeur de M par le scalaire λ. On a donc :

$$\forall (i,j) \in [\![1,n]\!] \times [\![1,p]\!], \quad \text{tq} \quad M = (m_{i,j}) \quad \text{et} \quad \lambda.M = (\lambda \times m_{i,j})$$

Remarque À noter que toutes les opérations précédement définies sur les matrices concernent des matrices de même dimensions !

Propriété (**Opérations et transposition**) . Soient deux matrices $A, B \in \mathcal{M}_{n,p}(\mathbb{K})$ et $\lambda \in \mathbb{K}$. On a les propriétés suivantes :

- ${}^t(A+B) = {}^tA + {}^tB$
- ${}^t(\lambda A) = \lambda\, {}^tA$

Définition (**Matrice Antisyméytrique**) . Nous avons définit les matrices symétriques comme les matrices involutives par transposition. Les matrices *antisymétriques*, sont de la forme $M \in \mathcal{M}_{n,p}(\mathbb{K})$ telle que ${}^tM = -M$.

Ces opérations nous permettent de donner une structure à l'espace $\mathcal{M}_{n,p}(\mathbb{K})$.

Propriété (**Structure**) . L'ensemble $\mathcal{M}_{n,p}(\mathbb{K})$ muni des opérations usuelles (addition et multiplication par un scalaire) possède une structure de $\mathbb{K}$-espace vectoriel.
Le vecteur nul $0_{\mathcal{M}_{n,p}(\mathbb{K})}$ correspond à la matrice nulle $0_{n,p}$. De même, l'opposé de la matrice $A = (a_{i,j})$ est la matrice $-A = (-a_{i,j})$.

Définition (**Produit Matriciel**) . Soient $A \in \mathcal{M}_{m,n}(\mathbb{K})$ et $B \in \mathcal{M}_{n,p}(K)$ on peut alors définir le produit matriciel de A et B, noté $A \times B \in \mathcal{M}_{m,p}(\mathbb{K})$ comme la matrice $A \times B = (m_{i,j})_{1 \leqslant i \leqslant m, 1 \leqslant j \leqslant p}$ définie par :

$$\forall i \in [\![1,m]\!], \forall j \in [\![1,p]\!], \quad m_{i,j} = \sum_{k=1}^{n} a_{ik} \times b_{kj}$$

Remarque Le produit matriciel de deux matrices A, B est bien défini lorsque le nombre de lignes de A est égal au nombre de colonnes de B.

Propriété (**Produit Matriciel**) .

- Le produit matriciel N'EST PAS commutatif sur $\mathcal{M}_n(\mathbb{K})$.
- Le produit matriciel est associatif.

2.2.3 Inverses, trace et matrices semblables

Dans cette sous-section, nous nous plaçons dans $\mathcal{M}_n(\mathbb{K})$ et considérons des *matrices carrées*.

Définition (**Inverse**) . Soient $P \in \mathcal{M}_n(\mathbb{K})$ on dit que P est *inversible* si il existe $Q \in \mathcal{M}_n(\mathbb{K})$ telle que :

$$PQ = QP = I_n$$

on dit alors que Q est *l'inverse* de P. On note $GL_n(\mathbb{K})$ l'ensemble des matrices inversibles de $\mathcal{M}_n(\mathbb{K})$.

Proposition L'ensemble $(GL_n(\mathbb{K}), +, \times)$ est un *anneau non commutatif*.

Définition (**Trace**) . Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice carrée. On définit la *trace de A* comme le scalaire noté $\mathrm{tr}(A) \in \mathbb{K}$ formé par la somme des coefficients diagonaux de A. Ainsi :

$$\forall A \in \mathcal{M}_n(\mathbb{K}), \quad \mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii}$$

Propriété (**Trace et opérations**) . Soient $A, B \in \mathcal{M}_n(\mathbb{K})$ et $\lambda \in \mathbb{K}$. La trace dispose de plusieurs propriétés :

1. $tr(A+B) = tr(A) + tr(B)$
2. $tr(\lambda A) = \lambda tr(A)$
3. $tr(AB) = tr(BA)$

Définition (**Matrices Semblables**) . Soient $A, B \in \mathcal{M}_n(\mathbb{K})$ on dit que A et B sont *semblables* s'il existe une matrice inversible $P \in GL_n(\mathbb{K})$ telle que $B = P^{-1}AP$.

Remarque La relation $\mathcal{R}$ sur $\mathcal{M}_n(\mathbb{K})$ définie par :

$$A \mathcal{R} B \iff A \text{ et } B \text{ sont semblables}$$

est une relation d'quivalence sur.

Proposition Deux matrices semblables de $\mathcal{M}_n(\mathbb{K})$ ont même trace.

2.2.4 Matrice d'une application linéaire

Comme dit en introduction, chaque application linéaire peut être représentée par une matrice pour une base fixée.

Théorème (**Représentation Matricielle**) . Soient E, F deux $\mathbb{K}$-espaces vectoriels de dimension finie, $\mathcal{B}_E = (e_1, \ldots, e_p)$ une base de E et $\mathcal{B}_F = (f_1, \ldots, f_n)$ une base de F. Alors toute application linéaire $f : E \longrightarrow F$ est représentée par une matrice $A \in (M)_{n,p}(\mathbb{K})$ telle que :

$$\forall x \in E, \quad [f(x)]_{\mathcal{B}_\mathcal{F}} = A[x]_{\mathcal{B}_E}$$

De plus chaque colonne $j \in [\![1, p]\!]$ de la matrice A est constituée des coordonnées de l'image du vecteur de base e_j par la fonction f dans la base $\mathcal{B}_F$. La matrice de f est donc indiscociable de la base dans laquelle on l'exprime. On la notera $A = [f]^{\mathcal{B}_\mathcal{F}}_{\mathcal{B}_\mathcal{E}}$ ou $A = \mathrm{Mat}(f, \mathcal{B}_E, \mathcal{B}_f)$.

Exemple (Cas particulier) L'application identité $\mathrm{Id} : \mathbb{K}^n \longrightarrow \mathbb{K}^p$ est représentée canoniquement par la matrice identité dans la base canonique :

$$[\mathrm{Id}]^{\mathcal{C}_p}_{\mathcal{C}_n} = \begin{pmatrix} 1 & 0 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & \ddots & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \cdots & 0 \\ \vdots & \vdots & \vdots & 0 & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots & 0 & 1 \end{pmatrix}$$

Remarque Dans le cas d'un endomorphisme, on ne spécifiera pas l'espace d'arrivée et on considèrera des matrices carrées. De plus, par mesure de simplicité, pour une application linéaire $f \in \mathcal{L}(E)$, lorsque l'on voudra représenter sa matrice dans la même base d'arrivée que celle de départ, on notera $A = \text{Mat}(\mathcal{B})$.

Proposition Soient E, F deux $\mathbb{K}$ espaces vectoriels de base $\mathcal{B}$ de dimension n et $\mathcal{F}$ de dimension m. L'application :

$$\begin{cases} \mathcal{L}(E,F) \longrightarrow \mathcal{M}_{n,m}(\mathbb{K}) \\ f \longrightarrow \text{Mat}(f, \mathcal{B}, \mathcal{F}) \end{cases}$$

est un isomorphisme d'espace vectoriels. Dans le cas où E est un $\mathbb{K}$-espace vectoriel de base $\mathcal{B}$ de dimension n alors l'application :

$$\Phi : \begin{cases} \mathcal{L}(E) \longrightarrow \mathcal{M}_n(\mathbb{K}) \\ f \longrightarrow \text{Mat}(u, \mathcal{B}) \end{cases}$$

est à la fois un isomorphisme d'espaces vectoriels et un isomorphisme d'anneaux. On pourra vérifier facilement que :

1. $\forall u \in \mathcal{L}(E), \forall v \in \mathcal{L}(E), \quad \text{Mat}(u+v, \mathcal{B}) = \text{Mat}(u, \mathcal{B}) + \text{Mat}(v, \mathcal{B})$
2. $\forall \lambda \in \mathbb{K}, \forall u \in \mathcal{L}(E), \quad \text{Mat}(\lambda u, \mathcal{B}) = \lambda \times \text{Mat}(u, \mathcal{B})$
3. $\forall u \in \mathcal{L}(E), \forall v \in \mathcal{L}(E), \quad \text{Mat}(u \circ v, \mathcal{B}) = \text{Mat}(u, \mathcal{B}) \times \text{Mat}(v, \mathcal{B})$

Remarque Soient $x \in E$ de base $\mathcal{B}$ de dimension n et $y \in F$ de base $\mathcal{F}$ de dimension m. Soit $f : E \longrightarrow F$ une application linéaire et $A = \text{Mat}(f, \mathcal{B}, \mathcal{F})$. Si on pose :

$$[x]^{\mathcal{B}} = X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \text{et} \quad [y]^{\mathcal{F}} = Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$$

Alors on a :

$$y = f(x) \iff AX = Y$$

Théorème (**Inversion**) . Soit E un $\mathbb{K}$-espace vectoriel de dimension n. Soit $f \in \mathcal{L}(E)$ représentée par la matrice A dans une base quelconque de E notée $\mathcal{B}$. On a alors :

$$\boxed{f \text{ est un automorphisme de } E \iff A \in GL_n(\mathbb{K})}$$

2.2.5 Matrices de Passage

Cette sous-section va nous permettre d'introduire le chapitre suivant : la réduction d'endomorphismes.

Définition (**Matrice de Passage**) . Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n et $\mathcal{B}, \mathcal{B}'$ deux bases de E. On définit la *matrice de passage* de la base $\mathcal{B}$ à $\mathcal{B}'$ comme la matrice de l'application identité Id_E de la base $\mathcal{B}'$ dans la base $\mathcal{B}$. On la note $P_{\mathcal{B}}^{\mathcal{B}'} = \text{Mat}(Id_E, \mathcal{B}, \mathcal{B}')$. Ainsi, si un même vecteur $x \in E$ a pour coordonées dans $\mathcal{B}$ la matrice colonne X et dans $\mathcal{B}'$ la martrice colonne X', alors

$$X = P_{\mathcal{B}}^{\mathcal{B}'} X'$$

Remarque L'endomorphisme représenté par P permet donc de transformer une base de E en une autre. Par expansion, c'est donc un automorphisme de E. La matrice P est donc inversible.

Corollaire (**Changement de base**) . Soit E un $\mathbb{K}$-espace vectoriel de dimension finie et $u \in \mathcal{L}(E)$. Soient $\mathcal{B}, \mathcal{B}'$ deux bases de E, $A = \text{Mat}(u, \mathcal{B})$ et $B = \text{Mat}(u, \mathcal{B}')$. Soit P la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$. On a alors :

$$B = P^{-1}AP$$

Définition (**Trace d'un endomorphisme**) . Soit E un $\mathbb{K}$-espace vectoriel de dimension finie et $u \in \mathcal{L}(E)$. Soit A la matrice de u dans une base quelconque de E. On définit alors la trace de u comme la trace de n'importe quelle matrice représentative de u.

$$tr(u) = tr(A)$$

2.2.6 Rang d'une application linéaire / rang d'une matrice

Définition (**Rang**) . Soit E un $\mathbb{K}$-espace vectoriel.

- Soit $(v_1, \ldots, v_n)$ un système de vecteurs de E. On définit le rang de $(v_1, \ldots, v_n)$ noté $rg((v_1, \ldots, v_n))$ par :

$$rg((v_1, \ldots, v_n)) = \dim(\text{vect}(v_1, \ldots, v_n))$$

- Soient E et F deux $\mathbb{K}$-espaces vectoriels de dimension finie et $f : E \longrightarrow F$. On définit le rang de f, noté $rg(f)$ comme la dimension de $\text{Im}(f)$.

$$rg(f) = \dim(\text{Im}(f))$$

- Soit $A \in \mathcal{M}_{n,m}(\mathbb{K})$, le rang de A est égal au rang de l'application linéaire définie par A :

$$rg(A) = \dim(\text{Im}(A))$$

Remarque Soit E un $\mathbb{K}$-espace vectoriel de dimension finie et F un $\mathbb{K}$-espace vectoriel. Soit $f : E \longrightarrow F$. D'après le théorème du rang, on a alors :

$$rg(f) = \dim(E) - \dim(\ker f)$$

Théorème (**Invariance du Rang**) . Le rang d'une application linéaire est invariant par changement de bases dans E et F. De même, le rang d'une matrice A est invariant par opérations élémentaires sur les lignes et colonnes.

Chapitre 3

Réduction d'Endormorphismes

Contents

Dans des domaines appliqués ou de gros calculs de matriciels sonr effectués, il serait intéressant de pouvoir disposer de matrices pour lesquelles le calcul de produits est moins difficile. La solution est de trouver des matrices semblables ayant une forme plus simple (i.e diagonale ou triangulaire). Pour cela, nous utuliserons les formules de changement de base vues précédement. L'objectif sera donc de nous placer dans une base permettant d'exprimer la même application linéaire par une matrice plus simple.

Ainsi, nous commencerons tout d'abord par aborder la diagonalisation. Ses limites nous pousseront à définir de nouveaux objets tels que les polynômes d'endomorphismes. Ces derniers nous permettront de trigonaliser de nouvelles matrices grâce à la méthode des drapeaux de Dunford.

Remarque (Notation) Dans ce chapitre, nous noterons f des endomorphismes et M les matrices représentant ces endormorphismes. Il est important de comprendre la différence entre un endomorphisme sa matrice représentative. En effet, un endormorphisme est une application linéaire et sa matrice représentative permet de le caractériser dans une certaine base. Lorsque l'on change de base (i.e de point de vue), l'endomorphisme reste le même mais sa matrice représentative peut changer.

3.1 Diagonalisation

3.1.1 Eléments Propres

Définition (**Vecteur et valeur propre**) . Soit $f \in \mathcal{L}(E)$, $v \in E$ est un **vecteur propre** de f associé à la **valeur propre** λ ssi :

$$\begin{cases} v \neq 0 \\ f(v) = \lambda v \end{cases}$$

Il est important de noter qu'un vecteur propre est toujours associé à une valeur propre.

Définition (**Sous-espace propre**) . On appelle **sous-espace propre** de f (représenté par la matrice M) associé à la valeur propre λ le sous-espace vectoriel de E :

$$\boxed{E_\lambda = ker(M - \lambda Id_E)}$$

On remarquera la praticité de l'utilisation du théorème du rang pour déterminer la dimension du sous-espace propre :

$$dim(E_\lambda) = dim(E) - rg(M - \lambda Id_E)$$

Définition (**Polynôme Caractéristique**) . Soit $A \in \mathcal{M}_n(\mathbb{K})$. On appelle **polynôme caractéristique** de A le polynôme :

$$P_A = det(A - X \times Id_E)$$

Le terme "polynôme caractéristique" d'une matrice/d'un endomorphisme prend tout son sens. En effet, il permet vraiment de caractériser un endormorphisme par sa matrice représentative dans une base.

Remarque Le polynôme caractéristique d'un endomorphisme f est dit **intrinsèque**. (i.e le polynôme caractéristique de f ne dépend pas du choix de la matrice représentant de f)

3.1.2 Définitions et caractérisations

On cherche ici à savoir si un endormorphisme ou une matrice peut être représenté par une matrice diagonale dans une certaine base et à avoir une méthode pour construire une telle base.

Définition (**Endomorphisme diagonalisable**) . Un endomorphime f est dit **diagonalisable** ssi il existe une base E dans laquelle il est représenté par une matrice D diagonale. Une matrice A est dite diagonalisable ssi elle est semblable à une matrice diagonale.

$$\boxed{\text{A est diagonalisable } \Leftrightarrow \exists P \in GL_n(\mathbb{K}),\ A = PDP^{-1}}$$

Proposition Un endomorphisme f est diagonalisable dans E ssi il existe une base $\mathcal{B}$ de E constituée des vecteurs propres de f. La matrice diagonale associé sera donc composée des valeurs propres associés.

Proposition Un endomorphime f représenté par une matrice M est diagonalisable ssi M l'est.

Proposition Soit f un endomorphisme représenté par une matrice A dans une base $\mathcal{B}$ de dimension n. Soit $\lambda \in \mathbb{R}^*$.

$$\begin{aligned}\lambda \text{ est valeur propre de } f &\iff E_\lambda \neq \{0\} \\ &\iff f - \lambda\mathrm{Id} \text{ n'est pas de rang } n \\ &\iff A - \lambda\mathrm{Id} \text{ n'est pas de rang } n \\ &\iff A - \lambda\mathrm{Id} \text{ n'est pas inversible} \\ &\iff \det(A - \lambda\mathrm{Id}) = 0\end{aligned}$$

Cette proposition nous donne donc une caractérisation très pratique des valeurs propres. En plus du théorème du rang, on dispose donc déjà d'outils puissant pour étudier la diagonalisation d'une matrice/d'un endomorphime.

3.1.3 Recherche des éléments propres

En pratique, pour diagonaliser un endormorphisme, il nous faut donc commencer par trouver ses éléments propres. D'après la proposition précédente, une fois les valeurs propres trouvées, il reste à déterminer les sous-espaces propres E_λ. On se ramène ici à calculer des noyaux par le pivot de Gauss.

Propriété (**Lien Valeurs Propres/Racines**) . Soient $A \in \mathcal{M}_n(\mathbb{K})$ et $\lambda \in \mathbb{K}$ une valeur propre de A. Alors λ est une racine de P_A. Autrement dit, les valeurs propres d'une matrice sont exactement les racines de son polynôme caractéristique.

Définition (**Multiplicité**) . Soit f un endormorphisme représenté par une matrice A dans un base $\mathcal{B}$. La multiplicité d'une valeur propre $\lambda \in \mathbb{K}$ de f est exactement la multiplicité de λ en tant que racine de P_A.

Exemple (Recherche des éléments propres) Soit $A \in \mathcal{M}_3(\mathbb{R})$ telle que

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix}$$

Déterminons les éléments propres de A. Commençons par calculer son polynôme caractéristique.

$$\begin{aligned} P_A &= \begin{vmatrix} 2-X & 1 & 1 \\ 1 & 2-X & 1 \\ 0 & 0 & 3-X \end{vmatrix} = (3-X)\begin{vmatrix} 2-X & 1 \\ 1 & 2-X \end{vmatrix} \\ &= (3-X)((2-X)^2 - 1) \\ &= (3-X)(4 - 4X + X^2 - 1) \\ &= (3-X)^2(1-X) \end{aligned}$$

D'où : $Sp(A) = \{3, 1\}$ et 3 est de multiplicité 2.

3.1.4 Critères de Diagonalisation

Stabilité et somme directe

Définition (**Sous espace stable**) . Un s.e.v F de E est dit stable par f ssi :

$$\forall v \in F, f(v) \in F \quad \text{i.e} \quad f(F) \subset F$$

Propriété (**Sous-espace propre et stabilité**) . Un sous espace propre E_λ de f est stable par f.

Définition (**Somme Directe**) . Soit E un espace vectoriel et $(F_i)_{i\in I}$ une famille finie de sous-espaces vectoriels de E. On dit que les F_i sont en somme directe ssi pour tout vecteur $v \in \sum_{i\in I} F_i$, il existe une **unique** décomposition de v comme somme de vecteurs $v_i \in F_i, \forall i \in I$.

Propriété (**Somme Directe**) . Toute somme de sous espaces propres d'un même endomorphime est directe.

Caractérisation d'un endormorphisme diagonalisable

Proposition Deux propositions très utiles pour déterminer si un endomorphime est diagonalisable : Un endomorphisme f est diagonalisable ssi ses sous espaces propres sont supplémentaires dans E *i.e* :

$$E = \bigoplus_{\lambda\in Sp(f)} E_\lambda \quad \Longleftrightarrow \quad \sum_{\lambda\in Sp(f)} dim(E_\lambda) = n$$

Le second point de la proposition précédente nous donne une approximation de la dimension d'un sous-espace propre.

Propriété (**Multiplicité et encadrement**) . Si λ est valeur propre de f de multiplicité α alors :

$$1 \leq \dim(E_\lambda) \leq \alpha$$

On peut donc déterminer un critère fondamental de diagonalisation d'endomorphismes (matrices).

Critère (**Critère Fondamental de Diagonalisation**) . Un endomorphime f est diagonalisable ssi

- son polynôme caractéristique est scindé sur $\mathbb{K}$ et pour chaque valeur propre, la dimension du sous-espace propre associé est égale à la multiplicité de la valeur propre.
- la somme de la dimension de ses sous-espaces propres est n.

Corollaire (**Critère de Diagonalisation**) . Si f admet n valeurs propres distinctes alors f est diagonalisable.

3.1.5 Méthode Pratique de Diagonalisation

Maintenant que nous savons déterminer si un endomorphisme est diagonalisable, mettons en place une méthode pratique nous donnant la base de diagonalisation, les sous-espaces propres, et la matrice diagonale.
Soit f un endomorphisme représenté par une matrice A dans une base $\mathcal{B}$. Pour diagonaliser f, il suffit de :

1. Calculer le polynôme caractéristique de A.
2. Factoriser le polynôme trouvé pour déterminer le spectre de f.
3. Pour chaque valeur propre λ distincte de f :
   (a) Calculer $A - \lambda\, Id$.
   (b) Echelonner la matrice obtenue par le pivot de Gauss. Avec mémorisation.
   (c) Extraire du pivot une base du sous-espace.
4. Maintenant on peut se trouver face à deux situations :
   - Si pour chaque valeur propre, la dimension du sous-espace engendré est égale à la multiplicité de celle-ci, alors f est digonalisable. On poursuit alors la diagonalisation.
   - Sinon, f n'est pas diagonalisable. On s'arrête ici...
5. La base de diagonalisation sera donc l'union des bases des sous-espaces propres. La matrice diagonale sera composée des valeurs propres de f. Pour effectuer le changement de base, on veillera bien à inverser la matrice de passage.

Exemple Ici un exemple pratique.

3.1.6 Introduction à la trigonalisation

La diagonalisation semble être un outil très fort pour effectuer toute sorte de calculs sur des matrices/endomorphismes. Mais en pratique, on se rend compte que très peut de matrices sont diagonalisables. On peut donc trouver une autre forme de réduction d'un endomorphisme représenté par une matrice mais, cette fois ci, pas sous forme triagonale, mais triangulaire. On va ainsi définir une nouvelle forme de réduction et trouver des critères pour savoir si une matrice est trigonalisable.

Définition (**Trigonalisation**) . Un endomorphisme f représenté par une matrice A dans une base $\mathcal{B}$ est dit trigonalisable si il existe une base $\mathcal{B}'$ dans laquelle il est représenté par une matrice triangulaire supérieure.

De même une matrice A est dite trigonalisable si elle est semblable à une matrice triangulaire supérieure.

Proposition Un endomorphime f représenté par une matrice A est trigonalisable ssi A l'est.

Critère (**Fondamental de Trigonalisation**) . Soient E un $\mathbb{K}$-espace vectoriel et $f \in \mathcal{L}(E)$. f est trigonalisable ssi son polynôme caractéristique est scindé sur $\mathbb{K}$.

Corollaire (**La puissance de** $\mathbb{C}$) . Tout endormorphisme est trigonalisable sur $\mathbb{C}$.

On remarque que l'élaboration d'une méthode exaustive de trigonalisation de matrices semble plus complexe que la diagonalisation. Nous en verrons une dans le chapitre dédié, mais pour cela, il nous faut définir de nouveaux object qui vont nous permettre d'aller plus loin dans la théorie de la réduction.

3.2 Polynômes d'endomorphismes

Nous savons que l'ensemble des matrices carrées à coefficients dans un corps $\mathbb{K}$ (réel ou complexe) est muni d'une structure d'espace vectoriel. Ayant déjà étudié en détail les polynômes, on pourrait se demander si il est possible d'étudier des polynômes de matrices et par extension des polynômes d'endomorphismes. Pourrait-on, le cas échéant, en fonction des propriétés d'un tel polynôme, en déduire des propriétés sur des endomorphismes/matrices tels que des critères de diagonalisation ou de trigonalisation ? C'est ce que nous allons étudier dans cette section.

3.2.1 Définition et premières propriétés

Définition (**Polynôme d'endomorphisme**) . Soit $P = \sum_{i=0}^{n} a_i X^i$ un polynôme et f un endomorphisme, on peut évaluer P en f de la façon suivante :

$$P(f) = \sum_{i=0}^{n} a_i f^i$$

Par convention, $f^0 = \mathrm{Id}$.

Définition (**Polynôme de matrice**) . Soit $P = \sum_{i=0}^{n} a_i X^i$ un polynôme et $A \in \mathcal{M}_n(\mathbb{K})$ une matrice carrée. On peut évaluer P en A de la façon suivante :

$$P(A) = \sum_{i=0}^{n} a_i A^i$$

De même que précédement, on pose $A^0 = \mathrm{Id}$.

Proposition On peut "résumer" ces deux définitions en posant une application qui, à un polynôme, lui associe ce même polynôme évalué en un endomorphisme prédéterminé.
Plus formellement, $\forall f \in \mathcal{L}(E)$, l'application

$$\Phi_f : \begin{cases} \mathbb{K}[X] \longrightarrow \mathcal{L}(E) \\ P \longmapsto P(f) \end{cases}$$

est un morphisme d'algèbre. Autrement dit, c'est une application entre deux algèbres qui respecte leur structure.
Elle vérifie donc $\forall \alpha \in \mathbb{K}, \forall P, Q \in \mathbb{K}[X]$:

- $\Phi(\alpha P) = \alpha \Phi(P)$
- $\Phi(P + Q) = \Phi(P) + \Phi(Q)$
- $\Phi(P \times Q) = \Phi(P) \circ \Phi(Q)$

Ces propriétés nous serviront tout au long du chapitre et durant les suivant et s'avèrent très utiles lors de calculs.

Remarque (Extension aux matrices) De même, on peut définir un morphisme d'algèbre pour évaluer des polynômes par des matrices :

$$\forall A \in \mathcal{M}_n(\mathbb{K}), \quad \Phi_A : \begin{cases} \mathbb{K}[X] \longrightarrow \mathcal{L}(E) \\ P \longmapsto P(A) \end{cases}$$

Cette application possède les mêmes propriétés que la précédente.

Rapidement, on s'apperçoit que les polynômes d'endomorphismes possèdent beaucoup de bonnes propriétés calculatoires, notamment pour les relations de similitudes entre matrices, autrement appelées "changement de base".

Propriété (**Matrices semblables et polynômes**) . Soient $A, B \in \mathcal{M}_{\mathbb{K}}$ deux matrices semblables liés par la matrice M. Soit $P \in \mathbb{K}[X]$, on a :

$$\forall k \in \mathbb{N}, A^k = MB^kM^{-1} \quad \text{et} \quad P(A) = MP(B)M^{-1}$$

Propriété (**Commutation**) . Soit $f \in \mathcal{L}(E)$, deux polynômes en f commutent.

3.2.2 Polynômes Annulateurs

Nous connaissons déjà bien le concept de racine d'un polynôme réel ou complexe. Qu'en est-il des matrices/endomorphismes annulant un polynôme ?

Définition (**Polynôme Annulateurs**) . Soit $f \in \mathcal{L}(E)$ un polynôme $P \in \mathbb{K}[X]$ est dit annulateur de f si $P(f) = 0_{\mathcal{L}(E)}$. De même pour les matrices, on dit que P annule $A \in \mathcal{M}_n(\mathbb{K})$ si $P(A) = 0_{\mathcal{M}_n(\mathbb{K})}$.

On remarquera facilement que le polynôme nul annule tout le monde...pas très intéressant.

Remarque (Pas le même 0 ?) Vous remarquerez qu'en fonction de l'espace sur lequel est défini un polynôme (endomorphisme ou matrice), si le polynôme est un annulateur d'un élément de cet espace, le 0 obtenu n'est pas le même. En effet, pour un polynôme défini sur les endomorphismes, le 0 obtenu sera **l'application nulle**. Alors que pour les matrices, on obtient la **matrice nulle**. Il est important de bien différencier les ensembles de définition des applications que l'on utilise.

Proposition Tout endomorphisme de E admet un polynôme annulateur non nul. Idem pour les matrices.

Propriété (**Valeur propre et racine**) . Soit P un polynôme annulateur de $f \in \mathcal{L}(E)$, alors toute valeur propre de f est racine de P (la réciproque est généralement fausse).

Remarque Dans le cas ou f est diagonalisable, on peut facilement construire un annulateur. Soit $P \in \mathbb{K}[X]$ et $f \in \mathcal{L}(E)$. Soit $\mathrm{Sp}(f) = \{\alpha_1, \alpha_2, \ldots, \alpha_p\}, p \in N$ le spectre de f. D'après la propriété précédente, tous les α_i sont racine de P. Autrement dit :

$$\forall i \in \{1, \ldots, p\}, \quad (X - \alpha_i) \mid P$$

Posons :

$$M_f := (X - \alpha_1)(X - \alpha_2) \ldots (X - \alpha_p)$$

On a donc :

$$M_f \mid P$$

Autrement dit, P **est un multiple de** $M_f = (X - \alpha_1)(X - \alpha_2) \ldots (X - \alpha_p)$.

3.2.3 Théorèmes Fondamentaux

Ci-dessous, un des théorèmes les plus forts d'algèbre linéaire cette année. La démonstration n'est absolument pas triviale...

Théorème (**Arthur Cayley et William Hamilton**) . Pour toute matrice et endomorphisme, **le polynôme caractéristique est annulateur**.

Maintenant que nous avons vu pas mal de propriétés sur les polynômes annulateurs, le polynôme caractéristique et quelques relations de divisibilité en fonction des valeurs propres d'un endomorphisme/matrice, intéressons nous plus profondément aux ordres de divibilités entre des polynômes.

Saut mention contraire, on notera M_f le polynôme minimal d'un endomorphisme f et P_f son polynôme caractéristique dans une certaine base.

Définition (**Polynôme Minimal**) . Soit $f \in \mathcal{L}(E)$. Il existe un unique polynôme, noté M_f, tel que l'ensemble des polynômes annulateurs de f soient des multiples de M_u.

Si on interprète cette définition de façon ensembliste, le polynôme minimal d'un endomorphisme peut être vu comme le minimum (au sens de la divisibilité/degré) de l'ensemble des annulateurs.

Proposition Soit $f \in \mathcal{L}(E)$, on a :

- Si $\mathrm{Sp}(f) = \{\alpha_1, \alpha_2, \ldots, \alpha_p\}, p \in N$, alors $(X-\alpha_1)(X-\alpha_2)\ldots(X-\alpha_p)$ divise M_f.
- M_f divise P_f

Propriété (**Noyau et sev stable**) . Pour tout $P \in \mathbb{K}[X]$, et pour tout $f \in \mathcal{L}(E)$, $\ker P(u)$ est un sev de E stable par f.

Autrement dit, le noyau d'un polynôme évalué en un endomorphisme est un sev de l'espace de départ et, est de plus stable par cet endomorphisme.

Lemme (Noyaux) Soient P et Q deux polynômes premiers entre eux et $f \in \mathcal{L}(E)$, alors :

$$\boxed{\ker(PQ(f)) = \ker(P(f)) \oplus \ker(Q(f))}$$

On peut donc en déduire un cas plus général...

Théorème (**Généralisation du Lemme des noyaux**) . Soit P un polynôme annulateur de $f \in \mathcal{L}(E)$, et $P = P_1 P_2 \ldots P_k$ une décomposition de P en produit de polynômes deux à deux premiers entre eux, on a donc :

$$E = \bigoplus_{i \in \{1,\ldots,k\}} \ker(P_i(f))$$

Autrement dit, une factorisation d'un polynôme annulateur de f en produit de polynômes deux à deux premiers entre eux nous donne un décomposition de E en sous-espaces, d'une part stables par f mais aussi en somme directe dans E.

On peut donc en déduire le théorème suivant, nous donnant de nouveaux critères de diagonalisation.

Théorème (**Nouveau critère de diagonalisation**) . $f \in \mathcal{L}(E)$ est diagonalisable ssi

- elle admet un **polynôme annulateur** scindé à racine simples.
- ssi son **polynôme minimal** est scindé à racines simples.

Critère (**Ultime Critère**) . $f \in \mathcal{L}(E)$ est diagonalisable ssi son polynôme minimal est scindé à racines simples.

3.2.4 Endomorphismes Nilpotents

Définition (**Endomorphisme/Matrice Nilpotent**) . Soient $f \in \mathcal{L}(E)$ et $A \in \mathcal{M}_n(\mathbb{K})$, on dit que f ou A est nilpotent ssi

$$\exists k \in \mathbb{N}^* \text{ tel que : } \begin{cases} f^k = 0_{\mathcal{L}(E)} \text{ et } L^{k-1} \neq 0_{\mathcal{L}(E)} \\ A^k = 0_{\mathcal{M}_n(\mathbb{K})} \text{ et } A^{k-1} \neq 0_{\mathcal{M}_n(\mathbb{K})} \end{cases}$$

On nomme alors l'entier k **indice** de nilpotence.

Propriété (**Matrice strictement triangulaire et nilpotence**) . Tout matrice strictement triangulaire supérieure de $\mathcal{M}_n(\mathbb{K})$ est nilpotente d'indice inférieur à n.

Propriété (**Nilpotence et polynôme minimal**) . Soient $f \in \mathcal{L}(E)$ f est nilpotente d'indice k ssi $M_f = X^k$.

La même propriété est vraie pour n'importe quelle matrice $A \in \mathcal{M}_n(\mathbb{K})$.

3.3 Trigonalisation

Redéfinissons proprement la trigonalisation.

Définition (**Trigonalisation**) . Un endomorphisme f représenté par une matrice A dans une base $\mathcal{B}$ est dit trigonalisable si il existe une base $\mathcal{B}'$ dans laquelle il est représenté par une matrice triangulaire supérieure.

De même une matrice A est dite trigonalisable si elle est semblable à une matrice triangulaire supérieure.

Propriété (**Lien matrice/endomorphisme**) . Un endomorphime f représenté par une matrice A est trigonalisable ssi A l'est.

Critère (**Fondamental de Trigonalisation**) . Soient E un $\mathbb{K}$-espace vectoriel et $f \in \mathcal{L}(E)$. f est trigonalisable ssi son polynôme caractéristique est scindé sur $\mathbb{K}$.

Corollaire (**La puissance de** $\mathbb{C}$) . Tout endormorphisme est trigonalisable sur $\mathbb{C}$.

3.3.1 Méthode de Dunford

Définition (**Sous-espace caractéristique**) . Soit f un endomorphisme trigonalisable de polynôme caractéristique

$$P_f = (\lambda_1 - X)^{\alpha_1} \dots (\lambda_p - X)^{\alpha_p}$$

On appelle sous-espace caractéristique de f associé à la valeur propre λ_i, le sous-espace vectoriel

$$\boxed{F_i = \ker((u - \lambda_i \operatorname{Id})^{\alpha_i})}$$

Remarque Attention à la différence entre les sous-espaces propres et les sous-espaces caractéristiques ! Les seconds sont définis en fonction de la multiplicité de la valeur propre alors que les premiers non.

Théorème (**Trigonalisation et sous-espace caractérisitique**) . Soit f un endomorphisme trigonalisable

- Les F_i sont stables par f et :
$$E = \bigoplus_{i\in\{1,\dots,p\}} F_i$$
- Pour toute valeur propre, le sous-espace **propre** correspondant est inclus dans le sous-espace **caractéristique** correspondant.
$$\forall i \in \{1,\dots,p\}, \quad E_i \subset F_i$$
- Pour toute valeur propre, la dimension du sous-espace **caractéristique** correspondant est égale à sa multiplicité.
$$\forall i \in \{1,\dots,p\}, \quad \dim(F_i) = \alpha_i$$

En résumé, la méthode pratique de trigonalisation se résume en quelques étapes. Soit f un endomorphisme trigonalisable représenté par une matrice A dans une base $\mathcal{B}'$. Pour trigonaliser f (ou A), il faut :

- Calculer le polynôme caractéristique $P_f = P_A$. En le factorisant, en déduire le spectre de f et, pour chaque valeur propre, son ordre de multiplicité. Si P_f est scindé, alors f est trigonalisable.
- Pour chaque valeur propre, on calcule le sous-espace propre associé, on en donne la dimension et une base.
- Si pour chaque valeur propre on a égalité entre la dimension du sous-espace propre et la multiplicité de la valeur propre, alors f est diagonalisable. Sinon, on continue la trigonalisation.
- Pour chaque valeur propre (sympathique) pour laquelle on a égalité entre multiplicité et dimension du sep, on détermine une base du sous-espace propre. Le bloc correspondant sera diagonal.
- Pour chaque valeur propre (pénible) pour laquelle on n'a pas égalité entre multiplicité et dimension du sep, on détermine le drapeau complet $E_\lambda = K_1 \subset \cdots \subset K_p = F_\lambda$. On construit une base de $\mathcal{B}_i$ de F_λ par la méthode des drapeaux.
- On calcule l'image de chaque $\mathcal{B}_i$ par f, ce qui fournit le bloc correspondant à la valeur propre λ_i.
- Une base de trigonalisation est obtenue en faisant l'union des bases obtenues précédement.
- La matrice triangulaire (de Dunford) est obtenue à partir des bloc précédents.

Exemple (Trigonalisation) Soit $A \in \mathcal{M}_3(\mathbb{R})$ telle que :
$$A = \begin{pmatrix} 2 & -1 & -1 \\ 2 & 1 & -2 \\ 3 & -1 & -2 \end{pmatrix}$$

Calculons son polynôme caractéristique $P_A \in \mathbb{R}[X]$:
$$\begin{aligned} P_A &= \begin{vmatrix} 2-X & -1 & -1 \\ 2 & 1-X & -2 \\ 3 & -1 & -2-X \end{vmatrix} \\ &= -X^3 + X^2 + X - 1 \\ &= -(X-1)^2(X+1) \end{aligned}$$

On a donc $Sp(A) = \{-1, 1\}$ où la valeur propre 1 est de multiplicité 2. Puisque P_A est scindé, A est *trigonalisabe* dans $\mathbb{R}$. Calculons ses sous-espaces caractéristiques :

- $\ker(A + I_3)$:

$$A + I_3 = \begin{array}{c} \begin{matrix} C_1 & C_2 & C_3 \end{matrix} \\ \begin{pmatrix} 3 & -1 & -1 \\ 2 & 2 & -2 \\ 3 & -1 & -1 \end{pmatrix} \end{array} ; \quad \begin{array}{c} \begin{matrix} C_2 & C_1+3C_2 & C_3-C_2 \end{matrix} \\ \begin{pmatrix} -1 & 0 & 0 \\ 2 & 8 & -4 \\ -1 & 0 & 0 \end{pmatrix} \end{array} ; \quad \begin{array}{c} \begin{matrix} C_2 & C_1+3C_2 & C_1+5C_2-2C_3 \end{matrix} \\ \begin{pmatrix} -1 & 0 & 0 \\ 2 & 8 & 0 \\ -1 & 0 & 0 \end{pmatrix} \end{array}$$

  On a donc $\ker(A + I_3) = \text{vect}((1, 1, 2))$. La dimension du sous-espace caractéristique E_{-1} associé à la valeur propre -1 est égale à sa multiplicité, nous n'avons plus rien à faire ici.

- $\ker(A - I_3)$:

$$A - I_3 = \begin{array}{c} \begin{matrix} C_1 & C_2 & C_3 \end{matrix} \\ \begin{pmatrix} 1 & -1 & -1 \\ 2 & 0 & -2 \\ 3 & -1 & -3 \end{pmatrix} \end{array} ; \quad \begin{array}{c} \begin{matrix} C_1 & C_2+C_1 & C_3+C_1 \end{matrix} \\ \begin{pmatrix} 1 & 0 & 0 \\ 2 & 2 & 0 \\ 3 & 2 & 0 \end{pmatrix} \end{array}$$

  On a donc $\ker(A - I_3) = \text{vect}((1, 0, 1))$. La dimension du sous-espace caractéristique associé à la valeur propre 1 est inférieure à sa multiplicité. Il nous faut donc construire une suite de noyau itérés. Déterminons $\ker((A - I_3)^2)$:

$$(A - I_3)^2 = \begin{pmatrix} -4 & 0 & 4 \\ -4 & 0 & 4 \\ -8 & 0 & 8 \end{pmatrix}$$

  On voit immédiatement que $\ker((A - I_3)^2) = \text{vect}((0, 1, 0))$. Le drapeau associé à la valeur propre 1 est donc :

$$\{0\} \subset \text{vect}((1, 0, 1)) \subset \text{vect}((0, 1, 0))$$

On peut donc construire une base de trigonalisation $\mathcal{B}$ de la forme :

$$\mathcal{B} = ((1, 0, 1), (0, 1, 0), (1, 1, 2))$$

De matrice de passage :

$$Pass(\mathcal{B}, \mathcal{C}) = P = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 2 \end{pmatrix}$$

Par inversion, on obtient :

$$Pass(\mathcal{C}, \mathcal{B}) = P^{-1} = \begin{pmatrix} 2 & 0 & -1 \\ 1 & 1 & -1 \\ -1 & 0 & 1 \end{pmatrix}$$

Il ne reste plus qu'à construire la matrice triangulaire supérieure semblable à A dans la base $\mathcal{B}$. Pour cela, il faut exprimer l'image des vecteurs de la base $\mathcal{B}$ par A dans la base $\mathcal{B}$:

$$\left[\begin{pmatrix} 2 & -1 & -1 \\ 2 & 1 & -2 \\ 3 & -1 & -2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right]^{\mathcal{B}} = \left[\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right]^{\mathcal{B}} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$$\left[\begin{pmatrix} 2 & -1 & -1 \\ 2 & 1 & -2 \\ 3 & -1 & -2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right]^{\mathcal{B}} = \left[\begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix} \right]^{\mathcal{B}} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$$

$$\left[\begin{pmatrix} 2 & -1 & -1 \\ 2 & 1 & -2 \\ 3 & -1 & -2 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}\right]^{\mathcal{B}} = \left[\begin{pmatrix} -1 \\ -1 \\ -2 \end{pmatrix}\right]^{\mathcal{B}} = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}$$

On a donc que A est semblable à la matrice $J \in \mathcal{M}_3(\mathbb{R})$ de matrice de passage P^{-1} telles que :

$$A = PJP^{-1} \quad \text{avec} \quad J = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Chapitre 4

Espaces Euclidiens

Contents

Maintenant que nous avons vu beaucoup de propriétés des espaces vectoriels (en particulier réels), plaçons nous dans certains de ces espaces vérifiant quelques propriétés supplémentaire. Ainsi, nous allons pouvoir définir de nouveaux objects, trouver de nouvelles propriétés, etc...
Une des application de ce que l'on verra dans ce chapitre est l'Intelligence Artificielle, en effet, on se place souvent dans des espaces euclidiens pour obtenir des propriétés "sympathiques" sur des endomorphismes/matrices.

Dans ce chapitre nous nommerons E un espace euclien.

4.1 Contexte

4.1.1 Produit Scalaire et Norme

Définition (**Produit Scalaire**) . Soit E un $\mathbb{R}$ espace vectoriel. L'application :

$$f : E \times E \longrightarrow \mathbb{R}$$

est un produit scalaire sur E ssi f vérifie :

- Linéarité à droite : $\forall x, y, z \in E, \forall \lambda \in \mathbb{R}$:

$$f(x, \lambda y + z) = \lambda f(x, y) + f(x, z)$$

- Linéarité à gauche : $\forall x, y, z \in E, \forall \lambda \in \mathbb{R}$:

$$f(\lambda x + y, z) = \lambda f(x, z) + f(y, z)$$

- Symétrique : $\forall x, y \in E, \quad f(x, y) = f(y, x)$
- Définie Positive : $\forall x \in E, \quad f(x, x) \geqslant 0$ et $f(x, x) = 0 \iff x = 0$

En général on note un produit scalaire $\langle ., . \rangle$ ou $(.|.)$. On dit que x et y sont **orthogonaux** (noté $x \perp y$) ssi $(x|y) = 0$.

L'orthogonalité est une notion clé de ce chapitre.

Définition (**Norme**) . Soit E un $\mathbb{R}$ espace vectoriel. L'application : $N : E \longrightarrow \mathbb{R}$ est une norme sur E ssi $\forall x, y \in E, \forall \lambda \in R$ elle vérifie les propriétés suivantes :

- $N(x) \geqslant 0$
- $N(x) = 0 \implies x = 0$
- $N(\lambda x) = |\lambda| N(x)$
- $N(x + y) \leqslant N(x) + N(y)$

En général, on note une norme $\|.\|$.

Remarque Quelques remarques concernant les normes et produits scalaires. L'application

$$\|.\| : \begin{cases} E \longrightarrow \mathbb{R} \\ x \longmapsto \sqrt{\langle x, x \rangle} \end{cases}$$

est la **norme euclidienne** de E. A partir d'un produit scalaire, on peut donc facilement définir une norme en prenant la racine carrée de celui-ci.

4.1.2 Espace Euclidien et Inégalités

Définissons plus formellement les espaces utilisés dans ce cours.

Définition (**Espace Préhilbertien Réel**) . On appelle espace préhilbertien réel un $\mathbb{R}$ espace vectoriel muni d'un produit scalaire. On appelle **espace euclidien** un $\mathbb{R}$ espace vectoriel de dimension finie muni d'un produit scalaire.

Théorème (**Inégalité de Cauchy-Schwarz**) . Soit E un espace préhilbertient réel, alors :

$$\forall x, y \in E, \quad \boxed{(x|y)^2 \leqslant (x|x)(y|y) \quad \text{ou} \quad |(x|y)| \leqslant ||x||.||y||}$$

On a aussi : $|(x|y)| = ||x||.||y||$ ssi les vecteurs x et y sont liés.

On peut reformler en français cette inégalité par : *"La valeur absolue du produit scalaire est inférieure au produit des normes"*.

Proposition (Inégalité Striangulaire) Soit E un espace préhilbertient réel, alors :

$$\forall x, y \in E, \quad \boxed{||x + y|| \leqslant ||x|| + ||y||}$$

De même que pour l'inégalité de Cauchy Schwarz, on a un cas d'égalité :

$$\|x + y\| = \|x\| + \|y\| \quad \iff \quad y = 0 \text{ ou } x = \lambda y \text{ avec } \lambda \geqslant 0$$

Remarque Un espace euclidien est donc, en somme, simplement un espace vectoriel en dimension finie avec lequel on a l'habitude de travailler pour lequel on a défini un produit scalaire. Les deux inégalités ci-dessus sont fondamentales dans ce chapitre et seront très régulièrement utilisées.

Proposition (Théorème de Pythagore) Grâce à cette notion d'orthogonalité entre deux vecteur, on peut généraliser le théorème de Pythagore de la façon suivante :

Soient $x, y \in E$, on a :

$$x \perp y \iff \|x\|^2 + \|y\|^2 = \|x+y\|^2$$

4.1.3 Expression Analytique du produit scalaire

On note ici $x = (x_1, \ldots, x_n)$ un vecteur de E. Analytiquement, on notera $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ ses coordonnées dans une base de E.

Proposition Soient $x, y \in E$, et $(e_1, \ldots, e_n)$ une base de E. On a alors :

$$x = \sum_{i=1}^{n} x_i e_i \quad \text{(idem pour y)}$$

Donc, par bilinéarité du produit scalaire, on obtient la forme suivante :

$$(x|y) = \sum_{i=1}^{n} x_i y_i (e_i|e_j)$$

Définition (**Matrice du produit scalaire**) . Soit E un espace euclidien de base $\mathcal{B}$. On appelle matrice du produit scalaire $(.|.)$ la matrice A de taille n symétrique telle que :

$$(A_{i,j})_{i,j \in [\![1,n]\!]} \quad \text{où} \quad \forall i, j \in [\![1, n]\!] a_{i,j} = (e_i|e_j)$$

Une fois la matrice du produit scalaire définie pour un certain produit scalaire, on peut calculer chaque produit sclaire à partir des vecteurs de base, ce qui simplifie largement les calculs.

Proposition Soit A la matrice d'un produit scalaire de E, pour tout $x, y \in E$, on a donc :

$$(x|y) = {}^t XAY$$

4.2 Orthogonalité

4.2.1 Définitions et Propriétés

Définition (**Famille Orthogonale/Orthonormée**) . Une famille $(e_1, \ldots, e_p)$ de vecteurs non nuls **deux à deux orthogonaux** est dite orthogonale. De plus, si $\forall i \in [\![1, n]\!],\ ||x_i|| = 1$ la famille est dite orthonormée.

Remarque Si $p = n$, une famille orthogonale/orthonormée $(e_1, \ldots, e_p)$ est donc une base orthogonale ou une base orthonormée de E. Dans le cas de bases orthogonales on parlera de BOG et de BON si elle est orthonormée.

Définition (**Orthogonal**) . Soit $A \subset E$.
On appelle orthogonal de A dans E tous les vecteurs de E orthogonaux à tous les vecteurs de A.

$$\boxed{A^{\perp} = \{x \in E,\ \forall y \in A,\ x \perp y\}}$$

Remarque Les propriétés de bilinéarité d'un produit scalaire nous permettent de calculer des orthogonaux efficacement. En effet si nous connaissons une famille génératrice $\mathcal{F}$ d'un sev $F \subseteq E$, alors $y \in F^{\perp}$ ssi pour tout vecteur $f \in \mathcal{F}$, on ait $y \perp f$.

Propriété (**Orthogonal**) . Un orthogonal $A^\perp \subseteq E$ vérifie les propriétés suivantes :

- $A^\perp$ est un s.e.v de E
- $E^\perp = \{0_E\}$
- $A \subset (A^\perp)^\perp$
- $A \subset B \implies B^\perp \subset A^\perp$
- $(A \cup B)^\perp = A^\perp \cap B^\perp$
- $E = A \oplus A^\perp$ (dans un espace euclidien)

4.2.2 Matrice d'un produit scalaire et changement de base

Comme précédement, on peut réutiliser la forme analytique d'un produit scalaire et remarquer qu'elle dispose de quelques propriétés supplémentaire dans le cas d'une BOG.

Propriété (**Matrice du produit scalaire et BOG**) . Soient E un espace euclidien, $\mathcal{B}$ sa base et A la matrice d'un de ses produit scalaires.

- $\mathcal{B}$ est **orthogonale** ssi A est diagonale.
- $\mathcal{B}$ est **orthonormale** ssi A est l'identité.

Analytiquement, dans une base orthonormé $\mathcal{B} = (e_1, \dots, e_n)$ on a :

- $(x|y) = {}^tXY = \sum_{i=1}^n x_i y_i$
- $(x|e_i) = x_i, \forall i \in [\![1, n]\!]$

D'autre part, on peut appliquer le théorème de changement de base pour déterminer la matrice d'un même produit scalaire de E dans une nouvelle base de E.

Proposition Soit E un espace euclidien, $\mathcal{B}, \mathcal{B}'$ deux bases de E et A la matrice d'un produit scalaire de E. Soit P la matrice de changement de base de $\mathcal{B}$ vers $\mathcal{B}'$, on a :

$$B = {}^tPAP$$

Et B est la matrice du même produit scalaire de E mais exprimé dans la base $\mathcal{B}'$.

4.3 Projection Orthogonale

Maintenant que nous avons défini l'orthogonalité, on peut définir une nouvelle application. Soit $F \subseteq E$ un sev et $F^\perp$ son orthogonal dans E. Par propriété de l'orthogonal, on sait que $F^\perp$ est un sev de E et, de plus :

$$E = F \oplus F^\perp$$

Par propriété de la supplémentarité, on peut décomposer chaque vecteur $x \in E$ en somme de deux vecteurs $x_F \in F$ et $x_{F^\perp} \in F^\perp$ de façon unique.
On peut donc penser à une application, qui, à chaque vecteur de E lui associe sa composante dans l'un des deux sev de E orthogonaux. Application que nous nommerons **projection orthogonale**.

Théorème (**Caractéristiques de l'Orthogonal**) . Soit E un espace euclidien et F un s.e.v de E, on a :

- $E = F \oplus F^\perp$
- $F = (F^\perp)^\perp$

Définition (**Projection Orthogonale**) . Soit F un s.e.v de E, on appelle projection orthogonale sur F parallèlement à $F^\perp$ l'application :

$$p_F : \begin{cases} E \longrightarrow F \\ x = x_F + x_{F^\perp} \mapsto x_F \end{cases}$$

Le projeté orthogonal de $x \in E$ est l'unique vecteur $p_F(x)$ qui vérifie : $x - p_F(x) \in F^\perp$.
On a aussi : $Im(p_F) = F$ et $ker(p_F) = F^\perp$

Proposition Le projeté orthogonal nous permet de caractériser des vecteurs "plus proches" d'un sev F car :

$$\forall x \in E, \quad \|x - p_F(x)\| = \min_{y \in F}(\|x - y\|)$$

Proposition (Calcul de projeté orthogonal) Soit $\mathcal{B} = (e_1, \ldots, e_p)$ une base orthogonale de F, pour tout $x \in E$ on peut calculer explicitement son projeté orthogonal sur F par la formule suivante :

$$\forall x \in E, \quad p_F(x) = \sum_{i=1}^{p} \frac{(x|e_i)}{(e_i|e_i)} e_i$$

Dans le cas d'une base orthonormale, $\forall i \in [\![1, p]\!],\ (e_i|e_i) = 1$, on a donc la formule :

$$\forall x \in E, \quad p_F(x) = \sum_{i=1}^{p} (x|e_i) e_i$$

Un base orthonormale apporte beaucoup de simplifications dans le calculs de projetés orthogonal et de produit scalaire (voir matrice de produit scalaire). Il serait donc utile de savoir passer d'une base quelconque d'un sev E à une base orthonormale. C'est ce que permet (en partie) le procédé de Gram-Schmidt.

4.4 Procédé de Gram-Schmidt - Orthogonalisation

Théorème (**Orthogonalisation par Gram-Schmidt**) . Soit $(x_1, \ldots, x_n)$ une famille libre de de vecteurs d'un sev E. $\forall i \in [\![1, p]\!]$, on note $H_i = vect(x_1, \ldots, x_p)$.
On définit une famille $(e_1, \ldots, e_p)$ orthogonale par récurrence :

$$e_1 = x_1$$

et $\forall i \in [\![1, p]\!]$

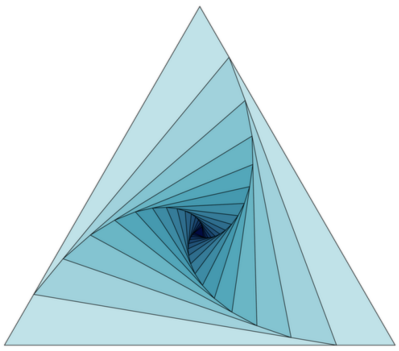
$$e_i = x_i - p_{H_{i-1}}(x_i) = x_i - \sum_{j=1}^{p} \frac{(x_i|e_j)}{(e_j|e_j)} e_j$$

Par propriété du projeté orthogonal, $\forall i \in [\![1, p]\!]$ la famille $(e_1, \ldots, e_i)$ est une base orthogonale de H_i.

Le procédé de Gram-Schmidt permet donc de "redresser" itérativement les vecteurs d'une base pour en faire une base orthogonale. Il ne manque plus qu'à "réduire" les vecteurs en vecteurs unitaires pour avoir une base orthonormale.

Corollaire (**Existence de bases orthonormales**) . Tous les espaces euclidiens admettent des bases orthonormés.

Topologie



Chapitre 1

Introduction - Les Réels

Contents

Rappellons les propriétés du principal espace que nous allons considérer dans ces chapitres, $\mathbb{R}$.

1.1 Majorant, Minorant, Supremum, Infimum

1.1.1 Définitions

Formellement, $\mathbb{R}$ est un corps totalement ordonné muni de 4 opérations compatibles avec cet ordre. Considérons ici un ensemble E et $A \subseteq E$ une partie de E.

Définition (**Majorant**) . On appelle majorant de A, un élément de $M \in E$ supérieur à tous les éléments de A. Plus formellement :

$$M \in E \text{ est un majorant de } A \iff \forall x \in A, M \geqslant x$$

Définition (**Minorant**) . De même que pour les majorants, on appelle minorant de A un élément $m \in E$ inférieur à tous les éléments de A. Plus formellement :

$$m \in E \text{ est un minorant de } A \iff \forall x \in A, m \leqslant x$$

Autrement dit, tous les éléments de A majorent m.

Exemple Soit $E = \mathbb{R}$ et $A = [0, 1] \subset E$. Alors $0, -1$ et $-\pi$ sont des minorants de A et $1, 7$ et e sont des majorants de A.

Définition (**Maximum/Minimum**) . Soit $A \subseteq E$. On appelle **maximum** de A un élément $x \in A$ qui majore tous les éléments de A. De même, un **minimum** de A est un élément $x \in A$ qui minore tous les éléments de A. On les notes généralement max(.) et min(.).

Remarque On repère rapidement la différence entre un maximum et un majorant. Un maximum a la propriété d'appartenir à la partie qu'il majore. Idem pour un minimum. Ces objets sont quand même limités. Si on prends $A =]0, 1[\subset \mathbb{R}$.
On remarque facilement que l'on ne peut pas trouver de minimum à cette partie. Il existe une infinité de minorants mais si l'on souhaite minimiser A de façon "plus fine", cela risque de ne pas suffire. On va donc définir les supremum et infimum.

Définition (**Supremum**) . Soit $A \subseteq E$, on appelle supremum de A le plus petit des majorants de A. Plus formellement :

$$x \in E \text{ est un supremum de } A \iff \begin{cases} \forall a \in A, x \geqslant a \\ m = \min(\{m \in E, \forall a \in A, m \geqslant a\}) \end{cases}$$

On le note généralement *sup* et on parle de "borne sup".

Définition (**Infimum**) . Soit $A \subseteq E$, on appelle infimum de A le plus grand des minorants de A. Plus formellement :

$$x \in E \text{ est un infimum } deA \iff \begin{cases} \forall a \in A, x \leqslant a \\ x = \max(\{m \in E, \forall a \in A, m \leqslant a\}) \end{cases}$$

On le note inf et on parle de "borne inf".

Proposition S'il exite, un supremum ou un infimum est unique.

Remarque Moins formellement, les bornes inf et sup permettent de résoudre beaucoup de problèmes de majoration/minoration fine en nous permettant de "regarder" de l'autre côté de notre partie $A \subseteq E$.
Les bornes inf et sup sont surtout utilisées dans des espaces tels que $\mathbb{R}$ et $\mathbb{Q}$ où les éléments sont "très proches" (nous définirons cette notion plus tard). Intuitivement dans des ensembles tels que $\mathbb{N}$ ou $\mathbb{Z}$ nous n'avons pas besoin de tels objets.

Exemple Il peut arriver que l'on considère des parties qui n'admettent pas de majorants/minorants. Par exemple, $A = \mathbb{N} \subset \mathbb{R}$.

1.1.2 Propriétés et caractérisations

Théorème (**Existence**) . Dans $\mathbb{R}$ toute partie non vide et majorée admet un supremum. De même, Toute partie non vide et minorée admet un infimum.

Proposition Soit $A \subseteq \mathbb{R}$. $x \in \mathbb{R}$ est un supremum de A ssi

- x majore A
- pour tout $\varepsilon > 0, \exists a \in A, \quad a > x - \varepsilon$

En français, un supremum de A est un élément $x \in E$ qui majore A et tel que pour tout réel positif ε, on peut trouver un élément $a \in A$ entre x et $x - \varepsilon$.
On a la même propriété pour les infimum.

1.1.3 Densité des rationnels dans les réels

Propriété (**Archimède**) . Pour tout réel $x \in \mathbb{R}$, il existe un entier $n \in \mathbb{N}$, tel que $n > x$.

Démonstration La démonstration se fait par l'absurde.

On peut donc démontrer la proprosition principale de la densité de $\mathbb{Q}$ dans $\mathbb{R}$. On veut montrer qu'entre deux réels distincts, il existe une infinité de rationnels. Soient $x, y \in \mathbb{R}$ distincts. Il suffit juste de montrer qu'il existe un rationnel r entre x et y et, par suite, puisqu'un rationnel est aussi un réel, on pourra trouver un autre rationnel entre x et r puis entre r et y et ainsi de suite...

Proposition (Densité) Entre deux réels distincts, il existe une infinité de rationnels.

Chapitre 2

Espaces Métriques

Contents

Dans tout le début de ce chapitre, nous nous placerons dans un ensemble E quelconque.

2.1 Espace Métrique

2.1.1 Produit Scalaire

Quand on parle de produit scalaire, on pense souvent à l'application dans le cas euclidien permettant de vérifier si deux vecteurs sont orthogonaux, en réalité, il existe une multitute de produits scalaires agissant sur tout autant d'espaces.

Définition (**Produit Scalaire**) . Soit E un $\mathbb{R}$-espace vectoriel. Une application $\phi : E \times E \to \mathbb{R}$ de E est un produit scalaire ssi c'est une forme :

- **Bilinéaire :** $\forall x, y, z \in E, \forall \lambda, \mu \in \mathbb{R}$ on a :

$$\phi(\lambda x + \mu y, z) = \lambda\phi(x, z) + \mu\phi(y, z) \tag{2.1}$$
$$\phi(x, \lambda y + \mu z) = \lambda\phi(x, z) + \mu\phi(x, z) \tag{2.2}$$

- **Symétrique :** $\forall x, y \in E, \quad \phi(x, y) = \phi(y, x)$
- **Définie :** $\forall x \in E, \quad \phi(x, x) = 0_{\mathbb{R}} \Longrightarrow x = 0_E$
- **Positive :** $\forall x, y \in E, \quad \phi(x, y) \geq 0$

On dit qu'un produit scalaire est une forme puisqu'elle est définie de E dans un corps telle que :

$$\phi : E \longrightarrow \mathbb{R}$$

On note généralement un produit scalaire $\langle ., . \rangle$ ou $(.|.)$.

Remarque En pratique, pour montrer qu'une application est un produit scalaire, il suffit juste de montrer qu'elle est symétrique, linéaire et définie positive. La bilinéarité découle de la symétrie et de la linéarité.

Exemple (Produits Sclaires) Regardons quelques exemples de produits scalaires...

- En génométrie euclidienne, on utilise un produit scalaire permettant de déterminer si deux vecteur sont orthogonaux. Soient $A, B, C, D \in \mathbb{R}^2$, on définit alors :

$$\vec{AB}.\vec{CD} = AB \times CD \times \cos(\widehat{\vec{AB}\vec{CD}})$$

- Dans l'espace des fonctions continues sur un intervalle $[a, b]$ on peut définir le produit scalaire suivant :

$$\forall f, g \in \mathcal{C}^0([a, b]), \quad \langle f, g \rangle = \int_a^b f(x)g(x)\, dx$$

- Enfin, dans $\mathbb{R}^n$ on a le produit scalaire dit **euclidien** défini par :

$$\forall x, y \in \mathbb{R}^n \text{ tels que : } \begin{cases} x = (x_1, \ldots, x_n) \\ y = (y_1, \ldots, y_n) \end{cases} \quad \text{on a : } \langle x, y \rangle = \sum_{i=1}^{n} (x_i y_i)^2$$

Propriété (**Inégalité sur un produit scalaire**) . Soient $x, y \in E$, on a l'égalité suivante, valable pour tout produit scalaire :

$$\boxed{\langle x, y \rangle^2 \leq \langle x, x \rangle \times \langle y, y \rangle}$$

Remarque Ce résultat découle de l'inégalité de Cauchy-Schwarz, vu plus tard.

2.1.2 Norme

Une fois un produit scalaire défini sur un espace, on peut définir une application supplémentaire nous donnant plus d'informations sur un élément de l'espace. Nous allons donc définir une norme de deux façon, à partir d'un produit scalaire mais aussi de façon axiomatique.

Définition (**Norme**) . Soit E un $\mathbb{R}$-espace vectoriel. Un application $\|.\| : E \to \mathbb{R}$ de E est appelée **norme** ssi elle est :

- **Définie :** $\forall x \in E, \quad \|x\| = 0_{\mathbb{R}} \Longrightarrow x = 0_E$
- **Positive :** $\forall x \in E, \quad \|x\| \geqslant 0$
- **Positivement Homogène :** $\forall x \in E, \forall \lambda \in \mathbb{R}, \quad \|\lambda x\| = |\lambda| \|x\|$
- **Sous-additivité :** $\forall x, y \in E, \quad \|x + y\| \leqslant \|x\| + \|y\|$

Remarque (Notation et vocabulaire) Tout comme le produit scalaire, une norme est une application vérifiant quelques propriétés. Généralement, on note une norme $\|.\|$. Un espace muni d'une norme est appelé **espace normé**.
Dans un espace euclidien, une norme sert à "mesurer" la distance d'un point à l'origine.

Exemple Si on se place dans $\mathbb{R}$, le produit scalaire usuel sera la multiplication et la norme, la valeur absolue.

Proposition (Norme à partir du produit scalaire) Dans un $\mathbb{R}$-espace vectoriel E, à partir d'un produit scalaire, on peut facilement définir une norme. Soit $\langle .,. \rangle$ un produit scalaire sur E, alors l'application :

$$\begin{cases} E \longrightarrow E \\ x \longmapsto \sqrt{\langle x, x \rangle} \end{cases}$$

est une norme sur E.

Exemple Dans $\mathbb{R}^n$, de même que le produit scalaire euclien, on peut définir la norme euclienne telle que :

$$\forall x = (x_1, \ldots, x_n) \in \mathbb{R}^n \quad \|x\| = \sqrt{\sum_{i=1}^{n} (x_i)^2}$$

Propriété (**Inégalité de Cauchy-Schwarz**) . Soit $\langle .,. \rangle$ une produit scalaire sur E. On a alors l'inégalité suivante :

$$\boxed{\forall x, y \in E, \quad |\langle x, y \rangle| \leqslant \|x\| \times \|y\|}$$

"La valeur absolue du produit scalaire est inférieure au produit des normes."

2.1.3 Distance et Espace Métrique

Maintenant que nous pouvons "mesurer" des "longueurs" de vecteurs dans notre espace, on peut se demander si il est possible de "calculer" la distance entre deux éléments de E. Grâce à une telle application, on pourrait déterminer si deux éléments sont plus ou moins proches. Dès que l'on a une notion de distance, on peut ensuite l'intéresser à la notion de limite, etc...

Définition (**Distance**) . Soit E un $\mathbb{R}$-espace vectoriel. Une application $d(.,.) : E \times E \to \mathbb{R}$ est appelée **distance** ssi elle est :

- **Symétrique :** $\forall x, y \in E, \quad d(x, y) = d(y, x)$
- **Définie :** $\forall x, y \in E, \quad d(x, y) = 0_{\mathbb{R}} \Longrightarrow x = y$
- **Positive :** $\forall x, y \in E, \quad d(x, y) \geqslant 0_{\mathbb{R}}$
- **Inégalité Triangulaire :** $\forall x, yz \in E, \quad d(x, y) \leqslant d(x, z) + d(z, y)$

Proposition (Distance à partir d'une norme) Comme précédement, à partir d'une norme sur E, on peut facilement définir une distance entre deux vecteur. Soit $\|.\|$ une norme sur E, l'application :

$$\begin{cases} E \times E \longrightarrow \mathbb{R} \\ (x, y) \longmapsto \|x - y\| \end{cases}$$

Est une distance sur E.

Définition (**Espace Métrique**) . Un espace métrique est un couple (E, d) où E est un ensemble quelconque et d une distance sur cet ensemble.

2.2 Boules, Intérieur et Adhérence

Maintenant que nous savons "mesurer" des "longueurs" et déterminer à quels points deux éléments d'un espace sont "proches", nous pouvons introduire de nouveaux object à la base de tous les raisonnements que nous auront par la suite.
Dans cette section, et pour la suite de ce cours, nous nous placerons dans des espaces métriques quelconques (E, d) comme définis plus haut.

2.2.1 Boules

Introduisons maintenant le concept de boule.

Définition (**Boule ouverte, fermée**) . Soit $a \in E, r \geqslant 0$ on appelle boule ouverte l'ensemble

$$B(a,r) = \{x \in E \mid \|x-a\| < r\}$$

De même on définit la boule fermé comme l'ensemble :

$$\overline{B}(a,r) = \{x \in E \mid \|x-a\| \leqslant r\}$$

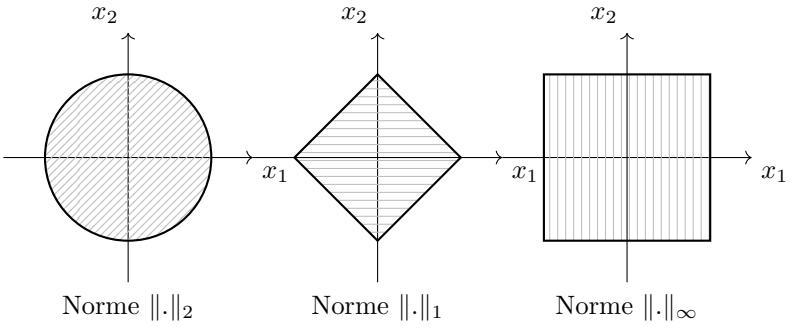
On dit alors que $B(a,r)$ est la boule ouverte de rayon r centrée en a et $\overline{B}(a,r)$ est la boule fermée de rayon r centrée en a.

Exemple Pour tout $a \in \mathbb{R}$ on a :

$$B(a,r) = \{x \in E \mid \|x-a\| < r\} =]a-r; a+r[$$

$$\overline{B}(a,r) = \{x \in E \mid \|x-a\| \leqslant r\} = [a-r; a+r]$$

Remarque En fonction de la norme (ou de la distance) choisie, une même boule peut avoir plusieurs formes.



2.2.2 Intérieur et Adhérence

Définition (**Intérieur**) . Soit $A \subset E$, soit $a \in E$, on dit que a est intérieur à A ssi

$$\boxed{\exists \varepsilon > 0, \quad B(a,\varepsilon) \subset A}$$

Autrement dit, a est intérieur à A ssi on peut construire une boule autour de a qui soit entièrement contenue dans A.
On dit alors que A est voisinnage de a. L'ensemble des points intérieurs d'une partie est

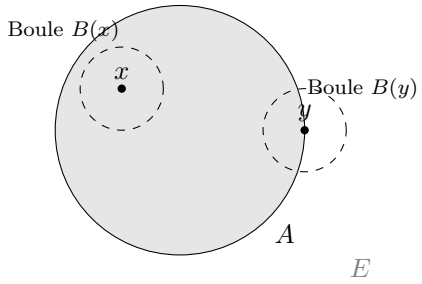
appelé l'intérieur de cette partie. On le note $int(.)$.

Définition (**Adhérence**) . Soit $A \subset E$, soit $a \in E$, on dit que a est adhérent à A ssi

$$\boxed{\forall \varepsilon > 0, \quad B(a, \varepsilon) \cap A \neq \emptyset}$$

Autrement dit, a est adhérent à A ssi pour toute boule autour de a, cette boule est intersectée avec A (i.e on ne peut pas construire de boule autour de a qui ne "déborde" pas sur A).
On dit alors que a est dans l'adhérence de A. L'adhérence d'une partie est l'ensemble de ses points adhérents. On le note $adh(.)$.

Remarque (Illustration) Soit $A \subseteq E$ et $x, y \in E$ tels que x soit intérieur à A et y soit adhérent à A. On pourrait représenter cela par le dessin ci-dessous :



L'intérieur d'une partie peut se voir comme l'ensemble des points "profonds" de cette partie. Au contraire l'adhérence peut se voir comme tous les points qui sont dedans et "très proches".

Définition (**Frontière**) . Soit $A \subseteq E$, on définit la frontière comme l'ensemble :

$$\partial A = adh(A)/int(A)$$

Exemple Soit A le disque ouvert de rayon 1 dans $\mathbb{R}^2$. Déterminons son intérieur, son adhérence et sa frontière.

$$A := \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1\}$$

On a :

- $adh(A) = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leqslant 1\}$
- $int(A) = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leqslant 1\} = A$
- $\partial A = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$

Proposition Soit $A \subset (E, d)$ $int(A)$ est le plus ouvert contenu dans A et $adh(A)$ est le plus petit fermé de E contenant A.

2.2.3 Propriétés

Proposition L'intérieur d'une boule ouverte est la boule fermée correspondante.

Propriété (**Inclusions et complémentaire**) . Soit E un espace métrique et $A, B \subseteq E$.

- Si $A \subset B$ on a alors :

$$int(A) \subset int(B) \quad adh(A) \subset adh(B)$$

- $int(A) = \overline{(adh(\overline{A}))}$ et $adh(A) = \overline{(int(\overline{A}))}$
- - $int(A \cup B) \supset int(A) \cup int(B)$
  - $int(A \cap B) = int(A) \cap int(B)$
  - $adh(A \cup B) = adh(A) \cup adh(B)$
  - $adh(A \cap B) \subset adh(A) \cap adh(B)$

Proposition (Adhérence et sup dans le cas réel) Soit $A \subset \mathbb{R}$, non vide et majorée alors, $\sup(A) \in adh(A)$.

2.3 Suites et Limites

2.3.1 Généralités

Définissons clairement la notion de suite numérique.

Définition (**Suite Numérique**) . Soit $I \subseteq \mathbb{N}$ une **partie infinie**. On appelle suite à valeurs dans (E, d), un espace métrique, d'ensemble d'indices I, toute application :

$$(u_n)_{n\in I} : n \longrightarrow u_n \in E$$

On dit alors que u_n est le *terme d'indice* $n \in I$. On note E^I l'ensemble des suites à valeurs dans E d'ensemble d'indices I.

Exemple On a par exemple :

- La suite $(u_n)_{n\in\mathbb{N}}$ définie $\forall n \in \mathbb{N}$ par $u_n = \sqrt{4n+1}$.
- La suite $(u_n)_{n\in\mathbb{N}^*}$ définie par $u_n = \ln(n + 2n^2)$.

Remarque Par abus de notation on notera souvent (u_n) pour désigner une suite. L'ensemble d'indices et les valeurs que prennent la suite dépendront du contexte.
Attention, il faut toutefois bien différentier la suite (u_n) de son terme général noté u_n qui correspond à une valeur de la suite pour un certain $n \in I$. En effet, le premier élément (u_n) appartient à E^I alors que le second appartient à E.

Proposition Soit $(u_n)_{n\in I}$ à valeur dans E. Si $E = \mathbb{R}$ on dira que la suite (u_n) est *à valeurs réelles*.

2.3.2 Propriétés

Nous allons ici nous concentrer sur les suites à *valeurs réelles*. En effet, la relation d'ordre $\leqslant$ dans $\mathbb{R}$ nous permettra de comparer différentes valeurs d'une suite et ainsi de définir le concept de monotonie.

Définition (**Monotonie**) . Soit $(u_n)_{n\in I}$ une suite à valeurs réelles d'ensemble d'indices $I \subseteq \mathbb{N}$.

1. On dit que (u_n) est *croissante* (resp. strictement croissante) si

$$\forall n \in I, \quad u_{n+1} \geqslant u_n \quad (\text{resp. } u_{n+1} > u_n)$$

2. On dit que (u_n) est *décroissante* (resp. strictement décroissante) si

$$\forall n \in I, \quad u_{n+1} \leqslant u_n \quad (\text{resp. } u_{n+1} < u_n)$$

3. On dit que (u_n) est *constante* si :

$$\exists C \in \mathbb{R}, \quad \forall n \in I, \quad u_n = C$$

4. On dit que (u_n) est *stationnaire* si :

$$\exists C \in \mathbb{R}, \quad \exists N \in \mathbb{N}, \quad \forall n \geqslant N, u_n = C$$

5. On dit que (u_n) est *périodique* si :

$$\exists p \in \mathbb{N}^*, \quad \forall n \in \mathbb{N}, \quad u_n = u_{n+p}$$

Remarque Dans le cadre de la définition suite :

1. Étudier la monotonie d'une suite revient donc à dire si elle est croissante ou décroissante.
2. On parlera de suite *monotone* lorsqu'elle sera uniquement croissante OU décroissante. Toutes les suites ne sont donc pas monotones.
3. Une suite stationnaire est une suite constante à partir d'un certain rang (noté N dans la définition).

Définition (**Suite Majorée, Minorée, Bornée**) . Soit $(u_n)_{n\in I}$ une suite à valeurs réelles d'ensemble d'indices $I \subseteq \mathbb{N}$.

1. On dit que (u_n) est *majorée* si :

$$\exists M \in \mathbb{R}, \quad \forall n \in I, \quad u_n \leqslant M$$

On dira que M est le majorant de (u_n).

2. On dit que (u_n) est *minorée* si :

$$\exists m \in \mathbb{R}, \quad \forall n \in I, \quad u_n \geqslant m$$

On dira que M est le minorant de (u_n).

3. On dit que (u_n) est *bornée* si :

$$\exists B \in \mathbb{R}, \quad \forall n \in I, \quad |u_n| \leqslant B$$

Proposition Soit $(u_n)_{n\in\mathbb{N}}$ une suite à valeurs réelles. Alors (u_n) est bornée ssi elle est majorée ET minorée.

2.3.3 Convergence et Limites de Suites Numériques

Définition (**Limite et Suite Convergente**) . Soit (u_n) une suite numérique.

1. On dit que (u_n) *converge* ou *tend* vers $l \in \mathbb{R}$ si

$$\forall \varepsilon > 0, \quad \exists N \in \mathbb{N}, \quad \forall n \in \mathbb{N}, \quad n \geqslant N \Longrightarrow |u_n - l| < \varepsilon$$

On notera alors $\lim\limits_{n\to\infty} u_n = l$ ou $u_n \underset{n\to\infty}{\longrightarrow} l$.

2. On dit que (u_n) est *convergente* si :

$$\exists l \in \mathbb{R}, \quad \forall \varepsilon > 0, \quad \exists N \in \mathbb{N}, \quad \forall n \in \mathbb{N}, \quad n \geqslant N \Longrightarrow |u_n - l| < \varepsilon$$

   On utilisera la même notation que précédement.

3. On dit que (u_n) est *divergente* si elle n'est pas convergente :

$$\forall l \in \mathbb{R}, \quad \exists \varepsilon > 0, \quad \forall N \in \mathbb{N}, \quad \exists n \in \mathbb{N}, \quad n \geqslant N \text{ et } |u_n - l| \geqslant \varepsilon$$

On peut facilement étendre cette définition aux limites infinies.

Propriété (**Unicité de la limite**) . Soit (u_n) une suite numérique. Si (u_n) converge vers une limite $l \in \mathbb{R}$ et qu'elle converge aussi vers une autre limite $l' \in \mathbb{R}$ alors $l = l'$. C'est ce que l'on appelle l'unicité de la limite.

Propriété (**Limites et Majoration/Minoration**) . Soit (u_n) une suite à valeurs réelles.

- Si (u_n) converge, alors (u_n) est bornée.
- Si $u_n \underset{n\to\infty}{\longrightarrow} +\infty$ (resp. $-\infty$) alors (u_n) n'est pas majorée (resp. minorée).

Propriété (**Convergence et suites partielles**) . Toute suite partielle d'une suite convergente est convergente et converge vers la même limite.

Définition (**Suite de Cauchy**) . Soit (E, d) un espace métrique et $(u_n)_{n\in\mathbb{N}}$ une suite à valeurs dans E. On dit que (u_n) est *de Cauchy* ou *une suite de Cauchy* lorsque ses termes se rapprochent uniformément les uns des autres lorsque n tend vers $+\infty$.

$$i.e \quad \lim_{p,q\to\infty} d(u_p, u_q) = 0$$

$$\iff \forall \varepsilon > 0, \ \exists N \in \mathbb{N}, \ \forall q \geqslant N, \ \forall q \geqslant N, \quad d(u_p, u_q) < \varepsilon$$

Remarque Attention : il ne suffit pas que la différence des termes consécutifs de la suite tendent vers zéro. Les suites de Cauchy ne sont pas à confondre avec les suites convergentes. En effet, une suite convergente est de Cauchy mais la réciproque est fausse.

Exemple Soit (u_n) une suite décroissante de rationnels positifs dont le carré tend vers 2 définie par :

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

La suite $(u_n^2)_{n\in\mathbb{N}}$ est convergente et minorée par 1. On en déduis facilement que la suite de rationnels $(u_n)_{n\in\mathbb{N}}$ est de Cauchy. Cependant elle n'a pas de limite rationnelle car une telle limite l vérifierait que $l^2 = 2$. Or $\sqrt{2} \notin \mathbb{Q}$. Donc $(u_n)_{n\in\mathbb{N}}$ ne converge pas dans $\mathbb{Q}$.

On a donc des suites de Cauchy qui ne convergent pas dans certains espaces. Il serait utile de définir des espaces dans lequels ce ne soit jamais le cas. Comme nous venons de le voir, $\mathbb{Q}$ ne suffit pas.

Définition (**Espaces Complets**) . Un espace métrique (E, d) est dit *complet* lors que toute suite de Cauchy de (E, d) converge dans (E, d).

La complétude est donc une notion pour caractériser des espaces "sans trous". Nous avons une propriété très utile dans $\mathbb{R}$:

Propriété (**Critère de Cauchy**) . Dans $\mathbb{R}$ toute suite de nombre réels converge si et seulement si elle est de Cauchy. (i.e les suites convergentes et de Cauchy sont les même dans $\mathbb{R}$).

Chapitre 3

Topologie

Contents

3.1 Ouverts et Fermés

Une fois définies les notions de boules, d'intérieur et d'adhérence, on peut maintenant "caractériser" des ensembles/parties en fonction des propriétés de leur adhérene/intérieur/frontière. Cela va nous permettre de définir les ouverts et les fermés, deux "catégories" d'ensembles essentielles pour la plupart des raisonnements analytiques de topologie.

3.1.1 Définitions et Conventions

Définition (**Ensemble Ouvert**) . Soit $A \subseteq E$, on dit que A est ouvert si $A = \int(A)$. Autrement dit si pour tout élément de A, il existe une boule autour de cet élément entièrement contenue dans A.

Définition (**Ensemble Fermé**) . Soit $A \subseteq E$, on dit que A est ouvert si $adh(A) = A$.

Remarque Quelques conventions sur les ouverts et les fermés.

- $\forall a \in E, \forall r > 0 \quad B(a,r)$ est un ouvert et $\overline{B}(a,r)$ est un fermé.
- $\emptyset$ et $\mathbb{R}^n$ dans $\mathbb{R}^n$ sont à la fois ouverts et fermés.
- $[a,b[\subset \mathbb{R}$ n'est ni ouvert, ni fermé.

Proposition Soit $A \subset \mathbb{R}^n$, A est ouvert ssi son complémentaire dans $\mathbb{R}^n$ est fermé.

Propriété (**Réunion et Intersection**) .

- Une réunion quelconque d'ouverts est ouverte.
- Une intersection finie d'ouverts est ouverte.

- Une réunion finie de fermés est fermée.
- Une intersection quelconque de fermés est fermés

Remarque (Moyen Mnémotechnique) Pour aider à la mémorisation, on peut s'aider de ces phrases :

- Les ouverts aiment s'étaler (réunion infinie), mais ils sont timides à se croiser (intersection finie).
- Les fermés aiment se serrer (intersection infinie), mais ne se dispersent pas trop (réunion finie).

Exemple (Réunion et Intersection) Quelques exemples pour retenir les propriétés :

- **Réunion infinie d'ouverts :** Soit $(A_n)_{n\in\mathbb{N}} =]-\frac{1}{n};\frac{1}{n}[$ une suite d'intervalles ouverts. Alors la réunion infinie de tout ces intervalles reste ouverte :
$$\bigcup_{n=1}^{\infty}]-\frac{1}{n};\frac{1}{n}[\quad \text{ouvert}$$
- **Intersection Infinie de fermés :** De même, soit $(B_n)_{n\in\mathbb{N}} = \left[-\frac{1}{n};\frac{1}{n}\right]$ une suite d'intervalles fermés. Alors leur intersection infinie reste fermée :
$$\bigcap_{n=1}^{\infty} \left[-\frac{1}{n};\frac{1}{n}\right] \quad \text{fermé}$$

Proposition Soit $A \subset E$, on dit que l'intérieur de A est le plus grand ouvert contenu dans A et l'adhérence de A est le plus petit fermé contenant A.

3.1.2 Ouverts/Fermés relativement

Définition (**Ouvert/Fermé relativement**) . Soient $A \subset E$ et $B \subset A$. On dit que B est **ouvert relativement** à A si il existe une ouvert V de E tel que $B = A \cap V$.
D'autre part, on dit que B est **fermé relativement** à A si il existe un fermé U de E tel que $B = A \cap U$.

3.2 Ensembles Compacts

Définition (**Recouvrement**) . Soit (E, d) un espace métrique et $A \subset E$. Soit I un ensemble quelconque et $(A_i)_{i\in I}$ une famille de sous-ensembles de E. On dit que la famille $(A_i)_{i\in I}$ est un *recouvrement de* A si :
$$A \subset \bigcup_{i\in I} A_i$$
Lorsque les A_i sont des ouverts, on parlera de recouvrement ouvert.

Définition (**Compact**) . Soit $K \subset E$. On dit que K est *compact dans* E si :

Toute suite à valeurs dans K admet une sous-suite convergente dans K.

$\iff$ De tout recouvrement ouvert de K on peut en extraire un recouvrement fini.

Proposition Un ensemble compact est *fermé et borné*.

Théorème (**Cas** $\mathbb{R}^n$) . Dans $\mathbb{R}^n$, les ensembles compacts sont exactement les fermés bornés.

Corollaire (**Théorème de Bolzano-Weierstraß**) . Dans $\mathbb{R}^n$, toute suite bornée possède une suite partielle convergente.

3.3 Ensembles Connexes

Dans cette section, nous allons détailler la notion de connexité chez les ensembles. Intuitivement, un ensemble connexe se résumera à un ensemble "en un seul morceau".

Définition (**Connexité**) . Soit $A \subset E$. On dit que A est connexe s'il est impossible de trouver $B, C \subset E$ tels que :

- $B \cap C = \emptyset$
- $E = B \cup C$
- $E \cap B \neq \{\emptyset\}$
- $E \cap C \neq \{\emptyset\}$

Chapitre 4

Fonctions Continues

Sûrement l'un des chapitres les plus important de ce cours, les notions et objets définis ici vont permettre de définir pleins de nouveaux objets et de proposer de nouveaux critères/caractérisations pour des propriétés déjà vues.
Fonctions continues
Définitions (séquentielle, par voisinages)
Propriétés : composition, opérations

Chapitre 5

Compacts

Contents

Vous voyez ce qu'est une Twingo ? Maintenant essayez d'y faire rentrer une équipe de rugby entière dedans... On pourrait dire que l'intérieur de la Twingo est compact. Voilà ce que l'on va essayer de définir dans ce chapitre, les ensembles compacts.

On nomme ici E un espace métrique muni d'une distance d.

5.1 Points d'accupulation et recouvrement

Avant de définir la notion de compact, il nous faire un effort théorique en définissant de nouveaux objects qui vont nous aider à caractériser les compacts.

Définition (**Point d'accumulation**) . Soient $A \subset E$ et $x \in E$. On dit que x est un point d'accumulation de A si toute boule de rayon non nul centrée en x contient une infinité de points de A. On remarquera qu'il suffit seulement que cette boule contienne un seul point de A différent de x.

Remarque Un point d'accumulation est un point adhérent. La réciproque est fausse en général. On remarquera que pour qu'une partie admette un point d'accumulation, elle doit contenir un nomre infini de points.

Définition (**Recouvrement**) . Soient $A \subset E$ et $(A_n)_{n\in\mathbb{N}}$ une suite de parties de E. On dit que $(A_n)_{n\in\mathbb{N}}$ constitue un recouvrement de A si $A \subset \bigcup_{n\in\mathbb{N}} A_n$.

Proposition Soit $A \subset E$, on a les trois propriétés suivantes :

- Toute suite d'éléments de A contient une suite partielle qui converge vers un élément de A.
- Tout ensemble infini d'éléments de A admet un point d'accumulation dans A.
- De tout recouvrement de A par des ensembles ouverts, on peut en extraire un recouvrement fini.

5.2 Compacts

5.2.1 Définition et caractérisations

Définition (**Ensemble Compacts**) . Soit $K \subset E$, on dit que K est compact si il satisfait au moins l'une des propriétés précedentes.

Proposition Un ensemble compact est borné et fermé.

Corollaire (**Caractérisation des compacts de** $\mathbb{R}^n$) . Dans $\mathbb{R}^n$ les compact sont exactement les fermés bornés.

5.2.2 Propriétés

Corollaire (**Théorème de Bolzano-Weierstraß**) . Dans $\mathbb{R}^n$ toute suite bornée possède une suite partielle convergente.

Corollaire (**Théorème de Bolzano-Weierstraß**) . Dans $\mathbb{R}^n$ tout ensemble infini borné admet un point d'accumulation.

Chapitre 6

Connexes

Contents

1793, Place de la Révolution, déconnexification de Louis XVI...

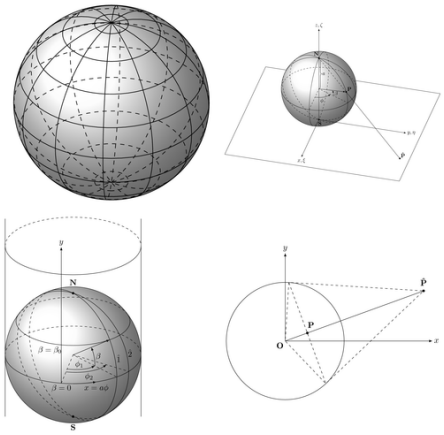
Blague à part, nous allons ici définir la notion de connexité pour un ensemble. Conceptuellement, un ensemble connexe est un ensemble "en une seule partie". Il reste à le définir proprement.

6.1 Connexité

Définition (**Connexité**) . Soit $A \subset E$, on dit que A est connexe (i.e "en une seule partie") si il est impossible de trouver deux ouverts B et C de E, disjoints, tels que leur intersection respective avec A soit non vide et que leur union soit égale à A.

6.2 Connexité et fonctions

Analyse



Chapitre 1

Dérivation sur $\mathbb{R}$

Chapitre 2

Séries Numériques

Chapitre 3

Calcul Différentiel Multivarié

Contents

3.1 Rappels de Topologie

3.2 Dérivée Partielles et Différentiabilité

Le but de ce chapitre est de généraliser les notions de dérivabilités aux fonctions de plusieurs variables. Nous nous considérerons des fonctions de la forme $f : E \longrightarrow F$ où E et F sont des espaces vectoriels normés.

Remarque (Notations) Si f est définie sur une partie de $\mathbb{R}^n$, on dit que f est une *fonction de n variables.*

- si $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ on dit que f est un *champ de scalaires.*
- pour $p \geqslant 2$, la fonction :

$$f : \begin{cases} \mathbb{R}^n \longrightarrow \mathbb{R}^p \\ (x_1, \ldots, x_n) \longmapsto (f_1(x_1), f_2(x_2), \ldots, f_n(x_n)) \end{cases}$$

  est appelée champ de vecteurs de composantes $f_1, \ldots, f_p$. Les composantes $f_i : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ sont des fonctions de n variables à valeurs dans $\mathbb{R}$.

Dans ce chapitre, nous nous placerons dans un *ouvert* $U \subset \mathbb{R}^n$.

3.2.1 Dérivées Partielles

Définition (**Application Partielle**) . Soit $f : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}^p$ et $a = (a_1, \ldots, a_n) \in \mathbb{U}$. Pour tout $i \in [\![1, n]\!]$ on pose :

$$U_i = \{t \in \mathbb{R} \mid (a_1, \ldots, a_{i-1}, t, a_{i+1,}, \ldots, a_n) \in \mathbb{U}\}$$

et on définit la i-ème application partielle de f en $a = (a_1, \ldots, a_n)$ comme le champ de vecteurs :

$$f_i : \begin{cases} U_i \longrightarrow \mathbb{R}^p \\ t \longmapsto f(a_1, \ldots, a_{i-1}, t, a_{i+1,}, \ldots, a_n) \end{cases}$$

Définition (**Dérivée Partielle**) . Dans le contexte précédent, $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ admet une i-ème dérivée partielle en a. Si f_i est dérivable en a_i et on note alors :

$$\frac{\partial f}{\partial x_i}(a) = f_i'(a_i)$$

Le calcul d'une dérivée partielle se ramène donc à un simple de calcul de dérivée (comme une fonction réelle simple). Attention à bien dériver en fonction du bon terme.

Proposition Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}^p$ un champ de vecteurs de composantes $f_1, \ldots, f_p$. f admet une dérivée partielle en $a = (a_1, \ldots, a_n)$ par rapport à x_i si chacune de ses composantes en admet une. On note alors :

$$\frac{\partial f}{\partial x_i}(a) = \left(\frac{\partial f_1}{\partial x}(a), \ldots, \frac{\partial f_p}{\partial x}(a)\right)$$

Définition (**i-ème dérivée partielle**) . Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}^p$ un champ de vecteurs de composantes $f_1, \ldots, f_p$. Si f admet en tout point a de l'ouvert U une ième dérivée partielle, on appelle *i-ème dérivée partielle* de f la fonction :

$$\frac{\partial f}{\partial x_i} : \begin{cases} U \longrightarrow \mathbb{R}^p \\ a \longmapsto \frac{\partial f}{\partial x_i}(a) \end{cases}$$

Définition (**Dérivées Partielles d'ordre supérieur**) . Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}^p$ et U un ouvert de $\mathbb{R}^n$. On définit les dérivées partielles secondes de f comme les dérivées partielles des dérivées partielles de f. On peut alors définir par récurrence dérivées partielles d'ordre p de f, notées :

$$\frac{\partial^p f}{\partial x_{i_1} \partial x_{i_2} \ldots \partial x_{i_p}}$$

comme les dérivées partielles des dérivées partielles d'ordre $p-1$ de f.

3.2.2 Classe $\mathcal{C}^1$ et Opérations

Tout comme en analyse réelle, nous pouvons définir les classes de continuité d'un champ de vecteurs.

Définition (**Classe $\mathcal{C}^1$**) . Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}^p$ et U un ouvert de $\mathbb{R}^n$. On dit que f est de classe $\mathcal{C}^1$ si elle admet des dérivées partielles par rapport à toutes ses variables sur U et si elles sont toutes continues sur U. On note alors $f \in \mathcal{C}^1(U, \mathbb{R}^p)$.

Passons maintenant à quelques propriétés sur les dérivées partielles, analogues à celles vues en analyse réelle.

Propriété (**Dérivées Partielles**) . Soient $f, g \in \mathcal{C}^1(U, \mathbb{R}^p)$ et $\lambda \in \mathcal{C}^1(U, \mathbb{R})$. Soient $a \in U$ et $i \in [\![1, n]\!]$. On a les propriétés suivantes :

- **Somme :** $f + g \in \mathcal{C}^1(U, \mathbb{R}^p)$ et :
$$\frac{\partial(f+g)}{\partial x_i}(a) = \frac{\partial f}{\partial x_i}(a) + \frac{\partial g}{\partial x_i}(a)$$
- **Multiplication :** $\lambda.fin\mathcal{C}^1(U, \mathbb{R}^p)$ et :
$$\frac{\partial \lambda.f}{\partial x_i}(a) = \lambda(a).\frac{\partial f}{\partial x_i}(a) + \lambda'(a).f(a)$$
- Si λ ne s'annule pas en a, on a : $\frac{1}{\lambda} \in \mathcal{C}^1(U, \mathbb{R})$ et :
$$\frac{\partial 1/\lambda}{\partial x_i}(a) = -\frac{1}{\lambda(a)^2}\frac{\partial \lambda}{\partial x_i}(a)$$
- On e déduit donc que si λ ne s'annule pas en a, on a $\lambda.f \in \mathcal{C}^1(U, \mathbb{R}^p)$ et :
$$\frac{\partial f/\lambda}{\partial x_i}(a) = \frac{1}{\lambda(a)^2}\left(\frac{\partial f}{\partial x_i}(a).\lambda(a) - f(a)\frac{\partial \lambda}{\partial x_i}(a)\right)$$

Proposition On peut déduire des propriétés précédentes que :

1. Toute application polynomiale est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}^n$.
2. Toute fraction rationnelle est de classe $\mathcal{C}^\infty$ sur son ensemble de définition.

3.2.3 Vecteur Gradient et Matrice Jacobienne

À partir des définitions précédentes, nous pouvons maintetant nous attaquer aux vecteurs gradients et aux matrices Jacobiennes.

Définition (**Vecteur Gradient**) . Soit $f : U \longrightarrow \mathbb{R}$ un champ de scalaires et $a \in \mathbb{U}$. Si f admet toutes ses dérivées partielles en a, on appelle *vecteur gradient* de f au point a le vecteur de $\mathbb{R}^n$:
$$\boxed{grad_a(f) = \nabla f = \left(\frac{\partial f}{\partial x_1}(a), \ldots, \frac{\partial f}{\partial x_n}(a)\right)}$$
Plus généralement, si f admet toutes ses dérivées partielles sur tout U on peut considérer

l'application gradient de f :

$$\nabla f : \begin{cases} U \subset \mathbb{R}^n \longrightarrow R^n \\ (a_1, \ldots, a_n) \longmapsto \nabla f(a) \end{cases}$$

Définition (**Matrice Jacobienne**) . Soit $f : U \longrightarrow \mathbb{R}^p, p \leqslant 2$ un champ de vecteurs admettant toutes ses dérivées partielles en $a \in \mathbb{U}$. On note $f_1, \ldots, f_p$ les composantes de f. On appelle *matrice jacobienne* de f en a, la matrice à p lignes et n colonnes définie par :

$$J_q(f) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \cdots & \frac{\partial f_1}{\partial x_n}(a) \\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) & \cdots & \frac{\partial f_2}{\partial x_n}(a) \\ \vdots & \ddots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1}(a) & \frac{\partial f_n}{\partial x_2}(a) & \cdots & \frac{\partial f_n}{\partial x_n}(a) \end{pmatrix}$$

On peut remarquer que dans le cas d'un champ de scalaires, la matrice Jacobienne se ramène à un gradient comme défini plus haut.
Si $n = p$ la matrice est donc carrée, on appelle alors le *Jacobien* le déterminant de la matrice jacobienne de f en a. On le note $|J_a(f)|$.

3.3 Applications Différentiables

3.3.1 Différentiabilité

Retournons brièvement du côté de l'analyse réelle pour bien comprendre le principe de différentielle. Pour une fonction $f : \mathbb{R} \longrightarrow \mathbb{R}$, le fait qu'elle soit dérivable en a signifie que :

$$\lim_{h \to 0} \frac{f(a+h) - f(a)}{h} = l$$

où de manière équivalente à :

$$\frac{f(a+h) - f(a) - l}{h} \underset{h \to 0}{\longrightarrow} 0$$

ce qui revient à approximer localement l'application $f(a+h) - f(a)$ par l'application linéaire $h \longmapsto lh$. Graphiquement, cela se correspond à une droite tangente de f en a.
L'objectif de cette section est de généraliser cette notion à une fonction de plusieurs variables $f : \mathbb{R}^n \longrightarrow \mathbb{R}$. Ici, au lieu d'une droite tangente, on cherchera une approximation linéaire $L : \mathbb{R}^n \longrightarrow \mathbb{R}$ qui approxime localement $f(a+h) - f(a)$. Cette application linéaire est la différentielle de f en a. Graphiquement, le graphe de L définit un hyperlan tangent au graphe de f en a.

Définition (**Différentielle**) . Soit $f : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}^p$. On dit que f est *différentiable* en $a \in \mathbb{U}$ s'il existe une application linéaire

$$L : \mathbb{R}^n \longrightarrow \mathbb{R}^p$$

un voisinnage V de $(0, \ldots, 0)$ dans $\mathbb{R}^n$ est une application $\varepsilon : V \longrightarrow \mathbb{R}$ telle que :

$$\boxed{\forall h \in V, \quad f(a+h) = f(a) + L(h) + ||h||\varepsilon(h)}$$

tel que :

$$\boxed{\varepsilon(h) \underset{h\to 0}{\longrightarrow} 0}$$

L'application linéaire L est nommée *différentielle de f en a* et notée df_a.

Définition (**Application Différentiable**) . Soit $f : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}^p$. Si f admet une différentielle pour tout $a \in U$, on dit que f est différentiable sur U et on note :

$$df : \begin{cases} U \longrightarrow \mathcal{L}(R^n, R^p) \\ a \longmapsto df_a \end{cases}$$

La différentielle est donc une application, qui à chaque valeur du domaine de f lui associe une application linéaire. Cette application linéaire, df_a est une approximation linéaire de f en a.

Exemple Prenons quelques exemples de différentielles :

- Si f est constante, alors $f(a+h) = f(a)$, sa différentielle est alors la fonction nulle, $df_a = 0$.
- Si f est une application linéaire, on a donc $f(a + h) = f(a) + f(h)$ donc $df_a = f$.
- Soit $f : \mathbb{R} \longrightarrow \mathbb{R}$ est une fonction réelle. Une application linéaire réelle est de la forme $x \longmapsto \alpha x, \alpha \in \mathbb{R}$. Donc f est différentiable s'il existe a au voisinnage de $0 \in \mathbb{R}$ tel que :

$$f(a + h) = f(a) + \alpha h + |h|\varepsilon(h) \quad \text{et} \quad \varepsilon(h) \underset{h\to 0}{\longrightarrow} 0$$

$$\iff \frac{f(a + h) - f(a)}{h} \underset{h\to 0}{\longrightarrow} 0$$

  ssi f est dérivable en a et $f'(a) = \alpha$. Dans ce cas, l'application linéaire tangente sera $df_a : h \longmapsto h.f'(a)$.

Proposition (Différentiabilité par composantes) Soit $f : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}^p, p \geqslant 2$. Soient $f_1, \ldots, f_p$ les composantes de f. On dit que f est *différentiable en $a \in U$* si $\forall i \in [\![1, p]\!], f_i$ est différentiable en a et :

$$df_a = (df_1(a), \ldots, df_p(a))$$

Remarque La différentielle permet donc de donner une *approximation locale* des petites variations de f de la forme $f(a + h) - f(a)$.

3.3.2 Implications de la différentiabilité

Étudions de plus près les applications différentiables. Soit $f : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}^p$ une application différentiable sur U. Soit L sa différentielle. Elle est continue et linéaire sur U donc par conséquent :

$$\begin{aligned} & \forall a \in U, \quad L(h) \underset{h\to 0}{\longrightarrow} 0 \\ \implies \quad & \forall a \in U, f(a + h) = f(a) + L(h) + ||h||\varepsilon(h) \underset{h\to 0}{\longrightarrow} 0 \\ \implies \quad & \forall a \in U, f(a + h) \underset{h\to 0}{\longrightarrow} a \end{aligned}$$

Donc f est continue en a pour tout $a \in U$.

Théorème (**Continuité**) . Une application différentiable en un point $a \in U$ est continue en ce point.

De nouveau, penchons nous sur les propriétés de la différentiabilité. Soit $f : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}^p$ une application différentiable sur U. Soit f_i sa ième application partielle et $(e_1, \ldots, e_n)$ la base canonique de $\mathbb{R}^n$. Étudions la dérivabilité de f_i en tant qu'application réelle :

$$\begin{aligned}\forall t \in \mathbb{R}, \forall a_i \in \mathbb{R}, \quad \frac{f_i(t) - f_i(a_i)}{t - a_i} &= \frac{f(a_1, \ldots, t, \ldots, a_n) - f(a_1, \ldots, a_n)}{t - a_i} \\ &= \frac{f(a + (t - a_i)e_i) - f(a)}{t - a_i} \\ &= \frac{df_a((t - a_i)e_i) + ||(t - a_i)e_i\varepsilon((t - a_i)e_i)}{t - a_i} \\ &= df_a(e_i) + \frac{|t - a_i}{t - a_i}||e_i||\varepsilon((t - a_i)e_i) \quad \underset{t \to a_i}{\longrightarrow} df_a(e_i)\end{aligned}$$

L'application f_i est donc dérivable en a de dérivée $df_a(e_i)$. On a donc $\forall i \in [\![1, n]\!]$, $\frac{\partial a}{\partial x_i}(a) = df_a(e_i)$. En utilisant les propriétés de linéarité, on a donc :

$$\begin{aligned}df_a(x) &= df_a\left(\sum_{i=1}^{n} x_i e_i\right) \\ &= \sum_{i=1}^{n} x_i df_a(e_i) \\ &= \sum_{i=1}^{n} x_i \frac{\partial f}{\partial x_i}(a).x_i\end{aligned}$$

Enfin en posant :

$$dx_i : \begin{cases}\mathbb{R}^n \longrightarrow \mathbb{R} \\ x \longmapsto x_i\end{cases} \qquad \text{on a} \quad df_a(x) = \left(\sum_{i=1}^{n} \frac{\partial f(a)}{\partial x_i}.dx_i\right)(x)$$

Et donc l'égalité entre :

$$df_a = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a).dx_i$$

Théorème (**Conséquences Différentiabilité**) . Soit $f : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}^p$ et U un ouvert. Si f est différentiable en $a \in U$, alors :

1. Toutes ses différentielles existent en a et :
$$\forall i \in [\![1, n]\!], \quad \frac{\partial f}{\partial x_i}(a) = df_a(e_i)$$
2. Pour tout $x \in \mathbb{R}^n$, on a :
$$df_a(x) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a).x_i$$
La différentielle de f en a est donc l'application linéaire :
$$\boxed{df_a = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a).d_{x_i} \quad \text{où} \quad d_{x_i} : \begin{cases} \mathbb{R}^n \Longrightarrow \mathbb{R} \\ (x_1, \ldots, x_n) \longmapsto x_i \end{cases}}$$
3. df_a est donc l'application linéaire représentée par la matrice jacobienne de f dans les bases canoniques de $\mathbb{R}^n$ et $\mathbb{R}^p$.

Théorème (**Différentiabilité et Classe $\mathcal{C}^1$**) . Soit $f : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}^p$ et U un ouvert. Si f est de classe $\mathcal{C}^1$ sur U alors f est différentiable sur U.

3.3.3 Méthode de l'étude de la différentiabilité

Détaillons la méthode de l'étude de la différentiabilité d'une application.

Soit $f : \begin{cases} U \subset \mathbb{R}^n \longrightarrow \mathbb{R}^p \\ x \longmapsto f(x) \end{cases}$

1. Si f est $\mathcal{C}^1$ sur U et que U est un ouvert, alors f est différentiable sur U. Il ne reste plus qu'à calculer les dérivées partielles de f
2. S'il reste des points a à étudier à la main :
   (a) Si f n'est pas continue en a, alors elle n'est pas différentiable en a.
   (b) Sinon on regarde si les dérivées partielles de f en a existent. Si ce n'est pas le cas, f n'est pas différentiable en a.
   (c) Si toutes les dérivées partielles existent, on pose :
   $$L = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) dx_i$$
   Si f est différentiable en a alors nécessairement, $df_a = L$. On en revient donc à la définition :
   $$\frac{f(a+h) - f(a) - L(h)}{||h||} \underset{h \to 0}{\longrightarrow} 0$$
   Alors f est différentiable. Sinon, non.

3.3.4 Schéma Récapitulatif

Résumons ce que nous venons de voir au travers d'un schéma :

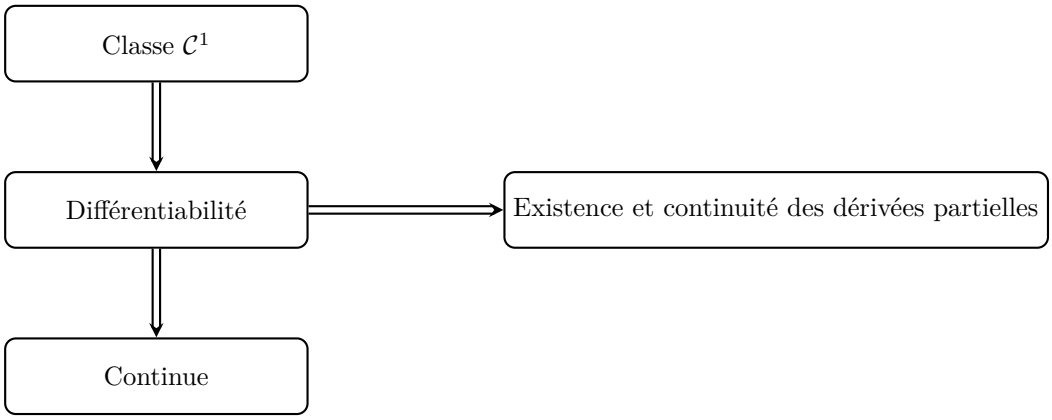


FIGURE 3.1 – Schéma des implications entre les notions de continuité, différentiabilité et classe $\mathcal{C}^1$.

3.3.5 Formule de Taylor à l'ordre 1

Essayons maintetant de voir ce que peut nous apporter la différentiaibilité dans le cas d'une fonction $f : U \subset \mathbb{R}^2 \longrightarrow \mathbb{R}$ pour laquelle le graphe est représenté en 3 dimensions. Supposons que f est de classe $\mathcal{C}^1$ sur U et soit $a = (a_1, a_2) \in U$. Le graphe de f est donc représenté par :

$$S = \{(x, y, z) \in \mathbb{R}^3 \mid f(x, y) = z\}$$

Or f est *différentiable* donc il existe V un voisinnage de $(0, 0) \in \mathbb{R}^2$ et $\varepsilon : V \longrightarrow \mathbb{R}$ tels que :

$$\forall h \in \mathbb{R}^2, \quad f(a + h) = f(a) + L(h) + ||h||\varepsilon(h)$$

avec $\varepsilon(h) \underset{h\to(0,0)}{\longrightarrow} 0$. De sorte que pour tout $(h_1, h_2) = (x - a_1, y - a_2) \in \mathbb{R}^2$, on ait :

$$L(h_1, h_2) = \langle \nabla f(a), h \rangle = \frac{\partial f}{\partial x_1}(a) \times h_1 + \frac{\partial f}{\partial x_2}(a) \times h_2$$

ainsi :

$$\begin{aligned} f(a + h) - f(a) &= \frac{\partial f}{\partial x_1}(a) \times h_1 + \frac{\partial f}{\partial x_2}(a) \times h_2 + ||h||\varepsilon(h) \\ \Longrightarrow z = f(x, y) &= f(a) + \frac{\partial f}{\partial x_1}(a)(x - a_1) + \frac{\partial f}{\partial x_2}(a)(y - a_2) + ||(x, y) - a||\varepsilon((x, y) - a) \end{aligned}$$

La différentiabilité de f au point a garantit l'existence d'un plan tangent au graphe de f.

Définition (**Plan Tangent**) . Soit $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$. On appelle plan tangent à la surface S en $a \in \mathbb{R}^2$ d'équation cartésienne $z = f(x, y)$ le plan de $\mathbb{R}^3$ d'équation cartésienne :

$$\boxed{z = f(a) + \frac{\partial f}{\partial x_1}(a)(x - a_1) + \frac{\partial f}{\partial x_2}(a)(y - a_2)}$$

3.4 Propriétés des applications différentiables

Maintenant que nous avons vu en quoi consiste la différentiabilité ainsi que ses implications, attardons nous sur les propriétés des applications différentiables.

3.4.1 Opérations Algébriques et Différentiabilité

Propriété (**Opérations**) . Soient $f, g : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}^p$ deux applications différentiables en $a \in \mathbb{U}$ et $\lambda : \mathbb{R}^n \longrightarrow \mathbb{R}$. Soit $\alpha \in \mathbb{R}$. On a les propriétés suivantes :

1. $f + g$ est diférentiable et :
$$d(f+g)_a = df_a + dg_a$$
2. λf est différentiable et :
$$d(\lambda f)_a = d(\lambda)_a f + \lambda df_a$$
3. Si λ ne s'annule pas au voisinnage de a, alors $1/\lambda$ est différentiable et :
$$d\left(\frac{1}{\lambda}\right)_a = \frac{-d\lambda_a}{\lambda^2}$$
4. si λ ne s'annule pas au voisinnage de a on en déduis donc que f/λ est différentiable et :
$$d\left(\frac{f}{\lambda}\right)_a = \frac{(df_a)\lambda - (d\lambda_a)f}{\lambda^2}$$
5. Si λ est strictement positive au voisinnage de a alors λ^α est différentiable et :
$$d(\lambda^\alpha)_a = \alpha\lambda^{\alpha-1}(d\lambda_\alpha)$$

3.4.2 Différentielle d'une composée : Règle de la chaine

3.4.3 Différentielle d'une réciproque

3.4.4 Théorème d'inversion globale

3.5 Théorème des accroissements finis

3.6 Extrema d'une fonction numérique

La fin de ce cours est consacrée à l'utilisation des outils étudiés dans le constexte de la recherche d'extrema. En effet, en dehors du cadre théorique de ce cours, il peut être utile de déterminer précisément (ou d'approximer) les extrema d'une fonction de plusieurs variables. C'est notamment le cas lors de l'application de l'algorithme de descente de gradient.
Commençons tout d'abord par définir les objets de ce cours.

Définition (**Extrema**) . Soit $f : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ et $a \in U$.

1. On dit que f *admet un minimum local en* a (resp. local strict) si il existe $\alpha > 0$ tel que :
$$\forall x \in U \cap B(a, \alpha), \quad f(x) \geqslant f(a)$$
$$(\text{resp. } \forall x \in U \cap B(a, \alpha), \quad f(x) > f(a))$$
2. On dit que f *admet un maximim local en* a (resp. local strict) si il existe $\alpha > 0$ tel que :
$$\forall x \in U \cap B(a, \alpha), \quad f(x) \leqslant f(a)$$
$$(\text{resp. } \forall x \in U \cap B(a, \alpha), \quad f(x) < f(a))$$
3. On parlera de la même façon de maximum/minimum global.

Pour désigner les maximums/minimums locaux/globaux d'une fonction, on parlera *d'extrema*.

Remarque Pour faire le lien avec le début de ce cours, on peut remarquer que f présente un extremum local en a si $f(a+h)-f(a)$ est de signe constant sur UN voisinnage de $0_{\mathbb{R}^n}$. Au contraire, f ne présentera pas d'extrema local en a si $f(a+h)-f(a)$ change de signe sur TOUT voisinnage de $0_{\mathbb{R}^n}$.

Nous allons tout d'abord chercher à déterminer des conditions nécessaires à la présence d'extrama d'un champ de vecteurs quelconques puis nous détaillerons des conditions suffisantes dans le cas général, puis dans le cas d'une fonction de deux variables.

3.6.1 Condition Nécessaire

Théorème (**Existence d'extrema**) . Soit $f : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ différentiable sur U ouvert et $a \in U$. Si f présente un extremum local en a alors :

$$df_a = 0$$

i.e

$$\forall i \in [\![1, n]\!], \quad \frac{\partial f}{\partial x_i}(a) = 0 \iff \nabla f(a) = 0$$

On appelera ces points des *singuliers*, *critiques* ou *stationnaires*.

Attention, ce n'est pas une condition suffisante !

3.6.2 Conditions suffisantes : cas général

Théorème (**Conditions Suffisantes (cas général)**) . Si $a \in U$ est un point critique de $f : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ de classe $\mathcal{C}^2$ sur U alors :

- Si toutes les valeurs propres de H_a sont strictement positives, alors f présente un minimum local strict en a.
- Si toutes les valeurs propres de H_a sont strictement négatives, alors f présente un maximum local strict en a.
- Si H_a admet deux valeurs propres non nulles de signe opposé, alors f admet un point selle en a.
- Dans le cas où 0 est une valeur propre de H_a et que toutes les autres valeurs propres sont de même signe, on ne peut conclure sans une étude approfondie.

3.6.3 Conditions Suffisantes : (cas $n = 2$)

Dans le cas d'une fonction de deux variables, la matrice hessienne de f est de taille 2. On peut alors poser :

$$\Delta = rt - s^2 \quad \text{où} \quad H_a = \begin{pmatrix} r & s \\ s & t \end{pmatrix}$$

On a alors les conditions suivantes :

Théorème (**Conditions Suffisantes (cas** $n = 2$)) . Soit f une fonction de classe $\mathcal{C}^2$ sur un ouvert U de $\mathbb{R}^2$ dans $\mathbb{R}$ et $a \in U$ un point critique de f.

1. Si $\Delta > 0$ alors f admet un extremum local strict en a :
   - Si $r > 0$ (ou $t > 0$) c'est un minimum.
   - Si $r < 0$ (ou $t < 0$) c'est un maximum.
2. Si $\Delta < 0$ alors f n'admet pas d'extrema en a, c'est un point selle.
3. Si $\Delta = 0$, il faut efectuer une étude plus fine à la main.

Chapitre 4

Théorie de la Mesure

Contents

Depuis le lycée, nous avons défini l'intégrale avec la définition de Riemann (ou Darboux pour les plus chauvins d'entre nous).

Remarque (Intégrale de Riemann) Soit $[a, b] \subseteq \mathbb{R}$ et soit f une fonction continue par morceaux sur $[a, b]$. On peut alors définir :

$$\int_a^b f = \int_a^b f(x)dx$$

Cependant cette définition se heurte à quelques problèmes dont le plus important est le fait qu'elle ne peut pas se généraliser à des fonctions non définies sur $[a, b]$, non continues par morceaux ou à plusieurs variables.
Nous allons essayer dans les deux chapitres suivants de définir des outils et de nouvelles structures pour être capable d'intégrer de telles fonctions.

4.1 Espaces Mesurables

4.1.1 Tribu Borélienne

Définition (**Tribu**) . Soit X un ensemble. Une *tribu* ou σ-algèbre sur X est une partie $\mathcal{B} \subseteq \mathcal{P}(X)$ telle que :

- $\emptyset \in \mathcal{B}$
- $\forall A \in \mathcal{B}, A^c \in \mathcal{B}$
- Stabilité par union dénombrable : $\forall (A_i)_{i\in I} A_i \in \mathcal{B}, \quad \forall i \in I, \quad \bigcup_{i\in I} A_i \in \mathcal{B}$

Un élément de $\mathcal{B}$ est appelé une "partie mesurable" de X.

Définition (**Tribu Borélienne**) . Soit E un espace topologique. La tribu borélienne est le plus petite tribu de E contenant tous les ouverts de E. On la note $\mathcal{B}_E$. Un borélien est un élément d'une tribu.

Exemple $\mathcal{B}_{\mathbb{R}}$ est la *tribu borélienne,* i.e la plus petite tribu contenant tous les ouverts de $\mathbb{R}$

Propriété (**Intersection**) . Une sigma-algèbre est stable par *intersection dénombrable.*

Le concept de tribu va nous servir plus tard lors de l'intégration de fontions pour "découper" leur domaine de définition de son ensemble de départ et permettre leur intégration.

Définition (**Espace Mesurable**) . Soit X un ensemble et $\mathcal{B}$ une tribu sur cet ensemble. On appelle *espace mesurable* le couple $(X, \mathcal{B})$.

4.1.2 Fonctions Mesurables

Dans toute la suite de ce cours, on se place dans un espace mesurable quelconque $(X, \mathcal{B})$.

Définition (**Fonction Mesurable**) . Soient $(X, \mathcal{B})$ et $(X', \mathcal{B}')$ deux espaces mesurables. On dit que $f : (X, \mathcal{B}) \longrightarrow (X', \mathcal{B}')$ est mesurable ssi :

$$\boxed{\forall A \in \mathcal{B}', \quad f^{-1}(A) \in \mathcal{B}}$$

On notera $\overline{\mathcal{M}(X)}$ l'ensemble des fonction mesurables étendues et $\overline{\mathcal{M}_+(X)}$ l'ensemble des fonction mesurables positives étendues.

Remarque On appelle *fonction borélienne* toute fonction mesurable au sens de la tribu borélienne.

Propriété (**Mesurabilité**) . Quelques propriétés (utiles) sur la mesurabilité :

- La mesurabilité est stable par composition.
- La mesurabilité est stable par recollement dénombrable i.e :

$$f : X \longrightarrow X' \quad X = \bigcup_{n \in \mathbb{N}} A_n \text{ parties deux à deux disjointes mesurables}$$

$$\text{si } \forall n \in \mathbb{N}, \quad f \restriction_A \text{ est mesurable}$$

- Toute fonction continue $f : E \longleftarrow E'$ ou E et E' sont deux espaces topologiques est mesurable.
- Toute fonction continue est mesurable (réciproque généralement fausse)
- De même, toute fonction continue par morceaux est mesurable.

Une fonction mesurable sera donc une fonction qui respecte la structure d'un espace mesurable et notamment sa tribu.

4.1.3 Fonctions mesurables réelles ou complexes

Propriété (**Lien Mesurabilité et Indicatrice**) . Soit X un ensemble et $A \subseteq X$ et soit

1_A la fonction indicatrice de A. Alors :

$$\forall A \subseteq X, \quad A \in \mathcal{B} \Longleftrightarrow 1_A \in \mathcal{M}_+(X)$$

Cette proposition établit un lien fondamental entre la mesurabilité des ensembles et celle des fonctions. Elle permet notamment d'identifier les ensembles mesurables comme ceux dont l'indicatrice est mesurable. Cela justifie l'étude des fonctions indicatrices comme briques de base pour construire des fonctions mesurables.

Théorème (**Stabilité**) . La mesurabilité est stable par opérations algébriques élémentaires $(+,.)$.

Théorème (**Convergence Simple**) . Soit (f_n) une suite de fonctions à valeurs dans $\mathcal{M}(X)$ qui converge simplement vers une fonction $f : X \longrightarrow \mathbb{C}$

$$\text{i.e } \forall x \in X, \lim_{n\to\infty} f_n(x) = f(x)$$

Alors : $f \in \mathcal{M}(X)$

Théorème (**Stabilité par inf/sup**) . Soit $(f_n)_{n\in\mathbb{N}} \in (\overline{\mathcal{M}_+(X)})^{\mathbb{N}}$ une suite de fonctions, on a :

$$\sup_{n\in\mathbb{N}} f_n : \begin{cases} X & \longrightarrow \overline{\mathbb{R}_+} \\ x & \longmapsto \sup_{n\in\mathbb{N}} f_n(x) = \sup\{f_n(x) : n \in \mathbb{N}\} \end{cases}$$

Alors : $\sup_{n\in\mathbb{N}} f_n \in \overline{\mathcal{M}_+(X)}$

Définition (**Fonction Etagée**) . Soit $(X, \mathcal{B})$ un espace mesurable. Une fonction $e : X \to \mathbb{R}$ est dite étagée si :

- $e \in \mathcal{M}_{\mathbb{R}}(X)$
- e prend un nombre fini de valeurs distinctes

Propriété (**Fonction étagée et parties mesurables**) . e est une fonction étagée ssi e est une combinaison linéaire d'indicatrices de parties mesurables.

Théorème (**Fondamental de l'intégrale de Lebesgue**) . si $f \in \overline{\mathcal{M}_+(X)}$, alors il existe une suite de fonctions croissantes étagées $(e_n)_{n\in\mathbb{N}}$ telle que :

$$\boxed{(e_n) \quad \overset{\text{CS}}{\longrightarrow} \quad f}$$

Les fonctions étagées sont donc les briques de base pour l'intégration de Lebesgue. En effet, grâce aux fonctions étagées définies sur un ensemble de boréliens, nous serons capables d'approcher très finement les fonctions à intégrer.

4.2 Mesures

4.2.1 Définitions et généralités

Soit $(X, \mathcal{B})$ un espace mesurable.

Définition (**Mesure**) . Une fonction sur $(X, \mathcal{B})$ $\mu : \mathcal{B} \longrightarrow \overline{\mathbb{R}_+}$ telle que :

1. $\mu(\emptyset) = 0$
2. $\forall (A_n)_{n\in\mathbb{N}}$ suites de parties mesurables deux à deux disjointes :

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n=0}^{\infty} \mu(A_n)$$

est appelée **mesure** sur l'espace $(X, \mathcal{B}, \mu)$, alors appelé espace mesuré.

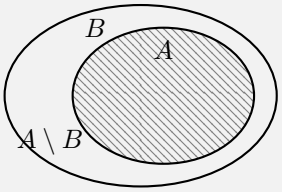
Exemple (Mesure de Comptage) Soient X un ensemble et $\mathcal{B} = \mathcal{P}(x)$. On définit alors :

$$c : \begin{cases} \mathcal{P}(X) \longrightarrow \overline{\mathbb{R}_+} \\ A \longmapsto \begin{cases} \infty \text{ si } A \text{ est infini} \\ |A| \text{ sinon} \end{cases} \end{cases}$$

est une mesure nommé **mesure de comptage**.

Propriété (**Mesures**) . Soit $(X, \mathcal{B}, \mu)$ un espace mesuré et μ sa mesure. Elle vérifie les propriétés suivantes :

- **Croissance :** La mesure est une fonction croissante i.e



$\forall A, B \in \mathcal{B}$ tels que $A \subseteq B$, on a
  - $\mu(A) \leq \mu(B)$
  - $\mu(B) = \mu(A) + \mu(B \setminus A)$

- **Convergence par union croissante :** Soit $(A_n)_{n\in\mathbb{N}}$ une suite croissante de parties mesurables

$$\text{i.e } \forall n \in \mathbb{N}, A_n \subseteq A_{n+1}, \text{ alors } \mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \lim_{n\to\infty} \mu(A_n)$$

- **Convergence par intersection décroissante :** Soit $(A_n)_{n\in\mathbb{N}}$ une suite décroissante de parties mesurables de mesure finie

$$\text{alors } \mu\left(\bigcap_{n\in\mathbb{N}} A_n\right) = \lim_{n\to\infty} \mu(A_n)$$

Définition (**Mesure Induite**) . Soit $A \subseteq C$ où $(X, \mathcal{B}, \mu)$ est un espace mesuré. On munit A d'une structure d'espace mesuré $(A, \mathcal{B}_{\restriction A}, \mu_A)$ où $\mathcal{B}_{\restriction A}$ est la tribu induite. Alors μ_A est appelée mesure induite.

Définition (**Mesure Pondéré**) . Soit $\omega \in \mathcal{M}_+(X)$ et $(X, \mathcal{B}, \mu)$ un espace mesuré.
Posons :

$$\mu_\omega = \omega\mu \quad \text{où} \quad \mu_\omega : A \longmapsto \int_A \omega \, d\mu$$

μ_ω est alors appelée mesure pondérée par ω. On a donc :

$$f \in (X, \mathcal{B}, \mu_\omega) \iff f\omega \in (X, \mathcal{B}, \mu)$$

Théorème (**Existence de la mesure de Lebesgue**) . Il existe une unique mesure sur $(\mathbb{R}, \mathcal{B}_\mathbb{R})$ notée λ telle que :

$$\boxed{\forall a, b \in \mathbb{R}, a \leq b, \quad \lambda(]a, b]) = b - a}$$

Cette mesure est appelée **mesure de Lebesgue**. En partiulier, on a $\lambda(\mathbb{R}) = \infty$

Remarque $\lambda(\mathbb{R}) = \infty$ car :

$$R = \bigcup_{n\in\mathbb{N}}]-n, n[\text{ donc } \lambda(\mathbb{R}) = \lambda\Big(\bigcup_{n\in\mathbb{N}}]-n, n[\Big) = \lim_{n\to\infty}(2n) = \infty$$

4.2.2 Parties négligeables et ensemble de Cantor

Définition (**Parties négligeables**) . Soit $A \in \mathcal{B}$ une partie mesurable, on dit que A est une **partie néligable** ssi $\lambda(A) = 0$. On dit aussi que $B \in \mathcal{B}$ est une **partie pleine** ssi $\lambda(B^c) = 0$. On remarquera que toute partie dénombrable est négligeable.

Démonstration Soit A une partie dénombrable. On a $A = \bigcup_{x\in A}\{x\}$ d'où :

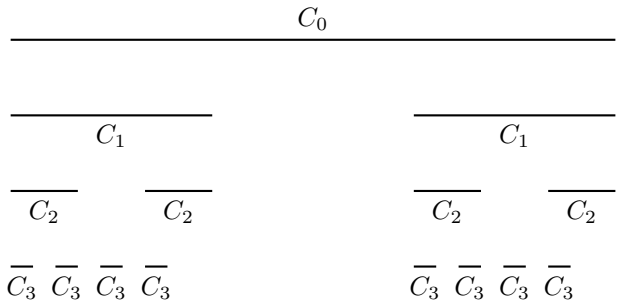
$$\lambda(A) = \lambda\left(\bigcup_{x\in A}\{x\}\right) = \lambda(x_1) + \cdots + \lambda(x_n) = 0$$

(?) Existe-t-il une partie négligeable non dénombrable ?

Exemple (Ensemble de Cantor) Soit $(C_n)_{n\in N}$ une suite de fermés non vides décroissante. Posons $C = \bigcap_{n\in N} C_n$ un fermé non vide.

$$\text{Alors : } \lambda(C) = \left(\frac{2}{3}\right)^n \underset{n\to\infty}{\longrightarrow} 0$$

C est donc négligeable mais on peut montrer que C est en bijection avec $\{0, 2\}^\mathbb{N} \simeq \mathbb{R}$ qui n'est pas dénombrable. Donc C n'est pas dénombrable.



Définition (**Presque Partout**) . Une propriété $\mathcal{P}(X)$ qui dépend d'une variable X variant

dans une espace mesuré $(X, \mathcal{B}, \mu)$ est dite vraie "presque partout", noté (p.p) si

$$\{x \in X : \neg\mathcal{P}(X)\} \text{ est négligeable}$$

Exemple Exemples de parties négligeables et propriétés vraies presque partout :

- Soient $f, g \in \overline{\mathcal{M}_+(X)}$, on a :

$$f = g \text{ presque partout} \quad \Longleftrightarrow \quad \mu(\{x \in X : f(x) = g(x)\}) = 0$$

- $\nVdash_{\mathbb{Q}} = 0$ p.p car $\mathbb{Q}$ est négligeable et $\mathbb{R}\backslash\mathbb{Q}$ est une partie pleine.

Chapitre 5

Intégration, Intégrales à paramètres et Intégration sur $\mathbb{R}^n$

Contents

5.1 Intégrale d'une fonction

5.1.1 Fonction étagée et fonction intégrale

Définition (**Intégrale d'une fonction étagée**) . Soit e une fonction étagée positive. On a $e = \sum_{i=1}^{n} \lambda_i 1_{A_i}$ où $A_i = e^{-1}(\{\lambda_i\})$ où $\lambda_i \in \mathbb{R}_+$. L'intégrale de e par rapport à μ est de la forme :

$$\boxed{\int e \, d\mu = \sum_{i=1}^{n} \lambda_i \mu(A_i)}$$

Définition (**Fonction intégrale**) . Soit $f \in \overline{\mathcal{M}_+}(X)$, alors :

$$\int f \, d\mu := \sup\Big(\Big\{ \int e \, d\mu : e \text{ fonction étagée positive tq } e \leq f \Big\}\Big)$$

Ainsi, l'intégrale d'une fonction mesurable positive existe toujours. On peut donc définir la fonction intégrale telle que :

$$\int . \, d\mu : \begin{cases} \overline{\mathcal{M}_+}(X) & \longrightarrow \overline{\mathbb{R}_+} \\ f & \longmapsto \int f \, d\mu \end{cases}$$

Propriété () . (Propriétés de l'intégrale) La fonction intégrale vérifie les propriétés suivantes :

- Extension de la mesure : $\forall A \in \mathcal{B}, \int 1_A \, d\mu = \mu(A)$
- Linéarité positive : $\forall f, g \in \overline{\mathcal{M}_+}(X), \forall \alpha \in \mathbb{R}_+$ on a :

$$\int (f + \alpha g) \, d\mu = \int f \, d\mu + \alpha \int g \, d\mu$$

- Convergence Monotone : Pour toute suite croissante de fonctions mesurables positives (étendues) (f_n), on a :

$$\boxed{\int \lim_{n\to\infty} f_n \, d\mu = \lim_{n\to\infty} \int f_n \, d\mu}$$

5.1.2 Intégrabilité (fonction positive)

Définition (**Fonction Intégrable (Cas Réel Positif)**) . Une fonction $f \in \overline{\mathcal{M}_+}(X)$ est dite intégrable si $\int f \, d\mu < \infty$.

Remarque Quelques remarques concernant la notation...

- $\overline{\mathcal{L}^1_+}(X)$ est l'ensemble des fonctions $f : X \to \overline{\mathbb{R}_+}$ intégrables.
- $\mathcal{L}^1_+(X)$ est l'ensemble des fonctions intégrables de X vers $\mathbb{R}_+$ mais non étendues.
- On remarquera que intégrable implique mesurable. La réciproques est généralement fause.

Propriété (**Croissance de l'intégrale**) . La fonction intégrale est croissante.

$$\text{i.e } \forall f, g \in \overline{\mathcal{M}_+}(X) \text{ si } f \leq g, \quad \text{alors } \int f \, d\mu \leq \int g \, d\mu$$

Démonstration Soient $f, g \in \overline{\mathcal{M}_+}(X)$ telles que $f \leq g$ alors $g = g - f + f$ où $g - f \geq 0$
D'où : $\int g \, d\mu = \int (g - f) \, d\mu + \int f \, d\mu \geq \int f \, d\mu$

Remarque Pour tout $f, g \in \overline{\mathcal{M}_+}(X)$ telles que $f \leq g$ si g est intégrable alors $f \in \overline{\mathcal{L}^1_+}(X)$.

5.1.3 Intégrabilité (fonction à valeurs complexes)

Définition (**Fonction Intégrable**) . Soit $f \in \mathcal{M}(x)$ une fonction mesurable d'un ensemble X vers $\mathbb{C}$. On dit que f est intégrable si la fonction $|f| \in \mathcal{M}_+(X)$ est intégrable (ici on parle du module).
On notera $\mathcal{L}^1(X)$ l'ensemble des fonction intégrables de X vers $\mathbb{C}$.

Définition (**Intégrale d'une fonction à valeurs complexes**) . Soit $f \in \mathcal{L}^1(X)$, distinguons deux cas :

- **Cas réel positif :** si $f \in \mathcal{L}^1_{\mathbb{R}}(X)$ alors soient :

$$\begin{aligned} f_+ : x &\longmapsto \max(0, f(x)) \\ f_+ &= \sup(0, f) \\ f_- : x &\longmapsto \min(0, f) \\ f_- &= -\inf(0, f) \end{aligned}$$

  on pose alors :

$$\int f \, d\mu = \int f_+ \, d\mu - \int f_- \, d\mu$$

- **Cas complexe :** si $f \in \mathcal{L}^1(X)$ on pose :

$$\int f \, d\mu = \int \Re f \, d\mu - i \int \Im f \, d\mu$$

Propriété (**Propriétés de l'intégrale...2**) . Quelques propriétés sur l'intégrale...

1. **Linéarité :** La fonction intégrale sur $\mathcal{L}^1(X)$ est linéaire et $\mathcal{L}^1(X)$ est un sev de $\mathbb{C}$.
2. **Croissance**
3. **Majoration par le module :** si $f \in \mathcal{L}^1(X)$ alors $\left| \int f \, d\mu \right| \leq \int |f| \, d\mu$

Théorème (**Critère d'intégrabilité des fonctions bornées**) . Si $(X, \mathcal{B}, \mu)$ est un espace mesuré **fini** alors toutes fonction $f \in \mathcal{M}(X)$ **bornée** est intégrable.

5.1.4 Extension par zéro et intégrale restreinte

Définition (**Fonction extension par zéro**) . Soit $f : A \to \mathbb{C}$ où $A \subset X$, alors l'extension par zéro de f est la fonction :

$$e_X(f) : \begin{cases} X \longrightarrow \mathbb{C} \\ x \longmapsto \begin{cases} f(x) \text{ si } x \in A \\ 0 \text{ sinon} \end{cases} \end{cases}$$

Définition (**Intégrale restreinte**) . Une fonction f intégrable sur X positive (resp. à valeurs complexes) est dite intégrable sur $A \subset X$ si la restriction de f (resp. la restriction $|f|$) à A est intégrable.

Propriété (**Intégrale restreinte**) . Soient $g \in \overline{\mathcal{M}_+(A)}, A \subset X$ et e_x son extension par zéro sur X. Soit 1_A la fonction indicatrice de A.

- **Intégrale et extension par zéro :**

$$\int_X e_x(g)\, d\mu = \int_A g\, d\mu$$

  g est donc intégrable sur A si $e_x(g)$ est intégrable sur X.

- **Intégrale restreinte et indicatrice :**
  - si $f \in \mathcal{M}_+(X)$ alors $\int_A f\, d\mu = \int_X f \times 1_A\, d\mu$
  - si $f \in \mathcal{M}(X)$ alors f est intégrable sur A si $f \times 1_A$ est intégrable sur X et dans ce cas :

$$\boxed{\int_A f\, d\mu = \int_X f\ \times 1_A\, d\mu}$$

- **Intégrabilité et restriction :** soit $f \in \mathcal{L}^1(X)$ alors $f \restriction_A \in \mathcal{L}^1(X)$ où $A \subset X$ et

$$\int_A |f|\, d\mu \leq \in_X |f|\, d\mu$$

- **Relation de Chasles généralisée :** Soit $(A_n)_{n\in\mathbb{N}}$ une suite de parties mesurables deux à deux disjointes et soit $f \in \overline{\mathcal{M}_+(X)}$ alors :

$$\int_{\bigcup_{n\in\mathbb{N}} A_n} f\, d\mu = \sum_{n=0}^{\infty} \int_{A_n} f\, d\mu$$

  Pour le cas complexe, si $f \in \overline{\mathcal{M}(X)}$ alors f est intégrable sur $(A_n)_{n\in\mathbb{N}}$ ssi

$$\sum_{n=0}^{\infty} \int_{A_n} |f|\, d\mu \leq \infty$$

  et dans ce cas on a l'égalité (dans $\mathbb{C}$) :

$$\int_{\bigcup_{n\in\mathbb{N}} A_n} f\, d\mu = \sum_{n=0}^{\infty} \int_{A_n} f\, d\mu$$

- **Indivisibilité des parties négligeables :** Si N est une partie négligeable, toute fonction mesurable est intégrable sur N et son intégrale est nulle.

$$\text{i.e } \int_N f\, d\mu = 0$$

5.1.5 Intervalles non compacts et rappels

Théorème (**Intégration sur un intervalle non compact**) .

1. **Cas positif :** Soit $f \in \overline{\mathcal{M}_+([a,c[)}$ on a

$$\boxed{\int_{[a,c[} f\, d\lambda = \sup_{b\in[a,c[} \int_{[a,c[} f\, d\lambda}$$

en particulier $f \in \mathcal{L}^1_+([a,c[)$:

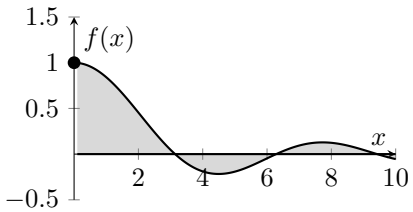
ssi $\left\{\int_{[a,b]} f\, d\lambda : b \in [a,c[\right\}$ est majoré dans $\mathbb{R}_+$

ssi il existe une suite (b_n) croissante de $[a,c[$ qui tend vers c telle que $(\int_{[a,bn]} f\, d\lambda)_n$ est majorée.

2. **Cas non positif :** Soit $f \in \mathcal{L}^1([a,c[)$ à valeurs complexes alors :

$$\boxed{\lim_{b\to c} \int_{[a,b]} d\,\lambda = \int_{[a,c[} f\, d\lambda}$$

Exemple :



$f : [a,c[\longrightarrow \mathbb{C}$ peut avoir une intégrale impropre convergente sans nécessairement être intégrable au sens de Lebesque. Par exemple la fonction ci-contre n'a pas d'intégrale définie au sens de Lebesgues mais au sens de Riemann oui.

$$f(x) = \frac{\sin(x)}{x} \text{ lorsque } x > 0, \quad f(0) = 1.$$

Exemple (Intégrale de Riemann)

$$f_\alpha : \begin{cases} \mathbb{R}_+^* \longrightarrow \mathbb{R} \\ t \longmapsto \frac{1}{t^\alpha}, \alpha \geq 0 \end{cases}$$

1. sur $[1,\infty[$, f_α est intégrable ssi $\alpha \geq 1$
2. sur $]0,1[$, f_α est intégrable ssi $\alpha \leq 1$

Exemple (Intégrales de Bertrand)

$$f_{\alpha,\beta} : \begin{cases} \mathbb{R}_+^* \setminus \{1\} \longrightarrow \mathbb{R} \\ t \longmapsto \frac{1}{t^\alpha |\ln t|^\beta} \end{cases}$$

1. sur $[0,\frac{1}{2}]$, $f_{\alpha,\beta}$ est intégrable ssi $\alpha \leq 1$ ou ($\alpha = 1$ et $\beta < 1$)
2. sur $[2,\infty[$, $f_{\alpha,\beta}$ est intégrable ssi $\alpha \geq 1$ ou ($\alpha = 1$ et $\beta > 1$)

5.1.6 Voisinnage

Définition (**Intégrabilité et voisinnage**) . Soit $f : I \longrightarrow \mathbb{C}$, avec I un intervalle et $a \in \overline{I}$. On dit que f est intégrable au voisinnage de a s'il existe une voisinnage V de a tel que :

$$f \restriction_{I \cap V} \text{ est intégrable}$$

Exemple Soit $f : [0, \infty[\longrightarrow \mathbb{C}$. Dire que f est intégrable au voisinnage de ∞ signifie que $\exists c \in \mathbb{R}_+, f \restriction_{[} c, \infty[$ est intégrable.

Définition (**Fonction localement intégrable**) . Soit $f : I \longrightarrow \mathbb{C}$, on dit que f est localement intégrable sur I si pour tout compact $K \subset I$, f est intégrable sur K.

Remarque (Critère d'intégrabilité) Si $f : [a, b] \longrightarrow \mathbb{C}$ est

- localement intégrable sur $[a, b]$
- intégrable au voisinnage de b

alors f est intégrable sur $[a, b]$.

Théorème (**Changement de variable**) . Soit $\Psi : I \to I'$ de classe $\mathcal{C}^1$ et bijective. Soit f une fonction positive, étendue, intégrable, alors :

$$\boxed{\int_{I'} f \, d\lambda = \int_I (f \circ \Psi) \times |\Psi'| \, d\lambda}$$

$$\begin{array}{ccc} X & \xrightarrow{\Psi} & I' \\ & \searrow^{f \circ \Psi} & \downarrow f \\ & & \mathbb{R}_+ \end{array}$$

Donc $f \in \mathcal{M}(X)$ est intégrable sur I' ssi $(f \circ \Psi)|\Psi'|$ est intégrable sur I.

5.2 Théorèmes fondamentaux de l'intégrale

Théorème (**Critère d'annulation presque partout**) . Soit $f \in \overline{\mathcal{M}_+(X)}$, on a :

$$\int d\,\mu = 0 \Longleftrightarrow f = 0 \text{ p.p}$$

Théorème (**Critère de finitude presque partout**) . Soit $f \in \overline{\mathcal{L}^1_+(X)}$, alors $f \leq \infty$ p.p

Théorème (**Inversion série intégrale(Cas positif)**) . Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonction dans $\overline{\mathcal{M}_+(X)}$

$$\text{alors } \int \sum_{n\geq 0} f_n \, d\mu = \sum_{n\geq 0} \int f_n \, d\mu$$

Attention : Cette propriété n'est pas à confondre avec la linéarité car ici, nous sommes en présence d'une limite de somme de fonctions.

Théorème (**Convergence Dominée**) . Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions dans $\overline{\mathcal{M}(X)}$. Supposons que $(f_n)_{n\in\mathbb{N}}$ converge simplement vers $f \in \mathcal{M}(X)$

si $\quad \exists g \in \mathcal{L}^1_+(X)$ telle que $\forall n \in \mathbb{N}, |f_n| \leq g$
alors $\quad \forall n \in \mathbb{N}, f_n$ est intégrable et f est intégrable sur$X$

$$\text{et } \int \Big| f_n - f \Big| \, d\mu \underset{n\to\infty}{\longrightarrow} 0 \quad \Longleftrightarrow \quad \int f_n \, d\mu \underset{n\to\infty}{\longrightarrow} \int f \, d\mu$$

$$\text{i.e } \lim_{n\to\infty} \int f_n \, d\mu = \int \lim_{n\to\infty} f_n \, d\mu$$

... dans $\mathbb{C}$

Théorème (**Inversion série-intégrale (Cas complexe)**) . Soit $\sum f_n$ une série de fonctions telle que $\forall n \in \mathbb{N}, f_n \in \mathcal{L}^1(X)$ converge absoluments pour tout $x \in X$ $\Big($ie $\quad \forall x \in X, \sum_{n\in\mathbb{N}} \Big| f_n(x) \Big| \leq \infty\Big)$

si $\quad \displaystyle\sum_{n\in\mathbb{N}} \int \Big| f_n \Big| \, d\mu \leq \infty$
alors $\quad \displaystyle\int \sum_{n\in\mathbb{N}} f_n \, d\mu = \sum_{n\in\mathbb{N}} \int f_n \, d\mu$

et $\sum_{n\in\mathbb{N}} f_n$ est intégrable sur X (en tant que fonction à valeurs complexes).

Démonstration Posons $g_n = \sum_{k=0}^n f_k$. Comme $\sum f_n$ converge, (g_n) converge simplement vers $\sum_{n=0}^\infty f_n = f$

Soit $n \in \mathbb{N}$, on sait que $|g_n| \leq \Big| \sum_{k=0}^n f_k \Big| \leq \sum_{k=0}^n |f_k| \leq \sum_{n=0}^\infty f_n := g$ Donc g est intégrable.
On peut appliquer le théorème de convergence dominée.
Alors $\sum_{n=0}^\infty f_n$ est intégrable et :

$$\lim_{n\to\infty} \int g_n \, d\mu = \int \sum_{n=0}^\infty f_n \, d\mu$$

$$\lim_{n\to\infty} \int \sum_{k=0}^n f_k \, d\mu = \int \sum_{n=0}^\infty f_n \, d\mu$$

$$\lim_{n\to\infty} \sum_{k=0}^n \int f_k \, d\mu = \int \sum_{n=0}^\infty f_n \, d\mu$$

$$\sum_{n=0}^\infty \int f_n \, d\mu = \int \sum_{n=0}^\infty f_n \, d\mu$$

5.3 Intégrales à paramètres

On cherche à savoir si les fonctions de la forme $t \longmapsto \int f(x,t)\, d\mu$ sont continues, dérivables par rapport à t ? Et si un théorème nous permet de dire que la dérivée de l'intégrale est égale à l'intégrale de la dérivée.

5.3.1 Définitions, Limites et continuité

Définition (**Intégrale à paramètre**) . Une intégrale à paramètre est une fonction

$$\begin{cases} I \longrightarrow \mathbb{C} \\ t \longmapsto \int f(\cdot,t)\, d\mu \end{cases} \qquad \text{où } f : X \times I \longrightarrow \mathbb{C}$$

telle que $\forall t \in I,\ f(\cdot,t) \in \mathcal{L}^1(X)$

Théorème (**Limite sous l'intégrale**) . Soient $F : I \to \mathbb{C}, t \longmapsto \int f(\cdot,t)\, d\mu$ une intégrale à paramètre et $a \in \overline{I}$.

si
1. $\forall x \in X, \lim_{t\to a} f(x,t)$ existe dans $\mathbb{C}$
2. Hypothèse de domination

$$\exists g \in \mathcal{L}^1_+(X) \text{ tq } \forall t \in I, |f(x,t)| \le g(x) \quad (\text{i.e } |f(\cdot,t) \le g)$$

alors F admet une limite dans $\mathbb{C}$ quand $t \to a$ et :

$$\boxed{\lim_{t\to a} F(t) = \int \lim_{t\to a} f(\cdot,t)\, d\mu}$$

Exemple

$$F : \begin{cases} \mathbb{R}_+^* \longrightarrow \mathbb{C} \\ t \longmapsto \int_{\mathbb{R}} e^{-tx} \frac{1}{1+x^2}\, dx \end{cases} \qquad \text{où } f(x,t) = e^{-tx}\frac{1}{1+x^2}$$

On a $\forall x \in \mathbb{R}_+^* \setminus \{1\}, \lim_{t\to\infty} f(x,t) = 0$ et $\forall t \in I, f(x,t) \le \frac{1}{1+x^2}$

$$\text{d'où } F \underset{t\to\infty}{\longrightarrow} \int_{\mathbb{R}_+^*} \lim_{t\to\infty} f(x,t)\, dx = 0$$

Corollaire (**Continuité sous l'intégrale**) . Soit F une intégrale à paramètre.

si
1. $\forall x \in X, f(x,\cdot)$ est continue sur I
2. Hypothèse de domination

alors F est $\mathcal{C}^0$ sur I

5.3.2 Dérivation

Théorème (**Dérivation sous l'intégrale**) . Soit F une intégrale à paramètre.

si 1. $\forall x \in X, f(x,\cdot) \in \mathcal{C}^1(I)$

2. Hypothèse de domination

$$\exists g \in \mathcal{L}^1_+(X), \forall t \in I, \left|\frac{\partial}{\partial t} f(\cdot,t)\right| \leq g$$

alors F est $\mathcal{C}^1(I)$ et on peut dériver F sous l'intégrale

$$\text{i.e } \forall t \in I, \quad F' = \int_X \frac{\partial}{\partial t} f(\cdot,t)\, d\mu = \frac{\partial}{\partial t} \int_X f(\cdot,t)\, d\mu$$

(Attention) : Les notions de continuité et de dérivabilité sont des définitions LOCALES

$$\text{i.e } f \in \mathcal{C}^0(I) \Longleftrightarrow \forall a \in I, f \in \mathcal{C}^0(\{a\})$$

5.4 Intégration sur R^n

5.4.1 Tribu Produit

Soit $(X_1, \mathcal{B}_1)$ et $(X_1, \mathcal{B}_2)$ deux espaces mesurables.

Définition (**Tribu Produit**) . La tribu produit sur $X_1 \times X_2$ est la plus petite tribu sur $X_1 \times X_2$ qui contient toutes les parties du type $A_1 \times A_2$ où $A_1 \in \mathcal{B}_1$ et $A_2 \in \mathcal{B}_2$.
On la note $\mathcal{B} := \mathcal{B}_1 \otimes \mathcal{B}_2$.

Propriété () . Soit $f : X \longleftrightarrow X_1 \times x_2, x \longmapsto (f_1(x), f_2(x))$ où X est un espace mesurable. Alors f est mesurable sur X ssi f_1 et f_2 sont mesurables sur X.

Théorème (**Produit de tribu**) . Soit $n \in \mathbb{N}$, on a :

$$\mathcal{B}_{\mathbb{R}^2} = \mathcal{B}_{\mathbb{R}} \otimes \mathcal{B}_{\mathbb{R}} \quad \text{et} \quad \mathcal{B}_{\mathbb{R}^n} = \bigotimes_{i=1}^{n} \mathcal{B}$$

Corollaire (**Produit d'un borélien**) . Le produit d'un borélien est un borélien.

5.4.2 Mesure Produit

Soit $(X_1, \mathcal{B}_1, \mu_1)$ et $(X_1, \mathcal{B}_2, \mu_2)$ deux espaces mesurés.

Définition (**Mesure Produit**) . Il existe une mesure μ sur $(X_1 \times X_2, \mathcal{B}_1 \times \mathcal{B}_2)$ telle que $\forall A_1 \in \mathcal{B}_1, \forall A_2 \in \mathcal{B}_2$, on ait :

$$\mu(A_1 \times A_2) = \mu_1(A_1) \times \mu_2(A_2)$$

elle est appelée mesure produit sur $X_1 \times X_2$ et notée $\mu_1 \otimes \mu_2$.

Définition (**Espace mesuré produit**) . Soit $B \in \mathcal{B}_1 \times \mathcal{B}_2$, on a :

$$\mu(B) = \inf \sum_{n=0}^{\infty} \mu(B_n)$$

où $(B_n)_{n\in\mathbb{N}}$ est la suite de l'ensemble des recouvrements dénombrables de B par des produits cartésiens.
$(X_1 \times X_2, \mathcal{B}_1 \otimes \mathcal{B}_2, \mu_1 \otimes \mu_2)$ est donc appelée espace mesuré produit.

5.4.3 Procédé d'intégration

Théorème (**Fubini-Tonelli**) .

1. Tonelli : Soit $F \in \overline{\mathcal{M}_+}(X)$ telle que :

$$f : \begin{cases} X_1 \times X_2 \longrightarrow \overline{\mathbb{R}_+} \\ (x_1, x_2) \longmapsto f(x_1, x_2) \end{cases}$$

Alors :

$$\begin{aligned} \int_{X_1\times X_2} f \, d\mu &= \int_{X_1} \Big(\int_{X_2} f(x_1, .) \, d\mu_2 \Big) \, d\mu_1(x_1) \\ &= \int_{X_2} \Big(\int_{X_1} f(., x_2) \, d\mu_1 \Big) \, d\mu_2(x_2) \end{aligned}$$

2. Fubini : Soit $f \in \mathcal{L}^1(X_1 \times X_2)$. Alors la formule précédente est vraie

$$\begin{aligned} \int_{X_1\times X_2} f \, d\mu &= \int_{X_1} \Big(\int_{X_2} f(x_1, .) \, d\mu_2 \Big) \, d\mu_1(x_1) \\ &= \int_{X_2} \Big(\int_{X_1} f(., x_2) \, d\mu_1 \Big) \, d\mu_2(x_2) \end{aligned}$$

ici $f(., x_2)$ est intégrable sur X_1 pour presque tout $x_2 \in X_2$ donc $\int f \, d\mu_1$ est bien définie p.p sur X_2

Définition (**Matrice Jacobienne**) . Soit $\phi : U \to V$ où U est un ouvert de R^k et V un ouvert de $\mathbb{R}^n$. La matrice Jacobiennne de ϕ est la fonction :

$$D_\phi : \begin{cases} U \longrightarrow \mathcal{M}_{n,k}(\mathbb{R}) \\ x \longmapsto \left(\frac{\partial \phi^i}{\partial x_j}\right)_{1\le i,j\le k,n} \end{cases}$$

Le jacobien de ϕ est la fonction :

$$J_\phi : \begin{cases} U \longrightarrow \mathbb{R}_+ \\ x \longmapsto \sqrt{\det({}^t D_\phi \times D_\phi)} \end{cases}$$

car ${}^t D_\phi \times D_\phi \in \mathcal{M}_{k,k}(\mathbb{R})$.

- si $k = n$: $\boxed{J_\phi(x) = |\det D_\phi|}$

- si $k = 1$: $J_\phi(x) = \sqrt{\sum_{i=1}^n \phi^i(x)} = \|\phi'(x)\|$

Définition (C^p**-difféomorphisme**) . Supposons que $k = n$, ϕ est un C^p-difféomorphisme si

$$\phi : U \longrightarrow V$$

où U et V sont deux ouverts de $\mathbb{R}^n$ et, de plus :

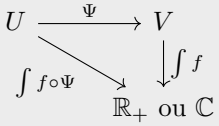
1. $\phi \in C^p(U)$
2. ϕ est bijective de réciproque ϕ^{-1}
3. $\phi^{-1} \in C^p(U)$

Théorème (**Inversion Globale**) . Soit $\Psi : U \subset \mathbb{R}^n \longrightarrow V \subset \mathbb{R}^n$ tel que $\Psi \in C^p$ et Ψ injective.

si $\forall x \in U, D_{\Psi_x}$ est inversible

alors $\Psi : U \longrightarrow \Psi(V)$ est un C^p-difféomorphisme et $\Psi(U)$ est un ouvert de $\mathbb{R}^n$

Théorème (**Changement de variable**) . Soient $U, V \subset \mathbb{R}^n$ deux ouverts et $\Psi : U \longrightarrow V$ un C^p-difféomorphisme.



où $f : V \longrightarrow \overline{R_+}$ où $\mathbb{C}$

Autrement dit, Ψ donne les anciennes variables en fonction des nouvelles. Le raisonnement est similaires aux matrices de passage.

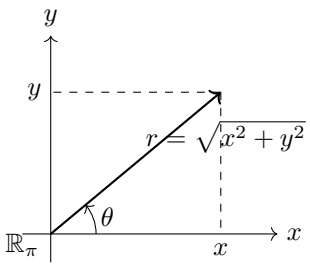
si $f \in \overline{\mathcal{M}_+}(V)$ alors :

$$\boxed{\int_V f \, d\lambda_n = \int_U f \circ \Psi \times J_\phi \, d\lambda_n}$$

si $f \in \mathcal{M}(V)$ on a : $f \in \mathcal{L}^1(U) \iff (f \circ \Psi) \times J_\phi \in \mathcal{L}^1(U)$ et dans ce cas :

$$\int_V f \, d\lambda_n = \int_U (f \circ \Psi) \times J_\phi \, d\lambda_n$$

Exemple (Changement de variable polaire) Soient $(r, \theta) \in \mathbb{R}_+^* \times]0, 2\pi[$ tels que $\Psi(r, \theta) = (r\cos\theta, r\sin\theta)$



$$\Psi : \begin{cases} \mathbb{R}_+^* \times]0, 2\pi[\to \mathbb{R}^2 \backslash \mathbb{R}_\pi \\ (r, \theta) \longmapsto (x, y) = (r\cos\theta, r\sin\theta) \end{cases}$$

Alors :

$$J_{\Psi(r,\theta)} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} \text{ et } |J_{\Psi(r,\theta)}| = r\cos^2\theta + r\sin^2\theta = r \in \mathbb{R}_+^*$$

Corollaire (**Passage en polaire**) . Soit Ψ le passage en coordonnées polaires. Alors Φ est un C^∞-difféomorphisme et pour tout $f \in \overleftarrow{\mathcal{M}_+}(\mathbb{R}^2)$, on a :

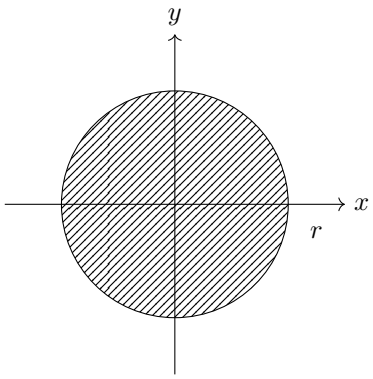
$$\int_U f(x,y)\ d\lambda_2 = \int_V f(r\cos\theta, r\sin\theta) \times r\ drd\theta$$

où $V = \mathbb{R}_+^* \times]0, 2\pi[$

Exemple (Aire d'un disque) Aire d'un disque avec le corollaire fraîchement énoncé :

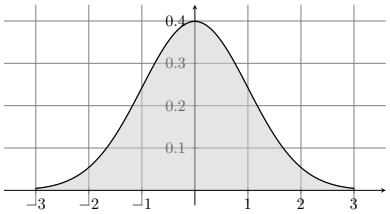
Soit A l'aire du disque de rayon R centré en $(0,0)$. Alors

$$\begin{aligned} A &= \lambda_2(Disque) \\ &= \int_{\mathbb{R}^2} 1_{Disque}(x,y)\ dxdy \end{aligned}$$



$$\begin{aligned} &= \int_{\mathbb{R}_+^* \times]0,2\pi[} 1_{Disque}(r\cos\theta, r\sin\theta) \times r\ drd\theta \\ &= \int_0^\infty \left(\int_0^{2\pi} 1_{[0,\mathbb{R}[}(r) \times r\ d\theta \right) dr \\ &= \left(\int_0^\infty 1_{[0,\mathbb{R}[}(r) \times r\ dr \right) \times 2\pi \\ &= 2\pi \int_0^R r\ dr = 2\pi \frac{R^2}{2} = \pi R^2 \end{aligned}$$

Exemple (Gaussienne) Soit γ la gaussienne sur $\mathbb{R}$ telle que $\gamma = \int_{\mathbb{R}} e^{-x^2}\ dx = \sqrt{\pi}$. Vérifions ce résultat.



$$\begin{aligned} \gamma &= \left(\int_{\mathbb{R}} e^{-x^2}\ dx \right) \left(\int_{\mathbb{R}} e^{-y^2}\ dy \right) \\ &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} e^{-y^2}\ dy \right) e^{-x^2}\ dx \\ &= \int_{\mathbb{R}} \int_R e^{-x^2-y^2}\ dxdy \end{aligned}$$

D'après le théorème de Tonelli :

$$\int_{\mathbb{R}} \int_R e^{-x^2-y^2}\ dxdy = \int_{R^2 \times \mathbb{R}^2} e^{-x^2-y^2}\ dxdy$$

D'après le changement de variable polaire :

$$\forall (x,y) \in \mathbb{R}^2 \backslash R_\pi, \quad \exists (x,\theta) \in \mathbb{R}_+^* \times]0, 2\pi[\text{ tq} \quad (x,y) = (r\cos\theta, r\sin\theta)$$

D'où :

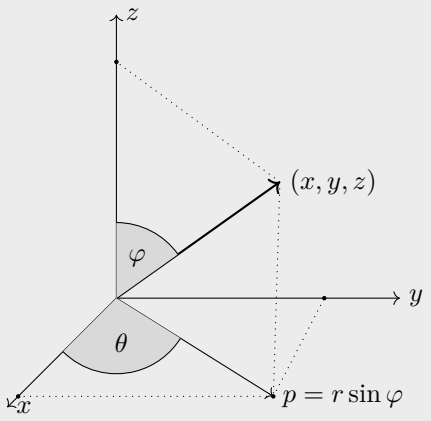
$$\begin{aligned}\ldots &= \int_0^{2\pi}\int_0^{\infty} re^{-r^2\cos^2\theta - r^2\sin^2\theta}\,drd\theta \\ &= \int_0^{2\pi}\int_0^{\infty} re^{-r^2}\,drd\theta = \int_0^{2\pi}\Big[-\frac{e^{-r^2}}{2}\Big]_0^{\infty}\,d\theta \\ &= \int_0^{2\pi}\frac{1}{2}\,d\theta = \Big[\frac{1}{2}\theta\Big]_0^{2\pi} = \pi\end{aligned}$$

donc $\gamma^2 = \pi \iff \gamma = \pi$

Théorème (**Changement de variable spérique**) .

$$\Psi : \begin{cases}\mathbb{R}_+\times]-\pi,\pi[\times]0,\pi[\longrightarrow \mathbb{R}^3\backslash P \\ (r,\theta,\varphi) \longmapsto \begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}r\sin\varphi\cos\theta\\ r\sin\varphi\sin\theta\\ r\cos\theta\end{pmatrix}\end{cases}$$

$$\begin{aligned}x &= p\cos\theta = r\sin\varphi\cos\theta\\ y &= p\sin\theta = r\sin\varphi\sin\theta\\ z &= r\cos\theta\end{aligned}$$



Ψ est un C^∞-difféomorphisme de Jacobien :

$$J_\Psi \begin{vmatrix}\sin\varphi\cos\theta & -r\sin\varphi\sin\Theta & r\sin\theta\cos\theta\\ \sin\theta & r\sin\varphi\cos\theta & r\sin\varphi\cos\theta\\ \cos\theta & 0 & -r\sin\varphi\end{vmatrix}\cdots = r^2\sin\theta > 0$$

Chapitre 6

Introduction à l'Analyse Fonctionnelle

Contents

Dans ce chapitre, nous allons essayer d'étudier les propriétés de fonctions en les assimilant à leur classe d'équivalence pour l'égalité presque partout. Nous allons d'abord commencer par montrer que c'est possible. Ensuite nous verrons quelques propriétés des espaces de ces "classes d'équivalences" et nous finirons par étudier le produit de convolution dans ces espaces.

6.1 Espaces L^p

6.1.1 Intégration et égalité presque partout

Soit $(X, \mathcal{B}, \mu)$ un espace mesuré et $\mathcal{M}(X) = \{f : X \to \mathbb{C}\}$ l'ensemble des fonctions mesurables de X. On sait que $f \sim_{\text{p.p}} g \iff f = g$ p.p $\iff \mu(\{x \in X,\ f(x) \neq g(x)\}) = 0$.
Donc l'égalité presque partout est une relation d'équivalence telle que :

$$M(X) := \mathcal{M}(X)/\sim_{\text{p.p}}$$

On différenciera bien les M dits "droits" et les $\mathcal{M}$ dit "courbes". Les premiers représentent classes d'équivalences des fonction de $\mathcal{M}$ pour la relation d'égalité presque partout.

Remarque Soit $f \in M(X)$ et x$\in X$, $f(x)$ n'a aucun sens cet f n'est pas une fonction mais une classe d'équivalence de fonctions.

Exemple (Fonctions de même classe) Plaçons nous dans $(\mathbb{R}, \mathcal{B}_{\mathbb{R}}, \lambda)$ avec $f = 0$ et $g = 1_{\{0\}}$, on remarquera que $f = g$ p.p. Donc f et g sont de même classe.

Définition (**Intégrale d'une classe**) . Soit $f \in \overline{\mathcal{M}_+}(X)$ (ici une classe), l'intégrale de f est définie par :

$$\int f \, d\mu := \int f_1 \, d\mu$$

Où f_1 est une fonction de la classe f.

Remarque On s'appuie sur la propriété suivante :

Soit $M(X)$ une classe d'équivalence de $\mathcal{M}(X)$, alors $\forall f_1, f_2 \in M(X)$, on a :

$$\int_X f_1 \, d\mu = \int_X f_2 \, d\mu$$

i.e l'intégrale de la classe ne dépend pas du représentant de la classe.

Définition (**Espace** L^p) . On définit de même les ensembles L^p, les classes d'équivalences des fonctions intégrables pour la relation d'égalité presque partout :

- si $p = 1$, on a : $L^1(X) = \{f \in M(X),\ \int_x |f| \, d\mu < \infty\}$
- si $p < \infty$, on a : $L^p(X) = \{f \in M(X),\ \int_x |f|^p \, d\mu < \infty\}$
- si $p = \infty$, on a : $L^\infty(X) = \{f \in M(X),\ \exists C \in R_+^*,\ |f| \leq C\}$

où f est la classe de f_1 pour tout fonction f_1 de la classe f.

Théorème (**Extension**) . Tous les théorèmes et définitions du cours d'intégration sont applicables aux classes de fonctions.

L^∞ est l'ensemble des classes de fonctions "grosso-modo" bornées.

6.1.2 Espace L^p

Définition (**Normes** L^p) . Soient $f \in M(X)$ et $p \in [1; \infty[$ on définit les normes suivantes :

$$\|f\|_p = \Big(\int_X |f|^p \, d\mu \Big)^{\frac{1}{p}} \qquad \|f\|_\infty = \inf \Big\{ C \in \overline{R_+},\ |f| \leq C \text{ p.p} \Big\}$$

Propriété () . Soit $f \in M(X)$ et $p \in [1, \infty]$ on a :

$$f \in L^p(X) \iff \|f\|_p < \infty$$

Théorème (**Inégalité de Holder**) . Soient $p, q, r \in [1, \infty]$ tels que :

$$\boxed{\frac{1}{p} + \frac{1}{q} = \frac{1}{r} \implies \forall f, g \in M(X),\ \|fg\|_r \leq \|f\|_p \times \|g\|_q}$$

Théorème (**Inégalité de Young**) . Pour tous $p, q, r \in [1, \infty]$, on a :

$$\forall a, b \in \mathbb{R}_+, \quad \frac{a^p}{p} + \frac{b^p}{q}$$

Remarque Quelques cas particuliers de l'inégalité de Holder :

- $\|fg\|_\infty \leq \|f\|_\infty \times \|g\|_\infty$
- $\|fg\|_p \leq \|f\|_\infty \times \|g\|_p$
- $\|fg\|_1 \leq \|f\|_2 \times \|g\|_2$

Remarque Conséquence de l'inégalité de Holder : Si $\forall f, g \in M(X),\ \|fg\|_r \leq \|f\|_p \times \|g\|_q$ alors on peut dire que :

$$L^p(X)L^q(X) \subseteq L^r(X)$$

Théorème (**Structure des espaces L^p**) . L'espace $(L^p(X), \|.\|_P)$ est un espace vetoriel normé, complet (i.e toutes les suites de Cauchy convergent).

Remarque (Rappel suites de Cauchy) Soit $(u_n) \in X^{\mathbb{N}}$ une suite de X. On dit que c'est une suite de Cauchy si :

$$\forall \varepsilon > 0, \exists n_0 \in \mathbb{N}, \forall n, p \geq n_0, \quad \|u_n - u_p\| < \varepsilon$$

i.e *à partir d'une certain rang tous les termes de la suite se rapprochent autant qu'on veut.*
On remarquera d'une suite convergente est "de Cauchy" mais la réciproques n'est pas tout le temps vraie.

Théorème (**Structure de L^2**) . L^2 est un **espace préhilbertien** complet muni du produit scalaire suivant :

$$\text{où} \quad \langle .,. \rangle : \begin{cases} L^2 \times L^2 \longrightarrow \mathbb{C} \\ (f,g) \longmapsto \int f\overline{g}\, d\mu \end{cases}$$

C'est un produit scalaire hermitien, de plus, bien défini car le produit de deux fonctions L^2 est L^1.

Remarque (Espaces de Banach, espaces de Hilbert) Posons quelques mots de vocabulaire sur les espaces définis :

- Un espace vectoriel muni d'un produit scalaire appelé **espace préhilbertien**.
- Un espace vectoriel normé, complet est appelé **espace de Banach**.
- Un espace préhilbertien complet est appelé **espace de Hilbert**.

Remarque (Inclusion des espaces) Un espace de Hilbert est un espace de Banach. La réciproque est, en général, fausse. Tout espace préhilbertien de dimension finie est un espace de Hilbert puisque la dimension finie implique la complétude.

Exemple Si $X = \mathbb{R}$ ou $X = \mathbb{R}^n$ ou même pour tout borélien de ces espaces muni de la mesure de Lebesgue se notera :

$$L^p(R^n) := L^p(\mathbb{R}^n, \lambda_n)$$

si $X = D$, un ensemble dénombrable muni de la mesure de comptage on notera $l^p(D) := L^p(D, c)$ et plus formellement :

$$\boxed{l^p(D) := \{(u_i)_{i \in D} \in \mathbb{C}^D \mid \|(u_i)\|_p < \infty\}}$$

Remarque (Rappel densité) Soit E un espace normé, et $A \subseteq E$. On dit que A est dense dans E ssi $\overline{A} = E$

$$\text{i.e} \quad \forall x \in E, x \in \overline{A} \iff \exists (x_n)_{n\in\mathbb{N}} \in A^{\mathbb{N}},\ \|x_n - x\| \underset{n\to\infty}{\longrightarrow} 0$$

Par exemple, $\mathbb{Q}$ est dense dans $\mathbb{R}$.

Théorème (**Fonctions à support compact**) . Si $p < \infty$ alors $C_c^0(\mathbb{R}^n)$ est dense dans $L^p(R^n)$ où $C_c^0(\mathbb{R}^n) = \{f \in \mathcal{C}^0(\mathbb{R}^n)$ à support compact $\}$ et $Supp(f) = \{x \in R^n \mid f(x) \neq 0\}$

6.1.3 Produit de convolution

Définitions et propriétés

Définition (**Cas positif**) . Soient $f, g \in \overline{\mathcal{M}_+}(R^n)$, on définit le produit de convolution comme l'application :

$$\boxed{\begin{cases} \overline{\mathcal{M}_+}(R^n) \times \overline{\mathcal{M}_+}(R^n) \to \overline{\mathcal{M}_+}(R^n) \\ (f,g) \longmapsto f * g = \int_{R^n} f(x-y)g(y)\, dy \end{cases}}$$

C'est une loi de composition interne cd $\overline{\mathcal{M}_+}(R^n)$ telle que $*$ est commutative et associative, et :

$$* : \begin{cases} \overline{\mathcal{L}_+^1}(R^n) \times \overline{\mathcal{L}_+^1}(R^n) \to \overline{\mathcal{L}_+^1}(R^n) \\ (f,g) \longmapsto f * g = \int_{\mathbb{R}^n} f\ \lambda_n \times \int_{\mathbb{R}^n} g\ \lambda_n \end{cases}$$

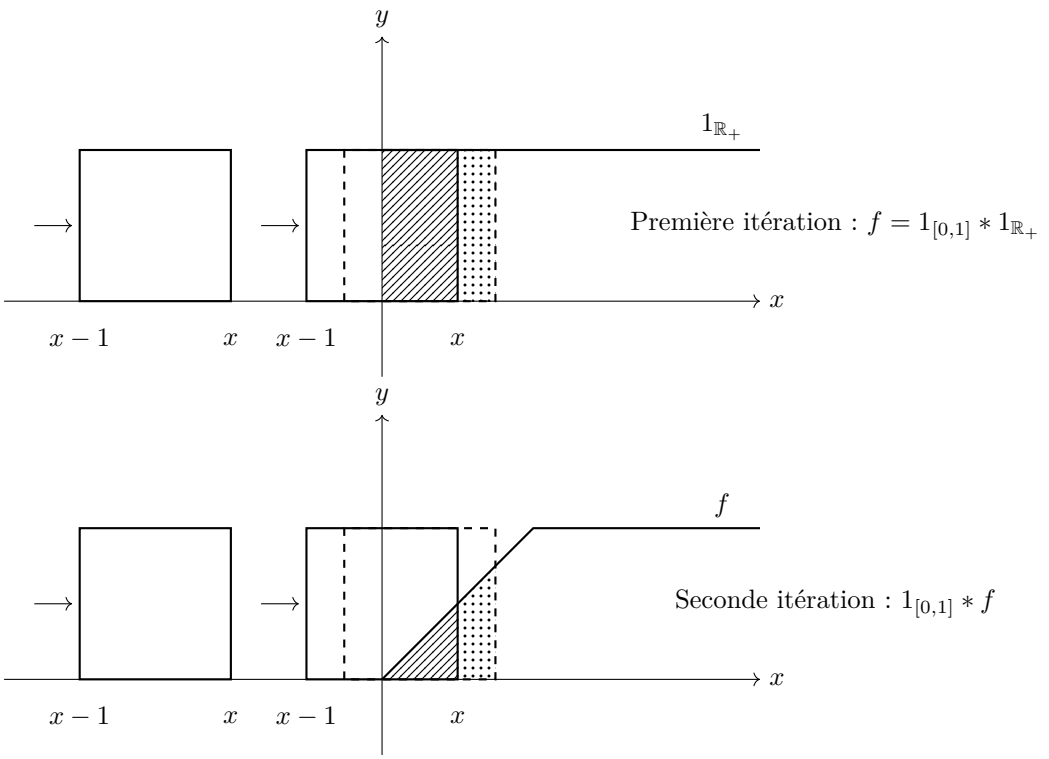
Exemple (Calcul de produit de convolution) Première itération du produit de convolution :

$$\begin{aligned} f = 1_{[0,1]} * 1_{\mathbb{R}_+} &= \int_{\mathbb{R}} 1_{[0,1]}(x-y) \times 1_{\mathbb{R}_+}(y)\, dy \\ &= \int_{R_+} 1_{[0,1]}(x-y)\, dy \\ &= \int_{-\infty}^{x} 1_{[0,1]}(u)\, du = \begin{cases} 0 \text{ si } x < 0 \\ 1 \text{ si } x \geq 1 \\ x \text{ si } x \in]0,1[\end{cases} \end{aligned}$$

Deuxième itération du produit de convolution :

$$\begin{aligned} 1_{[0,1]} * 1_{[0,1]} * 1_{\mathbb{R}_+} = 1_{[0,1]} * f &= \int_{\mathbb{R}} 1_{[0,1]}(x-y) \times f(y)\, dy \\ &= \int_{\mathbb{R}} 1_{[0,1]}(y) \times f(x-y)\, dy = \int_0^1 f(x-y) = \int_x^{x-1} -f(u)\, du \\ &= \int_{x-1}^{x} f(u)du = \begin{cases} 0 \text{ si } x \leq 0 \\ \frac{x^2}{2} \text{ si } x \in]0,1[\\ \frac{-x^2+4x-2}{2} \text{ si } x \in [0,2[\quad (1) \\ 1 \text{ si } x > 2 \end{cases} \end{aligned}$$

(1) : si $x \in [0,2[$ alors on a : $\int_{x-1}^{x} f(u)du = \int_{x-1}^{1} u\, du + \int_1^x 1\, du = \frac{-x^2+4x-2}{2}$



Théorème (**Convolution (Cas non positif)**) . Soient $f, g \in M(X)$. La convolution $f * g$ est la fonction :

$$x \longmapsto \int_{R^n} f(x-y)g(y)\,dy$$

Elle est définie dans les cas suivants :

- si $f \in L^1(X)$ et $g \in L^p(X)$ alors $fg \in L^p(X)$ et

$$\|f * g\|_p \leq \|f\|_1 \times \|g\|_p$$

- si $f \in L^2(X)$ et $g \in L^2(X)$ alors $fg \in L^\infty(X)$

Propriété () .

- Le produit de convolution est commutatif, associatif (quand il est bien défini) et bilinéaire.
- $(L^1(X), *)$ est une $\mathbb{C}$-algèbre sans unité.

Applications du produit de convolution

Théorème (**Régularisation**) . Soit $f \in \mathcal{C}^\infty$ bornée, telle que toutes ses dérivées soient bornées.

alors : $\forall g \in L^1,\ f * g \in \mathcal{C}^\infty$ et toutes ses dérivées sont bornées.

et : $\partial^\alpha (f * g) = (\partial^\alpha f) * g$

$$\text{où} \quad \partial^\alpha = \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \times \cdots \times \left(\frac{\partial}{\partial x_n}\right)^{\alpha_n}$$

Exemple Dans $\mathbb{R}^2$, $\alpha = (3,1)$, on a donc :

$$\partial^\alpha f(x,y) = \frac{\partial^3}{\partial x}\frac{\partial^1}{\partial y} f(x,y)$$

6.1.4 Apprixomation de l'unité

Ayant définit le produit de convolution sur les espaces L^p précédement, nous allons maintenant essayer de voir s'il existe une application de ces espaces candidate pour passer comme "neutre" vis-à-vis de ce produit. C'est à dire, une application laissant invariante toute autre application par le produit de convolution.

Définition (**Apprixomation de l'unité**) . Une suite (ρ_k) est appelée approximation de l'unité si :

1. $\forall k \in \mathbb{N}, \quad \rho_k \in L^1$
2. $\forall k \in \mathbb{N}, \quad \int \rho_k \, d\lambda_n = 1$
3. $\forall \varepsilon > 0, \quad \int_{\mathbb{R}^n \setminus B(0,\varepsilon)} \rho_k \, d\lambda_n \underset{k\to\infty}{\longrightarrow} 0$

Définition (**Suite régularisante**) . Une suite régularisante (U_k) est une approximation de l'unité $\mathcal{C}^\infty$ pour tout k.

Plus formellement, une suite (U_k) est une suite régularisante si $\forall k \in \mathbb{N}, u_k \in \mathcal{C}^\infty$ et (u_k) est une approximation de l'unité.

Théorème (**Intérêt des approximations de l'unité**) . Soit $f \in L^p, p < \infty$. Soit (ρ_k) une approximation de l'unité.

alors : $\rho_k * f \underset{k\to\infty}{\longrightarrow} f$

i.e $\|f - \rho_k * f\|_p \underset{k\to\infty}{\longrightarrow} 0$

Remarque On remarque rapidement l'intérêt du produit de convolution. En effet, il permet de "lisser" des fonctions non régulières L^p et de les approximer par des fonctions C^∞

6.2 Introduction à l'analyse de Fourrier

Ici, nous nous intéressons à l'analyse des fonctions non périodiques sur $\mathbb{R}$. Nous allons étudier une fonctions permettant d'isoler les principales composantes d'une fonction non périodique. L'analogie habituelle est de comparer une fonction, ici dite non régulière, avec un signal, par exemple sonore.

Un tel signal est composé de différentes fréquences. En effet, on peut le décomposer en différents signaux périodiques, dont chacun est responsable d'une fréquence précise du signal de départ. Une fois ce signal décomposé, on peut isoler les fréquences qui ne nous intéressent pas et effectuer l'opération inverse.
On obtient alors un nouveau signal issu du signal de départ avec, par exemple, une suppression du bruit. Une telle application est appelée tranformée de Fourier.

6.2.1 Transformée de Fourier

Définition (**Transformée de Fourier**) . Soit $f \in L^1(\mathbb{R})$, on définit l'application :

$$\mathcal{F} : \begin{cases} L^1 \longrightarrow \mathcal{C}_b \\ f \longmapsto \hat{f} \end{cases} \quad \text{où} \quad \hat{f} : \begin{cases} \mathbb{R} \longrightarrow \mathbb{C} \\ \xi \mapsto \int e^{-2i\pi\xi x} f(x)\, dx \end{cases}$$

Ici, $\mathcal{F}$ est la fonction transformée de Fourier. $\mathcal{C}_b$ représente l'ensemble des fonctions continues bornées.

Propriété (**Transformée de Fourier**) . La fonction transformée de Fourier possède énormément de propriétés, en voici quelques unes :

- Lemme de Riemann-Lebesgue :

$$\boxed{\forall f \in L^1, \quad \hat{f} \in \mathcal{C}^0}$$

- $\mathcal{F}$ est $\mathcal{C}^0$
- $\mathcal{F}$ est **injective**.
- $\mathcal{F}$ transforme les convolutions en produit

$$\boxed{\text{i.e} \quad \forall f, g \in L^1,\ \hat{f} * \hat{g} = \hat{f}\hat{g} \iff \mathcal{F}(f * g) = \mathcal{F}(f) \times \mathcal{F}(g)}$$

- $\mathcal{F}$ transforme les dérivations en produits par des polynômes :

$$\begin{aligned} &\text{si} \quad f \in L^1 \cap C^1 \text{ telle que } f \in L^1 \\ &\text{alors} \quad \widehat{f'} = M\hat{f} \text{ où } M : \xi \longrightarrow 2\pi i\xi \end{aligned}$$

- Formule de symétrie :

$$\forall f, g \in L^1, \quad \int f\hat{g}\, d\lambda = \int \hat{f} g\, d\lambda$$

Théorème (**Inversion**) .

$$\text{Posons } \mathcal{F} : \begin{cases} L^1 \longrightarrow C^0 \\ f \longmapsto \check{f} = R \circ \mathcal{F}(f) \end{cases}$$

telle que $\overline{\mathcal{F}} = R \circ \mathcal{F}$ où $R : f \longmapsto (\xi \mapsto f(-\xi))$. On a alors :

$$\boxed{\check{f}(\xi) = \widehat{f}(-\xi) \int e^{+2i\pi x\xi} f(x)\, dx}$$

Soit $f \in L^1$ telle que $\hat{f} \in L^1$ on a alors :

$$\forall x \in X,\ f(x) = \overline{\mathcal{F}}(\hat{f})(x) \text{ p.p}$$

$$\text{i.e } f = \check{\hat{f}} = \hat{\check{f}} \text{ p.p}$$

Remarque $\overline{\mathcal{F}(f)} = \overline{\mathcal{F}}(\overline{f})$ où $\overline{f}$ est le conjugué de f.

Corollaire (**Théorème d'inversion**) .

- $\mathcal{F}$ est injective
- $\mathcal{F}_{|\mathcal{F}^{-1}(L^1)}$ est une bijection de $\mathcal{F}^{-1}(L^1)$ vers $\mathcal{F}^{-1}(L^1)$. On note alors $\mathcal{A} := \mathcal{F}^{-1}(L^1)$.

Théorème (**Egalité de Plancherel**) . Soient $f, g \in \mathcal{A}$ telles que :

$$\langle f, g\rangle_{L^2} = \langle \hat{f}, \hat{g}\rangle_{L^2}$$
$$\text{i.e } \int f\overline{g}\, d\lambda = \int \hat{f} \times \overline{\hat{g}}\, d\lambda$$
$$\text{i.e } \langle f, g\rangle_{L^2} = \langle \mathcal{F}(f), \mathcal{F}(g)\rangle_{L^2}$$

Donc si $f = g$, alors $\| f \|_2 = \|\hat{f}\|_2$.
Donc $\mathcal{F}$ est une **isométrie** linéaire bijective de $\mathcal{A}$ vers $\mathcal{A}$ pour $\|.\|_2$.

Théorème (**Propriétés de $\mathcal{F}$ sur L^2**) . $\mathcal{F} : L^2 \longrightarrow L^2$ est une isométrie linéaire bijective et :

- $\forall f, g \in L^2, \quad \langle f, g\rangle = \langle \hat{f}, \hat{g}\rangle$
- $\forall f, g \in L^2, \quad \widehat{fg} = \hat{f} * \hat{g}$

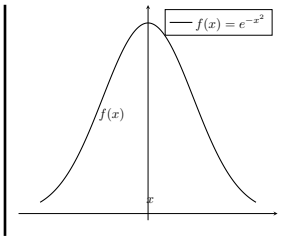
Remarque Soit $f \in L^2$ l'expression

$$\hat{f}(\xi) = \int e^{-2\pi i x\xi} f(x)\, dx$$

n'a pas de sens a priori dans le cas où $f \in L^2 \backslash L^1$. Mais on peut trouver une suite de fonction de $L^1 \cap L^2$ telle que f en soit la limite et ainsi écrire f comme une limite de fonctions de $L^1 \cap L^2$.

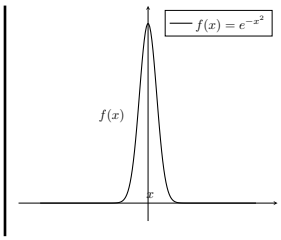
Exemple (Transformée de Fourier et Gaussienne)

$$\gamma_a : \begin{cases} \mathbb{R} \longrightarrow \mathbb{R} \\ x \longmapsto e^{-\pi a x^2} \end{cases}$$



Et en appliquant la transformée de Fourier, on obtient :

$$\hat{\gamma}_a = a^{-1/2}\gamma_{\frac{1}{a}}$$



Visuellement,la transformée de Fourier permet donc "d'isoler" les signaux d'une courbe.

Chapitre 7

Équations différentielles

Contents

Les équations différentielles font partie intégrante de la vie de tous les jours. En effet, elles permettent de modéliser des phénomènes qui varient dans le temps. Du plus simple comme le reffroidissement d'une tasse de café, au plus complexe tels que la modélisation de phénomène quantiques ou climatilogiques.
Nous allons ici définir formellement les équations différentielles. Nous nous attarderons ensuite sur les deux grands types d'équations : les équations différentielles linéaires et non linéaires. Nous finirons par une introduction à l'étude qualitative.

7.1 Équations différentielles - généralités

Nous allons chercher dans ce cours à définir proprement la notion d'équation différentielles, de solution d'une équation différentielle et voir quelques propriétés de ces deux objets.

7.1.1 Définitions et premières propriétés

Définition (**Equation Différentielle Scalaire**) . Soit $f : \Omega \longrightarrow \mathbb{K}$ où $\Omega \subseteq \mathbb{R} \times \mathbb{K}^N$ avec $N \in \mathbb{N}$. L'équation différentielle associée à f est :

$$(E) : y^{(n)} = f(t, y, \dots, y^{(n-1)})$$

On appelle cela une **équation différentielle d'ordre n scalaire**. On peut appeler f la fonction décrivant l'équation différentielle.

Définissons maintenant clairement le concept de "solution" d'une équation différentielle.

Définition (**Solution**) . Une fonction y est appelée solution d'une équation différentielle (E) d'ordre $n \in \mathbb{N}$ définie par la fonction f si :

- $y : I \longrightarrow \mathbb{K}$ est n fois dérivable
- $\forall t \in I, \quad (t, y(t), \dots, y^{(n-1)}(t)) \in \Omega$
- $\forall t \in I$, on a :

$$\boxed{y^{(n)}(t) = f(t, y(t), \dots, y^{(n-1)}(t))}$$

Remarque Si on change l'intervalle de définition I d'une solution d'une équation différentielle, alors on change aussi la solution. Autrement dit, une solution est très dépendante de son intervalle de définition. On peut donc parfois noter (I, y) une solution y définie sur un intervalle I.

Définition (**Equation Différentielle Vectorielle**) . Soit $f : \Omega \longrightarrow \mathbb{K}^N$ avec $n \in \mathbb{N}^*$ et $\Omega \subseteq \mathbb{R} \times (\mathbb{K}^N)^N$. L'équation différentielle ordinaire associée à f est :

$$(E) : y^{(n)} = f(t, y, \dots, y^{(n-1)})$$

On peut écrire cette équation différentielle sous forme de système d'équations différentielles en décomposant les différentes composantes de l'équation différentielle tel que :

$$\begin{cases} y_1^{(n)} = f_1(t, y_1, \dots, y_N^{(n-1)}) \\ \vdots \qquad\qquad\qquad \vdots \\ y_N^{(n)} = f_1(t, y_N, \dots, y_N^{(n-1)}) \end{cases}$$

Lors de l'étude d'un equation différentielle d'ordre n, un théorème que nous verrons plus tard nous permettra de nous ramener à l'étude d'une équation différentielle d'ordre 1. Il est donc important de bien comprendre ce qu'est une équation différentielle d'ordre 1.

Remarque (Equation Differentielle d'ordre 1) Soit $f : \Omega \longrightarrow \mathbb{K}^N$ où $\Omega \subseteq \mathbb{R} \times \mathbb{K}^N$. L'équation différentielle associée à f est :

$$(E) : y' = f(t, y)$$

Exemple Voyons quelques exemples d'équations différentielles :

- $\mathbb{K} = \mathbb{R}, N = 1, n = 1, \Omega = \mathbb{R} \times \mathbb{R}$ et $f(t, x) = x$

$$(E) : y' = y$$

- $\mathbb{K} = \mathbb{R}, N = 2, n = 1, \Omega = \mathbb{R} \times \mathbb{R}^2$

$$f : \begin{cases} \Omega \longrightarrow \mathbb{R}^2 \\ (t, x) \longmapsto Ax + B \end{cases} \qquad \text{où } B \in \mathbb{R}^2 \text{ et } A \in \mathcal{M}_2(\mathbb{R})$$

$$(E) : y' = Ay + B$$

$$(E) : \begin{cases} y_1' = a_{1,1}y_1 + a_{1,2}y_2 + b_1 \\ y_2' = a_{2,1}y_1 + a_{2,2}y_2 = b_2 \end{cases}$$

- $\mathbb{K} = \mathbb{C}, N = 1, n = 1, \Omega = \mathbb{R} \times \mathbb{C}$

$$f : \begin{cases} \Omega \longrightarrow \mathbb{C} \\ (t, z) \longmapsto z^2 + t \end{cases} \qquad \text{et} \quad (E) : y' = y^2 + t$$

Lors de la résolution d'une équation différentielle, on remarque que on trouve la plupart du temps une infinité de solutions. Parfois, une unique solution serait plus pratique (ex : en physique). On va donc chercher à donner des conditions supplémentaires sur la fonction solution en imposant une de ses valeurs.

Définition (**Problème de Cauchy**) . Soit (E) une équation différentielle ordinaire d'ordre 1, de fonction descriptive f et de la forme $y' = f(t, x)$. Résoudre un problème de Cauchy associé à (E) consiste à trouver une solution de (E) qui vérifie un condition initiale :

$$(C) : y(t_0) = y_0 \quad \text{où } (t_0, y_0) \in \Omega$$

Proposition (Exercice) Soit $f : \Omega \subset \mathbb{R} \times \mathbb{K}^N \longrightarrow \mathbb{K}^N$ et (E) l'équation associée d'ordre 1 :

$$y' = f(t, y)$$

Alors : $y : I \longrightarrow \mathbb{K}^N$ est solution du problème de Cauchy :

$$(P) : \begin{cases} (E) : y' = f(t, y) \\ (C) : y(t_0) = y_0 \end{cases}$$

si et seulement si

- $y \in \mathcal{C}^0(I)$
- $\forall t \in I, \quad (t, y(t)) \in \Omega$
- $\forall t \in I,$

$$y(t) = y_0 + \int_{t_0}^{t} f(s, y(s))\, ds$$

Définition (**Solution Maximale**) . y est une solution maximale de (E) si il n'existe pas de prolongement strict de y en une solution de (E).

Définition (**Solution Globale**) . Soit $(E) : y' = f(t, y)$ où $f : I \times \Omega \longrightarrow \mathbb{K}^N$. On dit que y est une solution globale de (E) si y est une solution de (E) et qu'elle est définie sur I.

Remarque Une solution globale est maximale. La réciproque est fausse, voyons un exemple.

$$(E) : y' = y^2 \quad \text{et} \quad y_s : \begin{cases} A \longrightarrow \mathbb{R} \\ t \longmapsto -\frac{1}{t} \end{cases} \quad \text{où } f : \begin{cases} \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R} \\ (t, x) \longmapsto x^2 \end{cases} \qquad \text{et} \quad A = \mathbb{R}_+^*$$

y_s n'est pas une solution globale car on ne peut pas la prolonger en une fonction continue solution de (E). En revanche, c'est une solution maximale.

Théorème (**Existence et Unicité de solution maximale**) . Il existe une solution maximale d'un problème de Cauchy.

7.1.2 Première méthode résolution

A partir des outils que nous avons à notre porté, nous pouvons dès à présent déterminer la forme de l'ensemble des solution d'une équation différentielle scalaire homogène d'ordre 1. Soient I un intervalle, $a : I \longrightarrow \mathbb{K}$ une fonction continue et (E) une équation différentielle de la forme :

$$(E) : y' = a(t)y = f(t,y) \quad \text{où} \quad f : \begin{cases} I \times \mathbb{K} \longrightarrow \mathbb{K} \\ (t,x) \longmapsto a(t)x \end{cases}$$

Soit le problème de Cauchy suivant :

$$(P) : \begin{cases} (E) : y' = a(t)y = f(t,y) \\ y(t_0) = y_0 \end{cases} \qquad \text{où} \quad (t_0, y_0) \in I \times \mathbb{K}$$

On cherche la forme générale de l'ensemble des solutions du problème de Cauchy.

Existence d'une solution

Soit $y : I \longrightarrow \mathbb{K}, t \longmapsto e^{\int_{t_0}^{t} a}$. Montrons que y est bien une solution maximale du problème de Cauchy.

- Soit $t \in I$ alors $y(t) \in \mathbb{K}$
- $\forall t \in I$, on a : $y'(t) = y_0 a(t) e^{\int_{t_0}^{t} a}$ Donc y est bien une solution de (E). De plus, y est une solution globale donc elle est maximale.
- Condition initiale :
$$y(t_0) = y_0 e^{\int_{t_0}^{t_0} a} = y_0 e^0 = y_0$$
On a donc que y est bien une solution maximale du problème de Cauchy énoncé.

Unicité de la solution

Soit $u : J \subset I \longrightarrow \mathbb{K}$ une solution maximale de (E) telle que $t_0 \in J$ et $u(t_0) = y_0$.

Montrons que la fonction $v : t \longmapsto u(t) e^{-\int_{t_0}^{t} a}$ est constante. v est dérivable par composition/-produit. On a donc $\forall t \in J$,

$$\begin{aligned} y'(t) &= u'(t) \times e^{\int_{t_0}^{t} a} - u(t)a(t)e^{\int_{t_0}^{t} a} \\ &= e^{\int_{t_0}^{t} a}(u'(t) - u(t)a(t)) \end{aligned}$$

or u est un solution maximale de (E) donc $\forall t \in J, u'(t) = a(t)u(t) \iff u'(t) - u(t)a(t) = 0$ d'où :

$$y'(t) = 0$$

Donc v est une fonction constante sur J. Déterminons cette constante.

$$v(t_0) = u(t_0) e^{-\int_{t_0}^{t_0} a} = u(t_0)e^0 = u(t_0) = y_0$$

Donc $\forall t \in J,\ v(t) = y_0$.

Montrons que $u = y_{|J}$ et que $I = J$. Soit $t \in J$, on a :

$$\begin{aligned} v(t) &= u(t)e^{-\int_{t_0}^{t_0} a} \\ \iff u(t) &= v(t)e^{\int_{t_0}^{t_0} a} \\ u(t) &= y_0 e^{\int_{t_0}^{t_0} a} \end{aligned}$$

Donc $u = y_{|J}$. De plus, y est un prolongement de u or u est supposée maximale (i.e elle n'admet pas de prolongement solution) donc $I = J$.

Conclusion

Enfin, ces calculs nous permettent d'énoncer le résultat suivant :

Théorème (**Solutions d'une ED scalaire homogène d'ordre 1**) . Soient I un intervalle, $a : I \longrightarrow \mathbb{K}$ une fonction continue et $f : I \times \mathbb{K} \longrightarrow \mathbb{K}$.
Soit $(E) : y' = a(t)y = f(t, y)$ l'équation différentielle homogène d'ordre 1 associée à f. Soient $(t_0, y_0) \in I \times \mathbb{K}$, et $(P) : \{(E), y(t_0) = y_0\}$ un problème de Cauchy. Alors les solutions maximales de (P) sont exactement les fonctions de la forme :

$$y : \begin{cases} I \longrightarrow \mathbb{K} \\ t \longmapsto y_0 \exp\left(\int_{t_0}^{t} a(s)\, ds\right) \end{cases}$$

7.2 Équations différentielles Linéaires

7.2.1 Notions d'analyse Matricielle

Soient E et F deux espaces vectoriels normés. $\mathcal{L}(E, F)$ désigne l'espace vectoriel des applications linéaires de E vers F. Il est muni d'une norme appelée **norme opérationnelle** :

$$\forall f \in \mathcal{L}(E, F), \quad \|f\| = \min\{C \in \mathbb{R}_+ \mid \forall x \in E, \|f(x)\|_F \leqslant C\|x\|_E\}$$

Définition (**Norme Subordonnée**) . Si $\|.\|$ est une norme sur $\mathbb{K}^n$ alors la norme subordonnée associée à $\|.\|$ sur $\mathcal{M}_n(\mathbb{K})$ est la norme opérationnelle associée à $\|.\|$ sur $\mathbb{K}^n = E$ et $\|.\|$ sur $\mathbb{K}^n = F$. On la note de la même façon.

$$\text{i.e} \quad \forall M \in \mathcal{M}_n(\mathbb{K}), \quad \|M\| = \min\{C \in \mathbb{R}_+ \mid \forall x \in \mathbb{K}^n, \|Mx\| \leqslant C\|x\|\}$$

Proposition Soient $M, N \in \mathcal{M}_n(\mathbb{K})$ alors :

$$\|M \times N\| \leqslant \|M\| \times \|N\|$$

Ceci d'écoule de l'inégalité de Cauchy-Schwarz.

Remarque Soit $f : I \longrightarrow \mathcal{M}_n(\mathbb{K})$ où I est un intervalle. Supposons que f soit continue. On peut alors définir $\int_I f(t)\, dt \in \mathcal{M}_n(\mathbb{K})$ composante par composante telle que :

$$\left(\int_I f(t)\, dt\right)_{i,j} = \int_I f(t)_{i,j}\, dt$$

Autrement dit, l'intégrale d'une matrice d'applications continues sur un intervalle est exactement la matrice les intégrales des applications.

Proposition Soit $f : I =]a, b[\xrightarrow{\mathcal{C}^0} \mathcal{M}_n(\mathbb{K})$. Alors :

$$\left\| \int_I f(t)\ dt \right\| \leqslant \int_a^b \|f(t)\|\ dt$$

Propriété (**Théorème Fondamental de l'Analyse**) . Dans ce cas ci, le théorème fondamental de l'analyse est toujours vrai. Autrement dit, soit $f : I =]a, b[\xrightarrow{\mathcal{C}^0} \mathcal{M}_n(\mathbb{K})$. Alors :

$$\int_a^b f'(t)\ dt = f(b) - f(a)$$

Remarque Soit $f \in \mathcal{C}^1(I, \mathcal{M}_n(\mathbb{K}))$. Alors la formule $(f^2)' = 2f'f$ est fausse puisque $\mathcal{M}_n(\mathbb{K})$ n'est pas un anneau commutatif. D'où :

$$(fg)' = f'g + fg' \quad \text{où} \quad g \in \mathcal{C}^1(I, \mathcal{M}_n(\mathbb{K}))$$

Remarque Pour rappel, on sait que $\forall z \in \mathbb{C},\ e^z = \sum_{k=0}^{\infty} \frac{z^k}{k!}$.

Définition (**Exponentielle d'une matrice**) . Soit $A \in \mathcal{M}_n(\mathbb{K})$ on a alors :

$$\exp(A) := \sum_{k=0}^{\infty} \frac{A^k}{k!}$$

Cette série converge toujours.

Proposition (Propriétés de l'exponentielle) Soient $A, B \in \mathcal{M}_n(\mathbb{K})$ alors

$$e^{A+B} = e^A e^B \iff AB = BA$$

Proposition (Inverse) Soit $A \in \mathcal{M}_n(\mathbb{K})$ alors e^A est inversible. Pour cela, on pose :

$$e^0 = \sum_{k=0}^{\infty} \frac{O^k}{k!} = 1_{\mathcal{M}_n(\mathbb{K})} = I_n \quad \text{et} \quad e^{(A+(-A))} = e^A e^{-A} = I_n$$

Propriété (**Dérivation de l'exponentielle matricielle**) .

- Soit $A \in \mathcal{A}_n(\mathbb{K})$. La fonction :

$$\begin{cases} \mathbb{R} \longrightarrow \mathcal{M}_n(\mathbb{K}) \\ t \longmapsto e^{tA} \end{cases}$$

  est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$ et :

$$\boxed{\frac{\partial e^{tA}}{\partial t} = Ae^{tA} = e^{tA}A}$$

- Soit $A \in \mathcal{C}^1(\mathbb{R}, \mathcal{M}_n(\mathbb{K}))$ telle que $\forall t \in \mathbb{R}, A(t)A'(t) = A'(t)A(t)$ alors :

$$t \longmapsto e^{A(t)} \in \mathcal{C}^1 \quad \text{et} \quad \frac{\partial e^{A(t)}}{\partial t} = A'(t)e^{A(t)}$$

Le théorème fondamental de l'analyse est toujours vrai.

Remarque Soit $f \in \mathcal{C}^1(I, \mathcal{M}_n(\mathbb{K}))$ la formule :

$$(f^2)' = 2f'f$$

est fausse puisque $(\mathcal{M}_n(\mathbb{K}), +, \times)$ n'est pas un anneau commutatif.

Définition (**Exponentielle**) . Soit $z \in \mathbb{C}$ on définit l'exponentielle de z par :

$$e^z = \exp(z) = \sum_{k=0}^{\infty} \frac{z^k}{k!}$$

Définition (**Exponentielle de Matrice**) . Soit $A \in \mathcal{M}_n(\mathbb{K})$ on a alors :

$$\exp(A) = e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!}$$

Cette série converge toujours.

Remarque Cette définition nous servira plus tard pour la résolution d'équations différentielles en dimension $n \in \mathbb{N}$.
On peut remarquer que la fonction $e^A : t \longmapsto e^{A(t)}$ est solution de

$$(E) : Y' = A'Y$$

7.2.2 Théorème de Cauchy Lipschitz Linéaire

Définition (**EQ linéaire du premier ordre**) . Soit $A \in \mathcal{C}^0(I, \mathcal{M}_n(\mathbb{K}))$ et $B \in \mathcal{C}^0(I, \mathbb{K}^n)$.
Une équation différentielle linéaire du premier ordre est une équation du type :

$$\boxed{\forall t \in I, \quad (E) : y'(t) = A(t)y(t) + B(t) = f(t, y)}$$

$$\text{où : } f : \begin{cases} I \times \mathbb{K}^n \longrightarrow \mathbb{K}^n \\ (t, x) \longmapsto A(t)x + B(t) \end{cases}$$

Théorème (**Cauchy-Lipschitz Linéaire**) . Soit (E) une équation diférentielle linéaire. Alors ile xiste une unique solution maximale au problème de Cauchy :

$$(P) : \begin{cases} (E) \\ y(t_0) = y_0 \end{cases}$$

avec $(t_0, y_0) \in I \times \mathbb{K}^n$. De plus, cette solution est **globale**.

Remarque Pour une équation d'ordre 2 scalaire :

$$(E) : y''(t) = a_0(t)y(t) + a_1(t)y'(t) + b(t) \quad \forall t \in U$$

où $a_0, a_1, b \in \mathcal{C}^1(I, \mathbb{K})$ équivalente à $Y'(t) = A(t)Y(t) + B(t)$ où

$$A(t) = \begin{pmatrix} 1 & 0 \\ a_0(t) & a_1(t) \end{pmatrix} \quad \text{et} \quad B(t) = \begin{pmatrix} 0 \\ b(t) \end{pmatrix} \quad \text{et} \quad Y(t) = \begin{pmatrix} y(t) \\ y'(t) \end{pmatrix}$$

D'après Cauchy-Lipschitz, il existe une unique solution globale à :

$$\begin{cases} Y'(t) = A(t)Y(t) + B(t) \\ Y(t_0) = Y_0 \end{cases} \qquad \text{où} \quad (t_0, Y_0) \in I \times K^n$$

Alors il existe une unique solution à l'équation globale d'où :

$$\begin{cases} y''(t) = a_0(t)y(t) + a_1(t)y'(t) + b(t) \\ y(t_0) = y_0 \\ y'(tt_0) = y_1 \end{cases}$$

Pour l'ordre 2, il faut donc fixer une valeur pour chaque ordre de l'équation. pour l'ordre 3, il faut en fixer 3, etc...

7.2.3 Structure de l'ensemble des solutions

Dans cette section, nous allons essayer de donner une structure à l'ensemble des solutions d'une équation différentielle. Cela nous permettra d'avoir de meilleures propriétés sur ces solutions et, à partir de solutions particulières d'en déduire des solutions plus générales.

Soit l'équation différentielle définie sur un intervalle I suivante :

$$(L) : y'(t) = A(t)y(t) + B(t) \quad (H) : y'(t) = A(t)y(t)$$

on appelle (H) l'équation homogène de (L). Soient y_1 et y_2 deux solutions de (H) alors $\forall \lambda \in \mathbb{K}$, on a :

$$(\lambda y_1 + y_2)'(t) = \lambda A(t)y_1(t) + A(t)y_2(t) = A(t)(\lambda y_1(t) + y_2(t))$$

Donc c'est aussi une solution de (H).

Théorème (**Structure**) . Soit (L) une équation différentielle linéaire d'ordre 1 et (H) son équation homogène associée. S_H est un sous-espace vectoriel de $\mathcal{C}^0(I, \mathbb{K}^n)$ de dimension finie égale à n (on peut donc trouver une base de S_H). On a donc :

$$\boxed{S = S_H + y_1}$$

où :

- S est l'ensemble des solutions de (L)
- y_1 est une solution particulière de (L)

On dit que S est un espace affine.

Remarque (Construction d'une solution particulière)

$$(L) : y(t) = a(t)y(t) + b(t) \quad \forall i \in I$$

Posons $\forall t \in I, y(t) = Ce^{\int_{t_0}^{t} a(t)\, ds}$ avec $C \in \mathbb{K}$. Modifions notre potentielle solution telle que :

$$\forall t \in I, \quad y(t) = C(t)e^{\int_{t_0} t a(s)\, ds}$$

où C est une fonction dérivable. On a donc que y est dérivable par produit/composition et d'après le théorème fondamental de l'analyse. D'où :

$$\begin{aligned} y'(t) &= C'(t)e^{\int_{t_0}^{t} a(s)\ ds} + C(t)a(t)e^{\int_{t_0}^{t} a(s)\ ds} \\ &= a(t)y(t) + b(t) \\ &\iff a(t)y(t) + b(t) = C'(t)e^{i(t)} + C(t)a(t)e^{i(t)} \\ &\iff b(t) = C'(t)e^{i(t)} \\ &\iff C'(t) = b(t)e^{-i(t)} \end{aligned}$$

Donc $\forall t \in I, C(t) = C(t_0) \int_{t_0}^{t} b(s)\ ds$.
Réciproquement :
La fonction $t \longmapsto \left(\lambda \int_{t_0}^{t} b(s)\ ds\right) e^{-i(t)}$ est bien solution de (L). On peut donc en déduire une solution particulière de (L) en fixant λ est t_0. De plus :

$$S_H = vect\left(t \longmapsto e^{\int_{t_0}^{t} a(s)\ ds}\right) = \left\{t \longmapsto \lambda e^{\int_{t_0}^{t} a(s)\ ds} \mid \lambda \in \mathbb{K}\right\}$$

Théorème (**Résolution Complète**) . Soit $(L) : y'(t) = a(t)y(t) + b(t)$ une équation différentielle linéaire scalaire d'ordre 1 où $a, b \in \mathcal{C}^0(I, \mathbb{K}^n)$. Alors l'ensemble des solutions de S de (L) est de la forme :

$$S = S_H + \left\{t \longmapsto \left(\int_{t_0}^{t} b(s)e^{-\int_{t_0}^{t} a(s)\ ds}\right) e^{\int_{t_0}^{t} a(s)\ ds}\right\}$$

7.2.4 Matrices Fondamentales

Généralités et propriétés

Nous savons que l'ensemble des solutions d'une équation différentielle linéaire est un espace affine composé d'un sev et d'une solution particulière. Le sev S_H est l'ensemble des solutions de l'équation homogène (H) associée à (L) on sait qu'il est de dimension $n \in \mathbb{N}$. On peut donc en trouver une base.

Soit $(H) : y'(t) = A(t)y(t) \quad \forall t \in I$. Une équation homogène d'ordre 1 en dimension $n \in \mathbb{N}$ et $A \in \mathcal{C}^0(I, \mathcal{M}_n(\mathbb{K}))$.

Définition (**Système Fondamental**) . Un système fondamental de solutions est une famille finie qui forme une base vectorielle de S_H.

Définition (**Matrice Fondamentale**) . Soit $\mathcal{F} = (y_1, \ldots, y_n)$ un système fondamental de (H). Une matrice fondamentale de (H) est une **fonction** Φ telle que :

$$\Phi : \begin{cases} I \longrightarrow \mathcal{M}_n(\mathbb{K}) \\ t \longmapsto (y_1(t), \ldots, y_n(t)) \end{cases}$$

Remarque Pour tout $i \in I$, Φ est inversible.

Définition (**Wronksien**) . Le wronksien associé à (H) est la fonction :

$$\omega : \begin{cases} I \longrightarrow \mathbb{K}\backslash\{0\} \\ t \longmapsto \det(\Phi(t)) \end{cases}$$

où Φ est une matrice fondamentale associée à (H). C'est le déterminant de la matrice fondamentale évaluée en $t \in I$.

Remarque Soit Φ une matrice fondamentale de (H) et $P \in GL_n(\mathbb{K})$ alors ΦP est aussi une matrice fondamentale de (H).

Soit Φ une matrice fondamentale de (H) et y une solution particulière de (H). Alors $\exists c_1, \ldots, n_1 \in \mathbb{K}$ tels que :

$$\forall i \in I, \quad y(t) = \sum_{i=1}^{n} c_i y_i(t) \quad \iff \quad y(t) = \Phi(t) \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$$

Proposition L'ensemble des solutions de (H) est de la forme :

$$S_H = \{\Phi C \mid C \in \mathbb{K}^n\}$$

où Φ est une matrice fondamentale de (H). De plus, l'unique solution du problème de Cauchy :

$$(P) : \begin{cases} (H) \\ y(t_0) = y_0 \quad (t_0, y_0) \in I \times \mathbb{K}^n \end{cases}$$

est la fonction $\Phi\Phi^{-1}(t_0)y_0 : t \longmapsto \Phi(t)\Phi^{-1}(t_0)y_0$.

On peut donc en conclure que dès que l'on connaît une matrice fondamentale d'une équation homogène (H), alors on en connaît toutes les solutions. Seulement, comment construire une telle matrice ?

Propriété (**Matrice Fondamentale et Base**) . Soit $\Phi : I \xrightarrow{\mathcal{C}^1} \mathcal{M}_n(\mathbb{K})$. Φ est une matrice fondamentale de (H) ssi :

1. $\forall t \in I, \quad \Phi(t) \in GL_n(\mathbb{K})$ (famille libre)
2. $\forall t \in I, \quad \Phi'(t) = A(t)\Phi(t)$ (famille génératrice de S_H)

Méthode du polynôme caractéristique

Soit (H_n) une équation linéaire homogène d'ordre $n \in \mathbb{N}$ en dimension 1 telle que :

$$\forall t \in (H_n) : a_n y^{(n)}(t) = a_0 y(t) + a_1 y'(t) + \cdots + a_{n-1} y^{(n-1)}(t)$$

où $a_n, \ldots, a_1 \in \mathbb{K}$.

Définition (**Polynôme caractéristique**) . Soit (H_n) une équation différentielle homogène d'ordre $n \in \mathbb{N}$. On définit le polynôme caractéristique de (H_N) comme le polynôme tel que :

$$P(X) = X^n + a_{n-1}X^{n-1} + \cdots + a_1 X + a_0$$

On dira que **l'équation caractérisitique de** (H_n) est $P(X) = 0$.

Théorème (**Construction des solutions**) . Plaçons nous dans $\mathbb{C}$. Soient $r_1, \dots, r_p,\ p \leqslant n$ les racines de $P(X)$ de multiplicités respectives $m_1, \dots, m_p$ dans $\mathbb{C}$. Alors :

$$\forall (k,l) \in \mathbb{N}^2,\ 1 \leqslant k \leqslant p,\ 0 \leqslant l \leqslant m_k - 1 \quad F_{k,l} : t \longmapsto t^l e^{r_k \times t}$$

donne une base (F, k, l) de solutions de (H_n).

7.2.5 Méthode de variation de la constante

Soient (L) une équation différentielle linéaire d'ordre 1 en dimension $n \in \mathbb{N}$ et (H) son équation homogène associée telle que :

$$\forall t \in I, \quad (L) : y'(t) = A(t)y(t) + B(t) \quad y'(t) = A(t)y(t)$$

Pour rappel, l'ensemble des solutions de cette équation est de la forme :

$$S = S_H + y$$

où y est une solution particulière de (L) et S_H l'ensemble des solutions de (H).
Ici, on cherche simplement **une solution particulière** y de la forme :

$$y = \Phi C \quad \text{où } C : \xrightarrow{\mathcal{C}^0} \mathbb{K}^n$$

et Φ est la matrice fondamentale associée à (H).
y est solution donc nécessairement, $\forall t \in I$:

$$\begin{aligned} y'(t) &= A(t)\Phi(t)C(t) + \Phi(t)C'(t) \\ &= A(t)y(t) + B(t) \\ &= A(t)\left(\Phi(t)C(t)\right) + B(t) \end{aligned}$$

$$\iff \begin{cases} \Phi'(t) = A(t)\Phi(t) \\ C'(t) = \Phi^{-1}(t)B(t) \end{cases}$$

D'où après intégration : $C(t) = C(t_0) + \int_{t_0}^{t} \Phi^{-1}(s)B(s)\ ds$. Une solution particulière de (L) est donc :

$$\boxed{y : t \longmapsto \Phi(t)\left(C(t_0) + \int_{t_0}^{t} \Phi^{-1}(s)B(s)\ ds\right)}$$

7.3 Équations différentielles non linéaires

7.3.1 Cauchy-Lipschitz

Problème : On cherche une condition suffisante d'existence et d'unicité d'une solution maximale au problème de Cauchy :

$$(P) : \begin{cases} (E) : y'(t) = f(t,y) \\ (C) : y(t_0) = y_0 \end{cases}$$

Définition (**Fonction Localement Lipschitzienne**) . Soit

$$f : \begin{cases} \Omega = I \times \Omega' \longrightarrow \mathbb{K}^n \\ (t,x) \longmapsto f(t,x) \end{cases}$$

On dit que f est LLVE (localement lipschitzienne en la variable d'état i.e x) <u>si</u> :

$$\forall (t_0, y_0) \in I \times \Omega', \exists [t_0 - \alpha, t_0 + \alpha] \times B_f(y_0, r)$$

Tel que :

$$\exists k \in \mathbb{R}_+, \forall t \in [t_0 - \alpha, t_0 + \alpha], \text{ on a :}$$

$$\boxed{\forall x, x' \in B_f(y_0, r), \|f(t,x) = f(t,x')\| \leqslant k\|x - x'\|}$$

On cherche donc un cylindre $[t_0 - \alpha, t_0 + \alpha] \times B_f(y_0, r) \subset I \times \Omega'$ sur lequel f soit lipschitzienne.

Définition (**Fonction Lipschitzienne en la variable d'état**) . Soit

$$f : \begin{cases} \Omega = I \times \Omega' \longrightarrow \mathbb{K}^n \\ (t,x) \longmapsto f(t,x) \end{cases}$$

On dit que f est LVE (lipschitzienne en la variable d'état i.e x) <u>si</u> :

$$\forall J \in I \text{ compact } \exists k \in \mathbb{R}_+ \text{ tel que :}$$

$$\forall t \in J, \forall x, x' \in \Omega', \quad \boxed{\|f(t,x) - f(t,x')\| \leqslant k\|x - x'\|}$$

Remarque Une fonction lipschitzienne en la variable d'état est localement lipschitzienne en la variable d'état.

$$\text{i.e} \quad LVE \Longrightarrow LLVE$$

Exemple Soit $f : (t,x) \longmapsto A(t)x + B(t)$. Supposons que $A, B \in \mathcal{C}^0(I, \mathcal{M}_n(\mathbb{K}))$.
f est LVE car $\forall t \in J, x, x' \in \mathbb{K}^n$ où $J \subset I$, on a :

$$\begin{aligned} \|f(t,x) - f(t,x')\| &= \|A(t)(x - x')\| \\ &\leqslant \|A(t)\| \times \|x - x'\| \\ &\leqslant \sup_{t \in J} \|A(t)\| \in \mathbb{R}_+ \end{aligned}$$

Théorème (**Cauchy-Lipschitz LVE**) . Soit

$$(P) : \begin{cases} (E) : y'(t) = f(t,y) \\ (C) : y(t_0) = y_0 \end{cases} \qquad \text{où } f : I \times \Omega \xrightarrow{\text{LVE}} \mathbb{K}^n$$

<u>Alors</u> il existe une unique solution maximale au problème de Cauchy (P). Cette solution est même globale.

Théorème (**Cauchy-Lipschitz LLVE**) . Soit

$$(P) : \begin{cases} (E) : y'(t) = f(t,y) \\ (C) : y(t_0) = y_0 \end{cases} \qquad \text{où } f : I \times \Omega \overset{\text{LLVE}}{\longrightarrow} \mathbb{K}^n$$

Alors il existe une unique solution **maximale** au problème de Cauchy (P).

Attention : Cette solution n'est pas forcément globale.

Exemple

$$f : \begin{cases} \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R} \\ (t,x) \longmapsto x^2 \end{cases} \quad \text{LLVE mais non LVE} \qquad g : \begin{cases} R \times R \longrightarrow \mathbb{R} \\ (t,x) \longmapsto \sqrt{|x|} \end{cases} \quad \mathcal{C}^0 \text{ mais pas LLVE}$$

En pratique nous utiliserons plutôt le résultat suivant pour invoquer le théorème de Cauchy-Lipschitz.

Proposition Toute fonction $\mathcal{C}^1$ par rapport à la variable d'état est LLVE.

7.3.2 Équations à variables séparables

Définition (**Équation différentielle à variable séparable**) . Une équation différentielle à variables séparables est une équation différentielle scalaire de la forme :

$$(E) : y'(t) = f(y)g(t)$$

où :

- $f : \Omega \overset{\mathcal{C}^0}{\longrightarrow} \mathbb{K}$
- $g : I \overset{\mathcal{C}^0}{\longrightarrow} \mathbb{K}$

Exemple Quelques exemples d'équation différentielles à variables séparables :

7.3.3 Théorème de sortie de tout compact

Théorème (**Sortie de tout compact**) . Soit $(E) : y'(t) = f(t,y)$ sur $]a,b[$ où :

$$f :]a,b[\times\Omega \longrightarrow \mathbb{K}^n$$

Soit $y :]c,d[\longrightarrow \mathbb{K}^n$ une solution maximale de (E) non globale telle que $d < b$.
Alors $\forall K \subseteq \Omega$ compact, il existe un voisinnage V de d tel que :

$$\forall t \in V, y(t) \notin K$$

Théorème (**Explosion**) . Avec les mêmes hypothèses et $\Omega = \mathbb{K}^n$, on a :

- Si $b > d$ alors $\|y(t)\| \underset{t\to d}{\longrightarrow} \infty$

  Par contraposée, si $\|y(t)\| \underset{t\to d}{\not\longrightarrow} \infty$ (i.e y est bornée au voisinnage de d) alors $b = d$
- Si y est bornée alors elle est globale.
- Si f est bornée alors toute fonction maximale est globale.

7.3.4 Flot d'une équation différentielle

Définition (**Flot**) . Soit $(E) : y'(t) = f(t, y)$ où

$$f : \Omega \subset \mathbb{R} \times \mathbb{K}^n \longrightarrow \mathbb{K}^n \quad \text{LLVE}$$

Le flot de cette équation est la fonction :

$$\Phi : \begin{cases} D \longrightarrow \mathbb{K}^n \\ (t, t_0, y_0) \longmapsto y_{t_0,y_0}(t) \end{cases}$$

où

- y_{t_0,y_0} est l'unique solution maximale de (E) qui vérifie $y(t_0) = y_0$.
- $D = \{(t, t_0, y_0) \mid t \in I_{t_0,y_0}\}$
- I_{t_0,y_0} est le domaine de définition de t_0 et y_0

Propriété (**Flot**) .

- D est un ouvert
- Φ est localement lipschitzienne (et donc $\mathcal{C}^0$)
- Si f est $\mathcal{C}^p$ alors Φ est $\mathcal{C}^p$.

Définition (**Equation Différentielle à paramètre**) . Une équation différentielle à paramètre est une famille d'équation différentielles

$$(E_\lambda) \text{ où } E_\lambda : y'(t) = f(t, y)$$

$$\text{et } \ f : \begin{cases} \Omega \times \Lambda \longrightarrow \mathbb{K}^n \\ (t, x, \lambda) \longmapsto f(t, x, \lambda) \end{cases}$$

Définition (**Flot d'une équation différentielle à paramètre**) . Soit (E_λ) une famille d'équation différentielles à parmaètre telle que $\forall \lambda \in \Lambda$ on ait :

$$f(., ., \lambda) \text{ LLVE}$$

Le flot paramétré est :

$$\Phi : \begin{cases} D \longrightarrow \mathbb{K}^n \\ (t, t_0, y_0, \lambda) \longmapsto y_{t_0,y_0,\lambda}(t) \end{cases}$$

où $y_{t_0,y_0,\lambda}$ est l'unique solution maximale de (E_λ) telle que $y(t_0) = y_0$.

Propriété (**Flot**) .

- D est un ouvert
- Φ est localement lipschitzienne (et donc $\mathcal{C}^0$)
- Si f est $\mathcal{C}^p$ alors Φ est $\mathcal{C}^p$.

7.4 Introduction à l'étude qualitative

En pratique, on ne peut que très rarement résoudre formellement une équation différentielle. On va donc cherche, à partir de la fonction qui la définit, de déduire des propriétés sur les solutions.

7.4.1 Équations Autonomes

Définition (**Équation Autonome**) . Une équation différentielle est dite autonomme si elle est de la forme :

$$(A) : y'(t) = f(y(t))$$

Autrement dit si la variable temporelle (i.e t) n'apparaît pas dans f.

On peut interpréter une équation autonome en physique comme une équation à la contrainte de temps n'apparaît pas.
Posons maintenant quelques définitions supplémentaires...

Définition (**Champ de vecteurs**) . Un champ de vecteurs est une fonction continue $f : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}^n$ où Ω est un ouvert. Intuitivement, elle associe à chaque point de l'espace Ω un vecteur.

Définition (**Courbe Intégrale**) . Une courbe intégrale d'un champ de vecteur $f : \Omega \longrightarrow \mathbb{R}^n$ est une **courbe paramétrée** $y : I \subset \mathbb{R} \longrightarrow \Omega$ telle que y est solution maximale de l'équation autonome :

$$(E) : y'(t) = f(y(t)) \quad \forall i \in I$$

Définition (**Trajectoire**) . Une trajectoire de $f : \Omega \longrightarrow \mathbb{R}^n$ est l'image d'une courbe intégrale de f.

Définition (**Point Stationnaire**) . On dit que $x_0 \in \Omega$ est un point stationnaire ou point d'quilibre si $f(x_0) = 0$. Un point stationnaire est un point où le champ de vecteur est nul.

Ainsi, un champ de vecteur va permettre de modéliser un système en mouvement tel que l'écoulement d'un fluide. Une trajectoire de ce champ de vecteur sera donc la trajectoire d'une particule dans le fluide.

Proposition Soit y une courbe intégrale d'une équation autonome (A) alors si $c \in \mathbb{R}$ alors la courbe :

$$y_c : \begin{cases} I + c \longrightarrow \mathbb{R}^n \\ t \longmapsto y(t - c) \end{cases}$$

est une courbe intégrale (i.e solution maximale de A). Elle a la même trajectoire que y à translation près.

Proposition Soit (A) une équation autonome définie par une fonction f. Supposons que f est localement lispchitzienne. Soient :

$$y_1 : I_1 \longrightarrow \mathbb{R}^n$$
$$y_2 : I_2 \longrightarrow \mathbb{R}^n$$

deux solutions maximales de (A) telles que

$$\exists t_1, t_2 \in I_1, I_2, \quad y_1(t_1) = y_2(t_2)$$

alors y_1 et y_2 sont égales à translation près.

Définition (**Orbite**) . Toutes les courbes intégrales y d'une équation autonome f vérifiant le même problème de Cauchy ont la même trajectoire. Cette trajectoire est appelée **orbite**. On appelle **portrait de phase** la partition de Ω en orbites.

Propriété (**Orbites**) . Deux orbites sont soient disjointes soient égales.

Proposition Soit y une courbe intégrale d'une équation autonome (A). On a :

- Si y admet une limite l à l'infini, alors l est un point stationnaire de.
- Si y n'est pas injective alors y est globale et périodique.

7.4.2 Intégrales Premières et Systèmes Hamiltoniens

Intégrales Premières et Courbes de Niveau

Définition (**Intégrale Première**) . Soit $(E) : y' = f(y)$ une équation autonome où $f : \Omega \xrightarrow{\text{L.L}} \mathbb{R}^n$. Une fonction $E : \Omega \xrightarrow{\mathcal{C}^1} \mathbb{R}$ telle que pour toute solution y de (A), on ait $E \circ y = c \in \mathbb{R}$

i.e pour toute courbe intégrale $(y, I), \forall t \in I, \quad \dfrac{\partial}{\partial t} E \circ y = 0$

i.e $\quad \langle \nabla E, f \rangle = 0$

i.e E est constante le long des orbites

est appelée **intégrale première** de (A).

En physique, les intégrales premières représentent les quantités invariances de systèmes dynnamiques représentés par des équations autonomes. Par exemple, dans un système fermé, un intégrale première peut représenter l'énergie du système.

Proposition (Lien équation autonome/système dynnamique) Dans la suite du chapitre, nous prendrons l'habitude d'écrire les équation autonomes sous forme de systèmes dynnamiques (système d'équation différentielles). Détaillons comment passer de l'un à l'autre dans la cas quelconque.
Soit $(A) : y' = f(y)$ une équation autonome où $f : \Omega \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^n$ et $y = (y_1, \ldots, y_n)$. Le système dynnamique associé est simplement l'équation différentielle noté :

$$\frac{\partial}{\partial t} \begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix} = \begin{pmatrix} f_1(y_1(t), \ldots, y_n(t)) \\ \vdots \\ f_n(y_1(t), \ldots, y_n(t)) \end{pmatrix}$$

Que l'on écrit plus généralement avec une lettre différente pour chaque composante de y.

Définition (**Courbe de Niveau**) . Soit $(A) : y' = f(y)$ une équation autonome où $f : \Omega \longrightarrow \mathbb{R}^n$. Soit $E : \mathbb{R}^n \longrightarrow \mathbb{R}$ une intégrale première de (A). Les courbes de niveaux de E sont les ensembles de points $y \in \mathbb{R}^n$ tels que

$$E(y) = c \in \mathbb{R}$$

Exemple (Oscillateur Harmonique) Soit l'équation différentielle autonome suivante :

$$(A) : \begin{cases} x'(t) = v(t) \\ v'(t) = -\omega^2 x(t) \end{cases} \qquad \forall (t, \omega) \in \mathbb{R} \times \mathbb{R}$$

ici, x représente la position d'un oscillateur et v sa vitesse. Une intégrale première de ce système est de la forme :

$$E : \begin{cases} R \times \mathbb{R} \xrightarrow{\mathcal{C}^1} \mathbb{R} \\ (x, y) \longmapsto E(x, y) \end{cases}$$

Multiplions (A) par v. On a alors $\forall t \in \mathbb{R}$:

$$\begin{aligned} v(t)v'(t) &= -\omega^2 x(t)v(t) \\ &= -\omega^2 x(t)x'(t) \\ \iff v(t)v'(t) + \omega^2 x(t)x'(t) &= 0 \end{aligned}$$

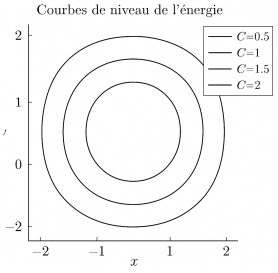
Supposons que E soit de la forme : $\forall t \in \mathbb{R}, E : (x(t), v(t)) \longmapsto E(x(t), v(t))$. D'après la règle de la chaîne, on a $\forall t \in \mathbb{R}$:

$$\begin{aligned} \frac{\partial}{\partial t} E(x(t), v(t)) = \frac{\partial x(t)}{\partial t}\frac{\partial E}{\partial x} + \frac{\partial v(t)}{\partial t}\frac{\partial E}{\partial y} &= x'(t)\frac{\partial E}{\partial x} + v'(t)\frac{\partial E}{\partial y} \\ \implies \frac{\partial E}{\partial x} = \omega^2 x(t) \quad \text{et} \quad \frac{\partial E}{\partial y} &= v(t) \end{aligned}$$

On peut donc en conclure que :

$$\forall t \in \mathbb{R}, \quad E(x(t), v(t)) = \frac{1}{2}\omega^2 x^2(t) + \frac{1}{2}v^2(t)$$

Réciproquement, par construction, $\frac{\partial E}{\partial t} = 0$. Traçons maintenant quelques courbes de niveau :



Ces courbes permettent de représenter l'évolution du mouvement de l'oscillateur dans le plan. Ici, plus l'énergie C est élevée, plus le rayon des cercles en grand, ce qui correspond à une plus grande amplitude.
On peut aussi les voir comme l'ensemble des orbites pour lesquelles l'énergie est égale à C.

Propriété (**Orbites et Ensemble de Niveau**) . Toute orbite est incluse dans un ensemble de niveau.

Ainsi, les intégrales premières d'équations autonomes (A) permettent de donner une idée de la forme des orbites formés par les solutions de (A).

Systèmes Hamiltoniens

Il existe un cas particulier des équations autonomes appelés systèmes hamiltoniens. Ils offrent une approche plus simple pour l'étude des systèmes dynnamiques conservatifs, où certaines quantités sont conservées au cours du temps.

Définition (**Système Hamiltonien**) . Un système hamiltonien est une équation différentielle autonome (ou système d'équations différentielles autonomes associé) pouvant être représentée par un fonction appelée **hamiltonien**. Plus formellement, soit $(A) : y' = f(y)$ une équation autonome telle que $f : \Omega \longrightarrow \mathbb{R}^{2n}$. On dit que (A) est un système hamiltonien s'il existe une fonction :

$$H : \begin{cases} \Omega \xrightarrow{\mathcal{C}^2} \mathbb{R} \\ (q,p) \longmapsto H(q,p) \end{cases} \qquad \text{où } q, p \in \mathbb{R}^n$$

telle que $\forall (q,p) \in \Omega$ on ait :

$$\boxed{f(q,p) = \left(\frac{\partial H}{\partial p}, -\frac{\partial H}{\partial q} \right)}$$

Cela revient donc à dire que (A) peut s'écrire de la forme suivante :

$$(A) : \begin{cases} p'(t) = -\partial_q H(q,p) \\ q'(t) = \partial_p H(q,p) \end{cases}$$

Remarque Tout système de dimension impaire n'est donc pas hamiltonien.

Exemple (Oscillateur Harmonique) Prenons comme exemple le cas classique d'un oscillateur harmonique gouverné par l'équation différentielle :

$$(A) : mx'' + kx = 0$$

où :

- m est la masse de l'oscillateur
- k est la constante de raideur du ressort
- $x(t)$ est la position de l'oscillateur à l'instant t.

Représentons cette équation autonome sous la forme d'un système hamiltonien. Posons $Y(t) = (x(t), y(t))$ on a alors $Y'(t) = (x'(t), y'(t))$. Soit la fonction suivante :

$$f : \begin{cases} \mathbb{R}^2 \longrightarrow \mathbb{R}^2 \\ (x,y) \longmapsto (y, -\frac{k}{m}x) \end{cases}$$

Cela nous donne que :

$$\begin{aligned} (A) &\iff Y'(t) = f(Y(t)) \\ &\iff \begin{cases} x'(t) = y(t) \\ y'(t) = -\frac{k}{m}x \end{cases} \iff \begin{cases} x'(t) = \frac{1}{m}p(t) \\ p'(t) = -kx(t) \end{cases} \end{aligned}$$

où $p(t) = mx'(t)$
Maintenant, posons :

$$H : \begin{cases} \mathbb{R}^2 \longrightarrow \mathbb{R} \\ (x,p) \longmapsto \frac{p^2}{2m} + \frac{1}{2}kx^2 \end{cases}$$

H est $\mathcal{C}^\infty$ par composantes donc calculons ses dérivées partielles :

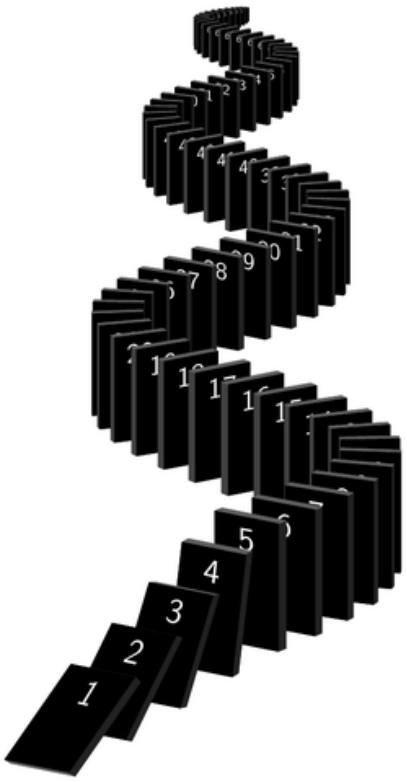
$$\frac{\partial H}{\partial p} = \frac{p}{m} = x'(t) \quad \frac{\partial H}{\partial x} = kx(t) = -p'(t)$$

Donc (A) est bien un système hamiltonien.

Propriété (**Hamiltonien et intégrale première**) . Tout hamiltonien est une intégrale première. Autrement dit, un hamiltonien est un cas particulier d'intégrale première.

Proposition (Condition Nécessaire) Soit $(A) = y' = f(y)$ un système hamiltonien, alors $\div f = 0$. En dimension 2, on a équivalence.

Probabilités et Statistiques



Chapitre 1

Espaces Probabilités et Mesures

Contents

1.1 Univers et espace probabilisé

Introduisons les concepts fondamentaux des probabilités, les univers et les espaces probabilisés.

Définition (**Univers**) . On appelle univers Ω pour une expérience aléatoire, l'ensemble de toutes les issues (situations finales) possibles de cette expérience aléatoire. Chaque élément $\omega \in \Omega$ représente une **issue** de cette expérience aléatoire.

Exemple
- Pour une expérience aléatoire de lancer de dé, il existe 6 issues possibles correspondant aux 6 faces du dé. On a donc $\Omega = \{1, 2, 3, 4, 5, 6\}$.
- Si on pioche un boule dans une urne contenant une boule rouge et deux boules noires, on a $\Omega = \{\text{rouge}, \text{noir}\}$.

A partir d'un univers, on peut définir la notion d'espace probabilisé. Plus complexe, la définition nécessaire les prérequis du cours d'intégration et de théorie de la mesure.

Définition (**Espace Probabilisé**) . Un espace probabilisé est un **triplet** $(\Omega, \mathcal{F}, \mathbb{P})$ où :
- Ω est un univers.
- $\mathcal{F}$ est une tribu (σ-algèbre) sur Ω.
- $\mathbb{P}$ est une mesure de probabilité sur $\mathcal{F}$ (voir plus loin).

1.2 Évènements, Issues et Mesure de Probabilité

Définition (**Évènement**) . Soit Ω un univers. On définit un évènement de Ω comme un sous-ensemble $A \subseteq \Omega$.

Remarque Comme définit au début, les **issues** $\omega \in \Omega$ correspondent à des résultats élémentaires de l'expérience aléatoire, à ne pas confondre avec les évènements. Dans notre expérience

de lancer de dé, $\{2\} \in \Omega$ est **l'issue** correspondant à "obtenir un 2" et $A = \{1, 2\} \subset \Omega$ est **l'évènement** correspondant à "le résultat est inférieur ou égal à 2".

Par construction, $\mathcal{F}$ contient donc tous les évènements et issues possibles de l'expérience aléatoire. Elle est dont "plus complète" que Ω.

Définition (**Mesure de Probabilité**) . Soit un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$. Une mesure de probabilité $\mathbb{P}$ sur $\mathcal{F}$ est une mesure (au sens de la théorie de la mesure) qui vérifie :

1. $\mathbb{P} : \mathcal{F} \longrightarrow [0, 1]$
2. $\mathbb{P}(\Omega) = 1$

1.3 Variable Aléatoire

Pour pouvoir quantifier des calculs de probabilités ou ce que nous appellerons plus des lois, nous devons définir les variables aléatoires.

Définition (**Variable Aléatoire**) . Une variable aléatoire est une fonction mesurable qui associe une valeur numérique à chaque issue d'un espace probabilisé.

Plus formellement, une variable aléatoire X sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ est une fonction $X : \Omega \longrightarrow \mathbb{R}$ telle que, pour tout ensemble $B \in \mathcal{B}_{\mathbb{R}}$ (tribu borélienne), on ait $X^{-1}(B) \subset \mathcal{F}$.

La mesurabilité d'une variable aléatoire permet donc garantir que les évènements associés aux valeurs de la variable aléatoire sont bien mesurables par la mesure de probabilité.

Proposition Maintetant que nous avons définit formellement le concept de variable aléatoire, on peut lier cette définition à celle des évènements. En effet, une variable aléatoire est une fonction mesurable sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ telle que :

$$X : \Omega \longrightarrow \mathbb{R} \quad \forall B \in \mathcal{B}_{\mathbb{R}}, X^{-1}(B) \subset \mathcal{F}$$

On peut alors caractériser un évènement A comme la préimage d'un sous-ensemble de $\mathbb{R}$ par X de la façon suivante.

$$A = X^{-1}(B) \text{ pour un certain } B \in \mathcal{B}_{\mathbb{R}}$$

Chapitre 2

Variables Aléatoires Réelles Discrètes

Contents

2.1 Variable Aléatoire

2.1.1 Univers et Evènement

Définition (**Univers**) . Soit Ω un ensemble. On dit que Ω est l'univers d'une expérience aléatoire s'il représente l'ensemble des issues possibles pour cette expérience.

Définition (**Variable Aléatoire**) . Soit Ω un univers d'une expérience aléatoire. Une variable aléatoire X sur Ω est une application $X : \Omega \longrightarrow \mathbb{R}$. On note $X(\Omega)$ l'ensemble des valeurs prises par X.

Remarque Si Ω est un ensemble fini ou dénombrable, on dit que $X : \Omega \longrightarrow \mathbb{R}$ est une variable aléatoire réelle discrète.

Exemple Soit l'expérience aléatoire du lancer d'une pièce de monnaie non truquée. On a alors $\Omega = \{$ Pile , Face $\}$ et $X : \Omega \longrightarrow \mathbb{R}$.

Une variable aléatoire associe donc des issues à des réels. Mais un événement d'une expérience aléatoire peut être caractérisée par plusieurs issues. Par exemple si on lance un dé à 6 faces, il y a 6 issues possibles correspondant à chaque face du dé. Mais on peut définir un événement telle que "le nombre obtenu est supérieur à 2". Dans ce cas ci, plusieurs issues de l'expérience aléatoire peuvent correspondre au même événement.

Définition (**Evènement d'une expérience aléatoire**) . Soit X une variable aléatoire réelle définie sur un univers Ω. On appelle évènement $[X = x]$ de l'expérience aléatoire l'ensemble des issues possibles correspondant à cet évènement $xin\mathbb{R}$. Autrement dit :

$$[X = x] = \{w \in \Omega \mid X(\omega) = x\} = X^{-1}(x)$$

Proposition Soit X une variable aléatoire réelle définie sur un univers Ω. L'ensemble des évènements possibles pour cette expérience aléatoire forme un système complet d'évènements. Autrement dit :

$$\sum_{x \in X(\Omega)} \mathbb{P}([X = x]) = 1$$

2.1.2 Loi d'une variable aléatoire

Définition (**Loi**) . Soit X une variable aléatoire réelle. On appelle loi de X la donnée de toutes les probabilités $\mathbb{P}(X = x)$ pour tout $x \in X(\Omega)$.

Pour donner la loi d'une variable aléatoire, il faut d'abord déterminer le support de la variable aléatoire puis en suite calculer la probabilité de chaque issue. On note le résultat dans un tableau pour plus de praticité.

Exemple Soit l'exprérience aléatoire du lancer d'un dé à 6 faces non truqué. On a :

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

Nous nous trouvons dans une situation d'équiprobabilité d'où :

$$\forall x \in X(\Omega), \quad \mathbb{P}(X = x) = \frac{1}{6}$$

D'où le tableau suivant :

Ω	1	2	3	4	5	6
$\mathbb{P}(X = x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

2.2 Espérance, Variance et écart-type

2.2.1 Espérance et propriétés

Définition (**Espérance**) . Soit $X : \Omega \longrightarrow \mathbb{R}$ une variable aléatoire. On appelle espérance l'application $\mathbb{E} : \mathcal{F}(\Omega) \longrightarrow \mathbb{R}$ qui calcule la moyenne de X pondérée par les valeurs qu'elle prend. Plus formellement :

$$\boxed{\mathbb{E}(X) = \sum_{x \in X(\Omega)} x\mathbb{P}(X = x)}$$

Propriété (**Espérance**) . L'espérance est une fonction linéaire. Autrement dit, pour toutes variables aléatoires X, Y sur un univers Ω, et pour tout $a, b \in \mathbb{R}$, on a :

$$\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y) \quad \mathbb{E}(aX+b) = a\mathbb{E}(X) + b$$

Lors d'une expérience aléatoire, par exemple un jeu d'argent l'espérance représente le gain moyen d'un joueur par partie s'il joue un grand nombre de fois. Son signe permet de savoir si le jeu est dit équitable (autant de chances de gagner que de perdre).

Il peut souvent arriver que l'on veuille appliquer une fonction à notre variable aléatoire. Un théorème nous permet alors simplement de calculer l'espérance de cette "nouvelle" variable aléatoire.

Théorème (**Transfert**) . Soit X une variable aléatoire sur un univers Ω et $g : \mathbb{R} \longrightarrow \mathbb{R}$ une application. L'espérance de la variable aléatoire $g(X) : \Omega \longrightarrow \mathbb{R}$ est l'application $\mathbb{E}(g(X)) : \mathcal{F}(\Omega) \longrightarrow \mathbb{R}$ telle que :

$$\boxed{\mathbb{E}(g(X)) = \sum_{x \in X(\Omega)} g(x)\mathbb{P}(X = x)}$$

2.2.2 Variance et écart-type

Définition (**Variance et écart-type**) . Soit X une variable aléatoire sur un univers Ω. On appelle variance l'application $\mathbb{V} : \mathcal{F}(\Omega) \longrightarrow \mathbb{R}$ telle que :

$$\boxed{\mathbb{V}(X) = \sum_{x \in X(\Omega)} (x - \mathbb{E}(X))^2 \mathbb{P}(X = x)}$$

De même, on appelle écart type l'application $\sigma : \mathcal{F}(\Omega) \longrightarrow \mathbb{R}$ telle que :

$$\boxed{\sigma(X) = \sqrt{\mathbb{V}(X)}}$$

La variance permet de mesurer la dispersion de la variable aléatoire autour de son espérance.

Remarque Les notions d'espérance, variance et écart-type sont définies par des sommes potentiellement infinies. Il se peut donc que dans le cas de variables aléatoires définies sur des univers infinies, leur espérance, variance et écart-type n'existent pas. Une étude de convergence de la somme est donc judicieuse.
En revanche pour les variables aléatoires finies (definies sur un univers fini) ces notions sont toujours bien définies.

Théorème (**Formule de König-Huygens**) . Soit X une variable aléatoire sur un univers Ω fini. On a alors :

$$\boxed{\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2}$$

Remarque Le calcul de la variance d'une variable aléatoire réelle finie est donc assez facile quand on le met en relation avec la formule de König-Huygens et la formule du transfert...

A partir de toutes ces formules, on peut en déduire quelques propriétés sympatiques sur la variance :

Propriété (**Variance**) . Soit X une variable aléatoire finie et $a, b \in \mathbb{R}$, on a :

$$\mathbb{V}(aX+b) = a^2\mathbb{V}(X) \quad \mathbb{V}(X+b) = \mathbb{V}(X)$$

La variance est donc assez similaire à une forme quadratique et est invariante par translation.

Théorème (**Inégalité de Markov**) . Si X est une variable aléatoire réelle discrète **positive** ou nulle sur Ω d'espérance $\mathbb{E}(X)$, alors :

$$\forall a \in]0, +\infty[\quad P(X \geqslant a) \leqslant \frac{\mathbb{E}(X)}{a}$$

En l'appliquant à la variable aléatoire $(X - \mathbb{E}(X))^2$ et en remarquant que son espérance vaut σ^2 et l'égalité des événements $|X - \mathbb{E}(X)| \geqslant \alpha$ et $(X - \mathbb{E}(X))^2 \geqslant \alpha^2$ on obtient :

$$\forall \alpha > 0, \quad P(|X - \mathbb{E}(X)| \geqslant \alpha) \leqslant \frac{\sigma^2}{\alpha^2}$$

2.3 Principales Lois

Abordons en détail maintenant quelques lois usuelles à connaître sur le bout des doigts. Ces lois permettent de modéliser la plupart des expériences aléatoires.

2.3.1 Loi Uniforme

Définition (**Loi Uniforme**) . Soit $n \in \mathbb{N}^*$. On dit qu'un variable aléatoire X suit une **loi uniforme** sur $[\![1, n]\!]$ lorsque son support est $X(\Omega) = [\![1, n]\!]$ et chaque issue a la même probabilité de se produire. Autrement dit :

$$\forall x \in [\![1, n]\!], \quad P(X = x) = \frac{1}{n}$$

On note alors $X \sim \mathcal{U}([\![1, n]\!])$.

Proposition Soit $X \sim \mathcal{U}([\![1, n]\!])$ alors l'espérance et la variance de X sont de la forme :

$$\mathbb{E}(X) = \frac{n+1}{2} \quad \text{et} \quad \mathbb{V}(X) = \frac{n^2-1}{12}$$

2.3.2 Loi de Bernouilli

Définition (**Loi de Bernouilli**) . Une variable aléatoire X suit une **loi de Bernoulli** de paramètre $p \in]0, 1[$ si il n'existe que deux issues possibles $X(\Omega) = \{0, 1\}$ telles que :

$$\mathbb{P}(X = 1) = p \quad \text{et} \quad \mathbb{P}(X = 0) = 1 - p$$

On note alors $X \sim \mathcal{B}(p)$.

Proposition Soit $X \sim \mathcal{B}(p)$ alors l'espérance et la variance de X sont de la forme :

$$\mathbb{E}(X) = p \quad \text{et} \quad \mathbb{V}(X) = p(1-p)$$

2.3.3 Loi Binomiale

L'expérience aléatoire consistant à répéter $n \in \mathbb{N}$ fois une expérience de Bernoulli de paramètre $p \in]0,1[$ de manière indépendante est appelée **schéma de Bernoulli** de paramètres n et p.

Définition (**Loi Binomiale**) . La variable aléatoire X égale au **nombre de succès** d'un schéma de Bernoulli suit une loi binomiale de paramètres n et p.
On note alors $X \sim \mathcal{B}(n,p)$.

Proposition Soit $X \sim \mathcal{B}(n,p)$. On a alors $X(\Omega) = [\![0,n]\!]$ et pour tout $k \in \mathbb{N}$ tel que $0 \leqslant k \leqslant n$, la probabilité d'obtenir k succès est donnée par :

$$\mathbb{P}(X=k) = \binom{k}{n} \times p^k \times (1-p)^{n-k}$$

Proposition Soit $X \sim \mathcal{B}(n,p)$ alors l'espérance et la variance de X sont de la forme :

$$\mathbb{E}(X) = np \quad \text{et} \quad \mathbb{V}(X) = np(1-p)$$

2.3.4 Loi géométrique

Définition (**Loi géométrique**) . Une variable aléatoire X suit une loi géométrique de paramètre $p \in]0,1[$ lorsque $X(\Omega) = \mathbb{N}^*$ et que

$$\forall k \in \mathbb{N}^*, \quad \mathbb{P}(X = k) = p(1-p)^{k-1}$$

On note alors $X \sim \mathcal{G}(p)$.

Une loi géométrique représente le temps d'attendre du premier succès d'un espérience de Bernoulli. Autrement dit X est le rang de l'épreuve ayant mené au premier succès.

Proposition Soit $X \sim \mathcal{G}(p)$ où $p \in]0,1[$ l'espérance et la variance de X sont de la forme :

$$\mathbb{E}(X) = \frac{1}{p} \quad \text{et} \quad \mathbb{V}(X) = \frac{1-p}{p^2}$$

2.3.5 Loi de Poisson

Définition (**Loi de Poisson**) . Une variable aléatoire X suit une loi de Poisson de paramètre $\lambda > 0$ lorsque $X(\Omega) = \mathbb{N}$ et

$$\forall k \in \mathbb{N}, \quad P(X = k) = e^{-\lambda} \times \frac{\lambda^k}{k!}$$

On note alors $X \sim \mathcal{P}(\lambda)$.

Une loi de poisson représente le nombre moyen d'évènements produits au cours d'un intervalle de temps donné ou d'une quantité donnée. Par exemple, elle peut modéliser le nombre moyen de voitures qui sont passées par un péage sur une journée ou le nombre moyen de fautes de frappes produits sur une page de texte.

Proposition Soit $X \sim \mathcal{P}(\lambda)$ alors l'espérance et la variance de X sont de la forme :

$$\mathbb{E}(X) = \lambda \quad \text{et} \quad \mathbb{V}(X) = \lambda$$

Chapitre 3

Variables Aléatoires Continues

Contents

On s'intéresse maintenant à des variables aléatoires qui prennent des valeurs réelles mais pas forcément en nombre fini ou dénombrable. Il est donc nécessaire de definir une probabilité sur $\mathbb{R}$ telle que la probabilité des singletons soit nulle.
Pour cela, nous allons très fortement nous appuyer sur l'intégrale de Lebesgue et la théorie de la mesure.

3.1 Tribu Borélienne et Mesure

3.1.1 Borélien vous dites ?

Définition (**Tribu**) . Soit X un ensemble. Une tribu sur X est une partie $\mathcal{B} \subseteq \mathcal{P}(X)$ telle que :

i) $\emptyset \in \mathcal{B}$

ii) $\mathcal{B}$ est stable par complémentaire

iii) $\mathcal{B}$ est stable par union dénombrable

Un élément de $\mathcal{B}$ est appelé **partie mesurable**. On appelle le couple $(X, \mathcal{B})$ un espace mesurable.

Remarque Si on prend $X = \mathbb{R}$, on appelle alors sa tribu la **tribu borélienne**. C'est la plus petite tribu contenant tous les ouverts de $\mathbb{R}$. On la note $\mathcal{B}_{\mathbb{R}}$.

Proposition Soit A un ensemble et X un ensemble de parties de A. Il existe une plus petite tribu sur A qui contienne X. On l'appelle **tribu engendrée** par X, notée $\sigma(C)$.
On définit ainsi la **tribu borélienne** sur $\mathbb{R}$ la tribu engendrée par les intervalles ouverts de $\mathbb{R}$. Les éléments de la tribu sont appelés les boréliens.

3.1.2 Mesure

Définition (**Mesure**) . Soit X un ensemble muni d'une tribu $\mathcal{B}$. On appelle mesure sur $\mathcal{B}$ toute application

$$\mu : \mathcal{B} \longrightarrow \overline{R_+}$$

telle que :

i) $\mu(\emptyset) = 0$

ii) $\forall (A_n)_{n\in\mathbb{N}}$ suite de parties mesurables deux à deux disjointes :

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in N} \mu(A_n)$$

On appelle le triplet $(X, \mathcal{B}, \mu)$ un espace mesuré.

Théorème (**Mesure de Lebesgue**) . Il existe une unique mesure λ sur $(\mathbb{R}, \mathcal{B}_\mathbb{R})$ appelée mesure de Lebesgue telle que :

$$\lambda : \begin{cases} \mathcal{B}_\mathbb{R} \longrightarrow \overline{\mathbb{R}_+} \\ ]a, b] \longmapsto b - a \end{cases}$$

On a $\lambda(\mathbb{R}) = \infty$.

3.1.3 Partie Négligeable et propriété vraie presque partout

Définition (**Partie Négligeable**) . Soit $(X, \mathcal{B}, \mu)$ un espace mesuré. On appelle partie négligeable de X toute partie A mesurable telle que $\mu(A) = 0$.

Définition (**Propriété vraie presque partout**) . Soit $(X, \mathcal{B}, \mu)$ un espace mesuré. Soient A une partie mesurable de X et $\mathcal{P}(A)$ une propriété sur A. On dit que $\mathcal{P}(A)$ est vraie presque partout ssi :

$$\mu(\{x \in A \mid \neg\mathcal{P}(A)\}) = 0$$

Autrement dit, une propriété sur une partie mesurable est vraie presque partout ssi l'ensemble des points où elle est fausse est négligeable.

3.2 Variables aléatoires continues

Définition (**Espace Probabilisé**) . Soit une expérience aléatoire. On appelle le triplet $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé si :

i) Ω est un ensemble d'évènements possibles

ii) $\mathcal{F}$ est une tribu des évènements mesurables.

iii) $\mathbb{P} : \mathcal{F} \longrightarrow [0, 1]$ est une mesure définie sur l'espace mesurable $(\Omega, \mathcal{F})$.

On appelle $\mathbb{P}$ une mesure de probabilité sur $(\Omega, \mathcal{F})$.

Définition (**Variable Aléatoire**) . Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé et $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ un espace mesuable. Une variable aléatoire de Ω vers $\mathbb{R}$ toute fonction mesurable $X : \Omega \longrightarrow \mathbb{R}$.

Définition (**Loi de probabilité**) . Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé et X une variable aléatoire sur Ω. La loi de X est la mesure image de $\mathbb{P}$ par X. Autrement dit :

$$\forall B \in \mathcal{F}, \quad \mathbb{P}(X \in B) = \mathbb{P}(\{\omega \in \Omega \mid X(\omega) \in \Omega\})$$

Définition (**Fonction de répartition**) . On appelle fonction de répartition de la variable aléatoire X la fonction

$$F_X : \begin{cases} \mathbb{R} & \longrightarrow [0,1] \\ a & \longmapsto P(X \in]-\infty, a[) \end{cases}$$

F_X est une fonction croissante admettant la limite 0 en $-\infty$ et la limite 1 en $+\infty$ et elle est continue à droite.

Définition (**Continuité d'une variable aléatoire réelle**) . Une variable aléatoire réelle X est dite continue si il existe une fonction F_X intégrable sur $\mathbb{R}$ positive ou nulle et continue par morceaux et telle que :

$$\forall x \in \mathbb{R}, \quad \int_{-\infty}^{X} f_X(t)dt = F_X(X)$$

f_X est alors la **densité** de la variable X.

Principales Lois

Définition (**Loi uniforme**) . La variable aléatoire X continue dont la densité est constante sur un intervalle borné I et nulle en dehors est appelée la loi uniforme sur l'intervalle I notée $\mathcal{U}(I)$. Ainsi, on a :

$$I = [a,b] \subset \mathbb{R}, \quad f_X(t) = \begin{cases} 0 \text{ si} & \text{t } \notin [a,b] \\ \dfrac{1}{b-a} & \text{si } t \in [a,b] \end{cases} \quad \text{et} \quad F_X(x) = \begin{cases} 0 & \text{si } x \leq a \\ \dfrac{x-a}{b-b} & \text{si } a \leq x \leq b \\ 1 & \text{si } x \geq b \end{cases}$$

Définition (**Loi Normale**) . La loi normale (ou gaussienne ou de Laplace-Gauss) notée $\mathcal{N}(\mu, \sigma^2)$ de moyenne μ et d'écart-type σ est la loi continue sur $\mathbb{R}$ de densité :

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

de fonction de répartition :

$$F_{\mu,\sigma^2}(a) = \int_{-\infty}^{a} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Définition (**Loi exponentielle**) . La loi exponentielle $\mathcal{E}(\lambda)$ de paramètre $\lambda > 0$ est la loi de densité nulle sur $\mathbb{R}_-$ et égale à $\lambda e^{-\lambda x}$ sur $\mathbb{R}_+$. Sa fonction de répartition est la suivante :

$$F(t) = \begin{cases} 0 \text{ si } t \leq 0 \\ 0^t \lambda e^{-\lambda x} dx = [-e^{-\lambda x}]_0^t = 1 - e^{-\lambda t} \text{ si } t \geq 0 \end{cases}$$

Propriétés des variables aléatoires continues

Définition (**Espérance**) . Soit X une variable aléatoire continue réelle absolument continue de densité f_X. On appelle espérance de X le nombre

$$E(X) = \int_{\mathbb{R}} t f_X(t) dt$$

Si cette espérance n'est pas absolument convergente, on dit que X n'a pas d'espérance.

Théorème (**Formule de Transfert**) . Soit X est une variable aléatoire réelle absolument continue de densité f_X et si φ est une fonction $\mathcal{C}^1$, alors $\varphi(X)$ est une variable aléatoire réelle. Si elle admet un espérance, on a alors :

$$E(\varphi(X)) = \int_{\mathbb{R}} \varphi(t) f_X(t) dt$$

Définition (**Variance**) . Soit X une variable aléatoire admettant une espérance. La variance de X est le nombre :

$$V(X) = E((E - E(X))^2) = \int_{\mathbb{R}} (t - E(X))^2 f(t) dt$$

Toujours comme chez les variables aléatoires réelles discrètes, l'acrt type de X est la racine carrée de $V(X)$ si $V(X)$ existe.

Lois conjointes continues

Définition (**Vecteur Aléatoire**) . Une vecteur aléatoire à n composantes est une al=pplication V d'un espace probabilisé $(\Omega, \mathcal{A}, P)$ dans $\mathbb{R}^n$ telle que l'image réciproque $V^{-1}(B)$ de tout borélien de $\mathbb{R}^n$ soit un élément de la tribu de Ω.

Théorème (**Fubini**) . Si $f : \mathbb{R}^2 \to \mathbb{R}$ est intégrable sur $[a, b] \times [c, d]$ alors pour presque tout $x \in [a, b]$, la fonction partielle $y \mapsto f(x, y)$ est intégrable sur $[c, d]$ et

$$\iint_{[a,b]\times[c,d]} f(x,y)dx\ dy = \int_a^b \left(\int_c^d f(x,y)dy \right) dx$$

Définition (**Densité Conjointe**) . Soit $f : \mathbb{R}^2 \to \mathbb{R}$ une application intégrable telle que $f \geq 0$ et

$$\iint_{\mathbb{R}^2} f(x,y)dx\ dy = 1$$

Alors la loi $\mathcal{L}(X, Y)$ du couple (X, Y) est absolument continue de densité conjointe f si, pour tout borélien B, on a :

$$P((X,Y) \in B) = \iint_B f(x,y)dx\ dy$$

Démonstrations

- intersection quelconque de tribu est une tribu - limite et continuité de la fonction de répartition

Chapitre 4

Vecteurs Aléatoires

Définition (**Vecteur Aléatoire**) . Une vecteur aléatoire à n composantes est une al=pplication V d'un espace probabilisé $(\Omega, \mathcal{A}, P)$ dans $\mathbb{R}^n$ telle que l'image réciproque $V^{-1}(B)$ de tout borélien de $\mathbb{R}^n$ soit un élément de la tribu de Ω.

Théorème (**Fubini**) . Si $f : \mathbb{R}^2 \to \mathbb{R}$ est intégrable sur $[a, b] \times [c, d]$ alors pour presque tout $x \in [a, b]$, la fonction partielle $y \mapsto f(x, y)$ est intégrable sur $[c, d]$ et

$$\iint_{[a,b]\times[c,d]} f(x,y)dx\ dy = \int_a^b \left(\int_c^d f(x,y)dy \right) dx$$

Définition (**Densité Conjointe**) . Soit $f : \mathbb{R}^2 \to \mathbb{R}$ une application intégrable telle que $f \geq 0$ et

$$\iint_{\mathbb{R}^2} f(x,y)dx\ dy = 1$$

Alors la loi $\mathcal{L}(X, Y)$ du couple (X, Y) est absolument continue de densité conjointe f si, pour tout borélien B, on a :

$$P((X,Y) \in B) = \iint_B f(x,y)dx\ dy$$

Si deux densité sont égales presque partout, alors elles définissent la même loi de probabilité. Donc si on modifie une densité sur un ensemble négligeable, elle définit toujours la même loi.

Démonstrations

- intersection quelconque de tribu est une tribu - limite et continuité de la fonction de répartition

Chapitre 5

Fonctions Génératrices

Définition Générale

On considère une variable aléatoire discrète entière positive non nulle X qui prend ses valeurs dans $\mathbb{N}$ et pour tout n dans $\mathbb{N}$ on pose $p_n = P(X = n)$, on a donc, d'après la définition d'une probabilité $\sum_{n=0}^{\infty} p_n = 1$

Définition (**Fonction Génératrice**) . On appelle fonction génératrice de X la fonction :

$$\boxed{g_X(z) = \sum_{n=0}^{\infty} P(X = n) z^n}$$

C'est une série entière de la variable $z \in \mathbb{C}$.

Remarque Le rayon de convergence cette série est supérieur à 1.

Propriétés

- g_X est continue sur $\overline{D(0,1)}$ et $\mathcal{C}^\infty$ sur $D(0,1)$.
- pour $|z| > 1$ on a $g_X(z) = E(z^X)$
- la fonction génératrice caractérise la loi

$$\text{i.e. } g_X = g_Y \text{ sur un voisinnage de } 0 \Longrightarrow \mathcal{L}(X = \mathcal{L}(Y)$$

- si $E(X)$ existe alors $E(X) = g_X(1)$
- si $E(X^k), k \in \mathbb{N}$ existe alors $g_X^{(k)}(z) = E(X(X-1)(X-2)\dots(X-k+1))$
  En particulier, si $E(X^2)$ existe, alors

$$g_X''(1) = E(X(X-1)) = E(X^2) - E(X)$$

Donc $V(X) = g_X''(1) + g_X'(1) - (g_X'(1))^2$

Somme de variables aléatoires entières positives indépendantes

Théorème () . Si X et Y sont indépendantes, alors

$$g_{X+Y} = g_X g_Y$$

Démonstration On a :

$$g_X(z) = \sum_{n\in\mathbb{N}} p_n z^n, \quad g_Y(z) = \sum_{n\in\mathbb{N}} q_n z^n, \quad g_{X+Y}(z) = \sum_{n\in\mathbb{N}} r_n z^n$$

D'après la formule des probabilités totales :

$$r_n = P(X+Y=n) = \sum_{k=0}^{n} P(X=k)P(Y=n-k) = \sum_{k=0}^{n} p_k q_{n_k}$$

et

$$\begin{aligned} g_{X+Y}(z) &= \sum_{n=0}^{\infty}\sum_{k=0}^{n} p_k q_{n-k} z^n = \sum_{n=0}^{\infty}\sum_{k=0}^{n} p_k q_{n-k} z^k z^{n-k} \\ &= \left(\sum_{i=0}^{\infty} p_i z^i\right)\left(\sum_{j=0}^{\infty} q_j z^j\right) \\ &= g_X g_Y \end{aligned}$$

Chapitre 6

Convergences

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires indépendantes suivant la même loi. On considère $Y_n = \frac{1}{n}(\sum_{k=1}^{n} X_k)$. On suppose que $V(X_n)$ est finie.

$$E(Y_n) = \frac{1}{n}\sum_{k=1}^{n} E(X_k) = E(X_n)$$

et

$$\begin{aligned} V(Y_n) &= V\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right) = \frac{1}{n^2} V\left(\sum_{k=1}^{n} X_k\right) \\ &= \frac{1}{n^2}\sum_{k=1}^{n} V(X_k) = \frac{1}{n} V(X_n) \end{aligned}$$

on applique ensuite l'inégalité de Bienaymé-Tchebychev à (Y_n)

$$\boxed{P(|Y_n - E(X_1) \geq b) \leq \frac{V(Y_n)}{b^2} = \frac{1}{n}\frac{V(X_0)}{b^2} \underset{n\to+\infty}{\longrightarrow} 0}$$

Convergence en probabilités

Soit $(Z_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires sur un même espace Ω. Soit Z_∞ une variable sur Ω.

Définition (**Convergence en Probabilités**) . On dit que la suite $(Z_n)_{n\in\mathbb{N}}$ converge en probabilité vers Z_∞ si

$$\forall \varepsilon > 0, \quad \boxed{P(|Z_n - Z_\infty| \geq \varepsilon) \underset{n\to+\infty}{\longrightarrow} 0}$$

On notera : $Z_n \xrightarrow{P} Z_\infty$

Remarque La loi faible des grands nombres affirme que $(Y_n)_{n\in\mathbb{N}}$ converge en probabilité vers la loi quasi-certaine $E(X_1)$.

Convergence presque-sûre

Si on considère nos Z_n comme des fonctions sur Ω à valeurs dans $\mathbb{R}$, on pourrait essayer de voir à quoi correspondrait la convergence simple.

Définition (**Convergence presque-sûre**) . Soit Ω un espace probabilisé, soit $(Z_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires sur Ω. Soit Z_∞ une variable aléatoire sur Ω. On dit que (Z_n) converge presque sûrement vers Z_∞ si

$$P(Z_n(\omega) \longrightarrow (Z_n(\omega))) = 1$$

i.e. l'ensemble des $\omega \in \Omega$ pour lesquels il n'y a pas de convergence est un ensemble négligeable.

Théorème (**Loi forte des grands nombres**) . Soient $(X_n)_{n\in\mathbb{N}}$ la suite de variables aléatoires réelles continues, indépendantes et suivant la même loi. Supposons que nous connaissons son espérance $(E(X_i))$ et sa variance $V(X_1)$. Alors $Y_n = \frac{1}{n}\sum_{k=1}^{n} X_k$ converge presque sûrement vers la constante $E(X_1)$.

Convergence en moyenne quadratique

Définition () . Soit $(Z_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires réelles sur un même espace probabilisé Ω. Soit Z_∞ une variable aléatoire réelle sur Ω. On dit que (Z_n) converge en moyenne quadratique vers Z_∞ si

$$E((Z_n - Z_\infty)^2) \underset{n\to\infty}{\longrightarrow} 0$$

On note $Z_n \xrightarrow{mq} Z_\infty$

Exemple Dans le contexte de la loi des grands nombres on avait $E(Y_n) = E(X_1)$ et $V(Y_n) = \frac{1}{n}V(X_1) \underset{n\to+\infty}{\longrightarrow} 0$ donc on avait une convergence en moyenne quadratique.

Remarque Sur l'ensemble des variables aléatoires sur Ω, la variance est une forme quadratique positive dont la forme bilinéaire symétrique associée est la covariance. La convergence en moyenne quadratique esr la convergence pour cette "norme".

Proposition Si la suite $(X_n)_{n\in\mathbb{N}}$ converge en moyenne cubique vers X alors :

$$\left.\begin{array}{l}\lim\limits_{n\to\infty} E(X_n) = E(X) \\ \lim\limits_{n\to+\infty} E({X_n}^2) = E(X^2)\end{array}\right\} \Longrightarrow \lim_{n\to+\infty} V(X_n) = V(X)$$

Démonstration X_n et X admettent une variance et $E((X_n - X)^2) \underset{n\to\infty}{\longrightarrow} 0$, d'après la seconde inégalité triangulaire, on a

$$|\sqrt{E({X_n}^2)} - \sqrt{E(X^2)}| \leq \sqrt{E((X_n - X)^2)} \underset{n\to\infty}{\longrightarrow} 0$$

pour la forme quadratique positive $(X, Y) \mapsto E(X, Y) =< X, Y >$. C'est une forme linéaire symétrique positive qui respecte l'inégalité triangulaire. D'où $\sqrt{E({X_n}^2) - E(X^2)} \underset{n\to\infty}{\longrightarrow} 0$ donc $E({X_n}^2) \underset{n\to\infty}{\longrightarrow} E(X^2)$

Remarque (Variance et Cauchy-Schwarz) D'après l'inégalité de Cauchy-Schwarz :

$$\begin{aligned}|E(X_n) - E(X)| = |E(X_n - X)| &= | < X_n - X, 1 > | \\ &\leq \sqrt{< X_n - X, X_n - X >}\sqrt{< 1, 1 >} \\ &= \sqrt{E((X_n - X)^2)} \underset{n\to\infty}{\longrightarrow} 0\end{aligned}$$

Donc $|E(X_n - E(X))| \underset{n\to\infty}{\longrightarrow} 0$ par encadrement. Or $V(X_n) = E({X_n}^2) - E(X_n)^2$ donc $V(X_n) \underset{n\to\infty}{\longrightarrow} 0$

Corollaire (**Critère de convergence quadratique**) . $(X_n)_{n\in\mathbb{N}}$ converge en moyenne quadratique vers la variable certaine $a \in \mathbb{R}$ ssi

$$E(X_n) \underset{n\to\infty}{\longrightarrow} a \quad \text{et} \quad V(X_n) \underset{n\to\infty}{\longrightarrow} 0$$

Chapitre 7

Lois Conjointes

Chapitre 8

Estimation Ponctuelle

Contents

> La statistique ou les statistiques1 est la discipline qui étudie des phénomènes à travers la collecte de données, leur traitement, leur analyse, l'interprétation des résultats et leur présentation afin de rendre ces données compréhensibles par tous. C'est à la fois une branche des mathématiques appliquées2, une méthode et un ensemble de techniques.
>
> — Source : Wikipédia

En pratique, lorsque l'on essaye de déterminer des caractéristiques d'une population, on réunit celle-ci et on "compte" le nombre d'occurences des propriétés qui nous intéressent. Une fois fait, il suffit d'appliquer quelques formules pour déterminer quelques caractéristiques de la population.

Cependant, lorsque l'on souhaite, par exemple savoir la proportion de Français droitier. Il faudrait donc mettre en place à très grande échelle un sondage pour que chaque Français dise s'il est droitier ou gaucher. Outre le fait que l'expérience n'ait aucun intérêt, elle semble très compliqué à mettre en place. Il faudrait donc sonder seulement une partir des Français (on appelera cela un échantillon) et à partir du résultat obtenu, en déduire le résultat pour l'ensemble des Français. On appelle ces processus l'échantillonnage et l'estimation d'un paramètre à partir d'un échantillon (ici une proportion).

8.1 Echantillonnage

8.1.1 Généralités et définitions

La théorie de l'échantillonnage a pour but de déterminer la distribution d'une caractéristique X dans une population P à partir de l'étude d'un sous-ensemble de cette population. On appelle ce sous-ensemble un échantillon de la population.

Définition (**Echantillonnage Simple**) . On appelle échantillonnage simple le procédé qui consiste, à partir d'une population P de choisir au hasard n individus de la population de manière aléatoire. Ainsi chaque individu a la même probabilité d'être sélectionné pour l'échantillonnage.

Plus formellement, nous allons représenter un échantillon par des variables aléatoires suivant une loi prédéfinie X correspondant au paramètre recherché. Ainsi, un échantillon de taille n d'une population sera un n-uplet de variables aléatoires

$$(X_1, \dots, X_n)$$

indépendantes et de même loi que X.

Définition (**Echantillon et réalisation**) . Soit X une variable aléatoire sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$. Un échantillon de X de taille $n \in \mathbb{N}$ est une n-uplet de variables aléatoires $(X_1, \dots, X_n)$ iid de même loi que X de paramètre $\theta \in \mathbb{R}$.

Une réalisation de cet échantillonage est un n-uplets de réel $(x_1, \dots, x_n)$ tels que $\forall i \in [\![1, n]\!],\ X_i(\omega) = x_i$ où $\omega \in \Omega$.

Remarque On appellera X la **loi mère**.

Exemple Ici un exemple du cours

8.1.2 Moyenne et Variance Empirique

Définition (**Statistique**) . Considérons un échantillon d'une population $(X_1, \dots, X_n)$ sur un espace probabilisé $(\Omega, \mathbb{P})$ tels que :

$$\forall i \in [\![1, n]\!], \quad X_i : \Omega \longrightarrow E$$

où $E \subseteq \mathbb{R}$ généralement. Une statistique est une fonction qui associe une valeur à chaque réalisation de l'échantillon telle que :

$$T : E^n \longrightarrow Im(T) \subseteq \mathbb{K}$$

8.1.3 Moyenne et loi normale

Définition (**Moyenne Empirique**) . Soit $(X_1, \dots, X_n)$ un échantillon d'une population sur un espace probabilisé. On appelle **moyenne empirique** de l'échantillon la moyenne arithmétique des variables de l'échantillon. On la note $\overline{X}_n$ telle que :

$$\overline{X}_n := \frac{1}{n} \sum_{i=0}^{n} X_i$$

Exemple Pour un échantillon $(X_1, \ldots, X_n)$, la moyenne empirique est une statistique de cet échantillon.

Proposition Soit $X_1, \ldots, X_n$ un échantillon élatoire de taille n de variables iid tiré d'une distribution X d'espérance μ et de variance σ^2. On a les propriétés suivantes :

$$\mathbb{E}(\overline{X}_n) = \mu \quad \mathbb{V}(\overline{X}_n) = \frac{\sigma^2}{n}$$

- L'espérance de la moyenne empirique reste égale à l'espérance de la variable aléatoire sous-jacente, car chaque observation apporte une estimation de μ.
- La variance de la moyenne empirique est plus petite que celle de la variable aléatoire initiale, et elle diminue avec la taille de l'échantillon n. Cela reflète le fait que la moyenne empirique devient plus précise quand on augmente le nombre d'observations.

Théorème (**Somme de Variables aléatoires et Loi Normale**) . Soient $X_1, \ldots, X_n$ des variables aléatoires indépendantes suivant toute une loi normale :

$$X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$$

Soit $S_n = X_1 + \cdots + X_n$ la somme de ces variables aléatoires. Alors cette variable aléatoire suit une loi normale :

$$S_n \sim \mathcal{N}(\sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \sigma_i^2)$$

Cas Particulier : si $X_1, \ldots, X_n$ son iid, autrement dit que :

$$\forall i \in \{1, \ldots, n\}\ X_i \sim \mathcal{N}(\mu, \sigma^2)$$

Alors S_n suit une loi normale :

$$S_n \sim \mathcal{N}(n \times \mu, n \times \sigma^2)$$

Propriété (**Conséquences**) . Ce théorème induit quelques conséquences sur les variables aléatoires suivant une loi normale :

- **Stabilité de la loi normale :** La somme de variables aléatoires suivant une loi normale suit aussi une loi normale.
- **Stabilité par combinaison linéaire :** Si $X_1, \ldots, X_n$ sont indépendantes et suivent une loi normale, alors toute combinaison linéaire des X_i suit aussi une loi normale.
- Même si les variables aléatoires ne suivent pas une loi normale initialement, sous certaines conditions, leur somme (lorsque n est grand) tend vers une distribution normale.

8.1.4 Variance et fréquence

Définition (**Variance Empirique**) . Soit $(X_1, \ldots, X_n)$ un échantillon aléatoire de taille n d'une distribution X dans une population. La **variance empirique** de cet échantillon est une statistique notée S_n^2 définie par :

$$\boxed{S_n^2 := \frac{1}{n} \sum_{i=1}^{n} (X_i - \overline{X}_n)^2}$$

où X_i est la ième observation de l'échantillon.

Remarque (Interprétation) La variance empirique est une statistique qui quantifie la dispersion des observations autour de la moyenne empirique. Elle joue un rôle fondamental dans l'inférence statistique, notamment dans les tests d'hypothèses et les intervalles de confiance.

En statistiques, la fréquence mesure la proportion d'un paramètre donné dans une population.

Définition (**Fréquence Empirique**) . Soit $(X_1, \ldots, X_n)$ une échantillon aléatoire et x_i une valeur particulière observée. La fréquence empirique $F_n(x_i)$ est définie telle que :

$$F_n(x_i) := \frac{1}{n} \sum_{j=0}^{n} 1_{X_j = x_i}$$

La fréquence empirique $F_n(x_i)$ est une estimation empirique de la probabilité que la variable aléatoire prenne la valeur x_i. Pour un échantillon suffisamment grand, $F_n(X_i)$ converge vers la probabilité réelle $\mathbb{P}(X = x_i)$.

8.2 Estimation Paramétrique Ponctuelle et Qualité

8.2.1 Contexte et définition

L'estimation paramétrique ponctuelle est une branche fondamentale de la statistique qui consiste à utiliser un échantillon de données pour fournir une estimation unique (ou "ponctuelle") d'un paramètre inconnu d'une population (souvent une proportion, une moyenne, un écart-type).

Plus formellement, soit une population décrite par une variable aléatoire X ayant une distribution qui dépend d'un ou plusieurs paramètres inconnus θ. L'objectif est d'estimer le paramètre θ à partir d'un échantillon aléatoire $X_1, \ldots, X_n$.
Pour cela, nous utiliserons des estimateurs. Un estimateur est une statistique utilisée pour approximer une caractéristique inconnue (ou paramètre) d'une population, en se basant sur un échantillon aléatoire.

Définition (**Estimateur**) . Soit $(X_1, \ldots, X_n)$ un échantillon aléatoire. Soit $f_X(x, \theta)$ une distribution de probabilité sur la popualtion. Un estimateur de θ, noté $\hat{\theta}$ est une fonction mesurable de l'échantillon telle que :

$$\hat{\theta} = g(X_1, \ldots, X_n)$$

où g est une fonction construite pour approximer θ et $\hat{\theta}$ est une variable aléatoire.

Exemple La moyenne empirique est un estimateur de l'espérance μ :

$$\hat{\mu} = \frac{1}{n} \sum_{i=0}^{n} X_i$$

8.2.2 Qualité d'un estimateur

Une fois que l'on a construit un estimateur d'un paramètre θ en fonction d'un échantillon $(X_1, \ldots, X_n)$, on souhaite estimer la qualité d'un estimateur. On voudrait savoir si celui-ci nous donne une "bonne" approximation du paramètre recherché. Pour cela, on définit plusieurs critères, le **biais**, la **variance**, la **consistance** et **l'efficacité** de cet estimateur.

Définition (**Biais d'un estimateur**) . Le biais b_θ d'un estimateur $\hat{\theta}$ mesure la différence entre l'espérance de l'estimateur et la valeur exacte du paramètre recherché.

$$\boxed{b_\theta(\hat{\theta}) := \mathbb{E}(\hat{\theta}) - \theta}$$

On dit alors qu'un estimateur est non biaisé ssi $\mathbb{E}(\hat{\theta}) = \theta$. Dans le cas contraire ($\mathbb{E}(\hat{\theta}) \neq \theta$), on dit qu'il est biaisé.

Définition (**Variance d'un estimateur**) . La variance $\mathbb{V}(\hat{\theta})$ d'un estimateur $\hat{\theta}$ d'un paramètre θ mesure la dispersion des valeurs possibles de l'estimateur autour de son espérance :

$$\boxed{\mathbb{V}(\hat{\theta}) := \mathbb{E}((\hat{\theta} - \mathbb{E}(\hat{\theta}))^2)}$$

Une variance faible indique que l'estimateur est stable. Au contraire, une variance élevée indique que l'estimateur est instable à la variantion des échantillons.

Définition (**Consistance**) . Un estimateur $\hat{\theta}$ d'un paramètre θ est dit **consistant** si, lorsque la taille de l'échantillon augmente, il converge vers le paramètre estimé. Autrement dit :

$$\hat{\theta} \text{ est consistant} \iff \hat{\theta} \xrightarrow[n\to\infty]{P} \theta$$

Proposition Entre deux estimateurs d'un paramètre θ sur un échantillon $(X_1, \ldots, X_n)$, on préfèrera celui dont la variance est minimale. On parle **d'efficacité**.

Définition (**Erreur Quadratique Moyenne**) . Soit $\hat{\theta}$ un estimateur d'un paramètre θ. L'erreur quadratique moyenne de $\hat{\theta}$ permet d'évaluer la précision globale de l'estimateur en fonction de sa variance et son biais. Elle est définit comme :

$$\begin{aligned} EQM(\hat{\theta}) &:= \mathbb{E}((\hat{\theta} - \theta)^2) \\ &= \mathbb{V}(\hat{\theta}) - b_\theta(\hat{\theta})^2 \end{aligned}$$

Remarque (Interprétation) L'estimateur d'un paramètre θ ayant la plus petite EQM est généralement considéré comme meilleur. L'EQM montre comment une réduction du biais peut être compensée par une augmentation de la variance, et inversement. Finalement, chercher à minimiser l'EQM revient à équilibrer biais et variance.

8.2.3 Estimation par la méthode du maximum de vraisemblance

La méthode d'estimation par le maximum de vraisemblance est une technique statistique largement utilisée pour estimer les paramètres d'un modèle probabiliste, à partir d'un échantillon de données observées. Cette méthode consiste à trouver les valeurs des paramètres qui maximisent la fonction de vraisemblance, c'est-à-dire les paramètres qui rendent les données observées les plus probables.

Définition (**Fonction de vraissemblance**) . Soit $(X_1, \ldots, X_n)$ un échantillon de variables aléatoires iid de densité $f_X(x, \theta)$ prenant ses valeurs sur $\mathcal{X} \subset \mathbb{R}^n$ et de paramètre $\theta \in \mathbb{R}$. La fonction de vraissemblance $L(x; \theta)$ donne la probabilité d'obtenir l'échantillon observé $\{x_1, \ldots, X_n\} \in \mathcal{X}$ étant donné un paramètre θ. Elle est définie par :

$$L : \begin{cases} \mathcal{X} \times \mathbb{R} \longrightarrow \mathbb{R}^+ \\ (x_1, \ldots, x_n, \theta) \longmapsto L(x_1, \ldots, x_n, \theta) \end{cases}$$

où :

$$L(x_1, \ldots, x_n, \theta) = \mathbb{P}(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n, \theta) = \prod_{i=1}^{n} f_X(x_i, \theta)$$

Dans le cas de données continues, on écrit la fonction de vraisemblance sous forme de produit des densités, et pour des données discrètes, ce sera un produit de masses de probabilité.

Remarque (Notations) En fonction du contexte et de ce qui est recherché, on notera la fonction de vraisemblance de plusieurs manières tout en parlant du même objet :

- On peut se fixer une réalisation d'un échantillon $(x_1, \ldots, x_n)$ on parlera alors :

$$L : \begin{cases} \mathbb{R} \longrightarrow R^+ \\ \theta \longmapsto L_{(x_1, \ldots, x_n)}(\theta) \end{cases}$$

- Dans des recherches de résultats plus formels, on parlera de :

$$L(\theta | x_1, \ldots, x_n)$$

- Ou lors de calculs plus analytiques où l'on considèrera que les variables θ et $x_1, \ldots, x_n$ varient, on écrira :

$$L(x_1, \ldots, x_n; \theta)$$

On notera aussi parfois la fonction de vraisemblance en fonction de variables aléatoires $X_1, \ldots, X_n$ ou en fonction de réalisations de variables aléatoires $x_1, \ldots, x_n$. Cela dépend du contexte et de ce que l'on a besoin de quantifier.

Proposition Pour faciliter, les calculs, nous utiliserons souvent la log-vraissemblance définie par :

$$l(\theta) = \log L(\theta) = \sum_{i=1}^{n} \log f_X(x_i, \theta)$$

où $f_X : \mathcal{F} \longrightarrow \mathbb{R}$ est la densité de X.

L'estimateur du maximum de vraissemblance $\hat{\theta}$ est la valeur des paramètres $\theta \in \mathbb{R}$ qui maximisent log-vraisemblance $l(\theta)$, c'est-à-dire l'ensemble des paramètres qui rendent l'échantillon observé le plus probable. Formellement, cela revient à résoudre le problème d'optimisation suivant :

$$\hat{\theta} = \arg\max_{\theta}(l(\theta)) = \arg\max_{\theta} \sum_{i=1}^{n} \log f_X(x_i, \theta)$$

Cela se résout en déterminant les points critiques de $l(\theta)$. Autrement dit en résolvant l'équation :

$$\frac{\partial l(\theta)}{\partial \theta} = 0 \quad \text{et} \quad \frac{\partial^2 l(\theta)}{\partial \theta^2} \leqslant 0$$

On peut aussi effectuer cette méthode avec la simple fonction de vraissemblance.

Remarque (Interprétation) La méthode du maximum de vraisemblance consiste à estimer les paramètres d'un modèle statistique en choisissant ceux qui rendent les données observées les plus probables. Cela repose sur une approche probabiliste où l'on cherche à maximiser la vraisemblance de l'échantillon en fonction des paramètres du modèle.

Exemple On souhaite exprimer le paramètre $\lambda \in \mathbb{R}$ d'une loi de Poisson à partir d'un échantillon $(X_1, \ldots, X_n)$ de taille n. On a alors par définition :

$$P_\lambda(X = x) = f(x, \lambda) = e^{-\lambda} \frac{\lambda^x}{x!}$$

La fonction de vraissemblance est donc :

$$L(\theta) = \prod_{i=1}^{n} e^{-\lambda} \frac{\lambda^{x_i}}{x_i!}$$

On peut utiliser la log-vraissemblance :

$$\begin{aligned} l(\theta) = \log L(\theta) &= \ln e^{-\lambda n} + \ln \left(\prod_{i=1}^{n} \frac{\lambda^{x_i}}{x_i!} \right) \\ &= -\lambda n + \sum_{i=1}^{n} \ln \left(\frac{\lambda^{x_i}}{x_i!} \right) \\ &= -\lambda n + ln\lambda \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \ln(x_i!) \end{aligned}$$

On a alors :

$$\begin{aligned} & \frac{\partial l(\theta)}{\partial \theta} = 0 \\ \iff & -n + \frac{\sum_{i=1}^{n} x_i}{\lambda} = 0 \\ \iff & \lambda = \frac{\sum_{i=1}^{n} x_i}{n} \end{aligned}$$

On a donc un estimateur $\hat{\lambda} = \frac{1}{n} \sum_{i=1}^{n} X_i = \overline{X}$. Il est normal de tomber sur la moyenne empirique puisque λ représente l'espérance de la loi de Poisson.

8.3 Information de Fisher

8.3.1 En dimension 1

On effectue un échantillonnage $(X_1, \ldots, X_n)$ d'une population possédant un caractère répartit par une variable aléatoire X pour déterminer un paramètre θ définissant cette répartition. On souhaiterai savoir à quel point notre échantillon est fiable pour déterminer avec précision le paramètre θ. On voudrait quantifier la quantité d'information que contient l'échantillon $(X_1, \ldots, X_n)$ sur le paramètre θ. Pour cela, on utilise **l'information de Fisher**.

Définition (**Information de Fisher**) . Soit $(X_1, \dots, X_n)$ un échantillon d'une population possédant un caractère répartit par une variable aléatoire X de paramètre θ. L'information de Fisher $I(\theta)$ est une mesure de la quantité d'information que contient l'échantillon sur θ. Elle est définie comme la variance de l'estimateur de ce paramètre. Plus formellement :

$$\boxed{I(\theta) := -\mathbb{E}\left(\frac{\partial^2}{\partial\theta^2} l(\theta)\right) = \mathbb{E}\left[\left(\frac{\partial l(\theta)}{\partial\theta}\right)^2\right]}$$

Elle est définie comme la variance de l'estimateur de ce paramètre, et elle est directement liée à la courbure de la fonction de log-vraisemblance $l(\theta)$.

Remarque (Interprétation) On peut voir différentes interprétations de l'information de Fisher d'un échantillon :

- **Courbure de la log-vraissemblance :** L'information de Fisher mesure la courbure de la fonction de log-vraisemblance $l(\theta)$ par rapport aux paramètres θ. Si la log-vraisemblance est fortement courbée autour de la valeur vraie de θ, l'information de Fisher sera grande, ce qui signifie que les paramètres sont bien estimés avec une faible incertitude. Si la courbure est faible, l'incertitude sur l'estimation de θ est élevée.
- **Estimation Précise :** Une grande information de Fisher implique que l'échantillon fournit beaucoup d'informations sur le paramètre, ce qui rend l'estimation plus précise. À l'inverse, une faible information de Fisher suggère que l'échantillon contient peu d'informations utiles sur le paramètre.

Propriété (**Information de Fisher**) . La quantité d'information de Fisher d'un échantillon est toujours **positive ou nulle**.

8.3.2 En dimension d

Depuis le début du chapitre, nous considérons un échantillon de variables iid distribué selon une loi paramétrée par un unique paramètre θ. Considérons maintenant que notre distribution X soit définie par $d \in \mathbb{N}$ paramètres $\theta \in \mathbb{R}^d$. On notera $\Theta \subseteq \mathbb{R}^d$ l'ensemble de définition de ces paramètres. Définissons alors la matrice d'information de Fisher pour cet échantillon...

Définition (**Matrice d'information de Fisher**) . Soit $(X_1, \dots, X_n)$ un échantillon de variables iid distribué par un loi X de paramètres $\theta = \theta_1, \dots, \theta_d \in \Theta$. La matrice d'information de Fisher de cet échantillon est définie par :

$$\boxed{I(\theta) = I(\theta_1, \dots, \theta_n) := \mathbb{E}\left(\langle \nabla_\theta \log L_{(x_1,\dots,x_n)}(\theta).\, {}^t\nabla_\theta \log L_{(x_1,\dots,x_n)}(\theta)\rangle\right)}$$

où :

- $L_{(x_1,\dots,x_n)}(\theta)$ est la fonction de vraisemblance de l'échantillon.
- $\nabla_\theta \log L_{(x_1,\dots,x_n)}(\theta)$ est le gradient de la log-vraissemblance par rapport à θ.

$$\nabla_\theta \log L_{(x_1,\dots,x_n)}(\theta) = \begin{pmatrix} \frac{\partial l_{(x_1,\dots,x_n)}(\theta)}{\partial\theta_1} \\ \frac{\partial l_{(x_1,\dots,x_n)}(\theta)}{\partial\theta_2} \\ \vdots \\ \frac{\partial l_{(x_1,\dots,x_n)}(\theta)}{\partial\theta_d} \end{pmatrix}$$

Propriété (**Matrice d'information de Fisher**) . La matrice d'information de Fisher pour une échantillon $(X_1, \ldots, X_n)$ d'une distribution X de paramètres $(\theta_1, \ldots, \theta_d) \in \Theta$ possède plusieurs propriétés :

- **Symétrie :** $I_{i,j}(\theta) = I_{j,i}(\theta) \quad \forall i, j \in [\![1, d]\!]$
- **Alternative :** Une autre façon de calculer cette matrice est de passer par les dérivées secondes de la log-vraisemblance :

$$I_{i,j}(\theta) = -\mathbb{E}\left(\frac{\partial^2 l(\theta)}{\partial\theta_i \partial\theta_j}\right) \quad \forall i, j \in [\![1, d]\!]$$

Proposition (CR) L'information de Fisher pour un échantillon $(X_1, \ldots, X_n)$ n'existe que sous certaines conditions :

1. La fonction de vraisemblance $L(\theta)$ doit être intégrable. Cela garantit que $f_\theta(X)$ soit une densité de probabilité.
2. Θ (l'ensemble de définition des paramètres de l'échantillon) est un ouvert de $\mathbb{R}^n$.
3. L'application $\theta \mapsto L_{(x_1,\ldots,x_n)}(\theta)$ doit être **différentiable** pour tout $x_i \in \mathcal{F}$ sur Θ

8.3.3 Cas particuliers et Exemples

Sous certaines conditions, il peut y avoir égalité de la matrice d'information de Fisher vectorielle et scalaire.
Pour que la matrice $I(\theta)$ en dimension $d > 1$ se réduise à une forme scalaire pour $d = 1$, il faut que :

- Le vecteur paramètre $\theta \in \Theta \subseteq R^d$ puisse se réduire à un seul paramètre $\theta' \in \mathbb{R}$.
- Ou si $\theta = (\theta_1, \ldots, \theta_d)$ est un vecteur, il faut que la matrice d'information se réduise à un scalaire.

  $$\text{i.e} \quad \forall i \neq j, \ \left\langle \frac{\partial l(\theta)}{\partial\theta_i} \cdot \frac{\partial l(\theta)}{\partial\theta_j} \right\rangle = 0$$

  Cela implique que les composantes θ_i sont statistiquement indépendantes dans l'information fournie par les données.
- La fonction de vraisemblance doit se factoriser de manière à dépendre liénairement d'un unique paramètre :

  $$L_{(X_1,\ldots,X_n)}(\theta) = g(X)h(\theta)$$

Chapitre 9

Estimation par Intervalle de Confiance

Contents

Dans le chapitre précédent, nous avons étudié l'estimation ponctuelle, qui consiste à proposer une valeur unique pour un paramètre inconnu d'une population à partir d'un échantillon. Cependant, une estimation ponctuelle est sujette à une certaine incertitude, car elle dépend du choix de l'échantillon observé.
Pour quantifier cette incertitude, on introduit la notion d'intervalle de confiance. Contrairement aux estimateurs ponctuels, un intervalle de confiance ne donne pas une valeur unique du paramètre, mais une plage de valeurs dans laquelle le paramètre a une forte probabilité de se situer.
L'objectif de ce chapitre est d'étudier les principes fondamentaux des intervalles de confiance, leur construction et leur interprétation dans le cadre des statistiques inférentielles.

9.1 Premières Définitions

Commençons par bien définir la notion d'intervalle de confiance...

Définition (**Intervalle de Confiance**) . Soit ϕ un paramètre inconnu d'une population (une moyenne, une proportion, etc..). Un intervalle de confiance au niveau $1-\alpha$ est un intervalle $[L(X); U(X)]$ définit à partir d'un échantillon $X = (X_1, \ldots, X_n)$, tel que :

$$\mathbb{P}(L(X) \leqslant \theta \leqslant U(X)) = 1 - \alpha$$

où α est le **risque** ou **niveau de signification** et $1-\alpha$ est le **niveau de confiance** de l'intervalle. En général, on choisit 0.95 ou 0.99.

Cette notion de risque signifie que si l'on répétait l'échantillonnage un grand nombre de fois, alors dans $100(1-\alpha)\%$ des cas, l'intervalle contiendrai la vraie valeur de θ.

En revanche, pour un intervalle donné, on ne peut pas affirmer que θ appartient à cet intervalle avec une probabilité $1-\alpha$, car θ est une quantité fixe, non aléatoire. Ce point est fondamental dans l'interprétation des intervalles de confiance.

9.2 Intervalle de confiance d'une moyenne

9.2.1 Principe

Soit μ la moyenne d'une population et une échantillon de cette population $X = (X_1, \ldots, x_n)$ de taille $n \in \mathbb{N}$. Ici, l'objectif est de construire un intervalle $[L, U]$ tel que :

$$\mathbb{P}(L \leqslant \mu \leqslant U) = 1 - \alpha$$

Pour construire cet intervalle, nous allons utiliser un estimateur ponctuel de la moyenne théorique, la moyenne empirique :

$$\overline{X} = \frac{1}{n}\left(\sum_{i=1}^{n} X_i\right)$$

La dispersion des valeurs autour de μ est caractérisée par l'écart-type σ et la taille de l'échantillon n. Grâce au théorème central limite, la distribution de $\overline{X}$ suit approximativement une loi normale si n est suffisament grand.

9.2.2 Premier Cas : si l'écart-type est connu

Si l'écart-type de la population σ est connu, alors $\overline{X}$ suit une loi normale de moyenne μ et d'écart-type $\frac{\sigma}{\sqrt{n}}$. On a donc :

$$\overline{X} \sim \mathcal{N}\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

On peut donc construire l'intervalle de confiance de μ suivant :

$$\boxed{IC_\mu := \left[\overline{X} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]}$$

où $z_{\alpha/2}$ est le quantille de la loi normale tel que :

$$\mathbb{P}(Z \leqslant z_{\alpha/2}) = 1 - \frac{\alpha}{2}$$

9.2.3 Second Cas : si l'écart-type est inconnu

Si l'écart-type de la population σ est connu, on doit l'estimer par l'écart-type empirique de l'échantillon :

$$S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2}$$

Dans ce cas, la statistique suivante suit une loi de Student à $n-1$ degrés de liberté :

$$T = \frac{\overline{X}\mu}{S/\sqrt{n}} \sim t_{n-1}$$

On peut alors construire l'intervalle de confiance suivant pour μ :

$$\boxed{IC_\mu := \left[\overline{X} \pm t_{\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n}}\right]}$$

où $t_{\frac{\alpha}{2},n-1}$ est le quantille de la loi de Student à $n-1$ degrés de liberté.

9.2.4 Interprétation et tableau récapitulatif

- Si l'on répétait de nombreuses fois l'expérience d'échantillonnage, et que l'on construisait à chaque fois l'intervalle alors dans $100(1-\alpha)\%$ des cas μ serait dans l'intervalle de confiance.
- **Attention :** Cela ne veut pas dire que μ a une probabilité de $1-\alpha$ d'être dans l'intervalle puisque μ n'est pas une variable aléatoire, c'est une constante inconnue.
- En étudiant la construction des intervalles, on peut remarquer que plus l'échantillon est grand, plus l'intervalle est petit. En théorie, on peut donc encadrer μ de façon aussi fine que l'on souhaite.

Taille	σ **connu ?**	**Statistique**	**Distribution**
$n > 30$	Oui	$Z = \dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}}$	Normale $\mathcal{N}(0,1)$
$n > 30$	Non	$T = \dfrac{\bar{X}-\mu}{S/\sqrt{n}}$	Approximativement $\mathcal{N}(0,1)$
$n < 30$	Oui	$Z = \dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}}$	Normale $\mathcal{N}(0,1)$
$n < 30$	Non	$T = \dfrac{\bar{X}-\mu}{S/\sqrt{n}}$	Student ($n-1$ ddl)

TABLE 9.1 – Synthèse et conseils d'application pour l'IC de μ

9.3 Intervalle de confiance d'une proportion

On considère maintenant un échantillon $X = (X_1, \ldots, X_n), n \in \mathbb{N}$ de fréquence f d'un paramètre observé. On souhaite déterminer la proportion p de ce même paramètre dans la population où l'échantillon a été prélevé. Pour cela, nous allons construire un intervalle de confiance pour obtenir un encadrement plus ou moins précis de la proportion (exacte) recherchée.

On suppose ici que $n > 30$. Alors la distribution de la proportion observée f peut être approchée par une loi normale d'après le théorème central limite.

Ainsi l'intervalle de confiance pour p à un niveau de confiance $1-\alpha$ est donné par :

$$\boxed{IC_p := \left[f \pm z_{\alpha/2}\sqrt{\frac{f(1-f)}{n}}\right]}$$

où :

- $z_{\alpha/2}$ est le quantile de la loi normale standard.
- $\sqrt{\frac{f(1-f)}{n}}$ représente l'écart-type de f.

9.4 Taille d'un échantillon

Depuis le début du chapitre, on construit des intervalles de confiance pour des moyennes et des proportions. Or on remarque bien que la précision de cet intervalle dépend énormément de la taille de l'échantillon prélevé dans la population. On pourrait alors se demander, si, en fonction des caractéristiques de la population et d'une précision voulue, on ne pourrait pas directement déterminer la taille minimale de l'échantillon à prélever.

9.4.1 Taille de l'échantillon pour estimer une moyenne

Avant de réaliser l'échantillonnage, on va poser une précision $d \in \mathbb{R}$, telle que :

$$\mu = \overline{x} \pm d$$

on a donc, d'après la formule de l'intervalle précédent :

$$\boxed{d = z_\alpha \frac{\sigma}{\sqrt{n}} \iff n = \left(z_\alpha \frac{\sigma}{d}\right)^2}$$

9.4.2 Taille de l'échantillon pour estimer une proportion

Avant de réaliser l'échantillonnage, on va poser une précision $d \in \mathbb{R}$ telle que

$$p = f \pm d$$

La taille de l'échantillon est donnée par :

$$\boxed{n = z_\alpha^2 \frac{f(1-f)}{d^2}}$$

Chapitre 10

Test d'hypothèses

Contents

10.1 Introduction Générale aux Tests d'Hypothèses

10.1.1 Hypothèses nulle et alternative

Avant tout test d'hypothèse, nous devons tous d'abord définir les hypothèses du test. Elles permettent de formaliser l'hypothèse à tester et dépendent du contexte.

Définition (**Hypothèse nulle et alternative**) . Soit $X = (X_1, \dots, X_n)$ un échantillon de taille $n \in \mathbb{N}$. On définit deux hypothèses pour tout test sur l'échantillon :

- **L'hypothèse nulle (H_0)** : on la considère a priori vraie. Il faut des observations sur X très éloignés de H_0 pour la rejeter.
- **L'hypothèse alternative (H_1)** : c'est l'hypothèse complémentaire à H_0, celle que l'on retient en cas de rejet de H_0.

En général H_0 est l'hypothèse que l'on préfère considérer (égalité de moyennes, appartenance à une loi, etc...), l'objectif du test est de confirmer ou infirmer cette hypothèse. Dans le second cas, on dit "qu'on rejette H_0". Cependant, on "n'accepte" pas forcément H_1 qui est l'hypothèse par défaut.

Exemple (Hypothèses nulles et alternative) Soient X_1 et X_2 deux échantillons de taux glycémie de deux populations. L'échantillon 1 correspond à une population ayant pris des médicaments et l'échantillon 2 à la population ayant pris des placebo.
On définit μ_1 et μ_2 les moyennes respectives des deux échantillons (moyennes empiriques). On souhaite savoir si la prise du médicemant effecte **sensible** la glyécmie de la population. Pour cela, nous allons faire un test d'égalité de moyennes (voir plus tard) dont les hypothèses sont :

$$H_0 : \mu_1 = \mu_2 \quad H_1 : \mu_1 \neq \mu_2$$

La description mathématique des hypothèses peut être biaisée, intuitivement, on comprendra plutôt :

- H_0 : il est crédible de penser que $\mu_1 = \mu_2$
- H_1 : μ_1 est signficativement différente de μ_2

Définition (**test bilatéral/unilatéral**) . En fonction de la formulation de l'hypothèse nulle, on définit deux types de test :

- $H_0 : \theta_1 \geqslant \theta_2$: **Test unilatéral**
- $H_0 : \theta_1 \leqslant \theta_2$: **Test unilatéral**
- $H_0 : \theta_1 = \theta_2$: **Test bilatéral**

10.1.2 Erreur de type I et erreur de type II

Lors d'un test d'hypothèse, on fixe que ce l'on appelle un **seuil de signification** α.

Définition (**Niveau de signification**) . Le niveau de signification d'un test d'hypothèse est un réel $\alpha \in [0, 1]$ qui correspond à la probabilité de rejeter H_0 alors qu'elle est vraie. Il est en général exprimé sous forme de pourcentage.

Ce paramètre est très important puisqu'il permet de définir ce que l'on appelera la région d'acceptation du test d'hypothèse.

Définition (**Erreurs**) . Pour un test d'hypothèse, on définit plusieurs erreurs :

- **Erreur de type I** (faussement rejeter H_0) : lorsque l'on rejette H_0 et que l'on accepte H_1 alors que H_0 est vraie. Cette probabilité est donnée par α.
- **Erreur de type II** (faussement accpeter H_0) : lors que l'on ne rejette pas H_0 alors qu'elle est fausse.

Les erreurs de type I et de type II permettent d'indentifier clairement le concept de "faux positif" ou "faux négatif". L'idéal serait de minimiser le risque de faire les deux erreurs mais cela entre parfois en contradiction. On va donc pour voir jouer sur la taille de **l'échantillon et le seuil** de **signification pour les minimiser**.

10.1.3 Statistique de test et région d'acceptation

Lors d'un test d'hypothèse, en fonction des caractéristiques de l'échantillon (variance, écart-type, moyenne, etc...), on définit une **statistique de test** souvent appelée **variable de décision**.

Définition (**Variable de décision**) . La variable de décision est une variable aléatoire dépendant de l'échantillon à tester qui permet de quantifier l'écart entre les données observés et ce que l'on attend sous l'hypothèse H_0.

La statistique de test mesure donc l'écart entre les observations de l'échantillon et ce que l'on attend sous l'hypothèse nulle. Elle est comparée à une valeur seuil (ou à une distribution théorique) pour déterminer la décision du test.

Définition (**Région d'acceptation**) . La région d'acceptation d'un test d'hypothèse est l'ensemble des valeurs possibles de la statistique de test (variable de décision) pour lesquelles l'hypothèse nulle (H_0) ne sera pas rejeté. Les bornes définissant la région d'acceptation sont appelées **valeurs critiques** il peut y en avoir une ou deux.

On appelle **région de rejet** le complémentaire de la région d'acceptation dans $\mathbb{R}$.
La région d'acceptation se représente donc sous la forme d'un intervalle qui dépend de l'hypothèse nulle H_0 :

- **Test bilatéral ($H_0 : \theta_1 = \theta_2$)** : la région d'acceptation est un intervalle fermé de la forme $[a, b] \subset \mathbb{R}$.
- **Test unilatéral droit** ($H_0 : \theta_1 \leqslant \theta_2$) : la région d'acceptation est de la forme : $]-\infty; a]$
- **Test unilatéral droit** ($H_0 : \theta_1 \geqslant \theta_2$) : la région d'acceptation est de la forme : $[a; +\infty[$

Pour un test bilatéral, on a donc deux valeurs critiques et pour un test unilatéral, nous en avons qu'un seule.

10.1.4 Processus d'un test d'hypothèse

A partir des définitions précedentes, on peut donc construire les tests d'hypothèses. En pratique un test d'hypothèses sur un échantillon $X = (X_1, \ldots, X_n)$ se fait en 5 étapes :

1. **Formulation des hypothèses :** en fonction du contexte, on fixe H_0 et H_1. Leur forme définira la forme de la région d'acceptation.
2. **Définition du seuil de signification** : on choisit la valeur de α en fonction de la rigueur que l'on veut apporter au test.
3. **Calcul de la variable de décision :** à partir des données de l'échantillon et des données réelles (si on les connaît) on calcule la valeur de la variable de décision (VD).
4. **Calcul de la région d'acceptation :** les hypothèses du test permettent de définir la forme de la région d'acceptation. Elle se calcule de la même façon qu'un intervalle de confiance.
5. **Conclusion :** en fonction de la valeur de la VD on peut observer plusieurs cas :
   - si $VD \in RA$ alors on **accepte** H_0
   - si $VD \notin RA$ alors on **rejette** H_0

   On conclut en fonction du contexte du test.

Définissons maintenant différents types de tests d'hypothèses sur des échantillons aléatoires.

10.2 Tests d'hypothèses pour une moyenne

Les tests d'hypothèses sur les moyennes permettent de vérifier la conformité de la moyenne d'un échantillon à une valeur théorique ou à la moyenne d'un autre échantillon indépendant.

10.2.1 Tests sur un échantillon unique

Soit $X = (X_1, \ldots, X_n)$ un échantillon aléatoire. On supppose que la distribution de X suit une loi normale de moyenne μ_0 et d'écart-type σ. On note $\overline{X} = \mu$ la moyenne de l'échantillon. L'hypothèse nulle est alors de la forme :

$$H_0 : \mu = \mu_0$$

contre une **hypothèse alternative** de la forme :

- **Bilatérale :** $H_1 : \mu \neq \mu_0$
- **Unilatérale droite :** $H_1 : \mu > \mu_0$
- **Unilatérale gauche :** $H_1 : \mu < \mu_0$

Ensuite, nous devons déterminer la statistique de test (variable de décision). Elle dépend de la connaissance ou non de l'écart type théorique σ.

- **Si σ est connu :** la statistique de test suit une **loi normale centrée réduite** :

$$VDR \sim \mathcal{N}(0,1) \quad \text{et} \quad VDR = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$$

- **Si σ est inconnu :** nous devons le déterminer par l'estimateur S correspondant à l'écart-type de l'échantillon. La variable de décision suit alors une **loi de Student à $n-1$ degrés de liberté**.

$$S = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2} \quad \text{et} \quad VDR = \frac{\overline{X} - \mu_0}{S = \sqrt{n}}$$

Enfin, il reste à définir la **région d'acceptation**, qui dépend de la loi de la variable de décision réduite (VDR).

- Pour un **test bilatéral** de valeur critique α on a :

$$RA := [z_{\alpha/2}\ ,\ z_{1-\alpha/2}]$$

- Pour un **test unilatéral droit** de valeur critique α on a :

$$RA := [-\infty\ ,\ z_{1-\alpha}]$$

- Pour un **test unilatéral gauche** de valeur critique α on a :

$$RA := [z_{\alpha}\ ,\ +\infty]$$

Remarque (Taille de l'échantillon) En fonction de la taille de l'échantillon, la variable de décision réduite peut être approchée par différentes lois :

- Si $n > 30$ alors $VDR \sim \mathcal{N}(0,1)$, on calcule donc la région d'acceptation avec la table de la loi normale.
- Si $n <= 40$ alors VDR suit une loi de Student à $n-1$ degrés de liberté. On utilise donc la table de la loi de Student.

10.2.2 Test de comparaison de deux moyennes (deux échantillons)

On cherche ici à déterminer si deux échantillons issus de deux populations ont sensiblement la même moyenne pour une erreur de α. On considère deux échantillons des moyennes μ_1 et μ_2, d'écart-types σ_1 et σ_2 de taille n_1 et n_2.
L'hypothèse nulle est alors de la forme :

$$H_0 : \mu_1 = \mu_2$$

La statistique de test dépend de l'égalité ou non des écart-types des deux échantillons σ_1 et σ_2 :

- Si les **écart-types sont différents**, alors on a la variable de décision suivante :

$$VDR = \frac{\mu_1 - \mu_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

- Si les **écart-types sont égaux**, la variance commune est définie par :

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n}(x_i - \mu_1^2) + \sum_{i=1}^{n}(x_i - \mu_2^2)}{n_1 + n_2 - 2}$$

  et la variable de décision réduite est définie par :

$$VDR = \frac{\mu_1 - \mu_2}{\sqrt{\hat{\sigma}^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Comme précédement, nous devons définir la région d'acceptation qui dépend toujours de H_1 :

- Pour un **test bilatéral** de valeur critique α on a :

$$RA := [z_{\alpha/2}\ ,\ z_{1-\alpha/2}]$$

- Pour un **test unilatéral droit** de valeur critique α on a :

$$RA := [-\infty\ ,\ z_{1-\alpha}]$$

- Pour un **test unilatéral gauche** de valeur critique α on a :

$$RA := [z_{\alpha}\ ,\ +\infty]$$

Remarque (Taille de l'échantillon) Comme précédement, la variable de décision suit une loi normale pour les échantillons supérieurs à 30 données. Dans le cas contraire, elle suite une loi de Student à $n_1 + n_2 - 2$ degrés de liberté.

10.2.3 Test de comparaison des moyennes de deux échantillons appariés

On considère maintenant deux échantillons $X = (X_1, \ldots, X_n)$ et $Y = (Y_1, \ldots, Y_n)$ de moyennes μ_1 et μ_2 et d'écart-type σ_1 et σ_2 où chaques valeurs des deux échantillons sont associées par paires. En pratique c'est le cas lorsque l'on soumet les mêmes individus d'une population à deux tests.
On commence par calculer la différence des deux observations d définie par :

$$\forall i \in [\![1, n]\!], d_i = x_i - y_i$$

La variable de décision correspond donc à la moyenne des différences :

$$VD = \overline{d}$$

Dans le cas où la distribution du caractère étudié suit une loi normale, la variable de décision est elle même normale. La variable de décision centrée réduite est donc :

$$VDR := \frac{\overline{d}}{\sigma_d/\sqrt{n}}$$

Comme précédement dans le cas d'un échantillon grand ($n \geqslant 30$) la VDR suit une loi normale centrée réduite ($\mathcal{N}(0,1)$). Dans le cas contraire ($n < 30$), elle suit une loi de Student à $n-1$ degrés de liberté.

Exemple Un ingénieur souhaite tester si la durée de vie moyenne d'une batterie est au moins de 500 heures. Il prélève un échantillon de $n = 20$ batteries et obtient les statistiques suivantes :

- Moyenne observée : $\overline{X} = 510$ heures.
- Écart-type empirique : $S = 15$ heures.
- Niveau de test : $\alpha = 5\%$.

Nous allons tester l'hypothèse suivante :

$$H_0 := 500$$
$$H_1 : \mu > 500$$

La statistique de test est donnée par :

$$T = \frac{\overline{X} - \mu_0}{S/\sqrt{n}} = \frac{510 - 500}{15/\sqrt{20}} \approx 2.99$$

La valeur critique de la loi de Student pour $n - 1 = 19$ degrés de liberté et un seuil de $\alpha = 5\%$ est $t_{0.05,19} \approx 1.73$
Nous comparons :

Condition	Résultat	Décision
$T > t_{0.05,19}$	$2.99 > 1.73$	Rejet de H_0

Puisque $T = 2.99$ est supérieur à $t_{0.05,19} = 1.73$, nous rejetons l'hypothèse nulle H_0 au seuil de 5%.
Il y a donc suffisamment d'évidence pour conclure que la durée de vie moyenne des batteries est significativement supérieure à 500 heures.

10.3 Tests d'hypothèses pour une variance

De même que pour les moyennes, on cherche à étudier la conformité de la variance d'un échantillon par rapport à une valeur théorie ou d'autres échantillon. Pour cela, on utilise le même procédé.

10.3.1 Test de conformité d'une variance

Soit $X = (X_1, \ldots, X_n)$ un échantillon de taille n, de moyenne μ et de variance σ^2. On cherche à savoir si la variance de l'échantillon est sensiblement la même qu'un variance théorie σ_0^2 pour un seuil de signification α.

On pose alors l'hypothèse nulle suivante :

$$H_0 : \sigma^2 = \sigma_0^2$$

La variable de décision est donnée par :

$$VD = \frac{(n-1)S^2}{\sigma_0^2} \quad \text{où} \quad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \mu)^2$$

Cette statistique suit une loi de chi-deux à $n-1$ degrés de liberté sous H_0.
On définit ensuite la région d'acceptation comme suit :

- Pour un **test bilatéral** de valeur critique α on a :

$$RA := [\chi^2_{\alpha/2}\ ,\ \chi^2_{1-\alpha/2}]$$

- Pour un **test unilatéral droit** de valeur critique α on a :

$$RA := [0\ ,\ \chi^2_{1-\alpha}]$$

- Pour un **test unilatéral gauche** de valeur critique α on a :

$$RA := [\chi^2_{\alpha}\ ,\ +\infty]$$

10.3.2 Test de comparaison de deux variances (échantillons indépendants)

On veut comparer la variance de deux échantillons indépendants X_1 et X_2 de variances respectives σ_1^2 et σ_2^2 et de tailles n_1 et n_2.
On a alors l'hypothèse nulle suivante :

$$H_0 : \sigma_1^2 = \sigma_2^2$$

La variable de décision du test correspond au rapport des deux variances observées des deux échantillons :

$$VD = \frac{\hat{\sigma_1}^2}{\hat{\sigma_2}^2}$$

Par convention on choisit l'échantillon 2 comme celui avec la plus grande variance. Elle suit une loi de Fisher à $n_1 - 1$ et $n_2 - 2$ degrés de liberté. On a tout le temps affaire à un test bilatéral d'où la région d'acceptation suivante :

$$RA := [F_{\alpha/2}\ ,\ F_{1-\alpha/2}]$$

10.4 Tests d'hypothèses pour une proportion

10.4.1 Test de conformité d'une proportion

Soit $X = (X_1, \ldots, X_n)$ un échantillon de taille n, où chaque X_i suit une loi de Bernoulli de paramètre p. On cherche à vérifier si la proportion observée est compatible avec une valeur théorique p_0, avec un seuil de signification α.
L'hypothèse nulle est donnée par :

$$H_0 : p = p_0$$

L'estimateur naturel de la proportion est :

$$\hat{p} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

La statistique de test est :

$$VD = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

Sous H_0, VD suit asymptotiquement une loi normale centrée réduite $\mathcal{N}(0,1)$ si n est suffisamment grand.
Les régions d'acceptation sont :

- Test bilatéral de niveau α :
$$RA := \left[-z_{1-\alpha/2}, z_{1-\alpha/2}\right]$$
- Test unilatéral droit :
$$RA :=]-\infty, z_{1-\alpha}]$$
- Test unilatéral gauche :
$$RA := [-z_{1-\alpha}, +\infty[$$

10.4.2 Test de comparaison de deux proportions

Considérons deux échantillons indépendants de tailles n_1 et n_2, suivant des lois de Bernoulli de paramètres respectifs p_1 et p_2. On souhaite tester si ces proportions sont égales.
L'hypothèse nulle est :

$$H_0 : p_1 = p_2$$

On estime les proportions par :

$$\hat{p}_1 = \frac{X_1}{n_1}, \quad \hat{p}_2 = \frac{X_2}{n_2}$$

L'estimateur global sous H_0 est :

$$\hat{p} = \frac{X_1 + X_2}{n_1 + n_2}$$

La variable de décision est donnée par :

$$VD = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Sous H_0, cette statistique suit asymptotiquement une loi normale centrée réduite.
Les régions d'acceptation sont définies comme suit :

- Test bilatéral de niveau α :
$$RA := \left[-z_{1-\alpha/2}, z_{1-\alpha/2}\right]$$
- Test unilatéral droit :
$$RA :=]-\infty, z_{1-\alpha}]$$
- Test unilatéral gauche :
$$RA := [-z_{1-\alpha}, +\infty[$$

Exemple On souhaite comparer les taux de réussite à un examen entre deux groupes d'étudiants. Le premier groupe, de taille $n_1 = 200$, a $X_1 = 140$ étudiants ayant réussi. Le second groupe, de taille $n_2 = 250$, a $X_2 = 175$ étudiants ayant réussi.
Nous voulons tester si la proportion de réussite est la même dans les deux groupes, au seuil de signification $\alpha = 5\%$.

Étape 1 : Hypothèses Les hypothèses sont les suivantes :

$$H_0 : p_1 = p_2 \quad H_1 : p_1 \neq p_2$$

(ce qui correspond à un test bilatéral).

Étape 2 : Estimation des proportions On estime les proportions observées :

$$\hat{p}_1 = \frac{X_1}{n_1} = \frac{140}{200} = 0.7, \quad \hat{p}_2 = \frac{X_2}{n_2} = \frac{175}{250} = 0.7$$

L'estimateur global sous H_0 est :

$$\hat{p} = \frac{X_1 + X_2}{n_1 + n_2} = \frac{140 + 175}{200 + 250} = 0.7$$

Étape 3 : Calcul de la statistique de test La variable de décision est donnée par :

$$VD = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{0.7 - 0.7}{\sqrt{0.7(1-0.7)\left(\frac{1}{200} + \frac{1}{250}\right)}} = 0$$

Étape 4 : Détermination de la région d'acceptation Le test est bilatéral au seuil $\alpha = 5\%$, donc la région d'acceptation est :

$$RA := \left[-z_{1-\alpha/2}, z_{1-\alpha/2}\right] = [-1.96, 1.96]$$

Avec $z_{1-\alpha/2} = 1.96$.

Étape 5 : Conclusion Comme $VD = 0 \in RA$, on ne rejette pas H_0. On conclut qu'il n'y a pas de différence significative entre les deux proportions au seuil de 5%.

10.5 Test de conformité à une loi

Les tests de conformité d'un échantillon à une distribution théorique visent à déterminer si les données observées suivent une certaine distribution théorique.
Le test du Chi-carré est utilisé pour comparer la fréquence observée des événements avec la fréquence attendue selon une hypothèse nulle.
Ainsi, soit $X = (X_1, \ldots, X_n)$ un échantillon aléatoire, on cherche à déterminer si il suit une loi

10.6 Tests d'indépendances

10.7 Interprétation des résultats

Annexe 1 - Graphes et Théorie des Langages

Chapitre 1

Théorie des Graphes

Contents

Fiche réalisée grâce au cours de Thierry Montaut et Laura Brillon.
Dans ce cours, on note $G = (X, E)$ un graphe.

1.1 Graphes, Représentations et Parcours

1.1.1 Définitions - Graphes Orientés et Non Orientés

Vocabulaire

Définition (**Graphe non orienté**) . On appelle graphe non orienté un couple d'ensembles finis $G = (X, E)$ où $X = \{1, \dots, n\}$ représente les sommets du graphe et $E = \{(x_i, y_j), \dots\}, x_i, y_j \in X$ l'ensemble des arrêtes du graphe. Une arrête est une liaison entre deux sommets.
Un graphe est dit **simple** s'il n'existe pas de double arrêtes entre deux sommets ou de bouble (i.e une arrête de la forme (x, x)). Autrement, on parle de **multigraphe.**

Définition (**Graphe Orienté**) . On appelle graphe orienté un couple d'ensembles finis $G = (X, E)$ où $X = \{1, \dots, n\}$ représente les sommets du graphe et $E = \{(x_i, y_j), \dots\}$, $x_i, y_j \in X$ l'ensemble ordonné des arrêtes du graphe. Pour un graphe orienté, les notions de graphe simple et multiraphe sont les même que pour le cas non orienté.

Exemple (Graphes et leur représentation graphique) Soient G_1 et G_2 deux graphes, en voici une représentation :

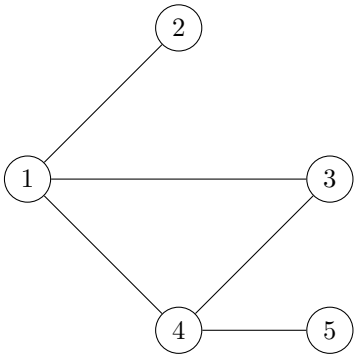


FIGURE 1.1 – Graphe non orienté

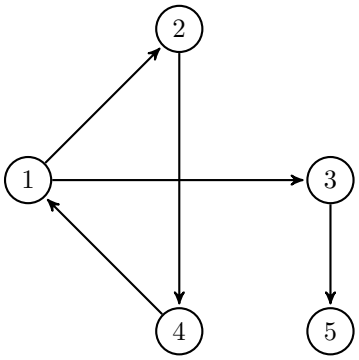


FIGURE 1.2 – Graphe orienté

Définition (**Planaire**) . Une graphe G est dit planaire s'il existe une représentation de G en deux dimensions telle qu'aucun de ses sommets ne se croisent.

Voisinnage et degré

Définition (**Voisinnage**) . Le voisinnage d'un sommet x de G est l'ensemble des sommets y de G tels qu'il existe une arrête entre x et y dans G. On le note $V(x)$.

Remarque Si x est dans le voisinnage de y, on dira que x est adlascent à y et inversement.

Définition (**Degré**) . Le degré d'un sommet x de G est le cardinal du voisinnage x. On le note $d(x)$. Un sommet de voisinnage nul est dit **isolé** et un sommet de voisinnage égal à 1 est dit **pendant**.

Définition (**Voisinnage entrant et sortant**) . Soit G un graphe orienté et x un sommet de G. On définit deux types de voisinnages :

- **Voisinnage entrant :** noté $V^-(X)$ est l'ensemble des prédécesseurs de x.
- **Voisinnage sortant :** noté $V^+(X)$ est l'ensemble des successeurs de x.

On définira comme précédemment le degré sortant et le degré entrant d'un sommet x.

Graphes Remarquables (non orientés)

Dans cette sous-section, nous ne parlerons que de graphes non-orientés.

Définition (**Graphe complet**) . On appelle graphe complet un graphe tel que pour tous sommets x et y de G il existe une arrête entre x et y dans G. Les graphes complets n sommets sont notés K_n

Remarque Autre définition de graphe complet et un petit peu d'histoire ne fera pas mal...

- On peut aussi définir un graphe complet comme étant un graphe dont tous ses sommets sont adjascents.
- *La notation K pourrait avoir deux origines, la première étant en hommage à Kazimierz Kuratowski, un éminent mathématiciens polonais ayant beaucoup contribué à la théorie des graphes. La seconde, plus simple, K proviendrait de sa traduction en Allemand komplett.*

Exemple Représentation des graphes complets K_4 et K_5 :

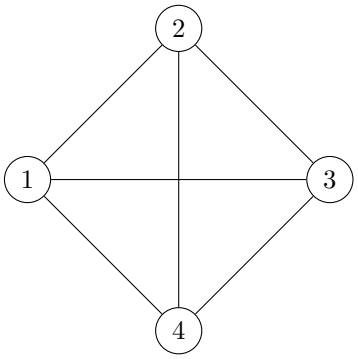


FIGURE 1.3 – Graphe complet K_4

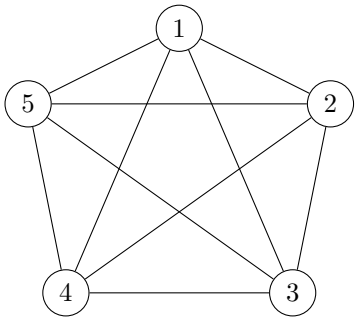


FIGURE 1.4 – Graphe complet K_5

Définition (**Bipartisme**) . Un graphe $G = (X, E)$ est dit biparti s'il existe une partition de X en ensembles X_1 et X_2 non vides et disjoints tels que pour toute arrête (x, y) de G, x et y soient des ensembles différentes.

Remarque G est dit k parti, s'il existe une partition en k ensembles de X vérifiant la définition ci-dessus.

Exemple Soit G un graphe à 5 sommets biparti, alors :

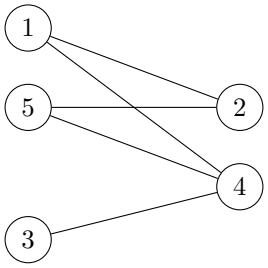


FIGURE 1.5 – Graphe biparti

Exemple (Graphe de Petersen) Sans doute l'un des graphes les plus connus en théorie des graphes, le graphe de Pertersen en hommage à Julius Petersen qui l'étudia en 1898, possédant 10 sommets et 15 arrêtes possède beaucoup de propriétés interressantes (notamment la connexité que nous verrons par la suite). Il est un contre-exemple pour beaucoup de propriétés et est très utile pour vérifier un algorithme en cours de développement ou une intuition.

Propriétés

Propriété () . Soit G un graphe non orienté simple à n sommets.

- Si G est complet, il possède $m = \frac{n(n-1)}{2}$ arrêtes.
- G vérifie donc toujours $m \leq \frac{1}{2}n(n-1)$
- $\sum_{x \in X} d(x)$ est le nombre d'extrémités d'arrêtes, c'est aussi deux fois le nombre d'arrêtes.
- Il y a un nombre pair de sommets de degré impairs.

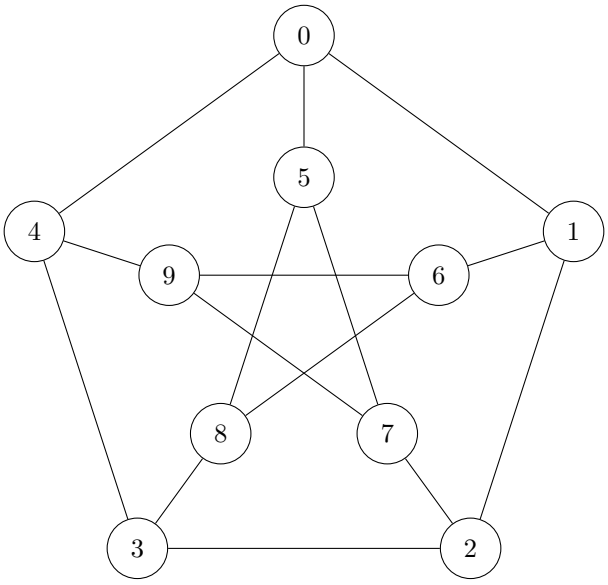


FIGURE 1.6 – Graphe de Petersen

Graphes partiels et sous-graphes

Définition (**Graphe partiel**) . Soit $G = (X, E)$ un graphe. Le graphe partiel $G' = (X, E')$ de G est tel que $E' \subset E$. Autrement dit, le graphe partiel d'un graphe G est le même graphe mais avec quelques arrêtes en moins.

Définition (**Sous-graphe**) . Soit $G = (X, E)$ un graphe. Le graphe $G' = (X', E')$ de G est tel que $X' \subset X$ et $E' = \{(x, y) : x \in X', y \in X', (x, y) \in E\}$. Autrement dit, le sous-graphe d'un graphe G est le même graphe mais avec quelques sommets en moins (et donc quelques arrêtes en moins aussi).

Définition (k-**clique**) . Soit G un graphe. On appelle k-clique, $k \leq n$, un sous-graphe complet de Gde taille k.

Chaines et Cycles d'un graphe non orienté

Définition (**Chaine**) . On appelle chaine de G de longueur n toute suite alternée de sommets et d'arrêtes de G telle que :

$$c = (x_0, a_1, x_1, \ldots, a_n, x_n), \text{ telle que } \forall i \in [\![1, n]\!], a_i = (x_{i-1}, x_i)$$

Ici, n représente le nombre d'arrêtes de la chaine. Dans le cas d'un graphe simple, on notera les chaines de la façon suivante :

$$c = (x_0, \ldots, x_n) \quad \text{ou} \quad c = x_0 - x_1 - \cdots - x_n$$

Définition (**Accessibilité**) . Soit $G = (X, E)$ un graphe. On a :

- Soient x et y deux sommets de G. On dit que y est accessible à partir de x s'il existe une chaine joignant x et y dans G.
- G est dit **connexe** ssi $\forall x \in X, \forall y \in X$, y est accessible à partir de x.
- L'accessibilité est un relation d'équivalence entre les sommets.
  Ses classes d'équivalences sont les composantes connexes de G.

Définition (**Chaîne Simple**) . Une chaine est **simple** si elle ne passe pas deux fois par le même arrête et elle est dite élémentaire si elle ne passe pas deux fois pas le même sommet. On remarquera facilement qu'une chaîne élémentaire est simple.

Définition (**Cycle**) . Un cycle de G est une chaine simple dont le départ et l'arrivée sont le même sommet. Un cycle est donc de la forme :

$$c = x_0 - x_1 - \cdots - x_{n-1} - x0$$

Théorème (**Propriétés des cycles et chaînes**) . Soit G un graphe.

- Toute chaine élémentaire a une longueur inférieure à $n - 1$
- Toute cycle élémentaire a une longueur inférieure à n
- De toute chaîne, on peut en extraire une chaîne élémentaire.

Chemins, circuits d'un graphe orienté et forte connexité

On utilise la même définition de chemin et circuit. Seulement, si le graphe est simple, il sera inutile de préciser les arrêtes par lesquelles on passe.

Définition (**Forte Connexité**) . Un graphe $G = (X, E)$ orienté est dit fortement connexe si pour tout somet x et y de G, y est accessible à partir de x.

Définition (**Arbre**) . On appelle arbre un graphe non orienté, connexe et sans cycle. Un graphe non orienté et sans cycle, i.e une union d'arbres et appelé **forêt**.

Théorème (**Caractérisation d'un arbre**) . Soit T un graphe à n sommets et m arrêtes. T est un arbre ssi :

T est sans cycle et $m = n - 1$
$\iff$ T est connexe et $m = n - 1$
$\iff$ T est sans cycle et maximal au sens des arrêtes
$\iff$ T est connexe et minimal au sens des arrêtes
$\iff$ Deux sommets quelconques de T sont reliés par un unique chemin.

1.1.2 Représentations d'un graphe

Représentation par liste d'arrêtes

Définition (**Liste d'arrêtes**) . On appelle liste d'arrêtes de G la liste des couples (x, y) avec $x \in X$ et $y \in X$ tels que $(x, y) \in E$

Exemple Représentations par liste d'arrêtes/d'arcs

- Le graphe orienté G_1 est représenté par la liste d'arcs :

$$[[1,5],[2,1],[2,4],[3,2],[4,3],[5,2],[5,4]]$$

- Le multigraphe non orienté G_2 est représente par la liste d'arrêtes :

$$[[1,2],[1,5],[1,5],[2,1],[2,4],[2,4],[2,3],[3,2],[3,3],[3,4],[4,2],[4,2],[4,3],[4,5],[5,1],[5,1],[5,4]]$$

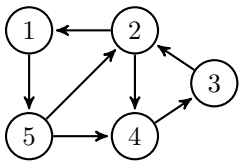


FIGURE 1.7 – Graphe Orienté G_1

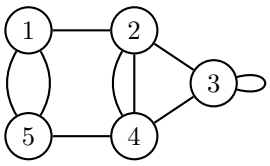


FIGURE 1.8 – Multigraphe non orienté G_2

Représentation Matricielle

Définition (**Matrice d'adjascence**) . On appelle matrice d'adjascence d'un graphe G la matrice $A = (a_{ij})$ telle que $\forall (i,j) \in X \times X$, on ait :

- Si G est non orienté, a_{ij} est égal à 1 si $(i,j) \in E$ et 0 sinon. La mrtice est donc symétrique.
- Si G est orienté, a_{ij} est égal à 1 si $(i,j) \in E$
- Dans le cas d'un multigraphe, le coefficient a_{ij} de la matrice d'adjascence de G représente le nombre d'arrêtes entre les sommets i et j de G. Les coefficients diagonaux de la matrice représentent donc les boucles de G.

Propriété () . Soit G un graphe non orienté, et A sa matrice d'adjascence.

- Puisque G est non orienté, alors $A \in \mathcal{S}_n(\mathbb{N})$ i.e A est symétrique.
- La somme des éléments de la ligne i est dégale à $d^-(i)$
- La somme des éléments de la colonne i esy égale à $d^+(i)$

Exemple (Matrices d'adjascences) La matrice d'adjascence de G_1 :

$$M_1 = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \end{pmatrix}$$

La matrice d'adjascence de G_2 :

$$M_2 = \begin{pmatrix} 0 & 1 & 0 & 0 & 2 \\ 1 & 0 & 1 & 2 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 2 & 1 & 1 & 1 \\ 2 & 0 & 0 & 1 & 0 \end{pmatrix}$$

Remarque (Algèbre Linéaire) Comme dans toutes les représentations matricielles de concepts (que ce soit des applications ou des graphes), elles nous permettent d'invoquer facilement tous les résultats d'algèbre linéaires tels que la réduction, les noyaux, rangs, la composition...tout en préservant les propriétés des objets étudiés.

Propriété () . Soit $A^k, k \in \mathbb{N}$ où A est la matrice d'adjascence d'un graphe G. Alors le coefficient d'indice (x, y) de A^k est le nomdre de chemins de longueur k entre les sommets x et y.

Remarque Toutefois, la représentation matricielle d'un graphe n'est pas optimale pour parcourir ces derniers puisque, ainsi, les algorithmes de parcours auront forcément une complexité temporielle/spatiale en $\Theta(n^2)$.

Représentation par liste d'adjascence

Définition (**Liste d'adjascence**) . La liste d'adjascence d'un graphe G est un vecteur L indexé par X tel que pour tout sommet $x \in X$, $L[x]$ est la liste des successeurs de x dans G.

Exemple Reprenons les graphes G_1 et G_2 représentés plus haut. On a alors :

- $G_1 = \{1 : [5], 2 : [1, 4], 3 : [2], 4 : [3], 5 : [2, 4]\}$
- $G_1 = \{1 : [2, 5, 5], 2 : [1, 3, 4, 4], 3 : [2, 3, 4], 4 : [2, 2, 3, 5], 5 : [1, 1, 4]\}$

Remarque On remarque vite que cette représentation sera plus optinale. Premièrement, le parcours d'un graphe se fera en $\Theta(n)$ et pour chaque sommet, on obtiendra sa liste de successeurs en $\Theta(1)$, si on représente une liste d'adjascence sous forme de dictionnaire en Python.

Propriété () . Pour chaque graphe à n sommets et m arcs, l'espace mémoire utilisé est un $\Theta(n + m) = \Theta(\max\{n, m\})$.

1.1.3 Parcours d'un graphe

Parcours en profondeur (Depth First Search)

Le parcours en profondeur est à voir comme le parcours d'un chien fou dans un labyrinthe. Celui-ci va partir dans un couloir et à chaque intersection va aller sur le chemin le plus proche jusqu'à arriver à une impasse. A chaque impasse, il va revenir sur ses pas pour parcourir les autre chemins les plus proches. Le tout jusqu'à parcourir tout le labyrinthe.
Ainsi, cet algorithme de parcours se concevra de manière récurisive où chaque appel récurisif représentera l'envoi du chien fou dans un chemin (ici une arrête). D'où l'algorithme suivant :

```
1 Visite est initialise a l'ensemble vide
2 Profond (G,x) :
3     x est visite
4     {Traiter x en premiere visite}
5     Pour chaque voisin y de x faire :
6         Si y n'est pas visite, alors :
7             Profond(G,y)
8         {Sinon on detecte une revisite de y}
9     {Traiter x en derniere visite}
```

A partir du parcours en profondeur, on peut définir un ordre de parcours qui sera représenté par une liste de sommets. Nous avons donc l'ordre de parcours en première visite (que l'on mettra à jour ligne 3 de l'algorithme) et l'ordre de parcours en derniere visite (mis à jour ligne 9).
Pour un graphe G non connexe ou non fortement connexe, le parcours en profondeur lancé à partir d'un unique sommet ne permettra pas d'accéder à tous ses sommets. D'où le parcours dit généralisé où l'on itère sur tous les sommets de G et s'il n'est pas visité, on lance un parcours en profondeur à partir de celui-ci.
Ainsi, le parcours en profondeur généralisé d'un graphe G permet de définir une arborescence de parcours.

Définition (**Arborescence de parcours - profondeur**) . L'arborescence de parcours en profondeur de G à partir d'un sommet $x \in X$ est une arbre A enraciné en x, orienté, tel qu'il existe un arc (x, y) dans A ssi l'appel à profond(G,x) a engendré récurisivement un appel à profond(G,y).

A la fin d'un parcours en profondeur généralisé d'un graphe G, on obtient donc une forêt de visite contenant tous les sommets de G, i.e tous les sommets de G ont étés visités.

Parcours en largeur (Breadth First Search)

Le parcours en largeur est fondamentalement différent de celui en profondeur. Tout d'abord, nous pouvons reprendre la comparaison avec le parcours d'un labyrinthe par un chien. Ici, notre chien sera vieux et plus malin. Pour chaque intersection (sommet), il va se rappeler de tous les chemins à proximité (sommets voisins). Il se dirigera donc vers l'intersection (sommet) la plus ancienne qu'il ait retenue pour ensuite lister toutes ses voisinnes, et ainsi de suite. Pour ne pas tourner en rond, à chaque fois qu'il voit une intersection (sommet) voisinne, il vérifie qu'il ne l'aie pas visitée avant de la retenir. Le vieux chien d'arrête donc dès qu'il n'a plus d'intersection (sommet) à visiter en mémoire.
On peut donc implémenter cette algorithme de façons itérative où un ensemble représentera les sommets déjà visités et une file d'attente les sommets à visiter. Le premier sommet de file sera donc le prochain sommet à visiter.

```
1  Visite est initilise a l'ensemble vide
2  Largeur(G,x):
3      F = [X]
4      Visite[x] = vrai
5      Tant que F n'est pas vide, faire :
6          Considerer la tete y de F et l'enlever de F
7          {Traiter y}
8          Pour chaque successeur z de y, faire :
9              Si z n'est pas visite, alors
10                 z est visite
11                 ajouter z a la fin de la file F
```

De même que le parcours en profondeur, on peut définir un ordre de parcours, en première ou derniere visite. Comme le parcours en profondeur, parfois nous auront besoin de lancer un parcours généralisé à tout le graphe pour parcourir tous ses sommets.

Définition (**Arborescence de parcours - largeur**) . L'arborescence de parcours en largeur de G à partir du sommet x est un arbre A enraciné en x, orienté tel qu'il existe un arc (y, z) dans A ssi le traitement de y ajoute le sommet z dans la liste d'attente F.

La tableau Visite est intilisé en $\Theta(n)$. Le parcours est appelé exactement une seule fois pour chaque sommet x de G et a pour complexité $\Theta(1)$ pour le traitement de x et $\Theta(d(x))$ pour l'exploration des successeurs de x. La complexité des deux parcours est donc :

$$\begin{aligned} C(n) &= \Theta(n) + \sum_{x=1}^{n} (d(x) + \Theta(1)) \\ &= \Theta(N) + \sum_{x=1}^{n} d(x) \\ &= \Theta(n) + \Theta(m) \\ &= \Theta(\max\{n, m\}) \end{aligned}$$

Maintenant que l'on sait parcourir un graphe, on peut étudier ses propriétés plus facilement. Voici quelques exemples d'applications interressantes :

Classification des arcs

Une fois le parcours réalisé, on remarque que la forêt de parcours est composée de différents types d'arcs. Soit (x, y) un arc de G, il est dit...

- **Couvrant** ssi (x, y) est un arbre de la forêt de visite.
- **En avant** ssi il existe un chemin de x à y dans la forêt de visite de G.
- **En arrière** ssi il est chemin de y à x dans la forêt de visite.
- **Transverse** ssi ce n'est pas un arc ci-dessus.

La présence de cycle dans un graphe se manifestera par un arc arrière dans l'arbre de visite et un arc avant peut être vu comme un raccourci dans G. Notons que pour chaque parcours, certains arcs sont présents et d'autres non.

Existence de chemins et connexité

Les deux parcours d'un graphe à partir d'un sommet x nous permettent de trouver tous les sommets accessibles à partir de x. Donc on peut facilement adapter un algorithme pour déterminer s'il existe un chemin entre deux sommets x et y de G, et même en trouver.
Dans le cas d'un graphe **non orienté**, si le parcours à partir d'un sommet x atteint tous les autres sommets y de G, alors G est connexe et réciproquement.

"A partir d'ici, vous ne vous perdez plus dans les labyrinthes." Thierry Montaut

1.2 Modélisation et Graphes

1.2.1 Chemins et circuits Eulérien

Les 7 ponts de Königsberg

Le problème des 7 ponts de Königsberg consiste à savoir s'il est possible de déterminer une promenade passant par tous les ponts de la ville de Königsberg en passant une et une seule fois par chaque ponts de la ville et en revenant à son point de départ. Ce problème est sûrement le plus connu de l'histoire de la théorie des graphes et fut résolut par **Leonhard Euler** en 1735. Il explique que le problème n'est pas résoluble pour la ville de Königsberg et établit ainsi l'un des tout premiers théorèmes de théorie des graphes.
La démonstration mathématique du théorème d'Euler ne fut énoncée qu'en 1873 par Carl Hierholzer.
Des problèmes similaires existent tels que celui du facteur chinois qui cherche à effectuer sa tournée de distribution en passant une et une seule fois par chaque rue et en revenant à son point de départ (Mai-Ko-Kwan, 1962).

Définition (**Graphe Eulérien/Circuit Eulérien**) . Soit $G = (X, E)$ un graphe non orienté, un circuit eulérien dans G est un circuit passant une et une seule fois par chaque arrête et revenant au sommet de départ.
Un graphe est dit eulérien ssi il possède un circuit eulérien.

Théorème d'Euler

Théorème (**Théorème d'Euler**) . Soit G un graphe non orienté **connexe**.

- G admet un circuit eulérien ssi tous ses sommets sont de degré pair.
- G admet un circuit eulérien ssi tous ses sommets ont un degré pair seuf deux sommets a et b alors, tous les circuits eulériens de G on a/b comme sommet de départ et b/a comme sommet d'arrivée.

Algorithme de recherche de circuits eulérien

L'objectif de l'algorithme est de construire un circuit eulérien dans un graphe en considérant un graphe partiel du graphe initial. A chaque étape, on cherche un chemin maximal partant du premier sommet non saturé du chemin précédent et qui revient à ce sommet. Puisque il existe un nombre pair de sommets de degré impair il existe un tel sommet à chaque itération.
A la fin de l'algorithme on "recolle" tous les chemins bout à bout en un seul chemin simple non extensible passant par toutes les arrêtes, créant ainsi un circuit eulérien.
L'idée est de créer une "copie" du graphe et, lors de la création d'un chemin à une certaine étape, d'enlever les arrêtes concernées par le chemin pour éviter qu'elles ne soient réutilisées.

1.2.2 Problème de coloration

Les problèmes de colorations de graphes, en plus d'êtres difficiles à résoudre, sont, cependant, très facile à imaginer. L'exemple le plus typique est celui de la coloration d'une carte géographique. En essayant de colorier une carte de l'Europe d'une telle façon que deux pays limitrophes n'aient pas la même couleur, on s'apperçoit vite que cela peut être compliqué, surtout si on essaye d'utiliser le moins de couleurs possibles.
C'est exactement ce que modélisent les problèmes de coloration de graphes. On essaye d'attribuer à chaque sommet une couleur de telle sorte que tous ses voisins n'aient pas la même couleur que lui en utilisant le moins de couleurs possibles.

Définition (k-**coloration**) . Soit G un graphe non orienté. On dit que G admet une k-coloration s'il existe k couleurs différentes telles que deux sommets adjacents de G n'aient pas la même couleur.

Propriété (**Propriétés des colorations**) .

- Si G est k-parti alors G est k-colorable.
- Si G est un graphe complet de taille n alors G ne peut pas être colorié avec moins de n couleurs.
- Si G admet une k-clique, alors G ne peut pas être colorié avec moins de k couleurs.

Définition (**Nombre Chromatique**) . On appelle nombre chromatique $\gamma(G) = k$ d'un graphe G le plus petit entier positif tel que G admette une k-coloration.

Coloration Naïve

Premier algorithme de coloration de graphe, il est très facile à comprendre et à mettre en oeuvre mais est particulièrement gourmand en couleurs.
L'idée est de parcourir le graphe dans l'ordre naturel des sommets et d'attribuer à chaque sommet la plus petite couleur non attribuée à ses voisins.
Il faut donc une fonction permettant de parcourir le graphe et une autre déterminant la plus petite couleur non attribuée aux voisins d'un sommet. La coloration est représentée par un dictionnaire dont les clés sont les sommets du graphe et les valeurs, les couleurs attibuées aux sommets.
C'est un algorithme très peut coûteux d'une complexité en $\Theta(m)$ mais pas très optimal, puisque en fonction de l'ordre de visite des sommets, le nombre de couleurs utilisées peut varier. Il est très peu efficace pour de gros graphes.

Coloration Gloutonne

Prenez vos crayons de couleurs préférés et essayer de colorier un graphe. Vous alez d'abord colorier tous les sommets possibles avec un certaine couleurs. Puis une fois que vous ne pouvez plus colorier de sommets de cette couleur tel qu'aucun de ses voisins n'est déjà colorié vous prenez une autre couleur et refaites de même. Le tout jusqu'à ce que tous les commets soient coloriés.
C'est le principe de l'algorithme de coloration Gloutonne. A chaque étape, on choisit une couleur et on essaye de colorier le maximum de sommets non adjacents avec cette même couleur. Pour cela, nous allons considérer le noyau d'une graphe.

Définition (**Noyau**) . Soit $G = (X, E)$ un graphe. On appelle noyau de G un ensemble maximal de sommets non adjacents deux à deux.

On va donc déterminer tous les noyaux de G en détruisant petit à petit le graphe. Et, pour chaque noyau de trouvé, on colorie tous les sommets du noyau avec la même couleur.

1.2.3 Ordonnancement

Ici, un graphe représente un système de dépendances entre différentes tâches d'un même projet. Chaque sommet repésente une tâche à effectuer et une arrête d'un sommet x vers un sommet y indique que la tâche y doit attendre que la tâche x est terminée avant d'être commencée.
On va donc essayer d'établir un graphe d'ordonnancement des tâches visant à indiquer dans quel ordre les différentes tâche devront être traitées. Mais pour cela, il faut définir certaines conditions sur le graphe de dépendances.

Proposition Soit $G = (X, E)$ un graphe.

- Si G est sans circuit, alors tous ses sous-graphes sont sans circuits.
- Si G est sans circuit, le graphe inverse de G est sans circuit.
- G est sans circuit ssi tous ses chemins sont élémentaires.

Définition (**Sommets remarquables d'un graphe sans circuit**) . Soit $G = (X, E)$ un graphe orienté sans circuit. On définit :

- une **source** comme étant un sommet dont le degré entrant est nul.
- un **puit** comme étant un sommet dont le degré sortant est nul.

Théorème (**Fondamental de l'ordonnancement**) . Tou graphe sans circuit possède une source et un puit.

Tri Topologique - Ordonnancement séquentiel

Définition (**Tri Topologique**) . On appelle tri topologique d'un graphe orienté une numérotation des des sommets respectant l'ordre des arcs.

Théorème (**Condition d'existence d'un tri topologique**) . Soit $G = (X, E)$ un graphe orienté. G admet un tri topologique ssi il est sans circuit.

Le principe de l'algorithme de tri topologique est **itératif**. Il repose sur le fait que tout graphe sans circuit **possède une source**. On va donc, à chaque étape, chercher une source du graphe et l'insérer dans notre tri topologique. L'étape suivante, on **considère le sous-graphe** du graphe initial auquel on a enlevé la source et tous les arcs sortants.
Si on ne souhaite pas détruire le graphe, on va considérer les **degrés entrants** de chaque sommets dans un **dictionnaire** des degrés.

```
1  S:=[];T:=[];
2  Pour x de 1 a n faire
3      Degre[x]:=d^-(x) dans G;
4      Si Degre[x]=0 alors ajouter x a S;
5  {S contient alors toutes les sources de G}
6
7  Tant que S<>[] faire
8      x:=enleverTete(S);
9      ajouterFin(x,T)
10
11     {Mettre a jour les degres}
12     Pour chaque successeur y de x dans G faire
13         Degre[y]:=Degre[y]-1;
14         si Degre[y] = 0 alors ajouterFin(y,S)
```

L'algorithme repose donc sur un **parcours en largeur**. On utilise une **file** pour stocker les sommets de degré entrant nul (i.e les sources). A chaque itération, on décrémente tous les degrés entrants des fils de la source, puis on ajoute la source à la liste représentant le tri topologique.

Tri par Niveaux - Ordonnancement en parallèle

Ici, comparé à un tri topologique, on va construire un graphe d'ordonnancement en considérant que l'on peut effectuer plusieurs tâches en même temps.

Définition (**Tri par niveaux**) . Soit $G = (X, E)$ un graphe orienté, sans circuit. Un tri par niveaux de G est une partition de G en niveaux tels que pour chaque sommet d'un même niveaux, son exécution n'est pas dépendante d'un sommet du même niveau ou d'un niveau suivant.

L'algorithme est sensiblement le même que le tri topologique. Il suffit juste de l'adapter pour **traiter en même temps** toutes les sources de G. On définit alors **deux listes N1 et N2** représentant deux niveaux successifs. Le traitement des sources de N1 permet de déterminer les sources de N2.

```
1  N1:=[];T:=[];
2  Pour x de 1 a n faire
3      Degre[x]:=d^-(x) dans G;
4      Si Degre[x]=0 alors ajouter x a N1;
5  {S contient alors toutes les sources de G}
6
7  tant que N1<>[] faire
8      ajouterFin(N1,T);
9      N2:=[];
10     pour tout x dans N1 faire
11         {Mettre a jour les degres des successeurs de x et calcul}
12         Pour chaque successeur y de x dans G faire
13         Degre[y]:=Degre[y]-1;
14         si Degre[y] = 0 alors ajouterFin(y,N2)
15     N1:=N2;
```

1.2.4 Arbre couvrant de poids minimum

Ici, on considère un **graphe non orienté valué** pour lequel on va chercher un sous-arbre couvrant de poids minimal. Cela peut éventuellement modéliser le problème d'un électricien qui doit relier différentes pièces par un câble et cherche à utiliser le moins de câble électrique possible.

Définition (**Valuation**) . Soit $G = (X, E)$ un graphe non orienté. Une valuation est une fonction

$$p : E \longrightarrow \mathbb{R}$$

qui, à chaque arrête lui associe une valeur appelée poids.
Un graphe muni qu'un valuation est appelé graphe valué.

Pour **représenter** un graphe valué, nous allons toujours utiliser une liste d'adjacence. Mais nous allons rajouter une matrice de poids représentée par un dictionnaire dont les clés sont les arrêtes (couple) et la valeur, le poids de l'arrête correspondante.
On pourra aussi représenter un graphe valué par une liste d'arrêtes composée de triplet représentant les deux sommets de l'extrémité de l'arrête et le poids de l'arrête.

Définition (**Arbres couvrant**) . Soit $G = (X, E)$ un arbre graphe non orienté valué. Un arbre couvrant de G est un graphe partiel de G, $A = (X, E')$ qui soit un arbre.

Théorème (**Condition d'existence**) . Soit $G = (X, E)$ un graphe non orienté valué. G admet un arbre couvrant ssi il est connexe.

Algorithme de Kruskal

L'algorithme de Kruskal construit de **manière itérative** un arbre couvrant de poids minimal. Pour cela, on va considérer les arrêtes par ordre de poids croissant. A chaque itération, on va essayer d'ajouter **la plus petite arrête ne créant pas de cycle** à l'arbre couvrant.
On s'arrête lorsque l'on a ajouté $n - 1$ arrêtes.

Remarque (Condition d'arrêt et vérification)

- A chaque itération, il faut vérifier que le graphe construit de possède pas de cycle, il faut donc effectuer un **parcours en largeur**.
- Si à une itération, on ne trouve pas d'arrête qui convienne, c'est que le graphe n'était initialement pas connexe.

Algorithme de Prim

L'algorithme de Prim cherche, contrairement à Kruskal, à **conserver la connexité à moindre coût**.
On va donc considérer deux ensembles :

- C : l'ensemble des arrêtes de T (arbre couvrant)
- M : le complémentaire de C dans X

A chaque étape, on va donc chercher une arrête **joignant un sommet de C à un sommet de M dans G.** Un arrête joignant deux composantes connexe ne pouvant pas créer de cycle, le graphe T reste bien un arbre.
On s'arrête lorsque **T contient tous les sommets de G**.

1.2.5 Plus courts chemins dans un graphe valué

On considère ici un **graphe orienté valué**. L'objectif est de trouver un chemins entre deux sommets x et y dans G de poids minimal. On retrouve le même problème pour le routage de paquets dans un réseau. En effet, lors de l'établissement des tables de routage, on va chercher le chemin le plus court (au sens de la durée de transfert) entre deux noeuds.

Définition (**Cycle Absorbant**) . Dans le cas d'un graphe orienté, valué à **poids négatifs**, on appelle cycle absorbant un cycle de coût total négatif.

Remarque Le passage par un cycle absorbant dans un tel graphe fait diminuer le coût du chemin de poids minimal. On comprend alors vite qu'un graphe avec un circuit absorbant ne possède pas de chemin de poids minimal.

Théorème (**Condition d'existence**) . Soit $G = (X, E)$ un grapge orienté valué. Soient x et y deux sommets de G. Alors il existe un chemin de poids minimal entre x et y ssi

- y est accessible à partir de x
- G ne possède pas de circuit absorbant

Algorithme de Dijkstra

L'algorithme de Dijkstra 1 se base sur un **parcours en largeur** du graphe. On cherche tous les plus courts chemins d'un sommet s vers les autres sommets du graphe. Pour cela on dispose d'un **dictionnaire des poids**, et d'un **dictionnaire des pères**.
A l'initialisation, on commence par parcourir tous les sommets du graphe à la recherche d'arrêtes partant de s. On va construire pas à pas un plus court chemin vers chaque sommet. A chaque étape, on considère tous les "nouveaux" chemins vers les sommets du voisinnage du sommet traité. Si on trouve un chemin plus court, on remplace le chemin actuel et on remet le sommet dans la file. Dans le cas contraire, on ne fait rien.

```
1  M:={s};
2  {Initialisation}
3  Pour i de 1 a n faire
4      Si (s,i) est une arete alors
5          D[i]:=cout(s,i);
6          P[i]:=s;
7          ajouter(i,M);
8      Sinon D[i]:=infini;
9
10 {Algorithme}
```

```
Tant que M<>{} faire
    x:=enleveTete(M);
    Pour tout successeur y de x faire
        d:=D[x]+cout(x,y);
        si d<D[y] alors
            D[y]:=d;
            P[y]:=x;
            ajouter(y,M)
```

L'algorithme de Dijkstra possède quand même quelques inconvénients. Un même sommet peut se retrouver plusieurs fois dans la file et on visite donc tous ses voisins plusieurs fois. Dans le cas **d'arrêtes à poids tous positifs**, on peut éviter tous ces parcours inutiles choisissant à chaque étape le sommet dont le chemin depuis s est de coût minimal.

Dijkstra Opt

A chaque fois que l'on récupérer un sommet dans la file, on va **récupérer le sommet de poids minimal.** Pour cela, on va **garder la file triée selon le poids des sommets**. Il faut donc implémenter un fonction `insere` qui ajoute un sommet dans la file en fonction de son degré d'accessibilité.

```
{Initialisations}
M:={};
Pour i de 1 a n faire
    Si (s,i) est une arete alors
        D[i]:=cout(s,i);
        P[i]:=s;
        ajouter(i,M);
    Sinon D[i]:=infini;

{Algorithme}
Tant que M<>{} faire
    x:=choisir_min(M,d);
    enlever(x,M);
    Pour tout successeur y de x faire
        Si y est dans M alors
            d:=D[x]+cout(x,y);
            si d<D[y] alors
                D[y]:=d;
                P[y]:=x;
                ajouter(y,M)
```

La **complexité** de l'algorithme de Dijkstra est donc en

$$\boxed{\Theta(n^2) + \Theta(m) = \Theta(n^2)}$$

Algorithme de Floyd

Cet algorithme permet de calculer **tous les plus courts chemins** de tous les sommets vers tous les autres. Il a un **complexité** en $\Theta(n^3)$, donc il reste très efficace.
L'algorithme repose sur une **idée itérative**. On commence donc par calculer tous les plus courts chemins de i vers j (deux voisins) sans intermédiaires. Puis on calcule tous les plus courts chemins avec **un intermédiaire.**

...

Arrivé à la n-ième itération, on a tous les plus courts chemins du graphe.

Par **récurence**, pour passer de l'étape k à l'étape k+1 :

$$\forall (i,j) \in \{1,n\}^2, \quad d = pcc[(i,k)] + pcc[(k,j)]$$
$$\text{si } d < pcc[(i,j)] :$$
$$pcc[(i,j)] = d$$

On utilise une **matrice de poids** (pcc) et un **dictionnaire des pères** P.

Algorithme de Bellman (poids quelconques)

Ici, on cherche les plus courts chemins d'un graphe orienté valué avec potentiellement des poids négatifs.
L'idée est que si on connaît tous les plus courts chemins d'un sommet s vers tous les prédécesseurs d'un sommet y de G, alors, le plus court chemin de s vers y est :

$$min(\{d_i + p_i, \quad i \in \{i,n\}\}) =: d_{i0} + pi0$$

et le prédécesseur de y dans le plus court chemin sera x_{i0}.
A **l'initialisation**, on connaît tous les plus courts chemins de s vers lui-même. On considère donc G' le **sous-graphe** de G constitué des sommets dont one ne connaît pas encore de plus courts chemins. G' est un sous-graphe d'un graphe sans cycle donc il admet toujours des **sources**. Une source de G' est un sommets de G' dont on connaît tous les prédécesseurs. On peut donc calculer son plus courts chemin comme vu précédement et il sort de G'.
On va construire un **dictionnaire des distances** `Dist` et un **dictionnaire des pères** dans le plus court chemin `Pred`. Pour l'algorithme, on utilise un dictionnaire représentant le nombre de prédécesseurs inconnus d'un sommet.

```
# Sort de G'
Pour x dans G[y] faire :
    Deg[x] -= 1
    si Deg[x] == 0 :
        ajouter(x,S)
```

```
# calcul du pcc vers y
{Soit H la liste des predecesseurs dans G}
Pour x dans H[y] faire :
    si Dist[x] + P[(x,y)] < D[y] faire :
        D[y] = Dist[x] + P[(x,y)]
        Pred[y] = x
```

Pour représenter le graphe on a donc besoin :

- La liste d'adjacence de G.
- Le dictionnaire des pères.
- La matrice de poids représentée par un dictionnaire.

L'algorithme se base donc sur un **parcours en largeur**, d'où :

```
{Initialisation}
...
{Algorithme}
Tant que S<>[] faire :
    y = tete de S
    # calcul pcc vers y
    # y sort de G'
```

Cet algorithme a donc un **complexité** en :

$$\Theta(m+n)$$

Chapitre 2

Mots et Langages

Contents

Si on connaît plusieurs langages de programmation, on remarque que chaque langage, ou plutôt chaque paradigme de langage est spécialisé pour la résolution d'une carégorie dde problèmes. On pourrait se demander s'il est possible de créer un langage permettant de résoudre tous les problèmes. Pour cela, il nous faudrait d'abord être capable de définir formellement un langage, des mots, etc...

2.1 Alphabets et Mots

2.1.1 Premières Définitions

Commençons tout d'abord par redéfinir correctement la notion d'alphabet et de mot.

Définition (**Alphabet**) . Un alphabet est simplement un ensemble fini, noté Σ. On nomme "lettre" ou "symbole" les éléments d'un alphabet.

Exemple Quelques exemples d'alphabets :

- $\Sigma = \{a, b\}$
- $\Sigma = \{a, b, \ldots, \%, \$\}$

Définition (**Mot**) . Un mot est une suite finie de lettres d'un alphabet.

Proposition Le mot vide est noté ε. On note l'ensemble des mots d'un alphabet Σ^*.

Définition (**Longueur d'un mot**) . On note $|w|$ la longueur d'un mot $w \in \Sigma^*$ qui correspond au nombre de lettres, avec répétition du mot w.

Définition (**Egalité de mots**) . On dit que deux mots $u, v \in \Sigma^*$ son égaux ssi :

- $|v| = |u|$
- $\forall i \in [\![1, |v|]\!], \quad u[i] = v[i]$ où $u[i]$ est la ième lettre du mot u.

Deux mot sont égaux ssi ils sont de même longueur et sont composés des mêmes lettres dans le même ordre.

2.1.2 Opérations sur les mots

Sur les mots, on ne définit qu'une seule opération, la **concaténation**.

Définition (**Concaténation**) . Soient $u, v \in \Sigma^*$ deux mots définis sur un même alphabet. On appelle la concaténation l'application :

$$\begin{cases} \Sigma^* \times \Sigma^* \longrightarrow \Sigma^* \\ (u, v) \longmapsto w = u.v \end{cases}$$

Elle est définie telle que $\forall u, v \in \Sigma^*$ de longueur $n, p \in \mathbb{N}$, on ait :

- $|u.v| = |u| + |v| = n + p$
- $\forall i \in [\![1, n]\!] u.v[i] = u[i]$ et $\forall i \in [\![1, p]\!], u.v[n + i] = v[i]$

On parlera identiquement de concaténation ou de produit.

Remarque Cette définition reprend bien la caractérisation de deux mots égaux.

Proposition (Propriétés de la concaténation) La concaténation est une application :

- **Associative :** $\forall u, v, w \in \Sigma^*, \quad w.(u.v) = (w.u).v$
- **Pas commutative** pour un alphabet de plus d'une lettre.
- admet pour **élément neutre** le mot vide ε.

2.1.3 Puissance d'un mot

Une fois la concaténation définie pour un mot, on peut alors parler de puissance de mot. Définissons celle-ci par récurrence.

Définition (**Puissance d'un mot**) . Soit Σ un alphabet et $u \in \Sigma^*$, on a :

- $u^0 = \varepsilon$
- $u^1 = u$
- $\forall n \in \mathbb{N}, u^{n+1} = u^n.u$

Exemple $(ba)^3 = bababa$

Proposition Soient $u \in \Sigma^*$ on peut appliquer les règles "connues" des puissances d'où :

$$\forall n, p \in \mathbb{N}, \quad u^{n+p} = u^n.u^p = u^p.u^n$$

On remarque que l'on peut effectuer des simplifications sur les égalités de mots.

Propriété (**Simplifications**) . L'ensemble Σ^* est simplifiable à gauche et à droite.

- $\forall u, v, w \in \Sigma^*, \quad u.w = v.w \Longrightarrow u = v$
- $\forall u, v, w \in \Sigma^*, \quad w.u = w.v \Longrightarrow u = v$

Ici, pas besoin d'inverse, la démonstration repose sur la définition de l'égalité entre deux mots.

2.2 Relations d'Ordre

Dans l'alphabet dit "classique" on possède un ordre lexicographique des mots permettant de les classer en fonction de leurs lettres et de la position de leurs lettres. Ici, nous allons définir deux types de relations d'ordre sur les mots.
Commençons par rappeler la définition de relation d'ordre.

Définition (**Relation d'Ordre**) . Soit $\lhd$ une relation sur un ensemble E. On dit que $\lhd$ est une **relation d'ordre** ssi pour tout $x, y, z \in E$, $\lhd$ est :

- **Réflexive :** $x \lhd x$
- **Anti-Symétrique :** $x \lhd y$ et $y \lhd x \Longrightarrow x = y$
- **Transitive :** $x \lhd z$ et $z \lhd y \Longrightarrow x \lhd y$

2.2.1 Ordre Préfixe

Naturellement, on munitΣ^* d'un ordre préfixe permettant de classer les mots en fonction de leur préfixe. Cette relation peut être vue comme un forme d'inclusion de mots.

Définition (**Ordre Préfixe**) . Soient $u, w \in \Sigma^*$, on définit la relation d'ordre préfixe $\sqsubseteq$ telle que :

$$\boxed{u \sqsubseteq w \iff \exists v \in \Sigma^*, w = u.v}$$

Autrement dit, u est un préfixe de w ssi il existe un mot v tel que w soit composé de la concaténation de u et v.

Remarque L'ordre préfixe ne nécessite pas de relation d'ordre directement sur l'alphabet Σ.

Propriété (**Ordre Préfixe et égalité**) . Soient $u, v \in \Sigma^*$ on a :

$$u \sqsubseteq v \text{ et } v \sqsubseteq u \Longrightarrow u = v$$

Démonstration Soient $u, v \in \Sigma^*$ tels que $\exists x, y \in \Sigma^*$ tels que

$$u = v.x \quad \text{et} \quad v = u.y$$

On a alors que :

$$\begin{cases} v = v.y.x \\ yx = \varepsilon \end{cases} \Longrightarrow \begin{cases} y = \varepsilon \\ x = \varepsilon \end{cases} \Longrightarrow u = v$$

Remarque Attention : l'ordre préfixe est une relation d'ordre partielle. Autrement dit, tous les éléments d'un même alphabet de sont pas comparables.

2.2.2 Ordre Lexicographique

Définition (**Ordre Lexicographique**) . Soit Σ un alphabet que l'on muni d'une relation d'ordre $\leqslant$. L'ordre lexicographique $\leqslant$ est une relation d'ordre totale sur Σ^*.

Remarque Ici, nous avons bien besoin de définir au préalable un ordre sur notre alphabet Σ.

Propriété (**Compatibilité**) . L'ordre lexicographique est compatible avec l'ordre préfixe. Plus formellement,

$$\forall u, v \in \Sigma^*, \quad u \sqsubseteq v \Longrightarrow u \leqslant v$$

2.3 Langage

Maintenant que nous sommes au clair sur la définition de lettre et de mot, on peut enfin définir l'objet principal de ce chapitre, les langages.

2.3.1 Definition

Définition (**Langage**) . Soit Σ un alphabet, on appelle langage sur Σ toute partie de Σ^*.

Remarque L'ensemble de tous les langages d'un alphabet Σ est donc $\mathcal{P}(\Sigma^*)$, l'ensemble de toutes les parties de Σ^*.

Définition (**Complémentaire**) . Soit L un langage sur Σ. On définit le complémentaire de L dans Σ^* le langage :

$$\overline{L} = \{w,\ w \notin L\}$$

2.3.2 Opérations

De même que pour les alphabets et les mots, on peut définir des opérations sur les langages.

Définition (**Opérations sur les langages**) . Soit Σ un alphabet et $L_1, L_2 \subseteq \Sigma^*$ deux langages de Σ. On définit 4 principales opérations sur des langages :

- **Somme :** notée $+$, la somme de deux langages d'apparente à l'union des ensembles.

  $$\boxed{L_1 + L_2 = \{w,\ w \in L_1 \text{ ou } w \in L_2\}}$$

  C'est une opération :
  - Commutative
  - Associative
  - dont $\emptyset$ est le neutre.
- **Intersection :** de même que pour les ensembles :

  $$\boxed{L_1 \cap L_2 = \{w,\ w \in L_1 \text{ et } w \in L_2\}}$$

  C'est une opération :
  - Commutative
  - Associative
  - dont Σ^* est le neutre

- **Différence :** comme les ensembles, on définit la différence de langages :

$$\boxed{L_1/L_2 = \{w,\ w \in L_1 \text{ et } w \notin L_2\} = L_1 \cap \overline{L_2}}$$

- **Produit de concaténation :** de même que pour les mots, on peut généraliser le produit de concaténation aux langages :

$$\boxed{L_1.L_2 = \{u.v,\ u \in L_1, v \in L_2\}}$$

  C'est une opération :
  - Associative
  - Distributive par rapport à l'union
  - D'élément neutre $\{\varepsilon\}$.

Définition (**Puissance de langage**) . Soit L un langage, on définit **par récurrence** la puissance de L par :

- $L^0 = \{\varepsilon\}$
- $L^1 = L$
- $\forall n \in \mathbb{N}^*,\ L^n = L^{n-1}.L$

Une fois définies des opérations "simples" sur les langages, on peut en définir des plus complexes, permettant de "générer" un langage infini à partir d'un langage fini ou infini.

Définition (**Langage plus et étoile**) . Soit L un langage sur un alphabet Σ. On définit le langage plus de L comme le langage :

$$L^+ = L^1 + L^2 + \dots$$

De même le langage étoile de L est défini par :

$$L^* = \{\varepsilon\} + L^1 + L^2 + \dots$$

2.3.3 Propriétés

Voyons quelques propriétés des langages...

Proposition Soient L_1 et L_2 deux langages sur un alphabet Σ, on a les propriétés suivantes :

- $\forall p \in \mathbb{N}$, on a :
$$(L_1)^p.(L_2)^p \subseteq (L_1)^*.(L_2)^*$$
- L'opération étoile est idempotente :
$$(L^*)^* = L^*$$
- $L^* = \{\varepsilon\} + L^+$
- $\varepsilon \in L \iff L^+ = L^*$

2.3.4 Expressions Régulières

Lorsque l'on manipule des langages infinis, il serait appréciable d'avoir une expression pratique pour un langage permettant de directement voir la forme des mots qu'il contient. On définit ainsi les expressions régulières.

Définition (**Expression Régulière**) . On définit récurisevement une expression régulière sur un alphabet Σ :

- ε est une expression régulière.
- $\forall w \in \Sigma$ est une expression régulière.
- Si E est un expression régulière alors (E) l'est aussi.
- Si E_1 et E_2 sont des expressions régulières, alors $E_1 + E_2$ l'est aussi.
- Si E_1 et E_2 sont des expressions régulières, alors $E_1.E_2$ l'est aussi.
- Si E est une expression régulière, alors E^* l'est aussi.

Exemple Voyons quelques exemples d'expressions régulières sur un alphabet $\Sigma = \{a, b\}$:

$$a^*b, \quad (a+b)^*, \quad (a+b)^*ba(a+b)^*$$

Définition (**Langage Régulier**) . On dit qu'un langage est régulier si il peut s'écrire sous la forme d'une expression régulière.

Il sera donc préférable de travailler avec des langages réguliers.

2.4 Langage Décidable

L'objectif de ce cours est bien entendu de comprendre comment fonctionne un compilateur, pour pouvoir en créer un par nous même. Pour rappel, on doit d'abord bien comprendre les notions de langage et de mot pour pouvoir ensuite déterminer si un ensemble de mots est syntaxiquement corrects lors de la compilation.
Lors de la compilation, il faut d'abord commencer par savoir si un mot traité appartient au langage défini ou pas. Pour des langages finis, l'opération n'est pas compliquée, pour chaque mot il suffit de vérifier si il appartient à un ensemble fini. Pour des langages infinis, l'opération semble plus complexe, il va falloir trouver une manière systématique et efficace de définir si un mot appartient au langage ou pas.
On appelle ce genre de problème un problème de **décision**.

Définition (**Langage Décidable**) . Un langage L est dit décidable si il existe un algorithme permettant de dire si un mot w appartient ou pas au langage L.

Théorème (**Nombre de Langages Décidables**) . Il existe un nombre fini de langages décidables.

Autrement dit, il existe un nombre infini de langages non décidables...

Chapitre 3

Automates (AFD, AFN, AFε)

Contents

Comme abordé dans le chapitre précédent, on cherche une méthode pratique et efficace pour déterminer si un mot appartient à un langage ou pas. On veut donc un modèle qui soit d'une part très pratique mathématiquement pour nous permettre de démontrer des choses dessus mais aussi facilemet implémentable algorithmiquement.
Alerte Spoiler : de solides connaissances en théorie des graphes seront plus qu'utiles...

3.1 Automates fini déterministes

3.1.1 Définition et représentation

Définition (**Automate fini déterministe**) . Un Automate Fini Déterministe est un quintuplet :

$$\boxed{\mathcal{A} = (\Sigma, Q, T, q_0, A)}$$

où :

- Σ est un alphabet

- Q est un ensemble fini d'états (souvent une partie finie de $\mathbb{N}$)
- $T : Q \times \Sigma \longrightarrow Q$ est une application qui, à un état et une lettre associe un autre état.
- q_0 un état initial
- $A \subseteq Q$ les états acceptants

On représentera ainsi un automate fini déterministe de plusieurs façons en fonction de son utilisation :

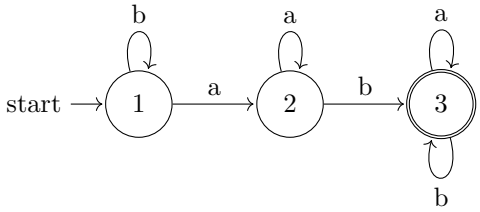
- **Mathématique :** $\mathcal{A} = (\Sigma, Q, T, q_0, A)$
- **Table de Transition :** Elle va permettre de trouver rapidement les différents types d'états.
- **Sagitale :** Sous forme de graphe
- **En Python :** Nous représenterons les automates finis déterministes sous la forme de quintuplet aussi.

Regardons en détail ces différentes représentations au travers d'un exemple.

Exemple Soit $\mathcal{A} = (\Sigma, Q, T, q_0, A)$ un automate fini déterministe.
Mathématiquement nous avons :

- $Q = \{1, 2, 3\}$
- $\Sigma = \{a, b\}$
- $q_0 = 1$
- $A = \{3\}$

La représentation **sagitale** de notre automate sera :



On représente de façon doublement cerclée les états acceptants. Les états sont les sommets du graphe. Les arcs valués sont les antécédants/images de la fonction T.
Enfin, la **table de transition** de l'automate est représentée par le tableau suivant :

Q/Σ	a	b
1	2	3
2	2	3
3	3	3

La première ligne présente les lettres de l'alphabet et la première colonne les différents états. Pour chaque état, le tableau donne l'état obtenu en fonction de la lettre suivante lue.

3.1.2 Mot et Langage Automatique

Définition (**Lecture d'un mot**) . Soit Σ un dictionnaire, $l \in \Sigma$ et $\mathcal{A}$ un automate. On lit la lettre l en dérivant d'un état $q \in Q$ vers un état $q' \in Q$ et si $T(q, l) = q'$. On notera

la lecture d'un mot de longueur p par la lecture successive de ses lettres :

$$q_1 \xrightarrow{l_1} q_2 \ \dots \ q_{p-1} \xrightarrow{l_p} q_p$$

Concrètement, pour la lecture d'un mot, on va partir de l'état initial et en fonction des valeurs de l'état courant et de la lettre lue, on va "bouger" d'un état (sommet) à un autre.

Définition (**Mot refusé**) . Un mot $w \in \Sigma^*$ est dit refusé par un automate $\mathcal{A}$ si sa lecture à partir de l'état initial se termine sur un état refusant ou ne se termine pas. Dans le cas contraire, w est dit accepté.

Définition (**Langage d'un Automate**) . Le langage d'un automate $\mathcal{A}$ est l'ensemble des mots acceptés par l'automate. On le note $L(\mathcal{A})$.

On parle de **langage automatique** si il est reconnaissable par un automate. Deux automates sont dits **équivalents** si ils reconnaissent le même langage.

Définition (**Automate Complet**) . Un automate $\mathcal{A}$ est dit complet si

$$\forall i \in Q, \forall l \in \Sigma, \quad T(i,l) \text{ est défini}$$

Autrement dit, un automate est dit complet si pour toute lettre et pour tout état fixés, il est possible de changer d'état dans l'automate.

Définition (**Puit**) . Un état $q \in Q$ est un puit ssi

$$\forall l \in \Sigma, \quad T(q,l) = q$$

Un puit est un état duquel on ne peut sortir.

On définit aussi la notion de piège comme un puit refusant (i.e un puit dont l'état est refusant). Puisque la notion de complémentaire existe pour les langages et que les automates semblent très étroitement liés aux langages, on peut se demander si un automate peut admettre un complémentaire...
Soit $\mathcal{A}$ un automate associé à un langage L. On cherche $\mathcal{A}' = \overline{\mathcal{A}}$.

$$w \in \overline{\mathcal{A}} \iff w \notin L \iff w \notin L(\mathcal{A})$$

Il semble falloir que $\mathcal{A}$ soit complet. Si c'est le cas, on pourrait inverser $\mathcal{A}$ en inversant les sommets accpetants/refusants.

Proposition Tout automate fini peut être complété par des puits refusants en un automate complet.

Théorème (**Complémentarité**) . L'ensemble des automates est stable par complémentarité.

3.2 Automates fini non déterministes

Dans la section précédente nous avons vu un modèle très efficace pour vérifier l'appartenance d'un mot à un langage. En plus d'être facilement représentable en mémoire (i.e Python), il est

facile à utiliser à la main et hérite de toute la théorie des graphes vue précédement ce qui en fait un très beau modèle mathématiquement parlant.
Malgré tout cela, nous ne savons pas comment, à partir de plusieurs langages simples, constituer un automate reconnaissant la somme de ces langages. Nous n'avons pas défini de somme/union d'automate et celles-ci semblent assez difficiles vu la rigidité de notre modèle.
Nous allons donc construire un modèle d'automate, appelé non déterministe, nous permettant de faire ces opérations d'union (que nous appelerons juxtaposition). Elles nous permettrons de construire des automates complexes à partir de somme de langages.

3.2.1 Généralités

Définition (**AFN**) . Un automate fini non déterministe $\mathcal{A}$ est un quintuplet :

$$\mathcal{A} := (Q, \Sigma, T, I, A)$$

tel que :

- Q est l'ensemble des états de l'automate
- Σ est un alphabet
- $T : Q \times \Sigma \longrightarrow \mathcal{P}(Q)$ est une application
- $I \subseteq Q$ est l'ensemble des états initiaux
- $A \subseteq Q$ est l'ensemble des états acceptants.

Remarque Plusieurs remarques concernant ce nouveau modèle. Premièrement, on remarque que l'on peut maintenant définir des transitions multiples entre les états de l'automate. Deuxièmement, il existe plusieurs états initiaux.

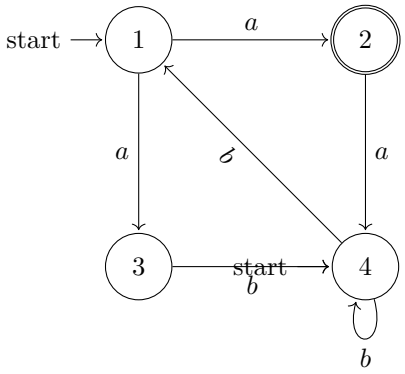
Définition (**Arbre de lecture**) . Soient $w \in \Sigma^*$, L un langage sur Σ et $\mathcal{A}$ un automate reconnaissant L. L'arbre de lecture de w par $\mathcal{A}$ est l'arbre résultat du parcours de $\mathcal{A}$ en fonction des lettres de w.
Autrement dit dans l'arbre de lecture G de w, un noeud a est le fils d'un noeud b si il existe une lettre l de w telle que $T(b, l) = a$. Une feuille de cet arbre est un état acceptant ou refusant de l'automate.

Définition (**Lecture acceptante**) . Soient $w \in \Sigma^*$, L un langage sur Σ et $\mathcal{A}$ un automate reconnaissant L. Une lecture de w par $\mathcal{A}$ est dite acceptante si il existe un chemin d'un état initial vers un état acceptant dans l'arbre de lecture de w par $\mathcal{A}$.

Proposition On peut dire plusieurs choses de la lecture d'un mot $w \in \Sigma^*$ par un automate $\mathcal{A}$:

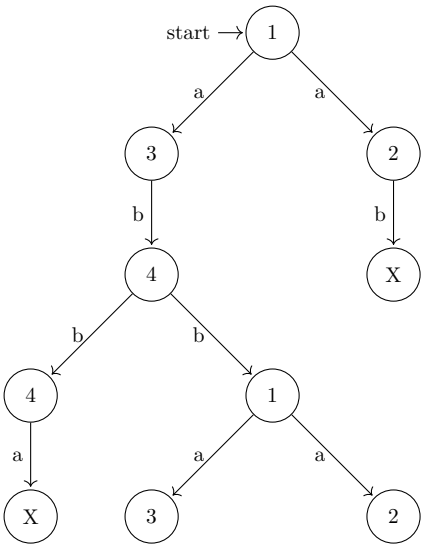
- Si l'automate possède plusieurs états initiaux, la lecture produit un arbre de lecture pour chaque état initial, nous auront donc une forêt d'arbres de lecture.
- Une lecture sera donc acceptante ssi il existe un arbre de la forêt dont au moins une des feuilles est un état acceptant.

Exemple Voyons tout cela sur un exemple. Soit $\mathcal{A} := (Q, \Sigma, T, I, A)$ un automate fini non déterministe $\mathcal{A}$ sur l'alphabet $\Sigma : \{a, b\}$. Représentons notre automate sous forme de graphe orienté valué et sa table de transition :



T	a	b
1	{2,3}	X
②	4	X
3	X	4
4	X	{1,4}

C'est un automate non déterministe puisqu'il contient deux transitions mutliples et deux états initiaux.
Posons $w := abba$, déterminons si ce mot appartient au langage $L(\mathcal{A})$. Nous allons construire un seul arbre permettant d'avoir une condition suffisante de validation du mot.



L'arbre de lecture du mot *abba* contient un état acceptant comme feuille.

Autrement dit, il existe un chemin menant d'un état initial à un état acceptant dans la forêt de lecture de *abba*. Donc $abba \in L(\mathcal{A})$.

3.2.2 Juxtaposition et construction d'un AFD

Juxtaposition

Rappelons la problématique principale du chapitre. On cherche un modèle dérivant des AFD nous permettant de définir des opérations dessus et qui puisse être convertit algorithmiquement vers un AFD pour construire des automates d'un langage complexe à partir de langages plus simple.
Autrement dit, on veut pouvoir appliquer l'opération de somme de langages sur les automates fini déterministes.

Définition (**Juxtaposition d'AFN**) . Soient L_1 et L_2 deux langages reconnus par deux automates $\mathcal{A}_1$ et $\mathcal{A}_2$. Le langage $L_1 + L_2$ est reconnu par la **juxtaposition disjointe** de $\mathcal{A}_1$ et $\mathcal{A}_2$.

Théorème (**Langages automatiques et stabilité**) . L'ensemble des langages automatiques est stable par somme.

On peut maintenant, à partir de deux langages automatiques, définir le nouveau langage résultant de la somme des deux qui sera lui aussi automatique. Il suffit de faire la juxtaposition disjointe des deux automates de départ.

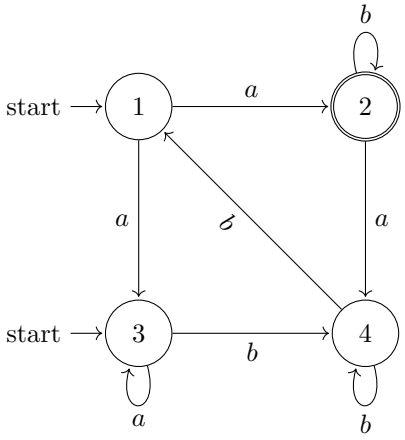
Déterminisation

Théorème (**Existence et équivalence**) . Pour tout automate fini non déterministe, il existe un automate déterministe équivalent.

Ce théorème est peut être un peu obscur mais permet de dire qu'il est toujours possible de passer d'un automate fini non déterministe (obtenu par exemple par juxtaposition) à un automate fini déterministe qui reconnaisse le même langage. En tout cas, il nous dit qu'il en existe un...
L'intérêt de vouloir repasser chez les automates fini déterministes vient du fait que la lecture d'un mot par un AFN est de complexité exponentielle alors que la lecture d'un mot par un AFD est polynômiale... Lors de la vérification syntaxique de très long mots pour des langages très complexes, cela fait une différence cruciale pour la compilation.
Ce processus est appelé **déterminisation** d'un AFN.

Proposition Soit $\mathcal{A} = (Q, \Sigma, T, I, A)$ un AFN. On cherche à construire un AFD $\mathcal{A}'$ équivalent à $\mathcal{A}$. L'idée est de raisonner sur l'application $T : Q \times \Sigma \longrightarrow \mathbb{Q}$. Dans un AFN, cette application n'est pas injective, on va donc poser une nouvelle application dans l'espace quotient de $Q \times \Sigma$ par le noyau de T. Nous obtiendrons donc une application injective et donc un AFD.

Exemple (Déterminisation) Soit $\mathcal{A}$ l'automate défini sur l'alphabet $\Sigma = \{a, b\}$ non déterministe et sa table de transition suivants :



T	a	b
1	{2, 3}	X
②	4	2
3	3	4
4	X	{1, 4}

Déterminisons cet automate. Pour cela, nous allons renommer tous les états de l'automate en prenant en compte les ensembles. L'algorithme consiste donc à construire la table de transition du nouvel automate. Pour chaque itération (i.e ajout d'une ligne dans la table), on effectue un parcours en largeur du nouvel automate pour "découvrir" de nouveau état. On créé ainsi un "automate des parties".

3.3 Automates fini à ε-transitions

Définissons un nouveau type d'automates non déterministes. Les automates non déterministes à ε-transition. Il diffèrent des premiers puisque l'on va permettre le changement d'état sans lecture

	a	b
$I = \{1,3\}$	II	III
$II = \{2,3\}$	IV	V
$III = \{4\}$	-	VI
$IV = \{3,4\}$	VII	VI
$V = \{2,4\}$	III	$VIII$
$VI = \{1,4\}$	II	VI
$VII = \{3\}$	VII	III
$VIII = \{1,2,4\}$	IX	$VIII$
$IX = \{2,3,4\}$	IV	$VIII$

FIGURE 3.1 – Table de l'AFD

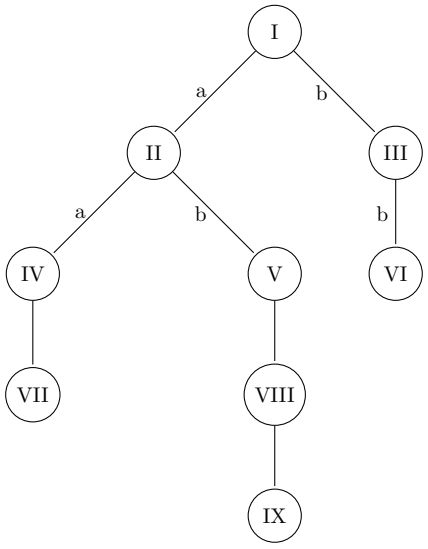


FIGURE 3.2 – Automate des parties

de lettres lors de la lecture d'un mot. Pour cela, nous allons définir une transition ε. Cela peut se voir comme une transition via le mot vide.

3.3.1 Généralités

Définition (**Automate Fini à ε-transitions (AFNε)**) . Un automate fini à ε-transitions est un quintuplet :

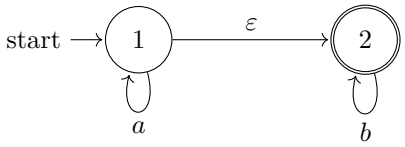
$$\mathcal{A} = (Q, \Sigma, T, I, A)$$

définit de la même façon que les automates précédents mais où :

$$T : Q \times \Sigma \cup \{\varepsilon\} \longrightarrow \mathcal{P}(Q)$$

Ici, le changement spontané d'état sans lecture de lettre sera donc caractérisé par une nouvelle entrée dans la table de transition ε.

Exemple Soit le langage $L := a^*b^*$. Un automate reconnaissant ce langage peut être écrit avec une ε-transition. Ecrivons aussi sa table de transition :



T	a	b	ε
1	1	-	2
②	-	2	-

Lors de la lecture d'un mot, les transitions peuvent être très aléatoires en fonction du nombre d'ε-transitions possibles de l'état courant. Un tel automate est donc hautement non déterministe.

Définition (**Lecture d'un mot**) . Soit $w = l_1 l_2 \dots l_n$ un mot sur Σ. Soit $w' = a_1 a_2 \dots a_p$ le mot w ε-complété (rembourré par des ε) tel que :

- $p \geqslant n$
- $\forall i \in [\![1, n]\!],\ a_i \in \{l_1, \dots, l_n\} \cup \{\varepsilon\}$

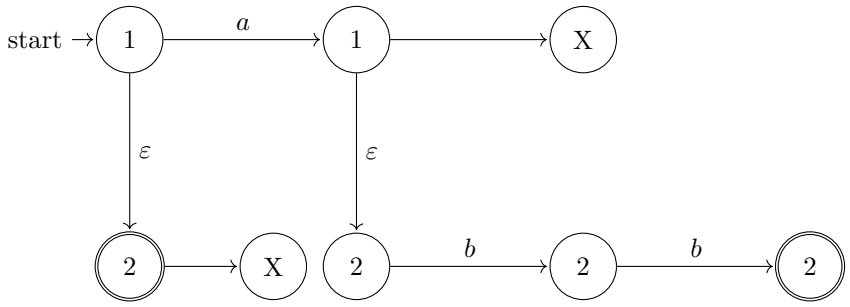
- $l_1 l_2 \dots l_n = a_1 a_2 \dots a_p$ (du point de vue du produit de concaténation)

Une lecture du mot w par $\mathcal{A}$ est une lecture par $\mathcal{A}$ de n'importe quel w', un ε-complété de w.

Remarque Tout comme pour les automates précédants, un mot appartient au langage d'un automate ssi la lecture de ce mot par celui-ci se finit sur au moins un état acceptant de l'automate.

La lecture d'un mot par un AFNε conduira donc à la construction d'une forêt de lecture de ce mot par l'automate.

Exemple (Lecture d'un mot) Soit l'automate fini non déterministe à ε-transitions précédent. Soit le mot *abb*. Construisons la forêt de lecture de *abb* par $\mathcal{A}$.



Il existe un chemin d'un racine vers une feuille acceptante dans la forêt de lecture :

$$1 \xrightarrow{a} 1 \xrightarrow{\varepsilon} 2 \xrightarrow{b} 2 \xrightarrow{b} 2$$

Donc par définition, $abb \in L(\mathcal{A})$.

3.3.2 Déterministation

Dans le chapitre précédent, un théorème nous permet de dire que pour tout automate fini non déterministe, il existe un automate fini déterministe équivalent. Ainsi, à chaque fois que l'on considère un AFNε, on est sûr qu'il existe un AFD équivalent.
Pour la déterminisation d'un AFNε, nous allons avoir besoin d'une définition supplémentaire...

Définition (Clôture) . Soit $\mathcal{A} = (Q, \Sigma, T, I, A)$ un AFNε. Pour tout $q \in Q$, on appelle clôture de q l'ensemble des états accessibles à partir de q sans lecture de lettre lors de la lecture d'un mot dans l'automate.
Autrement dit, la clôture de q est l'ensemble des états accessibles depuis q dans le sous-graphe de $\mathcal{A}$ restreint aux ε-transitions.
On note $cl(q)$ la clôture de q.

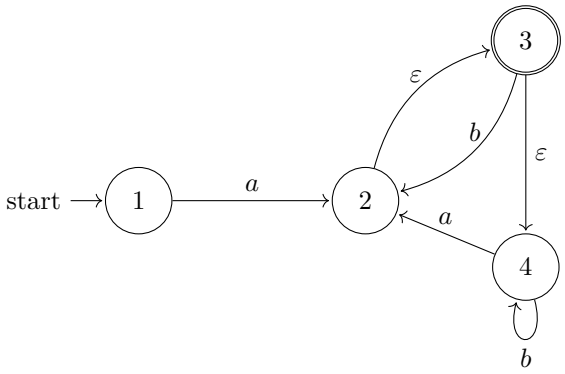
L'idée de la déterministation d'un AFNε est, non plus de regrouper des états, mais d'étendre les transitions de l'automate à tous les états accessibles par ε-transitions.
De plus, si un état acceptant est accessible uniquement par ε-transtion depuis un état, alors celui-ci hérite du caractère acceptant de l'état atteint.

Proposition (Algorithme de Déterministation) Soit $\mathcal{A} = (Q, \Sigma, T, I, A)$ un AFNε. On construit l'automate fini déterministe $\mathcal{A}'$ équivalent à $\mathcal{A}$ en :

1. Calculer les clôtures de $\mathcal{A}$
2. Héritage : Tous les états dont la clôture contient un état acceptant sont acceptants.

3. Calculer les transitions étendues
4. Déterminisation de $\mathcal{A}'$ par l'automate des parties (voir chap précédent)

Exemple (Déterminisation) Soit $\mathcal{A}$ l'AFNε suivant :



T	a	b	ε
1	2	-	-
2	-	-	3
③	-	2	4
4	2	4	-

1. **Calcul des clôtures et héritage**

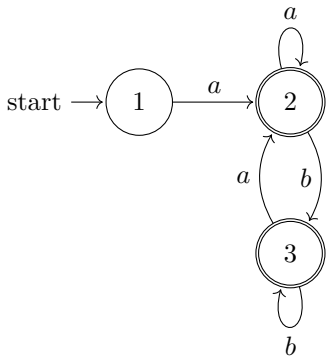
T	a	b	ε	$cl(q)$
1	2	-	-	$\{1\}$
②	-	-	3	$\{2,3,4\}$
③	-	2	4	$\{3,4\}$
4	2	4	-	$\{4\}$

2. **Calcul des transitions étendues**

$\widetilde{T}$	a	b	$cl(q)$
1	2	-	
②	2	$\{2,4\}$	$\{2,3,4\}$
③	2	$\{2,4\}$	$\{3,4\}$
4	2	4	

3. **Déterminisation par l'automate des parties**

	a	b
I = $\{1\}$	II	X
II = $\{2\}$	II	III
III = $\{2,4\}$	II	III



La déterminisation d'un AFNε nous permet donc de passer de la lecture d'un mot de complexité exponentielle (voire infinie) à un automate permettant de lire tous les mots avec une complexité linéaire.

==========

3.4 Opérations entre automates

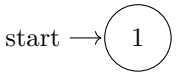
L'objectif de cette section est de nous permettre de construire un automate reconnaissant le langage $L_1 \ T \ L_2$ où T est une opération entre langages à partir des automates $\mathcal{A}_1$ et $\mathcal{A}_1$ reconnaissant respectivement les langages L_1 et L_2. Pour cela, nous aurons besoin de définir formellement les opérations entre langages et les processus nous permettant de "passer aux automates".

3.4.1 Langages Elémentaires

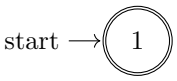
Toutes les opérations entre langages (par extension entre automates) se feront à partir de langages de élémentaires. Ces langages seront reconnus par des automates fixés, que nous connaissons d'avance. Nous en choisissons un nombre fini pour pouvoir les stocker en mémoire. Ils vont représenter les briques de base nous permettant de construire des automates plus complexes par la suite.

Proposition (Langages Elémentaires) Soit $\Sigma = \{a, b\}$ un alphabet. On définit les langages élémentaires de Σ comme :

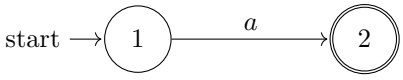
- $L = \emptyset$ reconnu par l'automate :



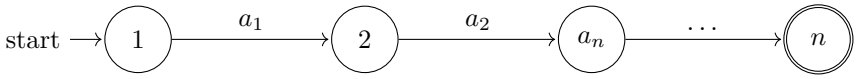
- $L = \{\varepsilon\}$ reconnu par l'automate :



- $L = \{a\}$ reconnu par le langage :



- $L = \{a_1 \dots a_n\}$ reconnu par le langage :



3.4.2 Automate Complémentaire

A partir d'un langage, on peut définir son langage complémentaire. De même, on peut définir l'automate complémantaire reconnaissant ce langage.

Définition (**Automate Complémentaire**) . Soit L un langage reconnu un automate $\mathcal{A} = (Q, \Sigma, T, q_0, A)$ **complet**. L'automate reconnaissant le langage complémentaire de L est $\overline{\mathcal{A}} = (Q, \Sigma, T, q_0, Q \backslash A)$.

Remarque Attention, pouvoir "passer au complémentaire" pour un automate, il faut que celuici soit complet.

3.4.3 Somme d'automates

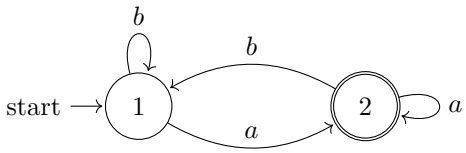
Définition (**Somme d'automates**) . Soient L_1 et L_2 deux langages respectivement reconnus par $\mathcal{A}_1$ et $\mathcal{A}_2$. L'automate reconnaissant L_1+L_2 est l'automate fini non déterministe construit par **l'union disjointe** de $\mathcal{A}_1$ et $\mathcal{A}_2$. On l'appelle **automate somme** des automates $\mathcal{A}_1$ et $\mathcal{A}_2$.

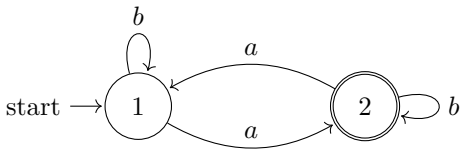
Cet automate n'est pas déterministe puisqu'il contient deux états initiaux mais que l'on peut déterminiser.

Exemple Soient les langages suivants sur $\Sigma = \{a, b\}$:

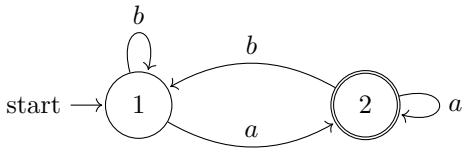
$$L_1 = \{\text{mots de } \Sigma \text{ terminant par } a\}$$
$$L_2 = \{\text{mots de } \Sigma\text{contenant un nombre pair de } a\}$$

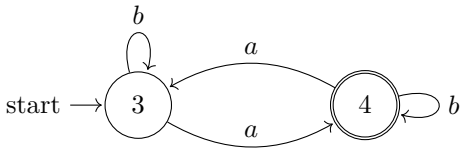
Ces langages sont reconnus par les automates suivants :





On construit l'automate $\mathcal{A}$ comme la juxtaposition disjointe des deux automates :





Que l'on doit ensuite déterminiser...

Ces définitions de langages élémentaires nous permettent de savoir que tout langage réduit à un mot est automatique. On en déduit le théorème suivant :

Théorème (**Langage Fini**) . Tout langage fini est automatique.

3.4.4 Intersection d'Automates

Proposition Soient L_1 et L_2 reconnus par les automates fini déterministes suivants $\mathcal{A}_1$ et $\mathcal{A}_2$. On a alors :

$$\overline{L_1 \cap L_2} = \overline{L_1} + \overline{L_2}$$

On peut donc en conclure que :

$$\boxed{L_1 \cap L_2 = \overline{\overline{L_1} + \overline{L_2}}}$$

Ainsi, l'intersection de deux automates peut être construite par somme et complémentarité.

En pratique nous utuliseront plutôt l'automate des couples :

Définition (**Automate des Couples**) . Soient $\mathcal{A}_1 = \{Q_1, \Sigma, T_1, Q_0, A_1\}$ reconnaissant le langage L_1 et $\mathcal{A}_2 = \{Q_2, \Sigma, T_2, q'_0, A_2\}$ reconnaissant le langage L_2. On définit l'automate des couples :

$$A = \{Q_1 \times Q_2, \Sigma, T, (q_0, q'_0), A_1 \times A_2\}$$

reconnaissant le langage $L_1 \cap L_2$ où

$$\forall i \in \Sigma, \forall q_1, q_2 \in Q_1, \forall q'_1, q'_2 \in Q_2 \text{ tels que } q_2 = T_1(q_1, j) \text{ et } q'_2 = T(q'_1, i)$$

$$\text{alors } T'((q_1, q_1'), i) = (q_2, q_2')$$

Proposition Si $\mathcal{A}_1$ possède $n \in \mathbb{N}$ états et que $\mathcal{A}_2$ possède $p \in \mathbb{N}$ états, alors $\mathcal{A}_1 \cap \mathcal{A}_2$ possède $n \times p$ états. Pour de gros automates, cette méthode peut donc engendrer des très gros. Même si la façon de les construire est assez simple et ressemble beaucoup à la déterminisation.
L'automate obtenu est, de plus, déterministe.

3.4.5 Différence d'Automates

Proposition Soient $\mathcal{A}_1$ et $\mathcal{A}_2$ deux automates reconnaissant respectivement les langages L_1 et L_2. Pour reconnaître le langages $L_1 \backslash L_2$, on peut simplement construire l'automate reconnaissant :

$$L_1 \backslash L_2 = L_1 \cap \overline{L_2}$$

3.4.6 Langages Automatiques

Essayons maintenant de déduire des conditions nécessaire pour qu'un langage soit automatique. D'après ce que l'on a vu grâce aux opérations, on peut déjà énoncer la proposition suivante :

Proposition Les langages réguliers sont tous automatiques.

Proposition que nous élargirons plus tard grâce au *théorème de Kleene*.

Théorème (**Pompage (faible)**) . Soit L un langage. Supposons L automatique. Soit $N \in \mathbb{N}$, alors pour tout $w \in L$ tel que $|w| \leqslant N$, w admet une décomposition de la forme :

$$\exists w_1, w_2, w_3 \in L, w = w_1 w_2 w_3 \quad \text{avec } w_2 \neq \varepsilon$$

telle que cette décomposition soit gonflable :

$$\text{i.e } \forall k \in \mathbb{N}, w_1 w_2^k w_3 \in L$$

Ce théorème n'est pas idéal pour montrer qu'un langage est automatique. En revanche, sa contraposé de la forme :

$$\boxed{\text{Non Pompable} \implies \text{Non Automatique}}$$

Est en pratique très utilisée pour montrer qu'un langage n'est pas automatique.

Proposition Ainsi, soit L un langage. Pour montrer que L n'est pas pompable, on montrera que pour tout $N \in \mathbb{N}$, il existe $w \in L$ tel que $|w| \geqslant N$ qui admette une décomposition de la forme :

$$w = w_1 w_2 w_3 \qquad \text{avec } w_2 \neq \varepsilon$$

telle que :

$$\exists k \in \mathbb{N}, w_1 w_2^k w_3 \notin L$$

Chapitre 4

Lemme d'Arden et Systèmes d'équations aux langages

Contents

Ce chapitre est dédié à l'étude du *Lemme d'Arden*. Ce lemme nous permettra de construire un automate déterministe à partir d'une expression régulière mais aussi à déterminer le langage d'un automate déterministe. Nous finirons par énoncer le théorème de Kleene caractérisant les langages automatiques. Il nous permettra de faire le lien entre langages automatiques et langages réguliers.

4.1 Lemme d'Arden

Le lemme d'Arden est présenté en 1961 par Dean N. Arden. En voici une de ces formes :

Lemme (Arden) Soient A et B deux langages. Le langage $L = A^*B$ est le plus petit langage (pour l'inclusion ensembliste) solution de l'équation :

$$(E) : X = (AX) \cup B$$

De plus, si A ne contient pas le mot vide ε, A^*B est l'unique solution de cette équation. Que l'on peut dériver en deux formes différentes : pour tout langage $L \in \Sigma^*$ et $a, b \in \Sigma$ tel que $\varepsilon \notin a$ on a :

Version Gauche :

$$\boxed{L = aL + b \iff L = a^*b}$$

Version Droite :

$$\boxed{L = La + b \iff L = ba^*}$$

4.2 Applications

4.2.1 Langage d'un automate fini

Maintenant que nous savons bien manipuler les automates fini déterministes et non déterministes, il serait utile, à partir d'un automate de pouvoir déterminer le langage qu'il reconnaît.

Pour cela nous avons besoin de définir les systèmes d'équations au langages.

Définition (**Langage d'arrivée**) . Soit $\mathcal{A} = \{Q, \Sigma, T, q_0, A\}$ un automate fini. On définit le langage d'arrivée à l'état $q \in Q$, noté L_q l'ensemble des mots de Σ^* dont la lecture par $\mathcal{A}$ débute par q_0 et finit en q.

Proposition Soit $\mathcal{A} = \{Q, \Sigma, T, q_0, A\}$ un automate fini. Soit $\{L_q \mid i \in A\}$ l'ensemble des langages d'arrivés aux états acceptants de l'automate $\mathcal{A}$. On a alors l'égalité suivante :

$$L(\mathcal{A}) = \sum_{q \in A} L_q$$

Autrement dit, le langage de $\mathcal{A}$ est la somme de tous ses langages d'arrivé aux états acceptants.

Définition (**Système d'équations aux langages**) . Soient un ensemble $X_1, \ldots, X_n$ de langages sur un même alphabet Σ. Un système d'équations aux langages est un ensemble d'équations de la forme :

$$X_i = \sum_{j=1}^{n} a_{ij} X_j + b_j \quad \forall i \in [\![1, n]\!]$$

où $\forall i, j \in [\![1, n]\!], a_{ij} \in \Sigma, b_i \in \Sigma^*$

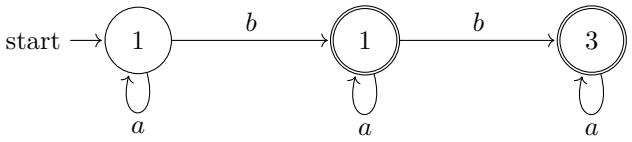
Proposition Soit $\mathcal{A} = \{Q, \Sigma, T, q_0, A\}$ un automate fini. À partir des définitions, on peut donc représenter le langage d'un automate par un système d'équations aux langages. Elles associent à chaque L_q une équation de langages dérivant de l'automate.

Ainsi si un état q a des transitions vers d'autres états selon les lettres $a \in \Sigma$ et si q est un état acceptant alors l'équation associée à L_q est de la forme

$$\begin{cases} L_q = \sum_{(q,a,q') \in T} a L_{q'} + \{\varepsilon\} & \text{si q est un état initial} \\ L_q = \sum_{(q,a,q') \in T} a L_{q'} & \text{sinon} \end{cases}$$

Il nous faut maintenant être capable de résoudre ces équations pour déterminer les L_q et ainsi le langage de l'automate. Équations que l'on peut maintenant résoudre grâce au lemme d'Arden.

Exemple Déterminons le langage de l'automate suivant :



Les langages d'arrivés vérifient donc le système d'équations aux langages suivant :

$$\begin{cases} L_1 = L_1 a + \varepsilon \\ L_2 = L_1 b + L_2 a \\ L_3 = L_2 b + L_3 b \end{cases}$$

Que l'on résout progressivement grâce au Lemme d'Arden :

$$\begin{cases} L_1 = \varepsilon a^* = a^* \\ L_2 = L_2 a + a^* b \\ L_3 = L_3 a + L_2 b \end{cases} \iff \begin{cases} L_1 = a^* \\ L_2 = a^* b a^* \\ L_3 = L_3 a + a^* b a^* \end{cases} \iff \begin{cases} L_1 = a^* \\ L_2 = a^* b a^* \\ L_3 = a^* b a^* b a^* \end{cases}$$

D'où :

$$\begin{aligned} L(\mathcal{A}) &= a^*ba^* + a^*ba^*ba^* \\ &= a^*ba^*(\varepsilon + ba^*) \end{aligned}$$

Donc $\mathcal{A}$ reconnaît les mots qui contiennent 1 ou 2 b.

Cet algorithme de résolution est très utile en terme pratique mais il apporte aussi une précisions supplémentaire sur les langages automatiques. En effet, à partir d'un automate (i.e langage automatique), on peut écrire le langage reconnu comme une expression régulière. D'où le résultat suivant :

Propriété (**Langages Automatiques**) . Tout langage automatique peut être décrit par une expression régulière.

On en déduit donc le théorème de Kleene établissant définitivement le lien entre les langages réguliers et automatique :

Théorème (**Kleene**) . Les langages réguliers sont les mêmes que les langages automatiques.

4.2.2 Construction d'un automate déterministe

Le lemme d'Arden et le théorème de Kleene nous permettent maintetant d'affirmer que toute expression régulière peut être exprimée comme un automate déterministe équivalent. Nous allons ici voir, autour de plusieurs exemples, comment construire un tel automate. Pour cela, nous aurons besoin de la version gauche du lemme d'Arden.

Remarque (Principe) L'idée de l'algorithme est de partir d'une expression régulière E et de la développer sous la forme :

$$E = x_1L_1 + x_2L_2 + \cdots + x_nL_n$$

où $x_1, \ldots, x_n \in \Sigma$ et $L_1, \ldots, L_n \in \Sigma^*$. Ensuite, on appelique récursivement cette opération sur tout les L_i pour obtenir une forme propice à appliquer Arden :

$$L_i = A^*B \iff L_i = AL_i + B$$

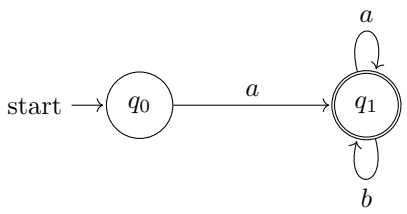
Enfin, chaque L_i représente un état de l'automate et tous les L_i composés d'un ε correspondent à un état acceptant.

Exemple Soit $\Sigma = \{a, b\}$ et le langage L décrit par l'expression régulière $a.(a+b)^*$. Déterminons l'automate de $\mathcal{A}$ reconnaîssant L :

$$\begin{aligned} L_0 &= a.(a+b)^* \\ &= a.L_1 \end{aligned}$$

$$\begin{aligned} L_1 &= (a+b)^*.\lambda \\ &= (a+b)L_1 + \lambda \\ &= a.L_1 + bL_1 + \lambda \end{aligned}$$

On a donc l'automate suivant :



Exemple Soit $\Sigma = \{a, b\}$ et le langage L décrit par l'expression régulière $ab^* + (a + b)c^*$. Déterminons l'automate de $\mathcal{A}$ reconnaîssant L :

$$\begin{aligned} L_0 &= ab^* + (a+b)c^* \\ &= ab^* + ac^* + bc^* \\ &= a(b^* + c^*) + bc^* \\ &= aL_1 + bL_2 \end{aligned}$$

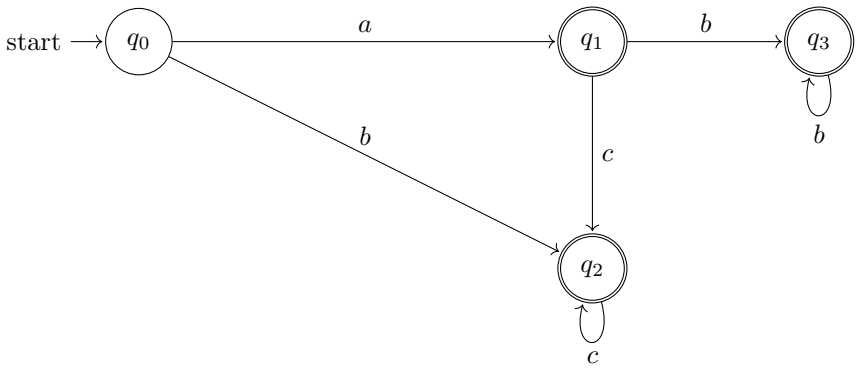
$$\begin{aligned} L_1 &= bb^* + \lambda + cc^* + \lambda \\ &= bL_3 + c_L 2 \end{aligned}$$

$$L_2 = c^* \iff L_2 = cL_2 + \lambda$$

$$L_3 = b^* if L_3 = bL_3 + \lambda$$

$$\begin{cases} L_0 = aL_1 + bL_2 \\ L_1 = bL_3 + cL_2 + \lambda \\ L_2 = cL_2 + \lambda \\ L_3 = bL_3 + \lambda \end{cases}$$

On a donc l'automate suivant :



Chapitre 5

Langages Algébriques

Contents

Précédement, nous avons vu que les langages automatiques sont de très bonnes propriétés. Ils sont stables pour la plupart des opérations définies sur les langages. De plus, le théorème de Kleene nous a permis d'établir le lien direct entre langages automatiques et langages réguliers. La simplicité de la représentation sagitale des automates permet de les implémenter facilement algorithmiquement. Il sont, de plus, facile à manipuler à la main et permettent de rapidement "voir" les langages reconnus.
Cependant, cette simplicité a un certain coût, celui de ne pas pouvoir reconnaître des langages "compliqués", notamment ceux où il faut "compter" les lettres. Un automate fini déterministe ne peut donc pas reconnaître le langage composé d'autant de a que de b. On va donc chercher à indroduire une nouvelle théorie, celle des **grammaires formelles** qui nous permettra de reconnaître de tels langages.

5.1 Grammaires Algébriques

5.1.1 Contexte et définition

Les grammaires formelles ont initialement été développés par des linguistes, notamment Noam Chomsky en 1955. L'objectif était de développer un méthode systématique de traduction entre différentes langues. Ils se sont alors heurtés au problème des mêmes mots qui admettent plusieurs traductions en fonction du contexte de la phrase et n'ont pas pu aboutir leur oeuvre.

Or en informatique, pour l'étude de la syntaxe de langages de programmation, le problème du contexte ne se pose pas. Leur théorie a donc été récupérée pour la vérification syntaxique.

L'idée est donc de représenter un langage *récursivement* par un ensemble de règles de production composées d'un axiome de départ et de différentes règles de productions ou de réécriture.
Nous utilisons souvent cette approche pour la gestion de types en Caml en définissant tes types récurisivement.

Définition (**Grammaire Algébrique**) . Une grammaire formelle est un quadruplet

$$G = (\Sigma, V, S, P)$$

où

- Σ est un *alphabet terminal* dont chaque élément ne peut se réécrire plus simplement.
- V est *l'alphabet auxiliaire* (disjoint de Σ) composé de variables, qui ne peuvent pas non plus se réécrire.
- S est la variable de départ, appelé axiome.
- P est un ensemble de règles dites *de production* ou de réécriture du type

$$X \longrightarrow w \quad X \in V \text{ et } w \in (V \cup \Sigma)^*$$

Par convention, on notera toujours les variables en majuscule et les éléments terminaux en minuscules. En pratique, on regroupera plusieurs réécritures d'une même variable sur la même ligne en les séparant par des barres verticales de la forme :

$$X \longrightarrow w_1|w_2|\ldots|w_p \iff \begin{cases} X \longrightarrow w_1 \\ X \longrightarrow w_2 \\ \vdots \\ W \longrightarrow w_p \end{cases}$$

5.1.2 Réécriture d'un mot et langages algébriques

L'objectif d'une grammaire algébrique, vous l'aurez compris, est de réécrire un mot récursivement jusqu'à arriver à des éléments terminaux.

Définition (**Réécriture d'un mot**) . Soit $G = (\Sigma, V, S, P)$ une grammaire formelle. Soient $u, v \in (V \cup \Sigma)^*$ deux mots. On dit que u *peut se réécrire en* v *en une étape* et on note :

$$u \vdash v$$

si il existe des décompositions de u et v en

$$u = u_1 X u_2 \text{ et } v = u_1 w u_2$$

et que G contient la règle de production :

$$X \longrightarrow w$$

Plus généralement, on peut définir la réécriture en plusieurs étapes de la forme :

Définition (**Réécriture (2)**) . Soit $G = (\Sigma, V, S, P)$ une grammaire algébrique. Soient $u, v \in (V \cup \Sigma)^*$ deux mots. On dit que u *peut se réécrire en* v ou que v *dérive en* u en un nombre quelconque de fois si il existe $u_1, \ldots, u_p \in (V \cup \Sigma)^*$ tels que

$$u \vdash u_1 \vdash u_2 \vdash \cdots \vdash u_p \vdash v$$

On note alors

$$u \overset{*}{\vdash} v$$

On peut maintenant définir les langages engendrés par des grammaires et les langages algébriques, le coeur de ce chapitre.

Définition (**Langage Engendré**) . La *langage engendré* par une grammaire algébrique $G = (\Sigma, V, S, P)$ est l'ensemble des mots de Σ^* qui dérivent de l'axiome S en un nombre quelconque d'étapes. On le note, comme pour les automates, $L(G)$.

Définition (**Langage Algébrique**) . Un langage engendré par une grammaire est appelé *langage algébrique* (d'où l'appelation de *grammaire algébrique*).

5.1.3 Arbre de dérivation d'un mot

Définition (**Arbre de dérivation d'un mot**) . Soit $G = (\Sigma, V, S, P)$ une grammaire algébrique. On appelle l'arbre de dérivation de $w \in Sigma^*$ l'arbre dont :

- La racine est S
- Tous les sommets intérieurs appartiennent à V
- Toutes les feuilles appartiennent à $\Sigma \cup \{\varepsilon\}$
- Si un sommet intérieur X a pour fils $X_1, \ldots, X_p$ alors la règle

$$X \longrightarrow X_1 | \ldots | X_p \in P$$

- Le mot obtenu en visitant les feuilles de l'arbre par un parcours profondeur préfixe de l'arbre est un mot de $L(G)$

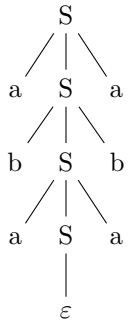
Définition (**Grammaire Ambiguë**) . Soit G une grammaire. On dit que G est ambiguë s'il existe un mot $w_{\in} L(G)$ possédant deux arbres de dérivation différents.

En pratique une telle grammaire est pas très utilisée. En effet, en informatique, il ne serait pas très pratique de pouvoir compiler un code en deux expressions différentes d'un autre langage. On ne saurait pas laquelle choisir. Il faut que la dérivation puisse se faire de façon unique.

Exemple (Grammaire et Arbre de dérivation) Soit $\Sigma = \{a, b\}$. Soit la grammaire algébrique G définie par les règles suivantes telles que $V = \{S\}$:

$$P : \begin{cases} S \longrightarrow aSa \\ S \longrightarrow SbS \\ S \longrightarrow \varepsilon \end{cases} \iff \Big\{ S \longrightarrow aSa \mid bSb \mid \varepsilon$$

Soit $abaaba \in G$ on a alors l'arbre de dérivation suivant pour ce mot :



Cette grammaire reconnaît bien les palindrômes pairs.

Remarque Lors de la dérivations de mots par une grammaire, on remarque qu'il est plus facile que les règles de dérivation possèdent des traces initiales ou finales uniques telles que les a et les b. Elles permettent d'identifier plus facilement les règles à utiliser pour les dérivations.

5.1.4 Grammaires Régulières

Nous allons ici faire le lien entre les deux modèles présentés précedement, les automates fini et les grammaires. Nous allons ainsi définir les *grammaires régulières* qui permettent de représenter les automates fini déterministes sous la forme que nous venons d'introduire.

Propriété (**Représentation d'un langage automatique**) . Soit $\mathcal{A} = (Q, \Sigma, T, q_0, A)$ un automate fini déterministe. Le langage L reconnu par cet automate peur être engendré par la grammaire :

$$G = (Q, \Sigma, q_0, P)$$

dont les variables auxiliaires sont les états de l'automate et où P est l'ensemble des productions de la forme :

$$q \longrightarrow x.T(q,x) \quad \text{où } q \in Q \text{ et } x \in \Sigma$$

$$q \longrightarrow \varepsilon \quad \text{si } q \in A$$

On peut donc représenter facilement n'importe quel langage automatique par une grammaire. D'où le théorème suivant.

Théorème (**Langage Automatique et Grammaire**) . Tout langage automatique (i.e reconnaissable par un automate fini) est algébrique (i.e reconnaissable par une grammaire formelle).

L'ensemble des langages automatiques est même strictement inclus dans l'ensemble des langages algébriques. Autrement dit, certains langages sont reconnaissables par une grammaire formelle mais pas par un automate.

On définit ainsi les grammaires régulières.

Définition (**Grammaire Régulière**) . Une *grammaire régulière* est une grammaire formelle dont toutes les règles de production de P sont de la forme :

$$X \longrightarrow a.Y \quad \text{ou } X \longrightarrow \varepsilon$$

où $X, Y \in V$ et $a \in \Sigma$.

Une grammaire régulière est donc conçue de façon à "laisser des traces" explicites de la structure des mots pour faciliter les dérivations. De même que précédement, on peut passer d'une grammaire régulière à un automate fini déterministe.

Proposition (Représentation d'une grammaire régulière) Soit G une grammaire régulière. Soit L le langage reconnu par G. L'automate fini déterministe reconnaissante aussi L est :

$$\mathcal{A} = (V, \Sigma, T, q_0 = S, A)$$

dont les états sont les variables auxiliaires de G et dont les transitions sont définies par :

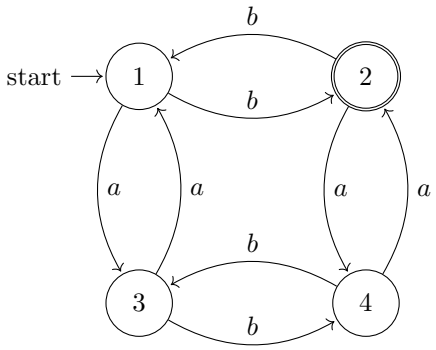
$$q' = T(q, x) \quad \text{si} \quad q \longrightarrow xq' \in P$$

et dont les états acceptants sont définis par :

$$q \in A \quad \text{si} \quad q \longrightarrow \varepsilon \in P$$

Remarque Grâce à cette propriété, les langages automatiques (réguliers) sont donc exactement les langages reconnus par des grammaires régulières. D'où le nom...

Exemple (Construction d'une grammaire régulière) Définissons le langage L reconnaissant les mots contenant un nombre pair de a et impair de b. Alors ce langage est reconnu par l'AFn suivant :



D'après la propriété précédente, L est reconnu par la grammaire régulière $G = (\Sigma, V, S, P)$ où $V = \{S, A, B, C\}$ et les états sont représentés par :

$$\begin{cases} 1 \longrightarrow S \\ 2 \longrightarrow A \\ 3 \longrightarrow B \\ 4 \longrightarrow C \end{cases}$$

On peut ensuite déterminer les règles de production à partir du voisinnage sortant de chaque état de l'automate. De plus, puisque A est un état acceptant, on y rajouter ε.

$$P : \begin{cases} S \longrightarrow bA \mid aB \\ A \longrightarrow bS \mid aC \mid \varepsilon \\ B \longrightarrow aS \mid cB \\ C \longrightarrow aA \mid bB \end{cases}$$

5.2 Forme Normale de Chomsky

Tout comme les automates, on va chercher à simplifier les grammaires. Or ici, pour un langage donné il n'existe pas de forme minimale de grammaire qui l'engendre.

On va donc chercher à simplifier les grammaires dans le but d'obtenir des formes dites normales pour réduire le nombre de dérivations à faire pour un mot donné. L'idée est de ramener les arbres de dérivation à des arbres binaires donc ekes dérivations sont seulement de deux formes. On pourra donc calculer directement la profondeur de l'arbre de dérivation de n'importe quel mot du langage engendré en fonction de son nombre de caractères.
Pour cela, il faut définir un certain nombre de règles qui serviront à cette simplification. Commençons par un règle très simple :

Définition (**Règle 0**) . On peut toujours éliminer une règle de la forme $X \longrightarrow X$.

5.2.1 Règle 1 : Suppression des epsilon-productions

On va chercher ici à supprimer toutes les ε productions qui ne produisent rien dans la dérivation d'un mot et prennent beaucoup de temps et de place à exécuter.

Définition (**Règle 1 : Suppression des espilon-productions**) . Soit G une grammaire formelle. On définit l'algorithme suivant pour supprimer toutes les ε-productions de G en une grammaire équivalente :

1. On cherche récursivement toutes les variables dont ε dérive (i.e toutes les variables qui peuvent nous donner ε à la fin).
2. On supprimer toutes les règles de la forme $X \longrightarrow \varepsilon$.
3. Pour toutes les variables X de la forme $X \longrightarrow w$ on rajoute toutes les productions $X \longrightarrow u$ avec $u \neq u$ et u est obtenu à partir de w en remplaçant une ou plusieurs variables indentifiées en 1.

Exemple Soit G une grammaire d'alphabet $\Sigma = \{a, b\}$ et $V = \{S, A, B\}$ tel que :

$$P : \begin{cases} S \longrightarrow AB|aS|A \\ A \longrightarrow Ab|\varepsilon \\ B \longrightarrow B|AS \end{cases}$$

On cherche G' telle sans ε-productions telle que :

$$L(G') \cup \{\varepsilon\} = L(G)$$

D'après la règle 1, toutes les variables produisent une ε-production. On obtient donc la grammaire équivalente à ε-production près :

$$P' : \begin{cases} S \longrightarrow AB|A|B|aS|a \\ A \longrightarrow Ab|b \\ B \longrightarrow AS|A|S \end{cases}$$

5.2.2 Règle 2 : Élimination des cycles

Dans les arbres de dérivation, les cycles peuveut conduire à des dérivations infinies. On va donc chercher à les supprimer.

Définition (**Règle 2 : Élimination des cycles**) . Soit G une grammaire formelle. On définit l'algorithme suivant pour supprimer tous les cycles de G en une grammaire équivalente. Soit un cycle de la forme :

$$X_1 \longrightarrow X_{n-1} \longrightarrow X_1$$

Alors on remplace dans P toutes les variables $X_i \forall i \in [\![1, n-1]\!]$ par X_1.

Exemple En reprennant l'exemple précédent :

$$P' : \begin{cases} S \longrightarrow AB|A|B|aS|a \\ A \longrightarrow Ab|b \\ B \longrightarrow AS|A|S \end{cases}$$

On détecte un seul cycle : $S \longrightarrow B \longrightarrow S$. On applique donc l'algorithme pour obtenir :

$$P'' : \begin{cases} S \longrightarrow AS|A|aS|a \\ A \longrightarrow Ab|b \end{cases}$$

5.2.3 Règle 3 : Suppression des changements de variable

Les changement de variable dans les arbres de dérivation font perdre du temps. En effet, ils augmentent la profondeur de l'arbre de dérivation sans produire de lettre. On va donc chercher à les supprimer avec la règle 3.

Définition (**Règle 3 : Suppression des changements de variable**) . Soit G une grammaire formelle. Soit une dérivation de la forme $A \longrightarrow B \longrightarrow C$. Alors on peut la remplacer en $A \longrightarrow C$.

Exemple En reprenant l'exemple précédent, on peut supprimer les changements de variable :

$$P'' : \begin{cases} S \longrightarrow AS|A|aS|a \\ A \longrightarrow Ab|b \end{cases}$$

On a donc les règles de productions suivantes en remplaçant :

$$P''' : \begin{cases} S \longrightarrow AS|Ab|b|aS|a \\ A \longrightarrow Ab|b \end{cases}$$

5.2.4 Forme Normale de Chomsky

Pour l'instant, on ne peut pas encore déterminer la profondeur de l'arbre de dérivation d'un mot. Il faut donc définir une forme, dite Normale de Chomsky, qui va permettre cette estimation.

Définition (**Forme Normale de Chomsky**) . Soit $G = (\Sigma, V, S, P)$ une grammaire formelle. Supposons que G ne contienne ni ε-production, si changements de variables, ni cycles. On définit alors la forme normale de Chomsky de ses règles de production P comme ses mêmes règles de production où seules deux formes sont autorisées :

- Les productions de lettres de la forme $X \longrightarrow a$
- Les dédoublements de variables de la forme $X \longrightarrow AB$

Exemple En reprenant l'exemplement précédent, on obtient la forme normale de Chomsky suivante :

$$\widetilde{P} : \begin{cases} S \longrightarrow AS|AY|b|XS|a \\ A \longrightarrow AY|b \\ X \longrightarrow a \\ Y \longrightarrow b \end{cases}$$

On en déduit donc le théorème suivant :

Théorème (**Majoration de la profondeur**) . Soit $w \in L(G)$ où G est une grammaire formelle sous forme normale de Chomsky. Notons p la profondeur de l'arbre de dérivation de w dans cette grammaire g. On a alors l'inégalité suivante :

$$\boxed{p \leqslant 2|w| - 1}$$

Chapitre 6

Automates à piles

Au début de ce cours, nous avons vu que l'ensemble des langages régulier est strictement inclus dans l'ensemble des langages algébriques. Il existe donc des langages reconnaissables par une grammaire mais pas par un automate... Or les automates sont des outils très pratiques pour l'analyse syntaxique d'un mot, en effet, ils permettent de vérifier rapidement et à moindre coût l'appartenance d'un mot à un alphabet. Nous allons ici définir un nouveau type d'automates réglant définitivement ce problème : les *automates à piles*.