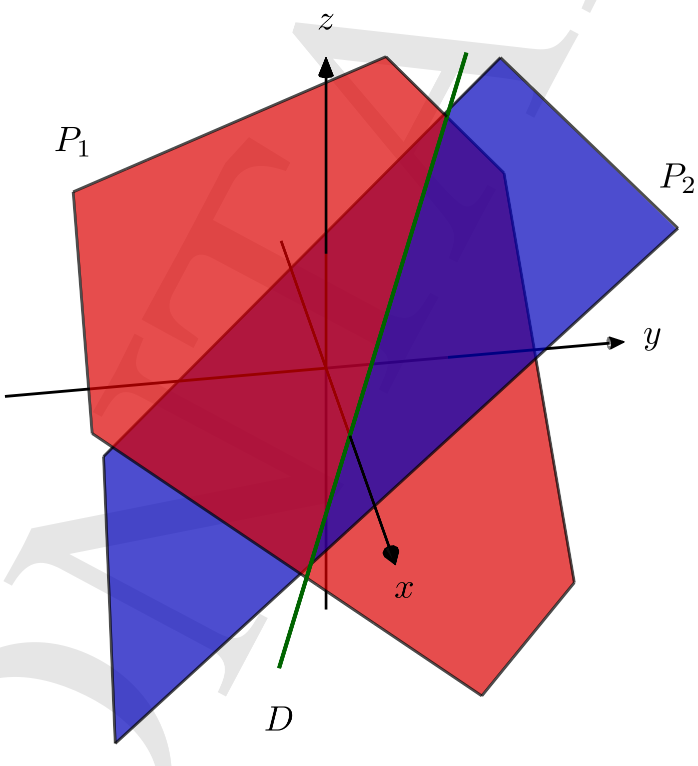
Cours de Mathématiques

MPSI

Lycée Montaigne 2021-22



Chapitre 1

Éléments de logique

Sommaire

I NOTIONS ENSEMBLISTES

1) Vocabulaire lié aux ensembles

Nous ne définirons pas rigoureusement la notion d'ensemble, celle-ci sera considérée comme intuitive. Nous nous contenterons de la « définition » suivante :

Définition 1.1

Un ensemble E *est une collection d'objets*[1]*, ceux-ci sont appelés éléments de* E*. Si* x *est un élément de* E *on écrira* $x \in \mathrm{E}$ *(se lit « x appartient à E »), dans le cas contraire on écrira* $x \notin \mathrm{E}$*. Si* E *n'a pas d'éléments on dira que c'est l'ensemble vide et on le notera* $\varnothing$*. Deux ensembles* E *et* F *sont dits égaux si et seulement si ils ont les mêmes éléments, on écrira alors* $\mathrm{E} = \mathrm{F}$*.*

☞**Exemples** :
- Les ensembles de nombres : $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{D}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$.
- L'ensemble des fonctions de $\mathbb{R}$ dans $\mathbb{R}$: $\mathcal{F}(\mathbb{R}, \mathbb{R})$.
- Ensembles définis en *extension*, comme : $\mathrm{E} = \{1; 8; 6; 2\}$ (éléments non ordonnés et devant apparaître une seule fois dans la liste).
- Ensembles définis en *compréhension*, comme : $\mathrm{E} = \{p \in \mathbb{N} \ / \ p \text{ est impair}\} = \{2k+1 \ / \ k \in \mathbb{N}\}$.

1. Cependant toute collection d'objets ne constitue pas forcément un ensemble. Par exemple, le paradoxe de Bertrand Russel a montré que l'ensemble des ensembles ne peut pas exister, sinon, la considération de l'ensemble $y = \{x \ / \ x \notin x\}$ conduit à une absurdité.

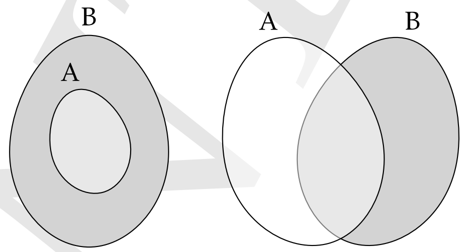
Attention !

L'écriture $E = \{p \in \mathbb{N} \,/\, p$ *est impair*$\}$ *ne signifie pas que* E *est un ensemble qui contient un seul élément qui s'appelle p, mais que* E *est l'ensemble de* **tous** *les entiers naturels p tels que p est impair, en langage courant on dit plutôt que* E *est l'ensemble de tous les entiers naturels impairs. Dans le langage mathématique, il faut donner un nom à ces entiers (variable muette, ici p) pour pouvoir ensuite les manipuler ou les décrire.*

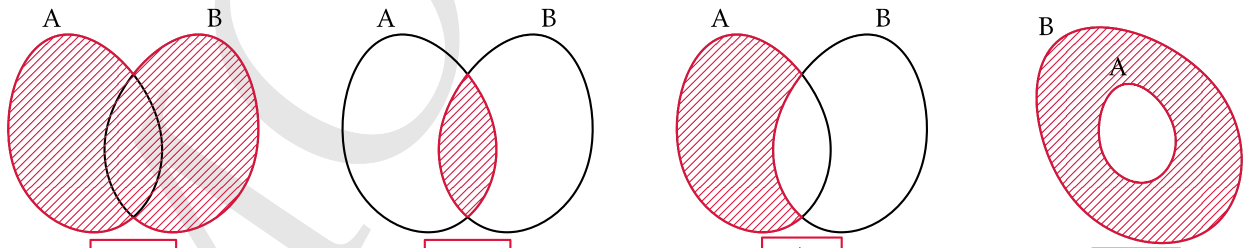
Définition 1.2

Soient A *et* B *deux ensembles :*

- ***L'inclusion*** *: on dit que* A *est inclus dans* B *tous les éléments de* A *sont également éléments de* B*, notation :* $A \subset B$.
- ***Ensemble des parties*** *: si* A *est inclus dans* B*, on dit que* A *est une* **partie** *de* B*. L'ensemble des parties de* B *est noté* $\mathcal{P}(B)$*, donc écrire « $A \subset E$ » revient à écrire « $A \in \mathcal{P}(B)$ ». Par exemple, l'ensemble vide et* B *sont des parties de* B*, donc* $\varnothing \in \mathcal{P}(B)$ *et* $B \in \mathcal{P}(B)$.
- ***La réunion*** *: on note* $A \cup B$ *(se lit « A union B »), l'ensemble que l'on obtient en regroupant les éléments de* A *avec ceux de* B*, par exemple* $\{2k+1 \,/\, k \in \mathbb{Z}\} \cup \{2n \,/\, n \in \mathbb{Z}\} = \mathbb{Z}$.
- ***L'intersection*** *: on note* $A \cap B$ *(se lit « A inter B »), l'ensemble des éléments* **communs** *à* A *et* B*. Par exemple* $\mathbb{N} \cap \mathbb{Z}^* = \mathbb{N}^*$*. On dit que deux ensembles sont* **disjoints** *lorsque leur intersection est l'ensemble vide.*
- ***La différence*** *: on note* $A \setminus B$ *(se lit « A moins B »), l'ensemble des éléments qui sont dans* A *mais pas dans* B*. Par exemple,* $\mathbb{R}^+ = \mathbb{R} \setminus]-\infty; 0[$.
- ***Le complémentaire d'une partie*** *: lorsque* A *est une* **partie** *de* B*, la différence* $B \setminus A$ *est appelé complémentaire de* A *dans* B*, noté* $\complement_B(A)$ *(ou bien* $\overline{A}$*). Par exemple* $\mathbb{R} \setminus \mathbb{Q}$ *est l'ensemble des irrationnels.*
- ***Le produit cartésien*** *: le produit cartésien de* A *par* B *est l'ensemble des couples* $(x; y)$ *avec* $x \in A$ *et* $y \in B$*, on le note* $A \times B$*, c'est à dire* $A \times B = \{(x; y) \,/\, x \in A,\ y \in B\}$*. On rappelle que* $(x; y) = (a; b)$ *si et seulement si* $x = a$ **et** $y = b$*. Attention à ne pas confondre un couple avec une paire (ensemble à deux éléments), par exemple :* $\{1; 2\} = \{2; 1\}$*, mais* $(1; 2) \neq (2; 1)$.



$A \subset B$ $\quad$ $A \not\subset B$



$A \cup B$ $\quad$ $A \cap B$ $\quad$ $A \setminus B$ $\quad$ $\complement_B(A)$

★**Exercice 1.1** *Décrire* $\mathcal{P}(E)$ *lorsque* $E = \{1; 2; 3\}$.

Remarque 1.1 :

- *Dire que deux ensembles* A *et* B *sont égaux, revient à dire que* A *est inclus dans* B*, et que* B *est inclus dans* A*. Donc* ***démontrer une égalité entre deux ensembles, peut se faire en démontrant une double inclusion****.*
- *Le produit cartésien se généralise à trois ensembles ou plus généralement à* n *ensembles :*

$$E_1 \times \cdots \times E_n = \{(x_1, \ldots, x_n) \,/\, x_1 \in E_1, \ldots, x_n \in E_n\}$$

Lorsque tous les ensembles sont égaux au même ensemble E*, on note* $E \times \cdots \times E = E^n$ *(ensemble des* n*-listes, ou* n*-uplets d'éléments de* E*). Par exemple,* $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$*, ensemble des couples de réels, peut servir à modéliser un plan muni d'un repère, de même,* $\mathbb{R}^3 = \mathbb{R} \times \mathbb{R} \times \mathbb{R}$*, ensemble des triplets de réels, peut servir à modéliser l'espace usuel muni d'un repère.*

2) Propriétés

Théorème 1.1 (Propriétés de la réunion et de l'intersection)

Soient A*,* B *et* C *trois ensembles, on a* $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$*. C'est la* ***distributivité*** *de la réunion sur l'intersection.*

De même, on a $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$*. C'est la* ***distributivité*** *de l'intersection sur la réunion.*

Preuve : Ceci sera démontré un peu plus loin. □

Théorème 1.2 (Propriétés du complémentaire)

Si A *et* B *sont deux parties d'un ensemble* E *:*

- $A \cup C_E(A) = E$.
- $C_E(E) = \emptyset$, $C_E(\emptyset) = E$.
- $C_E(C_E(A)) = A$.
- $C_E(A \cup B) = C_E(A) \cap C_E(B)$ *(loi de De Morgan*[2]*).*
- $C_E(A \cap B) = C_E(A) \cup C_E(B)$ *(2^e^ loi de De Morgan).*

Preuve : Ceci sera démontré un peu plus loin. □

II NOTIONS DE LOGIQUE

1) Propositions

Nous nous contenterons de la « définition » suivante :

Définition 1.3 (Proposition)

Une proposition est une phrase (ou assertion) qui a un sens mathématique et qui est soit vraie soit fausse[3]*. On dira qu'une proposition n'a que deux valeurs de vérité : vraie (notée V) et fausse (notée F). Si* P *désigne une assertion, on notera* $\neg P$ *sa négation (lire « non* P *»).*

☞**Exemples** :

- « 2 est un entier pair » est une proposition vraie.
- « 3 est un entier pair » est une proposition fausse.
- « n est un entier pair » n'est pas une proposition car sa valeur de vérité dépend de la valeur de n, un telle phrase est appelée **prédicat** portant sur la variable n à valeurs dans $\mathbb{Z}$, on pourrait noter ce prédicat $P(n)$ par exemple.
- Si A et B désignent deux ensembles, alors la phrase $A \subset B$ est une proposition (tous les éléments de A sont éléments de B). Sa négation s'écrit $A \not\subset B$ (un élément de A n'est pas forcément un élément de B).
- L'expression $\mathbb{N} \in \mathbb{R}$ » est une proposition, elle est fausse car $\mathbb{N}$ n'est pas un réel.

2. *DE MORGAN Augustus* (1806 – 1871) logicien anglais.
3. Il ne doit pas y avoir d'autre alternative selon le principe du tiers exclu.

– L'expression $\mathbb{N} \subset \mathbb{R}$ » est une proposition, elle est vraie car tout naturel est aussi un réel.

Si P est une proposition, la valeur de vérité de $\neg$P se déduit de celle de P conformément à la table de vérité suivante :

P	$\neg$P
V	F
F	V

Par convention :

Dans les raisonnements mathématiques on n'écrit que des propositions vraies. Si P est une proposition, au lieu d'écrire « P est vraie », on écrit plus simplement « P », et au lieu d'écrire « P est fausse », on écrit plus simplement « $\neg$P », c'est à dire la négation.

2) Connecteurs logiques

Ceux-ci permettent de relier deux propositions pour en donner une troisième.

Définition 1.4 (conjonction, disjonction inclusive)

Soient P *et* Q *deux propositions. On dit que :*
• *la proposition «* $P \wedge Q$ *» (lire «* P *et* Q *») est vraie lorsque les deux propositions le sont simultanément, sinon on dira qu'elle fausse.*
• *la proposition «* $P \vee Q$ *» (lire «* P *ou* Q *») est vraie lorsqu'au moins une des deux propositions est vraie (voire les deux), sinon on dira qu'elle fausse.*

P	Q	$P \wedge Q$	$P \vee Q$
V	V	V	V
V	F	F	V
F	V	F	V
F	F	F	F

Définition 1.5 (implication, équivalence)

Soient P *et* Q *deux propositions. On dit que :*
• *la proposition «* $P \implies Q$ *» (lire «* P *implication* Q *») est fausse lorsque* P ***est vraie mais pas*** Q.
• *la proposition «* $P \iff Q$ *» (lire «* P *équivalence* Q *») est vraie lorsque les deux propositions ont la* ***même valeur de vérité****.*

P	Q	$P \implies Q$	$P \iff Q$
V	V	V	V
V	F	F	F
F	V	V	F
F	F	V	V

☞**Exemples** :

– La proposition « (2 est pair) $\implies$ (3 est impair) » est vraie.
– La proposition « (2 est impair) $\implies$ (3 est pair) » est vraie [4].
– La proposition « (2 est pair) $\implies$ (3 est pair) » est fausse.
– La proposition « (2 est impair) $\iff$ (3 est pair) » est vraie.

Définition 1.6 (implique, équivaut)

Soient P *et* Q *deux propositions.*
• *Lorsque la proposition «* $P \implies Q$ *» est vraie on dit «* P *implique* Q *» (ou « si* P *alors* Q *»).*
• *Lorsque la proposition «* $P \iff Q$ *» est vraie on dit que «* P *équivaut à* Q *» (ou «* P *est équivalent à* Q *» ou encore «* P *si et seulement si* Q *»).*

4. Ceci peut paraître surprenant au premier abord, mais nous verrons qu'en écrivant la négation cela devient évident.

Principe de déduction

Soient P et Q deux propositions, si on sait que P implique Q, et que P est vraie, alors on a forcément Q vraie d'après la définition de l'implication. C'est le **principe de déduction** [5].

3) Propriétés

Maintenant que nous savons ce que sont des propositions équivalentes, nous allons pouvoir établir les propriétés suivantes :

Théorème 1.3

Soient P et Q deux propositions :
- *La proposition « $\neg\neg P$ » est équivalente à « P ».*
- *La proposition « $\neg(P \wedge Q)$ » est équivalente à « $(\neg P) \vee (\neg Q)$ » (1^re loi de De Morgan).*
- *La proposition « $\neg(P \vee Q)$ » est équivalente à « $(\neg P) \wedge (\neg Q)$ » (2^e loi de De Morgan).*
- *L'implication « $P \implies Q$ » est équivalente à « $(\neg P) \vee Q$ ».*
- *La proposition « $\neg(P \implies Q)$ » est équivalente à « $P \wedge (\neg Q)$ ».*
- *La proposition « $P \iff Q$ » est équivalente à « $(P \implies Q) \wedge (Q \implies P)$ ».*
- *La proposition « $P \iff Q$ » est équivalente à « $(\neg P) \iff (\neg Q)$ ».*

Preuve : La première propriété est évidente. Les autres se montrent avec une table de vérité (à compléter en exercice) :

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

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

□

☞**Exemple** : La négation de la proposition « (2 est impair) $\implies$ (3 est pair) » est équivalente à la proposition « (2 est impair) $\wedge$ (3 est impair) », or celle-ci est fausse, et donc la première est vraie.

★**Exercice 1.2** *Sans utiliser de table de vérité, redémontrer (en utilisant les autres propriétés) que : « $\neg(P \implies Q)$ » est équivalente à « $P \wedge (\neg Q)$ ».*

Définition 1.7 (réciproque, contraposée)

Soient P et Q deux propositions.
- *La réciproque de l'implication « $P \implies Q$ » est l'implication « $Q \implies P$ ».*
- *La contraposée de l'implication « $P \implies Q$ » est l'implication « $(\neg Q) \implies (\neg P)$ ».*

Il découle alors du théorème précédent :

Théorème 1.4

Une implication et sa contraposée sont équivalentes.
Deux propositions sont équivalentes si et seulement si les implications dans les deux sens sont vraies.

5. Par contre, si P implique Q, et que P est fausse, alors on ne peut rien dire de Q.

Remarque 1.2 – *Ce résultat est à connaître car très utilisé dans les raisonnements (raisonnements par contra-position, raisonnements par double implication).*

4) Quantificateurs

Les quantificateurs servent à construire des propositions à partir d'un prédicat $\mathrm{P}(x)$, dont la variable x prend ses valeurs dans un certain ensemble E. On rencontre :

- Le quantificateur **universel** : « $\forall x \in \mathrm{E}, \mathrm{P}(x)$ » (se lit « pour tout x dans E, $\mathrm{P}(x)$ [sous entendu est vraie] »). Par exemple, la proposition « $\forall x \in \mathbb{R}, x^2 \geqslant 0$ » se lit « pour tout réel x, le carré de x est positif ou nul », ou bien encore « le carré de tout réel est positif ».
- Le quantificateur **existentiel** : « $\exists x \in \mathrm{E}, \mathrm{P}(x)$ » (se lit « il existe au moins un x de E tel que $\mathrm{P}(x)$ [sous entendu est vraie] »). Par exemple, la proposition « $\exists x \in \mathbb{C}, x^2 = -1$ », se lit « il existe au moins un nombre complexe dont le carré vaut -1 ».

Remarque 1.3 :

- *On rencontre aussi parfois la proposition « $\exists! x \in \mathrm{E}, \mathrm{P}(x)$ » (se lit « il existe un unique x de E tel que $\mathrm{P}(x)$ [sous entendu est vraie] »). Par exemple, « $\exists! x \in \mathbb{R}^+, x^2 = 2$ ».*
- *On peut trouver plusieurs quantificateurs dans une même proposition. Par exemple, « $\forall y \in \mathbb{R}^+, \exists! x \in \mathbb{R}^+, x^2 = y$ » traduit que tout réel positif est le carré d'un unique réel positif.*

Attention !

*Les propositions « $\forall x \in \mathrm{A}, \exists y \in \mathrm{B}, \mathrm{P}(x, y)$ » et « $\exists y \in \mathrm{B}, \forall x \in \mathrm{A}, \mathrm{P}(x, y)$ », n'ont pas le même sens. En effet, dans la première le y **dépend** de x alors que dans la seconde il s'agit du **même** y pour tous les x.*

À retenir : utilisation des quantificateurs

Celle-ci est régie par les règles suivantes :

a) La négation de $\forall$ est $\exists$ (et vice - versa).

b) On peut intervertir deux quantificateurs de même nature.

c) On ne peut pas intervertir deux quantificateurs de nature différente.

☞**Exemples** :

- L'assertion « $\forall x \in \mathbb{R}^+, \exists y \in \mathbb{R}^+, y^2 = x$ » est vraie, elle traduit que tout réel positif (x) est le carré d'au moins un réel positif (y). Mais l'assertion « $\exists y \in \mathbb{R}^+, \forall x \in \mathbb{R}^+, y^2 = x$ » traduit que tout réel positif (x) est le carré d'un **même réel** (y), ce qui est évidemment faux. Sa négation est « $\forall y \in \mathbb{R}^+, \exists x \in \mathbb{R}^+, y^2 \neq x$ ».
- On a toujours l'implication suivante : $(\exists x \in \mathrm{A}, \forall y \in \mathrm{B}, \mathrm{P}(x, y)) \implies (\forall y \in \mathrm{B}, \exists x \in \mathrm{A}, \mathrm{P}(x, y))$.
- La négation de « $\forall x \in \mathrm{A}, \exists y \in \mathrm{B}, \mathrm{P}(x, y)$ » est « $\exists x \in \mathrm{A}, \forall y \in \mathrm{B}, \neg(\mathrm{P}(x, y))$ ».

★**Exercice 1.3**

1/ *Traduire dans le langage mathématique : la suite (u_n) est majorée. Écrire la négation. Qu'en est-il de la suite définie par $u_n = n^2$? Justifier.*

2/ *Traduire dans le langage courant les propositions suivantes :*

a) $\exists y \in \mathbb{R}, \forall x \in \mathbb{R}, x \leqslant y$.

b) $\exists y \in \mathbb{N}, \forall x \in \mathbb{N}, y \leqslant x$.

5) Retour sur les ensembles

Soient A et B deux parties d'un ensemble E.

Intersection d'ensembles

Si x désigne un élément de E, démontrer que $x \in \mathrm{A} \cap \mathrm{B}$, c'est démontrer la proposition « $(x \in \mathrm{A})$ et $(x \in \mathrm{B})$ ». On peut donc écrire : $\mathrm{A} \cap \mathrm{B} = \{x \in \mathrm{E} \,/\, x \in \mathrm{A} \text{ et } x \in \mathrm{B}\}$.

Réunion d'ensembles

Si x désigne un élément de E, démontrer que $x \in \mathrm{A} \cup \mathrm{B}$, c'est démontrer la proposition « $(x \in \mathrm{A})$ ou $(x \in \mathrm{B})$ ». On peut donc écrire : $\mathrm{A} \cup \mathrm{B} = \{x \in \mathrm{E} \,/\, x \in \mathrm{A} \text{ ou } x \in \mathrm{B}\}$.

Complémentaire

Si x désigne un élément de E, démontrer que $x \in C_E(A)$, c'est démontrer la proposition « ($x \notin A$). On peut donc écrire : $C_E(A) = \{x \in E \,/\, x \notin A\}$.

Différence d'ensembles

Si x désigne un élément de E, démontrer que $x \in A \setminus B$, c'est démontrer la proposition « $(x \in A)$ et $(x \notin B)$ ». On peut donc écrire : $A \setminus B = \{x \in E \,/\, x \in A \text{ et } x \notin B\}$, il en découle que $A \setminus B = A \cap C_E(B)$.

Inclusion d'ensembles

Démontrer que A est inclus dans B, c'est démontrer que les éléments de A sont également des éléments de B, c'est à dire, pour tout élément x de E, on a : $x \in A \implies x \in B$. Pour établir ceci, on prend un élément x **quelconque** de E, et on démontre la proposition $x \in A \implies x \in B$.

Égalité d'ensembles

Démontrer que A est égal à B, c'est démontrer la double inclusion : $A \subset B$ et $B \subset A$, c'est à dire, pour tout élément x de E, on a : $(x \in A \implies x \in B)$ et $(x \in B \implies x \in A)$, ce qui équivaut à : $x \in A \iff x \in B$. Pour établir ceci, on prend un élément x **quelconque** de E, et on démontre la proposition $x \in A \iff x \in B$.

À retenir

Démontrer que $A \subset B$, c'est démontrer : $\forall x \in E, x \in A \implies x \in B$.
Démontrer que $A = B$, c'est démontrer : $\forall x \in E, x \in A \iff x \in B$.

☞**Exemples** :

– Soient A, B et C trois parties d'un ensemble E, démontrer la distributivité de la réunion sur l'intersection c'est démontrer :

$$\forall x \in E, \big(x \in A \cup (B \cap C)\big) \iff \big(x \in [A \cup B] \cap [A \cup C]\big)$$

On considère un x quelconque dans E, on peut alors montrer l'équivalence avec une table de vérité (à compléter en exercice) :

$x \in A$	$x \in B$	$x \in C$	$x \in A \cup (B \cap C)$	$x \in [A \cup B] \cap [A \cup C]$
V	V	V		
V	V	F		
V	F	V		
V	F	F		
F	V	V		
F	V	F		
F	F	V		
F	F	F		

– Soient A et B deux parties d'un ensemble E, soit x un élément quelconque de E, alors :

$$\begin{aligned} x \in C_E(A \cup B) &\iff \neg(x \in A \cup B) \\ &\iff \neg([x \in A] \vee [x \in B]) \\ &\iff [x \notin A] \wedge [x \notin B] \\ &\iff [x \in C_E(A)] \wedge [x \in C_E(B)] \\ &\iff x \in C_E(A) \cap C_E(B) \end{aligned}$$

Ce qui prouve que $C_E(A \cup B) = C_E(A) \cap C_E(B)$.

★**Exercice 1.4** *En s'inspirant des deux exemples ci-dessus, démontrer le théorème 1.1 et le théorème 1.2.*

Résoudre une équation

Une équation dans $\mathbb{R}$, d'inconnue x réelle peut toujours se mettre sous la forme $f(x) = 0$. Résoudre cette équation dans $\mathbb{R}$ c'est déterminer la partie S de $\mathbb{R}$ telle que « $\forall x \in \mathbb{R}, f(x) = 0 \iff x \in S$ » (S est appelé l'ensemble des solutions réelles). La définition est la même pour une inéquation dans $\mathbb{R}$.

☞**Exemple** : Pour résoudre dans $\mathbb{R}$ l'inéquation $\frac{1}{x} < -1$, on écrit :

$$\begin{aligned}\frac{1}{x} < -1 &\iff \frac{1}{x} + 1 < 0\\ &\iff \frac{x+1}{x} < 0\\ &\iff (x+1)x < 0 \text{ et } x \neq 0\\ &\iff x \in \left]-1\,;0\right[\end{aligned}$$

L'ensemble des solutions réelles est donc $]-1\,;0[$.

III LE RAISONNEMENT

1) Raisonnement par l'absurde

Soit P une proposition dont on cherche à démontrer qu'elle est vraie. Raisonner par l'absurde c'est faire l'hypothèse que P est fausse, autrement dit, on suppose non(P), on cherche alors à obtenir une contradiction, c'est à dire une proposition Q dont sait qu'elle est fausse, et telle que non(P) $\implies$ Q. Si on n'y parvient, cela veut dire que « Q et non(Q) » est vraie, ce qui est impossible car une telle proposition est toujours fausse : **c'est le principe de non contradiction**. La conclusion est que P est vraie.

☞**Exemple** : montrons que $\sqrt{2}$ est un irrationnel par l'absurde :

On suppose que $\sqrt{2} \in \mathbb{Q}$, on peut écrire $\sqrt{2} = \frac{p}{q}$ avec p et q entiers strictement positifs et premiers entre eux. En élevant au carré on a $2q^2 = p^2$, ce qui entraîne que p est pair et donc $p = 2a$ avec a entier, d'où $2q^2 = 4a^2$ *i.e.* $q^2 = 2a^2$, donc q est pair lui aussi et par conséquent p et q ne sont pas premiers entre eux : contradiction.

2) Raisonnement par analyse-synthèse

Raisonner par analyse-synthèse lorsque l'on cherche la ou les solutions à un problème, c'est raisonner en deux étapes qui sont :

– **l'analyse** : on suppose que l'on a une solution du problème et on cherche à en déduire toutes les propriétés possibles de cette solution, l'objectif étant d'essayer de l'identifier au mieux,
– **la synthèse** : elle consiste à déterminer parmi tous les objets mathématiques ayant les propriétés requises (obtenues lors de l'analyse), ceux qui sont effectivement solutions du problème.

☞**Exemple** : Montrons que toute fonction $f\colon [0;1] \to \mathbb{R}$ est la somme d'une fonction affine et d'une fonction qui s'annule en 0 et en 1.

• **Analyse** : supposons qu'il existe une fonction g qui s'annule en 0 et en 1, ainsi que deux réels a et b tels que $\forall x \in \mathbb{R}, f(x) = g(x) + ax + b$. En évaluant en 0 on doit avoir $f(0) = b$, en évaluant en 1 on doit avoir $f(1) = a + b$, d'où $a = f(1) - b = f(1) - f(0)$. Maintenant que a et b sont connus, on en déduit que g est la fonction $x \mapsto f(x) - ax - b$

• **Synthèse** : posons $b = f(0)$, $a = f(1) - f(0)$ et $g\colon x \mapsto f(x) - ax - b$. Il est clair que $f(x) = g(x) + ax + b$, d'autre part $g(0) = f(0) - b = 0$ et $g(1) = f(1) - a - b = f(1) - f(1) + f(0) - f(0) = 0$. Donc a, b et g sont bien solution du problème et celle-ci est unique.

3) Démontrer une implication

Par définition, l'implication « P $\implies$ Q » est fausse lorsque P est vraie et Q fausse, elle est vraie dans les autres cas. En particulier, si P est fausse, l'implication est nécessairement vraie quelque soit la valeur de vérité de Q. Par contre lorsque P est vraie, tout dépend de Q. Nous savons également que l'implication « P $\implies$ Q » est équivalente à sa contraposée « $\neg$Q $\implies$ $\neg$P », donc démontrer l'une c'est démontrer l'autre.

À retenir : pour démontrer une implication.

• **Méthode directe** : on suppose que la proposition P est vraie (c'est **l'hypothèse**), on cherche alors à établir que nécessairement la proposition Q est vraie elle aussi.
• **Par l'absurde** : on suppose le contraire de P $\implies$ Q, c'est à dire on suppose « P $\wedge$ $\neg$Q » (*i.e.* P

est vraie et Q est fausse. On montre alors que ceci conduit à une contradiction, ce qui entraîne que l'hypothèse faite est fausse et par conséquent $P \implies Q$.
• **Par contraposition** : on cherche à établir que $\neg Q \implies \neg P$.

Remarque 1.4 – *Pour démontrer « $P \vee Q$ » : cette proposition est équivalente à « $\neg P \implies Q$ ». Par conséquent, démontrer « $P \vee Q$ » revient à démontrer « $\neg P \implies Q$ ».*

4) L'équivalence

Par définition, l'équivalence « $P \iff Q$ » est vraie lorsque P et Q ont même valeur de vérité. Nous savons qu'elle équivaut à « $(P \implies Q)$ et $(Q \implies P)$ », donc montrer l'équivalence c'est montrer une implication **et** réciproque.

À retenir : pour démontrer une équivalence.

• **Par double implication** : on établit dans un premier temps que $P \implies Q$, puis dans un deuxième temps on établit la réciproque, c'est à dire que $Q \implies P$.
• **Méthode directe** : on suppose que la proposition P est vraie (**hypothèse**) puis on cherche à établir que Q est vraie **en s'assurant à chaque étape du raisonnement que l'équivalence est conservée**[6].

5) La récurrence

À retenir : rappel du principe de récurrence

Soit $P(n)$ un prédicat portant sur une variable $n \in \mathbb{N}$.
Si on a $P(0)$ (initialisation) et si $\forall n \in \mathbb{N}, P(n) \implies P(n+1)$ (hérédité), alors nécessairement $\forall n \in \mathbb{N}, P(n)$.

Remarque 1.5 :

– *L'initialisation est juste une vérification, mais elle est indispensable.*
Par exemple, soit le prédicat $P(n)$:« $n = n+1$ », celui-ci vérifie bien l'hérédité ($n = n+1 \implies n+1 = n+2$), mais pour tout n, $P(n)$ est fausse.
– *Démontrer l'hérédité c'est démontrer une implication. En général on le fait par la méthode directe, on fait donc l'hypothèse $P(n)$, c'est ce que l'on appelle **l'hypothèse de récurrence**, et on essaie d'en déduire $P(n+1)$.*

Théorème 1.5 (variantes)

Soit $a \in \mathbb{Z}$ et $P(n)$ un prédicat portant sur une variable $n \in \mathbb{Z}$.
• *Si on a $P(a)$ et si $\forall n \geqslant a, P(n) \implies P(n+1)$, alors on peut conclure que $\forall n \in \mathbb{Z}, n \geqslant a \implies P(n)$.*
• *Si on a $P(a)$ et si $\forall n \leqslant a, P(n) \implies P(n-1)$, alors on peut conclure que $\forall n \in \mathbb{Z}, n \leqslant a \implies P(n)$ (récurrence descendante).*

Preuve : Pour le premier point, on applique le principe de récurrence au prédicat $Q(n) = P(n+a)$ avec $n \in \mathbb{N}$. Pour le deuxième, on applique le principe de récurrence au prédicat $Q(n) = P(a-n)$ avec $n \in \mathbb{N}$. □

★Exercice 1.5

1/ Montrer que $\forall n \in \mathbb{N}^*, \sum_{k=1}^{n} k = \frac{n(n+1)}{2}$.

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

3/ Montrer que $\forall n \in \mathbb{N}^*, \sum_{k=1}^{n} k^3 = \left(\frac{n(n+1)}{2}\right)^2$.

6. Cette méthode n'est pas toujours applicable, mais c'est celle que l'on utilise dans la mesure du possible pour résoudre une équation ou une inéquation.

Théorème 1.6 (récurrence forte)
Soit $\mathrm{P}(n)$ *un prédicat portant sur une variable* $n \in \mathbb{N}$.
Si on a $\mathrm{P}(0)$ *(initialisation) et si* $\forall \in \mathbb{N}, (\mathrm{P}(0) \wedge \mathrm{P}(1) \wedge \cdots \wedge \mathrm{P}(n)) \implies \mathrm{P}(n+1)$ *(hérédité), alors nécessairement* $\forall n \in \mathbb{N}, \mathrm{P}(n)$.

Preuve : Appliquer la principe de récurrence au prédicat $\mathrm{Q}(n) = \mathrm{P}(0) \wedge \mathrm{P}(1) \wedge \cdots \wedge \mathrm{P}(n)$. □

À retenir
La récurrence forte est utile lorsque le seul fait que $\mathrm{P}(n)$ soit vraie ne suffit pas à en déduire $\mathrm{P}(n+1)$. L'hypothèse de récurrence peut alors s'écrire : « supposons la propriété vraie **jusqu'au** rang n ».

★**Exercice 1.6** *Soit* (u_n) *la suite définie par* $u_0 = u_1 = 1$ *et* $\forall\ n \in \mathbb{N}, u_{n+2} = u_{n+1} + u_n$ *(suite de Fibonacci), montrer par récurrence que pour tout n :*

$$u_n = \frac{5+\sqrt{5}}{10}\left(\frac{1+\sqrt{5}}{2}\right)^n + \frac{5-\sqrt{5}}{10}\left(\frac{1-\sqrt{5}}{2}\right)^n$$

IV SOLUTION DES EXERCICES

Solution 1.1 *Dans l'ensemble* $\{1;2;3\}$ *il y a exactement 8 parties, donc :*
$$\mathcal{P}(\mathrm{E}) = \{\varnothing; \{1\}; \{2\}; \{3\}, \{1;2\}; \{1;3\}; \{2;3\}; \{1;2;3\}\}$$

Solution 1.2 « $\mathrm{P} \implies \mathrm{Q}$ » *équivaut à* « $(\neg\mathrm{P}) \vee \mathrm{Q}$ », *donc* « $\neg(\mathrm{P} \implies \mathrm{Q})$ » *équivaut à* « $\neg((\neg\mathrm{P}) \vee \mathrm{Q})$ », *ce qui équivaut encore à* « $(\neg\neg\mathrm{P}) \wedge (\neg\mathrm{Q})$ » *(loi de De Morgan), soit encore à* « $\mathrm{P} \wedge (\neg\mathrm{Q})$ ».

Solution 1.3

1/ $\exists \mathrm{M} \in \mathbb{R}, \forall n \in \mathbb{N}, u_n \leqslant \mathrm{M}$. *La négation est* $\forall \mathrm{M} \in \mathbb{R}, \exists n \in \mathbb{N}, u_n > \mathrm{M}$. *La suite* (n^2) *n'est pas majorée, en effet : soit* $\mathrm{M} \in \mathbb{R}$, *si* $n \in \mathbb{N}$ *avec* $n > \mathrm{M}$, *alors* $(n+1)^2 > n > \mathrm{M}$.

2/ *Traduction :*

a) *L'ensemble* $\mathbb{R}$ *a un maximum, cette proposition est fausse, sa négation s'écrit* $\forall y \in \mathbb{R}, \exists x \in \mathbb{R}, x > y$.

b) *L'ensemble* $\mathbb{N}$ *a un minimum, cette proposition est vraie.*

Solution 1.4

1/ *Démontrer la distributivité de la intersection sur la réunion, c'est démontrer :*

$$\forall x \in \mathrm{E}, \big(x \in \mathrm{A} \cap (\mathrm{B} \cup \mathrm{C})\big) \iff \big(x \in [\mathrm{A} \cap \mathrm{B}] \cup [\mathrm{A} \cap \mathrm{C}]\big)$$

On considère un x quelconque dans E, *on peut alors montrer l'équivalence avec une table de vérité :*

$x \in \mathrm{A}$	$x \in \mathrm{B}$	$x \in \mathrm{C}$	$x \in \mathrm{A} \cap (\mathrm{B} \cup \mathrm{C})$	$x \in (\mathrm{A} \cap \mathrm{B}) \cup (\mathrm{A} \cap \mathrm{C})$
V	V	V	V	V
V	V	F	V	V
V	F	V	V	V
V	F	F	F	F
F	V	V	F	F
F	V	F	F	F
F	F	V	F	F
F	F	F	F	F

2/ *Pour le théorème 1.2, les trois premiers résultats sont évidents. Montrons les lois de De Morgan à l'aide d'une table de vérité, soient* A *et* B *deux parties de* E *et x un élément quelconque de* E *:*

$x \in \mathrm{A}$	$x \in \mathrm{B}$	$x \in \complement_{\mathrm{E}}(\mathrm{A} \cap \mathrm{B})$	$x \in \complement_{\mathrm{E}}(\mathrm{A}) \cup \complement_{\mathrm{E}}(\mathrm{B})$
V	V	F	F
V	F	V	V
F	V	V	V
F	F	V	V

On en déduit que $\complement_{\mathrm{E}}(\mathrm{A} \cap \mathrm{B}) = \complement_{\mathrm{E}}(\mathrm{A}) \cup \complement_{\mathrm{E}}(\mathrm{B})$. *On précède de même pour l'autre loi.*

Solution 1.5

1/ *On vérifie la relation pour* $n=1$ *:* $\sum_{k=1}^{1} k = 1 = \frac{1(1+1)}{2}$. *On suppose ensuite que la relation est vérifiée pour un certain entier* $n \in \mathbb{N}^*$, *c'est à dire que* $\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$ *pour un certain* $n \in \mathbb{N}^*$, *on a alors au rang suivant :*

$$\sum_{k=1}^{n+1} k = \sum_{k=1}^{n} k + (n+1) = \frac{n(n+1)}{2} + (n+1) = \frac{(n+1)(n+2)}{2} \quad \text{: c'est la formule au rang } n+1$$

Donc la formule est vraie pour tout $n \in \mathbb{N}^*$.

2/ *Même méthode.*

3/ *Même méthode.*

Solution 1.6 *Comme la relation de récurrence est à deux pas, on procède à une récurrence forte avec une initialisation pour* $n=0$ *et* $n=1$. *Notons* $a = \frac{5+\sqrt{5}}{10}$, $r_1 = \frac{1+\sqrt{5}}{2}$, $b = \frac{5-\sqrt{5}}{10}$ *et* $r_2 = \frac{1-\sqrt{5}}{2}$. *On vérifie qu'au rang* $n=0$ *on a* $ar_1^0 + br_2^0 = a+b = 1 = u_0$, *et que au rang* $n=1$ *on a* $ar_1^1 + br_2^1 = ar_1 + br_2 = 1 = u_1$. *Pour un entier* $n \geqslant 1$, *on suppose la formule établie pour tous les entiers* $k \in [\![0\,;n]\!]$ *(hypothèse de récurrence forte), on a* $u_{n+1} = u_n + u_{n-1}$, *comme* n *et* $n-1$ *sont dans* $[\![0\,;n]\!]$, *l'hypothèse permet d'écrire* $u_{n+1} = ar_1^n + br_2^n + ar_1^{n-1} + br_2^{n-1} = ar_1^{n-1}[r_1+1] + br_2^{n-1}[r_2+1]$, *or on peut vérifier que* $r_1 + 1 = r_1^2$ *(de même pour* r_2*), on en déduit que* $u_{n+1} = ar_1^{n+1} + br_2^{n+1}$, *c'est la formule au rang* $n+1$. *La formule est donc établie pour tout entier* n.

Chapitre 2

Généralités sur les fonctions

Sommaire

Dans ce chapitre, les fonctions considérées sont définies sur un intervalle non trivial de $\mathbb{R}$.

I RAPPELS ET COMPLÉMENTS SUR LES FONCTIONS

Rappels :

a) Une fonction f est la donnée de : un ensemble de départ E, un ensemble d'arrivée F, et pour chaque élément x de E, une image dans F, que l'on note $f(x)$. Le graphe de f est l'ensemble des couples $\{(x, f(x)) \,/\, x \in \mathrm{E}\}$. L'ensemble des fonctions de E vers F est noté $\mathcal{F}(\mathrm{E},\mathrm{F})$, ou encore F^{E}.

b) Deux fonctions sont égales lorsqu'elles ont le même ensemble de départ E, le même ensemble d'arrivée F, et que tout élément de E a la même image par les deux fonctions.

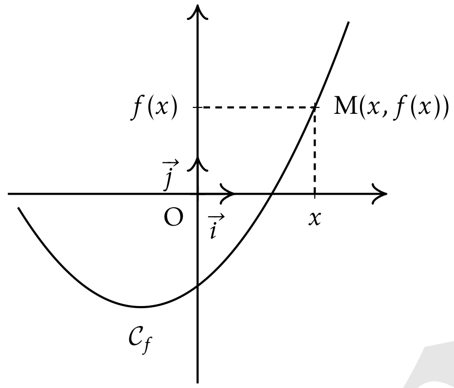
c) Si $f\colon \mathrm{I} \to \mathbb{R}$ est une fonction, on appelle « image de f », l'ensemble noté $f(\mathrm{I})$ et défini par :

$$f(\mathrm{I}) = \{y \in \mathbb{R} \,/\, \exists\, x \in \mathrm{I}, f(x) = y\},$$

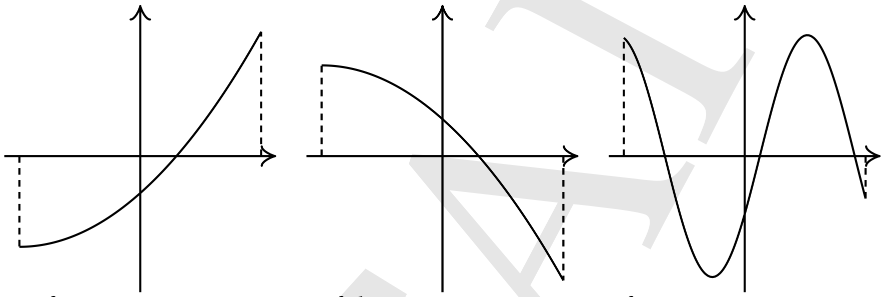
c'est l'ensemble des images par f des éléments de I. Celui-ci se détermine en général par l'étude de la fonction.

1) Vocabulaire

– **Représentation graphique**. Soit $\mathcal{P}$ un plan muni d'un repère $(\mathrm{O},\vec{\imath},\vec{\jmath})$, la représentation graphique de f est l'ensemble $\mathcal{C}_f = \{\mathrm{M}(x, f(x)) \mid x \in \mathrm{I}\}$. On dit que $y = f(x)$ est une équation de la courbe représentative de f car $\mathrm{M}(x;y) \in \mathcal{C}_f \iff y = f(x)$.



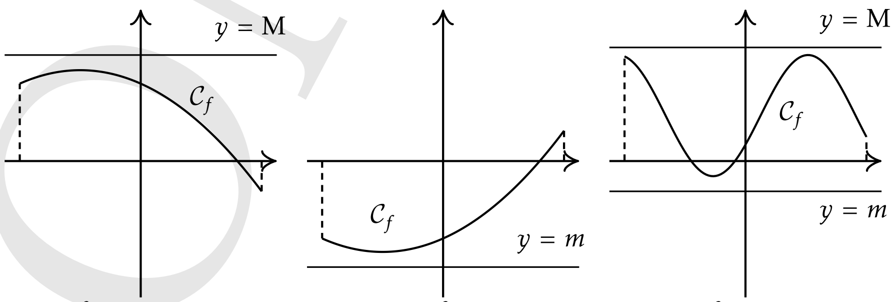
- **Sens de variation**. Soit $f\colon I \to \mathbb{R}$ une fonction, on dit que f est :
  - constante sur I lorsque : $\forall\, x, y \in I, f(x) = f(y)$.
  - croissante sur I lorsque : $\forall\, x, y \in I, x \leqslant y \implies f(x) \leqslant f(y)$.
  - strictement croissante sur I lorsque : $\forall\, x, y \in I, x < y \implies f(x) < f(y)$.
  - décroissante sur I lorsque : $\forall\, x, y \in I, x \leqslant y \implies f(x) \geqslant f(y)$.
  - strictement décroissante sur I lorsque : $\forall\, x, y \in I, x < y \implies f(x) > f(y)$.
  - monotone sur I lorsque : f est ou bien croissante ou bien décroissante.
  - strictement monotone sur I lorsque : f est ou bien strictement croissante ou bien strictement décroissante.



f croissante sur I $\qquad$ f décroissante sur I $\qquad$ f non monotone sur I

★**Exercice 2.1** *Étudier le sens de variation de la composée de deux fonctions monotones, puis de la somme, puis du produit.*

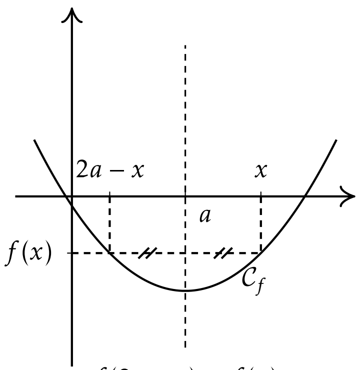
- **Fonction bornée**. Soit $f\colon I \to \mathbb{R}$ une fonction, on dit que f est :
  - majorée sur I lorsqu'il existe un réel M tel que $\forall\, x \in I, f(x) \leqslant M$. Si c'est le cas, la courbe de f est sous la droite $y = M$ (ex : la fonction $x \mapsto \frac{1}{x}$ est majorée sur $]-\infty\,; 0[$ par $M = 0$).
  - minorée sur I lorsqu'il existe un réel m tel que $\forall\, x \in I, f(x) \geqslant m$. Si c'est le cas, la courbe de f est au-dessus la droite $y = m$ (ex : la fonction $x \mapsto \frac{1}{x}$ est minorée sur $]0\,; +\infty[$ par $m = 0$).
  - bornée sur I lorsque f est à la fois minorée et majorée, **ce qui équivaut à : $|f|$ est majorée** (ex : la fonction $x \mapsto \sin(x)$ est bornée sur $\mathbb{R}$ car pour tout x, $|\sin(x)| \leqslant 1$).



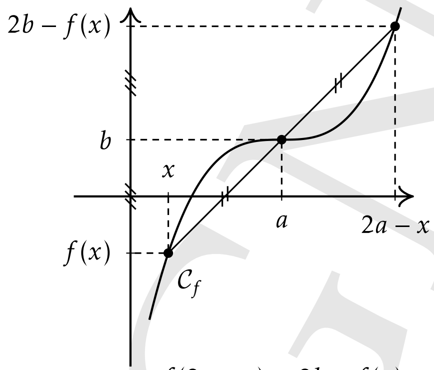
$f \leqslant M$ sur I $\qquad$ $m \leqslant f$ sur I $\qquad$ $m \leqslant f \leqslant M$ sur I

- **Parité**. Soit $f\colon I \to \mathbb{R}$ une fonction, on dit que f est :
  - paire lorsque : $\forall\, x \in I, -x \in I$ et $f(-x) = f(x)$. Dans ce cas la courbe représentative de f admet l'axe des ordonnées comme axe de symétrie (ex : la fonction carrée).
    Plus généralement, si $\forall\, x \in I, 2a - x \in I$ et $f(2a - x) = f(x)$, alors la courbe de f admet la droite d'équation $x = a$ comme axe de symétrie (ex : la fonction $f\colon x \mapsto \sin(x)$ vérifie $f(\pi - x) = f(x)$, donc la droite d'équation $x = \frac{\pi}{2}$ est un axe de symétrie pour la courbe).

- impaire lorsque : $\forall\, x \in \mathrm{I}, -x \in \mathrm{I}$ et $f(-x) = -f(x)$. Si c'est le cas, alors la courbe de f admet un centre de symétrie, l'origine du repère (ex : la fonction cube).
  Plus généralement, si $\forall\, x \in \mathrm{I}, 2a - x \in \mathrm{I}$ et $f(2a - x) = 2b - f(x)$, alors la courbe de f admet le point $\mathrm{A}(a, b)$ comme centre de symétrie (ex : la fonction $f\colon x \mapsto \cos(x)$ vérifie $f(\pi - x) = -f(x)$, donc le point de coordonnées $(\frac{\pi}{2}, 0)$ est un centre de symétrie pour la courbe).

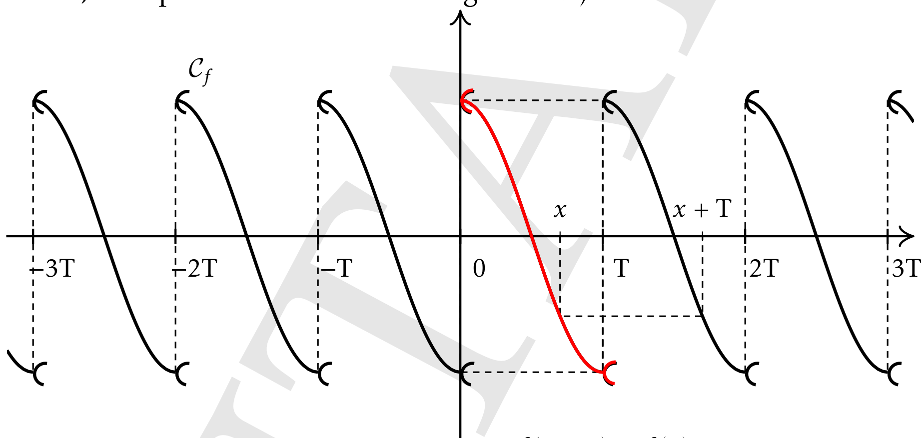


$f(2a - x) = f(x)$



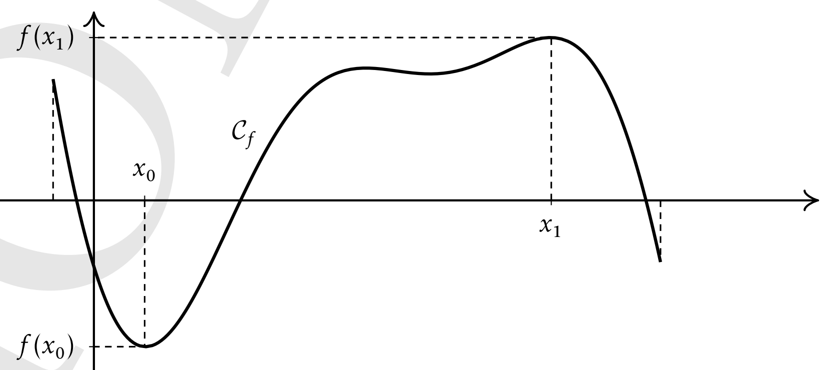
$f(2a - x) = 2b - f(x)$

– **Périodicité**. Soit $f\colon \mathrm{I} \to \mathbb{R}$ une fonction et soit $a \in \mathbb{R}^*$, on dit que a est une période de f lorsque : $\forall\, x \in \mathrm{I}, x \pm a \in \mathrm{I}$ et $f(x + a) = f(x)$. Si c'est le cas, le courbe de f est invariante par les translations de vecteurs $na\vec{i}$ où $n \in \mathbb{Z}$. Si f est périodique, on appelle période fondamentale de f la plus petite période strictement positive **si elle existe** (ex : les fonctions $x \mapsto \sin(x)$ et $x \mapsto \cos(x)$ sont périodique sur $\mathbb{R}$,et de période fondamentale égale à 2π).



$f(x + \mathrm{T}) = f(x)$

– Extremum global : on dit que $f\colon \mathrm{I} \to \mathbb{R}$ admet un :
  - maximum global en $x_0 \in \mathrm{I}$ lorsque : $\forall\, x \in \mathrm{I}, f(x) \leqslant f(x_0)$. Si c'est le cas, on pose $f(x_0) = \max\limits_{x \in \mathrm{I}} f(x)$.
  - minimum global en $x_0 \in \mathrm{I}$ lorsque : $\forall\, x \in \mathrm{I}, f(x) \geqslant f(x_0)$. Si c'est le cas, on pose $f(x_0) = \min\limits_{x \in \mathrm{I}} f(x)$.



f a un minimum en x_0 et un maximum en x_1

Attention !

Une fonction même bornée n'a pas forcément de maximum ou de minimum. Par exemple, la fonction $x \mapsto \frac{e^x-1}{e^x+1}$ a pour ensemble image $]-1;1[$, la fonction est donc bornée mais n'a ni maximum, ni minimum.

2) Opérations sur les fonctions

Définition 2.1

Soient $f, g \in \mathcal{F}(\mathrm{I}, \mathbb{R})$ et soit $\lambda \in \mathbb{R}$, on pose :
- *$f+g$ la fonction de I vers $\mathbb{R}$ définie par : $\forall x \in \mathrm{I}, (f+g)(x) = f(x)+g(x)$.*
- *$f \times g$ la fonction de I vers $\mathbb{R}$ définie par : $\forall x \in \mathrm{I}, (f \times g)(x) = f(x)g(x)$.*
- *$\lambda.f$ la fonction de I vers $\mathbb{R}$ définie par : $\forall x \in \mathrm{I}, (\lambda.f)(x) = \lambda f(x)$.*

Propriétés

a) Pour l'addition :
   - elle est commutative, associative,
   - elle admet un élément neutre : la fonction constamment nulle (notée 0),
   - toute fonction f de I vers $\mathbb{R}$ admet un opposé qui est la fonction $-f : x \mapsto -f(x)$,

b) Pour le produit par un réel : si $f, g \in \mathcal{F}(\mathrm{I}, \mathbb{R})$ et $\alpha, \beta \in \mathbb{R}$:
   - $1.f = f$,
   - $(\alpha+\beta).f = \alpha.f + \beta.f$,
   - $\alpha.(f+g) = \alpha.f + \alpha.g$,
   - $\alpha.(\beta.f) = (\alpha\beta).f$,

c) Pour la multiplication :
   - elle associative, commutative,
   - elle possède un élément neutre, la fonction constante qui à x donne 1 (notée 1),
   - elle est distributive sur l'addition,
   - seules les fonctions f qui **ne s'annulent jamais** ont un inverse (la fonction $\frac{1}{f}$).

Définition 2.2 (fonctions max et min)

Soient $f, g : \mathrm{I} \to \mathbb{R}$ deux fonctions, on pose $\max(f,g)$ et $\min(f,g)$ les fonctions de I vers $\mathbb{R}$ définies par : $\forall x \in \mathrm{I}, \max(f,g)(x) = \max(f(x), g(x))$ et $\min(f,g)(x) = \min(f(x), g(x))$. En particulier on pose $f^+ = \max(f,0)$ et $f^- = \max(-f,0)$, on a alors $f^+ - f^- = f$ et $f^+ + f^- = |f|$.

★**Exercice 2.2** *Montrer que $\max(f,g) = \frac{f+g+|f-g|}{2}$ et que $\min(f,g) = \frac{f+g-|f-g|}{2}$.*

II CONTINUITÉ, DÉRIVABILITÉ

1) Continuité

Définition 2.3

- Une fonction $f : \mathrm{I} \to \mathbb{R}$ est continue en $x_0 \in \mathrm{I}$, lorsque $\lim\limits_{x \to x_0} f(x) = f(x_0)$.

- On dit que f est continue sur I si elle est continue en tout point de I, et l'ensemble des fonctions continues sur I est noté $\mathcal{C}^0(\mathrm{I}, \mathbb{R})$.

Remarque 2.1 :
- *Si f a une limite à gauche en x_0 égale à $f(x_0)$, on dit que f est continue à gauche en x_0.*
- *Si f a une limite à droite en x_0 égale à $f(x_0)$, on dit que f est continue à droite en x_0.*
- *Si f est continue à gauche et à droite en x_0, alors f est continue en x_0.*

Théorème 2.1 (théorèmes généraux)

- *Les fonctions usuelles vues jusque là, sont continues sur leur ensemble de définition.*
- *Si $f, g\colon \mathrm{I} \to \mathbb{R}$ sont continues sur I alors les fonctions $f+g$, $f\times g$, $\lambda.f$ ($\lambda \in \mathbb{R}$) sont continues sur I.*
- *Si $f\colon \mathrm{I} \to \mathbb{R}$ est continue sur I **et ne s'annule pas**, alors $\frac{1}{f}$ est continue sur I.*
- *Si $f\colon \mathrm{I} \to \mathbb{R}$ est continue sur I, si $g\colon \mathrm{J} \to \mathbb{R}$ est continue sur J, et si $f(\mathrm{I}) \subset \mathrm{J}$, alors la composée $g\circ f$ est continue sur I.*

★**Exercice 2.3** *Étudier la continuité de la fonction f définie par $f(x) = |x\ln(|x|)|$.*

Théorème 2.2 (théorème des valeurs intermédiaires)

*Si f est continue sur **l'intervalle** I, alors l'ensemble $f(\mathrm{I})$ est un intervalle de $\mathbb{R}$. Plus précisément, soient $a, b \in \mathrm{I}$, pour tout réel α compris entre $f(a)$ et $f(b)$, il existe un réel c compris entre a et b tel que $f(c) = \alpha$.*

Théorème 2.3 (image d'un segment)

*Si f est continue sur un **segment** $[a;b]$ (avec $a < b$), alors f a un maximum (M) et un minimum (m). On a donc $f([a;b]) = [m;\mathrm{M}]$.*

Attention !

Une fonction continue sur un intervalle peut être majorée (ou minorée) sans avoir de maximum (ou de minimum). Par exemple, la fonction $x \mapsto \frac{1}{x}$ est continue sur $]0\,;+\infty[$, elle est minorée (par 0) mais n'a pas de minimum. La fonction $x \mapsto x$ et continue sur $[0\,;1[$, majorée, mais n'a pas de maximum.

Théorème 2.4 (de la bijection continue)

Si $f\colon \mathrm{I} \to \mathbb{R}$ est continue sur l'intervalle I et est strictement monotone sur I, alors elle réalise une bijection entre I et $f(\mathrm{I})$ (tout élément de $f(\mathrm{I})$ a un unique antécédent par f dans I). De plus, la bijection réciproque est continue sur l'intervalle $f(\mathrm{I})$. Dans un repère orthonormé, la courbe de la réciproque est l'image de la courbe de f par la symétrie orthogonale par rapport à la droite d'équation $y = x$.

Remarque 2.2 – *Ce théorème sera exploité dans le prochain chapitre pour définir de nouvelles fonctions.*

2) Dérivation

Définition 2.4

*Une fonction $f\colon \mathrm{I} \to \mathbb{R}$ est dérivable en $x_0 \in \mathrm{I}$, lorsque le taux d'accroissement $\frac{f(x)-f(x_0)}{x-x_0}$ admet une limite **finie** en x_0. Si c'est le cas, cette limite est notée $f'(x_0)$ et appelée nombre dérivé de f en x_0, et dans le plan muni d'un repère, la droite d'équation $y = f'(x_0)(x-x_0)+f(x_0)$ est appelé tangente à la courbe $\mathcal{C}_f$ au point d'abscisse x_0.*

Attention !

*Les fonctions usuelles vues jusque là sont dérivables sur leur ensemble de définition **SAUF** :*
- La valeur absolue qui n'est pas dérivable en 0.
- La racine carrée qui pas dérivable en 0.

Théorème 2.5 (théorèmes généraux)

- *Si f est dérivable en x_0 alors f est continue en x_0 (réciproque fausse).*
- *Si $f, g\colon \mathrm{I} \to \mathbb{R}$ sont dérivables sur I alors les fonctions $f+g$, $f\times g$, $\lambda.f$ ($\lambda \in \mathbb{R}$) sont dérivables sur I, et on a les formules de dérivation :*
$$(f+g)' = f'+g'\,;\,(fg)' = f'g+fg'\,;\,(\lambda f)' = \lambda f'.$$
- *Si $f\colon \mathrm{I} \to \mathbb{R}$ est dérivable sur I **et ne s'annule pas**, alors $\frac{1}{f}$ est dérivable sur I, et on a la formule :*

$$\left(\frac{1}{f}\right)' = \frac{-f'}{f^2}.$$

• *Si* $f\colon \mathrm{I} \to \mathbb{R}$ *est dérivable sur* I*, si* $g\colon \mathrm{J} \to \mathbb{R}$ *est dérivable sur* J*, et si* $f(\mathrm{I}) \subset \mathrm{J}$*, alors la composée* $g \circ f$ *est dérivable sur* I*, et on a la formule :*

$$\boxed{(g \circ f)' = f' \times g' \circ f} \text{ (ou encore } [g(f)]' = f' \times g'(f)).$$

Remarque 2.3 – *Si* f *et* g *sont dérivables et si* g *ne s'annule pas, alors* $\frac{f}{g}$ *est dérivable et* $\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}$.

Fonction	Dérivée
u^α (α constant)	$\alpha u' u^{\alpha-1}$
$\sqrt{u}$	$\frac{u'}{2\sqrt{u}}$
$\ln(\lvert u\rvert)$	$\frac{u'}{u}$
e^u	$u'e^u$
$\sin(u)$	$u'\cos(u)$
$\cos(u)$	$-u'\sin(u)$
$\tan(u) = \frac{\sin(u)}{\cos(u)}$	$u'(1 + \tan^2(u)) = \frac{u'}{\cos^2(u)}$

★**Exercice** 2.4 *Calculer la dérivée (si elle existe) des fonctions suivantes :* $\frac{1}{1+x^2}$; $e^{\frac{1}{\sqrt{x}}}$; $\cos^3(\frac{e^x}{x})$; $\ln(x + \sqrt{1+x^2})$.

Théorème 2.6 (de la bijection dérivable)
Si $f\colon \mathrm{I} \to \mathbb{R}$ *est continue, dérivable sur l'intervalle* I *et est strictement monotone sur* I*, alors elle réalise une bijection entre* I *et* $f(\mathrm{I})$*. On sait déjà que la bijection réciproque est continue sur l'intervalle* $f(\mathrm{I})$*. Si de plus la dérivée de* f *n'annule pas sur* I*, alors la bijection réciproque (notée* f^{-1}*) est dérivable sur* $f(\mathrm{I})$*, et on a la formule :* $\left(f^{-1}\right)' = \frac{1}{f' \circ f^{-1}}$.

Théorème 2.7 (sens de variation)
• *Une fonction dérivable* f *sur un intervalle* I *est constante si et seulement si sa dérivée est nulle sur* I.
• *Si* f *est dérivable sur un intervalle* I *et* $f' \geqslant 0$ *(respectivement* $f' \leqslant 0$ *alors* f *est croissante sur* I *(respectivement décroissante), et si de plus* f' *ne s'annule pas, alors la monotonie de* f *est stricte.*

3) Plan d'étude d'une fonction

Ensemble de définition, ensemble d'étude

– $\mathcal{D}_f$ est l'ensemble des réels de l'ensemble de départ ayant une image par f.
– Si $\mathcal{D}_f$ est symétrique par rapport à un réel a, il se peut que la courbe de f présente une symétrie :
  • un axe d'équation $x = a$ lorsque $\forall\, x \in \mathcal{D}_f, f(2a - x) = f(x)$.
  • un centre de symétrie de coordonnées (a, b) lorsque $\forall\, x \in \mathcal{D}_f, f(2a - x) = 2b - f(x)$.
  Dans les deux cas, on peut restreindre l'étude à $\mathcal{D}_f \cap [a\,;+\infty[$.
– S'il existe un réel $\mathrm{T} > 0$ tel que : $\forall\, x \in \mathcal{D}_f, x \pm \mathrm{T} \in \mathcal{D}_f, f(x + \mathrm{T}) = f(x)$, alors f est T-périodique. On peut restreindre l'étude à un intervalle de longueur une période : $\mathcal{D}_f \cap [a\,; a + \mathrm{T}[$ (a peut être quelconque), on complète ensuite la courbe avec les translations de vecteurs $n\mathrm{T}\vec{\imath}$, $n \in \mathbb{Z}$.

Continuité, dérivabilité

– On cherche à appliquer les théorèmes généraux, pour cela il faut regarder comment est faite la fonction (somme, produit, composée...).

– Il reste parfois des points où ces théorèmes ne s'appliquent as, on étudie alors la continuité en revenant à la définition (calcul de limite). S'il y a continuité, alors on étudie s'il y a dérivabilité en ce même point, il y a plusieurs méthodes : le théorème sur la limite de la dérivée, ou la définition (limite du taux d'accroissement).

Sens de variation

On rappelle que le théorème qui donne le sens de variation en fonction du signe de la dérivée, n'est valable que **sur un intervalle**.

– On peut parfois éviter l'étude du signe de la dérivée : sens de variation d'une somme, d'une composée, d'un produit... Par exemple, les fonctions $\ln(u)$, $\sqrt{u}$, e^u ont le même sens de variation que u.
– Lorsqu'on ne peut pas faire autrement, on étudie le signe de la dérivée (sur un intervalle).
– Les résultats sont consignés dans le tableau des variations, où doivent figurer :
  - l'ensemble d'étude,
  - les valeurs particulières qui sont intervenues dans l'étude de la continuité, la dérivabilité et l'étude du signe de la dérivée,
  - le signe de la dérivée (si on est passé par là),
  - les limites aux bornes de l'ensemble d'étude.

Étude des branches infinies

$\mathcal{C}_f$ désigne la courbe de f dans un repère orthogonal.

– Si x_0 est un réel de $\mathcal{D}_f$ ou une borne et si f a une limite infinie en x_0, alors on dit que $\mathcal{C}_f$ admet une asymptote verticale d'équation $x = x_0$ (ex : la fonction inverse en $x_0 = 0$ à gauche, et à droite).
– Si ∞ est une borne de $\mathcal{D}_f$, et si $\lim\limits_{\infty} f = \ell \in \mathbb{R}$, alors on dit que $\mathcal{C}_f$ admet une asymptote horizontale d'équation $y = \ell$ (ex : la fonction inverse en $\pm\infty$).
– Si ∞ est une borne de $\mathcal{D}_f$, si $\lim\limits_{\infty} f = \infty$, et si $\lim\limits_{x\to\infty} \frac{f(x)}{x} = \infty$, alors on dit que $\mathcal{C}_f$ admet une branche parabolique de direction $\mathrm{O}y$ (verticale) (ex : la fonction exponentielle en $+\infty$).
– Si ∞ est une borne de $\mathcal{D}_f$, si $\lim\limits_{\infty} f = \infty$, et si $\lim\limits_{x\to\infty} \frac{f(x)}{x} = 0$, alors on dit que $\mathcal{C}_f$ admet une branche parabolique de direction $\mathrm{O}x$ (horizontale) (ex : la fonction ln en $+\infty$).

Représentation graphique

– On commence par placer : les asymptotes, les tangentes remarquables, les points particuliers ;
– on donne ensuite **l'allure** de la courbe d'après le tableau de variation. Il est parfois nécessaire d'étudier la position de la courbe par rapport à certaines tangentes ou asymptotes.

III PRIMITIVES, INTÉGRALES

1) Généralités

Définition 2.5
*Soit $\mathrm{F}, f : \mathrm{I} \to \mathbb{R}$ deux fonctions, on dit que F est une **primitive** de f sur I lorsque F est dérivable sur I et $\mathrm{F}' = f$.*

Théorème 2.8
• Si F et G sont deux primitives de la fonction f sur l'intervalle I, alors il existe une constante $\alpha \in \mathbb{C}$ telle que : $\forall\, t \in \mathrm{I}, \mathrm{F}(t) = \mathrm{G}(t) + \alpha$.
• Si f admet des primitives sur l'intervalle I, si $x_0 \in \mathrm{I}$ et $a \in \mathbb{R}$, alors il existe une unique primitive F de f sur I telle que $\mathrm{F}(x_0) = a$.

Preuve : On a $\mathrm{F}' = \mathrm{G}' = f$, d'où $(\mathrm{F}-\mathrm{G})' = 0$ la fonction nulle, ce qui entraîne que la fonction $\mathrm{F}-\mathrm{G}$ est constantes sur l'intervalle I. □

Le théorème clé que nous établirons dans le chapitre sur l'intégration dit la chose suivante :

À retenir

Toute fonction $f\colon \mathrm{I} \to \mathbb{R}$ continue sur un intervalle I, admet des primitives sur cet intervalle.

Les primitives usuelles se lisent dans le tableau des dérivées usuelles, en faisant une lecture de droite à gauche.

★**Exercice 2.5** *Déterminer une primitive (en précisant l'intervalle) de :* $x \mapsto \tan(x)$; $x \mapsto \tan^2(x)$; $x \mapsto x\sqrt{1+x^2}$; $x \mapsto \frac{\ln(x)}{x}$; $x \mapsto \frac{1}{x\ln(x)}$.

2) Calculs d'intégrales et de primitives

Rappel

Si $f\colon \mathrm{I} \to \mathbb{R}$ est continue sur l'intervalle I, si F est une primitive de f sur I alors pour tous réels a et b de I, on a $\boxed{\int_a^b f(t)\,\mathrm{d}t = [\mathrm{F}(t)]_a^b = \mathrm{F}(b) - \mathrm{F}(a)}$.

À retenir

Soient $f, g\colon \mathrm{I} \to \mathbb{R}$ continues sur l'intervalle I, et $a, b \in \mathrm{I}$:

a) $\int_b^a f(t)\,\mathrm{d}t = -\int_a^b f(t)\,\mathrm{d}t$ et $\int_a^a f(t)\,\mathrm{d}t = 0$.

b) $\int_b^a [\alpha f(t) + \beta g(t)]\,\mathrm{d}t = \alpha \int_a^b f(t)\,\mathrm{d}t + \beta \int_a^b g(t)\,\mathrm{d}t$, c'est la **linéarité** de l'intégrale.

c) Si $0 \leqslant f$ sur $[a;b]$ (avec $a \leqslant b$), alors $0 \leqslant \int_a^b f(t)\,\mathrm{d}t$, c'est la **positivité** de l'intégrale. On en déduit que si $f \leqslant g$ sur $[a;b]$ (avec $a \leqslant b$) alors $\int_a^b f(t)\,\mathrm{d}t \leqslant \int_a^b g(t)\,\mathrm{d}t$ (c'est la **croissance** de l'intégrale).

d) Si a, b, c sont dans I, alors $\int_a^b f(t)\,\mathrm{d}t = \int_a^c f(t)\,\mathrm{d}t + \int_c^b f(t)\,\mathrm{d}t$, c'est la **relation de Chasles** pour l'intégrale.

e) Si $a \leqslant b$ alors $\left|\int_a^b f(t)\,\mathrm{d}t\right| \leqslant \int_a^b |f(t)|\,\mathrm{d}t$, c'est la **majoration en valeur absolue** de l'intégrale.

★**Exercice 2.6** *Montrer que* $\left|\int_{-1}^1 \frac{t^n}{1+t^2}\,\mathrm{d}t\right| \leqslant \frac{2}{n+1}$.

Théorème 2.9

Si $f\colon \mathrm{I} \to \mathbb{R}$ est continue sur I, si $\lambda \in \mathbb{R}$ et $x_0 \in \mathrm{I}$, alors l'unique primitive de f sur I qui prend la valeur λ en x_0 est la fonction F définie par : $\forall x \in \mathrm{I},\ \mathrm{F}(x) = \lambda + \int_{x_0}^x f(t)\,\mathrm{d}t$.

Preuve : Soit G une primitive de f sur I, on a $\mathrm{F}(x) = \lambda + \mathrm{G}(x) - \mathrm{G}(x_0) = \mathrm{G}(x) + \lambda - \mathrm{G}(x_0)$, donc F est une primitive de f car $\lambda - \mathrm{G}(x_0)$ est une constante, et $\mathrm{F}(x_0) = \lambda$. □

Les deux outils fondamentaux pour le calcul d'intégrales, sont : le théorème de l'intégration par parties et le théorème du changement de variable dont voici les énoncés :

Théorème 2.10 (IPP)

*Si f et g sont dérivables sur I avec leur **dérivée continue** (on dit qu'elles sont de classe $\mathcal{C}^1$), alors on a la formule **d'intégration par parties** (IPP) :*

$$\forall (a,b) \in \mathrm{I}^2,\ \int_a^b f'(t)g(t)\,\mathrm{d}t = [f(t)g(t)]_a^b - \int_a^b f(t)g'(t)\,\mathrm{d}t.$$

Preuve : Celle-ci est simple et laissée en exercice. □

Théorème 2.11 **(changement de variable)**
Soit $f\colon \mathrm{I} \to \mathbb{R}$ une fonction continue, $u : [a;b] \to \mathrm{I}$ une fonction dérivable à ***dérivée continue****, on a :*

$$\int_a^b f(u(t))u'(t)\,\mathrm{d}t = \int_{u(a)}^{u(b)} f(x)\,\mathrm{d}x$$

Preuve : Soit F une primitive de f sur I, alors $\mathrm{F}\circ u$ est une primitive de $u'f(u)$ et donc $\displaystyle\int_a^b f(u(t))u'(t)\,\mathrm{d}t = [\mathrm{F}(u)]_a^b = \mathrm{F}(u(b)) - \mathrm{F}(u(a)) = \int_{u(a)}^{u(b)} f(x)\,\mathrm{d}x$. □

Dans la pratique on rédige ainsi : posons $x = u(t)$ alors $\frac{dx}{dt} = u'(t)$ d'où $dx = u'(t)dt$ et $f(x) = f(u(t))$. Pour les bornes : lorsque $t = a$ on a $x = u(a)$ et pour $t = b$ on a $x = u(b)$, puis on remplace dans l'intégrale, ce qui donne : $\displaystyle\int_a^b f(u(t))u'(t)\,\mathrm{d}t = \int_{u(a)}^{u(b)} f(x)\,\mathrm{d}x$.

☞**Exemples** :

– Calculer $\displaystyle\int_0^1 \sqrt{1-x^2}\,\mathrm{d}x$. On pose $x = \sin(t)$ avec $t \in [0;\frac{\pi}{2}]$, sin est bien de classe $\mathcal{C}^1$ sur $[0;\frac{\pi}{2}]$. On a $\frac{dx}{dt} = \cos(t)$, d'où $dx = \cos(t)dt$. On a $x = 0$ lorsque $t = 0$, et $x = 1$ pour $t = \frac{\pi}{2}$, d'où :

$$\begin{aligned}\int_0^1 \sqrt{1-x^2}\,\mathrm{d}x &= \int_0^{\pi/2} \sqrt{1-\sin^2(t)}\cos(t)\,\mathrm{d}t\\ &= \int_0^{\pi/2} \cos^2(t)\,\mathrm{d}t\\ &= \int_0^{\pi/2} \frac{1+\cos(2t)}{2}\,\mathrm{d}t \quad (\text{car } \cos(2t) = 2\cos^2(t)-1)\\ &= \left[\frac{t}{2} + \frac{\sin(2t)}{4}\right]_0^{\pi/2} = \frac{\pi}{4} \quad (\text{car } \sin(0) = \sin(\pi) = 0)\end{aligned}$$

– Calculer une primitive de la fonction ln sur $]0\,;+\infty[$. Une primitive est (par exemple) $x \mapsto \displaystyle\int_1^x \ln(t)\,\mathrm{d}t$ pour $x > 0$, cette intégrale se calcule par parties en posant $f(t) = t$ et $g(t) = \ln(t)$, ces fonctions sont de classe $\mathcal{C}^1$ sur $]0\,;+\infty[$, d'où :

$$\begin{aligned}\int_1^x \ln(t)\,\mathrm{d}t = [f(t)g(t)]_1^x - \int_1^x f(t)g'(t)\,\mathrm{d}t = [t\ln(t)]_1^x - \int_1^x 1\,\mathrm{d}t\\ = x\ln(x) - (x-1) = x\ln(x) - x + 1\end{aligned}$$

donc une primitive de la fonction ln sur $]0\,;+\infty[$ est la fonction $x \mapsto x\ln(x) - x$ (on peut évidemment ajouter n'importe quelle constante).

★**Exercice** 2.7

1/ *Déterminer une primitive de $x \mapsto x^2e^x$ à l'aide d'IPPs.*

2/ *Déterminer une primitive de $x \mapsto \cos^5(x)$ avec la changement de variable $u = \sin(x)$.*

3) Primitives de certaines fractions rationnelles

– Fractions du type $f(x) = \frac{1}{(x-a)^n}$ avec $a \in \mathbb{R}$ et $n \in \mathbb{N}^*$.
Sur l'intervalle $\mathrm{I} =]-\infty\,;a[$ (ou $]a\,;+\infty[$), f est continue et admet donc des primitives. Si $n = 1$ alors une primitive est $\mathrm{F}(x) = \ln(|x-a|)$ et si $n \geqslant 2$, une primitive est $\mathrm{F}(x) = \frac{-1}{(n-1)(x-a)^{n-1}}$.

– Fractions du type $f(x) = \frac{\alpha x+\beta}{(x-a)(x-b)}$ avec α, β, a et b des réels tels que $\boxed{a \neq b}$.

À retenir

Il existe c et d réels tels que $\frac{\alpha x+\beta}{(x-a)(x-b)} = \frac{c}{x-a} + \frac{d}{x-b}$, ce qui nous ramène au cas précédent.

En effet : en réduisant au même dénominateur, on a au numérateur $(c+d)x-(bc+ad)$, il suffit donc de choisir c et d tels que $\begin{cases} c+d=\alpha \\ bc+ad=-\beta \end{cases}$ ce qui équivaut à $\begin{cases} c+d=\alpha \\ (b-a)d=\beta+b\alpha \end{cases}$ et on voit que ce système a une unique solution puisque $a \neq b$.

– Fractions du type $f(x) = \frac{\alpha x+\beta}{(x-a)^2}$ avec α, β et a des réels.

À retenir

On écrit $f(x) = \frac{\alpha(x-a+a)+\beta}{(x-a)^2} = \frac{\alpha}{x-a} + (\alpha a+\beta)\frac{1}{(x-a)^2}$, ce qui s'intègre ensuite facilement (on un terme en $\frac{u'}{u}$ et un autre en $\frac{u'}{u^2}$).

– Fractions du type $f(x) = \frac{ax+b}{x^2+px+q}$ avec a, b, p q des réels tels que $p^2-4q<0$. Le dénominateur n'a pas de racine réelle, f est donc définie sur $\mathbb{R}$. La méthode est la suivante :

À retenir

• on fait apparaître la dérivée du trinôme x^2+px+q au numérateur et on compense les x en multipliant par un facteur adéquat, puis on compense les constantes en ajoutant ce qu'il faut, ce qui donne :

$$\frac{ax+b}{x^2+px+q} = \frac{a}{2}\frac{2x+p}{x^2+px+q} + (b-\frac{ap}{2})\frac{1}{x^2+px+q}.$$

La première de ces deux fractions est facile à intégrer puisqu'elle est du type $\frac{u'}{u}$.
• Pour la deuxième fraction : on met le trinôme x^2+px+q sous forme canonique afin de mettre la fraction sous la forme : $\alpha\frac{u'}{1+u^2}$ où α est une constante et u est une fonction de x, cette fonction s'intègre en $\alpha\arctan(u)$[a].

a. La fonction arctan sera étudiée dans le chapitre suivant.

Par exemple, soit $f(x) = \frac{x-2}{x^2-x+1}$:

$$f(x) = \frac{x-2}{x^2-x+1} = \frac{1}{2}\frac{2x-1}{x^2-x+1} - \frac{3}{2}\frac{1}{x^2-x+1}$$

et :

$$\frac{1}{x^2-x+1} = \frac{1}{(x-\frac{1}{2})^2+\frac{3}{4}} = \frac{2}{\sqrt{3}}\frac{2/\sqrt{3}}{\left(\frac{2x-1}{\sqrt{3}}\right)^2+1}.$$

On en déduit qu'une primitive de f sur $\mathbb{R}$ est $\mathrm{F}: x \mapsto \frac{1}{2}\ln(x^2-x+1) - \sqrt{3}\arctan(\frac{2x-1}{\sqrt{3}})$.

IV SOLUTION DES EXERCICES

Solution 2.1

1/ Si $f: \mathrm{I} \to \mathbb{R}$ et $g: \mathrm{J} \to \mathbb{R}$ sont monotones et si $f(\mathrm{I}) \subset \mathrm{J}$, alors on vérifie que $g \circ f$ est :
– croissante si f et g ont le même sens de variation ;
– décroissante dans le cas contraire.

2/ Si $f: \mathrm{I} \to \mathbb{R}$ et $\mathrm{I}: \mathrm{J} \to \mathbb{R}$ sont monotones, alors on vérifie que $f+g$ est :
– croissante si f et g sont croissantes ;
– décroissante si f et g sont décroissantes.
On ne peut rien dire en général sinon.

3/ Si $f: \mathrm{I} \to \mathbb{R}^+$ et $\mathrm{I}: \mathrm{J} \to \mathbb{R}^+$ sont monotones et positives, alors on vérifie que $f+g$ est :
– croissante si f et g sont croissantes ;

– *décroissante si f et g sont décroissantes.*
On ne peut rien dire en général sinon.

Solution 2.2 *Simple vérification. Soit $x \in \mathrm{I}$, supposons $f(x) \leqslant g(x)$, alors $\frac{f(x)+g(x)+|f(x)-g(x)|}{2} = \frac{f(x)+g(x)-f(x)+g(x)}{2} = g(x)$ et c'est bien le maximum entre $f(x)$ et $g(x)$. L'autre cas se traite de la même façon en échangeant les rôles de f et de g, d'où $\max(f,g) = \frac{f+g+|f-g|}{2}$.*

L'autre égalité se traite sur le même principe, ou bien en remarquant que pour tous réels a et b, $\min(a,b) = -\max(-a,-b)$.

Solution 2.3 *La fonction f est paire et définie sur $\mathbb{R}^*$. Sur $]0\,;+\infty[$ on a $f(x) = x|\ln(x)|$ donc g est continue sur $]0\,;+\infty[$ (en appliquant les théorèmes généraux) ; f étant paire, elle continue sur $\mathbb{R}^*$. On a $\lim\limits_{x\to 0} f(x) = 0$, il y a donc un prolongement par continuité en 0 en posant $f(0) = 0$, ainsi prolongée, f est continue sur $\mathbb{R}$.*

Solution 2.4

1/ $\left[\frac{1}{1+x^2}\right]' = \frac{-2x}{(1+x^2)^2}$, *sur* $\mathbb{R}$.

2/ $\left[e^{\frac{1}{\sqrt{x}}}\right]' = -\frac{1}{2x\sqrt{x}}e^{\frac{1}{\sqrt{x}}}$, *sur* $\mathbb{R}^{*+}$.

3/ $\left[\cos^3(\frac{e^x}{x})\right]' = -\frac{e^x(x-1)}{x^2}\sin(\frac{e^x}{x})\cos^2(\frac{e^x}{x})$, *sur* $\mathbb{R}^{*+}$.

4/ $\left[\ln(x+\sqrt{1+x^2})\right]' = \frac{1}{\sqrt{1+x^2}}$, *sur* $\mathbb{R}$.

Solution 2.5

1/ *Une primitive sur* $]-\frac{\pi}{2}\,;\frac{\pi}{2}[$ *est* $\mathrm{F}(x) = -\ln(|\cos(x)|)$.

2/ *Une primitive sur* $]-\frac{\pi}{2}\,;\frac{\pi}{2}[$ *est* $\mathrm{F}(x) = \tan(x) - 1$.

3/ *Une primitive sur* $\mathbb{R}$ *est* $\mathrm{F}(x) = \frac{1}{3}(1+x^2)^{\frac{3}{2}}$.

4/ *Une primitive sur* $]0\,;+\infty[$ *est* $\mathrm{F}(x) = \frac{1}{2}\ln^2(x)$.

5/ *Une primitive sur* $]0\,;1[$ *est* $\mathrm{F}(x) = \ln(-\ln(x))$.

Solution 2.6 $\left|\int_{-1}^{1}\frac{t^n}{1+t^2}\,\mathrm{d}t\right| \leqslant \int_{-1}^{1}\left|\frac{t^n}{1+t^2}\right|\mathrm{d}t$ *(majoration en valeur absolue), or* $\left|\frac{t^n}{1+t^2}\right| = \frac{|t|^n}{1+t^2} \leqslant |t|^n$ *car* $\frac{1}{1+t^2} \leqslant 1$, *d'où* $\left|\int_{-1}^{1}\frac{t^n}{1+t^2}\,\mathrm{d}t\right| \leqslant \int_{-1}^{1}|t|^n\,\mathrm{d}t = \int_{0}^{1}t^n\,\mathrm{d}t + \int_{-1}^{0}(-1)^n t^n\,\mathrm{d}t = \frac{2}{n+1}$.

Solution 2.7

1/ *Soit $f(x) = x^2e^x$, f est continue sur $\mathbb{R}$, elle admet donc des primitives sur $\mathbb{R}$, la primitive qui s'annule en 0 (par exemple) est F définie par $\mathrm{F}(x) = \int_0^x t^2e^t\,\mathrm{d}t$. On fait une IPP en posant $u' = e^t$ et $v = t^2$, d'où $u = e^t$ et $v' = 2t$ (u et v sont bien dérivables à dérivée continue), on a $\mathrm{F}(x) = [uv]_0^x - \int_0^x uv' = x^2e^x - 2\int_0^x te^t\,\mathrm{d}t$. Pour la deuxième intégrale on refait une IPP en posant $u' = e^t$ et $v = t$, d'où $u = e^t$ et $v' = 1$ (u et v sont bien dérivables à dérivée continue), on a $\int_0^x te^t\,\mathrm{d}t = [uv]_0^x - \int_0^x uv' = xe^x - \int_0^x e^t\,\mathrm{d}t = xe^x - e^x + 1$, d'où $\mathrm{F}(x) = (x^2-2x+2)e^x - 2$.*

2/ *La fonction est continue sur $\mathbb{R}$, elle admet donc des primitives. La primitive qui s'annule en 0 (par exemple) est F définie par $\mathrm{F}(x) = \int_0^x \cos^5(t)\,\mathrm{d}t = \int_0^x \cos^4(t)\cos(t)\,\mathrm{d}t$. On pose $u = \sin(t)$ (continue dérivable, à dérivée continue sur $\mathbb{R}$), d'où $\mathrm{d}u = \cos(t)\,\mathrm{d}t$ et donc :*

$\mathrm{F}(x) = \int_{\sin(0)}^{\sin(x)}(1-u^2)^2\,\mathrm{d}u = \int_0^{\sin(x)}(u^4-2u^2+1)\,\mathrm{d}u\left[\frac{1}{5}u^5 - \frac{2}{3}u^3 + u\right]_0^{\sin(x)} = \frac{\sin^5(x)}{5} - \frac{2\sin^3(x)}{3} + \sin(x)$.

Chapitre 3

Nombres complexes

Sommaire

On suppose connu l'ensemble $\mathbb{R}$ des réels, les propriétés de ses deux opérations que sont l'addition et la multiplication, et la notion de valeur absolue d'un réel.

I ÉCRITURE ALGÉBRIQUE

1) L'ensemble des complexes

On admet l'existence d'un ensemble, que l'on notera $\mathbb{C}$, contenant $\mathbb{R}$ et tel que :

– $\mathbb{C}$ possède une addition et une multiplication, **prolongeant celles de** $\mathbb{R}$ et vérifiant les mêmes propriétés, c'est à dire :
  - L'addition est interne ($\forall z, z' \in \mathbb{C}, z+z' \in \mathbb{C}$), commutative ($\forall z, z' \in \mathbb{C}, z+z' = z'+z$), associative ($\forall z, z', z'' \in \mathbb{C}, (z+z')+z'' = z+(z'+z'')$), 0 est élément neutre ($\forall z \in \mathbb{C}, z+0 = 0+z = z$), chaque complexe admet un opposé dans $\mathbb{C}$ (l'opposé de $z \in \mathbb{C}$ est noté $-z$).
  - La multiplication est interne ($\forall z, z' \in \mathbb{C}, zz' \in \mathbb{C}$), commutative ($\forall z, z' \in \mathbb{C}, zz' = z'z$), associative ($\forall z, z', z'' \in \mathbb{C}, (zz')z'' = z(z'z'')$), 1 est élément neutre ($\forall z \in \mathbb{C}, z \times 1 = 1 \times z = z$), chaque complexe **autre que** 0 admet un inverse dans $\mathbb{C}$ (l'inverse de $z \in \mathbb{C}$ est noté z^{-1} ou $\frac{1}{z}$).
  - La multiplication est distributive sur l'addition à gauche et à droite ($\forall z, z', z'' \in \mathbb{C}, z(z'+z'') = zz'+zz''$, et $(z'+z'')z = z'z+z''z$).

On résume toutes ses propriétés en disant que $(\mathbb{C}, +, \times)$ est un corps commutatif.

– Il existe un complexe, que l'on notera i, qui vérifie $i^2 = -1$.
– Pour tout complexe z, il existe deux réels a et b **uniques** tels que $z = a + ib$.

★**Exercice** 3.1

1/ *Montrer que si $z \in \mathbb{C}$, alors $z \times 0 = 0 \times z = 0$.*

2/ *Soient $z, z' \in \mathbb{C}$ tels que $zz' = 0$, montrer que $z = 0$ ou $z' = 0$.*

2) Partie réelle, partie imaginaire

Définition 3.1

*Soit $z \in \mathbb{C}$, on sait qu'il existe un réel a unique et un réel b unique tel que $z = a + ib$. Le réel a est appelé partie réelle de z, notée $a = \mathrm{Re}(z)$, et b la partie imaginaire de z, notée $b = \mathrm{Im}(z)$. L'écriture $z = a + ib$ est appelée **forme algébrique** de z. L'ensemble des complexes dont la partie réelle est nulle, est appelé ensemble des imaginaires purs et noté $i\mathbb{R} = \{z \in \mathbb{C} \mid \mathrm{Re}(z) = 0\} = \{iy \mid y \in \mathbb{R}\}$.*

À retenir : quelques propriétés

a) $z = z' \iff \begin{cases} \mathrm{Re}(z) = \mathrm{Re}(z') \\ \mathrm{Im}(z) = \mathrm{Im}(z') \end{cases}$.

b) $z \in \mathbb{R} \iff \mathrm{Im}(z) = 0$.

c) $z \in i\mathbb{R} \iff \mathrm{Re}(z) = 0$.

d) $\mathrm{Re}(z + z') = \mathrm{Re}(z) + \mathrm{Re}(z')$ et $\mathrm{Im}(z + z') = \mathrm{Im}(z) + \mathrm{Im}(z')$.

e) Si α est un **réel**, alors $\mathrm{Re}(\alpha z) = \alpha\,\mathrm{Re}(z)$, et $\mathrm{Im}(\alpha z) = \alpha\,\mathrm{Im}(z)$.

★**Exercice** 3.2

1/ *Démontrer les propriétés ci-dessus.*

2/ *Déterminer la forme algébrique du produit zz'.*

Attention !

La dernière propriété est fausse lorsque α n'est pas réel.

3) Conjugué d'un nombre complexe

Définition 3.2

*Soit $z = x + iy$ un complexe sous forme algébrique, on appelle **conjugué** de z, le complexe noté $\overline{z}$ et défini par $\overline{z} = x - iy$. On a donc $\mathrm{Re}(\overline{z}) = \mathrm{Re}(z)$ et $\mathrm{Im}(\overline{z}) = -\mathrm{Im}(z)$.*

Théorème 3.1 (propriétés de la conjugaison)

Soient $z, z' \in \mathbb{C}$, on a :
- $\overline{z + z'} = \overline{z} + \overline{z'}$
- $\overline{zz'} = \overline{z}\,\overline{z'}$
- $\overline{\overline{z}} = z$.

Preuve : Celle-ci est simple et laissée en exercice. □

À retenir

- $z + \overline{z} = 2\,\mathrm{Re}(z)$ et $z - \overline{z} = 2i\,\mathrm{Im}(z)$;
- $z \in \mathbb{R} \iff z = \overline{z}$;
- z est un imaginaire pur si et seulement si $z = -\overline{z}$.

★**Exercice** 3.3 *Montrer que si z est non nul, alors $\overline{\frac{1}{z}} = \frac{1}{\overline{z}}$.*

II MODULE D'UN NOMBRE COMPLEXE

1) Définition

Soit $z = x + iy$ un complexe sous forme algébrique, on a $z \times \overline{z} = x^2 + y^2$ et cette quantité est un **réel positif**.

Définition 3.3

*Soit $z \in \mathbb{C}$, on appelle **module** de z, le réel positif noté $|z|$ et défini par : $|z| = \sqrt{z\overline{z}}$.*

À retenir : propriétés du module

a) $|z|^2 = z\overline{z}$.

b) $|z| = 0 \iff z = 0$.

c) $|\operatorname{Re}(z)| \leqslant |z|$ et $|\operatorname{Im}(z)| \leqslant |z|$.

d) Si z est réel, alors son module coïncide avec sa valeur absolue.

e) $|zz'| = |z||z'|$, en particulier, $\forall n \in \mathbb{N},\ |z^n| = |z|^n$ (ceci reste valable pour $n \in \mathbb{Z}$ si $z \neq 0$).

f) $|z| = |\overline{z}|$ et $|-z| = |z|$.

g) Pour mettre le complexe $\frac{z}{z'}$ sous forme algébrique, il suffit de multiplier en haut et en bas par $\overline{z'}$: $\frac{z}{z'} = \frac{z\overline{z'}}{|z'|^2} = \frac{\operatorname{Re}(z\overline{z'})}{|z'|^2} + i\frac{\operatorname{Im}(z\overline{z'})}{|z'|^2}$.

Théorème 3.2 (inégalité triangulaire)

Soient z et z' deux complexes : $\big||z| - |z'|\big| \leqslant |z + z'| \leqslant |z| + |z'|$.

Preuve : Pour l'inégalité de droite :

$$\begin{aligned}
|z + z'|^2 &= (z + z')(\overline{z} + \overline{z'}) \\
&= |z|^2 + |z'|^2 + 2\operatorname{Re}(z\overline{z'}) \\
&\leqslant |z|^2 + |z'|^2 + 2|\operatorname{Re}(z\overline{z'})| \\
&\leqslant |z|^2 + |z'|^2 + 2|z\overline{z'}| = |z|^2 + |z'|^2 + 2|z||\overline{z'}| \\
&\leqslant (|z| + |z'|)^2
\end{aligned}$$

d'où le résultat puisque les deux expressions sont des réels positifs.

Pour l'inégalité de gauche, on écrit $|z| = |z + z' - z'| \leqslant |z + z'| + |z'|$ d'où $|z| - |z'| \leqslant |z + z'|$, en inversant les rôles on a $|z'| - |z| \leqslant |z + z'|$, ce qui entraîne la première inégalité. □

Remarque 3.1 – *En remplaçant z' par $-z'$, on a $\big||z| - |z'|\big| \leqslant |z - z'| \leqslant |z| + |z'|$.*

Théorème 3.3 (cas d'égalité)

*Soient z et z' deux complexes non nuls, $|z + z'| = |z| + |z'|$ si et seulement si il existe un **réel** α **strictement positif** tel que $z = \alpha z'$.*

Preuve : Si on a $z = \alpha z'$, alors $|z + z'| = |\alpha z' + z'| = (1 + \alpha)|z'| = |z'| + \alpha|z'| = |z'| + |z|$. Réciproquement, si $|z+z'| = |z|+|z'|$, alors $|z+z'|^2 = (|z|+|z'|)^2$, ce qui donne en développant, $|z|^2+|z'|^2+2\operatorname{Re}(z\overline{z'}) = |z|^2+|z'|^2+2|z||z'|$, on en déduit que $\operatorname{Re}(z\overline{z'}) = |z\overline{z'}|$ ce qui prouve que $z\overline{z'}$ est un réel positif. Il suffit alors de prendre $\alpha = \frac{z\overline{z'}}{|z'|^2}$, c'est bien un réel strictement positif, et on a la relation voulue. □

2) Équation du second degré

Théorème 3.4

Soit $a \in \mathbb{C}$, l'équation $z^2 = a$ admet dans $\mathbb{C}$ deux solutions opposées (toutes deux nulles lorsque $a = 0$).

Preuve : Soit z_0 une solution, alors l'équation $z^2 = a$ équivaut à $z^2 = z_0^2$, c'est à dire à $(z - z_0)(z + z_0) = 0$, d'où $z = \pm z_0$, il reste à montrer l'existence d'une solution z_0. Posons $a = u + iv$ et $z = x + iy$, l'équation $z^2 = a$ est équivalente à $x^2 - y^2 = u$ et $2xy = v$. On doit avoir également $|z|^2 = |a|$, c'est à dire $x^2 + y^2 = |a|$, par conséquent on a : $x^2 = \frac{u+|a|}{2}$, $y^2 = \frac{|a|-u}{2}$ et $2xy = v$. Une solution $z_0 = x_0 + iy_0$ s'obtient en prenant : $x_0 = \sqrt{\frac{|a|+u}{2}}$, et $y_0 = \varepsilon\sqrt{\frac{|a|-u}{2}}$ avec $\varepsilon = 1$ si $v \geqslant 0$ et $\varepsilon = -1$ si $v < 0$, car on a $2x_0y_0 = \varepsilon|v| = v$. □

☞**Exemples** :
- Si a est un réel strictement positif, alors $v = 0$ et $u > 0$ d'où $|a| = u$ et donc $x_0 = \sqrt{a}$ et $y_0 = 0$, les deux solutions sont $\pm\sqrt{a}$, elles sont réelles.
- Si a est un réel strictement négatif, alors $v = 0$ et $u < 0$ d'où $|a| = -u$ et donc $x_0 = 0$ et $y_0 = \sqrt{-a}$, les deux solutions sont $\pm i\sqrt{-a}$, ce sont des **imaginaires purs**.

Théorème 3.5

Soient $a, b, c \in \mathbb{C}$ avec $a \neq 0$, l'équation $az^2 + bz + c = 0$ admet deux solutions complexes qui sont $z_1 = \frac{-b+\delta}{2a}$ et $z_2 = \frac{-b-\delta}{2a}$ avec $\delta \in \mathbb{C}$ tel que $\delta^2 = \Delta = b^2 - 4ac$ (discriminant). De plus, lorsque les coefficients a, b, c sont réels et que le discriminant $b^2 - 4ac$ est strictement négatif, ces deux solutions sont complexes ***non réelles et conjuguées****.*

Preuve : L'équation est équivalente à : $(z + \frac{b}{2a})^2 - \frac{b^2-4ac}{4a^2} = 0$. Posons $Z = z + \frac{b}{2a}$ et $\Delta = b^2 - 4ac$, on sait que Δ admet deux racines carrées dans $\mathbb{C}$, soit δ l'une d'elles ($\delta^2 = \Delta$), l'équation est équivalente à : $Z^2 = \frac{\delta^2}{4a^2}$, on en déduit que $Z = \pm\frac{\delta}{2a}$ et donc $z = \frac{-b\pm\delta}{2a}$. Lorsque les trois coefficients sont réels, le discriminant Δ est lui aussi un réel, s'il est strictement négatif, alors on peut prendre $\delta = i\sqrt{-\Delta}$ et les solutions sont dans ce cas $z = \frac{-b\pm i\sqrt{-\Delta}}{2a}$, on voit que celles - ci sont complexes non réelles et conjuguées. □

À retenir : somme et produit de racines d'une équation du second degré

• Si z_1 et z_2 sont racines de l'équation $az^2 + bz + c = 0$, alors on a les relations : $z_1 + z_2 = S = -\frac{b}{a}$ et $z_1z_2 = P = \frac{c}{a}$.
• De plus on a la factorisation :
$\forall z \in \mathbb{C}, az^2 + bz + c = a(z - z_1)(z - z_2)$.

III FORME TRIGONOMÉTRIQUE

1) Le groupe unité

Définition 3.4

On note $\mathbb{U}$ l'ensemble des complexes de module 1 : $\mathbb{U} = \{z \in \mathbb{C} \ / \ |z| = 1\}$, c'est une partie de $\mathbb{C}^*$.

Il est facile de vérifier que l'ensemble $\mathbb{U}$:
- est **stable pour la multiplication** : $\forall z, z' \in \mathbb{U}, zz' \in \mathbb{U}$.
- est **stable pour le passage à l'inverse** : $\forall z \in \mathbb{U}$, $z \neq 0$ et $z^{-1} \in \mathbb{U}$.
- contient 1.

De plus, la multiplication dans $\mathbb{U}$ est associative (car elle l'est dans $\mathbb{C}$), on dit alors que $(\mathbb{U}, \times)$ est un **groupe multiplicatif**. Comme la multiplication est en plus commutative, on dit que $(\mathbb{U}, \times)$ est un **groupe abélien** (ou commutatif), ce groupe est parfois appelé **groupe unité** de $\mathbb{C}$.

Théorème 3.6

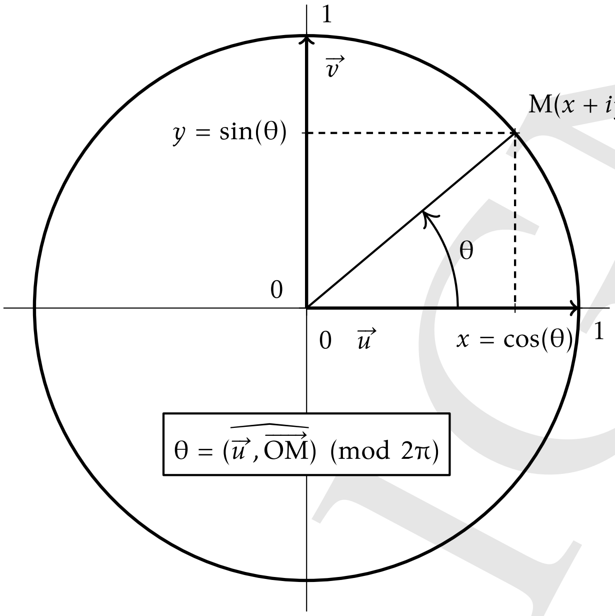
Tout complexe de module 1 peut se mettre sous la forme $\cos(\theta) + i\sin(\theta)$ avec θ un réel, plus précisément :

$$\mathbb{U} = \{\cos(\theta) + i\sin(\theta) \ / \ \theta \in \mathbb{R}\}.$$

Preuve : Soit $z = x + iy$ un complexe de module 1, on a $x^2 + y^2 = 1$, donc il existe un réel θ (unique à 2π près) tel que $x = \cos(\theta)$ et $y = \sin(\theta)$, c'est à dire $z = \cos(\theta) + i\sin(\theta)$. La réciproque est immédiate. □

2) Exponentielle d'un imaginaire pur

Pour tout **réel** x, on pose $e^{ix} = \cos(x) + i\sin(x)$. On a les propriétés suivantes :
- $\forall \theta \in \mathbb{R},\ \mathrm{Re}(e^{i\theta}) = \cos(\theta)$ et $\mathrm{Im}(e^{i\theta}) = \sin(\theta)$.
- $\forall \theta \in \mathbb{R},\ e^{-i\theta} = \cos(-\theta) + i\sin(-\theta) = \cos(\theta) - i\sin(\theta) = \overline{e^{i\theta}}$.
- $\forall \theta \in \mathbb{R},\ |e^{i\theta}| = \sqrt{\cos(\theta)^2 + \sin(\theta)^2} = 1$, donc $e^{i\theta} \in \mathbb{U}$.
- $\forall z \in \mathbb{U}, \exists \theta \in \mathbb{R},\ e^{i\theta} = z$.



- Soit $\theta, \theta' \in \mathbb{R}, e^{i\theta} = e^{i\theta'} \iff \begin{cases} \cos(\theta) & = \cos(\theta') \\ \sin(\theta) & = \sin(\theta') \end{cases} \iff \theta = \theta'\ (2\pi)$.

Théorème 3.7

$\forall \theta, \theta' \in \mathbb{R},\ e^{i\theta} \times e^{i\theta'} = e^{i(\theta+\theta')}$.

Preuve : $e^{i\theta}e^{i\theta'} = [\cos(\theta)+i\sin(\theta)][\cos(\theta')+i\sin(\theta')] = [\cos(\theta)\cos(\theta')-\sin(\theta)\sin(\theta')]+i[\cos(\theta)\sin(\theta')+\sin(\theta)\cos(\theta')] = \cos(\theta+\theta') + i\sin(\theta+\theta') = e^{i(\theta+\theta')}$ d'après les formules d'addition du sinus et du cosinus. □

Ce théorème permet de retrouver les formules trigonométriques.

À retenir

Pour tous réels x et y :
- $\cos(x+y) = \mathrm{Re}(e^{i(x+y)}) = \mathrm{Re}(e^{ix}e^{iy}) = \cos(x)\cos(y) - \sin(x)\sin(y)$.
- $\sin(x+y) = \mathrm{Im}(e^{i(x+y)}) = \mathrm{Im}(e^{ix}e^{iy}) = \cos(x)\sin(y) + \sin(x)\cos(y)$.
- On en déduit : $\cos(2x) = 2\cos^2(x) - 1 = 1 - 2\sin^2(x) =$ et $\sin(2x) = 2\sin(x)\cos(x)$.

$$\text{En posant } a = \frac{x+y}{2} \text{ et } b = \frac{x-y}{2} \text{ on obtient :}$$

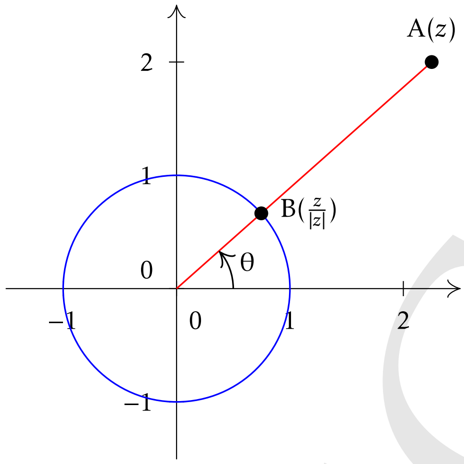
- $\cos(x) + \cos(y) = \cos(a+b) + \cos(a-b) = 2\cos(a)\cos(b)$, donc
$$\cos(x) + \cos(y) = 2\cos(\tfrac{x+y}{2})\cos(\tfrac{x-y}{2}).$$
- $\cos(x) - \cos(y) = \cos(a+b) - \cos(a-b) = -2\sin(a)\sin(b)$, donc
$$\cos(x) - \cos(y) = -2\sin(\tfrac{x+y}{2})\sin(\tfrac{x-y}{2}).$$
- $\sin(x) + \sin(y) = \sin(a+b) + \sin(a-b) = 2\sin(a)\cos(b)$, donc
$$\sin(x) + \sin(y) = 2\sin(\tfrac{x+y}{2})\cos(\tfrac{x-y}{2})$$
- etc

3) Argument d'un nombre complexe

Soit $z \in \mathbb{U}$, on sait qu'il existe un réel θ (unique à 2π près) tel que $z = e^{i\theta}$. Si maintenant z est un complexe non nul quelconque alors $\frac{z}{|z|} \in \mathbb{U}$ et donc il existe un réel θ (unique à 2π près) tel que $\frac{z}{|z|} = e^{i\theta}$, c'est à dire $z = |z|e^{i\theta}$.

Définition 3.5

Soit z un complexe non nul, on appelle **argument** *de z tout réel θ tel que $z = |z|e^{i\theta}$, cette égalité est appelée* **forme trigonométrique** *de z. L'ensemble des arguments de z est noté $\arg(z)$, on a donc $\arg(z)=\{\theta \in \mathbb{R} / \ z = |z|e^{i\theta}\}$, et si θ_0 est un argument de z, alors $\arg(z)=\{\theta_0 + 2k\pi / \ k \in \mathbb{Z}\}$.*



Définition 3.6

Soit $z \in \mathbb{C}^*$, z possède un unique argument dans l'intervalle $]-\pi;\pi]$, par définition cet argument est appelé **argument principal** *de z et noté $\mathrm{Arg}(z)$.*

☞**Exemples :**

- $\mathrm{Arg}(i) = \frac{\pi}{2}$, $\mathrm{Arg}(j) = \frac{2\pi}{3}$.
- si $x \in \mathbb{R}^{*+}$ alors $\mathrm{Arg}(x) = 0$ et si $x \in \mathbb{R}^{*-}$ alors $\mathrm{Arg}(x) = \pi$.

Attention !

Si $z = r\,e^{i\theta}$ avec r et θ réels, alors c'est la forme trigonométrique de z lorsque $r > 0$. Mais lorsque $r < 0$, la forme trigonométrique de z est $z = -re^{i(\theta+\pi)}$.

À retenir : factorisation par l'arc moitié

- Si $z = e^{ix} + e^{iy}$, alors $z = e^{i\frac{x+y}{2}}[e^{i\frac{x-y}{2}} + e^{-i\frac{x-y}{2}}] = 2\cos(\frac{x-y}{2})e^{i\frac{x+y}{2}}$. D'où $|z| = 2|\cos(\frac{x-y}{2})|$ et $\mathrm{Arg}(z) = \frac{x+y}{2}\ (\pi)$.
- Si $z = e^{ix} - e^{iy}$, alors $z = e^{i\frac{x+y}{2}}[e^{i\frac{x-y}{2}} - e^{-i\frac{x-y}{2}}] = 2i\sin(\frac{x-y}{2})e^{i\frac{x+y}{2}}$. D'où $|z| = 2|\sin(\frac{x-y}{2})|$ et $\mathrm{Arg}(z) = \frac{x+y+\pi}{2}\ (\pi)$.

Théorème 3.8 (propriétés)

Soient $z, z' \in \mathbb{C}^*$ avec $\theta = \mathrm{Arg}(z)$ et $\theta' = \mathrm{Arg}(z')$:

- $z = z' \iff \left\{\begin{array}{lcl} |z| & = & |z'| \\ \theta & = & \theta'\ (2\pi) \end{array}\right.$.
- $z \in \mathbb{R}^* \iff \theta = 0\ (\pi)$.
- $\bar{z} = |z|e^{-i\theta}$ *donc* $\mathrm{Arg}(\bar{z}) = -\theta\ (2\pi)$.
- $-z = |z|e^{i(\theta+\pi)}$ *donc* $\mathrm{Arg}(-z) = \theta + \pi\ (2\pi)$.
- $zz' = |zz'|e^{i(\theta+\theta')}$ *donc* $\mathrm{Arg}(zz') = \theta + \theta'\ (2\pi)$.
- $\frac{z}{z'} = \frac{|z|}{|z'|}e^{i(\theta-\theta')}$ *donc* $\mathrm{Arg}(\frac{z}{z'}) = \theta - \theta'\ (2\pi)$.
- $\forall n \in \mathbb{Z}, z^n = |z^n|e^{in\theta}$ *donc* $\mathrm{Arg}(z^n) = n\theta\ (2\pi)$.

Preuve : Celle-ci est simple et laissée en exercice. □

À retenir

Soient a,b deux réels non tous deux nuls et soit $x \in \mathbb{R}$, en posant $z = a + ib = |z|e^{i\theta}$ on obtient :

$$a\cos(x) + b\sin(x) = \mathrm{Re}(\bar{z}e^{ix}) = |z|\cos(x-\theta) = \sqrt{a^2+b^2}\cos(x-\theta).$$

4) Racines n^{es} d'un nombre complexe

Définition 3.7

*Soit a, z deux complexes et $n \in \mathbb{N}$, z est une **racine n^e** de a lorsque $z^n = a$.*

Cas particuliers des racines n^{es} de l'unité

Théorème 3.9

Soit n un entier supérieur ou égal à deux, l'équation $z^n = 1$ possède exactement n solutions qui sont les complexes $e^{2ik\pi/n}$ avec $0 \leqslant k \leqslant n-1$. On note $\mathbb{U}_n$ l'ensemble de ces solutions (appelées racines n^{es} de l'unité), on a donc :

$$\mathbb{U}_n = \left\{z \in \mathbb{U} \ / \ z^n = 1\right\} = \left\{e^{2ik\pi/n} \ / \ 0 \leqslant k \leqslant n-1\right\}$$

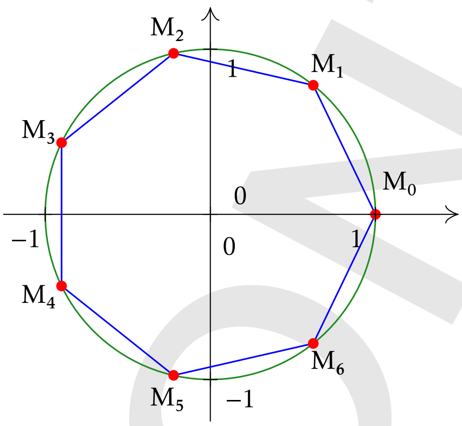
Preuve : Résolution trigonométrique : on pose $z = re^{i\theta}$ avec r réel positif et $\theta \in [0\,;2\pi[$, alors $z^n = r^n e^{in\theta}$, donc $z^n = 1$ si et seulement si $r^n = 1$ (égalité des modules) et $n\theta = 0 + 2k\pi$ (égalité des arguments avec k entier), ce qui équivaut encore à $r = 1$ (r étant réel positif) et $\theta = \frac{2k\pi}{n}$. Le choix de l'intervalle $[0\,;2\pi[$ pour l'argument θ entraîne que l'entier k est dans l'intervalle $[\![0\,;n-1]\!]$, donc $z^n = 1 \iff \exists k \in [\![0\,;n-1]\!], z = e^{2ik\pi/n}$, cela fait exactement n solutions distinctes (car tout complexe non nul possède un unique argument dans l'intervalle $[0\,;2\pi[$). □

★**Exercice 3.4**

1/ Décrire $\mathbb{U}_2$, $\mathbb{U}_3$, $\mathbb{U}_4$.

2/ Montrer que $(\mathbb{U}_n, \times)$ est un groupe commutatif (résultat important à connaître).

Représentation géométrique des racines n^{es} de l'unité (avec $n = 7$) :



M_k est le point d'affixe $e^{2ik\pi/n}$ ($n = 7$).

Cas général : résolution de l'équation $z^n = a$ ($a \neq 0$)

Considérons la forme trigonométrique de a : $a = |a|e^{i\theta_0}$, il est facile de voir alors que le complexe $z_0 = \sqrt[n]{|a|}e^{i\theta_0/n}$ est une solution particulière de l'équation $z_0^n = a$, par conséquent, comme $z_0 \neq 0$, on a :

$$z^n = a \iff z^n = z_0^n \iff \left(\frac{z}{z_0}\right)^n = 1 \iff \frac{z}{z_0} \in \mathbb{U}_n$$

On est ainsi ramené aux racines n^{es} de l'unité, on en déduit que $z = z_0 e^{i2k\pi/n}$ avec $0 \leqslant k \leqslant n-1$, d'où :

$$\boxed{z^n = a \iff z = \sqrt[n]{|a|}e^{i\frac{\theta_0+2k\pi}{n}} \quad \text{avec} \quad 0 \leqslant k \leqslant n-1 \quad \text{et} \quad \theta_0 = \mathrm{Arg}(a) \pmod{2\pi}}$$

Tout complexe non nul a donc exactement n racines n^{es} distinctes.

IV EXPONENTIELLE COMPLEXE

1) Définition

Définition 3.8

Soit $z = x + iy$ un nombre complexe, on appelle ***exponentielle*** *de z le complexe noté $\exp(z)$ et défini par : $\exp(z) = e^x[\cos(y) + i\sin(y)]$, c'est à dire $\exp(z) = e^x e^{iy}$ (forme trigonométrique).*

Remarque 3.2 :

- *Si z est réel (ie $y = 0$), alors l'exponentielle de z correspond à l'exponentielle* ***réelle*** *de z. De même, si z est imaginaire pur ($x = 0$), alors $\exp(z) = \exp(iy) = \cos(y) + i\sin(y)$.*
- $\exp(0) = 1$.
- $\mathrm{Re}(\exp(z)) = e^{\mathrm{Re}(z)}\cos(\mathrm{Im}(z))$ *et* $\mathrm{Im}(\exp(z)) = e^{\mathrm{Re}(z)}\sin(\mathrm{Im}(z))$.
- $|\exp(z)| = e^{\mathrm{Re}(z)}$ *et* $\mathrm{Arg}(\exp(z)) = \mathrm{Im}(z)$ (2π).
- $\overline{\exp(z)} = \exp(\overline{z})$.

Théorème 3.10 (propriété fondamentale)

La fonction $\exp : \mathbb{C} \to \mathbb{C}^*$ est $2i\pi$-périodique, vérifie :

$$\forall z, z' \in \mathbb{C},\ \exp(z + z') = \exp(z) \times \exp(z').$$

Et tout pour complexe non nul a, il existe $z \in \mathbb{C}$ tel que $a = \exp(z)$.

Preuve : Il est clair d'après la définition que $\exp(z)$ ne peut pas être nul, donc $\exp(z) \in \mathbb{C}^*$. Posons $z = x + iy$, $\exp(z + 2i\pi) = e^x[\cos(y + 2\pi) + i\sin(y + 2\pi)] = \exp(z)$. Soit a un complexe non nul, l'équation $\exp(z) = a$ équivaut à $|a| = e^x$ et $\mathrm{Arg}(a) = y \pmod{2\pi}$ (résolution trigonométrique), donc les complexes $z = \ln(|a|) + i(y + 2k\pi)$ (où k parcourt $\mathbb{Z}$) sont les antécédents de a, en particulier les solutions de l'équation $\exp(z) = 1$ sont les complexes $z = 2ik\pi, k \in \mathbb{Z}$. Soit $z' = x' + iy'$ un autre complexe, $\exp(z + z') = e^{x+x'}[\cos(y + y') + i\sin(y + y')]$, et $\exp(z)\exp(z') = e^{x+x'}[\cos(y)\cos(y') - \sin(y)\sin(y')] = e^{x+x'}[\cos(y + y') + i\sin(y + y')]$. On peut déduire de cette propriété le calcul suivant :

$$\begin{aligned}
\exp(z) = \exp(z') &\iff \frac{\exp(z)}{\exp(z')} = 1\\
&\iff \exp(z)\exp(-z') = 1\\
&\iff \exp(z - z') = 1\\
&\iff \exists k \in \mathbb{Z},\ z = z' + 2ik\pi.
\end{aligned}$$

□

Remarque 3.3 – *La propriété fondamentale de l'exponentielle complexe : $\exp(z + z') = \exp(z)\exp(z')$, est la même que celle de l'exponentielle réelle. On convient alors de noter $\exp(z) = e^z$. La propriété fondamentale devient :*

$$\boxed{e^{z+z'} = e^z \times e^{z'}}$$

Il découle également de cette relation que : $\boxed{e^{-z} = \dfrac{1}{e^z}}$.

★**Exercice 3.5** *Résoudre $e^z = 1 + i$.*

2) Formules d'Euler et de Moivre

Formule de *Moivre*[1] : $\forall n \in \mathbb{Z}$, $\forall z \in \mathbb{C}$, $e^{nz} = [e^z]^n$. En particulier pour $z = ix$ avec x réel, on a $e^{inx} = [\cos(x) + i\sin(x)]^n$. On en déduit que :

$$\cos(nx) = \mathrm{Re}([\cos(x) + i\sin(x)]^n) \text{ et } \sin(nx) = \mathrm{Im}([\cos(x) + i\sin(x)]^n).$$

À l'aide du binôme de *Newton* ces formules permettent d'exprimer $\cos(nx)$ et $\sin(nx)$ sous forme d'un polynôme en $\cos(x)$ et $\sin(x)$.

☞**Exemples :**

1. *MOIVRE Abraham* DE (1667 – 1754) : mathématicien français, il s'expatria à Londres à l'age de dix-huit ans.

– $\cos(4x) = \text{Re}([\cos(x)+i\sin(x)]^4) = \cos(x)^4 - 6\cos(x)^2\sin(x)^2 + \sin(x)^4$. En remplaçant $\sin(x)^2$ par $1-\cos(x)^2$, on pourrait obtenir $\cos(4x)$ en fonction de $\cos(x)$ uniquement.
– $\sin(4x) = \text{Im}([\cos(x)+i\sin(x)]^4) = 4\cos(x)^3\sin(x) - 4\cos(x)\sin(x)^3$.

Formules d'*Euler* [2] : $\forall x \in \mathbb{R} : \cos(x) = \frac{e^{ix}+e^{-ix}}{2}$ et $\sin(x) = \frac{e^{ix}-e^{-ix}}{2i}$.
Ces formules permettent la **linéarisation** de $\cos(x)^p\sin(x)^q$.

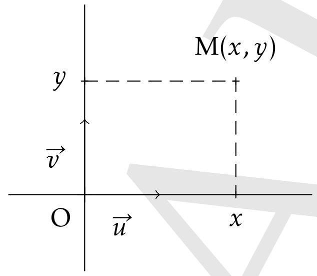
☞**Exemples :**
– $\cos(x)^3 = \frac{(e^{ix}+e^{-ix})^3}{8} = \frac{e^{i3x}+3e^{i2x}e^{-ix}+3e^{ix}e^{-i2x}+e^{-i3x}}{8} = \frac{\cos(3x)+3\cos(x)}{4}$.
– $\sin(x)^3 = \frac{(e^{ix}-e^{-ix})^3}{-8i} = \frac{e^{i3x}-3e^{i2x}e^{-ix}+3e^{ix}e^{-i2x}-e^{-i3x}}{-8i} = \frac{3\sin(x)-\sin(3x)}{4}$.

V REPRÉSENTATION GÉOMÉTRIQUE DES COMPLEXES, APPLICATIONS

Le plan complexe est un plan $\mathcal{P}$ muni d'un repère orthonormé direct que l'on note $\mathcal{R} = (\text{O}, \vec{u}, \vec{v})$.

1) Affixe

Chaque point M du plan complexe est repéré par ses coordonnées : une abscisse x et une ordonnée y, c'est à dire par le couple de réels (x, y). Autant dire que M est repéré par le **complexe** $z = x + iy$. Par définition, ce complexe est **l'affixe** du point M.



Réciproquement, tout complexe z est l'affixe d'un point M du plan que l'on appelle **image** de z. Les axes $(\text{O}, \vec{u})$ et $(\text{O}, \vec{v})$ sont appelés respectivement **axes des réels** et **axe des imaginaires**.

Par exemple, l'image de $\overline{z}$ est le symétrique de l'image de z par la réflexion d'axe $(\text{O}, \vec{u})$.

De la même façon, chaque vecteur du plan a des coordonnées dans la base $(\vec{u}, \vec{v})$. Si $\vec{w}$ a pour coordonnées (x, y), cela signifie que $\vec{w} = x\vec{u} + y\vec{v}$, là encore le vecteur $\vec{w}$ peut être représenté par le complexe $x + iy$, ce complexe est appelé **affixe** du vecteur $\vec{w}$. Réciproquement, tout complexe z est l'affixe d'un vecteur du plan. On remarquera que l'affixe d'un point M n'est autre que l'affixe du **vecteur** $\overrightarrow{\text{OM}}$.

$\ast$) L'affixe de la somme de deux vecteurs est la somme des affixes. Si $\alpha \in \mathbb{R}$ et si $\vec{w}$ est le vecteur d'affixe z, alors l'affixe du vecteur $\alpha\vec{w}$ est αz.

$\ast$) Soit M d'affixe z et M′ d'affixe z', l'affixe du vecteur $\overrightarrow{\text{MM}'}$ est $z' - z$.

2) Distances

Le module d'un complexe z représente dans le plan complexe la distance de l'origine O au point M d'affixe z, c'est à dire $|z| = \text{OM} = \|\overrightarrow{\text{OM}}\|$.

Si $\vec{w}$ est un vecteur d'affixe z, alors la norme de $\vec{w}$ est $\|\vec{w}\| = |z|$.

Soit M d'affixe z et M′ d'affixe z', la distance de M à M′ est $\text{MM}' = \|\overrightarrow{\text{MM}'}\| = |z' - z|$.

Définition 3.9

Soit $a \in \mathbb{C}$ et $\text{R} > 0$, on définit dans le plan complexe :
– *le disque fermé de centre a et de rayon R :* $\{\text{M} \in \mathcal{P} \ / \ |z - a| \leqslant \text{R}\}$.
– *le disque ouvert de centre a et de rayon R :* $\{\text{M} \in \mathcal{P} \ / \ |z - a| < \text{R}\}$.
– *le cercle de centre a et de rayon R :* $\{\text{M} \in \mathcal{P} \ / \ |z - a| = \text{R}\}$.

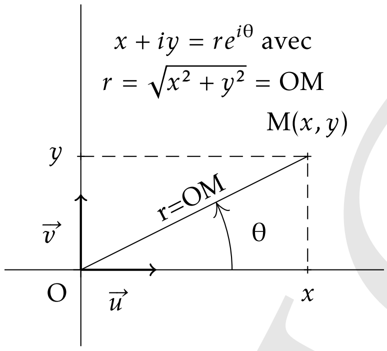
☞**Exemples :**

2. *EULER Léonhard* (1707 – 1783) : grand mathématicien suisse.

– La représentation géométrique du groupe unité $\mathbb{U} = \{z \in \mathbb{C} \ / \ |z| = 1\}$ est le cercle de centre O et de rayon 1 : le cercle **trigonométrique**.
– Les points d'affixe les racines n^es de l'unité ($n \geqslant 2$) sont les sommets d'un polygone régulier inscrit dans le cercle unité. La longueur du coté est $2\sin(\frac{\pi}{n})$, et la longueur du centre au milieu d'un coté (l'apothème) est $\cos(\frac{\pi}{n})$.

3) Angles orientés

Soit z un complexe non nul et M le point du plan d'affixe z, l'argument principal de z est une mesure de l'angle orienté $(\vec{u}, \overrightarrow{\mathrm{OM}})$, ce que l'on écrit $(\vec{u}, \overrightarrow{\mathrm{OM}}) = \mathrm{Arg}(z) \ (2\pi)$.



Soient $\vec{w}$ et $\vec{w'}$ deux vecteurs non nuls d'affixes respectifs z et z'. Désignons par M et M′ les points d'affixes respectifs z et z', l'angle orienté entre les deux vecteurs $\vec{w}$ et $\vec{w'}$ est :

$$\begin{aligned}(\vec{w}, \vec{w'}) &= (\overrightarrow{\mathrm{OM}}, \overrightarrow{\mathrm{OM'}}) \\ &= (\overrightarrow{\mathrm{OM}}, \vec{u}) + (\vec{u}, \overrightarrow{\mathrm{OM'}}) \\ &= -(\vec{u}, \overrightarrow{\mathrm{OM}}) + (\vec{u}, \overrightarrow{\mathrm{OM'}}) \\ &= -\mathrm{Arg}(z) + \mathrm{Arg}(z') \ (2\pi) \\ &= \mathrm{Arg}(\frac{z'}{z}) \ (2\pi)\end{aligned}$$

Conséquence : Soient A, B et C trois points distincts d'affixes respectifs a, b et c. L'affixe du vecteur $\overrightarrow{\mathrm{AB}}$ est $b - a$ et celui du vecteur $\overrightarrow{\mathrm{AC}}$ est $c - a$, par conséquent l'angle $(\overrightarrow{\mathrm{AB}}, \overrightarrow{\mathrm{AC}})$ est donné par :

$$\boxed{(\overrightarrow{\mathrm{AB}}, \overrightarrow{\mathrm{AC}}) = \mathrm{Arg}(\frac{c-a}{b-a}) \ (2\pi).}$$

Et on a le rapport des distances $\frac{\mathrm{AC}}{\mathrm{AB}} = \left|\frac{c-a}{b-a}\right|$.

Les trois points A, B et C sont alignés si et seulement si $(\overrightarrow{\mathrm{AB}}, \overrightarrow{\mathrm{AC}}) = 0 \pmod{\pi}$ ce qui équivaut donc à $\frac{c-a}{b-a} \in \mathbb{R}$.

Les vecteurs $\overrightarrow{\mathrm{AB}}$ et $\overrightarrow{\mathrm{AC}}$ sont orthogonaux si et seulement si $(\overrightarrow{\mathrm{AB}}, \overrightarrow{\mathrm{AC}}) = \frac{\pi}{2} \pmod{\pi}$ ce qui équivaut donc à $\frac{c-a}{b-a} \in i\mathbb{R}$.

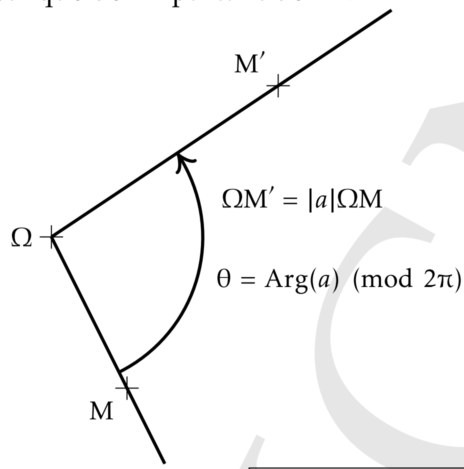
4) Similitudes directes

Soient $a, b \in \mathbb{C}$ avec $a \neq 0$, on note $\mathrm{S}_{a,b} : \mathcal{P} \to \mathcal{P}$ l'application qui au point d'affixe z faire correspondre le point d'affixe $z' = az + b$, c'est la similitude directe de paramètres a et b. On vérifie en appliquant la définition, les propriétés suivantes :

– $\mathrm{S}_{a,b}$ est une bijection et sa réciproque est aussi une similitude directe, plus précisément $\mathrm{S}_{a,b}^{-1} = \mathrm{S}_{\frac{1}{a}, \frac{-b}{a}}$.
– La composée de deux similitudes directes est une similitude directe, plus précisément, $\mathrm{S}_{c,d} \circ \mathrm{S}_{a,b} = \mathrm{S}_{ac, cb+d}$ avec $c \neq 0$.
– Soient $\mathrm{A}(z_\mathrm{A})$, $\mathrm{B}(z_\mathrm{B})$, $\mathrm{C}(z_\mathrm{C})$ et $\mathrm{D}(z_\mathrm{D})$ quatre points avec $\mathrm{A} \neq \mathrm{B}$ et $\mathrm{C} \neq \mathrm{D}$, alors il existe une unique similitude directe $\mathrm{S}_{a,b}$ qui transforme A en C et B en D, celle-ci se détermine en résolvant le système $\begin{cases} az_\mathrm{A} + b = z_\mathrm{C} \\ az_\mathrm{B} + b = z_\mathrm{D} \end{cases}$ pour déterminer a et b.
– Lorsque $a = 1$, la similitude $\mathrm{S}_{1,b}$ est la translation de vecteur $\vec{u}$ d'affixe b.

Étude lorsque $a \neq 1$

- Il existe un unique point fixe $\Omega(z_0)$ avec $z_0 = \frac{b}{1-a}$, on a alors pour tout z : $az + b = a(z - z_0) + z_0$. Ce point est appelé **centre** de la similitude.
- Soit $M(z)$ un point différent du centre Ω, et soit $M'(z')$ son image par $S_{a,b}$, alors : $\frac{z'-z_0}{z-z_0} = a$, on en déduit que : $\left|\frac{z'-z_0}{z-z_0}\right| = |a|$, c'est à dire $\Omega M' = |a|\Omega M$, et $(\overrightarrow{\Omega M}, \overrightarrow{\Omega M'}) = \mathrm{Arg}(a) \pmod{2\pi}$, ce qui permet une construction géométrique de M' partant de M :



- Cas particulier où $\mathrm{Arg}(a) = 0 \pmod{2\pi}$: on a alors $\boxed{z' = |a|(z - z_0) + z_0}$ d'où $\overrightarrow{\Omega M'} = |a|\overrightarrow{\Omega M}$, on dit que la similitude est **l'homothétie de centre** Ω **et de rapport** $|a|$, on la note $h_{(\Omega,|a|)}$.
- Cas particulier où $\mathrm{Arg}(a) = \pi \pmod{2\pi}$: on a alors $\boxed{z' = -|a|(z - z_0) + z_0}$ d'où $\overrightarrow{\Omega M'} = |a|\overrightarrow{\Omega M}$, on dit que la similitude est **l'homothétie de centre** Ω **et de rapport** $-|a|$, on la note $h_{(\Omega,-|a|)}$.
- Cas particulier où $|a| = 1$: on a alors $\boxed{z' = e^{i\theta}(z - z_0) + z_0}$ d'où $\Omega M' = \Omega M$ et $(\overrightarrow{\Omega M}, \overrightarrow{\Omega M'}) = \mathrm{Arg}(a) \pmod{2\pi}$, on dit que la similitude est la **rotation de centre** Ω **et d'angle** $\mathrm{Arg}(a)$, on la note $r_{(\Omega,\mathrm{Arg}(a))}$.
- Dans le cas général, on vérifie par le calcul que $S_{a,b}$ et la composée commutative :

$$S_{a,b} = h_{(\Omega,|a|)} \circ r_{(\Omega,\mathrm{Arg}(a))} = r_{(\Omega,\mathrm{Arg}(a))} \circ h_{(\Omega,|a|)}$$

On dit que $\boxed{S_{a,b} \text{ est la similitude de centre } \Omega\text{, de rapport } |a| \text{ et d'angle } \mathrm{Arg}(a)}$. On remarquera que la bijection réciproque est la similitude de centre Ω, de rapport $\frac{1}{|a|}$ et d'angle $-\mathrm{Arg}(a) \pmod{2\pi}$.

VI SOLUTION DES EXERCICES

Solution 3.1

1/ *Soit z un complexe, alors $z = z \times 1 = z \times (1+0) = z \times 1 + z \times 0 = z + z \times 0$, ce qui entraîne que $0 = z \times 0$. La démonstration est valable dans tout corps.*

2/ *Soient z et z' deux complexes tels que $zz' = 0$, supposons z non nul, alors z a un inverse, d'où $z^{-1}(zz') = z^{-1} \times 0 = 0$, ce qui donne $z' = 0$ par associativité.*

Solution 3.2

1/ a) *Si deux complexes sont égaux, alors ils ont la même partie réelle et la même partie imaginaire. Réciproquement, si les parties réelle et imaginaire sont identiques, alors les deux complexes ont la même écriture algébrique et sont donc égaux.*

b) *Si la partie imaginaire est nulle alors le complexe est réel car égal à sa partie réelle. Si le complexe est réel, alors l'unicité des parties réelle et imaginaire entraîne que cette dernière est nulle.*

c) *Même type d'argument que la question précédente.*

d) *$z + z' = \mathrm{Re}(z) + \mathrm{Re}(z') + i[\mathrm{Im}(z) + \mathrm{Im}(z')]$, c'est la forme algébrique de $z + z'$, donc par unicité, $\mathrm{Re}(z+z') = \mathrm{Re}(z) + \mathrm{Re}(z')$ et $\mathrm{Im}(z+z') = \mathrm{Im}(z) + \mathrm{Im}(z')$.*

e) *$\alpha z = \alpha\,\mathrm{Re}(z) + i\alpha\,\mathrm{Im}(z)$, c'est la forme algébrique de αz, donc par unicité, $\mathrm{Re}(\alpha z) = \alpha\,\mathrm{Re}(z)$ et $\mathrm{Im}(\alpha z) = \alpha\,\mathrm{Im}(z)$.*

2/ *Notons $z = a + ib$ et $z' = a' + ib'$ les deux formes algébriques, alors après développement, on a $zz' = aa' - bb' + i(ab' + ba')$, c'est la forme algébrique de zz', donc $\mathrm{Re}(zz') = aa' - bb'$ et $\mathrm{Im}(zz') = ab' + ba'$.*

Solution 3.3 *On effectue le produit* $\bar{z} \times \overline{\frac{1}{z}} = \overline{z \times \frac{1}{z}} = \bar{1} = 1$*, d'où* $\overline{\frac{1}{z}} = \frac{1}{\bar{z}}$.

Solution 3.4

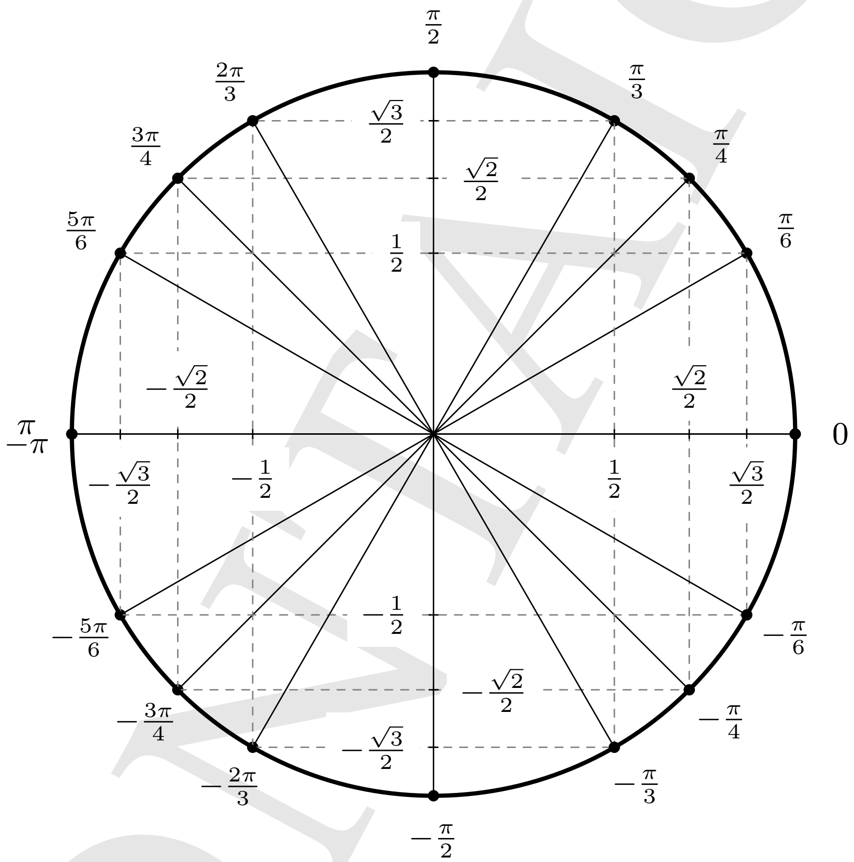
1/ $\mathbb{U}_2 = \{-1; 1\}$, $\mathbb{U}_3 = \left\{1; j; j^2\right\}$, $\mathbb{U}_4 = \{1; i; -1; -i\}$.

2/ *Si* z *et* z' *sont dans* $\mathbb{U}_n$ *alors* $(zz')^n = z^n \times z'^n = 1 \times 1 = 1$*, donc* zz' *est dans* $\mathbb{U}_n$ *: la multiplication et donc interne dans* $\mathbb{U}_n$*. Elle est associative et commutative dans* $\mathbb{C}$*, donc en particulier dans* $\mathbb{U}_n$*. L'élément neutre de la multiplication est dans* $\mathbb{U}_n$ *car* $1^n = 1$*. Si* z *est dans* $\mathbb{U}_n$ *alors* z *est non nul (sinon* $z^n = 0 \neq 1$*), et* $\left(\frac{1}{z}\right)^n = \frac{1}{z^n} = \frac{1}{1} = 1$*, donc* $\frac{1}{z}$ *est dans* $\mathbb{U}_n$ *: tout élément de* $\mathbb{U}_n$ *a un symétrique dans* $\mathbb{U}_n$ *pour la multiplication. En conclusion* $(\mathbb{U}_n, \times)$ *est un groupe commutatif.*

Solution 3.5 *On écrit* $z = x + iy$ *avec* x, y *réels, on a alors* $z = e^x e^{iy}$ *et* $1 + i = \sqrt{2}e^{i\pi/4}$*. Résolution trigonométrique : l'équation équivaut alors à* $e^x = \sqrt{2}$ *(égalité des modules) et* $y = \frac{\pi}{4} + 2k\pi$ *avec* $k \in \mathbb{Z}$ *(égalité des arguments à* 2π *près), l'ensemble des solutions est donc* $\left\{\frac{1}{2}\ln(2) + i(\frac{\pi}{4} + 2k\pi) \;/\; k \in \mathbb{Z}\right\}$.

VII FORMULAIRE DE TRIGONOMÉTRIE

Les valeurs angulaires écrites autour du cercle peuvent être modifiées en leur ajoutant un multiple de 2π.



- Formules découlant du cercle trigonométrique, pour tous réels x et y :
  - $\cos(x + 2\pi) = \cos(x)$, $\sin(x + 2\pi) = \sin(x)$ (2π est une période).
  - $\cos(x \pm \pi) = -\cos(x)$, $\sin(x \pm \pi) = -\sin(x)$ (π est une anti-période).
  - $\cos(-x) = \cos(x)$, $\sin(-x) = -\sin(x)$ (parité).
- Formules fondamentales :

À retenir

- $\cos^2(x) + \sin^2(x) = 1$;
- $\cos(x + y) = \cos(x)\cos(y) - \sin(x)\sin(y)$ (formule d'addition).
- $\sin(x + y) = \cos(x)\sin(y) + \sin(x)\cos(y)$ (formule d'addition).

- $\cos(2x) = \cos^2(x) - \sin^2(x) = 2\cos^2(x) - 1 = 1 - 2\sin^2(x)$ (formule de duplication).
- $\sin(2x) = 2\sin(x)\cos(x)$ (formule de duplication).

– Formules qui en découlent :
- $\cos(x + \frac{\pi}{2}) = -\sin(x)$ (dérivée), $\sin(x + \frac{\pi}{2}) = \cos(x)$ (dérivée).
- $\cos(x - \frac{\pi}{2}) = \sin(x)$, $\sin(x - \frac{\pi}{2}) = -\cos(x)$.
- En posant $a = \frac{x+y}{2}$ et $b = \frac{x-y}{2}$ on obtient :
  * $\cos(x) + \cos(y) = \cos(a+b) + \cos(a-b) = 2\cos(a)\cos(b)$, donc
  $$\cos(x) + \cos(y) = 2\cos(\tfrac{x+y}{2})\cos(\tfrac{x-y}{2}).$$
  * $\cos(x) - \cos(y) = \cos(a+b) - \cos(a-b) = -2\sin(a)\sin(b)$, donc
  $$\cos(x) - \cos(y) = -2\sin(\tfrac{x+y}{2})\sin(\tfrac{x-y}{2}).$$
  * $\sin(x) + \sin(y) = \sin(a+b) + \sin(a-b) = 2\sin(a)\cos(b)$, donc
  $$\sin(x) + \sin(y) = 2\sin(\tfrac{x+y}{2})\cos(\tfrac{x-y}{2})$$

– Lien avec les complexes : $\boxed{\cos(x) = \mathrm{Re}(e^{ix}) \text{ et } \sin(x) = \mathrm{Im}(e^{ix})}$.

Chapitre 4

Calculs algébriques

Sommaire

I SOMMES ET PRODUITS

On se place dans $\mathbb{C}$. Les sommes et les produits dont on va parler ne contiennent qu'un nombre fini de termes.

1) Définition

Définition 4.1 (somme et produit sur une partie finie)
Soit A une partie finie de $\mathbb{C}$, la somme des éléments de A est notée $\sum\limits_{a\in A} a$ et le produit des éléments de A est noté $\prod\limits_{a\in A} a$. Par convention, lorsque A est vide, la somme est nulle et le produit vaut 1.

Remarque 4.1 – *Ces opérations étant commutatives dans $\mathbb{C}$, l'ordre n'a pas d'importance. Comme elles sont également associatives, il est inutile de préciser un parenthésage pour la somme ou le produit.*

☞**Exemple** : La somme et le produit des racines n^{es} de l'unité se notent $\sum\limits_{\omega\in\mathbb{U}_n} \omega$ et $\prod\limits_{\omega\in\mathbb{U}_n} \omega$.

Définition 4.2 (somme et produit d'une famille finie)
Soit $(a_k)_{k\in I}$ une famille de complexes indexée par un ensemble fini I. La somme des éléments de la famille est notée $\sum\limits_{k\in I} a_k$ et le produit des éléments de la famille est noté $\prod\limits_{k\in I} a_k$. Par convention, lorsque I est vide, la somme est nulle et le produit vaut 1.

☞**Exemple** : La famille des racines n^{es} de l'unité est $(e^{2ik\pi/n})_{k\in[\![0;n-1]\!]}$, on peut donc écrire la somme et le produit ainsi : $\sum\limits_{k\in[\![0;n-1]\!]} e^{2ik\pi/n}$ et $\prod\limits_{k\in[\![0;n-1]\!]} e^{2ik\pi/n}$.

Définition 4.3 (lorsque I est un intervalle d'entiers)

Soit $(a_k)_{k\in \mathrm{I}}$ une famille de complexes indexée par un intervalle d'entiers $\mathrm{I} = [\![n\,;m]\!]$. La somme des éléments de la famille est notée $\sum\limits_{k=n}^{m} a_k$ et le produit des éléments de la famille est noté $\prod\limits_{k=n}^{m} a_k$. Par convention, lorsque $n > m$, la somme est nulle et le produit vaut 1 (car dans ce cas l'intervalle $[\![n\,;m]\!]$ est vide).

☞**Exemple** : La famille des racines n^{es} de l'unité est $(e^{2ik\pi/n})_{k\in[\![0;n-1]\!]}$, on peut donc écrire la somme et le produit ainsi : $\sum\limits_{k=0}^{n-1} e^{2ik\pi/n}$ et $\prod\limits_{k=0}^{n-1} e^{2ik\pi/n}$.

Attention !

*Dans la notation $\sum\limits_{k=n}^{m} a_k$, il est implicite que **l'indice** k **augmente de** 1 **lorsqu'on passe d'un terme au suivant** (idem pour le produit). Par exemple, la somme des entiers impairs de 1 à 15 ne s'écrit pas $\sum\limits_{k=1}^{15} k$ (qui correspond à la somme de tous les entiers de 1 à 15), mais $\sum\limits_{k=0}^{7} 2k+1$.*

Remarque 4.2 :

- *Le terme a_k est appelé terme général de la somme (ou produit), la valeur n est appelée valeur initiale de l'indice et m la valeur finale.*
- *Dans les formules donnant la somme ou le produit, l'indice est une variable dite **muette**, on peut lui donner le nom que l'on veut, **le résultat ne dépend pas de l'indice**.*

Cas particuliers

- **Terme général constant** : $\sum\limits_{k=n}^{m} \alpha = (m-n+1)\alpha$ et $\prod\limits_{k=n}^{m} \alpha = \alpha^{m-n+1}$ si $n \leqslant m$.
- **Sommes et produits télescopiques** : soit $(a_k)_{k\in[\![n;m+1]\!]}$ une famille de complexes, alors : $\sum\limits_{k=n}^{m}(a_{k+1}-a_k) = a_{m+1}-a_n$ et si aucun terme ne s'annule, alors $\prod\limits_{k=n}^{m} \frac{a_{k+1}}{a_k} = \frac{a_{m+1}}{a_n}$.
- **Sommes géométriques** : somme de termes consécutifs d'une suite u géométrique de raison $q \in \mathbb{C}$ (*i.e.* $\forall k \in \mathbb{N}$, $u_{k+1} = u_k \times q$) :

$$\sum_{k=n}^{m} u_k = \begin{cases} \frac{u_n - q\times u_m}{1-q} & \text{si } q \neq 1 \\ (m-n+1)u_p & \text{sinon} \end{cases}$$

On retiendra la formule $\boxed{\mathrm{S} = \dfrac{p-q\times d}{1-q}}$ où p désigne le premier terme de la somme et d le dernier.

- **Sommes arithmétiques** : somme de termes consécutifs d'une suite u arithmétique de raison $r \in \mathbb{C}$ (*i.e.* $\forall k \in \mathbb{N}$, $u_{k+1} = u_k + r$) :

$$\sum_{k=n}^{m} u_k = \frac{(m-n+1)(u_n+u_m)}{2}$$

On retiendra la formule $\boxed{\mathrm{S} = \dfrac{n(p+d)}{2}}$ où p désigne le premier terme de la somme, d le dernier et n le **nombre de termes**.

★**Exercice 4.1** *Démontrer les deux formules ci-dessus.*

2) Changement d'indice

Cas particuliers

Lorsque l'ensemble des indices est un intervalle d'entiers, les changements qui reviennent le plus souvent dans les sommes $\sum_{k=n}^{m} a_k$ (ou produits) sont les suivants (avec $n \leqslant m$) :

- **translation de l'indice** : soit à calculer $\sum_{k=n}^{m} a_k$, soit $p \in \mathbb{Z}$, posons un nouvel indice $k' = k + p$ (*i.e.* $k = k' - p$), on a alors :
$$\sum_{k=n}^{m} a_k = \sum_{k'=n+p}^{m+p} a_{k'-p}$$
En effet, $\sum_{k'=n+p}^{m+p} a_{k'-p} = a_n + a_{n+1} + \cdots + a_m = \sum_{k=n}^{m} a_k$.
- **symétrie de l'indice** : soit à calculer $\sum_{k=n}^{m} a_k$, soit $p \in \mathbb{Z}$, posons un nouvel indice $k' = p - k$ (*i.e.* $k = p - k'$) , on a alors :
$$\sum_{k=n}^{m} a_k = \sum_{k'=p-m}^{p-n} a_{p-k'}$$
En effet, $\sum_{k'=p-m}^{p-n} a_{p-k'} = a_m + a_{m-1} + \cdots + a_n = a_n + a_{n+1} + \cdots + a_m = \sum_{k=n}^{m} a_k$.

Attention !

Le nouvel indice doit également varier de 1 en 1. Considérons par exemple $\mathrm{S} = \sum_{k=1}^{n} 2k = 2 + 4 + 6 + \cdots + 2n$, *alors le changement d'indice* $k' = 2k$ *n'est pas « valable », car* $\sum_{k'=2}^{2n} k' = 2 + 3 + 4 + 5 + 6 + \cdots + 2n \neq \mathrm{S}$.

Formulation générale

Soit $(a_k)_{k\in \mathrm{I}}$ une famille finie de complexes indexée par I non vide, supposons qu'il existe une bijection $f : \mathrm{J} \to \mathrm{I}$ (où J désigne un autre ensemble), par composition on a ainsi une autre famille, indexée par J, et qui est $(a_{f(k')})_{k'\in \mathrm{J}}$. La fonction f **étant bijective** les deux familles comportent exactement les mêmes termes à l'ordre près, par conséquent la somme des termes des deux familles est la même, et le produit aussi (l'addition et la multiplication étant commutatives) :

$$\sum_{k\in \mathrm{I}} a_k = \sum_{k'\in \mathrm{J}} a_{f(k')} \text{ et } \prod_{k\in \mathrm{I}} a_k = \prod_{k'\in \mathrm{J}} a_{f(k')}$$

On dit qu'on a fait le changement d'indice $k = f(k')$ avec $k' \in \mathrm{J}$.

★**Exercice 4.2** *Résoudre dans $\mathbb{C}$ l'équation $(1-z)^{2n} = (1+z)^{2n}$ et calculer le produit des solutions **non nulles**.*

3) Propriétés

Théorème 4.1

Soient $\alpha, a_1, a_2, \ldots, a_n, b_1, \ldots, b_n, c_1, \ldots, c_m$ des nombres complexes, $p \leqslant q \leqslant n$ des entiers, on a :
- pour la somme :
$$\sum_{k=p}^{q} a_k + \sum_{k=q+1}^{n} a_k = \sum_{k=p}^{n} a_k \,;\ \sum_{k=p}^{n} \alpha \times a_k = \alpha \times \sum_{k=p}^{n} a_k \,;\ \sum_{k=p}^{n} (a_k + b_k) = \sum_{k=p}^{n} a_k + \sum_{k=p}^{n} b_k$$
- pour le produit :
$$\left(\prod_{k=p}^{q} a_k\right) \times \left(\prod_{k=q+1}^{n} a_k\right) = \prod_{k=p}^{n} a_k \,;\ \prod_{k=p}^{n} \alpha \times a_k = \alpha^{n-p+1} \times \prod_{k=p}^{n} a_k \,;\ \prod_{k=p}^{n} (a_k \times b_k) = \left(\prod_{k=p}^{n} a_k\right) \times \left(\prod_{k=p}^{n} b_k\right)$$

Preuve : Celle-ci est simple et laissée en exercice. □

4) Sommes doubles

Sur un rectangle

Soit $(a_{k,l})_{(k,l)\in[\![p;q]\!]\times[\![n;m]\!]}$ une famille indexée par $[\![p;q]\!]\times[\![n;m]\!]$ où $p \leqslant q$ et $n \leqslant m$ sont des entiers. La somme de la famille est appelée **somme double** et notée $\sum\limits_{(k,l)\in[\![p;q]\!]\times[\![n;m]\!]} a_{k,l} = \sum\limits_{\substack{p\leqslant k\leqslant q\\ n\leqslant l\leqslant m}} a_{k,l}$.

Disposons ces nombres dans un tableau en indexant les lignes de p à q, et les colonnes de n à m :

$a_{p,n}$	$a_{p,n+1}$	$\cdots$	$a_{p,m}$
$a_{p+1,n}$	$a_{p+1,n+1}$	$\cdots$	$a_{p+1,m}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$a_{q,n}$	$a_{q,n+1}$	$\cdots$	$a_{q,m}$

La somme des nombres figurant sur la ligne k est $S_k = a_{k,n} + a_{k,n+1} + \cdots + a_{k,m} = \sum\limits_{l=n}^{m} a_{k,l}$. Si maintenant nous faisons le total des sommes sur chaque ligne, nous obtenons $S_p + \cdots + S_q = \sum\limits_{k=p}^{q} S_k = \sum\limits_{k=p}^{q}\left(\sum\limits_{l=n}^{m} a_{k,l}\right)$, or ce total représente la somme des nombres de la famille (car l'addition est commutative et associative), d'où :

$$\sum_{\substack{p\leqslant k\leqslant q\\ n\leqslant l\leqslant m}} a_{k,l} = \sum_{k=p}^{q}\left(\sum_{l=n}^{m} a_{k,l}\right)$$

De même, la somme des nombres figurant sur la colonne l est $C_l = a_{p,l} + a_{p+1,l} + \cdots + a_{q,l} = \sum\limits_{k=p}^{q} a_{k,l}$. Si maintenant nous faisons le total des sommes sur chaque colonne, nous obtenons $C_n + \cdots + C_m = \sum\limits_{l=n}^{m} C_l = \sum\limits_{l=n}^{m}\left(\sum\limits_{k=p}^{q} a_{k,l}\right)$, or ce total représente la somme des nombres de la famille (car l'addition est commutative et associative), d'où finalement :

$$\boxed{\sum_{\substack{p\leqslant k\leqslant q\\ n\leqslant l\leqslant m}} a_{k,l} = \sum_{k=p}^{q}\left(\sum_{l=n}^{m} a_{k,l}\right) = \sum_{l=n}^{m}\left(\sum_{k=p}^{q} a_{k,l}\right)}$$

À retenir : produit de deux sommes simples

Soit $(a_k)_{k\in[\![p;q]\!]}$ et $(b_l)_{l\in[\![n;m]\!]}$ deux familles alors :

$$\sum_{\substack{p\leqslant k\leqslant q\\ n\leqslant l\leqslant m}} a_k b_l = \sum_{k=p}^{q}\left(\sum_{l=n}^{m} a_k b_l\right) = \sum_{k=p}^{q}\left(a_k \sum_{l=n}^{m} b_l\right) = \left(\sum_{k=p}^{q} a_k\right)\left(\sum_{l=n}^{m} b_l\right)$$

Sur un triangle

Soit $(a_{k,l})_{(k,l)\in A}$ une famille indexée par $A = \{(k,l) \mid 1 \leqslant k \leqslant l \leqslant n\}$ où n est un entier. La somme de la famille est notée $\sum\limits_{(k,l)\in A} a_{k,l} = \sum\limits_{1\leqslant k\leqslant l\leqslant n} a_{k,l}$.

Disposons ces nombres dans un tableau en indexant les lignes de 1 à n :

$a_{1,1}$	$a_{1,2}$	$\cdots$	$a_{1,n}$
	$a_{2,2}$	$\cdots$	$a_{2,2}$
		$\ddots$	$\vdots$
			$a_{n,n}$

La somme des nombres figurant sur la ligne k est $S_k = a_{k,k} + a_{k,k+1} + \cdots + a_{k,n} = \sum\limits_{l=k}^{n} a_{k,l}$. Si maintenant nous faisons le total des sommes sur chaque ligne, nous obtenons $S_1 + \cdots + S_n = \sum\limits_{k=1}^{n} S_k = \sum\limits_{k=1}^{n}\left(\sum\limits_{l=k}^{n} a_{k,l}\right)$, or ce total représente la somme des nombres de la famille (car l'addition est commutative et associative), d'où :

$$\sum_{1\leqslant k\leqslant l\leqslant n} a_{k,l} = \sum_{k=1}^{n}\left(\sum_{l=k}^{n} a_{k,l}\right)$$

De même, la somme des nombres figurant sur la colonne l est $C_l = a_{1,l} + a_{2,l} + \cdots + a_{l,l} = \sum\limits_{k=1}^{l} a_{k,l}$. Si maintenant nous faisons le total des sommes sur chaque colonne, nous obtenons $C_1 + \cdots + C_n = \sum\limits_{l=1}^{n} C_l = \sum\limits_{l=1}^{n}\left(\sum\limits_{k=1}^{l} a_{k,l}\right)$, or ce total représente la somme des nombres de la famille (car l'addition est commutative et associative), d'où finalement :

$$\boxed{\sum_{1\leqslant k\leqslant l\leqslant n} a_{k,l} = \sum_{k=1}^{n}\left(\sum_{l=k}^{n} a_{k,l}\right) = \sum_{l=1}^{n}\left(\sum_{k=1}^{l} a_{k,l}\right)}$$

II BINÔME DE NEWTON

1) Factorielle

Définition 4.4

Soit $n \in \mathbb{N}^*$, on pose $n! = \prod\limits_{k=1}^{n} k = 1 \times 2 \times \cdots \times n$. Par convention, on pose $0! = 1$.

Exemple : $1! = 1$, $2! = 2$, $3! = 6$, $4! = 24$, ...

À retenir

$\forall n \in \mathbb{N}$, $(n+1)! = (n+1) \times n!$

★**Exercice 4.3** *Montrer que* $\forall\, n \in \mathbb{N}^*, (n+1)! = 1 + \sum\limits_{k=1}^{n} k(k!)$.

2) Coefficients binomiaux

Définition 4.5

Soient n, p deux entiers positifs tels que $p \leqslant n$, on pose $\binom{n}{p} = \frac{n!}{p!(n-p)!}$ (lire p parmi n).
Par convention, lorsque $p > n$, ou lorsque $p < 0$, on pose $\binom{n}{p} = 0$. Le coefficient binomial $\binom{n}{p}$ est ainsi défini pour $n \in \mathbb{N}$ et $p \in \mathbb{Z}$.

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

Lorsque $p \in [\![0\,;n]\!]$, en simplifiant dans la formule $n!$ avec $(n-p)!$, il reste :

À retenir : formule pour le calcul pratique

Si $1 \leqslant p \leqslant n$ alors $\binom{n}{p} = \frac{n(n-1)\cdots(n-p+1)}{p!} = \frac{\prod\limits_{k=0}^{p-1}(n-k)}{p!}$ (formule valable si $p = 0$).

Cette formule pratique permet d'étendre la définition à tout réel x (et $p \in \mathbb{N}$) en posant : $\binom{x}{p} = \frac{\prod\limits_{k=0}^{p-1}(x-k)}{p!}$.

Attention !

*La fonction factorielle telle que nous l'avons définie, ne s'applique qu'à des **entiers positifs**.*

Théorème 4.2 (propriétés)

Soient $n \in \mathbb{N}$ et $p \in \mathbb{Z}$:
- Si $0 \leqslant p \leqslant n$ alors $\binom{n}{p} = \binom{n}{n-p}$ (symétrie).
- Si $n \geqslant 1$, $p\binom{n}{p} = n\binom{n-1}{p-1}$.

- $\binom{n}{p} + \binom{n}{p+1} = \binom{n+1}{p+1}$ *(relation de Pascal).*

Preuve : La première formule découle directement de la définition.

La deuxième formule est évidente si $p > n$ ou si $p \leqslant 1$, supposons $1 \leqslant p \leqslant n$, $p\binom{n}{p} = \frac{n!}{(n-p)!(p-1)!} = n\frac{(n-1)!}{(n-1-(p-1))!(p-1)!} = n\binom{n-1}{p-1}$.

Pour la relation de Pascal, elle est simple à vérifier si $p > n$ ou si $p = n$ ou si $p < 0$. Supposons $0 \leqslant p \leqslant n-1$, alors $\binom{n}{p} + \binom{n}{p+1} = \frac{n!}{(n-p)!p!} + \frac{n!}{(n-p-1)!(p+1)!} = \frac{n!(p+1)+n!(n-p)}{(n-p)(p+1)!} = \frac{(n+1)!}{(n-p)!(p+1)!} = \binom{n+1}{p+1}$. □

Triangle de Pascal

La relation de Pascal permet de calculer les coefficients binomiaux de proche en proche dans un tableau :

$n\backslash p$	0	1	2	3	4	5	6	7	8
0	1								
1	1	1							
2	1	2	1						
3	1	3	3	1					
4	1	4	6+	4 =	1				
5	1	5	10	10	5	1			
6	1	6	15	20	15	6	1		
7	1	7	21	35	35	21	7	1	
8	1	8	28	56	70	56	28	8	1

3) Formule du binôme

Théorème 4.3

Soient $a, b \in \mathbb{C}$ et $n \in \mathbb{N}$, alors : $(a+b)^n = \sum\limits_{k=0}^{n} \binom{n}{k} a^k b^{n-k} = \sum\limits_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$.

Preuve : Par récurrence sur n : au rang 0 la formule donne 1 ce qui correspond bien à $(a+b)^0$. Si la formule est démontrée au rang n, alors :

$$\begin{aligned}
(a+b)^{n+1} &= (a+b)(a+b)^n = a(a+b)^n + b(a+b)^n \\
&= \sum_{k=0}^{n} \binom{n}{k} a^{k+1} b^{n-k} + \sum_{k=0}^{n} \binom{n}{k} a^k b^{n+1-k} \\
&= \sum_{p=1}^{n+1} \binom{n}{p-1} a^p b^{n+1-p} + \sum_{k=0}^{n} \binom{n}{k} a^k b^{n+1-k} \text{ (changement d'indice } p = k+1 \text{ dans la première somme)} \\
&= \sum_{k=1}^{n+1} \binom{n}{k-1} a^k b^{n+1-k} + \sum_{k=0}^{n} \binom{n}{k} a^k b^{n+1-k} \\
&= a^{n+1} + b^{n+1} + \sum_{k=1}^{n} \left[\binom{n}{k-1} + \binom{n}{k} \right] a^k b^{n+1-k} \text{ (on regroupe les sommes sur } [\![1\,;n]\!]) \\
&= a^{n+1} + b^{n+1} + \sum_{k=1}^{n} \binom{n+1}{k} a^k b^{n+1-k} \text{ (relation de Pascal)} \\
&= \sum_{k=0}^{n+1} \binom{n+1}{k} a^k b^{n+1-k} \text{ (formule au rang } n+1)
\end{aligned}$$

La formule est donc établie pour tout $n \in \mathbb{N}$. □

☞**Exemples** :

- $\sum\limits_{k=0}^{n} \binom{n}{k} = (1+1)^n = 2^n$.
- Si $n \geqslant 1$, $\sum\limits_{k=1}^{n} k\binom{n}{k} = \sum\limits_{k=1}^{n} n\binom{n-1}{k-1} = n2^{n-1}$.
- $\sum\limits_{k=0}^{n} \binom{n}{k} \cos(k\frac{\pi}{3}) = \mathrm{Re}([1 + e^{i\pi/3}]^n) = \sqrt{3^n}\cos(n\frac{\pi}{6})$.

4) Une autre identité remarquable

Théorème 4.4
Soient $a, b \in \mathbb{C}$ et $n \in \mathbb{N}$, alors :
$$a^{n+1} - b^{n+1} = (a-b)(a^n + a^{n-1}b + \cdots + ab^{n-1} + b^n) = (a-b)\sum_{k=0}^{n} a^{n-k}b^k.$$

Preuve : Si $a = b$ il n'y a rien à faire, si $a = 0$ alors le terme de droite devient $-b \times b^n = -b^{n+1}$, la formule est encore vraie. Supposons maintenant que $a \neq 0$ et $a \neq b$, alors :

$$\frac{a^{n+1} - b^{n+1}}{a-b} = a^n \frac{1 - \left(\frac{b}{a}\right)^{n+1}}{1 - \frac{b}{a}} = a^n \sum_{k=0}^{n} \left(\frac{b}{a}\right)^k = \sum_{k=0}^{n} a^{n-k}b^k$$

Remarque : une preuve par récurrence est également possible. □

Cas particulier : si n est pair alors $n+1$ est impair et donc $a^{n+1} + b^{n+1} = a^{n+1} - (-b)^{n+1}$ ce qui donne alors en remplaçant b par $-b$:

$$a^{n+1} + b^{n+1} = (a+b)\sum_{k=0}^{n}(-1)^k a^{n-k}b^k \quad (\text{si } n+1 \text{ est impair})$$

À retenir : Une application
Si $\mathrm{P}(x)$ désigne une fonction polynomiale en la variable x, et si b est une racine de P (*i.e.* $\mathrm{P}(b) = 0$), alors $\mathrm{P}(x)$ est factorisable par $(x-b)$.

Preuve : En écrivant $\mathrm{P}(x) = \sum_{k=0}^{n} \alpha_k x^k$, alors :

$$\mathrm{P}(x) = \mathrm{P}(x) - \mathrm{P}(b) = \sum_{k=1}^{n} \alpha_k (x^k - b^k) = (x-b)\sum_{k=1}^{n} \alpha_k \left(\sum_{p=0}^{k-1} x^{k-1-p}b^p\right)$$

□

★**Exercice 4.4** *En écrivant $a^{n+1} - b^{n+1} = ((a-b)+b)^{n+1} - b^{n+1}$, trouver une autre factorisation par $(a-b)$. Quelle égalité en déduisez-vous ?*

III SYSTÈMES LINÉAIRES

1) Définition

Définition 4.6
Un système (S) *linéaire à n équations et p inconnues est un système d'équations de la forme :*

$$(\mathrm{S}) \left\{ \begin{array}{lcl} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,p}x_p & = & b_1 \\ \vdots & & \vdots \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,p}x_p & = & b_n \end{array} \right.$$

où $x_1, \ldots, x_p$ sont des nombres ***inconnus****, $b_1, \ldots, b_n$ sont des nombres donnés (appelés seconds membres) et la famille de nombres $(a_{i,j})_{(i,j) \in [\![1;n]\!] \times [\![1;p]\!]}$ est donnée (ce sont les coefficients, on dit qu'ils forment la matrice du système).*
Résoudre (S) *c'est trouver tous les p-uplets de nombres $(x_1, \ldots, x_p)$ vérifiant* (S)*. Deux systèmes sont dits équivalents si et seulement si ils ont les mêmes solutions.*

Codage

Les différentes équations d'un système linéaire sont notées du haut vers le bas $\mathrm{L}_1, \mathrm{L}_2, \ldots, \mathrm{L}_n$:

$$(\mathrm{S}) \left\{ \begin{array}{lcll} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,p}x_p & = & b_1 & (\mathrm{L}_1) \\ \vdots & & \vdots & \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,p}x_p & = & b_n & (\mathrm{L}_n) \end{array} \right.$$

Les différentes colonnes du système sont notées de gauche à droite $\mathrm{C}_1, \mathrm{C}_2, \ldots, \mathrm{C}_p$.

2) Interprétation géométrique

Lorsque $p = 2$: si les coefficients $a_{i,1}$ et $a_{1,2}$ ne sont pas nuls simultanément, alors l'équation $a_{i,1}x_1 + a_{i,2}x_2 = b_i$ peut être interprétée comme l'équation d'une droite $\mathcal{D}_i$ dans un plan muni d'un repère $(\mathrm{O}, \vec{\imath}, \vec{\jmath})$, les solutions étant les coordonnées des points $\mathrm{M}(x_1, x_2)$ de la droite $\mathcal{D}_i$. Résoudre (S) revient alors à chercher un point commun à plusieurs droites, c'est à dire $\mathcal{D}_1 \cap \cdots \cap \mathcal{D}_n$. Dans le cas où $n = 2$, il y a trois cas possibles :

- Soit $\mathcal{D}_1$ et $\mathcal{D}_2$ sont sécantes et alors le système a une unique solution.
- Soit $\mathcal{D}_1$ et $\mathcal{D}_2$ sont confondus et alors les solutions du système sont les coordonnées des points de $\mathcal{D}_1$.
- Soit $\mathcal{D}_1$ et $\mathcal{D}_2$ sont strictement parallèles et alors il n'y a pas de solutions au système.

Lorsque $p = 3$: si les coefficients $a_{i,1}$, $a_{1,2}$ et $a_{i,3}$ ne sont pas nuls simultanément, alors l'équation $a_{i,1}x_1 + a_{i,2}x_2 + a_{i,3}x_3 = b_i$ peut être interprétée comme l'équation d'un plan $\mathcal{P}_i$ dans un plan muni d'un repère $(\mathrm{O}, \vec{\imath}, \vec{\jmath}, \vec{k})$, les solutions étant les coordonnées des points $\mathrm{M}(x_1, x_2, x_3)$ du plan $\mathcal{P}_i$. Résoudre (S) revient alors à chercher un point commun à plusieurs plan c'est à dire $\mathcal{P}_1 \cap \cdots \cap \mathcal{P}_n$. Dans le cas où $n = 2$, il y a trois cas possibles :

- Soit $\mathcal{P}_1$ et $\mathcal{P}_2$ sont confondus et alors les solutions du système sont les coordonnées des points de $\mathcal{P}_1$.
- Soit $\mathcal{P}_1$ et $\mathcal{P}_2$ sont strictement parallèles et alors il n'y a pas de solutions au système.
- Soit $\mathcal{P}_1$ et $\mathcal{P}_2$ se coupent suivant une droite $\mathcal{D}_{1,2}$. S'il y a une troisième équation, il s'agit ensuite d'étudier l'intersection $\mathcal{D}_{1,2} \cap \mathcal{P}_3$, il y a plusieurs cas :
  - Soit la droite coupe le plan en un point I et ses coordonnées forment l'unique solution du système.
  - Soit la droite est incluse dans le plan et l'intersection est la droite.
  - Soit la droite est strictement parallèle au plan et il n'y a pas de solution.

3) Méthode du pivot de Gauss

Systèmes triangulaires

Un système linéaire (S) est dit triangulaire lorsque les coefficients situés sous la diagonale principale (les coefficients $a_{i,j}$ avec $i > j$) sont nuls. **Lorsqu'un système triangulaire a ses coefficients diagonaux non nuls, on le résout par substitutions remontantes**. Exemple :

$$\begin{cases} x + 2y - z + t = 1 \\ \quad y - z + 2t = 0 \\ \qquad z + t = 1 \end{cases} \iff \begin{cases} x + 2y - z + t = 1 \\ \quad y - z + 2t = 0 \\ \qquad z = 1 - t \end{cases} \iff \begin{cases} x + 2y - z + t = 1 \\ \quad y = z - 2t = 1 - 3t \\ \quad z = 1 - t \end{cases}$$

$$\iff \begin{cases} x = 1 - 2y + z - t = 4t \\ y = 1 - 3t \\ z = 1 - t \end{cases}$$

L'ensemble des solutions est donc $\{(4t, 1 - 3t, 1 - t, t) \ / \ \in \mathbb{R}\}$. Sur cet exemple, on voit que l'inconnue t peut-être quelconque et que les autres s'écrivent en fonction de t. On dit que t est une inconnue **auxiliaire** et que x, y, z sont les inconnues **principales**.

Opérations élémentaires

Les opérations élémentaires de la méthode de Gauss sont :

- L'échange de deux équations L_i et L_j avec $i \neq j$, notée $\mathrm{L}_i \leftrightarrow \mathrm{L}_j$, elle transforme le système en un système équivalent car le connecteur « et » est commutatif.
- L'échange de deux colonnes C_i et C_j avec $i \neq j$, notée $\mathrm{C}_i \leftrightarrow \mathrm{C}_j$, elle transforme le système en un système équivalent car l'addition est commutative.
- Multiplier une ligne L_i par un nombre α **non nul**, notée $\mathrm{L}_i \leftarrow \alpha\mathrm{L}_i$, elle transforme le système en un système équivalent car la nouvelle équation L_i est équivalente à l'ancienne puisque $\alpha \neq 0$.
- Ajouter à une ligne L_i un multiple d'une autre ligne L_j $(i \neq j)$, elle notée $\mathrm{L}_i \leftarrow \mathrm{L}_i + \lambda\mathrm{L}_j$, elle transforme le système en un système équivalent car l'opération $\mathrm{L}_i \leftarrow \mathrm{L}_i - \lambda\mathrm{L}_j$ nous permet de revenir à l'ancien système.

À retenir

La méthode du pivot de Gauss consiste à transformer, avec des opérations élémentaires, le système linéaire initial (S) en un système équivalent triangulaire dont les coefficients diagonaux sont non nuls.

La méthode est une succession d'étapes, voici le descriptif de l'étape k, celle-ci se déroule en trois temps :

1. On choisit un coefficient **non nul** dans les lignes L_k à L_n et les colonnes C_k à C_p. Ce coefficient sera appelé le pivot de l'étape k (noté p_k), pour les calculs à la main on essaie de choisir si possible un coefficient égal à 1 ou −1.
2. On amène le pivot p_k **à sa place**, c'est à dire ligne L_k et colonne C_k, il peut être nécessaire pour cela d'échanger deux lignes et/ou deux colonnes.
3. On « élimine » les coefficients situés sous le pivot dans les lignes L_{k+1} à L_n avec des opérations du type $L_i \leftarrow L_i + \lambda L_k$, c'est à dire qu'à chaque ligne (sous celle du pivot) on ajoute un certain nombre de fois la ligne du pivot pour faire disparaître le coefficient qui est dans la colonne du pivot, ceci est possible en faisant $L_i \leftarrow L_i - \frac{a_{i,k}}{p_k} L_k$ si $a_{i,k}$ est le coefficient de la ligne i et colonne k.

À l'issue de l'étape k on obtient un système de la forme :

$$\left\{\begin{array}{ccccccccc}
p_1x_1+ & a_{1,2}x_2+ & \cdots & & \cdots & \cdots & +a_{1,p}x_p & = & b_1 \\
 & p_2x_2+ & \cdots & & \cdots & \cdots & +a_{2,p}x_p & = & b_2 \\
 & \ddots & & & & & \vdots & & \vdots \\
 & & p_kx_k+ & & \cdots & \cdots & +a_{k,p}x_p & = & b_k \\
 & & & & a_{k+1,k+1}x_{k+1}+ & \cdots & +a_{k+1,p}x_p & = & b_{k+1} \\
 & & & & \vdots & & \vdots & & \vdots \\
 & & & & a_{n,k+1}x_{k+1}+ & \cdots & +a_{n,p}x_p & = & b_n
\end{array}\right. ,$$

La méthode s'arrête à l'issue de l'étape k lorsqu'on est arrivé à la dernière ligne, ou bien lorsqu'il n'est plus possible de trouver un pivot parce que tous les coefficients des lignes L_{k+1} à L_n et des colonnes C_{k+1} à C_p sont nuls. On peut alors procéder aux substitutions remontantes.

Exemple :

$$(S)\left\{\begin{array}{rcl}
\boxed{2x_1} + 3x_2 - 3x_3 + 4x_4 + 2x_5 & = & 1 \\
3x_1 + 6x_2 - 2x_3 + 5x_4 + 9x_5 & = & 0 \\
7x_1 + 18x_2 - 2x_3 + 7x_4 + 7x_5 & = & 2 \\
2x_1 + 4x_2 - 2x_3 + 3x_4 + x_5 & = & -1
\end{array}\right.$$

Étape 1 :

$$\iff \left\{\begin{array}{rcll}
\boxed{2x_1} + 3x_2 - 3x_3 + 4x_4 + 2x_5 & = & 1 & \\
3x_2 + 5x_3 - 2x_4 + 12x_5 & = & -3 & L_2 \leftarrow 2L_2 - 3L_1 \\
15x_2 + 17x_3 - 14x_4 & = & -3 & L_3 \leftarrow 2L_3 - 7L_1 \\
x_2 + x_3 - x_4 - x_5 & = & -2 & L_4 \leftarrow L_4 - L_1
\end{array}\right.$$

Étape 2 ($L_2 \leftrightarrow L_4$) :

$$\iff \left\{\begin{array}{rcll}
\boxed{2x_1} + 3x_2 - 3x_3 + 4x_4 + 2x_5 & = & 1 & \\
\boxed{x_2} + x_3 - x_4 - x_5 & = & -2 & L_2 \leftarrow L_4 \\
15x_2 + 17x_3 - 14x_4 & = & -3 & \\
3x_2 + 5x_3 - 2x_4 + 12x_5 & = & -3 & L_4 \leftarrow L_2
\end{array}\right.$$

$$\iff \left\{\begin{array}{rcll}
\boxed{2x_1} + 3x_2 - 3x_3 + 4x_4 + 2x_5 & = & 1 & \\
\boxed{x_2} + x_3 - x_4 - x_5 & = & -2 & \\
2x_3 + x_4 + 15x_5 & = & 27 & L_3 \leftarrow L_3 - 15L_2 \\
2x_3 + x_4 + 15x_5 & = & 3 & L_4 \leftarrow L_4 - 3L_2
\end{array}\right.$$

Étape 3 :

$$\iff \left\{\begin{array}{rcll}
\boxed{2x_1} + 3x_2 - 3x_3 + 4x_4 + 2x_5 & = & 1 & \\
\boxed{x_2} + x_3 - x_4 - x_5 & = & -2 & \\
\boxed{2x_3} + x_4 + 15x_5 & = & 27 & \\
0 & = & -24 & L_4 \leftarrow L_4 - L_3
\end{array}\right.$$

On voit donc que le système n'a pas de solution.

★**Exercice 4.5** *Résoudre en fonction du paramètre* λ *le système* $\begin{cases} (1-\lambda)x - y + z = 0 \\ -x + (1-\lambda)y - z = 0 \\ x - y + (1-\lambda)z = 0 \end{cases}$.

IV SOLUTION DES EXERCICES

Solution 4.1

1/ *Pour la suite arithmétique de raison r, si le premier terme est p, alors le suivant est* $p+r$*, puis* $p+2r$*, etc, et le dernier est* $d = p+(n-1)r$*, la somme s'écrit alors* $\mathrm{S} = np + r\times(1+2+\cdots+(n-1))$*, or* $1+2+\cdots+(n-1) = \frac{n(n-1)}{2}$ *(cela peut se vérifier par récurrence), d'où* $\mathrm{S} = np + \frac{nr(n-1)}{2} = \frac{n(2p+r(n-1))}{2} = \frac{n(p+d)}{2}$.

2/ *Pour la suite géométrique de raison q, si le premier terme est p, alors le suivant est pq, puis* pq^2*, etc, et le dernier est* $d = pq^{n-1}$*, la somme s'écrit alors* $\mathrm{S} = p\times(1+q+q^2+\cdots+q^{n-1})$*, or* $(1-q)\times(1+q+q^2+\cdots+q^{n-1}) = 1-q^n$ *après simplifications, donc si* $q\neq 1$ *alors* $1+q+q^2+\cdots+q^{n-1} = \frac{1-q^n}{1-q}$ *et donc* $\mathrm{S} = \frac{p-qd}{1-q}$*, par contre, si* $q=1$*, alors tous les termes de la somme sont égaux à p et dans ce cas,* $\mathrm{S} = np$.

Solution 4.2 *L'équation équivaut à* $\left(\frac{1+z}{1-z}\right)^{2n} = 1$*, ou encore* $\frac{1+z}{1-z} = e^{i\frac{k\pi}{n}}$ *avec* $k\in[\![0\,;2n-1]\!]$*, ce qui donne* $z = \frac{e^{i\frac{k\pi}{n}}-1}{e^{i\frac{k\pi}{n}}+1}$ *ou encore* $z = i\tan(\frac{k\pi}{2n})$ *avec* $k\in[\![0\,;n-1]\!]\cup[\![n+1\,;2n-1]\!]$.

Le produit des solutions non nulles est $\mathrm{P} = \prod_{k=1}^{n-1} i\tan(\frac{k\pi}{2n}) \times \prod_{k=n+1}^{2n-1} i\tan(\frac{k\pi}{2n})$*, en posant* $q = k-n$ *dans le deuxième produit, on obtient* $\mathrm{P} = \prod_{k=1}^{n-1} i\tan(\frac{k\pi}{2n}) \times \prod_{q=1}^{n-1} -\frac{i}{\tan(\frac{q\pi}{2n})}$*, d'où* $\mathrm{P} = i^{n-1}\times(-1)^{n-1}\times i^{n-1} = 1$.

Solution 4.3 *On a* $k(k!) = (k+1)! - k!$*, la somme est donc télescopique, d'où* $\sum_{k=1}^{n} k(k!) = \sum_{k=1}^{n}(k+1)! - \sum_{k=1}^{n} k! = \sum_{k=2}^{n+1} k! - \sum_{k=1}^{n} k! = (n+1)!-1$ *ce qui entraîne le résultat.*

Solution 4.4 *Par le binôme de Newton* $a^{n+1} - b^{n+1} = ((a-b)+b)^{n+1} - b^{n+1} = \left(\sum_{k=0}^{n+1}\binom{n+1}{k}(a-b)^{n+1-k}b^k\right) - b^{n+1} = \sum_{k=0}^{n}\binom{n+1}{k}(a-b)^{n+1-k}b^k = (a-b)\sum_{k=0}^{n}\binom{n+1}{k}(a-b)^{n-k}b^k$.

On en déduit pour $a\neq b$ *que* $\sum_{k=0}^{n}\binom{n+1}{k}(a-b)^{n-k}b^k = \sum_{k=0}^{n} a^{n-k}b^k$ *(on peut vérifier que c'est encore vrai si* $a=b$ *car on obtient* $(n+1)b^n$ *de chaque côté).*

Solution 4.5 *Notons* (S_λ) *le système* $\begin{cases} (1-\lambda)x - y + z = 0 \\ -x + (1-\lambda)y - z = 0 \\ x - y + (1-\lambda)z = 0 \end{cases}$*, alors :*

$$(\mathrm{S}_\lambda) \iff \begin{cases} \boxed{z} - y + (1-\lambda)x = 0 \\ -z + (1-\lambda)y - x = 0 \\ (1-\lambda)z - y + x = 0 \end{cases} \iff \begin{cases} \boxed{z} - y + (1-\lambda)x = 0 & \\ -\lambda y - \lambda x = 0 & \mathrm{L}_2 \leftarrow \mathrm{L}_2 + \mathrm{L}_1 \\ -\lambda y + \lambda(2-\lambda)x = 0 & \mathrm{L}_3 \leftarrow \mathrm{L}_3 + (\lambda-1)\mathrm{L}_1 \end{cases}$$

Si $\lambda = 0$ *: le système se réduit à une équation* $z - y + x = 0$ *d'où* $y = z + x$ *avec z et x quelconques.*

Lorsque $\lambda\neq 0$ *:* $(\mathrm{S}_\lambda) \iff \begin{cases} z - y + (1-\lambda)x = 0 \\ y + x = 0 \\ y + (\lambda-2)x = 0 \end{cases} \iff \begin{cases} z - y + (1-\lambda)x = 0 & \\ y + x = 0 & \\ (\lambda-3)x = 0 & \mathrm{L}_3 \leftarrow \mathrm{L}_3 - \mathrm{L}_2 \end{cases}$

Si $\lambda\notin\{0;3\}$ *alors l'unique solution est* $(0,0,0)$.

Si $\lambda = 3$ *: alors x est quelconque,* $y = -x$ *et* $z = y + 2x = x$.

Chapitre 5

Fonctions usuelles

Sommaire

I FONCTIONS LOGARITHME ET EXPONENTIELLE

1) Logarithme népérien

La fonction $x \mapsto \frac{1}{x}$ est continue sur $]0\,;+\infty[$, elle admet une unique primitive qui s'annule en 1.

Définition 5.1

*L'unique primitive de la fonction $x \mapsto \frac{1}{x}$ sur $]0\,;+\infty[$ qui s'annule en 1 est appelée **logarithme népérien** et notée $\ln$. On a donc $\forall x > 0, \ln(x) = \int_1^x \frac{dt}{t}$.*

Cette fonction est donc dérivable sur $I =]0\,;+\infty[$ et pour $x > 0$, $\boxed{\ln'(x) = \frac{1}{x}}$, elle est donc strictement croissante sur I.

Soit $a > 0$, la fonction $f : x \mapsto \ln(ax)$ et dérivable sur I par composition, et pour $x > 0$, $f'(x) = a\frac{1}{ax} = \frac{1}{x}$, on en déduit que $f(x) = \ln(x) + c$ où c est une constante, on a $\ln(a) = f(1) = \ln(1) + c = c$, par conséquent on obtient $\ln(ax) = \ln(a) + \ln(x)$:

Théorème 5.1 (Propriété fondamentale du logarithme)

$$\forall x, y > 0, \ln(xy) = \ln(x) + \ln(y).$$

Conséquences :

- Si u est une fonction dérivable qui ne s'annule pas, alors $\ln(|u|)$ est dérivable et $[\ln(|u|)]' = \frac{u'}{u}$.
- $\forall x, y \in \mathbb{R}^*, \ln(|xy|) = \ln(|x|) + \ln(|y|)$.
- $\forall x, y \in \mathbb{R}^*, \ln(|\frac{x}{y}|) = \ln(|x|) - \ln(|y|)$.
- $\forall n \in \mathbb{Z}^*, \forall x \in \mathbb{R}^*, \ln(|x^n|) = n\ln(|x|)$.

Théorème 5.2 (Limites du logarithme népérien)

$$\lim_{x\to+\infty} \ln(x) = +\infty;\ \lim_{x\to 0^+} \ln(x) = -\infty;\ \lim_{x\to+\infty} \frac{\ln(x)}{x} = 0;\ \lim_{x\to 0^+} x\ln(x) = 0;\ \lim_{x\to 1} \frac{\ln(x)}{x-1} = 1.$$

Preuve : $\forall n \in \mathbb{N}, \ln(2^n) = n\ln(2)$ or $\ln(2) > 0$, donc la suite $(\ln(2^n))$ tend vers $+\infty$ ce qui prouve que la fonction ln n'est pas majorée, par conséquent, comme elle est croissante, elle tend $+\infty$ en $+\infty$.

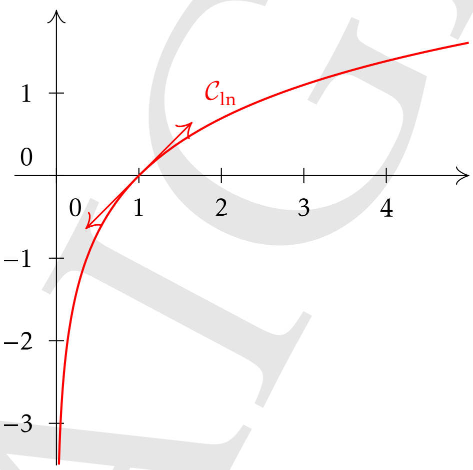
En posant $X = \frac{1}{x}$ on a $\lim\limits_{x\to 0^+} X = +\infty$ donc $\lim\limits_{x\to 0^+} \ln(x) = \lim\limits_{X\to+\infty} -\ln(X) = -\infty$.

$\lim\limits_{x\to 1} \frac{\ln(x)}{x-1} = \ln'(1) = 1$.

Pour $t \geqslant 1$ on a $\sqrt{t} \leqslant t$ et donc pour $x \geqslant 1$ on a $0 \leqslant \ln(x) \leqslant \int_1^x \frac{dt}{\sqrt{t}} = 2[\sqrt{x} - 1]$, le théorème des gendarmes entraîne $\lim\limits_{x\to+\infty} \frac{\ln(x)}{x} = 0$. □

Courbe représentative :

x	0		1		$+\infty$
$\ln'$	‖	$+$		$+$	
$\ln$	‖ $-\infty$	$\nearrow$	0	$\nearrow$	$+\infty$



La tangente au point d'abscisse 1 a pour équation $y = x - 1$.

Théorème 5.3 (Inégalité de convexité)

$$\forall x > 0, \ln(x) \leqslant x - 1.$$

Preuve : Il suffit d'étudier la fonction $f : x \mapsto \ln(x) - x + 1$. □

2) La fonction exponentielle

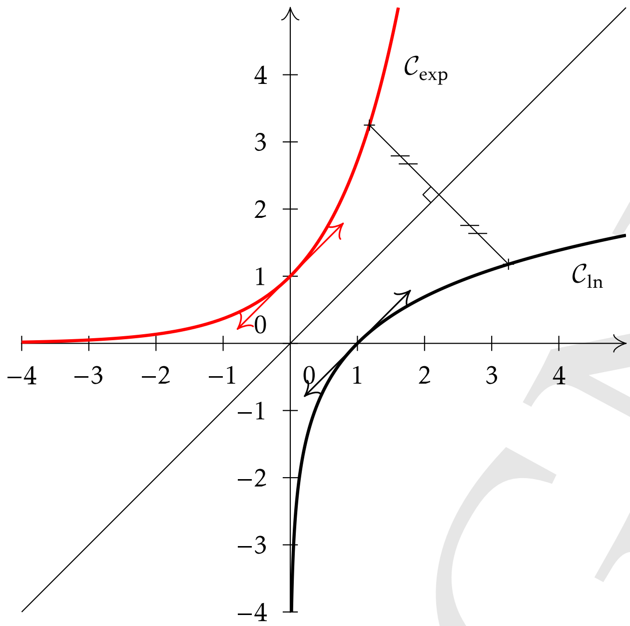
La fonction ln est strictement croissante sur $I =]0\,;+\infty[$, elle définit donc une bijection de I sur $J = \ln(I)$, comme elle est continue et strictement croissante, on a $\ln(I) =]\lim\limits_{0} \ln\,;\lim\limits_{+\infty} \ln[= \mathbb{R}$.

Définition 5.2

La réciproque est appelée ***fonction exponentielle*** *et notée* exp, *elle est définie par :*

$$\begin{array}{rcl} \exp : \mathbb{R} & \to &]0\,;+\infty[\\ x & \mapsto & \exp(x) = y \text{ tel que } y > 0 \text{ et } \ln(y) = x \end{array}.$$

Propriétés :
- La fonction exp est strictement croissante sur $\mathbb{R}$ et continue, de plus $\exp(0) = 1$.
- La fonction ln est dérivable sur $]0\,;+\infty[$ et sa dérivée ne s'annule pas, donc la fonction exp est dérivable sur $\mathbb{R}$ et $\forall x \in \mathbb{R}$, $\boxed{\exp'(x) = \frac{1}{\ln'(\exp(x))} = \exp(x)}$.
- Dans un repère orthonormé, la courbe de la fonction exp et celle de la fonction ln sont symétriques par rapport à la première bissectrice.



L'équation de la tangente à la courbe au point d'abscisse 0 est $y = x + 1$.

Soient $x, y \in \mathbb{R}$, notons $X = \exp(x)$ et $Y = \exp(y)$ alors X et Y sont dans $]0\,;+\infty[$ on peut donc écrire $\ln(XY) = \ln(X) + \ln(Y)$ ce qui donne $x + y = \ln(XY)$, par conséquent $\exp(x + y) = XY = \exp(x)\exp(y)$, on peut donc énoncer :

Théorème 5.4 (Propriété fondamentale de l'exponentielle)

$$\forall x, y \in \mathbb{R}, \exp(x + y) = \exp(x)\exp(y).$$

Il en découle en particulier que $\exp(-x) = \frac{1}{\exp(x)}$.

Notation : On déduit de ce théorème que pour tout $n \in \mathbb{Z}$ et pour tout x on a $\exp(nx) = [\exp(x)]^n$. En particulier on a pour $x = 1$, $\exp(n) = [\exp(1)]^n$. On pose alors $\boxed{e = \exp(1)}$, on peut donc écrire que $\exp(n) = e^n$. On convient alors d'écrire pour tout réel x :

$$\boxed{\exp(x) = e^x}.$$

Les propriétés deviennent alors :
- $e^{x+y} = e^x \times e^y$.
- $e^0 = 1$, $e^{-x} = \frac{1}{e^x}$, $\forall n \in \mathbb{Z}$, $e^{nx} = [e^x]^n$.
- Si u désigne une fonction dérivable alors eu est dérivable et $[e^u]' = u' \times e^u$.
- $\forall x \in \mathbb{R}, e^x \geqslant x + 1$ (inégalité de convexité).
  Preuve : Soit $X = e^x$, on sait que $\ln(X) \leqslant X - 1$ ce qui donne l'inégalité. □

Théorème 5.5 (Limites de la fonction exponentielle)

$$\lim_{x\to-\infty} e^x = 0,\ \lim_{x\to+\infty} e^x = +\infty,\ \lim_{x\to+\infty} \frac{e^x}{x} = +\infty,\ \lim_{x\to 0} \frac{e^x - 1}{x} = 1.$$

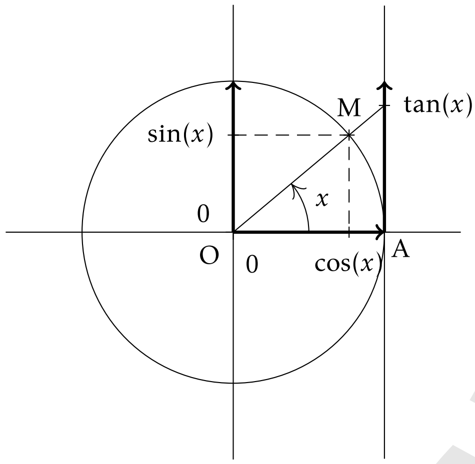
Preuve : La fonction exp est continue et strictement croissante sur $\mathbb{R}$ donc $\exp(\mathbb{R}) =]\lim_{-\infty} \exp\,; \lim_{+\infty} \exp[=]0\,;+\infty[$. Soit $X = e^x$ alors $\lim\limits_{x\to+\infty} \frac{e^x}{x} = \lim\limits_{X\to+\infty} \frac{X}{\ln(X)} = +\infty$. $\lim\limits_{x\to 0} \frac{e^x-1}{x} = \exp'(0) = 1$. □

Remarque 5.1 – *Il en découle que* $\lim\limits_{x\to+\infty} xe^{-x} = 0$, *ou encore que* $\lim\limits_{x\to+\infty} \frac{x}{e^x} = 0$

II FONCTIONS CIRCULAIRES - INVERSIONS

1) Fonctions circulaires : rappels

Le plan $\mathcal{P}$ est muni d'un repère orthonormé direct $(O, \vec{u}, \vec{v})$. Soit x un réel, et $M(x)$ le point du cercle trigonométrique tel que $(\vec{u}, \overrightarrow{OM}) = x \pmod{2\pi}$ alors les coordonnées de $M(x)$ sont $(\cos(x), \sin(x))$, lorsque $x \in \mathbb{R} \setminus \left\{\frac{\pi}{2} + k\pi \mid k \in \mathbb{Z}\right\}$, on pose $\tan(x) = \frac{\sin(x)}{\cos(x)}$.

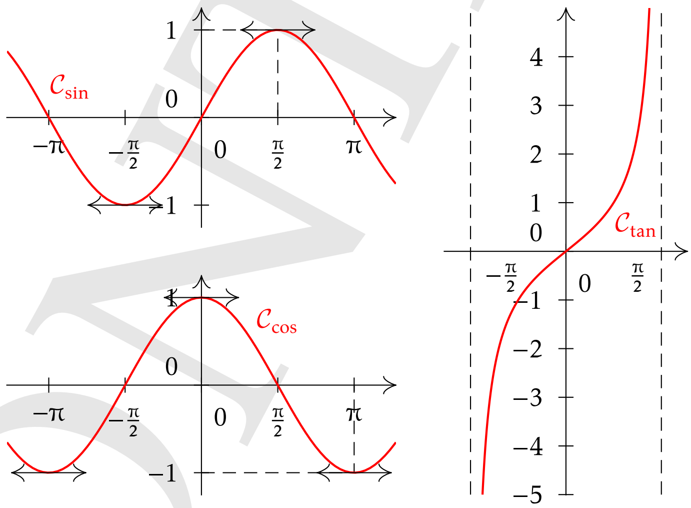


Remarque 5.2 – *Le réel x représente également la longueur de l'arc de cercle $\widehat{\mathrm{AM}}$ avec $\mathrm{A}(1,0)$, le cercle étant orienté dans le sens direct.*

Quelques propriétés :
- $\forall x \in \mathbb{R}, \cos^2(x) + \sin^2(x) = 1$.
- Les fonctions sinus et cosinus sont 2π-périodiques définies continues dérivables sur $\mathbb{R}$, à valeurs dans $[-1;1]$, et on a $\sin' = \cos$ et $\cos' = -\sin$.
- La fonction tangente est π-périodique, définie continue dérivable sur $\mathbb{R} \setminus \{\frac{\pi}{2} + k\pi \mid k \in \mathbb{Z}\}$ et on a $\tan'(x) = 1 + \tan^2(x) = \frac{1}{\cos^2(x)}$.
- Les fonctions sinus et tangente sont impaires alors que la fonction cosinus est paire.

À retenir

Si u est une fonction dérivable alors $\sin(u)$ et $\cos(u)$ sont dérivables avec les formules :
$$[\sin(u)]' = u'\cos(u) \text{ et } [\cos(u)]' = -u'\sin(u)$$
Si de plus la fonction $\cos(u)$ ne s'annule pas, alors la fonction $\tan(u)$ est dérivable et :
$$[\tan(u)]' = u'(1 + \tan^2(u)) = \frac{u'}{\cos^2(u)}$$



- On a les relations $\sin(\pi + x) = -\sin(x)$ et $\cos(\pi + x) = -\cos(x)$.
- On a les valeurs remarquables :

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

  comme $\sin(\pi - x) = \sin(x)$ et $\cos(\pi - x) = -\cos(x)$, on peut compléter le tableau avec les valeurs $\frac{2\pi}{3}$, $\frac{3\pi}{4}$, $\frac{5\pi}{6}$ et π, la parité permet ensuite d'avoir un tableau de $-\pi$ à π.
- Formules d'addition : $\forall x, y \in \mathbb{R}$ on a :
  - $\cos(x + y) = \cos(x)\cos(y) - \sin(x)\sin(y)$. En particulier, $\cos(2x) = 2\cos^2(x) - 1 = 1 - 2\sin^2(x)$.

- $\sin(x+y) = \sin(x)\cos(y) + \cos(x)\sin(y)$. En particulier, $\sin(2x) = 2\sin(x)\cos(x)$.
- $\tan(x+y) = \frac{\tan(x)+\tan(y)}{1-\tan(x)\tan(y)}$. En particulier, $\tan(2x) = \frac{2\tan(x)}{1-\tan^2(x)}$.
- En posant $u = \tan(\frac{x}{2})$, on a $\sin(x) = \frac{2u}{1+u^2}$ et $\cos(x) = \frac{1-u^2}{1+u^2}$.

★**Exercice** 5.1

1/ Montrer que $\forall x \in \mathbb{R}, |\sin(x)| \leqslant |x|, 0 \leqslant 1-\cos(x) \leqslant \frac{x^2}{2}$.

2/ Montrer que $\forall x \in]-\frac{\pi}{2};\frac{\pi}{2}[, |\tan(x)| \geqslant |x|$.

2) Inversion des fonctions circulaires

La fonction arctan

La fonction $f\colon]-\frac{\pi}{2};\frac{\pi}{2}[\to \mathbb{R}$ définie par $f(x) = \tan(x)$, est continue et strictement croissante, elle réalise donc une bijection entre $]-\frac{\pi}{2};\frac{\pi}{2}[$ et $f(]-\frac{\pi}{2};\frac{\pi}{2}[) = \mathbb{R}$. Par définition, la bijection réciproque est $f^{-1} = \arctan$ (lire arc tangente), elle est définie par :

$$\begin{array}{rccl} \arctan\colon & \mathbb{R} & \to &]-\frac{\pi}{2};\frac{\pi}{2}[\\ & x & \mapsto & \arctan(x) = y \text{ tel que } \left\{\begin{array}{l} y \in]-\frac{\pi}{2};\frac{\pi}{2}[\\ \tan(y) = x \end{array}\right. \end{array}.$$

☞**Exemple** : $\arctan(0) = 0$, $\arctan(1) = \frac{\pi}{4}$, $\arctan(\sqrt{3}) = \frac{\pi}{3}$, …

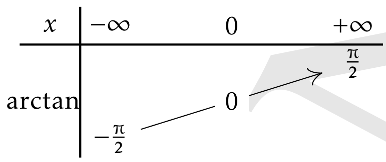
La fonction f étant strictement croissante et continue sur $\mathrm{I} =]-\frac{\pi}{2};\frac{\pi}{2}[$, la fonction $f^{-1} = \arctan$ est strictement croissante et continue sur $\mathbb{R}$; f est dérivable sur I et sa dérivée ne s'annule pas sur I, la réciproque est donc dérivable sur $\mathbb{R}$, et on a la formule suivante :

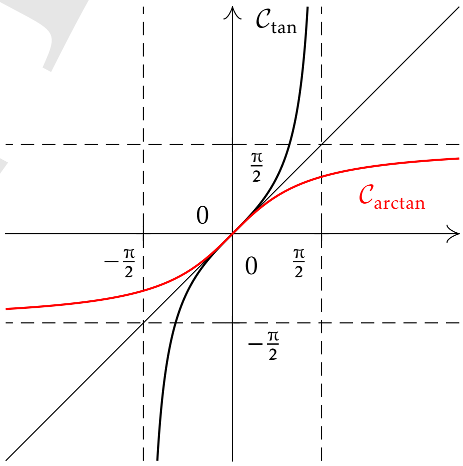
$$\forall x \in \mathbb{R}, \arctan'(x) = \frac{1}{\tan'(\arctan(x))} = \frac{1}{1+\tan^2(\arctan(x))} = \frac{1}{1+x^2}.$$

À retenir

Si u désigne un fonction dérivable, alors la fonction $\arctan(u)$ est dérivable et :

$$[\arctan(u)]' = \frac{u'}{1+u^2}$$





Propriétés :

- $\forall x \in \mathbb{R}, \tan(\arctan(x)) = x$.
- $\forall x \in]-\frac{\pi}{2};\frac{\pi}{2}[, \arctan(\tan(x)) = x$.
- $\forall x \in \mathbb{R}, \arctan(-x) = -\arctan(x)$.
- $\forall x \in \mathbb{R}_+^*, \arctan(x) + \arctan(\frac{1}{x}) = \frac{\pi}{2}$.
- $\forall x \in \mathbb{R}, \arctan(x) = \operatorname{Arg}(1+ix)$.

La fonction arcsin

La fonction $f\colon [-\frac{\pi}{2}\,;\frac{\pi}{2}] \to [-1\,;1]$ définie par $f(x) = \sin(x)$ est continue et strictement croissante sur $\mathrm{I} = [-\frac{\pi}{2}\,;\frac{\pi}{2}]$, elle réalise donc une bijection entre I et $f(\mathrm{I}) = \mathrm{J} = [\sin(-\frac{\pi}{2})\,;\sin(\frac{\pi}{2})] = [-1\,;1]$. La bijection réciproque est notée $f^{-1} = \arcsin$ (lire arcsinus), elle est définie par :

$$\begin{array}{rrcl} \arcsin\colon & [-1;1] & \to & [-\frac{\pi}{2};\frac{\pi}{2}] \\ & x & \mapsto & \arcsin(x) = y \text{ tel que } \left\{\begin{array}{l} y \in [-\frac{\pi}{2};\frac{\pi}{2}] \\ \sin(y) = x \end{array}\right. \end{array}.$$

☞**Exemple** : $\arcsin(-1) = -\frac{\pi}{2}$, $\arcsin(1) = \frac{\pi}{2}$, $\arcsin(0) = 0$, $\arcsin(\frac{1}{2}) = \frac{\pi}{6}$, …

La fonction f étant strictement croissante et continue sur I, la fonction $f^{-1} = \arcsin$ est strictement croissante et continue sur $[-1;1]$; f est dérivable sur I et sa dérivée s'annule en $-\frac{\pi}{2}$ et $\frac{\pi}{2}$, mais pas sur l'intervalle ouvert, la réciproque est donc dérivable sur $]-1;1[$ mais pas en -1 ni en 1 (il y a une tangente verticale en ces points), on a la formule suivante :

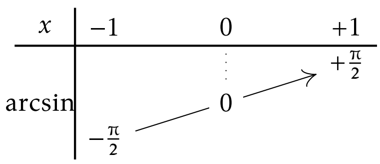
$$\forall x \in]-1;1[, \arcsin'(x) = \frac{1}{\cos(\arcsin(x))} = \frac{1}{\sqrt{1-x^2}}.$$

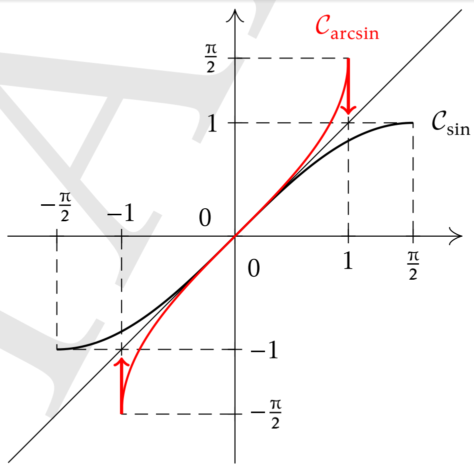
Car : $\cos^2(\arcsin(x)) + \sin^2(\arcsin(x)) = 1$, c'est à dire $\cos^2(\arcsin(x)) + x^2 = 1$ d'où $\cos(\arcsin(x)) = \pm\sqrt{1-x^2}$, mais ce cosinus est positif car $\arcsin(x) \in [-\frac{\pi}{2}\,;\frac{\pi}{2}]$, donc $\cos(\arcsin(x)) = \sqrt{1-x^2}$.

À retenir

Si u est une fonction dérivable à valeurs dans $]-1;1[$ alors la fonction $\arcsin(u)$ est dérivable et :

$$[\arcsin(u)]' = \frac{u'}{\sqrt{1-u^2}}$$





Propriétés :

- $\forall x \in [-1;1], \sin(\arcsin(x)) = x$.
- $\forall x \in [-\frac{\pi}{2};\frac{\pi}{2}], \arcsin(\sin(x)) = x$.
- $\forall x \in [-1;1], \arcsin(-x) = -\arcsin(x)$ [fonction impaire].
- $\forall x \in \mathbb{R}, \arctan(x) = \arcsin\left(\frac{x}{\sqrt{1+x^2}}\right)$.
- $\forall x \in [-1;1], \cos(\arcsin(x)) = \sqrt{1-x^2}$.
- $\forall x \in [-\pi;\pi], \arcsin(\cos(x)) = \frac{\pi}{2} - |x|$.

Attention !

La fonction $f\colon x \mapsto \arcsin(\sin(x))$ n'est pas l'identité, elle est 2π- périodique et impaire, il suffit donc l'étudier sur $[0;\pi]$, mais elle vérifie $f(\pi - x) = f(x)$, la droite $x = \frac{\pi}{2}$ est donc un axe de symétrie et l'étude se réduit à $[0;\frac{\pi}{2}]$, intervalle sur lequel $f(x) = x$.

La fonction arccos

La fonction $f\colon [0;\pi] \to [-1;1]$ définie par $f(x) = \cos(x)$, est continue et strictement décroissante, elle définit donc une bijection entre $[0;\pi]$ et $f([0;\pi]) = [f(\pi); f(0)] = [-1;1]$. Par définition, la bijection réciproque est notée $f^{-1} = \arccos$ (lire arc cosinus), elle est définie par :

$$\begin{array}{rrcl} \arccos\colon & [-1;1] & \to & [0;\pi] \\ & x & \mapsto & \arccos(x) = y \text{ tel que } \left\{\begin{array}{l} y \in [0;\pi] \\ \cos(y) = x \end{array}\right. \end{array}.$$

☞**Exemple** : $\arccos(-1) = \pi$, $\arccos(1) = 0$, $\arccos(0) = \frac{\pi}{2}$, $\arccos(-\frac{1}{2}) = \frac{2\pi}{3}$, …

La fonction f étant strictement décroissante et continue sur $I = [0;\pi]$, la fonction $f^{-1} = \arccos$ est strictement décroissante et continue sur $[-1;1]$; f est dérivable sur I et sa dérivée s'annule en 0 et et π, mais pas sur l'intervalle ouvert, la réciproque est donc dérivable sur $]-1;1[$ mais pas en -1 ni en 1 (il y a une tangente verticale en ces points), on a la formule suivante :

$$\forall x \in]-1;1[, \arccos'(x) = \frac{-1}{\sin(\arccos(x))} = \frac{-1}{\sqrt{1-x^2}}.$$

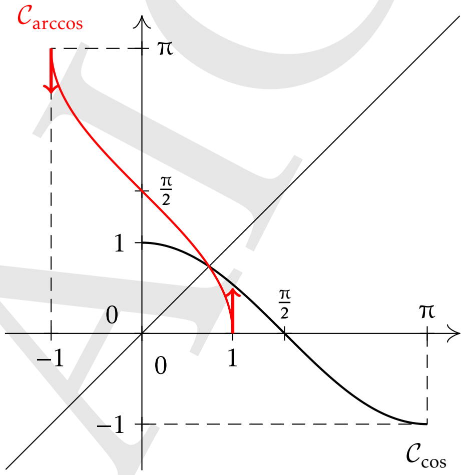
Car : $\sin^2(\arccos(x)) + \cos^2(\arccos(x)) = 1$, c'est à dire $\sin^2(\arccos(x)) + x^2 = 1$ d'où $\sin(\arccos(x)) = \pm\sqrt{1-x^2}$, mais ce sinus est positif car $\arccos(x) \in [0\,;\pi]$, donc $\sin(\arccos(x)) = \sqrt{1-x^2}$.

À retenir

Si u est une fonction dérivable à valeurs dans $]-1;1[$ alors la fonction $\arccos(u)$ est dérivable et :

$$[\arccos(u)]' = \frac{-u'}{\sqrt{1-u^2}}$$

x	-1		0		$+1$
arccos	π	↘	$\frac{\pi}{2}$	↘	0



Propriétés :

- $\forall x \in [-1;1], \cos(\arccos(x)) = x$.
- $\forall x \in [0;\pi], \arccos(\cos(x)) = x$.
- $\forall x \in [-1;1], \sin(\arccos(x)) = \sqrt{1-x^2}$.
- $\forall x \in [-1;1], \arccos(x) + \arcsin(x) = \frac{\pi}{2}$.
- $\forall x \in [-1;1], \arccos(-x) = \pi - \arccos(x)$.

Attention !

La fonction $f : x \mapsto \arccos(\cos(x))$ *n'est pas l'identité, elle est* 2π*- périodique et paire, il suffit donc l'étudier sur* $[0;\pi]$ *intervalle sur lequel* $f(x) = x$.

III FONCTIONS PUISSANCES

Les puissances entières sont supposées connues.

1) Puissance quelconque

Définition 5.3

Si α est un réel et si $x > 0$ alors on pose $x^\alpha = e^{\alpha\ln(x)}$.

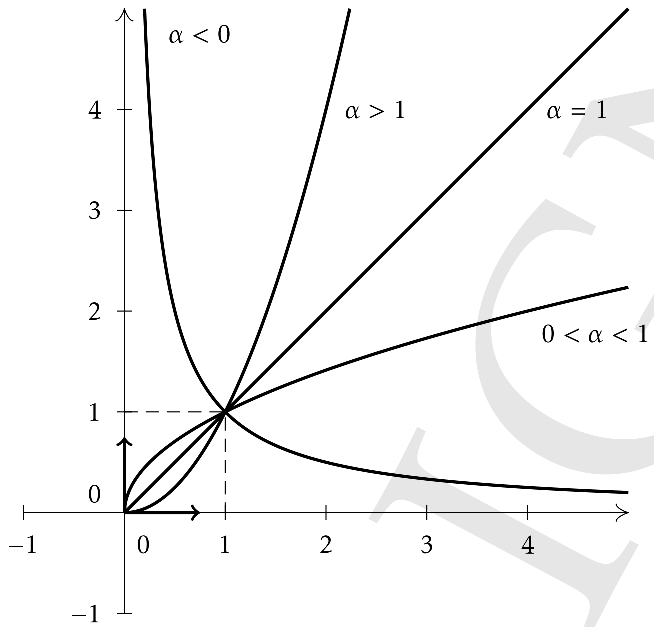
Cela définit une fonction f_α continue et dérivable sur $]0\,;+\infty[$ avec la formule : $\frac{d}{dx}[x^\alpha] = \alpha x^{\alpha-1}$.

Il en découle que si u est une fonction dérivable à valeurs strictement positives, alors la fonction u^α est dérivable et :

$$\boxed{(u^\alpha)' = \alpha \times u' \times u^{\alpha-1}}$$

On a $\lim\limits_{x\to 0} f_\alpha(x) = \begin{cases} 0 & \text{si } \alpha > 0 \\ +\infty & \text{si } \alpha < 0 \end{cases}$. Dans le premier cas on pose $0^\alpha = 0$, dans le second cas il y a une asymptote verticale.

Lorsque $\alpha > 0$: $\frac{x^\alpha - 0}{x} = e^{(\alpha-1)\ln(x)} \xrightarrow[x\to 0]{} \begin{cases} 0 & \text{si } \alpha > 1 \\ +\infty & \text{si } 0 < \alpha < 1 \end{cases}$, lorsque $\alpha > 1$ on a une tangente horizontale et lorsque $\alpha < 1$ on a une tangente verticale.



Cas particuliers (avec $x > 0$) :

a) Lorsque $\alpha = n \in \mathbb{Z}$, on retrouve bien les puissances entières car $\exp(n\ln(x)) = (\exp(\ln(x)))^n = x^n$.

b) Lorsque $\alpha = \frac{1}{n}$ avec $n \in \mathbb{N}^*$: soit $y = x^\alpha$, on a $y^n = \exp(\frac{n}{n}\ln(x)) = x$, comme y est positif, on dit que y est la racine n^e de x. Notation pour $x > 0$: $x^{1/n} = \sqrt[n]{x}$.

c) Lorsque $\alpha = \frac{p}{q} \in \mathbb{Q}$ avec $p \in \mathbb{Z}$ et $q \in \mathbb{N}^*$: soit $y = x^\alpha$, on a $y^q = \exp(q\frac{p}{q}\ln(x)) = x^p$, comme y est positif, y est la racine q^e de x^p. Autrement dit, pour $x > 0$, $x^{p/q} = \sqrt[q]{x^p}$.

Théorème 5.6 (Propriétés)

Avec $x,\ y > 0$ et $\alpha,\ \beta \in \mathbb{R}$:

- $\ln(x^\alpha) = \alpha\ln(x)$.
- $x^\alpha \times x^\beta = x^{\alpha+\beta}$, *et donc* $x^{-\alpha} = \frac{1}{x^\alpha}$, *et* $\frac{x^\alpha}{x^\beta} = x^{\alpha-\beta}$.
- $(x^\alpha)^\beta = x^{\alpha\beta}$.
- $(xy)^\alpha = x^\alpha \times y^\alpha$.
- *Pour α non nul,* $y = x^\alpha \iff x = y^{\frac{1}{\alpha}}$.

Preuve : Celle-ci est simple et laissée en exercice. □

★**Exercice** 5.2

1/ *Soient u et v deux fonctions dérivables avec $u > 0$, calculer la dérivée de la fonction $x \mapsto u(x)^{v(x)}$.*

2/ *Calculer* $\lim\limits_{x\to+\infty} \left(1 + \frac{1}{x}\right)^x$.

2) Croissance comparée de ces fonctions

Comparaison des puissances : si $\alpha < \beta$ alors x^α est négligeable devant x^β au voisinage de $+\infty$ et x^β est négligeable devant x^α au voisinage de 0, c'est à dire :

$$\boxed{\lim_{x\to+\infty} \frac{x^\alpha}{x^\beta} = 0 \text{ et } \lim_{x\to 0^+} \frac{x^\beta}{x^\alpha} = 0}$$

Comparaison des puissances et des logarithmes : si α et β sont des réels **strictement positifs**, alors $[\ln(x)]^\alpha$ est négligeable devant x^β au voisinage de $+\infty$ et $|\ln(x)|^\alpha$ est négligeable devant $\frac{1}{x^\beta}$ au voisinage de 0, c'est à dire :

$$\boxed{\lim_{x\to+\infty}\frac{[\ln(x)]^\alpha}{x^\beta}=0 \text{ et } \lim_{x\to 0^+} x^\beta|\ln(x)|^\alpha=0}$$

Preuve : $\frac{[\ln(x)]^\alpha}{x^\beta}=\left(\frac{\frac{\alpha}{\beta}\ln(u)}{u}\right)^\alpha=\left(\frac{\alpha}{\beta}\right)^\alpha\left(\frac{\ln(u)}{u}\right)^\alpha$ avec $u=x^{\frac{\beta}{\alpha}}$, ce qui donne la première limite. La deuxième en découle avec le changement de variable $u=\frac{1}{x}$. □

Comparaison des puissances et des exponentielles : si α est un réel et si $\beta>0$, alors x^α est négligeable devant $e^{\beta x}$ au voisinage de $+\infty$, c'est à dire :

$$\boxed{\lim_{x\to+\infty} x^\alpha e^{-\beta x}=0}.$$

Preuve : Lorsque $\alpha\leqslant 0$ il n'y a rien à démontrer. Lorsque $\alpha>0$, $u=e^x\xrightarrow[x\to+\infty]{}+\infty$ et on a $x^\alpha e^{-\beta x}=\frac{[\ln(u)]^\alpha}{u^\beta}\xrightarrow[u\to+\infty]{}0$.
□

★**Exercice 5.3** *Comparer x^α et e^{x^β} au voisinage de $+\infty$ avec α et β strictement positifs.*

IV FONCTIONS HYPERBOLIQUES

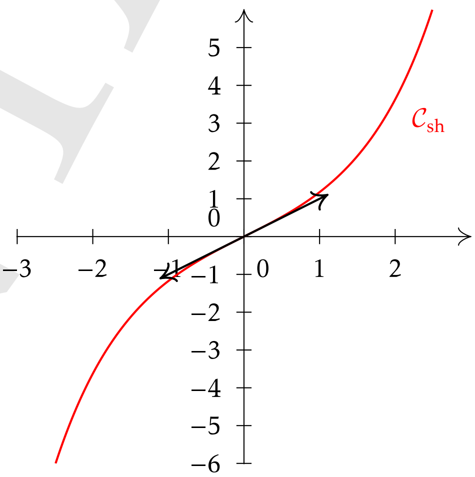
1) Définition

Définition 5.4

Pour $x\in\mathbb{R}$, on pose $\operatorname{ch}(x)=\frac{e^x+e^{-x}}{2}$ *[cosinus hyperbolique],* $\operatorname{sh}(x)=\frac{e^x-e^{-x}}{2}$ *[sinus hyperbolique] et* $\operatorname{th}(x)=\frac{\operatorname{sh}(x)}{\operatorname{ch}(x)}=\frac{e^x-e^{-x}}{e^x+e^{-x}}$ *[tangente hyperbolique].*

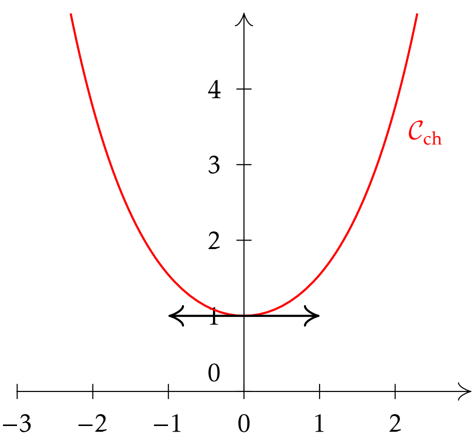
Le sinus hyperbolique : la fonction sh est impaire, définie continue dérivable sur $\mathbb{R}$ et $\operatorname{sh}'(x)=\operatorname{ch}(x)$, on en déduit le tableau de variation et la courbe :

x	$-\infty$		0		$+\infty$
sh	$-\infty$	↗	0	↗	$+\infty$



Le cosinus hyperbolique : la fonction ch est paire, définie continue dérivable sur $\mathbb{R}$ et $\operatorname{ch}'(x)=\operatorname{sh}(x)$, on en déduit le tableau de variation et la courbe :

x	$-\infty$		0		$+\infty$
ch	$+\infty$	↘	1	↗	$+\infty$



Quelques propriétés :

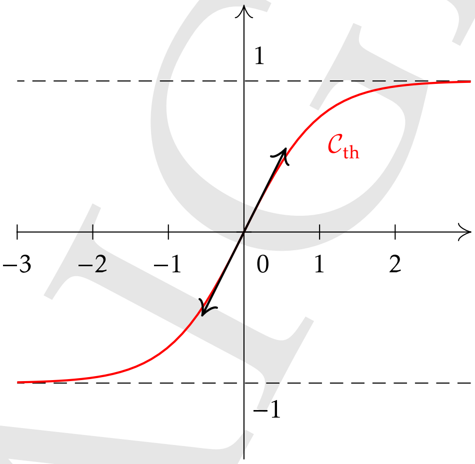
- $\forall x \in \mathbb{R}, \operatorname{ch}(x) \geqslant 1$.
- $\lim\limits_{x\to+\infty} \frac{\operatorname{ch}(x)}{x} = +\infty$ et $\lim\limits_{x\to+\infty} \frac{\operatorname{ch}(x)}{e^x} = \frac{1}{2}$.
- $\forall x \in \mathbb{R}, \operatorname{ch}(x) + \operatorname{sh}(x) = e^x$ et $\operatorname{ch}(x) - \operatorname{sh}(x) = e^{-x}$.
- $\forall x > 0, x < \operatorname{sh}(x) < \operatorname{ch}(x)$.
- $\lim\limits_{x\to+\infty} \frac{\operatorname{sh}(x)}{x} = +\infty$ et $\lim\limits_{x\to+\infty} \frac{\operatorname{sh}(x)}{e^x} = \frac{1}{2}$.

La tangente hyperbolique : la fonction th est impaire, définie continue dérivable sur $\mathbb{R}$ et

$$\operatorname{th}'(x) = \frac{\operatorname{ch}^2(x) - \operatorname{sh}^2(x)}{\operatorname{ch}^2(x)} = 1 - \operatorname{th}^2(x) = \frac{1}{\operatorname{ch}^2(x)}$$

d'où les variations et la courbe :

x	$-\infty$	0	$+\infty$
th	-1 $\nearrow$	0	$\nearrow$ $+1$



Quelques propriétés :

- $\forall x \in \mathbb{R}, -1 < \operatorname{th}(x) < 1$.
- $\forall x > 0, \operatorname{th}(x) < x$.

2) Trigonométrie hyperbolique

- $\forall x \in \mathbb{R}, \operatorname{ch}^2(x) - \operatorname{sh}^2(x) = 1$.
- Formules d'addition : $\forall x, y \in \mathbb{R}$ on a :
  - $\operatorname{ch}(x+y) = \operatorname{ch}(x)\operatorname{ch}(y) + \operatorname{sh}(x)\operatorname{sh}(y)$.
  - $\operatorname{sh}(x+y) = \operatorname{sh}(x)\operatorname{ch}(y) + \operatorname{ch}(x)\operatorname{sh}(y)$.
  - $\operatorname{th}(x+y) = \frac{\operatorname{th}(x)+\operatorname{th}(y)}{1+\operatorname{th}(x)\operatorname{th}(y)}$.

  En particulier : $\boxed{\begin{array}{l} \operatorname{ch}(2x) = 2\operatorname{ch}^2(x) - 1 = 1 + 2\operatorname{sh}^2(x) \\ \operatorname{sh}(2x) = 2\operatorname{sh}(x)\operatorname{ch}(x) \\ \operatorname{th}(2x) = \frac{2\operatorname{th}(x)}{1+\operatorname{th}^2(x)} \end{array}}$.
- Transformations de somme en produit : $\forall x, y \in \mathbb{R}$, en posant $p = \frac{x+y}{2}$ et $q = \frac{x-y}{2}$, on a $x = p+q$ et $y = p-q$, on obtient :
  - $\operatorname{ch}(x) + \operatorname{ch}(y) = 2\operatorname{ch}(\frac{x+y}{2})\operatorname{ch}(\frac{x-y}{2})$.
  - $\operatorname{ch}(x) - \operatorname{ch}(y) = 2\operatorname{sh}(\frac{x+y}{2})\operatorname{sh}(\frac{x-y}{2})$.
  - $\operatorname{sh}(x) + \operatorname{sh}(y) = 2\operatorname{sh}(\frac{x+y}{2})\operatorname{ch}(\frac{x-y}{2})$.
  - $\operatorname{th}(x) + \operatorname{th}(y) = \frac{\operatorname{sh}(x+y)}{\operatorname{ch}(x)\operatorname{ch}(y)}$.

V SOLUTION DES EXERCICES

Solution 5.1

1/ *Il suffit de le démontrer pour x positif en étudiant la fonction $x \mapsto x - \sin(x)$, puis on intègre de 0 à x ce qui donne la deuxième inégalité.*

2/ *On étudie $x \mapsto x - \tan(x)$ sur $[0\,;\frac{\pi}{2}[$. Pour $x \in]-\frac{\pi}{2}\,;0]$, on applique l'inégalité précédente à $-x$.*

Solution 5.2

1/ *Comme l'exposant varie, on passe à la forme exponentielle :* $u^v = e^{v\ln(u)}$*, d'où la dérivée :* $[u^v]' = [v\ln(u)]'e^{v\ln(u)} = v'\ln(u)u^v + v\frac{u'}{u}u^v = v'\ln(u)u^v + vu'u^{v-1}$.

2/ *Comme l'exposant varie, on passe à la forme exponentielle :* $\left(1+\frac{1}{x}\right)^x = e^{x\ln(1+\frac{1}{x})} = e^{\frac{\ln(1+X)}{X}}$ *avec* $X = \frac{1}{x} \xrightarrow[x\to+\infty]{} 0$*, or on sait que* $\lim\limits_{X\to 0}\frac{\ln(1+X)}{X} = 1$*, et donc la limite cherchée vaut* $e^1 = e$ *(par continuité de l'exponentielle).*

Solution 5.3 *Il faut étudier la limite du quotient, c'est à dire la limite en* $+\infty$ *de* $x^\alpha e^{-x^\beta}$.

On pose $X = x^\beta \xrightarrow[x\to+\infty]{} +\infty$*, on a alors* $x = X^{1/\beta}$ *et* $x^\alpha e^{-x^\beta} = X^{\alpha/\beta}e^{-X} \xrightarrow[X\to+\infty]{} 0$ *d'après les croissances comparées.*

Chapitre 6

Équations différentielles

Sommaire

I FONCTIONS À VALEURS COMPLEXES (ET VARIABLE RÉELLE)

1) Définition

Soit I un intervalle de $\mathbb{R}$, et $f\colon \mathrm{I} \to \mathbb{C}$ une fonction à valeurs complexes.
Pour $t \in \mathrm{I}$, on pose $u(t) = \mathrm{Re}(f(t))$ et $v(t) = \mathrm{Im}(f(t))$, on définit ainsi deux fonctions u et v à **valeurs réelles** telles que $\forall\, t \in \mathrm{I}$, $f(t) = u(t) + iv(t)$.

Définition 6.1
- *La fonction u est appelée* ***partie réelle*** *de f et la fonction v est appelée* ***partie imaginaire*** *de f.*
- *La fonction $\overline{f} = u - iv$ est appelée fonction conjuguée de f.*
- *La fonction $|f| = \sqrt{u^2 + v^2}$ est appelée fonction module de f.*
- *f est bornée si et seulement si $\exists \mathrm{M} \in \mathbb{R}, \forall t \in \mathrm{I}, |f(t)| \leqslant \mathrm{M}$ (i.e. les images sont dans le disque de centre l'origine et de rayon M).*

☞**Exemples** :
- Soit f définie sur $\mathbb{R}$ par $f(t) = e^{it}$, on a $\mathrm{Re}(f) = \cos$ et $\mathrm{Im}(f) = \sin$, elle est bornée.
- Soit f définie sur $\mathbb{R}$ par $f(t) = \frac{t^2}{1+it}$, on a $\mathrm{Re}(f)\colon t \mapsto \frac{t^2}{1+t^2}$ et $\mathrm{Im}(f)\colon t \mapsto \frac{-t^3}{1+t^2}$, elle n'est pas bornée.

2) Continuité, dérivation

Définition 6.2 (Continuité)

*• Soit $f : \mathrm{I} \to \mathbb{C}$ une fonction, soit u sa partie réelle et v sa partie imaginaire, on dit que f est continue en $t_0 \in \mathrm{I}$ lorsque **les fonctions u et v sont continues en** t_0.*
• On dit que f est continue sur I si elle est continue en tout point de I et l'ensemble des fonctions continues sur I est noté $\mathcal{C}^0(\mathrm{I}, \mathbb{C})$.

À retenir

• Les théorèmes généraux sont les mêmes que pour les fonctions continues à valeurs réelles.
• Le théorème des valeurs intermédiaires ne s'applique pas pour les fonctions continues complexes.
• Si f est continue sur le segment $[a; b]$, alors $|f|$ possède un maximum et un minimum.

Définition 6.3

*Soit $f : \mathrm{I} \to \mathbb{C}$ une fonction, soit u sa partie réelle et v sa partie imaginaire, on dit que f est dérivable en t_0 lorsque les fonctions u **et** v sont dérivables en t_0. Si c'est le cas, alors on pose $f'(t_0) = u'(t_0) + iv'(t_0)$. On remarquera que si f est dérivable sur I, alors $\mathrm{Re}(f') = (\mathrm{Re}(f))'$ et $\mathrm{Im}(f') = (\mathrm{Im}(f))'$.*

Les résultats suivants sont identiques au cas réel :

À retenir

- Si f est dérivable en t_0 alors f est continue en t_0 (réciproque fausse).
- Une fonction dérivable f sur un intervalle I est constante si et seulement si sa dérivée est nulle sur I.
- La somme, le produit et la composée de deux fonctions dérivables sont dérivables. Si f est dérivable et ne s'annule pas alors $\frac{1}{f}$ est dérivable (théorèmes généraux de la continuité).
- Formules de dérivation pour les opérations algébriques ($\lambda \in \mathbb{C}$) :

$$(f+g)' = f' + g' ; (\lambda f)' = \lambda f' ; (fg)' = f'g + fg' ; \left(\frac{1}{g}\right)' = \frac{-g'}{g^2}$$

- Dérivation d'une composée : si $f : \mathrm{I} \to \mathbb{R}$ est dérivable sur I, si $g : \mathrm{J} \to \mathbb{C}$ est dérivable sur J et si $f(\mathrm{I}) \subset \mathrm{J}$, alors $g \circ f$ est dérivable sur I et :

$$(g \circ f)' = f' \times g' \circ f \text{ (ou encore } [g(f)]' = f' \times g'(f)).$$

☞**Exemple** : La fonction f définie par $f(t) = \frac{1}{1+it}$ est dérivable sur $\mathbb{R}$ et $f'(t) = \frac{-i}{(1+it)^2}$.

Remarque 6.1 :

– *On remarquera que les formules de dérivation sont les mêmes pour les fonctions à valeurs complexes que pour les fonctions à valeurs réelles.*
– *Si f et g sont dérivables et que g ne s'annule pas, alors $\frac{f}{g}$ est dérivable et $\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}$.*

Théorème 6.1

Soit $f : \mathrm{I} \mapsto \mathbb{C}$ une fonction dérivable, alors la fonction $t \to e^{f(t)}$ est dérivable sur I (exponentielle complexe de $f(t)$) et :

$$\left(e^f\right)' = f' \times e^f$$

Preuve : On pose $f(t) = a(t) + ib(t)$ sous forme algébrique. $e^{f(t)} = e^{a(t)} \times [\cos(b(t)) + i\sin(b(t))]$, la partie réelle est donc $g(t) = e^{a(t)}\cos(b(t))$ et sa partie imaginaire est $h(t) = e^{a(t)}\sin(b(t))$. Ces fonctions sont dérivables sur I, donc e^f est dérivable sur I et sa dérivée est $g'(t) + ih'(t)$, il suffit alors de comparer $g'(t) + ih'(t)$ avec $f'(t)e^{f(t)}$ pour constater l'égalité. □

3) Primitives, intégrales

Définition 6.4

Soit F, f : I → ℂ *deux fonctions, on dit que* F *est une* ***primitive*** *de* f *sur* I *lorsque* F *est dérivable sur* I *et* $F' = f$.

★**Exercice 6.1** *Calculer une primitive de :* $h(t) = e^{\alpha t}$ *où* $\alpha \in \mathbb{C}^*$; $f(t) = \frac{1}{1+it}$ *et de* $g(t) = e^t \cos(t)$.

Les propriétés des primitives citées dans le cas réel restent valables. La définition de l'intégrale à partir d'une primitive est également la même et on retrouve les propriétés qui ne font pas intervenir la notion de signe : linéarité, relation de Chasles, majoration en module. En particulier la linéarité permet d'écrire que si $f = u + iv$ (forme algébrique de f), alors $\int_a^b f(t)\,\mathrm{d}t = \int_a^b u(t)\,\mathrm{d}t + i\int_a^b v(t)\,\mathrm{d}t$.

★**Exercice 6.2** *Montrer que* $\left|\int_0^{\pi} e^{(1+i)t}\,\mathrm{d}t\right| \leqslant e^{\pi} - 1$.

Les deux outils fondamentaux pour le calcul d'intégrales sont encore valables : le théorème de l'intégration par parties et le théorème du changement de variable.

Dans la suite de ce chapitre, la lettre 𝕂 désigne l'ensemble ℝ ou bien l'ensemble ℂ ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$), et les fonctions considérées sont définies sur un intervalle I de ℝ.

II ÉQUATIONS DIFFÉRENTIELLES LINÉAIRES DU PREMIER ORDRE

Une équation différentielle est une équation dont l'inconnue est une fonction y : I → 𝕂 intervenant sous forme dérivée (première ou supérieure). On rencontre ce genre d'équations en mécanique (lois de Newton), en électricité (circuits RLC), ... etc.

1) Définitions

Définition 6.5

Une équation différentielle scalaire linéaire d'ordre 1 est une équation différentielle de la forme :

$$(\mathrm{E}) : \forall t \in \mathrm{I},\ a(t)y'(t) + b(t)y(t) = c(t), \text{ notée plus simplement : } a(t)y' + b(t)y = c(t).$$

où a, b, c : I → 𝕂 *sont trois fonctions définies continues sur un intervalle* I *de* ℝ *et* y : I → 𝕂 *une fonction dérivable* ***inconnue****. On suppose de plus que la fonction* a *n'est pas la fonction nulle. On appelle* ***équation homogène associée*** *à* (E) *l'équation différentielle :*

$$(\mathrm{H}) : \forall t \in \mathrm{I},\ a(t)y' + b(t)y = 0.$$

La fonction c *est souvent appelée* ***second membre*** *de l'équation* (E).

Dans la pratique on a souvent en plus une condition sur la fonction inconnue y du type : $y(t_0) = \alpha$ où t_0 et α sont des données. Cette condition est appelée **condition initiale**, et on appelle **problème de** Cauchy [1] le système :

$$\begin{cases} \forall t \in \mathrm{I},\ a(t)y' + b(t)y = c(t) \\ y(t_0) = \alpha \end{cases}.$$

☞**Exemple** : L'équation différentielle : $y' - y = 0$ avec $y(0) = 1$ est utilisée en terminale pour introduire l'exponentielle.

2) Étude de l'équation homogène

Théorème 6.2

Soit $\mathrm{S_I(H)}$ *l'ensemble des solutions sur* I *de l'équation homogène* (H), *alors on a les propriétés :*

– $0 \in \mathrm{S_I(H)}$ *(la fonction nulle est dans* $\mathrm{S_I(H)}$*).*
– $\forall\, f, g \in \mathrm{S_I(H)}, f + g \in \mathrm{S_I(H)}$.
– $\forall\, \alpha \in \mathbb{K}, \forall\, f \in \mathrm{S_I(H)}, \alpha f \in \mathrm{S_I(H)}$.

Preuve : Celle-ci est simple et laissée en exercice. □

1. CAUCHY Augustin-Louis (1789 – 1857) : un des plus grands mathématiciens français.

Résolution de (H)

On se place sur un intervalle I où la fonction a **ne s'annule pas**, on a alors $\forall\, t \in \mathrm{I}, y' = -\frac{b}{a}y$. Soit F une primitive de la fonction $-\frac{b}{a}$ sur I, on a alors :

$$y \in \mathrm{S_I(H)} \iff y' = \mathrm{F}'y \iff y'e^{-\mathrm{F}} - \mathrm{F}'e^{-\mathrm{F}}y = 0 \iff \frac{d}{dt}\left(ye^{-\mathrm{F}}\right) = 0 \iff \exists\, \lambda \in \mathbb{K}, \forall\, t \in \mathrm{I}, y(t) = \lambda e^{\mathrm{F}(t)}.$$

On peut donc énoncer :

Théorème 6.3

Lorsque la fonction a ne s'annule pas sur l'intervalle I *alors les solutions de* (H) *sont les fonctions :*

$$y \colon t \mapsto \lambda e^{\mathrm{F}(t)},$$

où F *désigne une primitive de la fonction $-\frac{b}{a}$ sur* I, *et λ un élément quelconque de $\mathbb{K}$.*

À retenir

Si la fonction a ne s'annule pas sur I :
- Le problème de Cauchy pour l'équation (H) a une unique solution. Car la condition initiale détermine complètement la constante λ.
- L'unique solution sur I qui s'annule en un point donné est la fonction nulle. Par conséquent toutes les autres solutions ne s'annulent jamais sur I, lorsque $\mathbb{K} = \mathbb{R}$ elles ont toutes un signe constant (car elles sont continues).

☞**Exemples :**

– $y' + \omega y = 0$ où $\omega \in \mathbb{K}$ est une constante : les solutions sont les fonctions définies sur $\mathbb{R}$ par $y(t) = \lambda e^{-\omega t}$ avec $\lambda \in \mathbb{K}$ quelconque.

– $ty' = y$ sur $\mathbb{R}$: on se place d'abord sur $\mathrm{I} =]0; +\infty[$, sur cet intervalle on a $y' = \frac{1}{t}y$ d'où $y(t) = \alpha t$ ($\alpha \in \mathbb{K}$ quelconque). Puis on se place sur $\mathrm{J} =]-\infty; 0[$, sur cet intervalle on a encore $y' = \frac{1}{t}y$ d'où $y(t) = \lambda|t| = \beta t$ ($\beta = -\lambda \in \mathbb{K}$ quelconque). Soit maintenant y une solution sur $\mathbb{R}$, alors y est en particulier solution sur I donc il existe α tel que $\forall\, t > 0, y(t) = \alpha t$, de même y est solution sur J, donc il existe β tel que $\forall t < 0, y(t) = \beta t$, mais y doit être dérivable en 0, ce qui entraîne $\alpha = \beta$, finalement $\forall t \in \mathbb{R}, y(t) = \alpha t$. On vérifie pour terminer que cette fonction est bien solution.

3) Étude de l'équation avec second membre

On revient au cas général : (E) : $a(t)y' + b(t)y = c(t)$.

Théorème 6.4 (structure des solutions)

Si l'ensemble des solutions de (E) *n'est pas vide, et si y_1 est une solution de* (E), *alors les solutions de* (E) *sont les fonctions s'écrivant comme somme de y_1 avec une solution de* (H), *c'est à dire les fonctions de la forme : $y \colon t \mapsto y_1(t) + y_{\mathrm{H}}(t)$ avec y_{H} solution quelconque de* (H).

Preuve : Soit y une solution de (E), posons $f = y - y_1$, alors $af' + bf = ay' - ay_1' + by - by_1 = c - c = 0$ donc $f \in \mathrm{S_I(H)}$. Réciproquement, soit $f \in \mathrm{S_I(H)}$ et soit $y = y_1 + f$, alors $ay' + by = ay_1' + af + by_1 + bf = 0 + c = c$, donc $y \in \mathrm{S_I(E)}$. □

Pour déterminer toutes les solutions de (E) on est donc ramené à résoudre l'équation homogène puis à trouver une solution particulière de (E).

Recherche d'une solution particulière : on se place de nouveau sur un intervalle I **où la fonction** a **ne s'annule pas et on applique la méthode de la variation de la constante** :

Soit F une primitive de $-\frac{b}{a}$ sur I, on cherche une solution particulière sous la forme $y = \lambda e^{\mathrm{F}}$ où λ est **une fonction dérivable sur** I. La fonction y est solution de (E) si et seulement si $a[\lambda' e^{\mathrm{F}} + \lambda \mathrm{F}' e^{\mathrm{F}}] + b\lambda e^{\mathrm{F}} = c$, ce qui équivaut à $a\lambda' e^{\mathrm{F}} + \lambda[a\mathrm{F}'e^{\mathrm{F}} + be^{\mathrm{F}}] = c$ ou encore $\lambda' = \frac{c}{a}e^{-\mathrm{F}}$, car e^{F} est solution de (H). Donc λ doit être une primitive de la fonction $\frac{c}{a}e^{-\mathrm{F}}$, celle-ci est continue sur l'intervalle I, elle admet donc des primitives sur cet I, ce qui prouve l'existence de λ. Une solution de (E) est donc :

$$y_1 = \lambda e^{\mathrm{F}} \text{ avec } \lambda(t) = \int_{t_0}^{t} \frac{c(s)}{a(s)} e^{-\mathrm{F}(s)}\, ds \text{ et } \mathrm{F}(t) = \int_{t_0}^{t} -\frac{b(s)}{a(s)}\, ds,$$

et les solutions de (E) sont les fonctions :

$$y = y_1 + \alpha e^{\mathrm{F}} \text{ avec } \alpha \in \mathbb{K} \text{ quelconque [et } y(t_0) = \alpha\text{]}.$$

À retenir

Lorsque la fonction a ne s'annule pas sur l'intervalle I, le problème de Cauchy a une unique solution.

Exemples :

- $ty' + y = \sin(t)$ sur $\mathbb{R}$: sur l'intervalle $\mathrm{I} =]0\,;+\infty[$ les solutions de (H) sont les fonctions $y = \frac{\lambda}{t}$ avec $\lambda \in \mathbb{K}$ quelconque. On cherche une solution particulière de la forme $y = \frac{\lambda}{t}$ avec λ dérivable sur I ce qui donne $\lambda' = \sin(t)$, une solution particulière est donc $y_1 = -\frac{\cos(t)}{t}$, et les solutions de (E) sur I sont les fonctions $y = \frac{\lambda - \cos(t)}{t}$ avec $\lambda \in \mathbb{K}$ quelconque. On se place ensuite sur l'intervalle $\mathrm{J} =]-\infty;0[$ où le raisonnement est le même. On vérifie ensuite que la seule solution sur $\mathbb{R}$ est la fonction :
$$y : t \mapsto \frac{1-\cos(t)}{t} \text{ avec } y(0) = 0 \text{ et } y'(0) = \frac{1}{2}.$$
- $\cos(t)y' - \sin(t)y = t^2$ sur $\mathrm{I} =]-\frac{\pi}{2};\frac{\pi}{2}[$: les solutions de l'équation homogène sont les fonctions $y = \frac{\lambda}{\cos(t)}$ avec $\lambda \in \mathbb{K}$ quelconque. On cherche une solution particulière sous la forme $y = \frac{\lambda}{\cos(t)}$ avec λ dérivable sur I, ce qui donne $\lambda' = t^2$. On peut donc prendre comme solution particulière $y_1 = \frac{t^3}{3\cos(t)}$, et les solutions de (E) sont les fonctions :
$$y : t \mapsto \frac{\lambda + t^3}{3\cos(t)} \text{ avec } \lambda \in \mathbb{K}.$$

Remarque 6.2 – *Résoudre de telles équations différentielles revient donc à calculer des intégrales, d'où les expressions que l'on rencontre parfois comme : « intégrer une équation différentielle », ou « solution intégrale d'une équation différentielle ».*

★**Exercice 6.3** *Résoudre sur $\mathbb{R}$ l'équation $x^3y' - (3x^2+2)y = x^3$.*

III ÉQUATIONS DIFFÉRENTIELLES LINÉAIRES DU SECOND ORDRE

On s'intéressera uniquement au cas où les coefficients sont des constantes, c'est à dire aux équations différentielles de la forme : $ay'' + by' + cy = f$ où $a, b, c \in \mathbb{K}$, $a \neq 0$, et $f : \mathrm{I} \to \mathbb{K}$ une fonction continue (second membre). L'équation homogène associée est (H) : $ay'' + by' + cy = 0$.

Pour de telles équations, le problème de Cauchy est : $\begin{cases} ay'' + by' + cy = f \\ y(t_0) = \alpha \\ y'(t_0) = \beta \end{cases}$, où t_0, α et β sont des données.

1) Étude de l'équation homogène

Théorème 6.5

Soit $\mathrm{S_I(H)}$ l'ensemble des solutions sur I de l'équation homogène (H), alors on a les propriétés :
- $0 \in \mathrm{S_I(H)}$ *(fonction nulle).*
- $\forall\, f, g \in \mathrm{S_I(H)}, f + g \in \mathrm{S_I(H)}$.
- $\forall\, \alpha \in \mathbb{K}, \forall\, f \in \mathrm{S_I(H)}, \alpha f \in \mathrm{S_I(H)}$.

Preuve : Celle-ci est simple et laissée en exercice. □

Résolution de (H)

On cherche les solutions de la forme $y = e^{\lambda t}$ avec $\lambda \in \mathbb{K}$, on obtient alors $y \in S_{\mathbb{R}}(H) \iff a\lambda^2 + b\lambda + c = 0$, λ doit donc être solution de l'équation $ax^2 + bx + c = 0$, que l'on appelle **équation caractéristique** de (H). Il faut donc distinguer plusieurs cas :

- $\mathbb{K} = \mathbb{C}$:
  - Si $\Delta = b^2 - 4ac \neq 0$: il y a deux solutions distinctes à l'équation caractéristique : λ_1 et λ_2. On pose $\phi_1(t) = e^{\lambda_1 t}$. Soit y deux fois dérivable, posons $z = \frac{y}{\phi_1}$, on a alors $y = z\phi_1$, en remplaçant dans l'équation on obtient que y est solution de (H) si et seulement si $az'' + (2a\lambda_1 + b)z' + (a\lambda_1^2 + b\lambda_1 + c)z = 0$, c'est à dire $az'' + (2a\lambda_1 + b)z' = 0$ car λ_1 est racine de l'équation caractéristique. En posant $Z = z'$ on a une équation différentielle linéaire homogène en Z, dont les solutions sont $Z(t) = z'(t) = \gamma e^{-(2\lambda_1 + b/a)t} = \gamma e^{(\lambda_2-\lambda_1)t}$, ce qui équivaut à $z(t) = \beta e^{(\lambda_2-\lambda_1)t} + \alpha$, et donc $y(t) = \alpha e^{\lambda_1 t} + \beta e^{\lambda_2 t}$ avec $\alpha, \beta \in \mathbb{C}$ des constantes quelconques.
  - Si $\Delta = b^2 - 4ac = 0$: alors il y a une solution double à l'équation caractéristique : λ. Posons $\phi_1(t) = e^{\lambda t}$ et $z = \frac{y}{\phi_1}$ *i.e.* $y = z\phi_1$. Le calcul précédent montre que $y \in S(H) \iff z'' = 0$ c'est à dire il existe $\alpha, \beta \in \mathbb{C}$ tels que $z(t) = \beta t + \alpha$, ce qui donne $y(t) = (\alpha + \beta t)e^{\lambda t}$.
- $\mathbb{K} = \mathbb{R}$ ($a, b, c \in \mathbb{R}$, avec $a \neq 0$) : la démarche est la même, on cherche les solutions de l'équation caractéristique, d'où la discussion :
  - Si $\Delta > 0$: deux racines distinctes λ_1 et λ_2, comme dans le cas complexe, on montre que les solutions de H sont les fonctions $y\colon t \mapsto \alpha e^{\lambda_1 t} + \beta e^{\lambda_2 t}$ avec $\alpha, \beta \in \mathbb{R}$.
  - Si $\Delta = 0$: une racine double λ, comme dans le cas complexe, on montre que les solutions de H sont les fonctions $y\colon t \mapsto (\alpha + \beta t)e^{\lambda t}$ avec $\alpha, \beta \in \mathbb{R}$.
  - Si $\Delta < 0$: deux racines complexes **non réelles** conjuguées $\lambda = r + i\omega$ et $\overline{\lambda}$. Les solutions **complexes** de (H) sont les fonctions $y(t) = \alpha e^{\lambda t} + \beta e^{\overline{\lambda} t}$ avec $\alpha, \beta \in \mathbb{C}$, une telle solution est réelle ssi $y(t) = \overline{y}(t) = \overline{\alpha} e^{\overline{\lambda} t} + \overline{\beta} e^{\lambda t}$, ce qui équivaut à $\overline{\alpha} = \beta$. Les solutions réelles sont donc les fonctions $y(t) = \alpha e^{\lambda t} + \overline{\alpha} e^{\overline{\lambda} t} = 2\,\mathrm{Re}(\alpha e^{\lambda t}) = e^{rt}[u\cos(\omega t) + v\sin(\omega t)]$, avec $u = \mathrm{Re}(\alpha)/2$ et $v = -\mathrm{Im}(\alpha)/2$ réels quelconques (car α est un complexe quelconque). On a encore que les solutions de (H) sont les fonctions $y = u\phi_1 + v\phi_2$ avec $u, v \in \mathbb{R}$ et $\phi_1(t) = e^{rt}\cos(\omega t)$ et $\phi_2(t) = e^{rt}\sin(\omega t)$.

À retenir : solutions de l'équation homogène

Soit $x^2 + ax + b = 0$ l'équation caractéristique et $\Delta = b^2 - 4ac$ son discriminant :

- Si $\mathbb{K} = \mathbb{C}$ (a, b et c sont complexes avec $a \neq 0$) :
  - Si $\Delta \neq 0$, l'équation caractéristique à deux solutions distinctes : λ_1 et λ_2. Les solutions sont les fonctions : $t \mapsto \alpha e^{\lambda_1 t} + \beta e^{\lambda_2 t}$ avec $\alpha, \beta \in \mathbb{C}$.
  - Si $\Delta = 0$, l'équation caractéristique à une solution double : $\lambda_1 = \lambda_2$. Les solutions sont les fonctions : $t \mapsto (\alpha + \beta t)e^{\lambda_1 t}$ avec $\alpha, \beta \in \mathbb{C}$.
- Si $\mathbb{K} = \mathbb{R}$ (a, b et c sont **réels** avec $a \neq 0$) :
  - Si $\Delta > 0$, l'équation caractéristique à deux solutions distinctes : λ_1 et λ_2. Les solutions sont les fonctions : $t \mapsto \alpha e^{\lambda_1 t} + \beta e^{\lambda_2 t}$ avec $\alpha, \beta \in \mathbb{R}$.
  - Si $\Delta = 0$, l'équation caractéristique à une solution double : $\lambda_1 = \lambda_2$. Les solutions sont les fonctions : $t \mapsto (\alpha + \beta t)e^{\lambda_1 t}$ avec $\alpha, \beta \in \mathbb{R}$.
  - Si $\Delta < 0$, l'équation caractéristique possède deux solutions complexes non réelles et conjuguées : λ et $\overline{\lambda}$, en posant $\lambda = r + i\omega$ (forme algébrique), les solutions sont les fonctions :
  $$t \mapsto (\alpha\cos(\omega t) + \beta\sin(\omega t))e^{rt} = A\cos(\omega t + \varphi)e^{rt} \text{ avec } \alpha, \beta, A, \varphi \in \mathbb{R}.$$

Théorème 6.6

Soient $a, b, c \in \mathbb{R}$ avec $a \neq 0$, les solutions réelles de l'équation $ay'' + by' + cy = 0$ sont les parties réelles des solutions complexes.

Preuve : Celle-ci est simple et laissée en exercice. □

2) Étude de l'équation avec second membre

Théorème 6.7 (structure des solutions)

Si $f : \mathrm{I} \to \mathbb{K}$ est une fonction continue, alors l'équation $(\mathrm{E})\ :\ ay'' + by' + cy = f$ admet des solutions sur I. Si y_1 est une solution de (E), alors les solutions de (E) sont les fonctions définies sur I par $y : t \mapsto y_1(t) + y_{\mathrm{H}}(t)$ avec y_{H} solution quelconque de (H). De plus, le problème de Cauchy a une unique solution.

Preuve : L'existence dans le cas général d'une solution particulière est admise. Soit y_1 une solution de (E), soit $g \in \mathrm{S_I(H)}$, il est facile de vérifier que $y_1 + g$ est solution de (E), réciproquement, si g est solution de (E), il est facile de vérifier que $g - y_1$ est solution de (H). Les solutions au problème de Cauchy sont les fonctions de la forme $y = y_1 + \alpha\phi_1 + \beta\phi_2$ vérifiant $y(t_0) = c_1$ et $y'(t_0) = c_2$. Ce qui donne le système :

$$\begin{cases} \alpha\phi_1(t_0) + \beta\phi_2(t_0) & = c_1 - y_1(t_0) \\ \alpha\phi_1'(t_0) + \beta\phi_2'(t_0) & = c_2 - y_1'(t_0) \end{cases}.$$

Lorsque $\phi_1(t) = e^{\lambda_1 t}$ et $\phi_2(t) = e^{\lambda_2 t}$ avec λ_1 et λ_2 les racines distinctes de l'équation caractéristique, le déterminant du système est $\mathrm{D} = (\lambda_2 - \lambda_1)e^{(\lambda_1+\lambda_2)t_0} \neq 0$. Lorsque les deux racines sont confondues, alors $\phi_1(t) = e^{\lambda t}$ et $\phi_2(t) = t\phi_1(t)$, dans ce cas, le déterminant du système est $\mathrm{D} = e^{2\lambda t_0} \neq 0$, dans les deux cas, le système a une unique solution. □

Remarque 6.3 – *Dans le cas réel avec $\Delta < 0$, l'unique solution complexe au problème de Cauchy est une solution réelle.*

Dans la suite on s'intéressera seulement au cas où le second membre est de la forme $f(t) = \alpha e^{\lambda t}$ où α et λ sont des constantes dans $\mathbb{K}$.

Théorème 6.8

L'équation $ay'' + by' + cy = \alpha e^{\lambda t}$ admet une solution particulière de la forme $y_1(t) = \beta e^{\lambda t}$ ($\beta \in \mathbb{K}$) si λ n'est pas racine de l'équation caractéristique, de la forme $y_1(t) = \beta t e^{\lambda t}$ si λ est racine simple de l'équation caractéristique ($\Delta \neq 0$), et de la forme $y_1(t) = \beta t^2 e^{\lambda t}$ si λ est racine double de l'équation caractéristique ($\Delta = 0$),

Preuve : On pose $y(t) = \mathrm{Q}(t)e^{\lambda t}$, y est solution si et seulement si $a\mathrm{Q}''(t) + (2a\lambda + b)\mathrm{Q}'(t) + (a\lambda^2 + b\lambda + c)\mathrm{Q}(t) = \alpha$, d'où la discussion :

- Si λ n'est pas racine de l'équation caractéristique : on voit que la constante $\mathrm{Q}(t) = \frac{\alpha}{a\lambda^2+b\lambda+c}$ convient.
- Si λ est solution simple de l'équation caractéristique ($2a\lambda + b \neq 0$) : alors $\mathrm{Q}(t) = \frac{\alpha t}{2a+\lambda}$.
- Si λ est solution double de l'équation caractéristique ($2a\lambda + b = 0$) : on a $\mathrm{Q}(t) = \frac{\alpha t^2}{2a}$ convient.

□

Théorème 6.9

*• Lorsque le second membre est de la forme $f(t) = \sum_{i=1}^{n} \alpha_i e^{\lambda_i t}$, on cherche une solution particulière y_i à l'équation $ay'' + by' + cy = \alpha_i e^{\lambda_i t}$ pour i allant de 1 à n, la fonction $y = y_1 + \cdots + y_n$ est une solution particulière de $ay'' + by' + cy = f$, c'est le **principe de superposition**.*
• Si a, b, c sont réels et si y_0 est solution de $ay'' + by' + cy = f(t)$ alors $\mathrm{Re}(y_0)$ est solution de $ay'' + by' + cy = \mathrm{Re}(f)$ et $\mathrm{Im}(y_0)$ est solution de $ay'' + by' + cy = \mathrm{Im}(f)$.

Preuve : Celle-ci est simple et laissée en exercice. □

☞**Exemples** :

- $y'' + \omega^2 y = 1$ avec $\omega \in \mathbb{R}^*$, ici $\lambda = 0$, il y a une solution particulière évidente qui est $y_1 : t \mapsto \frac{1}{\omega^2}$
  Les solutions réelles de l'équation homogène sont les fonctions $y_{\mathrm{H}} : t \mapsto \alpha\cos(\omega t) + \beta\sin(\omega t)$, ($\alpha$, $\beta \in \mathbb{R}$), et les solutions de l'équation sont donc les fonctions $y : t \mapsto \frac{1}{\omega^2} + \alpha\cos(\omega t) + \beta\sin(\omega t)$.
- $y'' - 4y' + 4y = 3(1 + \sin(t) + e^{2t})$ sur $\mathbb{R}$: l'équation caractéristique est $x^2 - 4x + 4 = 0 = (x-2)^2$, il y a une solution double 2, les solutions de l'équation homogène sont les fonctions $y(t) = (\alpha + \beta t)e^{2t}$.
  Cherchons une solution particulière en prenant comme second membre :
  – $f_1(t) = 3$: il y a une solution particulière constante $y_1(t) = \frac{3}{4}$.
  – $f_2(t) = 3e^{2t}$: on chercher une solution particulière de la forme $y = \mathrm{Q}(t)e^{2t}$, ce qui donne $\mathrm{Q}''(t) = 3$ et donc on peut prendre $y_2(t) = \frac{3}{2}t^2e^{2t}$.
  – $f_3(t) = 3\sin(t)$: on prend en fait $f_3(t) = 3e^{it}$ puis on prendra la partie imaginaire d'une solution particulière. On cherche y sous la forme $y(t) = \mathrm{Q}(t)e^{it}$ ce qui donne $\mathrm{Q}''(t) + (2i - 4)\mathrm{Q}'(t) + (3 -$

$4i)Q(t) = 3$, d'où $Q(t) = \frac{3}{3-4i} = 3\frac{3+4i}{25}$. Une solution particulière est donc $y_3(t) = \mathrm{Im}(3\frac{3+4i}{25}e^{it}) = \frac{9}{25}\sin(t) + \frac{12}{25}\cos(t)$. Les solutions sont donc les fonctions :

$$y(t) = \frac{3}{4} + \frac{9}{25}\sin(t) + \frac{12}{25}\cos(t) + (\alpha + \beta t + \frac{3}{2}t^2)e^{2t}.$$

★**Exercice 6.4** *Résoudre* $y'' - y' - 6y = 10e^{3x} + e^{-2x}$.

IV COMPLÉMENTS

1) Changement de variable

Il arrive que pour certaines équations différentielles qui n'entrent pas dans le cadre étudié précédemment, un changement de variable (et non pas de fonction inconnue) permette de s'y ramener. En voici un exemple, soit à résoudre (E) $x^2y''-xy'+y = x$ pour $x > 0$, en faisant le changement de variable $x = e^t$ avec $t \in \mathbb{R}$. Soit y une fonction deux fois dérivable sur $I =]0\,;+\infty[$, pour $t \in \mathbb{R}$ posons $g(t) = y(e^t)$ $[= y(x)]$, alors g est deux fois dérivable sur $\mathbb{R}$ et on a $g'(t) = e^t y'(e^t)$, et $g''(t) = e^t y'(e^t) + e^{2t}y''(t)$, d'où :

$$\begin{aligned}
(\mathrm{E}) &\iff \forall x > 0,\ x^2y''(x) - xy'(x) + y(x) = x \\
&\iff \forall t \in \mathbb{R},\ e^{2t}y''(e^t) - e^t y'(e^t) + y(e^t) = e^t \\
&\iff \forall t \in \mathbb{R},\ g''(t) - 2g'(t) + g(t) = e^t \quad (\mathrm{E}')
\end{aligned}$$

La fonction g vérifie donc une EDL2 à coefficients constants, l'équation caractéristique est $X^2 - 2X + 1 = 0$ qui admet 1 comme racine double, les solutions de l'équation homogène sont les fonction $g_H\colon t \to (a + bt)e^t$ avec $(a, b) \in \mathbb{R}^2$. D'après le cours, il existe une solution particulière g_0 de la forme $g_0(t) = ct^2e^t$, d'où $g_0'(t) = ce^t(t^2 + 2t)$ et $g_0''(t) = ce^t(t^2 + 4t + 2)$, en reportant dans l'équation on obtient $ce^t[t^2+4t+2-2t^2-4t+t^2] = e^t$ ce qui donne $c = \frac{1}{2}$ et donc $g(t) = \frac{t^2}{2}e^t$, les solutions générales de (E′) sont les fonctions $g\colon t \mapsto (a+bt+\frac{t^2}{2})e^t$, et donc les solutions de (E) sont les fonctions $y\colon x \mapsto (a+b\ln(x)+\frac{\ln^2(x)}{2})x$ car $g(t) = y(x)$ avec $t = \ln(x)$.

2) Équations à variables séparées

Définition 6.6

Une équation différentielle à variables séparées est une équation de la forme : $y'b(y) = a(t)$ *où* a, b *sont deux fonctions continues données.*

À retenir : méthode de résolution

Si a est continue sur un intervalle I et b sur un intervalle J, on peut considérer une primitive A de a sur I et une primitive B de b sur J, dans ce cas l'équation équivaut à : $\frac{d}{dt}[B(y)] = A'(t)$, et donc $B(y) = A(t) + \lambda$ où λ désigne une constante. On regarde ensuite si la fonction B est localement ou globalement bijective, auquel cas on pourra écrire $y(t) = B^{-1}(A(t) + \lambda)$.

☞**Exemple** : $t^3y' + y^3 = 0$ avec $y(1) = -1$, y ne doit pas être constamment nulle, si une telle solution existe, il doit exister un intervalle I sur lequel y ne s'annule pas, un tel intervalle ne peut pas contenir 0 et sur I l'équation est équivalente à : $\frac{y'}{y^3} = \frac{-1}{t^3}$, c'est une équation à variable séparée. Elle est équivalente à : $-\frac{1}{y^2} = \frac{1}{t^2} + \lambda$, ce qui donne $y^2 = -\frac{t^2}{1+\lambda t^2}$, on voit que la condition initiale donne la constante $\lambda = -2$. Comme y ne s'annule pas sur I, y garde un signe constant et donc $\forall\ t \in I, y(t) = -\sqrt{\frac{t^2}{2t^2-1}}$. Cette solution est définie sur l'intervalle $]\frac{1}{\sqrt{2}}\,;+\infty[$.

3) Équation de Bernoulli

Définition 6.7

Une équation de Bernoulli[2] *est une équation différentielle de la forme $y' = a(t)y^{\lambda} + b(t)y$ où a et b sont deux fonctions continues sur un intervalle I, et $\lambda \in \mathbb{R}^* \setminus \{1\}$.*

À retenir : méthode de résolution

La fonction nulle est solution. S'il existe une solution y non constamment nulle, alors il doit exister un intervalle J sur lequel y ne s'annule pas, sur un tel intervalle y est de signe constant, on peut donc faire le changement de fonction $y = \varepsilon z^{\alpha}$ avec $\varepsilon = \pm 1$ suivant le signe de y, l'équation devient alors : $\alpha z' = b(t)z + a(t)z^{\alpha(\lambda-1)+1}$, en prenant $\alpha = \frac{1}{1-\lambda}$, on a une équation différentielle linéaire du premier ordre, on sait donc la résoudre.

☞**Exemple :** $t^2y'+y+y^2 = 0$ avec $y(1) = 1$: y est une solution non constamment nulle, on pose $z = \frac{1}{y}$ ce qui donne : $z' = \frac{1}{t^2}z + \frac{1}{t^2}$. Les solutions de l'équation homogène sont les fonctions $z(t) = \lambda e^{-\frac{1}{t}}$ et une solution particulière est $z_1(t) = -1$, les solutions générales sont donc les fonctions $z(t) = -1 + \lambda e^{-\frac{1}{t}}$, la condition initiale donne $\lambda = 2e$ d'où $y(t) = \frac{1}{2e^{1-\frac{1}{t}}-1}$. Cette solution est définie sur l'intervalle $]\frac{1}{1+\ln(2)}\,;+\infty[$.

★**Exercice 6.5** *Résoudre $y' = xy^2 + y$ avec $y(0) = 1$.*

V SOLUTION DES EXERCICES

Solution 6.1

1/ Une primitive de h est $\mathrm{H}\colon t \mapsto \frac{1}{\alpha}e^{\alpha t}$.

2/ $f(t) = \frac{1-it}{1+t^2} = \frac{1}{1+t^2} + i\frac{-t}{1+t^2}$, donc une primitive F de f sur $\mathbb{R}$ est définie par $\mathrm{F}(t) = \arctan(t) - \frac{i}{2}\ln(1+t^2)$.

3/ $g(t)$ est la partie réelle de $h(t) = e^te^{it} = e^{at}$ avec $a = 1+i$. Une primitive de h sur $\mathbb{R}$ est $\mathrm{H}(t) = \frac{1}{a}e^{at} = e^t\frac{e^{it}}{1+i}$. Une primitive G de g sur $\mathbb{R}$ est définie par $\mathrm{G}(t) = \mathrm{Re}(\mathrm{H}(t)) = e^t\,\mathrm{Re}(\frac{e^{it}(1-i)}{2}) = \frac{1}{2}e^t(\cos(t)+\sin(t))$.

Solution 6.2 $\left|\int_0^{\pi} e^{(1+i)t}\,\mathrm{d}t\right| \leqslant \int_0^{\pi}\left|e^{(1+i)t}\right|\mathrm{d}t$ *(majoration en module), or $e^{(1+i)t} = e^te^{it}$ donc $|e^{(1+i)t}| = e^t$, d'où* $\left|\int_0^{\pi} e^{(1+i)t}\,\mathrm{d}t\right| \leqslant \int_0^{\pi} e^t\,\mathrm{d}t = e^{\pi}-1$.

Solution 6.3 *Équation différentielle linéaire d'ordre 1, le coefficient de y' s'annule en 0.*

– *Résolution sur $]-\infty;0[$: $(\mathrm{H}) \iff y' = (\frac{3}{x}+\frac{2}{x^3})y \iff y = \lambda e^{3\ln(|x|)-\frac{1}{x^2}} = \alpha\varphi(x)$ avec $\alpha \in \mathbb{R}$ et $\varphi(x) = x^3e^{-\frac{1}{x^2}}$. La méthode de la variation de la constante donne $\lambda' = \frac{1}{\varphi(x)} = \frac{1}{x^3}e^{\frac{1}{x^2}}$, on peut donc prendre $\lambda = -\frac{1}{2}e^{\frac{1}{x^2}}$ ce qui donne la solution particulière $y_1(x) = -\frac{x^3}{2}$. Les solutions générales sur cet intervalle sont les fonctions $y : x \mapsto -\frac{x^3}{2} + \lambda x^3e^{-\frac{1}{x^2}}$.*

– *Sur $]0\,;+\infty[$: mêmes résultats.*

– *Sur $\mathbb{R}$: y est solution sur $\mathbb{R}$ ssi* $\boxed{\begin{cases} y(x) = -\frac{x^3}{2} + \lambda x^3e^{-\frac{1}{x^2}} & \text{si } x<0 \\ y(x) = 0 & \text{si } x=0 \\ y(x) = -\frac{x^3}{2} + \alpha x^3e^{-\frac{1}{x^2}} & \text{si } x>0 \end{cases}}$ *avec $\lambda, \alpha \in \mathbb{R}$, on vérifie qu'une telle fonction est bien dérivable en 0.*

Solution 6.4 *Équation différentielle linéaire d'ordre 2 à coefficients constants. L'équation caractéristique est $x^2-x-6 = 0$ et ses racines sont 3 et -2, les solutions de l'équation homogène sont les fonctions $y : x \mapsto \alpha e^{3x}\beta e^{-2x}$ avec $\alpha,\beta \in \mathbb{R}$. Cherchons une solution particulière en prenant comme second membre :*

– *$10e^{3x}$, en posant $y_1(x) = axe^{3x}$ où a est une constante. On obtient la condition $5a = 10$, et donc $y_1(x) = 2xe^{3x}$.*

– *e^{-2x}, en posant $y_2(x) = axe^{-2x}$. On obtient la condition $-5a = 1$, et donc $y_2(x) = -\frac{x}{5})e^{-2x}$.*

Les solutions générales sont les fonctions $y : x \mapsto (\alpha + 2x)e^{3x} + (\beta - \frac{x^2}{10} - \frac{x}{5})e^{-2x}$ (principe de superposition).

Solution 6.5 *Équation de Bernoulli, on cherche une solution y qui ne s'annule par sur un intervalle I contenant 0. On pose $z = \frac{1}{y}$ ce qui donne $-\frac{z'}{z^2} = \frac{x}{z^2} + \frac{1}{z}$ ou encore $z' = -z + x$ avec la condition $z(0) = 1$. C'est une équation différentielle linéaire d'ordre 1, les solutions de l'équation homogène sont les fonctions $z : x \mapsto \lambda e^{-x}$ et la fonction $z_1 : x \mapsto 1-x$ est*

2. BERNOULLI Jakob (1654 – 1705) : c'est le plus illustre d'une grande famille de mathématiciens suisses.

solution particulière. Les solutions générales sont les fonctions $z : x \mapsto 1 - x + \lambda e^{-x}$*, la condition* $z(0) = 1$ *donne* $\lambda = 0$*, d'où* $z(x) = 1 - x$ *et donc* $y : x \mapsto \frac{1}{1-x}$*, définie sur* $\mathrm{I} =]-\infty\,;1[$.

Chapitre 7

Applications - Relations

Sommaire

I APPLICATIONS

La notion d'application (ou fonction) entre deux ensembles E et F (non vides) est une notion clé en mathématiques. C'est l'idée d'associer, ou de faire correspondre, à chaque élément de E un élément de F.

1) Définitions

Définition 7.1

Une relation $\mathcal{R}$ est la donnée de :
- *Un ensemble de départ :* E *(non vide).*
- *Un ensemble d'arrivée :* F *(non vide).*
- *D'un graphe* G *qui est une partie de* $\mathrm{E}\times\mathrm{F}$ *(*$\mathrm{G}\subset\mathrm{E}\times\mathrm{F}$*).*

Soient $x\in\mathrm{E}$ *et* $y\in\mathrm{F}$*, on dira que* x *est relation avec* y *pour* $\mathcal{R}$ *lorsque* $(x,y)\in\mathrm{G}$*, on écrira* $x\mathcal{R}y$*. Si c'est le cas, on dira que* y *est une image de* x *par* $\mathcal{R}$ *et que* x *est un antécédent de* y *par* $\mathcal{R}$*. Lorsque tout élément de* E *a une et une seule image par* $\mathcal{R}$*, on dit que* $\mathcal{R}$ *est une* ***application*** *(ou fonction). Si c'est le cas, et si* $x\mathcal{R}y$*, alors on écrira plutôt* $y=\mathcal{R}(x)$*, on dira que* y *est* ***l'image*** *de* x *par* $\mathcal{R}$*. L'ensemble des applications de* E *vers* F *est noté* $\mathcal{F}(\mathrm{E},\mathrm{F})$ *ou encore* F^{E}*.*

Pour désigner une application on utilise en général une lettre minuscule. Si f est une application de E vers F on écrit : $f\colon \begin{array}{rcl}\mathrm{E} & \to & \mathrm{F}\\ x & \mapsto & f(x)\end{array}$, et le graphe de f est l'ensemble $\mathrm{G}_f=\{(x;f(x))\,/\,x\in\mathrm{E}\}$.

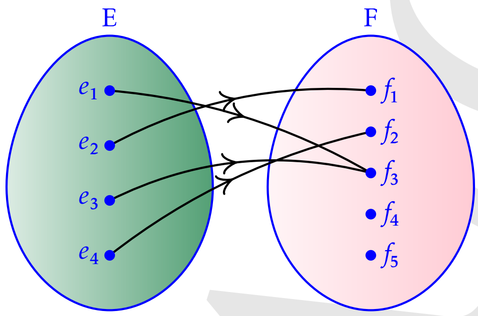
☞**Exemples** :

- L'exponentielle est une application de $\mathbb{R}$ vers $\mathbb{R}$.
- Le logarithme est une application de $]0\,;+\infty[$ vers $\mathbb{R}$.
- Soit $f\colon \mathbb{R} \to \mathbb{R}$ définie par $f(x) = \frac{1}{x}$ n'est pas une application car 0 n'a pas d'image, on dit que son ensemble de définition est $D_f = \mathbb{R}^*$. Par contre, la **restriction** de f à son ensemble de définition est une application, on la note $f|_{D_f}\colon D_f \to \mathbb{R}$. Lorsqu'une application g est une restriction d'une application f, on dit que f est un **prolongement** de g.

Définition 7.2 (identité d'un ensemble)
Soit E *un ensemble, l'identité de* E *est l'application de* E *dans* E *qui à chaque élément de* E *associe lui-même. On la note* $\mathrm{id}_E\colon \begin{array}{ccc} E & \to & E \\ x & \mapsto & \mathrm{id}_E(x) = x \end{array}$.

Diagramme sagittal

Lorsque les ensembles E et F ont très peu d'éléments, on peut représenter une application $f\colon E \to F$ sous forme d'un diagramme sagittal :



Dans cet exemple, le graphe de f est $G_f = \{(e_1, f_3); (e_2, f_1); (e_3, f_3); (e_4, f_2)\}$.

Attention ! (égalité de fonctions)
Deux fonctions f et g sont égales si et seulement si elles ont :
- *le même ensemble de départ* E,
- *le même ensemble d'arrivée* F,
- *le même graphe, c'est à dire :* $\forall x \in E, f(x) = g(x)$.

Par exemple, la fonction $f\colon \mathbb{R} \to \mathbb{R}$ *définie par* $f(x) = x^2$ *et la fonction* $g\colon \mathbb{R} \to \mathbb{R}^+$ *définie par* $g(x) = x^2$ *ne sont pas égales !*

Définition 7.3 (ensemble image)
Soit $f\colon E \to F$ *une application, on appelle ensemble image de* f *l'ensemble de toutes les images par* f*, on le note* $f(E)$ *(parfois* $\mathrm{Im}(f)$*). C'est donc une partie de* F*, plus précisément c'est l'ensemble des éléments de* F *qui ont au moins un antécédent par* f *dans* E *:*

$$f(E) = \{y \in F \,/\, \exists x \in E, y = f(x)\}$$

☞**Exemples** :
- Dans l'exemple du diagramme sagittal, l'ensemble image est $f(E) = \{f_1; f_2; f_3\}$.
- L'ensemble image de la fonction cosinus est $[-1; 1]$. Plus généralement, pour déterminer l'ensemble image d'une fonction $f\colon I \to \mathbb{R}$ où I est un intervalle de $\mathbb{R}$, on étudie les variations de f et sa continuité (en vue d'appliquer le théorème des valeurs intermédiaires).
- L'ensemble image de la fonction $g\colon \mathbb{C} \to \mathbb{C}$ définie par $\forall z \in \mathbb{C}, g(z) = z^2$ est $g(\mathbb{C}) = \mathbb{C}$.

2) Composition

Lorsque l'ensemble d'arrivée d'une application coïncide avec l'ensemble de départ d'une autre application, alors il est possible « d'enchaîner » les deux, c'est la composition :

$$\begin{array}{ccccc} E & \xrightarrow{f} & F & \xrightarrow{g} & G \\ x & \mapsto & f(x) & \mapsto & g(f(x)) \end{array}$$

Définition 7.4

Soient E, F, G *trois ensembles,* $f\colon \mathrm{E} \to \mathrm{F}$ *et* $g\colon \mathrm{F} \to \mathrm{G}$ *deux applications. La composée* $g \circ f$ *est l'application de* E *vers* G *définie par :* $\forall x \in \mathrm{E}, (g \circ f)(x) = g(f(x))$ *:*

$$g \circ f\colon \begin{array}{rcl} \mathrm{E} & \to & \mathrm{G} \\ x & \mapsto & (g \circ f)(x) = g(f(x)) \end{array}$$

Attention !

On prendra garde à l'ordre dans l'écriture de $g \circ f$.

Remarque 7.1 :

- *Lorsqu'on a deux applications* $f\colon \mathrm{E} \to \mathrm{F}$ *et* $g\colon \mathrm{H} \to \mathrm{G}$ *avec seulement* $\mathrm{F} \subset \mathrm{H}$ *(au lieu de* $\mathrm{F} = \mathrm{H}$*), alors on peut encore définir la composée et on la note encore* $g \circ f$ *même si théoriquement on devrait plutôt écrire* $(g|_{\mathrm{F}}) \circ f$.
- *Lorsque* f *est une application d'un ensemble* E *vers* ***lui-même****, alors on peut composer* f *avec elle-même, et autant de fois que l'on veut. Si* n *est un entier strictement positif, on notera* $f^n = f \circ \cdots \circ f$*. Par convention, on pose* $f^0 = \mathrm{id}_{\mathrm{E}}$.

☞**Exemple** : Si f est l'application de $\mathbb{R}$ vers $\mathbb{R}$ définie par $\forall x \in \mathbb{R}, f(x) = x + 1$, alors $\forall n \in \mathbb{N}$, f^n est l'application de $\mathbb{R}$ vers $\mathbb{R}$ définie par $\forall x \in \mathbb{R}, f^n(x) = x + n$.

★**Exercice 7.1** *Écrire la fonction* $f\colon]1\,;+\infty[\to \mathbb{R}$ *définie par* $f(x) = \frac{1}{\sqrt{x-1}}$*, comme composée de trois fonctions.*

Théorème 7.1 (propriétés)

Soient $f\colon \mathrm{E} \to \mathrm{F}$*,* $g\colon \mathrm{F} \to \mathrm{G}$ *et* $h\colon \mathrm{G} \to \mathrm{H}$ *trois applications.*

- $\mathrm{id}_{\mathrm{F}} \circ f = f$ *et* $f \circ \mathrm{id}_{\mathrm{E}} = f$.
- $(h \circ g) \circ f = h \circ (g \circ f)$*, c'est* ***l'associativité*** *de la composition.*

Preuve : Celle-ci est simple et laissée en exercice. □

3) Famille d'éléments d'un ensemble

Définition 7.5

Soit I *un ensemble non vide, et soit* E *un ensemble. On appelle famille d'éléments de* E *indexée par* I*, toute application* $u\colon \mathrm{I} \to \mathrm{E}$*, on note généralement cette famille* $(u_i)_{i \in \mathrm{I}}$*, et pour* $i \in \mathrm{I}$*, on note* $u(i) = u_i$ *(appelé terme d'indice* i*). L'ensemble de départ* I *est appelé ensemble des indices de la famille. Une famille d'éléments de* E *indexée par* $\mathbb{N}$ *est appelée* ***suite*** *d'éléments de* F*. L'ensemble des familles d'éléments de* E *indexées par* I *se note* $\mathcal{F}(\mathrm{I}, \mathrm{E})$ *ou* E^{I}.

Familles de parties d'un ensemble E

Conformément à la définition ci-dessus, une famille de parties de E indexée par un ensemble I non vide, est une application $\mathrm{A}\colon \mathrm{I} \to \mathcal{P}(\mathrm{E})$, que l'on peut noter $(\mathrm{A}_i)_{i \in \mathrm{I}}$ (A_i étant une partie de E pour tout $i \in \mathrm{I}$). On peut alors généraliser les notions d'intersection et de réunion de la manière suivante :

- La réunion de la famille $(\mathrm{A}_i)_{i \in \mathrm{I}}$ est $\bigcup\limits_{i \in \mathrm{I}} \mathrm{A}_i = \{x \in \mathrm{E} \,/\, \exists i \in \mathrm{I}, x \in \mathrm{A}_i\}$.
- L'intersection de la famille $(\mathrm{A}_i)_{i \in \mathrm{I}}$ est $\bigcap\limits_{i \in \mathrm{I}} \mathrm{A}_i = \{x \in \mathrm{E} \,/\, \forall i \in \mathrm{I}, x \in \mathrm{A}_i\}$.

★**Exercice 7.2**

1/ Pour $n \in \mathbb{N}$*, on pose* $\mathrm{A}_n = [n\,; n+1[$*. Déterminer* $\bigcup\limits_{n \in \mathbb{N}} \mathrm{A}_n$ *et* $\bigcap\limits_{n \in \mathbb{N}} \mathrm{A}_n$.

2/ Même question avec $n \in \mathbb{N}^*$ *et* $\mathrm{A}_n = [1 - \frac{1}{n}\,; 1 + \frac{1}{n}]$.

Les propriétés vues dans le chapitre I se généralisent :

Théorème 7.2

Soit $(A_i)_{i\in I}$ *une famille de parties de* E *et soit* B *une partie de* E*, alors :*

- $B\cap\left(\bigcup\limits_{i\in I} A_i\right) = \bigcup\limits_{i\in I}(B\cap A_i)$ *et* $B\cup\left(\bigcap\limits_{i\in I} A_i\right) = \bigcap\limits_{i\in I}(B\cup A_i)$ *(distributivité).*
- $C_E\left(\bigcup\limits_{i\in I} A_i\right) = \bigcap\limits_{i\in I} C_E(A_i)$ *et* $C_E\left(\bigcap\limits_{i\in I} A_i\right) = \bigcup\limits_{i\in I} C_E(A_i)$ *(lois de De Morgan).*

Preuve : Celle-ci est simple et laissée en exercice. □

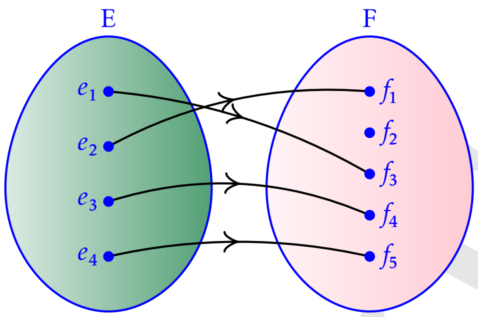
II INJECTION, SURJECTION, BIJECTION

Dans cette partie nous allons dégager des propriétés éventuelles des applications. Ces notions joueront un rôle important par la suite.

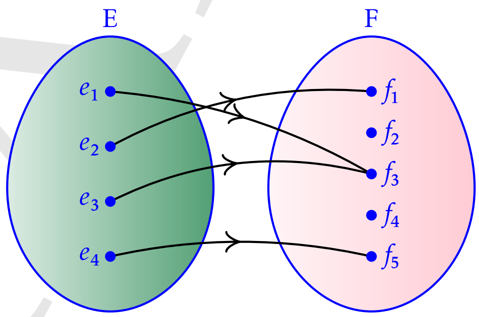
1) Injection

Définition 7.6

Soit $f\colon E\to F$ *une application, on dit que* f *est une injection (ou* f *est injective) lorsque* ***tout élément*** *de l'ensemble d'arrivée a* **au plus** *un antécédent dans l'ensemble de départ. Ce qui revient à dire : pour tout élément* y *de* F*, l'équation* $y = f(x)$ *a plus une solution* x *dans* E*.*



Injective



Non Injective

☞**Exemples** :

- Si E est un ensemble non vide, id_E est injective. Soit A une partie non vide de E, l'application $f\colon \begin{array}{ccc} A & \to & E \\ x & \mapsto & x \end{array}$ est injective, c'est **l'injection canonique** de A dans E.
- $f\colon \mathbb{R}\setminus\{1\}\to\mathbb{R}$ définie par $f(x) = \frac{x+1}{x-1}$ est une injection.
- $g\colon]0\,;+\infty[\to\mathbb{R}$ définie par $g(x) = \ln(x)$ est une injection.
- $h\colon \mathbb{R}\to\mathbb{R}$ définie par $h(x) = x^2$ n'est pas une injection.
- Une fonction $f\colon I\to\mathbb{R}$ **strictement monotone sur l'intervalle** I est injective.

Théorème 7.3 (définition équivalente)

$f\colon E\to F$ *est injective si et seulement si :* $\forall x,y\in E, f(x) = f(y) \implies x = y$*.*
Ce qui équivaut encore par contra-position à : $\forall x,y\in E, x\neq y \implies f(x)\neq f(y)$*.*

Preuve : Si f est injective : soit $x,y\in E$ tels que $f(x) = f(y)$, x et y sont donc deux antécédents d'un même élément de F, f étant injective ces deux éléments ne peuvent pas être distincts (sinon on a une contradiction) donc $x = y$.

Supposons que f vérifie : $\forall x,y\in E, f(x) = f(y) \implies x = y$. Soit $z\in F$ ayant deux antécédents distincts x et y dans E, alors $z = f(x) = f(y)$, on en déduit que $x = y$ ce qui est absurde, donc z ne peut pas avoir deux antécédents distincts, par conséquent f est injective. □

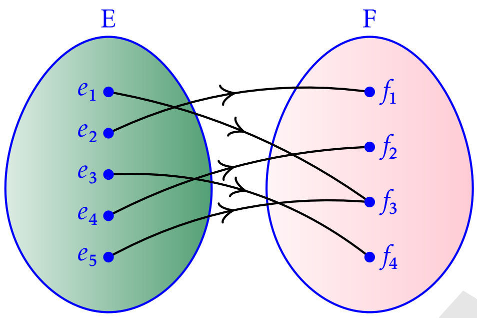
Théorème 7.4 (propriétés)
Soient $f\colon \mathrm{E}\to\mathrm{F}$ et $g\colon \mathrm{F}\to\mathrm{G}$ deux applications :
– Si f et g sont injectives alors $g\circ f$ est injective.
– Si $g\circ f$ est injective alors f est injective mais pas forcément g.

Preuve : Celle-ci est simple et laissée en exercice. □

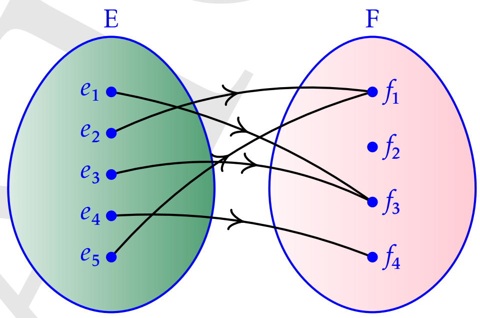
★**Exercice 7.3** *Soit $f\colon \mathrm{E}\to\mathrm{F}$ une application, montrer que f est injective si et seulement si il existe $h\colon \mathrm{F}\to\mathrm{E}$ telle que $h\circ f=\mathrm{id}_{\mathrm{E}}$.*

2) Surjection

Définition 7.7
*Soit $f\colon \mathrm{E}\to\mathrm{F}$ une application, on dit que f est une surjection (ou f est surjective) lorsque **tout élément** de l'ensemble d'arrivée a **au moins** un antécédent dans l'ensemble de départ. Ce qui revient à dire : pour tout élément y de F, l'équation $y=f(x)$ a moins une solution x dans E, ou encore $\forall y\in\mathrm{F}, \exists x\in\mathrm{E}, y=f(x)$.*



Surjective



Non surjective

À retenir
Dire que $f\colon \mathrm{E}\to\mathrm{F}$ est surjective, équivaut à $f(\mathrm{E})=\mathrm{F}$.

☞**Exemples** :
- Si E est un ensemble non vide, id_{E} est surjective.
- $f\colon \mathbb{C}\to\mathbb{C}$ définie par $f(z)=z^2$ et une surjection.
- $f\colon \mathbb{R}\to\mathbb{U}$ définie par $f(x)=e^{ix}$ est une surjection.
- $h\colon \mathbb{R}\setminus\{1\}\to\mathbb{R}$ définie par $h(x)=\frac{2x+1}{x-1}$ n'est pas surjective.
- Si $f\colon \mathrm{E}\to\mathrm{F}$ est une application, alors f induit une surjection entre E et $f(\mathrm{E})$ qui est l'application $g\colon \mathrm{E}\to f(\mathrm{E})$ définie par $g(x)=f(x)$.

Théorème 7.5 (propriétés)
Soient $f\colon \mathrm{E}\to\mathrm{F}$ et $g\colon \mathrm{F}\to\mathrm{G}$ deux applications :
– Si f et g sont surjectives alors $g\circ f$ est surjective.
– Si $g\circ f$ est surjective alors g est surjective mais pas forcément f.

Preuve : Celle-ci est simple et laissée en exercice. □

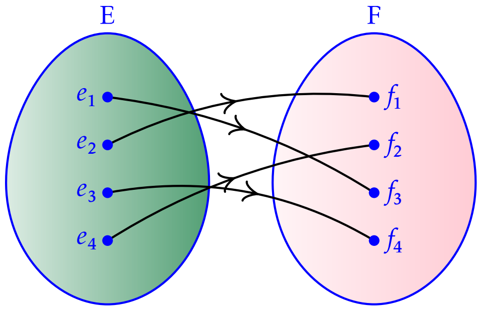
★**Exercice 7.4** *Soit $f\colon \mathrm{E}\to\mathrm{F}$ une application, montrer que f est surjective si et seulement si il existe $h\colon \mathrm{F}\to\mathrm{E}$ telle que $f\circ h=\mathrm{id}_{\mathrm{F}}$.*

★**Exercice 7.5** *Soit E un ensemble non vide, et f une application de E vers $\mathrm{P(E)}$. En considérant la partie $\mathrm{A}=\{x\in\mathrm{E}\,/\,x\notin f(x)\}$, montrer que f ne peut pas être surjective.*

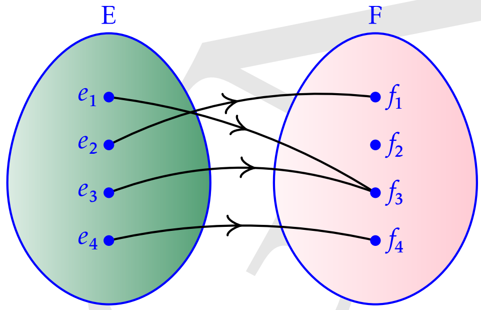
3) Bijection

Définition 7.8

Soient E, F *deux ensembles et* $f\colon \mathrm{E} \to \mathrm{F}$ *une application, on dit que* f *est une* ***bijection*** *(ou application bijective) lorsque* ***tout élément de*** F ***a un unique antécédent par*** f*, ce qui peut s'écrire de la manière suivante :* $\forall y \in \mathrm{F}, \exists!\, x \in \mathrm{E}, f(x) = y$.



Bijective



Non bijective

Dire que tout élément de F a un unique antécédent revient à dire que tout élément de F a au moins un antécédent et au plus un antécédent. Par conséquent dire que f est bijective revient à dire que f est surjective et injective.

À retenir

f est bijective $\iff$ f est surjective et injective.

Exemples :

- Si E est un ensemble non vide, alors id_E est une bijection.
- $f\colon [0\,;+\infty[\to [0\,;+\infty[$ définie par $f(x) = x^2$ est une bijection.
- $g\colon \mathbb{R} \to]0\,;+\infty[$ définie par $g(x) = e^x$ est une bijection.
- $h\colon \mathbb{R} \to \mathbb{R}$ définie par $h(x) = e^x$ n'est pas une bijection.

À retenir

Si $f\colon \mathrm{E} \to \mathrm{F}$ est injective, alors f induit une bijection de E vers $f(\mathrm{E})$ qui est $\tilde{f}\colon \mathrm{E} \to f(\mathrm{E})$ définie par $\tilde{f}(x) = f(x)$. Cela s'applique en particulier aux fonctions $f\colon \mathrm{I} \to \mathbb{R}$ strictement monotones sur l'intervalle I.

Définition 7.9 (bijection réciproque)

Si $f\colon \mathrm{E} \to \mathrm{F}$ *est une bijection, alors on peut considérer l'application qui va de* F *vers* E *et qui à tout élément* x *de* F *associe* ***son unique antécédent par*** f*, cette application est appelée* ***bijection réciproque de*** f*, on la note* f^{-1}*. Autrement dit,* $f^{-1}\colon \begin{array}{ccl} \mathrm{F} & \to & \mathrm{E} \\ x & \mapsto & y \text{ défini par } f(y) = x \end{array}$.

Attention !

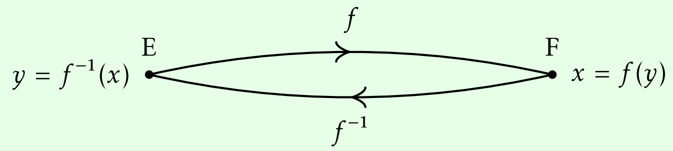
La notation f^{-1} *n'a de sens que lorsque* f *est bijective.*

Exemples :

- Si E est un ensemble non vide, alors id_E est une bijection et la bijection réciproque est $\mathrm{id}_\mathrm{E}^{-1} = \mathrm{id}_\mathrm{E}$.
- $f\colon [0\,;+\infty[\to [0\,;+\infty[$ définie par $f(x) = x^2$ est une bijection et la bijection réciproque est la fonction racine carrée.
- $g\colon \mathbb{R} \to]0\,;+\infty[$ définie par $g(x) = e^x$ est une bijection et la bijection réciproque est la fonction logarithme népérien.

À retenir

Lorsque $f\colon \mathrm{E} \to \mathrm{F}$ est bijective : $\forall x \in \mathrm{F}, \forall y \in \mathrm{E}, f^{-1}(x) = y \iff f(y) = x$.



Théorème 7.6

Soit $f\colon \mathrm{E} \to \mathrm{F}$ une bijection.

– On a $f^{-1} \circ f = \mathrm{id}_\mathrm{E}$ et $f \circ f^{-1} = \mathrm{id}_\mathrm{F}$. De plus f^{-1} est une bijection et $\left(f^{-1}\right)^{-1} = f$.

– Si $g\colon\ \colon \mathrm{F} \to \mathrm{G}$ est une autre bijection, alors la composée $g \circ f$ est une bijection de E vers G, et sa bijection réciproque est : $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

Preuve : – La composée $f^{-1} \circ f$ existe et va de E dans E. Si $x \in \mathrm{E}$, alors $(f^{-1} \circ f)(x) = f^{-1}(f(x)) = x$ car x est l'unique antécédent de $f(x)$ par f, on a donc $f^{-1} \circ f = \mathrm{id}_\mathrm{E}$. De même, $f \circ f^{-1}$ existe et va de F dans F. Si $x \in \mathrm{F}$, alors $(f \circ f^{-1})(x) = f(f^{-1}(x)) = x$ car $f^{-1}(x)$ est l'unique antécédent de x par f, on a donc $f \circ f^{-1} = \mathrm{id}_\mathrm{F}$. Le point suivant est évident.

– Si $g\colon \mathrm{F} \to \mathrm{G}$ est une autre bijection, soit $y \in \mathrm{G}$, alors pour tout $x \in \mathrm{E}$ on a :

$$\begin{aligned}(g \circ f)(x) = y &\iff g(f(x)) = y\\ &\iff f(x) = g^{-1}(y)\\ &\iff x = f^{-1}(g^{-1}(y)) = (f^{-1} \circ g^{-1})(y)\end{aligned}$$

le résultat en découle. □

★**Exercice 7.6** *Soient $f\colon \mathrm{E} \to \mathrm{F}$ et $g\colon \mathrm{F} \to \mathrm{E}$ deux applications telles que $f \circ g = \mathrm{id}_\mathrm{F}$ et $g \circ f = \mathrm{id}_\mathrm{E}$. Montrer que f et g sont bijectives et réciproques l'une de l'autre.*

Remarque 7.2 – *Le résultat de cet exercice est à connaître.*

Définition 7.10 (involution)

*Soit E un ensemble non vide. Une involution de E est une application f de E vers **lui-même** telle que $f \circ f = \mathrm{id}_\mathrm{E}$. Une telle application est bijective et elle est sa propre réciproque : $f^{-1} = f$.*

☞**Exemples** :
- Dans le plan, les symétries ponctuelles, les symétries axiales, sont des involutions du plan.
- La fonction $f\colon \mathbb{R}^* \to \mathbb{R}^*$ définie par $f(x) = \frac{1}{x}$ est une involution de $\mathbb{R}^*$.
- La conjugaison dans $\mathbb{C}$ est une involution de $\mathbb{C}$.

III IMAGES DIRECTES, IMAGES RÉCIPROQUES

1) Définitions

Définition 7.11

Soit $f\colon \mathrm{E} \to \mathrm{F}$ une application, A une partie de E et B une partie de F.

– On appelle image directe de A par f, l'ensemble des images des éléments de A par f, ou encore, l'ensemble des éléments F qui ont un antécédent dans A par f. Notation : $f(\mathrm{A}) = \{y \in \mathrm{F} \,/\, \exists x \in \mathrm{A}, f(x) = y\}$, c'est une partie de F.

– On appelle image réciproque de B par f, l'ensemble des antécédents des éléments de B par f. Notation : $f^{-1}(\mathrm{B}) = \{x \in \mathrm{E} \,/\, f(x) \in \mathrm{B}\}$, c'est une partie de E.

Attention !

*La notation $f^{-1}(\mathrm{B})$ **ne suppose pas que f est bijective**. Mais lorsque f est effectivement bijective, on peut vérifier que l'image réciproque de B par f correspond à l'image directe de B par f^{-1}.*

Remarque 7.3 :

- *On a $f^{-1}(\mathrm{F}) = \mathrm{E}$.*
- *Dans le cas d'une fonction $f\colon \mathrm{I} \to \mathbb{R}$ continue sur l'intervalle I, l'étude de la fonction permet de déterminer l'image directe d'une partie de I et l'image réciproque d'une partie de $\mathbb{R}$.*

☞**Exemple** : Soit f la fonction sinus, on a $f([0\,;\pi[) = [0;1]$, $f^{-1}([0\,;1]) = \bigcup\limits_{k\in\mathbb{Z}} [2k\pi;(2k+1)\pi]$.

2) Propriétés

Théorème 7.7

Soit $f\colon \mathrm{E} \to \mathrm{F}$ une application. Pour toutes parties A et B de E, on a :
- *Si $\mathrm{A} \subset \mathrm{B}$ alors $f(\mathrm{A}) \subset f(\mathrm{B})$.*
- *$f(\mathrm{A}\cup\mathrm{B}) = f(\mathrm{A}) \cup f(\mathrm{B})$.*
- *$f(\mathrm{A}\cap\mathrm{B}) \subset f(\mathrm{A}) \cap f(\mathrm{B})$.*
- *$\mathrm{A} \subset f^{-1}(f(\mathrm{A}))$.*

Preuve :
- Le premier point est évident.
- Si $x \in \mathrm{A}\cup\mathrm{B}$ alors $x \in \mathrm{A}$ ou $x \in \mathrm{B}$ donc $f(x) \in f(\mathrm{A})$ ou $f(x) \in f(\mathrm{B})$, c'est à dire $f(x) \in f(\mathrm{A}) \cup f(\mathrm{B})$. On a donc $f(\mathrm{A}\cup\mathrm{B}) \subset f(\mathrm{A}) \cup f(\mathrm{B})$. Réciproquement, on a $f(\mathrm{A}) \subset f(\mathrm{A}\cup\mathrm{B})$ et $f(\mathrm{B}) \subset f(\mathrm{A}\cup\mathrm{B})$ (d'après le premier point) et donc $f(\mathrm{A}) \cup f(\mathrm{B}) \subset f(\mathrm{A}\cup\mathrm{B})$, d'où l'égalité.
- Si $x \in \mathrm{A}\cap\mathrm{B}$ alors $f(x) \in f(\mathrm{A})$ et $f(x) \in f(\mathrm{B})$ d'où $f(x) \in f(\mathrm{A}) \cap f(\mathrm{B})$, donc $f(\mathrm{A}\cap\mathrm{B}) \subset f(\mathrm{A}) \cap f(\mathrm{B})$.
- Si $x \in \mathrm{A}$ alors $f(x) \in f(\mathrm{A})$, d'où l'inclusion. □

Remarque 7.4 :

- *Dans le cas de l'intersection (3^e propriété), on n'a pas l'égalité en général. Par exemple, si f est la fonction cosinus de $\mathbb{R}$ dans $\mathbb{R}$, si $\mathrm{A} = [-\frac{\pi}{2}\,;-\frac{\pi}{3}]$ et $\mathrm{B} = [0\,;\pi]$, alors $f(\mathrm{A}) \cap f(\mathrm{B}) = [0\,;\frac{1}{2}]$ alors que $f(\mathrm{A}\cap\mathrm{B}) = \varnothing$.*
- *De même pour la dernière propriété, par exemple, en reprenant la fonction cosinus avec $\mathrm{A} = [0\,;\pi]$, on a $f(\mathrm{A}) = [-1\,;1]$ et $f^{-1}(f(\mathrm{A})) = \mathbb{R}$.*

★**Exercice** 7.7

1/ *Montrer que les propriétés 2 et 3 du théorème précédent se généralisent à une famille quelconque de parties de E.*

2/ *Soient $f\colon \mathrm{E} \to \mathrm{F}$ et $g\colon \mathrm{F} \to \mathrm{G}$ deux applications, soit A une partie de E, montrer que $(g\circ f)(\mathrm{A}) = g(f(\mathrm{A}))$. Soit B une partie de G, montrer que $(g\circ f)^{-1}(\mathrm{B}) = f^{-1}(g^{-1}(\mathrm{B}))$.*

3/ *Montrer que $f\colon \mathrm{E} \to \mathrm{F}$ est injective si et seulement si pour tout partie A de E on a $f^{-1}(f(\mathrm{A})) = \mathrm{A}$.*

Théorème 7.8

Soit $f\colon \mathrm{E} \to \mathrm{F}$ une application. Pour toutes parties A et B de F, on a :
- *Si $\mathrm{A} \subset \mathrm{B}$ alors $f^{-1}(\mathrm{A}) \subset f^{-1}(\mathrm{B})$.*
- *$f^{-1}(\mathrm{A}\cup\mathrm{B}) = f^{-1}(\mathrm{A}) \cup f^{-1}(\mathrm{B})$.*
- *$f^{-1}(\mathrm{A}\cap\mathrm{B}) = f^{-1}(\mathrm{A}) \cap f^{-1}(\mathrm{B})$.*

Preuve :
- Si $x \in f^{-1}(\mathrm{A})$ alors $f(x) \in \mathrm{A}$ donc $f(x) \in \mathrm{B}$, d'où $x \in f^{-1}(\mathrm{B})$.
- $x \in f^{-1}(\mathrm{A}\cup\mathrm{B}) \iff f(x) \in \mathrm{A}$ ou $f(x) \in \mathrm{B} \iff x \in f^{-1}(\mathrm{A}) \cup f^{-1}(\mathrm{B})$.
- $x \in f^{-1}(\mathrm{A}\cap\mathrm{B}) \iff f(x) \in \mathrm{A}$ et $f(x) \in \mathrm{B} \iff x \in f^{-1}(\mathrm{A}) \cap f^{-1}(\mathrm{B})$. □

IV RELATIONS BINAIRES

Nous revenons dans cette partie à la notion générale de relation définie en début de chapitre. Mais on s'intéresse plus particulièrement aux relations d'un ensemble E dans lui-même.

1) Définitions

Définition 7.12

*Soit $\mathcal{R}$ une relation d'un ensemble E **vers lui - même**, on dit que $\mathcal{R}$ est :*
*– **réflexive** lorsque tout élément est en relation avec lui - même : $\forall\, x \in \mathrm{E}$, $x\mathcal{R}x$.*
*– **symétrique** lorsque : $\forall(x,y) \in \mathrm{E}^2$, si $x\mathcal{R}y$ alors $y\mathcal{R}x$ (le graphe de $\mathcal{R}$ est symétrique).*
*– **antisymétrique** lorsque : $\forall(x,y) \in \mathrm{E}^2$, si $x\mathcal{R}y$ et $y\mathcal{R}x$ alors $x = y$. On remarquera qu'il ne s'agit pas de la négation de symétrique.*
*– **transitive** lorsque : $\forall(x,y,z) \in \mathrm{E}^3$, si $x\mathcal{R}y$ et $y\mathcal{R}z$ alors $x\mathcal{R}z$.*

☞**Exemples** :
- Dans $\mathbb{R}$, la relation $\mathcal{R}$ définie par : $\forall(x,y) \in \mathbb{R}^2$, $x\mathcal{R}y \iff x \leqslant y$, est une relation réflexive, antisymétrique et transitive.
- Dans $\mathbb{Z}$, la relation $\mathcal{S}$ définie par : $\forall(x,y) \in \mathbb{Z}^2$, $x\mathcal{S}y \iff x - y \in 2\mathbb{Z}$, est une relation réflexive, symétrique et transitive.
- Soit E un ensemble, la relation $\mathcal{T}$ définie dans $\mathcal{P}(\mathrm{E})$ par : $\forall(\mathrm{A},\mathrm{B}) \in \mathcal{P}(\mathrm{E})^2$, $\mathrm{A}\mathcal{T}\mathrm{B} \iff \mathrm{A} \subset \mathrm{B}$. Cette relation $\mathcal{T}$ est réflexive antisymétrique et transitive.

2) Relation d'équivalence

Définition 7.13

*Soit E un ensemble et $\mathcal{R}$ une relation de E dans E, on dit que $\mathcal{R}$ est une relation d'équivalence lorsqu'elle est réflexive, symétrique et transitive. Si c'est le cas, alors pour tout élément a de E, on appelle **classe** de a l'ensemble des $x \in \mathrm{E}$ en relation avec a, notation : $\mathrm{Cl}(a) = \{x \in \mathrm{E} \,/\, x\mathcal{R}a\}$.*

☞**Exemples** :
- L'égalité dans un ensemble est une relation d'équivalence.
- Soit $n \in \mathbb{Z}$, la relation définie dans $\mathbb{Z}$, par $\forall(x,y) \in \mathbb{Z}^2$, $x\mathcal{R}y \iff x - y \in n\mathbb{Z}$, est une relation d'équivalence. Cette relation est appelée la congruence modulo n dans $\mathbb{Z}$, et on note $\forall(x,y) \in \mathbb{Z}^2$, $x \equiv y \pmod{n} \iff \exists k \in \mathbb{Z}, x - y = kn$.
- Soit $a \in \mathbb{R}$, la relation définie dans $\mathbb{R}$, par $\forall(x,y) \in \mathbb{R}^2$, $x\mathcal{R}y \iff x - y \in a\mathbb{Z}$, est une relation d'équivalence. Cette relation est appelée la congruence modulo a dans $\mathbb{R}$, et on note $\forall(x,y) \in \mathbb{R}^2$, $x \equiv y \pmod{a} \iff \exists k \in \mathbb{Z}, x - y = ka$.

Théorème 7.9

Si $\mathcal{R}$ est une relation d'équivalence dans E, alors :
- $\forall(a,b) \in \mathrm{E}^2$, $\mathrm{Cl}(a) = \mathrm{Cl}(b) \iff a\mathcal{R}b$.
- *Les classes d'équivalence forment une partition de E, c'est à dire :*
  - *Les classes d'équivalence sont des parties de E non vides et deux à deux disjointes.*
  - *La réunion des classes d'équivalence est égale à E.*

Preuve : Celle-ci est simple et laissée en exercice. □

☞**Exemple** : Considérons la relation de congruence modulo 5 dans $\mathbb{Z}$, soit $n \in \mathbb{Z}$, on a $\mathrm{Cl}(n) = \{n + 5k \,/\, k \in \mathbb{Z}\}$. On peut vérifier qu'il n'y a que cinq classes pour cette relation, celles de 0, de 1, de 2, de 3 et de 4.

3) Relation d'ordre

Définition 7.14

– Soit $\mathcal{R}$ une relation dans un ensemble E, on dit que $\mathcal{R}$ est une relation d'ordre lorsque cette relation est : réflexive, antisymétrique et transitive. Lorsque c'est le cas, on dit que $(\mathrm{E},\mathcal{R})$ est un ensemble ordonné.
*– Deux éléments x et y de E sont dits **comparables** pour l'ordre $\mathcal{R}$ lorsque l'on a $x\mathcal{R}y$ ou bien $y\mathcal{R}x$. Lorsque tous les éléments de E sont comparables deux à deux, on dit que l'ordre $\mathcal{R}$ est **total** et que $(\mathrm{E},\mathcal{R})$ est un ensemble totalement ordonné, sinon on dit que l'ordre est partiel et que $(\mathrm{E},\mathcal{R})$ est partiellement ordonné.*

– *Une relation d'ordre est en général notée $\leqslant$, c'est à dire que $x\mathcal{R}y$ est plutôt noté $x \leqslant y$.*

☞**Exemples** :

– L'ordre naturel sur les réels est une relation d'ordre total.
– Soit E un ensemble, $(\mathcal{P}(\mathrm{E}), \subset)$ est un ensemble partiellement ordonné (dès que $\mathrm{card}(\mathrm{E}) \geqslant 2$).
– Soit I un ensemble non vide, on pose $\mathrm{E} = \mathcal{F}(\mathrm{I}, \mathbb{R})$ l'ensemble des fonctions définies sur I et à valeurs réelles. On définit dans E la relation $\mathcal{R}$: pour $f, g \in \mathrm{E}$, $f\mathcal{R}g \iff \forall\, x \in \mathrm{I}, f(x) \leqslant g(x)$. On vérifie que $\mathcal{R}$ est une relation d'ordre **partiel** (dès que $\mathrm{card}(\mathrm{I}) > 1$), cette relation est appelée ordre fonctionnel et notée $\leqslant$.
– Pour (x, y) et $(x', y') \in \mathbb{R}^2$, on pose :

$$(x,y)\mathcal{R}(x',y') \iff \begin{cases} x < x' \\ \text{ou} \\ x = x' \text{ et } y \leqslant y' \end{cases}$$

On vérifie que $\mathcal{R}$ est une relation d'ordre total sur $\mathbb{R}^2$ (appelée ordre lexicographique et notée $\leqslant$).

Remarque 7.5 – *On prendra garde au fait que lorsque l'ordre est partiel, la négation de $x \leqslant y$ est :*

$$\begin{cases} x \textit{ et } y \textit{ ne sont pas comparables} \\ \textit{ou} \\ x \textit{ et } y \textit{ sont comparables et } x > y \end{cases}$$

Définition 7.15

Soit $(\mathrm{E}, \leqslant)$ *un ensemble ordonné et* A *une partie de* E, *on dit que :*
– A *est majoré dans* E *lorsque :* $\exists \mathrm{M} \in \mathrm{E}, \forall x \in \mathrm{A}, x \leqslant \mathrm{M}$.
– A *est minoré dans* E *lorsque :* $\exists m \in \mathrm{E}, \forall x \in \mathrm{A}, m \leqslant x$.
– A *est borné dans* E *lorsque* A *est à la fois majoré et minoré.*
– A *admet un maximum lorsque :* $\exists a \in \mathrm{A}, \forall x \in \mathrm{A}, x \leqslant a$. *Si c'est le cas, on note* $a = \max(\mathrm{A})$.
– A *admet un minimum lorsque :* $\exists a \in \mathrm{A}, \forall x \in \mathrm{A}, a \leqslant x$. *Si c'est le cas, on note* $a = \min(\mathrm{A})$.

Attention !

– *Une partie d'un ensemble ordonné n'est pas forcément majoré (ou minoré), par exemple* $\mathbb{N}$ *est non majoré dans* $\mathbb{R}$.
– *Une partie majorée (ou minorée) dans un ensemble ordonné n'a pas forcément de maximum (ou minimum). Par exemple* $[0;1[$ *dans* $\mathbb{R}$.

★**Exercice 7.8** *Montrer que si* A *admet un maximum dans* $(\mathrm{E}, \leqslant)$, *alors celui-ci est unique (même chose pour minimum).*

V SOLUTION DES EXERCICES

Solution 7.1 *On a :*

$$\begin{array}{rccccccc} f\colon &]1\,;+\infty[& \to &]0\,;+\infty[& \to &]0\,;+\infty[& \to &]0\,;+\infty[\\ & x & \mapsto & x-1 & \mapsto & \sqrt{x-1} & \mapsto & \frac{1}{\sqrt{x-1}} \end{array}$$

Solution 7.2

1/ *La réunion est* $]0\,;+\infty[$, *et l'intersection est vide.*

2/ *La réunion est* $]0\,;2[$ *et l'intersection est le singleton* $\{1\}$.

Solution 7.3 *Si h existe alors f est injective d'après le théorème.*

Réciproquement, si f est injective, soit $y \in \mathrm{F}$, on pose $h(y) = x$ avec x antécédent de y par f si y a un antécédent (on sait alors qu'il est unique) et si y n'a pas d'antécédent par f on choisit ce qu'on veut pour $h(y)$ dans E. *On vérifie alors que pour tout $x \in \mathrm{E}$, $h(f(x)) = x$ car x est l'unique antécédent de $f(x)$ par f.*

Solution 7.4 *Si h existe alors f est surjective d'après le théorème.*

Réciproquement, si f est surjective, soit $y \in \mathrm{F}$, on pose $h(y) = x$ avec x antécédent de y par f que l'on choisit car il en existe. On vérifie alors que pour tout $x \in \mathrm{E}$, $f(h(x)) = x$ car $h(x)$ est un antécédent de x par f.

Solution 7.5 *Par l'absurde : si f est surjective, alors il existe $x_0 \in \mathrm{E}$ tel que $f(x_0) = \mathrm{A}$, si $x_0 \in \mathrm{A}$, alors $x_0 \in f(x_0)$ et donc $x_0 \notin \mathrm{A}$ (par définition de A), ce qui est absurde, donc $x_0 \notin \mathrm{A}$, mais alors $x_0 \notin f(x_0)$ et donc $x_0 \in \mathrm{A}$, ce qui est de nouveau absurde. Par conséquent f ne peut pas être surjective.*

Solution 7.6 *On sait que id_F et id_E sont injectives, on en déduit que g et f sont injectives. On sait aussi que id_F et id_E sont surjectives, on en déduit que f et g sont surjectives. Ce sont donc deux bijections, d'où $f = g^{-1} \circ \mathrm{id}_\mathrm{E} = g^{-1}$.*

Solution 7.7

1/ *C'est le même raisonnement quand dans la preuve.*

2/ *Soient $f\colon \mathrm{E} \to \mathrm{F}$ et $g\colon \mathrm{F} \to \mathrm{G}$ deux applications. Soit A une partie de E :*

$$z \in g(f(\mathrm{A})) \iff \exists y \in f(\mathrm{A}),\ z = g(y) \iff \exists x \in \mathrm{A},\ z = g(f(x)) = (g \circ f)(x) \iff z \in (g \circ f)(\mathrm{A})$$

Soit B une partie de G :

$$x \in (g \circ f)^{-1}(\mathrm{B}) \iff g(f(x)) \in \mathrm{B} \iff f(x) \in g^{-1}(\mathrm{B}) \iff x \in f^{-1}(g^{-1}(\mathrm{B}))$$

3/ *Si f est injective, soit A une partie de E, on sait que $\mathrm{A} \subset f^{-1}(f(\mathrm{A}))$, si $x \in f^{-1}(f(\mathrm{A}))$, alors $f(x) \in f(\mathrm{A})$, donc il existe $y \in \mathrm{A}$ tel que $f(x) = f(y)$, mais alors $x = y$ (f étant injective) et donc $x \in \mathrm{A}$. Donc $f^{-1}(f(\mathrm{A})) = \mathrm{A}$.*
Réciproque, si pour tout partie A de E on a $f^{-1}(f(\mathrm{A})) = \mathrm{A}$. Sot $x \in \mathrm{E}$ alors $f^{-1}(f(\{x\})) = \{x\}$, donc si $f(y) = f(x)$ alors $y \in \{x\}$, d'où $y = x$. L'application f est donc injective.

Solution 7.8 *Soient m_1 et m_2 deux maximums de A, alors $m_1 \leqslant m_2$ car m_2 est un maximum, et inversement $m_2 \leqslant m_1$ car m_1 est un maximum, d'où $m_1 = m_2$.*

Chapitre 8

Nombres réels

Sommaire

L'existence des ensembles $\mathbb{Q}$ et $\mathbb{R}$ est admise.

I L'ENSEMBLE DES RÉELS

1) Rappels sur les rationnels

Un rationnel est un réel de la forme pq^{-1} (ou $\frac{p}{q}$) où p et q sont deux entiers avec $q \neq 0$. L'ensemble des rationnels est noté $\mathbb{Q}$. Tout rationnel peut s'écrire de différentes manières sous forme de fractions, par exemple : $\frac{p}{q} = \frac{2p}{2q} = \frac{-p}{-q}$. Mais tout nombre rationnel s'écrit de manière **unique** sous forme de fraction **irréductible**, c'est à dire sous la forme $\frac{p}{q}$ avec $p \in \mathbb{Z}$, $q \in \mathbb{N}^*$ et avec p et q **premiers entre eux** (*i.e.* sans autres diviseurs communs que 1 et -1).

Opérations sur les rationnels

On rappelle que : $\frac{p}{q} + \frac{a}{b} = \frac{aq+bp}{bq}$ et $\frac{p}{q} \times \frac{a}{b} = \frac{ap}{bq}$. L'addition et la multiplication sont donc des lois de composition internes dans $\mathbb{Q}$, on vérifie que $(\mathbb{Q}, +, \times)$ est un **corps commutatif**. On vérifie également que $(\mathbb{Q}, +)$, $(\mathbb{Q}^*, \times)$ et $(\mathbb{Q}^{*+}, \times)$ sont des groupes commutatifs.

L'ensemble des rationnels est insuffisant

En termes d'approximations numériques, $\mathbb{Q}$ peut paraître suffisant en sciences appliquées. Le problème se pose lorsqu'on a besoin de connaître la **valeur exacte** de certaines grandeurs. Par exemple, peut - on mesurer dans $\mathbb{Q}$ la longueur de la diagonale d'un carré de côté 1 ? D'après le théorème de *Pythagore* [1], cela revient à se demander s'il existe un rationnel dont le carré est égal à 2, or nous avons déjà établi que la réponse est négative ($\sqrt{2} \notin \mathbb{Q}$).

Cette lacune de $\mathbb{Q}$ avait été remarquée par les Pythagoriciens, ce qui a conduit les mathématiciens à introduire de nouveaux nombres, les **irrationnels**, en concevant un ensemble plus vaste que $\mathbb{Q}$, l'ensemble des nombres réels noté $\mathbb{R}$.

1. *PYTHAGORE De Samos* (569 av J.-C. – 500 av J.-C. (environ)) : mathématicien et philosophe grec dont la vie et l'œuvre restent entourées de mystères.

2) Opérations et ordre sur les réels

L'ensemble $\mathbb{R}$ contient $\mathbb{Q}$ et possède une addition et une multiplication (qui prolongent celles de $\mathbb{Q}$) qui font que $(\mathbb{R}, +, \times)$ est un corps commutatif. On admettra également qu'il existe deux parties de $\mathbb{R}$ que l'on note A et B et qui vérifient :

- A et B sont stables pour l'addition.
- $\mathbb{Q}^+ \subset \mathrm{A}$ et $\mathbb{Q}^- \subset \mathrm{B}$.
- $\mathbb{R} = \mathrm{A} \cup \mathrm{B}$.
- $\mathrm{A} \cap \mathrm{B} = \{0\}$.
- Si $x, y \in \mathrm{A}$ alors $xy \in \mathrm{A}$, si $x, y \in \mathrm{B}$ alors $xy \in \mathrm{A}$ et si $x \in \mathrm{A}$ et $y \in \mathrm{B}$, alors $xy \in \mathrm{B}$ (règle des signes).

On définit alors une relation $\mathcal{R}$ dans $\mathbb{R}$ en posant : $\forall\, x, y \in \mathbb{R}, x\mathcal{R}y \iff x - y \in \mathrm{B}$. Cette relation est :

- Réflexive : $\forall\, x \in \mathbb{R}, x\mathcal{R}x$.
- Antisymétrique : si $x\mathcal{R}y$ et $y\mathcal{R}x$ alors $x = y$.
- Transitive : si $x\mathcal{R}y$ et $y\mathcal{R}z$, alors $x\mathcal{R}z$.

Le relation $\mathcal{R}$ est donc une relation **d'ordre** sur $\mathbb{R}$. On la notera désormais $\leqslant$, c'est à dire que $x\mathcal{R}y$ sera noté $x \leqslant y$ (*i.e.* $x - y \in \mathrm{B}$).

On remarquera que $x \leqslant 0$ signifie que $x \in \mathrm{B}$, et que $0 \leqslant x$ signifie que $-x \in \mathrm{B}$ et donc $x \in \mathrm{A}$ car $x = (-1)(-x)$: produit de deux éléments de B. D'autre part, si $x \in \mathrm{A}$ et $y \in \mathrm{B}$, alors $x \leqslant y$ car $y - x = y + (-x)$: somme de deux éléments de B.

Si x et y sont deux réels quelconques, on a $x - y \in \mathrm{A}$ ou $x - y \in \mathrm{B}$, c'est à dire $x - y \in \mathrm{B}$ ou $y - x \in \mathrm{B}$, c'est à dire encore $x \leqslant y$ ou $y \leqslant x$. Deux réels sont donc toujours comparables, l'ordre est **total**.

Notation : On pose $\mathrm{A} = \mathbb{R}^+$ et $\mathrm{B} = \mathbb{R}^-$.

Théorème 8.1

La relation d'ordre $\leqslant$ est :

• Compatible avec l'addition, c'est à dire :

$$\forall\, x, y, z \in \mathbb{R}, \text{ si } x \leqslant y \text{ alors } x + z \leqslant y + z.$$

• Compatible avec la multiplication par un réel positif :

$$\forall\, x, y, z \in \mathbb{R}, \text{ si } 0 \leqslant z \text{ et } x \leqslant y, \text{ alors } xz \leqslant yz.$$

Preuve : Si $x \leqslant y$, alors $x - y \in \mathbb{R}^-$, mais $(x + z) - (y + z) = x - y$, donc $(x + z) - (y + z) \in \mathbb{R}^-$ *i.e.* $x + z \leqslant y + z$. Si $0 \leqslant z$ et $x \leqslant y$, alors $x - y \in \mathbb{R}^-$ donc $z(x - y) \in \mathbb{R}^-$, *i.e.* $zx \leqslant zy$. On remarquera que si $z \leqslant 0$ alors $z(x - y) \in \mathbb{R}^+$ donc $zy \leqslant zx$, l'inégalité change de sens. □

Conséquences

- Si $x \leqslant y$ et $a \leqslant b$, alors $x + a \leqslant y + b$.
- Si $0 \leqslant x \leqslant y$ et $0 \leqslant a \leqslant b$ alors $0 \leqslant ax \leqslant by$.

II BORNE INFÉRIEURE, BORNE SUPÉRIEURE

1) Propriété fondamentale de l'ensemble des réels

Soit A une partie non vide de $\mathbb{R}$ et soit a un réel, on dit que :

- A est majorée par a (ou a est un majorant de A), lorsque tout élément de A est inférieur ou égal à a : $\forall\, x \in \mathrm{A}, x \leqslant a$.
- A est minorée par a (ou a est un minorant de A), lorsque tout élément de A est supérieur ou égal à a : $\forall\, x \in \mathrm{A}, x \geqslant a$.
- A est bornée, lorsque A est à la fois minorée et majorée : $\exists\, m, \mathrm{M} \in \mathbb{R}, \forall\, x \in \mathrm{A}, m \leqslant x \leqslant \mathrm{M}$.

☞**Exemples** :

- L'ensemble $\mathrm{A} = \{\frac{x^2}{1+x^2} \ / \ x \in \mathbb{R}\}$ est borné (minoré par 0 et majoré par 1).
- L'ensemble $\mathrm{A} = \{\frac{x^2}{1+|x|} \ / \ x \in \mathbb{R}\}$ est minoré par 0, mais non majoré (étudier la fonction $x \mapsto \frac{x^2}{1+|x|}$).

Remarque 8.1 –

- *A est non majoré équivaut à :* $\forall\, \mathrm{M} \in \mathbb{R}, \exists\, x \in \mathrm{A}, x > \mathrm{M}$.
- *A est non minoré équivaut à :* $\forall\, m \in \mathbb{R}, \exists\, x \in \mathrm{A}, x < m$.
- *A est borné équivaut à :* $\exists\, \mathrm{M} \in \mathbb{R}, \forall\, x \in \mathrm{A}, |x| \leqslant \mathrm{M}$.

Définition 8.1

Soit A *une partie non vide de* $\mathbb{R}$*. Si l'ensemble des majorants de* A *n'est pas vide et s'il admet un plus petit élément, alors celui-ci est appelé* **borne supérieure** *de* A *et noté* sup(A)*. La borne supérieure (lorsqu'elle existe) est donc* **le plus petit des majorants***.*
Si l'ensemble des minorants de A *n'est pas vide et s'il admet un plus grand élément, alors celui-ci est appelé* **borne inférieure** *de* A *et noté* inf(A)*. La borne inférieure (lorsqu'elle existe) est donc* **le plus grand des minorants***.*

☞**Exemples** :
- $A =]0;1]$, l'ensemble des majorants est $[1\,;+\infty[$, celui-ci admet un plus petit élément qui est 1, donc $\sup(A) = 1$. L'ensemble des minorants de A est $]-\infty\,;0]$ qui admet un plus grand élément : 0, donc $\inf(A) = 0$.
- $A =]1\,;+\infty[$, l'ensemble des majorants est vide donc A **n'a pas de borne supérieure**. L'ensemble des minorants est $]-\infty\,;1]$, celui-ci admet un plus grand élément : 1, donc $\inf(A) = 1$.
- Soient $a < b$ deux réels, si $A = [a;b]$ ou $[a;b[$ ou $]a;b]$ ou $]a;b[$, alors comme ci-dessus on montrer que $\sup(A) = b$ et $\inf(A) = a$.

Attention !

On remarquera qu'une borne inférieure (ou supérieure) d'une partie A *de* $\mathbb{R}$*, n'a aucune raison d'appartenir à* A*.*

Voici le lien entre minimum et borne inférieure (ou maximum et borne supérieure) :

Théorème 8.2

Soit A *une partie non vide de* $\mathbb{R}$ *et soit* a *un réel :*
- $a = \min(A)$ *si et seulement si* $a \in A$ *et* $a = \inf(A)$.
- $a = \max(A)$ *si et seulement si* $a \in A$ *et* $a = \sup(A)$.

Preuve : Celle-ci est simple et laissée en exercice. □

Il découle de la définition :

Théorème 8.3

Soit A *une partie non vide de* $\mathbb{R}$ *et soit* m *un réel, alors :*

$$m = \sup(A) \iff \begin{cases} m \text{ majore } A \\ \forall \varepsilon > 0, m - \varepsilon \text{ ne majore pas } A, \textit{i.e. } \exists x \in A, m - \varepsilon < x \end{cases}.$$

$$m = \inf(A) \iff \begin{cases} m \text{ minore } A \\ \forall \varepsilon > 0, m + \varepsilon \text{ ne minore pas } A, \textit{i.e. } \exists x \in A, x < m + \varepsilon \end{cases}.$$

Théorème 8.4 (**Propriété fondamentale de $\mathbb{R}$ (admise)**)

Toute partie de $\mathbb{R}$ ***non vide et majorée*** *admet une borne supérieure.*
Toute partie de $\mathbb{R}$ ***non vide et minorée*** *admet une borne inférieure.*

☞**Exemple** : Soit a un réel positif, on pose $A = \left\{x \in \mathbb{R} \,/\, x^2 \leqslant a\right\}$. A est une partie non vide de $\mathbb{R}$ car $0 \in A$, d'autre part A est majoré par $a+1$ car $x > a+1 \implies x^2 > a^2 + 2a + 1 > a$. L'ensemble A admet donc une borne supérieure M. En raisonnant par l'absurde on peut montrer que $M^2 = a$, par conséquent $M = \sqrt{a}$, c'est une définition possible de la fonction racine carrée.

★**Exercice 8.1** *Soient* A *et* B *deux parties de* $\mathbb{R}$ *non vides et bornées telles que* $A \subset B$*. Montrer que* $\inf(B) \leqslant \inf(A)$ *et* $\sup(A) \leqslant \sup(B)$.

★**Exercice 8.2** *Soient* A *et* B *deux parties de* $\mathbb{R}$ *non vides et majorées, on pose* $A + B = \{a + b \,/\, a \in A, b \in B\}$*. Montrer que* $\sup(A+B) = \sup(A) + \sup(B)$.

★**Exercice 8.3** *Soit* $A = \{\frac{1}{n} + (-1)^n \,/\, n \in \mathbb{N}^*\}$*. Déterminer les bornes supérieure et inférieure de* A *si elles existent.*

2) Intervalles

Définition 8.2

Soit I *une partie non vide de* $\mathbb{R}$*, on dit que* I *est un intervalle lorsque :* ***tout réel compris entre deux éléments de*** I ***est lui-même élément de*** I*, c'est à dire :*

$$\forall\, x, y \in \mathrm{I}, \forall\, z \in \mathbb{R}, x \leqslant z \leqslant y \implies z \in \mathrm{I}.$$

Par convention, $\varnothing$ *est un intervalle de* $\mathbb{R}$*.*

Théorème 8.5

Si I *est un intervalle non vide de* $\mathbb{R}$ *alors on a :*
- *soit* $\mathrm{I} = \mathbb{R}$,
- *soit* $\mathrm{I} = [a; +\infty[$ *ou* $\mathrm{I} =]a, +\infty[$,
- *soit* $\mathrm{I} =]-\infty; b]$ *ou* $\mathrm{I} =]-\infty; b[$,
- *soit* $\mathrm{I} =]a; b[$ *ou* $\mathrm{I} =]a; b]$ *ou* $\mathrm{I} = [a; b[$ *ou* $\mathrm{I} = [a; b]$.

Preuve : Le premier correspond à I non borné, le deuxième à I minoré et non majoré, le troisième à I non minoré et majoré, le quatrième à I borné. □

☞**Exemple** : $\mathbb{Z}$ n'est pas un intervalle de $\mathbb{R}$ car $1, 2 \in \mathbb{Z}$ mais pas $\frac{3}{2}$. $\mathbb{Q}$ n'est pas un intervalle de $\mathbb{R}$.

Théorème 8.6

On a les propriétés suivantes :
- *L'intersection de deux intervalles de* $\mathbb{R}$ *est un intervalle de* $\mathbb{R}$*.*
- *La réunion de deux intervalles de* $\mathbb{R}$ ***non disjoints*** *est un intervalle de* $\mathbb{R}$*.*

Preuve : Soient I et J deux intervalles de $\mathbb{R}$, posons $\mathrm{K} = \mathrm{I} \cap \mathrm{J}$. Si K est vide, alors c'est un intervalle. Si K n'est pas vide, alors soit $x, y \in \mathrm{K}$ et soit z un réel tel que $x \leqslant z \leqslant y$. Comme I est un intervalle contenant x et y, I contient z, de même J contient z, finalement $z \in \mathrm{K}$ et donc K est un intervalle de $\mathbb{R}$.

Supposons I et J non disjoints et soit $\mathrm{K} = \mathrm{I} \cup \mathrm{J}$. K est non vide, soit $x, y \in \mathrm{K}$ et soit z un réel tel que $x \leqslant z \leqslant y$. Si x et y sont dans I, alors z est dans I et donc dans K, de même si x et y sont dans J. Si x est dans I et y dans J, soit $t \in \mathrm{I} \cap \mathrm{J}$, si $z \leqslant t$, alors z est compris entre x et t qui sont éléments de I, donc $z \in \mathrm{I}$. Si $t \leqslant z$, alors z est compris entre t et y qui sont éléments de J, donc z est élément de J. Dans les deux cas on a bien $z \in \mathrm{K}$ et donc K est un intervalle de $\mathbb{R}$. □

3) La droite numérique achevée

On ajoute à l'ensemble $\mathbb{R}$ deux éléments non réels (par exemple i et $-i$), l'un de ces deux éléments est noté $-\infty$ et l'autre $+\infty$.

Définition 8.3

L'ensemble $\mathbb{R} \cup \{-\infty, +\infty\}$ *est noté* $\overline{\mathbb{R}}$ *et appelé* ***droite numérique achevée****.*

On prolonge la relation d'ordre de $\mathbb{R}$ à $\overline{\mathbb{R}}$ en posant pour tout réel x : $-\infty < x < +\infty$. L'ensemble $\overline{\mathbb{R}}$ devient ainsi un ensemble totalement ordonné, de plus il possède un maximum ($+\infty$) et un minimum ($-\infty$).

Pour tout réel x on pose :
- $(+\infty) + x = x + (+\infty) = +\infty$.
- $(-\infty) + x = x + (-\infty) = -\infty$.
- $(+\infty) + (+\infty) = +\infty$.
- $(-\infty) + (-\infty) = -\infty$.
- Si $x > 0$: $x(+\infty) = (+\infty)x = +\infty$ et $(-\infty)x = x(-\infty) = -\infty$.
- si $x < 0$: $x(+\infty) = (+\infty) = -\infty$ et $(-\infty)x = x(-\infty) = +\infty$.
- $(+\infty)(+\infty) = +\infty$, $(-\infty)(-\infty) = +\infty$ et $(-\infty)(+\infty) = (+\infty)(-\infty) = -\infty$.

Remarque 8.2 – *On prendra garde au fait que nous n'avons pas défini de loi de composition interne dans* $\overline{\mathbb{R}}$ *puisque nous n'avons pas défini* $0 \times (\pm\infty)$ *ni* $(-\infty) + (+\infty)$*. Les règles de calculs définies ci-dessus auront leur utilité dans le chapitre sur les limites.*

4) Voisinages

Définition 8.4

Soit $x \in \mathbb{R}$*, tout intervalle de la forme* $]x - \varepsilon; x + \varepsilon[$ *où* $\varepsilon > 0$ *est appelé* ***voisinage*** *de* x *.*
Tout intervalle ouvert de la forme $]a; +\infty[$ *(*$a \in \mathbb{R}$*) est appelé voisinage de* $+\infty$ *.*
Tout intervalle ouvert de la forme $]-\infty; a[$ *(*$a \in \mathbb{R}$*) est appelé voisinage de* $-\infty$ *.*

Théorème 8.7

Soit V_1, V_2 *deux voisinages de* $x \in \overline{\mathbb{R}}$*, alors* $V_1 \cap V_2$ *est un voisinage de* x*. Soit* $a, b \in \overline{\mathbb{R}}$*, si* $a < b$ *alors il existe un voisinage* V *de* a *et un voisinage* V' *de* b *tels que* $\forall x \in V$ *et* $\forall y \in V'$*,* $x < y$*.*

Preuve : Celle - ci est simple et laissée en exercice. □

Définition 8.5

Soit $P(x)$ *une proposition dépendante de* $x \in \mathbb{R}$*, et soit* $a \in \overline{\mathbb{R}}$*, on dit que la propriété* P *est* ***vraie au voisinage*** *de* a ***lorsqu'il existe au moins un voisinage*** V *de* a *tel que :*

$$\forall\, x \in V, P(x) \text{ est vraie.}$$

☞**Exemple** : Soit $f(x) = x^2 + x - 1$, alors au voisinage de 0 on a $f(x) < 0$, et au voisinage de $+\infty$, $f(x) > 0$. En effet, le trinôme $x^2 + x - 1$ admet deux racines réelles : $x_1 < 0$ et $x_2 > 0$, posons $\varepsilon = \min(|x_1|, |x_2|)$, si $x \in]0 - \varepsilon; 0 + \varepsilon[$ alors $x \in]x_1; x_2[$ et donc $x^2 + x - 1 < 0$, $V =]0 - \varepsilon; 0 + \varepsilon[$ est donc un voisinage de 0 et sur ce voisinage on a bien $f(x) < 0$. Posons $W =]x_2; +\infty[$, alors W est un voisinage de $+\infty$ et sur ce voisinage on a bien $f(x) > 0$.

III APPROXIMATION D'UN RÉEL

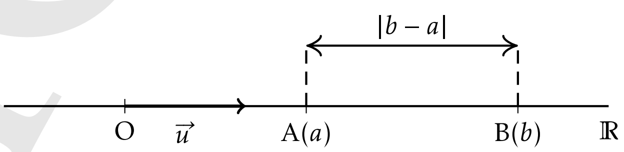
1) Valeur absolue

Soit x un réel, les deux nombres x et $-x$ sont comparables puisque l'ordre est total, ce qui donne un sens à la définition suivante :

Définition 8.6

Soit $x \in \mathbb{R}$*, on appelle* ***valeur absolue*** *de* x *le réel noté* $|x|$ *et défini par :* $|x| = \max(x, -x)$*. On a donc* $|x| = x$ *lorsque* $0 \leqslant x$*, et* $|x| = -x$ *lorsque* $x \leqslant 0$*.*

L'ensemble $\mathbb{R}$ peut être assimilé à une droite graduée (*i.e.* munie d'un repère $(O, \vec{u})$), les réels sont alors les abscisses des points de cette droite. Si $A(a)$ et $B(b)$ sont deux points de cette droite, alors le réel positif $|b - a|$ représente la **distance** de A à B, en particulier $|x|$ représente la distance de l'origine au point d'abscisse x.



Théorème 8.8

Soient x, y des réels :
- $|x| \in \mathbb{R}^+$, $|x| = |-x|$, $x \leqslant |x|$ *et* $-x \leqslant |x|$.
- $|x| = 0 \iff x = 0$.
- $|xy| = |x||y|$ *et si* $x \neq 0$ *alors* $|\frac{1}{x}| = \frac{1}{|x|}$. *On en déduit que* $\|frac1x| = \frac{1}{|x|}$ *et* $\forall n \in \mathbb{N}$, $|x^n| = |x|^n$.
- $||x| - |y|| \leqslant |x - y| \leqslant |x| + |y|$ *(**inégalité triangulaire**).*

Preuve : Celle-ci est simple et laissée en exercice. □

À retenir

Soient a, b, x trois réels avec b positif :
- $|a| \leqslant b \iff a \leqslant b$ et $-a \leqslant b \iff -b \leqslant a \leqslant b$.
- $|a| \geqslant b \iff a \geqslant b$ ou $-a \geqslant b$.
- $|a - x| \leqslant b \iff -b \leqslant a - x \leqslant b \iff a - b \leqslant x \leqslant a + b$.
- $|a - x| \geqslant b \iff x \geqslant a + b$ ou $x \leqslant a - b$.

Ces inégalités sont importantes, et peuvent se retrouver en raisonnant en termes de distance.

Définition 8.7

*Soit a un réel et $\varepsilon > 0$, on appelle intervalle **ouvert** de **centre a** et de **rayon** ε, l'intervalle $]a - \varepsilon\,; a + \varepsilon[$. C'est l'ensemble des réels x tels que $|x - a| < \varepsilon$. On définit de la même façon l'intervalle fermé de centre a et de rayon ε.*

On rappelle qu'un intervalle ouvert est un intervalle de la forme : $]a; b[$ ou $]a; +\infty[$ ou $]-\infty; b[$. L'ensemble vide et $\mathbb{R}$ sont des intervalles ouverts.

Théorème 8.9

Soit I un intervalle ouvert non vide, pour tout élément a de I il existe au moins un voisinage de a inclus dans I : $\forall\, a \in \mathrm{I}, \exists\, \varepsilon > 0,\ ]a - \varepsilon\,; a + \varepsilon[\, \subset \mathrm{I}$.

Preuve : Il suffit de passer en revue les différents cas pour I. Par exemple, si $\mathrm{I} =]\alpha\,;\beta[$ avec $\alpha < \beta$ (sinon I est vide), on peut prendre $\varepsilon = \min(a - \alpha, \beta - a)$. On remarquera que l'on peut remplacer intervalle ouvert de centre a par intervalle fermé de centre a. □

2) Partie entière

Théorème 8.10

*L'ensemble $\mathbb{R}$ est **archimédien**, c'est à dire :* $\forall\, x, y \in \mathbb{R}^{*+}, \exists\, n \in \mathbb{N}, x \leqslant ny$.

Preuve : Par l'absurde, supposons que $\forall\, n \in \mathbb{N}, x > ny$. Soit $A = \{ny \;/\; n \in \mathbb{N}\}$, A est non vide (contient y) et majoré par x, donc A admet une borne supérieure. Soit $b = \sup(A)$, on a $b - y < b$ donc il existe un entier $n_0 \in \mathbb{N}$ tel que $b - y < n_0 y$, d'où $b < (n_0 + 1)y$ ce qui est absurde car $(n_0 + 1)y \in A$. □

Théorème 8.11 (et définition)

*Soit $x \in \mathbb{R}$, il existe un **unique** entier $n \in \mathbb{Z}$ tel que $n \leqslant x < n + 1$, celui-ci est appelé **partie entière** de x, noté $\lfloor x \rfloor$ (ou $\mathrm{E}(x)$).*

Preuve : Montrons l'existence : si $x = 0$ il suffit de prendre $n = 0$. Supposons x non nul et soit $A = \{n \in \mathbb{Z} \;/\; x < n + 1\}$, on vérifie que A est non vide (si $x < 0$ alors $0 \in A$ et si $x > 0$ on utilise que $\mathbb{R}$ est archimédien), de plus A est minoré par $x - 1$, cet ensemble admet donc une borne inférieure c. On a $c < c + \frac{1}{2}$ donc le réel $c + \frac{1}{2}$ ne minore pas A donc il existe un entier $n_0 \in A$ tel que $c \leqslant n_0 < c + \frac{1}{2}$, si $c < n_0$, alors n_0 ne minore pas A, donc il existe $n_1 \in A$ tel que $c \leqslant n_1 < n_0 < c + \frac{1}{2}$ ce qui est absurde car n_0 et n_1 sont deux entiers distincts dans un intervalle de longueur $\frac{1}{2}$. On en déduit que $c = n_0 \in A$ d'où $n_0 = \min(A)$, mais alors $n_0 - 1 \notin A$, c'est à dire $x \geqslant n_0 - 1 + 1$, par conséquent on a $n_0 \leqslant x < n_0 + 1$.

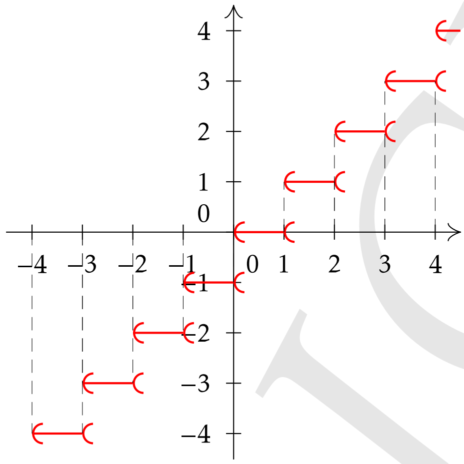
Montrons l'unicité : soient $n, n' \in \mathbb{Z}$ tels que $n \leqslant x < n+1$ et $n' \leqslant x < n'+1$, alors $|n-n'| = |(x-n)-(x-n')| < 1$ car $x - n$ et $x - n'$ sont dans l'intervalle $[0; 1[$, comme n et n' sont entiers, on en déduit que $|n - n'| = 0$ *i.e.* $n = n'$.
□

À retenir

La partie entière de $x \in \mathbb{R}$ est l'unique entier n tel que $n \leqslant x < 1+n$.
La partie entière de $x \in \mathbb{R}$ est l'unique entier n dans l'intervalle $]x-1\,;x]$.
Soit $n \in \mathbb{Z}$, $n = \lfloor x \rfloor \iff \exists \varepsilon \in [0\,;1[,\ x = n+\varepsilon$.

Propriétés

a) La fonction partie entière est une fonction croissante sur $\mathbb{R}$ et elle constante sur tout intervalle de la forme $[n; n+1[$ lorsque $n \in \mathbb{Z}$.



b) La fonction partie entière est continue sur $\mathbb{R} \setminus \mathbb{Z}$. Pour $n \in \mathbb{Z}$, elle est continue à droite mais pas à gauche en n.

c) Pour tout réel x et tout entier n, on a $\lfloor x+n \rfloor = \lfloor x \rfloor + n$.

d) La fonction $x \mapsto x - \lfloor x \rfloor$ est une fonction 1-périodique.

e) La partie entière de x est entièrement caractérisée par : $\begin{cases} \lfloor x \rfloor \in \mathbb{Z} \\ \lfloor x \rfloor \leqslant x < \lfloor x \rfloor + 1 \end{cases}$.

Théorème 8.12

Tout intervalle de la forme $]a; b[$ où $a < b$ contient au moins un rationnel, on dit que $\mathbb{Q}$ est ***dense*** *dans $\mathbb{R}$.*

Preuve : Il existe $q \in \mathbb{N}^*$ tel que $q(b-a) > 1$ (pendre $q = 1 + \left\lfloor \frac{1}{b-a} \right\rfloor$), l'intervalle $]qa; qb[$ a donc une longueur supérieure à 1, il contient donc au moins un entier p (prendre $p = qb-1$ si qb est entier, prendre $p = \lfloor qb \rfloor$ sinon). On a alors $qa < p < qb$ d'où $a < \frac{p}{q} < b$ et $\frac{p}{q} \in \mathbb{Q}$. □

Remarque 8.3 – *Ce théorème traduit que aussi près que l'on veut de n'importe quel réel, on peut trouver des rationnels. De plus la démonstration fournit une méthode de construction de $\frac{p}{q}$.*

Par exemple, avec $x = \sqrt{2}$ et $\varepsilon = 10^{-3}$, on peut prendre $q = 1000$ et $p = \left\lfloor 1000\sqrt{2} \right\rfloor = 1414$ (car $1414^2 \leqslant 2.10^6 < 1415^2$), d'où $\frac{p}{q} = 1,414$ et $|\sqrt{2} - 1,414| < 10^{-3}$.

Théorème 8.13

Tout intervalle $]a; b[$ où $a < b$ contient au moins un irrationnel, donc l'ensemble des irrationnels, $\mathbb{R} \setminus \mathbb{Q}$, est dense dans $\mathbb{R}$.

Preuve : D'après la densité de $\mathbb{Q}$ dans $\mathbb{R}$, il existe un rationnel $r \in]a-\sqrt{2}; b-\sqrt{2}[$, ce qui donne $r+\sqrt{2} \in]a; b[$, et on montre par l'absurde que $r+\sqrt{2}$ est irrationnel. □

Définition 8.8 (Généralisation)

Soit A une partie non vide de $\mathbb{R}$, on dit que A est dense dans $\mathbb{R}$ lorsque tout intervalle ouvert non vide de $\mathbb{R}$ contient au moins un élément de A, ce qui équivaut à :

$$\forall\, x \in \mathbb{R}, \forall\, \varepsilon > 0, \exists\, a \in A, |x-a| < \varepsilon.$$

Remarque 8.4 –
- *Ce qui signifie qu'aussi près que l'on veut de tout réel x, on peut trouver des éléments de A. Voici une autre définition équivalente (et très utile) :*
- *A est dense dans $\mathbb{R}$ ssi pour tout réel x il existe une suite (a_n) d'éléments de A qui converge vers x.*

3) Approximations décimales

Définition 8.9

*Soient a, x, ε trois réels avec $\varepsilon > 0$, on dit que a est une **valeur approchée** de x à ε près lorsque la distance entre a et x et inférieure ou égale à ε : $|a - x| \leqslant \varepsilon$. On dit que a est une valeur approchée de x par défaut (respectivement par excès) à ε près lorsque $a \leqslant x \leqslant a + \varepsilon$ (respectivement $a - \varepsilon \leqslant x \leqslant a$).*

Propriétés

a) Si a est une valeur approchée de x par défaut et b une valeur approchée de x par excès, alors $\frac{a+b}{2}$ est une valeur approchée de x à $\frac{b-a}{2}$ près.

b) Si a est une valeur approchée de x par défaut à ε près et b une valeur approchée de x par excès à ε près, alors $\frac{a+b}{2}$ est une valeur approchée de x à $\frac{\varepsilon}{2}$ près.

Soit $x \in \mathbb{R}$ et soit $n \in \mathbb{N}$, on a $\lfloor x10^n \rfloor \leqslant x10^n < 1 + \lfloor x10^n \rfloor$, en multipliant par 10^{-n} on obtient :

$$\frac{\lfloor x10^n \rfloor}{10^n} \leqslant x < \frac{\lfloor x10^n \rfloor}{10^n} + 10^{-n}.$$

Ce qui signifie que $\frac{\lfloor x10^n \rfloor}{10^n}$ est une valeur approchée de x par défaut à 10^{-n} près, et que $\frac{\lfloor x10^n \rfloor}{10^n} + 10^{-n}$ est une valeur approchée de x par excès à 10^{-n} près. Il faut remarquer que ces deux approximations de x sont des **nombres décimaux** (*i.e.* un entier sur une puissance de dix).

Définition 8.10

*On appelle **approximation décimale** de x par défaut à 10^{-n} près, le nombre : $\frac{\lfloor x10^n \rfloor}{10^n}$.*

Remarque 8.5 – *Posons $r_n = \lfloor 10^n x \rfloor$, alors $10r_n \leqslant 10^{n+1}x < 10r_n + 10$ et donc $10r_n \leqslant \left\lfloor 10^{n+1}x \right\rfloor \leqslant 10r_n + 9$, i.e. $r_{n+1} \in [10r_n; 10r_n + 9]$.*

☞**Exemples** :
- Prenons $x = \sqrt{2}$ et posons $a_n = \frac{\lfloor x10^n \rfloor}{10^n}$
  - $1 \leqslant x^2 < 2^2$, donc $1 \leqslant x < 2$ et $a_0 = \lfloor x \rfloor = 1$ (partie entière de x).
  - $(10x)^2 = 200$ et $14^2 = 196 \leqslant (10x)^2 < 15^2 = 225$, donc $14 \leqslant 10x < 15$ et $a_1 = \lfloor 10x \rfloor / 10 = 14/10 = 1,4$.
  - $(100x)^2 = 20000$ et $141^2 \leqslant (100x)^2 < 142^2$, donc $141 \leqslant 100x < 142$ et $a_2 = \lfloor 100x \rfloor / 100 = 141/100 = 1,41$...etc

  Si on continue le processus, on construit la suite (a_n) des approximations décimales de $\sqrt{2}$ à 10^{-n} près par défaut.

Si on pose pour $n \in \mathbb{N}$, $a_n = \frac{\lfloor x10^n \rfloor}{10^n}$, alors on a l'inégalité $|x - a_n| \leqslant 10^{-n}$, ce qui prouve que la suite (a_n) converge vers x. On a donc une suite de rationnels qui converge vers x, ce qui est une autre façon de prouver la densité de $\mathbb{Q}$ dans $\mathbb{R}$. On remarquera que la suite $(a_n + 10^{-n})$ (valeurs approchées décimales par excès) converge également vers x.

Théorème 8.14

Soit $x \in \mathbb{R}$ et $a_n = \frac{\lfloor x10^n \rfloor}{10^n}$, pour $n \in \mathbb{N}^*$ on pose $d_n = 10^n(a_n - a_{n-1})$, alors d_n est un entier compris entre 0 et 9.

Preuve : $10^n a_n = \lfloor 10^n x \rfloor \leqslant 10^n x < 1 + \lfloor 10^n x \rfloor$, d'autre part $10^n a_{n-1} = 10 \left\lfloor 10^{n-1}x \right\rfloor \leqslant 10^n x < 10 + 10 \left\lfloor 10^{n-1}x \right\rfloor$, d'où $-10 - \left\lfloor 10^{n-1}x \right\rfloor < -10^n x \leqslant -10^n a_{n-1}$, on en déduit que $d_n - 10 < 0 < d_n + 1$, par conséquent $0 \leqslant d_n < 10$, or d_n est un entier, donc $d_n \leqslant 9$. □

Définition 8.11

Pour $n \geqslant 1$, *l'entier* $d_n = 10^n(a_n - a_{n-1}) = \lfloor 10^n x \rfloor - 10\left\lfloor 10^{n-1}x \right\rfloor$ *est appelé* n^e ***décimale*** *de* x.

Remarquons que $d_n 10^{-n} = a_n - a_{n-1}$, ce qui entraîne que $a_0, d_1 d_2 \dots d_n = a_0 + \sum_{k=1}^{n} d_k 10^{-k} = a_n$, or la suite (a_n) converge vers x, on écrit alors :

$$x = a_0 + \sum_{k=1}^{+\infty} d_k 10^{-k} = a_0, d_1 d_2 \cdots \text{ (développement décimal infini de } x\text{)}$$

IV SOLUTION DES EXERCICES

Solution 8.1 $\inf(B)$ *est un minorant de* B *donc un minorant de* A, *par conséquent* $\inf(B) \leqslant \inf(A)$ *car* $\inf(A)$ *est le plus grand des minorants de* A. *De même*, $\sup(B)$ *majore* B , *donc majore* A *également, d'où* $\sup(A) \leqslant \sup(B)$ *car* $\sup(A)$ *est le plus petit des majorants de* A.

Solution 8.2 $\sup(A) + \sup(B)$ *majore* $A + B$, *donc* $A + B$ *admet une borne sup. et* $\sup(A + B) \leqslant \sup(A) + \sup(B)$. *Soient* $a \in A$ *et* $b \in B$, $a + b \leqslant \sup(A + B)$, *donc* $a \leqslant \sup(A + B) - b$, *ce qui signifie que* A *est majoré par* $\sup(A + B) - b$, *d'où* $\sup(A) \leqslant \sup(A + B) - b$, *mais alors* $b \leqslant \sup(A + B) - \sup(A)$, *donc* B *est majoré par* $\sup(A + B) - \sup(A)$, *d'où* $\sup(B) \leqslant \sup(A + B) - \sup(A)$ *et finalement* $\sup(A) + \sup(B) \leqslant \sup(A + B)$ *ce qui prouve bien l'égalité.*

Solution 8.3 *Il est clair que* A *est minoré par* -1 *et majoré par* 2. *Les éléments positifs de* A *sont de la forme* $1 + \frac{1}{2k}$ *(n pair), on voit que* A *possède un maximum et que celui-ci vaut* $\frac{3}{2}$, *donc* $\sup(A) = \frac{3}{2}$.
Les éléments négatifs de A *sont de la forme* $-1 + \frac{1}{2k+1}$ *(n impair), ceux-ci peuvent être très proches de* -1 *mais ne prennent jamais la valeur* -1. *Soit* $\varepsilon > 0$, $-1 + \frac{1}{2k+1} > -1 + \varepsilon \iff 2k + 1 > \frac{1}{\varepsilon}$, *on voit donc qu'il est possible de trouver une élément de* A *strictement inférieur à* $-1 + \varepsilon$ *qui ne ne minore donc pas* A, *donc* $\inf(A) = -1$.

Chapitre 9

Suites numériques

Sommaire

I SUITES RÉELLES, GÉNÉRALITÉS

1) Définitions

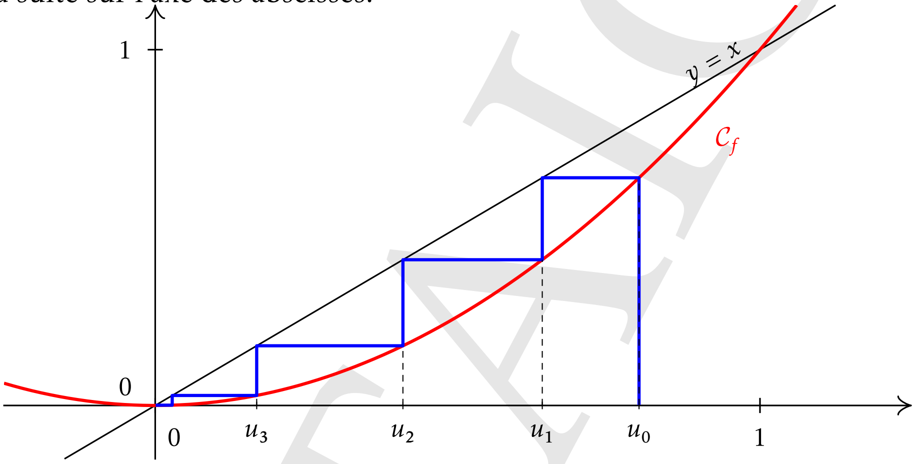
Définition 9.1

Une suite numérique u est une application de A vers $\mathbb{R}$: $u : A \to \mathbb{R}$, où A est une partie de $\mathbb{N}$. Par convention le réel $u(n)$ est noté u_n, et la suite u est parfois notée $(u_n)_{n\in A}$. Si la partie A est finie, on dit que la suite u est une suite finie. L'ensemble des suites réelles définies sur A est donc l'ensemble des applications de A vers $\mathbb{R}$, c'est à dire $\mathcal{F}(A, \mathbb{R})$.

Remarque 9.1 – *On prendra garde à ne pas confondre u_n qui est un réel (terme de rang n) avec $(u_n)_{n\in A}$ qui désigne la suite u. Les suites finies présentant peu d'intérêt, on étudiera seulement le cas où A est une partie infinie de $\mathbb{N}$. On peut alors montrer qu'il est toujours possible de se ramener au cas où $A = \mathbb{N}$, si bien que dans la suite de ce chapitre on étudiera $\mathcal{F}(\mathbb{N},\mathbb{R})$ l'ensemble des suites réelles définies sur $\mathbb{N}$.*

☞**Exemples** :

- Une suite u est **arithmétique** si et seulement si il existe un réel r (appelé **raison**), tel que $\forall n \in \mathbb{N}, u_{n+1} = u_n + r$. On a alors les formules suivantes : $\forall n, p \in \mathbb{N}, u_n = u_p + (n-p)r$. La somme de n termes consécutifs est $S = \frac{n(p+d)}{2}$ où p désigne le premier terme, et d le dernier.
- Une suite u est **géométrique** si et seulement si il existe $q \in \mathbb{R}$ (appelé **raison**), tel que $\forall n \in \mathbb{N}, u_{n+1} = qu_n$. On a alors les formules suivantes : $\forall n, p \in \mathbb{N}, u_n = u_p q^{n-p}$. La somme de n termes consécutifs est $S = \begin{cases} np \text{ si } q = 1 \\ \dfrac{p-qd}{1-q} \text{ si } q \neq 1 \end{cases}$, où p désigne le premier terme, d le dernier et q la raison.
- Suites récurrentes à un pas : ce sont les suites u définies par : $u_0 \in \mathbb{R}$, et $\forall n \in \mathbb{N}, u_{n+1} = f(u_n)$, où $f : I \to \mathbb{R}$ est une fonction donnée. Par exemple : $u_0 = \frac{1}{2}$ et $u_{n+1} = u_n^2$. Dans le plan, à l'aide de la courbe représentative de f et de la première bissectrice, on peut construire géométriquement les termes de la suite sur l'axe des abscisses.



- Suites récurrentes à deux pas : par exemple la suite de *Fibonacci*[1] qui est définie par : $u_0 = u_1 = 1$ et $\forall n \in \mathbb{N}, u_{n+2} = u_{n+1} + u_n$.

2) Vocabulaire

- Sens de variation : soit u une suite réelle et p un entier, on dit que la suite u est :
  - croissante à partir du rang p lorsque : $\forall n \geqslant p, u_n \leqslant u_{n+1}$.
  - strictement croissante à partir du rang p lorsque : $\forall n \geqslant p, u_n < u_{n+1}$.
  - décroissante à partir du rang p lorsque : $\forall n \geqslant p, u_{n+1} \leqslant u_n$.
  - strictement décroissante à partir du rang p lorsque : $\forall n \geqslant p, u_{n+1} < u_n$.
  - constante (ou stationnaire) à partir du rang p lorsque : $\forall n \geqslant p, u_{n+1} = u_n$.
  - monotone lorsque u est croissante ou bien décroissante.
  - strictement monotone lorsque u est strictement croissante ou bien strictement décroissante.

  Remarque 9.2 – *Étudier le sens de variation de u peut se faire en étudiant le signe de $u_{n+1} - u_n$, ou encore le signe de $f(u_{n+1}) - f(u_n)$ où f désigne une fonction monotone.*
- Suite bornée : on dit qu'une suite réelle u est :
  - majorée lorsque : $\exists M \in \mathbb{R}, \forall n \in \mathbb{N}, u_n \leqslant M$.
  - minorée lorsque : $\exists m \in \mathbb{R}, \forall n \in \mathbb{N}, m \leqslant u_n$.
  - bornée lorsque : $\exists m, M \in \mathbb{R}, \forall n \in \mathbb{N}, m \leqslant u_n \leqslant M$ (*i.e.* minorée et majorée).

  Remarque 9.3 – *Une suite u est bornée si et seulement si il existe un réel M positif tel que $\forall n \in \mathbb{N}, |u_n| \leqslant M$.*

  Par exemple, la suite $(u_n = \sin(n))$ est bornée, la suite $(v_n = n^2)$ est minorée mais non majorée, la suite $(w_n = (-2)^n)$ est ni minorée ni majorée.

1. *FIBONACCI Leonardo* (1180 – 1250 (environ)) : mathématicien italien (de son vrai nom *Leonardo da Pisa*) qui œuvra pour l'introduction de nombres arabes en Occident.

- Suite périodique : on dit qu'une suite u est p-périodique (où $p \in \mathbb{N}^*$) à partir du rang n_0 lorsque : $\forall n \geqslant n_0$, $u_{n+p} = u_n$. Par exemple, la suite $(u_n) = (-1)^n)$ est 2-périodique, la suite w définie par $w_0 = 1$, $w_1 = 1$ et pour tout n $w_{n+2} = -w_{n+1} - w_n$, est 3-périodique, mais la suite des décimales de π n'est pas périodique car π est irrationnel.
- Suite extraite : soit u une suite réelle et soit $\sigma : \mathbb{N} \to \mathbb{N}$ une **application strictement croissante**, alors la suite v définie par $v_n = u_{\sigma(n)}$ est appelée suite extraite de u (σ étant l'extraction). On remarquera que l'on a : $\forall n \in \mathbb{N}$, $n \leqslant \sigma(n)$. Par exemple, la suite (u_{2n}) est une suite extraite de u, c'est la suite des termes de rangs pairs, de même la suite (u_{2n+1}) est extraite de u, c'est la suite des termes de rangs impairs.

3) Opérations sur les suites

Soient u et v deux suites réelles et soit $\lambda \in \mathbb{R}$, on définit les suites :
- $u + v$: en posant pour tout $n \in \mathbb{N}$, $(u+v)_n = u_n + v_n$;
- $u \times v$: en posant $(u \times v)_n = u_n v_n$.
- λv : en posant $(\lambda v)_n = \lambda v_n$.
- $\frac{1}{v}$: si v ne s'annule pas à partir d'un certain rang n_0, en posant : $(\frac{1}{v})_n = \frac{1}{v_n}$.

On vérifie alors que :
- $(\mathcal{F}(\mathbb{N}, \mathbb{R}), +)$ est un groupe commutatif. Son élément neutre est la suite nulle (notée 0) et l'opposé d'une suite u est la suite $(-u_n)_{n\in\mathbb{N}}$ (notée $-u$).
- La multiplication est associative, commutative, admet comme élément neutre la suite constante $(u_n = 1)_{n\in\mathbb{N}}$ (notée 1), et elle est distributive sur l'addition. Mais il y a des suites non nulles qui n'ont pas d'inverse, par exemple la suite u définie par $u_n = 1 + (-1)^n$. Seules les suites u qui ne s'annulent jamais ont un inverse, et cet inverse est la suite $\frac{1}{u}$.

L'ensemble $(\mathcal{F}(\mathbb{N}, \mathbb{R}), +, \times)$ n'est donc pas un corps, mais seulement un **anneau commutatif**. Les deux suites u et v définies par $u_n = 1 + (-1)^n$ et $v_n = 1 - (-1)^n$ sont non nulles, mais leur produit est la suite nulle, ceci prouve que $(\mathcal{F}(\mathbb{N}, \mathbb{R}), +, \times)$ est un anneau **non intègre**.

II SUITES CONVERGENTES

1) Définition

Définition 9.2

Soit u une suite réelle et $\ell \in \mathbb{R}$, on dit que u admet comme limite ℓ lorsque u_n peut être aussi proche (ou voisin) que l'on veut de ℓ pourvu que n soit assez grand, c'est à dire :

$$\forall \varepsilon > 0, \exists \mathrm{N} \in \mathbb{N}, \forall n \in \mathbb{N}, n \geqslant \mathrm{N} \implies |u_n - \ell| < \varepsilon.$$

Notation : $\lim u = \ell$ ou $\lim u_n = \ell$ ou $u_n \to \ell$.

Remarque 9.4 –
- *Comme $|u_n - \ell| = |(u_n - \ell) - 0| = ||u_n - \ell| - 0|$, on a :*

$$\lim u_n = \ell \iff \lim u_n - \ell = 0 \iff \lim |u_n - \ell| = 0.$$

- *Comme $||u_n| - |\ell|| \leqslant |u_n - \ell|$, on a : $\lim u_n = \ell \implies \lim |u_n| = |\ell|$ (réciproque fausse).*

Théorème 9.1

Si à partir d'un certain rang on a : $|u_n - \ell| \leqslant v_n$, et si $v_n \to 0$, alors $\lim u_n = \ell$.

Preuve : Soit $\varepsilon >$, à partir d'un rang N_1 on a $|v_n| < \varepsilon$, et à partir d'un rang N_2 on a $|u_n - \ell| \leqslant v_n$, donc à partir du rang $\mathrm{Max}(\mathrm{N}_1, \mathrm{N}_2)$ on a $|u_n - \ell| < \varepsilon$. □

Définition 9.3

Lorsque la suite u admet une limite ***finie****, on dit que u est* ***convergente****, sinon on dit qu'elle est* ***divergente****.*

☞**Exemples** :

- Toute suite stationnaire (à partir d'un certain rang) est convergente.
- Soit $x \in \mathbb{R}$ et $v_n = \frac{\lfloor nx \rfloor}{n}$: on a $v_n \to x$. Soit $\varepsilon > 0$, $|v_n - x| = \frac{nx - \lfloor nx \rfloor}{n} < \frac{1}{n} < \varepsilon \iff n > \frac{1}{\varepsilon}$, il suffit donc de prendre $\mathrm{N} = 1 + \left\lfloor \frac{1}{\varepsilon} \right\rfloor$ pour avoir : $n \geqslant \mathrm{N} \implies |v_n - x| < \varepsilon$.
- $u_n = q^n$ avec $q = 1$: la suite est constante et $u_n \to 1$.
- $u_n = q^n$ avec $|q| < 1$ et $q \neq 0$: alors $q^n \to 0$. Soit $\varepsilon > 0$, comme $\frac{1}{|q|} > 1$, on a $\frac{1}{|q|} = 1 + p$ avec $p > 0$, on peut montrer alors que $\forall n \in \mathbb{N}$, $\frac{1}{|q|^n} \geqslant 1 + np$ (récurrence ou binôme de *Newton*), on a $\frac{1}{|q|^n} > \frac{1}{\varepsilon}$ dès que $1 + np > \frac{1}{\varepsilon}$ c'est à dire dès que $n \geqslant \mathrm{N} = 1 + \left\lfloor \frac{1}{p\varepsilon} - \frac{1}{p} \right\rfloor$, donc $n \geqslant \mathrm{N} \implies |q^n| < \varepsilon$.
- $u_n = (-1)^n$ alors la suite est divergente (2-périodique). Supposons qu'elle ait une limite finie ℓ alors à partir d'un certain rang N on aura $|u_n - \ell| < \frac{1}{3}$ par conséquent les valeurs -1 et 1 sont dans l'intervalle $]\ell - \frac{1}{3}; \ell + \frac{1}{3}[$ ce qui est absurde.

★**Exercice 9.1** *Montrer qu'une suite d'entiers convergente est stationnaire.*

2) Premières propriétés

Soit u une suite réelle :

- Si u admet une limite $\ell \in \mathbb{R}$, alors celle-ci est unique.
  Preuve : Supposons $u_n \to \ell$ et $u_n \to \ell'$ avec $\ell < \ell'$, Soit $\alpha \in]\ell; \ell'[$, $\varepsilon = \alpha - \ell$ et $\varepsilon' = \ell' - \alpha$, alors à partir d'un certain rang N on a $|u_n - \ell| < \varepsilon$, ce qui donne $u_n < \alpha$, et à partir d'un certain rang N′ on a $|u_n - \ell'| < \varepsilon'$, ce qui donne $\alpha < u_n$, donc à partir de $\max(\mathrm{N}, \mathrm{N}')$ on a une contradiction, donc $\ell = \ell'$. □

  On a démontré au passage :
- Si u converge vers ℓ et si $\alpha < \ell$, alors à partir d'un certain rang $\alpha < u_n$. De même, si $\alpha > \ell$, alors à partir d'un certain rang on a $\alpha > u_n$.
- Si u est convergente, alors u est bornée (la réciproque est fausse).
  Preuve : Si $u_n \to \ell \in \mathbb{R}$, il existe un entier N tel que $n \geqslant \mathrm{N} \implies |u_n - \ell| < 1$, ce qui entraîne $|u_n| < |\ell| + 1$. On a alors pour tout entier n : $|u_n| \leqslant \max(|u_0|, \ldots, |u_\mathrm{N}|, 1 + |\ell|)$. Pour voir que la réciproque est fausse, on peut considérer la suite u définie par $u_n = (-1)^n$, elle est bornée mais non convergente. □

  Conséquence : la suite (q^n) avec $|q| > 1$ est divergente car non bornée, en effet : $|q| = 1 + p$ avec $p > 0$ donc $|q^n| \geqslant 1 + np$ qui peut être aussi grand que l'on veut.
- Si u converge vers ℓ, alors toutes les suites extraites de u convergent vers ℓ.
  Preuve : Soit $v_n = u_{\sigma(n)}$ une suite extraite de u et supposons $u_n \to \ell \in \overline{\mathbb{R}}$. Soit W un voisinage de ℓ, il existe un entier N tel que $n \geqslant \mathrm{N} \implies u_n \in \mathrm{W}$. Mais σ étant strictement croissante, on a $\forall n \in \mathbb{N}$, $n \leqslant \sigma(n)$, donc $n \geqslant \mathrm{N} \implies \sigma(n) \geqslant \mathrm{N}$, mais alors $u_{\sigma(n)} \in \mathrm{W}$, c'est à dire $n \geqslant \mathrm{N} \implies v_n \in \mathrm{W}$ et donc $v_n \to \ell$. □

  Remarque 9.5 – *Cette propriété est souvent utilisée pour montrer qu'une suite u n'a pas de limite. Soit en trouvant une suite extraite qui diverge, soit en trouvant deux suites extraites qui ne convergent pas vers la même limite. Par exemple :* $u_n = \cos((n + \frac{1}{n})\pi)$.
- Si $\lim u_{2n} = \lim u_{2n+1} = \ell \in \mathbb{R}$, alors $\lim u = \ell$.
  Preuve : Soit $\varepsilon > 0$, il existe un entier N_1 tel que $k \geqslant \mathrm{N}_1 \implies |u_{2k} - \ell| < \varepsilon$, de même il existe un entier N_2 tel que $k \geqslant \mathrm{N}_2 \implies |u_{2k+1} - \ell| < \varepsilon$. Posons $\mathrm{N} = \max(2\mathrm{N}_1, 2\mathrm{N}_2 + 1)$, si $n \geqslant \mathrm{N}$ alors lorsque $n = 2k$ on a $k \geqslant \mathrm{N}_1$ et donc $|u_n - \ell| < \varepsilon$, lorsque $n = 2k + 1$ on a $k \geqslant \mathrm{N}_2$ et donc $|u_n - \ell| < \varepsilon$, finalement dès que $n \geqslant \mathrm{N}$ on a $|u_n - \ell| < \varepsilon$ et donc $u_n \to \ell$. □

3) Convergence et opérations

Théorème 9.2

Soient u et v deux suites qui convergent respectivement vers ℓ et ℓ', et soit $\lambda \in \mathbb{R}$ alors :
- *$(u_n + v_n)$ converge vers $\ell + \ell'$.*
- *(λu_n) converge vers $\lambda\ell$.*

Preuve : Soit $\varepsilon > 0$, il existe un entier N à partir duquel on a $|u_n - \ell| < \varepsilon/2$ et $|v_n - \ell'| < \varepsilon/2$, mais alors on a $|u_n + v_n - (\ell + \ell')| \leqslant |u_n - \ell| + |v_n - \ell'| < \varepsilon$, donc $u_n + v_n \to \ell + \ell'$.

Soit $\lambda \neq 0$, et soit $\varepsilon > 0$, à partir d'un certain rang on a $|u_n - \ell| < \frac{\varepsilon}{|\lambda|}$ d'où $|\lambda u_n - \lambda\ell| < \varepsilon$. □

Théorème 9.3

Si (u_n) converge vers ℓ et (v_n) vers ℓ' alors :
- *(u_nv_n) converge vers $\ell\ell'$.*
- *Si $\ell \neq 0$, alors à partir d'un certain rang les termes u_n sont non nuls et la suite $(\frac{1}{u_n})$ converge vers $\frac{1}{\ell}$.*

Preuve : $|u_nv_n - \ell\ell'| = |(u_n-\ell)v_n + \ell(v_n-\ell')| \leqslant |u_n-\ell||v_n| + |\ell||v_n-\ell'|$, mais la suite v est bornée donc il existe un réel M strictement positif tel que $|v_n| \leqslant \mathrm{M}$ et donc $|u_nv_n - \ell\ell'| < |u_n-\ell|\mathrm{M} + |\ell||v_n-\ell'|$, mais d'après le théorème précédent la deuxième suite tend vers 0, donc $u_nv_n \to \ell\ell'$.

La suite $(|u_n|)$ converge vers $|\ell| > 0$ donc à partir d'un certain rang on a $|u_n| > \frac{|\ell|}{2} > 0$, donc $u_n \neq 0$ et alors : $|\frac{1}{u_n} - \frac{1}{\ell}| = \frac{|\ell-u_n|}{|\ell u_n|} < \frac{2|\ell-u_n|}{\ell^2}$, or cette deuxième suite tend vers 0, donc $\frac{1}{u_n} \to \frac{1}{\ell}$. □

4) Convergence et relation d'ordre

Théorème 9.4

*Soient u, et v deux suites réelles. Si u converge vers ℓ, v converge vers ℓ', et si à partir d'un certain rang on a $u_n \leqslant v_n$, alors $\ell \leqslant \ell'$ (c'est le théorème du **passage à la limite**).*

Preuve : Supposons $\ell > \ell'$, alors il existe $\alpha \in]\ell';\ell[$ donc à partir d'un certain rang on doit avoir $u_n > \alpha$ et $v_n < \alpha$ ce qui est contradictoire, donc $\ell \leqslant \ell'$. □

Remarque 9.6 – *Pour le passage à la limite on peut avoir $u_n < v_n$ et $\ell = \ell'$, par exemple en prenant $u_n = 1 - \frac{1}{n}$ et $v_n = 1 + \frac{1}{n}$, donc dans un passage à la limite les inégalités deviennent larges.*

Théorème 9.5

*Soient u, v et w trois suites réelles. Si u et v convergent vers ℓ et si à partir d'un certain rang on a $u_n \leqslant w_n \leqslant v_n$, alors w converge vers ℓ (c'est le théorème **des gendarmes ou de l'étau**).*

Preuve : Soit $\varepsilon > 0$, il existe un entier N à partir duquel on a $u_n \leqslant w_n \leqslant v_n$ avec $u_n, v_n \in]\ell-\varepsilon;\ell+\varepsilon[$, donc $w_n \in]\ell-\varepsilon;\ell+\varepsilon[$ à partir du rang N, donc $w_n \to \ell$. □

Théorème 9.6

Soient u et v deux suites réelles. Si u converge vers 0 et si v est bornée, alors $\lim u \times v = 0$.

Preuve : Il existe un réel positif M tel que $|v_n| \leqslant \mathrm{M}$ pour tout n, d'où $|u_nv_n| \leqslant \mathrm{M}|u_n|$, c'est à dire $-\mathrm{M}|u_n| \leqslant u_nv_n \leqslant \mathrm{M}|u_n|$, on peut donc conclure que $u_nv_n \to 0$. □

Déterminer la limite des suites (si elle existe) :

☞**Exemples** :
- $a_n = \frac{\sin(n)}{n}$ $\quad b_n = \frac{n}{2n+(-1)^n}$ $\quad c_n = \sum_{k=1}^{n} \frac{1}{n+\sqrt{k}}$ $\quad d_n = n - \sqrt{n}$
- $e_n = \frac{n^3-1}{n^2+1}$ $\quad f_n = \sqrt{n^2+n+1} - n$ $\quad g_n = \left(1+\frac{1}{n}\right)^n$

5) Caractérisations séquentielles

Théorème 9.7 (caractérisation séquentielle de la borne supérieure)

Soit A une partie non vide de $\mathbb{R}$ et soit $\mathrm{M} \in \mathbb{R}$, on a :
M est la borne supérieure (inférieure) de A si et seulement si M majore (minore) A et il existe une suite d'éléments de A qui converge vers M.

Preuve : Si $\mathrm{M} = \sup \mathrm{A}$, alors pour tout $n \in \mathbb{N}$, $\exists a_n \in \mathrm{A}$ tel que $\mathrm{M} - \frac{1}{n+1} < a_n$ car $\mathrm{M} - \frac{1}{n+1}$ ne majore pas A, la suite (a_n) ainsi construite converge vers M car $\mathrm{M} - \frac{1}{n+1} < a_n \leqslant \mathrm{M}$.

Si M majore A et qu'il existe une suite (a_n) de A qui converge vers M, alors pour tout $\varepsilon > 0$, à partir d'un certain rang N on a $\mathrm{M} - \varepsilon < a_n$, donc $\mathrm{M} - \varepsilon$ ne majore pas A, M est donc le plus petit majorant de A. □

Théorème 9.8 (caractérisation séquentielle de la densité)

Soit A une partie non vide de $\mathbb{R}$, A est dense dans $\mathbb{R}$ si et seulement si pour tout réel x il existe une suite d'éléments de A qui converge vers x.

Preuve : Celle-ci est simple et laissée en exercice. □

III SUITES AYANT UNE LIMITE INFINIE

1) Définition

Définition 9.4

Soit u une suite réelle :
- *on dit que u admet comme limite $+\infty$ lorsque u_n peut être aussi grand que l'on veut pourvu que n soit assez grand, c'est à dire :* $\forall A \in \mathbb{R}, \exists N \in \mathbb{N}, \forall n \in \mathbb{N}, n \geqslant N \implies u_n > A$.
  Notation : $\lim u = +\infty$ *ou* $\lim u_n = +\infty$ *ou* $u_n \to +\infty$.
- *on dit que u admet comme limite $-\infty$ lorsque u_n peut être aussi petit que l'on veut pourvu que n soit assez grand, c'est à dire :* $\forall A \in \mathbb{R}, \exists N \in \mathbb{N}, \forall n \in \mathbb{N}, n \geqslant N \implies u_n < A$.
  Notation : $\lim u = -\infty$ *ou* $\lim u_n = -\infty$ *ou* $u_n \to -\infty$.

Remarque 9.7 –
- *Si $u_n \to +\infty$ alors u n'est pas majorée (mais elle est minorée).*
- *Si $u_n \to -\infty$ alors u n'est pas minorée (mais elle est majorée).*
- *On a l'équivalence :* $\lim u_n = -\infty \iff \lim -u_n = +\infty$.

☞**Exemple** : Si $q > 1$ alors $\lim q^n = +\infty$.

Comme pour les suites convergentes, on peut montrer :
- Si u admet une limite infinie, alors toutes les suites extraites de u ont la même limite que u.
- Si $u_{2n} \to +\infty$ et $u_{2n+1} \to +\infty$, alors $u_n \to +\infty$.

2) Limite infinie et ordre

Théorème 9.9

Soient u et v deux suites réelles :
- *Si* $\lim u = +\infty$ *et si à partir d'un certain rang on a* $u_n \leqslant v_n$*, alors* $\lim v = +\infty$.
- *Si* $\lim v = -\infty$ *et si* $u_n \leqslant v_n$ *à partir d'un certain rang, alors* $\lim u = -\infty$.
- *Si* $\lim u = +\infty$ *(respectivement $-\infty$) et si v est minorée (respectivement majorée), alors* $\lim u + v = +\infty$ *(respectivement $-\infty$).*

Preuve : Pour le premier point : il existe un entier N_1 à partir duquel on a $u_n \leqslant v_n$, soit A un réel, il existe un entier N_2 à partir duquel on a $A < u_n$, donc si $n \geqslant \max(N_1, N_2)$ alors $A < v_n$, donc $v_n \to +\infty$.

Pour le deuxième point : on peut appliquer le précédent aux suites $-u$ et $-v$.

Pour le troisième point : supposons $u_n \to +\infty$ et v minorée par un réel m, alors pour tout entier n on a $m + u_n \leqslant u_n + v_n$, or la suite $(m + u_n)$ tend vers $+\infty$, on peut donc appliquer le premier point, *i.e.* $u_n + v_n \to +\infty$. Dans l'autre cas on peut raisonner sur les suites $-u$ et $-v$. □

3) Limite infinie et opérations

Théorème 9.10

Soient u et v deux suites de limites respectives ℓ et ℓ' dans $\overline{\mathbb{R}}$, et soit $\lambda \in \mathbb{R}$.
- $\lim u + v = \ell + \ell'$ *sauf si* $\ell = +\infty$ *et* $\ell' = -\infty$ *(ou l'inverse).*
- $\lim \lambda u = \lambda \ell$ *(si $\lambda = 0$ alors la suite λu est nulle).*
- $\lim u \times v = \ell \ell'$ *sauf si* $\ell = 0$ *et* $\ell' = \pm\infty$ *(ou l'inverse).*

– *Si à partir d'un certain rang la suite u ne s'annule pas, alors la suite $\frac{1}{u}$:*

$$\begin{cases} \text{tend vers } \frac{1}{\ell} & \text{si } \ell \in \mathbb{R}^* \\ \text{tend vers } 0 & \text{si } \ell = \pm\infty \\ \text{tend vers } +\infty & \text{si } \ell = 0 \text{ et } u > 0 \\ \text{tend vers } -\infty & \text{si } \ell = 0 \text{ et } u < 0 \\ \text{n'a pas de limite dans les autres cas} & \end{cases}.$$

Preuve : Pour la somme : prenons par exemple le cas $\ell \in \mathbb{R}$ et $\ell' = +\infty$, la suite u_n est minorée par un certain réel m (car convergente) d'après le paragraphe précédent, $u_n + v_n \to +\infty$. Les autres cas non indéterminés se ramènent à celui-ci. Pour la forme indéterminée, on peut considérer les exemples suivants : $(n + (-n + a))$ qui converge, $(n + (-\frac{n}{2}))$ qui tend vers $+\infty$, $(n + (-2n))$ qui tend vers $-\infty$, et $(n + (-n + (-1)^n))$ qui n'a pas de limite.

Pour λu : il suffit de considérer le cas $\lambda > 0$ et $u_n \to +\infty$ (laissé en exercice). Les autres cas se ramènent à celui-ci.

Pour le produit : prenons par exemple le cas où ℓ est un réel strictement positif et $\ell' = +\infty$, alors à partir d'un certain rang, on a $v_n > 0$ et $u_n > \frac{\ell}{2}>0$, d'où $u_nv_n > v_n\frac{\ell}{2}$, or $v_n\frac{\ell}{2}$ tend vers $+\infty$, et donc u_nv_n aussi. Les autres cas non indéterminés se ramènent à celui-ci. Pour la forme indéterminée, on peut considérer les exemples suivants : $(\frac{a}{n} \times n)$ qui converge, $(n \times \frac{1}{\sqrt{n}})$ qui tend vers $+\infty$, $(-n \times \frac{1}{\sqrt{n}})$ qui tend vers $-\infty$, et $(n \times \frac{(-1)^n}{n})$ qui n'a pas de limite.

Pour l'inverse : supposons que $\ell = 0$ et $u > 0$, soit A un réel et $\varepsilon = \frac{1}{1+|A|}$, il existe un entier N à partir duquel on a $|u_n| < \varepsilon$, c'est à dire en fait, $0 < u_n < \varepsilon$ et donc $A < 1 + |A| = \frac{1}{\varepsilon} < \frac{1}{u_n}$, par conséquent $\frac{1}{u_n} \to +\infty$. Les autres cas non indéterminés se ramènent à celui-ci. Pour terminer prenons la suite $u_n = \frac{(-1)^n}{n}$, son inverse est la suite $((-1)^n n)$ et cette suite n'a pas de limite (distinguer les termes de rangs pairs et les termes de rangs impairs). □

IV THÉORÈMES D'EXISTENCE D'UNE LIMITE

1) Suites monotones

Théorème 9.11

Si u est une suite croissante majorée (respectivement décroissante minorée), alors(u_n) converge vers $\sup\limits_{n\in\mathbb{N}} u_n$ (respectivement vers $\inf\limits_{n\in\mathbb{N}} u_n$).
Si u est une suite croissante non majorée (respectivement décroissante non minorée), alors (u_n) tend vers $+\infty$ (respectivement vers $-\infty$).

Preuve : Supposons u croissante majorée, soit $\ell = \sup\limits_{n\in\mathbb{N}} u_n$, et soit $\varepsilon > 0$, alors il existe un entier N tel que $u_N > \ell - \varepsilon$ (car $\ell - \varepsilon$ ne majore pas la suite). Si $n \geqslant N$ alors, la suite étant croissante, $\ell - \varepsilon < u_N \leqslant u_n \leqslant \ell < \ell + \varepsilon$ et donc $|u_n - \ell| < \varepsilon$, ce qui prouve que $u_n \to \ell$.

Lorsque u est croissante non majorée : soit A un réel, alors il existe un entier N tel que $u_N > A$ (A ne majore pas la suite), si $n \geqslant N$ alors $A < u_N \leqslant u_n$, donc $u_n \to +\infty$. □

Conséquences :

a) Si (u_n) est croissante majorée, alors $u_n \to \ell = \sup u_n \in \mathbb{R}$ et donc $\forall n \in \mathbb{N}, u_n \leqslant \ell$. En fait si u est strictement croissante, alors $\forall n \in \mathbb{N}, u_n < \ell$ (car s'il y avait l'égalité au rang N, alors la suite serait constante à partir de l'indice N).

b) Si (u_n) est décroissante minorée, alors $u_n \to \ell = \inf u_n \in \mathbb{R}$ et donc $\forall n \in \mathbb{N}, u_n \geqslant \ell$. En fait si u est strictement décroissante, alors $\forall n \in \mathbb{N}, u_n > \ell$ (car s'il y avait l'égalité au rang N, alors la suite serait constante à partir de l'indice N).

c) Une suite monotone est donc convergente si et seulement si elle est bornée.

☞**Exemples** :

– Soit u la suite définie par : $u_n = \sum\limits_{k=1}^{n} \frac{1}{k^2}$. Cette suite est croissante ($u_{n+1} - u_n > 0$), en remarquant que pour $k \geqslant 2$ on a $\frac{1}{k^2} < \frac{1}{k(k-1)} = \frac{1}{k} - \frac{1}{k-1}$, on voit que $u_n < 2$, la suite u est donc convergente (de limite $\frac{\pi^2}{6}$).

– Soit v la suite définie par $v_0 = 1$ et $\forall n \in \mathbb{N}, v_{n+1} = \sin(v_n)$. Il s'agit d'une suite récurrente, la représentation graphique des premiers termes suggère que la suite est décroissante minorée par 0, ce qui est facile à vérifier par récurrence. La suite v est donc convergente de limite ℓ, la fonction sinus étant continue, on a $\sin(v_n) \to \sin(\ell)$, c'est à dire $v_{n+1} \to \sin(\ell)$, donc $\ell = \sin(\ell)$.

L'étude de la fonction $x \mapsto \sin(x) - x$ montre que l'unique solution de $\sin(x) = x$ est 0, donc $\ell = 0$, *i.e.* $v_n \to 0$.

2) Suites adjacentes

Définition 9.5
Soient u et v deux suites, on dit qu'elles sont adjacentes lorsque l'une est croissante, l'autre décroissante et $\lim u_n - v_n = 0$.

☞**Exemple** : Soient u et v les suites définies par : $u_n = \sum\limits_{k=0}^{n} \frac{1}{k!}$ et $v_n = u_n + \frac{1}{n \times n!}$, ces deux suites sont adjacentes.

Théorème 9.12
Deux suites adjacentes sont nécessairement convergentes et convergent vers la même limite.

Preuve : Supposons u croissante, v décroissante, et $\lim u_n - v_n = 0$. Soit $w_n = v_n - u_n$, alors $w_{n+1} - w_n = (v_{n+1} - v_n) - (u_{n+1} - u_n) \leqslant 0$, donc la suite w est décroissante, or $\lim w_n = 0$, donc $\forall n \in \mathbb{N}, w_n \geqslant 0$, *i.e.* $u_n \leqslant v_n$. Mais alors u est majorée par v_0 et v est minorée par u_0, donc u et v sont convergentes : $u_n \to \ell$ et $v_n \to \ell'$, par conséquent $w_n \to \ell' - \ell$, or $w_n \to 0$, donc $\ell = \ell'$. □

3) Le théorème de BOLZANO - WEIERSTRASS

Théorème 9.13 (de ***Bolzano***[2]- ***Weierstrass***[3].)
Si u est une suite réelle bornée, alors on peut en extraire une suite convergente.

Preuve : On applique le principe de dichotomie : il existe $a_0 < b_0$ deux réels tels que $\forall n \in \mathbb{N}, u_n \in [a_0; b_0]$. On pose $\mathrm{I}_0 = [a_0; b_0]$ et $\sigma(0) = 0$. On coupe cet intervalle en deux, soit $\mathrm{I}'_0 = [a_0; \frac{a_0+b_0}{2}]$ et $\mathrm{I}''_0 = [\frac{a_0+b_0}{2}; b_0]$, si $\{n \in \mathbb{N} \ / \ u_n \in \mathrm{I}'_0\}$ est infini alors on pose $\mathrm{I}_1 = \mathrm{I}'_0$, sinon on pose $\mathrm{I}_1 = \mathrm{I}''_0$. On a alors un nouveau segment $\mathrm{I}_1 = [a_1; b_1]$ inclus dans I_0 avec $b_1 - a_1 = \frac{b-a}{2}$ et $\{n \in \mathbb{N} \ / \ u_n \in \mathrm{I}_1\}$ infini. On peut donc choisir $n_1 > 0$ tel que $u_{n_1} \in \mathrm{I}_1$, on pose $\sigma(1) = n_1$. On recommence de la même façon avec I_1 ...

On construit ainsi une suite de segments $\mathrm{I}_n = [a_n; b_n]$, emboîtés ($\mathrm{I}_{n+1} \subset \mathrm{I}_n$), tels que $b_n - a_n = \frac{b-a}{2^n}$, et une application $\sigma : \mathbb{N} \to \mathbb{N}$ strictement croissante telles que pour tout n, et $u_{\sigma(n)} \in \mathrm{I}_n$, c'est à dire $a_n \leqslant u_{\sigma(n)} \leqslant b_n$. Or les suites (a_n) et (b_n) sont adjacentes, elles convergent donc vers une même limite ℓ, et donc par le théorème des gendarmes, on a $u_{\sigma(n)} \to \ell$: on a donc construit une suite extraite convergente. □

V EXTENSION AUX SUITES COMPLEXES

1) Définitions

On adopte la même définition et les mêmes notations que pour les suites réelles, une suite complexe est donc une application $u : \mathbb{N} \to \mathbb{C}$, l'ensemble des suites complexes est $\mathcal{F}(\mathbb{N}, \mathbb{C})$.

- Si u est une suite complexe, on pose pour tout entier n, $a_n = \mathrm{Re}(u_n)$ et $b_n = \mathrm{Im}(u_n)$, alors les suites a et b sont des **suites réelles**, avec $u_n = a_n + ib_n$. La suite a est appelée **partie réelle** de u et notée $a = \mathrm{Re}(u)$, la suite b est appelée **partie imaginaire** de u et notée $\mathrm{Im}(u)$. Par exemple, si $\theta \in \mathbb{R}$, la partie réelle que la suite $(e^{in\theta})$ est la suite $(\cos(n\theta))$, et sa partie imaginaire est la suite $(\sin(n\theta))$.
- La suite **conjuguée** de u est notée $\overline{u}$ et définie par $\overline{u}_n = a_n - ib_n$.
- La suite **module** de u est notée $|u|$ est définie par $|u|_n = |u_n| = \sqrt{a_n^2 + b_n^2}$.
- Soit $\sigma : \mathbb{N} \to \mathbb{N}$ une application strictement croissante, la suite $(u_{\sigma(n)})$ est appelée **suite extraite** de u et on a $u_{\sigma(n)} = a_{\sigma(n)} + ib_{\sigma(n)}$.
- On dit que la suite complexe u est bornée lorsque sa partie réelle a et sa partie imaginaire b sont des suites réelles bornées. Ceci revient à dire que la suite $|u|$ est majorée.
- On définit dans $\mathcal{F}(\mathbb{N}, \mathbb{C})$ les mêmes opérations que pour les suites réelles : addition, multiplication et produit par un complexe. On trouve de même que $(\mathcal{F}(\mathbb{N}, \mathbb{C}), +, \times)$ est un anneau commutatif non intègre.

3. *BOLZANO Bernhard* (1781 – 1848) : mathématicien et philosophe tchèque.
3. *WEIERSTRASS Karl* (1815 – 1897) : mathématicien allemand parfois surnommé *le père de l'analyse moderne*

2) Convergence

Définition 9.6

Soit u une suite complexe, et soit ℓ un complexe. On dira que la suite u converge vers ℓ lorsque la suite $(|u_n - \ell|)_{n\in\mathbb{N}}$ tend vers 0, c'est à dire :

$$\forall \varepsilon > 0, \exists \mathrm{N} \in \mathbb{N}, \forall n \in \mathbb{N}, n \geqslant \mathrm{N} \implies |u_n - \ell| < \varepsilon.$$

☞**Exemple** : Soit $u_n = \frac{e^{in\theta}}{n}$, alors $|u_n| = \frac{1}{n} \to 0$ donc u converge vers 0.

3) Propriétés

Théorème 9.14

Soit u une suite complexe et ℓ un complexe, alors la suite u converge vers ℓ si et seulement si la suite $(\mathrm{Re}(u_n))$ converge vers $\mathrm{Re}(\ell)$ ***et*** *la suite $(\mathrm{Im}(u_n))$ converge vers $\mathrm{Im}(\ell)$.*

Preuve : Notons $u_n = a_n + ib_n$ et $\ell = \alpha + i\beta$, (formes algébriques). Supposons que $a_n \to \alpha$ et $b_n \to \beta$, alors $|u_n - \ell| = \sqrt{(a_n-\alpha)^2 + (b_n-\beta)^2}$ qui tend donc vers 0, donc u converge vers ℓ.

Réciproquement, si u converge vers ℓ, on a $|a_n - \alpha| \leqslant |u_n - \ell|$ et $|b_n - \beta| \leqslant |u_n - \ell|$, or $|u_n - \ell|$ tend vers 0, par conséquent $a_n \to \alpha$ et $b_n \to \beta$. □

Connaissant les propriétés de suites réelles convergentes, on peut en déduire celles des suites complexes convergentes en raisonnant sur les parties réelles et imaginaires :

- Toute suite convergente est bornée.
- Si u converge vers $\ell \in \mathbb{C}$, alors toute suite extraite de u converge vers ℓ.
- Si u converge vers $\ell \in \mathbb{C}$ et v converge vers $\ell' \in \mathbb{C}$, alors $u+v \to \ell+\ell'$, $uv \to \ell\ell'$ et $\forall \lambda \in \mathbb{C}, \lambda u \to \lambda\ell$.
- Si $u \to \ell \in \mathbb{C}^*$, alors à partir d'un certain rang $u_n \neq 0$ et $\frac{1}{u} \to \frac{1}{\ell}$.
- Si u converge vers $\ell \in \mathbb{C}$, alors la suite $\overline{u}$ converge vers $\overline{\ell}$ et la suite $|u|$ converge vers $|\ell|$.
- Si u est bornée alors on peut en extraire une suite convergente (*Bolzano - Weierstrass*).

Remarque 9.8 – *Si $u_n \to \ell$ dans $\mathbb{C}$, et si u est à valeurs réelles, alors la suite (b_n) est la suite nulle, or $b_n \to \mathrm{Im}(\ell)$, donc $\mathrm{Im}(\ell) = 0$, c'est à dire $\ell \in \mathbb{R}$.*

★**Exercice 9.2** *Étude de la suite $(u_n = e^{in\theta})$.*

VI ÉTUDE DE SUITES PARTICULIÈRES

Nous avons déjà évoquées le cas des suites arithmétiques et des suites géométriques au début du chapitre, avec les formules à connaître. Voici deux autres types de suites pour lesquelles il faudra connaître la méthode pour exprimer le terme général en fonction de n.

1) Suite arithmético-géométrique

Ce sont les suites u définies par : $u_0 \in \mathbb{C}$, et $\forall n \in \mathbb{N}$, $u_{n+1} = au_n + b$ où $a \in \mathbb{C} \setminus \{1\}$ et $b \in \mathbb{C}^*$ (si $a = 0$ ou si $b = 0$ on retrouve les suites arithmétiques ou géométriques).

La méthode pour exprimer u_n en fonction de n est la suivante :

a) On cherche une suite constante $(\alpha)_{n\in\mathbb{N}}$ vérifiant la relation de récurrence, ce qui donne $\alpha = a\alpha + b$ c'est à dire $\alpha = \frac{b}{1-a}$.

b) On introduit la suite auxiliaire v définie par $v_n = u_n - \alpha$, et on vérifie que celle-ci est géométrique de raison a :

$$v_{n+1} u_{n+1} - \alpha = au_n + b - (a\alpha + b) = a(u_n - \alpha) = av_n$$

c) On exprime alors v_n en fonction de n, puis u_n :

$$\forall n \in \mathbb{N},\ u_n = v_n + \alpha = v_0 a^n + \alpha = (u_0 - \alpha)a^n + \alpha$$

2) Relation de récurrence linéaire à deux pas

Les suites u vérifiant une relation de récurrence linéaire à deux pas, sont les suites définies par : $(u_0, u_1) \in \mathbb{C}^2$ et $\forall n \in \mathbb{N}$, $u_{n+2} = au_{n+1} + bu_n$ (R), avec $(a, b) \in \mathbb{C}^2$ et $b \neq 0$ (si $b = 0$ la suite est géométrique à partir du rang 1).

Théorème 9.15

Si u et v sont deux suites qui vérifient la relation (R)*, alors pour tout* $(\lambda, \mu) \in \mathbb{C}^2$*, la suite* $w = (\lambda u_n + \mu v_n)_{n\in\mathbb{N}}$ *vérifie également* (R)*.*

Preuve : Il suffit pour $n \in \mathbb{N}$, d'évaluer $aw_{n+1} + bw_n$, on trouve w_{n+2}. □

Théorème 9.16

Soient u et v sont deux suites qui vérifient la relation (R)*, si $u_0 = v_0$ et $u_1 = v_1$, alors $\forall n \in \mathbb{N}$, $u_n = v_n$, i.e. les deux suites u et v sont égales.*

Preuve : C'est une récurrence double sur n. □

Théorème 9.17

Une suite géométrique $(q^n)_{n\in\mathbb{N}}$ avec $q \in \mathbb{C}$, vérifie la relation (R) *si et seulement si q est racine de l'équation $X^2 - aX - b = 0$. Cette équation est appelée équation caractéristique de la relation.*

Preuve : $(\forall n \in \mathbb{N}, q^{n+2} = aq^{n+1} + bq^n) \iff q^2 = aq + b$. □

De là découle la méthode :

– Si l'équation caractéristique $X^2 - aX - b = 0$ admet deux racines distinctes q_1 et q_2, alors les suites géométriques $(q_1^n)_{n\in\mathbb{N}}$ et $(q_2^n)_{n\in\mathbb{N}}$ vérifient la relation (R), donc pour tout $(\lambda, \mu) \in \mathbb{C}^2$, la suite $(\lambda q_1^n + \mu q_2^n)_{n\in\mathbb{N}}$ vérifie également (R).

  Soit u une suite vérifiant (R), on cherche λ et μ tels que $\begin{cases} \lambda q_1^0 + \mu q_2^0 = u_0 \\ \lambda q_1^1 + \mu q_2^1 = u_1 \end{cases}$, comme $q_1 \neq q_2$ on vérifie que ce système possède une unique solution. On en déduit que la suite $v = (\lambda q_1^n + \mu q_2^n)_{n\in\mathbb{N}}$ (avec λ et μ solutions du système précédent), est égale à la suite u, et donc :

$$\boxed{\forall n \in \mathbb{N}, u_n = \lambda q_1^n + \mu q_2^n}$$

– Si l'équation caractéristique $X^2 - aX - b = 0$ admet une racine double q ($q = \frac{-a}{2}$), alors la suite géométrique $(q^n)_{n\in\mathbb{N}}$ vérifie la relation (R). On peut facilement vérifier que la suite $(nq^n)_{n\in\mathbb{N}}$ vérifie également la relation (R), donc pour tout $(\lambda, \mu) \in \mathbb{C}^2$, la suite $((\lambda n\mu)q^n)_{n\in\mathbb{N}}$ vérifie également (R).

  Soit u une suite vérifiant (R), on cherche λ et μ tels que $\begin{cases} (\lambda + 0\mu)q^0 = u_0 \\ (\lambda + \mu)q^1 = u_1 \end{cases}$, comme $q \neq 0$ on vérifie que ce système possède une unique solution. On en déduit que la suite $v = ((\lambda + n\mu)q^n)_{n\in\mathbb{N}}$ (avec λ et μ solutions du système précédent), est égale à la suite u, et donc :

$$\boxed{\forall n \in \mathbb{N}, u_n = (\lambda + n\mu)q^n}$$

Cas réel.

Lorsque a, b sont réels et que l'on cherche le suites réelles vérifiant la relation (R), il y a trois cas suivant la discriminant $\Delta = a^2 + 4b$ de l'équation caractéristique :

– Si $\Delta > 0$, l'équation caractéristique $X^2 - aX - b = 0$ admet deux racines distinctes réelles q_1 et q_2, comme précédemment on montre qu'une suite u vérifie (R) si et seulement si il existe deux réels λ et μ tels que :

$$\boxed{\forall n \in \mathbb{N}, u_n = \lambda q_1^n + \mu q_2^n}$$

– Si $\Delta = 0$, l'équation caractéristique $X^2 - aX - b = 0$ admet une racine double réelle q, comme précédemment on montre qu'une suite u vérifie (R) si et seulement si il existe deux réels λ et μ tels que :

$$\boxed{\forall n \in \mathbb{N}, u_n = \lambda q_1^n + \mu q_2^n}$$

– Si $\Delta < 0$, l'équation caractéristique $X^2 - aX - b = 0$ admet deux racines complexes non réelles conjuguées q et $\overline{q}$, une suite complexe u vérifie (R) si et seulement si il existe deux réels λ et μ tels que $\forall n \in \mathbb{N}$, $u_n = \lambda q^n + \mu \overline{q}^n$, elle est réelle si et seulement si $\forall n \in \mathbb{N}$, $u_n = \overline{u_n}$ *i.e.* $\lambda q^n + \mu \overline{q}^n = \overline{\lambda}\overline{q}^n + \overline{\mu} q^n$, ce qui équivaut à $\mu = \overline{\lambda}$, et donc $u_n = 2\operatorname{Re}(\lambda q^n)$, on écrit $q = re^{i\theta}$ (forme trigonométrique), on obtient alors qu'il existe deux réels α, β tels que :

$$\boxed{\forall n \in \mathbb{N},\ u_n = r^n(\alpha \cos(n\theta) + \beta \sin(n\theta))}$$

VII SOLUTION DES EXERCICES

Solution 9.1 *C'est une suite géométrique de raison $e^{i\theta}$. Si $\theta = 0\ (2\pi)$, alors la suite est constante égale à 1, donc $u_n \to 1$. Si $\theta \neq 0\ (2\pi)$, supposons que $u_n \to \ell \in \mathbb{C}$, alors $|u_n| \to |\ell|$, or $|u_n| = 1$, donc $|\ell| = 1$. D'autre part, $u_{n+1} = e^{i\theta}u_n$, par passage à la limite, on a $\ell = \ell e^{i\theta}$, or $\ell \neq 0$ (car $|\ell| = 1$), donc $e^{i\theta} = 1$ ce qui est absurde, par conséquent si $\theta \neq 0\ (2\pi)$, la suite (u_n) est divergente.*

Chapitre 10

Arithmétique

Sommaire

I DIVISIBILITÉ

1) La propriété fondamentale

Théorème 10.1
Toute partie de $\mathbb{Z}$ non vide et minorée admet un plus petit élément.

Preuve : Soit A une partie de $\mathbb{Z}$ non vide et minorée par un entier n_0. Soit M l'ensemble des minorants de A, on a $n_0 \in \mathrm{M}$, supposons que $n \in \mathrm{M} \implies n+1 \in \mathrm{M}$, alors d'après le principe de récurrence, $\forall\, n \in \mathbb{Z}, n \geqslant n_0 \implies n \in \mathrm{M}$. Soit $p \in \mathrm{A}$, $p \geqslant n_0$, donc $p+1 \in \mathrm{M}$ ce qui entraîne que $p+1 \leqslant p$: absurde, donc il existe un entier n_1 tel que $n_1 \in \mathrm{M}$ et $n_1 + 1 \notin \mathrm{M}$, mais alors il existe un élément p_1 de A tel que $p_1 < n_1 + 1$, d'où $n_1 \leqslant p_1 < n_1 + 1$, ce qui entraîne $p_1 = n_1$, et donc $n_1 \in \mathrm{A}$, nécessairement n_1 est le plus petit élément de A. □

À retenir

- Toute partie non vide et majorée de $\mathbb{Z}$ admet un plus grand élément. En effet, si A est non vide majorée, alors $-\mathrm{A} = \{-a \mid a \in \mathrm{A}\}$ est non vide minorée, donc $-\mathrm{A}$ admet un plus petit élément $-n_0$, ce qui signifie que n_0 est le plus grand élément de A.
- Toute partie non vide de $\mathbb{N}$ admet un plus petit élément (propriété fondamentale de $\mathbb{N}$).

2) La division euclidienne

Théorème 10.2

Soient $a \in \mathbb{Z}$ et $b \in \mathbb{Z}^*$, il existe un unique couple d'entiers (q,r) tel que $a = bq+r$ avec $0 \leqslant r < |b|$, q est appelé le quotient, et r le reste.

Preuve : Supposons $b > 0$: soit $B = \{b(n+1) \ / \ n \in \mathbb{Z}\}$, alors B est non majoré et non minoré, donc il existe un entier n_1 tel que $a < b(n_1+1)$ et il existe un entier n_2 tel que $b(n_2+1) < a$. Soit $A = \{n \in \mathbb{Z} \ / \ a < b(n+1)\}$, alors A est non vide ($n_1 \in A$) et minoré par n_2, donc A admet un plus petit élément q, d'où $bq \leqslant a < b(q+1)$, en posant $r = a - bq$, on a $a = bq + r$ et $0 \leqslant r < b = |b|$.

Supposons $b < 0$: on applique ce qui précède à $-b > 0$, il existe un entier q et un entier r tels que $a = (-b)q + r = b(-q) + r$ avec $0 \leqslant r < -b = |b|$.

Montrons l'unicité : si $a = bq + r = bq' + r'$ avec $0 \leqslant r < |b|$ et $0 \leqslant r' < |b|$, alors $|r - r'| = |bq' - bq| = |b||q' - q| < |b|$, d'où $q' = q$ (ce sont des entiers) et donc $r' = r$. □

Définition 10.1

Soient $a, b \in \mathbb{Z}$, on dit que b divise a lorsqu'il existe $k \in \mathbb{Z}$ tel que $a = bk$. Notation : $b|a$.

Remarque 10.1 – *On a ainsi défini une relation dans $\mathbb{Z}$, elle est réflexive, non symétrique, non antisymétrique, et transitive.*

★**Exercice 10.1** *Soient $a, b \in \mathbb{Z}$ et $n \in \mathbb{N}$, montrer que $a - b \mid a^n - b^n$.*

Théorème 10.3

Soient $a, b \in \mathbb{Z}$ avec $b \neq 0$, alors $b|a$ ssi le reste dans la division euclidienne de a par b est nul.

Preuve : Celle-ci est simple et laissée en exercice. □

Notation : Soit $n \in \mathbb{Z}$, on note $n\mathbb{Z}$ l'ensemble des multiples de n, $n\mathbb{Z} = \{kn \ / \ k \in \mathbb{Z}\}$.

Théorème 10.4

- *$b \mid a \iff a \in b\mathbb{Z}$.*
- *Si $a \neq 0$, alors $b \mid a \implies |b| \leqslant |a|$.*
- *($a \mid b$ et $b \mid a$) $\iff a\mathbb{Z} = b\mathbb{Z} \iff a = \lambda b$ avec $\lambda = \pm 1$ [on dit que a et b sont associés].*
- *Si $b \mid a$ et $b \mid c$ alors $\forall\ u, v \in \mathbb{Z}, b \mid au + cv$.*
- *Si $nb \mid na$ et si $n \neq 0$, alors $b \mid a$.*

3) Congruences

Définition 10.2 (congruences)

Soient $a, b, n \in \mathbb{Z}$, on dit que a est congru à b modulo n lorsque $n \mid a - b$. Notation : $a \equiv b \pmod n$.

Théorème 10.5

- *La relation de congruence modulo n est une relation d'équivalence.*
- *Soient $a, b, c, d, n \in \mathbb{Z}$, si $a \equiv b \pmod n$ et $c \equiv d \pmod n$ alors :*

$$ac \equiv bd \pmod n \text{ et } a + c \equiv b + d \pmod n.$$

On dit que la relation de congruence est compatible avec les opérations.

Preuve : Celle-ci est simple et laissée en exercice. □

Remarque 10.2 – *Soit $a, n \in \mathbb{Z}$ avec $n \neq 0$, $a \equiv r \pmod n$ et $0 \leqslant r < |n|$ si et seulement si r est le reste de la division de a par n.*

☞**Exemple** : Dans $\mathbb{Z}$, si $n = a_0 + 10a_1 + \cdots + 10^p a_p$ (écriture décimale) alors $n \equiv a_0 + \cdots + a_p \pmod 3$ car $10^k \equiv 1 \pmod 3$

4) Diviseurs communs

Définition 10.3 (diviseurs communs)

Pour $a \in \mathbb{Z}$, on note D_a l'ensemble des diviseurs de a. Si $a, b \in \mathbb{Z}$, on note $D_{a,b}$ l'ensemble des diviseurs communs à a et b, on a donc $D_{a,b} = D_a \cap D_b$, cet ensemble contient toujours ± 1.

Remarque 10.3 –

- *Pour tout élément $a \in \mathbb{Z}$, $\pm 1 \mid a$.*
- *Si $a \neq 0$, alors D_a est un ensemble fini, plus précisément $D_a \subset [\![-|a| ; |a|]\!]$.*
- $D_0 = \mathbb{Z}$, $D_{\pm 1} = \{\pm 1\}$.
- *Si a et b sont dans $\mathbb{Z}$: $D_a = D_{|a|}$ (on en déduit que $D_{a,b} = D_{|a|,|b|}$).*

Théorème 10.6

Soient $a, b, q, r \in \mathbb{Z}$, si $a = bq + r$, alors $D_{a,b} = D_{b,r}$.

Preuve : Si $d \in D_{a,b}$, alors $d \mid a$ et $d \mid b$ donc $d \mid a - bq$ *i.e.* $d \mid r$, donc $d \in D_{b,r}$.
Réciproquement, si $d \in D_{b,r}$, alors $d \mid b$ et $d \mid r$ donc $d \mid bq + r$ *i.e.* $d \mid a$, d'où $d \in D_{a,b}$. □

Application – Le théorème ci-dessus fournit un algorithme pour la recherche des diviseurs communs à a et b (entiers naturels) basé sur la division euclidienne : c'est **l'algorithme d'***Euclide*[1], voici son principe :

On remarque que si $b = 0$ alors $D_{a,b} = D_a$. On peut supposer désormais que $b \neq 0$ et on cherche à calculer $D = D_{a,b}$:

Étape 1 : on effectue la division euclidienne de a par b, $a = bq_1 + r_1$ avec $0 \leqslant r_1 < b$. On a $D = D_{b,r_1}$, donc si $r_1 = 0$ alors $D = D_b$, sinon on passe :

Étape 2 : on effectue la division euclidienne de b par r_1, $b = r_1 q_2 + r_2$ avec $0 \leqslant r_2 < r_1$. On a $D = D_{r_1,r_2}$, donc si $r_2 = 0$ alors $D = D_{r_1}$, sinon on passe :

Étape 3 : on effectue la division euclidienne de r_1 par r_2 ; $r_1 = r_2 q_3 + r_3$ avec $0 \leqslant r_3 < r_2$. On a $D = D_{r_2,r_3}$, donc si $r_3 = 0$ alors $D = D_{r_2}$, sinon on passe à l'étape 4...

La suite des restes obtenus est une suite strictement décroissante d'entiers positifs, elle est donc nécessairement finie, *i.e.* il existe un entier $n \geqslant 1$ tel que $r_n = 0$, l'ensemble cherché est donc $D = D_{r_{n-1}}$ (avec la convention $r_0 = b$).

À retenir

$D_{a,b} = D_r$ où r est le dernier reste non nul dans l'algorithme d'Euclide.

☞**Exemple** : Cherchons les diviseurs communs à $a = 336$ et $b = 210$

– on effectue la division de a par b : $336 = 1 \times 210 + 126$, donc $D_{a,b} = D_{210,126}$.
– on effectue la division de 210 par 126 : $210 = 1 \times 126 + 84$, donc $D_{a,b} = D_{210,126} = D_{126,84}$.
– on effectue la division de 126 par 84 : $126 = 1 \times 84 + 42$, donc $D_{a,b} = D_{84,42}$.
– on effectue la division de 84 par 42 : $84 = 2 \times 42 + 0$, donc $D_{a,b} = D_{42,0} = D_{42}$, c'est à dire :

$$D_{336,210} = \{\pm 1, \pm 2, \pm 3, \pm 6, \pm 7, \pm 14, \pm 21, \pm 42\}.$$

II ÉLÉMENTS PREMIERS ENTRE EUX

1) Théorème de Bézout

1. *EUCLIDE* (300 av. J.C. – 275 av. J.C. environ) : on ne sait pratiquement rien de sa vie, il était vraisemblablement grec. Son œuvre est colossale et son ouvrage fondamental « Les éléments » regroupe toutes les connaissances de l'époque, il faudra près de vingt siècles pour dépasser son œuvre.

Définition 10.4

Soient $a, b \in \mathbb{Z}$, on dit que a et b sont premiers entre eux (ou a est premier avec b) lorsque le seul diviseur commun positif est 1, i.e. $D_{a,b} = \{\pm 1\}$.

Remarque 10.4 –

- *Dire que a est premier avec b revient à dire que le dernier reste non nul dans l'algorithme d'Euclide est 1.*
- *Si a est premier avec b, alors au moins un des deux est non nul (sinon l'ensemble des diviseurs communs est $\mathbb{Z}$).*
- *a est premier avec a si et seulement si $a \pm 1$.*

Théorème 10.7 (**théorème de *Bézout*** [2])

Soient $a, b \in \mathbb{Z}$, alors a et b sont premiers entre eux si et seulement si il existe $u, v \in \mathbb{Z}$ tels que $au + bv = 1$. Les entiers u et v sont appelés coefficients de Bézout (non uniques en général).

Preuve : Supposons que u et v existent et soit d un diviseur commun à a et b, alors $d \mid a$ et $d \mid b$, donc $d \mid au + bv$ *i.e.* $d \mid 1$, donc $d = \pm 1$ ce qui prouve que a et b sont premiers entre eux.

Réciproquement : si a est premier avec b. En appliquant l'algorithme d'*Euclide* on vérifie qu'à chaque étape le reste r_k peut se mettre sous la forme $r_k = au_k + bv_k$ avec u_k et v_k dans $\mathbb{Z}$ (récurrence) (algorithme d'Euclide étendu), comme le dernier reste non nul est 1, il existe bien u et v dans $\mathbb{Z}$ tels que $1 = au + bv$ (de plus on sait les calculer !). □

☞**Exemple** : $\forall\, n \in \mathbb{Z}$, n et $n+1$ sont premiers entre eux, puisque $n + 1 - n = 1$.

2) Conséquences

Théorème 10.8

Si a est premier avec b et si a est premier avec c, alors a est premier avec le produit bc. On en déduit que si a est premier avec $c_1, \ldots, c_n$, alors a est premier avec le produit $c_1 \times \ldots \times c_n$.

Preuve : Il existe $u, v \in \mathbb{Z}$ tels que $au + bv = 1$, il existe $p, q \in \mathbb{Z}$ tels que $ap + cq = 1$. On effectue le produit de ces deux relations, ce qui donne $a(ucq + uap + pbv) + bc(vq) = 1$, d'après le théorème de *Bézout*, a et bc sont premiers entre eux. Une simple récurrence sur n permet de démontrer la généralisation. □

Théorème 10.9

Si a est premier avec c, si $a \mid b$ et si $c \mid b$, alors $ac \mid b$.

Preuve : Il existe $u, v \in \mathbb{Z}$ tels que $au + cv = 1$, on multiplie par b, ce qui donne : $bau + bcv = b$, or $c \mid b$ donc $ac \mid bau$, et $a \mid b$ donc $ac \mid bcv$, ce qui entraîne $ac \mid bau + bcv$ *i.e.* $ac \mid b$.

Remarquons que ce théorème est faux lorsque a et c ne sont pas premiers entre eux, par exemple : $2 \mid 12$ et $4 \mid 12$ mais $2 \times 4 = 8 \nmid 12$. □

Théorème 10.10 (**théorème de *Gauss***)

Si $a \mid bc$ et si a est premier avec c, alors $a \mid b$.

Preuve : Il existe $u, v \in \mathbb{Z}$ tels que $au + cv = 1$, on multiplie par b, ce qui donne $bau + bvc = b$, or $a \mid bc$ donc $a \mid bau + bcv$, *i.e.* $a \mid b$. □

★**Exercice 10.2** *Résoudre dans $\mathbb{Z}$ l'équation $17x + 12y = 3$.*

III LE PLUS GRAND DIVISEUR COMMUN

1) Définition

Soient $a, b \in \mathbb{Z}$ non tous deux nuls (*i.e.* $a \neq 0$ ou $b \neq 0$), on sait que $D_{a,b} = D_r$ où r est le dernier reste non nul dans l'algorithme d'*Euclide*, on voit que les diviseurs communs à a et b ont une valeur absolue inférieure ou égale à celle de r et donc r est le plus grand diviseur commun.

2. *BÉZOUT Étienne*(1730 – 1783) : mathématicien français, l'un des précurseurs de la géométrie algébrique.

Définition 10.5

Soient $a, b \in \mathbb{Z}$ non tous deux nuls, on appelle pgcd de a et de b le plus grand diviseur commun. Notation : $\operatorname{pgcd}(a, b)$ *ou* $a \wedge b$*, c'est le dernier reste non nul dans l'algorithme d'Euclide.*

Remarque 10.5 – *Il en découle que deux éléments a et b de $\mathbb{Z}$, non tous deux nuls, sont premiers entre eux si et seulement si* $\operatorname{pgcd}(a, b) = 1$.

Théorème 10.11 (Calcul pratique d'un pgcd)

Si $a, b \in \mathbb{Z}$ sont non tous deux nuls alors $\forall q \in \mathbb{Z}, \operatorname{pgcd}(a, b) = \operatorname{pgcd}(a - bq, b)$.

Preuve : Soit $r = a - bq$, on a $a = bq + r$ et on sait alors que $D_{a,b} = D_{b,r}$, le résultat en découle. □

Listing 10.1 – *euclide pour a et b positifs*

```python
def pgcd(a,b):
    A,B = a,b
    while B != 0:
        A,B = B, A%B
    return(A) #dernier reste non nul
```

★**Exercice 10.3**

1/ Prouver la terminaison de l'algorithme.

2/ Montrer que l'on a l'invariant de boucle de boucle $P(k) =$ « $\operatorname{pgcd}(a, b) = \operatorname{pgcd}(A_k, B_k)$ ».

3/ Conclure.

☞**Exemple** : Soit à calculer $d = \operatorname{pgcd}(3282, 1281)$:
- $3282 = 2 \times 1281 + 720$, donc $d = \operatorname{pgcd}(1281, 720)$,
- $1281 = 1 \times 720 + 561$, donc $d = \operatorname{pgcd}(720, 561)$,
- $720 = 1 \times 561 + 159$, donc $d = \operatorname{pgcd}(561, 159)$,
- $561 = 3 \times 159 + 84$, donc $d = \operatorname{pgcd}(159, 84)$,
- $159 = 1 \times 84 + 75$, donc $d = \operatorname{pgcd}(84, 75)$,
- $84 = 1 \times 75 + 9$, donc $d = \operatorname{pgcd}(75, 9)$,
- $75 = 8 \times 9 + 3$, donc $d = \operatorname{pgcd}(9, 3)$,
- $9 = 3 \times 3 + 0$, donc $d = 3$.

2) Propriétés

Théorème 10.12 (caractérisations du pgcd)

Soient $a, b \in \mathbb{Z}$ non tous deux nuls, et soit $d \in \mathbb{N}^*$. On a alors :
- $d = \operatorname{pgcd}(a, b) \iff d \mid a, d \mid b$ *et* $\exists\, u, v \in \mathbb{Z}$ *tels que* $au + bv = d$.
- $d = \operatorname{pgcd}(a, b) \iff \exists\, u, v \in \mathbb{Z}$ *premiers entre eux, tels que* $a = du$ *et* $b = dv$.

Preuve : Pour le premier point : Si $d = \operatorname{pgcd}(a, b)$, cela découle d l'algorithme d'Euclide étendu.

Si $d \mid a, d \mid b$ et $\exists\, u, v \in \mathbb{Z}$ tels que $au + bv = d$, alors d est diviseur commun à a et b, et la deuxième relation entraîne que tout diviseur commun à a et b est un diviseur de d, donc $d = \operatorname{pgcd}(a, b)$ (d étant positif).

Pour le second point : Si $d = \operatorname{pgcd}(a, b)$ alors il existe $u, v \in \mathbb{Z}$ tels que $a = du$ et $b = dv$, soit $k = u \wedge v$, alors kd divise a et b, donc $|kd| \leqslant |d|$ ce qui entraîne $|k| \leqslant 1$ et donc $k = 1$.

Si $a = du, b = dv$ avec $u \wedge v = 1$, alors d est un diviseur commun à a et b, d'après le théorème de *Bézout*, il existe $\alpha, \beta \in \mathbb{Z}$ tels que $\alpha u + \beta v = 1$, d'où $d = \alpha a + \beta b$, et donc d'après le premier point, *i.e.* $d = a \wedge b$. □

Théorème 10.13 (quelques propriétés du pgcd)

Soient $a, b \in \mathbb{Z}$ non tous deux nuls :

a) $\forall\, n \in \mathbb{Z}$*, si* $n \mid a$ *et* $n \mid b$*, alors* $n \mid \operatorname{pgcd}(a, b)$.

b) $\forall\, k \in \mathbb{N}^*, \operatorname{pgcd}(ka, kb) = k \operatorname{pgcd}(a, b)$.

c) $\forall\, n \in \mathbb{N}, \operatorname{pgcd}(a^n, b^n) = \operatorname{pgcd}(a, b)^n$.

d) Si a et c sont premiers entre eux, alors $\text{pgcd}(a, bc) = \text{pgcd}(a, b)$.

Preuve : Pour le premier point : Soit $d = \text{pgcd}(a, b)$, alors $D_{a,b} = D_d$ donc tout diviseur commun à a et b est un diviseur de d.

Pour le deuxième point : soit $d = \text{pgcd}(a, b)$, alors il existe $u, v \in \mathbb{Z}$ premiers entre eux tels que $a = du$ et $b = dv$, d'où $ka = kdu$ et $kb = kdv$, donc $kd = \text{pgcd}(ka, kb)$.

Pour le troisième point : en reprenant les notations ci-dessus, $a^n = d^n u^n$ et $b^n = d^n v^n$, or u et v sont premiers entre eux, donc u^n et v^n aussi (conséquence du théorème de *Bézout*), par conséquent $d^n = \text{pgcd}(a^n, b^n)$.

Pour le dernier point : on reprend les notations ci-dessus, $a = du$ et $bc = dcv$ mais $u \mid a$ et a est premier avec c, donc u est premier avec c, d'où u est premier avec cv, et donc $d = \text{pgcd}(a, bc)$. □

3) Généralisation

Soient a, b, c trois entiers non tous nuls, l'ensemble des diviseurs communs à a, b et c est :

$$D_{a,b,c} = D_a \cap D_b \cap D_c = (D_a \cap D_b) \cap D_c = D_a \cap (D_b \cap D_c)$$

or on sait que $D_a \cap D_b = D_{a\wedge b}$, donc $D_{a,b,c} = D_{(a\wedge b)\wedge c} = D_{a\wedge(b\wedge c)}$. Ces deux entiers étant strictement positifs, on a $(a \wedge b) \wedge c = a \wedge (b \wedge c)$ et ce nombre est le plus grand diviseur positif commun à a, b et c. Par définition ce nombre est $\boxed{\text{le pgcd de } a,\ b \text{ et } c, \text{ on le note : } \text{pgcd}(a, b, c)}$.

Théorème 10.14 **(associativité du pgcd)**
Soient a, b, c trois entiers avec b non nul, alors $\text{pgcd}(a, b, c) = (a \wedge b) \wedge c = a \wedge (b \wedge c)$.

À retenir
L'associativité du pgcd permet de ramener le calcul au cas de deux entiers.

Notons $d' = a \wedge b$ et $d = \text{pgcd}(a, b, c)$, alors $d = d' \wedge c$, donc il existe deux entiers u' et w tels que $d = d'u' + cw$, de même, il existe deux entiers α et β tels que $d' = \alpha a + \beta b$, d'où en remplaçant, $d = a\alpha u' + b\beta u' + cw = au + bv + cw$ avec $u, v, w \in \mathbb{Z}$.

Réciproquement, si d est un diviseur commun positif, et si $d = au + bv + cw$, alors il est facile de voir que tout diviseur commun à a, b et c est un diviseur de d et donc $d = \text{pgcd}(a, b, c)$, d'où le théorème :

Théorème 10.15
Soient a, b, c trois entiers non tous nuls et $d \in \mathbb{N}^*$, alors :
$$d = \text{pgcd}(a, b, c) \iff d \in D_{a,b,c} \text{ et } \exists u, v, w \in \mathbb{Z}, d = au + bv + cw.$$

Définition 10.6
Soient a, b, c trois entiers non tous nuls, on dira que ces trois nombres sont :
- ***premiers entre eux dans leur ensemble*** *lorsque* $\text{pgcd}(a, b, c) = 1$.
- ***premiers entre eux deux à deux*** *lorsque* $\text{pgcd}(a, b) = \text{pgcd}(b, c) = \text{pgcd}(a, c) = 1$.

Attention !
Les deux notions ne sont pas équivalentes, ***la deuxième entraîne la première mais la réciproque est fausse*** *comme le montre l'exemple suivant :*
$$\text{pgcd}(6, 15, 20) = 1 \textit{ mais } \text{pgcd}(6, 15) = 3,\ \text{pgcd}(6, 20) = 2 \textit{ et } \text{pgcd}(15, 20) = 5.$$

Il découle du théorème précédent :

Théorème 10.16 **(de Bézout)**
Soient a, b, c trois entiers non tous nuls, alors a, b et c sont premiers entre eux dans leur ensemble si et seulement si :
$$\exists u, v, w \in \mathbb{Z}, au + bv + cw = 1.$$

Théorème 10.17 (caractérisation)
Soient a, b, c trois entiers non tous nuls et $d \in \mathbb{N}^*$, alors :
$$d = \text{pgcd}(a, b, c) \iff \exists u, v, w \in \mathbb{Z}, a = du, b = dv \text{ et } c = dw \text{ avec } \text{pgcd}(u, v, w) = 1.$$

Preuve : Si $d = \text{pgcd}(a, b, c)$ alors il existe $\exists u, v, w \in \mathbb{Z}$, $a = du$, $b = dv$ et $c = dw$. Il existe également des entiers α, β et γ tels que $d = \alpha a + \beta b + \gamma c$ d'où $1 = \alpha u + \beta v + \gamma c$ et donc $\text{pgcd}(u, v, w) = 1$.

Réciproquement, si $a = du$, $b = dv$ et $c = dw$ avec $\text{pgcd}(u, v, w) = 1$. Il existe des entiers α, β et γ tels que $1 = \alpha u + \beta v + \gamma w$, en multipliant par d il vient alors que $d = \alpha a + \beta b + \gamma c$, ce qui entraîne que $d = \text{pgcd}(a, b, c)$ (car $d \in \mathbb{N}^*$ et $d \in \text{D}_{a,b,c}$). □

Remarque 10.6 – *Nous avons étendu la notion de pgcd à trois entiers, mais on pourrait l'étendre de la même manière à n entiers.*

IV LE PLUS PETIT MULTIPLE COMMUN

1) Définition

Théorème 10.18
Si a et b sont non nuls, il existe un unique élément m positif dans $\mathbb{Z}$ tel que $(a\mathbb{Z}) \cap (b\mathbb{Z}) = m\mathbb{Z}$.

Preuve : $a\mathbb{Z} \cap b\mathbb{Z}$ contient $|ab| > 0$, on note m le plus petit élément strictement positif dans $a\mathbb{Z} \cap b\mathbb{Z}$, alors il est facile de voir que $m\mathbb{Z} \subset a\mathbb{Z} \cap b\mathbb{Z}$. Si $p \in a\mathbb{Z} \cap b\mathbb{Z}$, on effectue la division de p par m, $p = mq + r$ avec $0 \leqslant r < m$, d'où $r = p - mq$, on vérifie alors que r est aussi dans $a\mathbb{Z} \cap b\mathbb{Z}$ (car a et b divisent p et m donc $p - mq$). Si $r > 0$ alors $r \geqslant m$ car m est le plus petit élément strictement positif dans $a\mathbb{Z} \cap b\mathbb{Z}$, ceci est absurde, donc $r = 0$ et $p = mq$, d'où $a\mathbb{Z} \cap b\mathbb{Z} \subset m\mathbb{Z}$, et finalement on a bein l'inégalité.

Si $m'\mathbb{Z} = m\mathbb{Z}$ avec $m' > 0$, alors m et m' se divisent muutuellement, d'où $m = m'$. □

Il découle de ce théorème que c est un multiple commun à a et b si et seulement si $c \in (a\mathbb{Z}) \cap (b\mathbb{Z})$, ce qui équivaut à $c \in m\mathbb{Z}$, c'est à dire $m \mid c$. Ceci entraîne en particulier : $m \leqslant |c|$.

Définition 10.7
Soit $a, b \in \mathbb{Z}$, non nuls, et soit $m \in \mathbb{N}^*$, on dit que m est le ppcm de a et b lorsque $(a\mathbb{Z}) \cap (b\mathbb{Z}) = m\mathbb{Z}$.
Notation : $m = \text{ppcm}(a, b)$ ou encore $m = a \vee b$.

Théorème 10.19 (caractérisation du ppcm)
Soient $a, b \in \mathbb{Z}$, non nuls, et soit $m \in \mathbb{N}^*$ alors :
$$m = \text{ppcm}(a, b) \iff \exists\, u, v \in \mathbb{Z} \text{ premiers entre eux; tels que } m = au = bv.$$

Preuve : On suppose $a, b \in \mathbb{Z}$, non nuls. Si $m = \text{ppcm}(a, b)$: alors $a \mid m$ et $b \mid m$. Donc il existe $u, v \in \mathbb{Z}$ tels que $m = au = bv$, soit $d = \text{pgcd}(u, v)$ alors il existe $\alpha, \beta \in \mathbb{Z}$ premiers entre eux tels que $u = d\alpha$ et $v = d\beta$, d'où $m = ad\alpha = bd\beta$, mais alors $m' = a\alpha = b\beta$ est un multiple commun à a et b donc $|m| \leqslant |m'|$ ce qui entraîne $d = 1$.

Si $\exists u, v \in \mathbb{Z}$ premiers entre eux tels que $m = au = bv$, alors $a \mid m$ et $b \mid m$, il existe α, β tels que $u\alpha + v\beta = 1$, soit m' un multiple commun non nul, alors $m' = m'u\alpha + m'v\beta$, on en déduit que $m \mid m'$ et donc $|m| \leqslant |m'|$, ce qui prouve que $m = \text{ppcm}(a, b)$. □

2) Propriétés

Théorème 10.20
Soient $a, b \in \mathbb{Z}$, non nuls :

a) $\forall\, n \in \mathbb{Z}$, *si $a \mid n$ et $b \mid n$ alors* $\text{ppcm}(a, b) \mid n$.

b) Si a et b sont premiers entre eux, alors $\text{ppcm}(a, b) = |ab|$.

c) $\forall\, k \in \mathbb{N}$, *non nul,* $\text{ppcm}(ka, kb) = k\,\text{ppcm}(a, b)$.

d) $\text{ppcm}(a, b) \times \text{pgcd}(a, b) = |ab|$.

e) $\forall\, n \in \mathbb{N}, \text{ppcm}(a^n, b^n) = \text{ppcm}(a, b)^n$.

Preuve : Pour le deuxième point : a et b sont premiers entre eux, alors $ab = ba$ par conséquent $\text{ppcm}(a, b) = ab$ d'après le théorème précédent.

Pour le troisième point : soit $m = \text{ppcm}(a, b)$, alors $m = au = bv$ avec u et v premiers entre eux, d'où $km = kau = kbv$ et donc $km = \text{ppcm}(ka, kb)$.

Pour le quatrième point : soit $m = \text{ppcm}(a, b)$ et $d = \text{pgcd}(a, b)$, il existe u et v premiers entre eux tels que $a = dv$ et $b = du$, or $au = bv$ donc $m = au = bv$ par conséquent $md = adu = ab$.

Pour le cinquième point : soit $m = \text{ppcm}(a, b)$ on a $m = au = bv$ avec u et v premiers entre eux, donc $m^n = a^n u^n = b^n v^n$ avec u^n et v^n premiers entre eux, donc $m^n = \text{ppcm}(a^n, b^n)$. □

Remarque 10.7 – *Soient $n, m \in \mathbb{N}^*$, si $d = \text{pgcd}(n, m)$ alors il existe u et v premiers entre eux tels que $n = du$ et $m = dv$, on a alors $\text{ppcm}(n, m) = d\,\text{ppcm}(u, v) = duv$.*

V NOMBRES PREMIERS, DÉCOMPOSITION

1) Définition

Définition 10.8

*Un entier $p \in \mathbb{Z}$ est dit **premier** lorsque $p \geqslant 2$, et que ses seuls diviseurs positifs sont 1 et p. L'ensemble des nombres premiers est noté $\mathcal{P}$.*

Exemples :

- $2, 3, 5, 7, 11, 13, 17, 19, 23, \ldots$ sont des nombres premiers.
- Les nombres de *Fermat*[3] : $F_n = 2^{2^n} + 1$ sont premiers pour $n = 0, 1, 2, 3, 4$ mais pas pour $n = 5$, car $F_5 = 641 \times 6700417$.
- Les nombres de *Mersennes*[4] : $M_p = 2^p - 1$ où $p \in P$, sont premiers pour $p = 2, 3, 5, 7, 127, \ldots$ mais pas pour $p = 11$, car $M_{11} = 23 \times 89$.

★**Exercice 10.4**

1/ *Soient $a, b \in \mathbb{Z}$ et $n \in \mathbb{N}$, montrer que $a - b \mid a^n - b^n$. Si n est impair, montrer que $a + b \mid a^n + b^n$.*

2/ *Montrer que si $2^p + 1$ est un nombre premier alors p est une puissance de 2.*

3/ *Montrer que si $2^p - 1$ est un nombre premier, alors p est un nombre premier.*

Propriétés élémentaires :

a) Si p est premier, alors $\forall n \in \mathbb{Z}$, soit $p \mid n$, soit $\text{pgcd}(n, p) = 1$.
Preuve : Soit $d = \text{pgcd}(p, n)$, alors $d \mid p$ donc soit $d = 1$ soit $d = p$, c'est à dire, soit $p \wedge n = 1$, soit $p \mid n$. □

b) Si $n \geqslant 2$, alors n possède au moins un diviseur premier.
Preuve : Soit $B = \{|d| \ / \ d \mid n \text{ et } d \neq 1\}$, alors B est une partie de $\mathbb{N}$ non vide ($|n| \in B$), soit p un diviseur de n avec $|p| \in B$ **minimal**, si $d \mid p$ avec d positif et $d \neq 1$, alors $d \mid n$ et donc $|d| \in B$, d'où $|d| \geqslant |p|$, or $d \mid p$, donc $|d| \leqslant |p|$ et finalement $|d| = |p|$, d'où $d = p$ et donc p est premier. □

c) L'ensemble $\mathcal{P}$ est infini.
Preuve : Si $\mathcal{P}$ est fini, alors $\mathcal{P} = \{p_1, \ldots, p_n\}$, posons $N = 1 + p_1 \times \ldots \times p_n$, alors $N > 1$, donc N admet au moins un diviseur premier q, comme $q \in \mathcal{P}$, on a $q \mid p_1 \times \ldots \times p_n$, et comme $q \mid N$, on a $q \mid 1$ ce qui est absurde, donc $\mathcal{P}$ est infini. □

d) Si p est premier et si $p \mid nm$, alors $p \mid n$ ou $p \mid m$.
Preuve : Supposons que p ne divise pas n, alors $n \notin p\mathbb{Z}$ donc $\text{pgcd}(p, n) = 1$ et par conséquent $p \mid m$ (d'après le théorème de *Gauss*). □

e) Si $n > 1$ n'a pas de diviseur autre que 1 dans l'intervalle $[1; \sqrt{n}]$, alors n est premier.
Preuve : Si n est non premier alors on peut écrire $n = pq$ avec $p > 1$ et $q > 1$. Si les deux étaient strictement supérieurs à $\sqrt{n}$ alors on aurait $pq > n$ ce qui est absurde, donc un des deux est dans $[1; \sqrt{n}]$. Le résultat s'en déduit par contraposée. □

3. *FERMAT Pierre De* (1601 – 1665) : mathématicien amateur (éclairé !) l'un des plus féconds de son époque mais qui faisait peu de démonstrations et publiait peu.

4. *MERSENNES Marin* (1588 – 1648) : moine français qui entretenait une correspondance suivie avec les mathématiciens de son époque.

f) Si p est premier, alors $\forall\, k \in [\![1\,;p-1]\!], p \mid \binom{p}{k}$. On en déduit que pour tout entier a et b, on a $(a+b)^p \equiv a^p + b^p \pmod p$.
Preuve : On a $k\binom{p}{k} = p\binom{p-1}{k-1}$ qui est donc divisible par p, mais comme $k \in [\![1\,;p-1]\!]$, p est premier avec k, par conséquent, d'après le théorème de *Gauss*, $p \mid \binom{p}{k}$. Pour le second point, on développe le binôme. □

Compléments : Soit $(p_n)_{n\geqslant 1}$ la suite strictement croissante des nombres premiers, la répartition de ces nombres encore aujourd'hui mal connue, cependant on a les quelques résultats suivants :

- Tout segment de la forme $[\![n\,;2n]\!]$ contient au moins un nombre premier (théorème de *Bertrand*).
- Si $a, b \in \mathbb{N}^*$ sont premier entre eux, alors il existe une infinité de nombre premiers de la forme $an+b$ (théorème de *Dirichlet*).
- $p_n \underset{+\infty}{\sim} n\ln(n)$ (théorème de *Hadamard*).

Théorème 10.21 (petit théorème de Fermat)
Si p est un nombre premier, alors pour tout entier n on a $n^p \equiv n \pmod p$. Et si $n \notin p\mathbb{Z}$, alors $n^{p-1} \equiv 1 \pmod p$.

Preuve : Pour $n \in \mathbb{N}$ on fait une récurrence : la propriété est vraie au rang 0, supposons la vraie au rang n, alors $(n+1)^p \equiv n^p + 1 \pmod p$, en appliquant l'hypothèse de récurrence, on a $(n+1)^p \equiv n+1 \pmod p$.

On remarque ensuite que $(-1)^p \equiv -1 \pmod p$ car soit $p = 2$, soit p est premier impair, on en déduit que $(-n)^p \equiv -n \pmod p$. On a donc pour tout entier n, $p \mid n^p - n = n(n^{p-1}-1)$, si p ne divise pas n alors p est premier avec n, donc $p \mid n^{p-1}-1$, c'est à dire $n^{p-1} \equiv 1 \pmod p$. □

★**Exercice 10.5** *Soient u et v des entiers premiers entre eux, montrer que $u+v$ et uv sont premiers entre eux.*

2) Décomposition en facteurs premiers

Théorème 10.22 (décomposition en produit de facteurs premiers)
Tout élément $n \in \mathbb{Z}$, autre que ± 1, est un produit de nombres premiers. Plus précisément, il existe $r \geqslant 1$, il existe $p_1, \dots, p_r \in \mathcal{P}$ (distincts), il existe des entiers $\alpha_1, \dots, \alpha_r \in \mathbb{N}^*$, il existe $\lambda \in \{-1;1\}$ tels que :
$$n = \lambda \times p_1^{\alpha_1} \times p_2^{\alpha_2} \times \dots \times p_r^{\alpha_r}.$$

Preuve : On a $n = \lambda \times |n|$ avec $\lambda = \pm 1$. On se ramène ainsi au cas où n est positif.

Par récurrence sur n : pour $n = 2$ il n'y a rien à montrer car 2 est premier. Supposons le théorème démontré jusqu'au rang $n \geqslant 2$, alors $n+1$ admet au moins un diviseur premier p, donc $n+1 = pk$, si $k = 1$ alors $n+1$ est premier, sinon k est un produit de facteurs premiers (HR), donc $n+1$ aussi. □

Théorème 10.23 (unicité de la décomposition)
Si $n \in \mathbb{Z}$ s'écrit sous la forme :
$$n = \lambda \times p_1^{\alpha_1} \times \dots \times p_r^{\alpha_r} = \mu \times q_1^{\beta_1} \times \dots \times q_s^{\beta_s},$$
avec $p_1, \dots, p_r \in \mathcal{P}$ (distincts),$\alpha_1, \dots, \alpha_r \in \mathbb{N}^*, q_1, \dots, q_s \in \mathcal{P}$ (distincts), $\beta_1, \dots, \beta_s \in \mathbb{N}^*$, et $\lambda, \mu \in \{-1;1\}$ alors $r = s$, $\lambda = \mu$ et il existe une permutation σ de $[\![1\,;r]\!]$ telle que pour $i \in [\![1\,;r]\!], p_i = q_{\sigma(i)}, \alpha_i = \beta_{\sigma(i)}$. La décomposition est unique (à l'ordre près).

Preuve : Si $p_1 \notin \{q_1, \dots, q_s\}$, alors p_1 est premier avec $q_1, \dots, q_s$, donc p_1 est premier avec $q_1^{\beta_1} \times \dots \times q_s^{\beta_s}$, *i.e.* p_1 est premier avec n, ce qui est absurde puisque $p_1 \mid n$, donc $p_1 \in \{q_1, \dots, q_s\}$. Finalement on a $\{p_1, \dots, p_r\} \subset \{q_1, \dots, q_s\}$ et par symétrie on a l'égalité des deux ensembles, donc $r = s$. Quitte à permuter les indices que la famille (q_i), on peut supposer que $p_1 = q_1, \dots, p_r = q_r$.

Le théorème de *Gauss* entraîne que $p_k^{\alpha_k} \mid p_k^{\beta_k}$, donc $\alpha_k \leqslant \beta_k$, par symétrie on a $\beta_k \leqslant \alpha_k$, et donc $\alpha_k = \beta_k$, ce qui termine la preuve. □

3) Notion de valuation p-adique

Si n est un entier naturel non nul et p un nombre premier, alors l'ensemble $\left\{k \in \mathbb{N} \;/\; p^k \mid n\right\}$ est non vide (contient 0) et majoré par n (on peut montrer par récurrence que $p^n \geqslant$n), cet ensemble admet donc un maximum :

Définition 10.9

Soit $p \in \mathcal{P}$ et $n \in \mathbb{N}^*$ on appelle valuation p-adique de n, notée $v_p(n)$, le plus grand entier k tel que $p^k \mid n$. La définition s'étend à $\mathbb{Z}$, en posant $v_p(-n) = v_p(n)$ si $n < 0$ et $v_p(0) = +\infty$.

Remarque 10.8 :

- $v_p(n) = k \iff p^k \mid n$ et $p^{k+1} \nmid n \iff \boxed{\exists\, q \in \mathbb{N},\, n = p^k q \text{ avec } p \wedge q = 1}$.
- $v_p(n) \geqslant 1 \iff p \mid n$, auquel cas $v_p(n)$ est la puissance de p dans la décomposition de n en facteurs premiers.
- $\left\{k \in \mathbb{N} \;/\; p^k \mid n\right\} = [\![0\,;v_p(n)]\!]$.

À retenir

Pour tout entier $n \geqslant 2$, la décomposition de n en produit de facteurs premiers s'écrit : $n = \prod\limits_{p \in \mathcal{P}} p^{v_p(n)}$.

En effet, seul un nombre fini de valuations sont non nulles (les autres donnent un facteur égal à 1).

Théorème 10.24 (Propriétés)

$\forall n, m \in \mathbb{Z}$, on a :

a) $\forall p \in \mathcal{P}$, $v_p(nm) = v_p(n) + v_p(m)$.

b) $\forall p \in \mathcal{P}$, $v_p(n+m) \geqslant \min(v_p(n); v_p(m))$.

c) $n \mid m \iff \forall p \in \mathcal{P},\ v_p(n) \leqslant v_p(m)$.

d) Si n et m sont non nuls alors $\forall p \in \mathcal{P}$:

$$v_p(n \wedge m) = \min(v_p(n); v_p(m)) \text{ et } v_p(n \vee m) = \max(v_p(n); v_p(m)).$$

Preuve :

a) Si un des deux est nul, c'est évident. Supposons n et m non nuls, soit $n = p^k q$ avec $p \wedge q = 1$ et $m = p^{k'} q'$ avec $p \wedge q' = 1$, d'où $nm = p^{k+k'} qq'$ et $p \wedge (qq') = 1$, donc $v_p(nm) = kk'$.

b) Si un des deux est nul, c'est évident. Supposons n et m non nuls, et avec les mêmes notations, supposons $k \leqslant k'$, alors $n + m = p^k[q + p^{k'-k} q']$ donc $v_p(n+m) \geqslant k$. On remarque qu'il y a égalité lorsque $k \neq k'$.

c) Si $m = 0$ c'est évident, supposons $m \neq 0$ et $n \mid m$, alors pour tout premier p, $\left\{k \in \mathbb{N} \;/\; p^k \mid n\right\} \subset \left\{k \in \mathbb{N} \;/\; p^k \mid m\right\}$ donc $v_p(n) \leqslant v_p(m)$. La réciproque est évidente.

d) Soit $d = n \wedge m$ alors $v_p(d) \leqslant v_p(n)$ et $v_p(d) \leqslant v_p(m)$, donc $v_p(d) \leqslant \min(v_p(n); v_p(m))$. D'autre part $p^{\min(v_p(n);v_p(m))}$ divise n et m donc divise d d'où $\min(v_p(n); v_p(m)) \leqslant v_p(d)$, par conséquent $\min(v_p(n); v_p(m)) = v_p(d)$. Pour le ppcm on peut utiliser le fait que $(n \wedge m)(n \vee m) = |nm|$ et donc $v_p(n \vee m) = v_p(nm) - v_p(n \wedge m) = v_p(n) + v_p(m) - \min(v_p(n); v_p(m)) = \max(v_p(n); v_p(m))$.

□

À retenir : formules du pgcd et du ppcm

Il découle du théorème ci-dessus que :

$$\operatorname{pgcd}(n, m) = \prod_{p \in \mathcal{P}} p^{\min(v_p(n);v_p(m))} \text{ et } \operatorname{ppcm}(n, m) = \prod_{p \in \mathcal{P}} p^{\max(v_p(n);v_p(m))}.$$

★**Exercice 10.6** *Soient n et m deux naturels non nuls, premiers entre eux, tels que le produit nm est un carré, montrer que n et m sont des carrés.*

4) Applications

- Si $n \neq \pm 1$, alors la décomposition de n en produit de facteurs premiers permet de trouver tous les diviseurs de n.
  En effet : Si $n = \lambda \times p_1^{\alpha_1} \times \ldots \times p_r^{\alpha_r}$, soit d est un diviseur positif de n, si p est un diviseur premier de d, alors p est un diviseur premier de n, donc $p \in \{p_1, \ldots, p_r\}$, donc d s'écrit sous la forme :

$$\boxed{d = p_1^{\beta_1} \times \ldots \times p_r^{\beta_r} \text{ avec } 0 \leqslant \beta_k \leqslant \alpha_k}$$

– Si $n, m \notin \{-1; 1\}$, alors à partir de leur décomposition en produit de facteurs premiers, on peut calculer $\operatorname{pgcd}(n,m) = \prod\limits_{p\in\mathcal{P}} p^{\min(v_p(n);v_p(m))}$ et $\operatorname{ppcm}(n,m) = \prod\limits_{p\in\mathcal{P}} p^{\max(v_p(n);v_p(m))}$. Plus précisément : Si $n = \lambda \times p_1^{\alpha_1} \times \ldots \times p_r^{\alpha_r}$ et $m = \mu \times q_1^{\beta_1} \times \ldots \times q_s^{\beta_s}$, alors les diviseurs premiers communs à n et m doivent appartenir à $\{p_1, \ldots, p_r\} \cap \{q_1, \ldots, q_s\}$, d'où la discussion :
- $\{p_1, \ldots, p_r\} \cap \{q_1, \ldots, q_s\} = \varnothing$, alors n et m sont premiers entre eux, *i.e.* $\operatorname{pgcd}(n,m) = 1$ et donc $\operatorname{ppcm}(n,m) = |nm|$.
- $\{p_1, \ldots, p_r\} \cap \{q_1, \ldots, q_s\} = \{v_1, \ldots, v_t\}$, alors quitte à changer la numérotation, on peut supposer que $p_1 = q_1 = v_1, \ldots, p_t = q_t = v_t$ sont les diviseurs premiers communs à n et m.

$$\boxed{\operatorname{pgcd}(n,m) = v_1^{k_1} \times \ldots \times v_t^{k_t}} \text{ avec } k_i = \min(\alpha_i, \beta_i) \text{ pour } i \in [\![1\,;t]\!].$$

Et :

$$\boxed{\operatorname{ppcm}(n,m) = v_1^{k_1} \times \ldots \times v_t^{k_t} \times p_{t+1}^{\alpha_{t+1}} \times \ldots \times p_r^{\alpha_r} \times q_{t+1}^{\beta_{t+1}} \times \ldots \times q_s^{\beta_s}}$$
avec $k_i = \max(\alpha_i, \beta_i)$ pour $i \in [\![1\,;t]\!]$.

☞**Exemple** : $336 = 2^4 \times 3 \times 7$ et $420 = 2^2 \times 3 \times 5 \times 7$, donc $\operatorname{pgcd}(336, 420) = 2^2 \times 3 \times 7 = 84$, et $\operatorname{ppcm}(336, 420) = 2^4 \times 3 \times 5 \times 7$.

VI SOLUTION DES EXERCICES

Solution 10.1 *$a^n - b^n$ se factorise par $a - b$ (identité de Bernoulli), l'autre facteur est entier.*

Solution 10.2 *$a = 17$ et $b = 12$ sont premiers entre eux, appliquons l'algorithme d'Euclide :*
$a = b \times 1 + 5$ d'où $r_1 = 5 = a - b$
$b = r_1 \times 2 + 2$, d'où $r_2 = 2 = b - 2r_1 = b - 2(a - b) = -2a + 3b$
$r_1 = r_2 \times 2 + 1$, d'où $r_3 = 1 = r_1 - 2r_2 = a - b + 4a - 6b = 5a - 7b$
On a ainsi une relation de Bézout entre a et b, on en déduit une solution particulière en multipliant par 3 : $15a - 21b = 3$, donc $(x_0 = 15, y_0 = -21)$ est une solution particulière.
L'équation équivaut alors à $a(x - x_0) = b(y_0 - y)$, d'après le théorème de Gauss, a et b étant premiers entre eux, on a $b \mid x - x_0$ et $a \mid y_0 - y$, i.e. $x = x_0 + bk$ et $y = y_0 - bk'$, en reportant dans la relation on voit que $k = k'$ et donc les solutions sont les couples $(x_0 + bk, y_0 - ak)$ avec $k \in \mathbb{Z}$.

Solution 10.3

1/ *À l'itération 0 ,on a $B_0 = b \in \mathbb{N}$. Si la boucle ne se termine jamais, alors on a une suite B_i de nombres non nuls. À l'itération $i + 1$, on a $A_{i+1} = B_i$ et B_{i+1} est le reste de la division de A_i par B_i, on en déduit par récurrence sur i, que (B_i) est une suite d'entiers positifs ($B_0 = b \in \mathbb{N}$) et strictement décroissante car $B_{i+1} < B_i$, ce qui est absurde, donc la boucle while se termine.*

2/ *Montrons l'invariant de boucle $P(i)$: $\operatorname{pgcd}(a, b) = \operatorname{pgcd}(A_i, B_i)$, par récurrence sur i, au rang 0, on a $A_0 = a$ et $B_0 = b$, $P(0)$ est donc vraie. Supposons $P(i)$ vraie et qu'il y a une itération $i + 1$, donc $B_i \neq 0$, à l'itération $i + 1$ on a $A_{i+1} = B_i$, et $A_i = B_iQ_i + B_{i+1}$ (division euclidienne de A_i par B_i), donc $\operatorname{pgcd}(A_i, B_i) = \operatorname{pgcd}(B_i, B_{i+1}) = \operatorname{pgcd}(A_{i+1}, B_{i+1})$, ce qui montre $P(i + 1)$.*

3/ *Soit n le nombre d'itérations, à l'issue de l'itération n, on a $\operatorname{pgcd}(a, b) = \operatorname{pgcd}(A_n, B_n)$ (invariant), et comme il n'y a pas d'itération $n + 1$ cela signifie que $B_n = 0$, finalement $\operatorname{pgcd}(a, b) = \operatorname{pgcd}(A_n, 0) = A_n$. La valeur de la variable A qui est renvoyée est bien le pgcd cherché.*

Solution 10.4

1/ *$a \equiv b \pmod{a - b}$ d'où $a^n \equiv b^n \pmod{a - b}$, c'est à dire $a - b \mid a^n - b^n$.*
$a \equiv -b \pmod{a+b}$ d'où $a^n \equiv (-b)^n \pmod{a+b}$, si n est impair alors $a^n \equiv -b^n \pmod{a+b}$, et donc $a+b \mid a^n+b^n$.

2/ *Soit $N = 2^p + 1$ un nombre premier, on peut écrire $p = 2^r q$ avec q impair (r est valuation 2-adique de p), on a alors $N = 2^{2^r q} + 1 = \left(2^{2^r}\right)^q + 1^q$, q étant impair, N est divisible par $2^{2^r} + 1$, ce nombre est plus grand que 1 et N est premier, donc $2^{2^r} + 1 = N = 2^p + 1$, ce qui entraîne $p = 2^r$ (et donc $q = 1$).*

3/ *Soit $N = 2^p - 1$ un nombre premier, soit d un diviseur positif de p, on peut écrire $p = dq$ avec $q \in \mathbb{N}$, on a alors $N = 2^{dq} + 1 = 2^{d^q} - 1^q$, donc N est divisible par $2^d - 1$, or N est premier, donc $2^d - 1 = 1$ ou bien $2^d - 1 = N = 2^p - 1$, ce qui entraîne $d = 1$ ou $d = p$, donc p est premier.*

Solution 10.5 *Par l'absurde, si p est un premier qui divise $\operatorname{pgcd}(u + v, uv)$, alors $p \mid uv$ donc p divise u ou v, mais p divise aussi $u + v$, donc p divise u et v ce qui est absurde.*

Solution 10.6 *Soit p un nombre premier, comme $n \wedge m = 1$, on a $v_p(n \wedge m) = 0$, or $v_p(n \wedge m) = \min(v_p(n); v_p(m))$, on a donc $v_p(n) = 0$ ou $v_p(m) = 0$, d'autre part, nm étant un carré d'entier, $v_p(nm)$ est pair, or $v_p(nm) = v_p(n) + v_p(m)$ et une des deux valuations est nulle, donc l'autre est forcément paire. Finalement, pour tout premier p, $v_p(n)$ est pair et $v_p(m)$ est pair, donc n et m sont des carrés d'entiers.*

Chapitre 11

Limites - Continuité

Sommaire

Dans ce chapitre, les fonctions considérées sont définies sur un intervalle non trivial de $\mathbb{R}$.

I LIMITES

1) Définition

Soit $f\colon I \to \mathbb{R}$ une fonction, soit a un élément de I ou bien une extrémité de I ($a \in \overline{\mathbb{R}}$), et soit $b \in \overline{\mathbb{R}}$, intuitivement on dira que b est la limite de $f(x)$ quand x tend vers a lorsque $f(x)$ peut être aussi voisin que l'on veut de b pourvu que x soit suffisamment voisin de a, d'où la définition :

Définition 11.1

On dit que f admet pour limite b en a lorsque : $\forall$ W*, voisinage de* b*,* $\exists$ V*, voisinage de* a*, tel que* $\forall x \in I, x \in I \cap V \implies f(x) \in W$*. Si c'est le cas, on notera :*

$$\lim_{x\to a} f(x) = b = \lim_a f = b, \text{ ou encore } f(x) \xrightarrow[x\to a]{} b.$$

À retenir : $\lim_a f = b$ **signifie**

- Si $a, b \in \mathbb{R}$: $\forall\, \varepsilon > 0, \exists\, \alpha > 0, \forall\, x \in \mathrm{I}, |x-a| < \alpha \implies |f(x) - b| < \varepsilon$.
- Si $a \in \mathbb{R}$ et $b = +\infty$: $\forall\, \mathrm{A} \in \mathbb{R}, \exists\, \alpha > 0, \forall\, x \in \mathrm{I}, |x-a| < \alpha \implies f(x) > \mathrm{A}$.
- Si $a \in \mathbb{R}$ et $b = -\infty$: $\forall\, \mathrm{A} \in \mathbb{R}, \exists\, \alpha > 0, \forall\, x \in \mathrm{I}, |x-a| < \alpha \implies f(x) < \mathrm{A}$.
- Si $a = +\infty$ et $b \in \mathbb{R}$: $\forall\, \varepsilon > 0, \exists\, \mathrm{A} \in \mathbb{R}, \forall\, x \in \mathrm{I}, x > \mathrm{A} \implies |f(x) - b| < \varepsilon$.
- Si $a = +\infty$ et $b = +\infty$: $\forall\, \mathrm{A} \in \mathbb{R}, \exists\, \mathrm{B} \in \mathbb{R}, \forall\, x \in \mathrm{I}, x > \mathrm{B} \implies f(x) > \mathrm{A}$.
- Si $a = +\infty$ et $b = -\infty$: $\forall\, \mathrm{A} \in \mathbb{R}, \exists\, \mathrm{B} \in \mathbb{R}, \forall\, x \in \mathrm{I}, x > \mathrm{B} \implies f(x) < \mathrm{A}$.
- Si $a = -\infty$ et $b \in \mathbb{R}$: $\forall\, \varepsilon > 0, \exists\, \mathrm{B} \in \mathbb{R}, \forall\, x \in \mathrm{I}, x < \mathrm{B} \implies |f(x) - b| < \varepsilon$.
- Si $a = -\infty$ et $b = +\infty$: $\forall\, \mathrm{A} \in \mathbb{R}, \exists\, \mathrm{B} \in \mathbb{R}, \forall\, x \in \mathrm{I}, x < \mathrm{B} \implies f(x) > \mathrm{A}$.
- Si $a = -\infty$ et $b = -\infty$: $\forall\, \mathrm{A} \in \mathbb{R}, \exists\, \mathrm{B} \in \mathbb{R}, \forall\, x \in \mathrm{I}, x < \mathrm{B} \implies f(x) < \mathrm{A}$.

☞**Exemples** :

- Soit $f(x) = x^2$, montrons que $\lim_{+\infty} f = +\infty$: soit $\mathrm{A} \in \mathbb{R}$, posons $\mathrm{B} = \sqrt{|\mathrm{A}|}$, si $x > \mathrm{B}$ alors $x^2 > \mathrm{B}^2 = |\mathrm{A}| \geqslant \mathrm{A}$ donc $f(x) > \mathrm{A}$.
- Soit $f(x) = x^2$ et soit $a \in \mathbb{R}$, montrons que $\lim_a f = a^2$: soit $\varepsilon > 0$, $|x^2 - a^2| = |x-a||x+a|$, si $|x-a| < \alpha$, alors $|x^2 - a^2| < \alpha(\alpha + 2|a|)$, si on prend $\alpha = \min(1; \frac{\varepsilon}{1+2|a|})$, alors $\alpha(\alpha + 2|a|) \leqslant \alpha(1 + 2|a|) \leqslant \varepsilon$, donc $\forall\, x \in \mathbb{R}, |x-a| < \alpha \implies |x^2 - a^2| < \varepsilon$.

Remarque 11.1 –

a) *Si* $\lim_a f = b$ *alors* $\lim_a |f| = |b|$, *mais la réciproque est fausse sauf pour* $b = 0$.

b) *Lorsque* $b \in \mathbb{R}$, $\lim_a f = b \iff \lim_a |f(x) - b| = 0 \iff \lim_a f(x) - b = 0$.

La définition de la limite d'une suite, que nous avons vue dans un chapitre précédent, peut s'énoncer ainsi en terme de voisinage :

Définition 11.2 (Retour sur les suites)

On dit que la suite (u_n) *admet pour limite* $\ell \in \overline{\mathbb{R}}$ *lorsque :*

$$\forall\, \mathrm{W}, \text{ voisinage de } \ell, \exists\, \mathrm{N} \in \mathbb{N}, \forall\, n \in \mathbb{N}, n \geqslant \mathrm{N} \implies u_n \in \mathrm{W}.$$

2) Premières propriétés

Théorème 11.1

Soit $f : \mathrm{I} \to \mathbb{R}$ *une fonction et soit* a *un élément ou une extrémité de* I.

- *Si f admet une limite en a, alors celle - ci est unique.*
- *Si* f *admet une limite* ***finie*** *en a, alors* f *est* ***bornée au voisinage*** *de* a *(réciproque fausse).*
- *Si* $\lim_a f = b$ *et si* $\alpha < b$ *(respectivement* $b < \alpha$*), alors au voisinage de* a f *est strictement supérieure à* α *(respectivement* f *est strictement inférieure à* α*).*
- *Si* $\lim_a f = b$ *avec* $a \in \mathrm{I}$*, alors nécessairement* $b = f(a)$.

Preuve : Pour les trois premiers points, la preuve est tout à fait analogue à celle faite pour les suites.
Pour le quatrième point : tout voisinage de b doit contenir $f(a)$, on en déduit par l'absurde que $b = f(a)$. □

Théorème 11.2 (caractérisation séquentielle de la limite)

$\lim_a f = b \iff$ *pour toute suite* (u_n) *d'éléments de* I *qui tend vers* a *(dans* $\overline{\mathbb{R}}$*), la suite* $(f(u_n))$ *tend vers* b *(dans* $\overline{\mathbb{R}}$*).*

Preuve : Supposons que $\lim_a f = b$ et soit (u_n) une suite d'éléments de I telle que $u_n \to a$. Soit W un voisinage de b, il existe V un voisinage de a tel que $x \in \mathrm{I} \cap \mathrm{V} \implies f(x) \in \mathrm{W}$. Comme $u_n \to a$, il existe un entier $\mathrm{N} \in \mathbb{N}$ tel que $n \geqslant \mathrm{N} \implies u_n \in \mathrm{V}$, or les termes u_n sont dans I donc si $n \geqslant \mathrm{N}$ alors $u_n \in \mathrm{I} \cap \mathrm{V}$ et donc $f(u_n) \in \mathrm{W}$, ce qui prouve que $f(u_n) \to b$.

Supposons maintenant que pour toute suite (u_n) d'éléments de I qui tend vers a, la suite $(f(u_n))$ tend vers b. Si la fonction f n'a pas pour limite b en a, alors il existe un voisinage W de b tel que pour tout voisinage V de a, il existe $x \in \mathrm{I} \cap \mathrm{V}$ tel que $f(x) \notin \mathrm{W}$. En prenant pour $n \in \mathbb{N}^*$ des voisinages de la forme $\mathrm{V}_n =]a - \frac{1}{n}; a + \frac{1}{n}[$ si $a \in \mathbb{R}$, $\mathrm{V}_n =]n; +\infty[$ si $a = +\infty$, ou $\mathrm{V}_n =]-\infty; -n[$ si $a = -\infty$, on construit une suite (u_n) d'éléments de I telle que $u_n \in \mathrm{V}_n$ et $f(u_n) \notin \mathrm{W}$, il est facile de voir que la suite (u_n) tend vers a, donc la suite $(f(u_n))$ tend vers b, à partir d'un certain rang on doit donc avoir $f(u_n) \in \mathrm{W}$ ce qui est contradictoire, donc $\lim\limits_{a} f = b$. □

Applications :

- Ce théorème peut être utilisé pour montrer qu'une fonction f n'a pas de limite en a.
  Par exemple, la fonction $f : x \mapsto \sin(x)$ n'a pas de limite en $+\infty$ car la suite u définie par $u_n = \frac{\pi}{2} + n\pi$ tend vers $+\infty$ mais la suite $(f(u_n) = (-1)^n)$ n'a pas de limite.
- Ce théorème peut être également utilisé pour prouver les propriétés de la limite d'une fonction en se ramenant à celles des suites.

Voici une conséquence avec les suites (que l'on utilisait déjà de manière assez naturelle) :

Théorème 11.3

Soit $f :]\mathrm{A}; +\infty[\to \mathbb{R}$ une fonction telle que $\lim\limits_{+\infty} f = b \in \overline{\mathbb{R}}$, alors la suite (u_n) définie (à partir d'un certain rang) par $u_n = f(n)$ a pour limite b.

Preuve : On applique la caractérisation séquentielle à la suite u. □

3) Limite à gauche, limite à droite

Définition 11.3

*Soit $f : \mathrm{I} \to \mathbb{R}$ une fonction, soit a un élément de I ou une extrémité **réelle** de I, et soit $b \in \overline{\mathbb{R}}$.*

- *Si $\mathrm{I} \cap]-\infty; a[\neq \emptyset$: on dit que b est la limite à gauche en a de f lorsque :*
  $$\forall\ \mathrm{W}, \text{voisinage de } b, \exists\, \alpha > 0, \forall\, x \in \mathrm{I}, x \in]a - \alpha; a[\implies f(x) \in \mathrm{W} \quad .$$
  Notations : $\lim\limits_{a^-} f = \lim\limits_{x \to a^-} f(x) = \lim\limits_{\substack{x \to a \\ x<a}} f(x) = b$.
- *Si $\mathrm{I} \cap]a; +\infty[\neq \emptyset$: on dit que b est la limite à droite en a de f lorsque :*
  $$\forall\ \mathrm{W}, \text{ voisinage de } b, \exists\, \alpha > 0, \forall\, x \in \mathrm{I}, x \in]a; a + \alpha[\implies f(x) \in \mathrm{W}.$$
  Notations : $\lim\limits_{a^+} f = \lim\limits_{x \to a^+} f(x) = \lim\limits_{\substack{x \to a \\ x>a}} f(x) = b$.

☞**Exemple** : Soit $f(x) = \lfloor x \rfloor$ et soit $a \in \mathbb{Z}$, alors $\lim\limits_{a^+} f = a$ et $\lim\limits_{a^-} f = a - 1$.

Théorème 11.4

On a $\lim\limits_{\substack{x \to a \\ x \neq a}} f(x) = b \iff \lim\limits_{a^+} f = \lim\limits_{a^-} f = b$. *Et lorsque $a \in \mathrm{I}$:*

$$\lim_{x \to a} f(x) = b \iff \left(f(a) = b \text{ et } \lim_{\substack{x \to a \\ x \neq a}} f(x) = b\right).$$

Preuve : Celle - ci est simple et laissée en exercice. □

☞**Exemple** : Soit f la fonction définie sur $\mathbb{R}$ par $f(x) = \begin{cases} x+1 & \text{si } x < 0 \\ 0 & \text{si } x = 0 \\ 1 - x^2 & \text{si } x > 0 \end{cases}$. Il est facile de voir que : $\lim\limits_{\substack{x \to 0 \\ x \neq 0}} f(x) = 1$, mais la fonction f n'a pas de limite en 0 car $f(0) \neq 1$.

II PROPRIÉTÉS DES LIMITES

1) Limites et opérations

Soient $f, g \in \mathcal{F}(\mathrm{I}, \mathbb{R})$ et soit a un élément de I ou une extrémité de I. Si $\lim\limits_{a} f = \ell$ et $\lim\limits_{a} g = \ell'$ (dans $\overline{\mathbb{R}}$), alors :

Théorème 11.5

- $\lim\limits_a f+g=\ell+\ell'$ *sauf si* $\ell=+\infty$ *et* $\ell'=-\infty$ *(ou l'inverse) :* ***forme indéterminée****.*
- $\lim\limits_a f\times g=\ell\ell'$ *sauf si* $\ell=0$ *et* $\ell'=\pm\infty$ *(ou l'inverse) :* ***forme indéterminée****.*
- *Si* $\lambda\in\mathbb{R}^*$, $\lim\limits_a \lambda f=\lambda\ell$.

Preuve : Pour le premier point : soit (u_n) une suite d'éléments de I qui tend vers a, alors $f(u_n)\to\ell$ et $g(u_n)\to\ell'$ donc (propriétés des suites) $f(u_n)+g(u_n)\to\ell+\ell'$ car nous ne sommes pas dans le cas d'une forme indéterminée, par conséquent la fonction $f+g$ a pour limite $\ell+\ell'$ en a. Le raisonnement est le même pour tous les autres points jusqu'au dernier. □

Théorème 11.6

Si f ne s'annule pas au voisinage de a alors :

$$\lim_a \frac{1}{f}=\begin{cases}\frac{1}{\ell} & \text{si } \ell\in\mathbb{R}^*\\ 0 & \text{si } \ell=\pm\infty\\ +\infty & \text{si } \ell=0 \text{ et } f>0 \text{ au voisinage de } a\\ -\infty & \text{si } \ell=0 \text{ et } f<0 \text{ au voisinage de } a\\ \text{n'existe pas sinon}\end{cases}$$

Preuve : Dans le dernier cas on a $\ell=0$ et sur tout voisinage de a f prend des valeurs strictement positives et des valeurs strictement négatives (f n'est pas de signe constant), on peut donc construire deux suites (u_n) et (v_n) qui tendent vers a et telles que $f(u_n)>0$ et $f(v_n)<0$, mais alors $\frac{1}{f(u_n)}\to+\infty$ et $\frac{1}{f(v_n)}\to-\infty$, donc $\frac{1}{f}$ n'a pas de limite en a. □

☞**Exemples** :

- Soit $a\in\mathbb{R}$, $\lim\limits_a x=a$ d'où $\forall\ n\in\mathbb{N}^*$, $\lim\limits_a x^n=a^n$ (encore vrai pour $n=0$). On en déduit que si P est une fonction polynomiale, alors $\lim\limits_a \mathrm{P}=\mathrm{P}(a)$.
- Si $\mathrm{R}=\frac{\mathrm{P}}{\mathrm{Q}}$ est une fraction rationnelle et si $\mathrm{Q}(a)\neq 0$, alors $\lim\limits_a \mathrm{R}=\mathrm{R}(a)$.

2) Limite et relation d'ordre

Théorème 11.7

Soient $f,g,h:\mathrm{I}\to\mathbb{R}$ trois fonctions et soit a un élément de I ou une extrémité de I.

- *On suppose qu'au voisinage de a, $f\leqslant g$, alors :*
  - *Si* $\lim\limits_a f=+\infty$ *alors* $\lim\limits_a g=+\infty$.
  - *Si* $\lim\limits_a g=-\infty$ *alors* $\lim\limits_a f=-\infty$.
- *Si $f\leqslant h\leqslant g$ au voisinage de a et si* $\lim\limits_a f=\lim\limits_a g=\ell\in\overline{\mathbb{R}}$, *alors* $\lim\limits_a h=\ell$ *(théorème des gendarmes ou de l'étau).*
- *Si $f\leqslant g$ au voisinage de a et si f et g ont chacune une limite dans $\overline{\mathbb{R}}$, alors* $\lim\limits_a f\leqslant\lim\limits_a g$ *(théorème du passage à la limite).*
- *Si* $\lim\limits_a f=0$ *et si g est bornée* ***au voisinage de*** *a, alors* $\lim\limits_a f\times g=0$.
- *Si* $\lim\limits_a f=+\infty$ *(respectivement $-\infty$) et si g est minorée au voisinage de a (respectivement majorée), alors* $\lim\limits_a f+g=+\infty$ *(respectivement $-\infty$).*

Preuve : Supposons $\lim\limits_a f=+\infty$, soit (u_n) une suite d'éléments de I qui tend vers a, à partir d'un certain rang, on a $f(u_n)\leqslant g(u_n)$, or $f(u_n)\to+\infty$, donc $g(u_n)\to+\infty$ et par conséquent $\lim\limits_a g=+\infty$. Pour le deuxième cas, on raisonne sur $-f$ et $-g$.

Pour les autres points on procède de la même façon, en se ramenant aux suites. □

Remarque 11.2 –

- *Si $|f|\leqslant g$ au voisinage de a et si* $\lim\limits_a g=0$, *alors* $\lim\limits_a f=0$.
- *On peut avoir $f<g$ au voisinage de a et* $\lim\limits_a f=\lim\limits_a g$. *Dans un passage à la limite les inégalités deviennent larges.*

3) Limite et composition des fonctions

Soit $f : \mathrm{I} \to \mathbb{R}$ une fonction, soit a un élément de I ou une extrémité de I, et soit $g : \mathrm{J} \to \mathbb{R}$ une autre fonction avec $\mathrm{Im}(f) \subset \mathrm{J}$.

Théorème 11.8

Si $\lim\limits_{a} f = b$, alors b appartient à J ou b est une extrémité de J.

Preuve : Il suffit de distinguer les cas sur J, par exemple, si $\mathrm{J} =]\alpha\,;\beta[$, alors $\forall\, x \in \mathrm{I}, \alpha < f(x) < \beta$, par passage à la limite, on obtient $\alpha \leqslant b \leqslant \beta$. Les autres cas se traitent de la même façon. □

Théorème 11.9 (composition des limites)

Si $\lim\limits_{a} f = b$ et $\lim\limits_{b} g = \ell$ (dans $\overline{\mathbb{R}}$), alors $\lim\limits_{a} g \circ f = \ell$.

Preuve : Soit (u_n) une suite d'éléments de I qui tend vers a, la suite $(f(u_n))$ est une suite d'éléments de J qui tend vers b, donc la suite $(g[f(u_n)])$ tend vers ℓ, ce qui prouve que $\lim\limits_{a} g \circ f = \ell$. □

Dans la pratique, ce théorème est parfois appelé **changement de variable** dans une limite. Il dit en effet que si on pose $\mathrm{X} = f(x)$, alors comme $\mathrm{X} \underset{x \to a}{\longrightarrow} b$, on a $\lim\limits_{x \to a} g(f(x)) = \lim\limits_{\mathrm{X} \to b} g(\mathrm{X}) = \ell$.

★**Exercice 11.1** *Calculer $\lim\limits_{0^+} f$ avec $f(x) = \frac{e^{\sin(x)\ln(x)} - 1}{\sin(x)\ln(x)}$.*

4) Limite et sens de variation

Théorème 11.10

Si $f : \mathrm{I} \to \mathbb{R}$ est croissante, en notant a la borne de gauche de I et b la borne de droite :
• si f est majorée, alors f admet une limite finie à gauche en b qui est $\lim\limits_{b^-} f = \sup\limits_{x \in]a;b[} f(x)$. Si de plus si $b \in \mathrm{I}$, alors $\lim\limits_{b^-} f \leqslant f(b)$.
• si f est non majorée, alors f admet $+\infty$ comme limite à gauche en b.
• si f est minorée, alors f admet une limite finie à droite en a qui est $\lim\limits_{a^+} f = \inf\limits_{x \in]a;b[} f(x)$. Si de plus si $a \in \mathrm{I}$, alors $\lim\limits_{a^+} f \geqslant f(a)$.
• si f est non minorée, alors f admet $-\infty$ comme limite à droite en a.

Preuve : Démontrons deux cas :

Si f est majorée, soit $\mathrm{S} = \sup\limits_{]a;b[} f$, soit $\varepsilon > 0$, il existe un réel $x_0 \in]a;b[$ tel que $\mathrm{S} - \varepsilon < f(x_0)$. Si $x \in]x_0;b[$, alors f étant croissante, $\mathrm{S} - \varepsilon < f(x_0) \leqslant f(x) \leqslant \mathrm{S} < \mathrm{S} + \varepsilon$ ce qui entraîne $|f(x) - \mathrm{S}| < \varepsilon$ et donc $\lim\limits_{b^-} f = \mathrm{S}$. Si de plus $b \in \mathrm{I}$, alors comme f est croissante, f est majorée sur $]a;b[$ par $f(b)$, donc on a $\mathrm{S} \leqslant f(b)$.

Si f est croissante non minorée, soit $\mathrm{A} \in \mathbb{R}$ ce n'est pas un minorant de f, donc il existe un réel $x_0 \in]a;b[$ tel que $f(x_0) < \mathrm{A}$. Si $x \in]a;x_0[$, alors f étant croissante, $f(x) \leqslant f(x_0) < \mathrm{A}$ ce qui entraîne $f(x) < \mathrm{A}$ et donc $\lim\limits_{a^+} f = -\infty$. □

Remarque 11.3 –

a) Si f est croissante majorée sur I alors f admet une limite finie en b^- (limite non atteinte sur $]a;b[$ si la croissance est stricte).

b) Si f est croissante non majorée sur I alors f a pour limite $+\infty$ en b^-.

c) Si f est croissante minorée sur I alors f admet une limite finie en a^+ (limite non atteinte sur $]a;b[$ si la croissance est stricte).

d) Si f est croissante non minorée sur I alors f a pour limite $-\infty$ en a^+.

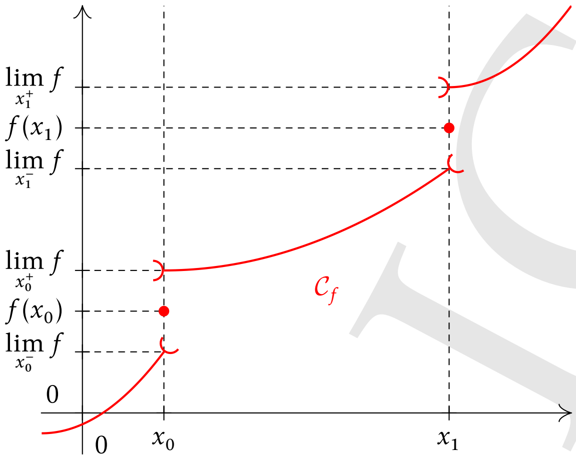
☞**Exemple** : Soit $f(x) = \ln(x)$. Pour $n \in \mathbb{N}^*, \ln(2^n) = n\ln(2)$, or $\ln(2) > 0$ car $1 < 2$ et f est strictement croissante, donc $n\ln(2) \to +\infty$, ce qui prouve que f est non majorée, comme elle est croissante, on a $\lim\limits_{+\infty} f = +\infty$.

Théorème 11.11 (de la limite monotone)

Si f est croissante sur I, soit a et b les bornes de I (dans $\overline{\mathbb{R}}$), pour tout réel $x_0 \in]a;b[$, f admet une limite finie à droite et à gauche en x_0, de plus on a : $\lim\limits_{x_0^-} f \leqslant f(x_0) \leqslant \lim\limits_{x_0^+} f$. Et si $x_1 \in]a;b[$ avec $x_0 < x_1$, alors : $\lim\limits_{x_0^+} f \leqslant \lim\limits_{x_1^-} f$.

Preuve : Sur l'intervalle $]a;x_0[$, la fonction f est croissante et majorée par $f(x_0)$, donc la fonction f a une limite finie à gauche en x_0 et d'après le théorème précédent : $\lim\limits_{x_0^-} f = \sup\limits_{t\in]a;x_0[} f(t) \leqslant f(x_0)$. Le raisonnement est le même à droite.
Si $x_0 < x_1 < b$, on applique le théorème précédent sur l'intervalle $]x_0;x_1[$: $\lim\limits_{x_0^+} f = \inf\limits_{t\in]x_0;x_1[} f(t) \leqslant \sup\limits_{t\in]x_0;x_1[} f(t) = \lim\limits_{x_1^-} f$.

□



Remarque 11.4 – *En changeant f et $-f$ et en utilisant que pour une partie non vide A de $\mathbb{R}$: $\inf(\mathrm{A}) = -\sup(-\mathrm{A})$, on obtient deux théorèmes analogues aux précédents pour les fonctions décroissantes.*

III CONTINUITÉ

1) Définitions

Définition 11.4

Soit $f : \mathrm{I} \to \mathbb{R}$ une fonction et soit $a \in \mathrm{I}$, on dit que f est

- *continue en a lorsque $\lim\limits_{t\to a} f(t) = f(a)$ (sinon on dit que a est un point de discontinuité de f).*
- *continue à gauche en a lorsque $\mathrm{I} \cap]-\infty; a[\neq \varnothing$ et $\lim\limits_{t\to a^-} f(t) = f(a)$.*
- *continue à droite en a lorsque $\mathrm{I} \cap]a; +\infty[\neq \varnothing$ et $\lim\limits_{t\to a^+} f(t) = f(a)$.*

Si f est continue en tout point de I, alors on dit que f est continue sur I. L'ensemble des fonctions continues sur I est noté $\mathcal{C}^0(\mathrm{I}, \mathbb{R})$.

Remarque 11.5 :

- *Les fonctions trigonométriques, logarithmes, exponentielles, puissances, polynomiales, rationnelles, ainsi que la fonction valeur absolue sont continues sur leur ensemble de définition.*
- *f est continue en $a \in \mathrm{I}$ si et seulement si $\forall \varepsilon > 0, \exists \alpha > 0, \forall x \in \mathrm{I}, |x-a| < \alpha \implies |f(x) - f(a)| < \varepsilon$.*
- *Si f est continue sur I et si $\mathrm{J} \subset \mathrm{I}$, alors f est continue sur J.*
- *f est continue en a si et seulement si $\lim\limits_{\substack{x\to a\\ x\neq a}} f(t) = f(a)$, lorsque a n'est pas une borne de I, ceci équivaut à $\lim\limits_{t\to a^+} f(t) = f(a)$ et $\lim\limits_{t\to a^-} f(t) = f(a)$, i.e. f est continue à gauche et à droite en a.*
- *Si f est continue en a, alors f est bornée au voisinage de a (car f a une limite finie en a).*
- *f est continue en a ssi pour toute suite (u_n) d'éléments de I, qui tend vers a, la suite $(f(u_n))$ tend vers $f(a)$.*

★**Exercice 11.2**

1/ Étudier la continuité de $x \mapsto \lfloor x \rfloor$ en $a \in \mathbb{R}$, distinguer $a \in \mathbb{Z}$ et $a \in \mathbb{R} \setminus \mathbb{Z}$.

2/ Montrer que la fonction caractéristique de $\mathbb{Q}$, $1_{\mathbb{Q}}$, est discontinue en tout point de $\mathbb{R}$.

Définition 11.5 (Prolongement par continuité)

Soit $f : \mathrm{I} \setminus \{a\} \to \mathbb{R}$ une fonction non définie en $a \in \mathrm{I}$, si f admet une limite finie ℓ en a, alors la fonction $\tilde{f} : \mathrm{I} \to \mathbb{R}$ définie par :

$$\tilde{f}(x) = \begin{cases} f(x) & \text{si } x \neq a \\ \ell & \text{si } x = a \end{cases}$$

est continue en a. Cette fonction est appelée ***prolongement de f par continuité en*** *a.*

☞**Exemple** : La fonction f définie sur $\mathbb{R}^*$ par $f(x) = \frac{\sin(x)}{x}$ admet un prolongement par continuité en 0 en posant $f(0) = 1$.

2) Théorèmes généraux

Théorème 11.12

Soient f, g deux fonctions continues sur I, et soit α un réel, alors :
- *$f + g$, $f \times g$ et αf sont continues sur I.*
- *Si g ne s'annule pas sur I alors $\frac{f}{g}$ est continue sur I.*
- *Si $h : \mathrm{J} \to \mathbb{R}$ est une fonction continue sur l'intervalle J et si $f(\mathrm{I}) \subset \mathrm{J}$, alors $h \circ f$ est continue sur I.*

Preuve : Ceci découle des propriétés des limites, par exemple : $\lim\limits_a f = f(a)$ et $\lim\limits_a g = g(a)$, donc $\lim\limits_a (f+g) = f(a) + g(a)$ (somme de limites finies), ce qui prouve que $f + g$ est continue en a. Les autres points se démontrent de la même façon. □

Conséquences :

a) Il découlent des théorèmes généraux que $\mathcal{C}^0(\mathrm{I}, \mathbb{R})$ est une $\mathbb{R}$-algèbre pour les opérations usuelles sur les fonctions.

b) Si f et g sont continues sur I alors $\sup(f, g)$ et $\inf(f, g)$ le sont (en particulier f^+ et f^- le sont), car $\sup(f, g) = \frac{f+g+|f-g|}{2}$ et $\inf(f, g) = \frac{f+g-|f-g|}{2}$.

★**Exercice 11.3** *Étudier la continuité de la fonction f avec $f(x) = \begin{cases} \frac{\ln(1+x)(x-\pi)}{\sin(x)} & \text{si } 0 < x < \pi \\ e^x - \cos(x) & \text{si } x \leqslant 0 \end{cases}$, y-a-t'il un prolongement par continuité en π ?*

IV FONCTIONS CONTINUES SUR UN INTERVALLE

1) Théorème des valeurs intermédiaires

Théorème 11.13

Soit $f : [a; b] \to \mathbb{R}$ une fonction continue sur le segment $[a; b]$ ($a < b$), si $f(a)$ et $f(b)$ sont de signes contraires, alors f s'annule au moins une fois, i.e. : $\exists\, \ell \in [a; b], f(\ell) = 0$.

Preuve : Méthode dichotomique : on construit deux suites (récurrentes) (a_n) et (b_n) en posant $a_0 = a$, $b_0 = b$, puis pour tout entier n :

si $f(\frac{a_n+b_n}{2})$ et $f(a_n)$ sont de signes contraires, alors on pose $a_{n+1} = a_n$ et $b_{n+1} = \frac{a_n+b_n}{2}$ (moitié de gauche), sinon, on pose $a_{n+1} = \frac{a_n+b_n}{2}$ et $b_{n+1} = b_n$ (moitié de droite).

On montre ensuite par récurrence, la propriété :

$$\mathrm{P}(n) : \boxed{\text{« } a_n, b_n \in [a; b],\ b_n - a_n = \tfrac{b-a}{2^n},\ f(a_n) \text{ et } f(b_n) \text{ sont de signes contraires »}.}$$

Pour $n = 0$: rien à faire. Si c'est vrai pour un entier $n \geqslant 0$: alors a_n et b_n sont dans $[a; b]$, donc $c_n = \frac{a_n+b_n}{2}$ aussi (milieu).

Si $f(c_n)$ et $f(a_n)$ sont de signes contraires, alors $b_{n+1} = \frac{a_n+b_n}{2}$ et $a_{n+1} = b_n$, on voit donc que a_{n+1} et b_{n+1} sont dans $[a, b]$, que $b_{n+1} - a_{n+1} = \frac{b_n-a_n}{2} = \frac{b-a}{2^{n+1}}$, que $f(a_{n+1})$ et $f(b_{n+1})$ sont de signes contraires, et que $a_n \leqslant a_{n+1}$ et $b_{n+1} \leqslant b_n$.

Si $f(c_n)$ et $f(a_n)$ ont le même signe, alors d'après l'hypothèse de récurrence, $f(c_n)$ et $f(b_n)$ sont de signes contraires, on a alors $a_{n+1} = \frac{a_n+b_n}{2}$ et $b_{n+1} = b_n$, on voit donc que a_{n+1} et b_{n+1} sont dans $[a, b]$, que $b_{n+1} - a_{n+1} = \frac{b_n-a_n}{2} = \frac{b-a}{2^{n+1}}$, que $f(a_{n+1})$ et $f(b_{n+1})$ sont de signes contraires, et que $a_n \leqslant a_{n+1}$ et $b_{n+1} \leqslant b_n$.

La propriété est donc vraie pour tout entier n, de plus la suite (a_n) est croissante, et la suite (b_n) est décroissante. Comme $b_n - a_n = \frac{b-a}{2^n} \to 0$, les deux suites sont adjacentes. Elles ont donc une limite commune $c \in [a;b]$ (passage à la limite). La fonction f étant continue en c, on a $f(a_n) \to f(c)$ et $f(b_n) \to f(c)$, donc $f(a_n) \times f(b_n) \to f(c)^2$, or pour tout n, $f(a_n) \times f(b_n) \leqslant 0$, donc $f(c)^2 \leqslant 0$ (passage à la limite), et donc $f(c) = 0$. □

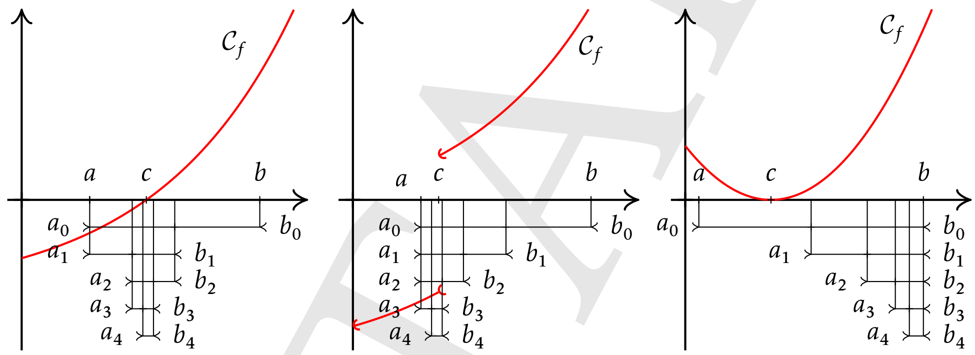
Remarque 11.6 –

– *Cette méthode permet de calculer des valeurs approchées de ℓ. Dans la preuve ci-dessus, on a pour tout n, a_n est une valeur approchée de ℓ (solution de $f(x) = 0$) par défaut à $\frac{b-a}{2^n}$ près car $|a_n - \ell| = \ell - a_n \leqslant b_n - a_n = \frac{b-a}{2^n}$. De même, pour tout n, b_n est une valeur approchée de ℓ par excès à $\frac{b-a}{2^n}$ près car $|b_n - \ell| = b_n - \ell \leqslant b_n - a_n = \frac{b-a}{2^n}$. Voici un algorithme en python :*

Listing 11.1 – *dichotomie*

```python
def dichotomie(f,a,b,epsilon): #f continue et f(a)*f(b)<=0 avec a<b
    while b-a >= epsilon:
        milieu = (a+b)/2.
        if f(a)*f(milieu) <= 0: #f s'annule dans la première moitié
            b = milieu
        else:
            a = milieu #f s'annule dans la deuxième moitié
    return (a+b)/2 #valeur approchée à epsilon/2 près
```

Invariant *: on peut vérifier que la proposition $\mathrm{P}(k)$: « $\mathrm{B}_k - \mathrm{A}_k = \frac{\mathrm{B}-\mathrm{A}}{2^k}$, et $f(\mathrm{A}_k) \times f(\mathrm{B}_k) \leqslant 0$ », est un invariant de la boucle while, qui permet de prouver la fonction, et sa terminaison. Quelques exemples d'utilisation de cet algorithme :*



f continue en c *f discontinue en c* *f ne change pas de signe.*

– *Il découle de ce théorème que si $f : \mathrm{I} \to \mathbb{R}$ est continue sur **l'intervalle** I et si f change de signe, alors f s'annule au moins une fois sur I.*
– *Une fonction continue sur un intervalle et qui ne s'annule pas, garde **un signe constant**. Ceci est faux si I n'est pas un intervalle, par exemple la fonction $x \mapsto \frac{1}{x}$ sur $\mathbb{R}^*$.*

★**Exercice 11.4** *Montrer que tout polynôme (réel) de degré impair admet au moins une racine réelle.*

Théorème 11.14 (des valeurs intermédiaires)
Si $f : \mathrm{I} \to \mathbb{R}$ est continue sur l'intervalle I, alors $f(\mathrm{I})$ est un intervalle.
Plus précisément, si $a, b \in \mathrm{I}$ et si α est un réel compris entre $f(a)$ et $f(b)$, alors il existe c entre a et b tel que $f(c) = \alpha$.

Preuve : Soient a, b deux réels distincts de I, supposons $a < b$, soit α un réel compris entre $f(a)$ et $f(b)$, posons $g(t) = f(t) - \alpha$, alors g est continue sur l'intervalle $[a;b]$ et $g(a)$ et $g(b)$ sont de signes contraires. D'après le théorème précédent, il existe $c \in [a;b]$ tel que $g(c) = 0$, *i.e.* $f(c) = \alpha$.

Posons $\mathrm{J} = f(\mathrm{I})$ et soient $u < v$ deux éléments de J, alors il existe $a, b \in \mathrm{I}$ (distincts) tels que $f(a) = u$ et $f(b) = v$. Soit $\alpha \in [u, v]$, alors il existe c entre a et b tel que $f(c) = \alpha$ donc $\alpha \in \mathrm{J}$, ce qui prouve que J est un intervalle. □

2) Continuité sur un segment

Théorème 11.15

L'image d'un segment $[a;b]$ par une fonction continue est un segment $[m;M]$.

Preuve : Soit $f\colon [a;b] \to \mathbb{R}$ une fonction continue sur le segment $[a;b]$, posons $J = f([a;b])$, on sait que J est un intervalle. Posons m = la borne de gauche de J, et M la borne de droite (dans $\overline{\mathbb{R}}$), il existe une suite (y_n) de J qui tend vers m, or $y_n \in f([a;b])$, donc il existe $x_n \in [a;b]$ tel que $f(x_n) = y_n$. La suite (x_n) est une suite de $[a;b]$, elle est donc bornée, d'après le théorème de *Bolzano-Weierstrass* on peut en extraire une suite convergente : $x_{\sigma(n)} \to \ell$, par passage à la limite on a $\ell \in [a;b]$, mais alors $f(x_{\sigma(n)}) \to f(\ell)$ car f est continue, c'est à dire $y_{\sigma(n)} \to f(\ell)$, or $y_{\sigma(n)} \to m$, donc $m = f(\ell)$. Ceci prouve que m est un réel et que $m \in J$, donc $m = \min(J)$. De même on montre que M est un réel que $M \in J$, finalement $J = [m;M]$. □

Remarque 11.7 – *Il en découle qu'une fonction continue sur un segment possède un maximum* (M) *et un minimum* (m). *On dit aussi parfois qu'une fonction continue sur un segment est bornée et atteint ses bornes.*

★**Exercice 11.5** *Montrer qu'une fonction continue sur un segment strictement positive, est minorée par un réel strictement positif.*

3) Uniforme continuité

Dans la définition de « f est continue en a », on a :

$$\forall\, \varepsilon > 0, \exists\, \alpha > 0, \forall\, x \in I, |x-a| < \alpha \implies |f(x) - f(a)| < \varepsilon.$$

Dans cette définition, le réel α **dépend de** a (et de ε bien entendu). On va distinguer dans la suite le cas où α ne dépend que de ε :

Définition 11.6

On dit que la fonction $f\colon I \to \mathbb{R}$ est uniformément continue sur I *lorsque :*

$$\forall\, \varepsilon > 0, \exists\, \alpha > 0, \forall\, a, x \in I, |x-a| < \alpha \implies |f(x) - f(a)| < \varepsilon.$$

Remarque 11.8 –

a) *Cette définition dépend aussi de l'ensemble* I, *on dit qu'elle a un caractère global, alors que la définition de la continuité en un point est locale car elle ne dépend que du point (pas de l'ensemble* I*).*

b) *La définition d'uniforme continuité est plus forte que la définition de continuité. Autrement dit, une fonction uniformément continue sur* I *est nécessairement continue sur* I. *Nous verrons que la réciproque est fausse en général.*

Définition 11.7 *(fonction lipschitzienne)*

On dit que la fonction $f\colon I \to \mathbb{R}$ est lipschitzienne lorsqu'il existe une constante $K \in \mathbb{R}^+$ tel que :

$$\forall\, x, y \in I, |f(x) - f(y)| \leqslant K|x-y|.$$

Ce qui signifie que tous les taux d'accroissements de f sont majorés en valeur absolue par K.

c) *Soit $K \in \mathbb{R}^+$, une fonction K-lipschitzienne sur* I *est nécessairement uniformément continue sur* I. *En effet, une telle fonction vérifie pour tout $x, y \in I$, $|f(x) - f(y)| \leqslant K|x-y|$, par conséquent, si on prend $\alpha = \frac{\varepsilon}{K+1}$ alors on a $|f(x) - f(y)| \leqslant \frac{K}{K+1}\varepsilon < \varepsilon$. Nous verrons dans le chapitre sur la dérivation, que si f est dérivable et si f' est majorée en valeur absolue par une constante* K, *alors f est* K*-lipschitzienne (par contre si f' n'est pas bornée, alors la fonction ne peut pas être lipschitzienne). Par exemple, les fonctions* sin *et* cos *sont* 1*-lipschitziennes.*

☞**Exemples** :

- La fonction $x \mapsto x^2$ est uniformément continue sur tout segment $[a;b]$ (car lipschitzienne). Pour la même raison, les fonctions sin et cos sont uniformément continues sur $\mathbb{R}$.
- La fonction $x \mapsto \sqrt{x}$ est uniformément continue sur $[0;+\infty[$. Car pour $x, y \geqslant 0$, $|\sqrt{x} - \sqrt{y}| \leqslant \sqrt{|x-y|}$, il suffit donc de prendre $\alpha = \varepsilon^2$ dans la définition. Cependant nous verrons que cette même fonction n'est pas lipschitzienne sur $[0;+\infty[$ (dérivée non bornée).

– La fonction $x \mapsto x^2$ n'est pas uniformément continue sur $\mathbb{R}$. Si c'était le cas : avec $\varepsilon = 1$ il existe $\alpha > 0$ tel que $\forall\, x, y \in \mathbb{R}, |x-y| < \alpha \implies |x^2 - y^2| < 1$. Prenons $x_n = \sqrt{n+1}$ et $y_n = \sqrt{n}$ alors $x_n - y_n \sim \frac{1}{2\sqrt{n}} \to 0$, donc pour n assez grand on aura $|x_n - y_n| < \alpha$ d'où $|x_n^2 - y_n^2| < 1$ *i.e.* $1 < 1$ ce qui est absurde.

★**Exercice 11.6** *Étudier l'uniforme continuité des fonctions* $x \mapsto \cos(\sqrt{x})$ *et* $x \mapsto \cos(x^2)$ *sur* $[0\,;+\infty[$.

Théorème 11.16 (de *Heine*[1])
Toute fonction continue sur un segment est uniformément continue.

Preuve : Par l'absurde : on suppose qu'il existe $\varepsilon > 0$ tel que :

$$\forall\, \alpha > 0, \exists\, x, y \in \mathrm{I}, |x-y| < \alpha \text{ et } |f(x) - f(y)| \geqslant \varepsilon.$$

En prenant $\alpha = \frac{1}{n+1}$ pour chaque valeur de $n \in \mathbb{N}$, on construit deux suites (x_n) et (y_n) de I telles que $|x_n - y_n| < \frac{1}{n+1}$ et $|f(x_n) - f(y_n)| \geqslant \varepsilon$. La suite (x_n) étant bornée (car I est un segment), on peut en extraire une suite convergente : $x_{\sigma(n)} \to \ell$. Par passage à la limite on $\ell \in \mathrm{I}$. L'inégalité $|x_n - y_n| < \frac{1}{n+1}$ pour tout n entraîne que $y_{\sigma(n)} \to \ell$. La fonction f étant continue en ℓ, on a $f(x_{\sigma(n)}) \to f(\ell)$ et $f(y_{\sigma(n)}) \to f(\ell)$, donc $|f(x_{\sigma(n)}) - f(y_{\sigma(n)})| \to 0$, ce qui donne par passage à la limite, $0 \geqslant \varepsilon$ ce qui est absurde. Donc f est uniformément continue sur I. □

V CONTINUITÉ ET FONCTIONS MONOTONES

1) Image d'un intervalle

Théorème 11.17
Si f est strictement croissante et continue sur l'intervalle I, *alors :*
- *lorsque* $\mathrm{I} = [a;b]$, *on a* $f(\mathrm{I}) = [f(a); f(b)]$,
- *lorsque* $\mathrm{I} = [a;b[$, *on a* $f(\mathrm{I}) = [f(a); \lim\limits_{b} f[$,
- *lorsque* $\mathrm{I} =]a;b]$, *on a* $f(\mathrm{I}) =]\lim\limits_{a} f; f(b)]$,
- *lorsque* $\mathrm{I} =]a;b[$, *on a* $f(\mathrm{I}) =]\lim\limits_{a} f; \lim\limits_{b} f[$.

Preuve : Montrons par exemple le cas où $\mathrm{I} = [a;b[$, on sait que $\mathrm{J} = f(\mathrm{I})$ est un intervalle car f est continue sur I. La borne de droite de J est la limite de f en b (finie ou $+\infty$, la limite non atteinte car la monotonie est stricte), la borne de gauche de J est $f(a)$ qui le minimum de f, donc $f(\mathrm{I}) = [f(a); \lim\limits_{b} f[$. Les autres cas se traitent de la même façon, on a évidemment un énoncé analogue lorsque f est strictement décroissante. □

Attention !
Lorsque la monotonie n'est pas stricte, il se peut que les limites aux bornes exclues soit atteintes.

2) Monotonie et continuité

Théorème 11.18
Si $f : \mathrm{I} \to \mathbb{R}$ *est monotone sur l'intervalle* I *et si* $f(\mathrm{I})$ *est un intervalle, alors f est nécessairement continue sur* I.

Preuve : Quitte à changer f en $-f$, on peut supposons f croissante sur I. Soit $a \in \mathrm{I}$ un élément de I qui n'est pas la borne inférieure de I. Si $x \in \mathrm{I}$ avec $x < a$, alors $f(x) \leqslant \lim\limits_{a^-} f \leqslant f(a)$, ce qui entraîne que $\lim\limits_{a^-} f \in f(\mathrm{I})$. D'autre part, si $x \geqslant a$, alors $f(x) \geqslant f(a)$. On en déduit que l'intervalle $]\lim\limits_{a^-} f, f(a)[$ est inclus dans $f(\mathrm{I})$ mais il ne contient aucun élément de $f(\mathrm{I})$, cet intervalle est donc vide, *i.e.* $\lim\limits_{a^-} f = f(a)$, ce qui prouve que f est continue à gauche en a. Le raisonnement est analogue pour montrer la continuité à droite en a (si a n'est pas la borne de droite de I). □

Remarque 11.9 – *Ce théorème énonce une réciproque du théorème des valeurs intermédiaires, mais elle n'est valable que pour les fonctions monotones.*

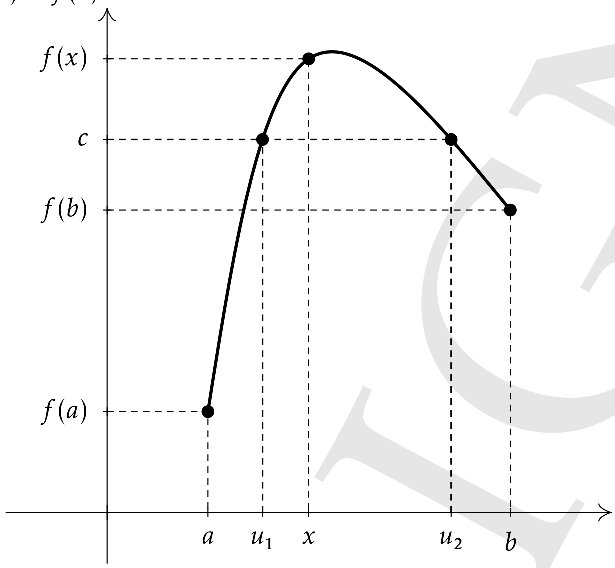
1. *HEINE Heinrich Eduard* (1821 – 1881) : mathématicien allemand qui a travaillé sur la théorie des fonctions.

Théorème 11.19

Si $f : \mathrm{I} \to \mathbb{R}$ est continue sur l'intervalle I *et injective, alors f est strictement monotone.*

Preuve : Soient $a < b$ deux éléments de I, f étant injective, $f(a) \neq f(b)$, quitte à changer f en $-f$, on peut supposer $f(a) < f(b)$, montrons alors que f est strictement croissante sur I :

- Étape 1 : soit $x \in]a;b[$, si $f(x) > f(b)$, alors un réel $c \in]f(b);f(x)[$ aura un antécédent dans $]x;b[$ (théorème des valeurs intermédiaires), et un antécédent dans $]a;x[$ car on a aussi $c \in]f(a);f(x)[$, ce qui contredit l'injectivité de f, donc $f(x) < f(b)$.



  De la même façon, on montre que $f(x) > f(a)$. En conclusion, si $x \in]a;b[$, alors $f(a) < f(x) < f(b)$.
- Étape 2 : soit $x \in \mathrm{I}$ avec $x < a$, si $f(x) > f(b)$ alors d'après l'étape 1 (appliquée à $-f$), on devrait avoir $f(x) > f(a) > f(b)$ ce qui est absurde, donc $f(x) < f(b)$, mais alors l'étape 1 (en échangeant a et x) nous dit que $f(x) < f(a) < f(b)$. En conclusion, si $x < a$ alors $f(x) < f(a)$.
- Étape 3 : soit $x \in \mathrm{I}$ avec $x > b$, comme ci-dessus, on montre que $f(x) > f(b)$.
- Étape 4 : soient $x < y$ deux éléments de I :
  - Si $x < y \leqslant a$: on sait que $f(x) < f(a)$, mais alors l'étape 1 entraîne que $f(x) < f(y)$.
  - Si $x \leqslant a < y$: on sait alors que $f(x) \leqslant f(a) < f(y)$, donc $f(x) < f(y)$.
  - Si $a < x < y$: alors on sait que $f(a) < f(y)$, mais alors l'étape 1 entraîne que $f(x) < f(y)$.

  Dans tous les cas, $f(x) < f(y)$, f est strictement croissante.

□

3) Théorème des bijections

Soit $f \colon \mathrm{I} \to \mathbb{R}$ une fonction strictement monotone, alors f est injective, donc f induit une bijection $\tilde{f}$ de I sur $f(\mathrm{I})$, la bijection réciproque est :

$$\tilde{f}^{-1} \colon f(\mathrm{I}) \to \mathrm{I}$$
$$x \mapsto y \text{ défini par } y \in \mathrm{I} \text{ et } f(y) = x$$

De plus, la bijection a le même sens de variation que f, en effet, supposons f croissante et soient $y < y'$ deux éléments de $f(\mathrm{I})$, alors il existe $x, x' \in \mathrm{I}$, tels que $f(x) = y$ et $f(x') = y'$; si on avait $x \geqslant x'$ alors on aurait $y \geqslant y'$ ce qui est contradictoire, donc $x < x'$ *i.e.* $\tilde{f}^{-1}(y) < \tilde{f}^{-1}(y')$.

D'autre part, dans un repère orthonormé du plan, on a :

$$\mathrm{M}(x,y) \in \mathrm{C}_{\tilde{f}^{-1}} \iff \begin{cases} x \in f(\mathrm{I}) \\ y = \tilde{f}^{-1}(x) \end{cases} \iff \begin{cases} y \in \mathrm{I} \\ f(y) = x \end{cases} \iff \mathrm{M}'(y,x) \in \mathrm{C}_f.$$

On en déduit que les courbes représentatives des fonctions f et $\tilde{f}^{-1}$ sont symétriques par rapport à la première bissectrice.

Le théorème suivant apporte une précision sur la continuité de la réciproque :

Théorème 11.20
Si $f\colon \mathrm{I} \to \mathbb{R}$ est strictement monotone sur l'intervalle I, alors f induit une bijection de I sur $\mathrm{J} = f(\mathrm{I})$. Si de plus f est continue sur I, alors la bijection réciproque est continue sur J.

Preuve : I étant un intervalle et f continue, l'ensemble $\mathrm{J} = f(\mathrm{I})$ est un intervalle, donc la bijection réciproque $\tilde{f}^{-1}$ est monotone et transforme l'intervalle J en l'intervalle I, d'après un des théorèmes précédents, $\tilde{f}^{-1}$ est continue sur J. □

VI EXTENSION AUX FONCTIONS À VALEURS COMPLEXES

1) Définition de la limite

Les fonctions à valeurs complexes ont été introduites au début du chapitre 6. Soit $f\colon \mathrm{I} \to \mathbb{C}$ une fonction, on note $u = \mathrm{Re}(f)$ (partie réelle de f) et $v = \mathrm{Im}(f)$ (partie imaginaire de f), on rappelle que u et v sont des fonctions de I vers $\mathbb{R}$, et $\forall\, t \in \mathrm{I}, f(t) = u(t) + iv(t)$.

La fonction **conjuguée** de f et la fonction $\overline{f}\colon t \mapsto u(t) - iv(t)$.

La fonction **module** de f est la fonction $|f|\colon t \mapsto |f(t)| = \sqrt{u(t)^2 + v(t)^2}$.

La fonction f est bornée sur I si et seulement si $\exists\, \mathrm{M} \in \mathbb{R}^+, \forall\, t \in \mathrm{I}, |f(t)| \leqslant \mathrm{M}$. Ceci équivaut à dire que les fonctions u et v sont bornées.

L'ensemble des fonctions de I vers $\mathbb{C}$ est notée $\mathcal{F}(\mathrm{I}, \mathbb{C})$, pour les opérations usuelles sur les fonctions, c'est un $\mathbb{C}$-espace vectoriel et un anneau commutatif non intègre.

Définition 11.8
Soit $\ell \in \mathbb{C}$, et soit a un élément de I ou une extrémité de I. On dira que la fonction f a pour limite ℓ en a lorsque $\lim\limits_{t\to a} |f(t) - \ell| = 0$. C'est à dire :
$$\forall \varepsilon > 0, \exists \mathrm{V}, \text{ voisinage de } a, \forall t \in \mathrm{I}, t \in \mathrm{V} \implies |f(t) - \ell| < \varepsilon.$$

★**Exercice 11.7** *Soit $f(t) = \frac{e^{it}}{i+t}$, déterminer la limite de f en $+\infty$.*

2) Propriétés

Théorème 11.21
$\lim\limits_{t\to a} |f(t) - \ell| = 0 \iff \lim\limits_{t\to a} \mathrm{Re}(f(t)) = \mathrm{Re}(\ell)$ ***et*** $\lim\limits_{t\to a} \mathrm{Im}(f(t)) = \mathrm{Im}(\ell)$.

Preuve : Celle - ci découle de l'inégalité : $\forall\, t \in \mathrm{I}$,
$$\max(|\mathrm{Re}(f(t)) - \mathrm{Re}(\ell)|, |\mathrm{Im}(f(t)) - \mathrm{Im}(\ell)|) \leqslant |f(t) - \ell| = \sqrt{|\mathrm{Re}(f(t)) - \mathrm{Re}(\ell)|^2 + |\mathrm{Im}(f(t)) - \mathrm{Im}(\ell)|^2}.$$ □

Connaissant les propriétés des limites (finies) des fonctions à valeurs réelles, on peut déduire celles des fonctions à valeurs complexes en raisonnant sur les parties réelles et imaginaires :

- $\lim\limits_a f = \ell \in \mathbb{C}$ si et seulement si pour toute suite (u_n) d'éléments de I qui tend vers a, la suite $(f(u_n))$ tend vers ℓ.
- Si $\lim\limits_a f = \ell \in \mathbb{C}$, alors f est bornée au voisinage de a.
- Si $\lim\limits_a f = \ell, \lim\limits_a g = \ell'$, alors $\lim\limits_a f + g = \ell + \ell', \lim\limits_a f \times g = \ell\ell', \forall\, \lambda \in \mathbb{C}, \lim\limits_a \lambda f = \lambda\ell$.
- Si $\lim\limits_a f = \ell$ alors $\lim\limits_a \overline{f} = \overline{\ell}$ et $\lim\limits_a |f| = |\ell|$.
- Si $\lim\limits_a f = \ell \in \mathbb{C}^*$, alors au voisinage de a f ne s'annule pas et $\lim\limits_a \frac{1}{f} = \frac{1}{\ell}$.

3) Continuité

Soit $f\colon \mathrm{I} \to \mathbb{C}$ une fonction à valeurs complexes, on pose $u = \mathrm{Re}(f)$ et $v = \mathrm{Im}(f)$.

Définition 11.9
On dira que f est continue sur I lorsque $\lim\limits_{t_0} f = f(t_0)$. L'ensemble des fonctions continues sur I est noté $\mathcal{C}^0(\mathrm{I}, \mathbb{C})$.

Remarque 11.10 – *D'après le chapitre sur les limites, on a vu que :*
$\lim_{t_0} f = f(t_0)$ *si et seulement si* $\lim_{t_0} u = u(t_0)$ *et* $\lim_{t_0} v = v(t_0)$*, par conséquent on peut dire que :*

$$\boxed{f \text{ est continue en } t_0 \text{ si et seulement si } \operatorname{Re}(f) \text{ et } \operatorname{Im}(f) \text{ sont continues en } t_0.}$$

☞**Exemple** : La fonction $t \mapsto e^{it}$ est continue sur $\mathbb{R}$, la fonction $t \mapsto \frac{1}{1+it}$ aussi.

4) Propriétés

Compte tenu de la définition, on retrouve des propriétés analogues au cas réel, à une exception près.

- On retrouve les **mêmes théorèmes généraux**, en particulier $\mathcal{C}^0(\mathrm{I}, \mathbb{C})$ est une $\mathbb{C}$-algèbre.
- Si f est continue sur I, alors les fonctions $\overline{f}$ et $|f|$ aussi.
- Si $f: [a;b] \to \mathbb{C}$ est continue sur le segment $[a;b]$, alors f est bornée et atteint ses bornes en module, c'est à dire, il existe $t_0, t_1 \in [a;b]$ tels que :
$$|f(t_0)| = \sup_{t\in[a;b]} |f(t)| \text{ et } |f(t_1)| = \inf_{t\in[a;b]} |f(t)|.$$
En effet : la fonction $|f|$ est continue sur $[a;b]$ et à valeurs réelles, on sait donc qu'elle admet un minimum et un maximum.
- Si $f: [a;b] \to \mathbb{C}$, est continue sur le segment $[a;b]$, alors f est uniformément continue (théorème de Heine).
En effet : cela découle de l'égalité : $|f(t)-f(t_0)| = \sqrt{|u(t)-u(t_0)|^2 + |v(t)-v(t_0)|^2}$, et du théorème de Heine pour les fonctions à valeurs réelles.

Attention ! (Le théorème des valeurs intermédiaires n'est plus vrai)
Par exemple, la fonction $f(t) = e^{it}$ *est continue sur* $[0; 2\pi]$, 0 *est compris entre* $f(\pi) = -1$ *et* $f(0) = 1$*, mais* $0 \notin f([0;2\pi])$ *car* f *ne s'annule pas (ici* $f[0;2\pi]$*) n'est pas un intervalle, mais un cercle !).*

VII SOLUTION DES EXERCICES

Solution 11.1 *On pose* $\mathrm{X} = \sin(x)\ln(x)$*, on a* $\mathrm{X} = \frac{\sin(x)}{x} x\ln(x)$*, donc* $\lim_{0^+} \mathrm{X} = 0$*, or* $\lim_{\mathrm{X}\to 0} \frac{e^{\mathrm{X}}-1}{\mathrm{X}} = 1$*, la limite cherchée vaut donc 1.*

Solution 11.2

1/ *Si a est non entier et* $n = \lfloor a \rfloor$*, alors* $n < a < n+1$*, et pour* $x \in]n\,;n+1[$ *on a Ex est constamment égal à n, et donc* $\lim_{x\to a} \lfloor x \rfloor = n = \lfloor a \rfloor$*.*
Si $a = n \in \mathbb{Z}$*, alors pour* $x \in]n-1\,;n[$ *on a* $\lfloor x \rfloor = n-1$ *et donc* $\lim_{x\to a^-} \lfloor x \rfloor = n-1 \neq \lfloor a \rfloor$*, et pour* $x \in]n\,;n+1[$ *on a* $\lfloor x \rfloor = n$ *et donc* $\lim_{x\to a^-} \lfloor x \rfloor = n = \lfloor a \rfloor$*. La fonction est donc continue à droite en a, mais pas à gauche.*

2/ *Soit a un irrationnel, par densité de* $\mathbb{Q}$ *dans* $\mathbb{R}$*, on sait qu'il existe une suite de rationnels* (r_n) *qui converge vers a, or* $1_{\mathbb{Q}}(r_n) = 1 \to 1$ *alors que* $1_{\mathbb{Q}}(a) = 0$*, la fonction n'est donc pas continue en a.*
Soit a un rationnel, par densité de $\mathbb{R} \setminus \mathbb{Q}$ *dans* $\mathbb{R}$*, on sait qu'il existe une suite de rationnels* (i_n) *qui converge vers a, or* $1_{\mathbb{Q}}(i_n) = 0 \to 0$ *alors que* $1_{\mathbb{Q}}(a) = 1$*, la fonction n'est donc pas continue en a. Finalement, la fonction* $1_{\mathbb{Q}}$ *est discontinue en tout point de* $\mathbb{R}$*.*

Solution 11.3 *Les théorème généraux s'appliquent sur l'intervalle* $]-\infty\,;0[$ *ainsi que sur l'intervalle* $]0\,;\pi[$*. La limite à gauche en 0 de f vaut* $e^0 - \cos(0) = 0 = f(0)$ *(continuité de* $\exp$ *et* $\cos$ *en 0). Pour la limite à droite en 0, on sait que* $\ln(1+x) \underset{0}{\sim} x$ *et* $\sin(x) \underset{0}{\sim} x$*, d'où* $f(x) \underset{0}{\sim} x - \pi$ *et donc la limite à droite en 0 de f vaut* $-\pi$*, la fonction est donc continue à gauche mais pas à droite en 0.*

La limite à gauche en π *vaut* $-\ln(1+\pi)$ *car* $\frac{\sin(x)}{x-\pi} = \frac{\sin(x)-\sin(\pi)}{x-\pi} \xrightarrow[x\to\pi]{} \sin'(\pi) = \cos(\pi) = -1$*. Il y a donc un prolongement par continuité en* π *en posant* $f(\pi) = -\ln(1+\pi)$*.*

Solution 11.4 *Soit* $f(x) = a_0 + a_1 x + \cdots + a_{2n+1} x^{2n+1}$ *avec* $a_{2n+1} \neq 0$*. La fonction f est fonction sur* $\mathbb{R}$ *et on a les équivalents* $f(x) \underset{\pm\infty}{\sim} a_{2n+1} x^{2n+1}$*, si* $a_{2n+1} > 0$ *alors f tend vers* $+\infty$ *en* $+\infty$ *et vers* $-\infty$ *en* $-\infty$ *(c'est l'inverse si* $a_{2n+1} < 0$*). Donc, au voisinage de* $+\infty$ *f n'a pas le même signe qu'au voisinage de* $-\infty$*, et donc f s'annule au moins une fois.*

Solution 11.5 *Si f est continue sur un segment* $[a\,;b]$ *alors elle admet un minimum en un certain réel* x_1 *de* $[a\,;b]$*, donc pour tout x de* $[a\,;b]$*, on* $f(x) \geqslant f(x_1)$*, mais comme f est strictement positive, on a* $f(x_1) > 0$*.*

Solution 11.6

1/ *Pour* $x, y \in [0\,;+\infty[$, $|\cos(\sqrt{x}) - \cos(\sqrt{y})| \leqslant |\sqrt{x} - \sqrt{y}| \leqslant \sqrt{|x-y|}$. *Soit* $\varepsilon > 0$, *si on prend* $\alpha = \varepsilon^2$, *alors pour* $[x-y| < \alpha$ *on aura* $|\cos(\sqrt{x}) - \cos(\sqrt{y})| < \sqrt{\alpha} = \varepsilon$. *La fonction* $x \mapsto \cos(\sqrt{x})$ *est donc uniformément continue sur* $[0\,;+\infty[$.

2/ *Pour* $n \in \mathbb{N}$, *posons* $x_n = \sqrt{2n\pi}$ *et* $y_n = \sqrt{2n\pi - \frac{\pi}{2}}$, *alors* $|x_n - y_n| = \sqrt{2n\pi}\left(1 - \sqrt{1 - \frac{1}{4n}}\right) \sim \frac{\sqrt{2\pi}}{8\sqrt{n}} \to 0$. *D'autre part* $|\cos(x_n^2) - \cos(y_n^2)| = 1 > \frac{1}{2}$, *ce qui prouve (par l'absurde) que la fonction* $x \mapsto \cos(x^2)$ *n'est pas uniformément continue sur* $\mathbb{R}$.

Solution 11.7 *On a* $|f(t)| = \frac{1}{\sqrt{1+t^2}} \xrightarrow[t\to+\infty]{} 0$, *donc* $\lim\limits_{t\to+\infty} f(t) = 0$.

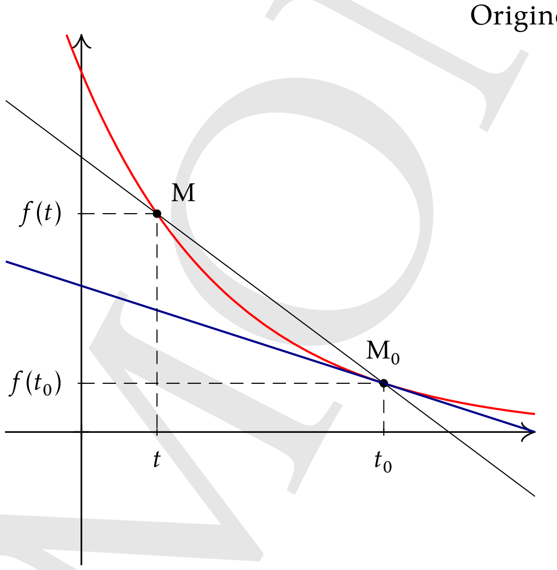
Chapitre 12

Dérivation

Sommaire

I DÉRIVÉE PREMIÈRE

1) Définition



Origine géométrique :

La droite qui joint les points $M(t, f(t))$ et $M_0(t_0, f(t_0))$ (sécante) a pour équation :

$$y = \frac{f(t) - f(t_0)}{t - t_0}(x - t_0) + f(t_0)$$

Lorsque l'on rapproche t de t_0, cette droite pivote autour du point M_0 et, lorsque la courbe est régulière, semble se rapprocher d'une position « limite » qui nous définirons comme la tangente au point M_0. Le coefficient directeur de cette droite « limite » doit être la limite lorsque t tend vers t_0 du coefficient directeur de la sécante, c'est dire $\lim\limits_{t \to t_0} \frac{f(t)-f(t_0)}{t-t_0}$.

Définition 12.1

Soit $f \colon \mathrm{I} \to \mathbb{R}$ une fonction et soit $t_0 \in \mathrm{I}$, on dit que f est ***dérivable en*** *t_0 lorsque la fonction : $t \mapsto \frac{f(t)-f(t_0)}{t-t_0}$ admet une limite* ***finie*** *en t_0. Si c'est le cas, cette limite est notée $f'(t_0)$ et appelée* ***nombre dérivé de*** *f* ***en*** *t_0. Lorsque f est dérivable en tout point de I on dit que f est dérivable sur I et la fonction de I vers $\mathbb{R}$ qui à t associe $f'(t)$ est appelée* ***dérivée de*** *f* ***sur*** *I, on la note f' ou bien $\frac{df}{dt}$. L'ensemble des fonctions dérivables sur I est noté $\mathrm{D}(\mathrm{I}, \mathbb{R})$. Si le plan est muni d'un repère orthonormé et si f est dérivable en t_0, la droite d'équation $y = f'(t_0)(x - t_0) + f(t_0)$ est appelée tangente à la courbe au point d'abscisse t_0. Si le taux d'accroissement de f en t_0 a une limite infinie et si f est continue en t_0, alors on dit que la courbe admet une tangente verticale au point d'abscisse t_0, d'équation $x = t_0$.*

Remarque 12.1 – *Les fonctions trigonométriques, logarithme, exponentielle, polynomiales et rationnelles sont dérivables sur leur ensemble de définition. Mais :*

Attention !

La fonction valeur absolue et la fonction $x \mapsto x^\alpha$ avec $0 < \alpha < 1$, ne sont pas dérivables en 0. La fonction partie entière n'est pas dérivable aux points entiers relatifs (pas continue).

Théorème 12.1 (définition équivalente)

f est dérivable en t_0 et $f'(t_0) = a$ si et seulement si $f(t) = f(t_0) + a(t - t_0) + (t - t_0)\underset{t_0}{o}(1)$.

On dit alors que f admet un développement limité d'ordre 1 en t_0.

Preuve : Celle-ci est simple et laissée en exercice. □

2) Théorème généraux

Théorème 12.2 (Dérivabilité et continuité)

Si f est dérivable en t_0, alors f est continue en t_0 mais la réciproque est fausse.

Preuve : Il suffit d'appliquer la définition équivalente ci-dessus pour voir que $\lim\limits_{t_0} f = f(t_0)$. Pour la réciproque, on a par exemple la fonction $t \mapsto |t|$ qui est continue en 0 mais non dérivable. □

Théorème 12.3 (Théorèmes généraux)

• Si f et g sont dérivables sur I et si $\alpha \in \mathbb{R}$ alors les fonctions $f+g$, $f \times g$ et αf sont dérivables sur I avec les formules :

– $(f+g)' = f' + g'$.

– $(f \times g)' = f' \times g + f \times g'$.

– $(\alpha f)' = \alpha f'$.

• Si f est dérivable sur I et ***ne s'annule pas*** *alors $\frac{1}{f}$ est dérivable sur et $\left(\frac{1}{f}\right)' = \frac{-f'}{f^2}$.*

• Si f est dérivable sur I et si g est dérivable sur J avec $\mathrm{Im}(f) \subset \mathrm{J}$, alors $g \circ f$ est dérivable sur I et $(g \circ f)' = f' \times [g' \circ f]$.

Preuve : Les deux premiers points ne posent pas de difficultés, passons au troisième : soit $x_0 = f(t_0)$, posons :

$$h(x) = \begin{cases} \frac{g(x)-g(x_0)}{x-x_0} & \text{si } x \neq x_0 \\ g'(x_0) & \text{si } x = x_0 \end{cases}$$

alors h est continue en x_0 et pour $t \neq t_0$ on a $\frac{g(f(t))-g(f(t_0))}{t-t_0} = h[f(t)] \times \frac{f(t)-f(t_0)}{t-t_0}$, même si $f(t) = f(t_0)$, comme f est continue en t_0, on a $\lim\limits_{t \to t_0} \frac{g(f(t))-g(f(t_0))}{t-t_0} = h(x_0) \times f'(t_0) = f'(t_0) \times g'(f(t_0))$. □

Du troisième point découlent les formules de dérivation usuelles :

Fonction	Dérivée
$\sin(u)$	$u'\cos(u)$
$\cos(u)$	$-u'\sin(u)$
$\tan(u)$	$u'(1+\tan(u)^2) = \frac{u'}{\cos(u)^2}$
$\mathrm{sh}(u)$	$u'\,\mathrm{ch}(u)$
$\mathrm{ch}(u)$	$u'\,\mathrm{sh}(u)$
$\mathrm{th}(u)$	$u'(1-\mathrm{th}(u)^2) = \frac{u'}{\mathrm{ch}(u)^2}$
e^u	$u'e^u$
$\ln(\lvert u\rvert)$	$\frac{u'}{u}$
u^α	$\alpha u' u^{\alpha-1}$

Remarque 12.2 – *Il découle des théorèmes généraux que pour les opérations usuelles sur les fonctions* $\mathrm{D}(\mathrm{I},\mathbb{R})$ *est un anneau et un* $\mathbb{R}$*-espace vectoriel.*

★**Exercice 12.1** *Soit* $f(x) = \begin{cases} x^2\sin(\frac{1}{x}) & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}$.

1/ *Montrer que f est dérivable sur $\mathbb{R}$.*

2/ *Montrer que f' n'est pas continue en* 0.

3) Dérivabilité à gauche et à droite

Définition 12.2

Soit $f\colon \mathrm{I}\to\mathbb{R}$ une fonction, et soit $t_0 \in \mathrm{I}$:

• *Si $t_0 \neq \inf(\mathrm{I})$: on dit que f est dérivable à gauche en t_0 lorsque le taux d'accroissement de f a une limite finie à gauche en t_0. Si c'est le cas, cette limite est notée $f_g'(t_0)$ et la demi-droite d'équation* $\begin{cases} y = f_g'(t_0)(x-t_0)+f(t_0) \\ x \leqslant t_0 \end{cases}$*, est appelée demi-tangente à la courbe au point d'abscisse t_0.*

• *Si $t_0 \neq \sup(\mathrm{I})$: on dit que f est dérivable à droite en t_0 lorsque le taux d'accroissement de f a une limite finie à droite en t_0. Si c'est le cas, cette limite est notée $f_d'(t_0)$ et la demi-droite d'équation* $\begin{cases} y = f_d'(t_0)(x-t_0)+f(t_0) \\ x \geqslant t_0 \end{cases}$*, est appelée demi-tangente à la courbe au point d'abscisse t_0.*

☞**Exemples** :

- La fonction valeur absolue est dérivable à gauche en 0, et $f_g'(0) = -1$, elle est dérivable à droite en 0 et $f_d'(0) = 1$, mais elle n'est pas dérivable en 0 car $-1 \neq 1$, on dit que le point de la courbe d'abscisse 0 est **un point anguleux**.
- La fonction $f(t) = \sqrt{|t|}$ n'est pas dérivable en 0, le taux d'accroissement tend vers $+\infty$ en 0^+ et vers $-\infty$ en 0^-, on dit que le point de la courbe d'abscisse 0 est un point **de rebroussement de première espèce**.

Théorème 12.4

Soit t_0 un point intérieur à I, f est dérivable en t_0 ssi f est dérivable à gauche et à droite en t_0 avec $f_g'(t_0) = f_d'(t_0)$.

Preuve : Cela découle des propriétés des limites. □

4) Dérivée d'une bijection réciproque

Théorème 12.5

Si $f\colon \mathrm{I}\to\mathbb{R}$ est une fonction continue strictement monotone, alors f induit une bijection de I sur $\mathrm{J} = \mathrm{Im}(f)$. Soit $y_0 = f(t_0) \in \mathrm{J}$ ($t_0 \in \mathrm{I}$), si f est dérivable en t_0 et si $f'(t_0) \neq 0$, alors la bijection réciproque, ϕ, est dérivable en y_0 et $\phi'(y_0) = \frac{1}{f'(t_0)} = \frac{1}{f'\circ\phi(y_0)}$. Si f est dérivable en t_0 et $f'(t_0) = 0$, alors ϕ n'est pas dérivable en y_0 mais la courbe représentative de ϕ admet une tangente verticale *au point d'abscisse y_0.*

Preuve : Soit $t_0 \in \mathrm{I}$ et $y_0 = f(t_0)$, pour $y \in \mathrm{J} \setminus \{y_0\}$, on a $\frac{\phi(y)-\phi(y_0)}{y-y_0} = \frac{t-t_0}{f(t)-f(t_0)}$ en posant $t = \phi(y)$, ϕ étant continue, lorsque $y \to y_0$, on a $t \to t_0$ et donc $\frac{t-t_0}{f(t)-f(t_0)} \to \frac{1}{f'(t_0)}$ car $f'(t_0) \neq 0$. Ce qui prouve le premier résultat.

Si $f'(t_0) = 0$, comme f est monotone la fraction $\frac{t-t_0}{f(t)-f(t_0)}$ garde un signe constant, donc sa limite lorsque $y \to y_0$ est infinie, ce qui prouve le second résultat. □

Remarque 12.3 –

- *Si $f\colon \mathrm{I} \to \mathrm{J}$ est bijective, continue, dérivable et si f' ne s'annule pas sur I, alors d'après le théorème précédent, f^{-1} est dérivable sur J et on a la formule :*

$$\left(f^{-1}\right)' = \frac{1}{f' \circ f^{-1}}.$$

- *Si f n'est pas dérivable en t_0 mais si sa courbe a une tangente verticale en ce point, alors f^{-1} est dérivable en $y_0 = f(t_0)$ et $\left(f^{-1}\right)'(y_0) = 0$ (car le taux d'accroissement de f en t_0 a une limite infinie en t_0).*

☞**Exemples** :

- La fonction $\ln\colon]0\,;+\infty[\to \mathbb{R}$ est une fonction continue, strictement croissante, dérivable et sa dérivée ne s'annule pas. Sa bijection réciproque, la fonction exponentielle, est donc dérivable sur $\mathbb{R}$ et :

$$\forall\, x \in \mathbb{R}, \exp(x)' = \frac{1}{\ln' \circ \exp(x)} = \exp(x).$$

- La fonction $f\colon [-\frac{\pi}{2}\,;\frac{\pi}{2}] \to [-1\,;1]$ définie par $f(x) = \sin(x)$ est bijective, continue, dérivable et sa dérivée ($f'(x) = \cos(x)$) ne s'annule pas sur $]-\frac{\pi}{2}\,;\frac{\pi}{2}[$, donc la bijection réciproque arcsin, est dérivable sur $]-1\,;1[$ et :

$$\arcsin'(x) = \frac{1}{f'(\arcsin(x))} = \frac{1}{\cos(\arcsin(x))} = \frac{1}{\sqrt{1-x^2}}.$$

Par contre la fonction arcsin n'est pas dérivable en ± 1 (une tangente verticale en ces points).

- La fonction $f\colon [0\,;\pi] \to [-1\,;1]$ définie par $f(x) = \cos(x)$ est bijective, continue, dérivable et sa dérivée ($f'(x) = -\sin(x)$) ne s'annule pas sur $]0\,;\pi[$, donc la bijection réciproque arccos, est dérivable sur $]-1\,;1[$ et :

$$\arccos'(x) = \frac{1}{f'(\arccos(x))} = \frac{-1}{\sin(\arccos(x))} = \frac{-1}{\sqrt{1-x^2}}.$$

Par contre la fonction arccos n'est pas dérivable en ± 1 (une tangente verticale en ces points).

- La fonction $f\colon]-\pi/2\,;\pi/2[\to \mathbb{R}$ définie par $f(x) = \tan(x)$ est bijective, continue, dérivable et sa dérivée ($f'(x) = 1 + \tan(x)^2$) ne s'annule pas, donc la bijection réciproque arctan, est dérivable sur $\mathbb{R}$ et :

$$\arctan'(x) = \frac{1}{f'(\arctan(x))} = \frac{1}{1+\tan^2(\arctan(x))} = \frac{1}{1+x^2}.$$

II APPLICATIONS DE LA DÉRIVATION

1) Théorème de Rolle

Théorème 12.6

Soit $f\colon [a;b] \to \mathbb{R}$ dérivable sur $]a;b[$ et soit $t_0 \in]a;b[$. Si f admet un extremum local en t_0, alors $f'(t_0) = 0$, mais la réciproque est fausse.

Preuve : Supposons que f présente un maximum local en t_0, alors à gauche en t_0 on a $\frac{f(t)-f(t_0)}{t-t_0} \geqslant 0$, d'où par passage à la limite en t_0 : $f'(t_0) \geqslant 0$. À droite en t_0 on a : $\frac{f(t)-f(t_0)}{t-t_0} \leqslant 0$, d'où par passage à la limite en t_0 : $f'(t_0) \leqslant 0$, par conséquent $f'(t_0) = 0$. Pour la réciproque il suffit de considérer la fonction $x \mapsto x^3$ en 0. □

Remarque 12.4 – *Dans le théorème ci-dessus, il est essentiel que t_0 ne soit pas une borne de l'intervalle. Par exemple la fonction $f(t) = 1 + t$ admet un maximum sur $[0;1]$ en $t_0 = 1$ mais $f'(t_0) \neq 0$.*

Théorème 12.7 (de Rolle [1])
Si $f\colon [a;b] \to \mathbb{R}$ est continue sur $[a;b]$, dérivable sur $]a;b[$ et si $f(a) = f(b)$, alors :
$$\text{il existe } c \in]a;b[,\ f'(c) = 0.$$

Preuve : Si f est constante alors il n'y a rien à montrer. Si f n'est pas constante, $\mathrm{Im}(f) = [m;\mathrm{M}]$ (f est continue sur le segment $[a;b]$) avec $m < \mathrm{M}$. Supposons $f(a) \neq \mathrm{M}$, alors $f(b) \neq \mathrm{M}$ or il existe $c \in [a;b]$ tel que $f(c) = \mathrm{M}$ donc $c \in]a;b[$, d'après la proposition précédente (maximum global en c) on a $f'(c) = 0$. Si $f(a) = \mathrm{M}$ alors $f(a) \neq m$ et le même raisonnement s'applique avec le minimum. □

Remarque 12.5 –
- *Ce théorème est faux si f n'est pas continue en a ou en b (prendre $f(x) = x$ sur $[0;1[$ et $f(1) = 0$).*
- *Ce théorème est faux si f est à valeurs complexes, par exemple $f(t) = e^{it}$, on a $f(0) = f(2\pi)$ mais $f'(t) = ie^{it}$ ne s'annule jamais.*

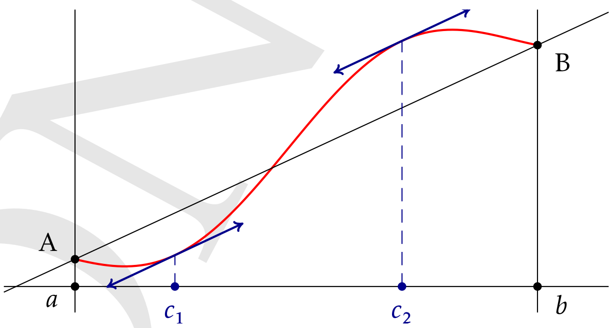
★**Exercice 12.2** *Soit $f\colon \mathbb{R} \to \mathbb{R}$ une fonction dérivable qui admet n racines distinctes, alors f' admet au moins $n-1$ racines distinctes.*

2) Les accroissements finis

Théorème 12.8 (égalité de accroissements finis)
Si $f\colon [a;b] \to \mathbb{R}$ est continue sur $[a;b]$ et dérivable sur $]a;b[$ alors :
$$\exists c \in]a;b[,\ f(b) - f(a) = (b-a)f'(c).$$

Preuve : Soit $\phi(t) = t(f(b) - f(a)) - (b-a)f(t)$, la fonction ϕ est continue sur $[a;b]$ et dérivable sur $]a;b[$, de plus $\phi(a) = af(b) - bf(a) = \phi(b)$, d'après le théorème de Rolle, il existe $c \in]a;b[$ tel que $\phi'(c) = 0$, ce qui donne la relation. □

Remarque 12.6 –
- *De même, si f et g sont continues sur $[a;b]$ et dérivables sur $]a;b[$, il existe $c \in]a;b[$ tel que :*
$$(f(b) - f(a))g'(c) = (g(b) - g(a))f'(c).$$
- *L'égalité s'écrit aussi : $f'(c) = \frac{f(b)-f(a)}{b-a}$, ce qui signifie géométriquement qu'il existe un point de la courbe (d'abscisse c) où la tangente est parallèle à la corde définie par le point d'abscisse a et le point d'abscisse b.*
- *Autre preuve : soit g la fonction affine prenant la même valeur que f en a et b, $g(x) = \frac{f(b)-f(a)}{b-a}(x-a) + f(a)$. On a $f(a) - g(a) = f(b) - g(b)$, d'après le théorème de Rolle il existe $c \in]a;b[$ tel que $f'(c) = g'(c)$ ce qui donne $f'(c) = \frac{f(b)-f(a)}{b-a}$.*



Théorème 12.9 (inégalité des accroissements finis)
Si $f\colon [a;b] \to \mathbb{R}$ est continue sur $[a;b]$, dérivable sur $]a;b[$ et s'il existe deux réels m et M tels que $\forall x \in]a;b[, m \leqslant f'(x) \leqslant \mathrm{M}$, alors :
$$m(b-a) \leqslant f(b) - f(a) \leqslant \mathrm{M}(b-a).$$

Preuve : Celle-ci découle directement de l'égalité des accroissement finis. □

1. *ROLLE Michel* (1652 – 1719) : mathématicien français.

À retenir

Si $\forall\, t \in]a;b[, |f'(t)| \leqslant \mathrm{M}$, alors $|f(b) - f(a)| \leqslant \mathrm{M}(b-a)$, et plus généralement :
$\forall\, x, y \in [a;b], |f(x) - f(y)| \leqslant \mathrm{M}|x-y|$, la fonction f est M-lipschitzienne.
Réciproquement, si f est M-lipschitzienne sur un intervalle I, alors tous les taux d'accroissements sont majorés en valeur absolue par M, et donc par passage à la limite, $|f'| \leqslant \mathrm{M}$.

★**Exercice** 12.3 *Soit $f\colon [a;b] \to [a;b]$ continue sur $[a;b]$, dérivable sur $]a;b[$ telle que $|f'| \leqslant k < 1$, on considère la suite définie par $u_0 \in [a;b]$ et $u_{n+1} = f(u_n)$.*

1/ Montrer que f admet au moins un point fixe ℓ et que celui-ci est unique.

2/ Montrer la suite u converge vers ℓ.

☞**Exemples** :
- Pour tout réel x, $\sin'(x) = \cos(x)$ et donc $|\sin'(x)| \leqslant 1$, on en déduit (IAF) que pour tous réels x et y on a $|\sin(x) - \sin(y)| \leqslant |x-y|$. De la même façon, on montre que $|\cos(x) - \cos(y)| \leqslant |x-y|$.
- Pour tout x, y de $[1\,;+\infty[$, on a $|\sqrt{x} - \sqrt{y}| \leqslant \frac{1}{2}|x-y|$.
- $\forall x > 0$, $\frac{1}{x+1} \leqslant \ln(x+1) - \ln(x) \leqslant \frac{1}{x}$.

Théorème 12.10 (limite de la dérivée)

Soit $f\colon [a;b] \to \mathbb{R}$ continue sur $[a;b]$ et dérivable sur $[a;b[$. Si f' admet une limite ℓ en b, alors :
• Si $\ell \in \mathbb{R}$ alors f est dérivable en b et $f'(b) = \ell$.
• Si $\ell = \pm\infty$ alors f n'est pas dérivable en b, mais il y a une tangente verticale pour la courbe représentative.

Preuve : D'après l'égalité des accroissements finis, pour $t \in [a;b[$, il existe $c_t \in]t;b[$ tel que $f(b) - f(t) = (b-t)f'(c_t) = (b-t)f'(c_t)$, d'où $\frac{f(t)-f(b)}{t-b} = f'(c_t)$, mais si t tend vers b, alors c_t tend vers b et donc $f'(c_t)$ tend vers ℓ, d'où : $\lim\limits_{t\to b} \frac{f(t)-f(b)}{t-b} = \ell$, ce qui termine la preuve. □

Remarque 12.7 – *Si f' n'a pas de limite en b, on ne peut rien dire en général.*
On a un résultat analogue pour $f\colon [a;b] \to \mathbb{R}$ continue sur $[a;b]$, dérivable sur $]a;b]$, avec $\lim\limits_{t\to a} f'(t) = \ell$.

☞**Exemple** : La fonction arcsin est dérivable sur $]-1;1[$ et $\arcsin'(x) = \frac{1}{\sqrt{1-x^2}}$, cette dérivée a pour limite $+\infty$ quand $x \to 1$. On retrouve ainsi que arcsin n'est pas dérivable en 1 et qu'il y a une tangente verticale en ce point pour la courbe.

3) Sens de variation

Théorème 12.11

Soit $f\colon \mathrm{I} \to \mathbb{R}$ une fonction continue sur l'intervalle I, et dérivable sur I privé des ses bornes (noté $\mathring{\mathrm{I}}$, intérieur de I), on a les résultats suivants :
• f est croissante si et seulement si $\forall\, t \in \mathring{\mathrm{I}}, f'(t) \geqslant 0$.
• f est décroissante si et seulement si $\forall\, t \in \mathring{\mathrm{I}}, f'(t) \leqslant 0$.
• f est constante si et seulement si $\forall\, t \in \mathring{\mathrm{I}}, f'(t) = 0$.
• f est strictement croissante si et seulement si $\forall\, t \in \mathring{\mathrm{I}}, f'(t) \geqslant 0$ et il n'existe aucun intervalle ouvert non vide inclus dans I sur lequel f' est constamment nulle.
• f est strictement décroissante si et seulement si $\forall\, t \in \mathring{\mathrm{I}}, f'(t) \leqslant 0$ et il n'existe aucun intervalle ouvert non vide inclus dans I sur lequel f' est constamment nulle.

Preuve : Si f est croissante sur I, soit $t_0 \in \mathring{\mathrm{I}}$, le taux d'accroissement de f en t_0 est toujours positif, donc par passage à la limite, on a $f'(t_0) \geqslant 0$. Réciproquement, si $f' \geqslant 0$ sur $\mathring{\mathrm{I}}$, soit $t < t'$ deux éléments de I, d'après l'égalité des accroissements finis, il existe c compris entre t et t' (strictement) tel que $f(t) - f(t') = f'(c)(t-t') \leqslant 0$, donc $f(t) \leqslant f(t')$ *i.e.* f est croissante. Pour f décroissante on applique ce qui précède à $-f$. Pour f constante, il suffit de dire que f est à la fois croissante et décroissante.

Si f est strictement croissante, alors on sait que $f' \geqslant 0$ sur $\mathring{\mathrm{I}}$. Si f' est nulle sur un intervalle $\mathrm{J} \subset \mathrm{I}$, alors f est constante sur J, ce qui est absurde. Réciproquement, si $\forall\, t \in \mathring{\mathrm{I}}, f'(t) \geqslant 0$ et il n'existe aucun intervalle ouvert non vide inclus dans I sur lequel f' est constamment nulle, soit $t < t'$ deux éléments de I, on sait que $f(t) \leqslant f(t')$, si on avait $f(t) = f(t')$ alors $\forall\, c \in [t;t'], f(t) = f(c) = f(t')$, donc f est constante sur $[t;t']$, ce qui entraîne que f' est nulle sur $]t;t'[$: absurde, donc $f(t) < f(t')$ *i.e.* f est strictement croissante. □

Remarque 12.8 – *Ce théorème est faux si* I *n'est pas intervalle, par exemple la fonction* $f(t) = \frac{1}{t}$ *est dérivable sur* $\mathbb{R}^*$ *avec* $f' < 0$*, mais* f *n'est pas monotone sur* $\mathbb{R}^*$.

III DÉRIVÉES SUCCESSIVES

1) Classe d'une application

Définition 12.3

Soit $f\colon \mathrm{I} \to \mathbb{R}$ *une fonction et soit* $n \in \mathbb{N}^*$. *On dit que* f *est de classe* $\mathcal{C}^n$ *sur* I *lorsque* f *est* n *fois dérivable sur* I *et que la dérivée* n^e *de* f *est continue sur* I. *L'ensemble des fonctions de classe* $\mathcal{C}^n$ *sur* I *est noté* $\mathcal{C}^n(\mathrm{I},\mathbb{R})$. *La dérivée* n^e *de* f *est notée* $f^{(n)}$ *où* $\frac{d^n f}{dt^n}$. *Par convention, on pose* $f^{(0)} = f$, *on a alors* $\forall\, n \in \mathbb{N}, f^{(n+1)} = \left(f^{(n)}\right)'$.

Remarque 12.9 –
- $\mathcal{C}^{n+1}(\mathrm{I},\mathbb{R}) \subset \mathcal{C}^n(\mathrm{I},\mathbb{R})$.
- *Si* $f \in \mathcal{C}^n(\mathrm{I},\mathbb{R})$ *avec* $n \geqslant 1$, *alors* $\forall\, k \in [\![0\,;n]\!], f^{(k)} \in \mathcal{C}^{n-k}(\mathrm{I},\mathbb{R})$.

☞**Exemples :**
- $\forall\, n \in \mathbb{N}, f\colon t \mapsto e^t$ est de classe $\mathcal{C}^n$ sur $\mathbb{R}$, et $f^{(n)}(t) = e^t$.
- Soit $n \in \mathbb{N}, \forall\, p \in \mathbb{N}, f\colon t \mapsto t^n$ est de classe $\mathcal{C}^p$ sur $\mathbb{R}$, et $f^{(p)}(t) = \begin{cases} 0 & \text{si } p > n \\ \frac{n!}{(n-p)!}t^{n-p} & \text{sinon} \end{cases}$.
- $\forall\, n \in \mathbb{N}, f\colon t \mapsto \frac{1}{t}$ est de classe $\mathcal{C}^n$ sur $\mathbb{R}^*$, et $f^{(n)}(t) = \frac{(-1)^n n!}{t^{n+1}}$.
- $\forall\, n \in \mathbb{N}, f\colon t \mapsto \ln(t)$ est de classe $\mathcal{C}^n$ sur $]0\,;+\infty[$, et pour $n \geqslant 1$, $f^{(n)}(t) = \frac{(-1)^{n-1}(n-1)!}{t^n}$.
- $\forall\, n \in \mathbb{N}$, cos et sin sont de classe $\mathcal{C}^n$ sur $\mathbb{R}$ et $\cos^{(n)}(t) = \cos(t + n\frac{\pi}{2})$, $\sin^{(n)}(t) = \sin(t + n\frac{\pi}{2})$.

★**Exercice 12.4**

1/ *Soit* $a \in \mathbb{R}$, *et* $f\colon x \mapsto \frac{1}{x-a}$, *montrer que* f *est de classe* $\mathcal{C}^n$ *sur* $\mathbb{R}\setminus\{a\}$ *pour tout* $n \in \mathbb{N}$, *et calculer* $f^{(n)}(x)$.

2/ *Soit* $f\colon x \mapsto \frac{1}{x^2-1}$, *montrer que* f *est de classe* $\mathcal{C}^n$ *sur* $\mathbb{R}\setminus\{\pm 1\}$ *pour tout* n, *et calculer* $f^{(n)}(x)$.

Définition 12.4

Lorsque f *est de classe* $\mathcal{C}^n$ *pour tout entier* n, *on dit que* f *est de classe* $\mathcal{C}^\infty$, *l'ensemble des ces fonctions est noté* $\mathcal{C}^\infty(\mathrm{I},\mathbb{R})$, *et on a donc* $\mathcal{C}^\infty(\mathrm{I},\mathbb{R}) = \bigcap\limits_{n\in\mathbb{N}} \mathcal{C}^n(\mathrm{I},\mathbb{R})$.

Remarque 12.10 –
- $\forall\, n \in \mathbb{N}, \mathcal{C}^\infty(\mathrm{I},\mathbb{R}) \subset \mathcal{C}^n(\mathrm{I},\mathbb{R})$.
- *Dire que* f *est* $\mathcal{C}^\infty$ *sur* I *revient à dire que* f *est dérivable autant de fois que l'on veut (infiniment dérivable), autrement dit* $\mathcal{C}^\infty(\mathrm{I},\mathbb{R}) = \bigcap\limits_{n\in\mathbb{N}} \mathcal{D}^n(\mathrm{I},\mathbb{R})$.

☞**Exemples :**
- Toute fonction polynomiale est $\mathcal{C}^\infty$ sur $\mathbb{R}$ (car la dérivée d'un polynôme est un polynôme).
- Toute fonction rationnelle est $\mathcal{C}^\infty$ sur son ensemble de définition (car la dérivée d'une fonction rationnelle est une fonction rationnelle).
- Les fonctions $\ln, \exp, \cos, \sin$ et $\tan$ sont $\mathcal{C}^\infty$ sur leur ensemble de définition.

★**Exercice 12.5** *Étudier la classe sur* $\mathbb{R}$ *de la fonction* $f\colon x \mapsto x^2|x|$.

Théorème 12.12 (prolongement de classe $\mathcal{C}^n$)

Soit $f\colon [a\,;b[\to \mathbb{R}$ *une fonction de classe* $\mathcal{C}^n$ *sur* $[a;b[$ *telle que toutes ses dérivées* k^e *ont une limite finie en* b : $\forall k \in [\![0\,;n]\!], \exists \ell_k \in \mathbb{R}, \lim\limits_{x\to b} f^{(k)}(x) = \ell_k$. *Alors le prolongement de* f *obtenu en* *posant $f(b) = \ell_0$, est un prolongement de classe $\mathcal{C}^n$ sur $[a\,;b]$, et on a $\forall k \in [\![0\,;n]\!]$, $f^{(k)}(b) = \ell_k$.*

Preuve : Par récurrence sur n. Pour $n = 0$, c'est un prolongement par continuité de f en b. Supposons le théorème établi au rang n et que f vérifie les hypothèses au rang $n+1$, en appliquant (HR), le prolongement de f obtenu en posant $f(b) = \ell_0$, est un prolongement de classe $\mathcal{C}^n$ sur $[a;b]$, et on a $\forall k \in [\![0\,;n]\!]$, $f^{(k)}(b) = \ell_k$. Soit $g = f^{(n)}$, alors g est continue sur $[a;b]$, de classe $\mathcal{C}^1$ sur $[a;b[$ et $\lim\limits_{x\to b} g'(x) = \ell_{n+1} \in \mathbb{R}$, on en déduit que g est dérivable en b (théorème sur la limite de la dérivée) et que $g'(b) = \ell_{n+1}$, ce qui entraîne que g' est continue en b. Finalement le prolongement de f est bien de classe $\mathcal{C}^{n+1}$ sur $[a\,;b]$, et $f^{(k)}(b) = \ell_k$ pour $k \in [\![0\,;n+1]\!]$. □

2) Formule de Leibniz

Théorème 12.13 (généraux)

Si f et g sont de classe $\mathcal{C}^n$ sur I alors :

- *$f + g$ est de classe $\mathcal{C}^n$ sur I et $(f+g)^{(n)} = f^{(n)} + g^{(n)}$.*
- *$\forall \lambda \in \mathbb{R}$, $\lambda.f$ est de classe $\mathcal{C}^n$ sur I et $(\lambda.f)^{(n)} = \lambda.f^{(n)}$.*
- *$f \times g$ est de classe $\mathcal{C}^n$ sur I et on a la formule (de Leibniz) : $(f\times g)^{(n)} = \sum\limits_{k=0}^{n} \binom{n}{k} f^{(k)} \times g^{(n-k)}$.*

Preuve : Pour le dernier point : pour $n = 0$ le résultat est vrai. Supposons le dernier point démontré au rang $n \geqslant 0$ avec la formule de Leibniz, et supposons que f et g sont de classe $\mathcal{C}^{n+1}$. En particulier f et g sont $\mathcal{C}^n$, donc $f\times g$ aussi et $(f\times g)^{(n)} = \sum\limits_{k=0}^{n} \binom{n}{k} f^{(k)} \times g^{(n-k)}$, on en déduit donc que $(f\times g)^{(n)}$ est dérivable sur I (somme de produits de fonctions dérivables) et sa dérivée est $(f\times g)^{(n+1)} = \sum\limits_{k=0}^{n} \binom{n}{k} f^{(k+1)} \times g^{(n-k)} + \sum\limits_{k=0}^{n} \binom{n}{k} f^{(k)} \times g^{(n+1-k)}$, ce qui donne $f^{(n+1)} \times g + f \times g^{(n+1)} + \sum\limits_{k=1}^{n} \left(\binom{n}{k} + \binom{n}{k-1}\right) f^{(k)} \times g^{(n+1-k)}$, c'est à dire $\sum\limits_{k=0}^{n+1} \binom{n+1}{k} f^{(k)} \times g^{(n+1-k)}$, ce qui donne la formule au rang $n+1$, de plus cette somme est une somme de fonctions continues, ce qui prouve que $f\times g$ est bien de classe $\mathcal{C}^{n+1}$ sur I. □

Théorème 12.14

$\forall\, n \in \mathbb{N} \cup \{+\infty\}$, $\mathcal{C}^n(\mathrm{I}, \mathbb{R})$ est un $\mathbb{R}$-espace vectoriel et un anneau.

Preuve : Cela découle du théorème précédent (s.e.v et sous-anneau de $\mathcal{F}(\mathrm{I}, \mathbb{R})$). □

★**Exercice 12.6** *Calculer de deux façons la dérivée n^e en 0 de la fonction $x \mapsto (1-x^2)^n$. Quelle relation obtient-on ?*

3) Classe d'une composée

Théorème 12.15

Soient $f\colon \mathrm{I} \to \mathbb{R}$ et $g : \mathrm{J} \to \mathbb{R}$ deux fonctions de classe $\mathcal{C}^n$ avec $\mathrm{Im}(f) \subset \mathrm{J}$, alors $g\circ f$ est de classe $\mathcal{C}^n$ sur I. En particulier, si f et g sont $\mathcal{C}^\infty$ alors $g \circ f$ aussi.

Preuve : Le théorème est vrai pour $n = 0$ (composée de deux fonctions continues), supposons le vrai au rang $n \geqslant 0$ et supposons f et g de classe $\mathcal{C}^{n+1}$, comme $n+1 \geqslant 1$, f et g sont dérivables, donc $g\circ f$ est dérivable avec la formule $(g\circ f)' = f' \times g' \circ f$, d'après l'hypothèse de récurrence, $g'\circ f$ est de classe $\mathcal{C}^n$ (car g' et f sont de classe $\mathcal{C}^n$), or f' est également de classe $\mathcal{C}^n$, par conséquent $f' \times g' \circ f$ est de classe $\mathcal{C}^n$, ce qui signifie que $g\circ f$ est de classe $\mathcal{C}^{n+1}$. □

Remarque 12.11 –

- *Il existe une formule qui exprime $(g\circ f)'$ en fonction des dérivées de f et de g, mais ce n'est pas une formule simple.*
- *La fonction inverse $g : x \mapsto \frac{1}{x}$ est $\mathcal{C}^\infty$ sur $\mathbb{R}^*$, si $f\colon \mathrm{I} \to \mathbb{R}$ est une fonction de classe $\mathcal{C}^n$ qui ne s'annule, alors la composée, i.e. la fonction $\frac{1}{f}$, est de classe $\mathcal{C}^n$ (même si $n = \infty$).*
- *On retrouve donc les mêmes théorèmes généraux que pour la continuité et la dérivabilité.*

4) Classe d'une réciproque

Théorème 12.16

Soit $f : \mathrm{I} \to \mathrm{J}$ une bijection de I sur $\mathrm{J} = \mathrm{Im}(f)$, de classe $\mathcal{C}^n$ avec $n \in \mathbb{N}^* \cup \{\infty\}$. Si f' ne s'annule pas sur I, alors la bijection réciproque f^{-1} est de classe $\mathcal{C}^n$ sur J (i.e. de même classe que f).

Preuve : On sait déjà que f^{-1} est dérivable sur J et que $(f^{-1})' = \frac{1}{f' \circ f^{-1}}$, on voit alors que f^{-1} est de classe $\mathcal{C}^1$ sur J, le théorème est donc vrai pour $n = 1$, supposons le vrai au rang $n \geqslant 1$ et supposons que f est $\mathcal{C}^{n+1}$, par hypothèse de récurrence f^{-1} est de classe $\mathcal{C}^n$, mais alors $f' \circ f^{-1}$ est une fonction de classe $\mathcal{C}^n$ qui ne s'annule pas, donc son inverse est de classe $\mathcal{C}^n$, *i.e.* $(f^{-1})'$ est $\mathcal{C}^n$, ce qui signifie que f^{-1} est de classe $\mathcal{C}^{n+1}$ sur J. □

☞**Exemples** :
- Les fonctions arcsin et arccos sont de classe $\mathcal{C}^\infty$ sur $]-1;1[$.
- La fonction arctan est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$.

IV EXTENSION AUX FONCTIONS À VALEURS COMPLEXES

1) Définition

On adopte la même définition que dans le cas réel :

Définition 12.5

On dira que $f : \mathrm{I} \to \mathbb{C}$ est dérivable en $t_0 \in \mathrm{I}$ si et seulement si la fonction $t \mapsto \frac{f(t)-f(t_0)}{t-t_0}$ définie sur $\mathrm{I} \setminus \{t_0\}$, admet une limite finie (dans $\mathbb{C}$) en t_0. Si celle-ci existe, elle est notée $f'(t_0)$. L'ensemble des fonctions dérivables sur I est noté $\mathcal{D}(\mathrm{I}, \mathbb{C})$.

2) Propriétés

Théorème 12.17 (caractérisation)

Soit $f : \mathrm{I} \to \mathbb{C}$ une fonction, soit $u = \mathrm{Re}(f)$ et $v = \mathrm{Im}(f)$, alors f est dérivable en $t_0 \in \mathrm{I}$ si et seulement si u et v sont dérivables en t_0. Si tel est le cas, alors $f'(t_0) = u'(t_0) + iv'(t_0)$.

Preuve : Il suffit d'écrire que :

$$\frac{f(t)-f(t_0)}{t-t_0} = \frac{u(t)-u(t_0)}{t-t_0} + i\frac{v(t)-v(t_0)}{t-t_0}$$

avec $u = \mathrm{Re}(f)$ et $v = \mathrm{Im}(f)$. □

À retenir

Il découle de ce théorème, que lorsque f est dérivable sur I, on a : $\mathrm{Re}(f') = \mathrm{Re}(f)'$ et $\mathrm{Im}(f') = \mathrm{Im}(f)'$.

Comme la caractérisation nous ramène aux fonctions à valeurs réelles, on peut déduire les propriétés des fonctions dérivables à valeurs complexes :
- On retrouve les mêmes théorèmes généraux, à savoir :
  - Toute fonction $f : \mathrm{I} \to \mathbb{C}$ dérivable est continue (réciproque fausse).
  - Si $f, g : \mathrm{I} \to \mathbb{C}$ sont dérivables, alors $f+g$, $f \times g$ et λf ($\lambda \in \mathbb{C}$) sont dérivables avec les formules : $(f+g)' = f'+g'$, $(f \times g)' = f' \times g + f \times g'$, $(\lambda f)' = \lambda f'$.
  - Si $g : \mathrm{I} \to \mathbb{C}$ est dérivable et ne s'annule pas, alors $\frac{1}{g}$ est dérivable sur I et $(\frac{1}{g})' = -\frac{g'}{g^2}$. On en déduit que si f est également dérivable sur I alors $\left(\frac{f}{g}\right)' = \frac{f' \times g - f \times g'}{g^2}$.
  - Si $f : \mathrm{I} \to \mathbb{R}$ et $g : \mathrm{J} \to \mathbb{C}$ sont dérivables avec $\mathrm{Im}(f) \subset \mathrm{J}$, alors $g \circ f$ est dérivable sur I et $(g \circ f)' = f' \times g' \circ f$.
  - Si $f : \mathrm{I} \to \mathbb{C}$ est dérivable alors $\exp(f)$ est dérivable sur I et $[\exp(f)]' = f' \times \exp(f)$.
- Cependant, le théorème de Rolle n'est plus valable, par exemple la fonction $f(t) = \exp(it)$ est dérivable sur $\mathbb{R}$ et $f'(t) = i\exp(it)$, on a $f(0) = f(2\pi)$ mais f' ne s'annule pas. Par conséquent l'égalité des accroissements finis n'est plus valable non plus, mais on conserve les inégalités.

Théorème 12.18 (inégalité des accroissements finis généralisée)
Si $f : \mathrm{I} \to \mathbb{C}$ est une fonction $\mathcal{C}^1$ sur I, et si $\forall\, t \in \mathrm{I}, |f'(t)| \leqslant g'(t)$ où $g : \mathrm{I} \to \mathbb{R}$ est une fonction $\mathcal{C}^1$ sur I, alors :

$$\forall a, b \in \mathrm{I}, |f(b) - f(a)| \leqslant |g(b) - g(a)|.$$

Remarque 12.12 –

– *Si $\forall\, t \in \mathrm{I}, |f'(t)| \leqslant \mathrm{M}$, alors en prenant la fonction $g(t) = \mathrm{M}t$, et en appliquant le théorème ci-dessus, on obtient $\forall a, b \in \mathrm{I}, |f(b) - f(a)| \leqslant \mathrm{M}|b - a|$.*

☞**Exemple** : Avec $f(t) = \exp(\alpha t)$ où $\alpha = a + ib \in \mathbb{C}$ avec $a \neq 0$, on a $|f'(t)| = |\alpha| \exp(at) = g'(t)$, par conséquent :

$$\forall\, t, t' \in \mathbb{R}, |\exp(\alpha t) - \exp(\alpha t')| \leqslant \frac{|\alpha|}{|a|} |\exp(at) - \exp(at')|.$$

3) Classe d'une fonction

On donne la même définition avec les mêmes notations que pour les fonctions à valeurs réelles, à savoir : $f : \mathrm{I} \to \mathbb{C}$ est de classe $\mathcal{C}^n$ ssi f est n fois dérivable et $f^{(n)}$ est continue sur I, ce qui revient à dire que les parties réelle et imaginaire de f sont de classe $\mathcal{C}^n$. L'ensemble des fonctions de classe $\mathcal{C}^n$ sur I est noté $\mathcal{C}^n(\mathrm{I}, \mathbb{C})$, et on pose $\mathcal{C}^\infty(\mathrm{I}, \mathbb{C}) = \bigcap\limits_{n \in \mathbb{N}} \mathcal{C}^n(\mathrm{I}, \mathbb{C})$: ensemble des fonctions de classe $\mathcal{C}^\infty$.

On retrouve les mêmes théorèmes généraux : $\mathcal{C}^n(\mathrm{I}, \mathbb{C})$ est une $\mathbb{C}$-algèbre ($n \in \mathbb{N} \cup \{\infty\}$). La formule de Leibniz reste valable, et la composée de deux fonctions de classe $\mathcal{C}^n$ est également de classe $\mathcal{C}^n$.

★**Exercice 12.7** *Soit $f(t) = \cos(t)\exp(t\sqrt{3})$, calculer $f^{(n)}(t)$.*

V SOLUTION DES EXERCICES

Solution 12.1

1/ *Les théorèmes généraux s'appliquent sur $\mathbb{R}^*$. Le taux d'accroissement en 0 s'écrit $x\sin(\frac{1}{x})$ qui tend vers 0 en 0, donc f est dérivable en 0 et $f'(0) = 0$.*

2/ *Pour $x \neq 0$, $f'(x) = 2x\sin(\frac{1}{x}) - \cos(\frac{1}{x})$, or si on pose $u_n = \frac{1}{n\pi}$, alors $u_n \to 0$ et $\cos(\frac{1}{u_n}) = (-1)^n$ n'a pas de limite. Donc la fonction $x \mapsto \cos(\frac{1}{x})$ n'a pas de limite en 0, et comme la fonction $x \mapsto 2x\sin(\frac{1}{x})$ tend vers 0 en 0, cela entraîne que $f'(x)$ ne peut pas avoir de limite en 0.*

Solution 12.2 *Il suffit d'appliquer le théorème de Rolle à la fonction f entre deux racines consécutives. On montre ainsi qu'entre deux racines de f il y a toujours une racine de f'.*

Solution 12.3

1/ *On montre que la fonction $g : x \mapsto f(x) - x$ s'annule en un point ℓ car elle est continue et change de signe puisque $g(a) = f(a) - a \geqslant 0$ et $g(b) = f(b) - b \leqslant 0$. Si ℓ et ℓ' sont deux points de f dans $[a\,;b]$, alors en appliquant l'inégalité des AF, on $|\ell - \ell'| = [f(\ell) - f(\ell')| \leqslant k|\ell - \ell'|$, or $k < 1$, ce qui entraîne $|\ell - \ell'| = 0$ et donc $\ell = \ell'$.*

2/ *Par récurrence tous les termes u_n existent dans $[a\,;b]$. On a alors (IAF) $|u_{n+1} - \ell| = |f(u_n) - f(\ell)| \leqslant k|u_n - \ell|$, on en déduit par récurrence que $|u_n - \ell| \leqslant k^n|u_0 - \ell|$, or $k^n \to 0$ car $|k| < 1$, et donc $u_n \to \ell$ (c'est le théorème du point fixe).*

Solution 12.4

Solution 12.5 *On vérifie que f est dérivable sur $\mathbb{R}$, avec $f'(x) = 3x^2$ si $x > 0$, $f'(x) = -3x^2$ si $x < 0$ et $f'(0) = 0$. On peut écrire $f'(x) = 3x|x|$ pour tout x, et donc f' est continue sur $\mathbb{R}$, f est donc au moins de classe $\mathcal{C}^1$ sur $\mathbb{R}$.*

De même, on vérifie que f' est dérivable sur $\mathbb{R}$, avec $f''(x) = 6x$ si $x > 0$, $f'(x) = -6x$ si $x < 0$ et $f''(0) = 0$. On peut écrire $f''(x) = 6|x|$ pour tout x, et donc f'' est continue sur $\mathbb{R}$, f' est donc au moins de classe $\mathcal{C}^1$ sur $\mathbb{R}$, c'est à dire f est au moins de classe $\mathcal{C}^2$ sur $\mathbb{R}$.

Par contre, on peut vérifier que f'' n'est pas dérivable en 0, et donc f n'est pas de classe $\mathcal{C}^3$ sur $\mathbb{R}$.

Solution 12.6 *On a $f(x) = \sum\limits_{k=0}^{n} \binom{n}{k}(-1)^k x^{2k}$ (Newton) et donc $f^{(n)}(x) = \sum\limits_{\frac{n}{2} \leqslant k \leqslant n} n\binom{n}{k}(-1)^k \frac{(2k)!}{(2k-n)!} x^{2k-n}$ et donc $f^{(n)}(0) = 0$ si n est impair, et $f^{(n)}(0) = (-1)^{\frac{n}{2}}\binom{n}{\frac{n}{2}} n!$.*

On a $f(x) = (1-x)^n \times (1+x)^n$, *et donc* $f^{(n)}(x) = \sum_{k=0}^{n} \binom{n}{k}(-1)^k \frac{n!}{(n-k)!}(1-x)^{n-k}\frac{n!}{k!}(1+x)^k$*(Leibniz), d'où* $f^{(n)}(x) = n!\sum_{k=0}^{n} \binom{n}{k}^2(-1)^k(1-x)^{n-k}(1+x)^k$ *et donc* $f^{(n)}(0) = n!\sum_{k=0}^{n} \binom{n}{k}^2(-1)^k$. *En égalant les deux résultats, on en déduit que :* $\sum_{k=0}^{n} \binom{n}{k}^2(-1)^k = (-1)^{\frac{n}{2}}\binom{n}{\frac{n}{2}}$ *si* n *est pair, et* 0 *sinon.*

Solution 12.7 *On a* $f(t) = \mathrm{Re}(g(t))$ *avec* $g(t) = \exp(t(i+\sqrt{3})) = \exp(\alpha t)$ *en posant* $\alpha = \sqrt{3} + i = 2\exp(i\frac{\pi}{6})$. *On a donc* $g^{(n)}(t) = \alpha^n \exp(\alpha t)$ *et* $f(t) = \mathrm{Re}(\alpha^n \exp(\alpha t)) = 2^n \cos(t + n\frac{\pi}{6})\exp(t\sqrt{3})$.

Chapitre 13

Fonctions convexes

Sommaire

Dans ce chapitre on désigne par I un intervalle non trivial de $\mathbb{R}$, et toutes les fonctions considérées vont de I vers $\mathbb{R}$. On notera $\mathring{\mathrm{I}}$ l'intervalle obtenu en ouvrant les crochets de I (intérieur de I).

I DÉFINITION

1) Paramétrage d'un segment

Dans $\mathbb{R}$.

Soient $a < b$ deux réels, la fonction $f\colon [0\,;1] \to [a\,;b]$ définie par $f(t) = (1-t)a + tb$, est une bijection (strictement croissante), on dit que f est un paramétrage du segment $[a\,;b]$, il en découle que $[a\,;b] = \{(1-t)a + tb \,/\, t \in [0\,;1]\}$.

Dans le plan muni d'un repère.

Soient $\mathrm{A}(x_\mathrm{A}, y_\mathrm{A})$ et $\mathrm{B}(x_\mathrm{B}, y_\mathrm{B})$ deux points du plan, on définit le segment $[\mathrm{A};\mathrm{B}]$ en posant :

$$[\mathrm{A};\mathrm{B}] = \left\{\mathrm{M}(x,y) \,/\, \exists t \in [0\,;1],\ \overrightarrow{\mathrm{AM}} = t\overrightarrow{\mathrm{AB}}\right\}$$

ce qui entraîne que $\mathrm{M}(x,y) \in [\mathrm{A};\mathrm{B}]$ si et seulement si il existe $t \in [0\,;1]$ tel que $\begin{cases} x - x_\mathrm{A} &= t(x_\mathrm{B} - x_\mathrm{A}) \\ y - y_\mathrm{A} &= t(y_\mathrm{B} - y_\mathrm{A}) \end{cases}$

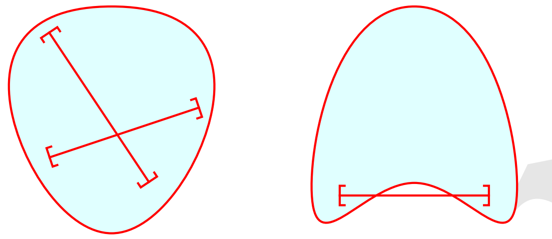
ou encore :

$$\begin{cases} x &= (1-t)x_\mathrm{A} + tx_\mathrm{B} \\ y &= (1-t)y_\mathrm{A} + ty_\mathrm{B} \end{cases}$$

La fonction $f\colon [0\,;1] \to [\mathrm{A};\mathrm{B}]$ définie par $f(t) = \mathrm{M}((1-t)x_\mathrm{A} + tx_\mathrm{B}, (1-t)y_\mathrm{A} + ty_\mathrm{B})$ est un paramétrage du segment $[\mathrm{A};\mathrm{B}]$.

2) Définition de la convexité

On définit la notion de convexité pour les parties du plan $\mathbb{R}^2$: une partie $\mathcal{V}$ non vide de $\mathbb{R}^2$ est *convexe* lorsque pour tous points A et B de $\mathcal{V}$, le segment $[A, B]$ est inclus dans $\mathcal{V}$.



Ensemble convexe Ensemble non convexe

Soit $f : \mathrm{I} \to \mathbb{R}$, on dira que f est convexe sur I lorsque la partie du plan située **au-dessus de la courbe** est une partie convexe.

Définition 13.1

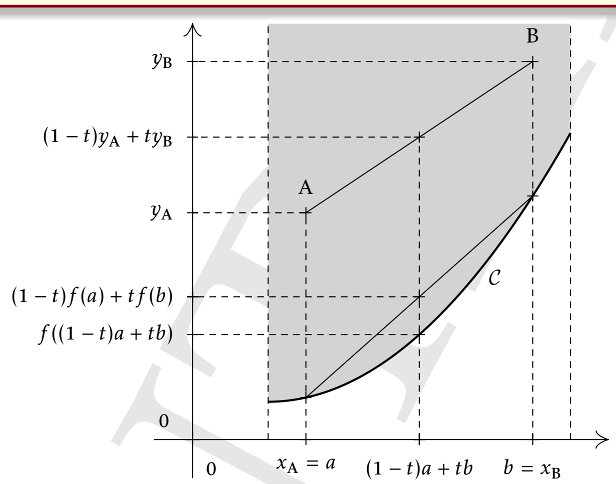
$f : \mathrm{I} \mapsto \mathbb{R}$ est dite convexe lorsque $\{(x, y) \mathbin{/} x \in \mathrm{I}$ et $y \geqslant f(x)\}$ est une partie convexe de $\mathbb{R}^2$. Ce qui revient à dire :

$f : \mathrm{I} \to \mathbb{R}$ est convexe si et seulement si pour tout réels a, b de I et tout réel t de $[0; 1]$ on a :

$$f((1-t)a + tb) \leqslant (1-t)f(a) + tf(b)$$

Cette inégalité signifie que sur l'intervalle I, ***chaque arc (de courbe) est sous sa corde****.*

On dira que f est concave sur I lorsque $-f$ est convexe.



Remarque 13.1 :

- *f est convexe sur I si et seulement si :*
  $\forall (a, b) \in \mathrm{I}^2, \forall x \in \mathbb{R}$, *si* $a \leqslant x \leqslant b$, *alors* $f(x) \leqslant \frac{f(b)-f(a)}{b-a}(x-a) + f(a) = \frac{f(b)-f(a)}{b-a}(x-b) + f(b)$.
- *f est concave sur I si et seulement si :*
  $$\forall x, y \in \mathrm{I},\ \forall t \in [0; 1],\ f((1-t)a + tb) \geqslant (1-t)f(a) + tf(b),$$
  ou encore :
  $\forall (a, b) \in \mathrm{I}^2, \forall x \in \mathbb{R}$, *si* $a \leqslant x \leqslant b$, *alors* $f(x) \geqslant \frac{f(b)-f(a)}{b-a}(x-a) + f(a) = \frac{f(b)-f(a)}{b-a}(x-b) + f(b)$.
- *La convexité est une notion globale, une fonction peut être convexe sur un intervalle et pas sur un autre. De même, une fonction peut être ni concave ni convexe sur un intervalle.*

★**Exercice 13.1**

1/ *Montrer que les fonctions affines sur $\mathbb{R}$ sont convexes.*

2/ *Montrer que la fonction $x \mapsto x^2$ est convexe sur $\mathbb{R}$.*

II PROPRIÉTÉS

1) Inégalité de Jensen

Théorème 13.1 (inégalité de Jensen)

Si $f\colon \mathrm{I} \to \mathbb{R}$ est convexe sur l'intervalle I, alors pour tous réels $a_1, \dots, a_n$ de I et tous réels $t_1, \dots, t_n$ de $[0;1]$ tels que $t_1 + \dots + t_n = 1$, on a :

$$t_1a_1 + \cdots t_na_n \in \mathrm{I} \text{ et } f(t_1a_1 + \cdots t_na_n) \leqslant t_1f(a_1) + \cdots + t_nf(a_n)$$

Preuve : Soit $m = \min\{a_1, \dots, a_n\} \in \mathrm{I}$ et $\mathrm{M} = \max\{a_1, \dots, a_n\} \in \mathrm{I}$, alors pour $i \in [\![1\,;n]\!]$, $t_im \leqslant t_ia_i \leqslant \mathrm{M}a_i$, en sommant on obtient $m \leqslant t_1a_1 + \cdots t_na_n \leqslant \mathrm{M}$ et donc $t_1a_1 + \cdots t_na_n \in \mathrm{I}$.

Pour l'inégalité, on procède par récurrence sur n, c'est vrai pour $n = 2$. Supposons le théorème établi pour un entier $n \geqslant 2$. Soient $a_1, \dots, a_n, a_{n+1}$ éléments de I et $t_1, \dots, t_n, t_{n+1}$ éléments de $[0;1]$ tels que $t_1 + \cdots + t_n + t_{n+1} = 1$, posons $s = t_1 + \cdots + t_n$, alors $s \in [0\,;1]$ et $f(t_1a_1 + \cdots t_{n+1}a_{n+1}) = f(su + t_{n+1}a_{n+1})$ en posant $u = \frac{t_1}{s}a_1 + \cdots \frac{t_n}{s}a_n$, comme $s + t_{n+1} = 1$ et que f est convexe, on a $f(su + t_{n+1}a_{n+1}) \leqslant sf(u) + t_{n+1}f(a_{n+1})$, d'autre part on a $f(u) = f(\frac{t_1}{s}a_1 + \cdots \frac{t_n}{s}a_n) \leqslant \frac{t_1}{s}f(a_1) + \cdots \frac{t_n}{s}f(a_n)$ en appliquant l'hypothèse de récurrence, car les réels $\frac{t_i}{s}$ pour $i \in [\![1\,;n]\!]$, sont dans $[0\,;1]$ et de somme égale à 1. En reportant dans l'inégalité précédente, s étant positif ou nul, on obtient $f(t_1a_1 + \cdots t_{n+1}a_{n+1}) \leqslant \sum\limits_{i=1}^{n} t_if(a_i)$. □

2) Convexité et pentes

Théorème 13.2

$f\colon \mathrm{I} \to \mathbb{R}$ est convexe sur I si et seulement si pour tout $a \in \mathrm{I}$, la fonction $t_a\colon x \mapsto \frac{f(x)-f(a)}{x-a}$ est croissante sur $\mathrm{I} \setminus \{a\}$.

Preuve : Si f est convexe sur I, soit $a \in \mathrm{I}$, envisageons plusieurs cas :

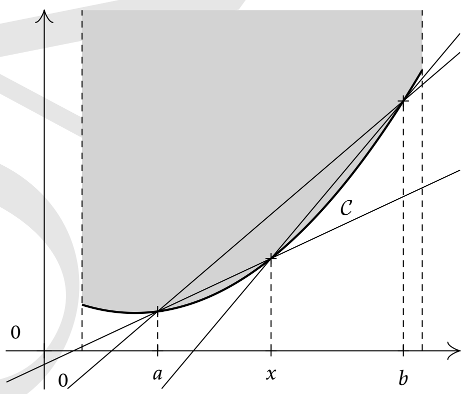
- Si $a < x < y$ sont dans I, l'arc sur $[a;y]$ est sous la corde, donc $f(x) \leqslant \frac{f(y)-f(a)}{y-a}(x-a) + f(a)$ ce qui entraîne $\frac{f(x)-f(a)}{x-a} \leqslant \frac{f(y)-f(a)}{y-a}$ car $x - a > 0$.
- Si $x < y < a$ sont dans I, on a $f(y) \leqslant \frac{f(x)-f(a)}{x-a}(y-a) + f(a)$ ce qui entraîne $\frac{f(y)-f(a)}{y-a} \geqslant \frac{f(x)-f(a)}{x-a}$ car $y - a < 0$.
- Si $x < a < y$ sont dans I alors $f(a) \leqslant \frac{f(y)-f(x)}{y-x}(a-x)+f(x)$ d'où $\frac{f(a)-f(x)}{a-x} \leqslant \frac{f(y)-f(x)}{y-x}$ car $a-x > 0$. Mais on a aussi $f(a) \leqslant \frac{f(y)-f(x)}{y-x}(a-y)+f(y)$ d'où $\frac{f(a)-f(y)}{a-y} \geqslant \frac{f(y)-f(x)}{y-x}$ car $a-y < 0$. Par conséquent $\frac{f(a)-f(x)}{a-x} \leqslant \frac{f(y)-f(x)}{y-x} \leqslant \frac{f(y)-f(a)}{y-a}$.

La fonction $t_a\colon x \mapsto \frac{f(x)-f(a)}{x-a}$ est donc croissante sur $\mathrm{I} \setminus \{a\}$.

Réciproquement, si la fonction $t_a\colon x \mapsto \frac{f(x)-f(a)}{x-a}$ est croissante sur $\mathrm{I} \setminus \{a\}$ pour tout $a \in \mathrm{I}$. Soit $a < x < b$ dans I, alors $\frac{f(x)-f(a)}{x-a} \leqslant \frac{f(b)-f(a)}{b-a}$ donc $f(x) \leqslant \frac{f(b)-f(a)}{b-a}(x-a) + f(a)$, ce qui signifie que sur $[a;b]$, l'arc est sous la corde, f est donc convexe. □

Remarque 13.2 :

- *La fonction t_a est la fonction taux d'accroissement de f en a.*
- *Si f est convexe sur I et si $a < x < b$ sont dans I, alors $t_a(x) \leqslant t_a(b) = t_b(a) \leqslant t_b(x)$ (inégalité des trois pentes).*



☞**Exemple** : Sachant que la fonction ln est concave sur $]0\,;+\infty[$, on peut affirmer que la fonction $x \mapsto \frac{\ln(x)}{x-1}$ est décroissante sur son ensemble de définition, car c'est le taux d'accroissement en 1 de la fonction ln.

★**Exercice 13.2** *Si f est convexe sur I et admet un minimum local en a, montrer que c'est un minimum global.*

3) Convexité des fonctions dérivables

Théorème 13.3 (Caractérisation de la convexité pour les fonctions dérivables)

Si $f : \mathrm{I} \to \mathbb{R}$ est continue sur l'intervalle I et dérivable sur $\mathring{\mathrm{I}}$, alors f est convexe sur I si et seulement si f' est croissante sur $\mathring{\mathrm{I}}$.

Preuve : Si f est convexe, soient $x_0 < x < x_1$ dans $\mathring{\mathrm{I}}$, alors on a $t_{x_0}(x) \leqslant t_{x_0}(x_1) = t_{x_1}(x_0) \leqslant t_{x_1}(x)$, en faisant tendre x vers x_0 par la droite on obtient $f'(x_0) \leqslant t_{x_0}(x_1)$ et si on fait tendre x vers x_1 par la gauche, on obtient $t_{x_0}(x_1) \leqslant f'(x_1)$, par conséquent $f'(x_0) \leqslant f'(x_1)$, donc f' est croissante.

Si f' est croissante sur $\mathring{\mathrm{I}}$, soient $a < b$ dans I et $g(x) = f(x) - \frac{f(b)-f(a)}{b-a}(x-a) + f(a)$, alors g est continue sur $[a\,;b]$, dérivable sur $]a\,;b[$ et $g'(x) = f'(x) - \frac{f(b)-f(a)}{b-a}$, or il existe $c \in\,]a\,;b[$ tel que $\frac{f(b)-f(a)}{b-a} = f'(c)$ (IAF), on en déduit que g est décroissante sur $[a;c]$ puis croissante sur $[c;b]$, or $g(a) = g(b) = 0$, donc $g \leqslant 0$ sur $[a\,;b]$, ce qui montre que f est convexe sur I.

x	a		c		b
g'		$-$	0	$+$	
g	0	$\searrow$	$g(c)$	$\nearrow$	0

□

Remarque 13.3 – *Il en découle qu'une fonction f continue sur l'intervalle I et deux fois dérivable sur $\mathring{\mathrm{I}}$ est convexe si et seulement si sa dérivée seconde et positive sur $\mathring{\mathrm{I}}$ (et concave si et seulement si $f'' \leqslant 0$ sur $\mathring{\mathrm{I}}$).*

☞**Exemples** :
- La fonction exp est convexe sur $\mathbb{R}$.
- La fonction ln est concave sur $]0;+\infty[$.
- La fonction sin est convexe sur $[-\pi;0]$ et concave sur $[0;\pi]$.
- La fonction $x \mapsto x^n$ avec $n \in \mathbb{N}^*$ est convexe sur $\mathbb{R}$ lorsque n est pair, sur $[0\,;+\infty[$ seulement si n est impair.

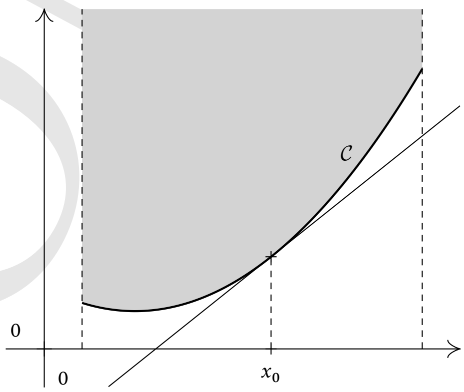
Théorème 13.4 (Caractérisation géométrique avec les tangentes)

Si $f : \mathrm{I} \to \mathbb{R}$ est continue sur l'intervalle I et dérivable sur $\mathring{\mathrm{I}}$, alors f est convexe sur I si et seulement si $\forall x_0 \in \mathring{\mathrm{I}}$, la courbe de f est au-dessus de la tangente au point d'abscisse x_0, ce qui se traduit par :

$$\forall x_0 \in \mathring{\mathrm{I}}, \forall x \in \mathrm{I},\ f(x) \geqslant f'(x_0)(x-x_0) + f(x_0).$$

Preuve : Si f est convexe, soit $x_0 < x_1 \in \mathring{\mathrm{I}}$, on pose pour $x \in \mathrm{I}$, $g(x) = f(x) - f'(x_0)(x-x_0) - f(x_0)$, g est dérivable sur $\mathring{\mathrm{I}}$ et $g'(x) = f'(x) - f'(x_0)$, comme f' est croissante sur $\mathring{\mathrm{I}}$, on a g décroissante à gauche de x_0 et croissante à droite, or $g(x_0) = 0$, donc g est positive sur I, la courbe de f est bien au-dessus de la tangente au point d'abscisse x_0.

Réciproquement : soit $x_0 < x_1 \in \mathring{\mathrm{I}}$, on a $f(x_0) \geqslant f'(x_1)(x_0-x_1)+f(x_1)$, d'où $t_{x_0}(x_1) \leqslant f'(x_1)$ car $x_0 - x_1 < 0$. D'autre part, on a $f(x_1) \geqslant f'(x_0)(x_1-x_0) + f(x_0)$, d'où $t_{x_0}(x_1) \geqslant f'(x_0)$ car $x_1 - x_0 > 0$, on a donc $f'(x) \leqslant t_{x_0}(x_1) \leqslant f'(x_1)$, f' donc croissante sur $\mathring{\mathrm{I}}$ et donc f est convexe sur I. □



4) Régularité des fonctions convexes

Théorème 13.5

Soit f une fonction convexe sur I et $a \in \mathring{I}$, alors f est dérivable à gauche et à droite en a. De plus, si $x < a < y$ alors :

$$\frac{f(x)-f(a)}{x-a} \leqslant f_g'(a) \leqslant f_d'(a) \leqslant \frac{f(y)-f(a)}{y-a}.$$

En particulier, si $a < b$ sont dans $\mathring{I}$, alors $f_d'(a) \leqslant f_g'(b)$.

Preuve : La fonction t_a est croissante sur $I \setminus \{a\}$ donc elle majorée sur $I\cap]-\infty;a[$ par $t_a(y)$, par conséquent elle admet une limite finie à gauche en a qui est $\sup_{x<a} t_a(x)$, cela signifie que f est dérivable à gauche en a et $f_g'(a) = \sup_{x<a} t_a(x) \leqslant t_a(y)$.

Sur $I \cap]a;+\infty[$ la fonction t_a est croissante minorée par $f_g'(a)$, elle admet une limite finie à droite en a qui est $\inf_{x>a} t_a(x)$, cela signifie que f est dérivable à droite en a et $f_g'(a) \leqslant f_d'(a) = \inf_{x>a} t_a(x) \leqslant t_a(y)$. □

Remarque 13.4 –
- *La fonction f n'a aucune raison d'être dérivable aux bornes* I*, par exemple, la fonction* arcsin *est convexe sur* $[0;1]$ *mais non dérivable à gauche en* 1.
- *Le théorème ne dit pas que f est dérivable en a ! Considérer par exemple $f(t) = |t|$ sur $[-1;1]$, elle est convexe mais non dérivable en* 0.

Application – Si f une fonction convexe sur I et $a \in \mathring{I}$, alors :
- sur $I\cap]-\infty;a[$ la courbe de f est au-dessus de la droite d'équation $y = f_g'(a)(x-a)+f(a)$ (tangente à la courbe à gauche au point d'abscisse c). En effet, pour $x < a$, l'inégalité $t_a(x) \leqslant f_g'(a)$ équivaut à $f(x) \geqslant f_g'(a)(x-a)+f(a)$.
- sur $I \cap]a;+\infty[$ la courbe de f est au-dessus de la droite d'équation $y = f_d'(a)(x-a)+f(a)$ (tangente à la courbe à droite au point d'abscisse c). En effet, pour $x > a$, l'inégalité $f_d'(a) \leqslant t_a(x)$ équivaut à $f(x) \geqslant f_d'(a)(x-a)+f(a)$.
- si f est dérivable en a alors on retrouve que la courbe de f est au-dessus de la tangente au point d'abscisse a.

Il découle du théorème précédent :

Théorème 13.6

Si f est convexe sur I *alors f est continue en tout point **intérieur** à* I.

Remarque 13.5 – *La fonction f n'a aucune raison d'être continue aux bornes* I*, par exemple, la fonction f définie sur* $[0;1]$ *par* $f(t) = 0$ *si* $t \in [0;1[$ *et* $f(1) = 1$ *est convexe mais discontinue en* 1.

★**Exercice 13.3** *Une fonction f **convexe** sur un **segment** $[a;b]$ admet un maximum global en a ou en b. Si ce maximum global est également atteint à l'intérieur de $[a;b]$, alors f est constante.*

III INÉGALITÉS DE CONVEXITÉ

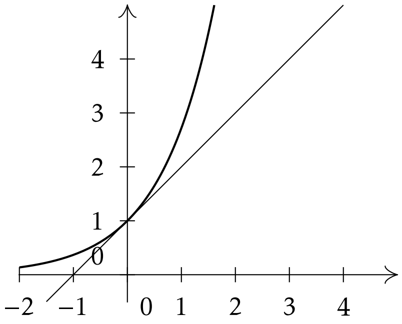
1) Tangentes ou sécantes

Lorsqu'une fonction f est dérivable et convexe sur I, on sait que la courbe est au-dessus chacune de ses tangentes :

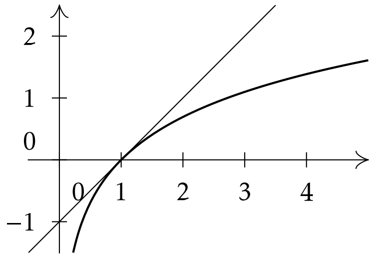
$$\forall a,\ x \in I,\ f(x) \geqslant f'(a)(x-a)+f(a)$$

Pour les fonctions concaves, l'inégalité est inversée.

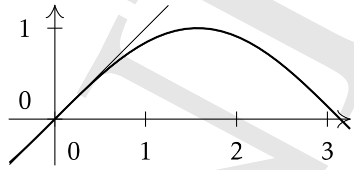
☞**Exemples** :
- La fonction exp est convexe sur $\mathbb{R}$ donc $\forall a,\ x \in \mathbb{R},\ e^x \geqslant e^a(x-a)+e^a$. En particulier avec $a=0$ on retrouve l'inégalité classique : $e^x \geqslant x+1$.
- La fonction ln est concave sur $]0;+\infty[$ donc $\forall a,\ x \in]0;+\infty[,\ \ln(x) \leqslant \frac{x-a}{a}+\ln(a)$. En particulier avec $a=1$ on retrouve l'inégalité classique : $\forall x > 0,\ \ln(x) \leqslant x-1$.
- La fonction sin est concave sur $[0;\pi]$ donc $\forall a,\ x \in [0;\pi],\ \sin(x) \leqslant \cos(a)(x-a)+\sin(a)$. En particulier avec $a=0$ on retrouve l'inégalité classique : $\forall x \in [0;\pi],\ \sin(x) \leqslant x$.
- …



$e^x \geqslant x+1$



$\ln(x) \leqslant x-1$



$\sin(x) \leqslant x$

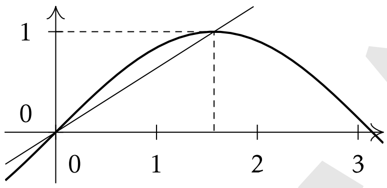
Lorsqu'une fonction f est convexe sur I, on sait que sur tout segment $[a;b] \subset$ I l'arc est sous la corde :

$$\forall x \in [a;b],\ f(x) \leqslant \frac{f(b)-f(a)}{b-a}(x-a)+f(a)$$

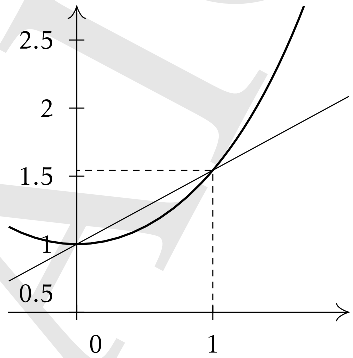
Pour les fonctions concaves, l'inégalité est inversée.

☞**Exemples** :

- La fonction sin est concave sur $[0;\frac{\pi}{2}]$ donc $\forall x \in [0;\frac{\pi}{2}],\ \sin(x) \geqslant \frac{2x}{\pi}$.
- La fonction ch est convexe sur $[0;a]$ $(a>0)$ donc $\forall x \in [0;a],\ \mathrm{ch}(x) \leqslant \frac{\mathrm{ch}(a)-1}{a}x+1$.
- …



$\sin(x) \geqslant \frac{2x}{\pi}$ sur $[0;\frac{\pi}{2}]$



$\mathrm{ch}(x) \leqslant \frac{\mathrm{ch}(a)-1}{a}x+1$ sur $[0;1]$

2) Inégalités de moyennes

Soient $x_1,\dots,x_n$ des réels strictement positifs, la fonction ln étant concave, on a :

$$\ln\left(\frac{x_1+\cdots+x_n}{n}\right) \geqslant \frac{\ln(x_1)+\cdots\ln(x_n)}{n} = \frac{1}{n}\ln\left(\prod_{i=1}^{n} x_i\right)$$

en passant à l'exponentielle on obtient une inégalité entre la moyenne géométrique et la moyenne arithmétique :

$$\sqrt[n]{\prod_{i=1}^{n} x_i} \leqslant \frac{x_1+\cdots x_n}{n}$$

en remplaçant x_i par son inverse, inégalité entre la moyenne harmonique et la moyenne géométrique :

$$\frac{n}{\frac{1}{x_1}+\cdots+\frac{1}{x_n}} \leqslant \sqrt[n]{\prod_{i=1}^{n} x_i}$$

et finalement :

$$\boxed{\frac{n}{\frac{1}{x_1}+\cdots+\frac{1}{x_n}} \leqslant \sqrt[n]{\prod_{i=1}^{n} x_i} \leqslant \frac{x_1+\cdots x_n}{n}}$$

3) Inégalités de Hölder et Minkowski

Soient u, v, p et q strictement positifs avec $\frac{1}{p}+\frac{1}{q}=1$, alors la concavité de la fonction ln permet d'écrire : $\ln(\frac{1}{p}u^p+\frac{1}{q}v^q)\geqslant\frac{1}{p}\ln(u^p)+\frac{1}{a}\ln(v^q)=\ln(uv)$, en passant à l'expoentielle on obtient :

$$\boxed{uv\leqslant\frac{1}{p}u^p+\frac{1}{q}v^q}$$

Soient $a_i,\ldots,a_n,\ b_1,\ldots,b_n$ strictement positifs, $p,\ q>0$ tels que $\frac{1}{p}+\frac{1}{q}=1$. Soit $\mathrm{A}=\left(\sum_{i=1}^n a_i^p\right)^{1/p}$ et $\mathrm{B}=\left(\sum_{i=1}^n b_i^q\right)^{1/q}$, on a $\frac{a_i}{\mathrm{A}}\frac{b_i}{\mathrm{B}}\leqslant\frac{1}{p}\frac{a_i^p}{\mathrm{A}^p}+\frac{1}{q}\frac{b_i^q}{\mathrm{B}^q}$, en sommant de 1 à n on obtient : $\frac{\sum_{i=1}^n a_ib_i}{\mathrm{AB}}\leqslant\frac{1}{p}+\frac{1}{q}=1$, d'où $\sum_{i=1}^n a_ib_i\leqslant\mathrm{AB}$, c'est à dire :

$$\boxed{\sum_{i=1}^n a_ib_i\leqslant\left(\sum_{i=1}^n a_i^p\right)^{1/p}\left(\sum_{i=1}^n b_i^q\right)^{1/q}\quad\text{(inégalité de Hölder)}}$$

Soient $a_i,\ldots,a_n,\ b_1,\ldots,b_n$ strictement positifs, $p>1$, on pose $q=\frac{1}{1-\frac{1}{p}}=\frac{p}{p-1}$ de tel sorte que $q>0$ et $\frac{1}{p}+\frac{1}{q}=1$. L'inégalité de Hölder donne $\sum_{i=1}^n a_i(a_i+b_i)^{p-1}\leqslant\left(\sum_{i=1}^n a_i^p\right)^{1/p}\left(\sum_{i=1}^n(a_i+b_i)^{(p-1)q}\right)^{1/q}$, de même, on a $\sum_{i=1}^n b_i(a_i+b_i)^{p-1}\leqslant\left(\sum_{i=1}^n b_i^p\right)^{1/p}\left(\sum_{i=1}^n(a_i+b_i)^{(p-1)q}\right)^{1/q}$, en ajoutant ces deux inégalités, obtient $\sum_{i=1}^n(a_i+b_i)^p\leqslant\left[\left(\sum_{i=1}^n a_i^p\right)^{1/p}+\left(\sum_{i=1}^n b_i^p\right)^{1/p}\right]\left(\sum_{i=1}^n(a_i+b_i)^{(p-1)q}\right)^{1/q}$, c'est à dire :

$$\sum_{i=1}^n(a_i+b_i)^p\leqslant\left[\left(\sum_{i=1}^n a_i^p\right)^{1/p}+\left(\sum_{i=1}^n b_i^p\right)^{1/p}\right]\left(\sum_{i=1}^n(a_i+b_i)^p\right)^{1-1/p}$$

d'où :

$$\boxed{\left(\sum_{i=1}^n(a_i+b_i)^p\right)^{1/p}\leqslant\left(\sum_{i=1}^n a_i^p\right)^{1/p}+\left(\sum_{i=1}^n b_i^p\right)^{1/p}\quad\text{(inégalité de Minkowski)}}$$

IV SOLUTION DES EXERCICES

Solution 13.1

1/ *Si $f(x)=ax+b$, alors pour u, v, réels et $t\in[0\,;1]$, $f(tu+(1-t)v)=a(tu+(1-t)v)+v=t[au+b]+(1-t)[av+b]=t(u)+(1-t)f(v)$, on a donc bien $f(tu+(1-t)v)\leqslant t(u)+(1-t)f(v)$.*

2/ *Soient a,b réels et $t\in[0\,;1]$, on étudie $g(t)=(ta+(1-t)b)^2-ta^2-(1-t)b^2$ sur $[0\,;1]$, on a $g'(t)=-(a-b)^2(1-2t)$, g est décroissante puis croissante avec un minimum pour $t=\frac{1}{2}$, et $g(0)=g(1)=0$, ce qui entraîne que $g\leqslant 0$.*

Solution 13.2 *Au voisinage à gauche de a la fonction t_a est négative, comme cette fonction est croissante on en déduit que pour $x<a$ on a $\frac{f(x)-f(a)}{x-a}\leqslant 0$ et donc $f(x)\geqslant f(a)$.*

Au voisinage à droite de a la fonction t_a est positive, comme cette fonction est croissante on en déduit que pour $x>a$ on a $\frac{f(x)-f(a)}{x-a}\geqslant 0$ et donc $f(x)\geqslant f(a)$.

Solution 13.3 *$f(ta+(1-t)b)\leqslant tf(a)(1-t)f(b)\leqslant\mathrm{M}$ où $\mathrm{M}=\max(f(a),f(b))$. Si $\mathrm{M}=f(c)$ avec $c\in]a;b[$, à gauche de c on a $t_c(x)\geqslant 0$, alors qu'à droite de c on a $t_c(x)\leqslant 0$, or la fonction t_c est croissante sur $[a;b]\setminus\{c\}$, par conséquent la fonction t_c est nulle, ce qui signifie que f est constante.*

Chapitre 14

Structures algébriques

Sommaire

I LOIS DE COMPOSITION INTERNE

1) Définitions

Définition 14.1 (loi de composition interne)
Soit E un ensemble non vide. On appelle loi de composition interne (ou lci) dans E toute application de $E \times E$ vers E. L'image d'un couple $(x, y) \in E^2$ par une telle application est en général noté sous la forme d'une opération avec un symbole : $x + y$ ou $x \times y$ ou $x * y$ ou $x.y$ ou $x\mathrm{T}y$, ... Et on écrira $(E, *)$ pour dire que l'ensemble E est muni d'une lci notée $*$.

☞**Exemples :**
- L'addition et la multiplication des nombres sont des lci dans $\mathbb{N}$, dans $\mathbb{Z}$, dans $\mathbb{Q}$, dans $\mathbb{R}$, dans $\mathbb{C}$ mais pas dans $[-2; 2]$ par exemple.
- L'addition et la multiplication des fonctions dans $(\mathcal{F}(A, \mathbb{C})$ sont des lci (A étant un ensemble non vide). En particulier l'addition et la multiplication des suites sont des lci dans $\mathcal{F}(\mathbb{N}, \mathbb{C})$.
- Soit E un ensemble non vide, dans $\mathcal{F}(E, E)$, la composition des applications (dite loi $\circ$) est une lci.

★**Exercice 14.1** *Soit $E =]-1; 1[$. Pour $x, y \in E$, on pose $x * y = \frac{x+y}{1+xy}$. Montrer que l'on définit ainsi une lci dans E.*

Définition 14.2 (associativité, commutativité)
Soit $(E, *)$ un ensemble muni d'une lci.
• On dit que la loi est associative lorsque : $\forall x, y, z \in E$, $x * (y * z) = (x * y) * z$.
• On dit que deux éléments x et y de E commutent lorsque $x * y = y * x$. Si tous les éléments commutent deux à deux, non dit que la lci est commutative.

☞**Exemples :**

– L'addition et la multiplication des nombres dans $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$, sont associatives et commutatives.
– L'addition et la multiplication des fonctions dans $\mathcal{F}(A, \mathbb{C})$ (A étant un ensemble non vide) sont associatives et commutatives.
– Dans $(\mathcal{F}(E, E), \circ)$, la loi est associative mais non commutative en général.
– Dans $(\mathbb{R}, *)$ avec $x * y = x - y$ pour tout $x, y \in \mathbb{R}$, la loi est une lci non commutative ($2 * 3 = -1$ mais $3 * 2 = 1$) et non associative ($(1 * 2) * 3 = (-1) * 3 = -4$ et $1 * (2 * 3) = 1 * (-1) = 2$).

★**Exercice 14.2** *Soit* $E =]-1;1[$. *Pour* $x, y \in E$, *on pose* $x * y = \frac{x+y}{1+xy}$. *Montrer que l'opération* $*$ *est associative et commutative.*

2) Élément neutre

Définition 14.3
Soit $(E, *)$ *un ensemble muni d'une lci et soit* $e \in E$, *on dit que* e *est un élément neutre pour la loi* $*$ *lorsque* $\forall x \in E, x * e = e * x = x$. *Dans le cas où l'opération est notée additivement, l'élément neutre (s'il existe) est appelé noté en général* 0_E *(zéro de* E*). Dans le cas où l'opération est notée multiplicativement, l'élément neutre (s'il existe) est noté en général* 1_E *(un de* E*).*

☞**Exemples :**
– L'élément neutre de l'addition des nombres est 0. Celui de la multiplication des nombres est 1.
– L'élément neutre de l'addition des fonctions dans $(\mathcal{F}(A, \mathbb{C})$ est la fonction constamment nulle. Celui de la multiplication est la fonction constante $x \mapsto 1$.
– Dans $(\mathcal{F}(E, E), \circ)$, id_E est élément neutre.
– Dans $(\mathbb{R}, *)$ avec $x * y = x - y$ il y a un élément neutre à droite qui est 0 car $\forall x \in \mathbb{R}, x - 0 = x$, mais il n'y a pas d'élément neutre à gauche.

★**Exercice 14.3** *Soit* $E =]-1;1[$. *Pour* $x, y \in E$, *on pose* $x * y = \frac{x+y}{1+xy}$. *Montrer qu'il y a un élément neutre pour cette opération.*

Remarque 14.1 – *Si* $(E, *)$ *possède un élément neutre, alors celui-ci est unique. En effet, si on a deux éléments neutre* e *et* e', *alors* $e * e' = e$ *car* e' *est neutre, et* $e * e' = e'$ *car* e *est neutre, d'où* $e = e'$.

Définition 14.4
Soit $(E, *)$ *un ensemble muni d'une lci possédant un élément neutre* e. *On dit que* $x \in E$ *est symétrisable lorsqu'il existe* $x' \in E$ *tel que* $x * x' = x' * x = e$. *Si c'est le cas, on dit que* x' *est un symétrique de* x. *Dans le cas où l'opération est notée additivement, le symétrique de* x *est appelé opposé de* x *et noté* $-x$. *Dans le cas où l'opération est notée multiplicativement, le symétrique de* x *est appelé inverse de* x *et noté* x^{-1}.

☞**Exemples :**
– Dans $(\mathbb{C}, +)$, $(\mathbb{R}, +)$, $(\mathbb{Q}, +)$ et $(\mathbb{Z}, +)$, chaque élément possède un opposé, mais pas dans $(\mathbb{N}, +)$.
– Dans $(\mathbb{C}^*, \times)$, $(\mathbb{R}^*, \times)$, $(\mathbb{Q}^*, \times)$, chaque élément possède un inverse, mais pas dans $(\mathbb{Z}^*, \times)$.
– Dans $(\mathcal{F}(A, \mathbb{C}), +)$ chaque fonction possède un opposé.
– Dans $(\mathcal{F}(A, \mathbb{C}), \times)$, seules les fonctions qui ne s'annulent jamais ont un inverse.
– Dans $(\mathcal{F}(E, E), \circ)$, un élément f de $\mathcal{F}(E, E)$ a un symétrique si et seulement si il existe g dans $\mathcal{F}(E, E)$ telle que $f \circ g = g \circ f = \mathrm{id}_E$, ce qui équivaut à dire que f est une bijection, auquel cas le symétrique de f est la bijection réciproque f^{-1}.

Théorème 14.1
Soit $(E, *)$ *un ensemble muni d'une lci associative et possédant un élément neutre* e. *Si un élément* x *possède un symétrique* x' *dans* E, *alors celui-ci est unique.*

Preuve : Si x' et x'' sont deux symétriques de x, alors $x' = x' * e = x' * (x * x'') = (x' * x) * x'' = e * x'' = x''$. □

★**Exercice 14.4** *Soit* $E =]-1;1[$. *Pour* $x, y \in E$, *on pose* $x * y = \frac{x+y}{1+xy}$. *Montrer que chaque élément a un symétrique dans* E.

Définition 14.5 (Partie stable)
Soit $(\mathrm{E},*)$ *un ensemble muni d'une lci et soit* A *une partie non vide de* E*, on dit que* A *est stable pour l'opération* $*$*, lorsque :*
$$\forall (x,y)\in \mathrm{A}^2, x*y\in \mathrm{A}.$$
Dans ce cas, l'opération $*$ *induit une lci dans* A.

☞**Exemples** :
- $\mathbb{Q}$ est une partie stable de $(\mathbb{R},+)$, mais pas $[0\,;1]$.
- $\mathbb{R}^{*+}$ est une partie stable de $(\mathbb{R},\times)$, mais pas $\mathbb{R}^{*-}$.

II STRUCTURE DE GROUPE

1) Définitions

Définition 14.6
Un groupe est un ensemble non vide G *muni d'une opération* $*$ *(ou loi de composition) qui vérifie les propriétés suivantes :*
- *elle doit être interne :* $\forall\ x,y\in \mathrm{G}, x*y\in \mathrm{G}$.
- *elle doit être associative :* $\forall\ x,y,z\in \mathrm{G}, x*(y*z)=(x*y)*z$.
- *elle doit posséder un élément neutre :* $\exists\ e\in \mathrm{G}, \forall\ x\in \mathrm{G}, e*x=x*e=x$. *Si la loi est une addition l'élément neutre sera noté* 0_{G} *et on parlera de groupe additif. Si la loi est une multiplication, l'élément neutre sera noté* 1_{G} *et on parlera de groupe multiplicatif. Dans le cas général l'élément neutre est souvent noté* e_{G}.
- *tout élément de* G *doit avoir un symétrique dans* G *:* $\forall\ x\in \mathrm{G}, \exists\ x'\in \mathrm{G}, x*x'=x'*x=e_{\mathrm{G}}$. *En notation additive, le symétrique de* x *est appelé* ***opposé de*** x *et noté* $-x$*, en notation multiplicative on l'appelle* ***inverse de*** x *et on le note* x^{-1}.

Lorsque toutes ces conditions sont remplies, on dit $(\mathrm{G},*)$ *est un groupe. Si en plus la loi* $*$ *est commutative* ($\forall\ x,y\in \mathrm{G}, x*y=y*x$)*, alors on dit que* $(\mathrm{G},*)$ *est un* ***groupe abélien*** *(ou groupe commutatif).*

☞**Exemples** :
- $(\mathbb{Z},+)$, $(\mathbb{Q},+)$, $(\mathbb{R},+)$, $(\mathbb{C},+)$, $(\mathbb{Q}^*,\times)$, $(\mathbb{R}^*,\times)$, $(\mathbb{C}^*,\times)$ sont des groupes abéliens.
- $(\mathbb{N},+)$ et $(\mathbb{Z}^*,\times)$ ne sont pas des groupes.
- $(\mathcal{F}(\mathrm{I},\mathbb{C}),+)$ est un groupe abélien pour l'addition des fonctions (ou des suites si $\mathrm{I}=\mathbb{N}$).
- Si E est un ensemble non vide, on note $\mathrm{S}_{\mathrm{E}}=\{f:\mathrm{E}\to\mathrm{E}\,/\,f \text{ est bijective}\}$, alors $(\mathrm{S}_{\mathrm{E}},\circ)$ est un groupe (non abélien en général), on l'appelle **groupe des permutations** de E.
- Soient $(\mathrm{G},\cdot)$ et $(\mathrm{G}',*)$ sont deux groupes, (a,a') et (b,b') dans $\mathrm{G}\times\mathrm{G}'$, on pose $(a,a')\mathrm{T}(b,b')=(a\cdot a',b*b')$, alors $(\mathrm{G}\times\mathrm{G}',\mathrm{T})$ est un groupe, c'est **le groupe produit** de G et G'. Ceci peut être généralisé au produit cartésien de n groupes.

À retenir : Règles de calculs
Soit $(\mathrm{G},*)$ un groupe :
- Soient $x,y\in\mathrm{G}$, le symétrique de $x*y$ est : $(x*y)'=y'*x'$.
- Soient $a,b\in\mathrm{G}$, l'équation $a*x=b$ admet comme unique solution dans G, $x=a'*b$.
- Soient $a,b,c\in\mathrm{G}$, si $a*b=a*c$ alors $b=c$ (régularité à gauche).
- Soient $a,b,c\in\mathrm{G}$, si $b*a=c*a$ alors $b=c$ (régularité à droite).

★**Exercice 14.5** *Soit* $n\geqslant 2$*, dans* $\mathbb{K}^n$ ($\mathbb{K}=\mathbb{R}$ *ou* $\mathbb{K}=\mathbb{C}$) *on définit l'addition :*
$$(x_1,\dots,x_n)+(y_1,\dots,y_n)=(x_1+y_1,\dots,x_n+y_n).$$
Montrer que $(\mathbb{K}^n,+)$ *est un groupe abélien.*

2) Morphisme de groupes

Définition 14.7

Soit $(G,\cdot)$ et $(G',*)$ deux groupes, un morphisme de groupes de G vers G' est une application $f\colon G \to G'$ telle que $\forall\, x,y \in G, f(x\cdot y) = f(x)*f(y)$.
Si de plus f est bijective, alors on dit que f est un isomorphisme de groupes.

☞**Exemples** :
- $\ln\colon (]0\,;+\infty[,\times) \to (\mathbb{R},+)$ est un morphisme de groupes, c'est même un isomorphisme de groupes.
- $\exp\colon (\mathbb{R},+) \to (]0\,;+\infty[,\times)$ est un morphisme de groupes, c'est même un isomorphisme de groupes.
- $f\colon (\mathbb{Z},+) \to (\mathbb{U}_n,\times)$ définie par $f(k) = \exp(2ik\pi/n)$ est un morphisme de groupes, mais non bijectif.

Théorème 14.2

On a les propriétés suivantes :
• Si $f\colon (G,\cdot) \to (G',*)$ est un morphisme de groupes, alors $f(e) = e'$ et $\forall\, x \in G, f(x^{-1}) = f(x)^{-1}$ (où e et e' désignent les éléments neutres respectifs de G et G').
• La composée de deux morphismes de groupes est un morphisme de groupes.
• Si $f\colon (G,\cdot) \to (G',*)$ est un isomorphisme de groupes, alors la bijection réciproque $f^{-1}\colon (G',*) \to (G,.)$ est un morphisme de groupes.

Preuve : Celle-ci est simple et laissée en exercice. □

☞**Exemple** : $\ln\colon (]0\,;+\infty[,\times) \to (\mathbb{R},+)$ est isomorphisme de groupes, donc $\ln(1) = 0, \forall\, x > 0, \ln(1/x) = -\ln(x)$ et la bijection réciproque, exp, est un morphisme de groupes, *i.e.* :

$$\forall\, x,y \in \mathbb{R}, \exp(x+y) = \exp(x)\times\exp(y)$$

Théorème 14.3

Soit $f\colon (\mathbb{R},+) \to (\mathbb{R},+)$ un morphisme de groupes continu en 0, alors :
$$\forall x \in \mathbb{R},\ f(x) = ax \text{ (avec } a = f(1)).$$

Preuve : On pose $a = f(1)$, on montre que $\forall n \in \mathbb{N}, f(n) = an$ (récurrence), on en déduit que $f(-n) = a(-n)$ car $f(-n) = -f(n)$, d'où : $\forall n \in \mathbb{Z}, f(n) = an$.

Soit $r = \frac{p}{q}$ un rationnel avec $q \in \mathbb{N}^*$, alors $f(qr) = f(p) = ap = qf(r)$ d'où $f(r) = ar$.

Soit $x \in \mathbb{R}$ et (r_n) une suite de rationnels qui converge vers x, alors $(x-r_n)$ converge vers 0 et donc $f(x-r_n)$ tend vers $f(0) = 0$, or $f(x-r_n) = f(x) - f(r_n)$ donc $(f(r_n))$ converge vers $f(x)$. Or $f(r_n) = ar_n \to ax$, par conséquent $f(x) = ax$. □

3) Sous-groupes d'un groupe

Définition 14.8

Soit $(G,\cdot)$ un groupe et H un ensemble, on dit que H est un sous-groupe de $(G,\cdot)$ lorsque :
$$H \subset G\,;\, H \neq \varnothing, \forall\, x,y \in H, x\cdot y \in H \text{ et } x^{-1} \in H.$$

☞**Exemples** :
- $\{e\}$ et G sont des sous-groupes de $(G,\cdot)$, ils sont appelés sous-groupes triviaux de $(G,\cdot)$.
- Si $(H,\cdot)$ est un groupe inclus dans un groupe $(G,\cdot)$ pour la même loi, alors H est un sous-groupe de G, car on vérifie facilement que l'élément neutre de G est forcément égal à l'élément neutre de H, ce qui entraîne pour $x \in H$, que son symétrique dans G et son symétrique dans H sont les mêmes.
- $(\mathbb{U},\times)$ et $(\mathbb{U}_n,\times)$ sont des sous-groupes de $(\mathbb{C}^*,\times)$.
- L'ensemble des fonctions définies sur $\mathbb{R}$ et 2π-périodiques est un groupe additif, car c'est un sous-groupe de $(\mathcal{F}(\mathbb{R},\mathbb{R}),+)$.
- L'ensemble des entiers pairs est un groupe additif, car c'est un sous-groupe de $(\mathbb{Z},+)$.
- $\mathbb{Z}[i] = \{a+ib\,/\,a,b \in \mathbb{Z}\}$ est un sous-groupe de $(\mathbb{C},+)$.

Remarque 14.2 – *Si* H *est un sous-groupe de* $(\mathrm{G},\cdot)$ *alors* $(\mathrm{H},\cdot)$ *lui-même est un groupe (de même élément neutre que* G*). Ceci est souvent utilisé dans la pratique pour montrer qu'un ensemble est un groupe pour une loi, on essaie de montrer (quand c'est possible) que c'est un sous-groupe d'un groupe connu pour cette même loi.*

Théorème 14.4 (sous-groupes de $(\mathbb{Z},+)$)
H est un sous-groupe de $(\mathbb{Z},+)$ si et seulement si il existe un entier n tel que $\mathrm{H}=n\mathbb{Z}$ (ensemble des multiples entiers de n). L'entier n est unique au signe près et si $\mathrm{H}\neq\{0\}$, alors $n=\min \mathrm{H}^{*+}$.

Preuve : Il est facile de vérifier que pour $n\in\mathbb{Z}$, $n\mathbb{Z}$ est un sous-groupe de $\mathbb{Z}$. Si $\mathrm{H}=\{0\}$, alors on peut prendre $n=0$ et c'est le seul entier qui convienne. Si $\mathrm{H}\neq\{0\}$, posons, $n=\min \mathrm{H}^{*+}$ (n existe dans $\mathbb{N}$, c'est la propriété fondamentale de $\mathbb{N}$), on a $n\in\mathrm{H}$, comme H est un sous-groupe de $(\mathbb{Z},+)$, tout multiple de n est dans H, *i.e.* $n\mathbb{Z}\subset\mathrm{H}$. Soit $k\in\mathrm{H}$ effectuons la division euclidienne de k par n ($n\neq 0$) : $k=nq+r$ avec $0\leqslant r<n$. On a donc $r=k-nq\in\mathrm{H}^{+}$, si $r\neq 0$ alors $r\geqslant n$ ce qui est absurde, donc $r=0$ ce qui donne $k=nq\in n\mathbb{Z}$, finalement $\mathrm{H}=n\mathbb{Z}$. Si on a aussi $\mathrm{H}=m\mathbb{Z}$ avec $m\in\mathbb{N}$, alors $n\mathbb{Z}=m\mathbb{Z}$, donc n et m se divisent mutuellement dans $\mathbb{N}$, donc $n=m$. □

Théorème 14.5 (sous-groupes de $(\mathbb{R},+)$)
Soit G un sous-groupe de $(\mathbb{R},+)$ non réduit à $\{0\}$, alors soit G est dense dans $\mathbb{R}$, soit G est de la forme $\mathrm{G}=\alpha\mathbb{Z}$ avec $\alpha>0$ réel.

Preuve : Il existe $x\in\mathrm{G}$ non nul, on a donc x et $-x$ dans G, par conséquent G_+^* est non vide (et minoré par 0), soit $\alpha=\inf \mathrm{G}_+^*$.

- Si $\alpha\neq 0$, alors $\alpha>0$: supposons $\alpha\notin\mathrm{G}$, alors on peut trouver deux éléments de G, g_1 et g_2 tels que $\alpha<g_1<g_2<2\alpha$, mais alors $0<g_2-g_1<\alpha$ avec $g_2-g_1\in\mathrm{G}_+^*$: absurde, donc $\alpha\in\mathrm{G}$. On en déduit alors que $\alpha\mathbb{Z}\subset\mathrm{G}$. Soit $g\in\mathrm{G}$ et $n=\left\lfloor\frac{g}{\alpha}\right\rfloor$, alors $n\alpha\leqslant g<(n+1)\alpha$ et donc $0\leqslant g-n\alpha<\alpha$, on en déduit que $g-n\alpha$ est nul car c'est un élément positif de G, d'où $\mathrm{G}=\alpha\mathbb{Z}$.
- Si $\alpha=0$: soit $x\in\mathbb{R}$ et $\varepsilon>0$, il existe $g\in\mathrm{G}$ tel que $0<g<\varepsilon$, soit $n=\left\lfloor\frac{x}{g}\right\rfloor$ alors $ng\leqslant x<ng+g<ng+\varepsilon$, donc $|x-ng|<\varepsilon$ avec $ng\in\mathrm{G}$, donc G est dense dans $\mathbb{R}$.

□

★**Exercice 14.6** *Soient $a,b\in\mathbb{R}^*$, montrer que $\mathrm{G}=a\mathbb{Z}+b\mathbb{Z}$ est un sous-groupe de $(\mathbb{R},+)$ dense dans $\mathbb{R}$ si et seulement si $\frac{a}{b}$ est irrationnel. En déduire que $\{\cos(n)\mid n\in\mathbb{N}\}$ est dense dans $[-1;1]$.*

Théorème 14.6 (propriétés des sous-groupes)
• *Soit $\mathrm{H}\subset(\mathrm{G},*)$ non vide, alors H est un sous-groupe de $(\mathrm{G},*)$ si et seulement si $\forall x,y\in\mathrm{H}$, $x*y^{-1}\in\mathrm{H}$.*
• *Une intersection quelconque de sous-groupes de $(\mathrm{G},*)$, est un sous-groupe de $(\mathrm{G},*)$, mais ceci n'est pas vrai pour la réunion.*
• *Si $f\colon(\mathrm{G},\cdot)\to(\mathrm{G}',*)$ est un morphisme de groupes, alors :*
– *L'image directe d'un sous-groupe de $(\mathrm{G},\cdot)$ est un sous-groupe de $(\mathrm{G}',*)$.*
– *L'image réciproque d'un sous-groupe de $(\mathrm{G}',*)$ est un sous-groupe de $(\mathrm{G},\cdot)$*

Preuve : Le premier point est laissé en exercice.

Soit $(\mathrm{H}_i)_{i\in\mathrm{I}}$ une famille de sous-groupes de $(\mathrm{G},*)$, posons $\mathrm{K}=\bigcap\limits_{i\in\mathrm{I}}\mathrm{H}_i$, ce sont des parties de G donc l'intersection aussi : $\mathrm{K}\subset\mathrm{G}$. Tous les sous-groupes H_i contiennent l'élément neutre e de G, donc l'intersection aussi et par conséquent K est non vide. Si $x,y\in\mathrm{K}$, alors x,y sont dans tous les sous-groupes H_i, donc $x*y^{-1}$ aussi, par conséquent $x*y^{-1}$ est dans l'intersection, *i.e.* $x*y^{-1}\in\mathrm{K}$, K est donc un sous-groupe de $(\mathrm{G},*)$.

Donnons un contre-exemple pour la réunion : $2\mathbb{Z}\cup 3\mathbb{Z}$ n'est pas un sous-groupe de $(\mathbb{Z},+)$, car 2 et 3 sont dans la réunion, mais pas $2+3=5$, cet ensemble n'est donc pas stable pour l'addition (les autres conditions sont néanmoins remplies).

Soit $f\colon(\mathrm{G},\cdot)\to(\mathrm{G}',*)$ un morphisme de groupes, si A est un sous-groupe de G, soient x',y' dans $f(\mathrm{A})$, alors il existe x et y dans A tels que $f(x)=x'$ et $f(y)=y'$, on a $x'*y'^{-1}=f(x\cdot y^{-1})\in f(\mathrm{A})$ car $x\cdot y^{-1}\in\mathrm{A}$, donc $f(\mathrm{A})$ est un sous-groupe de $(\mathrm{G}',*)$.

Si B est un sous-groupe de $(\mathrm{G}',*)$, soient x et y dans $f^{-1}(\mathrm{B})$, alors $f(x\cdots y^{-1})=f(x)*f(y)^{-1}\in\mathrm{B}$, car $f(x)$, et $f(y)$ sont dans B, donc $x\cdot y^{-1}\in f^{-1}(\mathrm{B})$, donc $f^{-1}(\mathrm{B})$ est un sous-groupe de $(\mathrm{G},\cdot)$.

□

4) Noyau et image d'un morphisme de groupe

Définition 14.9

Soit $f\colon (G,\cdot) \to (G',*)$ *un morphisme de groupes.*
• *On appelle image de* f *l'ensemble des images des éléments de* G *par* f*, c'est à dire* $f(G)$*, que l'on note* $\mathrm{Im}(f)$*.*
• *On appelle noyau de* f *l'ensemble noté* $\ker(f)$ *et défini par :*

$$\ker(f) = \{x \in G \,/\, f(x) = e'\}.$$

C'est donc l'ensemble des antécédents de e' *(neutre de* G'*) par* f*, c'est donc une partie de* G *qui contient toujours* e *(car* $f(e) = e'$*).*

☞**Exemples :**
- $f\colon (\mathbb{Z},+) \to (\mathbb{U}_n,\times)$ définie par $f(k) = \exp(2ik\pi/n)$, est un morphisme de groupes, son noyau est $\ker(f) = n\mathbb{Z}$, son image est $\mathrm{Im}(f) = \mathbb{U}_n$.
- $\exp\colon (\mathbb{C},+) \to (\mathbb{C}^*,\times)$ est un morphisme de groupes, son noyau est $\ker(\exp) = 2i\pi\mathbb{Z}$, son image est $\mathrm{Im}(f) = \mathbb{C}^*$.
- $\exp\colon (\mathbb{R},+) \to (\mathbb{R}^*,\times)$ est un morphisme de groupes, son noyau est $\{0\}$, son image est $\mathbb{R}_+^*$.
- $f\colon (\mathbb{C}^*,\times) \to (\mathbb{C}^*,\times)$ définie par $f(z) = z^n$ est un morphisme de groupes, son noyau est $\ker(f) = \mathbb{U}_n$, son image est $\mathrm{Im}(f) = \mathbb{C}^*$.

Théorème 14.7 (propriétés)

Soit $f\colon (G,.) \to (G',*)$ *un morphisme de groupes, alors :*
• $\mathrm{Im}(f)$ *est un sous-groupe de* $(G',*)$*, et* f *est surjective si et seulement si* $\mathrm{Im}(f) = G'$*.*
• $\ker(f)$ *est un sous-groupe de* $(G,\cdot)$ *et* f *est injective si et seulement si* $\ker(f) = \{e\}$*.*

Preuve : Le premier point est évident (image directe d'un sous-groupe de g).
D'autre part, $\ker(f) = f^{-1}(\{e'\})$, c'est donc un sous-groupe de G (image réciproque d'un sous-groupe de G').
Si f est injective, soit $x \in \ker(f)$, alors $f(x) = e' = f(e)$, donc $x = e$ d'où $\ker(f) = \{e\}$.
Réciproquement, si $\ker(f) = \{e\}$, supposons $f(x) = f(y)$ avec $x, y \in G$, on a $f(x)*f(y)^{-1} = e'$, *i.e.* $f(x)*f(y^{-1}) = e'$ et donc $f(x\cdot y^{-1}) = e'$, ce qui signifie que $x\cdot y^{-1} \in \ker(f)$, mais alors $x\cdot y^{-1} = e$, d'où $x = y$, f est injective. □

Remarque 14.3 – *Si* $f\colon (G,\cdot) \to (G',*)$ *est un morphisme de groupes, alors* f *est un isomorphisme de groupes si et seulement si* $\mathrm{Im}(f) = G'$ *(surjectivité)* ***et*** $\ker(f) = \{e\}$ *(injectivité).*

III ANNEAUX ET CORPS

1) Anneaux

Définition 14.10

Un anneau est un ensemble A *muni de deux lois de composition internes : une addition et une multiplication, qui vérifient :*
- $(A,+)$ *est un groupe abélien.*
- *La multiplication :*
  - *est associative,*
  - *admet un élément neutre (noté 1).*
  - *est distributive sur l'addition.*

Si de plus la multiplication est commutative, on dit que $(A,+,\times)$ *est un anneau commutatif.*

☞**Exemples :**
- $(\mathbb{Z},+,\times)$ est un anneau commutatif mais ce n'est pas un corps.
- $(\mathcal{F}(\mathbb{N},\mathbb{C}),+,\times)$ est un anneau commutatif.
- Si E est un ensemble non vide, l'ensemble des fonctions de E dans $\mathbb{C}$ muni des opérations usuelles sur les fonctions, est un anneau commutatif, *i.e.* $(\mathcal{F}(E,\mathbb{C}),+,\times)$ est un anneau commutatif.
- Plus généralement, si $(A,+,\times)$ est un anneau et E est un ensemble non vide, alors $(\mathcal{F}(E,A),+,\times)$ est un anneau.

À retenir : Règles de calculs dans un anneau

Soit $(A, +, \times)$ un anneau :

- $\forall\, x \in A, x \times 0 = 0 \times x = 0$.
- $\forall\, x, y \in A, -(x \times y) = (-x) \times y = x \times (-y)$.
- $\forall\, x, y \in A$, si x et y sont inversibles (pour la multiplication), alors $x \times y$ est inversible est $(x \times y)^{-1} = y^{-1} \times x^{-1}$.
- Si $x \in A$ et $p \in \mathbb{Z}$ alors on note $p.x = 0$ si $p = 0$, $p.x = x + \cdots + x$, p fois si $p > 0$, et $p.x = (-x) + \cdots + (-x)$, $-p$ fois si $p < 0$. On note également $x^p = 1_A$ si $p = 0$, $x^p = x \times \cdots \times x$, p fois si $p > 0$, et $x^p = x^{-1} \times \cdots \times x^{-1}$, $-p$ fois si $p < 0$ **et** x **inversible**.
- $\forall\, x, y \in A$, si x et y commutent ($x \times y = y \times x$), alors on peut utiliser la formule du binôme, c'est à dire : $\forall\, n \in \mathbb{N}, (x+y)^n = \sum_{k=0}^{n} \binom{n}{k}.x^k \times y^{n-k} = \sum_{k=0}^{n} \binom{n}{k}.x^{n-k} \times y^k$.
- Si $x, y \in A$ avec $x \times y = y \times x$ alors $\forall n \in \mathbb{N}^*$: $x^n - y^n = (x-y)\sum_{k=0}^{n-1} a^k b^{n-1-k}$.

★**Exercice 14.7** *Si $x \in A$, simplifier $(1-x)\sum_{k=0}^{n} x^k$ et $(1+x)\sum_{k=0}^{n}(-1)^k x^k$.*

Théorème 14.8 (groupe des inversibles)

*Soit $(A, +, \times)$ un anneau, l'ensemble des inversibles de A est noté $U(A)$, cet ensemble est un **groupe multiplicatif**. $(U(A), \times)$ est appelé **groupe des unités** de A.*

Preuve : Celle-ci est simple et laissée en exercice. □

Exemples :

- $U(\mathbb{Z}) = \{\pm 1\}$.
- Si A est l'anneau des suites complexes, alors $U(A)$ est l'ensemble des suites complexes qui ne s'annulent pas.

Définition 14.11 (anneau intègre)

Soit $(A, +, \times)$ un anneau. On dit que A est un anneau intègre lorsque le produit de deux éléments non nuls est toujours non nul, sinon on dit que A est un anneau non intègre.

Remarque 14.4 –

- *Dans un anneau intègre, un produit de facteurs est nul si et seulement si au moins un des facteurs est nul.*
- *$(\mathbb{Z}, +, \times)$ est un anneau intègre.*
- *L'ensemble des suites complexes est un anneau non intègre.*

Définition 14.12 (morphisme d'anneaux)

Soit $f\colon (A, +, \times) \to (B, +, \times)$ une application d'un anneau A dans un anneau B, on dit que f est un morphisme d'anneaux lorsque :

- *f est un morphisme de groupes additifs, i.e. $\forall\, x, y \in A, f(x+y) = f(x) + f(y)$.*
- *$\forall\, x, y \in A, f(x \times y) = f(x) \times f(y)$.*
- *$f(1_A) = 1_B$.*

Si de plus f est bijective, alors on dit que f est un isomorphisme d'anneaux.

Exemples :

- On note $\mathbb{Z}[i] = \{a + ib \,/\, a, b \in \mathbb{Z}\}$, il est facile de voir que pour les opérations usuelles sur les complexes,$\mathbb{Z}[i]$ est un anneau intègre (anneau des entiers de *Gauss*), et que le groupe des inversibles est $U(\mathbb{Z}[i]) = \{\pm 1, \pm i\}$. La conjugaison est isomorphisme d'anneaux de $\mathbb{Z}[i]$ dans lui-même.
- Soit $(A, +, \times)$ un anneau, l'application $f\colon \mathbb{Z} \to A$ définie par $f(n) = n.1_A$ est un morphisme d'anneaux, et que $\mathrm{Im}(f) = G(1_A)$ le sous-groupe additif engendré par 1_A.

Théorème 14.9 (propriétés des morphismes d'anneaux)
• La composée de deux morphismes d'anneaux est un morphisme d'anneaux.
• Si $f: \mathrm{A} \to \mathrm{B}$ est un morphisme d'anneaux, alors pour tout x inversible dans A, $f(x)$ est inversible dans B et $f(x^{-1}) = (f(x))^{-1}$.
• La réciproque d'un isomorphisme d'anneaux f est un morphisme d'anneaux, et induit un isomorphisme de groupes entre les groupes des inversibles.

Preuve : Celle-ci est simple et laissée en exercice. □

Remarque 14.5 – *Il découle du deuxième point que $\tilde{f}: \mathrm{U(A)} \to \mathrm{U(B)}$ définie par $\tilde{f}(x) = f(x)$, est un morphisme de groupes multiplicatifs.*

Définition 14.13 (sous-anneaux d'un anneau)
Soit $(\mathrm{A}, +, \times)$ un anneau, et soit H un ensemble, on dit que H est un sous-anneau de A lorsque :
– $\mathrm{H} \subset \mathrm{A}$.
– $1_\mathrm{A} \in \mathrm{H}$.
– $\forall\, x, y \in \mathrm{H}, x + y \in \mathrm{H}, x \times y \in \mathrm{H}$ *et* $-x \in \mathrm{H}$.
Si c'est le cas, alors $(\mathrm{H}, +, \times)$ est lui-même un anneau.

☞**Exemple** : $\mathbb{Z}[i]$ est un sous-anneau de $(\mathbb{C}, +, \times)$.

Théorème 14.10
Une intersection de sous-anneaux de $(\mathrm{A}, +, \times)$ est un sous-anneau de A.

Preuve : Celle-ci est simple et laissée en exercice. □

2) Corps

Définition 14.14
Un corps est un ensemble E muni de deux opérations (ou deux lois de composition), une addition et une multiplication. Ces deux opérations doivent vérifier les propriétés suivantes :
– *$(\mathrm{E}, +, \times)$ est un anneau.*
– *$\mathrm{U(E)} = \mathrm{E} \setminus \{0\}$, i.e. : $\forall\, x \in \mathrm{E} \setminus \{0\}, x$ a un inverse dans E. Si de plus la multiplication est commutative, on dit que $(\mathrm{E}, +, \times)$ est un corps commutatif.*

☞**Exemples** :
– $(\mathbb{R}, +, \times), (\mathbb{Q}, +, \times), (\mathbb{C}, +, \times)$ sont des corps commutatifs, mais $(\mathbb{Z}, +, \times)$ n'est pas un corps.
– Il existe des corps non commutatifs (corps des quaternions).

Remarque 14.6 –
– *Un corps est toujours intègre.*
– *Les règles de calculs sont les mêmes que dans un anneau.*

Définition 14.15 (sous-corps d'un corps)
Soit $(\mathrm{K}, +, \times)$ un corps et soit H un ensemble, on dit que H est un sous-corps de K lorsque :
– $\mathrm{H} \subset \mathrm{K}$.
– $1_\mathrm{K} \in \mathrm{H}$.
– $\forall\, x, y \in \mathrm{H}, x + y \in \mathrm{H}, -x \in \mathrm{H}$ *et* $x \times y \in \mathrm{H}$.
– $\forall\, x \in \mathrm{H} \setminus \{0\}, x^{-1} \in \mathrm{H}$.
Si c'est le cas alors $(\mathrm{H}, +, \times)$ est lui-même un corps.

☞**Exemples** :
– $\mathbb{Q}$ est un sous-corps de $\mathbb{R}$ qui est lui-même un sous-corps de $\mathbb{C}$.
– $\mathbb{Q}[i] = \{a + ib \;/\; a, b \in \mathbb{Q}\}$ est un sous-corps de $(\mathbb{C}, +, \times)$.
– $\mathbb{Q}[\sqrt{2}] = \{a + b\sqrt{2} \;/\; a, b \in \mathbb{Q}\}$ est un sous-corps de $\mathbb{R}$.

IV SOLUTION DES EXERCICES

Solution 14.1

Solution 14.2

Solution 14.3

Solution 14.4

Solution 14.5 *Produit cartésien de n groupes tous égaux à* $(\mathbb{K}, +)$.

Solution 14.6

Solution 14.7

Chapitre 15

Polynômes

Sommaire

Dans tout ce chapitre, $\mathbb{K}$ désigne un corps inclus dans $\mathbb{C}$ (par exemple, $\mathbb{K}=\mathbb{Q}$ ou $\mathbb{K}=\mathbb{R}$ ou $\mathbb{K}=\mathbb{C}$).

I ENSEMBLE DES POLYNÔMES

1) Définition

Définition 15.1

Une application $f\colon \mathbb{K}\to\mathbb{K}$ est dite polynomiale si et seulement si il existe $n\in\mathbb{N}$, et des nombres $a_0,\dots,a_n$ dans $\mathbb{K}$ tels que $\forall x\in\mathbb{K}$, $f(x)=a_0+a_1x+\cdots+a_nx^n$, i.e. $f(x)=\sum\limits_{k=0}^{n}a_kx^k$.

Notation : on note $X=\mathrm{id}_{\mathbb{K}}$ l'application identité de $\mathbb{K}$ (*i.e.* $\forall x\in\mathbb{K}, X(x)=x$). On sait que $(\mathcal{F}(\mathbb{K},\mathbb{K}),+,\times)$ est un anneau commutatif pour les opérations usuelles sur les fonctions ; en utilisant les notations habituelles dans un anneau, on a pour $k\in\mathbb{N}$, $X^k\colon \begin{array}{ccc}\mathbb{K} & \to & \mathbb{K}\\ x & \mapsto & x^k\end{array}$. Par conséquent une application $f\colon\mathbb{K}\to\mathbb{K}$ est polynomiale si et seulement si il existe $n\in\mathbb{N}$ et des nombres $a_0,\dots,a_n$ dans $\mathbb{K}$ tels que $f=a_0X^0+a_1X+\cdots+a_nX^n=\sum\limits_{k=0}^{n}a_kX^k$. Par abus écriture, on écrira $a_0X^0=a_0$, ce qui revient à identifier tout nombre a_0 avec la fonction constante $a_0X^0\colon \begin{array}{ccc}\mathbb{K} & \to & \mathbb{K}\\ x & \mapsto & a_0\times 1=a_0\end{array}$.

Définition 15.2

On appelle polynôme à coefficients dans $\mathbb{K}$ toute application polynomiale $P\colon\mathbb{K}\to\mathbb{K}$, c'est à dire toute application de la forme $P=a_0+a_1X+\cdots+a_nX^n$, où $n\in\mathbb{N}$, $a_0,\dots,a_n\in\mathbb{K}$, le polynôme X est appelé ***indéterminée****. Les a_i sont appelés* ***coefficients*** *du polynôme P. Si tous* *les coefficients sont nuls, P est l'application nulle, on dit que P est le polynôme nul . Si tous les coefficients sont nuls sauf un, le polynôme est appelé monôme. Si tous les coefficients sont nuls à partir de l'indice 1, on dit que le polynôme est constant ($\mathrm{P} = a_0$). L'ensemble des polynômes à coefficients dans $\mathbb{K}$ est noté $\mathbb{K}[\mathrm{X}]$.*

Remarque 15.1 :

- *Il est commode de noter les polynômes sous la forme $\sum\limits_{k\in\mathbb{N}} a_k\mathrm{X}^k$ avec la convention que* ***les coefficients a_k sont tous nuls à partir d'un certain rang****. Ainsi on peut dire que les coefficients d'un polynôme forment une suite (a_k) d'éléments de $\mathbb{K}$, nulle à partir d'un certain rang.*
- *Avec l'identification faite plus haut, tout nombre est un polynôme constant, donc $\mathbb{K} \subset \mathbb{K}[\mathrm{X}]$.*

Théorème 15.1 (Égalité de deux polynômes)

Deux polynômes sont égaux si et seulement si ils ont les mêmes coefficients :

$$\sum_k a_k\mathrm{X}^k = \sum_k b_k\mathrm{X}^k \iff \forall k \in \mathbb{N},\ a_k = b_k.$$

Preuve : Par récurrence : en évaluant en 0 on vérifie que $a_0 = b_0$. On suppose que pour un entier n on a $a_k = b_k$ pour $k \in [\![0\,;n]\!]$, après simplification on a $\sum\limits_{k\geqslant n+1} a_k\mathrm{X}^k = \sum\limits_{k\geqslant n+1} b_k\mathrm{X}^k$, donc pour tout nombre x on a $\sum\limits_{k\geqslant n+1} a_kx^k = \sum\limits_{k\geqslant n+1} b_kx^k$, pour $x \neq 0$ on peut simplifier par x^{n+1}, puis on fait tendre x vers 0 ce qui donne $a_{n+1} = b_{n+1}$. □

2) Structure de $\mathbb{K}[\mathrm{X}]$

Théorème 15.2

L'ensemble $\mathbb{K}[\mathrm{X}]$ est un sous-anneau de $(\mathcal{F}(\mathbb{K},\mathbb{K}), +, \times)$, et donc :

$$(\mathbb{K}[\mathrm{X}], +, \times) \text{ est un anneau commutatif.}$$

Preuve : Par définition on a l'inclusion $\mathbb{K}[\mathrm{X}] \subset \mathcal{F}(\mathbb{K},\mathbb{K})$. Les deux éléments neutres (application nulle et application constante $x \mapsto 1$) sont des polynômes (constants).

Soient $\mathrm{P} = \sum\limits_{k\in\mathbb{N}} a_k\mathrm{X}^k$ et $\mathrm{Q} = \sum\limits_{k\in\mathbb{N}} b_k\mathrm{X}^k$ deux polynômes. Il existe deux entiers N et N′ tels que $k \geqslant \mathrm{N} \implies a_k = 0$, et $k \geqslant \mathrm{N}' \implies b_k = 0$.

Soit $k \geqslant \max(\mathrm{N},\mathrm{N}')$, alors $a_k + b_k = 0$. La somme des deux polynômes est la somme des fonctions $\mathrm{P}+\mathrm{Q}$, les propriétés des opérations dans $(\mathcal{F}(\mathbb{K},\mathbb{K}), +, \times)$ permettent d'écrire $\mathrm{P}+\mathrm{Q} = \sum\limits_{k\in\mathbb{N}} (a_k + b_k)\mathrm{X}^k$ (la somme est bien finie), ce qui est un polynôme par définition, donc $\mathrm{P}+\mathrm{Q} \in \mathbb{K}[\mathrm{X}]$.

L'opposé de la fonction P est la fonction $-\mathrm{P} = \sum\limits_k -a_k\mathrm{X}^k$ qui est bien un polynôme.

En écrivant $\mathrm{P} = \sum\limits_{i=0}^{\mathrm{N}} a_i\mathrm{X}^i$ et $\mathrm{Q} = \sum\limits_{j=0}^{\mathrm{N}'} b_j\mathrm{X}^j$, le produit des polynômes est le produit des fonctions $\mathrm{P}\times\mathrm{Q}$, les propriétés des opérations dans $(\mathcal{F}(\mathbb{K},\mathbb{K}), +, \times)$ permettent d'écrire $\mathrm{P}\times\mathrm{Q} = \sum\limits_{i=0}^{\mathrm{N}}\sum\limits_{j=0}^{\mathrm{N}'} a_ib_j\mathrm{X}^{i+j}$, on reconnaît une somme de polynômes, donc $\mathrm{P}\times\mathrm{Q}$ est bien un polynôme, *i.e.* $\mathrm{P}\times\mathrm{Q} \in \mathbb{K}[\mathrm{X}]$.

□

Théorème 15.3 (Coefficients du produit)

Si $\mathrm{P} = \sum\limits_{k\in\mathbb{N}} a_k\mathrm{X}^k$ et $\mathrm{Q} = \sum\limits_{k\in\mathbb{N}} b_k\mathrm{X}^k$ sont deux polynômes, alors :

$$\mathrm{P}\times\mathrm{Q} = \sum_{n\in\mathbb{N}} c_n\mathrm{X}^n \text{ avec } c_n = \sum_{k=0}^{n} a_kb_{n-k}.$$

Preuve : En reprenant le calcul précédent, $\mathrm{P}\times\mathrm{Q} = \sum\limits_{i\in\mathbb{N}}\sum\limits_{j\in\mathbb{N}} a_ib_j\mathrm{X}^{i+j}$ (sommes finies), pour $n \in \mathbb{N}$, on regroupe les termes tels que $i + j = n$, ce qui correspond aux indices $i \in [\![0\,;n]\!]$ et $j = n - i$, le coefficient de X^n est donc bien $\sum\limits_{i=0}^{n} a_ib_{n-i}$. □

Remarque 15.2 :

- *Avec les notations précédentes, si $a_k = 0$ pour $k \geqslant \mathrm{N}$ et $b_k = 0$ pour $k \geqslant \mathrm{N}'$, pour $n \in \mathbb{N}$, on pose $c_n = \sum\limits_{k=0}^{n} a_kb_{n-k}$ (coefficient de X^n dans $\mathrm{P}\times\mathrm{Q}$), si $n \geqslant \mathrm{N}+\mathrm{N}'-1$, alors il est facile de voir que pour toute valeur de k dans $[\![0\,;n]\!]$, le produit a_kb_{n-k} est nul, et donc c_n est nul dès que $n \geqslant \mathrm{N}+\mathrm{N}'-1$.*
- *On a aussi $c_n = \sum\limits_{k=0}^{n} a_{n-k}b_k = \sum\limits_{p+q=n} a_pb_q$.*

Produit par un scalaire.

Si $P = \sum_k a_k X^k$ est un polynôme et $\lambda \in \mathbb{K}$, alors on sait par définition que le produit de P par le scalaire λ est la fonction $\lambda.P : x \mapsto \lambda \times \sum_k a_k x^k = \sum_k \lambda a_k x^k$, donc $\lambda.P$ est le polynôme $\sum_{k \in \mathbb{N}} \lambda a_k X^k$.

Conséquence : $(\mathbb{K}[X], +, .)$ est un $\mathbb{K}$-espace vectoriel (nous verrons en algèbre linéaire que c'est un sous-espace vectoriel de $(\mathcal{F}(\mathbb{K}, \mathbb{K}), +, .)$.

On vérifie également : $\forall P, Q \in \mathbb{K}[X], \forall \lambda \in \mathbb{K},\ \lambda.(P \times Q) = (\lambda.P) \times Q = P \times (\lambda.Q)$.

À retenir

$$\sum_{n \in \mathbb{N}} a_n X^n = \sum_{n \in \mathbb{N}} b_n X^n \iff \forall\, n \in \mathbb{N}, a_n = b_n.$$

$$\left(\sum_{n \in \mathbb{N}} a_n X^n\right) + \left(\sum_{n \in \mathbb{N}} b_n X^n\right) = \sum_{n \in \mathbb{N}} (a_n + b_n) X^n.$$

$$\left(\sum_{n \in \mathbb{N}} a_n X^n\right) \times \left(\sum_{n \in \mathbb{N}} b_n X^n\right) = \sum_{n \in \mathbb{N}} \left(\sum_{p+q=n} a_p b_q\right) X^n.$$

$$\sum_{n \in \mathbb{N}} a_n X^n \in \mathbb{K} \iff \forall\, n \geqslant 1, a_n = 0.$$

II DIVISION EUCLIDIENNE

1) Degré d'un polynôme

Soit $P \in \mathbb{K}[X]$, si $P = 0$ alors tous les coefficients de P sont nuls, si $P \neq 0$, alors l'ensemble des indices des coefficients non nuls de P n'est pas vide, et il est majoré (les coefficients sont nuls à partir d'un certain rang), donc cet ensemble admet un plus grand élément.

Définition 15.3

Soit $P \in \mathbb{K}[X]$*, si* $P = 0$ *alors on pose* $\deg(P) = -\infty$*, sinon on pose* $\deg(P) = \max\{k \in \mathbb{N} \ / \ a_k \neq 0\}$*.*
Si P *est non nul de degré* n*, alors le coefficient* a_n *est appelé coefficient dominant de* P*, noté* $\mathrm{dom}(P)$*, si ce coefficient vaut 1, alors on dit que le polynôme* P *est unitaire (ou normalisé).*

Remarque 15.3 – *Caractérisations du polynôme nul et des polynômes constants non nuls*

- $P = 0 \iff \deg(P) = -\infty$.
- $P \in \mathbb{K}^* \iff \deg(P) = 0$.

Théorème 15.4

Soient $P, Q \in \mathbb{K}[X]$, $\deg(P + Q) \leqslant \max(\deg(P), \deg(Q))$, $\deg(P \times Q) = \deg(P) + \deg(Q)$*, et si les polynômes sont non nuls, alors* $\mathrm{dom}(P \times Q) = \mathrm{dom}(P) \times \mathrm{dom}(Q)$.

Preuve : Si l'un des deux polynômes est nul, alors le théorème est évident.

Supposons les deux polynômes non nuls : $P = \sum_n a_n X^n$ et $Q = \sum_n b_n X^n$, si $a_n + b_n \neq 0$ alors $a_n \neq 0$ ou $b_n \neq 0$, donc $n \leqslant \deg(P)$ ou $n \leqslant \deg(Q)$ *i.e.* $n \leqslant \max(\deg(P), \deg(Q))$, ce qui prouve le premier résultat.

$P \times Q = \sum_n c_n X^n$ où $c_n = \sum_{p+q=n} a_p b_q$. Posons $N = \deg(P)$ et $N' = \deg(Q)$, il est clair que $c_{N+N'} = a_N b_{N'} \neq 0$, d'autre part si $n > N + N'$, alors si $p + q = n$ on a $p > N$ ou $q > N'$ donc $a_p b_q = 0$ ce qui entraîne $c_n = 0$. Par conséquent, $\deg(P \times Q) = N + N' = \deg(P) + \deg(Q)$. □

Remarque 15.4 – *Lorsque* P *et* Q *ont des degrés distincts, ou bien lorsque* P *et* Q *ont même degré mais des coefficients dominants non opposés, alors* $\deg(P + Q) = \max(\deg(P), \deg(Q))$.

Théorème 15.5

L'anneau $(\mathbb{K}[X], +, \times)$ *est un anneau intègre, et seuls les polynômes constants non nuls ont un inverse dans* $\mathbb{K}[X]$.

Preuve : Si P et Q sont deux polynômes non nuls, alors $\deg(P \times Q) = \deg(P) + \deg(Q) \in \mathbb{N}$, donc $P \times Q \neq 0$, ce qui prouve que $\mathbb{K}[X]$ est intègre.

Si P est inversible dans $\mathbb{K}[X]$, alors il existe un polynôme Q tel que $P \times Q = 1$, d'où $\deg(P) + \deg(Q) = 0$, ce qui entraîne $\deg(P) = \deg(Q) = 0$ et donc $P \in \mathbb{K}^*$. La réciproque est évidente. □

Notation : Soit $n \in \mathbb{N}$, on note $\mathbb{K}_n[X]$ l'ensemble des polynômes de degré inférieur ou égal à n :

$$\mathbb{K}_n[X] = \{P \in \mathbb{K}[X] \ / \ \deg(P) \leqslant n\}.$$

2) Algorithme de la division euclidienne

Théorème 15.6 (de la division euclidienne)
Soient A *et* B *deux polynômes avec* $\mathrm{B} \neq 0$*, alors il existe deux polynômes* Q *et* R ***uniques*** *tels que :*

$$\mathrm{A} = \mathrm{B} \times \mathrm{Q} + \mathrm{R} \text{ avec } \deg(\mathrm{R}) < \deg(\mathrm{B})$$

Preuve : Pour l'existence : on examine deux cas particuliers, si $\mathrm{A} = 0$ alors $\mathrm{A} = \mathrm{B} \times 0 + 0$, si $\mathrm{B} = \alpha \in \mathbb{K}^*$ alors $\mathrm{A} = \mathrm{B} \times (\frac{1}{\alpha}\mathrm{A}) + 0$. On suppose maintenant que A est non nul, et que B est non constant. Pour le cas général, on fait une récurrence sur le degré de A.

Si $\deg(\mathrm{A}) < \deg(\mathrm{B}) = d \in \mathbb{N}^*$, alors on peut prendre $\mathrm{Q} = 0$ et $\mathrm{R} = \mathrm{A}$. Supposons maintenant l'existence démontrée pour $\deg(\mathrm{A}) \leqslant n$ avec $n \geqslant d - 1$ (l'existence est établie au rang $n = d - 1$), et soit A de degré $n + 1$ (on a donc $n + 1 \geqslant d$), notons a_{n+1} son coefficient dominant, soit $\mathrm{Q}' = \frac{a_{n+1}}{b_d}\mathrm{X}^{n+1-d}$, alors $\deg(\mathrm{B} \times \mathrm{Q}') = n + 1$ et le coefficient dominant de $\mathrm{B} \times \mathrm{Q}'$ est a_{n+1}, donc $\deg(\mathrm{A} - \mathrm{B} \times \mathrm{Q}') \leqslant n$, d'après l'hypothèse de récurrence, il existe deux polynômes Q'' et R tels que $\mathrm{A} - \mathrm{B} \times \mathrm{Q}' = \mathrm{B} \times \mathrm{Q} + \mathrm{R}$ avec $\deg(\mathrm{R}) < \deg(\mathrm{B})$, mais alors $\mathrm{A} = \mathrm{B} \times (\mathrm{Q}' + \mathrm{Q}'') + \mathrm{R}$, ce qui prouve l'existence au rang $n + 1$.

Pour l'unicité : supposons que $\mathrm{A} = \mathrm{B} \times \mathrm{Q} + \mathrm{R} = \mathrm{B} \times \mathrm{Q}' + \mathrm{R}'$ avec $\deg(\mathrm{R}) < \deg(\mathrm{B})$ et $\deg(\mathrm{R}') < \deg(\mathrm{B})$, alors $\mathrm{B} \times (\mathrm{Q} - \mathrm{Q}') = \mathrm{R}' - \mathrm{R}$, d'où $\deg(\mathrm{B}) + \deg(\mathrm{Q} - \mathrm{Q}') = \deg(\mathrm{R}' - \mathrm{R}) < \deg(\mathrm{B})$, comme $\deg(\mathrm{B}) \geqslant 0$, on a nécessairement $\deg(\mathrm{Q} - \mathrm{Q}') = -\infty = \deg(\mathrm{R}' - \mathrm{R})$, et donc $\mathrm{Q} = \mathrm{Q}', \mathrm{R} = \mathrm{R}'$. □

Remarque 15.5 – *La démonstration est constructive, en ce sens qu'elle donne un algorithme de calcul du quotient* (Q) *et du reste* (R).

☞**Exemple** : Avec $\mathrm{A} = \mathrm{X}^4 + a\mathrm{X}^2 + b\mathrm{X} + c$ et $\mathrm{B} = \mathrm{X}^2 + \mathrm{X} + 1$, on obtient le quotient $\mathrm{Q} = \mathrm{X}^2 - \mathrm{X} + a$ et le reste $\mathrm{R} = (b - a + 1)\mathrm{X} + c - a$. On peut vérifier que $\mathrm{A} = \mathrm{B} \times (\mathrm{X}^2 - \mathrm{X} + a) + (b - a + 1)\mathrm{X} + c - a$.

3) Divisibilité

Définition 15.4
Soient $\mathrm{A}, \mathrm{B} \in \mathbb{K}[\mathrm{X}]$*, on dit que* B *divise* A *lorsqu'il existe un polynôme* Q *tel que* $\mathrm{A} = \mathrm{Q} \times \mathrm{B}$*, notation* $\mathrm{B}|\mathrm{A}$.

Remarque 15.6 – *On définit ainsi une relation dans* $\mathbb{K}[\mathrm{X}]$*, on peut vérifier que celle - ci est réflexive, transitive, mais elle n'est ni symétrique, ni antisymétrique. Plus précisément,* $\mathrm{B}|\mathrm{A}$ *et* $\mathrm{A}|\mathrm{B}$ *ssi il existe* $\lambda \in \mathbb{K}^*$ *tel que* $\mathrm{A} = \lambda\mathrm{B}$ *(on dit que* A *et* B *sont* ***associés****).*

Théorème 15.7
- *Si* $\mathrm{B} \neq 0$*, alors* $\mathrm{B}|\mathrm{A}$ *si et seulement si le reste de la division euclidienne de* A *par* B *est nul.*
- *Si* $\mathrm{A} \neq 0$ *et* $\mathrm{B}|\mathrm{A}$*, alors* $\deg(\mathrm{B}) \leqslant \deg(\mathrm{A})$.
- *Si* $\mathrm{B}|\mathrm{A}$ *et* $\mathrm{B}|\mathrm{C}$*, alors* $\forall\, \mathrm{U}, \mathrm{V} \in \mathbb{K}[\mathrm{X}], \mathrm{B}|\mathrm{A} \times \mathrm{U} + \mathrm{C} \times \mathrm{V}$.

Preuve : Celle-ci est simple et laissée en exercice. □

Remarque 15.7 – *Il découle du dernier point que si* $\mathrm{B}|\mathrm{A} - \mathrm{C}$ *et* $\mathrm{B}|\mathrm{D} - \mathrm{E}$*, alors* $\mathrm{B}|(\mathrm{A} + \mathrm{D}) - (\mathrm{C} + \mathrm{E})$ *et* $\mathrm{B}|\mathrm{AD} - \mathrm{EC}$*, en particulier, si* $\mathrm{B}|\mathrm{A} - \mathrm{C}$ *alors* $\forall\, n \in \mathbb{N}, \mathrm{B}|\mathrm{A}^n - \mathrm{C}^n$.

III RACINES

Dans la définition choisie, les polynômes de $\mathbb{K}[\mathrm{X}]$ sont les fonctions polynomiales de $\mathbb{K}$ dans $\mathbb{K}$, donc si $x \in \mathbb{K}$ et $\mathrm{P} = \sum\limits_k a_k\mathrm{X}^k \in \mathbb{K}[\mathrm{X}]$ (somme finie), alors $\mathrm{P}(x) = \sum\limits_k a_k\mathrm{X}^k(x) = \sum\limits_k a_k x^k$ est l'image de x par P, on dit aussi qu'on a évalué le polynôme P en x.

1) Racines d'un polynôme

Définition 15.5
Soit $\mathrm{P} = \sum\limits_{k=0}^{n} a_k\mathrm{X}^k \in \mathbb{K}[\mathrm{X}]$*, on appelle racine de* P *dans* $\mathbb{K}$ *tout nombre* $\alpha \in \mathbb{K}$ *tel que* $\mathrm{P}(\alpha) = 0$*,*

c'est à dire toute solution dans $\mathbb{K}$ à l'équation $\sum_{k=0}^{n} a_k x^k = 0$.

Théorème 15.8

Soit $P \in \mathbb{K}[X]$:
- *Soit $a \in \mathbb{K}$, a est racine de P ssi $X - a | P$.*
- *Si $\deg(P) \leqslant n$ et si P admet au moins $(n+1)$ racines dans $\mathbb{K}$, alors $P = 0$.*

Preuve : Soit $a \in \mathbb{K}$, on effectue la division euclidienne de P par $X - a$: $P = Q \times (X - a) + R$ avec $\deg(R) < 1$, donc R est un polynôme constant $R = \lambda$, finalement $P = Q \times (X - a) + \lambda$. Substituons a à X : $P(a) = Q(a) \times (a - a) + \lambda$, c'est à dire : $\lambda = P(a)$, ce qui prouve la première assertion.

La deuxième assertion se démontre par récurrence sur n : pour $n = 0$, l'hypothèse dit que P est une constante et que P a au moins une racine, donc cette constante est nulle, *i.e.* $P = 0$. Supposons le résultat démontré au rang n, et soit $\deg(P) \leqslant n + 1$ avec P ayant au moins $n + 2$ racines, soit a l'une d'elles, alors il existe $Q \in \mathbb{K}[X]$, tel que $P = Q \times (X - a)$, mais alors $\deg(Q) \leqslant n$ et Q a au moins $n + 1$ racines dans $\mathbb{K}$, donc $Q = 0$ (HR) et par conséquent, $P = 0$. □

Conséquences :

a) Si $a_1, \ldots, a_n$ sont des racines distinctes de P alors $(X - a_1) \cdots (X - a_n) | P$.

b) Si P est non nul de degré n, alors P admet au plus n racines distinctes.

★**Exercice 15.1** *Soit P un polynôme de degré 2, on pose :*

$$Q = (1 - X^2)P(0) + \frac{X(X-1)}{2}P(-1) + \frac{X(X+1)}{2}P(1)$$

Montrer que $P = Q$.

Remarque 15.8 – *Pour montrer qu'un polynôme P est nul on dispose de trois méthodes :*
- *Montrer que tous les coefficients de P sont nuls.*
- *Montrer que le degré de P est $-\infty$.*
- *Montrer que P a une infinité de racines.*

Soit P un polynôme non nul et soit $a \in \mathbb{K}$, on sait que que si $(X - a)^k | P$ alors $k \leqslant \deg(P)$ (car $P \neq 0$). Par conséquent l'ensemble $\{k \in \mathbb{N} \ / \ (X - a)^k | P\}$ est un ensemble non vide (contient 0) et majoré par $\deg(P)$, comme c'est une partie de $\mathbb{N}$, cet ensemble admet un plus grand élément.

Définition 15.6 (multiplicité d'un nombre dans un polynôme)

Soit $P \in \mathbb{K}[X]$ un polynôme non nul et soit $a \in \mathbb{K}$, on appelle multiplicité de a dans P le plus grand des entiers k tels que $(X - a)^k | P$. Notation : $m_P(a) = \max\left\{k \in \mathbb{N} \ / \ (X - a)^k \mid P\right\}$. Un nombre de multiplicité 1 est appelée une racine simple, un nombre de multiplicité 2 est appelée une racine double...etc

Remarque 15.9 –
- *a est racine de P équivaut à $m_P(a) \geqslant 1$.*
- *Il est facile de vérifier que si $q \in \{k \in \mathbb{N} \ / \ (X - a)^k | P\}$, alors tout entier inférieur ou égal à q est également dans l'ensemble, cela signifie que l'ensemble $\{k \in \mathbb{N} \ / \ (X - a)^k | P\}$ est un intervalle d'entiers, on peut donc énoncer : $m = m_P(a) \iff (X - a)^m$ divise P et $(X - a)^{m+1}$ ne divise pas P.*

★**Exercice 15.2** *Calculer la multiplicité de 1 dans les polynômes $P = X^3 - 3X^2 + 2$ et $Q = X^3 - 4X^2 + 5X - 2$.*

Théorème 15.9

Soit P un polynôme non nul, soit $a \in \mathbb{K}$, et soit $m \in \mathbb{N}$, on a alors :

$$m = m_P(a) \iff \exists\, Q \in \mathbb{K}[X], P = (X - a)^m \times Q \text{ et } Q(a) \neq 0.$$

Preuve : Si on a $P = (X - a)^m \times Q$ et $Q(a) \neq 0$, alors $m_P(a) \geqslant m$, mais si $(X - a)^{m+1} | P$, il est facile de voir que $X - a | Q$ ce qui est absurde, donc $m_P(a) = m$.

Réciproquement, si $m = m_P(a)$, alors il existe Q tel que $P = (X - a)^m \times Q$, si $Q(a) = 0$ alors $X - a | Q$ et donc $(X - a)^{m+1} | P$ ce qui contradictoire, donc $Q(a) \neq 0$. □

Théorème 15.10

Soient $P, Q \in \mathbb{K}[X]$, *non nuls, et* $a \in \mathbb{K}$

a) $m_{P\times Q}(a) = m_P(a) + m_Q(a)$.

b) *si* $P + Q \neq 0$, *alors* $m_{P+Q}(a) \geqslant \min(m_P(a); m_Q(a))$.

Preuve : Celle-ci est simple et laissée en exercice. □

2) Corps algébriquement clos

Soit P un polynôme non nul ayant des racines dans $\mathbb{K}$, soient $a_1, \ldots, a_n$ **toutes** les racines distinctes de P de multiplicités respectives : $m_1, \ldots, m_n$. D'après ce qui précède il existe un polynôme Q tel que $P = (X - a_1)^{m_1} \times Q$ avec $Q(a_1) \neq 0$, comme $a_2 \neq a_1$ on peut affirmer que a_2 est racine de Q : $Q = (X - a_2)^m \times T$ avec $T(a_2) \neq 0$, mais alors $P = (X - a_2)^m \times (X - a_1)^{m_1} \times T$, on en déduit que $m = m_2$, par conséquent on a $P = (X - a_1)^{m_1}(X - a_2)^{m_2} \times T$ avec a_1 et a_2 qui ne sont pas racines de T. De proche en proche (récurrence sur n) on a arrive à : il existe un polynôme S tel que $P = (X - a_1)^{m_1} \cdots (X - a_n)^{m_n} \times S$, avec $a_1, \ldots, a_n$ qui ne sont pas racines de S, mais comme P n'a pas d'autres racines on peut en déduire que S est **sans racine** dans $\mathbb{K}$.

Théorème 15.11 (factorisation d'un polynôme connaissant toutes ses racines)

Si $a_1, \ldots, a_n$ *sont les racines distinctes de* P *de multiplicités respectives* $m_1, \ldots, m_n$, *alors il existe un polynôme* Q *sans racine dans* $\mathbb{K}$ *tel que* $P = Q \times \prod_{k=1}^{n} (X - a_k)^{m_k}$.

Définition 15.7 (polynôme scindé)

Si $a_1, \ldots, a_n$ *sont les racines distinctes de* P *de multiplicités respectives* $m_1, \ldots, m_n$, *alors d'après le théorème précédent :* $\sum_{k=1}^{n} m_k \leqslant \deg(P)$. *La quantité* $\sum_{k=1}^{n} m_k$ *(somme des multiplicités des racines) est appelée* ***nombre de racines de*** P ***comptées avec leur multiplicité***. *On dira que le polynôme* P *est scindé sur* $\mathbb{K}$ *lorsque cette quantité est égale au degré de* P, *on dit aussi que* P *admet toutes ses racines dans* $\mathbb{K}$ *(toutes : signifie que le nombre de racines comptées avec leur multiplicité, est égal au degré)*

En reprenant la factorisation précédente : $P = Q\times \prod_{k=1}^{n} (X-a_k)^{m_k}$, on voit que lorsque P est scindé, alors $\deg(Q) = 0$, le polynôme Q est donc une constante non nulle, en comparant les coefficients dominants de chaque coté, on voit que Q est égal au coefficient dominant de P, d'où l'énoncé :

Théorème 15.12

Si P *est scindé et si* $a_1, \ldots, a_n$ *sont les racines distinctes de* P *de multiplicités respectives* $m_1, \ldots, m_n$, *alors* $P = \lambda \prod_{k=1}^{n} (X - a_k)^{m_k}$, *où* λ *est le coefficient dominant de* P.

☞**Exemples** :
- $X^2 - 2$ est scindé sur $\mathbb{R}$, mais pas sur $\mathbb{Q}$.
- $X^2 + 1$ est scindé sur $\mathbb{C}$, mais pas sur $\mathbb{R}$.

Définition 15.8

On dit que le corps $\mathbb{K}$ *est algébriquement clos lorsque tout polynôme non constant de* $\mathbb{K}[X]$ *admet au moins une racine dans* $\mathbb{K}$.

Remarque 15.10 – *D'après les exemples précédents, les corps* $\mathbb{Q}$ *et* $\mathbb{R}$ *ne sont pas algébriquement clos.*

Théorème 15.13

Si $\mathbb{K}$ est un corps algébriquement clos, alors tout polynôme non constant de $\mathbb{K}[X]$ est scindé sur $\mathbb{K}$.

Preuve : On montre par récurrence sur n que si $\deg(P) = n$ alors P admet n racines dans $\mathbb{K}$. Pour $n = 1$, $P = aX + b = a(X + b/a)$, une racine $-b/a$. Supposons le résultat démontré au rang n, et soit P de degré $n+1$: P est non constant, donc P admet au moins une racine a, d'où $P = (X - a) \times Q$, mais $\deg(Q) = n$, il suffit alors d'appliquer l'hypothèse de récurrence à Q pour terminer. □

Théorème 15.14 (de D'Alembert[1])

$\mathbb{C}$ est un corps algébriquement clos.

☞**Exemples** :

- Pour $n \geqslant 1$, $X^n - 1 = \prod_{k=0}^{n-1}(X - \exp(2ik\frac{\pi}{n}))$.
- Factoriser $X^{2n} - 1$ dans $\mathbb{C}[X]$ puis dans $\mathbb{R}[X]$.

$$\begin{aligned}X^{2n} - 1 &= \prod_{k=0}^{2n-1}(X - \exp(ik\frac{\pi}{n}))\\ &= (X-1)(X+1)\prod_{k=1}^{n-1}(X - \exp(ik\frac{\pi}{n}))(X - \exp(-ik\frac{\pi}{n}))\\ &= (X-1)(X+1)\prod_{k=1}^{n-1}(X^2 - 2\cos(k\frac{\pi}{n})X + 1).\end{aligned}$$

- Factoriser dans $\mathbb{R}[X]$: $X^4 + X^2 + 1$ et $X^4 + X^2 - 1$.

$$X^4 + X^2 + 1 = (X^2+1)^2 - X^2 = (X^2 - X + 1)(X^2 + X + 1)$$

$$\begin{aligned}X^4 + X^2 - 1 &= (X^2 + \frac{1}{2})^2 - \frac{5}{4} = (X^2 + \frac{1+\sqrt{5}}{2})(X^2 - \frac{\sqrt{5}-1}{2})\\ &= (X^2 + \frac{1+\sqrt{5}}{2})(X - \sqrt{\frac{\sqrt{5}-1}{2}})(X + \sqrt{\frac{\sqrt{5}-1}{2}})\end{aligned}$$

★**Exercice 15.3**

1/ *Soit $P \in \mathbb{C}[X]$ avec $P = \sum_k a_k X^k$, on appelle conjugué de P le polynôme $\overline{P} = \sum_k \overline{a_k} X^k$. Montrer que $\overline{P+Q} = \overline{P} + \overline{Q}$, que $\overline{PQ} = \overline{P} \times \overline{Q}$, et que $P \in \mathbb{R}[X]$ si et seulement si $\overline{P} = P$. Vérifier que pour $z \in \mathbb{C}$, $\overline{P(z)} = \overline{P}(\overline{z})$.*

2/ *Soit $P \in \mathbb{R}[X]$ non constant, soit z une racine complexe de P de multiplicité m. Montrer que $\overline{z}$ est racine de P de multiplicité m.*

3) Relations racines coefficients

Soit P un polynôme scindé sur $\mathbb{K}$, si $\deg(P) = n$ et si λ est le coefficient dominant de P, alors il existe $a_1, \ldots, a_n \in \mathbb{K}$ (racines de P) tels que $P = \lambda(X - a_1)\cdots(X - a_n)$, si on développe ensuite cette expression, on va obtenir les coefficients de P en fonction des a_k. Par exemple :

- $P = \lambda(X - a_1)(X - a_2) = \lambda X^2 - \lambda(a_1 + a_2)X + \lambda a_1 a_2$.
- $P = \lambda(X - a_1)(X - a_2)(X - a_3) = \lambda X^3 - \lambda(a_1 + a_2 + a_3)X^2 + \lambda(a_1a_2 + a_1a_3 + a_2a_3)X - \lambda a_1a_2a_3$.

Notation : On pose $\sigma_0 = 1$, et pour k compris entre 1 et n : $\sigma_k = \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} a_{i_1} \cdots a_{i_k}$.

σ_k est la somme des produits des racines (de P) par paquets de longueur k, par exemple : σ_1 est la somme des racines, σ_2 est la somme des produits deux à deux, $\cdots$, σ_n est le produit des racines.

Par récurrence on peut alors établir que :

$$(X - a_1)\cdots(X - a_n) = X^n - \sigma_1 X^{n-1} + \sigma_2 X^{n-2} - \cdots + (-1)^n \sigma_n = \sum_{k=0}^{n} (-1)^k \sigma_k X^{n-k}$$

On en déduit :

1. D'ALEMBERT Jean Le Rond (1717 – 1783) : mathématicien français qui contribua notamment à l'étude des nombres complexes, l'analyse et les probabilités.

Théorème 15.15

Soient $a_1, \ldots, a_n \in \mathbb{K}$, *si* $P = \sum_{k=0}^{n} \alpha_k X^k = \alpha_n(X - a_1)\cdots(X - a_n)$, *alors on a les relations racines - coefficients suivantes :* $\alpha_{n-k} = (-1)^k \alpha_n \sigma_k$.

En particulier, la somme des racines est $-\frac{\alpha_{n-1}}{\alpha_n}$ et le produit des racines est $(-1)^n \frac{\alpha_0}{\alpha_n}$.

★**Exercice 15.4** *Calculer la somme et le produit des racines n^{es} de l'unité ($n \geqslant 2$).*

IV FORMULE DE TAYLOR DES POLYNÔMES

1) Dérivation des polynômes

On reprend la dérivation usuelle des fonctions polynomiales :

Définition 15.9

Soit $P = \sum_k a_k X^k$, *on appelle polynôme dérivé de* P, *le polynôme noté* P' *ou* $\frac{dP}{dX}$, *et défini par :*

$$P' = \sum_{k \geqslant 1} k a_k X^{k-1}.$$

Par récurrence, la dérivée n*-ième de* P, *notée* $P^{(n)}$, *est :* $P^{(n)} = \begin{cases} P & \text{si } n = 0 \\ \left[P^{(n-1)}\right]' & \text{si } n \geqslant 1 \end{cases}$

Théorème 15.16 (propriétés)

Soient, $P, Q \in \mathbb{K}[X]$ *et soit* $\lambda \in \mathbb{K}$ *:*

- $P' = 0 \iff P$ *est constant.*
- $(P + Q)' = P' + Q'$ *et* $(\lambda P)' = \lambda P'$.
- $(P \times Q)' = P' \times Q + P \times Q'$, *plus généralement, on a la formule de* LEIBNIZ[2] *:*

$$(P \times Q)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} P^{(k)} \times Q^{(n-k)}.$$

- $P(Q)' = Q' \times P'(Q)$ *(dérivée d'une composée, une composée est une substitution de* X *par un autre polynôme).*

Preuve : La première propriété est simple à vérifier. Pour la deuxième propriété, on commence par montrer que $(X^n \times Q)' = nX^{n-1} \times Q + X^n \times Q'$, puis on applique la première propriété. La formule de LEIBNIZ se montre ensuite par récurrence sur n (exactement comme la formule du binôme de NEWTON). Quant à la troisième, on commence par le cas où $P = X^n$, c'est à dire on commence par montrer que $[Q^n]' = nQ' \times Q^{n-1}$, ce qui se fait par récurrence sur n, on utilise ensuite la première propriété pour le cas général. □

Théorème 15.17

Si $P = X^n$, *alors* $P^{(k)} = \begin{cases} \frac{n!}{(n-k)!} X^{n-k} & \text{si } k \leqslant n \\ 0 & \text{si } k > n \end{cases}$. *On en déduit que si* $P = \sum_n a_n X^n$, *alors*

$$P^{(k)} = \sum_{n \geqslant k} \frac{n!}{(n-k)!} a_n X^{n-k}$$

En particulier si $\deg(P) = n$ *alors* $P^{(n)} = a_n n!$ *et si* $k > \deg(P)$, *alors* $P^{(k)} = 0$. *D'autre part, lorsque* $k \leqslant \deg(P)$, *alors* $\deg(P^{(k)}) = \deg(P) - k$.

Preuve : Celle - ci est simple et laissée en exercice. □

2. *LEIBNIZ Gottfried* (1646 – 1716) : philosophe et mathématicien allemand.

2) Formule de Taylor

Soit $\mathrm{P} = \sum\limits_{k\in\mathbb{N}} a_k\mathrm{X}^k$, soit r un entier naturel, alors $\mathrm{P}^{(r)} = \sum\limits_{k\geqslant r}^{n} \frac{k!}{(k-r)!}a_k\mathrm{X}^{k-r}$, substituons 0 à X, on obtient alors $\mathrm{P}^{(r)}(0) = r!a_r$, on en déduit donc que :

$$\forall r \in [\![0\,;n]\!], a_r = \frac{\mathrm{P}^{(r)}(0)}{r!}.$$

On obtient ainsi la formule de Taylor[3] en 0 :

$$\mathrm{P} = \sum_k \frac{\mathrm{P}^{(k)}(0)}{k!}\mathrm{X}^k.$$

Soit $a \in \mathbb{K}$, posons $\mathrm{Q} = \mathrm{P}(\mathrm{X}+a)$ (composée de P avec le polynôme $\mathrm{X}+a$), d'après ce qui précède, on a :

$$\mathrm{Q} = \sum_k \frac{\mathrm{Q}^{(k)}(0)}{k!}\mathrm{X}^k.$$

Or, il est facile de montrer que $\mathrm{Q}^{(k)} = \mathrm{P}^{(k)}(\mathrm{X}+a)$, par conséquent $\mathrm{Q}^{(k)}(0) = \mathrm{P}^{(k)}(a)$, et comme $\mathrm{P} = \mathrm{Q}(\mathrm{X}-a)$, on obtient :

$$\mathrm{P} = \sum_k \frac{\mathrm{P}^{(k)}(a)}{k!}(\mathrm{X}-a)^k.$$

Théorème 15.18

Si $\mathrm{P} \in \mathbb{K}[\mathrm{X}]$ et $a \in \mathbb{K}$, alors $\mathrm{P} = \sum\limits_k \frac{\mathrm{P}^{(k)}(a)}{k!}(\mathrm{X}-a)^k$. C'est la formule de Taylor pour le polynôme P en a.

Applications :

– Division euclidienne d'un polynôme P par $(\mathrm{X}-a)^n$: d'après la formule de Taylor en a appliquée à P, on a :

$$\begin{aligned}
\mathrm{P} &= \sum_k \frac{\mathrm{P}^{(k)}(a)}{k!}(\mathrm{X}-a)^k\\
&= \sum_{k\geqslant n} \frac{\mathrm{P}^{(k)}(a)}{k!}(\mathrm{X}-a)^k + \sum_{k<n} \frac{\mathrm{P}^{(k)}(a)}{k!}(\mathrm{X}-a)^k\\
&= (\mathrm{X}-a)^n \times \sum_{k\geqslant n} \frac{\mathrm{P}^{(k)}(a)}{k!}(\mathrm{X}-a)^{k-n} + \sum_{k<n} \frac{\mathrm{P}^{(k)}(a)}{k!}(\mathrm{X}-a)^k,
\end{aligned}$$

comme $\deg(\sum\limits_{k<n} \frac{\mathrm{P}^{(k)}(a)}{k!}(\mathrm{X}-a)^k) < n$, on en déduit que le quotient Q et le reste R dans la division euclidienne par $(\mathrm{X}-a)^n$ sont :

$$\mathrm{Q} = \sum_{k\geqslant n} \frac{\mathrm{P}^{(k)}(a)}{k!}(\mathrm{X}-a)^{k-n} \text{ et } \mathrm{R} = \sum_{k<n} \frac{\mathrm{P}^{(k)}(a)}{k!}(\mathrm{X}-a)^k.$$

– Calcul de la multiplicité d'une racine :

Théorème 15.19

$a \in \mathbb{K}$ est une racine de P de multiplicité $n \geqslant 1$ si et seulement si :

$$\forall\, k \in [\![0\,;n-1]\!], \mathrm{P}^{(k)}(a) = 0 \text{ et } \mathrm{P}^{(n)}(a) \neq 0.$$

Preuve : En effet, d'après ce qui précède, $\mathrm{P} = (\mathrm{X}-a)^n\mathrm{Q}+\mathrm{R}$ avec $\mathrm{Q} = \sum\limits_{k\geqslant n} \frac{\mathrm{P}^{(k)}(a)}{k!}(\mathrm{X}-a)^{k-n}$ et $\mathrm{R} = \sum\limits_{k<n} \frac{\mathrm{P}^{(k)}(a)}{k!}(\mathrm{X}-a)^k$, d'où :

$$\begin{aligned}
n = m_{\mathrm{P}}(a) &\iff \mathrm{R} = 0 \text{ et } \mathrm{Q}(a) \neq 0\\
&\iff \mathrm{R}(\mathrm{X}+a) = 0 \text{ et } \mathrm{Q}(a) \neq 0\\
&\iff \forall k \in [\![0\,;n-1]\!], \mathrm{P}^{(k)}(a) = 0, \text{ et } \mathrm{P}^{(n)}(a) \neq 0
\end{aligned}$$

□

3. TAYLOR Brook (1685 – 1731) : mathématicien anglais qui a énoncé sa célèbre formule en 1715.

★**Exercice 15.5** *Montrer que si a est racine de* P *de multiplicité* $m \in \mathbb{N}^*$*, alors la multiplicité de a dans le polynôme* P' *est* $m-1$*. Montrer que la réciproque est fausse.*

V SOLUTION DES EXERCICES

Solution 15.1 *On pose* $\mathrm{R} = \mathrm{P} - \mathrm{Q}$*, alors* $\deg(\mathrm{R}) \leqslant \max(\deg(\mathrm{P}), \deg(\mathrm{Q})) \leqslant 2$*. On évalue le polynôme* R *en* -1*,* 0 *et* 1*, on trouve* $\mathrm{R}(-1) = 0$*,* $\mathrm{R}(0) = 0$*,* $\mathrm{R}(1) = 0$*, on a donc au moins 3 racines alors que* $\deg(\mathrm{R}) \leqslant 2$*, on en déduit que* R *et nul et donc* $\mathrm{P} = \mathrm{Q}$*.*

Solution 15.2 *On trouve que* $\mathrm{P} = (\mathrm{X}-1)(\mathrm{X}^2 - 2\mathrm{X} - 2)$ *et 1 n'est pas racine du deuxième facteur, donc* $m_{\mathrm{P}}(1) = 1$*.*
$\mathrm{Q} = (\mathrm{X}-1)^2(\mathrm{X}-2)$ *donc* $m_{\mathrm{Q}}(1) = 2$*.*

Solution 15.3

1/ *Simple vérification en appliquant la définition des opérations.*

2/ *Dans* $\mathbb{C}[\mathrm{X}]$ *on a* $\mathrm{P} = (\mathrm{X}-z)^m\mathrm{Q}$ *avec* $\mathrm{Q}(z) \neq 0$*, d'où en conjuguant (*P *et* Q *étant à coefficients réels)* $\mathrm{P} = \overline{\mathrm{P}} = (\mathrm{X}-\overline{z})^m\overline{\mathrm{Q}} = (\mathrm{X}-\overline{z})^m\mathrm{Q}$*, avec* $\mathrm{Q}(\overline{z}) = \overline{\mathrm{Q}}(z) \neq 0$*.*

Solution 15.4 *Soit* $\mathrm{P} = \mathrm{X}^n - 1$*, ses racines sont exactement les racines* n^{es} *de l'unité, notons* a_k *le coefficient de* X^k*, alors on sait que la somme des racines est donnée par la formule* $-\frac{a_{n-1}}{a_n} = 0$ *car* a_{n-1} *est nul (*$n-1 \geqslant 1$*).*

Le produit des racines est donné par la formule $(-1)^n\frac{a_0}{a_n} = (-1)^{n+1}$ *car* $a_0 = a_n = 1$*.*

Solution 15.5 *Pour* $k \in [\![0\,;m-1]\!]$ *on a* $\mathrm{P}^{(k)}(a) = 0$*, donc pour* $k \in [\![0\,;m-2]\!]$ *(intervalle vide si* $m = 1$*), on a* $\mathrm{P}'^{(k)}(a) = 0$*, on a aussi* $\mathrm{P}^{(m)}(a) \neq 0$*, donc* $\mathrm{P}'^{(m-1)}(a) \neq 0$*, donc* $m_{\mathrm{P}'}(a) = m-1$*.*

Pour la réciproque, considérons un contre-exemple, soit $\mathrm{P} = \mathrm{X}^2 + 1$*, on a* $\mathrm{P}' = 2\mathrm{X}$*, la multiplicité de* 0 *dans* P' *vaut* 1*, mais la multiplicité de* 0 *dans le polynôme* P *est nulle (et non pas* 2*).*

Chapitre 16

Calcul matriciel

Sommaire

$\mathbb{K}$ désigne un sous-corps de $\mathbb{C}$.

I MATRICES

1) Définitions

Définition 16.1

Soient $n, p \in \mathbb{N}^*$, on appelle matrice à n lignes et p colonnes à coefficients dans $\mathbb{K}$, toute application $\mathrm{M}\colon [\![1\,;n]\!] \times [\![1\,;p]\!] \to \mathbb{K}$. Pour $(i,j) \in [\![1\,;n]\!] \times [\![1\,;p]\!]$, on pose $\mathrm{M}(i,j) = \mathrm{M}_{i,j}$ (ou $m_{i,j}$), c'est le coefficient de la matrice M d'indices i et j, le premier indice est appelé indice de ligne, et le second indice de colonne.

L'ensemble des matrices à n lignes et p colonnes à coefficients dans $\mathbb{K}$ est noté $\mathcal{M}_{n,p}(\mathbb{K})$, on a donc $\mathcal{M}_{n,p}(\mathbb{K}) = \mathcal{F}([\![1\,;n]\!] \times [\![1\,;p]\!], \mathbb{K})$.

Notations : Si $\mathrm{M} \in \mathcal{M}_{n,p}(\mathbb{K})$, on peut écrire : $\mathrm{M} = (m_{i,j})_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant p}}$ ou bien $\mathrm{M} = \begin{pmatrix} m_{1,1} & \cdots & m_{1,p} \\ \vdots & & \vdots \\ m_{n,1} & \cdots & m_{n,p} \end{pmatrix}$.

Remarque 16.1 – *L'égalité entre deux matrices est en fait l'égalité entre deux fonctions, par conséquent deux matrices sont égales lorsqu'elles ont la même taille et les mêmes coefficients.*

Cas particuliers :

- Lorsque $n = p$ on dit que la matrice est **carrée**, l'ensemble des matrices carrées à n lignes est noté $\mathcal{M}_n(\mathbb{K})$ au lieu de $\mathcal{M}_{n,n}(\mathbb{K})$.
- Lorsque $n = 1$ on dit que M est une matrice **ligne**, ex. : $\begin{pmatrix} 1 & 2 & -3 \end{pmatrix} \in \mathcal{M}_{1,3}(\mathbb{K})$.
- Lorsque $p = 1$ on dit que M est une matrice **colonne**, ex. : $\begin{pmatrix} 2 \\ 0 \\ 5 \end{pmatrix} \in \mathcal{M}_{3,1}(\mathbb{K})$.

Définition 16.2

Soit $\mathrm{M} \in \mathcal{M}_{n,p}(\mathbb{K})$*, pour* $k \in [\![1\,;p]\!]$ *:*
On appelle k*-ième vecteur colonne de* M *le* n*-uplet* $c_k(\mathrm{M}) = (m_{1,k}, \ldots, m_{n,k})$*, c'est un élément de* $\mathbb{K}^n$*.*
On appelle k*-ième matrice colonne de* M *la matrice* $\mathcal{C}_k(\mathrm{M}) = \begin{pmatrix} m_{1,k} \\ \vdots \\ m_{n,k} \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbb{K})$.
On appelle k*-ième vecteur ligne de* M *le* p*-uplet* $\ell_k(\mathrm{M}) = (m_{k,1}, \ldots, m_{k,p})$*, c'est un élément de* $\mathbb{K}^p$*.*
On appelle k*-ième matrice ligne de* M *la matrice* $\mathcal{L}_k(\mathrm{M}) = \begin{pmatrix} m_{k,1} & \cdots & m_{k,p} \end{pmatrix} \in \mathcal{M}_{1,p}(\mathbb{K})$.

Définition 16.3 (transposition)

Soit $\mathrm{M} \in \mathcal{M}_{n,p}(\mathbb{K})$*, on appelle* ***transposée*** *de* M *la matrice de* $\mathcal{M}_{p,n}(\mathbb{K})$ *notée* M^{T} *et définie par :*

$$[\mathrm{M}^{\mathrm{T}}]_{i,j} = \mathrm{M}_{j,i} \ \text{ pour } i \in [\![1\,;p]\!] \text{ et } j \in [\![1\,;n]\!]$$

Autrement dit, la ligne i *de* M^{T} *est la colonne* i *de* M*.*

☞**Exemple** : Soit $\mathrm{M} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \in \mathcal{M}_{2,3}(\mathbb{K})$, on a $\mathrm{M}^{\mathrm{T}} = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} \in \mathcal{M}_{3,2}(\mathbb{K})$.

Théorème 16.1 (propriétés de la transposition)

On a les propriétés suivantes :
- $\mathcal{M}_n(\mathbb{K})$ *est stable pour la transposition.*
- $\left(\mathrm{M}^{\mathrm{T}}\right)^{\mathrm{T}} = \mathrm{M}$*, on en déduit en particulier que la transposition est une involution dans* $\mathcal{M}_n(\mathbb{K})$*.*
- $\mathrm{L}_k(\mathrm{M}^{\mathrm{T}}) = \mathrm{C}_k(\mathrm{M})$ *et* $\mathrm{C}_k(\mathrm{M}^{\mathrm{T}}) = \mathrm{L}_k(\mathrm{M})$*.*

Preuve : Celle-ci est simple et laissée en exercice. □

Matrices particulières :

- **Matrice nulle** : la matrice nulle à n lignes et p colonnes est la matrice de $\mathcal{M}_{n,p}(\mathbb{K})$ dont tous les coefficients sont nuls, celle-ci est notée $\mathrm{O}_{n,p}$. Lorsque $p = n$, la matrice $\mathrm{O}_{n,n}$ est notée simplement O_n, c'est la matrice nulle de $\mathcal{M}_n(\mathbb{K})$.
- **Matrice unité** : la matrice unité de $\mathcal{M}_n(\mathbb{K})$ est la matrice carrée de taille n, notée I_n et définie par $\mathrm{I}_n = (\delta_{i,j})_{1\leqslant i,j\leqslant n}$, c'est à dire, I_n est la matrice dont tous les coefficients diagonaux sont égaux à 1, les autres (coefficients extra-diagonaux) sont tous nuls.

☞**Exemple** : $\mathrm{I}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ est la matrice unité de $\mathcal{M}_3(\mathbb{K})$.

- **Matrice diagonale** : une matrice diagonale est une matrice carrée dont tous les coefficients extra-diagonaux sont nuls. C'est donc une matrice de la forme $\mathrm{M} = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & 0 \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & a_n \end{pmatrix}$, une telle matrice est notée parfois $\mathrm{M} = \mathrm{diag}(a_1, \ldots, a_n)$.

- **Matrice élémentaire** : une matrice élémentaire de $\mathcal{M}_{n,p}(\mathbb{K})$ est une matrice dont tous les coefficients sont nuls sauf un qui vaut 1. Il y a donc np matrices élémentaires dans $\mathcal{M}_{n,p}(\mathbb{K})$, pour $(i,j) \in [\![1\,;n]\!] \times [\![1\,;p]\!]$, on note $\mathrm{E}^{i,j}$ la matrice élémentaire qui possède un 1 ligne i colonne j, et des 0 ailleurs, plus précisément : $(\mathrm{E}^{i,j})_{k,l} = \delta_{i,k}\delta_{j,l}$.

☞**Exemple** : Dans $\mathcal{M}_{2,3}(\mathbb{K})$, on a $\mathrm{E}^{1,2} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

- **Matrice triangulaire supérieure** : c'est une matrice carrée dont tous les éléments situés sous la diagonale principale sont nuls. L'ensemble des matrices triangulaires supérieures de $\mathcal{M}_n(\mathbb{K})$ est noté $\mathcal{T}_n^s(\mathbb{K})$, on a donc :
$$\mathrm{M} \in \mathcal{T}_n^s(\mathbb{K}) \iff \forall\, i,j \in [\![1\,;n]\!], i > j \implies \mathrm{M}_{i,j} = 0.$$
- **Matrice triangulaire inférieure** : c'est une matrice carrée dont tous les éléments situés au-dessus de la diagonale principale sont nuls. L'ensemble des matrices triangulaires inférieures de $\mathcal{M}_n(\mathbb{K})$ est noté $\mathcal{T}_n^i(\mathbb{K})$, on a donc :
$$\mathrm{M} \in \mathcal{T}_n^i(\mathbb{K}) \iff \forall\, i,j \in [\![1\,;n]\!], i < j \implies \mathrm{M}_{i,j} = 0.$$
- **Matrice symétrique** : c'est une matrice qui est égale à sa transposée (elle est donc nécessairement carrée) : $\mathrm{M} = \mathrm{M}^{\mathrm{T}}$. L'ensemble des matrices symétriques de taille n est noté $\mathcal{S}_n(\mathbb{K})$, on a donc :
$$\mathrm{M} \in \mathcal{S}_n(\mathbb{K}) \iff \forall\, i,j \in [\![1\,;n]\!], \mathrm{M}_{i,j} = \mathrm{M}_{j,i}.$$
- **Matrice antisymétrique** : c'est une matrice qui est égale à l'opposé de sa transposée (elle est donc nécessairement carrée) : $\mathrm{M} = -\mathrm{M}^{\mathrm{T}}$. L'ensemble des matrices antisymétriques de taille n est noté $\mathcal{A}_n(\mathbb{K})$, on a donc :
$$\mathrm{M} \in \mathcal{A}_n(\mathbb{K}) \iff \forall\, i,j \in [\![1\,;n]\!], \mathrm{M}_{i,j} = -\mathrm{M}_{j,i},$$
on en déduit en particulier que $\mathrm{M}_{i,i} = 0$ (les coefficients diagonaux sont nuls).

2) Opérations sur les matrices

Définition 16.4 (somme de deux matrices)
Soient $\mathrm{A}, \mathrm{B} \in \mathcal{M}_{n,p}(\mathbb{K})$, on appelle somme de A et B la matrice de $\mathcal{M}_{n,p}(\mathbb{K})$ notée $\mathrm{A}+\mathrm{B}$ et définie par : $\forall\,(i,j) \in [\![1\,;n]\!] \times [\![1\,;p]\!], (\mathrm{A}+\mathrm{B})_{i,j} = \mathrm{A}_{i,j} + \mathrm{B}_{i,j}$. On additionne entre eux les éléments ayant les mêmes indices.

☞**Exemple** : $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} + \begin{pmatrix} -1 & 0 & 1 \\ -2 & 3 & 4 \end{pmatrix} = \begin{pmatrix} 0 & 2 & 4 \\ 2 & 8 & 10 \end{pmatrix}$.

Théorème 16.2
$(\mathcal{M}_{n,p}(\mathbb{K}), +)$ est un groupe commutatif. L'élément neutre est la matrice nulle, $\mathrm{O}_{n,p}$, et si $\mathrm{A} \in \mathcal{M}_{n,p}(\mathbb{K})$, l'opposé de A est la matrice $-\mathrm{A}$ définie par $\forall\,(i,j) \in [\![1\,;n]\!] \times [\![1\,;p]\!], (-\mathrm{A})_{i,j} = -\mathrm{A}_{i,j}$.

Preuve : Celle-ci est simple et laissée en exercice. □

Définition 16.5 (produit par un scalaire)
Soit $\mathrm{M} \in \mathcal{M}_{n,p}(\mathbb{K})$ et soit $\lambda \in \mathbb{K}$, on appelle produit de la matrice M par le scalaire λ la matrice de $\mathcal{M}_{n,p}(\mathbb{K})$ notée $\lambda.\mathrm{M}$ et définie par :$\forall\,(i,j) \in [\![1\,;n]\!] \times [\![1\,;p]\!], (\lambda.\mathrm{M})_{i,j} = \lambda \times \mathrm{M}_{i,j}$. C'est à dire, chaque coefficient de M est multiplié par λ.

☞**Exemple** : $2 \cdot \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ 6 & 8 \\ 10 & 12 \end{pmatrix}$.

Propriétés : On peut vérifier facilement, soient $\mathrm{A}, \mathrm{B} \in \mathcal{M}_{n,p}(\mathbb{K})$, soient $\lambda, \mu \in \mathbb{K}$:
- $1.\mathrm{A} = \mathrm{A}$.
- $\lambda.(\mathrm{A}+\mathrm{B}) = \lambda.\mathrm{A} + \lambda.\mathrm{B}$.
- $(\lambda+\mu).\mathrm{A} = \lambda.\mathrm{A} + \mu.\mathrm{A}$.
- $(\lambda\mu).\mathrm{A} = \lambda.(\mu.\mathrm{A})$.

★**Exercice 16.1** *Montrer que $\mathcal{T}_n^s(\mathbb{K}), \mathcal{T}_n^i(\mathbb{K}), \mathcal{S}_n(\mathbb{K})$ et $\mathcal{A}_n(\mathbb{K})$ sont des sous-groupes de $(\mathcal{M}_n(\mathbb{K}), +)$.*

> **Théorème 16.3 (propriétés de la transposition)**
> *La transposition est linéaire et bijective de $\mathcal{M}_{n,p}(\mathbb{K})$ sur $\mathcal{M}_{p,n}(\mathbb{K})$. C'est à dire si A et B sont dans $\mathcal{M}_{n,p}(\mathbb{K})$ et $\lambda \in \mathbb{K}$, alors $A + B^T = A^T + B^T$ et $(\lambda A)^T = \lambda A^T$.*

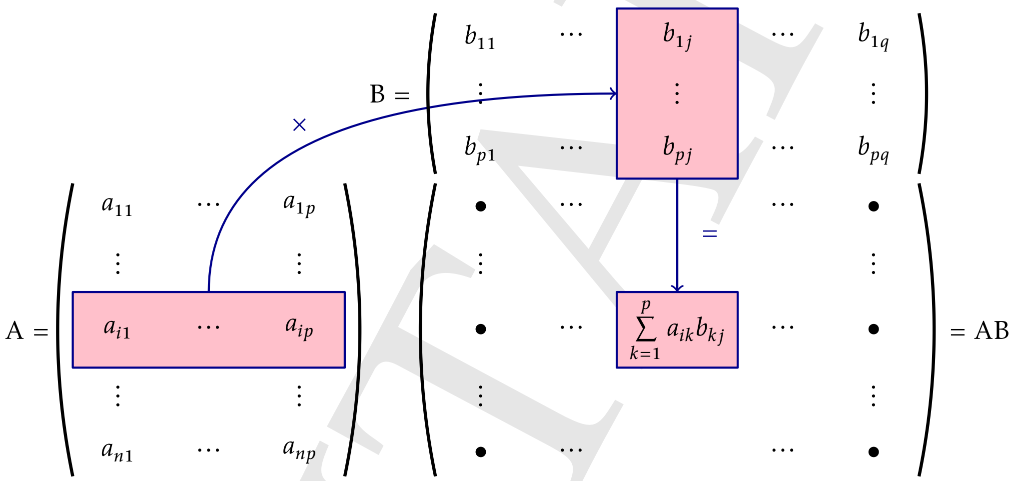
Preuve : Celle-ci est simple et laissée en exercice. □

II PRODUIT MATRICIEL

1) Définition

> **Définition 16.6**
> *Soient $A \in \mathcal{M}_{n,p}(\mathbb{K})$, soit $B \in \mathcal{M}_{p,q}(\mathbb{K})$, on appelle produit de A par B la matrice de $\mathcal{M}_{n,q}(\mathbb{K})$ notée $A \times B$ et définie par :*
> $$\forall\,(i,j) \in [\![1\,;n]\!] \times [\![1\,;q]\!], [A \times B]_{i,j} = \sum_{k=1}^{p} A_{i,k}B_{k,j}.$$
> *On retient ceci en disant que le coefficient $[A \times B]_{i,j}$ est le résultat du « produit de la ligne i de A avec la colonne j de B ».*

Disposition des calculs :



Remarque 16.2 :
- *Le produit $A \times B$ n'est possible que si le nombre de colonnes de A est égal au nombre de lignes de B. Le résultat a alors autant de lignes que A et autant de colonnes que B.*
- *Dans $\mathcal{M}_n(\mathbb{K})$ le produit matriciel est interne.*
- *En général $A \times B \neq B \times A$, il se peut même que $A \times B$ soit défini, mais pas $B \times A$.*

☞**Exemple** :

$$\begin{pmatrix} 1 & -1 & 3 \end{pmatrix} \times \begin{pmatrix} 0 & 1 \\ -1 & 2 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 7 & 2 \end{pmatrix} \qquad \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \times \begin{pmatrix} 1 & -1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 3 \\ 2 & -2 & 6 \\ 3 & -3 & 9 \end{pmatrix}$$

$$\begin{pmatrix} 1 & -1 & 2 \end{pmatrix} \times \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 5 \end{pmatrix} \qquad \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} \times \begin{pmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 1 & 3 \\ -1 & -2 & -1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} \times \begin{pmatrix} 0 & 1 \\ -1 & 2 \\ 0 & 1 \end{pmatrix} \text{ n'est pas défini} \qquad \begin{pmatrix} 0 & 1 \\ -1 & 2 \\ 0 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ -5 & 0 \\ -2 & 1 \end{pmatrix}$$

2) Propriétés du produit matriciel

- « **Associativité** » : Si $A \in \mathcal{M}_{n,p}(\mathbb{K})$, $B \in \mathcal{M}_{p,q}(\mathbb{K})$ et $C \in \mathcal{M}_{q,r}(\mathbb{K})$, alors :
$$A \times (B \times C) = (A \times B) \times C \in \mathcal{M}_{n,r}(\mathbb{K}).$$
- « **Élément neutre** » : Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$, on a : $A \times I_p = A$ et $I_n \times A = A$.
- « **Distributivité** » : Soient $A, B \in \mathcal{M}_{n,p}(\mathbb{K})$, soit $C \in \mathcal{M}_{p,q}(\mathbb{K})$ et $D \in \mathcal{M}_{r,n}$, on a :
$$(A + B) \times C = A \times C + B \times C \text{ et } D \times (A + B) = D \times A + D \times B$$
- **Transposée d'un produit** : Si $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B \in \mathcal{M}_{p,q}(\mathbb{K})$ alors : $(A \times B)^{T} = B^{T} \times A^{T}$.
Preuve : $[(A \times B)^{T}]_{i,j} = [A \times B]_{j,i} = \sum\limits_{k=1}^{p} a_{j,k} b_{k,i} = \sum\limits_{k=1}^{p} [B^{T}]_{i,k}[A^{T}]_{k,j} = [B^{T} \times A^{T}]_{i,j}$. □

★**Exercice 16.2** *Calculer le produit entre deux matrices carrées élémentaires de même taille.*

III MATRICES CARRÉES INVERSIBLES

1) Structure de $\mathcal{M}_n(\mathbb{K})$

Théorème 16.4
On a le résultat suivant : $(\mathcal{M}_n(\mathbb{K}), +, \times)$ est un anneau (non commutatif si $n \geqslant 2$).

Preuve : On sait déjà que $(\mathcal{M}_n(\mathbb{K}), +)$ est un groupe, le reste découle des propriétés du produit matriciel. Donnons un contre-exemple pour la non commutativité : soit $A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, on a $AB = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$, mais $BA = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$. □

Remarque 16.3 :

- *L'anneau $\mathcal{M}_n(\mathbb{K})$ n'est pas intègre lorsque $n \geqslant 2$. Par exemple : $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \times \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} = O_2$. De ce fait, il y a dans $\mathcal{M}_n(\mathbb{K})$ des éléments nilpotents, par exemple : $A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$.*
- *On peut utiliser dans $\mathcal{M}_n(\mathbb{K})$ les règles du calcul algébrique, en prenant garde toutefois au fait que le produit n'est pas commutatif. En particulier :*
  - *si $A, B \in \mathcal{M}_n(\mathbb{K})$ commutent (i.e. $AB = BA$), alors on peut utiliser le binôme de Newton :*
  $(A + B)^n = \sum\limits_{k=0}^{n} \binom{n}{k} A^k \times B^{n-k}$.
  Mais si $AB \neq BA$ on peut néanmoins développer, par exemple : $(A + B)^2 = A^2 + AB + BA + B^2$.
  - *si $A, B \in \mathcal{M}_n(\mathbb{K})$ commutent (i.e. $AB = BA$), alors pour tout entier naturel p on a :*
  $A^{p+1} - B^{p+1} = (A - B) \times \sum\limits_{k=0}^{p} A^k B^{p-k}$.

★**Exercice 16.3** *Soit $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$. En écrivant $A = I_3 + J$, calculer A^n pour $n \in \mathbb{N}$.*

Définition 16.7
L'ensemble $(\mathcal{M}_n(\mathbb{K}), +, \times)$ ayant structure d'anneau, on peut s'intéresser aux éléments inversibles de cet anneau. C'est à dire aux matrices $M \in \mathcal{M}_n(\mathbb{K})$ pour lesquelles il existe une matrice $N \in \mathcal{M}_n(\mathbb{K})$ telle que $M \times N = N \times M = I_n$. Si $M \in \mathcal{M}_n(\mathbb{K})$ est inversible, son inverse sera noté M^{-1}.
L'ensemble des inversibles de l'anneau $(\mathcal{M}_n(\mathbb{K}), +, \times)$ est noté $GL_n(\mathbb{K})$.

Théorème 16.5

• Le produit de deux matrices inversibles est inversible, ce qui entraîne que $(\mathrm{GL}_n(\mathbb{K}), \times)$ *est un groupe (appelé groupe linéaire de type* n *sur* $\mathbb{K}$*).*
• Si $\mathrm{M}, \mathrm{N} \in \mathrm{GL}_n(\mathbb{K})$*, alors* $(\mathrm{M} \times \mathrm{N})^{-1} = \mathrm{N}^{-1} \times \mathrm{M}^{-1}$.

Preuve : Celle-ci est simple et laissée en exercice. □

2) Cas particuliers

– **Matrices diagonales inversibles** : Soit $\mathrm{D} = \mathrm{diag}(a_1, \ldots, a_n) \in \mathcal{M}_n(\mathbb{K})$, alors D est inversible ssi **les coefficients diagonaux sont tous non nuls**, auquel cas on a : $\mathrm{D}^{-1} = \mathrm{diag}(\frac{1}{a_1}, \ldots, \frac{1}{a_n})$.
Preuve : Si les coefficients diagonaux sont tous non nuls, il est facile de vérifier que la matrice proposée est bien l'inverse de D.
Réciproquement, supposons $\mathrm{D} \in \mathrm{GL}_n(\mathbb{K})$, alors l'équation $\mathrm{DX} = \mathrm{O}_{n,1}$ d'inconnue $\mathrm{X} \in \mathcal{M}_{n,1}(\mathbb{K})$ admet comme unique solution $\mathrm{X} = \mathrm{D}^{-1} \times \mathrm{O}_{n,1} = \mathrm{O}_{n,1}$. Supposons $a_1 = 0$ et prenons $\mathrm{X} \in \mathcal{M}_{n,1}(\mathbb{K})$ définie par $\mathrm{X}_{i,1} = \delta_{i,1}$, il est facile de voir que le produit DX donne la première colonne de D, c'est à dire $\mathrm{O}_{n,1}$, pourtant $\mathrm{X} \neq \mathrm{O}_{n,1}$: contradiction, donc $a_1 \neq 0$. Le raisonnement est similaire pour les autres coefficients. □

– **Polynômes de matrices** : Soit P un polynôme et $\mathrm{A} \in \mathcal{M}_n(\mathbb{K})$, si $\mathrm{P}(\mathrm{X}) = \sum_{k=0}^{r} a_k \mathrm{X}^k$, alors la matrice P(A) est $\mathrm{P}(\mathrm{A}) = \sum_{k=0}^{r} a_k \mathrm{A}^k$ (substitution de X par A), on a alors le résultat suivant : Si $\mathrm{P}(\mathrm{A}) = \mathrm{O}_n$ et si $\mathrm{P}(0) \neq 0$, alors A est inversible.
Preuve : $\mathrm{P}(0) \neq 0$ signifie que $a_0 \neq 0$, on a alors :
$$\mathrm{I}_n = \mathrm{A} \times \left[\sum_{k=1}^{r} \frac{-a_k}{a_0} \mathrm{A}^{k-1}\right] = \left[\sum_{k=1}^{r} \frac{-a_k}{a_0} \mathrm{A}^{k-1}\right] \times \mathrm{A}.$$
Par exemple, si $\mathrm{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ avec $ad - bc \neq 0$, on vérifie que $\mathrm{A}^2 - (a+d)\mathrm{A} + (ad-bc)\mathrm{I}_2 = \mathrm{O}_2$, donc A est inversible et $\mathrm{A}^{-1} = \frac{1}{ad-bc}[(a+d)\mathrm{I}_2 - \mathrm{A}] = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$. □

3) Caractérisation

Théorème 16.6 (caractérisations des matrices carrées inversibles)

Soit $\mathrm{A} \in \mathcal{M}_n(\mathbb{K})$*, alors les assertions suivantes sont équivalentes :*

a) A est inversible.

b) Il existe une matrice $\mathrm{B} \in \mathcal{M}_n(\mathbb{K})$ *telle que* $\mathrm{BA} = \mathrm{I}_n$.

c) L'équation $\mathrm{AX} = \mathrm{O}_{n,1}$ *d'inconnue* $\mathrm{X} \in \mathcal{M}_{n,1}(\mathbb{K})$ *admet une* ***unique solution*** $\mathrm{X} = \mathrm{O}_{n,1}$.

d) $\forall\, \mathrm{Y} \in \mathcal{M}_{n,1}(\mathbb{K})$*, l'équation* $\mathrm{AX} = \mathrm{Y}$ *d'inconnue* $\mathrm{X} \in \mathcal{M}_{n,1}(\mathbb{K})$ *admet une unique solution.*

e) $\forall\, \mathrm{Y} \in \mathcal{M}_{n,1}(\mathbb{K})$*, l'équation* $\mathrm{AX} = \mathrm{Y}$ *d'inconnue* $\mathrm{X} \in \mathcal{M}_{n,1}(\mathbb{K})$ *admet au moins une solution.*

Preuve : L'implication $i) \implies ii)$ est évidente en prenant $\mathrm{B} = \mathrm{A}^{-1}$.
Montrons $ii) \implies iii)$: On a $\mathrm{BA} = \mathrm{I}_n$, d'où $\mathrm{AX} = \mathrm{O}_{n,1} \implies \mathrm{BAX} = \mathrm{O}_{n,1} = \mathrm{X}$.
Montrons $iii) \implies iv)$: si $\mathrm{AX} = \mathrm{Y}$ alors $\mathrm{BAX} = \mathrm{BY}$ d'où $\mathrm{X} = \mathrm{BY}$ (existence). Si $\mathrm{AX} = \mathrm{AZ} = \mathrm{Y}$, alors $\mathrm{A}(\mathrm{X}-\mathrm{Z}) = \mathrm{O}_{n,1}$, d'où $\mathrm{X} - \mathrm{Z} = \mathrm{BA}(\mathrm{X}-\mathrm{Z}) = \mathrm{B} \times \mathrm{O}_{n,1} = \mathrm{O}_{n,1}$ et donc $\mathrm{X} = \mathrm{Z}$ (unicité).
L'implication $iv) \implies v)$ est évidente.
Montrons $v) \implies i)$: en prenant pour Y les différentes colonnes de I_n, on en déduit qu'il existe une matrice carrée B telle que $\mathrm{AB} = \mathrm{I}_n$. Donc pour toute colonne Y il existe une colonne X telle que $\mathrm{X}^{\mathrm{T}} \times \mathrm{A} = \mathrm{Y}^{\mathrm{T}}$ (c'est $\mathrm{X} = \mathrm{B}^{\mathrm{T}} \times \mathrm{Y}$), en prenant pour Y les différentes colonnes de I_n, on en déduit qu'il existe une matrice carrée C telle que $\mathrm{CA} = \mathrm{I}_n$, on a donc $\mathrm{C} = \mathrm{C} \times \mathrm{I}_n = \mathrm{CAB} = \mathrm{I}_n\mathrm{B} = \mathrm{B}$, d'où A est inversible, d'inverse B. □

Remarque 16.4 :

– *Il découle en particulier de ce théorème que si* $\mathrm{BA} = \mathrm{I}_n$ *alors* $\mathrm{AB} = \mathrm{I}_n$ *(car* $\mathrm{A} \in \mathrm{GL}_n(\mathbb{K})$ *et donc* $\mathrm{B} = \mathrm{A}^{-1}$*), ce qui est remarquable puisqu'en général* $\mathrm{AB} \neq \mathrm{BA}$.
– *Une matrice carrée A ayant une colonne nulle ne peut pas être inversible, car si la colonne i est nulle alors* $\mathrm{A} \times \mathcal{C}_i(\mathrm{I}_n) = \mathcal{C}_i(\mathrm{A}) = \mathrm{O}_{n,1}$*, alors que* $\mathcal{C}_i(\mathrm{I}_n)$ *n'est pas nulle.*

★**Exercice** 16.4

1/ *Si* $A \in GL_n(\mathbb{K})$*, montrer que* A^T *est inversible et que* $(A^T)^{-1} = (A^{-1})^T$.

2/ *Soit* $A = \begin{pmatrix} 1 & \lambda & -1 \\ 0 & 2 & 1 \\ 1 & 0 & 1 \end{pmatrix}$*, déterminer en fonction de* λ *si* A *est inversible ou non, si c'est le cas, calculer* A^{-1}.

3/ *Soit* $T \in \mathcal{M}_n(\mathbb{K})$ *une matrice triangulaire supérieure, montrer que* $T \in GL_n(\mathbb{K})$ *ssi ses éléments diagonaux sont tous non nuls, si c'est le cas, montrer que* T^{-1} *est également triangulaire supérieure.*

IV OPÉRATIONS ÉLÉMENTAIRES SUR LES MATRICES

1) Définition

À retenir

Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B \in \mathcal{M}_{p,q}(\mathbb{K})$:
- soit $i \in [\![1\,;n]\!]$, on a $\mathcal{L}_i(A \times B) = \mathcal{L}_i(A) \times B$ (ligne i de $A \times B$).
- soit $j \in [\![1\,;q]\!]$, on a $\mathcal{C}_j(A \times B) = A \times \mathcal{C}_j(B)$ (colonne j de $A \times B$).

Définition 16.8

Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$*, on appelle opérations élémentaires sur* A *les opérations suivantes :*
- *Permuter deux lignes de* A *(ou deux colonnes), notation :* $L_i \leftrightarrow L_j$ *(respectivement* $C_i \leftrightarrow C_j$*).*
- *Multiplier une ligne (ou une colonne) par un scalaire* ***non nul****, notation :* $L_i \leftarrow \alpha L_i$ *(respectivement* $C_i \leftarrow \alpha C_i$*).*
- *Ajouter à une ligne (ou une colonne) un multiple d'une autre ligne (respectivement une autre colonne), notation :* $L_i \leftarrow L_i + \alpha L_j$*, avec* $i \neq j$ *(respectivement* $C_i \leftarrow C_i + \alpha C_j$*)).*

2) Propriétés

Théorème 16.7

Effectuer une opération élémentaire sur une matrice $A \in \mathcal{M}_{n,p}(\mathbb{K})$ *revient à multiplier* A *à gauche par une matrice* ***inversible*** *pour les opérations sur les lignes (à droite pour une opération sur les colonnes).*

Preuve : Pour l'opération $L_i \leftrightarrow L_j$ (avec $i \neq j$) : soit $P_{ij} \in \mathcal{M}_n(\mathbb{K})$ la matrice obtenue en effectuant cette opération sur la matrice I_n, alors $P_{ij} \times A$ est la matrice que l'on obtient en effectuant l'opération $L_i \leftrightarrow L_j$ dans A, en effet :
- si $k \notin \{i, j\}$, alors $\mathcal{L}_k(P_{ij} \times A) = \mathcal{L}_k(P_{ij}) \times A = \mathcal{L}_k(I_n) \times A = \mathcal{L}_k(I_n \times A) = \mathcal{L}_k(A)$,
- si $k = i$, alors $\mathcal{L}_i(P_{ij} \times A) = \mathcal{L}_i(P_{ij}) \times A = \mathcal{L}_j(I_n) \times A = \mathcal{L}_j(A)$,
- de même $\mathcal{L}_j(P_{ij} \times A) = \mathcal{L}_i(A)$.

De plus, par définition même, $P_{ij} \times P_{ij} = I_n$, donc P_{ij} est inversible et $P_{ij}^{-1} = P_{ij}$.

Pour l'opération $L_i \leftarrow \alpha L_i$, avec $\alpha \in \mathbb{K}^*$: soit $D_i(\alpha) \in \mathcal{M}_n(\mathbb{K})$ la matrice obtenue en effectuant cette opération sur I_n, alors $D_i(\alpha) \times A$ est la matrice que l'on obtient en effectuant cette même opération sur A, en effet :
- si $k \neq i$, alors $\mathcal{L}_k(D_i(\alpha) \times A) = \mathcal{L}_k(D_i(\alpha)) \times A = \mathcal{L}_k(I_n) \times A = \mathcal{L}_k(A)$,
- si $k = i$, $\mathcal{L}_i(D_i(\alpha) \times A) = \mathcal{L}_i(D_i(\alpha)) \times A = \alpha\mathcal{L}_i(I_n) \times A = \alpha\mathcal{L}_i(A)$.

De plus, il est clair que $D_i(\alpha) \times D_i(1/\alpha) = I_n$, donc cette matrice est inversible et $D_i(\alpha)^{-1} = D_i(1/\alpha)$.

Pour l'opération $L_i \leftarrow L_i + \alpha L_j$, avec $i \neq j$ et $\alpha \in \mathbb{K}$: soit $T_{ij}(\alpha) \in \mathcal{M}_n(\mathbb{K})$ la matrice obtenue en effectuant cette opération sur I_n, alors $T_{ij}(\alpha) \times A$ est la matrice que l'on obtient en effectuant cette même opération sur A, en effet :
- si $k \neq i$, $\mathcal{L}_k(T_{ij}(\alpha) \times A) = \mathcal{L}_k(T_{ij}(\alpha)) \times A = \mathcal{L}_k(I_n) \times A = \mathcal{L}_k(A)$,
- si $k = i$, $\mathcal{L}_i(T_{ij}(\alpha) \times A) = \mathcal{L}_i(T_{ij}(\alpha)) \times A = (\mathcal{L}_i(I_n) + \alpha\mathcal{L}_j(I_n)) \times A = \mathcal{L}_i(I_n) \times A + \alpha.\mathcal{L}_j(I_n) \times A = \mathcal{L}_i(A) + \alpha.\mathcal{L}_j(A)$.

De plus, il est clair que $T_{ij}(\alpha) \times T_{ij}(-\alpha) = I_n$, donc cette matrice est inversible et $T_{ij}(\alpha)^{-1} = T_{ij}(-\alpha)$. □

3) La méthode du pivot de Gauss

Celle-ci consiste à transformer la matrice A en la matrice $\begin{pmatrix} p_1 & * & \cdots & * & \cdots & * \\ 0 & \ddots & \ddots & \vdots & \cdots & \vdots \\ \vdots & \ddots & p_r & * & \cdots & * \\ \vdots & & 0 & 0 & \cdots & 0 \\ \vdots & & \boxtimes & \vdots & \cdots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}$, avec $p_1 \times \cdots \times p_r \neq 0$, à l'aide des opérations élémentaires sur les lignes ou les colonnes (méthode de Gauss), à chaque étape, la matrice obtenue s'écrit sous la forme $\mathrm{U}_k \times \mathrm{A} \times \mathrm{V}_k$ avec U_k, V_k inversibles. À l'étape n° k, le principe est le suivant :

a) On choisit un pivot (*i.e.* un coefficient non nul) dans les lignes L_k à L_n et dans les colonnes C_k à C_p.

b) On amène le pivot à sa place, c'est à dire sur la ligne L_k dans la colonne C_k en échangeant éventuellement deux lignes et/ou deux colonnes.

c) On fait des éliminations **en dessous** du pivot pour faire apparaître des 0, avec les opérations du type : $\mathrm{L}_i \leftarrow \mathrm{L}_i + \alpha \mathrm{L}_k$.

☞**Exemple** : Soit $\mathrm{A} = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 0 & 2 \\ -1 & 2 & -12 \end{pmatrix}$

Étape 1 : premier pivot : 1 (ligne L_1 colonne C_2)

$$\mathrm{C}_1 \leftrightarrow \mathrm{C}_2 \text{ et } \mathrm{L}_3 \leftarrow \mathrm{L}_3 - 2\mathrm{L}_1 \text{ donnent } \begin{pmatrix} \boxed{1} & 3 & 1 \\ 0 & 1 & 2 \\ 0 & -7 & -14 \end{pmatrix}.$$

Étape 2 : deuxième pivot : 1 (ligne L_2 colonne C_2)

$$\mathrm{L}_3 \leftarrow \mathrm{L}_3 + 7\mathrm{L}_2 \text{ donne } \begin{pmatrix} \boxed{1} & 3 & 1 \\ 0 & \boxed{1} & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

Étape 3 : pas de troisième pivot.

★**Exercice 16.5** *Avec la matrice* A *précédente, déduire de la méthode deux matrices inversibles* U *et* V *telles que* $\mathrm{UAV} = \mathrm{J}_{3,3,2}$.

4) Calcul pratique de l'inverse d'une matrice

Soit $\mathrm{A} \in \mathcal{M}_n(\mathbb{K})$, supposons qu'en r opérations **sur les lignes uniquement** de A on obtienne la matrice I_n, on a alors une relation du type $\mathrm{G}_r \times \cdots \times \mathrm{G}_1 \times \mathrm{A} = \mathrm{I}_n$, où G_i est la matrice correspondant à l'opération numéro i. On peut alors en déduire que la matrice A est inversible et que son inverse est $\mathrm{A}^{-1} = \mathrm{G}_r \times \cdots \times \mathrm{G}_1$, pour obtenir cette matrice, il suffit d'effectuer les mêmes opérations (dans le même ordre) sur la matrice I_n en même temps que sur A. La méthode consiste donc à écrire la matrice A suivie de la matrice I_n :

$$\begin{array}{ccc|ccc} & \mathrm{A} & & & \mathrm{I}_n & \\ a_{1,1} & \cdots & a_{1,n} & 1 & & \mathrm{O} \\ \vdots & & \vdots & & \ddots & \\ a_{n,1} & \cdots & a_{n,n} & \mathrm{O} & & 1 \end{array}$$

Les opérations sont effectuées sur toute la longueur de chaque ligne. L'objectif est d'obtenir I_n à la place de A, on conclura que A est inversible, et là où il y avait I_n on aura A^{-1}, on utilise la méthode de Gauss-Jordan :

À l'étape k :

– On choisit un pivot (*i.e.* un coefficient non nul) dans les lignes $\mathrm{L}_k \ldots \mathrm{L}_n$ et **dans la colonne** C_k.

– On amène le pivot à sa place : ligne L_k (en échangeant éventuellement deux lignes).

– On fait les éliminations (pour faire apparaître des zéros) **en dessous et au-dessus** du pivot avec les opérations : $L_i \leftarrow L_i + \alpha L_k$.

Il y a donc au plus n étapes. Il y a deux cas possibles au cours du processus :

– Si à chaque étape on peut trouver un pivot, alors après l'étape n, il ne reste plus qu'à diviser chaque ligne par le pivot correspondant pour obtenir la matrice I_n : c'est le cas où la matrice A **est inversible**.

– Si au cours de l'étape k on ne peut pas trouver de pivot dans la colonne C_k et dans les lignes $L_k \dots L_n$, alors on est dans la situation suivante, à l'issue de l'étape $k-1$:

$$\begin{array}{cccccc|cccc}
p_1 & 0 & 0 & * & * & * & * & \cdots & \cdots & * \\
0 & \ddots & 0 & \vdots & \vdots & \vdots & \vdots & & & \vdots \\
\vdots & & p_{k-1} & * & * & * & \vdots & & & \vdots \\
\vdots & & 0 & 0 & * & * & \vdots & & & \vdots \\
\vdots & & \vdots & \vdots & \vdots & \vdots & \vdots & & & \vdots \\
0 & \cdots & 0 & 0 & * & * & * & \cdots & \cdots & *
\end{array}$$

$p_1, \dots, p_{k-1}$ désignent les pivots des $k-1$ étapes précédentes, ces pivots étant non nuls, il est facile de voir qu'avec des opérations sur les colonnes, on peut faire apparaître des zéros dans la colonne k sur les lignes $L_1 \dots L_{k-1}$, sans changer les coefficients des lignes $L_k \dots L_n$ de cette même colonne. La matrice ainsi obtenue possède une colonne nulle, donc elle n'est pas inversible, et par conséquent la matrice initiale A est **non inversible** (les opérations élémentaires conservent l'inversibilité).

Remarque 16.5 – *On a donc établi que si* A *est non inversible, la méthode de Gauss-Jordan n'aboutit pas, donc il existe deux matrices inversibles* U *et* V *telles que* UAV *possède une colonne nulle, soit* j *son indice, alors la colonne* $X = D\mathcal{C}_j(I_n)$ *est non nulle et vérifie* $AX = O_{n,1}$. *Par contraposition, si l'équation* $AX = O_{n,1}$ *a pour unique solution* $X = O_{n,1}$, *alors* A *est inversible.*

☞**Exemple** : Soit $A = \begin{pmatrix} 2 & 4 & 2 \\ 0 & 1 & 1 \\ 2 & 2 & -1 \end{pmatrix}$, appliquons la méthode de *Gauss-Jordan* : $\begin{array}{ccc|ccc} 2 & 4 & 2 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 2 & 2 & -1 & 0 & 0 & 1 \end{array}$.

Étape 1 : pivot $p_1 = 2$, ligne L_1 colonne C_1, éliminations : $L_3 \leftarrow L_3 - L_1$, ce qui donne :

$$\begin{array}{ccc|ccc} 2 & 4 & 2 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & -2 & -3 & -1 & 0 & 1 \end{array}$$

Étape 2 : pivot $p_2 = 1$, ligne L_2, colonne C_2, éliminations : $L_1 \leftarrow L_1 - 4L_2$ et $L_3 \leftarrow L_3 + 2L_2$, ce qui donne :

$$\begin{array}{ccc|ccc} 2 & 0 & -2 & 1 & -4 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & -1 & -1 & 2 & 1 \end{array}$$

Étape 3 : pivot $p_3 = -1$, ligne L_3, colonne C_3, éliminations : $L_1 \leftarrow L_1 - 2L_3$ et $L_2 \leftarrow L_2 + L_3$, ce qui donne :

$$\begin{array}{ccc|ccc} 2 & 0 & 0 & 3 & -8 & -2 \\ 0 & 1 & 0 & -1 & 3 & 1 \\ 0 & 0 & -1 & -1 & 2 & 1 \end{array}$$

En conclusion, la matrice A est inversible et son inverse est : $A^{-1} = \begin{pmatrix} 3/2 & -4 & -1 \\ -1 & 3 & 1 \\ 1 & -2 & -1 \end{pmatrix}$.

V MATRICES ET SYSTÈMES LINÉAIRES

1) Définition

Définition 16.9

Un système linéaire (S) *est un système de la forme* (S) $\begin{cases} a_{11}x_1 + \dots + a_{1p}x_p = b_1 \\ \qquad\qquad\vdots \\ a_{n1}x_1 + \dots + a_{np}x_p = b_n \end{cases}$ *, où les scalaires*

*$x_1, \dots, x_p$ sont les inconnues, et les scalaires $b_1, \dots, b_n$ sont les seconds membres, et les scalaires $a_{i,j}$ sont les coefficients de la matrice du système. Lorsque $n = p$ on dit que le système est **carré**.*

Le système $(\mathrm{S_H})$ $\begin{cases} a_{11}x_1 + \dots + a_{1p}x_p = & 0 \\ \qquad\vdots & \\ a_{n1}x_1 + \dots + a_{np}x_p = & 0 \end{cases}$ *est appelé **système homogène associé** à* (S).

*Deux systèmes linéaires (S_1) et (S_2) sont dits **équivalents** lorsqu'ils ont exactement les mêmes solutions, on écrit alors $(\mathrm{S}_1) \iff (\mathrm{S}_2)$.*

– Le système (S) peut s'interpréter sous forme matricielle :

$$(\mathrm{S}) \iff \mathrm{AX} = \mathrm{B} \text{ où } \mathrm{X} = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} \in \mathcal{M}_{p,1}(\mathbb{K}),\ \mathrm{B} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}, \text{ et } \mathrm{A} = \begin{pmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{np} \end{pmatrix}$$

A $\in \mathcal{M}_{n,p}(\mathbb{K})$ est la **matrice du système** et B $\in \mathcal{M}_{n,1}(\mathbb{K})$ le second membre.

– Le système (S) peut s'interpréter sous forme vectorielle :

$$(\mathrm{S}) \iff x_1\mathcal{C}_1(\mathrm{A}) + \dots + x_p\mathcal{C}_p(\mathrm{A}) = \mathrm{B}$$

c'est une équation vectorielle dans $\mathcal{M}_{n,1}(\mathbb{K})$, où on cherche toutes les combinaisons linéaires des matrices colonnes de A qui sont égales à B.

Résolution d'un système linéaire.

– Dans le cas général on applique la méthode du pivot de Gauss. Une opération élémentaire sur les lignes ou les colonnes du système donne un nouveau système équivalent au premier (les opérations élémentaires sont des opérations inversibles, on peut revenir en arrière). On obtient à la fin du processus, un système triangulaire équivalent au premier, et on résout ce dernier en remontant.

– Dans le cas particulier d'un système carré, si la matrice A du système est inversible, alors le système possède une unique solution qui est $\mathrm{X} = \mathrm{A}^{-1} \times \mathrm{B}$.

2) Structures des solutions d'un système linéaire

Théorème 16.8 (Solutions d'un système homogène)

Soit A $\in \mathcal{M}_{n,p}(\mathbb{K})$, l'ensemble des solutions du système homogène $\mathrm{AX} = \mathrm{O}_{n,1}$ est un sous-groupe de $(\mathcal{M}_{p,1}(\mathbb{K}), +)$, de plus si X est solution, alors $\forall \lambda \in \mathbb{K}$, $\lambda \cdot \mathrm{X}$ est aussi solution, on dit que l'ensemble des solutions est stable pour le produit par les scalaires. Cet ensemble est noté ker(A) *:*

$$\ker(\mathrm{A}) = \left\{\mathrm{X} \in \mathcal{M}_{p,1}(\mathbb{K}) \ /\ \mathrm{AX} = \mathrm{O}_{n,1}\right\}.$$

Preuve : Celle-ci est simple et laissée en exercice. □

Théorème 16.9 (Structure des solutions)

Soit A $\in \mathcal{M}_{n,p}(\mathbb{K})$ et B $\in \mathcal{M}_{n,1}(\mathbb{K})$. Le système linéaire AX = B admet au moins une solution si et seulement si B est combinaison linéaire des colonnes de A. Si c'est le cas, et si $\mathrm{X}_0 \in \mathcal{M}_{p,1}(\mathbb{K})$ est une solution (i.e. $\mathrm{AX}_0 = \mathrm{B}$), alors l'ensemble des solutions de AX = B est $\{\mathrm{X}_0 + \mathrm{Y} \ /\ \mathrm{Y} \in \ker(\mathrm{A})\}$, ce que l'on écrit $\mathrm{X}_0 + \ker(\mathrm{A})$.

Preuve : Le premier point est évident. Soit X_0 une solution particulière, alors soit X $\in \mathcal{M}_{p,1}(\mathbb{K})$:

$$\mathrm{AX} = \mathrm{B} \iff \mathrm{AX} = \mathrm{AX}_0 \iff \mathrm{A}(\mathrm{X} - \mathrm{X}_0) = \mathrm{O}_{n,1} \iff \mathrm{X} - \mathrm{X}_0 \in \ker(\mathrm{A}) \iff \exists \mathrm{Y} \in \ker(\mathrm{A}),\ \mathrm{X} = \mathrm{X}_0 + \mathrm{Y}$$

□

Conséquence :

Un système linéaire AX = B peut avoir :

– soit aucune solution, lorsque B n'est pas combinaison linéaire des colonnes de A,

- soit une unique solution, lorsque B est combinaison linéaire des colonnes de A et que $ker(\mathrm{A}) = \left\{\mathrm{O}_{p,1}\right\}$,
- soit une infinité de solutions, lorsque B est combinaison linéaire des colonnes de A et que $ker(\mathrm{A})$ n'est pas réduit à $\left\{\mathrm{O}_{p,1}\right\}$, car dans ce cas $\ker(\mathrm{A})$ est infini (il est stable pour le produit par un scalaire et contient au moins une colonne non nulle).

VI SOLUTION DES EXERCICES

Solution 16.1

1/ $\mathcal{T}_n^s(\mathbb{K}) = \mathrm{Vect}\left[(\mathrm{E}^{ij})_{1\leqslant i\leqslant j\leqslant n}\right]$, *la famille* $(\mathrm{E}^{ij})_{1\leqslant i\leqslant j\leqslant n}$ *est libre, c'est une base de* $\mathcal{T}_n^s(\mathbb{K})$, *on a donc* $\dim(\mathcal{T}_n^s(\mathbb{K})) = \frac{n(n+1)}{2}$.

2/ $\mathcal{T}_n^i(\mathbb{K}) = \mathrm{Vect}\left[(\mathrm{E}^{ij})_{1\leqslant j\leqslant i\leqslant n}\right]$, *la famille* $(\mathrm{E}^{ij})_{1\leqslant j\leqslant i\leqslant n}$ *est libre, c'est une base de* $\mathcal{T}_n^i(\mathbb{K})$, *on a donc* $\dim(\mathcal{T}_n^i(\mathbb{K})) = \frac{n(n+1)}{2}$.

3/ $\mathcal{S}_n(\mathbb{K}) = \mathrm{Vect}\left[(\mathrm{E}^{ij}+\mathrm{E}^{ji})_{1\leqslant i\leqslant j\leqslant n}\right]$, *la famille* $(\mathrm{E}^{ij}+\mathrm{E}^{ji})_{1\leqslant i\leqslant j\leqslant n}$ *est libre, c'est une base de* $\mathcal{S}_n(\mathbb{K})$, *on a donc* $\dim(\mathcal{S}_n(\mathbb{K})) = \frac{n(n+1)}{2}$.

4/ $\mathcal{A}_n(\mathbb{K}) = \mathrm{Vect}\left[(\mathrm{E}^{ij}-\mathrm{E}^{ji})_{1\leqslant i<j\leqslant n}\right]$, *la famille* $(\mathrm{E}^{ij}-\mathrm{E}^{ji})_{1\leqslant i<j\leqslant n}$ *est libre, c'est une base de* $\mathcal{A}_n(\mathbb{K})$, *on a donc* $\dim(\mathcal{A}_n(\mathbb{K})) = \frac{n(n-1)}{2}$.

Solution 16.2 *Soient* $\mathrm{E}^{i,j}, \mathrm{E}^{k,l}$ *deux matrices élémentaires de* $\mathcal{M}_n(\mathbb{K})$, *soient* $(r,s)\in[\![1\,;n]\!]^2$:

$$[\mathrm{E}^{i,j}\times\mathrm{E}^{k,l}]_{r,s} = \sum_{p=1}^{n}[\mathrm{E}^{i,j}]_{r,p}[\mathrm{E}^{k,l}]_{p,s} = \sum_{p=1}^{n}\delta_{i,r}\delta_{p,j}\delta_{p,k}\delta_{l,s} = \delta_{j,k}\delta_{i,r}\delta_{l,s} = \delta_{j,k}[\mathrm{E}^{i,l}]_{r,s},$$

on a donc $\mathrm{E}^{i,j}\times\mathrm{E}^{k,l} = \delta_{j,k}\mathrm{E}^{i,l}$.

Solution 16.3 *On a* $\mathrm{A} = \mathrm{I}_3 + \mathrm{J}$ *avec* $\mathrm{J} = \begin{pmatrix}0&1&0\\0&0&1\\0&0&0\end{pmatrix}$, *de plus* $\mathrm{J}^2 = \mathrm{K} = \begin{pmatrix}0&0&1\\0&0&0\\0&0&0\end{pmatrix}$, *et* $\mathrm{J}^3 = \mathrm{O}_3$. *Comme* $\mathrm{I}_3\times\mathrm{J} = \mathrm{J}\times\mathrm{I}_3$, *on peut utiliser le binôme de Newton :*

$$\mathrm{A}^n = \sum_{k=0}^{n}\binom{n}{k}\mathrm{J}^k = \mathrm{I}_3 + n\mathrm{J} + \frac{n(n-1)}{2}\mathrm{K} = \begin{pmatrix}1&n&\frac{n(n-1)}{2}\\0&1&n\\0&0&1\end{pmatrix}.$$

Solution 16.4

1/ *Posons* $\mathrm{B} = (\mathrm{A}^{-1})^{\mathrm{T}}$, *alors* $\mathrm{B}\times\mathrm{A}^{\mathrm{T}} = (\mathrm{A}\times\mathrm{A}^{-1})^{\mathrm{T}} = \mathrm{I}_n^{\mathrm{T}} = \mathrm{I}_n$, *donc* A^{T} *est inversible et son inverse est* B.

2/ *Soit* $\mathrm{X} = \begin{pmatrix}x\\y\\z\end{pmatrix}$ *et soit* $\mathrm{Y} = \begin{pmatrix}a\\b\\c\end{pmatrix}$, *résolvons l'équation* $\mathrm{AX} = \mathrm{Y}$:

$$\mathrm{AX}=\mathrm{Y} \iff \left\{\begin{array}{ccccccc} x&+&\lambda y&-&z&=&a\\ &&2y&+&z&=&b\\ x&+&&&z&=&c\end{array}\right.$$
$$\iff \left\{\begin{array}{ccccccc} x&+&\lambda y&-&z&=&a\\ &&2y&+&z&=&b\\ &-&\lambda y&+&2z&=&c-a\ (\mathrm{L}_3\leftarrow\mathrm{L}_3-\mathrm{L}_1)\end{array}\right.$$
$$\iff \left\{\begin{array}{ccccccc} x&-&z&+&\lambda y&=&a\\ &&z&+&2y&=&b\\ &&&-&(4+\lambda)y&=&c-a-2b\ (\mathrm{L}_3\leftarrow\mathrm{L}_3-2\mathrm{L}_2)\end{array}\right. .$$

D'où la discussion :

– *Si* $\lambda = -4$: *alors le système n'a pas de solution lorsque* $c-a-2b\neq 0$, *la matrice* A *n'est donc pas inversible.*

– *Si* $\lambda\neq -4$: *le système admet une unique solution qui est :*

$$\left\{\begin{array}{ccc} y&=&\frac{a+2b-c}{4+\lambda}\\ z&=&\frac{-2a+\lambda b+2c}{4+\lambda}\\ x&=&\frac{2a-\lambda b+(\lambda+2)c}{4+\lambda}\end{array}\right. .$$

Or on sait que cette unique solution est $\mathrm{X} = \mathrm{A}^{-1}\mathrm{Y}$, *on en déduit alors que :*

$$\mathrm{A}^{-1} = \frac{1}{4+\lambda}\begin{pmatrix} 2 & -\lambda & \lambda+2 \\ 1 & 2 & -1 \\ -2 & \lambda & 2 \end{pmatrix}.$$

3/ *Supposons les coefficients diagonaux tous non nuls, soit* $\mathrm{X} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ *et* $\mathrm{Y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$*, lorsqu'on résout le système* $\mathrm{TX} = \mathrm{Y}$ *(d'inconnue* X*) par substitutions remontantes, on obtient une solution de la forme :*

$$\mathrm{X} = \begin{cases} x_n & = b_{n,n}y_n \\ x_{n-1} & = b_{n-1,n-1}y_{n-1} + b_{n-1,n}y_n \\ \vdots & \\ x_1 & = b_{1,1}y_1 + \cdots + b_{1,n}y_n \end{cases}.$$

Il y a une seule solution, donc T *est inversible, on sait alors que* $\mathrm{X} = \mathrm{T}^{-1}\mathrm{Y}$*, donc les coefficients de la matrice* T^{-1} *sont les coefficients* $b_{i,j}$ *ci-dessus, ce qui prouve que* T^{-1} *est triangulaire supérieure.*

Réciproquement, si T *est inversible, alors* $\mathrm{T}^{-1}\mathrm{T} = \mathrm{I}_n$*, notons* a_{ij} *les coefficients de* T^{-1}*, alors on doit avoir :* $a_{11}a_1 = 1$*, donc* $a_1 \neq 0$*. Puis* $a_{21}a_1 = 0$ *donc* $a_{21} = 0$*, puis* $a_{22}a_2 = 1$ *donc* $a_2 \neq 0$*. On montre ainsi de proche en proche que* T^{-1} *est triangulaire supérieure et que les coefficients diagonaux de* T *sont tous non nuls.*

Solution 16.5 *On transforme la matrice obtenue précédemment en* $\mathrm{J}_{3,3,2}$ *:*

$$\mathrm{C}_2 \leftarrow \mathrm{C}_2 - 3\mathrm{C}_1, \mathrm{C}_3 \leftarrow \mathrm{C}_3 - \mathrm{C}1 \ \textit{donnent} \ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}, \mathrm{C}_3 \leftarrow \mathrm{C}_3 - 2\mathrm{C}_2 \ \textit{donne} \ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

On a obtient la matrice U *en effectuant sur lignes de* I_3 *les mêmes opérations que sur les lignes de* A *(dans le même ordre), et la matrice* V *en effectuant sur colonnes de* I_3*, les mêmes opérations sur les colonnes de* A *(dans le même ordre) :*

$$\mathrm{U} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 7 & 1 \end{pmatrix} \ \textit{et} \ \mathrm{V} = \begin{pmatrix} 0 & 1 & -2 \\ 1 & -3 & 5 \\ 0 & 0 & 1 \end{pmatrix}$$

Chapitre 17

Arithmétique des polynômes

Sommaire

I DIVISIBILITÉ

1) La division euclidienne

Théorème 17.1

Soient $A \in \mathbb{K}[X]$ *et* $B \in \mathbb{K}[X]$ ***non nul****, il existe un unique couple de polynômes* (Q, R) *tel que* $A = BQ + R$ *avec* $\deg(R) < \deg(B)$*,* Q *est appelé le quotient, et* R *le reste.*

Définition 17.1

Soient $A, B \in \mathbb{K}[X]$*, on dit que* B *divise* A *lorsqu'il existe un polynôme* Q *tel que* $A = Q \times B$*, notation* $B|A$*.*

Remarque 17.1 – *On définit ainsi une relation dans* $\mathbb{K}[X]$*, on peut vérifier que celle - ci est réflexive, transitive, mais elle n'est ni symétrique, ni antisymétrique. Plus précisément,* $B|A$ *et* $A|B$ *ssi il existe* $\lambda \in \mathbb{K}^*$ *tel que* $A = \lambda B$ *(on dit que* A *et* B *sont* ***associés****).*

Théorème 17.2

– *Si* $B \neq 0$*, alors* $B|A$ *si et seulement si le reste de la division euclidienne de* A *par* B *est nul.*
– *Si* $A \neq 0$ *et* $B|A$*, alors* $\deg(B) \leqslant \deg(A)$.
– *Si* $B|A$ *et* $B|C$*, alors* $\forall\, U, V \in \mathbb{K}[X], B|A \times U + C \times V$.

Preuve : Celle-ci est simple et laissée en exercice. □

Remarque 17.2 – *Il découle du dernier point que si* $B|A - C$ *et* $B|D - E$*, alors* $B|(A + D) - (C + E)$ *et* $B|AD - EC$*, en particulier, si* $B|A - C$ *alors* $\forall\, n \in \mathbb{N}, B|A^n - C^n$.

Notation : Soit $P \in \mathbb{K}[X]$, on note $P\mathbb{K}[X]$ l'ensemble des multiples de P : $P\mathbb{K}[X] = \{P \times Q \ /\ Q \in \mathbb{K}[X]\}$.

On vérifie facilement la propriété suivante :

$(P\mathbb{K}[X], +)$ est un groupe commutatif et $\forall B \in \mathbb{K}[X], \forall U \in P\mathbb{K}[X], BU \in P\mathbb{K}[X]$.

2) Congruences

Définition 17.2 (congruences)

Soient $A, B, P \in \mathbb{K}[X]$*, on dit que* A *est congru à* B *modulo* P *lorsque* $P \mid A - B$.
Notation : $A \equiv B \pmod P$.

Théorème 17.3

• *La relation de congruence modulo* P *est une relation d'équivalence.*
• *Soient* $A, B, C, D, P \in \mathbb{K}[X]$*, si* $A \equiv B \pmod P$ *et* $C \equiv D \pmod P$ *alors :*

$$AC \equiv BD \pmod P \text{ et } A + C \equiv B + D \pmod P.$$

On dit que la relation de congruence est compatible avec les opérations.

3) Diviseurs communs

Définition 17.3 (polynôme normalisé)

Soit A *un polynôme non nul, on appelle* A *normalisé le polynôme noté* A *obtenu en divisant* A *par son coefficient dominant, ce polynôme est donc unitaire et associé à* A*. C'est l'unique polynôme unitaire associé à* A.

Définition 17.4 (diviseurs communs)

Pour $A \in \mathbb{K}[X]$*, on note* D_A *l'ensemble des diviseurs de* A*. Cet ensemble contient toujours* $\mathbb{K}^*$.
On notera $D_{A,B} = D_A \cap D_B$ *l'ensemble des diviseurs communs à* A *et* B.

Remarque 17.3 –

– *Si* $A \neq 0$*, alors* D_A *est un ensemble infini, mais* $\{\deg(P) \ /\ P \in D_A\}$ *est fini car inclus dans* $[\![0\,;\deg(A)]\!]$.
– $D_0 = \mathbb{K}[X]$*, si* $\lambda \in \mathbb{K}^*$, $D_\lambda = \mathbb{K}^*$.
– *Si* A *est non nul,* $D_A = D_A$.

Théorème 17.4

Soient $A, B, Q, R \in \mathbb{K}[X]$*, si* $A = BQ + R$*, alors* $D_A \cap D_B = D_B \cap D_R$.

Application – Le théorème ci-dessus fournit un algorithme pour la recherche des diviseurs communs à A et B basé sur la division euclidienne : c'est **l'algorithme d'***Euclide*, rappelons son principe :

On remarque que si $B = 0$ alors $D_{A,B} = D_A$. On peut supposer désormais que $B \neq 0$ et on cherche à calculer $D = D_{A,B}$:

Étape 1 : on effectue la division euclidienne de A par B : $A = BQ_1 + R_1$ avec $\deg(R_1) < \deg(B)$. On a $D = D_{B,R_1}$, donc si $R_1 = 0$ alors $D = D_B$, sinon on passe à l'étape 2 :

Étape 2 : on effectue la division euclidienne de B par R_1 : $B = R_1Q_2 + R_2$ avec $\deg(R_2) < \deg(B)$. On a $D = D_{R_1,R_2}$, donc si $R_2 = 0$ alors $D = D_{R_1}$, sinon on passe à l'étape 3 ...

La suite des degrés des restes obtenus est une suite strictement décroissante d'entiers positifs, elle est donc nécessairement finie, *i.e.* il existe nécessairement un entier $n \geqslant 1$ tel que $R_n = 0$, l'ensemble cherché est donc $D = D_{R_{n-1}}$ (avec la convention $R_0 = B$).

II ÉLÉMENTS PREMIERS ENTRE EUX

1) Théorème de Bézout

Définition 17.5

Soient $A, B \in \mathbb{K}[X]$, on dit que a et b sont premiers entre eux (ou A est premier avec B) lorsque le seul diviseur commun unitaire est 1, i.e. $D_{A,B} = \mathbb{K}^*$.

Remarque 17.4 –

- *Dire que A est premier avec B revient à dire que le dernier reste non nul dans l'algorithme d'Euclide est égal à 1 une fois normalisé.*
- *Si A est premier avec B, alors au moins un des deux est non nul (sinon l'ensemble des diviseurs communs est $\mathbb{K}[X]$).*
- *A est premier avec A si et seulement si $A \in \mathbb{K}^*$.*

Théorème 17.5 (théorème de Bézout)

Soient $A, B \in \mathbb{K}[X]$, alors A et B sont premiers entre eux si et seulement si il existe $U, V \in \mathbb{K}[X]$ tels que $AU + BV = 1$. Les polynômes U et V sont appelés coefficients de Bézout (non uniques en général).

Preuve : C'est l'algorithme d'Euclide étendu, comme dans $\mathbb{Z}$. □

★**Exercice 17.1** *Soient $A = X^3 + 1$ et $B = X^2 + 1$. Montrer que A et B sont premiers entre eux, et déterminer une relation de Bézout.*

2) Conséquences

Théorème 17.6

- *Si A est premier avec B et si A est premier avec C, alors A est premier avec le produit BC. On en déduit que si A est premier avec $C_1, \ldots, C_n$, alors A est premier avec le produit $C_1 \times \ldots \times C_n$.*
- *Si A est premier avec C, si $A \mid B$ et si $C \mid B$, alors $AC \mid B$.*
- *Si $A \mid BC$ et si A est premier avec C, alors $A \mid B$.*

Preuve : Identique à celle dans $\mathbb{Z}$. □

III LE PLUS GRAND DIVISEUR COMMUN

1) Définition

Soient $A, B \in \mathbb{K}[X]$ non tous deux nuls, on sait que $D_{A,B} = D_R$ où R est le dernier reste non nul dans l'algorithme d'Euclide, on voit que les diviseurs communs à A et B ont un degré inférieur ou égal à celui de R. Soit D un diviseur commun de même degré que R, alors comme $D \mid R$ on a $R = \lambda Q$ avec $\lambda \in \mathbb{K}^*$, on en déduit que les polynômes R et D **normalisés** sont égaux.

Définition 17.6

*Soient $A, B \in \mathbb{K}[X]$ non tous deux nuls, le pgcd de A et de B le plus grand diviseur commun **unitaire**. Notation : $\text{pgcd}(A, B)$ ou $A \wedge B$, c'est le dernier reste non nul dans l'algorithme d'Euclide, **une fois normalisé**.*

Remarque 17.5 – *Il en découle que deux éléments A et B de $\mathbb{K}[X]$, non tous deux nuls, sont premiers entre eux si et seulement si $\text{pgcd}(A, B) = 1$. On remarquera au passage qu'un pgcd entre deux polynômes est unitaire.*

Théorème 17.7

Soient $A, B \in \mathbb{K}[X]$ *non tous deux nuls, et* D *un polynôme unitaire, alors* $D = \text{pgcd}(A, B)$ *si et seulement si* $D \mid A$, $D \mid B$ *et il existe deux polynômes* U *et* V *tels que* $D = AU + BV$.

Preuve : Si $D = \text{pgcd}(A, B)$ cela découle de l'algorithme d'Euclide étendu.

Si D est diviseur commun et si on a la relation, alors tout diviseur commun de A et B divise D et a donc un degré inférieur ou égal à celui de D, comme D est unitaire, c'est le pgcd de A et B. □

Théorème 17.8 (Calcul pratique d'un pgcd)

Si $A, B \in \mathbb{K}[X]$ *sont non tous deux nuls alors* $\forall Q \in \mathbb{K}[X], \text{pgcd}(A, B) = \text{pgcd}(A - BQ, B)$.

★**Exercice 17.2** *Calculer le pgcd entre* $X^4 - 1$ *et* $X^{10} - 1$.

2) Propriétés

Théorème 17.9 (caractérisations du pgcd)

Soient $A, B \in \mathbb{K}[X]$ *non tous deux nuls, et soit* $D \in \mathbb{K}[X]$ *non nul et unitaire. On a alors :*

$D = \text{pgcd}(A, B) \iff \exists\, U, V \in \mathbb{K}[X]$ ***premiers entre eux*** *tels que* $A = DU$ *et* $B = DV$.

Preuve : Si $D = \text{pgcd}(A, B)$ alors il existe $U, V \in \mathbb{K}[X]$ tels que $A = DU$ et $B = DV$, de plus il existe des polynômes U_1 et V_1 tels que $D = AU_1 + BV_1$, *i.e.* $D = DUU_1 + DVV_1$, D étant non nul et $\mathbb{K}[X]$ intègre, on en déduit que $1 = UU_1 + VV_1$ et donc U et V sont premiers entre eux.

Si $A = DU, B = DV$ avec $U \wedge V = 1$, alors D est un diviseur commun à A et B, d'après le théorème de Bézout, il existe $\alpha, \beta \in \mathbb{K}[X]$ tels que $\alpha U + \beta V = 1$, d'où $D = \alpha A + \beta B$, comme D est unitaire, on a $D = A \wedge B$. □

Théorème 17.10 (quelques propriétés du pgcd)

Soient $A, B \in \mathbb{K}[X]$ *non tous deux nuls :*

a) $\forall\, P \in \mathbb{K}[X]$, *si* $P \mid A$ *et* $P \mid B$, *alors* $P \mid \text{pgcd}(A, B)$.

b) $\text{pgcd}(A, B) = \text{pgcd}(A, B)$.

c) $\forall\, K \in \mathbb{K}[X]$, *unitaire,* $\text{pgcd}(KA, KB) = K\,\text{pgcd}(A, B)$.

d) $\forall\, n \in \mathbb{N}, \text{pgcd}(A^n, B^n) = \text{pgcd}(A, B)^n$.

e) Si A *et* C *sont premiers entre eux, alors* $\text{pgcd}(A, BC) = \text{pgcd}(A, B)$.

Preuve : Pour le premier point : soit $D = \text{pgcd}(A, B)$, alors $D_{A,B} = D_D$ donc tout diviseur commun à A et B est un diviseur de D.

Pour le deuxième point : soit $D = \text{pgcd}(A, B)$, alors il existe $U, V \in \mathbb{K}[X]$ premiers entre eux tels que $A = DU$ et $B = DV$, d'où $KA = KAU$ et $KB = KDV$, donc $KD = \text{pgcd}(KA, KB)$ (KD est unitaire).

Pour le reste la preuve est identique à celle dans $\mathbb{Z}$. □

Théorème 17.11

Soient A *et* B *deux polynômes non tous deux nuls :*

- *les racines communes à* A *et* B *dans* $\mathbb{K}$, *sont exactement les racines dans* $\mathbb{K}$ *de* $\text{pgcd}(A, B)$.
- A *et* B *sont premiers entre eux si et seulement si ils n'ont pas de racine commune dans* $\mathbb{C}$.

Preuve : Soit $a \in \mathbb{K}$, alors $A(a) = B(a) = 0 \iff X - a \mid A$ et $X - a \mid B \iff X - a \mid \text{pgcd}(A, B)$. Si A et B sont premiers entre eux, alors $\text{pgcd}(A, B) = 1$ qui est sans racine dans $\mathbb{C}$, donc A et B n'ont pas de racine commune dans $\mathbb{C}$.

Si A et B n'ont pas de racine commune dans $\mathbb{C}$, alors leur pgcd n'a pas de racine dans $\mathbb{C}$, or $\mathbb{C}$ est algébriquement clos (théorème de D'Alembert Gauss), donc D est nécessairement constant, comme il est unitaire, $D = 1$.
□

3) Généralisation

Soient A, B, C trois polynômes non tous nuls, l'ensemble des diviseurs communs à A, B et C est :

$$D_{A,B,D} = D_A \cap D_B \cap D_C = (D_A \cap D_B) \cap D_C = D_A \cap (D_B \cap D_C)$$

or on sait que $D_A \cap D_B = D_{A\wedge B}$, donc $D_{A,B,C} = D_{(A\wedge B)\wedge C} = D_{A\wedge(B\wedge C)}$. Ces deux polynômes étant unitaires, on a $(A \wedge B) \wedge C = A \wedge (B \wedge C)$ et ce polynôme est le plus grand (en degré) diviseur unitaire commun à A, B et C. Par définition ce nombre est $\boxed{\text{le pgcd de A, B et C, on le note : } \mathrm{pgcd}(A, B, C)}$.

Théorème 17.12 (associativité du pgcd)
Soient A, B, C *trois polynômes avec* B *non nul, alors* $\mathrm{pgcd}(A, B, C) = (A \wedge B) \wedge C = A \wedge (B \wedge C)$.

À retenir
L'associativité du pgcd permet de ramener le calcul au cas de deux polynômes.

Notons $D_1 = A \wedge B$ et $R = \mathrm{pgcd}(A, B, C)$, alors $R = D_1 \wedge C$, donc il existe deux polynômes U_1 et W tels que $R = D_1U_1 + CW$, de même, il existe deux polynômes U_2 et V_1 tels que $D_1 = AU_2 + BV_1$, d'où en remplaçant, $R = AU_2U_1 + BV_1U_1 + cW = AU + BV + CW$ avec $U, V, W \in \mathbb{K}[X]$.

Réciproquement, si R est un diviseur commun unitaire, et si $R = AU + BV + CW$, alors il est facile de voir que tout diviseur commun à A, B et C est un diviseur de R et donc $R = \mathrm{pgcd}(A, B, C)$, d'où le théorème :

Théorème 17.13
Soient A, B, C *trois polynômes non tous nuls et* R *unitaire, alors :*
$$R = \mathrm{pgcd}(A, B, C) \iff R \in D_{A,B,C} \text{ et } \exists U, V, W \in \mathbb{K}[X], R = AU + BV + CW.$$

Définition 17.7
Soient A, B, C *trois polynômes non tous nuls, on dira que ces trois polynômes sont :*
- ***premiers entre eux dans leur ensemble*** *lorsque* $\mathrm{pgcd}(A, B, C) = 1$.
- ***premiers entre eux deux à deux*** *lorsque* $\mathrm{pgcd}(A, B) = \mathrm{pgcd}(B, C) = \mathrm{pgcd}(A, C) = 1$.

Attention !
Les deux notions ne sont pas équivalentes, ***la deuxième entraîne la première mais la réciproque est fausse*** *comme le montre l'exemple suivant :*
$\mathrm{pgcd}((X+1)X, (X+1)(X+2), X(X+2)) = 1$ *mais* $\mathrm{pgcd}(X(X+1), (X+1)(X+2)) = X+1$, $\mathrm{pgcd}(X(X+1), X(X+2)) = X$ *et* $\mathrm{pgcd}((X + 1)(X + 2), X(X + 2)) = X + 2$.

Il découle du théorème précédent :

Théorème 17.14 (de Bézout)
Soient A, B, C *trois polynômes non tous nuls, alors* A, B *et* C *sont premiers entre eux dans leur ensemble si et seulement si :*
$$\exists U, V, W \in \mathbb{K}[X], AU + BV + CW = 1.$$

Théorème 17.15 (caractérisation)
Soient A, B, C *trois polynômes non tous nuls et* R *unitaire, alors :*
$$R = \mathrm{pgcd}(A, B, C) \iff \exists U, V, W \in \mathbb{K}[X], A = RU, B = RV \text{ et } C = RW \text{ avec } \mathrm{pgcd}(U, V, W) = 1.$$

Preuve : Si $R = \mathrm{pgcd}(A, B, C)$ alors il existe $\exists U, V, W \in \mathbb{K}[X]$, $A = RU$, $B = RV$ et $C = RW$. Il existe également des polynômes U_1, V_1 et W_1 tels que $R = AU_1 + BV_1 + CW_1$ d'où $1 = UU_1 + VV_1 + WW_1$ et donc $\mathrm{pgcd}(U, V, W) = 1$.

Réciproquement, si $A = RU$, $B = RV$ et $C = RW$ avec $\mathrm{pgcd}(U, V, W) = 1$. Il existe des polynômes U_1, V_1 et W_1 tels $1 = UU_1 + VV_1 + WW_1$, en multipliant par R il vient alors que $R = RUU_1 + RVV_1 + RWW_1 = AU_1 + BV_1 + CW_1$, ce qui entraîne que $R = \mathrm{pgcd}(A, B, C)$ (car R est unitaire et diviseur commun à A, B et C). □

Remarque 17.6 – *La notion de pgcd s'étend de la même manière à n polynômes.*

IV LE PLUS PETIT MULTIPLE COMMUN

1) Définition

Théorème 17.16

Si A *et* B *sont non nuls, il existe un unique polynôme* M *unitaire dans* $\mathbb{K}[X]$ *tel que :*

$$A\mathbb{K}[X] \cap B\mathbb{K}[X] = M\mathbb{K}[X].$$

Preuve : $\{\deg(P)\ /\ P \in A\mathbb{K}[X] \cap B\mathbb{K}[X],\ P \neq 0\}$ contient $\deg(AB)$, il existe donc un multiple commun non nul de degré minimal, quitte à le normaliser, on peut le supposer unitaire, et on le note M. Il est facile de voir que $M\mathbb{K}[X] \subset A\mathbb{K}[X] \cap B\mathbb{K}[X]$. Si $P \in A\mathbb{K}[X] \cap B\mathbb{K}[X]$, on effectue la division de P par M, $P = MQ + R$ avec $\deg(R) < \deg(M)$, d'où $R = P - MQ$, on vérifie alors que R est aussi dans $A\mathbb{K}[X] \cap B\mathbb{K}[X]$. Si $R \neq 0$ alors $\deg(R) \geqslant \deg(M)$ car M est de degré minimal, ceci est absurde, donc $R = 0$ et $P = MQ$, d'où $A\mathbb{K}[X] \cap B\mathbb{K}[X] \subset M\mathbb{K}[X]$, et finalement on a bien l'égalité.

Si $M'\mathbb{K}[X] = M\mathbb{K}[X]$ avec M' unitaire, alors M et M' se divisent mutuellement, ils sont donc associés et unitaires d'où $M = M'$. □

Il découle de ce théorème que C est un multiple commun à A et B si et seulement si $C \in A\mathbb{K}[X] \cap B\mathbb{K}[X]$, ce qui équivaut à $C \in M\mathbb{K}[X]$, c'est à dire $M \mid C$. Ceci entraîne en particulier : $\deg(M) \leqslant \deg(C)$.

Définition 17.8

Soit $A, B \in \mathbb{K}[X]$, *non nuls, et soit* $M \in \mathbb{K}[X]$ ***unitaire****, on dit que* M *est le ppcm de* A *et* B *lorsque* $A\mathbb{K}[X] \cap B\mathbb{K}[X] = M\mathbb{K}[X]$. *Notation :* $M = \text{ppcm}(A, B)$ *ou encore* $M = A \vee B$.

Théorème 17.17 (caractérisation du ppcm)

Soient $A, B \in \mathbb{K}[X]$, *non nuls, et soit* $M \in \mathbb{K}[X]$ *unitaire alors :*

$$M = \text{ppcm}(A, B) \iff \exists\, U, V \in \mathbb{K}[X] \text{ \textbf{premiers entre eux} tels que } M = AU = BV.$$

Preuve : On suppose $A, B \in \mathbb{K}[X]$, non nuls.

Si $M = \text{ppcm}(A, B)$: alors $A \mid M$ et $B \mid M$. Donc il existe $U, V \in \mathbb{K}[X]$ tels que $M = AU = BV$, soit $D = \text{pgcd}(U, V)$ alors il existe $\alpha, \beta \in \mathbb{K}[X]$ premiers entre eux tels que $U = D\alpha$ et $V = D\beta$, d'où $M = AD\alpha = BD\beta$, mais alors $M_0 = A\alpha = B\beta$ est un multiple commun à A et B donc $\deg(M) \leqslant \deg(M_0)$ ce qui entraîne $D = 1$ (car D est unitaire et $M = M_0D$).

Si $\exists U, V \in \mathbb{K}[X]$ premiers entre eux tels que $M = AU = BV$, alors $A \mid M$ et $B \mid M$, il existe $\alpha, \beta \in \mathbb{K}[X]$ tels que $U\alpha + V\beta = 1$, soit M_0 un multiple commun non nul, alors $M_0 = M_0U\alpha + M_0V\beta$, on en déduit que $M \mid M_0$ et donc $\deg(M) \leqslant \deg(M_0)$, ce qui prouve que $M = \text{ppcm}(A, B)$. □

2) Propriétés

Théorème 17.18

Soient $A, B \in \mathbb{K}[X]$, *non nuls :*

a) $\forall\, P \in \mathbb{K}[X]$, *si* $A \mid P$ *et* $B \mid P$ *alors* $\text{ppcm}(A, B) \mid P$.

b) Si A *et* B *sont premiers entre eux, alors* $\text{ppcm}(A, B) = AB$.

c) $\forall\, K \in \mathbb{K}[X]$, *unitaire,* $\text{ppcm}(KA, KB) = K\,\text{ppcm}(A, B)$.

d) $\text{ppcm}(A, B) \times \text{pgcd}(A, B) = AB$.

e) $\forall\, n \in \mathbb{N}, \text{ppcm}(A^n, B^n) = \text{ppcm}(A, B)^n$.

Preuve : Analogue au cas des entiers. □

V POLYNÔMES IRRÉDUCTIBLES, DÉCOMPOSITION

1) Définition

Définition 17.9

Un polynôme $P \in \mathbb{K}[X]$ est dit ***irréductible*** *sur $\mathbb{K}$ lorsque P est non constant et que ses seuls diviseurs unitaires sont 1 et P. L'ensemble des éléments irréductibles normalisés de $\mathbb{K}[X]$ est noté $\mathcal{I}_{\mathbb{K}[X]}$.*

☞**Exemples** :

- Tout polynôme de degré 1 est irréductible, donc $\forall \lambda \in \mathbb{K}, X + \lambda \in \mathcal{I}_{\mathbb{K}[X]}$.
- Tout polynôme de degré 2 **sans racine dans** $\mathbb{K}$ est irréductible dans $\mathbb{K}[X]$. Cependant cette propriété ne se généralise pas au delà du degré 2, par exemple : $X^4 + 1$ est sans racine dans $\mathbb{R}$, mais ce polynôme est réductible car $X^4 + 1 = (X^2 - \sqrt{2}X + 1)(X^2 + \sqrt{2}X + 1)$.
- La notion de polynôme irréductible dépend du corps $\mathbb{K}$, par exemple, $X^2 + 1$ est irréductible dans $\mathbb{R}[X]$, mais pas dans $\mathbb{C}[X]$. De même, le polynôme $X^2 - 2$ est irréductible dans $\mathbb{Q}[X]$, mais pas dans $\mathbb{R}[X]$.

Théorème 17.19

Dans $\mathbb{C}[X]$, les polynômes irréductibles unitaires sont les polynômes unitaires de degré 1, c'est à dire :

$$\mathcal{I}_{\mathbb{C}[X]} = \{X + a \ / \ a \in \mathbb{C}\}.$$

Dans $\mathbb{R}[X]$, les polynômes irréductibles sont les polynômes unitaires de degré 1, plus les polynômes unitaires de degré 2 ***sans racines réelles****. C'est à dire :*

$$\mathcal{I}_{\mathbb{R}[X]} = \{X + a \ / \ a \in \mathbb{R}\} \cup \left\{X^2 + pX + q \ / \ p, q \in \mathbb{R}, p^2 - 4q < 0\right\}.$$

Preuve : Pour $\mathbb{C}[X]$ cela découle du théorème de *D'Alembert*.

Dans $\mathbb{R}[X]$: les polynômes annoncés sont bien irréductibles unitaires. Soit $P \in \mathcal{I}_{\mathbb{K}[X]}$, avec $\deg(P) \geqslant 2$ alors P admet des racines complexes, et celles-ci sont non réelles (P est irréductible de degré supérieur à 1), soit α l'une d'elles, alors $\overline{\alpha}$ est également racine de P (et distincte de α), donc dans $\mathbb{C}[X]$ le polynôme P est divisible par $(X - \alpha)(X - \overline{\alpha}) = X^2 + pX + q \in \mathbb{R}[X]$ avec $p^2 - 4q < 0$. Mais alors P est divisible dans $\mathbb{R}[X]$ par $X^2 + pX + q$ (unicité du quotient et du reste), or $P \in \mathcal{I}_{\mathbb{K}[X]}$, donc nécessairement $P = X^2 + pX + q$. □

Propriétés élémentaires :

a) Si P est irréductible, alors pour tout polynôme Q, soit $P \mid Q$ soit $\text{pgcd}(P, Q) = 1$.
Preuve : Soit $D = \text{pgcd}(P, Q)$, $D \mid P$ donc $D = 1$ ou $D = P$. □

b) Si $P \in \mathbb{K}[X]$ est non constant, alors P possède au moins un diviseur irréductible.
Preuve : Soit $B = \{\deg(D) \ / \ D \mid P \text{ et } D \notin \mathbb{K}^*\}$, alors B est une partie de $\mathbb{N}$ non vide ($\deg(P) \in B$), soit Q un diviseur de P avec $\deg(Q) \in B$ **minimal**, si $D \mid Q$ avec D unitaire et $D \neq 1$, alors $D \mid P$ et donc $\deg(D) \in B$, d'où $\deg(D) \geqslant \deg(Q)$, or $D \mid Q$, donc $\deg(D) \leqslant \deg(Q)$ et finalement $\deg(D) = \deg(Q)$, d'où $D = Q$ et donc Q est irréductible. □

c) L'ensemble $\mathcal{I}_{\mathbb{K}[X]}$ est infini, puisque tout polynôme $X + a$ où $a \in \mathbb{K}$ est irréductible unitaire.

d) Si P est irréductible et si $P \mid AB$, alors $P \mid A$ ou $P \mid B$.
Preuve : Supposons que P ne divise pas A, alors $\text{pgcd}(P, A) = 1$ et par conséquent $P \mid B$ (d'après le théorème de Gauss). □

2) Décomposition en facteurs irréductibles

Théorème 17.20 (décomposition en produit de facteurs irréductibles)

Tout élément $Q \in \mathbb{K}[X]$ non constant, est un produit d'éléments irréductibles. Plus précisément, il existe $r \geqslant 1$, il existe $P_1, \ldots, P_r \in \mathcal{I}_{\mathbb{K}[X]}$, il existe des entiers $\alpha_1, \ldots, \alpha_r \in \mathbb{N}^*$, il existe $\lambda \in \mathbb{K}^*$ tels que :

$$Q = \lambda \times P_1^{\alpha_1} \times P_2^{\alpha_2} \times \ldots \times P_r^{\alpha_r}.$$

Preuve : On a $Q = \lambda \times Q$ avec λ le coefficient dominant de Q. On se ramène ainsi au cas où Q est unitaire.

Par récurrence sur $\deg(Q)$: pour $\deg(Q) = 1$ il n'y a rien à montrer. Supposons le théorème démontré jusqu'au rang k, si $\deg(Q) = k + 1$ alors Q admet au moins un diviseur irréductible unitaire P, donc $Q = PR$, si $R = 1$ alors Q est irréductible, sinon R est un produit de facteurs irréductibles (HR), donc Q aussi. □

Théorème 17.21 (unicité de la décomposition)
Si $Q \in \mathbb{K}[X]$ s'écrit sous la forme : $Q = \lambda \times P_1^{\alpha_1} \times \ldots \times P_r^{\alpha_r} = \mu \times Q_1^{\beta_1} \times \ldots \times Q_s^{\beta_s}$, avec $P_1, \ldots, P_r \in \mathcal{I}_{\mathbb{K}[X]}, \alpha_1, \ldots, \alpha_r \in \mathbb{N}^*, Q_1, \ldots, Q_s \in \mathcal{I}_{\mathbb{K}[X]}, \beta_1, \ldots, \beta_s \in \mathbb{N}^*$, et $\lambda, \mu \in \mathbb{K}^*$, alors $r = s$, $\lambda = \mu$ et il existe une permutation σ de $[\![1\,;r]\!]$ telle que pour $i \in [\![1\,;r]\!], P_i = Q_{\sigma(i)}, \alpha_i = \beta_{\sigma(i)}$. La décomposition est unique [à l'ordre près].

Preuve : Identique à celle des entiers. □

À retenir
Décomposition en irréductibles dans $\mathbb{C}[X]$, si $P \in \mathbb{C}[X]$ est non constant alors :
$$P = \text{dom}(P) \prod_{k=1}^{r} (X - a_k)^{m_k}$$
où $a_1, \ldots, a_r$ sont les racines distinctes de P, et $m_1, \ldots, m_r$ les multiplicités respectives.

À retenir
Décomposition en irréductibles dans $\mathbb{R}[X]$, si $P \in \mathbb{R}[X]$ est non constant alors :
$$P = \text{dom}(P) \prod_{k=1}^{r} (X - a_k)^{m_k} \times \prod_{k=1}^{s} \left(X^2 + p_k X + q_k\right)^{\alpha_k}$$
où $a_1, \ldots, a_r$ sont les racines réelles distinctes de P, $m_1, \ldots, m_r$ les multiplicités respectives, et où les facteurs de degré 2 sont sans racines réelles.

3) Notion de P-valuation

Si Q est un polynôme non nul et P un polynôme irréductible, alors l'ensemble $\left\{k \in \mathbb{N} \;/\; P^k \mid Q\right\}$ est non vide (contient 0) et majoré par deg(Q), cet ensemble admet donc un maximum :

Définition 17.10
Soit $P \in \mathcal{I}_{\mathbb{K}[X]}$ et Q polynôme non nul, on appelle P-valuation de Q, notée $v_P(Q)$, le plus grand entier k tel que $P^k \mid Q$. La définition s'étend au polynôme nul en posant $v_P(0) = +\infty$.

Remarque 17.7 :
- $v_P(Q) = k \iff P^k \mid Q$ *et* $P^{k+1} \nmid Q \iff \exists\, T \in \mathbb{K}[X], Q = P^k T$ *avec* $P \nmid T$.
- $v_P(Q) \geqslant 1 \iff P \mid Q$. *$v_P(Q)$ est l'exposant de P dans la décomposition de Q en facteurs irréductibles.*
- $\left\{k \in \mathbb{N} \;/\; P^k \mid Q\right\} = [\![0\,;v_P(Q)]\!]$.

À retenir
Si Q est non constant, la décomposition de Q s'écrit : $Q = \lambda_Q \prod_{P \in \mathcal{I}_{\mathbb{K}[X]}} P^{v_P(Q)}$.

En effet, seul un nombre fini de valuations sont non nulles (les autres donnent un facteur égal à 1).

Théorème 17.22 (Propriétés)
Soient Q, R deux polynômes, on a :
a) $\forall P \in \mathcal{I}_{\mathbb{K}[X]}, v_P(QR) = v_P(Q) + v_P(R)$.
b) $\forall P \in \mathcal{I}_{\mathbb{K}[X]}, v_P(Q + R) \geqslant \min(v_P(Q); v_P(R))$.
c) $Q \mid R \iff \forall P \in \mathcal{I}_{\mathbb{K}[X]},\ v_P(Q) \leqslant v_P(R)$.
d) Si Q et R sont non nuls alors $\forall P \in \mathcal{I}_{\mathbb{K}[X]}$:
$$v_P(Q \wedge R) = \min(v_P(Q); v_P(R)) \text{ et } v_P(Q \vee R) = \max(v_P(Q); v_P(R)).$$

Preuve : Analogue au cas des entiers. □

À retenir : formules du pgcd et du ppcm

Il découle du théorème ci-dessus que :
$$\mathrm{pgcd}(Q,R) = \prod_{P\in\mathcal{I}_{\mathbb{K}[X]}} P^{\min(v_P(Q);v_P(R))} \text{ et } \mathrm{ppcm}(Q,R) = \prod_{P\in\mathcal{I}_{\mathbb{K}[X]}} P^{\max(v_P(Q);v_P(R))}.$$

4) Applications

Comme dans $\mathbb{Z}$:

– Si P est non constant, alors la décomposition de P en produit de facteurs irréductibles permet de trouver tous les diviseurs de P.

– Si P, Q sont non constants, alors à partir de leur décomposition en produit de facteurs irréductibles, on peut calculer $\mathrm{pgcd}(P,Q)$ et $\mathrm{ppcm}(P,Q)$.

★**Exercice 17.3** *Dans* $\mathbb{C}[X]$, *montrer que pour* $n,m \in \mathbb{N}^*$, *on a* $\mathrm{pgcd}(X^n-1, X^m-1) = X^d - 1$ *où* $d = \mathrm{pgcd}(n,m)$.

VI SOLUTION DES EXERCICES

Solution 17.1 *L'algorithme d'Euclide étendu donne un dernier reste non nul égal à 2 avec la relation :* $2 = (X+1)A - (X^2+X-1)B$, *il suffit de tout diviser par 2.*

Solution 17.2 *En appliquent l'algorithme d'Euclide, le dernier reste non nul normalisé est* $D = X^2 - 1$ *qui est donc le pgcd.*

Solution 17.3 *Il existe* $u, v \in \mathbb{Z}$ *tels que* $nu + mv = d$, *on en déduit que* $z^n = z^m = 1$ *si et seulement si* $z^d = 1$, *les racines du pgcd sont donc les racines* d^{es} *de l'unité, comme les racines de* $X^n - 1$ *sont simples, celles du pgcd le sont aussi, ce qui donne le résultat.*

Chapitre 18

Fractions rationnelles

Sommaire

I CONSTRUCTION DE L'ENSEMBLE DES FRACTIONS RATIONNELLES

Le corps $\mathbb{K}$ désigne un sous-corps de $\mathbb{C}$, *i.e.* un corps inclus dans $\mathbb{C}$.

1) Définition d'une fraction rationnelle

Dans l'ensemble $\mathbb{K}[X] \times (\mathbb{K}[X] \setminus \{0\}) = \{(P, Q) \ / \ P, Q \in \mathbb{K}[X], Q \neq 0\}$, on définit la relation $\mathcal{R}$ en posant :

$$(P, Q)\mathcal{R}(R, S) \iff P \times S = Q \times R.$$

On vérifie que la relation $\mathcal{R}$ est une relation d'équivalence dans $\mathbb{K}[X] \times (\mathbb{K}[X] \setminus \{0\})$. La transitivité de la relation utilise l'intégrité de $\mathbb{K}[X]$.

Définition 18.1

On appelle fraction rationnelle à coefficients dans $\mathbb{K}$ toute classe d'équivalence pour la relation

$\mathcal{R}$. La classe de (P, Q) est notée $\frac{P}{Q}$ [avec P le numérateur et Q le dénominateur], on a donc :

$$\frac{P}{Q} = \{(R, S) \in \mathbb{K}[X] \times (\mathbb{K}[X] \setminus \{0\}) \ / \ PS = QR\}.$$

*On dit que (P, Q) est un **représentant** de la fraction $\frac{P}{Q}$. L'ensemble des fractions rationnelles est noté $\mathbb{K}(X)$ et la relation $\mathcal{R}$ est appelée **égalité des fractions rationnelles**.*

2) Représentant irréductible

Définition 18.2

Soit $F = \frac{P}{Q}$ une fraction, on dit que $\frac{P}{Q}$ est un représentant irréductible lorsque $\text{pgcd}(P, Q) = 1$ et que Q est unitaire.

☞**Exemple** : Soit $F = \frac{X^3-1}{X^2-1}$, un représentant irréductible est $\frac{X^2+X+1}{X+1}$, c'est à dire $F = \frac{X^2+X+1}{X+1}$.

Théorème 18.1

Toute fraction admet un représentant irréductible unique.

Preuve : Celle-ci est simple et laissée en exercice. □

3) Opérations sur les fractions

Définition 18.3 (addition, multiplication, produit par un scalaire)

Soient $\frac{P}{Q}$, $\frac{R}{S}$ deux fractions rationnelles et soit $\lambda \in \mathbb{K}$, on pose :

$$\frac{P}{Q} + \frac{R}{S} = \frac{PS + QR}{QS}, \ \frac{P}{Q} \times \frac{R}{S} = \frac{PR}{QS}, \text{ et } \lambda\frac{P}{Q} = \frac{\lambda P}{Q}.$$

Pour que la définition ait un sens il faut le résultat ne dépende pas des représentants choisis pour les fractions, c'est à dire si $\frac{P}{Q} = \frac{P'}{Q'}$ et $\frac{R}{S} = \frac{R'}{S'}$, alors :

$$\frac{PS + QR}{QS} = \frac{P'S' + Q'R'}{Q'S'} \ ; \ \frac{PR}{QS} = \frac{P'R'}{Q'S'} \text{ et } \lambda\frac{P}{Q} = \lambda\frac{P'}{Q'}.$$

Cette vérification est simple et laissée en exercice.

Propriétés :

a) Pour l'addition :
  - elle est associative, commutative,
  - elle admet un élément neutre, la fraction $\frac{0}{Q}$ ($\forall Q \neq 0$), appelée fraction nulle. On remarquera qu'une fraction est nulle ssi son numérateur est nul.
  - toute fraction $\frac{P}{Q}$ admet un opposé et $-\frac{P}{Q} = \frac{-P}{Q} = \frac{P}{-Q}$.

b) Pour la multiplication :
  - elle est associative, commutative,
  - elle admet un élément neutre qui est la fraction $\frac{P}{P}$ ($\forall P \neq 0$), appelée fraction unité.
  - toute fraction $\frac{P}{Q}$ non nulle (*i.e.* $P \neq 0$) admet un inverse, et $\left(\frac{P}{Q}\right)^{-1} = \frac{Q}{P}$.
  - elle est distributive sur l'addition.

c) Pour le produit par un scalaire : $\forall \lambda, \mu \in \mathbb{K}, \forall F, G \in \mathbb{K}(X)$:

$$1.F = F \ ; \quad \lambda.(F + G) = \lambda.F + \lambda.G \ ; \quad (\lambda + \mu).F = \lambda.F + \lambda.G \ ; \quad \lambda.(\mu).F = (\lambda\mu).F$$

et

$$\lambda.(F \times G) = (\lambda.F) \times G = F \times (\lambda.G).$$

À retenir

Par conséquent, $(\mathbb{K}(X), +, \times)$ est un corps commutatif et $(\mathbb{K}(X), +\times, .)$ est une $\mathbb{K}$-algèbre commutative.

Théorème 18.2 (plongement des polynômes dans $\mathbb{K}(X)$)

L'application $\phi : \mathbb{K}[X] \to \mathbb{K}(X)$ définie par $\phi(P) = \frac{P}{1}$ est un morphisme d'algèbres injectif.

Preuve : Celle-ci est simple et laissée en exercice. □

Par conséquent on peut identifier le polynôme P avec la fraction $\frac{P}{1}$, ce qui fait que l'on peut considérer que $\mathbb{K}[X] \subset \mathbb{K}(X)$. En particulier la fraction nulle (en vertu de l'égalité des fractions) est identifiée au polynôme nul 0, et la fraction unité est identifiée au polynôme constant 1.

II DEGRÉ, PÔLES ET RACINES D'UNE FRACTION

1) Notion de degré

Soit F une fraction non nulle et $\frac{P}{Q}$, $\frac{R}{S}$ deux représentants de F (*i.e.* $F = \frac{P}{Q} = \frac{R}{S}$), on a donc $PS = QR$, d'où $\deg(P) - \deg(Q) = \deg(R) - \deg(S)$. Autrement dit, la différence entre le degré du numérateur et le degré du dénominateur, ne dépend pas du représentant de F, mais seulement de F.

Définition 18.4 (degré d'une fraction)

Soit $F = \frac{P}{Q}$ une fraction, on pose $\deg(F) = -\infty$ si $F = 0$, et $\deg(F) = \deg(P) - \deg(Q)$ sinon. Le degré d'une fraction est donc un élément de $\mathbb{Z} \cup \{-\infty\}$.

Remarque 18.1 – *Soit P un polynôme, en tant que polynôme son degré est $\deg(P)$, mais en tant que fraction, son degré est $\deg(\frac{P}{1}) = \deg(P) - \deg(1) = \deg(P)$, on trouve bien la même chose.*

☞**Exemple** : $\deg(\frac{X^2+X+1}{X+1}) = 1$ et $\deg(\frac{X}{X^3-X^2+2}) = -2$.

Théorème 18.3 (propriétés du degré)

Soient $F, G \in \mathbb{K}(X)$, on a : $\deg(F+G) \leqslant \max(\deg(F), \deg(G))$, et $\deg(F \times G) = \deg(F) + \deg(G)$. On retrouve les mêmes propriétés que pour les polynômes.

Preuve : Posons $F = \frac{P}{Q}$ et $G = \frac{R}{S}$, alors $F \times G = \frac{PR}{QS}$, donc $\deg(F \times G) = \deg(PR) - \deg(QS) = \deg(P) - \deg(Q) + \deg(R) - \deg(S) = \deg(F) + \deg(G)$. De même, $\deg(F+G) = \deg(PS+QR) - \deg(QS)$, or $\deg(PS+QR) \leqslant \max(\deg(PS), \deg(QR))$, donc on a $\deg(F+G) \leqslant \deg(PS) - \deg(QS)$ **ou** $\deg(F+G) \leqslant \deg(QR) - \deg(QS)$, c'est à dire $\deg(F+G) \leqslant \deg(F)$ ou $\deg(F+G) \leqslant \deg(G)$, finalement, $\deg(F+G) \leqslant \max(\deg(F), \deg(G))$. □

Remarque 18.2 :

- *Une fraction rationnelle constante non nulle a un degré nul, mais la réciproque est fausse, par exemple :* $F = \frac{X}{X+1}$.
- *Si* $\deg(F) \neq \deg(G)$ *alors* $\deg(F+G) = \max(\deg(F), \deg(G))$.
- *Une fraction F est nulle ssi son degré vaut $-\infty$.*

2) Pôles et racines

Définition 18.5

Soit $F \in \mathbb{K}(X)$ non nulle, et soit $\frac{P}{Q}$ un représentant ***irréductible*** *de F. On dit que $a \in \mathbb{K}$ est racine de F de multiplicité $m \in \mathbb{N}^*$ lorsque a est racine du numérateur P de multiplicité m. On dit que $a \in \mathbb{K}$ est pôle de F de multiplicité $m \in \mathbb{N}^*$ lorsque a est racine du dénominateur Q de multiplicité m.*

Remarque 18.3 :

- *Puisque* $\frac{P}{Q}$ *est irréductible, on voit qu'un scalaire* a *ne peut pas être à la fois pôle et racine de* F*, sinon* P *et* Q *seraient divisibles par* $X - a$*.*
- a *est un pôle de* F *de multiplicité* $m \in \mathbb{N}^*$ *revient à dire que* a *est racine de multiplicité* m *de* $\frac{1}{F}$*.*
- *Par exemple, la fraction* $F = \frac{X^3-1}{X^2-1}$ *possède deux racines complexes simples* j *et* j^2*, un pôle simple* -1*, mais pas de racine réelle.*
- *Si* F *est une fraction qui admet une infinité de racines, alors* F *est la fraction nulle.*

3) Dérivation d'une fraction rationnelle

Soit $F \in \mathbb{K}(X)$ une fraction rationnelle et $\frac{P}{Q} = \frac{R}{S}$ deux représentants de F, on a $PS = QR$, d'où $(P'Q - PQ')S^2 = P'QS^2 - PQ'S^2 = P'QS^2 - Q'QRS = QS(P'S - Q'R)$, mais en dérivant la relation polynomiale $PS = QR$ on obtient $P'S + PS' = Q'R + QR'$, d'où $(P'Q - PQ')S^2 = QS(QR' - PS') = Q^2SR' - QPSS' = Q^2SR' - Q^2RS' = Q^2(SR' - RS')$, ce qui traduit l'égalité des fractions :

$$\frac{P'Q - PQ'}{Q^2} = \frac{R'S - RS'}{S^2}.$$

Définition 18.6

Soit $F = \frac{P}{Q} \in \mathbb{K}(X)$*, on appelle* fraction dérivée *de* F *la fraction notée* F' *(ou* $\frac{dF}{dX}$*) définie par :*

$$F' = \frac{P'Q - PQ'}{Q^2},$$

Le résultat ne dépend pas du représentant de F *choisi. On définit également les dérivées successives de* F *en posant* $F^{(0)} = F$ *et pour tout* $n \in \mathbb{N}, F^{(n+1)} = \left(F^{(n)}\right)'$*.*

Remarque 18.4 –

- *Soit* P *un polynôme, la dérivée de* P *en tant que fraction rationnelle est* $\left(\frac{P}{1}\right)' = \frac{P'1-P1'}{1^2} = P'$*, on retrouve bien la dérivée de* P *en tant que polynôme.*
- *Contrairement aux polynômes le degré de* F' *n'est pas toujours égal à* $\deg(F) - 1$*, par exemple :* $F = \frac{X}{X+1}$*, on a* $\deg(F) = 0$ *et* $F' = \frac{1}{(X+1)^2}$ *donc* $\deg(F') = -2$*.*
  Par contre on a toujours $\deg(F') \leqslant \deg(F) - 1$*.*

★**Exercice 18.1** *Montrer que si* $F' = 0$ *alors* F *est une fraction constante.*

Théorème 18.4 (propriétés)

On retrouve les propriétés usuelles de la dérivation avec les formules usuelles : $(F + G)' = F' + G'; (F \times G)' = F' \times G + F \times G'; (\lambda.F)' = \lambda.F'; \left(\frac{1}{F}\right)' = \frac{-F'}{F^2}$*, et la formule de Leibniz :*

$$(F \times G)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} F^{(k)} \times G^{(n-k)}.$$

Preuve : Celle-ci est simple et laissée en exercice. □

Définition 18.7 (Dérivée logarithmique)

Soit F *une fraction non nulle, la* dérivée logarithmique *de* F *est la fraction* $\frac{F'}{F}$*.*

Théorème 18.5

Si F *est une fraction non nulle qui se factorise en* $F = F_1 \times \cdots \times F_n$ *dans* $\mathbb{K}(X)$ *avec* $(n \in \mathbb{N}^*)$*, alors :*

$$\frac{F'}{F} = \frac{F_1'}{F_1} + \cdots + \frac{F_n'}{F_n}.$$

Preuve : Par récurrence sur n en commençant par le cas $n = 2$. Si $F = F_1F_2$ alors $F' = F_1'F_2 + F_1F_2'$, d'où $\frac{F'}{F} = \frac{F_1'F_2+F_1F_2'}{F_1F_2} = \frac{F_1'}{F_1} + \frac{F_2'}{F_2}$. Le passage du rang n au rang $n+1$ se ramène au cas $n = 2$. □

III DÉCOMPOSITION D'UNE FRACTION RATIONNELLE

1) Partie entière

Soit $F = \frac{A}{B}$ une fraction, on effectue la division euclidienne de A par B : $A = BQ + R$ avec $\deg(R) < \deg(B)$. On a alors $F = Q + \frac{R}{B}$ avec $\deg(\frac{R}{B}) < 0$ et $Q \in \mathbb{K}[X]$. Supposons qu'il existe un autre polynôme S et une fraction G tels que $F = S + G$ avec $\deg(G) < 0$, alors $\deg(Q - S) = \deg(G - \frac{R}{B}) < 0$ donc $Q = S$ car ce sont des polynômes, et $G = \frac{R}{B}$. On peut donc énoncer :

Théorème 18.6

*Soit $F \in \mathbb{K}(X)$, il existe un **unique polynôme** Q tel que $\deg(F - Q) < 0$, celui-ci est appelé **partie entière** de F, c'est le quotient dans la division euclidienne du numérateur de F par le dénominateur.*

À retenir

Si $\deg(F) < 0$ alors la partie entière de F est nulle (à cause de l'unicité).

2) Éléments simples

Définition 18.8

*Un élément simple de $\mathbb{K}(X)$ est une fraction du type $\frac{A}{B^n}$ où B est un **polynôme irréductible unitaire** (i.e. $B \in I_{\mathbb{K}[X]}$), $\deg(A) < \deg(B)$, et $n \geqslant 1$.*

- Éléments simples dans $\mathbb{C}(X)$:
  on sait que $I_{\mathbb{C}[X]} = \{X - a \ / \ a \in \mathbb{C}\}$, donc les éléments simples de $\mathbb{C}(X)$ sont les fractions :

$$\frac{\alpha}{(X-a)^n} \text{ avec } \alpha, a \in \mathbb{C} \text{ et } n \geqslant 1.$$

- Éléments simples de $\mathbb{R}(X)$:
  on sait que $I_{\mathbb{R}[X]} = \{X - a \ / \ a \in \mathbb{R}\} \cup \{X^2 + pX + q \ / \ p, q \in \mathbb{R}, p^2 - 4q < 0\}$, donc les éléments simples de $\mathbb{R}(X)$ sont de deux types :
  - éléments simples de première espèce :

$$\frac{\alpha}{(X-a)^n} \text{ avec } \alpha, a \in \mathbb{R} \text{ et } n \geqslant 1.$$

  - éléments simples de seconde espèce :

$$\frac{aX+b}{(X^2+pX+q)^n} \text{ avec } a, b, p, q \in \mathbb{R}, p^2 - 4q < 0, \text{ et } n \geqslant 1.$$

Définition 18.9

Décomposer une fraction rationnelle F non nulle, c'est l'écrire comme somme de sa partie entière et d'éléments simples.

☞**Exemples** :
- $F = \frac{X^3}{X^2+1}$, sa partie entière est X, et on a $F = X + \frac{-X}{X^2+1}$: c'est la décomposition de F en éléments simples dans $\mathbb{R}(X)$, mais pas dans $\mathbb{C}(X)$.
- Dans $\mathbb{C}(X)$ on a : $F = X + \frac{-1/2}{X+i} + \frac{-1/2}{X-i}$.

3) Existence de la décomposition

Soit F une fraction non nulle et non polynomiale : $F = \frac{A}{B}$ (**forme irréductible**), on calcule sa partie entière : E, on a alors $= E + \frac{R}{B}$ avec $\deg(\frac{R}{B}) < 0$, on est alors amené à décomposer une fraction de degré strictement négatif en éléments simples.

On factorise le dénominateur B en produit de polynômes irréductibles unitaires : $B = \prod_{i=1}^{r} P_i^{m_i}$ (B est unitaire).

Théorème 18.7

Si T, S *sont deux polynômes* ***premiers entre eux*** *et si* $\deg(\frac{A}{TS}) < 0$*, alors il existe deux polynômes* U *et* V *tels que :*

$$\frac{A}{TS} = \frac{U}{T} + \frac{V}{S} \text{ avec } \deg(U) < \deg(T), \deg(V) < \deg(S).$$

Preuve : Il existe deux polynômes U', V' tels que $U'S + V'T = 1$ (théorème de Bézout), on a alors $\frac{A}{TS} = \frac{AU'}{T} + \frac{AV'}{S}$, soit E_1 la partie entière de $\frac{AU'}{T}$ et E_2 celle de $\frac{AV'}{S}$, il existe deux polynômes U et V tels que $\frac{AU'}{T} = E_1 + \frac{U}{T}$ avec $\deg(U) < \deg(T)$, et $\frac{AV'}{S} = E_2 + \frac{V}{S}$ avec $\deg(V) < \deg(S)$, d'où : $\frac{A}{TS} = E_1 + E_2 + \frac{U}{T} + \frac{V}{S}$, mais $\deg(\frac{U}{T} + \frac{V}{S}) < 0$, donc $E_1 + E_2$ est la partie entière de $\frac{A}{TS}$, or celle-ci est nulle, donc $E_1 + E_2 = 0$, ce qui donne le résultat. □

Conséquence : Par récurrence on en déduit que si $B_1, \ldots, B_n$ sont premiers entre eux deux à deux et si $\deg(\frac{A}{B_1 \times \ldots \times B_n}) < 0$, alors il existe des polynômes $U_1, \ldots, U_n$ tels que :

$$\frac{A}{B_1 \times \ldots \times B_n} = \sum_{i=1}^{n} \frac{U_i}{B_i} \text{ avec } \deg(U_i) < \deg(B_i).$$

En appliquant ceci à notre fraction F, on peut affirmer qu'il existe des polynômes $(U_i)_{1 \leqslant i \leqslant r}$ tels que :

$$F = E + \sum_{i=1}^{r} \frac{U_i}{P_i^{m_i}} \text{ avec } \deg(U_i) < \deg[P_i^{m_i}].$$

Théorème 18.8

Si T *est un polynôme irréductible unitaire et si* $\deg(\frac{A}{T^n}) < 0$ *(*$n \geqslant 1$*), alors il existe des polynômes* $V_1, \ldots, V_n$ *tels que :*

$$\frac{A}{T^n} = \sum_{k=1}^{n} \frac{V_k}{T^k} \text{ avec } \deg(V_k) < \deg(T).$$

C'est une décomposition en éléments simples.

Preuve : Par récurrence sur n : pour $n = 1$ il n'y a rien à faire. Si le théorème est vrai au rang n et si $\deg(\frac{A}{T^{n+1}}) < 0$, alors on effectue la division euclidienne de A par T : $A = QT + V_{n+1}$ avec $\deg(V_{n+1}) < \deg(T)$, ce qui donne $\frac{A}{T^{n+1}} = \frac{Q}{T^n} + \frac{V_{n+1}}{T^{n+1}}$, il est facile de voir que $\deg(\frac{Q}{T^n}) < 0$, on peut donc appliquer l'hypothèse de récurrence, ce qui donne le résultat. □

On peut appliquer ce théorème à chacune des fractions $\frac{U_i}{P_i^{m_i}}$: il existe des polynômes $V_{1,i}, \ldots, V_{m_i,i}$ tels que :

$$\frac{U_i}{P_i^{m_i}} = \sum_{j=1}^{m_i} \frac{V_{j,i}}{P_i^j} \text{ avec } \deg(V_{j,i}) < \deg(P_i).$$

Ce qui donne pour F :

$$F = E + \sum_{i=1}^{r} \left[\sum_{j=1}^{m_i} \frac{V_{j,i}}{P_i^j} \right].$$

C'est une décomposition de F en éléments simples.

Théorème 18.9 (admis)

La décomposition en éléments simples est unique.

Théorème 18.10 (décomposition de $\frac{P'}{P}$)
Soit P un polynôme non nul et $P = \lambda P_1^{m_1} \times \cdots \times P_n^{m_n}$ sa décomposition en facteurs irréductibles unitaires, alors $\frac{P'}{P} = \frac{m_1P_1'}{P_1} + \cdots + \frac{m_nP_n'}{P_n}$ (décomposition en éléments simples).

Preuve : Découle de la propriété de la dérivée logarithmique. □

IV DÉCOMPOSITION DANS LE CAS COMPLEXE

1) Forme de la décomposition

Soit $F = \frac{A}{B} \in \mathbb{C}(X)$, **sous forme irréductible**, soit E sa partie entière et soit $B = \prod_{k=1}^{r}(X - a_k)^{m_k}$ la factorisation du dénominateur. Les complexes a_k sont **les pôles** de F, et les entiers m_k ($\geqslant 1$) sont **les multiplicités respectives**.

D'après l'étude générale, la forme de la décomposition de F sera :

$$F = E + \sum_{k=1}^{r}\left[\sum_{j=1}^{m_k}\frac{b_{j,k}}{(X-a_k)^j}\right].$$

Chaque pôle de F va donc générer des éléments simples qui lui correspondent : ce sont les $\frac{b_{j,k}}{(X-a_k)^j}$ pour $j \in [\![1\,;m_k]\!]$.

Définition 18.10 (partie polaire)
*La somme des éléments simples relatifs au pôle a_k est appelée **partie polaire** de F relative au pôle a_k, elle est notée $P_F(a_k)$.*

On a donc $P_F(a_k) = \sum_{j=1}^{m_k}\frac{b_{j,k}}{(X-a_k)^j}$, et la forme de la décomposition de F est :

$$F = E + P_F(a_1) + \cdots + P_F(a_r).$$

C'est à dire : $\boxed{\text{partie entière plus les parties polaires relatives aux pôles de F.}}$

La décomposition dans $\mathbb{C}(X)$ consiste donc à calculer des parties polaires.

2) Calcul d'une partie polaire

Soit $F = \frac{A}{B} \in \mathbb{C}(X)$ (sous forme irréductible) et soit $a \in \mathbb{C}$ un pôle de F de multiplicité $m \geqslant 1$.

– **Cas d'un pôle simple** : on prend $m = 1$. On peut écrire $B = (X-a)Q$ avec $Q(a) \neq 0$. Comme $m = 1$, la partie polaire de F relative à a est $P_F(a) = \frac{c}{X-a}$, en regroupant les parties polaires relatives **aux autres pôles**, on peut écrire $F = E + \frac{c}{X-a} + \frac{U}{V}$ avec E la partie entière et $\deg(\frac{U}{V}) < 0$. En multipliant par $X - a$ on obtient : $\frac{A}{Q} = (X-a)E + c + (X-a)\frac{U}{V}$, mais a n'étant pas un pôle de $\frac{U}{V}$, on peut évaluer en a, ce qui donne : $c = \frac{A(a)}{Q(a)}$. Comme $B = (X-a)Q$, il est facile de voir que $Q(a) = B'(a)$, en conclusion :

Si a est un pôle simple de $F = \frac{A}{B}$, alors la partie polaire de F relative à a est : $P_F(a) = \frac{c}{X-a}$ avec $c = \frac{A(a)}{B'(a)} = \frac{A(a)}{Q(a)}$ où Q est tel que $B = (X-a)Q$.

☞ **Exemple** : Soit $F = \frac{1}{X^n-1}$ avec $n \geqslant 1$. On a $\deg(F) < 0$ donc la partie entière est nulle. Les pôles de F sont les racines n-ièmes de l'unité : $a_k = \exp(2ik\pi/n)$ avec $k \in [\![0\,;n-1]\!]$, et ce sont des pôles simples. La partie polaire de F relative à a_k est $\frac{c_k}{X-a_k}$ avec $c_k = \frac{1}{na_k^{n-1}} = \frac{a_k}{n}$. La décomposition de F est :

$$\frac{1}{X^n-1} = \sum_{k=0}^{n-1}\frac{a_k}{n(X-a_k)}.$$

– **Cas d'un pôle double** : on prend $m = 2$, on peut écrire $B = (X-a)^2Q$ avec $Q(a) \neq 0$, la partie polaire de F relative à a est $P_F(a) = \frac{\alpha}{X-a} + \frac{\beta}{(X-a)^2}$, en regroupant les parties polaires relatives aux autres pôles, on obtient :

$$F = E + \frac{\alpha}{X-a} + \frac{\beta}{(X-a)^2} + \frac{U}{V} \text{ avec } \deg(\frac{U}{V}) < 0.$$

Si on multiplie le tout par $(X-a)^2$ et que l'on évalue en a (a n'est pas un pôle de $\frac{U}{V}$), on obtient $\beta = \frac{A(a)}{Q(a)}$.
Pour obtenir α, on peut poser $G = F - \frac{\beta}{(X-a)^2}$, on a alors $G = E + \frac{\alpha}{X-a} + \frac{U}{V}$, donc a est un pôle simple de G, ce qui nous ramène au cas précédent.
Autre méthode : on pose $H = (X-a)^2 \times F = \frac{A}{Q}$, on a en fait $H = (X-a)^2E + \alpha(X-a) + \beta + (X-a)^2\frac{U}{V}$, en évaluant en a on trouve $\beta = H(a)$, et en évaluant la dérivée en a, on trouve $\alpha = H'(a)$. En conclusion :

> Si a est un pôle double de $F = \frac{A}{B}$, alors la partie polaire de F relative à a est :
>
> $P_F(a) = \frac{\alpha}{X-a} + \frac{\beta}{(X-a)^2}$ avec $\beta = H(a)$ et $\alpha = H'(a)$, en posant $H = (X-a)^2 \times F$.

Remarque 18.5 – *Cette autre méthode se généralise au cas d'un pôle a de multiplicité $m \geqslant 3$ en posant* $H = (X-a)^m \times F$.

☞**Exemple** : Soit $F = \frac{X^6}{(X-1)^2(X^3+1)}$.
La fraction est irréductible et son degré vaut 1, il y a donc une partie entière non nulle, on trouve $E = X + 2$ (le dénominateur est égal à $X^5 - 2X^4 + X^3 + X^2 - 2X + 1$). La fraction possède 4 pôles :
• 1 : c'est un pôle double, on pose $H = (X-1)^2 \times F = \frac{X^3}{X^3+1}$, la partie polaire relative à 1 est :

$$P_F(1) = \frac{9/4}{X-1} + \frac{1/2}{(X-1)^2}.$$

Car $H(1) = 1/2$ et $H'(1) = 9/4$.
• -1 : c'est un pôle simple, la partie polaire de F relative à -1 est :

$$P_F(-1) = \frac{1/12}{X+1}.$$

• $-j$: c'est un pôle simple, la partie polaire relative à $-j$ est :

$$P_F(-j) = \frac{1/3}{X+j}.$$

• $-j^2$: est un pôle simple, la partie polaire relative à $-j^2$ est :

$$P_F(-j^2) = \frac{1/3}{X+j^2}.$$

Finalement, la décomposition de F en éléments simples dans $\mathbb{C}(X)$ est :

$$F = X + 2 + \frac{9/4}{X-1} + \frac{1/2}{(X-1)^2} + \frac{1/12}{X+1} + \frac{1/3}{X+j} + \frac{1/3}{X+j^2}.$$

3) Cas particuliers

– Si F est à coefficients réels alors :

les parties polaires relatives aux pôles conjugués, sont conjuguées.

Preuve : Si a est un pôle complexe non réel de F de multiplicité m, alors on sait que $\overline{a}$ est un pôle de F de même multiplicité car $F \in \mathbb{R}(X)$, en regroupant les parties polaires relatives aux pôles autres que a, on obtient : $F = E + P_F(a) + \frac{U}{V}$, où $E \in \mathbb{R}[X]$ est la partie entière, si on conjugue l'expression, alors on obtient : $F = E + \overline{P_F(a)} + \frac{\overline{U}}{\overline{V}}$. Si on pose $P_F(a) = \sum_{k=1}^{m} \frac{c_k}{(X-a)^k}$, alors $\overline{P_F(a)} = \sum_{k=1}^{m} \frac{\overline{c_k}}{(X-\overline{a})^k}$ et donc :

$$F = E + \sum_{k=1}^{m} \frac{\overline{c_k}}{(X-\overline{a})^k} + \frac{\overline{U}}{\overline{V}},$$

mais $\overline{a}$ n'est pas un pôle de $\frac{\overline{U}}{\overline{V}}$, donc $\overline{P_F(a)}$ est la partie polaire de F relative à $\overline{a}$, *i.e.* $\overline{P_F(a)} = P_F(\overline{a})$. □

☞**Exemple** : Soit $\mathrm{F} = \frac{1}{(\mathrm{X}^2+\mathrm{X}+1)^2}$, $\deg(\mathrm{F}) < 0$ donc sa partie entière est nulle. F possède deux pôles doubles j et j^2. La partie polaire relative au pôle j est : $\mathrm{P_F}(j) = \frac{\mathrm{H}'(j)}{\mathrm{X}-j} + \frac{\mathrm{H}(j)}{(\mathrm{X}-j)^2}$ en posant $\mathrm{H} = (\mathrm{X}-j)^2 \times \mathrm{F} = \frac{1}{(\mathrm{X}-j^2)^2}$, on obtient $\mathrm{H}(j) = -1/3$ et $\mathrm{H}'(j) = -\frac{2i\sqrt{3}}{9}$. F étant à coefficients réels, la partie polaire relative à j^2 est la conjuguée de celle relative à j, la décomposition de F est donc :

$$\mathrm{F} = \frac{-1}{3(\mathrm{X}-j)^2} - \frac{2i\sqrt{3}}{9(\mathrm{X}-j)} + \frac{-1}{3(\mathrm{X}-\bar{j})^2} + \frac{2i\sqrt{3}}{9(\mathrm{X}-\bar{j})}.$$

– Si F est paire ou impaire, alors en utilisant la relation entre $\mathrm{F(X)}$ et $\mathrm{F(-X)}$ et avec l'unicité de la décomposition, on obtient des relations entre les coefficients à déterminer dans les parties polaires.

☞**Exemple** : Soit $\mathrm{F} = \frac{\mathrm{X}^4+1}{\mathrm{X}(\mathrm{X}^2-1)^2}$.
$\deg(\mathrm{F}) < 0$ donc la partie entière est nulle. La fraction est irréductible, impaire, et possède un pôle simple : 0, et deux pôles doubles : 1 et -1. La forme générale de la décomposition de F est :

$$\mathrm{F} = \frac{a}{\mathrm{X}} + \frac{b}{\mathrm{X}-1} + \frac{c}{(\mathrm{X}-1)^2} + \frac{d}{\mathrm{X}+1} + \frac{e}{(\mathrm{X}+1)^2}.$$

F étant impaire, on a $\mathrm{F(X)} = -\mathrm{F(-X)}$, ce qui donn e :

$$\mathrm{F} = \frac{a}{\mathrm{X}} + \frac{b}{\mathrm{X}+1} + \frac{-c}{(\mathrm{X}+1)^2} + \frac{d}{\mathrm{X}-1} + \frac{-e}{(\mathrm{X}-1)^2}.$$

L'unicité de la décomposition nous donne les relations : $\begin{cases} d &= b \\ e &= -c \end{cases}$, ce qui fait deux coefficients en moins à calculer. La partie polaire relative à 0 est $\mathrm{P_F}(0) = \frac{1}{\mathrm{X}}$ (pôle simple). En substituant 1 à X dans $(\mathrm{X}-1)^2 \times \mathrm{F}$, on obtient $c = 1/2$, et en faisant tendre x vers $+\infty$ dans la fonction rationnelle $x \mapsto x\mathrm{F}(x)$, on obtient la relation $1 = a+b+d$ *i.e.* $2b = 0$ d'où $b = 0$, finalement la décomposition de F est :

$$\mathrm{F} = \frac{1}{\mathrm{X}} + \frac{1}{2(\mathrm{X}-1)^2} - \frac{1}{2(\mathrm{X}+1)^2}.$$

V DÉCOMPOSITION DANS LE CAS RÉEL

1) Forme de la décomposition

Soit $\mathrm{F} = \frac{\mathrm{A}}{\mathrm{B}} \in \mathbb{R}(\mathrm{X})$ (sous forme irréductible), soit E sa partie entière et soit :

$$\mathrm{B} = \prod_{k=1}^{n} (\mathrm{X}-a_k)^{m_k} \times \prod_{k=1}^{r} (\mathrm{X}^2 + p_k\mathrm{X} + q_k)^{\alpha_k}$$

la factorisation de B en produit de facteurs irréductibles unitaires ($p_k^2 - 4q_k < 0$). D'après l'étude générale, la forme de la décomposition de F est :

$$\mathrm{F} = \mathrm{E} + \sum_{k=1}^{n}\left[\sum_{j=1}^{m_k} \frac{b_{j,k}}{(\mathrm{X}-a_k)^j}\right] + \sum_{k=1}^{r}\left[\sum_{j=1}^{\alpha_k} \frac{c_{j,k}\mathrm{X} + d_{j,k}}{(\mathrm{X}^2 + p_k\mathrm{X} + q_k)^j}\right].$$

La première somme est en fait la somme des parties polaires de F relatives aux pôles réels de F. Les techniques de calculs sont les mêmes dans le cas complexe.

La seconde somme est la somme des éléments simples de seconde espèce.

2) Calcul des éléments simples de seconde espèce

On se limitera au cas où $\mathrm{X}^2+p\mathrm{X}+q$ est un diviseur irréductible de B de **multiplicité** 1, en regroupant les autres éléments simples, on obtient :

$$\mathrm{F} = \mathrm{E} + \frac{a\mathrm{X}+b}{\mathrm{X}^2+p\mathrm{X}+q} + \frac{\mathrm{U}}{\mathrm{V}}.$$

Soient c et $\bar{c}$ les deux racines complexes (non réelles) de X^2+pX+q, alors c et $\bar{c}$ ne sont pas pôles de $\frac{U}{V}$, et c et $\bar{c}$ sont pôles simples de F, on peut calculer la partie polaire de F relative à c dans $\mathbb{C}(X)$: $P_F(c) = \frac{\alpha}{X-c}$, comme $F \in \mathbb{R}(X)$ on a $P_F(\bar{c}) = \frac{\bar{\alpha}}{X-\bar{c}}$, la somme de ces deux parties polaires donne : $P_F(c)+P_F(\bar{c}) = \frac{2\operatorname{Re}(\alpha)X-2\operatorname{Re}(\alpha\bar{c})}{X^2+pX+q}$, c'est un élément simple de $\mathbb{R}(X)$, comme la décomposition dans $\mathbb{R}(X)$ est unique, il en résulte que :

$$\frac{aX+b}{X^2+pX+q} = \frac{2\operatorname{Re}(\alpha)X-2\operatorname{Re}(\alpha\bar{c})}{X^2+pX+q}.$$

Autre méthode : soit $H = (X^2+pX+q)\times F$, on a : $H = (X^2+pX+q)\times E + aX + b + (X^2+pX+q)\times\frac{U}{V}$.

On obtient alors le système : $\left\{\begin{array}{rcl} H(c) & = & ac+b \\ H(\bar{c}) & = & a\bar{c}+b \end{array}\right.$, en résolvant on trouve a et b.

☞**Exemple** : Soit $F = \frac{X^4}{X^3-1}$.

On a $\deg(F) = 1$, il y a donc une partie entière non nulle, celle-ci vaut X, d'autre part on a $X^3-1 = (X-1)(X^2+X+1)$, d'où la forme de la décomposition :

$$F = X + \frac{a}{X-1} + \frac{bX+c}{X^2+X+1}.$$

La partie polaire relative à 1 est $P_F(1) = \frac{1}{3(X-1)}$. Dans $\mathbb{C}(X)$, la partie polaire relative à j est $P_F(j) = \frac{j^2}{3(X-j)}$, et la partie polaire relative à j^2 est la conjuguée, *i.e.* $P_F(j^2) = \frac{j}{3(X-j^2)}$, la somme de ces deux parties polaires donne : $\frac{-X+1}{3(X^2+X+1)}$, la décomposition de F est donc :

$$F = X + \frac{1}{3(X-1)} + \frac{-X+1}{3(X^2+X+1)}.$$

Remarque 18.6 – *En évaluant en 0 on obtient $c-a=0$ d'où $c=a=1/3$. En faisant tendre x vers $+\infty$ dans $x(F(x)-x) = \frac{x^2}{x^3-1}$, on obtient $a+b=0$ d'où $b=-a=-1/3$.*

Le résultat suivant est souvent utile lors du calcul des différents coefficients réels :

Théorème 18.11

Soit z un complexe ***non réel****, soient a, b, c et d quatre* ***réels*** *tels que $az+b = cz+d$, alors $a=c$ et $b=d$.*

Preuve : Par l'absurde, si $a \neq c$ alors on aurait $z = \frac{d-b}{c-a} \in \mathbb{R}$, or z est non réel, donc $a=c$, ce qui entraîne $b=d$. □

VI APPLICATIONS DE LA DÉCOMPOSITION

1) Calcul de la dérivée n-ième d'une fraction

☞**Exemple** : Soit $f(x) = \frac{1}{x^2+1}$, calculons $f^{(n)}(x)$. Dans $\mathbb{C}(X)$ on a $\frac{1}{X^2+1} = \frac{1}{2i(X-i)} - \frac{1}{2i(X+i)}$, d'où :

$$f^{(n)}(x) = \frac{1}{2i}\left[\frac{(-1)^n n!}{(x-i)^{n+1}} - \frac{(-1)^n n!}{(x+i)^{n+1}}\right].$$

Ce qui donne :

$$f^{(n)}(x) = (-1)^n n! \frac{\operatorname{Im}((x+i)^{n+1})}{(x^2+1)^{n+1}} = (-1)^n n! \frac{\sum\limits_{k=0}^{\lfloor \frac{n}{2} \rfloor}\binom{n+1}{2k+1}(-1)^k x^{n-2k}}{(x^2+1)^{n+1}}.$$

★**Exercice** 18.2 *Calculer la dérivée n-ième de la fonction $f(x) = \frac{x}{(x-1)(x^2+x+1)}$.*

2) Primitives d'une fraction rationnelle

Soit $F \in \mathbb{R}(X)$, on décompose F en éléments simples dans $\mathbb{R}(X)$, on est donc ramené à calculer des primitives de trois types :

- La partie entière : c'est un polynôme.
- Les éléments simples de première espèce : $\frac{1}{(X-a)^n}$ avec $n \geqslant 1$, on sait les intégrer, car :

$$\int^x \frac{dt}{(t-a)^n} = \begin{cases} \ln(|x-a|) & \text{si } n = 1 \\ \frac{-1}{(n-1)(x-a)^{n-1}} & \text{si } n \geqslant 2 \end{cases}.$$

- Les éléments simples de seconde espèce : $\frac{aX+b}{X^2+pX+q}$, pour ceux-là la méthode est la suivante :
  - on fait apparaître la dérivée du trinôme $X^2 + pX + q$ au numérateur et on compense les X en multipliant par un facteur adéquat, puis on compense les constantes en ajoutant ce qu'il faut, ce qui donne :
  $$\frac{aX+b}{X^2+pX+q} = \frac{a}{2}\frac{2X+p}{X^2+pX+q} + (b - \frac{ap}{2})\frac{1}{X^2+pX+q}.$$
  La première de ces deux fractions est facile à intégrer (du type $\frac{u'}{u}$).
  - Pour la deuxième fraction : on met le trinôme $X^2 + pX + q$ sous forme canonique afin de mettre la fraction sous la forme : $\alpha\frac{u'}{1+u^2}$ où u est une fonction de x, cette fonction est s'intègre en $\alpha \arctan(u)$.

☞**Exemple** : Calculons $F(x) = \int^x \frac{dt}{t^3+1}$ sur $]-1\,;+\infty[$:

On décompose la fraction rationnelle $\frac{1}{X^3+1}$ en éléments simples dans $\mathbb{R}(X)$, ce qui donne :

$$\frac{1}{X^3+1} = \frac{1}{3(X+1)} - \frac{X-2}{3(X^2-X+1)}.$$

On a :

$$\frac{X-2}{X^2-X+1} = \frac{1}{2}\frac{2X-1}{X^2-X+1} - \frac{3}{2}\frac{1}{X^2-X+1}$$

et :

$$\frac{1}{X^2-X+1} = \frac{1}{(X-1/2)^2+3/4} = \frac{2}{\sqrt{3}}\frac{2/\sqrt{3}}{\left(\frac{2X-1}{\sqrt{3}}\right)^2+1}.$$

On en déduit alors :

$$F(x) = \frac{1}{3}\ln(x+1) - \frac{1}{6}\ln(x^2-x+1) + \frac{\sqrt{3}}{3}\arctan(\frac{2x-1}{\sqrt{3}}).$$

C'est à dire :

$$F(x) = \frac{1}{3}\ln(\frac{x+1}{\sqrt{x^2-x+1}}) + \frac{\sqrt{3}}{3}\arctan(\frac{2x-1}{\sqrt{3}}) + \text{ cte}.$$

VII SOLUTION DES EXERCICES

Solution 18.1 *Soit $F = \frac{P}{Q}$ un représentant irréductible, $F' = 0$ entraîne $P'Q = PQ'$, mais $P \wedge Q = 1$, donc $Q \mid Q'$, d'où $Q' = 0$, donc Q est constant et unitaire, finalement $Q = 1$ et $P' = 0$, donc P est constant et F aussi.*

Solution 18.2 *Soit $F = \frac{X}{(X-1)(X^2+X+1)}$. La décomposition dans $\mathbb{C}(X)$ de F donne :*

$$F = \frac{1}{3(X-1)} + \frac{j^2}{3(X-j)} + \frac{j}{3(X-j^2)}.$$

On a donc :

$$f^{(n)}(x) = \frac{(-1)^n n!}{3}\left[\frac{1}{(x-1)^{n+1}} + 2\,\mathrm{Re}\left(\frac{j^2}{(x-j)^{n+1}}\right)\right],$$

or $\frac{j^2}{(x-j)^{n+1}} = \frac{j^2(x-j^2)^{n+1}}{(x^2+x+1)^{n+1}}$, ce qui donne finalement :

$$f^{(n)}(x) = \frac{(-1)^n n!}{3}\left[\frac{1}{(x-1)^{n+1}} + 2\frac{\sum_{k=0}^{n+1}\binom{n+1}{k}(-1)^k\cos(4(k+1)\pi/3)x^{n+1-k}}{(x^2+x+1)^{n+1}}\right].$$

Chapitre 19

Développements limités

Sommaire

I COMPARAISON DES FONCTIONS

1) Définitions

Définition 19.1

Soient $f, g : \mathrm{I} \to \mathbb{R}$ deux fonctions, et soit $a \in \mathrm{I}$ ou une extrémité de I. On dit que :

- *f est dominée par g au voisinage de a lorsqu'il existe un voisinage V de a, et une fonction $\varepsilon\colon \mathrm{V} \to \mathbb{R}$ tels que : $\forall\, x \in \mathrm{V} \cap \mathrm{I}, f(x) = g(x)\varepsilon(x)$ avec ε* ***bornée****. Notation : $f(x) = \underset{a}{\mathrm{O}}(g(x))$.*
- *f est négligeable devant g au voisinage de a lorsqu'il existe un voisinage V de a, et une fonction $\varepsilon\colon \mathrm{V} \to \mathbb{R}$ tels que : $\forall\, x \in \mathrm{V} \cap \mathrm{I}, f(x) = g(x)\varepsilon(x)$ avec $\lim\limits_{x\to a} \varepsilon(x) = 0$. Notation : $f(x) = \underset{a}{o}(g(x))$.*
- *f est équivalente à g au voisinage de a lorsqu'il existe un voisinage V de a, et une fonction $\varepsilon\colon \mathrm{V} \to \mathbb{R}$ tels que : $\forall\, x \in \mathrm{V} \cap \mathrm{I}, f(x) = g(x)\varepsilon(x)$ avec $\lim\limits_{x\to a} \varepsilon(x) = 1$. Notation : $f(x) \underset{a}{\sim} g(x)$.*

Théorème 19.1 (Caractérisations)

Lorsque la fonction g ***ne s'annule pas au voisinage de*** *a (sauf peut être en a) :*

- *$f(x) = \underset{a}{\mathrm{O}}(g(x))$ si et seulement si $\begin{cases} \frac{f}{g} \text{ est bornée au voisinage de } a \\ \text{si } a \in \mathrm{I} \text{ alors } g(a) = 0 \implies f(a) = 0 \end{cases}$.*

- $f(x) \underset{a}{=} o(g(x))$ *si et seulement si* $\begin{cases} \lim\limits_a \frac{f}{g} = 0 \\ \text{si } a \in \mathrm{I} \text{ alors } f(a) = 0 \end{cases}$.
- $f(x) \underset{a}{\sim} g(x)$ *si et seulement si* $\begin{cases} \lim\limits_a \frac{f}{g} = 1 \\ \text{si } a \in \mathrm{I} \text{ alors } g(a) = f(a) \end{cases}$.

Preuve : Celle - ci est simple et laissée en exercice. □

Remarque 19.1 –

a) $f(x) \underset{a}{=} \mathrm{O}(1)$ *signifie que la fonction f est bornée au voisinage de a.*

b) $f(x) \underset{a}{=} o(1)$ *signifie que* $\lim\limits_a f = 0$.

c) *Si* $f(x) \underset{a}{=} o(g(x))$ *alors* $f(x) \underset{a}{=} \mathrm{O}(g(x))$.

d) *Si* $f(x) \underset{a}{\sim} g(x)$ *alors* $f(x) \underset{a}{=} \mathrm{O}(g(x))$.

e) *Si* $f(x) \underset{a}{=} o(g(x))$ *et* $g(x) \underset{a}{=} o(h(x))$, *alors* $f(x) \underset{a}{=} o(h(x))$ *(transitivité).*

f) *Si* $f(x) \underset{a}{=} \mathrm{O}(g(x))$ *et* $g(x) \underset{a}{=} \mathrm{O}(h(x))$, *alors* $f(x) \underset{a}{=} \mathrm{O}(h(x))$ *(transitivité)*

g) $f(x) \underset{a}{\sim} g(x) \iff f(x) \underset{a}{=} g(x) + o(g(x))$.

Théorème 19.2

La relation « ... est équivalente à ... au voisinage de a » est une relation d'équivalence dans $\mathcal{F}(\mathrm{I}, \mathbb{R})$, c'est à dire qu'elle est réflexive, symétrique et transitive. De plus :

- *Si* $\ell \in \mathbb{R}^*$ *alors* $\lim\limits_a f = \ell$ *équivaut à* $f(x) \underset{a}{\sim} \ell$.
- *Si* $f(x) \underset{a}{=} o(g(x))$ *alors* $f(x) + g(x) \underset{a}{\sim} g(x)$.

Preuve : Celle - ci est simple et laissée en exercice. □

2) Les exemples classiques

Théorème 19.3 (croissances comparées)

Soient $\alpha, \beta \in]0; +\infty[$:

- *Si* $\alpha < \beta$ *alors* : $x^\alpha \underset{+\infty}{=} o\left(x^\beta\right)$ *et* $x^\beta \underset{0}{=} o(x^\alpha)$.
- $[\ln(x)]^\alpha \underset{+\infty}{=} o\left(x^\beta\right)$ *et* $|\ln(x)|^\alpha \underset{0}{=} o\left(\frac{1}{x^\beta}\right)$.
- $x^\alpha \underset{+\infty}{=} o\left(e^{x\beta}\right)$ *et* $x^\alpha \underset{+\infty}{=} o\left(e^{x^\beta}\right)$.
- *Si* $a > 1$ *alors* $x^\alpha \underset{+\infty}{=} o(a^x)$.

Preuve : Celle-ci est simple et laissée en exercice. □

Théorème 19.4 (les équivalents usuels)

- *Si* $f\colon]-a\,; a[\to \mathbb{R}$ *(où $a > 0$) est dérivable en 0, et si $f'(0) \neq 0$, alors* $f(x) - f(0) \underset{0}{\sim} f'(0)x$.
- $\sin(x) \underset{(0)}{\sim} x$; $e^x - 1 \underset{0}{\sim} x$; $\ln(1+x) \underset{0}{\sim} x$; $\tan(x) \underset{0}{\sim} x$; $1 - \cos(x) \underset{0}{\sim} \frac{1}{2}x^2$; $(1+x)^\alpha - 1 \underset{0}{\sim} \alpha x$.
- *Soit* $\mathrm{P}(x) = \sum\limits_{k=0}^{p} a_k x^k$ *une fonction polynomiale avec* $a_p \neq 0$, *alors* $\mathrm{P}(x) \underset{\pm\infty}{\sim} a_p x^p$ *(équivalence avec le terme de plus haut degré).*
- *Soit* $\mathrm{Q}(x) = \frac{\mathrm{P}(x)}{\mathrm{R}(x)}$ *une fraction rationnelle avec* $a_p x^p$ *le terme de plus haut degré de* P $(a_p \neq 0)$ *et* $b_r x^r$ *celui de* R $(b_r \neq 0)$, *alors* $\mathrm{Q}(x) \underset{\pm\infty}{\sim} \frac{a_p}{b_r} x^{p-r}$ *(équivalence avec le rapport des termes de plus haut degré).*

Preuve : Si f est une fonction dérivable en 0, alors il existe une fonction ε de limite nulle en 0 telle que $f(x) - f(0) = xf'(0) + x\varepsilon(x)$, si $f'(0) \neq 0$ alors $f(x) - f(0) = xf'(0)\left(1 + \frac{\varepsilon(x)}{f'(0)}\right) \underset{0}{\sim} xf'(0)$. □

Théorème 19.5 (composition ou changement de variable)
Soient $f, g\colon \mathrm{J} \to \mathbb{R}$, $h\colon \mathrm{I} \to \mathbb{R}$ *telle que* $\mathrm{Im}(h) \subset \mathrm{J}$ *et soit* $a \in \mathrm{I}$ *ou une extrémité de* I. *Si* $\lim\limits_{a} h = b$ *et si* $f(u) \underset{u\to b}{\sim} g(u)$, *alors* $f(h(x)) \underset{x\to a}{\sim} g(h(x))$.

Preuve : Celle - ci découle du théorème de composition des limites. □

Remarque 19.2 – *Pour la recherche d'un équivalent en a de* $k(x)$, *on peut toujours se ramener en 0 :*
- *Si* $a \in \mathbb{R}$, *on pose* $u = h(x) = x - a$, *on a alors* $h(x) \xrightarrow[x\to a]{} 0$, *on pose* $f(u) = k(x) = k(u+a)$. *Si* $f(u) \underset{0}{\sim} g(u)$, *alors* $f(h(x)) \underset{a}{\sim} g(h(x))$, *c'est à dire* $k(x) \underset{a}{\sim} g(x-a)$. *Dans la pratique on écrit plus simplement : posons* $u = x - a$ *(changement de variable), on a* $k(x) = k(u+a) = f(u)$ *(fonction de u), si* $f(u) \underset{0}{\sim} g(u)$, *alors* $k(x) \underset{a}{\sim} g(x-a)$.
- *Si* $a = \pm\infty$ *alors on pose* $u = h(x) = \frac{1}{x}$, *on a alors* $h(x) \xrightarrow[x\to a]{} 0$, *on pose* $f(u) = k(x) = k(\frac{1}{u})$. *Si* $f(u) \underset{0}{\sim} g(u)$, *alors* $f(h(x)) \underset{a}{\sim} g(h(x))$, *c'est à dire* $k(x) \underset{a}{\sim} g(\frac{1}{x})$. *Dans la pratique on écrit plus simplement : posons* $u = \frac{1}{x}$ *(changement de variable), on a* $k(x) = k(\frac{1}{u}) = f(u)$ *(fonction de u), si* $f(u) \underset{0}{\sim} g(u)$, *alors* $k(x) \underset{a}{\sim} g(\frac{1}{x})$.

3) Propriétés

Il découle de la définition :

Théorème 19.6
Soient $f, g : \mathrm{I} \to \mathbb{R}$ *deux fonctions, et soit* $a \in \mathrm{I}$ *ou une extrémité de* I *:*
- *Si* $f \underset{a}{\sim} g$ *alors f et g ont le même signe au voisinage de a.*
- *Si* $f \underset{a}{\sim} g$ *et si* $\lim\limits_{a} g = \ell \in \overline{\mathbb{R}}$, *alors* $\lim\limits_{a} f = \ell$.
- *Si* $f \underset{a}{\sim} g$ *et si* $h \underset{a}{\sim} k$, *alors* $f \times h \underset{a}{\sim} g \times k$ *(compatibilité avec la multiplication).*
- *Si* $f \underset{a}{\sim} g$ *et si g ne s'annule pas au voisinage de a, alors* $\frac{1}{f} \underset{a}{\sim} \frac{1}{g}$ *(compatibilité avec le passage à l'inverse).*
- *Si* $f \underset{a}{\sim} g$ *et si* $g > 0$ *au voisinage de a, alors* $f^{\alpha} \underset{a}{\sim} g^{\alpha}$ *pour tout réel* α.

Remarque 19.3 –
- *Il n'y a pas compatibilité avec l'addition en général. Par exemple :* $x + \frac{1}{x} \underset{+\infty}{\sim} x$ *et* $-x \underset{+\infty}{\sim} 1 - x$, *mais* $\frac{1}{x}$ *n'est pas équivalent à 1 au voisinage de* $+\infty$.
- *Ces propriétés sont utiles pour les calculs de limites qui ne peuvent pas être faits directement, on essaie de se ramener à un équivalent plus simple (s'il y en a ...) dont on sait calculer la limite.*

★**Exercice 19.1**

1/ *Limite en* $+\infty$ *de* $(1 + \frac{1}{x})^x$.

2/ *Calculer* $\lim\limits_{x\to 0^+} \frac{\sin(x)^x - 1}{\sqrt{x}\ln(x)}$.

4) Cas particulier : comparaison des suites

Une suite réelle étant une fonction de $\mathbb{N}$ vers $\mathbb{R}$, les définitions précédentes pour les fonctions peuvent s'étendre aux suites. La variable est en général noté n (indice) au lieu de x, et les comparaisons s'effectuent **toujours au voisinage de** $+\infty$, il est donc inutile d'indiquer le point où se fait la comparaison.

Définition 19.2
Soient (u_n), (v_n) *et* (ε_n) *trois suites telles qu'à partir d'un certain rang* $u_n = v_n\varepsilon_n$. *On dit que :*
- u_n *est dominée par* v_n *lorsque la suite* (ε_n) *est* ***bornée***. *Notation :* $u_n = \mathrm{O}(v_n)$.
- u_n *est négligeable devant* v_n *lorsque* $\varepsilon_n \to 0$. *Notation :* $u_n = o(v_n)$.
- u_n *est équivalente à* v_n *lorsque* $\varepsilon_n \to 1$. *Notation :* $u_n \sim v_n$.

Théorème 19.7 **(Caractérisations)**

*Lorsque la suite v **ne s'annule pas à partir d'un certain rang** :*
- $u_n = \mathrm{O}(v_n)$ *si et seulement si la suite* $\frac{u}{v}$ *est bornée.*
- $u_n = o(v_n)$ *si et seulement si* $\lim \frac{u_n}{v_n} = 0$.
- $u_n \sim v_n$ *si et seulement si* $\lim \frac{u_n}{v_n} = 1$.

Les propriétés sont évidemment les mêmes que pour les fonctions :

Théorème 19.8

La relation « ... est équivalente à ... » est une relation d'équivalence dans $\mathcal{F}(\mathbb{N}, \mathbb{R})$, c'est à dire qu'elle est réflexive, symétrique et transitive. De plus :
- *Si $\ell \in \mathbb{R}$ et si $u_n \sim \ell$ alors $u_n \to \ell$ (réciproque vraie lorsque $\ell \in \mathbb{R}^*$).*
- *Si $u_n = o(v_n)$ alors $\forall \lambda \in \mathbb{R}$, $\lambda u_n + v_n \sim v_n$.*
- *Si $u_n \sim v_n$ alors à partir d'un certain rang, u_n et v_n sont de même signe.*
- *Si $u_n \sim v_n$ et si $\lim v_n = \ell \in \overline{\mathbb{R}}$, alors $\lim u_n = \ell$.*
- *Si $u_n \sim v_n$ et si $a_n \sim b_n$, alors $u_n a_n \sim v_n b_n$ (compatibilité avec la multiplication). On en déduit que si $u_n \sim v_n$, alors $u_n^p \sim v_n^p$ pour tout naturel p.*
- *Si $u_n \sim v_n$ et si v ne s'annule pas à partir d'un certain rang, alors $\frac{1}{u_n} \sim \frac{1}{v_n}$ (compatibilité avec le passage à l'inverse).*
- *Si $u_n \sim v_n$ et si $v_n > 0$ à partir d'un certain rang, alors $u_n^\alpha \sim v_n^\alpha$ pour tout réel α (compatibilité avec les puissances constantes).*

Les comparaisons usuelles découlent également de celles des fonctions (au voisinage de $+\infty$) :

Théorème 19.9 **(comparaisons usuelles)**

- *Croissances comparées :soient $\alpha, \beta \in]0\,;+\infty[$:*
  - *Si $\alpha < \beta$ alors $n^\alpha = o\left(n^\beta\right)$ et $\frac{1}{n^\beta} = o\left(\frac{1}{n^\alpha}\right)$.*
  - $[\ln(n)]^\alpha = o\left(n^\beta\right)$.
  - $n^\alpha = o\left(e^{n\beta}\right)$ *et* $n^\alpha = o\left(e^{n^\beta}\right)$.
  - $\forall a \in \mathbb{R}, a^n = o(n!)$ *et donc* $n^\alpha = o(n!)$.
  - $n! = o(n^n)$.
- *Équivalents usuels : soit (u_n) une suite de **limite nulle**, alors ;*
  - *Si $f :]-a; a[\to \mathbb{R}$ (où $a > 0$) est dérivable en 0, et si $f'(0) \neq 0$, alors pour toute suite (u_n) de limite nulle, on a $f(u_n) - f(0) \sim f'(0)u_n$.*
  - $\sin(u_n) \sim u_n$; $e^{u_n} - 1 \sim u_n$; $\ln(1 + u_n) \sim u_n$; $\tan(u_n) \sim u_n$; $(1 + u_n)^\alpha - 1 \sim \alpha u_n$;
  - $1 - \cos(u_n) \sim \frac{1}{2}u_n^2$.
  - *(Stirling) $n! \sim n^n e^{-n}\sqrt{2\pi n}$ (admis).*
  - *Soit $\mathrm{P}(x) = \sum\limits_{k=0}^{p} a_k x^k$ une fonction polynomiale avec $a_p \neq 0$, alors $\mathrm{P}(n) \sim a_p n^p$ (équivalence avec le terme de plus haut degré).*
  - *Soit $\mathrm{Q}(x) = \frac{\mathrm{P}(x)}{\mathrm{R}(x)}$ une fraction rationnelle avec $a_p x^p$ le terme de plus haut degré de P $(a_p \neq 0)$ et $b_r x^r$ celui de R $(b_r \neq 0)$, alors $\mathrm{Q}(n) \sim \frac{a_p}{b_r} n^{p-r}$ (équivalence avec le rapport des termes de plus haut degré).*

Preuve : Pour les deux derniers points des croissances comparées :
• soit $u_n = \frac{|a|^n}{n!}$ avec $a \neq 0$, alors $\frac{u_{n+1}}{u_n} = \frac{|a|}{n+1} \leqslant \frac{1}{2}$ à partir d'un certain rang N, d'où pour $n \geqslant \mathrm{N}$, $0 \leqslant u_n \leqslant u_\mathrm{N} \frac{1}{2^{n-\mathrm{N}}}$ et donc $u_n \to 0$.
• soit $u_n = \frac{n!}{n^n}$ alors $0 \leqslant u_n \leqslant \frac{1}{n}$ (en écrivant que $\frac{k}{n} \leqslant 1$ pour $k > 1$).
Les équivalents usuels s'obtiennent par composition avec les équivalents usuels sur les fonctions. □

II DÉVELOPPEMENTS LIMITÉS

1) Définition

Définition 19.3

Soit $f : \mathrm{I} \to \mathbb{C}$ une fonction et soit $a \in \mathrm{I}$ ou une borne réelle de I. Soit $n \in \mathbb{N}$, on dit que f admet un développement limité d'ordre n en a (ou un $\mathrm{dl}_n(a)$) lorsqu'il existe un polynôme P de degré au plus n tel que :

$$f(x) = \mathrm{P}(x-a) + \underset{a}{o}((x-a)^n).$$

Si c'est le cas, alors le polynôme $\mathrm{P}(x-a)$ est appelé ***partie régulière*** *du $\mathrm{dl}_n(a)$.*

Remarque 19.4 –
- *On notera $\mathbb{C}_n[\mathrm{X}]$ l'ensemble des polynômes de degré au plus n et à coefficients complexes.*
- *On a $\mathrm{P}(x-a) = \sum_{k=0}^{n} a_k(x-a)^k$, dans la pratique on* ***ne développe jamais*** *les termes $(x-a)^k$.*
- *Le reste du $\mathrm{dl}_n(a)$, c'est à dire $\underset{a}{o}((x-a)^n)$ peut aussi se mettre sous la forme $(x-a)^n\varepsilon(x)$ où $\lim\limits_{x\to a}\varepsilon(x) = 0$.*
- *C'est le reste qui donne l'ordre du $\mathrm{dl}_n(a)$.*

☞**Exemples** :
- On sait que $\sin(x) \underset{0}{\sim} x$, donc $\sin(x) = x + \underset{0}{o}(x)$, c'est un $\mathrm{dl}_1(0)$ de $\sin(x)$, la partie régulière est x.
- Pour $x \neq 0$, $e^{-\frac{1}{x^2}} = \underset{0}{o}\left(x^2\right)$, c'est un $\mathrm{dl}_2(0)$, la partie régulière est nulle.

2) Formule de Taylor-Young

Définition 19.4 (polynôme de Taylor et reste)

Soit $f : \mathrm{I} \to \mathbb{C}$ une fonction, et soit $a \in \mathrm{I}$, si f possède des dérivées jusqu'à l'ordre n en a, alors on appelle polynôme de Taylor de f en a à l'ordre n, la fonction polynomiale notée $\mathrm{T}_{n,f,a}$ définie par :

$$\mathrm{T}_{n,f,a}(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k.$$

La différence $f(x) - \mathrm{T}_{n,f,a}(x)$ est notée $\mathrm{R}_{n,f,a}(x)$ est appelée ***reste*** *de f en a à l'ordre n.*

Remarque 19.5 – *Si f est $n+1$ fois dérivable en a, alors $\left[\mathrm{T}_{n+1,f,a}(x)\right]' = \mathrm{T}_{n,f',a}(x)$.*

Théorème 19.10

Si f est définie et continue au voisinage de a et si $f(x) = \underset{a}{o}((x-a)^n)$ avec $n \in \mathbb{N}$, alors au voisinage de a, on a $\int_a^x f(t)\,\mathrm{d}t = \underset{a}{o}\left((x-a)^{n+1}\right)$.

Preuve : On écrit $f(t) = (t-a)^n\alpha(t)$ avec $\lim\limits_{t\to 0}\alpha(t) = 0$. On se donne $\varepsilon > 0$, dans un voisinage V de a on aura $|f(t)| \leqslant |t-a|^n\varepsilon$, pour $x > a$ dans V, on obtient $\left|\int_a^x f(t)\,\mathrm{d}t\right| \leqslant \varepsilon\int_a^x (t-a)^n\,\mathrm{d}t = \varepsilon\frac{(x-a)^{n+1}}{n+1} < (x-a)^{n+1}\varepsilon$. De même pour $x < a$ dans V, on vérifie que $\left|\int_a^x f(t)\,\mathrm{d}t\right| < |x-a|^{n+1}\varepsilon$. Ce qui montre que $\lim\limits_{x\to a}\frac{\int_a^x f(t)\,\mathrm{d}t}{(x-a)^{n+1}} = 0$ et donc $\int_a^x f(t)\,\mathrm{d}t = \underset{a}{o}\left((x-a)^{n+1}\right)$. □

Théorème 19.11 (formule de Taylor-Young)

Soit $n \in \mathbb{N}$, soit $f : \mathrm{I} \to \mathbb{C}$ une fonction de classe $\mathcal{C}^n$ sur l'intervalle I et $a \in \mathrm{I}$, on a : $\mathrm{R}_{n,f,a}(x) = \underset{a}{o}((x-a)^n)$, c'est à dire : $f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + \underset{a}{o}((x-a)^n)$.

Preuve : Par récurrence sur n : pour $n = 0$ $f(x) = f(a) + (f(x) - f(a))$, et $f(x) - f(a) = \underset{a}{o}(1)$, car f est continue en a. Supposons le théorème démontré au rang n, et supposons f de classe $\mathcal{C}^{n+1}$ sur I, alors f' est de classe $\mathcal{C}^n$, l'hypothèse de récurrence s'applique à f' : $f'(t) = \mathrm{T}_{n,f',a}(t) + \underset{a}{o}((t-a)^n)$, en intégrant entre a et x, on obtient $f(x) - f(a) = \sum_{k=0}^{n} \frac{f'^{(k)}(a)}{(k+1)!}(x-a)^{k+1} + \underset{a}{o}\left((x-a)^{n+1}\right)$, en appliquant le théorème précédent pour l'intégration du o, un changement d'indice donne alors $f(x) = f(a) + \sum_{k=0}^{n} \frac{f'^{(k)}(a)}{(k+1)!}(x-a)^{k+1} + \underset{a}{o}\left((x-a)^{n+1}\right) = \sum_{k=0}^{n+1} \frac{f^{(k)}(a)}{k!}(x-a)^k + \underset{a}{o}\left((x-a)^{n+1}\right)$.
□

Remarque 19.6 – *Sous les mêmes hypothèses, on peut écrire qu'il existe une fonction* ε *telle que :*

$$f(x) = \mathrm{T}_{n,f,a}(x) + (x-a)^n\varepsilon(x) \text{ avec } \lim_{x\to a}\varepsilon(x) = 0.$$

Théorème 19.12

Soit $f : \mathrm{I} \to \mathbb{C}$ *et soit* $a \in \mathrm{I}$, *si* f *est de classe* $\mathcal{C}^n$ *sur l'intervalle* I, *alors* f *admet un* $\mathrm{dl}_n(a)$ *et sa partie régulière est* $\mathrm{T}_{n,f,a}(x)$, *c'est à dire son polynôme de Taylor en* a *à l'ordre* n. *Si* f *est de classe* $\mathcal{C}^\infty$ *sur* I, *alors* f *admet un* dl *en tout point de* I *et à n'importe quel ordre.*

3) Développements usuels en 0

À retenir

- $\exp(x) = \sum_{k=0}^{n} \frac{x^k}{k!} + \underset{0}{o}(x^n)$.
- $\ln(1+x) = \sum_{k=1}^{n} (-1)^{k-1}\frac{x^k}{k} + \underset{0}{o}(x^n)$.
- $\sin(x) = \sum_{k=0}^{n} (-1)^k \frac{x^{2k+1}}{(2k+1)!} + \underset{0}{o}\left(x^{2n+2}\right)$ (il s'agit du $\mathrm{dl}_{2n+2}(0)$).
- $\cos(x) = \sum_{k=0}^{n} (-1)^k \frac{x^{2k}}{(2k)!} + \underset{0}{o}\left(x^{2n+1}\right)$ (il s'agit d'un $\mathrm{dl}_{2n+1}(0)$).
- $\mathrm{sh}(x) = \sum_{k=0}^{n} \frac{x^{2k+1}}{(2k+1)!} + \underset{0}{o}\left(x^{2n+2}\right)$ (il s'agit du $\mathrm{dl}_{2n+2}(0)$).
- $\mathrm{ch}(x) = \sum_{k=0}^{n} \frac{x^{2k}}{(2k)!} + \underset{0}{o}\left(x^{2n+1}\right)$ (il s'agit d'un $\mathrm{dl}_{2n+1}(0)$).
- $(1+x)^\alpha = \sum_{k=0}^{n} \binom{\alpha}{k} x^k + \underset{0}{o}(x^n)$, où $\binom{\alpha}{k} = \dfrac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!}$, formule des coefficients du binôme généralisée. En particulier : $\frac{1}{1+x} = \sum_{k=0}^{n} (-1)^k x^k + \underset{0}{o}(x^n)$.

III PROPRIÉTÉS

1) Généralités

Théorème 19.13 (troncature)

Si f *admet un* $\mathrm{dl}_n(a)$ *alors pour tout entier* $p \in [\![0\,;n]\!]$, f *admet un* $\mathrm{dl}_p(a)$ *dont la partie régulière s'obtient en tronquant la partie régulière du* $\mathrm{dl}_n(a)$ *au degré* p.

Preuve : Celle-ci est simple et laissée en exercice. □

★**Exercice 19.2** *Comment s'écrit le* $\mathrm{dl}_n(0)$ *d'un polynôme ?*

Théorème 19.14 (unicité du dl)

Si f *admet un* $\mathrm{dl}_n(a)$ *alors celui-ci est unique.*

Preuve : Par récurrence sur n. Au rang 0 c'est évident avec un passage à la limite. Supposons l'unicité montrée au rang $n-1$. Si f a deux $\mathrm{dl}_n(a)$, alors par troncature elle a deux $\mathrm{dl}_{n-1}(a)$, ils sont donc égaux, ce qui donne après simplification $\alpha_n(x-a)^n + \underset{a}{o}((x-a)^n) = \beta_n(x-a)^n + \underset{a}{o}((x-a)^n)$, en simplifiant par $(x-a)^n$ (pour $x \neq a$), et par passage à la limite en a, on obtient $\alpha_n = \beta_n$ et on a bien l'unicité au rang n. □

Changement de variable

On peut toujours se ramener en $a = 0$:

- On pose $u = x - a$, on a alors $f(x) = f(u+a) = g(u)$, d'où :

$$\begin{aligned} f \text{ admet un } \mathrm{dl}_n(a) &\iff \exists\, \mathrm{P} \in \mathbb{C}_n[\mathrm{X}], f(x) = \mathrm{P}(x-a) + \underset{a}{o}((x-a)^n) \\ &\iff \exists\, \mathrm{P} \in \mathbb{C}_n[\mathrm{X}], g(u) = \mathrm{P}(u) + \underset{0}{o}(u^n) \\ &\iff g \text{ admet un } \mathrm{dl}_n(0). \end{aligned}$$

– Si f est définie au voisinage de $\pm\infty$: on pose $u = \frac{1}{x}$, on a alors $f(x) = f(\frac{1}{u}) = g(u)$. Si g admet un $\text{dl}_n(0)$; alors il existe un polynôme $P \in \mathbb{C}_n[X]$ tel que $g(u) = P(u) + \underset{0}{o}(u^n)$, ce qui donne $f(x) = P(\frac{1}{x}) + \underset{\pm\infty}{o}\left(\frac{1}{x^n}\right)$, on dit alors que f admet un développement asymptotique en $\frac{1}{x}$ d'ordre n en $\pm\infty$, on remarquera que la partie régulière n'est pas un polynôme en x mais en $\frac{1}{x}$.

Théorème 19.15

- *f admet un $\text{dl}_0(a)$ si et seulement si f admet une limite finie en a.*
- *f admet un $\text{dl}_1(a)$ si et seulement si f admet un prolongement continu dérivable en a.*
- *Si f admet un $\text{dl}_n(0)$, alors la partie régulière a la même parité que f.*

Preuve : Celle-ci est simple et laissée en exercice. □

Attention !

Une fonction peut avoir un $\text{dl}_2(a)$ sans être deux fois dérivable en a.

★**Exercice 19.3** *Montrer que la fonction définie par $f(x) = x^3\sin(\frac{1}{x})$ avec $f(0) = 0$, admet un $\text{dl}_2(0)$ (dont la partie régulière est nulle), mais f' n'est pas dérivable en 0, donc f n'est pas deux fois dérivable en 0 (et donc f' n'a pas de DL en 0).*

2) Règles de calculs

Théorème 19.16

Si f, g admettent un $\text{dl}_n(0)$, $f(x) = P(x) + \underset{0}{o}(x^n)$ et $g(x) = Q(x) + \underset{0}{o}(x^n)$, avec $P, Q \in \mathbb{C}_n[X]$.

- *DL d'une combinaison linéaire : $\forall\, \lambda \in \mathbb{C}, \lambda f + g$ admet un $\text{dl}_n(0)$ dont la partie régulière en $\lambda P(x) + Q(x)$.*
- *DL d'un produit : $f(x) \times g(x)$ admet un $\text{dl}_n(0)$ dont la partie régulière est $[P(x)Q(x)]_n$ (troncature du polynôme $P(x) \times Q(x)$ au degré n).*
- *DL d'une composée : si $\lim\limits_{x\to 0} g(x) = 0$, alors $f(g(x))$ admet un $\text{dl}_n(0)$ dont la partie régulière est $[P(Q(x))]_n$.*
- *Intégration des DL : si f' admet un $\text{dl}_n(0)$ dont la partie régulière est $P(x)$, alors f admet un $\text{dl}_{n+1}(0)$ dont la partie régulière est : $f(0) + \int_0^x P(t)\,dt$.*
- *Inversion d'un DL : si $\lim\limits_{x\to 0} f(x) = a \neq 0$, alors $\frac{1}{f(x)}$ admet un $\text{dl}_n(0)$ qui s'obtient en composant le $\text{dl}_n(0)$ de $\frac{1}{1+u}$ avec celui de $\frac{f(x)-a}{a}$, et en multipliant par $\frac{1}{a}$.*

Preuve : Donnons un exemple pour la composition avec $n = 2$: $f(x) = a + bx + cx^2 + x^2u(x)$ et $g(x) = \alpha x + \beta x^2 + x^2 v(x)$, avec u et v de limite nulle en 0, on en déduit en composant : $f(g(x)) = a + b(\alpha x + \beta x^2 + x^2 v(x)) + c(\alpha x + \beta x^2 + v(x))^2 + (\alpha x + \beta x^2 + x^2 v(x))^2 u(g(x))$, ce qui donne après avoir développer et regrouper les puissances de x strictement supérieures à 2 : $f(g(x)) = a + b\alpha x + [b\beta + c\alpha^2]x^2 + \underset{0}{o}\left(x^2\right)$, on peut vérifier que la partie régulière est la troncature au degré 2 de $P(Q(x))$.

Pour l'intégration cela découle de la propriété déjà établie : $\int_0^x \underset{0}{o}(t^n)\,dt = \underset{0}{o}\left(x^{n+1}\right)$.

Pour $\frac{1}{f(x)}$: comme $a \neq 0$, on a $\frac{1}{f(x)} = \frac{1}{a}\frac{1}{1+\frac{f(x)-a}{a}}$, et $\frac{f(x)-a}{a} \xrightarrow[x\to 0]{} 0$, donc la règle de composition s'applique. □

Attention !

Il n'y a pas de propriété de dérivation de DL. Par exemple, on vérifiera que la fonction définie par $f(x) = x^2\sin(\frac{1}{x})$ avec $f(0) = 0$, admet un $\text{dl}_1(0)$ (dont la partie régulière est nulle), mais f' n'a pas de limite finie en 0, donc f' n'a pas de $\text{dl}_0(0)$.

3) Développements usuels (compléments)

Pour $x \neq 1$, on a $\sum\limits_{k=0}^{n} x^k = \frac{1-x^{n+1}}{1-x}$, on en déduit :

$$\frac{1}{1-x} = \sum_{k=0}^{n} x^k + \frac{x^{n+1}}{1-x} = \sum_{k=0}^{n} x^k + \underset{0}{o}(x^n).$$

En substituant $-x$ à x, on obtient :

$$\frac{1}{1+x} = \sum_{k=0}^{n} (-1)^k x^k + \underset{0}{o}(x^n).$$

En intégrant ce dernier développement entre 0 et x, on obtient :

$$\ln(1+x) = \sum_{k=0}^{n} (-1)^k \frac{x^{k+1}}{k+1} + \underset{0}{o}\left(x^{n+1}\right).$$

En substituant x^2 à x dans l'avant dernier, on obtient :

$$\frac{1}{1+x^2} = \sum_{k=0}^{n} (-1)^k x^{2k} + \underset{0}{o}\left(x^{2n+1}\right) \text{ c'est un } \mathrm{dl}_{2n+1}(0).$$

En intégrant ce dernier développement entre 0 et x, on obtient :

$$\arctan(x) = \sum_{k=0}^{n} (-1)^k \frac{x^{2k+1}}{2k+1} + \underset{0}{o}\left(x^{2n+2}\right).$$

On a :

$$\frac{1}{\sqrt{1+x}} = \sum_{k=0}^{n} \binom{-1/2}{k} x^k + \underset{0}{o}(x^n).$$

Or $\binom{-1/2}{k} = (-1)^k \frac{1\times 3\times \ldots \times (2k-1)}{k!2^k} = (-1)^k \frac{\binom{2k}{k}}{4^k}$, on a finalement :

$$\frac{1}{\sqrt{1+x}} = \sum_{k=0}^{n} (-1)^k \frac{\binom{2k}{k}}{4^k} x^k + \underset{0}{o}(x^n).$$

On en déduit que :

$$\frac{1}{\sqrt{1-x^2}} = \sum_{k=0}^{n} \frac{\binom{2k}{k}}{4^k} x^{2k} + \underset{0}{o}\left(x^{2n+1}\right).$$

En intégrant entre 0 et x, obtient :

$$\arcsin(x) = \sum_{k=0}^{n} \frac{\binom{2k}{k}}{4^k(2k+1)} x^{2k+1} + \underset{0}{o}\left(x^{2n+2}\right).$$

et :

$$\arccos(x) = \frac{\pi}{2} - \sum_{k=0}^{n} \frac{\binom{2k}{k}}{4^k(2k+1)} x^{2k+1} + \underset{0}{o}\left(x^{2n+2}\right).$$

☞**Exemples** :

- Calculer un $\mathrm{dl}_3(0)$ de $\exp(\sin(x))$.
  On a $\sin(x) = x - \frac{x^3}{6} + \underset{0}{o}\left(x^3\right)$ et $\exp(u) = 1 + u + \frac{u^2}{2} + \frac{u^3}{6} + \underset{0}{o}\left(u^3\right)$. Comme $\lim\limits_{x\to 0} \sin(x) = 0$, on peut appliquer le théorème de composition, et composer les parties régulières jusqu'à l'ordre 3, ce qui donne $\exp(\sin(x)) = 1 + x + \frac{x^2}{2} + \underset{0}{o}\left(x^3\right)$.
- Calculer un $\mathrm{dl}_4(0)$ de $(1+\sin(x))^x$.
  L'expression est égale à $\exp[x\ln(1+\sin(x))]$. Un $\mathrm{dl}_3(0)$ de $\sin(x)$ est $\sin(x) = x - \frac{x^3}{6} + \underset{0}{o}\left(x^3\right)$, et $\ln(1+u) = u - \frac{u}{2} + \frac{u^3}{3} + \underset{0}{o}\left(u^3\right)$, on obtient par composition, $\ln(1+\sin(x)) = x - \frac{x^2}{2} + \frac{x^3}{6} + \underset{0}{o}\left(x^3\right)$ et donc $x\ln(1+\sin(x)) = x^2 - \frac{x^3}{2} + \frac{x^4}{6} + \underset{0}{o}\left(x^4\right)$. On a également $\exp(v) = 1 + v + \frac{v^2}{2} + \underset{0}{o}\left(v^2\right)$, d'où par composition : $\exp[x\ln(1+\sin(x))] = 1 + x^2 - \frac{x^3}{2} + 2\frac{x^4}{3} + \underset{0}{o}\left(x^4\right)$.

- Calculer un $dl_5(0)$ de $\tan(x)$:
  On a $\tan(x) = \sin(x)\times\frac{1}{\cos(x)}$. $\sin(x) = x-\frac{x^3}{6}+\frac{x^5}{120}+\underset{0}{o}\left(x^5\right)$. D'autre part $\frac{1}{\cos(x)} = \frac{1}{1+u}$ avec $u = \cos(x)-1$, comme $u \to 0$, on pourra donc composer, $\cos(x)-1 = -\frac{x^2}{2}+\frac{x^4}{24}+\underset{0}{o}\left(x^5\right)$ et $\frac{1}{1+u} = 1-u+u^2+\underset{0}{o}\left(u^2\right)$, ce qui donne : $\frac{1}{\cos(x)} = 1+\frac{x^2}{2}+5\frac{x^4}{24}+\underset{0}{o}\left(x^5\right)$. On effectue ensuite le produit avec le dl de $\sin(x)$, ce qui donne $\tan(x) = x+\frac{x^3}{3}+\frac{2x^5}{15}+\underset{0}{o}\left(x^5\right)$.
  Autre méthode : on a $\tan(x) = 0+\underset{0}{o}(1)$, d'où $1+\tan(x)^2 = 1+\underset{0}{o}(1)$, en intégrant, on obtient $\tan(x) = x+\underset{0}{o}(x) = x\times(1+\underset{0}{o}(1))$. Puis on recommence : $1+\tan(x)^2 = x^2\times(1+\underset{0}{o}(1))^2 = x^2\times(1+\underset{0}{o}(1)) = 1+x^2+\underset{0}{o}\left(x^2\right)$ et donc $\tan(x) = x+\frac{x^3}{3}+\underset{0}{o}\left(x^3\right) = x\times(1+\frac{x^2}{3}+\underset{0}{o}\left(x^2\right))$, mais alors $1+\tan(x)^2 = x^2\times(1+\frac{x^2}{3}+\underset{0}{o}\left(x^2\right))^2 = x^2\times(1+\frac{2x^2}{3}+\underset{0}{o}\left(x^2\right)) = 1+x^2+\frac{2x^4}{3}+\underset{0}{o}\left(x^4\right)$, et donc $\tan(x) = x+\frac{x^3}{3}+\frac{2x^5}{15}+\underset{0}{o}\left(x^5\right)$... etc

IV APPLICATIONS

1) Recherche d'une limite

☞**Exemple** : Soit $f(x) = \frac{x^2-1-2x\ln(x)}{x(x-1)\ln(x)}$, calculer $\lim\limits_{x\to 1} f(x)$.

Il s'agit bien d'une forme indéterminée, on ramène le problème en 0 en posant $u = x-1$, ce qui donne $f(x) = \frac{u^2+2u-2(1-u)\ln(1+u)}{u(1+u)\ln(1+u)} \underset{0}{\sim} \frac{u^2+2u-2(1+u)\ln(1+u)}{u^2}$. On cherche alors un $dl_2(0)$ du numérateur, ce qui donne $\underset{0}{o}\left(u^2\right)$, on a donc $f(x) = f(1+u) \underset{0}{\sim} \underset{0}{o}(1)$ et donc la limite cherchée est nulle.

★**Exercice 19.4** *Calculer* $\lim\limits_{x\to+\infty} x^2\ln(\frac{x}{1+x})+x-1$.

2) Étude locale d'une fonction au voisinage d'un point

Si f a un $dl_2(a)$, $f(x) = a_0+a_1(x-a)+a_2(x-a)^2+\underset{a}{o}\left((x-a)^2\right)$, alors on voit que $\lim\limits_{x\to a} f(x) = a_0$, on peut donc prolonger f par continuité en a, en posant $f(a) = a_0$ (si ce n'est pas déjà fait !). Le taux d'accroissement en a s'écrit : $\frac{f(x)-a_0}{x-a} = a_1+\underset{a}{o}(1)$, donc ce prolongement est dérivable en a et $f'(a) = a_1$. L'équation de la tangente à la courbe au point d'abscisse a est $y = a_1(x-a)+a_0$, et l'étude de la position courbe-tangente se fait en étudiant le signe de $f(x)-[a_0+a_1(x-a)] = (x-a)^2[a_2+\underset{a}{o}(1)]$, d'où la discussion :

- si $a_2 > 0$: alors au voisinage de a on a $[a_2+\underset{a}{o}(1)] > 0$ et donc $f(x) > a_0+a_1(x-a)$, *i.e.* la courbe est au-dessus de sa tangente **au voisinage** de a.
- si $a_2 < 0$: c'est la situation inverse.
- si $a_2 = 0$: on ne peut rien dire, il faut aller plus loin dans le développement limité. Dans la pratique on s'arrête au premier terme non nul de degré supérieur ou égal à 2.

★**Exercice 19.5** *Soit* $f(x) = \frac{x\ln(x)}{x^2-1}$, *effecter une étude locale en* $a = 1$.

3) Étude locale au voisinage de l'infini

Si f est définie au voisinage de ∞ et admet une limite infinie, alors on peut étudier la branche infinie de f de la manière suivante : on pose $g(x) = \frac{f(x)}{x}$, on se ramène en 0 en posant $u = 1/x$, ce qui donne $g(x) = g(1/u) = uf(1/u)$, et on cherche un $dl_2(0)$ de cette expression : $uf(1/u) = a_0+a_1u+a_2u^2+\underset{0}{o}\left(u^2\right)$, ce qui donne en revenant à x, $f(x) = a_0x+a_1+a_2\frac{1}{x}+\underset{\infty}{o}\left(\frac{1}{x}\right)$, d'où $\lim\limits_{x\to\infty} f(x)-[a_0x+a_1] = 0$, donc la droite d'équation $y = a_0x+a_1$ est asymptote à C_f au voisinage de ∞. Pour la position courbe-asymptote, on étudie la différence : $f(x)-[a_0x+a_1] = \frac{1}{x}[a_2+\underset{\infty}{o}(1)]$, l'étude du signe se fait comme dans le paragraphe précédent si $a_2 \neq 0$. Lorsque $a_2 = 0$ il faut aller plus loin dans le développement pour avoir le signe.

Remarque 19.7 – *Dans la pratique, on n'est pas obligé de passer par la fonction $\frac{f(x)}{x}$, en travaillant directement sur $f(x)$.*

★**Exercice 19.6** $f(x) = \sqrt[3]{\frac{x^4}{x-3}}$ *au voisinage de* $+\infty$.

4) Recherche d'un équivalent

Théorème 19.17
Si f admet un $\mathrm{dl}_n(a)$, alors $f(x)$ est équivalente en a au terme **non nul de plus bas degré** *de la partie régulière, s'il existe.*

Preuve : Soit $a_p(x-a)^p$ le premier terme non nul, on a alors $f(x) = a_p(x-a)^p + \underset{a}{o}((x-a)^p) = (x-a)^p[1 + \underset{a}{o}(1)]$, ce qui prouve l'équivalence annoncée. □

Remarque 19.8 –
- *En se ramenant en 0, on peut également trouver un équivalent d'une fonction en $\pm\infty$.*
- *Avec ce théorème, on retrouve tous les équivalents dits « classiques ».*

★**Exercice 19.7** *Équivalent en 0 de la fonction* $f(x) = \frac{\arcsin(x)}{\sqrt{1-x^2}} - \frac{3x}{3-2x^2}$.

V SOLUTION DES EXERCICES

Solution 19.1

1/ *On a $f(x) = \exp(x\ln(1+\frac{1}{x}))$, or $\ln(1+\frac{1}{x}) \underset{+\infty}{\sim} \frac{1}{x}$ car $\frac{1}{x} \underset{+\infty}{\to} 0$, donc $x\ln(1+\frac{1}{x}) \underset{+\infty}{\sim} 1$, la limite cherchée est donc égale à 1.*

2/ *On a $\sin(x)^x = \exp(x\ln(\sin(x)))$, or $\ln(\sin(x)) = \ln(\frac{\sin(x)}{x}) + \ln(x) = \ln(x)\left[1 + \frac{\ln(\frac{\sin(x)}{x})}{\ln(x)}\right] \underset{0}{\sim} \ln(x)$ et donc $x\ln(\sin(x)) \underset{0}{\sim} x\ln(x) \underset{0}{\to} 0$, d'où : $\exp(x\ln(\sin(x))) - 1 \underset{0}{\sim} x\ln(\sin(x)) \underset{0}{\sim} x\ln(x)$, par conséquent $f(x) \underset{0}{\sim} \frac{x}{\sqrt{x}} = \sqrt{x}$, et donc la limite cherchée est égale à 0.*

Solution 19.2 *Il suffit de le tronquer au degré n, la somme des termes suivants est un $\underset{0}{o}(x^n)$.*

Solution 19.3 *On a $f(x) = \underset{0}{o}\left(x^2\right)$ qui un $\mathrm{dl}_2(0)$ de f, on en déduit que f est dérivable en 0 avec $f'(0) = 0$. Pour $x \neq 0$, on a $f'(x) = 2x^2\sin(\frac{1}{x}) - x\cos(\frac{1}{x})$, d'où $\frac{f'(x)-f'(0)}{x-0} = 2x\sin(\frac{1}{x}) - \cos(\frac{1}{x})$, or cette fonction n'a pas de limite en 0 car si on prend $u_n = \frac{1}{n\pi}$, alors $\cos(\frac{1}{u_n}) = (-1)^n$ n'a pas de limite, alors que $2u_n\sin(\frac{1}{u_n}) \to 0$. Donc f n'est pas deux fois dérivable en 0 (f a un $\mathrm{dl}_2(0)$ mais f' n'a pas de $\mathrm{dl}_1(0)$).*

Solution 19.4 *Il s'agit bien d'une forme indéterminée, on se ramène en 0 en posant $u = 1/x$, on a alors $f(x) = f(\frac{1}{u}) = \frac{-1}{u^2}\ln(1+u) + \frac{1}{u} - 1$, ce qui donne $f(\frac{1}{u}) = \frac{-1}{2} + \underset{0}{o}(1)$, et donc la limite cherchée est $\frac{-1}{2}$.*

Solution 19.5 *On pose $u = x-1$ d'où $f(x) = f(1+u) = \frac{\ln(1+u)}{u}(1+u)\frac{1}{2[1+u/2]}$, le calcul donne $f(x) = f(1+u) = \frac{1}{2} - \frac{u^2}{12} + \underset{0}{o}\left(u^2\right)$. On en déduit que f se prolonge par continuité en 1 en posant $f(1) = \frac{1}{2}$, ce prolongement est dérivable en 1 et $f'(1) = 0$, de plus, au voisinage de 1, la courbe est en-dessous de la tangente. On a donc un maximum local en 1.*

Solution 19.6 *On voit que f est définie au voisinage de $+\infty$ et que $f(x) \underset{+\infty}{\sim} x$, il y a donc une branche infinie de direction asymptotique $y = x$. Posons $u = \frac{1}{x}$, on a alors $f(x) = f(\frac{1}{u}) = \frac{1}{u}(1-3u)^{-1/3} = \frac{1}{u}(1 + u + 2u^2 + \underset{0}{o}\left(u^2\right))$, d'où $f(x) = x + 1 + \frac{2}{x} + \underset{+\infty}{o}\left(\frac{1}{x}\right)$. Donc la droite d'équation $y = x+1$ est asymptote à la courbe de f en $+\infty$, et au voisinage de $+\infty$ la courbe de f est au-dessus.*

Solution 19.7 *On a $\frac{1}{\sqrt{1-x^2}} = 1 + \frac{x^2}{2} + \frac{3x^4}{8} + \underset{0}{o}\left(x^5\right)$, en intégrant on obtient $\arcsin(x) = x + \frac{x^3}{6} + \frac{3x^5}{40} + \underset{0}{o}\left(x^5\right)$, en effectuant le produit, il vient que : $\frac{\arcsin(x)}{\sqrt{1-x^2}} = x + \frac{2x^3}{3} + \frac{8x^5}{15} + \underset{0}{o}\left(x^5\right)$. D'un autre côté, on a $\frac{3x}{3-2x^2} = x\frac{1}{1-2x^2/3} = x[1 + \frac{2x^2}{3} + \frac{4x^4}{9} + \underset{0}{o}\left(x^4\right)] = x + \frac{2x^3}{3} + \frac{4x^5}{9} + \underset{0}{o}\left(x^5\right)$. Finalement, on a $f(x) = \frac{4x^5}{45} + \underset{0}{o}\left(x^5\right)$, et donc $f(x) \underset{0}{\sim} \frac{4x^5}{45}$.*

Chapitre 20

Espaces vectoriels

Sommaire

Dans ce chapitre, $\mathbb{K}$ désigne un sous-corps de $\mathbb{C}$.

I GÉNÉRALITÉS

1) Définition

Définition 20.1

Soit E *un ensemble non vide, on dit que* E *est un* $\mathbb{K}$ *- espace vectoriel (ou* $\mathbb{K}$*-e.v.) lorsque* E *possède une addition et un produit par les scalaires (loi de composition externe, notée « . »,* *c'est une application :* $\begin{array}{ccc} \mathbb{K}\times E & \to & E \\ (\lambda, x) & \mapsto & \lambda.x \end{array}$*), avec les propriétés suivantes :*

- $(E, +)$ *est un groupe abélien (l'élément neutre est noté* 0_E *ou* $\overrightarrow{0_E}$ *et appelé* **vecteur nul** *de* E*).*
- *La loi . (ou produit par les scalaires) doit vérifier :* $\forall\ \lambda, \mu \in \mathbb{K}, \forall\ x, y \in E$ *:*
  - $1.x = x$
  - $\lambda.(x + y) = \lambda.x + \lambda.y$
  - $(\lambda + \mu).x = \lambda.x + \mu.x$
  - $\lambda.(\mu.x) = (\lambda\mu).x$

Si ces propriétés sont vérifiées, on dit que $(E, +, .)$ *est un* $\mathbb{K}$ *- e.v., les éléments de* $\mathbb{K}$ *sont appelés*

les scalaires** et les éléments de* E *sont appelés* ***vecteurs *(parfois notés avec une flèche).*

2) Exemples de référence

Exemples :
- Un corps $\mathbb{K}$ est un $\mathbb{K}$-e.v..
- $\mathbb{R}$ est un $\mathbb{Q}$-e.v., $\mathbb{C}$ est un $\mathbb{Q}$-e.v., $\mathbb{C}$ est un $\mathbb{R}$-e.v. Plus généralement si $\mathbb{K}$ est corps inclus dans un autre corps $\mathbb{L}$, alors $\mathbb{L}$ est un $\mathbb{K}$-e.v..
- L'ensemble $\mathbb{K}^n$ muni des opérations suivantes :
$$(x_1,\dots,x_n)+(y_1,\dots,y_n)=(x_1+y_1,\dots,x_n+y_n) \text{ et } \lambda.(x_1,\dots,x_n)=(\lambda x_1,\dots,\lambda x_n),$$
est un $\mathbb{K}$-e.v., le vecteur nul est le n-uplet : $(0,\dots,0)$.
- Si I est un ensemble non vide, alors l'ensemble des applications de I vers $\mathbb{K}$: $\mathcal{F}(\mathrm{I},\mathbb{K})=\mathbb{K}^{\mathrm{I}}$, pour les opérations usuelles (addition de deux fonctions et produit par un scalaire) est un $\mathbb{K}$-e.v., le vecteur nul étant l'application nulle. En particulier $(\mathcal{C}^n(\mathrm{I},\mathbb{K}),+,.)$ sont des $\mathbb{K}$-e.v., ainsi que l'espace des suites à valeurs dans $\mathbb{K}$.
Plus généralement, si E est un $\mathbb{K}$-e.v., l'ensemble des applications de I vers E : $\mathcal{F}(\mathrm{I},\mathrm{E})=\mathrm{E}^{\mathrm{I}}$, pour les opérations usuelles sur les fonctions, est un $\mathbb{K}$-e.v..
- $(\mathbb{K}[\mathrm{X}],+,.)$ est un $\mathbb{K}$-espace vectoriel.
- $(\mathbb{K}(\mathrm{X}),+,.)$ est un $\mathbb{K}$-espace vectoriel.
- $\mathcal{M}_{n,p}(\mathbb{K}))$ est un $\mathbb{K}$-espace vectoriel.
- Espace produit : soient E et F deux $\mathbb{K}$-e.v., on définit sur $\mathrm{E}\times\mathrm{F}$ l'addition : $(x,y)+(x',y')=(x+x',y+y')$, et un produit par les scalaires : $\lambda.(x,y)=(\lambda.x,\lambda.y)$. On peut vérifier alors que $(\mathrm{E}\times\mathrm{F},+,.)$ est un $\mathbb{K}$-e.v., le vecteur nul étant $(0_{\mathrm{E}},0_{\mathrm{F}})$. Cela se généralise au produit cartésien d'un nombre fini de $\mathbb{K}$-e.v.

3) Règles de calculs

Soit E un $\mathbb{K}$-e.v.

- $\forall\,\vec{x}\in\mathrm{E},\ 0.\vec{x}=\vec{0}$, et $\forall\,\lambda\in\mathbb{K},\ \lambda.\vec{0}=\vec{0}$.
- $\forall\,\vec{x}\in\mathrm{E},\ \forall\,\lambda\in\mathbb{K},\ -(\lambda.\vec{x})=(-\lambda).\vec{x}=\lambda.(-\vec{x})$.
- $\forall\,\vec{x}\in\mathrm{E},\ \forall\,\lambda\in\mathbb{K},\ \lambda.\vec{x}=\vec{0}\implies\lambda=0 \text{ ou } \vec{x}=\vec{0}$.

II SOUS-ESPACES VECTORIELS D'UN ESPACE VECTORIEL

1) Définition

Définition 20.2
Soit E *un* $\mathbb{K}$*-e.v. et soit* H *un ensemble, on dit que* H *est un sous-espace vectoriel de* E *(ou s.e.v de* E*) lorsque :*
- $\mathrm{H}\subset\mathrm{E},\ \mathrm{H}\neq\varnothing$.
- $\forall\,x,y\in\mathrm{H},\ x+y\in\mathrm{H}$ *(*H *est stable pour l'addition).*
- $\forall\,x\in\mathrm{H},\ \forall\,\lambda\in\mathbb{K},\ \lambda.x\in\mathrm{H}$ *(*H *est stable pour la loi .).*

Si c'est le cas, alors il est facile de vérifier que $(\mathrm{H},+,.)$ *est lui-même un* $\mathbb{K}$*-e.v.*

Exemples :
- L'ensemble des fonctions paires (respectivement impaires, bornées, T-périodiques, lipschitziennes) définies sur $\mathbb{R}$ est un s.e.v. de $\mathcal{F}(\mathbb{R},\mathbb{R})$.
- L'ensemble $(\mathcal{C}^n(\mathrm{I},\mathbb{C}),+,.)$ est un sous-espace vectoriel de $(\mathcal{F}(\mathrm{I},\mathbb{C}),+,.)$.
- L'ensemble $(\mathbb{K}_n[\mathrm{X}],+,.)$ est un sous-espace vectoriel de $(\mathbb{K}[\mathrm{X}],+,.)$.
- L'ensemble des suites complexes de limite nulle et un s.e.v de l'espace des suites complexes convergentes, qui est lui-même un s.e.v de l'espace de suites complexes bornées, qui est lui-même un s.e.v de l'espace des suites complexes.
- Soient $a,b,c\in\mathbb{K}$, $\mathrm{F}=\{(x,y,z)\in\mathbb{K}^3\ /\ ax+by+cz=0\}$ est un s.e.v de $\mathbb{K}^3$.

– $S_n(\mathbb{K})$ (matrices symétriques), $A_n(\mathbb{K})$ (matrices antisymétriques), $T_n^s(\mathbb{K})$ (matrices triangulaire supérieures), $T_n^i(\mathbb{K})$ (matrices triangulaires inférieures), $D_n(\mathbb{K})$ (matrices diagonales), sont des s.e.v. de $\mathcal{M}_n(\mathbb{K})$.

Théorème 20.1 (intersection de sous-espaces vectoriels)
Soit $(H_i)_{i\in I}$ une famille de s.e.v de E (I est un ensemble d'indices), alors $\bigcap\limits_{i\in I} H_i$ est un s.e.v de E.

Preuve : Celle-ci est simple et laissée en exercice. □

2) Sous-espace engendré

Définition 20.3 (combinaisons linéaires d'un nombre fini de vecteurs)
Soit E un $\mathbb{K}$-e.v et soit $x_1, \ldots, x_n$ des vecteurs de E. On appelle combinaison linéaire de la famille $(x_i)_{1\leqslant i\leqslant n}$, tout vecteur x de E pour lequel il existe des scalaires $\lambda_1, \ldots, \lambda_n$ tels que :

$$x = \sum_{i=1}^{n} \lambda_i x_i.$$

L'ensemble des combinaisons linéaires de la famille $(x_i)_{1\leqslant i\leqslant n}$ est noté $\mathrm{Vect}\,[x_1, \ldots, x_n]$.
Deux vecteurs x et y de E sont dits colinéaires lorsque l'un des deux est combinaison linéaire de l'autre, i.e. $\exists\, \lambda \in \mathbb{K}, x = \lambda y$ ou $y = \lambda x$.

Théorème 20.2 (sous-espace engendré)
Soit $(x_i)_{1\leqslant i\leqslant n}$ une famille de vecteurs de E, $\mathrm{Vect}\,[x_1, \ldots, x_n]$ est un s.e.v de E. C'est même le plus petit (pour l'inclusion) s.e.v de E qui contient tous les vecteurs de cette famille. On l'appelle s.e.v engendré par $(x_1, \ldots, x_n)$.

Preuve : Celle-ci est simple et laissée en exercice. □

☞**Exemples** :
– $\mathrm{Vect}\,[0_E] = \{0_E\}$.
– Si $x \in E \setminus \{0_E\}$, alors $\mathrm{Vect}\,[x] = \{\lambda x \;/\; \lambda \in \mathbb{K}\}$, c'est un s.e.v de E appelé droite vectorielle engendrée par x. On dit que x est un vecteur directeur de cette droite. Les autres vecteurs directeurs sont les vecteurs de la forme λx avec $\lambda \neq 0$.
– Soient $x, y \in E$ deux vecteurs non nuls, si les deux vecteurs sont colinéaires, alors $\mathrm{Vect}\,[x, y] = \mathrm{Vect}\,[x] = \mathrm{Vect}\,[y]$ (droite vectorielle). Si ces deux vecteurs sont non colinéaires, alors :

$$\mathrm{Vect}\,[x, y] = \{\alpha x + \beta y \;/\; \alpha, \beta \in \mathbb{K}\}$$

c'est un s.e.v de E, on l'appelle plan vectoriel engendré par x et y, il contient (strictement) les deux droites engendrées par x et y.
– Dans $\mathbb{K}^3$ déterminer une équation cartésienne du plan vectoriel engendré par les vecteurs $x = (1, 1, 1)$ et $y = (0, -1, 1)$.

☞**Exemples** :
– Soit $E = \mathbb{K}^n$ pour $i \in [\![1\,;n]\!]$ on pose $e_i = (\delta_{i,1}, \ldots, \delta_{i,n})$, on a alors $E = \mathrm{Vect}\,[e_1, \ldots, e_n]$.
– Soit $H = \{u \in \mathbb{K}^3 \;/\; \exists\, \alpha, \beta, \gamma \in \mathbb{K}, u = (\alpha - \beta, 2\alpha - 2\beta + \gamma, -\alpha + \beta + 2\gamma)\}$. Posons $e_1 = (1, 2, -1)$, $e_2 = (-1, -2, 1)$ et $e_3 = (0, 1, 2)$, on a alors $H = \mathrm{Vect}\,[e_1, e_2, e_3]$, ce qui prouve que H est un s.e.v de $\mathbb{K}^3$. On remarque que $e_2 = -e_1$, donc finalement $H = \mathrm{Vect}\,[e_1, e_3]$, et comme e_1 et e_3 ne sont pas colinéaires, H est un plan vectoriel.
– Soit $E = \mathcal{F}(\mathbb{R}, \mathbb{R})$, les deux fonctions $\mathrm{id}_{\mathbb{R}}$ et 1 sont non colinéaires, donc elles engendrent un plan vectoriel dans E : $P = \mathrm{Vect}\,[\mathrm{id}_{\mathbb{R}}, 1]$. $f \in P$ équivaut à $\exists\, a, b \in \mathbb{R}, f = a.\mathrm{id}_{\mathbb{R}} + b.1$, et donc $f : x \mapsto ax + b$, P est donc l'ensemble des applications affines.

3) Somme de sous-espaces vectoriels

Définition 20.4 (somme de s.e.v)
Soient F *et* G *deux s.e.v de* E*, on appelle somme de* F *et* G *l'ensemble noté* $\mathrm{F}+\mathrm{G}$ *et défini par :*
$$\mathrm{F}+\mathrm{G} = \{x \in \mathrm{E} \ / \ \exists\, u \in \mathrm{F}, v \in \mathrm{G}, x = u+v\}.$$

Théorème 20.3
Une somme de s.e.v de E *est un s.e.v de* E.

Preuve : $\mathrm{F}+\mathrm{G}$ est inclus dans E et contient le vecteur nul puisque celui-ci est dans F et dans G. Si x et y sont dans $\mathrm{F}+\mathrm{G}$ alors on peut écrire $x = x_{\mathrm{F}} + x_{\mathrm{G}}$ et $y = y_{\mathrm{F}} + y_{\mathrm{G}}$ avec $x_{\mathrm{F}}, y_{\mathrm{F}} \in \mathrm{F}$ et $\mathrm{X}_{\mathrm{G}}, y_{\mathrm{G}} \in \mathrm{G}$, on a $x+y = (x_{\mathrm{F}}+y_{\mathrm{F}}) + (x_{\mathrm{G}}+y_{\mathrm{G}})$ et $\lambda x = \lambda x_{\mathrm{F}} + \lambda x_{\mathrm{G}}$, comme F et G sont stables pour l'addition et le produit par les scalaires, on voit que $x+y$ et λx sont dans $\mathrm{F}+\mathrm{G}$. □

Remarque 20.1 – $\mathrm{F}+\mathrm{G}$ *est un s.e.v. de* E *qui contient à la fois* F *et* G.

☞**Exemples** :
- Dans $\mathbb{K}^3$, posons $i = (1,0,0), j = (0,1,0), k = (0,0,1)$, on peut vérifier que $\mathbb{K}^3 = \mathrm{Vect}\,[i] + \mathrm{Vect}\,[j,k] = \mathrm{Vect}\,[i,j] + \mathrm{Vect}\,[k] = \mathrm{Vect}\,[i,k] + \mathrm{Vect}\,[j]$.
- Soient $x, y \in \mathrm{E}$ deux vecteurs, on a $\mathrm{Vect}\,[x] + \mathrm{Vect}\,[y] = \mathrm{Vect}\,[x,y]$. Plus généralement, on peut remplacer x et y par deux familles de vecteurs de E.

4) Sommes directes

Définition 20.5 (somme directe)
Soient F, G *deux s.e.v de* E*, on dit que la somme* $\mathrm{F}+\mathrm{G}$ *est directe lorsque tout vecteur* x *de cette somme s'écrit* ***de manière unique*** *sous la forme* $x = x_{\mathrm{F}} + x_{\mathrm{G}}$ *avec* $x_{\mathrm{F}} \in \mathrm{F}$*, et* $x_{\mathrm{G}} \in \mathrm{G}$*. Si c'est le cas, la somme est notée* $\mathrm{F} \oplus \mathrm{G}$.

Théorème 20.4 (caractérisation des sommes directes)
Soient F *et* G *deux s.e.v de* E*, les assertions suivantes sont équivalentes :*

a) *la somme* $\mathrm{F}+\mathrm{G}$ *est directe.*

b) $\forall\, (x_{\mathrm{F}}, x_{\mathrm{G}}) \in \mathrm{F}\times\mathrm{G}$*, si* $x_{\mathrm{F}} + x_{\mathrm{G}} = 0_{\mathrm{E}}$ *alors* $x_{\mathrm{F}} = x_{\mathrm{G}} = 0_{\mathrm{E}}$.

c) $\mathrm{F} \cap \mathrm{G} = \{0_{\mathrm{E}}\}$.

Preuve : Celle-ci est simple et laissée en exercice. □

☞**Exemples** :
- Dans $\mathcal{F}(\mathbb{R},\mathbb{R})$ le s.e.v des fonctions paires et le s.e.v des fonctions impaires sont en somme directe.
- Dans $\mathbb{K}^3$ le plan P d'équation $x+y+z = 0$ et la droite engendrée par le vecteur $i = (1,1,1)$ sont en somme directe, mais P n'est pas en somme directe avec le plan P' engendré par i et $j = (1,-1,1)$.

5) S.e.v. supplémentaires

Définition 20.6 (s.e.v supplémentaires)
Soient F *et* G *deux s.e.v de* E *, on dit que* F *et* G *sont supplémentaires lorsque* $\mathrm{F}\oplus\mathrm{G} = \mathrm{E}$*. Ce qui signifie que* $\mathrm{E} = \mathrm{F}+\mathrm{G}$ *et la somme* $\mathrm{F}+\mathrm{G}$ *est directe, ou encore : tout vecteur de* E *s'écrit de manière unique comme somme d'un vecteur de* F *et d'un vecteur de* G.

☞**Exemples** :
- Dans $\mathcal{F}(\mathbb{R},\mathbb{R})$ le s.e.v des fonctions paires et le s.e.v des fonctions impaires sont supplémentaires.
- Dans $\mathrm{E} = \mathcal{C}^0([a;b],\mathbb{R})$ le s.e.v $\mathrm{H} = \{f \in \mathrm{E} \ / \ \int_a^b f = 0\}$ et le s.e.v $\mathrm{G} = \mathrm{Vect}\,[\mathrm{id}_{\mathbb{R}}]$ sont supplémentaires.

Définition 20.7

Un hyperplan de E *est un s.e.v. de* E *qui admet une droite vectorielle comme supplémentaire :*
Un s.e.v. H *est un hyperplan de* E *si et seulement si* $\exists x_0 \in E \setminus \{0\}, E = H \oplus \text{Vect}[x_0]$.

Exemples :

- Dans $E = \mathcal{C}^0([a;b], \mathbb{R})$ le s.e.v $H = \{f \in E \;/\; \int_a^b f = 0\}$ et le s.e.v $G = \text{Vect}[\text{id}_\mathbb{R}]$ sont supplémentaires donc H est un hyperplan de E.
- Dans $E = \mathbb{K}^n$ le s.e.v $H = \{(x_1, \ldots, x_n) \,/\, x_1 + \cdots + x_n = 0\}$ et le s.e.v $G = \text{Vect}[(1, 0, \ldots, 0)]$ sont supplémentaires, donc H est un hyperplan de E.

Théorème 20.5 (propriété)

Soit H *un s.e.v de* E*, les assertions suivantes sont équivalentes :*

a) H *est un hyperplan de* E.

b) $\forall\, x_0 \in E \setminus H, H \oplus \text{Vect}[x_0] = E$.

Preuve : $b) \implies a)$: rien à faire.

$a) \implies b)$: il existe $y \in E \setminus \{0\}$ tel que $E = H \oplus \text{Vect}[y]$. Soit $x_0 \in E \setminus H$, alors il existe $h \in H$ et $\alpha \in \mathbb{K}$ tels que $x_0 = h + \alpha y$, avec $\alpha \neq 0$ (sinon on aurait $x_0 \in H$). Comme $x_0 \notin H$, il est facile de voir que $H \cap \text{Vect}[x_0] = \{0\}$. Soit $x \in E$, alors il existe $u \in H$ et $\beta \in \mathbb{K}$ tel que $x = u + \beta y$, d'où $x = u + \frac{\beta}{\alpha}(x_0 - h) = (u - \frac{\beta}{\alpha}h) + \frac{\beta}{\alpha}x_0 \in H + \text{Vect}[x_0]$, donc $E = H \oplus \text{Vect}[x_0]$. □

Remarque 20.2 – *Le théorème précédent montre que pour un hyperplan de* E*, il n'y a pas unicité d'une droite vectorielle supplémentaire.*

III APPLICATIONS LINÉAIRES

1) Définition, noyau

Définition 20.8

Soient E *et* F *deux* $\mathbb{K}$*-e.v. et soit* $f : E \to F$ *une application, on dit que* f *est une application linéaire (ou morphisme de* $\mathbb{K}$*-espaces vectoriels), lorsque :*

$$\forall\, x, y \in E, \forall\, \lambda \in \mathbb{K}, f(x+y) = f(x) + f(y) \text{ et } f(\lambda.x) = \lambda.f(x).$$

Si de plus, f *est bijective, alors on dit que* f *est un isomorphisme (d'espaces vectoriels). L'ensemble des applications linéaires de* E *vers* F *est noté* $\mathcal{L}(E, F)$.

Remarque 20.3 – *Les applications linéaires de* $\mathbb{K}$ *dans* $\mathbb{K}$ *sont les applications de la forme* $f(x) = ax$ $(a \in \mathbb{K})$*, car* $f(x) = xf(1)$.

Exemples :

- L'application nulle (notée 0) de E vers F est linéaire.
- L'application identité de E : $\text{id}_E : E \to E$ définie par $\text{id}_E(x) = x$, est linéaire bijective (et $(\text{id}_E)^{-1} = \text{id}_E$).
- Soit $\lambda \in \mathbb{K}^*$, l'homothétie de rapport λ : $h_\lambda : E \to E$, définie par $h_\lambda(x) = \lambda.x$, est linéaire et bijective. Sa réciproque est l'homothétie de rapport $1/\lambda$. L'ensemble des homothéties de E est un groupe pour la loi ∘ car c'est un sous-groupe du groupe des permutations de E.
- L'application $f : \mathbb{K}^2 \to \mathbb{K}^2$ définie par $f(x, y) = (x; -y)$ est un isomorphisme de $\mathbb{K}^2$ sur lui-même.

À retenir

$f \in \mathcal{L}(E, F)$ alors $f(0_E) = 0_F$ et $\forall x \in E$, $f(-x) = -f(x)$.

Définition 20.9 (vocabulaire)

- *Une application linéaire de* E *vers* E *est appelée un endomorphisme de* E*. L'ensemble des endomorphismes de* E *est noté* $\mathcal{L}(E)$ *(on a donc* $\mathcal{L}(E) = \mathcal{L}(E, E)$*).*
- *Un isomorphisme de* E *vers* E *est appelé un automorphisme de* E*. L'ensemble des automorphismes de* E *est noté* GL(E) *et appelé groupe linéaire de* E.

– *Une application linéaire de* E *vers* $\mathbb{K}$ *est appelée une forme linéaire sur* E*. L'ensemble des formes linéaires sur* E *est noté* E^* *et appelé dual de* E *(on a donc* $E^* = \mathcal{L}(E, \mathbb{K})$*).*

☞**Exemples :**

- $\mathrm{id}_E \in GL(E), \forall\, \lambda \in \mathbb{K}^*, h_\lambda \in GL(E)$.
- Soit $E = \mathcal{C}^0([0;1], \mathbb{R})$ et $\phi : E \to \mathbb{R}$ définie par $\phi(f) = \int_0^1 f(t)\, dt$, alors ϕ est une forme linéaire sur E.
- Soit $E = \{u \in \mathcal{F}(\mathbb{N}, \mathbb{C}) \ / \ (u_n) \text{ converge}\}$ est un $\mathbb{C}$-e.v. et l'application $\phi : E \to \mathbb{C}$ définie par $\phi(u) = \lim u_n$, est une forme linéaire sur E.
- Soient $a, b, c \in \mathbb{K}$, l'application $\phi : \mathbb{K}^3 \to \mathbb{K}$ définie par $\phi(x, y, z) = ax + by + cz$, est une forme linéaire sur $\mathbb{K}^3$. En exercice, montrer la réciproque, c'est à dire que toutes les formes linéaires sur $\mathbb{K}^3$ sont de ce type.

Définition 20.10 (Noyau d'une application linéaire)

Soit $f \in \mathcal{L}(E, F)$*, on appelle noyau de* f *l'ensemble noté* $\ker(f)$ *et défini par :*

$$\ker(f) = \{x \in E \ / \ f(x) = 0_F\}$$

Le noyau de f *contient toujours* 0_E*.*

☞**Exemples :**

- Le noyau d'une application linéaire bijective est $\{0_E\}$.
- Le noyau de l'application linéaire $d \colon \mathbb{K}[X] \to \mathbb{K}[X]$ définie par $d(P) = P'$ est $\ker(d) = \mathbb{K}$.
- Le noyau de l'application linéaire $f \colon \mathbb{K}^3 \to \mathbb{K}^2$ définie par $f(x, y, z) = (x + y + z, x - 2y - z)$ est $\ker(f) = \{(x, 2x, -3x) \ / \ x \in \mathbb{K}\} = \mathrm{Vect}\,[(1, 2, -3)]$.

2) Propriétés

Il est facile de vérifier les propriétés suivantes :

- $f \in \mathcal{L}(E, F)$ est injective si et seulement si $\ker(f) = \{0_E\}$.
- La composée de deux applications linéaires est linéaire. On en déduit que $GL(E)$ est stable pour la loi $\circ$.
- Si $f \in \mathcal{L}(E, F)$ est un isomorphisme, alors $f^{-1} \in \mathcal{L}(F, E)$. On en déduit que $GL(E)$ est stable par symétrisation, *i.e.* si $f \in GL(E)$, alors $f^{-1} \in GL(E)$.
- $(GL(E), \circ)$ est un groupe (non abélien en général), c'est en fait un sous-groupe du groupe des permutations de E : $(S_E, \circ)$.
- Si $f, g \in \mathcal{L}(E, F)$ et si $\lambda \in \mathbb{K}$, alors $f + g$ et $\lambda.f$ sont linéaires. On en déduit que $(\mathcal{L}(E, F), +, .)$ est un $\mathbb{K}$-e.v. (s.e.v. de $\mathcal{F}(E, F)$).
- $(\mathcal{L}(E), +, \circ)$ est un anneau, la loi $\circ$ jouant le rôle d'une multiplication.

Remarque 20.4 –

- *En général, l'anneau* $\mathcal{L}(E)$ *n'est pas commutatif. Le groupe des inversibles de cet anneau est* $GL(E)$*.*
- *La loi* $\circ$ *jouant le rôle d'une multiplication, on adopte les notations usuelles des anneaux pour les puissances, i.e. si* $u \in \mathcal{L}(E)$ *et si* n *est entier, alors :*

$$u^n = \begin{cases} \mathrm{id}_E & \textit{si } n = 0 \\ u \circ \cdots \circ u & n \textit{ fois si } n > 0 \\ u^{-1} \circ \cdots \circ u^{-1} & -n \textit{ fois si } u \textit{ est inversible et } n < 0 \end{cases},$$

de plus si $u, v \in \mathcal{L}(E)$ *commutent (i.e.* $u \circ v = v \circ u$*), alors on peut utiliser le binôme de Newton :*

$$(u + v)^n = \sum_{k=0}^{n} \binom{n}{k} u^k \circ v^{n-k}$$

- *Soit* $E = \mathbb{K}^2$ *et* $f : (x; y) \mapsto (y; 0)$*, on vérifie facilement que* $f \in \mathcal{L}(E)$ *et que* $f^2 = 0$ *(application nulle), pourtant* $f \neq 0$*. Cet exemple montre qu'en général* $\mathcal{L}(E)$ *n'est pas un anneau intègre.*

3) S.e.v. et applications linéaires

Remarque 20.5 – *Un s.e.v. de E est stable par combinaisons linéaires.*

Théorème 20.6 (image d'une combinaison linéaire par une application linéaire)
Soit E un $\mathbb{K}$-e.v et soit $(x_i)_{1\leqslant i\leqslant n}$ une famille de vecteurs de E. Soit $f \in \mathcal{L}(E,F)$, alors l'image par f d'une combinaison linéaire de la famille $(x_i)_{1\leqslant i\leqslant n}$ est une combinaison linéaire de la famille $(f(x_i))_{1\leqslant i\leqslant n}$ (dans F) avec les mêmes coefficients.

Preuve : Par récurrence sur n : pour $n = 1$ il n'y a rien à démontrer. Supposons le théorème vrai au rang n, et soit $x = \lambda_1 x_1 + \cdots + \lambda_{n+1}x_{n+1}$, f étant linéaire, on peut écrire $f(x) = f(\lambda_1 x_1 + \cdots + \lambda_n x_n) + \lambda_{n+1}.f(x_{n+1})$, on applique alors l'hypothèse de récurrence pour conclure. □

Théorème 20.7 (noyau et image d'une application linéaire)
Si $f \in \mathcal{L}(E,F)$ alors $\ker(f)$ est un s.e.v de E et $\mathrm{Im}(f)$ est un s.e.v de F.

Preuve : Celle-ci est simple et laissée en exercice. □

À retenir
Soit $f \in \mathcal{L}(E,F)$ alors f est un isomorphisme si et seulement si $\ker(f) = \{0_E\}$ et $\mathrm{Im}(f) = F$.

Théorème 20.8 (image d'un s.e.v par une application linéaire)
Soit H un s.e.v de E et $f \in \mathcal{L}(E,F)$, alors $f(H)$ (ensemble des images par f des éléments de H) est un s.e.v de F.

Preuve : Il suffit de considérer la restriction de f à H : $g : H \to F$ définie par $\forall x \in H, g(x) = f(x)$, il est clair que g est linéaire et que $f(H) = \mathrm{Im}(g)$, on peut appliquer alors le théorème précédent. □

Théorème 20.9 (image réciproque d'un s.e.v par une application linéaire)
Soit H un s.e.v de F et soit $f \in \mathcal{L}(E,F)$ alors $f^{-1}(H)$ (ensemble des antécédents des éléments de H par f) est un s.e.v de E.

Preuve : Celle-ci est simple et laissée en exercice. □

☞**Exemples** :
- $H = \{f \in \mathcal{C}^0([a;b],\mathbb{R}) \ / \ \int_a^b f = 0\}$ est un s.e.v de $\mathcal{C}^0([a;b],\mathbb{R})$, car c'est le noyau de la forme linéaire $\phi : f \mapsto \int_a^b f$.
- $H = \{(x,y,z) \in \mathbb{K}^3 \ / \ ax+by+cz = 0\}$ est un s.e.v de $\mathbb{K}^3$ car c'est le noyau de la forme linéaire $\phi : (x,y,z) \mapsto ax+by+cz$.
- $H = \{(x,y,z) \in \mathbb{K}^3 \ / \ 2x+y-z = 0 \text{ et } 3x-2z = 0\}$ est un s.e.v de $\mathbb{K}^3$ car c'est l'intersection des noyaux des deux formes linéaires : $\phi_1 : (x,y,z) \mapsto 2x+y-z$ et $\phi_2 : (x,y,z) \mapsto 3x-2z$.

Théorème 20.10
Soient $u \in \mathcal{L}(E,F)$ et $v \in \mathcal{L}(F,G)$, alors : $v \circ u = 0 \iff \mathrm{Im}(u) \subset \ker(v)$.

Preuve : Celle-ci est simple et laissée en exercice. □

4) Hyperplans et applications linéaires

Théorème 20.11 (caractérisation des hyperplans)
Soit H un s.e.v de E, H est un hyperplan de E si et seulement si H est le noyau d'une forme linéaire sur E non constamment nulle.

Preuve : Si H est un hyperplan de E, soit $x_0 \in E \setminus H$, on sait que $E = H \oplus \mathrm{Vect}[x_0]$. Pour $x \in E$, il existe $y_x \in H$ et $\lambda_x \in \mathbb{K}$, **uniques** tels que $x = y_x + \lambda_x x_0$. Posons $\phi(x) = \lambda_x$. On définit ainsi une application non nulle de E vers $\mathbb{K}$, on peut vérifier ensuite que ϕ est bien linéaire (laissé en exercice), $x \in \ker(\phi) \iff \lambda_y = 0 \iff x = y_x \iff x \in H$, donc $\ker(\phi) = H$, H est le noyau de la forme linéaire ϕ, celle-ci n'est pas constamment nulle puisque $\phi(x_0) = 1$.

Réciproque : soit ϕ une forme linéaire (non nulle) telle que $\ker(\phi) = H$, soit $x_0 \in E \setminus H$, on a $\phi(x_0) = \alpha \neq 0$, soit $x \in E$ et $\lambda = \phi(x)$, posons $y = x - \frac{\lambda}{\alpha}x_0$, on a $\phi(y) = 0$, donc $y \in H$ et de plus $x = y + \frac{\lambda}{\alpha}x_0$, ce qui prouve que $E = H + \mathrm{Vect}[x_0]$, et comme $x_0 \notin H$, la somme est directe, donc H est un hyperplan. □

IV PROJECTIONS, SYMÉTRIES

1) Projecteurs

Définition 20.11

Soit E un $\mathbb{K}$-e.v, une projection dans E (ou un projecteur de E) est un endomorphisme p de E tel que $p^2 = p$ (i.e. $p \circ p = p$).

☞**Exemples :**
- $E = \mathbb{K}^2$ et $p(x, y) = (x, 0)$.
- $E = \mathcal{F}(\mathbb{R}, \mathbb{R})$ et p qui à $f \in E$ associe $p(f) : x \mapsto \frac{f(x)+f(-x)}{2}$.

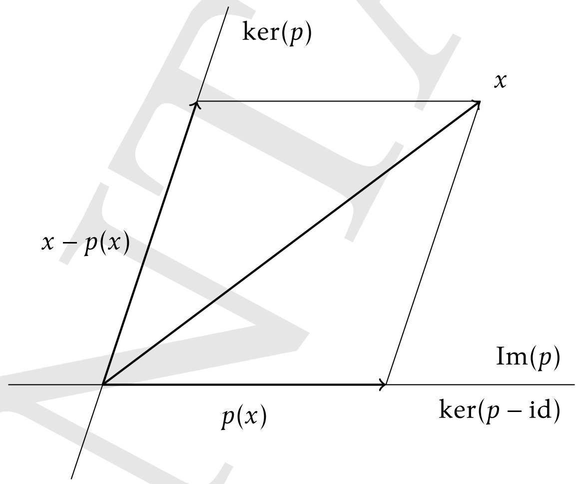
Remarque 20.6 – *Invariants d'un endomorphisme : si $f \in \mathcal{L}(E)$, alors $x \in E$ est invariant par f (ou un point fixe de f) si et seulement si $f(x) = x$, ce qui équivaut à $(f - \mathrm{id}_E)(x) = 0_E$, ou encore $x \in \ker(f - \mathrm{id}_E)$. L'ensemble des points fixes de f est donc le s.e.v $\ker(f - \mathrm{id}_E)$.*

Théorème 20.12 (caractérisation des projections)

$p \in \mathcal{L}(E)$ est un projecteur $\iff E = \ker(p) \oplus \ker(p - \mathrm{id}_E)$. Si c'est le cas, alors $\mathrm{Im}(p) = \ker(p - \mathrm{id}_E)$ et on dit que p est la projection sur $\mathrm{Im}(p)$ parallèlement à $\ker(p)$. Tout vecteur x de E se décompose de la manière suivante : $x = (x - p(x)) + p(x)$, avec $x - p(x) \in \ker(p)$ et $p(x) \in \ker(p - \mathrm{id}_E)$.

Preuve : Si p est un projecteur, soit $x \in \ker(p) \cap \ker(p - \mathrm{id}_E)$, alors $p(x) = 0_E = x$, donc la somme est directe. Soit $x \in E$, alors $p(x - p(x)) = p(x) - p^2(x) = 0_E$, donc $x - p(x) \in \ker(p)$, on a alors $x = (x - p(x)) + p(x)$ et $p(x) \in \ker(p - \mathrm{id}_E)$, donc $E = \ker(p) \oplus \ker(p - \mathrm{id}_E)$. De la définition, il découle que $\mathrm{Im}(p) \subset \ker(p - \mathrm{id}_E)$, l'inclusion étant évidente, on a $\mathrm{Im}(p) = \ker(p - id_E)$.

Réciproque : si $E = \ker(p) \oplus \ker(p - \mathrm{id}_E)$, soit $x \in E$, alors $x = y + z$ avec $y \in \ker(p)$ et $z \in \ker(p - \mathrm{id}_E)$, d'où $p(x) = p(y) + p(z) = p(z) = z$, et donc $p^2(x) = p(z) = z = p(x)$, ce qui prouve que p est un projecteur. □



☞**Exemples :**
- Dans le premier exemple, p est la projection sur la droite $\mathrm{Vect}\,[(1, 0)]$ et parallèlement à la droite $\mathrm{Vect}\,[(0, 1)]$.
- Dans le deuxième exemple, p est la projection sur le s.e.v des fonctions paires, parallèlement au s.e.v des fonctions impaires.

Théorème 20.13 (projection associée à une décomposition)

Si F et G sont deux s.e.v de E supplémentaires ($E = F \oplus G$), alors il existe une unique projection p telle que $\mathrm{Im}(p) = F$ et $\ker(p) = G$, i.e. qui soit la projection sur F parallèlement à G.

Preuve : Pour $x \in E$, il existe $x_F \in F$ et $x_G \in G$, **uniques** tels que $x = x_F + x_G$, on pose alors $p(x) = x_F$, ce qui définit une application de E dans E. On vérifie facilement que p est linéaire, et comme $x_F \in F$, on a par définition même de p, que $p^2(x) = x_F = p(x)$, donc p est bien un projecteur. On a $p(x) = 0_E \iff x_F = 0_E \iff x = x_G \iff x \in G$, donc $\ker(p) = G$, d'autre part, $p(x) = x \iff x = x_F \iff x \in F$, donc $\ker(p - \mathrm{id}_E) = F$, ce qui termine la preuve. □

★**Exercice 20.1**

1/ Soit $E = \mathbb{K}^3$, $F = \{(x, y, z) \in E \;/\; z = 0\}$ *et* $G = \text{Vect}[(1, 1, 1)]$. *Montrer que* F *et* G *sont supplémentaires, et déterminer l'expression analytique de la projection sur* F *parallèlement à* G.

2/ Soit p *un projecteur de* E, *montrer que* $q = \text{id}_E - p$ *est un projecteur, préciser ses éléments caractéristiques.*

2) Symétries

Définition 20.12

Soit E *un* $\mathbb{K}$*-e.v, une symétrie de* E *est un endomorphisme* s *tel que* $s^2 = \text{id}_E$ *(involution linéaire).*

☞**Exemples :**

- Dans $E = \mathbb{K}^2$, l'application s définie par $s(x, y) = (y, x)$ est une symétrie.
- Dans $E = \mathcal{F}(\mathbb{R}, \mathbb{R})$ l'application s définie par $s(f)$ est la fonction qui à $s(f) : x \mapsto f(-x)$, est une symétrie.

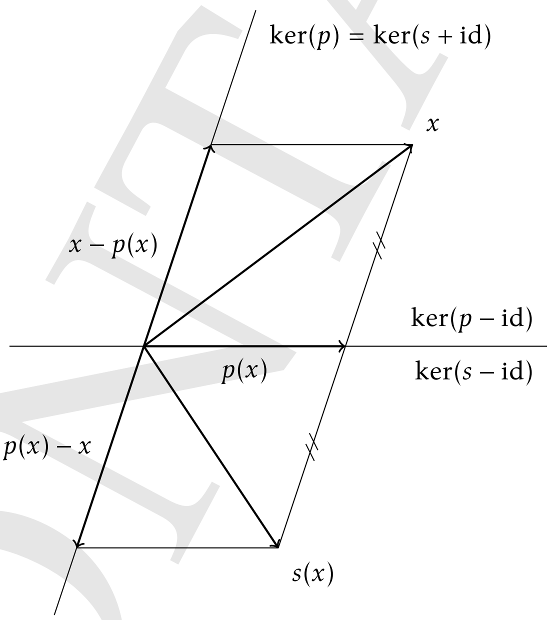
Théorème 20.14 (caractérisation des symétries)

Soit $s \in \mathcal{L}(E)$, *s est une symétrie* $\iff E = \ker(s - \text{id}_E) \oplus \ker(s + \text{id}_E)$. *Ce qui revient à dire que l'application* $p = \frac{1}{2}(\text{id}_E + s)$ *est une projection. Si c'est le cas, on dit que s est la symétrie par rapport à* $\ker(s - \text{id}_E)$ *(ensemble des invariants) et parallèlement à* $\ker(s + \text{id}_E)$, *et on dit que p est la projection associée à s. Tout vecteur x de* E *se décompose de la manière suivante :*

$$x = \tfrac{1}{2}(x + s(x)) + \tfrac{1}{2}(x - s(x)),$$

avec $\frac{1}{2}(x + s(x)) \in \ker(s - \text{id}_E)$ *et* $\frac{1}{2}(x - s(x)) \in \ker(s + \text{id}_E)$.

Preuve : Posons $p = \frac{1}{2}(\text{id}_E + s)$, s est une symétrie équivaut à $s^2 = \text{id}_E$, c'est à dire $(2p - \text{id}_E)^2 = \text{id}_E$, ou encore $p^2 = p$, ce qui équivaut à dire que $E = \ker(p) \oplus \ker(p - \text{id}_E)$, et donc $E = \ker(s + \text{id}_E) \oplus \ker(s - \text{id}_E)$. □



Théorème 20.15 (symétrie associée à une décomposition)

Si F *et* G *sont deux s.e.v de* E *supplémentaires* ($E = F \oplus G$), *alors il existe une unique symétrie s telle que* $\ker(s - \text{id}_E) = F$ *et* $\ker(s + \text{id}_E) = G$, *i.e. qui soit la symétrie par rapport à* F *et parallèlement à* G.

Preuve : Soit p la projection sur F parallèlement à G, posons $s = 2p - \text{id}_E$, on sait alors que s est une symétrie et $\ker(s - \text{id}_E) = \ker(p - \text{id}_E) = F$ et $\ker(s + \text{id}_E) = \ker(p) = G$, donc s existe. Réciproquement, si s existe, alors la projection associée est nécessairement la projection sur F parallèlement à G, or celle-ci est unique, c'est p, donc s est unique. □

☞**Exemples :**

- Dans le premier exemple ci-dessus, s est la symétrie par rapport à la droite $\mathrm{Vect}[(1,1)]$ et parallèlement à la droite $\mathrm{Vect}[(1,-1)]$.
- Dans le deuxième exemple, s est la symétrie par rapport au s.e.v des fonctions paires, et parallèlement au s.e.v des fonctions impaires.

V FAMILLE INFINIE DE VECTEURS

1) Définition

Définition 20.13 (généralisation)

Soit $(x_i)_{i\in I}$ une famille de vecteurs de E. On appelle combinaison linéaire de la famille, tout vecteur de E pouvant s'écrire comme combinaison linéaire d'un ***nombre fini*** *de vecteurs de la famille. Notation :*

$$\mathrm{Vect}\left[(x_i)_{i\in I}\right] = \left\{\sum_{j\in J} \alpha_j x_j \ / \ J \text{ partie finie de } I \text{ et } \forall j \in J, \alpha_j \in \mathbb{K}\right\}.$$

Remarque 20.7 – *Si X est une partie de E, on notera Vect [X] l'ensemble des combinaisons linéaires des vecteurs de X. On peut écrire :*

$$\mathrm{Vect}[X] = \left\{\sum_{x\in X} \alpha_x x \ / \ (\alpha_x)_{x\in X} \text{ est une famille de scalaires tous nuls sauf un nombre fini}\right\}.$$

De telles familles de scalaires sont appelées ***familles à support fini****. Le théorème 20.6 se généralise alors ainsi :*

$$\boxed{\text{Si } f \in \mathcal{L}(E,F) \text{ alors } f\Big(\sum_{x\in X} \alpha_x x\Big) = \sum_{x\in X} \alpha_x f(x), \text{ pour toute famille de scalaires } (\alpha_x)_{x\in X} \text{ à support fini.}}$$

2) Sous-espace engendré par une famille infinie

Théorème 20.16 (sous-espace engendré)

Soit $(x_i)_{i\in I}$ une famille de vecteurs de E, $\mathrm{Vect}\left[(x_i)_{i\in I}\right]$ est un s.e.v de E. C'est même le plus petit (pour l'inclusion) s.e.v de E qui contient tous les vecteurs de cette famille. On l'appelle s.e.v engendré par $(x_i)_{i\in I}$.

Preuve : Celle-ci est simple et laissée en exercice. □

☞**Exemples** :
- $\mathbb{K}[X] = \mathrm{Vect}\left[(X^n)_{n\in\mathbb{N}}\right]$.
- $\mathrm{Vect}\left[\left(\frac{1}{(X-a)^n}\right)_{(n,a)\in\mathbb{N}^*\times\mathbb{C}}\right] = \{F \in \mathbb{C}(X) \ / \ \deg(F) < 0\}$.

VI SOLUTION DES EXERCICES

Solution 20.1

1/ Soit $(x,y,z) \in E$, alors $(x,y,z) = (z,z,z) + (x-z, y-z, 0)$, le premier triplet est dans G, et le second dans F, donc $E = G + F$. Si $(x,y,z) \in F\cap G$ alors $x = y = z$ et $z = 0$ donc $F \cap G = \{(0,0,0)\}$ d'où $E = F \oplus G$ et le projeté de (x,y,z) sur F parallèlement à G est $p(x,y,z) = (x-z, x-z, 0)$ (composante sur F dans la décomposition).

2/ On sait que p est la projection sur $F = \mathrm{Im}(p) = \ker(p - \mathrm{id}_E)$ et parallèlement à $G = \ker(p)$, on a $E = F \oplus G$ et pour tout x de E, $x = p(x) + x - p(x)$ avec $x - p(x) \in G$ et $p(x) \in F$, donc par définition, $x - p(x)$ est le projeté de x sur G parallèlement à F, autrement dit, $q = \mathrm{id}_E - p$ est la projection sur G parallèlement à F.

Chapitre 21

Séries numériques

Sommaire

I GÉNÉRALITÉS

Dans tout le chapitre $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$. On cherche à généraliser la notion de somme d'éléments de $\mathbb{K}$ à un nombre infini de termes.

1) Définitions

Définition 21.1

*Soit $(u_n)_{n\in\mathbb{N}}$ une suite d'éléments de $\mathbb{K}$. On appelle **série de terme général** u_n la suite $(S_n)_{n\in\mathbb{N}}$ définie par $S_n = \sum_{k=0}^{n} u_k$. Le nombre S_n est appelé **somme partielle de rang** n (d'ordre n) de la série. La série de terme général u_n est notée $\sum_{n\in\mathbb{N}} u_n$.*

☞**Exemples** :

- $\sum_{n\geqslant 0} q^n$: série géométrique de raison q.
- $\sum_{n\in\mathbb{N}} \frac{z^n}{n!}$ avec $z \in \mathbb{K}$: série exponentielle.
- $\sum_{n\geqslant 1} \frac{1}{n^\alpha}$ où $\alpha \in \mathbb{R}$: séries de Riemann.

Remarque 21.1 – *L'ensemble des séries de $\mathbb{K}$ n'est autre que $\mathcal{F}(\mathbb{N}, \mathbb{K})$. Pour les opérations usuelles sur les suites, on sait que c'est un e.v. sur $\mathbb{K}$. Soient $\sum_{n\in\mathbb{N}} u_n$ et $\sum_{n\in\mathbb{N}} v_n$ deux séries, $\lambda \in \mathbb{K}$, la somme des deux séries $\sum_{n\in\mathbb{N}} u_n + \sum_{n\in\mathbb{N}} v_n$ est la série de terme général $u_n + v_n$, et la série $\lambda \sum_{n\in\mathbb{N}} u_n$ est la série de terme général λu_n.*

Définition 21.2 (convergence d'une série)

- *On dit que la série de terme général u_n est* **convergente** *lorsque la suite des sommes partielles $(S_n)_{n\in\mathbb{N}}$ est convergente. Si c'est le cas, la limite de la suite des sommes partielles est appelée* **somme de la série** *et notée $\sum\limits_{n=0}^{+\infty} u_n$. On a donc $\sum\limits_{n=0}^{+\infty} u_n = \lim\limits_{n\to+\infty} S_n = \lim\limits_{n\to+\infty} \sum\limits_{k=0}^{n} u_k$. La différence entre la somme de la série et la somme partielle S_n est appelée* **reste de rang** *n, noté $R_n = \sum\limits_{n=0}^{+\infty} u_n - \sum\limits_{k=0}^{n} u_k = \sum\limits_{k=n+1}^{+\infty} u_k$.*
- *Une série qui n'est pas convergente, est dite* **divergente**.
- *Déterminer la nature d'une série, c'est déterminer si elle est convergente ou divergente.*

Attention !

On prendra garde à ne pas confondre la notation $\sum\limits_{n\in\mathbb{N}} u_n$ qui désigne la série (c'est à dire la suite des sommes partielles), avec la notation $\sum\limits_{n=0}^{+\infty} u_n$ qui désigne la somme de la série, **si celle-ci converge** *(c'est la limite des sommes partielles lorsque celle-ci existe dans $\mathbb{K}$).*

☞**Exemples :**

- La série de terme général $u_n = n$: $S_n = \sum\limits_{k=0}^{n} k = \frac{n(n+1)}{2}$ donc $\lim S_n = +\infty$ et la série $\sum_{n\in\mathbb{N}} n$ est divergente.
- La série de terme général $u_n = (-\frac{1}{3})^n$: $S_n = \sum\limits_{k=0}^{n} (-\frac{1}{3})^k = \frac{1-(\frac{1}{3})^{n+1}}{1+\frac{1}{3}}$, or $|-\frac{1}{3}| < 1$ donc $\lim S_n = \frac{3}{4}$. La série $\sum_{n\in\mathbb{N}}(-\frac{1}{3})^n$ est convergente et sa somme est $\sum\limits_{n=0}^{+\infty}(-\frac{1}{3})^n = \frac{3}{4}$.

★**Exercice 21.1** *Étudier la nature de la série de terme général $u_n = (-1)^n$.*

2) Premières propriétés

Théorème 21.1 (lien suite-série)

L'application de $\varphi\colon \mathcal{F}(\mathbb{N},\mathbb{K}) \to \mathcal{F}(\mathbb{N},\mathbb{K})$ qui à chaque suite $(u_n)_{n\in\mathbb{N}}$ associe $\varphi(u) = S$ la suite des sommes partielles $\left(S_n = \sum\limits_{k=0}^{n} u_k\right)_{n\in\mathbb{N}}$ est un automorphisme.

Preuve : Soient u et v deux suites, et $\alpha \in \mathbb{K}$, soit (U_n) la suite des sommes partielles de la série $\sum\limits_{n\in\mathbb{N}} u_n$, *i.e.* $\varphi(u) = U$, et (V_n) la suite des sommes partielles de la série $\sum\limits_{n\in\mathbb{N}} v_n$, *i.e.* $\varphi(v) = V$, alors $\forall n \in \mathbb{N}$, $\sum\limits_{k=0}^{n} (\alpha u_k + v_k) = \alpha\left(\sum\limits_{k=0}^{n} \alpha u_k\right) + \left(\sum\limits_{k=0}^{n} +v_k\right)$, donc $(\alpha U_n + \beta V_n)$ est la suite des sommes partielles de la série $\sum\limits_{n\in\mathbb{N}} \alpha u_n + \beta v_n$, c'est à dire $\varphi(\alpha u + v) = \alpha\varphi(u) + \varphi(v)$. Ce qui prouve la linéarité.

Si $u \in \ker(\varphi)$, alors $\varphi(u) = 0$ *i.e.* $\forall n \in \mathbb{N}$, $U_n = 0$, ce qui entraîne que $u_0 = U_0 = 0$, $u_1 = U_1 - U_0 = 0$, et plus généralement $u_n = U_n - U_{n-1}0$, donc $u = 0$, φ est injective.

Soit $(s_n)_{n\in\mathbb{K}}$ une suite, soit $u_0 = s_0$ et pour $n > 0$, $u_n = s_n - s_{n-1}$, on a alors $\sum\limits_{k=0}^{n} u_k = s_0 + (s_1 - s_0) + \cdots + (s_n - s_{n-1}) = s_n$ (somme télescopique), ce qui montre que $\varphi(u) = s$, l'application φ est donc surjective. □

Théorème 21.2 (séries convergentes)

L'ensemble des séries convergentes de $\mathbb{K}$ est un espace vectoriel sur $\mathbb{K}$. De plus l'application qui à chaque série convergente associe sa somme, est (une forme) linéaire. Plus précisément, si $\sum\limits_{n\in\mathbb{N}} u_n$ et $\sum\limits_{n\in\mathbb{N}} v_n$ sont deux séries convergentes, alors $\forall \alpha, \beta \in \mathbb{K}$, la série $\sum\limits_{n\in\mathbb{N}} \alpha u_n + \beta v_n$ est convergente et :

$$\sum_{n=0}^{+\infty} \alpha u_n + \beta v_n = \alpha \sum_{n=0}^{+\infty} u_n + \beta \sum_{n=0}^{+\infty} v_n.$$

Preuve : Soit (U_n) la suite des sommes partielles de la série $\sum\limits_{n\in\mathbb{N}} u_n$ et (V_n) la suite des sommes partielles de la série $\sum\limits_{n\in\mathbb{N}} v_n$. Alors $(\alpha U_n + \beta V_n)$ est la suite des sommes partielles de la série $\sum\limits_{n\in\mathbb{N}} \alpha u_n + \beta v_n$. Le reste découle des propriétés des suites convergentes. □

Attention !

• *La somme de deux séries divergentes peut très bien donner une série convergente (prendre deux séries divergentes opposées par exemple).*

• *La somme d'une série divergente et d'une série convergente est forcément une série divergente (raisonner par l'absurde).*

Théorème 21.3 **(convergence d'une série complexe)**

Soit $\sum\limits_{n\in\mathbb{N}} u_n$ *une série à terme complexe, celle-ci est convergente si et seulement si les deux séries réelles* $\sum\limits_{n\in\mathbb{N}} \mathrm{Re}(u_n)$ *et* $\sum\limits_{n\in\mathbb{N}} \mathrm{Im}(v_n)$ *sont convergentes, et auquel cas on a :* $\sum\limits_{n=0}^{+\infty} u_n = \sum\limits_{n=0}^{+\infty} \mathrm{Re}(u_n) + i\sum\limits_{n=0}^{+\infty} \mathrm{Im}(u_n)$.

Preuve : Soit (A_n) la suite des sommes partielles de la série $\sum\limits_{n\in\mathbb{N}} \mathrm{Re}(u_n)$ et (B_n) la suite des sommes partielles de la série $\sum\limits_{n\in\mathbb{N}} \mathrm{Im}(v_n)$. Alors $(U_n + iV_n)$ est la suite des sommes partielles de la série $\sum\limits_{n\in\mathbb{N}} u_n$. Le reste découle des propriétés des suites complexes convergentes. □

Théorème 21.4 **(condition nécessaire de convergence)**

Si la série $\sum\limits_{n\in\mathbb{N}} u_n$ *converge alors on a nécessairement* $\lim u_n = 0$, *mais* ***la réciproque est fausse***.

Preuve : Soit (S_n) la suite des sommes partielles, alors la suite (S_n) converge vers S_∞ la somme de la série, d'où $\lim U_n - U_{n-1} = S_\infty - S_\infty = 0$, c'est à dire $\lim u_n = 0$.

Pour la réciproque prenons $u_n = \frac{1}{\sqrt{n}}$, alors $S_n = \sum\limits_{k=1}^{n} \frac{1}{\sqrt{k}} \geqslant n\frac{1}{\sqrt{n}} = \sqrt{n}$, on en déduit que $S_n \to +\infty$. La série est divergente bien que le terme général u_n tende vers 0. □

Définition 21.3

On dit que la série $\sum\limits_{n\in\mathbb{N}} u_n$ *diverge* ***grossièrement*** *lorsque le terme général ne tend pas vers* 0.

☞**Exemples** :

– La série $\sum\limits_{n\in\mathbb{N}} (-1)^n$ est grossièrement divergente.

– La série $\sum\limits_{n\in\mathbb{N}} \frac{1}{n}$ (série harmonique) est divergente (mais pas grossièrement), en effet $S_{2n} - S_n = \sum\limits_{k=n+1}^{2n} \frac{1}{k} \geqslant \frac{n}{2n} = \frac{1}{2}$. Si la série était convergente alors la suite (S_n) convergerait vers la somme S_∞, mais alors $S_{2n} - S_n \to S_\infty - S_\infty = 0$, on aurait donc $0 \geqslant \frac{1}{2}$ ce qui est absurde. La série harmonique est donc divergente.

Rappel : dans le TD sur les suites, nous avons établi que $\sum\limits_{k=1}^{n} \frac{1}{k} \sim \ln(n)$.

3) Séries géométriques

Théorème 21.5 **(nature des séries géométriques)**

La série géométrique $\sum\limits_{n\in\mathbb{N}} z^n$ *avec* $z \in \mathbb{C}$, *est convergente si et seulement si* $|z| < 1$, *auquel cas sa somme est* $\sum\limits_{n=0}^{+\infty} z^n = \frac{1}{1-z}$.

Preuve : Si $|z| \geqslant 1$ alors $|z^n| \geqslant 1$ et donc la suite (z^n) ne peut pas tendre vers 0, la série est donc grossièrement divergente dans ce cas.

Si $|z| < 1$, alors $S_n = \sum\limits_{k=0}^{n} z^k = \frac{1-z^{n+1}}{1-z}$, or $z^{n+1} \to 0$ car $|z| < 1$, il en découle que $S_n \to \frac{1}{1-z}$. □

☞**Exemples** :

– La série $\sum\limits_{n\in\mathbb{N}} (-\frac{1}{3})^n$ est convergente et $\sum\limits_{n=0}^{+\infty} (-\frac{1}{3})^n = \frac{1}{1+\frac{1}{3}} = \frac{3}{4}$.

– La série $\sum\limits_{n\in\mathbb{N}} (-1)^n$ est grossièrement divergente.

Attention !
Pour les séries géométriques convergentes on prendra garde à l'indice du premier terme avant d'appliquer la formule !

II SÉRIES À TERMES POSITIFS

Dans ce paragraphe toutes les séries sont à **termes réels positifs**. Pour ces séries, il faut remarquer que les sommes partielles forment **une suite croissante**. La définition peut s'étendre aux séries dont le terme général est réel positif **à partir d'un certain rang** seulement.

1) Critère de convergence

Théorème 21.6
Une série à termes positifs est convergente si et seulement si la suite des sommes partielles est majorée dans $\mathbb{R}$.

Preuve : La série converge si et seulement si la suite (S_n) converge, or cette suite est croissante, on sait donc qu'elle converge si et seulement si elle est majorée. La résultat en découle. □

Remarque 21.2 –
• *Lorsque la suite (S_n) des sommes partielles n'est pas majorée, on a $S_n \to +\infty$ puisque cette suite est croissante.*
• *Le critère s'applique aux séries à termes positifs **à partir d'un certain rang**.*

☞**Exemple** : La série $\sum\limits_{n\in\mathbb{N}} \frac{1}{n^2}$ est convergente. En effet, pour $n \geqslant 2$ on a $\frac{1}{k^2} \leqslant \frac{1}{k(k-1)} = \frac{1}{k-1} - \frac{1}{k}$, on en déduit que $S_n \leqslant 1 + \sum\limits_{k=2}^{n} \frac{1}{k-1} - \frac{1}{k} = 2 - \frac{1}{n}$, et par conséquent $S_n \leqslant 2$.

Attention !
Le critère est en défaut lorsque la série n'est pas à termes positifs. Par exemple la série $\sum\limits_{n\in\mathbb{N}} (-1)^n$ est divergente alors que ses sommes partielles sont bornées.

2) Théorèmes de comparaison

Théorème 21.7
Soient $\sum\limits_{n\in\mathbb{N}} u_n$ et $\sum\limits_{n\in\mathbb{N}} v_n$ deux séries telles qu'à partir d'un certain rang N on ait $0 \leqslant u_n \leqslant v_n$, alors on a :
• *Si la $\sum\limits_{n\in\mathbb{N}} v_n$ converge alors la série $\sum\limits_{n\in\mathbb{N}} u_n$ converge et on a $\sum\limits_{n=N}^{+\infty} u_n \leqslant \sum\limits_{n=N}^{+\infty} v_n$.*
• *Si la série $\sum\limits_{n\in\mathbb{N}} u_n$ diverge alors la série $\sum\limits_{n\in\mathbb{N}} v_n$ diverge.*

Preuve : D'après les hypothèses, on a $\sum\limits_{k=N}^{n} u_k \leqslant \sum\limits_{k=N}^{n} v_k$.
• Si la série $\sum\limits_{n\in\mathbb{N}} v_n$ converge, alors $(\sum\limits_{k=0}^{n} v_k)$ converge et donc $(\sum\limits_{k=N}^{n} v_k)$ converge, cette suite est donc majorée par un certain réel M, on en déduit que la somme partielle de la série $\sum\limits_{n\in\mathbb{N}} u_n$ est majorée, comme elle est à termes positifs à partir d'un certain rang, elle converge et l'inégalité sur les sommes en découle.
• Si la série $\sum\limits_{n\in\mathbb{N}} u_n$ diverge, alors $(\sum\limits_{k=N}^{n} u_k)$ est croissante divergente, donc de limite $+\infty$, on en déduit que $(\sum\limits_{k=0}^{n} v_k)$ tend vers $+\infty$, cette suite n'est donc pas majorée, donc la série $\sum\limits_{n\in\mathbb{N}} v_n$ est divergente puisqu'elle est à terme positif à partir d'un certain rang. □

☞**Exemples** :
– La série de terme général $\frac{1}{n^3}$ est convergente car $\frac{1}{n^3} \leqslant \frac{1}{n^2}$ et celle-ci est une série est convergente.
– Si $\alpha < 1$, alors la série de terme général $\frac{1}{n^\alpha}$ est divergente car $\frac{1}{n} \leqslant \frac{1}{n^\alpha}$ et on sait que la série harmonique est divergente.

Théorème 21.8 (utilisation des relations de comparaison)

Soient (u_n) et (v_n) deux suites à termes positifs (à partir d'un certain rang).

- *Si $u_n = \mathrm{O}(v_n)$ et si $\sum\limits_{n\in\mathbb{N}} v_n$ converge, alors la série $\sum\limits_{n\in\mathbb{N}} u_n$ converge.*
- *Si $u_n = \mathrm{O}(v_n)$ et si $\sum\limits_{n\in\mathbb{N}} u_n$ diverge, alors la série $\sum\limits_{n\in\mathbb{N}} v_n$ diverge.*
- *Si $u_n \sim v_n$ alors les deux séries $\sum\limits_{n\in\mathbb{N}} u_n$ et $\sum\limits_{n\in\mathbb{N}} v_n$ sont de même nature.*

Preuve : Si $u_n = \mathrm{O}(v_n)$, alors il existe un réel $\mathrm{M} > 0$ tel qu'à partir d'un certain rang on ait $u_n \leqslant \mathrm{M}v_n$, les deux premiers points en découlent compte tenu du théorème précédent.

Si $u_n \sim v_n$ alors on a à la fois $u_n = \mathrm{O}(v_n)$ **et** $v_n = \mathrm{O}(u_n)$. Le résultat en découle. □

Remarque 21.3 – *Les deux premiers points du théorème s'appliquent aussi lorsque $u_n = o(v_n)$.*

☞**Exemples** :

- Soit $u_n = \sin(\frac{1}{n})$ alors $u_n \sim \frac{1}{n}$, on en déduit que la série de terme général u_n est à termes positifs pour n assez grand et qu'elle est divergente, car la série harmonique diverge.
- Soit $u_n = 1 - \cos(\frac{1}{n})$, alors $u_n \sim \frac{1}{2n^2}$, on en déduit que la série de terme général u_n est à termes positifs pour n assez grand et qu'elle est convergente, car la série de terme général $\frac{1}{n^2}$ converge.

3) Comparaison avec une intégrale

Cette comparaison peut se faire pour les séries dont le terme général est de la forme $u_n = f(n)$ où f est une fonction continue monotone et à valeurs positives sur un intervalle $\mathrm{I} = [\mathrm{A}; +\infty[$.

Théorème 21.9

*Soit $f\colon [0;+\infty[\to \mathbb{R}$ une fonction **continue, monotone** et à **valeurs positives**.*

- *Si f est croissante alors $\forall n \in \mathbb{N}$, $f(0) + \int_0^n f(t)\,dt \leqslant \sum\limits_{k=0}^{n} f(k) \leqslant \int_0^{n+1} f(t)\,dt$.*
- *Si f est décroissante alors $\forall n \in \mathbb{N}$, $\int_0^{n+1} f(t)\,dt \leqslant \sum\limits_{k=0}^{n} f(k) \leqslant f(0) + \int_0^n f(t)\,dt$.*

Preuve : Supposons f croissante, alors sur l'intervalle $[k; k+1]$ on a $f(k) \leqslant f(t) \leqslant f(k+1)$, on en déduit en intégrant sur cet intervalle que $f(k) \leqslant \int_k^{k+1} f(t)\,dt \leqslant f(k+1)$, on somme alors ces inégalités pour k allant de 0 à n, ce qui donne $\sum\limits_{k=0}^{n} f(k) \leqslant \int_0^{n+1} f(t)\,dt \leqslant \sum\limits_{k=1}^{n+1} f(k)$, ce qui donne le premier encadrement. L'autre cas se traite de la même façon. □

☞**Exemple** : On considère la série de terme général $u_n = \frac{1}{n\ln(n)}$ pour $n \geqslant 2$. Soit $f(t) = \frac{1}{t\ln(t)}$ sur $[2;+\infty[$, cette fonction est continue, positive et décroissante, on en déduit que $\int_2^{n+1} f(t)\,dt \leqslant \sum\limits_{k=2}^{n} \frac{1}{k\ln(k)}) \leqslant f(2) + \int_2^n f(t)\,dt$, or $\int_2^n f(t)\,dt = [\ln(\ln(t))]_2^n = \ln(\ln(n)) - \ln(\ln(2) \to +\infty$, la série est donc divergente.

Théorème 21.10 (convergence des séries de Riemann)

Soit $\alpha \in \mathbb{R}$, la série de terme général $\frac{1}{n^\alpha}$ est convergente si et seulement si $\alpha > 1$.

Preuve : Si $\alpha \leqslant 0$ alors le terme général ne tend pas vers 0, la série est donc grossièrement divergente.

Supposons $\alpha > 0$ et soit $f(t) = \frac{1}{t^\alpha}$, c'est une fonction continue, positive et décroissante sur $[1;+\infty[$, on en déduit que que $\int_1^{n+1} f(t)\,dt \leqslant \sum\limits_{k=1}^{n} \frac{1}{k^\alpha} \leqslant f(1) + \int_1^n f(t)\,dt$. On sait que $\int_1^n \frac{dt}{t^\alpha} = \begin{cases} \ln(n) & \text{si } \alpha = 1 \\ \frac{n^{1-\alpha}-1}{1-\alpha} & \text{sinon} \end{cases}$. Si $\alpha \leqslant 1$ alors $\int_1^n f(t)\,dt \to +\infty$ et donc les sommes partielles tendent vers $+\infty$. Par contre, si $\alpha > 1$ alors $(\int_1^n f(t)dt)$ est convergente, donc majorée, et donc la série (qui est à termes positifs) converge. □

☞**Exemples** :

- Soit $u_n = e^{-\sqrt{n}}$, on sait que $n^2 = o\left(e^{\sqrt{n}}\right)$, on en déduit que $n^2 u_n = o(1)$ et donc $u_n = o\left(\frac{1}{n^2}\right)$, or la série de Riemann de terme général $\frac{1}{n^2}$ converge, donc la série $\sum\limits_{n\in\mathbb{N}} e^{-\sqrt{n}}$ est convergente.
- Soit $u_n = \frac{1}{\sqrt{n}\ln(n)}$, $n^\alpha u_n = \frac{n^{\alpha-1/2}}{\ln(n)} \to +\infty$ si on prend $\alpha > \frac{1}{2}$ (théorème des croissances comparées), donc pour n assez grand, on a $\frac{1}{n^\alpha} \leqslant u_n$, en prenant $\alpha \in]\frac{1}{2}; 1[$ on peut en déduire que série de terme général u_n est divergente.

– Soit $u_n = \frac{\ln^2(n)}{n^{3/2}}$, $n^\alpha u_n = n^{\alpha-3/2}\ln(n)^2 \to 0$ si on prend $\alpha < \frac{3}{2}$ (théorème des croissances comparées), donc pour n assez grand, on a $u_n \leqslant \frac{1}{n^\alpha}$, en prenant $\alpha \in]1; \frac{3}{2}[$ on peut en déduire que série de terme général u_n est convergente.

III SÉRIES À TERMES QUELCONQUES

Les séries considérées sont à termes complexes.

1) Séries absolument convergentes

Définition 21.4

*La série $\sum\limits_{n\in\mathbb{N}} u_n$ est dite **absolument convergente** lorsque la série $\sum\limits_{n\in\mathbb{N}} |u_n|$ est convergente.*

Remarque 21.4 – *La série $\sum\limits_{n\in\mathbb{N}} |u_n|$ est évidemment à termes positifs.*

☞**Exemple** : La série de terme général $u_n = \frac{e^{in}}{n^3}$ est absolument convergente car $|u_n| = \frac{1}{n^3}$: série de Riemann convergente.

Théorème 21.11

Si la série $\sum\limits_{n\in\mathbb{N}} u_n$ est absolument convergente, alors est elle convergente et on a $\left|\sum\limits_{n=0}^{+\infty} u_n\right| \leqslant \sum\limits_{n=0}^{+\infty} |u_n|$.
La réciproque est fausse.

Preuve : Posons $u_n = a_n + ib_n$ avec a_n et b_n réels. Posons $a_n^+ = \max(0, a_n)$ et $a_n^- = \max(0, -a_n)$ (idem pour b_n), alors on a $0 \leqslant a_n^+$, $0 \leqslant a_n^-$, $a_n^+ + a_n^- = |a_n|$ et $a_n^+ - a_n^- = a_n$. On en déduit que $0 \leqslant a_n^+ \leqslant |a_n| \leqslant |u_n|$ et $0 \leqslant a_n^- \leqslant |a_n| \leqslant |u_n|$. On en déduit que les séries à termes positifs a_n^+ et a_n^- sont convergentes, et par conséquent la série de terme général $a_n^+ - a_n^- = a_n$ est convergente. De même on montre que la série de terme général b_n converge et donc la série de terme général $a_n + ib_n = u_n$ est convergente.

L'inégalité triangulaire donne $\left|\sum\limits_{k=0}^{n} u_k\right| \leqslant \sum\limits_{k=0}^{n} |u_k|$, par passage à la limite, on obtient l'inégalité annoncée car on sait que les deux sommes ont une limite. □

Attention !

La réciproque du théorème ci-dessus est fausse, par exemple on montrera que la série de terme général $u_n = \frac{(-1)^n}{n}$ est convergente, mais on sait que la série de terme général $|u_n| = \frac{1}{n}$ est divergente. Une telle série est dite semi-convergente.

2) Séries alternées

Définition 21.5

Une série alternée est une série dont le terme général est de la forme $(-1)^n a_n$ où (a_n) est une suite réelle positive.

Théorème 21.12 **(critère spécial à certaines séries alternées)**

*Une série alternée $\sum\limits_{n\in\mathbb{N}} (-1)^n a_n$ telle que la suite (a_n) est **décroissante de limite nulle**, est convergente. De plus on a la majoration du reste d'ordre n : $|R_n| \leqslant a_{n+1}$.*

Preuve : Soit $S_n = \sum\limits_{k=0}^{n} (-1)^k a_k$ les sommes partielles, posons $u_n = S_{2n}$ et $v_n = S_{2n+1}$, alors $u_{n+1} - u_n = S_{2n+2} - S_{2n} = a_{2n+2} - a_{2n+1} \leqslant 0$ car la suite (a_k) est décroissante, on en déduit que la suite u est décroissante. De même, $v_{n+1} - v_n = S_{2n+3} - S_{2n+1} = a_{2n+2} - a_{2n+3} \geqslant 0$, la suite v est donc croissante. Or $u_n - v_n = a_{2n+1} \to 0$, les deux suites sont donc adjacentes, elles convergent vers une même limite ℓ, on a donc $S_{2n} \to \ell$ et $S_{2n+1} \to \ell$, on en déduit que $S_n \to \ell$, la série est donc bien convergente de somme ℓ.

De plus on a l'encadrement $v_n \leqslant \ell \leqslant u_n$, ce qui veut dire que ℓ est compris entre S_n et S_{n+1} pour tout n, par conséquent : $|R_n| = |\ell - S_n| \leqslant |S_{n+1} - S_n| = a_{n+1}$. □

Remarque 21.5 – *Lorsque le critère s'applique, la majoration du reste donne un renseignement sur la vitesse de convergence. Cette majoration est également utile lorsqu'on veut faire un calcul approché de la somme ℓ avec une précision donnée.*

☞**Exemple** : La série de terme général $\frac{(-1)^n}{n}$ est convergente, soit S sa somme, on a $|S - \sum\limits_{k=1}^{n} \frac{(-1)^k}{k}| \leqslant \frac{1}{n+1}$, à priori la convergence est lente. Nous verrons en exercice que $S = -\ln(2)$. Cette série n'est pas absolument convergente, mais seulement semi-convergente.

★**Exercice 21.2** ***(utilisation d'un développement asymptotique)***

1/ Nature de la série de terme général $u_n = \sin(\frac{(-1)^n}{\sqrt{n}})$.

2/ Même chose avec $u_n = \ln(1 + \frac{(-1)^n}{\sqrt{n}})$.

Attention !

Sur le deuxième exemple on a $\ln(1 + \frac{(-1)^n}{\sqrt{n}}) \sim \frac{(-1)^n}{\sqrt{n}}$, la série de gauche est divergente alors que la série de droite est convergente. Le théorème de comparaison est donc en défaut lorsque les séries ne sont pas à termes de signe constant.

3) Séries obtenues par une formule de Taylor

Théorème 21.13 (formule de Taylor avec reste intégrale)

Soit $f : \mathrm{I} \to \mathbb{C}$ une fonction de classe $\mathcal{C}^{n+1}$ sur l'intervalle I, alors on a la formule suivante :

$$\forall\, a, x \in \mathrm{I},\, f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + \int_a^x \frac{(x-t)^n}{n!} f^{(n+1)}(t)\, dt.$$

Preuve : Celle-ci est laissée en exercice, il s'agit d'une simple récurrence sur n. On en déduit le théorème suivant. □

Théorème 21.14 (inégalité de Taylor-Lagrange [1])

Si $f : \mathrm{I} \to \mathbb{C}$ est de classe $\mathcal{C}^{n+1}$ sur I et si $|f^{(n+1)}|$ est majorée par un réel M_{n+1}, alors :

$$\forall\, x, a \in \mathrm{I},\, \left|f(x) - \mathrm{T}_{n,f,a}(x)\right| \leqslant \mathrm{M}_{n+1} \frac{|x-a|^{n+1}}{(n+1)!}.$$

Applications :

- **La série exponentielle** : soit $z \in \mathbb{C}$, pour $t \in [0;1]$ on pose $f(t) = e^{tz}$, cette fonction est de classe $\mathcal{C}^\infty$ sur $[0;1]$, pour $n \in \mathbb{N}$, on a $f^{(n+1)}(t) = z^{n+1}e^{tz}$, en posant $z = a + ib$, on a $|f^{(n+1)}(t)| \leqslant |z|^{n+1}e^{ta} \leqslant |z|^{n+1}\mathrm{M}$ où M désigne le maximum de la fonction e^{ta} sur $[0;1]$, appliquons l'inégalité de Taylor-Lagrange entre 0 et 1 à l'ordre n :
$$\left|e^z - \sum_{k=0}^{n} \frac{z^k}{k!}\right| \leqslant \mathrm{M}\frac{|z|^{n+1}}{(n+1)!}.$$
On voit que le majorant tend vers 0 lorsque $n \to +\infty$, car $|z|^n = o(n!)$, par conséquent :
$$\forall\, z \in \mathbb{C}, e^z = \lim_{n\to+\infty} \sum_{k=0}^{n} \frac{z^k}{k!} \text{ ou encore } \boxed{\forall z \in \mathbb{C},\ e^z = \sum_{n=0}^{+\infty} \frac{z^n}{n!}}.$$
- **Développement de** sin **en série** : avec la fonction sin, toutes ses dérivées sont majorées par 1, on a donc pour $x \in \mathbb{R}$ et $n \in \mathbb{N}$, en appliquant l'inégalité de Taylor-Lagrange entre 0 et x à l'ordre $2n+1$:
$$\left|\sin(x) - \sum_{k=0}^{n} (-1)^k \frac{x^{2k+1}}{(2k+1)!}\right| \leqslant \frac{|x|^{2n+2}}{(2n+2)!}.$$
Là encore, on voit que le majorant tend vers 0 lorsque $n \to +\infty$ on en déduit donc que :
$$\forall\, x \in \mathbb{R}, \sin(x) = \lim_{n\to+\infty} \sum_{k=0}^{n} (-1)^k \frac{x^{2k+1}}{(2k+1)!} \text{ ou encore } \boxed{\forall x \in \mathbb{R},\ \sin(x) = \sum_{n=0}^{+\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}}.$$

1. *LAGRANGE Joseph Louis* (1736 – 1813) : mathématicien qui fut un précurseur dans de nombreux domaines scientifiques.

– **Développement de** $\cos$ **en série** : de la même façon on montre que

$$\forall\, x \in \mathbb{R}, \cos(x) = \lim_{n\to+\infty} \sum_{k=0}^{n} (-1)^k \frac{x^{2k}}{(2k)!} \text{ ou encore } \boxed{\forall x \in \mathbb{R},\ \cos(x) = \sum_{n=0}^{+\infty} (-1)^n \frac{x^{2n}}{(2n)!}}.$$

Remarque 21.6 –

– *On peut aussi déduire le développement en série de* $\sin$ *et* $\cos$ *en prenant les parties imaginaire et réelle du développement en série de* e^{ix}.

– *Ceci se généralise au cas où* f *est* $\mathcal{C}^\infty$ *et que ses dérivées en module sont* ***toutes majorées par une même constante****, car on sait que* $|x-a|^n = o(n!)$.

IV SOLUTION DES EXERCICES

Solution 21.1 *On vérifie par récurrence que* $S_{2n} = 1$ *et* $S_{2n+1} = 0$*, la suite* (S_n) *est donc divergente et par conséquent la série de terme général* u_n *aussi.*

Solution 21.2

1/ *On a* $\sin(\frac{(-1)^n}{\sqrt{n}}) = \frac{(-1)^n}{\sqrt{n}} - \frac{(-1)^n}{6n^{3/2}} + o\left(\frac{1}{n^{3/2}}\right)$*, on reconnaît trois séries convergentes : la première par le critère spécial des séries alternées, la deuxième est absolument convergente car en valeur absolue c'est une série de Riemann, et la troisième par comparaison avec une série de Riemann également. On en déduit que la série* $\sum\limits_{n\in\mathbb{N}} u_n$ *converge.*

2/ *On a pour* $n > 1$, $\ln(1 + \frac{(-1)^n}{\sqrt{n}}) = \frac{(-1)^n}{\sqrt{n}} - \frac{1}{2n} - \frac{(-1)^n}{3n^{3/2}} + o\left(\frac{1}{n^{3/2}}\right)$*, on reconnaît trois séries convergentes : la première par le critère spécial des séries alternées, la troisième est absolument convergente car en valeur absolue c'est une série de Riemann, la quatrième par comparaison avec une série de Riemann également, mais la deuxième est une série de Riemann divergente. On en déduit que la série* $\sum\limits_{n\in\mathbb{N}} u_n$ *diverge.*

Chapitre 22

La dimension finie

Sommaire

I ESPACES DE DIMENSION FINIE

1) Familles génératrices

Définition 22.1
Soit E *un* $\mathbb{K}$*-e.v, et soit* A *une famille de vecteurs de* E*, on dit que la famille* A *est une famille génératrice de* E *lorsque* $E = \text{Vect}[A]$*. Ce qui signifie que tout vecteur de* E *est combinaison linéaire d'un nombre fini de vecteurs de* A*.*

☞**Exemples :**

- Soit $E = \mathbb{K}^n$, pour $i \in [\![1\,;n]\!]$, on pose $e_i = (\delta_{i,1}, \ldots, \delta_{i,n})$, alors la famille $A = (e_1, \ldots, e_n)$ est une famille génératrice de E, car $(x_1, \ldots, x_n) = \sum\limits_{i=1}^{n} x_i e_i$.
- Soit $E = \mathbb{K}_n[X]$, la famille $A = (X^i)_{0 \leqslant i \leqslant n}$ est une famille génératrice de E.
- Soit $E = \mathbb{K}[X]$, la famille $A = (X^i)_{i \in \mathbb{N}}$ est une famille génératrice de E.
- Soit $E = \mathcal{M}_{n,p}(\mathbb{K})$, la famille des matrices élémentaires de taille (n,p), $A = (E^{i,j})_{(i,j) \in [\![1;n]\!] \times [\![1;p]\!]}$ est une famille génératrice de E.
- Soit $E = \mathcal{C}^0(\mathbb{R}, \mathbb{R})$, la famille $A = (f_k)_{0 \leqslant k \leqslant n}$ où $f_k : x \mapsto e^{kx}$, n'est pas une famille génératrice de E. Si elle l'était, il existerait des réels a_k tels que $\sin = \sum\limits_{k=0}^{n} a_k f_k$, en posant $P = \sum\limits_{k=0}^{n} a_k X^k$, on aurait alors pour tout réel x, $\sin(x) = P(e^x)$, la fonction sin s'annulant une infinité de fois et la fonction exp étant injective, on voit que P possède une infinité de racines, donc $P = 0$ *i.e.* tous les réels a_k sont nuls et donc la fonction sin est nulle, ce qui est absurde.

Théorème 22.1 (premières propriétés)
Soit A une famille génératrice de E :
- *Toute sur-famille de A est génératrice, i.e. si B est une famille de vecteurs de E telle que $A \subset B$, alors B est génératrice.*
- *Si $f \in \mathcal{L}(E,F)$, alors $B = (f(x))_{x\in A}$ est une famille génératrice de $\mathrm{Im}(f)$.*
- *Soit $f \in \mathcal{L}(E,F)$, alors f est surjective si et seulement si $f(A)$est une famille génératrice de F.*

Preuve : Le premier point est évident. Si $f \in \mathcal{L}(E,F)$, soit $y \in \mathrm{Im}(f)$, alors il existe $x \in E$ tel que $f(x) = y$, mais A est génératrice de E, donc il existe $x_1,\dots,x_n \in A$ et des scalaires $\alpha_1,\dots,\alpha_n$ tels que $x = \sum\limits_{i=1}^{n} \alpha_i x_i$, f étant linéaire, on a alors $f(x) = \sum\limits_{i=1}^{n} \alpha_i f(x_i)$, ce qui prouve que B est génératrice de $\mathrm{Im}(f)$. Le troisième point en découle. □

★**Exercice 22.1** *Les familles suivantes sont-elles génératrices dans $\mathbb{K}^3$?*
- $A = \{(1,1,1);(-1,0,2);(1,2,3)\}$.
- $A = \{(1,1,1),(1,2,3)\}$.

Définition 22.2 (espace de dimension finie)
*Soit E un $\mathbb{K}$-e.v, on dit que E est de dimension finie lorsque E possède une famille génératrice **finie**, c'est à dire lorsqu'il existe des vecteurs $x_1,\dots,x_n \in E$ tels que $E = \mathrm{Vect}\,[x_1,\dots,x_n]$. Si ce n'est pas le cas, on dit que E est de dimension infinie.*

☞**Exemples** :
- $\mathbb{K}^n$ est de dimension finie.
- $\mathbb{K}_n[X]$ est de dimension finie.
- $\mathcal{M}_{n,p}(\mathbb{K})$ est de dimension finie.
- $\mathbb{K}[X]$ est de dimension infinie. En effet : sinon, il existerait des polynômes $P_1,\dots,P_n$ tels que $\mathbb{K}[X] = \mathrm{Vect}\,[P_1,\dots,P_n]$, donc $\forall P \in \mathbb{K}[X]$, $\deg(P) \leqslant \max(\deg(P_1),\dots,\deg(P_n))$, ce qui est absurde. Nous verrons plus loin que $\mathcal{F}(\mathbb{R},\mathbb{R})$ est également de dimension infinie.

Théorème 22.2
Soit E un $\mathbb{K}$-e.v de dimension finie :
- *Si $f \in \mathcal{L}(E,F)$, alors $\mathrm{Im}(f)$ est de dimension finie. En particulier lorsque f est **surjective**, F est de dimension finie.*
- *Si F est de dimension finie, alors $E \times F$ est de dimension finie également.*

Preuve : Le premier point découle directement du théorème précédent. Si $(x_1,\dots,x_n)$ est une famille génératrice de E et si $(y_1,\dots,y_p)$ est une famille génératrice de F, alors pour $(x,y) \in E\times F$, on a :

$$(x,y) = \Big(\sum_{i=1}^{n} \alpha_i x_i, \sum_{k=1}^{p} \beta_k y_k\Big) = \sum_{1\leqslant i\leqslant n} \alpha_i(x_i,0_F) + \sum_{1\leqslant k\leqslant p} \beta_k(0_E,y_k),$$

ce qui prouve que la famille $((x_i,0_F);(0_E,y_k))_{1\leqslant i\leqslant n, 1\leqslant k\leqslant n}$ est une famille génératrice finie de $E\times F$. □

2) Familles libres, familles liées

Définition 22.3
Soit $(x_i)_{i\in I}$ une famille de vecteurs d'un $\mathbb{K}$-e.v E, on dit que :
- *La famille est libre lorsque la seule combinaison linéaire de la famille qui donne le vecteur nul est celle pour laquelle tous les coefficients sont nuls (on dit aussi que les vecteurs sont linéairement indépendants), c'est à dire :*
$$\sum_{i\in I} \alpha_i x_i = 0_E \implies \forall\, i \in I, \alpha_i = 0.$$
- *La famille est dite liée lorsqu'elle n'est pas libre (on dit aussi que les vecteurs sont linéairement dépendants), ce qui signifie qu'il existe une famille de scalaires à support fini $(\alpha_i)_{i\in I}$ **non tous nuls** tels que :*

$$\sum_{i\in I} \alpha_i x_i = 0_E.$$

Théorème 22.3 (caractérisation des familles liées)
La famille $(x_i)_{i\in I}$ est liée si et seulement si un des vecteurs de la famille est combinaison linéaire des autres.

Preuve : Supposons $x_k = \sum_{i\in I\setminus\{k\}} \alpha_i x_i$, on a alors $x_k - \sum_{i\in I\setminus\{k\}} \alpha_i x_i = 0_E$ avec le coefficient de x_k qui est non nul, donc la famille est liée.

Réciproquement : si la famille est liée, il existe une famille à support fini de scalaires $(\lambda_i)_{i\in I}$ non tous nuls tels que $\sum_{i\in I} \lambda_i x_i = 0_E$, supposons que $\lambda_{i_0} \neq 0$, on a alors $x_{i_0} = \sum_{i\in I\setminus\{i_0\}} \frac{-\lambda_i}{\lambda_{i_0}} x_i$. □

Remarque 22.1 – *Soient $x, y \in E$, alors la famille (x, y) est liée si et seulement si les deux vecteurs sont colinéaires.*

☞**Exemples** :
- Dans $E = \mathbb{K}^n$, la famille $A = (e_1, \dots, e_n)$ avec $e_i = (\delta_{i,1}, \dots, \delta_{i,n})$ est une famille libre.
- Dans $E = \mathbb{K}_n[X]$, la famille $A = (X^i)_{0\leqslant i\leqslant n}$ est une famille libre.
- Dans $E = \mathbb{K}[X]$, la famille $A = (X^i)_{i\in\mathbb{N}}$ est une famille libre.
- Dans $E = \mathcal{M}_{n,p}(\mathbb{K})$, la famille des matrices élémentaires $(E^{i,j})_{(i,j)\in[\![1;n]\!]\times[\![1;p]\!]}$ est une famille libre.
- Dans $E = \mathcal{C}^0([a;b], \mathbb{C})$, pour $k \in [\![0\,;n]\!]$ on pose $f_k : x \mapsto e^{kx}$, alors la famille $A = (f_k)_{k\in\mathbb{N}}$ est libre.
- Dans $E = \mathbb{K}^2$, soit $i = (1,1)$, $j = (1,2)$ et $k = (4,3)$, la famille $A = (i, j, k)$ est une famille liée.

★**Exercice 22.2** *Soit $E = \mathcal{C}^0(\mathbb{R}, \mathbb{R})$ pour $k \in [\![1\,;n]\!]$ on pose $f_k : x \mapsto \sin(kx)$. Montrer par récurrence sur n que la famille $A = (f_k)_{k\geqslant 1}$ est libre.*

Théorème 22.4 (propriétés des familles libres et des familles liées)
Soit E un $\mathbb{K}$-e.v
- *Si $x \in E$, alors la famille (x) est libre si et seulement si $x \neq 0_E$.*
- *Toute famille contenant le vecteur nul est liée.*
- *Toute partie d'une famille libre est une famille libre.*
- *Toute famille contenant une famille liée est une famille liée.*
- *L'image d'une famille liée par une application linéaire est une famille liée.*
- *Soit $f \in \mathcal{L}(E, F)$, alors f est injective si et seulement si f transforme toute famille libre de E en une famille libre de F.*
- *La famille $(x_i)_{i\in I}$ est libre si et seulement si $\forall\, x \in \mathrm{Vect}\left[(x_i)_{i\in I}\right]$, x s'écrit **de manière unique** comme combinaison linéaire des vecteurs $(x_i)_{i\in I}$.*
- *Soit $(x_i)_{i\in I}$ une famille libre, alors $\forall\, x \in E, x \in \mathrm{Vect}\left[(x_i)_{i\in I}\right] \iff \{x, (x_i)_{i\in I}\}$ est liée.*

Preuve : Démontrons les trois derniers points, il suffit de les démontrer pour toute famille finie :

Si f est injective et si $(x_1, \dots, x_n)$ est libre dans E, supposons que $\sum_{k=1}^n \alpha_k f(x_k) = 0_F$, alors $\sum_{k=1}^n \alpha_k x_k \in \ker(f)$, donc $\sum_{k=1}^n \alpha_k x_k = 0_E$, mais cette famille étant libre, les coefficients α_k sont tous nuls.

Si f transforme une famille libre en une famille libre : si $x \in \ker(f)$ et si $x \neq 0_E$, alors (x) est libre, donc $(f(x))$ est libre *i.e.* $f(x) \neq 0_F$ ce qui est absurde, donc $x = 0_E$ et f est injective.

Si $(x_1, \dots, x_n)$ est libre et si $x = \sum_{k=1}^n \alpha_k x_k = \sum_{k=1}^n \beta_k x_k$, alors $\sum_{k=1}^n (\alpha_k - \beta_k) x_k = 0_E$ et donc, la famille étant libre, $\alpha_k = \beta_k$ pour $k \in [\![1\,;n]\!]$.

Réciproquement, si l'écriture est unique alors $\sum_{k=1}^n \alpha_k x_k = 0_E \implies \forall\, k \in [\![1\,;n]\!], \alpha_k = 0$, *i.e.* la famille est libre.

Si $x \in \mathrm{Vect}\left[x_1, \dots, x_n\right]$, alors $x = \sum_{k=1}^n \alpha_k x_k$ ce qui prouve que la famille $(x_1, \dots, x_n, x)$ est liée.

Réciproque : si $(x_1, \dots, x_n, x)$ est liée, alors il existe des scalaires α_k pour $k \in [\![1\,;n+1]\!]$ tels que $\sum_{k=1}^n \alpha_k x_k + \alpha_{n+1} x = 0_E$, si $\alpha_{n+1} = 0$, alors il reste $\sum_{k=1}^n \alpha_k x_k = 0_E$ donc tous les scalaires α_k sont nuls, ce qui est contradictoire, d'où $\alpha_{n+1} \neq 0$, ce qui entraîne $x = -\sum_{k=1}^n \frac{\alpha_k}{\alpha_{n+1}} x_k$. □

3) Familles libres et familles liées en dimension finie

Théorème 22.5 (fondamental de la dimension finie)

Si E est de dimension finie et si $(x_1,\ldots,x_n)$ est une famille génératrice de E, alors toute famille de cardinal supérieur ou égal à $n+1$ est une famille liée.

Preuve : Par récurrence sur n : pour $n=1$ on a $\mathrm{E}=\mathrm{Vect}\left[x_1\right]$ soit $x,y\in\mathrm{E}$, $x=\alpha x_1$ et $y=\beta x_1$. Si $\alpha=0$ alors $x=0_{\mathrm{E}}$ et la famille (x,y) est liée, sinon on a $y=\frac{\beta}{\alpha}x$ et donc la famille (x,y) est liée.

Supposons le théorème vrai au rang n et soit $(x_1,\ldots,x_{n+1})$ une famille génératrice de E, soit $(y_1,\ldots,y_{n+2})$ une famille de vecteurs de E, on pose $\mathrm{F}=\mathrm{Vect}\left[x_1,\ldots,x_n\right]$. Si tous les vecteurs y_i sont dans F alors on peut appliquer l'hypothèse de récurrence et la famille $(y_1,\ldots,y_{n+2})$ est liée, dans le cas contraire, supposons $y_{n+2}\notin\mathrm{F}$, pour tout $j\in[\![1\,;n+2]\!]$ il existe un scalaire λ_j et un vecteur u_j de F tels que $y_j=u_j+\lambda_j x_{n+1}$, on a donc $\lambda_{n+2}\neq 0$, on en déduit x_{n+1} en fonction de u_{n+2} et y_{n+2}, d'où pour $j\in[\![1\,;n+1]\!]$: $y_j=u_j-\frac{\lambda_j}{\lambda_{n+2}}u_{n+2}+\frac{\lambda_j}{\lambda_{n+2}}y_{n+2}$, mais la famille $(u_j-\frac{\lambda_j}{\lambda_{n+2}}u_{n+2})_{j\in[\![1;n+1]\!]}$ est liée dans F (hypothèse de récurrence), donc la famille $(y_j-\frac{\lambda_j}{\lambda_{n+2}}y_{n+2})_{j\in[\![1;n+1]\!]}$ est liée dans E, ce qui entraîne que la famille $(y_1,\ldots,y_{n+2})$ est liée dans E. □

☞**Exemples** :

- $\mathrm{E}=\mathbb{K}^n$, pour $i\in[\![1\,;n]\!]$, on pose $e_i=(\delta_{i,1},\ldots,\delta_{i,n})$, on a vu en exemple que la famille $\mathrm{A}=(e_1,\ldots,e_n)$ est génératrice, donc dans $\mathbb{K}^n$, toute famille de $n+1$ vecteurs (ou plus) est liée.
- $\mathrm{E}=\mathbb{K}_n[\mathrm{X}]$, on sait que la famille $\mathrm{A}=(\mathrm{X}^0,\ldots,\mathrm{X}^n)$ est génératrice, donc dans $\mathbb{K}_n[\mathrm{X}]$, toute famille de $n+2$ vecteurs (ou plus) est liée.
- Dans $\mathrm{E}=\mathcal{M}_{n,p}(\mathbb{K})$, la famille des matrices élémentaires $(\mathrm{E}^{i,j})_{(i,j)\in[\![1;n]\!]\times[\![1;p]\!]}$ est une famille génératrice de cardinal np, donc dans $\mathcal{M}_{n,p}(\mathbb{K})$, toute famille de $np+1$ matrices (ou plus) est liée.
- $\mathrm{E}=\mathcal{F}(\mathbb{R},\mathbb{C})$ est de dimension infinie, sinon il existe une famille génératrice finie $\mathrm{A}=(f_1,\ldots,f_n)$, mais alors toute famille de $n+1$ vecteurs est liée, or la famille $(g_0,\ldots,g_n)$ où $g_k\colon x\mapsto e^{kx}$, est une famille libre, on a donc une contradiction, ce qui prouve que E est de dimension infinie.

Théorème 22.6

Si E est de dimension finie et si $(x_1,\ldots,x_n)$ est une famille libre, alors toute famille génératrice de E contient au moins n vecteurs.

Preuve : Soit $(y_1,\ldots,y_m)$ une famille génératrice de E, si $n>m$, alors d'après le théorème fondamental, la famille $(x_1,\ldots,x_n)$ est liée, ce qui est contradictoire, donc $m\geqslant n$. □

II PROPRIÉTÉS DE LA DIMENSION FINIE

1) Bases, coordonnées

Définition 22.4

Soit E un $\mathbb{K}$-e.v et soit $\mathrm{B}=(e_1,\ldots,e_n)$ une famille de vecteurs de E, on dit que B est une base de E lorsque B est à la fois ***libre et génératrice*** *de E. Si c'est le cas, alors $\forall\ x\in\mathrm{E}, \exists!(\lambda_1,\ldots,\lambda_n)\in\mathbb{K}^n$, tel que $x=\sum\limits_{k=1}^{n}\lambda_k x_k$. Par définition $(\lambda_1,\ldots,\lambda_n)$ sont* ***les coordonnées*** *de x dans la base B, et on pose $\mathrm{Coord}_{\mathrm{B}}(x)=(\lambda_1,\ldots,\lambda_n)\in\mathbb{K}^n$ (on fera attention à l'ordre sur les vecteurs de la base B, λ_i est la coordonnée de x sur le vecteur e_i).*

☞**Exemples** :

- Dans $\mathrm{E}=\mathbb{K}^n$, la famille $\mathrm{B}=(e_1,\ldots,e_n)$ où $e_i=(\delta_{i,1},\ldots,\delta_{i,n})$ pour $i\in[\![1\,;n]\!]$, est une famille libre et génératrice de $\mathbb{K}^n$, c'est donc une base, on l'appelle **base canonique** de $\mathbb{K}^n$. Si $u=(\lambda_1,\ldots,\lambda_n)\in\mathbb{K}^n$, alors on sait que $u=\sum\limits_{k=1}^{n}\lambda_k e_k$, donc $\mathrm{Coord}_{\mathrm{B}}(u)=(\lambda_1,\ldots,\lambda_n)$, c'est à dire : dans la base canonique de $\mathbb{K}^n$ les coordonnées d'un vecteur u sont les composantes de u.
- Dans $\mathrm{E}=\mathbb{K}_n[\mathrm{X}]$, on sait que la famille $\mathrm{B}=(\mathrm{X}^0,\ldots,\mathrm{X}^n)$ est libre et génératrice de $\mathbb{K}_n[\mathrm{X}]$, c'est donc une base de $\mathbb{K}_n[\mathrm{X}]$, on l'appelle **base canonique** de $\mathbb{K}_n[\mathrm{X}]$. Si $\mathrm{P}=\sum\limits_{k=0}^{n}\lambda_k\mathrm{X}^k\in\mathbb{K}_n[\mathrm{X}]$, alors $\mathrm{Coord}_{\mathrm{B}}(\mathrm{P})=(\lambda_0,\ldots,\lambda_n)$, c'est à dire : dans la base canonique de $\mathbb{K}_n[\mathrm{X}]$, les coordonnées d'un polynôme P sont les coefficients de P.

- Dans $E = \mathbb{K}[X]$, on sait que la famille $B = (X^k)_{k\in\mathbb{N}}$ est libre et génératrice de $\mathbb{K}[X]$, c'est donc une base de $\mathbb{K}_n[X]$, on l'appelle **base canonique** de $\mathbb{K}[X]$. Si $P = \sum\limits_{k\in\mathbb{N}} \lambda_k X^k \in \mathbb{K}[X]$, alors $\text{Coord}_B(P) = (\lambda_k)_{k\in\mathbb{N}}$, c'est à dire : dans la base canonique de $\mathbb{K}[X]$, les coordonnées d'un polynôme P sont les coefficients de P.
- Dans $E = \mathcal{M}_{n,p}(\mathbb{K})$, la famille des matrices élémentaires $B = (E^{i,j})_{(i,j)\in[\![1;n]\!]\times[\![1;p]\!]}$ est une famille libre et génératrice, donc c'est une base de $\mathcal{M}_{n,p}(\mathbb{K})$, on l'appelle **base canonique** de $\mathcal{M}_{n,p}(\mathbb{K})$. Si $A \in \mathcal{M}_{n,p}(\mathbb{K})$ alors $A = \sum\limits_{(i,j)\in[\![1;n]\!]\times[\![1;p]\!]} a_{i,j}E^{i,j}$, donc $\text{Coord}_B(A) = (a_{i,j})_{(i,j)\in[\![1;n]\!]\times[\![1;p]\!]}$, c'est à dire : dans la base canonique, les coordonnées d'une matrice sont ses coefficients.
- Dans $\mathbb{K}_n[X]$, soit $a \in \mathbb{K}$, d'après la formule de *Taylor* en a, la famille $B = ((X-a)^k)_{k\in[\![0;n]\!]}$ est une famille libre et génératrice de $\mathbb{K}_n[X]$, c'est donc une base. Pour $P \in \mathbb{K}_n[X]$, on sait que $P = \sum\limits_{k=0}^{n} \frac{P^{(k)}(a)}{k!}(X-a)^k$, donc $\text{Coord}_B(P) = (P(a), P'(a), \ldots, \frac{P^{(n)}(a)}{n!})$.

★**Exercice 22.3** $E = \mathbb{K}^2$, *on pose* $i = (1,1)$ *et* $j = (1,2)$, *montrer que* $B = (i,j)$ *est une base de* E, *et pour* $u = (x,y) \in E$, *calculer* $\text{Coord}_B(u)$.

Théorème 22.7 (existence de bases)

Tout $\mathbb{K}$-e.v non réduit au vecteur nul et de dimension finie, possède des bases. Plus précisément, toute famille génératrice contient une base.

Preuve : Soit $(x_1, \ldots, x_n)$ une famille génératrice de E, si tous ces vecteurs sont nuls, alors $E = \{0_E\}$ ce qui est exclu, donc au moins un des vecteurs de la famille est non nul, on peut donc considérer les sous-familles de $(x_1, \ldots, x_n)$ qui sont libres et en prendre une de **cardinal maximal** : $(x_{i_1}, \ldots, x_{i_r})$. Posons $y_j = x_{i_j}$ pour $j \in [\![1;r]\!]$, lorsque $k \notin \{i_1, \ldots, i_r\}$ on a $(y_1, \ldots, y_r, x_k)$ est liée (maximalité de r), donc $x_k \in \text{Vect}[y_1, \ldots, y_r]$, mais alors $\text{Vect}[x_1, \ldots, x_n] \subset \text{Vect}[y_1, \ldots, y_r]$, ce qui entraîne que $(y_1, \ldots, y_r)$ est une famille génératrice de E, comme elle est libre, c'est une base de E. □

Théorème 22.8 (propriété fondamentale des bases en dimension finie)

Si E *est un $\mathbb{K}$-e.v de dimension finie, alors toutes les bases de* E *ont le même cardinal. Ce cardinal est appelé* ***dimension de*** E *et noté* $\dim(E)$ *ou* $\dim_{\mathbb{K}}(E)$. *Par convention* $\{0_E\}$ *est de dimension* 0.

Preuve : Soit $B = (e_1, \ldots, e_n)$ une base de E, cette famille est libre, donc toute famille génératrice possède au minimum n vecteurs, mais cette famille est également génératrice, donc toute famille libre contient au maximum n vecteurs, finalement, toute base de E, étant à la fois libre et génératrice, contient exactement n vecteurs. □

☞**Exemples** :
- La base canonique de $\mathbb{K}^n$ contient n vecteurs, donc $\dim_{\mathbb{K}}(\mathbb{K}^n) = n$.
- La base canonique de $\mathbb{K}_n[X]$ contient $n+1$ vecteurs, donc $\dim_{\mathbb{K}}(\mathcal{M}_{n,p}(\mathbb{K})) = np$.
- La base canonique de $\mathcal{M}_{n,p}(\mathbb{K})$ contient np matrices, donc $\dim_{\mathbb{K}}(\mathbb{K}_n[X]) = n+1$.
- $\mathbb{C} = \mathbb{R}^2$ donc $\dim_{\mathbb{R}}(\mathbb{C}) = 2$, mais $\dim_{\mathbb{C}}(\mathbb{C}) = 1$.
- Une droite vectorielle est un espace de dimension 1 et un plan vectoriel est un espace de dimension 2.

Théorème 22.9 (application)

Si E *est de dimension* $n \geqslant 1$ *et si* B *est une famille de vecteurs de* E *alors les assertions suivantes sont équivalentes :*

a) B *est une base de* E.

b) B *est libre de cardinal* n *(on dit aussi libre maximale).*

c) B *est génératrice de cardinal* n *(on dit aussi génératrice minimale).*

Preuve : On a évidemment $a) \implies b)$ et $a) \implies c)$.

Montrons $b) \implies c)$: si B est libre de cardinal n : $B = (y_1, \ldots, y_n)$, alors pour tout $x \in E$, la famille $(y_1, \ldots, y_n, x)$ est liée (car de cardinal $n+1$ et E possède une base de cardinal n), ce qui entraîne que $x \in \text{Vect}[y_1, \ldots, y_n]$, ce qui prouve que B est génératrice.

Montrons $c) \implies a)$: si B est génératrice de cardinal n, on peut extraire de B une famille libre de cardinal maximal, cette famille sera alors une base de E, elle est donc de cardinal n, mais B est de cardinal n, donc cette famille extraite ne peut être que B elle-même, c'est à dire B est une base de E. □

★**Exercice 22.4** *Montrer que dans* $\mathbb{K}_n[X]$*, la famille* $B = (X^k(1-X)^{n-k})_{k\in[\![0;n]\!]}$ *est une base.*

On remarquera au passage la puissance du théorème précédent, car le fait que cette famille B soit génératrice n'est pas du tout évident à démontrer « à la main ».

Théorème 22.10

Soit $B = (e_1, \dots, e_n)$ *une base de* E*, l'application :*

$$\begin{array}{rccl} \mathrm{Coord}_B : & E & \to & \mathbb{K}^n \\ & x & \mapsto & \mathrm{Coord}_B(x) \end{array}$$

est un isomorphisme entre E *et* $\mathbb{K}^n$*.*

Preuve : En exercice. □

Remarque 22.2 – *Ce théorème permet de calculer facilement les coordonnées d'une combinaison linéaire de vecteurs connaissant les coordonnées de chacun d'eux.*

Théorème 22.11

Soient E *et* F *deux* $\mathbb{K}$*-e.v, soit* $B = (e_1, \dots, e_n)$ *une base de* E*, toute application linéaire de* E *vers* F *est entièrement déterminée par la donnée des images des vecteurs de la base* B*, plus précisément, si on se donne* $y_1, \dots, y_n \in F$*, alors il existe une unique application linéaire* $f : E \to F$ *telle que* $\forall\, i \in [\![1\,;n]\!], f(e_i) = y_i$*, de plus :*

– f *est injective si et seulement si* $(f(e_1), \dots, f(e_n))$ *est libre dans* F*.*
– f *est surjective si et seulement si* $(f(e_1), \dots, f(e_n))$ *est génératrice de* F*.*
– f *est bijective si et seulement si* $(f(e_1), \dots, f(e_n))$ *est une base de* F*.*

Preuve : Soit $x \in E$ et $(\lambda_1, \dots, \lambda_n) = \mathrm{Coord}_B(x)$, on définit $f : E \to F$ en posant $f(x) = \sum_{k=1}^{n} \lambda_k y_k$. Il est facile de vérifier que f est linéaire et que c'est la seule qui vérifie $f(e_k) = y_k$ pour $k \in [\![1\,;n]\!]$.

Les deux équivalences suivantes sont simples à démontrer et la troisième est la conjonction des deux précédentes. On retient la troisième en disant que f est un isomorphisme de E vers F si et seulement si f transforme une base de E en une base de F. □

À retenir

Il en découle que deux applications linéaires de E vers F qui coïncident sur une base de E sont égales.

Théorème 22.12 (de la base incomplète)

Soit $B = (e_1, \dots, e_n)$ *une base de* E *et soit* $(y_1, \dots, y_r)$ *une famille libre de* E *avec* $r < n$ *alors il existe des vecteurs* $e_{i_{r+1}}, \dots, e_{i_n}$ *de* B *tels que la famille* $(y_1, \dots, y_r, e_{i_{r+1}}, \dots, e_{i_n})$ *soit une base de* E*.*

Preuve : Comme $r < n$ la famille $(y_1, \dots, y_r)$ n'est pas génératrice de E, donc il existe au moins un vecteur e_k tel que $(y_1, \dots, y_r, e_k)$ soit libre. Posons $A = (y_1, \dots, y_r, e_1, \dots, e_n)$ alors A est une famille génératrice de E. La famille A contient des familles libres et parmi celles-ci il y en a qui contiennent les vecteurs $y_1, \dots, y_r$, on peut donc considérer une sous-famille libre de A qui contient ces vecteurs et qui soit de **cardinal maximal** : $C = (y_1, \dots, y_r, e_{i_{r+1}}, \dots, e_{i_s})$, si $k \notin \{i_{r+1}, \dots, i_s\}$ alors la famille $(y_1, \dots, y_r, e_{i_{r+1}}, \dots, e_{i_s}, e_k)$ est liée (maximalité de C), donc $e_k \in \mathrm{Vect}\,[C]$ ce qui entraîne que C est une famille génératrice de E, or C est libre donc C est une base de E, ce qui entraîne également que $s = n$ car toutes les bases de E ont n vecteurs. □

☞**Exemple** : Soit $E = \mathbb{K}^3$ et $B = (i, j, k)$ la base canonique de $\mathbb{K}^3$, soit $u = (1, 1, 1)$. La famille (u) est libre, on peut donc compléter cette famille en une base de E avec deux vecteurs de la base canonique. Les vecteurs i et u sont non colinéaires, donc (u, i) est libre. On a $j \notin \mathrm{Vect}\,[u, i]$ donc (u, i, j) est libre, c'est donc une base de $\mathbb{K}^3$. On remarquera que (u, i, k) et (u, j, k) conviennent également.

2) Sous-espaces vectoriels en dimension finie

Théorème 22.13
Si E *est de dimension finie, alors tout s.e.v* F *de* E *est de dimension finie et* $\dim(F) \leqslant \dim(E)$. *De plus, si* $\dim(F) = \dim(E)$ *alors* $F = E$.

Preuve : Si $\dim(E) = 0$ il n'y a rien à démontrer. On suppose donc $\dim(E) = n \geqslant 1$, si $F = \{0_E\}$ alors il n'y a rien à démontrer, on suppose donc $F \neq \{0_E\}$, mais alors F contient des familles libres et elles ont toute un cardinal inférieur ou égal à n, on peut donc considérer une famille $\mathcal{C}$ libre de F de **cardinal maximal**, on a alors $\forall x \in F, \mathcal{C} \cup (x)$ est liée, donc $x \in \text{Vect}[\mathcal{C}]$ ce qui prouve que $\mathcal{C}$ est une base de F, et le premier point est démontré. Si $\dim(F) = \dim(E)$, alors $\text{card}(\mathcal{C}) = n$, donc $\mathcal{C}$ est également une base de E, d'où $E = F$. □

Théorème 22.14
Si E *est de dimension finie, alors tout s.e.v* F *de* E *admet au moins un supplémentaire* G *dans* E.

Preuve : On écarte les trois cas triviaux : $E = \{0_E\}, F = E, F = \{0_E\}$. Dans le cas général, soit $(y_1, \ldots, y_p)$ une base de F alors on peut compléter cette famille en une base de E : $(y_1, \ldots, y_p, e_{p+1}, \ldots, e_n)$, on pose alors $G = \text{Vect}\left[e_{p+1}, \ldots, e_n\right]$ et on vérifie que G est un supplémentaire de F dans E. □

Théorème 22.15
Soient F *et* G *deux s.e.v d'un* $\mathbb{K}$*-e.v* E *de dimension finie, les assertions suivantes sont équivalentes :*

i) F *et* G *sont supplémentaires.*

ii) La somme $F + G$ *est directe et* $\dim(F) + \dim(G) = \dim(E)$.

iii) $E = F + G$ *et* $\dim(F) + \dim(G) = \dim(E)$.

Preuve : Il est clair que *i*) $\implies$ *ii*).

Supposons *ii*) : soit $B = (y_1, \ldots, y_p)$ une base de F, alors $\dim(G) = n - p$ (où $n = \dim(E)$), soit $B' = (y_{p+1}, \ldots, y_n)$ une base de G, il est facile de vérifier que $F \cap G = \{0_E\} \implies B \cup B'$ est libre, or $\text{card}(B \cup B') = n$, c'est donc une famille génératrice de E, ce qui entraîne $E = F + G$, donc *ii*) $\implies$ *iii*).

Supposons *iii*) : avec les notations précédentes, l'hypothèse $E = F + G$ implique que $B \cup B'$ est génératrice de E, mais elle est de cardinal n, c'est donc une famille libre ce qui entraîne ue la somme $F + G$ est directe et donc $E = F \oplus G$, donc *iii*) $\implies$ *i*). □

Théorème 22.16
Si E *est de dimension finie et si* F *et* G *sont deux s.e.v de* E*, alors (formule de Grassmann* [1]*) :*

$$\dim(F + G) = \dim(F) + \dim(G) - \dim(F \cap G).$$

Preuve : Soit H un supplémentaire de $F \cap G$ dans F : $F = H \oplus (F \cap G)$, on a donc d'après la propriété précédente, $\dim(H) = \dim(F) - \dim(F \cap G)$. On vérifie ensuite que $F + G = H \oplus G$, on en déduit alors que $\dim(F + G) = \dim(H) + \dim(G)$, ce qui donne la formule. □

Théorème 22.17
Si E *et* F *sont deux* $\mathbb{K}$*-e.v de dimension finie, alors* $E \times F$ *est de dimension finie et :*

$$\dim(E \times F) = \dim(E) + \dim(F).$$

Preuve : Soit $(e_1, \ldots, e_n)$ une base de E et $(y_1, \ldots, y_p)$ une base de F, alors on a déjà montré que la famille :

$B = ((e_1, 0_F), \ldots, (e_n, 0_F), (0_E, y_1), \ldots, (0_E, y_p))$ est un famille génératrice de $E \times F$, on peut vérifier alors que c'est aussi une famille libre, ce qui donne le résultat. □

☞**Exemple** : Si $\dim(E) = 3$, décrire les s.e.v de E et étudier les positions relatives (intersection de deux s.e.v de E).

3) Applications linéaires et dimension finie

1. *Hermann Günter GRASSMANN* (1809 – 1877) : mathématicien allemand qui a développé les notions fondamentales de l'algèbre linéaire.

Théorème 22.18 (dimension de $\mathcal{L}(E,F)$)
Si $\dim(E) = n$ *et* $\dim(F) = p$*, alors* $\mathcal{L}(E,F)$ *est de dimension finie et* $\dim(\mathcal{L}(E,F)) = np$*. En particulier,* $\dim(\mathcal{L}(E)) = n^2$.

Preuve : Ce théorème sera établi dans le chapître sur les matrices. □

☞**Exemples** :
- $\mathcal{L}(\mathbb{K}^n, \mathbb{K}^p)$ est de dimension np.
- Si $\dim(E) = n$, alors $\dim(E^*) = n$ (dual de E).

Théorème 22.19 (comparaison des dimensions)
Si E *et* F *sont de dimension finie et si* $f \in \mathcal{L}(E,F)$ *:*
- f *injective* $\implies \dim(E) \leqslant \dim(F)$.
- f *surjective* $\implies \dim(E) \geqslant \dim(F)$.
- f *bijective (isomorphisme)* $\implies \dim(E) = \dim(F)$.

Preuve : Soit $B = (e_1, \dots, e_n)$ une base de E et $B' = (y_1, \dots, y_p)$ une base de F.
Si f est injective, alors la famille $(f(e_1), \dots, f(e_n))$ est libre dans F donc $n \leqslant p = \dim(F)$.
Si f est surjective, alors la famille $(f(e_1), \dots, f(e_n))$ est génératrice de F donc $n \geqslant p = \dim(F)$.
Si f est un isomorphisme, alors la famille $(f(e_1), \dots, f(e_n))$ est une base de F donc $n = p = \dim(F)$. □

Remarque 22.3 –
- *Tout* $\mathbb{K}$*-e.v de dimension* n *est isomorphe à* $\mathbb{K}^n$.
- *Si* E *est de dimension finie, alors* E *est isomorphe à* E^**, son dual.*

Théorème 22.20 (caractérisations des isomorphismes)
Soient E *et* F *deux* $\mathbb{K}$*-e.v de dimension finie tels que* $\dim(E) = \dim(F)$*, et soit* $f \in \mathcal{L}(E,F)$*, les assertions suivantes sont équivalentes :*

a) f *est injective.*

b) f *est surjective.*

c) f *est bijective.*

d) $\exists\, g \in \mathcal{L}(F,E), f \circ g = \mathrm{id}_F$.

e) $\exists\, h \in \mathcal{L}(F,E), h \circ f = \mathrm{id}_E$.

Preuve : Soit $B = (e_1, \dots, e_n)$ une base de E.
a) $\implies$ *b)* : si f est injective alors $(f(e_1), \dots, f(e_n))$ est libre de cardinal n dans F, or $\dim(F) = \dim(E) = n$, c'est une famille génératrice de F, et donc f est surjective.
b) $\implies$ *c)* : f est surjective, donc la famille $(f(e_1), \dots, f(e_n))$ est génératrice de F de cardinal $\dim(F)$, c'est donc une base de F, ce qui signifie que f est un isomorphisme.
c) $\implies$ *d)* : f est bijective, il suffit donc de prendre $g = f^{-1}$.
d) $\implies$ *e)* : on a $f \circ g = \mathrm{id}_F$ donc f est surjective, ce qui entraîne que f est bijective, on a donc $g = f^{-1}$ et il suffit alors de prendre $h = g$.
e) $\implies$ *a)* : on a $h \circ f = \mathrm{id}_E$, donc f est injective. □

Remarque 22.4 –
- *Lorsque* $\dim(E) = \dim(F)$ *l'égalité* $f \circ g = \mathrm{id}_F$ *entraîne automatiquement* $g \circ f = \mathrm{id}_E$*, ce qui est tout à fait remarquable puisqu'en général la loi* $\circ$ *n'est pas commutative.*
- *Le théorème précédent est faux en dimension infinie comme le montre les exemples suivants :* $f : \mathbb{K}[X] \to \mathbb{K}[X]$ *définie par* $f(P) = P'$ *(surjective non injective) et* $g : \mathbb{K}[X] \to \mathbb{K}[X]$ *définie par* $g(P) = XP$ *(injective non surjective).*

☞**Exemples** :
- Polynômes d'endomorphismes : si $\dim(E) = n$ alors $\dim(\mathcal{L}(E)) = n^2$, donc si $u \in \mathcal{L}(E)$, alors la famille $(id, u, \dots, u^{n^2})$ est liée, donc il existe un polynôme $P = \sum \lambda_k X^k$ non nul de degré $\leqslant n^2$ tel que $P(u) = 0$. Si $P(0) \neq 0$, alors $\lambda_0 \neq 0$ d'où $\mathrm{id}_E = -\sum\limits_{k\geqslant 1} \frac{\lambda_k}{\lambda_0} u^k$, ou encore :

$$\mathrm{id}_E = u \circ \left[-\sum_{k\geqslant 1} \frac{\lambda_k}{\lambda_0} u^{k-1} \right].$$

Ce qui prouve que $u \in \mathrm{GL}(\mathrm{E})$ et $u^{-1} = -\sum\limits_{k\geqslant 1} \frac{\lambda_k}{\lambda_0} u^{k-1}$.

– Polynôme d'interpolation de Lagrange : soient $a_0, \dots, a_n$ $n+1$ scalaires distincts, on note $f\colon \mathbb{K}_n[\mathrm{X}] \to \mathbb{K}^{n+1}$ définie par $f(\mathrm{P}) = (\mathrm{P}(a_0), \dots, \mathrm{P}(a_n))$. Il est facile de vérifier que f est linéaire et injective, comme $\dim(\mathbb{K}_n[\mathrm{X}]) = n+1$, f est un **isomorphisme**, par conséquent pour $(\alpha_0, \dots, \alpha_n) \in \mathbb{K}^{n+1}$ il existe un unique polynôme $\mathrm{P} \in \mathbb{K}_n[\mathrm{X}]$ tel que $\forall\, i \in [\![0\,;n]\!], \mathrm{P}(a_i) = \alpha_i$. Soit $\mathrm{B} = (e_0, \dots, e_n)$ la base canonique de $\mathbb{K}^{n+1}$, on note L_i l'unique polynôme de degré $\leqslant n$ tel que $f(\mathrm{L}_i) = e_i$. On vérifie facilement que :

$$\mathrm{L}_i(\mathrm{X}) = \prod_{k=0, k\neq i}^{n} \frac{\mathrm{X} - a_k}{a_i - a_k}.$$

C'est le i^{e} polynôme d'interpolation de Lagrange aux points $(a_0, \dots, a_n)$, on en déduit que :

$$f^{-1}(\alpha_0, \dots, \alpha_n) = \sum_{k=0}^{n} \alpha_k \mathrm{L}_k(\mathrm{X}).$$

★**Exercice** 22.5 *Soit* E *de dimension finie, soit* $u \in \mathcal{L}(\mathrm{E})$, *et soit* F *un s.e.v de* E *stable par* u, *on appelle* v *la restriction de* u *à* F, *c'est à dire* $v\colon \mathrm{F} \to \mathrm{F}$ *définie par* $v(x) = u(x)$.

1/ *Montrer que si* $u \in \mathrm{GL}(\mathrm{E})$, *alors* $v \in \mathrm{GL}(\mathrm{F})$.

2/ *Montrer que si* $\ker(u) + \mathrm{F}$ *est directe, alors* $v \in \mathrm{GL}(\mathrm{F})$.

III NOTION DE RANG

1) Rang d'une application linéaire

Définition 22.5

Soit E *un espace vectoriel de dimension finie et* $f \in \mathcal{L}(\mathrm{E}, \mathrm{F})$, *on sait que* $\mathrm{Im}(f)$ *est de dimension finie, par définition, on appelle* ***rang*** *de* f *la dimension de* $\mathrm{Im}(f)$.
Notation : $\mathrm{rg}(f) = \dim(\mathrm{Im}(f))$.

Remarque 22.5 –

– *Si* $\mathrm{B} = (e_1, \dots, e_n)$ *est une base de* E *alors on sait que* $(f(e_1), \dots, f(e_n))$ *est une famille génératrice de* $\mathrm{Im}(f)$ *donc* $\dim(\mathrm{Im}(f)) \leqslant n$, *c'est à dire :*

$$\mathrm{rg}(f) \leqslant \dim(\mathrm{E}).$$

– *Si de plus* F *est de dimension finie, alors* $\dim(\mathrm{Im}(f)) \leqslant \dim(\mathrm{F})$, *donc :*

$$\mathrm{rg}(f) \leqslant \min(\dim(\mathrm{E}), \dim(\mathrm{F})).$$

– f *est surjective si et seulement si* $\mathrm{Im}(f) = \mathrm{F}$, *donc :*

$$f \text{ est surjective} \iff \mathrm{rg}(f) = \dim(\mathrm{F}).$$

Théorème 22.21 (théorème du rang)

Si E *et de dimension finie et si* $f \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ *alors* $\dim(\mathrm{E}) = \dim(\ker(f)) + \mathrm{rg}(f)$.

Preuve : Soit G un supplémentaire de $\ker(f)$ dans E : $\mathrm{E} = \ker(f) \oplus \mathrm{G}$. Il est facile de vérifier que l'application linéaire $v\colon \mathrm{G} \to \mathrm{Im}(f)$ définie par $v(x) = f(x)$ est un isomorphisme, ce qui entraîne que $\dim(\mathrm{G}) = \dim(\mathrm{Im}(f)) = \mathrm{rg}(f)$, or $\dim(\mathrm{E}) = \dim(\ker(f)) + \dim(\mathrm{G})$, d'où le résultat. □

Conséquences : Soit $f\colon \mathrm{E} \to \mathrm{F}$ linéaire,si E de dimension finie, alors f est injective $\iff \mathrm{rg}(f) = \dim(\mathrm{E})$.

☞**Exemples** :

– Soit $u\colon \mathbb{K}_{n+1}[\mathrm{X}] \to \mathbb{K}_n[\mathrm{X}]$ définie par $u(\mathrm{P}) = \mathrm{P}(\mathrm{X}+1) - \mathrm{P}(\mathrm{X})$. Il est clair que u est linéaire, on vérifie que $\ker(u) = \mathbb{K}$, le théorème du rang nous dit alors que $\mathrm{rg}(u) = \dim(\mathbb{K}_{n+1}[\mathrm{X}]) - \dim(\ker(u)) = n+2-1 = n+1 = \dim(\mathbb{K}_n[\mathrm{X}])$, ce qui prouve que u est surjective.

– Soient F et G deux s.e.v de E, on considère l'application $\varphi\colon F\times G \to E$ définie par $\varphi(x,y)=x+y$, on vérifie que φ est linéaire et que $\ker(\varphi)=\{(x,-x)\ /\ x\in F\cap G\}$. On voit que $\ker(\varphi)$ est isomorphe à $F\cap G$ et que $\mathrm{Im}(\varphi)=F+G$, le théorème du rang nous dit alors que $\dim(F\times G)=\dim(F\cap G)+\dim(F+G)$, d'où $\dim(F+G)=\dim(F)+\dim(G)-\dim(F\cap G)$, on retrouve ainsi la formule de *Grassmann*.

★**Exercice** 22.6 *Soit* $B=(i,j,k)$ *une base de* $\mathbb{K}^3$, *on considère* $u\colon \mathbb{K}^3\to\mathbb{K}^4$ *définie par* $\begin{cases} u(i)=(1,2,-1,0)\\ u(j)=(-1,0,1,-2)\\ u(k)=(1,4,-1,-2)\end{cases}$. *Calculer le rang de* u.

Théorème 22.22 (propriétés du rang)
Soient E, F, G *trois* $\mathbb{K}$*-e.v, soit* $f\in\mathcal{L}(E,F)$ *et* $g\in\mathcal{L}(F,G)$, *on a* $\mathrm{rg}(g\circ f)\leqslant\min(\mathrm{rg}(g),\mathrm{rg}(f))$.
De plus, si f *est surjective, alors* $\mathrm{rg}(g\circ f)=\mathrm{rg}(g)$, *et si* g *est injective, alors* $\mathrm{rg}(g\circ f)=\mathrm{rg}(f)$.

Preuve : Soit $h\colon \mathrm{Im}(f)\to G$ définie par $h(x)=g(x)$, alors h est linéaire et $\mathrm{Im}(h)=\mathrm{Im}(g\circ f)\subset\mathrm{Im}(g)$, on a donc $\mathrm{rg}(g\circ f)=\mathrm{rg}(h)\leqslant\dim(\mathrm{Im}(f))$ d'après le théorème du rang, de plus l'inclusion ci-dessus prouve que $\mathrm{rg}(h)\leqslant\mathrm{rg}(g)$, d'où le résultat.

Si f est surjective, alors $\mathrm{Im}(f)=F$ et donc $h=g$, d'où $\mathrm{rg}(g\circ f)=\mathrm{rg}(g)$.

Si g est injective, alors h est injective, donc $\mathrm{rg}(h)=\dim(\mathrm{Im}(f))=\mathrm{rg}(f)$. □

★**Exercice** 22.7 *Soient* E, F *deux* $\mathbb{K}$*-e.v de dimension finie et soit* $u,v\in\mathcal{L}(E,F)$, *montrer que :*

$$|\,\mathrm{rg}(u)-\mathrm{rg}(v)|\leqslant\mathrm{rg}(u+v)\leqslant\mathrm{rg}(u)+\mathrm{rg}(v).$$

2) Rang d'une famille de vecteurs

Définition 22.6
Soit $(x_1,\dots,x_n)$ *une famille de vecteurs d'un* $\mathbb{K}$*-.v* E. *On appelle rang de la famille* $(x_1,\dots,x_n)$ *la dimension du s.e.v engendré par cette famille.*
Notation : $\mathrm{rg}(x_1,\dots,x_n)=\dim(\mathrm{Vect}[x_1,\dots,x_n])$.

Remarque 22.6 –

– *La famille* $(x_1,\dots,x_n)$ *est génératrice de* $F=\mathrm{Vect}[x_1,\dots,x_n]$, *donc son rang est inférieur ou égal à* n. *De plus on sait que l'on peut extraire de cette famille une base de* F *en prenant une sous-famille libre de* ***cardinal maximal***, *ce cardinal maximal est donc le rang de la famille.*

– *La famille* $(x_1,\dots,x_n)$ *est libre si et seulement si c'est une base de* $F=\mathrm{Vect}[x_1,\dots,x_n]$ *ce qui équivaut à* $\dim(F)=n$, *ce qui équivaut encore à : le rang de la famille est égal à* n, *donc :*

$$(x_1,\dots,x_n)\ \textit{libre}\iff \mathrm{rg}(x_1,\dots,x_n)=n.$$

– *Soit* $B=(e_1,\dots,e_n)$ *une base de* E *et soit* $f\in\mathcal{L}(E,F)$ *alors :*

$$\mathrm{rg}(f)=\dim(\mathrm{Im}(f))=\dim(\mathrm{Vect}[f(e_1),\dots,f(e_n)])=\mathrm{rg}(f(e_1),\dots,f(e_n)).$$

☞**Exemples** :

– Dans $\mathbb{K}^3$ calculons le rang de la famille (i,j,k) avec $i=(1,1,1)$, $j=(-1,2,5)$ et $k=(3,0,-3)$. La famille (i) est libre, j est non colinéaire à i donc (i,j) est libre, mais $k=2i-j$ donc la famille (i,j,k) est liée, par conséquent (i,j) est libre maximale d'où $\mathrm{rg}(i,j,k)=2$ et $\mathrm{Vect}[i,j,k]=\mathrm{Vect}[i,j]$.

3) Propriétés du rang d'une famille de vecteurs

Soit E un $\mathbb{K}$-e.v de dimension finie, on considère une famille de vecteurs de E, $(x_1,\dots,x_p)$. On a les propriétés suivantes :

– Si $x_p=0_E$, alors $\mathrm{rg}(x_1,\dots,x_p)=\mathrm{rg}(x_1,\dots,x_{p-1})$. Autrement dit, si l'un des vecteurs de la famille est nul, on peut le retirer de la famille, cela ne change pas le rang (en fait cela ne change pas le s.e.v engendré).

– Si $\alpha \in \mathbb{K}^*$, alors $\text{rg}(x_1, \dots, x_p) = \text{rg}(\alpha x_1, x_2, \dots, x_p)$, autrement dit, multiplier l'un des vecteurs par un scalaire non nul ne change pas le rang.
– Si $\lambda_2, \dots, \lambda_p \in \mathbb{K}$, alors $\text{rg}(x_1, \dots, x_p) = \text{rg}(x_1 + \sum_{k=2}^{p} \lambda_k x_k, x_2, \dots, x_p)$. Autrement dit, ajouter à l'un des vecteurs une combinaison linéaire des autres ne change pas le rang.

À retenir

Pour déterminer le rang d'une famille de vecteurs, on étudie la liberté de la famille. Si elle est libre, son rang est son cardinal, si elle est liée, un des vecteurs est combinaison linéaire des autres, on le retire alors de la famille et on recommence.

IV COMPLÉMENTS : HYPERPLANS EN DIMENSION FINIE

Théorème 22.23

Soit E un espace de dimension n et H un s.e.v. de E, soit $B = (e_1, \dots, e_n)$ une base de E. Les assertions suivantes sont équivalentes :

a) H *est un hyperplan de* E.

b) Il existe des scalaires $(a_1, \dots, a_n)$ ***non tous nuls*** *tels que :*

$$H = \{x\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in E \ / \ a_1x_1 + \dots + a_nx_n = 0\}$$ *(on a une équation cartésienne de* H*).*

c) $\dim(H) = n - 1$.

Preuve : Montrons $i \implies ii$: H est le noyau d'une forme linéaire f sur E, non nulle. Posons $f(e_i) = a_i$, alors au moins un des a_i est non nul. Soit x de coordonnées $(x_1, \dots, x_n)$ dans la base B, alors $f(x) = a_1 f(e_1) + \dots + a_n f(e_n) = \sum_{i=1}^{n} a_i x_i$. Le résultat en découle.

Montrons que $ii \implies iii$: supposons $a_1 \neq 0$, quitte à tout diviser par a_1 on peut suppose $a_1 = 1$, l'équation de H équivaut alors à $x_1 = -a_2x_2 - \dots - a_nx_n$, on en déduit que $x \in H \iff x = x_2[-a_2e_1 + e_2] + \dots x_n[-a_ne_1 + e_n]$, par conséquent $H = \text{Vect}[-a_2e_1 + e_2; \dots; -a_ne_1 + e_n]$, on a une famille génératrice de H, on vérifie qu'elle est libre, c'est donc une base de H et donc $\dim(H) = n - 1$.

Montrons que $iii \implies i$: soit $(h_1, \dots, h_{n-1})$ une base de H, c'est une famille libre, on peut la compléter en une base de E : $(h_1, \dots, h_{n-1}, h_n)$, soit f la forme linéaire sur E définie par $f(h_i) = 0$ si $1 \leqslant i \leqslant n-1$ et $f(h_n) = 1$. Alors f est non nulle, et on vérifie que $\ker(f) = H$, donc H est un hyperplan. □

Théorème 22.24 (égalité de deux hyperplans)

Soient $H_1 = \ker(f_1)$ et $H_2 = \ker(f_2)$ deux hyperplans de E, alors $H_1 = H_2 \iff (f_1, f_2)$ est liée.

Preuve : Si (f_1, f_2) est liée, il existe un scalaire a tel que $f_1 = af_2$ (car f_2 est non nulle), comme f_1 est non nulle, on a $a \neq 0$, par conséquent $\forall x \in E, f_1(x) = 0 \iff f_2(x) = 0$, on en déduit que $H_1 = H_2$.

Si $H_1 = H_2$, soit $(h_1, \dots, h_{n-1})$ une base de H_1, que l'on complète en une base de E : $(h_1, \dots, h_{n-1}, h_n)$, h_n étant ni dans H_1 ni dans H_2, on a $f_1(h_n)$ et $f_2(h_n)$ non nuls, il existe donc un scalaire a tel que $f_1(h_n) = af_2(h_n)$, on vérifie alors que pour i dans $[\![1\,;n]\!]$, on a $f_1(h_i) = af_2(h_i)$ et donc $f_1 = af_2$, la famille (f_1, f_2) est liée. □

★**Exercice 22.8**

1/ *Montrer que les deux équations $a_1x_1 + \dots + a_nx_n = 0$ et $b_1x_1 + \dots + b_nx_n = 0$ dans une base B sont équivalentes si et seulement si $(a_1, \dots, a_n)$ et $(b_1, \dots, b_n)$ sont colinéaires dans $\mathbb{K}^n$.*

2/ *Soient H_1 et H_2 deux hyperplans distincts, montrer que $\dim(H_1 \cap H_2) = n - 2$ si $n = \dim(E)$.*

Théorème 22.25

- *Si $H_1, \dots, H_p$ sont des hyperplans de E, alors $\dim(H_1 \cap \dots \cap H_p) \geqslant n - p$.*
- *Tout s.e.v. de E de dimension $n - p$ est l'intersection de p hyperplans de E.*

Preuve :
• On écrit $H_i = \ker(f_i)$ où les f_i sont des formes linéaires non nulles sur E. L'application $g\colon E \to \mathbb{K}^p$ définie par $g(x) = (f_1(x), \dots, f_p(x))$ est linéaire, son noyau est l'intersection $H_1 \cap \dots \cap H_p$ et son rang est inférieure ou égale à p. Le théorème du rang donne le résultat.

• Soit $(e_1, \dots, e_{n-p})$ une base de ce s.e.v., on complète en une base de E : $B = (e_1, \dots, e_n)$, soit f_i la forme linéaire définie par $f_i(e_j) = \delta_{ij}$ pour $i \in [\![n-p+1\,;n]\!]$. On vérifie que $\ker(f_i)$ est l'ensemble des vecteurs dont la coordonnée dans la base B sur e_i est nulle, on en déduit que $\mathrm{Vect}\left[e_1, \dots, e_{n-p}\right] = \ker(f_{n-p+1}) \cap \dots \cap \ker(f_n)$. □

V SOLUTION DES EXERCICES

Solution 22.1

1/ *Soit $(a, b, c) \in \mathbb{K}^3$, on résout le système $x(1,1,1)+y(-1,0,2)+z(1,2,3) = (a,b,c)$ ce qui donne* $\begin{cases} -y + x + z = a \\ x + 2z = b \\ -z = c + 2a - 3b \end{cases}$, *ce système a une (unique) solution (x, y, z), donc la famille A est génératrice de $\mathbb{K}^3$.*

2/ $\mathrm{Vect}\,[A]$ *est un plan vectoriel, c'est le plan d'équation $x - 2y + z = 0$, il ne contient pas le triplet $(-1, 0, 2)$ par exemple, donc A n'est pas une famille génératrice de $\mathbb{K}^3$.*

Solution 22.2 *Il est clair que (f_1) est libre, supposons que $(f_1, \dots, f_{n-1})$ soit libre et supposons que $\sum_{k=1}^{n} \alpha_k f_k = 0$, on a donc pour tout réel x de $[a; b]$, $\sum_{k=0}^{n} \alpha_k \sin(kx) = 0$, en dérivant deux fois cette relation, on obtient $\sum_{k=1}^{n} k^2 \alpha_k \sin(kx) = 0$, en lui retranchant n^2 fois la première, on obtient $\sum_{k=1}^{n-1} (k^2 - n^2)\alpha_k f_k = 0$, de l'hypothèse de récurrence on déduit que $\alpha_k = 0$ pour $k \in [\![1\,; n-1]\!]$, il reste alors $\alpha_n f_n = 0$ et donc $\alpha_n = 0$.*

Solution 22.3 *La famille $B = (i, j)$ est libre et pour tout $(a, b) \in \mathbb{K}^2$ on a $(a, b) = (2a - b)i + (b - a)j$, donc la famille est génératrice de $\mathbb{K}^2$, c'est donc une base de $\mathbb{K}^2$ et $\mathrm{Coord}_B(a, b) = (2a - b, b - a)$.*

Solution 22.4 *La base canonique de $\mathbb{K}_n[X]$ possède $n + 1$ vecteurs et $\mathrm{card}(B) = n + 1$, il suffit donc de démontrer que B est libre : si on a $\sum_{k=0}^{n} \alpha_k X^k (1 - X)^{n-k} = 0$, en évaluant en 1, il reste $\alpha_n = 0$, donc $\sum_{k=0}^{n-1} \alpha_k X^k (1 - X)^{n-k} = 0$, c'est à dire $\sum_{k=0}^{n-1} \alpha_k X^k (1 - X)^{n-1-k} = 0$, on peut terminer par récurrence sur n pour montrer que B est libre, la propriété étant évidente pour $n = 0$. B est donc une base de $\mathbb{K}_n[X]$.*

Solution 22.5

1/ *On a $v \in \mathcal{L}(F)$ et F est de dimension finie, il suffit donc de vérifier que v est injective, or $v(x) = 0 \implies u(x) = 0 \implies x = 0$ car u est injective ($u \in GL(F)$), donc v est bien injective ce qui entraîne ici que $v \in GL(F)$.*

2/ $x \in \ker(v) \iff x \in F$ *et* $x \in \ker(u)$, *donc $\ker(v) = F \cap \ker(u)$. Si $\ker(u) + F$ est directe, alors $F \cap \ker(u) = \{0\}$ et donc $\ker(v) = \{0\}$, v est donc injective, d'où $v \in GL(F)$ (car F est de dimension finie).*

Solution 22.6 $x = ai + bj + ck \in \ker(u) \iff \begin{cases} b = -c \\ a = -2c \end{cases}$, *on en déduit que* $\ker(u) = \mathrm{Vect}\,[2i + j - k]$, *d'où* $\mathrm{rg}(u) = \dim(\mathbb{K}^3) - 1 = 2$, *donc* $\mathrm{Im}(u)$ *est un plan vectoriel, comme $u(i)$ et $u(j)$ sont non colinéaires, on a* $\mathrm{Im}(u) = \mathrm{Vect}\,[u(i), u(j)]$.

Solution 22.7 $\mathrm{Im}(u + v) \subset \mathrm{Im}(u) + \mathrm{Im}(v)$ *donc en prenant les dimensions,* $\mathrm{rg}(u + v) \leqslant \mathrm{rg}(u) + \mathrm{rg}(v) - \dim(\mathrm{Im}(u) \cap \mathrm{Im}(v)) \leqslant \mathrm{rg}(u) + \mathrm{rg}(v)$.

En écrivant $u = (u + v) - v$ et en appliquant ce qui précède, on a $\mathrm{rg}(u) \leqslant \mathrm{rg}(u + v) + \mathrm{rg}(-v)$, or $\mathrm{rg}(-v) = \mathrm{rg}(v)$ (car $\mathrm{Im}(-v) = \mathrm{Im}(v)$), d'où $\mathrm{rg}(u) - \mathrm{rg}(v) \leqslant \mathrm{rg}(u + v)$. En échangeant les rôles de u et v on a $\mathrm{rg}(v) - \mathrm{rg}(u) \leqslant \mathrm{rg}(u + v)$ et donc $|\,\mathrm{rg}(u) - \mathrm{rg}(v)| \leqslant \mathrm{rg}(u + v)$.

Solution 22.8

1/ *C'est une application directe du théorème ci-dessus.*

2/ *Il existe au moins un vecteur de H_2 qui n'est pas dans H_1, donc $H_1 + H_2$ contient une base de E, par conséquent $H_1 + H_2 = E$, en appliquant la formule de Grassmann on obtient le résultat.*

Chapitre 23

Matrices et dimension finie

Sommaire

$\mathbb{K}$ désigne un sous-corps de $\mathbb{C}$.

I MATRICES, LIENS AVEC LES APPLICATIONS LINÉAIRES

1) Matrice d'une application linéaire

Soit E un $\mathbb{K}$-e.v de dimension p, soit $\mathcal{B} = (e_1, \dots, e_p)$ une base de E. Soit F un $\mathbb{K}$-e.v de dimension n et soit $\mathcal{B}' = (u_1, \dots, u_n)$ une base de F. Soit $f \in \mathcal{L}(\mathrm{E}, \mathrm{F})$, on sait que f est entièrement déterminée par la donnée de $f(e_1), \dots, f(e_p)$, mais chacun de ces vecteurs est lui-même déterminé par ses coordonnées dans la base $\mathcal{B}'$ de F. Notons $\underset{\mathcal{B}'}{\mathrm{coord}}(f(e_j)) = (a_{1,j}, \dots, a_{n,j})$ pour $j \in [\![1\,;p]\!]$, c'est à dire :

$$\forall\, j \in [\![1\,;p]\!], f(e_j) = \sum_{i=1}^{n} a_{i,j} u_i.$$

On obtient ainsi une matrice $\mathrm{A} = (a_{i,j})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant p}}$ cette matrice est définie par : $c_j(\mathrm{A}) = \underset{\mathcal{B}'}{\mathrm{coord}}(f(e_j))$.

Définition 23.1

Soit $f \in \mathcal{L}(\mathrm{E}, \mathrm{F})$, soit $\mathcal{B} = (e_1, \dots, e_p)$ une base de E et soit $\mathcal{B}' = (u_1, \dots, u_n)$ une base de F, on appelle matrice de f relative aux bases $\mathcal{B}$ et $\mathcal{B}'$ la matrice de $\mathcal{M}_{n,p}(\mathbb{K})$ notée $\underset{\mathcal{B},\mathcal{B}'}{\mathrm{mat}}(f)$ et définie par : pour $j \in [\![1\,;p]\!]$, le j-ième vecteur colonne de cette matrice est $\underset{\mathcal{B}'}{\mathrm{coord}}(f(e_j))$, autrement dit, le coefficient de la ligne i colonne j est la coordonnée sur u_i du vecteur $f(e_j)$.

Construction de cette matrice :

$$\begin{array}{cccccc} & f(e_1) & \dots & f(e_p) & & \\ & \downarrow & & \downarrow & & \\ \underset{\mathcal{B},\mathcal{B}'}{\mathrm{mat}}(f) = & \left(\begin{matrix} a_{1,1} \\ \vdots \\ a_{n,1} \end{matrix}\right. & \begin{matrix} \cdots \\ \vdots \\ \cdots \end{matrix} & \left.\begin{matrix} a_{1,p} \\ \vdots \\ a_{n,p} \end{matrix}\right) & \begin{matrix} \rightarrow \\ \\ \rightarrow \end{matrix} & \begin{matrix} u_1 \\ \vdots \\ u_n \end{matrix} \end{array}$$

☞**Exemples :**

– Soit $\mathcal{B} = (e_1, e_2, e_3)$ la base canonique de $\mathbb{K}^3$ et soit $\mathcal{B}' = (u, v)$ la base canonique de $\mathbb{K}^2$, soit $f \in \mathcal{L}(\mathbb{K}^3, \mathbb{K}^2)$ définie par $\forall\ (x, y, z) \in \mathbb{K}^3, f(x, y, z) = (2x - y + z, x + 2y - 3z)$. Déterminons $A = \underset{\mathcal{B},\mathcal{B}'}{\mathrm{mat}}(f)$, on a $\begin{cases} f(e_1) = f(1,0,0) = (2,1) = 2u + v \\ f(e_2) = f(0,1,0) = (-1,2) = -u + v \\ f(e_3) = f(0,0,1) = (1,-3) = u - 3v \end{cases}$, donc la matrice de f est :

$$A = \begin{pmatrix} 2 & -1 & 1 \\ 1 & 2 & -3 \end{pmatrix}.$$

– Avec les notations précédentes, déterminons l'application linéaire $g : \mathbb{K}^3 \to \mathbb{K}^2$ donnée par :

$$\underset{\mathcal{B},\mathcal{B}'}{\mathrm{mat}}(g) = \begin{pmatrix} 6 & -2 & 1 \\ 4 & 5 & -1 \end{pmatrix}.$$

On a $g(x, y, z) = xg(e_1) + yg(e_2) + zg(e_3) = x(6,4) + y(-2,5) + z(1,-1) = (6x - 2y + z, 4x + 5y - z)$.

Remarque 23.1 – *Cas particuliers des endomorphismes : lorsque l'espace d'arrivée est le même que celui de départ ($F = E$), on choisit en général la même base à l'arrivée qu'au départ ($\mathcal{B}' = \mathcal{B}$), on note alors $\underset{\mathcal{B},\mathcal{B}}{\mathrm{mat}}(f) = \underset{\mathcal{B}}{\mathrm{mat}}(f)$, c'est une matrice carrée.*

★**Exercice 23.1**

1/ *Soit $E = \mathbb{K}_3[X]$ et soit $\mathcal{B}$ la base canonique de E :*

a) *On note D la dérivation dans E, calculer $\underset{\mathcal{B}}{\mathrm{mat}}(D)$.*

b) *Soit Δ définie par $\Delta(P) = P(X+1) - P(X)$, calculer $\underset{\mathcal{B}}{\mathrm{mat}}(\Delta)$.*

c) *Soit $P_0 = 1, P_1 = X, P_2 = \frac{X(X-1)}{2}$ et $P_3 = \frac{X(X-1)(X-2)}{6}$, montrer que $\mathcal{B}' = (P_0, P_1, P_2, P_3)$ est une base de E et calculer $\underset{\mathcal{B}'}{\mathrm{mat}}(\Delta)$.*

2/ *Calculer la matrice de la transposition dans la base canonique de $\mathcal{M}_2(\mathbb{K})$.*

Théorème 23.1 (caractérisation de l'identité et de l'application nulle)

Soit E un e.v de dimension n et soit $\mathcal{B}$ une base de E :

- *Soit $f \in \mathcal{L}(E)$, alors $f = \mathrm{id}_E \iff \underset{\mathcal{B}}{\mathrm{mat}}(f) = I_n$.*
- *Soit F un e.v de dimension p et soit $\mathcal{B}'$ une base de F, soit $f \in \mathcal{L}(E, F)$, alors :*

$$f = 0 \iff \underset{\mathcal{B},\mathcal{B}'}{\mathrm{mat}}(f) = O_{p,n}.$$

Preuve : Celle-ci est simple et laissée en exercice. □

Théorème 23.2

Soient E, F deux $\mathbb{K}$-e.v, soit $\mathcal{B}$ une base de E et soit $\mathcal{B}'$ une base de F, soient $f, g \in \mathcal{L}(E, F)$ et soit $\lambda \in \mathbb{K}$. On a :

$$\underset{\mathcal{B},\mathcal{B}'}{\mathrm{mat}}(f+g) = \underset{\mathcal{B},\mathcal{B}'}{\mathrm{mat}}(f) + \underset{\mathcal{B},\mathcal{B}'}{\mathrm{mat}}(g) \quad et \quad \underset{\mathcal{B},\mathcal{B}'}{\mathrm{mat}}(\lambda.f) = \lambda.\underset{\mathcal{B},\mathcal{B}'}{\mathrm{mat}}(f).$$

Autrement dit, l'application : $\underset{\mathcal{B},\mathcal{B}'}{\mathrm{mat}} : \mathcal{L}(E, F) \to \mathcal{M}_{n,p}(\mathbb{K})$ (avec $n = \dim(F)$ et $p = \dim(E)$), est une application linéaire. Plus précisément, cette application est un isomorphisme.
On en déduit que $\dim(\mathcal{L}(E, F)) = \dim(\mathcal{M}_{n,p}(\mathbb{K})) = np = \dim(E) \times \dim(F)$.

Preuve : La vérification de la linéarité est simple (elle découle de la linéarité de l'application coordonnées). Si $\underset{\mathcal{B},\mathcal{B}'}{\mathrm{mat}}(f) = O_{n,p}$, alors on sait que f est nulle, donc l'application $\underset{\mathcal{B},\mathcal{B}'}{\mathrm{mat}}$ est injective, la surjectivité étant évidente, on a donc bien un isomorphisme. □

2) Produit matriciel

Matrice d'une composée : Soient $\mathfrak{B} = (e_1, \dots, e_q)$ une base de E, $\mathfrak{B}' = (u_1, \dots, u_p)$ une base de F, et $\mathfrak{B}'' = (v_1, \dots, v_n)$ une base de G, soit $f \in \mathcal{L}(E, F), g \in \mathcal{L}(F, G)$, on pose $B = \underset{\mathfrak{B},\mathfrak{B}'}{\text{mat}}(f) \in \mathcal{M}_{p,q}(\mathbb{K}), A = \underset{\mathfrak{B}',\mathfrak{B}''}{\text{mat}}(g) \in \mathcal{M}_{n,p}(\mathbb{K})$ et $C = \underset{\mathfrak{B},\mathfrak{B}''}{\text{mat}}(g \circ f) \in \mathcal{M}_{n,q}(\mathbb{K})$. Il s'agit de calculer $g \circ f(e_j)$ dans la base $\mathfrak{B}''$, on a : $f(e_j) = \sum\limits_{k=1}^{p} B_{k,j}u_k$, donc : $g \circ f(e_j) = \sum\limits_{k=1}^{p} B_{k,j}g(u_k) = \sum\limits_{k=1}^{p}\sum\limits_{i=1}^{n} B_{k,j}A_{i,k}v_i$, c'est à dire : $g \circ f(e_j) = \sum\limits_{i=1}^{n}\left(\sum\limits_{k=1}^{p} A_{i,k}B_{k,j}\right)v_i$. On doit donc avoir : $C \in \mathcal{M}_{n,q}(\mathbb{K})$ avec $C_{i,j} = \sum\limits_{k=1}^{p} A_{i,k}B_{k,j}$. On voit que l'opération à effectuer sur les matrices A et B pour obtenir C est exactement le **produit entre les deux matrices** A et B, *i.e.* $C = A \times B$.

Théorème 23.3

Soit $\mathfrak{B}$ une base de E, soit $\mathfrak{B}'$ une base de F et soit $\mathfrak{B}''$ une base de G, soit $f \in \mathcal{L}(E, F)$ et soit $g \in \mathcal{L}(F, G)$ avec $A = \underset{\mathfrak{B}',\mathfrak{B}''}{\text{mat}}(g)$ et $B = \underset{\mathfrak{B},\mathfrak{B}'}{\text{mat}}(f)$, alors :

$$\underset{\mathfrak{B},\mathfrak{B}''}{\text{mat}}(g \circ f) = A \times B = \underset{\mathfrak{B}',\mathfrak{B}''}{\text{mat}}(g) \times \underset{\mathfrak{B},\mathfrak{B}'}{\text{mat}}(f).$$

Remarque 23.2 – *Cas particulier des endomorphismes : Soit E un $\mathbb{K}$-e.v, soit $\mathfrak{B}$ une base de E et soient $u, v \in \mathcal{L}(E)$ avec $A = \underset{\mathfrak{B}}{\text{mat}}(u)$ et $B = \underset{\mathfrak{B}}{\text{mat}}(v)$, on a alors $\underset{\mathfrak{B}}{\text{mat}}(u \circ v) = \underset{\mathfrak{B}}{\text{mat}}(u) \times \underset{\mathfrak{B}}{\text{mat}}(v) = A \times B$, en particulier :*

$$\forall\, n \in \mathbb{N}, \underset{\mathfrak{B}}{\text{mat}}(u^n) = \left[\underset{\mathfrak{B}}{\text{mat}}(u)\right]^n = A^n.$$

Théorème 23.4 (relation fondamentale)

*Soit $\mathfrak{B} = (e_1, \dots, e_p)$ une base de E, soit $\mathfrak{B}' = (u_1, \dots, u_n)$ une base de F et soit $f \in \mathcal{L}(E, F)$. Pour $x \in E$, on pose X la **matrice colonne** des coordonnées de x dans la base $\mathfrak{B}$, ce que l'on note : $X = \text{Coord}_{\mathfrak{B}}(x) \in \mathcal{M}_{p,1}(\mathbb{K})$) et Y la matrice colonne des coordonnées de $y = f(x)$ dans la base $\mathfrak{B}'$: $Y = \text{Coord}_{\mathfrak{B}'}(f(x)) \in \mathcal{M}_{n,1}(\mathbb{K})$. En posant $A = \underset{\mathfrak{B},\mathfrak{B}'}{\text{mat}}(f)$, on a alors la relation suivante :*

$$Y = A \times X \quad i.e. \quad \text{Coord}_{\mathfrak{B}'}(f(x)) = \underset{\mathfrak{B},\mathfrak{B}'}{\text{mat}}(f) \times \text{Coord}_{\mathfrak{B}}(x).$$

Preuve : Posons $X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$ et $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$, comme $A \in \mathcal{M}_{n,p}(\mathbb{K})$ on voit que le produit $A \times X$ est bien défini et que c'est une matrice colonne à n lignes. On a $f(x) = \sum\limits_{k=1}^{p} x_k f(e_k)$, mais on a $f(e_k) = \sum\limits_{i=1}^{n} a_{i,k}u_i$, ce qui donne :

$$f(x) = \sum_{i=1}^{n}\left(\sum_{k=1}^{p} a_{i,k}x_{k,1}\right)u_i = \sum_{i=1}^{n}[A \times X]_{i,1}u_i = \sum_{i=1}^{n} y_i u_i.$$

Ce qui prouve que $Y = A \times X$. □

★Exercice 23.2

1/ Soient $\mathfrak{B}$ la base canonique de $\mathbb{K}^3$ et $\mathfrak{B}'$ la base canonique de $\mathbb{K}^2$, soit $f \in \mathcal{L}(\mathbb{K}^3, \mathbb{K}^2)$ définie par sa matrice dans les bases $\mathfrak{B}$ et $\mathfrak{B}'$: $\underset{\mathfrak{B},\mathfrak{B}'}{\text{mat}}(f) = A = \begin{pmatrix} 1 & -2 & 3 \\ 2 & 1 & -5 \end{pmatrix}$, calculer $f(x, y, z)$.

2/ Soit $\mathfrak{B} = (i, j, k)$ la base canonique de $\mathbb{K}^3$, on pose $\mathfrak{B}' = (i, i + j, i + j + k)$, on vérifie que $\mathfrak{B}'$ est une base de $\mathbb{K}^3$. Soit $f \in \mathcal{L}(\mathbb{K}^3)$ défini par $\underset{\mathfrak{B}'}{\text{mat}}(f) = A = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 2 & -1 \\ 0 & 0 & 1 \end{pmatrix}$, calculer $f(x, y, z)$.

3/ Soient $A, B \in \mathcal{M}_{n,p}(\mathbb{K})$ telles que $\forall\, X \in \mathcal{M}_{p,1}(\mathbb{K}), AX = BX$, montrer que $A = B$.

Définition 23.2 (application linéaire canoniquement associée)

*Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$, on appelle application linéaire canoniquement associée à A l'application linéaire $f_A \in \mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$ dont la matrice dans **les bases canoniques** de $\mathbb{K}^p$ et $\mathbb{K}^n$, est A.*

3) Caractérisation des isomorphismes

Théorème 23.5

Soient E *et* F *deux* $\mathbb{K}$*-e.v de même dimension* n*, soit* $\mathfrak{B}$ *une base de* E *et* $\mathfrak{B}'$ *une base de* F*, soit* $u \in \mathcal{L}(E,F)$*, alors* u *est un isomorphisme de* E *vers* F *si et seulement si* $\underset{\mathfrak{B},\mathfrak{B}'}{\mathrm{mat}}(u) \in \mathrm{GL}_n(\mathbb{K})$*, si c'est le cas, alors* $\underset{\mathfrak{B}',\mathfrak{B}}{\mathrm{mat}}(u^{-1}) = \left[\underset{\mathfrak{B},\mathfrak{B}'}{\mathrm{mat}}(u)\right]^{-1}$.

Preuve : Si u est un isomorphisme, posons $A = \underset{\mathfrak{B},\mathfrak{B}'}{\mathrm{mat}}(u)$ et $B = \underset{\mathfrak{B}',\mathfrak{B}}{\mathrm{mat}}(u^{-1})$, on a $A, B \in \mathcal{M}_n(\mathbb{K})$. On sait que $u \circ u^{-1} = \mathrm{id}_F$, d'où $I_n = \underset{\mathfrak{B}'}{\mathrm{mat}}(\mathrm{id}_F) = \underset{\mathfrak{B},\mathfrak{B}'}{\mathrm{mat}}(u) \times \underset{\mathfrak{B}',\mathfrak{B}}{\mathrm{mat}}(u^{-1}) = A \times B$, de même $B \times A = \underset{\mathfrak{B}}{\mathrm{mat}}(\mathrm{id}_E) = I_n$.

Si la matrice de u est inversible, soit $v \in \mathcal{L}(F,E)$ telle que $\underset{\mathfrak{B}',\mathfrak{B}}{\mathrm{mat}}(v) = A^{-1}$n alors en considérant la matrice de $v \circ u$ dans la base $\mathfrak{B}$, on vérifie que $v \circ u = \mathrm{id}_E$, donc u est un isomorphisme (théorème de la dimension finie). □

Cas des endomorphismes : Si E est un $\mathbb{K}$-espace vectoriel de dimension n et $\mathfrak{B}$ une base de E, alors on sait déjà que l'application $\underset{\mathfrak{B}}{\mathrm{mat}} : \mathcal{L}(E) \to \mathcal{M}_n(\mathbb{K})$ est un isomorphisme d'espaces vectoriels, mais comme $(\mathcal{L}(E), +, \circ, .)$ et $(\mathcal{M}_n(\mathbb{K}), +, \times; .)$ sont des $\mathbb{K}$-algèbres et que $\underset{\mathfrak{B}}{\mathrm{mat}}(u \circ v) = \underset{\mathfrak{B}}{\mathrm{mat}}(u) \times \underset{\mathfrak{B}}{\mathrm{mat}}(v)$ et $\underset{\mathfrak{B}}{\mathrm{mat}}(\mathrm{id}_E) = I_n$, on peut affirmer que l'application $\underset{\mathfrak{B}}{\mathrm{mat}}$ est **un isomorphisme d'algèbres**. En particulier celui-ci induit un isomorphisme de groupes : $\underset{\mathfrak{B}}{\mathrm{mat}} : \mathrm{GL}(E) \to \mathrm{GL}_n(\mathbb{K})$.

Théorème 23.6 (caractérisations des matrices carrées inversibles)

Soit $A \in \mathcal{M}_n(\mathbb{K})$*, alors les assertions suivantes sont équivalentes :*

a) A *est inversible.*

b) Il existe une matrice $B \in \mathcal{M}_n(\mathbb{K})$ *telle que* $BA = I_n$.

c) L'équation $AX = O_{n,1}$ *d'inconnue* $X \in \mathcal{M}_{n,1}(\mathbb{K})$ *admet une* ***unique solution*** $X = O_{n,1}$.

d) $\forall\, Y \in \mathcal{M}_{n,1}(\mathbb{K})$*, l'équation* $AX = Y$ *d'inconnue* $X \in \mathcal{M}_{n,1}(\mathbb{K})$ *admet une unique solution.*

e) $\forall\, Y \in \mathcal{M}_{n,1}(\mathbb{K})$*, l'équation* $AX = Y$ *d'inconnue* $X \in \mathcal{M}_{n,1}(\mathbb{K})$ *admet au moins une solution.*

Preuve : L'implication $i) \implies ii)$ est évidente en prenant $B = A^{-1}$.

Montrons $ii) \implies iii)$: On a $BA = I_n$, d'où $AX = O_{n,1} \implies BAX = O_{n,1} = X$.

Montrons $iii) \implies iv)$: Soit $f \in \mathcal{L}(\mathbb{K}^n)$ l'endomorphisme de $\mathbb{K}^n$ canoniquement associé à A, soit $x \in \ker(f)$, posons $X = \mathrm{Coord}_{\mathfrak{B}}(x)$ où $\mathfrak{B}$ désigne la base canonique de $\mathbb{K}^n$, on a alors $\mathrm{Coord}_{\mathfrak{B}}(f(x)) = AX = O_{n,1}$, donc $X = O_{n,1}$ *i.e.* $x = \vec{0}$, l'application f est donc injective, mais alors elle est bijective : $\forall\, y \in \mathbb{K}^n, \exists!\, x \in \mathbb{K}^n, f(x) = y$, ce qui entraîne $\forall\, Y \in \mathcal{M}_{n,1}(\mathbb{K}), \exists!\, X \in \mathcal{M}_{n,1}(\mathbb{K}), AX = Y$ (remarquons que A est inversible puisque f est bijective, et que $X = A^{-1}Y$).

L'implication $iv) \implies v)$ est évidente.

Montrons $v) \implies i)$: Avec les notations précédentes, l'application f est surjective par hypothèse, donc f est bijective et par conséquent sa matrice A est inversible. □

Remarque 23.3 – *Il découle en particulier de ce théorème que si* $BA = I_n$ *alors* $AB = I_n$ *(car* $A \in \mathrm{GL}_n(\mathbb{K})$ *et donc* $B = A^{-1}$*), ce qui est remarquable.*

II CHANGEMENT DE BASES

1) Matrice de passage

Définition 23.3

Soit E *un* $\mathbb{K}$*-espace vectoriel, soit* $\mathfrak{B} = (e_1, \dots, e_n)$ *une base de* E*, soit* $S = (x_1, \dots, x_p)$ *une famille de vecteurs de* E*, on appelle* ***matrice de la famille*** S ***dans la base*** $\mathfrak{B}$*, la matrice* $A \in \mathcal{M}_{n,p}(\mathbb{K})$ *définie par :* $\forall\, (i,j) \in [\![1\,;n]\!] \times [\![1\,;p]\!], a_{i,j}$ *est la coordonnée sur* e_i *de* x_j*. Autrement dit, pour* $j \in [\![1\,;p]\!]$*, la* j^e *colonne de* A *est* $\mathcal{C}_j(A) = \mathrm{Coord}_{\mathfrak{B}}(x_j)$*. Cette matrice est notée* $\mathcal{P}_{\mathfrak{B},S}$ *et appelée* *matrice de la famille S dans la base $\mathcal{B}$ à S, elle exprime les vecteurs de S dans la base $\mathcal{B}$:*

$$\begin{array}{ccccl} & x_1 & \cdots & x_p & \rightarrow \text{ vecteurs de S} \\ & \downarrow & \cdots & \downarrow & \\ \mathcal{P}_{\mathcal{B},S} = & \left(\begin{matrix} a_{1,1} & \cdots & a_{1,p} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,p}\end{matrix}\right) & & & \begin{matrix} \rightarrow \text{ coordonnée sur } e_1 \text{ (premier vecteur de } \mathcal{B}) \\ \vdots \\ \rightarrow \text{ coordonnée sur } e_n \text{ (dernier vecteur de } \mathcal{B}) \end{matrix} \end{array}$$

☞**Exemples** :

– Soit $\mathcal{B}$ la base canonique de $\mathbb{K}^3$, soit $x_1 = (1,-1,0)$ et $x_2 = (2,-1,3)$, alors la matrice de la famille $S = (x_1, x_2)$ dans la base $\mathcal{B}$, est $\mathcal{P}_{\mathcal{B},S} = \begin{pmatrix} 1 & 2 \\ -1 & -1 \\ 0 & 3 \end{pmatrix}$.

– Soit $\mathcal{B} = (i, j, k)$ la base canonique de $\mathbb{K}^3$, soit $\mathcal{B}' = (i, i+j, i+j+k)$, on vérifie que $\mathcal{B}'$ est une base de $\mathbb{K}^3$. Déterminons la matrice de la famille S précédent dans la base $\mathcal{B}'$: on a $x_1 = i - j = 2i - (i+j)$ et $x_2 = 2i - j + 3k = 3i - 4(i+j) + 3(i+j+k)$, on a donc $\mathcal{P}_{\mathcal{B}',S} = \begin{pmatrix} 2 & 3 \\ -1 & -4 \\ 0 & 3 \end{pmatrix}$.

Interprétations de la matrice de passage :

a) Dans le cas général : soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E et soit $S = (x_1, \dots, x_p)$ une famille de p vecteurs de E. Soit $\mathcal{B}' = (u_1, \dots, u_p)$ la base canonique de $\mathbb{K}^p$, on définit l'application linéaire $f : \mathbb{K}^p \to E$ en posant pour $i \in [\![1\,;p]\!]$, $f(u_i) = x_i$, alors : $\mathcal{P}_{\mathcal{B},S} = \underset{\mathcal{B}',\mathcal{B}}{\text{mat}}(f)$.

b) Dans le cas où $p = n$: on a $S = (x_1, \dots, x_n)$, soit $u \in \mathcal{L}(E)$ défini par $\forall\, i \in [\![1\,;n]\!]$, $u(e_i) = x_i$, on a alors : $\mathcal{P}_{\mathcal{B},S} = \underset{\mathcal{B}}{\text{mat}}(u)$.

Théorème 23.7 (caractérisation des bases)

Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E, et soit $\mathcal{B}' = (x_1, \dots, x_n)$ une famille de n vecteurs de E, alors $\mathcal{B}'$ est une base de E si et seulement si la matrice de la famille $\mathcal{B}'$ dans la base $\mathcal{B}$ est inversible ,i.e. $\mathcal{P}_{\mathcal{B},\mathcal{B}'} \in \text{GL}_n(\mathbb{K})$.

Preuve : Soit $u \in \mathcal{L}(E)$ défini par $\forall\, i \in [\![1\,;n]\!]$, $u(e_i) = x_i$, on a alors : $\mathcal{P}_{\mathcal{B},\mathcal{B}'} = \underset{\mathcal{B}}{\text{mat}}(u)$. On sait que u est bijectif si et seulement si $\underset{\mathcal{B}}{\text{mat}}(u)$ est inversible, et u est bijectif si et seulement si u transforme une base en une base, c'est à dire $(u(e_1), \dots, u(e_n))$ est une base, autrement dit $\mathcal{B}'$ est une base de E. □

Interprétation de la matrice de passage entre deux bases : Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E, en considérant l'application $\text{id}_E : (E, \mathcal{B}') \to (E, \mathcal{B})$ avec $\mathcal{B}'$ comme base au départ et $\mathcal{B}$ comme base à l'arrivée, on a la relation : $\mathcal{P}_{\mathcal{B},\mathcal{B}'} = \underset{\mathcal{B}',\mathcal{B}}{\text{mat}}(\text{id}_E)$.

Théorème 23.8 (application)

Soient $\mathcal{B}, \mathcal{B}', \mathcal{B}''$ trois bases de E, on a : $\mathcal{P}_{\mathcal{B}',\mathcal{B}} = [\mathcal{P}_{\mathcal{B},\mathcal{B}'}]^{-1}$ et $\mathcal{P}_{\mathcal{B},\mathcal{B}''} = \mathcal{P}_{\mathcal{B},\mathcal{B}'} \times \mathcal{P}_{\mathcal{B}',\mathcal{B}''}$.

Preuve : On a $[\mathcal{P}_{\mathcal{B},\mathcal{B}'}]^{-1} = \left[\underset{\mathcal{B}',\mathcal{B}}{\text{mat}}(\text{id}_E)\right]^{-1} = \underset{\mathcal{B},\mathcal{B}'}{\text{mat}}(\text{id}_E^{-1}) = \underset{\mathcal{B},\mathcal{B}'}{\text{mat}}(\text{id}_E) = \mathcal{P}_{\mathcal{B}',\mathcal{B}}$, ce qui prouve le premier point.

Soit $f = g = \text{id}_E$, alors $\text{id}_E = g \circ f : (E, \mathcal{B}'') \xrightarrow{f} (E, \mathcal{B}') \xrightarrow{g} (E, \mathcal{B})$, ce qui donne $\underset{\mathcal{B}'',\mathcal{B}}{\text{mat}}(\text{id}_E) = \underset{\mathcal{B}',\mathcal{B}}{\text{mat}}(g) \times \underset{\mathcal{B}'',\mathcal{B}'}{\text{mat}}(f)$, c'est à dire $\mathcal{P}_{\mathcal{B},\mathcal{B}''} = \underset{\mathcal{B}',\mathcal{B}}{\text{mat}}(\text{id}_E) \times \underset{\mathcal{B}'',\mathcal{B}'}{\text{mat}}(\text{id}_E) = \mathcal{P}_{\mathcal{B},\mathcal{B}'} \times \mathcal{P}_{\mathcal{B}',\mathcal{B}''}$, ce qui prouve le second point. □

2) Formules du changement de bases

Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E, pour tout vecteur $x \in E$ on peut calculer ses coordonnées dans la base $\mathcal{B}$: $X = \text{Coord}_{\mathcal{B}}(x)$, ou bien ses coordonnées dans la base $\mathcal{B}'$: $X' = \text{Coord}_{\mathcal{B}'}(x)$, on cherche le lien entre X et X'.

Théorème 23.9 (formules du changement de bases)

Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E, soit $x \in E$, on pose $X = \mathrm{Coord}_{\mathcal{B}}(x)$ et $X' = \mathrm{Coord}_{\mathcal{B}'}(x)$, on a les formules suivantes : $X = \mathcal{P}_{\mathcal{B},\mathcal{B}'} \times X'$ *et* $X' = \mathcal{P}_{\mathcal{B}',\mathcal{B}} \times X$.

Preuve : Considérons $\mathrm{id}_E : (E, \mathcal{B}') \to (E, \mathcal{B})$, on sait que $\underset{\mathcal{B}',\mathcal{B}}{\mathrm{mat}}(\mathrm{id}_E) = \mathcal{P}_{\mathcal{B},\mathcal{B}'}$. On a $\mathrm{id}_E(x) = x$, la relation fondamentale donne alors $\mathrm{Coord}_{\mathcal{B}}(\mathrm{id}_E(x)) = \mathcal{P}_{\mathcal{B},\mathcal{B}'} \times \mathrm{Coord}_{\mathrm{B}'}(x)$, d'où la relation : $X = \mathcal{P}_{\mathcal{B},\mathcal{B}'} \times X'$. De même, on a $X' = \mathcal{P}_{\mathcal{B}',\mathcal{B}} \times X$, c'est à dire $X' = [\mathcal{P}_{\mathcal{B},\mathcal{B}'}]^{-1} \times X$. □

★**Exercice 23.3** *Soit $\mathcal{B}$ la base canonique de $\mathbb{K}_3[X]$, on pose $\mathcal{B}' = (1, X, X(X-1), X(X-1)(X-2))$, montrer que $\mathcal{B}'$ est une base de $\mathbb{K}_3[X]$ et pour $P \in \mathbb{K}_3[X]$ calculer $\underset{\mathcal{B}'}{\mathrm{coord}}(P)$.*

3) Changement de bases et applications linéaires

Soient E et F deux $\mathbb{K}$-espaces vectoriels, soit $\mathcal{B}_1$ une base de E et soit $\mathcal{B}_2$ une base de F. Si $u \in \mathcal{L}(E, F)$ on peut calculer $A = \underset{\mathcal{B}_1,\mathcal{B}_2}{\mathrm{mat}}(u)$. Si on prend une autre base dans E : $\mathcal{B}_1'$ et une autre base dans F, $\mathcal{B}_2'$, alors on peut calculer $A' = \underset{\mathcal{B}_1',\mathcal{B}_2'}{\mathrm{mat}}(u)$, on cherche le lien entre ces deux matrices.

Théorème 23.10 (effet d'un changement de bases sur la matrice d'une application linéaire)

Soient $\mathcal{B}_1, \mathcal{B}_1'$ deux bases de E et $P = \mathcal{P}_{\mathcal{B}_1,\mathcal{B}_1'}$ la matrice de passage, soient $\mathcal{B}_2, \mathcal{B}_2'$ deux bases de F et soit $Q = \mathcal{P}_{\mathcal{B}_2,\mathcal{B}_2'}$ la matrice de passage, soit $u \in \mathcal{L}(E, F)$, on pose $A = \underset{\mathcal{B}_1,\mathcal{B}_2}{\mathrm{mat}}(u)$, $A' = \underset{\mathcal{B}_1',\mathcal{B}_2'}{\mathrm{mat}}(u)$, on a alors la relation : $A' = Q^{-1} \times A \times P$.

Preuve : On a $u = \mathrm{id}_F \circ u \circ \mathrm{id}_E : (E, \mathcal{B}_1') \xrightarrow{\mathrm{id}_E} (E, \mathcal{B}_1) \xrightarrow{u} (F, \mathcal{B}_2) \xrightarrow{\mathrm{id}_F} (E, \mathcal{B}_2')$, d'où :

$$\begin{aligned}
A' = \underset{\mathcal{B}_1',\mathcal{B}_2'}{\mathrm{mat}}(u) &= \underset{\mathcal{B}_1',\mathcal{B}_2'}{\mathrm{mat}}(\mathrm{id}_F \circ u \circ \mathrm{id}_E)\\
&= \underset{\mathcal{B}_2,\mathcal{B}_2'}{\mathrm{mat}}(\mathrm{id}_F) \times \underset{\mathcal{B}_1,\mathcal{B}_2}{\mathrm{mat}}(u) \times \underset{\mathcal{B}_1',\mathcal{B}_1}{\mathrm{mat}}(\mathrm{id}_E)\\
&= \mathcal{P}_{\mathcal{B}_2',\mathcal{B}_2} \times \underset{\mathcal{B}_1,\mathcal{B}_2}{\mathrm{mat}}(u) \times \mathcal{P}_{\mathcal{B}_1,\mathcal{B}_1'}\\
&= Q^{-1} \times A \times P
\end{aligned}$$

□

Théorème 23.11 (cas des endomorphismes)

Soient $\mathcal{B}, \mathcal{B}'$ deux bases de E et soit $P = \mathcal{P}_{\mathcal{B},\mathcal{B}'}$ la matrice de passage, soit $u \in \mathcal{L}(E)$, si on pose $A = \underset{\mathcal{B}}{\mathrm{mat}}(u)$ et $A' = \underset{\mathcal{B}'}{\mathrm{mat}}(u)$, alors on a la relation : $A' = P^{-1} \times A \times P$.

Preuve : Cela découle du théorème précédent, puisque l'on a $Q = P$. □

★**Exercice 23.4** *Soit $\mathcal{B} = (i, j)$ la base canonique de $\mathbb{K}^2$ et soit $u \in \mathcal{L}(\mathbb{K}^2)$ défini par $\underset{\mathcal{B}}{\mathrm{mat}}(u) = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$. On pose $e_1 = (1, 1)$ et $e_2 = (1, -1)$, montrer que $\mathcal{B}' = (e_1, e_2)$ est une base de $\mathbb{K}^2$, et calculer la matrice de u dans la base $\mathcal{B}'$. En déduire l'expression de $u^n(x, y)$.*

Définition 23.4

Soient $A, B \in \mathcal{M}_n(\mathbb{K})$, on dit que les matrices A et B sont semblables si et seulement si il existe une matrice carrée inversible $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $A = P^{-1} \times B \times P$.

Remarque 23.4 :

- *Les matrices d'un endomorphisme de E calculées dans une base $\mathcal{B}$ pour l'une, et une base $\mathcal{B}'$ pour l'autre, sont semblables, d'après le théorème précédent.*
- *Deux matrices sont semblables lorsque ce sont deux matrices d'un même endomorphisme exprimées dans deux bases (P représentant alors la matrice de passage entre les deux bases).*
- *La relation « ..est semblable à .. » est une relation d'équivalence dans $\mathcal{M}_n(\mathbb{K})$.*

4) Trace d'un endomorphisme

Théorème 23.12
Soient $A, B \in \mathcal{M}_n(\mathbb{K})$*, on a la propriété :* $\operatorname{tr}(A \times B) = \operatorname{tr}(B \times A)$.

Preuve : Rappelons que le produit matriciel n'est pas commutatif ! On a :

$$\operatorname{tr}(A \times B) = \sum_{i=1}^{n} [A \times B]_{i,i} = \sum_{i=1}^{n} \left(\sum_{k=1}^{n} A_{i,k} B_{k,i} \right), \text{ ce qui donne } \operatorname{tr}(A \times B) = \sum_{k=1}^{n} \left(\sum_{i=1}^{n} B_{k,i} A_{i,k} \right) = \sum_{k=1}^{n} [B \times A]_{k,k} = \operatorname{tr}(B \times A).$$

□

Remarque 23.5 – *Plus généralement, si* $A \in \mathcal{M}_{n,p}(\mathbb{K})$ *et* $B \in \mathcal{M}_{p,n}(\mathbb{K})$ *alors les produits* AB *et* BA *sont définis, et on montre de la même façon que* $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.

Théorème 23.13 (conséquence)
Deux matrices carrées semblables ont la même trace. C'est à dire :
Si $A \in \mathcal{M}_n(\mathbb{K})$ *et si* $P \in GL_n(\mathbb{K})$*, alors* $\operatorname{tr}(A) = \operatorname{tr}(P^{-1} \times A \times P)$.

Preuve : $\operatorname{tr}(A) = \operatorname{tr}\left(P^{-1}AP\right) = \operatorname{tr}\left([P^{-1}A] \times P\right) = \operatorname{tr}\left(P \times [P^{-1}A]\right) = \operatorname{tr}(I_n A) = \operatorname{tr}(A)$. □

Soit E un espace vectoriel de dimension n, soient $\mathfrak{B}$ et $\mathfrak{B}'$ deux bases de E, et soit $u \in \mathcal{L}(E)$, on note $A = \underset{\mathfrak{B}}{\operatorname{mat}}(u)$ et $A' = \underset{\mathfrak{B}'}{\operatorname{mat}}(u)$, on sait alors que $A' = P^{-1} \times A \times P$ avec $P = \mathcal{P}_{\mathfrak{B},\mathfrak{B}'}$ la matrice de passage, A et A′ étant semblables, d'après le théorème précédent, on peut affirmer que $\operatorname{tr}(A) = \operatorname{tr}(A')$.

Définition 23.5
Soit $u \in \mathcal{L}(E)$ *et soit* $\mathfrak{B}$ *une base de* E*, on appelle trace de l'endomorphisme* u *le scalaire noté* $\operatorname{tr}(u)$ *et défini par* $\operatorname{tr}(u) = \operatorname{tr}(\underset{\mathfrak{B}}{\operatorname{mat}}(u))$*, ce scalaire est indépendant de la base* $\mathfrak{B}$ *choisie.*

Théorème 23.14
L'application trace, $\operatorname{tr} : \mathcal{L}(E) \to \mathbb{K}$*, est une forme linéaire non nulle sur* $\mathcal{L}(E)$*, qui vérifie :*

$$\forall\, u, v \in \mathcal{L}(E), \operatorname{tr}(u \circ v) = \operatorname{tr}(v \circ u).$$

Preuve : Soit $\mathfrak{B}$ une base de E, soit $A = \underset{\mathfrak{B}}{\operatorname{mat}}(u)$ et $B = \underset{\mathfrak{B}}{\operatorname{mat}}(v)$, on a par définition :

$$\operatorname{tr}(u + v) = \operatorname{tr}(\underset{\mathfrak{B}}{\operatorname{mat}}(u + v)) = \operatorname{tr}(\underset{\mathfrak{B}}{\operatorname{mat}}(u)) + \operatorname{tr}(\underset{\mathfrak{B}}{\operatorname{mat}}(v)) = \operatorname{tr}(u) + \operatorname{tr}(v).$$

De la même façon, on montre que $\operatorname{tr}(\lambda u) = \lambda \operatorname{tr}(u)$ avec $\lambda \in \mathbb{K}$. On a donc une forme linéaire sur $\mathcal{L}(E)$, celle-ci est non nulle, car $\operatorname{tr}(\mathrm{id}_E) = \operatorname{tr}(\underset{\mathfrak{B}}{\operatorname{mat}}(\mathrm{id}_E)) = \operatorname{tr}(I_n) = n = \dim(E) \geqslant 1$. D'autre part, $\operatorname{tr}(u \circ v) = \operatorname{tr}(\underset{\mathfrak{B}}{\operatorname{mat}}(u \circ v)) = \operatorname{tr}(\underset{\mathfrak{B}}{\operatorname{mat}}(u) \times \underset{\mathfrak{B}}{\operatorname{mat}}(v)) = \operatorname{tr}(\underset{\mathfrak{B}}{\operatorname{mat}}(v) \times \underset{\mathfrak{B}}{\operatorname{mat}}(u)) = \operatorname{tr}(\underset{\mathfrak{B}}{\operatorname{mat}}(v \circ u)) = \operatorname{tr}(v \circ u)$. □

★**Exercice 23.5** *Soit* E *un espace vectoriel de dimension* n*, et soit* $p \in \mathcal{L}(E)$ *un projecteur, montrer que* $\operatorname{tr}(p) = \operatorname{rg}(p)$.

III OPÉRATIONS ÉLÉMENTAIRES SUR LES MATRICES

1) Rang d'une matrice

Définition 23.6
Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$ *une matrice, on appelle rang de la matrice* A*, le rang dans* $\mathbb{K}^n$ *de la famille constituée par ses* p *vecteurs colonnes, notation :* $\operatorname{rg}(A) = \operatorname{rg}(c_1(A), \dots, c_p(A))$.

Théorème 23.15
Soit $u \in \mathcal{L}(E, F)$*, soit* $\mathfrak{B}$ *une base de* E*, soit* $\mathfrak{B}'$ *une base de* F*, et soit* $A = \underset{\mathfrak{B},\mathfrak{B}'}{\operatorname{mat}}(u)$*, alors* $\operatorname{rg}(u) = \operatorname{rg}(A)$.

Preuve : Posons $\mathfrak{B} = (e_1, \dots, e_p)$, $\mathfrak{B}' = (e'_1, \dots, e'_n)$ et soit $\mathfrak{B}'' = (e''_1, \dots, e''_n)$ la base canonique de $\mathbb{K}^n$. Soit $v \in \mathcal{L}(F, \mathbb{K}^n)$ défini par $\forall\, i \in [\![1\,;n]\!], v(e'_i) = e''_i$, alors v est bijective (transforme une base en une base), donc $\operatorname{rg}(u) = \operatorname{rg}(v \circ u) = \operatorname{rg}(v(u(e_1)), \dots, v(u(e_p)))$; or $v(u(e_j)) = \sum_{k=1}^{n} A_{k,j} e''_j = (A_{1,j}, A_{2,j}, \dots, A_{n,j}) = c_j(A)$, donc $\operatorname{rg}(u) = \operatorname{rg}(A)$ d'après la définition précédente. □

Théorème 23.16 (conséquence)
Soit E *un espace vectoriel de dimension* n*, soit* $\mathrm{S} = (x_1, \dots, x_p)$ *une famille de* p *vecteurs de* E *et soit* $\mathfrak{B}$ *une base de* E*, alors* ***le rang de la famille*** S ***est égal au rang de la matrice de cette famille dans la base*** $\mathfrak{B}$*.*

Preuve : Posons $\mathfrak{B} = (e_1, \dots, e_n)$, soit $\mathfrak{B}' = (e'_1, \dots, e'_p)$ la base canonique de $\mathbb{K}^p$, soit $u \in \mathcal{L}(\mathbb{K}^p, \mathrm{E})$ l'application linéaire définie par : $\forall\, i \in [\![1\,;p]\!], u(e'_i) = x_i$, alors $\mathrm{A} = \underset{\mathfrak{B}',\mathfrak{B}}{\mathrm{mat}}(u)$ est la matrice de la famille S dans la base $\mathfrak{B}$, or $\mathrm{rg}(\mathrm{A}) = \mathrm{rg}(u) = \mathrm{rg}(x_1, \dots, x_p)$, ce qui donne le résultat. □

Calculer le rang d'une application linéaire, d'une famille de vecteurs, c'est calculer le rang d'une matrice.

2) Propriétés du rang d'une matrice

Les propriétés suivantes découlent de celles du rang des applications linéaires.

a) Soit $f \in \mathcal{L}(\mathrm{E}, \mathrm{F})$, soit $\mathfrak{B}$ une base de E avec $\dim(\mathrm{E}) = p$, soit $\mathfrak{B}'$ une base de F avec $\dim(\mathrm{F}) = n$, et soit $\mathrm{A} = \underset{\mathfrak{B},\mathfrak{B}'}{\mathrm{mat}}(f) \in \mathcal{M}_{n,p}(\mathbb{K})$, on a :
   i) $\mathrm{rg}(\mathrm{A}) \leqslant \min(n, p)$.
   ii) $\mathrm{rg}(\mathrm{A}) = n \iff f$ est surjective.
   iii) $\mathrm{rg}(\mathrm{A}) = p \iff f$ est injective.

b) Si $\mathrm{A} \in \mathcal{M}_n(\mathbb{K})$, alors $\mathrm{A} \in \mathrm{GL}_n(\mathbb{K}) \iff \mathrm{rg}(\mathrm{A}) = n$.

c) Si $\mathrm{A} \in \mathcal{M}_{n,p}(\mathbb{K}), \mathrm{B} \in \mathcal{M}_{p,q}(\mathbb{K})$, alors $\mathrm{rg}(\mathrm{A} \times \mathrm{B}) \leqslant \min(\mathrm{rg}(\mathrm{A}), \mathrm{rg}(\mathrm{B}))$.

d) Si $\mathrm{A} \in \mathrm{GL}_n(\mathbb{K}), \mathrm{B} \in \mathcal{M}_{n,p}(\mathbb{K})$, alors $\mathrm{rg}(\mathrm{A} \times \mathrm{B}) = \mathrm{rg}(\mathrm{B})$.

e) Si $\mathrm{A} \in \mathcal{M}_{n,p}(\mathbb{K}), \mathrm{B} \in \mathrm{GL}_p(\mathbb{K})$, alors $\mathrm{rg}(\mathrm{A} \times \mathrm{B}) = \mathrm{rg}(\mathrm{A})$.

Notation : soient $n, p \in \mathbb{N}^*$ et soit $r \in \mathbb{N}$ avec $r \leqslant \min(n, p)$ on note $\mathrm{J}_{n,p,r} \in \mathcal{M}_{n,p}(\mathbb{K})$ définie par :

$$\mathrm{J}_{n,p,r} = \begin{pmatrix} 1 & 0 & & \cdots & & 0 \\ 0 & \ddots & & & & \vdots \\ \vdots & & 1 & & & \\ 0 & \cdots & & 0 & \cdots & 0 \\ \vdots & & & & \ddots & \\ 0 & & \cdots & & & 0 \end{pmatrix}$$

Il y a r fois le scalaire 1 sur la diagonale, si $r = 0$ alors cette matrice est nulle.

Théorème 23.17
Soit $\mathrm{A} \in \mathcal{M}_{n,p}(\mathbb{K})$*, alors :* $\mathrm{rg}(\mathrm{A}) = r \iff \exists\, \mathrm{U} \in \mathrm{GL}_n(\mathbb{K}), \exists\, \mathrm{V} \in \mathrm{GL}_p(\mathbb{K}), \mathrm{UAV} = \mathrm{J}_{n,p,r}$.

Preuve : Si $r = 0$ il n'y a rien à prouver, car $\mathrm{rg}(\mathrm{A}) = 0 \iff \mathrm{A} = \mathrm{O}_{n,p}$, toute matrices inversibles U et V (de la bonne taille) conviennent, supposons $r > 0$.

Si U et V existent alors, comme elles sont inversibles, $\mathrm{rg}(\mathrm{A}) = \mathrm{rg}(\mathrm{UAV}) = \mathrm{rg}(\mathrm{J}_{n,p,r}) = r$.

Réciproquement, si $\mathrm{rg}(\mathrm{A}) = r$, soit $\mathfrak{B}_1$ la base canonique de $\mathbb{K}^p$, soit $\mathfrak{B}_2$ la base canonique de $\mathbb{K}^n$, et soit $u \in \mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$ défini par $\underset{\mathfrak{B}_1,\mathfrak{B}_2}{\mathrm{mat}}(u) = \mathrm{A}$ ($u = f_{\mathrm{A}}$ l'application linéaire canoniquement associée à A), on a $\mathrm{rg}(u) = \mathrm{rg}(\mathrm{A}) = r$, d'après le théorème du rang, $\dim(\ker(u)) = p - r$, soit H un supplémentaire de $\ker(u)$ dans $\mathbb{K}^p$ et soit $(e_1, \dots, e_r)$ une base de H, soit $(e_{r+1}, \dots, e_p)$ une base de $\ker(u)$, alors $\mathfrak{B}'_1 = (e_1, \dots, e_p)$ est une base de $\mathbb{K}^p$. On sait que $(u(e_1), \dots, u(e_r))$ est une base de $\mathrm{Im}(u)$ (cf démo du théorème du rang), on peut compléter en une base de F : $\mathfrak{B}'_2 = (u(e_1), \dots, u(e_r), v_{r+1}, \dots, v_n)$ et la matrice de u dans les bases $\mathfrak{B}'_1$ et $\mathfrak{B}'_2$ est exactement $\mathrm{J}_{n,p,r}$. Soit $\mathrm{P} = \mathcal{P}_{\mathfrak{B},\mathfrak{B}'_1}$ et $\mathrm{Q} = \mathcal{P}_{\mathfrak{B}_2,\mathfrak{B}'_2}$, d'après les formules du changement de bases, on a $\mathrm{J}_{n,p,r} = \mathrm{Q}^{-1} \times \mathrm{A} \times \mathrm{P}$, ce qui termine la preuve, en prenant $\mathrm{U} = \mathrm{Q}^{-1}$ et $\mathrm{V} = \mathrm{P}$. □

★**Exercice 23.6** *Montrer qu'une matrice* $\mathrm{A} \in \mathcal{M}_{n,p}(\mathbb{K})$ *et sa transposée ont le même rang.*

Théorème 23.18
Les opérations élémentaires conservent le rang de la matrice.

Preuve : Multiplier une matrice par une autre matrice inversible ne change pas le rang de la première. □

3) Calcul pratique du rang d'une matrice

Définition 23.7

Soient $A, B \in \mathcal{M}_{n,p}(\mathbb{K})$*, on dit que* A *et* B *sont équivalentes lorqu'il existe* $Q \in GL_n(\mathbb{K})$ *et* $P \in GL_p(\mathbb{K})$ *telles que* $B = QAP$*.*

Remarque 23.6 –

- *On définit ainsi une relation d'équivalence dans* $\mathcal{M}_{n,p}(\mathbb{K})$*.*
- *Deux matrices équivalentes ont le même rang.*
- *Une opération élémentaire donne une matrice équivalente.*
- *Deux matrices carrées semblables sont équivalentes.*
- $A \in \mathcal{M}_{n,p}(\mathbb{K})$ *a un rang égal à* r *si et seulement si* A *est équivalente à* $J_{n,p,r}$*.*

Théorème 23.19

Deux matrices sont équivalentes si et seulement si elles ont le même rang.

Preuve : Celle-ci est simple et laissée en exercice. □

Théorème 23.20

Si $A \in \mathcal{M}_{n,p}(\mathbb{K})$ *est équivalente à* $B = \begin{pmatrix} p_1 & * & \cdots & * & \cdots & * \\ 0 & \ddots & \ddots & \vdots & \cdots & \vdots \\ \vdots & \ddots & p_r & * & \cdots & * \\ \vdots & & 0 & 0 & \cdots & 0 \\ \vdots & & \boxtimes & \vdots & \cdots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}$*, avec* $p_1 \times \cdots \times p_r \neq 0$*, alors* $\mathrm{rg}(A) = r$*. Les symboles* $*$ *désignent des scalaires quelconques.*

Preuve : Le rang de B est le rang de ses vecteurs lignes, d'où $\mathrm{rg}(B) = \mathrm{rg}(l_1(B), \ldots, l_r(B)) = r$ (on enlève les lignes qui sont nulles, et les r premiers vecteurs lignes forment une famille libre de $\mathbb{K}^p$), or A est équivalente à B donc $\mathrm{rg}(A) = \mathrm{rg}(B) = r$. □

La méthode

Celle-ci consiste à transformer la matrice A en la matrice B ci-dessus à l'aide des opérations élémentaires sur les lignes ou les colonnes (méthode de Gauss), à chaque étape, la matrice obtenue a le même rang que A, plus précisément, à chaque étape la nouvelle matrice s'écrit sous la forme $U_k \times A \times V_k$ avec U_k, V_k inversibles. À l'étape n° k, le principe est le suivant :

a) On choisit un pivot (*i.e.* un coefficient non nul) dans les lignes L_k à L_n et dans les colonnes C_k à C_p.

b) On amène le pivot à sa place, c'est à dire sur la ligne L_k dans la colonne C_k en échangeant éventuellement deux lignes et/ou deux colonnes.

c) On fait des éliminations **en dessous** du pivot pour faire apparaître des 0, avec les opérations du type : $L_i \leftarrow L_i + \alpha L_k$.

☞**Exemple** : Soit $A = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 0 & 2 \\ -1 & 2 & -12 \end{pmatrix}$

Étape 1 : premier pivot : 1 (ligne L_1 colonne C_2)

$$C_1 \leftrightarrow C_2 \text{ et } L_3 \leftarrow L_3 - 2L_1 \text{ donnent } \begin{pmatrix} \boxed{1} & 3 & 1 \\ 0 & 1 & 2 \\ 0 & -7 & -14 \end{pmatrix}.$$

Étape 2 : deuxième pivot : 1 (ligne L_2 colonne C_2)

$$L_3 \leftarrow L_3 + 7L_2 \text{ donne } \begin{pmatrix} \boxed{1} & 3 & 1 \\ 0 & \boxed{1} & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

Étape 3 : pas de troisième pivot, donc $\mathrm{rg}(A) = 2$.

★**Exercice** 23.7 *Avec la matrice* A *précédente, déduire de la méthode deux matrices inversibles* U *et* V *telles que* $UAV = J_{3,3,2}$.

☞**Exemple** : (Variante) Il peut être parfois avantageux de n'effectuer que des transformations sur les colonnes, les éliminations se font alors à droite du pivot avec les opérations du type $C_i \leftarrow C_i + \alpha C_k$ (à l'étape k). Voici quel peut être l'intérêt : soit $\mathfrak{B} = (i, j, k)$ une base de E et soit $u \in \mathcal{L}(E)$ défini par $\underset{\mathfrak{B}}{\operatorname{mat}}(u) = A$ (la matrice précédente), calculons le rang de A (donc le rang de u) en faisant uniquement des opérations sur les colonnes :

Étape 1 : premier pivot 1 (ligne L_1 colonne C_2)

$$C_1 \leftrightarrow C_2 \text{ donne } \begin{pmatrix} \boxed{1} & 3 & 1 \\ 0 & 1 & 2 \\ 2 & -1 & -12 \end{pmatrix}$$

La matrice obtenue est la matrice dans la base $\mathfrak{B}$ de la famille $(u(j), u(i), u(k))$.

$$C_2 \leftarrow C_2 - 3C_1 \text{ et } C_3 \leftarrow C_3 - C_1 \text{ donnent } \begin{pmatrix} \boxed{1} & 0 & 0 \\ 0 & 1 & 2 \\ 2 & -7 & -14 \end{pmatrix}.$$

La matrice obtenue est la matrice dans la base $\mathfrak{B}$ de la famille $(u(j), u(i-3j), u(k-j))$.

Étape 2 : deuxième pivot 1 (ligne L_2 colonne C_2)

$$C_3 \leftarrow C_3 - 2C_2 \text{ donne } \begin{pmatrix} \boxed{1} & 0 & 0 \\ 0 & \boxed{1} & 0 \\ 2 & -7 & 0 \end{pmatrix}$$

La matrice obtenue est la matrice dans la base $\mathfrak{B}$ de de la famille $(u(j), u(i-3j), u(-2i+5j+k))$.

On en déduit que le rang de u est égal à 2 et $\operatorname{Im}(u) = \operatorname{Vect}[u(j), u(i-3j)] = \operatorname{Vect}[u(j), u(i)]$. Le théorème du rang permet d'en déduire que $\dim(\ker(u) = 3 - 2 = 1$, or $-2i + 5j + k$ est dans $\ker(u)$ et non nul, donc $\ker(u) = \operatorname{Vect}[-2i + 5j + k]$

IV SOLUTION DES EXERCICES

Solution 23.1

1/ a) $D(1) = 0$, $D(X) = 1$, $D(X^2) = 2X$ *et* $D(X^3) = 3X^2$, *d'où* $\underset{\mathfrak{B}}{\operatorname{mat}}(D) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$.

b) $\Delta(1) = 0$, $\Delta(X) = 1$, $\Delta(X^2) = 2X + 1$ *et* $\Delta(X^3) = 3X^2 + 3X + 1$, *d'où* $\underset{\mathfrak{B}}{\operatorname{mat}}(\Delta) = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$.

c) $\mathfrak{B}' = (P_0, P_1, P_2, P_3)$ *est une base de* E *(degrés étagés), et* $\Delta(P_0) = 0$, $\Delta(P_1) = P_0$, $\Delta(P_2) = P_1$ *et* $\Delta(P_3) = P_2$, *d'où* $\underset{\mathfrak{B}'}{\operatorname{mat}}(\Delta) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$.

2/ *Dans* $\mathcal{M}_2(\mathbb{K})$ *on a* ${E^{11}}^T = E^{11}$, ${E^{12}}^T = E^{21}$, ${E^{21}}^T = E^{12}$, ${E^{22}}^T = E^{22}$, *la matrice demandée est donc* $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

Solution 23.2

1/ $\operatorname{Coord}_{\mathfrak{B}}(x, y, z) = X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, *d'où* $\operatorname{Coord}_{\mathfrak{B}'}(f(x,y,z)) = A \times X = \begin{pmatrix} x - 2y + 3z \\ 2x + y - 5z \end{pmatrix}$, *donc* $f(x, y, z) = (x - 2y + 3z, 2x + y - 5z)$ *(*$\mathfrak{B}'$ *est la base canonique de* $\mathbb{K}^2$*).*

2/ *On a* $(x,y,z) = xi + yj + zk = z(i+j+k) + (y-z)(i+j) + (x-y)i$, *donc* $\mathrm{Coord}_{\mathcal{B}'}(x,y,z) = \mathrm{X} = \begin{pmatrix} x-y \\ y-z \\ z \end{pmatrix}$, *d'où* $\mathrm{Coord}_{\mathcal{B}'}(f(x,y,z)) = \mathrm{A}\times\mathrm{X} = \begin{pmatrix} x-2y+z \\ 2y-3z \\ z \end{pmatrix}$, *c'est à dire* $f(x,y,z) = (x-2y+z)i + (2y-3z)(i+j) + z(i+j+k)$, *et donc* $f(x,y,z) = (x-z, 2y-2z, z)$.

3/ *Soit* $\mathcal{B}$ *la base canonique de* $\mathbb{K}^p$ *et* $\mathcal{B}'$ *la base canonique de* $\mathbb{K}^n$, *soit* $f \in \mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$ *l'application linéaire définie par* $\underset{\mathcal{B},\mathcal{B}'}{\mathrm{mat}}(f) = \mathrm{A}-\mathrm{B}$. *Pour* $x \in \mathbb{K}^p$, *posons* $\mathrm{X} = \mathrm{Coord}_{\mathcal{B}}(x)$, *on a alors* $\mathrm{Coord}_{\mathcal{B}'}(f(x)) = (\mathrm{A}-\mathrm{B})\mathrm{X} = \mathrm{O}_{n,1}$, *ce qui montre que* f *est l'application nulle, donc sa matrice est nulle, ce qui donne* $\mathrm{A} = \mathrm{B}$.

Solution 23.3 *La matrice de passage de* $\mathcal{B}$ *à* $\mathcal{B}'$ *est* $\mathrm{A} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 1 \end{pmatrix}$, *cette matrice est triangulaire et ses éléments diagonaux sont tous non nuls, elle est donc inversible, ce qui prouve que* $\mathcal{B}'$ *est une base de* $\mathbb{K}_3[\mathrm{X}]$. *On a la relation* $\mathrm{Coord}_{\mathcal{B}'}(\mathrm{P}) = \mathrm{A}^{-1}\times\mathrm{Coord}_{\mathcal{B}}(\mathrm{P})$, *il faut donc calculer* A^{-1}, *on peut résoudre l'équation* $\mathrm{A}\times\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}$, *ce qui donne* $\begin{cases} x &= a \\ y &= b+c+d \\ z &= c+3d \\ t &= d \end{cases}$, *on en déduit que* $\mathrm{A}^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix}$. *Finalement, si* $\mathrm{P} = a + b\mathrm{X} + c\mathrm{X}^2 + d\mathrm{X}^3$, *alors* $\mathrm{Coord}_{\mathcal{B}'}(\mathrm{P}) = \begin{pmatrix} a \\ b+c+d \\ c+3d \\ d \end{pmatrix}$.

Solution 23.4 *La matrice de passage de* $\mathcal{B}$ *à* $\mathcal{B}'$ *est* $\mathrm{P} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, *cette matrice est inversible et* $\mathrm{P}^{-1} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, *on en déduit que* $\mathcal{B}'$ *est bien une base de* $\mathbb{K}^2$ *et que* $\underset{\mathcal{B}'}{\mathrm{mat}}(u) = \mathrm{A}' = \mathrm{P}^{-1}\times\mathrm{A}\times\mathrm{P} = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$. *On en déduit que* $\mathrm{A}'^n = \begin{pmatrix} 1 & 0 \\ 0 & 3^n \end{pmatrix}$, *on a alors* $\mathrm{A}^n = [\mathrm{P}\times\mathrm{A}'\times\mathrm{P}^{-1}]^n = \mathrm{P}\times\mathrm{A}'^n\times\mathrm{P}^{-1}$ *ce qui donne :* $\mathrm{A}^n = \frac{1}{2}\begin{pmatrix} 1+3^n & 1-3^n \\ 1-3^n & 1+3^n \end{pmatrix}$, *or* $\mathrm{A}^n = \underset{\mathcal{B}}{\mathrm{mat}}(u^n)$, *par conséquent :*

$$\forall\,(x,y) \in \mathbb{K}^2, u^n(x,y) = \tfrac{1}{2}\left((1+3^n)x + (1-3^n)y; (1-3^n)x + (1+3^n)y\right).$$

Solution 23.5 $r = \mathrm{rg}(p)$, *on a* $\mathrm{E} = \mathrm{Im}(p)\oplus\ker(p)$, *soit* $(e_1,\ldots,e_r)$ *une base de* $\mathrm{Im}(p)$ *et soit* $(e_{r+1},\ldots,e_n)$ *une base de* $\ker(p)$, *alors* $\mathcal{B} = (e_1,\ldots,e_n)$ *est une base de* E *et il est clair que* $\underset{\mathcal{B}}{\mathrm{mat}}(p) = \mathrm{J}_{n,n,r}$ *(car* $\mathrm{Im}(p) = \ker(p-\mathrm{id}_{\mathrm{E}})$*), d'où* $\mathrm{tr}(p) = \mathrm{tr}(\mathrm{J}_{n,n,r}) = r = \mathrm{rg}(p)$.

Solution 23.6 *Il existe* $\mathrm{U} \in \mathrm{GL}_n(\mathbb{K}), \mathrm{V} \in \mathrm{GL}_p(\mathbb{K})$ *telles que* $\mathrm{UAV} = \mathrm{J}_{n,p,r}$ *avec* $r = \mathrm{rg}(\mathrm{A})$. *On a alors* $\mathrm{V}^{\mathrm{T}}\mathrm{A}^{\mathrm{T}}\mathrm{U}^{\mathrm{T}} = {\mathrm{J}_{n,p,r}}^{\mathrm{T}} = \mathrm{J}_{p,n,r}$, *ce qui donne le résultat.*

Solution 23.7 *On transforme la matrice obtonue précédemment en* $\mathrm{J}_{3,3,2}$:

$$\mathrm{C}_2 \leftarrow \mathrm{C}_2 - 3\mathrm{C}_1, \mathrm{C}_3 \leftarrow \mathrm{C}_3 - \mathrm{C}1 \ \textit{donnent} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}, \mathrm{C}_3 \leftarrow \mathrm{C}_3 - 2\mathrm{C}_2 \ \textit{donne} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

On a obtient la matrice U *en effectuant sur lignes de* I_3 *les mêmes opérations que sur les lignes de* A *(dans le même ordre), et la matrice* V *en effectuant sur colonnes de* I_3*, les mêmes opérations sur les colonnes de* A *(dans le même ordre) :*

$$\mathrm{U} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 7 & 1 \end{pmatrix} \ \textit{et}\ \mathrm{V} = \begin{pmatrix} 0 & 1 & -2 \\ 1 & -3 & 5 \\ 0 & 0 & 1 \end{pmatrix}$$

Chapitre 24

Dénombrement

Sommaire

I CARDINAL D'UN ENSEMBLE FINI

1) Rappels : injections, surjections, bijections, permutations

a) La composée de deux injections (respectivement surjections) est une injection (respectivement surjection).

b) Si $f\colon \mathrm{E} \to \mathrm{F}$ est injective, alors f **induit** une bijection de E sur $\mathrm{Im}(f)$.

c) Si $f \circ g$ est injective, alors g est injective.

d) Si $f \circ g$ est surjective, alors f est surjective.

e) Si $f\colon \mathrm{E} \to \mathrm{F}$ est une application, alors f induit une surjection de E sur $\mathrm{Im}(f)$.

f) Si $f\colon \mathrm{E} \to \mathrm{F}$ est surjective, alors il existe une application $g\colon \mathrm{F} \to \mathrm{E}$ telle que $f \circ g = \mathrm{id}_{\mathrm{F}}$.

Définition 24.1

Soit E *un ensemble, on appelle permutation de* E *toute bijection de* E *vers* E*. L'ensemble des permutations de* E *est noté* $\mathcal{S}(\mathrm{E})$.

Théorème 24.1

Soit E *un ensemble non vide, alors* $(\mathcal{S}(\mathrm{E}), \circ)$ *est un groupe (non commutatif en général), appelé groupe des permutations de* E.

Preuve : Celle - ci est simple et laissée en exercice. On vérifie que l'élément neutre est l'application identité de E : id_{E}, et que le symétrique de $f \in \mathcal{S}(\mathrm{E})$ est la bijection réciproque f^{-1}. □

2) Ensembles finis

Définition 24.2

Soit E *un ensemble non vide, on dit que* E *est fini lorsqu'il existe un entier* $n \in \mathbb{N}^*$ *et une bijection* $\phi\colon [\![1\,;n]\!] \to \mathrm{E}$*. Si c'est le cas, on pose* $\operatorname{card}(\mathrm{E}) = n$ *(ou* $|\mathrm{E}| = n$ *ou* $\#(\mathrm{E}) = n$*), sinon on dit que* E *est un ensemble infini. Par convention* $\varnothing$ *est un ensemble fini de cardinal nul.*

Remarque 24.1 :

– *Dire que* E *est fini de cardinal* $n \geqslant 1$ *revient à dire que l'on peut indexer les éléments de* E *de 1 à* n *:* $\mathrm{E} = \{e_1, \dots, e_n\}$ *(les éléments étant distincts deux à deux).*

– *Si* E *est fini de cardinal* $n+1$ *et si* $a \in \mathrm{E}$*, alors* $\mathrm{E} \setminus \{a\}$ *est fini de cardinal* n*.*
En effet : soit $\phi\colon [\![1\,;n+1]\!] \to \mathrm{E}$ *une bijection, soit* τ *la permutation de* E *qui échange* $\phi(n+1)$ *et* a*, alors* $\tau\circ\phi\colon [\![1\,;n+1]\!] \to \mathrm{E}$ *est une bijection qui envoie* $n+1$ *en* a*, elle induit donc une bijection de* $[\![1\,;n]\!]$ *sur* $\mathrm{E}\setminus\{a\}$*.*

– *Si* E *est fini de cardinal* n *et* $b \notin \mathrm{E}$*, alors* $\mathrm{E} \cup \{b\}$ *est fini de cardinal* $n+1$*.*

Théorème 24.2

Soit $n \in \mathbb{N}^*$*, toute partie de* $[\![1\,;n]\!]$ *est un ensemble fini de cardinal au plus égal à* n*. De plus, si* $\mathrm{F} \subset [\![1\,;n]\!]$ *et si* $\operatorname{card}(\mathrm{F}) = n$ *alors* $\mathrm{F} = [\![1\,;n]\!]$*.*

Preuve : Par récurrence sur n : pour $n = 1$ c'est évident. Supposons le théorème établi pour un entier $n \geqslant 1$ et soit F une partie de $[\![1\,;n+1]\!]$. Si $n+1 \notin \mathrm{F}$, alors F est une partie de $[\![1\,;n]\!]$ donc (hypothèse de récurrence) F est fini et $\operatorname{card}(\mathrm{F}) \leqslant n < n+1$. Si $n+1 \in \mathrm{F}$, alors $\mathrm{F}\setminus\{n+1\}$ est une partie de $[\![1\,;n]\!]$, donc $\mathrm{F}\setminus\{n+1\}$ est un ensemble fini de cardinal $p \leqslant n$, mais alors F est fini de cardinal $p+1 \leqslant n+1$. Supposons maintenant que $\operatorname{card}(\mathrm{F}) = n+1$, on a nécessairement $n+1 \in \mathrm{F}$, d'où $\mathrm{F}\setminus\{n+1\} \subset [\![1\,;n]\!]$ et $\operatorname{card}(\mathrm{F}\setminus\{n+1\}) = n$, donc $\mathrm{F}\setminus\{n+1\} = [\![1\,;n]\!]$ (hypothèse de récurrence) et finalement $\mathrm{F} = [\![1\,;n+1]\!]$. □

Théorème 24.3

Soient $n, p \in \mathbb{N}^*$*, et soit* $f\colon [\![1\,;n]\!] \to [\![1\,;p]\!]$ *une application :*

– *Si* f *est injective, alors* $n \leqslant p$*.*
– *Si* f *est surjective, alors* $n \geqslant p$*.*
– *Si* f *est bijective, alors* $n = p$*.*

Preuve : On remarque que la troisième propriété découle des deux précédentes. Montrons la première : on a $f\colon [\![1\,;n]\!] \to [\![1\,;p]\!]$ une injection, alors f induit une bijection de $[\![1\,;n]\!]$ sur $\operatorname{Im}(f)$, donc $\operatorname{Im}(f)$ est fini de cardinal n, or $\operatorname{Im}(f)$ est une partie de $[\![1\,;p]\!]$, donc $\operatorname{Im}(f)$ est fini de cardinal au plus p, *i.e.* $n \leqslant p$.

Montrons la deuxième : $f\colon [\![1\,;n]\!] \to [\![1\,;p]\!]$ est surjective, alors il existe une application $g\colon [\![1\,;p]\!] \to [\![1\,;n]\!]$ telle que $f\circ g = \mathrm{id}_{[\![1;p]\!]}$, donc g est injective et par conséquent $p \leqslant n$. □

Conséquence : soit E un ensemble fini non vide, il existe un entier $n \geqslant 1$ et une bijection $\phi\colon [\![1\,;n]\!] \to \mathrm{E}$, s'il existe un autre entier p et une bijection $\psi : [\![1\,;p]\!] \to \mathrm{E}$, alors l'application $\psi^{-1}\circ\phi$ est une bijection de $[\![1\,;n]\!]$ sur $[\![1\,;p]\!]$, donc $n = p$. Ce qui prouve l'unicité du nombre $\operatorname{card}(\mathrm{E})$ et justifie à posteriori la définition.

Théorème 24.4

Soit $n \geqslant 1$*, toute application injective (respectivement surjective) de* $[\![1\,;n]\!]$ *dans* $[\![1\,;n]\!]$ *est bijective.*

Preuve : Si $f : [\![1\,;n]\!] \to [\![1\,;n]\!]$ est injective, alors f induit une bijection de $[\![1\,;n]\!]$ sur $\operatorname{Im}(f)$, donc $\operatorname{Im}(f)$ est fini de cardinal n, mais $\operatorname{Im}(f) \subset [\![1\,;n]\!]$, donc $\operatorname{Im}(f) = [\![1\,;n]\!]$ *i.e.* f est surjective (et donc bijective).

Supposons maintenant que f est surjective, alors il existe $g\colon [\![1\,;n]\!] \to [\![1\,;n]\!]$ telle que $f\circ g = \mathrm{id}_{[\![1;n]\!]}$, mais alors g est injective, donc bijective d'après ce qui précède et $f = (f\circ g)\circ g^{-1}$ composée de bijections, donc f est bijective. □

3) Propriétés du cardinal

Théorème 24.5

Soient E *et* F *deux ensembles finis non vides, avec* $n = \mathrm{card}(E)$ *et* $p = \mathrm{card}(F)$ *et soit* $f\colon E \to F$ *une application :*
- *Si f est injective alors $n \leqslant p$.*
- *Si f est surjective alors $n \geqslant p$.*
- *Si f est bijective alors $n = p$.*

Preuve : Soient $\phi_1\colon [\![1\,;n]\!] \to E$ et $\phi_2\colon [\![1\,;p]\!] \to F$ deux bijections. Si f est injective alors $\phi_2^{-1} \circ f \circ \phi_1$ est une injection de $[\![1\,;n]\!]$ vers $[\![1\,;p]\!]$, donc $n \leqslant p$. Le raisonnement est le même pour les deux autres points. □

Remarque 24.2 – *Il en découle que si* F *est en bijection avec* E *et si* E *est fini, alors* F *est fini de même cardinal que* E.

Théorème 24.6

Soient E *et* F *deux ensembles finis non vides* ***de même cardinal*** *et soit* $f\colon E \to F$ *une application, les assertions suivantes sont équivalentes :*

a) f est injective.

b) f est surjective.

c) f est bijective.

Preuve : Soient $\phi\colon [\![1\,;n]\!] \to E$ et $\psi\colon [\![1\,;n]\!] \to F$ deux bijections, alors $g = \psi^{-1} \circ f \circ \phi$ est une application de $[\![1\,;n]\!]$ vers lui - même, avec $f = \psi \circ g \circ \phi^{-1}$. Si f est injective, alors g aussi, donc g est bijective et f aussi. Si f est surjective, alors g aussi et donc g est bijective et f aussi. □

★**Exercice 24.1** *Soit* A *un anneau intègre fini, montrer que* A *est un corps.*

Théorème 24.7

Si E *est un ensemble fini et si* F *est une partie de* E, *alors* F *est fini et* $\mathrm{card}(F) \leqslant \mathrm{card}(E)$. *De plus, si* $\mathrm{card}(F) = \mathrm{card}(E)$, *alors* $F = E$.

Preuve : On écarte le cas évident où $E = \varnothing$. Soit $n = \mathrm{card}(E)$ et $\phi\colon [\![1\,;n]\!] \to E$ une bijection. Notons $i\colon F \to E$ définie par $i(x) = x$, i est une injection donc $g = \phi^{-1} \circ i$ est une injection de F vers $[\![1\,;n]\!]$ qui induit donc une bijection de F sur $\mathrm{Im}(g)$, or $\mathrm{Im}(g)$ est une partie de $[\![1\,;n]\!]$, donc $\mathrm{Im}(g)$ est un ensemble fini de cardinal $p \leqslant n$, par conséquent F est fini de cardinal $p \leqslant \mathrm{card}(E)$. Si $n = p$, alors $\mathrm{Im}(g) = [\![1\,;n]\!]$ donc g est une bijection ce qui entraîne que i est une bijection, donc $\mathrm{Im}(i) = E$, c'est à dire $F = E$. □

Théorème 24.8

Soient E *et* F *deux ensembles finis, l'ensemble* $E \cup F$ *est fini et :*

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

Preuve : Si l'un des deux est vide, il n'y a rien à démontrer. Supposons E et F non vides, dans un premier temps on envisage le cas où $E \cap F = \varnothing$, soit $f\colon [\![1\,;n]\!] \to E$ et $g\colon [\![1\,;p]\!] \to F$ deux bijections, on considère l'application $\phi\colon [\![1\,;n+p]\!] \to E \cup F$ définie par $\phi(k) = f(k)$ si $1 \leqslant k \leqslant n$ et $\phi(k) = g(k-n)$ si $n+1 \leqslant k \leqslant n+p$, comme $E \cap F = \varnothing$ on voit que ϕ est injective, d'autre part la surjectivité est évidente, donc ϕ est bijective, ce qui montre que $E \cup F$ est fini de cardinal $n+p$.

Passons maintenant au cas général : posons $I = E \cap F$, on a $E \cup F = E \cup (F \setminus E)$ et ces deux ensembles sont disjoints et finis, donc $E \cup F$ est fini et $\mathrm{card}(E \cup F) = \mathrm{card}(E) + \mathrm{card}(F \setminus E)$, d'autre part $F = I \cup (F \setminus E)$ et ces deux ensembles sont disjoints et finis, donc $\mathrm{card}(F) = \mathrm{card}(I) + \mathrm{card}(F \setminus E)$, on a donc $\mathrm{card}(F \setminus E) = \mathrm{card}(F) - \mathrm{card}(I)$, ce qui donne la formule. □

Théorème 24.9

Si E *et* F *sont deux ensembles finis, alors l'ensemble* $E \times F$ *est fini et* $\mathrm{card}(E \times F) = \mathrm{card}(E) \times \mathrm{card}(F)$.

Preuve : Si l'un des deux est vide, alors $E \times F$ est vide et le résultat est évident. Soit $n = \mathrm{card}(E)$, si $n = 1$ alors $E = \{e\}$ et l'application $f\colon F \to E \times F$ définie par $f(x) = (e, x)$ est une bijection, donc $E \times F$ est fini de même cardinal que F, le théorème est donc vrai pour $n = 1$.

Supposons le théorème démontré pour un entier $n \geqslant 1$ et supposons $\mathrm{card}(E) = n+1$, on fixe un élément $e \in E$ et on pose $E' = E \setminus \{e\}$. On a $E \times F = (\{e\} \times F) \cup (E' \times F)$, ces deux ensembles sont disjoints et finis (hypothèse de récurrence), donc $E \times F$ est fini et $\mathrm{card}(E \times F) = \mathrm{card}(\{e\} \times F) + \mathrm{card}(E' \times F) = \mathrm{card}(F) + \mathrm{card}(E') \times \mathrm{card}(F) = (n+1) \times \mathrm{card}(F)$. Le théorème est démontré au rang $n+1$. □

Conséquence : si $p \in \mathbb{N}^*$, et si E est fini de cardinal $n \geqslant 1$, alors E^p (ensemble des p - uplets d'éléments de E) est fini et $\text{card}(E^p) = [\text{card}(E)]^p$.

II DÉNOMBREMENT

1) Préliminaires

Définition 24.3

Dénombrer un ensemble fini E *c'est calculer son cardinal. Dans la pratique, c'est le mettre en bijection avec un ensemble* F *dont on connaît le cardinal.*

La fonction factorielle : elle est définie sur $\mathbb{N}$ par : $n! = \begin{cases} 1 & \text{si } n = 0 \\ 1 \times \cdots \times n & \text{si } n > 0 \end{cases}$. On peut également en donner une définition récurrente : $0! = 1$ et $\forall\, n \in \mathbb{N}, (n+1)! = (n+1) \times n!$.

Théorème 24.10 (diviser pour mieux compter)

Soient E *un ensemble fini et soient* $A_1, \ldots, A_n$ *n parties de* E *deux à deux disjointes et dont la réunion est égale à* E*, alors :* $\text{card}(E) = \sum\limits_{k=1}^{n} \text{card}(A_k)$.

Preuve : Celle - ci est simple, c'est un raisonnement par récurrence sur n, sachant que la formule est vraie pour $n = 2$. □

Application – Si $f : E \to F$ est une application et si E est fini, alors :

$$\text{card}(E) = \sum_{y \in \text{Im}(f)} \text{card}(f^{-1}(\{y\}))$$

Dans le cas où les éléments de $\text{Im}(f)$ ont tous le même nombre d'antécédents p, alors $\text{card}(E) = p\,\text{card}(\text{Im}(f))$.

2) Le nombre d'applications

Théorème 24.11

Soit E *et* F *deux ensembles finis avec* $p = \text{card}(E)$ *et* $n = \text{card}(F)$*, l'ensemble des applications de* E *vers* F*,* $\mathcal{F}(E, F)$ *(ou* F^E*), est fini de cardinal* n^p*.*

Preuve : Posons $E = \{e_1, \ldots, e_p\}$, on vérifie que l'application $\phi : F^E \to F^p$ définie par $\phi(f) = (f(e_1), \ldots, f(e_p))$ est une bijection. Or F^p est un ensemble fini de cardinal n^p ce qui donne le résultat. □

Remarque 24.3 :

- *Le théorème justifie le raisonnement suivant : pour construire une application de* E *vers* F *on compte pour chaque élément de* E *le nombre de choix possibles pour son image (soit n choix), puis on fait le produit, soit* n^p *constructions possibles.*
- *Le nombre de façons de tirer avec remise p boules parmi n est* n^p*.*
- *Le nombre de façons de ranger p boules dans n boites est* n^p*.*

Complément : lorsque $p \leqslant n$, le nombre d'injections de E vers F est $n(n-1)\cdots(n-p+1)$.

3) Le nombre de parties d'un ensemble

Définition 24.4

Soit E *un ensemble et* A *une partie de* E*, on appelle* ***fonction indicatrice*** *de* A *l'application* $1_A : E \to \{0; 1\}$ *définie par* $1_A(x) = \begin{cases} 1 & \textit{si } x \in A \\ 0 & \textit{sinon} \end{cases}$.

Théorème 24.12
Si E est fini de cardinal n, alors $\mathrm{P(E)}$, l'ensemble des parties de E, est fini de cardinal 2^n.

Preuve : Il est facile de vérifier que l'application de $\mathcal{P}(\mathrm{E})$ vers $\mathrm{F}(\mathrm{E},\{0;1\})$ qui à toute partie de E associe sa fonction caractéristique, est une bijection. Or l'ensemble $\mathcal{F}(\mathrm{E},\{0;1\})$ est fini de cardinal 2^n ce qui donne le résultat. □

Remarque 24.4 – *Le théorème justifie le raisonnement suivant : pour construire une partie de E il y a deux choix possibles pour chaque élément de E (on le prend ou on ne le prend pas), comme il y a n éléments dans E cela fait 2^n constructions possibles, soit 2^n parties.*

4) Le nombre de bijections

Théorème 24.13
*Si E et de F sont deux ensembles finis **de même cardinal** $n > 0$, il y a $n!$ bijections de E vers F. En particulier, $\mathrm{card}(\mathcal{S}(\mathrm{E})) = n!$ (groupe des permutations de E).*

Preuve : Lorsque $\mathrm{card(E)} = \mathrm{card(F)} = n$, l'ensemble des bijections de E vers F est inclus dans l'ensemble des applications, c'est donc un ensemble fini de cardinal inférieur ou égal à n^n. On montre ensuite la formule par récurrence sur n, le résultat étant immédiat pour $n = 1$, supposons le vrai au rang $n-1$. Posons $\mathrm{F} = \{d_1,\dots,d_n\}$, soit $e \in \mathrm{E}$ fixé, on pose $\mathrm{D}_k = \{f \in \mathrm{Bij(E,F)}\ /\ f(e) = d_k\}$ pour $k \in [\![1\,;n]\!]$. Il est clair que $\mathrm{Bij(E,F)} = \mathrm{D}_1 \cup \ldots \cup \mathrm{D}_n$ et que $\mathrm{card}(\mathrm{D}_k) = \mathrm{card}(\mathrm{Bij}(\mathrm{E}\setminus\{e\},\mathrm{F}\setminus\{d_k\})$, on obtient ainsi que $\mathrm{card(Bij(E,F))} = n \times (n-1)! = n!$. □

5) Le nombre de p-parties (ou p-combinaisons)

Définition 24.5
Soit E un ensemble de cardinal $n \in \mathbb{N}$ et soit $p \in \mathbb{N}$, on appelle p - combinaison d'éléments de E (ou p - partie) toute partie de E de cardinal p. L'ensemble des p - parties de E est noté $\mathcal{P}_p(\mathrm{E})$.

Remarque 24.5 – $\mathcal{P}_p(\mathrm{E})$ *est un ensemble fini car il est inclus dans* $\mathcal{P}(\mathrm{E})$, *et son cardinal est majoré par* 2^n.

Cas particuliers :

a) Si $p = 0$ la seule partie de E à 0 élément est $\varnothing$, donc $\mathrm{card}(\mathcal{P}_0(\mathrm{E})) = 1$.

b) Si $p = n$, la seule partie de E à n éléments est E, donc $\mathrm{card}(\mathcal{P}_n(\mathrm{E})) = 1$.

c) Si $p > n$ il n'y a aucune partie de E à p éléments donc dans ce cas, $\mathrm{card}(\mathcal{P}_p(\mathrm{E})) = 0$.

Théorème 24.14
Si $n, p \in \mathbb{N}$, alors $\mathrm{card}(\mathcal{P}_p(\mathrm{E})) = \frac{\prod_{k=0}^{p-1}(n-k)}{p!} = \binom{n}{p}$ (avec la convention que le produit vaut 1 lorsque $p = 0$, et que $\binom{n}{p} = 0$ si $p > n$).

Preuve : Par récurrence sur n : pour $n = 0$ et $n = 1$, la vérification est immédiate.

Supposons le théorème vrai pour un entier $n \geqslant 1$ et supposons $\mathrm{card(E)} = n+1$, si $p = 0$ la formule est vraie, supposons $p \geqslant 1$, on fixe un élément $a \in \mathrm{E}$, soit A l'ensemble des p - parties de E contenant a et B l'ensemble des p - parties de E ne contenant pas a, alors $\mathcal{P}_p(\mathrm{E}) = \mathrm{A} \cup \mathrm{B}$ et $\mathrm{A} \cap \mathrm{B} = \varnothing$, donc $\mathrm{card}(\mathcal{P}_p(\mathrm{E})) = \mathrm{card(A)} + \mathrm{card(B)}$, or $\mathrm{card(B)} = \frac{\prod_{k=0}^{p-1}(n-k)}{p!}$ (car B est en bijection avec $\mathcal{P}_p(\mathrm{E}\setminus\{a\})$) et $\mathrm{card(A)} = \frac{\prod_{k=0}^{p-2}(n-k)}{(p-1)!}$ (car A est en bijection avec $\mathcal{P}_{p-1}(\mathrm{E}\setminus\{a\})$), d'où :

$$\begin{aligned}\mathrm{card}(\mathcal{P}_p(\mathrm{E})) &= \frac{n\times\cdots\times(n-p+1)}{p!} + \frac{n\times\cdots\times(n-p+2)}{(p-1)!}\\ &= \frac{n\times\cdots(n-p+2)[n-p+1+p]}{p!}\\ &= \frac{(n+1)n\times\cdots(n+1-p+1)}{p!}\end{aligned}$$

la formule est donc vraie au rang $n+1$. □

★**Exercice 24.2** *À l'aide d'un raisonnement de dénombrement, retrouver sans calcul les propriétés suivantes :*

1/ *Si* $p \leqslant n$, $\binom{n}{p} = \binom{n}{n-p}$.

2/ *Si* $0 \leqslant p \leqslant n-1$, $\binom{n}{p} + \binom{n}{p+1} = \binom{n+1}{p+1}$.

3/ *Binôme de Newton :* $\forall\, n \in \mathbb{N}, \forall\, x, y \in \mathbb{C}, (x+y)^n = \sum\limits_{k=0}^{n} \binom{n}{k} x^k y^{n-k} = \sum\limits_{k=0}^{n} \binom{n}{k} x^{n-k} y^k$.

4/ *Si* $1 \leqslant p \leqslant n$, $\binom{n}{p} = \dfrac{n}{p}\binom{n-1}{p-1}$.

III SOLUTION DES EXERCICES

Solution 24.1 *Soit a un élément non nul de* A*, l'application $f\colon \mathrm{A} \to \mathrm{A}$ définie par $f(x) = a \times x$ est injective (car* A *est intègre), or* A *est fini, donc f est bijective, par conséquent il existe $a' \in \mathrm{A}$ tel que $f(a') = 1$ i.e. $a \times a' = 1$. De même (en considérant $g\colon x \mapsto x \times a$) il existe $a'' \in \mathrm{A}$ tel que $a'' \times a = 1$, mais alors $a'' = a'' \times (a \times a') = (a'' \times a) \times a' = a'$. Finalement, tout élément non nul de* A *possède un inverse et donc* A *est un corps.*

Solution 24.2

1/ *L'application de $f\colon \mathcal{P}_p(\mathrm{E}) \to \mathcal{P}_{n-p}(\mathrm{E})$ définie par $f(\mathrm{A}) = \overline{\mathrm{A}}$ (complémentaire de* A *dans* E*) est bijective.*

2/ *Pour compter le nombre de $(p+1)$-parties de $[\![1\,;n+1]\!]$, on compte celles qui contiennent $n+1$ (il y en a $\binom{n}{p}$) et celles qui ne contiennent pas $n+1$ (il y en a $\binom{n}{p+1}$).*

3/ *Lorsqu'on développe $(x+y)^n = (x+y) \times \cdots \times (x+y)$ on obtient une somme de termes $f_1 \times \cdots \times f_n$ où f_i provient du facteur numéro i, on a $f_i = x$ ou y, par conséquent on a une somme de termes du type $x^k y^{n-k}$ avec $k \in [\![0\,;n]\!]$, et chacun de ces termes est obtenu $\binom{n}{k}$ fois (k facteurs parmi n égaux à x et les autres égaux à y).*

4/ *Considérons un ensemble constitué de p boules rouges et $n-p$ bleues. Le nombre de façons de ranger ces n boules dans n boites (une par boite) est : $\binom{n}{p}$ (p boites parmi n pour les rouges, celles qui restent sont pour les bleues), imaginons que pour chacun de ces rangements on peint une des boules rouges en blanc, on obtient alors $p\binom{n}{p}$ façons de ranger n boules dont 1 blanche, $p-1$ rouges et $n-p$ bleues, c'est à dire $n\binom{n-1}{p-1}$ (une boite pour la blanche, $p-1$ pour les rouges, celles qui restent pour les bleues).*

Chapitre 25

Déterminants

Sommaire

I LE GROUPE SYMÉTRIQUE

1) Décomposition des permutations

Définition 25.1

*Soit $E_n = [\![1\,;n]\!]$, avec $n \in \mathbb{N}^*$, on note $\mathfrak{S}_n$ l'ensemble des permutations de E_n, c'est à dire l'ensemble des bijections de E_n vers lui-même. On rappelle que $(\mathfrak{S}_n, \circ)$ est un groupe de cardinal $n!$ (non abélien dès que $n \geqslant 3$), ce groupe est appelé **groupe symétrique de type** n.*

Représentation d'une permutation :

On adoptera la notation suivante : soit $\sigma \in \mathfrak{S}_n, \sigma = \begin{pmatrix} 1 & \cdots & n \\ \sigma(1) & \cdots & \sigma(n) \end{pmatrix}$, la deuxième ligne étant les images de la première par σ (dans le même ordre).

Définition 25.2 (notion de cycle)

Soit $\sigma \in \mathfrak{S}_n$, on dit que σ est un ***cycle*** *lorsqu'il existe k entiers distincts dans $[\![1\,;n]\!]$: $i_1, \ldots, i_k$ ($k > 1$) tels que $\sigma(i_1) = i_2, \ldots, \sigma(i_{k-1}) = i_k$, $\sigma(i_k) = i_1$, les autres entiers de E_n* ***étant fixes*** *par σ. Dans ce cas, on notera $\sigma = (i_1\ \ i_2 \ldots i_k)$. Le nombre k est appelé* ***longueur du cycle****, et un cycle de longueur 2 est appelé* ***une transposition****. L'ensemble $\{i_1, \ldots, i_k\}$ est appelé* ***support du cycle*** *σ. On convient que l'identité de E_n est un cycle de longueur nulle et à support vide.*

★**Exercice** 25.1

1/ Décrire tous les cycles de $\mathfrak{S}_3$.

2/ Combien y-a-t'il de cycles de longueur k dans $\mathfrak{S}_n$?

3/ Si σ est un cycle dans $\mathfrak{S}_n$, déterminer σ^{-1}.

4/ Si σ est un cycle de longueur p dans $\mathfrak{S}_n$, montrer que $\sigma^p = \mathrm{id}$ et que $\sigma^i \neq \mathrm{id}$ si $i < p$. En déduire que si a est un élément du support de σ, alors $\sigma = (a\ \ \sigma(a) \ldots \sigma^{p-1}(a))$.

Théorème 25.1 (admis)

Toute permutation de E_n est un produit (pour la loi $\circ$) de cycles ***à supports disjoints****, cette décomposition est unique à l'ordre près.*

☞**Exemple** : Soit $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 2 & 1 & 6 & 4 & 5 \end{pmatrix}$, alors $\sigma = (1\ 3) \circ (4\ 6\ 5)$.

★**Exercice** 25.2 *Montrer que deux cycles à supports disjoints commutent. Et si les supports sont non disjoints ?*

Théorème 25.2

Toute permutation est un produit de transpositions.

Preuve : Il suffit de le prouver pour un cycle, soit $\sigma = (i_1\ \ldots\ i_k)$, il est facile de vérifier que $\sigma = (i_1\, i_2)(i_2\, i_3)\cdots(i_{k-1}\, i_k)$. On remarquera que les supports des transpositions ne sont pas disjoints deux à deux, et que le nombre de transpositions dans la décomposition est égal à $k-1$. □

2) Signature

Définition 25.3 (inversion)

Soit $\sigma \in \mathfrak{S}_n$, on appelle inversion de σ tout couple d'entiers $(i,j) \in [\![1\,;n]\!]^2$ tel que $i < j$ et $\sigma(i) > \sigma(j)$. Le nombre d'inversions de σ est noté I_σ.

Théorème 25.3 (Nombre d'inversions d'une transposition)

Soit $\sigma = (a\ b)$ une transposition de $\mathfrak{S}_n$ avec $a < b$, alors $\mathrm{I}_\sigma = 2(b-a)-1$.

Preuve : En effet, soit $i < j$ deux entiers de E_n :
- Si $i, j \notin \{a, b\}$, alors i et j sont fixes, donc (i,j) n'est pas une inversion de σ.
- Si $i = a$ et $j = b$, alors (i,j) est une inversion.
- Si $i = a$ et $j \neq b$, alors (i,j) est une inversion ssi $j \in [\![a+1\,;b-1]\!]$.
- Si $i \neq a$ et $j = b$, alors (i,j) est une inversion ssi $i \in [\![a+1\,;b-1]\!]$.

Ce qui nous fait au total $2(b-a)-1$ inversions pour la transposition $(a\ b)$. □

Définition 25.4 (signature d'une permutation)

Soit $\sigma \in \mathfrak{S}_n$, on appelle ***signature*** *de σ le nombre noté $\varepsilon(\sigma)$ et défini par $\varepsilon(\sigma) = (-1)^{\mathrm{I}_\sigma}$.*

On peut vérifier que $\varepsilon(\sigma) = \prod\limits_{1 \leqslant i < j \leqslant n} \frac{\sigma(i)-\sigma(j)}{i-j}$.

☞**Exemple** : Signature d'une transposition, si $\sigma = (a\ b)$ avec $a < b \in \mathrm{E}_n$, alors $\mathrm{I}_\sigma = 2(b-a)-1$, donc $\varepsilon(\sigma) = -1$.

Théorème 25.4 (admis)
L'application signature $\varepsilon\colon (\mathfrak{S}_n, \circ) \to (\{\pm 1\}, \times)$*, est un morphisme de groupes, c'est à dire :*
$$\forall\ \sigma, \sigma' \in \mathfrak{S}_n, \varepsilon(\sigma \circ \sigma') = \varepsilon(\sigma) \times \varepsilon(\sigma').$$

Application – Calcul de la signature d'une permutation. Pour calculer la signature d'une permutation σ, il suffit de connaître sa décomposition en produit de transpositions : $\sigma = \tau_1 \circ \ldots \circ \tau_k$, on a alors $\varepsilon(\sigma) = \varepsilon(\tau_1) \times \ldots \varepsilon(\tau_k) = (-1)^k$.

☞**Exemples** :

- Calculer la signature d'un cycle de longueur $p \geqslant 1$.
- Calculer la signature de $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 3 & 10 & 4 & 6 & 9 & 7 & 1 & 5 & 8 & 2 \end{pmatrix}$.

II APPLICATIONS N-LINÉAIRES

1) Définition

Définition 25.5
Soient E *et* F *deux* $\mathbb{K}$*-espaces vectoriels et* $f : \mathrm{E}^n \to \mathrm{F}$ *une application. On dit que* f *est* n*-linéaire lorsque* f *est* ***linéaire par rapport à chacune de ses variables*** *(les autres étant fixes), c'est à dire :* $\forall\ x_1, \ldots, x_n \in \mathrm{E}, \forall\ i \in [\![1\,; n]\!]$*, l'application* $f_i\colon \mathrm{E} \to \mathrm{F}$ *définie par :*
$$f_i(x) = f(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n) \text{ est linéaire.}$$

☞**Exemples** :

- L'application $f\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ définie par $f(u, v) = u_1v_1 + u_2v_2$ (avec $u = (u_1, u_2,)$ et $v = (v_1, v_2)$), est bilinéaire, plus précisément, c'est une forme bilinéaire sur $\mathbb{R}^2$.
- L'application $f\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ définie par $f(u, v) = u_1v_2 - u_2v_1$ (avec $u = (u_1, u_2,)$ et $v = (v_1, v_2)$), est une forme bilinéaire sur $\mathbb{R}^2$.
- Soit $\mathrm{E} = \mathcal{M}_n(\mathbb{K})$, l'application $f : \mathrm{E}^3 \to \mathrm{E}$ définie par $f(\mathrm{A}, \mathrm{B}, \mathrm{C}) = \mathrm{A} \times \mathrm{B} \times \mathrm{C}$, est trilinéaire.
- Soit $\mathrm{E} = \mathcal{C}^0([a; b], \mathbb{R})$. L'application $f : \mathrm{E}^2 \to \mathbb{R}$ définie par $f(u, v) = \int_a^b u(t)v(t)\,dt$, est une forme bilinéaire sur E.

Définition 25.6
Soit $f : \mathrm{E}^n \to \mathrm{F}$ *une application* n*-linéaire, on dit que* f *est :*

- ***symétrique*** *: lorsque* $\forall\ \sigma \in \mathfrak{S}_n, \forall\ x_1, \ldots, x_n \in \mathrm{E}, f(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) = f(x_1, \ldots, x_n)$*, autrement dit, changer l'ordre des vecteurs ne change pas le résultat.*
- ***antisymétrique*** *: lorsque* $\forall\ \sigma \in \mathfrak{S}_n, \forall\ x_1, \ldots, x_n \in \mathrm{E}$*,* $f(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) = \varepsilon(\sigma).f(x_1, \ldots, x_n)$*, autrement dit, échanger deux vecteurs change le signe du résultat.*
- ***alternée*** *: lorsque* $\forall\ x_1, \ldots, x_n \in \mathrm{E}$*, s'il existe* $i, j \in [\![1\,; n]\!]$ *tels que* $i \neq j$ *et* $x_i = x_j$*, alors* $f(x_1, \ldots, x_n) = 0$*, autrement dit, si deux des vecteurs sont égaux alors le résultat est nul.*

Notation : soit $f : \mathrm{E}^n \to \mathrm{F}$ une application n-linéaire, pour $\sigma \in \mathfrak{S}_n$ on pose $f_\sigma\colon \mathrm{E}^n \to \mathrm{F}$ définie par $f_\sigma(x_1, \ldots, x_n) = f(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$. On remarquera que f_σ est également n-linéaire.

☞**Exemple** : En reprenant les exemples ci-dessus, la première application et la quatrième sont symétriques, la deuxième est antisymétrique et alternée, la troisième est ni symétrique, ni antisymétrique, ni alternée.

Théorème 25.5
Il y a équivalence entre antisymétrique et alternée.

Preuve : Si $f : \mathrm{E}^n \to \mathrm{F}$ est n-linéaire alternée, soient $x_1, \ldots, x_n \in \mathrm{E}$, soit $\tau = (i\ j)$ avec $i, j \in [\![1\,; n]\!]$, posons $x_i' = x_j' = x_i + x_j$, et $\mathrm{S} = f(x_1, \ldots, x_{i-1}, x_i', x_{i+1}, \ldots, x_{j-1}, x_j', x_{j+1}, \ldots, x_n)$, comme f est alternée, $\mathrm{S} = 0$, mais $\mathrm{S} = f(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_{j-1}, x_j, x_{j+1}, \ldots, x_n)+$
$f(x_1, \ldots, x_{i-1}, x_j, x_{i+1}, \ldots, x_{j-1}, x_i, x_{j+1}, \ldots, x_n)$ (en développant par rapport à la i^{e} variable puis par rapport à la j^{e}), ce qui donne $f_\tau = -f = \varepsilon(\tau).f$, mais comme toute permutation est un produit de transpositions, on en déduit que $\forall\ \sigma \in \mathfrak{S}_n$, $f_\sigma = \varepsilon(\sigma).f$ c'est à dire f est antisymétrique.

Si f est antisymétrique : supposons $x_i = x_j$ avec $i \neq j$, alors soit $\tau = (i\ j)$, on a $f(x_1, \ldots, x_n) = f_\tau(x_1, \ldots, x_n) = -f(x_1, \ldots, x_n)$, donc $f(x_1, \ldots, x_n) = 0$, *i.e.* f est alternée. □

Théorème 25.6
L'ensemble des applications n-linéaires de E^n vers F est un $\mathbb{K}$-espace vectoriel (s.e.v. de $\mathcal{F}(E^n, F)$).
L'ensemble des applications n-linéaires alternées de E^n vers F est un $\mathbb{K}$-espace vectoriel (s.e.v. de $\mathcal{F}(E^n, F)$).

Preuve : Celle-ci est simple et laissée en exercice. □

2) Développement suivant une base

Soit E de dimension p, soit $\mathfrak{B} = (e_1, \dots, e_p)$ une base de E, soit $S = (x_1, \dots, x_n)$ une famille de n vecteurs de E et soit $A = \mathcal{P}_{\mathfrak{B},S} \in \mathcal{M}_{p,n}(\mathbb{K})$ la matrice de la famille S dans la base $\mathfrak{B}$. Soit $u = f(x_1, \dots, x_n)$, en développant par rapport à la première variable, on obtient :

$$u = \sum_{1 \leqslant j_1 \leqslant p} a_{j_1,1} f(e_{j_1}, x_2, \dots, x_n)$$

Puis on développe chacun de ces termes par rapport à la deuxième variable x_2, ce qui donne

$$u = \sum_{1 \leqslant j_1, j_2 \leqslant p} a_{j_1,1} a_{j_2,2} f(e_{j_1}, e_{j_2}, x_3, \dots, x_n)$$

etc..., ce qui donne à la fin :

$$f(x_1, \dots, x_n) = \sum_{1 \leqslant j_1, \dots, j_n \leqslant p} a_{j_1,1} \cdots a_{j_n,n} f(e_{j_1}, \dots, e_{j_n})$$

C'est le développement de $f(x_1, \dots, x_n)$ dans la base $\mathfrak{B}$, on voit ainsi qu'une application n-linéaire $f\colon E^n \to F$ est entièrement déterminée par la donnée des vecteurs $f(e_{j_1}, \dots, e_{j_n})$.

III DÉTERMINANT DANS UNE BASE

1) Formes n-linéaires en dimension n

Avec les notations précédentes, supposons que la dimension de E soit égale à n, soit $\mathfrak{B} = (e_1, \dots, e_n)$ une base de E, et f une forme n-linéaire alternée sur E, d'après ce qui précède, on a :

$$f(x_1, \dots, x_n) = \sum_{1 \leqslant j_1, \dots, j_n \leqslant n} a_{j_1,1} \cdots a_{j_n,n} f(e_{j_1}, \dots, e_{j_n})$$

comme f est alternée, dès que deux indices sont égaux, le terme correspondant est nul, il reste donc :

$$f(x_1, \dots, x_n) = \sum_{1 \leqslant j_1, \dots, j_n \leqslant n, \text{distincts}} a_{j_1,1} \cdots a_{j_n,n} f(e_{j_1}, \dots, e_{j_n})$$

Lorsque les indices sont distincts deux à deux, on a forcément, $\{j_1, \dots, j_n\} = [\![1\,;n]\!]$, en posant $\sigma(k) = j_k$ on définit une permutation des entiers de 1 à n et on peut alors écrire :

$$f(x_1, \dots, x_n) = \sum_{\sigma \in \mathfrak{S}_n} a_{\sigma(1),1} \cdots a_{\sigma(n),n} f(e_{\sigma(1)}, \dots, e_{\sigma(n)})$$

Or, $f(e_{\sigma(1)}, \dots, e_{\sigma(n)}) = \varepsilon(\sigma).f(e_1, \dots, e_n)$ (f est antisymétrique), finalement on a la formule :

$$f(x_1, \dots, x_n) = \left(\sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n} \right) f(e_1, \dots, e_n)$$

Donc si f est une forme n-linéaire alternée sur E, alors il existe un scalaire α tel que :

$$\forall\, x_1, \dots, x_n \in E, f(x_1, \dots, x_n) = \alpha \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1),1} a_{\sigma(2),2} \dots a_{\sigma(n),n}.$$

Avec $\alpha = f(e_1, \dots, e_n)$ (ce qui détermine entièrement f), et les scalaires $a_{i,j}$ étant les coefficients de la matrice de la famille $(x_1, \dots, x_n)$ dans la base $\mathfrak{B}$.

Si σ parcourt $\mathfrak{S}_n$, alors σ^{-1} aussi, on peut donc écrire :

$$\begin{aligned}\forall x_1, \dots, x_n \in \mathrm{E}, f(x_1, \dots, x_n) &= \alpha \sum_{\sigma^{-1} \in \mathfrak{S}_n} \varepsilon(\sigma) a_{1,\sigma^{-1}(1)} a_{2,\sigma^{-1}(2)} \dots a_{n,\sigma^{-1}(n)} \\ &= \alpha \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{1,\sigma(1)} \dots a_{n,\sigma(n)}.\end{aligned}$$

Théorème 25.7

On pose pour $x_1, \dots, x_n \in \mathrm{E}$, $f_1(x_1, \dots, x_n) = \sum\limits_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{1,\sigma(1)} \dots a_{n,\sigma(n)}$, alors f_1 est une forme n-linéaire alternée sur E et $f_1(e_1, \dots, e_n) = 1$ (en particulier f_1 est non nulle).

Preuve : On note c_i la forme linéaire qui à tout vecteur x de E associe sa coordonnée sur e_i.

On note alors $f(x_1, \dots, x_n) = c_1(x_1)c_2(x_2)\cdots c_n(x_n)$, on vérifie que f ainsi définie, est n-linéaire de E^n vers $\mathbb{K}$. On pose $g(x_1, \dots, x_n) = \sum\limits_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) f(x_{\sigma(1)}, \dots, x_{\sigma(n)})$, on sait que g est n-linéaire (combinaison linéaire de formes n-linéaires), soit τ une transposition de E_n, alors $g(x_{\tau(1)}, \dots, x_{\tau(n)}) = \sum\limits_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) f(x_{\tau(\sigma(1))}, \dots, x_{\tau(\sigma(n))})$, on effectue un changement d'indice en posant $\sigma' = \tau \circ \sigma$ (σ' parcourt $\mathfrak{S}_n$ une et seule fois lorsque σ parcourt $\mathfrak{S}_n$), on a $\varepsilon(\sigma') = -\varepsilon(\sigma)$, d'où $g(x_{\tau(1)}, \dots, x_{\tau(n)}) = -\sum\limits_{\sigma' \in \mathfrak{S}_n} \varepsilon(\sigma') f(x_{\sigma'(1)}, \dots, x_{\sigma'(n)}) = -g(x_1, \dots, x_n)$, donc g est alternée.

En revenant à la définition de f, et en posant $x_j = \sum\limits_{i=1}^{n} a_{ij} e_i$, on a :

$$g(x_1, \dots, x_n) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{1\sigma(1)} \times \dots \times a_{n\sigma(n)} = f_1(x_1, \dots, x_n)$$

Lorsqu'on calcule $f_1(e_1, \dots, e_n)$, tous les a_{ij} sont nuls sauf ceux de la diagonale qui valent 1, il reste seulement le terme correspondant à l'identité de E_n, ce qui donne $f_1(e_1, \dots, e_n) = a_{11} \times \dots \times a_{nn} = 1$. □

Théorème 25.8

Si $\dim(\mathrm{E}) = n$, alors l'ensemble des formes n-linéaires alternées sur E est un $\mathbb{K}$-espace vectoriel de dimension 1 (droite vectorielle), dont une base est l'application f_1 ci-dessus.

Preuve : D'après le début de ce paragraphe, toute forme f, n-linéaire alternée sur E, est du type $f = \alpha f_1$ avec $\alpha \in \mathbb{K}$. □

2) Déterminants de n vecteurs dans une base

Définition 25.7

*L'application f_1 définie ci-dessus est appelée **déterminant dans la base** $\mathfrak{B}$, elle est notée $\det_{\mathfrak{B}}$. C'est donc une forme n-linéaire alternée sur E, qui constitue une base de l'ensemble des formes n-linéaires alternées sur E, plus précisément, c'est l'unique forme n-linéaire alternée sur E qui vérifie :* $\det_{\mathfrak{B}}(e_1, \dots, e_n) = 1$.

Expression du déterminant : soit $\mathfrak{B} = (e_1, \dots, e_n)$ une base de E et soit $(x_1, \dots, x_n)$ une famille de n vecteurs de E, et soit $\mathrm{A} = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$ la matrice de cette famille dans la base $\mathfrak{B}$, on a l'expression suivante :

$$\det_{\mathfrak{B}}(x_1, \dots, x_n) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{1,\sigma(1)} \dots a_{n,\sigma(n)}$$

☞**Exemples** :

- Soit $\mathrm{E} = \mathbb{K}^2$, $\mathfrak{B} = (i, j)$, la base canonique, soient $u = (u_1, u_2)$, $v = (v_1, v_2) \in \mathrm{E}$, alors $\det_{\mathfrak{B}}(u, v) = a_{1,1}a_{2,2} - a_{1,2}a_{2,1} = u_1v_2 - u_2v_1$.
- Soit $\mathrm{E} = \mathbb{K}^3$, $\mathfrak{B} = (i, j, k)$, la base canonique, soient $u = (u_1, u_2, u_3)$, $v = (v_1, v_2, v_3)$ et $w = (w_1, w_2, w_3) \in \mathrm{E}$, alors $\det_{\mathfrak{B}}(u, v, w) = u_1v_2w_3 - u_2v_1w_3 - u_3v_2w_1 - u_1v_3w_2 + u_3v_1w_2 + u_2v_3w_1$.
- Si $\dim(\mathrm{E}) = n$, $\mathfrak{B} = (e_1, \dots, e_n)$ une base de E, $(x_1, \dots, x_n)$ une famille de E telle que la matrice A du système dans la base $\mathfrak{B}$ soit triangulaire supérieure, alors : $\det_{\mathfrak{B}}(x_1, \dots, x_n) = a_{1,1} \dots a_{n,n}$, produit des coefficients diagonaux de la matrice.
  En effet, pour que $a_{1,\sigma(1)} \dots a_{n,\sigma(n)}$ ait une chance d'être non nul, il faut que $1 \leqslant \sigma(1), \dots, n \leqslant \sigma(n)$, ce qui ne donne qu'une seule possibilité : $\sigma = \mathrm{id}$.

3) Propriétés du déterminant

Soit E un $\mathbb{K}$-espace vectoriel de dimension n et $\mathcal{B} = (e_1, \dots, e_n)$ une base de E.

– L'application $\boxed{\det_{\mathcal{B}} \text{ est une forme } n\text{-linéaire alternée}}$, donc :
  - elle est linéaire par rapport à chaque variable, par exemple : $\det_{\mathcal{B}}(\alpha x_1 + \beta x_1', x_2, \dots, x_n) = \alpha \det_{\mathcal{B}}(x_1, x_2, \dots, x_n) + \beta \det_{\mathcal{B}}(x_1', x_2, \dots, x_n)$.
  - si on échange deux vecteurs dans le déterminant, alors le résultat change de signe. Par exemple : $\det_{\mathcal{B}}(x_2, x_1, x_3, \dots, x_n) = -\det_{\mathcal{B}}(x_1, x_2, x_3, \dots, x_n)$. Plus généralement :
$$\forall\, \sigma \in \mathfrak{S}_n, \det_{\mathcal{B}}(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = \varepsilon(\sigma) \det_{\mathcal{B}}(x_1, \dots, x_n).$$
  - Si deux des vecteurs de la famille $(x_1, \dots, x_n)$ sont égaux, alors $\det_{\mathcal{B}}(x_1, \dots, x_n) = 0$. On en déduit si l'un des vecteurs de la famille est combinaison linéaire des autres (*i.e.* la famille est liée), alors le déterminant est nul, on en déduit également qu'ajouter à l'un des vecteurs une combinaison linéaire des autres, ne change pas le déterminant, par conséquent on peut calculer un déterminant avec la méthode de Gauss (pour obtenir un système de vecteurs dont la matrice dans la base $\mathcal{B}$ est triangulaire supérieure).

– $\boxed{\det_{\mathcal{B}}(e_1, \dots, e_n) = 1}$, ce que l'on note $\det_{\mathcal{B}}(\mathcal{B}) = 1$, c'est l'unique forme n-linéaire alternée sur E qui vérifie cette égalité.

– Si $\mathcal{B}'$ est une autre base de E, alors $\boxed{\det_{\mathcal{B}'} = \det_{\mathcal{B}'}(\mathcal{B}). \det_{\mathcal{B}}}$. On en déduit que $\det_{\mathcal{B}'}(\mathcal{B}) \det_{\mathcal{B}}(\mathcal{B}') = 1$.

  Preuve : L'ensemble des formes n-linéaires alternées est une droite vectorielle engendrée par $\det_{\mathcal{B}}$, donc il existe un scalaire α tel que $\det_{\mathcal{B}'} = \alpha. \det_{\mathcal{B}}$, pour obtenir α, il suffit d'appliquer cette égalité sur les vecteurs de $\mathcal{B}$. □

– $\boxed{\text{La famille } (x_1, \dots, x_n) \text{ est une base de E ssi } \det_{\mathcal{B}}(x_1, \dots, x_n) \neq 0}$.

  Preuve : Soit $\mathcal{B}' = (x_1, \dots, x_n)$, si $\mathcal{B}'$ est une base de E, alors l'ensemble des formes n-linéaires alternées sur E est une droite engendrée par $\det_{\mathcal{B}'}$ et $\det_{\mathcal{B}} = \det_{\mathcal{B}}(\mathcal{B}'). \det_{\mathcal{B}'}$, or $\det_{\mathcal{B}}$ n'est pas l'application nulle, donc $\det_{\mathcal{B}}(\mathcal{B}') \neq 0$.
  Réciproquement, si $\det_{\mathcal{B}}(\mathcal{B}') \neq 0$, alors la famille $\mathcal{B}'$ est libre, car si elle était liée, l'un de ses vecteurs serait combinaison linéaire des autres et donc le déterminant dans la base $\mathcal{B}$ serait nul, ce qui est absurde. Donc $\mathcal{B}'$ est une famille libre de n vecteurs en dimension n, c'est donc une base de E. □

IV DÉTERMINANT D'UN ENDOMORPHISME

1) Définition

Soit E un $\mathbb{K}$-espace vectoriel, soient $\mathcal{B} = (e_1, \dots, e_n)$ et $\mathcal{B}' = (e_1', \dots, e_n')$ deux bases de E, et soit $u \in \mathcal{L}(E)$, on note f la forme n-linéaire alternée définie par :
$$f(x_1, \dots, x_n) = \det_{\mathcal{B}}(u(x_1), \dots, u(x_n))$$

On sait qu'il existe un scalaire α tel que $f = \alpha \det_{\mathcal{B}'}$, avec $\alpha = f(e_1', \dots, e_n')$, or $\det_{\mathcal{B}} = \det_{\mathcal{B}}(\mathcal{B}') \det_{\mathcal{B}'}$, d'où $\alpha = \det_{\mathcal{B}}(\mathcal{B}') \det_{\mathcal{B}'}(u(e_1'), \dots, u(e_n'))$. En calculant $f(e_1, \dots, e_n)$, on obtient :
$$\boxed{\det_{\mathcal{B}}(u(e_1), \dots, u(e_n)) = \alpha \det_{\mathcal{B}'}(\mathcal{B}) = \det_{\mathcal{B}'}(u(e_1'), \dots, u(e_n'))}$$
car $\det_{\mathcal{B}'}(\mathcal{B}) \det_{\mathcal{B}}(\mathcal{B}') = 1$.

> **Définition 25.8**
> *Soit E un $\mathbb{K}$-espace vectoriel et soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E, soit $u \in \mathcal{L}(E)$, on appelle déterminant de u le scalaire, noté $\det(u)$ et défini par : $\det(u) = \det_{\mathcal{B}}(u(e_1), \dots, u(e_n))$. Ce scalaire est **indépendant de la base $\mathcal{B}$ choisie**.*

☞**Exemple** : Soit $u \in \mathcal{L}(\mathbb{R}^2)$ défini par $u(x, y) = (x + y; 2x - y)$, notons $\mathcal{B} = (i, j)$ la base canonique de $\mathbb{R}^2$, alors $\det(u) = \det_{\mathcal{B}}(i + 2j, i - j) = \det_{\mathcal{B}}(3i, i - j) = \det_{\mathcal{B}}(3i, -j) = -3 \det_{\mathcal{B}}(i, j) = -3$.

Expression du déterminant : soit $u \in \mathcal{L}(E)$ et soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E, la matrice du système $(u(e_1), \dots, u(e_n))$ est en fait la matrice de u dans la base $\mathcal{B}$, posons $A = (a_{i,j}) = \underset{\mathcal{B}}{\mathrm{mat}}(u)$, alors on peut écrire :
$$\det(u) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}$$

2) Propriétés du déterminant

– Soit $u \in \mathcal{L}(\mathrm{E})$, soit $\mathfrak{B}$ une base de E, alors :

$$\boxed{\forall\, x_1, \ldots, x_n \in \mathrm{E}, \det_{\mathfrak{B}}(u(x_1), \ldots, u(x_n)) = \det(u) \det_{\mathfrak{B}}(x_1, \ldots, x_n)}$$

Preuve : L'application f définie par $f(x_1, \ldots, x_n) = \det_{\mathfrak{B}}(u(x_1), \ldots, u(x_n))$ est n-linéaire alternée, donc il existe un scalaire α tel que $f = \alpha \det_{\mathfrak{B}}$, ce scalaire vaut $\alpha = f(e_1, \ldots, e_n)$ (en posant $\mathfrak{B} = (e_1, \ldots, e_n)$), donc $\alpha = \det(u)$ d'après la définition. □

– Soient $u, v \in \mathcal{L}(\mathrm{E})$, on a $\boxed{\det(u \circ v) = \det(v \circ u) = \det(u)\det(v)}$.
Preuve : Soit $\mathfrak{B} = (e_1, \ldots, e_n)$ une base de E, alors :

$$\det(u \circ v) = \det_{\mathfrak{B}}(u(v(e_1)), \ldots, u(v(e_n))) = \det(u) \det_{\mathfrak{B}}(v(e_1), \ldots, v(e_n)) = \det(u)\det(v),$$

de même, on obtient $\det(v \circ u) = \det(v)\det(u)$, ce qui donne l'égalité. □

– Soit $u \in \mathcal{L}(\mathrm{E})$ et soit $\mathfrak{B}$ une base de E, alors $\boxed{u \in \mathrm{GL}(\mathrm{E}) \text{ ssi } \det(u) \neq 0}$, et si c'est le cas, alors $\det(u^{-1}) = \frac{1}{\det(u)}$.
Preuve : Si $u \in \mathrm{GL}(\mathrm{E})$, alors $\det(u)\det(u^{-1}) = \det(u \circ u^{-1}) = \det(\mathrm{id}_{\mathrm{E}}) = 1$, donc $\det(u) \neq 0$, et on a la formule.
Réciproquement, si $\det(u) \neq 0$, alors la famille $(u(e_1), \ldots, u(e_n))$ est libre, donc u est bien un automorphisme de E. □

V DÉTERMINANT D'UNE MATRICE CARRÉE

1) Définition

Définition 25.9
Soit $\mathrm{A} = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$, et soit u l'endomorphisme de $\mathbb{K}^n$ canoniquement associé à A, on appelle déterminant de A le déterminant de u et on pose :

$$\det(\mathrm{A}) = \det(u) = \det_{\mathfrak{B}}(\mathrm{C}_1(\mathrm{A}), \ldots, \mathrm{C}_n(\mathrm{A})),$$

où $\mathfrak{B}$ désigne la base canonique de $\mathbb{K}^n$.

Expression du déterminant : A est la matrice de u dans la base canonique de $\mathbb{K}^n$, donc d'après le paragraphe précédent, on a :

$$\det(\mathrm{A}) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}$$

Notation : on écrit $\det(\mathrm{A}) = \begin{vmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{vmatrix}$

☞**Exemples** :

– Cas d'une matrice carrée de taille 2 : $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$.

– Cas d'une matrice triangulaire : si A est triangulaire, alors $\det(\mathrm{A}) = \prod_{i=1}^{n} a_{i,i}$ (produit des éléments diagonaux).

2) Propriétés du déterminant d'une matrice carrée

– Une matrice carrée A et sa transposée, ont le même déterminant. On en déduit que si $\mathfrak{B}$ est la base canonique de $\mathbb{K}^n$, alors $\det(\mathrm{A}) = \det_{\mathfrak{B}}(\mathrm{C}_1(\mathrm{A}), \ldots, \mathrm{C}_n(\mathrm{A})) = \det_{\mathfrak{B}}(\mathrm{L}_1(\mathrm{A}), \ldots, \mathrm{L}_n(\mathrm{A}))$.
Preuve : En effet, on a $\det(\mathrm{A}) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)} = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n} = \det(\mathrm{A}^{\mathrm{T}})$. □

– Soit $u \in \mathcal{L}(\mathrm{E})$, et soit $\mathfrak{B}$ une base quelconque de E, si $\mathrm{mat}_{\mathfrak{B}}(u) = \mathrm{A} \in \mathcal{M}_n(\mathbb{K})$, alors $\det(u) = \det(\mathrm{A})$.
On en déduit en particulier que si $\mathrm{P} \in \mathrm{GL}_n(\mathbb{K})$, alors $\det(\mathrm{P}^{-1} \times \mathrm{A} \times \mathrm{P}) = \det(\mathrm{A})$.
– Si $\mathrm{A}, \mathrm{B} \in \mathcal{M}_n(\mathbb{K})$, alors $\det(\mathrm{A} \times \mathrm{B}) = \det(\mathrm{A})\det(\mathrm{B})$, en particulier, on a $\det(\mathrm{A} \times \mathrm{B}) = \det(\mathrm{B} \times \mathrm{A})$.

– Une matrice carrée est inversible ssi son déterminant est non nul, si c'est le cas, alors $\det(A^{-1}) = \det(A)^{-1}$.
– Utilisation de la méthode de *Gauss* : on utilise celle-ci pour se ramener au calcul du déterminant d'une matrice triangulaire, plus précisément :
  - L'opération $L_i \leftrightarrow L_j$, change le signe du déterminant (idem avec les colonnes).
  - L'opération $L_i \leftarrow \alpha L_i$ ($\alpha \neq 0$), multiplie le déterminant par α (idem avec les colonnes).
  - L'opération $L_i \leftarrow L_i + \alpha L_j$, ne change pas le déterminant (idem avec les colonnes).

☞**Exemple** : Avec $A = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 2 & 0 \\ 2 & 3 & 1 \end{pmatrix}$.

$$\det(A) \overset{C_1 \leftrightarrow C_3}{=} -\begin{vmatrix} 1 & 1 & 1 \\ 0 & 2 & -1 \\ 1 & 3 & 2 \end{vmatrix} \overset{L_3 \leftarrow L_3 - L_1}{=} -\begin{vmatrix} 1 & 1 & 1 \\ 0 & 2 & -1 \\ 0 & 2 & 1 \end{vmatrix} \overset{L_3 \leftarrow L_3 - L_2}{=} -\begin{vmatrix} 1 & 1 & 1 \\ 0 & 2 & -1 \\ 0 & 0 & 2 \end{vmatrix} = -4.$$

★**Exercice 25.3** *Soit* $A \in \mathcal{M}_n(\mathbb{K})$ *et* $\lambda \in \mathbb{K}$, *montrer que* $\det(\lambda A) = \lambda^n \det(A)$.

3) Développement suivant une ligne ou une colonne

Soit $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$, soit $\mathfrak{B} = (e_1, \dots, e_n)$ la base canonique de $\mathbb{K}^n$, notons $C_1, \dots, C_n$ les vecteurs colonnes de A, pour $j \in [\![1 ; n]\!]$, on a $C_j = \sum\limits_{i=1}^{n} a_{i,j} e_i$, par linéarité par rapport à la j^e variable, on peut écrire :

$$\det(A) = \sum_{i=1}^{n} a_{i,j} \det_{\mathfrak{B}}(C_1, \dots, C_{j-1}, e_i, C_{j+1}, \dots, C_n).$$

En posant $\gamma_{i,j}(A) = \det_{\mathfrak{B}}(C_1, \dots, C_{j-1}, e_i, C_{j+1}, \dots, C_n)$, on a alors $\det(A) = \sum\limits_{i=1}^{n} a_{i,j}\gamma_{i,j}(A)$, c'est le développement de $\det(A)$ suivant la colonne j.

Définition 25.10
*Le scalaire $\gamma_{i,j}(A)$ est appelé **cofacteur** $(i\ j)$ de la matrice A, c'est le déterminant de la matrice A dans laquelle la colonne j a été remplacée par le i^e vecteur de la base canonique de $\mathbb{K}^n$.*

Calcul de $\gamma_{i,j}(A)$:

On a $\gamma_{i,j}(A) = \begin{vmatrix} a_{1,1} & \cdots & a_{1,j-1} & 0 & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & 0 & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i,1} & \cdots & a_{i,j-1} & 1 & a_{i,j+1} & \cdots & a_{i,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & 0 & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & 0 & a_{n,j+1} & \cdots & a_{n,n} \end{vmatrix}$, on échange les colonnes C_{j+1} et C_j, puis C_{j+1} et C_{j+2}, ...etc, pour amener la colonne j en dernière position, **sans changer l'ordre sur les autres colonnes**, le déterminant est multiplié par $(-1)^{n-j}$. De le même façon, on amène la ligne i en dernière position sans changer l'ordre des autres lignes, le déterminant est multiplié par $(-1)^{n-i}$. On obtient alors :

$$\gamma_{i,j}(A) = (-1)^{i+j} \begin{vmatrix} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} & 0 \\ \vdots & & \vdots & \vdots & & \vdots & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} & 0 \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} & 0 \\ \vdots & & \vdots & \vdots & & \vdots & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} & 0 \\ a_{i,1} & \cdots & a_{i,j-1} & a_{i,j+1} & \cdots & a_{i,n} & 1 \end{vmatrix},$$

notons $b_{i,j}$ les coefficients de la matrice à l'intérieur de ce déterminant, on a

$$\gamma_{i,j}(A) = (-1)^{i+j} \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) b_{\sigma(1),1} \cdots b_{\sigma(n),n},$$

on remarque que lorsque $\sigma(n) \neq n$, alors $b_{\sigma(n),n} = 0$, on peut donc ne retenir que les permutations ayant n comme point fixe, c'est à dire en fait les éléments de $\mathfrak{S}_{n-1}$, et comme $b_{n,n} = 1$, on a alors :

$$\gamma_{i,j}(A) = (-1)^{i+j} \sum_{\sigma \in \mathfrak{S}_{n-1}} \varepsilon(\sigma) b_{\sigma(1),1} \cdots b_{\sigma(n-1),n-1}.$$

C'est le déterminant de la matrice extraite de A **par suppression de la ligne** i **et de la colonne** j**, multiplié par** $(-1)^{i+j}$**.**

Définition 25.11

*Soit $A \in \mathcal{M}_n(\mathbb{K})$, on appelle **mineur** $(i\ j)$ de la matrice A, le déterminant $\Delta_{i,j}(A)$ de la matrice $A_{i,j}$ obtenue par suppression de ligne i et de la colonne j de A. Le lien entre cofacteur et mineur est : $\gamma_{i,j}(A) = (-1)^{i+j}\Delta_{i,j}(A)$.*

Le développement de det(A) suivant la colonne j s'écrit donc :

$$\det(A) = \sum_{i=1}^{n} a_{i,j}(-1)^{i+j}\Delta_{i,j}(A).$$

$\Delta_{i,j}(A)$ est un déterminant d'ordre $n-1$, c'est donc une formule **récurrente**.

☞**Exemple** : En développant suivant la deuxième colonne :

$$\begin{vmatrix} 1 & 2 & 1 \\ -1 & 0 & 2 \\ 1 & 3 & 1 \end{vmatrix} = -2\begin{vmatrix} -1 & 2 \\ 1 & 1 \end{vmatrix} - 3\begin{vmatrix} 1 & 1 \\ -1 & 2 \end{vmatrix} = (-2)(-3) - 9 = -3.$$

Développement suivant la ligne j : comme une matrice $A \in \mathcal{M}_n(\mathbb{K})$ a le même déterminant que sa transposée, développer det(A) suivant la ligne j revient à développer $\det(A^T)$ suivant la colonne j, ce qui donne $\det(A) = \sum\limits_{i=1}^{n}(A^T)_{i,j}(-1)^{i+j}\Delta_{i,j}(A^T)$, mais $\Delta_{i,j}(A^T) = \Delta_{j,i}(A)$, ce qui donne finalement :

$$\det(A) = \sum_{i=1}^{n} a_{j,i}(-1)^{i+j}\Delta_{j,i}(A)$$

Un exemple classique : le déterminant de Vandermonde [1]

Soit $A = \begin{pmatrix} 1 & \alpha_0 & \cdots & \alpha_0^n \\ \vdots & \vdots & & \vdots \\ 1 & \alpha_n & \cdots & \alpha_n^n \end{pmatrix}$, où $\alpha_0, \ldots, \alpha_n \in \mathbb{K}$, soit $V_n = \det(A)$, alors on a le résultat suivant :

$$V_n = \prod_{0 \leqslant i < j \leqslant n} (\alpha_j - \alpha_i)$$

Preuve : Si deux des scalaires sont égaux, le résultat est évident. Supposons les α_i distincts deux à deux. On peut vérifier que pour $n = 2$ le résultat est vrai. Supposons le vrai au rang n, et remplaçons dans V_{n+1} la dernière ligne par $(1\ X\ X^2 \ldots X^{n+1})$, le déterminant obtenu en développant suivant la dernière ligne, est un polynôme P(X) de degré au plus $n+1$, dont les racines sont $\alpha_0, \ldots, \alpha_n$, de plus son coefficient dominant est V_n, par conséquent on a $P(X) = V_n \prod\limits_{i=0}^{n}(X - \alpha_i)$, d'où $V_{n+1} = V_n \prod\limits_{i=0}^{n}(\alpha_{n+1} - \alpha_i) = \prod\limits_{0 \leqslant i < j \leqslant n+1}(\alpha_j - \alpha_i)$. □

4) Comatrice

Définition 25.12

Soit $A \in \mathcal{M}_n(\mathbb{K})$, on appelle comatrice de A la matrice de $\mathcal{M}_n(\mathbb{K})$ dont les coefficients sont les cofacteurs $\gamma_{i,j}(A)$. Notation $\mathrm{Com}(A) = (\gamma_{i,j}(A))_{1 \leqslant i,j \leqslant n}$.

☞**Exemple** : Si $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, alors $\mathrm{Com}(A) = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$.

1. VANDERMONDE Alexandre (1735 – 1796) : mathématicien français qui fut le premier à étudier les déterminants.

Théorème 25.9 (relation fondamentale)

$$\forall\, A \in \mathcal{M}_n(\mathbb{K}),\ A \times \mathrm{Com}(A)^T = \mathrm{Com}(A)^T \times A = \det(A)I_n$$

Preuve : Soit $B = A \times \mathrm{Com}(A)^T$, alors $b_{i,j} = \sum_{k=1}^{n} a_{i,k}(-1)^{j+k}\Delta_{jk}(A)$, qui est le déterminant de la matrice obtenue en remplaçant dans A, la ligne j par la ligne i, d'où $b_{i,j} = \det(A)\delta_{i,j}$.

Posons $C = \mathrm{Com}(A)^T \times A$, alors $c_{i,j} = \sum_{k=1}^{n} a_{k,j}(-1)^{i+k}\Delta_{k,i}(A)$, qui est le déterminant de la matrice obtenue en remplaçant dans A, la colonne i par la colonne j, on a donc $c_{i,j} = \det(A)\delta_{i,j}$. □

Application – Si $A \in \mathcal{M}_n(\mathbb{K})$ est inversible, alors :

$$A^{-1} = \frac{1}{\det(A)}\mathrm{Com}(A)^T$$

VI APPLICATIONS

1) Équation cartésienne d'un hyperplan dans une base

Soit E un espace de dimension n et $\mathcal{B} = (e_1, \dots, e_n)$ une base de E, soit H un hyperplan de E et $(u_1, \dots, u_{n-1})$ une base de H. On a alors :

$$\begin{aligned}
v\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}_{\mathcal{B}} \in H &\iff (u_1, \dots, u_{n-1}, v) \text{ est liée [car } (u_1, \dots, u_{n-1}) \text{ est libre]} \\
&\iff \det_{\mathcal{B}}(u_1, \dots, u_{n-1}, v) = 0 \qquad \text{(on est en dimension } n) \\
&\iff \begin{vmatrix} a_{11} & \cdots & a_{1n-1} & x_1 \\ \vdots & & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn-1} & x_n \end{vmatrix} = 0 \qquad (a_{i,j} \text{ est la coordonnée sur } e_i \text{ de } u_j) \\
&\iff (-1)^{n+1}\Delta_{1n}x_1 + \cdots + (-1)^{n+n}\Delta_{nn}x_n = 0 \text{ (développement suivant la colonne } n) \\
&\iff \boxed{\alpha_1 x_1 + \cdots + \alpha_n x_n = 0} \quad \text{c'est une équation cartésienne de H dans la base } \mathcal{B},
\end{aligned}$$

en posant $\alpha_i = (-1)^{n+i}\Delta_{i,n}$. Les coefficients α_i ne sont pas tous nuls, en effet, s'ils l'étaient alors pour tout vecteur v de E le déterminant ci-dessus serait nul, or il existe un vecteur v tel que la famille $(u_1, \dots, u_{n-1}, v)$ soit une base de E (théorème de la base incomplète), pour ce vecteur v on a $\det_{\mathcal{B}}(u_1, \dots, u_{n-1}, v) \neq 0$: contradiction, donc au moins un des coefficients α_i est non nul.

2) Orientation d'un espace vectoriel réel

Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases d'un espace vectoriel sur $\mathbb{R}$, de dimension finie, E, soit P la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$, alors P est une matrice inversible, donc $\det(P) \neq 0$, on a alors soit $\det(P) > 0$, soit $\det(P) < 0$ (les coefficients de cette matrice sont des réels).

Définition 25.13

On dit que la base $\mathcal{B}$ est en relation avec la base $\mathcal{B}'$ lorsque la matrice de passage $P_{\mathcal{B},\mathcal{B}'}$ a un déterminant strictement positif, on note alors $\mathcal{B}\mathcal{R}\mathcal{B}'$.

☞**Exemple** : Une base $\mathcal{B} = (e_1, \dots, e_n)$ est en relation avec elle-même (car $P_{\mathcal{B},\mathcal{B}} = I_n$), mais pas avec $\mathcal{B}' = (-e_1, e_2, \dots, e_n)$ (car $\det(P_{\mathcal{B},\mathcal{B}'}) = -\det(I_n)$ par linéarité sur la colonne 1).

Théorème 25.10

*La relation $\mathcal{R}$ est une relation d'équivalence dans l'ensemble des bases de E, et **il n'y a que deux classes d'équivalence**.*

Preuve : La réflexivité est évidente (car $\mathrm{P}_{\mathcal{B},\mathcal{B}} = \mathrm{I}_n$). La symétrie découle de la formule $\mathrm{P}_{\mathcal{B}',\mathcal{B}} = \mathrm{P}_{\mathcal{B},\mathcal{B}'}^{-1}$, donc $\det(\mathrm{P}_{\mathcal{B}',\mathcal{B}}) = \frac{1}{\det(\mathrm{P}_{\mathcal{B},\mathcal{B}'})}$, les deux déterminants ont donc le même signe. La transitivité découle de la formule : $\mathrm{P}_{\mathcal{B},\mathcal{B}''} = \mathrm{P}_{\mathcal{B},\mathcal{B}'} \times \mathrm{P}_{\mathcal{B}',\mathcal{B}''}$, donc $\det(\mathrm{P}_{\mathcal{B},\mathcal{B}''}) = \det(\mathrm{P}_{\mathcal{B},\mathcal{B}'})\det(\mathrm{P}_{\mathcal{B}',\mathcal{B}''})$.

On reprend l'exemple ci-dessus : $\mathcal{B} = (e_1, \dots, e_n)$ une base de E et $\mathcal{B}' = (-e_1, e_2, \dots, e_n)$, $\mathcal{B}$ et $\mathcal{B}'$ ne sont pas en relation, donc il y a au moins deux classes d'équivalence (celle de $\mathcal{B}$ et celle de $\mathcal{B}'$), soit $\mathcal{B}''$ une troisième base, supposons que $\mathcal{B}$ ne soit pas en relation avec $\mathcal{B}''$, alors $\mathrm{P}_{\mathcal{B}',\mathcal{B}''} = \mathrm{P}_{\mathcal{B}',\mathcal{B}} \times \mathrm{P}_{\mathcal{B},\mathcal{B}''}$, d'où $\det(\mathrm{P}_{\mathcal{B}',\mathcal{B}''}) = \det(\mathrm{P}_{\mathcal{B}',\mathcal{B}}) \times \det(\mathrm{P}_{\mathcal{B},\mathcal{B}''}) > 0$, car ces deux déterminants sont négatifs, par conséquent $\mathcal{B}''\mathcal{R}\mathcal{B}'$, et donc la classe de $\mathcal{B}''$ et égale à celle de $\mathcal{B}'$, il n'y a pas d'autres classe d'équivalence que celles de $\mathcal{B}$ et de $\mathcal{B}'$. □

Définition 25.14

Orienter l'espace vectoriel E *c'est choisir une des deux classes d'équivalence, que l'on appelle* ***classe des bases directes****, l'autre étant alors appelée* ***classe des bases indirectes****. Comme il n'y a que deux choix, il n'y a donc que deux orientations possibles. Dans la pratique, on choisit une base de* E*, et on décrète que celle-ci est directe.*

À retenir

Le déterminant de la matrice de passage entre deux bases directes (ou deux bases indirectes) est strictement positif. Le déterminant de la matrice de passage entre une base directe et une base indirecte est strictement négatif.

3) Systèmes linéaires

C'est un système de la forme (S) $\left\{\begin{array}{rcl} a_{11}x_1 + \dots + a_{1p}x_p & = & b_1 \\ \vdots & & \\ a_{n1}x_1 + \dots + a_{np}x_p & = & b_n \end{array}\right.$, où les scalaires $x_1, \dots, x_p$ sont les inconnues.

Le système (S) peut s'interpréter de plusieurs façons :

- Interprétation géométrique : chaque équation de (S) peut être vue comme l'équation d'un hyperplan de $\mathbb{K}^p$ lorsque les coefficients ne sont pas tous nuls. Les solutions de (S) sont les coordonnées des points de l'intersection de n hyperplans.
- Sous forme matricielle : (S) $\iff$ AX = B où $X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$, $B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$, et $A = \begin{pmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{np} \end{pmatrix}$, A est la **matrice du système**. Le rang de la matrice A est appelé le rang du système S, il peut être vu comme le nombre maximal d'équations indépendantes du système.
- Sous forme linéaire : soit $u \in \mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$ l'application linéaire canoniquement associée à A, soit $x = (x_1, \dots, x_p) \in \mathbb{K}^p$, et soit $b = (b_1, \dots, b_n) \in \mathbb{K}^n$, on a alors (S) $\iff u(x) = b$.
- Sous forme vectorielle : soient $c_1(\mathrm{A}), \dots, c_p(\mathrm{A})$ les vecteurs colonnes de A, on a (S) $\iff x_1c_1(\mathrm{A}) + \dots + x_pc_p(\mathrm{A}) = b$, c'est une équation vectorielle dans $\mathbb{K}^n$ (on cherche toutes les combinaisons linéaires des vecteurs colonnes de a égales à b).

Définition 25.15 (Systèmes de Cramer)

C'est un système linéaire ***carré possédant une unique solution****, ce qui revient à dire que la matrice du système est carrée inversible.*

Formules de Cramer[2] : on a ici $n = p$, notons $\mathrm{C}_1, \dots, \mathrm{C}_n$ les colonnes de A ,on a $\mathrm{B} = x_1\mathrm{C}_1 + \dots + x_n\mathrm{C}_n$, soit D_i la matrice obtenue en remplaçant dans A la colonne i par la colonne B, on a :

$$\det(\mathrm{D}_i) = \sum_{k=1}^{n} x_k \det(\mathrm{C}_1, \dots, \mathrm{C}_{i-1}, \mathrm{C}_k, \mathrm{C}_{i+1}, \dots, \mathrm{C}_n)$$

ce qui donne $\det(\mathrm{D}_i) = x_i \det(\mathrm{A})$ par linéarité sur la variable i, d'où les formules :

$$\forall\, i \in [\![1\,;n]\!], x_i = \frac{\det(\mathrm{C}_1, \dots, \mathrm{C}_{i-1}, \mathrm{B}, \mathrm{C}_{i+1}, \dots, \mathrm{C}_n)}{\det(\mathrm{A})} \qquad \text{(formules de Cramer)}$$

2. *CRAMER Gabriel* (1704 – 1752) : mathématicien français qui s'est intéressé aux systèmes linéaires et à la théorie des déterminants.

pour avoir x_i : on remplace la colonne i de A par le second membre B, on en calcule le déterminant, puis son divise par celui de A.

☞**Exemple** : Si $ad - bc \neq 0$, alors le système (S) $\begin{cases} ax + by &= u \\ cx + dy &= v \end{cases}$, d'inconnues x et y, possède une unique solution :

$$\begin{cases} x &= \frac{\begin{vmatrix} u & b \\ v & d \end{vmatrix}}{ad-bc} = \frac{ud-vb}{ad-bc} \\ y &= \frac{\begin{vmatrix} a & u \\ c & v \end{vmatrix}}{ad-bc} = \frac{av-cu}{ad-bc} \end{cases}$$

À partir de la taille $n = 3$ ces formules ne sont pas très efficaces en terme de complexité ($n + 1$ déterminants à calculer...).

VII SOLUTION DES EXERCICES

Solution 25.1

1/ *Cycles de* $\mathfrak{S}_3$: id_{E_3}, $(1\ 2)$, $(1\ 3)$, $(2\ 3)$, $(1\ 2\ 3)$ *et* $(1\ 3\ 2)$.

2/ *Pour faire un cycle de longueur k, on prend k entiers parmi n pour faire le support,* $\binom{n}{k}$ *choix, puis il faut faire un cycle avec ces k entiers, c'est comme placer k convives autour d'un table ronde, il y a* $(k-1)!$ *façons de faire. Soit au total* $(k-1)!\binom{n}{k} = \frac{n!}{(n-k)!k}$ *cycles de longueur k.*

3/ *Si* $\sigma = (i_1\ i_2\ \dots\ i_p)$ *alors* $\sigma^{-1} = (i_p\ i_{p-1}\ \cdots\ i_1)$.

4/ *Si* $\sigma = (j_0\ j_1\ \dots\ j_{p-1})$*, alors on peut vérifier que* $\sigma(j_k) = j_{k+1 \pmod p}$*, on en déduit que* $\sigma^i(j_k) = j_{k+i \pmod p}$ *(les autres entiers étant fixes). On a donc bien* $\sigma^p = \mathrm{id}$ *et si* $i < p$ *alors* $\sigma^i(j_k) = j_{k+i \pmod p} \neq j_k$ *donc* $\sigma^i \neq \mathrm{id}$*. Si a est un élément du support de σ, alors on a* $a = j_k$*, si* $\sigma^i(a) = \sigma^s(a)$ *alors* $j_{k+i \pmod p} = j_{k+s \pmod p}$ *et donc* $i \equiv s \pmod p$*, donc les entiers* $a, \sigma(a), \dots, \sigma^{p-1}(a)$ *sont distincts et* $\sigma^p(a) = a$*, d'où* $\sigma = (a\ \ \sigma(a) \dots \sigma^{p-1}(a))$.

Solution 25.2

1/ *Si* $c = (i_1\ i_2\ \dots\ i_p)$ *et* $c' = (j_1\ j_2\ \dots\ j_k)$ *deux cycles à support disjoints, si k est un entier, en distinguant trois cas :* $k \in \left\{i_1, \dots, i_p\right\}$, $k \in \{j_1, \dots, j_k\}$ *et* $k \notin \left\{i_1, \dots, i_p\right\} \cup \{j_1, \dots, j_k\}$*, on vérifie que* $c \circ c'(k) = c' \circ c(k)$.

2/ *Les cycles à support non disjoints ne commutent pas, par exemple :* $(1\ 2) \circ (2\ 3) = (1\ 2\ 3)$ *alors que* $(2\ 3) \circ (1\ 2) = (1\ 3\ 2)$.

Solution 25.3 *Par linéarité sur chaque vecteur colonne, on peut sortir le coefficient λ de chaque colonne.*

Chapitre 26

Probabilités sur un univers fini

Sommaire

Le calcul des probabilités est la branche des mathématiques qui modélise les phénomènes aléatoires [1].

I UNIVERS FINIS

1) Expérience aléatoire

Définition 26.1
Une expérience est dite aléatoire lorsque l'issue ne peut pas être prédite avec certitude (on parle aussi d'épreuve aléatoire). On considère qu'à une expérience aléatoire correspond un ensemble contenant tous les résultats possibles, cet ensemble est appelé **univers** *et noté* Ω.

Remarque 26.1 – *On ne donne pas de définition mathématique du terme* expérience. *Par contre, l'univers* Ω *associé à une expérience aléatoire, est un ensemble au sens mathématique du terme, sa détermination résulte d'une modélisation de l'expérience (avec éventuellement des choix simplificateurs).*

☞**Exemples** :
- Expérience : on lance un dé, on note le numéro de la face supérieure.
  Univers : $\Omega = [\![1\,;6]\!]$.
- Expérience : on lance trois dés distincts, on note les trois numéros de la face supérieure.
  Univers : $\Omega = [\![1\,;6]\!]^3$, ensemble des triplets de nombres entre 1 et 6.
- Expérience : on prélève trois cartes d'un jeu de 32.
  Univers : Ω est l'ensemble des parties à 3 éléments de l'ensemble des 32 cartes.

1. Andreï Kolmogorov (25 avril 1903 – 20 octobre 1987) : mathématicien russe, fondateur de la théorie des probabilités dans les années 30.

– Expérience : on lance une pièce n fois.
  Univers : $\Omega = \{P, F\}^n$ avec P pour pile et F pour face.

Conformément au programme on se limitera au cas où Ω est un ensemble non vide et fini.

2) Évènements

Définition 26.2

On considère une expérience dont l'univers est Ω (fini).
*• Les éléments de Ω sont appelés les **possibles** (ou éventualités), en général notés avec une lettre minuscule : $\Omega = \{\omega_1, \dots, \omega_n\}$.*
*• On appelle **événement** toute partie de Ω. Un singleton est appelé **événement élémentaire**. L'ensemble des événements est donc $\mathcal{P}(\Omega)$.*
*• L'ensemble vide est appelé **événement impossible** et Ω et appelé **événement certain**.*
*• Si A est un événement, on appelle **événement contraire** l'événement $\overline{A}$ (parfois noté A^c, c'est le complémentaire de A dans Ω).*
• On dit que l'événement A implique (ou entraîne) l'événement B lorsque $A \subset B$.

☞**Exemple** : Expérience : lancer un dé. Univers $\Omega = [\![1\,;6]\!]$. Les événements élémentaires sont $\{1\}$, $\{2\}$, $\{3\}$, $\{4\}$, $\{5\}$, $\{6\}$. L'événement « obtenir un chiffre pair » est représenté par $A = \{2, 4, 6\}$, l'événement « obtenir un chiffre impair » est l'événement contraire, représenté par $A^c = \{1, 3, 5\}$. L'événement « obtenir le chiffre 1 » implique l'événement « obtenir un chiffre impair ».

Définition 26.3 (Opérations sur les événements)

On considère une expérience dont l'univers est Ω (fini). Soient A et B deux événements.
• L'événement « A ou B » est représenté par la réunion $A \cup B$.
• L'événement « A et B » est représenté par l'intersection $A \cap B$ (conjonction des événements).
*• On dira que ces deux événements sont **incompatibles** (ou disjoints) lorsque $A \cap B = \emptyset$.*

☞**Exemple** : Tout événement est incompatible avec l'événement contraire.

Définition 26.4 (Système complet d'événements)

*On considère une expérience dont l'univers est Ω (fini). Soit $(A_i)_{i\in[\![1;n]\!]}$ une famille d'événements, on dit que cette famille est un **système complet d'événements** lorsque :*
• les événements sont deux à deux incompatibles : $\forall (i, j) \in [\![1\,;n]\!]^2,\ i \neq j \implies A_i \cap A_j = \emptyset$.
• la réunion des ces événements est l'événement certain : $\bigcup\limits_{i=1}^{n} A_i = \Omega$.

Remarque 26.2 – *On retrouve une partie de la définition de partition, la différence est qu'on ne demande pas à ce que les événements A_i soient non vides.*

☞**Exemples** :
– Si A est un événement alors $\{A; A^c\}$ est un système complet d'événements.
– La famille des événements élémentaires, $(\{\omega\})_{\omega\in\Omega}$, est un système complet d'événements.

II ESPACES PROBABILISÉS

Une fois l'expérience clairement définie et son univers déterminé, on cherche à définir le pourcentage de chance de réalisation des événements. L'approche statistique consiste à répéter n fois l'expérience et calculer la fréquence de réalisation de chaque événement A : $f_n(A) = \frac{\text{nombre de réalisations de A}}{n}$. La fréquence apparaît alors comme une application de $\mathcal{P}(\Omega)$ vers $[0; 1]$ qui vérifie $f_n(\Omega) = 1$ et $f_n(A \cup B) = f_n(A) + f_n(B)$ lorsque A et B sont incompatibles (il en découle qu'il suffit de connaître la fréquence de chaque événement élémentaire). C'est cette approche qui est à l'origine de la définition suivante.

1) Probabilité

Définition 26.5

*Soit Ω un univers fini. On appelle **probabilité** sur Ω toute application $\mathbb{P}$ de $\mathcal{P}(\Omega)$ vers $[0;1]$ vérifiant :*
- $\mathbb{P}(\Omega) = 1$;
- $\mathbb{P}$ *est **additive**, c'est à dire :*

$$\forall\ A, B \in \mathcal{P}(\Omega),\ \text{si } A \text{ et } B \text{ sont incompatibles, alors } \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B).$$

*Le couple $(\Omega, \mathbb{P})$ est appelé **espace probabilisé**.*

Exemples :
- Si Ω est fini, alors en posant pour tout événement A, $\mathbb{P}(A) = \frac{\operatorname{card}(A)}{\operatorname{card}(\Omega)}$, on définit une probabilité sur Ω.
- Si $\Omega = \{\omega_1, \dots, \omega_n\}$, alors en posant pour tout événement A, $\mathbb{P}(A) = \begin{cases} 1 & \text{si } \omega_1 \in A \\ 0 & \text{sinon} \end{cases}$, on définit une probabilité sur Ω.

Attention !

Il n'y a pas unicité de la probabilité sur un univers fini, le choix de celle-ci fait partie de la modélisation de l'expérience et résulte souvent d'une hypothèse (par exemple : toutes les issues sont équiprobables).

2) Propriétés

Théorème 26.1

Soit $(\Omega, \mathbb{P})$ un espace probabilisé, et soient A et B deux événements :
- $\mathbb{P}(\varnothing) = 0$;
- $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$;
- *si $A \subset B$ alors $\mathbb{P}(A) \leqslant \mathbb{P}(B)$ (on dit que $\mathbb{P}$ est **croissante**) et $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A)$;*
- $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$;
- *pour toute famille $(A_i)_{i \in [\![1;n]\!]}$ d'événements **incompatibles deux à deux** :*

$$\mathbb{P}(A_1 \cup \cdots \cup A_n) = \sum_{i=1}^{n} \mathbb{P}(A_i).$$

Preuve :
- On a $1 = \mathbb{P}(\Omega) = \mathbb{P}(\Omega \cup \varnothing) = \mathbb{P}(\Omega) + \mathbb{P}(\varnothing) = 1 + \mathbb{P}(\varnothing)$, d'où le résultat.
- On a $1 = \mathbb{P}(\Omega) = \mathbb{P}(A \cup A^c) = \mathbb{P}(A) + \mathbb{P}(A^c)$, d'où le résultat.
- On a $\mathbb{P}(B) = \mathbb{P}(A \cup B \setminus A) = \mathbb{P}(A) + \mathbb{P}(B \setminus A) \geqslant \mathbb{P}(A)$ car une probabilité est positive. Les deux résultats en découlent.
- On a $\mathbb{P}(B) = \mathbb{P}([A \cap B] \cup B \setminus A) = \mathbb{P}(A \cap B) + \mathbb{P}(B \setminus A)$, d'où $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A \cap B)$, puis on a $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B \setminus A)$, le résultat en découle.
- récurrence sur n. □

★**Exercice 26.1** ***(Inégalité de Boole)***

Montrer que toute famille d'événements $A_1, \dots, A_n$ on a $\mathbb{P}(A_1 \cup \cdots \cup A_n) \leqslant \sum_{i=1}^{n} \mathbb{P}(A_i)$.

Théorème 26.2

Soit $(A_i)_{1 \in [\![1;n]\!]}$ un système complet d'événements, on a :
- $\sum_{i=1}^{n} \mathbb{P}(A_i) = 1$;
- *Pour tout événement B : $\mathbb{P}(B) = \sum_{i=1}^{n} \mathbb{P}(B \cap A_i)$.*

Preuve : Les événements de la famille sont incompatibles deux à deux et leur réunion donne Ω, donc $1 = \mathbb{P}(\Omega) = \mathbb{P}(A_1 \cup \cdots \cup A_n) = \sum_{i=1}^{n} \mathbb{P}(A_i)$.

$B = B \cap \Omega = B \cap (A_1 \cup \cdots \cup A_n) = (B \cap A_1) \cup \cdots \cup (B \cap A_n)$ et ces événements sont incompatibles deux à deux, d'où $\mathbb{P}(B) = \sum_{i=1}^{n} \mathbb{P}(B \cap A_i)$. □

3) Probabilité des événements élémentaires

Une probabilité $\mathbb{P}$ sur un univers fini Ω est entièrement déterminée par la connaissance des probabilités des événements élémentaires.

Théorème 26.3

Soit $\Omega = \{\omega_1, \ldots, \omega_n\}$ un univers fini.

- *Soit $\mathbb{P}$ une probabilité sur Ω et soit $p_i = \mathbb{P}(\{\omega_i\})$, on a alors : $\forall i \in [\![1\,;n]\!]$, $p_i \geqslant 0$, et $\sum\limits_{i=1}^{n} p_i = 1$.*
- *Réciproquement, si $(p_i)_{i\in[\![1;n]\!]}$ est une famille de réels positifs et de somme égale à 1, alors il existe une probabilité $\mathbb{P}$ sur Ω telle que $\forall i \in [\![1\,;n]\!]$, $\mathbb{P}(\{\omega_i\}) = p_i$. Pour tout événement A, on a alors $\mathbb{P}(\mathrm{A}) = \sum\limits_{\omega_i\in \mathrm{A}} p_i$.*

Preuve : Celle-ci est simple et laissée en exercice. □

★**Exercice 26.2** *On lance un dé truqué à 6 faces, la probabilité d'obtenir la face k est proportionnelle à k. Quelle est la probabilité d'obtenir un chiffre pair ?*

Théorème 26.4

*Sur tout univers fini Ω, il existe une unique probabilité $\mathbb{P}$ qui prend la même valeur sur chaque événement élémentaire (on parle d'événements équiprobables). Elle est définie par $\forall \omega \in \Omega$, $\mathbb{P}(\{\omega\}) = \frac{1}{\mathrm{card}(\Omega)}$, on l'appelle **probabilité uniforme** sur Ω, et pour tout événement A on a $\mathbb{P}(\mathrm{A}) = \frac{\mathrm{card}(\mathrm{A})}{\mathrm{card}(\Omega)}$.*

Preuve : Il suffit de vérifier que la valeur commune est un réel positif, et que $\sum\limits_{\omega\in\Omega} \frac{1}{\mathrm{card}(\Omega)} = 1$. On a alors $\mathbb{P}(\mathrm{A}) = \sum\limits_{\omega\in \mathrm{A}} \mathbb{P}(\{\omega\}) = \sum\limits_{\omega\in \mathrm{A}} \frac{1}{\mathrm{card}(\Omega)} = \frac{\mathrm{card}(\mathrm{A})}{\mathrm{card}(\Omega)}$. □

Remarque 26.3 – *La formule $\mathbb{P}(\mathrm{A}) = \frac{\mathrm{card}(\mathrm{A})}{\mathrm{card}(\Omega)}$ s'énonce parfois « nombre de cas favorables sur nombre de cas possibles », dans ce cadre, calculer des probabilités se ramène à des calculs de dénombrement.*

★**Exercice 26.3**

1/ *On lance six fois un dé non truqué. Quelle est la probabilité d'obtenir une fois chaque numéro ?*

2/ *On prélève cinq cartes au hasard dans un jeu de 52. Quelle est la probabilité d'avoir au moins une paire ?*

3/ *On distribue 52 cartes à 4 joueurs (13 par joueur). Quelle est la probabilité que chacun reçoive un as ?*

III PROBABILITÉS CONDITIONNELLES

Soit $(\Omega, \mathbb{P})$ un univers fini muni de la probabilité uniforme (par exemple), soit A un événement dont on sait qu'il est réalisé, un événement B est réalisé si et seulement si $\mathrm{A}\cap\mathrm{B}$ est réalisé, donc la probabilité de B sachant que A est réalisé est $\frac{\mathrm{card}(\mathrm{A}\cap\mathrm{B})}{\mathrm{card}(\mathrm{A})} = \frac{\mathrm{card}(\mathrm{A}\cap\mathrm{B})/\mathrm{card}(\Omega)}{\mathrm{card}(\mathrm{A})/\mathrm{card}(\Omega)} = \frac{\mathbb{P}(\mathrm{A}\cap\mathrm{B})}{\mathbb{P}(\mathrm{A})}$.

1) Définition

Définition 26.6

*Soit $(\Omega, \mathbb{P})$ un espace probabilisé fini et A un événement tel que $\mathbb{P}(\mathrm{A}) \neq 0$. Pour tout événement B on appelle **probabilité de** B **sachant** A le réel noté $\mathbb{P}_{\mathrm{A}}(\mathrm{B})$ est défini par $\mathbb{P}_{\mathrm{A}}(\mathrm{B}) = \frac{\mathbb{P}(\mathrm{A}\cap\mathrm{B})}{\mathbb{P}(\mathrm{A})}$.*

Théorème 26.5 (probabilité conditionnelle)

*Si A est un événement tel que $\mathbb{P}(\mathrm{A}) \neq 0$, alors l'application $\mathbb{P}_{\mathrm{A}} : \mathcal{P}(\Omega) \to \mathbb{R}$ définie par $\mathbb{P}_{\mathrm{A}}(\mathrm{B}) = \frac{\mathbb{P}(\mathrm{A}\cap\mathrm{B})}{\mathbb{P}(\mathrm{A})}$ est une probabilité sur Ω, appelée **probabilité conditionnelle** sachant A.*

Preuve : Celle-ci est simple et laissée en exercice. □

Attention !

On veillera à ne pas confondre $\mathbb{P}_A(B)$ et $\mathbb{P}(A \cap B)$.

Remarque 26.4 –

- *Dans certains ouvrages la probabilité conditionnelle $\mathbb{P}_A(B)$ est notée $\mathbb{P}(B \mid A)$, mais nous éviterons cette notation qui pourrait laisser croire que $B \mid A$ est un événement.*
- *Comme $\mathbb{P}_A$ est une probabilité, on a en particulier $\mathbb{P}_A(B^c) = 1 - \mathbb{P}_A(B)$.*

★**Exercice 26.4** *Une famille a deux enfants, chaque enfant a une chance sur deux d'être une fille.*

1/ *Quelle est la probabilité que les deux enfants soient des filles sachant que l'aînée est une fille ?*

2/ *Quelle est la probabilité que les deux enfants soient des filles sachant qu'il y a au moins fille ?*

2) Probabilités composées

La plupart du temps on ne calcule pas $\mathbb{P}_A(B)$ à partir de $\mathbb{P}(A \cap B)$ car le plus souvent on connaît $\mathbb{P}_A(B)$ et $\mathbb{P}(A)$ ce qui permet d'en déduire $\mathbb{P}(A \cap B)$.

Théorème 26.6

Si A et B sont deux événements alors :
- $\mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}_A(B)$ *si* $\mathbb{P}(A) \neq 0$;
- $\mathbb{P}(A \cap B) = \mathbb{P}(B) \times \mathbb{P}_B(A)$ *si* $\mathbb{P}(B) \neq 0$.

Preuve : Il suffit de revenir à la définition de probabilité conditionnelle. □

★**Exercice 26.5** *Une urne contient 4 boules blanches et 2 noires. On tire une boule, on la remet dans l'urne en ajoutant une autre boule de la même couleur, puis on procède à un autre tirage. Quelle est la probabilité d'obtenir deux boules noires ?*

Théorème 26.7 (formules des probabilités composées)

Soient $(A_i)_{i \in [\![1;n]\!]}$ une famille d'événements ($n \geqslant 2$) telle que $\mathbb{P}(A_1 \cap \cdots \cap A_{n-1}) \neq 0$, alors :

$$\mathbb{P}(A_1 \cap \cdots \cap A_n) = \mathbb{P}(A_1) \times \mathbb{P}_{A_1}(A_2) \times \mathbb{P}_{A_1 \cap A_2}(A_3) \times \cdots \times \mathbb{P}_{A_1 \cap \cdots \cap A_{n-1}}(A_n).$$

Preuve : Remarquons que toutes les probabilités conditionnelles de la formule sont bien définies car $0 \neq \mathbb{P}(A_1 \cap \cdots \cap A_{n-1}) \leqslant \cdots \leqslant \mathbb{P}(A_1 \cap A_2) \leqslant \mathbb{P}(A_1)$. La formule se démontre par récurrence sur n, le théorème précédent est le cas $n = 2$. En appliquant l'hypothèse de récurrence on a :

$$\mathbb{P}(A_1 \cap \cdots \cap A_{n+1}) = \mathbb{P}(A_1) \times \mathbb{P}_{A_1}(A_2) \times \mathbb{P}_{A_1 \cap A_2}(A_3) \times \cdots \times \mathbb{P}_{A_1 \cap \cdots \cap A_{n-1}}(A_n \cap A_{n+1})$$

notons $B = A_1 \cap \cdots \cap A_{n-1}$, alors :

$$\mathbb{P}_B(A_n \cap A_{n+1}) = \frac{\mathbb{P}(B \cap A_n \cap A_{n+1})}{\mathbb{P}(B)} = \frac{\mathbb{P}(B \cap A_n) \times \mathbb{P}_{B \cap A_n}(A_{n+1})}{\mathbb{P}(B)} = \mathbb{P}_B(A_n) \times \mathbb{P}_{B \cap A_n}(A_{n+1})$$

en reportant dans la relation ci-dessus, on obtient la formule au rang $n+1$. □

Remarque 26.5 – *Le produit à droite de l'égalité dans le théorème est un produit télescopique.*

À retenir

Cette formule est utile lorsque les événements sont dans un ordre chronologique.

★**Exercice 26.6** *Une urne contient n boules : b blanches et r rouges. Soit $k \in [\![1\,;b+1]\!]$. On tire des boules successivement de cette urne sans remise. Quelle est la probabilité qu'une boule rouge apparaisse pour la première fois au k^e tirage ?*

★**Exercice 26.7** *Une puce se déplace sur les trois sommets A, B C d'un triangle en partant de A. À chaque instant elle fait un saut : si elle est en A alors elle va en B, si elle est en B alors elle a une chance sur deux d'aller en A et une chance sur deux d'aller en C, si elle est en C elle y reste.*

1/ *Montrer qu'on ne peut arriver en C qu'à des instants pairs.*

2/ *Quelle est la probabilité que la puce arrive en C pour la première fois à l'instant $2n$?*

3) Formule des probabilités totales

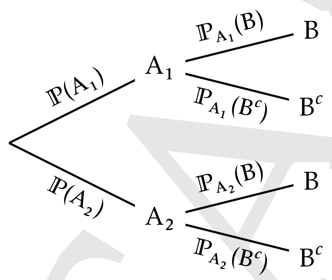
Soit $(A_1, \dots, A_n)$ un système complet d'événements sur un espace probabilisé fini $(\Omega, \mathbb{P})$, on a déjà vu que pour tout événement B, on a $\mathbb{P}(B) = \sum_{k=1}^{n} \mathbb{P}(B \cap A_k)$, or nous savons que $\mathbb{P}(B \cap A_k) = \mathbb{P}(A_k)\mathbb{P}_{A_k}(B)$ lorsque $\mathbb{P}(A_k) \neq 0$. Nous pouvons donc énoncer :

Théorème 26.8 (formule des probabilités totales)
Soit $(A_1, \dots, A_n)$ un système complet d'événements sur un espace probabilisé fini $(\Omega, \mathbb{P})$, tous de probabilité non nulle, alors pour tout événement B on a :
$$\mathbb{P}(B) = \sum_{k=1}^{n} \mathbb{P}(A_k)\mathbb{P}_{A_k}(B).$$

★**Exercice 26.8** *Le fonctionnement d'un appareil est régi par la règle suivante :*
- *s'il fonctionne à l'instant t_n alors il a la probabilité $a \in]0;1[$ de fonctionner à l'instant t_{n+1} ;*
- *s'il est en panne à l'instant t_n alors il a la probabilité $b \in]0;1[$ d'être en panne à l'instant t_{n+1}.*

Il est en état de marche à l'instant 0. Déterminer la probabilité p_n qu'il soit en état de marche à l'instant t_n.

Remarque 26.6 – *Lorsque le système complet comporte peu d'événements, on peut illustrer la formule des probabilités totales par un arbre pondéré, mais la justification est l'application de la formule des probabilités totales.*



4) Formule de Bayes

Soient A et B deux événements de probabilité non nulle de l'espace probabilisé $(\Omega, \mathbb{P})$, imaginons que l'on sache calculer la probabilité conditionnelle $\mathbb{P}_A(B)$ et que l'on souhaite calculer $\mathbb{P}_B(A)$. On a :

$$\boxed{\mathbb{P}_B(A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A)\mathbb{P}_A(B)}{\mathbb{P}(B)}}$$

Cela peut-être utile lorsqu'il y a une chronologie, (A antérieure à B) et que l'on cherche à « remonter le temps ». Dans le cas le plus général, on utilise la formule des probabilités totales pour exprimer $\mathbb{P}(B)$:

Théorème 26.9 (formule de Bayes (ou formule de probabilité des causes))
Soit $(A_1, \dots, A_n)$ un système complet d'événements sur un espace probabilisé fini $(\Omega, \mathbb{P})$, tous de probabilité non nulle, soit B un événement de probabilité non nulle, alors pour tout $i \in [\![1\,;n]\!]$:
$$\mathbb{P}_B(A_i) = \frac{\mathbb{P}(A_i)\mathbb{P}_{A_i}(B)}{\sum_{k=1}^{n} \mathbb{P}(A_k)\mathbb{P}_{A_k}(B)}.$$

Remarque 26.7 – *En particulier, lorsque le système est réduit à deux événements (A, A^c) (de probabilité non nulle), et si $\mathbb{P}(B) \neq 0$, alors $\mathbb{P}_B(A) = \frac{\mathbb{P}(A)\mathbb{P}_A(B)}{\mathbb{P}(A)\mathbb{P}_A(B)+\mathbb{P}(A^c)\mathbb{P}_{A^c}(B)}$.*

★**Exercice 26.9** *On a six urnes numérotées de 1 à 6, l'urne k contient k boules blanches et 6 – k boules noires. On lance un dé non truqué, si la face k sort alors on tire une boule de l'urne k. La boule tirée est blanche, quelle est la probabilité d'avoir fait un 6 ?*

IV INDÉPENDANCE

1) Indépendance de deux événements

Définition 26.7

On dit que deux événements A *et* B *d'un espace probabilisé* $(\Omega, \mathbb{P})$ *sont indépendants lorsque :*
$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}(B).$$

Remarque 26.8 –
- *La relation est symétrique.*
- *Si* $\mathbb{P}(A) = 0$ *alors* A *est indépendant avec tout autre événement* B.

À retenir

Lorsque $\mathbb{P}(A) \neq 0$, alors A et B sont indépendants si et seulement si $\mathbb{P}_A(B) = \mathbb{P}(B)$.

☞**Exemple** : On lance un dé non truqué, soit A:« obtenir un chiffre pair » et B : « obtenir un chiffre inférieur ou égale à 4 ». On a $\mathbb{P}(A) = \frac{1}{2}$, $\mathbb{P}(B) = \frac{2}{3}$, $\mathbb{P}(A \cap B) = \frac{1}{3} = \mathbb{P}(A)\mathbb{P}(B)$, ces deux événements sont indépendants. Soit C:« obtenir un chiffre inférieur ou égale à 3 », on a $\mathbb{P}(C) = \frac{1}{2}$ et $\mathbb{P}(A \cap C) = \frac{1}{6} \neq \mathbb{P}(A)\mathbb{P}(C)$, A et C ne sont pas indépendants.

Attention !

On veillera à ne pas confondre indépendance de deux événements (qui dépend de la probabilité) et l'incompatibilité de deux événements (qui ne dépend pas de la probabilité).

Théorème 26.10

Si A *est* B *sont deux événements* ***indépendants*** *de* $(\Omega, \mathbb{P})$ *alors :*
- A *et* B^c *sont indépendants ;*
- A^c *et* B *sont indépendants ;*
- A^c *et* B^c *sont indépendants.*

Preuve : $\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B^c) = \mathbb{P}(A)\mathbb{P}(B) + \mathbb{P}(A \cap B^c)$ d'où $\mathbb{P}(A \cap B^c) = \mathbb{P}(A)[1 - \mathbb{P}(B)] = \mathbb{P}(A)\mathbb{P}(B^c)$. Les deux autres résultats en découlent. □

2) Indépendance d'une famille d'événements

Définition 26.8

Soient $A_1, \ldots, A_n$ *des événements dans* $(\Omega, \mathbb{P})$ *:*
- *On dit ces événements sont* ***deux à deux indépendants*** *lorsque :*
$$\forall i, j \in [\![1\,;n]\!], i \neq j \implies \mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i)\mathbb{P}(A_j).$$
- *On dit ces événements sont* ***mutuellement indépendants*** *lorsque :*
$$\forall J \subset [\![1\,;n]\!],\ \mathbb{P}(\bigcap_{k \in J} A_k) = \prod_{k \in J} \mathbb{P}(A_k).$$

Remarque 26.9 – *La vérification de l'indépendance mutuelle nécessite beaucoup de vérifications :* $\sum_{k=2}^{n} \binom{n}{k} = 2^n - n - 1$. *Il découle de la définition que des événements mutuellement indépendants sont deux à deux indépendants.*

Attention !

Des événements deux à deux indépendants ne sont pas nécessairement mutuellement indépendants comme le montre l'exemple suivant.

☞**Exemple** : On lance deux dés parfaits, on note A_1 : « le premier dé amène un nombre pair », A_2 : « le deuxième dé amène un nombre pair » et A_3 : « la somme des nombres obtenus est paire ». On a $\mathbb{P}(A_1) = \frac{1}{2}$, $\mathbb{P}(A_2) = \frac{1}{2}$, $\mathbb{P}(A_3) = \frac{18}{36} = \frac{1}{2}$, $\mathbb{P}(A_1 \cap A_2) = \frac{1}{4}$, $\mathbb{P}(A_1 \cap A_3) = \frac{1}{4}$, $\mathbb{P}(A_2 \cap A_3) = \frac{1}{4}$, mais $\mathbb{P}(A_1 \cap A_2 \cap A_3) = \mathbb{P}(A_1 \cap A_2) = \frac{1}{4}$: les événements sont deux à deux indépendants mais pas mutuellement indépendants.

★**Exercice 26.10** *Soient* $A_1, \dots, A_n$ *des événements mutuellement indépendants dans* $(\Omega, \mathbb{P})$.

1/ Montrer que les événements $B_1, \dots, B_n$ *où* $B_i = A_i$ *ou* A_i^c*, sont mutuellement indépendants.*

2/ Soit $p \in [\![1\,;n-1]\!]$*, montrer que* $B_1 \cup \dots \cup B_p$ *et* $B_{p+1} \cap \dots \cup B_n$ *sont indépendants.*

V SOLUTION DES EXERCICES

Solution 26.1 *Pour* $n = 2$: $\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2) \leqslant \mathbb{P}(A_1) + \mathbb{P}(A_2)$. *On termine ensuite par récurrence sur* n.

Solution 26.2 *On a* $\mathbb{P}(\{k\}) = ak$ *avec k une constante positive, or* $\sum_{k=1}^{6} \mathbb{P}(\{k\}) = 1$*, on en déduit que* $a = \frac{1}{1+2+3+4+5+6} = \frac{1}{21}$*, la probabilité demandée est* $\mathbb{P}(\{2,4,6\}) = a(2+4+6) = \frac{12}{21} = \frac{4}{7}$.

Solution 26.3

1/ $\Omega = [\![1\,;6]\!]^6$*, les issues sont considérées équiprobables. L'événement cherché est l'ensemble 6-listes contenant une et une seule fois chaque chiffre, par conséquent* $\operatorname{card}(A) = 6!$ *et* $\mathbb{P}(A) = \frac{6!}{6^6} \approx 0{,}015$.

2/ Ω *est l'ensemble des 5-parties de l'ensemble des cartes. Dénombrons l'événement contraire (aucune paire), il faut donc 5 hauteurs de cartes différentes et pour chacune des hauteurs il faut choisir une carte parmi 4 ce qui donne* $\binom{13}{5}4^5$ *mains dans cet événement, d'où la probabilité cherchée* $1 - \frac{\binom{13}{5}4^5}{\binom{52}{5}} \approx 0{,}49$.

3/ La probabilité cherchée est $\frac{4\binom{48}{12}\times 3\binom{36}{12}\times 2\binom{24}{12}\times 1\binom{12}{12}}{\binom{52}{13}\binom{39}{13}\binom{26}{13}} = \frac{4!13^4}{52\cdot 51\cdot 50\cdot 49} \approx 0{,}105$.

Solution 26.4 *L'univers est* $\Omega = \{F, G\}^2$ *ensemble des couples de lettres* F *ou* G *(premier enfant et deuxième enfant).*

1/ L'événement « l'aînée est une fille » est $A = \{(F,G),(F,F)\}$ *et* $\mathbb{P}(A) = \frac{2}{4} = \frac{1}{2}$. *L'événement cherché est* $B = \{(F,F)\}$ *qui implique* A*, d'où* $\mathbb{P}_A(B) = \frac{\mathbb{P}(B)}{\mathbb{P}(A)} = \frac{1}{2}$.

2/ L'événement « au moins un enfant est une fille » est $A = \{(F,G);(F,F);(G,F)\}$ *et* $\mathbb{P}(A) = \frac{3}{4}$. *L'événement cherché est* $B = \{(F,F)\}$ *qui implique* A*, d'où* $\mathbb{P}_A(B) = \frac{\mathbb{P}(B)}{\mathbb{P}(A)} = \frac{1}{3}$.

Solution 26.5 *Notons* A *l'événement « la première boule tirée est noire », et* B *l'événement « la deuxième boule tirée est noire ». On cherche à calculer* $\mathbb{P}(A \cap B)$*, on a* $\mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}_A(B) = \frac{2}{6} \times \frac{3}{7} = \frac{1}{7}$.

Solution 26.6 *Soit* B_i *l'événement « au* i^e *tirage la boule est blanche (on définit de même* R_i*), on demande :*

$$p(B_1 \cap \dots \cap B_{k-1} \cap R_k) = p(B_1)p_{B_1}(B_2)\cdots p_{B_1\cap\dots B_{k-1}}(R_k) = \frac{b}{n} \times \dots \times \frac{b-k+2}{n-k+2} \times \frac{r}{n-k+1} = \frac{b(b-1)\cdots(b-k+2)r}{n(n-1)\cdots(n-k+1)}$$

Solution 26.7

1/ Pour arriver en C *la puce doit venir de* B *(un saut), pour être en* B *la puce doit venir de* A *et pour arriver en* A *la puce doit venir de* B*, donc avant d'arriver en* C *la puce fait un certain nombre d'aller-retours en* A *et* B*, ce qui donne un instant impair pour être en* B *et donc pair pour être en* C.

2/ Soit A_k *l'événement « la puce est en* A *à l'instant* k *» et* B_k *l'événement « la puce est en* B *à l'instant* k *», on cherche donc* $\mathbb{P}(B_1 \cap A_2 \cap \dots \cap B_{2n-1} \cap C_{2n})$*, d'après la formule des probabilités composées, c'est :*

$$\mathbb{P}(B_1 \cap A_2 \cap \dots \cap B_{2n-1} \cap C_{2n}) = \mathbb{P}(B_1) \times \mathbb{P}_{B_1}(A_2) \times \mathbb{P}_{B_1 \cap A_2}(B_3) \times \dots \times \mathbb{P}_{B_1 \cap A_2 \cap \dots \cap B_{2n-1}}(C_{2n})$$

ce qui donne $\left(\frac{1}{2}\right)^n = \frac{1}{2^n}$.

Solution 26.8 *Soit* M_n *l'événement correspondant, alors* (M_n, M_n^c) *est un système complet d'événements donc* $p_{n+1} = \mathbb{P}(M_n)\mathbb{P}_{M_n}(M_{n+1}) + \mathbb{P}(M_n^c)\mathbb{P}_{M_n^c}(M_{n+1}) = p_n a + (1-p_n)(1-b) = (a+b-1)p_n + 1 - b$*, on a une suite arithmético-géométrique, une solution constante est* $\ell = \frac{1-b}{2-a-b}$ *et la suite* $(p_n - \ell)$ *est géométrique de raison* $a+b-1$*, d'où* $p_n = (a+b-1)^n(p_0 - \ell) + \ell$ *avec* $p_0 = 1$*, ce qui donne* $p_n = \frac{1-b+(a+b-1)^n(1-a)}{2-a-b}$. *On voit que lorsque* $|a+b-1| < 1$*, cette suite tend vers* $\frac{1-b}{2-a-b}$.

Solution 26.9 *On a un système complet* $(D_1, \dots, D_6)$ *(résultat du jet de dé), soit* B *l'événement « obtenir une boule blanche », alors on demande* $\mathbb{P}_B(D_6)$*, d'après la formule de Bayes, on a :*

$$\mathbb{P}_B(D_6) = \frac{\mathbb{P}(D_6)\mathbb{P}_{D_6}(B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(D_6)\mathbb{P}_{D_6}(B)}{\sum_{k=1}^{6}\mathbb{P}(D_k)\mathbb{P}_{D_k}(B)} = \frac{1/6}{\sum_{k=1}^{6}\frac{k}{36}} = \frac{6}{21} = \frac{2}{7}.$$

Solution 26.10

1/ *Il suffit de traiter le cas où un seul des* B_i *est égal* A_i^c *(on peut supposer que c'est* B_n*), si* J *est une partie de* $[\![1\,;n-1]\!]$*, alors on sait que* $\bigcap_{j\in\mathrm{J}}\mathrm{A}_j$ *et* A_n *sont indépendants, donc* $\bigcap_{j\in\mathrm{J}}\mathrm{A}_j$ *et* A_n^c *aussi.*

2/ *On sait que* $\mathrm{B}_1^c,\dots,\mathrm{B}_p^c,\mathrm{B}_{p+1},\dots,\mathrm{B}_n$ *sont mutuellement indépendants, donc* $\mathrm{B}_1^c\cap\cdots\cap\mathrm{B}_p^c$ *et* $\mathrm{B}_{p+1}\cap\cdots\cap\mathrm{B}_n$ *sont indépendants, par conséquent,* $\left[\mathrm{B}_1^c\cap\cdots\cap\mathrm{B}_p^c\right]^c=\mathrm{B}_1\cup\cdots\cup\mathrm{B}_p$ *et* $\mathrm{B}_{p+1}\cap\cdots\cap\mathrm{B}_n$ *sont indépendants.*

Chapitre 27

Variables aléatoires sur un univers fini

Sommaire

Souvent, lors d'une expérience aléatoire, on ne s'intéresse pas vraiment à l'ensemble des résultats possibles, mais un aspect particulier du résultat, par exemple : la somme obtenue lors d'un lancer de deux dés, le temps d'attente du premier pile dans un lancer de pièce etc.

I NOTION DE VARIABLE ALÉATOIRE

1) Définition

Définition 27.1

Soit Ω un univers fini. Toute application $X\colon \Omega \to E$ est appelée ***variable aléatoire****, lorsque $E = \mathbb{R}$, on dit que X est une* ***variable aléatoire réelle*** *(notée VAR). Lorsque X est constante, on parle de VA constante ou de variable aléatoire* ***certaine****.*

Remarque 27.1 –

– *Notons qu'une variable aléatoire n'est une pas une variable en réalité, mais une application, et n'a rien d'aléatoire !*
– *L'ensemble des images $X(\Omega)$ est un ensemble fini qu'on note parfois $X(\Omega) = \{x_1, \ldots, x_n\}$.*

☞**Exemples** :

– Expérience : on lance deux dés, on note X la somme des chiffres obtenus.
Univers : $\Omega = [\![1\,;6]\!]$, $X(\Omega) = [\![2\,;12]\!]$.

– Expérience : on lance n fois une pièce, on note X le numéro du lancer où apparaît un pîle la première fois.
Univers : $\Omega = \{P, F\}^n$, $X(\Omega) = [\![0\,;n]\!]$.

Définition 27.2 (variable indicatrice d'un événement)

Soit A *un événement, l'application* $1_A : \Omega \to \mathbb{R}$ *définie par* $1_A(\omega) = \begin{cases} 1 \text{ si } \omega \in A \\ 0 \text{ sinon} \end{cases}$ *est la* ***variable indicatrice de*** A.

Remarque 27.2 –

– *L'ensemble des VAR sur* Ω *est* $\mathcal{F}(\Omega, \mathbb{R})$ *qui a une structure de* $\mathbb{R}$*-espace vectoriel et d'anneau pour les lois usuelles.*
– *Si* $X : \Omega \to E$ *est une VA et* $u : E \to F$ *est une application, alors* $u \circ X$ *est une VA sur* Ω*, elle est généralement notée* $u(X)$.

Événements liés à une VA

– Si $X : \Omega \to E$ est une VA, si A est une partie de E alors $X^{-1}(A)$ est un événement, on le note généralement $(X \in A)$, on a donc $(X \in A) = \{\omega \in \Omega \ / \ X(\omega) \in A\}$.
– Plus généralement, si $P(x)$ désigne une propriété définie sur E, l'événement $\{\omega \in \Omega \ / \ P(X(\omega))\}$ est noté (X vérifie P).
– Dans le cas particulier d'une VAR, on écrit pour tout $x \in \mathbb{R}$: $(X \leqslant x) = \{\omega \in \Omega \ / \ X(\omega) \leqslant x\}$; $(X \geqslant x) = \{\omega \in \Omega \ / \ X(\omega) \geqslant x\}$; etc. En particulier $(X \leqslant x) = (X < x) \cup (X = x)$.

2) Loi d'une VA

Théorème 27.1

Si X *est une VA sur* $(\Omega, \mathbb{P})$*, l'application* $\mathbb{P}_X : \mathcal{P}(X(\Omega)) \to [0;1]$ *définie par* $\mathbb{P}_X(A) = \mathbb{P}(X \in A)$ *pour toute partie* A *de* $X(\Omega)$*, est une probabilité sur* $X(\Omega)$*, on l'appelle* ***loi de*** X.

Preuve : On a $\mathbb{P}_X(X(\Omega)) = \mathbb{P}(X \in X(\Omega)) = \mathbb{P}(\Omega) = 1$. Soient A et B deux parties disjointes de $X(\Omega)$ alors $X^{-1}(A)$ et $X^{-1}(B)$ sont deux événements incompatibles de Ω, donc $\mathbb{P}_X(A \cup B) = \mathbb{P}(X^{-1}(A \cup B)) = \mathbb{P}(X^{-1}(A) \cup X^{-1}(B)) = \mathbb{P}(X^{-1}(A)) + \mathbb{P}(X^{-1}(B)) = \mathbb{P}_X(A) + \mathbb{P}_X(B)$. □

Remarque 27.3 – *Si* A *est une partie de* $X(\Omega)$ *alors* $\mathbb{P}(X \in A) = \mathbb{P}_X(A) = \sum\limits_{x \in A} \mathbb{P}_X(\{x\}) = \sum\limits_{x \in A} \mathbb{P}(X = x)$.

À retenir

Déterminer la loi de X, c'est déterminer $X(\Omega)$ et pour tout $x \in X(\Omega)$, la probabilité $\mathbb{P}(X = x)$. On remarquera que si X est une VA sur l'univers fini Ω, alors la famille d'événements $((X = x))_{x \in X(\Omega)}$ est un système complet d'événements.

Il est clair que ces événements sont incompatibles deux à deux et que :

$$\sum_{x \in (\Omega)} \mathbb{P}(X = x) = \mathbb{P}\Big(\bigcup_{x \in X(\Omega)} (X = x)\Big) = \mathbb{P}(\Omega) = 1$$

Remarque 27.4 – *Parfois on ne connaît pas exactement* $X(\Omega)$ *mais un ensemble* E *contenant* $X(\Omega)$*, dans ce cas, pour* $x \notin X(\Omega)$ *on trouve* $\mathbb{P}(X = x) = 0$.

★**Exercice 27.1** *On lance un dé deux fois, on note* X *la somme obtenue. Déterminer la loi de* X.

★**Exercice 27.2** *Soit* X *une VAR à valeurs entières, montrer que* $\mathbb{P}(X = k) = \mathbb{P}(X \leqslant k) - \mathbb{P}(X \leqslant k-1) = \mathbb{P}(X \geqslant k)) - \mathbb{P}(X \geqslant k+1)$.

3) Image d'une VA par une fonction

Théorème 27.2

Soit X une VA sur $(\Omega, \mathbb{P})$ et u une fonction définie sur $X(\Omega)$, notons $Y = u \circ X$, alors :

$$\forall y \in Y(\Omega), \mathbb{P}(Y = y) = \sum_{x \in u^{-1}(\{y\})} \mathbb{P}(X = x)$$

Preuve : $(Y = y) = \{\omega \in \Omega \ / \ u \circ X(\omega) = y\} = (X \in u^{-1}(\{y\}))$, d'où :
$\mathbb{P}(Y = y) = \mathbb{P}_X(u^{-1}(\{y\}) = \sum_{x \in u^{-1}(\{y\})} \mathbb{P}_X(\{x\}) = \sum_{x \in u^{-1}(\{y\})} \mathbb{P}(X = x)$. □

★**Exercice 27.3** *On lance un dé successivement deux fois, on note X le premier résultat moins le deuxième, déterminer la loi de X, de $|X|$ et de X^2.*

Remarque 27.5 – *Si X et Y sont deux VA sur $(\Omega, \mathbb{P})$ alors on peut définir les lois de $X + Y$ et de XY :*

$$\mathbb{P}(X + Y = z) = \sum_{(x,y) \in I_z} \mathbb{P}((X = x) \cap (Y = y)) \text{ où } I_z = \{(x, y) \in X(\Omega) \times Y(\Omega) \ / \ x + y = z\}$$

$$\mathbb{P}(XY = z) = \sum_{(x,y) \in I_z} \mathbb{P}((X = x) \cap (Y = y)) \text{ où } I_z = \{(x, y) \in X(\Omega) \times Y(\Omega) \ / \ xy = z\}$$

II LOIS USUELLES

1) Variables certaines

Définition 27.3

Une VA est dite certaine lorsqu'elle est constante, auquel cas on a $X(\Omega) = \{x_1\}$ et donc $\mathbb{P}(X = x_1) = 1$.

2) Loi uniforme

Définition 27.4

On dit que la VA X suit une loi uniforme sur $X(\Omega) = \{x_1, \dots, x_n\}$ lorsque $\forall k \in [\![1 ; n]\!], \mathbb{P}(X = x_k) = \frac{1}{n}$. Ce que l'on note $X \hookrightarrow \mathcal{U}(X(\Omega))$.

☞**Exemple** : On lance un dé parfait, on note X le chiffre obtenu alors $X \hookrightarrow \mathcal{U}([\![1 ; 6]\!])$.

3) Loi de Bernoulli

Définition 27.5

Soit $p \in [0; 1]$ et $q = 1 - p$. On dit que la VA X suit une loi de Bernoulli de paramètre p, lorsque sur $X(\Omega) = \{0, 1\}$, $\mathbb{P}(X = 1) = p$ et $\mathbb{P}(X = 0) = q$. Ce que l'on note $X \hookrightarrow \mathcal{B}(p)$.

Remarque 27.6 – *Cette loi est utile pour les expériences aléatoires ayant deux issues possibles : succès avec probabilité p ou bien échec avec une probabilité de $1 - p$, ce que l'on appelle des épreuves de Bernoulli.*

☞**Exemple** : La variable indicatrice d'un événement A suit une loi de Bernoulli de paramètre $p = \mathbb{P}(A)$.

Remarque 27.7 – *Si X et Y sont deux VA qui suivent une loi de Bernoulli, alors XY également.*

4) Loi binomiale

Définition 27.6

Soient $n \in \mathbb{N}$, $p \in [0; 1]$ et $q = 1 - p$. On dit que la VA X suit une loi binomiale de paramètres n et p, lorsque sur $X(\Omega) = [\![0 ; n]\!]$, et $\forall k \in [\![0 ; n]\!]$, $\mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$. Ce que l'on note

$X \hookrightarrow \mathcal{B}(n,p)$.

On vérifie qu'il s'agit bien d'une loi de probabilité sur $[\![0\,;n]\!]$.

À retenir

La loi binomiale apparaît quand on considère le nombre X de succès dans une suite de n épreuves de Bernoulli **mutuellement indépendantes** et dans lesquelles la probabilité du succès est $p \in [0;1]$.

En effet, on a $X(\Omega) = [\![0\,;n]\!]$ et l'événement $(X = k)$ est l'ensemble des n-uplets de 1 ou 0 contenant k fois le nombre 1, il y en a $\binom{n}{k}$, la probabilité de chacun d'eux est $p^k(1-p)^{n-k}$ car les épreuves sont mutuellement indépendantes, d'où $\mathbb{P}(X = k) = \binom{n}{k}p^k(1-p)^{n-k}$.

☞**Exemples** :

- On effectue n tirages successifs avec remise dans une urne contenant des boules blanches et noires dans des proportions de p et $q = 1 - p$. On note X le nombre de boules blanches obtenues, alors $X \hookrightarrow \mathcal{B}(n,p)$.
- Une pièce donne face avec une probabilité p et pile avec une probabilité $q = 1 - p$. On effectue n lancers dans les mêmes conditions à chaque fois, on note X le nombre de faces obtenus, alors $X \hookrightarrow \mathcal{B}(n,p)$.

III ESPÉRANCE ET VARIANCE D'UNE VARIABLE ALÉATOIRE

1) Espérance

Définition 27.7

L'espérance d'une variable aléatoire X est le réel noté $\mathbb{E}(X)$ et défini par :

$$\mathbb{E}(X) = \sum_{x \in X(\Omega)} x\mathbb{P}(X = x).$$

*Lorsque son espérance est nulle, on dit que la variable aléatoire X est **centrée**.*

Remarque 27.8 – *Il s'agit de la moyenne pondérée des valeurs prises par X, car la somme des probabilités vaut* 1.

☞**Exemples** :

- On lance un dé équilibré, X est le numéro de la face supérieure, $\mathbb{E}(X) = \sum_{i=1}^{6} \frac{i}{6} = \frac{21}{6} = \frac{7}{2}$.
- Si A est un événement, l'espérance de l'indicatrice de A est $\mathbb{E}(1_A) = \mathbb{P}(A)$.
- L'espérance d'une variable certaine est la constante.

Théorème 27.3

Si X est une variable aléatoire sur $(\Omega, \mathbb{P})$ alors $\mathbb{E}(X) = \sum_{\omega \in \Omega} X(\omega)\mathbb{P}(\{\omega\})$.

Preuve : Soit $X(\Omega) = \{x_1, \dots, x_p\}$, soit $A_k = X^{-1}(\{x_k\})$, alors $(A_1, \dots, A_p)$ est un système complet d'événements et $\mathbb{P}(A_k) = \sum_{\omega \in A_k} \mathbb{P}(\{\omega\})$. On a donc :

$$\begin{aligned}
\sum_{\omega \in \Omega} X(\omega)\mathbb{P}(\{\omega\}) &= \sum_{k=1}^{p}\left(\sum_{\omega \in A_k} X(\omega)\mathbb{P}(\{\omega\})\right) \\
&= \sum_{k=1}^{p}\left(\sum_{\omega \in A_k} x_k\mathbb{P}(\{\omega\})\right) \\
&= \sum_{k=1}^{p} x_k\left(\sum_{\omega \in A_k} \mathbb{P}(\{\omega\})\right) \\
&= \sum_{k=1}^{p} x_k\mathbb{P}(A_k) = \sum_{k=1}^{p} x_k\mathbb{P}(X = x_k) = \mathbb{E}(X).
\end{aligned}$$

□

Théorème 27.4 (**linéarité de l'espérance**)
Si X *et* Y *sont deux variables aléatoires sur* $(\Omega, \mathbb{P})$ *alors pour tout réel* λ *:*
$$\mathbb{E}(\lambda\mathrm{X}+\mathrm{Y}) = \lambda\mathbb{E}(\mathrm{X}) + \mathbb{E}(\mathrm{Y}).$$

Preuve : En utilisant le résultat du théorème précédent, on a :

$$\mathbb{E}(\lambda\mathrm{X}+\mathrm{Y}) = \sum_{\omega\in\Omega}[\lambda\mathrm{X}(\omega)+\mathrm{Y}(\omega)]\mathbb{P}(\{\omega\}) = \lambda\sum_{\omega\in\Omega}\mathrm{X}(\omega)\mathbb{P}(\{\omega\}) + \sum_{\omega\in\Omega}\mathrm{Y}(\omega)\mathbb{P}(\{\omega\}) = \lambda\mathbb{E}(\mathrm{X}) + \mathbb{E}(\mathrm{Y}).$$

□

À retenir
Si X est une variable aléatoire alors $\mathrm{X}-\mathbb{E}(\mathrm{X})$ est une variable aléatoire **centrée**. C'est la variable aléatoire centrée associée à X.

★**Exercice** 27.4 *Une urne contient n boules numérotées de 1 à n, on les tire successivement et sans remise, on dit qu'il y a rencontre au tirage i lorsque la boule tirée est la boule numéro i. Déterminer le nombre moyen de rencontres.*

Théorème 27.5 (**positivité de l'espérance**)
Si X *est une variable aléatoire positive ou nulle sur* $(\Omega, \mathbb{P})$ *alors :*
- $\mathbb{E}(\mathrm{X}) \geqslant 0$ *(positivité).*
- $\mathbb{E}(\mathrm{X}) = 0 \iff \mathbb{P}(\mathrm{X}=0) = 1$*, on dit que la variable* X *est* ***presque sûrement nulle****.*

Preuve : Le premier point découle de la définition d'espérance $\mathbb{E}(\mathrm{X}) = \sum_{k=1}^{n} x_k\mathbb{P}(\mathrm{X}=x_k)$, car chaque valeur x_k de X est positive ou nulle et les probabilités aussi. Si cette somme est nulle, alors tous les termes doivent être nuls, X prend donc nécessairement la valeur 0 (sinon $\mathbb{E}(\mathrm{X}) > 0$ car les probabilités ne peuvent pas être toutes nulles) et pour $x_k \neq 0$ on a forcément $\mathbb{P}(\mathrm{X}=x_k) = 0$), ce qui entraîne que $\mathbb{P}(\mathrm{X}=0) = 1$ puisque la somme des probabilités doit faire 1. La réciproque est immédiate. □

Théorème 27.6 (**croissance de l'espérance**)
Si X *et* Y *sont deux variables aléatoires réelles définies sur* $(\Omega, \mathbb{P})$ *telles que* $\mathrm{X} \leqslant \mathrm{Y}$*, alors* $\mathbb{E}(\mathrm{X}) \leqslant \mathbb{E}(\mathrm{Y})$.

Preuve : La variable $\mathrm{Z} = \mathrm{Y}-\mathrm{X}$ est positive, donc $\mathbb{E}(\mathrm{Z}) \geqslant 0$, la linéarité permet alors d'écrire $\mathbb{E}(\mathrm{Y})-\mathbb{E}(\mathrm{X}) \geqslant 0$ ce qui donne le résultat. □

Théorème 27.7 (**inégalité de Markov**)
Pour tout variable aléatoire positive X *on a :*
$$\forall a > 0, \mathbb{P}(\mathrm{X} \geqslant a) \leqslant \frac{\mathbb{E}(\mathrm{X})}{a}.$$

Preuve : Notons $\mathrm{A} = \{x \in \mathrm{X}(\Omega) \mathbin{/} x \geqslant a\}$, $\mathrm{A} \subset \mathrm{X}(\Omega)$ et $(\mathrm{X} \geqslant a) = (\mathrm{X} \in \mathrm{A})$. On a :
$\mathbb{E}(\mathrm{X}) = \sum_{x\in\mathrm{X}(\Omega)} x\mathbb{P}(\mathrm{X}=x) \geqslant \sum_{x\in\mathrm{A}} x\mathbb{P}(\mathrm{X}=x) \geqslant a\sum_{x\in\mathrm{A}}\mathbb{P}(\mathrm{X}=x) = a\mathbb{P}(\mathrm{X}\in\mathrm{A})$, d'où $\mathbb{P}(\mathrm{X} \geqslant a) \leqslant \frac{\mathbb{E}(\mathrm{X})}{a}$. □

Remarque 27.9 – *On peut aussi montrer l'inégalité de Markov en remarquant que* $1_{(\mathrm{X}\geqslant a)} \leqslant \frac{\mathrm{X}}{a}$ *et utiliser la croissance de l'espérance.*

Théorème 27.8 (**de transfert**)
Si X *est une variable aléatoire sur* $(\Omega, \mathbb{P})$ *et si* u *est une application de* $\mathrm{X}(\Omega)$ *vers* $\mathbb{R}$*, alors :*
$$\mathbb{E}(u(\mathrm{X})) = \sum_{x\in\mathrm{X}(\Omega)} u(x)\mathbb{P}(\mathrm{X}=x).$$

Preuve : Posons $\mathrm{X}(\Omega) = \{x_1, \ldots, x_n\}$ et $\mathrm{A}_k = \mathrm{X}^{-1}(\{x_k\})$, alors $(\mathrm{A}_1, \ldots, \mathrm{A}_n)$ est un système complet d'événements de

Ω, et on a :

$$\begin{aligned}
\mathbb{E}(u(X)) &= \sum_{\omega\in\Omega} u(X(\omega))\mathbb{P}(\{\omega\}) \\
&= \sum_{k=1}^{n}\left(\sum_{\omega\in A_k} u(X(\omega))\mathbb{P}(\{\omega\})\right) = \sum_{k=1}^{n} u(x_k)\left(\sum_{\omega\in A_k}\mathbb{P}(\{\omega\})\right) \\
&= \sum_{k=1}^{n} u(x_k)\mathbb{P}(A_k) = \sum_{k=1}^{n} u(x_k)\mathbb{P}(X=x_k) \\
&= \sum_{x\in X(\Omega)} u(x)\mathbb{P}(X=x)
\end{aligned}$$

□

Remarque 27.10 – *La formule de droite proposée dans le théorème n'est autre que l'espérance de la variable aléatoire u sur l'espace probabilisé $(X(\Omega), \mathbb{P}_X)$. Cette formule fait intervenir la loi de probabilité de X mais pas celle de $u(X)$, ce qui fait son interêt.*

★**Exercice** 27.5 *Si $X \hookrightarrow \mathcal{U}([\![1\,;n]\!])$, calculer $\mathbb{E}(X^2)$ et $\mathbb{E}(e^X)$.*

2) Variance et écart-type

Définition 27.8 (moments d'une variable aléatoire)

Soit X un variable aléatoire réelle et $r \in \mathbb{N}$, on appelle ***moment d'ordre r de la variable*** *X l'espérance de la variable aléatoire X^r, c'est à dire le nombre $\mathbb{E}(X^r) = \sum_{x\in X(\Omega)} x^r\mathbb{P}(X=x)$ (théorème de transfert), on le note parfois $m_r(X)$.*

Remarque 27.11 – *Le moment d'ordre 0 vaut 1, et le moment d'ordre 1 est l'espérance.*

Définition 27.9 (variance, écart-type)

On appelle variance de la variable aléatoire X le réel $\mathbb{V}(X) = \mathbb{E}\left([X-\mathbb{E}(X)]^2\right)$, et l'écart-type de X est le réel $\sigma(X) = \sqrt{\mathbb{V}(X)}$.

Remarque 27.12 –

- *La définition a bien un sens puisque $X - \mathbb{E}(X)$ est une variable aléatoire. Ces notions permettent de mesurer la dispersion de X autour de sa valeur moyenne.*
- *$\mathbb{V}(X) = 0 \iff \mathbb{P}(X = \mathbb{E}(X)) = 1$, la variable aléatoire X est presque sûrement constante.*
- *Il résulte du théorème de transfert que $\mathbb{V}(X) = \sum_{x\in X(\Omega)} (x-\mathbb{E}(X))^2\mathbb{P}(X=x)$, mais la formule la plus utilisée est celle du théorème suivant.*

Théorème 27.9 (formule de Kœnig-Huygens)

Si X est une variable aléatoire réelle sur $(\Omega, \mathbb{P})$ alors $\mathbb{V}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$.

Preuve : En effet :
$(X-\mathbb{E}(X))^2 = X^2 + \mathbb{E}(X)^2 - 2X\mathbb{E}(X)$, l'espérance étant linéaire, on a $\mathbb{V}(X) = \mathbb{E}(X^2) + \mathbb{E}(X)^2 - 2\mathbb{E}(X)\mathbb{E}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$. □

★**Exercice** 27.6

1/ Si $X(\Omega) = [\![1\,;n]\!]$ et $\mathbb{P}(X=k) = ak$. Calculer a, $\mathbb{E}(X)$ et $\mathbb{V}(X)$.

2/ Si X est une VAR, déterminer le minimum de la fonction $f : t \mapsto \mathbb{E}([X-t]^2)$ sur $\mathbb{R}$.

À retenir

Il est parfois judicieux de calculer $\mathbb{E}(X(X-1))$ ou $\mathbb{E}(X(X+1))$ pour en déduire $\mathbb{E}(X^2)$.

Théorème 27.10 (propriétés)

Soit X une variable aléatoire réelle sur $(\Omega, \mathbb{P})$ alors :
- $\mathbb{V}(X) \geqslant 0$.
- $\forall a, b \in \mathbb{R}$, $\mathbb{V}(aX+b) = a^2\mathbb{V}(X)$ *et* $\sigma(aX+b) = |a|\sigma(X)$.
- ***Inégalité de Bienaymé-Tchebychev** :* $\forall \varepsilon > 0$, $\mathbb{P}(|X - \mathbb{E}(X)| \geqslant \varepsilon) \leqslant \frac{\mathbb{V}(X)}{\varepsilon^2}$.

Preuve :
- Par définition, $\mathbb{V}(X) = \mathbb{E}([X-\mathbb{E}(X)]^2)$ qui est donc un réel positif par positivité de l'espérance.
- On a $\mathbb{E}(aX+b) = a\mathbb{E}(X)+b$ d'où $[(aX+b)-\mathbb{E}(aX+b)]^2 = a^2[X-\mathbb{E}(X)]^2$ et donc $\mathbb{V}(aX+b) = \mathbb{E}(a^2[X-\mathbb{E}(X)]^2) = a^2\mathbb{V}(X)$. En prenant la racine carrée on obtient $\sigma(aX+b) = |a|\sigma(X)$.
- Soit $Y = [X-\mathbb{E}(X)]^2$, alors $\mathbb{P}(|X-\mathbb{E}(X)| \geqslant \varepsilon) = \mathbb{P}(Y \geqslant \varepsilon^2) \leqslant \frac{\mathbb{E}(Y)}{\varepsilon^2}$ (inégalité de Markov), ce qui donne exactement le résultat. □

Applications :
- Une variable aléatoire dont l'écart-type vaut 1 est appelée **variable aléatoire réduite**. Si X est une VAR dont la variance est non nulle, alors la variable aléatoire $X^* = \frac{X-\mathbb{E}(X)}{\sigma(X)}$ est une variable aléatoire **centrée réduite**, appelée variable centrée réduite associée à X.
- L'inégalité de Bienaymé-Tchebychev apporte réellement un renseignement lorsque $\frac{\mathbb{V}(X)}{\varepsilon^2} \leqslant 1$, mais cette majoration est en général assez grossière et son intérêt est surtout théorique. Elle indique néanmoins que X prend des valeurs proches de sa moyenne avec une grande probabilité. On peut également l'écrire sous la forme : $\mathbb{P}(|X-\mathbb{E}(X)| < \varepsilon) \geqslant 1 - \frac{\mathbb{V}(X)}{\varepsilon^2}$.

3) Cas des lois usuelles

Théorème 27.11

- *Si X est une variable certaine de valeur $a \in \mathbb{R}$, alors $\mathbb{E}(X) = a$ et $\mathbb{V}(X) = 0$.*
- *Si X suit la loi uniforme sur $\{x_1, \dots, x_n\}$ alors $\mathbb{E}(X) = \frac{x_1+\cdots+x_n}{n}$. En particulier si $X \hookrightarrow \mathcal{U}([\![1\,;n]\!])$ alors $\mathbb{E}(X) = \frac{n+1}{2}$ et $\mathbb{V}(X) = \frac{n^2-1}{12}$.*
- *Si X suit une loi de Bernoulli de paramètre $p \in [0;1]$ alors $\mathbb{E}(X) = p$ et $\mathbb{V}(X) = p(1-p)$.*
- *Si X suit une loi binomiale de paramètres n et p, alors $\mathbb{E}(X) = np$ et $\mathbb{V}(X) = np(1-p)$.*

Preuve :
- Immédiat.
- Si $X \hookrightarrow \mathcal{U}([\![1\,;n]\!])$: $\mathbb{E}(X) = \sum_{k=1}^{n} \frac{k}{n} = \frac{n+1}{2}$, $\mathbb{E}(X^2) = \sum_{k=1}^{n} \frac{k^2}{n} = \frac{(2n+1)(n+1)}{6}$, d'où $\mathbb{V}(X) = \frac{(2n+1)(n+1)}{6} - \frac{(n+1)^2}{4} = \frac{(n+1)(n-1)}{12}$.
- Si $X \hookrightarrow \mathcal{B}(p)$, $\mathbb{E}(X) = \mathbb{P}(X=1) = p$, X^2 suit la même loi que X donc $\mathbb{V}(X) = p - p^2 = p(1-p)$.
- Si $X \hookrightarrow \mathcal{B}(n,p)$, il s'agit de calculer $\mathbb{E}(X) = \sum_{k=0}^{n} k\binom{n}{k}p^kq^{n-k}$ (avec $q = 1-p$), cette expression est la dérivée en 1 de la fonction $f(x) = \sum_{k=0}^{n} \binom{n}{k}x^kp^kq^{n-k} = (px+q)^n$, d'où $f'(x) = pn(px+q)^{n-1}$ et donc $f'(1) = np$.

Autre méthode : une pièce donne face avec une probabilité p et pile avec une probabilité $q = 1-p$. On effectue n lancers dans les mêmes conditions à chaque fois, on note X le nombre de faces obtenus, alors $X \hookrightarrow \mathcal{B}(n,p)$. Notons X_i la variable aléatoire valant 1 si on a obtenu face au lancer numéro i, 0 sinon, alors $X_i \hookrightarrow \mathcal{B}(p)$ et $X = X_1 + \cdots + X_n$, d'où $\mathbb{E}(X) = \mathbb{E}(X_1) + \cdots + \mathbb{E}(X_n) = np$.

Pour la variance, on calcule $\mathbb{E}(X(X-1)) = \sum_{k=0}^{n} k(k-1)\binom{n}{k}p^kq^{n-k} = f''(1)$ avec les notations ci-dessus, or $f''(x) = p^2n(n-1)(px+q)^{n-2}$ et donc $\mathbb{E}(X(X-1)) = n(n-1)p^2 = \mathbb{E}(X^2) - \mathbb{E}(X)$, d'où $\mathbb{E}(X^2) = n(n-1)p^2 + np = np[np-p+1]$ et $\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = np[np-p+1] - (np)^2 = np(1-p)$. □

IV COUPLES DE VARIABLES ALÉATOIRES

1) Définitions

Définition 27.10

Soient X et Y deux variables aléatoires sur $(\Omega, \mathbb{P})$ à valeurs dans E pour X et F pour Y, on appelle couple des variables X et Y, et on note $Z = (X, Y)$, l'application :

$$Z : \begin{array}{rcl} \Omega & \to & E \times F \\ \omega & \mapsto & Z(\omega) = (X(\omega), Y(\omega)) \end{array}.$$

Remarque 27.13 – *L'application Z est une variable aléatoire sur $(\Omega, \mathbb{P})$ et $Z(\Omega) \subset X(\Omega) \times Y(\Omega)$.*

Définition 27.11 (loi conjointe)

*Soient $X : \Omega \to E$ et $Y : \Omega \to F$ deux variables aléatoires sur $(\Omega, \mathbb{P})$. On appelle **loi conjointe du couple** des variables X et Y, la loi de la variable aléatoire $Z = (X, Y)$, c'est à dire l'application :*

$$\begin{array}{ccc} Z(\Omega) & \to [0;1] \\ (x,y) & \mapsto & \mathbb{P}([X = x] \cap [Y = y]) \end{array}.$$

Remarque 27.14 – *Déterminer la loi conjointe revient donc à déterminer $X(\Omega) = \{x_1, \dots, x_n\}$ et $Y(\Omega) = \{y_1, \dots, y_p\}$, puis à calculer les probabilités $p_{ij} = \mathbb{P}([X = x_i] \cap [Y = y_j])$ pour $(i, j) \in [\![1 ; n]\!] \times [\![1 ; p]\!]$. Dans la pratique on présente la loi conjointe sous forme d'un tableau :*

X \ Y	y_1	y_2	$\cdots$	y_p
x_1	p_{11}	p_{12}	$\cdots$	p_{1p}
x_2	p_{21}	p_{22}	$\cdots$	p_{2p}
$\vdots$	$\vdots$	$\vdots$		$\vdots$
x_n	p_{n1}	p_{n2}	$\cdots$	p_{np}

Définition 27.12 (lois marginales)

*Si (X, Y) est un couple de variables aléatoires sur $(\Omega, \mathbb{P})$, les lois de probabilités de X et de Y sont appelées **lois marginales** du couple.*

À retenir

La loi conjointe permet d'obtenir les lois marginales, car on peut écrire (avec les notations précédentes) :

$$\mathbb{P}(X = x_i) = \sum_{j=1}^{p} \mathbb{P}([X = x_i] \cap [Y = y_j]) = \sum_{j=1}^{p} p_{ij}$$

et :

$$\mathbb{P}(Y = y_j) = \sum_{i=1}^{n} \mathbb{P}([X = x_i] \cap [Y = y_j]) = \sum_{i=1}^{n} p_{ij}.$$

Notation : on écrit $p_{i\bullet} = \mathbb{P}(X = x_i)$ et $p_{\bullet j} = \mathbb{P}(Y = y_j)$.

★**Exercice 27.7** *On lance 2 dés parfaits, on note X le minimum des deux résultats et Y le maximum. Déterminer la loi conjointe et les lois marginales.*

Attention !

La connaissance des lois marginales ne permet pas en général de retrouver la loi conjointe, comme on peut s'en convaincre sur l'exemple précédent.

Définition 27.13

*Soient $X : \Omega \to E$ et $Y : \Omega \to F$ deux variables aléatoires sur $(\Omega, \mathbb{P})$. Soit $y \in Y(\Omega)$, on appelle **loi conditionnelle de** X **sachant** $(Y = y)$ la loi de X dans l'espace $(\Omega, \mathbb{P}_{(Y=y)})$, c'est à dire l'application :*

$$\begin{array}{ccl} X(\Omega) & \to & [0;1] \\ x & \mapsto & \mathbb{P}_{(Y=y)}(X = x) = \frac{\mathbb{P}([X=x]\cap[Y=y])}{\mathbb{P}(Y=y)} \end{array}.$$

*De même, soit $x \in X(\Omega)$, on appelle **loi conditionnelle de** Y **sachant** $(X = x)$ la loi de Y dans l'espace $(\Omega, \mathbb{P}_{(X=x)})$, c'est à dire l'application :*

$$\begin{array}{ccl} Y(\Omega) & \to & [0;1] \\ y & \mapsto & \mathbb{P}_{(X=x)}(Y = y) = \frac{\mathbb{P}([Y=y]\cap[X=x])}{\mathbb{P}(X=x)} \end{array}.$$

☞**Exemple** : En reprenant l'exercice précédent, la loi de X sachant $(Y = 4)$ est :

x_i	1	2	3	4	5	6
$\mathbb{P}_{(Y=4)}(X = x_i)$	$\frac{2}{7}$	$\frac{2}{7}$	$\frac{2}{7}$	$\frac{1}{7}$	0	0

Plus généralement $\mathbb{P}_{(Y=j)}(X = i) = \begin{cases} 0 & \text{si } i > j \\ \frac{1/36}{\frac{2j-1}{36}} = \frac{1}{2j-1} & \text{si } i = j \\ \frac{2/36}{\frac{2j-1}{36}} = \frac{2}{2j-1} & \text{si } i < j \end{cases}$

Il découle de la formule des probabilités totales :

À retenir

Si (X, Y) est un couple de variables aléatoires sur $(\Omega, \mathbb{P})$, on suppose que pour tout $x \in X(\Omega)$ et tout $y \in Y(\Omega)$, $\mathbb{P}(X = x) \neq 0$ et $\mathbb{P}(Y = y) \neq 0$, alors :

$$\mathbb{P}(X = x) = \sum_{y \in Y(\Omega)} \mathbb{P}(Y = y) \times \mathbb{P}_{(Y=y)}(X = x),$$

et :

$$\mathbb{P}(Y = y) = \sum_{x \in X(\Omega)} \mathbb{P}(X = x) \times \mathbb{P}_{(X=x)}(Y = y),$$

La notion de couple de variables aléatoires se généralise à un n-uplet de variables aléatoires :

Définition 27.14 (généralisation : vecteurs aléatoires)

Soient $X_1, \dots, X_n$ des variables aléatoires sur $(\Omega, \mathbb{P})$ à valeurs dans $E_1, \dots, E_n$ respectivement, le vecteur aléatoire $Z = (X_1, \dots, X_n)$ est l'application :

$$Z : \begin{array}{ccl} \Omega & \to & E_1 \times \cdots \times E_n \\ \omega & \mapsto & Z(\omega) = (X_1(\omega), \cdots, X_n(\omega)) \end{array}.$$

La loi conjointe du vecteur Z est la loi de probabilité de la variable aléatoire $Z = (X_1, \dots, X_n)$, c'est à dire l'application :

$$\begin{array}{ccl} Z(\Omega) & \to & \mathbb{R} \\ (x_1, \dots, x_n) & \mapsto & \mathbb{P}([X_1 = x_1] \cap \cdots \cap [X_n = x_n]) \end{array}.$$

et les lois des variables aléatoires $X_1, \dots, X_n$ sont appelées lois marginales du vecteur Z.

2) Indépendance de variables aléatoires

Définition 27.15

Soient X et Y deux variables aléatoires sur $(\Omega, \mathbb{P})$, on dit que X et Y sont indépendantes lorsque :

$$\forall A \subset X(\Omega), \forall B \subset Y(\Omega), \mathbb{P}([X \in A] \cap [Y \in B]) = \mathbb{P}(X \in A) \times \mathbb{P}(Y \in B),$$

c'est à dire les événements $(X \in A)$ et $(Y \in B)$ sont indépendants.

Théorème 27.12

Deux variables aléatoires X et Y sur $(\Omega, \mathbb{P})$ sont indépendantes si et seulement si :

$$\forall (x, y) \in X(\Omega) \times Y(\Omega), \mathbb{P}([X = x] \cap [Y = y]) = \mathbb{P}(X = x) \times \mathbb{P}(Y = y).$$

Preuve : Le sens $\Longrightarrow$ est évident. Montrons la réciproque : soient A une partie de $X(\Omega)$ et B une partie de $Y(\Omega)$, on a alors $[X \in A] \cap [Y \in B] = \bigcup_{(x,y) \in A \times B} [X = x] \cap [Y = y]$ et ces événements sont incompatibles deux à deux, d'où :

$$\begin{aligned} \mathbb{P}([X \in A] \cap [Y \in B]) &= \sum_{(x,y) \in A \times B} \mathbb{P}([X = x] \cap [Y = y]) \\ &= \sum_{(x,y) \in A \times B} \mathbb{P}(X = x) \times \mathbb{P}(Y = y) \text{ (d'après l'hypothèse)} \\ &= \sum_{x \in A} \sum_{y \in B} \mathbb{P}(X = x) \times \mathbb{P}(Y = y) \\ &= \left(\sum_{x \in A} \mathbb{P}(X = x) \right) \times \left(\sum_{y \in B} \mathbb{P}(Y = y) \right) \\ &= \mathbb{P}(X \in A) \times \mathbb{P}(Y \in B) \end{aligned}$$

donc les deux variables aléatoires sont indépendantes. □

Remarque 27.15 – *Lorsque X et Y sont deux variables aléatoires indépendantes sur $(\Omega, \mathbb{P})$, la connaissance des deux lois marginales permet de reconstituer la loi conjointe car $p_{ij} = p_{i\bullet} \times p_{\bullet j}$, et les lois conditionnelles sont égales aux lois marginales.*

☞**Exemple** : Dans l'exercice 7, les variables aléatoires X et Y ne sont pas indépendantes.

Théorème 27.13 (indépendance de fonctions de variables aléatoires)
Soient X et Y deux variables aléatoires indépendantes sur $(\Omega, \mathbb{P})$, si f est une application sur $X(\Omega)$ et g une application sur $Y(\Omega)$, alors les variables aléatoires $f(X)$ et $g(Y)$ sont indépendantes.

Preuve : Soient $a \in \mathrm{Im}(f)$ et $b \in \mathrm{Im}(g)$, soit $A = f^{-1}(\{a\})$ et $B = g^{-1}(\{b\})$, alors :

$$\mathbb{P}([f(X) = a] \cap [g(Y) = b]) = \mathbb{P}([X \in A] \cap [Y \in B]) = \mathbb{P}(X \in A) \times \mathbb{P}(Y \in B) = \mathbb{P}(f(X) = a) \times \mathbb{P}(g(Y) = b)$$

Les variables aléatoires $f(X)$ et $g(Y)$ sont donc indépendantes. □

Voici un autre résultat utile :

À retenir
Soit X et Y deux variables aléatoires indépendantes sur $(\Omega, \mathbb{P})$, et $f\colon X(\Omega) \times Y(\Omega) \to E$, alors :
$$\mathbb{P}(f(X, Y) = z) = \sum_{(x,y) \in f^{-1}(\{z\})} \mathbb{P}(X = x) \times \mathbb{P}(Y = y).$$

En effet, il suffit de remarquer que l'événement $(f(X, Y) = z)$ s'écrit $\bigcup_{(x,y) \in f^{-1}(\{z\})} [X = x] \cap [Y = y]$.

★**Exercice 27.8** *Soient X et Y deux variables aléatoires indépendantes sur $(\Omega, \mathbb{P})$ suivant la même loi uniforme sur $[\![1\,;n]\!]$. Déterminer les lois de $X + Y$ et de $X - Y$.*

Définition 27.16 (généralisation : indépendance de n variables aléatoires)
Soit $(X_1, \dots, X_n)$ un vecteur aléatoire sur $(\Omega, \mathbb{P})$ à valeurs dans $E_1 \times \cdots \times E_n$. On dit que ces n variables aléatoires sont :
*• **deux à deux indépendantes** : lorsque pour tout couple $(i, j) \in [\![1\,;n]\!]^2$, X_i et X_j sont indépendantes ;*
*• **mutuellement indépendantes** : lorsque pour toute partie A_1 de E_1, A_2 de $E_2, \dots, A_n$ de E_n, les événements $(X_1 \in A_1), \dots, (X_n \in A_n)$ sont mutuellement indépendants.*

Remarque 27.16 – *Comme pour les événements, l'indépendance mutuelle des variables aléatoires entraîne l'indépendance deux à deux, mais la réciproque est fausse.*

En généralisant la preuve du théorème 27.12

Théorème 27.14
Les variables aléatoires $X_1, \dots, X_n$ sur $(\Omega, \mathbb{P})$ sont mutuellement indépendantes si et seulement si :
pour tout $(x_1, \dots, x_n) \in X_1(\Omega) \times \cdots \times X_n(\Omega)$, les événements $(X_1 = x_1), \dots, (X_n = x_n)$ sont mutuellement indépendantes.

De même, on montrerait que $f_1(X_1), \dots, f_n(X_n)$ sont mutuellement indépendantes, avec $f_i\colon X_i(\Omega) \to F_i$.

Théorème 27.15
On considère des variables aléatoires $X_1, \dots, X_n$ sur $(\Omega, \mathbb{P})$ mutuellement indépendantes.
• Soit $1 < p < n$, les variables aléatoires $Y = (X_1, \dots, Y_p)$ et $Z = (X_{p+1}, \dots, X_n)$ sont indépendantes.
• Si f est une fonction à p variables et g à $n - p$ variables, alors $f(X_1, \dots, X_p)$ et $g(X_{p+1}, \dots, X_n)$ sont indépendantes.

Preuve : Celle-ci est simple et laissée en exercice. □

Remarque 27.17 – *Le résultat ci-dessus se généralise au cas où on partage l'ensemble des n variables aléatoires en plus de deux parties, on remplace alors « indépendantes » par « mutuellement indépendantes » dans les conclusions. Par exemple, si X, Y, Z et T sont mutuellement indépendantes, alors XY et ZT sont indépendantes, les variables aléatoires X, Y + Z et T^2 sont mutuellement indépendantes etc.*

3) Applications de l'indépendance

Théorème 27.16

Si $X_1, \ldots, X_n$ n variables aléatoires sur $(\Omega, \mathbb{P})$ suivent toutes une même loi de Bernoulli de paramètre $p \in [0;1]$, et sont mutuellement indépendantes, alors la somme $X_1 + \cdots + X_n$ suit une loi binomiale de paramètre n et p.

Preuve : On a $X_i \hookrightarrow \mathcal{B}(p)$, soit $X = X_1 + \cdots + X_n$, alors $X(\Omega) = [\![0\,;n]\!]$, l'événement $(X = k)$ se réalise si et seulement si k des variables X_i prennent la valeur 1 et les $n-k$ autres la valeur 0, il y a donc $\binom{n}{k}$ cas possibles incompatibles deux à deux, calculons la probabilité d'un cas : notons $X_{i_1}, \ldots, X_{i_k}$ celles qui prennent la valeur 1 et $X_{i_{k+1}}, \ldots, X_{i_n}$ les autres alors du fait de l'indépendance, on a :

$$\mathbb{P}([X_{i_1} = 1] \cap \cdots \cap [X_{i_k} = 1] \cap [X_{i_{k+1}} = 0] \cap \cdots \cap [X_{i_n} = 0]) =$$
$$\mathbb{P}(X_{i_1} = 1) \times \cdots \times \mathbb{P}(X_{i_k} = 1) \times \mathbb{P}(X_{i_{k+1}} = 0) \times \cdots \times \mathbb{P}(X_{i_n} = 0) = p^k(1-p)^{n-k}.$$

par conséquent $\mathbb{P}(X = k) = \binom{n}{k}p^k(1-p)^{n-k}$. □

À retenir

Si $X \hookrightarrow \mathcal{B}(n,p)$, alors on peut écrire $X = X_1 + \cdots + X_n$ où les X_i sont mutuellement indépendantes et suivent toutes la loi de Bernoulli de paramètre p. On a donc $\mathbb{E}(X) = \mathbb{E}(X_1) + \cdots + \mathbb{E}(X_n) = p + \cdots + p = np$.

★**Exercice 27.9** *Soit $(X_k)_{k\in[\![1;m]\!]}$ m variables aléatoires sur $(\Omega, \mathbb{P})$ qui suivent toutes une loi binomiale de paramètre (n_k, p) (respectivement) avec $p \in [0;1]$, et sont mutuellement indépendantes, alors :*

$$X_1 + \cdots + X_m \hookrightarrow \mathcal{B}(n_1 + \cdots + n_m, p).$$

Théorème 27.17 (espérance d'un produit)

Si X et Y sont deux variables aléatoires indépendantes sur $(\Omega, \mathbb{P})$, alors $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$, mais la réciproque est fausse. Le résultat se généralise au produit de n variables aléatoires mutuellement indépendantes.

Preuve : Soit $Z = (X, Y)$ et $u\colon (x,y) \mapsto xy$, alors :

$$\begin{aligned}
\mathbb{E}(XY) = \mathbb{E}(u(Z)) &= \sum_{(x,y)\in Z(\Omega)} u(x,y)\mathbb{P}(Z = (x,y)) \text{ (théorème de transfert)}\\
&= \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} xy\mathbb{P}(X = x) \times \mathbb{P}(Y = y)) \text{ (indépendance)}\\
&= \left(\sum_{x\in X(\Omega)} x\mathbb{P}(X = x)\right)\left(\sum_{y\in Y(\Omega)} y\mathbb{P}(Y = y)\right)\\
&= \mathbb{E}(X)\mathbb{E}(Y)
\end{aligned}$$

une récurrence permet ensuite la généralisation. □

★**Exercice 27.10** *Soit X une variable aléatoire qui suit une loi uniforme sur $\{-1;0;1\}$ et Y la fonction indicatrice de l'événement $(X = 0)$. Montrer que $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$ mais que X et Y ne sont pas indépendantes.*

Théorème 27.18 (variance et indépendance)

Si X et Y sont deux variables aléatoires indépendantes sur $(\Omega, \mathbb{P})$, alors $\mathbb{V}(X+Y) = \mathbb{V}(X) + \mathbb{V}(Y)$. Le résultat se généralise au cas de n variables aléatoires mutuellement indépendantes.

Preuve : D'après la formule de Kœnig-Huygens, $\mathbb{V}(X+Y) = \mathbb{E}([X+Y]^2) - [\mathbb{E}(X) + \mathbb{E}(Y)]^2$ avec la linéarité de l'espérance, on a donc $\mathbb{V}(X+Y) = \mathbb{E}(X^2) + \mathbb{E}(Y^2) + 2\mathbb{E}(XY) - \mathbb{E}(X)^2 - \mathbb{E}(Y)^2 - 2\mathbb{E}(X)\mathbb{E}(Y) = \mathbb{V}(X) + \mathbb{V}(Y)$ car les variables aléatoires étant indépendantes, on a $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$. Une récurrence permet ensuite la généralisation. □

À retenir

Si $X \hookrightarrow \mathcal{B}(n,p)$, alors on peut écrire $X = X_1 + \cdots + X_n$ où les X_i sont mutuellement indépendantes et suivent toutes la loi de Bernoulli de paramètre p. On a donc $\mathbb{V}(X) = \mathbb{V}(X_1) + \cdots + \mathbb{V}(X_n) = p(1-p) + \cdots + p(1-p) = np(1-p)$.

4) Covariance

Définition 27.17

Soient X *et* Y *deux variables aléatoires réelles sur* $(\Omega, \mathbb{P})$. *On appelle* ***covariance de*** X ***et*** Y *le réel défini par :*

$$\mathrm{cov}(X,Y) = \mathbb{E}([X - \mathbb{E}(X)][Y - \mathbb{E}(Y)]).$$

Remarque 27.18 – *C'est l'espérance du produit des variables centrées associées à* X *et* Y. *Lorsque* $Y = X$ *on a* $\mathrm{cov}(X,X) = \mathbb{V}(X) \geqslant 0$, *on dit que la covariance est positive.*

Théorème 27.19 (propriétés de la covariance)

Soient X *et* Y *deux variables aléatoires réelles sur* $(\Omega, \mathbb{P})$ *:*
• $\mathrm{cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$ *;*
• $\mathrm{cov}(Y,X) = \mathrm{cov}(X,Y)$, *la covariance est symétrique ;*
• *La covariance est bilinéaire :* $\mathrm{cov}(aX + Y, Z) = a\mathrm{cov}(X,Z) + \mathrm{cov}(Y,Z)$ *(même chose sur la deuxième variable) ;*
• $\mathbb{V}(X+Y) = \mathbb{V}(X) + \mathbb{V}(Y) + 2\mathrm{cov}(X,Y)$ *;*
• *Si* X *et* Y *sont indépendantes, alors* $\mathrm{cov}(X,Y) = 0$ *mais la réciproque est fausse.*

Preuve :
• On a $[X - \mathbb{E}(X)][Y - \mathbb{E}(Y)] = XY - \mathbb{E}(X)Y - \mathbb{E}(Y)X + \mathbb{E}(X)\mathbb{E}(Y)$, on applique ensuite la linéarité de l'espérance ce qui donne la formule ;
• immédiat ;
• On utilise la formule ci-dessus et la linéarité de l'espérance : $\mathbb{E}([aX+Y]Z) - \mathbb{E}(aX+Y)\mathbb{E}(Z) = a\mathbb{E}(XZ) + \mathbb{E}(YZ) - a\mathbb{E}(X)\mathbb{E}(Z) - \mathbb{E}(Y)\mathbb{E}(Z) = a\mathrm{cov}(X,Z) + \mathrm{cov}(Y,Z)$. Par symétrie on a la linéarité sur la deuxième variable.
• D'après un calcul fait plus haut : $\mathbb{V}(X+Y) = \mathbb{E}([X+Y]^2) - [\mathbb{E}(X) + \mathbb{E}(Y)]^2$ avec la linéarité de l'espérance, on a donc $\mathbb{V}(X+Y) = \mathbb{E}(X^2) + \mathbb{E}(Y^2) + 2\mathbb{E}(XY) - \mathbb{E}(X)^2 - \mathbb{E}(Y)^2 - 2\mathbb{E}(X)\mathbb{E}(Y) = \mathbb{V}(X) + \mathbb{V}(Y) + 2\mathrm{cov}(X,Y)$.
• On sait que si X et Y sont indépendantes alors $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$ ce qui donne $\mathrm{cov}(X,Y) = 0$. □

Remarque 27.19 – *La covariance est une forme bilinéaire symétrique et positive, mais ce n'est pas un produit scalaire car* $\mathbb{V}(X) = 0$ *n'entraîne pas* $X = 0$, *mais seulement* $\mathbb{P}(X = m) = 1$. *On peut cependant établir l'inégalité de Cauchy-Schwarz pour ce type d'applications, ce qui ce traduit ici par :*

$$\mathrm{cov}(X,Y)^2 \leqslant \mathbb{V}(X)\mathbb{V}(Y) \text{ ou encore } |\mathrm{cov}(X,Y)| \leqslant \sigma(X)\sigma(Y)$$

★**Exercice 27.11** *On lance deux fois un dé équilibré dans les mêmes conditions, on note* S *la somme obtenue et* D *la différence (premier moins deuxième). Calculer* $\mathrm{cov}(S,D)$. *Les variables aléatoires* S *et* S *sont-elles indépendantes ?*

Théorème 27.20

Si $X_1, \ldots, X_n$ *sont des variables aléatoires sur* $(\Omega, \mathbb{P})$ *alors :*

$$\mathbb{V}(X_1 + \cdots + X_n) = \sum_{i=1}^{n} \mathbb{V}(X_i) + 2 \sum_{1 \leqslant i < j \leqslant n} \mathrm{cov}(X_i, X_j).$$

Si les variables aléatoires X_i *sont* ***deux à deux indépendantes****, alors :*

$$\mathbb{V}(X_1 + \cdots + X_n) = \sum_{i=1}^{n} \mathbb{V}(X_i).$$

Preuve : La deuxième formule découle de la première, et la première se montre par récurrence sur n. □

Définition 27.18

Lorsque $\mathrm{cov}(X, Y) = 0$ *on sait que les variables aléatoires ne sont pas forcément indépendantes, on dit qu'elles sont* ***non corrélées***.
Plus généralement, lorsque $\mathbb{V}(X)$ *et* $\mathbb{V}(Y)$ *sont non nuls, on sait que* $\frac{\mathrm{cov}(X,Y)}{\sigma(X)\sigma(Y)} \in [-1;1]$, *ce nombre est appelé* ***coefficient de corrélation*** *entre* X *et* Y.

★**Exercice 27.12** *On suppose que* $\mathbb{V}(X)$ *et* $\mathbb{V}(Y)$ *sont non nuls et le coefficient de corrélation vaut* ± 1. *Montrer qu'il existe deux réels* a *et* b *tel que* $\mathbb{P}(Y = aX + b) = 1$, *c'est à dire il est presque sûr que* $Y = aX + b$.

★**Exercice 27.13** *Montrer que deux variables de Bernoulli sont indépendantes si et seulement si elles sont non corrélées.*

V SOLUTION DES EXERCICES

Solution 27.1 $\Omega = [\![1;6]\!]^2$, $X(\Omega) = [\![2;12]\!]$ *et* $\mathbb{P}(X = k) = \mathbb{P}(\{(i, j) \in [\![1;6]\!]^2 \ / \ i + j = k\})$. *On doit avoir* $j = k - i \in [\![1;6]\!]$, *d'où* $k - 6 \leqslant i \leqslant k - 1$ *avec* $i \in [\![1;6]\!]$, *d'où la discussion : si* $k \leqslant 6$, *alors* $\mathbb{P}(X = k) = \sum\limits_{i=1}^{k-1} \mathbb{P}(\{(i, k - i)\}) = \frac{k-1}{36}$, *et si* $7 \leqslant k \leqslant 12$ *alors* $\mathbb{P}(X = k) = \sum\limits_{i=k-6}^{6} \mathbb{P}(\{(i, k - i)\}) = \frac{13-k}{36}$.

Solution 27.2 *Il suffit d'écrire que* $(X \leqslant k) = (X = k) \cup (X \leqslant k - 1)$ *et que* $(X \geqslant k) = (X = k) \cup (X \geqslant k + 1)$ *(réunion d'événements incompatibles).*

Solution 27.3 $X(\Omega) = [\![-5;5]\!]$, *si* $k = i - j$ *alors* $i = k + j \in [\![1;6]\!]$ *d'où* $1 - k \leqslant j \leqslant 6 - k$, *d'où la discussion, si* $k \leqslant 0$, *alors* $\mathbb{P}(X = k) = \sum\limits_{j=1-k}^{6} \mathbb{P}(\{(k + j, j)\}) = \frac{6+k}{36}$, *et si* $k \geqslant 1$, $\mathbb{P}(X = k) = \sum\limits_{j=1}^{6-k} \mathbb{P}(\{(k + j, j)\}) = \frac{6-k}{36}$.

$\mathbb{P}(|X| = k) = \mathbb{P}((X = k) \cup (X = -k)) = \frac{1}{6}$ *si* $k = 0$ *et* $\frac{6-k}{18}$ *si* $k \in [\![1;5]\!]$.

$\mathbb{P}(X^2 = k^2) = \mathbb{P}((X = k) \cup (X = -k)) = \frac{1}{6}$ *et* $\frac{6-k}{12}$ *si* $k \in [\![1;5]\!]$.

Solution 27.4 *Soit* X *le nombre de rencontres, on demande* $\mathbb{E}(X)$, *notons* X_i *la variable aléatoire qui vaut* 1 *si le* i^e *tirage est une rencontre,* 0 *sinon (variable de Bernoulli), on a alors* $X = X_1 + \cdots + X_n$ *et donc* $\mathbb{E}(X) = \sum\limits_{i=1}^{n} \mathbb{E}(X_i) = \sum\limits_{i=1}^{n} \mathbb{P}(X_i = 1)$. *On peut modéliser l'expérience par le choix équiprobable d'une permutation des n boules, on alors* $\mathbb{P}(X_i = 1) = \frac{(n-1)!}{n!} = \frac{1}{n}$ *(autrement dit* $X_i \hookrightarrow \mathcal{B}(\frac{1}{n})$*) et donc* $\mathbb{E}(X) = 1$.

Solution 27.5 $\mathbb{E}(X^2) = \sum\limits_{k=1}^{n} k^2 \frac{1}{n} = \frac{(2n+1)(n+1)}{6}$, *et* $\mathbb{E}(e^X) = \sum\limits_{k=1}^{n} e^k \frac{1}{n} = \frac{e(e^n-1)}{n(e-1)}$.

Solution 27.6

1/ *On doit avoir* $\sum\limits_{k=1}^{n} ak = 1$ *d'où* $a = \frac{2}{n(n+1)}$. $\mathbb{E}(X) = \sum\limits_{k=1}^{n} ak^2 = \frac{2}{n(n+1)} \frac{(2n+1)n(n+1)}{6} = \frac{2n+1}{3}$. $\mathbb{E}(X^2) = \sum\limits_{k=1}^{n} ak^3 = \frac{n(n+1)}{2}$, $\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{n(n+1)}{2} - \frac{(2n+1)^2}{9} = \frac{(n-1)(n+2)}{18}$.

2/ $f(t) = \mathbb{E}(X^2) - 2t\mathbb{E}(X) + t^2 = (t - \mathbb{E}(X))^2 + \mathbb{E}(X^2) - \mathbb{E}(X)^2 = (t - \mathbb{E}(X))^2 + \mathbb{V}(X)$, *cette quantité est minimale lorsque* $t = \mathbb{E}(X)$ *et le minimum est* $\mathbb{V}(X)$.

Solution 27.7 *On a* $X(\Omega) = Y(\Omega) = [\![1;6]\!]$, *pour* $(i, j) \in [\![1;6]\!]$, *on a* $p_{ij} = 0$ *si* $i > j$, $p_{ii} = \frac{1}{36}$ *et* $p_{ij} = \frac{1}{18}$ *si* $i < j$, *d'où la table :*

X \ Y	1	2	3	4	5	6	$p_{i\bullet}$
1	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{11}{36}$
2	0	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{9}{36}$
3	0	0	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{7}{36}$
4	0	0	0	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{5}{36}$
5	0	0	0	0	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{12}$
6	0	0	0	0	0	$\frac{1}{36}$	$\frac{1}{36}$
$p_{\bullet j}$	$\frac{1}{36}$	$\frac{1}{12}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$	1 \ 1

Solution 27.8 *Soit* $Z = X + Y$, *alors* $Z(\Omega) = [\![2\,;2n]\!]$ *et :*

$$\begin{aligned}\mathbb{P}(Z = k) &= \sum_{i=\max(k-n,1)}^{\min(k-1,n)} \mathbb{P}(X = i) \times \mathbb{P}(Y = k - i)\\ &= \frac{\min(k-1,n) - \max(k-n,1) + 1}{n^2} = \frac{n - |n+1-k|}{n^2}\\ &= \begin{cases} \frac{k-1}{n^2} & \textit{si } k \leqslant n+1 \\ \frac{2n+1-k}{n^2} & \textit{sinon}\end{cases}\end{aligned}$$

De même, en posant $H = X - Y$, *on montre que* $H(\Omega) = [\![-(n-1)\,;(n-1)]\!]$ *et* $\mathbb{P}(H = k) = \frac{n-|k|}{n^2}$.

Solution 27.9 *Par récurrence sur* m, *il suffit de montrer le théorème pour* $m = 2$ *: soit* $Z = X_1 + X_2$, *alors* $Z(\Omega) = [\![0\,;n_1 + n_2]\!]$ *et :*

$$\begin{aligned}\mathbb{P}(Z = k) &= \sum_{i=0}^{k} \mathbb{P}([X_1 = i] \cap [X_2 = k - i])\\ &= \sum_{i=0}^{k} \mathbb{P}(X_1 = i) \times \mathbb{P}(X_2 = k - i) \textit{ (indépendance)}\\ &= \sum_{i=0}^{k} \binom{n_1}{i} p^i (1-p)^{n_1-i} \binom{n_2}{k-i} p^{k-i}(1-p)^{n_2-k+i} \textit{ (en convenant que } \binom{n}{k} = 0 \textit{ si } k > n \textit{ ou } k < 0)\\ &= \sum_{i=0}^{k} \binom{n_1}{i} p^i (1-p)^{n_1-i} \binom{n_2}{k-i} p^{k-i}(1-p)^{n_2-k+i}\\ &= \sum_{i=0}^{k} \binom{n_1}{i}\binom{n_2}{k-i} p^k (1-p)^{n_1+n_2-k}\\ &= p^k(1-p)^{n_1+n_2-k} \sum_{i=0}^{k} \binom{n_1}{i}\binom{n_2}{k-i}\\ &= \binom{n_1+n_2}{k} p^k (1-p)^{n_1+n_2-k} \textit{ (formule de Vandermonde)}\end{aligned}$$

donc $X_1 + X_2 \hookrightarrow \mathcal{B}(n_1 + n_2, p)$. *Le passage du rang* m *au rang* $m+1$ *se ramène au rang 2 en posant* $X_2' = X_2 + \cdots + X_{m+1}$.

Solution 27.10 *Ces variables ne sont pas indépendantes, car* $\mathbb{P}([X = 1] \cap [Y = 1]) = \mathbb{P}([X = 1] \cap [X = 0]) = 0$ *alors que* $\mathbb{P}(X = 1) \times \mathbb{P}(Y = 1) = \frac{1}{9}$. *D'autre part* $\mathbb{E}(X) = 0$, $\mathbb{E}(Y) = \frac{1}{3}$, *et* $\mathbb{E}(XY) = 0$ *car XY est une variable certaine égale à* 0, *on a donc* $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.

Solution 27.11 *Soit* X *le premier résultat et* Y *le second, alors* $S = X+Y$ *et* $D = X-Y$, *on calcule donc* $\operatorname{cov}(X+Y, X-Y) = \mathbb{V}(X) - \operatorname{cov}(X,Y) + \operatorname{cov}(Y,X) - \mathbb{V}(Y) = 0$ *car* X *et* Y *suivent la même loi uniforme sur* $[\![1\,;6]\!]$. *Les deux variables aléatoires ne sont pas indépendantes car par exemple* $\mathbb{P}([S = 3] \cap [D = 0]) = 0$ *alors que* $\mathbb{P}(S = 3)$ *et* $\mathbb{P}(D = 0)$ *ne sont pas nulles.*

Solution 27.12 *Reprendre la démonstration du cas d'égalité de Cauchy-Schwarz vue dans les espaces euclidiens en partant de* $\mathbb{V}(\lambda X + Y) \geqslant 0$ *pour tout réel* λ.

Solution 27.13 *Il y a déjà un sens connu : indépendantes implique non corrélées. Supposons* X *et* Y *non corrélées et suivant une loi de Bernoulli de paramètres respectifs* p_1 *et* p_2, *alors* XY *suit une loi de Bernoulli, de paramètre* $p = \mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y) = p_1 p_2$, *par conséquent* $\mathbb{P}([X = 1] \cap [Y = 1]) = \mathbb{P}(X = 1) \times \mathbb{P}(Y = 1)$, *les événements* $(X = 1)$ *et* $(Y = 1)$ *sont indépendants, donc* $(X = 0)$ *et* $(Y = 1)$ *aussi etc. Les deux variables aléatoires sont indépendantes.*

Chapitre 28

Espaces euclidiens

Sommaire

Dans tout le chapitre, E désigne un $\mathbb{R}$-espace vectoriel.

I PRODUIT SCALAIRE

1) Définitions

Définition 28.1

Un produit scalaire sur E *est une forme bilinéaire sur* E*, généralement notée* $(.|.)$*, qui à tout couple de vecteurs* (x,y) *associe le réel* $(x|y)$*, et qui vérifie :*
- $\forall\, x,y \in E, (x|y) = (y|x)$ *(symétrie).*
- $\forall\, x \in E, (x|x) \geqslant 0$ *(positive).*
- $\forall\, x \in E$*, si* $(x|x) = 0$*, alors* $x = 0$ *(définie).*

Lorsque E *est muni d'un produit scalaire* $(.|.)$*, on dit que* $(E,(.|.))$ *est un espace euclidien s'il est de dimension finie, ou un espace préhilbertien réel sinon.*

☞**Exemples :**
- Produit scalaire canonique de $\mathbb{R}^n$: $(x|y) = \sum_{i=1}^{n} x_i y_i$.
- $E = \mathcal{C}^0([a;b],\mathbb{R})$ et $\forall\, f,g \in E, (f|g) = \int_a^b f(t)g(t)\,dt$ (avec $a < b$).
- $E = \mathbb{R}[X]$, on définit un produit scalaire sur E en posant :

$$\forall\, P,Q \in E, (P|Q) = \int_a^b P(t)Q(t)\,dt \text{ (avec } a < b)$$

- Pour $x,y \in \mathbb{R}^2$, $\varphi(x,y) = x_1y_1 + x_1y_2 + x_2y_1 + 2x_2y_2$ est un produit scalaire sur $\mathbb{R}^2$, mais pas $\psi(x,y) = x_1y_1 + x_1y_2 + x_2y_1 + x_2y_2$.
- $E = \mathbb{R}_n[X]$, soient $a_0, \dots, a_n$ des réels distincts, on définit un produit scalaire sur E en posant :

$$\forall\, P,Q \in E, (P|Q) = \sum_{k=0}^{n} P(a_k)Q(a_k)$$

2) Cauchy-Schwarz

Théorème 28.1 (Inégalité de Cauchy-Schwarz)
$\forall\, x,y \in E, (x|y)^2 \leqslant (x|x)(y|y)$. De plus :
$(x|y)^2 = (x|x)(y|y) \iff (x,y)$ est liée (cas d'égalité).

Preuve : Si y est nul il n'y a rien à démontrer (même pour l'égalité) car on a forcément $(x|y) = 0$ par linéarité à droite. Supposons y non nul, alors $\forall\, \lambda \in \mathbb{R}, (x+\lambda y|x+\lambda y) \geqslant 0$, ce qui donne en développant : $\lambda^2(y|y) + 2\lambda(x|y) + (x|x) \geqslant 0$, on a un trinôme du second degré en λ de signe constant, le discriminant doit être négatif ou nul (s'il était strictement positif il y aurait deux racines réelles distinctes $r_1 < r_2$ et le trinôme serait strictement négatif dans l'intervalle ouvert $]r_1;r_2[$), ce qui donne l'inégalité.

S'il y a l'égalité alors cela signifie que le discriminant est nul, donc le trinôme admet une racine double λ_0, et donc $(x+\lambda_0 y|x+\lambda_0 y) = 0$, ce qui entraîne $x + \lambda_0 y = 0$ et donc la famille (x,y) est liée.

Réciproquement, si la famille est liée, il existe $\lambda \in \mathbb{R}$ tel que $x = \lambda y$, d'où $(x|y)^2 = \lambda^2(y|y)^2 = (\lambda y|\lambda y)(y|y) = (x|x)(y|y)$. □

Définition 28.2 (norme euclidienne)
*Soit $x \in (E, (.|.))$, on pose $\|x\| = \sqrt{(x|x)}$, c'est la norme euclidienne de x. Un vecteur de norme égale à 1 est dit **unitaire**. L'inégalité de Cauchy-Schwarz peut alors s'écrire : $|(x \mid y)| \leqslant \|x\|\|y\|$.*

Théorème 28.2 (Propriétés)
- $\|x\| = 0 \iff x = 0$.
- $\forall\, \lambda \in \mathbb{R}, \|\lambda x\| = |\lambda|\|x\|$.
- $\|x+y\| \leqslant \|x\| + \|y\|$ *(inégalité triangulaire).*

Preuve : Celle-ci est simple et laissée en exercice. □

Remarque 28.1 – *Si x est non nul alors le vecteur $\frac{1}{\|x\|}x$ est unitaire.*

Définition 28.3 (écart angulaire)
*Soient x et y deux vecteurs non nuls, on a $\frac{|(x|y)|}{\|x\|\|y\|} \leqslant 1$, donc en posant $\theta = \arccos\left(\frac{(x|y)}{\|x\|\|y\|}\right) \in [0;\pi]$, on peut écrire : $(x \mid y) = \|x\|\|y\|\cos(\theta)$, ce réel θ est appelé **écart angulaire** entre les vecteurs x et y. Lorsque $\theta \in [0;\frac{\pi}{2}]$ (c'est à dire $(x \mid y) \geqslant 0$) on dit que les vecteurs sont dans le même sens, et lorsque $\theta \in [\frac{\pi}{2};\pi]$ (c'est à dire $(x \mid y) \leqslant 0$) on dit que les vecteurs sont dans le sens contraire.*

★**Exercice 28.1** *Soient $x,y \in E$ deux vecteurs non nuls, montrer que :*

1/ $|\|x\| - \|y\|| \leqslant \|x+y\|$.

2/ x et y sont colinéaires si et seulement si leur écart angulaire vaut 0 ou π (pour x et y non nuls).

3/ $\|x+y\| = \|x\| + \|y\| \iff \exists\, \alpha > 0, x = \alpha y$.

☞**Exemples** :
- $E = \mathbb{R}^n$, avec le produit scalaire canonique, l'inégalité de Cauchy-Schwarz s'écrit :

$$\left(\sum_{i=1}^{n} x_iy_i\right)^2 \leqslant \left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} y_i^2\right) \text{ et } \|x\| = \sqrt{\sum_{i=1}^{n} x_i^2}$$

- $E = \mathcal{C}^0([a;b],\mathbb{R})$ avec le produit scalaire : $(f|g) = \int_a^b f(t)g(t)\,dt$, l'inégalité de Cauchy-Schwarz s'écrit : $\left(\int_a^b f(t)fg(t)\,dt\right)^2 \leqslant \left(\int_a^b f^2\right)\left(\int_a^b g^2\right)$ et $\|f\| = \sqrt{\int_a^b f^2}$.

À retenir : Relations entre le produit scalaire et la norme

- $\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2(x|y)$ et $\|x-y\|^2 = \|x\|^2 + \|y\|^2 - 2(x|y)$.
- $\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2)$ (théorème de la médiane ou identité du parallélogramme).
- $(x|y) = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right) = \frac{1}{2}\left(\|x+y\|^2 - \|x\|^2 - \|y\|^2\right) = \frac{1}{2}\left(\|x\|^2 + \|y\|^2 - \|x-y\|^2\right)$ (identités de polarisation).

II ORTHOGONALITÉ

1) Définition

Définition 28.4

Soient $x, y \in \mathrm{E}$, et soient F, G deux s.e.v de E, on dit que :
• x et y sont orthogonaux lorsque $(x|y) = 0$.
• F et G sont orthogonaux lorsque $\forall\, x \in \mathrm{F}, \forall\, y \in \mathrm{G}, (x|y) = 0$.
*On appelle **orthogonal de** A (une partie de E), l'ensemble des vecteurs de E orthogonaux à tous les vecteurs de A, notation : $\mathrm{A}^\perp = \{x \in \mathrm{E} \ /\ \forall\, y \in \mathrm{A}, (x|y) = 0\}$. On remarquera que dire que F et G sont orthogonaux équivaut à $\mathrm{F} \subset \mathrm{G}^\perp$, ou encore $\mathrm{G} \subset \mathrm{F}^\perp$.*

Remarque 28.2 :

- *Le seul vecteur orthogonal à tous les autres est le vecteur nul, i.e. $\mathrm{E}^\perp = \{0\}$, car le produit scalaire est défini.*
- *Deux vecteurs non nuls sont orthogonaux si et seulement si leur écart angulaire vaut $\frac{\pi}{2}$.*

Théorème 28.3 (Premières propriétés)

• Si A est une partie de E (non vide) alors $\mathrm{A}^\perp$ est un sev de E, et $\mathrm{A}^\perp = \mathrm{Vect}\,[\mathrm{A}]^\perp$.
• Si A et B sont deux parties de E telles que $\mathrm{A} \subset \mathrm{B}$, alors $\mathrm{B}^\perp \subset \mathrm{A}^\perp$.
• Si A est une partie de E, alors $\mathrm{A} \subset (\mathrm{A}^\perp)^\perp$.
• Si F et G sont deux sev de E, alors $(\mathrm{F}+\mathrm{G})^\perp = \mathrm{F}^\perp \cap \mathrm{G}^\perp$.

Preuve : Celle-ci est simple et laissée en exercice. □

Théorème 28.4 (de Pythagore)

Deux vecteurs x et y sont orthogonaux si et seulement si $\|x+y\|^2 = \|x\|^2 + \|y\|^2$.

Preuve : On a $\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2(x|y)$, l'équivalence en découle. □

Théorème 28.5

Si F est un s.e.v de E, alors $\mathrm{F}^\perp$ est un s.e.v de E en somme directe avec F.

Preuve : Pour $y \in \mathrm{E}$, on pose $f_y : \mathrm{E} \to \mathbb{R}$ définie par $f_y(x) = (x|y)$, alors f_y est une forme linéaire sur E, et il est facile de voir que $\mathrm{F}^\perp = \bigcap\limits_{y \in \mathrm{F}} \ker(f_y)$, ce qui prouve que $\mathrm{F}^\perp$ est un s.e.v de E. Si $x \in \mathrm{F} \cap \mathrm{F}^\perp$, alors on doit avoir $(x|x) = 0$, d'où $x = 0$. □

Théorème 28.6

Si E est un espace euclidien de dimension n, et si F est un s.e.v de E de dimension p, alors $\dim(\mathrm{F}^\perp) = n - p$, on a donc :

$$\mathrm{E} = \mathrm{F} \oplus \mathrm{F}^\perp.$$

Preuve : On sait que $\dim(\mathrm{F} \oplus \mathrm{F}^\perp) \leqslant n$, d'où $\dim(\mathrm{F}^\perp) \leqslant n - p$.

Soit $f : \mathrm{E} \to \mathbb{R}^p$ l'application définie par $f(x) = ((e_1|x), \ldots, (e_p|x))$ où $\mathrm{B} = (e_1, \ldots, e_p)$ désigne une base de F, alors il est facile de voir que f est linéaire et que $\ker(f) = \mathrm{F}^\perp$. D'après le théorème du rang, on a $n = \dim(\mathrm{F}^\perp) + \mathrm{rg}(f) \leqslant \dim(\mathrm{F}^\perp) + p$, ce qui donne $\dim(\mathrm{F}^\perp) \geqslant n - p$, et donc $\dim(\mathrm{F}^\perp) = n - p$. □

Attention !

Ce théorème est faux en dimension infinie. Par exemple dans $E = \mathbb{R}[X]$ *avec le produit scalaire* $(P \mid Q) = \int_0^1 P(t)Q(t)\,dt$*, l'orthogonal de* $F = \mathrm{Vect}\left[X, X^2, \dots\right]$ *est réduit au polynôme nul alors que* $F \neq E$.

À retenir : Quelques conséquences

Soit E un espace euclidien, F et G deux sev de E :
- $(F^{\perp})^{\perp} = F$.
- $(F \cap G)^{\perp} = F^{\perp} + G^{\perp}$.

2) Bases orthonormales dans un euclidien

Définition 28.5

Une famille $(e_1, \dots, e_p)$ *de* E *est dite orthonormale lorsque* $\forall\, i, j \in [\![1\,;p]\!], (e_i|e_j) = \delta_{ij}$*. Cette famille est dite orthogonale lorsque* $\forall\, i, j \in [\![1\,;p]\!], i \neq j \implies (e_i|e_j) = 0$.

Théorème 28.7

Une famille orthogonale ne contenant pas le vecteur nul est libre. En particulier, une famille orthonormale est libre.
Si $(e_1, \dots, e_p)$ *est une famille orthogonale, alors :* $\|\sum_{k=1}^{p} e_i\|^2 = \sum_{i=1}^{p} \|e_i\|^2$.

Preuve : Soit $(e_1, \dots, e_p)$ une famille orthogonale ne contenant pas le vecteur nul, si $\sum_{k=1}^{p} \lambda_k e_k = 0$, alors soit $i \in [\![1\,;p]\!]$, on a $(e_i|\sum_{k=1}^{p} \lambda_k e_k) = \sum_{k=1}^{p} \lambda_k (e_i|e_k) = \lambda_i \|e_i\|^2 = 0$, ce qui entraîne $\lambda_i = 0$.

Pour la deuxième partie, on a $\|\sum_{i=1}^{p} e_i\|^2 = \sum_{i,j=1}^{p} (e_i|e_j) = \sum_{i=1}^{p} \|e_i\|^2$. □

Cas particulier

Si E est un euclidien de dimension n, alors une famille orthonormale de n vecteurs est une base de E, on dit que l'on a une **base orthonormale** (b.o.n en abrégé). Par exemple, la base canonique que $\mathbb{R}^n$ est une base orthonormale pour le produit scalaire canonique.

Théorème 28.8 (interêt d'une b.o.n)

Soit $\mathcal{B} = (e_1, \dots, e_n)$ *une b.o.n de* E*, alors* $\forall\, x, y \in E$ *:*
- $x = \sum_{i=1}^{n} (x|e_i) e_i$*, autrement dit, la coordonnée de* x *sur* e_i *est* $(x|e_i)$ *;*
- $(x|y) = \sum_{i=1}^{n} x_i y_i$*, expression analogue au produit scalaire canonique de* $\mathbb{R}^n$ *;*
- $\|x\|^2 = \sum_{i=1}^{n} x_i^2$.

avec $x_i = (x|e_i)$ *et* $y_i = (y|e_i)$ *les coordonnés de* x *et de* y.

Preuve : Soit $\mathrm{Coord}_{\mathcal{B}}(x) = (\lambda_1, \dots, \lambda_n)$, on a $(x|e_k) = (\sum_{i=1}^{n} \lambda_i e_i | e_k) = \sum_{i=1}^{n} \lambda_i (e_i|e_k) = \lambda_k$. Pour les deux autres points, il suffit de développer le produit scalaire. □

Attention !

Le théorème ci-dessus est fondamental et démontre l'intérêt des b.o.n. dans les espaces euclidiens.

Théorème 28.9

Dans un espace euclidien, il existe toujours des bases orthonormales.

Preuve : Par récurrence sur $n = \dim(E)$: pour $n = 1$, on a $E = \mathrm{Vect}\,[e_1]$, une b.o.n de E est (e_1') avec $e_1' = \frac{e_1}{\|e_1\|}$.

Supposons le théorème vrai au rang $n-1$ ($n \geqslant 1$), et soit e_1 un vecteur unitaire de E, soit $F = \mathrm{Vect}\,[e_1]^{\perp}$, alors F est un s.e.v de dimension $n-1$, soit $(e_2, \dots, e_n)$ une b.o.n de F, il est facile de voir que $(e_1, e_2, \dots, e_n)$ est une b.o.n de E. □

Théorème 28.10 (représentation des formes linéaires sur E**)**

Soit E *un euclidien et* $f : \mathrm{E} \to \mathbb{R}$ *une forme linéaire, alors il existe un unique vecteur* $a \in \mathrm{E}$ *tel que* $\forall\, x \in \mathrm{E}, f(x) = (a|x)$.

Preuve : Pour l'existence : si f est nulle alors on peut prendre $a = 0$. Si f est non nulle, alors $\ker(f)$ est un hyperplan de E, donc $\ker(f)^{\perp} = \mathrm{Vect}\,[u]$ est une droite vectorielle. Posons $\lambda = f(u)$ et prenons $a = \frac{\lambda}{\|u\|^2} u$. Il est facile de vérifier que pour tout $x \in \mathrm{E}, f(x) = (a|x)$.

Pour l'unicité : si b est un autre vecteur qui convient, alors $\forall\, x \in \mathrm{E}, (a-b|x) = 0$, donc $a - b = 0$. □

III PROJECTIONS ORTHOGONALES

1) Définition

Définition 28.6

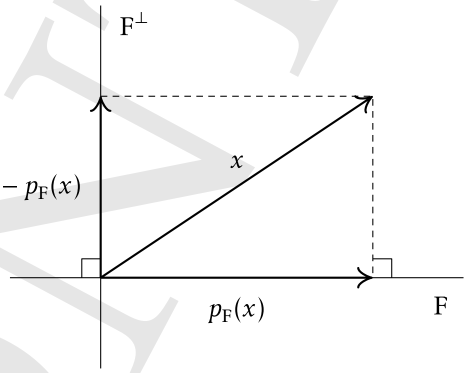
Soit E *un préhilbertien réel et* $p \in \mathcal{L}(\mathrm{E})$ *une projection* ($p \circ p = p$)*, on dit que* p *est une projection orthogonale lorsque* $\ker(p)$ *et* $\ker(p - \mathrm{id})$ *sont orthogonaux. Si* F *est un s.e.v de* E *tel que* $\mathrm{E} = \mathrm{F} \oplus \mathrm{F}^{\perp}$*, la projection orthogonale sur* F*, notée* p_{F}*, est la projection sur* F *parallèlement à* $\mathrm{F}^{\perp}$.

Remarque 28.3 – *Si* F *est un s.e.v de* E *tel que* $\mathrm{E} = \mathrm{F} \oplus \mathrm{F}^{\perp}$*, alors* $(\mathrm{F}^{\perp})^{\perp} = \mathrm{F}$ *et la projection orthogonale sur* $\mathrm{F}^{\perp}$ *est* $\mathrm{id} - p_{\mathrm{F}}$.

Théorème 28.11 (projeté orthogonal sur un s.e.v.)

Soit E *un préhilbertien réel et* F *un s.e.v de* E*. Si* F *est de dimension finie, alors* $\mathrm{E} = \mathrm{F} \oplus \mathrm{F}^{\perp}$*, et si* $(e_1, \dots, e_p)$ *est une b.o.n de* F*, alors* $\forall\, x \in \mathrm{E}, p_{\mathrm{F}}(x) = \sum_{i=1}^{p} (x|e_i)e_i$.

Preuve : Soit $x \in \mathrm{E}$, posons $y = \sum_{i=1}^{p} (x|e_i)e_i$, alors $y \in \mathrm{F}$, on a pour tout $j \in [\![1\,;p]\!]$, $(y|e_j) = \sum_{i=1}^{p} (x|e_i)(e_i|e_j) = (x|e_j)$, on en déduit que $(x - y|e_j) = 0$ et donc $x - y \in \mathrm{F}^{\perp}$, or $x = y + (x - y)$ ce qui prouve que $\mathrm{E} = \mathrm{F} + \mathrm{F}^{\perp}$, comme on sait que la somme est directe, on a $\mathrm{E} = \mathrm{F} \oplus \mathrm{F}^{\perp}$ et $p_{\mathrm{F}}(x) = y = \sum_{i=1}^{p} (x|e_i)e_i$. □



☞**Exemples** :
- Si $\mathrm{D} = \mathrm{Vect}\,[u]$ est une droite vectorielle de E, alors $(e_1 = \frac{u}{\|u\|})$ est une b.o.n de D, donc $\forall\, x \in \mathrm{E}, p_{\mathrm{D}}(x) = (x|e_1)e_1$, c'est à dire $p_{\mathrm{D}}(x) = \frac{(x|u)}{\|u\|^2} u$.
- Si E est euclidien et H un hyperplan, alors $\mathrm{H}^{\perp}$ est de dimension 1, donc il existe un vecteur non nul n tel que $\mathrm{H}^{\perp} = \mathrm{Vect}\,[n] = \mathrm{D}$ et $p_{\mathrm{H}}(x) = x - p_{\mathrm{D}}(x) = x - \frac{(x|n)}{\|n\|^2} n$ (le vecteur n est appelé un vecteur normal à H).

2) Orthonormalisation

Théorème 28.12 (procédé d'orthonormalisation de Schmidt [1])

Soit E *un espace préhilbertien réel, et soit* $(e_1,\dots,e_n)$ *une famille libre de* E*, alors il existe une unique famille orthonormale* $(v_1,\dots,v_n)$ *de* E *telle que :*

$$\forall\, i \in [\![1\,;n]\!], \left\{\begin{array}{l} \mathrm{Vect}\,[e_1,\dots,e_i] = \mathrm{Vect}\,[v_1,\dots,v_i] \\ (e_i|v_i) > 0 \end{array}\right. .$$

Preuve : On pose $v_1 = \frac{e_1}{\|e_1\|}$, on a bien $\mathrm{Vect}\,[e_1] = \mathrm{Vect}\,[v_1]$ et $(e_1|v_1) > 0$.
Supposons les vecteurs $v_1,\dots,v_i$ construits et vérifiant les conditions, on pose $e'_{i+1} = p_{\mathrm{F}_i^\perp}(e_{i+1})$ où $\mathrm{F}_i = \mathrm{Vect}\,[v_1,\dots,v_i] = \mathrm{Vect}\,[e_1,\dots,e_i]$, ce qui donne $e'_{i+1} = e_{i+1} - \sum\limits_{k=1}^{i}(e_{i+1}|v_k)v_k$, ce vecteur e'_{i+1} est non nul et dans F_{i+1}, on pose ensuite $v_{i+1} = \frac{e'_{i+1}}{\|e'_{i+1}\|}$, il est facile de voir que $\mathrm{Vect}\,[e_1,\dots,e_{i+1}] = \mathrm{Vect}\,[v_1,\dots,v_{i+1}]$. D'autre part, $(e_{i+1}|v_{i+1}) = (e'_{i+1}|v_{i+1}) = \|e'_{i+1}\| > 0$.

On remarque qu'à chaque étape, il y a deux choix pour v_i, mais la condition $(e_i|v_i) > 0$ élimine une des deux possibilités, ce qui entraîne l'unicité, car on doit prendre e'_{i+1} dans $\mathrm{F}_{i+1} \cap \mathrm{F}_i^\perp$ qui est une droite vectorielle. □

Remarque 28.4 – *Si* $(e_1,\dots,e_n)$ *une base de* E*, alors la famille* $(v_1,\dots,v_n)$ *obtenue par la méthode Schmidt est une b.o.n de* E.

★**Exercice 28.2** *Soit* $\mathrm{E} = \mathbb{R}^3$*, muni du produit scalaire canonique, on pose* $v_1 = (1,1,0)$*,* $v_2 = (1,0,1)$ *et* $v_3 = (0,1,1)$*. Appliquer la méthode de Schmidt à la base* (v_1,v_2,v_3)*.*

Dans la suite, $(\mathrm{E},(.|.))$ désigne un espace euclidien.

3) Distance d'un vecteur à un s.e.v

Soit F un s.e.v de E et soit $x \in \mathrm{E}$, pour tout vecteur $y \in \mathrm{F}$, on a :

$$\|x-y\|^2 = \|(p_\mathrm{F}(x)-y) + p_{\mathrm{F}^\perp}(x)\|^2 = \|p_\mathrm{F}(x)-y\|^2 + \|p_{\mathrm{F}^\perp}(x)\|^2$$

on voit donc que $\forall\, y \in \mathrm{E}$, $\|x-y\|^2 \geqslant \|p_{\mathrm{F}^\perp}(x)\|^2$, et que cette valeur est un minimum atteint uniquement pour $y = p_\mathrm{F}(x)$,

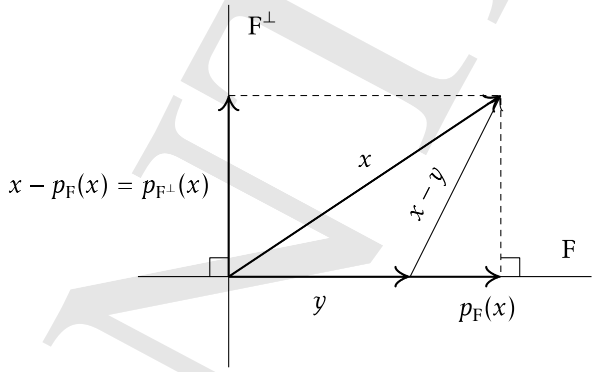


Figure 28.1 – Projection orthogonale et distance à F

d'où le théorème :

Théorème 28.13

Soit F *un s.e.v de* E*, pour* $x \in \mathrm{E}$*, l'ensemble* $\{\|x-y\| \;/\; y \in \mathrm{F}\}$ *admet un minimum, celui-ci est atteint uniquement pour* $y = p_\mathrm{F}(x)$*, et vaut* $\|p_{\mathrm{F}^\perp}(x)\|$*. Ce minimum est appelé distance de* x *à* F *et noté* $\mathrm{d}(x,\mathrm{F})$ *:* $\mathrm{d}(x,\mathrm{F}) = \min\limits_{y\in\mathrm{F}}\|x-y\| = \|p_{\mathrm{F}^\perp}(x)\| = \|x - p_\mathrm{F}(x)\|$.

☞**Exemples** :

– Distance d'un vecteur à une droite : soit $\mathrm{D} = \mathrm{Vect}\,[u]$ une droite vectorielle, on sait que $p_\mathrm{D}(x) = \frac{(x|u)}{\|u\|^2}u$, d'où $\mathrm{d}(x,\mathrm{D}) = \sqrt{\|x-p_\mathrm{D}(x)\|^2} = \sqrt{\|x\|^2 - \frac{(x|u)^2}{\|u\|^2}}$.

1. SCHMIDT Erhard (1876 – 1959) : mathématicien allemand.

– Distance d'un vecteur à un hyperplan : soit H un hyperplan de E, alors $H^\perp = \mathrm{Vect}[u]$ est une droite vectorielle, d'où $d(x, H) = \|p_D(x)\| = \frac{|(x|u)|}{\|u\|}$. Si $\mathcal{B}$ est une b.o.n de E et si u a pour coordonnées $(a_1, \dots, a_n)$ dans la base $\mathcal{B}$ alors un vecteur x de coordonnées dans la base $\mathcal{B}$ $(x_1, \dots, x_n)$ appartient à H si et seulement si $(x|u) = 0$, c'est à dire $x \in H \iff a_1x_1 + \cdots a_nx_n = 0$ (équation cartésienne de H dans la base $\mathcal{B}$). On remarque alors que $d(x, H) = \frac{|a_1x_1+\cdots a_nx_n|}{\sqrt{a_1^2+\cdots+a_n^2}}$. On remarquera que les coordonnées d'un vecteur normal à un hyperplan se lisent directement sur une équation cartésienne de cet hyperplan dans une b.o.n. (ce sont les coefficients).

IV SOUS-ESPACES AFFINES

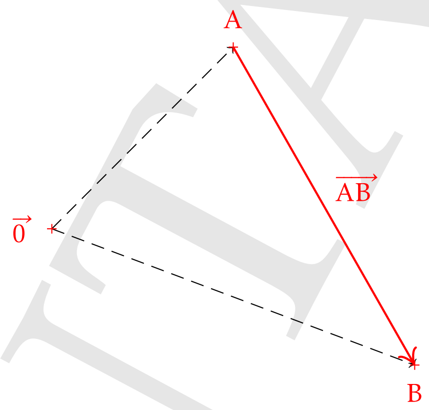
1) Définition

Soit $\vec{u} \in E$, la translation de vecteur $\vec{u}$ est l'application [2] $t_{\vec{u}} : E \to E$ définie par $t_{\vec{u}}(\vec{v}) = \vec{u} + \vec{v}$. L'ensemble de ces applications est noté $\mathcal{T}_E$ et il est facile de vérifier que $(\mathcal{T}_E, \circ)$ est un groupe abélien, c'est un sous-groupe du groupe des permutations de E.

Si $\vec{v}, \vec{w}$ sont deux vecteurs de E, alors il existe un unique vecteur $\vec{u} \in E$ tel que $t_{\vec{u}}(\vec{v}) = \vec{w}$. Cette propriété très simple, suggère un autre point de vue pour les éléments de E : la notion de points.

La propriété précédente peut alors s'énoncer sous la forme suivante : si A et B sont deux points de E, alors il existe un unique vecteur $\vec{u}$ tel que $B = t_{\vec{u}}(A)$. Ce vecteur $\vec{u}$ est noté : $\overrightarrow{AB}$, on remarquera que :

– $\overrightarrow{AB} = B - A$, et $A + \overrightarrow{AB} = B$,
– $\overrightarrow{AB} = 0 \iff A = B$,
– $\overrightarrow{BA} = -\overrightarrow{AB}$,
– $\overrightarrow{AC} = \overrightarrow{AB} + \overrightarrow{BC}$.



Définition 28.7
Un sous - espace affine de E est l'image d'un s.e.v. de E par une translation. Soit $\mathcal{V}$ une partie de E, $\mathcal{V}$ est un s.e.a. de E si et seulement si il existe un s.e.v V et un vecteur $\vec{u}$ tel que $\mathcal{V} = t_{\vec{u}} < V >$, si c'est le cas, V est appelé direction du s.e.a. $\mathcal{V}$ (ce que l'on écrira : $(\mathcal{V}, V)$), et on pose $\dim(\mathcal{V}) = \dim(V)$. On remarquera qu'un s.e.v. est un s.e.a. de direction lui-même.

2) Propriétés

– Si $\mathcal{V}$ est un s.e.a. de direction V, alors $\forall A \in \mathcal{V}$, $\mathcal{V} = t_A(V)$.
En effet : il existe $u \in E$ tel que $\mathcal{V} = t_u(V)$. Soit $A \in \mathcal{V}$, il existe $x_0 \in V$ tel que $A = u + x_0$. Si $x \in V$, alors $A + x = u + (x_0 + x) \in t_u(V)$, donc $t_A(V) \subset t_u(V) = \mathcal{V}$.
Réciproquement, si $x \in V$, alors $u + x = A + (-x_0 + x) \in t_A(V)$, donc $\mathcal{V} = t_u(V) \subset t_A(V)$, d'où l'égalité.
– Conséquence 1 : si $\mathcal{V}$ est un s.e.a. de direction V, alors V est unique (mais pas le vecteur de la translation), de plus $\mathcal{V}$ est un s.e.v. si et seulement si $\mathcal{V}$ contient le vecteur nul.

2. non linéaire si $\vec{u} \neq \vec{0}$.

- Conséquence 2 : si $(\mathcal{V}, \mathrm{V})$ est un s.e.a. alors $\mathrm{V} = \{\overrightarrow{\mathrm{AB}} \ / \ \mathrm{B} \in \mathcal{V}\}$, et $\forall\, \mathrm{B} \in \mathrm{E}, \mathrm{B} \in \mathcal{V} \iff \overrightarrow{\mathrm{AB}} \in \mathrm{V}$.
- Si $(\mathcal{V}, \mathrm{V})$ et $(\mathcal{V}', \mathrm{V}')$ sont deux s.e.a. de E, et si $\mathcal{V} \cap \mathcal{V}'$ n'est pas vide, alors $\mathcal{V} \cap \mathcal{V}'$ est un s.e.a de direction $\mathrm{V} \cap \mathrm{V}'$.
- Soient $(\mathcal{V}, \mathrm{V})$ et $(\mathcal{V}', \mathrm{V}')$ sont deux s.e.a. de E, soient $\mathrm{A} \in \mathcal{V}$ et $\mathrm{A}' \in \mathcal{V}'$, alors $\mathcal{V} \cap \mathcal{V}' \neq \varnothing \iff \overrightarrow{\mathrm{AA}'} \in \mathrm{V} + \mathrm{V}'$. On remarquera que la condition est nécessairement remplie lorsque $\mathrm{E} = \mathrm{V} + \mathrm{V}'$.

★**Exercice** 28.3 *Étudier les s.e.a. de* E *lorsque* $\dim(\mathrm{E}) = 3$.

Définition 28.8 (Repères cartésiens)

Un repère cartésien $\mathcal{R} = (\mathrm{O}, \overrightarrow{e_1}, \dots, \overrightarrow{e_n})$ de E, est la donnée d'un point O de E (appelé origine *du repère), et d'une base $\mathcal{B} = (\overrightarrow{e_1}, \dots, \overrightarrow{e_n})$ de E. Pour tout point M de E, on appelle* coordonnées *de M dans le repère $\mathcal{R}$, les coordonnées du vecteur $\overrightarrow{\mathrm{OM}}$ dans la base $\mathcal{B}$. On remarquera qu'il s'agit des coordonnées de $\mathrm{M} - \mathrm{O}$ dans la base $\mathcal{B}$.*

Propriété : si A et B ont pour coordonnées respectives $(u_1, \dots, u_n)$ et $(v_1, \dots, v_n)$ dans le repère $\mathcal{R}$, alors les coordonnées du vecteur $\overrightarrow{\mathrm{AB}}$ dans la base $\mathcal{B}$ sont $(v_1 - u_1, \dots, v_n - u_n)$. Il suffit pour cela d'écrire $\overrightarrow{\mathrm{AB}} = \overrightarrow{\mathrm{OB}} - \overrightarrow{\mathrm{OA}}$.

3) Hyperplans affines

Équation(s) cartésienne(s) d'un hyperplan affine : Soit $\mathcal{R} = (\mathrm{O}, \mathcal{B})$ un repère de E, avec $\dim(\mathrm{E}) = n$, soit $(\mathcal{H}, \mathrm{H})$ un hyperplan affine de E, soit $a_1x_1 + \cdots + a_nx_n = 0$ une équation de H dans la base $\mathcal{B}$ (un des a_i au moins est non nul), soit $\mathrm{A}(u_1, \dots, u_n) \in \mathcal{H}$. Un point M de coordonnées $(x_1, \dots, x_n)$ appartient à $\mathcal{H}$ si et seulement si $\overrightarrow{\mathrm{AM}} \in \mathrm{H}$, ce qui revient à dire que $a_1(x_1 - u_1) + \cdots + a_n(x_n - u_n) = 0$, on obtient une équation cartésienne de $\mathcal{H}$ de la forme $a_1x_1 + \cdots + a_nx_n = b$.

★**Exercice** 28.4 *Montrer que la réciproque est vraie, c'est à dire que les points de coordonnées $(x_1, \dots, x_n)$ vérifiant une équation du type $a_1x_1 + \cdots + a_nx_n = b$ avec au moins un des coefficients a_i non nul, forment un hyperplan affine de direction l'hyperplan vectoriel d'équation cartésienne $a_1x_1 + \cdots + a_nx_n = 0$.*

Hyperplans affines dans un euclidien

On se place dans un repère $\mathcal{R} = (\mathrm{O}, \mathfrak{B})$ orthonormal (ce qui signifie que $\mathfrak{B}$ est un b.o.n. de E).

Soit $\mathcal{H}$ un hyperplan affine de direction H, on sait que $\mathcal{H}$ admet une équation cartésienne de la forme $\alpha_1x_1 + \cdots + \alpha_nx_n = c$ où les α_i sont non tous nuls, soit $n(\alpha_1, \dots, \alpha_n)$, et soit $\mathrm{A}(a_1, \cdots, a_n)$ un point de $\mathcal{H}$, alors :

$$\begin{aligned} \mathrm{M}(x_1, \dots, x_n) \in \mathcal{H} &\iff \alpha_1x_1 + \cdots + \alpha_nx_n = c \\ &\iff \alpha_1x_1 + \cdots + \alpha_nx_n = \alpha_1a_1 + \cdots + \alpha_na_n \ (\text{car } \mathrm{A} \in \mathcal{H}) \\ &\iff \alpha_1(x_1 - a_1) + \cdots + \alpha_n(x_n - a_n) = 0 \\ &\iff \boxed{(n|\overrightarrow{\mathrm{AM}}) = 0} \end{aligned}$$

On dit que $\mathcal{H}$ est l'hyperplan affine passant par A et **normal au vecteur** n.

À retenir

Les coordonnées d'un vecteur normal à $\mathcal{H}$ se lisent directement sur une équation cartésienne de $\mathcal{H}$. La direction de l'hyperplan affine est $\mathrm{H} = \{x \in \mathrm{E} \ / \ (n|x) = 0\} = \mathrm{Vect}\,[n]^{\perp}$.

4) Équations linéaires

Définition 28.9

Une équation linéaire est une équation du type : $u(x) = b$ avec $u \in \mathcal{L}(\mathrm{E}, \mathrm{F}), b \in \mathrm{F}$ et $x \in \mathrm{E}$ (inconnue). L'équation $u(x) = 0_{\mathrm{F}}$ est appelée équation homogène associée.

☞**Exemples** :

– Tout système linéaire est une équation linéaire, par exemple, le système $\begin{cases} 2x - y = 1 \\ x + 2y = 3 \\ 3x + 5y = -1 \end{cases}$, peut se mettre sous la forme $u(\mathrm{X}) = b$ avec $u\colon \mathbb{R}^2 \to \mathbb{R}^3$ définie par $u(x,y) = (2x-y, x+2y, 3x+5y)$, avec $b = (1,3,-1) \in \mathbb{R}^3$ et $\mathrm{X} = (x,y) \in \mathbb{R}^2$, il est facile de vérifier que $u \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^3)$.
– Une équation différentielle linéaire est une équation linéaire, par exemple l'équation différentielle $y' + y = 1$ peut se mettre sous la forme $u(y) = b$ avec $u : \mathcal{C}^1(\mathbb{R},\mathbb{R}) \to \mathcal{C}^0(\mathbb{R},\mathbb{R})$ définie par $u(y) = y' + y$ (u est linéaire), et $b \in \mathcal{C}^0(\mathbb{R},\mathbb{R})$ la fonction constante 1.

Théorème 28.14 (structure des solutions d'une équation linéaire)
*Soit $u \in \mathcal{L}(\mathrm{E},\mathrm{F})$ et soit $b \in \mathrm{F}$, l'équation linéaire $u(x) = b$ avec $x \in \mathrm{E}$ a des solutions si et seulement si $b \in \mathrm{Im}(u)$. Si c'est le cas, et si $x_0 \in \mathrm{E}$ désigne une solution particulière (i.e. $u(x_0) = b$), alors l'ensemble de toutes les solutions est $\mathrm{S} = \{y + x_0 \ / \ y \in \ker(u)\} = x_0 + \ker(u)$. **C'est donc un sous-espace affine de direction** $\ker(u)$.*

Preuve : Il est clair qu'il y a des solutions si et seulement si $b \in \mathrm{Im}(u)$. Si on a $u(x_0) = b$, alors l'équation $u(x) = b$ équivaut à $u(x) = u(x_0)$, ou encore à $u(x - x_0) = 0_{\mathrm{F}}$ (u est linéaire), ce qui équivaut encore à $\exists\, y \in \ker(u), x = y + x_0$.
□

☞**Exemples** :
– Le système linéaire équivaut à (méthode de *Gauss*[3]) :

$$\begin{cases} -y + 2x = 1 \\ \phantom{-y + {}} 5x = 5 \ (\mathrm{L}_2 \leftarrow \mathrm{L}_2 + 2\mathrm{L}_1) \\ \phantom{-y + {}} 13x = 4 \ (\mathrm{L}_3 \leftarrow \mathrm{L}_3 + 5\mathrm{L}_1) \end{cases} \iff \begin{cases} -y + 2x = 1 \\ x = 1 \\ 0 = -9 \ (\mathrm{L}_3 \leftarrow \mathrm{L}_3 - 13\mathrm{L}_2) \end{cases}$$

La dernière équation montre que ce système n'a pas de solution (*i.e.* $b \notin \mathrm{Im}(u)$).
– Pour l'équation différentielle, il y a une solution particulière : $y_0 : t \mapsto 1$. L'équation homogène associée est $y' + y = 0$ dont les solutions sont les fonctions $y = \lambda\phi$ où $\phi\colon t \mapsto e^{-t}$. L'ensemble des solutions est donc $\mathrm{S} = \{1 + \lambda\phi \ / \ \lambda \in \mathbb{R}\}$.

V SOLUTION DES EXERCICES

Solution 28.1

1/ *On écrit $\|x\| = \|(x+y) - y\| \leqslant \|x+y\| + \|y\|$ et donc $\|x\| - \|y\| \leqslant \|x+y\|$, on échange les rôles de x et y, ce qui donne $\|y\| - \|x\| \leqslant \|x+y\|$, le résultat en découle.*

2/ *Les vecteurs sont colinéaires si et seulement si on a l'égalité de Cauchy-Schwarz, c'est à dire $(x \mid y) = \|x\|\|y\|\cos(\theta) = \pm\|x\|\|y\|$ ce qui équivaut à $\cos(\theta) = \pm 1$.*

3/ *En élevant au carré l'égalité équivaut à $(x|y) = \|x\|\|y\| > 0$, et donc la famille est liée (Cauchy-Schwarz) et les vecteurs de même sens.*

Solution 28.2 *On pose $e_1 = \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{2}}(1,1,0)$, puis $e_2' = v_2 - (v_2 \mid e_1)\cdot e_1 = (\frac{1}{2}, -\frac{1}{2}, 1)$ et $e_2 = \frac{e_2'}{\|e_2'\|} = \frac{1}{\sqrt{6}}(1,-1,2)$. Enfin, $e_3' = v_3 - (v_3 \mid e_1)\cdot e_1 - (v_3 \mid e_2)\cdot e_2 = \frac{2}{3}(-1,1,1)$ et $e_3 = \frac{e_3'}{\|e_3'\|} = \frac{1}{\sqrt{3}}(-1,1,1)$.*

Solution 28.3 *Soit $\mathcal{V}$ un s.e.a. de E de direction V et contenant un point A :*
– *Si $\dim(\mathrm{V}) = 0$, alors $\mathrm{V} = \{0_{\mathrm{E}}\}$ et donc $\mathcal{V} = \{\mathrm{A} + u \ / \ u \in \mathrm{V}\} = \{\mathrm{A}\}$. Un s.e.a. de dimension 0 est réduit à un point.*
– *Si $\dim(\mathrm{V}) = 1$, alors $\mathrm{V} = \overrightarrow{u}$ avec $u \neq 0_{\mathrm{E}}$, et donc $\mathcal{V} = \{\mathrm{A} + \lambda u \ / \ \lambda \in \mathbb{R}\}$, c'est la droite affine passant par A et de vecteur directeur u.*
– *Si $\dim(\mathrm{V}) = 2$, alors $\mathrm{V} = \overrightarrow{u,v}$ avec (u,v) libre, et donc $\mathcal{V} = \{\mathrm{A} + \alpha u + \beta v \ / \ \alpha,\beta \in \mathbb{R}\}$, c'est le plan affine passant par A et de base (u,v).*
– *Si $\dim(\mathrm{V}) = 3$, alors $\mathrm{V} = \mathrm{E}$, et donc $\mathcal{V} = \{\mathrm{A} + u \ / \ u \in \mathrm{E}\} = \mathrm{E}$.*

Solution 28.4 *Soit $\mathcal{V} = \{\mathrm{M}(x_1,\ldots,x_n) \ / \ a_1x_1 + \cdots + a_nx_n = b\}$, supposons le coefficient a_1 non nul, alors le point $\mathrm{A}(y_1,\ldots,y_n)$ avec $(y_1,\ldots,yn) = (\frac{b}{a_1}, 0, \ldots, 0)$, est dans $\mathcal{V}$. D'autre part :*

$$\mathrm{M}(x_1,\ldots,x_n) \in \mathcal{V} \iff a_1x_1 + \cdots + a_nx_n = a_1y_1 + \cdots + a_ny_n \iff a_1(x_1 - y_1) + \cdots + a_n(x_n - y_n) = 0$$

Soit V l'hyperplan vectoriel d'équation $a_1x_1 + \cdots + a_nx_n = 0$, alors $\mathrm{M} \in \mathcal{V} \iff \overrightarrow{\mathrm{AM}} \in \mathrm{V}$, donc $\mathcal{V}$ est l'hyperplan affine passant par A et de direction V.

3. *Carl Friedrich* (1777 – 1855) : mathématicien allemand de génie, sans doute l'un des plus grands de tous les temps.

Chapitre 29

Intégration sur un segment

Sommaire

Dans tout le chapitre $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$

I INTÉGRALE DES FONCTIONS EN ESCALIER

1) Fonctions en escalier

Définition 29.1

Soit $f: [a;b] \to \mathbb{K}$ une fonction, on dit que f est en escalier sur $[a;b]$ lorsqu'il existe un entier $n \in \mathbb{N}^*$, des réels $x_0 = a < x_1 < \cdots < x_n = b$, et des nombres $c_0, \ldots, c_{n-1}$ de $\mathbb{K}$ tels que sur chacun des intervalles ouverts : $]x_k;x_{k+1}[$ la fonction f est constante égale à c_k ($0 \leqslant k \leqslant n-1$). On dit aussi parfois que f est constante par morceaux. L'ensemble des fonctions en escalier sur $[a;b]$ est noté $\mathcal{E}([a;b],\mathbb{K})$.

☞**Exemples** :
- Une fonction constante sur $[a;b]$ est en escalier.
- La fonction partie entière restreinte à $[a;b]$ est en escalier.

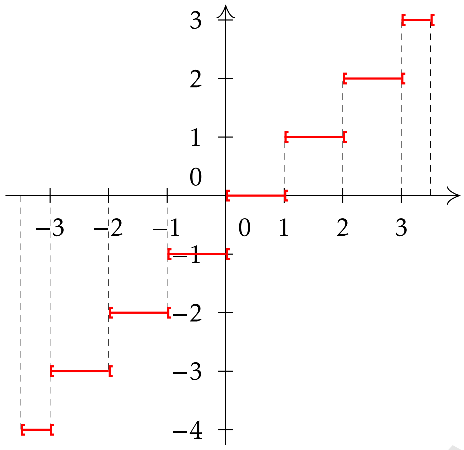


FIGURE 29.1 – La fonction partie entière sur $[-3.5, 3.5]$.

Remarque 29.1 –

- *La courbe représentative d'une fonction en escalier a la forme d'un escalier !*
- *Les réels $x_0 = a < \cdots < x_n = b$ de la définition constituent ce que l'on appelle une **subdivision** de l'intervalle $[a\,;b]$. Une telle subdivision est dite **adaptée** à la fonction en escalier f lorsque f est constante sur chacun des morceaux (ouverts) de la subdivision. Il est facile de voir qu'il y a une infinité de subdivisions adaptées à f lorsque f est en escalier. Le réel $\max\{x_{k+1} - x_k \,/\, 0 \leqslant k \leqslant n-1\}$ est appelé **le pas** de la subdivision. Lorsque le pas est constant, il vaut $\frac{b-a}{n}$, on dit que la subdivision est **régulière**.*
- *La définition ne fait pas intervenir la valeur de f aux points de la subdivision.*
- *Une fonction en escalier sur $[a\,;b]$ est bornée.*

Définition 29.2

Si σ, σ' sont deux subdivisions de $[a\,;b]$, on dit que σ' est plus fine que σ lorsque les points de la subdivision σ font partie de la subdivision σ', ce que l'on note $\sigma \subset \sigma'$, si de plus σ est adaptée à $f \in \mathcal{E}([a\,;b], \mathbb{K})$, alors σ' aussi.

Remarque 29.2 –

- *Si σ, σ' sont deux subdivisions de $[a\,;b]$, on note $\sigma \cup \sigma'$ la subdivision obtenue en réunissant les points de σ avec ceux de σ' (et en les rangeant dans l'ordre croissant). Cette nouvelle subdivision est plus fine que les deux précédentes.*
- *Il en découle que si f et g sont en escalier sur $[a\,;b]$ alors il existe une subdivision σ de $[a\,;b]$ qui est à la fois adaptée à f et à g.*

Théorème 29.1

L'ensemble $\mathcal{E}([a\,;b], \mathbb{K})$ est un $\mathbb{K}$-espace vectoriel (et même une $\mathbb{K}$-algèbre) pour les opérations usuelles sur les fonctions.

Preuve : Celle-ci est simple et laissée en exercice. □

2) Intégrale d'une fonction en escalier

Théorème 29.2

Soit $f \in \mathcal{E}([a\,;b], \mathbb{K})$ et soit $\sigma = (x_i)_{0 \leqslant i \leqslant n}$ une subdivision de $[a\,;b]$ adaptée à f, alors la quantité :

$$\mathrm{I}_\sigma(f) = \sum_{i=0}^{n-1} (x_{i+1} - x_i) c_i,$$

où c_i désigne la valeur de f sur l'intervalle $]x_i, x_{i+1}[$, est indépendante de la subdivision adaptée à f. Autrement dit, si σ' est une autre subdivision adaptée à f, alors $\mathrm{I}_\sigma(f) = \mathrm{I}_{\sigma'}(f)$.

Preuve : Si on rajoute un point d à la subdivision σ, on obtient une nouvelle subdivision σ' et il existe un indice $j \in [\![0\,;(n-1)]\!]$ tel que $d \in]x_j, x_{j+1}[$, mais alors $c_j(x_{j+1} - x_j) = c_j(d - x_j) + c_j(x_{j+1} - d)$, on voit donc que $\mathrm{I}_\sigma(f) = \mathrm{I}_{\sigma'}(f)$. Par récurrence, on en déduit que si σ' est une subdivision plus fine que σ, alors $\mathrm{I}_\sigma(f) = \mathrm{I}_{\sigma'}(f)$.

Soit σ' une autre subdivision de $[a\,;b]$ adaptée à f, la subdivision $\sigma'' = \sigma' \cup \sigma$ est adaptée à f et plus fine que σ et σ', donc $\mathrm{I}_{\sigma''}(f) = \mathrm{I}_{\sigma'}(f) = \mathrm{I}_\sigma(f)$. □

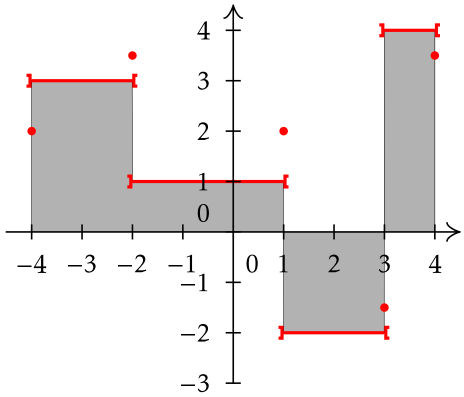


FIGURE 29.2 – Interprétation géométrique

Remarque 29.3 – *Géométriquement, si $f \in \mathcal{E}([a\,;b],\mathbb{K})$ et si σ est une subdivision de $[a\,;b]$ adaptée à f, alors dans un repère orthonormé, la quantité $\mathrm{I}_\sigma(f)$ représente l'aire algébrique de la portion de plan délimitée par la courbe de f, l'axe des abscisses, et les droites d'équation : $x = a$ et $x = b$, c'est une somme d'aires de rectangles.*

Définition 29.3 (intégrale d'une fonction en escalier)

Si $f \in \mathcal{E}([a\,;b],\mathbb{R})$, on appelle intégrale de f sur $[a\,;b]$ le nombre (complexe) noté $\int_{[a;b]} f$ et défini par :

$$\int_{[a;b]} f = \mathrm{I}_\sigma(f) = \sum_{i=0}^{n-1}(x_{i+1}-x_i)c_i,$$

où $\sigma = (x_i)_{0\leqslant i\leqslant n}$ est une subdivision adaptée à f.

À retenir

- L'intégrale de f sur $[a\,;b]$ ne dépend pas de la valeur de f aux points de la subdivision. Il en découle que si on modifie la valeur de f en un nombre fini de points, la valeur de l'intégrale reste inchangée.
- Si $f,g \in \mathcal{E}([a\,;b],\mathbb{K})$ et si f et g coïncident sur $[a\,;b]\setminus\{x_1,\ldots,x_n\}$, alors $\int_{[a;b]} f = \int_{[a;b]} g$ car il suffit de changer la valeur de f aux points $x_1,\ldots,x_n$ pour obtenir la fonction g, ce qui ne change pas l'intégrale de f.

Théorème 29.3 (Propriétés élémentaires)

Soient $f,g \in \mathcal{E}([a\,;b],\mathbb{K})$:

- ***linéarité** : $\int_{[a;b]} f+g = \int_{[a;b]} f + \int_{[a;b]} g$ et si $\lambda \in \mathbb{K}$, $\int_{[a,b]} \lambda f = \lambda \int_{[a;b]} f$.*
- ***positivité** : si f est à **valeurs réelles positives**, alors $\int_{[a;b]} f \geqslant 0$. On en déduit que si f et g sont à valeurs réelles et si $f \leqslant g$, alors $\int_{[a;b]} f \leqslant \int_{[a;b]} g$.*
- ***majoration** : $|\int_{[a;b]} f| \leqslant \int_{[a;b]} |f|$.*
- ***relation de CHASLES**[1] : si $a < c < b$, alors $\int_{[a;b]} f = \int_{[a;c]} f + \int_{[c;b]} f$.*

Preuve : Celle-ci est simple et laissée en exercice. □

II INTÉGRALE DES FONCTIONS CONTINUES PAR MORCEAUX

1) Fonctions continues par morceaux

1. CHASLES MICHEL (1793 – 1880) : mathématicien français, auteur d'importants travaux en géométrie.

Définition 29.4

Une fonction $f\colon [a\,;b] \to \mathbb{K}$ est dite continue par morceaux sur le segment $[a\,;b]$, lorsqu'il existe une subdivision $\sigma = (a = x_0, \dots, x_n = b)$ de $[a\,;b]$ telle que sur chaque morceau $]x_k; x_{k+1}[$ la fonction f est continue et ***prolongeable par continuité sur*** *$[x_k; x_{k+1}]$ (i.e. f a une limite finie à droite en x_k et à gauche en x_{k+1}). L'ensemble des fonctions continues par morceaux sur $[a\,;b]$ est noté $\mathcal{C}^0_{\mathrm{m}}([a\,;b], \mathbb{K})$. Plus généralement, on dit qu'une fonction est continue par morceaux sur un intervalle I, lorsque sa restriction à tout segment inclus dans I est continue par morceaux.*

☞**Exemples** :
- Une fonction continue sur $[a\,;b]$ est continue par morceaux.
- Une fonction en escalier sur $[a\,;b]$ est continue par morceaux.
- La fonction f définie par $f(x) = \sin(\frac{1}{x})$ sur $]0;1]$ et $f(0) = 0$ n'est pas continue par morceaux sur $[0;1]$.

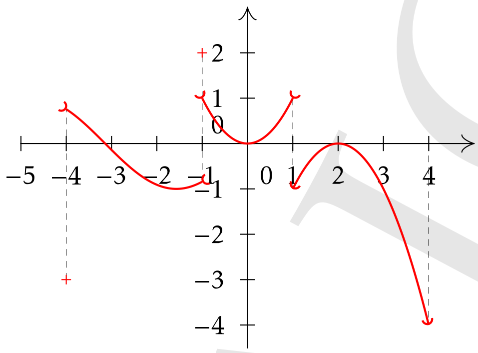


Figure 29.3 – Exemple de fonction continue par morceaux

Remarque 29.4 –
- *Une fonction continue par morceaux sur $[a\,;b]$ est bornée.*
- *La valeur de f aux points de la subdivision n'intervient pas dans la définition.*

Théorème 29.4

L'ensemble des fonctions continues par morceaux sur $[a\,;b]$, $\mathcal{C}^0_{\mathrm{m}}([a\,;b], \mathbb{K})$, est un $\mathbb{K}$-espace vectoriel (et même une $\mathbb{K}$-algèbre) pour les opérations usuelles sur les fonctions.

Preuve : Celle-ci est simple et laissée en exercice. □

2) Approximation uniforme

Théorème 29.5

Si $f\colon [a\,;b] \to \mathbb{K}$ est continue par morceaux sur le segment $[a\,;b]$, alors pour tout $\varepsilon > 0$, il existe une fonction φ en escalier sur $[a\,;b]$ telle que $|f - \varphi| \leqslant \varepsilon$, c'est à dire :

$$\forall t \in [a\,;b], |f(t) - \varphi(t)| \leqslant \varepsilon.$$

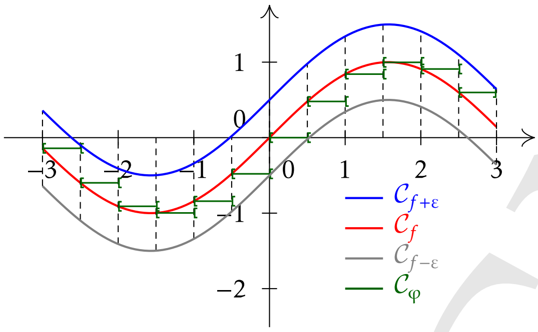
Preuve :
• Cas où f est continue : f est uniformément continue sur $[a\,;b]$ (théorème de Heine), il existe donc un réel $\alpha > 0$ tel que pour tout $x, y \in [a,b]$, $|x - y| < \alpha \implies |f(x) - f(y)| < \varepsilon$. Soit $n = 1 + \left\lfloor \frac{b-a}{\alpha} \right\rfloor$, on a alors $\dfrac{b-a}{n} < \alpha$. Découpons l'intervalle $[a\,;b]$ en n morceaux de longueur $\frac{b-a}{n}$, on obtient ainsi une subdivision dont les points sont les réels $x_k = a + k\dfrac{b-a}{n}$ pour $k \in [\![0\,;n]\!]$. On définit maintenant la fonction φ en posant :

$$\varphi(t) = \begin{cases} f(b) & \text{si } t = b \\ f(x_k) & \text{si } t \in [x_k\,; x_{k+1}[\end{cases}$$

il est alors facile de vérifier que pour tout réel t de $[a\,;b]$, on a $|f(t) - \varphi(t)| < \varepsilon$ en distinguant les cas $t = b$ et $t \in [x_k\,; x_{k+1}[$, on a donc $|f - \varphi| < \varepsilon$.

• Cas où f est continue par morceaux : sur chaque morceau $]x_k\,;x_{k+1}[$ la fonction f est prolongeable par continuité sur $[x_k;x_{k+1}]$ en une fonction f_k. On sait alors qu'il existe une fonction φ_k en escalier sur $[x_k\,;x_{k+1}]$ telle que $\forall\, t \in [x_k\,;x_{k+1}], |f_k(t)-\varphi_k(t)| < \varepsilon$, en particulier $\forall\, t \in]x_k\,;x_{k+1}[, |f(t)-\varphi_k(t)| < \varepsilon$.

On construit la fonction φ en posant : $\varphi(x_k) = f(x_k)$ et pour $t \in]x_k\,;x_{k+1}[, \varphi(t) = \varphi_k(t)$. Il est clair que φ est en escalier sur $[a\,;b]$ et que $\forall\, t \in [a\,;b], |f(t)-\varphi(t)| < \varepsilon$, donc $|f-\varphi| < \varepsilon$. □



Remarque 29.5 –

– *C'est l'uniforme continuité de f qui fait aboutir la démonstration.*

– *Si f est continue sur $[a,b]$, alors pour tout $\varepsilon > 0$ il existe deux fonctions en escalier ϕ et ψ telle que $\forall\, t \in [a\,;b], \psi(t) \leqslant f(t) \leqslant \phi(t)$ avec $\phi(t)-\psi(t) < \varepsilon$.*

En effet, on sait qu'il existe une fonction en escalier g telle que $|f-g| < \varepsilon/2$, on a donc pour $t \in [a\,;b], g(t)-\varepsilon/2 \leqslant f(t) \leqslant g(t)+\varepsilon/2$, il suffit donc de prendre $\psi = g-\varepsilon/2$ et $\phi = g+\varepsilon/2$.

– *Si $f \in \mathcal{C}^0_{\mathrm{m}}([a\,;b],\mathbb{K})$ alors il existe une suite (ψ_n) de fonctions en escalier sur $[a\,;b]$ telle que :*

$$\sup_{x\in[a;b]} |f(x)-\psi_n(x)| \xrightarrow[n\to+\infty]{} 0.$$

On dit que la suite (ψ_n) ***converge uniformément vers*** *f* ***sur*** *$[a\,;b]$.*

Définition 29.5

• *Si f est bornée sur $[a\,;b]$, on pose $\|f\|_\infty = \sup\limits_{x\in[a;b]} |f(x)|$, c'est la norme infinie de f.*

• *Soit (φ_n) une suite de fonctions bornées sur $[a\,;b]$ et soit f bornée sur $[a\,;b]$, on dit que la suite (φ_n) converge uniformément vers f sur $[a\,;b]$, lorsque : $\lim\limits_{n\to+\infty} \|\varphi_n - f\|_\infty = 0$.*

À retenir : Un exemple important

Si f est continue sur $[a\,;b]$, alors la suite de fonctions en escalier $(\varphi_n)_{n\in\mathbb{N}^*}$ définie par : $\forall t \in [a\,;b]$, $\varphi_n(t) = f(b)$ si $t = b$ et $\varphi_n(t) = f(x_k)$ si $t \in [x_k\,;x_{k+1}[$ où $x_k = a + k\frac{b-a}{n}$, converge uniformément vers f.

En effet, il suffit de reprendre la même démonstration que le théorème précédent, dans le cas où f est continue.

3) Définition de l'intégrale

Théorème 29.6

Soit $f \in \mathcal{C}^0_{\mathrm{m}}([a\,;b],\mathbb{K})$ et soit (ϕ_n) une suite de fonctions en escalier qui ***converge uniformément vers*** *f sur $[a\,;b]$, alors :*

• *la suite $(\int_{[a;b]} \phi_n)$ converge vers un nombre $\ell \in \mathbb{K}$;*

• *ce nombre ℓ ne dépend pas de la suite (ϕ_n) choisie.*

Preuve : Posons $u_n = \int_{[a;b]} \phi_n$, soit $\varepsilon > 0$, il existe un entier $\mathrm{N} \in \mathbb{N}$ tel que $n \geqslant \mathrm{N} \implies |\phi_n - f| < \varepsilon$, on en déduit que $|\phi_n| < |f|+\varepsilon \leqslant \mathrm{M}+\varepsilon$ où M est un majorant de $|f|$, on a alors $|u_n| \leqslant (b-a)(\mathrm{M}+\varepsilon)$: la suite u est bornée, on peut donc en extraire une suite convergente : $u_{\sigma(n)} \to \ell$, mais pour $n \geqslant \mathrm{N}$ on a $|\phi_n - \phi_{\sigma(n)}| \leqslant |\phi_n - f| + |f - \phi_{\sigma(n)}| \leqslant 2\varepsilon$, on en déduit que $|u_n - u_{\sigma(n)}| \leqslant (b-a)2\varepsilon$ ce qui entraîne que $u_n \to \ell$.

Soit (ψ_n) une autre suite de fonctions en escalier qui converge uniformément vers f, posons $v_n = \int_{[a;b]} \psi_n$, d'après ce qui précède, la suite (v_n) converge vers un nombre ℓ'. Soit (g_n) la suite de fonctions en escalier définie par $g_{2n} = \phi_n$ et $g_{2n+1} = \psi_n$, il est facile de voir que (g_n) converge uniformément vers f, donc la suite $(w_n = \int_{[a;b]} g_n)$ converge vers un nombre ℓ'', or $w_{2n} = u_n$ et $w_{2n+1} = v_n$, on en déduit que $\ell'' = \ell = \ell'$. □

Définition 29.6 (intégrale d'une fonction continue par morceaux)

Soit $f \in \mathcal{C}^0_{\mathrm{m}}([a\,;b], \mathbb{K})$, on appelle intégrale de f sur $[a\,;b]$ le nombre noté $\int_{[a;b]} f$ et défini par :

$$\int_{[a;b]} f = \lim_{n\to+\infty} \int_{[a;b]} \phi_n,$$

où (ϕ_n) est une suite de fonctions en escalier sur $[a\,;b]$ qui converge uniformément vers f.
Géométriquement, dans un repère orthonormé, si f est à valeurs réelles, on dit que $\int_{[a;b]} f$ représente l'aire algébrique de la portion de plan délimitée par la courbe de f, l'axe des abscisses, et les droites d'équation $x = a$ et $x = b$.

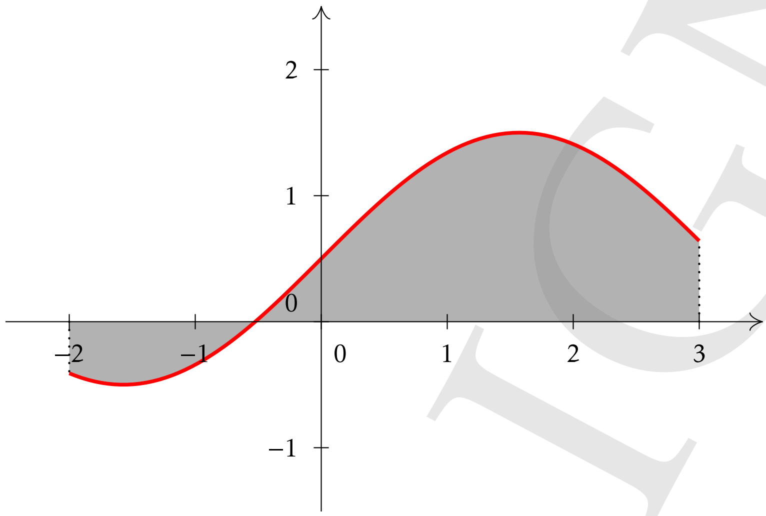


Figure 29.4 – Cas d'une fonction continue

Théorème 29.7 (Sommes de Riemann des rectangles de gauche)

Soit f continue sur $[a\,;b]$, pour $n \in \mathbb{N}^*$, on appelle somme de Riemann[2] des rectangles de gauche associée à f, le nombre $\mathrm{R}_n(f) = \frac{b-a}{n}\sum\limits_{k=0}^{n-1} f(a + k\frac{b-a}{n})$. On a alors le résultat suivant :

$$\lim_{n\to+\infty} \mathrm{R}_n(f) = \int_{[a;b]} f.$$

Preuve : On a $\mathrm{R}_n(f) = \sum\limits_{k=0}^{n-1} \frac{b-a}{n} f(a + k\frac{b-a}{n}) = \sum\limits_{k=0}^{n-1} f(x_k)(x_{k+1} - x_k) = \int_{[a;b]} \varphi_n$, où φ_n est la fonction en escalier définie par : $\forall t \in [a\,;b]$, $\varphi_n(t) = f(b)$ si $t = b$ et $\varphi_n(t) = f(x_k)$ si $t \in [x_k\,;x_{k+1}[$ où $x_k = a + k\frac{b-a}{n}$. Or on a vu que cette suite de fonctions en escalier converge uniformément vers f, d'après le théorème précédent on peut donc écrire que $\int_{[a;b]} f = \lim\limits_{n\to+\infty} \int_{[a;b]} \varphi_n$, ce qui donne le résultat. □

Remarque 29.6 :

– *On admettra que le résultat reste vrai pour une fonction continue par morceaux sur $[a\,;b]$.*
– *On peut montrer un résultat analogue pour les sommes de Riemann des rectangles de droite, si f est continue par morceaux sur $[a\,;b]$:*

$$\lim_{n\to+\infty} \frac{b-a}{n}\sum_{k=0}^{n-1} f(a + (k+1)\frac{b-a}{n}) = \lim_{n\to+\infty} \frac{b-a}{n}\sum_{k=1}^{n} f(a + k\frac{b-a}{n}) = \int_{[a;b]} f.$$

4) Propriétés de l'intégrale

Théorème 29.8 (linéarité de l'intégrale)

Soient $f, g \in \mathcal{C}^0_{\mathrm{m}}([a\,;b], \mathbb{K})$ et soit $\lambda \in \mathbb{K}$, alors :

$$\int_{[a;b]} (f+g) = \int_{[a;b]} f + \int_{[a;b]} g \quad \text{et} \quad \int_{[a;b]} \lambda.f = \lambda.\int_{[a;b]} f.$$

2. RIEMANN Georg Friedrich Bernhard (1826 – 1866) : mathématicien allemand dont l'œuvre est colossale.

Preuve : Soient (ϕ_n) et (ψ_n) deux suites de fonctions en escalier qui convergent uniformément respectivement vers f et g, il est facile de voir que la suite $(\phi_n + \psi_n)$ converge uniformément vers $f+g$. D'après la définition, on a $\int_{[a;b]}(f+g) = \lim\limits_{n\to+\infty} \int_{[a;b]}(\phi_n+\psi_n)$, la linéarité étant vérifiée pour les fonctions en escalier, on peut écrire que $\int_{[a;b]}(f+g) = \int_{[a;b]}\phi_n + \int_{[a;b]}\psi_n$, le résultat s'obtient alors par passage à la limite. La preuve est du même type pour le second point. □

Remarque 29.7 – *Soit $f \in \mathcal{C}^0_{\mathrm{m}}([a\,;b],\mathbb{K})$, posons $u = \mathrm{Re}(f)$ et $v = \mathrm{Im}(f)$. La linéarité de l'intégrale permet d'écrire : $\int_{[a;b]} f = \int_{[a;b]}\mathrm{Re}(f) + i\int_{[a;b]}\mathrm{Im}(f)$. On peut donc toujours se ramener à intégrer des fonctions à valeurs réelles (mais ce n'est pas toujours la meilleure solution). D'autre part, on a établi :*

$$\mathrm{Re}\Big(\int_{[a;b]} f\Big) = \int_{[a;b]}\mathrm{Re}(f) \ \text{ et } \ \mathrm{Im}\Big(\int_{[a;b]} f\Big) = \int_{[a;b]}\mathrm{Im}(f), \text{ d'où } \overline{\int_{[a;b]} f} = \int_{[a;b]}\overline{f}.$$

Théorème 29.9 (positivité)

*Si $f \in \mathcal{C}^0_{\mathrm{m}}([a\,;b],\mathbb{R})$ est à **valeurs positives**, alors $0 \leqslant \int_{[a;b]} f$. En particulier si $f, g \in \mathcal{C}^0_{\mathrm{m}}([a\,;b],\mathbb{R})$ et si $f \leqslant g$, alors $\int_{[a;b]} f \leqslant \int_{[a;b]} g$.*

Preuve : Si f est à valeurs positives, on peut construire une suite (ϕ_n) de fonctions en escalier **positives**, qui converge uniformément vers f, comme $\int_{[a;b]} f = \lim\limits_{n\to+\infty}\phi_n$, on a le résultat par passage à la limite.

Si $f \leqslant g$, on applique ce qui précède à la fonction $h = g - f$ et on conclut avec la linéarité. □

Théorème 29.10 (majoration en module)

Si $f \in \mathcal{C}^0_{\mathrm{m}}([a\,;b],\mathbb{K})$, alors $\left|\int_{[a;b]} f\right| \leqslant \int_{[a;b]} |f|$.

En particulier si $|f| \leqslant \mathrm{M}$ sur $[a\,;b]$, alors $\left|\int_{[a;b]} f\right| \leqslant \mathrm{M}(b-a)$.

Preuve : Si f est continue par morceaux sur $[a\,;b]$, alors $|f|$ aussi, et si (ϕ_n) est une suite de fonctions en escalier qui converge uniformément vers f, il est facile de vérifier que la suite $(|\phi_n|)$ est une suite de fonctions en escalier qui converge uniformément vers $|f|$, de plus, on sait que $\left|\int_{[a;b]}\phi_n\right| \leqslant \int_{[a;b]}|\phi_n|$, le résultat s'obtient par passage à la limite. □

Théorème 29.11 (relation de Chasles)

Si $f \in \mathcal{C}^0_{\mathrm{m}}([a\,;b],\mathbb{K})$ et si $a < c < b$, alors $\int_{[a;b]} f = \int_{[a;c]} f + \int_{[c;b]} f$.

Preuve : Même type de preuve que pour les résultats précédents. □

Théorème 29.12 (égalité d'intégrales)

Si $f, g \in \mathcal{C}^0_{\mathrm{m}}([a\,;b],\mathbb{K})$ coïncident sur $[a\,;b]\setminus\{x_1,\dots,x_n\}$, alors $\int_{[a;b]} f = \int_{[a;b]} g$.

Preuve : Posons $h = g - f$, alors h est une fonction en escalier qui coïncide avec la fonction nulle sauf éventuellement aux points $x_1,\dots,x_n$, on sait alors que $\int_{[a;b]} h = 0$, et la linéarité entraîne alors le résultat. □

Convention d'écriture

Si f est continue par morceaux sur un intervalle $[a\,;b]$, pour $x, y \in [a\,;b]$, on pose :

$$\int_x^y f(t)\,\mathrm{d}t = \begin{cases} \int_{[x;y]} f & \text{si } x < y \\ 0 & \text{si } x = y \\ -\int_{[y;x]} f & \text{si } y < x \end{cases}$$

Théorème 29.13

Si $f \in \mathcal{C}^0_{\mathrm{m}}([a\,;b],\mathbb{K})$ alors :

- *$\forall x, y \in [a\,;b]$, $\int_x^y f(t)\,\mathrm{d}t = -\int_y^x f(t)\,\mathrm{d}t$.*
- *$\forall x, y, z \in [a\,;b]$, $\int_x^y f(t)\,\mathrm{d}t = \int_x^z f(t)\,\mathrm{d}t + \int_z^y f(t)\,\mathrm{d}t$ (relation de Chasles généralisée).*

Preuve : Celle-ci est simple et laissée en exercice. □

5) Approximation par une somme de Riemann

Théorème 29.14

Soit $f\colon [a\,;b] \to \mathbb{C}$ une fonction continue par morceaux, on a :

$$\lim_{n\to+\infty} \mathrm{R}_n(f) = \lim_{n\to+\infty} \tfrac{b-a}{n} \sum_{k=0}^{n-1} f(a + k\tfrac{b-a}{n}) = \int_{[a;b]} f.$$

Et si f est K-lipschitzienne sur $[a\,;b]$, alors :

$$\left|\mathrm{R}_n(f) - \int_{[a;b]} f\right| \leqslant \mathrm{K}\tfrac{(b-a)^2}{n}.$$

Preuve : On a donc $\forall x, y \in [a\,;b], |f(x) - f(y)| \leqslant \mathrm{K}|x-y|$.

$$\begin{aligned}
\left|\mathrm{R}_n(f) - \int_{[a;b]} f\right| &= \left|\sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} f(x_k)\,\mathrm{d}t - \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} f(t)\,\mathrm{d}t\right| = \left|\sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} (f(x_k) - f(t))\,\mathrm{d}t\right| \\
&\leqslant \sum_{k=0}^{n-1} \left|\int_{x_k}^{x_{k+1}} (f(x_k) - f(t))\,\mathrm{d}t\right| \leqslant \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} |f(x_k) - f(t)|\,\mathrm{d}t \\
&\leqslant \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} \mathrm{K}|x_k - t|\,\mathrm{d}t \leqslant \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} \mathrm{K}\frac{b-a}{n}\,\mathrm{d}t \\
&\leqslant \sum_{k=0}^{n-1} \mathrm{K}\frac{(b-a)^2}{n^2} = \frac{\mathrm{K}(b-a)^2}{n}
\end{aligned}$$

on a donc $\left|\mathrm{R}_n(f) - \int_{[a;b]} f\right| \leqslant \frac{\mathrm{K}(b-a)^2}{n} \to 0$. □

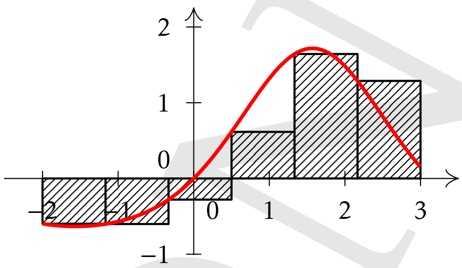
Remarque 29.8 – *Le théorème s'applique en particulier quand f est $\mathcal{C}^1$ sur $[a\,;b]$, en prenant $\mathrm{K} = \sup\limits_{t\in[a;b]} |f'(t)|$, f est K-lipschitzienne grâce à l'inégalité des accroissements finis.*

À retenir

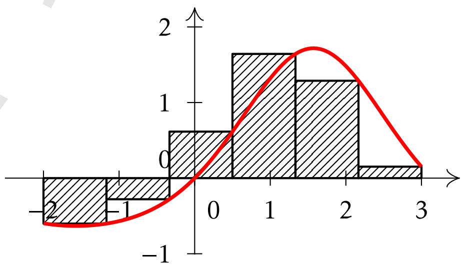
Méthode des rectangles de gauche pour le calcul approché d'une intégrale :

$\mathrm{R}_n(f) = \frac{b-a}{n} \sum\limits_{k=0}^{n-1} f(a + k\frac{b-a}{n})$ est une valeur approchée de $\int_{[a;b]} f$ à $\frac{\mathrm{K}(b-a)^2}{n}$ près où $\mathrm{K} = \sup\limits_{a\leqslant t\leqslant b} |f'(t)|$.

On a le même résultat pour les rectangles de droite.



Méthode des rectangles de gauche



Méthode des rectangles de droite

☞**Exemples** :

- Étude de certaines suites : soit $u_n = \sum\limits_{k=1}^{n} \frac{1}{n+k}$, on a alors $u_n = \frac{1}{n}\sum\limits_{k=1}^{n} \frac{1}{1+k/n}$, c'est la méthode des rectangles de droite appliquée à la fonction $f\colon t \mapsto \frac{1}{1+t}$ sur l'intervalle $[0; 1]$, la fonction étant continue sur cet intervalle, on a $\lim u_n = \int_{[0;1]} f = \ln(2)$.
- Calcul de certaines intégrales : $\int_0^{2\pi} \ln(1 - 2x\cos(t) + x^2)\,\mathrm{d}t$ pour $|x| \neq 1$ (cf. TD).

III CALCUL D'UNE INTÉGRALE

1) Primitives

Théorème 29.15 (fondamental de l'intégration)
*Si $f\colon \mathrm{I} \to \mathbb{C}$ est **continue**, alors f admet des primitives sur I. Plus précisément, si $t_0 \in \mathrm{I}$ et $y_0 \in \mathbb{C}$, alors la fonction F définie sur I par : $\mathrm{F}(t) = y_0 + \int_{t_0}^{t} f(u)\,\mathrm{d}u$, est l'unique primitive de f sur I qui prend la valeur y_0 en t_0.*

Preuve : Soit $t_1 \in \mathrm{I}$, on a $|\mathrm{F}(t)-\mathrm{F}(t_1)-(t-t_1)f(t_1)| = |\int_{t_1}^{t} f(u)\,\mathrm{d}u - \int_{t_1}^{t} f(t_1)\,\mathrm{d}u| = |\int_{t_1}^{t}(f(u)-f(t_1))\,\mathrm{d}u| \leqslant |\int_{t_1}^{t}|f(u)-f(t_1)|\,\mathrm{d}u|$. On se donne $\varepsilon > 0$, f étant continue en t_1, il existe $\alpha > 0$ tel que $\forall\, u \in \mathrm{I}, |u-t_1| < \alpha \implies |f(u)-f(t_1)| < \varepsilon$, donc si $|t-t_1| < \alpha$, alors $|\int_{t_1}^{t}|f(u)-f(t_1)|\,\mathrm{d}u| \leqslant |t-t_1|\varepsilon$, d'où :

$$\left|\frac{\mathrm{F}(t)-\mathrm{F}(t_1)}{t-t_1} - f(t_1)\right| \leqslant \varepsilon$$

on en déduit que F est dérivable en t_1 et que $\mathrm{F}'(t_1) = f(t_1)$. La fonction F est donc une primitive de f sur I, et il est clair que $\mathrm{F}(t_0) = y_0$. □

Remarque 29.9 – *La continuité de f est essentielle pour la démonstration, prenons $f(t) = \begin{cases} 1 & \text{si } t = 0 \\ 0 & \text{sinon} \end{cases}$, alors avec $t_0 = y_0 = 0$, on obtient que $\mathrm{F} = 0$, F est bien dérivable mais ce n'est pas une primitive de f sur* $[0;1]$.

Théorème 29.16 (calcul d'une intégrale)
Si $f\colon \mathrm{I} \to \mathbb{C}$ est continue sur l'intervalle I et si F désigne une primitive de f sur I, alors :
$$\forall a, b \in \mathrm{I}, \int_a^b f(t)\,\mathrm{d}t = [\mathrm{F}]_a^b = \mathrm{F}(b) - \mathrm{F}(a).$$

Preuve : F étant une primitive de f, on a $\forall\, t \in \mathrm{I}, \mathrm{F}(t) = \mathrm{F}(a) + \int_a^t f(u)\,\mathrm{d}u$, d'où $\mathrm{F}(b) - \mathrm{F}(a) = \int_a^b f(u)\,\mathrm{d}u$. □

Cas d'une fonction continue par morceaux

si $f\colon [a\,;b] \to \mathbb{C}$ est continue par morceaux, soit $\sigma = (x_i)_{0\leqslant i\leqslant n}$ une subdivision adaptée à f. Sur chacun des morceaux $]x_i;x_{i+1}[$ la fonction f admet un prolongement par continuité f_i sur le segment $[x_i\,;x_{i+1}]$, les deux fonctions coïncidant sur le segment $[x_i\,;x_{i+1}]$, sauf peut être en deux points, on a $\int_{x_i}^{x_{i+1}} f(t)\,\mathrm{d}t = \int_{x_i}^{x_{i+1}} f_i(t)\,\mathrm{d}t$, mais f_i admet une primitive F_i sur $[x_i\,;x_{i+1}]$, d'où $\int_{x_i}^{x_{i+1}} f = \mathrm{F}_i(x_{i+1}) - \mathrm{F}_i(x_i)$, la relation de Chasles donne alors :

$$\int_a^b f(t)\,\mathrm{d}t = \sum_{i=0}^{n-1} \mathrm{F}_i(x_{i+1}) - \mathrm{F}_i(x_i)$$

On peut donc toujours se ramener au cas des fonctions continues et donc à une recherche de primitive.

Théorème 29.17 (lien entre une fonction et sa dérivée)
Si f est de classe $\mathcal{C}^1$ sur l'intervalle I, alors $\forall\, t, t_0 \in \mathrm{I}, f(t) = f(t_0) + \int_{t_0}^{t} f'(u)\,\mathrm{d}u$.

Théorème 29.18 (inégalité des accroissements finis généralisée)
Si $f, g\colon \mathrm{I} \to \mathbb{C}$ sont de classe $\mathcal{C}^1$ sur l'intervalle I et si $\forall\, t \in \mathrm{I}, |f'(t)| \leqslant g'(t)$, alors :
$$\forall\, a, b \in \mathrm{I}, |f(b) - f(a)| \leqslant |g(b) - g(a)|.$$

Preuve : Supposons $a < b$, on a :
$|f(b) - f(a)| = |\int_{[a;b]} f'(u)\,\mathrm{d}u| \leqslant \int_{[a;b]} |f'(u)|\,\mathrm{d}u \leqslant \int_{[a;b]} g'(u)\,\mathrm{d}u = g(b) - g(a) = |g(b) - g(a)|$. □

Théorème 29.19 (cas d'une intégrale nulle)
*Si f est à **valeurs réelles, continue, positive** sur $[a\,;b]$, et si $\int_{[a;b]} f = 0$, alors f est nulle sur $[a\,;b]$.*

Preuve : Soit F une primitive de f sur $[a\,;b]$, alors $\mathrm{F}' = f \geqslant 0$, donc F est croissante, or $\mathrm{F}(b) - \mathrm{F}(a) = \int_0^b f(t)\,\mathrm{d}t = 0$, donc $\mathrm{F}(a) = \mathrm{F}(b)$ ce qui entraîne que F est constante sur $[a\,;b]$, et sa dérivée est nulle *i.e.* $f = 0$ sur $[a\,;b]$. □

Remarque 29.10 – *Le théorème ci-dessus est faux si f n'est pas continue sur $[a\,;b]$, on peut considérer par exemple la fonction f définie par $f(t) = \begin{cases} 1 & \text{si } t = a \\ 0 & \text{sinon} \end{cases}$, cette fonction est positive, non nulle et d'intégrale nulle.*

2) Techniques de calculs (rappels)

Théorème 29.20 (intégration par parties)
Soient $f, g : \mathrm{I} \to \mathbb{C}$ deux fonctions de classe $\mathcal{C}^1$, alors :
$$\forall\, a, b \in \mathrm{I}, \int_a^b f'(u)g(u)\,\mathrm{d}u = [f(u)g(u)]_a^b - \int_a^b f(u)g'(u)\,\mathrm{d}u.$$

Théorème 29.21 (changement de variable)
Soit $\theta : \mathrm{J} \to \mathrm{I}$ une fonction de classe $\mathcal{C}^1$ sur l'intervalle J, et soit $f : \mathrm{I} \to \mathbb{C}$ une fonction continue sur l'intervalle I, alors on a : $\forall\, a, b \in \mathrm{J}, \int_a^b f(\theta(u))\theta'(u)\,\mathrm{d}u = \int_{\theta(a)}^{\theta(b)} f(t)\,\mathrm{d}t$.

☞**Exemples :**

- Soit $\mathrm{I} = \int_0^1 \sqrt{1-t^2}\,\mathrm{d}t$.
  On effectue le changement de variable $t = \sin(u)$ avec $u \in [0\,;\frac{\pi}{2}]$, on a alors $\mathrm{d}t = \cos(u)\,\mathrm{d}u$, d'où $\mathrm{I} = \int_0^{\pi/2} \sqrt{1-\sin(u)^2}\cos(u)\,\mathrm{d}u = \int_0^{\pi/2} \cos(u)^2\,\mathrm{d}u = \int_0^{\pi/2} \frac{1+\cos(2u)}{2}\,\mathrm{d}u = \left[\frac{u}{2} + \frac{\sin(2u)}{4}\right]_0^{\pi/2} = \frac{\pi}{4}$. En particulier on en déduit que la surface du cercle trigonométrique vaut π.
- Soit $\mathrm{I} = \int_0^{\pi/3} \ln(\cos(t))\sin(t)\,\mathrm{d}t$.
  On effectue le changement de variable $u = \cos(t)$ avec $\cos : [0\,;\frac{\pi}{3}] \to [\frac{1}{2}\,;1]$ ($\mathcal{C}^1$), on a $\mathrm{d}u = -\sin(t)\,\mathrm{d}t$ et donc $\mathrm{I} = \int_{1/2}^1 \ln(u)\,\mathrm{d}u = [u\ln(u) - u]_{1/2}^1 = \frac{\ln(2)-1}{2}$.

★**Exercice 29.1** *Calculer $\mathrm{I} = \int_0^1 t\sqrt{1-t^2}\,\mathrm{d}t$.*

IV INÉGALITÉS

Théorème 29.22 (inégalité de Cauchy-Schwarz [3])
Soit f, g deux fonctions continues par morceaux sur $[a\,;b]$ et ***à valeurs réelles****, on a :*
$$\left(\int_{[a;b]} fg\right)^2 \leqslant \left(\int_{[a;b]} f^2\right)\left(\int_{[a;b]} g^2\right) \text{ (inégalité de Cauchy-Schwarz).}$$

Preuve : L'application $(f, g) \mapsto \int_{[a;b]} fg$ n'est pas un produit scalaire sur $\mathcal{C}^0_{\mathrm{m}}([a\,;b], \mathbb{R})$ c'est néanmoins une forme bilinéaire symétrique et positive, on peut donc lui appliquer l'inégalité de Cauchy-Schwarz, dont on rappelle une preuve : posons $a = \int_{[a;b]} f^2$, $b = \int_{[a;b]} fg$ et $c = \int_{[a;b]} g^2$. Pour tout réel λ on a $0 \leqslant \int_{[a;b]} (\lambda f + g)^2$ (intégrale d'une fonction positive), en développant on obtient par linéarité $a\lambda^2 + 2b\lambda + c \geqslant 0$. Si $a \neq 0$ alors on a un trinôme du second degré qui est toujours positif, donc son discriminant est négatif ou nul, *i.e.* $b^2 - ac \leqslant 0$ ce qui donne exactement l'inégalité de Cauchy-Schwarz. Si $a = 0$ alors pour tout réel λ on a $2b\lambda + c \geqslant 0$ ce qui entraîne $b = 0$ et donc $b^2 \leqslant ac$. □

Dans le cas des fonctions continues sur $[a\,;b]$, l'application $(f, g) \mapsto \int_{[a;b]} fg$ est un produit scalaire sur $\mathcal{C}^0([a\,;b], \mathbb{R})$, on a donc comme dans tout espace préhilbertien réel :

Théorème 29.23 (cas d'égalité de Cauchy-Schwarz)
Si f, g sont ***continues*** *et à valeurs* ***réelles****, alors :*
$$\left(\int_{[a;b]} fg\right)^2 = \left(\int_{[a;b]} f^2\right)\left(\int_{[a;b]} g^2\right) \iff f \text{ et } g \text{ sont colinéaires.}$$

3. SCHWARZ Hermann (1846 – 1921) : mathématicien allemand.

☞**Exemple** : Soit $f : [0;1] \to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$ telle que $f(0) = 0$. Pour $t \in [0;1]$ on a $f(t) = \int_0^t f'(u)\,du$, d'où $f(t)^2 = \left(\int_0^t f'(u)\,du\right)^2 \leqslant \left(\int_0^t 1\right)\left(\int_0^t f'(u)^2\,du\right)$, ce qui entraîne $f(t)^2 \leqslant t\int_0^1 f'(u)^2\,du$, il en découle alors que $\int_0^1 f^2(t)\,dt \leqslant \frac{1}{2}\int_0^1 f'(u)^2\,du$.

Théorème 29.24 (Inégalité de la moyenne)
Soi $f : [a\,;b] \to \mathbb{C}$ continue par morceaux, majorée par $M \in \mathbb{R}^+$ en module, alors :

$$\left|\int_{[a;b]} f\right| \leqslant M(b-a).$$

☞**Exemple** : Soient $0 < a < b$ et soit $\varepsilon > 0$. D'après le théorème ci-dessus :
$\cos(\varepsilon b)\ln(\frac{b}{a}) \leqslant \int_{\varepsilon a}^{\varepsilon b} \frac{\cos(t)}{t}\,dt \leqslant \cos(\varepsilon a)\ln(\frac{b}{a})$. D'où : $\lim\limits_{\varepsilon\to 0}\int_{\varepsilon a}^{\varepsilon b} \frac{\cos(t)}{t}\,dt = \ln(\frac{b}{a})$.

V SOLUTION DES EXERCICES

Solution 29.1 *La fonction à intégrer est de la forme $au'u^{1/2}$ et s'intègre donc en $\frac{2a}{3}u^{3/2}$, d'où :*
$I = [-\frac{1}{3}(1-t^2)^{3/2}]_0^1 = \frac{1}{3}$.

Chapitre 30

Fonctions de deux variables

Sommaire

I CONTINUITÉ

1) Ouverts

On rappelle que $\mathbb{R}^2$ est un $\mathbb{R}$-espace vectoriel muni du produit scalaire canonique : si $u = (x, y)$ et $v = (x', y')$ alors $(u|v) = xx' + yy'$, la norme euclidienne est : $\|u\| = \sqrt{x^2 + y^2}$, celle-ci ayant les propriétés suivantes :

- $\forall\ u = (x, y) \in \mathbb{R}^2, \|u\| \geqslant 0$.
- $\forall\ u \in \mathbb{R}^2, \|u\| = 0 \iff u = 0$.
- $\forall\ u \in \mathbb{R}^2, \lambda \in \mathbb{R}, \|\lambda u\| = |\lambda|.\|u\|$.
- $\forall\ u, v \in \mathbb{R}^2, \|u + v\| \leqslant \|u\| + \|v\|$ (inégalité triangulaire).

Définition 30.1

- *Distance : la distance de $u \in \mathbb{R}^2$ à $v \in \mathbb{R}^2$ est la norme de la différence :* $\mathrm{d}(u, v) = \|u - v\|$.
- *Partie bornée : une partie* A *de $\mathbb{R}^2$ est dite* ***bornée*** *lorsqu'il existe un réel* M *tel que :*
$$\forall x \in \mathrm{A}, \|x\| \leqslant \mathrm{M}.$$
- *Boule ouverte : soit $u \in \mathbb{R}^2$ et $r > 0$, la boule ouverte de centre u et de rayon r est l'ensemble* $\mathrm{B}(u, r) = \{v \in \mathbb{R}^2\ /\ \|u - v\| < r\}$. *De même on peut définir les boules fermées et les sphères.*

À retenir

Si $u = (x, y)$ alors le pavé ouvert $\mathrm{P} =]x - \frac{r}{\sqrt{2}}; x + \frac{r}{\sqrt{2}}[\times]y - \frac{r}{\sqrt{2}}; y + \frac{r}{\sqrt{2}}[$ est inclus dans la boule ouverte $\mathrm{B}(u, r)$.

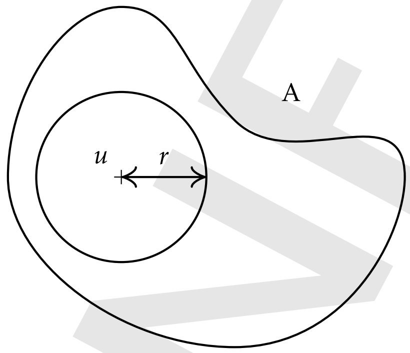
Théorème 30.1 (Ouverts de $\mathbb{R}^2$)

Une partie A *de $\mathbb{R}^2$ est un ouvert lorsque* A *est une réunion (quelconque) de boules ouvertes. Ce qui équivaut à :* $\forall u \in \mathrm{A}, \exists r > 0, \mathrm{B}(u, r) \subset \mathrm{A}$. *Par convention,* $\varnothing$ *est un ouvert.*

Preuve : Celle-ci est simple et laissée en exercice. □

Exemples :

☞**Exemples** :
- $\mathbb{R}^2$ est une partie ouverte de $\mathbb{R}^2$.
- Une boule ouverte est une partie ouverte de $\mathbb{R}^2$.
- Un demi-plan ouvert (*i.e.* bord exclu) est une partie ouverte.
- Une réunion quelconque de parties ouvertes est une partie ouverte.
- Une intersection finie de parties ouvertes est une partie ouverte.
- Une boule fermée n'est pas une partie ouverte de $\mathbb{R}^2$.



2) Fonctions de deux variables

Nous considérerons par la suite des fonctions définies sur une partie A de $\mathbb{R}^2$ et à valeurs dans $\mathbb{R}$.

☞**Exemples** :
- $f\colon (x,y) \mapsto x^2+y^2$ est définie sur $A=\mathbb{R}^2$.
- $f\colon (x,y) \mapsto \frac{1}{x^2-y^2}$ est définie sur $A=\mathbb{R}^2 \setminus \{(x,\pm x\ /\ x\in\mathbb{R}\}$ (c'est un ouvert de $\mathbb{R}^2$).

Soit $A\subset\mathbb{R}^2$, l'ensemble des fonctions de A vers $\mathbb{R}$ est noté $\mathcal{F}(A,\mathbb{R})$, il est facile de voir que pour les opérations usuelles sur les fonctions, c'est une $\mathbb{R}$-algèbre.

Définition 30.2 (représentation graphique)
Soit $f\colon A\to\mathbb{R}$ une fonction de deux variables, le graphe de la fonction est l'ensemble : $\{(x,y,f(x,y))\ /\ (x,y)\in A\}$. La représentation graphique de f est l'ensemble des points de $\mathbb{R}^3$ de coordonnées $(x,y,f(x,y))$, cet ensemble est appelé ***surface cartésienne*** *d'équation $z=f(x,y)$.*

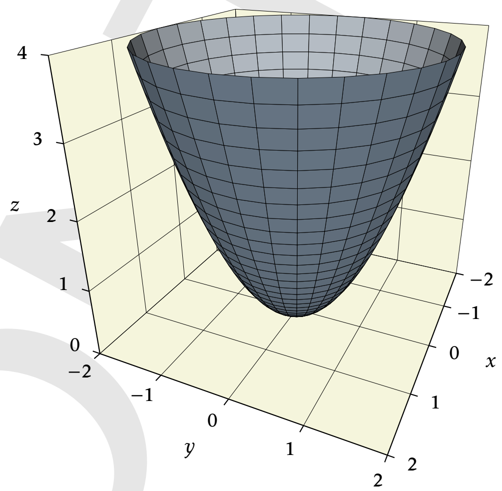


Figure 30.1 – Surface d'équation $z=x^2+y^2$ (tronquée à $z=4$)

Définition 30.3 (applications partielles)
Soit A une partie de $\mathbb{R}^2$, soit $f : A\to\mathbb{R}$ une fonction, et soit $a=(x_0,y_0)\in A$. La première application partielle de f en a est la fonction $f_{1,a} : t\mapsto f(t,y_0)$ (on fixe la deuxième variable à y_0), et la deuxième application partielle de f en a est la fonction $f_{2,a} : t\mapsto f(x_0,t)$ (on fixe la première variable à x_0).

☞**Exemple** : Soit $f(x,y)=\frac{x^2+y}{x^2+y^2+1}$, la première application partielle de f en $a=(0,0)$ est $f_{1,a}(t)=\frac{t^2}{1+t^2}$, et la deuxième application partielle de f en a est $f_{2,a}(t)=\frac{t}{1+t^2}$.

Remarque 30.1 – *Les applications partielles permettent de se ramener aux fonctions d'une variable réelle.*

3) Continuité

Définition 30.4 (continuité)
Soit $f\colon A \to \mathbb{R}$ et soit $a \in A$, on dit que f est continue en a lorsque :

$$\forall \varepsilon > 0, \exists r > 0, \forall u \in A, \|u - a\| < r \implies |f(u) - f(a)| < \varepsilon.$$

Si f est continue en tout point de A, on dit que f est continue sur A, l'ensemble des fonctions continues sur A est noté $\mathcal{C}^0(A, \mathbb{R})$.

Théorème 30.2
- *Si f est continue en $a \in A$, alors f est bornée au voisinage de a.*
- *Si f est continue en a et si $f(a) \neq 0$, alors au voisinage de a f ne s'annule pas.*
- *On retrouve les théorèmes généraux de la continuité (somme, produit, quotient). En particulier on en déduit que $\mathcal{C}^0(A, \mathbb{R})$ est une $\mathbb{R}$-algèbre.*
- *Si $f\colon A \to \mathbb{R}$ est continue sur A, et si $g\colon J \to \mathbb{R}$ est continue sur J avec $\operatorname{Im}(f) \subset J$, alors $g \circ f$ est continue sur A.*

Preuve : Celle-ci est simple et laissée en exercice. □

À retenir
Il en découle en particulier que toute fonction polynomiale ou rationnelle en x et y, est continue sur son ensemble de définition.

Théorème 30.3
Si f est continue en $a = (x_0, y_0) \in A$, alors la première application partielle de f en a est continue en x_0, et la deuxième est continue en y_0. Mais la réciproque est fausse.

Preuve : Soit $\varepsilon > 0$, il existe $r > 0$ tel que $\forall\, u \in A, \|u - a\| < r \implies |f(u) - f(a)| < \varepsilon$. Soit $t \in \mathbb{R}$, si $|t - x_0| < r$, alors $\|(t, y_0) - a\| = |t - x_0| < r$, donc $|f(t, y_0) - f(a)| < \varepsilon$, c'est à dire $|f_{1,a}(t) - f_{1,a}(x_0)| < \varepsilon$, ce qui prouve que $f_{1,a}$ est continue en x_0. Le raisonnement est similaire pour $f_{2,a}$.

Donnons un contre-exemple pour la réciproque : $f(x, y) = \frac{xy}{x^2+y^2}$ si $(x, y) \neq (0, 0)$, et $f(0, 0) = 0$. Les deux applications partielles de f en $(0, 0)$ sont continues en 0 car elles sont nulles. Par contre, on considérant les couples (x, x) non nuls, on a $f(x, x) = \frac{1}{2}$, donc on ne peut pas avoir par exemple $|f(x, x) - f(0, 0)| < \frac{1}{4}$ quand (x, x) est voisin de $(0, 0)$, donc f n'est pas continue en $(0, 0)$. □

4) Extension

Soit A un partie de $\mathbb{R}^2$ et $f\colon A \to \mathbb{R}^2$, alors pour tout couple (x, y) de A, $f(x, y)$ est un couple de réels dont les deux composantes sont fonctions de x et y, par conséquent il existe deux fonctions : $f_1, f_2\colon A \to \mathbb{R}$ telles que :

$$\forall (x, y) \in A, f(x, y) = (f_1(x, y), f_2(x, y)).$$

Par définition, les fonctions f_1 et f_2 sont les fonctions **composantes** de f.

Définition 30.5
*Une telle fonction f est dite **continue** en $a \in A$ lorsque les **fonctions composantes sont continues en** a.*

Remarque 30.2 :
- *Cela s'applique aussi aux fonctions à valeurs complexes.*
- *Cette définition se généralise aux fonctions à valeurs dans $\mathbb{R}^n$.*

II CALCUL DIFFÉRENTIEL

1) Dérivées partielles premières

Soit U un ouvert de $\mathbb{R}^2$ et soit $a = (x_0, y_0) \in \mathrm{U}$, il existe $\varepsilon > 0$ tel que $\mathrm{B}(a, \sqrt{2}\varepsilon) \subset \mathrm{A}$, par conséquent le pavé ouvert $]x_0 - \varepsilon\,; x_0 + \varepsilon[\times]y_0 - \varepsilon\,; y_0 + \varepsilon[$ est inclus dans U, donc la première application partielle de f en a est définie au moins sur l'intervalle $]x_0 - \varepsilon\,; x_0 + \varepsilon[$, et la deuxième sur $]y_0 - \varepsilon\,; y_0 + \varepsilon[$.

Définition 30.6

Si la première (respectivement la deuxième) application partielle de f en a est dérivable en x_0 (respectivement y_0), on dit que f admet une dérivée partielle par rapport à x (respectivement par rapport à y) en a, on la note : $\frac{\partial f}{\partial x}(a)$ (respectivement $\frac{\partial f}{\partial y}(a)$). Si f admet une dérivée partielle par rapport à x en tout point de U, alors on définit la fonction : $\frac{\partial f}{\partial x} : \begin{array}{ccc} \mathrm{U} & \to & \mathbb{R} \\ (x, y) & \mapsto & \frac{\partial f}{\partial x}(x, y) \end{array}$, *(même chose par rapport à y).*

Remarque 30.3 – *Les applications partielles sont des fonctions de* I *dans* $\mathbb{R}$ *où* I *est un intervalle de* $\mathbb{R}$, *on peut donc utiliser les théorèmes généraux pour étudier leur dérivabilité, et les règles de dérivation usuelles pour les calculs.*

☞**Exemple** : Soit $f(x, y) = \frac{x^2+y}{x^2+y^2+1}$ et soit $a = (x, y)$, on a $f_{1,a}(t) = \frac{t^2+y}{t^2+y^2+1}$ qui est dérivable sur $\mathbb{R}$ d'où $\frac{\partial f}{\partial x}(a) = \frac{2x(y^2-y+1)}{(x^2+y^2+1)^2}$; d'autre part $f_{2,a}(t) = \frac{x^2+t}{x^2+t^2+1}$ qui est dérivable sur $\mathbb{R}$, d'où $\frac{\partial f}{\partial y}(a) = \frac{x^2(1-2y)-y^2+1}{(x^2+y^2+1)^2}$.

2) Extremum

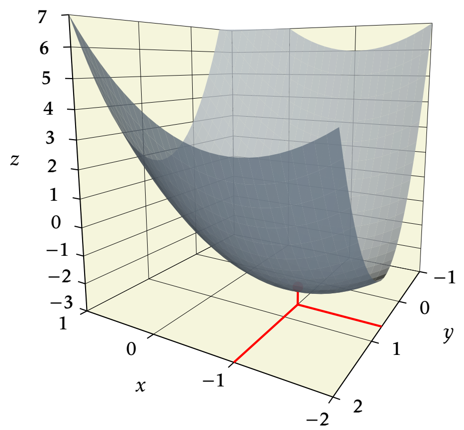
Définition 30.7

Soit $f : \mathrm{U} \to \mathbb{R}$ définie sur un ouvert U de $\mathbb{R}^2$. On dit que :
- *f admet un maximum global en $a = (x_0, y_0) \in \mathrm{U}$ lorsque $\forall u \in \mathrm{U}$, $f(u) \leqslant f(a)$.*
- *f admet un minimum global en $a = (x_0, y_0) \in \mathrm{U}$ lorsque $\forall u \in \mathrm{U}$, $f(a) \leqslant f(u)$.*
- *f admet un maximum local en $a = (x_0, y_0) \in \mathrm{U}$ lorsque $\exists r > 0$, $\forall u \in \mathrm{U} \cap \mathrm{B}(a, r)$, $f(u) \leqslant f(a)$.*
- *f admet un minimum local en $a = (x_0, y_0) \in \mathrm{U}$ lorsque $\exists r > 0$, $\forall u \in \mathrm{U} \cap \mathrm{B}(a, r)$, $f(a) \leqslant f(u)$.*

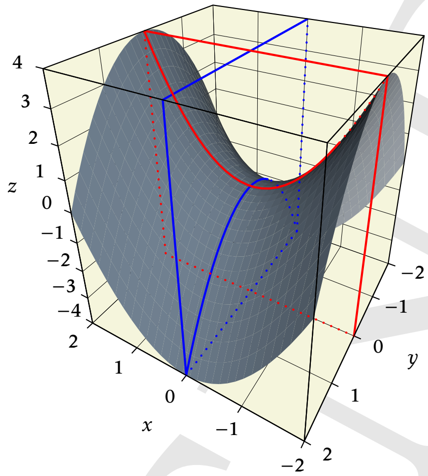
Théorème 30.4 (condition nécessaire pour un extremum)

Si $f : \mathrm{U} \to \mathbb{R}$ admet un extremum local en $a = (x_0, y_0) \in \mathrm{U}$, et si f admet ses deux dérivées partielles en a, alors $\frac{\partial f}{\partial x}(a) = 0$ et $\frac{\partial f}{\partial y}(a) = 0$, mais la réciproque est fausse.

Preuve : Supposons que a soit un maximum local, il existe donc $r > 0$ tel que $\mathrm{B}(a, r) \subset \mathrm{U}$ et $\forall\ u \in \mathrm{B}(a, r)$, $f(u) \leqslant f(a)$, par conséquent $\forall\ t \in]x_0 - \frac{r}{\sqrt{2}}; x_0 + \frac{r}{\sqrt{2}}[$, $f(t, y_0) \leqslant f(a)$, c'est à dire $f_{1,a}(t) \leqslant f_{1,a}(x_0)$, or la fonction $f_{1,a}(t)$ est dérivable en x_0 et x_0 est à 'intérieur de l'intervalle $]x_0 - \frac{r}{\sqrt{2}}; x_0 + \frac{r}{\sqrt{2}}[$, d'où $f'_{1,a}(x_0) = 0$, c'est à dire $\frac{\partial f}{\partial x}(a) = 0$, le raisonnement est le même pour la deuxième variable. □



$z = x^2 + 3y^2 + 2x - 4y$
minimum en $(-1, \frac{2}{3})$



$z = x^2 - y^2$,
pas d'extrêmum en $(0,0)$ (point col)

☞**Exemples :**

- Soit $f(x,y) = x^2+3y^2+2x-4y$, f admet ses deux dérivées partielles sur $\mathbb{R}^2$, qui sont $\frac{\partial f}{\partial x}(x,y) = 2x+2$ et $\frac{\partial f}{\partial y}(x,y) = 6y-4$, ces deux fonctions s'annulent pour $x = -1$ et $y = \frac{2}{3}$, donc le seul point où il peut y avoir un extremum est $a = (-1, \frac{2}{3})$. On a $f(x,y) = (x+1)^2 + 3(y-\frac{2}{3})^2 - \frac{7}{3}$, or $f(-1,\frac{2}{3}) = -\frac{7}{3}$, on voit donc que $f(x,y) \geqslant f(a)$, f présente donc un minimum global en a.
- Soit $f(x,y) = x^2-y^2$, f admet ses deux dérivées partielles sur $\mathbb{R}^2$ et $\frac{\partial f}{\partial x}(x,y) = 2x$ et $\frac{\partial f}{\partial y}(x,y) = -2y$, donc le seul point où f peut présenter un extremum est $a = (0,0)$, on a $f(a) = 0$, or si $t > 0$, on a $f(t,0) = t^2 > 0$ et $f(0,t) = -t^2 < 0$, donc f ne présente pas d'extremum en a (ce qui fournit un contre-exemple pour la réciproque du théorème).

Remarque 30.4 – *Soit $f\colon \mathrm{A} \to \mathbb{R}$ définie par $f(x,y) = 2x - y + 1$ avec $\mathrm{A} = \mathrm{B}'(0,1)$, alors en notant $u = (x,y)$ et $n = (2,-1)$ on a $f(x,y) = \langle u \mid n \rangle + 1$ et donc $1 - \|u\| \times \|n\| \leqslant f(u) \leqslant 1 + \|u\| \times \|n\|$, c'est à dire $1-\sqrt{5} \leqslant f(u) \leqslant 1+\sqrt{5}$, f est donc bornée, mais on voit que les bornes sont atteintes lorsque $u = \pm\frac{n}{\|n\|}$, f a donc un maximum et un minimum sur U. Mais si on observe les deux dérivées partielles : $\frac{\partial f}{\partial x}(x,y) = 2$ et $\frac{\partial f}{\partial y}(x,y) = -1$, ont voit qu'elles ne s'annulent jamais, **le théorème ne s'applique donc pas sur U, car ici, U n'est pas un ouvert**. Par contre, le théorème s'applique sur la boule ouverte $\mathrm{B}(0,1)$ et permet de dire que f ne présente pas d'extremum local sur la boule ouverte.*

3) Fonctions de classe $\mathcal{C}^1$

Définition 30.8

Soit U un ouvert de $\mathbb{R}^2$, soit $f\colon \mathrm{U} \to \mathbb{R}$ une fonction, on dit que f est de classe $\mathcal{C}^1$ sur U lorsque f admet ses deux dérivées partielles en tout point de U, et que celles-ci sont toutes deux continues sur U. L'ensemble des fonctions de classe $\mathcal{C}^1$ sur U est noté $\mathcal{C}^1(\mathrm{U}, \mathbb{R})$.

À retenir

Toute fraction rationnelle en x et y est de classe $\mathcal{C}^1$ sur son ensemble de définition.

Théorème 30.5 (DL1)

Si $f \in \mathcal{C}^1(\mathrm{U}, \mathbb{R})$, alors f admet un développement limité d'ordre 1 en tout point $a \in \mathrm{U}$, qui s'écrit :

$$f(a+h) = f(a) + h_1 \frac{\partial f}{\partial x}(a) + h_2 \frac{\partial f}{\partial y}(a) + \|h\|\varepsilon(h)$$

avec $\lim\limits_{\|h\| \to 0} \varepsilon(h) = 0$.

Preuve : On a (avec $a = (x, y)$ et $h = (h_1, h_2)$) :

$$\begin{aligned} f(x+h_1, y+h_2) - f(a) - h_1\frac{\partial f}{\partial x}(a) - h_2\frac{\partial f}{\partial y}(a) \\ &= f(x+h_1, y+h_2) - f(x, y+h_2) + f(x, y+h_2) - f(a) - h_1\frac{\partial f}{\partial x}(a) - h_2\frac{\partial f}{\partial y}(a) \\ &= h_1\frac{\partial f}{\partial x}(x+\theta h_1, y+h_2) + h_2\frac{\partial f}{\partial y}(x, y+\theta' h_2) - h_1\frac{\partial f}{\partial x}(a) - h_2\frac{\partial f}{\partial y}(a) \text{ avec } \theta, \theta' \in]0;1[\end{aligned}$$

d'où

$$\begin{aligned} |f(x+h_1, y+h_2) - f(a) - h_1\frac{\partial f}{\partial x}(a) - h_2\frac{\partial f}{\partial y}(a)| \\ &\leqslant |h_1||\frac{\partial f}{\partial x}(x+\theta h, y+h_2) - \frac{\partial f}{\partial x}(a)| + |h_2||\frac{\partial f}{\partial y}(x, y+\theta' h_2) - \frac{\partial f}{\partial y}(a)| \\ &\leqslant \|h\|\left(|\frac{\partial f}{\partial x}(x+\theta h, y+h_2) - \frac{\partial f}{\partial x}(a)| + |\frac{\partial f}{\partial y}(x, y+\theta' h_2) - \frac{\partial f}{\partial y}(a)|\right) \end{aligned}$$

les deux dérivées partielles étant continues, le terme entre parenthèses tend vers 0 lorsque h tend vers $(0, 0)$, ce qui termine la preuve. □

Définition 30.9 (gradient de f)

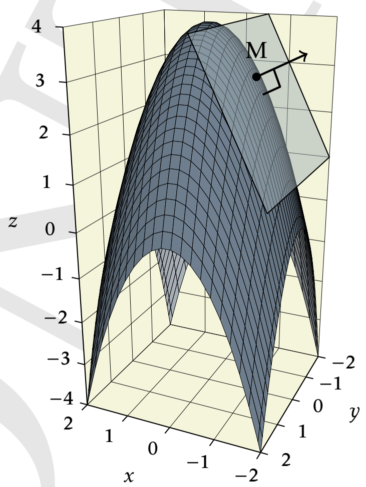
Si f est de classe $\mathcal{C}^1$ sur U, alors on pose pour $a \in$ U : $\text{Grad}_f(a) = \nabla f(a) = \left(\frac{\partial f}{\partial x}(a), \frac{\partial f}{\partial y}(a)\right)$*, c'est le* ***gradient de f en*** *a. En prenant le produit scalaire canonique de $\mathbb{R}^2$, le développement limité d'ordre 1 de f en a s'écrit :* $f(a+h) = f(a) + (\nabla f(a)|h) + \underset{0}{o}(h)$.

Définition 30.10 (plan tangent)

Si $f \in \mathcal{C}^1(\mathrm{U}, \mathbb{R})$, alors pour tout $a = (x_0, y_0) \in$ U, le plan d'équation :

$$z = f(x_0, y_0) + (x - x_0)\tfrac{\partial f}{\partial x}(a) + (y - y_0)\tfrac{\partial f}{\partial y}(a).$$

est appelé plan tangent à la surface $z = f(x, y)$ au point $\mathrm{M}(x_0, y_0, f(x_0, y_0))$.



Remarque 30.5 – *Le plan tangent en $a = (x_0, y_0)$ a donc pour équation :*

$$z - z_0 = (x - x_0)\tfrac{\partial f}{\partial x}(a) + (y - y_0)\tfrac{\partial f}{\partial y}(a).$$

un vecteur normal à ce plan est $\left(\frac{\partial f}{\partial x}(a), \frac{\partial f}{\partial y}(a), -1\right)$.

☞**Exemple** : $f(x, y) = \sqrt{1 - x^2 - y^2}$ avec $x^2 + y^2 \leqslant 1$, (demi-sphère de centre O de rayon 1).

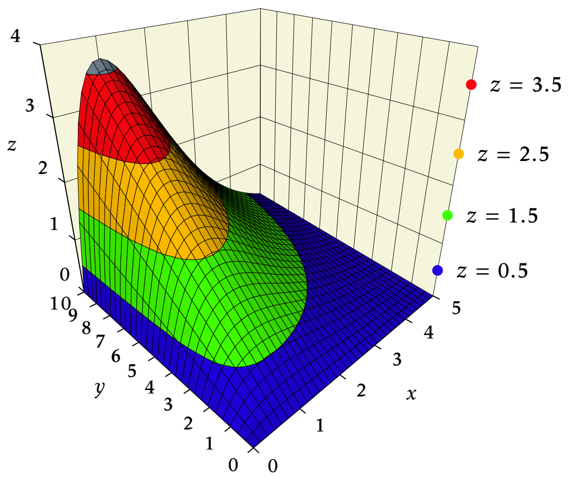
Soit $a = (x_0, y_0) \in \mathrm{B}(0, 1)$ et $z_0 = f(a)$, le plan tangent à la surface en $\mathrm{M} = (x_0, y_0, z_0)$ a pour équation $(x - x_0)\frac{\partial f}{\partial x}(a) + (y - y_0)\frac{\partial f}{\partial y}(a) = z - z_0$, ce qui donne $xx_0 + yy_0 + zz_0 = 0$, on remarque que le vecteur $\overrightarrow{\mathrm{OM}}$ est normal au plan tangent.

Définition 30.11 (courbe de niveau)

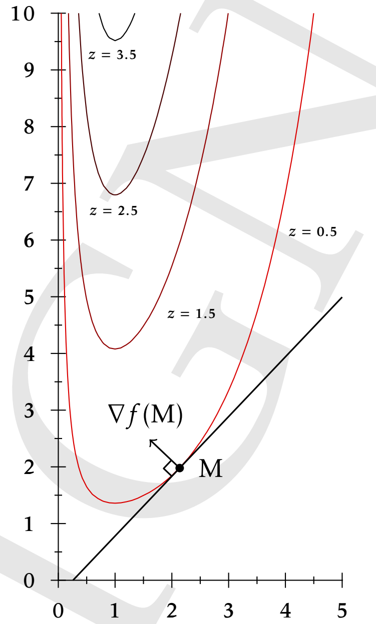
Soit $f\colon \mathrm{U} \to \mathbb{R}$ et soit $\lambda \in \mathbb{R}$, on appelle courbe de niveau d'équation $f(x, y) = \lambda$, l'ensemble :

$$\mathrm{C}_\lambda = \{(x, y, z) \,/\, z = f(x, y) = \lambda\}.$$

C'est l'intersection de la surface d'équation $z = f(x, y)$ avec le plan d'équation $z = \lambda$.



$z = f(x, y) = xye^{-x}$
courbes de niveau



même chose dans le plan $x\mathrm{O}y$

À retenir

Sur une courbe de niveau de f ($f(x, y) = \lambda$) le DL1 devient $(\nabla f(a) | \frac{h}{\|h\|}) = \underset{0}{o}(1)$ ce qui entraîne que dans le plan $z = \lambda$, la tangente à cette courbe « au point a » est la droite **orthogonale au vecteur gradient**.

4) Propriétés des fonctions $\mathcal{C}^1$

Théorème 30.6

- *Une fonction de classe $\mathcal{C}^1$ sur* U *est continue sur* U.
- $\mathcal{C}^1(\mathrm{U}, \mathbb{R})$ *est une $\mathbb{R}$-algèbre pour les lois usuelles sur les fonctions, c'est en fait une sous-algèbre de $\mathcal{C}^0(\mathrm{U}, \mathbb{R})$.*

Preuve : Soit $f\colon \mathrm{U} \to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$, et soit $a \in \mathrm{U}$, on peut écrire pour h voisin de $(0, 0)$: $f(a + h) = f(a) + h_1 \frac{\partial f}{\partial x}(a) + h_2 \frac{\partial f}{\partial y}(a) + \|h\|\varepsilon(h)$, on voit que $\lim\limits_{h \to (0,0)} f(a + h) = f(a)$, *i.e.* f est continue en a.

Montrons par exemple la stabilité pour l'addition : si f, g sont $\mathcal{C}^1$ sur U, soit $a = (x, y) \in \mathrm{U}$, la première application partielle de $f + g$ en a est $f_{1,a} + g_{1,a} \colon t \mapsto f(t, y) + g(t, y)$ or ces deux fonctions sont dérivables en x, donc $f + g$ admet une dérivée partielle par rapport à sa première variable et $\frac{\partial (f+g)}{\partial x}(a) = \frac{\partial f}{\partial x}(a) + \frac{\partial f}{\partial y}(a)$, or ces deux fonctions sont continues sur U et donc $\frac{\partial (g+h)}{\partial x}$ est continue sur U. Le raisonnement est le même pour la deuxième variable, finalement les deux dérivées partielles de $f + g$ sont continues sur U, donc $f + g$ est de classe $\mathcal{C}^1$ sur U.
□

Soit U un ouvert de $\mathbb{R}^2$, soit $a \in \mathrm{U}$, soit $f\colon \mathrm{U} \to \mathbb{R}$, et soit $h = (h_1, h_2) \in \mathbb{R}^2$ non nul, il existe $r > 0$ tel que $\mathrm{B}(a, r) \subset \mathrm{U}$, comme $\lim\limits_{t \to 0} a + th = a$, il existe $\varepsilon > 0$ tel que $t \in \left]-\varepsilon\,;\varepsilon\right[\implies a + th \in \mathrm{B}(a, r)$ et donc $a + th \in \mathrm{U}$, on peut alors considérer la fonction $g_{h,a} \colon t \mapsto f(a + th)$, c'est une fonction de $\mathbb{R}$ dans $\mathbb{R}$ définie au moins sur $\left]-\varepsilon\,;\varepsilon\right[$ (voisinage de 0).

Définition 30.12 (dérivée suivant un vecteur h)
Si la fonction $g_{h,a}$ ci-dessus est dérivable en 0, on dit que f ***admet une dérivée en a suivant le vecteur h****, et on pose $g'_{h,a}(0) = \mathrm{D}_h(f)(a)$.*

Exemple : Soit $f(x,y) = \sin(xy) + x - y$, soit $a = (0,0)$, et soit $h = (1,-2)$, on a alors $g_{h,a}(t) = f(t,-2t) = -\sin(2t^2) + 3t$, cette fonction est dérivable en 0 et $g'_{h,a}(0) = 3$, donc f admet une dérivée en a suivant le vecteur h et $\mathrm{D}_h(f)(a) = 3$.

Cas particuliers :

– f admet une dérivée partielle par rapport à la première variable en $a = (x_0, y_0)$ si et seulement si f admet une dérivée en a suivant le vecteur $(1,0)$.
Preuve : On a $g_{h,t} = f(x_0 + t, y_0) = f_{1,a}(x_0 + t)$, donc $g_{h,a}$ et dérivable en 0 ssi $f_{1,a}$ est dérivable en x_0. Si c'est le cas, alors $\mathrm{D}_h(f)(a) = \frac{\partial f}{\partial x}(a)$. □

– f admet une dérivée partielle par rapport à la deuxième variable en $a = (x_0, y_0)$ ssi f admet une dérivée en a suivant le vecteur $(0,1)$.
Preuve : On a $g_{h,t} = f(x_0, y_0 + t) = f_{2,a}(y_0 + t)$, donc $g_{h,a}$ et dérivable en 0 ssi $f_{2,a}$ est dérivable en y_0. Si c'est le cas, alors $\mathrm{D}_h(f)(a) = \frac{\partial f}{\partial y}(a)$. □

Théorème 30.7 (dérivée d'une composée : règle de la chaîne)
Soit U un ouvert de $\mathbb{R}^2$, soit I un intervalle de $\mathbb{R}$, soit $\varphi\colon \mathrm{I} \to \mathbb{R}^2$ définie par $\varphi(t) = (u_1(t), u_2(t))$ où u_1 et u_2 sont de classe $\mathcal{C}^1$ de I dans $\mathbb{R}$, avec $\mathrm{Im}(\varphi) \subset \mathrm{U}$, et soit $f\colon \mathrm{U} \to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$ sur U, alors la fonction $f \circ \varphi : \mathrm{I} \to \mathbb{R}$ est de classe $\mathcal{C}^1$ et :

$$\forall\, t \in \mathrm{I}, (f \circ \varphi)'(t) = u_1'(t)\frac{\partial f}{\partial x}(\varphi(t)) + u_2'(t)\frac{\partial f}{\partial y}(\varphi(t))$$

Preuve : $f \circ \varphi(t) = f(u_1(t), u_2(t))$, soit $t_0 \in \mathrm{I}$:

$$f[\varphi(t)] - f[\varphi(t_0)] = [u_1(t) - u_1(t_0)]\frac{\partial f}{\partial x}(\varphi(t_0)) + [u_2(t) - u_2(t_0)]\frac{\partial f}{\partial y}(\varphi(t_0)) + \mathrm{N}(\varphi(t) - \varphi(t_0))\varepsilon(\varphi(t) - \varphi(t_0)).$$

On divise tout par $t - t_0$, il est clair que la somme des deux premiers termes va tendre vers $u_1'(t_0)\frac{\partial f}{\partial x}(\varphi(t_0)) + u_2'(t_0)\frac{\partial f}{\partial y}(\varphi(t_0))$, et c'est une fonction continue de t_0, quant au reste, il devient : $\frac{|t-t_0|}{t-t_0}\left\|\frac{\varphi(t)-\varphi(t_0)}{t-t_0}\right\|\varepsilon(\varphi(t) - \varphi(t_0))$, il est facile de voir que cette expression a pour limite 0 lorsque t tend vers t_0, ce qui termine la preuve. □

★**Exercice 30.1** *La formule d'Euler. Soit $f\colon \mathrm{U} \to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$ sur U homogène de rapport $\alpha > 0$, i.e. $\forall\, a \in \mathrm{U}, f(ta) = t^\alpha f(a)$. On a alors : $x\frac{\partial f}{\partial x}(a) + y\frac{\partial f}{\partial y}(a) = \alpha f(a)$.*

Théorème 30.8
Soient U et V deux ouverts de $\mathbb{R}^2$, soit $\varphi\colon \mathrm{V} \to \mathrm{U}$ définie par $\varphi(x,y) = (\varphi_1(x,y), \varphi_2(x,y))$ où φ_1 et φ_2 sont de classe $\mathcal{C}^1$ à valeurs réelles, soit $f : \mathrm{U} \to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$, alors la fonction $f \circ \varphi\colon \mathrm{V} \to \mathbb{R}$ est de classe $\mathcal{C}^1$ sur V et $\forall\, a \in \mathrm{V}$:

$$\frac{\partial(f\circ\varphi)}{\partial x}(a) = \frac{\partial\varphi_1}{\partial x}(a) \times \frac{\partial f}{\partial x}(\varphi(a)) + \frac{\partial\varphi_2}{\partial x}(a) \times \frac{\partial f}{\partial y}(\varphi(a))$$

$$\frac{\partial(f\circ\varphi)}{\partial y}(a) = \frac{\partial\varphi_1}{\partial y}(a) \times \frac{\partial f}{\partial x}(\varphi(a)) + \frac{\partial\varphi_2}{\partial y}(a) \times \frac{\partial f}{\partial y}(\varphi(a))$$

Preuve : La première application partielle de $f \circ \varphi$ en $a = (x,y) \in \mathrm{V}$ est $(f\circ\varphi)_{1,a}(t) = f(\varphi_1(t,y), \varphi_2(t,y))$, il suffit alors d'appliquer le théorème précédent en prenant $u_1(t) = \varphi_1(t,y)$ et $u_2(t) = \varphi_2(t,y)$. □

Chapitre 31

Familles sommables

Sommaire

Problématique :
• Étant donné une famille de nombres $(a_i)_{i\in I}$, quel sens donner à $\sum\limits_{i\in I} a_i$ permettant de retrouver les propriétés usuelles (commutativité, associativité) ?
• Lorsque I est un ensemble fini la somme $\sum\limits_{i\in I} a_i$ a évidemment déjà un sens avec les propriétés bien connues.
• Lorsque $I = \mathbb{N}$, nous avons étudié la notion de série, cette notion va être généralisée.

I FAMILLES SOMMABLES DE RÉELS POSITIFS

1) Définition

Notation :
Si I désigne un ensemble, on notera $\mathcal{P}_f(I)$ l'ensemble des parties finies de I. Lorsque I est lui-même fini, on a bien sûr $\mathcal{P}_f(I) = \mathcal{P}(I)$.

Définition 31.1
*Une famille $(a_i)_{i\in I}$ de réels **positifs** est dite **sommable** lorsque l'ensemble :*
$$\left\{\sum_{i\in J} a_i \,/\, J \in \mathcal{P}_f(I)\right\} \text{ est un ensemble \textbf{majoré} dans } \mathbb{R}.$$
Si c'est le cas, la somme de la famille est : $\sum\limits_{i\in I} a_i = \sup\left\{\sum\limits_{i\in J} a_i \,/\, J \in \mathcal{P}_f(I)\right\}$.
Lorsque cet ensemble n'est pas majoré, on écrit $\sum\limits_{i\in I} a_i = +\infty$ *(c'est la borne sup. dans $\overline{\mathbb{R}}$ d'un ensemble non majoré).*

Remarque 31.1 :
- *Lorsque la famille est sommable, alors sa somme est dans* $\mathbb{R}+$.
- *Si la somme S de la famille est nulle, alors tous les termesa_i sont nuls (car $0 \leqslant a_i \leqslant S$ pour tout $i \in I$).*
- *Si I est fini, alors la famille $(a_i)_{i\in I}$ est sommable et sa somme est la somme usuelle des nombres a_i.*

☞**Exemples :**

– La famille $(2^{-n})_{n\in\mathbb{N}}$ est sommable de somme égale à 2.
En effet, si $J \in \mathcal{P}_f(\mathbb{N})$ alors il existe $n \in \mathbb{N}$ tel que $J \subset [\![0\,;n]\!]$, on a alors $\sum\limits_{k\in J} 2^{-k} \leqslant \sum\limits_{k=0}^{n} 2^{-k} = 2 - \frac{1}{2^n} < 2$, donc la famille est sommable et de somme $S \leqslant 2$, d'autre part $\sum\limits_{k=0}^{n} 2^{-k} = 2 - \frac{1}{2^n} \xrightarrow[n\to+\infty]{} 2$, donc $2 = \sup\left\{\sum\limits_{k\in J} 2^{-k} \,/\, J \in \mathcal{P}_f(\mathbb{N})\right\}$.

– Soit E un espace préhilbertien réel, soit $(e_i)_{i\in I}$ une famille orthonormale, pour $J \in \mathcal{P}_f(I)$ on pose $F_J = \mathrm{Vect}\left[(e_i)_{i\in J}\right]$. Soit $x \in E$, on sait que le projeté orthogonal de x sur F_J est $p_{F_J}(x) = \sum\limits_{i\in J} (x \mid e_i)\cdot e_i$. On a sait que $\|x\|^2 = \|p_{F_J}(x)\|^2 + \|x - p_{F_J}(x)\|^2$, et $\|p_{F_J}(x)\|^2 = \sum\limits_{i\in J} (x \mid e_i)^2$, on a donc $\sum\limits_{i\in J} (x \mid e_i)^2 \leqslant \|x\|^2$, par conséquent : la famille $((x \mid e_i)^2)_{i\in I}$ est sommable et $\sum\limits_{i\in I} (x \mid e_i)^2 \leqslant \|x\|^2$ (inégalité de Bessel).

Théorème 31.1

Soit $a(a_n)_{n\in\mathbb{N}}$ une suite de réels positifs, alors la famille $(a_n)_{n\in\mathbb{N}}$ est sommable si et seulement si la série de terme général a_n est convergente, auquel cas, la somme de la famille est la somme de la série. Dans le cas où la série est divergente, on écrira $\sum\limits_{n=0}^{+\infty} a_n = +\infty$.

Preuve : Si la famille est sommable, l'ensemble des sommes partielles est inclus dans $\left\{\sum\limits_{i\in J} a_i \,/\, J \in \mathcal{P}_f(\mathbb{N})\right\}$, cet ensemble est majoré par la somme S de la famille, la série et donc convergente (SATP) et la somme de la série est inférieure ou égal S. Soit $\varepsilon > 0$, il existe $J \subset \mathcal{P}_f(\mathbb{N})$ telle que $S - \varepsilon \leqslant \sum\limits_{k\in J} a_k$, or il existe $n \in \mathbb{N}$ tel que $J \in [\![0\,;n]\!]$ et donc $S - \varepsilon \leqslant \sum\limits_{k\in J} a_k \leqslant \sum\limits_{k=0}^{n} a_k \leqslant \sum\limits_{k=0}^{+\infty} a_k$, ce qui donne l'égalité des sommes.

Réciproquement, si la série et convergente, alors pour $J \in \mathcal{P}_f(\mathbb{N})$ il existe $n \in \mathbb{N}$ tel que $J \subset [\![0\,;n]\!]$ et donc $\sum\limits_{k\in J} a_k \leqslant \sum\limits_{k=0}^{n} a_k \leqslant \sum\limits_{k=0}^{+\infty} a_k \in \mathbb{R}$, ce qui entraîne que la famille est sommable. D'après ce qui précède, les deux sommes sont identiques. □

Théorème 31.2 (de comparaison)

Soient $(a_i)_{i\in I}$ et $(b_i)_{i\in I}$ deux familles de réels telles que $\forall i \in I$, $0 \leqslant a_i \leqslant b_i$. Si la famille $(b_i)_{i\in I}$ est sommable, alors la famille $(a_i)_{i\in I]}$ est sommable et $\sum\limits_{i\in I} a_i \leqslant \sum\limits_{i\in I} b_i$.

Preuve : Soit $J \in \mathcal{P}_f(I)$, alors $\sum\limits_{i\in J} a_i \leqslant \sum\limits_{j\in J} b_i \leqslant S_b$, on notant S_b la somme de la famille $(b_i)_{i\in I}$. Il en découle que $\left\{\sum\limits_{i\in J} a_i \,/\, J \in \mathcal{P}_f(I)\right\}$ est majoré par S_b, donc la famille $(a_i)_{i\in I}$ est sommable et de somme inférieure ou égale à S_b. □

2) Changement d'indice. Commutativité

Théorème 31.3 (changement d'indice)

*Soit $(a_i)_{i\in I}$ une famille de réels positifs, soit $\sigma\colon J \to I$ une **bijection**, alors la famille $(a_i)_{i\in I}$ est sommable si et seulement si la famille $(a_{\sigma(j)})_{j\in J}$ est sommable, auquel cas $\sum\limits_{i\in I} a_i = \sum\limits_{j\in J} a_{\sigma(j)}$.*

Preuve : Il faut d'abord remarquer que l'image d'une partie finie de J par σ est une partie finie de I. Supposons la famille $(a_i)_{i\in I}$ sommable de somme S_a.

Pour $j \in J$, posons $b_j = a_{\sigma(j)}$. Soit A une partie finie de J, alors $\sum\limits_{j\in A} b_j = \sum\limits_{j\in A} a_{\sigma(j)} = \sum\limits_{i\in\sigma(A)} a_i$ en posant $i = \sigma(j)$ (changement d'indice usuel dans une somme finie), et donc $\sum\limits_{j\in A} b_j \leqslant S_a$. On en déduit que la famille $(b_j)_{j\in J}$ est sommable et de somme $S_b \leqslant S_a$.

Soit $\varepsilon > 0$, il existe A partie finie de J telle que $S_b - \varepsilon \leqslant \sum\limits_{j\in A} b_j = \sum\limits_{i\in\sigma(A)} a_i \leqslant S_a$. Il en découle que $S_b = S_a$.

Réciproquement, si la famille $(b_j)_{j\in J}$ est sommable, alors on applique ce qui précède en considérant cette-fois ci $\sigma^{-1}\colon I \to J$ qui est bijective, en remarquant que $b_{\sigma^{-1}(i)} = a_i$. □

Théorème 31.4 (commutativité)
*Soit $(a_i)_{i\in I}$ une famille **sommable** de réels positifs, soit $\sigma : I \to I$ une **permutation de** I, alors la famille $(a_{\sigma(i)})_{i\in J}$ est sommable et $\sum\limits_{i\in I} a_i = \sum\limits_{i\in I} a_{\sigma(i)}$.*

Preuve : C'est une conséquence du théorème précédent avec $J = I$. □

Conséquence :
Si une série à termes positifs, de terme général a_n est convergente, alors pour toute permutation σ de $\mathbb{N}$, la série de terme général $a_{\sigma(n)}$ est convergente et $\sum\limits_{n=0}^{+\infty} a_n = \sum\limits_{n=0}^{+\infty} a_{\sigma(n)}$ (on dit que la série est commutativement convergente).

Théorème 31.5
Soient $(a_i)_{i\in I}$ et $(b_i)_{i\in I}$ deux familles sommables de réels positifs, et $\lambda \in \mathbb{R}^+$, alors les familles $(a_i + b_i)_{i\in I}$ et $(\lambda a_i)_{i\in I}$ sont sommables, avec les relations :

$$\sum_{i\in I} a_i + b_i = \sum_{i\in I} a_i + \sum_{i\in I} b_i \text{ et } \sum_{i\in I} \lambda a_i = \lambda \sum_{i\in I} a_i.$$

Preuve : Soit J une partie finie de I, alors $\sum\limits_{i\in J} a_i + b_i = \sum\limits_{i\in J} a_i + \sum\limits_{i\in J} b_i$ (sommes finies) et donc $\sum\limits_{i\in J} a_i + b_i \leqslant S_a + S_b$ (sommes respectives des deux familles), on en déduit que la famille $(a_i + b_i)_{i\in I}$ est sommable de somme inférieure ou égale à $S_a + S_b$.

Soit $\varepsilon > 0$, il existe J_1 partie finie de I telle que $S_a - \varepsilon \leqslant \sum\limits_{i\in J_1} a_i$ et il existe J_2 partie finie de I telle que $S_b - \varepsilon \leqslant \sum\limits_{i\in J_2} b_i$, il en découle que $S_a + S_b - 2\varepsilon \leqslant \sum\limits_{i\in J_1\cup J_2} a_i + b_i$, on en déduit que la somme de la famille $(a_i + b_i)_{i\in I}$ est supérieure ou égale à $S_a + S_b$, d'où le résultat.

La deuxième partie de la preuve est laissée en exercice. □

★**Exercice 31.1** *Soit $(a_i)_{i\in I}$ une famille de réels positifs, soit $J \subset I$, montrer que $\sum\limits_{i\in J} a_i \leqslant \sum\limits_{i\in I} a_i$.*

3) Sommation par paquets

Théorème 31.6 (premier résultat)
*Si $I = I_1 \cup I_2$ avec I_1 et I_2 **disjoints**, alors la famille de réels positifs $(a_i)_{i\in I}$ est sommable si et seulement si les sous-familles $(a_i)_{i\in I_1}$ et $(a_i)_{i\in I_2}$ sont sommables, auquel cas on a :*

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

Preuve : Si $(a_i)_{i\in I}$ est sommable de somme S_a, alors toute sous-famille est sommable (exercice précédent), notons S_1 la somme de la première et S_2 la deuxième. Soit $\varepsilon > 0$, il existe K_1 une partie finie de I_1 et K_2 une partie finie de I_2, telles que $S_1 - \varepsilon < \sum\limits_{i\in K_1} a_i$ et $S_2 - \varepsilon \leqslant \sum\limits_{i\in K_2} a_i$, $K = K_1 \cup K_2$ est une partie finie de I avec $K_1 \cap K_2 = \emptyset$, d'où $S_1 + S_2 - 2\varepsilon \leqslant \sum\limits_{i\in K_1} a_i + \sum\limits_{i\in K_2} a_i = \sum\limits_{i\in K_1\cup K_2} a_i \leqslant S_a$, on en déduit que $S_1 + S_2 \leqslant S_a$.

Supposons maintenant que les deux sous-familles sont sommables de sommes respectives S_1 et S_2, soit J une partie finie de I, alors $J = (J\cap I_1) \cup (J \cap I_1)$, $K_1 = J \cap I_1$ est une partie finie de I_1, et $K_2 = J \cap I_2$ est une partie finie de I_2 avec $K_1 \cap K_2 = \emptyset$, d'où $\sum\limits_{i\in J} a_i = \sum\limits_{i\in K_1} a_i + \sum\limits_{i\in K_2} a_i \leqslant S_1 + S_2$, on en déduit que la famille $(a_i)_{i\in I}$ est sommable de somme S_a inférieure ou égale à $S_1 + S_2$, et par conséquent $S_a = S_1 + S_2$. □

☞**Exemple** : On sait que la série de terme général $\frac{1}{n^2}$ est une SATP convergente, on admet que $\sum\limits_{n=1}^{+\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$. La famille $(\frac{1}{n^2})_{n\in\mathbb{N}^*}$ est sommable, on écrit $\mathbb{N} = I_1 \cup I_2$ avec $I_1 = \{2k+1 \,/\, k \in \mathbb{N}\}$ et $I_2 = \{2k \,/\, k \in \mathbb{N}^*\}$, ces deux ensembles sont disjoints. Les deux sous-familles $(\frac{1}{n^2})_{n\in I_1}$ et $(\frac{1}{n^2})_{n\in I_2}$ sont donc sommables de sommes S_1 et S_2, avec $S_1 + S_2 = \frac{\pi^2}{6}$.

Or $\sum\limits_{n\in I_2} \frac{1}{n^2} = \sum\limits_{k=1}^{+\infty} \frac{1}{4k^2} = \frac{1}{4}\sum\limits_{k=1}^{+\infty} \frac{1}{k^2} = \frac{\pi^2}{24}$. On en déduit que $\sum\limits_{k=0}^{+\infty} \frac{1}{(2k+1^2} = \sum\limits_{n\in I_1} \frac{1}{n^2} = \frac{\pi^2}{6} - \frac{\pi^2}{24} = \frac{\pi^2}{8}$.

Remarque 31.2 – *Par une récurrence immédiate, on peut généraliser le théorème précédent à une partition de I en p parties.*

Théorème 31.7 (sommation par paquets)
Si $I = \bigcup_{j\in J} I_j$ où les ensembles I_j sont deux à deux disjoints, si $(a_i)_{i\in I}$ est une famille de réels positifs, alors la famille est sommable si et seulement si :
- *$\forall j \in J$, la sous-famille $(a_i)_{i\in I_j}$ est sommable, de somme notée S_j,*
- *la famille $(S_j)_{j\in J}$ est sommable. On a alors :*

$$\sum_{i\in I} a_i = \sum_{j\in J} S_j = \sum_{j\in J}\left(\sum_{i\in I_j} a_i\right).$$

Preuve : Admis. □

★**Exercice 31.2** *Sommabilité et somme éventuelle de la famille $\left(\frac{1}{(p+q+1)^3}\right)_{(p,q)\in\mathbb{N}^2}$.*

★**Exercice 31.3** *Sommabilité d'une famille de réels positifs indexées par $\mathbb{Z}$, montrer que la famille $(a_n)_{n\in\mathbb{Z}}$ est sommable si et seulement si les séries $\sum a_n$ et $\sum a_{-n}$ convergent, que cela équivaut encore à a série $\sum(a_n + a_{-n})$ converge, et que cela équivaut encore à la suite $u_p = \sum_{n=-p}^{p} a_n$ a une limite finie quand $p \to +\infty$. Auquel cas on a :*

$$\sum_{n\in\mathbb{Z}} a_n = \sum_{n=0}^{+\infty} a_n + \sum_{n=1}^{+\infty} a_{-n} = a_0 + \sum_{n=1}^{+\infty}(a_n + a_{-n}) = \lim_{p\to+\infty}\sum_{n=-p}^{p} a_n.$$

Théorème 31.8 (Fubini)
Soit $(a_{i,j})_{(i,j)\in I\times J}$ une famille de réels positifs, alors cette famille est sommable si et seulement si :
- *$\forall i \in I$, la famille $(a_{i,j})_{j\in J}$ est sommable, de somme notée S_i*
- *la famille $(S_i)_{i\in I}$ est sommable. On a alors :*

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

On a bien sûr le même résultat en échangeant les rôles de I et de J.

Preuve : C'est une sommation par paquet en prenant comme partition de $I\times J$ les ensembles $J_i = \{(i,j) \,/\, j \in J\}$ pour $i \in I$. □

II FAMILLES SOMMABLES DE COMPLEXES

1) Définition

Définition 31.2
Une famille de complexes $(a_k)_{k\in I}$ est sommable lorsque la famille de réels positifs $(|a_k|)_{k\in I}$ est sommable. L'ensemble des familles sommables indexées par I est noté $\ell^1(I)$.

À retenir
Une série de terme général complexe a_n est donc une famille sommable si et seulement si cette série est absolument convergente. Par exemple la famille $(\frac{e^{in}}{n})_{n\in\mathbb{N}^*}$ n'est pas sommable alors que la famille $(\frac{e^{in}}{n^2})_{n\in\mathbb{N}^*}$ est sommable.

Cas particulier :
Si $(a_k)_{k\in I}$ est une famille à support fini, *i.e.* $\mathrm{Supp}(a) = \{k \in I \,/\, a_k \neq 0\}$ est un ensemble fini, alors cette famille est sommable puisque pour toute partie finie J de I, on a $\sum_{k\in J} |a_k| \leqslant \sum_{k\in\mathrm{Supp}(a)} |a_i|$ qui est une somme finie indépendante de J.

Théorème 31.9
L'ensemble $\ell^1(I)$ est un $\mathbb{C}$-e.v. (sev de $\mathbb{C}^I$).

Preuve : On a l'inclusion dans $\mathbb{C}^I$, la famille nulle est évidemment sommable. Si $(a_k)_{k\in I}$ et $(b_k)_{k\in I}$ sont sommables, alors en écrivant que $|a_k + b_k| \leqslant |a_k| + |b_k|$, on montre que $(a_k + b_k)_{k\in I}$ est sommable. Pour le dernier point on sait que si $(|a_k|)_{k\in I}$ est sommable alors $(|\lambda a_k|)_{k\in I}$ est sommable ($\lambda \in \mathbb{C}$). □

2) Somme d'une famille sommable

Rappel :

Si $x \in \mathbb{R}$, on note $x^+ = \max(0,x) = \frac{x+|x|}{2}$ et $x^- = \max(0,-x) = \frac{|x|-x}{2}$. On a alors que x^+ et x^- sont des réels positifs qui vérifient : $x^+ + x^- = |x|$ et $x^+ - x^- = x$.

Théorème 31.10

• *Une famille de* ***réels*** $(a_k)_{k\in I}$ *est sommable si et seulement si les familles* $(a_k^+)_{k\in I}$ *et* $(a_k^-)_{k\in I}$ *sont sommables.*

• *Une famille de* ***complexes*** $(a_k)_{k\in I}$ *est sommable si et seulement si les familles* $(\mathrm{Re}(a_k))_{k\in I}$ *et* $(\mathrm{Im}(a_k))_{k\in I}$ *sont sommables.*

Preuve : Pour le premier point, il suffit d'écrire que $0 \leqslant a_k^+ \leqslant |a_k|$ et $0 \leqslant a_k^- \leqslant |a_k|$, ce qui entraîne que si la famille $(a_k)_{k\in I}$ est sommable, alors les familles $(a_k^+)_{k\in I}$ et $(a_k^+)_{k\in I}$ sont sommables (par comparaison).

Réciproquement, si les familles $(a_k^+)_{k\in I}$ et $(a_k^+)_{k\in I}$ sont sommables, alors écrivant $|a_k| = a_k^+ + a_k^-$, on a que la famille $(a_k)_{k\in I}$ est sommable.

Si une famille de **complexes** $(a_k)_{k\in I}$ est sommable, alors en écrivant $|\mathrm{Re}(a_k)| \leqslant |a_k|$, cela montre que la famille $(\mathrm{Re}(a_k))_{k\in I}$ est sommable par comparaison (idem pour $(\mathrm{Im}(a_k))_{k\in I}$.

Réciproquement, si $(\mathrm{Re}(a_k))_{k\in I}$ et $(\mathrm{Im}(a_k))_{k\in I}$ sont sommables, alors en écrivant que $|a_k| \leqslant |\mathrm{Re}(a_k)| + |\mathrm{Im}(a_k)|$, on montre que la famille $(a_k)_{k\in I}$ est sommable. □

Définition 31.3

• *Si une famille de* ***réels*** $(a_k)_{k\in I}$ *est sommable, alors on pose :*
$$\sum_{k\in I} a_k = \sum_{k\in I} a_k^+ - \sum_{k\in I} a_k^-.$$
• *Si une famille de* ***complexes*** $(a_k)_{k\in I}$ *est sommable, alors on pose :*
$$\sum_{k\in I} a_k = \sum_{k\in I} \mathrm{Re}(a_k) + i\sum_{k\in I} \mathrm{Im}(a_k).$$

On remarquera que la définition 1 est cohérente avec le cas d'une famille de réels positifs, et que la définition 2 est cohérente avec le cas d'une famille réelle. D'autre part, si une famille de réels $(a_k)_{k\in I}$ est sommable, alors $\sum\limits_{k\in I} |a_k| = \sum\limits_{k\in I} a_k^+ + \sum\limits_{k\in I} a_k^-$ (d'après la partie I).

3) Propriétés

Théorème 31.11

Si $(a_k)_{k\in I}$ *est sommable de somme* S*, alors :*
$$\forall \varepsilon > 0, \exists K \in \mathcal{P}_f(I), \forall J \in \mathcal{P}_f(I), K \subset J \implies \left|\sum_{k\in J} a_k - S\right| < \varepsilon.$$

Preuve :

• Cas réel : $a_k = a_k^+ - a_k^-$, on sait que les familles $(a_k^+)_{k\in I}$ et $(a_k^+)_{k\in I}$ sont sommables de sommes respectives S_1 et S_2 (et donc la famille est de somme $S = S_1 - S_2$). Soit $\varepsilon > 0$, il existe une partie finie K_1 de I telle que $S_1 - \varepsilon < \sum\limits_{k\in K_1} a_k^+ \leqslant S_1$, de même, il existe une partie finie K_2 de I telle que $S_2 - \varepsilon < \sum\limits_{k\in K_2} a_k^- \leqslant S_2$, posons $K = K_1 \cup K_2$ (partie finie de I contenant K_1 et K_2), soit J une partie finie de I contenant K, alors on a $S_1 - \varepsilon < \sum\limits_{k\in K_1} a_k^+ \leqslant \sum\limits_{k\in J} a_k^+ \leqslant S_1$, et $S_2 - \varepsilon < \sum\limits_{k\in K_2} a_k^- \leqslant \sum\limits_{k\in J} a_k^- \leqslant S_2$, ce qui entraîne que $S_1 - S_2 - \varepsilon < \sum\limits_{k\in J} a_k^+ - \sum\limits_{k\in J} a_k^+ < S_1 - S_2 + \varepsilon$, et donc $S_1 - S_2 - \varepsilon < \sum\limits_{k\in J} a_k < S_1 - S_2 + \varepsilon$, d'où $\left|\sum\limits_{k\in J} a_k - S\right| < \varepsilon$.

• Cas complexe : soit $(a_k)_{k\in I}$ une famille sommable, on peut appliquer le théorème aux familles $(\mathrm{Re}(a_k))_{k\in I}$ et $(\mathrm{Im}(a_k))_{k\in I}$ qui sont sommables, de sommes respectives S_1 et S_2 (la famille a pour somme $S = S_1 + iS_2$), pour $\varepsilon < 0 <$, il existe une partie $K \in \mathcal{P}_f(I)$ tel que pour toute partie finie J de I contenant K, on a $\left|\sum\limits_{k\in J} \mathrm{Re}(a_k) - S_1\right| < \varepsilon$ et $\left|\sum\limits_{k\in J} \mathrm{Im}(a_k) - S_2\right| < \varepsilon$, ce qui entraîne : $\left|\sum\limits_{k\in J} a_k - S\right| < \left|\sum\limits_{k\in J} \mathrm{Re}(a_k) - S_1\right| + \left|\sum\limits_{k\in J} \mathrm{Im}(a_k) - S_2\right| < 2\varepsilon$. □

Théorème 31.12 (inégalité triangulaire)

Si $(a_k)_{k\in I}$ est sommable, alors $\left|\sum\limits_{k\in I} a_k\right| \leqslant \sum\limits_{k\in I} |a_k|$.

Preuve : Soit S la somme de la famille, soit $\varepsilon > 0$, il existe une partie finie J de I telle que $\left|S - \sum\limits_{k\in J} a_k\right| < \varepsilon$, d'où $|S| \leqslant \left|S - \sum\limits_{k\in J} a_k\right| + \sum\limits_{k\in J} |a_k| < \varepsilon + \sum\limits_{k\in I} |a_k|$, ce qui entraîne le résultat. □

Théorème 31.13 (linéarité de la somme)

Si $(a_k)_{k\in I}$ et $(b_k)_{k\in I}$ sont sommables, et $\lambda \in \mathbb{C}$, alors $\sum\limits_{k\in I} \lambda a_k + b_k = \lambda \sum\limits_{k\in I} a_k + \sum\limits_{k\in I} b_k$.

Preuve : Soit $S = \sum\limits_{k\in I} \lambda a_k + b_k$, $S_1 = \sum\limits_{k\in I} a_k$ et $S_2 = \sum\limits_{k\in I} b_k$, soit $\varepsilon > 0$, il existe une partie finie J de I telle que $\left|S_1 - \sum\limits_{k\in J} a_k\right| < \varepsilon$, $\left|S_2 - \sum\limits_{k\in J} b_k\right| < \varepsilon$, et $\left|S - \sum\limits_{k\in J} \lambda a_k + b_k\right| < \varepsilon$. D'où :

$$|S - (\lambda S_1 + S_2)| \leqslant \left|S - \sum_{k\in J} \lambda a_k + b_k\right| + |\lambda| \times \left|S_1 - \sum_{k\in J} a_k\right| + \left|S_2 - \sum_{k\in J} b_k\right| < (2+\lambda)\varepsilon,$$ d'où le résultat. □

Théorème 31.14 (changement d'indice, commutativité)

Si $(a_i)_{i\in I}$ est sommable alors :

- *Si $\sigma : J \to I$ est une bijection, alors la famille $(a_{\sigma(p)})_{p\in J}$ est sommable et $\sum\limits_{k\in I} a_k = \sum\limits_{p\in J} a_{\sigma(p)}$ (changement d'indice). On en déduit :*
- *Si $\sigma : I \to I$ est une bijection, alors la famille $(a_{\sigma(k)})_{k\in I}$ est sommable et $\sum\limits_{k\in I} a_k = \sum\limits_{k\in I} a_{\sigma(k)}$ (commutativité).*

Preuve : Pour les familles complexes on se ramène aux familles réelles en passant à la forme algébrique. Pour une famille sommable de réels $(a_i)_{i\in I}$, on sait que $\sum\limits_{i\in I} a_i = \sum\limits_{i\in I} a_i^+ - \sum\limits_{i\in I} a_i^-$, ce qui nous ramène aux familles de réels positifs pour lesquelles le théorème a déjà été démontré. □

À retenir

Ce théorème s'applique en particulier aux séries **absolument convergentes**, mais il est faux pour une série semi-convergente.

☞**Exemple** : La série de terme général $a_n = \frac{(-1)^n}{n}$ est semi-convergente et sa somme est $-\ln(2)$. Soit $\sigma : \mathbb{N}^* \to \mathbb{N}^*$ définie par $\sigma(3k+1) = 2k+1$, $\sigma(3k+2) = 4k+2$, $\sigma(3k+3) = 4k+4$, on vérifie que c'est bien une bijection. On vérifie ensuite que :

$u_{\sigma(3k+1)} + u_{\sigma(3k+2)} + u_{\sigma(3k+3)} = u_{2k+1} + u_{4k+2} + u_{4k+4} = \frac{1}{2}\left[\frac{-1}{2k+1} + \frac{1}{2k+2}\right]$, on peut alors en déduire que la série $\sum u_{\sigma(n)}$ converge et a pour somme $\frac{-\ln(2)}{2}$.

Théorème 31.15 (sommation par paquets)

Si $I = \bigcup\limits_{p\in J} I_p$ où les ensembles I_p sont deux à deux disjoints, si $(a_k)_{k\in I}$ est une famille sommable de complexes, alors :

- *$\forall p \in J$, la sous-famille $(a_k)_{k\in I_p}$ est sommable, de somme notée S_p,*
- *la famille $(S_p)_{p\in J}$ est sommable.*
- $\sum\limits_{k\in I} a_k = \sum\limits_{p\in p} S_p = \sum\limits_{p\in J} \left(\sum\limits_{k\in I_p} a_k\right)$.

Preuve : Admis. □

Théorème 31.16 (Fubini)

Soit $(a_{k,p})_{(k,p)\in I\times J}$ une famille sommable de complexes, alors :
- *$\forall k \in I$, la famille $(a_{k,p})_{p\in J}$ est sommable, de somme notée S_k*
- *la famille $(S_k)_{k\in I}$ est sommable.*
- $\sum\limits_{(k,p)\in I\times J} a_{k,p} = \sum\limits_{k\in I} S_k = \sum\limits_{k\in I}\left(\sum\limits_{p\in J} a_{k,p}\right)$.

On a bien sûr le même résultat en échangeant les rôles de I et de J.

Preuve : C'est une sommation par paquet en prenant comme partition de $I \times J$ les ensembles $J_k = \{(k,p) \,/\, p \in J\}$ pour $k \in I$. □

4) Applications

Théorème 31.17 (produit de Cauchy de séries AC)

Soient $\sum a_n$ et $\sum b_n$ deux séries absolument convergentes, pour $\in \mathbb{N}$, on pose $c_n = \sum\limits_{k=0}^{n} a_k \times b_{n-k}$.

Alors la série $\sum c_n$ est absolument convergente et $\sum\limits_{n=0}^{+\infty} c_n = \left(\sum\limits_{n=0}^{+\infty} a_n\right) \times \left(\sum\limits_{n=0}^{+\infty} b_n\right)$.

Preuve : On montre que la famille $(a_p b_q)_{(p,q)\in\mathbb{N}^2}$ est sommable : en appliquant le théorème de Fubini on peut écrire $\sum\limits_{(p,q)\in\mathbb{N}^2} |a_p b_q| = \sum\limits_{p\in\mathbb{N}}\left(\sum\limits_{q\in\mathbb{N}} |a_p b_q|\right)$ (propriété vérifiée pour les familles de réels positifs, c'est une égalité dans $\overline{\mathbb{R}}$). On a donc :

$$\sum_{(p,q)\in\mathbb{N}^2} |a_p b_q| = \sum_{p\in\mathbb{N}}\left(|a_p| \sum_{q\in\mathbb{N}} |b_q|\right) = \left(\sum_{p\in\mathbb{N}} |a_p|\right) \times \left(\sum_{q\in\mathbb{N}} |b_q|\right) < +\infty$$

ce qui montre que la famille $(a_p b_q)_{(p,q)\in\mathbb{N}^2}$ est effectivement sommable , on peut donc lui appliquer aussi le théorème de Fubini :

$$\sum_{(p,q)\in\mathbb{N}^2} a_p b_q = \sum_{p\in\mathbb{N}}\left(a_p \sum_{q\in\mathbb{N}} b_q\right) = \left(\sum_{p\in\mathbb{N}} a_p\right) \times \left(\sum_{q\in\mathbb{N}} b_q\right)$$

En considérant maintenant les ensembles $J_n = \{(p,q) \,/\, p+q = n\}$, qui forment une partition de $\mathbb{N}^2$ pour $n \in \mathbb{N}$, le théorème de sommation par paquets permet d'en déduire que les familles $(a_p b_q)_{(p,q)\in J_n}$ sont sommables, les sommes correspondantes sont les nombres c_n, que la famille $(c_n)_{n\in\mathbb{N}}$ est elle-même sommable, (donc la série $\sum c_n$ est bien absolument convergente), et que :

$$\sum_{(p,q)\in\mathbb{N}^2} a_p b_q = \sum_{n\in\mathbb{N}}\left(\sum_{(p,q)\in J_n} a_p b_q\right) = \sum_{n\in\mathbb{N}} c_n = \sum_{n\in\mathbb{N}}\left(\sum_{k=0}^{n} a_k b_{n-k}\right)$$

Et par conséquent, on a bien $\sum\limits_{n=0}^{+\infty} c_n = \left(\sum\limits_{n=0}^{+\infty} a_n\right) \times \left(\sum\limits_{n=0}^{+\infty} b_n\right)$. □

Théorème 31.18 (« vraie » définition, et propriété fondamentale, de la fonction exponentielle)

Soit $z \in \mathbb{C}$, la série $\sum \frac{z^n}{n!}$ est absolument convergente. On pose par définition $\exp(z) = \sum\limits_{n=0}^{+\infty} \frac{z^n}{n!}$, on a alors la propriété :

$$\forall (z,z') \in \mathbb{C}^2,\ \exp(z+z') = \exp(z) \times \exp(z').$$

Preuve :

- Convergence : pour $z = 0$ il n'y a rien à faire. Si $z \neq 0$, on pose $u_n = \frac{|z|^n}{n!}$, on a $\frac{u_{n+1}}{u_n} = \frac{|z|}{n+1} \to 0$, donc pour n supérieur à un certain entier N, on a $u_{n+1} \leqslant \frac{1}{2} u_n$, ce qui entraîne $0 \leqslant u_n \leqslant u_N \left(\frac{1}{2}\right)^{n-N}$, on a à droite le terme général d'une série géométrique convergente, par comparaison des SATP, la série $\sum u_n$ est convergente, ce qui entraîne le résultat.
- Propriété : les deux séries $\sum \frac{z^n}{n!}$ et $\sum \frac{z'^n}{n!}$ sont absolument convergentes, donc (produit de Cauchy) la série $\sum c_n$ où $c_n = \sum\limits_{k=0}^{n} \frac{z^k}{k!} \frac{z'^{n-k}}{(n-k)!}$, est absolument convergente, avec :

$$\sum_{n=0}^{+\infty} c_n = \left(\sum_{n=0}^{+\infty} \frac{z^n}{n!}\right) \times \left(\sum_{n=0}^{+\infty} \frac{z'^n}{n!}\right) = \exp(z) \times \exp(z')$$

or $c_n = \sum\limits_{k=0}^{n} \frac{z^k}{k!} \frac{z'^{n-k}}{(n-k)!} = \frac{1}{n!} \sum\limits_{k=0}^{n} \binom{n}{k} z^k z'^{n-k} = \frac{(z+z')^n}{n!}$, et donc $\sum\limits_{n=0}^{+\infty} c_n = \exp(z+z')$, ce qui donne le résultat. □

Table des matières