UNIVERSITÉ DE VERSAILLES
SAINT-QUENTIN-EN-YVELINES

DEUG MIAS premier niveau

Cours de mathématiques

année 2003/2004

Guillaume Legendre
legendre@math.uvsq.fr

(version révisée du 24 janvier 2018)

Table des matières

Chapitre 1

Éléments de logique

Nous rappelons ou introduisons dans ce chapitre préliminaire quelques éléments de logique et de vocabulaire mathématique indispensables à la compréhension des démonstrations des chapitres suivants.

1.1 Assertions

Dans le cadre d'une théorie mathématique donnée, une *assertion* est une phrase mathématique à laquelle on peut attribuer une, et une seule, *valeur booléenne*, à savoir *vrai* (V en abrégé) ou *faux* (F en abrégé). Certaines assertions sont déclarées vraies *a priori* : ce sont les *axiomes* ; sinon, la véracité d'une assertion doit résulter d'une *démonstration*.

Les assertions démontrées sont appelées *théorèmes* ou *propositions* suivant leur importance. Un *lemme* est un résultat préalable utile à une démonstration plus conséquente, tandis qu'un *corollaire* est une assertion vraie qui découle d'un résultat précédent.

1.1.1 Opérations sur les assertions

À une assertion, on peut associer sa *négation*, à deux assertions, leur *disjonction* ou leur *conjonction*. Les valeurs booléennes des ces assertions associées dépendent des valeurs des assertions de départ et sont données par les définitions suivantes.

Définition 1.1 *La **négation** d'une assertion* (P) *se note* $(\text{non } P)$*. L'assertion* $(\text{non } P)$ *est vraie si et seulement si* (P) *est fausse.*

Définition 1.2 *Étant données deux assertions* (P) *et* (Q)*, on appelle **disjonction** de ces deux assertions l'assertion* $(P \text{ ou } Q)$ *qui est vraie si et seulement si l'une au moins des assertions est vraie.*

Définition 1.3 *Étant données deux assertions* (P) *et* (Q)*, on appelle **conjonction** de ces deux assertions l'assertion* $(P \text{ et } Q)$ *qui est vraie si et seulement si les deux assertions sont simultanément vraies.*

À ces trois opérations ou *connecteurs*, on ajoute les opérations d'*implication* et d'*équivalence*.

Définition 1.4 *Étant données deux assertions* (P) *et* (Q)*, on définit l'assertion « (P) implique (Q) », appelée **implication** et notée* $(P \Rightarrow Q)$*, de la manière suivante : l'assertion* $(P \Rightarrow Q)$ *est vraie quand* (P) *est fausse ou lorsque* (P) *et* (Q) *sont vraies.*

On dit encore que « (P) *est une condition suffisante pour* (Q) » et que « (Q) *est une condition nécessaire pour* (P) ». L'assertion (P) est appelée l'*hypothèse* et (Q) la *conséquence*.

Définition 1.5 *Étant données deux assertions* (P) *et* (Q)*, on définit l'assertion « (P) équivaut à (Q) », appelée **équivalence** et notée* $(P \Longleftrightarrow Q)$*, de la manière suivante : l'assertion* $(P \Longleftrightarrow Q)$ *est vraie si* (P) *et* (Q) *sont toutes les deux vraies ou fausses.*

On dit encore que « (P) *et* (Q) *sont équivalentes* » ou que « (P) *est une condition nécessaire et suffisante pour* (Q) ».

1.1.2 Tableaux de vérité

La définition des assertions précédentes est résumée dans les *tableaux de vérité* ci-dessous.

(P)	$(\text{non } P)$
V	F
F	V

(P)	(Q)	$(P \text{ ou } Q)$
V	V	V
V	F	V
F	V	V
F	F	F

(P)	(Q)	$(P \text{ et } Q)$
V	V	V
V	F	F
F	V	F
F	F	F

(P)	(Q)	$(P \Rightarrow Q)$
V	V	V
V	F	F
F	V	V
F	F	V

(P)	(Q)	$(P \Leftrightarrow Q)$
V	V	V
V	F	F
F	V	F
F	F	V

Les opérations que nous venons d'introduire peuvent être répétées pour former des assertions dépendant d'assertions initiales (P_1), (P_2), (P_3), etc... Deux assertions ainsi construites sont dites *synonymes* si elles ont le même tableau de vérité.

Exemple. Montrons que lorsque l'on a simultanément $(P \Rightarrow Q)$ et $(Q \Rightarrow P)$, on a $(P \Leftrightarrow Q)$. On établit pour cela le tableau de vérité suivant

(P)	(Q)	$(P \Rightarrow Q)$	$(Q \Rightarrow P)$	$((P \Rightarrow Q) \text{ et } (Q \Rightarrow P))$	$(P \Leftrightarrow Q)$
V	V	V	V	V	V
V	F	F	V	F	F
F	V	V	F	F	F
F	F	V	V	V	V

Quelques synonymies classiques. Soient (P), (Q) et (R) trois propositions. On a alors
— (non $(P$ ou $Q)$) est équivalente à (non P et non Q),
— (non $(P$ et $Q)$) est équivalente à (non P ou non Q),
— $(P \Rightarrow Q)$ est équivalente à (non $Q \Rightarrow$ non P),
— (P ou (Q et R)) est équivalente à ((P ou Q) et (P ou R)),
— (P et (Q ou R)) est équivalente à ((P et Q) ou (P et R)).
La preuve des ces affirmations est laissée en exercice au lecteur (il suffit d'écrire les tableaux de vérité).

1.2 Quantificateurs

Les *quantificateurs* sont des symboles utilisés pour écrire des énoncés. Une *phrase quantifiée* est une assertion mathématique contenant un ou des quantificateurs.
Le symbole $\exists$ désigne le *quantificateur existentiel*. Ainsi, $\exists x$ se lit *« il existe au moins un élément x »* et $\exists! x$ signifie *« il existe un et un seul élément x »*.
Le symbole $\forall$ désigne le *quantificateur universel* et $\forall x$ signifie *« pour tout élément x »*.
La lettre affectée par un quantificateur est muette et peut être remplacée par n'importe quelle autre lettre n'ayant pas déjà une signification dans l'énoncé.

Exemple. La notation $x \in E$ signifie *« x appartient à E »*. Soit par ailleurs $P(x)$ une assertion dépendant de x. Alors

$$(\forall x \in E,\ P(x)) \Longleftrightarrow (\forall y \in E,\ P(y)),$$
$$(\exists x \in E,\ P(x)) \Longleftrightarrow (\exists y \in E,\ P(y)).$$

La *négation d'une phrase quantifiée* se définit comme suit :

$$(\text{non } (\forall x \in E,\ P(x))) \Longleftrightarrow (\exists x \in E,\ \text{non } P(x)),$$
$$(\text{non } (\exists x \in E,\ P(x))) \Longleftrightarrow (\forall x \in E,\ \text{non } P(x)).$$

1.3 La démonstration en mathématiques

Dans le cadre d'une théorie mathématique, une démonstration consiste, en général, à établir une proposition de la forme $(H \Rightarrow C)$ avec (H) vraie (sinon, il n'y a rien à faire), ou bien à montrer qu'une assertion (C) est vraie. On dispose pour cela non seulement de (H) mais aussi de toutes les propositions de la théorie concernée déjà établies. On met alors en évidence une suite d'assertions (P_0), (P_1),..., (P_n), où (P_0) est (H) et (P_n) est (C), telles que la vérité de chaque (P_i), $i = 1, \dots, n$, puisse être déduite des précédentes par l'intermédiaire des propositions établies et des règles de logique.

Il existe plusieurs types de démonstration : la *démonstration directe*, la *démonstration par l'absurde*, la *démonstration par contraposé* et la *démonstration par récurrence*.

1.3.1 Démonstration directe

La démonstration directe de $(H \Rightarrow C)$ consiste à établir successivement des assertions intermédiaires pour arriver à (C) en partant de (H).

Exemple de démonstration directe. Démontrons directement que, pour tout entier naturel n, on a n est impair ou n^2 est pair.
Notons (P) l'assertion *« l'entier n est impair »* et (Q) l'assertion *« l'entier n^2 est pair »*. Dans ce cas, (H) est *« n est un entier naturel »* et (C) est $(P$ ou $Q)$. Nous avons à établir que $(P$ ou $Q)$ est vraie. Supposons que l'assertion *« n est impair »* soit fausse. Alors n est pair, il existe un entier p tel que $n = 2p$. Il s'ensuit que $n^2 = 4p^2$, et donc n^2 est pair.

1.3.2 Démonstration par l'absurde

La démonstration par l'absurde qu'une assertion (P) est vraie consiste à supposer que (P) est fausse et à montrer que (non $P \Rightarrow Q$) est vraie, avec (Q) une assertion dont on sait qu'elle est fausse d'où une contradiction. Il en résulte que (P) est vraie car (non $P \Rightarrow Q$) est équivalente à (non $Q \Rightarrow P$) (on pourra vérifier cette affirmation en exercice).

Exemple de démonstration par l'absurde. Démontrons que le réel $\sqrt{2}$ est irrationnel.
Ici, l'assertion (P) est *« $\sqrt{2}$ est irrationnel »*. La négation de (P) est donc *« $\sqrt{2}$ est rationnel »*. Nous la supposons vraie et il existe alors deux entiers naturels p et q (avec $q \neq 0$) premiers entre eux[1] tels que $\sqrt{2} = \frac{p}{q}$. De $2q^2 = p^2$, nous déduisons que p^2 est pair et donc que p est pair (voir l'exemple suivant). Écrivant ensuite $p = 2k$ avec k entier, il s'ensuit que $q^2 = 2k^2$, montrant ainsi que q^2, et donc q, est pair. Les nombres p et q ont 2 comme diviseur commun, ce qui est contradictoire avec le fait que p et q sont premiers entre eux.

1.3.3 Démonstration par contraposée

La démonstration par contraposée s'appuie sur la règle logique :

$$(P \Rightarrow Q) \text{ est équivalente à } (\text{non } Q \Rightarrow \text{non } P).$$

Exemple de démonstration par contraposée. Soit n un entier naturel. Démontrons que si n^2 est pair alors n l'est aussi.
Ici, (P) est l'assertion *« n^2 est pair »* et (Q) est l'assertion *« n est pair »*. La négation de (Q) est alors *« n est impair »*. Supposons (non Q) vraie. Dans ce cas, $n = 2k + 1$, k étant un entier, et donc $n^2 = 4k^2 + 4k + 1$ est impair.

1.3.4 Démonstration par récurrence

La démonstration par récurrence s'utilise pour prouver des propositions dont l'énoncé dépend d'un entier naturel n. Elle s'appuie sur une propriété (admise) particulière de l'ensemble des entiers naturels $\mathbb{N}$.

1. C'est-à-dire sans diviseur commun autre que 1.

Théorème 1.6 *Toute partie de A de $\mathbb{N}$, contenant l'entier p et telle que, pour tout n appartenant à A, $n+1$ appartient à A, contient l'ensemble des entiers plus grands ou égaux à p.*

Corollaire 1.7 *Soient $(P(n))$ une proposition dépendant de l'entier n et p un entier naturel. Si on a*

i) $(P(p))$ *est vraie,*

ii) Pour tout entier naturel $n \geq p$, $(P(n) \Rightarrow P(n+1))$,

alors $(P(n))$ est vraie pour tout entier naturel $n \geq p$.

DÉMONSTRATION. On pose A l'ensemble des entiers naturels n pour lesquels la proposition $(P(n))$ est vraie et on applique le théorème 1.6. On en déduit que A contient l'ensemble des entiers plus grands ou égaux à p. □

Exemple de démonstration par récurrence. Démontrons par récurrence que, pour tout entier naturel n, l'entier $n^3 - n$ est divisible par 3.
Posons tout d'abord $(P(n))$ l'assertion « *$n^3 - n$ est divisible par* 3 ». Nous vérifions ensuite que

i) $0 = 0 \times 3$ donc $0^3 + 0 = 0$ est divisible par 3 et $(P(0))$ est vraie.

ii) Soit n un entier naturel tel que $(P(n))$ est vraie (on dit que c'est l'hypothèse de récurrence). Alors il existe un entier k tel que $n^3 - n = 3k$ et

$$(n+1)^3 - (n+1) = n^3 + 3n^2 + 3n + 1 - n - 1 = n^3 - n + 3(n^2 + n) = 3k + 3(n^2 + n).$$

Donc $(n+1)^3 - (n+1)$ est divisible par 3. On conclut que si $(P(n))$ est vraie, alors $(P(n+1))$ est vraie.

La démonstration est alors terminée.

1.3.5 Le contre-exemple

Pour infirmer une assertion, on peut utiliser un exemple ou un cas particulier qui la contredit, qu'on appelle alors un *contre-exemple*.

Première partie

Algèbre

Chapitre 2

Ensembles et relations

Le cadre dans lequel nous allons travailler est celui de la théorie des ensembles. L'objectif de ce chapitre est d'en présenter une théorie « naïve » : langage des ensembles, opérations sur les ensembles, etc...

2.1 Ensembles

En mathématiques, on étudie des objets de différents types : des *points*, des *nombres* ou encore des *vecteurs* par exemple. Ces objets ou *éléments* forment ou constituent, en vertu de certaines propriétés, des collections ou *ensembles*. Parmi les ensembles déjà rencontrés par l'étudiant, citons notamment

— $\mathbb{N}$ l'ensemble des entiers naturels,
— $\mathbb{Z}$ l'ensemble des entiers relatifs,
— $\mathbb{Q}$ l'ensemble des nombres rationnels,
— $\mathbb{R}$ l'ensemble des nombres réels,
— $\mathbb{C}$ l'ensemble des nombres complexes,
— $\mathscr{F}(\mathbb{N},\mathbb{R})$ l'ensemble des suites réelles ou des fonctions réelles définies sur $\mathbb{N}$,
— si a et b sont deux réels, $[a,b[$ est l'ensemble des réels x qui vérifient $a \leq x < b$.

On notera, en général, un élément par une lettre minuscule (l'élément x) et un ensemble par une lettre majuscule (l'ensemble E).

Un ensemble peut être *fini* ou *infini*. S'il est fini, il peut être donné en *extension*, c'est-à-dire par la liste (non ordonnée) de ses éléments, *a priori* supposés distincts. S'il est infini (ou même fini), l'ensemble peut être donné en *compréhension*, c'est-à-dire par une ou des propriétés définissant ses éléments. Nous reviendrons brièvement sur les notions d'ensembles finis et infinis dans la section 2.4.

Un ensemble formé d'un et un seul élément est appelé un *singleton*.

Définition 2.1 *On dit que* ***l'ensemble*** E ***est égal à l'ensemble*** F *(et l'on note $E = F$) si tout élément de E est un élément de F et si tout élément de F est un élément de E. Lorsque les ensembles E et F ne sont pas égaux, ils sont dits* ***distincts*** *et l'on note $E \neq F$.*

2.1.1 Appartenance

La *relation d'appartenance* se note $x \in E$ et s'énonce *« x appartient à l'ensemble E »*, ou encore *« x est un élément de E »*. La négation de cette relation est une autre relation, qui se note $x \notin E$ et s'énonce *« x n'appartient pas à l'ensemble E »*, ou encore *« x n'est pas un élément de E »*.

Exemples. Le nombre 1 est un entier naturel, ce qui se représente par $1 \in \mathbb{N}$. De même, le nombre π est un nombre réel, d'où $\pi \in \mathbb{R}$, mais il n'est pas un nombre rationnel, d'où $\pi \notin \mathbb{Q}$.

2.1.2 Inclusion

On définit la *relation d'inclusion* de la manière suivante.

Définition 2.2 *On dit qu'un ensemble E est* **inclus** *dans un ensemble F, ce que l'on note $E \subset F$, si et seulement si tout élément de E appartient à F :*

$$E \subset F \Longleftrightarrow (\forall x \in E,\ x \in F).$$

Cette relation est synonyme de $F \supset E$, qui se lit « *F contient E* ». Par ailleurs, on note $E \not\subset F$ la négation de $E \subset F$:

$$E \not\subset F \Longleftrightarrow (\exists x \in E,\ x \notin F).$$

Lorsque $E \subset F$ et qu'il existe au moins un élément de F qui n'appartient pas à E, on dit que E est un *sous-ensemble propre de F*, ce qui est noté $E \subsetneq F$.

Exemple. L'ensemble constitué de l'entier 1, ou ayant pour unique élément l'entier 1, est le singleton noté $\{1\}$ et l'on a $\{1\} \subset \mathbb{N}$.

Les propriétés suivantes sont immédiates. Soient E, F et G trois ensembles, on a
- — $E \subset E$.
- — ($E \subset F$ et $F \subset G$) $\Rightarrow E \subset G$ (transitivité de l'inclusion).

Proposition 2.3 *Étant donnés deux ensembles E et F, on a $E = F$ si et seulement si l'on a simultanément $E \subset F$ et $F \subset E$.*

2.1.3 Partie d'un ensemble. Ensemble des parties

Définition 2.4 *Soit E un ensemble. On appelle* **partie** *de E (ou* **sous-ensemble** *de E) tout ensemble A vérifiant $A \subset E$.*

Exemple. Dans l'ensemble $\mathbb{N}$ des entiers naturels, la propriété de divisibilité par 2 définit une partie A de $\mathbb{N}$: l'ensemble des nombres pairs. Un entier naturel n appartient à A s'il existe un entier naturel p tel que $n = 2p$. On note alors

$$A = \{n \in \mathbb{N} \mid \exists p \in \mathbb{N},\ n = 2p\}.$$

La partie de E n'ayant aucun élément se nomme l'*ensemble vide de E* et se note $\varnothing$. Si E est un ensemble alors $\varnothing \subset E$. En effet, sinon, en raisonnant par l'absurde, il existerait au moins un élément appartenant à $\varnothing$ qui n'appartiendrait pas à E. Or, ceci est impossible, puisque l'ensemble vide n'a pas d'élément. Donc, $\varnothing \subset E$ est une assertion vraie.

Enfin, toutes les parties de E constituent un nouvel ensemble, noté $\mathscr{P}(E)$, que l'on nomme *ensemble des parties de E*. Pour tout ensemble E, E et $\varnothing$ appartiennent à $\mathscr{P}(E)$.

Remarques. On a

1) $A \in \mathscr{P}(E) \Leftrightarrow A \subset E$.
2) $\{x\} \in \mathscr{P}(E) \Leftrightarrow x \in E$.

Exemple. Si $E = \{0, 1, 2\}$, on a

$$\mathscr{P}(E) = \{\varnothing, \{0\}, \{1\}, \{2\}, \{0,1\}, \{0,2\}, \{1,2\}, \{0,1,2\}\}.$$

Définition 2.5 *Soient E un ensemble et $\mathscr{P}$ une partie de $\mathscr{P}(E)$. On dit que $\mathscr{P}$ est une* **partition de** *E si et seulement si*

(i) $\forall A \in \mathscr{P}$, $A \neq \varnothing$,

(ii) $\forall A \in \mathscr{P}$, $\forall B \in \mathscr{P}$, $(A \neq B \Leftrightarrow A \cap B = \varnothing)$,

(iii) $\forall x \in E$, $\exists A \in \mathscr{P}$, $x \in A$.

Exemple. Pour tout ensemble non vide E, $\{E\}$ et $\{\{x\} \mid x \in E\}$ sont des partitions de E.

2.1.4 Opérations sur les ensembles

Considérons un ensemble E. Nous allons tout d'abord définir deux *lois de composition internes* dans $\mathscr{P}(E)$.

Intersection

Définition 2.6 *Soient A et B deux parties de E. On appelle* ***intersection des ensembles*** A ***et*** B *l'ensemble des éléments qui appartiennent à la fois à A et à B. Celui-ci est noté $A \cap B$. Lorsque $A \cap B = \varnothing$ (c'est-à-dire lorsque A et B n'ont aucun élément commun), on dit que A et B sont* ***disjoints****.*

Réunion

Définition 2.7 *Soient A et B deux parties de E. On appelle* ***réunion des ensembles*** A ***et*** B *l'ensemble des éléments qui appartiennent à A ou à B. Cet ensemble est noté $A \cup B$.*

Soient A, B et C trois sous-ensembles de E. L'intersection et la réunion sont des lois sont *commutatives* :

$$A \cap B = B \cap A \text{ et } A \cup B = B \cup A,$$

et *associatives* :

$$A \cup (B \cup C) = (A \cup B) \cup C,$$
$$A \cap (B \cap C) = (A \cap B) \cap C.$$

Elles sont également *distributives* l'une pour l'autre :

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C),$$
$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$$

La preuve de ces propriétés est basée sur les synonymies introduites dans la sous-section 1.1.2 du chapitre 1. Si l'on définit la proposition $(P(x))$ (resp. $(Q(x))$, resp. $(R(x))$) comme étant vraie si et seulement si $x \in A$ (resp. $x \in B$, resp. $x \in C$), la proposition $x \in B \cap C$ est alors équivalente à ($P(x)$ et $Q(x)$) et $x \in B \cup C$ à ($P(x)$ ou $Q(x)$). Ainsi, $x \in A \cup (B \cap C)$ est équivalente à ($P(x)$ ou ($Q(x)$ et $R(x)$)), ce qui est équivalent à (($P(x)$ ou $Q(x)$) et ($P(x)$ ou $R(x)$)), ou encore à $x \in (A \cup B) \cap (A \cup C)$. On procède de manière identique pour montrer les autres assertions de ce chapitre.

Nous introduisons à présent la notion de *complémentaire* d'un sous-ensemble.

Complémentaire

Définition 2.8 *Soit A un sous-ensemble de E. On appelle* ***complémentaire de*** A ***dans*** E*, et l'on note $C_E(A)$, l'ensemble des éléments de E qui n'appartiennent pas à A.*

Différence

Définition 2.9 *Soient A et B deux parties d'un ensemble E. On appelle* ***différence de*** A ***et de*** B*, et on note $A \backslash B$, l'ensemble des éléments de E appartenant à A mais pas à B.*

Toutes ces opérations sont résumées sur la figure suivante sous la forme de *diagrammes de Venn* [1].

Remarques. Les démonstrations des propriétés suivantes sont laissées en exercice au lecteur. Soient A et B deux parties d'un ensemble E. On a

- $A \cap \varnothing = \varnothing$,
  $A \cap A = A$,
  $A \cap E = A$ (on dit que E est *neutre* pour $\cap$),
  $A \cap B = B \Leftrightarrow A \subset B$.

1. John Venn (4 août 1834 - 4 avril 1923) était un mathématicien et logicien anglais. Il est célèbre pour avoir conçu les diagrammes portant son nom, qui sont notamment employés en théorie des ensembles, en probabilité, en logique, en statistique et en informatique.

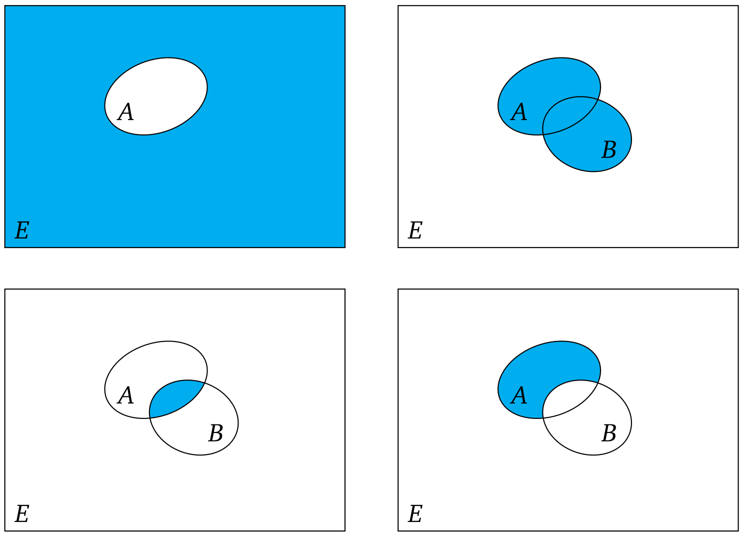

FIGURE 2.1 – Diagrammes de Venn représentant (en bleu) respectivement les ensembles $C_E(A)$, $A \cup B$, $A \cap B$ et $A \backslash B$, où A et B sont des sous-ensembles d'un ensemble E.

- $A \cup \varnothing = A$ (on dit que $\varnothing$ est neutre pour $\cup$),
  $A \cup A = A$,
  $A \cup E = E$,
  $A \cup B = B \Leftrightarrow A \subset B$.
- $A \cap (A \cup B) = A \cup (A \cap B) = A$.
- $C_E(\varnothing) = E$, $C_E(E) = \varnothing$,
  $C_E(C_E(A)) = A$,
  $A \cap C_E(A) = \varnothing$,
  $A \cup C_E(A) = E$.
  On déduit de ces deux dernières égalités que $\{A, C_E(A)\}$ est une partition de E.
- $C_E(A \cap B) = C_E(A) \cup C_E(B)$,
  $C_E(A \cup B) = C_E(A) \cap C_E(B)$ (*lois de De Morgan* [2]).
- $E \backslash A = C_E(A)$,
  $A \backslash B = \varnothing \Leftrightarrow A \subset B$,
  $A \backslash B = A \cap C_E(B) = A \backslash (A \cap B)$.

2.1.5 Produit d'ensembles

Étant donnés deux ensembles E et F, on peut associer à tous éléments $x \in E$ et $y \in F$ le nouvel objet (x, y) appelé *couple ordonné*. Ce couple est un élément d'un nouvel ensemble, que l'on nomme *ensemble produit de E par F*.

Définition 2.10 *Soient E et F deux ensembles. On appelle* ***ensemble produit de*** E ***par*** F *(ou* ***produit cartésien de*** E ***et de*** F*), et l'on note $E \times F$, l'ensemble des couples (x, y), tels que $x \in E$ et $y \in F$:*

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

L'égalité entre couples est définie par l'équivalence logique suivante :

$$(a, b) = (c, d) \iff (a = c \text{ et } b = d).$$

Remarque. Lorsque $E = F$, $E \times F = E^2$. Par extension, étant donnés un entier $n \geq 1$ et des ensembles $E_1, \ldots, E_n$, on appelle produit de $E_1, \ldots, E_n$ l'ensemble de tous les n-uplets $(x_1, \ldots, x_n)$ tels que $x_1 \in E_1, \ldots, x_n \in E_n$. Cet ensemble est noté $E_1 \times \cdots \times E_n$ ou encore $\prod_{i=1}^{n} E_i$. Lorsque $E_1 = \cdots = E_n = E$, l'ensemble produit est noté E^n.

2. Augustus De Morgan (27 juin 1806 - 18 mars 1871) était un mathématicien britannique. Il est considéré comme l'un des fondateurs de la logique moderne.

2.2 Relations

2.2.1 Généralités

Définition 2.11 *Soient E et F deux ensembles. Une **relation** (ou **correspondance**) $\mathscr{R}$ **de** E **vers** F est définie par tout triplet (E,R,F) où R est une partie de $E \times F$.*

Considérons un couple (x,y) de $E \times F$ vérifiant $(x,y) \in R$. On dit que l'élément x de E est *en relation par $\mathscr{R}$* avec l'élément y de F, ce que l'on note $x\mathscr{R}y$. L'ensemble E s'appelle l'*ensemble de départ* de $\mathscr{R}$, F l'*ensemble d'arrivée* de $\mathscr{R}$ et R est le *graphe* de $\mathscr{R}$. Enfin, on appelle *ensemble de définition* de $\mathscr{R}$ la partie A de E définie comme suit :

$$A = \{x \in E \mid \exists y \in F,\ x\mathscr{R}y\},$$

et *ensemble image* de $\mathscr{R}$ la partie B de F définie par :

$$B = \{y \in F \mid \exists x \in E,\ x\mathscr{R}y\}.$$

Définition 2.12 *Soient E, F, G trois ensembles, $\mathscr{R}$, (resp. $\mathscr{S}$) une relation de E vers F (resp. F vers G). On définit la **relation composée de $\mathscr{R}$ et $\mathscr{S}$**, notée $\mathscr{S} \circ \mathscr{R}$, de E vers G par :*

$$\forall (x,z) \in E \times G,\ (x\mathscr{S} \circ \mathscr{R}z \Longleftrightarrow (\exists y \in F,\ x\mathscr{R}y \text{ et } y\mathscr{S}z)).$$

Exemple. Soient $E = F = G$ l'ensemble des droites d'un plan affine euclidien et $\mathscr{R} = \mathscr{S} = \perp$, la relation d'orthogonalité. Alors $\mathscr{S} \circ \mathscr{R} = \parallel$, la relation de parallélisme, car, pour toutes droites D et D'' du plan E, on a :

$$(D \parallel D'') \Longleftrightarrow (\exists D' \in E,\ D \perp D' \text{ et } D' \perp D'')$$

On peut donc écrire ici : $\perp \circ \perp = \parallel$.

Proposition 2.13 *Soient E, F, G, H des ensembles et $\mathscr{R}$ (resp. $\mathscr{S}$, resp. $\mathscr{T}$) une relation de E vers F (resp. F vers G, resp. G vers H). On a alors :*

$$(\mathscr{T} \circ \mathscr{S}) \circ \mathscr{R} = \mathscr{T} \circ (\mathscr{S} \circ \mathscr{R}).$$

DÉMONSTRATION. On remarque tout d'abord que les relations $(\mathscr{T} \circ \mathscr{S}) \circ \mathscr{R}$ et $\mathscr{T} \circ (\mathscr{S} \circ \mathscr{R})$ ont le même ensemble de départ (E) et d'arrivée (H). Soit $(x,t) \in E \times H$. On a :

$$\begin{aligned}
x(\mathscr{T} \circ \mathscr{S}) \circ \mathscr{R}t &\Longleftrightarrow (\exists y \in F,\ (x\mathscr{R}y \text{ et } y\mathscr{T} \circ \mathscr{S}t)) \\
&\Longleftrightarrow (\exists y \in F,\ \exists z \in G,\ (x\mathscr{R}y \text{ et } y\mathscr{S}z \text{ et } z\mathscr{T}t)) \\
&\Longleftrightarrow (\exists z \in G,\ (x\mathscr{S} \circ \mathscr{R}z \text{ et } z\mathscr{T}t)) \\
&\Longleftrightarrow x\mathscr{T} \circ (\mathscr{S} \circ \mathscr{R})t.
\end{aligned}$$

□

Définition 2.14 *Soient E, F deux ensembles et $\mathscr{R}$ une relation de E vers F. On définit la **relation réciproque de $\mathscr{R}$**, notée $\mathscr{R}^{-1}$, de F vers E par :*

$$\forall (x,y) \in E \times F,\ (y\mathscr{R}^{-1}x \Longleftrightarrow x\mathscr{R}y).$$

Proposition 2.15 *1) Pour toute relation $\mathscr{R}$, on a : $(\mathscr{R}^{-1})^{-1} = \mathscr{R}$.*
2) Soient E, F, G trois ensembles et $\mathscr{R}$ (resp. $\mathscr{S}$) une relation de E vers F (resp. F vers G). On a :

$$(\mathscr{S} \circ \mathscr{R})^{-1} = \mathscr{R}^{-1} \circ \mathscr{S}^{-1}.$$

DÉMONSTRATION.
1) C'est immédiat.
2) $\forall (x,z) \in E \times G$, on a :

$$\begin{aligned}
z(\mathscr{S} \circ \mathscr{R})^{-1}x &\Longleftrightarrow x(\mathscr{S} \circ \mathscr{R})z \\
&\Longleftrightarrow (\exists y \in F,\ (x\mathscr{R}y \text{ et } y\mathscr{S}z)) \\
&\Longleftrightarrow (\exists y \in F,\ (z\mathscr{S}^{-1}y \text{ et } y\mathscr{R}^{-1}x)) \\
&\Longleftrightarrow z\mathscr{R}^{-1} \circ \mathscr{S}^{-1}x.
\end{aligned}$$

□

Définition 2.16 *Soit E un ensemble. Une correspondance de E vers E est appelée une **relation binaire dans** E.*

Exemples.

— Soit n un entier naturel. et $\mathscr{R} = \{(p,q) \in \mathbb{Z}^2 \mid \exists k \in \mathbb{Z},\ p-q = kn\}$. La relation ainsi définie est appelée congruence modulo n. Au lieu d'écrire $(p,q) \in \mathscr{R}$ ou encore $p\mathscr{R}q$, on note $p \equiv q \pmod n$.

— La relation d'égalité $a = b$ sur un ensemble E est une relation binaire dont le graphe est la diagonale de E^2, c'est-à-dire l'ensemble des (x,x) quand x parcourt E. Les ensembles de définition et image de cette relation coïncident avec E.

— Soient E un ensemble, A et B deux parties de E. La relation d'inclusion $A \subset B$ est une relation binaire dans $\mathscr{P}(E)$. Les ensembles de définition et image de cette relation coïncident avec $\mathscr{P}(E)$ tout entier.

Définition 2.17 *Une relation binaire $\mathscr{R}$ dans un ensemble E est dite :*
- ***réflexive** si et seulement si : $\forall x \in E,\ x\mathscr{R}x$,*
- ***symétrique** si et seulement si : $\forall (x,y) \in E^2,\ (x\mathscr{R}y \Rightarrow y\mathscr{R}x)$,*
- ***antisymétrique** si et seulement si : $\forall (x,y) \in E^2,\ ((x\mathscr{R}y \text{ et } y\mathscr{R}x) \Rightarrow x = y)$,*
- ***transitive** si et seulement si : $\forall (x,y,z) \in E^3,\ ((x\mathscr{R}y \text{ et } y\mathscr{R}z) \Rightarrow x\mathscr{R}z)$.*

Exemples.

— La relation d'égalité dans un ensemble quelconque est réflexive, symétrique et transitive.

— L'inclusion dans $\mathscr{P}(E)$ est réflexive, non symétrique, antisymétrique et transitive.

— Dans l'ensemble des droites du plan affine euclidien, la relation $\perp$ d'orthogonalité est symétrique, mais n'est ni réflexive, ni antisymétrique, ni transitive. Une droite n'est en effet pas perpendiculaire à elle-même et, d'autre part, $D \perp D'$ et $D' \perp D''$ entraînent $D \parallel D''$ (D parallèle à D'') et non $D \perp D''$.

Définition 2.18 *Soient E un ensemble, $\mathscr{R}$ une relation binaire sur E et A une partie de E. La relation binaire dans A, notée $\mathscr{R}_A$, définie par :*

$$\forall (x,y) \in A^2,\ (x\mathscr{R}_A y \Leftrightarrow x\mathscr{R}y),$$

*est appelée **relation induite par $\mathscr{R}$ sur A**.*

2.2.2 Relations d'équivalence

Définition 2.19 *Soit $\mathscr{R}$ une relation binaire dans un ensemble E. On dit que $\mathscr{R}$ est une **relation d'équivalence** si et seulement si $\mathscr{R}$ est réflexive, symétrique et transitive.*

Étant donnée une relation d'équivalence, on identifie les éléments qui sont en relation en introduisant les *classes d'équivalence*.

Définition 2.20 *Soit $\mathscr{R}$ une relation d'équivalence dans un ensemble E. Pour chaque x de E, on appelle **classe d'équivalence de x (modulo $\mathscr{R}$)** le sous-ensemble de E défini par :*

$$\mathscr{C}(x) = \{y \in E \mid x\mathscr{R}y\}.$$

*Tout élément de $\mathscr{C}(x)$ est appelé un **représentant de la classe** $\mathscr{C}(x)$. L'ensemble des classes d'équivalence modulo $\mathscr{R}$ se nomme **ensemble quotient de** E **par** $\mathscr{R}$ et se note $E/\mathscr{R}$.*

Exemples.

— Comme on l'a déjà vu, la relation d'égalité dans un ensemble E quelconque est une relation d'équivalence, d'ensemble quotient $\{\{x\} \mid x \in E\}$.

— Lorsque la relation d'équivalence est la congruence modulo n, avec n un entier positif, dans $\mathbb{Z}$, l'ensemble quotient est noté $\mathbb{Z}/n\mathbb{Z}$ et on a :

$$\mathbb{Z}/n\mathbb{Z} = \{\mathscr{C}(0), \mathscr{C}(1), \dots, \mathscr{C}(n-1)\},$$

avec $\mathscr{C}(i) = \{i + nk \mid k \in \mathbb{Z}\}$, $0 \le i \le n-1$.

— Dans l'ensemble des droites d'un plan affine, la relation de parallélisme $\parallel$ est une relation d'équivalence. Pour toute droite D, la classe de D modulo $\parallel$ est appelée la direction D.

Théorème 2.21 *À toute relation $\mathscr{R}$ d'équivalence dans un ensemble E correspond une partition de E en classes d'équivalence et réciproquement, toute partition de E définit sur E une relation d'équivalence $\mathscr{R}$, dont les classes coïncident avec les éléments de la partition donnée.*

DÉMONSTRATION. Soit $\mathscr{R}$ une relation d'équivalence dans E. Pour tout élément x de E, $\mathscr{C}(x)$ est non vide car x appartient à $\mathscr{C}(x)$. Soit $(x,y)\in E^2$ tel que $\mathscr{C}(x)\cap\mathscr{C}(y)\neq\varnothing$. Il existe donc $z\in\mathscr{C}(x)\cap\mathscr{C}(y)$. On a alors $x\mathscr{R}z$ et $y\mathscr{R}z$, d'où (par symétrie et transitivité de la relation) $x\mathscr{R}y$. On en déduit que $\mathscr{C}(x)\subset\mathscr{C}(y)$. Soit en effet t de $\mathscr{C}(x)$, on a $x\mathscr{R}t$ et $x\mathscr{R}y$, d'où $y\mathscr{R}t$, soit encore $t\in\mathscr{C}(y)$. Puis, x et y jouant des rôles symétriques, on a $\mathscr{C}(x)=\mathscr{C}(y)$. Comme, pour chaque élément x de E, x appartient à $\mathscr{C}(x)$, la réunion des éléments de $E/\mathscr{R}$ est E.
Réciproquement, soit $\mathscr{P}$ une partition de E et $\mathscr{R}$ la relation définie dans E par :

$$\forall(x,y)\in E^2,\ (x\mathscr{R}y\Longleftrightarrow(\exists P\in\mathscr{P},\ (x\in P \text{ et } y\in P))).$$

Par définition, il existe, pour chaque x de E, un élément P de $\mathscr{P}$ auquel x appartient, on a : $\forall x\in E$, $x\mathscr{R}x$, et donc $\mathscr{R}$ est réflexive.
Pour tout (x,y) de E^2, on a :

$$x\mathscr{R}y\Longleftrightarrow(\exists P\in\mathscr{P},\ (x\in P \text{ et } y\in P))\Longleftrightarrow(\exists P\in\mathscr{P},\ (y\in P \text{ et } x\in P))\Longleftrightarrow y\mathscr{R}x,$$

et donc $\mathscr{R}$ est symétrique.
Soit $(x,y,z)\in E^3$ tel que $x\mathscr{R}y$ et $y\mathscr{R}z$. Il existe P et $Q\in\mathscr{P}$ tels que :

$$((x\in P \text{ et } y\in P) \text{ et } (y\in Q \text{ et } z\in Q)).$$

Comme $P\cap Q\neq\varnothing$ et que $\mathscr{P}$ est une partition, on a $P=Q$, donc $(x\in P$ et $z\in P)$ d'où $x\mathscr{R}z$. Ainsi, $\mathscr{R}$ est transitive.

Enfin, soit x un élément de E. Il existe P de $\mathscr{P}$ tel que x appartient à P et l'on a alors $\mathscr{C}(x)=P$. En effet, pour tout y de P, $(x\in P$ et $y\in P)$ donc $x\mathscr{R}y$.
Pour tout élément y de $\mathscr{C}(x)$, il existe Q appartenant à $\mathscr{P}$ tel que $(x\in Q$ et $y\in Q)$ et $Q=P$ (pour les mêmes raisons que précédemment) et donc $y\in P$. Ceci prouve que $E/\mathscr{R}\subset\mathscr{P}$.
Réciproquement, soit $P\in\mathscr{P}$. Il existe x appartenant à P et l'on a alors $\mathscr{C}(x)=\mathscr{P}$. Ceci montre que $\mathscr{P}\subset E/\mathscr{R}$. □

2.2.3 Relations d'ordre

Généralités

Définition 2.22 *Soit $\mathscr{R}$ une relation binaire dans un ensemble E. On dit que $\mathscr{R}$ est une **relation d'ordre** si et seulement si $\mathscr{R}$ est réflexive, antisymétrique et transitive.*

Une relation d'ordre est souvent notée $\leq$ ou encore $\preceq$. Le couple $(E,\leq)$, où E est un ensemble et $\leq$ une relation d'ordre, est appelé *ensemble ordonné*. La relation $x\leq y$ et $x\neq y$ est notée $x<y$.

Définition 2.23 *Soit $(E,\leq)$ un ensemble ordonné. La relation $\leq$ est dite **relation d'ordre total** si deux éléments quelconques de E sont **comparables** :*

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

*Dans le cas contraire, l'ordre est dit **partiel**.*

Exemples.
- La relation $\leq$ usuelle sur $\mathbb{R}$ est une relation d'ordre total.
- Si E est un ensemble ayant au moins deux éléments, l'inclusion dans $\mathscr{P}(E)$ est une relation d'ordre partiel.

Soient E un ensemble, $\leq$ une relation d'ordre sur E et A une partie de E. La relation induite par $\leq$ dans A est une relation d'ordre appelée *relation d'ordre induite par $\leq$ sur A*.

Éléments remarquables d'un ensemble ordonné

Définition 2.24 *Soient $(E, \leq)$ un ensemble ordonné et A une partie de E.*

1) *Un élément x de E est appelé un **majorant** (resp. **minorant**) de A dans E si et seulement si :*
$$\forall a \in A,\ a \leq x \quad (\text{resp. } \forall a \in A,\ x \leq a).$$
2) *On dit que A est **majorée** (resp. **minorée**) dans E si et seulement si A admet au moins un majorant (resp. minorant) dans E, c'est-à-dire :*
$$\exists x \in E,\ \forall a \in A,\ a \leq x \quad (\text{resp. } \exists x \in E,\ \forall a \in A,\ x \leq a).$$
3) *Un élément x de E est appelé un **plus grand** (resp. **plus petit**) élément de A si et seulement s'il appartient à A et s'il majore (resp. minore) A, c'est-à-dire :*
$$(x \in A \text{ et } (\forall a \in A,\ a \leq x)) \quad (\text{resp. } (x \in A \text{ et } (\forall a \in A,\ x \leq a))).$$
4) *Soit x appartenant à A. On dit que x est un élément **maximal** (resp. **minimal**) de A si et seulement si :*
$$\forall a \in A,\ (x \leq a \Rightarrow x = a) \quad (\text{resp. } \forall a \in A,\ (a \leq x \Rightarrow x = a)).$$

Exemples.
- Dans $(\mathbb{R}, \leq)$, $\mathbb{R}_+$ n'a pas d'élément maximal.
- Dans $(\mathbb{R}, \leq)$, $[0, 1]$ admet un élément maximal et un seul, qui est 1. C'est aussi le plus grand élément de $[0, 1]$.

Définition 2.25 *Soient $(E, \leq)$ un ensemble ordonné et A une partie de E. On dit qu'un élément M de E est la **borne supérieure de** A **dans** E, notée $\sup A$, si l'ensemble des majorants de A dans E admet M comme plus petit élément. Un élément m de E sera appelé la **borne inférieure de** A **dans** E, notée $\inf A$, si l'ensemble des minorants de A dans E admet m comme plus grand élément.*

Exemples.
- Dans $\mathbb{R}$ muni de l'ordre usuel, $\mathbb{N}$ ne possède pas de borne supérieure, $[0, 1[$ possède une borne supérieure égale à 1 et une borne inférieure égale à 0.
- Dans $\mathbb{Q}$ muni de l'ordre usuel, on considère $A = \{x \in \mathbb{Q}_+ \mid x^2 < 2\}$. L'ensemble des majorants de A dans $\mathbb{Q}$ est $\{x \in \mathbb{Q}_+\ \ x^2 \geq 2\} = \{x \in \mathbb{Q}_+\ \ x^2 > 2\}$, qui n'a pas de plus petit élément (cela revient à $\sqrt{2} \notin \mathbb{Q}$) ; A n'admet donc pas de borne supérieure dans $\mathbb{Q}$.

2.3 Applications

Nous allons à présent étudier des relations particulières nommées *applications*.

2.3.1 Définitions

Définitions 2.26 *Soient deux ensembles E et F. On appelle **fonction** de E dans F une relation qui à x de E associe au plus un élément y de F. L'ensemble des éléments de E auxquels f associe exactement un élément dans F est appelé l'**ensemble** (ou **domaine**) **de définition de** f.*

On dit que y est l'*image* de x par f, ce que l'on note $f(x)$, tandis que x est un *antécédent* de y par f. On dit aussi que E (resp. F) est l'*ensemble de départ* (resp. l'*ensemble d'arrivée*) de f. Le *graphe* de la fonction est l'ensemble des couples $(x, f(x))$ lorsque x parcourt E.
Notons que tout élément de F n'est pas nécessairement l'image d'un élément de E.

Définitions 2.27 *Une fonction de E dans F est une **application** si et seulement si son domaine de définition est égal à E.*

Une application de E dans F est souvent notée $f : E \to F$. La définition de l'égalité de deux ensembles implique que deux applications $f, g : E \to F$ sont égales si, pour chaque $x \in E$, on a $f(x) = g(x)$.

Exemples d'applications.

- Soient E, F deux ensembles et $a \in F$. Si tout élément x de E a pour image $f(x) = a$, l'application est dite *constante* et égale à a.
- Soit E un ensemble. L'application qui à chaque élément x de E associe lui-même est appelée *application identique* ou *identité de E*. On la note I_E.
- Soient E un ensemble et A une partie de E. On appelle *fonction caractéristique de A dans E* ou *fonction indicatrice de A dans E* la fonction de E dans $\{0,1\}$ telle que $f(x) = 1$ si $x \in A$ et $f(x) = 0$ si $x \in C_E(A)$.
- La fonction qui à chaque nombre réel x associe 0 est appelée *fonction nulle* et on la note 0.
- Soient n un entier positif, a_0, a_1,..., a_n des réels, la fonction $p : \mathbb{R} \to \mathbb{R}$ qui à chaque réel x, associe :
$$p(x) = a_0 + a_1 x + \cdots + a_n x^n$$
est appelée une *fonction polynôme*.
- Soient p et q deux fonctions polynômes avec $q \neq 0$. Dans ce cas, l'ensemble des racines du polynôme q :
$$\{x \in \mathbb{R} \mid q(x) = 0\}$$
est ou bien égal à l'ensemble vide $\emptyset$, ou bien il existe un entier $k \geq 1$ et $x_1 \in \mathbb{R}$,..., $x_k \in \mathbb{R}$ tels que $\{x \in \mathbb{R} \mid q(x) = 0\} = \{x_1, \ldots, x_k\}$.
La fonction qui à $x \in \mathbb{R} \backslash \{x_1, \ldots, x_k\}$ associe $\dfrac{p(x)}{q(x)}$ est une fonction rationnelle.
- Les fonctions *exponentielle*, *cosinus* et *sinus*.
- La fonction *logarithme népérien* ln, définie sur $]0, +\infty[$.
- La fonction *tangente*, notée tan et définie sur $\mathbb{R} \backslash \left\{(2k+1)\dfrac{\pi}{2},\ k \in \mathbb{Z}\right\}$ par $\tan x = \dfrac{\sin x}{\cos x}$.

Définition 2.28 *Une partie A d'un ensemble E est dite **stable par** une application $f : E \to E$ si et seulement si : $\forall a \in A,\ f(a) \in A$.*

2.3.2 Surjection. Injection. Bijection

Soient E et F deux ensembles et f une application de E dans F.

Définition 2.29 *L'application f est dite **surjective** (on dit encore que f est une **surjection**) si et seulement si :*
$$\forall y \in F,\ \exists x \in E,\ y = f(x),$$
c'est-à-dire si tout élément de F est l'image par f d'au moins un élément de E.

Définition 2.30 *L'application f est dite **injective** (ou f est une **injection**) si et seulement si :*
$$f(x) = f(x') \;\Rightarrow\; x = x',\ \forall x, x' \in E,$$
c'est-à-dire si deux éléments distincts de E ont des images distinctes.

Définition 2.31 *L'application f est dite **bijective** (ou f est une **bijection**) si et seulement si elle est à la fois surjective et injective.*

Toute bijection de E dans E s'appelle une *permutation*.

Proposition 2.32 *L'application f est bijective si et seulement si tout élément y de F possède un unique antécédent x par f dans E :*
$$\forall y \in F,\ \exists! x \in E,\ f(x) = y.$$

DÉMONSTRATION. Si f est bijective, alors f est surjective. Par conséquent, tout élément y appartenant à F admet au moins un antécédent x par f dans E. Supposons maintenant que y ait deux antécédents x_1 et x_2. On a alors $y = f(x_1) = f(x_2)$, d'où $x_1 = x_2$ puisque f est injective. On en déduit que y admet un seul antécédent.
Réciproquement, si tout y de F admet un unique antécédent x par f dans E, alors f est surjective de E dans F. Soient x_1, x_2 de E tels que $f(x_1) = f(x_2)$. Posons $y = f(x_1) = f(x_2)$, alors x_1 et x_2 sont deux antécédents de y. Par unicité de l'antécédent, on a $x_1 = x_2$, ce qui prouve l'injectivité de f. L'application f est donc bijective de E dans F. □

Lorsqu'une application est bijective, il est possible d'introduire la notion d'*application réciproque*.

Définition 2.33 *Soit $f : E \to F$ une application bijective de E dans F. On définit alors une application de F vers E en associant à tout élément y de F son seul antécédent. Cette application, appelée* ***application réciproque de** f et notée f^{-1}, vérifie donc :*

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

L'application f^{-1} est également bijective et $(f^{-1})^{-1} = f$.

Exemples d'applications réciproques.

— Si E est un ensemble, I_E est bijective et $I_E^{-1} = I_E$.
— La fonction qui à $x \in \mathbb{R}$ associe $e^x \in]0, +\infty[$ est bijective et son application réciproque est la fonction ln.
— La fonction qui à $x \in [0, \pi]$ associe $\cos x \in [-1, 1]$ est bijective. Sa fonction réciproque est notée arccos. Pour chaque $y \in [-1, 1]$, $\arccos y$ est l'unique réel x appartenant à $[0, \pi]$ tel que $\cos x = y$.
— La fonction qui à $x \in [-\pi/2, \pi/2]$ associe $\sin x \in [-1, 1]$ est bijective. Sa fonction réciproque est notée arcsin. Pour chaque $y \in [-1, 1]$, $\arcsin y$ est l'unique réel x appartenant à $[-\pi/2, \pi/2]$ tel que $\sin x = y$.
— La fonction qui à $x \in]-\pi/2, \pi/2[$ associe $\tan x \in \mathbb{R}$ est bijective. Sa fonction réciproque est notée arctan. Pour chaque $y \in \mathbb{R}$, $\arctan y$ est l'unique réel x appartenant à $]-\pi/2, \pi/2[$ tel que $\tan x = y$.

2.3.3 Restrictions et prolongements

Définition 2.34 *Soient E, F deux ensembles, $f : E \to F$ une application et A une partie de E. On appelle* ***restriction de f à** A l'application, notée $f_{|_A}$, définie par :*

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

Définition 2.35 *Soient E et F deux ensembles, $f : E \to F$ une application et G un ensemble tel que $E \subset G$. On appelle* ***prolongement de f à** G toute application $\widetilde{f} : G \to F$ telle que :*

$$\forall x \in E,\ \widetilde{f}(x) = f(x).$$

2.3.4 Composition des applications

Définition 2.36 *Soient E, F, G trois ensembles, $f : E \to F$ et $g : F \to G$ deux applications. L'application $g \circ f$ de E dans G définie par $g \circ f(x) = g(f(x))$ est appelée* ***composée de g et de f***.

Pour pouvoir définir $g \circ f$, il est nécessaire que l'ensemble de départ de g soit égal à l'ensemble d'arrivée de f. L'ordre de composition est également important. Même dans le cas où l'on peut composer dans les deux sens, on a en général $g \circ f \neq f \circ g$.

Proposition 2.37 *Soient E, F, G et H quatre ensembles. Pour toutes applications $f : E \to F$, $g : F \to G$ et $h : G \to H$, on a :*

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

On se réfèrera à la preuve de la proposition 2.13 pour une démonstration, cette proposition en étant un cas particulier.

Proposition 2.38 *La composée de deux injections (resp. surjections, resp. bijections) est une injection (resp. surjection, resp. bijection).*

DÉMONSTRATION. Soient $f : E \to F$ et $g : F \to G$ deux applications.

1) On suppose f et g injectives. On a, pour tout $(x, x') \in E^2$:

$$(g \circ f)(x) = (g \circ f)(x') \Leftrightarrow g(f(x)) = g(f(x')) \Rightarrow f(x) = f(x') \Rightarrow x = x',$$

d'où $g \circ f$ est injective.

2) On suppose f et g surjectives. Soit $z \in G$. Puisque g est surjective, il existe $y \in F$ tel que $z = g(y)$. Puis, comme f est surjective, il existe $x \in E$ tel que $y = f(x)$. On a donc $z = g(f(x)) = (g \circ f)(x)$, ce qui montre que $g \circ f$ est surjective.

3)

$$(f \text{ et } g \text{ bijectives}) \Rightarrow \left\{ \begin{array}{l} f \text{ et } g \text{ injectives} \\ f \text{ et } g \text{ surjectives} \end{array} \right. \Rightarrow \left\{ \begin{array}{l} g \circ f \text{ injective} \\ g \circ f \text{ surjective} \end{array} \right. \Rightarrow (g \circ f \text{ bijective}).$$

□

Remarque. Si $f : E \to F$ est une application bijective, on a déjà vu que son application réciproque f^{-1} est aussi bijective. On a également $f \circ f^{-1} = I_F$ et $f^{-1} \circ f = I_E$.

Proposition 2.39 *Soient E, F, G trois ensembles, $f : E \to F$ et $g : F \to G$ deux applications.*

1) Si $g \circ f$ est injective, alors f est injective.

2) Si $g \circ f$ est surjective, alors g est surjective.

DÉMONSTRATION.

1) On suppose $g \circ f$ injective. On a, pour tout $(x, x') \in E^2$:

$$f(x) = f(x') \Rightarrow g(f(x)) = g(f(x')) \Leftrightarrow (g \circ f)(x) = (g \circ f)(x') \Rightarrow x = x',$$

ce qui montre que f est injective.

2) On suppose $g \circ f$ surjective. Soit $z \in G$. Il existe $x \in E$ tel que $z = (g \circ f)(x) = g(f(x))$. Ceci montre que g est surjective.

□

Proposition 2.40 *Soient E, F, G trois ensembles, $f : E \to F$ et $g : F \to G$ deux applications bijectives. On a alors $g \circ f : E \to G$ bijective et $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.*

Cette propriété résulte d'une part de la proposition 2.38 et d'autre part de la proposition 2.15.

Définition 2.41 *Soit E un ensemble. On appelle* ***involution de** E toute application de E dans E telle que $f \circ f = I_E$.*

2.3.5 Images directes ou réciproques de parties par une application

Définition 2.42 *Soient E et F deux ensembles, f une application de E dans F, $A \subset E$ et $B \subset F$. L'**image directe de** A **par** f (ou, plus simplement, **image de** A **par** f), notée $f(A)$, est le sous-ensemble de F contenant l'image des éléments de A par f :*

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

*L'**image réciproque de** B **par** f, notée $f^{-1}(B)$, est le sous-ensemble de E contenant les antécédents des éléments de B par f :*

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

Remarque. Si f est une application d'un ensemble E dans un ensemble F, il est toujours possible de définir $f^{-1}(B)$, même si f n'est pas bijective. Si c'est cependant le cas, et donc si f^{-1} existe, on pourra vérifier que l'image directe d'une partie B de F par f^{-1} est aussi l'image réciproque $f^{-1}(B)$ de B par f. En effet, $x \in f^{-1}(B)$ signifie que $f(x) \in B$ et réciproquement, si l'on pose $y = f(x)$, on a $x = f^{-1}(y)$ avec $y \in B$, ce qui équivaut à $x \in f^{-1}(B)$.

La proposition qui suit nous sera utile en pratique pour déterminer si une application est surjective ou non.

Proposition 2.43 *Soit f une application de E dans F. Elle est surjective si et seulement si $f(E) = F$.*

DÉMONSTRATION. On a toujours $f(E) \subset F$. D'autre part, $F \subset f(E)$ si et seulement si tout élément de F est l'image d'au moins un élément de E par f, ce qui signifie que l'application f est surjective de E dans F. □

2.3.6 Familles

Soit I un ensemble quelconque d'éléments, qui seront nommés *indices*.

Définition 2.44 *Soit E un ensemble. On appelle* ***famille d'éléments de** E toute application de I vers E.*

La famille $(x_i)_{i \in I}$ est dite *finie* ou *infinie*, selon que l'ensemble I est fini ou infini (voir section 2.4). Si I est un intervalle d'entiers $[m, n]$, la famille $(x_i)_{i \in I}$ est notée $(x_i)_{m \leq i \leq n}$.
On veillera à ne pas confondre la famille $(x_i)_{i \in I}$ et l'ensemble $\{\{x_i\} \mid i \in I\}$, qui est l'ensemble image de l'application en question.

Définition 2.45 *Soient E un ensemble, $(A_i)_{i \in I}$ une famille de parties de E. On définit :*

- *la* ***réunion de la famille*** $(A_i)_{i\in I}$, *notée* $\bigcup_{i\in I} A_i$, *par :*

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

- *l'****intersection de la famille*** $(A_i)_{i\in I}$, *notée* $\bigcap_{i\in I} A_i$, *par :*

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

Définition 2.46 *Soit E un ensemble. Une famille $(A_i)_{i\in I}$ de parties de E est appelée* ***partition de*** *E si et seulement si :*

(i) aucun des ensembles A_i n'est vide : $\forall \in I$, $A_i \neq \varnothing$,

(ii) les ensembles A_i sont disjoints deux à deux : $\forall (i,j) \in I^2$, $(i \neq j \Rightarrow A_i \cap A_j = \varnothing)$,

(iii) la réunion des ensembles A_i est égale à E : $\bigcup_{i\in I} A_i = E$.

Dans ces conditions, tout élément de l'ensemble E appartient à un sous-ensemble E_i unique. Il est à noter que cette définition est cohérente avec la définition 2.5, car pour que $(A_i)_{i\in I}$ soit une partition au sens ci-dessus, il faut, et il suffit, que $\{A_i \mid i \in I\}$ soit une partition de E au sens de la définition 2.5.

2.4 Ensembles finis et infinis

Le but de cette section est principalement de rappeler des propriétés élémentaires des ensembles finis, souvent considérées comme intuitivement évidentes.

2.4.1 Équipotence

Définition 2.47 *On dit qu'un ensemble E est* ***équipotent*** *à un ensemble F si et seulement s'il existe une bijection de E sur F.*

Exemples.
— $\{1,2,3\}$ est équipotent à $\{2,4,6\}$.
— $\mathbb{N}^*$ est équipotent à $\mathbb{N}$ car l'application $n \mapsto n-1$ de $\mathbb{N}^*$ dans $\mathbb{N}$ est une bijection.

La relation d'équipotence est une relation d'équivalence entre ensembles.

Définition 2.48 *On dit qu'un ensemble est* ***dénombrable*** *si et seulement s'il est équipotent à $\mathbb{N}$.*

2.4.2 Ensembles finis

Définition 2.49 *On dit qu'un ensemble E est* ***fini*** *si et seulement s'il existe un entier n tel que E est équipotent à $\{1,\dots,n\}$.*

Nous admettrons la proposition suivante, dont la preuve s'appuie sur les propriétés de l'ensemble des entiers naturels.

Proposition 2.50 *Soit $(n,p) \in \mathbb{N}^2$.*

1) Il existe une injection de $\{1,\dots,n\}$ dans $\{1,\dots,p\}$ si et seulement si $n \leq p$.

2) Il existe une surjection de $\{1,\dots,n\}$ sur $\{1,\dots,p\}$ si et seulement si $n \leq p$.

3) Il existe une bijection de $\{1,\dots,n\}$ sur $\{1,\dots,p\}$ si et seulement si $n = p$.

La troisième affirmation de la proposition ci-dessus amène la définition suivante.

Définition 2.51 *Soit E un ensemble fini. Il existe alors un entier n tel que E soit équipotent à $\{1,\dots,n\}$; n s'appelle le* ***cardinal*** *de E et est noté* $\operatorname{card} E$.

On convient de prendre 0 pour cardinal de l'ensemble vide.

Proposition 2.52 *Si E est un ensemble fini, toute partie F de E est finie, et l'on a :* $\operatorname{card} F \leq \operatorname{card} E$.

Proposition 2.53 *Si E et F sont deux ensembles finis, alors $E \cup F$ est fini et l'on a :*

$$\operatorname{card}(E \cup F) + \operatorname{card}(E \cap F) = \operatorname{card} E + \operatorname{card} F.$$

DÉMONSTRATION. Établissons tout d'abord un résultat préliminaire. Soient A et B deux ensembles finis disjoints ; notons $a = \operatorname{card} A$, $b = \operatorname{card} B$. Il existe des bijections $\alpha : \{1, \ldots, a\} \to A$ et $\beta : \{1, \ldots, b\} \to B$. Il est clair que l'application $\gamma : \{1, \ldots, a+b\} \to A \cup B$ définie par :

$$\forall n \in \{1, \ldots, a+b\},\ \gamma(n) = \begin{cases} \alpha(n) & \text{si } 1 \leq n < a \\ \beta(n-a) & \text{si } a+1 \leq n \leq a+b \end{cases}$$

est une bijection. Il en résulte que $A \cup B$ est fini et que $\operatorname{card}(A \cup B) = \operatorname{card} A + \operatorname{card} B$.
En appliquant ce résultat à E et $E \backslash F$ au lieu de A et B, on conclut $E \cup F$ est fini et :

$$\begin{aligned} \operatorname{card}(E \cup F) + \operatorname{card}(E \cap F) &= \operatorname{card}(E \cup (F \backslash E)) + \operatorname{card}(E \cap F) \\ &= (\operatorname{card} E + \operatorname{card}(F \backslash E)) + \operatorname{card}(E \cap F) \\ &= \operatorname{card} E + (\operatorname{card}(F \backslash E) + \operatorname{card}(E \cap F)) \\ &= \operatorname{card} E + \operatorname{card} F. \end{aligned}$$

□

Corollaire 2.54 *Soient E un ensemble fini et F une partie de E. Si* $\operatorname{card} F = \operatorname{card} E$ *alors $F = E$.*

Cette dernière propriété est très importante. Nous en verrons une analogue portant sur des dimensions d'espaces vectoriels en algèbre linéaire.
DÉMONSTRATION. Si $\operatorname{card} F = \operatorname{card} E$, comme $\operatorname{card} E = \operatorname{card} F + \operatorname{card}(E \backslash F)$, on en déduit $\operatorname{card}(E \backslash F) = 0$, d'où $E \backslash F = \varnothing$, $E = F$. □

Proposition 2.55 *Soient E et F des ensembles finis de même cardinal et f une application de E dans F. Les assertions suivantes sont deux à deux équivalentes.*

i) f est injective.

ii) f est surjective.

iii) f est bijective.

DÉMONSTRATION. Montrons tout d'abord que i) implique ii) et iii). Si f est injective, $\widetilde{f} : E \to f(E)$ qui à x associe $f(x)$ est bijective, donc $\operatorname{card} f(E) = \operatorname{card} E = \operatorname{card} F$, d'où $f(E) = F$ d'après le corollaire 2.54, d'où f est surjective et donc bijective.
Prouvons à présent que ii) implique i) et iii). Supposons l'application f surjective et non injective. Il existe $(x_1, x_2) \in E^2$ tel que $x_1 \neq x_2$ et $f(x_1) \neq f(x_2)$. L'application $g : E \backslash \{x_2\} \to F$ qui à x associe $f(x)$ est surjective, d'où $\operatorname{card}(E \backslash \{x_2\}) \leq \operatorname{card} F$. Mais on a $\operatorname{card}(E \backslash \{x_2\}) = \operatorname{card} E - 1$ et $\operatorname{card} E = \operatorname{card} F$, d'où une contradiction.
Enfin, iii) implique i) et ii) de manière triviale. □

2.4.3 Ensembles infinis

Définition 2.56 *Un ensemble est dit **infini** si et seulement s'il n'est pas fini.*

Exemple. L'ensemble $\mathbb{N}$ est infini.

Chapitre 3

Espaces vectoriels

Dans ce chapitre, nous allons étudier une structure algébrique omniprésente en mathématiques : la structure d'*espace vectoriel*. Nous allons procéder de manière abstraite et axiomatique ; il s'agira d'ensembles munis d'une loi interne et d'une loi externe qui vérifient certaines règles de calcul. Dans ce chapitre, $\mathbb{K}$ désigne un *corps commutatif*. En pratique, $\mathbb{K}$ est égal à $\mathbb{R}$ ou $\mathbb{C}$.

3.1 Généralités

3.1.1 Loi de composition interne

Définition 3.1 *On considère un ensemble E. On appelle* **loi de composition interne** *sur E toute application définie sur $E \times E$ à valeurs dans E.*

Exemples.

— Dans $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$ ou $\mathbb{C}$, l'addition et la multiplication sont des lois de composition internes.

— L'application

$$\begin{array}{ccc} \mathbb{R}^3 \times \mathbb{R}^3 & \to & \mathbb{R}^3 \\ ((x,y,z),(x',y',z')) & \mapsto & (x+x', y+y', z+z'), \end{array}$$

est une loi de composition interne sur $\mathbb{R}^3$.

— Soit $\mathscr{A}(\mathbb{R},\mathbb{R})$ l'ensemble des applications à valeurs réelles définies sur $\mathbb{R}$. On considère l'application qui à tout couple (f,g) de $\mathscr{A}(\mathbb{R},\mathbb{R})^2$ associe l'application $f+g$ définie par

$$\forall x \in I,\ (f+g)(x) = f(x) + g(x).$$

Cette application est une loi de composition interne sur $\mathscr{A}(\mathbb{R},\mathbb{R})$.

— On considère l'application qui à tout couple (f,g) de $\mathscr{A}(\mathbb{R},\mathbb{R})^2$ associe l'application $g \circ f$ définie par

$$\forall x \in \mathbb{R},\ (g \circ f)(x) = g(f(x)).$$

Cette application est une loi de composition interne sur $\mathscr{A}(\mathbb{R},\mathbb{R})$.

3.1.2 Loi de composition externe

Définition 3.2 *Soit E un ensemble. On appelle* **loi de composition externe** *sur E toute application définie sur $\mathbb{K} \times E$ à valeurs dans E.*

Exemples.

— L'application

$$\begin{array}{ccc} \mathbb{R} \times \mathbb{R}^3 & \to & \mathbb{R}^3 \\ (\lambda,(x,y,z)) & \mapsto & (\lambda x, \lambda y, \lambda z), \end{array}$$

est une loi de composition externe sur $\mathbb{R}^3$.

— L'application qui à tout élément (λ, f) de $\mathbb{R} \times \mathscr{A}(\mathbb{R},\mathbb{R})$ associe l'application notée $\lambda \cdot f$, définie par

$$\forall x \in I,\ (\lambda \cdot f)(x) = \lambda f(x),$$

est une loi de composition externe sur $\mathscr{A}(\mathbb{R},\mathbb{R})$.

— On considère $\mathscr{F}(\mathbb{N},\mathbb{R})$ l'ensemble des suites réelles définies sur $\mathbb{N}$. L'application qui à tout élément $(\lambda, (u_n)_{n\in\mathbb{N}})$ de $\mathbb{R} \times \mathscr{F}(\mathbb{N},\mathbb{R})$ associe la suite $(w_n)_{n\in\mathbb{N}}$, définie par

$$\forall n \in \mathbb{N},\ w_n = \lambda u_n,$$

est une loi de composition externe sur $\mathscr{F}(\mathbb{N},\mathbb{R})$.

3.1.3 Structure de groupe

Définition 3.3 *On considère un ensemble E muni d'une loi de composition interne $*$. On dit que $(E,*)$ est un* ***groupe*** *si les propriétés suivantes sont vérifiées :*

1. *la loi $*$ est* ***associative*** *:*
$$\forall (x,y,z) \in E^3,\ (x*y)*z = x*(y*z),$$
2. *il existe un élément de E, noté e et appelé* ***élément neutre*** *pour la loi $*$, vérifiant*
$$\forall x \in E,\ x*e = e*x = x,$$
3. *tout élément de E admet un* ***symétrique*** *:*
$$\forall x \in E,\ \exists y \in E,\ x*y = y*x = e.$$

Si de plus la loi $*$ est ***commutative*** *:*
$$\forall (x,y) \in E^2,\ x*y = y*x,$$
le groupe $(E,*)$ est dit ***commutatif*** *ou* ***abélien****.*

Remarque. Dans un groupe $(E,*)$, l'élément neutre et le symétrique de tout élément de E sont uniques.

Exemples.
— Pour tout entier $n \geq 1$, $(\mathbb{R}^n,+)$ est un groupe commutatif.
— $(\mathscr{A}(\mathbb{R},\mathbb{R}),+)$ est un groupe commutatif.
— $(\mathbb{R}^*,\times)$ est un groupe commutatif.

Lorsque la loi de composition interne est l'addition, l'élément neutre est noté 0_E (ou simplement 0 s'il n'y a pas de confusion possible) et le symétrique d'un élément x est appelé *opposé* et noté $-x$. Lorsque la loi interne est la multiplication, l'élément neutre est noté 1_E (ou simplement 1 s'il n'y a pas de confusion possible), le symétrique d'un élément x étant appelé son *inverse* et noté x^{-1} ou $\frac{1}{x}$.

3.2 Structure d'espace vectoriel

Soient E un ensemble quelconque et $\mathbb{K}$ un corps commutatif. Nous supposons que E est muni d'une une loi de composition interne, notée additivement :

$$E \times E \to E\ :\ (x,y) \mapsto x+y,$$

et d'une loi de composition externe sur $\mathbb{K}$, notée multiplicativement :

$$\mathbb{K} \times E \to E\ :\ (\lambda,y) \mapsto \lambda x.$$

On a alors la définition suivante.

Définition 3.4 *On dit que $(E,+,\cdot)$ est un* ***espace vectoriel sur*** $\mathbb{K}$ *si*

1. $(E,+)$ *est un groupe abélien,*
2. $\forall(\lambda,\mu)\in\mathbb{K}^2$ *et* $\forall x\in E$, $(\lambda+\mu)x=(\lambda x)+(\mu x)$,
3. $\forall(\lambda,\mu)\in\mathbb{K}^2$ *et* $\forall x\in E$, $\lambda(\mu x)=(\lambda\mu)x$,
4. $\forall\lambda\in\mathbb{K}$ *et* $\forall(x,y)\in E^2$, $\lambda(x+y)=\lambda x+\lambda y$,
5. $\forall x\in E$, $1_{\mathbb{K}}x=x$,

$1_{\mathbb{K}}$ *étant l'élément unitaire du corps* $\mathbb{K}$.

Souvent (et ce sera le cas dans ce cours), la lettre E désignera aussi bien l'espace vectoriel $(E,+,\cdot)$ que l'ensemble sous-jacent. On dit aussi que E est un $\mathbb{K}$-*espace vectoriel*, ou encore un espace vectoriel *réel* (resp. *complexe*) lorsque $\mathbb{K}=\mathbb{R}$ (resp. $\mathbb{K}=\mathbb{C}$). Les éléments de E sont les *vecteurs* (0_E est le *vecteur nul*), ceux de $\mathbb{K}$ les *scalaires*. Si $x_1,\ldots,$ x_n ($n\geq 1$) sont des vecteurs de E et $\lambda_1,\ldots,\lambda_n$ des scalaires, l'élément $\lambda_1x_1+\cdots+\lambda_nx_n$ est un vecteur appelé *combinaison linéaire de la famille* $(x_i)_{1\leq i\leq n}$. On définit plus généralement les combinaisons linéaires d'une famille de vecteurs $(x_i)_{i\in I}$ en imposant que la famille de scalaires $(\lambda_i)_{i\in I}$ ne comporte qu'un nombre fini de termes non nuls.

Exemples.

- Le corps $\mathbb{K}$, muni de l'addition pour loi interne et de la multiplication pour loi externe, est un $\mathbb{K}$-espace vectoriel. Les éléments de $\mathbb{K}$ sont alors simultanément considérés comme des vecteurs et des scalaires. En particulier, $\mathbb{R}$ (resp. $\mathbb{C}$) est un espace vectoriel sur $\mathbb{R}$ (resp. $\mathbb{C}$).
- Le plan vectoriel usuel est un espace vectoriel réel.
- L'ensemble $\mathbb{K}[X]$ des polynômes à coefficients dans $\mathbb{K}$, muni de l'addition usuelle et de la multiplication externe $\lambda P=\lambda a_0+\cdots+\lambda a_nX^n$, pour $\lambda\in\mathbb{K}$ et $P=a_0+\cdots+a_nX^n\in\mathbb{K}[X]$, est un espace vectoriel sur $\mathbb{K}$. Tout polynôme est une combinaison linéaire de la famille $\left(X^i\right)_{i\in\mathbb{N}}$.
- Si $E=\{0_E\}$, on parle d'espace vectoriel nul.

Rappelons quelques règles de calcul s'appliquant dans un espace vectoriel.

Proposition 3.5 *Soit E un espace vectoriel sur $\mathbb{K}$. On a, pour tous λ, μ de $\mathbb{K}$ et pour tous x, y de E,*

1. $\lambda x=0_E\Longleftrightarrow(\lambda=0_{\mathbb{K}}$ *ou* $x=0_E)$,
2. $(\lambda-\mu)x=\lambda x-\mu x$,
3. $\lambda(x-y)=\lambda x-\lambda y$.

DÉMONSTRATION.

1. $0_{\mathbb{K}}x=(0_{\mathbb{K}}+0_{\mathbb{K}})x=0_{\mathbb{K}}x+0_{\mathbb{K}}x$ d'où $0_{\mathbb{K}}x=0_E$ et $\lambda 0_E=\lambda(0_E+0_E)=\lambda 0_E+\lambda 0_E$ d'où $\lambda 0_E=0_E$. Si $\lambda x=0_E$ et si $\lambda\neq 0_{\mathbb{K}}$, alors, en notant λ^{-1} l'inverse de λ dans le corps $\mathbb{K}$,
$$x=1_{\mathbb{K}}x=(\lambda^{-1}\lambda)x=\lambda^{-1}(\lambda x)=\lambda^{-1}0_E=0_E.$$
2. $\lambda x=((\lambda-\mu)+\mu)x=(\lambda-\mu)x+\mu x$ d'où $(\lambda-\mu)x=(\lambda x)-(\mu x)$ qui est noté $\lambda x-\lambda y$.
3. $\lambda x=\lambda((x-y)+y)=\lambda(x-y)-\lambda y$ d'où $\lambda(x-y)=\lambda x-\lambda y$.

□

La démonstration du résultat suivant est laissée en exercice.

Proposition 3.6 *Soient E_1 et E_2 deux espaces vectoriels sur $\mathbb{K}$. L'ensemble produit $E=E_1\times E_2$, muni des lois suivantes : pour tous (x_1,x_2), (y_1,y_2) de E, pour tout λ de $\mathbb{K}$, $(x_1,x_2)+(y_1,y_2)=(x_1+y_1,x_2+y_2)$ et $\lambda(x_1,x_2)=(\lambda x_1,\lambda x_2)$, est un espace vectoriel sur $\mathbb{K}$. Le vecteur nul de E est $0_E=(0_{E_1},0_{E_2})$ et l'opposé d'un vecteur $x=(x_1,x_2)$ est le vecteur $-x=(-x_1,-x_2)$.*

Cette propriété se généralise au cas d'un produit fini d'espaces vectoriels $E_1\times\cdots\times E_n$. Si tous les espaces vectoriels sont égaux au même espace vectoriel E, on note plutôt E^n.

Exemples.

— Pour tout entier $n \geq 1$, $\mathbb{K}^n$, muni des lois usuelles données par

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n) \text{ et } \lambda(x_1, \dots, x_n) = (\lambda x_1, \dots, \lambda x_n),$$

est un espace vectoriel sur $\mathbb{K}$. En particulier, $\mathbb{R}^n$ (resp. $\mathbb{C}^n$) est un espace vectoriel réel (resp. complexe).

— On a déjà vu que $\mathbb{C}$ muni des lois usuelles est un espace vectoriel sur $\mathbb{C}$; la proposition précédente permet de munir $\mathbb{C} = \mathbb{R}^2$ d'une structure d'espace vectoriel sur $\mathbb{R}$.

3.3 Sous-espaces vectoriels

Soit E un espace vectoriel sur $\mathbb{K}$.

Définition 3.7 *On appelle **sous-espace vectoriel de** E toute partie non vide F de E telle que*

1. $0_E \in F$,
2. $\forall (x, y) \in F^2$ *et* $(\lambda, \mu) \in \mathbb{K}^2$, $\lambda x + \mu y \in F$.

On voit qu'un sous-espace vectoriel contient toutes les combinaisons linéaires de ses éléments ; on dit qu'il est *stable par combinaison linéaire*.
Dans la pratique, on pourra remplacer la condition *2.* par la condition équivalente

$$\forall (x, y) \in F^2 \text{ et } \lambda \in \mathbb{K}, \ \lambda x \in F \text{ et } x + y \in F.$$

En d'autres mots, une partie F non vide d'un $\mathbb{K}$-espace vectoriel E est dite sous-espace de E si et seulement si les conditions suivantes sont réalisées :

1. $(F, +)$ est un sous-groupe de $(E, +)$,
2. la restriction à $\mathbb{K} \times F$ de la loi externe définie sur $\mathbb{K} \times E$ est une application dans F.

Exemples.

— E et $\{0_E\}$ sont des sous-espaces vectoriels de E.

— $\mathbb{K} \times \{0_\mathbb{K}\} = \{(x, 0_\mathbb{K}) \mid x \in \mathbb{K}\}$ et $\{0_\mathbb{K}\} \times \mathbb{K} = \{(0_\mathbb{K}, x) \mid x \in \mathbb{K}\}$ sont des sous-espaces vectoriels de $\mathbb{K}^2$.

— $F = \{(x, y) \in \mathbb{R}^2 \mid 4x - y = 0\}$ est un sous-espace vectoriel de $\mathbb{R}^2$.

— Pour tout entier $n \geq 0$, l'ensemble $\mathbb{K}_n[X]$ des polynômes de degré inférieur ou égal à n est un sous-espace vectoriel de l'ensemble $\mathbb{K}[X]$ des polynômes sur $\mathbb{K}$.

— L'ensemble des suites complexes convergentes est un sous-espace vectoriel de $\mathscr{F}(\mathbb{N}, \mathbb{C})$.

— Pour tout $n \in \mathbb{N}$, l'ensemble $\mathscr{C}^n(\mathbb{R})$ des fonctions réelles n-fois continûment dérivables sur $\mathbb{R}$ est un sous-espace vectoriel de $\mathscr{F}(\mathbb{R}, \mathbb{R})$.

— L'ensemble $\{f \in \mathscr{C}^2(\mathbb{R}) \mid f'' + f' = 0\}$ est un sous-espace vectoriel de $\mathscr{C}^2(\mathbb{R})$.

Le résultat suivant est utile en pratique pour montrer de façon simple que l'on est en présence d'un espace vectoriel.

Proposition 3.8 *Soit F un sous-espace vectoriel de E. Alors F, muni des lois induites par E, est un espace vectoriel sur $\mathbb{K}$.*

3.3.1 Intersection de sous-espaces vectoriels

Proposition 3.9 *Soit $(F_i)_{i \in I}$ une famille de sous-espaces vectoriels de E. Alors $F = \bigcap_{i \in I} F_i$ est un sous-espace vectoriel de E.*

Démonstration. Comme $0_E \in F_i$ pour tout $i \in I$, on a $0_E \in F$. D'autre part, pour tous $(x, y) \in F^2$ et $(\lambda, \mu) \in \mathbb{K}^2$, $\lambda x + \mu y \in F_i$ pour tout $i \in I$, et donc $\lambda x + \mu y \in F$. □

Remarque. En général, l'union de deux sous-espaces vectoriels F_1 et F_2 de E n'est pas un sous-espace vectoriel (penser à deux droites dans le plan vectoriel). Plus précisément, on a $F_1 \cup F_2$ est un sous-espace vectoriel de E si et seulement si $F_1 \subset F_2$ ou $F_2 \subset F_1$.

3.3.2 Sous-espace engendré par une partie

Définition 3.10 *Soient E un $\mathbb{K}$-espace vectoriel et A une partie de E. On appelle* ***sous-espace vectoriel engendré par la partie*** *A, et l'on note* $\mathrm{Vect}A$*, l'intersection de tous les sous-espaces de E contenant A.*

Proposition 3.11 *Soient E un $\mathbb{K}$-espace vectoriel et A une partie de E.*

1) $\mathrm{Vect}A$ *est le plus petit sous-espace vectoriel de E contenant A.*

2) Si $A \neq \varnothing$, alors $\mathrm{Vect}A$ *est l'ensemble des combinaisons linéaires des éléments de A.*

On pourra convenir que $\mathrm{Vect}\,\varnothing = \{0_E\}$. Enfin, on définit le *sous-espace vectoriel engendré par une famille d'éléments* de la manière suivante.

Définition 3.12 *Soient E un $\mathbb{K}$-espace vectoriel et $(x_i)_{i\in I}$ une famille d'éléments de E. On appelle* ***sous-espace vectoriel engendré par la famille*** *$(x_i)_{i\in I}$, et l'on note* $\mathrm{Vect}(x_i)_{i\in I}$*, le sous-espace vectoriel engendré par la partie $\{x_i \mid i \in I\}$ de E.*

Exemples.
- Pour tout vecteur d non nul, $\mathrm{Vect}\{d\} = \{\lambda d \mid \lambda \in \mathbb{K}\}$ est la droite vectorielle engendrée par d, notée $\mathbb{K}d$.
- Dans $\mathbb{R}^2$, si $F = \{(x,y) \in \mathbb{R}^2 \mid 4x - y = 0\}$, alors $F = \mathbb{R}d = \mathrm{Vect}\{d\}$, avec $d = (1,4)$.
- Dans $\mathbb{K}[X]$, pour tout entier naturel n, $\mathbb{K}_n[X] = \mathrm{Vect}\{1, X, \dots, X^n\}$.

3.3.3 Somme de sous-espaces vectoriels

Soient F_1 et F_2 deux sous-espaces vectoriels de E.

Proposition et définition 3.13 *L'ensemble $F_1 + F_2 = \{x_1 + x_2 \mid (x_1, x_2) \in F_1 \times F_2\}$ est appelé* ***somme de F_1 et F_2*** *et est un sous-espace vectoriel de E.*

DÉMONSTRATION. Tout d'abord, $F_1 + F_2 \neq \varnothing$ car $0_E = 0_E + 0_E \in F_1 + F_2$. Par ailleurs, Soit $(x,y) \in (F_1+F_2)^2$. Il existe $(x_1,x_2) \in F_1 \times F_2$ et $(y_1,y_2) \in F_1 \times F_2$ tels que $x = x_1 + x_2$ et $y = y_1 + y_2$. On a alors

$$x + y = (x_1 + x_2) + (y_1 + y_2) = (x_1 + y_1) + (x_2 + y_2) \in F_1 + F_2.$$

Soit enfin $(\lambda, x) \in \mathbb{K} \times (F_1 + F_2)$. Il existe $(x_1, x_2) \in F_1 \times F_2$ tel que $x = x_1 + x_2$. On a

$$\lambda x = \lambda(x_1 + x_2) = \lambda x_1 + \lambda x_2 \in F_1 + F_2.$$

□

Remarque 3.14 *Il est clair que $F_1 + F_2 \subset$* $\mathrm{Vect}(F_1 \cup F_2)$*. Réciproquement, comme $F_1 \cup F_2 \subset F_1 + F_2$, on a* $\mathrm{Vect}(F_1 \cup F_2) \subset F_1 + F_2$ *et l'on a montré que*

$$F_1 + F_2 = \mathrm{Vect}(F_1 \cup F_2).$$

On définit de la même façon la somme de n sous-espaces vectoriels $F_1, \dots, F_n$ de E. On a alors

$$\sum_{i=1}^{n} F_i = \mathrm{Vect}\left(\bigcup_{i=1}^{n} F_i\right).$$

Définition 3.15 *On dit que les sous-espaces vectoriels F_1 et F_2 sont en* ***somme directe*** *lorsque, pour tout x de $F_1 + F_2$, il existe un unique couple $(x_1, x_2) \in F_1 \times F_2$ tel que $x = x_1 + x_2$. On note alors $F_1 \oplus F_2$. Lorsque $E = F_1 \oplus F_2$, on dit que F_1 et F_2 sont* ***supplémentaires dans*** *E.*

Théorème 3.16 *Les assertions suivantes sont équivalentes.*

(1) $E = F_1 \oplus F_2$.

(2) $E = F_1 + F_2$ *et,* $\forall (x_1, x_2) \in F_1 \times F_2$, *si* $x_1 + x_2 = 0_E$ *alors* $x_1 = x_2 = 0_E$.

(3) $E = F_1 + F_2$ *et* $F_1 \cap F_2 = \{0_E\}$.

DÉMONSTRATION. Montrons que (1) implique (2) : si $x_1 + x_2 = 0_E$, alors on dispose de deux décompositions de 0_E et donc $x_1 = x_2 = 0_E$. Montrons que (2) implique (3) : si x admet deux décompositions $x = x_1 + x_2 = x_1' + x_2'$ alors $x_1 - x_1' = x_2' - x_2 \in F_1 \cap F_2$, et donc $x_1 = x_1'$ et $x_2 = x_2'$. □

Dans la pratique, si on utilise la caractérisation (3), il suffit de démontrer que $E \subset F_1 + F_2$ et $F_1 \cap F_2 \subset \{0_E\}$, puisque les deux autres inclusions $F_1 + F_2 \subset E$ et $\{0_E\} \subset F_1 \cap F_2$ sont claires.

Exemples.

— Dans $\mathscr{F}(\mathbb{R},\mathbb{R})$, l'ensemble des fonctions paires et celui des fonctions impaires sont des sous-espaces vectoriels supplémentaires.

— Dans l'espace produit $\mathbb{K}^2$, $\mathbb{K}\times\{0_{\mathbb{K}}\}$ et $\{0_{\mathbb{K}}\}\times\mathbb{K}$ sont des sous-espaces vectoriels supplémentaires.

3.4 Familles génératrices, familles libres et bases

Soient E un $\mathbb{K}$-espace vectoriel, un entier $n\geq 1$ et $(x_i)_{1\leq i\leq n}$ une famille d'éléments de E.

Définition 3.17 *On dit que la famille* $(x_i)_{1\leq i\leq n}$ *est*

- ***génératrice*** *lorsque* $\mathrm{Vect}\{x_1,\ldots,x_n\}=E$ *:* $\forall x\in E,\ \exists(\lambda_1,\ldots,\lambda_n)\in\mathbb{K}^n,\ x=\lambda_1x_1+\cdots+\lambda_nx_n$,
- ***libre*** *lorsque* $\forall(\lambda_1,\ldots,\lambda_n)\in\mathbb{K}^n,\ (\lambda_1x_1+\cdots+\lambda_nx_n=0_E)\Rightarrow(\lambda_i=0_{\mathbb{K}},\ 1\leq i\leq n)$,
- ***une base de*** E *lorsqu'elle est génératrice et libre ; tout élément de* E *est alors combinaison linéaire d'éléments de* $\{x_i\}_{1\leq i\leq n}$*, de manière unique :* $\forall x\in E,\ \exists!(\lambda_1,\ldots,\lambda_n)\in\mathbb{K}^n,\ x=\lambda_1x_1+\cdots+\lambda_nx_n$.

Justifions ce dernier point. Si x appartient à E alors il existe $(\lambda_1,\ldots,\lambda_n)\in\mathbb{K}^n$ tel que $x=\lambda_1x_1+\cdots+\lambda_nx_n$, car la famille $(x_i)_{1\leq i\leq n}$ est génératrice. Si un autre n-uplet $(\mu_1,\ldots,\mu_n)\in\mathbb{K}^n$ convient, alors

$$(\lambda_1-\mu_1)x_1+\cdots+(\lambda_n-\mu_n)x_n=0_E,$$

et donc, comme $(x_i)_{1\leq i\leq n}$ est libre, $\lambda_i=\mu_i$ pour tout i. Dans ce cas, les λ_i, $1\leq i\leq n$, s'appellent les *coordonnées* ou *composantes* du vecteur x dans la base $(x_i)_{1\leq i\leq n}$.
Une famille qui n'est pas libre est dite *liée*. Lorsque la famille $(x_i)_{1\leq i\leq n}$ est liée (resp. libre), on dit aussi que les vecteurs x_i sont *linéairement dépendants* (resp. *linéairement indépendants*). On dit que deux vecteurs sont *colinéaires* lorsque la famille (x_1,x_2) est liée, c'est-à-dire qu'il existe $\lambda\in\mathbb{K}$ tel que $x_1=\lambda x_2$ ou $x_2=\lambda x_1$.

Exemples.

— Une famille contenant le vecteur nul est liée. En effet, on a $1_{\mathbb{K}}0_E=0_E$, d'où, pour tout $\lambda\neq 0$, $\lambda^{-1}(\lambda 0_E)=0_E$.

— Soient $A=\{e_1,\ldots,e_n\}$ une partie d'un $\mathbb{K}$-espace vectoriel E et $e\in\mathrm{Vect}A$. La famille $(e_1,\ldots,e_n,e)$ est liée puisque, par définition, il existe $\lambda_1,\ldots,\lambda_n$ des scalaires non tous nuls tels que $e=\lambda_1e_1+\cdots+\lambda_ne_n$, d'où $\lambda_1e_1+\cdots+\lambda_ne_n+(-1)e=0_E$.

— Dans l'espace usuel à trois dimensions, trois vecteurs non coplanaires forment une base.

— $\mathbb{C}$ est un $\mathbb{C}$-espace vectoriel, de base l'élément $1\in\mathbb{C}$, et un $\mathbb{R}$-espace vectoriel, de base 1 et i.

— Dans $\mathbb{R}^n$ ($n\in\mathbb{N}^*$), la famille constituée des vecteurs $e_1=(1,0,\ldots,0)$, $e_2=(0,1,0,\ldots,0)$,..., $e_n=(0,\ldots,0,1)$ est une base, appelée *base canonique* de $\mathbb{R}^n$.

Remarque. Le second des exemples précédents décrit en fait le cas général : si la partie $\{e_1,\ldots,e_n\}$, composée de vecteurs de E, forme une famille liée, alors il existe $i\in\{1,\ldots,n\}$ tel que e_i est combinaison linéaire des vecteurs e_j, $j\neq i$.

3.5 Théorie de la dimension

3.5.1 Généralités

Soit E un espace vectoriel sur $\mathbb{K}$. À toute famille $\mathscr{F}=(x_i)_{i\in I}$ d'éléments de E, on peut associer la partie $A=\{x_i\,|\,i\in I\}\subset E$. On dira qu'une famille est contenue dans une autre lorsqu'on aura l'inclusion des parties associées. On pose, lorsque I est fini, $\mathrm{card}\,\mathscr{F}=\mathrm{card}\,I$; en particulier, $\mathrm{card}A\leq\mathrm{card}\,\mathscr{F}$.

Définition 3.18 *Un* $\mathbb{K}$*-espace vectoriel est dit* ***de dimension finie*** *s'il admet une famille génératrice de cardinal fini. Sinon, il est dit* ***de dimension infinie****.*

Exemples.

— Les espaces vectoriels $\{0_E\}$ et $\mathbb{K}^n$ $(n \geq 1)$ sont de dimension finie.

— L'espace vectoriel $\mathbb{K}[X]$ n'est pas de dimension finie, car sinon il admettrait une famille génératrice finie $(P_1, \ldots, P_n)$ et on aurait alors, pour tout P de $\mathbb{K}[X]$, $\deg(P) \leq \max\limits_{1 \leq i \leq n} \deg(P_i)$.

Nous commençons par énoncer un résultat relatif à l'existence de bases de E.

Théorème 3.19 ***(« théorème de la base incomplète »)*** *Soit E un espace vectoriel de dimension finie, de famille génératrice $\mathscr{G}$.*

(1) Si $\mathscr{L} \subset \mathscr{G}$ est une famille libre, alors il existe une base $\mathscr{B}$ de E telle que $\mathscr{L} \subset \mathscr{B} \subset \mathscr{G}$.

(2) Plus généralement, si $\mathscr{L}'$ est une famille libre, alors il y a au moins une façon de la compléter en une base $\mathscr{B}'$, en utilisant uniquement des éléments de $\mathscr{G}$.

(3) En particulier, E admet au moins une base.

DÉMONSTRATION. Démontrons (1). Considérons pour cela l'ensemble des familles libres $\mathscr{C}$ telles que $\mathscr{L} \subset \mathscr{B} \subset \mathscr{G}$. Cet ensemble n'est pas vide car il contient $\mathscr{L}$. De plus, le nombre de familles de cet ensemble est fini, puisque la famille génératrice $\mathscr{G}$ a un nombre fini d'éléments. Soit alors $\mathscr{B}$ une de ces familles, choisie de telle sorte qu'elle contienne le plus grand nombre d'éléments ; montrons que $\mathscr{B}$ est une base de E. Tout d'abord, par définition, $\mathscr{B}$ est libre. Prouvons ensuite par l'absurde que $\mathscr{B}$ est génératrice : supposons que $F = \operatorname{Vect} \mathscr{B} \subsetneq E$. Il existe au moins un élément $g \in \mathscr{G}$ qui n'appartient pas à F (sinon $E = \operatorname{Vect} \mathscr{G} \subset F$). Mais alors, alors en notant $\mathscr{B} = (b_1, \ldots, b_n)$, la famille $\mathscr{C} = (b_1, \ldots, b_n, g)$ est libre : si $\lambda_1 b_1 + \cdots + \lambda_n b_n + \lambda_{n+1} g = 0$, alors $\lambda_{n+1} = 0$ (sinon $g \in F$) et donc $\lambda_i = 0$ pour $1 \leq i \leq n$ puisque $\mathscr{B}$ est libre. Ainsi, $\mathscr{C}$ appartient à l'ensemble des familles libres telles que $\mathscr{L} \subset \mathscr{B} \subset \mathscr{G}$, d'où une contradiction.
On en déduit (2) en utilisant (1) avec $\mathscr{L}'$ pour famille libre et $\mathscr{L}' \cup \mathscr{G}$ pour famille génératrice. Enfin, l'assertion (3) est obtenue en utilisant (2) avec $\mathscr{L}'$ choisie telle que $\operatorname{Vect} \mathscr{L}' = \{0_E\}$ (on pourra, dans le cas où E n'est pas l'espace vectoriel nul, prendre $\mathscr{L}' = (x)$ avec $x \neq 0_E$). □

Remarque. On a ici démontré que de toute famille génératrice d'un espace vectoriel E, on peut extraire une base de E.

Théorème 3.20 *Si E admet une famille génératrice de cardinal fini n, alors toute famille libre (et donc toute base) est finie et de cardinal au plus égal à n.*

DÉMONSTRATION. On montre par récurrence sur $n \geq 1$ que si $\mathscr{G} = (g_1, \ldots, g_n)$ est une famille génératrice de E et si $\mathscr{F} = (f_1, \ldots, f_n, f_{n+1})$ est une famille de $n+1$ éléments de E, alors $\mathscr{F}$ est liée.
Pour $n = 1$, on a $f_1 = \lambda_1 g_1$ et $f_2 = \lambda_2 g_1$. On en déduit que $\mathscr{F}$ est liée car ou $f_1 = 0_E$ ou bien $f_2 = \frac{\lambda_2}{\lambda_1} f_1$.
Soit maintenant $n \geq 2$. Il existe une famille $(a_{ij})_{1 \leq i \leq n+1,\ 1 \leq j \leq n}$ de scalaires telle que

$$\begin{array}{ccccccccc}
f_1 & = & a_{11} g_1 & + & \ldots & + & a_{1n-1} g_{n-1} & + & a_{1n} g_n, \\
f_2 & = & a_{21} g_1 & + & \ldots & + & a_{2n-1} g_{n-1} & + & a_{2n} g_n, \\
\vdots & & & \vdots & \ldots & \vdots & & \vdots & \\
f_n & = & a_{n1} g_1 & + & \ldots & + & a_{nn-1} g_{n-1} & + & a_{nn} g_n, \\
f_{n+1} & = & a_{n+11} g_1 & + & \ldots & + & a_{n+1n-1} g_{n-1} & + & a_{n+1n} g_n.
\end{array}$$

Si les coefficients a_{in}, $1 \leq i \leq n+1$, sont nuls, alors les vecteurs f_i, $1 \leq i \leq n+1$, sont dans $E' = \operatorname{Vect}(g_i)_{1 \leq i \leq n-1}$; de l'hypothèse de récurrence, on déduit que la famille $(f_1, \ldots, f_n)$ est liée et donc que $\mathscr{F}$ est liée.
Sinon, il existe un entier i compris entre 1 et $n+1$, disons $i = n+1$ tel que $a_{in} \neq 0$. On peut alors remplacer g_n par $\frac{1}{a_{n+1n}} \left(f_{n+1} - \sum_{j=1}^{n-1} a_{n+1j} g_j \right)$, de sorte que les vecteurs $f'_j = f_j - \frac{a_{jn}}{a_{n+1n}} f_{n+1}$, $1 \leq j \leq n$ sont encore dans E'. Par l'hypothèse de récurrence, la famille $(f'_1, \ldots, f'_n)$ est liée : il existe des scalaires $\lambda_1, \ldots, \lambda_n$ non tous nuls tels que $\sum_{i=1}^n \lambda_i f'_i = \sum_{i=1}^n \lambda_i f_i + \mu f_{n+1} = 0_E$. On en déduit que $\mathscr{F}$ est liée. □

Corollaire 3.21 *Si E est un espace vectoriel de dimension finie, alors toutes ses bases sont finies et ont le même cardinal.*

DÉMONSTRATION. Si $\mathscr{B}$ et $\mathscr{B}'$ sont deux bases, alors $\mathscr{B}$ est libre et $\mathscr{B}'$ est génératrice, donc $\operatorname{card} \mathscr{B} \leq \operatorname{card} \mathscr{B}'$ par le théorème précédent. On obtient l'autre inégalité en échangeant $\mathscr{B}$ et $\mathscr{B}'$. □

Définition 3.22 *Le cardinal d'une base quelconque d'un espace vectoriel E de dimension finie s'appelle la* ***dimension de*** *E et se note* $\dim_{\mathbb{K}} E$ *ou* $\dim E$.

Exemples.

- Pour tout $n \geq 1$, $\dim_{\mathbb{K}} \mathbb{K}^n = n$. En particulier, $\dim_{\mathbb{R}} \mathbb{R} = 1$ et $\dim_{\mathbb{C}} \mathbb{C} = 1$.
- La famille $(X^i)_{0\leq i\leq n}$ est une base de $\mathbb{R}^n[X]$ et donc $\dim_{\mathbb{R}} \mathbb{R}^n[x] = n+1$.
- Si E contient un vecteur x non nul, alors la famille (x) est libre et E est de dimension supérieure ou égale à 1.
- L'espace vectoriel $E = \{0_E\}$ est le seul espace vectoriel de dimension 0.

Corollaire 3.23 *Soit E un espace vectoriel de dimension n. Toute famille libre (resp. génératrice) a pour cardinal au plus (resp. au moins) n, et exactement n si et seulement si c'est une base.*

DÉMONSTRATION. Soit $\mathscr{L}$ une famille libre d'éléments de E. Par le théorème de la base incomplète, on peut la compléter en une base $\mathscr{B}$ de E, d'où $\operatorname{card} \mathscr{L} \leq \operatorname{card} \mathscr{B} = n$. Si $\operatorname{card} \mathscr{L} = n$, alors $\mathscr{L} \subset \mathscr{B} \iff \mathscr{L} = \mathscr{B}$ et $\mathscr{L}$ est bien une base.
Soit une famille génératrice $\mathscr{G}$ de E. On sait que l'on peut extraire de $\mathscr{G}$ une base $\mathscr{B}$ de E, donc $\operatorname{card} \mathscr{G} \geq \operatorname{card} \mathscr{B} = n$. Par ailleurs, si $\operatorname{card} \mathscr{G} = n$, alors $\mathscr{B} \subset \mathscr{G} \iff \mathscr{B} = \mathscr{G}$, donc $\mathscr{G}$ est une base. □

Ce corollaire est souvent utilisé en pratique dans le sens suivant : si E est un espace vectoriel de dimension finie égale à n, alors toute famille libre ou génératrice de E ayant n éléments est une base de E.

Sous-espaces vectoriels

Théorème 3.24 *Soit E un espace vectoriel de dimension finie $n \geq 1$. Tout sous-espace vectoriel F de E est de dimension finie $0 \leq p \leq n$. De plus, $p = n$ si et seulement si $F = E$.*

DÉMONSTRATION. Si $\mathscr{L}$ est une famille libre de F, alors $\operatorname{card} \mathscr{L} \leq n$. Considérons une famille libre $\mathscr{B} = (b_i)_{1\leq i\leq p}$ de F de cardinal maximal $p \leq n$ et montrons que $\mathscr{B}$ est une base de F, c'est-à-dire que $\operatorname{Vect} \mathscr{B} = F$. Si cela n'était pas le cas, il existerait un élément x de F qui ne serait pas une combinaison linéaire d'éléments de $\mathscr{B}$. Dans ce cas, $(b_1, \ldots, b_p, x)$ est une famille libre de F, ce qui contredit la définition de p.
Si de plus $p = n$, alors $\mathscr{B}$ est une base de E (une famille libre de cardinal n), et donc $E = F$. □

Définition 3.25 *Dans tout espace vectoriel de dimension $n \geq 1$, on appelle **droite vectorielle** (resp. **plan vectoriel**) tout sous-espace vectoriel de dimension 1 (resp. 2). Enfin, on appelle **hyperplan de** E tout sous-espace vectoriel de dimension $n-1$.*

La notion d'hyperplan en dimension n généralise celle de droite en dimension $n = 2$ et de plan en dimension $n = 3$.

Exemple. L'ensemble $\{(x, y, z) \in \mathbb{R}^3 \mid 4x - 3y + 5z = 0\}$ est un plan de $\mathbb{R}^3$.

Théorème 3.26 *Soit E un espace vectoriel de dimension finie. Tout sous-espace vectoriel F de E admet au moins un supplémentaire G dans E. De plus, on a $\dim F + \dim G = \dim E$.*

DÉMONSTRATION. En notant $p = \dim F$, on complète une base $(f_i)_{1\leq i\leq p}$ de F en une base $(f_i)_{1\leq i\leq n}$ de E, et on constate que $\operatorname{Vect}(f_i)_{p+1\leq i\leq n}$ est un supplémentaire de F de dimension $n-p$. □

Corollaire 3.27 *Soit E un espace vectoriel de dimension finie, F et G deux sous-espaces vectoriels de E en somme directe. On a alors $\dim(F \oplus G) = \dim F + \dim G$.*

Il est possible de généraliser cette dernière propriété.

Corollaire 3.28 *Soient E un espace vectoriel de dimension finie, $n \in \mathbb{N}^*$, $F_1, \ldots, F_n$ des sous-espaces vectoriels de E en somme directe. On a alors*

$$\dim\left(\bigoplus_{i=1}^{n} F_i\right) = \sum_{i=1}^{n} \dim F_i.$$

Théorème 3.29 *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie. On a, pour tous sous-espaces vectoriels F, G de E,*

$$\dim(F+G) = \dim F + \dim G - \dim(F \cap G).$$

DÉMONSTRATION. D'après le théorème 3.26, le sous-espace vectoriel $F \cap G$ admet au moins un supplémentaire F' dans F. Montrons que F' et G sont en somme directe et que $F' \oplus G = F + G$. D'une part, on a $F' \subset F$ d'où $F' \cap G = (F' \cap F) \cap G = F' \cap (F \cap G) = \{0_E\}$. D'autre part, $F + G = (F' + (F \cap G)) + G = F' + ((F \cap G)) + G) = F' + G$.
Par ailleurs, d'après le corollaire 3.27, on a les égalités suivantes :

$$\dim(F+G) = \dim(F' \oplus G) = \dim F' + \dim G \text{ et } \dim F = \dim(F' \oplus (F \cap G)) = \dim f' + \dim(F \cap G),$$

d'où la relation voulue. □

Proposition 3.30 *Soient E et F deux $\mathbb{K}$-espaces vectoriels de dimension finie. Alors $E \times F$ est de dimension finie et* $\dim(E \times F) = \dim E + \dim F$.

DÉMONSTRATION. D'après le théorème 3.19, les espaces E et F admettent des bases finies respectives $(e_1, \ldots, e_n)$ et $(f_1, \ldots, f_p)$, où $n = \dim E$ et $p = \dim F$. Nous allons montrer que la famille $\mathscr{B}$, constituée des vecteurs $(e_i, 0_F)$, $1 \leq i \leq n$, et $(0_E, f_j)$, $1 \leq j \leq p$, est une base de $E \times F$.
Soit $(\lambda_1, \ldots, \lambda_n, \mu_1, \ldots, \mu_p) \in \mathbb{K}^{n+p}$ tel que

$$\sum_{i=1}^{n} \lambda_i (e_i, 0_F) + \sum_{j=1}^{p} \mu_j (0_E, f_j) = (0_E, 0_F).$$

On a alors $\left(\sum_{i=1}^{n} \lambda_i e_i, \sum_{j=1}^{p} \mu_j f_j \right) = (0_E, 0_F)$, d'où $\sum_{i=1}^{n} \lambda_i e_i = 0_E$ et $\sum_{j=1}^{p} \mu_j f_j = 0_F$, et donc $\lambda_1 = \cdots = \lambda_n = \mu_1 = \cdots = \mu_p = 0$, puisque les familles $(e_1, \ldots, e_n)$ et $(f_1, \ldots, f_p)$ sont libres. Ainsi, $\mathscr{B}$ est libre.
Soit $(x, y) \in E \times F$. Les familles $(e_1, \ldots, e_n)$ et $(f_1, \ldots, f_p)$ engendrant respectivement E et F, il existe $(\lambda_1, \ldots, \lambda_n) \in \mathbb{K}^n$ et $(\mu_1, \ldots, \mu_p) \in \mathbb{K}^p$ tels que $x = \sum_{i=1}^{n} \lambda_i e_i$ et $y = \sum_{j=1}^{p} \mu_j f_j$. On a alors

$$(x, y) = \left(\sum_{i=1}^{n} \lambda_i e_i, \sum_{j=1}^{p} \mu_j f_j \right) = \sum_{i=1}^{n} \lambda_i (e_i, 0_F) + \sum_{j=1}^{p} \mu_j (0_E, f_j).$$

Ainsi, $\mathscr{B}$ engendre E.
La famille $\mathscr{B}$ est par conséquent une base de $E \times F$, qui est donc de dimension finie :

$$\dim(E \times F) = \operatorname{card} B = n + p = \dim E + \dim F.$$

□

Ce résultat se généralise par récurrence au produit de m espaces vectoriels.

Corollaire 3.31 *Soient $m \in \mathbb{N}^*$, E_1, ..., E_m des $\mathbb{K}$-espaces vectoriels de dimension finie. Alors $\prod_{i=1}^{m} E_i$ est de dimension finie et*

$$\dim\left(\prod_{i=1}^{m} E_i \right) = \sum_{i=1}^{m} \dim E_i.$$

3.5.2 Rang d'une famille de vecteurs

Définition 3.32 *Soient E un $\mathbb{K}$-espace vectoriel et $\mathscr{F}$ une famille finie d'éléments de E. On appelle **rang de $\mathscr{F}$**, et l'on note* $\operatorname{rang} \mathscr{F}$*, la dimension du sous-espace vectoriel de E engendré par $\mathscr{F}$:*

$$\operatorname{rang} \mathscr{F} = \dim(\operatorname{Vect} \mathscr{F}).$$

Proposition 3.33 *Soit E un $\mathbb{K}$-espace vectoriel. Pour toutes familles finies $\mathscr{F}$ et $\mathscr{F}'$ d'éléments de E, on a*

1) $\mathscr{F} \subset \mathscr{F}' \Rightarrow \operatorname{rang} \mathscr{F} \leq \operatorname{rang} \mathscr{F}'$,
2) $\max(\operatorname{rang} \mathscr{F}, \operatorname{rang} \mathscr{F}') \leq \operatorname{rang}(\mathscr{F} \cup \mathscr{F}') \leq \operatorname{rang} \mathscr{F} + \operatorname{rang} \mathscr{F}'$.

DÉMONSTRATION.

1) $\mathscr{F} \subset \mathscr{F}' \Rightarrow \operatorname{Vect} \mathscr{F} \subset \operatorname{Vect} \mathscr{F}' \Rightarrow \dim(\operatorname{Vect} \mathscr{F}) \subset \dim(\operatorname{Vect} \mathscr{F}')$.
2) D'une part, $\mathscr{F} \subset \mathscr{F} \cup \mathscr{F}'$ (resp. $\mathscr{F}' \subset \mathscr{F} \cup \mathscr{F}'$) $\Rightarrow \operatorname{rang} \mathscr{F} \leq \operatorname{rang}(\mathscr{F} \cup \mathscr{F}')$ (resp. $\operatorname{rang} \mathscr{F}' \leq \operatorname{rang}(\mathscr{F} \cup \mathscr{F}')$), d'où $\max(\operatorname{rang} \mathscr{F}, \operatorname{rang} \mathscr{F}') \leq \operatorname{rang}(\mathscr{F} \cup \mathscr{F}')$.
D'autre part, $\operatorname{rang}(\mathscr{F} \cup \mathscr{F}') = \dim(\operatorname{Vect}(\mathscr{F} \cup \mathscr{F}')) = \dim(\operatorname{Vect} \mathscr{F} + \operatorname{Vect} \mathscr{F}') \leq \dim(\operatorname{Vect} \mathscr{F}) + \dim(\operatorname{Vect} \mathscr{F}')$, d'après la remarque 3.14.

□

Chapitre 4

Applications linéaires

Après avoir introduit, dans le chapitre 3, la structure d'espace vectoriel, nous allons à présent étudier les applications linéaires, c'est-à-dire les applications entre deux espaces vectoriels qui conservent l'addition interne et la multiplication externe.
Dans ce chapitre, $\mathbb{K}$ désigne une nouvelle fois un corps commutatif (en pratique, $\mathbb{K}=\mathbb{R}$ ou $\mathbb{C}$), E et F sont deux espaces vectoriels sur $\mathbb{K}$.

4.1 Généralités

Définition 4.1 *On appelle* ***application linéaire de*** E ***dans*** F *(ou* ***morphisme de*** $\mathbb{K}$***-espaces vectoriels****) toute application de* $f : E \to F$ *telle que*

$$\forall (x,y)\in E^2,\ \forall(\lambda,\mu)\in\mathbb{K}^2,\ f(\lambda x+\mu y)=\lambda f(x)+\mu f(y). \tag{4.1}$$

On note $\mathscr{L}(E,F)$ *l'ensemble des applications linéaires de* E *dans* F*.*

Si $(E,+,\cdot)=(F,+,\cdot)$, on parle d'*endomorphisme* de E et on note simplement $\mathscr{L}(E)(=\mathscr{L}(E,E))$ l'ensemble des endomorphismes de E.
Dans la pratique, on remplace parfois la condition (4.1) par

$$\forall (x,y)\in E^2,\ \forall\lambda\in\mathbb{K},\ f(x+y)=f(x)+f(y) \text{ et } f(\lambda x)=\lambda f(x).$$

Exemples.
- L'application nulle ($x\mapsto 0_E$) et l'identité de E sont des endomorphismes de E.
- Pour tout λ de $\mathbb{K}$, l'homothétie de rapport λ, définie par $h_\lambda(x)=\lambda x$ pour tout x de E, est un endomorphisme de E.
- L'application $f\mapsto f'$ est une application linéaire de $\mathscr{C}^1(\mathbb{R})$ dans $\mathscr{C}^0(\mathbb{R})$.

Définition 4.2 *Toute application linéaire de* E *dans* $\mathbb{K}$ *est appelée* ***forme linéaire sur*** E*. On note* E^* *(ou encore* E'*) l'ensemble des formes linéaires sur* E*, appelé* ***dual de*** E*.*

Exemple. L'application $f\mapsto \displaystyle\int_a^b f(t)\,\mathrm{d}t$ est une forme linéaire sur $\mathscr{C}^0([a,b])$.

Proposition et définition 4.3 *Soit* f *une application linéaire de* E *dans* F*.*
Le ***noyau*** *(kernel en anglais) de* f*, défini par* $\operatorname{Ker} f=\{x\in E \mid f(x)=0_F\}$*, est un sous-espace vectoriel de* E*.*
*L'****image*** *de* f*, définie par* $\operatorname{Im} f=f(E)=\{y\in F\mid \exists x\in E, y=f(x)\}$*, est un sous-espace vectoriel de* F*.*

DÉMONSTRATION. On a $0_E\in\operatorname{Ker} f$, car $f(0_E)=0_F$. D'autre part, pour tous $(x,y)\in(\operatorname{Ker} f)^2$ et $(\lambda,\mu)\in\mathbb{K}^2$, on a $f(\lambda x+\mu y)=\lambda f(x)+\mu f(y)=0_F$, et donc $\lambda x+\mu y\in\operatorname{Ker} f$.

On a $0_F = f(0_E) \in \operatorname{Im} f$. D'autre part, pour tous $(x', y') \in (\operatorname{Im} f)^2$ et $(\lambda, \mu) \in \mathbb{K}^2$, on a $x' = f(x)$, $y' = f(y)$, et $\lambda x' + \mu y' = f(\lambda x + \mu y) \in \operatorname{Im} f$. □

Plus généralement, on vérifie aisément que l'image réciproque d'un sous-espace vectoriel de F par f est un sous-espace vectoriel de E et que l'image directe d'un sous-espace vectoriel de E par f est un sous-espace vectoriel de F.
Le résultat suivant sera très utile en pratique.

Proposition 4.4 *Soit f une application linéaire de E dans F.*

1) f est injective si et seulement si $\operatorname{Ker} f = \{0_E\}$.

2) f est surjective si et seulement si $\operatorname{Im} f = F$.

DÉMONSTRATION.

1) Supposons f injective. Si $x \in \operatorname{Ker} f$, alors $f(x) = 0_F = f(0_E)$ et donc $x = 0_E$ d'où $\operatorname{Ker} f = \{0_E\}$. Réciproquement, supposons $\operatorname{Ker} f = \{0_E\}$. Si $(x, y) \in E^2$ est tel que $f(x) = f(y)$, alors $f(x - y) = 0_F$ et donc $x - y = 0_E$, d'où $x = y$: f est injective.

2) f surjective $\Longleftrightarrow (\forall y \in F,\ \exists x \in E,\ y = f(x)) \Longleftrightarrow f(E) = F \Longleftrightarrow \operatorname{Im} f = F$.

□

Proposition 4.5 *Si l'application linéaire $f : E \to F$ est bijective, alors sa réciproque $f^{-1} : F \to E$ est linéaire (et bijective).*

DÉMONSTRATION. Pour tous $(x', y') \in F^2$ et $(\lambda, \mu) \in \mathbb{K}^2$, on a $x' = f(x)$, $y' = f(y)$ et $\lambda x' + \mu y' = f(\lambda x + \mu y)$, d'où $f^{-1}(\lambda x' + \mu y') = \lambda x + \mu y = \lambda f^{-1}(x') + \mu f^{-1}(y')$. □

Définitions 4.6 *Une application $f : E \to F$ est un **isomorphisme** de E dans F si et seulement si f est linéaire est bijective.*
*Si f est une application linéaire et bijective de E dans E, f est appelée un **automorphisme** de E.*

On dit que les espaces E et F sont *isomorphes* lorsqu'il existe un isomorphisme d'espaces vectoriels de E sur F.

Exemple. Pour tout λ de $\mathbb{K}^*$, l'homothétie h_λ de rapport λ est un automorphisme et l'on a ${h_\lambda}^{-1} = h_{\frac{1}{\lambda}}$.

4.2 Opérations sur les applications linéaires

Proposition 4.7 *L'ensemble $\mathscr{L}(E, F)$, muni des lois usuelles, est un $\mathbb{K}$-espace vectoriel.*

DÉMONSTRATION. Nous allons montrer que $\mathscr{L}(E, F)$ est un sous-espace vectoriel de $(\mathscr{F}(E, F), +, \cdot)$.
Tout d'abord, $\mathscr{L}(E, F)$ est non vide puisque l'application nulle est à l'évidence linéaire. Soient par ailleurs f, g appartenant à $\mathscr{L}(E, F)$ et $\alpha \in \mathbb{K}$. Pour tous $(x, y) \in E^2$ et $\lambda \in \mathbb{K}$, on a

$$\begin{aligned}(\alpha f + g)(\lambda x + y) &= \alpha f(\lambda x + y) + g(\lambda x + y)\\ &= \alpha(\lambda f(x) + f(y)) + (\lambda g(x) + g(y))\\ &= \lambda(\alpha f(x) + g(x)) + (\alpha f(y) + g(y))\\ &= \lambda(\alpha f + g)(x) + (\alpha f + g)(y).\end{aligned}$$

L'application $\alpha f + g$ est donc linéaire et $\alpha f + g \in \mathscr{L}(E, F)$. □

Composition

Proposition 4.8 *Soient E, F et G trois $\mathbb{K}$-espaces vectoriels. Pour tous f de $\mathscr{L}(E, F)$ et g de $\mathscr{L}(F, G)$, $g \circ f$ appartient à $\mathscr{L}(E, G)$.*

DÉMONSTRATION. On a

$$\forall \lambda \in \mathbb{K},\ \forall (x,y) \in E^2,\ (g \circ f)(\lambda x + y) = g(f(\lambda x + y)) = g(\lambda f(x) + f(y)) = \lambda g(f(x)) + g(f(y))$$
$$= \lambda g(f(x)) + g(f(y)) = \lambda g \circ f(x) + g \circ f(y).$$

□

En particulier, si $E = F = G$, $(f, g) \mapsto g \circ f$ est une loi de composition interne sur $\mathscr{L}(E)$. Elle est associative (pour tous f, g, h dans $\mathscr{L}(E)$, $h \circ (g \circ f) = (h \circ g) \circ f$), a pour élément neutre I_E (pour tout $f \in \mathscr{L}(E)$, $f \circ I_E = I_E \circ f = f$) et est distributive par rapport à l'addition (pour tous f, g, h dans $\mathscr{L}(E)$, $h \circ (g + f) = (h \circ g) + (h \circ f)$ et $(g + f) \circ h = (g \circ h) + (f \circ h)$). Cette loi de composition n'est en général pas commutative.

L'ensemble $\mathscr{GL}(E)$ des éléments inversibles de $\mathscr{L}(E)$, c'est-à-dire des automorphismes de E, est appelé *groupe linéaire de E*.

4.3 Projecteurs et symétries

Définition 4.9 *On appelle* ***projecteur de*** *E tout endomorphisme $p \in \mathscr{L}(E)$ tel que $p \circ p = p$.*

Théorème 4.10 *Si $E = E_1 \oplus E_2$, alors l'application $p_1 \in \mathscr{L}(E)$ définie par $p_1(x_1 + x_2) = x_1$ est un projecteur. Réciproquement, si $p \in \mathscr{L}(E)$ est un projecteur, alors en notant $E_1 = \operatorname{Im} p$ et $E_2 = \operatorname{Ker} p$, on a $E = E_1 \oplus E_2$ et $p = p_1$.*

DÉMONSTRATION. Il est clair que p_1 est linéaire et que $p_1 \circ p_1 = p_1$. Réciproquement, soit p un projecteur de E. Pour tout x de E, on a $x = p(x) + (x - p(x)) \in E_1 + E_2$. D'autre part, si $x \in E_1 \cap E_2$, alors $x = p(y)$ et $p(x) = 0_E$. On en déduit que $x = p(y) = p \circ p(y) = p(x) = 0_E$. Enfin, si $x = x_1 + x_2$, alors $x_1 = p(y)$ et $p_2(x_2) = 0_E$, donc $p(x) = p(p(y)) = p(y) = x_1 = p_1(x)$. □

On dit que p_1 est le projecteur sur E_1 parallèlement à E_2. On définit de même le projecteur p_2 sur E_2 parallèlement à E_1, et l'on a, dans $\mathscr{L}(E)$, $p_1 + p_2 = I_E$ et $p_1 \circ p_2 = p_2 \circ p_1 = 0_{\mathscr{L}(E)}$.

Définition 4.11 *On appelle* ***symétrie de*** *E tout endomorphisme de $s \in \mathscr{L}(E)$ tel que $s \circ s = I_E$.*

On observe que $s \in \mathscr{GL}(E)$ et que $s^{-1} = s$.

Théorème 4.12 *Si $E = E_1 \oplus E_2$, alors l'application $s_1 \in \mathscr{L}(E)$ définie par $s_1(x_1 + x_2) = x_1 - x_2$ est une symétrie. Réciproquement, si $s \in \mathscr{L}(E)$ est une symétrie, alors en notant $E_1 = \operatorname{Ker}(s - I_E)$ et $E_2 = \operatorname{Ker}(s + I_E)$, on a $E = E_1 \oplus E_2$ et $s = s_1$.*

DÉMONSTRATION. Il est clair que s_1 est linéaire et que $s_1 \circ s_1 = I_E$. Réciproquement, soit s une symétrie de E. Pour tout x de E, écrivons $x = \frac{1}{2}(x + s(x)) + \frac{1}{2}(x - s(x)) \in E_1 + E_2$. De plus, si $x \in E_1 \cap E_2$, alors $s(x) = x = -x$ et donc $x = 0_E$. On en déduit que $E = E_1 \oplus E_2$. Finalement, si $x = x_1 + x_2$, alors $s(x) = s(x_1) + s(x_2) = x_1 - x_2 = s_1(x)$. □

On dit que s_1 est la symétrie par rapport à E_1 parallèlement à E_2. On définit de même la symétrie s_2 par rapport à E_2 parallèlement à E_1. On vérifiera alors que $s_1 = p_1 - p_2 = 2p_1 - I_E$.

4.4 Familles de vecteurs et applications linéaires

Proposition 4.13 *Pour toute application f de $\mathscr{L}(E, F)$ et toute famille finie $\mathscr{F}$ d'éléments de E, on a*

$$f(\operatorname{Vect} \mathscr{F}) = \operatorname{Vect}(f(\mathscr{F})).$$

DÉMONSTRATION. Posons $\mathscr{F} = (x_i)_{1 \leq i \leq n}$. Soit $y \in f(\operatorname{Vect} \mathscr{F})$; il existe $x \in \operatorname{Vect} \mathscr{F}$ tel que $y = f(x)$ et des scalaires $(\lambda_i)_{1 \leq i \leq n}$ tels que $x = \sum_{i=1}^{n} \lambda_i x_i$. On a alors

$$y = f(x) = f\left(\sum_{i=1}^{n} \lambda_i x_i\right) = \sum_{i=1}^{n} \lambda_i f(x_i) \in \operatorname{Vect}(f(\mathscr{F})).$$

Ceci montre $f(\operatorname{Vect} \mathscr{F}) \subset \operatorname{Vect}(f(\mathscr{F}))$. L'inclusion réciproque se montre de manière analogue. □

Proposition 4.14 *Soit $f \in \mathscr{L}(E,F)$ et $\mathscr{F}$ une famille finie d'éléments de E.*

1) Si $\mathscr{F}$ est liée, alors $f(\mathscr{F})$ est liée.

2) Si $\mathscr{F}$ est libre, alors $f(\mathscr{F})$ est libre.

DÉMONSTRATION. Posons $\mathscr{F} = (x_i)_{1\leq i\leq n}$.

1) La famille $\mathscr{F}$ étant liée, il existe des scalaires $(\lambda_i)_{1\leq i\leq n}$ non tous nuls tels que $\sum_{i=1}^n \lambda_i x_i = 0_E$. On a alors $\sum_{i=1}^n \lambda_i f(x_i) = f\left(\sum_{i=1}^n \lambda_i x_i\right) = f(0_E) = 0_F$, et donc $f(\mathscr{F})$ est liée.
2) Se déduit de 1) par contraposée.

□

Supposons à présent que E admet une base $\mathscr{B} = (e_i)_{1\leq i\leq n}$ ($n \geq 1$). Nous allons montrer qu'une application linéaire $f : E \to F$ peut se définir par les images $f(e_i)$, $1 \leq i \leq n$, des vecteurs de la base $\mathscr{B}$.

Théorème 4.15 *Pour toute famille $\mathscr{C} = (f_i)_{1\leq i\leq n}$ d'éléments de F, il existe une unique application linéaire $f : E \to F$ telle que $f(\mathscr{B}) = \mathscr{C}$, c'est-à-dire que $f(e_i) = f_i$ pour tout i compris entre 1 et n. Cette application est surjective (resp. injective, resp. bijective) si et seulement si $\mathscr{C}$ est génératrice (resp. libre, resp. une base) dans F.*

DÉMONSTRATION. Tout vecteur x de E s'écrit de manière unique $x = x_1e_1 + \cdots + x_ne_n$ dans la base $\mathscr{B}$. En posant, $f(x) = x_1f_1 + \cdots + x_nf_n$, il est évident que $f(e_i) = f_i$ pour tout $1 \leq i \leq n$ et que $f : E \to F$ est linéaire. On a aussi l'unicité, car si $g : E \to F$ linéaire vérifie $g(e_i) = f_i$, alors pour tout $x \in E$, $g(x) = x_1f_1 + \cdots + x_nf_n = f(x)$.
Pour démontrer les équivalences, il suffit d'écrire les définitions. L'application f est surjective lorsque, pour tout y de F, il existe x de E tel que $y = f(x)$. Ceci est équivalent à l'existence de $(x_1,\dots,x_p) \in \mathbb{K}^n$ tel que $x = x_1e_1 + \cdots + x_ne_n$ et $y = x_1f_1 + \cdots + x_nf_n$, et donc à $\mathscr{C}$ est une famille génératrice. D'autre part, f est injective lorsque, pour tout x de E, $f(x) = 0_F$ implique $x = 0_E$. En écrivant encore $x = x_1e_1 + \cdots + x_ne_n$, ceci est équivalent à l'implication : si $x_1f_1 + \cdots + x_nf_n = 0_F$ alors $(x_1,\dots,x_p) = (0,\dots,0)$, et donc la famille $\mathscr{C}$ est libre. □

Nous en déduisons une propriété utile dans la pratique.

Corollaire 4.16 *Si $f : E \to F$ est une application linéaire et si $\mathscr{C} = f(\mathscr{B}) = (f(e_1),\dots,f(e_n))$, alors f est un isomorphisme si et seulement si $\mathscr{C}$ est une base de F.*

4.5 Cas de la dimension finie

Soient E, F deux espaces vectoriels de dimension finie et f une application linéaire de E dans F.

Définition 4.17 *On appelle **rang de l'application** f, et on note* $\operatorname{rang} f$, *la dimension de l'image de f :*

$$\operatorname{rang} f = \dim(\operatorname{Im} f).$$

Théorème 4.18 ***(« théorème du rang »)*** *On a*

$$\dim E = \dim(\operatorname{Ker} f) + \operatorname{rang} f.$$

DÉMONSTRATION. Notons $n = \dim E$. Le sous-espace-vectoriel $\operatorname{Ker} f$ de E admet au moins une base $(e_i)_{1\leq i\leq p}$ que l'on peut compléter en une base $(e_i)_{1\leq i\leq n}$ de E. Nous allons montrer que $(f(e_{p+1}),\dots,f(e_n))$ est une base de $\operatorname{Im} f$. Les vecteurs $f(e_i)$, $p+1 \leq i \leq n$, sont à l'évidence des éléments de $\operatorname{Im} f$. Soit $(\lambda_{p+1},\dots,\lambda_n) \in \mathbb{K}^{n-p}$ tel que

$$\sum_{i=p+1}^n \lambda_i f(e_i) = 0_F.$$

On a alors $f\left(\displaystyle\sum_{i=p+1}^n \lambda_i e_i\right) = 0_F$, et donc $\displaystyle\sum_{i=p+1}^n \lambda_i e_i \in \operatorname{Ker} f$.

Il existe donc $(\mu_1,\dots,\mu_p) \in \mathbb{K}^p$ tel que $\displaystyle\sum_{i=p+1}^n \lambda_i e_i = \sum_{i=1}^p \mu_i e_i$, d'où $\mu_1e_1 + \cdots + \mu_pe_p - \lambda_{p+1}e_{p+1} - \cdots - \lambda_ne_n = 0_E$.

Comme la famille $(e_1,\dots,e_p)$ est libre, on en déduit que $\lambda_{p+1}=\cdots=\lambda_n=0$, ce qui montre que $(f(e_{p+1}),\dots,f(e_n))$ est libre.
Soit maintenant $y\in\operatorname{Im} f$. Il existe $x\in E$ tel que $y=f(x)$. Comme $(e_1,\dots,e_n)$ engendre E, il existe $(\alpha_1,\dots,\alpha_n)\in\mathbb{K}^n$ tel que $x=\sum_{i=1}^n \alpha_i e_i$. On a alors

$$y=f(x)=f\left(\sum_{i=1}^n \alpha_i e_i\right)=\sum_{i=1}^n \alpha_i f(e_i)=\sum_{i=p+1}^n \alpha_i f(e_i),$$

puisque les vecteurs e_i, $1\le i\le p$, appartiennent au noyau de f. La famille $(f(e_{p+1}),\dots,f(e_n))$ engendre donc $\operatorname{Im} f$ et c'est une base de ce sous-espace de F. On conclut alors

$$\operatorname{rang} f=\dim(\operatorname{Im} f)=n-p=\dim E-\dim(\operatorname{Ker} f).$$

□

Exemple. Si $f:E\to\mathbb{K}$ est une forme linéaire non nulle, alors $\operatorname{rang} f=1$ et $\operatorname{Ker} f$ est un hyperplan de E. On dit que $f(x)=0$ est une équation de cet hyperplan. En particulier, si $\alpha_1,\dots,\alpha_n$ sont des réels non tous nuls, alors $\{(x_1,\dots,x_n)\in\mathbb{R}^n \mid \alpha_1x_1+\cdots+\alpha_nx_n=0\}$ est un hyperplan de $\mathbb{R}^n$.

Remarque. Même lorsque $E=F$, on n'a généralement pas $E=\operatorname{Ker} f\oplus\operatorname{Im} f$. Par exemple, , pour $f:\mathbb{R}\to\mathbb{R}^2$ définie par $f(x,y)=(0,x)$, on a $\operatorname{Ker} f=\operatorname{Im} f=\operatorname{Vect}\{(0,1)\}$. Toutefois, on pourra vérifier que $E=\operatorname{Ker} f+\operatorname{Im} f$ si et seulement si $\operatorname{Ker} f\cap\operatorname{Im} f=\{0_E\}$.

Quelques propriétés fondamentales du rang sont récapitulées dans le théorème suivant.

Théorème 4.19 *Soient E, F deux $\mathbb{K}$-espaces vectoriels et f une application linéaire de E dans F. Alors*

- *on a $\operatorname{rang} f\le\dim F$ avec $\operatorname{rang} f=\dim F$ si et seulement si f est surjective,*
- *on a $\operatorname{rang} f\le\dim E$ avec $\operatorname{rang} f=\dim E$ si et seulement si f est injective,*
- *si $\dim E=\dim F=n$, les assertions suivantes sont équivalentes :*
  (1) f est injective, (2) f est surjective, (3) f est bijective, (4) $\operatorname{rang} f=n$.

DÉMONSTRATION. Comme $\operatorname{Im} f$ est un sous-espace vectoriel de F, on a $\operatorname{rang} f\le\dim F$, avec égalité si et seulement si $\operatorname{Im} f=F$, c'est-à-dire si f est surjective.
D'autre part, d'après le théorème du rang, $\operatorname{rang} f=\dim E-\dim(\operatorname{Ker} f)$, et donc $\operatorname{rang} f\le\dim E$, avec égalité si et seulement si $\dim(\operatorname{Ker} f)=0$, c'est-à-dire f est est injective.
De plus, si $\dim E=\dim F=n$, alors, d'après ce qui précède, on a (1)$\Longleftrightarrow$(2) et chacune des assertions est équivalente à (3). Finalement, on a (2)$\Longleftrightarrow$(4), par définition du rang. □

Rappelons que si f est bijective, alors f^{-1} est linéaire. On en déduit la propriété suivante.

Corollaire 4.20 *Pour tout endomorphisme $f\in\mathscr{L}(E)$, les assertions suivantes sont équivalentes :*

(1) f est inversible dans $\mathscr{L}(E)$, c'est-à-dire qu'il existe $g\in\mathscr{L}(E)$ tel que $f\circ g=g\circ f=I_E$.

(2) f est inversible à droite dans $\mathscr{L}(E)$, c'est-à-dire qu'il existe $g\in\mathscr{L}(E)$ tel que $f\circ g=I_E$.

(3) f est inversible à gauche dans $\mathscr{L}(E)$, c'est-à-dire qu'il existe $g\in\mathscr{L}(E)$ tel que $g\circ f=I_E$.

On a alors $f\in\mathscr{GL}(E)$ et $f^{-1}=g$.

DÉMONSTRATION. Il est clair que (1) implique (2) et (1) implique (3).
Montrons que (2) implique (1). Si $f\circ g=I_E$, $f\circ g$ est surjective et donc f est surjective ; ainsi f est bijective d'après le théorème précédent.
Montrons que (3) implique (1). Si $g\circ f=I_E$, $g\circ f$ est injective et donc f est injective ; ainsi f est bijective d'après le théorème précédent. □

Chapitre 5

Matrices

Étant donnés deux espaces vectoriels E et F sur $\mathbb{K}$, une base $\mathscr{B} = (e_1, \dots, e_p)$ de E et une base $\mathscr{C} = (f_1, \dots, f_n)$ de F, nous avons vu que toute application linéaire f de E dans F est entièrement déterminée par la donnée des images $f(e_1), \dots, f(e_p)$ des vecteurs de $\mathscr{B}$ par f, c'est-à-dire par la donnée des coordonnées de chacun des vecteurs $f(e_1), \dots, f(e_p)$ sur la base $\mathscr{C}$ de F. Ces coordonnées peuvent être rangées en un tableau à n lignes et p colonnes, appelé *matrice*. Ceci va nous permettre d'utiliser des algorithmes de calcul en algèbre linéaire.
Dans ce chapitre, $\mathbb{K}$ désigne un corps commutatif (en pratique $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$), n et p sont deux entiers naturels non nuls.

5.1 Calcul matriciel

5.1.1 Notion de matrice

Définition 5.1 *Une* ***matrice*** A ***d'ordre*** (n, p) ***à éléments dans*** $\mathbb{K}$ *est un tableau rectangulaire, composé de n lignes et p colonnes, d'éléments $(a_{ij})_{1 \le i \le n,\ 1 \le j \le p}$ de $\mathbb{K}$, que l'on écrit sous la forme*

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{1p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{pmatrix}.$$

L'ensemble de ces matrices se note $M_{n,p}(\mathbb{K})$.

Le couple (n, p) est le *format* de la matrice A, également dite *de type* (n, p) ; l'entier n est le *nombre de lignes* de A et l'entier p le *nombre de colonnes* de A. Pour $(i, j) \in \{1, \dots, n\} \times \{1, \dots, p\}$, l'élément a_{ij}, situé à la i^{e} ligne et j^{e} colonne, s'appelle le $(i, j)^{\text{e}}$ *terme* ou *coefficient* de A.
Si $n = p$, on parle de *matrices carrées d'ordre n* et leur ensemble se note $M_n(\mathbb{K})$. La famille $(a_{ii})_{1 \le i \le n}$ est alors appelée la *diagonale* de A et les scalaires a_{ii}, $1 \le i \le n$, les *éléments diagonaux* de A. Une *matrice diagonale* est une matrice carrée dont tous les éléments hors de la diagonale sont nuls. On la note parfois $\mathrm{diag}(a_{11}, \dots, a_{nn})$. Lorsque de plus tous les éléments diagonaux sont égaux, on parle de *matrice scalaire*. Par exemple,

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix} \in M_n(\mathbb{K})$$

est une matrice scalaire qu'on appelle la *matrice identité d'ordre n*.
On dit que A est une *matrice-colonne* (ou *matrice unicolonne*) (resp. *matrice-ligne* (ou *matrice uniligne*)) si et seulement si $p = 1$ (resp. $n = 1$).
Soit A une matrice d'ordre (n, p). Pour $i \in \{1, \dots, n\}$, la matrice-ligne $(a_{i1} \ \dots \ a_{ip})$ est appelée la i^{e} *ligne* de A. Pour

$j \in \{1, \dots, p\}$, la matrice-colonne $\begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix}$ est appelée la j^{e} *colonne* de A.
Une matrice carrée est dite *triangulaire supérieure* lorsque les éléments situés en dessous de la diagonale sont nuls, c'est-à-dire $a_{ij} = 0$ pour tout (i, j) tel que $1 \leq j < i \leq n$. On définit de manière analogue les matrices *triangulaires inférieures*.
Pour toute matrice A de $M_{n,p}(\mathbb{K})$, on appelle *transposée* de A, et l'on note ${}^tA$ la matrice d'ordre (p, n) obtenue en échangeant les lignes et les colonnes de A. Ainsi, ${}^tA = (b_{ij})_{1 \leq i \leq p,\ 1 \leq n \leq p}$, avec $b_{ij} = a_{ji}$ pour tous i et j. On remarque que si une matrice carrée A est triangulaire supérieure (resp. inférieure), alors ${}^tA$ est une matrice triangulaire inférieure (resp. supérieure), et que la transposée d'une matrice-ligne (resp. matrice-colonne) est une matrice-colonne (resp. matrice-ligne).
Enfin, on dit qu'une matrice carrée est *symétrique* lorsque ${}^tA = A$ et *antisymétrique* lorsque ${}^tA = -A$; dans ce dernier cas, $a_{ii} = 0$ pour tout $i \in \{1, \dots, n\}$.

5.1.2 Matrice représentative d'une application linéaire

Dorénavant, E désignera un espace vectoriel sur $\mathbb{K}$ de dimension p muni d'une base $\mathscr{B} = (e_j)_{1 \leq j \leq p}$ et F sera un espace vectoriel sur $\mathbb{K}$ de dimension n muni d'une base $\mathscr{C} = (f_i)_{1 \leq i \leq n}$.
Pour tout vecteur x de E, la matrice-colonne $\begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$ est appelée la *matrice-colonne des composantes de x dans la base $\mathscr{B}$* et est notée $\mathrm{mat}_{\mathscr{B}}(x)$. Ainsi, pour toute famille $\mathscr{F} = (c_1, \dots, c_p)$ de p vecteurs de F, on note $\mathrm{mat}_{\mathscr{C}}(\mathscr{F})$ la matrice, appartenant à $M_{n,p}(\mathbb{K})$, de la famille $\mathscr{F}$ relativement à la base $\mathscr{C}$, c'est-à-dire la matrice dont les colonnes sont $\mathrm{mat}_{\mathscr{C}}(c_1)$, ..., $\mathrm{mat}_{\mathscr{C}}(c_p)$.
Considérons à présent une application linéaire $f : E \to F$. Pour tout vecteur x de E, on a : $x = \sum_{j=1}^{p} x_j e_j$ et alors : $f(x) = \sum_{j=1}^{p} x_j f(e_j)$. Pour tout $j \in \{1, \dots, p\}$, exprimons $f(e_j)$ dans la base $\mathscr{C}$; on a : $f(e_j) = \sum_{i=1}^{n} a_{ij} f_i$, d'où : $f(x) = \sum_{j=1}^{p} \sum_{i=1}^{n} x_j a_{ij} f_i$. L'unicité de la décomposition des vecteurs de F dans $\mathscr{C}$ permet alors de dire que $f(x) = \sum_{i=1}^{n} y_i f_i$, avec $y_i = \sum_{j=1}^{p} a_{ij} x_j$.

Définition 5.2 *Soit $f \in \mathscr{L}(E, F)$. On appelle **matrice représentative de f relativement aux bases $\mathscr{B}$ et $\mathscr{C}$** la matrice* $\mathrm{mat}_{\mathscr{C}}(f(\mathscr{B}))$ *appartenant à $M_{n,p}(\mathbb{K})$. On la notera* $\mathrm{mat}_{\mathscr{B},\mathscr{C}}(f)$.

On remarquera ici que les éléments de la j^{e} colonne de $\mathrm{mat}_{\mathscr{B},\mathscr{C}}(f)$ sont les composantes du vecteur $f(e_j)$ dans la base $\mathscr{C}$, c'est-à-dire $\mathrm{mat}_{\mathscr{C}}(f(e_j))$.
Lorsque $f \in \mathscr{L}(E)$ et que l'espace vectoriel E est muni de la base $\mathscr{B}$, on note $\mathrm{mat}_{\mathscr{B}}(f)$ au lieu de $\mathrm{mat}_{\mathscr{B},\mathscr{B}}(f)$.

Exemples.

— Soit $f : E \to F$ l'application nulle, *i.e.* pour tout $e \in E$, $f(e) = 0_F$. Par conséquent, pour tout vecteur e_j, $1 \leq j \leq p$ d'une base $\mathscr{B}$ de E, on a : $f(e_j) = 0_F = 0 f_1 + \cdots + 0 f_n$, où $(f_1, \dots, f_n) = \mathscr{C}$ est une base de F, c'est-à-dire que la j^{e} colonne de $\mathrm{mat}_{\mathscr{B},\mathscr{C}}(f)$ est nulle, pour tout indice j compris entre 1 et p. Cette application est donc représentée par la matrice nulle, dont tous les termes sont égaux à zéro.

— Soit $f : E \to E$ l'application identité. On a alors : $f(e_j) = e_j = 0e_1 + \cdots + 1e_j + \cdots + 0e_p$, $1 \leq j \leq p$. On en déduit que $\mathrm{mat}_{\mathscr{B}}(f) = I_p$, la matrice identité.
Plus généralement, si $f : E \to E$ homothétie de rapport $\lambda \in \mathbb{K}$, alors $f(e_j) = \lambda e_j$, $1 \leq j \leq p$, et :

$$\mathrm{mat}_{\mathscr{B}}(f) = \begin{pmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda \end{pmatrix} \in M_p(\mathbb{K}).$$

— Dans $\mathbb{R}^3$, muni du repère d'origine $(0, 0, 0)$ et de base orthonormée $\{e_1, e_2, e_3\}$, soit $f : \mathbb{R}^3 \to \mathbb{R}^3$ la rotation d'angle θ autour de l'axe dirigé selon e_3. On a alors : $f(e_1) = \cos\theta\, e_1 + \sin\theta\, e_2$, $f(e_2) = -\sin\theta\, e_1 + \cos\theta\, e_2$

et $f(e_3)=e_3$. On en déduit :

$$\mathrm{mat}_{(e_i)_{1\le i\le 3}}(f)=\begin{pmatrix}\cos\theta & -\sin\theta & 0\\ \sin\theta & \cos\theta & 0\\ 0 & 0 & 1\end{pmatrix}.$$

Théorème 5.3 *L'application $f\mapsto \mathrm{mat}_{\mathscr{B},\mathscr{C}}(f)$ est une bijection de $\mathscr{L}(E,F)$ sur $M_{n,p}(\mathbb{K})$.*

DÉMONSTRATION. Soit $A=(a_{ij})_{1\le i\le n,\ 1\le j\le p}\in M_{n,p}(\mathbb{K})$. Nous avons vu dans le chapitre précédent (théorème 4.15) qu'en posant posant $b_j=a_{1j}f_1+\cdots+a_{nj}f_n$, $1\le j\le p$, il existe une unique application linéaire $f\ :\ E\to F$ telle que, pour tout $j\in\{1,\dots,p\}$, $b_j=f(e_j)$, c'est-à-dire telle que $\mathrm{mat}_{\mathscr{B},\mathscr{C}}(f)=A$. □

5.1.3 L'espace vectoriel $(M_{n,p}(\mathbb{K}),+,\cdot)$

La bijection $\mathrm{mat}_{\mathscr{B},\mathscr{C}}$ que nous venons de mettre en évidence va nous permettre d'« amener » sur $M_{n,p}(\mathbb{K})$ la structure vectorielle existant sur $\mathscr{L}(E,F)$. En effet, si $\mathscr{B}=(e_1,\dots,e_p)$ et $\mathscr{C}=(f_1,\dots,f_n)$ sont respectivement des bases des espaces vectoriels E et F, on a, pour tout $\lambda\in\mathbb{K}$, f, $g\in\mathscr{L}(E,F)$ et $A=(a_{ij})_{1\le i\le n,\ 1\le j\le p}=\mathrm{mat}_{\mathscr{B},\mathscr{C}}(f)$, $B=(b_{ij})_{1\le i\le n,\ 1\le j\le p}=\mathrm{mat}_{\mathscr{B},\mathscr{C}}(g)$:

$$\forall j\in\{1,\dots,p\},\ f(e_j)=\sum_{i=1}^n a_{ij}f_i \text{ et } g(e_j)=\sum_{i=1}^n b_{ij}f_i,$$

d'où :

$$\forall j\in\{1,\dots,p\},\ (\lambda f+g)(e_j)=\sum_{i=1}^n(\lambda a_{ij}+b_{ij})f_i.$$

Ceci nous conduit aux définitions suivantes.

Définitions 5.4 *On appelle* ***addition*** *dans $M_{n,p}(\mathbb{K})$ la loi interne, notée $+$, définie par :*

$$\forall A\in M_{n,p}(\mathbb{K}),\ \forall B\in M_{n,p}(\mathbb{K}),\ A+B=(a_{ij}+b_{ij})_{1\le i\le n,\ 1\le j\le p}.$$

On appelle ***multiplication par les scalaires*** *dans $M_{n,p}(\mathbb{K})$ la loi externe $\mathbb{K}\times M_{n,p}(\mathbb{K})\to M_{n,p}(\mathbb{K})$, notée $\cdot$, définie par :*

$$\forall\lambda\in\mathbb{K},\ \forall A\in M_{n,p}(\mathbb{K}),\ \lambda\cdot A=(\lambda a_{ij})_{1\le i\le n,\ 1\le j\le p}.$$

On note souvent λA au lieu de $\lambda\cdot A$. Il est facile de vérifier que $(M_{n,p}(\mathbb{K}),+,\cdot)$ est un espace vectoriel sur $\mathbb{K}$, notamment en le voyant comme l'espace vectoriel $(\mathbb{K}^{n\times p},+,\cdot)$, les calculs étant disposés dans des tableaux. En particulier, l'élément neutre de $M_{n,p}(\mathbb{K})$ est la matrice dont tous les termes sont nuls, dite *matrice nulle*.
Pour tout $(i,j)\in\{1,\dots,n\}\times\{1,\dots,p\}$, on définit la matrice E_{ij} de $M_{n,p}(\mathbb{K})$ dont tous les termes sont nuls sauf le $(i,j)^{\text{ème}}$ qui vaut 1. Les matrices E_{ij} sont appelées les *matrices élémentaires*.

Proposition 5.5 *La famille de matrices $(E_{ij})_{1\le i\le n,\ 1\le j\le p}$ est une base de $M_{n,p}(\mathbb{K})$, appelée* ***base canonique*** *de $M_{n,p}(\mathbb{K})$, et $\dim M_{n,p}(\mathbb{K})=np$.*

DÉMONSTRATION. Pour toute matrice $A=(a_{ij})_{1\le i\le n,\ 1\le j\le p}$ de $M_{n,p}(\mathbb{K})$, on a : $A=\sum_{i=1}^n\sum_{j=1}^p a_{ij}E_{ij}$. Par ailleurs, la somme $\sum_{i=1}^n\sum_{j=1}^p a_{ij}E_{ij}$ est nulle si et seulement si tous les coefficients a_{ij} sont nuls.
La famille $(E_{ij})_{1\le i\le n,\ 1\le j\le p}$ est donc libre et engendre $M_{n,p}(\mathbb{K})$, c'est une base de cet espace vectoriel. □

On en déduit que $\dim M_n(\mathbb{K})=n^2$. Par ailleurs, nous avons démontré (cf. théorème 5.3) que l'application $f\mapsto\mathrm{mat}_{\mathscr{B},\mathscr{C}}(f)$ de $\mathscr{L}(E,F)$ sur $M_{n,p}(\mathbb{K})$ était un isomorphisme d'espaces vectoriels, et donc[1] $\dim\mathscr{L}(E,F)=np$, $\dim\mathscr{L}(E)=n^2$.

Le lecteur est invité à vérifier que les ensembles des matrices diagonales, triangulaires inférieures ou triangulaires supérieures sont des sous-espaces vectoriels de $M_n(\mathbb{K})$ et à donner leurs dimensions respectives.
On vérifiera également que la transposition est un isomorphisme entre les espaces vectoriels $M_{n,p}(\mathbb{K})$ et $M_{p,n}(\mathbb{K})$. Plus précisément, pour toutes matrices A et B de $M_{n,p}(\mathbb{K})$ et tout scalaire $\lambda\in\mathbb{K}$, on a :

$${}^t(A+B)={}^tA+{}^tB,\ {}^t(\lambda A)=\lambda\,{}^tA,\ {}^t({}^tA)=A.$$

1. On remarquera aussi à cette occasion que le dual de E a la même dimension que E, puisque $\dim E^*=\dim\mathscr{L}(E,\mathbb{K})=\dim E$.

5.1.4 Multiplication des matrices

Soient G un espace vectoriel de dimension $q \geq 1$, muni d'une base $\mathscr{D} = (g_1, \ldots, g_q)$, $f \in \mathscr{L}(E, F)$ et $g \in \mathscr{L}(G, E)$. Posons $A = \mathrm{mat}_{\mathscr{B},\mathscr{C}}(f)$ et $B = \mathrm{mat}_{\mathscr{D},\mathscr{B}}(g)$. Nous allons définir la matrice $C = \mathrm{mat}_{\mathscr{D},\mathscr{C}}(f \circ g)$ de l'application $f \circ g$ relativement aux bases $\mathscr{D}$ et $\mathscr{C}$.
Pour tout $j \in \{1, \ldots, q\}$, nous avons d'une part : $f \circ g(g_j) = \sum_{i=1}^n c_{ij} f_i$, d'autre part :

$$f \circ g(g_j) = \sum_{k=1}^p b_{kj} f(e_k) = \sum_{k=1}^p b_{kj} \left(\sum_{i=1}^n a_{ik} f_i \right) = \sum_{k=1}^p \sum_{i=1}^n a_{ik} b_{kj} f_i = \sum_{i=1}^n \left(\sum_{k=1}^p a_{ik} b_{kj} \right) f_i,$$

et ainsi $c_{ij} = \displaystyle\sum_{k=1}^p a_{ik} b_{kj}$, $\forall (i,j) \in \{1, \ldots, n\} \times \{1, \ldots, q\}$.
Ceci nous conduit à la définition suivante.

Définition 5.6 *Soient $A \in M_{n,p}(\mathbb{K})$ et $B \in M_{p,q}(\mathbb{K})$ deux matrices. On appelle **produit de** A **par** B, et l'on note AB, la matrice de $M_{n,q}(\mathbb{K})$ définie par :*

$$AB = (c_{ij})_{1 \leq i \leq n,\ 1 \leq j \leq q} \text{ où } c_{ij} = \sum_{k=1}^p a_{ik} b_{kj}.$$

Exemples.

- $\begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \left(\displaystyle\sum_{i=1}^n x_i y_i \right)$.
- $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \begin{pmatrix} y_1 & \cdots & y_p \end{pmatrix} = \begin{pmatrix} x_1 y_1 & \cdots & x_1 y_p \\ \vdots & & \vdots \\ x_n y_1 & \cdots & x_n y_p \end{pmatrix}$.

En pratique, pour effectuer le produit AB de deux matrices, on dispose A, B et AB de la manière suivante :

$$\begin{array}{cccc} & & & j^{\text{ème}}\text{ colonne} \\ & & & \downarrow \\ & & \times & \begin{pmatrix} & b_{1j} & \\ & \vdots & \\ \cdots & b_{kj} & \cdots \\ & \vdots & \\ & b_{pj} & \end{pmatrix} \\ i^{\text{ème}}\text{ ligne} \rightarrow & \begin{pmatrix} & & \vdots & & \\ a_{i1} & \cdots & a_{ik} & \cdots & a_{ip} \\ & & \vdots & & \end{pmatrix} & = & \begin{pmatrix} & \vdots & \\ \cdots & \displaystyle\sum_{k=1}^p a_{ik} b_{kj} & \cdots \\ & \vdots & \end{pmatrix} \end{array}$$

Par exemple :

$$\begin{array}{ccc} & \times & \begin{pmatrix} 1 & 2 & -1 & 4 \\ 1 & 0 & 0 & -2 \\ -1 & 1 & 1 & -2 \end{pmatrix} \\ \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & -1 \end{pmatrix} & = & \begin{pmatrix} 0 & 5 & 2 & -6 \\ 2 & -1 & -1 & 0 \end{pmatrix} \end{array}$$

L'application de $M_{n,p}(\mathbb{K}) \times M_{p,q}(\mathbb{K})$ dans $M_{n,q}(\mathbb{K})$ qui à (A, B) associe AB s'appelle la *multiplication des matrices*. On retiendra que la matrice représentative d'une composée d'applications est le produit des matrices représentatives de ces applications dans leurs bases respectives :

$$\mathrm{mat}_{\mathscr{D},\mathscr{C}}(f \circ g) = \mathrm{mat}_{\mathscr{B},\mathscr{C}}(f)\, \mathrm{mat}_{\mathscr{D},\mathscr{B}}(g).$$

L'associativité et la distributivité de la multiplication de matrices se vérifient en utilisant les applications linéaires. On a :

— $\forall A \in M_{n,p}(\mathbb{K})$, $\forall B, C \in M_{p,q}(\mathbb{K})$, $A(B+C) = AB + AC$ *(pseudo-distributivité à gauche)*,
— $\forall A, B \in M_{n,p}(\mathbb{K})$, $\forall C \in M_{p,q}(\mathbb{K})$, $(A+B)C = AC + BC$ *(pseudo-distributivité à droite)*,
— $\forall A \in M_{n,p}(\mathbb{K})$, $\forall B \in M_{p,q}(\mathbb{K})$ et $\forall C \in M_{q,r}(\mathbb{K})$, $(AB)C = A(BC)$ *(pseudo-associativité)*,
— $\forall \lambda \in \mathbb{K}$, $\forall A \in M_{n,p}(\mathbb{K})$ et $\forall B \in M_{p,q}(\mathbb{K})$, $\lambda(AB) = (\lambda A)B = A(\lambda B)$.

On vérifie également que, pour toutes matrices $\forall A \in M_{n,p}(\mathbb{K})$ et $\forall B \in M_{p,q}(\mathbb{K})$, ${}^t(AB) = {}^tB\,{}^tA$. En effet, en posant $C = {}^t(AB)$ et $D = {}^tB\,{}^tA$, on a directement :

$$c_{ij} = \sum_{k=1}^{p} a_{jk} b_{ki} = d_{ij}.$$

5.1.5 Matrices carrées

La multiplication des matrices définit une loi interne sur l'espace vectoriel $M_n(\mathbb{K})$, d'élément neutre I_n, associative et distributive par rapport à l'addition. Pour $n = 1$, c'est la multiplication usuelle dans le corps $\mathbb{K}$. Dans ce cas particulier, elle est commutative et vérifie la propriété suivante : si $a \in \mathbb{K}$ et $b \in \mathbb{K}$ sont tels que $ab = 0$, alors ou $a = 0$ ou $b = 0$. Nous allons voir que pour $n \leq 2$, ces deux propriétés ne subsistent pas.

Exemple. Si $A = \begin{pmatrix} 0 & 5 \\ 0 & 8 \end{pmatrix}$ et $B = \begin{pmatrix} -4 & 3 \\ 0 & 0 \end{pmatrix}$, alors $AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ et $BA = \begin{pmatrix} 0 & 4 \\ 0 & 0 \end{pmatrix}$. Ainsi, $AB \neq BA$. De plus, on observe que A et B ne sont pas des matrices nulles mais $AB = 0_{M_4(\mathbb{R})}$.

Proposition 5.7 *Si $A \in M_n(\mathbb{K})$ et $B \in M_n(\mathbb{K})$ commutent (c'est-à-dire si $AB = BA$), on dispose de la formule du binôme de Newton [2] :*

$$\forall m \geq 0,\ (A+B)^m = \sum_{k=0}^{m} C_m^k A^k B^{m-k} = \sum_{k=0}^{m} C_m^k A^{m-k} B^k.$$

En particulier, $(I_n + B)^m = \sum_{k=0}^{m} C_m^k B^k$.

On rappelle que $C_m^k = \dfrac{m!}{k!(m-k)!}$. Comme dans $\mathbb{K}$, la démonstration de cette proposition se fait par récurrence sur l'entier m.

Le groupe $GL_n(\mathbb{K})$

Définition 5.8 *Une matrice A de $M_n(\mathbb{K})$ est dite* ***inversible*** *si et seulement s'il existe $A' \in M_n(\mathbb{K})$ telle que $AA' = A'A = I_n$. Si A est inversible, alors A' est unique, appelée* ***inverse*** *de A et notée A^{-1}.*
On note $GL_n(\mathbb{K})$ l'ensemble des matrices inversibles de $M_n(\mathbb{K})$.

Proposition et définition 5.9 *1) La multiplication des matrices est interne dans $GL_n(\mathbb{K})$, qui, muni de cette loi de composition, est un groupe appelé* ***groupe linéaire****.*

2) Pour tout $\mathbb{K}$-espace vectoriel E muni d'une base $\mathscr{B}$, l'application $f \mapsto \mathrm{mat}_{\mathscr{B}}(f)$ définit un isomorphisme entre les groupes $(\mathscr{GL}(E), \circ)$ et $(GL_n(\mathbb{K}), \cdot)$.

DÉMONSTRATION.

1) Tout d'abord, $I_n \in GL_n(\mathbb{K})$. De plus, pour tout $(A,B) \in GL_n(\mathbb{K})^2$, $(AB)(B^{-1}A^{-1}) = (B^{-1}A^{-1})(AB) = I_n$, donc $AB \in GL_n(\mathbb{K})$, et $A^{-1}A = AA^{-1} = I_n$, donc $A^{-1} \in GL_n(\mathbb{K})$.
2) Pour tout $(f,g) \in \mathscr{GL}(E)^2$, $\mathrm{mat}_{\mathscr{B}}(f \circ g) = \mathrm{mat}_{\mathscr{B}}(f)\,\mathrm{mat}_{\mathscr{B}}(g)$.
Par ailleurs, pour toute application f de $\mathscr{GL}(E)$, comme

$$\mathrm{mat}_{\mathscr{B}}(f)\,\mathrm{mat}_{\mathscr{B}}(f^{-1}) = \mathrm{mat}_{\mathscr{B}}(f^{-1})\,\mathrm{mat}_{\mathscr{B}}(f) = \mathrm{mat}_{\mathscr{B}}(I_E) = I_n,$$

2. Sir Isaac Newton (4 janvier 1643 - 31 mars 1727) était un philosophe, mathématicien, physicien et astronome anglais. Figure emblématique des sciences, il est surtout reconnu pour sa théorie de la gravitation universelle et l'invention du calcul infinitésimal.

on a $\mathrm{mat}_{\mathscr{B}}(f) \in GL_n(\mathbb{K})$.
Réciproquement, pour toute matrice A de $GL_n(\mathbb{K})$, il existe $(f, g) \in \mathscr{L}(E)^2$ unique tel que $\mathrm{mat}_{\mathscr{B}}(f) = A$ et $\mathrm{mat}_{\mathscr{B}}(g) = A^{-1}$, et on a : $\mathrm{mat}_{\mathscr{B}}(f \circ g) = \mathrm{mat}_{\mathscr{B}}(f)\,\mathrm{mat}_{\mathscr{B}}(g) = AA^{-1} = I_n$ et $\mathrm{mat}_{\mathscr{B}}(g \circ f) = \mathrm{mat}_{\mathscr{B}}(g)\,\mathrm{mat}_{\mathscr{B}}(f) = A^{-1}A = I_n$, donc $f \circ g = g \circ f = I_E$, d'où $f \in \mathscr{GL}(E)$.

□

On déduit de cette proposition le résultat suivant.

Théorème 5.10 *Une matrice A de $M_n(\mathbb{K})$ est inversible si et seulement si ses colonnes, considérées comme des éléments de $\mathbb{K}^n$, sont linéairement indépendantes.*

DÉMONSTRATION. D'après la proposition 5.9, A inversible signifie que c'est la matrice d'un isomorphisme f sur E. Par injectivité de f, on a directement l'indépendance linéaire des vecteurs $f(e_1),\dots, f(e_n)$. □

Nous terminons en énonçant quelques propriétés vérifiées par les éléments de $GL_n(\mathbb{K})$.
- Pour toutes matrices A et B de $GL_n(\mathbb{K})$, on a : $(AB)^{-1} = B^{-1}A^{-1}$.
- Si $A \in GL_n(\mathbb{K})$, alors ${}^tA \in GL_n(\mathbb{K})$ et $({}^tA)^{-1} = {}^t(A^{-1})$, puisque ${}^t(A^{-1}){}^tA = {}^t(AA^{-1}) = I_n$.

Par ailleurs, si $A \in M_n(\mathbb{K})$ est inversible à droite (resp. à gauche), *i.e.* s'il existe $B \in M_n(\mathbb{K})$ telle que $AB = I_n$ (resp. $BA = I_n$), alors A est inversible.

Calcul pratique de l'inverse d'une matrice

Soit $A \in GL_n(\mathbb{K})$. En notant $AX = Y$, où X, $Y \in M_{n,1}(\mathbb{K})$, il suffit d'exprimer X en fonction de Y par résolution d'un système linéaire pour obtenir la matrice A^{-1}. En effet, si A est inversible, on a $AX = Y \Leftrightarrow X = A^{-1}Y$.

Exemple. Calculons l'inverse $A = \begin{pmatrix} 2 & 4 & -1 \\ 1 & -2 & 3 \\ 1 & 1 & 1/2 \end{pmatrix}$. En notant $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ et $Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$, nous avons

$$AX = Y \Leftrightarrow \left\{ \begin{array}{rcl} 2x_1 + 4x_2 - x_3 & = & y_1 \\ x_1 - 2x_2 + 3x_3 & = & y_2 \\ x_1 + x_2 + \dfrac{x_3}{2} & = & y_3 \end{array} \right. .$$

En résolvant ce système, nous trouvons

$$\left\{ \begin{array}{rcl} x_1 & = & 4y_1 + 3y_2 - 10y_3 \\ x_2 & = & -\dfrac{5y_1}{2} - 2y_2 + 7y_3 \\ x_3 & = & -3y_1 - 2y_2 + 8y_3 \end{array} \right. ,$$

d'où $A^{-1} = \begin{pmatrix} 4 & 3 & -10 \\ -5/2 & -2 & 7 \\ -3 & -2 & 8 \end{pmatrix}$.

5.1.6 Matrices écrites par blocs

Supposons que $n = n_1 + \cdots + n_r$ et $p = p_1 + \cdots + p_s$, avec $(r,s) \in \mathbb{N}^* \times \mathbb{N}^*$, et que, pour tout $(i, j) \in \{1,\dots,r\} \times \{1,\dots,s\}$, $(n_i, p_i) \in \mathbb{N}^* \times \mathbb{N}^*$. On définit la *matrice écrite par blocs* $A \in M_{n,p}(\mathbb{K})$ de la manière suivante :

$$A = (A_{ij})_{1\le i\le r,\ 1\le j\le s} = \begin{pmatrix} A_{11} & \cdots & A_{1s} \\ \vdots & A_{ij} & \vdots \\ A_{r1} & \cdots & A_{rs} \end{pmatrix}, \text{ où } A_{ij} \in M_{n_i,p_j}(\mathbb{K}).$$

5.1.7 Rang d'une matrice

On rappelle que si $\mathscr{F}$ est une famille d'éléments de F, alors $\mathrm{rang}\,\mathscr{F} = \dim(\mathrm{Vect}\,\mathscr{F})$, et que si $f : E \to F$ est une application linéaire, alors $\mathrm{rang}\, f = \dim(\mathrm{Im}\, f) = \mathrm{rang}\, f(\mathscr{B})$.

Définition 5.11 *Soit $A \in M_{n,p}(\mathbb{K})$. On appelle **rang** de A, et l'on note* $\mathrm{rang}A$*, le rang de la famille des colonnes de A dans $M_{n,1}(\mathbb{K})$.*

Ainsi, en notant $A=(a_{ij})_{1\leq i\leq n,\ 1\leq j\leq p}$ et $C_1=\begin{pmatrix}a_{11}\\ \vdots\\ a_{n1}\end{pmatrix},\dots,C_p=\begin{pmatrix}a_{1p}\\ \vdots\\ a_{np}\end{pmatrix}$ les colonnes de A, on a que $\operatorname{rang}A=\operatorname{rang}(C_1,\dots,C_p)$.
Cette définition est compatible avec l'identification usuelle entre $M_{n,p}(\mathbb{K})$ et $\mathscr{L}(\mathbb{K}^p,\mathbb{K}^n)$, le rang d'une matrice A étant dans ce cadre le rang de n'importe quelle application linéaire représentée par A et le rang d'une application linéaire étant le rang de n'importe quelle matrice la représentant. Ceci est résumé dans la

Proposition 5.12 *Soient $f\in\mathscr{L}(E,F)$ et $A=\operatorname{mat}_{\mathscr{B},\mathscr{C}}(f)$, on a*

$$\operatorname{rang}f=\operatorname{rang}A.$$

Proposition 5.13 *Pour toute matrice A de $M_{n,p}(\mathbb{K})$, on a*

$$\operatorname{rang}A\leq\min(n,p).$$

DÉMONSTRATION. On a d'une part que $\operatorname{rang}A=\operatorname{rang}(C_1,\dots,C_p)\leq p$, et d'autre part que $\operatorname{rang}A=\dim(\operatorname{Vect}(C_1,\dots,C_p))\leq\dim M_{n,1}(\mathbb{K})=n$. □

Proposition 5.14 *Pour toute matrice A de $M_n(\mathbb{K})$, $\operatorname{rang}A=n\Leftrightarrow A\in GL_n(\mathbb{K})$.*

DÉMONSTRATION. Soit f l'endomorphisme de $\mathscr{L}(\mathbb{R}^n)=M_{n,1}(\mathbb{K})$ représenté par A dans la base canonique de $M_{n,1}(\mathbb{K})$. Comme $(C_1,\dots,C_n)$ est une base de $M_{n,1}(\mathbb{K})$ si et seulement si f est bijectif, on en conclut que $\operatorname{rang}A=n\Leftrightarrow A\in GL_n(\mathbb{K})$. □

Nous montrons enfin qu'on ne modifie pas le rang d'une matrice en la multipliant par une matrice inversible.

Proposition 5.15 *Pour toute matrice A de $M_{n,p}(\mathbb{K})$, on a*

$$\forall P\in GL_p(\mathbb{K})\ (\textit{resp. } Q\in GL_n(\mathbb{K})),\ \operatorname{rang}(AP)=\operatorname{rang}A\ (\textit{resp. } \operatorname{rang}(QA)=\operatorname{rang}A).$$

Ce résultat se déduit aisément de la

Proposition 5.16 *Pour toutes applications linéaires $f:E\to F$ et $g:F\to G$, si f (resp. g) est un isomorphisme, alors $\operatorname{rang}g\circ f=\operatorname{rang}g$ (resp. $\operatorname{rang}g\circ f=\operatorname{rang}f$).*

DÉMONSTRATION. Tout d'abord, $\operatorname{Im}(g\circ f)=g(\operatorname{Im}f)\subset g(F)$, et donc $\operatorname{rang}(g\circ f)\leq\operatorname{rang}g$, avec égalité si f est surjective. De plus, $\operatorname{rang}(g\circ f)\leq\dim(\operatorname{Im}f)=\operatorname{rang}f$, avec égalité si g est injective. □

5.2 Changement de bases

Soient respectivement $\mathscr{B}'=(e'_1,\dots,e'_p)$ et $\mathscr{C}'=(f'_1,\dots,f'_p)$ d'autres bases de E et F.

Définition 5.17 *On appelle* ***matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$*** *la matrice $P_{\mathscr{B}}^{\mathscr{B}'}=\operatorname{mat}_{\mathscr{B}}(\mathscr{B}')$ de $M_p(\mathbb{K})$, dont les colonnes sont formées par les composantes des vecteurs de la base $\mathscr{B}'$ dans la base $\mathscr{B}$.*

Proposition 5.18 *On a $P_{\mathscr{B}}^{\mathscr{B}'}=\operatorname{mat}_{\mathscr{B}',\mathscr{B}}(I_E)$.*

DÉMONSTRATION. Notons $\mathscr{B}'=(e'_1,\dots,e'_p)$. Pour chaque j de $\{1,\dots,p\}$, la $j^{\text{ème}}$ colonne de $\operatorname{mat}_{\mathscr{B}',\mathscr{B}}(I_E)$ est formée par les composantes de $I_E(e'_j)$, soit e'_j, sur la base $\mathscr{B}$. □

Proposition 5.19 *Soient $\mathscr{B}$, $\mathscr{B}'$ et $\mathscr{B}''$ des bases d'un $\mathbb{K}$-espace vectoriel E de dimension p. On a*

1) $P_{\mathscr{B}}^{\mathscr{B}}=I_p$.

2) $P_{\mathscr{B}}^{\mathscr{B}''}=P_{\mathscr{B}}^{\mathscr{B}'}P_{\mathscr{B}'}^{\mathscr{B}''}$.

3) $P_{\mathscr{B}}^{\mathscr{B}'}$ est inversible et $(P_{\mathscr{B}}^{\mathscr{B}'})^{-1}=P_{\mathscr{B}'}^{\mathscr{B}}$.

DÉMONSTRATION.

1) $P_{\mathscr{B}}^{\mathscr{B}}=\operatorname{mat}_{\mathscr{B},\mathscr{B}}(I_E)=I_p$.

2) $P_{\mathscr{B}}^{\mathscr{B}''}=\operatorname{mat}_{\mathscr{B}'',\mathscr{B}}(I_E)=\operatorname{mat}_{\mathscr{B}',\mathscr{B}}(I_E)\operatorname{mat}_{\mathscr{B}'',\mathscr{B}'}(I_E)=P_{\mathscr{B}}^{\mathscr{B}'}P_{\mathscr{B}'}^{\mathscr{B}''}$.

3) $P_{\mathscr{B}}^{\mathscr{B}'} P_{\mathscr{B}'}^{\mathscr{B}} = P_{\mathscr{B}}^{\mathscr{B}} = I_p$.

□

On note à présent $P = P_{\mathscr{B}}^{\mathscr{B}'}$ (resp. $Q = P_{\mathscr{C}}^{\mathscr{C}'}$) la matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$ (resp. de $\mathscr{C}$ à $\mathscr{C}'$). Les formules dites *de changement de base* sont données par les théorèmes suivants.

Théorème 5.20 ***(changement de base pour un vecteur)*** *Pour tout $x \in E$, si $X = \mathrm{mat}_{\mathscr{B}}(x)$ et $X' = \mathrm{mat}_{\mathscr{B}'}(x)$, alors $X = PX'$.*

DÉMONSTRATION. $X = \mathrm{mat}_{\mathscr{B}}(x) = \mathrm{mat}_{\mathscr{B}',\mathscr{B}}(I_E)\,\mathrm{mat}_{\mathscr{B}'}(x) = PX'$. □

Théorème 5.21 ***(changement de base pour une application linéaire)*** *Soit $f : E \to F$ une application linéaire. Si $A = \mathrm{mat}_{\mathscr{B},\mathscr{C}}(f)$ et $A' = \mathrm{mat}_{\mathscr{B}',\mathscr{C}'}(f)$, alors $A = QAP^{-1}$. Dans le cas particulier où f est un endomorphisme, c'est-à-dire lorsque $E = F$, $\mathscr{B} = \mathscr{C}$ et $\mathscr{B}' = \mathscr{C}'$, on a $A = PA'P^{-1}$.*

DÉMONSTRATION. On a $A = \mathrm{mat}_{\mathscr{B},\mathscr{C}}(f) = \mathrm{mat}_{\mathscr{B}',\mathscr{C}'}(I_F \circ f \circ I_E) = \mathrm{mat}_{\mathscr{C},\mathscr{C}'}(I_F)\,\mathrm{mat}_{\mathscr{B}',\mathscr{C}'}(f)\,\mathrm{mat}_{\mathscr{B},\mathscr{B}'}(I_E) = QA'P$. □

Dans les conditions du théorème précédent, A et A' sont dites *équivalentes* dans $M_{n,p}(\mathbb{K})$. Dans le cas où f est un endomorphisme, on dit que ces matrices sont *semblables* dans $M_n(\mathbb{K})$.

Remarque. La proposition 5.15 nous permet directement d'affirmer que deux matrices équivalentes ont même rang.

Exemple. Dans l'espace vectoriel $E = \mathbb{R}^3$, on note $\mathscr{B} = (e_1, e_2, e_3)$ la base canonique, et on introduit les vecteurs $e_1' = e_2 = (0,1,0)$, $e_2' = e_2 + e_3 = (0,1,1)$ et $e_3' = e_1 = (1,0,0)$. On vérifie facilement que $\mathscr{B}' = (e_1', e_2', e_3')$ est une base et que les matrices de passage sont

$$P = P_{\mathscr{B}}^{\mathscr{B}'} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix} \text{ et } P_{\mathscr{B}'}^{\mathscr{B}} = P^{-1} = \begin{pmatrix} 0 & 1 & -1 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

Si $x \in E$, $X = \mathrm{mat}_{\mathscr{B}}(x) \in M_{3,1}(\mathbb{R})$ et $X' = mat_{\mathscr{B}'}(x) \in M_{3,1}(\mathbb{R})$, alors les formules de changement de coordonnées sont données par $X = PX'$ et $X' = P^{-1}X$. Si f est un endomorphisme de E tel que

$$A = \mathrm{mat}_{\mathscr{B}}(f) = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix},$$

alors la matrice A' de f dans la base $\mathscr{B}'$ est

$$A' = \mathrm{mat}_{\mathscr{B}'}(f) = P^{-1}AP = \begin{pmatrix} -1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 0 \end{pmatrix}.$$

Chapitre 6

Systèmes d'équations linéaires

Soit $\mathbb{K}$ un corps commutatif. Nous allons chercher dans ce chapitre à résoudre un système de n équations linéaires à p inconnues, de la forme

$$\left\{\begin{array}{ccccccc} a_{11}x_1 & + & \dots & + & a_{1p}x_p & = & b_1 \\ \vdots & & & & \vdots & & \vdots \\ a_{n1}x_1 & + & \dots & + & a_{np}x_p & = & b_n \end{array}\right. , \tag{6.1}$$

appelé *système d'équations linéaires* ou, plus simplement, *système linéaire*. Les scalaires $(a_{ij})_{1\leq i\leq p,\ 1\leq j\leq n}$ et $(b_i)_{1\leq i\leq n}$, nommés respectivement les *coefficients* et *seconds membres du système*, sont donnés, alors que les x_j, $1 \leq j \leq p$, sont les *inconnues du système*. On appelle *solution du système* tout p-uplet $(x_1, \dots, x_p) \in \mathbb{K}^p$ tel que les égalités de (6.1) sont toutes vérifiées. Résoudre un tel système, c'est déterminer l'ensemble des solutions, qui peut éventuellement être l'ensemble vide.
Enfin, le système linéaire est dit *homogène* lorsque $b_1 = \dots = b_n = 0$. Dans ce cas, il possède au moins $(0, \dots, 0) \in \mathbb{K}^p$ comme solution, appelé *solution triviale* du système.

6.1 Différentes interprétations d'un système d'équations linéaires

Un système d'équations linéaires et, par suite, sa résolution peuvent être vus de différentes manières, que nous allons expliciter ici.

6.1.1 Interprétation géométrique

Supposons momentanément que $\mathbb{K} = \mathbb{R}$ et considérons le système (6.1) de n équations à p inconnues. Si

— $n = 1$ et $p = 2$, le système se résume à une seule équation de la forme

$$ax_1 + bx_2 = c,$$

qui est l'équation d'une droite affine dans l'espace $\mathbb{R}^2$ si a ou b est non nul.

— $n = 1$ et $p = 3$, le système se réduit alors à une seule équation de la forme

$$ax_1 + bx_2 + cx_3 = d,$$

qui est l'équation d'un plan affine dans l'espace $\mathbb{R}^3$ si a, b ou c est non nul.

— $n = 2$ et $p = 3$, le système équivaut au système suivant

$$\left\{\begin{array}{l} ax_1 + bx_2 + cx_2 = d \\ a'x_1 + b'x_2 + c'x_2 = d' \end{array}\right. ,$$

dont l'ensemble des solutions est l'intersection de deux plans affines P et P' de $\mathbb{R}^3$, définis chacun par l'une des deux équations ci-dessus, c'est-à-dire

(i) l'ensemble vide si les plans P et P' sont parallèles,

(ii) le plan P si chacune des équations décrit ce même plan,

(iii) la droite $d = P \cap P'$ dans les autres cas.

Revenons à présent au problème concernant le système initial (6.1). Une équation de la forme

$$a_{i1}x_1 + \cdots + a_{ip}x_p = b_i, \ \forall i \in \{1,\ldots,n\},$$

définit de la même manière un hyperplan affine de l'espace vectoriel $\mathbb{K}^p$. La solution du système (6.1) est alors l'*intersection de n hyperplans affines de $\mathbb{K}^p$*.

6.1.2 Interprétation matricielle

Il est possible d'écrire le système (6.1) sous une *forme matricielle*. Pour cela, notons $A = (a_{ij})_{1\leq i\leq p,\ 1\leq j\leq n}$ la *matrice du système* appartenant à $M_{n,p}(\mathbb{K})$, $B = (b_i)_{1\leq i\leq n}$ et $X = (x_j)_{1\leq j\leq p}$ les matrices unicolonnes appartenant respectivement à $M_{n,1}(\mathbb{K})$ et $M_{p,1}(\mathbb{K})$. D'après la définition du produit matriciel (cf. définition 5.6), le système (6.1) équivaut à

$$AX = B. \tag{6.2}$$

La résolution du système revient alors à la résolution de cette équation matricielle : les matrices $A \in M_{n,p}(\mathbb{K})$ et $B \in M_{n,1}(\mathbb{K})$ étant données, il s'agit de déterminer l'ensemble des matrices $X \in M_{p,1}(\mathbb{K})$ vérifiant (6.2).
Si A est une matrice carrée inversible, le système est dit *de Cramer* [1] et admet une unique solution $X = A^{-1}B$.

6.1.3 Interprétation en termes de combinaisons linéaires

Soient $A \in M_{n,p}(\mathbb{K})$ la matrice du système et $B \in M_{n,1}(\mathbb{K})$ la matrice-colonne des seconds membres associées à (6.1) et définies dans la sous-section 6.1.2. Notons C_j, pour tout j de $\{1,\ldots,p\}$, la j^{e} colonne de A. Le système linéaire (6.1) s'écrit alors

$$\sum_{j=1}^{p} x_j C_j = B.$$

Résoudre le système revient alors à trouver toutes les manières d'écrire B comme combinaison linéaire (dans $\mathbb{K}^n$) de C_1, C_2,..., C_p ; il existe au moins une solution si et seulement si B appartient au sous-espace vectoriel $\mathrm{Vect}(C_1,\ldots,C_p)$. Si le système est homogène, on doit trouver le p-uplet $(x_1,\ldots,x_p) \in \mathbb{K}^p$ tel que

$$\sum_{j=1}^{p} x_j C_j = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Le système admet dans ce cas d'autres solutions que la solution triviale si et seulement si la famille des matrices-colonnes $(C_1,\ldots,C_p)$ est liée.

6.1.4 Interprétation en termes d'une application linéaire

Introduisons les espaces vectoriels E et F, de dimensions respectives p et n, et respectivement munis de la base $\mathscr{B} = (e_1,\ldots,e_p)$ et $\mathscr{C} = (f_1,\ldots,f_n)$. Considérons à présent $f : E \to F$ l'application linéaire de matrice $A = \mathrm{mat}_{\mathscr{B},\mathscr{C}}(f)$ respectivement aux bases $\mathscr{B}$ et $\mathscr{C}$. Si l'on désigne par x le vecteur de E qui a pour coordonnées dans la base $\mathscr{B}$ les inconnues x_1,..., x_p et par b le vecteur de F qui a pour coordonnées dans la base $\mathscr{C}$ les seconds membres b_1,..., b_n, (6.1) est alors équivalent à

$$f(x) = b. \tag{6.3}$$

1. Gabriel Cramer (31 juillet 1704 - 4 janvier 1752) était un mathématicien suisse. Le travail par lequel il est le mieux connu est son traité, publié en 1750, d'*Introduction à l'analyse des lignes courbes algébriques* dans lequel il démontra qu'une courbe algébrique de degré n est déterminée par $\frac{n(n+3)}{2}$ de ses points en position générale.

Résoudre le système revient alors à déterminer l'image réciproque du singleton $\{b\}$ par l'application f. Dans cette interprétation, il est aisé de discuter les situations envisageables et de déterminer la structure de l'ensemble des solutions.
Tout d'abord on voit que, lorsqu'il est homogène, le système linéaire équivaut à l'équation

$$f(x) = 0_F,$$

qui possède toujours au moins une solution (la solution triviale $x = 0_E$). Résoudre le système revient à déterminer le noyau de l'application f. L'ensemble des solutions du système est alors un $\mathbb{K}$-espace vectoriel de dimension $p - \operatorname{rang} f$ (d'après le théorème du rang).
Si le second membre du système est non nul, deux cas de figure se présentent :

i) Supposons que b n'appartient pas à l'image de E par f : $b \notin \operatorname{Im} f$. Alors, l'équation (6.3) n'a pas de solution et l'ensemble des solutions est donc vide.

ii) Supposons que b appartient à l'image de E par f : $b \in \operatorname{Im} f$. Il existe alors au moins un vecteur x' de E tel que $f(x') = b$. Pour toute solution x de (6.3), on aura

$$f(x) - f(x') = 0_F,$$

soit encore, par linéarité de f,

$$f(x - x') = 0_F.$$

Par conséquent $x - x' \in \operatorname{Ker} f$.
Réciproquement, tout x tel que $x = x' + y$, avec $y \in \operatorname{Ker} f$, est une solution de (6.3). En effet, on a

$$f(x) = f(x') + f(y) = f(x') = b.$$

Par conséquent, si x' est une solution de (6.3), l'ensemble des solutions est

$$\{x'\} + \operatorname{Ker} f.$$

L'ensemble des solutions du système s'obtient donc en ajoutant à une solution particulière du système l'ensemble des solutions du système homogène associé.

En résumé, le système linéaire n'a pas de solution si $b \notin \operatorname{Im} f$ et au moins une solution si $b \in \operatorname{Im} f$.
Amenons maintenant à son terme la discussion entamée en étudiant l'application f. Si f est surjective (ce qui exige $n \leq p$, cf. proposition 2.50), alors $\operatorname{Im} f = F$. Par conséquent, tout vecteur b de F appartient à $\operatorname{Im} f$ et le système admet toujours des solutions. Si f est injective (on a forcément $p \leq n$), alors $\dim(\operatorname{Im} f) = \dim E = p$ d'après le théorème du rang. Le système admet alors au plus une solution, puisque $\operatorname{Ker} f = \{0_E\}$. Enfin, si f est bijective (ce qui exige $p = n$), alors le système admet une solution (puisque f est surjective) et une seule (puisque f est injective). Dans ce dernier cas, on dit que le système est *de Cramer*.

6.2 Méthode d'élimination de Gauss

La méthode d'élimination de Gauss [2] permet entre autres choses de résoudre des systèmes d'équations linéaires de la forme (6.1) de manière systématique. Elle se sert de l'interprétation matricielle que l'on peut faire de ces systèmes et utilise plusieurs résultats de calcul matriciel. Elle se base en particulier sur la propriété suivante : pour toute matrice $P \in GL_n(\mathbb{K})$, les systèmes $AX = B$ et $PAX = PB$ ont les mêmes solutions, on dit alors qu'ils sont *équivalents*. L'idée est donc de chercher $P \in GL_n(\mathbb{K})$ de sorte que que le système linéaire $PAX = PB$ soit facile à résoudre. Par exemple, un système à matrice carrée triangulaire supérieure se résout très facilement par la méthode dite *de remontée* : on calcule x_p, puis x_{p-1}, etc, jusqu'à x_1.

Nous allons maintenant décrire les *opérations élémentaires* que l'on peut effectuer sur les lignes (resp. colonnes) d'une matrice A de $M_{n,p}(\mathbb{K})$. Elles conduiront toujours à une matrice $A' = PA$ (resp. $A' = AQ$), où $P \in GL_n$ (resp. $Q \in GL_p(\mathbb{K})$) est obtenue en effectuant l'opération correspondante sur la matrice I_n (resp. I_p).

2. Johann Carl Friedrich Gauß (30 avril 1777 - 23 février 1855) était un mathématicien, astronome et physicien allemand. Surnommé par ses pairs « *le prince des mathématiciens* », il fit des contributions significatives dans de nombreux domaines des sciences de son époque, notamment en théorie des nombres, en statistique, en analyse, en géométrie différentielle, en électrostatique, en astronomie et en optique.

6.2.1 Opérations élémentaires

Soient $(n,p) \in (\mathbb{N}\backslash\{0,1\})^2$ et $A = (a_{ij})_{1\le i\le n,\ 1\le j\le p} \in M_{n,p}(\mathbb{K})$. On appelle *opérations élémentaires sur les lignes* de A les transformations suivantes (où L_i, $1 \le i \le n$, et L_j, $1 \le j \le p$, désignent respectivement les i^e et j^e lignes de A) :

— échange (entre elles) de deux lignes L_i et L_j ($L_i \leftrightarrow L_j$),
— multiplication de la ligne L_i par le scalaire $\lambda \in \mathbb{K}\backslash\{0\}$ ($L_i \leftarrow \lambda L_i$),
— remplacement d'une ligne L_i par $L_i + \lambda L_j$, où $\lambda \in \mathbb{K}\backslash\{0\}$ et $j \in \{1,\dots,n\}\backslash\{i\}$, ($L_i \leftarrow L_i + \lambda L_j$).

On notera par ailleurs $L_i \leftarrow \lambda L_i + \mu L_j$ la composée des opérations élémentaires $L_i \leftarrow \lambda L_i$ et $L_i \leftarrow L_i + \mu L_j$, où λ et μ sont des scalaires non nuls.
Explicitons à présent les matrices carrées inversibles P correspondant à chacunes de ces opérations élémentaires.

— Dans le cas de l'échange des i^e et j^e lignes de la matrice, $(i,j) \in \{1\dots,n\}^2$ et $i < j$, nous avons

$$P = P_{ij} = \begin{pmatrix} 1 & 0 & & & \cdots & & & & 0 \\ 0 & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & 0 & 0 & \cdots & 0 & 1 & & & \\ & & & 0 & 1 & & & 0 & & & \\ \vdots & & & \vdots & & \ddots & & \vdots & & & \vdots \\ & & & 0 & & & 1 & 0 & & & \\ & & & 1 & 0 & \cdots & 0 & 0 & & & \\ & & & & & & & & 1 & & \\ & & & & & & & & & \ddots & 0 \\ 0 & & & & & \cdots & & & & 0 & 1 \end{pmatrix} = I_n + (E_{ij} + E_{ji} - E_{ii} - E_{jj}) \in M_n(\mathbb{K}).$$

Cette matrice est bien inversible puisqu'elle vérifie $P_{ij}^2 = I_n$.

— Dans le cas de la multiplication de la i^e ligne, $1 \le i \le n$, de la matrice par un scalaire non nul λ, nous avons

$$P = D_i(\lambda) = \begin{pmatrix} 1 & 0 & & \cdots & & & 0 \\ 0 & \ddots & & & & & \\ & & 1 & & & & \\ \vdots & & & \lambda & & & \vdots \\ & & & & 1 & & \\ & & & & & \ddots & 0 \\ 0 & & & \cdots & & 0 & 1 \end{pmatrix} = I_n + (\lambda - 1)E_{ii} \in M_n(\mathbb{K}).$$

Cette matrice est bien inversible, car $D_i(\lambda)D_i\left(\frac{1}{\lambda}\right) = I_n$.

— Enfin, dans le cas du remplacement de la i^e ligne, $1 \le i \le n$, de la matrice par $L_i + \lambda L_j$, avec $1 \le j \le n$, $j \ne i$, et $\lambda \ne 0$, nous avons

$$P = P_{ij} = \begin{pmatrix} 1 & 0 & & \cdots & & & 0 \\ 0 & \ddots & & & & & \\ & & 1 & \cdots & \lambda & & \\ & & \vdots & \ddots & \vdots & & \\ & & 0 & \cdots & 1 & & \\ & & & & & \ddots & 0 \\ 0 & & & \cdots & & 0 & 1 \end{pmatrix} = I_n + \lambda E_{ij} \in M_n(\mathbb{K}).$$

On définit de manière analogue les opérations élémentaires, ainsi que les matrices leur correspondant, sur les colonnes de A (qui sont les opérations élémentaires sur les lignes de la transposée de A).

6.2.2 Principe et mise en œuvre de la méthode

Par une suite d'opérations élémentaires, nous allons transformer le système linéaire (6.1) en un système équivalent dont la matrice sera triangulaire supérieure (si c'est une matrice carrée) ou *échelonnée supérieurement*[3]. La résolution « en cascades » de ce dernier système donne alors les solutions de (6.1).
Nous procédons comme suit. Considérons le système

$$\left\{\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1p}x_p & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \dots & + & a_{2p}x_p & = & b_2 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{n1}x_1 & + & a_{n2}x_2 & + & \dots & + & a_{np}x_p & = & b_n \end{array}\right.$$

et supposons dans un premier temps que le coefficient a_{11}, appelé le *pivot*, est non nul. Nous effectuons ensuite les opérations élémentaires

$$L_2 \leftarrow L_2 - \frac{a_{21}}{a_{11}}L_1,\ L_3 \leftarrow L_3 - \frac{a_{31}}{a_{11}}L_1, \dots,\ L_n \leftarrow L_n - \frac{a_{n1}}{a_{11}}L_1$$

qui ont pour effet l'annulation des coefficients de l'inconnue x_1 dans les équations L_2 à L_n. Nous obtenons le système équivalent

$$\left\{\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1p}x_p & = & b_1 \\ & & a'_{22}x_2 & + & \dots & + & a'_{2p}x_p & = & b'_2 \\ & & \vdots & & & & \vdots & & \vdots \\ & & a'_{n2}x_2 & + & \dots & + & a'_{np}x_p & = & b'_n \end{array}\right.$$

où $a'_{ij} = a_{ij} - \dfrac{a_{i1}}{a_{11}}a_{1j}$ pour tout $(i,j) \in \{2,\dots,n\} \times \{2,\dots,p\}$ et $b'_i = b_i - \dfrac{a_{i1}}{a_{11}}b_1$ pour tout $i \in \{2,\dots,n\}$.
Supposons à présent que le pivot a'_{22} est non nul. Les opérations élémentaires

$$L_3 \leftarrow L_3 - \frac{a'_{32}}{a'_{22}}L_2, \dots,\ L_n \leftarrow L_n - \frac{a'_{n2}}{a'_{22}}L_2$$

conduisent alors au système

$$\left\{\begin{array}{ccccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & a_{13}x_3 & + & \dots & + & a_{1p}x_p & = & b_1 \\ & & a'_{22}x_2 & + & a'_{23}x_3 & + & \dots & + & a'_{2p}x_p & = & b'_2 \\ & & & & a''_{33}x_3 & + & \dots & + & a''_{3p}x_p & = & b''_3 \\ & & & & \vdots & & & & \vdots & & \vdots \\ & & & & a''_{n3}x_3 & + & \dots & + & a''_{np}x_p & = & b''_n \end{array}\right. ,$$

avec $a''_{ij} = a'_{ij} - \dfrac{a'_{i2}}{a'_{22}}a'_{2j}$ pour tout $(i,j) \in \{3,\dots,n\} \times \{3,\dots,p\}$ et $b''_i = b'_i - \dfrac{a'_{i2}}{a'_{22}}b'_2$ pour tout $i \in \{3,\dots,n\}$. On poursuit ainsi la mise sous forme échelonnée de la matrice du système.
Il est possible qu'à un moment donné le pivot soit nul, il suffit alors d'échanger l'équation concernée par l'une des équations suivantes pour laquelle le pivot est non nul. Si tous les pivots « potentiels » sont nuls, il faut alors procéder à une permutation de colonnes, en n'oubliant pas que cela implique un changement dans l'ordre de numérotation des inconnues.
Le système équivalent final $A'X = B'$, obtenu en un nombre d'étapes r, où r correspond au nombre de pivots non

3. Une matrice A de $M_{n,p}(\mathbb{K})$ est dite échelonnée supérieurement s'il existe un entier r de $\{0,\dots,n\}$ tel que :
— pour tout indice $i \leq r$, la ligne L_i est non nulle,
— pour tout indice $i \geq r$, la ligne L_i est nulle,
— pour chaque ligne L_i, soit $d(i)$ le plus petit indice j tel que $a_{ij} \neq 0$; la suite $d(1),\dots, d(r)$ est strictement croissante.
Une telle matrice est de rang r.

nuls et est au plus égal à $\min(n,p)$, se présente sous la forme

$$\left\{\begin{array}{ccccccccccccc}
a'_{11}x_1 & + & a'_{12}x_2 & + & \dots & + & a'_{1r}x_r & + & \dots & + & a'_{1p}x_p & = & b'_1 \\
 & & a'_{22}x_2 & + & \dots & + & a'_{2r}x_r & + & \dots & + & a'_{2p}x_p & = & b'_2 \\
 & & & & \ddots & & \vdots & & & & \vdots & & \vdots \\
 & & & & & & a'_{rr}x_r & + & \dots & + & a'_{rp}x_p & = & b'_r \\
 & & & & & & & & & & 0 & = & b'_{r+1} \\
 & & & & & & & & & & \vdots & & \vdots \\
 & & & & & & & & & & 0 & = & b'_n
\end{array}\right. .$$

Si $r \geq 1$, on appelle $x_1,\dots, x_r$ les *inconnues principales* et, si $r+1 \leq p$, $x_{r+1},\dots, x_p$ sont les *inconnues secondaires*. Le nombre d'inconnues principales est le *rang* du système. Plusieurs cas de figure se présentent alors.
Si $r+1 \leq n$, le système admet des solutions si et seulement si $b'_i = 0$ pour tout i de $\{r+1,\dots,n\}$. Ces $n-r$ équations s'appellent les *conditions de compatibilité* (on conviendra qu'elles sont satisfaites lorsque $r = n$, c'est-à-dire si A représente est la matrice d'une application surjective). Dans ce cas, si $p > r$, l'ensemble des solutions est paramétré par les $p-r$ inconnues non principales et obtenu en résolvant par une méthode de remontée le système de Cramer suivant

$$\left\{\begin{array}{ccccccccccccc}
a'_{11}x_1 & + & \dots & + & a'_{1r}x_r & = & b'_1 & - & a'_{1r+1}x_{r+1} & - & \dots & - & a'_{1p}x_p \\
 & & \ddots & & \vdots & & \vdots & & \vdots & & & & \vdots \\
 & & & & a'_{rr}x_r & = & b'_r & - & a'_{rr+1}x_{r+1} & - & \dots & - & a'_{rp}x_p
\end{array}\right. ,$$

avec $(x_{r+1},\dots,x_p) \in \mathbb{K}^{p-r}$. En particulier, si $r = p$, c'est-à-dire si A est la matrice représentative d'une application injective, il y a existence et unicité de la solution.

Ce procédé se résume sous la forme d'un *algorithme*, ce qui facilite la mise en œuvre pratique de la méthode sur le plan numérique. Remarquons d'ailleurs que, puisque toutes les opérations sur les lignes concernent la matrice du système $A = (a_{ij})_{1\leq i\leq n,\ 1\leq j\leq p}$ et la matrice unicolonne des seconds membres $B = (b_i)_{1\leq i\leq n}$, on pourra appliquer cet algorithme à la matrice écrite par blocs $(A\,B) \in M_{n,p+1}(\mathbb{K})$, appelée parfois *matrice augmentée du système*.

$\boxed{\text{ALGORITHME DE LA MÉTHODE D'ÉLIMINATION DE GAUSS}}$

<u>Entrées :</u> $n \in \mathbb{N}^*$, $p \in \mathbb{N}^*$, $A \in M_{n,p}(\mathbb{K})$, $B \in M_{n,1}(\mathbb{K})$
<u>Sorties :</u> $r \in \mathbb{N}$, $A' \in M_{n,p+1}(\mathbb{K})$
$r := 0\,;\, A' := (A\,B)$
`tant que` $m < \min(n,p)$ `et que` $(a'_{ij})_{(i,j)\in\{r+1,\dots,n\}\times\{r+1,\dots,p\}} \neq 0$ `faire`
    $r := r+1$
    `si` $a'_{rr} = 0$ `alors`
        `trouver` $(i_0, j_0) \in \{r+1,\dots,n\}\times\{r+1,\dots,p\}$ `tel que` $a'_{i_0j_0} \neq 0$
        $L_r \leftrightarrow L_{i_0}$, $C_r \leftrightarrow C_{j_0}$
    `fin si`
    `pour` i `allant de` $r+1$ `à` n $L_i \leftarrow \dfrac{a'_{ir}}{a'_{rr}} L_r$
`fin tant que`

6.2.3 Exemples

Premier exemple. Considérons le système linéaire

$$\left\{\begin{array}{ccccccccc}
x_1 & + & 2x_2 & + & 3x_3 & + & 4x_4 & = & 11 \\
2x_1 & + & 3x_2 & + & 4x_3 & + & x_4 & = & 12 \\
3x_1 & + & 4x_2 & + & x_3 & + & 2x_4 & = & 13 \\
4x_1 & + & x_2 & + & 2x_3 & + & 3x_4 & = & 14
\end{array}\right. ,$$

et utilisons la méthode d'élimination de Gauss pour le résoudre. Nous trouvons

$$\left\{\begin{array}{rcrcrcrcr} x_1 & + & 2x_2 & + & 3x_3 & + & 4x_4 & = & 11 \\ & - & x_2 & - & 2x_3 & - & 7x_4 & = & -10 \\ & - & 2x_2 & - & 8x_3 & - & 10x_4 & = & -20 \\ & - & 7x_2 & - & 10x_3 & - & 13x_4 & = & -3 \end{array}\right. \quad \text{par} \quad \begin{array}{l} L_2 \leftarrow L_2 - 2L_1 \\ L_3 \leftarrow L_3 - 3L_1 \\ L_4 \leftarrow L_4 - 4L_1 \end{array},$$

$$\left\{\begin{array}{rcrcrcrcr} x_1 & + & 2x_2 & + & 3x_3 & + & 4x_4 & = & 11 \\ & - & x_2 & - & 2x_3 & - & 7x_4 & = & -10 \\ & & & - & 4x_3 & + & 4x_4 & = & 0 \\ & & & & 4x_3 & + & 36x_4 & = & 40 \end{array}\right. \quad \text{par} \quad \begin{array}{l} L_3 \leftarrow L_3 - 2L_2 \\ L_4 \leftarrow L_4 - 7L_2 \end{array},$$

$$\left\{\begin{array}{rcrcrcrcr} x_1 & + & 2x_2 & + & 3x_3 & + & 4x_4 & = & 11 \\ & - & x_2 & - & 2x_3 & - & 7x_4 & = & -10 \\ & & & - & 4x_3 & + & 4x_4 & = & 0 \\ & & & & & & 40x_4 & = & 40 \end{array}\right. \quad \text{par} \quad L_4 \leftarrow L_4 + L_3 ,$$

d'où

$$\left\{\begin{array}{l} x_4 = 1 \\ x_3 = x_4 = 1 \\ x_2 = 10 - 2x_3 - 7x_4 = 1 \\ x_1 = 11 - 2x_2 - 3x_3 - 4x_4 = 2 \end{array}\right. .$$

Ce système a pour unique solution $(2, 1, 1, 1)$.

Deuxième exemple. Considérons le système linéaire

$$\left\{\begin{array}{rcrcrcrcrcr} x_1 & + & 3x_2 & + & 5x_3 & - & 2x_4 & - & 7x_5 & = & 3 \\ 3x_1 & + & x_2 & + & x_3 & - & 2x_4 & - & x_5 & = & 1 \\ 2x_1 & - & x_2 & - & 3x_3 & + & 7x_4 & + & 5x_5 & = & 2 \\ 3x_1 & - & 2x_2 & - & 5x_3 & + & 7x_4 & + & 8x_5 & = & \alpha \end{array}\right. ,$$

avec α un paramètre réel, et appliquons la méthode d'élimination de Gauss pour le résoudre. Nous trouvons

$$\left\{\begin{array}{rcrcrcrcrcr} x_1 & + & 3x_2 & + & 5x_3 & - & 2x_4 & - & 7x_5 & = & 3 \\ & - & 8x_2 & - & 14x_3 & + & 4x_4 & + & 20x_5 & = & -8 \\ & - & 7x_2 & - & 13x_3 & + & 11x_4 & + & 19x_5 & = & -4 \\ & - & 11x_2 & - & 20x_3 & + & 13x_4 & + & 29x_5 & = & \alpha - 9 \end{array}\right. \quad \text{par} \quad \begin{array}{l} L_2 \leftarrow L_2 - 3L_1 \\ L_3 \leftarrow L_3 - 2L_1 \\ L_4 \leftarrow L_4 - 3L_1 \end{array},$$

$$\left\{\begin{array}{rcrcrcrcrcr} x_1 & + & 3x_2 & + & 5x_3 & - & 2x_4 & - & 7x_5 & = & 3 \\ & - & 8x_2 & - & 14x_3 & + & 4x_4 & + & 20x_5 & = & -8 \\ & & & - & 6x_3 & + & 60x_4 & + & 12x_5 & = & 24 \\ & & & - & 6x_3 & + & 60x_4 & + & 12x_5 & = & 8\alpha + 16 \end{array}\right. \quad \text{par} \quad \begin{array}{l} L_3 \leftarrow 8L_3 - 7L_2 \\ L_4 \leftarrow 8L_4 - 11L_2 \end{array},$$

$$\left\{\begin{array}{rcrcrcrcrcr} x_1 & + & 3x_2 & + & 5x_3 & - & 2x_4 & - & 7x_5 & = & 3 \\ & - & 8x_2 & - & 14x_3 & + & 4x_4 & + & 20x_5 & = & -8 \\ & & & - & 6x_3 & + & 60x_4 & + & 12x_5 & = & 24 \\ & & & & & & & & 0 & = & 8\alpha - 8 \end{array}\right. \quad \text{par} \quad L_4 \leftarrow L_4 + L_3 .$$

Nous constatons que si $\alpha \neq 1$, la condition de compatibilité $\alpha - 1 = 0$ n'est pas vérifiée et le système n'admet pas de solution. Supposons donc $\alpha = 1$; le système se réduit aux trois équations

$$\left\{\begin{array}{rcrcrcrcrcr} x_1 & + & 3x_2 & + & 5x_3 & - & 2x_4 & - & 7x_5 & = & 3 \\ & - & 8x_2 & - & 14x_3 & + & 4x_4 & + & 20x_5 & = & -8 \\ & & & - & 6x_3 & + & 60x_4 & + & 12x_5 & = & 24 \end{array}\right. .$$

Nous obtenons alors un système de Cramer par rapport aux inconnues x_1, $x-2$ et x_3 en traitant x_4 et x_5 comme des paramètres arbitraires :

$$\left\{\begin{array}{ccccccccc} x_1 & + & 3x_2 & + & 5x_3 & = & 3 & + & 2x_4 & + & 7x_5 \\ & & 4x_2 & + & 7x_3 & = & 4 & + & 2x_4 & + & 10x_5 \\ & & & & x_3 & = & -4 & + & 10x_4 & + & 2x_5 \end{array}\right. ,$$

d'où

$$\left\{\begin{array}{l} x_3 = -4 + 10x_4 + 2x_5 \\ x_2 = 8 - 17x_4 - x_5 \\ x_1 = -1 + 3x_4 \end{array}\right. .$$

Les solutions du système sont donc les 5-uplets s'écrivant

$$(-1+3x, 8-17x-y, -4+10x+2y, x, y) = (-1, 8, -4, 0, 0) + x(3, -17, 10, 1, 0) + y(0, -1, 2, 0, 1),$$

où x et y ont des valeurs arbitraires. La structure de l'ensemble des solutions a donc bien la forme d'une somme d'une solution particulière $(-1, 8, -4, 0, 0)$, obtenue pour $x = y = 0$, et de la solution générale du système homogène associé, qui est le plan vectoriel engendré par les vecteurs $(3, -17, 10, 1, 0)$ et $(0, -1, 2, 0, 1)$.

Troisième exemple. Soit le système linéaire

$$\left\{\begin{array}{ccccccccc} \alpha x_1 & + & x_2 & + & x_3 & + & x_4 & = & 1 \\ x_1 & + & \alpha x_2 & + & x_3 & + & x_4 & = & -2 \\ x_1 & + & x_2 & + & \alpha x_3 & + & x_4 & = & 0 \\ x_1 & + & x_2 & + & x_3 & + & \alpha x_4 & = & 3 \end{array}\right. ,$$

avec α un réel. Nous ne pouvons ici directement appliquer la méthode d'élimination de Gauss en raison du paramètre α (qui peut éventuellement être nul). Il va cependant être particulièrement utile d'adjoindre au système une équation supplémentaire, qui est une combinaison linéaire des équations initiales $L_5 = L_1 + L_2 + L_3 + L_4$. Nous obtenons

$$\left\{\begin{array}{ccccccccc} \alpha x_1 & + & x_2 & + & x_3 & + & x_4 & = & 1 \\ x_1 & + & \alpha x_2 & + & x_3 & + & x_4 & = & -2 \\ x_1 & + & x_2 & + & \alpha x_3 & + & x_4 & = & 0 \\ x_1 & + & x_2 & + & x_3 & + & \alpha x_4 & = & 3 \\ \multicolumn{9}{c}{(\alpha+3)(x_1+x_2+x_3+x_4) = 2} \end{array}\right. .$$

Nous voyons alors que si $\alpha = -3$, le système n'a pas de solution. Supposons donc $\alpha \neq -3$; nous trouvons alors

$$\left\{\begin{array}{ccc} (\alpha-1)x_1 & = & \dfrac{\alpha+1}{\alpha+3} \\ (\alpha-1)x_2 & = & -2\dfrac{\alpha+4}{\alpha+3} \\ (\alpha-1)x_3 & = & \dfrac{-2}{\alpha+3} \\ (\alpha-1)x_4 & = & \dfrac{3\alpha+7}{\alpha+3} \end{array}\right. \quad \text{par} \quad \begin{array}{l} L_1 \leftarrow L_1 - \dfrac{1}{\alpha+3} L_5 \\ L_2 \leftarrow L_2 - \dfrac{1}{\alpha+3} L_5 \\ L_3 \leftarrow L_3 - \dfrac{1}{\alpha+3} L_5 \\ L_4 \leftarrow L_4 - \dfrac{1}{\alpha+3} L_5 \end{array} ,$$

Nous constatons que si $\alpha = 1$, le système ne possède pas de solution.
Par conséquent, si le paramètre α n'est ni égal à -3, ni égal à 1, le système admet une unique solution donnée par le 4-uplet

$$(x_1, x_2, x_3, x_4) = \frac{1}{(\alpha+3)(\alpha-1)}(\alpha+1, -2\alpha-8, -2, 3\alpha+7).$$

6.2.4 Autres applications de la méthode

La méthode d'élimination de Gauss ne sert pas qu'à la résolution de systèmes linéaires. Pour toute matrice A de $M_{n,p}(\mathbb{K})$, elle permet de construite une matrice équivalente à A dont le rang s'obtient de façon évidente. Pour une matrice de $M_n(\mathbb{K})$ inversible, elle fournit un algorithme de calcul de A^{-1}.

Calcul du rang d'une matrice

Les opérations élémentaires sur une matrice ne modifiant pas son rang, on peut utiliser la méthode d'élimination de Gauss pour calculer le rang d'une matrice. Partant d'une matrice initiale A, nous appliquons l'algorithme de la méthode pour obtenir une matrice échelonnée. Le rang de A est alors simplement donné par le nombre de pivots non nuls.

Exemple. Calculons le rang de la matrice $A = \begin{pmatrix} 2 & 3 & 5 \\ 1 & 4 & 0 \\ -1 & -3 & -1 \\ 3 & 6 & 6 \end{pmatrix}$.
Nous obtenons successivement

$$r := 1 \quad \begin{pmatrix} 2 & 3 & 5 \\ 0 & 5/2 & -5/2 \\ 0 & -3/2 & 3/2 \\ 0 & 3/2 & -3/2 \end{pmatrix} \quad \text{par} \quad \begin{array}{l} L_2 \leftarrow L_2 - \frac{1}{2} L_1 \\ L_3 \leftarrow L_3 + \frac{1}{2} L_1 \\ L_4 \leftarrow L_4 - \frac{3}{2} L_1 \end{array}$$

$$r := 2 \quad \begin{pmatrix} 2 & 3 & 5 \\ 0 & 5/2 & -5/2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{par} \quad \begin{array}{l} L_3 \leftarrow L_3 + \frac{3}{5} L_2 \\ L_4 \leftarrow L_4 - \frac{3}{5} L_2 \end{array}$$

À l'issue de l'algorithme, nous avons obtenu une matrice échelonnée et trouvé deux pivots non nuls, 2 et $\dfrac{5}{2}$. Le rang de la matrice A est donc 2.

Calcul de l'inverse d'une matrice carrée

Soit $A \in GL_n(\mathbb{K})$. Commençons par remarquer, en notant e_j le j^e vecteur de la base canonique de $M_{n,1}(\mathbb{K})$ et $c_j \in M_{n,1}(\mathbb{K})$ la j^e colonne de A^{-1}, que $A^{-1}e_j = c_j$, et donc que $Ac_j = e_j$. On résoudra donc les n systèmes $Ac_j = e_j$ en utilisant l'algorithme du pivot de Gauss à la matrice A, en effectuant simultanément les mêmes opérations sur la matrice I_n. Remarquons que, s'il n'y a pas eu d'échange de colonnes et si la matrice A' obtenue à l'issue de n itérations de l'algorithme est égale à la matrice identité d'ordre n, alors la matrice obtenue à partir de I_n par la même suite d'opérations est l'inverse de A. Pour obtenir cela, on pourra utiliser la variante suivante de la méthode, appelée *méthode d'élimination de Gauss-Jordan* [4], appliquée à la matrice écrite par blocs $(A\ I_n) \in M_{n,2n}(\mathbb{K})$.

ALGORITHME DE LA MÉTHODE D'ÉLIMINATION DE GAUSS-JORDAN POUR LE CALCUL DE L'INVERSE

Entrées : $n \in \mathbb{N}^*$, $A \in GL_n(\mathbb{K})$
Sorties : $A' \in M_{n,2n}(\mathbb{K})$
$r := 0\,;\ A' := (A\ I_n)$
`tant que` $r < n$ `faire`
    $r := r + 1$
    `si` $a'_{rr} = 0$ `alors`
        `trouver` $i_0 \in \{r+1, \dots, n\}$ `tel que` $a'_{i_0 r} \neq 0$
        $L_r \leftrightarrow L_{i_0}$
    `fin si`
    $L_r \leftarrow \dfrac{1}{a'_{rr}} L_r$
    `pour` i `allant de` 1 `à` $r-1$ $L_i \leftarrow a'_{ir} L_r$
    `pour` i `allant de` $r+1$ `à` n $L_i \leftarrow a'_{ir} L_r$
`fin tant que`

4. Wilhelm Jordan (1er mars 1842 - 17 avril 1899) était un géodésiste allemand. Il est connu parmi les mathématiciens pour le procédé d'élimination portant son nom, publié en 1888 dans son *Handbuch der Vermessungskunde*, qu'il appliqua à la résolution de problèmes aux moindres carrés en géodésie.

Exemple. Soit $A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$. On trouve successivement

$$r := 0 \quad \left(\begin{array}{ccc|ccc} 2 & -1 & 0 & 1 & 0 & 0 \\ -1 & 2 & -1 & 0 & 1 & 0 \\ 0 & -1 & 2 & 0 & 0 & 1 \end{array}\right)$$

$$r := 1 \quad \left(\begin{array}{ccc|ccc} 1 & -1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 3/2 & -1 & 1/2 & 1 & 0 \\ 0 & -1 & 2 & 0 & 0 & 1 \end{array}\right) \qquad \text{par } L_1 \leftarrow \frac{1}{2}L_1 \text{ et } L_2 \leftarrow L_2 + \frac{1}{2}L_1$$

$$r := 2 \quad \left(\begin{array}{ccc|ccc} 1 & 0 & -1/3 & 2/3 & 1/3 & 0 \\ 0 & 1 & -2/3 & 1/3 & 2/3 & 0 \\ 0 & 0 & 4/3 & 1/3 & 2/3 & 1 \end{array}\right) \quad \text{par } L_1 \leftarrow L_1 + \frac{1}{3}L_2,\ L_2 \leftarrow \frac{2}{3}L_2 \text{ et } L_3 \leftarrow L_3 + \frac{2}{3}L_2$$

$$r := 3 \quad \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 3/4 & 1/2 & 1/4 \\ 0 & 1 & 0 & 1/2 & 1 & 1/2 \\ 0 & 0 & 1 & 1/4 & 1/2 & 3/4 \end{array}\right) \quad \text{par } L_1 \leftarrow L_1 + \frac{1}{4}L_3,\ L_2 \leftarrow L_2 + \frac{1}{2}L_3 \text{ et } L_3 \leftarrow \frac{3}{4}L_3$$

et donc

$$A^{-1} = \frac{1}{4}\begin{pmatrix} 3 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 3 \end{pmatrix}.$$

Deuxième partie

Analyse

Chapitre 7

Nombres réels

7.1 Rappels

Nous admettons l'existence et l'unicité de l'ensemble $\mathbb{R}$, muni des lois internes d'*addition* $+$ et de *multiplication* $\cdot$ et d'une relation $\leq$, ayant les propriétés suivantes :

(1) $(\mathbb{R}, +, \cdot)$ est un *corps commutatif*.

(2) $(\mathbb{R}, +, \cdot, \leq)$ est un *corps totalement ordonné*.

(3) Toute partie non vide et majorée de $\mathbb{R}$ admet une borne supérieure dans $\mathbb{R}$.

Rappelons que la propriété (1) signifie que $(\mathbb{R}, +)$ est un groupe commutatif, c'est-à-dire que

i) la loi $+$ est commutative : $\forall (a,b) \in \mathbb{R}^2$, $a + b = b + a$,

ii) la loi $+$ est associative : $\forall (a,b,c) \in \mathbb{R}^3$, $(a+b)+c = a+(b+c)$,

iii) la loi $+$ admet un élément neutre : $\forall a \in \mathbb{R}$, $a + 0 = 0 + a = a$,

iv) tout élément de $\mathbb{R}$ admet un symétrique pour la loi $+$: $\forall a \in \mathbb{R}$, $a + (-a) = (-a) + a = 0$,

et que $(\mathbb{R}^*, \cdot)$, où $\mathbb{R}^* = \mathbb{R} \backslash \{0\}$, est un groupe commutatif, c'est-à-dire que

i) la loi $\cdot$ est commutative : $\forall (a,b) \in \mathbb{R}^2$, $ab = ba$,

ii) la loi $\cdot$ est associative : $\forall (a,b,c) \in \mathbb{R}^3$, $(ab)c = a(bc)$,

iii) la loi $\cdot$ admet un élément neutre : $\forall a \in \mathbb{R}$, $a1 = 1a = a$,

iv) tout élément de $\mathbb{R}^*$ admet un symétrique pour la loi $\cdot$: $\forall a \in \mathbb{R}^*$, $a\frac{1}{a} = \frac{1}{a}a = 1$,

et que la multiplication est distributive par rapport à l'addition :

$$\forall (a,b,c) \in \mathbb{R}^3,\ a(b+c) = ab + ac.$$

La propriété (2) signifie pour sa part que $\leq$ est une relation d'ordre total dans $\mathbb{R}$, c'est-à-dire que

i) la relation $\leq$ est réflexive : $\forall a \in \mathbb{R}$, $a \leq a$,

ii) la relation $\leq$ est antisymétrique : $\forall (a,b) \in \mathbb{R}^2$, $(a \leq b \text{ et } b \leq a) \Rightarrow a = b$,

iii) la relation $\leq$ est transitive : $\forall (a,b,c) \in \mathbb{R}^3$, $(a \leq b \text{ et } b \leq c) \Rightarrow a \leq c$,

iv) la relation $\leq$ est totale : $\forall (a,b) \in \mathbb{R}^2$, $(a \leq b \text{ ou } b \leq a)$,

et que cette relation d'ordre est compatible avec les lois $+$ et $\cdot$, c'est-à-dire que

i) $\forall (a,b,c) \in \mathbb{R}^3$, $a \leq b \Rightarrow a + c \leq b + c$,

ii) $\forall (a,b,c) \in \mathbb{R}^3$, $(a \leq b \text{ et } 0 \leq c) \Rightarrow ac \leq bc$.

Pour $(a,b) \in \mathbb{R}^2$, $a < b$ signifie $a \leq b$ et $a \neq b$. On peut également noter $b \geq a$ (resp. $b > a$) au lieu de $a \leq b$ (resp. $a < b$).
Les éléments de $\mathbb{R}$ sont appelés les *nombres réels* et l'on note $\mathbb{R}_+ = \{x \in \mathbb{R} \mid x \geq 0\}$, $\mathbb{R}_- = \{x \in \mathbb{R} \mid x \leq 0\}$, $\mathbb{R}^* = \mathbb{R} \backslash \{0\}$, $\mathbb{R}_+^* = \mathbb{R}_+ \backslash \{0\}$ et $\mathbb{R}_-^* = \mathbb{R}_- \backslash \{0\}$.
Nous reviendrons dans la suite sur la propriété (3), qui est appelée l'*axiome de la borne supérieure*.

7.2 Majorant, minorant, bornes supérieure et inférieure

Définitions 7.1 *Soient A une partie de $\mathbb{R}$ et x un nombre réel. On dit que x est*
- *un **majorant de** A dans $\mathbb{R}$ si et seulement si il est supérieur ou égal à tous les éléments de A :*
$$\forall a \in A,\ a \leq x,$$
- *un **minorant de** A dans $\mathbb{R}$ si et seulement si il est inférieur ou égal à tous les éléments de A :*
$$\forall a \in A,\ x \leq a,$$
- *un **élément maximal**, ou **plus grand élément, de** A dans $\mathbb{R}$ si et seulement si c'est un majorant de A dans $\mathbb{R}$ appartenant à A,*
- *un **élément minimal**, ou **plus petit élément, de** A dans $\mathbb{R}$ si et seulement si c'est un minorant de A dans $\mathbb{R}$ appartenant à A.*

La partie A est dite *majorée* (resp. *minorée*) dans $\mathbb{R}$ si et seulement si elle possède au moins un majorant (resp. minorant) et *bornée* si et seulement si elle est à la fois majorée et minorée.
Si A admet un plus grand (resp. petit) élément, celui-ci est unique et on le note $\max(A)$ (resp. $\min(A)$).

Définitions 7.2 *On appelle*
- *la **borne supérieure de** A dans $\mathbb{R}$ le plus petit des majorants de A dans $\mathbb{R}$, s'il existe. Il est alors unique et noté $\sup(A)$,*
- *la **borne inférieure de** A dans $\mathbb{R}$ le plus grand des minorants de A dans $\mathbb{R}$, s'il existe. Il est alors unique et noté $\inf(A)$.*

Notons que si A possède un plus grand (resp. petit) élément $\max(A)$ (resp. $\min(A)$), alors $\max(A) = \sup(A)$ (resp. $\min(A) = \inf(A)$).

Exemples.
— On a $\max[0,1] = 1$, mais ni $[0,1[$ ni $[0,+\infty[$ n'ont un plus grand élément dans $\mathbb{R}$.
— Les intervalles $\varnothing$, $[0,1]$ et $[0,1[$ sont majorés dans $\mathbb{R}$, mais l'ensemble $\mathbb{N}$ n'est pas majoré dans $\mathbb{R}$.
— On a $\sup\limits_{\mathbb{R}}[0,1] = \sup\limits_{\mathbb{R}}[0,1[= 1$, mais $\sup\limits_{\mathbb{R}}[0,+\infty[$ n'existe pas.

Rappelons enfin une propriété que nous admettrons, à savoir l'axiome de la borne supérieure.

Proposition 7.3 *Toute partie non vide et majorée de $\mathbb{R}$ admet une borne supérieure dans $\mathbb{R}$.*

En considérant l'ensemble des opposés des éléments de la partie envisagée, on obtient à partir de cet axiome le résultat suivant.

Proposition 7.4 *Toute partie non vide et minorée de $\mathbb{R}$ admet une borne inférieure dans $\mathbb{R}$.*

7.3 Propriétés des nombres réels

7.3.1 Théorème d'Archimède

Théorème 7.5 *L'ensemble $\mathbb{R}$ est un corps **archimédien**, c'est-à-dire un corps vérifiant l'axiome d'Archimède [1] :*
$$\forall x \in \mathbb{R}_+^*,\ \forall y \in \mathbb{R}_+^*,\ \exists n \in \mathbb{N}^*,\ ny > x.$$

DÉMONSTRATION. Soient $x \in \mathbb{R}_+^*$ et $y \in \mathbb{R}_+^*$. Vérifions par l'absurde que x n'est pas un majorant de l'ensemble $E = \{ny \mid n \in \mathbb{N}^*\}$. Supposons donc que x est un majorant de E. L'ensemble E étant une partie non vide de $\mathbb{R}$ et de plus majorée par x, celui-ci admet, d'après la proposition 7.3, une borne supérieure réelle que l'on note M. Comme M est le plus petit des majorants de E, le réel $M - y$ n'est pas un majorant de E, ce qui signifie qu'il existe un entier relatif n tel que $ny > M - y$, et donc $(n+1)y > M$. Or, $(n+1)y \in E$, ce qui contredit le fait que $M = \sup E$. On en déduit que le réel donné x ne majore pas E et, par suite, qu'on peut trouver un élément n de E qui vérifie $ny > x$. □

1. Archimède de Syracuse (Αρχιμήδης en grec, 287 av. J.-C. - 212 av. J.-C.) était un physicien, mathématicien et ingénieur grec de l'Antiquité. Scientifique de grande envergure, il inventa la poulie, la roue dentée, la vis sans fin ainsi que des machines de guerre pour repousser les romains durant le siège de Syracuse. En physique, on lui doit en particulier les premières lois de l'hydrostatique et une étude précise sur l'équilibre des surfaces planes. En mathématiques, il établit notamment de nombreuses formules relatives aux mesures des surfaces et des volumes qui font de lui un précurseur du calcul intégral.

7.3.2 Partie entière d'un réel

Pour tout $x \in \mathbb{R}$, en appliquant le théorème précédent avec $y = 1$, on voit que $\{n \in \mathbb{Z} \mid n \leq x\}$ est une partie majorée et non vide de $\mathbb{Z}$, qui admet par conséquent un plus grand élément. Nous obtenons ainsi le résultat suivant.

Proposition et définition 7.6 *Pour tout réel x, il existe un unique entier relatif n vérifiant $n \leq x < n+1$. Cet entier est appelé **partie entière de** x et noté $E(x)$.*

Exemples. $E(\pi) = 3$, $E\left(\frac{3}{2}\right) = 1$, $E\left(-\frac{3}{2}\right) = -2$.

7.3.3 Valeur absolue d'un réel

Définition 7.7 *On appelle **valeur absolue du réel** x le réel noté $|x|$ défini par $|x| = \sup\{x, -x\}$.*

Propriétés

On dispose des propriétés suivantes pour la valeur absolue :

1) $\forall x \in \mathbb{R}$, $|x| = |-x|$,
2) $\forall (x,y) \in \mathbb{R}^2$, $|xy| = |x||y|$.
   On en déduit
$$\forall n \in \mathbb{N}^*,\ \forall (x_1,\ldots,x_n) \in \mathbb{R}^n,\ \left|\prod_{i=1}^n x_i\right| = \prod_{i=1}^n |x_i|.$$
3) $\forall x \in \mathbb{R}^*$, $\left|\frac{1}{x}\right| = \frac{1}{|x|}$.
4) $\forall (x,y) \in \mathbb{R}^2$, $|x+y| \leq |x| + |y|$ (*inégalité triangulaire*).
   Toutes les quantités étant positives, il suffit de comparer leurs carrés. On a
$$(|x|+|y|)^2 = |x|^2 + |y|^2 + 2|x||y| = x^2 + y^2 + 2|x||y|.$$
   Or, $xy \leq |x||y|$ d'où $|x+y|^2 \leq (|x|+|y|)^2$.
   On en déduit
$$\forall n \in \mathbb{N}^*,\ \forall (x_1,\ldots,x_n) \in \mathbb{R}^n,\ \left|\sum_{i=1}^n x_i\right| \leq \sum_{i=1}^n |x_i|.$$
5) $\forall (x,y) \in \mathbb{R}^2$, $||x|-|y|| \leq |x-y|$,
   car $|x| = |x-y+y| \leq |x-y| + |y|$, donc $|x|-|y| \leq |x-y|$ et de même $|y|-|x| \leq |y-x| = |x-y|$.

7.3.4 Densité de $\mathbb{Q}$ dans $\mathbb{R}$

Définition 7.8 *Une partie D de $\mathbb{R}$ est dite **dense dans** $\mathbb{R}$ si et seulement si*
$$\forall (x,y) \in \mathbb{R}^2,\ (x < y \Rightarrow (\exists d \in D,\ x < d < y)).$$

Théorème 7.9 *L'ensemble $\mathbb{Q}$ est dense dans $\mathbb{R}$.*

DÉMONSTRATION. Soient $(x,y) \in \mathbb{R}^2$, tel que $x < y$, et $\varepsilon = y - x > 0$. Puisque $\mathbb{R}$ est archimédien, il existe $n \in \mathbb{N}^*$ tel que $n\varepsilon > 1$, c'est-à-dire $\frac{1}{n} < \varepsilon$. En notant $m = E(nx)+1$, on obtient $m-1 \leq nx < m$, d'où $x < \frac{m}{n} \leq x + \frac{1}{n} < x + \varepsilon = y$. □

7.4 Intervalles

Définition 7.10 *Une partie A de $\mathbb{R}$ est un **intervalle** si, dès qu'elle contient deux réels, elle contient tous les réels intermédiaires, c'est-à-dire*
$$\forall (a,b) \in A^2,\ \forall x \in \mathbb{R},\ (a \leq x \leq b \Rightarrow x \in A).$$

Exemples.

— $\mathbb{R}_+$ est un intervalle. En effet, tout réel compris entre deux réels positifs est lui même positif.

— $\mathbb{R}^*$ n'est pas un intervalle. On a en effet $-1 \in \mathbb{R}^*$, $1 \in \mathbb{R}^*$ et $-1 \leq 0 \leq 1$, mais $0 \notin \mathbb{R}^*$.

Classification des intervalles

La propriété de la borne supérieure permet de classifier tous les intervalles de $\mathbb{R}$. Tout d'abord, $\emptyset$ est un intervalle, ainsi que tout singleton $\{x\}$ (avec x réel), puisque ces deux types d'ensemble vérifient la définition ci-dessus. Considérons maintenant un intervalle A contenant au moins deux éléments.

i) Si A est à la fois majoré et minoré, l'intervalle possède une borne supérieure, que l'on note $b = \sup(A)$, et une borne inférieure, notée $a = \inf(A)$. On a alors $\forall x \in A,\ a \leq x \leq b$ et donc $A \subset \{x \in \mathbb{R} \mid a \leq x \leq b\}$.
Réciproquement, soit x un réel tel que $a < x < b$; x n'est pas un majorant (resp. minorant) de A, donc il existe un réel z (resp. y) tel que $z > x$ (resp. $x > y$). Par conséquent, y et z appartiennent à A et on a $y < x < z$. En utilisant la définition d'un intervalle, il vient que x appartient à A, qui contient donc tous les éléments compris entre a et b.
Selon que les réels a et b appartiennent eux-mêmes à A ou non, on peut avoir
$A = \{x \in \mathbb{R} \mid a \leq x \leq b\} = [a, b]$, l'intervalle est dit *fermé borné* ou encore appelé *segment*,
$A = \{x \in \mathbb{R} \mid a < x \leq b\} =]a, b]$, l'intervalle est dit *borné semi-ouvert à gauche*,
$A = \{x \in \mathbb{R} \mid a \leq x < b\} = [a, b[$, l'intervalle est dit *borné semi-ouvert à droite*,
$A = \{x \in \mathbb{R} \mid a < x < b\} =]a, b[$, l'intervalle est dit *borné ouvert*.

ii) Si A est minoré et non majoré, l'intervalle admet une borne inférieure que l'on note a et $\forall x \in A,\ x \geq a$.
Réciproquement, soit x un réel tel que $x > a$; x n'est ni un majorant, ni un minorant de A, donc il existe y et z appartenant à A tels que $y < x < z$ et par conséquent $x \in A$.
Selon que le réel a appartient ou non à A, on peut avoir
$A = \{x \in \mathbb{R} \mid x \geq a\} = [a, +\infty[$, l'intervalle est dit *fermé non majoré*,
$A = \{x \in \mathbb{R} \mid x > a\} =]a, +\infty[$, l'intervalle est dit *ouvert non majoré*.

iii) Si A est majoré et non minoré, on peut avoir de la même façon
$A = \{x \in \mathbb{R} \mid x \leq b\} =]-\infty, b]$, l'intervalle est dit *fermé non minoré*,
$A = \{x \in \mathbb{R} \mid x < b\} =]-\infty, b[$, l'intervalle est dit *ouvert non minoré*.

iv) Si A est non majoré et non minoré, un réel x quelconque n'est ni minorant, ni majorant de A, il existe donc des réels y et z appartenant à A tels que $y < x < z$ et $x \in A$. Par conséquent, on a $A = \mathbb{R}$.

En définitive, tout intervalle est de l'un des onze types que nous venons d'énoncer.

7.5 Droite numérique achevée

Définition 7.11 *On appelle **droite numérique achevée**, et l'on note $\overline{\mathbb{R}}$, l'ensemble $\mathbb{R} \cup \{-\infty, +\infty\}$, où $-\infty$ et $+\infty$ sont deux éléments non réels, sur lequel sont prolongées les lois internes $+$ et $\cdot$, ainsi que la relation d'ordre $\leq$.*

La relation $\leq$ est étendue à $\overline{\mathbb{R}}$ de la manière suivante

$$\forall x \in \mathbb{R},\ -\infty < x < +\infty\ ;\ -\infty \leq -\infty,\ +\infty \leq +\infty.$$

On remarquera que les lois $+$ et $\cdot$ ne sont définies que partiellement sur $\overline{\mathbb{R}}$; on a en effet

$$\forall x \in \mathbb{R},\ x + (+\infty) = +\infty,\ x + (-\infty) = -\infty,\ (+\infty) + (+\infty) = +\infty,\ (-\infty) + (-\infty) = -\infty,$$

et d'autre part

$$\begin{aligned}
&\forall x > 0,\ x(+\infty) = (+\infty)x = +\infty,\ x(-\infty) = (-\infty)x = -\infty,\\
&\forall x < 0,\ x(+\infty) = (+\infty)x = -\infty,\ x(-\infty) = (-\infty)x = +\infty,\\
&(+\infty)(+\infty) = (-\infty)(-\infty) = +\infty,\ (+\infty)(-\infty) = (-\infty)(+\infty) = -\infty,
\end{aligned}$$

mais les sommes $(+\infty) + (-\infty)$ et $(-\infty) + (+\infty)$ et les produits $0(+\infty)$, $(+\infty)0$, $0(-\infty)$ et $(-\infty)0$ ne sont pas définis. Ils correspondent à des *formes indéterminées*.
Nous terminons par un résultat admis, analogue à la proposition 7.3.

Proposition 7.12 *Toute partie non vide de $\overline{\mathbb{R}}$ admet une borne supérieure et une borne inférieure dans $\overline{\mathbb{R}}$.*

Chapitre 8

Suites numériques

8.1 Généralités

8.1.1 Définitions et propriétés

Une *suite numérique* est une application de $\mathbb{N}$ dans $\mathbb{K}$ (où $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$). Plutôt que de noter

$$u : \begin{array}{ccc} \mathbb{N} & \to & \mathbb{K} \\ n & \mapsto & u(n) \end{array}$$

on emploie souvent les notations $(u_n)_{n\in\mathbb{N}}$, $(u_n)_{n\geq 0}$ ou bien encore $(u_n)_n$.
Une *suite réelle* (resp. *complexe*) est une suite numérique telle que

$$\forall n \in \mathbb{N},\ u_n \in \mathbb{R} \text{ (resp. } \mathbb{C}\text{)}.$$

Pour chaque entier n, u_n est appelé le n^e *terme* de la suite. On veillera à ne pas confondre la suite $(u_n)_{n\in\mathbb{N}}$ et son terme général u_n. L'ensemble des suites numériques est noté $\mathscr{F}(\mathbb{N},\mathbb{K})$, ou encore $\mathbb{K}^{\mathbb{N}}$.

Définitions 8.1 *Une suite numérique $(u_n)_{n\in\mathbb{N}}$ est dite* ***constante*** *si et seulement si*

$$\forall n \in \mathbb{N},\ u_{n+1} = u_n.$$

Elle est dite ***stationnaire*** *si et seulement si elle est constante à partir d'un certain rang, c'est-à-dire si*

$$\exists N \in \mathbb{N},\ \forall n \in \mathbb{N},\ (n \geq N \Rightarrow u_{n+1} = u_n).$$

Enfin, la suite est dite ***périodique*** *si et seulement s'il existe un entier p strictement positif tel que*

$$\forall n \in \mathbb{N},\ u_{n+p} = u_n.$$

Définition 8.2 *Une suite numérique $(u_n)_{n\in\mathbb{N}}$ est dite* ***bornée*** *si et seulement s'il existe un réel positif M tel que*

$$\forall n \in \mathbb{N},\ |u_n| \leq M.$$

Définitions 8.3 *Une suite réelle $(u_n)_{n\in\mathbb{N}}$ est dite* ***majorée*** *(resp.* ***minorée****) si et seulement s'il existe un réel M (resp. m), appelé* ***majorant*** *(resp.* ***minorant****) de la suite $(u_n)_{n\in\mathbb{N}}$, tel que*

$$\forall n \in \mathbb{N},\ u_n \leq M \ (\textit{resp. } u_n \geq m).$$

On voit qu'une suite réelle est bornée si et seulement si elle est majorée et minorée.

8.1.2 Opérations sur les suites

On peut définir sur l'ensemble des suites numériques $\mathbb{K}^{\mathbb{N}}$

— une *addition* :
$$(w_n)_{n\in\mathbb{N}} = (u_n)_{n\in\mathbb{N}} + (v_n)_{n\in\mathbb{N}} \Leftrightarrow \forall n \in \mathbb{N},\ w_n = u_n + v_n.$$
Cette opération est commutative, associative et admet pour élément neutre la suite constante nulle. Tout suite $(u_n)_{n\in\mathbb{N}}$ a une suite opposée $(-u_n)_{n\in\mathbb{N}}$. $(\mathbb{K}^{\mathbb{N}}, +)$ est un groupe commutatif.

— une *multiplication interne* :
$$(w_n)_{n\in\mathbb{N}} = (u_n)_{n\in\mathbb{N}}(v_n)_{n\in\mathbb{N}} \Leftrightarrow \forall n \in \mathbb{N},\ w_n = u_n v_n.$$
Cette opération est commutative, associative et admet pour élément neutre la suite constante égale à 1. Elle est distributive par rapport à l'addition. Il existe cependant des suites non nulles n'ayant pas d'inverse et $(\mathbb{K}^{\mathbb{N}}, \cdot)$ n'est pas un groupe.

— une *multiplication externe par les scalaires* :
$$\forall \lambda \in \mathbb{K},\ (w_n)_{n\in\mathbb{N}} = \lambda(u_n)_{n\in\mathbb{N}} \Leftrightarrow \forall n \in \mathbb{N},\ w_n = \lambda u_n.$$

8.1.3 Suites extraites

Définition 8.4 *Soit une suite $(u_n)_{n\in\mathbb{N}}$. On appelle* ***suite extraite de*** $(u_n)_{n\in\mathbb{N}}$ *toute suite $(u_{\sigma(n)})_{n\in\mathbb{N}}$, où $\sigma : \mathbb{N} \to \mathbb{N}$ est une application strictement croissante.*

On donne aussi le nom de *sous-suite de* $(u_n)_{n\in\mathbb{N}}$ à toute suite extraite de $(u_n)_{n\in\mathbb{N}}$.
L'application σ de la définition est parfois appelée une *extractrice*. On montre aisément, par récurrence, que pour toute extractrice σ, on a
$$\forall n \in \mathbb{N},\ \sigma(n) \geq n.$$

Exemples.
— $(u_{2n})_{n\in\mathbb{N}}$ et $(u_{2n+1})_{n\in\mathbb{N}}$ sont des suites extraites de $(u_n)_{n\in\mathbb{N}}$.
— $(u_{n^2})_{n\in\mathbb{N}}$ est extraite de $(u_n)_{n\in\mathbb{N}}$.
— $(u_{n^2-n})_{n\in\mathbb{N}}$ n'est pas une suite extraite de $(u_n)_{n\in\mathbb{N}}$ (le terme u_0 apparaît deux fois).

8.1.4 Suites de Cauchy

Définition 8.5 *Une suite $(u_n)_{n\in\mathbb{N}}$ est dite* ***de Cauchy***[1] *si et seulement si*
$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N},\ \forall (p,q) \in \mathbb{N}^2,\ (p \geq N,\ q \geq N \Rightarrow |u_p - u_q| \leq \varepsilon).$$

Proposition 8.6 *Toute suite de Cauchy est bornée.*

DÉMONSTRATION. Soit $(u_n)_{n\in\mathbb{N}}$ une suite de Cauchy. On a
$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N},\ \forall (p,q) \in \mathbb{N}^2,\ (p \geq N,\ q \geq N \Rightarrow |u_p - u_q| \leq \varepsilon).$$
Fixons une valeur particulière pour ε, par exemple $\varepsilon = 1$. Il existe alors un entier N' tel que $(p \geq N',\ q \geq N' \Rightarrow |u_p - u_q| \leq 1)$, ou encore, pour $q = N'$, $(p \geq N' \Rightarrow |u_p - u_{N'}| \leq 1)$, c'est-à-dire
$$\forall n \in \mathbb{N},\ (n \geq N' \Rightarrow u_{N'} - 1 \leq u_n \leq u_{N'} + 1).$$
Notons $A = \min\{u_0, \ldots, u_{N'-1}, u_{N'} - 1\}$ et $B = \max\{u_0, \ldots, u_{N'-1}, u_{N'} + 1\}$. Nous avons alors
$$\forall n \in \mathbb{N},\ A \leq u_n \leq B,$$
et, par suite, $(u_n)_{n\in\mathbb{N}}$ est bornée. □

1. Augustin-Louis Cauchy (21 août 1789 – 23 mai 1857) était un mathématicien français. Très prolifique, ses recherches couvrent l'ensemble des domaines mathématiques de son époque. On lui doit notamment en analyse l'introduction des fonctions holomorphes et des critères de convergence des séries. Ses travaux sur les permutations furent précurseurs de la théorie des groupes. Il fit aussi d'importantes contributions à l'étude de la propagation des ondes en optique et en mécanique.

8.1.5 Suites réelles monotones

Définitions 8.7 *Soit $(u_n)_{n\in\mathbb{N}}$ une suite réelle. On dit que $(u_n)_{n\in\mathbb{N}}$ est*

- ***croissante*** *si et seulement si*
$$\forall n \in \mathbb{N},\ u_n \leq u_{n+1},$$
- ***décroissante*** *si et seulement si*
$$\forall n \in \mathbb{N},\ u_{n+1} \leq u_n,$$
- ***strictement croissante*** *si et seulement si*
$$\forall n \in \mathbb{N},\ u_n < u_{n+1},$$
- ***strictement décroissante*** *si et seulement si*
$$\forall n \in \mathbb{N},\ u_{n+1} < u_n,$$
- ***monotone*** *si et seulement si elle est croissante ou décroissante,*
- ***strictement monotone*** *si et seulement si elle est strictement croissante ou strictement décroissante.*

8.2 Convergence d'une suite

8.2.1 Généralités

Définitions 8.8 *On dit qu'une suite numérique $(u_n)_{n\in\mathbb{N}}$*

- ***converge vers*** *$l \in \mathbb{K}$ si et seulement si*
$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N},\ \forall n \in \mathbb{N}, (n \geq N \Rightarrow |u_n - l| \leq \varepsilon).$$
- ***converge*** *si et seulement s'il existe $l \in \mathbb{K}$ tel que $(u_n)_{n\in\mathbb{N}}$ converge vers l, c'est-à-dire*
$$\exists l \in \mathbb{K},\ \forall \varepsilon > 0,\ \exists N \in \mathbb{N},\ \forall n \in \mathbb{N}, (n \geq N \Rightarrow |u_n - l| \leq \varepsilon).$$
- ***diverge*** *si et seulement si elle ne converge pas, c'est-à-dire*
$$\forall l \in \mathbb{K},\ \exists \varepsilon > 0,\ \forall N \in \mathbb{N},\ \exists n \in \mathbb{N}, (n \geq N \text{ et } |u_n - l| \geq \varepsilon).$$

On remarque que $(u_n)_{n\in\mathbb{N}}$ converge vers l si et seulement si la suite $(u_n - l)_{n\in\mathbb{N}}$ converge vers 0. Lorsque $(u_n)_{n\in\mathbb{N}}$ converge vers l, on dit aussi que $(u_n)_{n\in\mathbb{N}}$ *tend vers* l.

Proposition 8.9 *Si une suite numérique $(u_n)_{n\in\mathbb{N}}$ converge vers l, alors l est unique.*

DÉMONSTRATION. Supposons que $(u_n)_{n\in\mathbb{N}}$ converge à la fois vers l et vers l', avec $l \neq l'$. Posons $\varepsilon = \frac{1}{3}|l - l'|$. Par définition de la convergence, il existe des entiers N et N' tels que
$$\forall n \in \mathbb{N},\ (n \geq N \Rightarrow |u_n - l| \leq \varepsilon) \text{ et } (n \geq N' \Rightarrow |u_n - l'| \leq \varepsilon).$$
Soit $n \geq \max(N, N')$, nous avons alors $|u_n - l| \leq \varepsilon$ et $|u_n - l'| \leq \varepsilon$, d'où
$$|l - l'| \leq |l - u_n| + |u_n - l'| \leq 2\varepsilon = \frac{2}{3}|l - l'|,$$
ce qui est absurde. □

Il ne peut donc exister qu'un scalaire l tel que $(u_n)_{n\in\mathbb{N}}$ converge vers l. On dit que l est la *limite de* $(u_n)_{n\in\mathbb{N}}$, et l'on note
$$\lim_{n\to+\infty} u_n = l \text{ (ou } u_n \underset{n\to+\infty}{\longrightarrow} l).$$

8.2.2 Suites réelles tendant vers l'infini

Définition 8.10 *Soit $(u_n)_{n\in\mathbb{N}}$ une suite réelle.*

1) *On dit que $(u_n)_{n\in\mathbb{N}}$ **tend vers** $+\infty$ si et seulement si*

$$\forall A \in \mathbb{R},\ \exists N \in \mathbb{N},\ \forall n \in \mathbb{N},\ (n \geq N \Rightarrow u_n \geq A).$$

On note alors $\lim\limits_{n\to+\infty} u_n = +\infty$.

2) *On dit que $(u_n)_{n\in\mathbb{N}}$ **tend vers** $-\infty$ si et seulement si*

$$\forall B \in \mathbb{R},\ \exists N \in \mathbb{N},\ \forall n \in \mathbb{N},\ (n \geq N \Rightarrow u_n \leq B).$$

et l'on note $\lim\limits_{n\to+\infty} u_n = -\infty$.

8.2.3 Suites adjacentes

La limite d'une suite étant définie, nous pouvons introduire la notion de *suites adjacentes*.

Définition 8.11 *Deux suites réelles $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ sont dite **adjacentes** si et seulement si*

i) l'une est croissante et l'autre est décroissante,

ii) $\lim\limits_{n\to+\infty}(u_n - v_n) = 0$.

8.2.4 Propriétés d'une suite réelle convergente

Proposition 8.12 *Toute suite réelle convergente est bornée.*

DÉMONSTRATION. Soient $(u_n)_{n\in\mathbb{N}}$ une suite convergente de limite $l \in \mathbb{R}$ et $\varepsilon > 0$. Il existe alors un entier N tel que, pour tout $n \in \mathbb{N}$, si $n \geq N$ alors $|u_n - l| \leq \varepsilon$, c'est-à-dire $l - \varepsilon \leq u_n \leq l + \varepsilon$. L'ensemble $\{u_n \mid n > N\}$ est donc majoré par $l + \varepsilon$ et minoré par $l - \varepsilon$. D'autre part, l'ensemble $\{u_0, \ldots, u_N\}$ est fini et, par suite, il admet un maximum u_M et un minimum u_m, d'où

$$\forall n \in \mathbb{N},\ \inf(u_m, l - \varepsilon) \leq u_n \leq \max(u_M, l + \varepsilon).$$

La suite $(u_n)_{n\in\mathbb{N}}$ est donc bornée. □

Remarques.

— La réciproque de cette proposition est fausse, comme le montre l'exemple de la suite définie par $u_n = (-1)^n$ pour tout entier $n \geq 0$.

— Toute suite non bornée diverge.

Proposition 8.13 *Si une suite $(u_n)_{n\in\mathbb{N}}$ est convergente, toute suite extraite de $(u_n)_{n\in\mathbb{N}}$ est convergente et tend vers la même limite.*

DÉMONSTRATION. Soit $(u_n)_{n\in\mathbb{N}}$ une suite convergente de limite l. Nous avons alors

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N},\ \forall n \in \mathbb{N},\ (n \geq N \Rightarrow |u_n - l| \leq \varepsilon).$$

Soit σ une extractrice. Nous savons que, pour tout entier naturel n, $\sigma(n) \geq n$, donc si $n \geq N$ alors $\sigma(n) \geq \sigma(N) \geq N$ et, par suite, $|u_{\sigma(n)} - l| \leq \varepsilon$. On en conclut que

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N},\ \forall n \in \mathbb{N},\ (n \geq N \Rightarrow |u_{\sigma(n)} - l| \leq \varepsilon),$$

et donc que la suite extraite $(u_{\sigma(n)})_{n\in\mathbb{N}}$ converge vers l. □

La contraposée de cette proposition permet de montrer qu'une suite diverge : il suffit pour cela d'en extraire deux suites qui convergent vers deux limites différentes.

Exemple. Soit $(u_n)_{n\in\mathbb{N}}$ la suite définie par $u_n = (-1)^n$ pour tout entier $n \geq 0$. La suite extraite $(u_{2n})_{n\in\mathbb{N}}$ a pour limite 1 et la suite $(u_{2n+1})_{n\in\mathbb{N}}$ a pour limite -1. Cette suite diverge donc.

Par ailleurs, la réciproque de cette dernière proposition est en général fausse ; en effet, on peut trouver des suites divergentes qui admettent pourtant deux suites extraites convergeant vers une même limite.

Exemple. Soit la suite réelle définie par $u_n = \cos\left((n+(-1)^n)\frac{\pi}{3}\right)$ pour tout entier positif n. Nous avons $u_{6n} = \cos\left((6n+1)\frac{\pi}{3}\right) = \cos\left(2n\pi + \frac{\pi}{3}\right) = \frac{1}{2}$ et $u_{6n+4} = \cos\left((6n+5)\frac{\pi}{3}\right) = \cos\left(2n\pi + \frac{5\pi}{3}\right) = \frac{1}{2}$. Cependant, la suite $(u_n)_{n\in\mathbb{N}}$ est divergente car $u_{3n+6} = \cos\left((6n+2)\frac{\pi}{3}\right) = \cos\left(2n\pi + \frac{2\pi}{3}\right) = -\frac{1}{2}$.
Cependant, on a le résultat suivant.

Proposition 8.14 *Soient $(u_n)_{n\in\mathbb{N}}$ une suite réelle et $l \in \mathbb{R}$. Pour que $(u_n)_{n\in\mathbb{N}}$ converge vers l, il faut et il suffit que les suites $(u_{2n})_{n\in\mathbb{N}}$ et $(u_{2n+1})_{n\in\mathbb{N}}$ convergent toutes deux vers l.*

DÉMONSTRATION. Supposons que les suites extraites $(u_{2n})_{n\in\mathbb{N}}$ et $(u_{2n+1})_{n\in\mathbb{N}}$ convergent vers une limite l. Soit ε un réel strictement positif. Il existe des entiers N et N' tels que, pour tout entier n,

$$(n \geq N \Rightarrow |u_{2n} - l| \leq \varepsilon) \text{ et } (n \geq N' \Rightarrow |u_{2n+1} - l| \leq \varepsilon).$$

Notons $N'' = \max(2N, 2N'+1)$ et soit $p \in \mathbb{N}$ tel que $p \geq N''$. Si p est pair, il existe n tel que $p = 2n$. Dans ce cas, nous avons $2n \geq 2N$ donc $n \geq N$, d'où

$$|u_p - l| = |u_{2n} - l| \leq \varepsilon.$$

Si p est impair, il existe n tel que $p = 2n+1$. Nous avons alors $2n+1 \geq 2N'+1$ donc $n \geq N'$, d'où

$$|u_p - l| = |u_{2n+1} - l| \leq \varepsilon.$$

Ceci montre que la suite $(u_p)_{p\in\mathbb{N}}$ converge vers l. □

Proposition 8.15 *Toute suite convergente est de Cauchy.*

DÉMONSTRATION. Soient $(u_n)_{n\in\mathbb{N}}$ une suite convergente de limite l et ε un réel strictement positif. Il existe alors un entier N tel que

$$\forall n \in \mathbb{N}, \left(n \geq N \Rightarrow |u_n - l| \leq \frac{\varepsilon}{2}\right).$$

Soient p et q deux entiers supérieurs à N ; nous avons alors

$$|u_p - l| \leq \frac{\varepsilon}{2} \text{ et } |u_q - l| \leq \frac{\varepsilon}{2}.$$

Or,

$$|u_p - u_q| \leq |u_p - l| + |u_q - l| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

En conclusion, nous avons

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N},\ \forall (p,q) \in \mathbb{N}^2,\ (p \geq N,\ q \geq N \Rightarrow |u_p - u_q| \leq \varepsilon).$$

La suite $(u_n)_{n\in\mathbb{N}}$ est alors une suite de Cauchy. □

Nous montrerons dans la suite que la réciproque de cette proposition est vraie.

8.2.5 Propriétés d'ordre des suites réelles convergentes

Passage à la limite dans une inégalité

Proposition 8.16 *Soient $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ deux suites réelles convergentes. Si $u_n \leq v_n$ pour tout entier n, alors*

$$\lim_{n\to+\infty} u_n \leq \lim_{n\to+\infty} v_n.$$

DÉMONSTRATION. Supposons que $\lim\limits_{n\to+\infty} u_n > \lim\limits_{n\to+\infty} v_n$. Nous avons alors $\lim\limits_{n\to+\infty}(u_n - v_n) > 0$, ce qui entraîne

$$\exists N \in \mathbb{N},\ \forall n \in \mathbb{N},\ (n \geq N \Rightarrow (u_n - v_n) > 0),$$

ce qui contredit l'hypothèse. □

Même si l'on a $u_n < v_n$ pour tout $n \in \mathbb{N}$, on peut avoir $\lim\limits_{n\to+\infty} u_n \leq \lim\limits_{n\to+\infty} v_n$ car le passage à la limite élargit les inégalités, comme l'illustre l'exemple suivant.
Exemple. Soient les suites définies par

$$\forall n \in \mathbb{N},\ u_n = 1 - \frac{1}{n+1} \text{ et } v_n = 1 + \frac{1}{n+1}.$$

Nous avons, pour tout entier naturel n, $u_n < v_n$ et $\lim\limits_{n\to+\infty} u_n = \lim\limits_{n\to+\infty} v_n = 1$.

Encadrement par des suites de même limite

Proposition 8.17 ***(« théorème des gendarmes »)*** *Soient $(u_n)_{n\in\mathbb{N}}$, $(v_n)_{n\in\mathbb{N}}$ et $(w_n)_{n\in\mathbb{N}}$ trois suites réelles telles que*

$$\exists N \in \mathbb{N},\ \forall n \in \mathbb{N},\ (n \geq N \Rightarrow u_n \leq v_n \leq w_n).$$

Si $(u_n)_{n\in\mathbb{N}}$ et $(w_n)_{n\in\mathbb{N}}$ convergent vers la même limite l, alors la suite $(v_n)_{n\in\mathbb{N}}$ converge aussi vers l.

DÉMONSTRATION. Soit ε un réel strictement positif. Puisque $(u_n)_{n\in\mathbb{N}}$ et $(w_n)_{n\in\mathbb{N}}$ convergent toutes deux vers l, il existe des entiers N et N' tels que

$$\forall n \in \mathbb{N},\ (n \geq N \Rightarrow |u_n - l| \leq \varepsilon) \text{ et } (n \geq N' \Rightarrow |w_n - l| \leq \varepsilon).$$

En notant $N'' = \max(N, N')$, nous avons

$$\forall n \in \mathbb{N},\ n \geq N'' \Rightarrow \begin{cases} u_n \leq v_n \leq w_n \\ |u_n - l| \leq \varepsilon \\ |w_n - l| \leq \varepsilon \end{cases} \Rightarrow -\varepsilon \leq u_n - l \leq v_n - l \leq w_n - l \leq \varepsilon \Rightarrow |v_n - l| \leq \varepsilon,$$

ce qui montre la convergence de $(v_n)_{n\in\mathbb{N}}$ vers l. □

Exemple. Étant donné un réel x, soit la suite $(u_n)_{n\in\mathbb{N}}$ définie par

$$\forall n \in \mathbb{N}^*,\ u_n = \frac{\cos(x)}{n}.$$

On a $-1 \leq \cos x \leq 1$, d'où, $\forall n \in \mathbb{N}^*$, $-\frac{1}{n} \leq u_n \leq \frac{1}{n}$. Comme $\lim_{n\to+\infty}\left(-\frac{1}{n}\right) = \lim_{n\to+\infty}\frac{1}{n} = 0$, on a finalement $\lim_{n\to+\infty} u_n = 0$.

Proposition 8.18 *Soient $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ deux suites réelles telles que $u_n \leq v_n$ pour tout entier n. On a*

1) Si $\lim_{n\to+\infty} u_n = +\infty$ alors $\lim_{n\to+\infty} v_n = +\infty$.

2) Si $\lim_{n\to+\infty} v_n = -\infty$ alors $\lim_{n\to+\infty} u_n = -\infty$.

DÉMONSTRATION. Prouvons 1). Soit $A \in \mathbb{R}$. Comme $\lim_{n\to+\infty} u_n = +\infty$, il existe un entier N tel que

$$\forall n \in \mathbb{N},\ (n \geq N \Rightarrow u_n \geq A),$$

et, *a fortiori*, $v_n \geq A$ d'après l'hypothèse, ce qui implique que $\lim_{n\to+\infty} v_n = +\infty$.
La preuve de 2) s'obtient de manière analogue à celle de 1). □

8.2.6 Propriétés algébriques des suites convergentes

Proposition 8.19 *Soient $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ deux suites numériques, $\lambda \in \mathbb{K}$ et $(l, l') \in \mathbb{K}^2$. On a*

1) $\lim_{n\to+\infty} u_n = l \Rightarrow \lim_{n\to+\infty} |u_n| = |l|$,

2) $\lim_{n\to+\infty} u_n = l$ et $\lim_{n\to+\infty} v_n = l' \Rightarrow \lim_{n\to+\infty} (u_n + v_n) = l + l'$,

3) $\lim_{n\to+\infty} u_n = l \Rightarrow \lim_{n\to+\infty} \lambda u_n = \lambda l$,

4) $\lim_{n\to+\infty} u_n = 0$ et $(v_n)_{n\in\mathbb{N}}$ bornée $\Rightarrow \lim_{n\to+\infty} u_n v_n = 0$,

5) $\lim_{n\to+\infty} u_n = l$ et $\lim_{n\to+\infty} v_n = l' \Rightarrow \lim_{n\to+\infty} u_n v_n = ll'$,

6) $\lim_{n\to+\infty} v_n = l'$ et $l' \neq 0 \Rightarrow \frac{1}{v_n}$ est défini à partir d'un certain rang et $\lim_{n\to+\infty} \frac{1}{v_n} = \frac{1}{l'}$,

7) $\lim_{n\to+\infty} u_n = l$ et $\lim_{n\to+\infty} v_n = l'$ et $l' \neq 0 \Rightarrow \frac{u_n}{v_n}$ est défini à partir d'un certain rang et $\lim_{n\to+\infty} \frac{u_n}{v_n} = \frac{l}{l'}$.

DÉMONSTRATION.

1) Soit $\varepsilon > 0$. Puisque la suite $(u_n)_{n\in\mathbb{N}}$ tend vers l, il existe $N \in \mathbb{N}$ tel que
$$\forall n \in \mathbb{N},\ (n \geq N \Rightarrow |u_n - l| \leq \varepsilon).$$
Or, on a, $\forall n \in \mathbb{N}$, l'inégalité $||u_n| - |l|| \leq |u_n - l|$, dont on déduit que
$$\forall n \in \mathbb{N},\ (n \geq N \Rightarrow ||u_n| - |l|| \leq \varepsilon),$$
et donc $\lim\limits_{n\to+\infty} |u_n| = |l|$.

2) Soit $\varepsilon > 0$. Puisque les suites $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ convergent respectivement vers l et l', il existe des entiers N et N' tels que
$$\forall n \in \mathbb{N},\ \left(n \geq N \Rightarrow |u_n - l| \leq \frac{\varepsilon}{2}\right) \text{ et } \left(n \geq N' \Rightarrow |v_n - l'| \leq \frac{\varepsilon}{2}\right).$$
En notant $N'' = \max(N, N')$, nous avons
$$\forall n \in \mathbb{N},\ \left(n \geq N'' \Rightarrow |(u_n + v_n) - (l + l')| = |(u_n - l) + (v_n - l')| \leq |u_n - l| + |v_n - l'| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\right),$$
d'où $\lim\limits_{n\to+\infty} (u_n + v_n) = l + l'$.

3) Soit $\varepsilon > 0$. Puisque la suite $(u_n)_{n\in\mathbb{N}}$ tend vers l, il existe $N \in \mathbb{N}$ tel que
$$\forall n \in \mathbb{N},\ \left(n \geq N \Rightarrow |u_n - l| \leq \frac{\varepsilon}{|\lambda| + 1}\right),$$
d'où, $\forall n \in \mathbb{N}$, $\left(n \geq N \Rightarrow |\lambda u_n - \lambda l| = |\lambda||u_n - l| \leq \dfrac{|\lambda|\varepsilon}{|\lambda| + 1} \leq \varepsilon\right)$, et donc $\lim\limits_{n\to+\infty} \lambda u_n = \lambda l$.

4) Par hypothèse, il existe $M \in \mathbb{R}_+$ tel que, $\forall n \in \mathbb{N}$, $|v_n| \leq M$. Soit $\varepsilon > 0$. Puisque la suite $(u_n)_{n\in\mathbb{N}}$ tend vers 0, il existe $N \in \mathbb{N}$ tel que
$$\forall n \in \mathbb{N},\ \left(n \geq N \Rightarrow |u_n| \leq \frac{\varepsilon}{M + 1}\right).$$
Nous avons alors, $\forall n \in \mathbb{N}$, $\left(n \geq N \Rightarrow |u_n v_n| = |u_n||v_n| \leq \dfrac{M\varepsilon}{M + 1} \leq \varepsilon\right)$ et donc $\lim\limits_{n\to+\infty} u_n v_n = 0$.

5) Notons, pour tout $n \in \mathbb{N}$, $w_n = u_n - l$. Nous avons
$$\forall n \in \mathbb{N},\ u_n v_n = (w_n + l)v_n = w_n v_n + l v_n.$$
D'après 3), $\lim\limits_{n\to+\infty} l v_n = ll'$. D'autre part, $\lim\limits_{n\to+\infty} w_n = 0$ et $(v_n)_{n\in\mathbb{N}}$ est bornée puisque $(v_n)_{n\in\mathbb{N}}$ est convergente, donc $\lim\limits_{n\to+\infty} w_n v_n = 0$ d'après 4). Finalement, on a $\lim\limits_{n\to+\infty} u_n v_n = ll'$ d'après 2).

6) Puisque $(v_n)_{n\in\mathbb{N}}$ converge vers l', la suite $(|v_n|)_{n\in\mathbb{N}}$ converge vers $|l'|$. Il existe donc un entier N tel que, pour tout entier n, $\left(n \geq N \Rightarrow |v_n| \geq \dfrac{|l'|}{2}\right)$ (il suffit de choisir $\varepsilon = \dfrac{|l'|}{2}$). En particulier, $\forall n \in \mathbb{N}$, $(n \geq N \Rightarrow v_n \neq 0)$, et la suite $\left(\dfrac{1}{v_n}\right)_{n\in\mathbb{N}}$ est donc définie. Nous avons alors, pour tout entier n tel que $n \geq N$
$$0 \leq \left|\frac{1}{v_n} - \frac{1}{l'}\right| = \frac{|v_n - l'|}{|v_n||l'|} \leq \frac{2}{|l'|^2}|v_n - l'|.$$
Comme $\lim\limits_{n\to+\infty} v_n = l'$, on déduit $\lim\limits_{n\to+\infty} \left(\dfrac{2}{|l'|^2}|v_n - l'|\right) = 0$, puis que $\lim\limits_{n\to+\infty} \left|\dfrac{1}{v_n} - \dfrac{1}{l'}\right| = 0$, soit encore $\lim\limits_{n\to+\infty} \dfrac{1}{v_n} = \dfrac{1}{l'}$.

7) Il suffit d'appliquer 5) et 6) en remarquant que $\dfrac{u_n}{v_n} = u_n \dfrac{1}{v_n}$.

□

Proposition 8.20 *Soient $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ deux suites réelles.*

1) Si $\lim\limits_{n\to+\infty} u_n = +\infty$ et $(v_n)_{n\in\mathbb{N}}$ est minorée, alors $\lim\limits_{n\to+\infty} (u_n + v_n) = +\infty$.
En particulier, on a
— $\lim\limits_{n\to+\infty} u_n = +\infty$ *et* $\lim\limits_{n\to+\infty} v_n = +\infty \Rightarrow \lim\limits_{n\to+\infty} (u_n + v_n) = +\infty$,
— $\lim\limits_{n\to+\infty} u_n = +\infty$ *et* $\lim\limits_{n\to+\infty} v_n = l' \Rightarrow \lim\limits_{n\to+\infty} (u_n + v_n) = +\infty$.

2) Si $\lim\limits_{n\to+\infty} u_n = +\infty$ et $\big(\exists C \in \mathbb{R}_+^*,\ \exists N \in \mathbb{N},\ \forall n \in \mathbb{N},\ (n \geq N \Rightarrow v_n \geq C)\big)$, alors $\lim\limits_{n\to+\infty} u_n v_n = +\infty$.
En particulier, on a

— $\lim_{n\to+\infty} u_n = +\infty$ *et* $\lim_{n\to+\infty} v_n = +\infty \Rightarrow \lim_{n\to+\infty} u_n v_n = +\infty$,
— $\lim_{n\to+\infty} u_n = +\infty$ *et* $\lim_{n\to+\infty} v_n = l' \in \mathbb{R}_+^* \Rightarrow \lim_{n\to+\infty} u_n v_n = +\infty$.

3) $\lim_{n\to+\infty} u_n = +\infty \Rightarrow \lim_{n\to+\infty} \dfrac{1}{u_n} = 0$.

4) Si $\lim_{n\to+\infty} u_n = 0$ *et si* $(\exists N \in \mathbb{N},\ \forall n \in \mathbb{N},\ (n \geq N \Rightarrow u_n > 0))$*, alors* $\lim_{n\to+\infty} \dfrac{1}{u_n} = +\infty$.

DÉMONSTRATION.

1) Par hypothèse, il existe $m \in \mathbb{R}$ tel que
$$\forall n \in \mathbb{N},\ v_n \geq m.$$
Soit $A > 0$. Puisque $\lim_{n\to+\infty} u_n = +\infty$, il existe $N \in \mathbb{N}$ tel que, $\forall n \in \mathbb{N}$, $(n \geq N \Rightarrow u_n \geq A - m)$. Nous avons alors
$$\forall n \in \mathbb{N},\ (n \geq N \Rightarrow u_n + v_n \geq (A-m)+m),$$
et donc $\lim_{n\to+\infty} (u_n + v_n) = +\infty$.

2) Soit $A > 0$. Puisque $\lim_{n\to+\infty} u_n = +\infty$, il existe $N' \in \mathbb{N}$ tel que, $\forall n \in \mathbb{N}$, $\left(n \geq N' \Rightarrow u_n \geq \dfrac{A}{C}\right)$.
En notant $N'' = \max(N, N')$, nous avons alors
$$\forall n \in \mathbb{N},\ \left(n \geq N'' \Rightarrow \left(u_n \geq \frac{A}{C} \text{ et } v_n \geq C\right) \Rightarrow u_n v_n \geq A\right),$$
et donc $\lim_{n\to+\infty} u_n v_n = +\infty$.

3) Soit $\varepsilon > 0$. Puisque $\lim_{n\to+\infty} u_n = +\infty$, il existe $N \in \mathbb{N}$ tel que
$$\forall n \in \mathbb{N},\ \left(n \geq N \Rightarrow u_n \geq \frac{1}{\varepsilon}\right),$$
d'où $\lim_{n\to+\infty} \dfrac{1}{u_n} = 0$.

4) Soit $A > 0$. Puisque $\lim_{n\to+\infty} u_n = 0$, il existe $N' \in \mathbb{N}$ tel que, $\forall n \in \mathbb{N}$, $\left(n \geq N' \Rightarrow |u_n| \geq \dfrac{1}{A}\right)$.
En notant $N'' = \max(N, N')$, nous avons alors
$$\forall n \in \mathbb{N},\ \left(n \geq N'' \Rightarrow \left(|u_n| \geq \frac{1}{A} \text{ et } u_n \geq 0\right) \Rightarrow \frac{1}{u_n} \geq A\right),$$
et donc $\lim_{n\to+\infty} \dfrac{1}{u_n} = +\infty$.

□

Nous résumons certains ces résultats dans les tableaux suivants.

$(v_n)_{n\in\mathbb{N}}$ \ $(u_n)_{n\in\mathbb{N}}$	l	$+\infty$	$-\infty$	PL
l'	$l+l'$	$+\infty$	$-\infty$	PL
$+\infty$	$+\infty$	$+\infty$	?	?
$-\infty$	$-\infty$	?	$-\infty$	?
PL	PL	?	?	?

Tableau des limites possibles pour la suite $(u_n + v_n)_{n\in\mathbb{N}}$ en fonction des limites respectives des suites réelles $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$.

Les lettres PL signifient que la suite considérée n'a pas de limite. Le symbole « ? » correspond à une *forme indéterminée* ; tout est possible :

— la limite est finie ; par exemple, pour tout $n \in \mathbb{N}^*$, $u_n = n$ et $v_n = \dfrac{1}{n} - n$, $\lim_{n\to+\infty} (u_n + v_n) = 0$,
— la limite est infinie ; par exemple, pour tout $n \in \mathbb{N}$, $u_n = n^2$ et $v_n = -n$, $\lim_{n\to+\infty} (u_n + v_n) = +\infty$,

— il n'y a pas de limite ; par exemple, pour tout $n \in \mathbb{N}$, $u_n = n + (-1)^n$ et $v_n = -n$, alors $(u_n + v_n) = (-1)^n$ qui n'a pas de limite.

$(v_n)_{n\in\mathbb{N}}$ \ $(u_n)_{n\in\mathbb{N}}$	$l > 0$	$l = 0$	$l < 0$	$+\infty$	$-\infty$	PL
$l' > 0$	ll'	0	ll'	$+\infty$	$-\infty$	PL
$l' = 0$	0	0	0	?	?	?
$l' < 0$	ll'	0	ll'	$-\infty$	$+\infty$	PL
$+\infty$	$+\infty$	?	$-\infty$	$+\infty$	$-\infty$	?
$-\infty$	$-\infty$	?	$+\infty$	$-\infty$	$+\infty$	?
PL	PL	?	PL	?	?	?

Tableau des limites possibles pour la suite $(u_n v_n)_{n\in\mathbb{N}}$ en fonction des limites respectives des suites réelles $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$.

Là encore, les lettres PL signifient que la suite considérée n'a pas de limite et le symbole « ? » correspond à une forme indéterminée. Différents cas sont possibles :

— la limite est finie : par exemple, pour tout $n \in \mathbb{N}^*$, $u_n = n$ et $v_n = \dfrac{1}{n}$, $\lim\limits_{n\to+\infty} u_n v_n = 1$,

— la limite est infinie : par exemple, pour tout $n \in \mathbb{N}$, $u_n = n^2$ et $v_n = \dfrac{1}{n}$, $\lim\limits_{n\to+\infty} u_n v_n = +\infty$,

— il n'y a pas de limite : par exemple, pour tout $n \in \mathbb{N}^*$, $u_n = n$ et $v_n = \dfrac{(\sin n)^2}{n}$, alors $u_n v_n = (\sin n)^2$ qui n'a pas de limite.

8.3 Existence de limites

Nous rassemblons dans cette section plusieurs résultats relatifs à l'existence de limites.

Proposition 8.21 *1) Toute suite réelle croissante et majorée est convergente.*
2) Toute suite réelle décroissante et minorée est convergente.

DÉMONSTRATION.

1) Soit $(u_n)_{n\in\mathbb{N}}$ une suite réelle croissante et majorée. L'ensemble $\{u_k \mid k \in \mathbb{N}\}$ des termes de la suite est une partie de $\mathbb{R}$ non vide et majorée, qui admet donc une borne supérieure, notée l. On a alors $u_n \leq l$, pour tout entier n, et pour tout réel ε strictement positif, $l - \varepsilon$ n'est pas un majorant de l'ensemble $\{u_k \mid k \in \mathbb{N}\}$. Il existe alors $N \in \mathbb{N}$ tel que
$$l - \varepsilon \leq u_N \leq l.$$
La suite $(u_n)_{n\in\mathbb{N}}$ étant croissante, on en déduit
$$\forall n \in \mathbb{N},\ (n \geq N \Rightarrow l - \varepsilon \leq u_N \leq u_n \leq l),$$
et donc $|x_n - l| \leq \varepsilon$. On en conclut que $(u_n)_{n\in\mathbb{N}}$ converge vers l.
2) Il suffit d'appliquer le résultat 1) à la suite $(-u_n)_{n\in\mathbb{N}}$.

□

Proposition 8.22 *1) Toute suite réelle croissante et non majorée tend vers $+\infty$.*
2) Toute suite réelle décroissante et non minorée tend vers $-\infty$.

DÉMONSTRATION.

1) Soit $(u_n)_{n\in\mathbb{N}}$ une suite réelle croissante non majorée. L'ensemble $\{u_k \mid k \in \mathbb{N}\}$ des termes de la suite est une partie de $\mathbb{R}$ non majorée, et donc, quel que soit $A > 0$, il existe un entier N tel que $u_N > A$. La suite $(u_n)_{n\in\mathbb{N}}$ étant croissante, on en déduit
$$\forall n \in \mathbb{N},\ (n \geq N \Rightarrow u_n \geq u_N > A),$$
d'où $\lim\limits_{n\to+\infty} u_n = +\infty$.

2) Il suffit d'appliquer le résultat 1) à la suite $(-u_n)_{n\in\mathbb{N}}$.

□

Proposition 8.23 *Deux suites adjacentes convergent et ont même limite.*

DÉMONSTRATION. Soient $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ deux suites adjacentes. Supposons $(u_n)_{n\in\mathbb{N}}$ croissante et $(v_n)_{n\in\mathbb{N}}$ décroissante. La suite $(u_n-v_n)_{n\in\mathbb{N}}$ est donc croissante et tend par hypothèse vers 0, on en déduit que c'est une suite négative et, $\forall n\in\mathbb{N}$, $u_n\le v_n$. De plus, $\forall n\in\mathbb{N}$, $u_0\le u_n$ et $v_n\le v_0$. En combinant ces inégalités, nous obtenons

$$\forall n\in\mathbb{N},\ u_0\le u_n\le v_n\le v_0.$$

La suite $(u_n)_{n\in\mathbb{N}}$ est alors croissante et majorée par v_0, c'est donc une suite convergente. De même, la suite $(v_n)_{n\in\mathbb{N}}$ est décroissante et minorée par u_0, donc $(v_n)_{n\in\mathbb{N}}$ est convergente.
D'autre part, on a $\lim\limits_{n\to+\infty}(u_n-v_n)=0$ et, comme les deux suites sont convergentes, on en déduit $\lim\limits_{n\to+\infty}u_n=\lim\limits_{n\to+\infty}v_n$. □

Nous pouvons à présent établir le

Théorème 8.24 ***(« théorème des segments emboîtés »)*** *Soit $([a_n,b_n])_{n\in\mathbb{N}}$ une suite de segments emboîtés (c'est-à-dire, $\forall n\in\mathbb{N}$, $[a_{n+1},b_{n+1}]\subset[a_n,b_n]$) telle que $\lim\limits_{n\to+\infty}(b_n-a_n)=0$. Alors l'intersection $\bigcap_{n\in\mathbb{N}}[a_n,b_n]$ est un singleton.*

DÉMONSTRATION. Notons que les suites $(a_n)_{n\in\mathbb{N}}$ et $(b_n)_{n\in\mathbb{N}}$ sont adjacentes. On en déduit que qu'elles convergent vers une même limite l et l'on a, $\forall n\in\mathbb{N}$, $a_n\le l\le b_n$ donc $l\in[a_n,b_n]$ pour tout entier n, et, par suite, $l\in\bigcap_{n\in\mathbb{N}}[a_n,b_n]$.
D'autre part si $l'\in\bigcap_{n\in\mathbb{N}}[a_n,b_n]$, alors $l'\in[a_n,b_n]$ pour tout entier n. Comme on a également $l\in[a_n,b_n]$ pour tout entier n, nous obtenons $\forall n\in\mathbb{N}$, $b_n-a_n\ge|l-l'|$. En faisant tendre n vers l'infini, nous trouvons $l=l'$.
En conclusion, $\bigcap_{n\in\mathbb{N}}[a_n,b_n]=\{l\}$. □

Théorème 8.25 ***(« théorème de Bolzano [2]–Weierstrass [3] »)*** *De toute suite réelle bornée, on peut extraire une suite convergente.*

DÉMONSTRATION. Soit $(u_n)_{n\in\mathbb{N}}$ une suite réelle bornée. Il existe alors deux réels a_0 et b_0 tels que, pour tout entier n, $a_0\le u_n\le b_0$.
Il est clair que $\{k\in\mathbb{N}\mid u_k\in[a_0,b_0]\}=\mathbb{N}$ est infini.
Soit à présent $n\in\mathbb{N}$; nous supposons défini le couple $(a_n,b_n)\in\mathbb{R}^2$ tel que

i) $a_n\le b_n$,
ii) $\{k\in\mathbb{N}\mid u_k\in[a_n,b_n]\}$ est infini,
iii) $b_n-a_n=\dfrac{1}{2^n}(b_0-a_0)$.

En considérant alors le milieu $\frac{a_n+b_n}{2}$ de l'intervalle fermé $[a_n,b_n]$, il est clair que l'un des deux intervalles $\left[a_n,\frac{a_n+b_n}{2}\right]$, $\left[\frac{a_n+b_n}{2},b_n\right]$ est tel que l'ensemble des entiers k tels que u_k soit dans cet intervalle est infini. Il existe donc $(a_{n+1},b_{n+1})\in\mathbb{R}^2$ tel que

i) $a_{n+1}\le b_{n+1}$,
ii) $\{k\in\mathbb{N}\mid u_k\in[a_{n+1},b_{n+1}]\}$ est infini,
iii) $b_{n+1}-a_{n+1}=\dfrac{1}{2}(b_n-a_n)=\dfrac{1}{2^{n+1}}(b_0-a_0)$.

Il est alors évident que les intervalles $[a_n,b_n]$, $n\in\mathbb{N}$, forment une suite de segments emboîtés dont la longueur tend vers 0. On en déduit (d'après le théorème 8.24) qu'ils ont un seul point commun $l\in\mathbb{R}$, qui est la limite commune de $(a_n)_{n\in\mathbb{N}}$ et $(b_n)_{n\in\mathbb{N}}$.
D'autre part, il est aisé de construire une extractrice σ telle que $\sigma(0)=0$ et telle qu'il existe, pour tout entier n, un entier k tel que si $k>\sigma(n)$ alors $u_k\in[a_n,b_n]$ et $\sigma(n+1)=k$. Les inégalités $a_n\le u_{\sigma(n)}\le b_n$, valables pour tout entier n, montrent alors que la suite $(u_{\sigma(n)})_{n\in\mathbb{N}}$ tend vers l. □

Théorème 8.26 *Toute suite de Cauchy à valeurs réelles est convergente (on dit que $\mathbb{R}$ est* ***complet****).*

2. Bernardus Placidus Johann Nepomuk Bolzano (5 octobre 1781 - 18 décembre 1848) était un mathématicien, théologien et philosophe bohémien de langue allemande. Ses travaux portèrent essentiellement sur les fonctions et la théorie des nombres et il est considéré comme un des fondateurs de la logique moderne.

3. Karl Theodor Wilhelm Weierstraß (31 octobre 1815 - 19 février 1897) était un mathématicien allemand, souvent cité comme le « père de l'analyse moderne ». On lui doit l'introduction de plusieurs définitions et de formulations rigoureuses, comme les notions de limite et de continuité, et ses contributions au développement d'outils théoriques en analyse ouvrirent la voie à l'étude du calcul des variations telle que nous la connaissons aujourd'hui.

DÉMONSTRATION. Soit $(u_n)_{n\in\mathbb{N}}$ une suite réelle de Cauchy. D'après la proposition 8.6, nous savons que $(u_n)_{n\in\mathbb{N}}$ est bornée. Il existe alors, en vertu du théorème de Bolzano–Weierstrass, une suite extraite $(u_{\sigma(n)})_{n\in\mathbb{N}}$ qui converge vers une limite l. Montrons que la suite $(u_n)_{n\in\mathbb{N}}$ converge vers cette limite.
Soit $\varepsilon > 0$. Il existe des entiers N et N' tels que

$$\forall n\in\mathbb{N},\ \left(n\geq N \Rightarrow |u_{\sigma(n)}-l|\leq \frac{\varepsilon}{2}\right)\ (\text{car } (u_{\sigma(n)})_{n\in\mathbb{N}} \text{ converge vers } l)$$

et

$$\forall (p,q)\in\mathbb{N}^2,\ \left(p\geq N',\ q\geq N' \Rightarrow |u_p-u_q|\leq \frac{\varepsilon}{2}\right)\ (\text{car } (u_n)_{n\in\mathbb{N}} \text{ est une suite de Cauchy}).$$

Notons $N''=\max(N,N')$. Si $n\geq N''$, nous avons d'une part $\sigma(n)\geq n\geq N'$, d'où $|u_n-u_{\sigma(n)}|\leq \frac{\varepsilon}{2}$, et d'autre part $n\geq N$, d'où $|u_{\sigma(n)}-l|\leq \frac{\varepsilon}{2}$. En combinant ces deux inégalités, nous obtenons

$$\forall n\in\mathbb{N},\ \left(n\geq N'' \Rightarrow |u_n-l|\leq |u_n-u_{\sigma(n)}|+|u_{\sigma(n)}-l|\leq \frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon\right),$$

ce qui permet de conclure que la suite $(u_n)_{n\in\mathbb{N}}$ est convergente. □

En alliant ce théorème à la proposition 8.15, nous en déduisons qu'une suite réelle converge si et seulement si elle est une suite de Cauchy.

Nous terminons cette section avec un résultat relatif à la suite des moyennes arithmétiques des premiers termes d'un suite convergente.

Définition 8.27 ***(« moyenne de Cesàro*** [4] ***»)*** *Soit une suite $(u_n)_{n\in\mathbb{N}^*}$. On appelle* ***moyenne de Cesàro*** *de $(u_n)_{n\in\mathbb{N}^*}$ la suite $(s_n)_{n\in\mathbb{N}^*}$ de terme général*

$$\forall n\in\mathbb{N}^*,\ s_n=\frac{1}{n}\sum_{k=1}^{n}u_k.$$

Théorème 8.28 ***(« lemme de Cesàro »)*** *La moyenne de Cesàro d'une suite convergente de limite l converge vers l.*

DÉMONSTRATION. Soit $(u_n)_{n\in\mathbb{N}^*}$ une suite convergente de limite l et $(s_n)_{n\in\mathbb{N}^*}$ sa moyenne de Cesàro. Soit $\varepsilon>0$. Il existe un entier naturel non nul N tel que

$$\forall n\in\mathbb{N}^*,\ \left(n\geq N \Rightarrow |u_n-l|\leq \frac{\varepsilon}{2}\right).$$

Pour $n\geq N$, on a donc

$$|s_n-l|=\left|\frac{1}{n}\sum_{k=1}^{n}(u_k-l)\right|\leq \frac{1}{n}\sum_{k=1}^{n}|u_k-l|\leq \frac{1}{n}\sum_{k=1}^{N-1}|u_k-l|+\frac{1}{n}\sum_{k=N}^{n}|u_k-l|\leq \frac{1}{n}\sum_{k=1}^{N-1}|u_k-l|+\frac{n-N+1}{n}\,\frac{\varepsilon}{2}.$$

La somme $\sum_{k=1}^{N-1}|u_k-l|$ étant finie, on a

$$\lim_{n\to+\infty}\frac{1}{n}\sum_{k=1}^{N-1}|u_k-l|=0,$$

et il existe donc un entier naturel non nul N' tel que

$$\forall n\in\mathbb{N}^*,\ \left(n\geq N' \Rightarrow \frac{1}{n}\sum_{k=1}^{N-1}|u_k-l|\leq \frac{\varepsilon}{2}\right).$$

Notons $N''=\max(N,N')$. Il vient alors

$$\forall n\in\mathbb{N}^*,\ \left(n\geq N'' \Rightarrow |s_n-l|\leq \frac{\varepsilon}{2}+\frac{n-N+1}{n}\,\frac{\varepsilon}{2}\leq\varepsilon\right),$$

ce qui montre que $(s_n)_{n\in\mathbb{N}^*}$ converge vers l. □

4. Ernesto Cesàro (12 mars 1859 - 12 septembre 1906) était un mathématicien italien, connu pour ses contributions à la géométrie différentielle et son procédé de sommation des séries divergentes.

8.4 Quelques suites particulières

8.4.1 Suite arithmétique

Définition 8.29 *Une suite $(u_n)_{n\in\mathbb{N}}$ est dite **arithmétique de raison** r si et seulement s'il existe un scalaire $r \in \mathbb{K}$ tel que, pour tout entier naturel n, $u_{n+1} = u_n + r$.*

Si $(u_n)_{n\in\mathbb{N}}$ est une suite arithmétique de raison r, on a

$$\forall n \in \mathbb{N},\ u_n = u_0 + nr.$$

Si $r = 0$, la suite est constante. Si $\mathbb{K} = \mathbb{R}$, la suite est strictement croissante si $r > 0$ et strictement décroissante si $r < 0$.

Proposition 8.30 ***(somme des termes d'une suite arithmétique)*** *La somme des m premiers termes d'une suite arithmétique $(u_n)_{n\in\mathbb{N}}$ de raison r est*

$$\forall m \in \mathbb{N}^*,\ S_m = \sum_{k=0}^{m-1} u_k = m\,u_0 + \frac{m(m-1)}{2} r = \frac{m}{2}(u_0 + u_{m-1}).$$

8.4.2 Suite géométrique

Définition 8.31 *Une suite $(u_n)_{n\in\mathbb{N}}$ est dite **géométrique de raison** q si et seulement s'il existe un scalaire $q \in \mathbb{K}$ tel que, pour tout entier naturel n, $u_{n+1} = q\,u_n$.*

Si $(u_n)_{n\in\mathbb{N}}$ est une suite géométrique de raison q, on a

$$\forall n \in \mathbb{N},\ u_n = q^n\,u_0.$$

La suite $(u_n)_{n\in\mathbb{N}}$ est constante si $q = 1$ et stationnaire en 0 (à partir de $n = 1$) si $q = 0$.
Si $\mathbb{K} = \mathbb{R}$ et si $q > 0$, la suite $(u_n)_{n\in\mathbb{N}}$ garde un signe constant et est monotone. Plus précisément,
- si $u_0 > 0$ et $q > 1$, la suite est positive strictement croissante,
- si $u_0 > 0$ et $0 < q < 1$, la suite est positive strictement décroissante,
- si $u_0 < 0$ et $q > 1$, la suite est négative strictement décroissante,
- si $u_0 < 0$ et $0 < q < 1$, la suite est négative strictement croissante.

Si $\mathbb{K} = \mathbb{R}$ et si $q < 0$, alors, pour tout entier n, u_n et u_{n+1} sont de signes contraires et la suite $(u_n)_{n\in\mathbb{N}}$ n'est donc pas monotone.

Proposition 8.32 *Soit $(u_n)_{n\in\mathbb{N}}$ une suite géométrique réelle de premier terme $u_0 \in \mathbb{R}^*$ et de raison $q \in \mathbb{R}$. On a*

i) Si $|q| < 1$, $\lim\limits_{n\to+\infty} u_n = 0$.

ii) Si $q > 1$, $\lim\limits_{n\to+\infty} u_n = \begin{cases} +\infty & \text{si } u_0 > 0 \\ -\infty & \text{si } u_0 < 0 \end{cases}$.

iii) Si $q = 1$, $\lim\limits_{n\to+\infty} u_n = u_0$.

iv) Si $q \leq -1$, $(u_n)_{n\in\mathbb{N}}$ n'a pas de limite.

DÉMONSTRATION.

i) Si $q = 0$, la suite est stationnaire en 0 à partir du rang $n = 1$. Sinon, $0 < |q| < 1 \Rightarrow \dfrac{1}{|q|} > 1 \Rightarrow \exists h \in \mathbb{R}_+^*,\ \dfrac{1}{|q|} = 1 + h$. On a donc

$$\forall n \in \mathbb{N}^*,\ \left(\frac{1}{|q|}\right)^n = (1+h)^n = \sum_{k=0}^{n} C_n^k h^k \geq 1 + nh,$$

en utilisant la formule du binôme de Newton. Par ailleurs, on a $|u_n| = q^n |u_0| \leq \dfrac{|u_0|}{1+nh}$, d'où $\lim\limits_{n\to+\infty} u_n = 0$.

ii) La démonstration est similaire à celle de i).

iii) et iv) sont évidentes. □

Proposition 8.33 ***(somme des termes d'une suite géométrique)*** *La somme des m premiers termes d'une suite géométrique $(u_n)_{n\in\mathbb{N}}$ de raison q est*

$$\forall m \in \mathbb{N}^*,\ S_m = \sum_{k=0}^{m-1} u_k = u_0 \sum_{k=0}^{m-1} q^k = u_0 \frac{1-q^m}{1-q} \text{ si } q \neq 1,\ S_m = m\,u_0 \text{ si } q = 1.$$

8.4.3 Suite arithmético-géométrique

Définition 8.34 *Une suite $(u_n)_{n\in\mathbb{N}}$ est dite **arithmético-géométrique** si et seulement s'il existe des scalaires q et r tels que, pour tout entier naturel n, $u_{n+1} = q\,u_n + r$.*

Remarquons qu'une telle suite est arithmétique si $q = 1$ et géométrique si $r = 0$.

Méthode d'étude d'une suite arithmético-géométrique par utilisation d'un point fixe. Supposons $q \neq 1$. Soit α l'unique scalaire vérifiant $\alpha = q\,\alpha + r$, c'est-à-dire $\alpha = \dfrac{r}{1-q}$ (on dit que α est un *point fixe* de la fonction $f(x) = q\,x + r$). Nous avons

$$\forall n \in \mathbb{N}^*,\ u_n - \alpha = q\,u_{n-1} + r - (q\,\alpha + r) = q(u_{n-1} - \alpha).$$

La suite $(u_n - \alpha)_{n\in\mathbb{N}}$ est donc une suite géométrique de raison q. Ceci implique que $u_n - \alpha = q^n(u_0 - \alpha)$, soit encore

$$\forall n \in \mathbb{N},\ u_n = q^n(u_0 - \alpha) + \alpha.$$

On en déduit que, si $u_0 = \alpha$, la suite $(u_n)_{n\in\mathbb{N}}$ est constante et vaut α. Si $u_0 \neq \alpha$, on a alors

$$\lim_{n\to+\infty} u_n = \alpha \text{ si } |q| < 1 \text{ et } \lim_{n\to+\infty} |u_n| = +\infty \text{ si } |q| > 1.$$

8.4.4 Suite définie par récurrence

Suite récurrente d'ordre un

Définition 8.35 *Soient I un intervalle fermé de $\mathbb{R}$ et $f : I \to \mathbb{R}$ une application. On suppose que $f(I) \subset I$. La suite $(u_n)_{n\in\mathbb{N}}$ définie par $u_0 \in I$ et la relation de récurrence*

$$\forall n \in \mathbb{N},\ u_{n+1} = f(u_n),$$

*est appelée **suite récurrente**.*

Cette suite est bien définie car, pour tout entier n, on a $u_n \in I$ et $f(I) \subset I$. On remarque que si f est une application affine à coefficients constants, la suite récurrente est une suite arithmético-géométrique.
Pour l'étude de suites récurrentes du type $u_{n+1} = f(u_n)$, nous utiliserons les propriétés élémentaires des applications continues (chapitre 9) et des applications dérivables (chapitre 10).

Sens de variation de la suite $(u_n)_{n\in\mathbb{N}}$. Supposons que f soit une application monotone sur l'intervalle fermé I.

— Si f est croissante,

$$\forall n \in \mathbb{N}^*,\ u_{n+1} - u_n = f(u_n) - f(u_{n-1}),$$

et $u_{n+1} - u_n$ a le même signe que $u_1 - u_0 = f(u_0) - u_0$.
Ainsi, $(u_n)_{n\in\mathbb{N}}$ est monotone et son sens de variation dépend de la position relative de u_0 et u_1. Il reste à voir si la suite est minorée, majorée.

— Si f est décroissante, nous remarquons que, pour tout entier naturel n, $u_{n+1} - u_n$ a le signe opposé de $u_n - u_{n-1}$.
Nous étudions alors les suites extraites $(u_{2p})_{p\in\mathbb{N}}$ et $(u_{2p+1})_{p\in\mathbb{N}}$. Pour tout $p \in \mathbb{N}$, on a

$$u_{2p+2} = f(u_{2p+1}) = (f\circ f)(u_{2p}) \text{ et } u_{2p+3} = f(u_{2p+2}) = (f\circ f)(u_{2p+1}).$$

Comme f est décroissante, l'application $f\circ f$ est croissante et donc les suites extraites $(u_{2p})_{p\in\mathbb{N}}$ et $(u_{2p+1})_{p\in\mathbb{N}}$ sont toutes deux monotones et de sens contraires.

Limite de la suite $(u_n)_{n\in\mathbb{N}}$. Supposons à présent que f est continue sur l'intervalle I. Si la suite $(u_n)_{n\in\mathbb{N}}$ converge vers $l \in I$, alors, en passant à la limite quand n tend vers l'infini dans la relation de récurrence $u_{n+1} = f(u_n)$, nous déduisons que le réel l vérifie $l = f(l)$ (on dit encore que l est un *point fixe* de l'application f).
Pour déterminer les seules limites possibles d'une suite récurrente $(u_n)_{n\in\mathbb{N}}$ de type $u_{n+1} = f(u_n)$, on pourra donc chercher à résoudre l'équation $f(l) = l$, d'inconnue $l \in I$.

Suite récurrente d'ordre supérieur

On peut également définir une suite $(u_n)_{n\in\mathbb{N}}$ par une relation de récurrence d'ordre deux (ou plus), c'est-à-dire en se donnant les deux termes initiaux de la suite et la relation de récurrence

$$\forall n \in \mathbb{N},\ u_{n+2} = f(u_n, u_{n+1}).$$

Chapitre 9

Fonctions d'une variable réelle

Nous allons dans ce chapitre étudier des *fonctions numériques d'une variable réelle*, c'est-à-dire des applications définies sur une partie D de $\mathbb{R}$ et à valeurs dans le corps $\mathbb{K}$ (avec $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$). L'ensemble D, appelé le *domaine de définition de f*, est le plus souvent une réunion d'intervalles non vides de $\mathbb{R}$. L'étude d'une fonction s'effectuant cependant intervalle par intervalle, nous nous restreindrons ici le plus souvent à des applications dont les ensembles de définition sont contenus dans un intervalle non vide de $\mathbb{R}$, noté I.

9.1 Généralités sur les fonctions

Soit I une partie non vide de $\mathbb{R}$. On note $\mathscr{F}(I,\mathbb{K})$, ou encore $\mathbb{K}^I$, l'ensemble des applications définies sur I à valeurs dans $\mathbb{K}$.
Nous rappelons que si f et g appartiennent à $\mathscr{F}(I,\mathbb{K})$, alors

$$f = g \Leftrightarrow \forall x \in I,\ f(x) = g(x).$$

9.1.1 Opérations sur les fonctions

L'ensemble $\mathscr{F}(I,\mathbb{K})$ est muni de deux lois internes :

- une addition :

$$\forall f, g \in \mathscr{F}(I,\mathbb{K}),\ \forall x \in I,\ (f+g)(x) = f(x) + g(x),$$

- une multiplication :

$$\forall f, g \in \mathscr{F}(I,\mathbb{K}),\ \forall x \in I,\ (fg)(x) = f(x)g(x),$$

et d'une loi externe :

- une multiplication par les scalaires :

$$\forall \lambda \in \mathbb{K},\ \forall f \in \mathscr{F}(I,\mathbb{K}),\ \forall x \in I,\ (\lambda f)(x) = \lambda f(x).$$

9.1.2 Relation d'ordre pour les fonctions réelles

Lorsque les fonctions considérées sont à valeur réelles, il est possible d'utiliser la relation d'ordre total usuelle sur $\mathbb{R}$ pour comparer certaines fonctions entre elles.

Définition 9.1 *On définit dans l'ensemble $\mathscr{F}(I,\mathbb{R})$ une relation, notée $\leq$, par*

$$\forall f, g \in \mathscr{F}(I,\mathbb{R}),\ (f \leq g \Leftrightarrow (\forall x \in I,\ f(x) \leq g(x))).$$

Les résultats suivants sont immédiats.

Proposition 9.2 *1) La relation $\leq$ est une relation d'ordre sur $\mathscr{F}(I,\mathbb{R})$.*

2) La relation $\leq$ est compatible avec l'addition :

$$\forall f,g,h \in \mathscr{F}(I,\mathbb{R}),\ (f \leq g \Rightarrow f+h \leq g+h).$$

3) On a

$$\forall f,g,h \in \mathscr{F}(I,\mathbb{R}),\ (f \leq g \text{ et } 0 \leq h \Rightarrow fh \leq gh).$$

Remarque. L'ordre introduit sur $\mathscr{F}(I,\mathbb{R})$ par la relation $\leq$ définie ci-dessus n'est plus total dès que la partie I n'est plus réduite à un point. Supposons en effet que I contienne deux éléments distincts a et b. Considérons alors les applications f et g de I dans $\mathbb{R}$ définies par

$$\forall x \in I,\ f(x) = \begin{cases} 1 & \text{si } x = a \\ 0 & \text{si } x \neq a \end{cases},\ g(x) = \begin{cases} 1 & \text{si } x = b \\ 0 & \text{si } x \neq b \end{cases}.$$

On n'a dans ce cas ni $f \leq g$ (car $g(a) < f(a)$), ni $g \leq f$ (car $f(b) < g(b)$). On dit que f et g ne sont pas *comparables pour* $\leq$.

9.1.3 Propriétés globales des fonctions

Parité

Dans cette sous-section, I désigne une partie de $\mathbb{R}$ *symétrique par rapport à* 0, c'est-à-dire telle que

$$\forall x \in I,\ -x \in I.$$

Définitions 9.3 *Soit $f \in \mathscr{F}(I,\mathbb{K})$. On dit que f est*

- ***paire*** *si et seulement si*
$$\forall x \in I,\ f(-x) = f(x),$$
- ***impaire*** *si et seulement si*
$$\forall x \in I,\ f(-x) = -f(x).$$

Remarques.
— Toute application constante sur I est paire.
— Si f est une application impaire et $0 \in I$, alors $f(0) = 0$.
— Une application peut n'être ni paire, ni impaire.

Périodicité

Définition 9.4 *Soient $f \in \mathscr{F}(I,\mathbb{K})$ et T un réel strictement positif. L'application f est dite* **périodique de période** *T si et seulement si*

$$\forall x \in I,\ x+T \in I \text{ et } f(x+T) = f(x).$$

Remarques.
— Toute application constante d'un intervalle $[a,+\infty[$ ou $]a,+\infty[$ ($a \in \mathbb{R}$) dans $\mathbb{R}$ est périodique de période T pour tout réel T strictement positif.
— Si f est une application périodique de période T, alors, pour tout n de $\mathbb{N}^*$, f est périodique de période nT.

Exemples.
— L'application de $\mathbb{R}$ dans $\mathbb{R}$ définie par $f(x) = x - E(x)$, où $E(x)$ désigne la partie entière du réel x, est périodique de période 1.
— Les applications sin et cos sont périodiques de période 2π sur $\mathbb{R}$.
— L'application tan est périodique de période π sur $\mathbb{R} \backslash \left\{\frac{\pi}{2} + n\pi,\ n \in \mathbb{Z}\right\}$.

Monotonie

Définition 9.5 *Soit I une partie de $\mathbb{R}$ et $f \in \mathscr{F}(I,\mathbb{R})$. On dit que f est*

- ***croissante*** *si et seulement si*
$$\forall (x,x') \in I^2,\ (x \leq x' \Rightarrow f(x) \leq f(x')),$$
- ***décroissante*** *si et seulement si*
$$\forall (x,x') \in I^2,\ (x \leq x' \Rightarrow f(x) \geq f(x')),$$
- ***strictement croissante*** *si et seulement si*
$$\forall (x,x') \in I^2,\ (x < x' \Rightarrow f(x) < f(x')),$$
- ***strictement décroissante*** *si et seulement si*
$$\forall (x,x') \in I^2,\ (x < x' \Rightarrow f(x) > f(x')),$$
- ***monotone*** *si et seulement si elle est croissante ou décroissante,*
- ***strictement monotone*** *si et seulement si elle est strictement croissante ou strictement décroissante.*

Les résultats de la proposition suivante se démontrent de façon immédiate.

Proposition 9.6 *1) Si $f, g \in \mathscr{F}(I,\mathbb{R})$ sont croissantes, alors $f+g$ est croissante.*

2) Si $f \in \mathscr{F}(I,\mathbb{R})$ est croissante et $\lambda \in \mathbb{R}_+$, alors λf est croissante.

3) Si $f, g \in \mathscr{F}(I,\mathbb{R})$ sont croissantes et positives, alors fg est croissante.

4) Si $f \in \mathscr{F}(I,\mathbb{R})$ et $g \in \mathscr{F}(J,\mathbb{R})$ sont croissantes et si $f(I) \subset J$, alors l'application composée $g \circ f \in \mathscr{F}(I,\mathbb{R})$ est croissante.

Majoration, minoration

Définitions 9.7 *Une application $f : I \to \mathbb{R}$ est dite*

- ***majorée*** *si et seulement s'il existe $M \in \mathbb{R}$ tel que*
$$\forall x \in I,\ f(x) \leq M,$$
- ***minorée*** *si et seulement s'il existe $m \in \mathbb{R}$ tel que*
$$\forall x \in I,\ m \leq f(x),$$
- ***bornée*** *si et seulement s'il existe $(m,M) \in \mathbb{R}^2$ tel que*
$$\forall x \in I,\ m \leq f(x) \leq M.$$

Remarques.

- Une application f est majorée (resp. minorée, resp. bornée) si et seulement si la partie $f(I)$ de $\mathbb{R}$ est majorée (resp. minorée, resp. bornée).
- Toute application constante est bornée.
- L'application f est bornée si et seulement $|f|$ est majorée, c'est-à-dire si et seulement si $\exists M \in \mathbb{R}_+$, $\forall x \in I$, $|f(x)| \leq M$.

Proposition et définition 9.8 *Si l'application $f : I \to \mathbb{R}$ est majorée (resp. minorée), alors $f(I)$ admet une borne supérieure (resp. inférieure) dans $\mathbb{R}$, appelée **borne supérieure** (resp. **inférieure**) **de** f et notée $\sup\limits_{x\in I} f(x)$ (resp. $\inf\limits_{x\in I} f(x)$).*

Cette proposition résulte directement de l'axiome de la borne supérieure (proposition 7.3). Ainsi, par définition, $\sup\limits_{x\in I} f(x) = \sup(\{f(x) \mid x \in I\}) = \sup(f(I))$.

9.2 Limites

Dans cette section, I désigne un intervalle de $\mathbb{R}$, non vide et non réduit à un point.

9.2.1 Notion de limite

Notion de voisinage d'un point

Définitions 9.9 *Soit une propriété dépendant du point x de I. On dit que cette propriété est vraie **au voisinage d'un point** a **de** I si elle est vraie sur l'intersection de I avec un intervalle non vide, ouvert et centré en a.*
*Dans le cas où l'intervalle I est non majoré (resp. non minoré), on dit que la propriété est **vraie au voisinage de** $+\infty$ (resp. $-\infty$) s'il existe un réel M tel qu'elle est vraie sur l'intersection de I avec un intervalle $]M,+\infty[$ (resp. $]-\infty,M[$).*

Limite d'une fonction en un point

Définition 9.10 *Soient $a \in I$ et f une fonction définie sur I, sauf peut-être au point a, et à valeurs dans $\mathbb{K}$. On dit que f **admet** $l \in \mathbb{K}$ **pour limite en** a si et seulement si*

$$\forall \varepsilon > 0,\ \exists \alpha > 0,\ \forall x \in I,\ (0 < |x-a| \leq \alpha \Rightarrow |f(x)-l| \leq \varepsilon).$$

On voit que le fait que f ne soit pas définie en a n'empêche pas de considérer sa limite en ce point. On dit alors que f admet une limite lorsque x tend vers *a par valeurs différentes*. Lorsque f a pour limite $l \in \mathbb{K}$ en a, on dit aussi que *f admet une limite finie en a*.

Exemples.

— La fonction de $\mathbb{R}$ dans $\mathbb{R}$ qui à chaque réel x associe $x \sin \dfrac{1}{x}$ a pour limite 0 en 0.

— Soit f l'application de $\mathbb{R}$ dans $\mathbb{R}$ définie par

$$\forall x \in \mathbb{R},\ f(x) = \begin{cases} 0 & \text{si } x \neq 0 \\ 1 & \text{si } x = 0 \end{cases}.$$

Cette fonction tend vers 0 quand x tend vers 0 par valeurs différentes.

Remarque. Comme le montre ce dernier exemple, une application f peut admettre l pour limite en a sans que l'on ait pour autant $l = f(a)$.

Définitions 9.11 *Soient $a \in I$ et f une fonction définie sur I, sauf peut-être au point a, et à valeurs réelles. On dit que f admet*

- $+\infty$ ***pour limite en*** *a si et seulement si*

$$\forall A \in \mathbb{R},\ \exists \alpha > 0,\ \forall x \in I,\ (0 < |x-a| \leq \alpha \Rightarrow f(x) \geq A),$$

- $-\infty$ ***pour limite en*** *a si et seulement si*

$$\forall B \in \mathbb{R},\ \exists \alpha > 0,\ \forall x \in I,\ (0 < |x-a| \leq \alpha \Rightarrow f(x) \leq B).$$

Unicité de la limite

Proposition 9.12 *Si une application admet une limite finie en un point, alors celle-ci est unique.*

DÉMONSTRATION. Raisonnons par l'absurde et supposons que l'application $f : I \to \mathbb{K}$ admet l et l' appartenant à $\mathbb{K}$ pour limites en un point a de I, avec $l \neq l'$. Posons $\varepsilon = \dfrac{1}{3}|l-l'|$. Il existe $\alpha > 0$ et $\alpha' > 0$ tels que

$$\forall x \in I,\ (0 < |x-a| \leq \alpha \Rightarrow |f(x)-l| \leq \varepsilon) \text{ et } (0 < |x-a| \leq \alpha' \Rightarrow |f(x)-l'| \leq \varepsilon).$$

Alors, pour tout x de I tel que $0 < |x-a| \leq \min(\alpha, \alpha')$, nous avons

$$|l-l'| \leq |l-f(x)| + |f(x)-l'| \leq 2\varepsilon = \frac{2}{3}|l-l'|,$$

d'où une contradiction. □

En vertu de cette unicité, si l'application f admet l pour limite en a, on dit que *l est la limite de f en a* et l'on note

$$\lim_{x\to a} f(x) = l \text{ ou encore } f(x) \underset{x\to a}{\longrightarrow} l.$$

Extension de la notion de limite

Il est possible d'étendre les définitions de limites finie et infinie si la fonction f est définie sur un intervalle non majoré ou non minoré.

Définitions 9.13 *Soient $f \in \mathscr{F}(I,\mathbb{K})$ et $l \in \mathbb{K}$. Si I admet $+\infty$ comme extrémité, on dit que f **admet** l **pour limite en** $+\infty$ si et seulement si*

$$\forall \varepsilon > 0,\ \exists A \in \mathbb{R},\ \forall x \in I,\ (x \geq A \Rightarrow |f(x)-l| \leq \varepsilon),$$

et l'on note $\lim\limits_{x\to+\infty} f(x) = l$.
*Si I admet $-\infty$ comme extrémité, on dit que f **admet** l **pour limite en** $-\infty$ si et seulement si*

$$\forall \varepsilon > 0,\ \exists B \in \mathbb{R},\ \forall x \in I,\ (x \leq B \Rightarrow |f(x)-l| \leq \varepsilon),$$

et l'on note $\lim\limits_{x\to-\infty} f(x) = l$.

Définitions 9.14 *Soit $f \in \mathscr{F}(I,\mathbb{R})$. Si I admet $+\infty$ comme extrémité, on dit que f **admet** $+\infty$ (resp. $-\infty$) **pour limite en** $+\infty$ si et seulement si*

$$\forall A \in \mathbb{R},\ \exists A' \in \mathbb{R},\ \forall x \in I,\ (x \geq A' \Rightarrow f(x) \geq A)$$

$$(\textit{resp. } \forall B \in \mathbb{R},\ \exists A' \in \mathbb{R},\ \forall x \in I,\ (x \geq A' \Rightarrow f(x) \leq B)),$$

et l'on note $\lim\limits_{x\to+\infty} f(x) = +\infty$ (resp. $\lim\limits_{x\to+\infty} f(x) = -\infty$).
*Si I admet $-\infty$ comme extrémité, on dit que f **admet** $+\infty$ (resp. $-\infty$) **pour limite en** $-\infty$ si et seulement si*

$$\forall A \in \mathbb{R},\ \exists B' \in \mathbb{R},\ \forall x \in I,\ (x \leq B' \Rightarrow f(x) \geq A)$$

$$(\textit{resp. } \forall B \in \mathbb{R},\ \exists B' \in \mathbb{R},\ \forall x \in I,\ (x \leq B' \Rightarrow f(x) \leq B)),$$

et l'on note $\lim\limits_{x\to-\infty} f(x) = +\infty$ (resp. $\lim\limits_{x\to-\infty} f(x) = -\infty$).

Dans toute la suite de ce chapitre et afin d'unifier la présentation des définitions et résultats, nous considérons le point a comme élément de $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$.

Nous établissons à présent la

Proposition 9.15 *Si $f : I \to \mathbb{K}$ admet une limite finie en $a \in \overline{\mathbb{R}}$, alors f est bornée au voisinage de a.*

Démonstration. Supposons $a \in \mathbb{R}$, les cas $a = +\infty$ et $a = -\infty$ étant analogues.
Il existe $\alpha > 0$ tel que l'on a

$$\forall x \in I,\ (0 < |x-a| \leq \alpha \Rightarrow |f(x)-l| \leq 1 \Rightarrow |f(x)| \leq |f(x)-l| + |l| \leq 1 + |l|),$$

et donc f est bornée au voisinage de a. □

Limite à droite, limite à gauche

Définition 9.16 *On dit que f* ***admet une limite à droite*** *(resp.* ***à gauche)*** ***en*** *$a \in \mathbb{R}$ si et seulement si la restriction de f à $I \cap]a, +\infty[$ (resp. $]-\infty, a[\cap I$), notée $f_{|_{I\cap]a,+\infty[}}$ (resp. $f_{|_{]-\infty,a[\cap I}}$), admet une limite en a.*

Lorsque f admet l pour limite à droite (resp. à gauche) en a, on note

$$\lim_{x\to a,\ x>a} f(x) = l \text{ ou } \lim_{x\to a^+} f(x) = l \text{ (resp. } \lim_{x\to a,\ x<a} f(x) = l \text{ ou } \lim_{x\to a^-} f(x) = l),$$

et l'on a

$$\lim_{x\to a^+} f(x) = l \Leftrightarrow (\forall \varepsilon > 0,\ \exists \alpha > 0,\ \forall x \in I,\ (0 < x - a \leq \alpha \Rightarrow |f(x) - l| \leq \varepsilon))$$
$$\text{(resp. } \lim_{x\to a^-} f(x) = l \Leftrightarrow (\forall \varepsilon > 0,\ \exists \alpha > 0,\ \forall x \in I,\ (0 < a - x \leq \alpha \Rightarrow |f(x) - l| \leq \varepsilon))).$$

On dit également que f tend vers l lorsque x tend vers a par *valeurs supérieures* (resp. *inférieures*).

Exemple. Considérons la fonction $x \mapsto \dfrac{|x|}{x}$ définie sur $\mathbb{R}^*$. Si $x > 0$, $f(x) = 1$ et l'on a $\lim_{x\to 0^+} f(x) = 1$. Si $x < 0$, alors $f(x) = -1$ et l'on a $\lim_{x\to 0^-} f(x) = -1$.

Caractérisation séquentielle de la limite

Proposition 9.17 *Pour qu'une application $f : I \to \mathbb{K}$ admette $l \in \mathbb{K}$ pour limite en $a \in \overline{\mathbb{R}}$, il faut et il suffit que, pour toute suite $(u_n)_{n\in\mathbb{N}}$ d'éléments de I et ayant a pour limite, on ait*

$$\lim_{n\to+\infty} f(u_n) = l.$$

DÉMONSTRATION. Faisons l'hypothèse que $a \in \mathbb{R}$, les cas $a = +\infty$ et $a = -\infty$ étant analogues.
Supposons tout d'abord que f admet l pour limite en a. Soient $(u_n)_{n\in\mathbb{N}}$ une suite dans I, telle que $\lim_{n\to+\infty} u_n = a$, et $\varepsilon > 0$. Puisque $\lim_{x\to a} f(x) = l$, il existe $\alpha > 0$ tel que

$$\forall x \in I,\ (0 < |x - a| \leq \alpha \Rightarrow |f(x) - l| \leq \varepsilon).$$

Puisque $(u_n)_{n\in\mathbb{N}}$ tend vers a, il existe $N \in \mathbb{N}$ tel que

$$\forall n \in \mathbb{N},\ (n \geq N \Rightarrow |u_n - a| \leq \alpha).$$

On a alors

$$\forall n \in \mathbb{N},\ (n \geq N \Rightarrow |u_n - a| \leq \alpha \Rightarrow |f(u_n) - l| \leq \varepsilon),$$

d'où $\lim_{n\to+\infty} f(u_n) = l$.
Supposons à présent que f n'admet pas l pour limite en a. Il existe donc $\varepsilon > 0$ tel que

$$\forall \alpha > 0,\ \exists x \in I,\ (0 < |x - a| \leq \alpha \text{ et } |f(x) - l| > \varepsilon).$$

En particulier, en prenant $\alpha = \dfrac{1}{n}$ pour tout n de $\mathbb{N}^*$, il existe $u_n \in I$ tel que

$$|u_n - a| \leq \frac{1}{n} \text{ et } |f(u_n) - l| > \varepsilon.$$

On constate alors que la suite $(u_n)_{n\in\mathbb{N}}$ dans I ainsi construite satisfait $\lim_{n\to+\infty} u_n = a$, mais est telle que $(f(u_n))_{n\in\mathbb{N}}$ ne converge pas vers l. □

9.2.2 Ordre et limite

Passage à la limite dans une inégalité

Proposition 9.18 *Soient f et g deux fonctions de $\mathscr{F}(I,\mathbb{R})$ admettant une limite en $a \in \overline{\mathbb{R}}$. Si l'on a $f(x) \leq g(x)$ au voisinage de a, alors*

$$\lim_{x\to a} f(x) \leq \lim_{x\to a} g(x).$$

DÉMONSTRATION. Supposons que f et g tendent respectivement vers l et l' lorsque x tend vers a. Soit $\varepsilon > 0$. Il existe $\alpha > 0$ et $\alpha' > 0$ tels que

$$\forall x \in I,\ \left(0 < |x-a| \leq \alpha \Rightarrow l - \frac{\varepsilon}{2} \leq f(x) \leq l + \frac{\varepsilon}{2}\right) \text{ et } \left(0 < |x-a| \leq \alpha' \Rightarrow l' - \frac{\varepsilon}{2} \leq g(x) \leq l' + \frac{\varepsilon}{2}\right).$$

Nous avons donc, pour tout $x \in I$ tel que $0 < |x-a| \leq \min(\alpha,\alpha')$,

$$l - \frac{\varepsilon}{2} \leq f(x) \leq g(x) \leq l + \frac{\varepsilon}{2},$$

d'où $l - l' \leq \varepsilon$. On a finalement $l \leq l'$ car $\varepsilon > 0$ est arbitraire. □

Théorème d'encadrement

Proposition 9.19 ***(« théorème des gendarmes »)*** *Soient f, g et h trois fonctions de $\mathscr{F}(I,\mathbb{R})$ telles que $f(x) \leq g(x) \leq h(x)$ au voisinage de $a \in \overline{\mathbb{R}}$. Si f et h admettent une même limite l en a, alors g admet l pour limite en a.*

DÉMONSTRATION. Supposons $a \in \mathbb{R}$, les cas $a = +\infty$ et $a = -\infty$ étant analogues.
Soit $\varepsilon > 0$. Puisque f et h admettent l pour limite en a, il existe $\alpha > 0$ et $\alpha' > 0$ tels que

$$\forall x \in I,\ (0 < |x-a| \leq \alpha \Rightarrow |f(x)-l| \leq \varepsilon) \text{ et } (0 < |x-a| \leq \alpha' \Rightarrow |h(x)-l| \leq \varepsilon).$$

Nous avons donc, pour tout $x \in I$ tel que $0 < |x-a| \leq \min(\alpha,\alpha')$,

$$((|f(x)-l| \leq \varepsilon \text{ et } |h(x)-l| \leq \varepsilon) \Rightarrow -\varepsilon \leq f(x)-l \leq g(x)-l \leq h(x)-l \leq \varepsilon \Rightarrow |g(x)-l| \leq \varepsilon).$$

Donc g admet l pour limite en a. □

Ce théorème d'encadrement s'avère très utile en pratique puisqu'il permet notamment de conclure à l'existence d'une limite.

9.2.3 Opérations algébriques sur les limites

Cas des limites finies

Proposition 9.20 *Soient $\lambda \in \mathbb{K}$, $f,g : I \to \mathbb{K}$, $a \in \overline{\mathbb{R}}$ et $(l,l') \in \mathbb{K}^2$. On a*

1) $\lim_{x\to a} f(x) = l \Rightarrow \lim_{x\to a} |f(x)| = |l|$.

2) $\lim_{x\to a} f(x) = l$ *et* $\lim_{x\to a} g(x) = l' \Rightarrow \lim_{x\to a} (f(x)+g(x)) = l + l'$.

3) $\lim_{x\to a} f(x) = l \Rightarrow \lim_{x\to a} \lambda f(x) = \lambda l$.

4) $\lim_{x\to a} f(x) = 0$ *et g est bornée au voisinage de a* $\Rightarrow \lim_{x\to a} f(x)g(x) = 0$.

5) $\lim_{x\to a} f(x) = l$ *et* $\lim_{x\to a} g(x) = l' \Rightarrow \lim_{x\to a} f(x)g(x) = ll'$.

6) $\lim_{x\to a} g(x) = l'$ *et* $l' \neq 0 \Rightarrow \lim_{x\to a} \dfrac{1}{g(x)} = \dfrac{1}{l'}$.

7) $\lim_{x\to a} f(x) = l$ *et* $\lim_{x\to a} g(x) = l'$ *et* $l' \neq 0 \Rightarrow \lim_{x\to a} \dfrac{f(x)}{g(x)} = \dfrac{l}{l'}$.

DÉMONSTRATION. Supposons $a \in I \subset \mathbb{R}$, les cas $a = +\infty$ et $a = -\infty$ étant analogues.

1) Soit $\varepsilon > 0$. Puisque $\lim\limits_{x\to a} f(x) = l$, il existe $\alpha > 0$ tel que

$$\forall x \in I,\ (|x-a| \le \alpha \Rightarrow |f(x)-l| \le \varepsilon).$$

Comme $\forall x \in I$, $||f(x)|-|l|| \le |f(x)-l|$, on déduit

$$\forall x \in I,\ (|x-a| \le \alpha \Rightarrow ||f(x)|-|l|| \le \varepsilon),$$

et finalement $\lim\limits_{x\to a} |f(x)| = |l|$.

2) Soit $\varepsilon > 0$. Puisque $\lim\limits_{x\to a} f(x) = l$ et $\lim\limits_{x\to a} g(x) = l'$, il existe $\alpha > 0$ et $\alpha' > 0$ tels que

$$\forall x \in I,\ \left(|x-a| \le \alpha \Rightarrow |f(x)-l| \le \frac{\varepsilon}{2}\right) \text{ et } \left(|x-a| \le \alpha' \Rightarrow |g(x)-l'| \le \frac{\varepsilon}{2}\right)$$

En notant $\alpha'' = \min(\alpha, \alpha') > 0$, nous avons, $\forall x \in I$,

$$\left(|x-a| \le \alpha'' \Rightarrow (|f(x)-l| \le \frac{\varepsilon}{2} \text{ et } |g(x)-l'| \le \frac{\varepsilon}{2}\right) \Rightarrow |(f(x)+g(x))-(l+l')| \le |f(x)-l| + |g(x)-l'| \le \varepsilon),$$

d'où $\lim\limits_{x\to a} (f(x)+g(x)) = l + l'$.

3) Soit $\varepsilon > 0$. Puisque $\lim\limits_{x\to a} f(x) = l$, il existe $\alpha > 0$ tel que

$$\forall x \in I,\ \left(|x-a| \le \alpha \Rightarrow |f(x)-l| \le \frac{\varepsilon}{|\lambda|+1}\right),$$

d'où $\forall x \in I$, $\left(|x-a| \le \alpha \Rightarrow |\lambda f(x) - \lambda l| = |\lambda||f(x)-l| \le \dfrac{|\lambda|\varepsilon}{|\lambda|+1} \le \varepsilon\right)$, et donc $\lim\limits_{x\to a} \lambda f(x) = \lambda l$.

4) Par hypothèse, il existe $\alpha > 0$ et $C \in \mathbb{R}_+$ tels que

$$\forall x \in I,\ (|x-a| \le \alpha \Rightarrow |g(x)| \le C).$$

Soit $\varepsilon > 0$. Puisque $\lim\limits_{x\to a} f(x) = 0$, il existe $\alpha' > 0$ tel que

$$\forall x \in I,\ \left(|x-a| \le \alpha' \Rightarrow |f(x)| \le \frac{\varepsilon}{C+1}\right).$$

En notant $\alpha'' = \min(\alpha, \alpha') > 0$, nous avons alors

$$\forall x \in I,\ \left(|x-a| \le \alpha'' \Rightarrow |f(x)g(x)| = |f(x)||g(x)| \le \frac{C\varepsilon}{C+1} \le \varepsilon\right),$$

et donc $\lim\limits_{x\to a} f(x)g(x) = 0$.

5) Notons h l'application de I dans $\mathbb{K}$ telle que $h(x) = f(x) - l$, $\forall x \in I$. Nous avons

$$\forall x \in I,\ f(x)g(x) = h(x)g(x) + lg(x).$$

D'après 3), $\lim\limits_{x\to a} lg(x) = ll'$. D'autre part, $\lim\limits_{x\to a} h(x) = 0$, donc, d'après 4), $\lim\limits_{x\to a} h(x)g(x) = 0$, puisque g est bornée au voisinage de a. Finalement, on a $\lim\limits_{x\to a} f(x)g(x) = ll'$ d'après 2).

6) Puisque $\lim\limits_{x\to a} g(x) = l'$, on a, d'après 1), $\lim\limits_{x\to a} |g(x)| = |l'|$. Comme $|l'| > 0$, il existe $\alpha > 0$ tel que

$$\forall x \in I,\ \left(|x-a| \le \alpha \Rightarrow |g(x)| > \frac{|l'|}{2}\right),$$

En particulier, $\forall x \in I$, $(|x-a| \le \alpha \Rightarrow g(x) \ne 0)$. La fonction $\left(\dfrac{1}{g}\right)$ est donc définie, au moins sur $I \cap]a-\alpha, a+\alpha[$. Nous avons alors, pour tout x de $I \cap]a-\alpha, a+\alpha[$,

$$0 \le \left|\frac{1}{g(x)} - \frac{1}{l'}\right| = \frac{|g(x)-l'|}{|g(x)||l'|} \le \frac{2}{|l'|^2}|g(x)-l'|.$$

Comme $\lim\limits_{x\to a} g(x) = l'$, on en déduit que $\lim\limits_{x\to a}\left(\dfrac{2}{|l'|^2}|g(x)-l'|\right) = 0$, puis que $\lim\limits_{x\to a}\left|\dfrac{1}{g(x)} - \dfrac{1}{l'}\right| = 0$, soit encore $\lim\limits_{x\to a} \dfrac{1}{g(x)} = \dfrac{1}{l'}$.

7) Il suffit d'appliquer 5) et 6) en remarquant que $\dfrac{f}{g} = f\dfrac{1}{g}$.

$\square$

Cas des limites infinies

Proposition 9.21 *Soient $f, g : I \to \mathbb{R}$ et $a \in \overline{\mathbb{R}}$.*

1) Si $\lim\limits_{x\to a} f(x) = +\infty$ et si g est minorée au voisinage de a, alors

$$\lim_{x\to a}(f(x) + g(x)) = +\infty.$$

2) Si $\lim\limits_{x\to a} f(x) = +\infty$ et si g est minorée au voisinage de a par une constante strictement positive, alors

$$\lim_{x\to a} f(x)g(x) = +\infty.$$

DÉMONSTRATION. Les preuves sont analogues à celles de 1) et 2) dans la proposition 8.20. □

Composition des limites

Proposition 9.22 *Soient $f : I \to \mathbb{R}$, J un intervalle de $\mathbb{R}$ tel que $f(I) \subset J$, $g : J \to \mathbb{K}$, $a \in \overline{\mathbb{R}}$, $b \in \overline{\mathbb{R}}$ et $l \in \mathbb{K} \cup \{-\infty, +\infty\}$. Si f admet b pour limite en a et g admet l pour limite en b, alors $g \circ f$ admet l pour limite en a.*

DÉMONSTRATION. Supposons $a \in \mathbb{R}$, $b \in \mathbb{R}$ et $l \in \mathbb{K}$, les autres cas étant analogues.
Soit $\varepsilon > 0$. Puisque $\lim\limits_{y\to b} g(y) = l$, il existe $\alpha > 0$ tel que

$$\forall y \in J,\ (|y - b| \le \alpha \Rightarrow |g(y) - l| \le \varepsilon).$$

Puis, comme $\lim\limits_{x\to a} f(x) = b$, il existe $\alpha' > 0$ tel que

$$\forall x \in I,\ (|x - a| \le \alpha' \Rightarrow |f(x) - b| \le \alpha_1).$$

Nous avons alors

$$\forall x \in I,\ (|x - a| \le \alpha' \Rightarrow |f(x) - b| \le \alpha \Rightarrow |g(f(x)) - l| \le \varepsilon),$$

d'où $\lim\limits_{x\to a} g \circ f(x) = l$. □

9.2.4 Cas des fonctions monotones

Nous considérons dans cette sous-section des fonctions à valeurs réelles.

Théorème 9.23 *Soient $(a, b) \in \overline{\mathbb{R}}^2$, tel que $a < b$, et f une application croissante définie sur l'intervalle $]a, b[$.*

1) Si f est majorée, alors f admet une limite finie en b et $\lim\limits_{x\to b} f(x) = \sup\limits_{x\in]a,b[} f(x)$.

2) Si f n'est pas majorée, alors f admet $+\infty$ pour limite en b.

DÉMONSTRATION.

1) La partie $f(]a, b[)$ de $\mathbb{R}$ est non vide et majorée, elle admet par conséquent une borne supérieure l dans $\mathbb{R}$. Soit $\varepsilon > 0$. Puisque $l - \varepsilon$ n'est pas un majorant de $f(]a, b[)$ dans $\mathbb{R}$, il existe $x_0 \in]a, b[$ tel que $l - \varepsilon \le f(x_0) \le l$. Alors, pour tout $x \in]a, b[$, nous avons
$$x_0 \le x \Rightarrow f(x_0) \le f(x) \Rightarrow l - \varepsilon \le f(x) \Rightarrow |f(x) - l| \le \varepsilon.$$
Supposons $b \in \mathbb{R}$ (le cas $b = +\infty$ étant analogue). En posant $\alpha = b - x_0 > 0$, nous avons ainsi $\forall x \in]a, b[$, $(0 < b - x \le \alpha \Rightarrow |f(x) - l| \le \varepsilon)$, d'où $\lim\limits_{x\to b} f(x) = l$.

2) Soit $A \in \mathbb{R}$. Puisque f n'est pas majorée, il existe $x_0 \in]a, b[$ tel que $f(x_0) \ge A$. Alors, pour $x \in]a, b[$, nous avons
$$x_0 \le x \Rightarrow f(x_0) \le f(x) \Rightarrow f(x) \ge A.$$
Supposons $b \in \mathbb{R}$ (le cas $b = +\infty$ étant analogue). En posant $\alpha = b - x_0 > 0$, nous avons ainsi $\forall x \in]a, b[$, $(0 < b - x \le \alpha \Rightarrow |f(x)| \ge A)$, d'où $\lim\limits_{x\to b} f(x) = +\infty$.

□

Lorsque $b \in \mathbb{R}$, on peut parler de limite à gauche en b dans le théorème précédent.
On déduit de ce dernier résultat qu'une application croissante admet toujours une limite, finie ou infinie, en b. Un résultat analogue est obtenu pour les applications décroissantes en considérant $-f$ dans cette démonstration.

9.2.5 Applications équivalentes au voisinage d'un point

Définition 9.24 *On dit que deux applications f et g définies sur I sont* ***équivalentes au voisinage de*** *a si et seulement s'il existe une application h telle que $\lim\limits_{x\to a} h(x) = 1$ et*

$$\exists \delta > 0,\ \forall x \in I,\ (0 < |x-a| \leq \delta \Rightarrow f(x) = g(x)h(x)).$$

On note alors $f \underset{x\to a}{\sim} g$.

Définition 9.25 *Soit I un intervalle de $\mathbb{R}$ non majoré (resp. non minoré). On dit que deux applications f et g définies sur I sont* ***équivalentes au voisinage de*** *$+\infty$ (resp. $-\infty$) si et seulement s'il existe un réel A (resp. B) et une fonction h, définie sur l'ensemble des points de I qui vérifient $x \geq A$ (resp. $x \leq B$), telle que $\lim\limits_{x\to+\infty} h(x) = 1$ (resp. $\lim\limits_{x\to-\infty} h(x) = 1$) et que $f(x) = g(x)h(x)$. On note alors $f \underset{x\to+\infty}{\sim} g$ (resp. $f \underset{x\to-\infty}{\sim} g$).*

La notion d'applications équivalentes permet de ramener localement l'étude d'une application à celle d'une autre application dont le comportement est connu.

Exemples.
- Les applications de $\mathbb{R}$ dans $\mathbb{R}$ $x \mapsto x$ et $x \mapsto \sin x$ sont équivalentes au voisinage de 0.
- Les applications de $\mathbb{R}$ dans $\mathbb{R}$ $x \mapsto x$ et $x \mapsto e^x - 1$ sont équivalentes au voisinage de 0.
- Les applications de $\mathbb{R}$ dans $\mathbb{R}$ $x \mapsto x$ et $x \mapsto \ln(1+x)$ sont équivalentes au voisinage de 0.
- Les applications de $\mathbb{R}$ dans $\mathbb{R}$ $x \mapsto \frac{x^2}{2}$ et $x \mapsto 1 - \cos x$ sont équivalentes au voisinage de 0.
- Les applications de $\mathbb{R}$ dans $\mathbb{R}$ qui à x associent $x^4 + x^2 + x + 1$ et $x^4 + x^3$ sont équivalentes au voisinage de $+\infty$ ou de $-\infty$.

Proposition 9.26 *Si deux applications f et g sont équivalentes au voisinage de $a \in \overline{\mathbb{R}}$ et que f admet une limite en a, alors g admet également cette limite en a.*

DÉMONSTRATION. Par définition, on a $g = fh$ dans un voisinage de a. Le résultat se déduit alors facilement du 5) de la proposition 9.20, puisque $\lim\limits_{x\to a} h(x) = 1$. □

9.3 Continuité

Dans cette section, I désigne un intervalle de $\mathbb{R}$, non vide et non réduit à un point.

9.3.1 Définitions

Continuité en un point

Soient f une application définie sur I à valeurs dans $\mathbb{K}$ et $a \in I$.

Définition 9.27 *On dit que f* ***est continue en*** *a si et seulement si*

$$\forall \varepsilon > 0,\ \exists \alpha > 0,\ \forall x \in I,\ (|x-a| \leq \alpha \Rightarrow |f(x) - f(a)| \leq \varepsilon).$$

À la différence de la notion de limite, on ne parle de continuité qu'en des points où la fonction est définie.
On dit que f *est discontinue en a* si et seulement si f n'est pas continue en a, qui est alors appelé un *point de discontinuité de f*.

Définition 9.28 *On dit qu'une application f admet une* ***discontinuité de première espèce en*** *a si et seulement si elle n'est pas continue en a et possède une limite à droite et une limite à gauche en a. Lorsque f n'est pas continue et n'admet pas de discontinuité de première espèce en a, on dit qu'elle admet une* ***discontinuité de seconde espèce en*** *a.*

Exemples.

— L'application f de $\mathbb{R}$ dans $\mathbb{R}$, définie pour chaque réel x par

$$f(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 & \text{si } x \geq 0 \end{cases},$$

admet une discontinuité de première espèce en 0.

— L'application f de $\mathbb{R}$ dans $\mathbb{R}$, définie pour chaque réel x par

$$f(x) = \begin{cases} \dfrac{1}{x} & \text{pour } x \neq 0 \\ 0 & \text{pour } x = 0 \end{cases},$$

admet une discontinuité de seconde espèce en 0.

La proposition suivante est immédiate.

Proposition 9.29 *Pour qu'une application f soit continue en a, il faut et il suffit qu'elle admette $f(a)$ pour limite en a.*

Proposition 9.30 *Si une application f est continue en a, alors elle est bornée au voisinage de a.*

DÉMONSTRATION. Si l'application f est continue en a, alors il existe $\alpha > 0$ tel que

$$\forall x \in I, (|x-a| \leq \alpha \Rightarrow |f(x)-f(a)| \leq 1 \Rightarrow |f(x)| \leq |f(x)-f(a)| + |f(a)| \leq |f(a)| + 1),$$

donc f est bornée au voisinage de a. □

Continuité à droite, continuité à gauche

Définition 9.31 *Soient f une application définie sur I à valeurs dans $\mathbb{K}$ et $a \in I$. On dit que f est **continue à droite** (resp. **à gauche**) **en** a si et seulement si la restriction de f à $I \cap]a, +\infty[$ (resp. $]-\infty, a[\cap I$) est continue en a.*

Remarque. L'application f est continue en a si et seulement si elle est continue à droite et à gauche en a.

Exemple. Considérons l'application qui à tout réel x associe la partie entière de x, $E(x)$. Pour tout entier naturel n, cette application est continue à droite en n, mais n'est pas continue à gauche en n.

Caractérisation séquentielle de la continuité

Comme dans le cas de la limite, on peut définir la notion de continuité en se servant de suites réelles.

Proposition 9.32 *Soient f une application définie sur I à valeurs dans $\mathbb{K}$ et $a \in I$; f est continue en a si et seulement si, pour toute suite $(u_n)_{n\in\mathbb{N}}$ d'éléments de I tendant vers a, la suite $(f(u_n))_{n\in\mathbb{N}}$ tend vers $f(a)$.*

DÉMONSTRATION. La preuve découle directement de la proposition 9.29 et de la caractérisation séquentielle de la limite (proposition 9.17). □

Prolongement par continuité

Définition 9.33 *Soit f une fonction définie sur I sauf en un point a de I et admettant une limite finie l en a. On appelle **prolongement par continuité de** f **en** a la fonction g, définie sur I par*

$$g(x) = \begin{cases} l & \text{si } x = a \\ f(x) & \text{sinon} \end{cases}.$$

Cette fonction est continue en a.

Il est facile de vérifier que lorsqu'une fonction admet un prolongement par continuité en un point, celui-ci est unique. S'il n'y a pas de risque d'ambiguïté, on désigne alors la fonction et son prolongement par le même symbole. Notons qu'on peut aussi définir un prolongement par continuité à droite de a ou à gauche de a.

Exemples.

— On peut prolonger par continuité en 0 la fonction $f(x) = \dfrac{\sin x}{x}$ en posant $f(0) = 1$.

— On peut prolonger par continuité en 0 la fonction $f(x) = e^{-\frac{1}{x^2}}$ en posant $f(0) = 0$.

Continuité sur un intervalle

Définition 9.34 *Soit f une application définie sur I à valeurs dans $\mathbb{K}$. On dit que f* **est continue sur** *I si et seulement si f est continue en tout point de I.*

On note $C(I, \mathbb{K})$ l'ensemble des applications de I dans $\mathbb{K}$ continues sur I.

Continuité par morceaux

Définition 9.35 *Soient $(a, b) \in \mathbb{R}^2$, tel que $a < b$, et $f : [a, b] \to \mathbb{K}$. On dit que f* **est continue par morceaux sur** *$[a, b]$ si et seulement s'il existe $n \in \mathbb{N}^*$ et $(a_0, \dots, a_n) \in [a, b]^{n+1}$ tels que $a = a_0 < \cdots < a_n = b$ et, pour tout $i \in \{0, \dots, n-1\}$, f est continue sur $]a_i, a_{i+1}[$ et admet une limite finie à droite en a_i et une limite finie à gauche en a_{i+1}.*

9.3.2 Opérations algébriques sur les applications continues

Proposition 9.36 *Soient $\lambda \in \mathbb{K}$, $f, g : I \to \mathbb{K}$ et $a \in I$.*

1) Si f est continue en a, alors $|f|$ est continue en a.

2) Si f et g sont continues en a, alors $f + g$ est continue en a.

3) Si f est continue en a, alors λf est continue en a.

4) Si f et g sont continues en a, alors fg est continue en a.

5) Si g est continue en a et si $g(a) \neq 0$, alors $\dfrac{1}{g}$ est continue en a.

6) Si f et g sont continues en a et si $g(a) \neq 0$, alors $\dfrac{f}{g}$ est continue en a.

DÉMONSTRATION. Les preuves sont similaires à celles de la proposition 9.20. □

Proposition 9.37 *Soient J un intervalle de $\mathbb{R}$, $f : I \to \mathbb{R}$ telle que $f(I) \subset J$, $g : J \to \mathbb{K}$ et $a \in I$. Si f est continue en a et g est continue en $f(a)$, alors $g \circ f$ est continue en a.*

DÉMONSTRATION. La preuve est analogue à celle de la proposition 9.22. □

De ces deux propositions, nous déduisons aisément des résultats de continuité globale sur l'intervalle I.

Proposition 9.38 *Soient $\lambda \in \mathbb{K}$, $f, g : I \to \mathbb{K}$.*

1) Si f est continue sur I, alors $|f|$ est continue sur I.

2) Si f et g sont continues sur I, alors $f + g$ est continue sur I.

3) Si f est continue sur I, alors λf est continue sur I.

4) Si f et g sont continues sur I, alors fg est continue sur I.

5) Si g est continue sur I et si $(\forall x \in I,\ g(x) \neq 0$, alors $\dfrac{1}{g}$ est continue sur I.

6) Si f et g sont continues sur I et si $(\forall x \in I,\ g(x) \neq 0$, alors $\dfrac{f}{g}$ est continue sur I.

Proposition 9.39 *Soient J un intervalle de $\mathbb{R}$, $f : I \to \mathbb{R}$ telle que $f(I) \subset J$, $g : J \to \mathbb{K}$. Si f est continue sur I et g est continue sur $f(I)$, alors $g \circ f$ est continue sur I.*

9.3.3 Fonctions continues sur un intervalle

Théorème des valeurs intermédiaires

Théorème 9.40 *Soient f une fonction continue sur I et $(a,b) \in I^2$ tels que $f(a) \le f(b)$. Alors f atteint toutes les valeurs intermédiaires entre $f(a)$ et $f(b)$, c'est-à-dire*

$$\forall y \in [f(a), f(b)],\ \exists c \in I,\ f(c) = y.$$

DÉMONSTRATION. Si $y = f(a)$ ou $y = f(b)$, le résultat est immédiat. Soit donc $y \in]f(a), f(b)[$. Considérons l'ensemble $E = \{x \in [a,b] \mid f(x) \le y\}$; E est une partie de $\mathbb{R}$ non vide (car $a \in E$) et majorée (par b), qui admet donc une borne supérieure, notée c. Nous allons montrer que $f(c) = y$.
Par définition de la borne supérieure, il existe une suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de E telle que $\lim\limits_{n\to+\infty} x_n = c$. L'application f étant continue en c, on a $\lim\limits_{n\to+\infty} f(x_n) = f(c)$. Or, pour tout $n \in \mathbb{N}$, $f(x_n) \le y$ donc $f(c) \le y$. D'autre part, $f(b) > y$, donc $c \neq b$. Pour tout $x \in]c, b[$, $f(x) > y$ donc $\lim\limits_{x\to c,\ x>c} f(x) = f(c) \ge y$ d'où $f(c) = y$. □

Remarque. Le réel c n'est pas forcément unique.

Corollaire 9.41 *Soit f une fonction continue sur un intervalle $[a,b]$ de $\mathbb{R}$. Si $f(a)f(b) \le 0$, alors il existe $c \in [a,b]$ tel que $f(c) = 0$.*

DÉMONSTRATION. Puisque $f(a)$ et $f(b)$ ont des signes opposés, 0 est compris entre $f(a)$ et $f(b)$. D'après le théorème des valeurs intermédiaires, il existe donc $c \in [a,b]$ tel que $f(c) = 0$ □

Continuité sur un segment

Proposition 9.42 *Toute application réelle définie sur un segment $[a,b]$, $(a,b) \in \mathbb{R}^2$, et continue est bornée et atteint ses bornes.*

DÉMONSTRATION. Montrons que f est bornée. Supposons f non majorée. Il existe alors une suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de $[a,b]$ telle que, pour chaque entier n, on a $f(x_n) > n$. Puisque la suite est bornée, il existe, d'après le théorème de Bolzano–Weierstrass (théorème 8.25), une suite extraite de $(x_n)_{n\in\mathbb{N}}$, notée $(x_{\sigma(n)})_{n\in\mathbb{N}}$, qui converge vers un point c de $[a,b]$. Puisque f est continue sur $[a,b]$, on déduit que la suite $(f(x_{\sigma(n)}))_{n\in\mathbb{N}}$ tend vers $f(c)$. Mais d'autre part, on a $\forall n \in \mathbb{N}$, $f(x_{\sigma(n)}) > \sigma(n) \le n$, donc $\lim\limits_{n\to+\infty} f(x_{\sigma(n)}) = +\infty$, d'où une contradiction. L'application f est donc majorée. En appliquant ce résultat à $-f$ au lieu de f, on en déduit que f est minorée. Finalement, f est bornée.
Montrons à présent que f atteint ses bornes. Notons $M = \sup\limits_{x\in[a,b]} f(x)$. Pour chaque entier $n \ge 1$, il existe un réel x_n dans $[a,b]$ tel que

$$M - \frac{1}{n} < f(x_n) \le M.$$

La suite $(x_n)_{n\in\mathbb{N}}$ ainsi construite étant bornée, on peut en extraire, en vertu du théorème de Bolzano–Weierstrass, une sous-suite de $(x_n)_{n\in\mathbb{N}}$, notée $(x_{\tau(n)})_{n\in\mathbb{N}}$, qui converge vers un élément d de $[a,b]$. Puisque f est continue sur $[a,b]$, on en déduit que la suite $(f(x_{\tau(n)}))_{n\in\mathbb{N}}$ tend vers $f(d)$. D'autre part, on a

$$\forall n \in \mathbb{N}^*,\ M - \frac{1}{\tau(n)} < f(x_{\tau(n)}) \le M,$$

d'où, par passage à la limite, $M = f(d)$. Ceci montre que M est atteint par f : $\exists d \in [a,b]$, $M = f(d)$. En appliquant ce résultat à $-f$ au lieu de f, on montre que f atteint aussi $\inf\limits_{x\in[a,b]} f(x)$. □

Image d'un intervalle par une application continue

Proposition 9.43 *L'image d'un intervalle par une application continue à valeurs réelles est un intervalle.*

DÉMONSTRATION. Soient I un intervalle de $\mathbb{R}$ et f une application continue sur I à valeurs dans $\mathbb{R}$. D'après le théorème des valeurs intermédiaires (théorème 9.40), $f(I)$ est un intervalle de $\mathbb{R}$. □

Exemples.

— Soient $I =]0,1]$ et $f : x \mapsto \dfrac{1}{x}$. L'application f est continue sur I et $f(I) = [1,+\infty[$. L'intervalle I est borné alors que $f(I)$ n'est pas borné.

— Soient $I =]0,2\pi[$ et $f : x \mapsto \sin x$. L'application f est continue sur I et $f(I) = [-1,1]$. Dans ce cas, l'intervalle I est ouvert alors que $f(I)$ est fermé.

Comme le montrent les exemples ci-dessus, le caractère ouvert, fermé ou borné d'un intervalle n'est pas toujours conservé par une application. On a cependant le résultat suivant.

Proposition 9.44 *Soit $(a,b) \in \mathbb{R}^2$, tel que $a \le b$, et $f : [a,b] \to \mathbb{R}$ une application. Si f est continue, alors $f([a,b])$ est un segment de $\mathbb{R}$.*

DÉMONSTRATION. D'après la proposition précédente, $f([a,b])$ est un intervalle. D'après la proposition 9.42, c'est une partie bornée de $\mathbb{R}$ qui contient ses bornes. □

Application réciproque d'une application continue strictement monotone

Théorème 9.45 ***(« théorème de la bijection »)*** *Soit f une application continue et strictement monotone sur I ; on note $\widetilde{f} : I \to f(I)$ l'application qui à tout x de I associe $f(x)$. Alors*

1) *L'ensemble $f(I)$ est un intervalle dont les bornes sont les limites de f aux bornes de I.*
2) *L'application $\widetilde{f}$ est bijective.*
3) *La bijection réciproque $\widetilde{f}^{-1}$ est continue sur $f(I)$ et strictement monotone de même sens que f.*

DÉMONSTRATION. Supposons, par exemple, que f est strictement croissante et posons $I =]a,b[$, avec $(a,b) \in \overline{\mathbb{R}}^2$ tel que $a < b$.

1) Pour tout x de $]a,b[$, on a $\lim\limits_{t\to a} f(t) < f(x) < \lim\limits_{t\to b} f(t)$, donc $f(I) \subset]\lim\limits_{t\to a} f(t), \lim\limits_{t\to b} f(t)[$. Réciproquement, soit $y \in]\lim\limits_{x\to a} f(x), \lim\limits_{x\to b} f(x)[$; y n'est ni un majorant, ni un minorant de $f(I)$, il existe donc des éléments x_1 et x_2 de I tels que $f(x_1) < y < f(x_2)$. D'après le théorème des valeurs intermédiaires (théorème 9.40), il existe $x_0 \in I$ tel que $f(x_0) = y$, d'où $y \in f(I)$. Nous en concluons $f(I) =]\lim\limits_{x\to a} f(x), \lim\limits_{x\to b} f(x)[$.
2) Par définition, $\widetilde{f}$ est une surjection. Montrons qu'elle est également injective. Soient x_1 et x_2 des éléments de I tels que $\widetilde{f}(x_1) = \widetilde{f}(x_2)$. Si $x_1 < x_2$, alors $\widetilde{f}(x_1) < \widetilde{f}(x_2)$ et si $x_1 > x_2$, alors $\widetilde{f}(x_1) > \widetilde{f}(x_2)$, d'où une contradiction dans les deux cas. Par conséquent, $x_1 = x_2$ et $\widetilde{f}$ est injective.
3) Soient y_1 et y_2 appartenant à $f(I)$, tels que $y_1 < y_2$. Posons $x_1 = \widetilde{f}^{-1}(y_1)$ et $x_2 = \widetilde{f}^{-1}(y_2)$. Si $x_1 \ge x_2$, alors $\widetilde{f}(x_1) \ge \widetilde{f}(x_2)$, comme f est croissante, soit encore $y_1 \ge y_2$, ce qui est absurde, donc $x_1 < x_2$ et $\widetilde{f}^{-1}$ est strictement croissante.
Montrons qu'elle est également continue. Soit $y_0 \in f(I)$; posons $x_0 = \widetilde{f}^{-1}(y_0)$ et donnons-nous un réel $\varepsilon > 0$ tel que $x_0 - \varepsilon$ et $x_0 + \varepsilon$ appartiennent à I. Posons alors $y_1 = \widetilde{f}(x_0 - \varepsilon)$ et $y_2 = \widetilde{f}(x_0 - \varepsilon)$. L'application f étant strictement croissante, on a $y_1 < y_0 < y_2$ et, pour tout $y \in]y_1, y_2[$, $\widetilde{f}^{-1}(y_1) < \widetilde{f}^{-1}(y) < \widetilde{f}^{-1}(y_2)$, c'est-à-dire $x_0 - \varepsilon < \widetilde{f}^{-1}(y) < x_0 + \varepsilon$. Il existe donc $\beta > 0$ tel que $|y - y_0| \le \beta \Rightarrow |\widetilde{f}^{-1}(y) - \widetilde{f}^{-1}(y_0)| \le \varepsilon$. Par conséquent, $\widetilde{f}$ est continue en y_0.

□

Exemples de fonctions réciproques.

- La fonction sinus est continue et strictement croissante sur $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$. Elle induit donc une bijection de $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ sur $[-1,1]$. Sa bijection réciproque est appelée *arc sinus* et notée arcsin. C'est une fonction continue strictement croissante de $[-1,1]$ sur $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$.
- La fonction cosinus est continue et strictement décroissante sur $[0,\pi]$. Elle induit donc une bijection de $[0,\pi]$ sur $[-1,1]$. Sa bijection réciproque est appelée *arc cosinus* et notée arccos. C'est une fonction continue strictement décroissante de $[-1,1]$ sur $[0,\pi]$.
- La fonction tangente est continue et strictement croissante sur $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$. Elle induit donc une bijection de $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ sur $\mathbb{R}$. Sa bijection réciproque est appelée *arc tangente* et notée arctan. C'est une fonction continue strictement croissante de $\mathbb{R}$ sur $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$.

9.3.4 Continuité uniforme

Définition 9.46 *Soit f une fonction définie sur I à valeurs dans $\mathbb{R}$. On dit que f est **uniformément continue sur** I si et seulement si*

$$\forall \varepsilon > 0,\ \exists \alpha > 0,\ \forall (x,x') \in I^2,\ (|x - x'| \le \alpha \Rightarrow |f(x) - f(x')| \le \varepsilon).$$

Exemples.

— L'application qui à tout réel x associe $|x|$ est uniformément continue.

— L'application qui à tout réel x associe x^2 n'est pas uniformément continue.

Le résultat suivant est immédiat.

Proposition 9.47 *Si f est uniformément continue sur I, alors f est continue sur I.*

Comme on l'a vu plus haut, il existe des fonctions continues non uniformément continues. Cependant, lorsque I est un segment de $\mathbb{R}$, c'est-à-dire un intervalle borné et fermé, nous disposons du

Théorème 9.48 ***(« théorème de Heine[1] »)*** *Toute fonction continue sur un segment $[a,b]$ est uniformément continue sur $[a,b]$.*

DÉMONSTRATION. Raisonnons par l'absurde. Soit f une fonction continue et non uniformément continue sur $[a,b]$. Il existe donc $\varepsilon > 0$ tel que

$$\forall \alpha > 0,\ \exists (x,x') \in [a,b]^2,\ |x-x'| \leq \alpha \text{ et } |f(x)-f(x')| > \varepsilon.$$

En particulier, en prenant $\alpha = \dfrac{1}{n}$, il existe $(x_n, x'_n) \in [a,b]^2$ tel que

$$\forall n \in \mathbb{N}^*,\ |x_n - x'_n| \leq \frac{1}{n} \text{ et } |f(x_n)-f(x'_n)| > \varepsilon.$$

La suite $(x_n)_{n\in\mathbb{N}^*}$ étant bornée, elle admet, en vertu du théorème de Bolzano–Weierstrass (théorème 8.25), une sous-suite, notée $(x_{\sigma(n)})_{n\in\mathbb{N}^*}$, convergente vers un réel, noté l, appartenant à $[a,b]$. Comme

$$\forall n \in \mathbb{N}^*,\ |x_{\sigma(n)} - x'_{\sigma(n)}| \leq \frac{1}{\sigma(n)} \leq \frac{1}{n},$$

on déduit que la suite extraite $(x'_{\sigma(n)})_{n\in\mathbb{N}^*}$ converge aussi vers l. L'application f étant continue en l, les suites $(f(x_{\sigma(n)}))_{n\in\mathbb{N}^*}$ et $(f(x'_{\sigma(n)}))_{n\in\mathbb{N}^*}$ convergent vers $f(l)$ et, par conséquent, $\lim\limits_{n\to+\infty} |f(x_{\sigma(n)}) - f(x'_{\sigma(n)})| = 0$, ce qui contredit le fait que $|f(x_{\sigma(n)}) - f(x'_{\sigma(n)})| > \varepsilon$. □

9.3.5 Applications lipschitziennes

Nous allons à présent introduire une propriété de régularité des applications plus forte que la notion de continuité.

Définition 9.49 *Soient f une fonction définie sur I à valeurs réelles et un réel k strictement positif. On dit que f est **lipschitzienne** si et seulement si*

$$\forall (x,x') \in I^2,\ |f(x)-f(x')| \leq k\,|x-x'|.$$

Lorsque $k \in]0,1[$, on dit que l'application f est *contractante*. Le plus petit réel k tel que f soit k-lipschitzienne est appelé la *constante de Lipschitz[2] de f*.

Proposition 9.50 *Si $f : I \to \mathbb{R}$ est lipschitzienne, alors f est uniformément continue.*

DÉMONSTRATION. Supposons f k-lipschitzienne ($k \in \mathbb{R}_+^*$) et soit $\varepsilon > 0$. Si $k = 0$, f est constante sur I et donc uniformément continue sur I. Si $k > 0$, en prenant $\alpha = \dfrac{\varepsilon}{k}$, nous obtenons

$$\forall (x,x') \in I^2,\ (|x-x'| \leq \alpha \Rightarrow |f(x)-f(x')| \leq \varepsilon),$$

ce qui montre que f est uniformément continue sur I. □

1. Heinrich Eduard Heine (15 mars 1821 - 21 octobre 1881) était un mathématicien allemand. Il est célèbre pour ses résultats en analyse réelle et sur les fonctions spéciales.
2. Rudolph Otto Sigismund Lipschitz (14 mai 1832 - 7 octobre 1903) était un mathématicien allemand. Son travail s'étend sur des domaines aussi variés que la théorie des nombres, l'analyse, la géométrie différentielle et la mécanique classique.

9.4 Fonctions hyperboliques

Définitions 9.51 *On appelle **sinus hyperbolique** et **cosinus hyperbolique** les parties impaire et paire de la fonction exponentielle de base e :*

$$\forall x \in \mathbb{R},\ \sinh(x) = \frac{e^x - e^{-x}}{2}\ \textit{et}\ \cosh(x) = \frac{e^x + e^{-x}}{2}.$$

*On définit également les applications **tangente hyperbolique** et **cotangente hyperbolique** par*

$$\tanh(x) = \frac{\sinh(x)}{\cosh(x)},\ \forall x \in \mathbb{R},\ \textit{et}\ \coth(x) = \frac{\cosh(x)}{\sinh(x)},\ \forall x \in \mathbb{R}^*.$$

Les preuves des résultats suivants sont laissées en exercice au lecteur.

Proposition 9.52 *Pour tout $(x,y) \in \mathbb{R}^2$, on a*

1) $\cosh^2(x) - \sinh^2(x) = 1$,

2) $\sinh(x+y) = \sinh(x)\cosh(y) + \cosh(x)\sinh(y)$,

3) $\sinh(x-y) = \sinh(x)\cosh(y) - \cosh(x)\sinh(y)$,

4) $\cosh(x+y) = \cosh(x)\cosh(y) + \sinh(x)\sinh(y)$,

5) $\cosh(x-y) = \cosh(x)\cosh(y) - \sinh(x)\sinh(y)$.

Proposition 9.53 *1) L'application* sinh *est une bijection de* $\mathbb{R}$ *dans* $\mathbb{R}$*, de bijection réciproque notée* argsinh *telle que*

$$\operatorname{argsinh}(x) = \ln\left(x + \sqrt{x^2+1}\right),\ \forall x \in \mathbb{R}.$$

2) L'application cosh *est une bijection de* $\mathbb{R}_+$ *dans* $[1, +\infty[$*, de bijection réciproque notée* argcosh *telle que*

$$\operatorname{argcosh}(x) = \ln\left(x + \sqrt{x^2-1}\right),\ \forall x \in [1, +\infty[.$$

3) L'application tanh *est une bijection de* $\mathbb{R}$ *dans* $]-1, 1[$*, de bijection réciproque notée* argtanh *telle que*

$$\operatorname{argtanh}(x) = \frac{1}{2}\ln\left(\frac{1+x}{1-x}\right),\ \forall x \in]-1, 1[.$$

Chapitre 10

Dérivabilité

Dans ce chapitre, $\mathbb{K}$ désigne un corps ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$), I un intervalle de $\mathbb{R}$ non vide et non réduit à un point et $\mathscr{F}(I, \mathbb{K})$ est l'ensemble des applications définies sur I et à valeurs dans $\mathbb{K}$.

10.1 Dérivées

10.1.1 Dérivabilité en un point

Définition 10.1 *Soient $f \in \mathscr{F}(I, \mathbb{K})$ et $a \in I$. On dit que f est **dérivable en** a si et seulement si $\lim\limits_{x \to a} \dfrac{f(x) - f(a)}{x - a}$ existe et est finie ; cette limite est alors appelée **dérivée de** f **en** a et notée $f'(a)$.*

En posant $h = x - a$, on obtient une autre écriture très souvent employée :

$$f'(a) = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h},$$

dans laquelle le rapport $\dfrac{f(a+h) - f(a)}{h}$ s'appelle le *taux d'accroissement de f entre a et $a+h$.*
On note également $\dfrac{\mathrm{d}f}{\mathrm{d}x}(a)$ au lieu de $f'(a)$.

Exemple. Toute application constante de I dans $\mathbb{R}$ est dérivable en tout point de I, de dérivée nulle.

Définition 10.2 *Soient $f \in \mathscr{F}(I, \mathbb{K})$ et $a \in I$.*

1) *On dit que f est **dérivable à droite en** a si et seulement si $\lim\limits_{x \to a^+} \dfrac{f(x) - f(a)}{x - a}$ existe et est finie ; cette limite est alors appelée **dérivée de** f **à droite en** a et notée $f'_d(a)$.*
2) *On dit que f est **dérivable à gauche en** a si et seulement si $\lim\limits_{x \to a^-} \dfrac{f(x) - f(a)}{x - a}$ existe et est finie ; cette limite est alors appelée **dérivée de** f **à gauche en** a et notée $f'_g(a)$.*

Exemple. L'application $x \mapsto |x|$ de $\mathbb{R}$ dans $\mathbb{R}$ est dérivable à gauche en 0 et dérivable à droite en 0, et $f'_g(0) = -1$, $f'_d(0) = 1$.
Le résultat suivant est immédiat.

Proposition 10.3 *Soient $f \in \mathscr{F}(I, \mathbb{K})$ et $a \in I$. Pour que l'application f soit dérivable en a, il faut et il suffit que f soit dérivable à gauche et à droite en a et que $f'_g(a) = f'_d(a)$. Dans ces conditions, on a $f'(a) = f'_g(a) = f'_d(a)$.*

Proposition 10.4 *Soient $f \in \mathscr{F}(I, \mathbb{K})$ et $a \in I$. Si l'application f est dérivable en a, alors elle est continue en a.*

DÉMONSTRATION. On sait d'une part que
$$\lim_{x\to a}\frac{f(x)-f(a)}{x-a}=f'(a).$$
D'autre part, on a
$$\lim_{x\to a}f(x)-f(a)=\left(\lim_{x\to a}x-a\right)\left(\lim_{x\to a}\frac{f(x)-f(a)}{x-a}\right),$$
d'où $\lim_{x\to a}f(x)=f(a)$. □

Remarque. La réciproque de cette proposition est fausse : une application peut être continue en a sans pour autant être dérivable en ce point. Par exemple, l'application $x\mapsto |x|$ de $\mathbb{R}$ dans $\mathbb{R}$, déjà étudiée plus haut, est continue en 0 sans y être dérivable.

10.1.2 Propriétés algébriques des fonctions dérivables en un point

Théorème 10.5 *Soient $a\in I$, $\lambda\in\mathbb{K}$ et $f,g\in\mathscr{F}(I,\mathbb{K})$ deux applications dérivables en a. Alors on a*

1) $f+g$ est dérivable en a et $(f+g)'(a)=f'(a)+g'(a)$.

2) λf est dérivable en a et $(\lambda f)'(a)=\lambda f'(a)$.

3) fg est dérivable en a et $(fg)'(a)=f'(a)g(a)+f(a)g'(a)$.

4) Si $g(a)\neq 0$, $\dfrac{f}{g}$ est dérivable en a et $\left(\dfrac{f}{g}\right)'(a)=\dfrac{f'(a)g(a)-f(a)g'(a)}{g(a)^2}$.

DÉMONSTRATION.

1) On a
$$\frac{(f+g)(x)-(f+g)(a)}{x-a}=\frac{f(x)-f(a)}{x-a}+\frac{g(x)-g(a)}{x-a}\underset{x\to a}{\longrightarrow}f'(a)+g'(a).$$

2) On a
$$\frac{(\lambda f)(x)-(\lambda f)(a)}{x-a}=\lambda\frac{f(x)-f(a)}{x-a}\underset{x\to a}{\longrightarrow}\lambda f'(a).$$

3) On a
$$\frac{(fg)(x)-(fg)(a)}{x-a}=\frac{f(x)g(x)-f(a)g(a)}{x-a}=\frac{(f(x)-f(a))g(a)}{x-a}+\frac{f(a)(g(x)-g(a))}{x-a}.$$
Puisque f et g sont dérivables en a, ces applications sont continues en a, donc $\lim_{x\to a}f(x)=f(a)$ et $\lim_{x\to a}g(x)=g(a)$, d'où $\lim_{x\to a}\dfrac{(f+g)(x)-(f+g)(a)}{x-a}=f'(a)g(a)+f(a)g'(a)$.

4) Puisque $g(a)\neq 0$ et que g est continue en a, on a, au voisinage de a, $g(x)\neq 0$. La fonction $\dfrac{1}{g}$ est alors définie au voisinage de a. De plus, on a
$$\frac{1}{x-a}\left(\left(\frac{1}{g}\right)(x)-\left(\frac{1}{g}\right)(a)\right)=\frac{1}{x-a}\left(\frac{1}{g(x)}-\frac{1}{g(a)}\right)=\frac{g(a)-g(x)}{x-a}\frac{1}{g(x)g(a)},$$
d'où $\lim_{x\to a}\dfrac{1}{x-a}\left(\left(\dfrac{1}{g}\right)(x)-\left(\dfrac{1}{g}\right)(a)\right)=-\dfrac{g'(a)}{g(a)^2}$. Le résultat se déduit alors de 2) en utilisant que $\dfrac{f}{g}=f\,\dfrac{1}{g}$.

□

Dérivée d'une composée de fonctions

Théorème 10.6 *Soient J un intervalle de $\mathbb{R}$, $f: I\to\mathbb{K}$ telle que $f(I)\subset J$, $g: J\to\mathbb{K}$ et $a\in I$. Si f est dérivable en a et si g est dérivable en $f(a)$, alors $g\circ f$ est dérivable en a et $(g\circ f)'(a)=f'(a)g'(f(a))$.*

DÉMONSTRATION. On a
$$\frac{(g\circ f)(x)-(g\circ f)(a)}{x-a}=\frac{g(f(x))-g(f(a))}{x-a}=\frac{g(f(x))-g(f(a))}{f(x)-f(a)}\frac{f(x)-f(a)}{x-a},$$
d'où
$$\lim_{x\to a}\frac{(g\circ f)(x)-(g\circ f)(a)}{x-a}=\left(\lim_{x\to a}\frac{g(f(x))-g(f(a))}{f(x)-f(a)}\right)\left(\lim_{x\to a}\frac{f(x)-f(a)}{x-a}\right)=g'(f(a))f'(a).$$
□

Dérivée d'une fonction réciproque

Théorème 10.7 *Soient $a \in I$ et $f : I \to \mathbb{R}$ une application continue et strictement monotone sur I, dérivable en a et telle que $f'(a) \neq 0$; on note $\widetilde{f} : I \to f(I)$ l'application qui à tout x de I associe $f(x)$. Alors la fonction réciproque $\widetilde{f}^{-1}$ est dérivable en $f(a)$ et l'on a $(\widetilde{f}^{-1})'(f(a)) = \dfrac{1}{f'(a)}$.*

DÉMONSTRATION. D'après le théorème 9.45, l'application $\widetilde{f} : I \to f(I)$ est bijective et sa bijection réciproque $\widetilde{f}^{-1}$ est strictement monotone, de même sens que f, et continue sur $f(I)$. Pour tout y de $f(I) \setminus \{f(a)\}$, on a alors

$$\frac{\widetilde{f}^{-1}(y) - \widetilde{f}^{-1}(f(a))}{y - f(a)} = \frac{\widetilde{f}^{-1}(y) - a}{f(\widetilde{f}^{-1}(y)) - f(a)}.$$

Comme f est dérivable en a, de dérivée non nulle en ce point, et que $\lim\limits_{y \to f(a)} \widetilde{f}^{-1}(y) = a$, on obtient, après composition des limites,

$$\lim_{y \to f(a)} \frac{\widetilde{f}^{-1}(y) - a}{f(\widetilde{f}^{-1}(y)) - f(a)} = \frac{1}{f'(a)}.$$

L'application $\widetilde{f}^{-1}$ est donc dérivable en $f(a)$. □

Exemples.
- On sait que la restriction de la fonction sinus à l'intervalle $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ est une bijection continue de cet intervalle sur $[-1, 1]$. Sa dérivée, la fonction cosinus, ne s'annule qu'aux points $-\frac{\pi}{2}$ et $\frac{\pi}{2}$, d'images respectives -1 et 1. Sa bijection réciproque arc sinus est donc continue sur $[-1, 1]$ et dérivable sur $]-1, 1[$, de dérivée
$$(\arcsin x)' = \frac{1}{\cos(\arcsin x)} = \frac{1}{\sqrt{1 - x^2}}, \ \forall x \in]-1, 1[.$$
- On sait que la restriction de la fonction cosinus à l'intervalle $[0, \pi]$ est une bijection continue de cet intervalle sur $[-1, 1]$. De la même façon que pour la fonction arc sinus, on montre que sa bijection réciproque arc cosinus est donc continue sur $[-1, 1]$ et dérivable sur $]-1, 1[$, de dérivée
$$(\arccos x)' = \frac{-1}{\sin(\arccos x)} = \frac{-1}{\sqrt{1 - x^2}}, \ \forall x \in]-1, 1[.$$
- On sait que la restriction de la fonction tangente à l'intervalle $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ est une bijection continue de cet intervalle sur $\mathbb{R}$. Sa dérivée, la fonction $x \mapsto 1 + \tan^2 x$, ne s'annule pas. Sa bijection réciproque arc tangente est donc dérivable sur $\mathbb{R}$, de dérivée
$$(\arctan x)' = \frac{1}{1 + \tan^2(\arctan x)} = \frac{1}{1 + x^2}, \ \forall x \in \mathbb{R}.$$

10.1.3 Application dérivée

Définition 10.8 *Soit $f \in \mathscr{F}(I, \mathbb{K})$. On appelle **dérivée de** f l'application qui à chaque x de I tel que $f'(x)$ existe associe $f'(x)$.*

10.1.4 Dérivées successives

Définitions 10.9 *Soit $f \in \mathscr{F}(I, \mathbb{K})$. On définit les **dérivées successives de** f par récurrence pour tout n de $\mathbb{N}^*$:*
- *pour $a \in I$, $f^{(n)}(a)$ est, si elle existe, la dérivée de $f^{(n-1)}$ en a,*
- *$f^{(n)}$ est l'application dérivée de $f^{(n-1)}$.*

*On appelle **dérivée n^e de f en a** le réel $f^{(n)}(a)$ et **application dérivée n^e de f** l'application $x \mapsto f^{(n)}(x)$.*
*On dit que f est n **fois dérivable sur** I si et seulement si $f^{(n)}$ est définie sur I. Enfin, on dit que f est **indéfiniment dérivable sur** I si et seulement si f est n fois dérivable sur I pour tout entier positif n.*

Par convention, $f^{(0)} = f$ et l'on note aussi $\dfrac{\mathrm{d}^n f}{\mathrm{d}x^n}$ au lieu de $f^{(n)}$. Comme on l'a déjà vu, on écrit souvent $f' = f^{(1)}$ et, de la même manière, $f'' = f^{(2)}$ et $f''' = f^{(3)}$.

Proposition 10.10 *Soient $\lambda \in \mathbb{K}$, $n \in \mathbb{N}^*$ et $f, g : I \to \mathbb{K}$ des applications n fois dérivables sur l'intervalle I. On a*

1) $f+g$ est n fois dérivable sur I et $(f+g)^{(n)} = f^{(n)} + g^{(n)}$.

2) λf est n fois dérivable sur I et $(\lambda f)^{(n)} = \lambda f^{(n)}$.

3) fg est n fois dérivable sur I et on a la ***formule de Leibniz*** [1] *suivante*

$$(fg)^{(n)} = \sum_{k=0}^{n} C_n^k f^{(k)} g^{(n-k)}.$$

4) Si $(\forall x \in I,\ g(x) \neq 0)$, alors $\dfrac{f}{g}$ est n fois dérivable.

DÉMONSTRATION. Tous ces résultats se démontrent par récurrence sur n. Nous laissons au lecteur les preuves (aisées) de 1) et de 2).

3) Le cas $n = 1$ a été traité dans le théorème 10.5. Supposons la propriété vraie au rang $n > 1$. Soient f et g deux fonctions de I dans $\mathbb{K}$, $(n+1)$ fois dérivables sur I. D'après l'hypothèse de récurrence, fg est n fois dérivable sur I et

$$(fg)^{(n)} = \sum_{k=0}^{n} C_n^k f^{(k)} g^{(n-k)}.$$

Ainsi, $(fg)^{(n)}$ apparaît comme somme de produits d'applications dérivables sur I et est donc dérivable sur I. On a alors

$$\begin{aligned}
\left((fg)^{(n)}\right)' &= \left(\sum_{k=0}^{n} C_n^k f^{(k)} g^{(n-k)}\right)' \\
&= \sum_{k=0}^{n} C_n^k f^{(k+1)} g^{(n-k)} + \sum_{k=0}^{n} C_n^k f^{(k)} g^{(n-k+1)} \\
&= \sum_{k=1}^{n+1} C_n^{k-1} f^{(k)} g^{(n-k+1)} + \sum_{k=0}^{n} C_n^k f^{(k)} g^{(n-k+1)} \\
&= f^{(n+1)} g + \sum_{k=1}^{n} \left(C_n^{k-1} + C_n^k\right) f^{(k)} g^{(n-k+1)} + f g^{(n+1)} \\
&= f^{(n+1)} g + \sum_{k=1}^{n} C_{n+1}^k f^{(k)} g^{(n-k+1)} + f g^{(n+1)} \\
&= \sum_{k=0}^{n+1} C_{n+1}^k f^{(k)} g^{(n+1-k)}.
\end{aligned}$$

4) Le cas $n = 1$ a déjà été vu dans le théorème 10.5. Supposons la propriété vraie au rang $n > 1$. Soient f et g deux fonctions de I dans $\mathbb{K}$, $(n+1)$ fois dérivables sur I et telles que $(\forall x \in I,\ g(x) \neq 0)$. L'application $\dfrac{f}{g}$ étant dérivable sur I, nous avons

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

Puisque f, f', g et g' sont n fois dérivables sur I, $f'g - fg'$ et g^2 le sont aussi. Il résulte alors de l'hypothèse de récurrence que $\dfrac{f'g - fg'}{g^2}$ est n fois dérivable sur I. Finalement, $\dfrac{f}{g}$ est $(n+1)$ fois dérivable sur I.

□

Nous introduisons enfin la notion de *classe* d'une fonction.

Définition 10.11 *Soit $f \in \mathscr{F}(I, \mathbb{K})$.*

1) Soit $n \in \mathbb{N}$. On dit que f est ***de classe $\mathscr{C}^n$ sur*** *I si et seulement si f est n fois dérivable sur I et $f^{(n)}$ est continue sur I.*

2) On dit que f est ***de classe $\mathscr{C}^\infty$ sur*** *I si et seulement si f est indéfiniment dérivable sur I.*

Pour $n \in \mathbb{N} \cup \{+\infty\}$, on note $\mathscr{C}^n(I, \mathbb{K})$ l'ensemble des applications de classe $\mathscr{C}^n$ de I dans $\mathbb{K}$.

1. Gottfried Wilhelm von Leibniz (1er juillet 1646 - 14 novembre 1716) était un philosophe, mathématicien (et plus généralement scientifique), bibliothécaire, diplomate et homme de loi allemand. Il inventa, indépendamment de Newton, le calcul intégral et différentiel et introduisit plusieurs notations mathématiques en usage aujourd'hui.

Remarques.
- $f \in \mathscr{C}^0(I,\mathbb{K})$ si et seulement si f est continue sur I.
- Pour tout $(p,n) \in (\mathbb{N} \cup \{+\infty\})^2$ tel que $p \leq n$, on a $\mathscr{C}^n(I,\mathbb{K}) \subset \mathscr{C}^p(I,\mathbb{K})$.

Exemples d'applications de classe $\mathscr{C}^\infty$.
- Les fonctions polynomiales sont de classe $\mathscr{C}^\infty$.
- Les fonctions exponentielle ($x \mapsto e^x$) et logarithme ($x \mapsto \ln x$) sont de classe $\mathscr{C}^\infty$ respectivement sur $\mathbb{R}$ et $\mathbb{R}_+^*$.
- Les égalités $(\sin)' = \cos$ et $(\cos)' = -\sin$ montrent que les applications sinus et cosinus sont de classe $\mathscr{C}^\infty$ sur $\mathbb{R}$.

10.2 Propriétés des fonctions dérivables

10.2.1 Extrema locaux d'une fonction réelle dérivable

Définitions 10.12 *Soient $a \in I$ et $f \in \mathscr{F}(I,\mathbb{R})$. On dit que f admet*
- *un **maximum local en** a si et seulement si, au voisinage de a, $f(x) \leq f(a)$,*
- *un **minimum local en** a si et seulement si, au voisinage de a, $f(x) \geq f(a)$,*
- *un **maximum local strict en** a si et seulement si, au voisinage de a sauf en a, $f(x) < f(a)$,*
- *un **minimum local strict en** a si et seulement si, au voisinage de a sauf en a, $f(x) > f(a)$,*
- *un **extremum local en** a si et seulement si f admet un maximum local ou un minimum local en a,*
- *un **extremum local strict en** a si et seulement si f admet un maximum local strict ou un minimum local strict en a.*

Exemples.
- Toute application constante admet en tout point un maximum et un minimum local.
- L'application de $\mathbb{R}$ de $\mathbb{R}$ qui à x associe $|x|$ admet un minimum local strict en 0.

Proposition 10.13 *Soit $f \in \mathscr{F}(I,\mathbb{R})$. Si f admet en un point intérieur a de I un extremum local et si f est dérivable en a, alors $f'(a) = 0$.*

DÉMONSTRATION. Supposons, pour fixer les idées, que f admet un maximum local en a. Puisque f est dérivable en a, on a

$$\begin{aligned} f'(a) &= \lim_{x\to a} \frac{f(x)-f(a)}{x-a} \\ &= \lim_{x\to a^+} \frac{f(x)-f(a)}{x-a} \ \geq\ 0 \\ &= \lim_{x\to a^-} \frac{f(x)-f(a)}{x-a} \ \leq\ 0, \end{aligned}$$

d'où $f'(a) = 0$. □

Remarques.
- La réciproque de cette proposition est fausse. Par exemple, l'application $x \mapsto x^3$ de $\mathbb{R}$ dans $\mathbb{R}$ à une dérivée nulle en 0 mais ne possède pas d'extremum en ce point.
- De fait, les extrema locaux d'une fonction définie sur un intervalle I seront recherchés aux points intérieurs de I où la dérivée de la fonction s'annule ou bien aux extrémités de I, où la fonction n'est pas dérivable.

10.2.2 Théorème de Rolle[2]

Théorème 10.14 *Soient $(a,b) \in \mathbb{R}^2$, tel que $a < b$, et f une application de $[a,b]$ dans $\mathbb{R}$. Si f est continue sur $[a,b]$, dérivable sur $]a,b[$ et telle que $f(a) = f(b)$, alors il existe $c \in]a,b[$ tel que $f'(c) = 0$.*

2. Michel Rolle (21 avril 1652 - 8 novembre 1719) était un mathématicien français. S'il inventa la notation $\sqrt[n]{x}$ pour désigner la racine n^e d'un réel x, il reste principalement connu pour avoir établi en 1691, dans le cas particulier des polynômes réels à une variable, une première version du théorème portant aujourd'hui son nom.

DÉMONSTRATION. Puisque l'application f est continue sur le segment $[a,b]$, elle est bornée et atteint ses bornes (proposition 9.42). Notons $m = \inf_{x\in[a,b]} f(x)$ et $M = \sup_{x\in[a,b]} f(x)$. Si $M = m$, alors f est constante et $f'(x) = 0$ pour tout $x \in]a,b[$. Supposons $m < M$. Comme $f(a) = f(b)$, on a soit $M \neq f(a)$, soit $m \neq f(a)$. Ramenons-nous au cas $M \neq f(a)$. Il existe alors un point $c \in]a,b[$ tel que $f(c) = M$. Soit $x \in [a,b]$ tel que $f(x) \leq M = f(c)$. Si $x > c$, on a $\frac{f(x)-f(c)}{x-c} \leq 0$, et si $x < c$, on obtient $\frac{f(x)-f(c)}{x-c} \geq 0$. L'application f étant dérivable en c, nous obtenons, en passant à la limite, $f'(c) \leq 0$ et $f'(c) \geq 0$, d'où $f'(c) = 0$. □

Remarque. Le réel c n'est pas nécessairement unique.

10.2.3 Théorème des accroissements finis

Théorème 10.15 *Soient $(a,b) \in \mathbb{R}^2$, tel que $a < b$, et f une application de $[a,b]$ dans $\mathbb{R}$. Si f est continue sur $[a,b]$ et dérivable sur $]a,b[$, alors il existe $c \in]a,b[$ tel que*

$$f'(c) = \frac{f(b)-f(a)}{b-a}.$$

DÉMONSTRATION. Considérons la fonction $\varphi : [a,b] \to \mathbb{R}$ définie par

$$\varphi(x) = f(x) - \frac{f(b)-f(a)}{b-a}(x-a).$$

Il est clair que φ est continue sur $[a,b]$, dérivable sur $]a,b[$ et que $\varphi(a) = \varphi(b)$. En appliquant le théorème de Rolle à φ, on obtient qu'il existe $c \in]a,b[$ tel que $\varphi'(c) = 0$, c'est-à-dire tel que

$$f'(c) = \frac{f(b)-f(a)}{b-a}.$$

□

Remarque. Là encore, le réel c n'est pas forcément unique.

Inégalité des accroissements finis

Nous déduisons directement du théorème 10.15 l'*inégalité des accroissements finis*. Celle-ci est plus générale que le théorème du même nom, dans la mesure où elle s'étend à d'autres fonctions que les fonctions d'une variable réelle à valeurs dans $\mathbb{R}$, comme par exemple les fonctions de $\mathbb{R}$ dans $\mathbb{C}$ ou de $\mathbb{R}^n$ ($n \in \mathbb{N}^*$) dans $\mathbb{R}$.

Théorème 10.16 *Soit $(a,b) \in \mathbb{R}^2$, tel que $a < b$. Si f est une fonction continue sur $[a,b]$, dérivable sur $]a,b[$ et qu'il existe un réel $M > 0$ tel que*

$$\forall x \in]a,b[,\ |f'(x)| \leq M,$$

alors on a

$$|f(b)-f(a)| \leq M|b-a|.$$

10.2.4 Sens de variation d'une fonction dérivable

Les résultats précédents permettent d'établir un lien entre le sens de variation d'une fonction et le signe de sa dérivée. Nous avons la

Proposition 10.17 *Soient $(a,b) \in \overline{\mathbb{R}}^2$, tel que $a < b$, et f une fonction dérivable sur $]a,b[$. Alors*

i) f est croissante si et seulement si $(\forall x \in]a,b[,\ f'(x) \geq 0)$,

ii) f est décroissante si et seulement si $(\forall x \in]a,b[,\ f'(x) \leq 0)$,

iii) f est constante si et seulement si $(\forall x \in]a,b[,\ f'(x) = 0)$.

DÉMONSTRATION.

i) Supposons f croissante. Soit $x_0 \in]a,b[$, pour tout $x \in]a,b[$ tel que $x > x_0$, on a

$$\frac{f(x)-f(x_0)}{x-x_0} \geq 0.$$

En passant à la limite $x \to x_0$, on déduit que $f'(x_0) \geq 0$. Réciproquement, supposons que, pour tout $x \in]a,b[$, $f'(x) \geq 0$. Soient x_1 et x_2 deux éléments de $]a,b[$ tels que $x_1 < x_2$. En appliquant le théorème des accroissements finis à f sur $[x_1,x_2]$, on voit qu'il existe $c \in [x_1,x_2]$ tel que

$$f(x_2)-f(x_1) = f'(c)(x_2-x_1) \geq 0,$$

et f est par conséquent croissante sur]a,b[.

ii) L'application f est décroissante si et seulement si $-f$ est croissante ; ii) résulte donc de i).

iii) L'application f est constante si et seulement si elle est à la fois croissante et décroissante ; iii) résulte donc de i) et de ii).

□

10.3 Formule de Taylor–Lagrange

Nous terminons ce chapitre sur le théorème de Taylor[3]–Lagrange[4], qui constitue une généralisation du théorème des accroissements finis et que nous présentons sous différentes formes.

Théorème 10.18 ***(« formule de Taylor–Lagrange »)*** *Soient $(a,b) \in \mathbb{R}^2$, $a < b$, et $f : [a,b] \to \mathbb{R}$ une fonction de classe $\mathscr{C}^n$ sur $[a,b]$. On suppose de plus que $f^{(n)}$ est dérivable sur $]a,b[$. Alors, il existe $c \in]a,b[$ tel que*

$$f(b) = f(a) + \frac{f'(a)}{1!}(b-a) + \frac{f''(a)}{2!}(b-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(b-a)^n + \frac{f^{(n+1)}(c)}{(n+1)!}(b-a)^{n+1},$$

soit encore

$$f(b) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(b-a)^k + \frac{f^{(n+1)}(c)}{(n+1)!}(b-a)^{n+1}.$$

Comme dans le cas de l'égalité du théorème des accroissements finis, la preuve de la formule de Taylor–Lagrange consiste en l'application du théorème de Rolle à une certaine fonction.

DÉMONSTRATION. Soit A le réel tel que

$$\frac{(b-a)^{n+1}}{(n+1)!}A = f(b) - \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(b-a)^k$$

Il s'agit de montrer que $A = f^{(n+1)}(c)$, avec $c \in]a,b[$. On définit pour cela la fonction $\varphi : [a,b] \to \mathbb{R}$ comme suit

$$\varphi(x) = f(b) - \sum_{k=0}^{n} \frac{f^{(k)}(x)}{k!}(b-x)^k - \frac{(b-x)^{n+1}}{(n+1)!}A.$$

Cette fonction est continue sur $[a,b]$, dérivable sur $]a,b[$ et vérifie d'autre part $\varphi(a) = \varphi(b) = 0$. D'après le théorème de Rolle, il existe donc $c \in]a,b[$ tel que $\varphi'(c) = 0$. Or, pour tout $x \in]a,b[$, on a

$$\begin{aligned}\varphi'(x) &= \sum_{k=1}^{n} \frac{f^{(k)}(x)}{(k-1)!}(b-x)^{k-1} - \sum_{k=0}^{n} \frac{f^{(k+1)}(x)}{k!}(b-x)^k + \frac{(b-x)^n}{n!}A \\ &= \frac{(b-x)^n}{n!}\left(-f^{(n+1)}(x) + A\right).\end{aligned}$$

Par conséquent, on déduit de $\varphi'(c) = 0$ que $A = f^{(n+1)}(c)$. □

3. Brook Taylor (18 août 1685 - 30 novembre 1731) était un mathématicien, artiste peintre et musicien anglais. Il inventa le calcul aux différences finies et découvrit l'intégration par parties.

4. Joseph Louis Lagrange (Giuseppe Lodovico Lagrangia en italien, 25 janvier 1736 - 10 avril 1813) était un mathématicien et astronome franco-italien. Fondateur, avec Euler, du calcul des variations, il a également produit d'importantes contributions tant en analyse, en géométrie et en théorie des groupes qu'en mécanique.

Remarques.

— Le terme $\dfrac{f^{(n+1)}(c)}{(n+1)!}(b-a)^{n+1}$ est appelé *reste de Lagrange*.

— Dans le cas particulier $n = 0$, on retrouve l'égalité du théorème des accroissements finis.

Considérons à présent un intervalle I non nécessairement fermé ou borné. Dans ce cas, le théorème de Taylor–Lagrange s'énonce encore de la manière suivante.

Théorème 10.19 *Soient f une fonction $(n+1)$ fois dérivable sur un intervalle I et x un point de I. Alors, pour tout réel h tel que $x+h \in I$, il existe $\theta \in]0,1[$ tel que*

$$f(x+h) = \sum_{k=0}^{n} \frac{f^{(k)}(x)}{k!}h^k + \frac{f^{(n+1)}(x+\theta h)}{(n+1)!}h^{n+1}.$$

DÉMONSTRATION. Dans le cas $h > 0$, il suffit d'appliquer le théorème précédent avec $a = x$ et $b = x+h$; sinon, on reprend la démonstration ci-dessus. □

Remarque. Si l'intervalle I contient l'origine, on déduit de ce dernier théorème la *formule de Maclaurin* [5] : pour tout point x de I, il existe $\theta \in]0,1[$ tel que

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!}x^k + \frac{f^{(n+1)}(\theta x)}{(n+1)!}x^{n+1}.$$

Inégalité de Taylor–Lagrange

Comme pour l'égalité des accroissements finis, il existe une *inégalité de Taylor–Lagrange* plus générale que la formule du même nom.

Théorème 10.20 *Soit f une fonction $(n+1)$ fois dérivable sur un intervalle I. On suppose qu'il existe un réel M tel que $|f^{(n+1)}(x)| \le M$ pour tout x de I. Soient a et b deux points de I ; on a alors*

$$\left| f(b) - \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(b-a)^k \right| \le M\frac{|b-a|^{n+1}}{(n+1)!}.$$

Remarque. Pour $n = 0$, on constate que l'on retrouve l'inégalité des accroissements finis.

Exemples d'applications de l'inégalité de Taylor–Lagrange. Considérons l'application $x \mapsto \sin x$ sur l'intervalle $[0,h]$, $h \in \mathbb{R}$. Nous obtenons

$$n = 2,\ |\sin h - h| \le \frac{|h|^3}{3!}.$$

$$n = 4,\ \left|\sin h - h + \frac{h^3}{3!}\right| \le \frac{|h|^5}{5!}.$$

$$n = 6,\ \left|\sin h - h + \frac{h^3}{3!} - \frac{h^5}{5!}\right| \le \frac{|h|^7}{7!}.$$

Considérons l'application $x \mapsto \cos x$ sur l'intervalle $[0,h]$, $h \in \mathbb{R}$. Nous obtenons

$$n = 1,\ |\cos h - 1| \le \frac{|h|^2}{2!}.$$

$$n = 3,\ \left|\cos h - 1 + \frac{h^2}{2!}\right| \le \frac{|h|^4}{4!}.$$

$$n = 5,\ \left|\cos h - 1 + \frac{h^2}{2!} - \frac{h^4}{4!}\right| \le \frac{|h|^6}{6!}.$$

5. Colin Maclaurin (février 1698 - 14 juin 1746) était un mathématicien écossais. Il fit des travaux remarquables en géométrie, plus précisément dans l'étude de courbes planes, et écrivit un important mémoire sur la théorie des marées.