Cours Mathématiques MPSI 2021/2022

Première partie

Algèbre

Chapitre 1

Eléments de mathématiques

1.1 Les objets

1.1.1 Ensembles et éléments

Définition

On appelle ensemble toute collection d'objets appelés éléments de cet ensemble. Pour signifier qu'un élément x appartient à un ensemble E, on écrit $x \in E$. Sinon, on écrit $x \notin E$.

Exemple $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ et $\mathbb{C}$ sont les ensembles de nombres usuels.

Exemple $\{a, b, ..., s\}$ désigne l'ensemble constitué des éléments $a, b, ..., s$ et uniquement cela.

Exemple $\{2k/k \in \mathbb{Z}\}$ désigne l'ensemble des éléments de la forme $2k$ avec k décrivant $\mathbb{Z}$, à savoir les nombres pairs.

Définition

Deux ensembles E et F sont dits égaux s'ils sont constitués des mêmes éléments. On note alors $E = F$.

Exemple $\{a, b\} = \{b, a\}$.

Définition

On appelle ensemble vide l'ensemble constitué d'aucun élément, on le note $\emptyset$.

Remarque La notation $\{\}$ est caduque. La notation $\{\emptyset\}$ ne désigne par l'ensemble vide mais un ensemble constitué d'un élément qui est l'ensemble vide.

Définition

Etant donnés deux ensembles E et F, on appelle intersection de E et F l'ensemble $E \cap F$ formés des éléments communs à E et F.

Définition

Etant donnés deux ensembles E et F, on appelle union de E et F l'ensemble $E \cup F$ formés des éléments de l'un et de l'autre ensemble.

1.1.2 Inclusion

E désigne un ensemble.

Définition

Un ensemble F est dit inclus dans E si tous les éléments de F sont aussi éléments de E. On note alors $F \subset E$.

Exemple $\{a, c\} \subset \{a, b, c\}$.

Définition

On appelle partie (ou sous-ensemble) de E, tout ensemble F dont les éléments sont tous éléments de E c'est-à-dire tout ensemble inclus dans F.

Exemple $\emptyset$ et E sont des parties de E.

Définition

On appelle ensemble des parties de E l'ensemble, noté $\mathcal{P}(E)$, formé des sous-ensembles de E.

Exemple Si $E = \{a, b, c\}$ alors

$$\mathcal{P}(E) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{b, c\}, \{c, a\}, \{a, b, c\}\}$$

1.1.3 Produit cartésien

1.1.3.1 Couple

Définition

A partir de deux éléments a et b, on forme un nouvel élément appelé couple (a, b) défini de sorte que :

$$(a, b) = (a', b') \Leftrightarrow a = a' \text{ et } b = b'$$

Remarque Lorsque $a \neq b$, $(a, b) \neq (b, a)$.

Définition

On appelle produit cartésien de E par F l'ensemble formé des couples (a, b) avec a dans E et b dans F. On le note $E \times F$.

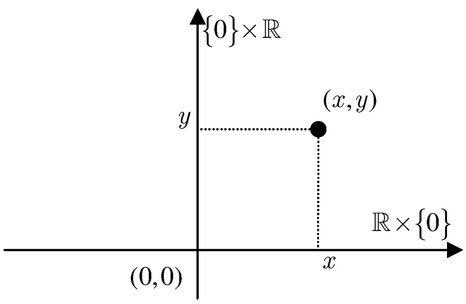
Exemple Pour $E = \{a, b, c\}$ et $F = \{1, 2\}$

$$E \times F = \{(a, 1), (a, 2), (b, 1), (b, 2), (c, 1), (c, 2)\}$$

Remarque Lorsque E et F sont des ensembles distincts non vides : $E \times F \neq F \times E$.

Remarque Lorsque $E = F$, il est usuel de noter E^2 au lieu de $E \times E$.

Exemple $\mathbb{R}^2 = \{(x, y)/x \in \mathbb{R}, y \in \mathbb{R}\}$ est usuellement visualisé comme un plan.



1.1.3.2 Multiplet

Définition

A partir d'éléments $a_1, ..., a_n$ (avec $n \in \mathbb{N}^\star$), on forme le n uplet $(a_1, ..., a_n)$ défini de sorte que :

$$(a_1, \ldots, a_n) = (a'_1, \ldots, a'_n) \Leftrightarrow \forall i \in \{1, \ldots, n\}, a_i = a'_i$$

Définition

On appelle produit cartésien des ensembles $E_1, ..., E_n$ (avec $n \in \mathbb{N}^\star$) l'ensemble formé des n uplets $(a_1, ..., a_n)$ avec pour tout $i \in \{1, 2, \ldots, n\}$, $a_i \in E_i$.

On le note $E_1 \times \cdots \times E_n$ ou encore $\prod\limits_{i=1}^{n} E_i$.

Remarque Si $E_1 = \cdots = E_n = E$ alors on note E^n au lieu de $E \times \cdots \times E$ (n termes).

Exemple $\mathbb{R}^3 = \{(x, y, z)/x \in \mathbb{R}, y \in \mathbb{R}, z \in \mathbb{R}\}$ est usuellement visualisé comme un espace de dimension 3.

Exemple On pressent que $\mathbb{R}^n$ permet de visualiser la dimension n...

1.1.4 Fonctions et applications

E et F désignent des ensembles

Définition

> Une application (ou fonction) f de E vers F est une « manipulation »qui à chaque élément x de E associe un et un seul élément y de F.
> L'élément y est alors noté $f(x)$ et est appelé image de x par f.
> On note $f : E \to F$ pour signifier que f est une application de E vers F.
> On note $\mathcal{F}(E, F)$ l'ensemble des applications de E vers F.

Remarque Pour définir une application $f : E \to F$ il suffit de préciser comment à chaque élément x de E est associé son image $f(x)$ dans F.
C'est le principe des notations :
-$f : \begin{cases} E \to F \\ x \mapsto ... \end{cases}$;
- $f : E \to F$ définie par $f(x) = \dots$;
- $f : x \mapsto \dots$ définie sur E et à valeurs dans F.

Exemple $\begin{cases} \mathbb{R} \to \mathbb{R} \\ x \mapsto \dfrac{\mathrm{e}^x}{x^2+1} \end{cases}$ est une application.

Exemple $\begin{cases} \mathbb{R} \to \mathbb{R} \\ x \mapsto \begin{cases} x \text{ si } x \geqslant 0 \\ -x \text{ sinon} \end{cases} \end{cases}$ est une application.

Exemple $\begin{cases} \mathbb{N} \to \mathbb{Z} \\ n \mapsto n^2 - n + 1 \end{cases}$ est une application.

Exemple $\begin{cases} \mathbb{R}^2 \to \mathbb{R} \\ (x, y) \mapsto x + y \end{cases}$ est une application.

Exemple $\begin{cases} \mathcal{F}(\mathbb{R}, \mathbb{R}) \to \mathbb{R} \\ f \mapsto f(0) \end{cases}$ est une application.

1.2 Notions de logique

1.2.1 Assertion

Définition

On appelle assertion toute phrase mathématique significative susceptible d'être vraie (V) ou fausse (F).

Exemple $\mathcal{P} =$ « $3 \geqslant 2$ »est une assertion vraie.
$\mathcal{Q} =$ « $2 + 2 = 5$ »est une assertion fausse.

Remarque Lorsque la valeur de vérité d'une assertion $\mathcal{P}$ dépend des valeurs prises par un paramètre x (resp. par plusieurs paramètres $x, y, \dots$) on note souvent celle-ci $\mathcal{P}(x)$ (resp. $\mathcal{P}(x, y, \dots)$) pour le souligner.
On parle parfois de prédicat plutôt que d'assertion.

Exemple $\mathcal{P}(x) =$ « $x \geqslant 0$ »est une assertion dépendant d'un paramètre x réel.
$\mathcal{Q}(x) =$ « $x + y = z$ »est une assertion dépendant de x, y, z réels.
$\mathcal{P}(2)$ et $\mathcal{Q}(1, 2, 3)$ sont vraies.

Définition

Deux assertions $\mathcal{P}$ et $\mathcal{Q}$ ayant mêmes valeurs de vérité sont dites équivalentes et on note $\mathcal{P} \sim \mathcal{Q}$.

Exemple « $x \geqslant 0$ »$\sim$ « $-x \leqslant 0$ »

Définition

Soit $\mathcal{P}(x)$ une assertion dépendant d'un paramètre x élément de E.
On note $\{x \in E \text{ tel que } \mathcal{P}(x)\}$ ou $\{x \in E / \mathcal{P}(x)\}$ la partie de E formée des éléments x qui rendent l'assertion $\mathcal{P}(x)$ vraie. On dit que l'ensemble est défini en compréhension par opposition avec un ensemble dont on liste les éléments qui est dit défini en extension.

Exemple $\{x \in \mathbb{R} / 1 \leqslant x < 2\} = [1, 2[$,
$\left\{x \in \mathbb{R} / x^2 > 0\right\} = \mathbb{R}^\star$.
$\{n \in \mathbb{Z} / n \text{ est pair}\} = \{2k / k \in \mathbb{Z}\}$.

Remarque Résoudre une équation consiste à décrire un ensemble défini en compréhension (i.e. défini par l'équation étudiée) en un ensemble défini par la description de ses éléments.

1.2.2 Négation

Soit $\mathcal{P}$ une assertion.

Définition

On appelle négation de $\mathcal{P}$, l'assertion notée non$(\mathcal{P})$ définie comme étant vraie lorsque $\mathcal{P}$ est fausse et inversement.
On peut aussi dire que l'assertion non$(\mathcal{P})$ est définie par la table de vérité :

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

Exemple Pour $\mathcal{P}(x) =$ « $x \geqslant 0$ »on a non$(\mathcal{P}(x)) \sim$ « $x < 0$ »

Proposition

non(non$(\mathcal{P})) \sim \mathcal{P}$.

dém. :

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

□

Remarque On note parfois $\neg\mathcal{P}$ ou $\bar{\mathcal{P}}$ au lieu de non$(\mathcal{P})$.

1.2.3 Conjonction et disjonction

Soient $\mathcal{P}$, $\mathcal{Q}$ et $\mathcal{R}$ des assertions.

Définition

On appelle conjonction (resp. disjonction) de ces deux assertions, l'assertion notée $\mathcal{P}$ et $\mathcal{Q}$ (resp. $\mathcal{P}$ ou $\mathcal{Q}$) définie comme étant vraie si, et seulement si, $\mathcal{P}$ et $\mathcal{Q}$ le sont toutes les deux (resp. lorsqu'au moins l'une des deux l'est). On a donc :

$\mathcal{P}$	$\mathcal{Q}$	$\mathcal{P}$ et $\mathcal{Q}$	$\mathcal{P}$ ou $\mathcal{Q}$
V	V	V	V
V	F	F	V
F	V	F	V
F	F	F	F

Exemple « $0 \leqslant x \leqslant 1$ »$\sim$ « $x \geqslant 0$ et $x \leqslant 1$ »
« $x \geqslant 0$ ou $x \leqslant 0$ » est une assertion vraie pour tout x réel.

*(!)*Le ou français est souvent exclusif.
Le ou mathématique est inclusif.

Proposition

non($\mathcal{P}$ et $\mathcal{Q}$) $\sim$ non($\mathcal{P}$) ou non($\mathcal{Q}$).
non($\mathcal{P}$ ou $\mathcal{Q}$) $\sim$ non($\mathcal{P}$) et non($\mathcal{Q}$).

dém. :
Ces propriétés s'obtiennent par étude de tables de vérité.
□

Proposition

$\mathcal{P}$ et $\mathcal{P} \sim \mathcal{P}$, $\mathcal{P}$ ou $\mathcal{P} \sim \mathcal{P}$.
$\mathcal{P}$ et $\mathcal{Q} \sim \mathcal{Q}$ et $\mathcal{P}$, $\mathcal{P}$ ou $\mathcal{Q} \sim \mathcal{Q}$ ou $\mathcal{P}$,
($\mathcal{P}$ et $\mathcal{Q}$) et $\mathcal{R} \sim \mathcal{P}$ et ($\mathcal{Q}$ et $\mathcal{R}$) (que l'on note alors $\mathcal{P}$ et $\mathcal{Q}$ et $\mathcal{R}$),
($\mathcal{P}$ ou $\mathcal{Q}$) ou $\mathcal{R} \sim \mathcal{P}$ ou ($\mathcal{Q}$ ou $\mathcal{R}$) (que l'on note alors $\mathcal{P}$ ou $\mathcal{Q}$ ou $\mathcal{R}$),
$\mathcal{P}$ et ($\mathcal{Q}$ ou $\mathcal{R}$) $\sim$ ($\mathcal{P}$ et $\mathcal{Q}$) ou ($\mathcal{P}$ et $\mathcal{R}$),
$\mathcal{P}$ ou ($\mathcal{Q}$ et $\mathcal{R}$) $\sim$ ($\mathcal{P}$ ou $\mathcal{Q}$) et ($\mathcal{P}$ ou $\mathcal{R}$).

dém. :
Ces propriétés s'obtiennent par étude de tables de vérité.
□

Attention : Ecrire $\mathcal{P}$ ou $\mathcal{Q}$ et $\mathcal{R}$ sans parenthèses n'est pas compréhensible !

Remarque On note parfois $\mathcal{P} \wedge \mathcal{Q}$ (resp. $\mathcal{P} \vee \mathcal{Q}$) au lieu de $\mathcal{P}$ et $\mathcal{Q}$ (resp. $\mathcal{P}$ ou $\mathcal{Q}$).

1.2.4 Implications

Soient $\mathcal{P}$ et $\mathcal{Q}$ deux assertions.

Définition

On définit l'assertion $\mathcal{P} \Rightarrow \mathcal{Q}$ comme étant vraie si $\mathcal{Q}$ ne peut pas être fausse quand $\mathcal{P}$ est vraie.
En français l'implication est traduite pas les expressions : « si... alors », « donc », « par suite »etc.
Plus précisément, la valeur de vérité de l'assertion $\mathcal{P} \Rightarrow \mathcal{Q}$ est donnée par :

$\mathcal{P}$	$\mathcal{Q}$	$\mathcal{P} \Rightarrow \mathcal{Q}$
V	V	V
V	F	F
F	V	V
F	F	V

Exemple Pour x réel : $x \geqslant 1 \Rightarrow x^2 \geqslant 1$ est une implication vraie.

Attention : Lorsque $\mathcal{P} \Rightarrow \mathcal{Q}$ est vraie, on ne présume rien sur la valeur de vérité de $\mathcal{P}$.

Remarque Lorsque $\mathcal{P} \Rightarrow \mathcal{Q}$ est vraie :
- si $\mathcal{P}$ est vraie alors $\mathcal{Q}$ l'est aussi ;
- si $\mathcal{P}$ est fausse alors on ne sait rien sur la valeur de vérité de $\mathcal{Q}$.

Définition

Lorsque $\mathcal{P} \Rightarrow \mathcal{Q}$ est vraie on dit que :
- $\mathcal{P}$ est une condition suffisante (CS) pour $\mathcal{Q}$;
- $\mathcal{Q}$ est une condition nécessaire (CN) pour $\mathcal{P}$.

Définition

$\mathcal{Q} \Rightarrow \mathcal{P}$ est appelée implication réciproque de $\mathcal{P} \Rightarrow \mathcal{Q}$.

Proposition

$(\mathcal{P} \Rightarrow \mathcal{Q}) \sim$ non$(\mathcal{P})$ ou $\mathcal{Q}$.

dém. :

$\mathcal{P}$	$\mathcal{Q}$	$\mathcal{P} \Rightarrow \mathcal{Q}$	non$(\mathcal{P})$	non$(\mathcal{P})$ ou $\mathcal{Q}$
V	V	V	F	V
V	F	F	F	F
F	V	V	V	V
F	F	V	V	V

□

Proposition

$\mathcal{P} \Rightarrow \mathcal{Q} \sim$ non$(\mathcal{Q}) \Rightarrow$ non$(\mathcal{P})$

dém. :
non$(\mathcal{Q}) \Rightarrow$ non$(\mathcal{P}) \sim \mathcal{Q}$ ou non$(\mathcal{P}) \sim \mathcal{P} \Rightarrow \mathcal{Q}$.
□

Définition

non$(\mathcal{Q}) \Rightarrow$ non$(\mathcal{P})$ est appelée contraposée de l'implication $\mathcal{P} \Rightarrow \mathcal{Q}$.

Exemple La contraposée de $x \geqslant 1 \Rightarrow x^2 \geqslant 1$ est $x^2 < 1 \Rightarrow x < 1$.

Proposition

non$(\mathcal{P} \Rightarrow \mathcal{Q}) \sim \mathcal{P}$ et non$(\mathcal{Q})$.

dém. :
Par négation de non$(\mathcal{P})$ ou $\mathcal{Q}$.
□

*(!)*non$(\mathcal{P} \Rightarrow \mathcal{Q})$ n'a pas la sens de $\mathcal{P} \Rightarrow$ non$(\mathcal{Q})$
Par passage à la négation le symbole d'implication disparaît.

Exemple $x \geqslant 0 \Rightarrow x^2 \geqslant 1$ et $x \geqslant 0 \Rightarrow x^2 < 1$ sont des implications toutes deux fausses.

1.2.5 Equivalence

Soient $\mathcal{P}$ et $\mathcal{Q}$ deux assertions.

Définition

On note $\mathcal{P} \Leftrightarrow \mathcal{Q}$ l'assertion $\mathcal{P} \Rightarrow \mathcal{Q}$ et $\mathcal{Q} \Rightarrow \mathcal{P}$.
En français l'équivalence se traduit par les expressions : « si, et seulement si »(ssi), « il faut et il suffit »,...
La table de vérité de $\mathcal{P} \Leftrightarrow \mathcal{Q}$ est donnée par :

$\mathcal{P}$	$\mathcal{Q}$	$\mathcal{P} \Rightarrow \mathcal{Q}$	$\mathcal{Q} \Rightarrow \mathcal{P}$	$\mathcal{P} \Leftrightarrow \mathcal{Q}$
V	V	V	V	V
V	F	F	V	F
F	V	V	F	F
F	F	V	V	V

Exemple Pour x, y réels : $x^2 = y^2 \Leftrightarrow x = y$ ou $x = -y$.

Remarque Lorsque $\mathcal{P} \Leftrightarrow \mathcal{Q}$ est vraie, on peut dire que $\mathcal{P}$ et $\mathcal{Q}$ ont mêmes valeurs de vérité et donc $\mathcal{P} \sim \mathcal{Q}$

Définition

Lorsque $\mathcal{P} \Leftrightarrow \mathcal{Q}$ est vraie, on dit que $\mathcal{P}$ et $\mathcal{Q}$ sont équivalentes et que $\mathcal{P}$ est un condition nécessaire et suffisante (CNS) pour $\mathcal{Q}$.

Proposition

$\mathcal{P} \Leftrightarrow \mathcal{Q} \sim$ non($\mathcal{P}$) $\Leftrightarrow$ non($\mathcal{Q}$).

*(!)*Ne pas écrire d'équivalences abusives !
Chaque équivalence correspond à deux implications et nécessite donc une double réflexion !

Exemple Pour x, y réels : $x = y \not\Leftrightarrow x^2 = y^2$.

1.2.6 Quantificateurs

Soit $\mathcal{P}(x)$ une assertion dépendant d'un élément $x \in E$.

Définition

On définit l'assertion

$$\forall x \in E, \mathcal{P}(x)$$

comme étant vraie lorsque $\mathcal{P}(x)$ est vraie pour tout x dans E.
Cette assertion se lit : « Quel que soit x dans E on a $\mathcal{P}(x)$ »

Exemple $\forall x \in \mathbb{R}, x^2 \geqslant 0$ est vraie.
$\forall x \in [-1, 1], x^2 \leqslant 2$ est vraie.
$\forall x \in \mathbb{R}, x^2 \leqslant 2$ est fausse.

Remarque Ecrire $\forall x, \mathcal{P}(x)$ est insuffisant !

Remarque Dans l'assertion

$$\forall x \in E, \mathcal{P}(x)$$

la lettre x à un rôle muet i.e. qu'elle peut être remplacée par n'importe quelle autre lettre.

Définition

> On définit l'assertion
>
> $$\exists x \in E, \mathcal{P}(x)$$
>
> comme étant vraie lorsque $\mathcal{P}(x)$ est vraie pour au moins un x dans E.
> Cette assertion se lit : « Il existe x dans E tel que $\mathcal{P}(x)$ ».

Exemple $\exists x \in \mathbb{R}, x^2 = 1$ est vraie.
$\exists x \in \mathbb{R}, x^2 < 0$ est fausse.

Remarque Les remarques précédentes sont encore valables.

Définition

> On définit l'assertion
>
> $$\exists! x \in E, \mathcal{P}(x)$$
>
> comme étant vraie lorsque $\mathcal{P}(x)$ est vraie pour un et un seul élément x dans E.
> Cette assertion se lit : « Il existe un unique x dans E tel que $\mathcal{P}(x)$ ».

Exemple $\exists! x \in \mathbb{R}^{+\star}, \ln(x) = 1$ (c'est le nombre de Neper)

Exemple $\forall x \in \mathbb{R}, \exists! n \in \mathbb{Z}, n \leqslant x < n + 1$.
Dans cette assertion, l'entier n apparaît après le x et est par suite susceptible de dépendre de x, on le note parfois n_x afin de le souligner.
Cette assertion est vraie et pour chaque x, l'entier n introduit est appelé partie entière de x.

*(!)*Il ne faut par intervertir sans justification les $\forall$ et les $\exists$:
« $\forall x \in \mathbb{R}, \exists n \in \mathbb{Z}, x \leqslant n$ »est vraie alors que « $\exists n \in \mathbb{Z}, \forall x \in \mathbb{R}, x \leqslant n$ »est fausse.

Remarque Abusivement, on écrit :
$\forall x \geqslant 0$ au lieu de $\forall x \in \mathbb{R}^+$,
$\forall 0 \leqslant x \leqslant 1$ au lieu de $\forall x \in [0, 1]$,
$\forall x, y \in E$ au lieu de $\forall x \in E, \forall y \in E$ ou de $\forall (x, y) \in E^2$,...

Proposition

non$(\forall x \in E, \mathcal{P}(x)) \sim \exists x \in E,$ non$(\mathcal{P}(x))$,
non$(\exists x \in E, \mathcal{P}(x)) \sim \forall x \in E,$ non$(\mathcal{P}(x))$.

dém. :
C'est du bon sens !
□
Convention :
Toute assertion commençant par : $\exists x \in \emptyset$ est fausse.
Par négation : toute assertion commençant $\forall x \in \emptyset$ est vraie.

Exemple Soit $f : \mathbb{R} \to \mathbb{R}$ une application.
- la fonction f est la fonction nulle :
$$\forall x \in \mathbb{R}, f(x) = 0\,;$$
- la fonction f s'annule :
$$\exists x \in \mathbb{R}, f(x) = 0\,;$$
- la fonction f s'annule une seule fois :
$$\exists! x \in \mathbb{R}, f(x) = 0\,;$$
- la fonction f s'annule sur $\mathbb{R}^+$:
$$\exists x \in \mathbb{R}^+, f(x) = 0\,;$$
- la fonction f ne s'annule que sur $\mathbb{R}^+$:
$$\forall x \in \mathbb{R}, f(x) = 0 \Rightarrow x \geqslant 0\,;$$
- la fonction f ne prend que des valeurs positives :
$$\forall x \in \mathbb{R}, f(x) \geqslant 0\,;$$
- la fonction f ne prend des valeurs positives que sur $\mathbb{R}^+$:
$$\forall x \in \mathbb{R}, f(x) \geqslant 0 \Rightarrow x \geqslant 0 :$$
- la fonction f est constante :
$$\forall x, y \in \mathbb{R}, f(x) = f(y)$$
ou encore
$$\exists C \in \mathbb{R}, \forall x \in \mathbb{R}, f(x) = C\,;$$
- la fonction f est croissante :
$$\forall x, y \in \mathbb{R}, x \leqslant y \Rightarrow f(x) \leqslant f(y)\,;$$
- tout réel possède un antécédent par f :
$$\forall y \in \mathbb{R}, \exists x \in \mathbb{R}, f(x) = y\,;$$
- la fonction f prend des valeurs deux à deux distincts :
$$\forall x, y \in \mathbb{R}, x \neq y \Rightarrow f(x) \neq f(y)$$
ou encore
$$\forall x, y \in \mathbb{R}, f(x) = f(y) \Rightarrow x = y$$

1.3 Raisonnements

Une assertion vraie est appelée énoncé, proposition ou théorème.
La véracité d'une assertion se justifie par une démonstration.
Certaines assertions sont postulées vraies sans démonstration, ce sont les axiomes.

1.3.1 Démonstration d'une assertion

Pour démontrer la véracité d'une assertion $\mathcal{P}$ on peut procéder de trois manières :
(1) Montrer que $\mathcal{P}$ découle de résultats antérieurs i.e. déterminer un énoncé $\mathcal{Q}$ tel que $\mathcal{Q} \Rightarrow \mathcal{P}$ soit vraie.

Exemple Montrons

$$\forall x \in \mathbb{R}, x^2 + 1 > 0$$

Soit $x \in \mathbb{R}$.
On sait que $x^2 \geqslant 0$ et $1 > 0$.
Or

$$a \geqslant 0 \text{ et } b > 0 \Rightarrow a + b > 0$$

donc $x^2 + 1 > 0$.

(2) Opérer par disjonction de cas i.e. déterminer un énoncé $\mathcal{Q}$ tel que $\mathcal{Q} \Rightarrow \mathcal{P}$ et non$(\mathcal{Q}) \Rightarrow \mathcal{P}$ soient vraies.

Exemple Montrons que

$$\forall n \in \mathbb{N}, \frac{n(n+1)}{2} \in \mathbb{N}$$

Soit $n \in \mathbb{N}$.
Si n est pair alors on peut écrire $n = 2k$ avec $k \in \mathbb{N}$ et alors

$$\frac{n(n+1)}{2} = k(2k+1) \in \mathbb{N}$$

Si n est impair alors on peut écrire $n = 2k + 1$ avec $k \in \mathbb{N}$ et alors

$$\frac{n(n+1)}{2} = (2k+1)(k+1) \in \mathbb{N}$$

Dans les deux cas $\dfrac{n(n+1)}{2} \in \mathbb{N}$.

(3) Raisonner par l'absurde i.e. montrer que non$(\mathcal{P})$ implique un résultat faux.

Exemple Montrons qu'il n'existe pas d'entier naturel supérieur à tout autre.
Par l'absurde : Supposons qu'il existe $N \in \mathbb{N}$ tel que $\forall n \in \mathbb{N}, n \leqslant N$.
Pour $n = N + 1 \in \mathbb{N}$, on a $N + 1 \leqslant N$ donc $1 \leqslant 0$. Absurde.

*(!)*En aucun cas, on ne commence le raisonnement par « Supposons $\mathcal{P}$ ».

1.3.2 Démonstration d'une implication

Pour démontrer la véracité d'une implication $\mathcal{P} \Rightarrow \mathcal{Q}$ on peut procéder de deux manières :
(1) Par déduction : on détermine une assertion $\mathcal{R}$ telle que : $\mathcal{P} \Rightarrow \mathcal{R}$ et $\mathcal{R} \Rightarrow \mathcal{Q}$.
(avec possibilité d'enchaîner plusieurs assertions intermédiaires)

Exemple Soient $x, y \in \mathbb{R}$. Montrons

$$x^2 = y^2 \Rightarrow |x| = |y|$$

Supposons $x^2 = y^2$.
On a $x^2 - y^2 = (x-y)(x+y) = 0$ donc $x - y = 0$ ou $x + y = 0$.
Par suite $x = y$ ou $x = -y$ et donc $|x| = |y|$.

(2) Par contraposée : on établit non$(\mathcal{Q}) \Rightarrow$ non$(\mathcal{P})$.

Exemple Montrons

$$x \notin \mathbb{Q} \Rightarrow 1 + x \notin \mathbb{Q}$$

Par contraposée, montrons : $1 + x \in \mathbb{Q} \Rightarrow x \in \mathbb{Q}$.
Supposons $1 + x \in \mathbb{Q}$. Puisque $x = (1 + x) - 1$, on a $x \in \mathbb{Q}$.

1.3.3 Démonstration par récurrence

Théorème

Soit $n_0 \in \mathbb{N}$ et $\mathcal{P}(n)$ une assertion dépendant d'un entier $n \geqslant n_0$.
Si
1) $\mathcal{P}(n_0)$ est vraie ;
2) $\forall n \geqslant n_0, \mathcal{P}(n)$ vraie $\Rightarrow \mathcal{P}(n+1)$ vraie.
alors

$$\forall n \geqslant n_0, \mathcal{P}(n) \text{ est vraie}$$

Exemple Montrons

$$\forall n \in \mathbb{N}^\star, 1 + 2 + \cdots + n = \frac{n(n+1)}{2}$$

Procédons par récurrence sur $n \in \mathbb{N}^\star$.
Pour $n = 1$: $1 = \dfrac{1(1+1)}{2}$.
Supposons la propriété vraie au rang $n \geqslant 1$.

$$1 + 2 + \cdots + (n+1) = (1 + 2 + \cdots + n) + (n+1)$$

Par hypothèse de récurrence, on obtient

$$1 + 2 + \cdots + (n+1) = \frac{n(n+1)}{2} + (n+1) = \frac{(n+1)(n+2)}{2}$$

Récurrence établie.

Exemple Montrons

$$\forall n \in \mathbb{N}, 2^n \geqslant n$$

Procédons par récurrence sur $n \in \mathbb{N}$.
Pour $n = 0$: $2^0 = 1 \geqslant 0$.
Pour $n = 1$: $2^1 = 2 \geqslant 1$.
Supposons la propriété établie au rang $n \geqslant 1$.

$$2^{n+1} = 2 \times 2^n$$

Par hypothèse de récurrence, $2^n \geqslant n$, donc

$$2^{n+1} \geqslant 2n = n + n \geqslant n + 1$$

car $n \geqslant 1$.
Récurrence établie.
Noter qu'ici la récurrence est amorcée à partir du rang $n_0 = 1$ et l'étude de $\mathcal{P}(0)$ peut être considérées comme à part.

Chapitre 2

Théorie des ensembles

Les ensembles ont déjà été brièvement présentés, dans ce chapitre on reprend l'étude de ceux-ci de manière plus approfondie.
E, F, G, H désignent des ensembles.

2.1 Ensembles

2.1.1 Inclusion

Définition

On dit que E est inclus dans F, et on note $E \subset F$, si tout élément de E est aussi élément de F.
Ainsi

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

Exemple $\emptyset \subset E, E \subset E$

Remarque

$$E \not\subset F \Leftrightarrow \exists x \in E, x \notin F$$

Proposition

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

dém. :
Si $E = F$ alors il est immédiat que E est inclus dans F et F inclus dans E.
Inversement, si $E \subset F$ et $F \subset E$ alors E et F sont formés des mêmes éléments ce qui permet d'affirmer $E = F$.
□

Proposition

$E \subset F$ et $F \subset G \Rightarrow E \subset G$,

2.1.2 Sous ensemble

Définition

On appelle partie (ou sous-ensemble) d'un ensemble E tout ensemble A inclus dans E.
L'ensemble formé des parties de E est noté $\mathcal{P}(E)$.

Remarque $\mathcal{P}(E)$ est un ensemble d'ensembles, $A \in \mathcal{P}(E) \Leftrightarrow A \subset E$.

Exemple $\emptyset \in \mathcal{P}(E), E \in \mathcal{P}(E), \emptyset \subset \mathcal{P}(E)$ et en général $E \not\subset \mathcal{P}(E)$.

Exemple Pour $E = \{a\}, \mathcal{P}(E) = \{\emptyset, \{a\}\}$.
Pour $E = \{a, b\}, \mathcal{P}(E) = \{\emptyset, \{a\}, \{b\}, \{a, b\}\}$.
Pour $E = \{a, b, c\}, \mathcal{P}(E) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{b, c\}, \{c, a\}, \{a, b, c\}\}$.
Pour $E = \emptyset, \mathcal{P}(E) = \{\emptyset\}$.

2.1.3 Opérations dans $\mathcal{P}(E)$

Soient A, B, C trois parties d'un ensemble E.

2.1.3.1 Union et intersection

Définition

On appelle union de A et B l'ensemble noté $A \cup B$ formé des éléments de E qui appartiennent à A ou à B. Ainsi

$$A \cup B = \{x \in E / x \in A \text{ ou } x \in B\}$$

Définition

On appelle intersection de A et B l'ensemble noté $A \cap B$ formé des éléments de E qui appartiennent
à A et à B. Ainsi

$$A \cap B = \{x \in E / x \in A \text{ et } x \in B\}$$

Remarque On a toujours les inclusions :

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

Proposition

$A \cup A = A, A \cap A = A,$
$A \cup E = E, A \cap E = A,$
$A \cup \emptyset = A, A \cap \emptyset = \emptyset,$
$A \cup B = B \cup A, A \cap B = B \cap A,$
$A \cup (B \cup C) = (A \cup B) \cup C$ noté $A \cup B \cup C,$
$A \cap (B \cap C) = (A \cap B) \cap C$ noté $A \cap B \cap C,$
$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ et $A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$

dém. :
On établit, en raisonnant par équivalence, qu'un élément de E appartient à la partie du premier membre si, et seulement si, il appartient à la partie du second membre.
par équivalence.
□

Attention : Ecrire $A \cup B \cap C$ n'a pas de sens ; un parenthèsage s'impose !

Proposition

Si $A \subset C$ et $B \subset C$ alors $A \cup B \subset C$.
Si $C \subset A$ et $C \subset B$ alors $C \subset A \cap B$.

2.1.3.2 Complémentaire

Définition

On appelle complémentaire d'une partie A de E l'ensemble noté $\mathcal{C}_E A$ formé des éléments de E qui ne sont pas dans A. Ainsi

$$\mathcal{C}_E A = \{x \in E / x \notin A\}$$

Remarque Lorsqu'il n'y a pas d'ambiguïté sur l'ensemble E à l'intérieur duquel on travaille, il est fréquent de noter $\bar{A}$ au lieu de $\mathcal{C}_E A$.

Remarque A et $\mathcal{C}_E A$ sont des parties disjointes de E.

Exemple $\mathcal{C}_E E = \emptyset, \mathcal{C}_E \emptyset = E$.

Proposition

$\mathcal{C}_E(\mathcal{C}_E A) = A$,
$\mathcal{C}_E(A \cup B) = \mathcal{C}_E A \cap \mathcal{C}_E B$,
$\mathcal{C}_E(A \cap B) = \mathcal{C}_E A \cup \mathcal{C}_E B$,
$A \subset B \Leftrightarrow \mathcal{C}_E B \subset \mathcal{C}_E A$.

dém. :
On établit, en raisonnant par équivalence, qu'un élément de E appartient à la partie du premier membre si, et seulement si, il appartient à la partie du second membre.
□

2.1.3.3 Différence

Définition

On appelle ensemble A privé de B l'ensemble noté $A\backslash B$ (ou $A - B$) constitué des éléments de E qui sont dans A sans être dans B. Ainsi :

$$A\backslash B = \{x \in E/x \in A \text{ et } x \notin B\}$$

Remarque En général : $A\backslash B \neq B\backslash A$.

Proposition

$A\backslash B = A \cap \mathcal{C}_E(B)$.

Définition

On appelle différence symétrique de A et B l'ensemble noté $A\Delta B$ déterminé par $A\Delta B = (A\backslash B) \cup (B\backslash A)$.

Proposition

$A\Delta B = (A \cup B)\backslash(A \cap B)$.

dém. :

$x \in A$	$x \in B$	$x \in A\Delta B$	$x \in (A \cup B)\backslash(A \cap B)$
V	V	F	F
V	F	V	V
F	V	V	V
F	F	F	F

□

Proposition

$A\Delta A = \emptyset$, $A\Delta\emptyset = A$ et $A\Delta E = \mathcal{C}_E A$.
$A\Delta B = B\Delta A$ et $(A\Delta B)\Delta C = A\Delta(B\Delta C)$.

dém. :
Les premières propriétés sont immédiates.
La dernière peut s'obtenir par un tableau de vérité comme dans la démonstration précédente.
□

2.1.4 Familles

I désigne un ensemble.

2.1.4.1 Définition

Définition

On appelle famille d'éléments de E indexée sur I la donnée d'un élément de E pour chaque indice $i \in I$. Si cet élément de E est noté a_i, la famille correspondante est notée $(a_i)_{i\in I}$.
On note E^I l'ensemble des familles d'éléments de E indexées sur I.

Remarque Une famille s'interprète comme étant une liste d'éléments indexés par les éléments d'un ensemble I.

Remarque L'indice joue un rôle muet : $(a_i)_{i \in I} = (a_j)_{j \in I}$.

Définition

Soit J une partie de I.
$(a_i)_{i \in J}$ est appelée sous famille de $(a_i)_{i \in I}$.
$(a_i)_{i \in I}$ est appelée sur famille de $(a_i)_{i \in J}$.

2.1.4.2 Famille finie

Définition

Lorsque I est un ensemble fini, on dit que la famille est finie.
Lorsque $I = \{1, ..., n\}$ on note souvent $(a_i)_{1 \leqslant i \leqslant n}$ au lieu de $(a_i)_{i \in I}$.
Cette famille est alors usuellement confondue avec le n-uplet : $(a_1, ..., a_n)$.

Exemple $E = \{a, b, c, d\}$, $x_1 = a, x_2 = b, x_3 = a, x_4 = x$.
$(x_1, x_2, x_3, x_4) = (x_i)_{1 \leqslant i \leqslant 4}$ est une famille finie d'éléments de E.

2.1.4.3 Suite

Définition

Lorsque $I = \mathbb{N}$, la famille $(a_n)_{n \in \mathbb{N}}$ est appelée suite d'éléments de E.
On note $E^{\mathbb{N}}$ l'ensemble de ces suites.

Remarque Ce vocabulaire s'étant au cas où $I = \{n \in \mathbb{N} / n \geqslant n_0\}$ (avec $n_0 \in \mathbb{N}$ fixé) ou à $I = \mathbb{Z}$, la suite correspondante étant alors notée $(a_n)_{n \geqslant n_0}$ ou $(a_n)_{n \in \mathbb{Z}}$.

Exemple Posons $u_n = 2n + 1$ pour tout $n \in \mathbb{N}$. $(u_n)_{n \in \mathbb{N}}$ est une suite d'entiers.

Exemple Posons $I_n = [1/n, n]$ pour tout $n \in \mathbb{N}^\star$. $(I_n)_{n \geqslant 1}$ est une suite d'intervalles.

2.1.4.4 Famille de parties d'un ensemble

Définition

On appelle famille de parties d'un ensemble E, toute famille $(A_i)_{i \in I}$ formée d'éléments de $\mathcal{P}(E)$ i.e. telle que

$$\forall i \in I, A_i \subset E$$

Exemple Soit $E = \{a, b, c, d\}$.
Posons $A_1 = \{a, b\}$, $A_2 = \{b, c\}$ et $A_3 = \{b, c, d\}$.
$(A_i)_{1 \leqslant i \leqslant 3} = (A_1, A_2, A_3)$ est une famille de parties de E.

Définition

Soit $(A_i)_{i\in I}$ une famille de parties de E. On pose :
- $\bigcup_{i\in I} A_i = \{x \in E / \exists i \in I, x \in A_i\}$ appelée union de la famille $(A_i)_{i\in I}$;
- $\bigcap_{i\in I} A_i = \{x \in E / \forall i \in I, x \in A_i\}$ appelée intersection de la famille $(A_i)_{i\in I}$.

Exemple Avec l'exemple ci-dessus

$$\bigcup_{i=1}^{3} A_i = E \text{ et } \bigcap_{i=1}^{3} A_i = \{b\}$$

Exemple $\bigcup_{n\in\mathbb{N}^\star} [1/n, n] = \mathbb{R}^{+\star}$ et $\bigcap_{n\in\mathbb{N}^\star} \left[-\frac{1}{n}, \frac{1}{n}\right] = \{0\}$.

Remarque Si $I = \emptyset$ alors : $\bigcup_{i\in I} A_i = \emptyset$ et $\bigcap_{i\in I} A_i = E$.

Définition

Soit $(A_i)_{i\in I}$ une famille de parties de E.
On dit que $(A_i)_{i\in I}$ est un recouvrement de E si $\bigcup_{i\in I} A_i = E$.

Exemple $([-n, n])_{n\in\mathbb{N}}$ est un recouvrement de $\mathbb{R}$.

Définition

On dit que $(A_i)_{i\in I}$ est une partition de E si c'est un recouvrement formé de parties non vides deux à deux disjointes.

Exemple $([n, n+1[)_{n\in\mathbb{Z}}$ est une partition de $\mathbb{R}$.

2.2 Applications

2.2.1 Définition

Définition

On appelle graphe de E vers F toute partie Γ de $E \times F$.
E est appelé ensemble de départ et F ensemble d'arrivée du graphe Γ.

Exemple $E = \{a, b, c, d\}$, $F = \{1, 2, 3, 4\}$ et $\Gamma = \{(a, 1), (b, 2), (c, 2), (d, 4)\}$.
$\Gamma = \{(a, 1), (b, 2), (c, 2), (d, 4)\}$ est un graphe de E vers F.
$\Gamma' = \{(a, 1), (b, 2), (b, 3), (c, 3)\}$ est un graphe de E vers F.

Définition

On dit qu'un graphe Γ de E vers F est une application f de E vers F si

$$\forall x \in E, \exists! y \in F \text{ tel que } (x, y) \in \Gamma$$

Pour tout $x \in E$, l'unique $y \in F$ tel que $(x, y) \in \Gamma$ est appelé image de x par l'application f, on la note $f(x)$.
Pour tout $y \in F$, les $x \in E$, s'il en existe, tels que $y = f(x)$ sont appelés antécédents de y par l'application f.
On note $f : E \to F$ pour signifier que f est une application de E vers F.
On note $\mathcal{F}(E, F)$ l'ensemble des applications de E vers F.

Remarque On parle indifféremment d'application ou de fonction.

Exemple Reprenons l'exemple précédent.
Le graphe $\Gamma = \{(a, 1), (b, 2), (c, 3), (d, 2)\}$ est une application f pour laquelle :
$f(a) = 1, f(b) = 2, f(c) = 3$ et $f(d) = 2$.
Le graphe $\Gamma' = \{(a, 1), (b, 2), (b, 3), (c, 3)\}$ n'est pas une application pour deux raisons :
- b a deux images ;
- d n'a pas d'images.

Remarque Si $E = \emptyset$ alors tout graphe Γ de E vers F est vide.
Puisque la phrase quantifiée $\forall x \in \emptyset, \exists! y \in F, (x, y) \in \emptyset$ est vraie, on peut dire que $\emptyset$ est une application de $E = \emptyset$ vers F : celle-ci est appelée application vide.

Proposition

Soient $f, g : E \to F$. On a $f = g \Leftrightarrow \forall x \in E, f(x) = g(x)$.

dém. :
Par égalités des graphes définissant f et g.
□

Remarque Pour définir une application $f : E \to F$ il suffit de préciser comment à chaque élément x de E est associé son image $f(x)$ dans F.
C'est le principe de la syntaxe

$$f : \begin{cases} E \to F \\ x \mapsto ... \end{cases}$$

C'est désormais ainsi que nous définirons et manipulerons les applications.

Exemple $\mathrm{Id}_E : \begin{cases} E \to E \\ x \mapsto x \end{cases}$ est une application appelée identité de E.

Exemple Soit $C \in F$. $\tilde{C} : \begin{cases} E \to F \\ x \mapsto C \end{cases}$ est appelée application constante égale à C.

Remarque La problématique de bonne définition d'une fonction $f : E \to F$ définie par une association $x \mapsto f(x)$ est double :
- au départ : tout élément de E doit posséder une image et une seule ;
- à l'arrivée : cette image doit être dans F.

Exemple L'application

$$f : \begin{cases} \mathbb{R} \to \mathbb{R}^+ \\ x \mapsto \sqrt{x^2 + x + 1} \end{cases}$$

est bien définie car :
- $\forall x \in \mathbb{R}, x^2 + x + 1 \geqslant 0$ et donc $\sqrt{x^2 + x + 1}$ existe ;
- $\sqrt{x^2 + x + 1} \in \mathbb{R}^+$.

Exemple Soit A une partie de E.
L'application

$$f : \begin{cases} \mathcal{P}(E) \to \mathcal{P}(A) \\ X \mapsto X \cap A \end{cases}$$

est bien définie.

Exemple Soit A une partie de E.
Considérons

$$\chi_A : \begin{cases} E \to \{0, 1\} \\ x \mapsto \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{sinon} \end{cases} \end{cases}$$

χ_A est appelée application caractéristique de A.

Remarque Les familles d'éléments de E indexée sur I sont en fait des applications de I vers E. En effet, une famille $(a_i)_{i \in I}$ se comprend comme une application qui à $i \in I$ associe $a_i \in E$.
En particulier les suites d'éléments de E indexées sur $\mathbb{N}$, correspondent aux fonctions de $\mathbb{N}$ vers E.

2.2.2 Composition d'applications

Définition

Soient $f : E \to F$ et $g : F \to G$.
On appelle composée de f par g l'application $g \circ f : E \to G$ définie par :

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

Symboliquement :

$$E \xrightarrow{f} F \xrightarrow{g} G$$
$$\underset{g \circ f}{\longrightarrow}$$

Remarque $g \circ f$ est bien définie car l'ensemble d'arrivée de f coïncide avec l'ensemble de départ de g.

Remarque Lorsque $G = E$, on peut à la fois définir $f \circ g$ et $g \circ f$ mais en général : $f \circ g \neq g \circ f$.
D'ailleurs $g \circ f : E \to E$ et $f \circ g : F \to F$.

Exemple Soient

$$f : \begin{cases} \mathbb{R} \to \mathbb{R} \\ x \mapsto x^2 \end{cases} \text{ et } g : \begin{cases} \mathbb{R} \to \mathbb{R} \\ x \mapsto x + 1 \end{cases}$$

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

Proposition

Soient $f : E \to F, g : F \to G$ et $h : G \to H$.
On a

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

Cette application est encore notée $h \circ g \circ f$.

dém. :
Commençons par notons que les composées proposées sont effectivement possibles.
Ensuite, pour tout $x \in E$,

$$[(h \circ g) \circ f] (x) = (h \circ g)(f(x)) = h (g(f(x)))$$

et

$$[h \circ (g \circ f)] (x) = h (g \circ f(x)) = h (g(f(x)))$$

Puisque les applications $(h \circ g) \circ f$ et $h \circ (g \circ f)$ coïncidents pour chaque $x \in E$, elles sont égales.
□

Proposition

Soit $f : E \to F$.
On a $f \circ \mathrm{Id}_E = f$ et $\mathrm{Id}_F \circ f = f$.

dém. :
Les composées proposées ci-dessus sont effectivement possibles.
On vérifie ensuite l'égalités des applications étudiées en chaque élément de leur ensemble de départ.
□

2.2.3 Injection et surjection

2.2.3.1 Injection

Définition

On dit que $f : E \to F$ est injective si f ne prend jamais deux fois la même valeur i.e. :

$$\forall x, x' \in E, x \neq x' \Rightarrow f(x) \neq f(x')$$

Remarque La contraposée de $x \neq x' \Rightarrow f(x) \neq f(x')$ est $f(x) = f(x') \Rightarrow x = x'$.
Pour établir l'injectivité de f, on montre l'une ou l'autre de ces deux implications, en pratique, c'est plutôt la deuxième.

Exemple Id_E est injective.

Exemple L'application

$$f : \begin{cases} \mathbb{R} \to \mathbb{R} \\ x \mapsto 2\mathrm{e}^x + 1 \end{cases}$$

est injective.
Soient $x, x' \in \mathbb{R}$.
Supposons $f(x) = f(x')$.
On a $2\mathrm{e}^x + 1 = 2\mathrm{e}^{x'} + 1$ donc $\mathrm{e}^x = \mathrm{e}^{x'}$ puis, en composant avec le logarithme népérien, $x = x'$.
Ainsi f est injective.

Exemple Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction strictement croissante.
Montrons que f est injective.
Soient $x, x' \in \mathbb{R}$.
Supposons $x \neq x'$.
Quitte à échanger x et x', on peut supposer $x < x'$.
Par la stricte croissance de f, on a alors $f(x) < f(x')$ et donc $f(x) \neq f(x')$.
Ainsi f est injective.

Attention : L'implication d'injectivité n'est pas :

$$x = x' \Rightarrow f(x) = f(x')$$

Cette dernière implication est d'ailleurs vérifiée par toute application f !

Remarque Par négation

$$f : E \to F \text{ non injective } \Leftrightarrow \exists x, x' \in E, f(x) = f(x') \text{ et } x \neq x'$$

Exemple L'application

$$f : \begin{cases} \mathbb{R} \to \mathbb{R} \\ x \mapsto x^2 \end{cases}$$

n'est pas injective.
En effet $f(1) = f(-1)$ alors que $1 \neq -1$

Proposition

Si $f : E \to F$ et $g : F \to G$ sont injectives alors $g \circ f$ l'est aussi.

dém. :
Soient $x, x' \in E$.
Supposons $(g \circ f)(x) = (g \circ f)(x')$.
On a $g(f(x)) = g(f(x'))$.
Par l'injectivité de g, on obtient $f(x) = f(x')$.
Par l'injectivité de f, on obtient $x = x'$.
Finalement, $g \circ f$ est injective.
□

2.2.3.2 Surjection

Définition

Soit $f : E \to F$. On dit que f est surjective si chaque élément de F possède au moins un antécédent par f i.e. :

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

Remarque Pour justifier la surjectivité de $f : E \to F$, il suffit, pour chaque $y \in F$, de justifier l'existence d'une solution à l'équation $y = f(x)$.

Exemple Id_E est surjective.

Exemple L'application

$$s : \begin{cases} \mathbb{Z} \to \mathbb{N}^\star \\ n \mapsto |n| + 1 \end{cases}$$

est surjective.
En effet pour tout $m \in \mathbb{N}^\star$, $f(m-1) = |m-1| + 1 = m - 1 + 1 = m$.
Ainsi chaque $m \in \mathbb{N}^\star$ possède au moins un antécédent dans $\mathbb{Z}$.

Exemple L'application

$$f : \begin{cases} \mathbb{C} \to \mathbb{C} \\ z \mapsto z^2 \end{cases}$$

est surjective.

En effet, considérons $Z \in \mathbb{C}$.
On peut écrire $Z = r\mathrm{e}^{i\theta}$ avec $r \geqslant 0$ et $\theta \in \mathbb{R}$.
Pour $z = \sqrt{r}\mathrm{e}^{i\theta/2} \in \mathbb{C}$, on a $f(z) = z^2 = r\mathrm{e}^{i\theta} = Z$.
Ainsi, chaque $Z \in \mathbb{C}$ possède au moins un antécédent dans $\mathbb{C}$.

Attention : Ne pas confondre la surjectivité avec :

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

assertion qui est vraie pour toute application f.

Remarque Par négation

$$f : E \to F \text{ non surjective } \Leftrightarrow \exists y \in F, \forall x \in E, y \neq f(x)$$

Exemple L'application

$$f : \begin{cases} \mathbb{R} \to \mathbb{R} \\ x \mapsto x^2 \end{cases}$$

n'est pas surjective. En effet l'élément $-1 \in \mathbb{R}$ ne possède par d'antécédent par f.

Proposition

> Si $f : E \to F$ et $g : F \to G$ sont surjectives alors $g \circ f$ l'est aussi.

dém. :
Soit $z \in G$.
Par la surjectivité de g, il existe $y \in F$ tel que $g(y) = z$.
Par la surjectivité de f, il existe $x \in E$ tel que $f(x) = y$.
On a alors $(g \circ f)(x) = g(f(x)) = g(y) = z$.
Ainsi chaque élément de G possède au moins un antécédent par $g \circ f$, l'application $g \circ f$ est surjective.
□

2.2.4 Bijection

2.2.4.1 Définition

Définition

> Soit $f : E \to F$. On dit que f est bijective si chaque élément de F possède un unique antécédent par f dans E i.e. :
>
> $$\forall y \in F, \exists! x \in E, y = f(x)$$

Remarque Pour montrer que $f : E \to F$ est bijective il suffit d'établir que, pour chaque $y \in F$, l'équation $y = f(x)$ possède une unique solution $x \in E$.

Exemple L'application

$$f : \begin{cases} \mathbb{R}^+ \to [1, +\infty[\\ x \mapsto x^2 + 1 \end{cases}$$

est bijective.
En effet, soit $x \in \mathbb{R}^+$ et $y \in [1, +\infty[$.

$$y = f(x) \Leftrightarrow y = x^2 + 1 \Leftrightarrow x^2 = y - 1 \Leftrightarrow x = \sqrt{y - 1}$$

(car $x \geqslant 0$ et $y \geqslant 1$).
Par cette chaîne d'équivalence, on peut affirmer que l'équation $y = f(x)$ possède une solution unique $x \in \mathbb{R}^+$ pour chaque $y \in [1, +\infty[$; l'application f est donc bien bijective.

Attention : Ne pas confondre la bijectivité avec
$\forall x \in E, \exists! y \in F, y = f(x)$
propriété qui est vérifiée par toute fonction.

Proposition

> Soit $f : E \to F$. On a équivalence entre :
> (i) f est bijective ;
> (ii) f est injective et surjective.

dém. :
(i) $\Rightarrow$ (ii)
Si f est bijective alors f est clairement surjective. Etablissons l'injectivité :
Soient $x, x' \in E$. Supposons $f(x) = f(x')$.
Pour $y = f(x) = f(x') \in F$, x et x' sont des antécédents de y.
Or f étant bijective, y ne possède qu'un antécédent, donc $x = x'$.
(ii) $\Rightarrow$ (i)
Supposons f injective et surjective.
Soit $y \in F$.
Comme f est surjective, il existe $x \in E$ tel que $y = f(x)$
Soit de plus $x' \in E$ tel que $y = f(x')$.
On a alors $f(x) = f(x')$, or f est injective, donc $x = x'$.
Par suite il existe un unique $x \in E$ tel que $y = f(x)$. Ainsi f est bijective.
□

Exemple Id_E est bijective.

Exemple Soit $f : \mathbb{N} \to \mathbb{Z}$ définie par

$$f(n) = \begin{cases} n/2 & \text{si } n \text{ est pair} \\ -(n+1)/2 & \text{sinon} \end{cases}$$

Visualisons les premières valeurs de f :

n	0	1	2	3	4	5
$f(n)$	0	-1	1	-2	2	-3

La fonction f est bien définie.
En effet, quand n est pair et quand n est impair, les formules déterminant $f(n)$ donnent un entier.
De plus, on remarque que pour n pair, $f(n) \geqslant 0$ alors que pour n impair : $f(n) < 0$..
Etudions la bijectivité de f via injectivité et surjectivité.
Soient $n, n' \in \mathbb{N}$. Supposons $f(n) = f(n')$.
Puisque $f(n)$ et $f(n')$ sont égaux, ils ont le même signe et donc n et n' ont la même parité.
Si n et n' sont pairs alors l'égalité $f(n) = f(n')$ donne $n/2 = n'/2$ et donc $n = n'$.
Si n et n' sont impairs alors l'égalité $f(n) = f(n')$ donne $(n + 1)/2 = (n' + 1)/2$ et on retrouve encore $n = n'$.
Ainsi f est injective. Etudions maintenant la surjectivité.
Soit $m \in \mathbb{Z}$.
Si $m \geqslant 0$ alors, pour $n = 2m \in \mathbb{N}$, on a $f(n) = n/2 = m$ car n est pair.
Si $m < 0$ alors, pour $n = -(2m + 1) \in \mathbb{N}$, on a $f(n) = -(n + 1)/2 = m$ car n est impair.
Dans tous les cas, on observe que m possède au moins un antécédent par f.
Ainsi f est surjective et finalement f est bijective.
ont le

Proposition

Si $f : E \to F$ et $g : F \to G$ sont bijectives alors $g \circ f$ l'est aussi.

dém. :
Via injectivité et surjectivité.
□

2.2.4.2 Application réciproque

Proposition

Soient $f : E \to F$ et $g : F \to G$.
Si $g \circ f$ est injective alors f est injective.
Si $g \circ f$ est surjective alors g est surjective.

dém. :
Supposons $g \circ f$ injective.
Soient $x, x' \in E$. Supposons $f(x) = f(x')$.
On a alors $g(f(x)) = g(f(x'))$ i.e. $(g \circ f)(x) = (g \circ f)(x')$.
Par l'injectivité de $g \circ f$, on obtient $x = x'$. Ainsi f est injective.
Supposons maintenant $g \circ f$ surjective.
Soit $z \in G$. Par la surjectivité de $g \circ f$, il existe $x \in E$ tel que $z = (g \circ f)(x)$.
Pour $y = f(x) \in F$, on a $g(y) = g(f(x)) = (g \circ f)(x) = z$.
Ainsi chaque $z \in G$ possède au moins un antécédent $y \in F$ pour l'application g et on peut affirmer que g est surjective.
□

Théorème

Soit $f : E \to F$. On a équivalence entre :
(i) f est bijective ;
(ii) il existe une application $g : F \to E$ telle que $g \circ f = \mathrm{Id}_E$ et $f \circ g = \mathrm{Id}_F$.
De plus, si tel est le cas, l'application g ci-dessus est unique.
On l'appelle application réciproque de f et on la note f^{-1}.

dém. :
Les compositions proposées dans le (ii) sont possibles.
(ii) $\Rightarrow$ (i) Supposons (ii)
L'application $g \circ f = \mathrm{Id}_E$ est injective donc f est injective.
L'application $f \circ g = \mathrm{Id}_F$ est surjective donc f est surjective.
Ainsi f est bijective.
(i) $\Rightarrow$ (ii)
Supposons f bijective. On a

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

Posons alors $g(y) = x$, ce qui définit une application $g : F \to E$.

$$\forall x \in E, g(f(x)) = x$$

car si $y = f(x)$ alors $g(y) = x$.

$$\forall y \in F, f(g(y)) = y$$

car si $x = g(y)$ alors $f(x) = y$.
Par suite

$$g \circ f = \mathrm{Id}_E \text{ et } f \circ g = \mathrm{Id}_F$$

De plus :
Soit $h : F \to E$ une application telle que $h \circ f = \mathrm{Id}_E$ et $f \circ h = \mathrm{Id}_F$.
On a

$$h = h \circ \mathrm{Id}_F = h \circ (f \circ g) = (h \circ f) \circ g = \mathrm{Id}_E \circ g = g$$

Ainsi, il y a unicité de l'application g introduite en (ii).
□

Exemple L'application réciproque de Id_E est Id_E.
En effet $\mathrm{Id}_E \circ \mathrm{Id}_E = \mathrm{Id}_E$ et $\mathrm{Id}_E \circ \mathrm{Id}_E = \mathrm{Id}_E$.

Corollaire

Si $f : E \to F$ est bijective alors on peut introduire $f^{-1} : F \to E$ et on a : $f^{-1} \circ f = \mathrm{Id}_E$ et $f \circ f^{-1} = \mathrm{Id}_F$.

Corollaire

Soit $f : E \to F$. Si on détermine $g : F \to E$ telle que $g \circ f = \mathrm{Id}_E$ et $f \circ g = \mathrm{Id}_F$ alors on peut conclure : f bijective et $f^{-1} = g$.

Exemple Soit

$$s : \begin{cases} \mathbb{N} \to \mathbb{N}^\star \\ n \mapsto n+1 \end{cases}$$

Montrons s bijective et déterminons s^{-1}.
Considérons

$$p : \begin{cases} \mathbb{N}^\star \to \mathbb{N} \\ n \mapsto n-1 \end{cases}$$

L'application p est bien définie.
On vérifie aisément $s \circ p = \mathrm{Id}_{\mathbb{N}^\star}$ et $p \circ s = \mathrm{Id}_{\mathbb{N}}$; on peut donc conclure que s est bijective et $s^{-1} = p$.

Attention : Il faut observer deux composées égales à l'identité avant de conclure à la bijectivité.

Exemple Soient $f : \mathbb{R} \to \mathbb{R}^+$ et $g : \mathbb{R}^+ \to \mathbb{R}$ définies par $f(x) = x^2$ et $g(x) = \sqrt{x}$.
On a $g \circ f = \mathrm{Id}_{\mathbb{R}^+}$ alors que ni f, ni g, ne sont bijectives !

Proposition

Si $f : E \to F$ est bijective alors f^{-1} est bijective et

$$(f^{-1})^{-1} = f$$

dém. :
Supposons f bijective. On peut introduire $f^{-1} : F \to E$.
Comme $f \circ f^{-1} = \mathrm{Id}_F$ et $f^{-1} \circ f = \mathrm{Id}_E$, par le deuxième corollaire appliqué à f^{-1} (en prenant $g = f$) on peut conclure f^{-1} bijective et $(f^{-1})^{-1} = f$.
□

Proposition

Si $f : E \to F$ et $g : F \to G$ sont bijectives alors $g \circ f$ aussi

$$(g \circ f)^{-1} = f^{-1} \circ g^{-1}$$

dém. :
On sait déjà que $g \circ f$ est bijective.
Puisque $(g \circ f)^{-1} \circ (g \circ f) = \mathrm{Id}_E$ on a, en composant avec f^{-1}, $(g \circ f)^{-1} \circ g = f^{-1}$ puis, en composant avec g^{-1}, $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.
□

2.2.4.3 Permutation

Définition

On appelle permutation de E toute application bijective de E dans lui-même.
On note $\mathfrak{S}(E)$ l'ensemble des permutations de E.

Exemple Id_E est une permutation de E.

Proposition

$\forall f, g \in \mathfrak{S}(E), f \circ g \in \mathfrak{S}(E)$ et $g \circ f \in \mathfrak{S}(E)$.
$\forall f \in \mathfrak{S}(E), f^{-1} \in \mathfrak{S}(E)$.

Définition

On appelle involution de E toute application $f : E \to E$ telle que $f \circ f = \mathrm{Id}_E$.

Proposition

Soit $f : E \to E$. On a équivalence entre :
(i) f est une involution ;
(ii) f est bijective et $f^{-1} = f$.

Exemple Id_E est une involution.

Exemple L'application

$$f : \begin{cases} \mathbb{C} \to \mathbb{C} \\ z \mapsto \bar{z} \end{cases}$$

est une involution.

Exemple L'application

$$f : \begin{cases} \mathcal{P}(E) \to \mathcal{P}(E) \\ X \mapsto \mathcal{C}_E X \end{cases}$$

est une involution.

2.2.5 Image directe, image réciproque d'une partie.

2.2.5.1 Image directe

Définition

On appelle image directe de $A \in \mathcal{P}(E)$ par $f : E \to F$ l'ensemble noté $f(A)$ formé des valeurs prises par f sur A.
Ainsi

$$f(A) = \{f(x) \text{ avec } x \in A\} = \{f(x)/x \in A\}$$

Exemple Pour $A = \{x\}$, $f(A) = \{f(x)\}$.
Pour $A = \{x, y\}$, $f(A) = \{f(x), f(y)\}$.
Pour $A = \emptyset$, $f(A) = \emptyset$.

Exemple Soit $f : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = x^2$. $f([-1, 2]) = [0, 4]$, $f(\mathbb{R}) = \mathbb{R}^+$.

Remarque $f(A)$ peut aussi se voir comme étant formé des $y \in F$ qui possèdent au moins un antécédent dans A. Ainsi :

$$y \in f(A) \Leftrightarrow \exists x \in A, y = f(x)$$

Cette équivalence est fondamentale : elle caractérise l'appartenance à la partie $f(A)$.

Exemple Soient $A, B \subset E$. Montrons

$$f(A \cap B) \subset f(A) \cap f(B)$$

Soit $y \in f(A \cap B)$.
Il existe $x \in A \cap B$ tel que $y = f(x)$.
Puisque $x \in A$, $y = f(x) \in f(A)$. De même $y \in f(B)$ et donc $y \in f(A) \cap f(B)$.

Définition

Soit $f : E \to F$.
On appelle image de f l'ensemble noté $\mathrm{Im} f$ constitué des valeurs prises par f sur E.
Ainsi

$$\mathrm{Im} f = f(E) = \{f(x)/x \in E\}$$

Exemple Soit $f : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = x^2$. On a $\mathrm{Im} f = \mathbb{R}^+$.

Exemple Soit $f : \mathbb{C} \to \mathbb{C}$ définie par $f(z) = \mathrm{e}^z$. Déterminons $\mathrm{Im} f$.
On a pour tout $z \in \mathbb{C}$, $\mathrm{e}^z \neq 0$ donc $\mathrm{Im} f \subset \mathbb{C}^\star$.
Inversement, soit $Z \in \mathbb{C}^\star$.
On peut écrire $Z = \rho \mathrm{e}^{i\theta}$ avec $\rho = |Z|$ et $\theta \in \mathbb{R}$.
Pour $z = \ln(\rho) + i\theta \in \mathbb{C}$, on a $\mathrm{e}^z = Z$ donc $Z \in \mathrm{Im} f$. Ainsi $\mathbb{C}^\star \subset \mathrm{Im} f$.
Finalement

$$\mathrm{Im} f = \mathbb{C}^\star$$

Proposition

$f : E \to F$ est surjective si, et seulement si, $\mathrm{Im} f = F$.

dém. :
Les éléments de $\mathrm{Im} f$ sont ceux de F possédant au moins un antécédent par $f\dots$
□

2.2.5.2 Image réciproque

Définition

On appelle image réciproque de $B \in \mathcal{P}(F)$ par $f : E \to F$ l'ensemble noté $f^{-1}(B)$ formé des antécédents des éléments de B.
Ainsi

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

Attention : La notation $f^{-1}(B)$ ne signifie pas l'existence de l'application réciproque de f.

Exemple Considérons

$$f : \begin{cases} \mathbb{R} \to \mathbb{R} \\ x \mapsto x^2 \end{cases}$$

$f^{-1}([0,1]) = [-1,1]$, $f^{-1}(\mathbb{R}^+) = \mathbb{R}$,...

Exemple Pour $y \in F$, $f^{-1}(\{y\})$ correspond à l'ensemble des antécédents de y.
Ainsi

$$y \in \mathrm{Im} f \Leftrightarrow f^{-1}(\{y\}) \neq \emptyset$$

Exemple $f^{-1}(\emptyset) = \emptyset$, $f^{-1}(F) = E$, $f^{-1}(\mathrm{Im} f) = E$.

Remarque Par définition de $f^{-1}(B)$, on a, pour $x \in E$:

$$x \in f^{-1}(B) \Leftrightarrow f(x) \in B$$

Cette équivalence est fondamentale : elle caractérise l'appartenance à la partie $f^{-1}(B)$.

Exemple Soient $A, B \subset F$. Montrons que

$$f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B)$$

Soit $x \in E$.

$$x \in f^{-1}(A \cup B) \Leftrightarrow f(x) \in A \cup B \Leftrightarrow f(x) \in A \text{ ou } f(x) \in B$$

puis

$$x \in f^{-1}(A \cup B) \Leftrightarrow x \in f^{-1}(A) \text{ ou } x \in f^{-1}(B) \Leftrightarrow x \in f^{-1}(A) \cup f^{-1}(B)$$

On peut alors conclure

$$f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B)$$

2.2.6 Prolongement et restriction d'une application

Définition

Soient $E, \tilde{E}, F, \tilde{F}$ quatre ensembles tels que $E \subset \tilde{E}$ et $F \subset \tilde{F}$.
Considérons $f : E \to F$ et $\tilde{f} : \tilde{E} \to \tilde{F}$.
On dit que $\tilde{f}$ prolonge f si

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

Exemple Le module complexe est un prolongement de la valeur absolue.
L'exponentielle complexe est un prolongement de l'exponentielle réelle.

Définition

Soient $f : E \to F$, $A \subset E$ et $B \subset F$ vérifiant

$$\forall x \in A, f(x) \in B$$

On appelle restriction de f de A vers B l'application

$$g : \begin{cases} A \to B \\ x \mapsto g(x) = f(x) \end{cases}$$

Remarque La condition

$$\forall x \in A, f(x) \in B$$

(i.e. $f(A) \subset B$) assure la bonne définition de l'application g.

Exemple La restriction de la fonction sinus au départ de $[-\pi/2, \pi/2]$ et à valeurs dans $[-1, 1]$ est bijective.

Exemple La restriction d'une application injective est injective.

Remarque Soient $f : E \to F$ et $A \subset E$.
L'application restreinte

$$\begin{cases} A \to F \\ x \mapsto f(x) \end{cases}$$

est appelée restriction de f à A (au départ) et est notée $f_{\restriction A}$.
Soient $f : E \to F$ et $B \subset F$ telle que $\text{Im}f \subset B$.
L'application restreinte

$$\begin{cases} E \to B \\ x \mapsto f(x) \end{cases}$$

est appelée restriction de f à l'arrivée dans B.
Généralement on la note encore f.

Exemple La restriction d'une application à l'arrivée dans $\text{Im}f$ est surjective.

Exemple La restriction d'une application injective à l'arrivée dans $\text{Im}f$ est bijective.

2.3 Les ensembles finis

2.3.1 Equipotence d'ensembles

Définition

On dit qu'un ensemble E est équipotent à un ensemble F s'il existe une bijection de E vers F.
On note alors $E \approx F$.

Proposition

$E \approx E$,
$E \approx F \Rightarrow F \approx E$,
$E \approx F$ et $F \approx G \Rightarrow E \approx G$.

dém. :
Puisque Id_E est une bijection de E vers E, $E \approx E$.
Puisque l'application réciproque d'une bijection de E vers F est une bijection de F vers E,

$$E \approx F \Rightarrow F \approx E$$

Puisque la composée d'une bijection de E vers F par une bijection de F vers G est un bijection de E vers G,

$$E \approx F \text{ et } F \approx G \Rightarrow E \approx G$$

□

Exemple $\{a, b, c\}$ et $\{1, 2, 3\}$ sont équipotents via la bijection f définie par $f(a) = 1$, $f(b) = 2$ et $f(c) = 3$.

Exemple $\mathbb{N}$ et $\mathbb{N}^\star$ sont équipotents via la bijection $s : n \mapsto n + 1$.

Exemple $\mathbb{N}$ et $\mathbb{Z}$ sont équipotents via la bijection

$$n \mapsto \begin{cases} n/2 & \text{si } n \text{ pair} \\ -(n+1)/2 & \text{sinon} \end{cases}$$

Exemple On peut montrer que $\mathbb{N}$ et $\mathbb{N}^2$ sont équipotents.
On peut montrer que $\mathbb{N}$ et $\mathbb{Q}$ sont équipotents.
On peut montrer que $\mathbb{N}$ et $\mathbb{R}$ ne le sont pas.

Définition

Un ensemble est dit dénombrable s'il est équipotent à $\mathbb{N}$.

Exemple $\mathbb{N}$, $\mathbb{N}^\star$, $\mathbb{Z}$, $\mathbb{N}^2$ et $\mathbb{Q}$ sont dénombrables, $\mathbb{R}$ ne l'est pas.

2.3.2 Cardinal d'un ensemble

Définition

Pour $n \in \mathbb{N}^\star$, on note

$$\mathbb{N}_n = [\![1, n]\!] = \{1, 2, ..., n\} \text{ et } \mathbb{N}_0 = \emptyset$$

Théorème

Soient $n, p \in \mathbb{N}$.
S'il existe une injection de $\mathbb{N}_p$ dans $\mathbb{N}_n$ alors $p \leqslant n$.
S'il existe une surjection de $\mathbb{N}_p$ sur $\mathbb{N}_n$ alors $p \geqslant n$.
S'il existe une bijection de $\mathbb{N}_p$ vers $\mathbb{N}_n$ alors $p = n$.

dém. :
La propriété relative à l'existence de bijection de $\mathbb{N}_p$ vers $\mathbb{N}_n$ découle immédiatement des deux précédentes, il ne reste qu'à établir celles-ci...
Par récurrence sur $p \in \mathbb{N}$, montrons que pour tout $n \in \mathbb{N}$, s'il existe une injection de $\mathbb{N}_p$ dans $\mathbb{N}_n$ alors $p \leqslant n$.

Pour $p = 0$, la propriété $p \leqslant n$ est vérifiée.
Supposons la propriété établie au rang $p \geqslant 0$.
Supposons qu'il existe une injection f de $\mathbb{N}_{p+1}$ dans $\mathbb{N}_n$ pour un certain $n \in \mathbb{N}$.
A partir de celle-ci construisons une injection g de $\mathbb{N}_{p+1}$ dans $\mathbb{N}_n$ vérifiant $g(p+1) = n$.
Si $f(p+1) = n$ alors $g = f$ convient.
Si $f(p+1) \neq n$ introduisons l'application τ de $\mathbb{N}_n$ dans lui-même qui a pour seul effet d'échanger les éléments n et $f(p+1)$. L'application τ est clairement bijective. Considérons ensuite $g = \tau \circ f$. Par composition d'applications injectives, g est injective et par construction $g(p+1) = \tau(f(p+1)) = n$.
Nous sommes donc parvenu à construire une application g comme voulue. Exploitons celle-ci.
Considérons l'application $h : \mathbb{N}_p \to \mathbb{N}_{n-1}$ définie par $h(x) = g(x)$ pour tout $x \in \mathbb{N}_p$.
Comme g est injective, n ne peut avoir d'autres antécédents que $p+1$ par g et l'application h ci-dessus est bien définie à valeurs dans $\mathbb{N}_{n-1}$.
De plus g étant injective, $h : \mathbb{N}_p \to \mathbb{N}_{n-1}$, qui est une restriction de g, l'est aussi.
On peut alors appliquer l'hypothèse de récurrence pour conclure $p \leqslant n-1$ i.e. $p+1 \leqslant n$.
Récurrence établie.
Etudions maintenant les surjections de $\mathbb{N}_p$ sur $\mathbb{N}_n$
Supposons qu'il existe une surjection f de $\mathbb{N}_p$ sur $\mathbb{N}_n$. A partir de celle-ci, nous allons construire une injection de $\mathbb{N}_n$ dans $\mathbb{N}_p$ ce qui permettra de conclure.
Pour chaque $y \in \mathbb{N}_n$, on peut dire que l'ensemble $f^{-1}(\{y\})$ est non vide car f est surjective.
Considérons alors x_y un élément quelconque de cet ensemble (par exemple $x_y = \min(f^{-1}(\{y\}))$).
Considérons ensuite l'application $g : \mathbb{N}_n \to \mathbb{N}_p$ définie par $g(y) = x_y$ pour tout $y \in \mathbb{N}_n$ et montrons qu'elle est injective.
Pour tout $y \in \mathbb{N}_n$, $f(g(y)) = f(x_y) = y$ car x_y est un antécédent de y ; ainsi $f \circ g = \mathrm{Id}_{\mathbb{N}_n}$. L'application $f \circ g$ étant injective, l'application g l'est aussi. On peut donc conclure à l'existence d'une injection de $\mathbb{N}_n$ dans $\mathbb{N}_p$ et par l'étude qui précède on a $n \leqslant p$.
$\square$

Définition

> On dit qu'un ensemble E est fini s'il existe $n \in \mathbb{N}$ tel que i $E \approx \mathbb{N}_n$.
> En vertu du théorème ci-dessus, cet entier n est unique, on l'appelle cardinal de E et on le note $\mathrm{Card}\, E$ (ou $|E|$, $\#E$)
> Lorsqu'un ensemble E n'est pas fini, on dit qu'il est infini et on pose $\mathrm{Card}\, E = +\infty$.

Exemple $\mathrm{Card}\mathbb{N}_n = n$,
$\mathrm{Card}\mathbb{N} = +\infty$,
$\mathrm{Card}\ \{0, 1, ..., n\} = n + 1$,
Pour $a \leqslant b \in \mathbb{Z}$, $\mathrm{Card}\ [\![a, b]\!] = b - a + 1$.

Exemple $\mathrm{Card}\emptyset = 0$,
$\mathrm{Card}\,\{a\} = 1$ et

$$\mathrm{Card}\,\{a, b\} = \begin{cases} 1 & \text{si } a = b \\ 2 & \text{sinon} \end{cases}.$$

Remarque Soit E un ensemble fini.
Si $\mathrm{Card}\, E = 0$ alors $E = \emptyset$.
Si $\mathrm{Card}E = n \in \mathbb{N}^\star$ alors il existe une bijection $\varphi : \mathbb{N}_n \to E$.
Pour tout $1 \leqslant i \leqslant n$, posons $x_i = \varphi(i)$.
Comme φ est injective, les $x_1, ..., x_n$ sont deux à deux distincts.

Comme φ est surjective, $E = \{x_1, ..., x_n\}$.
Ainsi, lorsque E est un ensemble fini à $n \in \mathbb{N}^\star$ éléments, on peut indexer ceux-ci de sorte d'écrire

$$E = \{x_1, ..., x_n\}$$

avec des x_i deux à deux distincts.

2.3.3 Cardinal d'une réunion

Proposition

Soient A et B deux ensembles disjoints.
Si A et B sont finis alors $A \cup B$ l'est aussi et

$$\text{Card}(A \cup B) = \text{Card}A + \text{Card}B$$

dém. :
Posons $n = \text{Card}A$ et $p = \text{Card}B$.
Si $n = 0$ ou $p = 0$: ok
Sinon écrivons $A = \{x_1, ..., x_n\}$ avec des x_i deux à deux distincts et $B = \{y_1, ..., y_p\}$ avec des y_j deux à deux distincts.
Comme A et B sont supposés disjoints, les x_i sont distincts des y_j.
Considérons alors $\varphi : \mathbb{N}_{n+p} \to A \cup B$ définie par :

i	1	2	...	n	$n+1$	...	$n+p$
$\varphi(i)$	x_1	x_2	...	x_n	y_1	...	y_p

Il est immédiate d'observer que φ est une bijection de $\mathbb{N}_{n+p}$ vers $A \cup B$ ce qui permet de conclure.
□

Corollaire

Soit $(A_i)_{1 \leqslant i \leqslant n}$ une famille finie d'ensemble deux à deux disjoints.
Si tous les A_i sont des ensembles finis alors $\bigcup_{i=1}^{n} A_i$ l'est aussi et

$$\text{Card} \bigcup_{i=1}^{n} A_i = \sum_{i=1}^{n} \text{Card}A_i$$

Théorème

Toute partie d'un ensemble fini est elle-même finie.

dém. :
Par récurrence sur $n \in \mathbb{N}$, montrons que toute partie d'un ensemble à n éléments est finie.
Pour $n = 0$: ok
Supposons la propriété établie au rang $n \geqslant 0$.
Soit E un ensemble fini a $n + 1$ éléments.

$$E = \{x_1, ..., x_n, x_{n+1}\}$$

avec des x_i deux à deux distincts.
Posons

$$E' = \{x_1, ..., x_n\}$$

On a Card $E' = n$.
Soit A une partie de E.
Si $x_{n+1} \notin A$ alors A est une partie de E' et elle donc finie par hypothèse de récurrence.
Si $x_{n+1} \in A$. Posons $A' = A \backslash \{x_{n+1}\}$. A' est une partie de E', donc par hypothèse de récurrence A' est finie et puisque $A = A' \cup \{x_{n+1}\}$, A l'est aussi car réunion de deux ensembles finis disjoints.
Récurrence établie
□

Corollaire

Soit A une partie d'un ensemble fini E.

$$\mathrm{Card} C_E A = \mathrm{Card} E - \mathrm{Card} A$$

$\mathrm{Card} A \leqslant \mathrm{Card} E$ avec égalité si, et seulement si, $A = E$

dém. :
E est la réunion des deux parties A et $C_E A$ qui sont disjointes et finies.
On en déduit

$$\mathrm{Card} A + \mathrm{Card} C_E A = \mathrm{Card} E$$

Puisque $\mathrm{Card} C_E A \geqslant 0$, on obtient $\mathrm{Card} A \leqslant \mathrm{Card} E$ avec égalité si, et seulement si, $\mathrm{Card} C_E A = 0$ i.e. $C_E A = \emptyset$ ce qui correspond au cas $A = E$.
□

Remarque Il est fréquent d'établir l'égalité de deux ensembles pas une inclusion et une égalité de cardinaux (finis).

Théorème

Soient A et B deux ensembles.
Si A et B sont finis alors $A \cup B$ l'est aussi et

$$\mathrm{Card}(A \cup B) = \mathrm{Card} A + \mathrm{Card} B - \mathrm{Card}(A \cap B)$$

dém. :
On peut écrire $A \cup B = A \cup (B \backslash A)$ avec les ensembles A et $B \backslash A$ finis et disjoints.
On en déduit que $A \cup B$ est un ensemble fini et

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

On peut aussi écrire $B = (B \backslash A) \cup (A \cap B)$ avec les ensembles $B \backslash A$ et $A \cap B$ finis et disjoints.
On en déduit

$$\mathrm{Card} B = \mathrm{Card}(B \backslash A) + \mathrm{Card}(A \cap B)$$

puis la relation proposée.
□

Corollaire

Soit $(A_i)_{i\in I}$ une famille finie d'ensembles.

Si tous les A_i sont des ensembles finis est fini alors $\bigcup_{i=1}^{n} A_i$ l'est aussi et

$$\text{Card}\bigcup_{i=1}^{n} A_i \leqslant \sum_{i=1}^{n} \text{Card}A_i$$

2.3.4 Applications entre ensembles finis

Théorème

Soient E et F deux ensembles finis.
S'il existe une injection de E dans F alors Card$E \leqslant$ Card$F$.
S'il existe une surjection de E sur F alors Card$E \geqslant$ Card$F$.
S'il existe une bijection de E vers F alors Card$E =$ Card$F$.

dém. :
Posons $p =$ Card$E$ et $n =$ Card$F$.
Il existe $\varphi : \mathbb{N}_p \to E$ et $\psi : \mathbb{N}_n \to F$ bijectives.
S'il existe une injection f de E vers F alors $(\psi^{-1} \circ f \circ \varphi) : \mathbb{N}_p \to \mathbb{N}_n$ est injective et donc $p \leqslant n$.
S'il existe une surjection f de E vers F alors $(\psi^{-1} \circ f \circ \varphi) : \mathbb{N}_p \to \mathbb{N}_n$ est surjective et donc $p \geqslant n$.
S'il existe une bijection f de E vers F alors $(\psi^{-1} \circ f \circ \varphi) : \mathbb{N}_p \to \mathbb{N}_n$ est bijective et donc $p = n$.
□

Proposition

Soient E et F deux ensembles et $f : E \to F$.
Si A est une partie finie de E alors $f(A)$ est une partie finie de F et

$$\text{Card}f(A) \leqslant \text{Card}A$$

De plus Card$f(A) =$ Card$A$ si f est injective.

dém. :
Posons $n =$ Card$A$.
Si $n = 0$ alors $A = \emptyset$ et $f(A) = \emptyset$: ok
Si $n \neq 0$, écrivons $A = \{x_1, ..., x_n\}$ avec des x_i deux à deux distincts.
On a alors $f(A) = \{f(x_1), ..., f(x_n)\}$ et donc $f(A)$ est un ensemble fini avec Card$f(A) \leqslant n$.
Si de plus f est injective, les $f(x_i)$ sont deux à deux distincts et donc Card$f(A) = n$.
□

Théorème

Soient E et F deux ensembles finis tels que

$$\text{Card}E = \text{Card}F$$

Toute application injective de E dans F est bijective.
Toute application surjective de E sur F est bijective.

dém. :
Soit $f : E \to F$ une application injective.

On a $f(E) \subset F$ et Card$f(E) = $ Card$E = $ Card$F$ donc $f(E) = F$.
Par suite f est surjective puis bijective.
Soit $f : E \to F$ une application surjective.
Par l'absurde, si f n'est pas injective alors il existe $x, x' \in E$ tels que $x \neq x'$ et $f(x) = f(x')$.
On a alors $f(E \backslash \{x\}) = f(E) = F$ car f est surjective et $f(x') = f(x)$.
Or Card$f(E \backslash \{x\}) <$ Card$E$ d'où Card$F <$ Card$E$. C'est absurde.
□

Corollaire

Soit E un ensemble fini et $f : E \to E$.
On a équivalence entre :
(i) f est bijective ;
(ii) f est injective ;
(iii) f est surjective.

Attention : Ces résultats ne valent que pour les ensembles finis.

2.4 Dénombrement

2.4.1 Principe des bergers

Théorème

Soient E un ensemble, F un ensemble fini et $\varphi : E \to F$.
S'il existe $p \in \mathbb{N}$ tel que chaque $y \in F$ possède exactement p antécédents par φ alors l'ensemble E est fini et

$$\text{Card}E = p.\text{Card}F$$

dém. :
Si $F = \emptyset$ alors il n'y a qu'une seule application à valeurs dans F, c'est l'application vide qui est au départ de $E = \emptyset$. La propriété est vraie.
Si $F \neq \emptyset$, posons $n = \text{Card}\,F$.
On peut écrire $F = \{y_1, ..., y_n\}$ avec les y_i deux à deux distincts.
Posons, pour tout $1 \leqslant i \leqslant n$, $A_i = \varphi^{-1}(\{y_i\})$.
Par hypothèse, chaque partie A_i est finie et Card$A_i = p$.
Puisque les parties A_i sont deux à deux disjointes et que $\bigcup_{i=1}^{n} A_i = E$, on peut affirmer que E est finie et

$$\text{Card}E = \sum_{i=1}^{n} \text{Card}A_i = np$$

□

2.4.2 Produit cartésien

Théorème

Si E et F sont deux ensembles finis alors $E \times F$ l'est aussi et

$$\text{Card}E \times F = \text{Card}E \times \text{Card}F$$

dém. :
Considérons $\varphi : E \times F \to F$ définie par $\varphi(x, y) = y$.
Tout élément y de F possède exactement Card$E$ antécédents par φ et le principe de bergers appliqué à φ permet de conclure.
□

Corollaire

Soient $E_1, ..., E_n$ une liste d'ensembles finis.

$$\prod_{i=1}^{n} E_i = E_1 \times ... \times E_n$$

est fini et

$$\text{Card} \prod_{i=1}^{n} E_i = \prod_{i=1}^{n} \text{Card}E_i$$

dém. :
Par récurrence, en observant que $\prod_{i=1}^{n+1} E_i$ n'est pas très différent de $\prod_{i=1}^{n} E_i \times E_{n+1} \ldots$
□

Exemple Si E est un ensemble fini et $n \in \mathbb{N}^\star$ alors E^n est fini et Card$E^n = (\text{Card}E)^n$.

2.4.3 Dénombrement

Pour dénombrer le nombre d'objets construits par une démarche :
- on multiple lorsque passe d'une étape à l'étape suivante dans la construction ;
- on somme lorsqu'il y a une alternative strice dans la construction.

Exemple Combien y a-t-il de couples $(x, y) \in \{-2, -1, 0, 1, 2\}^2$ tels que $xy \geqslant 0$:
Couples solutions avec $x = 0$:
5 choix de y et autant de possibilités.
Couples solusions avec $x > 0$:
2 choix de x et 3 choix de y, soit 6 possibilités.
Couples solutions avec $x < 0$:
2 choix de x et 3 choix de y, soit 6 possibilités.
Au total : $5 + 6 + 6 = 17$ possibilités.

2.4.4 Ensembles d'applications

Théorème

Si E et F sont deux ensembles finis alors $\mathcal{F}(E, F)$ est fini et

$$\text{Card}\mathcal{F}(E, F) = (\text{Card}F)^{\text{Card}E}$$

dém. :
Si $E = \emptyset$ alors il n'y a qu'une seule application au départ de E, l'application vide.
Or $(\text{Card}F)^0 = 1$ donc la relation proposée est exacte.
Si $E \neq \emptyset$ alors, en posant $n = \text{Card}E$, on peut écrire $E = \{x_1, \dots, x_n\}$ avec les x_i deux à deux distincts.
Considérons l'application $\varphi : \mathcal{F}(E, F) \to F^n$ définie par $\varphi(f) = (f(x_1), \dots, f(x_n))$.
On vérifie aisément que l'application φ est bijective et on en déduit que $\mathcal{F}(E, F)$ est fini et

$$\text{Card}(\mathcal{F}(E, F)) = \text{Card}F^n = (\text{Card}F)^{\text{Card}E}$$

□

Remarque Démonstration plus simple :
Pour construire une application $f : E \to F$ avec $E = \{x_1, \dots, x_n\}$ (et les $x_1, \dots, x_n$ deux à deux distincts) :
- on choisit $f(x_1)$ dans F : Card$F$ possibilités ;
- on choisit $f(x_2)$ dans F : Card$F$ possibilités ;
...
- on choisit $f(x_n)$ dans F : Card$F$ possibilités.

Au total, il y a $(\text{Card}F)^n$ possibilités et autant d'applications de E vers F.

2.4.5 Ensemble de parties

Théorème

Si E est un ensemble fini alors $\mathcal{P}(E)$ l'est aussi et

$$\text{Card}\mathcal{P}(E) = 2^{\text{Card}E}$$

dém. :
Soit E un ensemble fini à n éléments.

Pour A partie de E, considérons son application caractéristique

$$\chi_A : \begin{cases} E \to \{0, 1\} \\ x \mapsto \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{sinon} \end{cases} \end{cases}$$

L'application $A \mapsto \chi_A$ réalise une bijection de $\mathcal{P}(E)$ vers $\{0, 1\}^E$.
Il y a 2^n applications de E dans $\{0, 1\}$ d'où le résultat.
□

Remarque Démonstration par dénombrement
Pour construire une partie A d'un ensemble $E = \{x_1, \dots, x_n\}$ avec $x_1, \dots, x_n$ deux à deux distintcs :
- on choisit si $x_1 \in A$ ou non : 2 possibilités ;
- on choisit si $x_2 \in A$ ou non : 2 possibilités ;
...
- on choisit si $x_n \in A$ ou non : 2 possibilités.
Au total, il y a 2^n possibilités de construction et autant parties de E.

2.4.6 Permutation

Théorème

Il y a exactement $n!$ bijections entre deux ensembles finis à n éléments.

dém. :
Par récurrence sur $n \in \mathbb{N}$.
Pour $n = 0$: $0! = 1$
Il n'existe qu'une application de $\emptyset$ vers $\emptyset$, c'est l'application vide, qui est bijective.
Supposons la propriété établie au rang $n \geqslant 0$.
Soit E et F deux ensembles $n + 1$ éléments.
Notons $\mathrm{Bij}(E, F)$ l'ensemble des bijections de E vers F.
Soit $a \in E$ et $\varphi : \mathrm{Bij}(E, F) \to F$ définie par $\varphi(f) = f(a)$.
Tout $b \in F$, les antécédents de b par φ correspondent aux applications bijectives de $E \backslash \{a\}$ vers $F \backslash \{b\}$.
Par hypothèse de récurrence, il y en a $n!$.
Par le principe des bergers,

$$\mathrm{Card}(\mathrm{Bij}(E, F)) = (n + 1)!$$

Récurrence établie.

□

Remarque Démonstration plus simple :
Pour construire une permutation de $E = \{x_1, \dots, x_n\}$ avec $x_1, \dots, x_n$ deux à deux distincts vers F à n éléments :
- on choisit $f(x_1)$ dans F : n possibilités ;
- on choisit $f(x_2)$ dans $F \backslash \{f(x_1)\}$: $n - 1$ possibilités ;
...
- on choisit $f(x_n)$ dans $F \backslash \{f(x_1), \dots, f(x_{n-1})\}$: 1 possibilité.
Au final, il y a $n!$ possibilités de construction et autant de bijection de E vers F.

Corollaire

Si E est un ensemble fini alors $\mathfrak{S}(E)$ est fini et $\mathrm{Card}\mathfrak{S}(E) = (\mathrm{Card}E)!$.

2.4.7 Coefficients combinatoires

2.4.7.1 Définition

Définition

Soient $n \in \mathbb{N}$ et $p \in \mathbb{Z}$. On appelle coefficient combinatoire p parmi n le nombre

$$\binom{n}{p} = \begin{cases} \dfrac{n!}{p!(n-p)!} & \text{si } 0 \leqslant p \leqslant n \\ 0 & \text{sinon} \end{cases}$$

Exemple $\binom{n}{0} = 1$, $\binom{n}{1} = n$, $\binom{n}{2} = \dfrac{n(n-1)}{2}$, $\binom{n}{3} = \dfrac{n(n-1)(n-2)}{3 \times 2}$,..., $\binom{n}{n} = 1$.

Proposition

$\forall n \in \mathbb{N}, \forall p \in \mathbb{Z}, \binom{n}{p} = \binom{n}{n-p}$.

dém. :
Les cas $p < 0$ ou $p > n$ sont immédiats.
Pour $0 \leqslant p \leqslant n$,

$$\binom{n}{n-p} = \frac{n!}{(n-p)!(n-(n-p))!} = \frac{n!}{(n-p)!p!} = \binom{n}{p}$$

□

Théorème

$\forall n \in \mathbb{N}, \forall p \in \mathbb{Z}, \binom{n}{p} + \binom{n}{p+1} = \binom{n+1}{p+1}$.

dém. :
Cas $0 \leqslant p \leqslant n-1$

$$\binom{n}{p} + \binom{n}{p+1} = \frac{n!}{p!(n-p)!} + \frac{n!}{(p+1)!(n-p-1)}$$
$$= \frac{n!(p+1+n-p)}{(p+1)!(n-p)!} = \frac{(n+1)!}{(p+1)!(n-p)!} = \binom{n+1}{p+1}$$

Cas $p = n$

$$\binom{n}{n} + \binom{n}{n+1} = 1 + 0 = 1 = \binom{n+1}{n+1}$$

Cas $p = -1$

$$\binom{n}{-1} + \binom{n}{0} = 0 + 1 = 1 = \binom{n+1}{0}$$

Cas $p > n$ ou $p < -1$
$$\binom{n}{p} + \binom{n}{p+1} = 0 + 0 = 0 = \binom{n+1}{p+1}$$

□

Remarque On peut facilement calculer les premiers coefficients combinatoires grâce au triangle de Pascal :

$n\backslash p$	0	1	2	3	4
0	1	0	0	0	0
1	1	1	0	0	0
2	1	2	1	0	0
3	1	3	3	1	0
4	1	4	6	4	1

où l'on visualise :
$$\begin{array}{ccc} \binom{n}{p} & + & \binom{n}{p+1} \\ & & \| \\ & & \binom{n+1}{p+1} \end{array}$$

Exemple Soient $p, n \in \mathbb{N}$, calculons
$$\sum_{k=0}^{n} \binom{p+k}{k}$$

On a
$$\sum_{k=0}^{n} \binom{p+k}{k} = \binom{p}{0} + \binom{p+1}{1} + \binom{p+2}{2} + \cdots + \binom{p+n}{n}$$

Or
$$\binom{p}{0} + \binom{p+1}{1} = \binom{p+1}{0} + \binom{p+1}{1} = \binom{p+2}{1}$$

puis
$$\binom{p+2}{1} + \binom{p+2}{2} = \binom{p+3}{2}, \ldots$$

et enfin
$$\binom{p+n}{n-1} + \binom{p+n}{n} = \binom{p+n+1}{n}$$

Proposition

$\forall n \in \mathbb{N}^\star, \forall p \in \mathbb{Z}, p\binom{n}{p} = n\binom{n-1}{p-1}.$

dém. :
Cas $1 \leqslant p \leqslant n$
$$p\binom{n}{p} = p\frac{n!}{p!(n-p)!} = n\frac{(n-1)!}{(p-1)!(n-p)!} = n\binom{n-1}{p-1}$$
Cas $p = 0$
$$0\binom{n}{0} = 0 = n \times 0 = n\binom{n-1}{-1}$$
Cas $p < 0$ ou $p > n$
$$p\binom{n}{p} = p \times 0 = n \times 0 = n\binom{n-1}{p-1}$$
□

Remarque On retient
$$\binom{n}{p} = \frac{n}{p}\binom{n-1}{p-1}$$

2.4.7.2 Nombre de combinaisons

Définition

Soit E un ensemble et $p \in \mathbb{N}$.
On appelle combinaison de p éléments de E toute partie de E à p éléments.

Théorème

Soit E un ensemble fini à $n \in \mathbb{N}$ éléments et $p \in \mathbb{Z}$.
Il y a exactement $\binom{n}{p}$ combinaisons possibles de p éléments de E.
Autrement dit : il y a exactement $\binom{n}{p}$ parties à p éléments dans un ensemble à n éléments.

dém. :
Par récurrence sur $n \in \mathbb{N}$.
Pour $n = 0$: l'ensemble vide ne possède qu'une partie qui est l'ensemble vide.
Ceci est cohérent avec $\binom{0}{0} = 1$ et $\binom{0}{p} = 0$ pour $p \neq 0$.
Supposons la propriété établie au rang $n \geqslant 0$.
Soit E un ensemble fini à $n+1$ éléments et $p \in \{0, 1, \ldots, n+1\}$.
Si $p < 0$ ou $p > n$, il n'existe pas de parties de E à p éléments et $\binom{n}{p} = 0$.
Si $p = 0$, il n'existe qu'une partie de E à 0 élément (c'est l'ensemble vide) et $\binom{n}{0} = 1$.

Sinon, considérons $a \in E$ fixé.
Il y a $\binom{n}{p}$ parties à p éléments de E ne contenant pas a et $\binom{n}{p-1}$ parties contenant a donc il y a

$$\binom{n}{p} + \binom{n}{p-1} = \binom{n+1}{p}$$

parties à p éléments dans E.
Récurrence établie
□

Remarque Il y a dans E de cardinal n, autant de parties à p éléments que de parties à $n-p$ éléments

$$\binom{n}{p} = \binom{n}{n-p}$$

Cette propriété peut être jusitifié par la bijectivité du passage au complémentaire qui échange les parties à p éléments avec celles à $n-p$ éléments.

Proposition

$$\forall n \in \mathbb{N}, \sum_{p=0}^{n} \binom{n}{p} = 2^n.$$

dém. :
Si Card$E = n$ alors

$$\text{Card}\mathcal{P}(E) = 2^n$$

Or toute partie de E est constituée d'un nombre p d'éléments avec $0 \leqslant p \leqslant n$ et il existe $\binom{n}{p}$ parties de E à p éléments.
Par suite

$$\sum_{p=0}^{n} \binom{n}{p} = \text{Card}\mathcal{P}(E) = 2^n$$

□

Exemple Pour $n \in \mathbb{N}^\star$, calculons $\sum_{p=0}^{n} p\binom{n}{p}$.

$$\sum_{p=0}^{n} p\binom{n}{p} = \sum_{p=1}^{n} p\binom{n}{p} = \sum_{p=1}^{n} n\binom{n-1}{p-1} = n\sum_{q=0}^{n-1}\binom{n-1}{q} = n2^{n-1}$$

2.4.7.3 Formule du binôme de Newton

Théorème

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

dém. :
Par récurrence sur $n \in \mathbb{N}$.
Pour $n = 0$, la propriété est immédiate sachant que $\forall a \in \mathbb{C}, a^0 = 1$ (même pour $a = 0\dots$)
Supposons la propriété établie au rang $n \geqslant 0$.

$$(a+b)^{n+1} = (a+b)^n(a+b)$$

Par hypothèse de récurrence

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

Par décalage d'indice, on peut réécrire la deuxième somme

$$\sum_{k=0}^{n}\binom{n}{k}a^{n-k}b^{k+1} = \sum_{k=1}^{n+1}\binom{n}{k-1}a^{n+1-k}b^k$$

En adjoignant un terme nul à chacune des deux sommes

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

En combinant les deux sommes en une seule

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

en vertu de la formule du triangle de Pascal.
Récurrence établie.
□

Remarque On peut aussi énoncer :

$$(a+b)^n = (b+a)^n = \sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k}$$

Ainsi réapparaît la symétrie entre a et b.

Exemple $(a+b)^2 = a^2 + 2ab + b^2$,
$(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$,
$(a+b)^4 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4$.
En changeant b en $-b$:
$(a-b)^2 = a^2 - 2ab + b^2$
$(a-b)^3 = a^3 - 3a^2b + 3ab^2 - b^3$

Exemple Pour tout $x \in \mathbb{R}$,

$$(1+x)^n = \sum_{k=0}^{n}\binom{n}{k}x^k$$

Pour $x = 1$, on obtient

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n$$

Pour $x = -1$, on obtient

$$\sum_{k=0}^{n} \binom{n}{k} (-1)^k = 0^n = \begin{cases} 0 & \text{si } n \in \mathbb{N}^\star \\ 1 & \text{si } n = 0 \end{cases}$$

Pour $x = 2$,

$$\sum_{k=0}^{n} \binom{n}{k} 2^k = (1+2)^n = 3^n$$

En dérivant la relation initiale

$$n(1+x)^{n-1} = \sum_{k=1}^{n} k \binom{n}{k} x^{k-1}$$

Pour $x = 1$, on obtient

$$\sum_{k=0}^{n} k \binom{n}{k} = \sum_{k=1}^{n} k \binom{n}{k} = n2^{n-1}$$

En dérivant

$$nx(1+x)^{n-1} = \sum_{k=1}^{n} k \binom{n}{k} x^k$$

puis en évaluant en $x = 1$, on obtient

$$\sum_{k=0}^{n} k^2 \binom{n}{k} = n(n+1)2^{n-2}$$

Exemple Soient $n, p, q \in \mathbb{N}$ tels que $n \leqslant p + q$. Montrons

$$\sum_{k=0}^{n} \binom{p}{k} \binom{q}{n-k} = \binom{p+q}{n}$$

Exploitons la relation

$$(1+x)^n = \sum_{k=0}^{n} \binom{n}{k} x^k :$$

- le coefficient de x^n dans $(1+x)^{p+q}$ est

$$\binom{p+q}{n}$$

- le coefficient de x^n dans le développement de $(1+x)^p(1+x)^q$ est

$$\sum_{k=0}^{n} \binom{p}{k} \binom{q}{n-k}$$

Par identification des coefficients d'une fonction polynomiale, on obtient la relation proposée.

Chapitre 3

Ensemble ordonné

3.1 Relation d'ordre

E désigne un ensemble.

3.1.1 Définition

Définition

> On appelle relation binaire $\mathcal{R}$ sur E toute propriété vraie pour certains couples (x, y) d'éléments de E et fausse pour les autres.
> Lorsqu'un couple (x, y) vérifie la relation $\mathcal{R}$, on écrit $x\mathcal{R}y$. Sinon, on écrit $x\mathcal{R}y$ en barrant le $\mathcal{R}$.

Exemple L'égalité est une relation binaire sur E notée $=$.

Exemple Sur $E = \mathbb{R}$, la relation inférieur ou égal est une relation binaire notée $\leqslant$.

Exemple Sur $\mathcal{P}(E)$, l'inclusion est une relation binaire notée $\subset$.

Exemple Sur $E = \mathbb{R}$, on définit une relation binaire $\mathcal{R}$ par :

$$x\mathcal{R}y \Leftrightarrow \sin xy = y\mathrm{e}^{x+y^2}$$

Définition

Soit $\mathcal{R}$ une relation binaire sur E.
On dit que $\mathcal{R}$ est réflexive si
$$\forall x \in E, x\mathcal{R}x$$
On dit que $\mathcal{R}$ est symétrique si
$$\forall x, y \in E, x\mathcal{R}y \Leftrightarrow y\mathcal{R}x$$
On dit que $\mathcal{R}$ est antisymétrique si
$$\forall x, y \in E, x\mathcal{R}y \text{ et } y\mathcal{R}x \Rightarrow x = y$$
On dit que $\mathcal{R}$ est transitive si
$$\forall x, y, z \in E, x\mathcal{R}y \text{ et } y\mathcal{R}z \Rightarrow x\mathcal{R}z$$

Exemple Sur E, l'égalité est à la fois réflexive, symétrique, antisymétrique et transitive.

Exemple Sur E, la relation $\neq$ est symétrique. Elle n'est ni réflexive ni transitive.

Exemple Sur $E = \mathbb{R}$,
$$x\mathcal{R}y \Leftrightarrow \sin x = \sin y$$
définit une relation réflexive, symétrique et transitive.

Définition

Une relation binaire à la fois réflexive, symétrique et transitive est appelée une relation d'équivalence.

Exemple L'égalité est une relation d'équivalence sur E.

Exemple Sur $E = \mathbb{R}$,
$$x\mathcal{R}y \Leftrightarrow \sin x = \sin y$$
définit une relation d'équivalence.

Définition

On appelle relation d'ordre sur un ensemble E, toute relation binaire à la fois réflexive, antisymétrique et transitive.
A défaut d'autres notations, une relation d'ordre est usuellement notée $\preccurlyeq$ à défaut d'autres notations.

Exemple La relation $\leqslant$ est une relation d'ordre sur $\mathbb{R}$.

Exemple L'inclusion est une relation d'ordre $\mathcal{P}(E)$.

Exemple On dit que $m \in \mathbb{N}^\star$ divise $n \in \mathbb{N}^\star$, et note $m \mid n$, s'il existe $k \in \mathbb{N}$ tel que $n = mk$.
$\mid$ est une relation d'ordre sur $\mathbb{N}^\star$.

Exemple La relation $\mathcal{R}$ sur $\mathcal{P}(E)$ définie par

$$A\mathcal{R}B \Leftrightarrow \text{Card}A \leqslant \text{Card}B$$

n'est pas une relation d'ordre : cette relation n'est pas antisymétrique

3.1.2 Ensemble ordonné

Définition

On appelle ensemble ordonné tout couple $(E, \preccurlyeq)$ formé d'un ensemble E et d'une relation d'ordre $\preccurlyeq$ sur E.

Exemple $(\mathbb{R}, \leqslant)$, $(\mathcal{P}(E), \subset)$ et $(\mathbb{N}^\star, \mid)$ sont des ensembles ordonnés.

Définition

Soit $(E, \preccurlyeq)$ un ensemble ordonné.
On appelle ordre inverse associé à $\preccurlyeq$, la relation $\succcurlyeq$ définie par :

$$x \succcurlyeq y \Leftrightarrow y \preccurlyeq x$$

Remarque $\succcurlyeq$ est aussi une relation d'ordre sur E.

Exemple $\geqslant$ et $\supset$ sont les ordres inverses sur $\leqslant$ et $\subset$.

Définition

Soit $(E, \preccurlyeq)$ un ensemble ordonné.
On appelle ordre strict associé à $\preccurlyeq$ la relation $\prec$ définie par :

$$x \prec y \Leftrightarrow x \preccurlyeq y \text{ et } x \neq y.$$

Remarque $\prec$ n'est pas une relation d'ordre car non réflexive.

Exemple $<$ et $\subsetneq$ sont les ordres stricts associés à $\leqslant$ et $\subset$.

3.1.3 Ordre total, ordre partiel

Définition

Soit $(E, \preccurlyeq)$ un ensemble ordonné.
Deux éléments x et y de E sont dits comparables si $x \preccurlyeq y$ ou $y \preccurlyeq x$.

Exemple Dans $(\mathbb{R}, \leqslant)$, tous les réels sont deux à deux comparables.

Exemple Dans $(\mathbb{N}^\star, |)$, 2 et 3 ne sont pas comparables.

Définition

Soit $(E, \preccurlyeq)$ un ensemble ordonné.
On dit que l'ordre $\preccurlyeq$ est total si tous les éléments de E sont deux à deux comparables. On dit alors que $(E, \preccurlyeq)$ est un ensemble totalement ordonné.
Sinon, on parle d'ordre partiel et d'ensemble partiellement ordonné.

Exemple $(\mathbb{R}, \leqslant)$ est un ensemble totalement ordonné.

Exemple $(\mathbb{N}^\star, |)$ est un ensemble partiellement ordonné.

Exemple Si E contient au moins deux éléments distincts a et b alors $(\mathcal{P}(E), \subset)$ est un ensemble qui n'est que partiellement ordonné ; en effet $\{a\}$ et $\{b\}$ ne sont pas comparables.

3.1.4 Deux relations d'ordre sur $\mathbb{R}^2$

Exemple Sur $\mathbb{R}^2$, on définit une relation binaire $\preccurlyeq$ par :

$$(x, y) \preccurlyeq (x', y') \Leftrightarrow x \leqslant x' \text{ et } y \leqslant y'$$

C'est une relation d'ordre sur $\mathbb{R}^2$.
En effet :
Puisque $x = x$ et $y = y$, on a $(x, y) \preccurlyeq (x, y)$; la relation est réflexive.
Si $(x, y) \preccurlyeq (x', y')$ et $(x', y') \preccurlyeq (x, y)$ alors $x \leqslant x'$, $y \leqslant y'$ et $x' \leqslant x$, $y' \leqslant y$. On en déduit $x = x'$ et $y = y'$ donc $(x, y) = (x', y')$; la relation est antisymétrique.
Si $(x, y) \preccurlyeq (x', y')$ et $(x', y') \preccurlyeq (x'', y'')$ alors $x \leqslant x'$, $y \leqslant y'$ et $x' \leqslant x''$, $y' \leqslant y''$. On en déduit $x \leqslant x''$ et $y \leqslant y''$ donc $(x, y) \preccurlyeq (x'', y'')$; la relation est transitive.
La relation d'ordre ainsi définie n'est que partielle. En effet les couples $(0, 1)$ et $(1, 0)$ ne sont par comparables.

Exemple Sur $\mathbb{R}^2$, on définit une relation binaire $\preccurlyeq$ par :

$$(x, y) \preccurlyeq (x', y') \Leftrightarrow x < x' \text{ ou } (x = x' \text{ et } y \leqslant y')$$

Montrons qu'il s'agit d'une relation d'ordre total.
La relation $\preccurlyeq$ est évidemment réflexive. Pour obtenir son antisymétrie et sa transitivité, remarquons que :
- si $(x, y) \preccurlyeq (x', y')$ alors $x \leqslant x'$;
- si $(x, y) \preccurlyeq (x', y')$ et $x = x'$ alors $y \leqslant y'$.
Ainsi, si $(x, y) \preccurlyeq (x', y')$ et $(x', y') \preccurlyeq (x, y)$, on a $x \leqslant x'$ et $x' \leqslant x$ donc $x = x'$ puis sachant $(x, y) \preccurlyeq (x', y')$ et $(x', y') \preccurlyeq (x, y)$ avec $x = x'$, on a $y \leqslant y'$ et $y' \leqslant y$ donc $y = y'$. Au final $(x, y) = (x', y')$; la relation est antisymétrique.
Aussi, si $(x, y) \preccurlyeq (x', y')$ et $(x', y') \preccurlyeq (x'', y'')$, on a $x \leqslant x'$ et $x' \leqslant x''$ donc $x \leqslant x''$.
Si $x < x''$, on peut directement conclure $(x, y) \preccurlyeq (x'', y'')$.
Sinon, $x = x''$ puis par encadrement $x = x' = x''$. Sachant $(x, y) \preccurlyeq (x', y')$ et $(x', y') \preccurlyeq (x'', y'')$ avec $x = x'$ et $x' = x''$, on a $y \leqslant y'$ et $y' \leqslant y''$ donc $y \leqslant y''$ et on obtient encore $(x, y) \preccurlyeq (x'', y'')$. Ainsi la relation est transitive.
De plus, cette relation définit un ordre total puisque pour deux couples $(x, y), (x', y')$:
si $x < x'$ alors $(x, y) \preccurlyeq (x', y')$;
si $x' < x$ alors $(x', y') \preccurlyeq (x, y)$;
si $x = x'$ alors, selon $y \leqslant y'$ ou $y' < y$, on a $(x, y) \preccurlyeq (x', y')$ ou $(x', y') \preccurlyeq (x, y)$.
Finalement, dans tous les cas, les couples $(x, y), (x', y')$ sont comparables.

Remarque Plus généralement, si $(E, \preccurlyeq)$ et $(F, \preccurlyeq)$ sont deux ensembles ordonnés, les deux démarches qui précèdent permettent de définir des relations d'ordre sur le produit cartésien $E \times F$. Pour l'ordre produit, on obtient a priori un ordre partiel, pour l'ordre lexicographique, on obtient un ordre total si les relations d'ordre sur E et F sont totales.
On peut aussi généraliser ce qui précède à un produit cartésien de plusieurs ensembles ordonnés.

3.2 Relation d'ordre et sous ensembles

3.2.1 Partie minorée, partie majorée

Soit $(E, \preccurlyeq)$ un ensemble ordonné et A une partie de E.

Définition

On appelle majorant de A (resp. minorant), s'il en existe, tout élément $M \in E$ tel que $\forall a \in A, a \preccurlyeq M$ (resp. $M \preccurlyeq a$).
On note Majo(A) (resp. Mino(A)) l'ensemble de ces éléments.

Remarque Un majorant/minorant doit pouvoir être comparé à tout élément de la partie.

Exemple Dans $(\mathbb{R}, \leqslant)$.
Pour $A =]0, 1] \cup [2, 3]$, Majo$(A) = [3, +\infty[$ et Mino$(A) =]-\infty, 0]$.
Pour $A = \mathbb{N}$, Majo$(A) = \emptyset$ et Mino$(A) = \mathbb{R}^-$.

Exemple Dans $(\mathbb{N}^\star, |)$:
Pour $A = \{4, 6\}$, Majo$(A) = \{12, 24, 36, \ldots\}$ et Mino$(A) = \{1, 2\}$.
Pour $A = \{12, 18\}$, Majo$(A) = \{36, 72, \ldots\}$ et Mino$(A) = \{1, 2, 3, 6\}$.

Exemple Dans $(\mathcal{P}(E), \subset)$ avec $E = \{a, b, c, d\}$.
Pour $A = \{\{a, b\}, \{b\}, \{a, c\}\}$: Majo$(A) = \{\{a, b, c\}, E\}$ et Mino$(A) = \{\emptyset\}$.

Définition

La partie A est dite majorée (resp. minorée) si elle possède un majorant (resp. minorant). Une partie est dite bornée si elle est majorée et minorée.

Exemple Les segments de $\mathbb{R}$ sont des parties bornées.

Exemple Dans $(\mathbb{N}^\star, |)$, toute partie est minorée par 1.

Exemple Dans $(\mathcal{P}(E), \subset)$, toute partie est majorée par E et minorée par $\emptyset$.

3.2.2 Extremum d'une partie

Soit $(E, \preccurlyeq)$ un ensemble ordonné et A une partie de E.

Définition

On appelle plus grand élément de A (resp. plus petit élément), s'il en existe, tout élément $M \in A$ tel que

$$\forall a \in A, a \preccurlyeq M \text{ (respectivement } M \preccurlyeq a \text{).}$$

Remarque Un plus grand élément est un majorant qui appartient à la partie.

Proposition

Si A admet un plus grand élément (resp. plus petit élément) celui-ci est unique.
On le note

$$\max(A) \text{ (respectivement } \min(A) \text{)}$$

dém. :
Soient M, M' deux plus grands éléments de A.
On peut écrire $M \preccurlyeq M'$ car M' majore A et $M \in A$ mais aussi par symétrie $M' \preccurlyeq M$.
Par l'antisymétrie de la relation $\preccurlyeq$, on obtient $M = M'$.
□

Exemple Dans $(\mathbb{R}, \leqslant)$.
Pour $A = [0, 1[$, $\min(A) = 0$ et $\max(A)$ n'existe pas.
Pour $A = \{1/n / n \in \mathbb{N}^\star\} = \{1, 1/2, 1/3, \ldots\}$, $\max(A) = 1$ et $\min(A)$ n'existe pas.

3.2.3 Propriétés fondatrices des ensembles de nombres entiers

On considère la relation $\leqslant$ sur $\mathbb{N}$ ou $\mathbb{Z}$:
- toute partie non vide de $\mathbb{N}$ admet un plus petit élément ;
- toute partie non vide et majorée de $\mathbb{N}$ admet un plus grand élément.
Par extension :
- toute partie non vide et minorée de $\mathbb{Z}$ possède un plus petit élément ;
- toute partie non vide et majorée de $\mathbb{Z}$ possède un plus grand élément.

Exemple Soit $n \in \mathbb{N}^\star$. Montrons qu'il existe un plus grand $p \in \mathbb{N}$ tel que $2^p \mid n$.
Considérons

$$A = \{p \in \mathbb{N}/2^p \mid n\}$$

A est une partie non vide de $\mathbb{N}$ car $p = 0 \in A$ puisque $2^0 = 1$ divise n.
Pour tout $p \in A$, $2^p \leqslant n$ donc $p \leqslant \log_2 n$. On en déduit que la partie A est majorée par $E(\log_2 n) + 1$.
Puisque A est une partie de $\mathbb{N}$ non vide et majorée, A possède un plus grand élément.

Proposition

Soit $E \subset \mathbb{N}$.
Si $0 \in E$ et si l'on a la propriété $\forall p \in \mathbb{N}, p \in E \Rightarrow p + 1 \in E$
alors $E = \mathbb{N}$.

dém. :
Soit $m \in \mathbb{N}$.
Pour montrer que $m \in E$ considérons

$$A = \{n \in \mathbb{N}/n \in E \text{ et } n \leqslant m\}$$

A admet un plus grand élément p.
Comme $p \in A$ on a $p \in E$ et $p \leqslant m$.
Comme $p \in E$ on a $p + 1 \in E$.
Or $p + 1 \notin A$ donc $p + 1 > m$ d'où $p \leqslant m < p + 1$.
Ainsi $p = m$ et puisque $p \in E$ on a $m \in E$.
Finalement $\mathbb{N} \subset E$ puis $E = \mathbb{N}$.
□

Théorème

Soient $n_0 \in \mathbb{N}$ et $\mathcal{P}(n)$ une assertion dépendant d'un entier $n \geqslant n_0$.
Si
1) $\mathcal{P}(n_0)$ est vraie et
2) $\forall n \geqslant n_0, \mathcal{P}(n)$ vraie $\Rightarrow \mathcal{P}(n+1)$ vraie
alors
$\forall n \geqslant n_0, \mathcal{P}(n)$ est vraie.

dém. :
On applique le principe de récurrence à l'ensemble

$$E = \{p \in \mathbb{N}/\mathcal{P}(n_0 + p) \text{ est vraie}\}$$

□

Théorème

Soient $n_0 \in \mathbb{N}$ et $\mathcal{P}(n)$ une assertion dépendant d'un entier $n \geqslant n_0$.
Si
1) $\mathcal{P}(n_0)$ et $\mathcal{P}(n_0+1)$ sont vraies et
2) $\forall n \geqslant n_0$, (($\mathcal{P}(n)$ et $\mathcal{P}(n+1)$ vraies) $\Rightarrow \mathcal{P}(n+2)$ vraie
alors
$\forall n \geqslant n_0, \mathcal{P}(n)$ est vraie.

dém. :
On applique le principe de récurrence à l'assertion $\mathcal{Q}(n) =$ « $\mathcal{P}(n)$ et $\mathcal{P}(n+1)$ vraies »
□

Exemple Soit (u_n) une suite réelle telle que

$$\begin{cases} u_0 = 0, u_1 = 1 \\ \forall n \in \mathbb{N}^\star, u_{n+1} = 2u_n - u_{n-1} \end{cases}$$

Montrons que $\forall n \in \mathbb{N}, u_n = n$.
On procède par récurrence double...
Pour $n = 0$ et $n = 1$: la propriété $u_n = n$ est vraie.
Supposons $u_n = n$ et $u_{n+1} = n+1$ pour un rang $n \geqslant 0$.
Par définition de la suite (u_n), on a

$$u_{n+2} = 2u_{n+1} - u_n = 2(n+1) - n = n+2$$

Récurrence établie.

Remarque La récurrence double peut évidemment être généraliser aux récurrences triples, quadruples,... On parle ici de récurrence multiple. Pour que celles-ci s'enchaînent correctement, il conviendra de vérifier une initialisation suffisante.

Théorème

Soient $n_0 \in \mathbb{N}$ et $\mathcal{P}(n)$ une assertion dépendant d'un entier $n \geqslant n_0$.
Si
1) $\mathcal{P}(n_0)$ est vraie et
2) $\forall n \geqslant n_0$; $(\mathcal{P}(n_0), \ldots, \mathcal{P}(n)$ vraies) $\Rightarrow \mathcal{P}(n+1)$ vraie.
alors $\forall n \geqslant n_0, \mathcal{P}(n)$ est vraie.

dém. :
On applique la récurrence simple à l'assertion $\mathcal{Q}(n) =$ « $\forall n_0 \leqslant k \leqslant n, \mathcal{P}(k)$ vraie ».
□

Exemple Soit $(u_n) \in \mathbb{R}^{\mathbb{N}}$ telle que

$$\forall n \in \mathbb{N}, u_{n+1} = u_0 + u_1 + \ldots + u_n$$

Montrer que

$$\forall n \in \mathbb{N}^\star, u_n = 2^{n-1} u_0$$

Procédons par récurrence forte sur $n \in \mathbb{N}^\star$...

Pour $n = 1$, $u_1 = u_0 = 2^0 u_0$, la propriété est vraie.
Supposons la propriété vraie jusqu'au rang $n \geqslant 1$.
On a alors

$$u_{n+1} = u_0 + 2^0 u_0 + 2^1 u_1 + \cdots + 2^{n-1} u_n$$

Par sommation géométrique

$$u_{n+1} = u_0 + u_0 \frac{1 - 2^n}{1 - 2} = u_0 + (2^n - 1)u_0 = 2^n u_0$$

Récurrence établie.

Théorème

Soient $n_0, n_1 \in \mathbb{N}$ tels que $n_0 \leqslant n_1$ et $\mathcal{P}(n)$ une assertion dépendant d'un entier $n_0 \leqslant n \leqslant n_1$.
Si
1) $\mathcal{P}(n_0)$ est vraie et
2) $\forall n_0 \leqslant n < n_1$ on a $\mathcal{P}(n)$ vraie $\Rightarrow \mathcal{P}(n+1)$ vraie
alors $\forall n_0 \leqslant n \leqslant n_1, \mathcal{P}(n)$ est vraie.

dém. :
On applique la récurrence simple à $\mathcal{Q}(n) =$ « $\mathcal{P}(n)$ vraie ou $n > n_1$ ».
□

Remarque Dans le cadre de la récurrence finie, on peut aussi envisager des récurrences descendantes.

3.2.4 Borne supérieure, borne inférieure

Soit $(E, \preccurlyeq)$ un ensemble ordonné et A une partie de E.

Définition

On appelle borne supérieure de A, si elle existe, le plus petit des majorants de A. On la note $\sup A$.
On appelle borne inférieure de A, si elle existe, le plus grand des minorants de A. On la note $\inf A$.
Sous réserve d'existence :

$$\sup A = \min(Majo(A)) \text{ et } \inf A = \max(Mino(A))$$

Exemple Dans $(\mathbb{R}, \leqslant)$:
Pour $A = [0, 1[$, on a $Majo(A) = [1, +\infty[$ et $Mino(A) =]-\infty, 0]$ donc $\sup A$ et $\inf A$ existent avec $\sup A = 1$ et $\inf A = 0$.
Pour $A =]-\infty, 1]$, on a $Majo(A) = [1, +\infty[$ $Mino(A) = \emptyset$ donc $\sup A$ existe et $\sup A = 1$. En revanche $\inf A$ n'existe pas.

Attention : $\sup A$ et $\inf A$, lorsqu'ils existent, ne sont pas a priori des éléments de A.
Cependant :

Proposition

Si A admet un plus grand élément alors A admet une borne supérieure et $\sup A = \max A$.
Si A admet un plus petit élément alors A admet une borne inférieure et $\inf A = \min A$.

dém. :
Supposons : A admet un plus grand élément M.
M est un majorant de A et pour tout majorant M' de A, on a $M \preccurlyeq M'$ puisque $M \in A$.
Par suite M est le plus petit des majorants de A, c'est sa borne supérieure.
□

Attention : $\sup A$ existe $\not\Rightarrow$ $\max A$ existe.
En revanche : $\sup A$ existe et $\sup A \in A \Rightarrow \max A = \sup A$.

3.2.5 Propriétés fondatrices des nombres réels

Par construction de la droite réelle :
- toute partie non vide et majorée de $\mathbb{R}$ admet une borne supérieure ;
- toute partie non vide et minorée de $\mathbb{R}$ admet une borne inférieure.

Remarque L'ensemble $\mathbb{Q}$ ne possède pas cette propriété.
Par exemple la partie $A = \left\{r \in \mathbb{Q}/r^2 \leqslant 2\right\}$, qui est non vide et majorée, n'a pas de borne supérieure dans $\mathbb{Q}$; en effet $\sqrt{2} \notin \mathbb{Q}$.

Remarque Pour manipuler correctement $\sup A$ on retient :
- si A est une partie non vide et majorée de $\mathbb{R}$ alors $\sup A$ existe ;
- $\sup A$ est caractérisée par :
a) $\sup A$ est un majorant de A,
b) tout majorant de A est supérieur à $\sup A$.

Exemple Soit

$$A = \left\{\frac{1}{n}/n \in \mathbb{N}^\star\right\} = \left\{1, \frac{1}{2}, \frac{1}{3}, ...\right\}$$

Déterminons $\sup A$ et $\inf A$.
La partie A admet un plus grand élément donc $\sup A = \max A = 1$.
La partie A est une partie de $\mathbb{R}$ non vide et minorée par 0.
On en déduit que $\inf A$ existe et $\inf A \geqslant 0$.
De plus, puisque $\inf A$ minore A , on a la propriété :

$$\forall n \in \mathbb{N}^\star, 1/n \geqslant \inf A$$

En passant à la limite quand $n \to +\infty$, on obtient $0 \geqslant \inf A$.
On conclut $\inf A = 0$.

Exemple Soient A et B deux parties non vides et bornées de $\mathbb{R}$ telles que $A \subset B$.
Montrons

$$\inf B \leqslant \inf A \leqslant \sup A \leqslant \sup B$$

Puisque les parties A et B sont des parties de $\mathbb{R}$ non vides, minorées et majorées, on est assuré de l'existence des bornes proposées.
Pour tout $x \in A$, on a $x \in B$ car $A \subset B$. Or la partie B est majorée par $\sup B$ donc $x \leqslant \sup B$. Ainsi $\sup B$ est un majorant de la partie A. Or par définition $\sup A$ est le plus petit des majorants de A donc $\sup A \leqslant \sup B$.
Par un raisonnement symétrique, on obtient aussi $\inf B \leqslant \inf A$.
Enfin, en considérant un élément $a \in A$, on peut affirmer $\inf A \leqslant a$ et $a \leqslant \sup A$ donc $\inf A \leqslant \sup A$.

Définition

Si A est une partie de $\mathbb{R}$ non vide et non majorée, on pose $\sup A = +\infty$.
Si A est une partie de $\mathbb{R}$ non vide et non minorée, on pose $\inf A = -\infty$.
Si $A = \emptyset$, on pose $\sup A = -\infty$ et $\inf A = +\infty$.

Exemple Si I est un intervalle non vide, ses extrémités dans $\bar{\mathbb{R}}$ correspondent à $\inf I$ et $\sup I$.

Théorème

Soit A une partie non vide de $\mathbb{R}$.
Il existe une suite $(u_n) \in A^{\mathbb{N}}$ telle que $u_n \to \sup A \in \bar{\mathbb{R}}$.
Il existe une suite $(v_n) \in A^{\mathbb{N}}$ telle que $v_n \to \inf A \in \bar{\mathbb{R}}$.

dém. :
Supposons $\sup A = +\infty$.
Pour tout $n \in \mathbb{N}$, n n'est pas majorant de A donc il existe $a \in A$ tel que $a > n$.
Posons $u_n = a$. En faisant varier n, ce qui précède définit une suite $(u_n) \in A^{\mathbb{N}}$ qui par comparaison tend vers $+\infty$ puisque $u_n > n$ pour tout $n \in \mathbb{N}$.
Supposons $M = \sup A \in \mathbb{R}$.
Pour tout $n \in \mathbb{N}$, $M - 1/(n+1)$ n'est pas majorant de A donc il existe $a \in A$ tel que $M - 1/(n+1) \leqslant a$, de plus $a \leqslant M$. Posons $u_n = a$.
En faisant varier n, ce qui précède définit une suite $(u_n) \in A^{\mathbb{N}}$ qui par le théorème des gendarmes tend vers M.
□

Exemple Soient A et B deux parties non vides de $\mathbb{R}$ telles que

$$\forall a \in A, \forall b \in B, a \leqslant b$$

Montrons $\sup A \leqslant \inf B$.
Soient $(a_n)^{\mathbb{N}}$ et $(b_n) \in B^{\mathbb{N}}$ telles que $a_n \to \sup A$ et $b_n \to \inf B$.
Pour tout $n \in \mathbb{N}$, $a_n \leqslant b_n$ car $a_n \in A$ et $b_n \in B$.
En passant cette comparaison à la limite, on obtient $\sup A \leqslant \inf B$.

3.3 Fonctions et relation d'ordre

Les définitions qui suivent, présentées dans un cadre général, s'appliqueront en particulier aux fonctions à valeurs réelles et aux suites de nombres réels.

3.3.1 Comparaison de fonction

Soient E un ensemble et $(F, \preccurlyeq)$ un ensemble ordonné

Définition

On définit une relation binaire notée $\preccurlyeq$ sur $\mathcal{F}(E, F)$ par :

$$f \preccurlyeq g \Leftrightarrow \forall x \in E, f(x) \preccurlyeq g(x)$$

Exemple Soient $f, g : E \to \mathbb{R}$.

$$f \leqslant g \Leftrightarrow \forall x \in E, f(x) \leqslant g(x)$$

Exemple Soient $(u_n), (v_n) \in \mathbb{R}^{\mathbb{N}}$.

$$(u_n) \leqslant (v_n) \Leftrightarrow \forall n \in \mathbb{N}, u_n \leqslant v_n$$

Proposition

$\preccurlyeq$ est une relation d'ordre sur $\mathcal{F}(E, F)$.

dém. :
Pour tout $x \in E$, $f(x) \preccurlyeq f(x)$ donc $f \preccurlyeq f$.
Supposons $f \preccurlyeq g$ et $g \preccurlyeq f$. Pour tout $x \in E$, $f(x) \preccurlyeq g(x)$ et $g(x) \preccurlyeq f(x)$ donc $f(x) = g(x)$. Par suite $f = g$.
Supposons $f \preccurlyeq g$ et $g \preccurlyeq h$. Pour tout $x \in E$, $f(x) \preccurlyeq g(x)$ et $g(x) \preccurlyeq h(x)$ donc $f(x) \preccurlyeq h(x)$. Par suite $f \preccurlyeq h$.
Puisque réflexive, antisymétrique et transitive la relations $\preccurlyeq$ sur $\mathcal{F}(E, F)$ est une relation d'ordre.
□

Remarque Si E et F contiennent tous deux au moins deux éléments, on peut montrer que l'ordre n'est que partiel.

3.3.2 Monotonie de fonctions

Soient $(E, \preccurlyeq)$, $(F, \preccurlyeq)$ et $(G, \preccurlyeq)$ trois ensembles ordonnés.

Définition

On dit que $f : E \to F$ est croissante (resp. décroissante) si

$$\forall x, y \in E, x \preccurlyeq y \Rightarrow f(x) \preccurlyeq f(y) \text{ (resp. } x \preccurlyeq y \Rightarrow f(y) \preccurlyeq f(x) \text{)}$$

On dit que $f : E \to F$ est strictement croissante (resp. décroissante) si

$$\forall x, y \in E, x \prec y \Rightarrow f(x) \prec f(y) \text{ (resp. } x \prec y \Rightarrow f(y) \prec f(x) \text{)}$$

On dit que f est monotone ssi f est croissante ou décroissante.

Exemple La fonction $f : x \mapsto [-x, x]$ est une application croissante de $(\mathbb{R}^+, \leqslant)$ vers $(\mathcal{P}(\mathbb{R}), \subset)$

Exemple La fonction $f : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = x^3$ est strictement croissante.
f est dérivable, $f'(x) \geqslant 0$ et ne s'annule qu'en $x = 0$.

Exemple La fonction $f : \mathbb{R}^\star \to \mathbb{R}$ définie par $f(x) = 1/x$ n'est pas décroissante.
En revanche, ses restrictions à $]-\infty, 0[$ et $]0, +\infty[$ le sont.

Proposition

> Soient $f : E \to F$ et $g : F \to G$.
> Si f et g ont même monotonie alors $g \circ f$ est croissante.
> Si f et g sont de monotonies contraires alors $g \circ f$ est décroissante.

dém. :
C'est immédiat !
□

Exemple $f : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = \ln(\mathrm{e}^{-x} + 1)$ est strictement décroissante par composition de monotonie.

Remarque Par la définition qui précède, une suite $(u_n) \in E^{\mathbb{N}}$ est dite croissante si

$$\forall n, m \in E, n \leqslant m \Rightarrow u_n \preccurlyeq u_m$$

Plus efficacement, la monotonie de (u_n) s'étudie en comparant u_n et u_{n+1} grâce au résultat suivant :

Proposition

> Soit $(u_n)_{n \in \mathbb{N}}$ une suite d'éléments de E.
> La suite (u_n) est croissante (resp. décroissante) si, et seulement si,
>
> $$\forall n \in \mathbb{N}, u_{n+1} \succcurlyeq u_n \text{ (resp. } u_{n+1} \preccurlyeq u_n \text{)}$$
>
> La suite (u_n) est strictement croissante (resp. décroissante) si, et seulement si,
>
> $$\forall n \in \mathbb{N}, u_{n+1} \prec u_n \text{ (resp. } u_{n+1} \prec u_n \text{)}$$

dém. :
$(\Rightarrow)$ ok
$(\Leftarrow)$ Supposons $\forall n \in \mathbb{N}, u_{n+1} \succcurlyeq u_n$.
Par récurrence sur $n \in \mathbb{N}$, on montre que pour tout $0 \leqslant m \leqslant n$, on a $u_m \preccurlyeq u_n$.
□

Remarque Pour étudier la monotonie d'une suite réelle (u_n), on peut regarder le signe $u_{n+1} - u_n$.

Exemple La suite (u_n) de terme général $u_n = \sum_{k=1}^{n} 1/k$ est croissante.
En effet
$$u_{n+1} - u_n = \frac{1}{n+1} \geqslant 0$$

Exemple Pour $n \in \mathbb{N}$, posons $I_n = [-n, n]$.
La suite $(I_n)_{n\in\mathbb{N}}$ est une suite croissante de segments de $\mathbb{R}$.
En effet
$$\forall n \in \mathbb{N}, I_n \subset I_{n+1}$$

Remarque Pour étudier la monotonie d'une suite (u_n) de réels strictement positifs, on peut comparer le rapport u_{n+1}/u_n à 1.

Exemple La suite de terme général
$$u_n = \prod_{k=1}^{n} \left(1 - \frac{1}{2k^2}\right)$$
est décroissante.
En effet, cette suite est formé de termes strictement positifs (par produit de facteurs strictement positifs) et
$$\frac{u_{n+1}}{u_n} = 1 - \frac{1}{2(n+1)^2} < 1$$
donc $u_{n+1} < u_n$.

3.3.3 Fonction minorée, majorée

Soient E un ensemble, $(F, \preccurlyeq)$ un ensemble ordonné.

Définition

On appelle majorant de $f : E \to F$, s'il en existe, tout majorant de $\mathrm{Im} f$ i.e. tout élément $M \in F$ tel que
$$\forall x \in E, f(x) \preccurlyeq M$$
On appelle minorant de $f : E \to F$, s'il en existe, tout minorant de $\mathrm{Im} f$ i.e. tout élément $M \in F$ tel que
$$\forall x \in E, f(x) \succcurlyeq M$$
La fonction $f : E \to F$ est dite majorée (resp. minorée) si elle possède au moins un majorant (resp. minorant).
Une fonction minorée et majorée est dite bornée.

Exemple $f : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = \mathrm{e}^x$ est minorée, non majorée.

Exemple $u_n = \sin \mathrm{e}^n$ est le terme général d'une suite (u_n) bornée.

3.3.4 Extremum d'une fonction

Soient E un ensemble et $(F, \preccurlyeq)$ un ensemble ordonné.

Définition

On dit que $f : E \to F$ admet un maximum en $a \in E$ si

$$\forall x \in E, f(x) \preccurlyeq f(a)$$

$f(a)$ apparaît alors comme étant le plus grand élément de Imf, on l'appelle valeur au maximum de f et on la note $\max f$ ou $\max\limits_{x \in E} f(x)$.
On dit que f admet un minimum en $a \in E$ si

$$\forall x \in E, f(x) \succcurlyeq f(a)$$

$f(a)$ apparaît alors comme étant le plus petit élément de Imf, on l'appelle valeurs au minimum de f et on la note $\min f$ ou $\min\limits_{x \in E} f(x)$.
On appelle extremum, un minimum ou un maximum d'une fonction.

Exemple Soient $a, b, c \in \mathbb{R}$ avec $a \neq 0$.
La fonction $f : x \mapsto ax^2 + bx + c$ admet un extremum en $-b/2a$.

3.3.5 Borne supérieure et borne inférieure d'une fonction réelle

Soit E un ensemble.

Définition

On appelle borne supérieure de $f : E \to \mathbb{R}$, si elle existe, la borne supérieure de Imf. On la note

$$\sup f \text{ ou } \sup_{x \in E} f(x)$$

On appelle borne inférieure de $f : E \to \mathbb{R}$, si elle existe, la borne inférieure de Imf. On la note

$$\inf f \text{ ou } \inf_{x \in E} f(x)$$

Proposition

Si une fonction réelle présente un maximum (resp. un minimum) alors la valeur en celui-ci est aussi borne supérieure (resp. inférieure) de cette fonction.

Exemple $\sup\limits_{x \in \mathbb{R}} \mathrm{e}^x = +\infty$ et $\inf\limits_{x \in \mathbb{R}} \mathrm{e}^x = 0$.

Exemple $\sup\limits_{n \in \mathbb{N}^\star} \dfrac{1}{n} = 1$ et $\inf\limits_{n \in \mathbb{N}^\star} \dfrac{1}{n} = 0$.

Remarque Il est aisé de déterminer les éventuels extrema et les bornes supérieure et inférieure d'une fonction réelle d'une variable réelle lorsqu'on connaît sont tableau de variation.

Exemple Etudions $f : \mathbb{R} \to \mathbb{R}$ définie par

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

f est dérivable et

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

Les variations de f sont données par

x	$-\infty$		$-1/2$		$+\infty$
$f'(x)$		$+$	0	$-$	
$f(x)$	0	$\nearrow$	$4/3$	$\searrow$	0

Sur ce tableau, on lit

$$\sup f = \max f = 4/3 \text{ et } \inf f = 0$$

Chapitre 4

Structures algébriques

On connaît plusieurs additions (sur les nombres, les vecteurs, les suites, les fonctions,. . .), ces additions se ressemblent beaucoup bien qu'opérant sur des objets très différents. . . Dans ce chapitre nous allons voir en quoi certaines opérations ont des propriétés commune et plus généralement on va isoler les propriétés calculatoires des opérations sans s'intéresser à la nature des objets qu'elles manipulent ; c'est ce qui rend ce chapitre assez abstrait. . .

4.1 Loi de composition interne

E désigne un ensemble.

4.1.1 Définition

Définition

On appelle loi de composition interne (l.c.i.) ou opération sur E toute application de $E \times E$ vers E. Lorsque l'on convient de noter $\star$ cette loi de composition interne, on note $x \star y$ l'image du couple (x, y) par l'application précédente.
L'élément $x \star y$ est appelé composé de x par y via $\star$.
Les lois de composition interne sont généralement notées $\star, \top, \perp, +, \times, \circ, ...$

Exemple L'addition et la multiplication sur $\mathbb{C}$ sont des lois de composition interne.
En effet $(x, y) \mapsto x + y$ et $(x, y) \mapsto xy$ sont des applications (bien définies) de $\mathbb{C} \times \mathbb{C}$ vers $\mathbb{C}$.

Exemple L'union et l'intersection sont des lois de composition interne sur $\mathcal{P}(E)$.

Exemple La composition des applications est une loi de composition interne sur $\mathcal{F}(E, E)$.

Définition

On appelle magma tout couple $(E, \star)$ formé d'un ensemble E et d'une loi de composition interne $\star$ sur E.

Exemple $(\mathbb{C}, +), (\mathbb{C}, \times), (E^E, \circ), (\mathcal{P}(E), \cup), (\mathcal{P}(E), \cap)$ sont des magmas usuels.

4.1.2 Partie stable

Définition

On appelle partie stable d'un magma $(E, \star)$ toute partie A de E vérifiant

$$\forall x, y \in A, x \star y \in A$$

Exemple E et $\emptyset$ sont des parties stables de $(E, \star)$.

Exemple $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ et $\mathbb{R}$ sont des parties stables de $(\mathbb{C}, +)$ et $(\mathbb{C}, \times)$.

Exemple $\mathfrak{S}(E)$ est une partie stable de $(\mathcal{F}(E, E), \circ)$ car la composée de deux permutations est une permutation.

Définition

Soit A une partie stable d'un magma $(E, \star)$. L'application restreinte

$$\begin{cases} A \times A \to A \\ (x, y) \mapsto x \star y \end{cases}$$

définit une loi de composition interne sur A appelée loi de composition interne induite par $\star$ sur A.
On la note $\star_{\restriction A}$, ou plus couramment $\star$, et on peut ainsi donner un sens au magma $(A, \star)$.

Exemple $(\mathbb{N}, +), (\mathbb{N}, \times), (\mathbb{Z}, +), (\mathbb{Z}, \times), (\mathbb{Q}, +), (\mathbb{Q}, \times), (\mathbb{R}, +), (\mathbb{R}, \times)$ et $(\mathfrak{S}(E), \circ)$ sont de nouveaux magmas usuels.

Exemple On peut aussi donner un sens à $(\mathbb{R}^\star, \times)$ mais pas à $(\mathbb{R}^\star, +)$!

4.1.3 Propriétés d'une loi de composition interne

4.1.3.1 Commutativité

Définition

Soit $\star$ une loi de composition interne sur E. On dit que deux éléments a et b de E commutent pour la loi $\star$ si

$$a \star b = b \star a$$

Exemple Dans $(\mathbb{C}, +)$ et dans $(\mathbb{C}, \times)$ tous les éléments commutent deux à deux.

Exemple Dans $(\mathcal{F}(E,E),\circ)$ ce n'est plus le cas mais néanmoins on peut dire que tout élément de $\mathcal{F}(E,E)$ commute avec Id_E.

Définition

Une loi de composition interne $\star$ sur E est dite commutative si tous les éléments de E commutent deux à deux.
Le magma $(E,\star)$ est alors dit commutatif.

Exemple $(\mathbb{C},+),(\mathbb{C},\times),(\mathcal{P}(E),\cup),(\mathcal{P}(E),\cap)$ sont des magmas commutatifs.

Proposition

Si A est une partie stable d'un magma commutatif $(E,\star)$ alors $(A,\star)$ est aussi commutatif.

dém. :
Si on a la propriété

$$\forall a,b\in E, a\star b=b\star a$$

on a a fortiori la suivante

$$\forall a,b\in A, a\star b=b\star a$$

car $A\subset E$.
□

4.1.3.2 Associativité

Remarque Si $\star$ est une loi de composition interne sur E, considérer l'élément $a\star b\star c$ est ambigu car l'opération $\star$ n'engage a priori que deux éléments. Ainsi $a\star b\star c$ ne peut être compris qu'avec un parenthèsage $(a\star b)\star c$ ou $a\star(b\star c)$ spécifiant comment est organisée l'opération.

Définition

Une loi de composition interne $\star$ sur E est dite associative si

$$\forall a,b,c\in E,(a\star b)\star c=a\star(b\star c)$$

Le magma $(E,\star)$ est alors dit associatif.

Remarque Si $\star$ est associative, les parenthèses n'étant plus nécessaires à la compréhension de l'opération, on note $a\star b\star c$ au lieu de $(a\star b)\star c$ ou $a\star(b\star c)$.
De manière plus générale, on peut aussi donner un sens à l'opération

$$a_1\star a_2\star\ldots\star a_n \text{ (encore noté } \mathop{\star}_{i=1}^{n} a_i\text{)}$$

sans avoir à préciser le parenthèsage

Exemple $(\mathbb{C},+),(\mathbb{C},\times),(\mathcal{F}(E,E),\circ),(\mathcal{P}(E),\cup),(\mathcal{P}(E),\cap)$ sont des magmas associatifs.

Exemple Le produit vectoriel définit une loi de composition interne ni commutative, ni associative sur l'ensemble des vecteurs de l'espace.

Proposition

Si A est une partie stable d'un magma associatif $(E, \star)$ alors $(A, \star)$ est aussi associatif.

dém. :
Si on a la propriété

$$\forall a, b, c \in E, (a \star b) \star c = a \star (b \star c)$$

on a a fortiori celle-ci

$$\forall a, b, c \in A, (a \star b) \star c = a \star (b \star c)$$

puisque $A \subset E$.
□

Exemple Etudions l'opération $\star$ définie sur $\mathbb{R}$ par

$$x \star y = \sqrt{x^2 + y^2}$$

Pour tout $x, y \in \mathbb{R}$, $x^2 + y^2 \geqslant 0$ et donc $\sqrt{x^2 + y^2}$ est bien définie dans $\mathbb{R}$.
Ainsi $\star : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ définie une loi de composition interne sur $\mathbb{R}$.
Pour tout $x, y \in \mathbb{R}$,

$$y \star x = \sqrt{y^2 + x^2} = \sqrt{x^2 + y^2} = x \star y$$

La loi $\star$ est commutative.
Pour tout $x, y, z \in \mathbb{R}$,

$$x \star (y \star z) = x \star \left(\sqrt{y^2 + z^2}\right) = \sqrt{x^2 + (y^2 + z^2)}$$

donc

$$x \star (y \star z) = \sqrt{(x^2 + y^2) + z^2} = \sqrt{x^2 + y^2} \star z = (x \star y) \star z$$

La loi $\star$ et associative.
Les parties $\mathbb{R}^+$ et $[1, +\infty[$ sont des exemples de parties stables de $(\mathbb{R}, \star)$.

4.1.4 Eléments particuliers

Soit $(E, \star)$ un magma

4.1.4.1 Élément régulier

Définition

On appelle élément régulier de $(E, \star)$ tout élément x de E vérifiant

$$\forall a, b \in E, x \star a = x \star b \Rightarrow a = b \text{ [régularité à gauche]}$$

et

$$a \star x = b \star x \Rightarrow a = b \text{ [régularité à droite]}$$

Exemple Dans $(\mathbb{C}, +)$ tout élément est régulier,
Dans $(\mathbb{C}, \times)$ tout élément non nul est régulier alors que 0 est irrégulier.

Exemple Dans $(\mathcal{F}(E,E),\circ)$ toute permutation est régulière.
Plus généralement, une application injective est régulière à gauche alors qu'une application surjective est régulière à droite.

4.1.4.2 Élément neutre

Définition

On appelle élément neutre de $(E,\star)$ tout élément e de E vérifiant

$$\forall x \in E, e\star x = x \text{ et } x\star e = x$$

Exemple 0 est élément neutre de $(\mathbb{C},+)$.
1 est élément neutre de $(\mathbb{C},\times)$.
$\emptyset$ est élément neutre de $(\mathcal{P}(E),\cup)$.
E est élément neutre de $(\mathcal{P}(E),\cap)$.
Id_E est élément neutre de $(\mathcal{F}(E,E),\circ)$.

Proposition

Si $(E,\ \star)$ possède un élément neutre celui-ci est unique.

dém. :
Soient e,e' deux éléments neutres pour $\star$.
On a $e\star e' = e$ car e' est neutre à droite et $e\star e' = e'$ car e est neutre à gauche.
On en déduit que $e = e'$.
□

Définition

On appelle monoïde tout magma $(E,\star)$ associatif et possédant un élément neutre.
Si de plus la loi $\star$ est commutative, le monoïde $(E,\star)$ est dit commutatif.

Exemple $(\mathbb{C},+)$ est un monoïde commutatif d'élément neutre 0,
$(\mathbb{C},\times)$ est un monoïde commutatif d'élément neutre 1,
$(\mathcal{P}(E),\cup)$ est un monoïde commutatif d'élément neutre $\emptyset$,
$(\mathcal{P}(E),\cap)$ est un monoïde commutatif d'élément neutre E,
$(\mathcal{F}(E,E),\circ)$ est un monoïde d'élément neutre Id_E.

4.1.4.3 Élément symétrisable

Soit $(E,\star)$ un monoïde d'élément neutre e.

Définition

On appelle élément symétrisable de $(E,\star)$ tout élément x de E tel qu'il existe $y\in E$ pour lequel

$$x\star y = e \text{ et } y\star x = e$$

Proposition

Si x est symétrisable alors l'élément $y \in E$ vérifiant $x \star y = y \star x = e$ est unique.

dém. :
Soient y et y' solutions.
On a

$$y = y \star e = y \star (x \star y') = (y \star x) \star y' = e \star y' = y'$$

□

Définition

Si x est symétrisable, l'unique élément y de E tel que $x \star y = y \star x = e$ est appelé symétrique de x et on le note

$$\text{sym}(x)$$

Exemple Dans $(\mathbb{C}, +)$, tout x est symétrisable et $\text{sym}(x) = -x$.

Exemple Dans $(\mathbb{C}, \times)$, tout x non nul est symétrisable et $\text{sym}(x) = 1/x$.
En revanche 0 n'est pas symétrisable.

Exemple Dans $(\mathcal{F}(E, E), \circ)$ toute permutation de f est symétrisable et $\text{sym}(f) = f^{-1}$.
Inversement si $f : E \to E$ est symétrisable alors il existe $g : E \to E$ telle que $f \circ g = g \circ f = \text{Id}_E$ et par suite f est une permutation de E.

Exemple Dans $(E, \star)$, e est symétrisable et $\text{sym}(e) = e$.
En effet $e \star e = e$ et $e \star e = e$.

Proposition

Si x est symétrisable alors $\text{sym}(x)$ l'est aussi et

$$\text{sym}(\text{sym}(x)) = x$$

dém. :
Si x est symétrisable on a $x \star \text{sym}(x) = \text{sym}(x) \star x = e$ donc $\text{sym}(x)$ est symétrisable et $\text{sym}(\text{sym}(x)) = x$.
□

Proposition

Si x et y sont symétrisables alors $x \star y$ l'est aussi et

$$\text{sym}(x \star y) = \text{sym}(y) \star \text{sym}(x)$$

dém. :
Posons $z = \mathrm{sym}(y) \star \mathrm{sym}(x)$.
On a

$$(x \star y) \star z = x \star y \star \mathrm{sym}(y) \star \mathrm{sym}(x) = x \star \mathrm{sym}(x) = e$$

et de même $z \star (x \star y) = e$.
□

Attention : Il faut être attentif à l'inversion des termes lorsque la loi $\star$ n'est pas commutative !

Remarque L'ensemble des éléments symétrisables est une partie stable d'un monoïde.

Proposition

Si x est un élément symétrisable de $(E, \star)$ alors x est régulier.

dém. :
Soient $a, b \in E$. Supposons $x \star a = x \star b$.
En composant à gauche avec le symétrique de x, on obtient

$$\mathrm{sym}(x) \star x \star a = \mathrm{sym}(x) \star x \star b$$

et donc

$$e \star a = e \star b$$

i.e. $a = b$. Ainsi x est régulier à gauche et de même on obtient x régulier à droite.
□

4.1.5 Itéré d'un élément

Soit $(E, \star)$ un monoïde de neutre e.
Soit $x \in E$. Pour $n \in \mathbb{N}^\star$, on note

$$x^{\star n} = x \star x \star \ldots \star x \text{ (} n \text{ termes)}$$

Ainsi $x^{\star 1} = x$, $x^{\star 2} = x \star x$,...
De plus on pose $x^{\star 0} = e$.
Ainsi on donne un sens à $x^{\star n}$ pour $n \in \mathbb{N}$.

Définition

$x^{\star n}$ est appelé itéré d'ordre n de l'élément x.

Proposition

$\forall p, q \in \mathbb{N}, x^{\star p} \star x^{\star q} = x^{\star (p+q)}$ et $(x^{\star p})^{\star q} = x^{\star (pq)}$.

dém. :
Il suffit de dénombrer le nombre de terme x composé dans chacun des membres.
□

Attention : En général $(x \star y)^{\star p} \neq x^{\star p} \star y^{\star p}$.
En effet : $(x \star y)^{\star p} = (x \star y) \star (x \star y) \star \ldots \star (x \star y)$
et $x^{\star p} \star y^{\star p} = (x \star x \star \ldots \star x) \star (y \star y \star \ldots \star y)$.

Supposons maintenant que x est un élément symétrisable.
Pour $n \in \mathbb{N}^\star$, on note $x^{\star(-n)} = \text{sym}(x) \star \text{sym}(x) \star \ldots \star \text{sym}(x)$ (n termes)
Ainsi $x^{\star(-1)} = \text{sym}(x)$, $x^{\star(-2)} = \text{sym}(x) \star \text{sym}(x)$,...
On donne ainsi un sens à $x^{\star n}$ avec $n \in \mathbb{Z}$ lorsque x est symétrisable.

Proposition

Soit x un élément symétrisable de E.
$\forall n \in \mathbb{Z}, x^{\star n}$ est symétrisable et $\text{sym}(x^{\star n}) = x^{\star(-n)}$.
$\forall p, q \in \mathbb{Z}, x^{\star p} \star x^{\star q} = x^{\star(p+q)}$ et $(x^{\star p})^{\star q} = x^{\star(pq)}$.

dém. :
On discute selon les signes des puissances d'itération et on étudie chaque cas. . .
□

4.1.6 Structures produits

4.1.6.1 structure sur $E \times F$

Soient $(E, \top)$ et $(F, \perp)$ deux magmas.

Définition

On définit une loi de composition interne notée $\star$ sur $E \times F$ par

$$(x, y) \star (x', y') = (x \top x', y \perp y')$$

Cette loi $\star$ est appelé loi produit sur $E \times F$.

Exemple On peut définit une loi $\star$ sur $\mathbb{R}^2$ par produit des structures $(\mathbb{R}, +)$ et $(\mathbb{R}, \times)$
La loi $\star$ est alors définie par

$$(x, y) \star (x', y') = (x + x', yy')$$

Proposition

Si $(E, \top)$ et $(F, \perp)$ sont des monoïdes (resp. des monoïdes commutatifs) de neutre e et f alors $(E \times F, \star)$ est un monoïde (resp. un monoïde commutatif) d'élément neutre $\varepsilon = (e, f)$.
De plus, un élément (x, y) de $E \times F$ est symétrisable si, et seulement si, x et y le sont et alors

$$\text{sym}\left((x, y)\right) = (\text{sym}(x), \text{sym}(y))$$

dém. :
Par définition de la loi $\star$ sur $E \times F$, on a

$$((x, y) \star (x', y')) \star (x'', y'') = ((x \top x') \top x'', (y \perp y') \perp y'')$$

Par associativité des lois $\top$ et $\perp$ sur E et F, on obtient :

$$((x, y) \star (x', y')) \star (x'', y'') = (x \top (x' \top x''), y \perp (y' \perp y''))$$

Par définition de la loi $\star$ sur $E \times F$, on parvient à

$$((x, y) \star (x', y')) \star (x'', y'') = (x, y) \star ((x', y') \star (x'', y''))$$

Ainsi la loi $\star$ est associative sur $E \times F$.
Par définition de la loi $\star$ sur $E \times F$, on a

$$(x, y) \star (e, f) = (x \top e, y \perp f) = (x, y)$$

car e et f sont neutres pour les lois $\top$ et $\perp$ sur E et F.
De même on a aussi $(e, f) \star (x, y) = (x, y)$.
L'élément (e, f) est donc neutre pour la loi $\star$ sur $E \times F$.
Ainsi $(E \times F, \star)$ est un monoïde.
Si les lois $\top$ et $\perp$ sont commutatives sur E et F alors

$$(x, y) \star (x', y') = (x \top x', y \perp y') = (x' \top x, y' \perp y) = (x', y') \star (x, y)$$

et la loi $\star$ est commutative sur $E \times F$.
Enfin, si (x, y) est symétrisable dans $E \times F$ et si (x', y') désigne son symétrique alors $x \top x' = x' \top x = e$ et $y \perp y' = y' \perp y = f$ donnent x et y symétrisable et $sym(x) = x'$, $sym(y) = y'$.
Inversement, si (x, y) est un élément de $E \times F$ avec x et y symétrisables alors

$$(x, y) \star (sym(x), sym(y)) = (e, f) \text{ et } (sym(x), sym(y)) \star (x, y) = (e, f)$$

donc (x, y) est symétrisable dans $E \times F$.
$\square$

Exemple Pour la loi $\star$ définie sur $\mathbb{R}^2$ dans l'exemple ci-dessus, on obtient que $(\mathbb{R}^2, \star)$ est un monoïde commutatif de neutre $(0, 1)$ et dont les éléments symétrisables sont les (x, y) avec $y \neq 0$, de symétrique $(-x, 1/y)$.

4.1.6.2 structure sur E^n

Soient $(E, \star)$ un magma et n un naturel non nul.

Définition

On définit une loi de composition interne, encore notée $\star$, sur E^n par

$$(x_1, \ldots, x_n) \star (y_1, \ldots, y_n) = (x_1 \star y_1, \ldots, x_n \star y_n)$$

Cette loi $\star$ est appelé loi produit sur E^n.

Exemple On définit une addition sur $\mathbb{R}^2$ par structure produit de la façon suivante :

$$(x, y) + (x', y') = (x + x', y + y')$$

Exemple On définit une multiplication sur $\mathbb{R}^3$ par structure produit de la façon suivante :

$$(x, y, z) + (x', y', z') = (x + x', y + y', z + z')$$

Proposition

Si $(E, \star)$ est un monoïde (resp. un monoïde commutatif) d'élément neutre e alors $(E^n, \star)$ est un monoïde (resp. un monoïde commutatif) d'élément neutre $\varepsilon = (e, \ldots, e)$.
De plus, un élément $x = (x_1, \ldots, x_n)$ est symétrisable si, et seulement si, chaque x_i l'est, et alors

$$\text{sym}(x) = (\text{sym}(x_1), \ldots, \text{sym}(x_n))$$

dém. :
Par définition de la loi $\star$ sur E^n, on a

$$((x_1,\dots,x_n)\star(y_1,\dots,y_n))\star(z_1,\dots,z_n)=((x_1\star y_1)\star z_1,\dots,(x_n\star y_n)\star z_n)$$

Par associativité de la loi $\star$ sur E, on obtient :

$$((x_1,\dots,x_n)\star(y_1,\dots,y_n))\star(z_1,\dots,z_n)=(x_1\star(y_1\star z_1),\dots,x_n\star(y_n\star z_n))$$

Par définition de la loi $\star$ sur E^n, on parvient à

$$((x_1,\dots,x_n)\star(y_1,\dots,y_n))\star(z_1,\dots,z_n)=(x_1,\dots,x_n)\star((y_1,\dots,y_n)\star(z_1,\dots,z_n))$$

Ainsi la loi $\star$ est associative sur E^n.
Par définition de la loi $\star$ sur E^n, on a

$$(x_1,\dots,x_n)\star(e,\dots,e)=(x_1\star e,\dots,x_n\star e)=(x_1,\dots,x_n)$$

car e est neutre pour la loi $\star$ sur E.
De même on a aussi $(e,\dots,e)\star(x_1,\dots,x_n)=(x_1,\dots,x_n)$.
L'élément $\varepsilon=(e,\dots,e)$ est donc neutre pour la loi $\star$ sur E^n.
Ainsi $(E^n,\star)$ est un monoïde.
Si la loi $\star$ est commutative sur E alors

$$(x_1,\dots,x_n)\star(y_1,\dots,y_n)=(x_1\star y_1,\dots,x_n\star y_n)=(y_1\star x_1,\dots,y_n\star x_n)=(y_1,\dots,y_n)\star(x_1,\dots,x_n)$$

et la loi $\star$ est commutative sur E^n.
Enfin, si $x=(x_1,\dots,x_n)$ est symétrisable dans E^n et si $y=(y_1,\dots,y_n)$ désigne son symétrique alors $x\star y=y\star x=\varepsilon$ donne $x_i\star y_i=y_i\star x_i=e$ pour chaque $i\in\{1,\dots,n\}$ et donc chaque x_i est symétrisable et $y_i=sym(x_i)$.
Inversement, si $x=(x_1,\dots,x_n)$ avec chaque x_i symétrisable, on peut introduire $y=(sym(x_1),\dots,sym(x_n))$ et on vérifie $x\star y=y\star x=\varepsilon$ ce donne x symétrisable.
□

Exemple $(\mathbb{R}^n,+)$ et $(\mathbb{C}^n,+)$ sont des monoïdes commutatif de neutres $(0,\dots,0)$.

4.1.6.3 structure sur $\mathcal{F}(X,E)$

Soient $(E,\star)$ un magma et X un ensemble non vide.

Définition

> On définit une loi de composition interne, encore notée $\star$, sur $\mathcal{F}(X,E)$ par
>
> $$\forall x\in X,(f\star g)(x)=f(x)\star g(x)$$
>
> Cette loi $\star$ est appelé loi produit sur $\mathcal{F}(X,E)$.

Exemple Pour $X=\mathcal{D}\subset\mathbb{R}$ et $(E,\star)=(\mathbb{R},+)$ ou $(\mathbb{R},\times)$, ce qui précède définit l'addition et la multiplication sur les fonctions réelles.

Exemple Pour $X = \mathbb{N}$ et $(E, \star) = (\mathbb{R}, +)$ ou $(\mathbb{R}, \times)$, ce qui précède définit l'addition et la multiplication des suites réelles.

Proposition

Si $(E, \star)$ est un monoïde (resp. un monoïde commutatif) d'élément neutre e alors $(\mathcal{F}(X, E), \star)$ est un monoïde (resp. un monoïde commutatif) d'élément neutre $\varepsilon : x \mapsto e$.
De plus, un élément $f \in \mathcal{F}(X, E)$ est symétrisable si, et seulement si, $f(x)$ l'est pour chaque $x \in X$ et alors

$$(\text{sym} f)(x) = \text{sym}(f(x))$$

dém. :
Pour tout $x \in X$,

$$[(f \star g) \star h](x) = (f \star g)(x) \star h(x) = (f(x) \star g(x)) \star h(x)$$

Par associativité de la loi $\star$ sur E, on obtient

$$[(f \star g) \star h](x) = f(x) \star (g(x) \star h(x)) = [f \star (g \star h)](x)$$

et ainsi $(f \star g) \star h = f \star (g \star h)$.
La loi $\star$ est donc associative sur $\mathcal{F}(X, E)$.
Pour tout $x \in X$,

$$(f \star \varepsilon)(x) = f(x) \star e = f(x) \text{ et } (\varepsilon \star f)(x) = e \star f(x) = f(x)$$

Ainsi $f \star \varepsilon = \varepsilon \star f = f$ est donc ε est neutre pour la loi $\star$ sur $\mathcal{F}(X, E)$.
Ainsi $(\mathcal{F}(X, E), \star)$ est un monoïde.
Si la loi $\star$ est commutative sur E, pour tout $x \in X$,

$$(f \star g)(x) = f(x) \star g(x) = g(x) \star f(x) = (g \star f)(x)$$

et donc $f \star g = g \star f$.
Ainsi la loi $\star$ est commutative sur $\mathcal{F}(X, E)$.
Enfin, si f est un élément symétrisable de $\mathcal{F}(X, E)$ et si g désigne son symétrique alors pour tout $x \in X$,

$$(f \star g)(x) = (g \star f)(x) = \varepsilon(x)$$

donne $f(x) \star g(x) = g(x) \star f(x) = e$ et donc $f(x)$ est symétrisable dans E et $g(x)$ est son symétrique.
Inversement, si f est un élément de $\mathcal{F}(X, E)$ tel que pour chaque $x \in X$, $f(x)$ est symétrisable alors en introduisant $g : x \mapsto sym(f(x))$ on vérifie aisément $f \star g = g \star f = \varepsilon$ ce qui donne f symétrisable.
□

Exemple $(\mathcal{F}(\mathcal{D}, \mathbb{R}), +), (\mathcal{F}(\mathcal{D}, \mathbb{R}), \times), (\mathbb{R}^{\mathbb{N}}, +)$ et $(\mathbb{R}^{\mathbb{N}}, \times)$ sont des monoïdes commutatifs.

4.1.7 Notation additive et multiplicative

Définition

Un monoïde est dit noté additivement (resp. multiplicativement) si sa loi de composition interne est notée $+$ (resp. $\times$)

Attention : La notation additive n'est exploitée que pour les monoïdes commutatifs.
En revanche, la notation multiplicative ne sous entend pas la commutativité du produit : plus tard, on aura l'occasion dans le cadre matriciel de manipuler un produit non commutatif.

Lorsqu'on adopte la notation additive ou multiplicative d'un monoïde, on adopte les conventions de notations du tableau ci-dessous :

Notation par défaut	Notation additive	Notation multiplicative
$\star$	$+$	$\times$ ou .
e	0	1
$x \star y$	$x+y$	xy ou $x.y$
sym(x)	$-x$	x^{-1}
$\underset{i=1}{\overset{n}{\star}} x_i$	$\displaystyle\sum_{i=1}^{n} x_i$	$\displaystyle\prod_{i=1}^{n} x_i$
$x^{\star n}$	$n.x$	x^n

Remarque En notation additive :
Le symétrique d'un élément x est appelé opposé de x.
Les itérés additifs d'un élément sont donnés par les relations

$$\forall n \in \mathbb{N}^\star, n.x = x + x + \cdots + x \text{ (} n \text{ termes) et } 0.x = 0$$

Les propriétés calculatoires sur les itérés se relisent

$$\forall p,q \in \mathbb{N}, p.x + q.x = (p+q).x \text{ et } p.(q.x) = (pq).x$$

Si x est symétrisable, on peut introduire les itérés d'ordre négatif et en particulier

$$(-1).x = -x$$

Remarque En notation multiplicative :
Le symétrique d'un élément x est appelé inverse de x.
Les itérés multiplicatifs d'un élément sont donnés par les relations

$$\forall n \in \mathbb{N}^\star, x^n = x.x...x \text{ (} n \text{ termes) et } x^0 = 1$$

Les propriétés calculatoires sur les itérés se relisent

$$\forall p,q \in \mathbb{N}, x^p.x^q = x^{p+q} \text{ et } (x^p)^q = x^{pq}$$

Si x est inversible, on peut introduire les itérés d'ordre négatif et en particulier

$$x^{\star -1} = x^{-1}$$

Attention : L'inverse de x est noté $1/x$ seulement si $\times$ est commutative.

Attention : Si la $\times$ n'est pas commutative, faisons attention à la relation

$$(xy)^{-1} = y^{-1}x^{-1}$$

Remarque En notation par défaut, il est très fréquent, si le contexte le permet, de noter x^n au lieu de $x^{\star n}$. En particulier le symétrique de x se retrouve noté x^{-1} et l'on a la formule

$$(x \star y)^{-1} = y^{-1} \star x^{-1}$$

4.2 Groupes

4.2.1 Définition

Définition

On appelle groupe tout magma $(G, \star)$ tel que
1) $\star$ est associative ;
2) $(G, \star)$ possède un élément neutre e ;
3) tout élément de $(G, \star)$ est symétrisable.
Si de plus $\star$ est commutative, le groupe $(G, \star)$ est dit commutatif ou plus couramment abélien.

Remarque Un groupe n'est jamais vide, il contient e.

Remarque Dans un groupe tout élément est symétrisable, donc régulier.

Exemple $(\mathbb{C}, +)$ est un groupe abélien de neutre 0. En effet l'addition est commutative, associative, 0 en est élément neutre et tout élément est symétrisable dans $(\mathbb{C}, +)$.
De même $(\mathbb{R}, +)$, $(\mathbb{Q}, +)$ et $(\mathbb{Z}, +)$ sont des groupes abéliens.
En revanche $(\mathbb{N}, +)$ n'en est pas un, les naturels non nuls ne sont pas symétrisables dans $(\mathbb{N}, +)$.

Exemple $(\mathbb{C}, \times)$ n'est pas un groupe car 0 n'est pas symétrisable.
En revanche $(\mathbb{C}^\star, \times)$ est un groupe abélien de neutre 1.
De même $(\mathbb{Q}^\star, \times)$, $(\mathbb{R}^\star, \times)$ sont des groupes abéliens.

Exemple $(\mathfrak{S}(E), \circ)$ est un groupe.
En effet $\circ$ est associative, Id_E est élément neutre de $(\mathfrak{S}(E), \circ)$ et toute permutation de E est symétrisable dans $(\mathfrak{S}(E), \circ)$.

Proposition

Si $(G, \top)$ et $(G', \perp)$ sont des groupes de neutres e et e' alors $G \times G'$ muni de la loi produit $\star$ est un groupe de neutre (e, e').

dém. :
Grâce aux propriétés démontrées sur la loi produit sur $G \times G'$.
$\square$

Proposition

> Si $(G, \star)$ est un groupe de neutre e alors $(G^n, \star)$ est un groupe de neutre $(e, \dots, e)$.

dém. :
Grâce aux propriétés démontrées sur la loi produit sur G^n.
$\square$

Exemple $(\mathbb{R}^n, +)$ et $(\mathbb{C}^n, +)$ sont des groupes abéliens de neutre $(0, \dots, 0)$.

Proposition

> Si $(G, \star)$ est un groupe de neutre e alors $(\mathcal{F}(X, G), \star)$ est un groupe de neutre $x \mapsto e$.

dém. :
Grâce aux propriétés démontrées sur la loi produit sur $\mathcal{F}(X, G)$.
$\square$

Exemple $(\mathcal{F}(\mathcal{D}, \mathbb{R}), +)$ et $(\mathbb{R}^{\mathbb{N}}, +)$ sont des groupes abéliens de neutres la fonction nulle $\tilde{0} : x \mapsto 0$ et la suite nulle $(0)_{n \in \mathbb{N}}$.

Exemple L'addition définit une loi de composition interne sur l'ensemble P des vecteurs du plan.
$(P, +)$ est alors un groupe abélien de neutre le vecteur nul $\vec{o}$.
On a la même propriété en considérant les vecteurs de l'espace.

Exemple Soit $G = \mathbb{R} \setminus \{1\}$. Pour $a, b \in G$, on pose

$$a \star b = a + b - ab$$

Montrons que $(G, \star)$ est un groupe abélien.
$\star$ définit bien une loi de composition interne sur G, en effet pour $a, b \in G$, l'élément $a \star b$ existe dans $\mathbb{R}$ et puisque

$$a + b - ab = 1 \Leftrightarrow (a - 1)(1 - b) = 0$$

on peut affirmer que pour $a \neq 1$ et $b \neq 1$ on a $a + b - ab \neq 1$.
Ainsi $a \star b \in G$ et donc $\star$ associe à deux éléments de G un élément de G.
Soient $a, b \in G$.

$$b \star a = b + a - ba = a + b - ab = a \star b$$

donc la loi $\star$ est commutative.
Soient $a, b, c \in G$.

$$a \star (b \star c) = a + (b + c - b) - a\,(b + c - bc) = a + b + c - (ab + bc + ac) + abc$$

et de même

$$(a \star b) \star c) = a + b + c - (ab + bc + ac) + abc$$

donc la loi $\star$ est associative.

Soit $a \in G$, $a \star 0 = a + 0 - a \times 0 = a$ donc 0 est neutre pour la loi $\star$
Enfin, considérons $a, b \in G$.

$$a \star b = 0 \Leftrightarrow a + b - ab = 0$$

Après résolution de cette équation en l'inconnue b, on peut affirmer que pour tout $a \in G$, en posant $b = \dfrac{a}{a-1} \in G$, on a $a \star b = 0$. Ainsi tout élément de $(G, \star)$ est symétrisable.

4.2.2 Sous-groupe

4.2.2.1 Définition

Soit $(G, \star)$ un groupe d'élément neutre e.

Définition

> On appelle sous-groupe de $(G, \star)$ toute partie H de G vérifiant :
> 1) $e \in H$;
> 2) $\forall x \in H, \text{sym}(x) \in H$ [stabilité par passage au symétrique] ;
> 3) $\forall x, y \in H, x \star y \in H$ [stabilité par composition].

Exemple $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ sont des sous-groupes de $(\mathbb{C}, +)$.
$\mathbb{N}$ n'est pas un sous-groupe de $(\mathbb{C}, +)$.

Exemple $\mathbb{Q}^\star$, $\mathbb{R}^\star$, $\mathbb{R}^{+\star}$ sont des sous-groupes de $(\mathbb{C}^\star, \times)$.

Exemple G et $\{e\}$ sont des sous-groupes de $(G, \star)$.

Théorème

> Si H est un sous-groupe de $(G, \star)$ alors $(H, \star)$ est un groupe.
> Si de plus si le groupe $(G, \star)$ est abélien alors $(H, \star)$ l'est aussi.

dém. :
H est stable pour la loi $\star$, cela permet de donner un sens à $(H, \star)$ en considérant la loi obtenue par restriction de la loi sur G.
$\star$ est associative sur G donc aussi sur H.
e est élément neutre de $(G, \star)$ et $e \in H$ donc e est aussi neutre de $(H, \star)$.
Enfin, pour tout $x \in H$, comme $\text{sym}(x) \in H$ et puisque $x \star \text{sym}(x) = \text{sym}(x) \star x = e$, on peut dire que x est un élément symétrisable de $(H, \star)$.
□

Remarque Pour montrer qu'une structure est un groupe :
- soit on reconnaît la loi et alors on montre que la structure est un sous-groupe d'une structure connue ;
- soit on ne reconnaît pas la loi et on revient à la définition de groupe.

Proposition

> Soit H une partie de G.
> On a équivalence entre :
> (i) H est un sous-groupe de $(G, \star)$;
> (ii) $H \neq \emptyset$ et $\forall x, y \in H, x \star \text{sym}(y) \in H$.

dém. :
(i) $\Rightarrow$ (ii) immédiat
(ii) $\Rightarrow$ (i) Supposons $H \neq \emptyset$ et $\forall x, y \in H, x \star \text{sym}(y) \in H$.
Puisque $H \neq \emptyset$, on peut introduire un élément x dans H et on a alors $e = x \star \text{sym}(x) \in H$.
Puisque $e \in H$, pour tout $x \in H$, $\text{sym}(x) = e \star \text{sym}(x) \in H$. Enfin, pour tout $x, y \in H$, $x \star y = x \star \text{sym}(\text{sym}(y)) \in H$ car $\text{sym}(y) \in H$.
□

Exemple Pour $a \in \mathbb{R}$, on note $a\mathbb{Z} = \{ak/k \in \mathbb{Z}\}$.
Montrons que $a\mathbb{Z}$ est un sous-groupe de $(\mathbb{R}, +)$.
On a évidemment $a\mathbb{Z} \subset \mathbb{R}$.
$0 = a \times 0$ avec $0 \in \mathbb{Z}$ donc $0 \in a\mathbb{Z}$ (et par suite $a\mathbb{Z} \neq \emptyset$).
Pour $x, y \in a\mathbb{Z}$, on peut écrire $x = ak$ et $y = a\ell$ avec $k, \ell \in \mathbb{Z}$ et alors $x - y = ak - a\ell = a(k - \ell)$ avec $k - \ell \in \mathbb{Z}$ donc $x - y \in a\mathbb{Z}$.

Exemple Soit $U = \{z \in \mathbb{C}/ |z| = 1\}$.
Montrons que U est un sous-groupe de $(\mathbb{C}^\star, \times)$.
On a évidemment $U \subset \mathbb{C}^\star$.
$1 \in U$ car $|1| = 1$.
Pour $z, z' \in U$,

$$\left|zz'^{-1}\right| = \frac{|z|}{|z'|} = \frac{1}{1} = 1$$

donc $zz'^{-1} \in U$.

Exemple Soit $a \in E$ et $H = \{f \in \mathfrak{S}(E)/f(a) = a\}$.
Montrons que H est un sous-groupe de $(\mathfrak{S}(E), \circ)$.
On a évidemment $H \subset \mathfrak{S}(E)$.
$\text{Id}_E \in H$ car $\text{Id}_E(a) = a$.
Pour $f, g \in H$, $(f \circ g^{-1})(a) = (f \circ g^{-1})(g(a)) = f(a) = a$ donc $f \circ g^{-1} \in H$.

Proposition

Soient H_1, H_2 deux sous-groupes de $(G, \star)$.
$H_1 \cap H_2$ est un sous-groupe de $(G, \star)$.

dém. :
$H_1 \cap H_2 \subset G$.
$e \in H_1 \cap H_2$ car $e \in H_1$ et $e \in H_2$ puisque H_1 et H_2 sont des sous-groupes.
Pour $x, y \in H_1 \cap H_2$, $x \star y^{-1} \in H_1 \cap H_2$ car $x \star y^{-1} \in H_1$ puisque H_1 est un sous-groupe et de même $x \star y^{-1} \in H_2$.
□

Remarque Ce résultat est faux pour l'union :
Pour $(G, \star) = (\mathbb{C}^\star, \times)$, $H_1 = \mathbb{R}^\star$, $H_2 = U$ et $H = H_1 \cup H_2$, on a $2, i \in H$ et $2i \notin H$.
Ainsi H n'est pas stable pour $\times$ et ce n'est donc pas un sous-groupe.

4.2.2.2 Groupe des racines n-ième de l'unité

Soit $n \in \mathbb{N}^\star$ et $U_n = \{z \in \mathbb{C}/z^n = 1\}$.
Rappelons $U_n = \{\omega_0, \omega_1, \dots, \omega_{n-1}\}$ avec $\omega_k = \mathrm{e}^{2ik\pi/n}$.

Proposition

$(U_n, \times)$ est un groupe abélien.

dém. :
Montrons que U_n est un sous-groupe de $(\mathbb{C}^\star, \times)$.
$U_n \subset \mathbb{C}^\star$.
$1 \in U_n$ car $1^n = 1$.
Pour $z, z' \in U_n$, $(zz'^{-1})^n = z^n \dfrac{1}{z'^n} = 1$ donc $zz'^{-1} \in U_n$.
□

Exemple Pour $n = 1$, $U_1 = \{1\}$.
Pour $n = 2$, $U_2 = \{1, -1\}$.

$\times$	1	-1
1	1	-1
-1	-1	1

Pour $n = 3$, $U_3 = \left\{1, j, j^2\right\}$.

$\times$	1	j	j^2
1	1	j	j^2
j	j	j^2	1
j^2	j^2	1	j

Pour $n = 4$, $U_4 = \{1, i, -1, -i\}$.

$\times$	1	i	-1	$-i$
1	1	i	-1	$-i$
i	i	-1	$-i$	1
-1	-1	$-i$	1	i
$-i$	$-i$	1	i	-1

4.2.2.3 Groupes géométriques

On note :
- $\mathcal{P}$ le plan géométrique de direction P ;
- $t_{\vec{u}}$ la translation de vecteur $\vec{u}$;
- $H_{O,\lambda}$ l'homothétie de centre O et de rapport λ ;
- $\mathrm{Rot}_{O,\theta}$ la rotation de centre O et d'angle θ.

Proposition

$\mathcal{T} = \{t_{\vec{u}} / \vec{u} \in P\}$ est un sous-groupe de $(\mathfrak{S}(\mathcal{P}), \circ)$.

dém. :
$\mathcal{T} \subset \mathfrak{S}(\mathcal{P})$, $\mathrm{Id}_{\mathcal{P}} = t_{\vec{0}}$ et pour $\vec{u}, \vec{v} \in P$, $t_{\vec{u}} \circ (t_{\vec{v}})^{-1} = t_{\vec{u}-\vec{v}} \in \mathcal{T}$.
□

Proposition

Pour $O \in \mathcal{P}$, $\mathcal{H}_O = \{H_{O,\lambda} / \lambda \in \mathbb{R}^\star\}$ est un sous-groupe de $(\mathfrak{S}(\mathcal{P}), \circ)$.

dém. :
$\mathcal{H}_0 \subset \mathfrak{S}(\mathcal{P})$, $\mathrm{Id}_{\mathcal{P}} = H_{O,1}$ et pour $\lambda, \mu \neq 0$, $H_{O,\lambda} \circ (H_{O,\mu})^{-1} = H_{O,\lambda/\mu}$ avec $\lambda/\mu \neq 0$ donc $H_{O,\lambda} \circ (H_{O,\mu})^{-1} \in \mathcal{H}_O$.
□

Proposition

> Pour $O \in \mathcal{P}$, $\mathcal{R}_O = \{\mathrm{Rot}_{O,\theta}/\theta \in \mathbb{R}\}$ est un sous-groupe de $(\mathfrak{S}(\mathcal{P}), \circ)$.

dém. :
$\mathcal{R}_O \subset \mathfrak{S}(\mathcal{P})$, $\mathrm{Id}_{\mathcal{P}} = \mathrm{Rot}_{O,0}$ et pour $\theta, \theta' \in \mathbb{R}$, $\mathrm{Rot}_{O,\theta} \circ (\mathrm{Rot}_{O,\theta'})^{-1} = \mathrm{Rot}_{O,\theta-\theta'} \in \mathcal{R}_O$.
On appelle isométrie du plan toute permutation $f \in \mathfrak{S}(\mathcal{P})$ telle que $\forall A, B \in \mathcal{P}, d(f(A), f(B)) = d(AB)$
□

Proposition

> L'ensemble $\mathcal{I}$ des isométries du plan est un sous-groupe de $(\mathfrak{S}(\mathcal{P}), \circ)$.

dém. :
$\mathcal{I} \subset \mathfrak{S}(\mathcal{P})$, $\mathrm{Id}_{\mathcal{P}}$ est évidemment une isométrie et pour $f, g \in \mathcal{I}$, on a pour tout $A, B \in \mathcal{P}$, $d((f \circ g^{-1})(A), (f \circ g^{-1})(B)) = d((f \circ g^{-1})(g(A)), (f \circ g^{-1})(g(B)))$ car g est une isométrie et ainsi $d((f \circ g^{-1})(A), (f \circ g^{-1})(B)) = d(f(A), f(B) = d(A, B)$ car f est une isométrie.
Ainsi $f \circ g^{-1}$ est une isométrie.
□

4.2.3 Morphisme de groupes

Soit $(G, \star)$, $(G', \top)$ et $(G'', \perp)$ trois groupes d'éléments neutres e, e' et e''.

4.2.3.1 Définition

Définition

> On appelle morphisme du groupe $(G, \star)$ vers $(G', \top)$ toute application $\varphi : G \to G'$ vérifiant :
>
> $$\forall x, y \in G, f(x \star y) = f(x) \top f(y)$$
>
> Si f est bijective, on dite que f est un isomorphisme.
> Si $(G', \top) = (G, \star)$, on dit que f est un endomorphisme.
> Si $(G', \top) = (G, \star)$ et f est bijective on dit que f est un automorphisme.

Exemple $\ln$ est un isomorphisme de $(\mathbb{R}^{+\star}, \times)$ vers $(\mathbb{R}, +)$.
En effet, pour tout $a, b > 0$, $\ln(ab) = \ln(a) + \ln(b)$.

Exemple $\exp$ est un morphisme de $(\mathbb{C}, +)$ vers $(\mathbb{C}^\star, \times)$.
En effet, pour tout $z, z' \in \mathbb{C}$, $\exp(z + z') = \exp(z)\exp(z')$.

Exemple Id_G est un automorphisme de $(G, \star)$.

Exemple L'application constante $f : G \to G$ définie par $f(x) = e$ est un endomorphisme de $(G, \star)$.

Proposition

Soit a un élément d'un groupe $(G, \star)$.
L'application $\varphi : \mathbb{Z} \to G$ définie par $\varphi(n) = a^{\star n}$ est un morphisme de groupes.
En effet $\varphi(n+p) = a^{\star(n+p)} = a^{\star n} \star a^{\star p} = \varphi(n) \star \varphi(p)$.

4.2.3.2 Propriétés

Proposition

Pour $f : G \to G'$ un morphisme de groupes, on a

$$f(e) = e'$$

$$\forall x \in G, f(\text{sym}(x)) = \text{sym}(f(x)),$$

$$\forall x_1, \ldots, x_n \in G, f(\mathop{\star}_{i=1}^{n} x_i) = \mathop{\top}_{i=1}^{n} f(x_i) \text{ et}$$

$$\forall x \in G, \forall p \in \mathbb{Z}, f(x^{\star p}) = (f(x))^{\top p}$$

dém. :
D'une part $f(e \star e) = f(e)$ et d'autre part $f(e \star e) = f(e) \top f(e)$ donc

$$f(e) \top f(e) = f(e) = f(e) \top e'$$

d'où $f(e) = e'$.
On a

$$f(x) \top f(\text{sym}(x)) = f(x \star \text{sym}(x)) = f(e) = e'$$

En composant cette relation à droite avec $\text{sym}(f(x))$, on obtient $f(\text{sym}(x)) = \text{sym}(f(x))$.
Par récurrence sur $n \in \mathbb{N}^\star$, on montre facilement

$$f(\mathop{\star}_{i=1}^{n} x_i) = \mathop{\top}_{i=1}^{n} f(x_i)$$

Pour $p = 0$, $f(x^{\star 0}) = f(e) = e' = (f(x))^{\top 0}$.
Pour $p \in \mathbb{N}^\star$, $f(x^{\star p}) = f(\mathop{\star}_{i=1}^{p} x) = \mathop{\top}_{i=1}^{p} f(x) = f(x)^{\top p}$.
Pour $p \in \mathbb{Z}^{\star -}$, on peut écrire $p = -n$ avec $n \in \mathbb{N}^\star$ et on a
$f(x^{\star p}) = f\left(\text{sym}(x^{\star n})\right) = \text{sym}\left(f(x^{\star n})\right) = \text{sym}\left(f(x)^{\top n}\right) = f(x)^{\top p}$.
□

Proposition

Si $f : G \to G'$ et $g : G' \to G''$ sont deux morphismes de groupes alors $g \circ f : G \to G''$ est aussi un morphisme de groupes.

dém. :
Pour $x, y \in G$,

$$(g \circ f)(x \star y) = g\left(f(x \star y)\right) = g\left(f(x) \top f(y)\right) = g(f(x)) \perp g(f(y)) = (g \circ f)(x) \perp (g \circ f)(y)$$

□

Remarque On peut en particulier souligner que la composée de deux endomorphismes de groupe (resp. isomorphismes, automorphismes) en est un.

Proposition

Si $f : G \to G'$ est un isomorphisme de groupes alors $f^{-1} : G' \to G$ l'est aussi.

dém. :
Pour tout $x', y' \in G'$, il existe $x, y \in G$ tel que $f(x) = x'$ et $f(y) = y'$.
On a alors

$$f^{-1}(x' \top y') = f^{-1}(f(x) \top f(y)) = f^{-1}(f(x \star y)) = x \star y = f^{-1}(x') \star f^{-1}(y')$$

Ainsi f^{-1} est un morphisme de groupes et il est de plus bien connu que f^{-1} est bijective.
□

Remarque En particulier l'application réciproque d'un automorphisme en est un.

4.2.3.3 Noyau et image

Proposition

Soit $f : G \to G'$ un morphisme de groupes.
Si H est un sous-groupe de $(G, \star)$ alors $f(H)$ est un sous-groupe de $(G', \top)$.
Si H' est un sous-groupe de $(G', \top)$ alors $f^{-1}(H)$ est un sous-groupe de $(G, \star)$.

dém. :
Soit H un sous-groupe de $(G, \star)$.
$f(H) = \{f(x)/x \in H\}$ est une partie de G'.
D'une part $e' \in f(H)$ car $e' = f(e)$ avec $e \in H$.
D'autre part, pour $x', y' \in f(H)$, on peut écrire $x' = f(x)$ et $y' = f(y)$ avec $x, y \in H$ et $x' \top y'^{-1} = f(x \star y^{-1}) \in f(H)$ car $x \star y^{-1} \in H$.
Ainsi $f(H)$ est un sous-groupe de $(G', \top)$.
Soit H' un sous-groupe de $(G', \top)$.
$f^{-1}(H') = \{x \in G/f(x) \in H'\}$ est une partie de G.
D'une part $e \in f^{-1}(H')$ car $f(e) = e' \in H'$.
D'autre part, pour $x, y \in f^{-1}(H')$, on a $f(x \star y^{-1}) = f(x) \top f(y)^{-1} \in H'$ car $f(x), f(x') \in H'$.
Ainsi $f^{-1}(H')$ est un sous-groupe de $(G, \star)$.
□

Définition

Soit $f : G \to G'$ un morphisme de groupes.
On appelle image de f, l'ensemble $\text{Im} f = f(G)$. C'est un sous-groupe de $(G', \top)$.
On appelle noyau de f, l'ensemble $\ker f = f^{-1}(\{e'\})$. C'est un sous-groupe de $(G, \star)$.

Remarque Pour déterminer $\text{Im} f$ on étudie les valeurs prises par f.
Pour déterminer $\ker f$, on résout l'équation $f(x) = e'$ d'inconnue $x \in G$.

Exemple Soit $f : \mathbb{C}^\star \to \mathbb{C}^\star$ définie par $f(z) = |z|$.
f est un morphisme de groupes multiplicatifs pour lequel $\text{Im} f = \mathbb{R}^{+\star}$ et $\ker f = U$.

Exemple Soit $f : \mathbb{R} \to \mathbb{C}^\star$ définie par $f(\theta) = \mathrm{e}^{i\theta}$.
f est un morphisme du groupe $(\mathbb{R}, +)$ vers $(\mathbb{C}^\star, \times)$ pour lequel $\text{Im} f = U$ et $\ker f = 2\pi\mathbb{Z}$.

Théorème

> Soit $f : G \to G'$ un morphisme de groupes.
> f est surjective si, et seulement si, $\mathrm{Im} f = G'$,
> f est injective si, et seulement si, $\ker f = \{e\}$.

dém. :
La première propriété immédiate par définition de la surjectivité.
Pour la deuxième propriété, étudions les deux implications.
Supposons f injective.
Puisque $f(e) = e'$ et puisque e' ne peut avoir d'autres antécédents que e à cause de l'injectivité de f, on peut affirmer $\ker f = \{e\}$.
Inversement, supposons $\ker f = \{e\}$.
Soient $x, y \in G$. Si $f(x) = f(y)$ alors $f(x) \top f(y)^{-1} = e'$ et donc $f(x \star y^{-1}) = e'$. Ainsi $x \star y^{-1} \in \ker f = \{e\}$ et donc $x \star y^{-1} = e$ ce qui entraîne $x = y$.
Ainsi f est injective.
□

4.2.3.4 Quelques morphismes géométriques

$\mathcal{P}$ désigne le plan géométrique

Proposition

> L'application $\vec{u} \mapsto t_{\vec{u}}$ est un morphisme de $(P, +)$ vers $(\mathfrak{S}(\mathcal{P}), \circ)$ d'image $\mathcal{T}$ et de noyau $\left\{\vec{0}\right\}$.

dém. :
$t_{\vec{u}} \circ t_{\vec{v}} = t_{\vec{u}+\vec{v}}$.
□

Proposition

> Pour $O \in \mathcal{P}$, l'application $\lambda \mapsto H_{O,\lambda}$ est un morphisme de $(\mathbb{R}^\star, \times)$ vers $(\mathfrak{S}(\mathcal{P}), \circ)$ d'image $\mathcal{H}_O$ et de noyau $\{1\}$.

dém. :
$H_{O,\lambda} \circ H_{O,\mu} = H_{O,\lambda\mu}$.
□

Proposition

> Pour $O \in \mathcal{P}$, l'application $\theta \mapsto \mathrm{Rot}_{O,\theta}$ est un morphisme de $(\mathbb{R}, +)$ vers $(\mathfrak{S}(\mathcal{P}), \circ)$ d'image $\mathcal{R}_O$ et de noyau $2\pi\mathbb{Z}$.

dém. :
$\mathrm{Rot}_{O,\theta} \circ \mathrm{Rot}_{O,\theta'} = \mathrm{Rot}_{O,\theta+\theta'}$.
□

4.3 Etude du groupe symétrique

4.3.1 Permutation de $\mathbb{N}_n = \{1, 2, ..., n\}$

Définition

> Pour $n \in \mathbb{N}^\star$, on note $\mathfrak{S}_n$ l'ensemble des permutations de $\mathbb{N}_n$.
> $(\mathfrak{S}_n, \circ)$ est un groupe d'élément neutre $\mathrm{Id}_{\mathbb{N}_n} = \mathrm{Id}$ appelé groupe symétrique d'ordre n.

Remarque $\mathfrak{S}_n$ est un groupe fini à $n!$ éléments.

Définition

Pour $\sigma \in \mathfrak{S}_n$, on note
$$\sigma = \begin{pmatrix} 1 & 2 & ... & n \\ \sigma(1) & \sigma(2) & ... & \sigma(n) \end{pmatrix}$$
pour visualiser l'action de σ.

Exemple Dans $(\mathfrak{S}_6, \circ)$ considérons la permutation
$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 3 & 2 & 4 & 1 & 5 \end{pmatrix}$$
Pour celle-ci 4 est point fixe de σ et
$$\sigma^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 3 & 2 & 4 & 6 & 1 \end{pmatrix}, \sigma^2 = \sigma \circ \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 2 & 3 & 4 & 6 & 1 \end{pmatrix}$$

Remarque Pour $n = 1$: $\mathfrak{S}_1 = \{\text{Id}\}$. $(\mathfrak{S}_1, \circ)$ est un groupe abélien.
Pour $n = 2$: $\mathfrak{S}_2 = \{\text{Id}, \tau\}$ où
$$\tau = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$
$(\mathfrak{S}_2, \circ)$ est groupe abélien.

Proposition

Pour $n \geqslant 3$ le groupe $(\mathfrak{S}_n, \circ)$ n'est pas commutatif.

dém. :
Soient
$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & ... & n \\ 2 & 1 & 3 & 4 & ... & n \end{pmatrix} \text{ et } \sigma' = \begin{pmatrix} 1 & 2 & 3 & 4 & ... & n \\ 3 & 2 & 1 & 4 & ... & n \end{pmatrix}$$
éléments de $\mathfrak{S}_n$.
$(\sigma \circ \sigma')(1) = 3$ et $(\sigma' \circ \sigma)(1) = 2$ donc $\sigma \circ \sigma' \neq \sigma' \circ \sigma$.
Ainsi le groupe $(\mathfrak{S}_n, \circ)$ n'est pas commutatif.
□

4.3.2 Cycles

Soient $p \in \mathbb{N}$ tel que $2 \leqslant p \leqslant n$ et $a_1, ..., a_p$ une liste de p éléments deux à deux distincts de $\mathbb{N}_n$.
Soit $c : \mathbb{N}_n \to \mathbb{N}_n$ définie par :
$$c(a_1) = a_2, c(a_2) = a_3, ..., c(a_{p-1}) = a_p, c(a_p) = a_1$$
et
$$\forall x \in \mathbb{N}_n \setminus \{a_1, ..., a_p\}, c(x) = x$$
L'application c est une permutation de $\mathbb{N}_n$.

Définition

La permutation c est appelée cycle de longueur p (ou p-cycle).
Ce cycle est noté

$$c = \begin{pmatrix} a_1 & a_2 & \dots & a_p \end{pmatrix}$$

L'ensemble $S = \{a_1, ..., a_p\}$ est appelé support du cycle c.

Exemple Dans $(\mathfrak{S}_6, \circ)$, pour $c = \begin{pmatrix} 2 & 4 & 3 & 5 \end{pmatrix}$, on a

$$c = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 4 & 5 & 3 & 2 & 6 \end{pmatrix}$$

Remarque Si $c = \begin{pmatrix} a_1 & a_2 & \dots & a_p \end{pmatrix}$ alors $c^{-1} = \begin{pmatrix} a_p & \dots & a_2 & a_1 \end{pmatrix}$.

Remarque La description d'un cycle n'est pas unique :
Si $c = \begin{pmatrix} a_1 & a_2 & \dots & a_p \end{pmatrix}$ alors $c = \begin{pmatrix} a_2 & a_3 & \dots & a_p & a_1 \end{pmatrix}$, $c = \begin{pmatrix} a_3 & \dots & a_p & a_1 & a_2 \end{pmatrix}$,...,
$c = \begin{pmatrix} a_p & a_1 & a_2 & \dots & a_{p-1} \end{pmatrix}$.

Définition

Les cycles de longueur 2 sont appelés transpositions.
Une transposition $\tau = \begin{pmatrix} i & j \end{pmatrix}$ a pour effet d'échanger i et j.

Remarque $\begin{pmatrix} i & j \end{pmatrix} = \begin{pmatrix} j & i \end{pmatrix}$ donc toute transposition peut être visualisée sous la forme $\begin{pmatrix} i & j \end{pmatrix}$ avec $1 \leqslant i < j \leqslant n$.

Remarque Si τ est une transposition alors $\tau^2 = \text{Id}$ et $\tau^{-1} = \tau$.
Plus généralement

Proposition

Si c est un cycle de longueur p alors $c^p = \text{Id}$ et $c^{-1} = c^{p-1}$.

dém. :
Soit $c = \begin{pmatrix} a_1 & a_2 & \dots & a_p \end{pmatrix}$ un p cycle (avec des $a_1, ..., a_p$ deux à deux distincts).
$c^0(a_1) = a_1, c(a_1) = a_2, c^2(a_1) = a_3, ..., c^{p-1}(a_1) = a_p$ donc pour tout $1 \leqslant k \leqslant p-1, c^k(a_1) = a_{k+1}$ et par suite $c^p(a_1) = c(a_p) = a_1$.
Pour $2 \leqslant k \leqslant p, c^p(a_k) = c^p(c^{k-1}(a_1)) = c^{p+k-1}(a_1) = c^{k-1}(c^p(a_1)) = c^{k-1}(a_1) = a_k$
Enfin pour $x \in \mathbb{N}_n \setminus \{a_1, ..., a_p\}, c(x) = x$ donc $c^p(x) = x$.
Finalement $c^p = \text{Id}$.
□

Remarque Pour $n = 3$, on peut décrire les éléments de $\mathfrak{S}_3$

$$\mathfrak{S}_3 = \left\{\text{Id}, \begin{pmatrix} 1 & 2 \end{pmatrix}, \begin{pmatrix} 2 & 3 \end{pmatrix}, \begin{pmatrix} 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 3 & 2 & 1 \end{pmatrix}\right\}$$

Pour $n \geqslant 4$, il existe dans $\mathfrak{S}_n$ des éléments qui ne sont pas des cycles.

4.3.3 Décomposition d'une permutation en produit de transpositions

Proposition

Tout cycle de longueur p peut se décomposer en un produit de $p-1$ transpositions.

dém. :
$p=2$: ok $p=3$: $\begin{pmatrix} a_1 & a_2 & a_3 \end{pmatrix} = \begin{pmatrix} a_1 & a_2 \end{pmatrix} \circ \begin{pmatrix} a_2 & a_3 \end{pmatrix}$ $p=4$: $\begin{pmatrix} a_1 & a_2 & a_3 & a_4 \end{pmatrix} = \begin{pmatrix} a_1 & a_2 \end{pmatrix} \circ \begin{pmatrix} a_2 & a_3 \end{pmatrix} \circ \begin{pmatrix} a_3 & a_4 \end{pmatrix}$
et plus généralement :

$$\begin{pmatrix} a_1 & a_2 & \dots & a_p \end{pmatrix} = \begin{pmatrix} a_1 & a_2 \end{pmatrix} \circ \begin{pmatrix} a_2 & a_3 \end{pmatrix} \circ \dots \circ \begin{pmatrix} a_{p-1} & a_4 \end{pmatrix}$$

□

Théorème

Toute permutation de $\mathbb{N}_n$ peut se décomposer en un produit d'au plus $n-1$ transpositions.

dém. :
Démontrons la propriété par récurrence sur $n \in \mathbb{N}^\star$.
Pour $n=1$: ok
Supposons la propriété établie au rang $n \geqslant 1$.
Soit $\sigma \in \mathfrak{S}_{n+1}$ et posons $k = \sigma(n+1)$.
Si $k = n+1$ alors $\sigma_{\restriction \mathbb{N}_n} \in \mathfrak{S}_n$ et peut s'écrire comme produit d'au plus $n-1$ transpositions éléments de $\mathfrak{S}_n$. Cette décomposition permet aussi d'écrire σ comme produit de $n-1 \leqslant n$ transpositions éléments de $\mathfrak{S}_{n+1}$.
Si $k \neq n+1$ alors considérons $\sigma' = \begin{pmatrix} k+1 & n+1 \end{pmatrix} \circ \sigma$. On a $\sigma'(n+1) = n+1$ et comme ci-dessus, σ' peut s'écrire comme produit d'au plus $n-1$ transpositions éléments de $\mathfrak{S}_{n+1}$. Puisque $\sigma = \begin{pmatrix} k+1 & n+1 \end{pmatrix} \circ \sigma'$, σ s'écrit comme produit d'au plus n transpositions éléments de $\mathfrak{S}_{n+1}$.
□

Proposition

Toute permutation de $\mathbb{N}_n$ peut se décomposer en un produit de transpositions de la forme $\begin{pmatrix} 1 & k \end{pmatrix}$ avec $2 \leqslant k \leqslant n$.

dém. :
Il suffit de savoir décomposer une transposition $\begin{pmatrix} i & j \end{pmatrix}$ pour conclure.
Si $i=1$ ou $j=1$: ok
Sinon $\begin{pmatrix} i & j \end{pmatrix} = \begin{pmatrix} 1 & i \end{pmatrix} \circ \begin{pmatrix} 1 & j \end{pmatrix} \circ \begin{pmatrix} 1 & i \end{pmatrix}$.
□

4.3.4 Signature d'une permutation

Définition

Soient $\sigma \in \mathfrak{S}_n$ et un couple (i,j) avec $1 \leqslant i < j \leqslant n$.
On dit σ réalise une inversion sur le couple (i,j) si $\sigma(i) > \sigma(j)$.
On note $I(\sigma)$ le nombre de couples (i,j) (avec $1 \leqslant i < j \leqslant n$) sur lesquels σ réalise une inversion.

Exemple $I(\mathrm{Id}) = 0$.

Exemple Dans $(\mathfrak{S}_4, \circ)$ pour

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix}$$

$I(\sigma) = 4$.
Ici les couples sur lesquels σ opère une inversion sont $(1,4), (2,3), (2,4)$ et $(3,4)$.

Exemple Dans $(\mathfrak{S}_8, \circ)$ considérons

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 4 & 7 & 5 & 2 & 1 & 8 & 6 \end{pmatrix}$$

Pour déterminer $I(\sigma)$ on compte, pour chaque terme de la seconde ligne, le nombre de termes qui le suivent et qui lui sont inférieurs.
Ici $I(\sigma) = 2 + 2 + 4 + 2 + 1 + 0 + 1 + 0 = 12$.

Définition

On appelle signature d'une permutation σ de $\mathfrak{S}_n$ le réel $\varepsilon(\sigma) = (-1)^{I(\sigma)}$.

Exemple $\varepsilon(\mathrm{Id}) = 1$.

Exemple Dans $(\mathfrak{S}_n, \circ)$,

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & \dots & n \\ n & n-1 & n-2 & \dots & 1 \end{pmatrix}$$

$$I(\sigma) = (n-1) + (n-2) + \cdots + 1 = \frac{n(n-1)}{2} \text{ et } \varepsilon(\sigma) = (-1)^{\frac{n(n-1)}{2}}$$

Proposition

La signature d'une transposition vaut -1.

dém. :
Soit $\tau = \begin{pmatrix} i & j \end{pmatrix}$ avec $1 \leqslant i < j \leqslant n$ une transposition.

$$\tau = \begin{pmatrix} 1 & 2 & \dots & i & \dots & j & \dots & n \\ 1 & 2 & \dots & j & \dots & i & \dots & n \end{pmatrix}$$

$$I(\tau) = \underset{1}{0} + \cdots + \underset{i-1}{0} + \underset{i}{(j-i)} + \underset{i+1}{1} + \cdots + \underset{j-1}{1} + \underset{j}{0} + \cdots + \underset{n}{0} = 2(j-i) - 1$$

Donc $\varepsilon(\tau) = -1$.
□

Théorème

L'application $\varepsilon : \mathfrak{S}_n \to \{-1, 1\}$ est un morphisme du groupe $(\mathfrak{S}_n, \circ)$ sur $(\{-1, 1\}, \times)$.

dém. :
Soient σ et σ' deux éléments de $\mathfrak{S}_n$.
Soit (i, j) un couple avec $1 \leqslant i < j \leqslant n$.

$\sigma' \circ \sigma$ réalise une inversion sur (i,j) si, et seulement si :
σ réalise une inversion sur (i,j) et σ' ne réalise pas d'inversion sur $(\sigma(j),\sigma(i))$
ou bien
σ ne réalise pas d'inversion sur (i,j) mais σ' en réalise une sur $(\sigma(i),\sigma(j))$.
En sommant $I(\sigma)$ et $I(\sigma')$ on dénombre les cas précédemment étudiés ainsi que deux fois les cas où il y a à la fois inversion sur (i,j) et sur $(\sigma(j),\sigma(i))$.
Par suite $I(\sigma' \circ \sigma)$ à même parité que $I(\sigma)+I(\sigma')$, ainsi

$$\varepsilon(\sigma' \circ \sigma) = (-1)^{I(\sigma' \circ \sigma)} = (-1)^{I(\sigma)+I(\sigma')} = \varepsilon(\sigma')\varepsilon(\sigma)$$

On peut donc conclure que ε est un morphisme de groupes.
□

Corollaire

$\varepsilon(\sigma_1 \circ \cdots \circ \sigma_p) = \varepsilon(\sigma_1) \times \cdots \times \varepsilon(\sigma_p)$.
$\forall p \in \mathbb{Z}, \varepsilon(\sigma^p) = \varepsilon(\sigma)^p$ et en particulier $\varepsilon(\sigma^{-1}) = \varepsilon(\sigma)$.

Proposition

La signature d'un p-cycle est $(-1)^{p-1}$.

dém. :
Car un p-cycle s'écrit comme produit de $p-1$ transpositions, chacune de signature -1.
□

Définition

Une permutation de signature 1 est dite paire.
Une permutation de signature -1 est dite impaire.
On note $\mathfrak{A}_n$ l'ensemble des permutations paires de $\mathfrak{S}_n$.

Remarque Une permutation paire (resp. impaire) se décompose en un nombre pair (resp. impair) de transpositions.

Proposition

$\mathfrak{A}_n$ est un sous-groupe de $(\mathfrak{S}_n,\circ)$.

dém. :
C'est le noyau du morphisme signature, c'est donc un sous-groupe.
□

Définition

$(\mathfrak{A}_n,\circ)$ est appelé groupe alterné d'ordre n.

Remarque Pour $n=1$: $\mathfrak{A}_n = \{\mathrm{Id}\}$.
Pour $n=2$: $\mathfrak{A}_n = \{\mathrm{Id}\}$.
Pour $n=3$: $\mathfrak{A}_n = \left\{\mathrm{Id}, \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 3 & 2 \end{pmatrix}\right\}$.

Proposition

Pour $n \geqslant 2$,

$$\mathrm{Card}\mathfrak{A}_n = n!/2$$

dém. :
$\mathfrak{S}_n$ est la réunion disjointe des parties $\mathfrak{A}_n$ et $\mathfrak{S}_n \backslash \mathfrak{A}_n$.
Or ces dernières sont en bijection via l'application $\sigma \mapsto \tau \circ \sigma$ où τ désigne une certaine transposition.
On en déduit $\text{Card}\mathfrak{A}_n = \text{Card}\mathfrak{S}_n \backslash \mathfrak{A}_n$ puis $\text{Card}\mathfrak{S}_n = 2\text{Card}\mathfrak{A}_n$.
□

4.3.5 Hors programme : Décomposition d'une permutation en produit de cycles

Lemme

Pour tout $\sigma \in \mathfrak{S}_n$ et tout $i \in \{1, 2, ..., n\}$ il existe $p \in \mathbb{N}^\star$ tel que :
$i, \sigma(i), ..., \sigma^{p-1}(i)$ soient deux à deux distincts et $\sigma^p(i) = i$.

dém. :
L'application $\varphi : \mathbb{N} \to \{1, 2, ..., n\}$ définie par $\varphi(k) = \sigma^k(i)$ ne peut-être injective car $\mathbb{N}$ est un ensemble infini.
Par suite, il existe $k \neq \ell \in \mathbb{N}$ tels que $\sigma^k(i) = \sigma^\ell(i)$ et on peut, quitte à échanger, supposer $k < \ell$.
On a alors $\sigma^{\ell-k}(i) = \sigma^{-k} \circ \sigma^\ell(i) = i$ avec $\ell - k \in \mathbb{N}^\star$.
Considérons maintenant l'ensemble $A = \{q \in \mathbb{N}^\star / \sigma^q(i) = i\}$.
D'après l'étude précédente on peut affirmer que A est une partie non vide de $\mathbb{N}^\star$ et elle possède donc un plus petit élément p.
Pour celui-ci on a $\sigma^p(i) = i$. De plus les éléments $i, \sigma(i), ..., \sigma^{p-1}(i)$ sont nécessairement deux à deux distincts.
En effet s'il existait $0 \leqslant k < \ell \leqslant p - 1$ tels que $\sigma^k(i) = \sigma^\ell(i)$ alors on aurait $\sigma^{\ell-k}(i) = i$ avec $0 < \ell - k < p$ ce qui contredit la minimalité de p.
□

Théorème

Toute permutation σ de $\mathfrak{S}_n$ (avec $n \geqslant 2$) se décompose en un produit de cycles de supports disjoints.
De plus cette décomposition est unique à l'ordre près des facteurs.

dém. :
Existence :
Procédons par récurrence sur $n \geqslant 2$.
Pour $n = 2$, les permutations éléments de $\mathfrak{S}_2$ sont Id et $(1, 2)$ ce qui permet de conclure.
Supposons l'existence établie au rang $n \geqslant 2$.
Soit $\sigma \in \mathfrak{S}_{n+1}$.
Cas $\sigma(n + 1) = n + 1$
Considérons la restriction σ' de σ au départ de $\{1, 2, ..., n\}$.
Cette restriction σ' réalise une permutation de $\{1, 2, ..., n\}$ car σ est une permutation de $\{1, 2, ..., n + 1\}$ qui laisse $n + 1$ fixe. En appliquant l'hypothèse de récurrence à σ' on peut décomposer cette dernière en produit de cycles de $\mathfrak{S}_n$. Cette décomposition de σ' dans $\mathfrak{S}_n$ se prolonge en une décomposition de σ en cycles de $\mathfrak{S}_{n+1}$ ce qui résout le problème posé.
Cas $\sigma(n + 1) \neq n + 1$.
Considérons $p \in \mathbb{N}^\star$ tel que les éléments $n + 1, \sigma(n + 1), ..., \sigma^{p-1}(n + 1)$ soient deux à deux distincts et $\sigma^p(n + 1) = n + 1$. Formons le cycle $c = (n + 1, \sigma(n + 1), ..., \sigma^{p-1}(n + 1))$.
Par construction, la permutation $c^{-1} \circ \sigma$ laisse invariants les éléments $n + 1, \sigma(n + 1), ..., \sigma^{p-1}(n + 1)$.
Comme $n+1$ est invariant, on peut, par l'étude précédente, décomposer $c^{-1} \circ \sigma$ en $c_1 \circ c_2 \circ ... \circ c_r$ produit de cycles de supports disjoints. De plus les supports de $c_1, c_2, ..., c_r$ sont disjoints de $\left\{n + 1, \sigma(n + 1), \ldots, \sigma^{p-1}(n + 1)\right\}$ car ces éléments sont invariants par $c^{-1} \circ \sigma$. Par suite l'écriture

$\sigma = c \circ c_1 \circ c_2 \circ ... \circ c_r$ apparaît alors comme étant une décomposition de c en produit de cycles de supports disjoints.
Récurrence établie
Unicité à l'ordre près des facteurs :
Supposons que σ présente deux décompositions de la forme voulue :
$\sigma = c_1 \circ c_2 \circ ... \circ c_r$ et $\sigma = d_1 \circ d_2 \circ ... \circ d_s$.
Quitte à échanger, on peut supposer $r \geqslant s$.
Si $r = 0$ alors $s = 0$.
Sinon, considérons i un élément du support de c_r.
Par le lemme, il existe $p \in \mathbb{N}^\star$ tel que $i, \sigma(i), ..., \sigma^{p-1}(i)$ soient deux à deux distincts et $\sigma^p(i) = i$.
Or $\sigma(i) = c_r(i), \sigma^2(i) = c_r^2(i), ..., \sigma^{p-1}(i) = c_r^{p-1}(i)$ et $\sigma^p(i) = c_r^p(i) = i$ ceci car les éléments $i, c_r(i), c_r^2(i), ..., c_r^{p-1}(i)$ évoluent dans le support du cycle c_r.
Par suite $c_r = (i, \sigma(i), \sigma^2(i), ..., \sigma^{p-1}(i))$.
Comme i est modifié par $\sigma = d_1 \circ d_2 \circ ... \circ d_s$, il existe t entre 1 et s tel que i appartient au support de d_t.
Quitte à réorganiser l'ordre de la décomposition $\sigma = d_1 \circ d_2 \circ ... \circ d_s$ on peut supposer $t = s$.
En suivant la même démarche que ci-dessus, on obtient $d_s = (i, \sigma(i), \sigma^2(i), ..., \sigma^{p-1}(i))$ puis $d_s = c_r$.
On peut alors simplifier ce facteur commun afin d'obtenir $c_1 \circ c_2 \circ \cdots \circ c_{r-1} = d_1 \circ d_2 \circ \cdots \circ d_{s-1}$
Il suffit alors de réitérer ce processus pour conclure.
□

4.4 Anneaux

4.4.1 Définition

Définition

Soient $\top$ et $\star$ deux lois de composition internes sur un ensemble E.
On dit que $\top$ est distributive sur $\star$ si

$$\forall a, b, c \in E : a \star (b \top c) = (a \star b) \top (a \star c) \text{ [distributivité à gauche]}$$

et

$$(b \top c) \star a = (b \star a) \top (c \star a) \text{ [distributivité à droite]}$$

Exemple Dans $\mathbb{C}$, $\times$ est distributive sur $+$.
Dans $\mathcal{P}(E)$, $\cup$ est distributive sur $\cap$ et inversement.

Définition

On appelle anneau tout triplet $(A, \top, \star)$ formé d'un ensemble A et de deux lois de composition internes $\top$ et $\star$ tels que :
1) $(A, \top)$ est un groupe abélien ;
2) $(A, \star)$ est un monoïde ;
3) $\star$ est distributive sur $\top$.
Si de plus $\star$ est commutative, l'anneau $(A, \top, \star)$ est dit commutatif.

Remarque Les lois $\top$ et $\star$ sont généralement notées $+$ et $\times$.
Leurs neutres sont quant à eux notés 0_A et 1_A.
De plus pour l'évaluation de $a + b \times c$ on donne priorité à la multiplication.

Exemple $(\mathbb{Z},+,\times)$, $(\mathbb{Q},+,\times)$, $(\mathbb{R},+,\times)$ et $(\mathbb{C},+,\times)$ sont des anneaux commutatifs.

Exemple Si $A=\{0\}$ alors $(A,+,\times)$ est un anneau appelé anneau nul.

Proposition

Si $(A,+,\times)$ est un anneau et $n\in\mathbb{N}^\star$ alors $(A^n,+,\times)$ est un anneau de neutres $0_{A^n}=(0_A,\dots,0_A)$ et $1_{A^n}=(1_A,\dots,1_A)$.

dém. :
On sait déjà que $(A^n,+)$ est un groupe abélien de neutre $0_{A^n}=(0_A,\dots,0_A)$ et $(A^n,\times)$ un monoïde de neutre $1_{A^n}=(1_A,\dots,1_A)$. Il ne reste qu'à vérifier la propriété de distributivité ce qui est assez immédiat.
□

Exemple $(\mathbb{R}^2,+,\times)$ est un anneau commutatif.
Dans celui-ci rappelons les opérations :

$$(x,y)+(x',y')=(x+x',y+y') \text{ et } (x,y)\times(x',y')=(xx',yy')$$

Exemple $(\mathbb{Z}^n,+,\times)$, $(\mathbb{R}^n,+,\times)$ et $(\mathbb{C}^n,+,\times)$ sont des anneaux commutatifs

Proposition

Si $(A,+,\times)$ est un anneau et X un ensemble alors $(\mathcal{F}(X,A),+,\times)$ est un anneau de neutres la fonction nulle constante égale à 0_A et la fonction constante égale à 1_A.

dém. :
On sait déjà que $(\mathcal{F}(X,A),+)$ est un groupe abélien de neutre $\tilde{0}: x\mapsto 0_A$ et $(\mathcal{F}(X,A),\times)$ un monoïde de neutre $\tilde{1}: x\mapsto 1_A$. Il ne reste qu'à vérifier la propriété de distributivité ce qui est assez immédiat.
□

Exemple $(\mathcal{F}(X,\mathbb{R}),+,\times)$ est un anneau commutatif.
En particulier $(\mathbb{R}^{\mathbb{N}},+,\times)$ et $(\mathcal{F}(\mathcal{D},\mathbb{R}),+,\times)$.

4.4.2 Sous-anneau

Définition

On appelle sous-anneau d'un anneau $(A,+,\times)$ toute partie B incluse dans A telle que :
1) $1_A\in B$;
2) $\forall x,y\in B, x-y\in B$;
3) $\forall x,y\in B, xy\in B$.

Exemple $\mathbb{Z}$ est un sous-anneau de $(\mathbb{R},+,\times)$.

Théorème

Si B est un sous-anneau de $(A, +, \times)$ alors $(B, +, \times)$ est un anneau.
Si de plus $(A, +, \times)$ est commutatif alors $(B, +, \times)$ l'est aussi.

dém. :
Par 1) et 2) on obtient que $(B, +)$ est un groupe abélien.
Par 3) B est stable pour $\times$ et on peut donc donner un sens à $(B, \times)$.
$\times$ est associative sur A donc elle l'est aussi sur B.
1_A est neutre pour $\times$ et $1_A \in B$ donc $(B, \times)$ possède un neutre.
$\times$ est distributive sur $+$ sur A donc aussi sur B.
Finalement $(B, +, \times)$ est un anneau.
$\square$

Exemple Considérons

$$\mathbb{Z}\left[\sqrt{2}\right] = \left\{a + b\sqrt{2}/a, b \in \mathbb{Z}\right\}$$

et montrons que $(\mathbb{Z}\left[\sqrt{2}\right], +, \times)$ est un anneau commutatif.
Montrons que $\mathbb{Z}\left[\sqrt{2}\right]$ un sous-anneau de l'anneau commutatif $(\mathbb{R}, +, \times)$.
On a évidemment $\mathbb{Z}\left[\sqrt{2}\right] \subset \mathbb{R}$.
$1 = 1 + 0.\sqrt{2} \in \mathbb{Z}\left[\sqrt{2}\right]$.
Pour $x, y \in \mathbb{Z}\left[\sqrt{2}\right]$, on peut écrire $x = a + b\sqrt{2}$ et $y = c + d\sqrt{2}$ avec $a, b, c, d \in \mathbb{Z}$.
On a

$$x - y = (a - c) + (b - d)\sqrt{2} \in \mathbb{Z}\left[\sqrt{2}\right]$$

car $a - c, b - d \in \mathbb{Z}$
et

$$xy = (ac + 2bd) + (ad + bc)\sqrt{2} \in \mathbb{Z}\left[\sqrt{2}\right]$$

Ainsi $\mathbb{Z}\left[\sqrt{2}\right]$ est un sous-anneau de $(\mathbb{R}, +, \times)$ et donc $(\mathbb{Z}\left[\sqrt{2}\right], +, \times)$ est un anneau commutatif.

Exemple On note $\mathcal{C}$ l'ensemble des suites réelles convergentes.
Montrons que $\mathcal{C}$ est un sous-anneau de $(\mathbb{R}^{\mathbb{N}}, +, \times)$.
On a évidemment $\mathcal{C} \subset \mathbb{R}^{\mathbb{N}}$, la suite constante égale à 1 est convergente et la différence et le produit de deux suites convergentes est convergente.

Exemple Soit $\mathcal{D}$ une partie de $\mathbb{R}$.
Montrer que $\mathcal{C}(\mathcal{D}, \mathbb{R})$ est un sous-anneau de $(\mathcal{F}(\mathcal{D}, \mathbb{R}), +, \times)$.
On a évidemment $\mathcal{C}(\mathcal{D}, \mathbb{R}) \subset \mathcal{F}(\mathcal{D}, \mathbb{R})$, la fonction constante égale à 1 est convergente et la différence et le produit de deux fonctions continues est continue.

4.4.3 Règles de calculs dans un anneau

Soit $(A, +, \times)$ un anneau de neutres 0_A et 1_A.

Proposition

$\forall a \in A, 0_A \times a = a \times 0_A = 0_A$.

dém. :
$0_A \times a = (0_A + 0_A) \times a = 0_A \times a + 0_A \times a$.
En ajoutant de part et d'autre, l'opposé de l'élément $0_A \times a$, on obtient $0_A = 0_A \times a$.
De même, on démontre $a \times 0_A = 0_A$
□

Remarque Si les neutres additifs et multiplicatifs de l'anneau A sont égaux (i.e. $0_A = 1_A$) alors la neutralité de 1_A donne $\forall x \in A, x = x \times 1_A = x \times 0_A = 0_A$. Par suite $A = \{0_A\}$ est l'anneau nul.

Proposition

$\forall a, b \in A, (-a)b = -(ab) = a(-b)$.

dém. :
$(-a)b + ab = (-a + a)b = 0.b = 0$.
En ajoutant l'opposé de ab de part et d'autre, on obtient $(-a)b = -(ab)$.
De même, on obtient $a(-b) = -(ab)$.
□

Proposition

$\forall a_1, \ldots, a_n \in A, \forall b \in B, (a_1 + \cdots + a_n)b = a_1 b + \cdots + a_n b$,
$\forall a \in A, \forall b_1, \ldots, b_n \in A, a(b_1 + \cdots + b_n) = ab_1 + \cdots + ab_n$.

dém. :
Par récurrence sur $n \in \mathbb{N}^\star$…
□

Proposition

$\forall a, b \in A, \forall p \in \mathbb{Z}, (p.a)b = p.(ab) = a(p.b)$.

dém. :
Rappelons que $p.a$ désigne l'itéré additif d'ordre p de l'élément a.
Pour $p = 0$: c'est immédiat.
Pour $p \in \mathbb{N}^\star$, $(p.a)b = (a + \cdots + a)b = ab + \cdots + ab = p.(ab)$.
Pour $p \in \mathbb{Z}^{\star -}$, on peut écrire $p = -n$ avec $n \in \mathbb{N}^\star$ et $(p.a)b = (-n.a)b = -(n.a)b = -n.(ab) = p.(ab)$.
De même, on obtient $a(p.b) = p.(ab)$.
□

Attention : Dans un anneau A non commutatif, $(ab)^n \neq a^n b^n$ en général.
En effet $(ab)^n$ se comprend $(ab)(ab) \ldots (ab)$.

Proposition

$\forall a, b \in A, (a + b)^2 = a^2 + ab + ba + b^2$
et $(a + b)^3 = a^3 + a^2 b + aba + ba^2 + ab^2 + bab + ba^2 + b^3$.

dém. :
Il suffit de développer les produits...
□

Remarque Si a et b commutent (i.e. $ab = ba$) alors $(a+b)^2 = a^2 + 2ab + b^2$ et $(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$.

Théorème

Soient $a, b \in A$ tels que a et b commutent.

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

dém. :
Par récurrence sur $n \in \mathbb{N}$.
Pour $n = 0$, la propriété est immédiate sachant que $\forall a \in A, a^0 = 1_A$ (même pour $a = 0_A$)
Supposons la propriété établie au rang $n \geqslant 0$.

$$(a+b)^{n+1} = (a+b)^n(a+b)$$

Par hypothèse de récurrence

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

Par décalage d'indice, on peut réécrire la deuxième somme

$$\sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^{k+1} = \sum_{k=1}^{n+1} \binom{n}{k-1} a^{n+1-k} b^k$$

En adjoignant un terme nul à chacune des deux sommes

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

En combinant les deux sommes en une seule

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

en vertu de la formule du triangle de Pascal.
Récurrence établie.
□

Exemple Comme 1 et a commutent

$$(1+a)^n = \sum_{k=0}^{n} \binom{n}{k} a^k$$

Théorème

Soient $a, b \in A$ tels que a et b commutent.

$$\forall n \in \mathbb{N}^\star, a^n - b^n = (a-b)\sum_{k=0}^{n-1} a^{n-1-k}b^k = (a-b)(a^{n-1} + a^{n-2}b + ... + ab^{n-2} + b^{n-1})$$

dém. :
En développant

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

En procédant à un changement de d'indice sur la deuxième somme

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

On en déduit

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

On simplifie les portions communes des deux sommes

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

□

Exemple Comme 1_A et a commutent

$$1_A - a^n = (1_A - a)(1_A + a + a^2 + ... + a^{n-1})$$

4.4.4 Eléments inversibles

Définition

Un élément $a \in A$ est dit inversible s'il est symétrisable pour $\times$ i.e. s'il existe $b \in A$ tel que $ab = ba = 1_A$.
Cet élément b est alors unique, on l'appelle inverse de a, on le note a^{-1}.

Exemple Les inversibles de $(\mathbb{C}, +, \times)$ sont les éléments non nuls.
Les inversibles de $(\mathbb{Z}, +, \times)$ sont 1 et -1.

Exemple 1_A est inversible.
On verra que 0_A n'est pas inversible.

Exemple Si A n'est pas l'anneau nul alors $0_A \neq 1_A$ et puisque pour tout $a \in A$, $0_A \times a = 0_A$, l'élément 0_A n'est pas inversible.

Exemple Les inversibles de $(\mathbb{R}^2, +, \times)$ sont les $a = (x, y)$ avec $x \neq 0$ et $y \neq 0$.

Proposition

Si x est inversible alors x^{-1} est inversible et $(x^{-1})^{-1} = x$.

dém. :
C'est une propriété déjà connu en terme d'éléments symétrisables.
□

Proposition

Si x et y sont inversibles alors xy est inversible et $(xy)^{-1} = y^{-1}x^{-1}$.

dém. :
C'est une propriété déjà connue en terme d'éléments symétrisables.
□

4.4.5 Diviseurs de zéro

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

4.4.5.1 Définition

Attention : On a $a = 0_A$ ou $b = 0_A \Rightarrow ab = 0_A$.
En revanche la réciproque n'est pas vraie dans tous les anneaux.

Exemple Dans $(\mathbb{R}^2, +, \times)$, pour $a = (1, 0)$ et $b = (0, 1)$, on a $ab = 0_{\mathbb{R}^2}$ alors que $a, b \neq 0_{\mathbb{R}^2}$.

Définition

On appelle diviseurs de zéro dans l'anneau $(A, +, \times)$ tous éléments $a, b \in A$ vérifiant $ab = 0_A$ avec $a, b \neq 0_A$. On dit ici que a est un diviseur de zéro à gauche et b in diviseur de zéro à droite.

Remarque Par cette définition, on ne considère pas que 0_A soit un diviseur de zéro.

Exemple Dans $(\mathbb{R}^2, +, \times)$ les diviseurs de zéros sont les $(x, 0)$ et $(0, x)$ avec $x \neq 0$.

Proposition

Un diviseur de zéro est non régulier pour $\times$.

dém. :
Si a est diviseur de zéro à gauche alors il existe $b \in A \backslash \{0_A\}$ tel que $ab = 0_A$.
On a alors $ab = a0_A$ et $b \neq 0_A$ donc a n'est pas régulier à gauche...
□

Proposition

Les éléments inversibles de A ne sont pas diviseurs de zéro.

dém. :
Tout élément inversible est régulier.
□

4.4.5.2 Anneau sans diviseurs de zéro

Exemple $(\mathbb{Z}, +, \times), (\mathbb{Q}, +, \times), (\mathbb{R}, +, \times)$ et $(\mathbb{C}, +, \times)$ n'ont pas de diviseurs de zéro.

Exemple $(\mathbb{R}^2, +, \times)$ possède des diviseurs de zéro.

Proposition

Si $(A, +, \times)$ ne possède pas de diviseurs de zéro alors
$\forall a, b \in A, ab = 0_A \Rightarrow a = 0_A$ ou $b = 0_A$ [implication d'intégrité]

Exemple Dans un anneau $(A, +, \times)$ sans diviseurs de zéro, l'équation $x^2 = 1_A$ possède uniquement deux solutions 1_A et -1_A.
En effet

$$x^2 = 1_A \Leftrightarrow (x - 1_A)(x + 1_A) = 0_A \Leftrightarrow x = 1_A \text{ ou } x = -1_A$$

Dans $(\mathbb{R}^2, +, \times)$ l'équation $x^2 = 1_{\mathbb{R}^2}$ possède quatre solutions :

$$(1, 1), (1, -1), (-1, 1) \text{ et } (-1, -1)$$

Proposition

Dans un anneau $(A, +, \times)$ sans diviseurs de zéro tout élément non nul est régulier.

dém. :
Soient $a \neq 0_A$ et $b, c \in A$.
Si $ab = ac$ alors $ab - ac = 0_A$ puis $a(b - c) = 0_A$.
Par l'implication d'intégrité, sachant $a \neq 0_A$, on obtient $b = c$.
Ainsi a est régulier à gauche. De même on obtient la régularité à droite.
□

4.4.5.3 Idempotent et nilpotent

Définition

Un élément $a \in A$ est dit idempotent si $a^2 = a$.

Exemple Dans un anneau sans diviseurs de zéro, seuls 0 et 1 sont idempotents.

Exemple Dans $(\mathbb{R}^2, +, \times)$, $(1, 0)$ et $(0, 1)$ sont aussi idempotents.

Définition

> Un élément $a \in A$ est dit nilpotent s'il existe $n \in \mathbb{N}^\star$ tel que $a^n = 0_A$.

Exemple Dans un anneau sans diviseurs de zéro, seul 0_A est nilpotent.

Exemple Si a est nilpotent alors $1 - a$ est inversible.
En effet, il existe $n \in \mathbb{N}^\star$ tel que $a^n = 0_A$ et alors

$$1_A = 1_A - a^n = (1_A - a)(1_A + a + a^2 + \cdots + a^{n-1})$$

Aussi

$$(1_A + a + a^2 + \cdots + a^{n-1})(1_A - a) = 1_A$$

donc $1_A - a$ est inversible et

$$(1_A - a)^{-1} = 1_A + a + a^2 + \cdots + a^{n-1}$$

4.5 Corps

4.5.1 Définition

Définition

> On appelle corps tout anneau commutatif $(K, +, \times)$ non réduit à $\{0_K\}$ dont tous les éléments, sauf 0_K, sont inversibles.

Exemple $(\mathbb{C}, +, \times)$, $(\mathbb{R}, +, \times)$ et $(\mathbb{Q}, +, \times)$ sont des corps (fameux).

Exemple Soit $\mathbb{F}_2 = \{\bar{0}, \bar{1}\}$.
On définit les opérations $+$ et $\times$ par :

$$\begin{array}{|c|cc|}\hline + & \bar{0} & \bar{1} \\ \hline \bar{0} & \bar{0} & \bar{1} \\ \bar{1} & \bar{1} & \bar{0} \\ \hline \end{array} \text{ et } \begin{array}{|c|cc|}\hline \times & \bar{0} & \bar{1} \\ \hline \bar{0} & \bar{0} & \bar{0} \\ \bar{1} & \bar{0} & \bar{1} \\ \hline \end{array}$$

$(\mathbb{F}_2, +, \times)$ est un corps.

Exemple Soit $\mathbb{F}_3 = \{\bar{0}, \bar{1}, \bar{2}\}$.
On définit les opérations $+$ et $\times$ par :

$$\begin{array}{|c|ccc|}\hline + & \bar{0} & \bar{1} & \bar{2} \\ \hline \bar{0} & \bar{0} & \bar{1} & \bar{2} \\ \bar{1} & \bar{1} & \bar{2} & \bar{0} \\ \bar{2} & \bar{2} & \bar{0} & \bar{1} \\ \hline \end{array} \text{ et } \begin{array}{|c|ccc|}\hline \times & \bar{0} & \bar{1} & \bar{2} \\ \hline \bar{0} & \bar{0} & \bar{0} & \bar{0} \\ \bar{1} & \bar{0} & \bar{1} & \bar{2} \\ \bar{2} & \bar{0} & \bar{2} & \bar{1} \\ \hline \end{array}$$

Proposition

> Un corps n'a pas de diviseurs de zéro.

dém. :
Un corps ne possède pas de diviseurs de zéro car tout élément non nul y est inversible donc régulier.
□

4.5.2 Sous-corps

Soit $(K, +, \times)$ un corps.

Définition

On appelle sous-corps d'un $(K, +, \times)$ toute partie L de K telle que :
1) L est un sous-anneau de $(K, +, \times)$;
2) $\forall x \in L \setminus \{0_K\}, x^{-1} \in L$.

Exemple $\mathbb{Q}$ est un sous-corps de $(\mathbb{R}, +, \times)$.

Théorème

Si L est un sous-corps de $(K, +, \times)$ alors $(L, +, \times)$ est un corps.

dém. :
Puisque L est un sous-anneau de l'anneau commutatif $(K, +, \times)$, on peut affirmer que $(L, +, \times)$ est un anneau commutatif. Puisque $1_K \in L$, on peut affirmer que l'anneau $(L, +, \times)$ n'est pas réduit à 0. Enfin puisque l'inverse d'un élément non nul de L est élément de L, on peut affirmer que tout élément non nul de l'anneau L est inversible dans celui-ci.
□

Exemple Considérons $\mathbb{Q}[i] = \{a + ib/a, b \in \mathbb{Q}\}$. Montrons que $(\mathbb{Q}[i], +, \times)$ est un corps.
Pour cela montrons que $\mathbb{Q}[i]$ est un sous-corps du corps $(\mathbb{C}, +, \times)$.
On a évidemment $\mathbb{Q}[i] \subset \mathbb{C}$.
$1 = 1 + i \times 0 \in \mathbb{Q}[i]$.
Pour $x, y \in \mathbb{Q}[i]$, on peut écrire $x = a + ib$ et $y = c + id$ avec $a, b, c, d \in \mathbb{Q}$.
On a alors

$$x - y = (a - c) + i(b - d) \in \mathbb{Q}[i]$$

et

$$xy = (ab - dc) + i(ad + bc) \in \mathbb{Q}[i]$$

Enfin, si $x \neq 0$,

$$x^{-1} = \frac{1}{a + ib} = \frac{a - ib}{(a + ib)(a - ib)} = \frac{a}{a^2 + b^2} - i\frac{b}{a^2 + b^2} \in \mathbb{Q}[i]$$

car $\frac{a}{a^2 + b^2}, -\frac{b}{a^2 + b^2} \in \mathbb{Q}$.

Chapitre 5

Arithmétique dans Z

5.1 Divisibilité

5.1.1 Divisibilité dans Z

Définition

Soient $a, b \in \mathbb{Z}$. On dit que a divise b s'il existe $k \in \mathbb{Z}$ tel que $b = ak$.
On note alors $a \mid b$ et on dit que a est un diviseur de b et que b est un multiple de a.

Exemple 2 divise 6 mais 2 ne divise pas 3.

Exemple $1, a, -1, -a$ divisent a.

Exemple $\forall a \in \mathbb{Z}, a \mid 0$ et $0 \mid a \Rightarrow a = 0$.

Proposition

$\forall a, b \in \mathbb{Z}, a \mid b \Rightarrow (-a) \mid b, a \mid (-b)$ et $(-a) \mid (-b)$.

dém. :
Si a divise b alors il existe $k \in \mathbb{Z}$ tel que $b = ak$ et alors $b = (-a) \times (-k)$, $-b = a \times (-k)$ et $-b = -a \times k$.
□

Remarque Par la deuxième propriété, on voit que le signe des entiers n'influe pas dans la relation de divisibilité. Cela permet de se ramener systématiquement au cadre des entiers naturels.

Proposition

$\forall a, b \in \mathbb{Z}$ avec $b \neq 0$. $a \mid b \Rightarrow |a| \leqslant |b|$.

dém. :
Si $a \mid b$ alors il existe $k \in \mathbb{Z}$ tel que $b = ak$
Puisque $b \neq 0$ on a $k \neq 0$ et donc $|k| \geqslant 1$ d'où $|b| = |ak| \geqslant |a|$.
□

Définition

Pour $a \in \mathbb{Z}$, on note $\text{Div}(a)$ l'ensemble des diviseurs de a et $\text{Mul}(a)$ l'ensemble des multiples de a. Ainsi

$$\text{Div}(a) = \{k \in \mathbb{Z}/k \mid a\} \text{ et } \text{Mul}(a) = \{ak/k \in \mathbb{Z}\} = a\mathbb{Z}$$

Remarque Pour $a \neq 0$, $\text{Div}(a) \subset [\![-|a|, |a|]\!]$ et donc $\text{Div}(a)$ est un ensemble fini.

Exemple $\text{Div}(6) = \{-6, -3, -2, -1, 1, 2, 3, 6\}$, $\text{Div}(1) = \{1, -1\}$ et $\text{Div}(0) = \mathbb{Z}$.

Proposition

$\forall a, b, c, d \in \mathbb{Z}$.
$a \mid b$ et $b \mid c \Rightarrow a \mid c$, $a \mid b$ et $b \mid a \Rightarrow |a| = |b|$,
$a \mid b$ et $a \mid c \Rightarrow a \mid (b + c)$, $a \mid b$ et $c \mid d \Rightarrow ac \mid bd$,
$a \mid b \Rightarrow \forall p \in \mathbb{N}, a^p \mid b^p$.

Exemple Soient $a \in \mathbb{Z}$ et $d \in \mathbb{N}$. Montrons

$$d \mid a \text{ et } d \mid (a^2 + a + 1) \Rightarrow d = 1$$

Puisque d divise a, d divise aussi $a^2 + a = a(a + 1)$ donc d divise encore $1 = (a^2 + a + 1) - (a^2 + a)$. Or le seul naturel divisant 1 est 1 donc $d = 1$.

Exemple Déterminons les $x \in \mathbb{Z}$ tels que $(x - 2) \mid (x + 2)$.
Si x est solution alors $x \neq 2$ et

$$\frac{x+2}{x-2} = 1 + \frac{4}{x-2} \in \mathbb{Z}$$

donc $x - 2 \in \text{Div}(4)$.
On peut alors décrire l'ensemble solution

$$\mathcal{S} = \{3, 4, 6, 1, 0, -2\}$$

Exemple Résolvons dans $\mathbb{Z}^2$ l'équation $xy + 1 = 3x + y$.

$$xy + 1 = 3x + y \Leftrightarrow (x - 1)(y - 3) = 2$$

et $\text{Div}(2) = \{-2, -1, 1, 2\}$.
On peut alors décrire l'ensemble des couples (x, y) solutions

$$\mathcal{S} = (1, 3) + \{(-2, -1), (-1, -2), (1, 2), (2, 1)\}$$

5.1.2 Division euclidienne

Théorème

Pour tout $a \in \mathbb{Z}$ et $b \in \mathbb{N}^\star$ il existe un unique couple (q, r) vérifiant

$$a = bq + r \text{ et } 0 \leqslant r < b$$

q et r sont respectivement appelés quotient et reste de la division euclidienne de a par b.
dém. :
Unicité : Soient (q, r) et (q', r') deux couples solutions. On a $b(q-q') = r'-r$ et donc $-b < b(q'-q) < b$ puis $-1 < q' - q < 1$ d'où $q = q'$ puis $r = r'$.
Existence : $q = E(a/b)$ et $r = a - bq$ conviennent

Exemple Pour $a = 16, b = 3$: $q = 5, r = 1$.
Pour $a = 23, b = 6$: $q = 3, r = 5$.
Pour $a = 12, b = 3$: $q = 4, r = 0$.
Pour $a = 5, b = 9$: $q = 0, r = 5$.
Pour $a = -12, b = 5$: $q = -3, r = 3$.
Pour $a = -7, b = 10$: $q = -1, r = 3$.

Exemple Pour $a = 2^{n+1} + 1, b = 2$: $q = 2^n, r = 1$
Pour $a = 2^{n+1} + 3, b = 2$: $q = 2^n + 1, r = 1$
Pour $a = 2^{n+1} - 1, b = 2$. $q = 2^n - 1, r = 1$.

Attention : Pour identifier q et r il faut observer l'identité $a = bq + r$ mais aussi l'encadrement $0 \leqslant r < b$

Proposition

Soit $a \in \mathbb{Z}$ et $b \in \mathbb{N}^\star$, on a équivalence entre :
(i) $b \mid a$;
(ii) le reste de la division euclidienne de a par b est nul.

dém. :
(i) $\Rightarrow$ (ii) : Si $b \mid a$ alors il existe $k \in \mathbb{Z}$ tel que $a = bk$
On peut alors écrire $a = bq + r$ avec $0 \leqslant r < b$ en prenant $q = k$ et $r = 0$.
(ii) $\Rightarrow$ (i) : Si (ii) alors la division euclidienne de a par b s'écrit $a = bq + 0$ d'où $b \mid a$.
□

5.1.3 Calculs en congruence

Soit $n \in \mathbb{N}^\star$.

Définition

Soient $a, b \in \mathbb{Z}$. On dit que a est congru à b modulo n si n divise $b - a$. On note alors

$$a = b \quad [n]$$

Ainsi

$$a = b \quad [n] \Leftrightarrow \exists k \in \mathbb{Z}, a = b + k.n$$

Exemple $7 = 2 \quad [5], 13 = 3 \quad [5], 20 = 0 \quad [5]$.

Remarque $n \mid a \Leftrightarrow a = 0 \quad [n]$

Proposition

Pour tout $a \in \mathbb{Z}$, il existe un unique $r \in \{0, 1, \ldots, n-1\}$ tel que $a = r \quad [n]$.
De plus r correspond au reste de la division euclidienne de a par n.

dém. :
Existence : Réaliser la division euclidienne de a par n.
Unicité : Si r, r' sont deux solutions alors $r = r' \quad [n]$ donc il existe $k \in \mathbb{Z}$ tel que $r = r' + kn$.
Par suite $r - r' = kn$, or $-n < r - r' < n$, donc $k = 0$ puis $r = r'$.
□

Proposition

$\forall a, b, c \in \mathbb{Z}$
$a = b \quad [n] \Leftrightarrow b = a \quad [n]$
$a = b \quad [n]$ et $b = c \quad [n] \Rightarrow a = c \quad [n]$.

dém. :
$a = b \quad [n] \Leftrightarrow b = a \quad [n]$ car $n \mid (b - a) \Leftrightarrow n \mid (a - b)$.
et $a = b \quad [n]$ et $b = c \quad [n] \Rightarrow a = c \quad [n]$ car $n \mid (b - a)$ et $n \mid (c - b) \Rightarrow n \mid (c - a)$
□

Proposition

Soient $a, a', b, b' \in \mathbb{Z}$, tels que $a = a' \quad [n]$ et $b = b' \quad [n]$.
On a $a + b = a' + b' \quad [n]$, $ab = a'b' \quad [n]$, $-a = -a' \quad [n]$ et $\forall p \in \mathbb{N}, a^p = a'^p \quad [n]$.

Exemple On a $1513^5 + 1514^4 = 3^5 + (-1)^4 = 4 \quad [5]$.

Exemple Montrons

$$\forall n \in \mathbb{N}, 5 \mid 2^{2n+1} + 3^{2n+1}$$

On a

$$2^{2n+1} + 3^{2n+1} = 2.4^n + 3.4^n = 0.4^n = 0 \quad [5]$$

Exemple Montrons

$$\forall n \in \mathbb{N}, 9 \mid (4^n - 1 + 6n)$$

(1ère méthode) On peut procéder par récurrence.
La propriété est vrai pour $n = 0$ car $9 \mid 0$.
Supposons la propriété vraie au rang $n \geqslant 0$.

$$4^{n+1} - 1 + 6(n+1) = 4.4^n + 6n + 5$$

Par hypothèse de récurrence, on peut écrire $4^n = (9k + 1 - 6n)$ avec $k \in \mathbb{Z}$, donc

$$4^{n+1} - 1 + 6(n+1) = 4.(9k + 1 - 6n) + 6n + 5 = 36k - 18n + 9$$

et finalement $9 \mid \left(4^{n+1} - 1 + 6n\right)$.
Récurrence établie
(2ème méthide) On peut factoriser

$$4^n - 1 = (4-1)(1 + 4 + \cdots + 4^{n-1})$$

On a alors

$$4^n - 1 + 6n = 3\left(1 + 4 + \cdots + 4^{n-1} + 2n\right)$$

Or

$$\left(1 + 4 + \cdots + 4^{n-1} + 2n\right) = n + 2n = 0 \quad [3]$$

donc

$$3 \mid \left(1 + 4 + \cdots + 4^{n-1} + 2n\right)$$

puis

$$9 \mid (4^n - 1 + 6n)$$

Exemple Considérons $n \in \mathbb{N}^\star$ s'écrivant en nombre décimal à l'aide des chiffres $c_m, \dots, c_0$.
On a

$$n = c_m 10^m + \cdots + c_0 10^0$$

Or $10 = 1 \quad [9], 10^2 = 1 \quad [9], \dots, 10^m = 1 \quad [9]$.
Donc

$$n = c_m + \cdots + c_1 + c_0 \quad [9]$$

Par exemple $n = 1234567 = 28 = 10 = 1 \quad [9]$.
Réaliser une preuve par 9 pour vérifier la validité d'un calcul sur les entiers, consiste à exploiter ce qui précède pour réaliser à nouveau ce calcul modulo 9.
Par exemple, supposons avoir obtenu $123 \times 456 = 67158$.
On reproduit le calcul modulo 9 : $6 \times 6 = 36 = 0 \quad [9]$ et on vérifie $67158 = 0 \quad [9]$.

Exemple En exploitant l'écriture décimal d'un nombre, on peut énoncer les critères de divisibilités suivants :
- un nombre est divisible par 9 si, et seulement si, la somme de ses chiffres l'est ;
- un nombre est divisible par 3 si, et seulement si, la somme de ses chiffres l'est ;
- un nombre est divisible par 10 si, et seulement si, son dernier chiffre est un 0 ;
- un nombre est divisible par 5 si, et seulement si, son dernier chiffre est un 0 ou un 5 ;

- un nombre est divisible par 2 si, et seulement si, son dernier chiffre est divisible par 2 ;
- un nombre est divisible par 4 si, et seulement si, ses deux derniers chiffres donne un nombre divisible par 4 ;
- un nombre est divisible par 11 si, et seulement si, la somme alternée de ses chiffre est divisible par 11. Par exemple 6567 est divisible par 11 car $6 - 5 + 6 - 7 = 0$ divisible par 11.

5.1.4 Exponentiation rapide

Soient $x \in \mathbb{R}$ et $n \in \mathbb{N}^\star$. On désire calculer x^n.
De façon élémentaire

$$x^n = x \times x \times \cdots \times x$$

ce qui fournit $n - 1$ multiplications
On peut être plus efficace de la façon suivante :
On décompose n en somme de puissances de 2.
Pour $n = 53$ on a $n = 32 + 16 + 4 + 1$.
On calcule ensuite

$$x^2 = x.x, x^4 = x^2.x^2, x^8 = x^4.x^4, x^{16} = x^8.x^8 \text{ puis } x^{32} = x^{16}.x^{16}$$

et enfin

$$x^{53} = x^{32}.x^{16}.x^4.x$$

Ainsi x^{53} est déterminé en 8 multiplications.
Mettons en place l'algorithme sous-jacent.
Pour décomposer n en somme de puissance de 2 :
- à la main : on recherche la plus grand puissance de 2 inférieure àn, on la retire de n et on recommence avec le nombre obtenu ;
- avec l'ordinateur : on procède à l'envers :
si n est pair on écrit $n = 2m$,
sinon on écrit $n = 1 + 2m$.
Dans les deux cas, on décompose m.
Algorithme :
Argument : n.
Tant que $n \neq 0$ faire :
Si n pair alors $n \leftarrow n/2$, afficher(0)
Sinon $n \leftarrow (n - 1)/2$, afficher(1)
Fin Si
Fin Tant que.
Fin.

n	53	26	13	6	3	1
Affichage	1	0	1	0	1	1

L'affichage permet de comprendre :

$$53 = 1.2^0 + 0.2^1 + 1.2^2 + 0.2^3 + 1.2^4 + 1.2^5$$

En fait, ici on a l'écriture binaire de 53 à l'envers.
Pour l'algorithme d'exponentiation rapide, on calcule les puissances de x parallèlement à la décomposition de n en puissance de deux.

Arguments : x, n.
$p \leftarrow 1$
Tant que $n \neq 0$ faire :
Si n est pair alors $n \leftarrow n/2$
Sinon $n \leftarrow (n-1)/2$, $p \leftarrow p \times x$
Fin Si.
$x \leftarrow x \times x$.
Fin Tant que.
Afficher p.
Fin.

n	53	26	13	6	3	1	0
p	1	α	α	$\alpha.\alpha^4$	$\alpha.\alpha^4$	$\alpha.\alpha^4.\alpha^{16}$	$\alpha.\alpha^4.\alpha^{16}.\alpha^{32}$
x	α	α^2	α^4	α^8	α^{16}	α^{32}	α^{64}

Remarque En général il faut au plus $2E\left(\dfrac{\ln n}{\ln 2}\right)$ multiplications.

5.2 PGCD et PPCM

5.2.1 PGCD de deux entiers

Définition

Soient $a, b \in \mathbb{Z}$.
Tout $d \in \mathbb{Z}$ tel que $d \mid a$ et $d \mid b$ est appelé diviseur commun à a et b.
On note $\text{Div}(a, b)$ l'ensemble des diviseurs communs à a et b.
Ainsi

$$\text{Div}(a, b) = \text{Div}(a) \cap \text{Div}(b)$$

Exemple $1, -1 \in \text{Div}(a, b)$.

Exemple $\text{Div}(24, 18) = \{1, 2, 3, 6, -1, -2, -3, -6\} = \text{Div}(6)$.

Proposition

Soient $a, b \in \mathbb{Z}$ tels que $(a, b) \neq (0, 0)$ l'ensemble $\text{Div}(a, b)$ possède un plus grand élément.

dém. :
$\text{Div}(a, b) \subset \mathbb{Z}$ et $\text{Div}(a, b) \neq \emptyset$ car $1 \in \text{Div}(a, b)$.
$\forall d \in \text{Div}(a, b)$ on a $d \mid a$ et $d \mid b$.
Si $a \neq 0$ alors $|d| \leqslant |a|$. Si $b \neq 0$ alors $|d| \leqslant |b|$.
Dans les deux cas $\text{Div}(a, b)$ est une partie majorée de $\mathbb{N}$, elle possède donc un plus grand élément.
□

Définition

Soient $a, b \in \mathbb{Z}$ tels que $(a, b) \neq (0, 0)$. Le plus grand élément de $\text{Div}(a, b)$ est appelé pgcd de a et b. On le note $\text{pgcd}(a, b)$ ou encore $a \wedge b$.

Exemple $\text{pgcd}(24, 18) = 6$.

Définition

On pose $\text{pgcd}(0, 0) = 0$.

Proposition

$\forall a, b \in \mathbb{Z}$: $\text{pgcd}(a, b) \in \mathbb{N}$, $\text{pgcd}(a, b) = \text{pgcd}(b, a)$ et $\text{pgcd}(a, b) = \text{pgcd}(|a|, |b|)$.

dém. :
Les diviseurs communs à considérer sont les mêmes.
□

Proposition

$\forall a, b \in \mathbb{Z}, a \mid b \Rightarrow \text{pgcd}(a, b) = |a|$.

dém. :
Pour $a = 0$: $a \mid b \Rightarrow b = 0 \Rightarrow \text{pgcd}(a, b) = |a|$.
Pour $a \neq 0$: Si $a \mid b$ alors $\text{Div}(a) \subset \text{Div}(b)$ et $\text{Div}(a, b) = \text{Div}(a)$ de plus grand élément $|a|$.
□

Exemple $\text{pgcd}(1, a) = 1$ et $\text{pgcd}(a, 0) = |a|$.

5.2.2 Algorithme d'Euclide

Proposition

Soient $a \in \mathbb{Z}$ et $b \in \mathbb{N}^\star$.
En notant r le reste de la division euclidienne de a par b, on a

$$\text{pgcd}(a, b) = \text{pgcd}(b, r)$$

dém. :
Cas $a = b = 0$: ok
Cas $(a, b) \neq (0, 0)$: il suffit d'observer par opérations sur la divisibilité que $\text{Div}(a, b) = \text{Div}(b, r)$.
□

Remarque Si $r = 0$ alors $\text{pgcd}(a, b) = \text{pgcd}(b, 0) = b$.
Si $r \neq 0$ alors $\text{pgcd}(a, b) = \text{pgcd}(b, r) = \text{pgcd}(r, s)$ avec s le reste de la division euclidienne de b par r...
C'est le principe exploité par l'algorithme suivant :

Soient $a, b \in \mathbb{Z}$. On veut calculer $\delta = \text{pgcd}(a, b)$.
On pose $a_0 = \max(|a|, |b|)$ et $a_1 = \min(|a|, |b|)$.
On a $\delta = \text{pgcd}(a_0, a_1)$.
Etape 1 :
Si $a_1 = 0$ alors $\delta = \text{pgcd}(a_0, 0) = a_0$.
Si $a_1 \neq 0$ alors on pose a_2 le reste de la division euclidienne de a_0 par a_1.
On a $\delta = \text{pgcd}(a_1, a_2)$ et $a_2 < a_1$.

Etape 2 :
Si $a_2 = 0$ alors $\delta = \text{pgcd}(a_1, 0) = a_1$.
Si $a_2 \neq 0$ alors on pose a_3 le reste de la division euclidienne de a_1 par a_2.
On a $\delta = \text{pgcd}(a_2, a_3)$ et $a_3 < a_2$.
et ainsi de suite...
Ce processus s'arrête nécessairement car : $a_0 \geqslant a_1 > a_2 > a_3 > ... \geqslant 0$.
Ainsi il existe $m \in \mathbb{N}$ tel que $a_{m+1} = 0$ et alors $\delta = \text{pgcd}(a_m, a_{m+1}) = \text{pgcd}(a_m, 0) = a_m$.
Finalement le pgcd cherché est le dernier reste non nul.

Exemple Pour $a = 24$ et $b = 9$:
$24 = 9 \times 2 + 6$, $9 = 6 \times 1 + 3$ et $6 = 3 \times 2 + 0$. Le pgcd de 24 et 9 vaut 3

Remarque D'un point de vue informatique, l'algorithme d'Euclide peut être rédigé ainsi :
Arguments : a et b.
$x \leftarrow \max(|a|, |b|)$, $y \leftarrow \min(|a|, |b|)$.
Tant que $y \neq 0$ faire :
$r \leftarrow$ reste de la division euclidienne de x par y,
$x \leftarrow y$ et $y \leftarrow r$
Fin tant que
Afficher x.
Fin.

5.2.3 Egalité de Bézout

Théorème

> Soient $a, b \in \mathbb{Z}$, il existe $u, v \in \mathbb{Z}$ tels que
>
> $$\text{pgcd}(a, b) = au + bv$$
>
> Une telle relation est appelé égalité de Bézout.

dém. :
Reprenons l'algorithme d'Euclide

$$a_0 = \max(|a|, |b|),\ a_1 = \min(|a|, |b|)$$

et

$$a_0 = q_1 a_1 + a_2, \ldots, a_{m-2} = q_{m-1} a_{m-1} + a_m \text{ et } a_{m-1} = q_m a_m + 0$$

On peut écrire $a_0 = au_0 + bv_0$ avec $u_0, v_0 \in \mathbb{Z}$
On peut écrire $a_1 = au_1 + bv_1$ avec $u_1, v_1 \in \mathbb{Z}$.

$$a_2 = a_0 - a_1 q_1 = a(u_0 - q_1 u_1) + b(v_0 - v_1 q_1)$$

permet d'écrire $a_2 = au_2 + bv_2$ avec u_2, v_2 et ainsi de suite...
A la fin $a_m = \text{pgcd}(a, b)$ s'écrit $au + bv$ avec $u, v \in \mathbb{Z}$.
□

Remarque La démonstration nous donne ici la démarche à suivre pour obtenir une égalité de Bézout.

Exemple Pour $a = 150, b = 54$:
$150 = 54 \times 2 + 42$, $54 = 42 \times 1 + 12$, $42 = 12 \times 3 + 6$, $12 = 6 \times 2 + 0$ donc $\text{pgcd}(150, 54) = 6$.
En reversant les divisions euclidiennes :
$42 = a - 2b$, $12 = b - 42 = 3b - a$, $6 = 42 - 3 \times 12 = (a - 2b) - 3(3b - a) = 4a - 11b$.

Exemple Pour $a = -11, b = 25$, on obtient $1 = 4 \times 25 + 9 \times (-11)$.

Attention : Les coefficients u et v de l'égalité de Bézout ne sont pas uniques.

5.2.4 Propriété arithmétique du pgcd

Théorème

$\forall a, b \in \mathbb{Z}, \forall d \in \mathbb{Z}$, $d \mid a$ et $d \mid b \Leftrightarrow d \mid \text{pgcd}(a, b)$.

dém. :
($\Leftarrow$) Puisque le pgcd de a et b divisent a et b, la transitivité de la relation de divisibilité donne que si $d \mid \text{pgcd}(a, b)$ alors $d \mid a$ et $d \mid b$.
($\Rightarrow$) Si $d \mid a$ et $d \mid b$ alors d divise le reste de la division euclidienne de a par b. En suivant l'algorithme d'Euclide, on obtient que d divise le pgcd de a et b car d divise chaque reste apparaissant dans l'algorithme d'Euclide calculant ce pgcd.
□

Corollaire

$\text{Div}(a, b) = \text{Div}(\text{pgcd}(a, b))$.

Proposition

$\forall a, b \in \mathbb{Z}, \forall \lambda \in \mathbb{N}, \text{pgcd}(\lambda a, \lambda b) = \lambda \text{pgcd}(a, b)$.

dém. :
Posons $\delta = \text{pgcd}(a, b)$ et $d = \text{pgcd}(\lambda a, \lambda b)$.
$\delta \mid a$ et $\delta \mid b$ donc $\lambda\delta \mid d$.
$\delta = au + bv, \lambda\delta = \lambda au + \lambda bv$ et donc $d \mid \lambda\delta$.
□

5.2.5 PPCM de deux entiers

Définition

Soient $a, b \in \mathbb{Z}$. Tout $m \in \mathbb{Z}$ tel que $a \mid m$ et $b \mid m$ est appelé multiple commun à a et b.
L'ensemble des multiples communs à a et b est noté $\text{Mul}(a, b)$. Ainsi

$$\text{Mul}(a, b) = \text{Mul}(a) \cap \text{Mul}(b)$$

Exemple $0, ab \in \text{Mul}(a, b)$

Exemple Les multiples communs à 6 et 8 sont $\text{Mul}(6, 8) = \{0, 24, 48, \ldots, -24, -48, \ldots\} = \text{Mul}(24)$

Proposition

Soient $a, b \in \mathbb{Z}^\star$, l'ensemble $\mathrm{Mul}(a, b) \cap \mathbb{N}^\star$ possède un plus petit élément.

dém. :
Puisque $|ab| \in \mathrm{Mul}(a, b) \cap \mathbb{N}^\star$, l'ensemble $\mathrm{Mul}(a, b) \cap \mathbb{N}^\star$ est une partie non vide de $\mathbb{N}$, elle admet donc un plus petit élément.
□

Définition

Soient $a, b \in \mathbb{Z}$ tels que $a \neq 0$ et $b \neq 0$, le plus petit élément de $\mathrm{Mul}(a, b) \cap \mathbb{N}^\star$ est appelé ppcm de a et b. On le note $\mathrm{ppcm}(a, b)$ ou $a \vee b$.

Exemple ppcm(6,8) = 24.

Définition

Si $a = 0$ ou $b = 0$, on pose $\mathrm{ppcm}(a, b) = 0$.

Proposition

$\forall a, b \in \mathbb{Z}$, $\mathrm{ppcm}(a, b) \in \mathbb{N}$, $\mathrm{ppcm}(a, b) = \mathrm{ppcm}(b, a)$ et $\mathrm{ppcm}(a, b) = \mathrm{ppcm}(|a|, |b|)$.

dém. :
Les multiples communs à considérer sont les mêmes…
□

Proposition

$\forall a, b \in \mathbb{Z}$, $a \mid b \Rightarrow \mathrm{ppcm}(a, b) = |b|$

dém. :
Si $b = 0$ alors $\mathrm{ppcm}(a, b) = \mathrm{ppcm}(a, 0) = 0 = |b|$.
Si $b \neq 0$ alors si $a \mid b$ on a $\mathrm{Mul}(b) \subset \mathrm{Mul}(a)$ et $\mathrm{Mul}(a, b) = \mathrm{Mul}(b)$ $\mathrm{Mul}(a, b) \cap \mathbb{N}^\star = \mathrm{Mul}(b) \cap \mathbb{N}^\star$ de plus petit élément $|b|$.
□

Exemple $\mathrm{ppcm}(1, a) = |a|$ et $\mathrm{ppcm}(a, 0) = 0$.

5.2.6 Propriétés arithmétiques du ppcm

Théorème

$\forall a, b \in \mathbb{Z}, \forall m \in \mathbb{Z}, a \mid m$ et $b \mid m \Leftrightarrow \mathrm{ppcm}(a, b) \mid m$.

dém. :
$(\Leftarrow)$: Immédiat par transitivité de la relation de divisibilité $(\Rightarrow)$: Si $a = 0$ ou $b = 0$: ok.
Sinon, posons $\delta = \mathrm{pgcd}(\mathrm{ppcm}(a, b), m)$.
$a \mid m$ et $a \mid \mathrm{ppcm}(a, b)$ donc $a \mid \delta$. De même $b \mid \delta$ donc $\mathrm{ppcm}(a, b) \leqslant \delta$.
Or $\delta \mid \mathrm{ppcm}(a, b) \neq 0$ donc $\delta = \mathrm{ppcm}(a, b)$ puis $\mathrm{ppcm}(a, b) = \delta \mid m$.
□

Corollaire

$\mathrm{Mul}(a, b) = \mathrm{Mul}(\mathrm{ppcm}(a, b))$.

Proposition

$\forall a, b \in \mathbb{Z}, \forall \lambda \in \mathbb{N}, \text{ppcm}(\lambda a, \lambda b) = \lambda \text{ppcm}(a, b)$.

dém. :
Cas $\lambda = 0$: ok
Cas $\lambda \neq 0$:
Posons $m = \text{ppcm}(\lambda a, \lambda b)$ et $\mu = \text{ppcm}(a, b)$.
$a \mid \mu$ et $b \mid \mu$ donc $\lambda a \mid \lambda\mu$ et $\lambda b \mid \lambda\mu$ puis $m \mid \lambda\mu$.
Puisque $\lambda a \mid m$ on a $\lambda \mid m$ et donc on peut écrire $m = \lambda m'$.
$\lambda a \mid \lambda m'$ et $\lambda \neq 0$ donne $a \mid m'$. De même $b \mid m'$ d'où $\mu \mid m'$ puis $\lambda\mu \mid m$.
□

Théorème

$\forall a, b \in \mathbb{Z}, \text{pgcd}(a, b)\text{ppcm}(a, b) = |ab|$.

dém. :
Si $a = 0$ ou $b = 0$ alors $\text{ppcm}(a, b) = 0$ et la relation proposée est vraie.
Supposons désormais $a, b \neq 0$.
Posons $\delta = \text{pgcd}(a, b)$ et $\mu = \text{ppcm}(a, b)$.
Puisque δ divise a et b, on peut écrire $a = \delta a'$ et $b = \delta b'$ avec $a', b' \in \mathbb{Z}$. Considérons alors l'entier $m = \delta a'b' = ab' = a'b$. Les entiers a et b divisent m donc μ divise m puis $\delta\mu$ divise $\delta m = ab$.
Inversement, puisque ab est multiple de μ, on peut écrire $ab = \mu n$. Or μ est un multiple de a donc on peut écrire $\mu = ac$. La relation $ab = acn$ donne alors $b = cn$ et donc n divise b. De façon analogue, on obtient n divise a et donc n divise δ. On en déduit que $ab = \mu n$ divise $\mu\delta$.
Puisque ab et $\mu\delta$ se divisent mutuellement, on a $|ab| = \delta\mu$.
Rq :Ce résultat permet de calculer $\text{ppcm}(a, b)$ puisqu'on sait déjà calculer $\text{pgcd}(a, b)$
□

5.3 Nombres premiers entre eux

5.3.1 Définition

Définition

Soient $a, b \in \mathbb{Z}$. On dit que a et b sont premiers entre eux s'ils n'ont pas d'autres diviseurs communs que 1 et -1.

Exemple 3 et 4 sont premiers entre eux.
En revanche 6 et 4 ne le sont pas.

Proposition

Soient $a, b \in \mathbb{Z}$. On a équivalence entre :
(i) a et b sont premiers entre eux ;
(ii) $\text{pgcd}(a, b) = 1$;
(iii) $\text{ppcm}(a, b) = |ab|$.

dém. :
(i) $\Leftrightarrow$ (ii) car $\text{Div}(a, b) = \text{Div}(\text{pgcd}(a, b))$
(ii) $\Leftrightarrow$ (iii) car $\text{pgcd}(a, b)\text{ppcm}(a, b) = |ab|$
□

Remarque Désormais, on note $a \wedge b = 1$ pour signifier que a et b premiers entre eux.

Remarque Si $a \wedge b = 1$ alors

Exemple $1 \wedge a = 1$.

Attention : a ne divise pas b et b ne divise pas a n'entraîne pas que a et b sont premiers entre eux.

Proposition

Soient $a, b, a', b' \in \mathbb{Z}$.
Si $a' \mid a, b' \mid b$ et $a \wedge b = 1$ alors $a' \wedge b' = 1$.

dém. :
On introduit $\delta = \text{pgcd}(a', b')$.
Par transitivité de la divisibilité, on obtient $\delta \mid a$, $\delta \mid b$ donc $\delta \mid \text{pgcd}(a, b) = 1$.
□

5.3.2 Le théorème de Bézout

Théorème

Soient $a, b \in \mathbb{Z}$. On a équivalence entre :
(i) a et b sont premiers entre eux ;
(ii) $\exists u, v \in \mathbb{Z}, au + bv = 1$.

dém. :
(i) $\Rightarrow$ (ii) via égalité de Bézout.
(ii) $\Rightarrow$ (i) Supposons (ii). Soit $\delta = \text{pgcd}(a, b)$. Par opérations, $\delta \mid (au + bv) = 1$ donc $\delta = 1$.
□

Exemple $\forall n \in \mathbb{Z}, n \wedge (n+1) = 1$ car $1 \times (n+1) - 1 \times n = 1$.

Exemple $\forall n \in \mathbb{Z}, (2n+1) \wedge n = 1$ car $1 \times (2n+1) - 2 \times n = 1$.

Proposition

Soient $a, b, c \in \mathbb{Z}$.

$$a \wedge b = 1 \text{ et } a \wedge c = 1 \Rightarrow a \wedge bc = 1$$

dém. :
Supposons a premier avec b et c.
Il existe des entiers u, v, w, t tel que $au + bv = 1$ et $aw + ct = 1$.
En faisant le produit de ces deux relations, on obtient $a(auw + bvw + cut) + bc(wt) = 1$.
Par le théorème de Bézout, on peut conclure $a \wedge (bc) = 1$.
□

Exemple On a

$$\forall n \in \mathbb{Z}, n \wedge (n^2 - 1) = 1$$

En effet $n \wedge (n-1) = 1$ et $n \wedge (n+1) = 1$.

Proposition

Si $a \wedge b_1 = \ldots = a \wedge b_n = 1$ alors $a \wedge (b_1 \ldots b_n) = 1$.
Si $a \wedge b = 1$ alors
$$\forall n, m \in \mathbb{N}, a^n \wedge b^m = 1$$

dém. :
Il suffit de raisonner par récurrence.
□

5.3.3 Le théorème de Gauss

Attention : $a \mid bc$ et a ne divise pas b n'entraînent pas que a divise c.
Prendre par exemple : $a = 6, b = 4, c = 9$.

Théorème

Soient $a, b, c \in \mathbb{Z}$.
$$a \mid bc \text{ et } a \wedge b = 1 \Rightarrow a \mid c$$

dém. :
Puisque a et b sont premiers entre eux, on peut écrire $au + bv = 1$.
On a alors $c = acu + bcv$ et puisque $a \mid bc$, on obtient $a \mid c$.
□

Exemple $4 \mid 3n \Rightarrow 4 \mid n$ car $4 \wedge 3 = 1$.
En revanche $6 \mid 8n$ n'entraîne pas que $6 \mid n$..

Exemple Soient $n, p \in \mathbb{N}^\star$ tels que $n \wedge p = 1$. Montrons que
$$n \mid \binom{n}{p}$$
On sait
$$\binom{n}{p} = \frac{n}{p}\binom{n-1}{p-1}$$
donc
$$p\binom{n}{p} = n\binom{n-1}{p-1}$$
puis
$$n \mid p\binom{n}{p}$$
Or $p \wedge n = 1$ donc, par le théorème de Gauss,
$$n \mid \binom{n}{p}$$

Exemple Soient $a, b, c \in \mathbb{Z}$ tels que $a \wedge b = 1$. Montrons que pgcd$(a, bc) =$ pgcd(a, c).
Posons $d =$ pgcd(a, bc) et $\delta =$ pgcd(a, c).
Puisque $\delta \mid a$ et $\delta \mid bc$, on a $\delta \mid d$.
Inversement $d \mid a$ et $d \mid bc$.
Puisque $a \wedge b = 1$, on a $d \wedge b = 1$ car d divise a.
Sachant $d \mid bc$, par le théorème de Gauss, on obtient $d \mid c$, puis finalement $d \mid \delta$.
Par double divisibilité $d = \delta$.

Attention : $a \mid c, b \mid c$ et $a \neq b$ n'entraînent pas que $ab \mid c$.
Prendre par exemple : $a = 4, b = 2$ et $c = 4$.

Théorème

Soient $a, b, c \in \mathbb{Z}$.

$$a \mid c, b \mid c \text{ et } a \wedge b = 1 \Rightarrow ab \mid c$$

dém. :
$a \mid c = bk$ et $a \wedge b = 1$ donc $a \mid k$ ce qui permet d'écrire $k = a\ell$ puis $c = ab\ell$.
□

Exemple $4 \mid n$ et $3 \mid n \Rightarrow 12 \mid n$ car $4 \wedge 3 = 1$.
En revanche $4 \mid n$ et $6 \mid n$ n'entraînent pas que $24 \mid n$.

Proposition

Si $a_1 \mid b, \ldots, a_n \mid b$ et $a_1, \ldots, a_n$ deux à deux premiers entre eux alors $a_1 \ldots a_n \mid b$.

dém. :
Par récurrence sur $n \in \mathbb{N}^\star$.
Pour $n = 1$: ok.
Supposons la propriété établie au rang $n \geqslant 1$.
Soient $a_1, \ldots, a_n, a_{n+1}$ deux à deux premiers entre eux tels que $a_1 \mid b, \ldots, a_n \mid b, a_{n+1} \mid b$.
Par l'hypothèse de récurrence : $a_1 ... a_n \mid b$.
De plus $a_{n+1} \mid b$ et $a_1 \ldots a_n \wedge a_{n+1} = 1$ donc $a_1 ... a_n a_{n+1} \mid b$.
Récurrence établie.
□

5.3.4 Applications

5.3.4.1 Factorisation du pgcd

Théorème

Soient $a, b \in \mathbb{Z}$, en notant $\delta =$ pgcd(a, b) on peut écrire :
$a = \delta a'$ et $b = \delta b'$ avec $a', b' \in \mathbb{Z}$ tels que pgcd$(a', b') = 1$.

dém. :
Si $(a, b) = (0, 0)$: ok
Si $(a, b) \neq (0, 0)$ alors $\delta \neq 0$.

Comme $\delta \mid a$ et $\delta \mid b$, il existe $a', b' \in \mathbb{Z}$ tels que $a = \delta a'$ et $b = \delta b'$.
On a alors $\delta = \text{pgcd}(\delta a', \delta b') = \delta \text{pgcd}(a', b')$ donc $\text{pgcd}(a', b') = 1$.
□

Exemple Montrons

$$\forall a, b \in \mathbb{Z}, \forall n \in \mathbb{N}, \text{pgcd}(a^n, b^n) = (\text{pgcd}(a, b))^n$$

Soient $a, b \in \mathbb{Z}$ et $n \in \mathbb{N}$.
Posons $\delta = \text{pgcd}(a, b)$ et introduisons a', b' tels que $a = \delta a'$ et $b = \delta b'$.
$\text{pgcd}(a^n, b^n) = \text{pgcd}(\delta^n a'^n, \delta^n b'^n) = \delta^n \times 1$ car $a' \wedge b' = 1$.

Exemple Résolvons dans $\mathbb{N}^2$ le système

$$\begin{cases} \text{pgcd}(x, y) = 10 \\ \text{ppcm}(x, y) = 120 \end{cases}$$

Soit (x, y) un couple solution.
On peut écrire $x = 10x', y = 10y'$ avec $x' \wedge y' = 1$.
$\text{ppcm}(x, y) = 10x'y' = 120$ donc $x'y' = 12$.

$$\text{Div}(12) \cap \mathbb{N} = \{1, 2, 3, 4, 6, 12\}$$

Les couples (x', y') solutions de $x'y' = 12$ sont $(1, 12), (2, 6), (3, 4), ...$
Parmi ceux-ci, $(2, 6)$ et $(6, 2)$ sont à exclure car $x' \wedge y' = 1$.
Finalement

$$(x, y) \in \{(10, 120), (30, 40), (40, 30), (120, 10)\}$$

Vérification : ok

5.3.4.2 Représentant irréductible d'un nombre rationnel

Remarque Pgcd et ppcm sont utiles à la manipulation des nombres rationnels.
- pour réduire au même dénominateur la somme de deux nombres rationnels, il est intéressant de considérer le ppcm des dénominateurs pour trouver un dénominateur commun aussi petit que possible ;
- le pgcd est quand à la suite utile pour réduire le rapport représentant un nombre rationnel…

Théorème

$$\forall r \in \mathbb{Q}, \exists!(p, q) \in \mathbb{Z} \times \mathbb{N}^\star, r = p/q \text{ et } p \wedge q = 1$$

Ce rapport p/q est appelé représentant irréductible du nombre rationnel r.

dém. :
Existence : Soit $r \in \mathbb{Q}$, il existe $(a, b) \in \mathbb{Z} \times \mathbb{N}^\star$ tel que $r = a/b$.
Posons $\delta = \text{pgcd}(a, b)$. On peut écrire $a = \delta p$ et $b = \delta q$ avec $p \wedge q = 1$.
On a alors

$$r = \frac{a}{b} = \frac{\delta p}{\delta q} = \frac{p}{q}$$

de la forme voulue.
Unicité : Soient (p, q) et (p', q') deux couples solutions.

$p/q = p'/q'$ donne $pq' = p'q$.
Puisque $q' \mid p'q$ et $q' \wedge p' = 1$ on a $q' \mid q$.
De même $q \mid q'$ donc $q = q'$ puis $p = p'$.
□

Exemple $\forall n \in \mathbb{N}, \sqrt{n} \in \mathbb{Q} \Leftrightarrow \sqrt{n} \in \mathbb{N}$.
Ainsi la racine carrée d'un entier est soit entière, soit irrationnelle.
($\Leftarrow$) ok
($\Rightarrow$) Si $\sqrt{n} = p/q$ alors $q^2 n = p^2$ et donc $q^2 \mid p^2 = p^2 \times 1$.
Or $q^2 \wedge p^2 = 1$ donc $q^2 \mid 1$ puis $q^2 = 1$.

Remarque Ainsi $\sqrt{2} \notin \mathbb{Q}$ car il n'existe pas de $m \in \mathbb{N}$ tel que $2 = m^2$.
En effet $2 = m^2 \Rightarrow 1 < m^2 < 4$ puis $1 < m < 2$ ce qui est impossible avec $m \in \mathbb{N}$.
De même $\sqrt{3}, \sqrt{5}, \sqrt{6}, \sqrt{7}, \sqrt{8}, \sqrt{10} \notin \mathbb{Q}$.

5.4 Décomposition primaire d'un entier

5.4.1 Nombres premiers

Définition

> Soit $p \in \mathbb{N}$ tel que $p \geqslant 2$.
> On dit que p est premier si ses seuls diviseurs positifs sont 1 et p.
> Sinon, l'entier p est dit composé.
> On note $\mathcal{P}$ l'ensemble des nombres premiers.

Exemple $2, 3, 5, 7, 11$ sont premiers alors que $4, 6, 8, 9, 10, 12$ sont composés.
1 n'est pas un nombre premier, ni un nombre composé.

Définition

> On appelle facteur premier d'un entier n tout $p \in \mathcal{P}$ tel que $p \mid n$.

Proposition

> Tout $n \geqslant 2$ possède au moins un facteur premier.

dém. :
Par récurrence forte sur $n \geqslant 2$.
Pour $n = 2$: ok
Supposons la propriété établie jusqu'au rang $n \geqslant 2$.
Si $n + 1$ est premier, le résultat est établie au rang $n + 1$.
Si $n + 1$ est composé alors $n + 1$ est divisible par un certain d tel que $2 \leqslant d \leqslant n$.
Par hypothèse de récurrence forte, d est divisible par un nombre premier et donc $n + 1$ est divisible par ce même nombre premier.
Récurrence établie.
□

Remarque Si n est composé alors il possède un facteur premier $p \leqslant \sqrt{n}$.

Proposition

> $\mathcal{P}$ est infini.

dém. :
Par l'absurde, supposons que $\mathcal{P}$ soit un ensemble fini.
En notant N sont cardinal, on peut indexer les éléments de $\mathcal{P}$ pour écrire : $\mathcal{P} = \{p_1, \ldots, p_N\}$.
Considérons alors $n = p_1 \ldots p_N + 1 \geqslant 2$.
Par la proposition précédente, n est divisible par un certain nombre premier p.
Puisque $p \in \mathcal{P}$, il existe $i \in \{1, \ldots, N\}$ tel que $p_i \mid n$. Or $p_i \mid p_1 \ldots p_N$ donc $p_i \mid n - p_1 \ldots p_N = 1$.
C'est absurde.
□

Proposition

> Soit $n \geqslant 2$, si n est composé alors n possède un diviseur d avec $2 < d \leqslant \sqrt{n}$.

dém. :
Si $n = ab$ alors, quitte à permuter, on peut supposer $a \leqslant b$ et alors a est un diviseur de n vérifiant $a^2 \leqslant ab \leqslant n$.
□

Remarque Si n est composé alors il possède un facteur premier $p \leqslant \sqrt{n}$.

Le crible d'Eratosthène permet de déterminer les premiers nombres premiers par élimination des nombres composés. On figure dans un tableau, les entiers allant par exemple de 1 à 100.
- on élimine 1 qui est à part ;
- 2 est un nombre premier et on élimine tous les multiples de 2 qui sont, de fait, des nombres composés ;
- le premier entier restant, ici 3, est alors un nombre premier et on élimine tous ses multiples ;
- le premier entier restant, maintenant 5, est un nombre premier, on élimine tous ses multiples ;
- le premier entier restant, désormais 7, est un nombre premier, on élimine tous ses multiples ;
- enfin puisque l'entier qui suit est $11 > 10 = \sqrt{100}$, on est assuré que tous les entiers restant sont premiers !

En effet les entiers composé inférieur à 100 possède un facteur premier inférieur à $\sqrt{100}$ et on donc été éliminés.

5.4.2 Propriétés arithmétiques des nombres premiers

Proposition

> Soit $a \in \mathbb{Z}$ et p un nombre premier. Si p ne divise pas a alors $p \wedge a = 1$.

dém. :
Par contraposée :
Si $p \wedge a \neq 1$ alors introduisons $\delta = \mathrm{pgcd}(a, p) \neq 1$.
Puisque $\delta \in \mathbb{N}$ et $\delta \mid p$ on a $\delta = p$ car les seuls diviseurs positifs de p sont 1 lui-même.
Par suite $p \mid a$ car δ est diviseur de a.
□

Attention : Ce résultat ne vaut que pour p nombre premier !

Théorème

> Soient $a, b \in \mathbb{Z}$ et p un nombre premier.
> Si $p \mid ab$ alors $p \mid a$ ou $p \mid b$.

dém. :
Supposons $p \mid ab$.
Si $p \mid a$: ok
Sinon $p \wedge a = 1$ or $p \mid ab$ donc $p \mid b$ en vertu du théorème de Gauss.
□

Exemple $2 \mid ab \Rightarrow 2 \mid a$ ou $2 \mid b$.

Attention : Ce résultat ne vaut que pour p nombre premier !

Corollaire

Soient $a_1, \dots, a_n \in \mathbb{Z}$ et p un nombre premier.
Si $p \mid a_1 \dots a_n$ alors il existe $1 \leqslant i \leqslant n$ tel que $p \mid a_i$.

dém. :
Par récurrence sur $n \in \mathbb{N}^\star$.
□

5.4.3 Décomposition primaire

Théorème

Pour tout $n \in \mathbb{N}$ tel que $n \geqslant 2$, il existe $N \in \mathbb{N}^\star$, $p_1, \dots, p_N$ nombres premiers deux à deux distincts et $\alpha_1, \dots, \alpha_n \in \mathbb{N}^\star$ tels que

$$n = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_N^{\alpha_N}$$

De plus cette décomposition est unique à l'ordre près des facteurs.

dém. :
Existence :
Par récurrence forte sur $n \geqslant 2$, montrons que n peut s'écrire comme produit de nombres premiers. Il suffira ensuite de regrouper entre eux les nombres premiers égaux pour obtenir la formule proposée.
La propriété est vraie pour $n = 2$.
Supposons la propriété vraie jusqu'au rang $n \geqslant 2$.
Si $n + 1$ est un nombre premier alors la propriété est vraie.
Si $n + 1$ est un nombre composé alors on peut écrire $n + 1 = ab$ avec $2 \leqslant a, b \leqslant n$.
Par hypothèse de récurrence forte, on peut écrire a et b comme produit de nombres premiers et l'on obtient donc $n + 1 = ab$ produit de nombres premiers.
Récurrence établie.
Unicité (à l'ordre près des facteurs) :
Supposons deux écritures $n = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_N^{\alpha_N}$ et $n = q_1^{\beta_1} q_2^{\beta_2} \dots q_M^{\beta_M}$ de la forme annoncée.
Pour tout $1 \leqslant i \leqslant n$ on a $p_i \mid n$ donc p_i divise nécessairement l'un des q_j et dès lors, lui est égal car ce sont des nombres premier. Ainsi $\{p_1, \dots, p_n\} \subset \{q_1, \dots, q_M\}$.
Par un raisonnement symétrique, on obtient l'autre inclusion et donc l'égalité. En particulier $M = N$.
Quitte à permuter les couples (q_j, β_j) on peut supposer $q_1 = p_1, \dots, q_N = p_N$.
On a alors $n = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_N^{\alpha_N} = p_1^{\beta_1} p_2^{\beta_2} \dots p_N^{\beta_N}$. Pour tout $1 \leqslant i \leqslant n$, $p_i^{\alpha_i} \mid n = p_1^{\beta_1} p_2^{\beta_2} \dots p_N^{\beta_N} = p_i^{\beta_i} k$ avec $p_i^{\alpha_i} \wedge k = 1$ donc $p_i^{\alpha_i} \mid p_i^{\beta_i}$ puis $\alpha_i \leqslant \beta_i$.
De manière symétrique $\beta_i \leqslant \alpha_i$ et finalement l'égalité $\alpha_i = \beta_i$.
□

Définition

Cette écriture est appelée décomposition primaire de l'entier $n \geqslant 2$.
Les $p_1, \dots, p_N$ correspondent aux facteurs premiers de n.

Exemple $12 = 2^2 \times 3, 50 = 2 \times 5^2, 84 = 2^2 \times 3 \times 7$.

5.4.4 Diviseurs d'un entier

Théorème

Soit $n \in \mathbb{N}^\star$ s'écrivant
$$n = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_N^{\alpha_N}$$
avec $p_1, \dots, p_N$ nombres premiers deux à deux distincts et $\alpha_1, \alpha_2, \dots, \alpha_N \in \mathbb{N}$.
Les diviseurs positifs de n sont les entiers d pouvant s'écrire
$$d = p_1^{\beta_1} p_2^{\beta_2} \dots p_N^{\beta_N}$$
avec $0 \leqslant \beta_i \leqslant \alpha_i$ pour tout $1 \leqslant i \leqslant N$.

Exemple Les diviseurs positifs de $12 = 2^2 \times 3^1$ sont :
$$2^0 \times 3^0, 2^1 \times 3^0, 2^2 \times 3^0, 2^0 \times 3^1, 2^1 \times 3^1 \text{ et } 2^2 \times 3^1$$

dém. :
Les nombres considérés sont bien diviseurs de n.
Inversement, soit d un diviseur positif de n.
Puisque $n \neq 0$ on a $d \neq 0$.
Si $d = 1$ alors on peut écrire $d = p_1^{\beta_1} p_2^{\beta_2} \dots p_N^{\beta_N}$ avec $\forall 1 \leqslant i \leqslant N, \beta_i = 0$.
Si $d \geqslant 2$ alors soit p un facteur premier de d. Puisque $p \mid d$ on a $p \mid n$ donc p est un facteur premier de n.
Ainsi il existe $1 \leqslant i \leqslant n$ tel que $p = p_i$.
Puisque tous les facteurs premiers de d sont aussi facteurs premiers de n, la décomposition primaire de d permet d'écrire $d = p_1^{\beta_1} p_2^{\beta_2} \dots p_N^{\beta_N}$ avec $\forall 1 \leqslant i \leqslant N, \beta_i \in \mathbb{N}$.
De plus $\forall 1 \leqslant i \leqslant N, p_i^{\beta_i} \mid d$ donc $p_i^{\beta_i} \mid n = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_N^{\alpha_N}$ puis $p_i^{\beta_i} \mid p_i^{\alpha_i}$ d'où $\beta_i \leqslant \alpha_i$.
Finalement d est bien de la forme voulue.
□

5.4.5 Pgcd, ppcm et décomposition primaire

Soient $a, b \in \mathbb{N}$ tels que $a, b \geqslant 2$.
A partir des décompositions primaires de a et b on peut écrire simultanément :
$$a = p_1^{\alpha_1} \dots p_N^{\alpha_N} \text{ et } b = p_1^{\beta_1} \dots p_N^{\beta_N}$$
avec $p_1, \dots, p_N$ nombre premiers deux à deux distincts et $\alpha_1, \dots, \alpha_N, \beta_1, \dots, \beta_N \in \mathbb{N}$.

Exemple $a = 308 = 2^2 \times 7 \times 11 = 2^2 \times 3^0 \times 5^0 \times 7^1 \times 11^1 \times 13^0$.
$b = 2340 = 2^2 \times 3^2 \times 5 \times 13 = 2^2 \times 3^2 \times 5^1 \times 7^0 \times 11^0 \times 13^1$.

Théorème

Avec les décompositions ci-dessus :

$$\text{pgcd}(a,b) = \prod_{i=1}^{N} p_i^{\min(\alpha_i,\beta_i)} \text{ et ppcm}(a,b) = \prod_{i=1}^{N} p_i^{\max(\alpha_i,\beta_i)}$$

dém. :

Soit $\delta = \text{pgcd}(a,b)$ et $d = \prod_{i=1}^{N} p_i^{\min(\alpha_i,\beta_i)}$.

Puisque $d \mid a$ et $d \mid b$, on a $d \mid \delta$.

Inversement, comme $\delta \mid a$ et $\delta \mid b$ on peut écrire : $\delta = \prod_{i=1}^{N} p_i^{\gamma_i}$ avec $\forall 1 \leqslant i \leqslant n, 0 \leqslant \gamma_i \leqslant \alpha_i, \beta_i$

Ainsi $\gamma_i \leqslant \min(\alpha_i, \beta_i)$ pour tout $i \in \{1, \ldots, N\}$ et donc $\delta \mid d$.

Par double divisibilité $\delta = d$

Pour obtenir la formule relative au ppcm, on exploite les propriétés $\text{pgcd}(a,b)\text{ppcm}(a,b) = ab$ et $\min(\alpha,\beta) + \max(\alpha,\beta) = \alpha + \beta$.

□

Exemple Pour $a = 308$ et $b = 2340$: $\text{pgcd}(a,b) = 4$ et $\text{ppcm}(a,b) = 180180$.

Chapitre 6

Espaces vectoriels

$\mathbb{K}$ désigne le corps $\mathbb{R}$ ou $\mathbb{C}$.

6.1 Structure d'espace vectoriel

6.1.1 Loi de composition externe

Définition

On appelle loi de composition externe (ou produit extérieur) opérant de $\mathbb{K}$ sur un ensemble E toute application :

$$\begin{cases} \mathbb{K} \times E \to E \\ (\lambda, \vec{x}) \mapsto \lambda.\vec{x} \end{cases}$$

Une loi de composition externe est usuellement notée par un point.
Les éléments de $\mathbb{K}$ sont appelés scalaires.
Les éléments de E sont appelés vecteurs et seront, dans un premier temps, notés surmontés d'une flèche.

Exemple Le produit extérieur réel sur la direction du plan ou de l'espace géométrique est une loi de composition externe qui à $\lambda \in \mathbb{R}$ et $\vec{x}$ vecteur associe le vecteur $\lambda.\vec{x}$.

Définition

Soit E un ensemble muni d'un produit extérieur (.) de $\mathbb{K}$ sur E.
Une partie A de E est dite stable pour ce produit extérieur si

$$\forall \lambda \in \mathbb{K}, \forall \vec{x} \in A, \lambda.\vec{x} \in A$$

On peut alors considérer l'application restreinte

$$\begin{cases} \mathbb{K} \times A \to A \\ (\lambda, \vec{x}) \mapsto \lambda.\vec{x} \end{cases}$$

qui définit un produit extérieur de $\mathbb{K}$ sur A appelé produit extérieur induit.

Exemple $\emptyset$ et E sont des parties stables pour n'importe quel produit extérieur sur E.

Exemple Sur la direction du plan ou de l'espace géométrique, l'ensemble A formé des vecteurs colinéaires à un vecteur $\vec{u}$ donné est une partie stable pour le produit extérieur sur E précédemment introduit.

6.1.2 Définition d'un espace vectoriel

Définition

Soient E un ensemble, $+$ une loi de composition interne sur E et . une loi de composition externe opérant de $\mathbb{K}$ sur E.
On dit que le triplet $(E, +, .)$, ou plus brièvement E, est un $\mathbb{K}$-espace vectoriel si :
(1) $(E, +)$ est un groupe abélien ;
(2) $\forall \lambda, \mu \in \mathbb{K}, \forall \vec{u}, \vec{v} \in E$:
(a) $(\lambda + \mu).\vec{u} = \lambda.\vec{u} + \mu.\vec{u}$;
(b) $\lambda.(\vec{u} + \vec{v}) = \lambda.\vec{u} + \lambda.\vec{v}$;
(c) $\lambda.(\mu.\vec{u}) = (\lambda\mu).\vec{u}$;
(d) $1.\vec{u} = \vec{u}$.
L'élément neutre de $(E, +)$ est alors appelé vecteur nul, on le note $\vec{0}$.

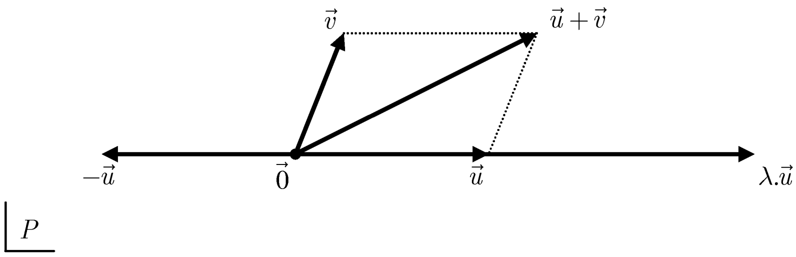
Remarque Il faut faire la distinction entre :
- les scalaires (qui sont des éléments du corps $\mathbb{K}$) ;
- les vecteurs (éléments de l'espace E).
En particulier, on distingue le scalaire nul ($0 \in \mathbb{K}$) et le vecteur nul ($\vec{0} \in E$).

Remarque Dans un $\mathbb{K}$-espace vectoriel on peut :
- sommer deux vecteurs $\vec{u}$ et $\vec{v}$ i.e. calculer $\vec{u} + \vec{v}$;
- faire la différence deux vecteurs $\vec{u}$ et $\vec{v}$ i.e. calculer $\vec{u} - \vec{v}$;
- multiplier un vecteur $\vec{u}$ par un scalaire λ i.e. calculer $\lambda.\vec{u}$ (et on évitera d'écrire $\vec{u}.\lambda$) ;
- diviser un vecteur $\vec{u}$ par un scalaire $\lambda \neq 0$ i.e. calculer $\dfrac{1}{\lambda}.\vec{u}$ (parfois abusivement noté $\dfrac{\vec{u}}{\lambda}$).

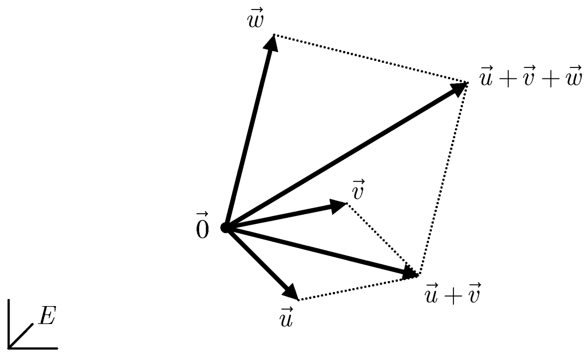
Attention : On ne peut pas multiplier deux vecteurs, ni diviser par un vecteur car les lois correspondantes ne sont pas a priori définie.

6.1.3 Visualisation géométrique d'un espace vectoriel

Notons P la direction du plan géométrique $\mathcal{P}$.
P est un $\mathbb{R}$-espace vectoriel pour les lois usuelles.
Pour visualise P, considérons O un point du plan géométrique $\mathcal{P}$. L'application $M \mapsto \overrightarrow{OM}$ est une bijection de $\mathcal{P}$ vers P. Cette bijection permet de visualiser les vecteurs du plan P en représentant systématiquement ceux-ci à partir du point O. Ce point O correspond alors au vecteur nul et on peut visualiser l'addition, la soustraction et le produit extérieur comme dans la figure ci-dessous



Notons E la direction de l'espace géométrique $\mathcal{E}$.
E est un $\mathbb{R}$-espace vectoriel pour les lois usuelles.
Comme ci-dessus, en choisissant un point central jouant le rôle du vecteur nul puis en convenant que tout vecteur de E est représenté en prenant ce point pour origine, on peut visualiser géométriquement le $\mathbb{R}$-espace vectoriel E.



Plus généralement, pour visualiser géométriquement un espace vectoriel, on choisira un élément central correspondant au vecteur nul et on représente tous les vecteurs en prenant le vecteur nul pour origine ; on peut alors visualiser les opérations dans l'espace vectoriel à l'instar des figures précédentes.

6.1.4 Exemples d'espaces vectoriels

6.1.4.1 structure sur $\mathbb{K}^n$

Soient $n \in \mathbb{N}^\star$ et $E = \mathbb{K}^n$. Pour $\lambda \in \mathbb{K}$ et $\vec{x} = (x_1, \dots, x_n), \vec{y} = (y_1, \dots, y_n) \in \mathbb{K}^n$ on pose

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

On définit ainsi un produit extérieur de $\mathbb{K}$ sur $\mathbb{K}^n$ et une loi de composition interne additive sur $\mathbb{K}^n$.

Théorème

$(\mathbb{K}^n, +, .)$ est un $\mathbb{K}$-espace vectoriel de vecteur nul $\vec{0} = (0, \dots, 0)$.

dém. :
$(\mathbb{K}, +)$ est un groupe abélien de neutre 0 donc $(\mathbb{K}^n, +)$ est aussi un groupe abélien de neutre $(0, \dots, 0)$

car l'addition sur $\mathbb{K}^n$ est la loi produit définie à partir de l'addition sur $\mathbb{K}$.
Pour $\lambda, \mu \in \mathbb{K}$ et $\vec{x} = (x_1, \ldots, x_n), \vec{y} = (y_1, \ldots, y_n) \in \mathbb{K}^n$,

$$(\lambda + \mu).\vec{x} = ((\lambda + \mu)x_1, \ldots, (\lambda + \mu)x_n) = \lambda\,(x_1, \ldots, x_n) + \mu\,(x_1, \ldots, x_n) = \lambda.\vec{x} + \mu.\vec{x}$$

$$\lambda.(\vec{x} + \vec{y}) = (\lambda(x_1 + y_1), \ldots, \lambda(x_n + y_n)) = \lambda\,(x_1, \ldots, x_n) + \lambda\,(y_1, \ldots, y_n) = \lambda.\vec{x} + \lambda.\vec{y}$$

$$\lambda.(\mu.\vec{x}) = (\lambda(\mu x_1), \ldots, \lambda(\mu x_n)) = (\lambda\mu)\,(x_1, \ldots, x_n) = (\lambda\mu).\vec{x}$$

$$\text{et } 1.\vec{x} = (x_1, \ldots, x_n) = \vec{x}$$

Finalement $(\mathbb{K}^n, +, .)$ est bien un $\mathbb{K}$-espace vectoriel.
□

Exemple $\mathbb{R}^n$ est un $\mathbb{R}$-espace vectoriel, $\mathbb{C}^n$ est un $\mathbb{C}$-espace vectoriel.

Exemple $\mathbb{K}$ est un $\mathbb{K}$-espace vectoriel.
Dans cette situation, scalaires et vecteurs désignent des éléments de $\mathbb{K}$ et seul l'interprétation voulue permet de distinguer scalaires et vecteurs. Notons aussi que le produit extérieur correspond alors à la multiplication sur $\mathbb{K}$.

6.1.4.2 structure sur $E \times F$

Soient E et F deux $\mathbb{K}$-espaces vectoriels. Pour $\lambda \in \mathbb{K}$, $(\vec{x}, \vec{y}), (\vec{x}', \vec{y}') \in E \times F$ on pose

$$\lambda.(\vec{x}, \vec{y}) = (\lambda.\vec{x}, \lambda.\vec{y}) \text{ et } (\vec{x}, \vec{y}) + (\vec{x}', \vec{y}') = (\vec{x} + \vec{x}', \vec{y} + \vec{y}')$$

On définit ainsi un produit extérieur de $\mathbb{K}$ sur $E \times F$ et une loi de composition interne additive sur $E \times F$.

Théorème

$E \times F$ est un $\mathbb{K}$-espace vectoriel de vecteur nul $\vec{0}_{E \times F} = \left(\vec{0}_E, \vec{0}_F\right)$.

dém. :
L'addition sur $E \times F$ correspond à la loi produit obtenue à partir des additions sur E et F. Puisque $(E, +)$ et $(F, +)$ sont des groupes abéliens de neutres $\vec{0}_E$ et $\vec{0}_F$, on peut affirmer que $(E \times F, +)$ est un groupe abélien de neutre $\left(\vec{0}_E, \vec{0}_F\right)$.
Pour $\lambda, \mu \in \mathbb{K}$ et $(\vec{x}, \vec{y}), (\vec{x}', \vec{y}') \in E \times F$,

$$(\lambda + \mu).(\vec{x}, \vec{y}) = ((\lambda + \mu).\vec{x}, (\lambda + \mu).\vec{y}) = (\lambda.\vec{x} + \mu.\vec{x}, \lambda.\vec{y} + \mu.\vec{y}) = \lambda.(\vec{x}, \vec{y}) + \mu.(\vec{x}, \vec{y})$$

$$\lambda.\left((\vec{x}, \vec{y}) + (\vec{x}', \vec{y}')\right) = (\lambda.(\vec{x} + \vec{x}'), \lambda.(\vec{y} + \vec{y}')) = (\lambda.\vec{x} + \lambda.\vec{x}', \lambda.\vec{y} + \lambda.\vec{y}') = \lambda.(\vec{x}, \vec{y}) + \lambda.(\vec{x}', \vec{y}')$$

$$\lambda.\left(\mu.(\vec{x}, \vec{y})\right) = \lambda.\left(\mu.\vec{x}, \mu.\vec{y}\right) = (\lambda.(\mu.\vec{x}), \lambda.(\mu.\vec{y})) = ((\lambda\mu).\vec{x}, (\lambda\mu).\vec{y}) = (\lambda\mu).(\vec{x}, \vec{y}) \text{ et } 1.(\vec{x}, \vec{y}) = (1.\vec{x}, 1.\vec{y}) = (\vec{x}, \vec{y})$$

Finalement $(E \times F, +, .)$ est bien un $\mathbb{K}$-espace vectoriel.
□

6.1.4.3 structure sur $\mathcal{F}(X, \mathbb{K})$

Soient X un ensemble et $E = \mathcal{F}(X, \mathbb{K})$. Pour $\lambda \in \mathbb{K}$ et $f, g : X \to \mathbb{K}$, on définit $\lambda.f : X \to \mathbb{K}$ et $f + g : X \to \mathbb{K}$ par :

$$\forall x \in X, (\lambda.f)(x) = \lambda f(x) \text{ et } (f + g)(x) = f(x) + g(x).$$

On définit ainsi un produit extérieur de $\mathbb{K}$ sur $\mathcal{F}(X,\mathbb{K})$ et une loi de composition interne additive sur $\mathcal{F}(X,\mathbb{K})$.

Théorème

$(\mathcal{F}(X,\mathbb{K}),+,.)$ est un $\mathbb{K}$-espace vectoriel dont le vecteur nul est la fonction nulle.

dém. :
$(\mathbb{K},+)$ est un groupe abélien de neutre 0 donc $(\mathcal{F}(X,\mathbb{K}),+)$ est aussi un groupe abélien de neutre la fonction nulle car l'addition sur $\mathcal{F}(X,\mathbb{K})$ est la loi produit définie à partir de l'addition sur $\mathbb{K}$.
Soient $\lambda,\mu\in\mathbb{K}$ et $f,g\in\mathcal{F}(X,\mathbb{K})$.
Pour tout $x\in X$,

$$[(\lambda+\mu).f](x)=(\lambda+\mu)f(x)=\lambda f(x)+\mu g(x)=(\lambda.f)(x)+(\mu.g)(x)=[\lambda f+\mu g](x)$$

$$[\lambda.(f+g)](x)=\lambda(f+g)(x)=\lambda\left(f(x)+g(x)\right)=\lambda f(x)+\lambda g(x)=(\lambda.f)(x)+(\lambda.g)(x)=[\lambda.f+\lambda.g](x)$$

$$[\lambda.(\mu.f)](x)=\lambda(\mu.f)(x)=\lambda(\mu f(x))=(\lambda\mu)f(x)=[(\lambda\mu).f](x) \text{ et } (1.f)(x)=1\times f(x)=f(x)$$

Ainsi $(\lambda+\mu).f=\lambda.f+\mu.f$, $\lambda.(f+g)=\lambda.f+\lambda.g$, $\lambda.(\mu.f)=(\lambda\mu).f$ et $1.f=f$.
Finalement $(\mathcal{F}(X,\mathbb{K}),+,.)$ est bien un $\mathbb{K}$-espace vectoriel.
□

Exemple Pour $X=\mathcal{D}\subset\mathbb{R}$, les ensembles $\mathcal{F}(\mathcal{D},\mathbb{R})$ et $\mathcal{F}(\mathcal{D},\mathbb{C})$ des fonctions réelles ou complexes d'une variable réelle définies sur $\mathcal{D}$ sont des $\mathbb{R}$ et $\mathbb{C}$-espaces vectoriels
Pour $X=\mathbb{N}$, les ensembles $\mathbb{R}^{\mathbb{N}}$ et $\mathbb{C}^{\mathbb{N}}$ des suites réelles et complexes sont des $\mathbb{R}$ et $\mathbb{C}$-espaces vectoriels.

6.1.4.4 structure sur $\mathcal{F}(X,F)$

Soient X un ensemble et F un $\mathbb{K}$-espace vectoriel.
Pour $\lambda\in\mathbb{K}$ et $f,g\in\mathcal{F}(X,F)$ on définit $\lambda.f:X\to F$ et $f+g:X\to F$ par :

$$\forall x\in X, (\lambda.f)(x)=\lambda.f(x) \text{ et } (f+g)(x)=f(x)+g(x).$$

Théorème

$(\mathcal{F}(X,F),+,.)$ est un $\mathbb{K}$-espace vectoriel de vecteur nul égal à la fonction constante égale à $\vec{o}_F$.

dém. :
Il suffit de reprendre la démonstration qui précède avec ici des fonctions à valeurs dans l'espace vectoriel F ce qui au final ne change pas grand-chose.
□

6.1.4.5 Les espaces complexes sont aussi réels

Soit E un $\mathbb{C}$-espace vectoriel. La loi de composition externe opérant de $\mathbb{C}$ sur E définit aussi par restriction, une loi de composition externe opérant de $\mathbb{R}$ sur E. Les propriétés calculatoires étant conservées, on peut affirmer que E est alors aussi un $\mathbb{R}$-espace vectoriel.

Exemple $\mathbb{C}^n$ est un $\mathbb{R}$-espace vectoriel.
$\mathcal{F}(X,\mathbb{C})$ et $\mathbb{C}^{\mathbb{N}}$ sont des $\mathbb{R}$-espaces vectoriels.

Exemple Il est fréquent de percevoir $\mathbb{C}$ comme un $\mathbb{R}$-espace vectoriel. Dans ce cas les vecteurs sont les nombres complexes, les scalaires sont les nombres réels et la loi de composition externe est le produit usuel.

6.1.5 Calculs dans un $\mathbb{K}$-espace vectoriel

Soit $(E,+,.)$ un $\mathbb{K}$-espace vectoriel.

Proposition

$\forall \lambda_1,\ldots,\lambda_n \in \mathbb{K}, \forall \vec{u} \in E,$

$$\sum_{i=1}^{n} (\lambda_i.\vec{u}) = \left(\sum_{i=1}^{n} \lambda_i\right).\vec{u}$$

dém. :
Par récurrence sur $n \in \mathbb{N}^\star$ en exploitant $\lambda.\vec{u}+\mu.\vec{u} = (\lambda+\mu).\vec{u}$.
□

Proposition

$\forall \lambda \in \mathbb{K}, \forall \vec{u}_1,\ldots,\vec{u}_n,$

$$\sum_{i=1}^{n} (\lambda.\vec{u}_i) = \lambda.\left(\sum_{i=1}^{n} \vec{u}_i\right)$$

dém. :
Par récurrence sur $n \in \mathbb{N}^\star$ en exploitant $\lambda.\vec{u}+\lambda.\vec{v} = \lambda.(\vec{u}+\vec{v})$.
□

Proposition

$\forall \lambda \in \mathbb{K}, \forall \vec{u} \in E,$

$$0.\vec{u} = \vec{0} \text{ et } \lambda.\vec{0} = \vec{0}$$

dém. :
$0.\vec{u} = (0+0).\vec{u} = 0.\vec{u}+0.\vec{u}$ et en ajoutant l'opposé du vecteur $0.\vec{u}$ de part et d'autre, obtient $\vec{0} = 0.\vec{u}$.
$\lambda.\vec{0} = \lambda.(\vec{0}+\vec{0}) = \lambda.\vec{0}+\lambda.\vec{0}$ et en ajoutant l'opposé du vecteur $\lambda.\vec{0}$ de part et d'autre, obtient $\vec{0} = \lambda.\vec{0}$.
□

Proposition

$\forall \vec{u} \in E,$

$$(-1).\vec{u} = -\vec{u}$$

dém. :
$\vec{u}+(-1).\vec{u} = 1.\vec{u}+(-1).\vec{u} = (1+(-1)).\vec{u} = 0.\vec{u} = \vec{0}$.
□

Proposition

$\forall \lambda \in \mathbb{K}, \forall \vec{u} \in E,$

$$\lambda.\vec{u} = \vec{0} \Rightarrow \lambda = 0 \text{ ou } \vec{u} = \vec{0}$$

dém. :
Supposons $\lambda.\vec{u} = \vec{0}$ et $\lambda \neq 0$.
Dans le corps $\mathbb{K}$, le scalaire λ est inversible et l'on peut donc introduire son inverse λ^{-1}. En multipliant par celui-ci, $\lambda^{-1}.(\lambda.\vec{u}) = (\lambda^{-1}\lambda).\vec{u} = 1.\vec{u} = \vec{u}$ or $\lambda^{-1}.(\lambda\vec{u}) = \lambda^{-1}.\vec{0} = \vec{0}$ donc $\vec{u} = \vec{0}$.
□

Exemple Soient $\vec{u} \in E \backslash \left\{ \vec{0} \right\}$ et $\lambda, \mu \in \mathbb{K}$. Montrons $\lambda.\vec{u} = \mu.\vec{u} \Rightarrow \lambda = \mu$.
Supposons $\lambda.\vec{u} = \mu.\vec{u}$.
On a $\lambda.\vec{u} - \mu.\vec{u} = \vec{0}$ puis $\lambda.\vec{u} + (-\mu).\vec{u} = \vec{0}$ et enfin $(\lambda - \mu).\vec{u} = \vec{0}$.
Puisque $\vec{u} \neq \vec{0}$, on obtient $\lambda - \mu = 0$ i.e. $\lambda = \mu$.

6.1.6 Combinaison linéaire

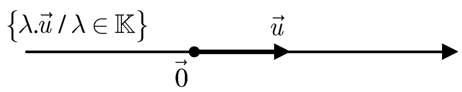
Définition

On dit qu'un vecteur $\vec{x}$ de E est combinaison linéaire des vecteurs $\vec{e}_1, \dots, \vec{e}_n$ de E si on peut écrire

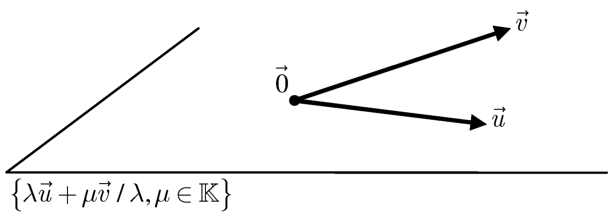
$$\vec{x} = \lambda_1 \vec{e}_1 + \cdots + \lambda_n \vec{e}_n = \sum_{i=1}^{n} \lambda_i \vec{e}_i$$

avec $\lambda_1, \dots, \lambda_n \in \mathbb{K}$.

Exemple Les vecteurs combinaisons linéaires d'un vecteur $\vec{u}$ sont ceux de la forme $\lambda.\vec{u}$ avec $\lambda \in \mathbb{K}$.



Exemple Les vecteurs combinaisons linéaires de deux vecteurs $\vec{u}$ et $\vec{v}$ sont ceux de la forme $\lambda.\vec{u} + \mu.\vec{v}$ avec $\lambda, \mu \in \mathbb{K}$.



Exemple Dans $E = \mathbb{R}^3$, considérons les vecteurs $\vec{u} = (1, 2, 1)$ et $\vec{v} = (1, -1, 0)$.
Déterminons pour quel(s) $a \in \mathbb{R}$, le vecteur $\vec{w}_a = (a, a+1, a+2)$ est combinaison linéaire de $\vec{u}$ et $\vec{v}$.
Si $\vec{w}_a$ est solution alors il existe $\lambda, \mu \in \mathbb{R}$ tels que $\vec{w} = \lambda.\vec{u} + \mu.\vec{v}$ ce qui donne le système

$$\begin{cases} a = \lambda + \mu \\ a + 1 = 2\mu - \mu \\ a + 2 = \lambda \end{cases}$$

puis

$$\begin{cases} \lambda = a + 2 \\ \mu = -2 \\ a = -5 \end{cases}$$

On en déduit $a = -5$.
Inversement, pour $a = -5$, on a $\vec{w}_a = \lambda.\vec{u} + \mu.\vec{v}$ avec $\lambda = -3$ et $\mu = -2$.

Exemple Dans $E = \mathcal{F}(\mathbb{R}, \mathbb{R})$, considérons les fonctions $f : x \mapsto \cos x$ et $g : x \mapsto \sin x$.
Pour tout $A, \varphi \in \mathbb{R}$, la fonction $h : x \mapsto A\cos(x - \varphi)$ est combinaison linéaire des fonctions f et g.
En effet par développement, $h(x) = A\cos\varphi\cos x + A\sin\varphi\sin x$ et donc $h = \lambda f + \mu g$ avec $\lambda = A\cos\varphi$ et $\mu = A\sin\varphi$.
En revanche, on peut montrer que la fonction $x \mapsto \sin(2x)$ n'est pas combinaison linéaire des fonctions f et g.

6.2 Sous espace vectoriel

$(E, +, .)$ désigne un $\mathbb{K}$-espace vectoriel.

6.2.1 Définition

Définition

On appelle sous-espace vectoriel de E toute partie F de E vérifiant
1) $F \neq \emptyset$;
2) $\forall \vec{x}, \vec{y} \in F, \vec{x} + \vec{y} \in F$ [stabilité par addition] ;
3) $\forall \vec{x} \in F, \forall \lambda \in \mathbb{K}, \lambda.\vec{x} \in F$ [stabilité par produit extérieur].

Exemple $\left\{\vec{0}\right\}$ et E sont des sous-espaces vectoriels de E dits triviaux.

Théorème

Si F est un sous-espace vectoriel de $(E, +, .)$ alors $(F, +, .)$ est un $\mathbb{K}$-espace vectoriel.

dém. :
Comme F est stable pour $+$ et ., on peut munir F des lois induites.
Par les propriétés 1), 2) et 3) avec $\lambda = -1$, on peut affirmer que F est un sous-groupe de $(E, +)$ et donc $(F, +)$ est un groupe abélien.
De plus les propriétés calculatoires engageant le produit extérieure étant vraies sur E, elle le sont aussi sur F.
□

Proposition

Une partie F de E est un sous-espace vectoriel si, et seulement si,
1) $\vec{0} \in F$;
2) $\forall \vec{x}, \vec{y} \in F, \forall \lambda, \mu \in \mathbb{K}, \lambda.\vec{x} + \mu.\vec{y} \in F$ [stabilité par combinaison linéaire].

dém. :
Si F est un sous-espace vectoriel alors F est un sous-groupe de $(E, +)$ et donc $\vec{0} \in F$.
De plus, F étant stable par addition et par produit extérieur, il l'est par combinaison linéaire.
Inversement, si F contient le vecteur nul et est stable par combinaison linéaire, il est évidemment non vide et stable par addition et produit extérieur.
□

Remarque Un sous-espace vectoriel contient toujours le vecteur nul.

Proposition

Si $\vec{e}_1, \ldots, \vec{e}_n$ sont des vecteurs d'un sous-espace vectoriel F alors tout vecteur combinaison linéaire des vecteurs $\vec{e}_1, \ldots, \vec{e}_n$ est encore élément de F.

dém. :
Par récurrence sur $n \in \mathbb{N}^\star$.
Pour $n = 1$.
Un vecteur combinaison linéaire de $\vec{e}_1 \in F$ s'écrit sous la forme $\lambda.\vec{e}_1$ et est donc élément de F par stabilité par produit extérieur.
Supposons la propriété établie au rang $n \geqslant 1$.
Soit $\vec{x}$ un vecteur combinaison linéaire de vecteurs $\vec{e}_1, \dots, \vec{e}_n, \vec{e}_{n+1} \in F$.
On peut écrire

$$\vec{x} = \lambda_1.\vec{e}_1 + \cdots + \lambda_n.\vec{e}_n + \lambda_{n+1}.\vec{e}_{n+1}$$

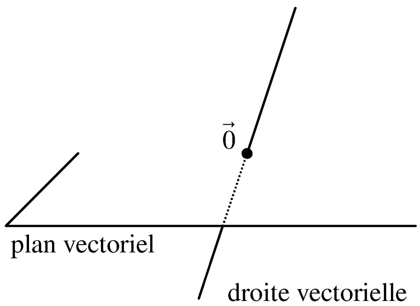
Par hypothèse de récurrence $\lambda_1.\vec{e}_1 + \cdots + \lambda_n.\vec{e}_n \in F$.
De plus $\lambda_{n+1}.\vec{e}_{n+1} \in F$ et par stabilité de F par addition

$$\vec{x} = (\lambda_1.\vec{e}_1 + \cdots + \lambda_n.\vec{e}_n) + (\lambda_{n+1}.\vec{e}_{n+1}) \in F$$

Récurrence établie.
□

6.2.2 Exemples

Exemple Géométriquement, les sous-espaces vectoriels non triviaux se visualisent comme étant des droites ou des plans passant tous par le vecteur $\vec{0}$.



Exemple Montrons que

$$D = \{(a, 2a, 3a)/a \in \mathbb{R}\}$$

est un sous-espace vectoriel de $\mathbb{R}^3$.
$\mathbb{R}^3$ est un $\mathbb{R}$-espace vectoriel de vecteur nul $\vec{0} = (0, 0, 0)$.
$D \subset \mathbb{R}^3$ et $\vec{0} = (a, 2a, 3a) \in D$ avec $a = 0 \in \mathbb{R}$.
Soient $\lambda, \mu \in \mathbb{R}$ et $\vec{x}, \vec{y} \in D$.
On peut écrire $\vec{x} = (a, 2a, 3a)$ et $\vec{y} = (b, 2b, 3b)$ avec $a, b \in \mathbb{R}$.
On a alors

$$\lambda.\vec{x} + \mu.\vec{y} = (\lambda a + \mu b, 2\lambda a + 2\mu b, 3\lambda a + 3\mu b) = (c, 2c, 3c) \in D$$

avec $c = \lambda a + \mu b \in \mathbb{R}$.

Exemple Montrons que

$$H = \{(x_1, ..., x_n) \in \mathbb{R}^n / x_1 + \cdots + x_n = 0\}$$

est un sous-espace vectoriel de $\mathbb{R}^n$.
$\mathbb{R}^n$ est un $\mathbb{R}$-espace vectoriel de vecteur nul $\vec{0} = (0, \ldots, 0)$.
$H \subset \mathbb{R}^n$ et $\vec{0} = (0, \ldots, 0) \in H$ car $0 + \cdots + 0 = 0$.
Soient $\lambda, \mu \in \mathbb{R}$ et $\vec{x} = (x_1, \ldots, x_n), \vec{y} = (y_1, \ldots, y_n) \in H$.
On a $\lambda.\vec{x} + \mu.\vec{y} = (\lambda x_1 + \mu y_1, \ldots, \lambda x_n + \mu y_n)$
Or

$$(\lambda x_1 + \mu y_1) + \cdots + (\lambda x_n + \mu y_n) = \lambda(x_1 + \cdots + x_n) + \mu(y_1 + \cdots + y_n) = 0$$

car $x_1 + \cdots + x_n = y_1 + \cdots + y_n = 0$ puisque $\vec{x}, \vec{y} \in H$
donc $\lambda.\vec{x} + \mu.\vec{y} \in H$.

Exemple Soient I un intervalle non singulier de $\mathbb{R}$ et $n \in \mathbb{N} \cup \{\infty\}$.
Montrons que $\mathcal{C}^n(I, \mathbb{R})$ est un sous-espace vectoriel de $\mathcal{F}(I, \mathbb{R})$.
$\mathcal{F}(I, \mathbb{R})$ est un $\mathbb{R}$-espace vectoriel de vecteur nul la fonction nulle $\tilde{0} : x \in I \mapsto 0$.
$\mathcal{C}^n(I, \mathbb{R}) \subset \mathcal{F}(I, \mathbb{R})$ et $\tilde{0} \in \mathcal{C}^n(I, \mathbb{R})$ car il est connu que la fonction nulle est de classe $\mathcal{C}^n$.
Soient $\lambda, \mu \in \mathbb{R}$ et $f, g \in \mathcal{C}^n(I, \mathbb{R})$.
Par opérations sur les fonctions de classe $\mathcal{C}^n$, on peut affirmer que la fonction $\lambda.f + \mu.g$ est de classe $\mathcal{C}^n$ et donc $\lambda.f + \mu.g \in \mathcal{C}^n(I, \mathbb{R})$.

Exemple Dans $E = \mathcal{C}^0([0, 1], \mathbb{R})$, montrons que

$$F = \left\{ f \in E / \int_0^1 f(t) \,\mathrm{d}t = 0 \right\}$$

est un sous-espace vectoriel.
$E = \mathcal{C}^0([0, 1], \mathbb{R})$ est un $\mathbb{R}$-espace vectoriel de vecteur nul la fonction nulle $\tilde{0} : x \in [0, 1] \to 0$.
$F \subset E$ et $\tilde{0} \in F$ car la fonction nulle est d'intégrale nulle sur $[0, 1]$.
Soient $\lambda, \mu \in \mathbb{R}$ et $f, g \in F$.

$$\int_0^1 (\lambda.f + \mu.g)(t) \,\mathrm{d}t = \int_0^1 \lambda f(t) + \mu g(t) \,\mathrm{d}t = \lambda \int_0^1 f(t) \,\mathrm{d}t + \mu \int_0^1 g(t) \,\mathrm{d}t = 0$$

car f et g sont d'intégrales nulles sur $[0, 1]$.
Ainsi $\lambda.f + \mu.g \in F$.

Remarque Les parties suivantes ne sont pas des sous-espaces vectoriels de $\mathbb{R}^2$:
- $\left\{(x, y) \in \mathbb{R}^2 / x + y = 1\right\}$ car ne contient pas le vecteur nul ;
- $\left\{(x, y) \in \mathbb{R}^2 / xy = 0\right\}$ car non stable par addition ;
- $\left\{(x, y) \in \mathbb{R}^2 / x + y \in \mathbb{Z}\right\}$ car non stable par produit extérieur.

6.2.3 Opérations sur les sous-espaces vectoriels

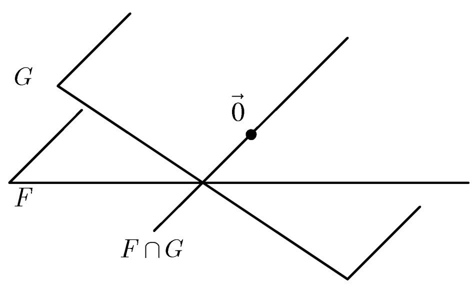
6.2.3.1 Intersection

Proposition

Si F et G sont des sous-espaces vectoriels de E alors $F \cap G$ est un sous-espace vectoriel de E.

dém. :
$F \cap G$ est une partie de E.
$\vec{0} \in F$ et $\vec{0} \in G$ donc $\vec{0} \in F \cap G$.
Soient $\lambda, \mu \in \mathbb{K}$ et $\vec{x}, \vec{y} \in F \cap G$.
$\lambda.\vec{x} + \mu.\vec{y} \in F$ car F est un sous-espace vectoriel et $\vec{x}, \vec{y} \in F$.
Aussi $\lambda.\vec{x} + \mu.\vec{y} \in G$ par des arguments semblables.
Par suite $\lambda.\vec{x} + \mu.\vec{y} \in F \cap G$.
□

Exemple



Proposition

Une intersection (finie ou infinie) de sous-espaces vectoriels d'un espace E est un sous-espace vectoriel de E.

dém. :
Soient $(F_i)_{i \in I}$ une famille de sous-espaces vectoriels de E et $F = \bigcap_{i \in I} F_i$.
F est une partie de E.
Pour tout $i \in I$, $\vec{0} \in F_i$ car F_i est un sous-espace vectoriel donc $\vec{0} \in F$.
Soient $\lambda, \mu \in \mathbb{K}$ et $\vec{x}, \vec{y} \in F$.
Pour tout $i \in I$, $\lambda.\vec{x} + \mu.\vec{y} \in F_i$ car F_i est un sous-espace vectoriel de E et $\vec{x}, \vec{y} \in F_i$.
Par suite $\lambda.\vec{x} + \mu.\vec{y} \in F$.
□

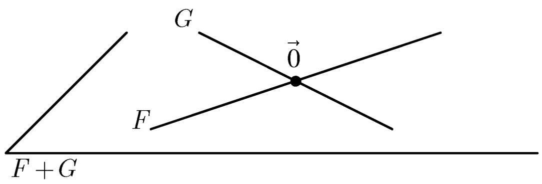
6.2.3.2 Somme de sous-espaces vectoriels

Attention : La réunion de deux sous-espaces vectoriels n'est généralement pas un sous-espace vectoriel.

Définition

On appelle somme de deux sous-espaces vectoriels F et G de E l'ensemble

$$F + G = \left\{\vec{a} + \vec{b} / \vec{a} \in F, \vec{b} \in G\right\}$$



Proposition

La somme $F + G$ de deux sous-espaces vectoriels F et G de E est un sous-espace vectoriel de E.
De plus, $F + G$ contient F et G et est inclus dans tout sous-espace vectoriel contenant F et G.

dém. :
$F + G$ est une partie de E.
$\vec{0} = \vec{0} + \vec{0}$ avec $\vec{0} \in F$ et $\vec{0} \in G$ donc $\vec{0} \in F + G$.
Soient $\lambda, \mu \in \mathbb{K}$ et $\vec{x}, \vec{y} \in F + G$.
On peut écrire $\vec{x} = \vec{a} + \vec{b}$ et $\vec{y} = \vec{c} + \vec{d}$ avec $\vec{a}, \vec{c} \in F$ et $\vec{b}, \vec{d} \in G$
On a alors $\lambda.\vec{x} + \mu.\vec{y} = (\lambda.\vec{a} + \mu.\vec{c}) + (\lambda.\vec{b} + \mu.\vec{d})$ avec $\lambda.\vec{a} + \mu.\vec{c} \in F$ et $(\lambda.\vec{b} + \mu.\vec{d}) \in G$.
Ainsi $\lambda.\vec{x} + \mu.\vec{y} \in F + G$.
De plus, $F \subset F + G$ car pour tout $\vec{x} \in F$, on peut écrire $\vec{x} = \vec{x} + \vec{0}$ avec $\vec{0} \in G$.
De même $G \subset F + G$.
Enfin, si un sous-espace vectoriel H de E contient F et G alors pour tout $\vec{a} \in F$ et tout $\vec{b} \in G$, on a $\vec{a} + \vec{b} \in H$ car $\vec{a}, \vec{b} \in H$ et donc $F + G \subset H$.
□

Remarque $F + G$ se comprend comme étant le plus petit sous-espace vectoriel contenant à la fois F et G.

Remarque Pour F, G, H trois sous-espaces vectoriels de E, on vérifie aisément les propriétés :
- $F + G = G + F$;
- $(F + G) + H = F + (G + H)$;
- $F + \left\{\vec{0}\right\} = F$, $F + E = E$ et $F + F = F$.

6.2.4 Sous-espaces vectoriels supplémentaires

Définition

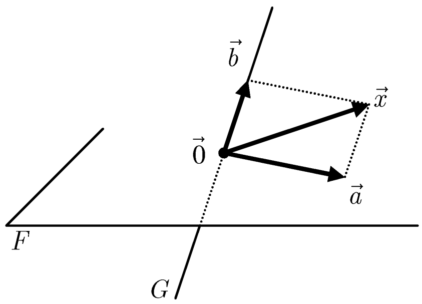
On dit que deux sous-espaces vectoriels F et G de E sont supplémentaires si

$$F \cap G = \left\{\vec{0}\right\} \text{ et } F + G = E$$

Théorème

Si F et G sont deux sous-espaces vectoriels supplémentaires de E alors

$$\forall \vec{x} \in E, \exists!(\vec{a}, \vec{b}) \in F \times G, \vec{x} = \vec{a} + \vec{b}$$



dém. :
Soit $\vec{x} \in E$.
Existence : Puisque $E = F + G$ tout vecteur de E peut se voir comme la somme d'un vecteur de F et d'un vecteur de G dont il existe $\vec{a} \in F$ et $\vec{b} \in G$ tel que $\vec{x} = \vec{a} + \vec{b}$.
Unicité : Supposons $\vec{x} = \vec{a} + \vec{b} = \vec{c} + \vec{d}$ avec $\vec{a}, \vec{c} \in F$ et $\vec{b}, \vec{d} \in G$.
On a $\vec{a} - \vec{c} = \vec{d} - \vec{b}$ or $\vec{a} - \vec{c} \in F$ et $\vec{d} - \vec{c} \in G$ car F et G sont des sous-espaces vectoriels.
Par suite $\vec{a} - \vec{c} = \vec{d} - \vec{b} \in F \cap G$, or $F \cap G = \left\{\vec{0}\right\}$ donc $\vec{a} = \vec{c}$ et $\vec{b} = \vec{d}$.
□

Exemple E et $\left\{\vec{0}\right\}$ sont des sous-espaces vectoriels supplémentaires de E.

Exemple Soient

$$H = \{(x_1, \dots, x_n) \in \mathbb{R}^n / x_1 + \dots + x_n = 0\} \text{ et } D = \{(\alpha, \dots, \alpha) / \alpha \in \mathbb{R}\}$$

H et D sont des sous-espaces vectoriels de $\mathbb{R}^n$, montrons qu'ils sont supplémentaires.
Etudions $H \cap D$.
Soit $\vec{x} = (x_1, \dots, x_n) \in H \cap D$.
On a $x_1 + \dots + x_n = 0$ car $\vec{x} \in H$ et $x_1 = \dots = x_n$ car $\vec{x} \in D$.
Par suite $x_1 + \dots + x_n = nx_1 = 0$ donc $x_1 = 0$ puis $\vec{x} = \vec{0}$.
Ainsi $H \cap D \subset \{\vec{o}\}$.
L'autre inclusion est toujours vraie (on ne l'évoque pas toujours) donc $H \cap D = \left\{\vec{0}\right\}$.
Montrons que $H + D = E$ i.e. :

$$\forall \vec{x} \in \mathbb{R}^n, \exists(a, b) \in H \times D, \vec{x} = \vec{a} + \vec{b}$$

Pour justifier l'existence des vecteurs $\vec{a}$ et $\vec{b}$, on va exhiber ceux-ci en procédant à un raisonnement par analyse/synthèse.
Analyse :
Soit $\vec{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$.
Supposons $\vec{x} = \vec{a} + \vec{b}$ avec $\vec{a} = (a_1, \dots, a_n) \in H$ et $\vec{b} = (\alpha, \dots, \alpha) \in D$.
Déterminons $\vec{a}$ et $\vec{b}$ en fonction de $\vec{x}$.

$\vec{x} = \vec{a} + \vec{b}$ donne $(x_1, \ldots, x_n) = (a_1 + \alpha, \ldots, a_n + \alpha)$
Donc $x_1 + \cdots + x_n = n\alpha$ car $\vec{a} \in H$.
Par suite

$$\alpha = \frac{1}{n}(x_1 + \cdots + x_n)$$

Ceci détermine α, donc $\vec{b} = (\alpha, \ldots, \alpha)$ puis $\vec{a} = \vec{x} - \vec{b}$.
Synthèse :
Soit $\vec{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$.
Posons

$$\alpha = \frac{1}{n}(x_1 + \cdots + x_n), b = (\alpha, \ldots, \alpha) \text{ et } \vec{a} = \vec{x} - \vec{b}$$

On a immédiatement $\vec{b} \in D$ et $\vec{x} = \vec{a} + \vec{b}$.
De plus $\vec{a} = (x_1 - \alpha, \ldots, x_n - \alpha) \in H$ puisque $(x_1 - \alpha) + \cdots + (x_n - \alpha) = 0$.
Ainsi $E \subset H + D$ puis $E = H + D$.
Finalement H et D sont des sous-espaces vectoriels supplémentaires de E.

Exemple Soient

$$\mathcal{P} = \{f : \mathbb{R} \to \mathbb{R}/f \text{ paire}\} \text{ et } \mathcal{I} = \{f : \mathbb{R} \to \mathbb{R}/f \text{ impaire}\}$$

Montrons que $\mathcal{P}$ et $\mathcal{I}$ sont des sous-espaces vectoriels supplémentaires de l'espace $E = \mathcal{F}(\mathbb{R}, \mathbb{R})$.
On montre facilement que $\mathcal{P}$ et $\mathcal{I}$ sont des sous-espaces vectoriels de E.
Etudions $\mathcal{P} \cap \mathcal{I}$.
Soit $f \in \mathcal{P} \cap \mathcal{I}$.
Pour tout $x \in \mathbb{R}$, on a $f(-x) = f(x)$ car $f \in \mathcal{P}$ et $f(-x) = -f(x)$ car $f \in \mathcal{I}$.
Par suite $f(x) = 0$ et puisque ceci vaut pour tout $x \in \mathbb{R}$, on obtient $f = \tilde{0}$.
Ainsi $\mathcal{P} \cap \mathcal{I} \subset \{\tilde{0}\}$ puis $\mathcal{P} \cap \mathcal{I} = \{\tilde{0}\}$.
Montrons $\mathcal{P} + \mathcal{I} = E$.
Analyse :
Soit $f \in E$ telle que $f = g + h$ avec $g \in \mathcal{P}$ et $h \in \mathcal{I}$. Déterminons g et h en fonction de f.
Pour tout $x \in \mathbb{R}$, on a $f(x) = g(x) + h(x)$ et $f(-x) = g(-x) + h(-x) = g(x) - h(x)$.
Ainsi on a le système

$$\begin{cases} f(x) = g(x) + h(x) \\ f(x) = g(x) - h(x) \end{cases}$$

On en déduit

$$g(x) = \frac{1}{2}\left(f(x) + f(-x)\right) \text{ et } h(x) = \frac{1}{2}\left(f(x) - f(-x)\right)$$

ce qui déterminer g et h.
Synthèse :
Soient $f \in E$ et $g, h \in E$ les fonctions définies par

$$g(x) = \frac{1}{2}\left(f(x) + f(-x)\right) \text{ et } h(x) = \frac{1}{2}\left(f(x) - f(-x)\right)$$

On vérifie aisément $f(x) = g(x) + h(x)$, $g(-x) = g(x)$ et $h(-x) = -h(x)$ pour tout $x \in \mathbb{R}$.
Ainsi $f = g + h$ avec $g \in \mathcal{P}$ et $h \in \mathcal{I}$.
On a donc établi $E \subset F + G$ puis $E = F + G$.
Finalement $\mathcal{P}$ et $\mathcal{I}$ sont des sous-espaces vectoriels supplémentaires de E.

Remarque En conséquence, toute fonction $f : \mathbb{R} \to \mathbb{R}$ s'écrit de manière unique comme somme d'une fonction paire et d'une fonction impaire appelées les parties paire et impaire de f.
Par exemple, la fonction $x \mapsto \mathrm{e}^x$ se décompose en la somme de sa partie paire, qui est la fonction $x \mapsto \mathrm{ch}x$, et de sa partie impaire, $x \mapsto \mathrm{sh}x$.

6.2.5 Espace vectoriel engendré par une partie

Théorème

Etant donnée une partie A de E, il existe un unique sous-espace vectoriel de E, noté Vect$A$ vérifiant :
1) $A \subset \mathrm{Vect}A$;
2) Vect$A$ est inclus dans tout sous-espace vectoriel contenant A.
Le sous-espace vectoriel Vect$A$ se comprend comme étant le plus petit sous-espace vectoriel contenant A, on l'appelle espace vectoriel engendré par A.

dém. :
Unicité :
Soient G et G' deux sous-espaces vectoriels de E vérifiant les propriétés 1) et 2)
Puisque A est inclus dans le sous-espace vectoriel G' et puisque G est inclus dans tout sous-espace vectoriel contenant A, on a déjà $G \subset G'$. De façon symétrique on a aussi $G \subset G'$ et donc $G = G'$.
Existence :
Soient

$$\mathcal{S} = \{F \text{ sous - espace vectoriel de } E/A \subset F\}$$

et

$$\mathrm{Vect}A = \bigcap_{F \in \mathcal{S}} F$$

Puisque qu'une intersection de sous-espaces vectoriels est un sous-espace vectoriel, l'ensemble Vect$A$ est un sous-espace vectoriel de E. De plus, puisque Vect$A$ est une intersection de parties contenant toutes A, Vect$A$ contient aussi A. Enfin, si F est un sous-espace vectoriel de E contenant A alors $F \in \mathcal{S}$ et donc $\mathrm{Vect}A \subset F$ car Vect$A$ est l'intersection de toutes les sous-espaces vectoriels éléments de $\mathcal{S}$.
□

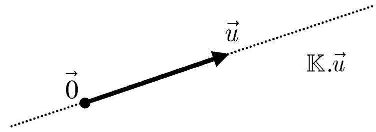
Exemple $\mathrm{Vect}\emptyset = \left\{\vec{0}\right\}$ car l'espace nul est le plus petit sous-espace vectoriel de E.
$\mathrm{Vect}E = E$ car Vect$E$ est un sous-espace vectoriel contenant E.

Exemple Soit $A = \{\vec{u}\}$. Montrons par double inclusion

$$\mathrm{Vect}\,\{\vec{u}\} = \{\lambda.\vec{u}/\lambda \in \mathbb{K}\} = \mathbb{K}.\vec{u}$$

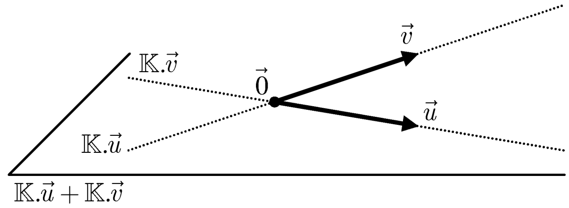
Puisque $\vec{u} \in A \subset \mathrm{Vect}A$ et puisque Vect$A$ est un sous-espace vectoriel, on a $\lambda.\vec{u} \in \mathrm{Vect}A$ pour tout $\lambda \in \mathbb{K}$. Ainsi $\mathbb{K}.\vec{u} \subset \mathrm{Vect}A$.
Inversement, on vérifie aisément que $\mathbb{K}.\vec{u}$ est un sous-espace vectoriel de E et puisque $\vec{u}$ est élément de $\mathbb{K}.\vec{u}$, on a $A \subset \mathbb{K}.\vec{u}$ puis $\mathrm{Vect}A \subset \mathbb{K}.\vec{u}$.
Par double inclusion, on obtient

$$\mathrm{Vect}\,\{\vec{u}\} = \mathbb{K}.\vec{u}$$



Exemple Soient $A = \{\vec{u}, \vec{v}\}$. Par double inclusion, on montre comme ci-dessus que

$$\text{Vect}\,\{\vec{u}, \vec{v}\} = \{\lambda\vec{u} + \mu\vec{v}/\lambda, \mu \in \mathbb{K}\} = \mathbb{K}.\vec{u} + \mathbb{K}.\vec{v}$$



Proposition

Si A et B sont deux parties de E alors

$$A \subset B \Rightarrow \text{Vect}A \subset \text{Vect}B$$

dém. :
Supposons $A \subset B$.
On a alors $A \subset \text{Vect}B$ or $\text{Vect}B$ est un sous-espace vectoriel donc $\text{Vect}A \subset \text{Vect}B$
□

Proposition

Si A et B sont deux parties de E alors

$$\text{Vect}(A \cup B) = \text{Vect}A + \text{Vect}B$$

dém. :
$A \subset \text{Vect}A$ donc $A \subset \text{Vect}A + \text{Vect}B$. De même, $B \subset \text{Vect}A + \text{Vect}B$ donc $A \cup B \subset \text{Vect}A + \text{Vect}B$.
Or $\text{Vect}A + \text{Vect}B$ est un sous-espace vectoriel donc $\text{Vect}(A \cup B) \subset \text{Vect}A + \text{Vect}B$.
Inversement $A \subset A \cup B$ donc $\text{Vect}A \subset \text{Vect}(A \cup B)$. Aussi $\text{Vect}B \subset \text{Vect}(A \cup B)$. Or $\text{Vect}(A \cup B)$ est un sous-espace vectoriel donc $\text{Vect}A + \text{Vect}B \subset \text{Vect}(A \cup B)$.
□

Proposition

Si F est un sous-espace vectoriel de E alors $\text{Vect}F = F$.

dém. :
$F \subset \text{Vect}F$ et puisque F est un sous-espace vectoriel contenant F, on a aussi $\text{Vect}F \subset F$.
□

Exemple Pour F et G sous-espaces vectoriels

$$\text{Vect}(F \cup G) = F + G$$

Ainsi $F + G$ apparaît comme étant le plus petit sous-espace vectoriel contenant F et G.

6.3 Applications linéaires

E, F et G désignent des $\mathbb{K}$-espaces vectoriels.

6.3.1 Définition

Définition

On dit qu'une application $f : E \to F$ est linéaire (ou est un morphisme d'espace vectoriel) si :
1) $\forall \vec{x}, \vec{y} \in E, f(\vec{x} + \vec{y}) = f(\vec{x}) + f(\vec{y})$;
2) $\forall \lambda \in \mathbb{K}, \forall \vec{x} \in E, f(\lambda.\vec{x}) = \lambda.f(\vec{x})$.
On note $\mathcal{L}(E, F)$ l'ensemble de ces applications.

Proposition

Soit $f : E \to F$. On a équivalence entre :
(i) f est une application linéaire ;
(ii) $\forall \lambda, \mu \in \mathbb{K}, \forall \vec{x}, \vec{y} \in E, f(\lambda.\vec{x} + \mu.\vec{y}) = \lambda.f(\vec{x}) + \mu.f(\vec{y})$.

dém. :
(i) $\Rightarrow$ (ii) Supposons (i)
Soient $\lambda, \mu \in \mathbb{K}$ et $\vec{x}, \vec{y} \in E$.
Par 1) $f(\lambda.\vec{x} + \mu.\vec{y}) = f(\lambda.\vec{x}) + f(\mu.\vec{y})$
puis par 2) $f(\lambda.\vec{x}) + f(\mu.\vec{y}) = \lambda.f(\vec{x}) + \mu.f(\vec{y})$
et ainsi $f(\lambda.\vec{x} + \mu.\vec{y}) = \lambda.f(\vec{x}) + \mu.f(\vec{y})$.
(ii) $\Rightarrow$ (i) car (ii) entraîne 1) en prenant $\lambda = \mu = 1$ et (ii) entraîne 2) en prenant $\mu = 0$.
□

Exemple Soit $f : E \to F$ définie par $f : \vec{x} \mapsto \vec{0}_F$.
f est une application linéaire.
En effet, pour tout $\lambda, \mu \in \mathbb{K}$ et $\vec{x}, \vec{y} \in F$, $f(\lambda.\vec{x} + \mu.\vec{y}) = \vec{0}_F = \lambda.f(\vec{x}) + \mu.f(\vec{y})$.
L'application f est appelée application linéaire nulle.

Exemple Montrons que $f : \mathbb{R}^2 \to \mathbb{R}^3$ définie par $f(x, y) = (x + y, x - y, 2y)$ est une application linéaire.
Soient $\lambda, \mu \in \mathbb{K}$ et $\vec{a}, \vec{b} \in \mathbb{R}^2$, $\vec{a} = (x, y)$, $\vec{b} = (x', y')$.
$f(\lambda.\vec{a} + \mu.\vec{b}) = f(\lambda x + \mu x', \lambda y + \mu y')$
donc $f(\lambda.\vec{a} + \mu.\vec{b}) = (\lambda x + \mu x' + \lambda y + \mu y', \lambda x + \mu x' - \lambda y - \mu y', 2\lambda y + 2\mu y')$
puis $f(\lambda.\vec{a} + \mu.\vec{b}) = \lambda(x + y, x - y, 2y) + \mu(x' + y', x' - y', 2y') = \lambda f(\vec{a}) + \mu f(\vec{b})$

Exemple Soient $E = \mathcal{C}^1(\mathbb{R}, \mathbb{R})$ et $F = \mathcal{C}^0(\mathbb{R}, \mathbb{R})$.
Montrons que l'application de dérivation $D : E \to F$ définie par $D : f \mapsto f'$ est linéaire.
Soient $\lambda, \mu \in \mathbb{R}$ et $f, g \in E$.

$$D(\lambda.f + \mu.g) = (\lambda.f + \mu.g)' = \lambda.f' + \mu.g' = \lambda.D(f) + \mu.D(g)$$

Remarque En cas d'ambiguïtés, il peut être nécessaire de préciser le corps de base des espaces vectoriels E et F entre lesquels opère l'application, on parle alors d'application $\mathbb{K}$-linéaire et on note $\mathcal{L}_{\mathbb{K}}(E, F)$ l'ensemble de ces applications.

Exemple Soient $E = F = \mathbb{C}$ vu comme $\mathbb{R}$-espace vectoriel.
Montrons que l'application conjugaison $f : \mathbb{C} \to \mathbb{C}$ définie par $f : z \mapsto \bar{z}$ est $\mathbb{R}$-linéaire.
Soient $\lambda, \mu \in \mathbb{R}$ et $z, z' \in \mathbb{C}$.

$$f(\lambda z + \mu z') = \overline{\lambda z + \mu z'} = \bar{\lambda}\bar{z} + \bar{\mu}\bar{z}' = \lambda\bar{z} + \mu\bar{z}' = \lambda f(z) + \mu f(z')$$

En revanche, on peut observer que f n'est pas $\mathbb{C}$-linéaire.

Proposition

Si $f : E \to F$ est linéaire alors $f(\vec{0}_E) = \vec{0}_F$.

dém. :
$f(\vec{0}_E) = f(\vec{0}_E + \vec{0}_E) = f(\vec{0}_E) + f(\vec{0}_E)$.
En ajoutant $-f(\vec{0}_E)$ de part et d'autre, on obtient $\vec{0}_F = f(\vec{0}_E)$.
□

Proposition

Si $f : E \to F$ est une application linéaire et si $\vec{e}_1, \dots, \vec{e}_n$ sont des vecteurs de E alors

$$\forall \lambda_1, \dots, \lambda_n \in \mathbb{K}, f(\lambda_1\vec{e}_1 + \cdots + \lambda_n\vec{e}_n) = \lambda_1 f(\vec{e}_1) + \cdots + \lambda_n f(\vec{e}_n)$$

Ainsi, l'image d'un combinaison linéaire est la combinaison linéaire des images.

dém. :
Par récurrence sur $n \in \mathbb{N}^\star$.
Pour $n = 1$, la propriété est immédiate par définition de la linéarité.
Supposons la propriété vraie au rang $n \geqslant 1$.
Soient $\vec{e}_1, \dots, \vec{e}_n, \vec{e}_{n+1}$ des vecteurs de E et $\lambda_1, \dots, \lambda_n, \lambda_{n+1} \in \mathbb{K}$.
Par image d'une somme

$$f(\lambda_1\vec{e}_1 + \cdots + \lambda_n\vec{e}_n + \lambda_{n+1}\vec{e}_{n+1}) = f(\lambda_1\vec{e}_1 + \cdots + \lambda_n\vec{e}_n) + f(\lambda_{n+1}\vec{e}_{n+1})$$

Par hypothèse de récurrence et image d'un produit extérieur

$$f(\lambda_1\vec{e}_1 + \cdots + \lambda_n\vec{e}_n + \lambda_{n+1}\vec{e}_{n+1}) = f(\lambda_1\vec{e}_1) + \cdots + f(\lambda_n\vec{e}_n) + f(\lambda_{n+1}\vec{e}_{n+1})$$

Récurrence établie.
□

Proposition

Si $f : E \to F$ et $g : F \to G$ sont linéaires alors $g \circ f : E \to F$ est aussi linéaire.

dém. :
Soient $\lambda, \mu \in \mathbb{K}$ et $\vec{x}, \vec{y} \in E$.

$$(g \circ f)(\lambda.\vec{x} + \mu.\vec{y}) = g\left(f(\lambda.\vec{x} + \mu.\vec{y})\right)$$

Par linéarité de f

$$(g \circ f)(\lambda.\vec{x} + \mu.\vec{y}) = g\left(\lambda.f(\vec{x}) + \mu.f(\vec{y})\right)$$

Par linéarité de g

$$(g \circ f)(\lambda.\vec{x} + \mu.\vec{y}) = \lambda.g\left(f(\vec{x})\right) + \mu.g\left(f(\vec{y})\right)$$

et donc

$$(g \circ f)(\lambda.\vec{x} + \mu.\vec{y}) = \lambda.(g \circ f)(\vec{x}) + \mu.(g \circ f)(\vec{y})$$

□

6.3.2 Application linéaires particulières

6.3.2.1 Formes linéaires

Définition

On appelle forme linéaire sur un $\mathbb{K}$-espace vectoriel E, toute application linéaire de E vers $\mathbb{K}$.
On note $E^\star$, au lieu de $\mathcal{L}(E, \mathbb{K})$, l'ensemble des formes linéaires sur E.
$E^\star$ est appelé dual de E.

Exemple Pour $a_1, ..., a_n \in \mathbb{K}$ fixé, l'application $f : \mathbb{K}^n \to \mathbb{K}$ définie par

$$f : (x_1, \dots, x_n) \mapsto a_1 x_1 + \dots + a_n x_n$$

est une forme linéaire sur $\mathbb{K}^n$.
En effet, c'est une application de $\mathbb{K}^n$ vers $\mathbb{K}$ et c'est aussi une application linéaire car on vérifie aisément $f(\lambda.\vec{x} + \mu.\vec{y}) = \lambda.f(\vec{x}) + \mu f(\vec{y})$.

Exemple Soient $a \leqslant b \in \mathbb{R}$ et $I : \mathcal{C}([a, b], \mathbb{C}) \to \mathbb{C}$ l'application définie par

$$I(f) = \int_a^b f(t)\,\mathrm{d}t$$

L'application I est une forme linéaire sur $\mathcal{C}([a, b], \mathbb{C})$ car c'est une application à valeurs dans $\mathbb{C}$ et c'est une application linéaire car

$$I(\lambda.f + \mu.g) = \int_a^b \lambda f(t) + \mu g(t)\,\mathrm{d}t = \lambda \int_a^b f(t)\,\mathrm{d}t + \mu \int_a^b g(t)\,\mathrm{d}t = \lambda I(f) + \mu I(g)$$

6.3.2.2 Endomorphisme

Définition

On appelle endomorphisme de E, toute application linéaire de E dans lui-même.
On note $\mathcal{L}(E)$, au lieu de , $\mathcal{L}(E, E)$ l'ensemble des endomorphismes de E.

Exemple L'application identité $\mathrm{Id}_E : E \to E$ est un endomorphisme de E.
C'est effectivement une application linéaire de E vers E.
L'endomorphisme identité est souvent noté Id ou encore I.

Exemple Soient $a, b, c, d \in \mathbb{R}$ fixés et $f : \mathbb{R}^2 \to \mathbb{R}^2$ l'application définie par

$$f : (x, y) \mapsto (ax + by, cx + dy)$$

On vérifie aisément que f est une application linéaire et puisque celle-ci est définie de $\mathbb{R}^2$ vers $\mathbb{R}^2$, c'est un endomorphisme de $\mathbb{R}^2$.

Exemple Soit $D : \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}) \to \mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ définie par $D(f) = f'$.
L'application D est bien défini car la dérivée d'une fonction de classe $\mathcal{C}^\infty$ est $\mathcal{C}^\infty$.
De plus

$$D(\lambda.f + \mu.g) = (\lambda.f + \mu.g)' = \lambda.f' + \mu.g' = \lambda.D(f) + \mu.D(g)$$

donc D est linéaire et c'est alors un endomorphisme de $\mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$.

Proposition

Si f et g sont des endomorphismes E alors la composée $g \circ f$ aussi.

dém. :
La composée de deux applications de E vers E est une application de E vers E et la composée de deux applications linéaires est linéaire.
□

6.3.2.3 Isomorphisme

Définition

On appelle isomorphisme d'un $\mathbb{K}$-espace vectoriel E vers un $\mathbb{K}$-espace vectoriel F toute application linéaire bijective de E vers F.

Exemple L'application $f : \mathbb{R}^2 \to \mathbb{C}$ définie par $f(a, b) = a + i.b$ est un isomorphisme de $\mathbb{R}$-espace vectoriels.
En effet cette application est $\mathbb{R}$-linéaire et bijective.

Proposition

Si $f : E \to F$ et $g : F \to G$ sont des isomorphismes alors la composée $g \circ f : E \to G$ est un isomorphisme.

dém. :
La composée d'une bijection de E vers F et d'une bijection de F vers G est une bijection de E vers G et la composée de deux applications linéaires est une application linéaire.
□

Proposition

Si $f : E \to F$ est un isomorphisme alors son application réciproque $f^{-1} : F \to E$ est un isomorphisme.

dém. :
L'application réciproque d'une bijection de E vers F est une bijection de F vers E.
Pour conclure il suffit alors de justifier que l'application f^{-1} est linéaire.
Soient $\lambda, \mu \in \mathbb{K}$ et $\vec{y}, \vec{y}' \in F$.

Posons les éléments $\vec{x}, \vec{x}' \in E$ déterminés par $\vec{x} = f^{-1}(\vec{y})$ et $\vec{x}' = f^{-1}(\vec{y}')$ de sorte que $f(\vec{x}) = \vec{y}$ et $f(\vec{x}') = \vec{y}'$.
Par linéarité de f,

$$f(\lambda.\vec{x} + \mu.\vec{x}') = \lambda.f(\vec{x}) + \mu.f(\vec{x}') = \lambda.\vec{y} + \mu.\vec{y}'$$

Par bijectivité de f,

$$\lambda.\vec{x} + \mu.\vec{x}' = f^{-1}(\lambda.\vec{y} + \mu.\vec{y'})$$

et ainsi

$$\lambda.f^{-1}(\vec{y}) + \mu.f^{-1}(\vec{y}') = f^{-1}(\lambda.\vec{y} + \mu.\vec{y}')$$

Finalement f^{-1} est linéaire.
□

Exemple L'application $g : \mathbb{C} \to \mathbb{R}^2$ définie par $g : z \mapsto (\mathrm{Re}(z), \mathrm{Im}(z))$ est l'isomorphisme réciproque de l'application $f : (a, b) \in \mathbb{R}^2 \mapsto a + ib \in \mathbb{C}$

Définition

Deux $\mathbb{K}$-espace vectoriels sont dits isomorphes s'il existe un isomorphisme entre ceux-ci.

Exemple $\mathbb{R}^2$ et $\mathbb{C}$ sont des $\mathbb{R}$-espaces vectoriels isomorphes.

6.3.2.4 Automorphisme

Définition

On appelle automorphisme de E, toute permutation linéaire de E i.e. toute bijection linéaire de E vers E. On note $\mathrm{GL}(E)$ l'ensemble des automorphismes de E.

Exemple L'identité est un automorphisme de E.

Proposition

Si f et g sont des automorphismes de E alors la composée $f \circ g$ est un automorphisme de E.

dém. :
La composée de deux permutations de E est une permutation de E et la composée de deux applications linéaires est linéaire.
□

Proposition

Si f est un automorphisme de E alors son application réciproque f^{-1} est un automorphisme de E.

dém. :
L'application f^{-1} est un permutation de E et est linéaire car isomorphisme réciproque de f.
□

Remarque L'ensemble $\mathrm{GL}(E)$ apparaît donc comme un sous-groupe du groupe $(\mathfrak{S}(E), \circ)$ des permutations de E.

Définition

$(\mathrm{GL}(E), \circ)$ est appelé groupe linéaire de E.

6.3.3 Noyau et image d'une application linéaire

Théorème

Soit $f : E \to F$ une application linéaire.
Si V est une sous-espace vectoriel de E alors $f(V)$ est un sous-espace vectoriel de F.
Si W est un sous-espace vectoriel de F alors $f^{-1}(W)$ est un sous-espace vectoriel de E.

dém. :
Soit V un sous-espace vectoriel de E.
$f(V) \subset F$ et $\vec{0}_F = f(\vec{0}_E) \in f(V)$.
Soient $\lambda, \mu \in \mathbb{K}$ et $\vec{x}, \vec{y} \in f(V)$. On peut écrire $\vec{x} = f(\vec{a})$ et $\vec{y} = f(\vec{b})$ avec $\vec{a}, \vec{b} \in V$.
On a alors $\lambda.\vec{x} + \mu.\vec{y} = f(\lambda.\vec{a} + \mu.\vec{b}) \in f(V)$ avec $\lambda.\vec{a} + \mu\vec{b} \in V$ puisque V est un sous-espace vectoriel.
Ainsi $f(V)$ est un sous-espace vectoriel de F.
Considérons W un sous-espace vectoriel de F et étudions $f^{-1}(W)$.
$f^{-1}(W) \subset E$ et $\vec{0}_E \in f^{-1}(W)$ car $f(\vec{0}_E) = \vec{0}_F \in W$.
Soient $\lambda, \mu \in \mathbb{K}$ et $\vec{x}, \vec{y} \in f^{-1}(W)$. $f(\lambda.\vec{x} + \mu.\vec{y}) = \lambda.f(\vec{x}) + \mu.f(\vec{y}) \in W$ donc $\lambda.\vec{x} + \mu.\vec{y} \in f^{-1}(W)$.
Ainsi $f^{-1}(W)$ est un sous-espace vectoriel de E.
□

Définition

Soit $f : E \to F$ une application linéaire.
On appelle image de f l'espace $\mathrm{Im} f = f(E)$.
On appelle noyau de f l'espace $\ker f = f^{-1}(\left\{\vec{0}_F\right\})$.

Proposition

$\mathrm{Im} f$ est un sous-espace vectoriel de F.
$\ker f$ est un sous-espace vectoriel de E.

dém. :
E est un sous-espace vectoriel de E et f est linéaire donc $f(E)$ est un sous-espace vectoriel de F.
$\left\{\vec{0}_F\right\}$ est un sous-espace vectoriel de F et f est linéaire donc $f^{-1}\left\{\vec{0}_F\right\}$ est un sous-espace vectoriel de E.
□

Remarque Pour déterminer l'image d'une application linéaire f, on détermine les valeurs prises par f i.e. les $\vec{y} \in F$ tels qu'il existe $\vec{x} \in E$ pour lequel $\vec{y} = f(\vec{x})$.

Remarque Pour déterminer le noyau d'une application linéaire f, on résout l'équation $f(\vec{x}) = \vec{0}_F$ d'inconnue $\vec{x} \in E$.

Exemple Déterminons noyau et image de l'endomorphisme $f : \mathbb{R}^2 \to \mathbb{R}^2$ définie par $f : (x, y) \mapsto (x - y, y - x)$.

Soit $\vec{a} = (x, y) \in \mathbb{R}^2$.

$$f(\vec{a}) = \vec{0}_{\mathbb{R}^2} \Leftrightarrow f(x,y) = (0,0) \Leftrightarrow \begin{cases} x - y = 0 \\ y - x = 0 \end{cases} \Leftrightarrow x = y$$

Par suite

$$\ker f = \{(x,x)/x \in \mathbb{R}\}$$

Soient $\vec{a} = (x, y) \in \mathbb{R}^2$ et $\vec{b} = (X, Y) \in \mathbb{R}^2$.

$$f(\vec{a}) = \vec{b} \Leftrightarrow f(x,y) = (X,Y) \Leftrightarrow \begin{cases} x - y = X \\ y - x = Y \end{cases} \Leftrightarrow \begin{cases} x = X + y \\ X + Y = 0 \end{cases}$$

Ainsi l'équation $f(\vec{a}) = \vec{b}$ possède au moins une solution $\vec{a}$ si, et seulement si, $X = Y$.
Par suite

$$\mathrm{Im} f = \{(X, -X)/X \in \mathbb{R}\} \text{ et } \ker f = \{(x,x)/x \in \mathbb{R}\}$$

Exemple Soit I un intervalle non singulier de $\mathbb{R}$.
Déterminons noyau et image de l'application linéaire $D : \mathcal{C}^1(I, \mathbb{R}) \to \mathcal{F}(I, \mathbb{R})$ définie par $D(f) = f'$.
Soit $f \in \mathcal{C}^1(I, \mathbb{R})$.

$$D(f) = \tilde{0} \Leftrightarrow f' = \tilde{0} \Leftrightarrow f \text{ est constante}$$

Par suite

$$\ker D = \{x \mapsto C/C \in \mathbb{R}\}$$

Soient $f \in \mathcal{C}^1(I, \mathbb{R})$ et $g \in \mathcal{F}(I, \mathbb{R})$.

$$D(f) = g \Leftrightarrow g = f'$$

Si g appartient à $\mathrm{Im}D$ alors g est la dérivée d'une fonction de classe $\mathcal{C}^1$ et c'est donc une fonction continue.
Inversement, si g est une fonction continue, on sait que g admet au moins une primitive et celle-ci est de classe $\mathcal{C}^1$.
Ainsi les valeurs prises par D sont exactement les fonctions continues de I vers $\mathbb{R}$.
Par suite

$$\mathrm{Im}D = \mathcal{C}^0(I, \mathbb{R})$$

Théorème

Si $f : E \to F$ est une application linéaire alors
1) f est surjective $\Leftrightarrow \mathrm{Im} f = F$;
2) f est injective $\Leftrightarrow \ker f = \left\{\vec{0}_E\right\}$.

dém. :
1) est immédiate par définition de la surjectivité.
2) ($\Rightarrow$) Si f est injective alors $\vec{0}_F$ possède au plus un antécédent, or $f(\vec{0}_E) = \vec{0}_F$ donc $\vec{0}_F$ possède exactement un antécédent, à savoir $\vec{0}_E$. On peut alors affirmer $\ker f = \left\{\vec{0}_E\right\}$.
($\Leftarrow$) Supposons $\ker f = \left\{\vec{0}_E\right\}$. Soient $\vec{x}, \vec{y} \in E$ tels que $f(\vec{x}) = f(\vec{y})$.
On a alors $f(\vec{x} - \vec{y}) = f(\vec{x}) - f(\vec{y}) = \vec{0}_F$ donc $\vec{x} - \vec{y} \in \ker f$ puis $\vec{x} - \vec{y} = \vec{0}_E$ et finalement $\vec{x} = \vec{y}$.
Ainsi f est injective.
□

6.3.4 L'espace vectoriel $\mathcal{L}(E,F)$

Théorème

$(\mathcal{L}(E,F),+,.)$ est un $\mathbb{K}$-espace vectoriel.

dém. :
Montrons que $\mathcal{L}(E,F)$ est un sous-espace vectoriel de $(\mathcal{F}(E,F),+,.)$.
$\mathcal{L}(E,F)\subset\mathcal{F}(E,F)$ et l'application nulle $\tilde{0}: x\in E\mapsto\vec{0}_F$ est linéaire.
Soient $\lambda,\mu\in\mathbb{K}$ et $f,g\in\mathcal{L}(E,F)$. Montrons que l'application $\lambda.f+\mu.g$ est linéaire.
Soient $\alpha,\beta\in\mathbb{K}$ et $\vec{x},\vec{y}\in E$.

$$(\lambda.f+\mu.g)\,(\alpha.\vec{x}+\beta.\vec{y})=\lambda.f(\alpha.\vec{x}+\beta.\vec{y})+\mu.g(\alpha.\vec{x}+\beta.\vec{y})$$

Par linéarité de f et g

$$(\lambda.f+\mu.g)\,(\alpha.\vec{x}+\beta.\vec{y})=\alpha\lambda.f(\vec{x})+\beta\lambda f(\vec{y})+\alpha\mu.g(\vec{x})+\beta\mu.g(\vec{y})$$

et en réorganisant ce calcul

$$(\lambda.f+\mu.g)\,(\alpha.\vec{x}+\beta.\vec{y})=\alpha.\,(\lambda.f+\mu.g)\,(\vec{x})+\beta.\,(\lambda.f+\mu.g)\,(\vec{y})$$

Ainsi $\lambda.f+\mu.g\in\mathcal{L}(E,F)$.
□

Exemple $(\mathcal{L}(E),+,.)$ et $(E^\star,+.)$ sont des $\mathbb{K}$-espaces vectoriels.

Proposition

$\forall\lambda,\mu\in\mathbb{K},\forall f,g\in\mathcal{L}(E,F)$ et $\forall h\in\mathcal{L}(F,G)$,

$$h\circ(\lambda.f+\mu.g)=\lambda.(h\circ f)+\mu.(h\circ g)$$

dém. :
Soient $\lambda,\mu\in\mathbb{K}$, $f,g\in\mathcal{L}(E,F)$ et $h\in\mathcal{L}(F,G)$.
Pour tout $\vec{x}\in E$,

$$[h\circ(\lambda.f+\mu.g)]\,(\vec{x})=h\,(\lambda.f(\vec{x})+\mu.g(\vec{x}))$$

Par linéarité de h,

$$[h\circ(\lambda.f+\mu.g)]\,(\vec{x})=\lambda.h(f(\vec{x}))+\mu.h(g(\vec{x}))=[\lambda.(h\circ f)+\mu.(h\circ g)]\,(\vec{x})$$

Ceci valant pour tout $\vec{x}\in E$, on a $h\circ(\lambda.f+\mu.g)=\lambda.(h\circ f)+\mu.(h\circ g)$.
□

Proposition

$\forall\lambda,\mu\in\mathbb{K},\forall h\in\mathcal{L}(E,F)$ et $\forall f,g\in\mathcal{L}(F,G)$,

$$(\lambda.f+\mu.g)\circ h=\lambda.(f\circ h)+\mu.(g\circ h)$$

dém. :
Soient $\lambda,\mu\in\mathbb{K}$, $h\in\mathcal{L}(E,F)$ et $f,g\in\mathcal{L}(F,G)$.
Pour tout $\vec{x}\in E$,

$$[(\lambda.f+\mu.g)\circ h]\,(\vec{x})=(\lambda.f+\mu.g)(h(\vec{x}))=\lambda.f(h(\vec{x}))+\mu.g(h(\vec{x}))$$

puis

$$[(\lambda.f + \mu.g) \circ h]\,(\vec{x}) = \lambda.(f \circ h)(\vec{x}) + \mu.(g \circ h)(\vec{x}) = [\lambda.(f \circ h) + \mu.(g \circ h)]\,(\vec{x})$$

Ceci valant pour tout $\vec{x} \in E$, on a $(\lambda.f + \mu.g) \circ h = \lambda.(f \circ h) + \mu.(g \circ h)$
□

6.3.5 L'anneau des endomorphismes

Théorème

$(\mathcal{L}(E), +, \circ)$ est un anneau d'élément nul l'endomorphisme nul (noté 0 ou $\tilde{0}$) et d'élément unité l'endomorphisme identité (noté I ou Id).
De plus, les éléments inversibles de cet anneau sont les automorphismes de E.

dém. :
On sait déjà que $(\mathcal{L}(E), +)$ est un groupe abélien de neutre $\tilde{0}$ car $(\mathcal{L}(E), +, .)$ est un espace vectoriel.
La loi de composition $\circ$ définit une loi de composition interne sur $\mathcal{L}(E)$ car la composée de deux endomorphismes de E est un endomorphisme de E.
La loi $\circ$ est associative et possède un neutre Id_E qui est un endomorphisme de E.
La loi $\circ$ est distributive sur $+$ dans $\mathcal{L}(E)$. En effet, par linéarité à droite et gauche et du produit de composition, on a

$$\forall f, g, h \in \mathcal{L}(E), f \circ (g + h) = f \circ g + f \circ h \text{ et } (g + h) \circ f = g \circ f + h \circ f$$

Enfin, un élément inversible de l'anneau $(\mathcal{L}(E), +, \circ)$ est un endomorphisme f pour lequel il existe $g \in \mathcal{L}(E)$ vérifiant $f \circ g = g \circ f = \mathrm{Id}_E$; un tel endomorphisme est bijectif.
Inversement un automorphisme est un élément inversible de l'anneau et son inverse est son isomorphisme réciproque.
□

Remarque L'anneau $(\mathcal{L}(E), +, \circ)$ n'est généralement pas commutatif.

Remarque Puisque $(\mathcal{L}(E), +, \circ)$ est un anneau, on peut exploiter les définitions et les résultats relatifs à cette structure. En particulier, on peut introduire les itérés de composition d'un endomorphisme.

Définition

Pour $f \in \mathcal{L}(E)$, on note

$$f^0 = I,\, f^1 = f,\, f^2 = f \circ f, \ldots, f^n = f \circ f \circ \ldots \circ f \text{ (} n \text{ facteurs)}$$

Exemple Soit $f \in \mathcal{L}(E)$ et $n \in \mathbb{N}$.
Montrons

$$\ker f^n \subset \ker f^{n+1} \text{ et } \mathrm{Im} f^n \supset \mathrm{Im} f^{n+1}$$

Soit $\vec{x} \in \ker f^n$. On a $f^n(\vec{x}) = \vec{0}$ donc $f^{n+1}(\vec{x}) = f(f^n(\vec{x})) = f(\vec{0}) = \vec{0}$ et ainsi $\vec{x} \in \ker f^{n+1}$.
Par conséquent $\ker f^n \subset \ker f^{n+1}$.
Soit $\vec{y} \in \mathrm{Im} f^{n+1}$. Il existe $\vec{x} \in E$ tel que $\vec{y} = f^{n+1}(\vec{x})$ et alors $\vec{y} = f^n(f(\vec{x})) = f^n(\vec{a})$ avec $\vec{a} = f(\vec{x}) \in E$ et ainsi $\vec{y} \in \mathrm{Im} f^n$.
Par conséquent $\mathrm{Im} f^{n+1} \subset \mathrm{Im} f^n$.

Attention : Puisque l'anneau $(\mathcal{L}(E),+,\circ)$ n'est pas a priori commutatif
- l'itéré $(f\circ g)^n$ ne correspond pas à $f^n\circ g^n$, il doit être compris $(f\circ g)\circ(f\circ g)\circ\cdots\circ(f\circ g)$;
- le carré d'une somme se développe ainsi $(f+g)^2=f^2+f\circ g+g\circ f+g^2$.

Proposition

Si $f,g\in\mathcal{L}(E)$ commutent (i.e. $f\circ g=g\circ f$) alors pour tout $n\in\mathbb{N}$,

$$(f\circ g)^n=f^n\circ g^n,\ (f+g)^n=\sum_{k=0}^{n}\binom{n}{k}f^k\circ g^{n-k}\text{ et }f^n-g^n=(f-g)\sum_{k=0}^{n-1}f^kg^{n-k-1}$$

dém. :
Par opérations dans un anneau sachant que les deux éléments f et g commutent.
□

Exemple Puisque les endomorphismes f et Id commutent

$$(\mathrm{Id}+f)^n=\sum_{k=0}^{n}\binom{n}{k}f^k\text{ et }f^n-\mathrm{Id}=(f-\mathrm{Id})\circ\left(\sum_{k=0}^{n-1}f^k\right)=\left(\sum_{k=0}^{n-1}f^k\right)\circ(f-\mathrm{Id})$$

Attention : L'anneau $(\mathcal{L}(E),+,\circ)$ n'est pas intègre, ainsi si un produit de composition est nul, on ne pour autant affirmer qu'un des facteur est nul :
$f\circ g=\tilde{0}$ n'implique par $f=\tilde{0}$ ou $g=\tilde{0}$.
Cependant, on peut souvent exploiter à profit le résultat suivant

Proposition

Si f et g sont des endomorphismes de E alors on a équivalence entre :
(i) $g\circ f=\tilde{0}$;
(ii) $\mathrm{Im}f\subset\ker g$.

dém. :
(i) $\Rightarrow$ (ii) Supposons $g\circ f=\tilde{0}$.
Pour tout $\vec{y}\in\mathrm{Im}f$, il existe $\vec{x}\in E$ tel que $\vec{y}=f(\vec{x})$.
On a alors $g(\vec{y})=g(f(\vec{x}))=(g\circ f)(\vec{x})=\vec{0}$ donc $\vec{y}\in\ker g$ et ainsi $\mathrm{Im}f\subset\ker g$.
(ii) $\Rightarrow$ (i) Supposons $\mathrm{Im}f\subset\ker g$.
Pour tout $\vec{x}\in E$, $(g\circ f)(\vec{x})=g(f(\vec{x}))=\vec{0}$ car $f(\vec{x})\in\mathrm{Im}f\subset\ker g$. Ainsi $g\circ f=\vec{0}$
□

Définition

Un endomorphisme $f\in\mathcal{L}(E)$ est dit idempotent si $f^2=f$.
Un endomorphisme $f\in\mathcal{L}(E)$ est dit nilpotent s'il existe $n\in\mathbb{N}$ tel que $f^n=\tilde{0}$.

Exemple Soit $f\in\mathcal{L}(E)$ nilpotent. Montrons que $\mathrm{Id}-f\in\mathrm{GL}(E)$.
Soit $n\in\mathbb{N}$ tel que $f^n=\tilde{0}$. On a $\mathrm{Id}=\mathrm{Id}-f^n$.
Puisque Id et f commutent, on peut factoriser et on obtient

$$\mathrm{Id}=(\mathrm{Id}-f)\circ\left(\sum_{k=0}^{n-1}f^k\right)=\left(\sum_{k=0}^{n-1}f^k\right)\circ(\mathrm{Id}-f)$$

Par suit $\mathrm{Id} - f$ est inversible et

$$(\mathrm{Id} - f)^{-1} = \sum_{k=0}^{n-1} f^k$$

Exemple Soit $f \in \mathcal{L}(E)$ tel que

$$f^2 - 5f + 6\mathrm{Id} = \tilde{0}$$

Montrons que $\ker(f - 2\mathrm{Id})$ et $\ker(f - 3\mathrm{Id})$ sont des sous-espaces vectoriels supplémentaires de E.
Puisque $f - 2\mathrm{Id}$ et $f - 3\mathrm{Id}$ sont des endomorphismes de E, leurs noyaux sont des sous-espaces vectoriels de E et ainsi $\ker(f - 2\mathrm{Id})$ et $\ker(f - 3\mathrm{Id})$ sont bien des sous-espaces vectoriels de E.
Etudions $\ker(f - 2\mathrm{Id}) \cap \ker(f - 3\mathrm{Id})$
Soit $\vec{x} \in \ker(f - 2\mathrm{Id}) \cap \ker(f - 3\mathrm{Id})$.
On a $(f - 2\mathrm{Id})(\vec{x}) = \vec{0}$ et $(f - 3\mathrm{Id})(\vec{x}) = \vec{0}$ donc $f(\vec{x}) = 2\vec{x}$ et $f(\vec{x}) = 3\vec{x}$.
On en déduit $\vec{x} = \vec{0}$ et ainsi $\ker(f - 2\mathrm{Id}) \cap \ker(f - 3\mathrm{Id}) \subset \left\{\vec{0}\right\}$ puis

$$\ker(f - 2\mathrm{Id}) \cap \ker(f - 3\mathrm{Id}) = \left\{\vec{0}\right\}$$

Etudions $\ker(f - 2\mathrm{Id}) + \ker(f - 3\mathrm{Id})$.
Analyse : Soit $\vec{x} \in E$. Supposons $\vec{x} = \vec{a} + \vec{b}$ avec $\vec{a} \in \ker(f - 2\mathrm{Id})$ et $\vec{b} \in \ker(f - 3\mathrm{Id})$.
On a $f(\vec{x}) = f(\vec{a}) + f(\vec{b})$ avec $f(\vec{a}) = 2\vec{a}$ et $f(\vec{b}) = 3\vec{b}$.
On obtient ainsi le système

$$\begin{cases} \vec{a} + \vec{b} = \vec{x} \\ 2\vec{a} + 3\vec{b} = f(\vec{x}) \end{cases}$$

Après résolution on obtient $\vec{a} = 3\vec{x} - f(\vec{x}) = (3\mathrm{Id} - f)(\vec{x})$ et $\vec{b} = f(\vec{x}) - 2\vec{x} = (f - 2\mathrm{Id})(\vec{x})$.
Synthèse :
Soient $\vec{x} \in E$ et les vecteurs

$$\vec{a} = (3\mathrm{Id} - f)(\vec{x}) \text{ et } \vec{b} = (f - 2\mathrm{Id})(\vec{x})$$

On a immédiatement $\vec{x} = \vec{a} + \vec{b}$.

$$(f - 2\mathrm{Id})(\vec{a}) = [(f - 2\mathrm{Id}) \circ (3\mathrm{Id} - f)](\vec{x}) = \left[-f^2 + 5f - 6\mathrm{Id}\right](\vec{x}) = \vec{0}$$

et

$$(f - 3\mathrm{Id})(\vec{b}) = [(f - 3\mathrm{Id}) \circ (f - 2\mathrm{Id})](\vec{x}) = \left[f^2 - 5f + 6\mathrm{Id}\right](\vec{x}) = \vec{0}$$

Ainsi $\vec{a} \in \ker(f - 2\mathrm{Id})$ et $\vec{b} \in \ker(f - 3\mathrm{Id})$.
On en déduit $E \subset \ker(f - 2\mathrm{Id}) + \ker(f - 3\mathrm{Id})$ et

$$E = \ker(f - 2\mathrm{Id}) + \ker(f - 3\mathrm{Id})$$

Finalement $\ker(f - 2\mathrm{Id})$ et $\ker(f - 3\mathrm{Id})$ sont des sous-espaces vectoriels supplémentaires de E.

6.4 Transformations vectorielles

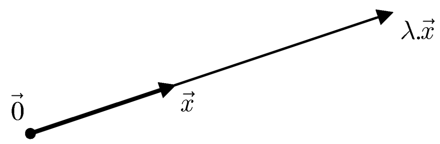
E désigne un $\mathbb{K}$-espace vectoriel.

6.4.1 Homothétie vectorielle

Soit $\lambda \in \mathbb{K}$.

Définition

On appelle homothétie (vectorielle) de rapport $\lambda \in \mathbb{K}$ l'application $h_\lambda : E \to E$ définie par $h_\lambda(\vec{x}) = \lambda.\vec{x}$.



Exemple Si $\lambda = 1$ alors $h_\lambda = \text{Id}$.
Si $\lambda = 0$ alors $h_\lambda = \tilde{0}$.

Proposition

Une homothétie vectorielle est un endomorphisme commutant avec tout endomorphisme.

dém. :
Considérons h une homothétie vectorielle de rapport λ.
On a $h = \lambda.\text{Id}$, or Id est une endomorphisme donc, par opération, h aussi.
De plus, pour tout endomorphisme f,

$$h \circ f = (\lambda.\text{Id}) \circ f = \lambda.(\text{Id} \circ f) = \lambda.f \text{ et } f \circ h = f \circ (\lambda.\text{Id}) = \lambda.(f \circ \text{Id}) = \lambda f \text{ donc } h \circ f = f \circ h$$

□

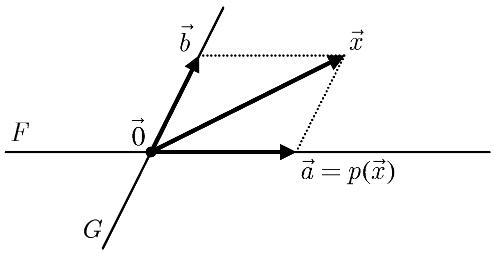
Proposition

$\forall \lambda, \mu \in \mathbb{K}, h_\lambda \circ h_\mu = h_{\lambda\mu}$ et
$\forall \lambda \in \mathbb{K}^\star, h_\lambda \in \text{GL}(E)$ avec $(h_\lambda)^{-1} = h_{1/\lambda}$.

dém. :
$h_\lambda \circ h_\mu = (\lambda.\text{Id}) \circ (\mu.\text{Id}) = (\lambda\mu).\text{Id} = h_{\lambda\mu}$ et $h_\lambda \circ h_{1/\lambda} = \text{Id} = h_{1/\lambda} \circ h_\lambda$.
□

6.4.2 Projection vectorielle

Soient F et G deux sous-espaces vectoriels supplémentaires de E.
Pour tout $\vec{x} \in E$, il existe un unique couple $(\vec{a}, \vec{b}) \in F \times G$ vérifiant $\vec{x} = \vec{a} + \vec{b}$
Posons $p(\vec{x}) = \vec{a}$; on définit ainsi une application $p : E \to E$.



Définition

L'application p est appelée projection (vectorielle) sur F parallèlement à G.

Exemple Si $F = E$ et $G = \left\{\vec{0}\right\}$ alors $p = \text{Id}$.
Si $F = \left\{\vec{0}\right\}$ et $G = E$ alors $p = \tilde{0}$.

Théorème

p est un endomorphisme de E vérifiant $p^2 = p$.
De plus

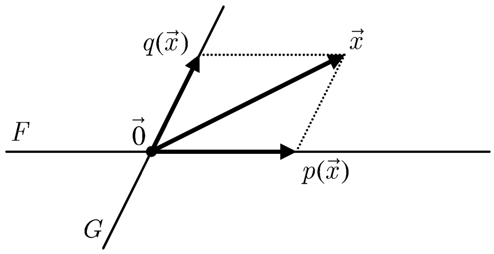
$$\text{Im}p = \ker(p - \text{Id}) = F \text{ et } \ker p = G$$

dém. :
p est une application de E vers E, vérifions qu'elle est linéaire.
Soient $\lambda, \mu \in \mathbb{K}$ et $\vec{x}, \vec{y} \in E$.
On peut écrire $\vec{x} = \vec{a} + \vec{b}$ et $\vec{y} = \vec{c} + \vec{d}$ avec $\vec{a}, \vec{c} \in F$ et $\vec{b}, \vec{d} \in G$.
On a alors $p(\vec{x}) = \vec{a}$ et $p(\vec{y}) = \vec{c}$.
On a aussi $\lambda.\vec{x} + \mu.\vec{y} = (\lambda.\vec{a} + \mu.\vec{c}) + (\lambda.\vec{b} + \mu.\vec{d})$ avec $(\lambda.\vec{a} + \mu.\vec{c}) \in F$ et $(\lambda.\vec{b} + \mu.\vec{d}) \in G$ et donc $p(\lambda.\vec{x} + \mu.\vec{y}) = \lambda.\vec{a} + \mu.\vec{b} = \lambda.p(\vec{x}) + \mu.p(\vec{y})$
Finalement p est un endomorphisme de E.
Montrons $p^2 = p$ avec $p^2 = p \circ p$.
Soit $\vec{x} \in E$. On a par construction $p(\vec{x}) \in F$, on peut alors écrire $p(\vec{x}) = p(\vec{x}) + \vec{0}$ avec $p(\vec{x}) \in F$ et $\vec{0} \in G$ et donc $p(p(\vec{x})) = p(\vec{x})$. Ainsi $p^2 = p \circ p$.
Montrons que $\text{Im}p = \ker(p - I) = F$
Par construction, les valeurs prises par p appartiennent à F. Ainsi $\text{Im}p \subset F$.
Soit $\vec{x} \in F$. On a $\vec{x} = \vec{x} + \vec{0}$ avec $\vec{x} \in F$ et $\vec{0} \in G$ donc $p(\vec{x}) = \vec{x}$ puis $(p - \text{Id})(\vec{x}) = \vec{0}$. Ainsi $F \subset \ker(p - \text{Id})$.
Enfin, si $\vec{x} \in \ker(p - \text{Id})$ alors $(p - \text{Id})(\vec{x})$ donc $p(\vec{x}) = \vec{x}$ puis $\vec{x} \in \text{Im}p$ car le vecteur $\vec{x}$ est une valeur prise par p. Ainsi $\ker(p - \text{Id}) \subset F$.
Finalement, par inclusions successives, on a montré $\text{Im}p = \ker(p - I) = F$.
Montrons $\ker p = G$.
Soit $\vec{x} \in \ker p$. On a $p(\vec{x}) = \vec{0}$ donc la décomposition de $\vec{x}$ en somme d'un vecteur de F et d'un vecteur de G s'écrit $\vec{x} = \vec{a} + \vec{b}$ avec $\vec{a} = \vec{0}$ et $\vec{b} \in G$. On en déduit $\vec{x} = \vec{b} \in G$.
Inversement, si $\vec{x} \in G$ alors on peut écrire $\vec{x} = \vec{0} + \vec{x}$ avec $\vec{0} \in F$ et $\vec{x} \in G$ et donc $p(\vec{x}) = \vec{0}$.
Par double inclusion $\ker p = G$.
□

Remarque $F = \text{Im}p = \ker(p - \text{Id})$ est l'espace des vecteurs invariants par p.

Définition

La projection vectorielle q sur G et parallèlement à F est appelée projection complémentaire de p.



Proposition

$q = \text{Id} - p$.

dém. :
Soit $\vec{x} \in E$. On peut écrire $\vec{x} = \vec{a} + \vec{b}$ avec $\vec{a} \in F$ et $\vec{b} \in G$.
On a alors $p(\vec{x}) = \vec{a}$ et $q(\vec{x}) = \vec{b}$ donc $(p+q)(\vec{x}) = p(\vec{x}) + q(\vec{x}) = \vec{a} + \vec{b} = \vec{x}$.
Ainsi $p + q = \text{Id}$.
□

Remarque On observe $p \circ q = q \circ p = \tilde{0}$ ce qui peut être utile.

6.4.3 Projecteur

Définition

On appelle projecteur de E, tout endomorphisme p de E vérifiant $p^2 = p$.

Exemple Les projections vectorielles sont des projecteurs.

Théorème

Si p est un projecteur de E alors
1) $\text{Im}p$ et $\ker p$ sont supplémentaires ;
2) p est la projection vectorielle sur $F = \text{Im}p$, parallèlement à $G = \ker p$.

dém. :
Soit p un projecteur de E.
Préalablement à notre étude montrons $p(\vec{x}) = \vec{x}$ pour tout $\vec{x} \in \text{Im}p$.
Soit $\vec{x} \in \text{Im}p$, il existe $\vec{a} \in E$ vérifiant $\vec{x} = p(\vec{a})$ et alors $p(\vec{x}) = p^2(\vec{a}) = p(\vec{a}) = \vec{x}$.
Ainsi on a vérifié

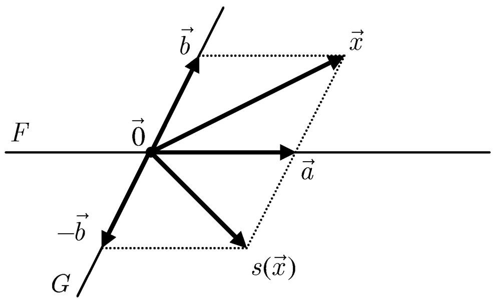
$$\forall \vec{x} \in \text{Im}p, p(\vec{x}) = \vec{x}$$

Montrons que $\text{Im}p$ et $\ker p$ sont supplémentaires.
Etudions $\text{Im}p \cap \ker p$.

Soit $\vec{x} \in \mathrm{Im}p \cap \ker p$.
On a $p(\vec{x}) = \vec{x}$ et $p(\vec{x}) = \vec{0}$ donc $\vec{x} = \vec{0}$.
Ainsi $\mathrm{Im}p \cap \ker p \subset \left\{\vec{0}\right\}$ puis $\mathrm{Im}p \cap \ker p = \left\{\vec{0}\right\}$.
Montrons $\mathrm{Im}p \cap \ker p = E$.
Analyse :
Supposons $\vec{x} = \vec{a} + \vec{b}$ avec $\vec{a} \in \mathrm{Im}p$ et $\vec{b} \in \ker p$.
On a $p(\vec{x}) = p(\vec{a}) + p(\vec{b}) = \vec{a}$ car $\vec{a} \in \mathrm{Im}p$ et $\vec{b} \in \ker p$.
Ainsi $\vec{a} = p(\vec{x})$ et $\vec{b} = \vec{x} - p(\vec{x})$.
Synthèse :
Posons $\vec{a} = p(\vec{x})$ et $\vec{b} = \vec{x} - p(\vec{x})$.
On a immédiatement $\vec{x} = \vec{a} + \vec{b}$ et $\vec{a} \in \mathrm{Im}p$.
On a aussi $p(\vec{b}) = p(\vec{x}) - p^2(\vec{x}) = \vec{0}$ car $p^2 = p$ et donc $\vec{b} \in \ker p$.
Ainsi $E \subset \mathrm{Im}p + \ker p$ puis $E = \mathrm{Im}p + \ker p$.
Finalement $\mathrm{Im}p$ et $\ker p$ sont supplémentaires dans E.
Enfin, pour tout $\vec{x} \in E$, on peut écrire de façon unique $\vec{x} = \vec{a} + \vec{b}$ avec $\vec{a} \in \mathrm{Im}p$ et $\vec{b} \in \ker p$ et on a alors $p(\vec{x}) = p(\vec{a}) + p(\vec{b}) = \vec{a}$. L'action de p se reconnaît alors comme étant la projection sur $F = \mathrm{Im}p$ parallèlement à $G = \ker p$.
□

6.4.4 Symétrie vectorielle

Soient F et G deux sous-espaces vectoriels supplémentaires de E.
Pour tout $\vec{x} \in E$, il existe un unique couple $(\vec{a}, \vec{b}) \in F \times G$ vérifiant $\vec{x} = \vec{a} + \vec{b}$
Posons $s(\vec{x}) = \vec{a} - \vec{b}$; on définit ainsi une application $s : E \to E$.



Définition

s est appelée symétrie (vectorielle) par rapport à F parallèlement à G.

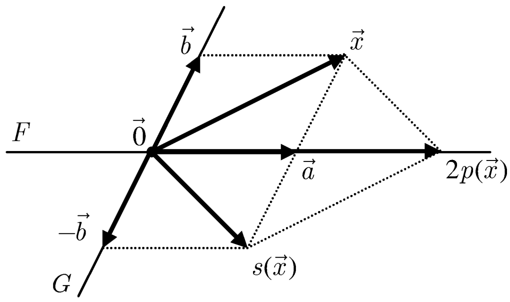
Exemple Si $F = E$ et $G = \left\{\vec{0}\right\}$ alors $s = \mathrm{Id}$.
Si $F = \left\{\vec{0}\right\}$ et $G = E$ alors $s = -\mathrm{Id}$.

Théorème

L'application s est un endomorphisme de E vérifiant $s^2 = I$.
De plus
$$F = \ker(s - \mathrm{Id}) \text{ et } G = \ker(s + \mathrm{Id})$$

dém. :
Introduisons la projection vectorielle p sur F parallèlement à G et $q = \mathrm{Id} - p$ sa projection complémentaire.
Pour tout $\vec{x} \in E$, on peut écrire $\vec{x} = \vec{a} + \vec{b}$ avec $\vec{a} \in F$ et $\vec{b} \in G$.
On a alors $p(\vec{x}) = \vec{a}$ et $q(\vec{x}) = \vec{b}$.
Par définition de l'application s, on a aussi $s(\vec{x}) = \vec{a} - \vec{b}$ et donc $s(\vec{x}) = p(\vec{x}) - q(\vec{x}) = (p - q)(\vec{x})$.
On en déduit que $s = p - q = 2p - \mathrm{Id}$ et donc s est un endomorphisme de E par opération sur les endomorphismes de E.
De plus p et Id commutent, on a $s^2 = (2p - \mathrm{Id})^2 = 4p^2 - 4p + \mathrm{Id} = \mathrm{Id}$ car $p^2 = p$.
Enfin $\ker(s - \mathrm{Id}) = \ker(2p - 2\mathrm{Id}) = \ker(p - \mathrm{Id}) = F$ et $\ker(s + \mathrm{Id}) = \ker 2p = \ker p = G$.
□

Remarque Durant la démonstration, on a vu la relation remarquable et utile $s = 2p - \mathrm{Id}$.



Remarque $F = \ker(s - \mathrm{Id})$ est l'espace des vecteurs invariants par s.
$G = \ker(s + \mathrm{Id})$ est l'espace des vecteurs changés en leur opposé par s.

Corollaire

s est un automorphisme de E et $s^{-1} = s$.

Définition

La symétrie s' par rapport à G et parallèlement à F est appelée symétrie complémentaire de s.

Proposition

$s' = -s$.

dém. :
Soit $\vec{x} \in E$. On peut écrire $\vec{x} = \vec{a} + \vec{b}$ avec $\vec{a} \in F$ et $\vec{b} \in G$.
On a $s(\vec{x}) = \vec{a} - \vec{b}$ et $s'(\vec{x}) = \vec{b} - \vec{a}$ donc $s'(\vec{x}) = -s(\vec{x})$ et ainsi $s' = -s$.
□

Théorème

Si s est un endomorphisme de E vérifiant $s^2 = \mathrm{Id}$ alors
1) $\ker(s - \mathrm{Id})$ et $\ker(s + \mathrm{Id})$ sont supplémentaires dans E ;
2) s est la symétrie vectorielle par rapport à $F = \ker(s - \mathrm{Id})$, parallèlement à $G = \ker(s + \mathrm{Id})$.

dém. :
Posons $p = \dfrac{1}{2}(s + \mathrm{Id}) \in \mathcal{L}(E)$.

Puisque s et Id commutent, on a $p^2 = \frac{1}{4}\left(s^2 + 2s + \mathrm{Id}\right) = \frac{1}{2}\left(s + \mathrm{Id}\right) = p$.
L'endomorphisme p est un projecteur de E. On en déduit que $F = \mathrm{Im}p = \ker(p - \mathrm{Id})$ et $G = \ker p$ sont des sous-espaces vectoriels supplémentaires de E.
Or

$$F = \ker(p - \mathrm{Id}) = \ker\left[\frac{1}{2}\left(s - \mathrm{Id}\right)\right] = \ker(s - \mathrm{Id})$$

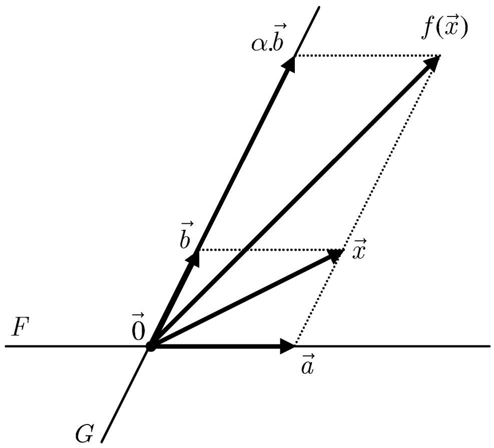
et

$$G = \ker p = \ker\left[\frac{1}{2}\left(s + \mathrm{Id}\right)\right] = \ker(s + \mathrm{Id})$$

Ainsi $\ker(s - \mathrm{Id})$ et $\ker(s + \mathrm{Id})$ sont des sous-espaces vectoriels supplémentaires de E.
De plus, pour tout $\vec{x} \in E$, on peut écrire de façon unique $\vec{x} = \vec{a} + \vec{b}$ avec $\vec{a} \in F = \ker(s - \mathrm{Id})$ et $\vec{b} \in G = \ker(s + \mathrm{Id})$. On a alors $s(\vec{x}) = s(\vec{a}) + s(\vec{b})$. Or $s(\vec{a}) = \vec{a}$ car $\vec{a} \in \ker(s - \mathrm{Id})$ et $s(\vec{b}) = -\vec{b}$ car $\vec{b} = \ker(s + \mathrm{Id})$. Ainsi $s(\vec{x}) = \vec{a} - \vec{b}$ et on reconnaît à travers l'action de s, la symétrie vectorielle par rapport à F est parallèlement à G.
□

6.4.5 Affinités vectorielles

Soient F et G deux sous-espaces vectoriels supplémentaires de E.
Pour tout $\vec{x} \in E$, il existe un unique couple $(\vec{a}, \vec{b}) \in F \times G$ vérifiant $\vec{x} = \vec{a} + \vec{b}$
Posons $f(\vec{x}) = \vec{a} + \alpha.\vec{b}$; on définit ainsi une application $f : E \to E$.



Définition

| f est appelée affinité (vectorielle) par rapport à F parallèlement à G et de rapport α.

Exemple Si $\alpha = 1$ alors $f = \mathrm{Id}$.
Si $\alpha = 0$ alors $f = p$.
Si $\alpha = -1$ alors $f = s$.

Proposition

| f est un endomorphisme de E.

dém. :
Par construction de f, on vérifie $f = p + \alpha.q$ avec p la projection sur F parallèlement à G et q sa projection complémentaire.
□

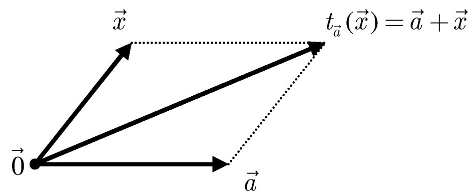
6.5 Notions affines

E désigne un $\mathbb{K}$-espace vectoriel.

6.5.1 Translation

Définition

On appelle translation de vecteur $\vec{a} \in E$ l'application $t_{\vec{a}} : E \to E$ définie par $\vec{x} \mapsto \vec{a} + \vec{x}$.



Exemple $t_{\vec{o}} = \mathrm{Id}$.

Attention : En dehors de la translation de vecteur nul, les translations ne sont pas des applications linéaires.

Proposition

$\forall \vec{a}, \vec{b} \in E, t_{\vec{a}} \circ t_{\vec{b}} = t_{\vec{a}+\vec{b}} = t_{\vec{b}} \circ t_{\vec{a}}$
$\forall \vec{a} \in E$, t_a est une permutation de E et $t_a^{-1} = t_{-a}$.

dém. :
Soient $\vec{a}, \vec{b} \in E$.
Pour tout $\vec{x} \in E$, $(t_{\vec{a}} \circ t_{\vec{b}})(\vec{x}) = t_{\vec{a}}(\vec{b} + \vec{x}) = \vec{a} + \vec{b} + \vec{x} = t_{\vec{a}+\vec{b}}(\vec{x})$.
Ainsi $t_{\vec{a}} \circ t_{\vec{b}} = t_{\vec{a}+\vec{b}}$ et puisque $t_{\vec{a}+\vec{b}} = t_{\vec{b}+\vec{a}}$, on a aussi $t_{\vec{b}} \circ t_{\vec{a}} = t_{\vec{a}+\vec{b}}$.
Enfin, $t_{\vec{a}} \circ t_{-\vec{a}} = t_{\vec{0}} = \mathrm{Id}$ et $t_{-\vec{a}} \circ t_{\vec{a}} = \mathrm{Id}$ donc $t_{\vec{a}}$ est une bijection d'application réciproque $t_{\vec{a}}^{-1} = t_{-\vec{a}}$
□

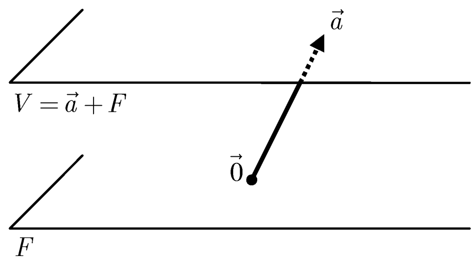
Remarque L'ensemble des translations de E est un sous-groupe de $(\mathfrak{S}(E), \circ)$

6.5.2 Sous-espace affine

Définition

On appelle sous-espace affine passant par $\vec{a} \in E$ et dirigé par un sous-espace vectoriel F l'ensemble

$$\vec{a} + F = \{\vec{a} + \vec{x} / \vec{x} \in F\}$$



Remarque $V = \vec{a} + F$ est l'image du sous-espace vectoriel par la translation de vecteur $\vec{a}$.

Remarque $\vec{a} \in \vec{a} + F$ car $\vec{a} = \vec{a} + \vec{0}$ et $\vec{0} \in F$, par suite un sous-espace affine n'est jamais vide.
En revanche, il est incertain que le vecteur nul appartienne à $\vec{a} + F$, ainsi un sous-espace affine peut ne pas être un sous-espace vectoriel.

Exemple $\{\vec{a}\} = \vec{a} + \left\{\vec{0}\right\}$ est un sous-espace affine.

Exemple Pour $\vec{u} \neq \vec{0}$,

$$\{\vec{a} + \lambda\vec{u} / \lambda \in \mathbb{K}\} = \vec{a} + \text{Vect}(\vec{u})$$

est un sous-espace affine, on parle de droite affine passant par $\vec{a}$ et dirigée par $\vec{u}$.

Exemple Géométriquement les sous-espaces affines se visualisent comme étant des points, des droites ou des plans ne passant pas nécessairement par $\vec{o}$.

Proposition

Si $V = \vec{a} + F$ est un sous-espace affine alors pour tout $\vec{b} \in E$ on a équivalence entre :
(i) $\vec{b} \in V$;
(ii) $\vec{b} - \vec{a} \in F$.
De plus, si tel est le cas,

$$V = \vec{b} + F$$

dém. :
(i) $\Rightarrow$ (ii) Si $\vec{b} \in V$ alors on peut écrire $\vec{b} = \vec{a} + \vec{x}$ avec $\vec{x} \in F$ donc $\vec{b} - \vec{a} = \vec{x} \in F$.

(ii) $\Rightarrow$ (i) Si $\vec{b}-\vec{a}\in F$ alors $\vec{b}=\vec{a}+(\vec{b}-\vec{a})\in\vec{a}+F$.
De plus, si $\vec{b}\in V$ alors pour tout $\vec{x}\in F$, $\vec{b}+\vec{x}=\vec{a}+(\vec{b}-\vec{a}+\vec{x})\in\vec{a}+F$.
Ainsi $\vec{b}+F\subset\vec{a}+F$. De façon semblable on obtient $\vec{a}+F\subset\vec{b}+F$ et on peut conclure $V=\vec{b}+F$
□

Remarque Ainsi, il n'y a pas unicité du vecteur $\vec{a}$ définissant un sous-espace affine $V=\vec{a}+F$.
En revanche, en vertu du résultat suivant, il y a unicité du sous-espace vectoriel F définissant un sous-espace affine.

Proposition

Si V est un sous-espace affine dirigé par F alors

$$F=\{\vec{y}-\vec{x}/\vec{x},\vec{y}\in V\}$$

dém. :
Soit $\vec{a}\in V$. On peut écrire $V=\vec{a}+F$.
Pour tout $\vec{x},\vec{y}\in V$, on peut écrire $\vec{x}=\vec{a}+\vec{u}$ et $\vec{y}=\vec{a}+\vec{v}$ avec $\vec{u},\vec{v}\in F$ et alors $\vec{y}-\vec{x}=\vec{v}-\vec{u}\in F$.
Inversement, soit $\vec{u}\in F$. Pour $\vec{x}=\vec{a}+\vec{0}\in V$ et $\vec{y}=\vec{a}+\vec{u}\in V$, $\vec{u}=\vec{y}-\vec{x}$ avec $\vec{x},\vec{y}\in V$.
□

Définition

Le sous-espace vectoriel F est alors appelé direction du sous-espace affine V.

Remarque Pour décrire un sous-espace affine V, il suffit de connaître sa direction F et l'un de ses points $\vec{a}$ car alors $V=\vec{a}+F$.

6.5.3 Incidence de sous-espaces affines

Proposition

Si V et W sont deux sous-espaces affines de directions F et G alors $V\cap W$ est soit vide, soit égal à un sous-espace affine de direction $F\cap G$.

dém. :
Si $V\cap W$ n'est pas vide, introduisons $\vec{a}\in V\cap W$.
On peut alors décrire les sous-espaces affines V et W : $V=\vec{a}+F$ et $W=\vec{a}+G$.
Pour $\vec{x}\in E$, on a
$\vec{x}\in V$ et $\vec{x}\in W\Leftrightarrow\vec{x}-\vec{a}\in F$ et $\vec{x}-\vec{a}\in G$
et donc
$\vec{x}\in V\cap W\Leftrightarrow\vec{x}-\vec{a}\in F\cap G$.
Par suite $V\cap W=\vec{a}+F\cap G$ est un sous-espace affine de direction $F\cap G$.
□

Définition

Soient V et W deux sous-espaces affines de directions F et G.
On dit que V est parallèle à W si $F\subset G$

Proposition

> Soient V et W deux sous-espaces affines de direction F et G.
> Si V est parallèle à W alors $V \cap W = \emptyset$ ou $V \subset W$.

dém. :
Supposons V parallèle à W i.e. $F \subset G$
Si $V \cap W$ n'est pas vide, introduisons $\vec{a} \in V \cap W$.
On peut alors décrire les sous-espaces affines V et W : $V = \vec{a} + F$ et $W = \vec{a} + G$.
Puisque $F \subset G$, on a $V \subset W$.
□

6.5.4 Equations linéaires

Définition

> On appelle équation linéaire toute équation de la forme $f(\vec{x}) = \vec{y}$ avec $f \in \mathcal{L}(E, F)$, $\vec{y} \in F$ et d'inconnue $\vec{x} \in E$.
> On appelle équation homogène associée l'équation $f(\vec{x}) = \vec{0}$.

Théorème

> L'ensemble solution de l'équation linéaire $f(\vec{x}) = \vec{y}$ est soit vide, soit égal à un sous-espace affine de dirigé par $\ker f$.

dém. :
S'il existe une solution $\vec{a}$ à l'équation étudiée alors pour tout $\vec{x} \in E$,

$$f(\vec{x}) = \vec{y} \Leftrightarrow f(\vec{x}) = f(\vec{a})$$

Or

$$f(\vec{x}) = f(\vec{a}) \Leftrightarrow f(\vec{x}) - f(\vec{a}) = \vec{0}$$

et puisque f est linéaire

$$f(\vec{x}) - f(\vec{a}) = f(\vec{x} - \vec{a})$$

Ainsi

$$f(\vec{x}) = \vec{y} \Leftrightarrow \vec{x} - \vec{a} \in \ker f$$

L'ensemble des solutions de l'équation étudiée est alors $\vec{a} + \ker f$, c'est un sous-espace affine dirigé par $\ker f$.
□

Remarque Il est fréquent, pour résoudre une équation linéaire $f(\vec{x}) = \vec{y}$, de procéder ainsi :
- on résout l'équation homogène associée ce qui donne $\ker f$;
- on cherche un élément solution $\vec{a}$, appelé solution particulière ;
- on exprimer l'ensemble solution $\vec{a} + \ker f$.

Ainsi la solution générale est somme d'une solution particulière est de la solution générale de l'équation homogène.

Exemple Soit $y' + a(x)y = b(x)$ équation différentielle linéaire d'ordre 1.
Considérons $E = \mathcal{C}^1(I, \mathbb{R})$, $F = \mathcal{C}^0(I, \mathbb{R})$ et l'application linéaire $f : E \to F$ déterminée par

$$f(y) : x \mapsto y'(x) - a(x)y(x)$$

L'équation différentielle étudiée équivaut à l'équation linéaire $f(y) = b$ d'inconnue $y \in \mathcal{C}^1(I, \mathbb{R})$.

Pour la résoudre :
- on résout l'équation homogène $f(y) = 0$ i.e. l'équation différentielle $y' + a(x)y = 0$;
- on détermine une solution particulière ;
- on exprime la solution générale comme somme de la solution particulière et de la solution générale homogène.

6.5.5 Barycentre

Soient $n \in \mathbb{N}^\star$, $\vec{u}_1, \dots, \vec{u}_n \in E$ et $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ tel que $\lambda_1 + \dots + \lambda_n \neq 0$.

Définition

On appelle barycentre des vecteurs $\vec{u}_1, \dots, \vec{u}_n$ affectés respectivement des masses $\lambda_1, \dots, \lambda_n$ le vecteur

$$\vec{v} = \frac{\lambda_1.\vec{u}_1 + \dots + \lambda_n.\vec{u}_n}{\lambda_1 + \dots + \lambda_n}$$

Remarque Si G est le point barycentre de points $A_1, \dots, A_n$ affectés des masses $\lambda_1, \dots, \lambda_n$ alors le vecteur $\overrightarrow{OG}$ est le vecteur barycentre des vecteurs $\overrightarrow{OA_1}, \dots, \overrightarrow{OA_n}$ affectés des mêmes masses $\lambda_1, \dots, \lambda_n$.

Remarque On ne modifie par le barycentre $\vec{v}$ lorsque :
- on permute les couples $(\vec{u}_i, \lambda_i)$;
- on supprime les couples $(\vec{u}_i, \lambda_i)$ de masse nulle ;
- on multiplie chaque masse par un même scalaire non nul.

Cette dernière propriété permet notamment de ramener la masse totale à 1.

Définition

Lorsque les $\lambda_1, \dots, \lambda_n$ sont égaux non nuls, alors le vecteur

$$\vec{v} = \frac{1}{n}(\vec{u}_1 + \dots + \vec{u}_n)$$

est appelé isobarycentre des vecteurs $\vec{u}_1, \dots, \vec{u}_n$.

Proposition

Si tous les vecteurs $\vec{u}_i$ appartiennent à un même sous-espace affine $V = \vec{a} + F$ alors le barycentre $\vec{v}$ des vecteurs $\vec{u}_1, \dots, \vec{u}_n$ affectés des masses $\lambda_1, \dots, \lambda_n$ appartient à V.

dém. :
Pour tout $i \in \{1, \dots, n\}$, on peut écrire $\vec{u}_i = \vec{a} + \vec{x}_i$ avec $\vec{x}_i$.
On a alors

$$\vec{v} = \frac{\lambda_1.\vec{a} + \dots + \lambda_n.\vec{a}}{\lambda_1 + \dots + \lambda_n} + \frac{\lambda_1.\vec{x}_1 + \dots + \lambda_n.\vec{x}}{\lambda_1 + \dots + \lambda_n} = \vec{a} + \vec{x}$$

avec

$$\vec{x} = \frac{\lambda_1.\vec{x}_1 + \dots + \lambda_n.\vec{x}}{\lambda_1 + \dots + \lambda_n} \in F$$

□

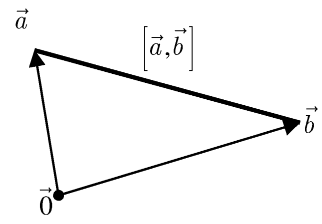
6.5.6 Convexité

Ici $\mathbb{K} = \mathbb{R}$.

Définition

On appelle segment d'extrémités $\vec{a}$ et $\vec{b} \in E$ l'ensemble :

$$\left[\vec{a}, \vec{b}\right] = \left\{(1-\lambda)\vec{a} + \lambda\vec{b} / \lambda \in [0,1]\right\}$$



Remarque $\left[\vec{a}, \vec{b}\right] = \left[\vec{b}, \vec{a}\right]$

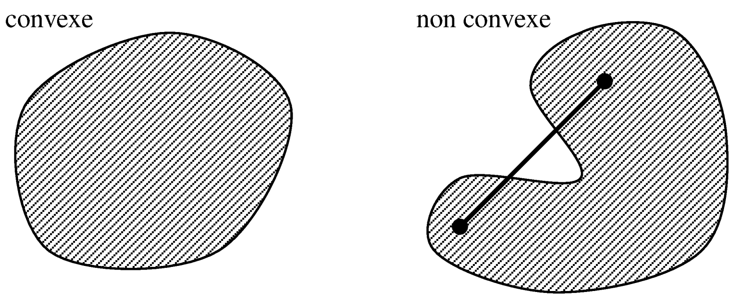
Définition

Une partie C de E est dite convexe si

$$\forall \vec{a}, \vec{b} \in C, \left[\vec{a}, \vec{b}\right] \subset C$$

Exemple $\emptyset$ et E sont des parties convexes.

Exemple



Exemple Les sous-espaces affines et les segments sont des parties convexes.

Définition

On appelle combinaison convexe d'une famille de vecteurs de E, tout barycentre obtenu avec des masses positives.

Exemple Un isobarycentre est une combinaison convexe.

Exemple Pour $\lambda \in [0,1]$, $(1-\lambda).\vec{a} + \lambda.\vec{b}$ est une combinaison convexe de $\vec{a}$ et $\vec{b}$.

Proposition

Une partie C de E est convexe si, et seulement si, elle est stable par combinaison convexe.

dém. :
Si une partie C de E est stable par combinaison convexe alors pour tout $\vec{a}, \vec{b} \in C$, le segment $\left[\vec{a}, \vec{b}\right]$ est inclus dans C car ses éléments sont des combinaisons convexes d'éléments de C.
Inversement, soit C une partie convexe.
Montrons par récurrence sur $n \in \mathbb{N}^\star$ que C est stable par combinaison convexe de n vecteurs.
Pour $n = 1$: ok
Supposons la propriété établie au rang $n \geqslant 1$.
Soient $\vec{u}_1, \ldots, \vec{u}_n, \vec{u}_{n+1} \in C$ et $\lambda_1, \ldots, \lambda_n, \lambda_{n+1} \geqslant 0$ non tous nuls.
Montrons que

$$\vec{x} = \frac{\lambda_1.\vec{u}_1 + \cdots + \lambda_{n+1}.\vec{u}_{n+1}}{\lambda_1 + \cdots + \lambda_{n+1}} \in C$$

Si $\lambda_1 = \ldots = \lambda_n = 0$ alors $\vec{x} = \vec{u}_{n+1} \in C$
Sinon, on peut écrire

$$\vec{x} = (1-\mu).\vec{a} + \mu.\vec{u}_{n+1}$$

avec

$$\vec{a} = \frac{1}{\lambda_1 + \cdots + \lambda_n}\left(\lambda_1.\vec{u}_1 + \cdots + \lambda_n.\vec{u}_n\right) \text{ et } \mu = \frac{\lambda_{n+1}}{\lambda_1 + \cdots + \lambda_{n+1}} \in [0,1]$$

Par suite $\vec{x} \in [\vec{a}, \vec{u}_{n+1}]$.
Or par hypothèse de récurrence $\vec{a} \in C$, donc $[\vec{a}, \vec{u}_{n+1}] \subset C$ puis $\vec{x} \in C$.
Récurrence établie.
□

Chapitre 7

Dimension d'un espace vectoriel

$\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$.

7.1 Famille de vecteurs

Soient E un $\mathbb{K}$-espace vectoriel et $\mathcal{F} = (\vec{e}_i)_{i \in I}$ une famille finie de vecteurs de E. Quitte à réindexer celle-ci, on suppose que les éléments de cette famille sont indexés sur l'ensemble $I = \{1, \ldots, n\}$. On considère donc une famille $\mathcal{F} = (\vec{e}_i)_{1 \leqslant i \leqslant n}$ de vecteurs de E qu'on peut encore écrire $\mathcal{F} = (\vec{e}_1, \ldots, \vec{e}_n)$. Dans le cas particulier où $n = 0$, on dit que la famille $\mathcal{F}$ est vide.

7.1.1 Sous-espace vectoriel engendré par une famille finie de vecteurs

Définition

On appelle combinaison linéaire des vecteurs de la famille $\mathcal{F} = (\vec{e}_i)_{1 \leqslant i \leqslant n}$ tout vecteur $\vec{x}$ de E pouvant s'écrire $\vec{x} = \sum_{i=1}^{n} \lambda_i \vec{e}_i$ avec $\lambda_1, \ldots, \lambda_n$ scalaires bien choisis.

Exemple Une somme sur le vide étant égale au vecteur nul, le vecteur nul est le seul vecteur combinaison linéaire de la famille vide.

Définition

On appelle espace vectoriel engendré par la famille $\mathcal{F} = (\vec{e}_i)_{1 \leqslant i \leqslant n}$, le sous-espace vectoriel engendré par la partie $\{\vec{e}_1, \ldots, \vec{e}_n\}$. On le note Vect$\mathcal{F}$, Vect$(\vec{e}_i)_{1 \leqslant i \leqslant n}$ ou Vect$(\vec{e}_1, \ldots, \vec{e}_n)$.

Exemple Le sous-espace vectoriel engendré par la famille vide est l'espace nul $\left\{\vec{0}\right\}$.

Théorème

Si $(\vec{e}_1, \ldots, \vec{e}_n)$ est une famille de vecteurs de E alors Vect$(\vec{e}_1, \ldots, \vec{e}_n)$ est l'ensemble des combinaisons linéaires des vecteurs $\vec{e}_1, \ldots, \vec{e}_n$ i.e. :

$$\text{Vect}(\vec{e}_1, \ldots, \vec{e}_n) = \left\{ \sum_{i=1}^{n} \lambda_i \vec{e}_i / \lambda_1, \ldots, \lambda_n \in \mathbb{K} \right\}$$

dém. :
Posons
$$F = \left\{ \sum_{i=1}^{n} \lambda_i \vec{e}_i / \lambda_1, \ldots, \lambda_n \in \mathbb{K} \right\}$$
On montre facilement que F est un sous-espace vectoriel de E et puisque les vecteurs $\vec{e}_1, \ldots, \vec{e}_n$ sont tous éléments de F, on a Vect$(\vec{e}_1, \ldots, \vec{e}_n) \subset F$.
Inversement, les vecteurs $\vec{e}_1, \ldots, \vec{e}_n$ sont tous éléments de Vect$(\vec{e}_1, \ldots, \vec{e}_n)$ et puisque Vect$(\vec{e}_1, \ldots, \vec{e}_n)$ est un sous-espace vectoriel, Vect$(\vec{e}_1, \ldots, \vec{e}_n)$ contient toutes les combinaisons linéaires des vecteurs $\vec{e}_1, \ldots, \vec{e}_n$ i.e. tous les éléments de F.
□

Exemple Cas $n = 1$
$$\text{Vect}(\vec{u}) = \{\lambda.\vec{u}/\lambda \in \mathbb{K}\} = \mathbb{K}.\vec{u}$$

Exemple Cas $n = 2$
$$\text{Vect}(\vec{u}, \vec{v}) = \{\lambda.\vec{u} + \mu.\vec{v}/\lambda, \mu \in \mathbb{K}\} = \mathbb{K}.\vec{u} + \mathbb{K}.\vec{v}$$

Exemple Dans $\mathbb{R}^3$, considérons $\vec{u} = (1, 1, 1), \vec{v} = (0, -1, 2)$.
$$\text{Vect}(\vec{u}, \vec{v}) = \{(\lambda, \lambda + \mu, 2\mu)/\lambda, \mu \in \mathbb{R}\}$$

Remarque Il est efficace d'établir qu'une partie est un sous-espace vectoriel en observant que celle-ci s'apparente à un espace vectoriel engendré par une famille.

Exemple Dans $\mathbb{R}^3$, considérons
$$P = \{(a + b, a - b, 2b)/a, b \in \mathbb{R}\}$$
Puisque $P =$ Vect$(\vec{u}, \vec{v})$ avec $\vec{u} = (1, 1, 0)$ et $\vec{v} = (1, -1, 2)$, P est un sous-espace vectoriel de $\mathbb{R}^3$.

Exemple Dans $\mathbb{R}^3$, considérons
$$P = \{(x, y, z)/x + y - z = 0\}$$
Puisque
$$x + y - z = 0 \Leftrightarrow z = x + y$$
on a
$$P = \{(x, y, x + y)/x, y \in \mathbb{R}\} = \text{Vect}(\vec{u}, \vec{v})$$
avec $\vec{u} = (1, 0, 1)$ et $\vec{v} = (0, 1, 1)$. Ainsi P est un sous-espace vectoriel de $\mathbb{R}^3$.

7.1.2 Famille génératrice

Définition

On dit qu'une famille $\mathcal{F} = (\vec{e}_1, \dots, \vec{e}_n) = (\vec{e}_i)_{1 \leqslant i \leqslant n}$ de vecteurs de E est génératrice de E si tout vecteur de E est combinaison linéaire des vecteurs de la famille $\mathcal{F}$ i.e.

$$\forall \vec{x} \in E, \exists (\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n, \vec{x} = \lambda_1 \vec{e}_1 + \cdots + \lambda_n \vec{e}_n = \sum_{i=1}^{n} \lambda_i \vec{e}_i$$

Remarque La famille $\mathcal{F}$ est génératrice de E si, et seulement si, Vect$\mathcal{F} = E$.

Exemple Dans $E = \mathbb{K}^n$, on pose $\vec{e}_i = (0, ..., 0, 1, 0, ..., 0) \in \mathbb{K}^n$ où 1 se situe en $i$ème position.
La famille $\mathcal{B} = (\vec{e}_i)_{1 \leqslant i \leqslant n}$ est une famille génératrice de $\mathbb{K}^n$.
En effet, pour tout $\vec{x} = (x_1, \dots, x_n) \in \mathbb{K}^n$, on peut écrire $\vec{x} = x_1 \vec{e}_1 + \cdots + x_n \vec{e}_n$.

Exemple Dans $E = \mathbb{K}$, la famille (1) est génératrice.
En effet, tout $x \in \mathbb{K}$ peut s'écrire $x = x.1$

Exemple Dans $E = \mathbb{C}$ vu comme $\mathbb{R}$-espace vectoriel, la famille $\mathcal{F} = (1, i)$ est génératrice.
En effet, pour tout $z \in \mathbb{C}$, on peut écrire $z = a.1 + b.i$ avec $a = \text{Re}z$ et $b = \text{Im}z$.

Exemple Dans l'espace $E = \mathcal{F}(\mathbb{R}, \mathbb{R})$, il n'existe pas de famille génératrice finie.

Remarque Toute sur-famille d'une famille génératrice est génératrice.
En revanche, une sous famille d'une famille génératrice peut ne pas être génératrice.

Proposition

Si $(\vec{e}_1, \dots, \vec{e}_n, \vec{e}_{n+1})$ est une famille génératrice et si $\vec{e}_{n+1} \in \text{Vect}(\vec{e}_1, \dots, \vec{e}_n)$ alors la sous-famille $(\vec{e}_1, \dots, \vec{e}_n)$ est génératrice.

dém. :
Tout vecteur de E est combinaison linéaire des vecteurs $\vec{e}_1, \dots, \vec{e}_n, \vec{e}_{n+1}$ et puisque $\vec{e}_{n+1}$ est combinaison linéaire des vecteurs $\vec{e}_1, ..., \vec{e}_n$, tout vecteur de E est combinaison linéaire des vecteurs $\vec{e}_1, \dots, \vec{e}_n$.
□

7.1.3 Famille libre, famille liée

Définition

Un vecteur $\vec{u}$ est dit colinéaire à un vecteur $\vec{v}$ s'il existe $\alpha \in \mathbb{K}$ tel que $\vec{u} = \alpha \vec{v}$.
Deux vecteurs $\vec{u}$ et $\vec{v}$ sont dits colinéaires si l'un des deux est colinéaire à l'autre.

Attention : $\vec{u}$ colinéaire à $\vec{v}$ n'équivaut pas à $\vec{v}$ colinéaire à $\vec{u}$.
En effet, le vecteur nul est colinéaire à tout vecteur mais tout vecteur n'est pas colinéaire au vecteur nul !

Définition

On dit qu'une famille $(\vec{e}_1, \dots, \vec{e}_n) = (\vec{e}_i)_{1 \leqslant i \leqslant n}$ de vecteurs de E est libre si elle vérifie

$$\forall \lambda_1, \dots, \lambda_n \in \mathbb{K}, \lambda_1 \vec{e}_1 + \dots + \lambda_n \vec{e}_n = \vec{0} \Rightarrow \lambda_1 = \dots = \lambda_n = 0$$

On dit alors que les vecteurs $\vec{e}_1, \dots, \vec{e}_n$ sont linéairement indépendants.

Définition

On dit que la famille $(\vec{e}_1, \dots, \vec{e}_n) = (\vec{e}_i)_{1 \leqslant i \leqslant n}$ est liée si elle n'est pas libre ce qui signifie

$$\exists \lambda_1, \dots, \lambda_n \in \mathbb{K}, \lambda_1 \vec{e}_1 + \dots + \lambda_n \vec{e}_n = \vec{0} \text{ et } (\lambda_1, \dots, \lambda_n) \neq (0, \dots, 0)$$

Une égalité $\lambda_1 \vec{e}_1 + \dots + \lambda_n \vec{e}_n = \vec{0}$ avec $\lambda_1, \dots, \lambda_n$ non tous nuls est appelée relation linéaire sur les vecteurs $\vec{e}_1, \dots, \vec{e}_n$.

Exemple Soit $\vec{u} \in E$, étudions la liberté de la famille $(\vec{u})$.
Si $\vec{u} \neq \vec{0}$ alors

$$\forall \lambda \in \mathbb{K}, \lambda.\vec{u} = \vec{0} \Rightarrow \lambda = 0$$

Par suite, la famille $(\vec{u})$ est libre.
Si $\vec{u} = \vec{0}$ alors on peut écrire $\lambda.\vec{u} = \vec{0}$ avec $\lambda = 1 \neq 0$.
Par suite, la famille $(\vec{0})$ est liée.

Proposition

Soient $n \geqslant 2$ et $(\vec{e}_1, \dots, \vec{e}_n)$ une famille de vecteurs de E.
On a équivalence entre :
(i) $(\vec{e}_1, \dots, \vec{e}_n)$ est liée ;
(ii) l'un des vecteurs $\vec{e}_1, \dots, \vec{e}_n$ est combinaison linéaire des autres.

dém. :
(i) $\Rightarrow$ (ii) Supposons la famille $(\vec{e}_1, \dots, \vec{e}_n)$ liée.
Il existe $\lambda_1, \dots, \lambda_n$ non tous nuls vérifiant

$$\lambda_1 \vec{e}_1 + \dots + \lambda_n \vec{e}_n = \vec{0}$$

Puisqu'il existe $i \in \{1, \dots, n\}$ tel que $\lambda_i \neq 0$, on peut écrire

$$\vec{e}_i = -\frac{1}{\lambda_i} \left(\lambda_1 \vec{e}_1 + \dots + \lambda_{i-1} \vec{e}_{i-1} + \lambda_i \vec{e}_i + \dots + \lambda_n \vec{e}_n \right)$$

Ainsi l'un des vecteurs de la famille $(\vec{e}_1, \dots, \vec{e}_n)$ est combinaison linéaire des autres.
(ii) $\Rightarrow$ (i) Supposons l'un des vecteurs $\vec{e}_1, \dots, \vec{e}_n$ combinaison linéaire des autres.
Il existe $i \in \{1, \dots, n\}$ tel que

$$\vec{e}_i = \lambda_1 \vec{e}_1 + \dots + \lambda_{i-1} \vec{e}_{i-1} + \lambda_i \vec{e}_i + \dots + \lambda_n \vec{e}_n$$

En posant $\lambda_i = -1$, on peut alors écrire la relation $\lambda_1 \vec{e}_1 + \dots + \lambda_n \vec{e}_n = \vec{0}$ avec $\lambda_1, \dots, \lambda_n$ non tous nuls.
Ainsi la famille $(\vec{e}_1, \dots, \vec{e}_n)$ est liée.
□

Exemple Soient $\vec{u}, \vec{v} \in E$.

$$(\vec{u}, \vec{v}) \text{ est liée si, et seulement si, } (\exists \alpha \in \mathbb{K}, \vec{u} = \alpha.\vec{v}) \text{ ou } (\exists \beta \in \mathbb{K}, \vec{v} = \beta.\vec{u})$$

Ainsi, la famille $(\vec{u}, \vec{v})$ est liée si, et seulement si, $\vec{u}$ et $\vec{v}$ sont colinéaires.

Exemple Dans $E = \mathbb{R}^3$, considérons les vecteurs $\vec{u} = (1, 2, 1), \vec{v} = (1, -1, 1), \vec{w} = (1, 1, 0)$ et la famille $\mathcal{F} = (\vec{u}, \vec{v}, \vec{w})$.
Etudions la liberté de la famille $\mathcal{F}$.
Soient $\alpha, \beta, \gamma \in \mathbb{R}$.

$$\alpha.\vec{u} + \beta.\vec{v} + \gamma.\vec{w} = \vec{0} \Leftrightarrow \begin{cases} \alpha + \beta + \gamma = 0 \\ 2\alpha - \beta + \gamma = 0 \\ \alpha + \beta = 0 \end{cases}$$

Après résolution du système, on obtient

$$\alpha.\vec{u} + \beta.\vec{v} + \gamma.\vec{w} = \vec{0} \Leftrightarrow \alpha = \beta = \gamma = 0$$

la famille $\mathcal{F}$ est donc libre.

Exemple Dans $E = \mathbb{R}^3$, considérons les vecteurs $\vec{u} = (1, -1, 0), \vec{v} = (2, -1, 1), \vec{w} = (0, 1, 1)$ et la famille $\mathcal{F} = (\vec{u}, \vec{v}, \vec{w})$.
Etudions la liberté de la famille $\mathcal{F}$.
Soient $\alpha, \beta, \gamma \in \mathbb{R}$.

$$\alpha.\vec{u} + \beta.\vec{v} + \gamma.\vec{w} = \vec{0} \Leftrightarrow \begin{cases} \alpha + 2\beta = 0 \\ -\alpha - \beta + \gamma = 0 \\ \beta + \gamma = 0 \end{cases}$$

Après résolution du système, on obtient

$$\alpha.\vec{u} + \beta.\vec{v} + \gamma.\vec{w} = \vec{0} \Leftrightarrow \begin{cases} \alpha = -2\beta \\ \gamma = -\beta \end{cases}$$

On en déduit que la famille $\mathcal{F}$ est liée car on a notamment la relation linéaire $-2\vec{u} + \vec{v} - \vec{w} = \vec{0}$.

Exemple Dans $E = \mathcal{F}(\mathbb{R}, \mathbb{R})$, considérons les fonctions $f : x \mapsto 1, g : x \mapsto \cos x, h : x \mapsto \sin x$ et montrons que la famille (f, g, h) est libre.
Soient $\alpha, \beta, \gamma \in \mathbb{R}$.
Supposons $\alpha f + \beta g + \gamma h = \tilde{0}$.
Pour tout $x \in \mathbb{R}$, on a $\alpha + \beta \cos x + \gamma \sin x = 0$.
Pour $x = 0$, on obtient l'équation $\alpha + \beta = 0$ (1).
Pour $x = \pi/2$, on obtient l'équation $\alpha + \gamma = 0$ (2).
Pour $x = \pi$, on obtient l'équation $\alpha - \beta = 0$ (3)
(1) et (3) donnent $\alpha = \beta = 0$ et par (2) on obtient $\gamma = 0$.
Finalement la famille (f, g, h) est effectivement libre.

Remarque Toute sous-famille d'une famille libre est libre.
Toute sur-famille d'une famille liée est liée, en particulier toute famille contenant le vecteur nul est liée.

Remarque Une sur-famille d'une famille libre n'est pas nécessairement libre.

Proposition

> Si $(\vec{e}_1, \ldots, \vec{e}_n)$ est une famille libre et si $\vec{e}_{n+1} \notin \text{Vect}(\vec{e}_1, \ldots, \vec{e}_n)$ alors la sur-famille $(\vec{e}_1, \ldots, \vec{e}_n, \vec{e}_{n+1})$ est libre.

dém. :
Soit $\lambda_1, \ldots, \lambda_n, \lambda_{n+1} \in \mathbb{K}$.
Supposons

$$\lambda_1\vec{e}_1 + \cdots + \lambda_n\vec{e}_n + \lambda_{n+1}\vec{e}_{n+1} = \vec{0}\,(1)$$

Si $\lambda_{n+1} \neq 0$ alors on peut écrire $\vec{e}_{n+1} = \mu_1\vec{e}_1 + \cdots + \mu_n\vec{e}_n$ avec $\mu_i = -\lambda_i/\lambda_{n+1}$.
Ceci est exclu car $\vec{e}_{n+1} \notin \text{Vect}(\vec{e}_1, \ldots, \vec{e}_n)$.
On en déduit $\lambda_{n+1} = 0$.
La relation (1) se réécrit alors

$$\lambda_1\vec{e}_1 + \cdots + \lambda_n\vec{e}_n = \vec{0}$$

Or la famille $(\vec{e}_1, \ldots, \vec{e}_n)$ est supposée libre donc $\lambda_1 = \ldots = \lambda_n = 0$.
Finalement $\lambda_1 = \ldots = \lambda_n = \lambda_{n+1} = 0$.
Ainsi la famille $(\vec{e}_1, \ldots, \vec{e}_n, \vec{e}_{n+1})$ est libre.
□

7.1.4 Base d'un espace vectoriel

Définition

> On dit qu'une famille $\mathcal{B} = (\vec{e}_i)_{1\leqslant i\leqslant n} = (\vec{e}_1, \ldots, \vec{e}_n)$ de vecteurs de E est une base de E si celle-ci est libre et génératrice.

Exemple Dans $E = \mathbb{K}^n$, on pose $\vec{e}_i = (0, ..., 0, 1, 0, ..., 0) \in \mathbb{K}^n$ où 1 se situe en $i$ème position.
On a déjà vu que la famille $\mathcal{B} = (\vec{e}_i)_{1\leqslant i\leqslant n}$ est génératrice de $\mathbb{K}^n$; montrons qu'elle est libre
Soient $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$
Supposons $\lambda_1\vec{e}_1 + \cdots + \lambda_n\vec{e}_n = \vec{0}$.
On a $(\lambda_1, \ldots, \lambda_n) = (0, \ldots, 0)$ et donc $\lambda_1 = \ldots = \lambda_n = 0$.
Finalement, la famille $\mathcal{B}$ est libre et génératrice de $\mathbb{K}^n$, c'est une base de $\mathbb{K}^n$.
Cas $n = 1$: $\vec{e}_1 = 1$, (1) est base de $\mathbb{K}$.
Cas $n = 2$: $\vec{e}_1 = (1, 0) = \vec{i}, \vec{e}_2 = (0, 1) = \vec{j}$, $(\vec{i}, \vec{j})$ est base de $\mathbb{K}^2$.
Cas $n = 3$: $\vec{e}_1 = (1, 0, 0) = \vec{i}, \vec{e}_2 = (0, 1, 0) = \vec{j}, \vec{e}_3 = (0, 0, 1) = \vec{k}$, $(\vec{i}, \vec{j}, \vec{k})$ est base de $\mathbb{K}^3$.

Exemple Considérons la famille $(1, i)$ d'éléments du $\mathbb{R}$-espace vectoriel $\mathbb{C}$.
On a déjà vu que cette famille est génératrice ; montrons qu'elle est libre.
Soient $\lambda, \mu \in \mathbb{R}$.
Supposons $\lambda.1 + \mu.i = 0$.
En identifiant parties réelles et imaginaires, on obtient $\lambda = \mu = 0$.
Finalement, la famille $\mathcal{B}$ est libre est génératrice du $\mathbb{R}$-espace vectoriel $\mathbb{C}$, c'est une base de $\mathbb{C}$.

Remarque La famille $(1, i)$ est liée dans le $\mathbb{C}$-espace vectoriel $\mathbb{C}$.
Elle n'est donc pas une base du $\mathbb{C}$-espace vectoriel $\mathbb{C}$.

Exemple La famille vide est base de l'espace vectoriel nul $\left\{\vec{0}\right\}$.
En effet, l'espace engendré par la famille vide est l'espace nul et la famille vide est libre (car toute assertion commençant par $\forall x \in \emptyset$ est vraie...)

Exemple Dans l'espace $E = \mathcal{F}(\mathbb{R}, \mathbb{R})$; il n'existe pas de bases au sens précédent.

Remarque Un $\mathbb{K}$-espace vectoriel E peut admettre une infinité de bases...

7.1.5 Composantes dans une base

Théorème

Si $\mathcal{B} = (\vec{e}_i)_{1 \leqslant i \leqslant n}$ est une base d'un $\mathbb{K}$-espace vectoriel E alors

$$\forall \vec{x} \in E, \exists!(\lambda_1, \ldots, \lambda_n) \in \mathbb{K}^n, \vec{x} = \lambda_1 \vec{e}_1 + \cdots + \lambda_n \vec{e}_n$$

dém. :
Existence : car la famille $\mathcal{B}$ est génératrice.
Unicité : Supposons $\vec{x} = \lambda_1 \vec{e}_1 + \cdots + \lambda_n \vec{e}_n$ et $\vec{x} = \mu_1 \vec{e}_1 + \cdots + \mu_n \vec{e}_n$ avec $\lambda_i, \mu_i \in \mathbb{K}$.
On a

$$(\lambda_1 - \mu_1)\vec{e}_1 + \cdots + (\lambda_n - \mu_n)\vec{e}_n = \vec{0}$$

Or la famille $\mathcal{B} = (\vec{e}_i)_{1 \leqslant i \leqslant n}$ est libre donc $\lambda_1 - \mu_1 = \ldots = \lambda_n - \mu_n = 0$ puis

$$(\lambda_1, \ldots, \lambda_n) = (\mu_1, \ldots, \mu_n)$$

□

Définition

Avec les notations ci-dessus, les scalaires $\lambda_1, \ldots, \lambda_n$ sont appelés le composantes de $\vec{x}$ dans la base $\mathcal{B}$ (ou encore les coordonnées de $\vec{x}$).

Remarque Les composantes d'un vecteur dépendent de la base dans laquelle on travaille.

Exemple Dans $E = \mathbb{K}^n$, considérons la base canonique $\mathcal{B} = (\vec{e}_i)_{1 \leqslant i \leqslant n}$ et le vecteur $\vec{x} = (x_1, \ldots, x_n)$.
Puisque $\vec{x} = x_1.\vec{e}_1 + \cdots + x_n.\vec{e}_n$, les composantes du vecteur $\vec{x}$ dans la base $\mathcal{B}$ sont les scalaires $x_1, \ldots, x_n$.

Exemple Dans le $\mathbb{R}$-espace vectoriel $\mathbb{C}$, les composantes de $z \in \mathbb{C}$ dans la base canonique $(1, i)$ sont $\mathrm{Re}(z)$ et $\mathrm{Im}(z)$.

Exemple Si $\mathcal{B} = (\vec{e}_i)_{1 \leqslant i \leqslant n}$ est une base d'un $\mathbb{K}$-espace vectoriel E alors
- le vecteur nul est le vecteur dont toutes les composantes sont nulles ;
- $\vec{e}_i$ est le vecteur dont les composantes sont $0, \ldots, 0, 1, 0, \ldots 0$ (où le 1 situé en i-ème position).

Théorème

> Si $\mathcal{B} = (\vec{e}_i)_{1 \leqslant i \leqslant n}$ est une base de E alors pour tout vecteur $\vec{x}$ et $\vec{y}$ de composantes $x_1, \ldots, x_n$ et $y_1, \ldots, y_n$ dans $\mathcal{B}$, les composantes de $\vec{x} + \vec{y}$ sont $x_1 + y_1, \ldots, x_n + y_n$ et celles de $\lambda\vec{x}$ sont $\lambda x_1, \ldots, \lambda x_n$.
> Ainsi l'application $\vec{x} \in E \mapsto x_i \in \mathbb{K}$ est une forme linéaire sur E.

dém. :
On a $\vec{x} = x_1\vec{e}_1 + \cdots + x_n\vec{e}_n$ et $\vec{y} = y_1\vec{e}_1 + \cdots + y_n\vec{e}_n$
donc $\vec{x} + \vec{y} = (x_1 + y_1)\vec{e}_1 + \cdots + (x_n + y_n)\vec{e}_n$ et $\lambda\vec{x} = (\lambda x_1)\vec{e}_1 + \cdots + (\lambda x_n)\vec{e}_n$.
□

Remarque Par calculs sur les composantes, la manipulation des vecteurs d'un $\mathbb{K}$-espace vectoriel rapporté à une base $\mathcal{B}$ formée de n vecteurs devient similaire à la manipulation d'éléments de $\mathbb{K}^n$.

7.2 Dimension d'un espace vectoriel

7.2.1 Espace vectoriel de dimension finie

Définition

> Un $\mathbb{K}$-espace vectoriel E est dit de dimension finie s'il possède une famille génératrice finie.
> Sinon, il est dit de dimension infinie.

Exemple Les espaces $\mathbb{K}^n$ et $\mathbb{C}$ sont de dimensions finies.

Exemple L'espace $E = \mathcal{F}(\mathbb{R}, \mathbb{R})$ n'est pas de dimension finie.

Exemple Tout $\mathbb{K}$-espace vectoriel possédant une base est de dimension finie.

Théorème

> Tout $\mathbb{K}$-espace vectoriel de dimension finie possède au moins une base.

dém. :
Soient E un $\mathbb{K}$-espace vectoriel de dimension finie et $\mathcal{G} = (\vec{u}_1, \ldots, \vec{u}_m)$ une famille génératrice de vecteurs de E.
Si la famille $\mathcal{G}$ est libre c'est une base de E.
Sinon, la famille $\mathcal{G}$ est liée, et alors un des vecteurs de $\mathcal{G}$ est combinaison linéaire des autres. Quitte à

permuter ceux-ci, supposons que le vecteur $\vec{u}_m$ est combinaison linéaire des vecteurs $\vec{u}_1, \ldots, \vec{u}_{m-1}$. Par le principe de réduction des familles génératrices, on peut affirmer que la famille $\mathcal{G}' = (\vec{u}_1, \ldots, \vec{u}_{m-1})$ est encore génératrice de E. On reprend alors la discussion précédente : si $\mathcal{G}'$ est libre, c'est une base de E, sinon on réduit la famille $\mathcal{G}'$. Nécessairement ce processus s'arrête car la famille initiale est finie et si, au pire, on retire tous les vecteurs de $\mathcal{G}$, on obtient la famille vide qui est libre.
$\square$

7.2.2 Dimension

Théorème

> Soit E un $\mathbb{K}$-espace vectoriel de dimension finie.
> Une famille libre de vecteurs de E ne peut avoir strictement plus de vecteurs qu'une famille génératrice.

dém. :
Par récurrence sur $n \in \mathbb{N}$, montrons la propriété :
« Si un $\mathbb{K}$- espace vectoriel E possède une famille génératrice à n vecteurs, alors toute famille constituée de $n+1$ vecteurs est liée ».
Le résultat du théorème découle de manière immédiate de cette propriété.
Pour $n = 0$
Si E possède une famille génératrice à 0 vecteurs c'est qu'alors $E = \text{Vect}(\emptyset) = \left\{\vec{0}\right\}$.
Par suite toute famille constituée d'un vecteur de E est liée car celle-ci contient le vecteur nul.
Supposons la propriété établie au rang $n-1 \in \mathbb{N}$.
Soit E un $\mathbb{K}$- espace vectoriel possédant une famille génératrice $\mathcal{G} = (\vec{e}_1, \ldots, \vec{e}_n)$ formée de n vecteurs et $\mathcal{U} = (\vec{u}_1, \ldots, \vec{u}_{n+1})$ une famille formée de $n+1$ vecteurs de E.
Nous voulons montrons que $\mathcal{U}$ est liée.
Posons $F = \text{Vect}(\vec{e}_1, \ldots, \vec{e}_{n-1})$.
$\mathcal{G}$ étant une famille génératrice de E, chaque vecteur de la famille $\mathcal{U}$ peut s'écrire comme combinaison linéaire des vecteurs de E, ceci permet d'écrire $\vec{u}_i = \vec{v}_i + \alpha_i \vec{e}_n$ avec $\vec{v}_i \in F$ et $\alpha_i \in \mathbb{K}$ pour tout $1 \leqslant i \leqslant n+1$.
Distinguons deux cas :
1er cas : Les α_i sont tous nuls.
Les vecteurs $\vec{u}_1, \ldots, \vec{u}_{n+1}$ sont alors tous éléments de F, or F est un sous-espace vectoriel engendré par $n-1$ vecteurs et la famille $\mathcal{U}$ est formée de $n+1$ vecteurs de celui-ci, l'hypothèse de récurrence permet alors d'affirmer qu'une sous-famille de n vecteurs de $\mathcal{U}$ est liée et donc qu'a fortiori la famille $\mathcal{U}$ est liée.
2ème cas : il existe au moins un α_i non nul.
Quitte à réindexer les vecteurs de la famille $\mathcal{U}$ (ce qui ne changera rien à son caractère liée ou non) on peut supposer $\alpha_{n+1} \neq 0$.
Posons alors

$$\vec{x}_i = \vec{u}_i - \frac{\alpha_i}{\alpha_{n+1}}.\vec{u}_{n+1}$$

pour tout $1 \leqslant i \leqslant n$.
Par construction, tous les vecteurs $\vec{x}_1, \ldots, \vec{x}_n$ appartiennent à F car

$$\vec{x}_i = (\vec{v}_i + \alpha_i \vec{e}_n) - \frac{\alpha_i}{\alpha_{n+1}} (\vec{v}_{n+1} + \alpha_{n+1} \vec{e}_n) = \vec{v}_i - \frac{\alpha_i}{\alpha_{n+1}} \vec{v}_{n+1}$$

La famille $(\vec{x}_1, \ldots, \vec{x}_n)$ apparaît alors comme étant une famille formée de n vecteurs de l'espace F, or F est engendré par $n-1$ vecteurs, l'hypothèse de récurrence permet alors d'affirmer que cette famille est liée.

Par suite il existe $(\lambda_1, \dots, \lambda_n) \neq (0, ..., 0)$ vérifiant $\lambda_1 \vec{x}_1 + \cdots + \lambda_n \vec{x}_n = \vec{0}$.
En reprenant l'expression de chaque $\vec{x}_i$, la relation ci-dessus donne

$$\lambda_1 \vec{u}_1 + \cdots + \lambda_n \vec{u}_n + \lambda_{n+1} \vec{u}_{n+1} = \vec{0}$$

avec

$$\lambda_{n+1} = -\frac{\lambda_1 \alpha_1 + \cdots + \lambda_n \alpha_n}{\alpha_{n+1}}$$

Cette dernière relation étant écrite avec des scalaires λ_i non tous nuls, on peut conclure que la famille $\mathcal{U}$ est liée.
Récurrence établie.
□

Corollaire

Les bases d'un $\mathbb{K}$-espace vectoriel E de dimension finie sont toutes constituées du même nombre de vecteurs.

dém. :
Soient $\mathcal{B} = (\vec{e}_1, ..., \vec{e}_n)$ et $\mathcal{B}' = (\vec{\varepsilon}_1, ..., \vec{\varepsilon}_m)$ deux bases de E.
Puisque $\mathcal{B}$ est libre et $\mathcal{B}'$ génératrice on a $n \leqslant m$.
Puisque $\mathcal{B}$ est génératrice et $\mathcal{B}'$ libre on a aussi $n \geqslant m$.
On en déduit $n = m$.
□

Définition

Si E est un $\mathbb{K}$-espace vectoriel de dimension finie, on appelle dimension de E le nombre de vecteurs constituant les bases de E. Elle est notée $\dim E$ ou encore $\dim_{\mathbb{K}} E$ s'il y a ambiguïté sur le corps de base $\mathbb{K}$.
Si l'espace E n'est pas de dimension finie, on pose $\dim E = +\infty$.

Remarque Pour déterminer la dimension d'un espace E, il suffit de déterminer une base de E et de dénombrer le nombre de vecteur la constituant.

Exemple $\dim \mathbb{K}^n = n$ car la base canonique de $\mathbb{K}^n$ est formée de n vecteurs.
En particulier $\dim \mathbb{K} = 1$.

Exemple $\dim_{\mathbb{R}} \mathbb{C} = 2$ et $\dim_{\mathbb{C}} \mathbb{C} = 1$.

Exemple $\dim \mathcal{F}(\mathbb{R}, \mathbb{R}) = +\infty$.

Exemple $\dim \left\{\vec{0}\right\} = 0$ car la famille vide est base de l'espace nul.

Exemple Les visualisations géométriques planes correspondent à la dimension 2, les visualisation dans l'espace correspondent à la dimension 3.

Remarque dimension = nombre de degré de liberté dans le choix d'un élément.

Définition

Un $\mathbb{K}$-espace vectoriel de dimension 1 est appelé une droite vectorielle.
Un $\mathbb{K}$-espace vectoriel de dimension 2 est appelé un plan vectoriel.

Théorème

Soient E et F deux $\mathbb{K}$-espaces vectoriels de dimensions finies.
L'espace $E \times F$ est de dimension finie et $\dim E \times F = \dim E + \dim F$.

dém. :
Posons $p = \dim E$ et $q = \dim F$.
Soient $\mathcal{B} = (\vec{e}_j)_{1\leqslant j\leqslant p}$ et $\mathcal{C} = (\vec{f}_k)_{1\leqslant k\leqslant q}$ des bases de E et F.
Pour $i \in \{1,\dots,p\}$, posons $\vec{g}_i = (\vec{e}_i, \vec{0}_F)$ et pour $i \in \{1,\dots,q\}$, posons $\vec{g}_{p+i} = (\vec{0}_E, \vec{f}_i)$.
On détermine ainsi une famille $\mathcal{D} = (\vec{g}_i)_{1\leqslant i\leqslant p+q}$ de vecteur de $E \times F$. Montrons que celle-ci est une base de $E \times F$.
Soit $(\vec{x}, \vec{y}) \in E \times F$. Puisque $\mathcal{B}$ et $\mathcal{C}$ sont des familles génératrices de E et F, on peut écrire

$$\vec{x} = x_1\vec{e}_1 + \cdots + x_p\vec{e}_p \text{ et } \vec{y} = y_1\vec{f}_1 + \cdots + y_q\vec{f}_q$$

avec $x_j, y_k \in \mathbb{K}$. On a alors

$$(\vec{x}, \vec{y}) = (\vec{x}, \vec{0}_F) + (\vec{0}_E, \vec{y}) = x_1\vec{g}_1 + \cdots + x_p\vec{g}_p + y_1\vec{g}_{p+1} + \cdots + y_q\vec{g}_{p+q}$$

Ainsi la famille $\mathcal{D}$ est génératrice de $E \times F$.
Montrons maintenant la liberté de $\mathcal{D}$.
Supposons

$$\lambda_1\vec{g}_1 + \cdots + \lambda_{p+q}\vec{g}_{p+q} = \vec{0}_{E\times F}$$

On a alors

$$(\lambda_1\vec{e}_1 + \cdots + \lambda_p\vec{e}_p, \lambda_{p+1}\vec{f}_1 + \cdots + \lambda_{p+q}\vec{f}_q) = (\vec{0}_E, \vec{0}_F)$$

donc

$$\lambda_1\vec{e}_1 + \cdots + \lambda_p\vec{e}_p = \vec{0}_E \text{ et } \lambda_{p+1}\vec{f}_1 + \cdots + \lambda_{p+q}\vec{f}_q = \vec{0}_F$$

Or les familles $\mathcal{B}$ et $\mathcal{C}$ sont libres donc $\lambda_1 = \ldots = \lambda_p = 0$ et $\lambda_{p+1} = \ldots = \lambda_{p+q} = 0$.
On peut donc conclure que la famille $\mathcal{D}$ est libre et finalement $\mathcal{D}$ est une base de $E \times F$.
□

7.2.3 Caractérisation de bases en dimension finie connue

Proposition

Soit E un $\mathbb{K}$-espace vectoriel de dimension $n \in \mathbb{N}$.
1) Toute famille libre a au plus n éléments.
2) Toute famille génératrice a au moins n éléments.
3) Toute base a exactement n éléments.

dém. :
Un espace de dimension n possède une base, c'est-à-dire une famille libre et génératrice formée de n vecteurs et on sait que toute famille libre a au plus autant de vecteurs qu'une famille génératrice.
□

Théorème

Soit E un $\mathbb{K}$-espace vectoriel de dimension $n \in \mathbb{N}$ et $\mathcal{B} = (\vec{e}_1, \ldots, \vec{e}_n)$ une famille formée d'exactement $n = \dim E$ vecteurs de E. On a équivalence entre :
(i) $\mathcal{B}$ est une base ;
(ii) $\mathcal{B}$ est une famille libre ;
(iii) $\mathcal{B}$ est une famille génératrice.

dém. :
Les implications (i) $\Rightarrow$ (ii) et (i) $\Rightarrow$ (iii) sont immédiates.
(ii) $\Rightarrow$ (i) Supposons la famille $\mathcal{B}$ libre et montrons qu'elle est génératrice.
Soit $\vec{x} \in E$. Si par l'absurde $\vec{x} \notin \text{Vect}(\vec{e}_1, \ldots, \vec{e}_n)$ alors par le principe d'extension des familles libres, on peut affirmer que la famille $(\vec{e}_1, \ldots, \vec{e}_n, \vec{x})$ est libre. Or celle-ci est formée de $n + 1 > \dim E$ vecteurs. C'est absurde et on peut donc conclure $\vec{x} \in \text{Vect}(\vec{e}_1, \ldots, \vec{e}_n)$.
Ainsi tout vecteur de E est combinaison linéaire des vecteurs de la famille $\mathcal{B}$ et celle-ci est donc génératrice.
(iii) $\Rightarrow$ (i) Supposons la famille $\mathcal{B}$ génératrice et montrons qu'elle est libre.
Par l'absurde, supposons que la famille $\mathcal{B}$ est liée.
L'un des vecteurs de $\mathcal{B}$ est combinaison linéaire des autres, notons i son indice. Par le principe de réduction des familles génératrices, la famille $(\vec{e}_1, \ldots, \widehat{\vec{e}_i}, \ldots, \vec{e}_n)$ est génératrice. Or cette famille est constituée de $n - 1 < \dim E$ vecteurs. C'est absurde et donc on peut affirmer que la famille $\mathcal{B}$ est libre.
□

Exemple Dans $\mathbb{R}^3$, considérons la famille $\mathcal{B} = (\vec{u}, \vec{v}, \vec{w})$ formée des vecteurs
$\vec{u} = (1, 1, 1), \vec{v} = (1, 2, 1), \vec{w} = (0, 1, 1)$ et $\mathcal{B} = (\vec{u}, \vec{v}, \vec{w})$.
Montrons que la famille $\mathcal{B}$ est une base de $\mathbb{R}^3$.
Supposons $\alpha\vec{u} + \beta\vec{v} + \gamma\vec{w} = \vec{0}$.
On a

$$(\alpha + \beta, \alpha + 2\beta + \gamma, \alpha + \beta + \gamma) = (0, 0, 0)$$

En résolvant le système formé par identification des éléments, on obtient $\alpha = \beta = \gamma = 0$.
La famille $\mathcal{B}$ est libre et formée de $3 = \dim \mathbb{R}^3$ vecteurs de $\mathbb{R}^3$, c'est donc une base de $\mathbb{R}^3$.
Déterminons les composantes du vecteur $\vec{x} = (1, 2, 3)$ dans cette base.
Il s'agit de déterminer $\alpha, \beta, \gamma \in \mathbb{R}$ vérifiant $\alpha\vec{u} + \beta\vec{v} + \gamma\vec{w} = \vec{x}$ i.e.

$$(\alpha + \beta, \alpha + 2\beta + \gamma, \alpha + \beta + \gamma) = (1, 2, 3)$$

En résolvant le système formé par identification des éléments, on obtient $\alpha = 2, \beta = -1, \gamma = 2$.

Exemple Soit E un $\mathbb{K}$-espace vectoriel de dimension 3 muni d'une base $\mathcal{B} = (\vec{e}_1, \vec{e}_2, \vec{e}_3)$.
Considérons $\mathcal{B}' = (\vec{\varepsilon}_1, \vec{\varepsilon}_2, \vec{\varepsilon}_3)$ la famille formée des vecteurs

$$\vec{\varepsilon}_1 = \vec{e}_2 + \vec{e}_3, \vec{\varepsilon}_2 = \vec{e}_1 + \vec{e}_3, \vec{\varepsilon}_3 = \vec{e}_1 + \vec{e}_2$$

Montrons que $\mathcal{B}' = (\vec{\varepsilon}_1, \vec{\varepsilon}_2, \vec{\varepsilon}_3)$ est une base de E.
Supposons $\lambda_1\vec{\varepsilon}_1 + \lambda_2\vec{\varepsilon}_2 + \lambda_3\vec{\varepsilon}_3 = \vec{0}_E$.
On a

$$(\lambda_2 + \lambda_3)\vec{e}_1 + (\lambda_1 + \lambda_2)\vec{e}_2 + (\lambda_1 + \lambda_3)\vec{e}_3 = \vec{0}$$

Or la famille $\mathcal{B}$ est libre donc $\lambda_2 + \lambda_3 = \lambda_1 + \lambda_3 = \lambda_1 + \lambda_2 = 0$.
Après résolution du système correspondant, on obtient $\lambda_1 = \lambda_2 = \lambda_3 = 0$.
La famille $\mathcal{B}'$ est libre et formée de $3 = \dim E$ vecteurs de E, c'est donc une base de E.
Soit $\vec{x}$ un vecteur de composantes $\lambda_1, \lambda_2, \lambda_3$ dans $\mathcal{B}$, déterminons ses composantes μ_1, μ_2, μ_3 dans $\mathcal{B}'$.

On veut déterminer les réels μ_1, μ_2, μ_3 vérifiant $\vec{x} = \mu_1\vec{\varepsilon}_1 + \mu_2\vec{\varepsilon}_2 + \mu_3\vec{\varepsilon}_3$ i.e.

$$\vec{x} = (\mu_2 + \mu_3)\vec{e}_1 + (\mu_1 + \mu_2)\vec{e}_2 + (\mu_1 + \mu_3)\vec{e}_3$$

Par unicité des composantes d'un vecteur dans une base, on a

$$\begin{cases} \mu_2 + \mu_3 = \lambda_1 \\ \mu_1 + \mu_2 = \lambda_2 \\ \mu_1 + \mu_3 = \lambda_3 \end{cases}$$

Après résolution de ce système, on obtient

$$\mu_1 = \frac{-\lambda_1 + \lambda_2 + \lambda_3}{2}, \mu_2 = \frac{\lambda_1 - \lambda_2 + \lambda_3}{2} \text{ et } \mu_3 = \frac{\lambda_1 + \lambda_2 - \lambda_3}{2}$$

Exemple Considérons le $\mathbb{R}$-espace vectoriel $\mathbb{C}$ et la famille $\mathcal{B} = (1, j)$.
Montrons que $\mathcal{B} = (1, j)$ est une base de $\mathbb{C}$.
Supposons $\alpha.1 + \beta.j = 0$ avec $\alpha, \beta \in \mathbb{R}$.
Par identification de partie imaginaire, on a $\beta = 0$ et on en déduit $\alpha = 0$.
Ainsi $\mathcal{B}$ est une famille libre formée de $2 = \dim_{\mathbb{R}} \mathbb{C}$ éléments de $\mathbb{C}$, c'est donc une base de $\mathbb{C}$.

7.2.4 Théorème de complétion de la base

Théorème

> Soit E un $\mathbb{K}$-espace vectoriel de dimension finie, $\mathcal{L}$ une famille libre et $\mathcal{G}$ une famille génératrice.
> On peut former une base de E en complétant la famille libre $\mathcal{L}$ par des vecteurs bien choisis dans la famille génératrice $\mathcal{G}$.

dém. :
Notons $\mathcal{L} = (\vec{e}_1, \ldots, \vec{e}_p)$ et $\mathcal{G} = (\vec{u}_1, \ldots, \vec{u}_m)$.
Si la famille $\mathcal{L}$ est génératrice, c'est une base de E et la complétion est inutile.
Sinon, la famille $\mathcal{L}$ n'est pas génératrice de E et par suite il existe $i \in \{1, \ldots, m\}$ tel que

$$\vec{u}_i \notin \text{Vect}(\vec{e}_1, ..., \vec{e}_p)$$

Posons alors $\vec{e}_{p+1} = \vec{u}_i$, par le principe d'extension des familles libres, la famille $(\vec{e}_1, ..., \vec{e}_p, \vec{e}_{p+1})$ est libre.
On a ainsi complété $\mathcal{L}$ à partir d'un vecteur de $\mathcal{G}$ de sorte d'obtenir une famille libre à $p + 1$ vecteurs.
On réitère ce processus jusqu'à obtention d'une famille génératrice ; ce processus s'arrête nécessairement car les il n'y a qu'un nombre fini de vecteurs dans la famille $\mathcal{G}$.
□

Corollaire

> Dans un $\mathbb{K}$-espace vectoriel de dimension finie :
> 1) Toute famille libre peut être complétée en une base.
> 2) De toute famille génératrice, on peut extraire une base.

dém. :
1) Il suffit de former une base en complétant la famille libre considérée à l'aide de vecteurs bien choisis

dans une famille génératrice de l'espace.
2) On peut former une base en complétant la famille vide par des vecteurs bien choisis dans la famille génératrice considérée. On peut aussi procéder en retirant de la famille génératrice les vecteurs combinaisons linéaires des autres jusqu'à obtention d'une famille libre.
$\square$

Exemple Dans l'espace $E = \mathbb{R}^3$, considérons les vecteurs $\vec{u} = (1, 2, 1)$ et $\vec{v} = (1, 1, 0)$.
Les vecteurs $\vec{u}$ et $\vec{v}$ ne sont pas colinéaires et donc la famille $(\vec{u}, \vec{v})$ est libre. Complétons-la en une base de E ce qui nécessite d'y adjoindre un vecteur puisque $\dim \mathbb{R}^3 = 3$.
Considérons la famille $(\vec{i}, \vec{j}, \vec{k})$ avec $\vec{i} = (1, 0, 0)$, $\vec{j} = (0, 1, 0)$ et $\vec{k} = (0, 0, 1)$.
La famille $(\vec{i}, \vec{j}, \vec{k})$ est génératrice de $\mathbb{R}^3$, on peut donc compléter $(\vec{u}, \vec{v})$ en une base de E à l'aide d'un vecteur bien choisi dans la famille $(\vec{i}, \vec{j}, \vec{k})$. Dans le cas présent, n'importe lequel convient et on vérifie en particulier que $(\vec{u}, \vec{v}, \vec{k})$ est une base de E car on démontre aisément que c'est une famille libre.

Exemple Soit E un $\mathbb{K}$-espace vectoriel muni d'une base $\mathcal{B} = (\vec{e}_1, \vec{e}_2, \vec{e}_3)$.
Considérons le vecteur $\vec{u} = \vec{e}_1 + \vec{e}_2$.
Puisque $\vec{u} \neq \vec{0}$, la famille $(\vec{u})$ est libre. Complétons-la en une base de E ce qui nécessite d'y adjoindre deux vecteurs puisque $\dim E = 3$.
Comme la famille $\mathcal{B} = (\vec{e}_1, \vec{e}_2, \vec{e}_3)$ est génératrice de E, on peut compléter $(\vec{u})$ à l'aide de vecteurs bien choisi dans cette famille.
Commençons par considérer le vecteur $\vec{e}_1$.
Les vecteurs $\vec{u}$ et $\vec{e}_1$ sont clairement non colinéaires et donc la famille $(\vec{u}, \vec{e}_1)$ est libre.
Pour poursuivre la complétion, le vecteur $\vec{e}_2$ n'est pas un bon choix car $\vec{u} = \vec{e}_1 + \vec{e}_2$. Prenons alors le vecteur $\vec{e}_3$, on peut affirmer que la famille $(\vec{u}, \vec{e}_1, \vec{e}_3)$ est base de E notamment parce que l'on sait qu'on peut compléter la famille $(\vec{u}, \vec{e}_1)$ en une base de E à l'aide d'un vecteur bien choisi dans $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ et que seul le vecteur $\vec{e}_3$ est possible.

7.2.5 Applications

7.2.5.1 Équation différentielle linéaire d'ordre 2 à coefficients constants

Soient $p, q \in \mathbb{K}$. On note $E(p, q)$ l'ensemble des fonctions de $\mathbb{R}$ vers $\mathbb{K}$ solutions de l'équation différentielle

$$y'' + py' + qy = 0$$

Théorème

L'ensemble $\mathcal{S}$ des fonctions de $\mathbb{R}$ vers $\mathbb{K}$ solutions de l'équation différentielle

$$E : y'' + py' + qy = 0 \text{ (avec } p, q \in \mathbb{K}\text{)}$$

est un plan vectoriel

dém. :
Montrons que $\mathcal{S}$ est un sous-espace vectoriel de $\mathcal{C}^2(\mathbb{R}, \mathbb{K})$.
Par définition les solutions de l'équation différentielle sont deux fois dérivables et de plus leur dérivée seconde est continue car s'exprimant en fonction de y et y'. On en déduit que les solutions de l'équation différentielle sont de classe $\mathcal{C}^2$ et donc $\mathcal{S} \subset \mathcal{C}^2(\mathbb{R}, \mathbb{K})$. On peut même, en approfondissant, affirmer que les solutions sont des fonctions de classe $\mathcal{C}^\infty$.

La fonction nulle est évidemment solution de l'équation E et donc $\tilde{0} \in \mathcal{S}$.
Soient $\lambda_1, \lambda_2 \in \mathbb{K}$ et $y_1, y_2 \in \mathcal{S}$.
Par linéarité de la dérivation, on a

$$(\lambda_1 y_1 + \lambda_2 y_2)'' + p(\lambda_1 y_1 + \lambda_2 y_2)' + q(\lambda_1 y_1 + \lambda_2 y_2) = \lambda_1(y_1'' + py_1' + qy_1) + \lambda_2(y_2'' + py_2' + qy_2) = 0$$

et donc $\lambda_1 y_1 + \lambda_2 y_2 \in \mathcal{S}$.
Ainsi $\mathcal{S}$ est un sous-espace vectoriel de $\mathcal{C}^2(\mathbb{R}, \mathbb{K})$ et c'est donc un $\mathbb{K}$-espace vectoriel.
Pour déterminer sa dimension, rappelons que pour tout $x_0 \in \mathbb{R}$ et tout $a, b \in \mathbb{K}$, il existe une unique solution à l'équation différentielle E vérifiant les conditions initiales $y(x_0) = a$ et $y'(x_0) = b$.
Considérons alors y_1 et y_2 les fonctions solutions déterminées par les conditions initiales

$$\begin{cases} y_1(0) = 1 \\ y_1'(0) = 0 \end{cases} \text{ et } \begin{cases} y_2(0) = 0 \\ y_2'(0) = 1 \end{cases}$$

Montrons que la famille (y_1, y_2) est base de $\mathcal{S}$.
Supposons $\lambda_1 y_1 + \lambda_2 y_2 = \tilde{0}$.
En évaluant en 0, on obtient $\lambda_1 = 0$.
En dérivant et en évaluant en 0, on obtient $\lambda_2 = 0$.
Ainsi la famille (y_1, y_2) est libre.
Montrons que la famille (y_1, y_2) est génératrice de $\mathcal{S}$.
Soit $y \in \mathcal{S}$. Considérons la fonction $\tilde{y} = y(0)y_1 + y'(0)y_2$.
Les fonctions y et $\tilde{y}$ sont solutions de E et vérifient par construction

$$\begin{cases} y(0) = \tilde{y}(0) \\ y'(0) = \tilde{y}'(0) \end{cases}$$

Puisque des conditions initiales données déterminent une solution unique, on peut affirmer $y = \tilde{y}$.
Ainsi la famille (y_1, y_2) est génératrice et c'est donc une base de $\mathcal{S}$.
Finalement $\mathcal{S}$ est un $\mathbb{K}$-espace vectoriel de dimension 2.
□

Remarque Pour décrire complètement les éléments de $\mathcal{S}$, il suffit de connaître deux solutions linéairement indépendantes.

Exemple Supposons $\mathbb{K} = \mathbb{C}$.
Introduisons l'équation caractéristique $r^2 + pr + q = 0$ de discriminant Δ.
Si $\Delta \neq 0$ alors l'équation caractéristique admet deux racines distinctes α et $\beta \in \mathbb{C}$.
On vérifie alors aisément que $x \mapsto \mathrm{e}^{\alpha x}$ et $x \mapsto \mathrm{e}^{\beta x}$ sont solutions de E et puisque celles-ci sont linéairement indépendante, on peut affirmer que la solution générale de E est

$$y(x) = \lambda \mathrm{e}^{\alpha x} + \mu \mathrm{e}^{\beta x} \text{ avec } \lambda, \mu \in \mathbb{C}$$

Si $\Delta = 0$ alors l'équation caractéristique admet une racine double $\alpha \in \mathbb{C}$.
On vérifie par le calcul des les fonctions $x \mapsto \mathrm{e}^{\alpha x}$ et $x \mapsto x\,\mathrm{e}^{\alpha x}$ sont solutions de E et puisque celles-ci sont linéairement indépendante, on peut affirmer que la solution générale de E est

$$y(x) = (\lambda + \mu x)\mathrm{e}^{\beta x} \text{ avec } \lambda, \mu \in \mathbb{C}$$

Remarque On peut transposer ce qui précède au cadre $\mathbb{K} = \mathbb{R}$.

7.2.5.2 Suites récurrentes linéaires d'ordre 2

Soient $(p, q) \in \mathbb{K} \times \mathbb{K}^\star$ et $u = (u_n) \in \mathbb{K}^\mathbb{N}$ vérifiant

$$\forall n \in \mathbb{N}, u_{n+2} + pu_{n+1} + qu_n = 0.$$

Définition

On dit que (u_n) est une suite récurrente linéaire d'ordre 2.

Notons $\mathcal{S}$ l'ensemble des suites de $\mathbb{K}^\mathbb{N}$ vérifiant

$$\forall n \in \mathbb{N}, u_{n+2} + pu_{n+1} + qu_n = 0$$

Proposition

Pour tout $a, b \in \mathbb{K}$, il existe une unique suite $u = (u_n) \in \mathcal{S}$ vérifiant les conditions initiales $u_0 = a$ et $u_1 = b$.

dém. :
Ces conditions initiales déterminent entièrement les termes successifs de la suite.
□

Théorème

$\mathcal{S}$ est un plan vectoriel.

dém. :
Montrons pour commencer que $\mathcal{S}$ est un sous-espace vectoriel de $\mathbb{K}^\mathbb{N}$.
On a $\mathcal{S} \subset \mathbb{K}^\mathbb{N}$ et on vérifie aisément $(0)_{n\in\mathbb{N}} \in \mathcal{S}$.
Soient $\lambda, \mu \in \mathbb{K}$ et $u, v \in \mathcal{S}$. Puisque $\lambda u + \mu v = (\lambda u_n + \mu v_n)_{n\in\mathbb{N}}$, on a

$$(\lambda u+\mu v)_{n+2}+p(\lambda u+\mu v)_{n+1}+q(\lambda u+\mu v)_n = \lambda(u_{n+2}+pu_{n+1}+qu_n)+\mu(v_{n+2}+pv_{n+1}+qv_n) = 0$$

Ainsi $\lambda u + \mu v \in \mathcal{S}$.
$\mathcal{S}$ est donc un sous-espace vectoriel de $\mathbb{K}^\mathbb{N}$ et c'est donc un $\mathbb{K}$-espace vectoriel.
Pour déterminer sa dimension, formons une base de $\mathcal{S}$. Pour cela considérons les deux suites $v, w \in \mathcal{S}$ déterminée par les conditions initiales

$$\begin{cases} v_0 = 1 \\ v_1 = 0 \end{cases} \text{ et } \begin{cases} w_0 = 0 \\ w_1 = 1 \end{cases}$$

Supposons $\lambda v + \mu w = 0$.
Les termes d'indices 0 et 1 de cette égalité de suites donnent $\lambda = 0$ et $\mu = 0$.
La famille (v, w) est donc libre.
Soit $u \in \mathcal{S}$. Les suites u et $u_0 v + u_1 w$ sont éléments de $\mathcal{S}$ et, par construction, elles ont les mêmes termes de rang 0 et de rang 1. On en déduit qu'elles sont donc égales et donc la famille (v, w) est un génératrice de $\mathcal{S}$.
Finalement (v, w) est une base de $\mathcal{S}$ et donc $\mathcal{S}$ est un $\mathbb{K}$-espace vectoriel de dimension 2.
□

Remarque Il n'est pas aisé d'exprimer les termes généraux des suites v et w. Pour cette raison, nous allons chercher des suites plus simples éléments de $\mathcal{S}$, à commencer par les suites géométrique.

Proposition

On a équivalence entre :
(i) la suite géométrique (r^n) appartient à $\mathcal{S}$;
(ii) r est solution de l'équation $r^2 + pr + q = 0$.

dém. :
(i) $\Rightarrow$ (ii) Si la suite (r^n) vérifie $\forall n \in \mathbb{N}, r^{n+2} + pr^{n+1} + qr^n = 0$ alors pour $n = 0$ on a $r^2 + pr + q = 0$.
(ii) $\Rightarrow$ (i) Si $r^2 + pr + q = 0$ alors pour tout $n \in \mathbb{N}$, $r^n(r^2 + pr + q) = 0$ et donc $r^{n+2} + pr^{n+1} + qr^n = 0$.
□

Définition

L'équation $r^2 + pr + q = 0$ d'inconnue $r \in \mathbb{C}$ est appelée équation caractéristique associée à la relation de récurrence $u_{n+2} + pu_{n+1} + qu_n = 0$.

Exemple On suppose $\mathbb{K} = \mathbb{C}$.
Notons Δ le discriminant de l'équation caractéristique $r^2 + pr + q = 0$.
Si $\Delta \neq 0$ alors l'équation caractéristique possède deux solutions distinctes r et $s \in \mathbb{C}^\star$.
Considérons alors les suites $v = (r^n)$ et $w = (s^n)$.
Les suites v et w sont éléments de $\mathcal{S}$ et on vérifie aisément que la famille (v, w) est libre. La famille (v, w) est donc une base de $\mathcal{S}$ et par suite

$$\forall u \in \mathcal{S}, \exists (\lambda, \mu) \in \mathbb{C}^2, \forall n \in \mathbb{N}, u_n = \lambda r^n + \mu s^n.$$

Si $\Delta = 0$ alors l'équation caractéristique possède une solution double $r = -p/2 \in \mathbb{C}^\star$
Considérons alors les suites $v = (r^n)$ et $w = (nr^n)$.
La suite v est élément de $\mathcal{S}$ et on vérifie par le calcul que la suite w l'est aussi. De plus, la famille (v, w) est libre et c'est donc une base de $\mathcal{S}$:

$$\forall u \in \mathcal{S}, \exists (\lambda, \mu) \in \mathbb{C}^2, \forall n \in \mathbb{N}, u_n = (\lambda + \mu n) r^n.$$

Exemple Déterminons le terme général de la suite (u_n) définie par :

$$\begin{cases} u_0 = 2, u_1 = 1 \\ \forall n \in \mathbb{N}, u_{n+2} - u_{n+1} + (1+i)u_n = 0 \end{cases}$$

(u_n) est une suite récurrente linéaire d'ordre 2 d'équation caractéristique $r^2 - r + (1+i) = 0$ de racines $1 - i$ et i.
Le terme général de la suite (u_n) est donc de la forme

$$u_n = \lambda(1-i)^n + \mu i^n \text{ avec } \lambda, \mu \in \mathbb{C}$$

Les conditions initiales $u_0 = 2$ et $u_1 = 1$ déterminent λ et μ.
On obtient

$$u_n = (1-i)^n + i^n$$

Exemple On suppose $\mathbb{K} = \mathbb{R}$.
Notons Δ le discriminant de l'équation caractéristique $r^2 + pr + q = 0$.
Si $\Delta > 0$ alors l'équation caractéristique possède deux solutions distinctes r et $s \in \mathbb{R}^\star$.

Comme ci-dessus, on obtient

$$\forall u \in \mathcal{S}, \exists(\lambda,\mu) \in \mathbb{R}^2, \forall n \in \mathbb{N}, u_n = \lambda r^n + \mu s^n.$$

Si $\Delta = 0$ alors l'équation caractéristique possède une solution double $r = -p/2 \in \mathbb{R}^\star$
Comme ci-dessus

$$\forall u \in \mathcal{S}, \exists(\lambda,\mu) \in \mathbb{C}^2, \forall n \in \mathbb{N}, u_n = (\lambda + \mu n) r^n.$$

Si $\Delta < 0$ alors l'équation caractéristique possède deux racines complexes conjuguées $z = \rho \mathrm{e}^{\pm i\theta}$ (avec $\rho > 0$ et $\theta \neq 0 \quad [\pi]$).
Puisque $z^{n+2} + pz^{n+1} + qz^n = 0$ en considérant les parties réelle et imaginaire de cette égalité, on obtient

$$v_{n+2} + pv_{n+1} + qv_n = 0 \text{ et } w_{n+2} + pw_{n+1} + qw_n = 0$$

avec $v_n = \mathrm{Re}(z^n)$ et $w_n = \mathrm{Im}(z^n)$. Ainsi les suites réelles $v = (\rho^n \cos n\theta)$ et $w = (\rho^n \sin n\theta)$ appartiennent à $\mathcal{S}$. De plus, on vérifie que la famille (v, w) est libre et c'est donc une base de $\mathcal{S}$. Ainsi

$$\forall u \in \mathcal{S}, \exists(\lambda,\mu) \in \mathbb{R}^2, \forall n \in \mathbb{N}, u_n = (\lambda \cos n\theta + \mu \sin n\theta)\rho^n.$$

Exemple Soit (u_n) la suite réelle déterminée par :

$$\begin{cases} u_0 = 1, u_1 = -2 \\ \forall n \in \mathbb{N}, u_{n+2} + 2u_{n+1} + u_n = 0 \end{cases}$$

(u_n) est une suite récurrente linéaire d'ordre 2 d'équation caractéristique $r^2 + 2r + 1 = 0$ de racine double -1.
Le terme général de la suite (u_n) est donc de la forme

$$u_n = (\lambda n + \mu)(-1)^n \text{ avec } \lambda, \mu \in \mathbb{R}$$

Par les conditions initiales, on a $\lambda = 1$ et $\mu = 1$.
Finalement

$$\forall n \in \mathbb{N}, u_n = (-1)^n(n+1)$$

Exemple Soit (u_n) la suite réelle déterminée par :

$$\begin{cases} u_0 = 1, u_1 = 1 \\ \forall n \in \mathbb{N}, u_{n+2} + u_{n+1} + u_n = 0 \end{cases}$$

(u_n) est une suite récurrente linéaire d'ordre 2 d'équation caractéristique $r^2 + r + 1 = 0$ de racines complexes conjuguées $j = \exp(2i\pi/3)$ et j^2.
Le terme général de la suite (u_n) est donc de la forme

$$u_n = \lambda \cos \frac{2n\pi}{3} + \mu \sin \frac{2n\pi}{3} \text{ avec } \lambda, \mu \in \mathbb{R}$$

Par les conditions initiales, on obtient $\lambda = 1$ et $\mu = \sqrt{3}$.
Finalement

$$\forall n \in \mathbb{N}, u_n = 2\cos \frac{(2n-1)\pi}{3}$$

7.3 Sous-espace vectoriel de dimension finie

7.3.1 Dimension d'un sous espace vectoriel d'un espace de dimension finie

Théorème

Tout sous-espace vectoriel F d'un $\mathbb{K}$-espace vectoriel est de dimension finie.

dém. :
Soit F un sous-espace vectoriel de E.
Si $F = \left\{\vec{0}\right\}$ alors F est de dimension finie.
Sinon, il existe un vecteur non nul $\vec{e}_1 \in F$ et la famille $(\vec{e}_1)$ est libre.
Si elle génératrice de F, c'est une base de F. Sinon, il existe un vecteur $\vec{e}_2 \in F$ qui n'est pas colinéaire à $\vec{e}_1$ et alors la famille $(\vec{e}_1, \vec{e}_2)$ est libre.
On peut alors reprendre la discussion précédente, si $(\vec{e}_1, \vec{e}_2)$ est génératrice de F, c'est une base de F, sinon on la complète en une famille libre à l'aide d'un vecteur bien choisi dans F etc.
Le processus s'arrête nécessairement car une famille libre d'éléments de F est une famille libre de vecteurs de E et possède donc au plus $\dim E$ vecteurs.
□

Corollaire

Si F est un sous-espace vectoriel d'un $\mathbb{K}$-espace vectoriel E de dimension finie alors

$$\dim F \leqslant \dim E$$

avec égalité si, et seulement si, $F = E$.

dém. :
Soit F un sous-espace vectoriel d'un $\mathbb{K}$-espace vectoriel E de dimension finie.
L'espace F admet une base $\mathcal{B} = (e_1, \ldots, e_p)$ avec $p = \dim F$.
Puisque la famille $\mathcal{B}$ est une famille libre de p vecteurs de E on a $p \leqslant \dim E$.
Ainsi $\dim F \leqslant \dim E$.
De plus, il y a égalité si, et seulement si, $\mathcal{B}$ est une base de E ce qui revient à dire que les espaces E et F sont égaux.
□

Exemple Si $\dim F = 0$ alors $F = \left\{\vec{0}\right\}$, c'est l'espace nul.

Exemple Si $\dim F = 1$ alors F est une droite vectorielle.
On a alors $F = \text{Vect}(\vec{u})$ pour tout vecteur non nul $\vec{u} \in F$.
On dit que $\vec{u}$ est un vecteur directeur de F.

Exemple Si $\dim F = 2$ alors F est un plan vectoriel.
On a alors $F = \text{Vect}(\vec{u}, \vec{v})$ pour tous vecteurs $\vec{u}, \vec{v} \in F$ non colinéaires.
On dit que $\vec{u}$ et $\vec{v}$ sont des vecteurs directeurs de F.

Exemple Si $\dim F = \dim E - 1$ alors on dit que F est un hyperplan vectoriel de E.

7.3.2 Construction de base d'un sous-espace vectoriel

7.3.2.1 Espace vectoriel engendré par une famille de vecteurs

On désire déterminer une base d'un espace $F = \text{Vect}(\vec{e}_1, \dots, \vec{e}_p)$.
La famille $(\vec{e}_1, \dots, \vec{e}_p)$ est génératrice de F, on peut donc en extraire une base de la manière suivante :
- si $(\vec{e}_1, \dots, \vec{e}_p)$ est libre alors c'est une base de F ;
- sinon, l'un des vecteurs de $(\vec{e}_1, \dots, \vec{e}_p)$ est combinaison linéaire des autres et on peut retirer celui-ci sans perdre le caractère générateur de la famille ;
- si besoin, on reprend ce processus jusqu'à obtention d'une famille libre donc d'une base.

Exemple Dans l'espace $\mathbb{R}^3$, considérons

$$F = \{(a + 2b + c, b + c, a + b)/a, b, c \in \mathbb{R}\}$$

On a

$$F = \{a\vec{u} + b\vec{v} + c\vec{w}/a, b, c \in \mathbb{R}\}$$

avec $\vec{u} = (1, 0, 1)$, $\vec{v} = (2, 1, 1)$ et $\vec{w} = (1, 1, 0)$.
Par suite $F = \text{Vect}(\vec{u}, \vec{v}, \vec{w})$.
La famille $(\vec{u}, \vec{v}, \vec{w})$ est génératrice de F.
On remarque la relation linéaire $\vec{u} - \vec{v} + \vec{w} = \vec{0}$.
Par suite le vecteur $\vec{w}$, par exemple, est combinaison linéaire des vecteurs $\vec{u}$ et $\vec{v}$ et donc $F = \text{Vect}(\vec{u}, \vec{v})$.
Or les vecteurs $\vec{u}$ et $\vec{v}$ ne sont pas colinéaires donc la famille $(\vec{u}, \vec{v})$ est une base de F.

7.3.2.2 Espace défini par des équations linéaires

Si un espace est défini par des équations linéaires, on peut, en résolvant celles-ci, déterminer une famille génératrice puis une base de cet espace.

Exemple Dans l'espace $\mathbb{R}^3$, considérons

$$F = \left\{(x, y, z) \in \mathbb{R}^3/x + y + z = 0, 2x - y + z = 0\right\}$$

Résolvons le système

$$\begin{cases} x + y + z = 0 \\ 2x - y + z = 0 \end{cases}$$

On obtient pour solution

$$\begin{cases} x = 2y \\ z = -3y \end{cases}$$

On en déduit

$$F = \{(2y, y, -3y)/y \in \mathbb{R}\} = \text{Vect}\vec{u}$$

avec $\vec{u} = (2, 1, -3)$.
Puisque $\vec{u} \neq \vec{0}$, la famille $(\vec{u})$ est une base de F.

Exemple Dans l'espace $\mathbb{R}^4$, considérons

$$F = \left\{(x, y, z, t) \in \mathbb{R}^4/x + 2y + 3z + 4t = 0\right\}$$

L'équation définissant F équivaut à l'équation résolue $x = -2y - 3y - 4t$.
On en déduit

$$F = \{(-2y - 3z - 4t, y, z, t)/y, z, t \in \mathbb{R}\} = \text{Vect}(\vec{u}, \vec{v}, \vec{w})$$

avec $\vec{u} = (-2, 1, 0, 0)$, $\vec{v} = (-3, 0, 1, 0)$ et $\vec{w} = (-4, 0, 0, 1)$.
Il est aisé de vérifier que la famille $(\vec{u}, \vec{v}, \vec{w})$ est libre et donc c'est une base de F.

7.3.3 Supplémentaire en dimension finie

Théorème

> Si F et G sont des sous-espaces vectoriels supplémentaires d'un $\mathbb{K}$-espace vectoriel E de dimension finie alors on peut former une base de E en accolant une base de F et une base de G.
> En particulier $\dim E = \dim F + \dim G$.

dém. :
Soient $\mathcal{B} = (\vec{e}_1, ..., \vec{e}_p)$ et $\mathcal{C} = (\vec{f}_1, ..., \vec{f}_q)$ des bases des espaces F et G.
Formons la famille $\mathcal{D} = (\vec{e}_1, ..., \vec{e}_p, \vec{f}_1, ..., \vec{f}_q)$ et montrons que c'est une base de E.
Supposons

$$\lambda_1\vec{e}_1 + \cdots + \lambda_p\vec{e}_p + \mu_1\vec{f}_1 + \cdots + \mu_q\vec{f}_q = \vec{0}$$

On a

$$\lambda_1\vec{e}_1 + \cdots + \lambda_p\vec{e}_p = -\left(\mu_1\vec{f}_1 + \cdots + \mu_q\vec{f}_q\right) \in F \cap G$$

Or les espaces F et G sont supplémentaires donc $F \cap G = \left\{\vec{0}\right\}$ et par suite

$$\lambda_1\vec{e}_1 + \cdots + \lambda_p\vec{e}_p = \mu_1\vec{f}_1 + \cdots + \mu_q\vec{f}_q = \vec{0}$$

Puisque les familles $\mathcal{B}$ et $\mathcal{C}$ sont libres on a alors $\lambda_1 = \ldots = \lambda_p = 0$ et $\mu_1 = \ldots = \mu_q = 0$.
Ainsi la famille $\mathcal{D}$ est libre. Montrons maintenant qu'elle est génératrice.
Soit $\vec{x} \in E$. Puisque $E = F + G$, on peut écrire $\vec{x} = \vec{a} + \vec{b}$ avec $\vec{a} \in F$ et $\vec{b} \in G$.
Puisque $\vec{a} \in F$ et que $\mathcal{B}$ est base de F, on peut écrire $\vec{a} = \lambda_1\vec{e}_1 + \cdots + \lambda_p\vec{e}_p$.
De même, on peut écrire $\vec{b} = \mu_1\vec{f}_1 + \cdots + \mu_q\vec{f}_q$ et ainsi

$$\vec{x} = \lambda_1\vec{e}_1 + \cdots + \lambda_p\vec{e}_p + \mu_1\vec{f}_1 + \cdots + \mu_q\vec{f}_q$$

Par suite $\mathcal{D}$ est une famille génératrice et finalement c'est une base de E.
□

Définition

> Toute base de E obtenue en accolant une base de F et une base de G est dite adaptée à la supplémentarité de F et G.

Théorème

> Soit E un $\mathbb{K}$-espace vectoriel de dimension finie.
> Tout sous-espace vectoriel F de E possède au moins un supplémentaire.

dém. :
Soit F un sous-espace vectoriel de E et $\mathcal{B} = (\vec{e}_1, ..., \vec{e}_p)$ une base de F avec $p = \dim F$.
La famille $\mathcal{B}$ est une famille libre de vecteurs de E, on peut la compléter en une base de E de la forme

$\mathcal{B}' = (\vec{e}_1, ..., \vec{e}_p, \vec{e}_{p+1}, ..., \vec{e}_n)$.
Considérons alors l'espace $G = \text{Vect}(\vec{e}_{p+1}, ..., \vec{e}_n)$ et montrons que G est un supplémentaire de F.
Soit $\vec{x} \in F \cap G$.
On peut écrire $\vec{x} = \lambda_1\vec{e}_1 + \cdots + \lambda_p\vec{e}_p$ et $\vec{x} = \lambda_{p+1}\vec{e}_{p+1} + \cdots + \lambda_n\vec{e}_n$.
On a alors

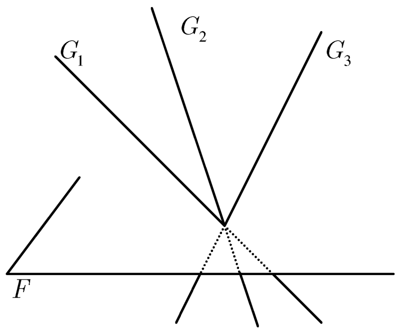
$$\lambda_1\vec{e}_1 + \cdots + \lambda_p\vec{e}_p - (\lambda_{p+1}\vec{e}_{p+1} + \cdots + \lambda_n\vec{e}_n) = \vec{0}$$

Or la famille $\mathcal{B}'$ est libre donc $\lambda_1 = \ldots = \lambda_p = -\lambda_{p+1} = \ldots = -\lambda_n = 0$.
On en déduit $\vec{x} = \vec{0}$ et ainsi $F \cap G = \left\{\vec{0}\right\}$.
Soit $\vec{x} \in E$.
Puisque $\mathcal{B}'$ est génératrice de E, on peut écrire

$$\vec{x} = \lambda_1\vec{e}_1 + \cdots + \lambda_p\vec{e}_p + \lambda_{p+1}\vec{e}_{p+1} + \cdots + \lambda_n\vec{e}_n$$

On a alors $\vec{x} = \vec{a} + \vec{b}$ avec $\vec{a} = \lambda_1\vec{e}_1 + \cdots + \lambda_p\vec{e}_p \in F$ et $\vec{b} = \lambda_{p+1}\vec{e}_{p+1} + \cdots + \lambda_n\vec{e}_n \in G$.
Ainsi $E = F + G$ et finalement F et G sont supplémentaires dans E.
□

Attention : Il y a existence d'un supplémentaire mais pas unicité !



Remarque Pour former un supplémentaire de F, il suffit de compléter une base de F en une base de E et de considérer l'espace engendré par les vecteurs complétant.

Exemple Dans l'espace $\mathbb{R}^3$ considérons $F = \text{Vect}(\vec{u}, \vec{v})$ avec $\vec{u} = (1, 1, 1)$ et $\vec{v} = (0, 1, 1)$.
Pour déterminer un supplémentaire de F, il suffit de compléter la famille libre $(\vec{u}, \vec{v})$ en une base de $\mathbb{R}^3$.
On vérifie aisément que le vecteur $\vec{j} = (0, 1, 0)$ convient (aussi $\vec{k}$ mais pas $\vec{i}$)
Par suite $G = \text{Vect}(\vec{j})$ est un sous-espace vectoriel supplémentaire de F dans $\mathbb{R}^3$.

7.3.4 Théorème des quatre dimensions

Théorème

Si F et G sont deux sous-espaces vectoriels de dimension finie d'un $\mathbb{K}$-espace vectoriel E alors $F + G$ et $F \cap G$ sont de dimensions finies et

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

dém. :
La réunion d'une famille génératrice de F et d'une famille génératrice de G donne une famille génératrice de $F+G$. Par suite l'espace $F+G$ est de dimension finie.
Puisque $F\cap G$ est un sous-espace vectoriel de F, l'espace $F\cap G$ possède un supplémentaire H dans F.
Montrons que H et G sont supplémentaires dans $F+G$.
Puisque $H\subset F$, on a $H=H\cap F$ et donc

$$H\cap G=H\cap F\cap G=H\cap(F\cap G)=\left\{\vec{0}\right\}$$

Puisque $F+G=H+(F\cap G)+G$ avec $F\cap G\subset G$, on a aussi $F+G=H+G$.
Ainsi H et G sont supplémentaires dans $F+G$.
Par supplémentarité de H et $F\cap G$ dans F on a $\dim F=\dim H+\dim(F\cap G)$.
Par supplémentarité de H et G dans $F+G$ on a aussi $\dim(F+G)=\dim H+\dim G$
On en déduit

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

□

7.3.5 Applications de la dimension

Théorème

> Soient F et G deux sous-espaces vectoriels de dimensions finies d'un $\mathbb{K}$-espace vectoriel E.
> Si $F\subset G$ et si $\dim F=\dim G$ alors $F=G$.

dém. :
Posons $p=\dim F$ et considérons $\mathcal{B}=(\vec{e}_1,...,\vec{e}_p)$ une base de F.
$\mathcal{B}$ est une famille libre formée de $p=\dim G$ vecteurs de G.
C'est donc une base de G et par suite $G=\text{Vect}(\vec{e}_1,...,\vec{e}_p)=F$.
□

Exemple Soient P_1 et P_2 deux plans vectoriels distincts d'une $\mathbb{K}$-espace vectoriel E de dimension 3.
Montrons que $P_1\cap P_2$ est une droite vectorielle.
Etudions P_1+P_2.
P_1+P_2 est un sous-espace vectoriel de E donc $\dim P_1+P_2\leqslant 3$.
P_1+P_2 contient P_1 donc $\dim(P_1+P_2)\geqslant 2$.
Si $\dim(P_1+P_2)=2$ alors par inclusion et égalité des dimensions, on a $P_1=P_1+P_2$. Par un argument identique on a aussi $P_2=P_1+P_2$ et donc $P_1=P_2$ ce qui est exclu.
On en déduit $\dim(P_1+P_2)=3$ et par la formule de Grassman

$$\dim(P_1\cap P_2)=\dim P_1+\dim P_2-\dim(P_1+P_2)=1$$

Théorème

> Soient F et G deux sous-espaces vectoriels d'un $\mathbb{K}$-espace vectoriel E de dimension finie vérifiant
>
> $$\dim F+\dim G=\dim E$$
>
> On a équivalence entre :
> (i) F et G sont supplémentaires dans E ;
> (ii) $F+G=E$;
> (iii) $F\cap G=\left\{\vec{0}\right\}$.

dém. :
(i) $\Rightarrow$ (ii) ok
(ii) $\Rightarrow$ (iii) Supposons $F+G=E$.
Par la formule de Grassman $\dim(F\cap G)=\dim F+\dim G-\dim E=0$ et donc $F\cap G=\left\{\vec{0}\right\}$.
(iii) $\Rightarrow$ (i) Supposons $F\cap G=\left\{\vec{0}\right\}$.
Par la formule de Grassman $\dim(F+G)=\dim F+\dim G-0=\dim E$. Par inclusion et égalité des dimensions on a $F+G=E$ et donc F et G sont supplémentaires dans E.
□

Exemple Soit H un hyperplan d'un $\mathbb{K}$-espace vectoriel E de dimension finie.
Montrons que si $\vec{a}\in E\backslash H$ alors H et Vect$(\vec{a})$ sont supplémentaires dans E.
Supposons $\vec{a}\in E\backslash H$.
Puisque $\vec{0}\in H$, on a $\vec{a}\neq\vec{0}$ et donc $\dim$ Vect$(\vec{a})=1$.
On en déduit $\dim H+\dim$ Vect$(\vec{a})=\dim E$ car H est un hyperplan.
Etudions alors $H\cap$ Vect$(\vec{a})$.
Soit $\vec{x}\in H\cap$ Vect$(\vec{a})$.
On peut écrire $\vec{x}=\lambda\vec{a}$ avec $\lambda\in\mathbb{K}$.
Si $\lambda\neq 0$ alors $\vec{a}=\dfrac{1}{\lambda}\vec{x}\in H$ ce qui est exclu.
Nécessairement $\lambda=0$ et donc $\vec{x}=\vec{0}$.
Ainsi $H\cap$ Vect$(\vec{a})=\left\{\vec{0}\right\}$ et par l'hypothèse de dimension précédemment soulignée, on peut affirmer que H et Vect$(\vec{a})$ sont supplémentaires dans E.

7.3.6 Rang d'une famille de vecteurs

Soit $\mathcal{F}=(\vec{e}_j)_{1\leqslant j\leqslant p}$ une famille de vecteurs d'un $\mathbb{K}$-espace vectoriel E.

7.3.6.1 Définition

Définition

On appelle rang de la famille $\mathcal{F}$ la dimension, notée rg$\mathcal{F}$, de l'espace vectoriel engendré par cette famille. Ainsi

$$\mathrm{rg}\mathcal{F}=\mathrm{rg}(\vec{e}_1,\ldots,\vec{e}_p)=\dim \mathrm{Vect}(\vec{e}_1,\ldots,\vec{e}_p)$$

Exemple

$$\mathrm{rg}(\vec{u})=\begin{cases}1 & \text{si } \vec{u}\neq\vec{0}\\ 0 & \text{si } \vec{u}=\vec{0}\end{cases}$$

Exemple

$$\mathrm{rg}(\vec{u},\vec{v})=\begin{cases}2 & \text{si } \vec{u},\vec{v} \text{ non colinéaires}\\ 1 & \text{si } \vec{u},\vec{v} \text{ colinéaires, non tous deux nuls}\\ 0 & \text{si } \vec{u}=\vec{v}=\vec{0}\end{cases}$$

Théorème

On a $\text{rg}(\vec{e}_1, \ldots, \vec{e}_p) \leqslant p$ avec égalité si, et seulement si, la famille est libre.

dém. :
$(\vec{e}_1, \ldots, \vec{e}_p)$ est une famille génératrice de l'espace $\text{Vect}(\vec{e}_1, \ldots, \vec{e}_p)$ donc $\dim \text{Vect}(\vec{e}_1, \ldots, \vec{e}_p) \leqslant p$.
Si la famille $(\vec{e}_1, \ldots, \vec{e}_p)$ est libre, c'est alors une base de $\text{Vect}(\vec{e}_1, \ldots, \vec{e}_p)$ et donc $\dim \text{Vect}(\vec{e}_1, \ldots, \vec{e}_p) = p$.
Inversement si $\dim \text{Vect}(\vec{e}_1, \ldots, \vec{e}_p) = p$ alors la famille $(\vec{e}_1, \ldots, \vec{e}_p)$ est génératrice et formée de $p = \dim \text{Vect}(\vec{e}_1, \ldots, \vec{e}_p)$ vecteurs de $\text{Vect}(\vec{e}_1, \ldots, \vec{e}_p)$, c'est donc une base de $\text{Vect}(\vec{e}_1, \ldots, \vec{e}_p)$ et en particulier cette famille est libre.
□

Théorème

On a $\text{rg}(\vec{e}_1, \ldots, \vec{e}_p) \leqslant \dim E$ avec égalité si, et seulement si, la famille est génératrice.

dém. :
L'espace $\text{Vect}(\vec{e}_1, \ldots, \vec{e}_p)$ est inclus dans E donc $\dim \text{Vect}(\vec{e}_1, \ldots, \vec{e}_p) \leqslant \dim E$.
De plus, puisqu'il y a inclusion, il y a égalité des dimensions si, et seulement si, les espaces sont égaux i.e. si, et seulement si, la famille $(\vec{e}_1, \ldots, \vec{e}_p)$ est génératrice de E.
□

Corollaire

$(\vec{e}_1, \ldots, \vec{e}_p)$ est une base de E si, et seulement si, $\text{rg}(\vec{e}_1, \ldots, \vec{e}_p) = p = \dim E$.

7.3.6.2 Opérations

Proposition

On ne modifie pas le rang d'une famille de vecteur lorsque :
- on retire le vecteur nul de la famille si celui-ci y apparaît ;
- on permute les vecteurs de la famille ;
- on multiple un vecteur par un scalaire $\lambda \neq 0$;
- on ajoute à un vecteur une combinaison linéaire des autres.

dém. :
Les différentes manipulations proposées ne modifient pas l'espace vectoriel engendré par la famille ni a fortiori son rang. Cela est immédiat pour les trois premières manipulations, détaillons la quatrième manipulation.
Soit $(\vec{e}_1, \ldots, \vec{e}_p)$ une famille de vecteurs d'un $\mathbb{K}$-espace vectoriel E. Quitte à permuter les vecteurs, supposons qu'on ajoute au vecteur $\vec{e}_p$ une combinaison linéaire $\vec{x} = \lambda_1 \vec{e}_1 + \cdots + \lambda_{p-1}\vec{e}_{p-1}$.
Puisqu'une combinaison linéaire des vecteurs $\vec{e}_1, \ldots, \vec{e}_{p-1}$ et $\vec{e}_p + \vec{x}$ est aussi une combinaison linéaire des vecteurs $\vec{e}_1, \ldots, \vec{e}_{p-1}, \vec{e}_p$ on a déjà

$$\text{Vect}(\vec{e}_1, \ldots, \vec{e}_{p-1}, \vec{e}_p + \vec{x}) \subset \text{Vect}(\vec{e}_1, \ldots, \vec{e}_p)$$

Par le même argument, on a aussi

$$\text{Vect}(\vec{e}_1, \ldots, \vec{e}_p) = \text{Vect}(\vec{e}_1, \ldots, (\vec{e}_p + \vec{x}) - \vec{x}) \subset \text{Vect}(\vec{e}_1, \ldots, \vec{e}_{p-1}, \vec{e}_p + \vec{x})$$

et on peut donc conclure à l'égalité

$$\text{Vect}(\vec{e}_1, \ldots, \vec{e}_{p-1}, \vec{e}_p + \vec{x}) = \text{Vect}(\vec{e}_1, \ldots, \vec{e}_p)$$

□

Exemple Soit E un $\mathbb{K}$-espace vectoriel muni d'une base $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$.
Déterminons le rang de la famille $(\vec{\varepsilon}_1, \vec{\varepsilon}_2, \vec{\varepsilon}_3)$ avec

$$\vec{\varepsilon}_1 = \vec{e}_2 + 2\vec{e}_1, \vec{\varepsilon}_2 = \vec{e}_1 \text{ et } \vec{\varepsilon}_3 = \vec{e}_1 - \vec{e}_2 + 3\vec{e}_3$$

Par permutation des vecteurs

$$\text{rg}(\vec{\varepsilon}_1, \vec{\varepsilon}_2, \vec{\varepsilon}_3) = \text{rg}(\vec{e}_1, \vec{e}_2 + 2\vec{e}_1, \vec{e}_1 - \vec{e}_2 + 3\vec{e}_3)$$

On ajoutant $-2\vec{e}_1$ au deuxième vecteur et $-\vec{e}_1$ au troisième

$$\text{rg}(\vec{\varepsilon}_1, \vec{\varepsilon}_2, \vec{\varepsilon}_3) = \text{rg}(\vec{e}_1, \vec{e}_2, -\vec{e}_2 + 3\vec{e}_3)$$

On ajoutant $\vec{e}_2$ au troisième vecteur

$$\text{rg}(\vec{\varepsilon}_1, \vec{\varepsilon}_2, \vec{\varepsilon}_3) = \text{rg}(\vec{e}_1, \vec{e}_2, 3\vec{e}_3)$$

Enfin, en multipliant par $1/3$ le troisième vecteur

$$\text{rg}(\vec{\varepsilon}_1, \vec{\varepsilon}_2, \vec{\varepsilon}_3) = \text{rg}(\vec{e}_1, \vec{e}_2, \vec{e}_3) = 3$$

car la famille $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ est une base de E.
On en déduit que la famille $(\vec{\varepsilon}_1, \vec{\varepsilon}_2, \vec{\varepsilon}_3)$ est aussi une base de E.

Remarque Avec le calcul matriciel, cet outil deviendra très efficace.

7.4 Applications linéaires en dimension finie

E, F et G désignent des $\mathbb{K}$-espaces vectoriels de dimensions finies.

7.4.1 Image d'une famille de vecteurs

Soient $f \in \mathcal{L}(E, F)$ et $\mathcal{B} = (\vec{e}_1, \ldots, \vec{e}_n)$ une famille de vecteurs de E.

Définition

On appelle image de la famille $\mathcal{B}$ par l'application f, la famille de vecteur

$$f(\mathcal{B}) = (f(\vec{e}_1), \ldots, f(\vec{e}_n))$$

Théorème

$f(\text{Vect}(\mathcal{B})) = \text{Vect}(f(\mathcal{B}))$ i.e. : $f\left(\text{Vect}(\vec{e}_1, \ldots, \vec{e}_n)\right) = \text{Vect}\left(f(\vec{e}_1), \ldots, f(\vec{e}_n)\right)$.

dém. :

$$\text{Vect}(\vec{e}_1, \ldots, \vec{e}_n) = \{\lambda_1\vec{e}_1 + \cdots + \lambda_n\vec{e}_n / \lambda_1, \ldots, \lambda_n \in \mathbb{K}\}$$

donc

$$f\left(\text{Vect}(\vec{e}_1,\ldots,\vec{e}_n)\right) = \{f(\lambda_1\vec{e}_1 + \cdots + \lambda_n\vec{e}_n)/\lambda_1,\ldots,\lambda_n \in \mathbb{K}\}$$

Par linéarité de f

$$f\left(\text{Vect}(\vec{e}_1,\ldots,\vec{e}_n)\right) = \{\lambda_1 f(\vec{e}_1) + \cdots + \lambda_n f(\vec{e}_n)/\lambda_1,\ldots,\lambda_n \in \mathbb{K}\}$$

et donc

$$f\left(\text{Vect}(\vec{e}_1,\ldots,\vec{e}_n)\right) = \text{Vect}\left(f(\vec{e}_1),\ldots,f(\vec{e}_n)\right)$$

□

Corollaire

Si V est un sous-espace vectoriel de E alors $\dim f(V) \leqslant \dim V$

dém. :
Soit $(\vec{e}_1,\ldots,\vec{e}_n)$ une base de V.

$$f(V) = f\left(\text{Vect}(\vec{e}_1,\ldots,\vec{e}_n)\right) = \text{Vect}\left(f(\vec{e}_1),\ldots,f(\vec{e}_n)\right)$$

donc $\dim f(V) \leqslant \dim V$.
□

Théorème

Si la famille $\mathcal{B}$ est libre et si $f \in \mathcal{L}(E,F)$ est injective alors la famille $f(\mathcal{B})$ est libre.

dém. :
Supposons

$$\lambda_1 f(\vec{e}_1) + \cdots + \lambda_n f(\vec{e}_n) = \vec{0}_F$$

Par linéarité de f on a

$$f\left(\lambda_1\vec{e}_1 + \cdots + \lambda_n\vec{e}_n\right) = \vec{0}_E$$

donc

$$\lambda_1\vec{e}_1 + \cdots + \lambda_n\vec{e}_n \in \ker f$$

Or $\ker f = \left\{\vec{0}_E\right\}$ donc

$$\lambda_1\vec{e}_1 + \cdots + \lambda_n\vec{e}_n = \vec{0}_E$$

Enfin puisque la famille $\mathcal{B}$ est libre, on obtient $\lambda_1 = \ldots = \lambda_n = 0$.
□

Corollaire

Si $f \in \mathcal{L}(E,F)$ est injective alors $\dim E \leqslant \dim F$.

dém. :
Une base de E est transformée par l'application f en une famille libre de F. Il y a donc moins de vecteurs dans une base de E que la dimension de F.
□

Remarque Si V est un sous-espace vectoriel de E et si f est injective alors $\dim f(V) = \dim V$.

Théorème

Si la famille $\mathcal{B}$ est génératrice de E et si $f \in \mathcal{L}(E,F)$ est surjective alors la famille $f(\mathcal{B})$ est génératrice de F.

dém. :
Soit $\vec{y} \in F$.
Puisque f est surjective, il existe $\vec{x} \in E$ vérifiant $\vec{y} = f(\vec{x})$.
Puisque la famille $\mathcal{B}$ est génératrice, il existe $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ vérifiant

$$\vec{x} = \lambda_1 \vec{e}_1 + \cdots + \lambda_n \vec{e}_n$$

et alors

$$f(\vec{x}) = \lambda_1 f(\vec{e}_1) + \cdots + \lambda_n f(\vec{e}_n)$$

Ainsi la famille $f(\mathcal{B})$ est génératrice de F.
□

Corollaire

Si $f \in \mathcal{L}(E, F)$ est surjective alors $\dim F \leqslant \dim E$.
dém. :
Une base de E est transformée par l'application f en une famille génératrice de F. Il y a donc plus de vecteurs dans une base de E que la dimension de F.

Théorème

Si la famille $\mathcal{B}$ est une base de E et si f est un isomorphisme alors $f(\mathcal{B})$ est une base de F.

dém. :
C'est immédiat par les deux théorèmes qui précèdent.
□

Corollaire

Si $f \in \mathcal{L}(E, F)$ est un isomorphisme alors $\dim E = \dim F$.

7.4.2 Image d'une base

Théorème

Soient E et F deux $\mathbb{K}$-espaces vectoriels.
Etant données $\mathcal{B} = (\vec{e}_1, \ldots, \vec{e}_p)$ une base de E et $\mathcal{F} = (\vec{y}_1, \ldots, \vec{y}_p)$ une famille de vecteurs de F, il existe une unique application linéaire $f \in \mathcal{L}(E, F)$ vérifiant

$$\forall 1 \leqslant j \leqslant p, f(\vec{e}_j) = \vec{y}_j$$

dém. :
Unicité :
Soit f solution. Pour tout $\vec{x} \in E$, on peut écrire $\vec{x} = \lambda_1 \vec{e}_1 + \cdots + \lambda_p \vec{e}_p$ avec $\lambda_j \in \mathbb{K}$
On a alors

$$f(\vec{x}) = \lambda_1 f(\vec{e}_1) + \cdots + \lambda_p f(\vec{e}_p) = \lambda_1 \vec{y}_1 + \cdots + \lambda_p \vec{y}_p$$

ce qui détermine $f(\vec{x})$ de manière unique.
Existence :
Pour tout $\vec{x} \in E$, posons $f(\vec{x}) = \lambda_1 \vec{y}_1 + \cdots + \lambda_p \vec{y}_p$ avec $\lambda_1, \ldots, \lambda_p$ les composantes de x dans la base $\mathcal{B}$. On définit ainsi une application $f : E \to F$.
Puisque les applications composantes dans une base sont linéaires, l'application f est linéaire et on vérifie aisément que $f(\vec{e}_j) = \vec{y}_j$ car les composantes de $\vec{e}_j$ dans $\mathcal{B}$ sont nulles sauf la j-ème égale à 1.
□

Corollaire

Une application linéaire est entièrement déterminée par l'image d'une base.

Exemple On vérifie aisément que les vecteurs $(1,1,1)$, $(0,1,1)$ et $(0,0,1)$ forment une base de $\mathbb{R}^3$.
Déterminons l'unique $f \in \mathcal{L}(\mathbb{R}^3,\mathbb{R}^2)$ vérifiant

$$f(1,1,1)=(1,0),\ f(0,1,1)=(0,1) \text{ et } f(0,0,1)=(1,1)$$

Pour $(x,y,z)\in\mathbb{R}^3$ déterminer $\alpha,\beta,\gamma\in\mathbb{R}$ tel que

$$(x,y,z)=\alpha.(1,1,1)+\beta(0,1,1)+\gamma(0,0,1)$$

Après résolution, on obtient $\alpha=x$, $\beta=y-x$ et $\gamma=z-y-x$.
On a alors

$$f(x,y,z)=xf(1,1,1)+(y-x)f(0,1,1)+(z-y-x)f(0,0,1)$$

donc

$$f(x,y,z)=(z-y,z-2x)$$

Théorème

Soient E et F deux $\mathbb{K}$-espaces vectoriels, $f\in\mathcal{L}(E,F)$ et $\mathcal{B}=(\vec{e}_1,\dots,\vec{e}_p)$ une base de E.
1) f est injective si, et seulement si, la famille $(f(\vec{e}_1),\dots,f(\vec{e}_p))$ est libre.
2) f est surjective si, et seulement si, la famille $(f(\vec{e}_1),\dots,f(\vec{e}_p))$ est génératrice de F.
3) f est un isomorphisme si, et seulement si, la famille $f(\mathcal{B})$ est une base de F.

dém. :
1) ($\Rightarrow$) déjà vue
($\Leftarrow$) Supposons la famille $(f(\vec{e}_1),\dots,f(\vec{e}_p))$ libre
Soit $\vec{x}\in\ker f$. On peut écrire $\vec{x}=\lambda_1\vec{e}_1+\dots+\lambda_p\vec{e}_p$.
On a alors

$$f(\vec{x})=\lambda_1 f(\vec{e}_1)+\dots+\lambda_p f(\vec{e}_p)=\vec{0}$$

Or la famille $(f(\vec{e}_1),\dots,f(\vec{e}_p))$ est libre donc $\lambda_1=\ldots=\lambda_p=0$ puis $\vec{x}=\vec{0}$.
Ainsi $\ker f=\left\{\vec{0}\right\}$ est donc f est injective.
2) ($\Rightarrow$) déjà vue
($\Leftarrow$) Supposons la famille $(f(\vec{e}_1),\dots,f(\vec{e}_p))$ génératrice de F.
Soit $\vec{y}\in F$. On peut écrire

$$\vec{y}=\lambda_1 f(\vec{e}_1)+\dots+\lambda_p f(\vec{e}_p)$$

Posons alors $\vec{x}=\lambda_1\vec{e}_1+\dots+\lambda_p\vec{e}_p\in E$. Par linéarité de f, on vérifie $\vec{y}=f(\vec{x})$.
Par suite f est surjective.
3) immédiat par ce qui précède.
□

Corollaire

Deux $\mathbb{K}$-espaces vectoriels de dimensions finies sont isomorphes si, et seulement si, ils ont même dimension

dém. :
Si deux espaces sont isomorphes, un isomorphisme transforme une base de l'un en une base de l'autre et donc les espaces ont même dimension.
Inversement, s'il y a égalité des dimensions, il existe une application linéaire transformant une base de l'un en une base de l'autre et celle-ci est un isomorphisme.
□

7.4.3 Rang d'une application linéaire

Définition

On appelle rang d'une application linéaire $f \in \mathcal{L}(E, F)$ la dimension de son image.
On note $\text{rg} f = \dim \text{Im} f$.

Exemple $\text{rg}(f) = 0 \Leftrightarrow f = \tilde{0}$.

Proposition

Si $\mathcal{B} = (\vec{e}_1, \dots, \vec{e}_p)$ est une base de E alors

$$\text{rg} f = \text{rg}(f(\vec{e}_1), \dots, f(\vec{e}_p))$$

dém. :

$$\text{rg} f = \dim \text{Im} f = \dim f(E)$$

or

$$f(E) = f(\text{Vect}(\vec{e}_1, \dots, \vec{e}_n)) = \text{Vect}(f(\vec{e}_1), \dots, f(\vec{e}_n))$$

donc

$$\text{rg} f = \text{rg}(f(\vec{e}_1), \dots, f(\vec{e}_p))$$

□

Proposition

$\text{rg} f \leqslant \dim E$ avec égalité si, et seulement si, f est injective.
$\text{rg} f \leqslant \dim F$ avec égalité si, et seulement si, f est surjective.

dém. :
Introduisons $(\vec{e}_1, \dots, \vec{e}_p)$ une base de E avec $p = \dim E$
$\text{rg} f = \text{rg}(f(\vec{e}_1), \dots, f(\vec{e}_p)) \leqslant p$ avec égalité si, et seulement si, la famille $(f(\vec{e}_1), \dots, f(\vec{e}_p))$ est libre i.e. f injective.
Aussi $\text{rg} f = \text{rg}(f(\vec{e}_1), \dots, f(\vec{e}_p)) \leqslant \dim F$ avec égalité si, et seulement si, la famille $(f(\vec{e}_1), \dots, f(\vec{e}_p))$ est génératrice de F i.e. f surjective.
□

Théorème

Soient $f \in \mathcal{L}(E, F)$ et $g \in \mathcal{L}(F, G)$.
On a

$$\text{rg}(g \circ f) \leqslant \min(\text{rg} f, \text{rg} g)$$

De plus, si f surjective $\text{rg}(g \circ f) = \text{rg} g$ et si g injective $\text{rg}(g \circ f) = \text{rg} f$.

dém. :

$$\text{rg}(g \circ f) = \dim \text{Im}(g \circ f) = \dim g(f(E))$$

D'une part, $g(f(E)) = \text{Im} g_{\restriction f(E)}$ donc $\text{rg}(g \circ f) = \text{rg} g_{\restriction f(E)} \leqslant \dim f(E) = \text{rg} f$ avec égalité quand g est injective.
D'autre part, $g(f(E)) \subset g(F) = \text{Im} g$ donc $\text{rg}(g \circ f) \leqslant \text{rg} g$ avec égalité quand f est surjective.
□

Corollaire

On ne modifie pas le rang d'une application linéaire en composant celle-ci avec un isomorphisme.

7.4.4 Théorème du rang

Théorème

Soit une application linéaire $f \in \mathcal{L}(E, F)$.
1) f réalise un isomorphisme de tout supplémentaire de $\ker f$ dans E vers Imf.
2) On a rg$f + \dim \ker f = \dim E$ [Formule du rang]

dém. :
1) Soit H un supplémentaire de $\ker f$ dans E.
Considérons la restriction $\tilde{f} : H \to \mathrm{Im} f$ définie par $\tilde{f}(\vec{x}) = f(\vec{x})$.
L'application $\tilde{f}$ est bien définie au départ de H et à valeurs dans Imf.
L'application $\tilde{f}$ est linéaire car restriction d'une application linéaire au départ d'un sous-espace vectoriel.
$\ker \tilde{f} = \ker f \cap H = \left\{\vec{0}\right\}$ donc l'application linéaire $\tilde{f}$ est injective.
Soit $\vec{y} \in \mathrm{Im} f$. Il existe $\vec{x} \in E$ vérifiant $\vec{y} = f(\vec{x})$. Or on peut écrire $\vec{x} = \vec{a} + \vec{b}$ avec $\vec{a} \in \ker f$ et $\vec{b} \in H$ et alors $\vec{y} = f(\vec{x}) = f(\vec{a}) + f(\vec{b}) = \tilde{f}(\vec{b}) \in \mathrm{Im}\tilde{f}$. Par suite l'application $\tilde{f}$ est surjective.
Finalement $\tilde{f}$ est un isomorphisme.
2) Puisque $\tilde{f}$ est un isomorphisme on a $\dim H = \dim \mathrm{Im} f = \mathrm{rg} f$.
Or $\dim \ker f + \dim H = \dim E$ car $\ker f$ et H sont supplémentaires dans E donc

$$\dim \ker f + \mathrm{rg} f = \dim E$$

□

Exemple Soit $f : \mathbb{R}^3 \to \mathbb{R}^3$ définie par

$$f : (x, y, z) \mapsto (x + 2y + z, y - x, 2x + y + z)$$

On vérifie aisément que l'application f est linéaire. Déterminons une base de Imf et $\ker f$.
Commençons par étudier $\ker f$.

$$(x, y, z) \in \ker f \Leftrightarrow f(x, y, z) = 0$$

Or

$$\begin{cases} x + 2y + z = 0 \\ y - x = 0 \\ 2x + y + z = 0 \end{cases} \Leftrightarrow \begin{cases} x = y \\ z = -3y \end{cases}$$

donc

$$\ker f = \{(y, y, -3y)/y \in \mathbb{R}\} = \mathrm{Vect}\vec{u}$$

avec $\vec{u} = (1, 1, -3)$.
La famille $(\vec{u})$ est base de $\ker f$.
Etudions maintenant Imf.
Par la formule du rang, on a rg$f = \dim \mathbb{R}^3 - \dim \ker f = 2$.
Pour former une base de Imf, il suffit de déterminer deux vecteurs indépendants dans l'image de f.
$\vec{v} = f(1, 0, 0) = (1, -1, 2)$ et $\vec{w} = f(0, 1, 0) = (2, 1, 1)$ conviennent et donc la famille $(\vec{v}, \vec{w})$ est base de l'image de f.

Exemple Soient E un $\mathbb{K}$-espace vectoriel de dimension finie et $f, g \in \mathcal{L}(E)$ vérifiant $f \circ g = \tilde{0}$.
Montrons

$$\text{rg} f + \text{rg} g \leqslant \dim E$$

Puisque $f \circ g = \tilde{0}$ on a $\text{Im} g \subset \ker f$ puis $\text{rg} g \leqslant \dim \ker f$.
Or par la formule du rang $\dim \ker f = \dim E - \text{rg} f$ donc $\text{rg} g \leqslant \dim E - \text{rg} f$.

Exemple Soient E un $\mathbb{K}$-espace vectoriel de dimension finie et $f \in \mathcal{L}(E)$ vérifiant $\ker f = \ker f^2$.
Montrons que Imf et $\ker f$ sont supplémentaires dans E.
Par la formule du rang on a déjà $\dim \text{Im} f + \dim \ker f = \dim E$. Il suffit alors d'observer
$\text{Im} f \cap \ker f = \left\{\vec{0}\right\}$ pour conclure.
Soit $\vec{x} \in \text{Im} f \cap \ker f$. Il existe $\vec{a} \in E$ tel que $\vec{x} = f(\vec{a})$. Puisque $f(\vec{x}) = \vec{0}$, on a $f^2(\vec{a}) = \vec{0}$ i.e.
$\vec{a} \in \ker f^2$. Or par hypothèse $\ker f = \ker f^2$ donc $a \in \ker f$ puis $\vec{x} = \vec{0}$.
Ainsi $\text{Im} f \cap \ker f = \left\{\vec{0}\right\}$ et donc les espaces $\text{Im} f$ et $\ker f$ sont supplémentaires dans E.

7.4.5 Théorème d'isomorphisme

Théorème

Soient E et F deux $\mathbb{K}$-espaces vectoriels de dimensions finies tels que $\dim E = \dim F = n$.
Pour $f \in \mathcal{L}(E, F)$, on a équivalence entre :
(i) f est un isomorphisme ;
(ii) f est injective ;
(iii) f est surjective ;
(iv) $\text{rg} f = n$;
(v) $\exists g \in \mathcal{L}(F, E), g \circ f = \text{Id}_E$;
(vi) $\exists h \in \mathcal{L}(F, E), f \circ h = \text{Id}_F$;
De plus, si tel est le cas,

$$g = h = f^{-1}$$

dém. :
(i) $\Rightarrow$ (ii) ok
(ii) $\Rightarrow$ (iii) Supposons f injective.
Par la formule du rang $\text{rg} f = \dim E - \dim \ker f = \dim E = \dim F$ donc f est surjective.
(iii) $\Rightarrow$ (iv) Supposons f surjective.
On a $\text{rg} f = \dim F = n$.
(iv) $\Rightarrow$ (i) Supposons $\text{rg} f = n$. On a $\text{rg} f = \dim E$ donc f est injective. On a aussi $\text{rg} f = \dim F$ donc f est surjective. Ainsi f est un isomorphisme.
(i) $\Rightarrow$ (v) et (vi) ok
(v) $\Rightarrow$ (ii) Supposons qu'il existe $g \in \mathcal{L}(F, E)$ vérifiant $g \circ f = \text{Id}_E$.
Puisque l'application $g \circ f$ injective alors f est injective.
(vi) $\Rightarrow$ (iii) Supposons qu'il existe $h \in \mathcal{L}(F, E)$ vérifiant $f \circ h = \text{Id}_F$.
Puisque l'application $f \circ h$ est surjective alors f est surjective.
Enfin si $g \circ f = \text{Id}_E$, sachant f bijective, on a $g = g \circ (f \circ f^{-1}) = (g \circ f) \circ f^{-1} = f^{-1}$ et de même si $f \circ h = \text{Id}_F$ alors $h = f^{-1}$.
□

Corollaire

Si E est un $\mathbb{K}$-espace vectoriel de dimension finie, le théorème ci-dessus caractérise les automorphismes de E parmi les endomorphismes de E.

Attention : Ces équivalences ne valent qu'en dimension finie.

Exemple Soit $f : \mathbb{R}^3 \to \mathbb{R}^3$ définie par

$$f : (x, y, z) \mapsto (y + z, z + x, x + y)$$

Montrons que $f \in \mathrm{GL}(\mathbb{R}^3)$.
On vérifie aisément que f est un endomorphisme de $\mathbb{R}^3$.
Puisque $\dim \mathbb{R}^3 < +\infty$, il suffit de vérifier l'injectivité pour affirmer que f est un automorphisme.
Soit $(x, y, z) \in \ker f$.

$$\begin{cases} y + z = 0 \\ z + x = 0 \\ x + y = 0 \end{cases} \Leftrightarrow \begin{cases} y = -z \\ x = -z \\ -2z = 0 \end{cases} \Leftrightarrow \begin{cases} x = 0 \\ y = 0 \\ z = 0 \end{cases}$$

Ainsi $\ker f = \left\{\vec{0}\right\}$ donc f est un automorphisme de $\mathbb{R}^3$.

7.4.6 Hyperplan

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n non nulle.

Définition

On appelle hyperplan de E tout sous-espace vectoriel H de dimension $n - 1$.

Exemple Dans $E = \mathbb{R}^3$, considérons $P = \mathrm{Vect}(\vec{u}, \vec{v})$ avec $\vec{u}, \vec{v}$ non colinéaires.
On a $\dim P = 2 = \dim E - 1$ donc P est un hyperplan de E.

Proposition

Le noyau d'une forme linéaire non nulle est un hyperplan.

dém. :
Soit φ une forme linéaire non nulle sur E, i.e. une forme linéaire différente de l'application nulle.
On a $\mathrm{Im}\varphi \subset \mathbb{K}$ donc $\mathrm{rg}\varphi = 0$ ou $\mathrm{rg}\varphi = 1$.
Or la forme linéaire φ est non nulle donc $\mathrm{rg}\varphi = 1$ puis par la formule du rang $\dim \ker \varphi = \dim E - \mathrm{rg}\varphi = n - 1$.
□

Exemple Dans $E = \mathbb{K}^n$, considérons

$$H = \{(x_1, \dots, x_n) \in \mathbb{K}^n / x_1 + \dots + x_n = 0\}$$

H est un hyperplan car noyau de la forme linéaire non nulle $(x_1, \dots, x_n) \in \mathbb{K}^n \mapsto x_1 + \dots + x_n$.

Proposition

Tout hyperplan peut se voir comme noyau d'une forme linéaire non nulle.

dém. :
Soient H un hyperplan et $(\vec{e}_1, \ldots, \vec{e}_{n-1})$ une base de H que l'on complète en une base de E

$$\mathcal{B} = (\vec{e}_1, \ldots, \vec{e}_{n-1}, \vec{e}_n)$$

Considérons φ l'application donnant la n-ième composante d'un vecteur dans $\mathcal{B}$.
L'application φ est une forme linéaire non nulle sur E et par construction $H \subset \ker \varphi$.
Or H est $\ker \varphi$ sont des hyperplans donc par inclusion et égalité des dimensions, on a $H = \ker \varphi$.
f est une forme linéaire non nulle et $H \subset \ker f$ donc $H = \ker f$.
On suppose E muni d'une base $\mathcal{B} = (\vec{e}_1, ..., \vec{e}_n)$.
Pour tout $\vec{x} \in E$, notons $x_1, ..., x_n$ ses composantes dans $\mathcal{B}$.
□

Théorème

Les hyperplans de E sont les ensembles H constitués des vecteurs $\vec{x} \in E$ dont les composantes $x_1, \ldots, x_n$ dans une base $\mathcal{B}$ de E sont solutions d'une équation de la forme

$$a_1 x_1 + \cdots + a_n x_n = 0$$

avec $a_1, \ldots, a_n$ scalaires non tous nuls.

dém. :
Notons $\vec{e}_1, \ldots, \vec{e}_n$ les vecteurs de la base $\mathcal{B}$ de E introduite dans l'énoncé et notons $x_1, \ldots, x_n$ les composantes d'un vecteurs $\vec{x}$ générique de E.
Soient $(a_1, \ldots, a_n) \neq (0, \ldots, 0)$ et H l'ensemble des vecteurs $\vec{x} \in E$ vérifiant l'équation

$$a_1 x_1 + \cdots + a_n x_n = 0$$

Considérons alors l'application

$$\varphi : \vec{x} \mapsto a_1 x_1 + \cdots + a_n x_n$$

Par linéarité des fonctions composantes dans une base, on peut affirmer que φ est une forme linéaire sur E et puisque $(a_1, \ldots, a_n) \neq (0, \ldots, 0)$, cette forme linéaire est non nulle.
Par définition de H, $H = \ker \varphi$ et donc H est un hyperplan.
Inversement, soit H un hyperplan de E. Il existe une forme linéaire non nulle φ sur E telle que $H = \ker \varphi$. Or pour $\vec{x} = x_1 \vec{e}_1 + \cdots + x_n \vec{e}_n$, on a

$$\varphi(\vec{x}) = x_1 \varphi(\vec{e}_1) + \cdots + x_n \varphi(\vec{e}_n) = a_1 x_1 + \cdots + a_n x_n$$

avec $a_i = \varphi(\vec{e}_i) \in \mathbb{K}$.
On a alors

$$\vec{x} \in H \Leftrightarrow \varphi(\vec{x}) = 0 \Leftrightarrow a_1 x_1 + \cdots + a_n x_n = 0$$

Enfin, les $a_1, \ldots, a_n$ ne sont pas tous nuls car la forme linéaire φ est non nulle.
□

Définition

Avec les notations qui précèdent, l'équation

$$a_1 x_1 + \cdots + a_n x_n = 0$$

est alors appelée équation de l'hyperplan H relative à la base $\mathcal{B}$.

Remarque Si $a_1x_1 + \cdots + a_nx_n = 0$ est une équation de H dans $\mathcal{B}$ alors pour tout $\lambda \neq 0$, l'équation $(\lambda a_1)x_1 + \cdots + (\lambda a_n)x_n = 0$ définit aussi H dans la base $\mathcal{B}$.
Inversement, on peut montrer qu'une équation définissant H dans $\mathcal{B}$ est nécessairement de la forme précédente.

Exemple Soit E une $\mathbb{R}$-espace vectoriel muni d'une base $\mathcal{B} = (\vec{i}, \vec{j}, \vec{k})$.
On convient de noter x, y, z les composantes des vecteurs de E.
Soient $\vec{u}(1, 0, 1)$ et $\vec{v}(1, 2, -1)$ et $H = \text{Vect}(\vec{u}, \vec{v})$.
Puisque les vecteurs $\vec{u}$ et $\vec{v}$ ne sont pas colinéaires, on a $\dim H = 2$ et donc H est un hyperplan de E.
Déterminons une équation de H.
Celle-ci est de la forme : $ax + by + cz = 0$ avec $(a, b, c) \neq (0, 0, 0)$.
Puisque $\vec{u}, \vec{v} \in H$ on a

$$\begin{cases} a + c = 0 \\ a + 2b - c = 0 \end{cases}$$

On en déduit

$$\begin{cases} a = -c \\ b = c \end{cases}$$

En prenant $c = 1$, l'équation $-x + y + z = 0$ est une équation d'un hyperplan contenant les vecteur $\vec{u}$ et $\vec{v}$, c'est une hyperplan ne peut être que H.
Finalement $-x + y + z = 0$ est une équation définissant H.

Chapitre 8

Polynômes en une indéterminée

Dans ce chapitre on étudie la manipulation des fonctions polynomiales dans un cadre plus abstrait. Les propriétés obtenues s'appliqueront aux fonctions polynomiales mais plus généralement à tout type d'expression polynomiale
$\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$.

8.1 Construction de l'anneau des polynômes

8.1.1 Polynômes

Définition

On appelle polynôme à coefficients dans $\mathbb{K}$ en l'indéterminée X tout objet noté

$$P = \sum_{n=0}^{+\infty} a_n X^n = a_0 + a_1 X + ... + a_n X^n + ...$$

où $(a_n)_{n\in\mathbb{N}}$ est une suite d'éléments de $\mathbb{K}$ nulle à partir d'un certain rang, appelée suite des coefficients de P.
On note $\mathbb{K}[X]$ l'ensemble de ces éléments.

Remarque Puisqu'il existe $p \in \mathbb{N}$ tel que

$$\forall n > p, a_n = 0$$

la somme $\sum\limits_{n=0}^{+\infty} a_n X^n$ ne contient qu'un nombre fini de termes non nuls. On peut aussi l'écrire

$$P = a_0 + a_1 X + a_2 X^2 + \cdots + a_p X^p$$

Exemple $2 + X - X^2$ est un polynôme de $\mathbb{R}[X]$ dont la suite des coefficients est : $2, 1, -1, 0, 0, \ldots$

Exemple $X^3 + X - 1$ est un polynôme de $\mathbb{R}[X]$ dont la suite des coefficients est : $-1, 1, 0, 1, 0, 0, \ldots$

Remarque L'indéterminée X n'est pas une variable, c'est une lettre qui permet de désigner les coefficients respectifs d'un polynôme.

Définition

Deux polynômes $P = \sum_{n=0}^{+\infty} a_n X^n$ et $Q = \sum_{n=0}^{+\infty} b_n X^n \in \mathbb{K}[X]$ sont dits égaux s'ils ont les mêmes coefficients.
Ainsi :

$$P = Q \Leftrightarrow \forall n \in \mathbb{N}, a_n = b_n$$

Définition

On appelle polynôme constant égal à $C \in \mathbb{K}$ le polynôme $C + 0.X + \cdots = C$.

Exemple On appelle polynôme nul le polynôme constant égal à 0.

Définition

On appelle monôme, tout polynôme de la forme : $0 + 0.X + \cdots + 0.X^{n-1} + aX^n + 0.X^{n+1} + \cdots = aX^n$
avec $a \in \mathbb{K}$, $n \in \mathbb{N}$.

Exemple Les polynômes constants sont des monômes.

Définition

On dit qu'un polynôme $P = \sum_{n=0}^{+\infty} a_n X^n \in \mathbb{K}[X]$ est pair (resp. impair) si

$$\forall p \in \mathbb{N}, a_{2p+1} = 0 \text{ (respectivement } a_{2p} = 0 \text{)}$$

Exemple Le polynôme nul est le seul polynôme à la fois pair et impair.

8.1.2 L'espace vectoriel des polynômes

Définition

Soient $P = \sum_{n=0}^{+\infty} a_n X^n \in \mathbb{K}[X]$ et $Q = \sum_{n=0}^{+\infty} b_n X^n \in \mathbb{K}[X]$
On définit le polynôme $P + Q \in \mathbb{K}[X]$ par

$$P + Q = \sum_{n=0}^{+\infty} (a_n + b_n) X^n$$

Remarque $P+Q$ est bien un polynôme car il existe $p,q\in\mathbb{N}$ tels que

$$\forall n>p, a_n=0 \text{ et } \forall n>q, b_n=0$$

Par suite $r=\max(p,q)$ on a

$$\forall n>r, a_n+b_n=0$$

Exemple $(2+X-X^2)+(X^3+X-1)=1+2X-X^2+X^3$.

Définition

Soient $P=\sum_{n=0}^{+\infty}a_nX^n\in\mathbb{K}[X]$ et $\lambda\in\mathbb{K}$. On définit le polynôme $\lambda.P\in\mathbb{K}[X]$ par :

$$\lambda.P=\sum_{n=0}^{+\infty}(\lambda.a_n)X^n$$

Remarque $\lambda.P$ est bien un polynôme car la suite des ses coefficients est nulle à partir d'un certain rang.

Exemple $2.(X^2-1)=2X^2-2$.

Théorème

$(\mathbb{K}[X],+,.)$ est un $\mathbb{K}$ -espace vectoriel dont l'élément nul est le polynôme nul.

dém. :
L'addition est une loi de composition interne sur $\mathbb{K}[X]$.
On vérifie aisément qu'elle est commutative et qu'elle est associative.
Pour tout $P=\sum_{n=0}^{+\infty}a_nX^n\in\mathbb{K}[X]$, on a

$$P+0=\sum_{n=0}^{+\infty}(a_n+0)X^n=\sum_{n=0}^{+\infty}a_nX^n=P$$

et pour $Q=\sum_{n=0}^{+\infty}-a_nX^n\in\mathbb{K}[X]$, on a

$$P+Q=\sum_{n=0}^{+\infty}(a_n-a_n)X^n=\sum_{n=0}^{+\infty}0.X^n=0$$

Ainsi $(\mathbb{K}[X],+)$ est un groupe abélien d'élément neutre le polynôme nul.
On vérifie ensuite aisément les propriétés calculatoires suivantes : $\forall\lambda,\mu\in\mathbb{K},\forall P,Q\in\mathbb{K}[X]$,

$$(\lambda+\mu).P=\lambda.P+\mu.P,\ \lambda.(P+Q)=\lambda.P+\lambda.Q$$

$$\lambda.(\mu.P)=(\lambda\mu).P \text{ et } 1.P=P$$

□

8.1.3 Degré

Définition

Soit $P = \sum_{n=0}^{+\infty} a_n X^n \in \mathbb{K}[X]$ tel que $P \neq 0$.
On appelle degré de P le plus grand $n \in \mathbb{N}$ tel que $a_n \neq 0$, on le note $n = \deg P$.

Remarque On peut alors écrire $P = a_0 + a_1 X + ... + a_n X^n$ avec $a_n \neq 0$.

Définition

Le coefficient a_n est alors appelé coefficient dominant de P.

Convention Si $P = 0$, on pose $\deg P = -\infty$.

Exemple $\deg(X^3 + X + 2) = 3$

Exemple $\deg(aX + b) = \begin{cases} 1 & \text{si } a \neq 0 \\ 0 & \text{si } a = 0, b \neq 0 \\ -\infty & \text{sinon} \end{cases}$.

Exemple $\forall n \in \mathbb{N}, \deg(X^n) = n$.

Exemple Les polynômes de degré 0 sont les polynômes constants non nuls.

Proposition

$\forall P \in \mathbb{K}[X], \forall \lambda \in \mathbb{K}$,

$$\deg(\lambda.P) = \begin{cases} \deg P \text{ si } \lambda \neq 0 \\ -\infty \text{ si } \lambda = 0 \end{cases}$$

dém. :
Si $\lambda = 0$ ou $P = 0$ alors $\lambda.P = 0$ et donc $\deg(\lambda.P) = -\infty$.
Si $\lambda \neq 0$ et $P \neq 0$ alors en posant $n = \deg P$, on peut écrire $P = a_0 + a_1 X + ... + a_n X^n$ avec $a_n \neq 0$ et alors $\lambda.P = \lambda a_0 + \lambda a_1 X + \cdots + \lambda a_n X^n$ avec $\lambda a_n \neq 0$ donc $\deg(\lambda.P) = n$.
□

Proposition

$$\forall P, Q \in \mathbb{K}[X], \deg(P + Q) \leqslant \max(\deg P, \deg Q)$$

avec égalité lorsque $\deg(P) \neq \deg(Q)$.

dém. :
Si $P = 0$ ou $Q = 0$: c'est immédiat

Sinon, posons $p = \deg P$ et $q = \deg Q$.
On peut écrire $P = \sum_{n=0}^{+\infty} a_n X^n$ avec $\forall n > p, a_n = 0$ et $Q = \sum_{n=0}^{+\infty} b_n X^n$ avec $\forall n > q, b_n = 0$.
On a alors $P + Q = \sum_{n=0}^{+\infty} (a_n + b_n) X^n$ et pour tout $n > \max(p, q)$, $a_n + b_n = 0$ donc $\deg(P + Q) \leqslant \max(p, q)$.
De plus, si $p \neq q$, en notant $r = \max(p, q)$, $a_r + b_r \neq 0$ par somme d'un terme nul et d'un terme non nul.
On en déduit qu'alors $\deg(P + Q) = r$.
□

Remarque Si $\deg P = \deg Q$ et que les coefficients dominants de P et Q s'annulent alors

$$\deg(P + Q) < \max(\deg P, \deg Q)$$

8.1.4 Le sous-espace vectoriel $\mathbb{K}_n[X]$

Définition

Pour $n \in \mathbb{N}$, on note $\mathbb{K}_n[X]$ l'ensemble des polynômes de $\mathbb{K}[X]$ de degré inférieur à n.
Ainsi

$$\mathbb{K}_n[X] = \{P \in \mathbb{K}[X] \,/\, \deg P \leqslant n\}$$

Exemple $\mathbb{K}_0[X]$ est l'ensemble des polynômes constants.

Exemple $\mathbb{K}_1[X] = \{aX + b/a, b \in \mathbb{K}\}$

Théorème

$\mathbb{K}_n[X]$ est un sous-espace vectoriel de $\mathbb{K}[X]$ de dimension $n + 1$ dont la famille $\mathcal{B} = (1, X, ..., X^n)$ est une base, dite base canonique.

dém. :
$\mathbb{K}_n[X] =$ Vect$(1, X, ..., X^n)$ donc $\mathbb{K}_n[X]$ est un sous-espace vectoriel de $\mathbb{K}[X]$ et $\mathcal{B} = (1, X, ..., X^n)$ en est une famille génératrice.
Montrons que $\mathcal{B}$ est libre.
Supposons $\lambda_0.1 + \lambda_1 X + \cdots + \lambda_n.X^n = 0$.
Par unicité des coefficients décrivant un polynôme, $\lambda_0 = \ldots = \lambda_n = 0$.
Ainsi la famille $\mathcal{B}$ est libre et c'est donc une base de $\mathbb{K}_n[X]$.
□

Corollaire

$\dim \mathbb{K}[X] = +\infty$.

dém. :
En effet $\mathbb{K}[X]$ contient des sous-espaces vectoriels de dimension arbitrairement grange.
□

Exemple Soit $\mathcal{B} = (P_k)_{0 \leqslant k \leqslant n}$ une famille de polynômes de $\mathbb{K}[X]$ telle que $\forall 0 \leqslant k \leqslant n$, $\deg P_k = k$.
Montrons que $\mathcal{B}$ est une base de $\mathbb{K}_n[X]$.
La famille $\mathcal{B}$ est formée de $n+1 = \dim \mathbb{K}_n[X]$ vecteurs tous éléments de $\mathbb{K}_n[X]$.
Montrons que la famille $\mathcal{B}$ est libre pour conclure.
Supposons

$$\lambda_0 P_0 + \cdots + \lambda_n P_n = 0$$

On peut écrire

$$\lambda_n P_n = -(\lambda_0 P_0 + \cdots + \lambda_{n-1} P_{n-1})$$

Or $\deg P_0, \ldots, \deg P_{n-1} \leqslant n-1$ donc $\deg(\lambda_0 P_0 + \cdots + \lambda_{n-1} P_{n-1}) \leqslant n-1$.
Cependant, le degré de $\lambda_n P_n$ vaut n ou $-\infty$ selon la nullité de λ_n, on peut donc conclure que $\lambda_n = 0$.
La relation initiale se réécrit $\lambda_0 P_0 + \cdots + \lambda_{n-1} P_{n-1} = 0$ et en reprenant l'idée précédente on obtient successivement $\lambda_{n-1} = 0, \ldots, \lambda_0 = 0$.

8.1.5 L'anneau des polynômes

On définit une multiplication sur $\mathbb{K}[X]$ en imitant le principe de multiplication des expressions polynomiales :

Définition

Pour $P = \sum_{n=0}^{+\infty} a_n X^n \in \mathbb{K}[X]$ et $Q = \sum_{n=0}^{+\infty} b_n X^n \in \mathbb{K}[X]$, on définit le polynôme PQ par :

$$PQ = \sum_{n=0}^{+\infty} c_n X^n \text{ avec } c_n = \sum_{k=0}^{n} a_k b_{n-k} = \sum_{k+\ell=n} a_k b_\ell = a_0 b_n + a_1 b_{n-1} + \cdots + a_n b_0$$

Remarque PQ est bien un polynôme car il existe $p, q \in \mathbb{N}$ tels que

$$\forall k > p, a_k = 0 \text{ et } \forall k > q, b_k = 0,$$

et alors pour $n > p+q$ on a

$$c_n = \sum_{k=0}^{n} a_k b_{n-k} = \sum_{k=0}^{p} a_k b_{n-k} + \sum_{k=p+1}^{n} a_k b_{n-k} = 0$$

Puisque les coefficients c_n sont nuls à partir d'un certain rang, on peut affirmer que PQ est un polynôme.

Exemple $(2 + X - X^2)(1 + X + X^3) = 2 + 3X + X^4 - X^5$.

Proposition

$\forall p, q \in \mathbb{N}$,

$$X^p \times X^q = X^{p+q}$$

dém. :
Introduisons le symbole de Kronecker :

$$\delta_{i,j} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{sinon} \end{cases}$$

On peut écrire $X^p = \sum_{n=0}^{+\infty} \delta_{n,p} X^n$ et $X^q = \sum_{n=0}^{+\infty} \delta_{n,q} X^n$.

On a alors $X^p \times X^q = \sum_{n=0}^{+\infty} c_n X^n$ avec

$c_n = \sum_{k=0}^{n} \delta_{k,p}\delta_{n-k,q}$, $c_n = \begin{cases} 1 & \text{si } n = p+q \\ 0 & \text{sinon} \end{cases} = \delta_{n,p+q}$ donc $X^p \times X^q = X^{p+q}$.

□

Théorème

$(\mathbb{K}[X], +, \times)$ est un anneau commutatif d'élément nul le polynôme nul et d'élément unité le polynôme constant égal à 1.

dém. :
On sait déjà que $(\mathbb{K}[X], +)$ est un groupe abélien de neutre le polynôme nul.

Soient $P = \sum_{n=0}^{+\infty} a_n X^n$ et $Q = \sum_{n=0}^{+\infty} b_n X^n$ des polynômes de $\mathbb{K}[X]$.

$PQ = \sum_{n=0}^{+\infty} c_n X^n$ avec $c_n = \sum_{k=0}^{n} a_k b_{n-k}$ et $QP = \sum_{n=0}^{+\infty} d_n X^n$ avec $d_n = \sum_{\ell=0}^{n} b_\ell a_{n-\ell}$.

Par le changement d'indice $\ell = n - k$, on obtient $c_n = d_n$ et on peut conclure $PQ = QP$.
Ainsi la loi $\times$ est commutative.
Notons que, de plus, on peut écrire le coefficient général de PQ

$c_n = \sum_{k+\ell=n} a_k b_\ell$

ce qui va nous être utile pour justifier maintenant l'associativité de la loi $\times$.

Soient $P = \sum_{n=0}^{+\infty} a_n X^n$, $Q = \sum_{n=0}^{+\infty} b_n X^n$ et $R = \sum_{n=0}^{+\infty} c_n X^n$ des polynômes de $\mathbb{K}[X]$.

$PQ = \sum_{n=0}^{+\infty} d_n X^n$ avec $d_n = \sum_{k+\ell=n} a_k b_\ell$ et $(PQ)R = \sum_{n=0}^{+\infty} e_n X^n$ avec

$$e_n = \sum_{k+\ell=n} d_k c_\ell = \sum_{k+\ell=n} \sum_{i+j=k} (a_i b_j) c_\ell = \sum_{i+j+\ell=n} a_i b_j c_\ell$$

Le calcul de $P(QR)$ conduit au même polynôme et donc $(PQ)R = P(QR)$.
La loi $\times$ est donc associative.

Soit $P = \sum_{n=0}^{+\infty} a_n X^n$ un polynôme de $\mathbb{K}[X]$ et $1 = \sum_{n=0}^{+\infty} \delta_{n,0} X^n$ le polynôme constant égal à 1.

$P \times 1 = \sum_{n=0}^{+\infty} b_n X^n$ avec $b_n = \sum_{k=0}^{n} a_k \delta_{n-k,0} = a_n$ donc $P \times 1 = P$ et on peut conclure que le polynôme constant égal à 1 est élément neutre pour la loi $\times$.
Enfin, la distributivité de $\times$ sur $+$ ne pose pas de difficultés.

□

Remarque On peut introduit la notion d'itéré d'un élément et on observe aisément que :
$X^n = X \times X \times \cdots \times X$ ce qui justifie a posteriori la notation X^n.

Remarque Puisque $\mathbb{K}[X]$ est un anneau commutatif, on peut y exploiter les formules :

$$(P+Q)^n = \sum_{k=0}^{n} \binom{n}{k} P^k Q^{n-k} \text{ et } P^n - Q^n = (P-Q)\sum_{k=0}^{n-1} P^k Q^{n-1-k}$$

8.1.6 Degré d'un produit

Théorème

$\forall P, Q \in \mathbb{K}[X]$,

$$\deg PQ = \deg P + \deg Q$$

dém. :
Si $P = 0$ ou $Q = 0$: c'est immédiat en convenant que $-\infty + n = n + (-\infty) = -\infty$.
Sinon, posons $p = \deg P$ et $q = \deg Q$. On peut écrire
$P = \sum_{n=0}^{+\infty} a_n X^n$ avec $a_p \neq 0$ et $\forall n > p, a_n = 0$
$Q = \sum_{n=0}^{+\infty} b_n X^n$ avec $b_q \neq 0$ et $\forall n > q, b_n = 0$.
On alors $PQ = \sum_{n=0}^{+\infty} c_n X^n$ avec $\forall n \in \mathbb{N}, c_n = \sum_{k=0}^{n} a_k b_{n-k}$.
Pour $n > p+q$, on peut affirmer que $c_n = 0$ en tant que somme de terme nul car pour tout $k \in \{0, \ldots, p\}$, $a_k b_{n-k} = a_k \times 0 = 0$ (puisque $n - k > q$) et pour tout $k \in \{p+1, \ldots, n\}$, $a_k b_{n-k} = 0 \times b_{n-k} = 0$ (puisque $k > p$)
Pour $n = p + q$, on a $c_{p+q} = \sum_{k=0}^{p+q} a_k b_{p+q-k} = a_p b_q \neq 0$ car pour $k \in \{0, \ldots, p-1\}$, $a_k b_{p+q-k} = a_k \times 0 = 0$ et pour $k \in \{p+1, \ldots, p+q\}$, $a_k b_{p+q-k} = 0 \times b_{p+q-k} = 0$.
Par suite le polynôme PQ est de degré exactement $p + q$.
□

Remarque Le coefficient dominant de PQ est le produit des coefficients dominants de P et Q.

Corollaire

Les polynômes inversibles sont les polynômes constants non nuls.

dém. :
Les polynômes constants non nuls sont inversibles car si $C \in \mathbb{K} \setminus \{0\}$ on a $C \times \frac{1}{C} = 1$.
Inversement, si P est un polynôme est inversible dans $\mathbb{K}[X]$ et si Q désigne son inverse, l'égalité $PQ = 1$ donne $\deg P + \deg Q = 0$ d'où $\deg P = \deg Q = 0$.
□

Corollaire

$$\forall P, Q \in \mathbb{K}[X], PQ = 0 \Rightarrow P = 0 \text{ ou } Q = 0$$

Ainsi, dans $\mathbb{K}[X]$, il n'y a pas de diviseurs de zéro.

dém. :
Si $P \neq 0$ et $Q \neq 0$ alors on a vu dans les calculs précédents que $PQ \neq 0$.
□

Remarque Dans $\mathbb{K}[X]$, pour $P \neq 0$

$$PQ = PR \Rightarrow Q = R$$

Autrement dit tout polynôme non nul est régulier.

8.1.7 Composition de polynômes

Définition

Soient $P = \sum_{n=0}^{+\infty} a_n X^n \in \mathbb{K}[X]$ et $Q \in \mathbb{K}[X]$.
On définit le polynôme composé $P \circ Q$ (aussi noté $P(Q)$) par :

$$P \circ Q = P(Q) = \sum_{n=0}^{+\infty} a_n Q^n \in \mathbb{K}[X]$$

Exemple Pour $Q = X + 1$, on a $P(X) = \sum_{n=0}^{+\infty} a_n (X+1)^n$.
Pour $Q = X^2$, on a $P(X^2) = \sum_{n=0}^{+\infty} a_n X^{2n}$.

Exemple Pour $Q = X$, on a $P(X) = \sum_{n=0}^{+\infty} a_n X^n = P$.

Remarque Un polynôme en l'indéterminé X peut indifféremment être noté P ou $P(X)$.

Exemple P est pair si, et seulement si, $P(-X) = P(X)$.

Proposition

$\forall P, Q, R \in \mathbb{K}[X]$ et $\forall \lambda, \mu \in \mathbb{K}$

$$(\lambda.P + \mu.Q) \circ R = \lambda.P \circ R + \mu.Q \circ R \text{ et } (PQ) \circ R = (P \circ R) \times (Q \circ R)$$

dém. :
Il suffit d'introduire les coefficients des polynômes P et Q pour décrire les polynômes étudiés et observer leur égalité.
□

Exemple Soit $\varphi : \mathbb{K}_n[X] \to \mathbb{K}_n[X]$ définie par

$$\varphi(P) = P(X+1) + P(X)$$

Montrer que φ est un automorphisme.
L'application φ est bien définie car pour $P \in \mathbb{K}_n[X]$, le polynôme $P(X+1) + P(X)$ est bien élément de $\mathbb{K}[X]$.
Pour $P, Q \in \mathbb{K}[X]$ et $\lambda, \mu \in \mathbb{K}$,

$$\varphi(\lambda P + \mu Q) = (\lambda P + \mu Q)(X) + (\lambda P + \mu Q)(X+1)$$

Par les propriétés calculatoires qui précèdent

$$\varphi(\lambda P + \mu Q) = \lambda P(X) + \lambda P(X+1) + \mu Q(X) + \mu Q(X+1)$$

et donc $\varphi(\lambda P + \mu Q) = \lambda\varphi(P) + \mu\varphi(Q)$
Ainsi φ est endomorphisme de $\mathbb{K}_n[X]$.
Pour $P \in \ker\varphi$, on a $\left((X^2-1)^n\right)^{(n)}$ donc $P(X+1) = -P(X)$.
Si P n'est pas le polynôme nul, le coefficient dominant de $P(X+1)$ est d'une part égal à celui de P et d'autre part égal à celui de $-P$. C'est absurde et donc $P = 0$.
Ainsi $\ker\varphi = \{0\}$.
L'endomorphisme φ est injectif et puisque

$$((X+1)^n)^{(n-k)} = \frac{n!}{k!}(X+1)^k$$

on peut conclure que φ est un automorphisme.

8.1.8 Fonctions polynomiales

8.1.8.1 Valeur d'un polynôme en un point

Définition

On appelle valeur de $P = a_0 + a_1X + \cdots + a_nX^n \in \mathbb{K}[X]$ en $x \in \mathbb{K}$ le scalaire

$$P(x) = a_0 + a_1x + \cdots + a_nx^n$$

Proposition

Soient $P, Q \in \mathbb{K}[X]$, $\lambda, \mu \in \mathbb{K}$ et $x \in \mathbb{K}$.

$$(\lambda.P + \mu.Q)(x) = \lambda P(x) + \mu Q(x),\ (PQ)(x) = P(x)Q(x) \text{ et } (P \circ Q)(x) = P(Q(x))$$

dém. :
Les opérations sur $\mathbb{K}[X]$ ont été définies de sorte d'obtenir ces propriétés.
□

Définition

On appelle racine (ou zéro) d'un polynôme $P \in \mathbb{K}[X]$ tout $x \in \mathbb{K}$ tel que $P(x) = 0$.

Exemple a est la seule racine du polynôme $X - a$.

Exemple Les racines n-ième de l'unité sont les racines de $X^n - 1 \in \mathbb{C}[X]$.

Définition

On appelle équation algébrique, tout équation de la forme $P(x) = 0$ d'inconnue $x \in \mathbb{K}$ et où $P \in \mathbb{K}[X]$.
Le degré de P est alors appelé degré de l'équation $P(x) = 0$.

8.1.8.2 Fonction polynomiale

Définition

On appelle fonction polynomiale associée à $P \in \mathbb{K}[X]$ définie sur $\mathcal{D} \subset \mathbb{K}$ l'application :

$$\tilde{P} : \begin{cases} \mathcal{D} \to \mathbb{K} \\ x \mapsto \tilde{P}(x) = P(x) \end{cases}$$

Lorsque $\mathcal{D} = \mathbb{K}$, on parle de fonction polynomiale associée à P.

Exemple La fonction polynomiale associée à $P = X^2 + 1$ est

$$\begin{cases} \mathbb{R} \to \mathbb{R} \\ x \to x^2 + 1 \end{cases}$$

Exemple Les fonctions polynomiales associée à $X^2 + 1$ et $X^4 + 1$ sur $\mathcal{D} = \{-1, 0, 1\}$ sont égales.

(!)Il faut distinguer :
- la fonction polynomiale $\tilde{P}$ (application) du polynôme sous-jacent P (expression) ;
- l'élément x (de $\mathbb{K}$) de l'indéterminée X (lettre permettant de décrire le polynôme).

Proposition

Soient $P, Q \in \mathbb{K}[X]$ et $\lambda, \mu \in \mathbb{K}$.

$$\lambda.\widetilde{P + \mu}.Q = \lambda.\widetilde{P} + \mu.\widetilde{Q}, \widetilde{P \times Q} = \widetilde{P} \times \widetilde{Q}$$

dém. :
Pour tout $x \in \mathcal{D}$,

$$\left[\lambda.\widetilde{P + \mu}.Q\right](x) = (\lambda P + \mu Q)(x) = \lambda P(x) + \mu Q(x) = \left[\lambda.\widetilde{P} + \mu.\widetilde{Q}\right](x)$$

Ainsi $\widetilde{\lambda.P + \mu.Q} = \lambda.\widetilde{P} + \mu.\widetilde{Q}$.
De la même façon, on obtient $\widetilde{P \times Q} = \widetilde{P} \times \widetilde{Q}$.
□

8.1.8.3 Schéma de Hörner

Comment calculer $a_0 + a_1 x + \cdots + a_n x^n$ en un minimum d'opérations ?
Idée : on réorganise le produit à l'instar du calcul suivant

$$a_0 + a_1 x + a_2 x^2 + a_3 x^3 = a_0 + (a_1 + (a_2 + a_3 x)x)x$$

Ainsi, on calcule $a_0 + a_1 x + \cdots + a_n x^n$ en évaluant

$$a_0 + (a_1 + (\cdots (a_{n-1} + a_n x)x) \cdots)x\,;$$

cela génère n additions et n multiplications.
Algorithme de Horner :
Arguments : n, $a_0, \ldots, a_n$ et x.
$S \leftarrow a_n$
Pour k allant de $n-1$ à 0 par pas de -1.
$S \leftarrow a_k + S \times x$
Fin pour.

8.2 Dérivation

8.2.1 Dérivée première

Définition

Soit $P = \sum_{n=0}^{+\infty} a_n X^n = a_0 + a_1 X + \cdots + a_n X^n + \cdots \in \mathbb{K}[X]$.
On appelle polynôme dérivé de P, le polynôme noté P' défini par :

$$P' = a_1 + 2a_2 X + \cdots + n a_n X^{n-1} + \cdots = \sum_{n=1}^{+\infty} n a_n X^{n-1} = \sum_{n=0}^{+\infty} (n+1) a_{n+1} X^n$$

Exemple Si $P = 3X^3 + 2X + 1$ alors $P' = 9X^2 + 2$.

Proposition

Si $P \in \mathbb{K}[X]$ est constant alors $P' = 0$.
Si $P \in \mathbb{K}[X]$ est non constant alors $\deg P' = \deg P - 1$.
Dans les deux cas : $\deg P' \leqslant \deg P - 1$.

dém. :
Si P constant : ok
Si P non constant, soit $p = \deg P$.
On peut écrire $P = a_0 + a_1 X + \cdots + a_p X^p$ avec $a_p \neq 0$.
On a alors $P' = a_1 + 2a_2 X + \cdots + p a_{p-1} X^{p-1}$ et on peut conclure.
□

Corollaire

$P' = 0 \Leftrightarrow P$ est un polynôme constant.

Proposition

$\forall \lambda, \mu \in \mathbb{K}, \forall P, Q \in \mathbb{K}[X]$,

$$(\lambda P + \mu Q)' = \lambda P' + \mu Q' \text{ et } (PQ)' = P'Q + PQ'$$

dém. :

Soient $P = \sum_{n=0}^{+\infty} a_n X^n$, $Q = \sum_{n=0}^{+\infty} b_n X^n$ polynôme de $\mathbb{K}[X]$.

On a $\lambda P + \mu Q = \sum_{n=0}^{+\infty} (\lambda a_n + \mu b_n) X^n$ donc

$$(\lambda P + \mu Q)' = \sum_{n=0}^{+\infty} (n+1)(\lambda a_{n+1} + \mu b_{n+1}) X^n$$
$$= \lambda \sum_{n=0}^{+\infty} (n+1) a_{n+1} X^n + \mu \sum_{n=0}^{+\infty} (n+1) b_{n+1} X^n = \lambda P' + \mu Q'$$

On a aussi $PQ = \sum_{n=0}^{+\infty} c_n X^n$ avec $c_n = \sum_{k=0}^{n} a_k b_{n-k}$ donc

$(PQ)' = \sum_{n=0}^{+\infty} (n+1) c_{n+1} X^n$.

D'autre part $P' = \sum_{n=0}^{+\infty} (n+1) a_{n+1} X^n$ donc $P'Q = \sum_{n=0}^{+\infty} \alpha_n X^n$ avec $\alpha_n = \sum_{k=0}^{n} (k+1) a_{k+1} b_{n-k}$,

et $Q' = \sum_{n=0}^{+\infty} (n+1) b_{n+1} X^n$ donc $PQ' = \sum_{n=0}^{+\infty} \beta_n X^n$ avec $\beta_n = \sum_{k=0}^{n} a_k (n+1-k) b_{n+1-k}$,

Ainsi $P'Q + PQ' = \sum_{n=0}^{+\infty} \gamma_n X^n$ avec $\gamma_n = \alpha_n + \beta_n$.

Or

$$\gamma_n = \sum_{k=0}^{n} (k+1) a_{k+1} b_{n-k} + \sum_{k=0}^{n} (n+1-k) a_k b_{n+1-k}$$

En procédant à un changement de d'indice dans la première somme

$$\gamma_n = \sum_{k=1}^{n+1} k a_k b_{n+1-k} + \sum_{k=0}^{n} (n+1-k) a_k b_{n+1-k}$$

En combinant les deux sommes en une seule

$$\gamma_n = \sum_{k=0}^{n+1} (n+1) a_k b_{n+1-k} = (n+1) c_{n+1}$$

Ainsi les polynômes $(PQ)'$ et $P'Q + PQ'$ sont égaux car de mêmes coefficients.
□

Corollaire

$$(P_1P_2...P_n)' = \sum_{i=1}^{n} (P_1...\hat{P}_i...P_n)P_i' \text{ et } (P^n)' = nP'P^{n-1}.$$

dém. :
Par récurrence sur $n \in \mathbb{N}^\star$, on obtient la première relation et la deuxième s'obtient en particularisant avec $P_1 = \ldots = P_n = P$.
□

Proposition

$\forall P, Q \in \mathbb{K}[X]$,

$$(P \circ Q)' = Q' \times P' \circ Q$$

dém. :
Ecrivons $P = \sum_{n=0}^{+\infty} a_n X^n$.

On a alors $P(Q) = \sum_{n=0}^{+\infty} a_n Q^n$.

Par ce qui précède $P(Q)' = \sum_{n=1}^{+\infty} n a_n Q' Q^{n-1}$ et donc $P(Q) = Q' \sum_{n=1}^{+\infty} n a_n Q^{n-1} = Q'P'(Q)$
□

Définition

On appelle polynôme primitif de $P \in \mathbb{K}[X]$ tout polynôme $Q \in \mathbb{K}[X]$ tel que $Q' = P$.

Remarque Tout polynôme P possède au moins un polynôme primitif Q et l'ensemble des polynômes primitifs de P est constitué des polynômes de la forme $Q + C$ avec $C \in \mathbb{K}$.

8.2.2 Dérivée d'ordre supérieur

Soit $D : \mathbb{K}[X] \to \mathbb{K}[X]$ défini par $D(P) = P'$.
D est un endomorphisme de $\mathbb{K}[X]$. On peut introduire ses itérés :
$D^0 = \mathrm{Id}, D^1 = D, D^2 = D \circ D, \ldots, D^n = D \circ D \circ \ldots \circ D$ (n termes)

Définition

Pour $n \in \mathbb{N}$ et $P \in \mathbb{K}[X]$, le polynôme $P^{(n)} = D^n(P)$ est appelé polynôme dérivé d'ordre n de P.

Exemple $P^{(0)} = P, P^{(1)} = P', P^{(2)} = (P')' = P''$ etc.

Proposition

Soit $P \in \mathbb{K}[X]$ et $n \in \mathbb{N}$.
Si $\deg P < n$ alors $\deg P^{(n)} = -\infty$.
Si $\deg P \geqslant n$ alors $\deg P^{(n)} = \deg P - n$.

dém. :
Par récurrence sur $n \in \mathbb{N}$.
□

Proposition

$\forall \lambda, \mu \in \mathbb{K}, \forall P, Q \in \mathbb{K}[X]$,

$$(\lambda P + \mu Q)^{(n)} = \lambda P^{(n)} + \mu Q^{(n)} \text{ et } (PQ)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} P^{(k)} Q^{(n-k)}$$

dém. :
La relation $(\lambda P + \mu Q)^{(n)} = \lambda P^{(n)} + \mu Q^{(n)}$ s'obtient par linéarité de D^n.
La relation $(PQ)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} P^{(k)} Q^{(n-k)}$ s'obtient comme pour la démonstration de la formule de Leibniz dans le cours de dérivation des fonctions numériques.
□

Exemple Soient $n, k \in \mathbb{N}$. Exprimons $(X^n)^{(k)}$.
On a $(X^n)' = nX^{n-1}$, $(X^n)'' = n(n-1)X^{n-2}$,...
Pour $0 \leqslant k \leqslant n$,

$$(X^n)^{(k)} = n(n-1)\dots(n-k+1)X^{n-k} = \frac{n!}{(n-k)!}X^{n-k}$$

En particulier $(X^n)^{(n)} = n!$.
Pour $k > n$, $(X^n)^{(k)} = 0$.

Remarque Par ce qui précède et la linéarité de D^k, on obtient les formules de dérivation à l'ordre k d'un polynôme :
Si $P = \sum_{n=0}^{+\infty} a_n X^n$ alors

$$P^{(k)} = \sum_{n=k}^{+\infty} a_n \frac{n!}{(n-k)!} X^{n-k} = \sum_{n=0}^{+\infty} a_{n+k} \frac{(n+k)!}{n!} X^n$$

Exemple Montrons

$$\sum_{k=0}^{n} \binom{n}{k}^2 = \binom{2n}{n}$$

en étudiant le coefficient de X^n dans $\left((X^2-1)^n\right)^{(n)}$.
On a $(X^2-1)^n = X^{2n} + \cdots$ (où $\cdots$ désigne des puissances inférieures de X)
Par suites

$$\left((X^2-1)^n\right)^{(n)} = \frac{(2n)!}{n!} X^n + \cdots$$

D'autre part
$$\left((X^2-1)^n\right)^{(n)} = ((X-1)^n(X+1)^n)^{(n)}$$
Par la formule de Leibniz
$$\left((X^2-1)^n\right)^{(n)} = \sum_{k=0}^{n}\binom{n}{k}((X-1)^n)^{(k)}\,((X+1)^n)^{(n-k)}$$
Or
$$((X-1)^n)^{(k)} = \frac{n!}{(n-k)!}(X-1)^{n-k} \text{ et } ((X+1)^n)^{(n-k)} = \frac{n!}{k!}(X+1)^k$$
Le coefficient de X^n dans ce calcul de $\left((X^2-1)^n\right)^{(n)}$ est
$$\sum_{k=0}^{n}\binom{n}{k}\frac{n!}{(n-k)!}\frac{n!}{k!} = n!\sum_{k=0}^{n}\binom{n}{k}^2$$
On en déduit
$$\sum_{k=0}^{n}\binom{n}{k}^2 = \frac{(2n)!}{(n!)^2} = \binom{2n}{n}$$

8.2.3 Formule de Taylor

Théorème

$\forall P \in \mathbb{K}[X], \forall a \in \mathbb{K}$,
$$P = \sum_{n=0}^{+\infty}\frac{P^{(n)}(a)}{n!}(X-a)^n$$

dém. :
Si $P=0$, la propriété est immédiate car
$$\forall n \in \mathbb{N}, P^{(n)}(a) = 0$$
Si $P \neq 0$ alors posons $p = \deg P$.
Puisque la famille $(1, X-a, (X-a)^2, ..., (X-a)^p)$ est une famille de polynômes de degrés étagés, c'est une base de $\mathbb{K}_p[X]$ ce qui permet d'écrire :
$P = \lambda_0 + \lambda_1(X-a) + ... + \lambda_p(X-a)^p$ avec $\lambda_0, \ldots, \lambda_p$ les composantes du polynôme P dans cette base.
Pour tout $0 \leqslant n \leqslant p$, on a alors
$$P^{(n)} = 0 + \lambda_n n! + \mu_{n+1}(X-a) + ... + \mu_p(X-a)^{p-n}$$
avec
$$\mu_{n+k} = \lambda_{n+k}\frac{(n+k)!}{k!}$$
On en déduit $P^{(n)}(a) = n!\lambda_n$ ce qui déterminer λ_n.
De plus, pour tout $n > p$, on a $P^{(n)}(a) = 0$ et on obtient donc la formule proposée.
□

Remarque La formule de Taylor s'écrit encore

$$P(a+X) = \sum_{n=0}^{+\infty} \frac{P^{(n)}(a)}{n!} X^n$$

En particulier, on a $P = \sum_{n=0}^{+\infty} a_n X^n$ avec

$$\forall n \in \mathbb{N}, a_n = \frac{P_n(0)}{n!}$$

8.3 Arithmétique des polynômes

8.3.1 Divisibilité

8.3.1.1 Polynômes associés

Définition

On dit qu'un polynôme $P \in \mathbb{K}[X]$ est associé à un polynôme $Q \in \mathbb{K}[X]$ s'il existe $\lambda \in \mathbb{K}^\star$ tel que $P = \lambda Q$.

Proposition

L'association de polynôme définit une relation binaire sur $\mathbb{K}[X]$ à la fois réflexive, symétrique et transitive ; c'est une relation d'équivalence sur $\mathbb{K}[X]$.

dém. :
P est associé à P car $P = 1 \times P$.
Si P est associé à Q alors il existe $\lambda \neq 0$ tel que $P = \lambda Q$ et alors $Q = \mu P$ avec $\mu = 1/\lambda \neq 0$ et donc Q est associé à P.
Enfin, si P est associé à Q et Q associé à P alors il existe $\lambda, \mu \neq 0$ tel que $P = \lambda Q$ et $Q = \mu R$ ce qui permet d'écrire $P = (\lambda\mu)R$ avec $\lambda\mu \neq 0$ et donc P est associé à R.
□

Proposition

Si P et Q sont associés et ont mêmes coefficients dominants alors $P = Q$.

dém. :
Si $P = \lambda Q$ et si P et Q ont même coefficients dominants alors $\lambda = 1$ et donc $P = Q$.
□

Définition

Un polynôme $P \in \mathbb{K}[X]$ est dit unitaire (ou normalisé) si son coefficient dominant est égal à 1.

Proposition

Tout polynôme $P \in \mathbb{K}[X]$ non nul est associé à unique polynôme unitaire.

dém. :
L'unicité provient de ce que deux polynômes unitaires associés sont nécessairement égaux car ayant le même coefficient dominant.

L'existence provient du fait que $P \neq 0$ est associé au polynôme unitaire λP en prenant pour λ l'inverse du coefficient dominant de P.

□

8.3.1.2 Relation de divisibilité

Définition

On dit que $A \in \mathbb{K}[X]$ divise $B \in \mathbb{K}[X]$ s'il existe $U \in \mathbb{K}[X]$, $B = AU$.
On note alors $A \mid B$.

Définition

Pour $A \in \mathbb{K}[X]$, on note Div(A) l'ensemble des diviseurs de A et Mul(A) l'ensemble de ses multiples.
Ainsi

$$\text{Div}(A) = \{D \in \mathbb{K}[X] / D \mid A\} \text{ et } \text{Mul}(A) = \{AU/U \in \mathbb{K}[X]\} = A.\mathbb{K}[X]$$

Exemple $(X+1) \mid (X^3 + X^2 + X + 1)$ car $(X^3 + X^2 + X + 1) = (X+1)(X^2+1)$.

Exemple $(X-1) \mid (X^n - 1)$ car $X^n - 1 = (X-1)(1 + X + \cdots + X^{n-1})$.

Exemple Pour tout $A \in \mathbb{K}[X]$ on a $A \mid 0$ car $0 = A \times 0$.

Exemple Les polynômes constants non nuls et les polynômes associés à $A \in \mathbb{K}[X]$ divisent A.

Proposition

Soient A et B des polynômes respectivement associés à C et D. On a

$$A \mid B \Leftrightarrow C \mid D$$

dém. :
On peut écrire $A = \lambda C$ et $B = \mu D$ avec $\lambda, \mu \in \mathbb{K}^\star$.
Si $C \mid D$ alors il existe $U \in \mathbb{K}[X]$ tel que $D = CU$ et alors $B = AV$ avec $V = \dfrac{\mu}{\lambda} U$. Ainsi $A \mid B$.
De façon symétrique, $A \mid B$ entraîne $C \mid D$.

□

Remarque Tout problème de divisibilité peut se ramener à un problème entre polynômes unitaires ou nuls.

8.3.1.3 Propriétés de la divisibilité

Proposition

$\forall A, B, C \in \mathbb{K}[X]$
$A \mid B$ et $B \mid C \Rightarrow A \mid C$,
$A \mid B$ et $B \mid A \Rightarrow A$ et B sont associés.

dém. :
Si $A \mid B$ et $B \mid C$ alors il existe $U, V \in \mathbb{K}[X]$ tels que $B = AU$ et $C = BV$. On a alors $C = A(UV)$ donc $A \mid C$.
Si $A \mid B$ et $B \mid A$ alors il existe $U, V \in \mathbb{K}[X]$ tels que $B = AU$ et $A = BV$. On a alors $B = B(UV)$.
Cas $B \neq 0$:
On obtient $UV = 1$ et donc les polynômes U et V sont constants non nuls. On en déduit que A et B sont associés.
Cas $B = 0$:
On a $A = BV = 0$ et donc A et B sont associés.

□

Proposition

$\forall A, B \in \mathbb{K}[X]$
$A \mid B$ et $B \neq 0 \Rightarrow \deg A \leqslant \deg B$,
$A \mid B$ et $\deg A = \deg B \Rightarrow A$ et B sont associés.

dém. :
Si $A \mid B$ et $B \neq 0$ alors il existe $U \in \mathbb{K}[X]$ non nul tel que $B = AU$. On a alors $\deg B = \deg A + \deg U \geqslant \deg A$ car $\deg U \in \mathbb{N}$.
Si $A \mid B$ et $\deg A = \deg B$ alors il existe $U \in \mathbb{K}[X]$ tel que $B = AU$.
Cas $B \neq 0$:
On a $\deg B = \deg A + \deg U$ avec $\deg B = \deg A \in \mathbb{N}$ donc $\deg U = 0$.
Ainsi U est un polynôme constant non nul et donc la relation $B = AU$ avec $U \in \mathbb{K}^\star$ donne A et B associés.
Cas $B = 0$:
$\deg A = \deg B$ entraîne $A = 0$ et donc A et B sont associés.

□

Remarque En arithmétique des polynômes on montre souvent $A = B$ via :
- $A \mid B, B \mid A$ et A et B ont le même coefficient dominant ;
- $A \mid B, \deg A = \deg B$ et A et B ont le même coefficient dominant.

Proposition

$\forall A, B, C, D \in \mathbb{K}[X]$
$A \mid B$ et $A \mid C \Rightarrow A \mid (B + C)$,
$A \mid B$ et $C \mid D \Rightarrow AC \mid BD$,
$A \mid B \Rightarrow \forall n \in \mathbb{N}, A^n \mid B^n$.

dém. :
Il suffit d'adapter les démonstrations des résultats semblables vu pour l'arithmétique des entiers.

□

8.3.2 Division euclidienne

8.3.2.1 Énoncé

Théorème

[Division euclidienne polynomiale]
Pour tout $A, B \in \mathbb{K}[X]$ avec $B \neq 0$ il existe un unique couple $(Q, R) \in \mathbb{K}[X]^2$ vérifiant

$$A = BQ + R \text{ et } \deg R < \deg B$$

Les polynômes Q et R sont appelés quotient et reste de la division euclidienne de A par B.

dém. :
Unicité :
Supposons (Q_1, R_1) et (Q_2, R_2) solutions.
On a $A = BQ_1 + R_1 = BQ_2 + R_2$ avec $\deg R_1, \deg R_2 < \deg B$.
Par suite $B(Q_1 - Q_2) = R_2 - R_1$ et donc $\deg B + \deg(Q_1 - Q_2) = \deg(R_2 - R_1) < \deg B$.
On en déduit $Q_1 - Q_2 = 0$ puis $R_2 - R_1 = 0$.
Finalement $(Q_1, R_1) = (Q_2, R_2)$.
Existence :
Posons $p \in \mathbb{N}$ le degré du polynôme B et b son coefficient dominant de B.
Montrons, par récurrence sur $n \geqslant p$ la propriété.

$$\forall A \in \mathbb{K}_{n-1}[X], \exists (Q, R) \in \mathbb{K}[X]^2, A = BQ + R \text{ avec } \deg R < p$$

Pour $n = p$: $Q = 0$ et $R = A$ conviennent.
Supposons la propriété établie au rang $n \geqslant p$.
Soit $A \in \mathbb{K}_n[X]$. On peut écrire $A = aX^n + \hat{A}$ avec $\deg \hat{A} < n$.
Or on a

$$\frac{a}{b} B X^{n-p} = aX^n + \hat{B} \text{ avec } \deg \hat{B} < n$$

Donc

$$A = \frac{a}{b} B X^{n-p} + \hat{A} - \hat{B}$$

Puisque $\deg(\hat{A} - \hat{B}) < n$, l'hypothèse de récurrence donne l'existence de $(\hat{Q}, \hat{R}) \in \mathbb{K}[X]$ tel que

$$\hat{A} - \hat{B} = \hat{Q}B + \hat{R} \text{ avec } \deg \hat{R} < p$$

On peut alors écrire

$$A = B\left(\frac{a}{b} X^{n+1-p} + \hat{Q}\right) + \hat{R} = BQ + R$$

avec

$$Q = \frac{a}{b} X^{n+1-p} + \hat{Q} \text{ et } R = \hat{R}$$

de sorte que $\deg R < p$.
Récurrence établie.
□

8.3.2.2 Détermination pratique

Déterminons pratiquement Q et R.
Si $\deg A < \deg B$ alors $Q = 0$ et $R = A$ comme on l'a vu lors de la démonstration du théorème de division euclidienne.
Si $\deg A \geqslant \deg B$ alors on pose une division euclidienne :

- on cherche par quel monôme multiplier pour égaler la plus grande puissance de A ;
- on retire de A le multiple de B correspondant ;
- on reprend le processus avec le résultat obtenu jusqu'à obtention d'un polynôme de degré strictement inférieur à celui de B ;
- le reste de la division euclidienne et ce dernier polynôme et le quotient la somme des monômes qui ont multiplié B.

Exemple Pour $A = X^4 - 3X^3 + 2X^2 + X - 1, B = X^2 - X + 1$, on obtient $Q = X^2 - 2X - 1$, $R = 2X$.

Exemple Pour $A = X^5 - X^4 - 2X^2 - 3X + 1, B = X^2 + 2$, on obtient
$Q = X^3 - X^2 - 2X, R = X + 1$.

Arguments : A, B
$Q \leftarrow 0$
Tant que $\deg A \geqslant \deg B$ faire :
$T \leftarrow \dfrac{\text{coeff}(A)}{\text{coeff}(B)} X^{\deg A - \deg B}$, $Q \leftarrow Q + T$, $A \leftarrow A - BT$
Fin tant que
$R \leftarrow A$
Fin.

8.3.2.3 Applications

Proposition

Soient $A, B \in \mathbb{K}[X]$ tels que $B \neq 0$. On a

$$B \mid A \Leftrightarrow \text{ le reste de la division euclidienne de } A \text{ par } B \text{ est nul}$$

dém. :
Si $B \mid A$ alors il existe $Q \in \mathbb{K}[X]$ tel que $B = AQ$.
On peut alors écrire $B = AQ + R$ avec $R = 0$ vérifiant $\deg R < \deg B$.
Par cette relation, on peut identifier le quotient et le reste de la division euclidienne de A et B, et justement ce reste est nul.
Inversement, si le reste de la division euclidienne de A par B est nul alors on peut écrire $A = BQ$ avec Q le quotient de la division euclidienne de A par B et donc B divise A.
□

Proposition

$a \in \mathbb{K}$ est racine de $P \in \mathbb{K}[X]$ si, et seulement si, $X - a$ divise P.

dém. :
La division euclidienne de P par $X - a$ s'écrit :
$P = (X - a)Q + R$ avec $\deg R < 1$.
R est donc un polynôme constant égal à λ et on peut écrire $P = (X - a)Q + \lambda$
Evaluons cette relation en a, on obtient $P(a) = 0 \times Q(a) + \lambda$ donc $R = \lambda = P(a)$.
Par suite $X - a$ divise P si, et seulement si, $P(a) = 0$.
□

Exemple Considérons $P = X^3 + X^2 - 3X + 1$.
1 est racine de P donc on peut factoriser P par $X - 1$.
On obtient $P = (X - 1)(X^2 + 2X - 1)$.

Exemple Considérons $P = X^3 + 1$
-1 est racine de P donc on peut factoriser P par $X + 1$.
On obtient $P = (X + 1)(X^2 - X + 1)$.

Exemple $P = X^4 + 3X^3 - 3X - 1$.
1 et -1 sont racines de P.
On obtient $P = (X - 1)(X + 1)(X^2 + 3X + 1)$.

8.3.3 Pgcd et ppcm de deux polynômes

8.3.3.1 Pgcd

Définition

On note $\text{Div}(A, B) = \text{Div}(A) \cap \text{Div}(B)$ l'ensemble des diviseurs communs à A et B.

Exemple Les polynômes constants non nuls sont diviseurs communs à A et B.

Remarque Si $P \in \text{Div}(A, B)$ tout polynôme associé à P est aussi dans $\text{Div}(A, B)$.

Proposition

Si $A = BQ + R$ alors
$$\text{Div}(A, B) = \text{Div}(B, R)$$

Théorème

Soient $A, B \in \mathbb{K}[X]$, il existe un unique polynôme $D \in \mathbb{K}[X]$ unitaire ou nul, tel que
$$\text{Div}(A, B) = \text{Div}(D)$$

dém. :
Unicité : Si D_1, D_2 sont solutions alors $\text{Div}(D_1) = \text{Div}(D_2)$ donc $D_1 \mid D_2$ et $D_2 \mid D_1$ et par suite ils sont associés.
Si l'un est nul alors l'autre aussi et $D_1 = D_2$.
S'ils ne sont pas nuls, ils sont unitaires et associés donc égaux.
Existence :
Si $A = B = 0$ alors $D = 0$ convient
Sinon, quitte à échanger A et B on peut supposer $B \neq 0$.
Posons $A_0 = A$ et $A_1 = B$. On réalise ensuite les divisions euclidiennes suivantes tant que les restes

obtenus sont non nuls (c'est l'algorithme d'Euclide) :
$A_0 = A_1Q_1 + A_2$ avec $\deg A_2 < \deg A_1$,
...
$A_{m-2} = A_{m-1}Q_{m-1} + A_m$ avec $\deg A_m < \deg A_{m-1}$,
$A_{m-1} = A_mQ_m + 0$.
Ce processus s'arrête puisque $\deg A_1 > \deg A_2 > \cdots$ et ces quantités sont des entiers naturels.
On a alors

$$\text{Div}(A, B) = \text{Div}(A_0, A_1) = \ldots = \text{Div}(A_m, 0) = \text{Div}(A_m)$$

Le polynôme D unitaire associé à A_m convient.
□

Définition

Ce polynôme D est appelé pgcd des polynômes A et B.
On note $D = \text{pgcd}(A, B)$ ou $D = A \wedge B$.

Exemple Déterminons $D = \text{pgcd}(X^3 + X^2 - 2, X^3 + X - 2)$.
Par divisions euclidiennes successives :
$X^3 + X^2 - 2 = (X^3 + X - 2) \times 1 + X^2 - X$,
$X^3 + X - 2 = (X^2 - X)(X + 1) + 2X - 2$,
$X^2 - X = (2X - 2) \times \dfrac{1}{2}X + 0$.
Donc D est le polynôme unitaire associé au dernier reste non nul, i.e. associé à $2X - 2$.
Ainsi $D = X - 1$.

Théorème

Si $D = \text{pgcd}(A, B)$ alors il existe $U, V \in \mathbb{K}[X]$ tels que

$$D = AU + BV$$

dém. :
Si $A = B = 0$ alors $D = 0$ et U, V quelconques conviennent.
Sinon, on réalise comme ci-dessus l'algorithme d'Euclide puis on écrit successivement les A_i sous la forme $AU_i + BV_i$ avec $U_i, V_i \in \mathbb{K}[X]$.
A terme on parvient à écrire D sous la forme $AU + BV$.
□
(!)Les polynômes U et V ne sont pas uniques.

Exemple Reprenons $A = X^3 + X^2 - 2$ et $B = X^3 + X - 2$ pour lesquels $D = \text{pgcd}(A, B) = X - 1$.
En renversant les divisions euclidiennes précédentes
$X^2 - X = A - B$,
$2X - 2 = B - (X + 1)(X^2 - X) = B - (X + 1)(A - B)$,
puis

$$D = \frac{1}{2}(X + 2)B - \frac{1}{2}(X + 1)A = AU + BV$$

avec

$$U = -\frac{1}{2}(X + 1) \text{ et } V = \frac{1}{2}(X + 2)$$

8.3.3.2 Propriétés du pgcd

Proposition

$$\text{pgcd}(A, B) = \text{pgcd}(B, A)$$

dém. :
$\text{Div}(A, B) = \text{Div}(B, A)$
□

Proposition

Si $A \mid B$ alors $\text{pgcd}(A, B)$ est associé à A.

dém. :
Si $A \mid B$ alors $\text{Div}(A, B) = \text{Div}(A)$
□

Proposition

$$P \mid A \text{ et } P \mid B \Leftrightarrow P \mid \text{pgcd}(A, B)$$

dém. :
Par définition, les diviseurs communs à A et B sont les diviseurs de $\text{pgcd}(A, B)$.
□

Proposition

Si C est unitaire alors
$$\text{pgcd}(AC, BC) = \text{pgcd}(A, B) \times C$$

dém. :
Posons $\Delta = \text{pgcd}(AC, BC)$ et $D = \text{pgcd}(A, B)$.
$DC \mid AC$ et $DC \mid BC$ donc $DC \mid \Delta$.
$D = AU + BV$ donc $DC = ACU + BCV$ d'où $\Delta \mid DC$.
□

8.3.3.3 Ppcm

Définition

On note $\text{Mul}(A, B) = \text{Mul}(A) \cap \text{Mul}(B)$ l'ensemble des multiples communs à A et $B \in \mathbb{K}[X]$.

Exemple 0 et AB sont multiples communs à A et B.

Remarque Si P est multiple commun à A et B alors tout polynôme associé à P l'est encore.

Théorème

Soient $A, B \in \mathbb{K}[X]$, il existe un unique polynôme unitaire ou nul tel que $\text{Mul}(A, B) = \text{Mul}(M)$.

dém. :
Unicité : Si M_1 et M_2 sont solutions alors Mul(A,B) = Mul(M_1) = Mul(M_2). Par suite $M_1 \in$ Mul(M_2) et donc $M_2 \mid M_1$. De même $M_1 \mid M_2$ et donc M_1 et M_2 sont associés. Puisque M_1 et M_2 sont unitaires ou nuls et qu'ils sont associés, ils sont égaux.
Existence : Si $A = 0$ ou $B = 0$ alors $M = 0$ convient.
Sinon, A et B possède des multiples communs non nul (par exemple AB), considérons en un qui soit unitaire et de degré minimal, notons le M.
On a Mul$(A,B) \supset$ Mul(M) car $M \in$ Mul(A,B).
Considérons $N \in$ Mul(A,B) et réalisons la division euclidienne de N par M : $N = MQ + R$ avec $\deg R < \deg M$.
$A \mid N, A \mid M$ donc $A \mid R$ et de même $B \mid R$ donc $R \in$ Mul(A,B).
Or M est un multiple commun non nul de degré minimal donc $R = 0$.
Par suite $M \mid N$ et donc $N \in$ Mul(A,B).
Finalement Mul(A,B) = Mul(M).
□

Définition

M est alors appelé ppcm de A et B.
On note $M = \text{ppcm}(A,B)$ ou $A \vee B$.

8.3.3.4 Propriété du ppcm

Proposition

$$\text{ppcm}(A,B) = \text{ppcm}(B,A)$$

dém. :
Mul(A,B) = Mul(B,A).
□

Proposition

Si $A \mid B$ alors ppcm(A,B) est associé à B.

dém. :
Si $A \mid B$ alors Mul(A,B) = Mul(B).
□

Proposition

$$A \mid P \text{ et } B \mid P \Leftrightarrow \text{ppcm}(A,B) \mid P$$

dém. :
Par définition, les multiples communs à A et B sont les multiples de ppcm(A,B).
□

Proposition

Si C est unitaire alors ppcm(AC,BC) = ppcm$(A,B) \times C$.

dém. :
Posons $M = \text{ppcm}(A,B)$ et $N = \text{ppcm}(AC,BC)$.

Comme $AC \mid MC$ et $BC \mid MC$ on a $N \mid MC$.
Comme $AC \mid N$, on a $C \mid N$ d'où $N = PC$.
Comme $AC \mid PC, BC \mid PC$ et $C \neq 0$ on a $A \mid P, B \mid P$ donc $M \mid P$ puis $MC \mid N$.
□

Théorème

$\forall A, B \in \mathbb{K}[X]$, pgcd$(A,B) \times$ ppcm(A,B) est associé à AB.

dém. :
Si $A = 0$ ou $B = 0$ alors ppcm$(A,B) = 0$ et l'affirmation proposée est vraie.
Supposons désormais $A, B \neq 0$.
Posons $D =$ pgcd(A,B) et $M =$ ppcm(A,B).
Puisque D divise A et B, on peut écrire $A = DA'$ et $B = DB'$ avec $A', B' \in \mathbb{K}[X]$. Considérons alors le polynôme $P = DA'B' = A'B = AB'$.
Les polynômes A et B divisent P donc M divise P puis DM divise $DP = AB$.
Inversement, puisque AB est un multiple commun à A et B, on peut écrire $AB = MQ$ avec $Q \in \mathbb{K}[X]$.
Or M est un multiple de A donc on peut écrire $M = AC$. La relation $AB = ACQ$ donne alors $B = CQ$.
Ainsi Q divise B. De façon analogue, on obtient que Q divise A et donc Q divise D.
On en déduit que $AB = MQ$ divise MD.
Puisque AB et MD se divisent mutuellement, ces deux polynômes sont associés.
□

8.3.4 Polynômes premiers entre eux

8.3.4.1 Définition

Définition

Deux polynômes A et B sont dits premiers entre eux si leurs seuls diviseurs communs sont les polynômes constants non nuls.
Ceci signifie encore pgcd$(A,B) = 1$. On note alors $A \wedge B = 1$.

Exemple Soient $a, b \in \mathbb{K}$ tels que $a \neq b$.
Montrons

$$(X - a) \wedge (X - b) = 1$$

Soit $D =$ pgcd$(X - a, X - b)$.
Puisque $D \mid X - a$ on a $\deg D \leqslant 1$ et puisque $D \neq 0$, $\deg D = 0$ ou 1.
Si $\deg D = 1$ alors D est associé à $X - a$ et $X - b$ par divisibilité et égalité de degré.
On en déduit que $X - a$ et $X - b$ sont associés, or ils sont unitaires, donc $X - a = X - b$.
Ceci est exclu puisque $a \neq b$.
Finalement $\deg D = 0$, puis $D = 1$.

8.3.4.2 Le théorème de Bézout

Théorème

On a équivalence entre :
(i) A et B sont premiers entre eux ;
(ii) $\exists U, V \in \mathbb{K}[X]$, $AU + BV = 1$.

dém. :
(i) $\Rightarrow$ (ii) Par égalité de Bézout.

(ii) $\Rightarrow$ (i) Supposons qu'il existe $U, V \in \mathbb{K}[X]$ tels que $AU + BV = 1$.
Si l'on pose $D = \text{pgcd}(A, B)$ alors $D \mid A$, $D \mid B$ donc $D \mid 1 = AU + BV$. Par suite $D = 1$.
□

Corollaire

$A \wedge B = 1$ et $A \wedge C = 1 \Rightarrow A \wedge (BC) = 1$;
$A \wedge B_1 = 1, \ldots, A \wedge B_n = 1 \Rightarrow A \wedge (B_1 \ldots B_n) = 1$;
$A \wedge B = 1 \Rightarrow \forall m, n \in \mathbb{N}, A^m \wedge B^n = 1$.

dém. :
Si A et B, d'une part, et A et C, d'autre part, sont premiers entre eux alors il existe U, V, W, T tels que $AU + BV = 1$ et $AW + CT = 1$. En faisant le produit de ces deux relations, on obtient
$A(AUW + BVW + CUT) + BC(VT) = 1$ et donc A et BC sont premiers entre eux.
La deuxième implication se démontre par récurrence à partir de la première.
En prenant les B_i égaux à B, on obtient

$$A \wedge B = 1 \Rightarrow A \wedge B^n = 1$$

puis par un argument de symétrie

$$A \wedge B^n = 1 \Rightarrow A^m \wedge B^n = 1$$

et ainsi on obtient la troisième implication.
□

Exemple Soient $a, b \in \mathbb{K}$ tels que $a \neq b$ et $\alpha, \beta \in \mathbb{N}$. On a

$$(X - a)^\alpha \wedge (X - b)^\beta = 1$$

8.3.4.3 Le théorème de Gauss

(!)$A \mid BC$ n'entraîne pas $A \mid B$ ou $A \mid C$.

Théorème

$A \mid BC$ et $A \wedge B = 1 \Rightarrow A \mid C$.

dém. :
Supposons $A \wedge B = 1$. Il existe $U, V \in \mathbb{K}[X]$ tels que $AU + BV = 1$ et par suite $C = ACU + BCV$.
Si de plus $A \mid BC$ alors, par opérations, $A \mid C$.
□
(!)$A \mid C$ et $B \mid C$ n'entraînent pas $AB \mid C$.

Théorème

$A \mid C, B \mid C$ et $A \wedge B = 1 \Rightarrow AB \mid C$.

dém. :
Puisque B divise C, il existe $U \in \mathbb{K}[X]$ tel que $C = BU$.
Puisque $A \mid C = BU$ et $A \wedge U = 1$, le théorème de Gauss entraîne $A \mid U$ puis $AB \mid BU = C$.
□

Corollaire

Si $A_1, \ldots, A_n$ sont des diviseurs de P deux à deux premiers entre eux
alors $A_1 \ldots A_n \mid P$.

dém. :
Par récurrence sur $n \in \mathbb{N}^\star$.
□

Exemple Si $a_1, \dots, a_n$ sont des racines deux à deux distinctes de P alors $(X - a_1) \dots (X - a_n) \mid P$.

8.3.5 Décomposition en produit de facteurs irréductibles

Définition

> On dit qu'un polynôme non constant $P \in \mathbb{K}[X]$ est irréductible dans $\mathbb{K}[X]$ si ses seuls diviseurs sont ses diviseurs triviaux.
> Sinon, le polynôme non constant P est dit composé.

Exemple Si P est irréductible, tout polynôme associé à P l'est encore.

Exemple $\forall a \in \mathbb{K}$, $X - a$ est irréductible.
En effet, soit D un diviseur de $X - a$.
Puisque $X - a \neq 0$, $\deg D \leqslant 1$ et $D \neq 0$.
Si $\deg D = 0$ alors D est constant non nul.
Si $\deg D = 1$ alors D est associé à $X - a$.

Exemple Dans $\mathbb{R}[X]$, le polynôme $X^2 + 1$ est irréductible.
En effet, si D est une diviseur de $X^2 + 1$ alors $\deg D \leqslant 2$ et $D \neq 0$.
Si $\deg D = 0$ ou 2 alors D est diviseur trivial de $X^2 + 1$.
Sinon D est degré 1 et donc de la forme $aX + b$ avec $a \neq 0$.
Par suite D possède une racine $\alpha = -b/a \in \mathbb{R}$ qui sera aussi racine de $X^2 + 1$.
C'est impossible

(!)L'irréductibilité d'un polynôme dépend du corps de base $\mathbb{K}$:
$X^2 + 1$ est irréductible dans $\mathbb{R}[X]$ et composé dans $\mathbb{C}[X]$ puisque

$$X^2 + 1 = (X - i)(X + i)$$

Proposition

> Soient $A \in \mathbb{K}[X]$ et P un polynôme irréductible de $\mathbb{K}[X]$.
> Si P ne divise pas A alors $P \wedge A = 1$.

dém. :
Par contraposée :
Si $P \wedge A \neq 1$ alors introduisons $D = \text{pgcd}(P, A) \neq 1$.
Puisque $D \mid P$ et que D n'est pas constant, D est associé à P.
Or $D \mid A$ donc $P \mid A$.
□

Proposition

Soient $A, B \in \mathbb{K}[X]$ et P un polynôme irréductible dans $\mathbb{K}[X]$.

$$P \mid AB \Rightarrow P \mid A \text{ ou } P \mid B$$

dém. :
Supposons $P \mid AB$ et P ne divise pas A.
Puisque P est irréductible, on a $A \wedge P = 1$ et, par le théorème de Gauss, $P \mid B$.
□

Théorème

Si A est un polynôme non constant de $\mathbb{K}[X]$ alors il existe $\lambda \in \mathbb{K}^\star$, $N \in \mathbb{N}^\star$, $P_1, \dots, P_N$ polynômes irréductibles de $\mathbb{K}[X]$ unitaires deux à deux distincts et il existe $\alpha_1, ..., \alpha_N \in \mathbb{N}^\star$ tels que

$$A = \lambda P_1^{\alpha_1} P_2^{\alpha_2} ... P_N^{\alpha_N}$$

De plus cette décomposition est unique à l'ordre près des facteurs, on l'appelle décomposition en facteurs irréductible de A.

dém. :
Unicité à l'ordre près des facteurs :
Supposons

$$A = \lambda P_1^{\alpha_1} \dots P_N^{\alpha_N} = \mu Q_1^{\beta_1} \dots Q_M^{\beta_M}$$

avec les hypothèses décrites dans le théorème.
Puisque les P_i et Q_j sont unitaires, les scalaires λ et μ correspondent tous deux au coefficient dominant de A. On en déduit $\lambda = \mu$.
Soit $1 \leqslant i \leqslant N$. On a $P_i \mid A$ donc $P_i \mid Q_1^{\beta_1} \dots Q_M^{\beta_M}$.
Puisque P_i est irréductible, il existe $1 \leqslant j \leqslant M$ tel que $P_i \mid Q_j$ et donc $P_i = Q_j$ puisque Q_j est irréductible. De plus l'indice j est unique car les $Q_1, \dots, Q_M$ sont deux à deux distincts.
Ce qui précède permet d'introduire une application $i \mapsto j$ de $\{1, \dots, N\}$ vers $\{1, \dots, M\}$ telles que $P_i = Q_j$.
Puisque les $P_1, \dots, P_N$ sont deux à deux distincts, cette application est injective et donc $N \leqslant M$.
Par un argument de symétrie $M \leqslant N$ et donc $M = N$.
L'application injective précédent s'avère donc être une permutations de $\{1, \dots, N\}$.
Ainsi, à l'ordre près des termes, $P_1, \dots, P_N = Q_1, \dots, Q_N$.
Quitte à réindexer, on peut supposer $P_1 = Q_1, \dots, P_N = Q_N$ et on a donc $A = \lambda P_1^{\alpha_1} \dots P_N^{\alpha_N} = \lambda P_1^{\beta_1} \dots P_N^{\beta_N}$
Comme les polynômes considérés sont irréductibles, pour $i \neq j$, $P_i \wedge P_j = 1$.
Puisque $P_i^{\alpha_i} \mid P_1^{\beta_1} \dots P_N^{\beta_N}$ et $P_i^{\alpha_i} \wedge P_j^{\beta_j} = 1$ pour tout $j \neq i$, on a $P_i^{\alpha_i} \mid P_i^{\beta_i}$ et donc $\alpha_i \leqslant \beta_i$.
Par un argument de symétrie $\beta_i \leqslant \alpha_i$ puis $\beta_i = \alpha_i$.
Existence
Montrons par récurrence forte sur $n \in \mathbb{N}^\star$ que tout polynôme de degré n est associé à un produit de polynômes irréductibles unitaires. Il suffira ensuite de réunir entre eux les éventuels polynômes irréductibles égaux pour obtenir la décomposition proposée.
Pour $n = 1$, tout polynôme de degré est associé à polynôme de la forme $X - a$ avec $a \in \mathbb{K}$ et un tel polynôme est irréductible.
Supposons la propriété établie jusqu'au rang $n \geqslant 1$.
Soit P un polynôme de degré $n + 1$.
Si P est irréductible alors on peut dire que P est associé à un polynôme irréductible unitaire.

Si P est composé alors on peut écrire $P = AB$ avec A, B polynômes non constants de degrés inférieurs à n.
Par hypothèse de récurrence A et B sont associés à un produit de polynômes irréductibles unitaires et donc $P = AB$ est aussi associé à un produit de polynômes irréductibles unitaires.
Récurrence établie.
□

Remarque On précisera par la suite quels sont exactement les polynômes irréductibles de $\mathbb{R}[X]$ et $\mathbb{C}[X]$.

8.4 Racines d’un polynôme

8.4.1 Racines et degré

Soient $P \in \mathbb{K}[X]$ et $a \in \mathbb{K}$. On a vu précédemment que a est racine de P (i.e. $P(a) = 0$) si, et seulement si, le polynôme $X - a$ divise P.

Proposition

Si $a_1, ..., a_n$ sont des racines deux à deux distinctes de $P \in \mathbb{K}[X]$ alors

$$(X - a_1)...(X - a_n) \mid P$$

dém. :
Les polynômes $X - a_1, ..., X - a_n$ sont des diviseurs de P deux à deux premiers entre eux dont leur produit divise encore P.
□

Théorème

Si P est un polynôme non nul alors P ne peut avoir plus de racines que son degré.

dém. :
Soient $P \in \mathbb{K}[X]$ tel que $P \neq 0$ et $a_1, ..., a_n$ ses racines distinctes.
On a $(X - a_1)...(X - a_n) \mid P$ et $P \neq 0$ donc $n \leqslant \deg P$.
□

Corollaire

Un polynôme de degré n possède au plus n racines.
Une équation algébrique de degré $n \in \mathbb{N}$ possède au plus n solutions.

Corollaire

Soit $P \in \mathbb{K}_n[X]$.
Si P possède au moins $n + 1$ racines alors $P = 0$.

Exemple Soient $P, Q \in \mathbb{K}_n[X]$ prenant la même valeur en $n + 1$ points distincts. Montrons $P = Q$.
Considérons pour cela $R = P - Q$.
$R \in \mathbb{K}_n[X]$ et R possède au moins $n + 1$ racines (correspondant aux points où P et Q prennent la même valeur) donc R est le polynôme nul. Par suite $P = Q$.

Corollaire

Si $P \in \mathbb{K}[X]$ possède une infinité de racines alors $P = 0$.

dém. :
Si P possède une infinité de racine, il en a plus que son degré...
□

Exemple Soit $P \in \mathbb{R}[X]$. Si $\int_0^1 P(t)^2 \, \mathrm{d}t = 0$ alors $P = 0$.

La fonction $t \mapsto P(t)^2$ est continue et positive sur $[0,1]$ donc, si $\int_0^1 P(t)^2 \, \mathrm{d}t = 0$, c'est la fonction nulle et le polynôme P possède alors une infinité de racines et c'est donc le polynôme nul.

Exemple Soit $P \in \mathbb{K}[X]$ tel que $P(X+1) = P(X)$. Montrons que P est constant.
Considérons $Q(X) = P(X) - P(0)$.
On a $Q(X+1) = Q(X)$. Or $Q(0) = 0$ donc par récurrence $Q(n) = 0$ pour tout $n \in \mathbb{N}$.
Le polynôme Q possède une infinité de racine, c'est donc le polynôme nul et par suite le polynôme P est constant (égal à $P(0)$).

Exemple Montrons qu'il existe un unique $P \in \mathbb{R}[X]$ tel que

$$\forall x \in \mathbb{R}, P(\sin x) = \sin 3x$$

Existence : on sait $\sin 3x = 3\sin x - 4\sin^3 x$.
Par suite le polynôme $P = 3X - 4X^3$ convient.
Unicité : Supposons P, Q solutions.
Pour tout $x \in \mathbb{R}$,

$$P(\sin x) = \sin(3x) = Q(\sin x)$$

donc pour tout $t \in [-1,1]$, $P(t) = Q(t)$. Le polynôme $P - Q$ possède une infinité de racines donc $P = Q$.

8.4.2 Polynôme et fonction polynomiale

Soit $\mathcal{D}$ une partie infinie de $\mathbb{K}$.

Définition

On note $\mathcal{P}(\mathcal{D}, \mathbb{K})$ l'ensemble des fonctions polynomiales de $\mathcal{D}$ vers $\mathbb{K}$.

Théorème

L'application qui à P associe $\tilde{P}$ est un isomorphisme du $\mathbb{K}$-espace vectoriel $\mathbb{K}[X]$ vers $\mathcal{P}(\mathcal{D}, \mathbb{K})$.

dém. :
Notons $\varphi : \mathbb{K}[X] \to \mathcal{P}(\mathcal{D}, \mathbb{K})$ l'application considérée.
On vérifie aisément la linéarité $\varphi(\lambda P + \mu Q) = \lambda\varphi(P) + \mu\varphi(Q)$.

Déterminons $\ker \varphi$.
Si $\varphi(P) = 0$ alors

$$\forall x \in \mathcal{D}, \tilde{P}(x) = 0$$

P possède donc une infinité de racine et donc $P = 0$.
Enfin, par définition, $\mathcal{P}(\mathcal{D}, \mathbb{K}) = \mathrm{Im}\varphi$.
Ainsi φ est un isomorphisme d'espaces vectoriels.
□

Corollaire

Si deux fonctions polynomiales prennent les mêmes valeurs une infinité de fois, c'est qu'elles sont issues du même polynôme. Ainsi

$$(\forall x \in \mathcal{D}, a_0 + a_1 x + a_2 x^2 + \cdots = b_0 + b_1 x + b_2 x^2 + \cdots) \Rightarrow \forall n \in \mathbb{N}, a_n = b_n$$

C'est le principe d'identification des coefficients des fonctions polynomiales.

Remarque Suite à ce résultat il est fréquent d'identifier polynômes et fonctions polynomiales associés définies sur une partie $\mathcal{D}$ infinie.

8.4.3 Multiplicité des racines

Définition

Soient $P \in \mathbb{K}[X]$ non nul et $a \in \mathbb{K}$.
On appelle ordre de multiplicité de a en tant que racine de P le plus grand $\alpha \in \mathbb{N}$ tel que $(X - a)^\alpha \mid P$.
Si $\alpha = 0$ alors a n'est pas racine de P.
Si $\alpha = 1$ on dit que a est une racine simple.
Si $\alpha \geqslant 2$ on dit que a est une racine multiple (double, triple,...)

Convention Si $P = 0$ alors tout $a \in \mathbb{K}$ est dit racine de multiplicité $+\infty$ de P.

Exemple Si $(X - a)^\alpha \mid P$ alors a est racine de multiplicité au moins α de P.

Proposition

Soient $P \in \mathbb{K}[X]$ non nul, $a \in \mathbb{K}$ et $\alpha \in \mathbb{N}$.
On a équivalence entre :
(i) a est racine de multiplicité α de P ;
(ii) $(X - a)^\alpha \mid P$ et $(X - a)^{\alpha+1}$ ne divise pas P ;
(iii) il existe $Q \in \mathbb{K}[X]$ tel que $P = (X - a)^\alpha Q$ et $Q(a) \neq 0$.

dém. :
(i) $\Rightarrow$ (ii) Si a est racine de multiplicité α de P alors α est le plus grand naturel tel que $(X - a)^\alpha$ divise P. Par suite $(X - a)^\alpha$ divise P et $(X - a)^{\alpha+1}$ ne divise pas P.
(ii) $\Rightarrow$ (iii) Supposons (ii).
Puisque $(X - a)^\alpha$ divise P, on peut écrire $P = (X - a)^\alpha Q$.
Si $Q(a) = 0$ alors $(X - a) \mid Q$ et donc $(X - a)^{\alpha+1}$ divise P, ce qui est exclu. Par suite $Q(a) \neq 0$.
(iii) $\Rightarrow$ (i) Supposons $P = (X - a)^\alpha Q$ avec $Q(a) \neq 0$.

On a immédiatement $(X-a)^\alpha \mid P$.
Pour $\beta \in \mathbb{N}^\star$, si $(X-a)^{\alpha+\beta}$ divise $P=(X-a)^\alpha Q$ alors $(X-a)^\beta$ divise Q et donc $Q(a)=0$ ce qui exclu.
Par suite, pour tout $\beta \in \mathbb{N}^\star$, $(X-a)^{\alpha+\beta}$ ne divise pas P. Ainsi α est le plus grand naturel tel que $(X-a)^\alpha$ divise P.
□

Exemple Soient $a,b,c \in \mathbb{K}$ deux à deux distincts et $P=(X-a)(X-b)^2(X-c)^3$.
Par ce qui précède on peut directement affirmer :
- a est racine simple de P ;
- b est racine double de P ;
- c est racine triple de P.

Proposition

Si $a_1,...,a_n$ sont des racines deux à deux distinctes de $P \in \mathbb{K}[X]$ de multiplicités respectives au moins égales à $\alpha_1,...,\alpha_n$ alors

$$(X-a_1)^{\alpha_1}...(X-a_n)^{\alpha_n} \mid P$$

dém. :
Les polynômes $(X-a_i)^{\alpha_i}$ sont des diviseurs de P deux à deux premiers entre eux.
□

Théorème

Si P est un polynôme non nul alors la somme des multiplicités de ses racines ne peut excéder son degré.

dém. :
Si $a_1,\dots,a_n$ sont les deux à deux distinctes de P de multiplicités respectives $\alpha_1,\dots,\alpha_n$ alors $(X-a_1)^{\alpha_1}...(X-a_n)^{\alpha_n} \mid P$ et puisque $P \neq 0$ on obtient $\alpha_1+...+\alpha_n \leqslant \deg P$.
□

Corollaire

Si la somme des multiplicités des racines de $P \in \mathbb{K}_n[X]$ est au moins égale à $n+1$ alors $P=0$.

Exemple Montrons qu'il existe un unique polynôme P tel que :
1) $\deg P = 4$;
2) P est unitaire ;
3) 0 est racine double et $1,-1$ racines simples.
Existence : Le polynôme $P=X^2(X-1)(X+1)$ est évidemment solution.
Unicité : Soient P et Q deux polynômes solutions.
Puisque P et Q sont unitaires et de degré 4 on a $\deg(P-Q) \leqslant 3$.
0 est racines double de P et de Q donc 0 est racine au moins double de $P-Q$.
De plus 1 et -1 sont racines au moins simples de $P-Q$.
La somme des multiplicités des racines de $P-Q$ vaut au moins $4>3$ donc $P-Q=0$ puis $P=Q$.

Corollaire

Soit P un polynôme de degré $n \in \mathbb{N}$.
Si P admet au moins n racines distinctes alors il n'y en n'a pas d'autres et celles-ci sont toutes simples.

dém. :
Notons $a_1, \ldots, a_n$ les n racines distinctes de P.
On a $(X - a_1) \ldots (X - a_n) \mid P$ et $\deg(X - a_1) \ldots (X - a_n) = \deg P$ donc $(X - a_1) \ldots (X - a_n)$ et P sont associés.
On en déduit que $a_1, \ldots, a_n$ sont racines simples de P.
□

Exemple Les racines n-ième l'unité sont des racines simples de $X^n - 1 \in \mathbb{C}[X]$.
En effet, il y a exactement n racines $n$ème de l'unité.

8.4.4 Multiplicité et dérivation

Théorème

Si a est racine de multiplicité $\alpha \in \mathbb{N}^\star$ de P alors a est racine de multiplicité $\alpha - 1$ de P'.

dém. :
On peut écrire $P = (X - a)^\alpha Q$ avec $Q(a) \neq 0$.
En dérivant $P' = (X - a)^{\alpha-1}\tilde{Q}$ avec $\tilde{Q} = \alpha Q + (X - a)Q'$.
Puisque $\tilde{Q}(a) = \alpha Q(a) \neq 0$, a est racine de multiplicité exactement $\alpha - 1$ de P'.
□

Corollaire

Les racines de P sont simples si, et seulement si, P et P' n'ont pas de racines communes.

dém. :
Les racines communes à P et P' sont exactement les racines multiples de P.
□

Exemple Montrons que les racines de $P = X^3 + 3X + 1 \in \mathbb{C}[X]$ sont simples.
Pour cela calculons son polynôme dérivé $P' = 3(X^2 + 1)$.
Les racines de P' sont i et $-i$. Celles-ci ne sont pas racines de P donc les racines de P sont simples.

Théorème

Soit $P \in \mathbb{K}[X]$ non nul, $a \in \mathbb{K}$ et $\alpha \in \mathbb{N}^\star$.
On a équivalence entre :
(i) a est racine d'ordre de multiplicité α de P ;
(ii) $P(a) = P'(a) = ... = P^{(\alpha-1)}(a) = 0$ et $P^{(\alpha)}(a) \neq 0$.

dém. :
(i) $\Rightarrow$ (ii) Si a est racine de multiplicité α de P alors par la proposition ci-dessus a est racine de P', de P'',..., et de $P^{(\alpha-1)}$ mais pas de $P^{(\alpha)}$.
(ii) $\Rightarrow$ (i) Supposons (ii)
Par la formule de Taylor :

$$P = \sum_{n=0}^{+\infty} \frac{P^{(n)}(a)}{n!}(X - a)^n = \sum_{n=\alpha}^{+\infty} \frac{P^{(n)}(a)}{n!}(X - a)^n = (X - a)^\alpha Q$$

avec

$$Q = \sum_{n=\alpha}^{+\infty} \frac{P^{(n)}(a)}{n!}(X-a)^{n-\alpha} \in \mathbb{K}[X]$$

De plus $Q(a) = \dfrac{P^{(\alpha)}(a)}{\alpha!} \neq 0$ donc a est racine de multiplicité exactement α de P.
□

Exemple Montrons

$$\forall n \in \mathbb{N}^\star, (X-1)^3 \mid nX^{n+2} - (n+2)X^{n+1} + (n+2)X - n$$

Posons $P_n = nX^{n+2} - (n+2)X^{n+1} + (n+2)X - n$.
On a

$$P_n(1) = n - (n+2) + (n+2) - n = 0$$

$$P_n'(1) = n(n+2) - (n+2)(n+1) + (n+2) = 0$$

et

$$P_n''(1) = n(n+2)(n+1) - (n+2)(n+1)n = 0$$

donc 1 est racine au moins triple de P_n d'où la divisibilité affirmée.

8.5 Polynômes scindés

8.5.1 Définition

Définition

Un polynôme $P \in \mathbb{K}[X]$ non constant est dit scindé dans $\mathbb{K}[X]$ si l'on peut écrire :

$$P = \lambda(X - x_1)\dots(X - x_n) \text{ avec } \lambda \in \mathbb{K}^\star \text{ et } x_1, \dots, x_n \in \mathbb{K}$$

Remarque Si tel est le cas :
- λ est le coefficient dominant de P ;
- n est le degré de P ;
- $x_1, \dots, x_n$ sont ses racines.

Exemple $X^2 + 1$ est scindé dans $\mathbb{C}[X]$ puisque $X^2 + 1 = (X - i)(X + i)$.
$X^2 + 1$ n'est pas scindé dans $\mathbb{R}[X]$ car ce polynôme n'a pas de racines dans $\mathbb{R}$.

Théorème

Soit $P \in \mathbb{K}[X]$ un polynôme non constant.
On a équivalence entre :
(i) P est scindé dans $\mathbb{K}[X]$;
(ii) la somme des multiplicités des racines de P égale son degré.

dém. :
(ii) $\Rightarrow$ (i) Supposons (ii)
Notons $a_1, ..., a_p$ les racines deux à deux distinctes de P et $\alpha_1, \dots, \alpha_p$ leurs multiplicités respectives.
On a $(X - a_1)^{\alpha_1} \dots (X - a_p)^{\alpha_p} \mid P$ et par (ii) $\deg(X - a_1)^{\alpha_1} \dots (X - a_p)^{\alpha_p} = \deg P$ donc P et $(X - a_1)^{\alpha_1} \dots (X - a_p)^{\alpha_p}$ sont associés et par suite P est scindé.
(i) $\Rightarrow$ (ii) Supposons P scindé. On peut écrire $P = \lambda(X - x_1) \dots (X - x_n)$.
En regroupant entre eux les x_i non distincts, on peut écrire :

$$P = \lambda(X - a_1)^{\alpha_1} \dots (X - a_p)^{\alpha_p}$$

avec $a_1, ..., a_p$ deux à deux distincts et $\alpha_1, \dots, \alpha_p \in \mathbb{N}^\star$ (α_i correspond au nombre d'occurrence de a_i dans la liste $x_1, \dots, x_n$).
Les $a_1, ..., a_p$ sont alors exactement les racines de P et leurs multiplicités respectives sont les $\alpha_1, ..., \alpha_p$.
L'égalité $P = \lambda(X - a_1)^{\alpha_1} \dots (X - a_p)^{\alpha_p}$ entraîne alors $\alpha_1 + \dots + \alpha_p = \deg P$ d'où (ii).
□

Remarque Si $P = \lambda(X - x_1) \dots (X - x_n)$ alors les racines de P sont les $x_1, ..., x_n$ comptées avec multiplicité (i.e. que chacune apparaît dans la liste autant de fois qu'elle est racine multiple de P).

8.5.2 Polynôme complexe

8.5.2.1 Le théorème de d'Alembert-Gauss

Théorème

> Tout polynôme non constant de $\mathbb{C}[X]$ admet au moins une racine.
> On dit que $\mathbb{C}$ algébriquement clos.

dém. :
Soit $P \in \mathbb{C}[X]$ de degré $p \in \mathbb{N}^\star$. $P = a_p X^p + \dots + a_1 X + a_0$ avec $a_p \neq 0$.
$A = \{|P(z)| / z \in \mathbb{C}\}$ est une partie non vide et minorée (par 0) de $\mathbb{R}$ donc $\alpha = \inf A = \inf\limits_{z \in \mathbb{C}} |P(z)|$ existe.
Nous allons montrer que α est une valeur prise par P puis que $\alpha = 0$. Ceci permettra alors de conclure.
Pour tout $z \in \mathbb{C}$ tel que $|z| \geqslant 1$, on a

$$a_p z^p = P(z) + \sum_{k=0}^{p-1} -a_k z^k$$

donc

$$|a_p z^p| \leqslant |P(z)| + \sum_{k=0}^{p-1} |a_k| \left|z^k\right| \leqslant |P(z)| + pM |z|^{p-1}$$

en notant $M = \max\limits_{0 \leqslant k \leqslant p-1} |a_k|$.
Par suite

$$|P(z)| \geqslant |a_p| |z|^p - pM |z|^{p-1} \xrightarrow[|z| \to +\infty]{} +\infty$$

Il en découle que l'ensemble des $z \in \mathbb{C}$ tel que $|P(z)| \leqslant \alpha + 1$ est borné.
Considérons maintenant (z_n) suite complexe telle que $P(z_n) \to \alpha$. Soulignons qu'une telle suite existe par le principe de réalisation d'une borne inférieure et qu'elle est bornée en vertu de l'étude ci-dessus.

Par le théorème de Bolzano-Weierstrass, on peut extraire de (z_n) une suite convergente $(z_{\varphi(n)})$.
Posons ω la limite de cette dernière.
Par opérations sur les limites

$$P(z_{\varphi(n)}) = a_p z_{\varphi(n)}^p + \cdots + a_1 z_{\varphi(n)} + a_0 \to a_p \omega^p + \cdots + a_1 \omega + a_0 = P(\omega)$$

D'autre part, par extraction, $P(z_{\varphi(n)}) \to \alpha$ donc par unicité de la limite $P(\omega) = \alpha$.
Ainsi nous venons d'établir que α est une valeur prise par P.
Il ne reste plus qu'à montrer $\alpha = 0$ et pour cela nous allons raisonner par l'absurde en supposant $\alpha > 0$.
Quitte à considérer le polynôme $\dfrac{P(X+\omega)}{\alpha}$, on peut supposer $\alpha = 1$ et $\omega = 0$.
On a alors $P = a_0 + a_1 X + \cdots + a_p X^p$ avec $a_0 = 1 = \min\limits_{z \in \mathbb{C}} |P(z)|$, $p > 1$ et $a_p \neq 0$.
En introduisant q le premier indice tel que $a_q \neq 0$, on obtient

$$P = 1 + a_q X^q + a_{q+1} X^{q+1} + \cdots + a_p X^p$$

Ecrivons $a_q = \rho \mathrm{e}^{i\theta}$ avec $\rho > 0$ et $\theta \in \mathbb{R}$.
Pour $z = r\mathrm{e}^{i(\theta+\pi)/q}$ avec $r > 0$, on a

$$P(z) = 1 - \rho r^q + \sum_{k=q+1}^{p} a_k r^k \mathrm{e}^{ik(\theta+\pi)/q}$$

et donc

$$|P(z)| \leqslant |1 - \rho r^q| + \sum_{k=q+1}^{p} |a_k| \, r^k$$

Pour r suffisamment petit, on obtient

$$|P(z)| \leqslant 1 - \rho r^q + \sum_{k=q+1}^{p} |a_k| \, r^k = 1 + f(r)$$

avec

$$f(r) = -\rho r^q + \sum_{k=q+1}^{p} |a_k| \, r^k \underset{r \to 0}{\sim} -\rho r^q < 0$$

Par suite, il existe des $z \in \mathbb{C}$ tel que

$$|P(z)| < 1 = \min_{z \in \mathbb{C}} |P(z)|$$

Absurde !
□

Corollaire

Les polynôme irréductibles de $\mathbb{C}[X]$ sont ceux de degré 1.

dém. :
Les polynômes de degré 1 sont irréductibles dans $\mathbb{C}[X]$.
Inversement, si P est irréductible dans $\mathbb{C}[X]$ alors P est non constant et possède donc une racine a. On a alors $X - a \mid P$ puis P associé à $X - a$.
□

8.5.2.2 Décomposition en facteurs irréductibles dans $\mathbb{C}[X]$

Théorème

Soit $P \in \mathbb{C}[X]$ non constant. Il existe $\lambda \in \mathbb{K}^\star$, $N \in \mathbb{N}^\star$, $a_1, \ldots, a_N \in \mathbb{C}$ deux à deux distincts et $\alpha_1, ..., \alpha_N \in \mathbb{N}^\star$ tels que

$$P = \lambda(X - a_1)^{\alpha_1} ...(X - a_N)^{\alpha_N}$$

De plus cette décomposition est unique à l'ordre près des facteurs.

dém. :
C'est la décomposition en facteurs irréductibles qui est ici particularisée sachant la forme des polynômes irréductibles dans $\mathbb{C}[X]$.
□

Remarque Avec une telle décomposition :
- λ est le coefficient dominant de P ;
- $a_1, ..., a_N$ sont les racines deux à deux distinctes de P ;
- $\alpha_1, ..., \alpha_N$ sont les multiplicités respectives des précédentes racines de P.

Corollaire

Tout polynôme non constant de $\mathbb{C}[X]$ est scindé.

Corollaire

Tout polynôme de $\mathbb{C}[X]$ de degré $n \in \mathbb{N}$ possède exactement n racines comptées avec multiplicité.
Toute équation algébrique complexe de degré $n \in \mathbb{N}^\star$ possède n solutions complexes comptées avec multiplicité.

Exemple Soit $a \in \mathbb{R}$. Factorisons dans $\mathbb{C}[X]$ le polynôme

$$X^2 - 2X \cos a + 1$$

Il suffit de détermines les racines de ce polynôme, ce sont e^{ia} et e^{-ia}.
On en déduit

$$X^2 - 2X \cos a + 1 = (X - \mathrm{e}^{ia})(X - \mathrm{e}^{-ia})$$

Exemple Factorisons dans $\mathbb{C}[X]$ le polynôme

$$X^n - 1$$

Les racines de ce polynômes sont les racines n-ième de l'unité, on sait que ce sont des racines simples.
Par suite

$$X^n - 1 = \prod_{k=0}^{n-1} (X - \mathrm{e}^{2ik\pi/n})$$

Exemple Factorisons dans $\mathbb{C}[X]$ le polynôme

$$X^4 + X^2 + 1$$

Posons $Y = X^2$.

$$X^4 + X^2 + 1 = Y^2 + Y + 1 = (Y - j)(Y - j^2) = (X^2 - j)(X^2 - j^2)$$

Or

$$(X^2 - j) = (X^2 - j^4) = (X - j^2)(X + j^2)$$

et

$$(X^2 - j^2) = (X - j)(X + j)$$

Par suite

$$X^4 + X^2 + 1 = (X - j)(X + j)(X - j^2)(X + j^2)$$

8.5.2.3 Arithmétique et racines

Proposition

Soient $A, B \in \mathbb{C}[X]$. On a équivalence entre :
(i) $A \mid B$;
(ii) les racines de A sont racines de B de multiplicité au moins égale.

dém. :
Si A est constant, l'équivalence est immédiate.
Sinon on peut écrire $A = \lambda(X - a_1)^{\alpha_1} \dots (X - a_n)^{\alpha_n}$ avec $a_1, \dots, a_n$ les racines de A et $\alpha_1, \dots, \alpha_n$ leurs multiplicités respectives.
(i) $\Rightarrow$ (ii) Si $A \mid B$ alors $(X - a_1)^{\alpha_1} \dots (X - a_n)^{\alpha_n} \mid B$ et donc les $a_1, \dots, a_n$ sont racines de B de multiplicités respectivement au moins égales à $\alpha_1, \dots, \alpha_n$.
(ii) $\Rightarrow$ (i) Si les $a_1, \dots, a_n$ sont racines de B de multiplicité au moins égales à $\alpha_1, \dots, \alpha_n$ alors $(X - a_1)^{\alpha_1} \dots (X - a_n)^{\alpha_n} \mid B$ et donc $A \mid B$.
□

Exemple On a $X^4 + X^2 + 1 \mid X^{18} - 1$.
En effet les racines de $X^4 + X^2 + 1$ sont $j, -j, j^2$ et $-j^2$. Ce sont des racines simples, toutes racines de $X^{18} - 1$.

Proposition

Soient $A, B \in \mathbb{C}[X]$. On a équivalence entre :
(i) $A \wedge B = 1$;
(ii) A et B n'ont pas de racines en commun.

dém. :
(i) $\Rightarrow$ (ii) Par contraposée.
Supposons que A et B ont une racine en commun a.
On a alors $X - a$ diviseur commun à A et B et donc A et B ne sont pas premiers entre eux.
(ii) $\Leftarrow$ (i) Par contraposée.
Supposons que A et B ne sont pas premiers entre eux.
Posons alors $D = \text{pgcd}(A, B)$.

Ce polynôme D est nul ou bien non constant.
Dans les deux cas D possède une racine qui sera racine commune à A et B.
□

Exemple Montrons $(X^2 + 1) \wedge (X^3 + X + 1) = 1$.
Les racines de $X^2 + 1$ sont i et $-i$.
Celles-ci ne sont pas racines de $X^3 + X + 1$.
Les polynômes $X^2 + 1$ et $X^3 + X + 1$ n'ont pas de racines complexes en commun, ils sont donc premiers entre eux.

Exemple Les racines de $P \in \mathbb{C}[X]$ sont simples si, et seulement si, $P \wedge P' = 1$.
En effet, les racines de P sont simples si, et seulement si, il n'y a pas de racines communes à P et P'.

8.5.2.4 Polynôme conjugué

Définition

On appelle polynôme conjugué de $P = a_n X^n + ... + a_1 X + a_0 \in \mathbb{C}[X]$ le polynôme

$$\bar{P} = \bar{a}_n X^n + ... + \bar{a}_1 X + \bar{a}_0 \in \mathbb{C}[X]$$

Proposition

$\forall P, Q \in \mathbb{C}[X]$,

$$\bar{\bar{P}} = P, \overline{P + Q} = \bar{P} + \bar{Q}, \overline{PQ} = \bar{P}\bar{Q}$$

dém. :
C'est immédiat en décrivant les polynômes par leurs coefficients.
□

Proposition

$\forall P, Q \in \mathbb{C}[X]$,

$$P \mid Q \Leftrightarrow \bar{P} \mid \bar{Q}$$

dém. :
Si $P \mid Q$ alors on peut écrire $Q = PU$ avec $U \in \mathbb{C}[X]$.
On alors $\bar{Q} = \overline{PU} = \bar{P}\bar{U}$ donc $\bar{P} \mid \bar{Q}$.
Si $\bar{P} \mid \bar{Q}$ alors $P = \bar{\bar{P}}$ divise $Q = \bar{\bar{Q}}$.
□

Proposition

$\forall P \in \mathbb{C}[X], \forall a \in \mathbb{C}$,

$$\overline{P(a)} = \bar{P}(\bar{a})$$

dém. :
C'est immédiat en décrivant le polynôme P par ses coefficients.
□

Proposition

Soit $P \in \mathbb{C}[X]$, $a \in \mathbb{C}$ et $\alpha \in \mathbb{N}$.
On a équivalence entre :
(i) a est racine de multiplicité α de P ;
(ii) $\bar{a}$ est racine de multiplicité α de $\bar{P}$.

dém. :
En effet

$$\forall k \in \mathbb{N}, (X-a)^k \mid P \Leftrightarrow (X-\bar{a})^k \mid \bar{P}$$

Ainsi la multiplicité de a en tant que racine de P égale celle de $\bar{a}$ en tant que racine de $\bar{P}$.
□

8.5.3 Polynôme réel

8.5.3.1 Racine complexe

Remarque Tout polynôme réel peut se voir comme un polynôme complexe.

Définition

On appelle racine complexe d'un polynôme $P \in \mathbb{R}[X]$ toute racine de P vu comme polynôme complexe.

Exemple j est une racine complexe du polynôme réel $X^2 + X + 1$.

Proposition

Soit $P \in \mathbb{R}[X]$ un polynôme de degré $n \in \mathbb{N}$.
P admet exactement n racines complexes comptées avec multiplicité.

dém. :
C'est la transposition avec le vocabulaire en cours du résultat assurant que tout polynôme complexe de degré n admet exactement n racines comptées avec multiplicité.
□

Proposition

Les racines complexes de $P \in \mathbb{R}[X]$ sont deux à deux conjuguées et deux racines conjuguées ont même multiplicité.

dém. :
Si a est racine complexe de multiplicité α de P alors $\bar{a}$ est aussi racine complexe de multiplicité α de $\bar{P} = P$.
□

8.5.3.2 Décomposition en facteurs irréductibles dans $\mathbb{R}[X]$

Proposition

Les polynômes irréductibles de $\mathbb{R}[X]$ sont :
- les polynômes de degré 1 ;
- les polynômes de degré 2 de discriminant < 0 (i.e. sans racines réelles).

dém. :
Ce sont des polynômes irréductibles :
En effet, on sait déjà que les polynômes de degré 1 sont irréductibles.
Considérons maintenant un polynôme P de degré 2 sans racines réelles.
Soit D un diviseur de P. On a $\deg D \leqslant 2$ et $D \neq 0$.
Si $\deg D = 0$ ou 2 alors D est un diviseur trivial de P.
Si $\deg D = 1$ alors D possède une racine réelle qui sera racine de P. C'est exclu.
Ainsi P est irréductible.
Inversement :
Soit P un polynôme irréductible dans $\mathbb{R}[X]$.
P est non constant donc P possède une racine complexe a.
Si $a \in \mathbb{R}$ alors $X - a \mid P$ et donc P est associé à $X - a$ puis c'est un polynôme de degré 1.
Si $a \in \mathbb{C}\backslash\mathbb{R}$ alors $a, \bar{a}$ sont racines distinctes de P puis $A = (X - a)(X - \bar{a})$ divise P dans le cadre des polynômes complexes. Ainsi il existe $B \in \mathbb{C}[X]$ tel que $P = AB$.
Or $P \in \mathbb{R}[X]$ et $A = X^2 - 2\text{Re}(a)X + |a|^2 \in \mathbb{R}[X]$ donc $\bar{P} = \bar{A}\bar{B}$ donne $P = A\bar{B}$.
On en déduit $B = \bar{B}$ d'où $B \in \mathbb{R}[X]$. Ainsi A divise P dans le cadre des polynômes réels.
Or P est irréductible, il est donc associé à A et apparaît comme étant un polynôme de degré 2 sans racines réelles.
□

Exemple Le polynôme $X^2 + 2X + 2$ est irréductible dans $\mathbb{R}[X]$.
Les polynômes $X^2 - 3X + 2$ et $X^4 + X^2 + 1$ sont composés dans $\mathbb{R}[X]$.

Théorème

Soit $P \in \mathbb{R}[X]$ non constant. Il existe $\lambda \in \mathbb{R}^\star$, $N, M \in \mathbb{N}$, $a_1, \dots, a_N \in \mathbb{R}$ deux à deux distincts,
$(p_1, q_1), \dots, (p_M, q_M) \in \mathbb{R}^2$ deux à deux distincts, tels que $\Delta_j = p_j^2 - 4q_j < 0$ et il existe $\alpha_1, \dots, \alpha_N, \beta_1, \dots, \beta_M \in \mathbb{N}^\star$ tels que :

$$P = \lambda \prod_{i=1}^{N} (X - a_i)^{\alpha_i} \prod_{j=1}^{M} (X^2 + p_j X + q_j)^{\beta_j}$$

De plus cette décomposition est unique à l'ordre près des facteurs.

dém. :
C'est la décomposition en facteurs irréductibles qui est ici particularisée sachant la forme des polynômes irréductibles dans $\mathbb{R}[X]$.
□

Remarque Avec une telle décomposition
- λ est le coefficient dominant de P ;
- $a_1, ..., a_N$ sont les racines réelles de multiplicité $\alpha_1, ..., \alpha_N$ de P ;
- les racines complexes conjuguées de P se retrouvent dans les termes $X^2 + p_j X + q_j$, leur multiplicité étant alors β_j.

Corollaire

Les polynômes de degré impair possèdent au moins une racine réelle.

dém. :
Si $N = 0$ dans l'écriture qui précède alors P est un polynôme de degré pair.
□

Exemple Factorisons dans $\mathbb{R}[X]$ le polynôme

$$X^5 - 1$$

Les racines complexe de $X^5 - 1$ sont $\omega_0, \omega_1, ..., \omega_4$ avec $\omega_k = \mathrm{e}^{2ik\pi/5}$.
$\omega_0 = 1$ est réel, $\omega_4 = \bar{\omega}_1$ et $\omega_3 = \bar{\omega}_2$.
Par suite

$$X^5 - 1 = (X-1)\left((X-\omega_1)(X-\bar{\omega}_1)\right)\left((X-\omega_2)(X-\bar{\omega}_2)\right)$$

ce qui donne

$$X^5 - 1 = (X-1)(X^2 - 2\cos\frac{2\pi}{5}X + 1)(X^2 - 2\cos\frac{4\pi}{5}X + 1)$$

Exemple Factorisation dans $\mathbb{R}[X]$ le polynôme

$$X^6 - 1$$

Comme ci-dessus, en isolant les racines réelles et en regroupant entre elles les racines complexes conjuguées, on obtient

$$X^6 - 1 = (X-1)(X+1)(X^2+X+1)(X^2-X+1)$$

8.5.4 Relations entre racines et coefficients d'un polynôme scindé

Soit $P = a_nX^n + \cdots + a_1X + a_0$ un polynôme scindé de degré $n \in \mathbb{N}^\star$ réel ou complexe.
Notons $x_1, ..., x_n$ ses racines de P comptées avec multiplicité.
On a $P = a_nX^n + \cdots + a_1X + a_0 = a_n(X - x_1)\ldots(X - x_n)$.
En développant le second membre, on peut exprimer les coefficients de P en fonction de ses racines ; explicitons ces expressions.

Définition

Pour $n \in \mathbb{N}^\star$ et $x_1, \ldots, x_n \in \mathbb{K}$, on appelle expressions symétriques élémentaires en les $x_1, ..., x_n$ les quantités suivantes :

$$\sigma_1 = \sum_{i=1}^{n} x_i,\ \sigma_2 = \sum_{1\leqslant i<j\leqslant n} x_ix_j,\ \sigma_3 = \sum_{1\leqslant i<j<k\leqslant n} x_ix_jx_k, \ldots$$

$$\sigma_p = \sum_{1\leqslant i_1<i_2<\ldots<i_p\leqslant n} x_{i_1}x_{i_2}...x_{i_p} \text{ (pour } 1 \leqslant p \leqslant n\text{)}, \ldots$$

$$\sigma_n = x_1x_2...x_n.$$

Remarque Ainsi σ_p apparaît comme étant la somme de tous les produits possibles de p éléments d'indices distincts choisis dans $x_1, ..., x_n$; on dit que σ_p est la somme des p-produits.

Exemple Pour $n = 2$

$$\sigma_1 = x_1 + x_2 \text{ et } \sigma_2 = x_1 x_2$$

Exemple Pour $n = 3$

$$\sigma_1 = x_1 + x_2 + x_3, \sigma_2 = x_1x_2 + x_1x_3 + x_2x_3 \text{ et } \sigma_3 = x_1x_2x_3$$

Proposition

Soient $n \in \mathbb{N}^\star$, $x_1, \dots, x_n \in \mathbb{K}$ et $\sigma_1, \dots, \sigma_n$ les expressions symétriques élémentaires en les $x_1, ..., x_n$.
On a

$$(X - x_1)\dots(X - x_n) = X^n - \sigma_1 X^{n-1} + \dots + (-1)^k \sigma_k X^{n-k} + \dots + (-1)^n \sigma_n$$

dém. :
Le coefficient de X^{n-k} dans le développement de $(X - x_1)\dots(X - x_n)$ s'obtient en considérant les termes obtenus lorsqu'on choisit k facteurs $-x_i$ et $n - k$ facteurs X. Ce coefficient vaut donc $(-1)^k \times$ la somme des produits de la forme $x_{i_1} x_{i_2} \dots x_{i_k}$ avec $i_1 < i_2 < \dots < i_k$ i.e. $(-1)^k \sigma_k$.
□

Théorème

Soient $P = a_n X^n + ... + a_1 X + a_0$ un polynôme de degré $n \in \mathbb{N}^\star$ et $x_1, ..., x_n \in \mathbb{K}$.
On a équivalence entre :
(i) $x_1, ..., x_n$ sont les racines de P comptées avec multiplicité ;
(ii) $\forall 1 \leqslant k \leqslant n, \sigma_k = (-1)^k a_{n-k}/a_n$.

dém. :
Sachant que le coefficient dominant de P est a_n, on a

$$\text{(i)} \Leftrightarrow P = a_n(X - x_1)...(X - x_n)$$

En développant ce produit, on obtient

$$\text{(i)} \Leftrightarrow a_n X^n + \dots + a_1 X + a_0 = a_n(X^n - \sigma_1 X^{n-1} + \dots + (-1)^k \sigma_k X^{n-k} + \dots + (-1)^n \sigma_n)$$

En identifiant les coefficients de ces polynômes

$$\text{(i)} \Leftrightarrow \forall 1 \leqslant k \leqslant n, a_{n-k} = a_n (-1)^k \sigma_k$$

et finalement on obtient l'équivalence de (i) et de (ii).
□

Remarque Le coefficient de X^{n-1} d'un polynôme scindé donne la somme de ses racines.
Le coefficient constant d'un polynôme scindé donne le produit de ses racines.

Exemple Le cas $n = 2$.
Soit $P = aX^2 + bX + c$ avec $a \neq 0$.
x_1 et x_2 sont les racines de P comptées avec multiplicité si, et seulement si,

$$\begin{cases} x_1 + x_2 = -b/a \\ x_1 x_2 = c/a \end{cases}$$

Remarque On peut résoudre un système de la forme

$$\begin{cases} x_1 + x_2 = \sigma_1 \\ x_1 x_2 = \sigma_2 \end{cases}$$

en déterminant les racines de $X^2 - \sigma_1 X + \sigma_2$.
Par exemple, le système

$$\begin{cases} x + y = 2 \\ xy = 7 \end{cases}$$

a pour couples solutions, les couples (x, y) formés des deux racines du polynôme $X^2 - 2X + 7$.

Exemple Le cas $n = 3$.
Soit $P = aX^3 + bX^2 + cX + d$ avec $a \neq 0$.
x_1, x_2, x_3 sont les racines de P comptées avec multiplicité si, et seulement si,

$$\begin{cases} x_1 + x_2 + x_3 = -b/a \\ x_1x_2 + x_2x_3 + x_3x_1 = c/a \\ x_1x_2x_3 = -d/a \end{cases}$$

Exemple Déterminons les racines de $P = X^3 - 4X^2 + 6X - 4 \in \mathbb{C}[X]$ sachant que la somme de deux des racines est égale à la troisième.
Notons x_1, x_2, x_3 les trois racines comptées avec multiplicité du polynôme P.
Quitte à les réindexer, on peut supposer $x_1 + x_2 = x_3$.
Les relations coefficients/racines donnent en particulier

$$x_1 + x_2 + x_3 = 4 \text{ et } x_1x_2x_3 = 4$$

Ainsi $2x_3 = 4$ d'où $x_3 = 2$ et

$$\begin{cases} x_1 + x_2 = 2 \\ x_1 x_2 = 2 \end{cases}$$

Les racines de $X^2 - 2X + 2$ étant $1 + i$ et $1 - i$, on peut dire que, à l'ordre près, x_1, x_2, x_3 valent $1 + i, 1 - i, 2$.

Remarque On peut résoudre un système de la forme

$$\begin{cases} x_1 + x_2 + x_3 = \sigma_1 \\ x_1x_2 + x_2x_3 + x_3x_1 = \sigma_2 \\ x_1x_2x_3 = \sigma_3 \end{cases}$$

en déterminant les racines de $X^3 - \sigma_1 X^2 + \sigma_2 X - \sigma_3$.
Cependant, comme il s'agit ici d'un polynôme de degré 3, une telle résolution n'est pas immédiate ; on pourra voir à ce sujet la méthode de Cardan à la fin de ce cours.

Exemple Résolvons le système

$$\begin{cases} x + y + z = 2 \\ xy + yz + zx = -5 \\ xyz = -6 \end{cases}$$

Les triplets (x, y, z) solutions de ce système sont ceux formés des racines comptées avec multiplicité du polynôme $X^3 - 2X^2 - 5X + 6$.
1 est racine apparente de ce polynôme et cela permet la factorisation :

$$X^3 - 2X^2 - 5X + 6 = (X - 1)(X^2 - X - 6) = (X - 1)(X + 2)(X - 3)$$

Les triplets (x, y, z) solutions sont $(1, -2, 3)$ et ses permutations.

Exemple Résolvons le système

$$\begin{cases} x_1 + x_2 + x_3 = 1 \\ x_1^2 + x_2^2 + x_3^2 = 9 \\ x_1^3 + x_2^3 + x_3^3 = 1 \end{cases}$$

Introduisons

$$S_1 = x_1 + x_2 + x_3,\ S_2 = x_1^2 + x_2^2 + x_3^2 \text{ et } S_3 = x_1^3 + x_2^3 + x_3^3$$

$$\sigma_1 = x_1 + x_2 + x_3,\ \sigma_2 = x_1x_2 + x_2x_3 + x_3x_1 \text{ et } \sigma_3 = x_1x_2x_3$$

Soit (x_1, x_2, x_3) un triplet solution, exprimons $\sigma_1, \sigma_2, \sigma_3$.
$\sigma_1 = S_1 = 1$. $S_1^2 = S_2 + 2\sigma_2$ donc $\sigma_2 = -4$. $S_1^3 = S_3 + 3t + 6\sigma_3$ et $S_1S_2 = S_3 + t$ avec

$$t = x_1^2x_2 + x_2x_1^2 + x_2^2x_3 + x_3^2x_2 + x_3^2x_1 + x_1^2x_3$$

On en déduit $\sigma_3 = -4$.
Par suite le triple (x_1, x_2, x_3) est formé des racines comptées avec multiplicité du polynôme

$$X^3 - X^2 - 4X + 4$$

Puisque 1 est racine apparent de ce polynôme, on peut facilement factoriser celui-ci :

$$(X^3 - X^2 - 4X + 4) = (X - 1)(X - 2)(X + 2)$$

On en déduit que, à l'ordre près, $x_1, x_2, x_3 = 1, 2, -2$.
Inversement, les triplets correspondants sont bien solutions du système posé.

Exemple Soient x_1, x_2, x_3 les trois racines comptées avec multiplicité de l'équation $x^3 + px + q = 0$. Calculons

$$S_2 = x_1^2 + x_2^2 + x_3^2, S_3 = x_1^3 + x_2^3 + x_3^3$$

Introduisons les quantités symétriques élémentaires
$\sigma_1 = x_1 + x_2 + x_3$, $\sigma_2 = x_1x_2 + x_2x_3 + x_3x_1$ et $\sigma_3 = x_1x_2x_3$.
On a $\sigma_1 = 0$, $\sigma_2 = p$ et $\sigma_3 = -q$.
Puisque a $\sigma_1^2 = S_2 + 2\sigma_2$, on a $S_2 = -2p$.
Puisque $x_i^3 = -px_i - q$, en sommant $S_3 = -p\sigma_1 - 3q = -3q$
Supposons de plus $q \neq 0$ et calculons

$$I_1 = \frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} \text{ et } I_2 = \frac{1}{x_1^2} + \frac{1}{x_2^2} + \frac{1}{x_3^2}$$

En réduisant au même dénominateur

$$I_1 = \frac{\sigma_2}{\sigma_3} = -\frac{p}{q}$$

De plus

$$I_1^2 = I_2 + \frac{2}{x_1x_2} + \frac{2}{x_2x_3} + \frac{2}{x_3x_1} = I_2 + \frac{2\sigma_1}{\sigma_3}$$

donc $I_2 = p^2/q^2$.

8.5.5 Musculation : Résolution de l'équation du troisième degré par la méthode de Cardan

Historiquement, la démarche suivie a été découverte par Tartaglia (1499-1557) mais celui-ci ne voulut pas la dévoiler afin de rester seul capable de résoudre les équations du troisième degré (cela constituait un jeu mathématique de l'époque...) Cardan le persuade de lui révéler sa méthode tout en promettant de ne pas la divulguer. Néanmoins il le fît, ce qui fâcha définitivement les deux hommes...
Soient $a, b, c \in \mathbb{C}$. On cherche ici à résoudre l'équation

$$\mathcal{E} : t^3 + at^2 + bt + c = 0$$

d'inconnue $t \in \mathbb{C}$.
1ère étape :
On réalise le changement d'inconnue $t = x - a/3$.
L'équation $\mathcal{E}$ se transforme alors en une équation du type :

$$x^3 + px + q = 0$$

d'inconnue $x \in \mathbb{C}$ où $p, q \in \mathbb{C}$ s'expriment en fonction de a, b, c.
Dans le cas où $p = 0$, on sait résoudre l'équation $x^3 = -q$:
- si $q \in \mathbb{R}$ les solutions sont $\sqrt[3]{-q}$, $\sqrt[3]{-q}.j$ et $\sqrt[3]{-q}.j^2$;
- si $q \in \mathbb{C}$ on écrit $q = \rho e^{i\theta}$ avec $\rho > 0$ et $\theta \in \mathbb{R}$, les solutions sont alors $\sqrt[3]{\rho}e^{i\theta/3}$, $\sqrt[3]{\rho}e^{i\theta/3}.j$ et $\sqrt[3]{\rho}e^{i\theta/3}.j^2$.
Supposons désormais $p \neq 0$.
2ème étape :
Soit x une solution. On écrit $x = u + v$ avec $u, v \in \mathbb{C}^\star$ tels que $uv = -p/3$. Une telle décomposition est

assurément possible car tout système somme-produit complexe possède un couple solution.
L'équation $x^3 + px + q = 0$ devient alors

$$(u+v)^3 + p(u+v) + q = 0$$

puis, après simplification,

$$u^3 + v^3 + q = 0$$

3ème étape :
En posant $U = u^3$ et $V = v^3$ on observe que (U, V) est solution du système

$$\begin{cases} U + V = -q \\ UV = -p^3/27 \end{cases}$$

On sait résoudre un tel système et ceci permet de choisir un couple (U, V).parmi les deux couples solutions. Ce choix n'a pas d'incidence compte tenu de la symétrie existant entre u et v, symétrie qui permet d'échanger u et v et donc U et V. Après résolution des équations $u^3 = U$ et $v^3 = V$, on obtient trois valeurs possibles pour u et trois valeurs possibles pour v soit, a priori, neuf couples (u, v) possibles. Cependant la relation $uv = -p/3$ permet d'exprimer v en fonction de u et, par suite, seuls trois des neufs couples (u, v) sont à considérer.
La relation $x = u + v$ permet alors d'exprimer x et ceci conduit alors à trois solutions possibles.
Résumons : Si x est solution de l'équation $x^3 + px + q = 0$ alors x est l'une des 3 solutions précédentes.
4ème étape :
Inversement, si x est l'une des trois valeurs précédentes, en remontant le raisonnement, on observe bien que x est solution de l'équation $x^3 + px + q = 0$. La résolution est achevée.

Exemple Résolvons l'équation $t^3 - 3t^2 - 3t - 4 = 0$ d'inconnue $t \in \mathbb{C}$.
On pose $x = t - 1$ ce qui conduit à l'équation $x^3 - 6x - 9 = 0$.
On écrit $x = u + v$ avec $uv = 2$.
On obtient alors $u^3 + v^3 = 9$.
On résout le système

$$\begin{cases} U + V = 9 \\ UV = 8 \end{cases}$$

en introduisant l'équation $y^2 - 9y + 8 = 0$ de racines 1 et 8.
On prend alors $U = 1, V = 8$ d'où $u = 1, j, j^2$ et $v = 2, 2j, 2j^2$.
La relation $uv = 2$ nous dit que les couples (u, v) à considérer sont $(1, 2), (j, 2j^2)$ et $(j^2, 2j)$ et les valeurs correspondantes de x sont $3, j(1 + 2j), j(j + 2)$ puis on obtient $t = 4, 1 + j(1 + 2j)$ ou $t = 1 + j(j + 2)$. Enfin ses valeurs sont bien solutions puisqu'on peut remonter le raisonnement.
Finalement

$$\mathcal{S} = \{4, 1 + j(1 + 2j), 1 + j(j + 2)\}$$

Chapitre 9

Les fractions rationnelles

$\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$.

9.1 Le corps des fractions rationnelles

9.1.1 Construction

Définition

On appelle fraction rationnelle à coefficients dans $\mathbb{K}$ et en l'indéterminée X tout élément représenté par un rapport A/B formé par $A, B \in \mathbb{K}[X]$ avec $B \neq 0$.
On note $\mathbb{K}(X)$ l'ensemble de ces éléments.

Exemple

$$\frac{X+1}{X} \in \mathbb{R}(X)$$

Définition

On dit deux rapport A/B et C/D (avec $B, D \neq 0$) représentent la même fraction rationnelle si $AD = BC$.
Ainsi

$$\frac{A}{B} = \frac{C}{D} \Leftrightarrow AD = BC$$

Exemple $\dfrac{X+1}{X^2-1} = \dfrac{1}{X-1}$ car $(X-1)(X+1) = (X^2-1) \times 1$.

Définition

Tout polynôme $P \in \mathbb{K}[X]$ est dit égal à la fraction rationnelle $P/1 \in \mathbb{K}(X)$.
En ce sens $\mathbb{K}[X] \subset \mathbb{K}(X)$.

Définition

Soient $F = A/B$ et $G = C/D$ deux éléments de $\mathbb{K}(X)$ et $\lambda \in \mathbb{K}$.
On définit les fractions rationnelles $\lambda.F, F+G, FG$ par :

$$\lambda.F = \frac{\lambda.A}{B}, F+G = \frac{AD+BC}{BD}, FG = \frac{AC}{BD}$$

Exemple Pour

$$F = \frac{X+1}{X-1} \text{ et } G = \frac{X^2-2X-1}{X(X+1)}$$

on a

$$F+G = \frac{2X^3-X^2+2X+1}{X(X^2-1)}$$

et

$$FG = \frac{X^2-2X-1}{X(X-1)}$$

Proposition

Les résultats de ces opérations sont indépendant des rapports A/B et C/D choisis pour représentés F et G.

dém. :
Supposons

$$F = \frac{A_1}{B_1} = \frac{A_2}{B_2} \text{ et } G = \frac{C_1}{D_1} = \frac{C_2}{D_2}$$

On a $A_1B_2 = A_2B_1$ et $C_1D_2 = C_2D_1$.
Puisque $\lambda A_1B_2 = \lambda A_2B_1$, on a $\dfrac{\lambda A_1}{B_1} = \dfrac{\lambda A_2}{B_2}$.
Ainsi le résultat de l'opération $\lambda.F$ ne dépend pas du rapport choisi pour représenter F.
On a $(A_1D_1 + B_1C_1)B_2D_2 = A_1B_2D_1D_2 + C_1D_2B_1B_2$
donc $(A_1D_1 + B_1C_1)B_2D_2 = A_2B_1D_1D_2 + C_2D_1B_1B_2$
puis $(A_1D_1 + B_1C_1)B_2D_2 = (A_1D_2 + B_2C_2)B_1D_1$.
Ainsi le résultat de l'opération $F+G$ ne dépend pas des rapports choisis pour représenter F et G.
Enfin $(A_1C_1)(B_2D_2) = (A_1B_2)(C_1D_2) = (A_2B_1)(C_2D_1) = (A_2C_2)(B_1D_1)$
Ainsi le résultat de l'opération FG ne dépend pas des rapports choisis pour représenter F et G.
□

Proposition

Les opérations sur $\mathbb{K}(X)$ prolongent celles connues $\mathbb{K}[X]$.

dém. :
Pour $P, Q \in \mathbb{K}[X]$,

$$\lambda.\frac{P}{1} = \frac{\lambda.P}{1}, \frac{P}{1} + \frac{Q}{1} = \frac{P+Q}{1} \text{ et } \frac{P}{1}\frac{Q}{1} = \frac{PQ}{1}$$

□

Théorème

$(\mathbb{K}(X), +, .)$ est un $\mathbb{K}$-espace vectoriel d'élément nul la fraction nulle $0 = \dfrac{0}{1}$.

dém. :
L'addition sur $\mathbb{K}(X)$ est évidemment commutative. Etudions son associativité.
D'une part

$$\left(\frac{A}{B}+\frac{C}{D}\right)+\frac{E}{F}=\frac{AD+BC}{BD}+\frac{E}{F}=\frac{(AD+BC)F+BDE}{BDF}=\frac{ADF+BCF+BDE}{BDF}$$

D'autre part

$$\frac{A}{B}+\left(\frac{C}{D}+\frac{E}{F}\right)=\frac{A}{B}+\frac{CF+ED}{DF}=\frac{ADF+B(CF+DE)}{BDF}=\frac{ADF+BCF+BDE}{BDF}$$

Ainsi l'addition sur $\mathbb{K}(X)$ est associative.
La fraction nulle $0=\dfrac{0}{1}$ est élément neutre pour l'addition ; en effet

$$\frac{A}{B}+\frac{0}{1}=\frac{A\times 1+B\times 0}{B}=\frac{A}{B}$$

Puisque

$$\frac{A}{B}+\frac{-A}{B}=\frac{AB-AB}{B^2}=\frac{0}{B^2}=\frac{0}{1}=0$$

toute fraction de $\mathbb{K}(X)$ est symétrisable
Ainsi $(\mathbb{K}(X),+)$ est un groupe abélien de neutre la fraction nulle.
Enfin les propriétés calculatoires

$$\lambda.(F+G)=\lambda.F+\lambda.G,\ (\lambda+\mu).F=\lambda.F+\mu.F,$$

$$\lambda.\left(\mu.F\right)=(\lambda\mu).F \text{ et } 1.F=F$$

sont immédiates à obtenir.
□

Théorème

$(\mathbb{K}(X),+,\times)$ est un corps de neutre multiplicatif la fraction $1=\dfrac{1}{1}$.

dém. :
On a vu ci-dessus que $(\mathbb{K}(X),+)$ est un groupe abélien.
La multiplication sur $\mathbb{K}(X)$ est évidemment commutative, associative et la fraction $1=1/1$ en est élément neutre.
Etudions la distributivité de la multiplication sur l'addition.

$$\begin{aligned}\frac{A}{B}\left(\frac{C}{D}+\frac{E}{F}\right)&=\frac{A}{B}\frac{CF+DE}{DF}=\frac{A(CF+DE)}{BDF}\\&=\frac{ACF+ADE}{BDF}=\frac{AC}{BD}+\frac{AE}{BF}=\frac{A}{B}\frac{C}{D}+\frac{A}{B}\frac{E}{F}\end{aligned}$$

Ainsi $\mathbb{K}(X)$ est un anneau commutatif.
De plus on peut affirmer que $\mathbb{K}(X)$ est non réduit à $\{0\}$.
Soit $F=\dfrac{A}{B}\in\mathbb{K}(X)$ différent de la fraction nulle.
On peut affirmer que $A\neq 0$ et introduire $G=\dfrac{B}{A}\in\mathbb{K}(X)$.
On a $FG=\dfrac{AB}{BA}=\dfrac{1}{1}$ donc F est inversible et $F^{-1}=\dfrac{B}{A}$.
□

Remarque L'inverse d'une fraction F non nulle est parfois notée $\frac{1}{F}$.

Exemple Soit $n \in \mathbb{N}$.
On a $X^{n+1} - 1 = (X-1)(1 + X + \cdots + X^n)$ dans $\mathbb{K}[X]$.
En multipliant par l'inverse de $X - 1$ dans $\mathbb{K}(X)$, on obtient la relation

$$\frac{X^{n+1}-1}{X-1} = 1 + X + \cdots + X^n$$

dans $\mathbb{K}(X)$.

Remarque Ici on peut diviser par $X - 1$ dans $\mathbb{K}(X)$ car $X - 1$ n'est pas la fraction nulle.
On peut faire cette manipulation sans se soucier de ce qui se passe en 1 car X est une indéterminée et non une variable. Le problème en 1 se posera seulement lorsqu'on voudra évaluer l'égalité précédente en 1.

9.1.2 Représentant irréductible

Théorème

Pour tout $F \in \mathbb{K}(X)$, il existe un unique couple $(P, Q) \in \mathbb{K}[X]^2$ tel que :
(1) Q est unitaire ;
(2) $F = P/Q$;
(3) P et Q sont premiers entre eux.
La fraction P/Q est alors appelée représentant irréductible de F.

dém. :
Unicité :
Soient (P_1, Q_1) et (P_2, Q_2) deux couples solutions.
Puisque $F = \frac{P_1}{Q_1} = \frac{P_2}{Q_2}$, on a $P_1Q_2 = P_2Q_1$.
Ainsi Q_1 divise P_1Q_2, or Q_1 est premier avec P_1 donc par le théorème de Gauss, Q_1 divise Q_2.
De façon symétrique, Q_2 divise Q_1. Les polynômes Q_1 et Q_2 sont donc associés, or ils sont tous deux unitaires, ils sont donc égaux. L'égalité $P_1Q_2 = P_2Q_1$ avec $Q_1 = Q_2 \neq 0$ donne alors $P_1 = P_2$.
Finalement $(P_1, Q_1) = (P_2, Q_2)$.
Existence :
Soit A/B un rapport représentant la fraction rationnelle F.
Quitte à multiplier en haut et en bas par un scalaire, on peut supposer que le polynôme B est unitaire.
Considérons ensuite D le pgcd de A et B.
Puisque D divise A et B, on peut écrire $A = DP$ et $B = DQ$.
On a alors

$$F = \frac{A}{B} = \frac{DP}{DQ} = \frac{P}{Q}$$

De plus $D = \text{pgcd}(A, B) = \text{pgcd}(DP, DQ) = D\text{pgcd}(P, Q)$ donc $\text{pgcd}(P, Q) = 1$.
Enfin puisque B et D sont unitaires, Q l'est aussi.
Finalement le couple (P, Q) ainsi construit est solution.
□

Exemple Pour $P \in \mathbb{K}[X]$, $\frac{P}{1}$ est le représentant irréductible de P.

Remarque Pour vérifier qu'un rapport représentant une fraction rationnelle est formé par deux polynômes premiers entre eux, il suffit de vérifier que ceux-ci n'ont pas de racines complexe en commun ; ceci est facile à faire dès que l'on connaît les racines de l'un d'entre eux.

Exemple Le représentant irréductible de $\dfrac{X^6-1}{X^4-1}$ est $\dfrac{X^4+X^2+1}{X^2+1}$.
En effet ces deux fractions rationnelles sont égales (via factorisation pas X^2-1) et la seconde est formée par un rapport irréductible car les polynômes X^4+X^2+1 et X^2+1 sont premiers entre eux puisque sans racines complexes en commun.

9.1.3 Degré

Définition

Soit $F \in \mathbb{K}(X)$ de représentant irréductible P/Q.
On appelle degré de F le nombre

$$\deg F = \deg P - \deg Q \in \mathbb{Z} \cup \{-\infty\}$$

Exemple Pour $P \in \mathbb{K}[X]$

$$\deg \frac{P}{1} = \deg P - \deg 1 = \deg P$$

Ainsi la notion de degré d'une fraction rationnelle prolonge celle de degré d'un polynôme.

Proposition

Pour $F = A/B \in \mathbb{K}(X)$,

$$\deg F = \deg A - \deg B$$

dém. :
Soit P/Q le représentant irréductible de F.
Puisque $F = \dfrac{A}{B} = \dfrac{P}{Q}$, on a $AQ = BP$ donc $\deg A + \deg Q = \deg B + \deg P$.
Puisque B, Q sont non nuls, $\deg B, \deg Q \in \mathbb{N}$ et la relation qui précède donne $\deg A - \deg B = \deg P - \deg Q$.
□

Exemple $\deg \dfrac{X+1}{X^2+2} = -1$, $\deg \dfrac{X^3+X+1}{X^2-X+2} = 1$, $\deg \dfrac{X^2}{X^2+1} = 0$.

Proposition

Soient $F, G \in \mathbb{K}(X)$ et $\lambda \in \mathbb{K}$.

$$\deg \lambda.F = \begin{cases} \deg F \text{ si } \lambda \neq 0 \\ -\infty \text{ si } \lambda = 0 \end{cases}$$

$$\deg(F+G) \leqslant \max(\deg F, \deg G) \text{ et}$$

$$\deg FG = \deg F + \deg G$$

dém. :
Soient A/B et C/D des rapports représentant F et G.
Si $\lambda = 0$ alors $\lambda.F = 0$ donc $\deg(\lambda F) = -\infty$.
Si $\lambda \neq 0$ alors $\lambda.F = \dfrac{\lambda A}{B}$ donc

$$\deg(\lambda F) = \deg(\lambda A) - \deg B = \deg A - \deg B = \deg F$$

$F + G = \dfrac{AD + BC}{BD}$ donc

$$\deg(F+G) = \deg(AD+BC) - \deg(BD)$$

Or

$$\deg(AD+BC) \leqslant \max(\deg AD, \deg BC)$$

donc

$$\deg(F+G) \leqslant \max\left(\deg AD - \deg BD, \deg BC - \deg BD\right)$$

puis

$$\deg(F+G) \leqslant \max\left(\deg A - \deg B, \deg C - \deg D\right) = \max(\deg F, \deg G)$$

Enfin $FG = \dfrac{AC}{BD}$ donc

$$\deg(FG) = \deg AC - \deg BD = \deg A + \deg C - \deg B - \deg D = \deg F + \deg G$$

□

9.1.4 Dérivation

Définition

Soit $F \in \mathbb{K}(X)$ de représentant irréductible P/Q.
On appelle fraction rationnelle dérivée de F la fraction

$$F' = \frac{P'Q - PQ'}{Q^2}$$

Exemple Pour $P \in \mathbb{K}[X]$,

$$\left(\frac{P}{1}\right)' = \frac{P' \times 1 - P \times 0}{1^2} = P'$$

Ainsi la notion de dérivation d'une fraction rationnelle prolonge la notion de dérivation d'un polynôme.

Proposition

Pour $F = A/B \in \mathbb{K}(X)$, on a

$$F' = \frac{A'B - AB'}{B^2}$$

dém. :
Soit P/Q le représentant irréductible de F.
Puisque $F = \dfrac{A}{B} = \dfrac{P}{Q}$, on a $PB = QA$.
En dérivant cette relation on obtient $P'B + PB' = Q'A + QA'$ (*).
On a

$$F' = \frac{P'Q - PQ'}{Q^2} = \frac{P'QB^2 - PQ'B^2}{Q^2B^2} = \frac{P'BQB - (PB)Q'B}{Q^2B^2}$$

donc

$$F' = \frac{P'BQB - (QA)Q'B}{Q^2B^2} = \frac{(P'Q - Q'A)QB}{Q^2B^2}$$

En exploitant la relation (*), on obtient

$$F' = \frac{(PB' - QA')QB}{Q^2B^2} = \frac{(PB)QB' - A'BQ^2}{Q^2B^2} = \frac{(AQ)B'Q - A'BQ^2}{Q^2B^2}$$

Enfin, en simplifiant par Q^2,

$$F' = \frac{A'B - AB'}{B^2}$$

□

Exemple Pour $n \in \mathbb{N}^\star$ et $a \in \mathbb{K}$,

$$\left(\frac{1}{(X-a)^n}\right)' = -\frac{n}{(X-a)^{n+1}}$$

Proposition

Soient $F, G \in \mathbb{K}(X)$ et $\lambda \in \mathbb{K}$.

$$(\lambda F)' = \lambda F', (F+G)' = F' + G', (FG)' = F'G + FG'$$

et, lorsque $G \neq 0$,

$$\left(\frac{F}{G}\right)' = \frac{F'G - FG'}{G^2}$$

dém. :
Soient A/B et C/D des rapports représentants F et G.
L'égalité $(\lambda F)' = \lambda F'$ est immédiate.

$$(F+G)' = \left(\frac{AD + BC}{BD}\right)' = \frac{(A'D + AD' + B'C + BC')BD + (AD+BC)(B'D + BD')}{B^2D^2}$$

et

$$F' + G' = \frac{A'B - AB'}{B^2} + \frac{C'D - CD'}{D^2} = \frac{(A'B - AB')D^2 + B^2(C'D - CD')}{B^2D^2}$$

Après simplification du terme $B'CBD$ dans la première, ces deux fractions rationnelles se correspondent et donc $(F+G)' = F' + G'$.

$$(FG)' = \left(\frac{AC}{BD}\right)' = \frac{(A'C + AC')BD + AC(B'D + BD')}{B^2D^2}$$

et

$$F'G + FG' = \frac{(A'B - AB')C}{B^2D} + \frac{A(C'D - CD')}{BD^2} = \frac{(A'B - AB')CD + AB(C'D - CD')}{B^2D^2}$$

Ces deux fractions rationnelles se correspondent et donc $(FG)' = F'G + FG'$
Enfin, puisque $G \times \frac{1}{G} = 1$, on obtient en dérivant, $G' \times \frac{1}{G} + G \times \left(\frac{1}{G}\right)' = 0$ et donc $\left(\frac{1}{G}\right)' = -\frac{G'}{G^2}$.
On en déduit

$$\left(\frac{F}{G}\right)' = \left(F \times \frac{1}{G}\right) = \frac{F'}{G} - \frac{FG'}{G^2} = \frac{F'G - FG'}{G^2}$$

□

Proposition

> Pour $F \in \mathbb{K}(X)$ tel que $F \neq 0$, on a $\deg F' \leqslant \deg F - 1$.

dém. :
Soit A/B un rapport représentant F.

$$F' = \frac{A'B - AB'}{B^2}$$

donc

$$\deg F' = \deg(A'B - AB') - 2\deg B$$

Or

$$\deg(A'B - AB') \leqslant \max(\deg A'B, \deg AB') \leqslant \deg A + \deg B - 1$$

donc $\deg F' \leqslant \deg F - 1$.
□

Attention : $\deg F'$ n'est pas toujours égal à $\deg F - 1$.

Exemple Pour $F = \frac{X+1}{X} = 1 + \frac{1}{X}$, $\deg F = 0$ et $\deg F' = -2$!

9.1.5 Racines et pôles d'une fraction rationnelle

Définition

Soit $F \in \mathbb{K}(X)$ de représentant irréductible P/Q.
On appelle racine de F toute racine de P.
On appelle pôle de F toute racine de Q.

Remarque Comme P et Q sont premiers entre eux, un même élément a ne peut pas être à la fois racine et pôle de F.

Exemple Pour $P \in \mathbb{K}[X]$, les racines de $P/1$ sont les racines de P et $P/1$ n'a pas de pôles.

Remarque Une fraction rationnelle n'a qu'un nombre fini de pôles. En effet son dénominateur est un polynôme non nul et celui-ci n'a qu'un nombre fini de racines.

Exemple Considérons

$$F = \frac{X^2 - 1}{X^3 - 1} \in \mathbb{C}(X)$$

Pour déterminer racines et pôles de F, calculons un représentant irréductible.

$$F(X) = \frac{(X-1)(X+1)}{(X-1)(X^2+X+1)} = \frac{X+1}{X^2+X+1}$$

Par suite : -1 est racine, j et j^2 sont pôles.
1 n'est ni racine ni pôle.

Définition

Soit $F \in \mathbb{K}(X)$ de représentant irréductible P/Q et $a \in \mathbb{K}$.
Si a est racine de F (resp. pôle de F), la multiplicité de a en tant que racine de P (resp. racine de Q) est appelée multiplicité de la racine a dans F (resp. du pôle a dans F).

Exemple Considérons

$$F = \frac{(X-1)(X+2)^2}{X^2(X+1)} \in \mathbb{R}(X)$$

La fraction rationnelles F est écrite ici sous forme irréductible.
1 est racine simple et -2 racine double de F.
-1 est pôle simple et 0 est pôle double de F.

Proposition

Soient $F = A/B \in \mathbb{K}(X) - \{0\}$ et $a \in \mathbb{K}$.
Posons $\alpha, \beta \in \mathbb{N}$ les multiplicités de a en tant que racine de A et B.
Si $\alpha > \beta$ alors a est racine de F de multiplicité $\alpha - \beta$.
Si $\alpha = \beta$ alors a n'est ni racine, ni pôle de F.
Si $\alpha < \beta$ alors a est pôle de F de multiplicité $\beta - \alpha$.

dém. :
On peut écrire $A = (X - a)^\alpha \tilde{A}$ avec $\tilde{A}(a) \neq 0$ et $B = (X - a)^\beta \tilde{B}$ avec $\tilde{B}(a) \neq 0$.
Soit P/Q le représentant irréductible de F.
On a $PB = QA$ donc

$$(X - a)^\beta P\tilde{B} = (X - a)^\alpha Q\tilde{A} \text{ (*)}$$

Cas $\alpha > \beta$:
En simplifiant dans la relation (*), on obtient $P\tilde{B} = (X - a)^{\alpha-\beta} Q\tilde{A}$
En évaluant en a, $P(a)\tilde{B}(a) = 0$ donc $P(a) = 0$ car $\tilde{B}(a) \neq 0$.
Ainsi a est racine de P. Cependant a n'est pas racine de Q car P et Q n'ont pas de racines en commun.
Il reste à déterminer la multiplicité de a en tant que racine de P.
Puisque $\tilde{B}$ divise $(X - a)^{\alpha-\beta} Q\tilde{A}$ et que $\tilde{B}$ est premier avec $X - a$ (car $\tilde{B}(a) \neq 0$), le théorème de Gauss donne que $\tilde{B}$ divise $Q\tilde{A}$ ce qui permet d'écrire $Q\tilde{A} = \tilde{B}C$.
La relation $P\tilde{B} = (X - a)^{\alpha-\beta} Q\tilde{A}$ donne alors $P = (X - a)^{\alpha-\beta} C$.
Comme $Q(a) \neq 0$ et $\tilde{A}(a) \neq 0$, on a nécessairement $C(a) \neq 0$ car $Q\tilde{A} = \tilde{B}C$.
Ainsi $P = (X - a)^{\alpha-\beta} C$ avec $C(a) \neq 0$, a est racine de multiplicité exactement $\alpha - \beta$ de P.
Cas $\alpha = \beta$:
En simplifiant dans la relation (*), on obtient $P\tilde{B} = Q\tilde{A}$.
En évaluant en a on obtient $P(a)\tilde{B}(a) = Q(a)\tilde{A}(a)$.
Puisque $\tilde{A}(a), \tilde{B}(a) \neq 0$, on peut affirmer $P(a) = 0 \Leftrightarrow Q(a) = 0$.
Or P et Q n'ont pas de racine commune donc $P(a) \neq 0$ et $Q(a) \neq 0$.
Ainsi a n'est ni racine, ni pôle de F.
Cas $\alpha < \beta$:
Il suffit de transposer l'étude menée dans le cadre $\alpha > \beta$.

□

Exemple Soient $p, q \in \mathbb{N}^\star$ tels que $p \wedge q = 1$ et

$$F = \frac{X^p - 1}{X^q - 1}$$

Déterminons les racines et les pôles de F.
Les racines du numérateur sont les racines p-ième de l'unité, ce sont des racines simples.
Les racines du dénominateur sont les racines q-ième de l'unité, ce sont des racines simples.
Déterminons les racines communes.
Soit ω une racine commune.
Puisque p et q sont premiers entre eux, par l'égalité de Bézout, il existe $u, v \in \mathbb{Z}$ tels que $pu + qv = 1$.
On a alors $\omega = \omega^1 = \omega^{pu+qv} = (\omega^p)^u (\omega^q)^v = 1$.
Ainsi 1 est la seule racine commune au numérateur et au dénominateur.
On peut conclure :
Les racines de F sont les racines $p$ème de l'unité autre que 1, ce sont des racines simples.
Les pôle de F sont les racines $q$ème de l'unité autre que 1, ce sont des pôles simples.
Enfin, 1 n'est ni racine, ni pôle de F.

9.1.6 Evaluation

Définition

Soient $F \in \mathbb{K}(X)$ de représentant irréductible P/Q et $a \in \mathbb{K}$.
On dit que F est définie en a si a n'est pas pôle de F i.e. $Q(a) \neq 0$.
On pose alors

$$F(a) = \frac{P(a)}{Q(a)}$$

appelée valeur de F en a.

Proposition

Soient $F \in \mathbb{K}(X)$ de représentant A/B et $a \in \mathbb{K}$.
Si $B(a) \neq 0$ alors F est définie en a et $F(a) = A(a)/B(a)$.

dém. :
Soit P/Q le représentant irréductible de F. On a $PB = QA$.
Si $B(a) \neq 0$ alors nécessairement $Q(a) \neq 0$ car $Q(a) = 0 \Rightarrow P(a)B(a) = 0 \Rightarrow P(a) = 0$ alors que P et Q n'ont pas de racines communes.
Ainsi F est définie en a et puisque l'égalité $PB = QA$ entraîne $P(a)B(a) = Q(a)A(a)$ on obtient

$$\frac{P(a)}{Q(a)} = \frac{A(a)}{B(a)}$$

sachant $B(a), Q(a) \neq 0$.
□

Proposition

Soient $a \in \mathbb{K}, F, G \in \mathbb{K}(X)$ et $\lambda \in \mathbb{K}$.
Si F et G sont définies en a alors $\lambda F, F + G$ et FG le sont aussi et on a les relations

$$(\lambda F)(a) = \lambda F(a),\ (F + G)(a) = F(a) + G(a) \text{ et } (FG)(a) = F(a)G(a)$$

dém. :
Soient P/Q et R/S les représentants irréductible de F et G.
a n'est pas racines de Q ni de S.
Puisque $\lambda F = \dfrac{\lambda P}{Q}$, $F + G = \dfrac{PS + QR}{QS}$ et $FG = \dfrac{PR}{QS}$ avec $Q(a) \neq 0$ et $(QS)(a) = Q(a)S(a) \neq 0$, les fractions λF, $F + G$ et FG sont définies en a et

$$(\lambda F)(a) = \frac{\lambda P(a)}{Q(a)} = \lambda F(a)$$

$$(F + G)(a) = \frac{P(a)S(a) + Q(a)R(a)}{Q(a)S(a)} = \frac{P(a)}{Q(a)} + \frac{R(a)}{S(a)} = F(a) + G(a)$$

et

$$(FG)(a) = \frac{P(a)R(a)}{Q(a)S(a)} = \frac{P(a)}{Q(a)}\frac{R(a)}{S(a)} = F(a)G(a)$$

□

Proposition

Soient $a \in \mathbb{K}$ et $F \in \mathbb{K}(X)$.
Si F est définie en a alors F' est définie en a.

dém. :
Soit P/Q le représentant irréductible de F, a n'est pas racine de Q.
Puisque $F' = \dfrac{P'Q - PQ'}{Q^2}$ avec $Q^2(a) = [Q(a)]^2 \neq 0$, F' est définie en a.
□

9.1.7 Fonctions rationnelles

Définition

On appelle ensemble de définition de $F \in \mathbb{K}(X)$, l'ensemble $\mathcal{D}_F$ formé des $a \in \mathbb{K}$ tels que F définie en a.

Exemple Pour $P \in \mathbb{K}[X]$, $\mathcal{D}_P = \mathbb{K}$
Pour $F = 1/X(X-1)$, $\mathcal{D}_F = \mathbb{K} - \{0, 1\}$.

Remarque Pour $F \in \mathbb{K}(X)$ de représentant irréductible P/Q, $\mathcal{D}_F = \mathbb{K} \backslash Q^{-1}(\{0\})$.

Définition

Soit $F \in \mathbb{K}(X)$ et $\mathcal{D}$ une partie de $\mathbb{K}$ incluse dans $\mathcal{D}_F$.
On appelle fonction rationnelle associée à F définie sur $\mathcal{D}$ l'application $\tilde{F} : \mathcal{D} \to \mathbb{K}$ qui à $a \in \mathcal{D}$ associe $\tilde{F}(a) = F(a)$.
Quand $\mathcal{D} = \mathcal{D}_F$, on parle simplement de fonction rationnelle associée à F.

Exemple La fonction rationnelle associée à $1/X \in \mathbb{C}[X]$ est l'application $z \mapsto 1/z$ définie de $\mathbb{C}^\star$ vers $\mathbb{C}$.

Proposition

Soient $F, G \in \mathbb{K}(X)$ et $\mathcal{D}$ une partie infinie de $\mathbb{K}$.
Si
$$\forall x \in \mathcal{D} \cap \mathcal{D}_F \cap \mathcal{D}_G, \tilde{F}(x) = \tilde{G}(x)$$
alors $F = G$.

dém. :
Soient P/Q et R/S les représentants irréductibles de F et G.
Pour tout $x \in \mathcal{D} \cap \mathcal{D}_F \cap \mathcal{D}_G$, l'égalité $\tilde{F}(x) = \tilde{G}(x)$ donne $\dfrac{P(x)}{Q(x)} = \dfrac{R(x)}{S(x)}$ et donc $(PS)(x) = (QR)(x)$.
Les polynômes PS et QR coïncident sur $\mathcal{D} \cap \mathcal{D}_F \cap \mathcal{D}_G$ qui est une partie infinie (car $\mathcal{D}$ est infinie et $\mathcal{D} \cap \mathcal{D}_F \cap \mathcal{D}_G$ correspond à $\mathcal{D}$ privé des pôles de F et G qui sont en nombre fini) donc $PS = QR$ puis $F = G$.
□

Remarque Suite à ce résultat, on peut identifier la fraction rationnelle F à la fonction rationnelle $\tilde{F}$ dès que l'ensemble de départ est infini.

Exemple Montrons que pour tout $n \in \mathbb{N}$, il existe un unique fraction rationnelle $F_n \in \mathbb{R}(X)$ telle que
1) $]-1,1[\subset \mathcal{D}_{F_n}$;
2) $\forall x \in]-1,1[$, $(\arcsin x)^{(n+1)} = \dfrac{F_n(x)}{\sqrt{1-x^2}}$.
Unicité :
Supposons F_n et G_n solutions.
Pour tout $x \in]-1,1[$, on a $F_n(x) = G_n(x)$.
Puisque les fractions rationnelles F_n et G_n coïncident sur une partie infinie, elles sont égales.
Existence :
Raisonnons par récurrence sur $n \in \mathbb{N}$.
Pour $n = 1$, $F_1 = 1$ convient.
Supposons l'existence établie au rang $n \geqslant 0$.
Par hypothèse de récurrence, il existe un fraction rationnelle $F_n \in \mathbb{R}(X)$ telle que $]-1,1[\subset \mathcal{D}_{F_n}$ et

$$(\arcsin x)^{(n+1)} = \frac{F_n(x)}{\sqrt{1-x^2}}$$

pour tout $x \in]-1,1[$.
En dérivant cette dernière relation, on obtient

$$(\arcsin x)^{(n+2)} = \frac{F_n'(x)}{\sqrt{1-x^2}} + \frac{xF_n(x)}{\sqrt{1-x^2}^3}$$

Posons alors

$$F_{n+1} = F_n' + \frac{X}{1-X^2}F_n \in \mathbb{R}(X)$$

Puisque F_n est définie sur $]-1,1[$, par opérations, F_{n+1} est aussi définie sur $]-1,1[$.
De plus, par construction, on a

$$(\arcsin x)^{(n+2)} = \frac{F_{n+1}(x)}{\sqrt{1-x^2}}$$

Récurrence établie.

9.2 Décomposition en éléments simples

9.2.1 Partie entière

Théorème

Pour tout $F \in \mathbb{K}(X)$ il existe un unique couple $(E,G) \in \mathbb{K}[X] \times \mathbb{K}(X)$ tel que

$$F = E + G \text{ et } \deg G < 0$$

E est alors appelé partie entière de F et on note $E = \text{Ent}(F)$.

dém. :
Unicité :
Soient (E_1, G_1) et (E_2, G_2) deux couples solutions.
On a $F = E_1 + G_1 = E_2 + G_2$ et donc $E_1 - E_2 = G_2 - G_1$ puis $\deg(E_1 - E_2) = \deg(G_2 - G_1) < 0$.
Or $E_1 - E_2 \in \mathbb{K}[X]$ donc $E_1 = E_2$ puis $G_1 = G_2$.
Existence :

Soit A/B un rapport représentant de F.
En réalisant la division euclidienne de A par B, on peut écrire $A = BQ + R$ avec $\deg R < \deg B$ et alors $F = \dfrac{A}{B} = Q + \dfrac{R}{B} = Q + G$ avec $Q \in \mathbb{K}[X]$ et $G \in \mathbb{K}(X)$ tel que $\deg G < 0$.
□

Remarque Pour calculer la partie entière de $F = A/B$, il suffit de poser la division euclidienne de A par B.

Exemple Calculons

$$\text{Ent}\left(\frac{X^3+X-1}{X^2-X+1}\right)$$

Par division euclidienne

$$X^3 + X - 1 = (X^2 - X + 1)(X + 1) + (X - 2)$$

donc

$$\text{Ent}\left(\frac{X^3+X-1}{X^2-X+1}\right) = X + 1$$

Remarque Si $\deg F < 0$ alors $\text{Ent}(F) = 0$.

Remarque Si on retranche à une fraction F sa partie entière, la fraction obtenue est de degré < 0.

9.2.2 Partie polaire

Théorème

Soit a un pôle de multiplicité $\alpha \in \mathbb{N}^\star$ de $F \in \mathbb{K}(X)$.
Il existe un unique couple $(R, G) \in \mathbb{K}[X] \times \mathbb{K}(X)$ tel que :
- $F = G + \dfrac{R}{(X-a)^\alpha}$;
- a n'est pas pôle de G ;
- $\deg R \leqslant \alpha - 1$.

La fraction $\dfrac{R}{(X-a)^\alpha}$ est alors appelée partie polaire de F en a.

dém. :
Unicité :
Soient (R_1, G_1) et (R_2, G_2) deux couples solutions.
On a

$$G_1 - G_2 = \frac{R_2 - R_1}{(X-a)^\alpha}$$

Comme a n'est pas pôle de $G_1 - G_2$, a est racine de multiplicité au moins n de $R_2 - R_1$.
Or $\deg(R_2 - R_1) < \alpha$ donc $R_2 - R_1 = 0$ car la somme des multiplicités de ses racines est strictement supérieure à son degré. Par suite $R_1 = R_2$ puis $G_1 = G_2$.

Existence :
Soit P/Q le représentant irréductible de F.
On a $Q = (X-a)^\alpha \hat{Q}$ avec $\hat{Q}(a) \neq 0$.
$(X-a)^\alpha$ et $\hat{Q}$ sont premiers entre eux donc par le théorème de Bézout, il existe $U, V \in \mathbb{K}[X]$ tels que $(X-a)^\alpha U + AV = 1$. On a alors

$$F = \frac{P}{Q} = \frac{P(X-a)^\alpha U + P\hat{Q}V}{(X-a)^\alpha \hat{Q}} = \frac{PU}{\hat{Q}} + \frac{PV}{(X-a)^\alpha}$$

Réalisons la division euclidienne de PV par $(X-a)^\alpha$:
$PV = (X-a)^\alpha T + R$ avec $\deg R < \alpha$.
On peut alors écrire

$$F = \frac{PU}{\hat{Q}} + T + \frac{R}{(X-a)^\alpha} = G + \frac{R}{(X-a)^\alpha}$$

avec

$$G = \frac{PU}{\hat{Q}} + T$$

qui n'a pas de pôle en a.
□

Remarque La partie polaire relative à un pôle a est de degré < 0.

Remarque Lorsqu'on retranche à une fraction sa partie polaire relative à un pôle a donné, la fraction obtenue ne présente plus de pôle en a.

Proposition

Soient $a \in \mathbb{K}, \alpha \in \mathbb{N}^\star$ et $R \in \mathbb{K}_{\alpha-1}[X]$.
Il existe un unique $(\lambda_1, \ldots, \lambda_\alpha) \in \mathbb{K}^\alpha$ tel que

$$\frac{R}{(X-a)^\alpha} = \frac{\lambda_\alpha}{(X-a)^\alpha} + \cdots + \frac{\lambda_2}{(X-a)^2} + \frac{\lambda_1}{X-a}$$

dém. :
On a l'équivalence

$$\frac{R}{(X-a)^\alpha} = \sum_{k=1}^{\alpha} \frac{\lambda_k}{(X-a)^k} \Leftrightarrow R = \sum_{k=1}^{n} \lambda_k (X-a)^{\alpha-k}$$

Puisque la famille $(1, X-a, (X-a)^2, ..., (X-a)^{\alpha-1})$ est une famille de polynôme de degrés étagés, elle forme une base de $\mathbb{K}_{\alpha-1}[X]$ et donc tout polynôme $R \in \mathbb{K}_{\alpha-1}[X]$ s'écrit de façon unique sous la forme $\sum_{k=1}^{n} \lambda_k (X-a)^{\alpha-k}$.
□

Remarque Une partie polaire relative à un pôle a de multiplicité α peut donc d'écrire sous la forme d'une combinaison linéaire des fractions

$$\frac{1}{X-a}, \frac{1}{(X-a)^2}, \ldots, \frac{1}{(X-a)^\alpha}$$

9.2.3 Décomposition en éléments simples dans $\mathbb{C}(X)$

Lemme

Dans $\mathbb{C}(X)$, une fraction rationnelle qui n'a pas de pôle est un polynôme.

dém. :
Soit P/Q le représentant irréductible d'une fraction rationnelle $F \in \mathbb{C}(X)$ qui n'a pas de pôles.
Si le polynôme Q n'est pas constant alors il possède au moins une racine et celle-ci est alors pôle de F ; c'est exclu. Le polynôme Q est donc constant et, puisqu'il est unitaire, $Q = 1$ et finalement $F = P \in \mathbb{C}[X]$.
□

Théorème

Toute fraction rationnelle de $\mathbb{C}(X)$ est égale à la somme de sa partie entière et de ses différentes parties polaires.

dém. :
Si on retranche à F sa partie entière on obtient une fraction de degré strictement négatif. Si de plus on y retranche ses parties polaires, la fraction obtenue ne présente plus de pôles et reste de degré strictement négatif. Cette fraction est donc un polynôme de degré strictement négatif, c'est le polynôme nul.
□

Corollaire

Pour $F \in \mathbb{C}(X)$ de représentant irréductible P/Q.
Si la factorisation de Q dans $\mathbb{C}[X]$ est

$$Q = (X - a_1)^{\alpha_1} \dots (X - a_n)^{\alpha_n}$$

avec $a_1, \dots, a_n$ deux à deux distincts alors les pôles de F sont les a_i de multiplicité α_i et on peut écrire

$$F = \text{Ent}(F) + \sum_{i=1}^{n} \sum_{j=1}^{\alpha_i} \frac{\lambda_{i,j}}{(X - a_i)^j} \text{ avec } \lambda_{i,j} \in \mathbb{C}$$

De plus cette écriture est unique.

dém. :
Unicité :
L'écriture proposée permet d'identifier la partie entière et les différentes parties polaires de F qui sont déterminées de façon unique.
Existence :
F est égale à la somme de sa partie entière et de ses différentes parties polaires qui se décrivent sous la forme proposée.
□

Exemple Soit $P = \lambda(X - a_1)^{\alpha_1} \dots (X - a_n)^{\alpha_n}$ un polynôme de $\mathbb{C}[X]$.
Réalisons la DES de la fraction $F = \dfrac{P'}{P}$.
Puisque

$$P' = \lambda \sum_{k=1}^{n} (X - a_1)^{\alpha_1} \dots [(X - a_k)^{\alpha_k}]' \dots (X - a_n)^{\alpha_n}$$

on a

$$P' = \lambda \sum_{k=1}^{n} \alpha_k (X - a_1)^{\alpha_1} \dots (X - a_k)^{\alpha_k - 1} \dots (X - a_n)^{\alpha_n}$$

puis

$$F = \frac{P'}{P} = \sum_{k=1}^{n} \frac{\alpha_k}{X - a_k} = \frac{\alpha_1}{X - a_1} + \cdots + \frac{\alpha_n}{X - a_n}$$

Cette relation s'interprète comme étant la décomposition en éléments simples de F.

9.2.4 Obtenir une décomposition en éléments simples

9.2.4.1 Démarche

Soit $F \in \mathbb{C}(X)$. Pour former la DES de F :
- on exprime F sous forme irréductible P/Q ;
- on détermine $\text{Ent}(F)$ en réalisant la division euclidienne de P par Q ;
- on factorise Q dans $\mathbb{C}(X)$ de sorte de déterminer les pôles de F ainsi que leurs multiplicités ;
- on exprime la DES de F à l'aide de coefficients inconnus : $a, b, c, d, \ldots$;
- on détermine ces coefficients par diverses méthodes

Exemple Considérons

$$F = \frac{X^2}{X^2 - 3X + 2} \in \mathbb{C}(X)$$

La fraction F est écrite sous forme irréductible car numérateur et dénominateur n'ont pas de racine en commun.
Il est immédiat que le partie entière de F vaut 1.
Puisque $X^2 - 3X + 2 = (X - 1)(X - 2)$, les pôles de F sont 1 et 2 ; ce sont des pôles simples.
Par le théorème de décomposition en éléments simples de F, on peut écrire

$$F = 1 + \frac{a}{X - 1} + \frac{b}{X - 2} \text{ avec } a, b \in \mathbb{C}$$

En évaluant cette relation en 0 et en -1, on obtient les équations

$$0 = 1 - a - \frac{b}{2} \text{ et } \frac{1}{6} = 1 - \frac{a}{2} - \frac{b}{3}$$

On a ainsi

$$\begin{cases} 2a + b = 2 \\ 3a + 2b = 5 \end{cases}$$

d'où l'on tire $a = -1$ et $b = 4$.
Ainsi

$$\frac{X^2}{X^2 - 3X + 2} = 1 - \frac{1}{X - 1} + \frac{4}{X - 2}$$

Ici, cette démarche est laborieuse et n'est pas à reprendre en pratique... Nous allons voir des techniques beaucoup plus efficace !

9.2.4.2 Détermination de la partie polaire relative à un pôle simple

Soient $F \in \mathbb{C}(X)$ de représentant irréductible P/Q et a un pôle simple de F.
On peut écrire $Q = (X-a)\hat{Q}$ avec $\hat{Q}(a) \neq 0$.
La décomposition en éléments simples de F permet d'écrire

$$F = \frac{P}{(X-a)\hat{Q}} = G + \frac{\lambda}{X-a} \text{ (1)}$$

avec G n'ayant pas de pôle en a.
En multipliant la relation (1) par $X-a$ on obtient

$$(X-a)F = \frac{P}{\hat{Q}} = (X-a)G + \lambda \text{ (2)}$$

En évaluant (2) en a, on parvient à $\dfrac{P(a)}{\hat{Q}(a)} = 0 + \lambda$ et donc

$$\lambda = \left(\frac{P}{\hat{Q}}\right)(a)$$

Remarque Puisque $Q = (X-a)\hat{Q}$, on a $Q' = (X-a)\hat{Q}' + R$ et donc $Q'(a) = \hat{Q}(a)$.
On a donc aussi la formule

$$\lambda = \left(\frac{P}{Q'}\right)(a)$$

Cette dernière est très utile quand on ne souhaite pas réaliser la factorisation du polynôme Q, c'est notamment le cas quand $Q = X^n - 1\ldots$

Exemple Considérons à nouveau

$$F = \frac{X^2}{X^2 - 3X + 2} \in \mathbb{C}(X)$$

Comme on l'a vu ci-dessus, la décomposition en éléments simples de F s'écrit

$$F = \frac{X^2}{(X-1)(X-2)} = 1 + \frac{a}{X-1} + \frac{b}{X-2}$$

Par ce qui précède

$$a = \left.\frac{X^2}{X-2}\right|_{X=1} = -1, b = \left.\frac{X^2}{X-1}\right|_{X=2} = 4$$

On retrouve

$$\frac{X^2}{X^2-3X+2} = 1 - \frac{1}{X-1} + \frac{4}{X-2}$$

Ici les calculs ont été plus efficace que lors du calcul précédent.

Exemple Soit $n \in \mathbb{N}^\star$ et $F = \dfrac{1}{X^n - 1} \in \mathbb{C}(X)$.
F est exprimée sous forme irréductible et Ent$(F) = 0$ car $\deg F < 0$.
Les pôles de F sont les racines n-ième de l'unité $\omega_0, ..., \omega_{n-1}$ avec $\omega_k = \mathrm{e}^{2ik\pi/n}$, ce sont des pôles simples.
La décomposition en éléments simples de F est de la forme

$$F = \frac{1}{X^n - 1} = \frac{c_0}{X - \omega_0} + \cdots + \frac{c_{n-1}}{X - \omega_{n-1}}$$

Par ce qui précède

$$c_k = \left.\frac{1}{(X^n - 1)'}\right|_{X=1} = \left.\frac{1}{nX^{n-1}}\right|_{X=1} = \frac{1}{n\omega_k^{n-1}} = \frac{1}{n}\omega_k$$

Ainsi

$$\frac{1}{X^n - 1} = \frac{1}{n}\sum_{k=0}^{n-1} \frac{\omega_k}{X - \omega_k}$$

9.2.4.3 Détermination de la partie polaire relative à un pôle double

Soient $F \in \mathbb{C}(X)$ de représentant irréductible P/Q et a un pôle double de F.
On peut écrire $Q = (X - a)^2\hat{Q}$ avec $\hat{Q}(a) \neq 0$.
La décomposition en éléments simples de F permet d'écrire :

$$F = \frac{P}{(X - a)^2\hat{Q}} = G + \frac{\lambda}{(X - a)^2} + \frac{\mu}{X - a} \ (1)$$

avec G n'ayant pas de pôle en a.
En multipliant (1) par $(X - a)^2$, on obtient

$$(X - a)^2 F = \frac{P}{\hat{Q}} = (X - a)^2 G + \lambda + \mu(X - a) \ (2)$$

En évaluant (2) en a, on parvient à
$\dfrac{P(a)}{\hat{Q}(a)} = 0 + \lambda$ et donc

$$\lambda = \left(\frac{P}{\hat{Q}}\right)(a)$$

En dérivant (2) puis en évaluant en a, on obtient

$$\mu = \left(\frac{P}{\hat{Q}}\right)'(a)$$

Exemple Considérons $F = \dfrac{X - 2}{X(X - 1)^2} \in \mathbb{C}(X)$.
F est exprimée sous forme irréductible et Ent$(F) = 0$ car $\deg F < 0$.
Les pôles de F sont 0 et 1 avec 0 pôle simple et 1 pôle double.

La décomposition en éléments simples de F est de la forme

$$\frac{X-2}{X(X-1)^2} = \frac{a}{X} + \frac{b}{(X-1)^2} + \frac{c}{X-1}$$

avec

$$a = \left.\frac{X-2}{(X-1)^2}\right|_{X=0} = -2,\, b = \left.\frac{X-2}{X}\right|_{X=1} = -1 \text{ et } c = \left.\left(\frac{X-2}{X}\right)'\right|_{X=1} = \left.\frac{2}{X^2}\right|_{X=1} = 2$$

Ainsi

$$\frac{X-2}{X(X-1)^2} = \frac{-2}{X} + \frac{-1}{(X-1)^2} + \frac{2}{X-1}$$

Exemple Considérons

$$F = \frac{X+1}{X^2(X-1)^2} \in \mathbb{C}(X)$$

F est exprimée sous forme irréductible et Ent$(F) = 0$ car $\deg F < 0$.
Les pôles de F sont 0 et 1 et ce sont des pôles doubles.
La décomposition en éléments simples de F est de la forme

$$\frac{X+1}{X^2(X-1)^2} = \frac{a}{X^2} + \frac{b}{X} + \frac{c}{(X-1)^2} + \frac{d}{X-1}$$

avec

$$a = \left.\frac{X+1}{(X-1)^2}\right|_{X=0} = 1,\, b = \left.\left(\frac{X+1}{(X-1)^2}\right)'\right|_{X=0} = \left.\frac{(X-1)^2 - 2(X+1)(X-1)}{(X-1)^4}\right|_{X=0} = 3$$

$$c = \left.\frac{X+1}{X^2}\right|_{X=1} = 2 \text{ et } d = \left.\left(\frac{X+1}{X^2}\right)'\right|_{X=1} = \left.-\frac{1}{X^2} - \frac{2}{X^3}\right|_{X=1} = -3$$

Ainsi

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

9.2.4.4 Démarche générale

Soient $F \in \mathbb{C}(X)$ de représentant irréductible P/Q et a un pôle de multiplicité α de F.
On peut écrire $Q = (X-a)^\alpha \hat{Q}$ avec $\hat{Q}(a) \neq 0$.
La décomposition en éléments simples de F permet d'écrire :

$$F = \frac{P}{(X-a)^\alpha \hat{Q}} = G + \frac{\lambda_\alpha}{(X-a)^\alpha} + \cdots + \frac{\lambda_1}{X-a} \text{ (1)}$$

avec G n'ayant pas de pôle en a.
En multipliant (1) par $(X-a)^\alpha$, on obtient

$$(X-a)^\alpha F = \frac{P}{\hat{Q}} = (X-a)^\alpha G + \lambda_\alpha + \lambda_{\alpha-1}(X-a) + \cdots + \lambda_1 (X-a)^{\alpha-1} \text{ (2)}$$

En évaluant (2) en a, on obtient

$$\frac{P(a)}{\hat{Q}(a)} = 0 + \lambda_\alpha + 0 \text{ et donc } \lambda_\alpha = \left(\frac{P}{\hat{Q}}\right)(a)$$

On peut alors calculer $G = F - \dfrac{\lambda_\alpha}{(X-a)^\alpha}$ et reprendre le processus précédent à partir de G qui présente en a un pôle d'ordre $< \alpha$.

Exemple Considérons

$$F = \frac{X^2}{(X+1)^3} \in \mathbb{C}(X)$$

F est exprimée sous forme irréductible et $\text{Ent}(F) = 0$ car $\deg F < 0$.
-1 est le seul pôle de F et c'est un pôle triple.
La décomposition en éléments simples de F est de la forme

$$\frac{X^2}{(X+1)^3} = \frac{a}{(X+1)^3} + \frac{b}{(X+1)^2} + \frac{c}{X+1}$$

avec $a = X^2\big|_{X=-1} = 1$.
Considérons alors

$$\frac{X^2}{(X+1)^3} - \frac{1}{(X+1)^3} = \frac{X^2-1}{(X+1)^3} = \frac{X-1}{(X+1)^2} = \frac{b}{(X+1)^2} + \frac{c}{X+1}$$

On a

$$b = X-1\big|_{X=-1} = -2 \text{ et } c = (X-1)'\big|_{X=-1} = 1$$

Finalement

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

Exemple Considérons

$$F = \frac{X^2+1}{X(X-1)^3} \in \mathbb{C}(X)$$

F est exprimée sous forme irréductible et $\text{Ent}(F) = 0$ car $\deg F < 0$.
Les pôles de F sont 0 et 1 ; 0 est pôle simple et 1 est pôle triple.
La décomposition en éléments simples de F est de la forme

$$\frac{X^2+1}{X(X-1)^3} = \frac{a}{X} + \frac{b}{(X-1)^3} + \frac{c}{(X-1)^2} + \frac{d}{X-1}$$

avec

$$b = \left.\frac{X^2+1}{X}\right|_{X=1} = 2$$

Considérons alors

$$\frac{X^2+1}{X(X-1)^3} - \frac{2}{(X-1)^3} = \frac{X^2-2X+1}{X(X-1)^3} = \frac{1}{X(X-1)} = \frac{a}{X} + \frac{c}{(X-1)^2} + \frac{d}{X-1}$$

On a

$$a = \left.\frac{1}{X-1}\right|_{X=0} = -1,\ c = 0 \text{ et } d = \left.\frac{1}{X}\right|_{X=1}$$

Finalement

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

9.2.4.5 Parties polaires complexes d'une fraction rationnelle réelle

Soit $F \in \mathbb{R}(X)$ de représentant irréductible P/Q.
Les pôles de F sont les racines du polynôme Q, or celui-ci est réel, donc les pôles complexes d'une fraction rationnelles sont deux à deux conjuguées. De plus

Proposition

Les parties polaires complexes d'une fraction rationnelle réelle sont deux à deux conjuguées.

dém. :
Si $F \in \mathbb{R}(X)$ alors par conjugaison complexe, la fraction rationnelle F est inchangée alors que les parties polaires relatives aux pôles complexes non réels sont échangées.
Ainsi, si

$$\frac{\lambda_1}{X-a} + \cdots + \frac{\lambda_\alpha}{(X-a)^\alpha}$$

est la partie polaire relative à un pôle $a \in \mathbb{C}$, par conjugaison,

$$\frac{\overline{\lambda_1}}{X-\bar{a}} + \cdots + \frac{\overline{\lambda_\alpha}}{(X-\bar{a})^\alpha}$$

est la partie polaire relative au pôle $\bar{a}$.
□

Exemple Considérons

$$F = \frac{X^2+X-1}{X(X^2+1)} \in \mathbb{R}(X)$$

F est écrite sous forme irréductible et sa partie entière est nulle car $\deg F < 0$.
Puisque $X(X^2+1) = X(X-i)(X+i)$, les pôles de F sont $0, i, -i$ et ce sont des pôles simples.
La décomposition en éléments simples de F est de la forme

$$\frac{X^2+X-1}{X(X^2+1)} = \frac{a}{X} + \frac{b}{X-i} + \frac{\bar{b}}{X+i}$$

car les parties polaires complexes de cette fraction réelle sont conjuguées.
On a

$$a = \left.\frac{X^2+X-1}{X^2+1}\right|_{X=0} = -1 \text{ et } b = \left.\frac{X^2+X-1}{X(X+i)}\right|_{X=i} = \frac{i-2}{-2} = 1 - \frac{i}{2}$$

puis $c = 1 + i/2$.
Finalement

$$\frac{X^2+X-1}{X(X^2+1)} = -\frac{1}{X} + \frac{1-\frac{i}{2}}{X-i} + \frac{1+\frac{i}{2}}{X+i}$$

Exemple Considérons
$$F = \frac{1}{X^2(X^2+X+1)} \in \mathbb{R}(X)$$
F est écrite sous forme irréductible et sa partie entière est nulle car $\deg F < 0$.
Puisque $X(X^2+X+1) = X^2(X-j)(X-j^2)$, les pôles de F sont $0, j, j^2$; 0 est pôles double et j, j^2 sont des pôles simples conjugués.
La décomposition en éléments simples de F est de la forme
$$\frac{1}{X^2(X^2+X+1)} = \frac{a}{X^2} + \frac{b}{X} + \frac{c}{X-j} + \frac{\bar{c}}{X-j^2}$$
car les parties polaires complexes de cette fraction réelle sont conjuguées.
On a
$$a = \left.\frac{1}{X^2+X+1}\right|_{X=0} = 1, b = \left.\left(\frac{1}{X^2+X+1}\right)'\right|_{X=0} = -\left.\frac{2X+1}{(X^2+X+1)^2}\right|_{X=0} = -1$$
et
$$c = \left.\frac{1}{X^2(X-j^2)}\right|_{X=j} = \frac{1}{j^2(j-j^2)} = \frac{1}{1-j} = \frac{1-j^2}{3}$$
Finalement
$$\frac{1}{X^2(X^2+X+1)} = \frac{1}{X^2} - \frac{1}{X} + \frac{1}{3}\frac{1-j^2}{X-j} + \frac{1}{3}\frac{1-j}{X-j^2}$$

9.2.4.6 Quelques astuces de calculs

On peut parfois obtenir rapidement une décomposition en éléments simples en écrivant astucieusement le numérateur…

Exemple En réécrivant le 1 du numérateur
$$\frac{1}{X(X+1)} = \frac{(X+1)-X}{X(X+1)} = \frac{1}{X} - \frac{1}{X+1}$$

Exemple Dans un esprit analogue
$$\frac{X^2}{(X+1)^3} = \frac{(X+1)^2-2X-1}{(X+1)^3} = \frac{1}{X+1} - \frac{2(X+1)-1}{(X+1)^2} = \frac{1}{X+1} - \frac{2}{(X+1)^2} + \frac{1}{(X+1)^3}$$

Quand une fraction est de degré strictement négatif, la limite $xF(x)$ quand $x \to +\infty$, permet d'obtenir une relation sur les coefficients des termes de la forme $1/(X-a)$ d'une décomposition en éléments simples ; c'est particulièrement efficace lorsqu'il y a un pôle double !

Exemple Considérons
$$F = \frac{X^2+X+1}{X(X+1)^2} \in \mathbb{C}(X)$$
F est écrite sous forme irréductible et sa partie entière est nulle car $\deg F < 0$.
Les pôles de F sont 0 et -1 ; 0 est pôle simple alors que -1 est pôle double.

La décomposition en éléments simples de F est de la forme

$$\frac{X^2+X+1}{X(X+1)^2}=\frac{a}{X}+\frac{b}{(X+1)^2}+\frac{c}{X+1}$$

On a

$$a=\left.\frac{X^2+X+1}{(X+1)^2}\right|_{X=0}=1 \text{ et } b=\left.\frac{X^2+X+1}{X}\right|_{X=-1}=-1$$

$xF(x)\xrightarrow[x\to+\infty]{}1$ donne la relation $a+c=1$, on en déduit $c=0$.
Finalement

$$\frac{X^2+X+1}{X(X+1)^2}=\frac{1}{X}-\frac{1}{(X+1)^2}$$

Exemple Considérons

$$F=\frac{3X+1}{(X-1)^2(X+1)}\in\mathbb{C}(X)$$

F est écrite sous forme irréductible et sa partie entière est nulle car $\deg F<0$.
Les pôles de F sont -1 et 1 ; -1 est pôle simple alors que 1 est pôle double.
La décomposition en éléments simples de F est de la forme

$$\frac{3X+1}{(X-1)^2(X+1)}=\frac{a}{X+1}+\frac{b}{X-1}+\frac{c}{(X-1)^2}$$

On a

$$a=\left.\frac{3X+1}{(X-1)^2}\right|_{X=-1}=-\frac{1}{2} \text{ et } b=\left.\frac{3X+1}{X+1}\right|_{X=1}=2$$

$xF(x)\xrightarrow[x\to+\infty]{}0$ donne la relation $a+c=0$, on en déduit $c=\dfrac{1}{2}$.
Finalement

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

Si la fraction rationnelle décomposée présente une parité, on retrouve celle-ci sur les éléments simples de sa décomposition…

Exemple Considérons

$$F=\frac{X}{(X^2-1)^2}\in\mathbb{C}(X)$$

F est écrite sous forme irréductible et sa partie entière est nulle car $\deg F<0$.
$(X^2-1)^2=(X-1)^2(X+1)^2$, les pôles de F sont -1 et 1 ; ce sont des pôles doubles.
La décomposition en éléments simples de F est de la forme

$$\frac{X}{(X+1)^2(X-1)^2}=\frac{a}{(X+1)^2}+\frac{b}{(X+1)}+\frac{c}{(X-1)^2}+\frac{d}{X-1}$$

Exploitons l'imparité de F en remplaçant X par $-X$

$$\frac{-X}{(-X+1)^2(-X-1)^2}=\frac{a}{(-X+1)^2}+\frac{b}{(-X+1)}+\frac{c}{(-X-1)^2}+\frac{d}{-X-1}$$

On en déduit

$$\frac{X}{(X+1)^2(X-1)^2} = \frac{-c}{(X+1)^2} + \frac{d}{(X+1)} + \frac{-a}{(X-1)^2} + \frac{b}{X-1}$$

Par unicité des coefficients d'une décomposition en éléments simples, on obtient
$a = -c$ et $b = d$.

$$a = \left.\frac{X}{(X-1)^2}\right|_{X=-1} = -\frac{1}{4}$$

donc $c = 1/4$
$xF(x) \xrightarrow[x\to+\infty]{} 0$ donne la relation $b + d = 0$ et on en déduit $b = d = 0$.
Finalement

$$\frac{X}{(X^2-1)^2} = \frac{1/4}{(X-1)^2} + \frac{1/4}{(X+1)^2}$$

Chapitre 10

Calcul matriciel

$\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$ et m, n, p, q, r désigne des naturels.
E, F, G désignent des $\mathbb{K}$-espace vectoriel de dimensions finies.
On introduit le symbole de Kronecker défini par

$$\delta_{i,j} = \begin{cases} 0 \text{ si } i \neq j \\ 1 \text{ si } i = j \end{cases}$$

Dans ce chapitre, on abandonne la notation fléchée des vecteurs.

10.1 Opérations sur les matrices

10.1.1 Matrice

Définition

On appelle matrice de type (n, p) (pour n lignes et p colonnes) à coefficients dans $\mathbb{K}$ tout famille $A = (a_{i,j})_{1 \leqslant i \leqslant n, 1 \leqslant j \leqslant p}$ d'éléments de $\mathbb{K}$.
Une telle matrice est généralement représentée sous la forme d'un tableau à n lignes et p colonnes :

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,p} \\ a_{2,1} & a_{2,2} & \dots & a_{2,p} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,p} \end{pmatrix}$$

Le terme $a_{i,j}$ est appelé coefficient d'indice (i, j) de la matrice A, il est positionné à la i-ème ligne et j-ème colonne dans le tableau représentant A.

Exemple Pour

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix}$$

$A = (a_{i,j})_{1 \leqslant i \leqslant 3, 1 \leqslant j \leqslant 4}$ avec $a_{1,1} = 1, a_{2,3} = 7, a_{3,1} = 9$ et plus généralement $a_{i,j} = 4(i-1) + j$.

Définition

On note $\mathcal{M}_{n,p}(\mathbb{K})$ l'ensemble des matrices de type (n, p) à coefficients dans $\mathbb{K}$.

Remarque Deux matrices de $\mathcal{M}_{n,p}(\mathbb{K})$ sont égales si, et seulement si, elles sont constituées des mêmes coefficients.

Remarque Le 1er indice est toujours l'indice de ligne.
Le 2nd indice est toujours l'indice de colonne.
Quand il n'y a pas d'ambiguïtés sur les indices, il est fréquent de présenter une matrice en écrivant seulement $A = (a_{i,j}) \in \mathcal{M}_{n,p}$. Il est alors entendu que le premier indice, ici i, varie entre 1 et n et le second, j, varie entre 1 et p.
S'il y a ambiguïté sur le rôle des indices, on peut écrire $A = (a_{i,j})_{i,j} \in \mathcal{M}_{n,p}(\mathbb{K})$ pour signifier que l'indice de ligne est i et celui de colonne est j.

Exemple Pour $A = (ij^2)_{i,j} \in \mathcal{M}_{n,p}(\mathbb{K})$ on a

$$A = \begin{pmatrix} 1 & 4 & 9 & \dots & p^2 \\ 2 & 8 & 18 & \dots & 2p^2 \\ 3 & 12 & 27 & \dots & 3p^2 \\ \vdots & \vdots & \vdots & & \vdots \\ n & 4n & 9n & \dots & np^2 \end{pmatrix}$$

Exemple La matrice $O_{n,p} = (0)_{i,j} \in \mathcal{M}_{n,p}(\mathbb{K})$ est appelée matrice nulle de type (n, p).

$$O_{n,p} = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{pmatrix}$$

Définition

Pour $n = p = 1$, les matrices de $\mathcal{M}_{1,1}(\mathbb{K})$ sont appelées matrices (uni-) coefficient. Elles sont de la forme (x).
Il est usuel d'identifier cette matrice et son coefficient x élément de $\mathbb{K}$.

10.1.2 Lignes et colonnes

Définition

Pour n quelconque et $p = 1$, les matrices de $\mathcal{M}_{n,1}(\mathbb{K})$ sont appelées matrices (uni-) colonnes. Elles sont de la forme

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

Il est usuel d'identifier cette matrice colonne avec le n uplet $(a_1, \dots, a_n)$ élément de $\mathbb{K}^n$.
Pour $n = 1$ et p quelconque, les matrices de $\mathcal{M}_{1,p}(\mathbb{K})$ sont appelées matrices (uni-) lignes. Elles sont de la forme

$$\begin{pmatrix} a_1 & \cdots & a_p \end{pmatrix}$$

Définition

Soit $A = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{K})$ (présentation de l'abus de notation correspondant).

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,p} \\ a_{2,1} & a_{2,2} & \dots & a_{2,p} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,p} \end{pmatrix}$$

Pour $1 \leqslant j \leqslant p$, la matrice

$$C_j = \begin{pmatrix} a_{1,j} \\ \vdots \\ a_{n,j} \end{pmatrix}$$

est appelée j-ème colonne de A.
Pour $1 \leqslant i \leqslant n$, la matrice

$$L_i = \begin{pmatrix} a_{i,1} & \dots & a_{i,p} \end{pmatrix}$$

est appelée i-ème ligne de A.

10.1.3 Matrice carrée

10.1.3.1 Définition

Définition

Les matrices de type (n, n) sont appelées matrices carrées d'ordre n.
On note $\mathcal{M}_n(\mathbb{K})$, au lieu de $\mathcal{M}_{n,n}(\mathbb{K})$, l'ensemble de ces matrices.

Exemple Une matrice carrée d'ordre n est de la forme

$$A = (a_{i,j})_{1 \leqslant i,j \leqslant n} = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{pmatrix}$$

Exemple Pour $A = (\min(i,j))_{1 \leqslant i,j \leqslant n} \in \mathcal{M}_n(\mathbb{K})$,

$$A = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 2 & 2 & \dots & 2 \\ 1 & 2 & 3 & \dots & 3 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 2 & 3 & \dots & n \end{pmatrix}$$

10.1.3.2 Matrice diagonale

Définition

Soit $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$.
Les coefficients d'indice (i, i) de A sont appelées coefficients diagonaux de A.
La famille $(a_{1,1}, a_{2,2}, \dots, a_{n,n}) = (a_{i,i})_{1 \leqslant i \leqslant n}$ est appelée diagonale de la matrice A.

Définition

Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est dite diagonale si tous ses coefficients hors de la diagonale sont nuls.
On note $D_n(\mathbb{K})$ l'ensemble de ces matrices.

Exemple Les matrices diagonales de taille sont de la forme

$$D = \begin{pmatrix} a_{1,1} & & 0 \\ & \ddots & \\ 0 & & a_{n,n} \end{pmatrix}$$

Remarque Pour $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$, on a

$$A \in D_n(\mathbb{K}) \Leftrightarrow \forall 1 \leqslant i \neq j \leqslant n, a_{i,j} = 0$$

Définition

Pour $\lambda_1, \dots, \lambda_n \in \mathbb{K}$, on note diag$(\lambda_1, \dots, \lambda_n)$ la matrice diagonale dont la diagonale est la famille $(\lambda_1, ..., \lambda_n)$ c'est-à-dire la matrice

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

Exemple On note

$$I = I_n = \text{diag}(1, 1, \dots, 1) = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$$

appelée matrice identité.

Remarque Le coefficient d'indice (i, j) de la matrice I_n est $\delta_{i,j}$.
Ainsi $I_n = (\delta_{i,j})_{1 \leqslant i,j \leqslant n}$

10.1.3.3 Matrice triangulaire

Définition

Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est dite triangulaire supérieure (resp. inférieure) si tous les coefficients en dessous (resp. au dessus) de la diagonale sont nuls.
On note $T_n^+(\mathbb{K})$ (resp. $T_n^-(\mathbb{K})$) l'ensemble de ces matrices.

Exemple Les matrices triangulaire supérieures de taille n sont de la forme

$$\begin{pmatrix} a_{11} & & \star \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix}$$

Dans cette écriture, l'étoile signifie « des coefficients quelconques).
Les matrices triangulaires inférieures de taille n sont de la forme

$$\begin{pmatrix} a_{11} & & 0 \\ & \ddots & \\ \star & & a_{nn} \end{pmatrix}$$

Remarque Pour $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$, on a les équivalences

$$A \in T_n^+(\mathbb{K}) \Leftrightarrow \forall 1 \leqslant j < i \leqslant n, a_{i,j} = 0 \; A \in T_n^-(\mathbb{K}) \Leftrightarrow \forall 1 \leqslant i < j \leqslant n, a_{i,j} = 0$$

Proposition

$D_n(\mathbb{K}) = T_n^+(\mathbb{K}) \cap T_n^-(\mathbb{K})$.

10.1.4 Espace vectoriel $(\mathcal{M}_{n,p}(\mathbb{K}), +, .)$

10.1.4.1 Opérations

Définition

Soit $A = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B = (b_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{K})$.
On définit la matrice $A + B \in \mathcal{M}_{n,p}(\mathbb{K})$ par $A + B = (a_{i,j} + b_{i,j})_{1 \leqslant i \leqslant n, 1 \leqslant j \leqslant p}$.
Ainsi

$$\begin{pmatrix} a_{1,1} & \dots & a_{1,p} \\ \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,p} \end{pmatrix} + \begin{pmatrix} b_{1,1} & \dots & b_{1,p} \\ \vdots & & \vdots \\ b_{n,1} & \dots & b_{n,p} \end{pmatrix} = \begin{pmatrix} a_{1,1} + b_{1,1} & \dots & a_{1,p} + b_{1,p} \\ \vdots & & \vdots \\ a_{n,1} + b_{n,1} & \dots & a_{n,p} + b_{n,p} \end{pmatrix}$$

Attention : On ne somme que des matrices de même type.

Définition

Soit $A = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{K})$ et $\lambda \in \mathbb{K}$.
On définit la matrice $\lambda.A \in \mathcal{M}_{n,p}(\mathbb{K})$ par $\lambda.A = (\lambda a_{i,j})_{1 \leqslant i \leqslant n, 1 \leqslant j \leqslant p}$.
Ainsi

$$\lambda. \begin{pmatrix} a_{1,1} & \dots & a_{1,p} \\ \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,p} \end{pmatrix} = \begin{pmatrix} \lambda a_{1,1} & \dots & \lambda a_{1,p} \\ \vdots & & \vdots \\ \lambda a_{n,1} & \dots & \lambda a_{n,p} \end{pmatrix}$$

Théorème

$(\mathcal{M}_{n,p}(\mathbb{K}), +, .)$ est un $\mathbb{K}$-espace vectoriel d'élément nul $O = O_{n,p}$.

dém. :
En identifiant une matrice $A = (a_{i,j})_{1 \leqslant i \leqslant n, 1 \leqslant j \leqslant p} \in \mathcal{M}_{n,p}(\mathbb{K})$ avec le multi-uplet

$$(a_{1,1}, \dots, a_{1,p}, a_{2,1}, \dots, a_{2,p}, \dots, a_{n,1}, \dots, a_{n,p}) \in \mathbb{K}^{np}$$

correspondant à une lecture continue des lignes, on observe que les opérations précédemment définies correspondent aux opérations sur $\mathbb{K}^{np}$. Or $(\mathbb{K}^{np},+,.)$ est un $\mathbb{K}$-espace vectoriel de vecteur nul le mutli-uplet nul, donc $(\mathcal{M}_{n,p}(\mathbb{K}),+,.)$ est un $\mathbb{K}$-espace vectoriel d'élément nul $O = O_{n,p}$.
□

Exemple On a

$$a.\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b.\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} + c.\begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} = \begin{pmatrix} a+c & -b+c \\ -b-c & a-c \end{pmatrix}$$

10.1.4.2 Dimension

Définition

Soient $1 \leqslant k \leqslant n$ et $1 \leqslant \ell \leqslant p$. On appelle matrice élémentaire d'indice (k,ℓ) de $\mathcal{M}_{n,p}(\mathbb{K})$, la matrice $E_{k,\ell}$ dont tous les coefficients sont nuls sauf celui d'indice (k,ℓ) qui vaut 1. Ainsi

$$E_{k,\ell} = \begin{pmatrix} 0 & & 0 \\ & 1 & \\ 0 & & 0 \end{pmatrix} \in \mathcal{M}_{n,p}(\mathbb{K})$$

Exemple Dans $\mathcal{M}_2(\mathbb{K})$, les matrices élémentaires sont

$$E_{1,1} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, E_{1,2} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, E_{2,1} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \text{ et } E_{2,2} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Exemple Dans $\mathcal{M}_{n,1}(\mathbb{K})$, les colonnes élémentaires sont

$$E_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, E_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

Remarque Le coefficient d'indice (i,j) de la matrice élémentaire $E_{k,\ell} \in \mathcal{M}_{n,p}(\mathbb{K})$ est $\delta_{(i,j),(k,\ell)} = \delta_{i,k}\delta_{j,\ell}$.

Théorème

La famille $\mathcal{B} = (E_{k,\ell})_{1\leqslant k\leqslant n, 1\leqslant \ell\leqslant p}$ est une base de $\mathcal{M}_{n,p}(\mathbb{K})$.
On l'appelle base canonique de l'espace $\mathcal{M}_{n,p}(\mathbb{K})$.

dém. :
Pour toute matrice $A = (a_{i,j})_{1\leqslant i\leqslant n, 1\leqslant j\leqslant p} \in \mathcal{M}_{n,p}(\mathbb{K})$, on peut écrire

$$A = \sum_{k=1}^{n}\sum_{\ell=1}^{p} a_{k,\ell}E_{k,\ell}$$

La famille $\mathcal{B}$ est donc génératrice.
Supposons $\sum_{k=1}^{n}\sum_{\ell=1}^{p}\lambda_{k,\ell}E_{k,\ell}=O_{n,p}$ avec $\lambda_{k,\ell}\in\mathbb{K}$.
On a alors

$$\begin{pmatrix}\lambda_{1,1} & \cdots & \lambda_{1,p}\\ \vdots & & \vdots\\ \lambda_{n,1} & \cdots & \lambda_{n,p}\end{pmatrix}=\begin{pmatrix}0 & \cdots & 0\\ \vdots & & \vdots\\ 0 & \cdots & 0\end{pmatrix}$$

et par identification des coefficients d'une matrice, on obtient $\lambda_{k,\ell}=0$ pour tous indices k,ℓ.
Ainsi la famille $\mathcal{B}$ est libre.
□

Corollaire

$\dim\mathcal{M}_{n,p}(\mathbb{K})=np$.
En particulier $\dim\mathcal{M}_n(\mathbb{K})=n^2$, $\dim\mathcal{M}_{n,1}(\mathbb{K})=n$ et $\dim\mathcal{M}_{1,p}(\mathbb{K})=p$.

10.1.4.3 Sous-espaces des matrices diagonales et triangulaires

Proposition

$D_n(\mathbb{K})$ est un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{K})$ de dimension n.

dém. :
Par définition

$$D_n(\mathbb{K})=\{\text{diag}(\lambda_1,\dots,\lambda_n)/\lambda_i\in\mathbb{K}\}$$

On a donc

$$D_n(\mathbb{K})=\{\lambda_1E_{1,1}+\cdots+\lambda_nE_{n,n}/\lambda_i\in\mathbb{K}\}=\text{Vect}(E_{1,1},\dots,E_{n,n})$$

$D_n(\mathbb{K})$ est donc un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{K})$ et la famille $(E_{1,1},\dots,E_{n,n})$ en est une base.
□

Proposition

$T_n^+(\mathbb{K})$ et $T_n^-(\mathbb{K})$ sont des sous-espaces vectoriels de $\mathcal{M}_n(\mathbb{K})$ de dimension $\dfrac{n(n+1)}{2}$.

dém. :
On observe

$$T_n^+(\mathbb{K})=\text{Vect}\left\{E_{i,j}/1\leqslant i\leqslant j\leqslant n\right\}\text{ et }T_n^-(\mathbb{K})=\text{Vect}\left\{E_{i,j}/1\leqslant j\leqslant i\leqslant b\right\}$$

ce qui permet de conclure.
□

10.1.4.4 Quelques exemples de manipulations vectorielles

Exemple Soient

$$A_1=\begin{pmatrix}1&0\\0&1\end{pmatrix},A_2=\begin{pmatrix}1&0\\0&-1\end{pmatrix},A_3=\begin{pmatrix}1&1\\1&1\end{pmatrix},A_4=\begin{pmatrix}0&-1\\1&0\end{pmatrix}$$

Montrons que la famille $(A_i)_{1\leqslant i\leqslant 4}$ est une base de l'espace $\mathcal{M}_2(\mathbb{R})$.
Les éléments de la famille $(A_i)_{1\leqslant i\leqslant 4}$ sont des matrices éléments de $\mathcal{M}_2(\mathbb{R})$ et il y en a exactement $4=\dim\mathcal{M}_2(\mathbb{R})$. Pour conclure il suffit, par exemple, d'étudier la liberté de cette famille.
Supposons $\lambda_1A_1+\lambda_2A_2+\lambda_3A_3+\lambda_4A_4=O_{2,2}$.

Matriciellement

$$\begin{pmatrix} \lambda_1+\lambda_2+\lambda_3 & \lambda_3-\lambda_4 \\ \lambda_3+\lambda_4 & \lambda_1+\lambda_2+\lambda_3 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

En identifiant les coefficients et en résolvant le système obtenu par les équations formées, on obtient $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0$ ce qui permet de conclure.

Exemple Montrons que

$$F = \left\{ \begin{pmatrix} a+b & -a+b \\ 2a+b & -a+2b \end{pmatrix} / a, b \in \mathbb{K} \right\}$$

est un sous-espace vectoriel de $\mathcal{M}_2(\mathbb{K})$.
On a

$$F = \left\{ a \begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix} + b \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} / a, b \in \mathbb{K} \right\}$$

donc

$$F = \text{Vect}\left\{ \begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \right\}$$

On en déduit que F est un sous-espace vectoriel de $\mathcal{M}_2(\mathbb{K})$ et on peut même préciser qu'il est de dimension 2.

Exemple Soit

$$H = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbb{K}) / a+b+c+d = 0 \right\}$$

Montrons que H est un hyperplan de $\mathcal{M}_2(\mathbb{K})$.
Considérons l'application $\varphi : \mathcal{M}_2(\mathbb{K}) \to \mathbb{K}$ définie par

$$\varphi\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix} \right) = a+b+c+d$$

On vérifie aisément que φ est une forme linéaire sur $\mathcal{M}_2(\mathbb{K})$ et que celle-ci est non nulle. On en déduit que $H = \ker \varphi$ est un hyperplan de $\mathcal{M}_2(\mathbb{K})$.

10.1.5 Produit matriciel

Définition

Soient $A = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B = (b_{j,k}) \in \mathcal{M}_{p,q}(\mathbb{K})$.
On définit la matrice $C = A \times B = (c_{i,k}) \in \mathcal{M}_{n,q}(\mathbb{K})$ par

$$\forall 1 \leqslant i \leqslant n, \forall 1 \leqslant k \leqslant q, c_{i,k} = \sum_{j=1}^{p} a_{i,j} b_{j,k}$$

Remarque Pour calculer la matrice C il est usuel de disposer les matrices comme ci-dessous

$$\begin{array}{cc} & B \\ A & C \end{array}$$

Pour déterminer le coefficient d'indice (i, k) de la matrice C, on repère la $i$ème ligne de A et la $k$ème colonne de B comme ci-dessous

$$\begin{array}{cc} & \overbrace{\begin{pmatrix} \quad b_{1,k} \quad \\ \vdots \\ b_{p,k} \end{pmatrix}}^{q} \\ n\left\{ \begin{pmatrix} a_{i,1} & \dots & a_{i,p} \end{pmatrix} \right. & \begin{pmatrix} \quad \\ \quad c_{i,k} \quad \\ \quad \end{pmatrix} \end{array}$$

On peut alors évaluer $c_{i,k}$ en procédant à la somme des produits des coefficients respectifs.

Attention : Pour que cette multiplication matricielle soit possible, il est nécessaire que le nombre de colonnes de A soit égal au nombre de lignes de B.
De plus, on peut retenir
type $(n, p) \times$ type (p, q) = type (n, q).

Exemple Pour

$$A = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} \text{ et } B = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & -1 \end{pmatrix}$$

on obtient

$$A \times B = \begin{pmatrix} 5 & 2 & -1 \\ -1 & 1 & -2 \end{pmatrix}$$

Exemple Pour

$$A = \begin{pmatrix} 2 & 1 & -1 \\ 0 & 1 & 2 \end{pmatrix} \text{ et } B = \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 0 & 1 \end{pmatrix}$$

on obtient

$$A \times B = \begin{pmatrix} 4 & 2 \\ 2 & 3 \end{pmatrix}$$

Exemple Pour

$$A = \begin{pmatrix} a_{1,1} & ... & a_{1,p} \\ \vdots & & \vdots \\ a_{n,1} & ... & a_{n,p} \end{pmatrix} \text{ et } X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$$

on obtient

$$AX = \begin{pmatrix} a_{1,1}x_1 + ... + a_{1,p}x_p \\ \vdots \\ a_{n,1}x_1 + ... + a_{n,p}x_p \end{pmatrix}$$

10.1.6 Propriétés du produit matriciel

Remarque Si les types de A et B permettent de calculer AB et BA alors, en général, $AB \neq BA$.
Par exemple

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

alors que

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} !$$

Proposition

Pour tout $A \in \mathcal{M}_{n,p}(\mathbb{K}), B \in \mathcal{M}_{p,q}(\mathbb{K}), C \in \mathcal{M}_{q,r}(\mathbb{K})$, on a

$$(AB)C = A(BC)$$

dém. :
On introduit les coefficients des matrices A, B, C
$A = (a_{i,j}), B = (b_{j,k}), C = (c_{k,\ell})$
en convenant de noter les indices en fonction du domaine où ceux-ci varient de sorte que $1 \leqslant i \leqslant n, 1 \leqslant j \leqslant p, 1 \leqslant k \leqslant q, 1 \leqslant \ell \leqslant r$
Posons $D = AB = (d_{i,k})$ et $E = (AB)C = DC = (e_{i,\ell})$.
On a $d_{i,k} = \sum_{j=1}^{p} a_{i,j}b_{j,k}$ et $e_{i,\ell} = \sum_{k=1}^{n} d_{i,k}c_{k,\ell} = \sum_{j=1}^{p}\sum_{k=1}^{q} a_{i,j}b_{j,k}c_{k,\ell}$.
Posons $F = BC = (f_{j,\ell})$ et $G = A(BC) = AF = (g_{i,\ell})$.
On a $f_{j,\ell} = \sum_{k=1}^{q} b_{j,k}c_{k,\ell}$ et $g_{i,\ell} = \sum_{j=1}^{p} a_{i,j}f_{j,\ell} = \sum_{j=1}^{p}\sum_{k=1}^{q} a_{i,j}b_{j,k}c_{k,\ell}$.
En réorganisant l'ordre des sommes, on obtient $g_{i,\ell} = e_{i,\ell}$ pour tous indices i, ℓ.
Prop :$\forall A \in \mathcal{M}_{n,p}(\mathbb{K})$, $AI_p = A$ et $I_nA = A$.
□
dém. :
Notons $a_{i,j}$ le coefficient général de la matrice A.
Le produit AI_p est possible et si l'on note $b_{i,j}$ le coefficient d'indice (i, j) de ce calcul, on a

$$b_{i,j} = \sum_{k=1}^{p} a_{i,k}\delta_{k,j} = a_{i,j}$$

Ainsi $AI_p = A$.
L'étude de I_nA est similaire.
□
Proposition

$\forall A, B \in \mathcal{M}_{n,p}(\mathbb{K})$, $\forall C \in \mathcal{M}_{p,q}(\mathbb{K})$, $(A + B)C = AC + BC$.
$\forall A \in \mathcal{M}_{n,p}(\mathbb{K})$, $\forall B, C \in \mathcal{M}_{p,q}(\mathbb{K})$, $A(B + C) = AB + AC$.

dém. :
Soient $A = (a_{i,j}), B = (b_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{K})$ et $C = (c_{j,k}) \in \mathcal{M}_{p,q}(\mathbb{K})$.

Les produits $(A+B)C$, AC et BC sont possibles et

$$(A+B)C = \left(\sum_{j=1}^{p}(a_{i,j}+b_{i,j})c_{j,k}\right)_{i,k},\ AC+BC = \left(\sum_{j=1}^{p}a_{i,j}c_{j,k}+\sum_{j=1}^{p}b_{i,j}c_{j,k}\right)_{i,k}$$

Par égalité des coefficients respectifs, on obtient $(A+B)C = AC+BC$.
L'étude de l'identité $A(B+C) = AB+AC$ est similaire.
□

Proposition

$\forall A\in\mathcal{M}_{n,p}(\mathbb{K}), \forall B\in\mathcal{M}_{p,q}(\mathbb{K}), \forall\lambda\in\mathbb{K}$, $(\lambda.A)B = \lambda.(AB) = A(\lambda.B)$.

dém. :
Soient $A = (a_{i,j})\in\mathcal{M}_{n,p}(\mathbb{K})$ et $B = (b_{j,k})\in\mathcal{M}_{p,q}(\mathbb{K})$.
Les coefficients d'indice (i,k) des matrices $(\lambda.A)B$, $\lambda.(AB)$ et $A(\lambda.B)$ sont respectivement

$$\sum_{j=1}^{p}(\lambda a_{i,j})b_{j,k},\ \lambda\sum_{j=1}^{p}a_{i,j}b_{j,k} \text{ et } \sum_{j=1}^{p}a_{i,j}(\lambda b_{j,k})$$

Ces coefficients sont égaux et donc $(\lambda.A)B = \lambda.(AB) = A(\lambda.B)$
□

10.1.7 L'anneau $(\mathcal{M}_n(\mathbb{K}),+,\times)$

10.1.7.1 Présentation

Remarque Le produit matriciel définit une loi de composition interne sur $\mathcal{M}_n(\mathbb{K})$.

Théorème

$(\mathcal{M}_n(\mathbb{K}),+,\times)$ est un anneau, généralement non commutatif, d'élément nul $O = O_n$ et d'élément unité $I = I_n$.
De plus, $\forall\lambda\in\mathbb{K},\forall A,B\in\mathcal{M}_n(\mathbb{K}) : (\lambda.A)B = \lambda.(AB) = A(\lambda.B)$.

dém. :
On sait déjà que $(\mathcal{M}_n(\mathbb{K}),+)$ est un groupe abélien de neutre O_n.
Par les propriétés du produit matriciel ci-dessus établies, on peut affirmer que la multiplication est associative, que la matrice I_n est neutre et que la multiplication est distributive sur l'addition. On en déduit que $(\mathcal{M}_n(\mathbb{K}),+,\times)$ est un anneau.
□

Remarque Pour $n=1$, l'anneau $(\mathcal{M}_1(\mathbb{K}),+,\times)$ s'identifie avec $(\mathbb{K},+,\times)$, c'est donc un corps.

Remarque Pour $n\geqslant 2$, l'anneau $\mathcal{M}_n(\mathbb{K})$ n'est pas commutatif.

Définition

On dit que deux matrices A et B de $\mathcal{M}_n(\mathbb{K})$ commutent si $AB = BA$.

Exemple Les matrices $A\in\mathcal{M}_n(\mathbb{K})$ et I_n commutent car $A\times I_n = A = I_n\times A$.

10.1.7.2 Puissances d'une matrice

Définition

> Pour $A \in \mathcal{M}_n(\mathbb{K})$, on note $A^0 = I_n, A^1 = A, A^2 = A \times A, \dots, A^m = A \times A \times \cdots \times A$ (m termes)

Exemple Calculons les puissance de

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

Par le calcul, on observer

$$A^2 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, A^3 = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$$

...
Par récurrence, on montre aisément

$$A^m = \begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix}$$

Attention : $(A+B)^2 = A^2 + AB + BA + B^2$.
$(A+B)^3 = A^3 + A^2B + ABA + BA^2 + AB^2 + BAB + B^2A + B^3$.

Théorème

> Si A et B commutent alors pour tout $m \in \mathbb{N}$,
>
> $$(AB)^m = A^m B^m,$$
>
> $$(A+B)^m = \sum_{k=0}^{m} \binom{m}{k} A^k B^{m-k} \text{ et}$$
>
> $$A^m - B^m = (A-B) \sum_{k=0}^{m-1} A^k B^{m-1-k}$$

dém. :
En vertu des règles de calculs dans un anneau...
□

Exemple Puisque les matrices $A \in \mathcal{M}_n(\mathbb{K})$ et I_n commutent, on a

$$(A+I_n)^m = \sum_{k=0}^{m} \binom{m}{k} A^k \text{ et } A^m - I_n = (A - I_n) \sum_{k=0}^{m-1} A^k$$

Attention : Ne pas écrire

$$\sum_{k=0}^{m-1} A^k = \frac{A^m - I}{A - I}$$

il n'y a pas de division matricielle !

Remarque Pour $n \geqslant 2$, l'anneau $\mathcal{M}_n(\mathbb{K})$ possède des diviseurs de zéro.
Ainsi : $AB = 0$ n'implique pas $A = 0$ ou $B = 0$.
En conséquence, l'équation $A^2 = I_2$ possède dans $\mathcal{M}_2(\mathbb{K})$ d'autres solutions que les matrices I_2 et $-I_2$ comme par exemple :

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 2 \\ 0 & 1 \end{pmatrix}, \dots$$

Définition

On dit qu'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est idempotente si $A^2 = A$.

Exemple La matrice suivante est idempotente

$$A = \begin{pmatrix} 2 & -1 \\ 2 & -1 \end{pmatrix}$$

Définition

On dit qu'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est nilpotente s'il existe $m \in \mathbb{N}$ tel que $A^m = O_n$.

Exemple La matrice

$$B = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{pmatrix}$$

est nilpotente.
En effet

$$B^2 = \begin{pmatrix} 0 & 0 & 8 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \text{ et } B^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Plus généralement, on peut montrer qu'une matrice triangulaire supérieure de diagonale nulle est nilpotente.

Exemple Calculons les puissances de

$$A = \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix} \in \mathcal{M}_3(\mathbb{R})$$

On peut écrire $A = I_3 + B$ avec

$$B = \begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{pmatrix}, B^2 = \begin{pmatrix} 0 & 0 & b(a+c) \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, B^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Par la formule du binôme

$$A^m = (I_3 + B)^m = \sum_{k=0}^{m} \binom{m}{k} B^k = \binom{m}{0} I_3 + \binom{m}{1} B + \binom{m}{2} B^2$$

car $B^k = O_3$ pour tout $k \geqslant 3$.
Ainsi

$$A^m = \begin{pmatrix} 1 & ma & mb + \dfrac{m(m-1)}{2}b(a+c) \\ 0 & 1 & mc \\ 0 & 0 & 1 \end{pmatrix}$$

10.1.7.3 Matrices inversibles

Définition

Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est dite inversible s'il existe $B \in \mathcal{M}_n(\mathbb{K})$ vérifiant $AB = BA = I_n$.
Cette matrice B est alors unique, c'est l'inverse de A noté A^{-1}.

Exemple La matrice I_n est inversible et $I_n^{-1} = I_n$.

Proposition

Soient $A, B \in \mathcal{M}_n(\mathbb{K})$
Si A et B sont inversibles alors AB l'est aussi et $(AB)^{-1} = B^{-1}A^{-1}$.
Si A est inversible alors A^{-1} l'est aussi et $(A^{-1})^{-1} = A$.

dém. :
Ce sont des propriétés relatives aux éléments inversibles d'un anneau.
□

Définition

On note $\mathrm{GL}_n(\mathbb{K})$ l'ensemble des matrices inversibles de $\mathcal{M}_n(\mathbb{K})$.

Proposition

$(GL_n(\mathbb{K}), \times)$ est un groupe appelé groupe linéaire d'ordre n.

dém. :
C'est le groupe des éléments inversibles de l'anneau $(\mathcal{M}_n(\mathbb{K}), +, \times)$.
□

Théorème

Pour $A \in M_n(\mathbb{K})$, on a équivalence entre
(i) A est inversible ;
(ii) A est inversible à droite i.e. $\exists B \in \mathcal{M}_n(\mathbb{K}), AB = I_n$;
(iii) A est inversible à gauche i.e. $\exists C \in \mathcal{M}_n(\mathbb{K}), CA = I_n$.
De plus si tel est le cas $A^{-1} = B = C$.

dém. :
(i) $\Rightarrow$ (ii) et (iii) immédiat.
(ii) $\Rightarrow$ (i) Supposons qu'il existe $B \in \mathcal{M}_n(\mathbb{K})$ vérifiant $AB = I_n$.
Soit $\varphi : \mathcal{M}_n(\mathbb{K}) \to \mathcal{M}_n(\mathbb{K})$ l'application définie par $\varphi(M) = MA$.
Il est immédiat de vérifier que φ est un endomorphisme de $\mathcal{M}_n(\mathbb{K})$.
Soit $M \in \ker \varphi$. On a $MA = O_n$ donc $MAB = O_n \times B$ puis $M = O_n$ car $AB = I_n$.
Par suite $\ker \varphi = \{O_n\}$.
Ainsi φ est un endomorphisme injectif de $\mathcal{M}_n(\mathbb{K})$, or $\dim \mathcal{M}_n(\mathbb{K}) < +\infty$, donc φ est un automorphisme de $\mathcal{M}_n(\mathbb{K})$. Par surjectivité, il existe $C \in \mathcal{M}_n(\mathbb{K})$ vérifiant $CA = I_n$ et alors $CAB = B$ d'où $C = B$

car $AB = I_n$.
Finalement $AB = BA = I_n$ et donc A est inversible d'inverse B.
(iii) $\Rightarrow$ (i) s'obtient de façon semblable en introduisant l'endomorphisme $\psi : M \to AM$.
□

Exemple Soit

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

On vérifie par le calcul que

$$A^2 - 5A = 2I_2$$

Par suite

$$A\left(\frac{1}{2}A - \frac{5}{2}I_2\right) = I_2$$

Par le théorème d'inversibilité, on peut conclure que A est inversible et

$$A^{-1} = \frac{1}{2}A - \frac{5}{2}I$$

10.1.7.4 Détermination pratique de l'inverse d'une matrice carrée inversible

Lemme

Soient $A, B \in \mathcal{M}_{n,p}(\mathbb{K})$.
Si $AX = BX$ pour toute colonne $X \in \mathcal{M}_{p,1}(\mathbb{K})$ alors $A = B$

dém. :
Supposons $AX = BX$ pour toute colonne $X \in \mathcal{M}_{p,1}(\mathbb{K})$.
Pour $X = E_j$ colonne élémentaire, le produit AE_j est égale à la $j$ème colonne de A. Or $AE_j = BE_j$ donc A et B admettent la même $j$ème colonne. Ainsi, colonne par colonne, les matrices A et B sont égales.
Comment établir l'inversibilité et calculer l'inverse de $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$?
On introduit

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbb{K}) \text{ et } Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = AX \in \mathcal{M}_{n,1}(\mathbb{K})$$

On a

$$\begin{cases} y_1 = a_{1,1}x_1 + \cdots + a_{1,n}x_n \\ \vdots \\ y_n = a_{n,1}x_1 + \cdots + a_{n,n}x_n \end{cases}$$

Si cela est possible, on résout ce système en les inconnues $x_1, ..., x_n$ et on obtient

$$\begin{cases} x_1 = b_{1,1}y_1 + \cdots + b_{1,n}y_n \\ \vdots \\ x_n = b_{n,1}y_1 + \cdots + b_{n,n}y_n \end{cases}$$

Soit B la matrice dont les coefficients apparaissent dans ce système c'est-à-dire $B = (b_{i,j}) \in \mathcal{M}_n(\mathbb{K})$.
Le système précédent donne $X = BY$ et ainsi $X = BAX$ et ce pour tout $X \in \mathcal{M}_{n,1}(\mathbb{K})$.

En vertu du lemme d'identification matricielle, on peut affirmer $BA = I_n$. Par le théorème d'inversibilité, on peut alors affirmer que A est inversible et que B est son inverse.
□

Remarque Si on ne parvient pas à résoudre le système, c'est que la matrice A n'est pas inversible…

Exemple Etudions l'inversibilité de

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \in \mathcal{M}_3(\mathbb{R})$$

Soient

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \text{ et } Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = AX$$

On a

$$\begin{cases} y_1 = x_2 + x_3 \\ y_2 = x_1 + x_3 \\ y_3 = x_1 + x_2 \end{cases}$$

En inversant ce système on obtient

$$\begin{cases} x_1 = (-y_1 + y_2 + y_3)/2 \\ x_2 = (y_1 - y_2 + y_3)/2 \\ x_3 = (y_1 + y_2 - y_3)/2 \end{cases}$$

On en déduit que A est inversible et

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

Exemple Etudions l'inversibilité de

$$A = \begin{pmatrix} 1 & -a & & (0) \\ & \ddots & \ddots & \\ & & \ddots & -a \\ (0) & & & 1 \end{pmatrix} \in \mathcal{M}_n(\mathbb{K})$$

(avec $a \in \mathbb{K}$)
Soient

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \text{ et } Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = AX$$

On a

$$\begin{cases} y_1 = x_1 - ax_2 \\ y_2 = x_2 - ax_3 \\ \quad\vdots \\ y_n = x_n \end{cases}$$

En inversant ce système on obtient

$$\begin{cases} x_1 = y_1 + ay_2 + \cdots + a^{n-1}y_n \\ \quad\vdots \\ x_{n-1} = y_{n-1} + ay_n \\ x_n = y_n \end{cases}$$

On en déduit que A est inversible et

$$A^{-1} = \begin{pmatrix} 1 & a & a^2 & \dots & a^{n-1} \\ & 1 & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & a^2 \\ & & & \ddots & a \\ (0) & & & & 1 \end{pmatrix}$$

10.1.7.5 Anneau des matrices diagonales

Proposition

$D_n(\mathbb{K})$ est un sous-anneau commutatif de $\mathcal{M}_n(\mathbb{K})$.

dém. :
$D_n(\mathbb{K}) \subset \mathcal{M}_n(\mathbb{K})$ et $I_n \in D_n(\mathbb{K})$.
Soient

$$A = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \text{ et } B = \begin{pmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_n \end{pmatrix} \in D_n(\mathbb{K})$$

On a

$$A - B = \begin{pmatrix} \lambda_1 - \mu_1 & & \\ & \ddots & \\ & & \lambda_n - \mu_n \end{pmatrix} \in D_n(\mathbb{K}) \text{ et } AB = \begin{pmatrix} \lambda_1\mu_1 & & \\ & \ddots & \\ & & \lambda_n\mu_n \end{pmatrix} \in D_n(\mathbb{K})$$

□

Exemple Pour

$$A = \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix} \in D_n(\mathbb{K})$$

on a

$$A^2 = \begin{pmatrix} a_1^2 & & \\ & \ddots & \\ & & a_n^2 \end{pmatrix}, A^3 = \begin{pmatrix} a_1^3 & & \\ & \ddots & \\ & & a_n^3 \end{pmatrix}, \ldots, A^m = \begin{pmatrix} a_1^m & & \\ & \ddots & \\ & & a_n^m \end{pmatrix}$$

Proposition

Pour

$$A = \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix} \in D_n(\mathbb{K})$$

on a équivalence entre :
(i) A est inversible ;
(ii) $\forall 1 \leqslant i \leqslant n, a_i \neq 0$.
De plus si tel est le cas

$$A^{-1} = \begin{pmatrix} 1/a_1 & & \\ & \ddots & \\ & & 1/a_n \end{pmatrix}$$

dém. :
(i) $\Rightarrow$ (ii) Par contraposée :
Supposons qu'il existe $i \in \{1, \ldots, n\}$ tel que $a_i = 0$.
Pour toute matrice $B \in \mathcal{M}_n(\mathbb{K})$, la i ème ligne du produit AB est nulle car la $i$ème ligne de la matrice A est nulle. Ainsi, il n'existe pas de matrice $B \in \mathcal{M}_n(\mathbb{K})$ vérifiant $AB = I_n$.
(ii) $\Rightarrow$ (i) Supposons que pour tout $i \in \{1, \ldots, n\}$, $a_i \neq 0$. Posons

$$B = \begin{pmatrix} 1/a_1 & & \\ & \ddots & \\ & & 1/a_n \end{pmatrix}$$

On a $AB = I_n$ donc A est inversible et B est son inverse.
□

10.1.7.6 Anneau des matrices triangulaires

Proposition

$T_n^+(\mathbb{K})$ est un sous-anneau de $\mathcal{M}_n(\mathbb{K})$.

dém. :
$T_n^+(\mathbb{K}) \subset \mathcal{M}_n(\mathbb{K})$ et $I_n \in T_n^+(\mathbb{K})$.
Soient $A = (a_{i,j})$ et $B = (b_{i,j}) \in T_n^+(\mathbb{K})$.
Pour tout $i > j$, $a_{i,j} = b_{i,j} = 0$ car A et B sont triangulaires supérieures.
On a $A - B = (a_{i,j} - b_{i,j})$ avec $a_{i,j} - b_{i,j} = 0$ pour tout $i > j$ donc $A - B \in T_n^+(\mathbb{K})$.
Etudions maintenant $C = AB = (c_{i,j})$. On a

$$c_{i,j} = \sum_{k=1}^{n} a_{i,k} b_{k,j}$$

Pour $i > j$, on peut écrire

$$c_{i,j} = \sum_{k=1}^{j} a_{i,k}b_{k,j} + \sum_{k=j+1}^{n} a_{i,k}b_{k,j}$$

avec pour tout $k \in \{1, \dots, j\}$, $k \leqslant j < i$ et donc $a_{i,k} = 0$
et pour tout $k \in \{j+1, \dots, n\}$, $k > j$ et donc $b_{k,j} = 0$.
Ainsi $c_{i,j} = 0$ pour tout $i > j$ et donc $AB \in T_n^+(\mathbb{K})$.
□

Remarque Les coefficients diagonaux du produit de matrices triangulaires supérieures sont remarquables

$$\begin{pmatrix} a_{11} & & \star \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix} \begin{pmatrix} b_{11} & & \star \\ & \ddots & \\ 0 & & b_{nn} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} & & \star \\ & \ddots & \\ 0 & & a_{nn}b_{nn} \end{pmatrix}$$

Proposition

Soit $A \in T_n^+(\mathbb{K})$. On a équivalence entre :
(i) A est inversible ;
(ii) les coefficients diagonaux de A sont tous non nuls.
De plus, si tel est le cas, $A^{-1} \in T_n^+(\mathbb{K})$.

dém. :
(ii) $\Rightarrow$ (i) Supposons les coefficients diagonaux de A non nuls.
Soient $X \in \mathcal{M}_{n,1}(\mathbb{K})$ et $Y = AX \in \mathcal{M}_{n,1}(\mathbb{K})$.
On peut facilement inverser le système correspondant à l'équation $Y = AX$ car celui-ci est triangulaire et car les coefficients diagonaux de A sont non nuls. On en déduit que A est inversible.
(i) $\Rightarrow$ (ii) Supposons A inversible. Commençons par montrer que $A^{-1} \in T_n^+(\mathbb{K})$ en considérant l'application $\varphi : T_n^+(\mathbb{K}) \to T_n^+(\mathbb{K})$ définie par $\varphi(M) = AM$. On vérifie aisément que φ est un endomorphisme de $T_n^+(\mathbb{K})$. Pour $M \in \ker\varphi$, on a $M = A^{-1}AM = A^{-1}\varphi(M) = O_n$ donc $\ker\varphi = \{O_n\}$. L'endomorphisme φ est injectif, or $\dim T_n^+(\mathbb{K}) < +\infty$ donc φ est un automorphisme. Par surjectivité, il existe $B \in T_n^+(\mathbb{K})$ vérifiant $\varphi(B) = I_n$ et alors $B = (A^{-1}A)B = A^{-1}(AB) = A^{-1}$. Ainsi l'inverse de A est une matrice triangulaire supérieure.
Notons $b_{i,i}$ les coefficients diagonaux de A^{-1}.
Puisque $AA^{-1} = I_n$, on a, par étude de la diagonale du produit de matrices triangulaires $a_{i,i}b_{i,i} = 1$ pour tout $1 \leqslant i \leqslant n$. On en déduit $a_{i,i} \neq 0$ pour tout $1 \leqslant i \leqslant n$.
□

Remarque Les résultats qui précèdent se transposent aux matrices triangulaires inférieures.

10.1.8 Transposition

10.1.8.1 Définition

Définition

On appelle matrice transposée de $A = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{K})$ la matrice ${}^tA = (a'_{j,i}) \in \mathcal{M}_{p,n}(\mathbb{K})$ définie par

$$\forall 1 \leqslant i \leqslant n, \forall 1 \leqslant j \leqslant p, a'_{j,i} = a_{i,j}$$

Ainsi le coefficient d'indice (j,i) de ${}^tA$ est égal au coefficient d'indice (i,j) de A.

Remarque La transposée d'une matrice de type (n, p) et une matrice de type (p, n).

Remarque Matriciellement, pour

$$A = \begin{pmatrix} a_{11} & \dots & a_{1p} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{np} \end{pmatrix}$$

on a

$$^tA = \begin{pmatrix} a_{11} & \dots & a_{n1} \\ \vdots & & \vdots \\ a_{1p} & \dots & a_{np} \end{pmatrix}$$

Remarque Les colonnes et lignes de ^tA correspondent respectivement aux lignes et colonnes de A.

Exemple Pour

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$$

on a

$$^tA = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix}$$

Exemple Pour

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

on a

$$^tX = \begin{pmatrix} x_1 & \dots & x_n \end{pmatrix}$$

Remarque Pour $A = (a_{i,j})_{i,j} \in \mathcal{M}_{n,p}(\mathbb{K})$ on a $^tA = (a_{i,j})_{j,i} \in \mathcal{M}_{p,n}(\mathbb{K})$.
La transposition correspond à l'inversion du rôle joué par les indices de lignes et de colonnes.

Proposition

$\forall A \in \mathcal{M}_{n,p}(\mathbb{K}), {}^t({}^tA) = A$.

dém. :
Soit $A = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{K})$.
$^tA = (a'_{j,i}) \in \mathcal{M}_{n,p}(\mathbb{K})$ avec $a'_{j,i} = a_{i,j}$ et $^t({}^tA) = (a''_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{K})$ avec $a''_{i,j} = a'_{j,i}$.
Puisque $a''_{i,j} = a'_{j,i} = a_{i,j}$ pour tous indices i et j, on obtient $^t({}^tA) = A$.
□

Proposition

$\forall A, B \in \mathcal{M}_{n,p}(\mathbb{K}), \forall \lambda, \mu \in \mathbb{K}, {}^t(\lambda A + \mu B) = \lambda^t A + \mu^t B$.

dém. :
Soient $\lambda, \mu \in \mathbb{K}$, $A = (a_{i,j}), B = (b_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{K})$ et $C = \lambda A + \mu B = (c_{i,j})$
${}^tA = (a'_{j,i}) \in \mathcal{M}_{n,p}(\mathbb{K})$ avec $a'_{j,i} = a_{i,j}$, ${}^tB = (b'_{j,i}) \in \mathcal{M}_{n,p}(\mathbb{K})$ avec $b'_{j,i} = b_{i,j}$ et ${}^tC = (c'_{j,i}) \in \mathcal{M}_{n,p}(\mathbb{K})$ avec $c'_{j,i} = c_{i,j}$.
Puisque $c'_{j,i} = c_{i,j} = \lambda a_{i,j} + \mu b_{i,j} = \lambda a'_{j,i} + \mu b'_{j,i}$ pour tous indices i et j, on a ${}^tC = \lambda^t A + \mu^t B$.
□

Remarque L'application $T : \mathcal{M}_{n,p}(\mathbb{K}) \to \mathcal{M}_{p,n}(\mathbb{K})$ définie par $T(M) = {}^tM$ est un isomorphisme de $\mathbb{K}$-espace vectoriel.

Proposition

$\forall A \in \mathcal{M}_{n,p}(\mathbb{K}), \forall B \in \mathcal{M}_{p,q}(\mathbb{K}), {}^t(AB) = {}^tB^tA$.
On dit que la transposition est inversive pour le produit matriciel.

dém. :
Soient $A = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B = (b_{j,k}) \in \mathcal{M}_{p,q}(\mathbb{K})$
On obverse que les produits matriciels proposés dans la relation ${}^t(AB) = {}^tB^tA$ sont possibles.
Posons $C = AB = (c_{i,k}) \in \mathcal{M}_{n,q}(\mathbb{K})$ avec $c_{i,k} = \sum_{k=1}^{n} a_{i,j} b_{j,k}$.
Posons encore ${}^tA = (a'_{j,i}) \in \mathcal{M}_{p,n}(\mathbb{K})$ avec $a'_{j,i} = a_{i,j}$, ${}^tB = (b'_{k,j}) \in \mathcal{M}_{q,p}(\mathbb{K})$ avec $b'_{k,j} = b_{j,k}$, ${}^tC = (c'_{k,i}) \in \mathcal{M}_{q,n}(\mathbb{K})$ avec $c'_{k,i} = c_{i,k}$.
Considérons $D = {}^tB^tA = (d_{k,i}) \in \mathcal{M}_{q,n}(\mathbb{K})$.
On a $d_{k,i} = \sum_{j=1}^{p} b'_{k,j} a'_{j,i} = \sum_{j=1}^{p} a_{i,j} b_{j,k} = c_{i,k} = c'_{k,i}$ pour tous indices i et k donc ${}^t(AB) = {}^tB^tA$
□

Proposition

Si $A \in \mathcal{M}_n(\mathbb{K})$ est inversible alors ${}^tA$ l'est aussi et $({}^tA)^{-1} = {}^t(A^{-1})$.

dém. :
Si A est inversible alors on peut introduire son inverse A^{-1}.
Puisque $AA^{-1} = I_n$ on a ${}^t(AA^{-1}) = {}^tI_n$ c'est-à-dire ${}^t(A^{-1})^tA = I_n$.
Par le théorème d'inversibilité, on peut affirmer que ${}^tA$ est inversible et ${}^t(A^{-1})$ est son inverse.
□

10.1.8.2 Matrices symétriques et antisymétriques

Définition

On dit qu'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est symétrique si ${}^tA = A$.
On dit que la matrice A est antisymétrique si ${}^tA = -A$.
On note $\mathcal{S}_n(\mathbb{K})$ et $\mathcal{A}_n(\mathbb{K})$ les ensembles formés des matrices symétriques et antisymétrique de $\mathcal{M}_n(\mathbb{K})$.

Exemple

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{pmatrix} \in \mathcal{S}_3(\mathbb{R}) \text{ et } \begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & 3 \\ -2 & -3 & 0 \end{pmatrix} \in \mathcal{A}_3(\mathbb{R})$$

Proposition

Les matrices symétriques de $\mathcal{M}_n(\mathbb{K})$ sont celles de la forme

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{12} & a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1,n} \\ a_{1n} & \dots & a_{n-1,n} & a_{n,n} \end{pmatrix}$$

Par suite $\mathcal{S}_n(\mathbb{K})$ est un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{K})$ de dimension $\dfrac{n(n+1)}{2}$.

dém. :
Soit $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$.
La matrice A est symétrique si, et seulement si, $\forall 1 \leqslant i, j \leqslant n, a_{i,j} = a_{j,i}$.
Les matrices symétriques sont donc celles de la forme proposées.
En introduisant les matrices élémentaires de $\mathcal{M}_n(\mathbb{K})$, on peut écrire

$$\mathcal{S}_n(\mathbb{K}) = \text{Vect}\,\{E_{i,i}/1 \leqslant i \leqslant n\} \cup \{E_{i,j} + E_{j,i}/1 \leqslant i < j \leqslant n\}$$

On en déduit que $\mathcal{S}_n(\mathbb{K})$ est un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{K})$ et puisqu'il est facile d'établir que la famille formée des $E_{i,i}$ pour $1 \leqslant i \leqslant n$ et des $E_{i,j} + E_{j,i}$ pour $1 \leqslant i < j \leqslant n$ est libre, c'est une base de $\mathcal{S}_n(\mathbb{K})$ ce qui permet d'en calculer la dimension.
□

Proposition

Les matrices antisymétriques de $\mathcal{M}_n(\mathbb{K})$ sont celles de la forme

$$A = \begin{pmatrix} 0 & a_{12} & \dots & a_{1n} \\ -a_{12} & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1,n} \\ -a_{1n} & \dots & -a_{n-1,n} & 0 \end{pmatrix}$$

Par suite $\mathcal{A}_n(\mathbb{K})$ est un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{K})$ de dimension $\dfrac{n(n-1)}{2}$.

dém. :
On adapte la preuve précédente et on obtient

$$\mathcal{A}_n(\mathbb{K}) = \text{Vect}\,\{E_{i,j} - E_{j,i}/1 \leqslant i < j \leqslant n\}$$

□

Théorème

Les espaces $\mathcal{S}_n(\mathbb{K})$ et $\mathcal{A}_n(\mathbb{K})$ sont supplémentaires dans $\mathcal{M}_n(\mathbb{K})$.

dém. :
Soit $A \in \mathcal{S}_n(\mathbb{K}) \cap \mathcal{A}_n(\mathbb{K})$.
On a ${}^tA = A$ et ${}^tA = -A$ donc $A = O_n$. Ainsi $\mathcal{S}_n(\mathbb{K}) \cap \mathcal{A}_n(\mathbb{K}) = \{O_n\}$.
Soit $M \in \mathcal{M}_n(\mathbb{K})$.
On peut écrire

$$M = \frac{1}{2}\left(M + {}^tM\right) + \frac{1}{2}\left(M - {}^tM\right)$$

avec
$$\frac{1}{2}\left(M+{}^tM\right)\in\mathcal{S}_n(\mathbb{K}) \text{ et } \frac{1}{2}\left(M-{}^tM\right)\in\mathcal{A}_n(\mathbb{K})$$
donc $\mathcal{S}_n(\mathbb{K})+\mathcal{A}_n(\mathbb{K})=\mathcal{M}_n(\mathbb{K})$.
□

10.2 Représentations matricielles

10.2.1 Matrice colonne des composantes d'un vecteur

Soit E un $\mathbb{K}$-espace vectoriel muni d'une base $\mathcal{B}=(e_1,\dots,e_n)$.
Pour tout $x\in E$, on peut écrire de façon unique $x=\alpha_1e_1+\dots+\alpha_ne_n$ avec $(\alpha_1,\dots,\alpha_n)\in\mathbb{K}^n$.

Définition

On appelle matrice des composantes dans $\mathcal{B}$ du vecteurs x la matrice colonne dont les coefficients sont les composantes $\alpha_1,\dots,\alpha_n$ du vecteur x dans la base $\mathcal{B}$. On note
$$\text{Mat}_{\mathcal{B}}(x)=\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{pmatrix}\in\mathcal{M}_{n,1}(\mathbb{K})$$

Remarque Puisque les composantes d'un vecteur dépendent de la base choisie, il est nécessaire de préciser celle-ci.

Exemple Puisque les composantes du vecteur e_i dans la base $\mathcal{B}=(e_1,\dots,e_n)$ sont $0,\dots,0,1,0,\dots,0$ on a
$$\text{Mat}_{\mathcal{B}}(e_i)=\begin{pmatrix}(0)\\ 1\\ (0)\end{pmatrix}=E_i$$

Théorème

L'application $M_{\mathcal{B}}: x\mapsto \text{Mat}_{\mathcal{B}}x$ est un isomorphisme du $\mathbb{K}$-espace vectoriel E vers $\mathcal{M}_{n,1}(\mathbb{K})$.

dém. :
L'application $M_{\mathcal{B}}$ est linéaire car l'application qui à un vecteur associe l'une des ses composantes dans $\mathcal{B}$ est linéaire.
L'application $M_{\mathcal{B}}$ est bijective car pour toute colonne X de coefficients $\alpha_1,\dots,\alpha_n$, il existe un unique vecteur dans les composantes dans $\mathcal{B}$ sont $\alpha_1,\dots,\alpha_n$ à savoir le vecteur $x=\alpha_1e_1+\dots+\alpha_ne_n$.
□

10.2.2 Matrice des composantes d'une famille de vecteurs

Soit $\mathcal{F}=(x_1,x_2,\dots,x_p)$ une famille de vecteurs d'un $\mathbb{K}$-espace vectoriel E muni d'une base $\mathcal{B}=(e_1,\dots,e_n)$.
Pour tout $1\leqslant j\leqslant p$, notons C_j la colonne des composantes dans $\mathcal{B}$ du vecteur x_j.

Définition

On appelle matrice des composantes dans la base $\mathcal{B}$ de la famille $\mathcal{F}$, la matrice de $\mathcal{M}_{n,p}(\mathbb{K})$ dont les colonnes sont les colonnes $C_1, C_2, \ldots, C_p$ des composantes dans $\mathcal{B}$ des vecteurs $x_1, \ldots, x_p$. On note

$$A = \mathrm{Mat}_{\mathcal{B}}\mathcal{F} = \mathrm{Mat}_{\mathcal{B}}(x_1, \ldots, x_p) \in \mathcal{M}_{n,p}(\mathbb{K})$$

Remarque Dans le cas $p = 1$, on retrouve la notion de matrice des composantes d'un vecteur.

Exemple Connaissant les composantes des vecteurs $e_1, \ldots, e_n$ dans la base $\mathcal{B} = (e_1, \ldots, e_n)$, on observe $\mathrm{Mat}_{\mathcal{B}}\mathcal{B} = \mathrm{Mat}_{\mathcal{B}}(e_1, \ldots, e_n) = I_n$

Exemple Considérons $E = \mathbb{K}^3$ muni de sa base canonique $\mathcal{B} = (e_1, e_2, e_3)$.
Soient $x_1 = (1, 2, 3)$, $x_2 = (2, 0, 1)$, $x_3 = (1, 0, -1)$ et $x_4 = (-1, 1, 1)$ vecteurs de E.
Les composantes des vecteurs x_i étant faciles à déterminer dans la base canonique $\mathcal{B}$, on obtient

$$\mathrm{Mat}_{\mathcal{B}}(x_1, x_2, x_3, x_4) = \begin{pmatrix} 1 & 2 & 1 & -1 \\ 2 & 0 & 0 & 1 \\ 3 & 1 & -1 & 1 \end{pmatrix} \in \mathcal{M}_{3,4}(\mathbb{R})$$

Exemple Considérons $E = \mathbb{R}_3[X]$ muni de sa base canonique $\mathcal{B} = (1, X, X^2, X^3)$.
Formons la matrice des composantes dans $\mathcal{B}$ de la famille $(P_k)_{0 \leqslant k \leqslant 3}$ avec $P_k = (X+1)^k$.
Pour cela, on réécrit les polynômes de sorte d'en déterminer les composantes dans $\mathcal{B}$.
$P_0 = 1$, $P_1 = 1 + X$, $P_2 = 1 + 2X + X^2$ et $P_3 = 1 + 3X + 3X^2 + X^3$
On en déduit

$$\mathrm{Mat}_{\mathcal{B}}(P_0, P_1, P_2, P_3) = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \in \mathcal{M}_4(\mathbb{R})$$

10.2.3 Matrice d'une application linéaire

Soient E et F des $\mathbb{K}$-espaces vectoriels de dimensions p et n et munis de bases $\mathcal{B} = (e_1, \ldots, e_p)$ et $\mathcal{C} = (f_1, \ldots, f_n)$.

Définition

On appelle matrice représentative dans les bases $\mathcal{B}$ et $\mathcal{C}$ d'une application linéaire $u \in \mathcal{L}(E, F)$ la matrice des composantes dans $\mathcal{C}$ de la famille image $u(\mathcal{B}) = (u(e_1), \ldots, u(e_p))$. On note

$$\mathrm{Mat}_{\mathcal{B},\mathcal{C}}u = \mathrm{Mat}_{\mathcal{C}}u(\mathcal{B}) = \mathrm{Mat}_{\mathcal{C}}\left(u(e_1), \ldots, u(e_p)\right) \in \mathcal{M}_{n,p}(\mathbb{K})$$

Remarque La matrice représentative de u dépend du choix des bases $\mathcal{B}$ et $\mathcal{C}$, il est donc nécessaire de préciser celles-ci.

Exemple Soit $u : \mathbb{K}^3 \to \mathbb{K}^2$ l'application linéaires définie par $u(x, y, z) = (x + 2y - z, x - y)$.
Formons la matrice de u relative aux bases canoniques $\mathcal{B} = (e_1, e_2, e_3)$ et $\mathcal{C} = (f_1, f_2)$ des espaces $\mathbb{K}^3$ et $\mathbb{K}^2$.
Pour cela on calcule les images des vecteurs de la base de départ
$u(e_1) = u(1, 0, 0) = (1, 1)$,
$u(e_2) = u(0, 1, 0) = (2, -1)$,
$u(e_3) = u(0, 0, 1) = (-1, 0)$.
On détermine ensuite les composantes de ces images dans la base d'arrivée ce qui est ici immédiat car la base d'arrivée est la base canonique de $\mathbb{K}^2$.
On en déduit

$$\mathrm{Mat}_{\mathcal{B},\mathcal{C}} u = \begin{pmatrix} 1 & 2 & -1 \\ 1 & -1 & 0 \end{pmatrix}$$

Exemple Soient $a, b, c \in \mathbb{K}$ fixés et $u : \mathbb{K}_3[X] \to \mathbb{K}^3$ l'application linéaire définie par $u(P) = (P(a), P(b), P(c))$.
Formons la matrice de u relative aux bases canoniques $\mathcal{B} = (1, X, X^2, X^3)$ et $\mathcal{C} = (f_1, f_2, f_3)$ des espaces $\mathbb{K}_3[X]$ et $\mathbb{K}^3$.
On commence par calculer les images des vecteurs de la base de départ
$u(1) = (1, 1, 1)$, $u(X) = (a, b, c)$, $u(X^2) = (a^2, b^2, c^2)$ et $u(X^3) = (a^3, b^3, c^3)$
puis on exprime les composantes de ces vecteurs dans la base d'arrivée ce qui est ici immédiat.
On en déduit

$$\mathrm{Mat}_{\mathcal{B},\mathcal{C}} = \begin{pmatrix} 1 & a & a^2 & a^3 \\ 1 & b & b^2 & b^3 \\ 1 & c & c^2 & c^3 \end{pmatrix}$$

10.2.4 Matrice d'un endomorphisme

Soit E un $\mathbb{K}$-espace vectoriel de dimension n muni d'une base $\mathcal{B} = (e_1, \ldots, e_n)$.

Définition

On appelle matrice représentative dans la base $\mathcal{B}$ d'un endomorphisme $u \in \mathcal{L}(E)$ la matrice représentative dans la base $\mathcal{B}$ au départ et à l'arrivée de l'application linéaire u. Celle matrice est notée $\mathrm{Mat}_{\mathcal{B}} f$ de préférence à $\mathrm{Mat}_{\mathcal{B},\mathcal{B}}(f)$. C'est une matrice carrée d'ordre n.

Exemple La matrice de l'endomorphisme Id_E dans la base $\mathcal{B}$ est

$$\mathrm{Mat}_{\mathcal{B}} \mathrm{Id}_E = \mathrm{Mat}_{\mathcal{B}} \left(\mathrm{Id}_E(e_1), \ldots, \mathrm{Id}_E(e_n)\right) = \mathrm{Mat}_{\mathcal{B}}(e_1, \ldots, e_n) = I_n$$

Exemple Soit $u : \mathbb{K}^3 \to \mathbb{K}^3$ l'endomorphisme défini par $u(x, y, z) = (y + z, z + x, x + y)$.
Formons la matrice de u relative à la base canonique $\mathcal{B} = (e_1, e_2, e_3)$ de $\mathbb{K}^3$.
$u(e_1) = (0, 1, 1)$, $u(e_2) = (1, 0, 1)$ et $u(e_3) = (1, 1, 0)$.
La matrice de u dans $\mathcal{B}$ est

$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

Considérons maintenant la famille de vecteurs $\mathcal{C} = (\varepsilon_1, \varepsilon_2, \varepsilon_3)$ avec $\varepsilon_1 = (1, 1, 1)$, $\varepsilon_2 = (1, 1, 0)$ et $\varepsilon_3 = (1, 0, 0)$. On vérifie aisément que la famille $\mathcal{C}$ est une base de $\mathbb{K}^3$.

Formons la matrice de u relative à la base $\mathcal{C}$.

$$u(\varepsilon_1) = (2,2,2),\ u(\varepsilon_2) = (1,1,2) \text{ et } u(\varepsilon_3) = (0,1,1)$$

Pour former la matrice voulue, on détermine les composantes des précédents vecteurs dans la base $\mathcal{C}$ (et non celles dans la base canonique).

$$u(\varepsilon_1) = 2\varepsilon_1 + 0.\varepsilon_2 + 0.\varepsilon_3,\ u(\varepsilon_2) = 2\varepsilon_1 - \varepsilon_2 \text{ et } u(\varepsilon_3) = \varepsilon_1 - \varepsilon_3$$

La matrice de u dans $\mathcal{C}$ est

$$\begin{pmatrix} 2 & 2 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

10.3 Application du calcul matriciel aux applications linéaires

10.3.1 Image d'un vecteur

Soient E et F deux $\mathbb{K}$-espaces vectoriels munis de bases $\mathcal{B} = (e_1, \ldots, e_p)$ et $\mathcal{C} = (f_1, \ldots, f_n)$.
Pour $x \in E$ et $y \in F$, on convient de noter X et Y les colonnes des composantes de x et y dans les bases $\mathcal{B}$ et $\mathcal{C}$.

Théorème

Pour $u \in \mathcal{L}(E,F)$, la matrice de u dans les bases $\mathcal{B}$ et $\mathcal{C}$ est l'unique matrice $A \in \mathcal{M}_{n,p}(\mathbb{K})$ vérifiant

$$\forall x \in E, \forall y \in F, y = u(x) \Leftrightarrow Y = AX$$

dém. :
Unicité : Soient A et A' deux matrices solutions.
Pour tout vecteur x de E, on a $Y = AX = AX'$.
On en déduit que pour toute colonne X, $AX = A'X$.
On en déduit $A = A'$ en vertu du lemme d'identification matricielle.
Convenance :
Posons $A = (a_{i,j}) = \text{Mat}_{\mathcal{B},\mathcal{C}} u$.
Notons $x_1, \ldots, x_p$ les composantes de x dans $\mathcal{B}$.
$x = \sum_{j=1}^{p} x_j e_j$ donc $u(x) = \sum_{j=1}^{p} x_j u(e_j)$
Puisque les colonnes de A sont formées des composantes des vecteurs $u(e_1), \ldots, u(e_p)$ on a pour tout $1 \leqslant j \leqslant p$, $u(e_j) = \sum_{i=1}^{n} a_{i,j} f_i$.
On en déduit

$$u(x) = \sum_{j=1}^{p} \sum_{i=1}^{n} a_{i,j} x_j f_i$$

En permutant les deux sommes

$$u(x) = \sum_{i=1}^{n} \left(\sum_{j=1}^{p} a_{i,j} x_j \right) f_i$$

Notons $y_1, \dots, y_n$ les composantes de y dans $\mathcal{C}$, $y = \sum_{i=1}^{n} y_i f_i$
Par identification des composantes, on a

$$y = u(x) \Leftrightarrow \forall i \in \{1, \dots, n\}, y_i = \sum_{j=1}^{p} a_{i,j} x_j$$

Parallèlement

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \text{ et } AX = \begin{pmatrix} a_{1,1}x_1 + \cdots + a_{1,p}x_p \\ \vdots \\ a_{n,1}x_1 + \cdots + a_{n,p}x_p \end{pmatrix}$$

donc

$$Y = AX \Leftrightarrow \forall i \in \{1, \dots, n\}, y_i = \sum_{j=1}^{p} a_{i,j} x_j$$

Ainsi

$$y = u(x) \Leftrightarrow Y = AX$$

□

Exemple Soit E un $\mathbb{R}$-espace vectoriel muni d'une base $\mathcal{B} = (e_1, e_2, e_3)$.
Soit u un endomorphisme de E dont la matrice dans $\mathcal{B}$ est

$$\text{Mat}_{\mathcal{B}} u = \begin{pmatrix} 1 & 1 & 2 \\ 1 & -1 & 0 \\ 1 & 0 & 1 \end{pmatrix} = A$$

Soit $x = x_1 e_1 + x_2 e_2 + x_3 e_3 \in E$. On peut calculer le vecteur $u(x)$ par produit matriciel.

$$\text{Mat}_{\mathcal{B}} u(x) = AX = \begin{pmatrix} x_1 + x_2 + 2x_3 \\ x_1 - x_2 \\ x_1 + x_3 \end{pmatrix}$$

On peut alors étudier le noyau de u en résolvant l'équation matricielle $AX = O_{3,1}$.

$$AX = O_{3,1} \Leftrightarrow \begin{cases} x_1 + x_2 + 2x_3 = 0 \\ x_1 - x_2 = 0 \\ x_1 + x_3 = 0 \end{cases} \Leftrightarrow \begin{cases} x_2 = x_1 \\ x_3 = -x_1 \end{cases}$$

Ainsi

$$\ker u = \{x_1(e_1 + e_2 - e_3) / x_1 \in \mathbb{K}\} = \text{Vect}(e_1 + e_2 - e_3)$$

On peut aussi facilement déterminer l'image de u.
En effet par le théorème du rang, on a déjà $\text{rg}u = \dim E - \dim \ker u = 2$. On peut donc déterminer une base de l'image de u en considérant deux vecteurs non colinéaires de l'image de u. Or les colonnes de A sont formées des composantes de vecteurs images par u et déterminent donc des éléments de l'image de u. Il est donc facile de déterminer deux vecteurs non colinéaires dans $\text{Im}u$. Par exemple avec la première et la deuxième colonne de A, les vecteurs $u(e_1) = e_1 + e_2 + e_3$ et $u(e_2) = e_1 - e_2$ forment une base de $\text{Im}u$.

Exemple Soit E un $\mathbb{K}$-espace vectoriel muni d'une base $\mathcal{B} = (e_1, \dots, e_n)$.
Les formes linéaires sur E sont les applications linéaires à valeurs dans $F = \mathbb{K}$. Il est usuel de munir $\mathbb{K}$ de sa base canonique (1) ce qui permet d'identifier un scalaire et sa composante dans cette base.
Pour φ une forme linéaire sur E, la matrice de φ dans $\mathcal{B}$ est une matrice ligne

$$L = \begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix} \in \mathcal{M}_{1,n}(\mathbb{K})$$

Pour tout vecteur $x \in E$ de composantes $x_1, \dots, x_n$ dans $\mathcal{B}$, on a

$$\varphi(x) = LX = a_1 x_1 + \cdots + a_n x_n$$

10.3.2 L'isomorphisme de représentation matricielle

Soient E et F deux $\mathbb{K}$-espaces vectoriels munis de bases $\mathcal{B} = (e_1, \dots, e_p)$ et $\mathcal{C} = (f_1, \dots, f_n)$.

Théorème

L'application $M_{\mathcal{B},\mathcal{C}} : \mathcal{L}(E, F) \to \mathcal{M}_{n,p}(\mathbb{K})$ définie par $M_{\mathcal{B},\mathcal{C}}(u) = \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(u)$ est un isomorphisme de $\mathbb{K}$-espaces vectoriels.

dém. :
Soient $\lambda, \mu \in \mathbb{K}$ et $u, v \in \mathcal{L}(E, F)$.
Posons $A = \mathrm{Mat}_{\mathcal{B},\mathcal{C}} u$ et $B = \mathrm{Mat}_{\mathcal{B},\mathcal{C}} v$.
Pour $x \in E$, on a $(\lambda u + \mu v)(x) = \lambda u(x) + \mu v(x)$.
Or les vecteurs $u(x)$ et $v(x)$ ont pour colonnes composantes AX et BX dans $\mathcal{C}$.
On en déduit que le vecteur $(\lambda u + \mu v)(x)$ a pour colonne composante $\lambda AX + \mu BX = (\lambda A + \mu B)X$.
Or il y a unicité de la matrice permettant de calculer les composantes d'un vecteur image par $\lambda u + \mu v$ et celle-ci est $\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(\lambda u + \mu v)$.
On en déduit

$$\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(\lambda u + \mu v) = \lambda A + \mu B = \lambda \mathrm{Mat}_{\mathcal{B},\mathcal{C}} u + \mu \mathrm{Mat}_{\mathcal{B},\mathcal{C}} v$$

i.e.

$$M_{\mathcal{B},\mathcal{C}}(\lambda u + \mu v) = \lambda M_{\mathcal{B},\mathcal{C}}(u) + \mu M_{\mathcal{B},\mathcal{C}}(v)$$

Ainsi l'application $M_{\mathcal{B},\mathcal{C}}$ est linéaire.
Soit $u \in \ker M_{\mathcal{B},\mathcal{C}}$. On a $\mathrm{Mat}_{\mathcal{B},\mathcal{C}} u = O_n$.
Pour tout $x \in E$ de composantes X dans $\mathcal{B}$, le vecteur $u(x)$ a pour composantes $O_n X = O_{n,1}$ et donc c'est le vecteur nul. On en déduit $u = \tilde{0}$ puis $\ker M_{\mathcal{B},\mathcal{C}} = \{\tilde{0}\}$. Ainsi l'application linéaire $M_{\mathcal{B},\mathcal{C}}$ est injective.
Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$. Considérons ensuite l'application $u : E \to F$ qui envoie un vecteur $x \in E$ de composantes X dans $\mathcal{B}$ sur le vecteur y de composantes AX dans $\mathcal{C}$. Vérifions que l'application u est linéaire de matrice A relatives aux bases $\mathcal{B}$ et $\mathcal{C}$.
Soient $\lambda_1, \lambda_2 \in \mathbb{K}$ et $x_1, x_2 \in E$. Notons X_1 et X_2 les colonnes des composantes des vecteurs x_1 et x_2 dans $\mathcal{B}$. Les vecteurs $u(x_1)$ et $u(x_2)$ ont pour colonnes composantes AX_1 et AX_2 dans $\mathcal{B}$ et puisque le vecteur $\lambda_1 x_1 + \lambda_2 x_2$ a pour colonne composante $\lambda_1 X_1 + \lambda_2 X_2$ dans $\mathcal{B}$, le vecteur $u(\lambda_1 x_1 + \lambda_2 x_2)$ a pour colonne composante $A(\lambda_1 X_1 + \lambda_2 X_2)$ dans $\mathcal{C}$.
Sachant $A(\lambda_1 X_1 + \lambda_2 X_2) = \lambda_1 A X_1 + \lambda_2 A X_2$, on a $u(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1 u(x_1) + \lambda_2 u(x_2)$. Ainsi l'application u est linéaire. De plus, par construction, la matrice de u relatives aux bases $\mathcal{B}$ et $\mathcal{C}$ est la matrice A.
Ainsi l'application $M_{\mathcal{B},\mathcal{C}}$ est surjective et finalement c'est un isomorphisme.
□

Corollaire

Si E et F sont deux $\mathbb{K}$-espaces vectoriels de dimensions finies alors l'espace $\mathcal{L}(E,F)$ est un $\mathbb{K}$-espace vectoriel de dimension finie et $\dim \mathcal{L}(E,F) = \dim E \times \dim F$.
En particulier, $\dim \mathcal{L}(E) = (\dim E)^2$ et $\dim E^\star = \dim E$.

dém. :
Par isomorphisme, $\dim \mathcal{L}(E,F) = \dim \mathcal{M}_{n,p}(\mathbb{K}) = np$ avec $p = \dim E$ et $n = \dim F$.
□

Remarque Par l'isomorphisme de représentation matricielle, introduire une application linéaire u de E vers F équivaut à introduire sa représentation matricielle relative à des bases données de E et F. C'est très souvent ainsi que sont introduit des applications linéaires en dimension finie.

10.3.3 Composition d'applications linéaires

Soient E, F et G trois $\mathbb{K}$-espaces vectoriels munis de bases $\mathcal{B} = (e_1, \ldots, e_p)$, $\mathcal{C} = (f_1, \ldots, f_n)$, $\mathcal{D} = (g_1, \ldots, g_m)$.

Théorème

Pour tout $u \in \mathcal{L}(E,F)$ et $v \in \mathcal{L}(F,G)$, on a

$$\mathrm{Mat}_{\mathcal{B},\mathcal{D}}(v \circ u) = \mathrm{Mat}_{\mathcal{C},\mathcal{D}} v \times \mathrm{Mat}_{\mathcal{B},\mathcal{C}} u$$

(bases organisées en « Chasles inversé »)

dém. :
Soient $u \in \mathcal{L}(E,F)$ et $v \in \mathcal{L}(F,G)$.
Posons $A = \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(u)$ et $B = \mathrm{Mat}_{\mathcal{C},\mathcal{D}}(v)$.
Soit $x \in E$ de colonne composante X dans la base $\mathcal{B}$.
La colonne des composantes de $y = u(x)$ dans $\mathcal{C}$ est $Y = AX$.
La colonne des composantes de $z = v(y)$ dans $\mathcal{D}$ est $Z = BY$.
Ainsi la colonne des composantes de $z = (v \circ u)(x)$ dans $\mathcal{D}$ est $Z = (BA)X$.
Par unicité de la matrice permettant de calculer les composantes d'un vecteur image, on a $\mathrm{Mat}_{\mathcal{B},\mathcal{D}}(v \circ u) = BA$.
□

Corollaire

Soit E un $\mathbb{K}$-espace vectoriel muni d'une base $\mathcal{B} = (e_1, \ldots, e_n)$
L'application $M_{\mathcal{B}} : \mathcal{L}(E) \to \mathcal{M}_n(\mathbb{K})$ défini par $M_{\mathcal{B}}(u) = \mathrm{Mat}_{\mathcal{B}} u$ est un isomorphisme d'anneaux.
En particulier, pour tout $u, v \in \mathcal{L}(E)$, on a

$$\mathrm{Mat}_{\mathcal{B}}(u \circ v) = \mathrm{Mat}_{\mathcal{B}} u \times \mathrm{Mat}_{\mathcal{B}} v$$

et

$$\mathrm{Mat}_{\mathcal{B}}(u^m) = [\mathrm{Mat}_{\mathcal{B}} u]^m$$

pour tout $m \in \mathbb{N}$.

dém. :
$M_{\mathcal{B}}(\mathrm{Id}_E) = I_n$, $M_{\mathcal{B}}(u - v) = M_{\mathcal{B}}(u) - M_{\mathcal{B}}(v)$ et par ce qui précède $M_{\mathcal{B}}(u \circ v) = M_{\mathcal{B}}(u) M_{\mathcal{B}}(v)$.
□

Exemple Si A est la matrice de $u \in \mathcal{L}(E)$ dans une certaine base de E alors

u est nilpotent $\Leftrightarrow A$ est nilpotente.
u est un projecteur $\Leftrightarrow A^2 = A$.
u est une symétrie $\Leftrightarrow A^2 = I_n$.

10.3.4 Isomorphisme et matrice inversible

Soient E et F deux $\mathbb{K}$-espaces vectoriels de même dimension n munis de bases $\mathcal{B} = (e_1, \dots, e_n)$ et $\mathcal{C} = (f_1, \dots, f_n)$.

Théorème

> Soient $u \in \mathcal{L}(E, F)$ et $A = \text{Mat}_{\mathcal{B},\mathcal{C}}(u)$. On a équivalence entre :
> (i) u est un isomorphisme ;
> (ii) A est inversible.
> De plus, si tel est le cas, $\text{Mat}_{\mathcal{C},\mathcal{B}}(u^{-1}) = A^{-1}$.

dém. :
(i) $\Rightarrow$ (ii) Supposons que u est un isomorphisme et introduisons son isomorphisme inverse u^{-1}. Puisque $u^{-1} \circ u = \text{Id}_E$, on a $\text{Mat}_{\mathcal{B}}(u^{-1} \circ u) = \text{Mat}_{\mathcal{B}}\text{Id}_E = I_n$. Or $\text{Mat}_{\mathcal{B}}(u^{-1} \circ u) = \text{Mat}_{\mathcal{C},\mathcal{B}}u^{-1} \times \text{Mat}_{\mathcal{B},\mathcal{C}}u$. Ainsi en posant $B = \text{Mat}_{\mathcal{C},\mathcal{B}}u^{-1}$, on a $BA = I_n$. Par le théorème d'inversibilité, on peut affirmer que A est inversible et que B est son inverse.
(ii) $\Rightarrow$ (i) Supposons A inversible et introduisons sa matrice inverse A^{-1}. On peut aussi introduire l'application linéaire $v \in \mathcal{L}(F, E)$ représentée par A^{-1} relativement aux bases $\mathcal{C}$ et $\mathcal{B}$. Puisque $A^{-1}A = I_n$, on a $\text{Mat}_{\mathcal{B}}(v \circ u) = \text{Mat}_{\mathcal{B}}(\text{Id}_E)$ et donc $v \circ u = \text{Id}_E$. Par le théorème d'isomorphisme, on peut affirmer que u est un isomorphisme et que v est son isomorphisme inverse.
□

Exemple Soit $u : \mathbb{R}_2[X] \to \mathbb{R}^3$ l'application linéaire définie par $u(P) = (P(0), P(1), P(2))$.
Introduisons $\mathcal{B} = (1, X, X^2)$ la base canonique de $\mathbb{R}_2[X]$ et $\mathcal{C} = (e_1, e_2, e_3)$ celle de $\mathbb{R}^3$.
La matrice de u relative aux bases $\mathcal{B}$ et $\mathcal{C}$ est

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{pmatrix}$$

On vérifie par le calcul que cette matrice est inversible et

$$A^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -3/2 & 2 & -1/2 \\ 1/2 & -1 & 1/2 \end{pmatrix}$$

On en déduit que l'application linéaire u est un isomorphisme et l'on connaît son isomorphisme inversible par le biais d'une représentation matricielle. Cet isomorphisme inverse résout un problème d'interpolation, à savoir déterminer un polynôme $P \in \mathbb{R}_2[X]$ prenant des valeurs donnée en $0, 1, 2$. Par exemple le polynôme $P \in \mathbb{R}_2[X]$ vérifiant $P(0) = 1$, $P(1) = 0$ et $P(2) = 2$ a pour composante dans $\mathcal{B}$

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

C'est le polynôme

$$P(X) = 1 - \frac{5}{2}X + \frac{3}{2}X^2$$

Corollaire

Soit E un $\mathbb{K}$-espace vectoriel muni d'une base $\mathcal{B} = (e_1, \ldots, e_n)$.
Pour $u \in \mathcal{L}(E)$ on a

$$u \in \mathrm{GL}(E) \Leftrightarrow \mathrm{Mat}_{\mathcal{B}} u \in \mathrm{GL}_n(\mathbb{K})$$

De plus, si tel est le cas :

$$\mathrm{Mat}_{\mathcal{B}}(u^{-1}) = [\mathrm{Mat}_{\mathcal{B}} u]^{-1}$$

10.3.5 Tableau des correspondances

Vecteur	Matrice colonne
$x \in E$	$X \in \mathcal{M}_{p,1}(\mathbb{K})$
0	$O_{p,1}$
$\lambda x + \mu y$	$\lambda X + \mu Y$
Application linéaire	Matrice rectangle
$u \in \mathcal{L}(E, F)$	$A \in \mathcal{M}_{n,p}(\mathbb{K})$
$\tilde{o}$	$O_{n,p}$
$y = u(x)$	$Y = AX$
$\lambda u + \mu v$	$\lambda A + \mu B$
$v \circ u$	BA
u isomorphisme, u^{-1}	A inversible, A^{-1}
Endomorphisme	Matrice carrée
$u \in \mathcal{L}(E)$	$A \in \mathcal{M}_n(\mathbb{K})$
Id_E	I_n
u^m	A^m
$u \in \mathrm{GL}(E)$, u^{-1}	$A \in \mathrm{GL}_n(\mathbb{K})$, A^{-1}
Formes linéaires	Matrice ligne
$\varphi \in E^\star$	$L \in \mathcal{M}_{1,n}(\mathbb{K})$
$\varphi(x)$	LX

10.3.6 Différentes représentation d'un endomorphisme

10.3.6.1 Matrice d'une homothétie vectorielle

Soit E un $\mathbb{K}$-espace vectoriel de dimension n.

Proposition

Dans toute base de E, la matrice de l'homothétie vectorielle de rapport $\lambda \in \mathbb{K}$ est λI_n.

dém. :
Notons h_λ l'homothétie vectorielle de rapport λ et considérons $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E.
Pour tout $1 \leqslant i \leqslant n$, $h(e_i) = \lambda e_i$ donc

$$\mathrm{Mat}_{\mathcal{B}} h_\lambda = \begin{pmatrix} \lambda & & (0) \\ & \ddots & \\ (0) & & \lambda \end{pmatrix}$$

□

Remarque On peut montrer que les homothéties vectorielles sont les endomorphismes ayant la même matrice dans toute base.

10.3.6.2 Matrice d'une projection, d'une symétrie

Soient F et G deux sous-espaces vectoriels supplémentaires de E de dimensions r et $n - r$.

Théorème

La matrice de la projection p sur F parallèlement à G dans une base $\mathcal{B}$ adaptée à la supplémentarité de F et G est

$$\mathrm{Mat}_{\mathcal{B}} p = \begin{pmatrix} I_r & 0 \\ 0 & 0_{n-r} \end{pmatrix}$$

dém. :
Notons $e_1, \ldots, e_n$ les vecteurs constituant la base $\mathcal{B}$.
Puisque la base $\mathcal{B}$ est adaptée à la supplémentarité de F et G, on a $e_1, \ldots, e_r \in F$ et $e_{r+1}, \ldots, e_n \in G$.
On en déduit

$$\forall i \in \{1, \ldots, r\}, p(e_i) = e_i$$

et

$$\forall i \in \{r+1, \ldots, n\}, p(e_i) = 0$$

d'où la représentation matricielle annoncée.
□

Théorème

La matrice de la symétrie s par rapport à F et parallèlement à G dans une base $\mathcal{B}$ adaptée à la supplémentarité de F et G est

$$\mathrm{Mat}_{\mathcal{B}} s = \begin{pmatrix} I_r & 0 \\ 0 & -I_{n-r} \end{pmatrix}$$

dém. :
En reprenant les notations de la preuve précédente, on a

$$\forall i \in \{1, \ldots, r\}, s(e_i) = e_i$$

et

$$\forall i \in \{r+1, \ldots, n\}, s(e_i) = -e_i$$

□

Remarque Ces représentations matricielles sont simples car relatives à une base adaptée au problème étudié.
Dans une base quelconque, la matrice d'une projection est plus compliquée…

Exemple Soit E un $\mathbb{R}$-espace vectoriel de dimension 3 muni d'une base $\mathcal{B} = (e_1, e_2, e_3)$.
Soient $P : x_1 + x_2 + x_3 = 0$ et $D = \text{Vect}u$ avec $u = e_1 + e_2 - e_3$.
Les espaces P et D sont supplémentaires dans E car P est un hyperplan de E et w est un vecteur ne lui appartenant pas.
Formons la matrice A de la projection p sur P parallèlement à D.
Soient $x = x_1e_1 + x_2e_2 + x_3e_3$ un vecteur de E et $y = p(x) = y_1e_1 + y_2e_2 + y_3e_3$ son image.
Exprimons y_1, y_2, y_3 en fonction de x_1, x_2, x_3.
Par définition de l'action de la projection p, on sait $p(x) \in P$ et $x - p(x) \in D$. Ces deux propriétés caractérisent le vecteur $p(x)$ en fonction du vecteur x et on vont donc permettre d'expliciter y_1, y_2, y_3 en fonction de x_1, x_2, x_3.
$p(x) \in P$ donne $y_1 + y_2 + y_3 = 0$.
$x - p(x) \in D$ assure l'existence d'un scalaire λ vérifiant $x - p(x) = \lambda u$ i.e.

$$\begin{cases} x_1 - y_1 = \lambda \\ x_2 - y_2 = \lambda \\ x_3 - y_3 = -\lambda \end{cases}$$

On est ainsi amené à résoudre le système

$$\begin{cases} y_1 + y_2 + y_3 = 0 \\ y_1 + \lambda = x_1 \\ y_2 + \lambda = x_2 \\ y_3 - \lambda = x_3 \end{cases}$$

en les inconnues y_1, y_2, y_3 et λ.
Au terme de cette résolution, on obtient les expressions de x', y', z' en fonction de x, y, z

$$\begin{cases} y_1 = -x_2 - x_3 \\ y_2 = -x_1 - x_3 \\ y_3 = x_1 + x_2 + 2x_3 \end{cases}$$

Il est alors facile d'exprimer la matrice A de p :
- soit en calculant les images des vecteurs de base e_1, e_2, e_3 ;
- soit en déterminant l'unique matrice A vérifiant $Y = AX$.
Finalement, on obtient

$$A = \begin{pmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ 1 & 1 & 2 \end{pmatrix}$$

Remarque On en déduit aisément la matrice de la symétrie s par rapport à P et parallèlement à D car $s = 2p - \text{Id}$.

$$\text{Mat}_{\mathcal{B}} s = 2A - I = \begin{pmatrix} -1 & -2 & -2 \\ 0 & -3 & -2 \\ 2 & 2 & 3 \end{pmatrix}$$

10.3.6.3 Réduction d'une projection, d'une symétrie

Soit f un endomorphisme d'un $\mathbb{K}$-espace vectoriel E connu par sa représentation matricielle A relative à une certaine base $\mathcal{B}$ de E.
Si $A^2 = A$ alors $f^2 = f$ et on sait :
- $F = \mathrm{Im} f$ et $G = \ker f$ sont supplémentaires ;
- f est la projection vectorielle sur F parallèlement à G.
Si $A^2 = I$ alors $f^2 = \mathrm{Id}$ et on sait :
- $F = \ker(f - \mathrm{Id})$ et $G = \ker(f + \mathrm{Id})$ sont supplémentaires ;
- f est la symétrie vectorielle par rapport à F et parallèlement à G.

Exemple Soient E un $\mathbb{R}$-espace vectoriel muni d'une base $\mathcal{B} = (e_1, e_2, e_3)$ et f l'endomorphisme de E dont la matrice dans $\mathcal{B}$ est

$$A = \begin{pmatrix} 0 & 1 & 1 \\ -1 & 2 & 1 \\ 1 & -1 & 0 \end{pmatrix}$$

On vérifie par le calcul $A^2 = A$ et on en déduit que f est une projection, plus précisément, c'est la projection sur $\mathrm{Im} f$ et parallèlement à $\ker f$.
Pour déterminer $\ker f$, on résout l'équation matricielle $AX = O_{3,1}$ i.e. le système

$$\begin{cases} x_2 + x_3 = 0 \\ -x_1 + 2x_2 + x_3 = 0 \\ x_1 - x_2 = 0 \end{cases}$$

On obtient $\ker f = \mathrm{Vect}(e_1 + e_2 - e_3)$
Par la formule du rang, on en déduit que $\mathrm{Im} f$ est un hyperplan.
Les vecteurs $f(e_1) = -e_2 + e_3$ et $f(e_2) = e_1 + 2e_2 - e_3$ appartiennent à cet hyperplan et donc $-x + y + z = 0$ est une équation de l'image de f.
Finalement f est la projection sur le plan $P : -x + y + z = 0$ parallèlement à la droite $\mathrm{Vect}(e_1 + e_2 - e_3)$.

Exemple Soient E un $\mathbb{R}$-espace vectoriel muni d'une base $\mathcal{B} = (e_1, e_2, e_3)$ et f l'endomorphisme de E dont la matrice dans $\mathcal{B}$ est

$$A = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ -2 & -2 & -3 \end{pmatrix}$$

On vérifie par le calcul $A^2 = I_3$ et on en déduit que f est une symétrie, plus précisément, c'est la symétrie par rapport à $\ker(f - \mathrm{Id})$ et parallèlement à $\ker(f + \mathrm{Id})$.
Pour déterminer $\ker(f - \mathrm{Id})$, on résout l'équation matricielle $(A - I_3)X = O_{3,1}$, pour déterminer $\ker(f + \mathrm{Id})$, c'est l'équation $(A + I_3)X = O_{3,1}$ que l'on résout.
Au terme des calculs, on obtient que f est la symétrie par rapport à la droite $D = \mathrm{Vect}(e_1 + e_2 - e_3)$ et parallèlement au plan $P : x + y + z = 0$.

10.3.6.4 Matrice d'un endomorphisme dans une base bien choisie

Exemple Soient E un $\mathbb{R}$-espace vectoriel muni d'une base $\mathcal{B} = (e_1, e_2, e_3)$ et f l'endomorphisme de E dont la matrice dans $\mathcal{B}$ est

$$A = \begin{pmatrix} -1 & 3 & -3 \\ -2 & 2 & -1 \\ -2 & 0 & 1 \end{pmatrix}$$

Considérons la famille $\mathcal{B}' = (\varepsilon_1, \varepsilon_2, \varepsilon_3)$ avec

$$\begin{cases} \varepsilon_1 = e_1 + e_2 + e_3 \\ \varepsilon_2 = e_1 - e_2 - 2e_3 \\ \varepsilon_3 = e_2 + e_3 \end{cases}$$

On vérifie aisément que $\mathcal{B}'$ est une base de E, par exemple en observant que cette une famille libre de trois vecteurs en dimension 3.
Formons la matrice de f dans cette base $\mathcal{B}'$.
Pour cela nous calculons les composantes dans $\mathcal{B}'$ des vecteurs $f(\varepsilon_1), f(\varepsilon_2), f(\varepsilon_3)$.

$f(\varepsilon_1)$ se déduit du calcul matriciel AX_1 avec $X_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ la colonne des composantes de ε_1 dans $\mathcal{B}$.

On obtient $f(\varepsilon_1) = -e_1 - e_2 - e_3 = -\varepsilon_1$.
De même, on obtient

$$f(\varepsilon_2) = 2e_1 - 2e_2 - 4e_3 = 2\varepsilon_2 \text{ et } f(\varepsilon_3) = e_2 + e_3 = \varepsilon_3$$

La matrice de f dans $\mathcal{B}'$ est donc la matrice diagonale

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Exemple Soit E un $\mathbb{R}$-espace vectoriel de dimension 3 muni d'une base $\mathcal{B} = (e_1, e_2, e_3)$ et f l'endomorphisme E dont la matrice dans $\mathcal{B}$ est

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

Montrons qu'il existe une base $\mathcal{B}' = (\varepsilon_1, \varepsilon_2, \varepsilon_3)$ de E telle que la matrice de f dans $\mathcal{B}'$ soit la matrice diagonale

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

Analyse : Supposons qu'une telle base existe.
On a $f(\varepsilon_1) = \varepsilon_1, f(\varepsilon_2) = -\varepsilon_2, f(\varepsilon_3) = 3\varepsilon_3$.
Cherchons de tels vecteurs. . . .
Pour déterminer un vecteur ε_1 convenable, on cherche les vecteurs $x = x_1e_1 + x_2e_2 + x_3e_3$ vérifiant $f(x) = x$.

Matriciellement, ce problème revient à résoudre l'équation $AX = X$ i.e. le système

$$\begin{cases} x_1 - 4x_2 - 4x_3 = x_1 \\ -2x_1 + 3x_2 + 4x_3 = x_2 \\ 2x_1 - x_3 = x_3 \end{cases}$$

Après résolution, on parvient à

$$\begin{cases} x_1 = x_3 \\ x_2 = -x_3 \end{cases}$$

Le vecteur $\varepsilon_1 = e_1 - e_2 + e_3$ est un donc un vecteur vérifiant $f(\varepsilon_1) = \varepsilon_1$.
On procédant de même avec les équations $AX = -X$ et $AX = 3X$, on obtient que les vecteurs $\varepsilon_2 = e_2 - e_3$ et $\varepsilon_3 = 2e_1 - 2e_2 + e_3$ vérifient $f(\varepsilon_2) = -\varepsilon_2$ et $f(\varepsilon_3) = 3\varepsilon_3$.
Synthèse : Considérons la famille $\mathcal{B}' = (\varepsilon_1, \varepsilon_2, \varepsilon_3)$ formée des vecteurs

$$\begin{cases} \varepsilon_1 = e_1 - e_2 + e_3 \\ \varepsilon_2 = e_2 - e_3 \\ \varepsilon_3 = 2e_1 - 2e_2 + e_3 \end{cases}$$

Par l'étude qui précède, on peut déjà affirmer que $f(\varepsilon_1) = \varepsilon_1, f(\varepsilon_2) = -\varepsilon_2$ et $f(\varepsilon_3) = 3\varepsilon_3$. Vérifions maintenant que la famille $\mathcal{B}'$ est libre.
Supposons $\lambda_1\varepsilon_1 + \lambda_2\varepsilon_2 + \lambda_3\varepsilon_3 = 0$.
On a $(\lambda_1 + 2\lambda_3)e_1 + (-\lambda_1 + \lambda_2 - 2\lambda_3)e_2 + (\lambda_1 - \lambda_2 + \lambda_3)e_3 = 0$.
Puisque la famille (e_1, e_2, e_3) est libre, on obtient le système

$$\begin{cases} \lambda_1 + 2\lambda_3 = 0 \\ -\lambda_1 + \lambda_2 - 2\lambda_3 = 0 \\ \lambda_1 - \lambda_2 + \lambda_3 = 0 \end{cases}$$

Après résolution, on parvient à $\lambda_1 = \lambda_2 = \lambda_3 = 0$.
La famille $\mathcal{B}' = (\varepsilon_1, \varepsilon_2, \varepsilon_3)$ est une famille libre formée de $3 = \dim E$ vecteurs de E, c'est donc une base de E. Par construction, la matrice de E est celle voulue

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

Remarque La représentation d'un endomorphisme par une matrice diagonale n'est pas toujours possible et, quand elle a lieu, celle-ci se fait avec des coefficients diagonaux bien précis. Cette problématique sera soulevée et résolue dans le cours de seconde année.

10.4 Formules de changement de base

10.4.1 Matrice de passage

Soit E un $\mathbb{K}$-espace vectoriel de dimension n muni de deux bases $\mathcal{B} = (e_1, \ldots, e_n)$ et $\mathcal{B}' = (e'_1, \ldots, e'_n)$.

Définition

On appelle matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$ la matrice $P = \text{Mat}_{\mathcal{B}}(\mathcal{B}') = \text{Mat}_{\mathcal{B}}(e'_1, \ldots, e'_n)$.

Exemple Soit E un $\mathbb{K}$-espace vectoriel de dimension 3 muni d'une base $\mathcal{B} = (e_1, e_2, e_3)$
Soit $\mathcal{B}' = (e'_1, e'_2, e'_3)$ la famille de vecteurs de E définie par :

$$\begin{cases} e'_1 = e_1 - e_2 + e_3 \\ e'_2 = e_2 - e_3 \\ e'_3 = -2e_1 + 2e_2 - e_3 \end{cases}$$

On vérifie aisément que la famille $\mathcal{B}'$ est libre et c'est donc une base de E.
La matrice de passage P de $\mathcal{B}$ à $\mathcal{B}'$ est

$$P = \begin{pmatrix} 1 & 0 & -2 \\ -1 & 1 & 2 \\ 1 & -1 & -1 \end{pmatrix}$$

Proposition

Si P est la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$ alors $P = \text{Mat}_{\mathcal{B}',\mathcal{B}}(\text{Id}_E)$.

dém. :
Par définition
$\text{Mat}_{\mathcal{B}',\mathcal{B}}(\text{Id}_E) = \text{Mat}_{\mathcal{B}}\left(\text{Id}_E(\mathcal{B}')\right) = \text{Mat}_{\mathcal{B}}(e'_1, \ldots, e'_n) = P$
□

Attention : Ici la matrice de l'endomorphisme Id_E n'est pas la matrice de l'identité car la représentation matricielle de l'identité est formée en choisissant une base à l'arrivée qui n'est pas a priori la même que la base au départ.

Proposition

Si P est la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$ alors P est inversible et son inverse P^{-1} est la matrice de passage de $\mathcal{B}'$ à $\mathcal{B}$.

dém. :
Notons $Q = \text{Mat}_{\mathcal{B}'}\mathcal{B}$ la matrice de passage de $\mathcal{B}'$ à $\mathcal{B}$.
On a $P = \text{Mat}_{\mathcal{B},\mathcal{B}'}(\text{Id}_E)$ et $Q = \text{Mat}_{\mathcal{B},\mathcal{B}'}(\text{Id}_E)$ donc $PQ = \text{Mat}_{\mathcal{B},\mathcal{B}}(\text{Id}_E \circ \text{Id}_E) = \text{Mat}_{\mathcal{B}}(\text{Id}_E) = I_n$
En vertu du théorème d'inversibilité, P est inversible et Q est son inverse.
□

Exemple Reprenons les notations de l'exemple précédent

$$\begin{cases} e'_1 = e_1 - e_2 + e_3 \\ e'_2 = e_2 - e_3 \\ e'_3 = -2e_1 + 2e_2 - e_3 \end{cases} \text{ et } P = \text{Mat}_{\mathcal{B}}\mathcal{B}' = \begin{pmatrix} 1 & 0 & -2 \\ -1 & 1 & 2 \\ 1 & -1 & -1 \end{pmatrix}$$

Pour former la matrice de passage inverse P^{-1}, il suffit d'exprimer les vecteurs de la base $\mathcal{B}$ en fonction de ceux de la base $\mathcal{B}'$. A l'aide du système précédent, et après calculs, on obtient

$$\begin{cases} e_1 = e'_1 + e'_2 \\ e_2 = 2e'_1 + e'_2 + e'_3 \\ e_3 = 2e'_1 + e'_3 \end{cases} \text{ et donc } P^{-1} = \begin{pmatrix} 1 & 2 & 2 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

Remarque Cette méthode est la méthode usuelle pour inverser une matrice de passage.

10.4.2 Nouvelles composantes d'un vecteur

Théorème

Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases d'un $\mathbb{K}$-espace vectoriel E de dimension n.
Si x est un vecteur de E dont on note X et X' les colonnes des composantes dans $\mathcal{B}$ et $\mathcal{B}'$ alors on a $X = PX'$ en notant P la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$.

dém. :
$X = \mathrm{Mat}_{\mathcal{B}}(x) = \mathrm{Mat}_{\mathcal{B}}(\mathrm{Id}_E(x)) = \mathrm{Mat}_{\mathcal{B}',\mathcal{B}}(\mathrm{Id}_E)\mathrm{Mat}_{\mathcal{B}'}(x) = PX'$.
□

Remarque On retient $\mathrm{Mat}_{\mathcal{B}}(x) = \mathrm{Mat}_{\mathcal{B}}\mathcal{B}'\mathrm{Mat}_{\mathcal{B}'}(x)$.

Corollaire

On a aussi $X' = P^{-1}X$

10.4.3 Nouvelle représentation d'une application linéaire

Théorème

Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases d'un $\mathbb{K}$-espace vectoriel E et $\mathcal{C}$ et $\mathcal{C}'$ deux bases d'un $\mathbb{K}$-espace vectoriel F.
Si u est une application linéaire de E vers F dont on note A la matrice relative aux bases $\mathcal{B}$ et $\mathcal{C}$, et A' celle relative aux bases $\mathcal{B}'$ et $\mathcal{C}'$ alors on a

$$A' = Q^{-1}AP$$

en notant P la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$ et Q celle de $\mathcal{C}$ à $\mathcal{C}'$.

dém. :
Soit $x \in E$ de colonnes composantes X et X' dans les bases $\mathcal{B}$ et $\mathcal{B}'$, et soit $y \in F$ de colonnes composantes Y et Y' dans les bases $\mathcal{C}$ et $\mathcal{C}'$. On a $X = PX'$ et $Y = QY'$.
On a alors la chaîne d'équivalences

$$y = u(x) \Leftrightarrow Y = AX \Leftrightarrow QY' = APX' \Leftrightarrow Y' = Q^{-1}APX'$$

Or la matrice A' est l'unique matrice telle que

$$y = u(x) \Leftrightarrow Y' = A'X'$$

et donc

$$A' = Q^{-1}AP$$

□

Remarque On « retient »$A' = \mathrm{Mat}_{\mathcal{B}',\mathcal{C}'}(u) = \mathrm{Mat}_{\mathcal{C}'}\mathcal{C}.\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(u).\mathrm{Mat}_{\mathcal{B}}\mathcal{B}'$.

10.4.4 Nouvelle représentation d'un endomorphisme

Théorème

Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases d'un $\mathbb{K}$-espace vectoriel E.
Si f est un endomorphisme de E dont on note A la matrice dans $\mathcal{B}$ et A' celle dans $\mathcal{B}'$ alors on a

$$A' = P^{-1}AP$$

avec P la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$.

dém. :
C'est un cas particulier de la formule de changement de base relative à une application linéaire.
□

Remarque On retient Ainsi $\mathrm{Mat}_{\mathcal{B}'}(f) = \mathrm{Mat}_{\mathcal{B}'}\mathcal{B}.\mathrm{Mat}_{\mathcal{B}}(f).\mathrm{Mat}_{\mathcal{B}}\mathcal{B}'$.

Attention : Erreur courante : n'écrire $A = PA'$ ce qui correspond à une transformation incomplète !

Exemple Soit

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

Nous allons calculer les puissances A^n pour tout $n \in \mathbb{N}$ en procédant à une transformation de la matrice A.
Soient E un $\mathbb{R}$-espace vectoriel muni d'une base $\mathcal{B} = (e_1, e_2, e_3)$ et f l'endomorphisme de E déterminé par

$$\mathrm{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} 2 & 1 & 1 \\ -3 & -2 & -1 \\ 3 & 5 & 4 \end{pmatrix} = A$$

Considérons la base $\mathcal{B}' = (\varepsilon_1, \varepsilon_2, \varepsilon_3)$ définie par

$$\begin{cases} e_1' = -e_2 + e_3 \\ e_2' = e_1 - e_2 + e_3 \\ e_3' = -e_1 + e_2 - 2e_3 \end{cases}$$

On vérifie aisément que la famille $\mathcal{B}'$ est libre est c'est donc une base de E.
On forme la matrice D de f dans la base $\mathcal{B}'$ en calculant $f(e_1')$, $f(e_2')$ et $f(e_3')$.
En procédant au calcul matriciel $Y = AX$, on obtient $f(e_1') = -e_1'$, $f(e_2') = 2e_2'$ et $f(e_3') = 3e_3'$.
On en déduit $D = \mathrm{Mat}_{\mathcal{B}'} f = \mathrm{diag}(-1, 2, 3)$
Par formule de changement de base, on a $A = PDP^{-1}$ avec P la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$.

$$P = \begin{pmatrix} 0 & 1 & -1 \\ -1 & -1 & 1 \\ 1 & 1 & -2 \end{pmatrix}$$

En exprimant les vecteurs de $\mathcal{B}$ en fonction de ceux de $\mathcal{B}'$, on obtient

$$\begin{cases} e_1 = -e_1' + e_2' \\ e_2 = -e_1' - e_2' - e_3' \\ e_3 = -e_2' - e_3' \end{cases}$$

et on en déduit la matrice de passage inverse

$$P^{-1} = \begin{pmatrix} -1 & -1 & 0 \\ 1 & -1 & -1 \\ 0 & -1 & -1 \end{pmatrix}$$

On peut alors calculer les puissances de A.
Pour tout $n \in \mathbb{N}$,

$$A^n = (PDP^{-1})^n = (PDP^{-1})(PDP^{-1})\dots(PDP^{-1}) = PD^nP^{-1}$$

avec

$$D^n = \text{diag}\left((-1)^n, 2^n, 3^n\right)$$

Au terme des calculs, on obtient

$$A^n = (-1)^n \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 0 \\ -1 & -1 & 0 \end{pmatrix} + 2^n \begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & 1 \\ 1 & -1 & -1 \end{pmatrix} + 3^n \begin{pmatrix} 0 & 1 & 1 \\ 0 & -1 & -1 \\ 0 & 2 & 2 \end{pmatrix}$$

10.4.5 Application : la trace

10.4.5.1 Trace d'une matrice carrée

Définition

> On appelle trace d'une matrice carrée $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$ le scalaire tr$(A) = \sum_{i=1}^{n} a_{i,i}$.

Proposition

> L'application tr : $\mathcal{M}_n(\mathbb{K}) \to \mathbb{K}$ est une forme linéaire.

dém. :
On vérifie immédiatement que tr$(\lambda A + \mu B) = \lambda$tr$A + \mu$tr$B$.
□

Théorème

> $\forall A \in \mathcal{M}_{n,p}(\mathbb{K}), \forall B \in \mathcal{M}_{p,n}(\mathbb{K})$, tr$(AB) =$ tr(BA).

dém. :
Introduisons les coefficients des matrices A et B : $A = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B = (b_{j,i}) \in \mathcal{M}_{p,n}(\mathbb{K})$.
Les matrices AB et BA sont carrées donc on peut calculer leur trace et on a

$$\text{tr}(AB) = \sum_{i=1}^{n} [AB]_{i,i} = \sum_{i=1}^{n}\sum_{j=1}^{p} a_{i,j}b_{j,i}$$

et

$$\text{tr}(BA) = \sum_{j=1}^{p} [BA]_{j,j} = \sum_{j=1}^{p}\sum_{i=1}^{n} b_{j,i}a_{i,j}$$

En permutant les deux sommes, on obtient tr$(BA) =$ tr(AB).
□

10.4.5.2 Trace d'un endomorphisme

Soient f un endomorphisme d'un $\mathbb{K}$-espace vectoriel E de dimension finie et $\mathcal{B}$, $\mathcal{B}'$ deux bases de E
Notons $P = \text{Mat}_{\mathcal{B}}\mathcal{B}'$, $A = \text{Mat}_{\mathcal{B}}(f)$, $A' = \text{Mat}_{\mathcal{B}'}(f)$.
La relation de changement de base permet d'écrire : $A = PA'P^{-1}$.
On a alors

$$\text{tr}(A) = \text{tr}(P(A'P^{-1})) = \text{tr}((A'P^{-1})P = \text{tr}(A')$$

Par suite, la trace de la matrice représentative de l'endomorphisme f est indépendante de la base choisie.

Définition

Cette quantité est appelée trace de l'endomorphisme f et est notée tr$f$.

Exemple $\text{tr}(\text{Id}_E) = n$

Théorème

La trace définit une forme linéaire sur $\mathcal{L}(E)$ vérifiant

$$\forall f, g \in \mathcal{L}(E), \text{tr}(f \circ g) = \text{tr}(g \circ f)$$

10.5 Rang d'une matrice

Rappel Si $\mathcal{F} = (x_1, \ldots, x_p)$ est une famille de vecteurs d'un $\mathbb{K}$-espace vectoriel E alors on appelle rang de la famille $\mathcal{F}$ la dimension de l'espace engendrée par $\mathcal{F}$
$\text{rg}\mathcal{F} = \dim \text{Vect}(x_1, \ldots, x_p)$.
Si E et F sont deux $\mathbb{K}$-espace vectoriels de dimensions finies et $u \in \mathcal{L}(E, F)$ alors on appelle rang de l'application linéaire u la dimension de l'image de u
$\text{rg}u = \dim \text{Im}u$.
Ces deux concepts sont liés puisque si $\mathcal{B} = (e_1, ..., e_p)$ est une base de E alors
$\text{rg}(u) = \text{rg}(u(e_1), ..., u(e_p))$.

10.5.1 Définition

Soit $A = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{K})$ de colonnes $C_1, C_2, \ldots, C_p$.
Les C_j sont des vecteurs du $\mathbb{K}$-espace vectoriel $\mathcal{M}_{n,1}(\mathbb{K})$, on peut donc considérer le rang de la famille de vecteurs $(C_1, C_2, \ldots, C_p)$.

Définition

On appelle rang d'une matrice $A \in \mathcal{M}_{n,p}(\mathbb{K})$ le rang de la famille $(C_1, C_2, \ldots, C_p)$ des colonnes de A. On note

$$\text{rg}A = \text{rg}(C_1, C_2, \ldots, C_p)$$

Théorème

Si $\mathcal{F} = (x_1, x_2, \ldots, x_p)$ une famille de vecteurs de d'un $\mathbb{K}$-espace vectoriel E et si A est la matrice de la famille $\mathcal{F}$ dans une certaine base $\mathcal{B}$ de E alors

$$\text{rg}A = \text{rg}(x_1, x_2, \ldots, x_p)$$

dém. :
Soit $\varphi : E \to \mathcal{M}_{n,1}(\mathbb{K})$ définie par $\varphi(x) = \text{Mat}_{\mathcal{B}}(x)$.
φ est un isomorphisme de $\mathbb{K}$-espace vectoriel.
Puisque les colonnes de A sont les $C_j = \varphi(x_j)$, on a

$$\text{rg}A = \text{rg}(\varphi(x_1), \ldots, \varphi(x_p)) = \dim \text{Vect}(\varphi(x_1), \ldots, \varphi(x_p))$$

Or puisque φ est une application linéaire injective on a

$$\dim \text{Vect}(\varphi(x_1), \ldots, \varphi(x_p)) = \dim \text{Vect}(x_1, \ldots, x_p) = \text{rg}(x_1, \ldots, x_p)$$

Ainsi

$$\text{rg}(A) = \text{rg}(x_1, x_2, \ldots, x_p)$$

□

Théorème

Si u est une application d'un $\mathbb{K}$-espace vectoriel E vers un $\mathbb{K}$-espace vectoriel F et si A est la matrice de u relative à des bases $\mathcal{B}$ et $\mathcal{C}$ de E et F alors
$\text{rg}A = \text{rg}u$.

dém. :
Si l'on note $e_1, \ldots, e_p$ les vecteurs constituant la base $\mathcal{B}$ alors A est la matrice de la famille $(u(e_1), \ldots, u(e_p))$ dans la base $\mathcal{C}$ et par suite $\text{rg}A = \text{rg}(u(e_1), \ldots, \text{rg}(u(e_p)) = \text{rg}u$.
□

10.5.2 Propriétés du rang d'une matrice

Soient $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et $u \in \mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$ l'application linéaire canoniquement associée à la matrice A c'est-à-dire l'application linéaire représentée par la matrice A relativement aux base canoniques de $\mathbb{K}^p$ et $\mathbb{K}^n$.
Puisque $\text{rg}A = \text{rg}u$, les propriétés relatives au rang d'applications linéaires se transposent aux matrices.

Proposition

$\forall A \in \mathcal{M}_{n,p}(\mathbb{K}), \text{rg}(A) \leqslant \min(n, p)$.

dém. :
Car $\text{rg}u \leqslant \min(\dim E, \dim F)$ pour $u \in \mathcal{L}(E, F)$.
□

Proposition

$\forall A \in \mathcal{M}_{n,p}(\mathbb{K}), \forall B \in \mathcal{M}_{p,q}(\mathbb{K}), \text{rg}(AB) \leqslant \min(\text{rg}(A), \text{rg}(B))$.
De plus :
Si A est une matrice carrée inversible alors $\text{rg}(AB) = \text{rg}(B)$
Si B est une matrice carrée inversible alors $\text{rg}(AB) = \text{rg}(A)$.

dém. :
Car rg$(u \circ v) \leqslant \min(\text{rg}u, \text{rg}v)$.
De plus rg$(u \circ v) =$ rg$u$ si v surjective et rg$(u \circ v) =$ rg$v$ si u injective.
□

Remarque On ne modifie par le rang d'une matrice en multipliant celle-ci par une matrice inversible.

Théorème

Soit $A \in \mathcal{M}_n(\mathbb{K})$. On a équivalence entre :
(i) A est inversible ;
(ii) rg$(A) = n$.

dém. :
A est inversible si, et seulement si, u est automorphisme de $\mathbb{K}^n$ i.e. si, et seulement si, rg$u = n$
□

10.5.3 Caractérisation théorique du rang

Pour $0 \leqslant r \leqslant \min(n, p)$, on note J_r la matrice de $\mathcal{M}_{n,p}(\mathbb{K})$ définie par

$$J_r = \left(\begin{array}{ccc|c} 1 & & 0 & \\ & \ddots & & 0 \\ 0 & & 1 & \\ \hline & 0 & & 0 \end{array}\right)$$

Précisément, les coefficients de $J_r \in \mathcal{M}_{n,p}(\mathbb{K})$ sont nuls sauf le coefficient d'indice (i, i) est égal à 1 pour $i \in \{1, \ldots, r\}$.

Exemple Si $r = 0$ alors $J_r = O_{n,p}$.
Si $r = n \leqslant p$ alors

$$J_r = \left(\begin{array}{ccc|c} 1 & & 0 & \\ & \ddots & & 0 \\ 0 & & 1 & \end{array}\right)$$

Si $r = p \leqslant n$ alors

$$J_r = \left(\begin{array}{ccc} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \\ \hline & 0 & \end{array}\right)$$

Si $r = p = n$ alors $J_r = I_r$.

Proposition

rg$(J_r) = r$.

dém. :
Si $\mathcal{B} = (e_1, \ldots, e_n)$ désigne la base canonique de $\mathbb{K}^n$ alors la matrice J_r peut se voir comme la matrice dans $\mathcal{B}$ de la famille $(e_1, \ldots, e_r, 0, \ldots, 0)$ formée de p vecteurs. Cette dernière est de rang r car la sous-famille $(e_1, \ldots, e_r)$ est libre et donc rg$J_r = r$.
□

Théorème

Soient $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et $r \in \mathbb{N}$ tel que $0 \leqslant r \leqslant \min(n,p)$.
On a équivalence entre :
(i) $\text{rg}(A) = r$;
(ii) $\exists P \in \text{GL}_p(\mathbb{K}), \exists Q \in \text{GL}_n(\mathbb{K}), A = QJ_rP$.

dém. :
(ii) $\Rightarrow$ (i) : C'est immédiat car $\text{rg}J_r = r$ et on sait qu'on ne modifie pas le rang en multipliant par une matrice inversible.
(i) $\Rightarrow$ (ii) Soient E et F deux $\mathbb{K}$-espaces vectoriels de dimensions p et n munis de bases $\mathcal{B}$ et $\mathcal{C}$ et $u \in \mathcal{L}(E,F)$ l'application linéaire déterminée par $\text{Mat}_{\mathcal{B},\mathcal{C}}u = A$. Puisque $r = \text{rg}A = \text{rg}u$, on a $\dim \ker u = p - r$ en vertu de la formule du rang.
Soit H un supplémentaire de $\ker u$ dans E et $\mathcal{B}' = (e'_1, \dots, e'_r, e'_{r+1}, \dots, e'_p)$ une base adaptée à la supplémentarité de H et $\ker u$ dans E..
Posons $f'_1 = u(e'_1), \dots, f'_r = u(e'_r)$.
Supposons $\lambda_1 f'_1 + \dots + \lambda_r f'_r = 0$.
On a $u(\lambda_1 e'_1 + \dots + \lambda_r e'_r) = 0$ donc $\lambda_1 e'_1 + \dots + \lambda_r e'_r \in \ker u$.
Or $\lambda_1 e'_1 + \dots + \lambda_r e'_r \in H$ et $\ker u \cap H = \{0\}$ donc $\lambda_1 e'_1 + \dots + \lambda_r e'_r = 0$.
Puisque la famille $(e'_1, \dots, e'_r)$ est libre, on obtient $\lambda_1 = \ldots = \lambda_r = 0$.
Ainsi la famille $(f'_1, \dots, f'_r)$ est libre. Complétons-la en une base de F de la forme $\mathcal{C}' = (f'_1, \dots, f'_n)$.
Par construction, la matrice de u relative aux bases $\mathcal{B}'$ et $\mathcal{C}'$ est la matrice J_r et par formule de changement de base, on obtient $A = QJ_rP$ avec $P = \text{Mat}_{\mathcal{B}}\mathcal{B}' \in \text{GL}_p(\mathbb{K})$ et $Q = \text{Mat}_{\mathcal{C}'}\mathcal{C} \in \text{GL}_n(\mathbb{K})$.
□

Corollaire

$\forall A \in \mathcal{M}_{n,p}(\mathbb{K}), \text{rg}({}^tA) = \text{rg}A$.

dém. :
Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$. On peut écrire $A = QJ_rP$ avec P, Q inversibles et $r = \text{rg}A$.
On a alors ${}^tA = {}^tP{}^tJ_r{}^tQ$ avec ${}^tP, {}^tQ$ inversibles donc $\text{rg}{}^tA = \text{rg}{}^tJ_r = r$ puisque la matrice ${}^tJ_r$ est analogue à la matrice J_r sauf qu'elle est de type (p,n) au lieu d'être de type (n,p).
□

PropSoit $A \in \mathcal{M}_{n,p}(\mathbb{K})$ dont on note $C_1, \dots, C_p$ les colonnes et $L_1, \dots, L_n$ les lignes.
On a $\text{rg}(A) = \text{rg}(C_1, \dots, C_p) = \text{rg}(L_1, \dots, L_n)$.
dém. :
Par définition, on a $\text{rg}(A) = \text{rg}(C_1, \dots, C_p)$
Les lignes $L_1, \dots, L_n$ sont des vecteurs de l'espace $\mathcal{M}_{1,p}(\mathbb{K})$.
Notons $\mathcal{B} = (E_1, \dots, E_p)$ la base canonique de l'espace $\mathcal{M}_{1,p}(\mathbb{K})$.
Si L est une ligne $\begin{pmatrix} \alpha_1 & \dots & \alpha_p \end{pmatrix}$ de $\mathcal{M}_{1,p}(\mathbb{K})$ alors la colonne des composantes de L dans $\mathcal{B}$ est

$$\text{Mat}_{\mathcal{B}}(L) = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_p \end{pmatrix} =^t L$$

Par suite la matrice représentative de la famille de vecteurs $(L_1, \dots, L_n)$ dans la base $\mathcal{B}$ est

$$\text{Mat}_{\mathcal{B}}(L_1, \dots, L_n) = {}^tA$$

On en déduit $\text{rg}(L_1, \dots, L_n) = \text{rg}({}^tA) = \text{rg}A$.
□

10.5.4 Opérations élémentaires sur les matrices

10.5.4.1 Préliminaire

Proposition

> Les matrices de $\mathcal{M}_n(\mathbb{K})$ suivantes sont inversibles :
> 1) $C = I_n + \lambda E_{i,j}$ avec $\lambda \in \mathbb{K}$ et $1 \leqslant i \neq j \leqslant n$;
> 2) $D = I_n - E_{i,i} + \lambda E_{i,i}$ avec $\lambda \in \mathbb{K}^\star$ et $1 \leqslant i \leqslant n$;
> 3) $E = I_n - E_{i,i} - E_{j,j} + E_{i,j} + E_{j,i}$ avec $1 \leqslant i \neq j \leqslant n$.

dém. :
1) La matrice C est triangulaire supérieure si $i < j$ et triangulaire inférieure si $i > j$. Dans les deux cas, ses coefficients diagonaux sont tous non nuls car égaux à 1, cette matrice est donc inversible.
2) La matrice D est diagonale et ses coefficients diagonaux sont non nuls car valent 1 ou λ. La matrice D est donc inversible.
3) La matrice E vérifie $E^2 = I_n$, elle est donc inversible et son inverse lui est égale.
□

10.5.4.2 Opérations élémentaires sur les lignes

Soit $A = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{K})$.
Pour $E_{i,j}$ la matrice élémentaire d'indice (i, j) de $\mathcal{M}_n(\mathbb{K})$, on a

$$E_{i,j}A = \begin{pmatrix} & 0 & \\ a_{j,1} & \cdots & a_{j,p} \\ & 0 & \end{pmatrix} = \begin{pmatrix} 0 \\ L_j \\ 0 \end{pmatrix}$$

avec la ligne L_j positionnée en i ème ligne.
1) Soit $C = I_n + \lambda E_{i,j}$ avec $\lambda \in \mathbb{K}$ et $1 \leqslant i \neq j \leqslant n$.
$CA = A + \lambda E_{i,j}A$ est la matrice obtenue en ajoutant λ fois la $j$ème colonne de A à sa $i$ème ligne.
Cette manipulation est notée $L_i \leftarrow L_i + \lambda.L_j$.
2) Soit $D = I_n - E_{i,i} + \lambda E_{i,i}$ avec $\lambda \in \mathbb{K}^\star$ et $1 \leqslant i \leqslant n$.
$DA = A - E_{i,i}A + \lambda E_{i,i}A$ est la matrice obtenue en multipliant par $\lambda \neq 0$ la $i$ème ligne de A.
Cette manipulation est notée $L_i \leftarrow \lambda.L_i$.
3) Soit $E = I_n - E_{i,i} - E_{j,j} + E_{i,j} + E_{j,i}$ avec $1 \leqslant i \neq j \leqslant n$.
$EA = A - E_{i,i}A - E_{j,j}A + E_{i,j}A + E_{j,i}A$ est la matrice obtenue en échangeant les $i$ème et $j$ème lignes de A.
Cette manipulation est notée $L_i \leftrightarrow L_j$.

Définition

> Les manipulations $L_i \leftarrow L_i + \lambda.L_j$, $L_i \leftarrow \lambda.L_i$ et $L_i \leftrightarrow L_j$ sont appelées opérations élémentaires sur les lignes de A.

Proposition

> Ces manipulations conservent le rang de A.

dém. :
Ces manipulations correspondent à des multiplications par des matrices inversibles et multiplier par une matrice inversible conserve le rang.
□

10.5.4.3 Opérations élémentaires sur les colonnes

Soit $A = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{K})$.

Pour $E_{i,j}$ la matrice élémentaire d'indice (i,j) de $\mathcal{M}_p(\mathbb{K})$, on a

$$AE_{i,j} = \begin{pmatrix} & a_{1,i} & \\ 0 & \vdots & 0 \\ & a_{n,i} & \end{pmatrix} = \begin{pmatrix} 0 & C_i & 0 \end{pmatrix}$$

avec la colonne C_i positionnée en $j$ème colonne.
1) Soit $C = I_p + \lambda E_{i,j}$ avec $\lambda \in \mathbb{K}$ et $1 \leqslant i \neq j \leqslant p$.
$AC = A + \lambda AE_{i,j}$ est la matrice obtenue en ajoutant λ fois la $i$ème colonne de A à sa $j$ème colonne.
Cette manipulation est notée $C_j \leftarrow C_j + \lambda.C_i$
2) Soit $D = I_p - E_{i,i} + \lambda E_{i,i}$ avec $\lambda \in \mathbb{K}^\star$ et $1 \leqslant i \leqslant p$.
$AD = A - AE_{i,i} + \lambda AE_{i,i}$ est la matrice obtenue en multipliant par $\lambda \neq 0$ la $i$ème colonne de A.
Cette manipulation est notée $C_i \leftarrow \lambda.C_i$
3) Soit $E = I_p - E_{i,i} - E_{j,j} + E_{i,j} + E_{j,i}$ avec $1 \leqslant i \neq j \leqslant p$.
$AE = A - AE_{i,i} - AE_{j,j} + AE_{i,j} + AE_{j,i}$ est la matrice obtenue en échangeant les $i$ème et $j$ème colonnes de A.
Cette manipulation est notée $C_i \leftrightarrow C_j$

Définition

Les manipulations $C_j \leftarrow C_j + \lambda.C_i$, $C_i \leftarrow \lambda.C_i$ et $C_i \leftrightarrow C_j$ sont appelées opérations élémentaires sur les colonnes de A.

Proposition

Ces manipulation conservent le rang de A.

Remarque On peut observer que lorsqu'on multiplie A
- par la droite, on opère sur les lignes de A ;
- par la gauche, on opère sur les colonnes de A.

Exemple Soient $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$ et $D = \text{diag}(\lambda_1, ..., \lambda_n)$. On a

$$DA = \begin{pmatrix} \lambda_1 a_{1,1} & \cdots & \lambda_1 a_{1,n} \\ \vdots & & \vdots \\ \lambda_n a_{n,1} & \cdots & \lambda_n a_{n,n} \end{pmatrix} = (\lambda_i a_{i,j})$$

et

$$AD = \begin{pmatrix} \lambda_1 a_{1,1} & \cdots & \lambda_n a_{1,n} \\ \vdots & & \vdots \\ \lambda_1 a_{n,1} & \cdots & \lambda_n a_{n,n} \end{pmatrix} = (\lambda_j a_{i,j})$$

10.5.5 Méthode du pivot de Gauss

Soit $A = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{K})$. On désire calculer le rang de A de façon algorithmique.
Si $A = O_{n,p}$ alors $\text{rg}A = 0$.
Sinon A possède un coefficient non nul $p_1 = a_{i,j}$.

Par les opérations $L_i \leftrightarrow L_1$ et $C_j \leftrightarrow C_1$ on amène le coefficient p_1 en position $(1,1)$ et on dispose alors d'une matrice de la matrice de la forme :

$$\left(\begin{array}{c|c} p_1 & \star \\ \hline \alpha_2 & \\ \vdots & \star \\ \alpha_n & \end{array}\right) \text{ avec } p_1 \neq 0 \text{ appelé 1er pivot}$$

On annule les coefficients en dessous de p_1 par des opérations élémentaires $L_i \leftarrow L_i - \dfrac{\alpha_i}{p_1} L_1$ avec $2 \leqslant i \leqslant n$. On parvient alors à une matrice de la forme

$$\left(\begin{array}{c|c} p_1 & \star \\ \hline 0 & \\ \vdots & B \\ 0 & \end{array}\right) \text{ avec } p_1 \neq 0$$

On recommence ce processus avec la matrice B tant que la matrice obtenue est non nulle. A terme, on obtient :

$$\left(\begin{array}{ccc|c} p_1 & & \star & \\ & \ddots & & \star \\ 0 & & p_r & \\ \hline & & & O_{n-r,p-r} \end{array}\right) \text{ avec } p_1, \dots, p_r \neq 0$$

On peut alors conclure que la matrice A est de rang r.
En effet, en suivant le principe qui précédent, on peut, en opérant sur les colonnes transformer A en

$$\left(\begin{array}{ccc|c} p_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & p_r & \\ \hline & 0 & & O_{n-r,p-r} \end{array}\right) \text{ puis en } \left(\begin{array}{ccc|c} 1 & & 0 & \\ & \ddots & & 0 \\ 0 & & 1 & \\ \hline & 0 & & O_{n-r,p-r} \end{array}\right) = J_r$$

qui est de rang r

Remarque En opérant essentiellement sur les lignes, ce processus a transformé la matrice A en

$$\left(\begin{array}{ccc|c} p_1 & & \star & \\ & \ddots & & \star \\ 0 & & p_r & \\ \hline & 0 & & O_{n-r,p-r} \end{array}\right)$$

De manière symétrique, mais en opérant essentiellement sur les colonnes, on peut aussi transformer la matrice A en

$$\left(\begin{array}{ccc|c} p_1 & & 0 & \\ & \ddots & & 0 \\ \star & & p_r & \\ \hline & \star & & O_{n-r,p-r} \end{array}\right)$$

10.5.6 Calculs de rang

10.5.6.1 Rang d'une matrice

Exemple Calculons le rang de la matrice

$$\begin{pmatrix} 1 & 2 & 1 \\ 1 & 3 & 2 \\ 1 & 1 & 0 \end{pmatrix}$$

On a

$$\operatorname{rg}\begin{pmatrix} 1 & 2 & 1 \\ 1 & 3 & 2 \\ 1 & 1 & 0 \end{pmatrix} = \operatorname{rg}\begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & -1 & -1 \end{pmatrix} = \operatorname{rg}\begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} = 2$$

Exemple Calculons le rang de

$$\begin{pmatrix} 1 & 1 & 1 \\ -1 & 1 & -1 \\ 2 & 1 & 1 \\ 1 & 2 & -1 \end{pmatrix}$$

En opérant par les lignes

$$\operatorname{rg}\begin{pmatrix} 1 & 1 & 1 \\ -1 & 1 & -1 \\ 2 & 1 & 1 \\ 1 & 2 & -1 \end{pmatrix} = \operatorname{rg}\begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & -1 & -1 \\ 0 & 1 & -2 \end{pmatrix} = \operatorname{rg}\begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & -2 \end{pmatrix} = \operatorname{rg}\begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix} = 3$$

En opérant pas les colonnes

$$\operatorname{rg}\begin{pmatrix} 1 & 1 & 1 \\ -1 & 1 & -1 \\ 2 & 1 & 1 \\ 1 & 2 & -1 \end{pmatrix} = \operatorname{rg}\begin{pmatrix} 1 & 0 & 0 \\ -1 & 2 & 0 \\ 2 & -1 & -1 \\ 1 & 1 & -2 \end{pmatrix} = 3$$

Attention : On prend garde à effectuer les opérations élémentaires successivement et non simultanément :

Exemple Via

$$\begin{cases} L_2 \leftarrow L_2 + L_1 \\ L_1 \leftarrow L_1 + L_2 \end{cases}$$

la matrice

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

devient

$$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \text{ et non } \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \text{ !}$$

10.5.6.2 Rang d'une famille de vecteurs

Rappel Soient E un $\mathbb{K}$-espace vectoriel de dimension n et $\mathcal{F}$ une famille de p vecteurs de E. On a :
$\mathcal{F}$ est libre si, et seulement si, $\text{rg}(\mathcal{F}) = p$,
$\mathcal{F}$ est génératrice si, et seulement si, $\text{rg}(\mathcal{F}) = n$,
$\mathcal{F}$ est une base si, et seulement si, $\text{rg}(\mathcal{F}) = n = p$.

Exemple Soient $\mathcal{B} = (e_1, e_2, e_3)$ la base canonique de $\mathbb{R}^3$ et $\mathcal{B}' = (e'_1, e'_2, e'_3)$ avec
$e'_1 = e_2 + e_3$, $e'_2 = e_1 + e_2 + 2e_3$ et $e'_3 = e_1 + 2e_3$.

$$\text{rg}\mathcal{B}' = \text{rg}(\text{Mat}_{\mathcal{B}}\mathcal{B}') = \text{rg}\begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 2 & 2 \end{pmatrix}$$

En échangeant les deux premières lignes

$$\text{rg}\begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 2 & 2 \end{pmatrix} = \text{rg}\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 2 & 2 \end{pmatrix} = \text{rg}\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix} = \text{rg}\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} = 3$$

Puisque la famille $\mathcal{B}'$ est de rang 3, c'est une base de $\mathbb{R}^3$.

10.5.6.3 Rang d'une application linéaire

Rappel Soient E un $\mathbb{K}$-espace vectoriel de dimension p, F un $\mathbb{K}$-espace vectoriel de dimension n et $u \in \mathcal{L}(E, F)$. On a :
u injective si, et seulement si, $\text{rg}(u) = p$
u surjective si, et seulement si, $\text{rg}(u) = n$
u isomorphisme si, et seulement si, $\text{rg}(u) = n = p$.

Exemple Soit f l'endomorphisme de $\mathbb{R}^3$ dont la matrice dans la base canonique est

$$\begin{pmatrix} 1 & 1 & 2 \\ 0 & 2 & 1 \\ 2 & 1 & -1 \end{pmatrix}$$

On a

$$\text{rg}f = \text{rg}\begin{pmatrix} 1 & 1 & 2 \\ 0 & 2 & 1 \\ 2 & 1 & -1 \end{pmatrix} = \text{rg}\begin{pmatrix} 1 & 1 & 2 \\ 0 & 2 & 1 \\ 0 & -1 & -5 \end{pmatrix} = \text{rg}\begin{pmatrix} 1 & 1 & 2 \\ 0 & 2 & 1 \\ 0 & 0 & -9/2 \end{pmatrix} = 3$$

L'endomorphisme f est un automorphisme de $\mathbb{R}^3$

10.5.6.4 Rang dépendant d'un paramètre

Pour déterminer le rang d'une matrice dépendant d'un paramètre, on cherche à calculer celui-ci de la façon la plus générale possible avant de traiter les valeurs particulières des paramètres.

Exemple Déterminons le rang de

$$M = \begin{pmatrix} 1 & 2 & 1 \\ 1 & m & 2 \\ m & 1 & 2 \end{pmatrix}$$

en fonction de $m \in \mathbb{R}$.

$$\text{rg}\begin{pmatrix} 1 & 2 & 1 \\ 1 & m & 2 \\ m & 1 & 2 \end{pmatrix} = \text{rg}\begin{pmatrix} 1 & 2 & 1 \\ 0 & m-2 & 1 \\ 0 & 1-2m & 2-m \end{pmatrix}$$

En échangeant les deux dernières colonnes

$$\text{rg}\begin{pmatrix} 1 & 2 & 1 \\ 1 & m & 2 \\ m & 1 & 2 \end{pmatrix} = \text{rg}\begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & m-2 \\ 0 & 2-m & 1-2m \end{pmatrix} = \text{rg}\begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & m-2 \\ 0 & 0 & x \end{pmatrix}$$

avec $x = (1-2m) + (m-2)(m-2) = (m-1)(m-5)$
On en déduit

$$\text{rg}M = \begin{cases} 3 & \text{si } m \neq 1, 5 \\ 2 & \text{si } m = 1 \text{ ou } 5 \end{cases}$$

Exemple Soient $a, b, c \in \mathbb{R}$.
On considère les vecteurs de $\mathbb{R}^3$ suivant $u = (1, a, a^2)$, $v = (1, b, b^2)$ et $w = (1, c, c^2)$.
A quelle condition la famille (u, v, w) forme-t-elle une base de $\mathbb{R}^3$?

$$\text{rg}(u,v,w) = \text{rg}\begin{pmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{pmatrix} = \text{rg}\begin{pmatrix} 1 & 1 & 1 \\ 0 & b-a & c-a \\ 0 & b^2-a^2 & c^2-a^2 \end{pmatrix} = \text{rg}\begin{pmatrix} 1 & 1 & 1 \\ 0 & b-a & c-a \\ 0 & 0 & x \end{pmatrix}$$

avec $x = c^2 - a^2 - (a+b)(c-a) = (c-a)(c-b)$.
On en déduit $\text{rg}(u, v, w) = 3 \Leftrightarrow a, b, c$ deux à deux distincts.
Ainsi la famille (u, v, w) est une base de $\mathbb{R}^3$ si, et seulement si, a, b, c sont deux à deux distincts.

Exemple Soit f l'endomorphisme de $\mathbb{R}^3$ défini par $f(x, y, z) = (y + z, z + x, x + y)$.
Pour quelles valeurs $\lambda \in \mathbb{R}$, l'équation vectorielle $f(u) = \lambda.u$ possède-t-elle une solution autre que le vecteur nul ?
Les solutions de l'équation $f(u) = \lambda.u$ sont les vecteurs u vérifiant $(f - \lambda\text{Id})(u) = 0$.
Par suite, l'ensemble des solutions de l'équation $f(u) = \lambda.u$ est l'espace $\ker(f - \lambda\text{Id})$
Déterminer les $\lambda \in \mathbb{K}$ tels que l'équation $f(u) = \lambda.u$ possède d'autres solutions que le vecteur nul revient à chercher les $\lambda \in \mathbb{K}$ vérifiant $\ker(f - \lambda\text{Id}) \neq \{0\}$ i.e. $f - \lambda\text{Id}$ non injectif. Nous déterminons ceux-ci en recherchant les $\lambda \in \mathbb{K}$ tels que $\text{rg}(f - \lambda\text{Id}) < 3$

$$\text{rg}(f - \lambda\text{Id}) = \text{rg}\begin{pmatrix} -\lambda & 1 & 1 \\ 1 & -\lambda & 1 \\ 1 & 1 & -\lambda \end{pmatrix}$$

En échangeant la première et la dernière colonne

$$\mathrm{rg}(f-\lambda\mathrm{Id}) = \mathrm{rg}\begin{pmatrix} 1 & 1 & -\lambda \\ 1 & -\lambda & 1 \\ -\lambda & 1 & 1 \end{pmatrix} = \mathrm{rg}\begin{pmatrix} 1 & 1 & -\lambda \\ 0 & -\lambda-1 & 1+\lambda \\ 0 & 1+\lambda & 1-\lambda^2 \end{pmatrix} = \mathrm{rg}\begin{pmatrix} 1 & 1 & -\lambda \\ 0 & -\lambda-1 & 1+\lambda \\ 0 & 0 & x \end{pmatrix}$$

avec $x = (1-\lambda^2) + (1+\lambda) = (1+\lambda)(2-\lambda)$
On en déduit $\mathrm{rg}(f-\lambda\mathrm{Id}) < 3 \Leftrightarrow \lambda = 1$ ou 2.

10.5.7 Inversion de matrice

On désire calculer l'inverse de $A \in \mathrm{GL}_n(\mathbb{K})$.
En suivant la méthode du pivot de Gauss, il est possible en opérant uniquement sur les lignes de transformer la matrice A en une matrice triangulaire supérieure

$$\begin{pmatrix} p_1 & & \star \\ & \ddots & \\ 0 & & p_n \end{pmatrix} \text{ avec } p_1, \dots, p_n \neq 0$$

En opérant à nouveau sur ces lignes on peut poursuivre la transformation de A en

$$\begin{pmatrix} 1 & & \star \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$$

puis en la matrice identité

$$\begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$$

Effectuer les opérations élémentaires sur les lignes transformant la matrice A en l'identité, revient à multiplier la matrice A à gauche par des matrices $T_1, \dots, T_m$. Ainsi la transformation de la matrice A en l'identité correspond à écrire $T_m \dots T_1.A = I_n$.
En multipliant à droite par A^{-1}, on en déduit $T_m \dots T_1.I_n = A^{-1}$
Ainsi la série d'opérations élémentaires sur les ligne qui transforme la matrice A en I_n transforme parallèlement la matrice I_n en A^{-1}.
Nous pouvons exploiter cette idée pour calculer A^{-1} de façon algorithmique.

Exemple Etudions l'éventuel inverse de $A = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & 0 \\ 2 & 1 & 1 \end{pmatrix}$

$$\begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & 0 \\ 2 & 1 & 1 \end{pmatrix} \left| \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right.$$

$$\begin{pmatrix} 1 & 1 & -1 \\ 0 & 0 & 1 \\ 0 & -1 & 3 \end{pmatrix} \left| \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix} \right.$$

$$\begin{pmatrix} 1 & 1 & -1 \\ 0 & -1 & 3 \\ 0 & 0 & 1 \end{pmatrix} \Bigg| \begin{pmatrix} 1 & 0 & 0 \\ -2 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix}$$
$$\begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & -3 \\ 0 & 0 & 1 \end{pmatrix} \Bigg| \begin{pmatrix} 1 & 0 & 0 \\ 2 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}$$
$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \Bigg| \begin{pmatrix} 0 & 1 & 0 \\ -1 & 3 & -1 \\ -1 & 1 & 0 \end{pmatrix}$$
$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \Bigg| \begin{pmatrix} 1 & -2 & 1 \\ -1 & 3 & -1 \\ -1 & 1 & 0 \end{pmatrix}$$

On en déduit que A est inversible et

$$A^{-1} = \begin{pmatrix} 1 & -2 & 1 \\ -1 & 3 & -1 \\ -1 & 1 & 0 \end{pmatrix}$$

Remarque On peut aussi transformer A en I_n, en ne manipulant que les colonnes de A, les opérations correspondantes transforme alors aussi I_n en A^{-1}.

Attention : On manipule lignes ou colonnes mais jamais les deux !

10.6 Systèmes d'équations linéaires

10.6.1 Positionnement du problème.

Soient $a_{i,j} \in \mathbb{K}$ et $b_i \in \mathbb{K}$ pour tous $i \in \{1, \dots, n\}$, $j \in \{1, \dots, p\}$.
On cherche tous les p uplets $(x_1, \dots, x_p) \in \mathbb{K}^p$ solutions du système :

$$\Sigma : \begin{cases} a_{11}x_1 + \cdots + a_{1p}x_p = b_1 \\ \vdots \\ a_{n1}x_1 + \cdots + a_{np}x_p = b_n \end{cases}$$

Définition

Σ est appelé système d'équations linéaires à n équations et p inconnues.
On note $\mathcal{S}_\Sigma \subset \mathbb{K}^p$ l'ensemble des solutions du système Σ.

Posons $x = (x_1, \dots, x_p) \in \mathbb{K}^p$, $b = (b_1, \dots, b_n) \in \mathbb{K}^n$ et considérons l'application linéaire

$$u : (x_1, \dots, x_p) \mapsto (a_{1,1}x_1 + \cdots + a_{1,p}x_p, \dots, a_{n,1}x_1 + \cdots + a_{n,p}x_p)$$

Le système Σ équivaut à l'équation $u(x) = b$.

Définition

L'équation $u(x) = b$ est appelée équation vectorielle du système Σ.

Posons $A = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{K})$, $B = (b_i) \in \mathcal{M}_{n,1}(\mathbb{K})$ et $X = (x_j) \in \mathcal{M}_{p,1}(\mathbb{K})$.
Le système Σ équivaut à l'équation $AX = B$.

Définition

L'équation $AX = B$ est appelée équation matricielle du système Σ.

Définition

On appelle rang du système Σ le naturel $\text{rg}\Sigma = \text{rg}A = \text{rg}u$.

10.6.2 Compatibilité d'un système

Définition

On dit que le système Σ est compatible si $\mathcal{S}_\Sigma \neq \emptyset$.

Exemple Le système

$$\begin{cases} x + y = 0 \\ x + y = 1 \end{cases}$$

est visiblement un système incompatible.

Considérons le système

$$\Sigma : \begin{cases} a_{11}x_1 + \cdots + a_{1p}x_p = b_1 \\ \vdots \\ a_{n1}x_1 + \cdots + a_{np}x_p = b_n \end{cases}$$

d'équation matricielle $AX = B$.

Définition

On note $\left(\begin{array}{c|c} A & B \end{array} \right)$ la matrice de $\mathcal{M}_{n,p+1}(\mathbb{K})$ constituée des colonnes de la matrice A suivies de la colonne B.

Théorème

Le système Σ est compatible si, et seulement si, $\text{rg} \left(\begin{array}{c|c} A & B \end{array} \right) = \text{rg}A$.

dém. :
Notons $C_1, \ldots, C_p$ les colonnes de A.

$$(x_1, \ldots, x_p) \in \mathcal{S}_\Sigma \Leftrightarrow x_1C_1 + \cdots + x_pC_p = B$$

On a $\text{rg}A = \text{rg}(C_1, \ldots, C_p)$ et $\text{rg}\left(\begin{array}{c|c} A & B \end{array} \right) = \text{rg}(C_1, \ldots, C_p, B)$.
Si Σ est compatible, soit $(x_1, \ldots, x_p)$ un p uplet solution.
Puisque $B = x_1C_1 + \cdots + x_pC_p$, on a $\text{rg}(C_1, \ldots, C_p, B) = \text{rg}(C_1, \ldots, C_p)$ et donc $\text{rg} \left(\begin{array}{c|c} A & B \end{array} \right) = \text{rg}A$
Si $\text{rg} \left(\begin{array}{c|c} A & B \end{array} \right) = \text{rg}(A)$ alors $\text{rg}(C_1, \ldots, C_p, B) = \text{rg}(C_1, \ldots, C_p)$. Or

$$\text{Vect}(C_1, \ldots, C_p) \subset \text{Vect}(C_1, \ldots, C_p, B)$$

donc par inclusion et égalité des dimensions

$$\text{Vect}(C_1, \ldots, C_p) = \text{Vect}(C_1, \ldots, C_p, B)$$

On en déduit $B \in \text{Vect}(C_1, \dots, C_p)$ et donc il existe $x_1, \dots, x_p \in \mathbb{K}$ vérifiant $B = x_1 C_1 + \cdots + x_p C_p$.
□

Corollaire

Tout système de rang $r = n$ est compatible.

Exemple Etudions la compatibilité du système

$$\begin{cases} -2x_1 + x_2 + x_3 = 1 \\ x_1 - 2x_2 + x_3 = 2 \\ x_1 + x_2 - 2x_3 = 3 \end{cases}$$

En échangeant la première et la dernière ligne

$$\text{rg}\left(\begin{array}{ccc|c} -2 & 1 & 1 & 1 \\ 1 & -2 & 1 & 2 \\ 1 & 1 & -2 & 3 \end{array}\right) = \text{rg}\left(\begin{array}{ccc|c} 1 & 1 & -2 & 3 \\ 1 & -2 & 1 & 2 \\ -2 & 1 & 1 & 1 \end{array}\right)$$

$$\text{rg}\left(\begin{array}{ccc|c} 1 & 1 & -2 & 3 \\ 1 & -2 & 1 & 2 \\ -2 & 1 & 1 & 1 \end{array}\right) = \text{rg}\left(\begin{array}{ccc|c} 1 & 1 & -2 & 3 \\ 0 & -3 & 3 & -1 \\ 0 & 3 & -3 & 7 \end{array}\right) = \text{rg}\left(\begin{array}{ccc|c} 1 & 1 & -2 & 3 \\ 0 & -3 & 3 & -1 \\ 0 & 0 & 0 & 6 \end{array}\right)$$

On en déduit $\text{rg}\,(A \mid B) = 3$ et $\text{rg}A = 2$.
Le système étudié est incompatible.

10.6.3 Description de l'ensemble solution

Considérons le système

$$\Sigma : \begin{cases} a_{11}x_1 + \cdots + a_{1p}x_p = b_1 \\ \vdots \\ a_{n1}x_1 + \cdots + a_{np}x_p = b_n \end{cases}$$

d'équation vectorielle $u(x) = b$.

Définition

On appelle système homogène (ou sans second membre) associé au système Σ le système :

$$\Sigma_0 : \begin{cases} a_{11}x_1 + \cdots + a_{1p}x_p = 0 \\ \vdots \\ a_{n1}x_1 + \cdots + a_{np}x_p = 0 \end{cases}$$

Proposition

L'ensemble solution du système Σ_0 est un sous-espace vectoriel de $\mathbb{K}^p$ de dimension $p - r$ avec $r = \text{rg}\Sigma$.

dém. :
L'ensemble solution du système Σ_0 est le noyau de l'application linéaire u et celui-ci est un sous-espace vectoriel de $\mathbb{K}^p$ de dimension $p - r$ avec $r = \text{rg}u = \text{rg}\Sigma$ en vertu de la formule du rang.
□

Théorème

Si le système Σ est compatible alors son ensemble solution $\mathcal{S}_\Sigma$ est un sous-espace affine de $\mathbb{K}^p$ de direction $\mathcal{S}_{\Sigma_0}$.

dém. :
dém. :
Soit $\bar{x}$ solution particulière du système Σ.
$x \in \mathcal{S}_\Sigma \Leftrightarrow u(x) = u(\bar{x}) \Leftrightarrow x - \bar{x} \in \ker u$.
Par suite $\mathcal{S}_\Sigma$ est le sous-espace affine donc $\bar{x} + \mathcal{S}_{\Sigma_0}$.
□

Corollaire

Tout système de rang $r = p$ possède au plus une solution.

10.6.4 Résolution pratique

Considérons le système

$$\Sigma : \begin{cases} a_{11}x_1 + \cdots + a_{1p}x_p = b_1 \\ \vdots \\ a_{n1}x_1 + \cdots + a_{np}x_p = b_n \end{cases}$$

d'équation matricielle $AX = B$.
Par la méthode du pivot de Gauss, on peut en opérant sur les lignes et en permutant éventuellement les colonnes, transformer la matrice A en

$$\left(\begin{array}{ccc|c} p_1 & & \star & \\ & \ddots & & \star \\ 0 & & p_r & \\ \hline & 0 & & 0 \end{array}\right)$$

avec $p_1, ..., p_r \neq 0$.
En suivant ce procédé, mais en opérant sur les équations au lieu des lignes et en permutant les inconnues au lieu des colonnes, on peut transformer Σ en le système équivalent Σ' suivant :

$$\begin{cases} p_1x_1' + \cdots + \star x_r' + \star x_{r+1}' + \ldots + \star x_p' = b_1' \quad (1) \\ \vdots \\ p_rx_r' + \star x_{r+1}' + \cdots + \times x_p' = b_r' \quad (r) \\ 0 = b_{r+1}' \quad (r+1) \\ \vdots \\ 0 = b_n' \quad (n) \end{cases}$$

Définition

Les équations $(1)', \ldots, (r)'$ sont appelées équations principales.
Les équations $(r+1)', \ldots, (n)'$ sont appelées équations de compatibilité.

Si l'une des équations de compatibilité est fausse, le système est incompatible.

Si les équations de compatibilité sont vérifiées, le système Σ est équivalent à :

$$\begin{cases} p_1 x'_1 + \cdots + \star x'_r = b'_1 - (\star x'_{r+1} + \cdots + \star x'_p) \\ \qquad\qquad\qquad \cdots \\ \qquad p_r x'_r = b'_r - (\star x'_{r+1} + \cdots + \star x'_p) \end{cases}$$

Ce système triangulaire se résout en cascade et permet d'exprimer les inconnues $x'_1, \ldots, x'_r$ en fonction des inconnues $x'_{r+1}, \ldots, x'_p$.

Définition

Les inconnues $x'_1, \ldots, x'_r$ sont appelées inconnues principales, c'est en exprimant celles-ci qu'on achève la résolution du système.
Les inconnues $x'_{r+1}, \ldots, x'_p$ sont appelée inconnues paramètres, ce sont elles qui permettent la description de l'ensemble solution $\mathcal{S}_\Sigma$.

Remarque Cette démarche, appelée méthode du pivot, est une résolution par équivalence du système Σ. La méthode par substitution en est une variante dans la mise en forme.

Exemple Etudions

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

On a successivement les systèmes équivalents

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

$$\begin{cases} x + y - z + t = 1 & (1) \\ y + z + t = 1 & (2) \\ t = 3 & (3) - (2) \end{cases}$$

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

$$\begin{cases} x = 2z \\ y = -2 - z \\ t = 3 \end{cases}$$

On peut alors décrire l'ensemble des solutions

$$\mathcal{S} = \{(2z, -2 - z, z, 3)/z \in \mathbb{K}\}$$

Exemple Etudions

$$\begin{cases} x+y+z=2 \\ x+y+2z=3 \\ 2x+2y+z=3 \end{cases}$$

On a successivement les systèmes équivalents

$$\begin{cases} x+y+z=2 & (1) \\ \quad z=1 & (2)-(1) \\ \quad -z=-1 & (3)-2\times(1) \end{cases}$$

$$\begin{cases} x+y+z=2 & (1) \\ \quad z=1 & (2) \\ \quad 0=0 & (3)-(2) \end{cases}$$

$$\begin{cases} x=1-y \\ z=1 \\ 0=0 \end{cases}$$

Et l'ensemble solution est donc

$$\mathcal{S}=\{(1-y,y,1)/y\in\mathbb{K}\}$$

Exemple Etudions

$$\begin{cases} x+y+z=1 \\ x+2y=a \\ x+2z=0 \end{cases}$$

On a successivement les systèmes équivalents

$$\begin{cases} x+y+z=1 & (1) \\ y-z=a-1 & (2)-(1) \\ -y+z=-1 & (3)-(1) \end{cases}$$

$$\begin{cases} x+y+z=1 & (1) \\ y-z=a-1 & (2) \\ \quad 0=a-2 & (3)+(2) \end{cases}$$

Si $a\neq 2$ alors $\mathcal{S}=\emptyset$.
Si $a=2$ alors on obtient

$$\begin{cases} x=-2z \\ y=z+1 \end{cases}$$

puis

$$\mathcal{S}=\{(-2z,1+z,z)/z\in\mathbb{R}\}$$

Exemple Etudions

$$\begin{cases} x+y+mz=m \\ x+my-z=1 \\ x+y-z=1 \end{cases}$$

On a successivement les systèmes équivalents

$$\begin{cases} x+y+mz=m & (1) \\ (m-1)y-(m+1)z=1-m & (2)-(1) \\ -(m+1)z=1-m & (3)-(1) \end{cases}$$

Si $m \neq 1$ et $m \neq -1$ alors

$$\mathcal{S}=\left\{(\frac{2m}{m+1},0,\frac{m-1}{m+1})\right\}$$

Si $m=1$ alors on obtient

$$\begin{cases} x+y+z=1 \\ -2z=0 \\ -2z=0 \end{cases}$$

et on en déduit

$$\mathcal{S}=\{(1-y,y,0)/y\in\mathbb{R}\}$$

Si $m=-1$ alors on obtient

$$\begin{cases} x+y-z=-1 \\ -2y=2 \\ 0=2 \end{cases}$$

donc

$$\mathcal{S}=\emptyset$$

10.6.5 Système de Cramer

Définition

On appelle système de Cramer d'ordre n tout système à n équations, n inconnues et de rang n.

Remarque La matrice associée à un système de Cramer est une matrice carrée inversible.

Théorème

Un système de Cramer possède une et une seule solution.

dém. :
L'équation matricielle associée à un système de Cramer d'ordre n est de la forme $AX=B$ avec $A\in \mathrm{GL}_n(\mathbb{R})$ et $B\in\mathcal{M}_{n,1}(\mathbb{R})$. Cette équation possède une unique solution $X=A^{-1}B$.
□

Remarque La méthode du pivot s'applique aux systèmes de Cramer et permet de déterminer l'unique solution. Cependant, si on sait préalablement que le système étudié est de Cramer, on peut le résoudre plus efficacement :
- en déterminant une solution évidente ;
- ou en déterminant sa solution par des combinaisons judicieuses d'équations.

Exemple Si Σ est un système de Cramer homogène alors $\mathcal{S} = \{(0, \dots, 0)\}$.

Exemple Considérons le système

$$\begin{cases} x + y = 1 \\ x - 2y = 2 \end{cases}$$

Puisque $\operatorname{rg}\begin{pmatrix} 1 & 1 \\ 1 & -2 \end{pmatrix} = 2$, ce système est de Cramer
$(1) - (2)$ donne $y = -1/3$
$2 \times (1) + (2)$ donne $x = 4/3$.
L'unique solution de ce système est donc $(4/3, 1/3)$.

Exemple Soient $\theta, \varphi \in \mathbb{R}$. Résolvons le système

$$\begin{cases} \cos\theta x + \sin\theta y = \cos\varphi \\ -\sin\theta x + \cos\theta y = \sin\varphi \end{cases}$$

La matrice

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

est inversible d'inverse

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

$\cos\theta \times (1) - \sin\theta \times (2)$ donne $x = \cos\theta\cos\varphi - \sin\theta\sin\varphi = \cos(\theta + \varphi)$.
$\sin\theta \times (1) + \cos\theta \times (2)$ donne $y = \sin\theta\cos\varphi + \cos\theta\sin\varphi = \sin(\theta + \varphi)$.
L'unique solution de ce système est donc $(\cos(\theta + \varphi), \sin(\theta + \varphi))$.

Exemple Résolvons le système

$$\begin{cases} x + y + z = a \\ x + jy + j^2 z = b \\ x + j^2 y + jz = c \end{cases}$$

Ce système est de Cramer car

$$\operatorname{rg}\begin{pmatrix} 1 & 1 & 1 \\ 1 & j & j^2 \\ 1 & j^2 & j \end{pmatrix} = \operatorname{rg}\begin{pmatrix} 1 & 1 & 1 \\ 0 & j-1 & j^2-1 \\ 0 & j^2-1 & j-1 \end{pmatrix} = \operatorname{rg}\begin{pmatrix} 1 & 1 & 1 \\ 0 & j-1 & j^2-1 \\ 0 & 0 & -(j-1)^2 \end{pmatrix} = 3$$

$(1) + (2) + (3)$ donne $3x = a + b + c$.
$(1) + j^2 \times (2) + j \times (3)$ donne $3y = a + bj^2 + cj$.
$(1) + j^2 \times (2) + j \times (3)$ donne $3z = a + bj + cj^2$
L'unique solution du système est donc

$$\left(\frac{a+b+c}{3}, \frac{a+bj^2+cj}{3}, \frac{a+bj+cj^2}{3}\right)$$

Chapitre 11

Déterminants

Le déterminant est un outil qui caractérise, par un calcul, l'inversibilité d'une matrice carrée.
$\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$.
n et p désignent des naturels non nuls.

11.1 Applications multilinéaires

Soient E et F deux $\mathbb{K}$-espaces vectoriels.

11.1.1 Définition

Définition

On appelle application n-linéaire de E vers F toute application

$$\varphi : \begin{cases} E^n \to F \\ (x_1, \dots, x_n) \mapsto \varphi(x_1, \dots, x_n) \end{cases}$$

linéaire par rapport à chacune de ses variables. Ceci signifie que pour toute famille $(x_1, \dots, \hat{x}_i, \dots, x_n) \in E^{n-1}$, il y a linéarité de l'application partielle $\varphi_i : E \to F$ définie par

$$\varphi_i : x_i \mapsto \varphi(x_1, \dots, x_i, \dots, x_n)$$

Quand $n = 2$, on parle d'application bilinéaire.
Quand $F = \mathbb{K}$, on parle de forme n linéaire.

Exemple Si E désigne l'ensemble des vecteurs du plan (ou de l'espace géométrique), l'application produit scalaire $(\vec{u}, \vec{v}) \mapsto \vec{u} \cdot \vec{v}$ est une forme bilinéaire.
En effet pour tout $\vec{u}, \vec{v}, \vec{w} \in E$ et tout $\lambda, \mu \in \mathbb{R}$, on a
$(\lambda\vec{u} + \mu\vec{v}) \cdot \vec{w} = \lambda\vec{u} \cdot \vec{w} + \mu\vec{v} \cdot \vec{w}$ (linéarité en la première variable) et
$\vec{u} \cdot (\lambda\vec{v} + \mu\vec{w}) = \lambda\vec{u} \cdot \vec{v} + \mu\vec{u} \cdot \vec{w}$ (linéarité en la seconde variable).

Exemple Si E désigne l'ensemble des vecteurs de l'espace géométrique, l'application produit vectoriel $(\vec{u}, \vec{v}) \mapsto \vec{u} \wedge \vec{v}$ est bilinéaire.

Exemple Si E désigne l'ensemble des vecteurs de l'espace géométrique, l'application produit mixte $(\vec{u}, \vec{v}, \vec{w}) \mapsto [\vec{u}, \vec{v}, \vec{w}]$ est une forme trilinéaire.

Exemple Soient $E = \mathbb{K}, F = \mathbb{K}$ et $n \in \mathbb{N}^\star$.
L'application $\varphi : \mathbb{K} \times \cdots \times \mathbb{K} \to \mathbb{K}$ définie par $\varphi(x_1, \ldots, x_n) = x_1 \times \cdots \times x_n$ est une forme n linéaire.
En effet, pour chaque $i \in \{1, \ldots, n\}$, l'application partielle $x_i \mapsto \varphi(x_1, \ldots, x_i, \ldots, x_n)$ est linéaire car de la forme $x_i \mapsto a x_i$.

Remarque Pour $n = 1$, une application 1-linéaire de E vers F n'est autre qu'application linéaire de E vers F.
Pour $n \geqslant 2$, les applications n-linéaires de E vers F ne correspondent pas à des applications linéaires.

Proposition

Soit φ une application n-linéaire de E vers F et $(x_1, \ldots, x_n) \in E^n$.
Si l'un des $x_1, \ldots, x_n$ est nul alors $\varphi(x_1, \ldots, x_n) = 0_F$.

dém. :
Soit $i \in \{1, \ldots, n\}$ tel que $x_i = 0_E$.
Puisque l'application $x_i \mapsto \varphi(x_1, \ldots, x_i, \ldots, x_n)$ est linéaire, elle s'annule sur le vecteur nul et donc $\varphi(x_1, \ldots, 0_E, \ldots, x_n) = 0_F$
□

11.1.2 Applications multilinéaires symétriques et antisymétriques

Définition

On dit qu'une application n-linéaire φ de E vers F est symétrique si

$$\forall \sigma \in \mathfrak{S}_n, \forall (x_1, \ldots, x_n) \in E^n, \varphi\left(x_{\sigma(1)}, \ldots, x_{\sigma(n)}\right) = \varphi(x_1, \ldots, x_n)$$

On note $S_n(E, F)$ l'ensemble de ces applications.

Définition

On dit qu'une application n-linéaire φ de E vers F est antisymétrique si

$$\forall \sigma \in \mathfrak{S}_n, \forall (x_1, \ldots, x_n) \in E^n, \varphi\left(x_{\sigma(1)}, \ldots, x_{\sigma(n)}\right) = \varepsilon(\sigma)\varphi(x_1, \ldots, x_n)$$

On note $A_n(E, F)$ l'ensemble de ces applications.

Exemple Soient $E = \mathbb{K}, F = \mathbb{K}$ et $n \in \mathbb{N}^\star$.
L'application $\varphi : \mathbb{K}^n \to \mathbb{K}$ définie par $\varphi(x_1, \ldots, x_n) = x_1 \ldots x_n$ est une forme n linéaire symétrique.
En effet par commutativité de la multiplication, on vérifie $x_{\sigma(1)} \ldots x_{\sigma(n)} = x_1 \ldots x_n$ pour tout $\sigma \in \mathfrak{S}_n$.

Exemple Le produit scalaire du plan ou de l'espace est une forme bilinéaire symétrique. En effet, pour tout $\sigma \in \mathfrak{S}_2 = \{\mathrm{Id}, \tau\}$ avec $\tau = \begin{pmatrix} 1 & 2 \end{pmatrix}$, on a
Si $\sigma = \mathrm{Id}$, $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$,
Si $\sigma = \begin{pmatrix} 1 & 2 \end{pmatrix}$, $\vec{v} \cdot \vec{u} = \vec{u} \cdot \vec{v}$.

Exemple Le produit vectoriel dans l'espace est une application bilinéaire antisymétrique. En effet, pour tout $\sigma \in \mathfrak{S}_2 = \{\text{Id}, \tau\}$ avec $\tau = \begin{pmatrix} 1 & 2 \end{pmatrix}$, on a
Si $\sigma = \text{Id}$, $\vec{u} \wedge \vec{v} = \varepsilon(\sigma)\vec{u} \wedge \vec{v}$,
Si $\sigma = \begin{pmatrix} 1 & 2 \end{pmatrix}$, $\vec{v} \wedge \vec{u} = -\vec{u} \wedge \vec{v} = \varepsilon(\sigma)\vec{u} \wedge \vec{v}$.

Proposition

Soit φ une application n-linéaire de E vers F.
On a équivalence entre :
(i) φ est symétrique (resp. antisymétrique) ;
(ii) $\forall (x_1, \ldots, x_n) \in E^n$, $\forall \tau$ transposition de $\{1, \ldots, n\}$,

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

(respectivement $\varphi\left(x_{\tau(1)}, \ldots, x_{\tau(n)}\right) = -\varphi(x_1, \ldots, x_n)$)

dém. :
(i) $\Rightarrow$ (ii) est immédiat car une transposition de $\{1, \ldots, n\}$ est en particulier une permutation de $\{1, \ldots, n\}$
(ii) $\Rightarrow$ (i) Supposons φ symétrique.
Soit $\sigma \in \mathfrak{S}_n$. Puisque toute permutation peut s'écrire comme un produit de transpositions, on peut écrire $\sigma = \tau_1 \circ \ldots \circ \tau_m$ avec τ_k transposition de $\{1, \ldots, n\}$. On a alors pour $(x_1, \ldots, x_n) \in E^n$,

$$\varphi\left(x_{\sigma(1)}, \ldots, x_{\sigma(n)}\right) = \varphi\left(x_{\tau_1 \circ \cdots \circ \tau_m(1)}, \ldots, x_{\tau_1 \circ \cdots \circ \tau_m(n)}\right) = \varphi\left(x_{\tau_1 \circ \cdots \circ \tau_{m-1}(1)}, \ldots, x_{\tau_1 \circ \cdots \circ \tau_{m-1}(n)}\right)$$

Puis en répétant le procédé

$$\varphi\left(x_{\sigma(1)}, \ldots, x_{\sigma(n)}\right) = \varphi\left(x_{\tau_1 \circ \cdots \circ \tau_m(1)}, \ldots, x_{\tau_1 \circ \cdots \circ \tau_m(n)}\right) = \varphi(x_1, \ldots, x_n)$$

L'étude de φ est antisymétrique est identique sachant $\varepsilon(\sigma) = \prod_{k=1}^{m} \varepsilon(\tau_k) = (-1)^m$

□

11.1.3 Application multilinéaire alternée

Définition

On dit qu'une application n-linéaire φ de E vers F est alternée si celle-ci s'annule sur toute famille contenant deux fois le même vecteur i.e.

$$\forall (x_1, \ldots, x_n) \in E^n, \exists 1 \leqslant i \neq j \leqslant n, x_i = x_j \Rightarrow \varphi(x_1, \ldots, x_n) = 0_F$$

On note $\Lambda_n(E, F)$ l'ensemble de ces applications.
Si $F = \mathbb{K}$, on notera $\Lambda_n^\star(E)$ au lieu de $\Lambda_n(E, \mathbb{K})$.

Exemple Le produit vectoriel est une application bilinéaire alternée.
En effet, si $\vec{u} = \vec{v}$ alors $\vec{u} \wedge \vec{v} = \vec{u} \wedge \vec{u} = \vec{0}$.

Exemple Le produit mixte est une forme trilinéaire alternée.
En effet si $\vec{u} = \vec{v}$, $\vec{v} = \vec{w}$ ou $\vec{u} = \vec{w}$ alors $[\vec{u}, \vec{v}, \vec{w}] = 0$.

Proposition

> Soit φ une application n-linéaire alternée de E vers F.
> Si $(x_1,\dots,x_n)$ est une famille liée de vecteurs de E alors $\varphi(x_1,\dots,x_n)=0_F$.

dém. :
Si la famille $(x_1,\dots,x_n)$ est alternée alors l'un de ses vecteurs est combinaison linéaire des autres. En notant i l'indice correspondant à celui-ci, on peut écrire $x_i=\sum_{j\neq i}\lambda_j x_j$.
Puisque l'application φ est linéaire en sa $i$ème variable, on peut écrire

$$\varphi(x_1,\dots,x_i,\dots,x_n)=\sum_{j\neq i}\lambda_j\varphi(x_1,\dots,x_{i-1},x_j,x_{i+1},\dots,x_n)$$

Or la famille $(x_1,\dots,x_{i-1},x_j,x_{j+1},\dots,x_n)$ comporte deux fois le vecteur x_j donc puisque φ est alternée

$$\varphi(x_1,\dots,x_{i-1},x_j,x_{j+1},\dots,x_n)=0_F$$

puisque φ est alternée et alors

$$\varphi(x_1,\dots,x_i,\dots,x_n)=0_F$$

□

Théorème

> Soit φ une application n-linéaire de E vers F.
> On a équivalence entre :
> (i) φ est antisymétrique ;
> (ii) φ est alternée.

dém. :
(i) $\Rightarrow$ (ii) Supposons φ antisymétrique.
Considérons $(x_1,\dots,x_n)\in E^n$ tel qu'il existe $i\neq j\in\{1,\dots,n\}$ vérifiant $x_i=x_j$. Pour la permutation $\tau=(i,j)$, on a $(x_{\tau(1)},\dots,x_{\tau(n)})=(x_1,\dots,x_n)$ car $x_i=x_j$. Or $\varphi\left(x_{\tau(1)},\dots,x_{\tau(n)}\right)=-\varphi(x_1,\dots,x_n)$ car φ est antisymétrique. On en déduit $\varphi(x_1,\dots,x_n)=0_F$.
Ainsi l'application n-linéaire φ est alternée.
(ii) $\Rightarrow$ (i) Supposons φ alternée.
Soit $\tau=(i,j)$ (avec $i<j$) une transposition de $\{1,\dots,n\}$
Puisque la famille

$$(x_1,\dots,x_{i-1},(x_i+x_j),x_{i+1},\dots,x_{j-1},(x_i+x_j),x_{j+1},\dots,x_n)$$

comporte deux fois le même vecteur on a

$$\varphi\left(x_1,\dots,(x_i+x_j),\dots,(x_i+x_j),\dots,x_n\right)=0_F$$

En développant grâce aux linéarités en la $i$ème et en la $j$ème variable, on obtient

$$\begin{aligned}&\varphi\left(x_1,\dots,(x_i+x_j),\dots,(x_i+x_j),\dots,x_n\right)\\&=\varphi\left(x_1,\dots,x_i,\dots,x_j,\dots,x_n\right)+\varphi\left(x_1,\dots,x_j,\dots,x_i,\dots,x_n\right)\end{aligned}$$

car

$$\varphi\left(x_1,\dots,x_i,\dots,x_i,\dots,x_n\right)=\varphi\left(x_1,\dots,x_j,\dots,x_j,\dots,x_n\right)=0_F$$

(deux fois le même vecteur).
On obtient ainsi $\varphi\left(x_{\tau(1)},\dots,x_{\tau(n)}\right)=-\varphi(x_1,\dots,x_n)$ et donc φ est antisymétrique.
□

Exemple Le produit mixte est une application linéaire antisymétrique car alternée.

11.2 Déterminants

Soit E un $\mathbb{K}$-espace vectoriel de dimension n.

11.2.1 Déterminant d'une famille de vecteurs

11.2.1.1 Définition

Définition

Soient $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E, $\mathcal{F} = (x_1, \ldots, x_n)$ une famille de vecteurs de E et

$$A = (a_{i,j}) = \mathrm{Mat}_{\mathcal{B}}(x_1, \ldots, x_n) \in \mathcal{M}_n(\mathbb{K})$$

On appelle déterminant de la famille $\mathcal{F}$ dans la base $\mathcal{B}$ le scalaire

$$\det_{\mathcal{B}} \mathcal{F} = \det_{\mathcal{B}}(x_1, \ldots, x_n) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{\sigma(i),i}$$

Remarque Si la formule est complexe, elle est néanmoins parfois utile et il est recommandé de la connaître.

Proposition

Si $\mathcal{B} = (e_1, \ldots, e_n)$ est une base de E alors $\det_{\mathcal{B}} \mathcal{B} = 1$.

dém. :
$A = \mathrm{Mat}_{\mathcal{B}}\mathcal{B} = I_n = (\delta_{i,j})$ donc

$$\det_{\mathcal{B}}(e_1, ..., e_n) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} \delta_{\sigma(i),i}$$

Or

$$\prod_{i=1}^{n} \delta_{\sigma(i),i} = \begin{cases} 1 & \text{si } \sigma = \mathrm{Id} \\ 0 & \text{sinon} \end{cases}$$

donc $\det_{\mathcal{B}}(e_1, ..., e_n) = \varepsilon(\mathrm{Id}) \times 1 = 1$.
□

11.2.1.2 Propriété fondatrice

Théorème

Soit E un $\mathbb{K}$-espace vectoriel de dimension n muni d'une base $\mathcal{B} = (e_1, \ldots, e_n)$.
L'application $\det_{\mathcal{B}} : E^n \to \mathbb{K}$ est une forme n-linéaire alternée non nulle.
De plus, l'ensemble $\Lambda_n^\star(E)$ des formes n linéaires alternées sur E est une droite vectorielle.

dém. :
Pour alléger ce qui suit nous noterons $\Delta = \det_{\mathcal{B}}$.

On peut déjà affirmer que l'application Δ est non nulle puisque $\Delta(\mathcal{B}) = 1$.
Montrons que l'application $\Delta : E^n \to \mathbb{K}$ est n-linéaire.
Soit $1 \leqslant j \leqslant n$, justifions la linéarité en la $j$ème variable.
Soient $\alpha, \beta \in \mathbb{K}$, $(x_1, \dots, \hat{x}_j, \dots, x_n) \in E^{n-1}$ et $x_j, y_j \in E$ vecteurs de matrices composantes dans $\mathcal{B}$

$$\begin{pmatrix} a_{1,1} \\ \vdots \\ a_{n,1} \end{pmatrix}, \dots, \widehat{\begin{pmatrix} a_{1,j} \\ \vdots \\ a_{n,j} \end{pmatrix}}, \dots, \begin{pmatrix} a_{1,n} \\ \vdots \\ a_{n,n} \end{pmatrix} \text{ et } \begin{pmatrix} a_{1,j} \\ \vdots \\ a_{n,j} \end{pmatrix}, \begin{pmatrix} b_{1,j} \\ \vdots \\ b_{n,j} \end{pmatrix}$$

On a alors

$$\Delta(x_1, \dots, \alpha x_j + \beta y_j, \dots, x_n) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1),1} \dots (\alpha a_{\sigma(j),j} + \beta b_{\sigma(j),j}) \dots a_{\sigma(n),n}$$

En développant

$$\begin{aligned} \Delta(x_1, \dots, \alpha x_j + \beta y_j, \dots, x_n) = \alpha & \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1),1} \dots a_{\sigma(j),j} \dots a_{\sigma(n),n} \\ + \beta & \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1),1} \dots b_{\sigma(j),j} \dots a_{\sigma(n),n} \end{aligned}$$

et donc

$$\Delta(x_1, \dots, \alpha x_j + \beta y_j, \dots, x_n) = \alpha \Delta(x_1, \dots, x_j, \dots, x_n) + \beta \Delta(x_1, \dots, y_j, \dots, x_n)$$

Ainsi Δ est une forme n-linéaire.
Montrons que Δ est alternée
Soit $(x_1, ..., x_n) \in E^n$ tel que $x_i = x_j$ pour un certain couple (i, j) avec $i < j$.
Notons $A = (a_{i,j}) = \mathrm{Mat}_{\mathcal{B}}(x_1, \dots, x_n)$.
Considérons la transposition $\tau = (i, j)$ et l'application $\varphi : \mathfrak{A}_n \to \mathfrak{S}_n \backslash \mathfrak{A}_n$ définie par $\varphi(\sigma) = \sigma \circ \tau$.
L'application φ est bien définie et est bijective.

$$\Delta(x_1, \dots, x_n) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1),1} \dots a_{\sigma(n),n}$$

En décomposant la somme

$$\Delta(x_1, \dots, x_n) = \sum_{\sigma \in \mathfrak{A}_n} \varepsilon(\sigma) a_{\sigma(1),1} \dots a_{\sigma(n),n} + \sum_{\sigma \in \mathfrak{S}_n - \mathfrak{A}_n} \varepsilon(\sigma) a_{\sigma(1),1} \dots a_{\sigma(n),n}$$

Par la bijectivité de φ

$$\Delta(x_1, \dots, x_n) = \sum_{\sigma \in \mathfrak{A}_n} \varepsilon(\sigma) a_{\sigma(1),1} \dots a_{\sigma(n),n} + \sum_{\sigma \in \mathfrak{A}_n} \varepsilon(\sigma \circ \tau) a_{(\sigma \circ \tau)(1),1} \dots a_{(\sigma \circ \tau)(n),n}$$

Or $x_i = x_j$ donc $a_{(\sigma \circ \tau)(i),k} = a_{\sigma(i),k}$ et $a_{(\sigma \circ \tau)(j),k} = a_{\sigma(j),k}$ donc

$$\Delta(x_1, \dots, x_n) = \sum_{\sigma \in \mathfrak{A}_n} \varepsilon(\sigma) a_{\sigma(1),1} \dots a_{\sigma(n),n} + \sum_{\sigma \in \mathfrak{A}_n} \varepsilon(\sigma \circ \tau) a_{\sigma(1),1} \dots a_{\sigma(n),n}$$

Enfin $\varepsilon(\sigma \circ \tau) = \varepsilon(\sigma) \varepsilon(\tau) = -\varepsilon(\sigma)$
donc

$\Delta(x_1,\dots,x_n) = 0$
Ainsi Δ est alternée.
Montrons enfin que $\Lambda_n^\star(E) = \{\lambda.\Delta/\lambda \in \mathbb{K}\}$.
Soient $\varphi \in \Lambda_n^\star(E)$ et $(x_1,\dots,x_n) \in E^n$. Notons $A = (a_{i,j}) = \text{Mat}_\mathcal{B}(x_1,\dots,x_n)$.
Par multilinéarité

$$\varphi(x_1,\dots,x_n) = \varphi(\sum_{i_1=1}^n a_{i_1,1}e_{i_1},\dots,\sum_{i_n=1}^n a_{i_n,n}e_{i_n}) = \sum_{i_1=1}^n \dots \sum_{i_n=1}^n a_{i_1,1}\dots a_{i_n,n}\varphi(e_{i_1},\dots,e_{i_n})$$

Or φ est alternée, donc $\varphi(e_{i_1},\dots,e_{i_n})$ est nul dès que les $i_1,\dots,i_n$ ne sont pas deux à deux distincts. Après simplification des termes correspondants, il ne reste plus que les termes pour lesquels $\{i_1,\dots,i_n\} = \{1,\dots,n\}$ c'est-à-dire ceux qui correspondent aux applications $\sigma : j \mapsto i_j$ bijective. Ainsi

$$\varphi(x_1,\dots,x_n) = \sum_{\sigma\in\mathfrak{S}_n} a_{\sigma(1),1}\dots a_{\sigma(n),n}\varphi(e_{\sigma(1)},\dots,e_{\sigma(n)})$$

Or φ est alternée donc antisymétrique et par conséquent

$$\varphi(x_1,\dots,x_n) = \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma)a_{\sigma(1),1}\dots a_{\sigma(n),n}\varphi(e_1,\dots,e_n)$$

Finalement $\varphi = \lambda.\Delta$ avec $\lambda = \varphi(e_1,\dots,e_n)$.
Inversement, il est immédiate que pour tout $\lambda \in \mathbb{K}$, on a $\lambda.\Delta \in \Lambda_n^\star(E)$.
On peut donc conclure que $\Lambda_n^\star(E)$ est une droite vectorielle dont Δ est un vecteur directeur.
□

Corollaire

L'application $\det_\mathcal{B}$ est linéaire en chacune de ses variables,
L'application $\det_\mathcal{B}$ s'annule sur les familles liées,
L'application $\det_\mathcal{B}$ est antisymétrique
L'application $\det_\mathcal{B}$ est vecteur directeur de $\Lambda_n^\star(E)$ i.e. :

$$\forall\varphi \in \Lambda_n^\star(E), \exists!\lambda \in \mathbb{K}, \varphi = \lambda.\det_\mathcal{B}$$

11.2.1.3 Changement de base

Le déterminant d'une famille de vecteurs dépend de la base choisie pour représenter celle-ci.

Théorème

Soient $\mathcal{B} = (e_1,\dots,e_n)$ et $\mathcal{B}' = (e'_1,\dots,e'_n)$ deux bases de E.

$$\forall(x_1,\dots,x_n) \in E^n, \det_{\mathcal{B}'}(x_1,\dots,x_n) = \det_{\mathcal{B}'}\mathcal{B}.\det_\mathcal{B}(x_1,\dots,x_n)$$

dém. :
L'application $\det_{\mathcal{B}'}$ est une forme n-linéaire alternée sur E, or l'ensemble de ces formes multilinéaires est une droite vectorielle engendrée par $\det_\mathcal{B}$ donc il existe $\lambda \in \mathbb{K}$ vérifiant $\det_{\mathcal{B}'} = \lambda\det_\mathcal{B}$ i.e.

$$\forall(x_1,\dots,x_n) \in E^n, \det_{\mathcal{B}'}(x_1,\dots,x_n) = \lambda.\det_\mathcal{B}(x_1,\dots,x_n)$$

Pour $(x_1,\dots,x_n)=(e_1,\dots,e_n)$, on obtient $\lambda=\det_{\mathcal{B}'}(e_1,\dots,e_n)=\det_{\mathcal{B}'}\mathcal{B}$ car $\det_{\mathcal{B}}\mathcal{B}=1$.

□

Corollaire

Si $\mathcal{B}$ et $\mathcal{B}'$ sont deux bases de E alors $\det_{\mathcal{B}'}\mathcal{B}\times\det_{\mathcal{B}}\mathcal{B}'=1$.

dém. :
Il suffit d'appliquer la relation qui précède à $(x_1,\dots,x_n)=(e'_1,\dots,e'_n)$ sachant $\det_{\mathcal{B}'}\mathcal{B}'=1$.

□

11.2.1.4 Déterminant et caractérisation des bases

Théorème

Soit E un $\mathbb{K}$-espace vectoriel muni d'une base $\mathcal{B}=(e_1,\dots,e_n)$.
Soit $\mathcal{B}'=(e'_1,\dots,e'_n)$ une famille de vecteurs de E.
On a équivalence entre :
(i) $\mathcal{B}'$ est une base de E ;
(ii) $\det_{\mathcal{B}}\mathcal{B}'\neq 0$.

dém. :
(i) $\Rightarrow$ (ii) Car si $\mathcal{B}'$ est une base de E alors $\det_{\mathcal{B}'}\mathcal{B}\times\det_{\mathcal{B}}\mathcal{B}'=1$.
(ii) $\Rightarrow$ (i) Car si $\mathcal{B}'$ n'est pas une base de E, la famille $\mathcal{B}'$ est liée et donc $\det_{\mathcal{B}}\mathcal{B}'=0$ car l'application multilinéaire $\det_{\mathcal{B}}$ est alternée.

□

11.2.2 Déterminant d'un endomorphisme

11.2.2.1 Définition

Soient u un endomorphisme de E et $\mathcal{B}=(e_1,\dots,e_n)$ une base de E. Posons

$$\delta=\det_{\mathcal{B}}(u(e_1),\dots,u(e_n))$$

Théorème

$$\forall\varphi\in\Lambda_n^\star(E),\forall(x_1,\dots,x_n)\in E^n,\varphi(u(x_1),\dots,u(x_n))=\delta\varphi(x_1,\dots,x_n)$$

dém. :
Soit $\varphi\in\Lambda_n^\star(E)$. Puisque l'espace $\Lambda_n^\star(E)$ est une droite vectorielle dirigée par l'application $\det_{\mathcal{B}}$, il existe un scalaire $\alpha\in\mathbb{K}$ tel que $\varphi=\alpha.\det_{\mathcal{B}}$ i.e. :

$$\forall(x_1,\dots,x_n)\in E^n,\varphi(x_1,\dots,x_n)=\alpha.\det_{\mathcal{B}}(x_1,\dots,x_n)\quad(1)$$

Soit $\psi:E^n\to\mathbb{K}$ définie par $\psi(x_1,\dots,x_n)=\varphi(u(x_1),\dots,u(x_n))$.
On vérifie aisément que $\psi\in\Lambda_n^\star(E)$ et donc il existe un scalaire $\beta\in\mathbb{K}$ tel que $\psi=\beta.\det_{\mathcal{B}}$ i.e. :

$$\forall(x_1,\dots,x_n)\in E^n,\psi(x_1,\dots,x_n)=\beta.\det_{\mathcal{B}}(x_1,\dots,x_n)\quad(2)$$

Or par (1) on a aussi

$$\psi(x_1,\dots,x_n)=\varphi(u(x_1),\dots,u(x_n))=\alpha.\det_{\mathcal{B}}(u(x_1),\dots,u(x_n))\quad(3)$$

En prenant $(x_1, \dots, x_n) = (e_1, \dots, e_n)$ dans (2) et (3), on obtient :

$$\varphi(u(e_1), \dots, u(e_n)) = \beta. \det_{\mathcal{B}}(e_1, \dots, e_n) = \beta$$

et

$$\varphi(u(e_1), \dots, u(e_n)) = \alpha. \det_{\mathcal{B}}(u(e_1), \dots, u(e_n)) = \alpha\delta$$

On en déduit $\beta = \alpha\delta$ puis

$$\forall (x_1, \dots, x_n) \in E^n, \varphi(u(x_1), \dots, u(x_n)) = \delta\alpha \det_{\mathcal{B}}(x_1, \dots, x_n) = \delta\varphi(x_1, \dots, x_n)$$

□

Corollaire

La quantité $\det_{\mathcal{B}}(u(e_1), \dots, u(e_n))$ ne dépend pas du choix de la base $\mathcal{B}$.

dém. :
Soit $\mathcal{B}' = (e'_1, \dots, e'_n)$ une base de E.
En appliquant le résultat précédent à $\varphi = \det_{\mathcal{B}'}$ et $(x_1, \dots, x_n) = (e'_1, \dots, e'_n)$, on obtient
$\det_{\mathcal{B}'}(u(e'_1), \dots, u(e'_n)) = \delta \det_{\mathcal{B}'}(e'_1, \dots, e'_n) = \delta$
□

Définition

On appelle déterminant de l'endomorphisme u le scalaire $\det u = \det_{\mathcal{B}}(u(e_1), \dots, u(e_n))$
où $\mathcal{B} = (e_1, \dots, e_n)$ désigne une base quelconque de E.

11.2.2.2 Propriétés

Proposition

$\det(\mathrm{Id}_E) = 1$.

dém. :
Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E.

$$\det(\mathrm{Id}_E) = \det_{\mathcal{B}}(\mathrm{Id}_E(e_1), \dots, \mathrm{Id}_E(e_n)) = \det_{\mathcal{B}}(e_1, \dots, e_n) = 1$$

□

Proposition

$$\forall \lambda \in \mathbb{K}, \forall u \in \mathcal{L}(E), \det(\lambda.u) = \lambda^n \det u$$

dém. :
Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E. On a

$$\det(\lambda.u) = \det_{\mathcal{B}}(\lambda.u(e_1), \dots, \lambda.u(e_n))$$

Or l'application $\det_{\mathcal{B}}$ est n-linéaire alternée donc en exploitant la linéarité en chacune de ses variables
$\det(\lambda.u) = \lambda^n \det_{\mathcal{B}}(u(e_1), \dots, u(e_n))$
□

Attention : En général

$$\det(\lambda u + \mu v) \neq \lambda \det u + \mu \det v$$

L'application déterminant n'est pas linéaire…

Théorème

$$\forall u, v \in \mathcal{L}(E), \det(v \circ u) = \det v \times \det u$$

dém. :
Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E.
Considérons $\varphi : E^n \to \mathbb{K}$ l'application définie par

$$\varphi(x_1, \ldots, x_n) = \det_{\mathcal{B}}(v(x_1), \ldots, v(x_n))$$

L'application φ est une forme n-linéaire alternée sur E donc on peut écrire

$$\forall (x_1, \ldots, x_n) \in E^n, \varphi(u(x_1), \ldots, u(x_n)) = \det u \times \varphi(x_1, \ldots, x_n)$$

En appliquant cette relation à $(x_1, \ldots, x_n) = (e_1, \ldots, e_n)$ on obtient

$$\det_{\mathcal{B}} (v(u(e_1)), \ldots, v(u(e_n))) = \det u \times \det_{\mathcal{B}}(v(e_1), \ldots, v(e_n))$$

ce qui signifie $\det(v \circ u) = \det u \times \det v$
□

Corollaire

$$\forall u \in \mathcal{L}(E), \forall m \in \mathbb{N}, \det(u^m) = (\det u)^m$$

dém. :
Par simple récurrence.
□

11.2.2.3 Déterminant et automorphisme

Théorème

Soit $u \in \mathcal{L}(E)$. On a équivalence entre :
(i) u est un automorphisme de E ;
(ii) $\det u \neq 0$.
De plus, si tel est le cas,

$$\det u^{-1} = 1/\det u$$

dém. :
(i) ⇒ (ii) Supposons $u \in \mathrm{GL}(E)$. On peut introduire l'automorphisme inverser u^{-1} et on a $u \circ u^{-1} = \mathrm{Id}_E$.
On a alors $\det(u \circ u^{-1}) = \det \mathrm{Id}_E$ puis $\det u \times \det(u^{-1}) = 1$.
On en déduit $\det u \neq 0$ et $\det u^{-1} = 1/\det u$

(ii) $\Rightarrow$ (i) Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E.
Si $\det u \neq 0$ alors $\det_{\mathcal{B}}(u(e_1), \dots, u(e_n)) \neq 0$ et donc la famille $(u(e_1), \dots, u(e_n))$ est une base de E.
Puisque l'endomorphisme u transforme une base en une base, c'est un automorphisme de E.
□

Corollaire

L'application $\det : \mathrm{GL}(E) \to \mathbb{K}^\star$ est un morphisme de groupes.

dém. :
$(\mathrm{GL}(E), \circ)$ et $(\mathbb{K}^\star, \times)$ sont deux groupes et $\det(u \circ v) = \det u \times \det v$ pour tout $u, v \in \mathrm{GL}(E)$.
□

Définition

Le noyau du morphisme $\det : \mathrm{GL}(E) \to \mathbb{K}^\star$ est noté $\mathrm{SL}(E)$, on l'appelle groupe spécial linéaire de E.
$\mathrm{SL}(E) = \{u \in \mathcal{L}(E) / \det u = 1\}$.

11.2.3 Déterminant d'une matrice carrée

11.2.3.1 Définition

Définition

On appelle déterminant d'une matrice carrée $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$ le scalaire :

$$\det A = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{\sigma(i),i}$$

On note encore

$$\det A = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix}_{[n]}$$

Attention : On ne calcule le déterminant que de matrices carrées !

Remarque La formule est compliquée mais à connaître.
Par celle-ci, on montre de façon immédiate :
- $\det \bar{A} = \overline{\det A}$ avec $\bar{A} = (\overline{a_{i,j}})_{1 \leqslant i,j \leqslant n}$;
- $\det A \in \mathbb{Z}$ quand $a_{i,j} \in \mathbb{Z}$ pour tout $i, j \in \{1, \dots, n\}$ etc.

Proposition

Soient $\mathcal{B} = (e_1, \dots, e_n)$ une base de E et $(x_1, \dots, x_n)$ une famille de vecteurs de E.
Si $A = \mathrm{Mat}_{\mathcal{B}}(x_1, \dots, x_n)$ alors $\det_{\mathcal{B}}(x_1, \dots, x_n) = \det A$.

dém. :
Les formules définissant $\det_{\mathcal{B}}(x_1, \dots, x_n)$ et $\det A$ sont identiques.
□

Proposition

Soient $u \in \mathcal{L}(E)$ et $\mathcal{B} = (e_1, \dots, e_n)$ une base de E.
Si $A = \mathrm{Mat}_{\mathcal{B}} u$ alors $\det u = \det A$.

dém. :
$A = \text{Mat}_{\mathcal{B}} u = \text{Mat}_{\mathcal{B}}(u(e_1), \dots, u(e_n))$ donc $\det A = \det_{\mathcal{B}}(u(e_1), \dots, u(e_n)) = \det u$.
□

Remarque Tout problème de calcul de déterminant d'un endomorphisme, ou d'une famille de vecteurs dans une base, se ramène à un problème de calcul de déterminant d'une matrice carrée.

11.2.3.2 Propriétés

Proposition

$\det I_n = 1$.

dém. :
$\det I_n = \det \text{Id}_E = 1$
□

Proposition

$$\forall \lambda \in \mathbb{K}, \forall A \in \mathcal{M}_n(\mathbb{K}), \det(\lambda.A) = \lambda^n \det A$$

dém. :
Soient $\mathcal{B} = (e_1, \dots, e_n)$ une base de E et $u \in \mathcal{L}(E)$ de matrice A dans $\mathcal{B}$. L'endomorphisme λu est alors représenté par la matrice λA et donc
$\det(\lambda A) = \det(\lambda u) = \lambda^n \det u = \lambda^n \det A$.
□

Attention : En général

$$\det(\lambda.A + \mu.B) \neq \lambda. \det A + \mu. \det B$$

L'application déterminant n'est pas linéaire.

Théorème

$$\forall A, B \in \mathcal{M}_n(\mathbb{K}), \det(AB) = \det A \times \det B$$

dém. :
Soient $\mathcal{B} = (e_1, \dots, e_n)$ une base de E et $u, v \in \mathcal{L}(E)$ de matrices A, B dans $\mathcal{B}$. L'endomorphisme $u \circ v$ est alors représenté par la matrice AB et donc
$\det(AB) = \det(u \circ v) = \det u \times \det v = \det A \times \det B$.
□

Corollaire

$$\forall A \in \mathcal{M}_n(\mathbb{K}), \forall m \in \mathbb{N}, \det(A^m) = (\det A)^m$$

dém. :
Par récurrence simple.
□

Théorème

Soit $A \in \mathcal{M}_n(\mathbb{K})$.
On a équivalence entre :
(i) A est inversible ;
(ii) $\det A \neq 0$.
De plus, si tel est le cas,

$$\det A^{-1} = 1/\det A$$

dém. :
Soient $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E et $u \in \mathcal{L}(E)$ de matrice A dans $\mathcal{B}$. Le résultat énoncé s'obtient en exploitant le résultat analogue relatif aux endomorphismes sachant que la matrice A est inversible si, et seulement si, u est un automorphisme de E et qu'alors A^{-1} est la matrice de u^{-1}.
□

Corollaire

L'application $\det$: $\mathrm{GL}_n(\mathbb{K}) \to \mathbb{K}^\star$ est un morphisme de groupes.

dém. :
$(\mathrm{GL}_n(\mathbb{K}), \times)$ et $(\mathbb{K}^\star, \times)$ sont deux groupes et $\det(AB) = \det A \times \det B$ pour tout $A, B \in \mathrm{GL}_n(\mathbb{K})$.
□

Définition

Le noyau de ce morphisme de groupes est noté $\mathrm{SL}_n(\mathbb{K})$, on l'appelle groupe spécial linéaire d'ordre n.

$$\mathrm{SL}_n(\mathbb{K}) = \{A \in \mathcal{M}_n(\mathbb{K}) / \det A = 1\}$$

Théorème

$$\forall A \in \mathcal{M}_n(\mathbb{K}), \det {}^tA = \det A$$

dém. :
$A = (a_{i,j}), {}^tA = (a'_{j,i})$ avec $a'_{j,i} = a_{i,j}$.
Par définition

$$\det {}^tA = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{j=1}^n a'_{\sigma(j),j} = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{j=1}^n a_{j,\sigma(j)}$$

En procédant au changement d'indice $i = \sigma(j)$ dans le produit

$$\det {}^tA = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^n a_{\sigma^{-1}(i),i}$$

En procédant au changement d'indice $\sigma' = \sigma^{-1}$ dans la somme

$$\det {}^tA = \sum_{\sigma' \in \mathfrak{S}_n} \varepsilon(\sigma'^{-1}) \prod_{i=1}^n a_{\sigma'(i),i}$$

Or $\varepsilon(\sigma'^{-1}) = \varepsilon(\sigma')$ donc

$$\det {}^tA = \sum_{\sigma' \in \mathfrak{S}_n} \varepsilon(\sigma') \prod_{i=1}^n a_{\sigma(i),i} = \det A$$

□

Remarque On a donc indifféremment

$$\det A = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{\sigma(i),i} = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)}$$

11.3 Calculs de déterminants

11.3.1 Premiers résultats

Soit $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$. Par définition

$$\det A = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{\sigma(i),i}$$

Pour $n = 1$, $A = (a)$, $\mathfrak{S}_1 = \{\text{Id}\}$ et on obtient

$$\det A = a.$$

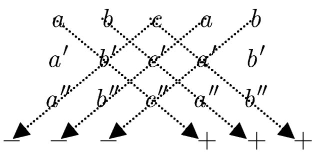
Pour $n = 2$, $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $\mathfrak{S}_2 = \{\text{Id}, (1,2)\}$ et on obtient $\det A = \varepsilon(\text{Id})ad + \varepsilon((1,2))cb$ i.e.

$$\det A = ad - bc \text{ d.}$$

Pour $n = 3$, $A = \begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix}$, $\mathfrak{S}_3 = \{\text{Id}, (1,2), (2,3), (3,1), (1,2,3), (1,3,2)\}$ et on obtient

$$\det A = (ab'c'' + a'b''c + a''bc') - (a''b'c + a'b''c + ab''c').$$

Remarque Pour mettre en place ce dernier calcul, on peut exploiter la règle de Sarrus
- on reproduit les deux premières colonnes de la matrice à sa suite ;
- on somme les produits des coefficients diagonaux obliquant sur la gauche avec un signe moins, sur la droite avec un signe plus.



Exemple Calculons

$$D = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix}$$

Par la règle de Sarrus

$$D = (1 \times 5 \times 3 + 2 \times 6 \times 7 + 3 \times 4 \times 8) - (3 \times 5 \times 7 + 1 \times 6 \times 8 + 2 \times 4 \times 9) = 0$$

11.3.2 Déterminant d'une matrice triangulaire

Proposition

Si $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$ est triangulaire supérieure alors

$$\det A = \prod_{i=1}^{n} a_{i,i}$$

dém. :
Puisque A est triangulaire supérieure on a $a_{i,j} = 0$ pour tout (i, j) tel que $i > j$.
Par définition

$$\det A = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{\sigma(i),i}$$

Soit $\sigma \in \mathfrak{S}_n$. S'il existe $1 \leqslant i \leqslant n$ tel que $\sigma(i) > i$ alors $\prod_{i=1}^{n} a_{\sigma(i),i} = 0$ car le produit contient un facteur nul.
Après simplification des termes correspondants dans la somme définissant $\det A$, il ne reste plus dans celle-ci que les termes associés aux permutations σ vérifiant

$$\forall 1 \leqslant i \leqslant n, \sigma(i) \leqslant i$$

Pour une telle permutation, on a $\sigma(1) = 1, \sigma(2) = 2, \ldots, \sigma(n) = n$ et donc $\sigma = \mathrm{Id}$.
Finalement

$$\det A = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{\sigma(i),i} = \varepsilon(\mathrm{Id}) \prod_{i=1}^{n} a_{i,i} = \prod_{i=1}^{n} a_{i,i}$$

□

Remarque Par transposition, on peut aussi énoncer que le déterminant d'une matrice triangulaire inférieure est le produit de ses coefficients diagonaux.

Remarque En particulier le déterminant d'une matrice diagonale est le produit de ses coefficients diagonaux.

Exemple Calculons le déterminant de l'endomorphisme $f : \mathbb{R}_n [X] \to \mathbb{R}_n [X]$ défini par

$$f(P) = ((X + 1)P)'$$

Formons la matrice de f dans la base canonique $\mathcal{B} = (1, X, \ldots, X^n)$.

$$f(1) = 1, f(X) = 2X + 1, \ldots, f(X^n) = (n + 1)X^n + nX^{n-1}$$

La matrice de f dans $\mathcal{B}$ est triangulaire supérieure de coefficients diagonaux $1, 2, \ldots, n + 1$.
On en déduit $\det f = (n + 1)!$.

11.3.3 Opérations sur les déterminants

Soit $A \in \mathcal{M}_n(\mathbb{K})$ de colonnes $C_1, \dots, C_n$.
Les colonnes $C_1, \dots, C_n$ sont des vecteurs de l'espace $\mathcal{M}_{n,1}(\mathbb{K})$. Notons $\mathcal{B} = (E_1, \dots, E_n)$ la base canonique de cet espace. La matrice des composantes de C_j dans la base $\mathcal{B}$ est C_j. On peut donc interpréter A comme la matrice représentant la famille de vecteurs $(C_1, \dots, C_n)$ dans la base $\mathcal{B}$.
Puisque $A = \mathrm{Mat}_{\mathcal{B}}(C_1, \dots, C_n)$, on a $\det A = \det\limits_{\mathcal{B}}(C_1, \dots, C_n)$. On en déduit le résultat suivant

Théorème

Soit $A \in \mathcal{M}_n(\mathbb{K})$ de colonnes $C_1, \dots, C_n$.
Si l'une des colonnes de A est nulle alors $\det A = 0$.
Si les colonnes de A sont liées alors $\det A = 0$.
Si on permute les colonnes de A selon une permutation σ alors $\det A$ est multiplié par $\varepsilon(\sigma)$.
Si on multiplie une colonne de A par un scalaire λ alors $\det A$ est multiplié par λ.
Si on scinde une colonne C_j de A en $C_j' + C_j''$ alors $\det A$ est la somme des deux déterminants obtenus.
Si on ajoute à une colonne de A une combinaison linéaire des autres colonnes de A, le déterminant de A est inchangé.

dém. :
Puisque $\det A = \det\limits_{\mathcal{B}}(C_1, ..., C_n)$, $A \mapsto \det A$ est une application n-linéaire alternée en les colonnes de A.
□

Corollaire

Par transposition, le théorème ci-dessus se généralise à la manipulation de lignes au lieu de colonnes.

Attention : La manipulation $C_1 \leftarrow C_1 - C_2$ ne modifie pas le déterminant (car correspond à l'ajout à une colonne d'une combinaison linéaire des autres).
En revanche, la manipulation $C_1 \leftarrow C_2 - C_1$ change le déterminant en son opposé car correspond à la précédent manipulation suivie de $C_1 \leftarrow -C_1$.

11.3.4 Calculs par triangulation

Par opérations élémentaires, on sait triangulariser une matrice et donc calculer son déterminant.

Exemple Calculons

$$\begin{vmatrix} 5 & 2 & 1 \\ 10 & 4 & 3 \\ 15 & 8 & 1 \end{vmatrix}$$

En factorisant 5 et 2 sur les deux premières colonnes

$$\begin{vmatrix} 5 & 2 & 1 \\ 10 & 4 & 3 \\ 15 & 8 & 1 \end{vmatrix} = 5 \times 2 \times \begin{vmatrix} 1 & 1 & 1 \\ 2 & 2 & 3 \\ 3 & 4 & 1 \end{vmatrix}$$

Par les opérations $L_2 \leftarrow L_2 - 2L_1$ et $L_3 \leftarrow L_3 - 3L_1$

$$\begin{vmatrix} 5 & 2 & 1 \\ 10 & 4 & 3 \\ 15 & 8 & 1 \end{vmatrix} = 10 \times \begin{vmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & -3 \end{vmatrix}$$

Par l'opération $L_2 \leftrightarrow L_3$

$$\begin{vmatrix} 5 & 2 & 1 \\ 10 & 4 & 3 \\ 15 & 8 & 1 \end{vmatrix} = -10 \times \begin{vmatrix} 1 & 1 & 1 \\ 0 & 1 & -3 \\ 0 & 0 & 1 \end{vmatrix} = -10 \times 1 \times 1 \times 1 = -10$$

Exemple Calculons

$$\begin{vmatrix} 1 & a+b & ab \\ 1 & b+c & bc \\ 1 & c+a & ca \end{vmatrix}$$

Par les opérations $L_2 \leftarrow L_2 - L_1$ et $L_3 \leftarrow L_3 - L_1$

$$\begin{vmatrix} 1 & a+b & ab \\ 1 & b+c & bc \\ 1 & c+a & ca \end{vmatrix} = \begin{vmatrix} 1 & a+b & ab \\ 0 & c-a & b(c-a) \\ 0 & c-b & a(c-b) \end{vmatrix}$$

En factorisant sur les deux premières lignes

$$\begin{vmatrix} 1 & a+b & ab \\ 1 & b+c & bc \\ 1 & c+a & ca \end{vmatrix} = (c-a)(c-b) \begin{vmatrix} 1 & a+b & ab \\ 0 & 1 & b \\ 0 & 1 & a \end{vmatrix}$$

Par l'opération $L_3 \leftarrow L_3 - L_2$

$$\begin{vmatrix} 1 & a+b & ab \\ 1 & b+c & bc \\ 1 & c+a & ca \end{vmatrix} = (c-a)(c-b) \begin{vmatrix} 1 & a+b & ab \\ 0 & 1 & b \\ 0 & 0 & a-b \end{vmatrix} = (c-a)(c-b)(a-b)$$

Exemple Soient $(a,b) \in \mathbb{K}^2$ et $n \geqslant 2$. Calculons

$$D_n = \begin{vmatrix} a & & b \\ & \ddots & \\ b & & a \end{vmatrix}_{[n]}$$

Par l'opération $C_1 \leftarrow C_1 + C_2 + ... + C_n$

$$D_n = \begin{vmatrix} a+(n-1)b & b & \cdots & b \\ \vdots & a & & (b) \\ \vdots & & \ddots & \\ a+(n-1)b & (b) & & a \end{vmatrix}_{[n]}$$

Par les opérations $L_2 \leftarrow L_2 - L_1, L_3 \leftarrow L_3 - L_1, \dots, L_n \leftarrow L_n - L_1$

$$D_n = \begin{vmatrix} a+(n-1)b & b & \cdots & b \\ 0 & a-b & & (0) \\ \vdots & & \ddots & \\ 0 & (0) & & a-b \end{vmatrix}_{[n]}$$

et donc $D_n = (a+(n-1)b)(a-b)^{n-1}$.

Exemple Calculons

$$\begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 2 & 2 & \dots & 2 \\ 1 & 2 & 3 & \dots & 3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 2 & 3 & \dots & n \end{vmatrix}$$

A chaque ligne, retirons la précédente en commençant par la dernière

$$\begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 2 & 2 & \dots & 2 \\ 1 & 2 & 3 & \dots & 3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 2 & 3 & \dots & n \end{vmatrix} = \begin{vmatrix} 1 & & 1 \\ & \ddots & \\ 0 & & 1 \end{vmatrix} = 1$$

Attention : Il faut être attentif à l'ordre des opérations élémentaires, chacune modifiant la matrice sur laquelle opère l'opération élémentaire suivante. Dans l'exemple précédent, on retirait à chaque ligne la précédente et ce en commençant de la dernière et absolument dans cet ordre pour obtenir la matrice proposée..

Exemple Calculons pour $n \geqslant 2$, le déterminant

$$\begin{vmatrix} 1 & & (0) & 1 \\ & \ddots & & (0) \\ & & \ddots & \\ (1) & & & 1 \end{vmatrix}_{[n]}$$

En écrivant la dernière colonne comme somme de deux colonnes

$$\begin{vmatrix} 1 & & 0 & 1 \\ & \ddots & & 0 \\ & & \ddots & \\ (1) & & & 1 \end{vmatrix} = \begin{vmatrix} 1 & & & (0) \\ & \ddots & & \\ & & \ddots & \\ (1) & & & 1 \end{vmatrix} + \begin{vmatrix} 1 & & (0) & 1 \\ & \ddots & & 0 \\ & & 1 & \vdots \\ (1) & & & 0 \end{vmatrix} = 1 + 0$$

le dernier déterminant étant nul car ses deux dernières lignes sont égales.

11.3.5 Développement du déterminant selon une rangée

Commençons par présenter un résultat transformant un déterminant de taille $n+1$ en un déterminant de taille n (donc plus petit…)

Lemme

$$\begin{vmatrix} a_{1,1} & \cdots & a_{1,n} & 0 \\ \vdots & & \vdots & \vdots \\ a_{n,1} & \cdots & a_{n,n} & 0 \\ a_{n+1,1} & \cdots & a_{n+1,n} & 1 \end{vmatrix}_{[n+1]} = \begin{vmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{vmatrix}_{[n]}$$

dém. :
Posons $A = (a_{i,j})_{1\leqslant i,j\leqslant n+1} \in \mathcal{M}_{n+1}(\mathbb{K})$ avec $a_{i,j} = 0$ pour $1 \leqslant i \leqslant n$, $j = n+1$ et $a_{n+1,n+1} = 1$.
Par définition du déterminant

$$\det A = \sum_{\sigma\in\mathfrak{S}_{n+1}} \varepsilon(\sigma) \prod_{i=1}^{n+1} a_{\sigma(i),i}$$

Soit $\sigma \in \mathfrak{S}_{n+1}$.
Si $\sigma(n+1) \neq n+1$ alors $\prod_{i=1}^{n+1} a_{\sigma(i),i} = 0$ car le dernier facteur de ce produit est nul.
Si $\sigma(n+1) = n+1$ alors $\prod_{i=1}^{n+1} a_{\sigma(i),i} = \prod_{i=1}^{n} a_{\sigma(i),i}$ car le dernier facteur de ce produit vaut 1.
Par suite

$$\det A = \sum_{\sigma\in\mathfrak{S}_{n+1},\sigma(n+1)=n+1} \varepsilon(\sigma) \prod_{i=1}^{n} a_{\sigma(i),i}$$

Or l'application qui prolonge une permutation $\sigma \in \mathfrak{S}_n$ en posant $\sigma(n+1) = n+1$ est une bijection de $\mathfrak{S}_n$ vers $\{\sigma \in \mathfrak{S}_{n+1}/\sigma(n+1) = n+1\}$. De plus cette bijection conserve la signature car elle conserve le nombre d'inversions. On en déduit

$$\det A = \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{\sigma(i),i}$$

et on obtient ainsi la formule énoncée.
Pour exploiter le résultat précédent, réinterprétons le déterminant d'une matrice comme le déterminant de la famille de ses colonnes.
Soit $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$ de colonnes $C_1, \dots, C_n$. On a $\det A = \det_{\mathcal{B}}(C_1, \dots, C_n)$ en notant $\mathcal{B} = (E_1, \dots, E_n)$ la base canonique de l'espace $\mathcal{M}_{n,1}(\mathbb{K})$.
Soit $1 \leqslant j \leqslant n$ fixé.
On a

$$C_j = \sum_{i=1}^{n} a_{i,j} E_i.$$

Par linéarité de l'application $\det_{\mathcal{B}}$ en sa j-ème variable, on peut écrire

$$\det A = \sum_{i=1}^{n} a_{i,j} \det_{\mathcal{B}}(C_1, \dots, C_{j-1}, E_i, C_{j+1}, \dots, C_n)$$

soit encore $\det A = \sum_{i=1}^{n} a_{i,j} A_{i,j}$ avec

$$A_{i,j} = \begin{vmatrix} a_{1,1} & \cdots & 0 & \cdots & a_{1,n} \\ & & \vdots & & \\ & & 1 & & \\ & & \vdots & & \\ a_{n,1} & \cdots & 0 & \cdots & a_{n,n} \end{vmatrix} \text{ où le 1 est en position } (i,j).$$

En permutant les colonnes selon le cycle $\sigma = (j, j+1, ..., n)$:

$$A_{i,j} = (-1)^{n-j} \begin{vmatrix} a_{1,1} & \cdots & \hat{a}_{1,j} & \cdots & a_{1,n} & 0 \\ & & \vdots & & & \vdots \\ & & \vdots & & & 1 \\ & & \vdots & & & \vdots \\ a_{n,1} & \cdots & \hat{a}_{n,j} & \cdots & a_{n,n} & 0 \end{vmatrix} \text{ car } \varepsilon(\sigma) = (-1)^{n-j}.$$

En permutant les lignes selon le cycle $\sigma' = (i, i+1, \ldots, n)$:

$$A_{i,j} = (-1)^{n-i} \times (-1)^{n-j} \begin{vmatrix} a_{1,1} & \cdots & a_{1,n} & 0 \\ \vdots & \hat{a}_{i,j} & \vdots & \vdots \\ a_{n,1} & \cdots & a_{n,n} & 0 \\ a_{i,1} & \cdots & a_{i,n} & 1 \end{vmatrix} \text{ car } \varepsilon(\sigma') = (-1)^{n-i}.$$

et donc, en vertu de précédent lemme

$$A_{i,j} = (-1)^{i+j} \times \begin{vmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \hat{a}_{i,j} & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{vmatrix}_{[n-1]}$$

Finalement

$$\det A = \sum_{i=1}^{n} a_{i,j} A_{i,j} = \sum_{i=1}^{n} (-1)^{i+j} a_{i,j} \Delta_{i,j}$$

avec

$$\Delta_{i,j} = \begin{vmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \hat{a}_{i,j} & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{vmatrix}_{[n-1]}.$$

□

Définition

Cette relation est appelée développement de $\det A$ selon la $j$ème colonne.
Le coefficient $\Delta_{i,j}$ est appelé mineur d'indice (i, j) de A.
Le coefficient $A_{i,j} = (-1)^{i+j}\Delta_{i,j}$ est appelé cofacteur d'indice (i, j) de A.

Remarque Cette formule transforme le calcul d'un déterminant d'ordre n à celui de n déterminants d'ordre $n-1$.

Remarque Le signe de $(-1)^{i+j}$ est donné par le tableau

$$\begin{pmatrix} + & - & + & \cdots & (-1)^{n+1} \\ - & + & - & & \\ + & - & + & & \\ \vdots & & & \ddots & \\ (-1)^{n+1} & & & & + \end{pmatrix}.$$

Exemple Calculons

$$\Delta = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix}$$

En développant selon la première colonne

$$\Delta = 1 \times \begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} - 4 \times \begin{vmatrix} 2 & 3 \\ 8 & 9 \end{vmatrix} + 7 \times \begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix} = -3 + 24 - 21 = 0$$

Par transposition, le développement d'un déterminant selon une colonne devient le développement d'un déterminant selon une ligne. Pour $1 \leqslant i \leqslant n$ fixé, on obtient la relation :

$$\det A = \sum_{j=1}^{n} a_{i,j} A_{i,j} = \sum_{j=1}^{n} (-1)^{i+j} a_{i,j} \Delta_{i,j}$$

qui est appelée développement de $\det A$ selon la i-ème ligne de A.

Exemple Calculons

$$\Delta = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix}$$

En développant selon la troisième ligne

$$\Delta = 7 \times \begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix} - 8 \times \begin{vmatrix} 1 & 3 \\ 4 & 6 \end{vmatrix} + 9 \times \begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix} = -21 + 48 - 27 = 0$$

11.3.6 Calculs par développements

Exemple Calculons

$$\Delta = \begin{vmatrix} 1 & 2 & -1 & 1 \\ -1 & 2 & 1 & 3 \\ 0 & 1 & 0 & -1 \\ 2 & 1 & 1 & 1 \end{vmatrix}$$

En développant selon la troisième ligne

$$\Delta = -1 \times \begin{vmatrix} 1 & -1 & 1 \\ -1 & 1 & 3 \\ 2 & 1 & 1 \end{vmatrix} - (-1) \times \begin{vmatrix} 1 & 2 & -1 \\ -1 & 2 & 1 \\ 2 & 1 & 1 \end{vmatrix}$$

Par opérations élémentaires, on fait apparaître des zéros

$$\Delta = -\begin{vmatrix} 0 & -1 & 1 \\ 0 & 1 & 3 \\ 3 & 1 & 1 \end{vmatrix} + \begin{vmatrix} 1 & 2 & 0 \\ -1 & 2 & 0 \\ 2 & 1 & 3 \end{vmatrix}$$

Puis en développant

$$\Delta = -3 \begin{vmatrix} -1 & 1 \\ 1 & 3 \end{vmatrix} + 3 \times \begin{vmatrix} 1 & 2 \\ -1 & 2 \end{vmatrix} = 12 + 12 = 24$$

Exemple Calculons

$$D_n = \begin{vmatrix} 2 & 1 & & (0) \\ 1 & \ddots & \ddots & \\ & \ddots & \ddots & 1 \\ (0) & & 1 & 2 \end{vmatrix}_{[n]}$$

En développant selon la première ligne

$$D_n = 2D_{n-1} - 1 \times \begin{vmatrix} 1 & 1 & 0 & \cdots & 0 \\ 0 & 2 & 1 & & (0) \\ \vdots & 1 & \ddots & \ddots & \\ \vdots & & \ddots & \ddots & 1 \\ 0 & (0) & & 1 & 2 \end{vmatrix}_{[n-1]}$$

et en développant le deuxième déterminant selon la première colonne

$$D_n = 2D_{n-1} - D_{n-2}$$

Ainsi (D_n) est une suite récurrente linéaire d'ordre 2 d'équation caractéristique $r^2 - 2r + 1 = 0$ de racine double 1.
On peut donc écrire

$$D_n = (\lambda n + \mu) \times 1^n$$

$D_1 = 2$ et $D_2 = 3$ donne $\lambda = \mu = 1$ et ainsi $D_n = n + 1$ pour tout $n \in \mathbb{N}^\star$

Exemple Soit $\theta \in]0, \pi[$.
Calculons

$$D_n = \begin{vmatrix} 2\cos\theta & 1 & & (0) \\ 1 & \ddots & \ddots & \\ & \ddots & \ddots & 1 \\ (0) & & 1 & 2\cos\theta \end{vmatrix}_{[n]}$$

En développant selon la première ligne

$$D_n = 2D_{n-1} - 1 \times \begin{vmatrix} 1 & 1 & 0 & \cdots & 0 \\ 0 & 2\cos\theta & 1 & & (0) \\ \vdots & 1 & \ddots & \ddots & \\ \vdots & & \ddots & \ddots & 1 \\ 0 & (0) & & 1 & 2\cos\theta \end{vmatrix}_{[n-1]}$$

et en développant le deuxième déterminant selon la première colonne

$$D_n = 2\cos\theta . D_{n-1} - D_{n-2}$$

Ainsi (D_n) est une suite récurrente linéaire d'ordre 2 d'équation caractéristique $r^2 - 2\cos\theta . r + 1 = 0$ de racines $e^{i\theta}$ et $e^{-i\theta}$ (distinctes car $\theta \in]0, \pi[$)

On peut donc écrire

$$D_n = (\lambda \cos n\theta + \mu \sin n\theta) \times 1^n$$

$D_1 = 2\cos\theta$ et $D_2 = 4\cos^2\theta - 1$ donne $\lambda = 1$ et $\mu = \cos\theta / \sin\theta$ et ainsi pour tout $n \in \mathbb{N}^\star$

$$D_n = \frac{\sin(n+1)\theta}{\sin\theta}$$

Exemple Soit $(x_1, ..., x_n) \in \mathbb{K}^n$.
Calculons

$$V_n(x_1, ..., x_n) = \begin{vmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{vmatrix}$$

Pour $n = 1$, $V_1(x_1) = 1$.
Pour $n = 2$,

$$V_2(x_1, x_2) = \begin{vmatrix} 1 & x_1 \\ 1 & x_2 \end{vmatrix} = x_2 - x_1$$

Pour $n = 3$. En procédant aux opérations élémentaires $C_3 \leftarrow C_3 - x_1 C_2$ et $C_2 \leftarrow C_2 - x_1 C_1$ on obtient

$$V_3(x_1, x_2, x_3) = \begin{vmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 1 & x_2 - x_1 & x_2^2 - x_1 x_2 \\ 1 & x_3 - x_1 & x_3^2 - x_1 x_3 \end{vmatrix}$$

En développant selon la première ligne puis en factorisant par ligne

$$V_3(x_1, x_2, x_3) = \begin{vmatrix} x_2 - x_1 & x_2^2 - x_1 x_2 \\ x_3 - x_1 & x_3^2 - x_1 x_3 \end{vmatrix} = (x_2 - x_1)(x_3 - x_1) V_2(x_2, x_3)$$

et donc
$V_3(x_1, x_2, x_3) = (x_2 - x_1)(x_3 - x_1)(x_3 - x_2)$.
Plus généralement, pour $n \geqslant 2$, en procédant aux opérations élémentaires
$C_n \leftarrow C_n - x_1 C_{n-1}, \ldots, C_2 \leftarrow C_2 - x_1 C_1$

$$V_n(x_1, ..., x_n) = \begin{vmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{vmatrix}$$

$$= \begin{vmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & x_2 - x_1 & x_2^2 - x_1 x_2 & \cdots & x_2^{n-1} - x_1 x_2^{n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n - x_1 & x_n^2 - x_1 x_n & \cdots & x_n^{n-1} - x_1 x_n^{n-2} \end{vmatrix}$$

En développant selon la première ligne puis en factorisant par ligne

$$V_n(x_1, ..., x_n) = \begin{vmatrix} x_2 - x_1 & x_2^2 - x_1 x_2 & \cdots & x_2^{n-1} - x_1 x_2^{n-2} \\ \vdots & \vdots & & \vdots \\ x_n - x_1 & x_n^2 - x_1 x_n & \cdots & x_n^{n-1} - x_1 x_n^{n-2} \end{vmatrix}$$

$$= (x_2 - x_1) \ldots (x_n - x_1) V_{n-1}(x_2, \ldots, x_n)$$

Autrement dit

$$V_n(x_1,\dots,x_n)=\prod_{i=2}^{n}(x_i-x_1)V_{n-1}(x_2,\dots,x_n)$$

En répétant le processus

$$V_n(x_1,\dots,x_n)=\prod_{i=2}^{n}(x_i-x_1)\prod_{i=3}^{n}(x_i-x_2)V_{n-2}(x_3,\dots,x_n)$$

et ainsi de suite pour obtenir finalement

$$V_n(x_1,...,x_n)=\prod_{1\leqslant i<j\leqslant n}(x_j-x_i)$$

Remarque La matrice sous-jacente est inversible si, et seulement si, les x_i sont 2 à 2 distincts.

11.4 Applications des déterminants

11.4.1 Formules de Cramer

Théorème

Soit Σ un système à n équations et n inconnues

$$\begin{cases} a_{1,1}x_1+\dots+a_{1,n}x_n=b_1 \\ \vdots \\ a_{n,1}x_1+\dots+a_{n,n}x_n=b_n \end{cases}$$

d'équation matricielle $AX=B$.
Le système Σ est de Cramer si, et seulement si, $\det A\neq 0$.
De plus, si tel est le cas, son unique solution est le n uplet $(x_1,...,x_n)$ avec

$$\forall 1\leqslant j\leqslant n, x_j=\frac{\begin{vmatrix} a_{1,1} & \cdots & b_1 & \cdots & a_{1,n} \\ \vdots & & \vdots & & \vdots \\ a_{n,1} & \cdots & b_n & \cdots & a_{n,n}\end{vmatrix}}{\begin{vmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,n}\end{vmatrix}}$$

dém. :
Le système Σ est de Cramer si, et seulement si, la matrice A inversible i.e. $\det A\neq 0$.
Supposons que tel est le cas. Σ admet une unique solution $(x_1,\dots,x_n)$.
Si on note $C_1,\dots,C_n$ les colonnes de A, on a

$$B=x_1C_1+\dots+x_nC_n=\sum_{i=1}^{n}x_iC_i$$

Notons $\mathcal{B} = (E_1, \dots, E_n)$ la base canonique de $\mathcal{M}_{n,1}(\mathbb{K})$.
$\det A = \det_{\mathcal{B}}(C_1, \dots, C_n)$.
Pour $j \in \{1, \dots, n\}$, considérons $\det_{\mathcal{B}}(C_1, \dots, C_{j-1}, B, C_{j+1}, \dots, C_n)$.
Par linéarité du déterminant en chacune de ses variables,

$$\det_{\mathcal{B}}(C_1, \dots, C_{j-1}, B, C_{j+1}, \dots, C_n) = \sum_{i=1}^{n} x_i \det_{\mathcal{B}}(C_1, \dots, C_{j-1}, C_i, C_{j+1}, \dots, C_n)$$

Or

$$\det_{\mathcal{B}}(C_1, \dots, C_{j-1}, C_i, C_{j+1}, \dots, C_n) = 0$$

si $i \neq j$ (car deux colonnes égales) et

$$\det_{\mathcal{B}}(C_1, \dots, C_{j-1}, C_j, C_{j+1}, \dots, C_n) = \det A$$

donc

$$\det_{\mathcal{B}}(C_1, \dots, C_{j-1}, B, C_{j+1}, \dots, C_n) = x_j \det A$$

□

Exemple Considérons le système

$$\begin{cases} ax + by = c \\ a'x + b'y = c' \end{cases}$$

Ce système est de Cramer si, et seulement si, $ab' - a'b \neq 0$.
Si tel est le cas, sa solution est

$$x = \frac{cb' - c'b}{ab' - a'b}, y = \frac{ac' - a'c}{ab' - a'b}$$

Exemple Considérons le système

$$\begin{cases} x + y + z = 1 \\ ax + by + cz = d \\ a^2x + b^2y + c^2z = d^2 \end{cases}$$

avec a, b, c deux à deux distincts.
Le déterminant (de Vandermonde) de ce système est $(b-a)(c-a)(c-b) \neq 0$.
Ce système est de Cramer et sa solution (x, y, z) est donnée par

$$x = \frac{(b-d)(c-d)}{(b-a)(c-a)}, y = \frac{(d-a)(c-d)}{(b-a)(c-b)}, z = \frac{(d-a)(d-b)}{(c-a)(c-b)}$$

11.4.2 Comatrice

Soit $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$.
Pour $1 \leqslant i, j \leqslant n$, on note $A_{i,j}$ le cofacteur d'indice (i, j) de la matrice A.

$$A_{i,j} = (-1)^{i+j} \Delta_{i,j} = (-1)^{i+j} \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \hat{a}_{ij} & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}.$$

Définition

On appelle comatrice de A la matrice des cofacteurs de A, on la note

$$\text{com}A = (A_{i,j})_{1\leqslant i,j\leqslant n} \in \mathcal{M}_n(\mathbb{K})$$

Exemple Pour

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

on obtient

$$\text{com}A = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$$

Exemple Pour

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

on obtient

$$\text{com}A = \begin{pmatrix} -3 & 6 & -3 \\ 6 & -12 & 6 \\ -3 & 6 & -3 \end{pmatrix}$$

Théorème

$$\forall A \in \mathcal{M}_n(\mathbb{K}), {}^t(\text{com}A)\, A = A^t\,(\text{com}A) = \det(A) I_n$$

dém. :
Notons $a_{i,j}$ les coefficients de la matrice A.
$\text{com}A = (A_{i,j}), {}^t(\text{com}A) = (A'_{j,i})$ avec $A'_{j,i} = A_{i,j}$.
${}^t(\text{com}A)\, A = (b_{i,j})$ avec

$$b_{i,j} = \sum_{k=1}^{n} A'_{i,k} a_{k,j} = \sum_{k=1}^{n} a_{k,j} A_{k,i}$$

Si $i = j$ alors en développant le déterminant de la matrice A selon la j-ème colonne

$$\det A = \begin{vmatrix} a_{1,1} & \cdots & a_{1,j} & \cdots & a_{1,n} \\ \vdots & & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,j} & \cdots & a_{n,n} \end{vmatrix} = \sum_{k=1}^{n} a_{k,j} A_{k,j} = b_{j,j}$$

Si $i \neq j$ alors en développant selon la i-ème colonne le déterminant de la matrice obtenue en remplaçant la i-ème colonne de A par sa $j$ème colonne, on obtient

$$\begin{vmatrix} a_{1,1} & \cdots & a_{1,i-1} & a_{1,j} & a_{1,i+1} & \cdots & a_{1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,i-1} & a_{n,j} & a_{n,i+1} & \cdots & a_{n,n} \end{vmatrix} = \sum_{k=1}^{n} a_{k,j} A_{k,i} = b_{i,j}$$

Or ce dernier déterminant est nul car la matrice considérée possède deux colonnes identiques.
Finalement $b_{i,j} = \det(A)\delta_{i,j}$ puis ${}^t(\text{com}A)\,A = \det(A)I_n$.
L'identité $A^t(\text{com}A) = \det(A)I_n$ s'obtient de façon semblable en considérant les coefficients de la matrice produit $A^t(\text{com}A)$ comme des développements de déterminants selon des lignes.
□

Corollaire

Si $\det A \neq 0$ alors

$$A^{-1} = \frac{1}{\det A}{}^t(\text{com}A)$$

Exemple Pour

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbb{K})$$

A est inversible si, et seulement si, $ad - bc \neq 0$. Si tel est le cas

$$A^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

11.4.3 Détermination du rang d'une matrice

Définition

On appelle matrice extraite de $A \in \mathcal{M}_{n,p}(\mathbb{K})$ toute matrice formée en retirant un certain nombre de lignes et/ou de colonnes de la matrice A.

Exemple $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 2 & 2 \end{pmatrix}$ est extraite de $\begin{pmatrix} 1 & 0 & 2 & 3 \\ -2 & 0 & 1 & 2 \\ 4 & 1 & 2 & 2 \end{pmatrix}$ via $\hat{L}_2$ et $\hat{C}_2$.

Proposition

Toute matrice extraite de $A \in \mathcal{M}_{n,p}(\mathbb{K})$ est de rang inférieur à celui de A.

dém.
Notons $C_1, \ldots, C_p$ les colonnes de A. On a $\text{rg}A = \text{rg}(C_1, \ldots, C_p)$.
Si on retire des colonnes à la matrice A, la matrice B obtenue est de rang inférieur.
Notons $L_1, \ldots, L_n$ les colonnes de B. On a $\text{rg}B = \text{rg}(L_1, \ldots, L_n)$.
Si on retire des lignes de B, la matrice extraite C obtenue est de rang inférieur au rang de B et donc de rang inférieur à celui de A.

Définition

On appelle déterminant extrait de $A \in \mathcal{M}_{n,p}(\mathbb{K})$ d'ordre r, tout déterminant d'une matrice carrée d'ordre r extraite de A.

Exemple $\begin{vmatrix} 1 & 2 \\ 1 & 0 \end{vmatrix} = -2$ est déterminant extrait de $\begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 0 \end{pmatrix}$ via $\hat{C}_2$

Exemple Les mineurs de $A \in \mathcal{M}_n(\mathbb{K})$ sont exactement ses déterminants extraits d'ordre $n-1$.

Théorème

Le rang d'une matrice non nulle $A \in \mathcal{M}_{n,p}(\mathbb{K})$ est l'ordre maximal des déterminants non nuls extraits de A.

dém. :
Si A possède un déterminant extrait d'ordre r non nul alors $\text{rg}A \geqslant r$ car A admet une matrice extraite de rang r.
Inversement, posons $r \geqslant 1$ le rang de $A \in \mathcal{M}_{n,p}(\mathbb{K})$.
A possède r colonnes indépendantes.
Notons B la matrice extraite formée par ces colonnes. $B \in \mathcal{M}_{n,r}(\mathbb{K})$ et B est de rang r.
B possède r lignes indépendantes.
Notons C la matrice extraite de B constituée pas ces lignes. $C \in \mathcal{M}_r(\mathbb{K})$ et C est de rang r.
La matrice C est une matrice carrée inversible donc $\det C \neq 0$ et par suite A possède un déterminant extrait d'ordre r non nul.
□

Exemple Calculons le rang de

$$M = \begin{pmatrix} 1 & 1 & & (0) \\ & \ddots & \ddots & \\ (0) & & \ddots & 1 \\ 1 & (0) & & 1 \end{pmatrix} \in \mathcal{M}_n(\mathbb{K})$$

En développant selon la première colonne

$$\det M = \begin{vmatrix} 1 & 1 & & (0) \\ & \ddots & \ddots & \\ & & 1 & 1 \\ 0 & & & 1 \end{vmatrix}_{[n-1]} + (-1)^{n+1} \begin{vmatrix} 1 & & & (0) \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ (0) & & 1 & 1 \end{vmatrix}_{[n-1]}$$

Ainsi $\det M = 1 + (-1)^{n+1}$.
Si n est impair alors $\det M = 2$ et donc $\text{rg}M = n$.
Si n est pair alors $\det M = 0$ et donc $\text{rg}M < n$.
Or M possède un mineur non nul donc $\text{rg}M \geqslant n-1$ puis $\text{rg}M = n-1$.

11.4.4 Orientation d'un espace vectoriel réel de dimension finie

E désigne un $\mathbb{R}$-espace vectoriel de dimension $n \in \mathbb{N}^\star$.

Définition

Soient $\mathcal{B} = (e_1, ..., e_n)$ et $\mathcal{B}' = (e'_1, ..., e'_n)$ deux bases de E.
On dit que la base $\mathcal{B}'$ a même orientation que la base $\mathcal{B}$ si $\det_{\mathcal{B}} \mathcal{B}' > 0$.
Sinon, on dit que $\mathcal{B}'$ est d'orientation opposée à celle de $\mathcal{B}$.

Proposition

Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E.
Si $\mathcal{B}$ a même orientation que $\mathcal{B}'$ alors $\mathcal{B}'$ a même orientation que $\mathcal{B}$.
On dit alors que $\mathcal{B}$ et $\mathcal{B}'$ ont même orientation.

dém. :
$\det_{\mathcal{B}} \mathcal{B}' \times \det_{\mathcal{B}'} \mathcal{B} = 1 > 0$ donc $\det_{\mathcal{B}'} \mathcal{B}$ et $\det_{\mathcal{B}} \mathcal{B}'$ ont même signe.
□

Proposition

Soient $\mathcal{B}, \mathcal{B}', \mathcal{B}''$ trois bases de E.
$\mathcal{B}'$ et $\mathcal{B}''$ ont même orientation si, et seulement si, elles ont toutes les deux la même orientation que $\mathcal{B}$ ou sont toutes les deux d'orientation opposée à celle de $\mathcal{B}$.

dém. :
Par formule de changement de base $\det_{\mathcal{B}'} \mathcal{B}'' = \det_{\mathcal{B}'} \mathcal{B} \times \det_{\mathcal{B}} \mathcal{B}''$.
Par suite $\det_{\mathcal{B}'} \mathcal{B}'' > 0$ si, et seulement si, $\det_{\mathcal{B}'} \mathcal{B}$ et $\det_{\mathcal{B}} \mathcal{B}''$ ont même signe.
□

Définition

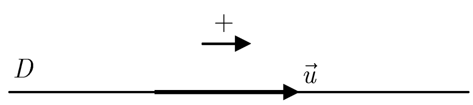
Orienter le $\mathbb{R}$-espace vectoriel E à partir du choix d'une base $\mathcal{B}$ c'est adopter le vocabulaire suivant :
Toute base de E de même orientation que $\mathcal{B}$ est dite directe.
Toute base de E d'orientation opposée à celle de $\mathcal{B}$ est dite indirecte.
La base $\mathcal{B}$ est dite base orientée de référence.

Remarque Il n'y a que deux orientations possibles sur l'espace E.
En effet l'ensemble des bases de E se scinde en deux sous-ensembles formés de bases qui sont de même orientation. Orienter E revient à choisir l'un de ces sous-ensembles et de qualifier de directes les bases de celui-ci et d'indirectes les bases de l'autre sous-ensemble.

Remarque L'espace E ne possède pas d'orientation privilégiée a priori.

Exemple En général, on oriente $\mathbb{R}^n$ par l'intermédiaire de sa base canonique. Si tel est le cas, on dit que l'espace $\mathbb{R}^n$ est canoniquement orienté.

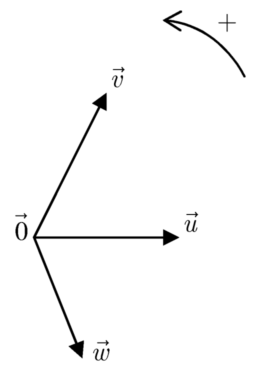
Exemple Pour orienter une droite D, il suffit de choisir un vecteur $u \neq 0$ de D.
Tout vecteur non nul de D ayant même sens que u est direct. Les autres sont indirects.



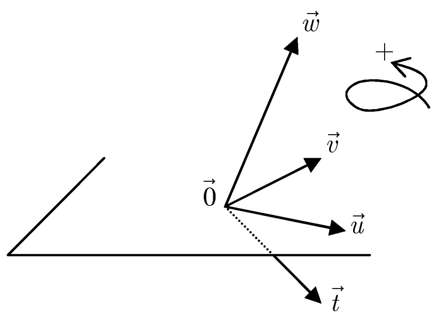
Définition

Une droite vectoriel orientée est appelé un axe vectoriel.

Exemple Le plan géométrique est usuellement orienté par le sens trigonométrique.
Dans la figure ci-dessous la famille $(\vec{u}, \vec{v})$ est directe et la famille $(\vec{u}, \vec{w})$ est indirecte



Exemple L'espace géométrique est usuellement orienté par la règle du tire-bouchon.
Dans la figure ci-dessous, la famille $(\vec{u}, \vec{v}, \vec{w})$ est directe et la famille $(\vec{u}, \vec{v}, \vec{t})$ est indirecte



Exemple Soit E un $\mathbb{R}$-espace vectoriel orienté par une base $\mathcal{B} = (e_1, ..., e_n)$.
Pour tout $\sigma \in \mathfrak{S}_n$, notons $\mathcal{B}_\sigma = (e_{\sigma(1)}, ..., e_{\sigma(n)})$.
On a $\det_{\mathcal{B}} \mathcal{B}_\sigma = \varepsilon(\sigma)$.
Si la permutation σ est paire alors $\mathcal{B}_\sigma$ est directe.
Si la permutation σ est impaire alors $\mathcal{B}_\sigma$ est indirecte.
En particulier, dans l'espace géométrique orienté $\mathcal{B} = (\vec{i}, \vec{j}, \vec{k})$
- les bases $(\vec{i}, \vec{j}, \vec{k}), (\vec{k}, \vec{i}, \vec{j}), (\vec{j}, \vec{k}, \vec{i})$ sont directes ;
- les bases $(\vec{i}, \vec{k}, \vec{j}), (\vec{k}, \vec{j}, \vec{i}), (\vec{j}, \vec{i}, \vec{k})$ sont indirectes.

Chapitre 12

Espaces vectoriels euclidiens

Dans ce chapitre on redéfinit la notion de produit scalaire dans un cadre plus général qui ne sera pas nécessairement géométrique. A partir de cette notion, on redéfinit les notions de distance, d'orthogonalité, d'angles,...
E désigne un $\mathbb{R}$-espace vectoriel.
n et p désignent des entiers naturels.

12.1 Produit scalaire

12.1.1 Définition

Définition

On appelle produit scalaire sur E toute application $\varphi : E \times E \to \mathbb{R}$ vérifiant
1) φ est bilinéaire ;
2) φ est symétrique i.e.

$$\forall x, y \in E, \varphi(x, y) = \varphi(y, x)$$

3) φ est positive i.e.

$$\forall x \in E, \varphi(x, x) \geqslant 0$$

4) φ est définie i.e.

$$\forall x \in E, \varphi(x, x) = 0 \Rightarrow x = 0_E$$

On dit qu'un produit scalaire est un forme bilinéaire symétrique définie positive.

Remarque Dans la pratique pour montrer que $\varphi : E \times E \to \mathbb{R}$ est bilinéaire symétrique (1. et 2.) on se contente d'établir la linéarité par rapport à une variable (par exemple, la deuxième) et la symétrie ce qui est suffisant pour conclure.

Exemple Le produit scalaire défini sur l'ensemble des vecteurs du plan (resp. de l'espace) est un produit scalaire au sens précédent.

Exemple Considérons $E = \mathbb{R}^n$.
Pour $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ et $y = (y_1, \dots, y_n) \in \mathbb{R}^n$ on pose

$$\varphi(x, y) = x_1 y_1 + \cdots + x_n y_n = \sum_{k=1}^{n} x_k y_k$$

Montrons que φ définit un produit scalaire sur $\mathbb{R}^n$.
$\varphi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ est bien définie
Soient $\lambda, \mu \in \mathbb{R}$ et $x, y, z \in \mathbb{R}^n$.
On écrit

$$x = (x_1, \dots, x_n), y = (y_1, \dots, y_n), z = (z_1, \dots, z_n)$$

et alors

$$\varphi(x, \lambda y + \mu z) = \sum_{k=1}^{n} x_k(\lambda y_k + \mu z_k) = \lambda \sum_{k=1}^{n} x_k y_k + \mu \sum_{k=1}^{n} x_k z_k = \lambda\varphi(x, y) + \mu\varphi(x, z)$$

Donc φ est linéaire en sa deuxième variable

$$\varphi(y, x) = \sum_{k=1}^{n} y_k x_k = \sum_{k=1}^{n} x_k y_k = \varphi(x, y)$$

Donc φ est symétrique puis bilinéaire

$$\varphi(x, x) = \sum_{k=1}^{n} x_k^2 \geqslant 0$$

Donc φ est positive
Si $\varphi(x, x) = 0$ alors $\sum_{k=1}^{n} x_k^2 = 0$.
Or la somme de quantités positives n'est nulle que si chacune des quantité est nulle. Par suite $x_1 = \ldots = x_n = 0$ et donc $x = 0_{\mathbb{R}^n}$.
Ainsi φ est définie.
Finalement φ est un produit scalaire sur $\mathbb{R}^n$, on l'appelle produit scalaire canonique de $\mathbb{R}^n$.
Sauf mention du contraire, $\mathbb{R}^n$ sera considéré muni de ce produit scalaire.

Remarque $\varphi(x, y) = \sum_{k=1}^{n} k x_k y_k$ définit aussi un produit scalaire sur $\mathbb{R}^n$.

Exemple Considérons $E = \mathbb{C}$ vu comme $\mathbb{R}$-espace vectoriel.
Pour $z, z' \in \mathbb{C}$ on pose

$$\varphi(z, z') = \text{Re}(\bar{z}z')$$

$\varphi : \mathbb{C} \times \mathbb{C} \to \mathbb{R}$ est bien définie.
Soient $\lambda, \mu \in \mathbb{R}$ et $z, z', z'' \in \mathbb{C}$.

$$\varphi(z, \lambda z' + \mu z'') = \text{Re}\left(\bar{z}(\lambda z' + \mu z'')\right) = \lambda\text{Re}(\bar{z}z') + \mu\text{Re}(\bar{z}z'') = \lambda\varphi(z, z') + \mu\varphi(z, z'')$$

$$\varphi(z', z) = \text{Re}(\bar{z'}z) = \text{Re}(\overline{\bar{z'}z}) = \text{Re}(z'\bar{z}) = \text{Re}(\bar{z}z') = \varphi(z, z')$$

$$\varphi(z, z) = \text{Re}(\bar{z}z) = \text{Re}(|z|^2) = |z|^2 \geqslant 0$$

Si $\varphi(z, z) = 0$ alors $|z|^2 = 0$ et donc $z = 0$.
φ est un produit scalaire sur le $\mathbb{R}$-espace vectoriel $\mathbb{C}$.

Remarque En fait, pour $z = a + ib$ et $z' = a' + ib'$ (avec $a, b, a', b' \in \mathbb{R}$), on a

$$\varphi(z, z') = aa' + bb'$$

Ce produit scalaire est appelé produit scalaire canonique sur $\mathbb{C}$.

Exemple Considérons $E = \mathcal{C}^0([a, b], \mathbb{R})$ (avec $a < b \in \mathbb{R}$).
Pour $f, g \in E$ on pose

$$\varphi(f, g) = \int_a^b f(t)g(t)\,\mathrm{d}t$$

$\varphi : E \times E \to \mathbb{R}$ est bien définie.
Soient $\lambda, \mu \in \mathbb{R}$ et $f, g, h \in E$.

$$\varphi(f, \lambda g + \mu h) = \int_a^b f(t)(\lambda g(t) + \mu h(t))\,\mathrm{d}t$$

$$= \lambda \int_a^b f(t)g(t)\,\mathrm{d}t + \mu \int_a^b f(t)h(t)\,\mathrm{d}t = \lambda\varphi(f, g) + \mu\varphi(f, h)$$

$$\varphi(g, f) = \int_a^b g(t)f(t)\,\mathrm{d}t = \int_a^b f(t)g(t)\,\mathrm{d}t = \varphi(f, g)$$

$$\varphi(f, f) = \int_a^b f^2(t)\,\mathrm{d}t \geqslant 0$$

Si $\varphi(f, f) = 0$ alors $\displaystyle\int_a^b f^2(t)\,\mathrm{d}t = 0$.
Or la fonction f^2 est continue positive sur $[a, b]$ donc $f = 0$.
Finalement φ est un produit scalaire sur E.

Exemple Considérons $E = \mathbb{R}[X]$
Pour $P, Q \in \mathbb{R}[X]$ on pose

$$\varphi(P, Q) = \int_{-1}^1 P(t)Q(t)\,\mathrm{d}t$$

$\varphi : E \times E \to \mathbb{R}$ est bien définie.
Soient $\lambda, \mu \in \mathbb{R}$ et $P, Q, R \in E$.
Par linéarité de l'intégrale

$$\varphi(P, \lambda Q + \mu R) = \lambda\varphi(P, Q) + \mu\varphi(P, R)$$

$$\varphi(Q, P) = \int_{-1}^1 Q(t)P(t)\,\mathrm{d}t = \int_{-1}^1 P(t)Q(t)\,\mathrm{d}t = \varphi(P, Q)$$

$$\varphi(P, P) = \int_{-1}^1 P^2(t)\,\mathrm{d}t \geqslant 0$$

Si $\varphi(P, P) = 0$ alors $\displaystyle\int_{-1}^b P^2(t)\,\mathrm{d}t = 0$.
Or la fonction $t \mapsto P^2(t)$ est continue positive sur $[-1, 1]$ donc c'est la fonction nulle et puisque le polynôme P admet alors une infinité de racines, c'est le polynôme nul.
Finalement φ est un produit scalaire sur E.

Exemple Considérons $E = \mathcal{M}_{n,p}(\mathbb{R})$
Pour $A, B \in E$ on pose

$$\varphi(A, B) = \text{tr}({}^t AB)$$

$\varphi : E \times E \to \mathbb{R}$ est bien définie.
Soient $\lambda, \mu \in \mathbb{R}$ et $A, B, C \in E$.

$$\varphi(A, \lambda B + \mu C) = \text{tr}\left({}^t A(\lambda B + \mu C)\right) = \lambda \text{tr}\left({}^t AB\right) + \mu \text{tr}\left({}^t AC\right) = \lambda\varphi(A, B) + \mu\varphi(A, C)$$

$$\varphi(B, A) = \text{tr}\left({}^t BA\right) = \text{tr}^t\left({}^t BA\right) = \text{tr}({}^t AB) = \varphi(A, B)$$

Si les $a_{i,j}$ désignent les coefficients de A,

$$\varphi(A, A) = \sum_{i=1}^{n} \left[{}^t AA\right]_{i,i} = \sum_{i=1}^{n}\sum_{j=1}^{p} a_{i,j}^2 \geqslant 0$$

Si $\varphi(A, A) = 0$ alors $\displaystyle\sum_{i=1}^{n}\sum_{j=1}^{p} a_{i,j}^2 = 0$.
Or la somme de quantités positives n'est nulle que si chacune des quantité est nulle et donc $a_{i,j} = 0$ pour tous i, j. Ainsi $A = O_{n,p}$.
Finalement φ est un produit scalaire sur E.
Ce produit scalaire est appelé produit scalaire canonique sur $\mathcal{M}_{n,p}(\mathbb{R})$.

Remarque En fait, pour $A = (a_{i,j})$ et $B = (b_{i,j})$,

$$\varphi(A, B) = \sum_{i=1}^{n}\sum_{j=1}^{p} a_{i,j} b_{i,j}$$

Remarque Le produit scalaire de deux vecteurs x et y est souvent noté $(x \mid y)$.
On trouve aussi les notations $x.y$, $\langle x, y\rangle$,...

Désormais $(. \mid .)$ désigne un produit scalaire sur l'espace vectoriel réel E.
Par définition on a les propriétés :

$$\forall x, y \in E, (x \mid y) = (y \mid x)$$

$$\forall x \in E, (x \mid x) \geqslant 0 \text{ et } (x \mid x) = 0 \Rightarrow x = 0$$

De plus par linéarité en l'une des variables, on a $(0_E \mid 0_E) = 0$ et donc $(x \mid x) = 0 \Leftrightarrow x = 0_E$
Enfin par bilinéarité : $\forall x, x', y, y' \in E, \forall \lambda, \lambda', \mu, \mu' \in \mathbb{R}$

$$(\lambda x + \lambda' x' \mid \mu y + \mu' y') = \lambda\mu(x \mid y) + \lambda\mu'(x \mid y') + \lambda'\mu(x' \mid y) + \lambda'\mu'(x' \mid y')$$

12.1.2 Notions métriques

12.1.2.1 Inégalité de Cauchy-Schwarz

Théorème

$$\forall x, y \in E, (x \mid y)^2 \leqslant (x \mid x)(y \mid y)$$

De plus, il y a égalité si, et seulement si, x et y sont colinéaires.

dém. :
Si $x = 0_E$, la propriété est immédiate avec égalité et colinéarité des vecteurs x et y.
Si $x \neq 0_E$, considérons pour $\lambda \in \mathbb{R}$, le produit scalaire $(\lambda x + y \mid \lambda x + y)$.
D'une part, par positivité, $(\lambda x + y \mid \lambda x + y) \geqslant 0$.
D'autre part, par bilinéarité et symétrie,

$$(\lambda x + y \mid \lambda x + y) = \lambda^2 (x \mid x) + 2\lambda (x \mid y) + (y \mid y)$$

En posant $a = (x \mid x)$ ($a \neq 0$ car $x \neq 0$), $b = 2(x \mid y)$ et $c = (y \mid y)$, le trinôme $a\lambda^2 + b\lambda + c$ est de signe constant donc de discriminant négatif. Ainsi $b^2 - 4ac \leqslant 0$ ce qui donne

$$(x \mid y)^2 \leqslant (x \mid x)(y \mid y)$$

De plus, s'il y a égalité dans cette inégalité, le discriminant est nul et donc il existe $\lambda \in \mathbb{R}$ tel que

$$(\lambda x + y \mid \lambda x + y) = 0$$

ce qui entraîne $\lambda x + y = 0_E$. La famille (x, y) est alors liée.
Inversement, supposons la famille (x, y) liée. Sachant $x \neq 0$, on peut écrire $y = \mu x$ et vérifier par bilinéarité du produit scalaire l'égalité $(x \mid y)^2 = (x \mid x)(y \mid y)$.
□

Corollaire

Pour $x_1, \ldots, x_n, y_1, \ldots, y_n \in \mathbb{R}$ on a

$$\left(\sum_{i=1}^n x_i y_i\right)^2 \leqslant \left(\sum_{i=1}^n x_i^2\right)\left(\sum_{i=1}^n y_i^2\right)$$

dém. :
Il suffit d'appliquer l'inégalité de Cauchy-Schwarz au produit scalaire canonique sur

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

□

12.1.2.2 Norme euclidienne

Définition

On appelle norme euclidienne (ou longueur) d'un vecteur $\mathcal{B}$ le réel

$$\|x\| = \sqrt{(x \mid x)}$$

Exemple La norme euclidienne associée au produit scalaire du plan ou de l'espace géométrique est la norme usuelle.

Exemple Sur $E = \mathbb{R}^n$ muni du produit scalaire canonique, pour $x = (x_1, \dots, x_n)$, on a

$$\|x\| = \sqrt{x_1^2 + \cdots + x_n^2}$$

En particulier, pour $n = 1$, $\|x\| = \sqrt{x^2} = |x|$.

Exemple Sur $E = \mathbb{C}$ muni du produit scalaire canonique, pour $z \in \mathbb{C}$ on a

$$\|z\| = |z|$$

Proposition

$\forall x, y \in E, \forall \lambda \in \mathbb{R}$,

$$\|x\| \geqslant 0,\ \|x\| = 0 \Leftrightarrow x = 0,\ \|\lambda.x\| = |\lambda|\,\|x\|$$

et

$$|(x \mid y)| \leqslant \|x\| \,.\, \|y\|$$

avec égalité si, et seulement si, x et y sont colinéaires.

dém. :
$\|x\| = \sqrt{(x \mid x)} \geqslant 0$.
$\|x\| = 0 \Leftrightarrow \sqrt{(x \mid x)} = 0 \Leftrightarrow (x \mid x) = 0 \Leftrightarrow x = 0_E$
$\|\lambda.x\| = \sqrt{(\lambda x \mid \lambda x)} = \sqrt{\lambda^2 (x \mid x)} = |\lambda| \sqrt{(x \mid x)} = |\lambda|\,\|x\|$
et enfin $|(x \mid y)| \leqslant \|x\| \,.\, \|y\| \Leftrightarrow (x \mid y)^2 \leqslant \|x\|^2 \|y\|^2$ qui est l'inégalité de Cauchy-Schwarz.
□

Exemple $\|-x\| = \|x\|$. En effet $\|-x\| = \|(-1).x\| = |-1|\,\|x\| = \|x\|$.

Théorème

$$\forall x, y \in E, \|x + y\| \leqslant \|x\| + \|y\|$$

De plus, il y a égalité si, et seulement si, x et y colinéaires et $(x \mid y) \geqslant 0$ (ce qui signifie que x et y ont même direction et même sens)

dém. :
Soient $x, y \in E$.

$$\|x + y\|^2 = (x + y \mid x + y)$$

Par bilinéarité et symétrie du produit scalaire

$$\|x + y\|^2 = (x \mid x) + 2(x \mid y) + (y \mid y) = \|x\|^2 + 2(x \mid y) + \|y\|^2$$

Par l'inégalité de Cauchy-Schwarz

$$\|x+y\|^2 \leqslant \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\|+\|y\|)^2$$

et on en déduit l'inégalité triangulaire.
De plus, il y a égalité si, et seulement si, $(x \mid y) = \|x\|\|y\|$ ce qui signifie

$$(x \mid y) \geqslant 0 \text{ et } |(x \mid y)| = \|x\|\|y\|$$

c'est-à-dire que les vecteurs x et y sont colinéaire et de même sens.
□
Corollaire

$$\forall x, y \in E, |\|x\| - \|y\|| \leqslant \|x-y\| \leqslant \|x\| + \|y\|$$

dém. :
Soient $x, y \in E$.

$$\|x\| = \|(x-y)+y\| \leqslant \|x-y\| + \|y\|$$

donc

$$\|x\| - \|y\| \leqslant \|x-y\|$$

De façon symétrique, on a aussi

$$\|y\| - \|x\| \leqslant \|y-x\| = \|x-y\|$$

□
Proposition

$\forall x, y \in E$, $\|x+y\|^2 = \|x\|^2 + 2(x \mid y) + \|y\|^2$
$\forall x, y \in E$, $\|x-y\|^2 = \|x\|^2 - 2(x \mid y) + \|y\|^2$
$\forall x, y \in E$, $(x+y \mid x-y) = \|x\|^2 - \|y\|^2$.

dém. :
Soient $x, y \in E$.

$$\|x+y\|^2 = (x+y \mid x+y)$$

Par bilinéarité du produit scalaire.

$$\|x+y\|^2 = (x \mid x+y) + (y \mid x+y) = (x \mid x) + (x \mid y) + (y \mid x) + (y \mid y)$$

Par symétrie du produit scalaire

$$\|x+y\|^2 = \|x\|^2 + 2(x \mid y) + \|y\|^2$$

En changeant y en $-y$, on obtient

$$\|x-y\|^2 = \|x\|^2 + 2(x \mid -y) + \|-y\|^2 = \|x\|^2 - 2(x \mid y) + \|y\|^2$$

Enfin, on a encore par bilinéarité du produit scalaire

$$(x+y \mid x-y) = (x \mid x) - (x \mid y) + (y \mid x) - (y \mid y)$$

puis par symétrie du produit scalaire

$$(x + y \mid x - y) = \|x\|^2 - \|y\|^2$$

□

Proposition

$\forall x, y \in E, \|x + y\|^2 + \|x - y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right)$

dém. :
Il suffit de sommer les deux premières identités remarquables précédentes
□

Remarque L'identité du parallélogramme signifie que, dans un parallélogramme donné, la somme des carrés des longueurs des diagonales est égale à la somme des carrés des longueurs des côtés.

Proposition

$\forall x, y \in E, \|x + y\|^2 - \|x - y\|^2 = 4(x \mid y)$

dém. :
Il suffit de faire la différence des deux premières identités remarquables précédentes
□

Remarque Par l'identité de polarisation, on peut calculer un produit scalaire à partir de la norme euclidienne associée.

12.1.2.3 Distance euclidienne

Définition

On appelle distance euclidienne séparant deux vecteurs x et y de E le réel

$$d(x, y) = \|y - x\|$$

Exemple Dans $E = \mathbb{R}^n$ muni du produit scalaire canonique, on a, pour $x = (x_1, \ldots, x_n)$ et $y = (y_1, \ldots, y_n)$,

$$d(x, y) = \sqrt{\sum_{k=1}^{n} (y_k - x_k)^2}$$

Proposition

$\forall x, y \in E, d(x, y) = 0 \Leftrightarrow x = y$ [séparation]
$\forall x, y \in E, d(x, y) = d(y, x)$ [symétrie]
$\forall x, y, z \in E, d(x, z) \leqslant d(x, y) + d(y, z)$ [inégalité triangulaire]
$\forall x, y, z \in E, d(x + z, y + z) = d(x, y)$ [invariance par translation]

dém. :
Soient $x, y, z \in E$.
$d(x, y) = 0 \Leftrightarrow \|y - x\| = 0 \Leftrightarrow y - x = 0_E \Leftrightarrow y = x$.
$d(y, x) = \|x - y\| = \|y - x\| = d(x, y)$.
$d(x, z) = \|z - x\| = \|(z - y) + (y - x)\| \leqslant \|z - y\| + \|y - x\| = d(y, z) + d(x, y)$,
$d(x + z, y + z) = \|(y + z) - (x + z)\| = \|y - x\| = d(x, y)$.
□

12.1.2.4 Écart angulaire formé par deux vecteurs non nuls

Soient u et v deux vecteurs non nuls de E.
Par l'inégalité de Cauchy Schwarz,

$$\frac{(u \mid v)}{\|u\| \|v\|} \in [-1, 1]$$

et donc il existe un unique $\theta \in [0, \pi]$ vérifiant

$$(u \mid v) = \|u\| \|v\| \cos\theta$$

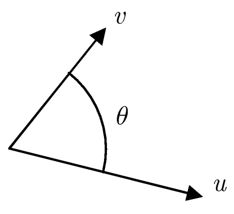
Définition

Le réel θ ainsi défini est appelé écart angulaire entre les vecteurs u et v.
On note

$$\theta = \text{Ecart}(u, v)$$

Remarque Nous avons ici une définition rigoureuse de la notion d'écart angulaire, définition réalisée à partir de la notion de produit scalaire. Dans le cas du plan ou de l'espace géométrique, on retrouve la notion d'écart angulaire usuelle. Cette notion angulaire est définie en dimension quelconque : 2,3 voire supérieure…

Remarque Cette notion n'est à pas confondre avec la notion d'angle orientée qui sera introduite ultérieurement, dans le cadre des plans euclidiens orientés.
Un écart angulaire entre deux vecteurs est figuré par un arc de cercle (non fléché) d'amplitude inférieure à π.



Proposition

Ecart$(u, v) =$ Ecart(v, u)
Ecart$(-u, v) = \pi -$ Ecart(u, v)
Ecart$(u, v) = 0$ si, et seulement si, u et v ont même direction et même sens.
Ecart$(u, v) = \pi$ si, et seulement si, u et v ont même direction et sont de sens opposés.

dém. :
Ecart$(u, v) =$ Ecart(v, u) car $(u \mid v) = (v \mid u)$.
Ecart$(-u, v) = \pi -$ Ecart(u, v) car $(-u \mid v) = -(u \mid v)$ et $\arccos(-x) = \pi - \arccos(x)$.
Ecart$(u, v) = 0$ si, et seulement si, $(u \mid v) = \|u\| \|v\|$ ce qui signifie $(u \mid v) \geqslant 0$ et $|(u \mid v)| = \|u\| \|v\|$.
En vertu du cas d'égalité dans l'inégalité de Cauchy-Schwarz, cela revient à u et v colinéaire et de même sens.
Ecart$(u, v) = \pi$ si, et seulement si, $(u \mid v) = -\|u\| \|v\|$ ce qui signifie $(u \mid v) \leqslant 0$ et $|(u \mid v)| = \|u\| \|v\|$ et on peut conclure comme ci-dessus.
□

Définition

On dit que les vecteurs u et v forment un angle aigu si Ecart$(u, v) \in [0, \pi/2]$.
On dit que les vecteurs u et v forment un angle obtus si Ecart$(u, v) \in [\pi/2, \pi]$.
On dit que les vecteurs u et v forment un angle droit si Ecart$(u, v) = \pi/2$.

12.1.3 Vecteurs orthogonaux

12.1.3.1 Famille orthogonale

Définition

Deux vecteurs x et y de E sont dits orthogonaux si

$$(x \mid y) = 0$$

Exemple Le vecteur nul est orthogonal à tout vecteur de l'espace car

$$\forall x \in E, (x \mid 0_E) = 0$$

Inversement, puisqu'un vecteur orthogonal à tout vecteur est orthogonal à lui-même et puisque $(x \mid x) = 0 \Rightarrow x = 0_E$, on peut affirmer que le vecteur nul est le seul vecteur orthogonal à tout vecteur.

Définition

On dit qu'une famille $\mathcal{F} = (e_1, \dots, e_n)$ de vecteurs de E est orthogonale si $\mathcal{F}$ est constituée de vecteurs deux à deux orthogonaux i.e.

$$\forall 1 \leqslant i \neq j \leqslant n, (e_i \mid e_j) = 0$$

Exemple Dans $\mathbb{R}^n$ muni du produit scalaire canonique, la base canonique est une famille orthogonale.

Théorème

Si $\mathcal{F} = (e_1, \dots, e_n)$ une famille orthogonale alors

$$\|e_1 + \dots + e_n\|^2 = \|e_1\|^2 + \dots + \|e_n\|^2$$

dém. :
Par récurrence sur $n \in \mathbb{N}^\star$.
Pour $n = 1$: ok
Supposons la propriété établie au rang $n \geqslant 1$.
Soit $(e_1, \dots, e_n, e_{n+1})$ une famille orthogonale.
En exploitant

$$\|a + b\|^2 = \|a\|^2 + 2(a \mid b) + \|b\|^2$$

avec $a = e_1 + \cdots + e_n$ et $b = e_{n+1}$, on obtient

$$\|e_1 + \cdots + e_n + e_{n+1}\|^2 = \|e_1 + \cdots + e_n\|^2 + \|e_{n+1}\|^2$$

car $(a \mid b) = 0$.
Par hypothèse de récurrence

$$\|e_1 + \cdots + e_n + e_{n+1}\|^2 = \|e_1\|^2 + \cdots + \|e_n\|^2 + \|e_{n+1}\|^2$$

Récurrence établie.
□

Corollaire

Toute famille orthogonale ne comportant pas le vecteur nul est libre.

dém. :
Soit $(e_1, \dots, e_n)$ une famille orthogonale ne comportant pas le vecteur nul.
Supposons $\lambda_1 e_1 + \cdots + \lambda_n e_n = 0_E$.
Puisque la famille $(\lambda_1 e_1, \dots, \lambda_n e_n)$ est elle aussi orthogonale, la formule de Pythagore donne

$$\|\lambda_1 e_1 + \cdots + \lambda_n e_n\|^2 = \|\lambda_1 e_1\|^2 + \cdots + \|\lambda_n e_n\|^2$$

Or $\|\lambda_1 e_1 + \cdots + \lambda_n e_n\|^2 = 0$ donc, par somme nulle de quantités positives, on a

$$\forall 1 \leqslant k \leqslant n, \|\lambda_k e_k\|^2 = 0$$

puis

$$\forall 1 \leqslant k \leqslant n, \lambda_k e_k = 0_E$$

Or les vecteurs $e_1, \dots, e_n$ sont non nuls, donc

$$\forall 1 \leqslant k \leqslant n, \lambda_k = 0$$

□

12.1.3.2 Famille orthonormée

Définition

On dit qu'un vecteur x de E est unitaire (ou normé) si $\|x\| = 1$.

Exemple La base canonique de $\mathbb{R}^n$ est constituée de vecteurs unitaires.

Définition

On dit qu'une famille $\mathcal{F} = (e_1, \ldots, e_n)$ de vecteurs de E orthonormée si $\mathcal{F}$ est constituée de vecteurs unitaires deux à deux orthogonaux i.e.

$$\forall 1 \leqslant i, j \leqslant n, (e_i \mid e_j) = \delta_{i,j}$$

Exemple La base canonique de $\mathbb{R}^n$ est une famille orthonormée.

Définition

Normer un vecteur non nul x, c'est considérer le vecteur unitaire $u = x/\|x\|$.

Exemple Si $\mathcal{F} = (e_1, \ldots, e_n)$ est une famille orthogonale ne comportant pas le vecteur nul alors, en normant chacun de ses vecteurs, on forme une famille orthonormée $\mathcal{G} = (e_1/\|e_1\|, \ldots, e_n/\|e_n\|)$.

Proposition

Toute famille orthonormée est libre.

dém. :
Car une famille orthonormée et est orthogonale et ne comporte pas le vecteur nul.
□

12.1.3.3 Procédé d'orthonormalisation de Schmidt

Théorème

Soit $\mathcal{F} = (e_1, \ldots, e_n)$ une famille libre de vecteurs de E.
On peut construire une famille orthonormée $(v_1, \ldots, v_n)$ de vecteurs de E vérifiant

$$\forall 1 \leqslant k \leqslant n, \text{Vect}(e_1, \ldots, e_k) = \text{Vect}(v_1, \ldots, v_k)$$

dém. :
Par récurrence sur $n \in \mathbb{N}^\star$.
Pour $n = 1$.
Soit (e_1) une famille libre.
On a $e_1 \neq 0$. Le vecteur $v_1 = e_1/\|e_1\|$ convient
Supposons la propriété établie au rang $n \in \mathbb{N}^\star$.
Soit $(e_1, ..., e_n, e_{n+1})$ une famille libre.
Par hypothèse de récurrence, il existe $(v_1, \ldots, v_n)$ famille orthonormée telle que

$$\forall 1 \leqslant k \leqslant n, \text{Vect}(e_1, \ldots, e_k) = \text{Vect}(v_1, \ldots, v_k)$$

Il reste à déterminer un vecteur v_{n+1} de sorte que la famille $(v_1, \ldots, v_{n+1})$ est orthonormée et

$$\text{Vect}(e_1, \ldots, e_n, e_{n+1}) = \text{Vect}(v_1, \ldots, v_n, v_{n+1})$$

Analyse :
Si v_{n+1} convient alors $v_{n+1} \in \text{Vect}(e_1, \dots, e_{n+1})$ et on peut donc écrire

$$v_{n+1} = \lambda_1 e_1 + \cdots + \lambda_n e_n + \lambda e_{n+1}$$

On peut aussi écrire

$$v_{n+1} = \mu_1 v_1 + \cdots + \mu_n v_n + \lambda e_n$$

car $\text{Vect}(e_1, \dots, e_n) = \text{Vect}(v_1, \dots, v_n)$.
Puisque la famille $(v_1, \dots, v_{n+1})$ est orthonormée, on a

$$\forall 1 \leqslant i \leqslant n, (v_i \mid v_{n+1}) = 0$$

On en déduit

$$\forall 1 \leqslant i \leqslant n, \mu_i = -\lambda(v_i \mid e_{n+1})$$

et par suite

$$v_{n+1} = \lambda \left(e_{n+1} - \sum_{i=1}^{n} (v_i \mid e_{n+1}) v_i \right) = \lambda w$$

avec

$$w = e_{n+1} - \sum_{i=1}^{n} (v_i \mid e_{n+1})$$

Enfin $\|v_{n+1}\| = 1$ donne $|\lambda| \|w\| = 1$ d'où $|\lambda| = 1/\|w\|$.
Synthèse :
Posons

$$w = e_{n+1} - \sum_{i=1}^{n} (e_{n+1} \mid v_i) v_i$$

Le vecteur w est non nul car la famille $(e_1, \dots, e_n, e_{n+1})$ est supposé libre et donc le vecteur e_{n+1} n'est pas combinaison linéaire des vecteurs $e_1, \dots, e_n$.
Posons $v_{n+1} = w/\|w\|$.
Par construction, on vérifie aisément que la famille a $(v_1, \dots, v_{n+1})$ est orthonormée.
De plus, $\text{Vect}(v_1, \dots, v_n, v_{n+1}) \subset \text{Vect}(e_1, \dots, e_n, e_{n+1})$ et par égalité des dimensions (les deux familles considérées étant libres), on a $\text{Vect}(v_1, \dots, v_n, v_{n+1}) = \text{Vect}(e_1, \dots, e_n, e_{n+1})$
Récurrence établie.
□

Remarque Dans la pratique pour orthonormaliser une famille libre $(e_1, \dots, e_n)$:
- on pose $V_1 = e_1$;
- pour $1 \leqslant p \leqslant n-1$, lorsque $V_1, \dots, V_p$ sont trouvés, on cherche V_{p+1} de la forme :
$V_{p+1} = e_{p+1} + \lambda_1 V_1 + \cdots + \lambda_p V_p$ de sorte que $\forall 1 \leqslant k \leqslant p, (V_{p+1} \mid V_k) = 0$ ce qui fournit la valeur de λ_k.
- une fois obtenue la famille $(V_1, \dots, V_n)$, on normalise chaque vecteur en posant $v_k = V_k/\|V_k\|$.
La famille ainsi formée est appelée famille orthonormalisée de $\mathcal{F} = (e_1, \dots, e_n)$ selon le procédé de Schmidt.

Exemple Dans $\mathbb{R}^3$ muni du produit scalaire canonique.
Soient $e_1 = (0, 1, 1), e_2 = (1, 0, 1), e_3 = (1, 1, 0)$.

La famille (e_1, e_2, e_3) est libre puisque

$$\begin{vmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{vmatrix} = 2$$

Formons son orthonormalisée selon le procédé de Schmidt.
$V_1 = e_1 = (0, 1, 1)$
$V_2 = e_2 + \lambda V_1$
$(V_2 \mid V_1) = 0$ donne $\lambda = -1/2$ puis $V_2 = (1, -1/2, 1/2)$.
$V_3 = e_3 + \lambda V_1 + \mu V_2$
$(V_3 \mid V_1) = 0$ donne $\lambda = -1/2$,
$(V_3 \mid V_2) = 0$ donne $\mu = -1/3$ puis $V_3 = (2/3, 2/3, -2/3)$.
Enfin

$$v_1 = (0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}), v_2 = (\frac{2}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}), v_3 = (\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}})$$

12.1.4 Orthogonal d'une partie

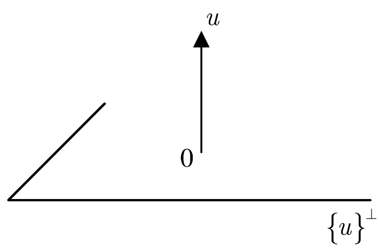
Définition

On appelle orthogonal d'une partie A de E, l'ensemble noté $A^\perp$ constitué des vecteurs orthogonaux a tout vecteur de A i.e.

$$A^\perp = \{x \in E / \forall a \in A, (a \mid x) = 0\}$$

Exemple $\{0_E\}^\perp = E$ et $E^\perp = \{0_E\}$.

Exemple



Proposition

Soit A une partie de E.
$A^\perp$ est un sous-espace vectoriel de E

dém. :
$A^\perp \subset E$.
$0_E \in A^\perp$ car 0_E est orthogonal à tout vecteur de E et en particulier à tout vecteur de A.

Soient $\lambda, \mu \in \mathbb{R}$ et $x, y \in A^{\perp}$.
Pour tout $a \in A$, $(a \mid \lambda x + \mu y) = \lambda(a \mid x) + \mu(a \mid y) = 0$.
Ainsi $\lambda x + \mu y \in A^{\perp}$.
□

Proposition

Soient A et B deux parties de E.
1) $A \subset A^{\perp\perp}$;
2) $A \subset B \Rightarrow B^{\perp} \subset A^{\perp}$;
3) $A^{\perp} = \text{Vect}(A)^{\perp}$.

dém. :
1) Soit $a \in A$. Pour tout $x \in A^{\perp}$, on a $(x \mid a) = (a \mid x) = 0$ donc $a \in A^{\perp\perp}$. Ainsi $A \subset A^{\perp\perp}$.
2) Supposons $A \subset B$.
Soit $x \in B^{\perp}$. Pour tout $a \in A$, $(a \mid x) = 0$ car $a \in B$ et $x \in B^{\perp}$. Ainsi $B^{\perp} \subset A^{\perp}$.
3) Enfin $A \subset \text{Vect}(A)$ donc $\text{Vect}(A)^{\perp} \subset A^{\perp}$.
Inversement $A \subset A^{\perp\perp}$ donc Vect$A \subset A^{\perp\perp}$ (car $A^{\perp\perp}$ est un sous-espace vectoriel) puis $A^{\perp} \subset A^{\perp\perp\perp} \subset \text{Vect}(A)^{\perp}$.
□

12.1.5 Sous-espaces vectoriels orthogonaux

Définition

Deux sous-espaces vectoriels F et G de E sont dits orthogonaux si

$$\forall (x, y) \in F \times G, (x \mid y) = 0$$

Exemple Si F est un sous-espace vectoriel de E alors F et $F^{\perp}$ sont des sous-espaces vectoriels orthogonaux.

Proposition

Soient F et G deux sous-espaces vectoriels de E.
Si F et G sont orthogonaux alors $F \cap G = \{0_E\}$.

dém. :
Soit $x \in F \cap G$.
Puisque F et G sont orthogonaux, $(x \mid x) = 0$ et donc $x = 0_E$.
Ainsi $F \cap G \subset \{0_E\}$ puis l'égalité car F et G sont des sous-espaces vectoriels et donc contiennent le vecteur nul.
□

Proposition

Soient F et G deux sous-espaces vectoriels de E.
On a équivalence entre :
(i) F et G sont orthogonaux ;
(ii) $F \subset G^{\perp}$;
(iii) $G \subset F^{\perp}$.

dém. :
(i) $\Rightarrow$ (ii) Supposons F et G orthogonaux.
Soit $x \in F$. Pour tout $y \in G$, on a $(y \mid x) = 0$ donc $x \in G^{\perp}$.
Ainsi $F \subset G^{\perp}$.
(ii) $\Rightarrow$ (iii) Supposons $F \subset G^{\perp}$.
On a alors $G^{\perp\perp} \subset F^{\perp}$ or $G \subset G^{\perp\perp}$ donc $G \subset F^{\perp}$.
(iii) $\Rightarrow$ (i) Supposons $G \subset F^{\perp}$.
Soient $x \in F$ et $y \in G$. On a $(x \mid y) = 0$ car $x \in F$ et $y \in F^{\perp}$.
Ainsi les espaces F et G sont orthogonaux.
□

12.2 Espaces vectoriels euclidiens

12.2.1 Définition

Définition

On appelle espace vectoriel euclidien tout $\mathbb{R}$-espace vectoriel de dimension finie muni d'un produit scalaire.

Exemple L'ensemble des vecteurs du plan (resp. de l'espace) géométrique est un espace vectoriel euclidien.

Exemple $\mathbb{R}^n$ muni du produit scalaire canonique est un espace vectoriel euclidien. On dit alors que $\mathbb{R}^n$ est muni de sa structure vectorielle euclidienne canonique.

Exemple Si F est un sous-espace vectoriel d'un espace vectoriel euclidien E alors F est un espace vectoriel euclidien lorsqu'on le munit de la restriction du produit scalaire existant sur E.

Désormais, E désigne un espace vectoriel euclidien.

12.2.2 Base orthonormée

Définition

On appelle base orthonormée d'un espace vectoriel euclidien E toute base de E constituée de vecteurs unitaires deux à deux orthogonaux.

Exemple La base canonique de $\mathbb{R}^n$ est une base orthonormée de $\mathbb{R}^n$.

Théorème

Tout espace vectoriel euclidien E possède une base orthonormée.

dém. :
Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E.
La famille $\mathcal{B}$ est libre, on peut donc l'orthonormalisée selon le procédé de Schmidt et la famille obtenue

est une famille orthonormée, donc aussi libre, formée de $n = \dim E$ vecteurs de E, c'est donc une base orthonormée de E.
□

Corollaire

Tout sous-espace vectoriel d'un espace euclidien possède une base orthonormée.

dém. :
Un sous-espace vectoriel d'un espace euclidien est un espace euclidien pour le produit scalaire défini par restriction.
□

12.2.3 Composantes dans une base orthonormée

Soit E un espace vectoriel euclidien muni d'une base orthonormée $\mathcal{B} = (e_1, \dots, e_n)$

Théorème

Pour tout $x \in E$, on a

$$x = \sum_{k=1}^{n} (e_k \mid x) e_k$$

Par suite les composantes de x dans $\mathcal{B}$ sont les

$$(e_1 \mid x), \dots, (e_n \mid x)$$

dém. :
Puisque $\mathcal{B}$ est une base de E, on peut écrire $x = \sum_{i=1}^{n} \lambda_i e_i$ et par bilinéarité du produit scalaire,

$$(e_k \mid x) = \sum_{i=1}^{n} \lambda_i (e_k \mid e_i) = \sum_{i=1}^{n} \lambda_i \delta_{k,i} = \lambda_k$$

□

Exemple Si $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{R})$ est la matrice dans la base $\mathcal{B}$ d'un endomorphisme f de E alors pour tout $i, j \in \{1, \dots, n\}$, $a_{i,j} = (e_i \mid f(e_j))$.
En effet $a_{i,j}$ est la i-ème composante dans la base $\mathcal{B}$ de l'image du j-ème vecteur de base.

Théorème

Si x et y sont de vecteurs de E de composantes respectives $x_1, \dots, x_n$ et $y_1, \dots, y_n$ dans la base orthonormée $\mathcal{B}$ alors

$$(x \mid y) = x_1 y_1 + \dots + x_n y_n \text{ et } \|x\| = \sqrt{x_1^2 + \dots + x_n^2}$$

dém. :
On a

$$x = \sum_{k=1}^{n} x_k e_k \text{ et } y = \sum_{\ell=1}^{n} y_\ell e_\ell$$

Par bilinéarité du produit scalaire

$$(x \mid y) = \sum_{k=1}^{n}\sum_{\ell=1}^{n} x_k y_\ell (e_k \mid e_\ell) = \sum_{k=1}^{n}\sum_{\ell=1}^{n} x_k y_\ell \delta_{k,\ell} = \sum_{k=1}^{n} x_k y_k$$

En particulier

$\|x\|^2 = (x \mid x) = \sum_{k=1}^{n} x_k^2.$

□

Remarque Si on introduit les matrices composantes $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ des vecteurs x, y, on peut écrire

$$(x \mid y) = {}^tXY \text{ et } \|x\|^2 = {}^tXX$$

Remarque L'application $\varphi : E \to \mathbb{R}^n$ définie par $\varphi(x) = (x_1, \ldots, x_n)$ est un isomorphisme de $\mathbb{R}$-espace vectoriel et

$$\forall x, y \in E, (x \mid y)_E = (\varphi(x) \mid \varphi(y))_{\mathbb{R}^n}$$

Ainsi E muni d'une base orthonormée devient semblable à $\mathbb{R}^n$ muni du produit scalaire canonique.

12.2.4 Supplémentaire orthogonal

Théorème

Si F un sous-espace vectoriel d'un espace vectoriel euclidien E alors les espaces F et $F^\perp$ sont supplémentaires dans E.

dém. :
Soit $(e_1, \ldots, e_m)$ une base orthonormée de F.
Analyse :
Soit $x \in E$. Supposons $x = a + b$ avec $a \in F$ et $b \in G$.
On a $a = \sum_{k=1}^{m} (e_k \mid a)e_k$ car $(e_1, \ldots, e_m)$ est une base orthonormée de F.
Or $(e_k \mid a) = (e_k \mid x) - (e_k \mid b) = (e_k \mid x)$ car $e_k \in F$ et $b \in F^\perp$.
On en déduit $a = \sum_{k=1}^{m} (e_k \mid x)e_k$ et $b = x - a$ ce qui détermine ces vecteurs de façon unique.
Synthèse
Soit $x \in E$.
Posons $a = \sum_{k=1}^{m} (e_k \mid x)e_k$ et $b = x - a$.
On a évidemment $a \in F$ et $x = a + b$.
Il reste à vérifier $b \in F^\perp$.
Pour tout $1 \leqslant k \leqslant m$, $(e_k \mid b) = (e_k \mid x) - (e_k \mid a)$.
Or les $(e_k \mid a)$ sont les composantes de a dans la base orthonormée $(e_1, \ldots, e_m)$ et, par construction,

celles-ci valent $(e_k \mid x)$.
Ainsi $(e_k \mid b) = 0$ et don c $b \in \{e_1, \dots, e_m\}^\perp = \text{Vect}(e_1, \dots, e_m)^\perp = F^\perp$.
$F = \text{Vect}(e_1, \dots, e_m)$ et $(e_k \mid b) = (e_k \mid x) - (e_k \mid a) = 0$ donc $b \in F^\perp$.

□

Définition

L'espace $F^\perp$ est appelé supplémentaire orthogonal de F.

Corollaire

Toute famille orthonormée de vecteurs de E peut être complétée en une base orthonormée de E.

dém. :
Soit $(e_1, \dots, e_m)$ une famille orthonormée de E.
Posons $F = \text{Vect}(e_1, \dots, e_m)$ et considérons $(e_{m+1}, \dots, e_n)$ une base orthonormée de l'espace $F^\perp$.
Puisque F et $F^\perp$ sont supplémentaires, la famille $(e_1, \dots, e_m, e_{m+1}, \dots, e_n)$ est une base de E et on vérifie aisément qu'elle est orthonormée.

□

Corollaire

$\dim F^\perp = \dim E - \dim F$ et $F^{\perp\perp} = F$.

dém. :
Puisque F et $F^\perp$ sont supplémentaires dans E, on a $\dim F + \dim F^\perp = \dim E$.
On en déduit $\dim F^\perp = \dim E - \dim F$ puis $\dim F^{\perp\perp} = \dim F$. Or $F \subset F^{\perp\perp}$ donc $F = F^{\perp\perp}$

□

Proposition

Si F et G sont deux sous-espaces vectoriels d'un espace vectoriel euclidien E alors

$$(F+G)^\perp = F^\perp \cap G^\perp \text{ et } (F \cap G)^\perp = F^\perp + G^\perp$$

dém. :
Puisque $F \subset F+G$ et $G \subset F+G$, on a $(F+G)^\perp \subset F^\perp$, $(F+G)^\perp \subset G^\perp$ et donc $(F+G)^\perp \subset F^\perp \cap G^\perp$.
Inversement, soit $x \in F^\perp \cap G^\perp$. Pour tout $y \in F+G$, on peut écrire $y = a+b$ avec $a \in F$ et $b \in G$ et alors $(x \mid y) = (x \mid a) + (x \mid b) = 0$. Ainsi $F^\perp \cap G^\perp \subset (F+G)^\perp$ puis l'égalité $(F+G)^\perp = F^\perp \cap G^\perp$.
En appliquant cette relation à $F^\perp$ et $G^\perp$ au lieu de F et G, on obtient $(F^\perp + G^\perp)^\perp = F \cap G$ puis en passant à l'orthogonal on obtient $F^\perp + G^\perp = (F \cap G)^\perp$.

□

12.2.5 Vecteur normal à un hyperplan

E désigne un espace vectoriel euclidien de dimension $n \in \mathbb{N}^\star$.

Définition

Si H est un hyperplan de E alors $H^\perp$ est une droite vectorielle appelée droite normale à H.
Tout vecteur non nul de celle-ci est appelé vecteur normal à H.

Remarque Si a est vecteur normal à H alors $H^\perp = \text{Vect}(a)$ et donc $H = \text{Vect}(a)^\perp = \{a\}^\perp$.

Proposition

Soient $\mathcal{B} = (e_1, \dots, e_n)$ une base orthonormée de E et H un hyperplan de E.
Un vecteur a de composantes $a_1, \dots, a_n$ est normal à H si, et seulement si, $a_1x_1 + \dots + a_nx_n = 0$ est une équation de H dans $\mathcal{B}$ (en notant $x_1, \dots, x_n$ les composantes dans $\mathcal{B}$ d'un vecteur générique $x \in E$).

dém. :
Si n est un vecteur normal à H de composantes $a_1, \dots, a_n$ dans $\mathcal{B}$ alors

$$x \in H \Leftrightarrow (a \mid x) = 0 \Leftrightarrow a_1x_1 + \dots + a_nx_n = 0$$

Ainsi $a_1x_1 + \dots + a_nx_n = 0$ est une équation de H dans $\mathcal{B}$.
Inversement
Supposons que $a_1x_1 + \dots + a_nx_n = 0$ est une équation de H dans $\mathcal{B}$ (avec $(a_1, \dots, a_n) \neq (0, \dots, 0)$).
Soit n le vecteur de composantes $a_1, \dots, a_n$.
Le vecteur n est non nul et pour tout $x \in H$, $(n \mid x) = a_1x_1 + \dots + a_nx_n = 0$.
Ainsi n est un vecteur normal à H.
□

Exemple Dans $\mathbb{R}^3$ muni de sa structure canonique, le plan de vecteur normal $n = (1, 2, -1)$ a pour équation $x + 2y - z = 0$.

12.2.6 Représentation d'une forme linéaire

Soit E un espace vectoriel euclidien.
Pour $a \in E$, on note $\varphi_a : E \to \mathbb{R}$ l'application déterminée par

$$\forall x \in E, \varphi_a(x) = (a \mid x)$$

Proposition

φ_a est une forme linéaire sur E de noyau $\{a\}^\perp$.

dém. :
φ_a est une forme linéaire sur E car $\varphi_a : E \to \mathbb{R}$ et, avec des notations immédiates

$$\varphi_a(\lambda x + \mu y) = (a \mid \lambda x + \mu y) = \lambda(a \mid x) + \mu(a \mid y) = \lambda\varphi_a(x) + \mu\varphi_a(y)$$

Aussi

$$x \in \ker \varphi_a \Leftrightarrow (a \mid x) = 0 \Leftrightarrow x \in \{a\}^\perp$$

□

Théorème

$\forall f \in E^\star, \exists! a \in E, f = \varphi_a$.

dém. :
Considérons l'application $\Phi : E \to E^\star$ définie par $\Phi(a) = \varphi_a$.
Montrons que Φ est un isomorphisme de $\mathbb{R}$-espace vectoriel ce qui permettra de conclure.
Pour commencer, notons que E et $E^\star$ sont des espaces de même dimension finie.

Avec des notation immédiates, on a $\Phi(\lambda a+\mu b)=\lambda\Phi(a)+\mu\Phi(b)$ puisque pour tout $x\in E$,
$\varphi_{\lambda a+\mu b}(x)=(\lambda a+\mu b\mid x)=\lambda(a\mid x)+\mu(b\mid x)=(\lambda\varphi_a+\mu\varphi_b)(x)$.
De plus, l'application Φ est injective car si $a\in\ker\Phi$ alors $\varphi_a=\tilde{0}$ et en particulier $\varphi_a(a)=(a\mid a)=0$ ce qui entraîne $a=0$.
Par le théorème d'isomorphisme, on peut alors affirmer que Φ est un isomorphisme de $\mathbb{R}$-espace vectoriel.
□

12.2.7 Produit mixte

Soit E un espace vectoriel euclidien orienté de dimension $n\in\mathbb{N}^\star$.

Proposition

Si $\mathcal{B}=(e_1,\dots,e_n)$ et $\mathcal{B}'=(e_1',\dots,e_n')$ sont deux bases orthonormées de E alors a $\det_{\mathcal{B}}\mathcal{B}'=\pm1$.

dém. :
Notons $P=(p_{i,j})=\mathrm{Mat}_{\mathcal{B}}\mathcal{B}'=\mathrm{Mat}_{\mathcal{B}}(e_1',\dots,e_n')$.
Le coefficient d'indice (i,j) de P est la $i$ème composante dans $\mathcal{B}$ du $j$ème vecteur de $\mathcal{B}'$. Ainsi $p_{i,j}=(e_i\mid e_j')$.
Puisque $P^{-1}=(q_{i,j})=\mathrm{Mat}_{\mathcal{B}'}\mathcal{B}=\mathrm{Mat}_{\mathcal{B}'}(e_1,...,e_n)$, on a aussi $q_{i,j}=(e_i'\mid e_j)=(e_j\mid e_i')=p_{j,i}$.
Par suite $P^{-1}={}^tP$ et donc $1/\det P=\det P$ d'où $(\det P)^2=1$.
□

Remarque Si $\mathcal{B}$ et $\mathcal{B}'$ sont deux bases orthonormées directes alors $\det_{\mathcal{B}}\mathcal{B}'=1$.

Théorème

Si $(x_1,...,x_n)$ est une famille de $n=\dim E$ vecteurs de E et si $\mathcal{B}$ est une base orthonormée directe de E alors la quantité $\det_{\mathcal{B}}(x_1,\dots,x_n)$ ne dépend pas du choix de la base orthonormée $\mathcal{B}$.

dém. :
Si $\mathcal{B}'$ désigne une autre base orthonormée directe de E alors la formule de changement de base relative aux déterminants donne $\det_{\mathcal{B}}(x_1,...,x_n)=\det_{\mathcal{B}}\mathcal{B}'\det_{\mathcal{B}'}(x_1,...,x_n)=\det_{\mathcal{B}'}(x_1,...,x_n)$ car $\det_{\mathcal{B}}\mathcal{B}'=1$.
□

Définition

Cette quantité est appelée produit mixte des vecteurs $(x_1,\dots,x_n)$, on la note $\mathrm{Det}(x_1,...,x_n)$ ou $[x_1,...,x_n]$.

Exemple Si E désigne l'ensemble des vecteurs du plan ou de l'espace géométrique, on retrouve la notion de produit mixte déjà définie.

Proposition

L'application produit mixte $(x_1,...,x_n)\in E^n\mapsto[x_1,...,x_n]\in\mathbb{R}$ est une forme n linéaire alternée et antisymétrique.
De plus
$[x_1,...,x_n]=0\Leftrightarrow(x_1,...,x_n)$ est liée.
$[x_1,...,x_n]>0\Leftrightarrow(x_1,\dots,x_n)$ base directe.
$[x_1,...,x_n]<0\Leftrightarrow(x_1,\dots,x_n)$ base indirecte.

dém. :
Les propriétés affirmées découlent de ce que le produit mixte est un déterminant dans une base directe.
□

12.2.8 Produit vectoriel en dimension 3

Soit E un espace vectoriel euclidien orienté de dimension 3.
Pour u et v deux vecteurs de E, l'application $f : E \to \mathbb{R}$ définie par

$$f(x) = \text{Det}(u, v, x)$$

est une forme linéaire sur E.
Par le théorème de représentation des formes linéaires, on peut affirmer qu'il existe un unique vecteur $w \in E$ vérifiant

$$\forall x \in E, \text{Det}(u, v, x) = (w \mid x)$$

Définition

Ce vecteur w est appelé produit vectoriel de u par v, on le note $u \wedge v$.

Remarque Par définition $u \wedge v$ est l'unique vecteur vérifiant $\forall x \in E, (u \wedge v \mid x) = \text{Det}(u, v, x)$.

Proposition

L'application produit vectoriel $(u, v) \in E \times E \mapsto u \wedge v \in E$ est bilinéaire antisymétrique.

dém. :
Soient $u, v, w \in E$ et $\lambda, \mu \in \mathbb{R}$.
Pour tout $x \in E$,

$$((\lambda u + \mu v) \wedge w \mid x) = \text{Det}(\lambda u + \mu v, w, x)$$

Par linéarité du produit mixte en la première variable

$$((\lambda u + \mu v) \wedge w \mid x) = \lambda \text{Det}(u, w, x) + \mu \text{Det}(v, w, x)$$

et ainsi

$$((\lambda u + \mu v) \wedge w \mid x) = \lambda \, (u \wedge w \mid x) + \mu \, (v \wedge w \mid x) = (\lambda(u \wedge w) + \mu(v \wedge w) \mid x)$$

En posant $a = (\lambda u + \mu v) \wedge w$ et $b = \lambda(u \wedge w) + \mu(v \wedge w)$, on a obtenu

$$\forall x \in E, (a \mid x) = (b \mid x)$$

On en déduit $a = b$ car on a $(a - b \mid x) = 0$ pour tout $x \in E$ et en particulier pour $x = a - b$ cela entraîne $a - b = 0$ i.e. $a = b$.
Ainsi

$$(\lambda u + \mu v) \wedge w = \lambda(u \wedge w) + \mu(v \wedge w)$$

Le produit vectoriel est donc linéaire en sa première variable.
De plus $(v \wedge u \mid x) = \text{Det}(v, u, x) = -\text{Det}(u, v, x) = (-(u \wedge v) \mid x)$ pour tout $x \in E$ et on en déduit comme ci-dessus que $v \wedge u = -(u \wedge v)$.
Par suite le produit vectoriel est antisymétrique et donc bilinéaire.
□

Proposition

$\forall u, v \in E, (u \wedge v \mid u) = (u \wedge v \mid v) = 0.$

dém. :
$\text{Det}(u, v, u) = \text{Det}(u, v, v) = 0$ car le produit mixte d'une famille liée est nul.
□

Proposition

$\forall u, v \in E$, (u, v) est libre si, et seulement si, $u \wedge v \neq 0$.
De plus, si tel est le cas, la famille $(u, v, u \wedge v)$ est une base directe.

dém. :
Si (u, v) liée alors $\forall x \in E, \text{Det}(u, v, x) = 0 = (0 \mid x)$ et donc $u \wedge v = 0$.
Si (u, v) est libre, complétons cette famille en une base (u, v, w).
$(u \wedge v \mid w) = \text{Det}(u, v, w) \neq 0$ donc $u \wedge v \neq 0$.
De plus $\|u \wedge v\|^2 = \text{Det}(u, v, u \wedge v) > 0$ donc la famille $(u, v, u \wedge v)$ est une base directe.
□

Proposition

Si (i, j, k) est une base orthonormée directe de E alors
$i \wedge j = k, j \wedge k = i$ et $k \wedge i = j$.

dém. :
Calculons les composantes de du vecteur $i \wedge j$ dans la base (i, j, k).
$(i \mid i \wedge j) = (i \wedge j \mid i) = 0, (j \mid i \wedge j) = (i \wedge j \mid j) = 0$ et $(k \mid i \wedge j) = (i \wedge j \mid k) = \text{Det}(i, j, k) = 1$.
On en déduit $i \wedge j = 0.i + 0.j + 1.k = k$.
Les calculs sont semblables pour $j \wedge k$ et $k \wedge i$ en exploitant $\text{Det}(j, k, i) = \text{Det}(k, i, j) = 1$.
□

Proposition

Soit (i, j, k) une base orthonormée directe de E.
Si $u = xi + yj + zk$ et $v = x'i + y'j + z'k$ alors

$$u \wedge v = (yz' - y'z)i + (zx' - z'x)j + (xy' - x'y)k$$

dém. :
Il suffit de calculer les composantes du vecteur $i \wedge j$ dans la base (i, j, k).

$$(i \mid u \wedge v) = (u \wedge v \mid i) = \text{Det}(u, v, i) = \begin{vmatrix} x & x' & 1 \\ y & y' & 0 \\ z & z' & 0 \end{vmatrix} = yz' - y'z$$

Les calculs $(j \mid u \wedge v)$ et de $(k \mid u \wedge v)$ sont analogues.
□

Remarque Par coïncidence de la formule, on peut affirmer que le produit vectoriel dans l'espace géométrique précédemment présenté se confond avec le produit vectoriel en cours d'étude.

Proposition

$\forall u, v, w \in E, u \wedge (v \wedge w) = (u \mid w)v - (u \mid v)w.$

dém. :
Soit (i, j, k) une base orthonormée directe de E.
Soient $v = xi + yj + zk$ et $w = x'i + y'j + z'k$ deux vecteurs quelconques de E. On a $v \wedge w = (yz' - y'z)i + (zx' - z'x)j + (xy' - x'y)k$.
Pour $u = i$, on obtient
$i \wedge (v \wedge w) = (x'y - xy')j + (zx' - x'z)k$
et $(i \mid w)v - (i \mid v)w = x'v - xw = (x'y - xy')j + (zx' - x'z)k$.
De même, on vérifie que la formule du double produit vectoriel est aussi valable pour $u = j$ et pour $u = k$, puis par linéarité du produit vectoriel et du produit scalaire en leur première variable, on peut affirmer que la formule énoncée est valable pour tout $u \in E$.
□

Théorème

$\forall u, v \in E, (u \mid v)^2 + \|u \wedge v\|^2 = \|u\|^2 \|v\|^2$.

dém. :
$\|u \wedge v\|^2 = \text{Det}(u, v, u \wedge v)$
Puisque le produit mixte est une application antisymétrique
$\|u \wedge v\|^2 = \text{Det}(v, u \wedge v, u) = (v \wedge (u \wedge v) \mid u)$
Par la formule du double produit vectoriel, on obtient
$\|u \wedge v\|^2 = \left(\|v\|^2 u - (v \mid u)v \mid u\right) = \|v\|^2 \|u\|^2 - (v \mid u)^2$
□

Corollaire

Soient $u, v \in E$ non nuls et $\theta = \text{Ecart}(u, v)$.
On a
$$\|u \wedge v\| = \|u\| \|v\| \sin\theta$$

dém. :
$(u \mid v) = \|u\| \|v\| \cos\theta$ donc par l'identité de Lagrange,
$\|u \wedge v\|^2 = \|u\|^2 \|v\|^2 - \|u\|^2 \|v\|^2 \cos^2\theta = \|u\|^2 \|v\|^2 \sin^2\theta$
puis $\|u \wedge v\| = \|u\| \|v\| \sin\theta$ car $\theta \in [0, \pi]$ et donc $\sin\theta \geqslant 0$.
□

Exemple Si (u, v) est une famille orthonormée de vecteurs de E alors $(u, v, u \wedge v)$ est une base orthonormée directe de E.

Exemple Soit $\mathcal{B} = (i, j, k)$ une base orthonormée directe de E.
Soit $P = \text{Vect}(u, v)$ avec $u = i + j$ et $v = j - k$.
Formons une équation de P.
Il suffit de déterminer un vecteur normal à P.
$n = u \wedge v = -i + j + k$ convient.
$-x + y + z = 0$ est une équation de P.

Exemple Soit $\mathcal{B} = (i, j, k)$ une base orthonormée directe de E.
Soient les plans $P : x + y - z = 0$ et $P' : x + 2y - z = 0$.
Les plans P et P' ne sont pas confondus, déterminons un vecteur directeur de la droite $D = P \cap P'$.

$n = i + j - k$ et $n' = i + 2j - k$ sont des vecteurs normaux à P et P'.
$u = n \wedge n' = i + k$ est orthogonal à n et n' donc appartient aux plans P et P' et ainsi dirige D.

12.3 Projection et symétries orthogonales

E désigne un espace vectoriel euclidien.

12.3.1 Projection orthogonale

Soit F un sous-espace vectoriel de E.
Les espaces F et $F^\perp$ sont supplémentaires dans E.

Définition

On appelle projection orthogonale sur F la projection vectorielle p_F sur F parallèlement à $F^\perp$.

Exemple Si $F = \{0\}$ alors $p = 0$.
Si $F = E$ alors $p = \text{Id}$.

Proposition

p_F est un endomorphisme de E vérifiant $p_F^2 = p_F, \text{Im} p_F = F$ et $\ker p_F = F^\perp$.
De plus $\text{Id} - p_F$ est la projection orthogonale sur $F^\perp$.

dém. :
Puisque p_F est la projection sur F parallèlement à $F^\perp$, p_F est un projecteur de E, c'est-à-dire un endomorphisme de E vérifiant $p_F^2 = p_F$. Aussi $\text{Im} p_F = F$ (espace sur lequel on projette) et $\ker p_F = F^\perp$ (espace parallèlement auquel on projette).
Enfin $\text{Id} - p_F$ est la projection complémentaire de p_F c'est-à-dire la projection sur $F^\perp$ parallèlement à $F = F^{\perp\perp}$, autrement dit, c'est la projection orthogonale sur $F^\perp$.
□

Théorème

Si $\mathcal{B} = (e_1, \ldots, e_p)$ est une base orthonormée de F alors

$$\forall x \in E, p_F(x) = \sum_{j=1}^{p} (e_j \mid x) e_j$$

dém. :
Complétons la famille orthonormée $\mathcal{B} = (e_1, \ldots, e_p)$ en une base orthonormée $(e_1, \ldots, e_n)$.
Puisque $x = \sum_{k=1}^{n} (e_k \mid x) e_k$, on a par linéarité $p_F(x) = \sum_{k=1}^{n} (e_k \mid x) p_F(e_k)$.
Or $p_F(e_k) = e_k$ pour $k \in \{1, \ldots, p\}$ car $e_k \in F$ et $p_F(e_k) = 0$ pour $k > p$ car $e_k \in F^\perp$.
On en déduit $p_F(x) = \sum_{k=1}^{p} (e_k \mid x) e_k$
□

Exemple Soit $D = \text{Vect}(a)$ avec $a \neq 0$.
La famille formée du seul vecteur $a/\|a\|$ est une base orthonormée de D.

Par la formule précédente, on obtient alors
$\forall x \in E, p_D(x) = \dfrac{(a \mid x)}{\|a\|^2} a.$

Exemple Soit $H = \text{Vect}(a)^\perp$ avec $a \neq 0$.
Puisque $\text{Id} - p_H$ est la projection orthogonale sur la droite $D = H^\perp = \text{Vect}(a)$, on obtient

$$\forall x \in E, p_H(x) = x - \frac{(a \mid x)}{\|a\|^2} a$$

Exemple Soit E un espace vectoriel euclidien de dimension 3 muni d'une base orthonormée $\mathcal{B} = (i, j, k)$.
Soit p la projection orthogonale sur le plan $P : -x + y + z = 0$.
Formons la matrice A de p dans la base $\mathcal{B}$.
$a = -i + j + k$ est vecteur normal à P.
Pour tout $x \in E$, on a

$$p(x) = x - \frac{(a \mid x)}{\|a\|^2} a$$

On en déduit

$$p(i) = i + \frac{1}{3}(-i + j + k), p(j) = j - \frac{1}{3}(-i + j + k) \text{ et } p(k) = k - \frac{1}{3}(-i + j + k)$$

Ainsi

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

Théorème

Soit $p \in \mathcal{L}(E)$. On a équivalence entre :
(i) p est une projection orthogonale ;
(ii) $p^2 = p$ et $\forall x, y \in E, (p(x) \mid y) = (x \mid p(y))$.

dém. :
(i) $\Rightarrow$ (ii) Supposons que p est une projection orthogonale.
On sait déjà que $p^2 = p$.
Soit $x, y \in E$
On a

$$(p(x) \mid y) = (p(x) \mid y - p(y)) + (p(x) \mid p(y))$$

Or $(p(x) \mid y - p(y)) = 0$ car $p(x) \in F$ et $y - p(y) \in F^\perp$
donc

$$(p(x) \mid y) = (p(x) \mid p(y))$$

De façon semblable

$$(x \mid p(y)) = (p(x) \mid p(y))$$

et donc

$$(p(x) \mid y) = (x \mid p(y))$$

(ii) $\Rightarrow$ (i) Supposons (ii)
Puisque $p^2 = p$, les espaces $F = \text{Im}p$ et $G = \ker p$ sont supplémentaires dans E et p est la projection vectorielle sur F parallèlement à G. Pour conclure, il suffit de montrer $G = F^\perp$.
Soient $x \in F$ et $y \in G$. On a $(x \mid y) = (p(x) \mid y) = (x \mid p(y)) = (x \mid 0) = 0$.
On en déduit que F et G sont orthogonaux.
Ainsi $G \subset F^\perp$, or $\dim G = \dim E - \dim F = \dim F^\perp$ donc $G = F^\perp$.
□

Corollaire

La matrice représentative d'une projection orthogonale dans une base orthonormée est symétrique.

dém. :
Notons $A = (a_{i,j})$ la matrice d'une projection orthogonale p dans une base orthonormée $\mathcal{B} = (e_1, ..., e_n)$.
Le coefficient d'indice (i, j) de A est la i-ème composante dans $\mathcal{B}$ de l'image du j-ème vecteur de base par p. Ainsi $a_{i,j} = (e_i \mid p(e_j))$.
Or p est une projection orthogonale donc $a_{i,j} = (p(e_i) \mid e_j) = (e_j \mid p(e_i)) = a_{j,i}$.
Par suite la matrice A est symétrique.
□

12.3.2 Distance à un sous-espace vectoriel

Définition

On appelle distance d'un vecteur x de A à un sous-espace vectoriel F de E le réel

$$d(x, F) = \inf_{y \in F} d(x, y)$$

Théorème

$\forall y \in F, \|x - y\| \geqslant \|x - p_F(x)\|$ avec égalité si, et seulement si, $y = p(x)$.

dém. :
$x - y = (x - p_F(x)) + (p_F(x) - y)$ avec $x - p_F(x) \in F^\perp$ et $p_F(x) - y \in F$.
Par Pythagore $\|x - y\|^2 = \|x - p_F(x)\|^2 + \|p_F(x) - y\|^2 \geqslant \|x - p_F(x)\|^2$ avec égalité si, et seulement si, $y = p_F(x)$.
□

Corollaire

$d(x, F) = \|x - p_F(x)\|$.

dém. :
$d(x, F) = \inf_{y \in F} \|x - y\| = \min_{y \in F} \|x - y\| = \|x - p_F(x)\|$.
□

Exemple Soit $D = \text{Vect}(a)$ avec $a \neq 0$.
Pour tout $x \in E$, $p_D(x) = \dfrac{(a \mid x)}{\|a\|^2} a$ donc

$$d(x, D) = \left\| x - \frac{(a \mid x)}{\|a\|^2} a \right\|$$

Exemple Soit $H = \text{Vect}(a)^\perp$ avec $a \neq 0$.
Pour tout $x \in E$, $p_H(x) = x - \dfrac{(a \mid x)}{\|a\|^2} a$ donc

$$d(x, H) = \left\| x - \left(x - \frac{(a \mid x)}{\|a\|^2} a \right) \right\| = \frac{|(a \mid x)|}{\|a\|}$$

Proposition

Si $\mathcal{B} = (e_1, ..., e_n)$ est une base orthonormée de E et si H est un hyperplan d'équation $a_1x_1 + \cdots + a_nx_n = 0$ (avec $(a_1, \ldots, a_n) \neq (0, \ldots, 0)$) alors pour tout vecteur x de E de composantes $x_1, \ldots, x_n$ dans $\mathcal{B}$ on a

$$d(x, H) = \frac{|a_1x_1 + \cdots + a_nx_n|}{\sqrt{a_1^2 + \cdots + a_n^2}}$$

dém. :
Le vecteur $a = a_1e_1 + \cdots + a_ne_n$ est vecteur normal à H et la formule proposée est alors la concrétisation de la formule $d(x, H) = \dfrac{|(a \mid x)|}{\|a\|}$
□

Exemple Soit E un espace vectoriel euclidien de dimension 3 muni d'une base orthonormée $\mathcal{B} = (i, j, k)$.
Soient $P : x + 2y - 3z = 0$ et $D = \text{Vect}(u)$ avec $u = i + k$.
Soit $x = i + 2j + 3k$. Calculons $d(x, P)$ et $d(x, D)$.
Par les formules précédentes $d(x, P) = \dfrac{|1 + 4 - 9|}{\sqrt{1 + 4 + 9}} = \dfrac{4}{\sqrt{14}}$ et $d(x, D) = \|-i + 2j + k\| = \sqrt{6}$

Exemple Calculons

$$m = \inf_{(a,b) \in \mathbb{R}^2} \int_0^1 \left(t^2 - (at + b)\right)^2 \, \mathrm{d}t$$

Considérons $E = \mathbb{R}[X]$ muni du produit scalaire : $(P, Q) \mapsto \displaystyle\int_0^1 P(t)Q(t) \, \mathrm{d}t$.
On a $m = d(X^2, \mathbb{R}_1[X])^2$.
En introduisant la projection orthogonale $p = p_{\mathbb{R}_1[X]}$, on a $m = \left\|X^2 - p(X^2)\right\|^2$.
Il reste à déterminer $p(X^2)$.
Pour cela exploitons $p(X^2) \in \mathbb{R}_1[X]$ (1) et $X^2 - p(X^2) \in \mathbb{R}_1[X]^\perp$ (2)
Par (1), on peut écrire $p(X^2) = aX + b$.
Par (2), on a $(X^2 - p(X^2) \mid 1) = (X^2 - p(X^2) \mid X) = 0$.
On en déduit $(aX + b \mid 1) = (X^2 \mid 1)$ et $(aX + b \mid X) = (X^2 \mid X)$.
Cela fournit le système

$$\begin{cases} \dfrac{a}{2} + b = \dfrac{1}{3} \\ \dfrac{a}{3} + \dfrac{b}{2} = \dfrac{1}{4} \end{cases}$$

Après résolution, on obtient $p(X^2) = X - 1/6$ et on en déduit après calculs $m = 1/180$.

12.3.3 Symétrie orthogonale

Soit F un sous-espace vectoriel d'un espace vectoriel euclidien E.
Les espaces F et $F^{\perp}$ sont supplémentaires dans E.

Définition

On appelle symétrie orthogonale par rapport à F la symétrie vectorielle s_F par rapport à F parallèlement à la direction $F^{\perp}$.
Si F est un hyperplan de E, on dit que s_F est la réflexion par rapport à F.

Exemple Si $F = \{0\}$ alors $s = -\mathrm{Id}$.
Si $F = E$ alors $s = \mathrm{Id}$.

Proposition

s_F est un endomorphisme vérifiant $s_F^2 = \mathrm{Id}$, $\ker(s_F - I) = F$ et $\ker(s_F + I) = F^{\perp}$.
De plus $-s_F = s_{F^{\perp}}$ et $s_F = 2p_F - \mathrm{Id}$.

dém. :
Puisque s_F est la symétrie par rapport à F et parallèlement à $F^{\perp}$, s_F est un endomorphisme de E vérifiant $s_F^2 = \mathrm{Id}$. Aussi $\ker(s_F - \mathrm{Id}) = F$ (espace par rapport auquel a lieu la symétrie) et $\ker(s_F + \mathrm{Id}) = F^{\perp}$ (espace parallèlement auquel a lieu la symétrie).
De plus $-s_F$ est la symétrie complémentaire de s_F c'est-à-dire la symétrie par rapport à $F^{\perp}$ parallèlement à $F = F^{\perp\perp}$, autrement dit la symétrie par rapport à $F^{\perp}$.
Enfin $s_F = 2p_F - \mathrm{Id}$ car p_F est la projection sur F parallèlement à $F^{\perp}$.
□

Théorème

Si $\mathcal{B} = (e_1, \dots, e_p)$ est une base orthonormée de F alors

$$\forall x \in E, s_F(x) = 2\sum_{j=1}^{p}(e_j \mid x)e_j - x$$

dém. :
$s_F = 2p_F - \mathrm{Id}$ et $p_F(x) = \sum_{j=1}^{p}(e_j \mid x)e_j$.
□

Exemple Soit $D = \mathrm{Vect}(a)$ avec $a \neq 0$.

$$\forall x \in E, s_D(x) = 2\frac{(a \mid x)}{\|a\|^2}a - x$$

Exemple Soit $H = \mathrm{Vect}(a)^{\perp}$ avec $a \neq 0$.

$$\forall x \in E, s_H(x) = x - 2\frac{(a \mid x)}{\|a\|^2}a$$

Exemple Considérons $\mathbb{R}^3$ muni de sa base orthonormée canonique $\mathcal{B} = (i, j, k)$.
Soit s la réflexion par rapport au plan $P : x + y - z = 0$.
Formons la matrice A de s dans la base $\mathcal{B}$.
$a = i + j - k$ est vecteur normal à P.
Pour tout $x \in E$, on a $s(x) = x - 2\dfrac{(a \mid x)}{\|a\|^2}a$.
On en déduit $s(i) = i - \dfrac{2}{3}(i + j - k)$, $s(j) = j - \dfrac{2}{3}(i + j - k)$ et $s(k) = k + \dfrac{2}{3}(i + j - k)$
Ainsi

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

Théorème

Soit $s \in \mathcal{L}(E)$. On a équivalence entre :
(i) s est un symétrie orthogonale ;
(ii) $s^2 = \mathrm{Id}$ et $\forall x, y \in E, (s(x) \mid y) = (x \mid s(y))$.

dém. :
(i) $\Rightarrow$ (ii) Supposons que s est une symétrie orthogonale.
On sait déjà $s^2 = \mathrm{Id}$.
Soit $p = \dfrac{1}{2}(s + I)$ la projection associée à la symétrie s.
Pour tout $x, y \in E$. On a
$(s(x) \mid y) = (2p(x) - x \mid y) = 2\,(p(x) \mid y) - (x \mid y)$.
Or $(p(x) \mid y) = (x \mid p(y))$ car la projection p est orthogonale
donc
$(s(x) \mid y) = 2\,(x \mid p(y)) - (x \mid y) = (x \mid 2p(y) - y) = (x \mid s(y))$
(ii) $\Rightarrow$ (i) Supposons (ii)
Puisque $s^2 = \mathrm{Id}$, les espaces $F = \ker(s - \mathrm{Id})$ et $G = \ker(s + \mathrm{Id})$ sont supplémentaires dans E et q est la symétrie par rapport à F et parallèlement à G. Pour conclure, il suffit de montrer que $G = F^\perp$.
Soient $x \in F$ et $y \in G$.
$(x \mid y) = (s(x) \mid y) = (x \mid s(y)) = (x \mid -y) = -\,(x \mid y)$ donc $(x \mid y) = 0$.
On en déduit que les espaces F et G sont orthogonaux.
Ainsi $G \subset F^\perp$, or $\dim G = \dim E - \dim F = \dim F^\perp$ donc $G = F^\perp$.

□

Corollaire

La matrice d'une symétrie orthogonale dans une base orthonormée est symétrique.

dém. :
Notons $A = (a_{i,j})$ la matrice d'une symétrie orthogonale s dans une base orthonormée $\mathcal{B} = (e_1, ..., e_n)$.
Le coefficient d'indice (i, j) de A est la $i$ème composante dans $\mathcal{B}$ de l'image du $j$ème vecteur de base par s. Ainsi $a_{i,j} = (e_i \mid s(e_j))$.
Or s est une symétrie orthogonale donc $a_{i,j} = (s(e_i) \mid e_j) = (e_j \mid s(e_i)) = a_{j,i}$.
Par suite la matrice A est symétrique.

□

Corollaire

Les symétries orthogonales conservent le produit scalaire

$$\forall x, y \in E, (s(x) \mid s(y)) = (x \mid y)$$

En particulier, les symétries orthogonales conservent la norme

$$\forall x \in E, \|s(x)\| = \|x\|$$

dém. :
Soient $x, y \in E$.
$(s(x) \mid s(y)) = \left(x \mid s^2(y)\right) = (x \mid y)$ car $s^2 = \text{Id}$.
□

Exemple Soient $a, b \in E$ tels que $\|a\| = \|b\|$ et $a \neq b$.
Montrons qu'il existe une unique réflexion qui échange a et b.
Analyse :
Soit s une réflexion d'hyperplan H solution.
On a $s(a) = b$ et $s(b) = a$ donc $s(b - a) = a - b = -(b - a)$.
Par suite $b - a \in H^\perp$. Or $b - a \neq 0$ donc $H = \text{Vect}(b - a)^\perp$.
Ceci déterminer H, et donc s, de façon unique.
Synthèse :
Soit s la réflexion d'hyperplan $H = \text{Vect}(b - a)^\perp$
Puisque $\|a\| = \|b\|$, on a $(a + b \mid a - b) = \|a\|^2 - \|b\|^2 = 0$ donc $a + b \in H$.
Or $a = \frac{1}{2}(a + b) + \frac{1}{2}(a - b)$ donc $s(a) = \frac{1}{2}(a + b) - \frac{1}{2}(a - b) = b$ et $s(b) = s^2(a) = a$.
Ainsi s est solution.

12.4 Automorphismes orthogonaux

E désigne un espace vectoriel euclidien de dimension $n \in \mathbb{N}^\star$.

12.4.1 Matrices orthogonales

Définition

Une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est dite orthogonale si A est inversible et $A^{-1} = {}^tA$.
On note $\mathcal{O}_n(\mathbb{R})$ l'ensemble de ces matrices.

Exemple $I_n \in \mathcal{O}_n(\mathbb{R})$ et $-I_n \in \mathcal{O}_n(\mathbb{R})$

Remarque Par le théorème d'inversibilité il y a équivalence entre
(i) $A \in \mathcal{O}_n(\mathbb{R})$;
(ii) ${}^tAA = I_n$;
(iii) $A^tA = I_n$.

Proposition

$\mathcal{O}_n(\mathbb{R})$ est un sous-groupe de $\text{GL}_n(\mathbb{R})$.

dém. :
$\mathcal{O}_n(\mathbb{R}) \subset \mathrm{GL}_n(\mathbb{R})$, $I_n \in \mathcal{O}_n(\mathbb{R})$.
Soient $A, B \in \mathcal{O}_n(\mathbb{R})$.
AB est inversible et $(AB)^{-1} = B^{-1}A^{-1} = {}^tB{}^tA = {}^t(AB)$ donc $AB \in \mathcal{O}_n(\mathbb{R})$.
Soit $A \in \mathcal{O}_n(\mathbb{R})$.
A^{-1} est inversible et $(A^{-1})^{-1} = ({}^tA)^{-1} = {}^t(A^{-1})$ donc $A^{-1} \in \mathcal{O}_n(\mathbb{R})$.
□

Définition

$(\mathcal{O}_n(\mathbb{R}), \times)$ est appelé groupe orthogonal d'ordre n.

Proposition

$\forall A \in \mathcal{O}_n(\mathbb{R}), \det A = \pm 1$.

dém. :
${}^tA = A^{-1}$ donc $\det({}^tA) = \det(A^{-1})$ i.e. $\det A = (\det A)^{-1}$ d'où $(\det A)^2 = 1$.
□

Attention : La réciproque est fausse, $\det A = \pm 1$ n'implique pas $A \in \mathcal{O}_n(\mathbb{R})$.

Définition

Les matrices orthogonales de déterminant 1 (resp. -1) sont qualifiées de positives (resp. négatives).

Exemple I_n est une matrice orthogonale positive.
$-I_n$ est une matrice orthogonale positive si n est pair et négative sinon.

Proposition

$\mathrm{SO}_n(\mathbb{R}) = \{A \in \mathcal{O}_n(\mathbb{R}) / \det A = 1\}$ est un sous-groupe de $(\mathrm{GL}_n(\mathbb{R}), \times)$.

dém. :
$\mathrm{SO}_n(\mathbb{R}) = \mathcal{O}_n(\mathbb{R}) \cap \mathrm{SL}_n(\mathbb{R})$ est l'intersection de deux sous-groupes de $(\mathrm{GL}_n(\mathbb{R}), \times)$
□

Définition

$(\mathrm{SO}_n(\mathbb{R}), \times)$ est appelé groupe spécial orthogonal d'ordre n.

12.4.2 Caractérisation des matrices orthogonales

Considérons l'espace vectoriel réel $\mathcal{M}_{n,1}(\mathbb{R})$.

Pour $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbb{R})$, on pose

$$(X \mid Y) = {}^tXY = x_1y_1 + \cdots + x_ny_n$$

On vérifie aisément qu'on définit ainsi un produit scalaire sur l'espace $\mathcal{M}_{n,1}(\mathbb{R})$ des matrices colonnes.
Considérons aussi l'espace vectoriel réel $\mathcal{M}_{1,n}(\mathbb{R})$.

Pour $X = \begin{pmatrix} x_1 & \dots & x_n \end{pmatrix}$ et $Y = \begin{pmatrix} y_1 & \dots & y_n \end{pmatrix} \in \mathcal{M}_{1,n}(\mathbb{R})$ on pose $(X \mid Y) = X^tY = x_1y_1 + \dots + x_ny_n$.
On vérifie aisément qu'on définit ainsi un produit scalaire sur l'espace $\mathcal{M}_{1,n}(\mathbb{R})$ des matrices lignes.

Théorème

Soit $A \in \mathcal{M}_n(\mathbb{R})$ de colonnes $C_1, \dots, C_n$ et de lignes $L_1, \dots, L_n$.
On a équivalence entre :
(i) la matrice A est orthogonale ;
(ii) la famille $(C_1, \dots, C_n)$ est orthonormale ;
(iii) la famille $(L_1, \dots, L_n)$ est orthonormale.

dém. :
Calculons ${}^tAA$.
$A = (a_{i,j})$, ${}^tA = (a'_{j,i})$ avec $a'_{j,i} = a_{i,j}$ et ${}^tAA = (b_{i,j})$ avec $b_{i,j} = \sum_{k=1}^{n} a_{k,i}a_{k,j}$
Or $(C_i \mid C_j) = \sum_{k=1}^{n} a_{k,i}a_{k,j}$ donc ${}^tAA = I_n \Leftrightarrow \forall 1 \leqslant i, j \leqslant n, (C_i \mid C_j) = \delta_{i,j}$ et ainsi (i) $\Leftrightarrow$ (ii).
De même, en calculant A^tA, on obtient (i) $\Leftrightarrow$ (iii)
□

Exemple Soit

$$A = \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & -2/\sqrt{6} \end{pmatrix} \in \mathcal{M}_3(\mathbb{R})$$

En notant C_1, C_2, C_3 les colonnes de A on a $(C_1 \mid C_2) = (C_2 \mid C_3) = (C_3 \mid C_1) = 0$ et $\|C_1\| = \|C_2\| = \|C_3\| = 1$.
On en déduit $A \in \mathcal{O}_3(\mathbb{R})$.
De plus $\det A = 1$ donc $A \in \mathrm{SO}_3(\mathbb{R})$.

Exemple Pour $\theta \in \mathbb{R}$,

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \in \mathrm{SO}_2(\mathbb{R}) \text{ et } \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \in \mathrm{SO}_3(\mathbb{R})$$

12.4.3 Matrice de passage entre deux bases orthonormée

Théorème

Soient $\mathcal{B} = (e_1, \dots, e_n)$ une base orthonormée de E, $\mathcal{B}' = (e'_1, \dots, e'_n)$ une famille de vecteurs de E et P la matrice représentative de $\mathcal{B}'$ dans $\mathcal{B}$. On a équivalence entre :
(i) $\mathcal{B}'$ est une base orthonormée de E ;
(ii) $P \in \mathcal{O}_n(\mathbb{R})$.
De plus, si tel est le cas, P apparaît comme étant la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$ tandis que ${}^tP = P^{-1}$ est la matrice de passage de $\mathcal{B}'$ à $\mathcal{B}$.

dém. :
Notons $p_{i,j}$ le coefficient général de la matrice $P = \mathrm{Mat}_{\mathcal{B}}(e'_1, \dots, e'_n)$.

$p_{i,j}$ est la $i$ème composante dans la base $\mathcal{B}$ du vecteur e'_j.
Notons $p'_{i,j}$ le coefficient général de la matrice ${}^tP$.
On a $p'_{i,j} = p_{j,i}$
Enfin, notons $a_{i,j}$ le coefficient général de la matrice ${}^tPP$.
On a

$$a_{i,j} = \sum_{k=1}^{n} p'_{i,k}p_{k,j} = \sum_{k=1}^{n} p_{k,i}p_{k,j}$$

Or par calcul du produit scalaire de deux vecteurs par ses composantes dans une base orthonormée, on a aussi

$$(e'_i \mid e'_j) = \sum_{k=1}^{n} p_{k,i}p_{k,j}$$

Ainsi $a_{i,j} = (e'_i \mid e'_j)$ et donc ${}^tPP = I_n$ si, et seulement si, la famille $\mathcal{B}'$ est orthonormée et cette dernière est alors une base orthonormée de E.
□

Remarque Si les bases $\mathcal{B}$ et $\mathcal{B}'$ ont même orientation alors $P \in \mathrm{SO}_n(\mathbb{R})$ et donc $\det P = 1$. C'est cet argument qui a précédemment été utilisé pour définir le produit mixte d'une famille de vecteurs d'un espace euclidien orienté.

12.4.4 Automorphismes orthogonaux

Définition

On appelle endomorphisme orthogonal de E tout endomorphisme f de E conservant le produit scalaire i.e. vérifiant :

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

On note $\mathcal{O}(E)$ l'ensemble de ces applications.

Exemple Id et $-$Id sont des endomorphismes orthogonaux.

Exemple Les symétries orthogonales, et en particulier les réflexions, sont des endomorphismes orthogonaux.

Attention : Les projections orthogonales ne sont pas des endomorphismes orthogonaux.

Théorème

Soit f un endomorphisme de E.
On a équivalence entre :
(i) f est orthogonal ;
(ii) f conserve la norme i.e. $\forall x \in E, \|f(x)\| = \|x\|$.

dém. :
(i) $\Rightarrow$ (ii) Immédiat car la conservation du produit scalaire entraîne la conservation de la norme.
(ii) $\Rightarrow$ (i) Supposons que f conserve la norme.

Soient $x, y \in E$.
Par polarisation
$$(f(x) \mid f(y)) = \frac{1}{4}\left(\|f(x)+f(y)\|^2 - \|f(x)-f(y)\|^2\right)$$
par linéarité de f
$$(f(x) \mid f(y)) = \frac{1}{4}\left(\|f(x+y)\|^2 - \|f(x-y)\|^2\right)$$
et puisque f conserve la norme
$$(f(x) \mid f(y)) = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right) = (x \mid y)$$
Ainsi f conserve le produit scalaire.
□

Remarque Un endomorphisme orthogonal f conserve aussi :
- l'orthogonalité :
$$\forall x, y \in E, (x \mid y) = 0 \Rightarrow (f(x) \mid f(y)) = 0\,;$$
- les écarts angulaires :
$$\forall x, y \in E \text{ non nuls, } \mathrm{Ecart}(f(x), f(y)) = \mathrm{Ecart}(x, y)$$

Corollaire

> Un endomorphisme orthogonal de E est un automorphisme de E.
> On parle indifféremment d'endomorphisme ou d'automorphisme orthogonal voire d'isométrie.

dém. :
Si $f \in \mathcal{O}(E)$ alors par conservation de la norme $f(x) = 0 \Rightarrow x = 0$ et donc $\ker f = \{0\}$.
□

Proposition

> $\mathcal{O}(E)$ est un sous-groupe de $(\mathrm{GL}(E), \circ)$.

dém. :
$\mathcal{O}(E) \subset \mathrm{GL}(E)$ et $\mathrm{Id} \in \mathcal{O}(E)$.
Soient $f, g \in \mathcal{O}(E)$.
Pour tout $x \in E$, $\|(g \circ f)(x)\| = \|g(f(x))\|$.
Or g conserve la norme donc $\|(g \circ f)(x)\| = \|f(x)\|$ et puisque f converse la norme $\|(g \circ f)(x)\| = \|x\|$.
Par suite $g \circ f \in \mathcal{O}(E)$.
Soit $f \in \mathcal{O}(E)$.
Pour tout $x \in E$, $\left\|f(f^{-1}(x))\right\| = \left\|f^{-1}(x)\right\|$ car f conserve la norme et donc $\|x\| = \left\|f^{-1}(x)\right\|$.
Par suite $f^{-1} \in \mathcal{O}(E)$.
□

Définition

> $(\mathcal{O}(E), \times)$ est appelé groupe orthogonal de E.

Théorème

> Soient $f \in \mathcal{L}(E)$ et $\mathcal{B} = (e_1, \ldots, e_n)$ une base orthonormée de E.
> On a équivalence entre :
> (i) $f \in \mathcal{O}(E)$;
> (ii) $f(\mathcal{B})$ est une base orthonormée ;
> (iii) $\mathrm{Mat}_{\mathcal{B}} f \in \mathcal{O}_n(\mathbb{R})$.

dém. :
(i) $\Rightarrow$ (ii) Supposons f orthogonal.
Par conservation du produit scalaire $(f(e_i) \mid f(e_j)) = (e_i \mid e_j) = \delta_{i,j}$ et donc $(f(e_1), \dots, f(e_n))$ est une famille orthonormée. Puisque celle-ci est formée de $n = \dim E$ vecteurs de E, c'est une base orthonormée de E.
(ii) $\Rightarrow$ (iii) Si $f(\mathcal{B})$ est une base orthonormée alors $\text{Mat}_{\mathcal{B}} f = \text{Mat}_{\mathcal{B}} f(\mathcal{B})$ est orthogonal car cette matrice s'interprète comme la matrice de passage entre deux bases orthonormées de E.
(iii) $\Rightarrow$ (i) Supposons $A = \text{Mat}_{\mathcal{B}} f \in \mathcal{O}_n(\mathbb{R})$.
Soient $x, y \in E$ et notons X et Y leurs matrices composantes dans $\mathcal{B}$.
Les vecteurs $f(x)$ et $f(y)$ ont pour matrices composantes AX et AY dans $\mathcal{B}$.
On a alors $(f(x) \mid f(y)) = {}^t(AX)AY = {}^tX{}^tAAY = {}^tXY = (x \mid y)$ et donc f conserve le produit scalaire.
□

Corollaire

> Etant donné deux bases orthonormées $\mathcal{B}$ et $\mathcal{B}'$, il existe un unique automorphisme orthogonal transformant $\mathcal{B}$ et $\mathcal{B}'$.

dém. :
L'existence et l'unicité d'un endomorphisme transformant $\mathcal{B}$ en $\mathcal{B}'$ est assurée par le fait que le choix de l'image d'une base détermine entièrement un endomorphisme. Celui-ci est nécessairement orthogonal en vertu de (ii) $\Rightarrow$ (i).
□

Corollaire

> Si $f \in \mathcal{O}(E)$ alors $\det f = \pm 1$.

dém. :
Le déterminant de f est celui d'une matrice orthogonale.
□

Définition

> Les automorphismes orthogonaux de déterminant 1 (resp. -1) sont qualifiés de positifs (resp. négatifs).

Exemple Id est un automorphisme orthogonal positif.
$-$Id est un automorphisme orthogonal positif en dimension paire et négatif en dimension impaire.

Exemple Les réflexions sont des automorphismes orthogonaux négatifs.
En effet, soit s une réflexion d'hyperplan H et $\mathcal{B}$ une base adaptée à la supplémentarité de H et $H^{\perp}$. La matrice de s dans la base $\mathcal{B}$ est diag$(1, \dots 1, -1)$ et donc $\det s = -1$.

Proposition

> SO$(E) = \{f \in O(E) / \det f = 1\}$ est un sous-groupe de (GL$(E), \circ$).

dém. :
SO$(E) = \mathcal{O}(E) \cap$ SL(E) est l'intersection de deux sous-groupes de (GL$(E), \circ$).
□

Définition

> (SO$(E), \times$) est appelé groupe des isométries positives de E.

12.5 Automorphismes orthogonaux du plan euclidien

E désigne un plan euclidien orienté, par exemple E peut être l'ensemble des vecteurs du plan géométrique ou $\mathbb{R}^2$.

12.5.1 Matrice de rotation

Définition

Pour $\theta \in \mathbb{R}$, on appelle matrice de rotation d'angle θ la matrice

$$R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \in \mathrm{SO}_2(\mathbb{R})$$

Exemple $R(0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2$, $R(\pi) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -I_2$ et $R(\pi/2) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

Proposition

$\forall \theta, \theta' \in \mathbb{R}, R(\theta) = R(\theta') \Leftrightarrow \theta = \theta' \quad [2\pi]$

dém. :
Car

$$\begin{cases} \cos\theta = \cos\theta' \\ \sin\theta = \sin\theta' \end{cases} \Leftrightarrow \theta = \theta' \quad [2\pi]$$

□

Proposition

$\forall \theta, \theta' \in \mathbb{R}, R(\theta)\, R(\theta') = R(\theta + \theta') = R(\theta')R(\theta)$
En particulier, les matrices de rotation commutent entre elles.

dém. :
Par le calculs et en exploitant les formules de développement de $\cos(\theta + \theta')$ et $\sin(\theta + \theta')$.
□

Proposition

$\forall \theta \in \mathbb{R}\ R(\theta)^{-1} = {}^tR(\theta) = R(-\theta)$.

dém. :
$R(\theta)^{-1}$ se détermine aisément sachant $R(\theta) \in \mathcal{O}_2(\mathbb{R})$
□

Théorème

Pour toute matrice $M \in \mathrm{SO}_2(\mathbb{R})$, il existe un réel θ unique à 2π près vérifiant

$$M = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

dém. :
Unicité à 2π près car $R(\theta) = R(\theta') \Leftrightarrow \theta = \theta' \quad [2\pi]$.
Existence :

Soit

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SO}_2(\mathbb{R})$$

Puisque $a^2 + c^2 = 1$, il existe $\theta \in \mathbb{R}$ tel que $a = \cos\theta$ et $b = \sin\theta$.
On a $(a-d)^2 + (b+c)^2 = a^2 + d^2 - 2ad + b^2 + c^2 - 2bc$.
Or $a^2 + c^2 = b^2 + d^2 = 1$ et $\det M = ad - bc = 1$ donc $(a-d)^2 + (b+c)^2 = 0$
On en déduit $a = d$, $c = -b$ puis $M = R(\theta)$.
□

Corollaire

$\mathrm{SO}_2(\mathbb{R}) = \{R(\theta)/\theta \in \mathbb{R}\}$ est un groupe commutatif.

dém. :
Car les matrices de rotation commutent entre elles.
□

12.5.2 Rotation du plan euclidien orienté

Théorème

La matrice d'un endomorphisme $r \in \mathrm{SO}(E)$ est la même dans toute base orthonormée directe du plan E.
Plus précisément, il existe $\theta \in \mathbb{R}$, unique à 2π près, tel que matrice de r dans toute base orthonormée directe du plan E est

$$R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

dém. :
Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases orthonormées directes de E et P la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$.
Puisque $\mathcal{B}$ et $\mathcal{B}'$ sont orthonormée directe $P \in \mathrm{SO}_2(\mathbb{R})$.
Notons A et A' les matrices de l'endomorphisme r dans les bases $\mathcal{B}$ et $\mathcal{B}'$.
Ces matrices appartiennent à $\mathrm{SO}_2(\mathbb{R})$.
Par la formule de changement de base, $A' = P^{-1}AP$.
Or $A, P \in \mathrm{SO}_2(\mathbb{R})$ et le groupe $(\mathrm{SO}_2(\mathbb{R}), \times)$ est commutatif donc $A' = P^{-1}PA = A$.
Ainsi, la représentation de l'endomorphisme r est la même dans toute base orthonormée directe choisie et puisque celle-ci est une matrice de $\mathrm{SO}_2(\mathbb{R})$, c'est une matrice $R(\theta)$ avec θ déterminé de façon unique à 2π près.
□

Définition

L'automorphisme orthogonale positif représenté par la matrice $R(\theta)$ dans les bases orthonormées directes du plan E est appelée rotation d'angle θ et est noté Rot_θ.

Exemple $\mathrm{Id}_E = \mathrm{Rot}_0, -\mathrm{Id}_E = \mathrm{Rot}_\pi$.

Proposition

$\forall \theta, \theta' \in \mathbb{R}$, $\mathrm{Rot}_\theta = \mathrm{Rot}_{\theta'} \Leftrightarrow \theta = \theta' \quad [2\pi]$
$\forall \theta, \theta' \in \mathbb{R}$, $\mathrm{Rot}_\theta \circ \mathrm{Rot}_{\theta'} = \mathrm{Rot}_{\theta+\theta'} = \mathrm{Rot}_{\theta'} \circ \mathrm{Rot}_\theta$
$\forall \theta \in \mathbb{R}$, $\mathrm{Rot}_\theta^{-1} = \mathrm{Rot}_{-\theta}$.

dém. :
Il suffit de transposer matriciellement les énoncés.
□

Corollaire

$\mathrm{SO}(E) = \{\mathrm{Rot}_\theta/\theta \in \mathbb{R}\}$ est un groupe abélien.

12.5.3 Angle orienté dans le plan

12.5.3.1 Angle de deux vecteurs

Soient u et v deux vecteurs non nuls du plan euclidien orienté E.
Posons $U = u/\|u\|$ et $V = v/\|v\|$ les vecteurs unitaires de même direction et de même sens que u et v.
Montrons qu'il existe $\theta \in \mathbb{R}$ unique à 2π près vérifiant

$$V = \mathrm{Rot}_\theta(U)$$

Soit $\mathcal{B}$ une base orthonormée directe du plan de la forme $\mathcal{B} = (U, U')$.
Puisque les vecteur V est unitaire, les composantes x, y de V dans $\mathcal{B}$ vérifie

$$x^2 + y^2 = 1$$

et donc il existe $\alpha \in \mathbb{R}$ tel que

$$V = \cos(\alpha)U + \sin(\alpha)U'$$

Puisque la matrice dans la base orthonormée directe $\mathcal{B}$ de la rotation d'angle θ est

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

on a

$$\mathrm{Rot}_\theta(U) = \cos(\theta)U + \sin(\theta)U'$$

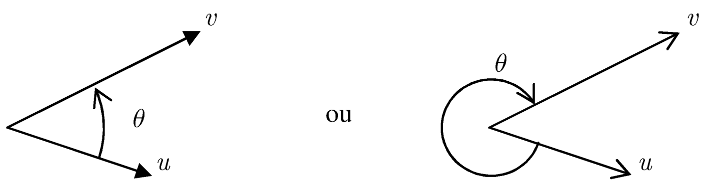
On en déduit

$$\mathrm{Rot}_\theta(U) = V \Leftrightarrow \theta = \alpha \quad [2\pi]$$

Ce qui détermine θ et ce de manière unique à 2π près.

Définition

Le réel θ, ainsi défini à 2π près, est appelé mesure de l'angle orienté de u à v.
On note $\theta = (u, v) \quad [2\pi]$ et on le visualise de la façon suivante



Remarque Nous avons ici une définition rigoureuse de la notion d'angle orienté formé par deux vecteurs non nuls d'un plan. Cette définition a été possible grâce à l'introduction d'un produit scalaire et d'une orientation du plan.

Attention : La notion d'angle orienté ne peut être introduite que dans un plan euclidien et celui-ci doit être préalablement orienté.

Remarque Si l'on inverse l'orientation du plan, les bases directes deviennent indirectes et on peut montrer que l'angle d'une rotation est transformée en son opposé. En conséquence, les angles orientés sont eux aussi transformés en leur opposé.

Exemple Pour tout vecteur non nul u du plan E, on a $(u, \mathrm{Rot}_\theta(u)) = \theta \quad [2\pi]$.
En particulier $(u, u) = 0 \quad [2\pi]$ et $(u, -u) = \pi \quad [2\pi]$.

12.5.3.2 Propriétés

Proposition

$\forall u, v \in E\setminus\{0\}, \forall \lambda, \mu \in \mathbb{R}^{+\star}, (\lambda u, \mu v) = (u, v) \quad [2\pi]$.

dém. :
Les vecteurs U et V sont les mêmes dans le cas où l'on étudie l'angle (u, v) ou dans le cas où on étudie l'angle $(\lambda u, \lambda v)$.
□
Remarque La notion d'angle orientée est liée à la direction et au sens des vecteurs en jeux et non à leur longueur.

Théorème

$\forall u, v, w \in E\setminus\{0\}, (u, v) = (u, w) + (w, v) \quad [2\pi]$

dém. :
Posons $U = u/\|u\|, V = v/\|v\|, W = w/\|w\|$.
Notons $\alpha = (u, w) \quad [2\pi]$ et $\beta = (w, v) \quad [2\pi]$.
On a $W = \mathrm{Rot}_\alpha(U)$ et $V = \mathrm{Rot}_\beta(W)$ donc $V = (\mathrm{Rot}_\beta \circ \mathrm{Rot}_\alpha)(U) = \mathrm{Rot}_{\beta+\alpha}(U)$.
On en déduit $(u, v) = \alpha + \beta \quad [2\pi]$.
□
Corollaire

$\forall u, v \in E\setminus\{0\}, (v, u) = -(u, v) \quad [2\pi]$.

dém. :
$(u, v) + (v, u) = (u, u) = 0 \quad [2\pi]$
□
Proposition

Soient $u, v \in E\setminus\{0\}$ et $\theta = (u, v) \quad [2\pi]$
On a $(u \mid v) = \|u\| \|v\| \cos\theta$ et $\mathrm{Det}(u, v) = \|u\| \|v\| \sin\theta$.
De plus ces deux relations déterminent θ à 2π près.

dém. :
Posons $U = u/\|u\|, V = v/\|v\|$.
Notons $\theta = (u, v) \quad [2\pi]$ de sorte que $V = \mathrm{Rot}_\theta(U)$.

Soit $\mathcal{B}$ une base orthonormée directe de E de la forme $\mathcal{B} = (U, U')$.
On a $u = \|u\| \, U$ et $v = \|v\| \operatorname{Rot}_\theta(U) = \|v\| (\cos(\theta) U + \sin(\theta) U')$.
Connaissant les composantes de u et v dans une base orthonormée directe, on peut alors calculer $(u \mid v)$ et $\operatorname{Det}(u, b)$ et on obtient $(u \mid v) = \|u\| \, \|v\| \cos\theta$ et $\operatorname{Det}(u, v) = \|u\| \, \|v\| \sin\theta$.
De plus ces deux relations déterminent $\cos\theta$ et $\sin\theta$ et permettent donc d'obtenir θ à 2π près.
□

Remarque Dans le cadre du plan géométrique, la notion d'angle orienté ci-dessus se confond avec la notion usuelle précédemment présentées.

Exemple Pour u, v vecteurs non nuls du plan E,
u et v sont orthogonaux si, et seulement si, $(u, v) = \pi/2 \quad [\pi]$
u et v sont colinéaires si, et seulement si, $(u, v) = 0 \quad [\pi]$.

Exemple Soient $\mathcal{B} = (i, j)$ une base orthonormée directe de E et $u = i + 2j$, $v = 3i - 4j$.
Déterminons une mesure θ de l'angle orienté de u à v.
On a $\|u\| = \sqrt{5}$, $\|v\| = 5$, $(u \mid v) = -5$ et $\operatorname{Det}(u, v) = -10$.
On en déduit $\cos\theta = -1/\sqrt{5}$ et $\sin\theta < 0$
Par suite $\theta = -\arccos(-1/\sqrt{5}) \quad [2\pi]$.

12.5.3.3 Lien entre angle orienté et écart angulaire

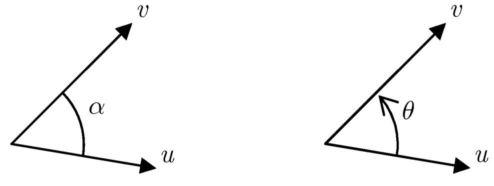
Soient $u, v \in E \setminus \{0\}$.
Notons $\alpha = \operatorname{Ecart}(u, v)$.
α est déterminé par

$$(u \mid v) = \|u\| \, \|v\| \cos\alpha \text{ et } \alpha \in [0, \pi]$$

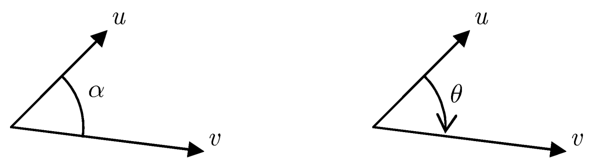
Notons $\theta = (u, v) \quad [2\pi]$.
θ est déterminé à 2π près par

$$\begin{cases} (u \mid v) = \|u\| \, \|v\| \cos\theta \\ \operatorname{Det}(u, v) = \|u\| \, \|v\| \sin\theta \end{cases}$$

On a donc $\cos\theta = \cos\alpha$ puis $\theta = \alpha \quad [2\pi]$ ou $\theta = -\alpha \quad [2\pi]$.
Si (u, v) est une famille liée alors $\theta = \alpha \quad [2\pi]$
Si (u, v) est une base directe alors $\operatorname{Det}(u, v) > 0$ et alors $\theta = \alpha \quad [2\pi]$.



Si (u, v) est une base indirecte alors $\theta = -\alpha \quad [2\pi]$.



Au final un écart angulaire, ou son opposé, est une mesure de l'angle orienté de deux vecteurs.

12.5.3.4 Angle orienté de deux droites

Soient D et D' deux droites vectorielles
Montrons qu'il existe $\theta \in \mathbb{R}$, unique à π près vérifiant

$$D' = \mathrm{Rot}_\theta(D)$$

Soient u et v deux vecteurs unitaires de D et D'.
On a

$$D' = \mathrm{Rot}_\theta(D) \Leftrightarrow \mathrm{Rot}_\theta(u) = v \text{ ou } \mathrm{Rot}_\theta(u) = -v$$

On en déduit

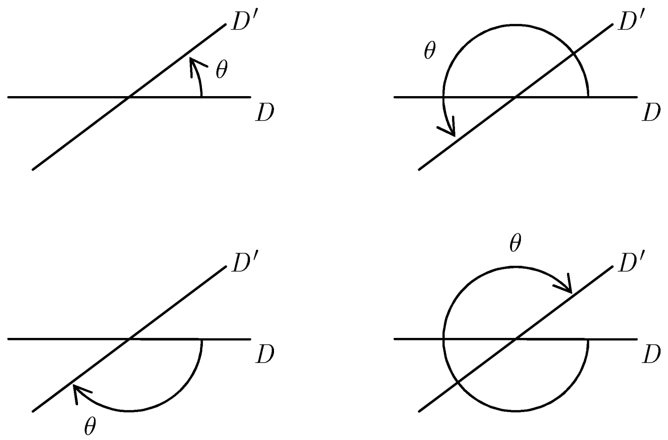
$$D' = \mathrm{Rot}_\theta(D) \Leftrightarrow \theta = (u, v) \text{ ou } \theta = (u, v) + \pi \quad [2\pi]$$

et finalement

$$D' = \mathrm{Rot}_\theta(D) \Leftrightarrow \theta = (u, v) \quad [\pi]$$

Définition

Ce réel θ, défini à π près, est appelé mesure de l'angle orienté de D à D'.
On note $\theta = (D, D') \quad [\pi]$ et on visualise cet angle de droite comme dans l'une des quatre figures suivantes



Remarque Pour déterminer θ, on mesure à π près l'angle orienté entre deux vecteurs directeurs de D et D'.

Proposition

Soient D, D', D'' des droites vectorielles de E.
$(D, D') + (D', D'') = (D, D'') \quad [\pi]$,
$(D', D) = -(D, D') \quad [\pi]$,
$D \perp D' \Leftrightarrow (D, D') = \pi/2 \quad [\pi]$,
$D = D' \Leftrightarrow (D, D') = 0 \quad [\pi]$.

dém. :
Soient u, v, w des vecteurs directeurs des droites D, D', D'' respectivement. On a
$(D, D') + (D', D'') = (u, v) + (v, w) = (u, w) = (D, D'') \quad [\pi]$
$(D', D) = (v, u) = -(u, v) = (D, D') \quad [\pi]$
$D \perp D' \Leftrightarrow (u \mid v) = 0 \Leftrightarrow (u, v) = \pi/2 \quad [\pi] \Leftrightarrow (D, D') = \pi/2 \quad [\pi]$
et
$D = D' \Leftrightarrow (u, v)$ liée $\Leftrightarrow (u, v) = 0 \quad [\pi] \Leftrightarrow (D, D') = 0 \quad [\pi]$
□

12.5.4 Automorphismes orthogonaux négatifs du plan

Théorème

Soit $\mathcal{B} = (i, j)$ une base orthonormée de E.
Pour tout automorphisme orthogonal négatif s de E, il existe $\varphi \in \mathbb{R}$ tel que

$$\operatorname{Mat}_{\mathcal{B}}(s) = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix}$$

De plus s correspond alors à la réflexion par rapport à la droite vectorielle dirigée par le vecteur $u = \cos(\varphi/2)i + \sin(\varphi/2)j$.

dém. :
Soit

$$M = \operatorname{Mat}_{\mathcal{B}}(s) = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{O}_2(\mathbb{R}) \backslash \mathrm{SO}_2(\mathbb{R})$$

Comme $a^2 + b^2 = 1$, il existe $\varphi \in \mathbb{R}$ tel que $a = \cos\varphi$ et $b = \sin\varphi$.
Or $(a + d)^2 + (b - c)^2 = a^2 + b^2 + c^2 + d^2 + 2(ad - bc)$
avec $a^2 + b^2 = 1$, $c^2 + d^2 = 1$ et $ad - bc = \det M = -1$ donc
$(a + d)^2 + (b - c)^2 = 0$.
On en déduit $c = \sin\varphi$, $d = -\cos\varphi$ puis

$$M = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix}$$

De plus
Considérons σ la réflexion par rapport à la droite vectorielle engendrée par le vecteur unitaire $u = \cos(\varphi/2)i + \sin(\varphi/2)j$. Pour tout vecteur x de E, on a $\sigma(x) = 2(x \mid u)u - x$ donc $\sigma(i) = \left(2\cos^2(\varphi/2) - 1\right) i + 2\cos(\varphi/2)\sin(\varphi/2)j = \cos(\varphi)i + \sin(\varphi)j$
et $\sigma(j) = 2\cos(\varphi/2)\sin(\varphi/2)i + \left(2\sin^2(\varphi/2) - 1\right) j = \sin(\varphi)i - \cos(\varphi)j$.
On en déduit que les applications linéaires σ et s prennent les mêmes valeurs sur i et j et sont donc égales.
□

12.5.5 Classification des automorphismes orthogonaux du plan

Les résultats qui précèdent se synthétisent de la façon suivante

Théorème

Les automorphismes orthogonaux positifs du plan sont les rotations vectorielles, celles-ci commutent entres elles et ont même représentation dans toute base orthonormée directe.
Les automorphismes orthogonaux négatifs du plan sont les réflexions.

Corollaire

La composée de deux rotations est une rotation.
La composée de deux réflexions est une rotation.
La composée d'une rotation et d'une réflexion est une réflexion.

dém. :
La composée de deux automorphismes orthogonaux et un automorphisme orthogonal et le signe de son déterminant s'obtient par le produit des signes des déterminants des automorphismes orthogonaux composés.
□

Corollaire

Toute rotation du plan peut s'écrire comme produit de deux réflexions, l'une d'elle étant choisie de manière quelconque.

dém. :
Soient r une rotation et σ une réflexion quelconque.
Posons $\sigma' = r \circ \sigma$ et $\sigma'' = \sigma \circ r$.
σ' et σ'' sont des réflexions et on a $r = \sigma' \circ \sigma^{-1} = \sigma' \circ \sigma$ et $r = \sigma^{-1} \circ \sigma'' = \sigma \circ \sigma''$..
□

Corollaire

Tout automorphisme orthogonal du plan peut s'écrire comme un produit d'au plus deux réflexions.

dém. :
Un automorphisme orthogonal du plan est :
- soit une réflexion ;
- soit une rotation, et cette dernière peut s'écrire comme produit de deux réflexions.
□

12.5.6 Composition d'automorphismes orthogonaux.

Théorème

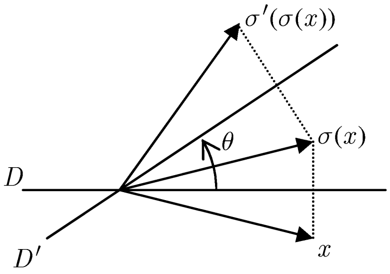
Les rotations du plan conservent les angles orientés tandis que les réflexions du plan les changent en leur opposé.

dém. :
Soient u, v deux vecteur unitaires de E et $\theta = (u, v) \quad [2\pi]$. On a $v = \mathrm{Rot}_\theta(u)$.
Soit r une rotation du plan.
$r(v) = (r \circ \mathrm{Rot}_\theta)(u) = (\mathrm{Rot}_\theta \circ r)(u)$ donc $(r(u), r(v)) = \theta \quad [2\pi]$.
Soit σ une réflexion du plan.
Posons $\sigma' = \sigma \circ \mathrm{Rot}_\theta$. σ' est une réflexion et $\mathrm{Rot}_\theta = \sigma \circ \sigma'$.
$\sigma(v) = \sigma \circ \mathrm{Rot}_\theta(u) = \sigma'(u) = (\sigma' \circ \sigma)(\sigma(u))$.
Or $\sigma' \circ \sigma = (\sigma'^{-1} \circ \sigma^{-1}) = (\sigma \circ \sigma')^{-1} = \mathrm{Rot}_\theta^{-1} = \mathrm{Rot}_{-\theta}$.
donc $(\sigma(v), \sigma(u)) = -\theta \quad [2\pi]$.
□

Proposition

Soient σ et σ' deux réflexions par rapport à deux droites D et D'.
On a $\sigma' \circ \sigma = \text{Rot}_{2\theta}$ avec $\theta = (D, D') \quad [\pi]$.

dém. :
$\sigma' \circ \sigma$ est une rotation, reste à déterminer son angle.
Soient u et v des vecteurs directeurs de D et D'.
On a $(u, v) = \theta \quad [\pi]$.
Calculons l'angle de la rotation $\sigma' \circ \sigma$ en évaluant
$(u, \sigma' \circ \sigma(u)) = (u, \sigma'(u)) \quad [2\pi]$
Par la relation de Chasles
$(u, \sigma' \circ \sigma(u)) = (u, v) + (v, \sigma'(v)) \quad [2\pi]$
Or
$(v, \sigma'(u)) = (\sigma'(v), \sigma'(u)) = -(v, u) \quad [2\pi]$
donc
$(u, \sigma' \circ \sigma(u)) = 2(u, v) = 2\theta \quad [2\pi]$.



□

Proposition

Soient r une rotation d'angle θ et σ une réflexion par rapport à D.
$r \circ \sigma$ est la réflexion par rapport à $D' = \text{Rot}_{\theta/2}(D)$
$\sigma \circ r$ est la réflexion par rapport à $D'' = \text{Rot}_{-\theta/2}(D)$.

dém. :
Posons σ' et σ'' les réflexions par rapport aux droites D' et D'' respectivement.
Puisque $(D, D') = \theta/2 \quad [\pi]$, $\sigma' \circ \sigma = \text{Rot}_\theta = r$ et donc $\sigma' = r \circ \sigma$.
Puisque $(D'', D) = \theta/2 \quad [\pi]$, $\sigma \circ \sigma'' = \text{Rot}_\theta = r$ et donc $\sigma'' = \sigma \circ r$.
□

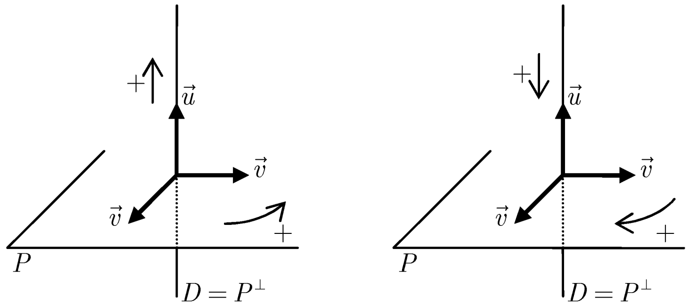
12.6 Automorphismes orthogonaux de l'espace de dimension 3

E désigne un espace vectoriel euclidien orienté de dimension 3, par exemple E peut être l'ensemble des vecteurs de l'espace géométrique ou $\mathbb{R}^3$.

12.6.1 Orientation induite

Soient P un plan de l'espace orienté E et $D = P^\perp$ sa droite normale.
Il n'existe pas a priori d'orientation préférentielle ni sur P, ni sur D.
Choisissons une orientation sur D et soit u un vecteur unitaire direct de D.
Complétons u en une base orthonormée directe $\mathcal{B} = (u, v, w)$ de E.
La famille (v, w) est une base orthonormée de P. En choisissant celle-ci pour base orientée de référence,

on dit qu'on a muni le plan P de l'orientation induite de celle de D. En effet on peut montrer que cette orientation est indépendante de la manière dont on a complété u en une base orthonormée directe.



Remarque Si l'on inverse l'orientation sur D, l'orientation induite sur P est-elle aussi inversée.

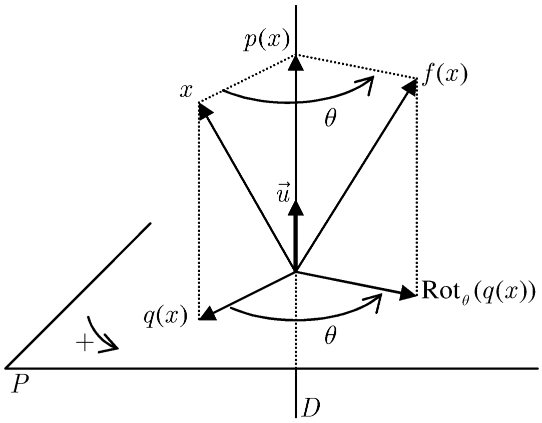
12.6.2 Rotation de l'espace de dimension 3

12.6.2.1 Définition

Soient $\theta \in \mathbb{R}$ et D une droite vectorielle orientée par un vecteur unitaire u.
Notons $P = D^{\perp}$ muni de l'orientation induite.
Soit p la projection orthogonale sur D et celle sur P.
Pour tout $x \in E$, on pose

$$f(x) = p(x) + \mathrm{Rot}_{\theta}(q(x))$$

où Rot_{θ} est la rotation d'angle θ du plan euclidien orienté P.



Définition

L'application f ainsi définie est appelée rotation d'axe dirigé et orienté par u et d'angle θ.
On note $f = \mathrm{Rot}_{u,\theta}$.

Proposition

f est un endomorphisme vérifiant $f_{\restriction D} = \mathrm{Id}_D$ et $f_{\restriction P} = \mathrm{Rot}_{\theta}$.

dém. :
f est linéaire par opération sur les applications linéaires.
Les restrictions affirmées sont immédiates sachant $p_{\upharpoonright D} = \mathrm{Id}_D$, $p_{\upharpoonright P} = \tilde{0}$, $q_{\upharpoonright D} = \tilde{0}$ et $q_{\upharpoonright P} = \mathrm{Id}_P$.
□

Proposition

Si $\theta = 0 \quad [2\pi]$ alors $f = \mathrm{Id}$.
Si $\theta \neq 0 \quad [2\pi]$ alors les vecteurs invariants par f sont ceux de D.

dém. :
Si $\theta = 0 \quad [2\pi]$ alors $\mathrm{Rot}_\theta = \mathrm{Id}_P$ donc $f(x) = p(x) + q(x) = x$ pour tout $x \in E$.
Si $\theta \neq 0 \quad [2\pi]$ alors pour $x \in E$,
$f(x) = x \Leftrightarrow \mathrm{Rot}_\theta(q(x)) = q(x)$.
Or le vecteur nul est le seul vecteur invariant par Rot_θ donc
$f(x) = x \Leftrightarrow q(x) = 0 \Leftrightarrow x \in D$
□

Théorème

La matrice de l'endomorphisme f dans une base orthonormée directe de E de la forme $\mathcal{B} = (u, v, w)$ est

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$$

dém. :
$f(u) = u$, $f(v) = \mathrm{Rot}_\theta(v)$ et $f(w) = \mathrm{Rot}_\theta(w)$.
Or (v, w) est une base orthonormée directe du plan $\mathcal{P}$ donc la matrice de Rot_θ dans la base (v, w) est

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

donc $f(v) = \cos(\theta)v + \sin(\theta)w$ et $f(w) = -\sin(\theta)v + \cos(\theta)w$.
□

Corollaire

Les rotations sont des automorphismes orthogonaux positifs de E.

dém. :
Car

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \in \mathrm{SO}_3(\mathbb{R})$$

□

Proposition

$\forall \theta, \theta' \in \mathbb{R}$, $\mathrm{Rot}_{u,\theta} = \mathrm{Rot}_{u,\theta'} \Leftrightarrow \theta = \theta' \quad [2\pi]$
$\forall \theta, \theta' \in \mathbb{R}$, $\mathrm{Rot}_{u,\theta} \circ \mathrm{Rot}_{u,\theta'} = \mathrm{Rot}_{u,\theta+\theta'} = \mathrm{Rot}_{u,\theta'} \circ \mathrm{Rot}_{u,\theta}$.
$\forall \theta, \theta' \in \mathbb{R}$, $\mathrm{Rot}_{u,\theta}^{-1} = \mathrm{Rot}_{u,-\theta}$.

dém. :
Immédiat par calcul matriciel.
□

Remarque Si on inverse l'orientation de D, l'orientation induite sur P l'est aussi et les mesures angulaires dans P sont alors changées en leur opposée. Par suite $\text{Rot}_{u,\theta} = \text{Rot}_{-u,-\theta}$.

12.6.2.2 Représentation matricielle dans une base orthonormée quelconque

La représentation matricielle précédente de l'endomorphisme $f = \text{Rot}_{u,\theta}$ a été obtenue dans une base adaptée à f, plus précisément une base orthonormée directe dont le premier vecteur est u. Voyons un résultat permettant d'obtenir la matrice de f dans une base orthonormée différente.

Proposition

Soit f la rotation d'axe D dirigé et orienté par un vecteur unitaire u et d'angle θ. $\forall x \in D^{\perp}, f(x) = \cos(\theta)x + \sin(\theta)u \wedge x$.

dém. :
Cas $x = 0$: ok.
Cas $x \neq 0$:
Posons $i = \dfrac{x}{\|x\|}$ et $j = u \wedge i = \dfrac{u \wedge x}{\|x\|}$.
La famille $\mathcal{B} = (i, j)$ est une base orthonormée directe du plan $P = D^{\perp}$.
Pour $x \in P$, $f(x) = \text{Rot}_\theta(x) = \|x\| \text{Rot}_\theta(i) = \|x\| (\cos\theta.i + \sin\theta.j) = \cos\theta.x + \sin\theta.(u \wedge x)$.
□

Exemple Soient $\mathcal{B} = (i, j, k)$ une base orthonormée directe de E et f la rotation d'axe D dirigé et orienté par $u = \dfrac{1}{\sqrt{3}}(i + j - k)$ et d'angle $\theta = \pi/3$.
Formons la matrice de f dans la base $\mathcal{B}$.
Notons p la projection orthogonale sur D et q celle sur $D^{\perp}$.

$$\forall x \in E, p(x) = (x \mid u)u \text{ et } q(x) = x - (x \mid u)u$$

Puisque

$$f(x) = f(p(x)) + f(q(x)) = p(x) + f(q(x))$$

avec $q(x) \in D^{\perp}$, la formule précédente donne

$$f(x) = p(x) + \cos(\theta)q(x) + \sin(\theta)u \wedge q(x)$$

puis

$$f(x) = \cos(\theta)x + (1 - \cos\theta)(x \mid u)u + \sin(\theta)u \wedge x$$

et au final

$$f(x) = \frac{1}{2}x + \frac{1}{6}(x \mid i + j + k)(i + j + k) + \frac{1}{2}(i + j - k) \wedge x$$

On en déduit

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

12.6.2.3 Retournement

Définition

On appelle retournement (ou demi-tour) autour d'une droite vectorielle D la rotation d'axe D et d'angle π. On la note Ret_D.

Remarque L'orientation de D n'a ici aucune incidence car $\text{Rot}_{u,\pi} = \text{Rot}_{-u,\pi}$.

Proposition

Ret_D est aussi la symétrie orthogonale par rapport à la droite D.

dém. :
Soit $\mathcal{B}$ une base orthonormée directe adaptée à la supplémentarité de D et $D^\perp$.
Le retournement autour de la droite D et la symétrie considérée ont tous deux la même matrice dans $\mathcal{B}$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

□

Exemple Montrons que Toute rotation f d'axe D et d'angle θ peut s'écrire comme un produit de deux retournements autour de droites orthogonales à D.
Considérons $P = D^\perp$ muni de l'orientation induite.
Dans le plan P, la rotation d'angle θ peut s'écrire $\text{Rot}_\theta = \sigma_2 \circ \sigma_1$ où σ_1 et σ_2 sont des réflexions du plan P par rapport à des droites D_1 et D_2 de ce plan.
Notons s_1 et s_2 les retournements de l'espaces autour de ces deux droites.
Pour tout $x \in D$, $f(x) = x$ et $(s_2 \circ s_1)(x) = s_2(-x) = x$.
Pour tout $x \in P$, $f(x) = \text{Rot}_\theta(x)$ et $(s_2 \circ s_1)(x) = s_2(\sigma_1(x)) = \sigma_2(\sigma_1(x)) = \text{Rot}_\theta(x)$.
Puisque tout vecteur x de E peut s'écrire $a + b$ avec $a \in D$ et $b \in P$, on a
$f(x) = f(a) + f(b) = (s_2 \circ s_1)(a) + (s_2 \circ s_1)(b) = (s_2 \circ s_1)(x)$.

12.6.3 Classification des automorphismes orthogonaux de l'espace

12.6.3.1 Préliminaires

Lemme

Soient F et G deux sous-espaces vectoriels supplémentaires d'un $\mathbb{K}$-espace vectoriel E et $f, g \in \mathcal{L}(E)$.
Si $\forall x \in F, f(x) = g(x)$ et $\forall x \in G, f(x) = g(x)$ alors $f = g$.

dém. :
Pour tout $x \in E$, on peut écrire $x = a + b$ avec $a \in F$ et $b \in G$.
On a alors $f(x) = f(a) + f(b) = g(a) + g(b) = g(x)$.
Ainsi $f = g$.
□

Lemme

> Soient f un automorphisme orthogonal d'un espace euclidien E et $F = \ker(f - \mathrm{Id})$ le sous-espace vectoriel formé des vecteurs invariants par f.
> L'espace $F^\perp$ est stable par f et $f_{\restriction F^\perp} : F^\perp \to F^\perp$ est un automorphisme orthogonal de $F^\perp$ dont le vecteur nul est le seul vecteur invariant.

dém. :
Affirmer que $F^\perp$ est stable par f signifie $f(F^\perp) \subset F^\perp$.
Soit $y \in f(F^\perp)$. Il existe $x \in F^\perp$ tel que $y = f(x)$.
Pour tout $a \in F$, on a alors $(y \mid a) = (f(x) \mid f(a)) = (x \mid a) = 0$ car $x \in F^\perp$ et $a \in F$.
Ainsi $f(F^\perp) \subset F^\perp$ et donc $F^\perp$ est stable par f.
L'application restreinte $f_{\restriction F^\perp} : F^\perp \to F^\perp$ est alors un endomorphisme de $F^\perp$ qui conserve le produit scalaire donc $f_{|F^\perp} \in \mathcal{O}(F^\perp)$.
Enfin, si $x \in F^\perp$ est un vecteur invariant par $f_{\restriction F^\perp}$ alors $x = f_{|F^\perp}(x) = f(x)$ donc $x \in F$ puis $x \in F \cap F^\perp = \{0\}$.
□

Lemme

> Tout automorphisme orthogonal positif d'un espace vectoriel euclidien de dimension 3 admet au moins un vecteur unitaire invariant.

dém. :
Soit $f \in \mathrm{SO}(E)$. Il suffit de montrer $\ker(f - \mathrm{Id}) \neq \{0\}$ pour conclure à l'existence d'au moins un vecteur unitaire invariant. Pour montrer $\ker(f - \mathrm{Id}) \neq \{0\}$, il suffit de vérifier $\det(f - \mathrm{Id}) = 0$ ce que nous allons justifier matriciellement.
Soient $\mathcal{B}$ une base orthonormée de E et A la matrice de f dans $\mathcal{B}$, $A \in \mathrm{SO}_3(\mathbb{R})$.
La matrice de $f - \mathrm{Id}$ est $M = A - I_3$.
On a ${}^tAM = {}^tAA - {}^tA = I - {}^tA = -{}^tM$ donc $\det({}^tAM) = \det(-{}^tM)$.
Or $\det({}^tAM) = \det A \det M = \det M$ et $\det(-{}^tM) = (-1)^3 \det M = -\det M$
donc $\det M = 0$.
Par suite M n'est pas inversible et donc $\ker(f - \mathrm{Id}) \neq \{0\}$.
□

12.6.3.2 Classification des automorphismes orthogonaux de l'espace

Théorème

> Soit $f \in \mathcal{O}(E)$ et $F = \ker(f - \mathrm{Id})$ le sous-espace vectoriel formé des vecteurs invariants par f.
> Si $\dim F = 3$ alors $f = \mathrm{Id}$.
> Si $\dim F = 2$ alors f est la réflexion de plan $P = F$.
> Si $\dim F = 1$ alors f est une rotation vectorielle autour de $D = F$.
> Si $\dim F = 0$ alors f est la composée commutative d'une réflexion par rapport à un plan P et d'une rotation autour de sa droite normale $D = P^\perp$

dém. :
Puisque F est formé des vecteurs invariants par f, on a $f_{\restriction F} = \mathrm{Id}_F$.
D'autre part $f_{\restriction F^\perp} \in \mathcal{O}(F^\perp)$ dont le seul vecteur invariant est le vecteur nul.
Cas $\dim F = 3$.
On a $F = E$ et donc $f = \mathrm{Id}$.
Cas $\dim F = 2$.
$P = F$ est un plan et $D = F^\perp$ est une droite vectorielle.
Pour tout $x \in D$, $f(x)$ appartient à la droite D et $\|f(x)\| = \|x\|$. On en déduit $f(x) = x$ ou $f(x) = -x$.

Or, dans $D = F^\perp$, seul le vecteur nul est invariant. On en déduit que $f(x) = -x$ pour tout $x \in D$.
Puisque de plus, $f(x) = x$ pour tout $x \in P$, on peut affirmer que f est la réflexion de plan P.
Cas $\dim F = 1$.
$D = F$ est une droite et $P = F^\perp$ est un plan vectoriel.
Puisque $f_{\restriction P}$ est un automorphisme orthogonal du plan P qui ne possède pas de vecteur invariant autre que le vecteur nul, on peut affirmer que $f_{\restriction P}$ est une rotation du plan P (rappelons que les automorphismes orthogonaux d'un plan euclidien sont les réflexions et les rotations)
Puisque de plus, $f(x) = x$ pour tout $x \in D$, on peut affirmer que f est une rotation autour de la droite D.
Cas $\dim F = 0$.
L'automorphisme orthogonal f ne peut être positif car un automorphisme orthogonal positif possède au moins un vecteur invariant unitaire. Puisque $\det(-f) = (-1)^3 \det f = -\det f$, l'endomorphisme orthogonal $-f$ est quant à lui positif et donc il existe un vecteur $x \in E$ unitaire tel que $f(x) = -x$.
Considérons alors σ la réflexion par rapport au plan $P = \text{Vect}(x)^\perp$.
On a $\sigma \circ f \in \text{SO}(E)$ et $(\sigma \circ f)(x) = x$.
De par l'étude ci-dessus, on peut affirmer :
- $\sigma \circ f = \text{Id}_E$ (exclu car alors $f = \sigma$ et $F \neq \{0\}$) ;
- $\sigma \circ f$ est une réflexion (exclu car alors $f \in$ SO(E)) ;
- ou $\sigma \circ f$ est une rotation autour de la droite $D = \text{Vect}(x)$.
Les deux premiers cas étant exclus, on obtient $f = \sigma \circ r$ avec r une rotation autour de D.
Enfin, puisque pour tout $x \in D$, $(\sigma \circ r)(x) = \sigma(x) = (r \circ \sigma)(x)$ et, pour tout $x \in P$, $(\sigma \circ r)(x) = r(x) = (r \circ \sigma)(x)$, on peut affirmer $\sigma \circ r = r \circ \sigma$ et donc f est la composée commutative d'une réflexion par rapport à P et d'une rotation autour de $D = P^\perp$.
□

Corollaire

> Les automorphismes orthogonaux positifs de l'espace E sont les rotations.
> Les automorphismes orthogonaux négatifs possédant d'autres vecteurs invariants que le vecteur nul sont les réflexions.

Corollaire

> La composée de deux rotations de l'espace est une rotation.
> La composée de deux réflexions distinctes de l'espace est une rotation autour de la droite intersection des plans de réflexion.

dém. :
La composée de deux automorphismes orthogonaux positifs est positive donc la composée de deux rotations est une rotation.
La composée de deux automorphisme orthogonaux négatifs est positive donc la composée de deux réflexions distinctes est une rotation. Puisque la droite intersection des deux plans de réflexion est invariante, cette rotation opère autour de cette droite.
□

Corollaire

> Toute rotation f d'axe D peut s'écrire comme produit de deux réflexions par rapport à des plans contenant D l'une d'elles pouvant être choisie de manière quelconque.

dém. :
Soit σ une réflexion par rapport à un plan contenant l'axe D.
Posons $\sigma' = \sigma \circ f$ et $\sigma'' = f \circ \sigma$.
σ' et σ'' sont des automorphismes orthogonaux négatifs vérifiant $\forall x \in D, \sigma'(x) = x = \sigma''(x)$.
σ' et σ'' sont donc des réflexions par rapport à des plans contenant D.
Puisque $f = \sigma \circ \sigma'$ et $f = \sigma'' \circ \sigma$, f apparaît comme la composée de deux réflexions par rapport à des

plans contenant D, l'une d'elles étant choisie de façon quelconque.
□

Corollaire

Tout automorphisme orthogonal de l'espace peut s'écrire comme un produit d'au plus trois réflexions.

dém. :
Car une rotation s'écrit comme produit de deux réflexions et qu'un automorphisme orthogonal autre qu'une rotation est une réflexion ou une composée commutative d'une réflexion et d'une rotation.
□

12.6.4 Réduction d'automorphismes orthogonaux

12.6.4.1 Principe

Soit f un endomorphisme de E connu par sa matrice A dans une base orthonormée directe $\mathcal{B} = (i, j, k)$.
Si $A \in \mathcal{O}_3(\mathbb{R})$ alors $f \in \mathcal{O}(E)$.
Pour préciser l'automorphisme orthogonal f, on détermine l'espace $F = \ker(f - \mathrm{Id})$ des vecteurs invariants par f.
Cas $\dim F = 3$
On a $f = \mathrm{Id}$.
Cas $\dim F = 2$
f est la réflexion par rapport au plan F.
Cas $\dim F = 1$
f est une rotation autour de la droite $D = F$.
Précisons celle-ci. On choisit un vecteur unitaire u de D et on oriente D par celui-ci. Déterminons l'angle θ de la rotation f autour de l'axe D.
Dans une base orthonormée directe de la forme $\mathcal{B}' = (u, v, w)$, la matrice de f est de la forme

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$$

On en déduit tr$f = 2\cos\theta + 1$. Or on a aussi tr$f =$ tr$A$ et donc $2\cos\theta + 1 =$ tr$A$ ce qui détermine $\cos\theta$.
Pour conclure, il suffit de connaître le signe de $\sin\theta$.
Soit $x = \alpha u + \beta v + \gamma w \notin D$. On a

$$\mathrm{Det}(u, x, f(x)) = \begin{vmatrix} 1 & \alpha & \alpha \\ 0 & \beta & \beta\cos\theta - \gamma\sin\theta \\ 0 & \gamma & \beta\sin\theta + \gamma\cos\theta \end{vmatrix} = (\beta^2 + \gamma^2)\sin\theta$$

Ainsi le signe de $\sin\theta$ est celui de Det$(u, x, f(x))$.
En pratique, on déterminer le signe de $\sin\theta$ en étudiant celui de Det$(u, i, f(i))$
Cas $\dim F = 0$ (hors-programme)
L'endomorphisme $-f$ est une rotation dont on peut préciser l'axe D et l'angle θ. f est alors la composée commutative de la réflexion de plan $P = D^{\perp}$ et de la rotation d'axe D et d'angle $\pi + \theta$.

12.6.4.2 Exemples

Soit f l'endomorphisme de E dont la matrice dans une base orthonormée directe $\mathcal{B} = (i, j, k)$ est la matrice A suivante :

Exemple Soit

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

Les colonnes de A sont unitaires et deux à deux orthogonales donc $A \in \mathcal{O}_3(\mathbb{R})$ puis $f \in \mathcal{O}(E)$.
Soit $u = xi + yj + zk \in E$. Après résolution
$f(u) = u \Leftrightarrow x + y - z = 0$.
L'ensemble des vecteurs invariants par f est un plan, on en déduit que f est la réflexion par rapport au plan d'équation $x + y - z = 0$.

Exemple Soit

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

Les colonnes de A sont unitaires et deux à deux orthogonales donc $A \in \mathcal{O}_3(\mathbb{R})$ puis $f \in \mathcal{O}(E)$.
Soit $u = xi + yj + zk \in E$. Après résolution

$$f(u) = u \Leftrightarrow x = y = z$$

L'ensemble des vecteurs invariants par f est la droite D dirigé par le vecteur unitaire

$$u = \frac{1}{\sqrt{3}}(i + j + k)$$

f est une rotation autour de la droite D.
Orientons la droite D par le vecteur u et déterminons l'angle θ de cette rotation.
Puisque tr$A = 2\cos\theta + 1$ et tr$A = 0$, on obtient $\cos\theta = -1/2$.
Déterminons le signe de Det$(u, i, f(i))$.

$$\text{Det}(u, i, f(i)) = \frac{1}{\sqrt{3}} \begin{vmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{vmatrix} = \frac{1}{\sqrt{3}}$$

On en déduit $\sin\theta > 0$ puis $\theta = -2\pi/3 \quad [2\pi]$.
Finalement f est la rotation d'axe dirigé et orienté par $u = \dfrac{1}{\sqrt{3}}(i + j + k)$ et d'angle $2\pi/3$.

Exemple Soit

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

Les colonnes de A sont unitaires et deux à deux orthogonales donc $A \in \mathcal{O}_3(\mathbb{R})$ puis $f \in \mathcal{O}(E)$.
Soit $u = xi + yj + zk \in E$. Après résolution

$$f(u) = u \Leftrightarrow x = z \text{ et } y = 2z$$

L'ensemble des vecteurs invariants par f est la droite D dirigé par le vecteur unitaire
$u = \dfrac{1}{\sqrt{6}}(i + 2j + k)$.
f est une rotation autour de la droite D.
Orientons la droite D par le vecteur u et déterminons l'angle θ de cette rotation.
Puisque tr$A = 2\cos\theta + 1$ et tr$A = -1$, on obtient $\cos\theta = -1$.
Il est alors inutile de déterminer le signe de $\sin\theta$, on peut directement affirmer $\theta = \pi \quad [2\pi]$ et conclure que f est le retournement autour de la droite $D = $ Vect$(i + 2j + k)$.

Deuxième partie

Analyse

Chapitre 13

Nombres réels et complexes

13.1 Nombres réels

Weierstrass (1863) et Dedekind (1872) furent les premiers à proposer une construction satisfaisante de $\mathbb{R}$... que nous ne présenterons pas.
Parmi les nombres réels, ceux qui s'écrivent comme rapport de deux entiers sont dits rationnels, ils forment l'ensemble $\mathbb{Q}$. Les autres sont dits irrationnels.

13.1.1 Opérations dans $\mathbb{R}$

$\mathbb{R}$ est muni de deux opérations $+$ et $\times$ c'est-à-dire de deux applications :

$$\begin{array}{l} \mathbb{R}\times\mathbb{R}\to\mathbb{R} \\ (x,y)\mapsto x+y \end{array} \text{ et } \begin{array}{l} \mathbb{R}\times\mathbb{R}\to\mathbb{R} \\ (x,y)\mapsto x\times y \end{array}$$

Ces opérations présentent des propriétés calculatoires qu'il est important de souligner.
Propriétés de l'addition :
Pour tout $a,b,c\in\mathbb{R}$,
- $a+b=b+a$ (commutativité) ;
- $(a+b)+c=a+(b+c)$ noté $a+b+c$ (associativité) ;
- $a+0=0+a=a$ (0 est élément neutre) ;
- Pour tout $a\in\mathbb{R}$ il existe un unique $d\in\mathbb{R}$ tel que $a+d=d+a=0$ (tout élément est symétrisable).

Cet élément d est noté $-a$ et cela permet de définir l'opération de soustraction.
Propriétés de la multiplication :
Pour tout $a,b,c\in\mathbb{R}$,
- $ab=ba$ (commutativité) ;
- $(ab)c=a(bc)$ noté abc (associativité) ;
- $a\times 1=1\times a=a$ (1 est élément neutre) ;
- $a(b+c)=ab+ac$ (la multiplication est distributive sur l'addition) ;
- Pour tout $a\neq 0$ il existe un unique $d\in\mathbb{R}$ tel que $ad=da=1$ (tout élément non nul est inversible).

Cet élément d est noté $1/a$ et cela permet de définir l'opération de division.

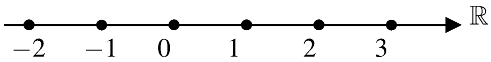
Attention : Il faut acquérir le réflexe : étudier la non nullité de ce par quoi on divise.

Remarque Les propriétés élémentaires précédentes permettent de retrouver les propriétés calculatoires usuelles connues. Par exemple :

$\forall a \in \mathbb{R}$, $a \times 0 = 0$ car $a \times 0 = a \times (0+0) = a \times 0 + a \times 0$ donc $a \times 0 = 0$.
$\forall a, b \in \mathbb{R}, (-a)b = -(ab)$ car $(-a)b + (ab) = (-a+a)b = 0b = 0$.

13.1.2 Relation d'ordre

$\mathbb{R}$ est ordonné par une relation $\leqslant$.
Celle-ci permet de visualiser $\mathbb{R}$ comme une droite



La relation $\leqslant$ est compatible avec les opérations $+$ et $\times$ dans le sens où pour tout $a, b, c \in \mathbb{R}$,

$$a \leqslant b \Leftrightarrow a + c \leqslant b + c \text{ et } (a \leqslant b \text{ et } c \geqslant 0) \Rightarrow ac \leqslant bc$$

Remarque Les propriétés élémentaires précédentes permettent de retrouver les propriétés calculatoires usuelles connues. Par exemple :
Si $a \leqslant b$ alors $-b \leqslant -a$.
En effet $-a = a - (a+b) \leqslant b - (a+b) = -a$.
Si $a > 0$ alors $1/a > 0$.
En effet $a > 0$ et $1/a \leqslant 0$ donne $1 \leqslant 0$ qui est absurde.
Si $c \leqslant 0$ alors $-c \geqslant 0$ donc $-ac \leqslant -bc$ puis $bc \leqslant ac$.

Attention : Il faut acquérir le réflexe : étudier le signe de la quantité par laquelle on multiplie une inégalité.

Proposition

Pour tout $a, b \in \mathbb{R}$, si $0 < a \leqslant b$ ou si $a \leqslant b < 0$ alors $\dfrac{1}{b} \leqslant \dfrac{1}{a}$.

dém. :
Si $0 < a \leqslant b$ ou $a \leqslant b < 0$ alors $ab > 0$ et donc $1/ab > 0$.
Par suite

$$\frac{1}{b} = \frac{a}{ab} \leqslant \frac{b}{ab} = \frac{1}{a}$$

□

Attention : Le passage à l'inverse de quantité de même signe, renverse les inégalités : c'est la décroissance de la fonction inverse.

Proposition

$\forall a \in \mathbb{R}, a^2 \geqslant 0$.

dém. :
Si $a \geqslant 0$ alors $a^2 \geqslant a \times 0 = 0$ via multiplication d'une inégalité par un facteur positif.

Si $a \leqslant 0$ alors $a^2 \geqslant a \times 0 = 0$ via multiplication d'une inégalité par un facteur négatif.
Dans les deux cas $a^2 \geqslant 0$.
□

Proposition

$\forall a, b \in \mathbb{R}, ab \leqslant \dfrac{1}{2}(a^2 + b^2)$.

dém. :
$(a - b)^2 = a^2 - 2ab + b^2 \geqslant 0$ donne $2ab \leqslant a^2 + b^2$.
□

Remarque En mathématique, il est usuel de manipuler des inégalités larges (car plus sûres) et de réserver la manipulation d'inégalités strictes au cas où celles-ci sont significatives.

Proposition

Pour $a \in \mathbb{R}^+$, $(\forall \varepsilon > 0, a \leqslant \varepsilon) \Rightarrow a = 0$

dém. :
Par contraposée.
Supposons $a \neq 0$. On a $a > 0$.
Pour $\varepsilon = a/2$, on a $\varepsilon > 0$ et $a > \varepsilon$.
Ainsi, il existe $\varepsilon > 0$ tel que $a > \varepsilon$.
□

13.1.3 Valeur absolue

Définition

On appelle valeur absolue d'un réel x , le réel positif

$$|x| = \begin{cases} x & \text{si } x \geqslant 0 \\ -x & \text{sinon} \end{cases}$$

Proposition

Pour $x, y \in \mathbb{R}$,
$|x| \geqslant 0$, $|-x| = |x|$, $x \leqslant |x|$,
$|x| = 0 \Leftrightarrow x = 0$,
$|xy| = |x|\,|y|$ et si $x \neq 0$, $|1/x| = 1/|x|$.

dém. :
Les propriétés énoncés sont vraies que x soit positif ou non.
□

Proposition

$$\forall x, y \in \mathbb{R}, |x + y| \leqslant |x| + |y|$$

De plus, il y a égalité si, et seulement si, x et y ont même signe.

dém. :

$$|x + y|^2 = (x + y)^2 = x^2 + 2xy + y^2 = |x|^2 + 2xy + |y|^2$$

Or

$$2xy \leqslant 2\,|xy| = 2\,|x|\,|y|$$

donc

$$|x+y|^2 \leqslant |x|^2 + 2\,|x|\,|y| + |y|^2$$

Ainsi

$$|x+y|^2 \leqslant (|x|+|y|)^2$$

Cette comparaison de carrés engageant des quantités positives, on en déduit $|x+y| \leqslant |x|+|y|$.
De plus il y a égalité si, et seulement si, $xy = |xy|$ i.e. $xy \geqslant 0$.
□

Corollaire

$\forall x, y \in \mathbb{R}, ||x|-|y|| \leqslant |x-y|$.

dém. :
$|x| = |x-y+y| \leqslant |x-y|+|y|$ donc $|x|-|y| \leqslant |x-y|$.
Un raisonnement symétrique donne $|y|-|x| \leqslant |y-x| = |x-y|$.
On en déduit que $||x|-|y|| \leqslant |x-y|$.
□

Attention : $|x-y|$ n'est pas inférieur à $|x|-|y|$.
En revanche $|x-y| = |x+(-y)| \leqslant |x|+|-y| \leqslant |x|+|y|$.
Cette majoration a transformé le signe $-$ en le signe $+$.

Définition

On appelle distance de x à y le réel $d(x,y) = |y-x| \in \mathbb{R}^+$.

Proposition

Pour tout $x, y, z \in \mathbb{R}$,
$d(x,y) = 0 \Leftrightarrow x = y$ [séparation]
$d(y,x) = d(x,y)$ [symétrie]
$d(x,y) \leqslant d(x,z) + d(z,y)$ [inégalité triangulaire]

dém. :
$d(x,y) = 0 \Leftrightarrow |x-y| = 0 \Leftrightarrow x-y = 0 \Leftrightarrow x = y$,
$d(y,x) = |x-y| = |y-x| = d(x,y)$,
$d(x,y) = |y-x| = |y-z+z-x| \leqslant |y-z|+|z-x| = d(z,y)+d(x,z)$.
□

Définition

On appelle valeur approchée de $a \in \mathbb{R}$ à $\varepsilon \in \mathbb{R}^+$ près tout réel x vérifiant $|x-a| \leqslant \varepsilon$.
On note $a = x$ à ε près.
Si $x \leqslant a$ on dit que x est une valeur approchée par défaut.
Si $x \geqslant a$ on dit que x est une valeur approchée par excès.

Exemple $\pi = 3,14$ à 10^{-2} près par défaut.

Remarque On notera que ce qui vient d'être écris est plus pertinent qu'écrire $\pi \simeq 3,14$.

13.1.4 Partie entière d'un réel

Définition

Pour $x \in \mathbb{R}$, le plus grand entier n vérifiant $n \leqslant x$ est appelé partie entière de x.
On le note $E(x)$, $[x]$ ou $\lfloor x \rfloor$.

Exemple $E(\pi) = 3$ et $E(-\pi) = -4$.

Proposition

La fonction $x \mapsto E(x)$ est croissante.

dém. :
Supposons $x \leqslant y$.
Puisque $E(x) \leqslant x$ on $E(x) \leqslant y$.
Ainsi $E(x)$ est un entier inférieur à y.
Or $E(y)$ est le plus grand entier inférieur à y donc $E(x) \leqslant E(y)$.
□

Proposition

Soient $x \in \mathbb{R}$ et $n \in \mathbb{Z}$.
On a équivalence entre :
(i) $n = E(x)$;
(ii) $n \leqslant x < n + 1$;
(iii) $x - 1 < n \leqslant x$.

dém. :
(i) $\Rightarrow$ (ii) Supposons $n = E(x)$.
Puisque n est le plus grand entier inférieur à x, on a d'une part $n \leqslant x$ et d'autre part $n + 1$ qui n'est pas inférieur à x donc $n + 1 > x$.
(ii) $\Rightarrow$ (iii) Supposons $n \leqslant x < n + 1$.
On a $n \leqslant x$ et $x - 1 < (n + 1) - 1 = n$ donc $x - 1 < n \leqslant x$.
(iii) $\Rightarrow$ (i) Supposons $x - 1 < n \leqslant x$.
L'entier n est inférieur à x et pour tout $m \in \mathbb{Z}$, si $m > n$ alors $m \geqslant n + 1 > (x - 1) + 1 = x$.
Ainsi n est le plus grand entier inférieur à x et donc $n = E(x)$.
□

Remarque Le positionnement des inégalités larges et strictes nécessite réflexion !

Proposition

Pour tout $n \in \mathbb{Z}$ et tout $x \in \mathbb{R}$,

$$\frac{1}{10^n} E(10^n x) \leqslant x < \frac{1}{10^n} E(10^n x) + \frac{1}{10^n}$$

dém. :
Il suffit d'appliquer la propriété $E(y) \leqslant y < E(y) + 1$ avec $y = 10^n x$.
□

Remarque Les réels $\frac{1}{10^n}E(10^n x)$ et $\frac{1}{10^n}E(10^n x) + \frac{1}{10^n}$ sont appelés des nombres décimaux car ils correspondent au rapport d'un entier par une puissance de 10 ; les nombres décimaux correspondent aux rationnels dont l'écriture décimale est finie.

Définition

Les nombres décimaux $\frac{1}{10^n}E(10^n x)$ et $\frac{1}{10^n}E(10^n x) + \frac{1}{10^n}$ sont appelés parties décimales par défaut et par excès du réel x à la précision $\frac{1}{10^n} = 10^{-n}$.

13.1.5 Intervalles de $\mathbb{R}$

Définition

Soient $a, b \in \mathbb{R}$ tels que $a \leqslant b$.
On appelle intervalles de $\mathbb{R}$ les ensembles suivants :
$[a, b] = \{x \in \mathbb{R}/a \leqslant x \leqslant b\}$, $]a, b] = \{x \in \mathbb{R}/a < x \leqslant b\}$,
$[a, b[= \{x \in \mathbb{R}/a \leqslant x < b\}$, $]a, b[= \{x \in \mathbb{R}/a < x < b\}$,
$]-\infty, a] = \{x \in \mathbb{R}/x \leqslant a\}$, $]-\infty, a[= \{x \in \mathbb{R}/x < a\}$,
$[a, +\infty[= \{x \in \mathbb{R}/x \geqslant a\}$, $]a, +\infty[= \{x \in \mathbb{R}/x > a\}$,
$\mathbb{R}$ et $\emptyset$.
Les intervalles du type $[a, b]$, $[a, +\infty[$, $]-\infty, a]$, $\mathbb{R}$ et $\emptyset$ sont dits des intervalles fermés.
Les intervalles du type $]a, b[$, $]-\infty, a[$, $]a, +\infty[$, $\mathbb{R}$ et $\emptyset$ sont dits des intervalles ouverts.
Les intervalles du type $[a, b[$ et $]a, b]$ sont dits semi-ouverts.
Les intervalles du type $[a, b]$ sont appelés segments.

Théorème

Soit $I \subset \mathbb{R}$. On a équivalence entre :
(i) I est un intervalle de $\mathbb{R}$;
(ii) $\forall a, b \in I$ tels que $a \leqslant b$, on a $[a, b] \subset I$.

dém. :
(i) $\Rightarrow$ (ii) C'est immédiat
(ii) $\Rightarrow$ (i) Supposons (ii)
Si $I = \emptyset$ alors I est un intervalle.
Supposons désormais $I \neq \emptyset$ et considérons un élément $x \in I$.
Posons ensuite $I^+ = I \cap [x, +\infty[$ et $I^- = I \cap]-\infty, x]$ de sorte que $I = I^+ \cup I^-$.
Etudions la partie I^+.
Cas I^+ n'est pas majoré :
Nous allons montrer que $I^+ = [x, +\infty[$.
On a déjà $I^+ \subset [x, +\infty[$.
Inversement, soit $t \in [x, +\infty[$.
Comme I^+ n'est pas majoré par t, il existe $y \in I^+$ tel que $t < y$.
Mais on a alors $x \in I, y \in I$ avec $x \leqslant y$ donc $[x, y] \subset I$.
Or $t \in [x, y]$, donc $t \in I$ puis $t \in I^+$.
Ainsi $[x, +\infty[\subset I^+$ ce qui permet, par double inclusion de conclure $I^+ = [x, +\infty[$.
Cas I^+ est majoré :
Posons $b = \sup I^+$ et montrons que $I^+ = [x, b[$ ou $I^+ = [x, b]$.
Puisque I^+ est majoré par b, on a déjà $I^+ \subset [x, b]$.
Montrons maintenant que $[x, b[\subset I^+$.

Soit $t \in [x, b[$. Comme $t < b$, t n'est pas un majorant de I^+ et donc il existe $y \in I^+$ tel que $t < y$.
Mais on a alors $x \in I, y \in I$ avec $x \leqslant y$ donc $[x, y] \subset I$.
Or $t \in [x, y]$, donc $t \in I$ puis $t \in I^+$. Ainsi $[x, b[\subset I^+ \subset [x, b]$.
Selon que b appartient ou non à I^+ on a : $I^+ = [x, b[$ ou $I^+ = [x, b]$.
Finalement on a montré que I^+ est de l'une des trois formes suivantes : $[x, +\infty[$, $[x, b[$ ou $[x, b]$ (avec $b \in \mathbb{R}$).
En procédant de manière symétrique, on montre que I^- est de l'une des trois formes suivantes :
$]-\infty, x]$, $]a, x]$ ou $[a, x]$ (avec $a \in \mathbb{R}$).
Dans chacun des neuf cas alors possibles, $I = I^+ \cup I^-$ est un intervalle.
□

Remarque Les intervalles correspondent aux parties de $\mathbb{R}$ « sans trous » ; on dit que les intervalles de $\mathbb{R}$ sont ses parties convexes.

Définition

Un intervalle I est dit non singulier s'il contient au moins deux points, i.e. qu'il n'est ni vide, ni réduit à un point.

Proposition

Tout intervalle non singulier contient des nombres rationnels et irrationnels.

dém. :
Soit I un intervalle non singulier et $a < b$ deux points de I.
Déterminer $r = p/q \in \mathbb{Q}$ tel que $a < r < b$ revient à déterminer q tel que l'intervalle $]qa, qb[$ contient un entier.
Soit $q \in \mathbb{N}^\star$ tel que $q(b - a) > 1$ et $p = E(qa) + 1$.
On a $qa < p < qb$ donc $r = p/q \in]a, b[\subset I$.
Ainsi, l'intervalle I contient un nombre rationnel.
Pour montrer l'existence d'irrationnels dans I, exploitons l'irrationalité connue de $\sqrt{2}$.
Par le même raisonnement qu'au dessus, on peut affirmer qu'il existe $r' \in \mathbb{Q}$ tel que $a + \sqrt{2} < r' < b + \sqrt{2}$.
En considérant $s = r' - \sqrt{2}$ on a $a < s < b$ donc $s \in I$ avec $s \in \mathbb{R} \backslash \mathbb{Q}$.
□

13.1.6 Congruence dans $\mathbb{R}$

Soit $\alpha > 0$.

Définition

On dit qu'un réel x est congru à un réel y modulo α s'il existe $k \in \mathbb{Z}$ tel que $x = y + k\alpha$.
On note alors
$$x = y \quad [\alpha]$$

Exemple $7 = 1 \quad [3]$, $-3 = 2 \quad [5]$ et $-\pi = \pi \quad [2\pi]$.

Exemple Les mesures d'angles orientés sont des réel égaux modulo 2π.

Proposition

Soient $x, y, z \in \mathbb{R}$.
On a $x = x \quad [\alpha]$
Si $x = y \quad [\alpha]$ alors $y = x \quad [\alpha]$.
Si $x = y \quad [\alpha]$ et $y = z \quad [\alpha]$ alors $x = z \quad [\alpha]$.

dém. :
$x = x + 0.\alpha$ donc $x = x \quad [\alpha]$.
Si $x = y \quad [\alpha]$ alors on peut écrire $x = y + k\alpha$ avec $k \in \mathbb{Z}$ et alors $y = x - k\alpha$ avec $-k \in \mathbb{Z}$ donc $y = x \quad [\alpha]$.
Si $x = y \quad [\alpha]$ et $y = z \quad [\alpha]$ alors on peut écrire $x = y + k\alpha$ et $y = z + \ell\alpha$ avec $k, \ell \in \mathbb{Z}$ et alors $x = z + (k + \ell)\alpha$ avec $k + \ell \in \mathbb{Z}$ donc $x = z \quad [\alpha]$.
□

Proposition

Soient $x, y, x', y' \in \mathbb{R}$.
Si $x = x' \quad [\alpha]$ et $y = y' \quad [\alpha]$ alors

$$x + y = x' + y' \quad [\alpha]\,,\ -x = -x' \quad [\alpha] \text{ et } \forall \lambda > 0, \lambda x = \lambda x' \quad [\lambda\alpha]$$

dém. :
On peut écrire $x = x' + k\alpha$ et $y = y' + \ell\alpha$ avec $k, \ell \in \mathbb{Z}$.
On a alors $x + y = x' + y' + (k + \ell)\alpha$ avec $k + \ell \in \mathbb{Z}$ donc $x + y = x' + y' \quad [\alpha]$.
Aussi $-x = -x' + (-k)\alpha$ avec $-k \in \mathbb{Z}$ donc $-x = -x' \quad [\alpha]$.
Enfin $\lambda x = \lambda x' + k(\lambda\alpha)$ avec $k \in \mathbb{Z}$ donc $\lambda x = \lambda x' \quad [\lambda\alpha]$.
□

Attention : On ne peut pas multiplier entre elles les relations de congruences réelles.

Exemple Résolvons l'équation

$$5x = x - 2 \quad [4]$$

En ajoutant $-x$ de part et d'autre,

$$5x = x - 2 \quad [4] \Leftrightarrow 4x = -2 \quad [4]$$

En multipliant par $1/4$,

$$5x = x - 2 \quad [4] \Leftrightarrow x = -1/2 \quad [1]$$

La solution générale de l'équation étudiée est donc $x = -\dfrac{1}{2} + k$ avec $k \in \mathbb{Z}$.

13.1.7 Droite numérique achevée

On forme un nouvel ensemble $\bar{\mathbb{R}}$ en adjoignant à $\mathbb{R}$ deux nouveaux éléments notés $+\infty$ et $-\infty$.

Définition

$\bar{\mathbb{R}}$ est appelé droite numérique achevée.

On prolonge partiellement l'opération d'addition à $\bar{\mathbb{R}}$ en posant :

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

$$(+\infty) + (+\infty) = +\infty \text{ et } (-\infty) + (-\infty) = -\infty$$

Remarque L'opération $(+\infty) + (-\infty)$ n'est pas définie.

On prolonge partiellement l'opération de multiplication à $\bar{\mathbb{R}}$ en posant :

$$\forall x \in \mathbb{R}^\star, x \times (+\infty) = \begin{cases} +\infty \text{ si } x > 0 \\ -\infty \text{ si } x < 0 \end{cases}, x \times (-\infty) = \begin{cases} -\infty \text{ si } x > 0 \\ +\infty \text{ si } x < 0 \end{cases}$$

$$(+\infty) \times (+\infty) = (-\infty) \times (-\infty) = +\infty \text{ et } (+\infty) \times (-\infty) = -\infty$$

Remarque Les opérations $0 \times (+\infty)$ et $0 \times (-\infty)$ ne sont pas définies.

Enfin on prolonge à $\bar{\mathbb{R}}$ la relation d'ordre $\leqslant$ en posant :

$$\forall x \in \mathbb{R}, -\infty \leqslant x, x \leqslant +\infty \text{ et } -\infty \leqslant +\infty$$

13.2 Nombres complexes

13.2.1 Présentation de $\mathbb{C}$

$\mathbb{C}$ est un ensemble contenant $\mathbb{R}$ dont les éléments sont appelés nombres complexes.
$\mathbb{C}$ est muni de deux opérations $+$ et $\times$ prolongeant les opérations existant sur $\mathbb{R}$ et conservant leurs propriétés.
$\mathbb{C}$ ne peut pas être muni d'une relation d'ordre « intéressante »mais en revanche, sur $\mathbb{C}$ l'équation $x^2 = -1$ possède deux solutions opposées notées i et $-i$.
Par construction de $\mathbb{C}$, on a

$$\forall z \in \mathbb{C}, \exists!(a, b) \in \mathbb{R}^2, z = a + i.b$$

Définition

L'écriture $z = a + i.b$ est appelée forme algébrique du complexe z.
Les réels a et b sont respectivement appelés parties réelle et imaginaire de z.
On note $a = \mathrm{Re}(z)$ et $b = \mathrm{Im}(z)$.

Remarque La partie imaginaire d'un complexe est un nombre réel !

Proposition

$\forall z, z' \in \mathbb{C}, \mathrm{Re}(z + z') = \mathrm{Re}(z) + \mathrm{Re}(z')$ et $\mathrm{Im}(z + z') = \mathrm{Im}(z) + \mathrm{Im}(z')$.
$\forall z, z' \in \mathbb{C}, \mathrm{Re}(zz') = \mathrm{Re}(z)\mathrm{Re}(z') - \mathrm{Im}(z)\mathrm{Im}(z')$ et $\mathrm{Im}(zz') = \mathrm{Re}(z)\mathrm{Im}(z') + \mathrm{Re}(z')\mathrm{Im}(z)$.

dém. :
On peut écrire $z = a + ib$ et $z' = a' + ib'$ avec $a, b, a', b' \in \mathbb{R}$.
On a alors $z + z' = (a + a') + i(b + b')$ avec $a + a', b + b' \in \mathbb{R}$.
On en déduit $\text{Re}(z + z') = \text{Re}(z) + \text{Re}(z')$ et $\text{Im}(z + z') = \text{Im}(z) + \text{Im}(z')$.
On a aussi $zz' = (aa' - bb') + i(ab' + a'b)$ avec $aa' - bb', ab' + a'b \in \mathbb{R}$.
On en déduit $\text{Re}(zz') = \text{Re}(z)\text{Re}(z') - \text{Im}(z)\text{Im}(z')$ et $\text{Im}(zz') = \text{Re}(z)\text{Im}(z') + \text{Re}(z')\text{Im}(z)$.
□

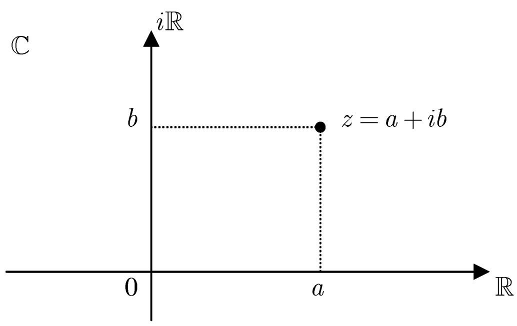
13.2.2 Le plan complexe

Soit $\mathcal{P}$ un plan géométrique rapporté à un repère orthonormé direct $\mathcal{R} = (O; \vec{i}, \vec{j})$.

Définition

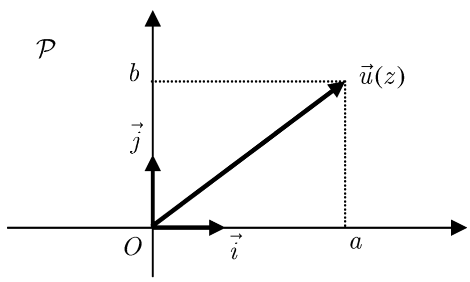
On appelle point d'affixe $z \in \mathbb{C}$ le point M de coordonnées (a, b) avec $a = \text{Re}(z)$ et $b = \text{Im}(z)$.
On note $M(z)$ pour signifier que M est le point d'affixe z.

Remarque Par la notion d'affixe, à tout nombre complexe correspond un point du plan et à tout point correspond un nombre complexe. Cela permet de visualiser $\mathbb{C}$ tel un plan :



Définition

On appelle vecteur d'affixe z le vecteur $\vec{u}$ de composantes (a, b) avec $a = \text{Re}(z)$ et $b = \text{Im}(z)$.
On note $\vec{u}(z)$ pour signifier que $\vec{u}$ est le vecteur d'affixe z.



Proposition

Si $\vec{u}$ est d'affixe z et $\vec{v}$ d'affixe z' alors $\vec{u}+\vec{v}$ est d'affixe $z+z'$.
Si $\vec{u}$ est d'affixe z et λ un réel alors $\lambda.\vec{u}$ est d'affixe λz.
Si M est d'affixe est z alors $\overrightarrow{OM}$ est d'affixe z et inversement.
Si M est d'affixe z et M' d'affixe z' alors $\overrightarrow{MM'}$ est d'affixe $z'-z$.

dém. :
C'est immédiat en raisonnant via les composantes.

□

Proposition

Si M est d'affixe z alors le point M' d'affixe $-z$ est le symétrique de M par rapport à O.

dém. :
Par les affixes, $\overrightarrow{OM} = -\overrightarrow{OM'}$.

□

Proposition

Si M est d'affixe z et $a \in \mathbb{C}$ alors le point M' est l'image de M par la translation de vecteur $\vec{u}$ d'affixe a.

dém. :
Par les affixes, $\overrightarrow{OM'} = \overrightarrow{OM} + \vec{u}$.

□

Proposition

Si M est d'affixe z et $\lambda \in \mathbb{R}$ alors le point M' d'affixe λz est l'image de M par l'homothétie de centre O et de rapport λ.

dém. :
Par les affixes, $\overrightarrow{OM'} = \lambda\overrightarrow{OM}$.

□

13.2.3 Conjugaison

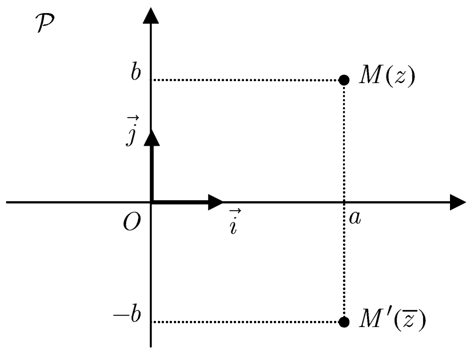
Définition

On appelle conjugué de $z = a + ib \in \mathbb{C}$ (avec $a, b \in \mathbb{R}$) le complexe $\bar{z} = a - i.b$.

Proposition

Si M est d'affixe z alors le point M' d'affixe $\bar{z}$ est le symétrique de M par rapport à l'axe (Ox).

dém. :
Le point M a pour coordonnées (a, b) et le point M' a pour coordonnées $(a, -b)$.



□

Proposition

$\forall z \in \mathbb{C}, \overline{\overline{z}} = z.$

dém. :
On peut écrire $z = a + ib$ avec $a, b \in \mathbb{R}$.
On a alors $\overline{z} = a - ib = a + i(-b)$ avec $a, -b \in \mathbb{R}$ et donc $\overline{\overline{z}} = a + ib = z$.
□

Proposition

$\forall z, z' \in \mathbb{C}, \overline{z + z'} = \overline{z} + \overline{z}', \overline{zz'} = \overline{z}\overline{z}'$ et, si $z \neq 0$, $\overline{\left(\frac{1}{z}\right)} = \frac{1}{\overline{z}}$.

dém. :
On peut écrire $z = a + ib$ et $z' = a' + ib'$ avec $a, b, a', b' \in \mathbb{R}$.
On a $z + z' = (a + a') + i(b + b')$ avec $a + a' \in \mathbb{R}$ et $b + b' \in \mathbb{R}$,
donc $\overline{z + z'} = (a + a) - i(b + b') = \overline{z} + \overline{z}'$.
On a $zz' = (aa' - bb') + i(a'b + ab')$ avec $aa' - bb' \in \mathbb{R}$ et $ab' + a'b \in \mathbb{R}$,
donc $\overline{zz'} = (aa' - bb') - i(a'b + ab') = (a - ib)(a' - ib') = \overline{z}\overline{z}'$.
Enfin, puisque $z \times \frac{1}{z} = 1$, on a $\overline{z} \times \overline{\left(\frac{1}{z}\right)} = \overline{1} = 1$ donc $\overline{\left(\frac{1}{z}\right)} = \frac{1}{\overline{z}}$.
□

Proposition

$\forall z \in \mathbb{C}$,

$$\mathrm{Re}(z) = \frac{1}{2}(z + \overline{z}) \text{ et } \mathrm{Im}(z) = \frac{1}{2i}(z - \overline{z})$$

dém. :
Si $z = a + ib$ avec $a, b \in \mathbb{R}$ alors $\overline{z} = a - ib$ et donc $z + \overline{z} = 2a$ et $z - \overline{z} = 2ib$.
□

Remarque Pour montrer qu'un complexe z est réel, on peut montrer que $\overline{z} = z$.
Pour montrer qu'un complexe z est imaginaire pur, on peut montrer que $\overline{z} = -z$.

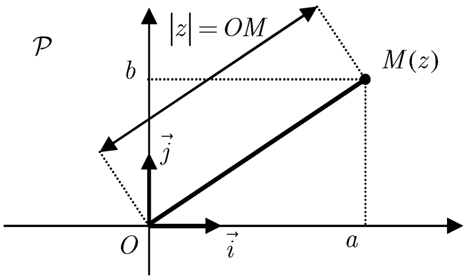
13.2.4 Module

Définition

On appelle module d'un complexe $z = a + ib$ (avec $a, b \in \mathbb{R}$) le réel positif

$$|z| = \sqrt{a^2 + b^2}$$

Remarque Si M est d'affixe z alors le module de z correspond à la longueur OM.



Remarque La notion de module d'un complexe prolonge celle de valeur absolue d'un réel, il n'y a donc pas de conflit dans les notations.

Proposition

$\forall z \in \mathbb{C}, |z| = 0 \Leftrightarrow z = 0.$

dém. :
On peut écrire $z = a + ib$ avec $a, b \in \mathbb{R}$ et on a alors

$$|z| = 0 \Leftrightarrow \sqrt{a^2 + b^2} = 0 \Leftrightarrow a^2 + b^2 = 0 \Leftrightarrow a = b = 0 \Leftrightarrow z = 0$$

□

Proposition

$\forall z, z' \in \mathbb{C}, |z|^2 = z\bar{z} = |\bar{z}|^2, |zz'| = |z|\,|z'|$ et, si $z \neq 0, \left|\dfrac{1}{z}\right| = \dfrac{1}{|z|}$.

dém. :
On peut écrire $z = a + ib$ avec $a, b \in \mathbb{R}$.
On a alors $z\bar{z} = (a + ib)(a - ib) = a^2 + b^2 = |z|^2$ et aussi $z\bar{z} = \bar{z}\bar{\bar{z}} = |\bar{z}|^2$.
Ainsi $|z|^2 = z\bar{z} = |\bar{z}|^2$.
Par suite $|zz'|^2 = zz'\overline{zz'} = z\bar{z}z'\bar{z}' = |z|^2\,|z'|^2$.
Cette égalité de carrés engageant des quantités positives, on en déduit $|zz'| = |z|\,|z'|$.
De plus, si $z \neq 0$, $z \times \dfrac{1}{z} = 1$ donne $|z| \times \left|\dfrac{1}{z}\right| = 1$ et donc $\left|\dfrac{1}{z}\right| = \dfrac{1}{|z|}$.
□

Remarque Par une récurrence immédiate,

$$\forall z \in \mathbb{C}, \forall n \in \mathbb{N},\ |z^n| = |z|^n$$

Par passage à l'inverse,

$$\forall z \in \mathbb{C}^\star, \forall n \in \mathbb{Z},\ |z^n| = |z|^n$$

Proposition

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

De plus, il y a égalité si, et seulement si, les points d'affixe z et z' figurent sur une même demi-droite d'origine O.

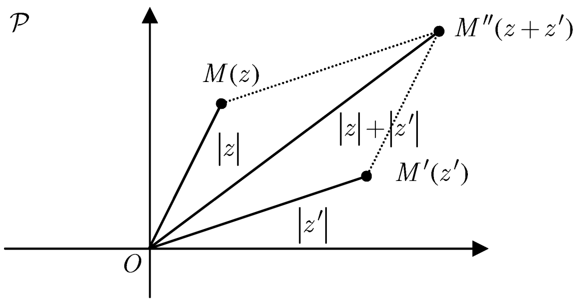
dém. :
Commençons par établir la propriété particulière :
$\forall u \in \mathbb{C}, |1 + u| \leqslant 1 + |u|$ avec égalité si, et seulement si, $u \in \mathbb{R}^+$
Soit $u \in \mathbb{C}$.
On peut écrire $u = a + ib$ avec $a, b \in \mathbb{R}$.
On a alors

$$|1 + u|^2 = (1 + a)^2 + b^2 = 1 + 2a + a^2 + b^2 + 2a + 1 = 1 + 2a + |u|^2$$

Puisque $a = \text{Re}u \leqslant |\text{Re}(u)| \leqslant |u|$, on obtient $|1 + u|^2 \leqslant 1 + 2\,|u| + |u|^2 = (1 + |u|)^2$.
Cette comparaison de carrés engageant des quantités positives, on en déduit $|1 + u| \leqslant 1 + |u|$.
De plus il y a égalité si, et seulement si, $\text{Re}(u) = |\text{Re}(u)|$ et $|\text{Re}(u)| = |u|$ c'est-à-dire $u \in \mathbb{R}^+$.
Démontrons maintenant la propriété de façon générale.
Soient $z, z' \in \mathbb{C}$.
Cas $z = 0$:
On a $|z + z'| = |z'| = |z| + |z'|$.
L'inégalité est donc vraie.
De plus c'est une égalité et les points d'affixes z et z' figurent sur une même demi-droite d'origine O.
Cas $z \neq 0$:
En introduisant $u = z'/z$, on a

$$|z + z'| = |z| \left|1 + \frac{z'}{z}\right| = |z|\,|1 + u| \leqslant |z|\,(1 + |u|) = |z| + |z'|$$

De plus, il y a égalité si, et seulement si, $u \in \mathbb{R}^+$ i.e. $z'/z \in \mathbb{R}^+$.
Ceci signifie que le point d'affixe z' figure sur la demi-droite d'origine O passant par le point d'affixe z.



□

Corollaire

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

dém. :

Par l'inégalité triangulaire,

$$|z| = |z - z' + z'| \leqslant |z - z'| + |z'|$$

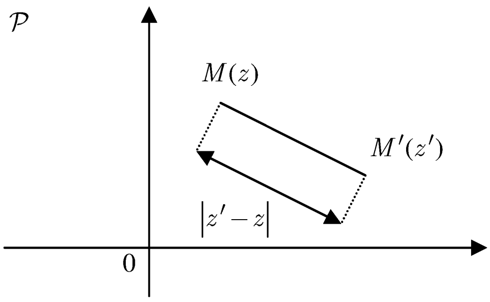
donc $|z| - |z'| \leqslant |z - z'|$. De façon symétrique, $|z'| - |z| \leqslant |z' - z| = |z - z'|$.

□

Définition

On appelle distance entre deux complexes z et z' le réel positif $d(z, z') = |z' - z|$.

Remarque Si M et M' sont les points d'affixes z et z' alors la distance entre z et z' est égale à la longueur MM'.



Proposition

Soient $z, z', z'' \in \mathbb{C}$.
$d(z, z') = 0 \Leftrightarrow z = z'$ [séparation],
$d(z', z) = d(z, z')$ [symétrie],
$d(z, z') \leqslant d(z, z'') + d(z'', z')$ [inégalité triangulaire].

dém. :

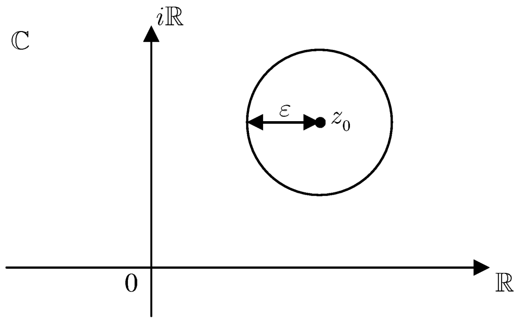
$d(z, z') = 0 \Leftrightarrow |z - z'| = 0 \Leftrightarrow z - z' = 0 \Leftrightarrow z = z'$.
$d(z', z) = |z - z'| = |z' - z| = d(z, z')$.
$d(z, z') = |z' - z| = |z' - z'' + z'' - z| \leqslant |z' - z''| + |z'' - z| = d(z'', z') + d(z, z'')$.

□

Définition

On appelle valeur approchée d'un complexe z_0 à $\varepsilon \in \mathbb{R}^+$ près tout complexe z vérifiant $|z - z_0| \leqslant \varepsilon$.

Remarque Si $z = z_0$ à ε près alors le point d'affixe z appartient au disque de centre le point Ω d'affixe z_0 et de rayon ε.



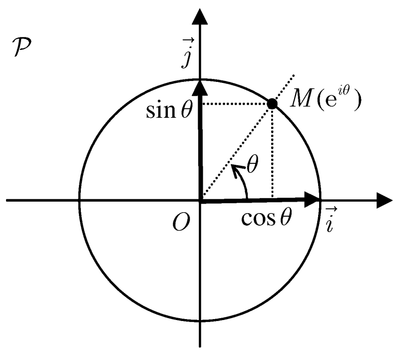
13.2.5 Argument

13.2.5.1 Exponentielle imaginaire

Définition

On appelle exponentielle imaginaire d'angle $\theta \in \mathbb{R}$ le complexe : $\mathrm{e}^{i\theta} = \cos\theta + i.\sin\theta$.

Remarque Le point M d'affixe $\mathrm{e}^{i\theta}$ est le point du cercle trigonométrique déterminé par $(\vec{i}, \overrightarrow{OM}) = \theta \quad [2\pi]$.



Exemple $\mathrm{e}^{i0} = 1, \mathrm{e}^{i\pi/2} = i, \mathrm{e}^{i\pi} = -1, \mathrm{e}^{3i\pi/2} = -i$ et $\mathrm{e}^{2i\pi} = 1$.

Proposition

$\forall\theta \in \mathbb{R}$, $\left|\mathrm{e}^{i\theta}\right| = 1$ et $\dfrac{1}{\mathrm{e}^{i\theta}} = \mathrm{e}^{-i\theta} = \overline{\mathrm{e}^{i\theta}}$.

dém. :

$$\left|\mathrm{e}^{i\theta}\right|^2 = \cos^2\theta + \sin^2\theta = 1$$

et

$$\frac{1}{\mathrm{e}^{i\theta}} = \frac{\overline{\mathrm{e}^{i\theta}}}{|\mathrm{e}^{i\theta}|^2} = \overline{\mathrm{e}^{i\theta}} = \cos\theta - i\sin\theta = \mathrm{e}^{-i\theta}$$

□

Proposition

$\forall \theta, \theta' \in \mathbb{R}$, $\mathrm{e}^{i\theta} = \mathrm{e}^{i\theta'} \Leftrightarrow \theta = \theta' \quad [2\pi]$

dém. :

$$\mathrm{e}^{i\theta} = \mathrm{e}^{i\theta'} \Leftrightarrow \begin{cases} \cos\theta = \cos\theta' \\ \sin\theta = \sin\theta' \end{cases} \Leftrightarrow \theta = \theta' \quad [2\pi]$$

□

Proposition

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

dém. :

$$\mathrm{e}^{i\theta}.\mathrm{e}^{i\theta'} = (\cos\theta + i\sin\theta)(\cos\theta' + i\sin\theta')$$

En développant

$$\mathrm{e}^{i\theta}.\mathrm{e}^{i\theta'} = (\cos\theta\cos\theta' - \sin\theta\sin\theta') + i(\cos\theta\sin\theta' + \sin\theta\cos\theta')$$

puis

$$\mathrm{e}^{i\theta}.\mathrm{e}^{i\theta'} = \cos(\theta+\theta') + i\sin(\theta+\theta') = \mathrm{e}^{i(\theta+\theta')}$$

□

Proposition

$\forall n \in \mathbb{Z}, \forall \theta \in \mathbb{R}$, $\left(\mathrm{e}^{i\theta}\right)^n = \mathrm{e}^{in\theta}$ i.e.

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

dém. :
Par récurrence, montrons que pour $n \in \mathbb{N}$, $\mathrm{e}^{in\theta} = \left(\mathrm{e}^{i\theta}\right)^n$.
Pour $n = 0$, $\mathrm{e}^{i0} = 1 = \left(\mathrm{e}^{i\theta}\right)^0$.
Supposons la propriété établie au rang $n \geqslant 0$.

$$\left(\mathrm{e}^{i\theta}\right)^{n+1} = \left(\mathrm{e}^{i\theta}\right)^n \mathrm{e}^{i\theta} = \mathrm{e}^{in\theta}\mathrm{e}^{i\theta} = \mathrm{e}^{i(n+1)\theta}$$

Récurrence établie.
Pour $n \in \mathbb{Z}^-$, $n = -p$ avec $p \in \mathbb{N}$.

$$\left(\mathrm{e}^{i\theta}\right)^n = \left(\mathrm{e}^{i\theta}\right)^{-p} = \frac{1}{\left(\mathrm{e}^{i\theta}\right)^p} = \frac{1}{\mathrm{e}^{ip\theta}} = \mathrm{e}^{-ip\theta} = \mathrm{e}^{in\theta}$$

□

Attention : Cette formule n'est valable que pour n entier

$$-1 = \mathrm{e}^{i\pi} = (\mathrm{e}^{2i\pi})^{1/2} = \sqrt{1} = 1\,!$$

13.2.5.2 Complexe de module 1

Définition

On note U l'ensemble des complexes de module 1. Ainsi

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

Remarque L'ensemble des point d'affixe dans U correspond géométriquement au cercle trigonométrique du plan.

Exemple $\forall \theta \in \mathbb{R}, \mathrm{e}^{i\theta} \in U$.

Exemple Si $z \in U$ alors $1/z = \bar{z}$.
En effet $z\bar{z} = |z|^2 = 1$.

Théorème

Pour tout $z \in U$, il existe un unique réel θ unique à 2π près tel que $z = \mathrm{e}^{i\theta}$.

dém. :
Unicité à 2π près :
Si $z = \mathrm{e}^{i\theta}$ et $z = \mathrm{e}^{i\theta'}$ alors $\mathrm{e}^{i\theta} = \mathrm{e}^{i\theta'}$ et donc $\theta = \theta' \quad [2\pi]$.
Existence :
Soit $z \in U$.
On peut écrire $z = a + i.b$ avec $a, b \in \mathbb{R}$.
Puisque $|z|^2 = a^2 + b^2 = 1$, on a $a^2 \leqslant 1$ et donc $a \in [-1, 1]$.
Par suite, il existe $\alpha \in \mathbb{R}$ tel que $a = \cos\alpha$.
Puisque $b^2 = 1 - a^2 = \sin^2\alpha$ donc $b = \sin\alpha$ ou $b = -\sin\alpha$.
Si $b = \sin\alpha$ alors $z = \mathrm{e}^{i\theta}$ avec $\theta = \alpha$.
Si $b = -\sin\alpha$ alors $z = \mathrm{e}^{i\theta}$ avec $\theta = -\alpha$.
□

13.2.5.3 Argument d'un nombre complexe non nul

Théorème

Pour tout $z \in \mathbb{C}^\star$, il existe un réel θ unique à 2π près tel que

$$z = |z|\,\mathrm{e}^{i\theta}$$

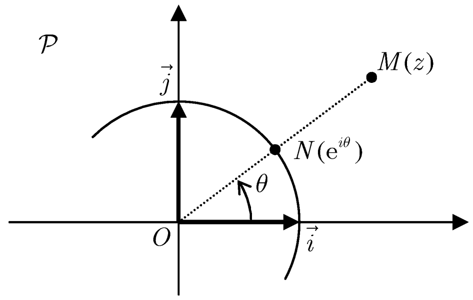
Cette écriture est appelée forme trigonométrique du complexe z.

dém. :
Unicité à 2π près :
Si $z = |z|\,\mathrm{e}^{i\theta}$ et $z = |z|\,\mathrm{e}^{i\theta'}$ alors $\mathrm{e}^{i\theta} = \mathrm{e}^{i\theta'}$ car $|z| \neq 0$ et donc $\theta = \theta' \quad [2\pi]$.
Existence :
Puisque $z/|z| \in U$, il existe $\theta \in \mathbb{R}$ tel que $z/|z| = \mathrm{e}^{i\theta}$ et donc $z = |z|\,\mathrm{e}^{i\theta}$.
□

Définition

Tout réel θ tel que $z = |z|\,\mathrm{e}^{i\theta}$ est appelé argument du complexe z non nul.
On note $\arg(z) = \theta \quad [2\pi]$.

Remarque Si M est le point d'affixe $z \neq 0$ alors $\arg(z) = (\vec{i}, \overrightarrow{OM}) \quad [2\pi]$.



Exemple $\arg(\mathrm{e}^{i\theta}) = \theta \quad [2\pi]$.
En effet $\mathrm{e}^{i\theta} = \left|\mathrm{e}^{i\theta}\right| \mathrm{e}^{i\theta}$

Exemple Déterminons un argument de $z = 1 + i$.
$|z| = \sqrt{2}$ donc

$$1 + i = \sqrt{2}\left(\frac{1}{\sqrt{2}} + i\frac{1}{\sqrt{2}}\right) = \sqrt{2}\left(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4}\right) = \sqrt{2}\mathrm{e}^{i\pi/4}$$

On en déduit

$$\arg(1 + i) = \frac{\pi}{4} \quad [2\pi]$$

Exemple Déterminons un argument de $z = 1 - i\sqrt{3}$.
$|z| = 2$ donc

$$1 - i\sqrt{3} = 2\left(\frac{1}{2} - i\frac{\sqrt{3}}{2}\right) = 2\left(\cos\left(-\frac{\pi}{3}\right) + i\sin\left(-\frac{\pi}{3}\right)\right) = 2\mathrm{e}^{-i\pi/3}$$

On en déduit $\arg(1 + i\sqrt{3}) = -\dfrac{\pi}{3} \quad [2\pi]$.

Exemple Pour $z \in \mathbb{C}^\star$,

$$z \in \mathbb{R}^{+\star} \Leftrightarrow \arg(z) = 0 \quad [2\pi] \text{ et } z \in \mathbb{R}^{-\star} \Leftrightarrow \arg(z) = \pi \quad [2\pi]$$

Par suite

$$z \in \mathbb{R}^\star \Leftrightarrow \arg(z) = 0 \quad [\pi]$$

Proposition

Soient $z, z' \in \mathbb{C}^\star$.
$\arg(\bar{z}) = -\arg(z) \quad [2\pi]$,
$\arg(zz') = \arg(z) + \arg(z') \quad [2\pi]$,
$\arg(-z) = \arg(z) + \pi \quad [2\pi]$,
$\arg(1/z) = -\arg(z) \quad [2\pi]$ et
$\forall n \in \mathbb{Z}, \arg(z^n) = n\arg(z) \quad [2\pi]$.

dém. :
Soient $\theta = \arg(z)$ et $\theta' = \arg(z') \quad [2\pi]$ de sorte que $z = |z|\,\mathrm{e}^{i\theta}$ et $z' = |z'|\,\mathrm{e}^{i\theta'}$.
$\bar{z} = \overline{|z|\mathrm{e}^{i\theta}} = |z|\,\mathrm{e}^{-i\theta} = |\bar{z}|\,\mathrm{e}^{-i\theta}$ donc $\arg(\bar{z}) = -\theta \quad [2\pi]$.
$zz' = |z|\,|z'|\,\mathrm{e}^{i\theta}\mathrm{e}^{i\theta'} = |zz'|\,\mathrm{e}^{i(\theta+\theta')}$ donc $\arg(zz') = \theta + \theta' \quad [2\pi]$.
$-z = -|z|\,\mathrm{e}^{i\theta} = |z|\,\mathrm{e}^{i(\theta+\pi)} = |-z|\,\mathrm{e}^{i(\theta+\pi)}$ donc $\arg(-z) = \theta + \pi \quad [2\pi]$.
$\dfrac{1}{z} = \dfrac{1}{|z|}\dfrac{1}{\mathrm{e}^{i\theta}} = \left|\dfrac{1}{z}\right|\mathrm{e}^{-i\theta}$ donc $\arg(1/z) = -\theta \quad [2\pi]$.
$z^n = |z|^n\left(\mathrm{e}^{i\theta}\right)^n = |z^n|\,\mathrm{e}^{in\theta}$ donc $\arg(z^n) = n\theta \quad [2\pi]$
□

Exemple Déterminons module et argument du complexe $z = \mathrm{e}^{i\theta} + 1$.

$$z = \mathrm{e}^{i\theta} + 1 = (\mathrm{e}^{i\theta/2} + \mathrm{e}^{-i\theta/2})\mathrm{e}^{i\theta/2} = 2\cos\frac{\theta}{2}\mathrm{e}^{i\theta/2}$$

[Factorisation de l'exponentielle équilibrée]
On en déduit

$$|z| = 2\left|\cos\frac{\theta}{2}\right|$$

Si $\cos\theta/2 > 0$ alors $\arg(z) = \theta/2 \quad [2\pi]$.
Si $\cos\theta/2 < 0$ alors $\arg(z) = \theta/2 + \pi \quad [2\pi]$.

13.2.6 Exponentielle complexe

Définition

On appelle exponentielle complexe de $z = a + i.b \in \mathbb{C}$ (avec $a, b \in \mathbb{R}$) le nombre complexe

$$\exp(z) = \mathrm{e}^a.\mathrm{e}^{ib} = \mathrm{e}^a(\cos b + i.\sin b)$$

Remarque Si $z = a \in \mathbb{R}$ alors $\exp(z) = \mathrm{e}^a.\mathrm{e}^{i0} = \mathrm{e}^a$.
L'exponentielle complexe prolonge l'exponentielle réelle connue.
Si $z = i.b \in i.\mathbb{R}$ alors $\exp(z) = \mathrm{e}^0.\mathrm{e}^{ib} = \mathrm{e}^{ib}$.
L'exponentielle complexe prolonge aussi l'exponentielle imaginaire.
On note souvent e^z au lieu de $\exp(z)$.

Théorème

Soient $z, z' \in \mathbb{C}$.
$\overline{\exp(z)} = \exp(\bar{z})$, $\exp(z + z') = \exp(z)\exp(z')$,
$\exp(z) \neq 0$ et $1/\exp(z) = \exp(-z)$.

dém. :
On peut écrire $z = a + ib$ et $z' = a' + ib'$ avec $a, b, a', b' \in \mathbb{R}$.
$\overline{\exp(z)} = \overline{e^a.e^{ib}} = e^a e^{-ib} = \exp(\bar{z})$.
$\exp(z + z') = \exp((a + a') + i(b + b')) = e^{a+a'} e^{i(b+b')} = e^a e^{ib} e^{a'} e^{ib'} = e^z e^{z'}$.
On en déduit $1 = \exp(0) = \exp(z)\exp(-z)$ et donc $\exp(z) \neq 0$ avec $1/\exp(z) = \exp(-z)$.
□

Théorème

Soient $z, z' \in \mathbb{C}$.

$$\exp(z) = \exp(z') \Leftrightarrow \exists k \in \mathbb{Z}, z = z' + 2ik\pi$$

dém. :
($\Leftarrow$) Si $z = z' + 2ik\pi$ avec $k \in \mathbb{Z}$ alors $\exp(z) = \exp(z')\exp(2ik\pi) = \exp(z')$ car $\exp(2ik\pi) = 1$.
($\Rightarrow$) Supposons $\exp(z) = \exp(z')$.
On peut écrire $z = a + ib$ et $z' = a' + ib'$ avec $a, b, a', b' \in \mathbb{R}$.
On a alors $e^a e^{ib} = e^{a'} e^{ib'}$.
En passant cette relation au module, on obtient $e^a = e^{a'}$ et donc $a = a'$ car l'exponentielle réelle prend des valeurs deux à deux distinctes.
Sachant $e^a = e^{a'} \neq 0$, la relation $e^a e^{ib} = e^{a'} e^{ib'}$ donne $e^{ib} = e^{ib'}$ et donc $b = b' \quad [2\pi]$.
Ainsi, il existe $k \in \mathbb{Z}$ tel que $b = b' + 2k\pi$ et donc $z = z' + 2ik\pi$.
□

13.3 Equations et systèmes numériques

13.3.1 Résolution d'une équation

Pour résoudre une équation, on transforme celle-ci par équivalences ou par implications en une ou plusieurs équations dont les solutions sont immédiates.

Attention : Si on procède par implications, il se peut que des solutions « parasites »apparaissent. Il est alors indispensable de vérifier l'exactitude des solutions proposées.

Exemple Résolvons l'équation $x = \sqrt{2 - x^2}$ d'inconnue $x \in \left[-\sqrt{2}, \sqrt{2}\right]$.
Erreur à ne pas commettre, écrire $x = \sqrt{2 - x^2} \Leftrightarrow x^2 = 2 - x^2$.
En effet l'élévation est au carré donne une implication $\Rightarrow$ et non une équivalence.
1ère méthode : Démarche par implications.

$$\begin{aligned} x = \sqrt{2 - x^2} \quad &\Rightarrow x^2 = 2 - x^2 \\ &\Leftrightarrow 2x^2 = 2 \\ &\Leftrightarrow x^2 = 1 \\ &\Leftrightarrow x = 1 \text{ ou } x = -1 \end{aligned}$$

Vérification : $x = 1$ est solution mais $x = -1$ ne l'est pas.
2ème méthode : Démarche par équivalences :

$$\begin{aligned} x = \sqrt{2 - x^2} \quad &\Rightarrow x^2 = 2 - x^2 \text{ et } x \geqslant 0 \\ &\Leftrightarrow 2x^2 = 2 \text{ et } x \geqslant 0 \\ &\Leftrightarrow x^2 = 1 \text{ et } x \geqslant 0 \\ &\Leftrightarrow x = 1 \end{aligned}$$

Exemple Résolvons l'équation $\mathrm{e}^{2x} - \mathrm{e}^x - 2 = 0$ d'inconnue $x \in \mathbb{R}$.
Posons $t = \mathrm{e}^x$.

$$\mathrm{e}^{2x} - \mathrm{e}^x - 2 = 0 \Leftrightarrow t^2 - t - 2 = 0$$

Les solutions de l'équation $t^2 - t - 2 = 0$ sont $t = 2$ et $t = -1$.
Ainsi

$$\begin{aligned} \mathrm{e}^{2x} - \mathrm{e}^x - 2 = 0 \quad &\Leftrightarrow t = 2 \text{ ou } t = -1 \\ &\Leftrightarrow \mathrm{e}^x = 2 \text{ ou } \mathrm{e}^x = -1 \\ &\Leftrightarrow x = \ln 2 \end{aligned}$$

Exemple Résolvons l'équation $\dfrac{z+i}{z-1} = i$ d'inconnue $z \in \mathbb{C} \setminus \{1\}$.

$$\begin{aligned} \frac{z+i}{z-1} = i \quad &\Leftrightarrow z + i = iz - i \\ &\Leftrightarrow z(1-i) = -2i \\ &\Leftrightarrow z = \frac{-2i}{1-i} = 1 - i \end{aligned}$$

La résolution d'une équation de la forme $\dfrac{az+b}{cz+d} = Z$ est analogue, une telle équation est appelée équation homographique.

Remarque Pour rédiger la résolution d'une équation, il est parfois plus habile de rédiger verbalement celle-ci plutôt que d'enchaîner les symboles $\Rightarrow$ et $\Leftrightarrow$. Voyons un exemple proposant une telle mise en forme :

Exemple Résolvons l'équation $x^2 + y^2 = 0$ d'inconnue $(x, y) \in \mathbb{R}^2$.
Soit (x, y) un couple solution.
Puisque $x^2 + y^2 = 0$ avec $x^2, y^2 \geqslant 0$, on a $x^2 = y^2 = 0$ et donc $x = y = 0$
Inversement, le couple $(0, 0)$ est solution.

13.3.2 Résolution d'un système

Pour résoudre un système d'équations, on transforme celui-ci par équivalences ou par implications en un ou plusieurs systèmes permettant d'exprimer l'ensemble solution.
Deux méthodes sont couramment utilisée pour réduire un système
- la méthode par substitution : on exploite une équation pour exprimer une inconnue en fonction des autres et remplacer cette expression dans les autres équations ;
- la méthode par combinaison : en combinant judicieusement plusieurs équations, on en forme de plus simples.

Attention : Si, on procède par implications, il faut vérifier l'exactitude des solutions obtenues.

Exemple Résolvons le système

$$\begin{cases} x^2 + y^2 = 1 \\ xy + y^2 = 0 \end{cases}$$

d'inconnue $(x, y) \in \mathbb{R}^2$.

$$\begin{aligned}
\begin{cases} x^2 + y^2 = 1 \\ xy + y^2 = 0 \end{cases} &\Leftrightarrow \begin{cases} x^2 + y^2 = 1 \\ (x+y)y = 0 \end{cases} \\
&\Leftrightarrow \begin{cases} x^2 + y^2 = 1 \\ x + y = 0 \end{cases} \text{ou} \begin{cases} x^2 + y^2 = 1 \\ y = 0 \end{cases} \\
&\Leftrightarrow \begin{cases} 2x^2 = 1 \\ y = -x \end{cases} \text{ou} \begin{cases} x^2 = 1 \\ y = 0 \end{cases} \\
&\Leftrightarrow \begin{cases} x = 1/\sqrt{2} \text{ ou } x = -1/\sqrt{2} \\ y = -x \end{cases} \text{ou} \begin{cases} x = 1 \text{ ou } x = -1 \\ y = 0 \end{cases}
\end{aligned}$$

Finalement les couples solutions sont

$$\left(\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right), \left(\frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), (1, 0) \text{ et } (-1, 0)$$

Exemple Résolvons le système

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

d'inconnue $(x, y, z) \in \mathbb{R}^3$.
Procédons par substitution.

$$\begin{aligned}
\begin{cases} x + y + z = 1 \\ 2x - y + z = 2 \\ x + y - z = 3 \end{cases} &\Leftrightarrow \begin{cases} x = 1 - y - z \\ -3y - z = 0 \\ -2z = 2 \end{cases} \\
&\Leftrightarrow \begin{cases} x = 1 - y - z \\ y = -z/3 \\ z = -1 \end{cases} \\
&\Leftrightarrow \begin{cases} x = 5/3 \\ y = 1/3 \\ z = -1 \end{cases}
\end{aligned}$$

Exemple Résolvons le système

$$\begin{cases} (1+i)x + iy = 1 \\ (1-i)x - y = 1 \end{cases}$$

d'inconnue $(x, y) \in \mathbb{C}^2$.

Procédons par combinaison.

$$\begin{cases}(1-i)x-y=1\\(1+i)x+iy=1\end{cases}\Leftrightarrow\begin{cases}(1-i)x-y=1\\2(1+i)x=1+i\quad(2)+i(1)\end{cases}$$
$$\Leftrightarrow\begin{cases}y=-(1+i)/2\\x=1/2\end{cases}$$

13.3.3 Résolution de l'équation $z^n = Z$

13.3.3.1 Racine n-ième de l'unité

Dans cette étude, n désigne un naturel non nul.

Définition

On appelle racine n-ième de l'unité tout complexe z vérifiant $z^n = 1$.
On note U_n l'ensemble des racines n-ième de l'unité.

Exemple Pour $n = 1$, $U_1 = \{1\}$.
Pour $n = 2$,

$$z^2 = 1 \Leftrightarrow (z-1)(z+1) = 0$$

donc $U_2 = \{1, -1\}$.
Pour $n = 4$,

$$z^4 = 1 \Leftrightarrow (z^2-1)(z^2+1) = 0 \Leftrightarrow (z-1)(z+1)(z-i)(z+i) = 0$$

donc $U_4 = \{1, i, -1, -i\}$.

Proposition

Si $z \in U_n$ alors $|z| = 1$, $\bar{z} \in U_n$ et $\forall k \in \mathbb{Z}, z^k \in U_n$.

dém. :
$z^n = 1$ donne $|z|^n = 1$ or $|z| \in \mathbb{R}^+$ donc $|z| = 1$.
$z^n = 1$ donne $\overline{z^n} = 1$ et donc $\bar{z}^n = 1$.
Enfin $z^n = 1$ donc $(z^n)^k = 1^k$ puis $(z^k)^n = 1$.
□

Définition

On note $\omega = \mathrm{e}^{\frac{2i\pi}{n}}$ et pour tout $k \in \mathbb{Z}$, $\omega_k = \omega^k = \mathrm{e}^{\frac{2ik\pi}{n}}$.

Proposition

Pour tout $k \in \mathbb{Z}$, les complexes ω_k sont des racines n-ième de l'unité.

dém. :
$\omega^n = \mathrm{e}^{2i\pi}$ donc $\omega \in U_n$ puis pour tout $k \in \mathbb{Z}$, $\omega_k = \omega^k \in U_n$.
□

Rq$\omega_0 = 1, \omega_1 = \omega, \omega_2 = \omega^2, ..., \omega_{n-1} = \omega^{n-1}$, puis
$\omega_n = 1, \omega_{n+1} = \omega, \omega_{n+2} = \omega^2, \ldots, \omega_{2n-1} = \omega^{n-1}$,...

Proposition

$\forall k, \ell \in \mathbb{Z}, \omega_k = \omega_\ell \Leftrightarrow k = \ell \quad [n]$.
Par suite $\{\omega_k / k \in \mathbb{Z}\} = \{\omega_0, \ldots, \omega_{n-1}\}$ et les $\omega_0, ..., \omega_{n-1}$ sont deux à deux distincts.

dém. :

$$\omega_k = \omega_\ell \Leftrightarrow \frac{2k\pi}{n} = \frac{2\ell\pi}{n} \quad [2\pi] \Leftrightarrow k = \ell \quad [n]$$

□

Théorème

Les racines n-ième de l'unité sont exactement les n complexes deux à deux distincts

$$\omega_0, \omega_1, ..., \omega_{n-1}$$

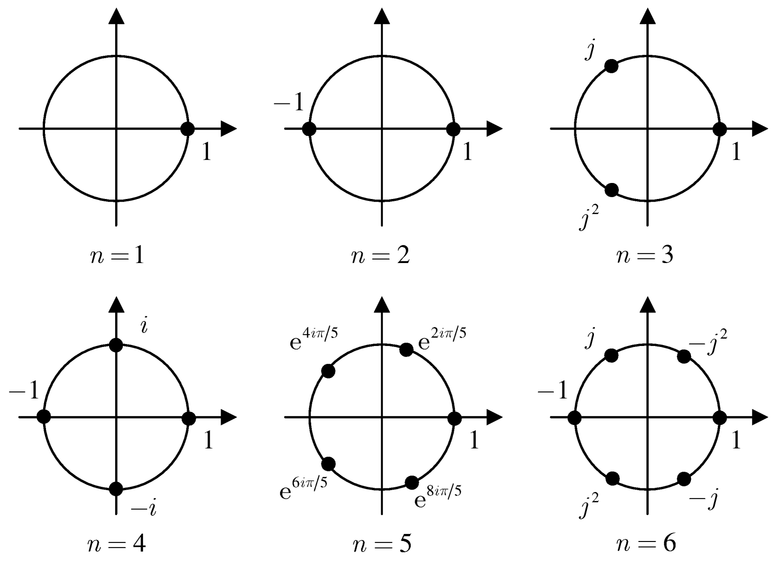
Ainsi U_n est un ensemble à n éléments que l'on peut décrire

$$U_n = \left\{ \mathrm{e}^{2ik\pi/n} / k \in [\![0, n-1]\!] \right\}$$

dém. :
Les $\omega_0, \omega_1, \ldots, \omega_{n-1}$ sont des racines de l'unité deux à deux distinctes.
Inversement, soit $z \in U_n$ et θ un argument de z.
Comme $|z| = 1$, on a $z = \mathrm{e}^{i\theta}$.
La relation $z^n = 1$ donne alors $\mathrm{e}^{in\theta} = 1 = \mathrm{e}^{i0}$ donc $n\theta = 0 \quad [2\pi]$ puis $\theta = 0 \quad [2\pi/n]$.
Ainsi il existe $k \in \mathbb{Z}$ tel que $\theta = \dfrac{2k\pi}{n}$ puis $z \in \{\omega_k / k \in \mathbb{Z}\} = \{\omega_0, \ldots, \omega_{n-1}\}$.

□

Exemple Pour $n = 1$, $\omega = 1$ et $U_1 = \{1\}$.
Pour $n = 2$, $\omega = \mathrm{e}^{i\pi} = -1$ et $U_2 = \{1, -1\}$.
Pour $n = 3$, $\omega = \mathrm{e}^{\frac{2i\pi}{3}} = j$ et $U_3 = \{1, j, j^2\}$.
Pour $n = 4$, $\omega = \mathrm{e}^{\frac{i\pi}{2}} = i$ et $U_4 = \{1, i, -1, -i\}$.
Pour $n = 5$, $\omega = \mathrm{e}^{\frac{2i\pi}{5}}$, $U_5 = \left\{1, \mathrm{e}^{\frac{2i\pi}{5}}, \mathrm{e}^{\frac{4i\pi}{5}}, \mathrm{e}^{\frac{6i\pi}{5}}, \mathrm{e}^{\frac{8i\pi}{5}}\right\}$.
Pour $n = 6$, $\omega = \mathrm{e}^{\frac{i\pi}{3}} = -j^2$, $U_6 = \{1, -j^2, j, -1, j^2, -j\}$.



Remarque Pour $n \geqslant 3$ les racines n-ième de l'unité sont les n sommets d'un polygone régulier inscrit dans le cercle trigonométrique.

Proposition

Pour tout $n \geqslant 2$, la somme des racines n-ième de l'unité est nulle.

dém. :
Posons $S = \omega_0 + \cdots + \omega_{n-1}$.
$\omega S = \omega_1 + \cdots + \omega_n = S$ car $\omega_n = \omega_0$ donc $S = 0$ puisque $\omega \neq 1$.
□

Exemple Calculons $\alpha = \cos \frac{2\pi}{5}$.
Puisque la somme des racines $5^{\text{ème}}$ de l'unité est nulle, on a

$$1 + e^{2i\pi/5} + e^{4i\pi/5} + e^{6i\pi/5} + e^{8i\pi/5} = 0$$

En passant à la partie réelle

$$1 + \cos\frac{2\pi}{5} + \cos\frac{4\pi}{5} + \cos\frac{6\pi}{5} + \cos\frac{8\pi}{5} = 0$$

Or par trigonométrie

$$\cos\frac{8\pi}{5} = \cos\frac{2\pi}{5} \text{ et } \cos\frac{6\pi}{5} = \cos\frac{4\pi}{5} = 2\cos^2\frac{2\pi}{5} - 1$$

En posant $\alpha = \cos\frac{2\pi}{5}$, on obtient

$$2\alpha^2 + 2\alpha - 1 = 0$$

On en déduit $\alpha = \dfrac{-1 \pm \sqrt{5}}{4}$.
Sachant $\cos \dfrac{2\pi}{5} \geqslant 0$, on conclut

$$\cos \frac{2\pi}{5} = \frac{-1+\sqrt{5}}{4}$$

Définition

Comme vu précédemment, on note j la racine cubique de l'unité $\mathrm{e}^{2i\pi/3}$.

Proposition

$j = -\dfrac{1}{2} + i\dfrac{\sqrt{3}}{2}$, $j^3 = 1$, $\bar{j} = j^2$ et $1 + j + j^2 = 0$.

dém. :
$j = \mathrm{e}^{2i\pi/3} = \cos \dfrac{2\pi}{3} + i \sin \dfrac{2\pi}{3} = -\dfrac{1}{2} + i\dfrac{\sqrt{3}}{2}$.
$j^3 = 1$ car j est racine cubique de l'unité.
$\bar{j} = \dfrac{j\bar{j}}{j} = \dfrac{|j|^2}{j} = \dfrac{1}{j} = \dfrac{j^3}{j} = j^2$ et enfin $1 + j + j^2 = 0$ car la somme des racines cubiques de l'unité est nulle.
□

Exemple Simplifions $\dfrac{1+j}{1+\bar{j}}$
$\dfrac{1+j}{1+\bar{j}} = \dfrac{1+j}{1+j^2} = \dfrac{-j^2}{-j} = j$.

13.3.3.2 L'équation $z^n = Z$ avec $Z \in \mathbb{C}$ et $n \in \mathbb{N}^\star$

Remarque Si $Z = 0$ alors l'équation $z^n = 0$ possède pour seule solution $z = 0$.

Théorème

Soit $n \in \mathbb{N}^\star$ et $Z \in \mathbb{C}^\star$.
L'équation $z^n = Z$ possède exactement n solutions distinctes.

dém. :
On peut écrire $Z = |Z|\,\mathrm{e}^{i\theta}$ avec $|Z| \neq 0$ et $\theta = \arg Z \quad [2\pi]$.
Pour $z_0 = \sqrt[n]{|Z|}\mathrm{e}^{i\theta/n} \neq 0$, on a $z_0^n = Z$.
Soit $z \in \mathbb{C}$.

$$\begin{aligned} z^n = Z \quad &\Leftrightarrow z^n = z_0^n \\ &\Leftrightarrow (z/z_0)^n = 1 \\ &\Leftrightarrow \exists k \in \{0, 1, \dots n-1\}, z = z_0\omega_k \end{aligned}$$

Notons $z_k = z_0\omega_k$ pour $k \in \{0, 1, \dots, n-1\}$
Les solutions de l'équation $z^n = Z$ sont les complexes $z_0, z_1, ..., z_{n-1}$ et ces derniers sont deux à deux distinctes car $z_0 \neq 0$ et les $\omega_0, \dots, \omega_{n-1}$ sont deux à deux distincts.
□

Exemple Résolvons dans $\mathbb{C}$ l'équation
$$z^3 = 2 + 2i$$
On a $2 + 2i = 2^{3/2}\mathrm{e}^{i\pi/4}$.
Pour $z_0 = \sqrt{2}\mathrm{e}^{i\pi/12}$, on vérifie $z_0^3 = 2 + 2i$
$$z^3 = 2 + 2i \Leftrightarrow (z/z_0)^3 = 1 \Leftrightarrow z/z_0 = 1, j \text{ ou } j^2$$
Les solutions de l'équation étudiée sont
$$\sqrt{2}\mathrm{e}^{i\pi/12},\ \sqrt{2}\mathrm{e}^{i3\pi/4} \text{ et } \sqrt{2}\mathrm{e}^{i17\pi/12}$$

Exemple Résolvons dans $\mathbb{C}$ l'équation
$$z^n + 1 = 0$$
On a
$$z^n = 1 \Leftrightarrow z^n = -1$$
$-1 = \mathrm{e}^{i\pi}$ donc $z_0 = \mathrm{e}^{i\pi/n}$ est solution de l'équation étudiée.
$$z^n = 1 \Leftrightarrow (z/z_0)^n = 1 \Leftrightarrow z/z_0 = \omega_0, \omega_1, \ldots, \omega_{n-1}$$
avec $\omega_k = \mathrm{e}^{2ik\pi/n}$.
Les solutions de l'équation étudiée sont les $\mathrm{e}^{i(2k+1)\pi/n}$ avec $k \in \{0, 1, \ldots, n-1\}$.

13.3.3.3 Cas particulier : l'équation $z^2 = Z$

Soit $Z \in \mathbb{C}^\star$.
D'après l'étude ci-dessus l'équation $z^2 = Z$ possède deux solutions opposées.
Cas : Z s'exprime sous forme trigonométrique $Z = |Z|\,\mathrm{e}^{i\theta}$ avec $\theta \in \mathbb{R}$.
Les complexes $\sqrt{|Z|}\mathrm{e}^{i\theta/2}$ et $-\sqrt{|Z|}\mathrm{e}^{i\theta/2}$ sont les deux solutions cherchées.
Cas : Z est exprimé sous forme algébrique $Z = a + i.b$ avec $a, b \in \mathbb{R}$.
Cherchons z solution sous la forme $z = x + i.y$ avec $x, y \in \mathbb{R}$.
L'équation $z^2 = Z$ donne $x^2 - y^2 + 2ixy = a + i.b$ d'où
$$\begin{cases} x^2 - y^2 = a & (1) \\ 2xy = b & (2) \end{cases}$$
Ce système est délicat à résoudre sans faire intervenir une troisième équation...
L'équation $z^2 = Z$ donne $|z|^2 = |Z|$ puis l'équation
$$x^2 + y^2 = \sqrt{a^2 + b^2} \quad (3)$$
La combinaison $(1) + (3)$ donne x^2 puis x au signe près.
La combinaison $(3) - (1)$ donne y^2 puis y au signe près.
Enfin, l'équation (2) donne la compatibilité des signes de x et y.
A la fin, on parvient on obtient deux complexes qui sont les seules solutions possibles, or nous savons qu'il y a exactement deux solutions à l'équation étudiée, les complexes déterminés sont donc solutions.

Exemple Résolvons l'équation $z^2 = -3 + 4i$.

On écrit $z = x + iy$ avec $x, y \in \mathbb{R}$.
z est solution de l'équation étudiée si, et seulement si,

$$\begin{cases} x^2 - y^2 = -3 & (1) \\ 2xy = 4 & (2) \\ x^2 + y^2 = 5 & (3) \end{cases}$$

$(1) + (3)$ donne $x^2 = 1$, $(3) - (1)$ donne $y^2 = 4$ et (2) donne $xy > 0$.
On en déduit que les solutions de l'équation sont $z = 1 + 2i$ et $z = -(1 + 2i)$

Exemple Résolvons l'équation $z^2 = 5 - 12i$ ($3 - 2i$).
On écrit $z = x + iy$ avec $x, y \in \mathbb{R}$.
z est solution de l'équation étudiée si, et seulement si,

$$\begin{cases} x^2 - y^2 = 5 & (1) \\ 2xy = -12 & (2) \\ x^2 + y^2 = 13 & (3) \end{cases}$$

$(1) + (3)$ donne $x^2 = 9$, $(3) - (1)$ donne $y^2 = 4$ et (2) donne $xy < 0$.
On en déduit que les solutions de l'équation sont $z = 3 - 2i$ et $z = -3 + 2i$

13.3.4 Résolution de l'équation du second degré

Soient $a, b, c \in \mathbb{C}$ avec $a \neq 0$.
On veut résoudre l'équation du second degré $az^2 + bz + c = 0$ d'inconnue $z \in \mathbb{C}$.
Ecrivons le trinôme $az^2 + bz + c$ sous forme canonique

$$az^2 + bz + c = a\left(z^2 + \frac{bz}{2a} + \frac{c}{a}\right) = a\left(\left(z + \frac{b}{2a}\right)^2 - \frac{\Delta}{4a^2}\right)$$

avec $\Delta = b^2 - 4ac$.
Si $\Delta = 0$ alors

$$z = -\frac{b}{2a}$$

est la seule solution à l'équation étudiée, on dit que c'est une solution double.
Si $\Delta \neq 0$ alors, en considérant $\delta \in \mathbb{C}$ tel que $\Delta = \delta^2$, l'équation possède deux solutions qui sont

$$z_1 = \frac{-b + \delta}{2a} \text{ et } z_2 = \frac{-b - \delta}{2a}$$

Exemple Résolvons l'équation

$$z^2 - (1 + 4i)z + 5(i - 1) = 0$$

On a $\Delta = 5 - 12i = (3 - 2i)^2$.
On en déduit que les solutions de l'équation sont $\dfrac{1 + 4i \pm (3 - 2i)}{2}$ i.e. $2 + i$ et $-1 + 3i$.

Proposition

Soient S et P deux nombres complexes.
Les couples $(x, y) \in \mathbb{C}^2$ solution du système

$$\begin{cases} x + y = S \\ xy = P \end{cases}$$

sont ceux tels que x et y sont les deux racines de l'équation du second degré

$$z^2 - Sz + P = 0$$

dém. :
Si x et y sont les deux solutions de l'équation $z^2 - Sz + P = 0$ alors, à l'ordre près, $x = \dfrac{S-\delta}{2}$ et $y = \dfrac{S+\delta}{2}$ avec $\delta^2 = S^2 - 4P$. On en déduit $x + y = S$ et $xy = P$.
Inversement, si (x, y) est solution du système $\begin{cases} x + y = S \\ xy = P \end{cases}$ alors $(z - x)(z - y) = z^2 - Sz + P$ et donc x et y sont les deux solutions de l'équation $z^2 - Sz + P = 0$.
□

Exemple Résolvons dans $\mathbb{C}^2$ le système

$$\begin{cases} x + y = 3 \\ xy = 2 \end{cases}$$

On introduit l'équation

$$z^2 - 3z + 2 = 0$$

Les solutions de cette équation sont 1 et 2.
Les solutions du système étudié sont donc les couples $(1, 2)$ et $(2, 1)$

Exemple Résolvons dans $\mathbb{C}^\star \times \mathbb{C}^\star$ le système

$$\begin{cases} x + y = i \\ \dfrac{1}{x} + \dfrac{1}{y} = -2i \end{cases}$$

On a

$$\begin{cases} x + y = i \\ \dfrac{1}{x} + \dfrac{1}{y} = -2i \end{cases} \Leftrightarrow \begin{cases} x + y = i \\ \dfrac{x+y}{xy} = -2i \end{cases} \Leftrightarrow \begin{cases} x + y = i \\ xy = -\dfrac{1}{2} \end{cases}$$

On introduit l'équation $z^2 - iz - \dfrac{1}{2} = 0$.
$\Delta = -1 + 2 = 1$, les solutions de cette équation sont $\dfrac{i \pm 1}{2}$.
Les solutions du système étudié sont donc les couples

$$\left(\frac{1-i}{2}, \frac{1+i}{2}\right) \text{ et } \left(\frac{1+i}{2}, \frac{1-i}{2}\right)$$

13.4 Sommes et produits numériques

13.4.1 Sommes numériques

Définition

Si I est un ensemble fini non vide et si pour tout $i \in I$, a_i désigne un nombre complexe (on dit que l'objet $(a_i)_{i\in I}$ est une famille de nombres complexes indexée sur I) alors on note $\sum_{i\in I} a_i$ la somme des a_i pour l'indice i parcourant I.

Remarque Lorsque $I = \emptyset$, on convient que $\sum_{i\in\emptyset} a_i = 0$.

Remarque Lorsque $I = \{m, m+1, \dots, n\}$ (avec $m \leqslant n \in \mathbb{Z}$), cette somme est encore notée

$$\sum_{i=m}^{n} a_i \text{ ou } \sum_{m\leqslant i\leqslant n} a_i$$

Exemple $\sum_{i=1}^{n} i = 1 + 2 + \cdots + n = \frac{n(n+1)}{2}$.
En effet $(1 + 2 + \cdots + n) + (n + (n-1) + \cdots + 1) = (n+1) + (n+1) + \cdots + (n+1) = n(n+1)$ donc $1 + 2 + \cdots + n = \frac{n(n+1)}{2}$

Remarque Dans l'écriture $\sum_{i\in I} a_i$, l'indice i joue un rôle muet. Ainsi $\sum_{i\in I} a_i = \sum_{j\in I} a_j$.

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

Proposition

Avec des notations immédiates,

$$\sum_{i\in I} (a_i + b_i) = \sum_{i\in I} a_i + \sum_{i\in I} b_i \text{ et } \sum_{i\in I} (\lambda a_i) = \lambda \sum_{i\in I} a_i$$

dém. :
La relation $\sum_{i\in I} (a_i + b_i) = \sum_{i\in I} a_i + \sum_{i\in I} b_i$ s'obtient en réorganisant l'ordre des termes sommés.
La relation $\sum_{i\in I} (\lambda a_i) = \lambda \sum_{i\in I} a_i$ se justifie en factorisant par λ les termes de la première somme.
□

Proposition

Si I et J sont des ensembles d'indices disjoints

$$\sum_{i \in I \cup J} a_i = \sum_{i \in I} a_i + \sum_{i \in J} a_i$$

dém. :
La relation $\sum_{i \in I \cup J} a_i = \sum_{i \in I} a_i + \sum_{i \in J} a_i$ se justifie en regroupant les termes sommés par paquets.
□

Attention : En général $\sum_{i \in I} a_i b_i \neq \left(\sum_{i \in I} a_i\right)\left(\sum_{i \in I} b_i\right)$.

13.4.2 Somme arithmétique et géométrique

Proposition

La somme des termes consécutifs d'une suite arithmétique s'obtient par

$$\frac{\text{(1er teme) + (dernier terme)}}{2} \times \text{(nb de termes)}$$

dém. :
Soit $u_0, \ldots, u_n$ les termes consécutifs d'une suite arithmétique de raison r.
Pour $k \in \{0, 1, \ldots, n\}$, $u_k = u_0 + kr$ donc

$$\sum_{k=0}^{n} u_k = \sum_{k=0}^{n} (u_0 + kr) = (n+1)u_0 + r \sum_{k=0}^{n} k = (n+1)u_0 + r\frac{n(n+1)}{2} = \frac{u_0 + u_n}{2}(n+1)$$

□

Proposition

La somme des termes consécutifs d'une suite géométrique de raison $q \in \mathbb{C} \setminus \{1\}$ se détermine par

$$\text{(1er terme)} \times \frac{1 - q^{\text{(nb de termes)}}}{1 - q}$$

dém. :
Soit $u_0, \ldots, u_n$ les termes consécutifs d'une suite géométrique de raison q.
Pour $k \in \{01, \ldots, n\}$, $u_k = u_0 q^k$.

$$(1-q)\sum_{k=0}^{n} u_k = \sum_{k=0}^{n} u_0 q^k - \sum_{k=0}^{n} u_0 q^{k+1} = u_0 - u_0 q^{n+1}$$

donc

$$\sum_{k=0}^{n} u_k = u_0 \frac{1 - q^{n+1}}{1 - q}$$

□

Exemple Pour $q \in \mathbb{C}$,

$$\sum_{k=0}^{n} q^k = \begin{cases} \dfrac{1-q^{n+1}}{1-q} & \text{si } q \neq 1 \\ n+1 & \text{si } q = 1 \end{cases}$$

Exemple Pour $x \in \mathbb{R}$,

$$\sum_{k=0}^{n} (-1)^k x^{2k} = \sum_{k=0}^{n} \left(-x^2\right)^k = \frac{1+(-1)^n x^{2(n+1)}}{1+x^2}$$

car $-x^2 \neq 1$.

Exemple Pour $\theta \in]0, 2\pi[$,

$$\sum_{k=0}^{n} \mathrm{e}^{ik\theta} = \sum_{k=0}^{n} \left(\mathrm{e}^{i\theta}\right)^k = \frac{1-\mathrm{e}^{i(n+1)\theta}}{1-\mathrm{e}^{i\theta}}$$

car $\mathrm{e}^{i\theta} \neq 1$.
On peut en déduire les valeurs des sommes

$$C_n = \sum_{k=0}^{n} \cos(k\theta) \text{ et } S_n = \sum_{k=0}^{n} \sin(k\theta)$$

En effet

$$C_n = \mathrm{Re}\left(\sum_{k=0}^{n} \mathrm{e}^{ik\theta}\right) \text{ et } S_n = \mathrm{Im}\left(\sum_{k=0}^{n} \mathrm{e}^{ik\theta}\right)$$

Or par factorisation d'exponentielle équilibrée,

$$\sum_{k=0}^{n} \mathrm{e}^{ik\theta} = \frac{\mathrm{e}^{i(n+1)\theta}-1}{\mathrm{e}^{i\theta}-1} = \frac{\mathrm{e}^{i(n+1)\theta/2}}{\mathrm{e}^{i\theta/2}} \frac{2i\sin\frac{(n+1)\theta}{2}}{2i\sin\frac{\theta}{2}} = \mathrm{e}^{in\theta} \frac{\sin\frac{(n+1)\theta}{2}}{\sin\frac{\theta}{2}}$$

donc

$$C_n = \cos\frac{n\theta}{2} \frac{\sin\frac{(n+1)\theta}{2}}{\sin\frac{\theta}{2}} \text{ et } S_n = \sin\frac{n\theta}{2} \frac{\sin\frac{(n+1)\theta}{2}}{\sin\frac{\theta}{2}}$$

13.4.3 Exemples de calculs de sommes

Exemple Montrons par récurrence

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

pour tout $n \in \mathbb{N}$.

Pour $n = 0$, la propriété est vérifiée (sachant qu'une somme vide est nulle).
Supposons la propriété vérifiée au rang $n \geqslant 0$.

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

puis en factorisant

$$\sum_{k=1}^{n+1} k^2 = \frac{(n+1)(n+2)(2n+3)}{6}$$

Récurrence établie.

Exemple Calculons $1.2 + 2.3 + \cdots + n.(n+1)$

$$1.2 + 2.3 + \cdots + n.(n+1) = \sum_{k=1}^{n} k(k+1) = \sum_{k=1}^{n} k^2 + \sum_{k=1}^{n} k = \frac{n(n+1)(n+2)}{3}$$

Exemple Calculons $\dfrac{1}{1.2} + \dfrac{1}{2.3} + \cdots + \dfrac{1}{n(n+1)}$.

$$\frac{1}{1.2} + \frac{1}{2.3} + \cdots + \frac{1}{n(n+1)} = \sum_{k=1}^{n} \frac{1}{k(k+1)} = \sum_{k=1}^{n} \left(\frac{1}{k} - \frac{1}{k+1}\right) = \sum_{k=1}^{n} \frac{1}{k} - \sum_{k=1}^{n} \frac{1}{k+1}$$

Après simplification des termes communs aux deux sommes (on parle de télescopage)

$$\frac{1}{1.2} + \frac{1}{2.3} + \cdots + \frac{1}{n(n+1)} = 1 - \frac{1}{n+1} = \frac{n}{n+1}$$

13.4.4 Somme double

Définition

Si I et J sont des ensembles finis non vides et si pour tout $(i,j) \in I \times J$, $a_{i,j}$ désigne un nombre réel ou complexe, la somme $\displaystyle\sum_{(i,j)\in I\times J} a_{i,j}$ est appelée somme double.

Proposition

Avec les notations qui précèdent

$$\sum_{(i,j)\in I\times J} a_{i,j} = \sum_{i\in I} \left(\sum_{j\in J} a_{i,j}\right) = \sum_{j\in J} \left(\sum_{i\in I} a_{i,j}\right)$$

dém. :
Les trois membres de cette identité somment exactement les mêmes termes, ils sont donc égaux.
□

Exemple Pour $m, n \in \mathbb{N}^{\star}$, calculons $\displaystyle\sum_{1 \leqslant i \leqslant n, 1 \leqslant j \leqslant m} (i+j)$.

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

Or $\displaystyle\sum_{j=1}^{m} i = i + i + \cdots + i = mi$ donc

$$\sum_{i=1}^{n} \sum_{j=1}^{m} i = \sum_{i=1}^{n} mi = m \sum_{i=1}^{n} i = \frac{mn(n+1)}{2}$$

De même

$$\sum_{i=1}^{n} \sum_{j=1}^{m} j = \frac{nm(m+1)}{2}$$

On en déduit

$$\sum_{1 \leqslant i \leqslant n, 1 \leqslant j \leqslant m} (i+j) = \frac{mn(m+n+2)}{2}$$

Exemple Pour $n \in \mathbb{N}^{\star}$, calculons $\displaystyle\sum_{1 \leqslant i,j \leqslant n} ij$.

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

Or $\displaystyle\sum_{j=1}^{n} ij = i \sum_{j=1}^{n} j$ donc

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

Or

$$\sum_{i=1}^{n} \left(i \sum_{j=1}^{n} j\right) = \sum_{i=1}^{n} \left(i \times C^{te}\right) = \left(\sum_{i=1}^{n} i\right) C^{te}$$

donc

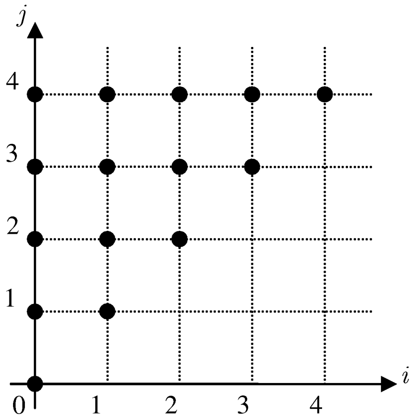
$$\sum_{1 \leqslant i,j \leqslant n} ij = \left(\sum_{i=1}^{n} i\right) \left(\sum_{j=1}^{n} j\right) = \left(\sum_{i=1}^{n} i\right)^{2} = \frac{n^2(n+1)^2}{4}$$

Exemple Calculons $\displaystyle\sum_{1 \leqslant i \neq j \leqslant n} ij$.

$$\sum_{1 \leqslant i \neq j \leqslant n} ij = \sum_{1 \leqslant i,j \leqslant n} ij - \sum_{1 \leqslant i = j \leqslant n} ij$$

Or
$$\sum_{1\leqslant i,j\leqslant n} ij = \frac{n^2(n+1)^2}{4} \text{ et } \sum_{1\leqslant i=j\leqslant n} ij = \sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$$
donc
$$\sum_{1\leqslant i\neq j\leqslant n} ij = \frac{1}{12}(n-1)n(n+1)(3n+2)$$

Exemple Calculons $\displaystyle\sum_{1\leqslant i\leqslant j\leqslant n} (i+j)$.
Les indices (i,j) sur lesquels la sommation est réalisée correspondent au couple suivant



Ces couples peuvent être décrits en faisant varier j de 1 à n puis en faisant varier i de 1 à j.
Ainsi
$$\sum_{1\leqslant i\leqslant j\leqslant n} (i+j) = \sum_{j=1}^{n}\sum_{i=1}^{j}(i+j)$$
Or
$$\sum_{i=1}^{j}(i+j) = \sum_{i=1}^{i} i + \sum_{i=1}^{j} j = \frac{j(j+1)}{2} + j^2 = \frac{3j^2+j}{2}$$
donc
$$\sum_{1\leqslant i\leqslant j\leqslant n} (i+j) = \sum_{j=1}^{n}\frac{3j^2+j}{2} = \frac{3}{2}\sum_{j=1}^{n} j^2 + \frac{1}{2}\sum_{j=1}^{n} j = \frac{n(n+1)^2}{2}$$

Remarque La somme $\displaystyle\sum_{1\leqslant i\leqslant j\leqslant n} \ldots$ aurait aussi pu être réorganisée en $\displaystyle\sum_{i=1}^{n}\sum_{j=i}^{n}\ldots$
car les couples (i,j) peuvent être obtenus en faisant varier i de 1 à n puis en faisant varier j de i à n.
En revanche, on ne pouvait par réorganiser la somme
- en $\displaystyle\sum_{i=1}^{n}\sum_{j=1}^{n}\ldots$ car il y a adjonction de nouveaux termes ;

- ni en $\sum_{i=1}^{j}\sum_{j=1}^{n}$... car celle-ci n'a pas de sens à cause de l'indice j extérieur à la somme en j.

13.4.5 Produits numériques

Définition

Si I est un ensemble fini non vide et si pour tout $i \in I$, a_i désigne un nombre complexe alors on note $\prod_{i\in I} a_i$ le produit des a_i pour l'indice i parcourant I.

Remarque Lorsque $I = \emptyset$, on convient que $\prod_{i\in\emptyset} a_i = 1$.

Remarque Lorsque $I = \{m, m+1, \dots, n\}$ (avec $m \leqslant n \in \mathbb{Z}$), ce produit est encore noté

$$\prod_{i=m}^{n} a_i \text{ ou } \prod_{m\leqslant i\leqslant n} a_i$$

Exemple $n! = 1 \times 2 \times 3 \times \cdots \times n = \prod_{i=1}^{n} i$.
En particulier, $1! = 1, 2! = 2, 3! = 6, 4! = 24, 5! = 120, 6! = 720$.
Aussi $0! = 1$.

Proposition

Avec des notations immédiates,

$$\prod_{i\in I} a_i b_i = \left(\prod_{i\in I} a_i\right)\left(\prod_{i\in I} b_i\right) \text{ et } \prod_{i\in I} a_i^\alpha = \left(\prod_{i\in I} a_i\right)^\alpha$$

dém. :
La relation $\prod_{i\in I} a_i b_i = \left(\prod_{i\in I} a_i\right)\left(\prod_{i\in I} b_i\right)$ s'obtient en réorganisant l'ordre des facteurs multipliés.
La relation $\prod_{i\in I} a_i^\alpha = \left(\prod_{i\in I} a_i\right)^\alpha$ s'obtient en exploitant la propriété $a^\alpha b^\alpha = (ab)^\alpha$.
□

Attention : En général $\prod_{i\in I}(a_i + b_i) \neq \prod_{i\in I} a_i + \prod_{i\in I} b_i$.

Exemple $\prod_{i=1}^{n}(\lambda a_i) = \prod_{i=1}^{n}\lambda \prod_{i=1}^{n} a_i = \lambda^n \prod_{i=1}^{n} a_i$.

13.4.6 Exemples de calculs de produit numérique

Exemple Calculons
$$\prod_{k=1}^{n}\left(1+\frac{1}{k}\right)$$
On a
$$\prod_{k=1}^{n}\left(1+\frac{1}{k}\right)=\prod_{k=1}^{n}\frac{k+1}{k}=\frac{2}{1}\times\frac{3}{2}\times\cdots\times\frac{n+1}{n}$$
Après simplification (on parle de télescopage)
$$\prod_{k=1}^{n}\left(1+\frac{1}{k}\right)=\frac{n+1}{1}=n+1$$

Exemple Calculons $P=\displaystyle\prod_{k=0}^{n}(1+a^{2^k})$ pour $a\in\mathbb{C}\setminus\{1\}$.
$$P=(1+a)(1+a^2)(1+a^4)(1+a^8)\dots(1+a^{2^n})$$
En exploitant $(1-a)(1+a)=1-a^2$, on obtient
$$(1-a)P=(1-a^2)(1+a^2)(1+a^4)\dots(1+a^{2^n})$$
En répétant le procédé
$$(1-a)P=1-a^{2^{n+1}}$$
et finalement $P=\dfrac{1-a^{2^{n+1}}}{1-a}$ car $a\neq 1$.

Exemple Exprimons $\displaystyle\prod_{k=1}^{n}(n+k)$ à l'aide de nombres factoriels
$$\prod_{k=1}^{n}(n+k)=(n+1)(n+2)\dots(2n)$$
En multipliant et en divisant par $1\times 2\times\ldots\times n$
$$\prod_{k=1}^{n}(n+k)=\frac{1\times 2\times\ldots\times(2n)}{1\times 2\times\ldots\times n}=\frac{(2n)!}{n!}$$

Exemple Exprimons $\displaystyle\prod_{k=0}^{n}(2k+1)=\frac{(2n)!}{2^n n!}$ à l'aide de nombres factoriels
$$\prod_{k=0}^{n}(2k+1)=1\times 3\times 5\times\ldots\times(2n+1)$$

En multipliant et en divisant par les entiers pairs intermédiaires

$$\prod_{k=0}^{n}(2k+1)=\frac{1\times 2\times 3\times 4\times 5\times\ldots\times(2n)\times(2n+1)}{2\times 4\times\ldots\times(2n)}$$

Or

$$2\times 4\times\ldots\times(2n)=\prod_{k=1}^{n}(2k)=2^n\prod_{k=1}^{n}k=2^n n!$$

et donc

$$\prod_{k=0}^{n}(2k+1)=\frac{(2n)!}{2^n n!}$$

13.5 Fonctions numériques

I désigne un intervalle de $\mathbb{R}$ non vide et non réduit à un point.
L'objet de cette partie est de mettre en place brièvement des définitions et des résultats qui seront approfondies dans d'autres chapitre de ce cours

13.5.1 Fonctions réelles

13.5.1.1 Définition

Définition

On appelle fonction réelle définie sur I toute application $f : I \to \mathbb{K}$.
On note $\mathcal{F}(I,\mathbb{R})$ l'ensemble des applications de I vers $\mathbb{K}$.

Exemple $f : x \mapsto \sin(1/x)$ est une fonction réelle définie sur $]0,+\infty[$.

Définition

On appelle alors graphe (ou courbe représentative) de f la courbe du plan géométrique d'équation cartésienne $y = f(x)$.

Définition

Une fonction $f : I \to \mathbb{R}$ est dite paire (resp. impaire) si

$$\forall x \in I, -x \in I \text{ et } f(-x) = f(x) \text{ (resp. } f(-x) = -f(x) \text{).}$$

Remarque Le graphe d'une fonction paire est symétrique par rapport à l'axe des ordonnées alors que le graphe d'une fonction impaire est symétrique par rapport à l'origine du repère.

Définition

Une fonction $f : I \to \mathbb{R}$ est dite T périodique (avec $T \in \mathbb{R}$) si

$$\forall x \in I, x + T \in I \text{ et } f(x + T) = f(x)$$

Exemple La fonction $x \mapsto \cos\dfrac{x}{2}$ est 4π périodique.

Définition

Une fonction $f : I \to \mathbb{R}$ est dite croissante si

$$\forall x, y \in I, x \leqslant y \Rightarrow f(x) \leqslant f(y)$$

Une fonction $f : I \to \mathbb{R}$ est dite décroissante si

$$\forall x, y \in I, x \leqslant y \Rightarrow f(x) \geqslant f(y)$$

Une fonction $f : I \to \mathbb{R}$ est dite strictement croissante si

$$\forall x, y \in I, x < y \Rightarrow f(x) < f(y)$$

Une fonction $f : I \to \mathbb{R}$ est dite strictement décroissante si

$$\forall x, y \in I, x < y \Rightarrow f(x) > f(y)$$

13.5.1.2 Limite

Soient une fonction $f : I \to \mathbb{R}$ et a un point de I ou une extrémité, éventuellement infinie, de I.
On définira ultérieurement la notion de limite finie ou infinie de f en a.
Retenons qu'on note indifféremment $\ell = \lim\limits_a f$, $\ell = \lim\limits_{x \to a} f(x)$ ou $f(x) \xrightarrow[x \to a]{} \ell$ pour signifier que $\ell \in \bar{\mathbb{R}}$ est la limite de f en a.
Nous admettons, pour le moment, les résultats « intuitifs »d'opérations et de comparaison de limite.

13.5.1.3 Continuité et dérivabilité

Définition

Une fonction $f : I \to \mathbb{R}$ est dite continue si pour tout $a \in I$, $\lim\limits_{x \to a} f(x)$ existe et vaut $f(a)$.

Remarque Le graphe d'une fonction continue sur un intervalle peut être tracé « sans lever le crayon ».

Théorème

Soient $f : I \to \mathbb{R}$ une fonction continue et $a \leqslant b \in I$.
Tout réel compris entre $f(a)$ et $f(b)$ possède un antécédent par f.

Définition

Une fonction $f : I \to \mathbb{R}$ est dite dérivable si pour tout $a \in I$, la limite quand $x \to a$ de

$$\frac{f(x) - f(a)}{x - a}$$

existe dans $\mathbb{R}$. On la note $f'(a)$ ce qui permet de définir la fonction $f' : I \to \mathbb{R}$.

Remarque $f'(a)$ apparaît comme étant la pente de la tangente à f en a.
Cette dernière a pour équation $y = f'(a)(x - a) + f(a)$.

Remarque Toute fonction dérivable est continue ; l'inverse est faux.

Remarque Rappelons les formules de dérivation

$$(u+v)' = u' + v',\ (uv)' = u'v + uv',\ \left(\frac{u}{v}\right)' = \frac{u'v - uv'}{v^2}$$

$$\forall n \in \mathbb{Z}, (u^n)' = nu'u^{n-1},\ (u \circ v)' = v' \times u' \circ v$$

et en particulier

$$(\mathrm{e}^u)' = u'\mathrm{e}^u,\ (\ln u)' = \frac{u'}{u},\ (\cos u)' = -u' \sin u,\ (\sin u)' = u' \cos u \text{ et } (\sqrt{u})' = \frac{u'}{2\sqrt{u}}$$

Proposition

Soit $f : I \to \mathbb{R}$ une fonction continue.
Si pour tout $x \in I$, $f'(x) \geqslant 0$ (resp. $f'(x) \leqslant 0$) alors f est croissante (resp. décroissante).
Si pour tout $x \in I$, $f'(x) > 0$ (resp. $f'(x) < 0$) sauf pour un nombre fini de valeurs alors f est strictement monotone.

Exemple La fonction $f : x \mapsto x - \ln(1 + x^2)$ est strictement croissante.
En effet f est dérivable et

$$f'(x) = 1 - \frac{2x}{1+x^2} = \frac{(x-1)^2}{1+x^2} > 0$$

sauf pour $x = 1$.

Définition

En itérant le processus de dérivation, on définit la notion de dérivée d'ordre $k \in \mathbb{N}^\star$ de $f : I \to \mathbb{R}$ par $f^{(k)} = ((f')' \ldots)'$ (k fois). On note aussi $f^{(0)} = f$ appelée dérivée d'ordre 0 de f.

Définition

Pour $k \in \mathbb{N}$, on dit que $f : I \to \mathbb{R}$ est de classe $\mathcal{C}^k$ si $f^{(k)}$ existe et est continue.
On dit que f est de classe $\mathcal{C}^\infty$ si f est indéfiniment dérivable.

13.5.1.4 Primitives et intégrales

Définition

On appelle primitive d'une fonction $f : I \to \mathbb{R}$, s'il en existe, toute fonction $F : I \to \mathbb{R}$ dérivable telle que $F' = f$.

Remarque Si F est primitive de f alors les autres primitives de f sont les fonctions de la forme $t \mapsto F(t) + C^{te}$. On note

$$\int f(t)\,\mathrm{d}t = F(t) + C^{te}$$

Exemple En reconnaissant une forme u'/u, on a

$$\int \frac{\mathrm{e}^t}{1+\mathrm{e}^t}\,\mathrm{d}t = \ln(1+\mathrm{e}^t) + C^{te}$$

Théorème

Toute fonction $f : I \to \mathbb{R}$ continue, possède au moins une primitive F.

Définition

Soient $f : I \to \mathbb{R}$ continue et $a, b \in I$. On appelle intégrale de f de a à b le réel

$$\int_a^b f(t)\,\mathrm{d}t = [F(t)]_a^b = F(b) - F(a)$$

où F désigne une primitive de f.

Exemple On a

$$\int_0^1 x^n\,\mathrm{d}x = \left[\frac{1}{n+1}x^{n+1}\right]_0^1 = \frac{1}{n+1}$$

Remarque Le réel $\displaystyle\int_a^b f(t)\,\mathrm{d}t$ se comprend comme l'aire algébrique comprise entre la courbe Γ_f et l'axe des abscisses (Ox).

Théorème

Soient $f, g : I \to \mathbb{R}$ continues et $a, b \in I$.

$$\int_a^b f(t) + g(t)\,\mathrm{d}t = \int_a^b f(t)\,\mathrm{d}t + \int_a^b g(t)\,\mathrm{d}t$$

$$\int_a^b \lambda f(t)\,\mathrm{d}t = \lambda \int_a^b f(t)\,\mathrm{d}t$$

Si $f \leqslant g$ et $a \leqslant b$ alors

$$\int_a^b f(t)\,\mathrm{d}t \leqslant \int_a^b g(t)\,\mathrm{d}t$$

13.5.2 Fonctions complexes

13.5.2.1 Définition

Définition

On appelle fonction complexe définie sur I toute application $f : I \to \mathbb{C}$.
On note $\mathcal{F}(I, \mathbb{C})$ l'ensemble des fonctions complexes définies sur I.

Exemple $f : t \mapsto \dfrac{\mathrm{e}^{it}}{t+i}$ est une fonction complexe définie sur $\mathbb{R}$.

Définition

Pour étudier une fonction $f : I \to \mathbb{C}$, on introduit les fonctions réelles $\text{Re}(f) : I \to \mathbb{R}$ et $\text{Im}(f) : I \to \mathbb{R}$ définie par

$$\forall t \in I, \text{Re}(f)(t) = \text{Re}(f(t)) \text{ et } \text{Im}(f)(t) = \text{Im}(f(t))$$

Exemple Soient $a \in \mathbb{C}$ et $f : \mathbb{R} \to \mathbb{C}$ définie par $f(t) = \mathrm{e}^{at}$.
f est une fonction complexe pour laquelle $\text{Re}(f)(t) = \cos \omega t \, \mathrm{e}^{\alpha t}$ et $\text{Im}(f)(t) = \sin \omega t \, \mathrm{e}^{\alpha t}$ avec $\alpha = \text{Re}(a)$ et $\omega = \text{Im}(a)$.

Définition

On appelle conjuguée d'une fonction complexe $f : I \to \mathbb{C}$, la fonction $\bar{f} : I \to \mathbb{C}$ définie par $\bar{f}(t) = \overline{f(t)}$.

Remarque Pour les fonctions complexes, on peut aussi définir les notions de parité et de périodicité mais on ne peut parler de monotonie.

13.5.2.2 Limite

Soit $f : I \to \mathbb{C}$ une fonction complexe et a un point de I, ou une extrémité éventuellement infinie de I.

Définition

On dit que f tend vers $\ell \in \mathbb{C}$ en a si $|f(t) - \ell| \xrightarrow[t \to a]{} 0$.
On note alors $f(t) \xrightarrow[t \to a]{} \ell$, $\ell = \lim\limits_{t \to a} f(t)$ ou $\ell = \lim\limits_{a} f$.

Exemple On a $\lim\limits_{t \to +\infty} \dfrac{\mathrm{e}^{it}}{t + i} = 0$.
En effet

$$\left| \frac{\mathrm{e}^{it}}{t + i} - 0 \right| = \frac{1}{\sqrt{t^2 + 1}} \to 0$$

Proposition

On a

$$f(t) \xrightarrow[t \to a]{} \ell \Leftrightarrow \begin{cases} \text{Re}(f(t)) \xrightarrow[t \to a]{} \text{Re}(\ell) \\ \text{Im}(f(t)) \xrightarrow[t \to a]{} \text{Im}(\ell) \end{cases}$$

Exemple On a $\lim\limits_{t \to 0} \dfrac{\mathrm{e}^{it} - 1}{t} = i$.
En effet $\text{Re}\left(\dfrac{\mathrm{e}^{it} - 1}{t}\right) = \dfrac{\cos t - 1}{t} \xrightarrow[t \to 0]{} (\cos)'(0) = 0$ et $\text{Im}\left(\dfrac{\mathrm{e}^{it} - 1}{t}\right) = \dfrac{\sin t}{t} \xrightarrow[t \to 0]{} (\sin)'(0) = 1$.

13.5.2.3 Continuité et dérivabilité

A l'aide de ce concept de limite, on peut, tout comme pour les fonctions réelles, définir les notions de continuité, de dérivabilité et de classe d'une fonction complexe. On obtient alors les résultats suivants :

Proposition

> Une fonction complexe $f : I \to \mathbb{C}$ est continue si, et seulement si, les deux fonctions réelles $\mathrm{Re}(f)$ et $\mathrm{Im}(f)$ le sont

Exemple Pour $a = \alpha + i\omega \in \mathbb{C}$ (avec $\alpha, \omega \in \mathbb{R}$), la fonction $f : t \mapsto \mathrm{e}^{at}$ est continue car $t \mapsto \mathrm{e}^{\alpha t} \cos \omega t$ et $t \mapsto \mathrm{e}^{\alpha t} \sin \omega t$ le sont.

Proposition

> Une fonction complexe $f : I \to \mathbb{C}$ est dérivable si, et seulement si, les deux fonctions réelles $\mathrm{Re}(f)$ et $\mathrm{Im}(f)$ le sont.
> De plus, si tel est le cas,
> $$f'(t) = (\mathrm{Re} f)'(t) + i.(\mathrm{Im} f)'(t)$$

Exemple Pour $a = \alpha + i\omega \in \mathbb{C}$ (avec $\alpha, \omega \in \mathbb{R}$), la fonction $f : t \mapsto \mathrm{e}^{at}$ est dérivable car $t \mapsto \mathrm{e}^{\alpha t} \cos \omega t$ et $t \mapsto \mathrm{e}^{\alpha t} \sin \omega t$ le sont.
De plus
$$f'(t) = \left(\mathrm{e}^{\alpha t} \cos \omega t\right)' + i \left(\mathrm{e}^{\alpha} \sin \omega t\right)'$$
En développant ce calcul, on obtient $f'(t) = a\mathrm{e}^{at}$.

Remarque Les formules de dérivation
$$(u + v)' = u' + v',\ (uv)' = u'v + uv',\ \left(\frac{u}{v}\right)' = \frac{u'v - uv'}{v^2}$$
$$(u^n)' = nu'u^{n-1} \text{ (avec } n \in \mathbb{Z} \text{) et } (\mathrm{e}^u) = u'\mathrm{e}^u$$
restent valables pour les fonctions complexes.
On a aussi la formule
$$(\bar{u})' = \overline{u'}$$

Proposition

> Une fonction $f : I \to \mathbb{C}$ est de classe $\mathcal{C}^k$ (avec $k \in \mathbb{N} \cup \{\infty\}$) si, et seulement si, les deux fonctions réelles $\mathrm{Re}(f)$ et $\mathrm{Im}(f)$ le sont.

13.5.2.4 Primitives et intégrales

Définition

> On appelle primitive d'une fonction $f : I \to \mathbb{C}$, s'il en existe, toute fonction $F : I \to \mathbb{C}$ dérivable telle que $F' = f$.

Remarque Si F est primitive de f alors les autres primitives de f sont les fonctions de la forme $t \mapsto F(t) + C^{te}$. On note :

$$\int f(t)\,\mathrm{d}t = F(t) + C^{te}$$

Exemple Pour $a \in \mathbb{C}^\star$,

$$\int \mathrm{e}^{at}\,\mathrm{d}t = \frac{1}{a}\mathrm{e}^{at} + C^{te}$$

Proposition

Toute fonction $f : I \to \mathbb{C}$ continue, possède au moins une primitive F.

Définition

Soient $f : I \to \mathbb{C}$ continue et $a, b \in I$. On appelle intégrale de f de a à b le complexe

$$\int_a^b f(t)\,\mathrm{d}t = [F(t)]_a^b$$

où F désigne une primitive de f.

Exemple $\displaystyle\int_0^{2\pi} \mathrm{e}^{it}\,\mathrm{d}t = \left[\frac{1}{i}\mathrm{e}^{it}\right]_0^{2\pi} = 0.$

13.5.3 Fonctions polynomiales et rationnelles

Ici $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$.

Définition

Une fonction $f : I \to \mathbb{K}$ est dite polynomiale s'il existe $n \in \mathbb{N}$ et $a_0, a_1, \dots, a_n \in \mathbb{K}$ tels que

$$\forall x \in I, f(x) = a_n x^n + \cdots + a_1 x + a_0$$

On note $\mathcal{P}(I, \mathbb{K})$ l'ensemble de ces fonctions.
Lorsque $\mathbb{K} = \mathbb{R}$, on parle de fonction polynomiale réelle.
Lorsque $\mathbb{K} = \mathbb{C}$, on parle de fonction polynomiale complexe.

Exemple $f : x \mapsto x^3 - x + 2$ est une fonction polynomiale réelle

Exemple $f : t \mapsto ix^2 + (i + 1)x - 1$ est une fonction polynomiale complexe.

Exemple Les fonctions constantes sont des fonctions polynomiales.

Proposition

Soient $n \in \mathbb{N}, a_0, a_1, \dots, a_n \in \mathbb{C}$ et $b_0, b_1, \dots, b_n \in \mathbb{C}$.
Si
$$\forall x \in I, a_0 + a_1 x + \cdots + a_n x^n = b_0 + b_1 x + \cdots + b_n x^n$$
alors
$$\forall k \in \{0, 1, \dots, n\}, a_k = b_k$$
Par suite, il n'y a qu'une seule manière de décrire une fonction polynomiale par ses coefficients.

Définition

On appelle degré d'une fonction polynomiale non nulle, la plus grande puissance de la variable intervenant dans sont expression. On le note $\deg f$.

Exemple Pour $f : x \mapsto x^3 + 1000x - 1$, $\deg f = 3$

Définition

Une fonction $f : I \to \mathbb{K}$ est dite rationnelle s'il existe deux fonctions polynomiales $g, h : I \to \mathbb{K}$ telles que
$$\forall x \in I, h(x) \neq 0 \text{ et } f(x) = \frac{g(x)}{h(x)}$$

Exemple $x \mapsto \dfrac{x^3 - 1}{x^2 + 1}$ est une fonction rationnelle réelle définie sur $\mathbb{R}$.

Chapitre 14

Fonctions usuelles

14.1 Bijection

I et J désignent des intervalles non singuliers de $\mathbb{R}$.

14.1.1 Définition

Définition

On dit qu'une fonction $f : I \to J$ est une bijection si tout $y \in J$ admet un unique antécédent par f dans I. En notant $f^{-1}(y)$ cet antécédent, on définit une application $f^{-1} : J \to I$ appelée bijection réciproque de f.

Proposition

Si f est une bijection de I vers J alors

$$\forall x \in I, f^{-1}(f(x)) = x \text{ et } \forall y \in J, f(f^{-1}(y)) = y$$

dém. :
Pour $x \in I$, x est antécédent de $f(x) \in J$ par f et donc $f^{-1}(f(x)) = x$.
Pour $y \in J$, $f^{-1}(y)$ est antécédent de y par f et donc $f(f^{-1}(y)) = y$.
□

Proposition

Si f est une bijection de I vers J alors f^{-1} réalise une bijection de J vers I et

$$(f^{-1})^{-1} = f$$

dém. :
Soit $x \in I$. Montrons que x possède un unique antécédent par f^{-1}.
Unicité : Si $y \in J$ est antécédent de x par f^{-1} alors $x = f^{-1}(y)$ et donc $y = f(x)$.
Existence : Pour $y = f(x) \in J$, $f^{-1}(y) = f^{-1}(f(x)) = x$.
Ainsi f^{-1} réalise une bijection de J vers I.
De plus, ce qui précède donne que pour tout $x \in I$, $(f^{-1})^{-1}(x) = f(x)$ et donc $(f^{-1})^{-1} = f$.
□

Remarque Une fonction continue strictement monotone réalise une bijection entre deux intervalles que l'on peut déterminer à l'aide d'un tableau de variation.

Exemple Soit $f : [0,+\infty[\to \mathbb{R}$ définie par $f(x) = x^2 + 1$.
La fonction f est continue, dérivable et $f'(x) = 2x > 0$ sauf pour $x = 0$; la fonction f est donc strictement croissante.
Son tableau de variation est

x	0		$+\infty$
$f(x)$	1	$\nearrow$	$+\infty$

Par suite f est une bijection de $[0,+\infty[$ vers $[1,+\infty[$.
On peut exprimer à l'aide des fonctions usuelles son application réciproque.
Pour $x \in [0,+\infty[$ et $y \in [1,+\infty[$,

$$y = f(x) \Leftrightarrow y = x^2 + 1 \Leftrightarrow x = \sqrt{y-1}$$

Ainsi

$$f^{-1} : y \mapsto \sqrt{y-1}$$

14.1.2 Propriétés

Théorème

> Si f est une bijection de I vers J alors les graphes de f et f^{-1} sont symétriques par rapport à la droite d'équation $y = x$.

dém. :
Soient $x \in I$ et $y \in J$.
Le point M de coordonnées (x, y) appartient au graphe de f si, et seulement si, $y = f(x)$.
Ceci équivaut encore à l'équation $x = f^{-1}(y)$ qui signifie que le point M' de coordonnées (y, x) appartient au graphe de f^{-1}.
Puisque l'application qui transforme le point M en le point M' est la symétrie par rapport à la droite d'équation $y = x$, les graphes de f et f^{-1} sont symétriques par rapport à cette droite.
□

Corollaire

> Si f est impaire alors f^{-1} l'est aussi.

dém. :
« Par symétrie des graphes par rapport à l'origine ».
□

Corollaire

> Si f est une bijection continue et strictement monotone alors f^{-1} est aussi continue et de même monotonie que f.

dém. :
« Par symétries des graphes de f et f^{-1} ».
□

Remarque Le tableau de variation d'une bijection réciproque f^{-1} ce déduit de celui de f.

Exemple Soit $f : [0, +\infty[\to \mathbb{R}$ définie par

$$f(x) = x^2 + 1$$

On a

x	0		$+\infty$
$f(x)$	1	$\nearrow$	$+\infty$

On en déduit

x	1		$+\infty$
$f^{-1}(x)$	0	$\nearrow$	$+\infty$

Exemple Soit $f :]0, +\infty[\to \mathbb{R}$ définie par

$$f(x) = \frac{1}{x} + 1$$

On a

x	0		$+\infty$
$f(x)$	$+\infty$	$\searrow$	1

On en déduit

x	1		$+\infty$
$f^{-1}(x)$	$+\infty$	$\searrow$	0

Proposition

Si f est une bijection de I vers J dérivable en $x \in I$ et si $f'(x) \neq 0$ alors f^{-1} est dérivable en $y = f(x)$ et

$$(f^{-1})'(y) = \frac{1}{f'(x)} = \frac{1}{f'(f^{-1}(y))}$$

dém. :
La symétrie par rapport à la droite d'équation $y = x$ transforme une droite de pente α en une droite de pente $1/\alpha$.
Par suite la tangente en x au graphe de f qui est de pente $f'(x)$ est transformée en une tangente en $f(x)$ au graphe de f^{-1} de pente $1/f'(x)$.
□

14.1.3 Radicaux

Soit $n \in \mathbb{N}$ tel que $n \geqslant 2$.
Posons $I = \mathbb{R}^+$ lorsque n est pair et $I = \mathbb{R}$ lorsque n est impair
Considérons la fonction $f : I \to \mathbb{R}$ définie par $f(x) = x^n$.
La fonction f est continue, dérivable et $f'(x) > 0$ sauf pour $x = 0$.
Ainsi la fonction f est strictement croissante et on a les tableaux de variations suivants :
Si n est pair

x	0		$+\infty$
$f(x)$	0	$\nearrow$	$+\infty$

Si n est impair

x	$-\infty$		0		$+\infty$
$f(x)$	$-\infty$	$\nearrow$	0	$\nearrow$	$+\infty$

Dans les deux cas f est une bijection de I vers I.

Définition

On note $x \mapsto \sqrt[n]{x}$ l'application réciproque de la bijection $x \mapsto x^n$ de I vers I.

Remarque $x \mapsto \sqrt[n]{x}$ est définie sur $\mathbb{R}^+$ si n est pair et sur $\mathbb{R}$ si n est impair.

Attention : Ne pas simplifier trop rapidement : $\forall x \in \mathbb{R}$, $\sqrt{x^2} = |x|$ et non $= x$!

Théorème

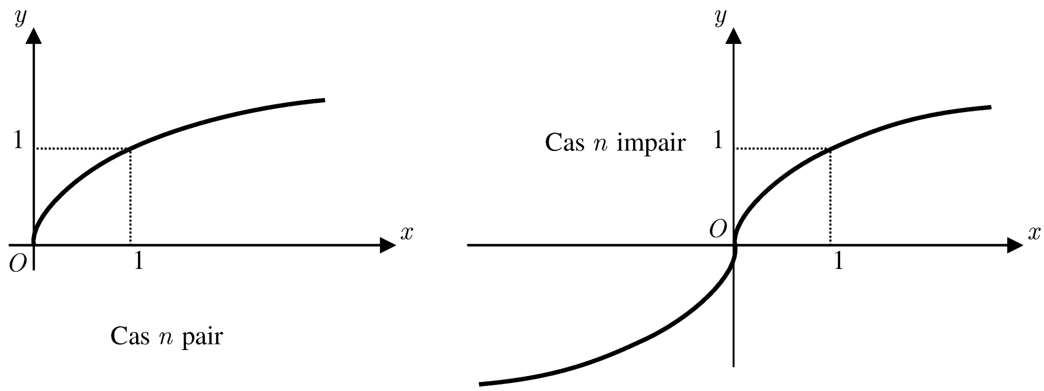
$x \mapsto \sqrt[n]{x}$ est définie et continue sur I.
$x \mapsto \sqrt[n]{x}$ est strictement croissante, $\lim\limits_{x \to +\infty} \sqrt[n]{x} = +\infty$ et pour $x \neq 0$,

$$\left(\sqrt[n]{x}\right)' = \frac{1}{n\left(\sqrt[n]{x}\right)^{n-1}}$$

De plus, quand n est impair, $x \mapsto \sqrt[n]{x}$ est impaire.

dém. :
Ces propriétés découlent de ce que la fonction $x \mapsto \sqrt[n]{x}$ est la bijection réciproque de la fonction $x \mapsto x^n$.
En particulier, pour $t \neq 0$, $(t^n)' = nt^{n-1} \neq 0$ donc $\sqrt[n]{.}$ est dérivable en $x = \sqrt[n]{t}$ et

$$\left(\sqrt[n]{x}\right)' = \frac{1}{(t^n)'} = \frac{1}{nt^{n-1}} = \frac{1}{n\left(\sqrt[n]{x}\right)^{n-1}}$$



□

14.2 Puissances et logarithmes

14.2.1 Logarithme népérien

Définition

On appelle logarithme népérien la fonction $\ln :]0,+\infty[\to \mathbb{R}$ primitive s'annulant en 1 de la fonction continue $x \mapsto 1/x$. Ainsi

L'ensemble demeure fragile mais a progressé

Théorème

La fonction $\ln$ est définie, continue, strictement croissante sur $]0,+\infty[$ et

$$(\ln x)' = 1/x$$

dém. :
Par définition, la fonction $\ln$ est dérivable sur $]0,+\infty[$ et $(\ln x)' = 1/x > 0$.
□

Proposition

$\forall x, y > 0, \ln(xy) = \ln x + \ln y$,
$\forall x > 0, \ln(1/x) = -\ln x$,
$\forall x > 0, \forall n \in \mathbb{Z}, \ln(x^n) = n \ln x$.

dém. :
Soient $y > 0$ et la fonction $f : x \mapsto \ln(xy) - \ln y$ définie sur $]0,+\infty[$.
La fonction f est dérivable,

$$f'(x) = \frac{y}{xy} = \frac{1}{x}$$

et $f(1) = \ln y - \ln y = 0$.
On en déduit que la fonction f est le logarithme népérien.
Ainsi, pour tout $x > 0$, $f(x) = \ln x$ et donc $\ln(xy) = \ln x + \ln y$.
En appliquant cette relation à $y = 1/x$, on obtient $\ln 1 = \ln x + \ln(1/x)$.
Or $\ln 1 = 0$ donc $\ln(1/x) = -\ln x$.
Enfin, par récurrence sur $n \in \mathbb{N}$, on obtient facilement

$$\ln(x^n) = n \ln x$$

et pour $n \in \mathbb{Z}^-$, en écrivant $n = -p$ avec $p \in \mathbb{N}$, on a

$$\ln(x^n) = \ln(1/x^p) = -\ln(x^p) = -p \ln x = n \ln x$$

□

Exemple Soit $n \in \mathbb{N}$ tel que $n \geqslant 2$ et $x > 0$.
On a

$$\ln \sqrt[n]{x} = \frac{1}{n} \ln(\sqrt[n]{x})^n = \frac{1}{n} \ln x$$

Proposition

$\lim_{x\to+\infty} \ln x = +\infty$ et $\lim_{x\to 0+} \ln x = -\infty$.

dém. :
Puisque la fonction $x \mapsto \ln x$ est croissante, celle-ci admet une limite $\ell \in \mathbb{R} \cup \{+\infty\}$ en $+\infty$.
D'une part $\ln 2^n = n \ln 2 \to +\infty$ car $\ln 2 > \ln 1 = 0$.
D'autre part $\ln 2^n \to \ell$ par composition de limites.
Par unicité de la limite, on obtient $\ell = +\infty$ et donc on peut affirmer $\lim_{x\to+\infty} \ln x = +\infty$.
Quand $x \to 0^+$, $1/x \to +\infty$ or $\ln x = -\ln(1/x)$ donc par composition de limites $\ln x \to -\infty$.
□

Remarque Ainsi on a le tableau de variation :

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

On en déduit que la fonction $x \mapsto \ln x$ réalise une bijection de $]0, +\infty[$ sur $\mathbb{R}$.

14.2.2 Exponentielle népérienne

Définition

On appelle exponentiel népérien la fonction $x \mapsto \exp x$ application réciproque de la bijection $x \mapsto \ln x$ de $]0, +\infty[$ vers $\mathbb{R}$.

Théorème

La fonction $\exp$ est définie, continue, strictement croissante sur $\mathbb{R}$, $\lim_{x\to+\infty} \exp x = +\infty$, $\lim_{x\to-\infty} \exp x = 0^+$ et

$$(\exp x)' = \exp x$$

dém. :
En tant qu'application réciproque de la bijection continue strictement croissante $\ln :]0, +\infty[\to \mathbb{R}$, on peut affirmer que la fonction exponentielle est continue, strictement croissante et ses limites en $-\infty$ et $+\infty$ sont respectivement 0 et $+\infty$.
De plus, pour tout $t > 0$, la fonction $\ln$ est dérivable en t et $(\ln)'(t) = 1/t \neq 0$ donc la fonction $\exp$ est dérivable en tout $x = \mathrm{e}^t$ et

$$(\exp x)' = \frac{1}{(\ln)'(t)} = t = \exp(x)$$

□

Proposition

$\forall x, y \in \mathbb{R}, \exp(x + y) = \exp(x)\exp(y)$,
$\forall x \in \mathbb{R}, \exp(-x) = \dfrac{1}{\exp x}$,
$\forall x \in \mathbb{R}, \forall n \in \mathbb{Z}, \exp(nx) = \exp(x)^n$.

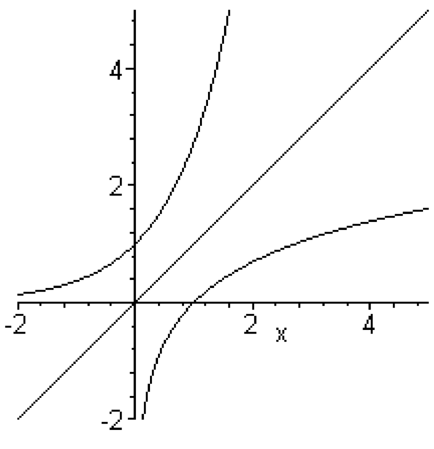
dém. :
$\ln(\exp(x)\exp(y)) = \ln(\exp x) + \ln(\exp y) = x + y$ donc $\exp(x)\exp(y) = \exp(x + y)$.

$\ln(1/\exp(x)) = -\ln(\exp(x)) = -x$ donc $1/\exp(x) = \exp(-x)$.
Enfin, $\ln(\exp(x)^n) = n\ln(\exp x) = nx$ donc $\exp(x)^n = \exp(nx)$.

□

Définition

On appelle nombre de Neper,
le nombre e $= \exp(1) = 2,71828$à 10^{-5} près.



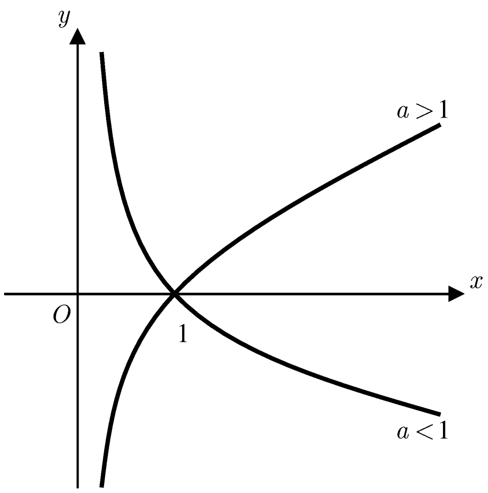
14.2.3 Logarithme en base a

Définition

Soit $a \in]0,+\infty[\setminus \{1\}$. On appelle logarithme en base a l'application $\log_a :]0,+\infty[\to \mathbb{R}$ définie par

$$\log_a(x) = \frac{\ln x}{\ln a}$$

En particulier $\ln = \log_{\mathrm{e}}$.



Exemple Pour $n \in \mathbb{N}^\star$, $E(\log_{10} n) + 1$ est le nombre de décimales permettant décrire le nombre n en base 10.

Proposition

$\forall x, y > 0, \log_a(xy) = \log_a(x) + \log_a(y)$ et $\log_a(a) = 1$.

dém. :
$\log_a(xy) = \dfrac{\ln(xy)}{\ln a} = \dfrac{\ln x + \ln y}{\ln a} = \log_a x + \log_a y$ et $\log_a(a) = \dfrac{\ln a}{\ln a} = 1$.

□

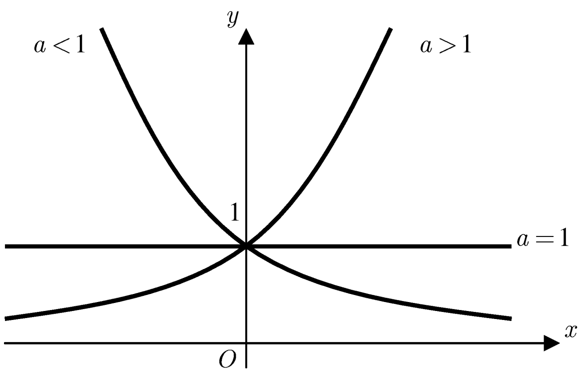
14.2.4 Exponentielle en base a

Définition

On appelle exponentielle en base $a > 0$ l'application $\exp_a : \mathbb{R} \to \mathbb{R}$ définie par

$$\exp_a(x) = \exp(x \ln a)$$

En particulier $\exp = \exp_e$.



Remarque Pour $a \neq 1, \exp_a = (\log_a)^{-1}$ car pour $x > 0$ et $y \in \mathbb{R}$,

$$y = \log_a x \Leftrightarrow \ln x = y \ln a \Leftrightarrow x = \exp_a y$$

Remarque Pour $n \in \mathbb{Z}$, $\exp_a(n) = \exp(n \ln a) = \exp(\ln a^n) = a^n$.
Cette propriété invite à utiliser une nouvelle notation pour manipuler l'exponentielle en base a.

Définition

Pour $x \in \mathbb{R}$ et $a > 0$, on note $a^x = \exp_a(x) = \exp(x \ln a)$.

Exemple $e^x = \exp_e(x) = \exp(x)$.

Exemple Pour $n \in \mathbb{N}$ tel que $n \geqslant 2$ et $a > 0$, $\sqrt[n]{a} = \exp(\ln \sqrt[n]{a}) = \exp(\frac{1}{n} \ln a) = a^{1/n}$.

Attention : Le nombre a^x est défini dans les cas suivants :
- $a > 0$ et $x \in \mathbb{R}$ via $a^x = \exp(x \ln a)$;
- $a \in \mathbb{R}$ et $x = 0$ via la convention $a^0 = 1$;
- $a \in \mathbb{R}$ et $x \in \mathbb{N}^\star$ via $a^x = a \times a \times \cdots \times a$ (x termes) ;
- $a \in \mathbb{R}^\star$ et $x \in \mathbb{Z}^-$ via $a^x = 1/a^{-x}$.

Proposition

$\forall x, y \in \mathbb{R}, \forall a, b > 0,$

$$a^0 = 1,\, 1^x = 1,\, a^x \times a^y = a^{x+y},\, \frac{1}{a^x} = a^{-x}\,,\, (a^x)^y = a^{xy},$$

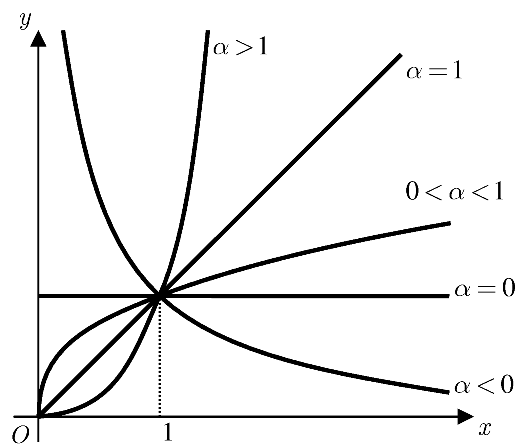
$$(ab)^x = a^x b^x \text{ et } \left(\frac{a}{b}\right)^x = \frac{a^x}{b^x}$$

dém. :
$a^0 = \exp(0 \ln a) = \exp(0) = 1.$
$1^x = \exp(x \ln 1) = \exp(0) = 1.$
$a^x \times a^y = \exp(x \ln a) \exp(y \ln a) = \exp((x+y) \ln a) = a^{x+y}.$
$\dfrac{1}{a^x} = \dfrac{1}{\exp(x \ln a)} = \exp(-x \ln a) = a^{-x}.$
$(a^x)^y = \exp(y \ln a^x) = \exp(xy \ln a) = a^{xy}.$
$(ab)^x = \exp(x \ln(ab)) = \exp(x \ln a + x \ln b) = \exp(x \ln a) \exp(x \ln b) = a^x b^x.$
$\left(\dfrac{a}{b}\right)^x = \exp(x \ln a/b) = \exp(x \ln a - x \ln b) = \exp(x \ln a) \exp(-x \ln b) = a^x b^{-x} = \dfrac{a^x}{b^x}.$
□

14.2.5 Fonctions puissances

Définition

On appelle fonction puissance d'exposant $\alpha \in \mathbb{R}$ l'application $x \mapsto x^\alpha = \mathrm{e}^{\alpha \ln x}$ de $]0, +\infty[$ vers $\mathbb{R}$.



Remarque Si $\alpha = 0$ alors la fonction f est constante égale à 1.
Si $\alpha = 1$ alors f est la fonction identité.

Proposition

Si $\alpha > 0$ alors $\lim\limits_{x \to +\infty} x^\alpha = +\infty$ et $\lim\limits_{x \to 0} x^\alpha = 0$.
Si $\alpha < 0$ alors $\lim\limits_{x \to +\infty} x^\alpha = 0$ et $\lim\limits_{x \to 0} x^\alpha = +\infty$.

dém. :
Pour $\alpha > 0$.
Quand $x \to +\infty$, $x^\alpha = \exp(\alpha \ln x) \to +\infty$ par composition de limites sachant $\alpha \ln x \to +\infty$.
Quand $x \to 0^+$, $x^\alpha = \exp(\alpha \ln x) \to 0$ par composition de limites sachant $\alpha \ln x \to -\infty$.
Pour $\alpha < 0$, c'est analogue
□

Remarque Il est usuel de poser $0^\alpha = 0$ pour $\alpha > 0$ afin de prolonger la fonction $x \mapsto x^\alpha$ par continuité en 0.

14.2.6 Limites de référence

Proposition

$\lim\limits_{x \to +\infty} \dfrac{\ln x}{x} = 0^+$.

dém. :
Pour $t \geqslant 1$, on a $\sqrt{t} \leqslant t$ et donc $\dfrac{1}{t} \leqslant \dfrac{1}{\sqrt{t}}$.
Par suite,

$$0 < \frac{\ln x}{x} = \frac{1}{x}\int_1^x \frac{\mathrm{d}t}{t} \leqslant \frac{1}{x}\int_1^x \frac{\mathrm{d}t}{\sqrt{t}} = \frac{2\sqrt{x}-1}{x}$$

Quand $x \to +\infty$, $\dfrac{2\sqrt{x}-1}{x} \to 0$ et par encadrement $\dfrac{\ln x}{x} \to 0$.
□

Proposition

$\forall \alpha > 0$, $\lim\limits_{x \to +\infty} \dfrac{\ln x}{x^\alpha} = 0^+$ et $\lim\limits_{x \to 0^+} x^\alpha \ln x = 0^-$.

dém. :
Quand $x \to +\infty$, posons $y = x^\alpha \to +\infty$.
$\dfrac{\ln x}{x^\alpha} = \dfrac{\ln y^{1/\alpha}}{y} = \dfrac{1}{\alpha}\dfrac{\ln y}{y} \to 0$ par composition de limites.
Quand $x \to 0^+$, posons $y = 1/x \to +\infty$.
$x^\alpha \ln x = \dfrac{\ln(1/y)}{y^\alpha} = -\dfrac{\ln y}{y^\alpha} \to 0$ par composition de limites.
□

Proposition

$\forall \alpha > 0$, $\lim\limits_{x \to +\infty} \dfrac{\mathrm{e}^x}{x^\alpha} = +\infty$ et $\lim\limits_{x \to +\infty} x^\alpha \mathrm{e}^{-x} = 0^+$.

dém. :
Quand $x \to +\infty$,

$$\frac{\mathrm{e}^x}{x^\alpha} = \exp(x - \alpha \ln x) = \exp\left(x\left(1 - \alpha\frac{\ln x}{x}\right)\right) \to +\infty$$

et

$$x^\alpha e^{-x} = \exp(-x + \alpha \ln x) = \exp\left(-x\left(1 - \alpha \frac{\ln x}{x}\right)\right) \to 0^+$$

□

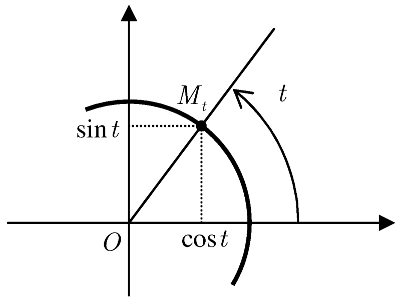
14.3 Fonctions trigonométriques

14.3.1 Fonctions sinus et cosinus

Soit $t \in \mathbb{R}$ et M_t le point du cercle trigonométrique déterminé par $(\vec{i}, \overrightarrow{OM}) = t \quad [2\pi]$.

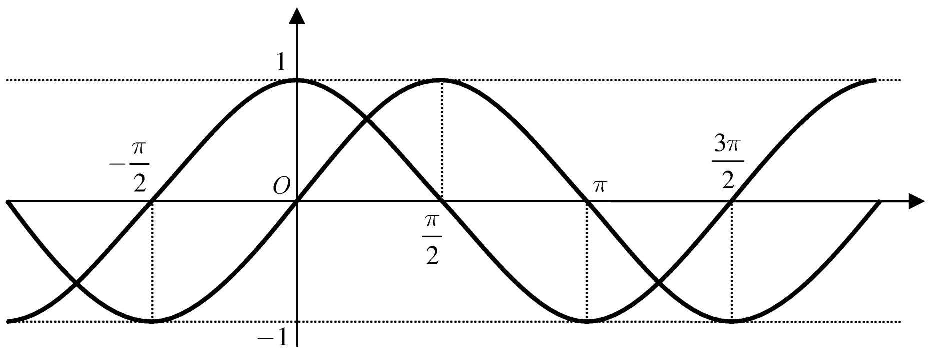
Définition

L'abscisse et l'ordonnée du point M_t sont notées $\cos t$ et $\sin t$.
On définit ainsi deux fonctions cos et sin de $\mathbb{R}$ vers $[-1, 1]$.



$OM_t^2 = 1$ donne $\cos^2 t + \sin^2 t = 1$
$M_{t+2\pi} = M_t$ donne $\cos(t + 2\pi) = \cos t$ et $\sin(t + 2\pi) = \sin t$.
Ainsi les fonctions cos et sin sont 2π périodiques.
$M_{-t} = s_{(Ox)}(M_t)$ donne $\cos(-t) = \cos t$ et $\sin(-t) = -\sin t$.
Ainsi la fonction cos est paire alors que la fonction sin est impaire.
$M_{\pi-t} = s_{(Oy)}(M_t)$ donne $\cos(\pi - t) = -\cos t$ et $\sin(\pi - t) = \sin t$.
$M_{t+\pi} = s_O(M_t)$ donne $\cos(\pi + t) = -\cos t$ et $\sin(\pi + t) = -\sin t$.
On dit que les fonctions cos et sin sont π antipériodiques.
En notant $\Delta : y = x$, $M_{\pi/2-t} = s_\Delta(M_t)$ donne $\cos(\pi/2 - t) = \sin t$ et $\sin(\pi/2 - t) = \cos t$.
On en déduit par les parités des fonctions cos et sin, $\cos(\pi/2 + t) = -\sin t$ et $\sin(\pi/2 + t) = \cos t$.
On admet que les sin et cos sont de classe $\mathcal{C}^\infty$ avec $(\cos t)' = -\sin t$ et $(\sin t)' = \cos t$.
On a le tableau des valeurs

t	0	$\pi/6$	$\pi/4$	$\pi/3$	$\pi/2$	π	$3\pi/2$
$\cos t$	1	$\sqrt{3}/2$	$\sqrt{2}/2$	$1/2$	0	-1	0
$\sin t$	0	$1/2$	$\sqrt{2}/2$	$\sqrt{3}/2$	1	0	-1

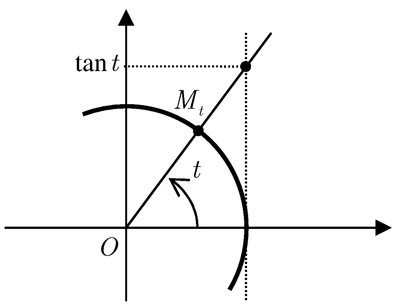


14.3.2 Fonctions tangente et cotangente

Définition

La fonction $\tan$ est définie pour tout $t \neq \pi/2 \quad [\pi]$ par

$$\tan t = \frac{\sin t}{\cos t}$$



Proposition

La fonction $\tan$ est impaire, π périodique, de classe $\mathcal{C}^\infty$ et

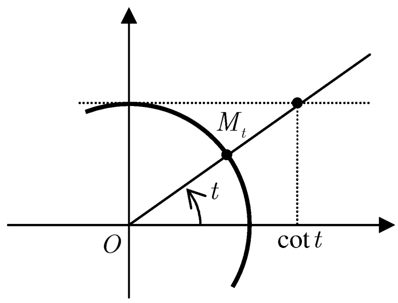
$$(\tan t)' = \frac{1}{\cos^2 t} = 1 + \tan^2 t$$

dém. :
$\tan(-x) = \dfrac{\sin(-x)}{\cos(-x)} = \dfrac{-\sin x}{\cos x} = -\tan x$.
$\tan(x+\pi) = \dfrac{\sin(x+\pi)}{\cos(x+\pi)} = \dfrac{-\sin x}{-\cos x} = -\tan x$.
La fonction $\tan$ est de classe $\mathcal{C}^\infty$ comme rapport de deux fonctions qui le sont
et $(\tan t)' = \left(\dfrac{\sin t}{\cos t}\right)' = \dfrac{\cos^2 t + \sin^2 t}{\cos^2 t} = \dfrac{1}{\cos^2 t}$ ou $= 1 + \tan^2 t$.
□

Définition

La fonction $\cot$ est définie pour tout $t \neq 0 \quad [\pi]$ par

$$\cot t = \frac{\cos t}{\sin t} = \tan\left(\frac{\pi}{2} - t\right)$$



Proposition

La fonction $\cot$ est une fonction impaire, π périodique, de classe $\mathcal{C}^\infty$ et

$$(\cot t)' = \frac{-1}{\sin^2 t} = -1 - \cot^2 t$$

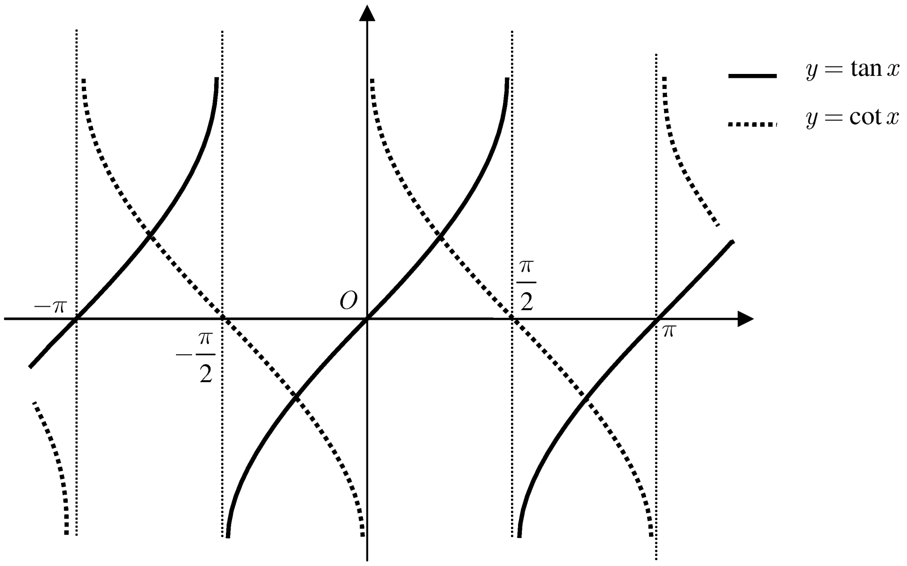
dém. :
Analogue à l'étude de la fonction $\tan$.

□

Remarque Notons que $\forall t \neq 0 \quad [\pi/2]\,, \cot t = \dfrac{1}{\tan t}$ et pour $t \neq \pi/2 \quad [\pi]$, $\cot t = \tan(\pi/2 - t)$.

On a le tableau de valeurs

t	0	$\pi/6$	$\pi/4$	$\pi/3$	$\pi/2$
$\tan t$	0	$1/\sqrt{3}$	1	$\sqrt{3}$	//
$\cot t$	//	$\sqrt{3}$	1	$1/\sqrt{3}$	0



14.3.3 Formules de trigonométrie

14.3.3.1 Développement

Pour tout $a, b \in \mathbb{R}$

$$\cos(a+b) = \cos a \cos b - \sin a \sin b,\ \cos(a-b) = \cos a \cos b + \sin a \sin b$$

$$\sin(a+b) = \sin a \cos b + \sin b \cos a,\ \sin(a-b) = \sin a \cos b - \sin b \cos a$$

En prenant $a = b$, on obtient

$$\cos 2a = 2\cos^2 a - 1 = \cos^2 a - \sin^2 a = 1 - 2\sin^2 a$$

$$\sin 2a = 2 \sin a \cos a$$

Aussi, et sous réserve d'existence

$$\tan(a+b) = \frac{\tan a + \tan b}{1 - \tan a \tan b},\ \tan(a-b) = \frac{\tan a - \tan b}{1 + \tan a \tan b} \text{ et } \tan 2a = \frac{2\tan a}{1 - \tan^2 a}$$

Exemple Développons $\cos 3a$ et $\sin 3a$.

$$\cos 3a = \cos(2a + a) = \cos 2a \cos a - \sin(2a) \sin a$$

donc

$$\cos 3a = 2\cos^3 a - \cos a - 2\sin^2 a \cos a$$

puis

$$\cos 3a = 4\cos^3 a - 3\cos a$$

Aussi

$$\sin 3a = \sin(2a + a) = \sin(2a)\cos a + \cos(2a)\sin a$$

donc
$$\sin 3a = 2\sin a\cos^2 a + \sin a - 2\sin^3 a$$
puis
$$\sin 3a = 3\sin a - 4\sin^3 a$$

14.3.3.2 Factorisation

On a
$$\cos(a+b) + \cos(a-b) = 2\cos a\cos b$$
pour $a = \dfrac{p+q}{2}$ et $b = \dfrac{p-q}{2}$ on obtient,
$$\cos p + \cos q = 2\cos\frac{p+q}{2}\cos\frac{p-q}{2}$$
De même
$$\cos p - \cos q = -2\sin\frac{p+q}{2}\sin\frac{p-q}{2}$$
$$\sin p + \sin q = 2\sin\frac{p+q}{2}\cos\frac{p-q}{2} \text{ et}$$
$$\sin p - \sin q = 2\sin\frac{p-q}{2}\cos\frac{p-q}{2}$$
En particulier, on retient
$$1 + \cos x = 2\cos^2\frac{x}{2} \text{ et } 1 - \cos x = 2\sin^2\frac{x}{2}$$

14.3.3.3 Formules de l'angle moitié

Proposition

Soient $x \neq \pi \quad [2\pi]$ et $t = \tan x/2$.
On a
$$\cos x = \frac{1-t^2}{1+t^2},\ \sin x = \frac{2t}{1+t^2}$$
et quand $x \neq \pi/2 \quad [\pi]$
$$\tan x = \frac{2t}{1-t^2}$$

dém. :
On a
$$\cos^2\frac{x}{2} = \frac{1}{1+\tan^2\frac{x}{2}} = \frac{1}{1+t^2}$$
donc
$$\cos x = 2\cos^2\frac{x}{2} - 1 = \frac{1-t^2}{1+t^2}$$
Aussi
$$\sin x = 2\sin\frac{x}{2}\cos\frac{x}{2} = 2t\cos^2\frac{x}{2} = \frac{2t}{1+t^2}$$

Enfin
$$\tan x = \tan\left(2 \times \frac{x}{2}\right) = \frac{2t}{1-t^2}$$
pour $x \neq \pi/2 \quad [\pi]$.
□

14.3.3.4 Formules d'Euler

$$\forall t \in \mathbb{R}, \mathrm{e}^{it} = \cos t + i.\sin t \text{ et } \mathrm{e}^{-it} = \cos t - i \sin t$$

donc
$$\cos t = \frac{\mathrm{e}^{it} + \mathrm{e}^{-it}}{2} \text{ et } \sin t = \frac{\mathrm{e}^{it} - \mathrm{e}^{-it}}{2i}$$

Exemple Simplifions $\dfrac{\mathrm{e}^{it}+1}{\mathrm{e}^{it}-1}$ pour $t \neq 0 \quad [2\pi]$.
En factorisant par l'exponentielle équilibrée numérateur et dénominateur
$$\frac{\mathrm{e}^{it}+1}{\mathrm{e}^{it}-1} = \frac{\mathrm{e}^{it/2}}{\mathrm{e}^{it/2}} \frac{\mathrm{e}^{it/2} + \mathrm{e}^{-it/2}}{\mathrm{e}^{it/2} - \mathrm{e}^{-it/2}} = \frac{2\cos(t/2)}{2i\sin(t/2)} = -i \cot \frac{t}{2}$$

14.3.3.5 Formule de Moivre

$$\forall n \in \mathbb{Z}, \forall t \in \mathbb{R}, \left(\mathrm{e}^{it}\right)^n = \mathrm{e}^{int}$$

donc
$$\forall n \in \mathbb{Z}, \forall t \in \mathbb{R}, (\cos t + i.\sin t)^n = \cos(nt) + i \sin(nt)$$

Exemple Développons $\cos 3a$ et $\sin 3a$.
$$\cos 3a + i \sin 3a = (\cos a + i \sin a)^3 = \cos^3 a + 3i \cos^2 a \sin a - 3 \cos a \sin^2 a - i \sin^3 a$$
En identifiant parties réelles et imaginaires
$$\cos 3a = \cos^3 a - 3 \cos a \sin^2 a = 4\cos^3 a - 3 \cos a$$
et
$$\sin 3a = 3\cos^2 a \sin a - \sin^3 a = 3 \sin a - 4 \sin^3 a$$

14.3.3.6 Linéarisation

Linéariser une expression trigonométrique consiste à faire disparaître les produits de fonctions cosinus et sinus. Pour cela :
- on réécrit, à l'aide des formules d'Euler, les cosinus et sinus en exponentielles imaginaires ;
- on développe les produits ;
- on recombine termes obtenus en cosinus et sinus.

Exemple Linéarisons $\cos^4 t$.

$$\cos^4 t = \left(\frac{\mathrm{e}^{it} + \mathrm{e}^{-it}}{2}\right)^4 = \frac{1}{16}\left(\mathrm{e}^{4it} + 4\mathrm{e}^{2it} + 6 + 4\mathrm{e}^{-2it} + \mathrm{e}^{-4it}\right)$$

car

$$(a+b)^4 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4$$

Ainsi

$$\cos^4 t = \frac{1}{8}\cos(4t) + \frac{1}{2}\cos(2t) + \frac{3}{8}$$

Exemple Les expressions $\cos^2 t$ et $\sin^2 t$ se linéarisent plus rapidement en exploitant les formules

$$\cos 2a = 2\cos^2 a - 1 = 1 - 2\sin^2 a$$

Ainsi on obtient

$$\cos^2 t = \frac{1}{2}\cos(2t) + \frac{1}{2} \text{ et } \sin^2 t = \frac{1}{2} - \frac{1}{2}\cos(2t)$$

Exemple Les expressions $\cos a \cos b$ et $\sin a \sin b$ se linéarisent plus rapidement en exploitant les formules
$\cos(a+b) + \cos(a-b) = 2\cos a\cos b$ et $\cos(a-b) - \cos(a+b) = 2\sin a \sin b$.
Ainsi on obtient

$$\cos a \cos b = \frac{1}{2}\left(\cos(a+b) + \cos(a-b)\right)$$

et

$$\sin a \sin b = \frac{1}{2}\left(\cos(a-b) - \cos(a+b)\right)$$

14.3.4 Equations trigonométriques

Pour $x, \theta \in \mathbb{R}$,

$$\cos x = \cos\theta \Leftrightarrow \begin{cases} x = \theta \quad [2\pi] \\ \text{ou} \\ x = -\theta \quad [2\pi] \end{cases}$$

$$\sin x = \sin\theta \Leftrightarrow \begin{cases} x = \theta \quad [2\pi] \\ \text{ou} \\ x = \pi - \theta \quad [2\pi] \end{cases}$$

et

$$\tan x = \tan\theta \Leftrightarrow x = \theta \quad [\pi]$$

Exemple Résolvons l'équation $\cos x = \sin 3x$ d'inconnue $x \in \mathbb{R}$.

$$\cos x = \sin 3x \Leftrightarrow \cos x = \cos\left(\frac{\pi}{2} - 3x\right) \Leftrightarrow \begin{cases} x = \frac{\pi}{2} - 3x \quad [2\pi] \\ \text{ou} \\ x = -\frac{\pi}{2} + 3x \quad [2\pi] \end{cases} \Leftrightarrow \begin{cases} x = \frac{\pi}{8} \quad \left[\frac{\pi}{2}\right] \\ \text{ou} \\ x = \frac{\pi}{4} \quad [\pi] \end{cases}$$

Exemple Résolvons l'équation $2\cos^2 x + \sin x = 1$ d'inconnue $x \in \mathbb{R}$.

Ensemble honorable mais qui peut néanmoins s'avérer fragile

Posons $r = \sin x$.

$$2\cos^2 x + \sin x = 1 \Leftrightarrow 2r^2 - r - 1 = 0$$

Les racines de l'équation $2r^2 - r - 1 = 0$ sont $-1/2$ et 1.
On en déduit

$$2\cos^2 x + \sin x = 1 \Leftrightarrow \sin x = -1/2 \text{ ou } \sin x = 1$$
$$\Leftrightarrow x = -\frac{\pi}{6} \quad [2\pi]\,, x = -\frac{5\pi}{6} \quad [2\pi] \text{ ou } x = \frac{\pi}{2} \quad [2\pi]$$

Soient $a, b, c \in \mathbb{R}$, on veut résoudre l'équation $a\cos x + b\sin x = c$ d'inconnue $x \in \mathbb{R}$...

Proposition

On peut transformer $a\cos x + b\sin x$ en $A\cos(x - \varphi)$ avec $A = \sqrt{a^2 + b^2}$ et $\varphi \in \mathbb{R}$ bien choisi.

dém. :
Si $a = b = 0$: ok.
Sinon, on peut écrire $a\cos x + b\sin x = \sqrt{a^2 + b^2}(\alpha\cos x + \beta\sin x)$ avec

$$\alpha = \frac{a}{\sqrt{a^2 + b^2}} \text{ et } \beta = \frac{b}{\sqrt{a^2 + b^2}}$$

Comme $\alpha^2 + \beta^2 = 1$, il existe $\varphi \in \mathbb{R}$ tel que $\alpha = \cos\varphi$ et $\beta = \sin\varphi$ et alors

$$a\cos x + b\sin x = A(\cos x\cos\varphi - \sin x\sin\varphi) = A\cos(x - \varphi)$$

□

Exemple Résolvons l'équation $\cos x + \sin x = \sqrt{2}/2$ d'inconnue $x \in \mathbb{R}$.
On peut écrire

$$\cos x + \sin x = \sqrt{2}\left(\frac{1}{\sqrt{2}}\cos x + \frac{1}{\sqrt{2}}\sin x\right) = \sqrt{2}\cos\left(x - \frac{\pi}{4}\right)$$

Par suite

$$\cos x + \sin x = \frac{\sqrt{2}}{2} \Leftrightarrow \cos\left(x - \frac{\pi}{4}\right) = \frac{1}{2} \Leftrightarrow \begin{cases} x - \frac{\pi}{4} = \frac{\pi}{3} \quad [2\pi] \\ \text{ou} \\ x - \frac{\pi}{4} = -\frac{\pi}{3} \quad [2\pi] \end{cases} \Leftrightarrow \begin{cases} x = \frac{7\pi}{12} \quad [2\pi] \\ \text{ou} \\ x = -\frac{\pi}{12} \quad [2\pi] \end{cases}$$

Exemple Résolvons l'équation $\sqrt{3}\cos(2x) - 2\sin x\cos x = 1$ d'inconnue $x \in \mathbb{R}$.
On peut écrire

$$\sqrt{3}\cos(2x) - 2\sin x\cos x = \sqrt{3}\cos(2x) - \sin(2x) = 2\left(\frac{\sqrt{3}}{2}\cos(2x) - \frac{1}{2}\sin(2x)\right) = 2\cos\left(2x + \frac{\pi}{6}\right)$$

Par suite

$$\sqrt{3}\cos(2x) - 2\sin x \cos x = 1 \Leftrightarrow \cos\left(2x + \frac{\pi}{6}\right) = \frac{1}{2} \Leftrightarrow \begin{cases} 2x + \dfrac{\pi}{6} = \dfrac{\pi}{3} & [2\pi] \\ \text{ou} & \\ 2x + \dfrac{\pi}{6} = -\dfrac{\pi}{3} & [2\pi] \end{cases} \Leftrightarrow \begin{cases} x = -\dfrac{\pi}{12} & [\pi] \\ \text{ou} & \\ x = -\dfrac{\pi}{4} & [\pi] \end{cases}$$

14.4 Fonctions trigonométriques réciproques

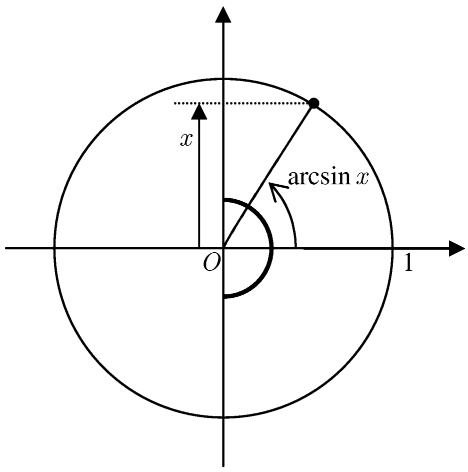
14.4.1 Fonction arc sinus

La fonction sinus est continue et strictement croissante de l'intervalle $[-\pi/2, \pi/2]$ vers l'intervalle $[-1, 1]$, elle réalise donc une bijection et par suite

$$\forall x \in [-1, 1], \exists! t \in [-\pi/2, \pi/2], \sin t = x.$$

Définition

Pour tout $x \in [-1, 1]$, on note $\arcsin x$ l'unique angle de l'intervalle $[-\pi/2, \pi/2]$ donc le sinus vaut x.
Ceci définit une application $\arcsin : [-1, 1] \to [-\pi/2, \pi/2]$ qui est la bijection réciproque de la restriction
$\sin : [-\pi/2, \pi/2] \to [-1, 1]$.



On a le tableau de valeurs

x	-1	$-\sqrt{3}/2$	$-\sqrt{2}/2$	$-1/2$	0	$1/2$	$\sqrt{2}/2$	$\sqrt{3}/2$	1
$\arcsin x$	$-\pi/2$	$-\pi/3$	$-\pi/4$	$-\pi/6$	0	$\pi/6$	$\pi/4$	$\pi/3$	$\pi/2$

Proposition

Pour $x \in [-1, 1]$, on a les simplifications

$$\sin(\arcsin x) = x, \ \cos(\arcsin x) = \sqrt{1 - x^2}$$

et quand $x \neq \pm 1$

$$\tan(\arcsin x) = \frac{x}{\sqrt{1 - x^2}}$$

dém. :
Soient $x \in [-1, 1]$ et $\theta = \arcsin x$.
Par définition $\theta \in [-\pi/2, \pi/2]$ et $\sin\theta = x$
On a déjà

$$\sin(\arcsin x) = x$$

Puisque $\cos^2\theta = 1 - \sin^2\theta = 1 - x^2$ et $\cos\theta \geqslant 0$ car $\theta \in [-\pi/2, \pi/2]$, on obtient

$$\cos(\arcsin x) = \sqrt{1 - x^2}$$

Enfin

$$\tan(\arcsin x) = \frac{\sin(\arcsin x)}{\cos(\arcsin x)} = \frac{x}{\sqrt{1 - x^2}}$$

pour $x \neq \pm 1$ de sorte que $\arcsin x \in \left]-\pi/2, \pi/2\right[$.
□

Théorème

La fonction $\arcsin$ est impaire, strictement croissante, continue sur $[-1, 1]$, de classe $\mathcal{C}^\infty$ sur $\left]-1, 1\right[$ et

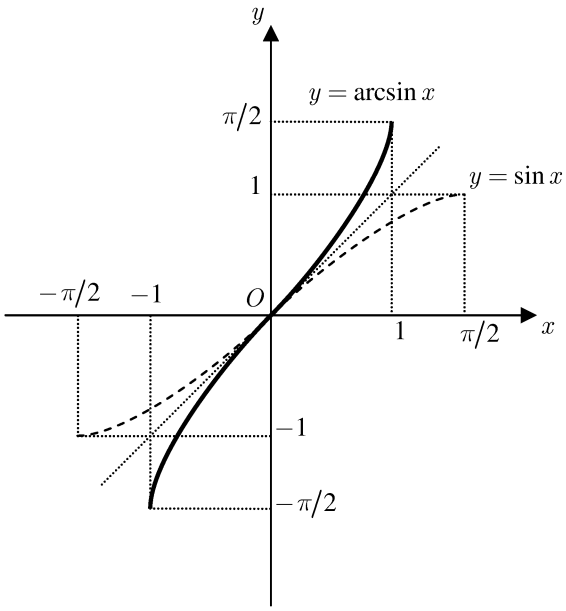
$$(\arcsin x)' = \frac{1}{\sqrt{1 - x^2}}$$

dém. :
La bijection $\sin : [-\pi/2, \pi/2] \to [-1, 1]$ est impaire continue et strictement croissante, il en est donc de même de sa bijection réciproque $\arcsin$.
La fonction $\sin : [-\pi/2, \pi/2] \to [-1, 1]$ est dérivable et pour tout $t \in \left]-\pi/2, \pi/2\right[$, $(\sin t)' = \cos t \neq 0$ donc la fonction $\arcsin$ est dérivable sur $\left]-1, 1\right[$ et

$$(\arcsin x)' = \frac{1}{\cos(\arcsin x)} = \frac{1}{\sqrt{1 - x^2}}$$

Enfin, la fonction $x \mapsto \dfrac{1}{\sqrt{1 - x^2}}$ étant indéfiniment dérivable sur $\left]-1, 1\right[$, la fonction $\arcsin$ l'est encore, c'est donc une fonction de classe $\mathcal{C}^\infty$ sur $\left]-1, 1\right[$.
□

Attention : La fonction $\arcsin$ n'est pas dérivable en 1 et -1, en fait la fonction $\arcsin$ présente des tangentes verticales en ces points.



Proposition

$\forall t \in [-\pi/2, \pi/2]\,, \arcsin(\sin t) = t.$

dém. :
Pour $t \in [-\pi/2, \pi/2]$, l'unique angle de $[-\pi/2, \pi/2]$ dont le sinus vaut $\sin t$ est t.
□

Attention : Cette formule ne vaut que pour $t \in [-\pi/2, \pi/2]$.

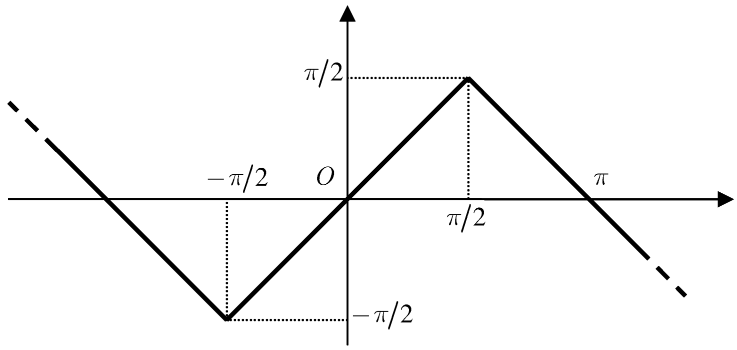
Exemple Simplifions

$$\arcsin\left(\sin\frac{2\pi}{3}\right)$$

En effet $\sin\dfrac{2\pi}{3} = \sin\dfrac{\pi}{3}$ et $\dfrac{\pi}{3} \in \left[-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right]$ donc

$$\arcsin(\sin\frac{2\pi}{3}) = \frac{\pi}{3}$$

Exemple Etudions la fonction $f : t \mapsto \arcsin(\sin t)$
La fonction f est définie et continue sur $\mathbb{R}$, elle est 2π périodique et impaire, il nous suffit de l'étudier sur l'intervalle $[0, \pi]$.
Sur $[0, \pi/2]$, $f(t) = t$ car on peut simplifier immédiatement $\arcsin(\sin t)$ sur cet intervalle.
Sur $[\pi/2, \pi]$, on a $\sin(t) = \sin(\pi - t)$ avec $\pi - t \in [0, \pi/2] \subset [-\pi/2, \pi/2]$ donc $f(t) = \pi - t$.
Finalement, on peut donner l'allure de f.



Exemple Résolvons l'équation $\arcsin x + \arcsin 2x = \pi/2$ d'inconnue $x \in [-1/2, 1/2]$
Soit x une solution, s'il y en a.
On a

$$\arcsin 2x = \pi/2 - \arcsin x$$

En appliquant la fonction sinus à chaque membre, on obtient

$$2x = \sin\left(\frac{\pi}{2} - \arcsin x\right)$$

Or

$$\sin\left(\frac{\pi}{2} - \arcsin x\right) = \cos(\arcsin x) = \sqrt{1 - x^2}$$

Ainsi

$$2x = \sqrt{1 - x^2}$$

En élevant au carré et sachant $x \geqslant 0$, on obtient $x = 1/\sqrt{5}$.
Il reste à vérifier si $x = 1/\sqrt{5}$ est, ou non, solution.
Pour cela considérons la fonction $f : [-1/2, 1/2] \to \mathbb{R}$ définie par $f(x) = \arcsin x + \arcsin 2x$.
La fonction f est continue, strictement croissante et $f(-1/2) \leqslant \frac{\pi}{2} \leqslant f(1/2)$.
Par suite l'équation étudiée possède une solution et une seule, à savoir $1/\sqrt{5}$.

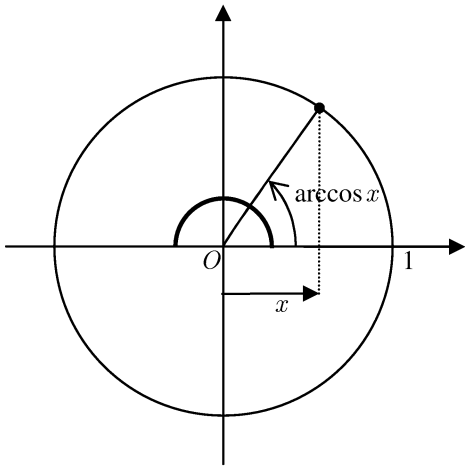
14.4.2 Fonction arc cosinus

La fonction cosinus est continue et strictement décroissante de l'intervalle $[0, \pi]$ vers l'intervalle $[-1, 1]$, elle réalise donc une bijection et par suite

$$\forall x \in [-1, 1], \exists! t \in [0, \pi], \cos t = x.$$

Définition

Pour tout $x \in [-1,1]$, on note $\arccos x$ l'unique angle de l'intervalle $[0,\pi]$ donc le cosinus vaut x.
Ceci définit une application $\arccos : [-1,1] \to [0,\pi]$ qui est la bijection réciproque de la restriction
$\cos : [0,\pi] \to [-1,1]$.



On a le tableau de valeurs

x	-1	$-\sqrt{3}/2$	$-\sqrt{2}/2$	$-1/2$	0	$1/2$	$\sqrt{2}/2$	$\sqrt{3}/2$	1
$\arccos x$	π	$5\pi/6$	$3\pi/4$	$2\pi/3$	$\pi/2$	$\pi/3$	$\pi/4$	$\pi/6$	0

Proposition

Pour $x \in [-1,1]$, on a les simplifications

$$\cos(\arccos x) = x,\ \sin(\arccos x) = \sqrt{1-x^2}$$

et quand $x \neq 0$

$$\tan(\arccos x) = \frac{\sqrt{1-x^2}}{x}$$

dém. :
Soient $x \in [-1,1]$ et $\theta = \arccos x$.
Par définition $\theta \in [0,\pi]$ et $\cos\theta = x$
On a déjà

$$\cos(\arccos x) = x$$

Puisque $\sin^2\theta = 1 - \cos^2\theta = 1 - x^2$ et $\sin\theta \geqslant 0$ car $\theta \in [0,\pi]$, on obtient

$$\sin(\arccos x) = \sqrt{1-x^2}$$

Enfin

$$\tan(\arccos x) = \frac{\sin(\arccos x)}{\cos(\arccos x)} = \frac{\sqrt{1-x^2}}{x}$$

pour $x \neq 0$ de sorte que $\arccos x \in [0, \pi/2[\cup]\pi/2, \pi]$.
□

Théorème

La fonction arccos est strictement décroissante, continue sur $[-1, 1]$, de classe $\mathcal{C}^\infty$ sur $]-1, 1[$ et

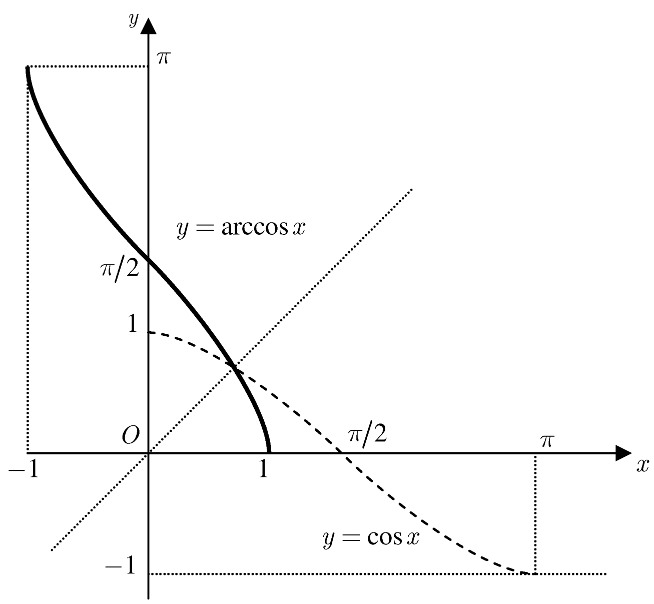
$$(\arccos x)' = -\frac{1}{\sqrt{1 - x^2}}$$

dém. :
La bijection $\cos : [0, \pi] \to [-1, 1]$ est continue et strictement décroissante, il en est donc de même de sa bijection réciproque arccos.
La fonction $\cos : [0, \pi] \to [-1, 1]$ est dérivable et pour tout $t \in]0, \pi[$, $(\cos t)' = -\sin t \neq 0$ donc la fonction arccos est dérivable sur $]-1, 1[$ et

$$(\arccos x)' = \frac{1}{-\sin(\arccos x)} = -\frac{1}{\sqrt{1 - x^2}}$$

Enfin, la fonction $x \mapsto -\dfrac{1}{\sqrt{1 - x^2}}$ étant indéfiniment dérivable sur $]-1, 1[$, la fonction arccos l'est encore, c'est donc une fonction de classe $\mathcal{C}^\infty$ sur $]-1, 1[$.
□

Attention : La fonction arccos n'est pas dérivable en 1 et -1 mais présente des tangentes verticales en ces points.



Proposition

$\forall t \in [0, \pi] , \arccos(\cos t) = t.$

dém. :
Pour $t \in [0, \pi]$, l'unique angle de $[0, \pi]$ dont le cosinus vaut $\cos t$ est t.
□

Attention : Cette formule ne vaut que pour $t \in [0, \pi]$.

Exemple $\forall t \in [-\pi, 0]\,, \arccos(\cos t) = -t$.
En effet $\cos t = \cos(-t)$ avec $-t \in [0, \pi]$.
$\forall t \in [\pi, 2\pi]\,, \arccos t = 2\pi - t$.
En effet $\cos t = \cos(2\pi - t)$ et $2\pi - t \in [0, \pi]$.

Proposition

$\forall x \in [-1, 1]$, $\arcsin x + \arccos x = \pi/2$.

dém. :
Posons $t = \pi/2 - \arcsin x$.
On a $t \in [0, \pi]$ et $\cos t = \sin(\arcsin x) = x$ donc $t = \arccos x$.
□

Exemple Résolvons l'équation $\arcsin x - \arccos x = \pi/6$ d'inconnue $x \in [-1, 1]$.
Puisque

$$\arccos x = \frac{\pi}{2} - \arcsin x$$

on a

$$\arcsin x - \arccos x = \pi/6 \Leftrightarrow 2\arcsin x = 2\pi/3 \Leftrightarrow \arcsin x = \frac{\pi}{3} \Leftrightarrow x = \frac{\sqrt{3}}{2}$$

Exemple Simplifions l'expression $\arccos(2x^2 - 1)$.
Cette expression a un sens pour $2x^2 - 1 \in [-1, 1]$ i.e. $x \in [-1, 1]$
Posons $\theta = \arccos x$, de sorte que $\cos\theta = x$ et $\theta \in [0, \pi]$.
On a $2x^2 - 1 = \cos 2\theta$ et donc

$$\arccos(2x^2 - 1) = \arccos(\cos 2\theta)$$

Si $x \in [0, 1]$ alors $\theta \in [0, \pi/2]$ et $\arccos(\cos 2\theta) = 2\theta$ car $2\theta \in [0, \pi]$.
Par suite $\arccos(2x^2 - 1) = 2\arccos x$.
Par parité, on a

$$\arccos(2x^2 - 1) = 2\arccos|x|$$

pour tout $x \in [-1, 1]$.

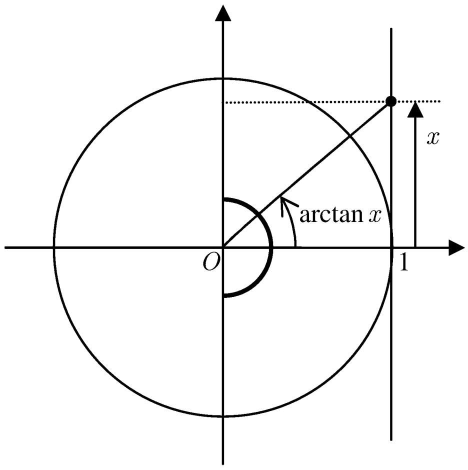
14.4.3 Fonction arc tangente

La fonction tangente est continue et strictement croissante de l'intervalle $]-\pi/2, \pi/2[$ vers $\mathbb{R}$, elle réalise donc une bijection et par suite

$$\forall x \in \mathbb{R}, \exists! t \in \left]-\pi/2, \pi/2\right[, \tan t = x.$$

Définition

Pour tout $x \in \mathbb{R}$, on note $\arctan x$ l'unique angle de l'intervalle $]-\pi/2, \pi/2[$ donc la tangente vaut x.
Ceci définit une application $\arctan : \mathbb{R} \to]-\pi/2, \pi/2[$ qui est la bijection réciproque de la restriction
$\tan :]-\pi/2, \pi/2[\to \mathbb{R}$.



Remarque La fonction $\arctan$ n'a pas de lien avec $\dfrac{\arcsin}{\arccos}$!

On a le tableau de valeurs

x	$-\infty$	$-\sqrt{3}$	-1	$-1/\sqrt{3}$	0	$1/\sqrt{3}$	1	$\sqrt{3}$	$+\infty$
$\arcsin x$	$-\pi/2$	$-\pi/3$	$-\pi/4$	$-\pi/6$	0	$\pi/6$	$\pi/4$	$\pi/3$	$\pi/2$

Proposition

Pour $x \in \mathbb{R}$, on a les simplifications

$$\tan(\arctan x) = x,\ \cos(\arctan x) = \frac{x}{\sqrt{1+x^2}} \text{ et } \sin(\arctan x) = \frac{x}{\sqrt{1+x^2}}$$

dém. :
Soient $x \in \mathbb{R}$ et $\theta = \arctan x$.

Par définition $\theta \in \left]-\pi/2, \pi/2\right[$ et $\tan\theta = x$
On a déjà

$$\tan(\arctan x) = x$$

Puisque $\cos^2\theta = \dfrac{1}{1+\tan^2\theta} = \dfrac{1}{1+x^2}$ et $\cos\theta > 0$ car $\theta \in \left]-\pi/2, \pi/2\right[$, on obtient

$$\cos(\arctan x) = \frac{1}{\sqrt{1+x^2}}$$

Enfin

$$\tan(\arctan x) = \frac{\sin(\arctan x)}{\cos(\arctan x)} = x$$

donne

$$\sin(\arctan x) = \frac{x}{\sqrt{1+x^2}}$$

□

Théorème

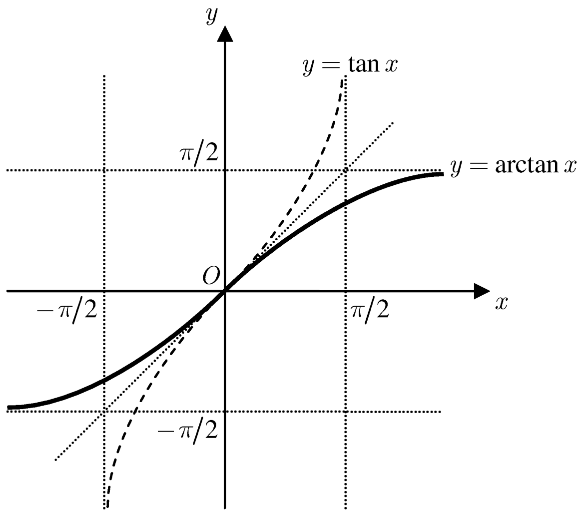
La fonction $\arctan$ est impaire, strictement croissante, continue et même $\mathcal{C}^\infty$ sur $\mathbb{R}$ et

$$(\arctan x)' = \frac{1}{1+x^2}$$

dém. :
La bijection $\tan : \left]-\pi/2, \pi/2\right[\to \mathbb{R}$ est impaire continue et strictement croissante, il en est donc de même de sa bijection réciproque $\arctan$.
La fonction $\tan : \left]-\pi/2, \pi/2\right[\to \mathbb{R}$ est dérivable et pour tout $t \in \left]-\pi/2, \pi/2\right[$, $(\tan t)' = 1 + \tan^2 t \neq 0$ donc la fonction $\arctan$ est dérivable sur $\mathbb{R}$ et

$$(\arctan x)' = \frac{1}{1+\tan^2(\arctan x)} = \frac{1}{1+x^2}$$

Enfin, la fonction $x \mapsto \dfrac{1}{1+x^2}$ étant indéfiniment dérivable sur $\mathbb{R}$, la fonction $\arctan$ l'est encore, c'est donc une fonction de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$.



□

Proposition

$\forall t \in \left]-\pi/2, \pi/2\right[, \arctan(\tan t) = t.$

dém. :

Pour $t \in \left]-\pi/2, \pi/2\right[$, l'unique angle de $\left]-\pi/2, \pi/2\right[$ dont la tangente vaut $\tan t$ est t.

□

Attention : Cette formule ne vaut que pour $t \in \left]-\pi/2, \pi/2\right[$.

Exemple $\forall t \in \left]\pi/2, 3\pi/2\right[, \arctan(\tan t) = t - \pi$.

En effet $\tan(t) = \tan(t - \pi)$ et $t - \pi \in \left]-\pi/2, \pi/2\right[$.

Proposition

$\forall x > 0, \arctan x + \arctan \dfrac{1}{x} = \pi/2$ et $\forall x < 0, \arctan x + \arctan \dfrac{1}{x} = -\pi/2.$

dém. :

Considérons la fonction $f : \mathbb{R}^\star \to \mathbb{R}$ définie par $f(x) = \arctan x + \arctan 1/x$.

La fonction f est dérivable et

$$f'(x) = \frac{1}{1+x^2} - \frac{1/x^2}{1+1/x^2} = 0$$

La fonction f est donc constante sur chaque intervalle formant $\mathbb{R}^\star$.

Puisque $f(1) = 2\arctan 1 = \pi/2$, on obtient $\forall x > 0, \arctan x + \arctan(1/x) = \pi/2$.

Par imparité, on obtient $\forall x < 0, \arctan x + \arctan(1/x) = -\pi/2$

□

Exemple Calculons

$$\theta = \arctan \frac{1}{2} + \arctan \frac{1}{3}$$

Sous réserve que $\theta \neq \pi/2 \quad [\pi]$, la formule

$$\tan(a+b) = \frac{\tan a + \tan b}{1 - \tan a \tan b}$$

donne $\tan\theta = 1$.
Pour déterminer θ, il n'y plus qu'à localiser cet angle dans un intervalle de longueur inférieure à π.
Puisque $0 < \arctan\frac{1}{2} < \frac{\pi}{4}$ et $0 < \arctan\frac{1}{3} < \frac{\pi}{4}$, on a $0 < \theta < \pi/2$.
Par suite $\theta = \pi/4$.

14.5 Fonctions hyperboliques

14.5.1 Fonction cosinus et sinus hyperboliques

Définition

On définit la fonction cosinus hyperbolique ch : $\mathbb{R} \to \mathbb{R}$ et la fonction sinus hyperbolique sh : $\mathbb{R} \to \mathbb{R}$ par

$$\mathrm{ch}t = \frac{\mathrm{e}^t + \mathrm{e}^{-t}}{2} \text{ et } \mathrm{sh}t = \frac{\mathrm{e}^t - \mathrm{e}^{-t}}{2}$$

Théorème

La fonction ch est paire, de classe $\mathcal{C}^\infty$ et $(\mathrm{ch}t)' = \mathrm{sh}t$.
La fonction sh est impaire, de classe $\mathcal{C}^\infty$ et $(\mathrm{sh}t)' = \mathrm{ch}t$.
Les tableaux de variation des fonctions ch et sh sont

t	$-\infty$		0		$+\infty$
$\mathrm{ch}t$	$+\infty$	$\searrow$	1	$\nearrow$	$+\infty$

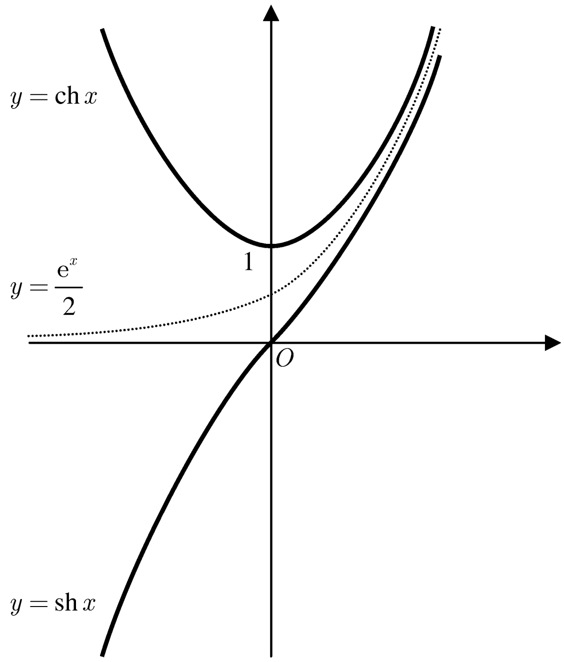
et

t	$-\infty$		0		$+\infty$
$\mathrm{sh}t$	$-\infty$	$\nearrow$	0	$\nearrow$	$+\infty$

dém. :
$\mathrm{ch}(-t) = \frac{\mathrm{e}^{-t} + \mathrm{e}^t}{2} = \mathrm{ch}t$ et $\mathrm{sh}(-t) = \frac{\mathrm{e}^{-t} - \mathrm{e}^t}{2} = -\mathrm{sh}t$.
Par opérations sur les fonctions, les fonctions ch et sh sont de classe $\mathcal{C}^\infty$ avec
$(\mathrm{ch}t)' = \frac{\mathrm{e}^t - \mathrm{e}^{-t}}{2} = \mathrm{sh}t$ et $(\mathrm{sh}t)' = \frac{\mathrm{e}^t + \mathrm{e}^{-t}}{2} = \mathrm{ch}t$.
La fonction ch étant positive, on en déduit la croissance de la fonction sh. Sachant que celle-ci s'annule en 0, on obtient le signe de la fonction sh et donc les variations de la fonction ch. Enfin les limites des fonctions ch et sh en $\pm\infty$ sont immédiates.
□

Remarque La fonction ch présente un minimum en 0 : $\forall t \in \mathbb{R}, \mathrm{ch}t \geqslant 1$.

Remarque Pour représenter les fonctions ch et sh, on peut prendre appui sur le graphe de la fonction $f : t \mapsto \mathrm{e}^t/2$.
En effet $\mathrm{sh}t \leqslant f(t) \leqslant \mathrm{ch}t$ et $\mathrm{ch}t - f(t) = f(t) - \mathrm{sh}t \xrightarrow[t\to+\infty]{} 0$.



Exemple Pour $n \in \mathbb{N}$ et $t \in \mathbb{R}$, calculons

$$S = \sum_{k=0}^{n} \mathrm{ch}(kt)$$

En réorganisant la somme

$$S = \frac{1}{2}\sum_{k=0}^{n}\left(\mathrm{e}^{kt} + \mathrm{e}^{-kt}\right) = 1 + \frac{1}{2}\sum_{k=-n}^{n} \mathrm{e}^{kt}$$

Pour $t = 0$, $S = n + 1$.
Pour $t \neq 0$, par sommation géométrique et factorisation d'exponentielle équilibrée,

$$S = 1 + \frac{1}{2}\mathrm{e}^{-nt}\frac{\mathrm{e}^{(2n+1)t} - 1}{\mathrm{e}^{t} - 1} = 1 + \frac{1}{2}\frac{\mathrm{sh}\frac{(2n+1)t}{2}}{\mathrm{sh}\frac{t}{2}}$$

On aurait aussi pu mener le calcul en introduisant $S' = \sum_{k=0}^{n} \mathrm{sh} kt$ et en exprimant $S + S'$ et $S - S'$.

14.5.2 Trigonométrie hyperbolique

Théorème

$\forall t \in \mathbb{R}, \mathrm{ch}^2 t - \mathrm{sh}^2 t = 1.$

dém. :
$\text{ch}^2 t - \text{sh}^2 t = (\text{ch}t + \text{sh}t)(\text{ch}t - \text{sh}t) = \mathrm{e}^t \mathrm{e}^{-t} = 1$.
□

Corollaire

$\forall t \in \mathbb{R}, \text{ch}t = \sqrt{1 + \text{sh}^2 t}$.

dém. :
Par ce qui précède $\text{ch}^2 t = 1 + \text{sh}^2 t$, or $\text{ch}t \geqslant 0$ donc $\text{ch}t = \sqrt{1 + \text{sh}^2 t}$.
□

Proposition

Pour tout $a, b \in \mathbb{R}$,

$$\text{ch}(a+b) = \text{ch}a\text{ch}b + \text{sh}a\text{sh}b,\ \text{ch}(a-b) = \text{ch}a\text{ch}b - \text{sh}a\text{sh}b$$

$$\text{sh}(a+b) = \text{sh}a\text{ch}b + \text{sh}b\text{ch}a,\ \text{sh}(a-b) = \text{sh}a\text{ch}b - \text{sh}b\text{ch}a$$

$$\text{ch}2a = \text{ch}^2 a - \text{sh}^2 a = 2\text{ch}^2 a - 1 = 1 + 2\text{sh}^2 a \text{ et } \text{sh}2a = 2\text{sh}a\text{ch}a.$$

dém. :

$$\text{ch}a\text{ch}b + \text{sh}a\text{sh}b = \frac{\mathrm{e}^a + \mathrm{e}^{-a}}{2}\frac{\mathrm{e}^b + \mathrm{e}^{-b}}{2} + \frac{\mathrm{e}^a - \mathrm{e}^{-a}}{2}\frac{\mathrm{e}^b - \mathrm{e}^{-b}}{2} = \frac{2\mathrm{e}^{a+b} + 2\mathrm{e}^{-(a+b)}}{4} = \text{ch}(a+b)$$

Par parité, on en déduit $\text{ch}(a-b) = \text{ch}a\text{ch}b - \text{sh}a\text{sh}b$.
La formule $\text{sh}(a+b) = \text{sh}a\text{ch}b + \text{sh}b\text{ch}a$ s'obtient comme ci-dessus et on en déduit par parité $\text{sh}(a-b) = \text{sh}a\text{ch}b - \text{sh}b\text{ch}a$.
Enfin, en prenant $a = b$, ce qui précède donne les développements de $\text{ch}2a$ et $\text{sh}2a$.
□

Remarque Ces formules se déduisent de celles connues en trigonométrie circulaires par les transformations $\cos \to$ ch et $\sin \to i.$sh.

Exemple Calculons $P_n = \prod_{k=1}^{n} \text{ch}(2^k t)$ pour $t \in \mathbb{R}$.
Pour $t = 0$, $P_n = \prod_{k=1}^{n} 1 = 1$.
Pour $t \neq 0$, la formule $\text{sh}2a = 2\text{sh}a\text{ch}a$ donne

$$\text{ch}(2^k t) = \frac{1}{2}\frac{\text{sh}(2^{k+1}t)}{\text{sh}(2^k t)}$$

On en déduit après télescopage

$$P_n = \frac{\text{sh}(2^n t)}{2^n \text{sh}t}$$

14.5.3 Fonction tangente hyperbolique

Définition

On définit la fonction tangente hyperbolique th : $\mathbb{R} \to \mathbb{R}$ par

$$\mathrm{th}t = \frac{\mathrm{sh}t}{\mathrm{ch}t}$$

Proposition

$\forall t \in \mathbb{R}$, $\mathrm{th}t = \dfrac{\mathrm{e}^t - \mathrm{e}^{-t}}{\mathrm{e}^t + \mathrm{e}^{-t}} = \dfrac{\mathrm{e}^{2t} - 1}{\mathrm{e}^{2t} + 1}$.

dém. :
$\mathrm{sh}t = \dfrac{\mathrm{e}^t - \mathrm{e}^{-t}}{2}$ et $\mathrm{ch}t = \dfrac{\mathrm{e}^t + \mathrm{e}^{-t}}{2}$ donne $\mathrm{th}t = \dfrac{\mathrm{sh}t}{\mathrm{ch}t} = \dfrac{\mathrm{e}^t - \mathrm{e}^{-t}}{\mathrm{e}^t + \mathrm{e}^{-t}}$.
En multipliant numérateur et dénominateur par e^t, on obtient $\mathrm{th}t = \dfrac{\mathrm{e}^{2t} - 1}{\mathrm{e}^{2t} + 1}$.

□

Théorème

La fonction th est impaire, de classe $\mathcal{C}^\infty$ et $(\mathrm{th}t)' = 1 - \mathrm{th}^2 t = \dfrac{1}{\mathrm{ch}^2 t}$.
Le tableaux de variation de la fonction th est

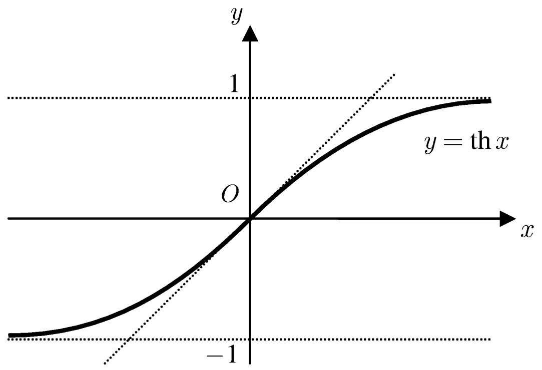
t	$-\infty$		0		$+\infty$
$\mathrm{th}t$	-1	$\nearrow$	0	$\nearrow$	1

dém. :
Par opérations sur les fonctions, la fonction th est impaire et de classe $\mathcal{C}^\infty$.
De plus,

$$(\mathrm{th}t)' = \frac{\mathrm{ch}^2 t - \mathrm{sh}^2 t}{\mathrm{ch}^2 t} = 1 - \mathrm{th}^2 t \text{ ou } (\mathrm{th}t)' = \frac{\mathrm{ch}^2 t - \mathrm{sh}^2 t}{\mathrm{ch}^2 t} = \frac{1}{\mathrm{ch}^2 t}$$

Cette deuxième formule de dérivation donne la croissance de la fonction th et les limites de cette fonction en $\pm\infty$ s'obtiennent facilement par les écritures

$$\mathrm{th}t = \frac{\mathrm{e}^{2t} - 1}{\mathrm{e}^{2t} + 1} = \frac{1 - \mathrm{e}^{-2t}}{1 + \mathrm{e}^{-2t}}$$



□

Remarque $\forall t \in \mathbb{R}, -1 < \text{th}t < 1$.

14.5.4 Fonctions hyperboliques réciproques

14.5.4.1 Fonction argument sinus hyperbolique

La fonction sh est continue et strictement croissante de $\mathbb{R}$ vers $\mathbb{R}$, elle réalise donc une bijection et par suite

$$\forall x \in \mathbb{R}, \exists! t \in \mathbb{R}, \text{sh}t = x.$$

Définition

Pour tout $x \in \mathbb{R}$, on note argsh$x$ l'unique réel t vérifiant sh$t = x$.
Ceci définit une application argsh : $\mathbb{R} \to \mathbb{R}$ qui est la bijection réciproque de sh : $\mathbb{R} \to \mathbb{R}$.

Proposition

Pour $x \in \mathbb{R}$, on a les simplifications

$$\text{sh}(\text{argsh}x) = x, \text{ch}(\text{argsh}x) = \sqrt{1+x^2} \text{ et } \text{th}(\text{argsh}x) = \frac{x}{\sqrt{1+x^2}}$$

dém. :
Soient $x \in \mathbb{R}$ et $t = \text{argsh}x$.
Par définition, on a sh$t = x$ et donc

$$\text{sh}(\text{argsh}x) = x$$

Puisque ch$t = \sqrt{1+\text{sh}^2 t}$, on a

$$\text{ch}(\text{argsh}x) = \sqrt{1+x^2}$$

Enfin th$t = \dfrac{\text{sh}t}{\text{ch}t}$ donne

$$\text{th}(\text{argsh}x) = \frac{x}{\sqrt{1+x^2}}$$

□

Théorème

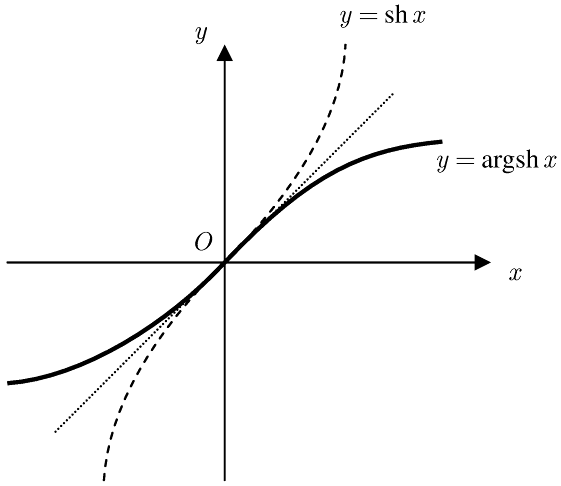
La fonction argsh est impaire, strictement croissante, de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$ et

$$(\text{argsh}x)' = \frac{1}{\sqrt{1+x^2}}$$

dém. :
La bijection sh : $\mathbb{R} \to \mathbb{R}$ est impaire et strictement croissante, il en est donc de même de sa bijection réciproque argsh.
La fonction sh : $\mathbb{R} \to \mathbb{R}$ est dérivable et $(\text{sh}t)' = \text{ch}t \neq 0$ donc la fonction argsh est dérivable sur $\mathbb{R}$ et

$$(\text{argsh}x)' = \frac{1}{\text{ch}(\text{argsh}x)} = \frac{1}{\sqrt{1+x^2}}$$

Enfin, la fonction $x \mapsto \dfrac{1}{\sqrt{1+x^2}}$ étant indéfiniment dérivable sur $\mathbb{R}$, la fonction argsh l'est encore, c'est donc une fonction de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$.



□

Proposition

$\forall x \in \mathbb{R}, \text{argsh}(x) = \ln(x + \sqrt{x^2+1})$

dém. :
Soient $x \in \mathbb{R}$ et $t = \text{argsh}x$.
On a $\text{sh}t = x$ donc $\text{e}^t - \text{e}^{-t} = 2x$ puis $\text{e}^{2t} - 2x\,\text{e}^t - 1 = 0$.
Posons $r = \text{e}^t$.
Les solutions de l'équation du second degré $r^2 - 2xr - 1 = 0$ sont $x + \sqrt{1+x^2}$ et $x - \sqrt{1+x^2}$.
Or $r > 0$ et $x - \sqrt{1+x^2} \leqslant 0$ donc $\text{e}^t = x + \sqrt{1+x^2}$ puis $t = \ln(x + \sqrt{1+x^2})$.
□

14.5.4.2 Fonction argument cosinus hyperbolique

La fonction ch est continue et strictement croissante de $[0, +\infty[$ vers $[1, +\infty[$, elle réalise donc une bijection et par suite

$$\forall x \in [1, +\infty[\,, \exists! t \in [0, +\infty[\,, \text{ch} t = x.$$

Définition

Pour tout Semestre irrégulier mais avec quelques bons résultats., on note argch$x$ l'unique $t \in [0, +\infty[$ vérifiant ch$t = x$.
Ceci définit une application argch : $[1, +\infty[\; \to \mathbb{R}^+$ bijection réciproque de la restriction ch : $[0, +\infty[\; \to [1, +\infty[$

Proposition

Pour $x \in [1, +\infty[$, on a les simplifications

$$\text{ch}(\text{argch} x) = x,\ \text{sh}(\text{argch} x) = \sqrt{x^2 - 1} \text{ et } \text{th}(\text{argch} x) = \frac{\sqrt{x^2 - 1}}{x}$$

dém. :
Soient $x \in [1, +\infty[$ et $t = \text{argch} x$.
Par définition, $t \geqslant 0$ et ch$t = x$.
On a déjà

$$\text{ch}(\text{argch} x) = x$$

Puisque $\text{sh}^2 t = \text{ch}^2 t - 1 = x^2 - 1$ et sh$t \geqslant 0$, on obtient

$$\text{sh}(\arg \text{ch} x) = \sqrt{x^2 - 1}$$

Enfin th$t = \dfrac{\text{sh} t}{\text{ch} t}$ donne

$$\text{th}(\text{argch} x) = \frac{\sqrt{x^2 - 1}}{x}$$

□

Théorème

La fonction argch est strictement croissante, continue sur $[1, +\infty[$, de classe $\mathcal{C}^\infty$ sur $]1, +\infty[$ et

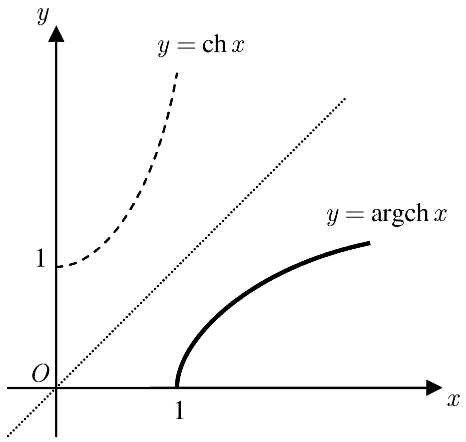
$$(\text{argch} x)' = \frac{1}{\sqrt{x^2 - 1}}$$

dém. :
La bijection ch : $[0, +\infty[\; \to [1, +\infty[$ et continue et strictement croissante, il en est donc de même de sa bijection réciproque argch : $[1, +\infty[\; \to \mathbb{R}^+$.
La fonction ch : $[0, +\infty[\; \to [1, +\infty[$ est dérivable et $\forall t > 0$, $(\text{ch} t)' = \text{sh} t \neq 0$ donc la fonction argch est dérivable sur $]1, +\infty[$ et

$$(\text{argch} x)' = \frac{1}{\text{sh}(\text{argch} x)} = \frac{1}{\sqrt{x^2 - 1}}$$

Enfin, la fonction $x \mapsto \dfrac{1}{\sqrt{x^2-1}}$ étant indéfiniment dérivable sur $]1,+\infty[$, la fonction argsh l'est encore, c'est donc une fonction de classe $\mathcal{C}^\infty$ sur $]1,+\infty[$.
□

Attention : La fonction argch n'est pas dérivable en 1, en fait la fonction argch présente une tangente verticale en ce point.



Proposition

$\forall x \in [1,+\infty[\,, \operatorname{argch}(x) = \ln(x+\sqrt{x^2-1})$

dém. :
Soient $x \in [1,+\infty[$ et $t = \operatorname{argch} x \in [0,+\infty[$.
On a $\operatorname{ch} t = x$ donc $\mathrm{e}^t + \mathrm{e}^{-t} = 2x$ puis $\mathrm{e}^{2t} - 2x\,\mathrm{e}^t + 1 = 0$.
Posons $r = \mathrm{e}^t$.
Les solutions de l'équation du second degré $r^2 - 2xr + 1 = 0$ sont $x + \sqrt{x^2-1}$ et $x - \sqrt{x^2-1}$.
Quand $x = 1$, on obtient deux fois la même racine et on peut donc affirmer $r = x + \sqrt{x^2-1}$.
Quand $x > 1$, on a $x - \sqrt{x^2-1} = \dfrac{1}{x+\sqrt{x^2-1}} < 1$ alors que $r > 1$ donc $r = x + \sqrt{x^2-1}$.
Dans les deux cas $\mathrm{e}^t = x + \sqrt{x^2-1}$ puis $t = \ln(x+\sqrt{x^2-1})$.
□

14.5.4.3 Fonction argument tangente hyperbolique

La fonction th est continue et strictement croissante de $\mathbb{R}$ vers $]-1,1[$, elle réalise donc une bijection et par suite

$$\forall x \in \,]-1,1[\,, \exists! t \in \mathbb{R}, \operatorname{th} t = x$$

Définition

Pour tout $x \in \,]-1,1[$, on note $\operatorname{argth} x$ l'unique $t \in \mathbb{R}$ vérifiant $\operatorname{th} t = x$.
Ceci définit une application $\operatorname{argth} :]-1,1[\, \to \mathbb{R}$ qui est la bijection réciproque de $\operatorname{th} : \mathbb{R} \to \,]-1,1[$.

Théorème

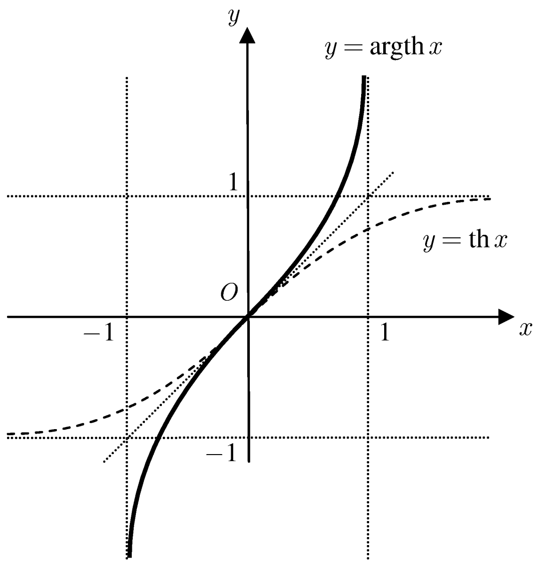
La fonction argth est impaire, strictement croissante, de classe $\mathcal{C}^\infty$ sur $]-1,1[$ et

$$(\text{argth}x)' = \frac{1}{1-x^2}$$

dém. :
La bijection th : $\mathbb{R} \to]-1,1[$ est impaire et strictement croissante, il en est donc de même de sa bijection réciproque argth.
La fonction th : $\mathbb{R} \to]-1,1[$ est dérivable et $(\text{th}t)' = 1 - \text{th}^2 t \neq 0$ donc la fonction argth est dérivable sur $]-1,1[$ et

$$(\text{argth}x)' = \frac{1}{1-\text{th}^2(\text{argth})} = \frac{1}{1-x^2}$$

Enfin, la fonction $x \mapsto \dfrac{1}{1-x^2}$ étant indéfiniment dérivable sur $]-1,1[$, la fonction argth l'est encore, c'est donc une fonction de classe $\mathcal{C}^\infty$ sur $]-1,1[$.



□

Proposition

$\forall x \in]-1,1[\,, \text{argth}(x) = \dfrac{1}{2}\ln\left(\dfrac{1+x}{1-x}\right).$

dém. :
Soient $x \in]-1,1[$ et $t = \text{argth}x$.
On a th$t = x$ donc $\dfrac{\mathrm{e}^{2t}-1}{\mathrm{e}^{2t}+1} = x$ puis $\mathrm{e}^{2t}(1-x) = 1+x$ et enfin $\mathrm{e}^{2t} = \dfrac{1+x}{1-x}$.
On en déduit

$$\text{argth}(x) = \frac{1}{2}\ln\left(\frac{1+x}{1-x}\right)$$

□

Chapitre 15

Suites numériques

15.1 Suites réelles

15.1.1 Définitions générales

Définition

On appelle suite réelle toute application $u : \mathbb{N} \to \mathbb{R}$.
Pour tout $n \in \mathbb{N}$, $u(n)$, usuellement notée u_n, est appelé terme d'indice n de la suite. Une telle suite est notée u, (u_n) ou $(u_n)_{n\in\mathbb{N}}$.
On note $\mathbb{R}^{\mathbb{N}}$ ou $\mathcal{F}(\mathbb{N}, \mathbb{R})$ l'ensemble de ces suites.

Attention : Il ne faut pas confondre la suite (u_n) avec son terme u_n.

Exemple $((-1)^n)$ est une suite réelle bornée.

Exemple $(n.(-1)^n)$ est une suite réelle ni minorée, ni majorée.

Définition

On appelle suite réelle définie à partir d'un rang $n_0 \in \mathbb{N}$ toute application $u : \{n \in \mathbb{N}/n \geqslant n_0\} \to \mathbb{R}$.
Une telle suite est notée u, $(u_n)_{n\geqslant n_0}$ ou (u_n).

Exemple La suite $(1/n)$ est définie à partir du rang 1.

Remarque Ce qui suit est présenté dans le cadre des suites définies à partir du rang 0 mais peut aisément se prolonger aux suites définies à partir d'un rang n_0, quitte à définir arbitrairement le terme de la suite pour les rangs $0, \ldots, n_0 - 1$.

Définition

Une suite réelle $u = (u_n)$ est dite constante s'il existe $C \in \mathbb{R}$ tel que $\forall n \in \mathbb{N}, u_n = C$.
Cette suite est alors dite constante égale à C et est souvent notée C.

Exemple La suite nulle est la suite constante égale à 0.

Définition

Une suite réelle $u = (u_n)$ est dite stationnaire s'il existe $C \in \mathbb{R}$ et $N \in \mathbb{N}$ vérifiant $\forall n \geqslant N, u_n = C$.
Cette suite est alors dite stationnaire égale à C.

Définition

On appelle valeur absolue d'une suite réelle $u = (u_n)$ la suite notée $|u|$ de terme général $|u_n|$.

Proposition

Soit $u = (u_n)$ une suite réelle. On a équivalence entre :
(i) u est bornée ;
(ii) $\exists M \in \mathbb{R}, \forall n \in \mathbb{N}, |u_n| \leqslant M$;
(iii) la suite $|u|$ est majorée.

dém. :
(i) $\Rightarrow$ (ii) Supposons la suite réelle u minorée et majorée.
Il existe $m_1, m_2 \in \mathbb{R}$ tels que pour tout $n \in \mathbb{N}$, $m_1 \leqslant u_n \leqslant m_2$.
Posons alors $M = \max(-m_1, m_2)$. En discutant selon signe de u_n, on vérifie que pour tout $n \in \mathbb{N}$, $|u_n| \leqslant M$.
(ii) $\Rightarrow$ (i) Supposons qu'il existe $M \in \mathbb{R}$ tel que pour tout $n \in \mathbb{N}$, $|u_n| \leqslant M$.
On a alors $-M \leqslant u_n \leqslant M$ donc la suite u est minorée et majorée.
(ii) $\Leftrightarrow$ (iii) Par définition de la majoration d'une suite.
□

Exemple Soit (u_n) la suite réelle définie par

$$u_n = \frac{\sin(n)}{3 - \cos(n)}$$

Montrons que (u_n) est bornée.
On a

$$|u_n| \leqslant \frac{|\sin(n)|}{|3 - \cos(n)|} \leqslant \frac{1}{|3 - \cos n|} \leqslant \frac{1}{2}$$

Ici, procéder par encadrement n'est pas le plus pratique car le signe de $\sin(n)$ changeant les inégalités multipliées doivent être renversées quand $\sin(n) \leqslant 0$; l'introduction des valeurs absolues évite ce problème en rendant positifs les facteurs.

15.1.2 Opérations sur les suites réelles

Définition

Soient $u = (u_n)$ et $v = (v_n)$ deux suites réelles et $\lambda \in \mathbb{R}$.
On note $\lambda.u$ la suite de terme général λu_n.
On note $u + v$ la suite de terme général $u_n + v_n$.
On note uv la suite de terme général $u_n v_n$.

Exemple $-u$ désigne la suite $(-1).u$ de terme général $-u_n$.

Proposition

Soient u et v des suites numériques et $\lambda \in \mathbb{K}$.
Si u et v sont bornées alors $\lambda.u, u + v$ et uv le sont aussi.

dém. :
On suppose qu'il existe $M, M' \in \mathbb{R}$ tels que pour tout $n \in \mathbb{N}$,

$$|u_n| \leqslant M \text{ et } |v_n| \leqslant M'$$

On a alors

$$|\lambda u_n| \leqslant |\lambda| M,\ |u_n + v_n| \leqslant |u_n| + |v_n| \leqslant M + M' \text{ et } |u_n v_n| = |u_n| |v_n| \leqslant MM'$$

Ainsi les suites $\lambda.u, u + v$ et uv sont bornées.
□

15.1.3 Suite arithmético-géométrique

Définition

On appelle suite arithmético-géométrique de raisons q et $r \in \mathbb{R}$ toute suite $u = (u_n)$ telle que

$$\forall n \in \mathbb{N}, u_{n+1} = qu_n + r$$

Exprimons le terme général d'une telle suite :
Cas $r = 0$:
On parle ici de suite géométrique de raison q, on a $u_n = q^n u_0$ pour tout $n \in \mathbb{N}$.
Cas $q = 1$:
On parle de suite arithmétique de raison r, on a $u_n = u_0 + nr$ pour tout $n \in \mathbb{N}$.
Cas $q \neq 1$:
Cherchons $\alpha \in \mathbb{R}$ tel que $v_n = u_n - \alpha$ soit géométrique.
Puisque $v_{n+1} = u_{n+1} - \alpha = qv_n + (q - 1)\alpha + r$, pour $\alpha = -\dfrac{r}{q-1}$ la suite (v_n) est géométrique de raison q et donc $v_n = q^n v_0$ pour tout $n \in \mathbb{N}$.
Par suite

$$u_n = q^n u_0 + \frac{q^n - 1}{q - 1} r$$

Exemple Déterminons le terme général de la suite (u_n) définie par $u_0 = 1$ et $u_{n+1} = 3u_n - 1$ pour tout $n \in \mathbb{N}$.

Considérons $v_n = u_n - \alpha$.

$$v_{n+1} = 3u_{n+1} - \alpha = 3(v_n + \alpha) - 1 - \alpha = 3v_n + 2\alpha - 1$$

Pour $\alpha = 1/2$, on obtient $v_{n+1} = 3v_n$.
Puisque $v_0 = u_0 - 1/2 = 1/2$, on a $v_n = 3^n/2$ puis

$$u_n = \frac{1}{2}(3^n + 1)$$

15.2 Limite d'une suite réelle

$u = (u_n)$, $v = (v_n)$ et $w = (w_n)$ désignent des suites réelles

15.2.1 Limite finie

15.2.1.1 Définition

Définition

On dit que la suite u tend vers $\ell \in \mathbb{R}$ si l'on a la propriété

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \in \mathbb{N}, n \geqslant N \Rightarrow |u_n - \ell| \leqslant \varepsilon$$

On note alors $u \to \ell$, $u_n \to \ell$ ou $u_n \xrightarrow[n\to+\infty]{} \ell$.

Remarque L'inégalité $|u_n - \ell| \leqslant \varepsilon$ équivaut à l'encadrement $\ell - \varepsilon \leqslant u_n \leqslant \ell + \varepsilon$.
Ceci signifie que u_n est une valeur approchée de ℓ à ε près.

Remarque Verbalement, dire que la suite u tend vers ℓ signifie que pour tout $\varepsilon > 0$, aussi petit soit-il, il existe un rang au-delà duquel tous les termes de la suite sont dans l'intervalle $[\ell - \varepsilon, \ell + \varepsilon]$.

Remarque L'entier introduit N dépend de ε (puisqu'il apparaît après celui-ci dans la définition quantitfiée).
Pour le souligner, on le note parfois N_ε.

Remarque On obtient une définition équivalente en prenant $\forall \varepsilon \in]0, 1]$ au lieu de $\forall \varepsilon > 0$.
En effet un rang N convenant pour $\varepsilon = 1$, convient aussi pour tout $\varepsilon \geqslant 1$.

Remarque Si la suite u n'est définie qu'à partir d'un rang n_0, on adapte la définition en remplaçant « $\forall n \in \mathbb{N}$ »par « $\forall n \geqslant n_0$ ».

Exemple Considérons $u = (u_n)$ constante égale à C. Montrons $u_n \to C$.
Soit $\varepsilon > 0$. Déterminons un rang $N \in \mathbb{N}$ au delà duquel $|u_n - C| \leqslant \varepsilon$.
Ici $|u_n - C| = |C - C| = 0$ donc n'importe quel N convient.
Ainsi pour $N = 42$, on a

$$\forall n \in \mathbb{N}, n \geqslant N \Rightarrow |u_n - C| \leqslant \varepsilon$$

Exemple Considérons $u = (1/n)_{n\geqslant 1}$ définie par $u_n = 1/n$. Montrons $u_n \to 0$.
Soit $\varepsilon > 0$. Déterminons un rang $N \in \mathbb{N}$ au delà duquel $|u_n - 0| \leqslant \varepsilon$.
Ici $|u_n - 0| = \dfrac{1}{n}$. Pour $N = E(1/\varepsilon) + 1$, on a pour tout $n \geqslant N$, $n \geqslant 1/\varepsilon$ donc $1/n \leqslant \varepsilon$.
Ainsi

$$\forall n \in \mathbb{N}^\star, n \geqslant N \Rightarrow |u_n - 0| \leqslant \varepsilon$$

Proposition

Soient $u = (u_n)$ et $\ell \in \mathbb{R}$. On a équivalence entre :
(i) $u_n \to \ell$;
(ii) $|u_n - \ell| \to 0$.

dém. :
$u_n \to \ell$ signifie

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \in \mathbb{N}, n \geqslant N \Rightarrow |u_n - \ell| \leqslant \varepsilon$$

alors que $|u_n - \ell| \to 0$ signifie

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \in \mathbb{N}, n \geqslant N \Rightarrow ||u_n - \ell| - 0| \leqslant \varepsilon$$

Sachant $||u_n - \ell| - 0| = |u_n - \ell|$, ces deux phrases quantifiées sont parfaitement équivalentes.
□

Exemple Si $u_n \to \ell$ alors $-u_n \to -\ell$.
En effet

$$|-u_n - (-\ell)| = |\ell - u_n| = |u_n - \ell| \to 0$$

Remarque Quand on veut montrer que (u_n) tend vers ℓ, il est très fréquent de se « ramener à 0 »en montrant que $|u_n - \ell|$ tend vers 0.

15.2.1.2 Convergence et divergence

Définition

On dit que la suite u converge s'il existe $\ell \in \mathbb{R}$ tel que $u_n \to \ell$.
Sinon on dit que la suite u diverge.
Déterminer la nature d'une suite, c'est savoir si celle-ci est convergente ou divergente.

Exemple La suite $(1/n)_{n\geqslant 1}$ converge.

Exemple $((-1)^n)$ diverge.
En effet, supposons, par l'absurde que $(-1)^n \to \ell$.
Pour $\varepsilon = 1/2$, il existe $N \in \mathbb{N}$ tel que pour tout $n \geqslant N$, on ait $|(-1)^n - \ell| \leqslant 1/2$.
Pour $n \geqslant N$ pair, on obtient $|1 - \ell| \leqslant 1/2$ donc $\ell \in [1/2, 3/2]$.
Pour $n \geqslant N$ impair on obtient $|-1 - \ell| \leqslant 1/2$ donc $\ell \in [-3/2, -1/2]$.
C'est absurde car $[1/2, 3/2] \cap [-3/2, -1/2] = \emptyset$.

Théorème

> Si la suite u converge alors il existe un unique $\ell \in \mathbb{R}$ tel que $u_n \to \ell$.
> Ce réel ℓ est alors appelé limite de la suite u et on note $\ell = \lim u$, $\ell = \lim u_n$ ou $\ell = \lim\limits_{n \to +\infty} u_n$.

dém. :
L'existence est acquise par définition de la convergence.
Montrons l'unicité.
Supposons $u_n \to \ell$ et $u_n \to \ell'$.
Pour tout $\varepsilon > 0$, il existe n assez grand tel que

$$|u_n - \ell| \leqslant \varepsilon/2 \text{ et } |u_n - \ell'| \leqslant \varepsilon/2$$

On a alors

$$|\ell - \ell'| \leqslant |\ell - u_n| + |u_n - \ell'| \leqslant \varepsilon$$

Ainsi

$$\forall \varepsilon > 0, |\ell - \ell'| \leqslant \varepsilon$$

Si par l'absurde $\ell \neq \ell'$ alors $|\ell - \ell'| > 0$ et en prenant $\varepsilon = \frac{1}{2}|\ell - \ell'|$, la propriété qui précède donc $|\ell - \ell'| \leqslant \frac{1}{2}|\ell - \ell'|$ puis $1 \leqslant \frac{1}{2}$ car $|\ell - \ell'| > 0$. C'est absurde et donc $\ell = \ell'$.
□

Exemple On peut écrire $\lim\limits_{n \to +\infty} C = C$ et $\lim\limits_{n \to +\infty} \frac{1}{n} = 0$.

15.2.1.3 Limites finies et relation d'ordre

Théorème

> Toute suite réelle convergente est bornée.

dém. :
Soit u une suite convergeant vers un réel $\ell \in \mathbb{R}$.
Pour $\varepsilon = 1$, il existe un rang $N \in \mathbb{N}$ tel que pour tout $n \geqslant N$, $|u_n - \ell| \leqslant 1$.
On a alors

$$|u_n| = |u_n - \ell + \ell| \leqslant |u_n - \ell| + |\ell| \leqslant 1 + |\ell|$$

pour tout $n \geqslant N$.
Posons alors

$$M = \max\{|u_0|, |u_1|, \ldots, |u_{N-1}|, 1 + |\ell|\}$$

(M existe en tant que max d'une partie finie de $\mathbb{R}$).
Pour tout $n \in \mathbb{N}$, on vérifie $|u_n| \leqslant M$ en discutant selon $n < N$ ou $n \geqslant N$.
On en déduit que la suite u est bornée.
□

Attention : La réciproque est fausse, la suite $((-1)^n)$ est bornée mais divergente.

Proposition

Soient u une suite réelle telle que $u_n \to \ell$ et $a, b \in \mathbb{R}$.
Si $a < \ell$ alors
$$\exists N_0 \in \mathbb{N}, \forall n \geqslant N_0, u_n > a$$
Si $\ell < b$ alors
$$\exists N_1 \in \mathbb{N}, \forall n \geqslant N_1, u_n < b$$
Si $a < \ell < b$ alors
$$\exists N_2 \in \mathbb{N}, \forall n \geqslant N_1, a < u_n < b$$

dém. :
Supposons $a < \ell$ et posons $\varepsilon = (\ell - a)/2 > 0$.
Puisque $u_n \to \ell$, il existe un rang N_0 à partir duquel $|u_n - \ell| \leqslant \varepsilon$.
Pour $n \geqslant N_0$, on a alors
$$u_n \geqslant \ell - \varepsilon = \frac{a + \ell}{2} > a$$
Le cas $\ell < b$, se traite de façon semblable en introduisant $\varepsilon = (b - \ell)/2 > 0$.
Enfin, le cas $a < \ell < b$ se résout en prenant $N_2 = \max(N_0, N_1)$.
□

Exemple Si $u_n \to \ell$ avec $\ell > 0$ alors à partir d'un certain rang $u_n > 0$.

Attention : Ce résultat est faux en terme d'inégalités larges.
Une suite peut tendre vers une limite nulle (donc positive au sens large) sans être nécessairement à termes positifs à partir d'un certain rang.

Exemple Si $u_n \to \ell \neq 0$ alors la suite $(1/u_n)$ est définie à partir d'un certain rang.
En effet, selon le signe de ℓ, $u_n > 0$ ou $u_n < 0$ à partir d'un certain rang et donc $u_n \neq 0$.

Théorème

On suppose que $u_n \to \ell$ et $v_n \to \ell'$.
Si à partir d'un certain rang $u_n \leqslant v_n$ alors $\ell \leqslant \ell'$.

dém. :
Par contraposée, supposons $\ell > \ell'$ et introduisons le milieu $a = (\ell + \ell')/2$.
Puisque $u_n \to \ell > a$, il existe un rang N_1 à partir duquel $u_n > a$.
Puisque $v_n \to \ell' < a$, il existe un rang N_2 à partir duquel $v_n < a$.
Pour tout n au-delà du rang $\max(N_1, N_2)$, on a $v_n < a < u_n$.
Ainsi, il faux qu'à partir d'un certain rang $u_n \leqslant v_n$.
□

Attention : Ce théorème est faux en terme d'inégalités strictes.
Pour $u_n = 1/n$ et $v_n = -1/n$, on a $u_n < v_n$ mais à la limite $\lim u_n = \lim v_n$.
On retient que, par passage à limite, les inégalités strictes deviennent larges.

Exemple On suppose que u est une suite convergente vérifiant

$$\forall n \in \mathbb{N}, u_n \in [a, b]$$

Montrons que $\ell \in [a, b]$.
Pour tout $n \in \mathbb{N}$, on a $u_n \geqslant a$. Sachant que la suite constante égale à a converge vers a, on obtient par passage à la limite $\ell \geqslant a$. De même on montre $\ell \leqslant b$ et on conclut.

Exemple On suppose que u est une suite croissante de réels strictement positifs et on suppose que u converge. Montrons que sa limite ℓ est strictement positive.
Puisque par passage à la limite, les inégalités strictes deviennent larges, passer à la limite la minoration $u_n > 0$ donne seulement $\ell \geqslant 0$.
Pour obtenir $\ell > 0$, nous allons « prendre un appui »en exploitant la croissance de u.
Puisque u est croissante, on a pour tout $n \in \mathbb{N}$, $u_n \geqslant u_0$.
A la limite, on obtient $\ell \geqslant u_0$. Or $u_0 > 0$ donc $\ell > 0$.

15.2.2 Limites infinies

15.2.2.1 Définition

Définition

On dit que la suite u tend vers $+\infty$ si l'on a la propriété

$$\forall A \in \mathbb{R}, \exists N \in \mathbb{N}, \forall n \in \mathbb{N}, n \geqslant N \Rightarrow u_n \geqslant A$$

On note alors $u \to +\infty$, $u_n \to +\infty$ ou $u_n \xrightarrow[n \to +\infty]{} +\infty$.

Remarque La suite u tend vers $+\infty$ si pour tout A, aussi grand soit-il, il existe un rang au-delà duquel tous les termes de la suite sont supérieurs à A.

Remarque On obtient une définition équivalente en remplaçant « $\forall A \in \mathbb{R}$ »par « $\forall A \in \mathbb{R}^+$ », « $\forall A \in [1, +\infty[$ »,...
En effet un rang N fonctionnant pour $A = 1$ conviendra aussi pour tout $A \leqslant 1$.

Exemple Considérons u définie par $u_n = n$ pour tout $n \in \mathbb{N}$ et montrons que $u_n \to +\infty$.
Soit $A \in \mathbb{R}^+$. Déterminons un rang N à partir duquel $u_n \geqslant A$.
Ici $u_n = n$ et donc pour $n \geqslant E(A) + 1$, on a $u_n \geqslant A$.

Définition

On dit que la suite u tend vers $-\infty$ si l'on a la propriété

$$\forall B \in \mathbb{R}, \exists N \in \mathbb{N}, \forall n \in \mathbb{N}, n \geqslant N \Rightarrow u_n \leqslant B$$

On note alors $u \to -\infty$, $u_n \to -\infty$ ou $u_n \xrightarrow[n \to +\infty]{} -\infty$.

Remarque On obtient une définition équivalente en remplaçant « $\forall B \in \mathbb{R}$ »par « $\forall B \in \mathbb{R}^-$ », « $\forall B \in]-\infty, -1]$ »,...

Remarque Si la suite u n'est définie qu'à partir d'un rang n_0, on adapte les définitions précédentes en remplaçant « $\forall n \in \mathbb{N}$ »par « $\forall n \geqslant n_0$ ».

Proposition

$u_n \to +\infty \Leftrightarrow -u_n \to -\infty$.

dém. :
($\Rightarrow$) Supposons $u_n \to +\infty$.
Soit $B \in \mathbb{R}$. Pour $A = -B$, il existe $N \in \mathbb{N}$ tel qu'à partir du rang N, $u_n \geqslant A$.
On a alors $-u_n \leqslant -A = B$ à partir du même rang N et donc on peut affirmer que $-u_n \to +\infty$.
($\Leftarrow$) Même principe.
□

Exemple $-n \to -\infty$

15.2.2.2 Limites infinies et relation d'ordre

Proposition

Si $u_n \to +\infty$ alors u est minorée, non majorée.
Si $u_n \to -\infty$ alors u est majorée, non minorée.

dém. :
Supposons $u_n \to +\infty$.
Pour $A = 0$, il existe un rang $N \in \mathbb{N}$ à partir duquel $u_n \geqslant 0$.
Pour $m = \min \{u_0, u_1, \ldots, u_{N-1}, 0\}$ (m existe en tant que min d'une partie finie de $\mathbb{R}$).
Pour tout $n \in \mathbb{N}$, on vérifie aisément $u_n \geqslant m$ en discutant selon $n < N$ ou $n \geqslant N$.
Ainsi la suite u est minorée.
Montrons maintenant qu'elle n'est pas majorée.
Soit $M \in \mathbb{R}$. Pour $A = M + 1 \in \mathbb{R}$, il existe un rang N à partir duquel $u_n \geqslant A$.
En particulier $u_N \geqslant A > M$. Ainsi la suite u n'est pas majorée par M et ce, quel que soit $M \in \mathbb{R}$.
Par passage à l'opposé, on transpose l'étude aux suites tendant vers $-\infty$.
□

Attention : Les réciproques sont fausses.
La suite de terme général $u_n = n(1 + (-1)^n)$ est minorée (par 0) et non majorée, cependant elle ne tend pas vers $+\infty$ car tous ses termes d'indices impairs sont nuls.

Remarque Si $u_n \to +\infty$ ou $u_n \to -\infty$ alors la suite u n'est pas bornée et est donc divergente.

Définition

Si $u_n \to +\infty$ (resp. $u_n \to -\infty$) alors on dit que la suite u diverge vers $+\infty$ (resp. $-\infty$).
On dit alors que $+\infty$ (resp. $-\infty$) est la limite de la suite u et on note $\lim u = +\infty$ (resp. $\lim u = -\infty$).

Exemple $\lim\limits_{n\to+\infty} n = +\infty$ et $\lim\limits_{n\to+\infty} -n = -\infty$.

Remarque Les comportements possibles d'une suite sont :
- converger vers une limite finie ;
- diverger vers une limite infinie ;
- diverger sans limite.

15.2.3 Opération sur les limites

15.2.3.1 Somme

Lemme

Si u est minorée et $v_n \to +\infty$ alors $u_n + v_n \to +\infty$.

dém. :
Supposons u minorée par m et $v_n \to +\infty$.
On a $u_n + v_n \geqslant m + v_n$.
Soit $A \in \mathbb{R}$. Pour $A' = A - M$, il existe $N \in \mathbb{N}$ tel que pour tout $n \geqslant N$, $v_n \geqslant A'$ et donc $u_n + v_n \geqslant M + A' = A$.
Ainsi $\forall A \in \mathbb{R}, \exists N \in \mathbb{N}, \forall n \in \mathbb{N}, n \geqslant N \Rightarrow u_n + v_n \geqslant A$.
□

Théorème

1) Si $u_n \to \ell \in \mathbb{R}$ et $v_n \to \ell' \in \mathbb{R}$ alors $u_n + v_n \to \ell + \ell'$.
2) Si $u_n \to \ell \in \mathbb{R}$ et $v_n \to +\infty$ alors $u_n + v_n \to +\infty$.
3) Si $u_n \to \ell \in \mathbb{R}$ et $v_n \to -\infty$ alors $u_n + v_n \to -\infty$.
4) Si $u_n \to +\infty$ et $v_n \to +\infty$ alors $u_n + v_n \to +\infty$.
5) Si $u_n \to -\infty$ et $v_n \to -\infty$ alors $u_n + v_n \to -\infty$.

dém. :
1) Supposons $u_n \to \ell \in \mathbb{R}$ et $v_n \to \ell' \in \mathbb{R}$.
On remarque que

$$|(u_n + v_n) - (\ell + \ell')| \leqslant |u_n - \ell| + |v_n - \ell'|$$

Il suffit donc de rendre $|u_n - \ell|$ et $|v_n - \ell'|$ inférieurs à $\varepsilon/2$ pour conclure...
Soit $\varepsilon > 0$.
Puisque $u_n \to \ell$ et $v_n \to \ell'$, il existe $N_1, N_2 \in \mathbb{N}$ tels que

$$\forall n \geqslant N_1, |u_n - \ell| \leqslant \varepsilon/2 \text{ et } \forall n \geqslant N_2, |v_n - \ell'| \leqslant \varepsilon/2$$

Pour $N_0 = \max(N_1, N_2)$ on a pour tout $n \geqslant N_0$,

$$|(u_n + v_n) - (\ell + \ell')| \leqslant |u_n - \ell| + |v_n - \ell'| \leqslant \varepsilon/2 + \varepsilon/2 = \varepsilon$$

Ainsi

$$\forall \varepsilon > 0, \exists N_0 \in \mathbb{N}, \forall n \in \mathbb{N}, n \geqslant N_0 \Rightarrow |(u_n + v_n) - (\ell + \ell')| \leqslant \varepsilon$$

2) 3) 4) et 5) découlent du résultat précédent après un éventuel un passage à l'opposé.
□

Remarque Si $u_n \to +\infty$ et $v_n \to -\infty$ on ne peut rien dire a priori ; c'est une forme indéterminée. Pour lever l'indétermination il convient d'exprimer autrement la somme $u_n + v_n$, par exemple « en factorisant le terme qui l'emporte »ou « en simplifiant ce que les termes ont en commun »...

Corollaire

Par passage à l'opposé, on obtient les règles relatives aux différences de limites.

Exemple On suppose que u et $u+v$ convergent.
Montrons que v converge.
Pour tout $n \in \mathbb{N}$, on peut écrire $v_n = (u_n + v_n) - u_n$ donc par différence de suites convergentes, la suite v est convergente.

Attention : L'implication $u + v$ converge $\Rightarrow u$ et v convergent est fausse.
Il suffit de prendre $u_n = (-1)^n$ et $v_n = -u_n$ pour obtenit un contre-exemple.

15.2.3.2 Produit

Lemme

Supposons qu'à partir d'un certain rang, (u_n) est minorée par un réel $\rho > 0$.
Si $v_n \to +\infty$ alors $u_n v_n \to +\infty$.
Si $v_n \to -\infty$ alors $u_n v_n \to -\infty$.

dém. :
Par hypothèse, il existe un rang $N_1 \in \mathbb{N}$ tel que $\forall n \geqslant N_1, u_n \geqslant \rho$.
Supposons $v_n \to \infty$.
On remarque $u_n v_n \geqslant \rho v_n$.
Soit $A \in \mathbb{R}^+$. Pour $A' = A/\rho$, il existe $N_2 \in \mathbb{N}$ tel que pour tout $n \geqslant N_2$, $v_n \geqslant A'$ et alors pour $n \geqslant N_0 = \max(N_1, N_2)$, on a $u_n v_n \geqslant \rho A' = A$.
Ainsi

$$\forall A \in \mathbb{R}, \exists N_0 \in \mathbb{N}, \forall n \in \mathbb{N}, n \geqslant N_0 \Rightarrow u_n v_n \geqslant A$$

Le cas $v_n \to -\infty$ se déduit du précédent par passage à l'opposé.
$\square$

Théorème

1) Si $u_n \to \ell \in \mathbb{R}$ et $v_n \to \ell' \in \mathbb{R}$ alors $u_n v_n \to \ell\ell'$.
2) Si $u_n \to \ell \neq 0$ et $v_n \to +\infty$ alors

$$u_n v_n \to \begin{cases} +\infty \text{ si } \ell > 0 \\ -\infty \text{ si } \ell < 0 \end{cases}$$

3) Si $u_n \to \ell \neq 0$ et $v_n \to -\infty$ alors

$$u_n v_n \to \begin{cases} -\infty \text{ si } \ell > 0 \\ +\infty \text{ si } \ell < 0 \end{cases}$$

4) Si $u_n \to +\infty$ et $v_n \to +\infty$ alors $u_n v_n \to +\infty$.
5) Si $u_n \to +\infty$ et $v_n \to -\infty$ alors $u_n v_n \to -\infty$.
6) Si $u_n \to -\infty$ et $v_n \to -\infty$ alors $u_n v_n \to +\infty$.

dém. :
1) Supposons $u_n \to \ell \in \mathbb{R}$ et $v_n \to \ell' \in \mathbb{R}$.

On remarque

$$|u_n v_n - \ell\ell'| = |u_n v_n - u_n\ell' + u_n\ell' - v_n\ell'| \leqslant |u_n|\,|v_n - \ell'| + |\ell'|\,|u_n - \ell|$$

Puisque la suite (u_n) converge, elle est bornée et donc il existe $M \in \mathbb{R}^+$ tel que

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

et alors

$$|u_n v_n - \ell\ell'| \leqslant M\,|v_n - \ell'| + |\ell'|\,|u_n - \ell|$$

Il suffit alors de rendre $|u_n - \ell|$ et $|v_n - \ell'|$ assez petit...
Soit $\varepsilon > 0$.
Puisque $u_n \to \ell$ et $v_n \to \ell'$, il existe $N_1, N_2 \in \mathbb{N}$ tel que

$$\forall n \geqslant N_1, |u_n - \ell| \leqslant \frac{\varepsilon}{2(|\ell'|+1)} \text{ et } \forall n \geqslant N_2, |v_n - \ell'| \leqslant \frac{\varepsilon}{2(M+1)}.$$

□
Pour $N_0 = \max(N_1, N_2)$ on a

$$\forall n \geqslant N_0, |u_n v_n - \ell\ell'| \leqslant \frac{M\varepsilon}{2(M+1)} + \frac{|\ell'|\,\varepsilon}{2(|\ell'|+1)} \leqslant \varepsilon$$

Ainsi

$$\forall \varepsilon > 0, \exists N_0 \in \mathbb{N}, \forall n \in \mathbb{N}, n \geqslant N_0 \Rightarrow |u_n v_n - \ell\ell'| \leqslant \varepsilon$$

2) 3) 4) 5) et 6) découlent directement du résultat précédent après un éventuel passage à l'opposé.

Remarque Si $u_n \to 0$ et $v_n \to +\infty$ (ou $-\infty$) on ne peut rien dire a priori sur $u_n v_n$; c'est une forme indéterminée.
Pour résoudre l'indéterminatrion, il convient de réécrire le produit.

Exemple Etudions

$$\lim_{n \to +\infty} (n^2 - n)$$

Quand $n \to +\infty$, $n^2 \to +\infty$ et $n \to +\infty$, c'est une forme indéterminée.
Cependant

$$n^2 - n = n^2\left(1 - \frac{1}{n}\right)$$

avec $n^2 \to +\infty$ et $1 - 1/n \to 1$ donc $n^2 - n \to +\infty$.

Remarque On notera sur cet exemple la mise en forme de la rédaction. Eviter autant que possible, les rédactions de la forme $\lim () = \lim () = \ldots$ En effet ce type de rédaction présume d'entrée de jeu l'existence de la limite alors que celle-ci n'est acquise qu'au terme de l'étude. La rédaction précédente ne pose pas ce problème et présente encore bien d'autres avantages !

15.2.3.3 Passage à l'inverse

Définition

Soit u une suite réelle et $\ell \in \mathbb{R}$.
On note $u_n \to \ell^+$ pour signifier que $u_n \to \ell$ et $u_n > \ell$ à partir d'un certain rang.
On note aussi $u_n \to \ell^-$ pour signifier $u_n \to \ell$ et $u_n < \ell$ pour n assez grand.

Exemple On peut écrire $\lim\limits_{n\to+\infty} 1/n = 0^+$.

Théorème

1) Si $u_n \to \ell \neq 0$ alors $1/u_n \to 1/\ell$.
2) Si $u_n \to 0^+$ alors $1/u_n \to +\infty$.
3) Si $u_n \to 0^-$ alors $1/u_n \to -\infty$.
4) Si $u_n \to +\infty$ alors $1/u_n \to 0^+$.
5) Si $u_n \to -\infty$ alors $1/u_n \to 0^-$.

dém. :
1) Supposons $u_n \to \ell$.
Cas $\ell > 0$:
Pour n assez grand $u_n > 0$ et donc la suite $(1/u_n)$ est définie à partir d'un certain rang.
On remarque que

$$\left|\frac{1}{u_n} - \frac{1}{\ell}\right| = \frac{|\ell - u_n|}{|u_n|\,|\ell|}$$

Sachant $\ell > \ell/2$, il existe un rang $N_1 \in \mathbb{N}$ tel que

$$\forall n \geqslant N_1, u_n > \ell/2$$

et alors

$$\left|\frac{1}{u_n} - \frac{1}{\ell}\right| \leqslant \frac{2\,|\ell - u_n|}{\ell^2}$$

Soit $\varepsilon > 0$, pour $\varepsilon' = \ell^2\varepsilon/2 > 0$, il existe $N_2 \in \mathbb{N}$ tel que

$$\forall n \geqslant N_2, |u_n - \ell| \leqslant \varepsilon'$$

Pour $N_0 = \max(N_1, N_2)$ on a

$$\forall n \geqslant N_0, \left|\frac{1}{u_n} - \frac{1}{\ell}\right| \leqslant \varepsilon$$

Ainsi

$$\frac{1}{u_n} \to \frac{1}{\ell}$$

Cas $\ell < 0$:
Il suffit de transiter par la suite $(-u_n)$.
2) Supposons $u_n \to 0^+$.
Il existe un rang $N_1 \in \mathbb{N}$ tel que

$$\forall n \geqslant N_1, u_n > 0$$

et donc la suite $(1/u_n)$ est définie à partir d'un certain rang.
Soit $A \in \mathbb{R}^{+\star}$. Pour $\varepsilon = 1/A > 0$, il existe $N_2 \in \mathbb{N}$ tel que

$$\forall n \geqslant N_2, |u_n| \leqslant \varepsilon$$

Pour $n \geqslant N_0 = \max(N_1, N_2)$, on a $u_n > 0$ et $|u_n| \leqslant \varepsilon$ donc

$$\frac{1}{u_n} \geqslant \frac{1}{\varepsilon} = A$$

Ainsi

$$\frac{1}{u_n} \to +\infty$$

3) Se déduit de l'étude précédente par passage à l'opposé.
4) Supposons $u_n \to +\infty$.
Pour $A = 1$, il existe un rang $N_1 \in \mathbb{N}$ tel que $\forall n \geqslant N_1, u_n \geqslant 1$ et donc la suite $(1/u_n)$ est définie à partir d'un certain rang.
Soit $\varepsilon > 0$, pour $A = 1/\varepsilon > 0$, il existe $N_0 \in \mathbb{N}$ tel que

$$\forall n \geqslant N_0, u_n \geqslant A > 0$$

On a alors

$$0 < \frac{1}{u_n} < \frac{1}{A}$$

Ainsi $u_n \to 0^+$
5) Se déduit de l'étude précédente par passage à l'opposé.
□

Corollaire

On en déduit les règles de calculs relatives aux rapports de limites.

Exemple Etudions

$$\lim_{n \to +\infty} \frac{n^3 + n + 1}{n^2 + 1}$$

Quand $n \to +\infty$, $n^3 + n + 1 \to +\infty$ et $n^2 + 1 \to +\infty$.
C'est une forme indéterminée.

$$\frac{n^3 + n + 1}{n^2 + 1} = \frac{n^3}{n^2} \frac{1 + 1/n^2 + 1/n^3}{1 + 1/n^2}$$

Or

$$\frac{n^3}{n} = n \to +\infty \text{ et } \frac{1 + 1/n^2 + 1/n^3}{1 + 1/n^2} \to 1$$

donc

$$\frac{n^3 + n + 1}{n^2 + 1} \to 1$$

Exemple Etudions

$$\lim_{n \to +\infty} \frac{n^2 - 1}{n^2 + 1}$$

Quand $n \to +\infty$,

$$\frac{n^2 - 1}{n^2 + 1} = \frac{1 - 1/n^2}{1 + 1/n^2} \to 1$$

15.2.3.4 Composition de limites

Théorème

Soit $f : \mathcal{D} \subset \mathbb{R} \to \mathbb{R}$ et $u = (u_n)$ une suite réelle telle qu'à partir d'un certain rang $u_n \in \mathcal{D}$.
Si $u_n \to a \in \bar{\mathbb{R}}$ et $f(x) \xrightarrow[x \to a]{} \ell \in \bar{\mathbb{R}}$ alors $f(u_n) \to \ell$.

dém. :
Cette démonstration ne peut être menée sans avoir préalablement défini $f(x) \xrightarrow[x \to a]{} \ell \in \bar{\mathbb{R}}$. Cela sera fait dans le chapitre « Fonctions Numériques »et l'on y trouvera la démonstration dans une version plus forte appelée « caractérisation séquentielle des limites ». Dans la suite de ce cours nous anticipons ce résultat...
□

Exemple Etudions de

$$\lim_{n \to +\infty} \sqrt{n+1}$$

Quand $n \to +\infty$, $n + 1 \to +\infty$.
Or $\sqrt{t} \xrightarrow[t \to +\infty]{} +\infty$ donc par composition de limites $\sqrt{n+1} \to +\infty$.

Remarque Rappelons quelques limites utiles :

$$\lim_{x \to 0} \frac{\sin x}{x} = \lim_{x \to 0} \frac{\ln(1+x)}{x} = 1$$

Exemple Etudions de

$$\lim_{n \to +\infty} \sqrt{n} \sin \frac{1}{\sqrt{n}}$$

Quand $n \to +\infty$, $\sqrt{n} \to +\infty$ donc $1/\sqrt{n} \to 0$ puis sachant

$$\frac{\sin x}{x} \xrightarrow[x \to 0]{} 1$$

on obtient

$$\sqrt{n} \sin \frac{1}{\sqrt{n}} \to 1$$

Exemple Etudions

$$\lim_{n \to +\infty} (\sqrt{n+1} - \sqrt{n})$$

Quand $n \to +\infty$,

$$\sqrt{n+1} - \sqrt{n} = \frac{(n+1) - n}{\sqrt{n+1} + \sqrt{n}} = \frac{1}{\sqrt{n+1} + \sqrt{n}}$$

Or $\sqrt{n+1} + \sqrt{n} \to +\infty$ donc

$$\sqrt{n+1} - \sqrt{n} \to 0$$

Exemple Etudions
$$\lim_{n\to+\infty} n^{1/n}$$
Quand $n \to +\infty$, $n \to +\infty$ et $1/n \to 0$, c'est une forme indéterminée.
$$n^{1/n} = \exp\left(\frac{1}{n}\ln n\right)$$
or par limite de référence $\dfrac{\ln n}{n} \to 0$ donc par composition
$$n^{1/n} \to 1$$

Exemple Etudions
$$\lim_{n\to+\infty}\left(1+\frac{1}{n}\right)^n$$
Quand $n \to +\infty$, $1+\dfrac{1}{n} \to 1$ et $n \to +\infty$, c'est une forme indéterminée.
$$\left(1+\frac{1}{n}\right)^n = \exp\left(n\ln\left(1+\frac{1}{n}\right)\right)$$
Or
$$n\ln\left(1+\frac{1}{n}\right) = \frac{\ln\left(1+1/n\right)}{1/n}$$
avec $1/n \to 0$. Sachant
$$\frac{\ln(1+x)}{x} \xrightarrow[x\to 0]{} 1$$
on a par composition de limites $n\ln\left(1+\dfrac{1}{n}\right) \to 1$ puis
$$\left(1+\frac{1}{n}\right)^n \to \mathrm{e}$$

15.2.4 Etude de limite par comparaison

15.2.4.1 Théorème d'encadrement

Théorème

On suppose qu'à partir d'un certain rang
$$v_n \leqslant u_n \leqslant w_n$$

Si (v_n) et (w_n) convergent vers $\ell \in \mathbb{R}$ alors $u_n \to \ell$.
dém. :
Par hypothèse, il existe un rang $N_1 \in \mathbb{N}$ tel que
$$\forall n \geqslant N_1, v_n \leqslant u_n \leqslant w_n$$

Supposons que (v_n) et (w_n) convergent vers $\ell \in \mathbb{R}$.
Soit $\varepsilon > 0$.
Il existe $N_2, N_3 \in \mathbb{N}$ tel que

$$\forall n \geqslant N_2, |v_n - \ell| \leqslant \varepsilon \text{ et } \forall n \geqslant N_3, |w_n - \ell| \leqslant \varepsilon$$

Rappelons que $|x - \ell| \leqslant \varepsilon$ signifie $\ell - \varepsilon \leqslant x \leqslant \ell + \varepsilon$.
Pour $n \geqslant \max(N_1, N_2, N_3)$, on a $v_n \leqslant u_n \leqslant w_n$ avec $v_n \geqslant \ell - \varepsilon$ et $w_n \leqslant \ell + \varepsilon$ donc $\ell - \varepsilon \leqslant u_n \leqslant \ell + \varepsilon$ puis $|u_n - \ell| \leqslant \varepsilon$.
Ainsi $u_n \to \ell$.

Exemple Soit $x \in \mathbb{R}$, étudions la limite de

$$u_n = \frac{E(nx+1)}{n+1}$$

Pour tout $n \in \mathbb{N}$,

$$nx \leqslant E(nx+1) \leqslant nx + 1$$

donc

$$\frac{nx}{n+1} \leqslant u_n \leqslant \frac{nx+1}{n+1}$$

Or $\dfrac{nx}{n+1} \to x$ et $\dfrac{nx+1}{n+1} \to x$ donc par encadrement

$$u_n \to x$$

Attention : Ne pas rédiger :
« $v_n \leqslant u_n \leqslant w_n$ donc $\lim v_n \leqslant \lim u_n \leqslant \lim w_n$ or $\lim v_n = \lim w_n = \ell \in \mathbb{R}$ donc $\lim u_n = \ell$. »
Ce raisonnement est maladroit car présuppose déjà acquise l'existence de limite étudiée ; par le théorème des gendarmes l'argumentation est différente car on y démontrer l'existence de la limite tout en donnant sa valeur.

Exemple Etudions la limite de

$$u_n = (2 + \sin n)^{1/n}$$

Pour tout $n \in \mathbb{N}$,

$$1 \leqslant 2 + \sin n \leqslant 3$$

donc

$$1 \leqslant u_n \leqslant 3^{1/n}$$

Or $3^{1/n} \to 1$ donc par encadrement

$$u_n \to 1$$

Exemple Etudions la limite de

$$u_n = \sum_{k=1}^{n} \frac{1}{n + \sqrt{k}}$$

Pour étudier la limite d'une somme il est fréquent d'encadre le terme sommé par le plus petit et le plus grand des termes.

Ici, pour tout $1 \leqslant k \leqslant n$,

$$\frac{1}{n+\sqrt{n}} \leqslant \frac{1}{n+\sqrt{k}} \leqslant \frac{1}{n+1}$$

donc

$$\sum_{k=1}^{n} \frac{1}{n+\sqrt{n}} \leqslant u_n \leqslant \sum_{k=1}^{n} \frac{1}{n+1}$$

puis

$$\frac{n}{n+\sqrt{n}} \leqslant u_n \leqslant \frac{n}{n+1}$$

Or $\dfrac{n}{n+\sqrt{n}} \to 1$ et $\dfrac{n}{n+1} \to 1$ donc par encadrement

$$u_n \to 1$$

15.2.4.2 Obtention de convergence

Théorème

> Si à partir d'un certain rang $|u_n - \ell| \leqslant v_n$ et si $v_n \to 0$ alors $u_n \to \ell$.

dém. :
On suppose qu'il existe $N_1 \in \mathbb{N}$ tel que

$$\forall n \geqslant N_1, |u_n - \ell| \leqslant v_n$$

et on suppose $v_n \to 0$.
Soit $\varepsilon > 0$.
Puisque $v_n \to 0$, il existe $N_2 \in \mathbb{N}$ tel que

$$\forall n \geqslant N_2, |v_n - 0| \leqslant \varepsilon$$

et donc $v_n \leqslant \varepsilon$.
Pour $N_0 = \max(N_1, N_2)$, on a

$$\forall n \geqslant N_0, |u_n - \ell| \leqslant \varepsilon$$

□

Remarque Lorsque l'on présume $u_n \to \ell$ et qu'on ne peut l'établir par opération, il est fréquent d'étudier $|u_n - \ell|$ et de majorer $|u_n - \ell|$ par le terme d'une suite de limite nulle.
Cette démarche est souvent plus efficace que le théorème des gendarmes car on n'y utilise qu'une inégalité au lieu de deux, car les valeurs absolues rendent positives les quantités manipulées, ce qui est pratique pour la multiplication d'inégalités. Cependant cette démarche nécessite de l'intuition car pour l'initier il faut avoir deviner quelle est la limite de (u_n).

Exemple Etudions la limite de

$$u_n = \frac{n+\sin n}{n+1}$$

On présume $u_n \to 1$ mais on ne peut le démontrer par opération car la suite $(\sin n)$ diverge.

Etudion alors $|u_n - 1|$.

$$|u_n - 1| = \frac{|\sin n - 1|}{n+1} \leqslant \frac{|\sin n| + 1}{n+1} \leqslant \frac{2}{n+1} \to 0$$

donc

$$u_n \to 1$$

Remarque La démarche par comparaison est souvent utilisé pour étudier les limites de sommes ou d'intégrales.

Exemple Etudions la limite de

$$u_n = \sum_{k=1}^{n} \frac{\sin(k)}{n^2+k}$$

On présume que $u_n \to 0$ donc étudions $|u_n - 0|$.

$$|u_n - 0| \leqslant \sum_{k=1}^{n} \frac{|\sin k|}{n^2+k} \leqslant \sum_{k=1}^{n} \frac{1}{n^2+1} = \frac{n}{n^2+1} \to 0$$

donc

$$u_n \to 0$$

Ici la démarche $\dfrac{\sin k}{n^2+k^2} \to 0$ donc $\displaystyle\sum_{k=1}^{n} \frac{\sin k}{n^2+k^2} \to 0$ est incorrecte car la somme contient un nombre $n \to +\infty$ de termes.

Exemple Etudions la limite de

$$u_n = \int_0^1 t^n \cos(nt)\,\mathrm{d}t$$

On présume que $u_n \to 0$ donc étudions $|u_n - 0|$.

$$|u_n - 0| \leqslant \int_0^1 |t^n|\,|\cos(nt)|\,\mathrm{d}t \leqslant \int_0^1 t^n\,\mathrm{d}t = \frac{1}{n+1} \to 0$$

donc $u_n \to 0$.

Ici la démarche $\forall t \in [0,1[\,, t^n \cos(nt) \to 0$ donc $\displaystyle\int_0^1 t^n \cos(nt)\,\mathrm{d}t \to 0$ est incorrecte, les théorèmes qui permettent de passer des limites à l'intégrales nécessitent plus d'hypothèses, ils seront présentés en seconde année.

Proposition

Si (u_n) est bornée et $v_n \to 0$ alors $u_n v_n \to 0$.

dém. :
Soit $M \in \mathbb{R}$ tel que $|u_n| \leqslant M$ pour tout $n \in \mathbb{N}$.
On a $|u_n v_n - 0| = |u_n|\,|v_n| \leqslant M\,|v_n| \to 0$ donc $u_n v_n \to 0$
□

Exemple $u_n = \dfrac{(-1)^n}{n} \to 0$ par produit d'une suite bornée par une suite de limite nulle.

15.2.4.3 Obtention de limite infinie

Théorème

On suppose qu'à partir d'un certain rang $u_n \leqslant v_n$.
Si $u_n \to +\infty$ alors $v_n \to +\infty$.
Si $v_n \to -\infty$ alors $u_n \to -\infty$.

dém. :
Par hypothèse, il existe $N_1 \in \mathbb{N}$ tel que

$$\forall n \geqslant N_1, u_n \leqslant v_n$$

Supposons $u_n \to +\infty$.
Soit $A \in \mathbb{R}$.
Il existe $N_2 \in \mathbb{N}$ tel que

$$\forall n \geqslant N_2, u_n \geqslant A$$

Pour $n \geqslant N_0 = \max(N_1, N_2)$, $u_n \leqslant v_n$ et $u_n \geqslant A$ donc $v_n \geqslant A$.
Ainsi $v_n \to +\infty$.
L'étude quand $v_n \to -\infty$ se déduit de la précédent par passage à l'opposé.
□

Remarque Lorsqu'on présume $u_n \to +\infty$ et qu'on ne peut l'établir par un calcul direct, il est fréquent d'essayer de minorer u_n par le terme d'une suite de limite $+\infty$.

Exemple Etudions la limite de $u_n = n!$.
Pour $n \geqslant 1$,

$$u_n = 1 \times 2 \times \ldots \times (n-1) \times n \geqslant 1 \times \ldots \times 1 \times n = n$$

or $n \to +\infty$ donc

$$u_n \to +\infty$$

Exemple Etudions la limite de

$$u_n = \sum_{k=1}^{n} \frac{1}{\sqrt{k}}$$

Pour $1 \leqslant k \leqslant n$, $1 \leqslant \sqrt{k} \leqslant \sqrt{n}$ donc

$$\frac{1}{\sqrt{k}} \geqslant \frac{1}{\sqrt{n}}$$

Par suite,

$$u_n \geqslant \sum_{k=1}^{n} \frac{1}{\sqrt{n}} = \frac{n}{\sqrt{n}} = \sqrt{n}$$

or $\sqrt{n} \to +\infty$ donc

$$u_n \to +\infty$$

15.2.5 Limite des suites monotones

Théorème

Si u est une suite croissante et majorée alors u converge.
Si u est une suite croissante non majorée alors u diverge vers $+\infty$.
Si u est une suite décroissante et minorée alors u converge.
Si u est une suite décroissante non minorée alors u diverge vers $-\infty$.

dém. :
Considérons une suite u croissante et majorée.
Posons $\ell = \sup_{n\in\mathbb{N}} u_n \in \mathbb{R}$ et montrons $u_n \to \ell$.
On a déjà

$$\forall n \in \mathbb{N}, u_n \leqslant \ell$$

car $\ell = \sup_{n\in\mathbb{N}} u_n$ majore la suite u.
Soit $\varepsilon > 0$.
Comme $\ell - \varepsilon < \ell = \sup_{n\in\mathbb{N}} u_n$, $\ell - \varepsilon$ n'est pas majorant de la suite u et donc il existe $N \in \mathbb{N}$ vérifiant $u_N > \ell - \varepsilon$.
Par croissance de la suite u, on a alors

$$\forall n \geqslant N, u_n \geqslant u_N > \ell - \varepsilon$$

Par suite, pour tout $n \geqslant N$, $\ell - \varepsilon \leqslant u_n \leqslant \ell$ donc $|u_n - \ell| \leqslant \varepsilon$.
Finalement $u_n \to \ell$.
Soit maintenant une suite u croissante non majorée.
Soit $A \in \mathbb{R}$.
La suite u n'est pas majorée par A donc il existe $N \in \mathbb{N}$ vérifiant $u_N > A$.
Par croissance de la suite u on a alors $\forall n \geqslant N, u_n \geqslant u_N$ et donc $u_n \geqslant A$.
Ainsi $u_n \to +\infty$
Les deux autres implications du théorème s'obtiennent par passage à l'opposé.
□

Remarque La monotonie de la suite u assure l'existence de sa limite.
De plus
Si (u_n) est croissante alors $\lim_{n^\infty} u_n = \sup_{n\in\mathbb{N}} u_n$.
Si (u_n) est décroissante alors $\lim_{n^\infty} u_n = \inf_{n\in\mathbb{N}} u_n$.

Remarque L'énoncé du théorème se généralise aux suites monotones à partir d'un certain rang.

Exemple Si (u_n) est croissante, $u_n \leqslant v_n$ et (v_n) converge alors (u_n) converge.
En effet, puisque (v_n) converge, la suite (v_n) est majorée et donc puisque $u_n \leqslant v_n$, la suite (u_n) est aussi majorée, or on suppose que (u_n) est croissante donc (u_n) converge.

Exemple Etudions la nature de la suite (u_n) définie par

$$u_n = \sum_{k=0}^{n} \frac{1}{k!}$$

On a

$$u_{n+1} - u_n = \sum_{k=0}^{n+1} \frac{1}{k!} - \sum_{k=0}^{n} \frac{1}{k!} = \frac{1}{(n+1)!} \geqslant 0$$

donc (u_n) est croissante.
Montrons que (u_n) est majorée.
Pour cela exploitons la propriété

$$\forall k \geqslant 1, k! = 1 \times 2 \times \ldots \times k \geqslant 1 \times 2 \times \ldots \times 2 = 2^{k-1}$$

On a alors

$$u_n = \sum_{k=0}^{n} \frac{1}{k!} = 1 + \sum_{k=1}^{n} \frac{1}{k!} \leqslant 1 + \sum_{k=1}^{n} \frac{1}{2^{k-1}} = 1 + \frac{1 - 1/2^n}{1 - 1/2} \leqslant 1 + 2 = 3$$

La suite (u_n) est croissante et majorée donc convergente.
On peut montrer que $u_n \to \mathrm{e}$ mais c'est une autre histoire...

Exemple Soit $n \in \mathbb{N}$. Montrons que l'équation $nx + \ln x = 0$ possède une unique solution $x_n \in \mathbb{R}^{+\star}$ et étudions la limite de la suite $(x_n)_{n\in\mathbb{N}}$ ainsi définie.
Considérons la fonction $f_n : x \mapsto nx + \ln x$ définie sur $\mathbb{R}^{+\star}$.
f_n est continue et strictement croissante par opérations.
Sa limite en 0^+ est $-\infty$ et sa limite en $+\infty$ est $+\infty$.
On peut donc affirmer que f_n réalise une bijection de $]0, +\infty[$ vers $\mathbb{R}$.
En particulier l'équation $f_n(x) = 0$ admet une unique solution $x_n \in \mathbb{R}^{+\star}$, à savoir $x_n = f_n^{-1}(0)$.
Etudions la monotonie de (x_n).
On a

$$(n+1)x_{n+1} + \ln x_{n+1} = 0$$

donc

$$f_n(x_{n+1}) = -x_{n+1} < 0 = f_n(x_n)$$

Puisque la fonction f_n est strictement croissante, on a $x_{n+1} < x_n$.
La suite (x_n) est donc (x_n) est strictement décroissante.
De plus, cette suite est minorée par 0, elle est donc convergente. Posons ℓ sa limite.
Puisque pour tout $n \in \mathbb{N}$, $u_n > 0$, à la limite on obtient $\ell \geqslant 0$.
Si par l'absurde $\ell > 0$ alors en passant la relation $nx_n + \ln x_n = 0$ à la limite, on obtient $+\infty = 0$.
C'est absurde et donc nécessairement $\ell = 0$.

15.2.6 Suites adjacentes

Définition

> Deux suites réelles a et b sont dites adjacentes si a est croissante, b décroissante et si $b_n - a_n \to 0$

Exemple Une suite croissante et une suite décroissante de même limite sont adjacentes.

Théorème

Si a et b sont deux suites réelles adjacentes alors elles convergent vers une même limite ℓ et encadrent celle-ci

$$\forall n \in \mathbb{N}, a_n \leqslant \ell \leqslant b_n$$

dém. :
Montrons

$$\forall n \in \mathbb{N}, a_n \leqslant b_n$$

Soit $m \geqslant n$.
Par monotonie, on a $a_m \geqslant a_n$ et $b_m \leqslant b_n$ donc

$$b_n - a_n \geqslant b_m - a_m$$

A la limite quand $m \to +\infty$ on obtient $b_n - a_n \geqslant 0$. Ainsi

$$\forall n \in \mathbb{N}, a_n \leqslant b_n$$

Puisque (b_n) est décroissant, on en déduit que pour tout $n \in \mathbb{N}$, $a_n \leqslant b_n \leqslant b_0$.
La suite (a_n) est donc majorée et puisqu'aussi croissante, elle converge vers une limite $\ell \in \mathbb{R}$
De plus $b_n = a_n + (b_n - a_n) \to \ell$ et ainsi (a_n) et (b_n) convergent vers une même limite ℓ.
Enfin puisque (a_n) est croissante et $a_n \to \ell$, on a $\forall n \in \mathbb{N}, a_n \leqslant \ell$ et aussi puisque (b_n) est décroissante et $b_n \to \ell$, on a $\forall n \in \mathbb{N}, \ell \leqslant b_n$.
□

Exemple Considérons les suites (a_n) et (b_n) de termes généraux

$$a_n = \sum_{k=n+1}^{2n} \frac{1}{k} \text{ et } b_n = \sum_{k=n}^{2n} \frac{1}{k}$$

On a

$$a_{n+1} - a_n = \sum_{k=n+2}^{2n+2} \frac{1}{k} - \sum_{k=n+1}^{2n} \frac{1}{k}$$

Après simplification des termes communs aux deux sommes,

$$a_{n+1} - a_n = \frac{1}{2n+2} + \frac{1}{2n+1} - \frac{1}{n+1} = \frac{1}{2n+1} - \frac{1}{2n+2}$$

Or $\dfrac{1}{2n+1} \geqslant \dfrac{1}{2n+2}$ donc $a_{n+1} - a_n \geqslant 0$.
La suite (a_n) est croissante.

$$b_{n+1} - b_n = \sum_{k=n+1}^{2n+2} \frac{1}{k} - \sum_{k=n}^{2n} \frac{1}{k} = \frac{1}{2n+2} + \frac{1}{2n+1} - \frac{1}{n}$$

Or $\dfrac{1}{2n+2} \leqslant \dfrac{1}{2n}$ et $\dfrac{1}{2n+1} \leqslant \dfrac{1}{2n}$ donc

$$b_{n+1} - b_n \leqslant \frac{1}{2n} + \frac{1}{2n} - \frac{1}{n} = 0$$

La suite (b_n) est décroissante.
Enfin

$$b_n - a_n = \sum_{k=n}^{2n} \frac{1}{k} - \sum_{k=n+1}^{2n} \frac{1}{k} = \frac{1}{n} \to 0$$

Ainsi les deux suites (a_n) et (b_n) sont adjacentes.
Elles convergent donc vers une limite commune, on peut montrer que celle-ci vaut $\ln(2)$, mais c'est une autre histoire…

Exemple Soient

$$a_n = \frac{E(10^n x)}{10^n} \text{ et } b_n = \frac{E(10^n x) + 1}{10^n}$$

a_n et b_n sont des nombres décimaux car chacun rapport d'un entier par une puissance de 10
Puisque $E(10^n x) \leqslant 10^n x$, on a $10E(10^n x) \leqslant 10^{n+1} x$ et donc $10E(10^n x) \leqslant E(10^{n+1} x)$.
On en déduit $a_n \leqslant a_{n+1}$ et donc (a_n) croissante.
Puisque $10^n x < E(10^n x) + 1$, on a $10^{n+1} x < 10E(10^n x) + 10$ et donc
$E(10^{n+1} x) + 1 \leqslant 10(E(10^n x) + 1)$.
On en déduit $b_{n+1} \leqslant b_n$ et donc (b_n) décroissante.
Enfin

$$b_n - a_n = \frac{1}{10^n} \to 0$$

Les suites (a_n) et (b_n) sont adjacentes.
Notons ℓ leur limite commune.
Comme

$$E(10^n x) \leqslant 10^n x < E(10^n x) + 1$$

on a $a_n \leqslant x \leqslant b_n$ et donc à la limite on obtient $\ell \leqslant x \leqslant \ell$ d'où $\ell = x$.
La suite (a_n) est la suite des valeurs décimales par défaut de x et la suite (b_n) est la suite des valeurs décimales par excès de x.

Théorème

> Si $([a_n, b_n])$ est une suite décroissante de segments tels que $b_n - a_n \to 0$
> alors $\bigcap_{n \in \mathbb{N}} [a_n, b_n]$ est un singleton.

dém. :
Par décroissance de la suite de segments $([a_n, b_n])$, on a $[a_{n+1}, b_{n+1}] \subset [a_n, b_n]$.
On en déduit $a_{n+1} \geqslant a_n$ et $b_{n+1} \leqslant b_n$. Puisque de plus, on suppose $b_n - a_n \to 0$, on peut affirmer que les suites (a_n) et (b_n) sont adjacentes. Notons ℓ leur limite commune.
Puisque pour tout $n \in \mathbb{N}$, $a_n \leqslant \ell \leqslant b_n$, on a $\ell \in \bigcap_{n \in \mathbb{N}} [a_n, b_n]$.
Inversement, soit $x \in \bigcap_{n \in \mathbb{N}} [a_n, b_n]$.
On a pour tout $n \in \mathbb{N}$, $a_n \leqslant x \leqslant b_n$ donc à la limite $\ell \leqslant x \leqslant \ell$ et donc $x = \ell$.
Finalement

$$\bigcap_{n \in \mathbb{N}} [a_n, b_n] = \{\ell\}$$

□

Exemple Soit $f : [a, b] \to \mathbb{R}$ une fonction continue vérifiant $f(a) \leqslant 0$ et $f(b) \geqslant 0$.
Posons $[a_0, b_0] = [a, b]$. On vérifier $f(a_0) \leqslant 0$ et $f(b_0) \geqslant 0$.
Considérons ensuite le milieu
$$c = \frac{a_0 + b_0}{2}$$
Si $f(c) \geqslant 0$, on pose $[a_1, b_1] = [a_0, c]$ et sinon, on pose $[a_1, b_1] = [c, b_0]$ de sorte que $f(a_1) \leqslant 0$ et $f(b_1) \geqslant 0$.
Et on reprend le processus. . .
Par ce schéma, on construit une suite $([a_n, b_n])$ de segments emboités avec
$$b_n - a_n = \frac{b - a}{2^n} \to 0$$
L'intersection de ces segments détermine une valeur x limite commune aux suites (a_n) et (b_n).
Puisque par construction $f(a_n) \leqslant 0$ et $f(b_n) \geqslant 0$, en passant ces comparaisons à la limite, on obtient $f(x) \leqslant 0$ et $f(x) \geqslant 0$ car f est supposée continue.
On en déduit $f(x) = 0$.
Le procédé dichotomique permet d'accéder par encadrement aux valeurs approchées d'une solution de l'équation $f(x) = 0$.

15.2.7 Suites extraites

Définition

On appelle suite extraite (ou sous-suite) d'une suite $u = (u_n)$ toute suite $v = (v_n)$ telle qu'il existe $\varphi : \mathbb{N} \to \mathbb{N}$ strictement croissante (appelée extractrice) pour laquelle
$$\forall n \in \mathbb{N}, v_n = u_{\varphi(n)}$$
Une telle suite est alors notée $(u_{\varphi(n)})_{n\in\mathbb{N}}$ ou encore $(u_{\varphi(n)})$.

Remarque Une suite extraite se comprend comme une sélection d'une infinité de termes d'une suite.

Exemple $v_0 = u_2, v_1 = u_3, v_2 = u_{200}, v_3 = u_{203}$ sont les premiers termes d'une suite extraite de (u_n)

Exemple Pour $\varphi(n) = 2n$, la suite extraite (u_{2n}) est la suite formée des termes d'indice pair de (u_n).
Pour $\varphi(n) = 2n + 1$, la suite extraite (u_{2n+1}) est la suite formée des termes d'indice impair de (u_n).

Proposition

Si w est une suite extraite d'une suite v elle-même extraite de u alors w est une suite extraite de u.

dém. :
Par hypothèse, il existe $\varphi, \psi : \mathbb{N} \to \mathbb{N}$ strictement croissantes telles que
$$\forall n \in \mathbb{N}, v_n = u_{\varphi(n)} \text{ et } \forall n \in \mathbb{N}, w_n = v_{\psi(n)}$$

On a alors pour tout $n \in \mathbb{N}$, $w_n = v_{\psi(n)} \underset{m=\psi(n)}{=} v_m = u_{\varphi(m)} = u_{\varphi(\psi(m))} = u_{(\varphi\circ\psi)(m)}$.
L'application $\varphi \circ \psi : \mathbb{N} \to \mathbb{N}$ étant strictement croissante par composition, on peut dire que w est une suite extraite de u.
□

Proposition

> Toute suite extraite d'une suite de limite $\ell \in \bar{\mathbb{R}}$ tend aussi vers ℓ.

dém. :
Cas $\ell \in \mathbb{R}$:
Soit (v_n) une suite extraite d'une suite (u_n) convergeant vers ℓ.
Il existe $\varphi : \mathbb{N} \to \mathbb{N}$ strictement croissante telle que

$$\forall n \in \mathbb{N}, v_n = u_{\varphi(n)}$$

Par récurrence, on montre facilement

$$\forall n \in \mathbb{N}, \varphi(n) \geqslant n$$

Soit $\varepsilon > 0$.
Puisque $u_n \to \ell$, il existe $N \in \mathbb{N}$ tel que

$$\forall n \geqslant N, |u_n - \ell| \leqslant \varepsilon$$

Pour tout $n \geqslant N$, on a $\varphi(n) \geqslant n \geqslant N$ donc

$$\left|u_{\varphi(n)} - \ell\right| \leqslant \varepsilon \text{ i.e. } |v_n - \ell| \leqslant \varepsilon$$

Ainsi $v_n \to \ell$.
Cas $\ell = +\infty$ ou $-\infty$:
La démarche est semblable, seule la traduction quantifiée de la limite change.
□

Remarque Ce résultat permet de justifier qu'une suite n'a pas de limite.

Exemple Considérons la suite (u_n) de terme général $u_n = (-1)^n$.
On a $u_{2n} = 1 \to 1$ et $u_{2n+1} = -1 \to -1 \neq 1$ donc (u_n) n'a pas de limite.

Remarque Nous avons ici que la convergence de (u_{2n}) et (u_{2n+1}) ne suffit pas à justifier celle de (u_n).
Cependant :

Théorème

> Si (u_{2n}) et (u_{2n+1}) tendent vers une même limite $\ell \in \bar{\mathbb{R}}$ alors $u_n \to \ell$.

dém. :
Cas $\ell \in \mathbb{R}$:
Soit $\varepsilon > 0$.
Puisqu'on suppose $u_{2n} \to \ell$ et $u_{2n+1} \to \ell$, il existe $N_1, N_2 \in \mathbb{N}$ tels que

$$\forall k \geqslant N_1, |u_{2k} - \ell| \leqslant \varepsilon \text{ et } \forall k \geqslant N_2, |u_{2k+1} - \ell| \leqslant \varepsilon$$

Posons alors $N_0 = \max(2N_1, 2N_2 + 1)$.
Pour tout $n \geqslant N_0$.
Si n est pair alors $n = 2k$ avec $k \geqslant N_1$ et donc $|u_n - \ell| = |u_{2k} - \ell| \leqslant \varepsilon$.
Si n est impair alors $n = 2k + 1$ avec $k \geqslant N_1$ et donc $|u_n - \ell| \leqslant |u_{2k+1} - \ell| \leqslant \varepsilon$.
Dans les deux cas $|u_n - \ell| \leqslant \varepsilon$.
Ainsi $u_n \to \ell$.
Cas $\ell = +\infty$ ou $-\infty$.
La démarche est semblable, seule la traduction quantifiée de la limite change.
□

Exemple Etudions la nature de la suite (S_n) de terme général

$$S_n = \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k}$$

Nous allons montrer que les suites (S_{2n}) et (S_{2n+1}) sont adjacentes

$$S_{2(n+1)} - S_{2n} = S_{2n+2} - S_{2n} = \sum_{k=1}^{2n+2} \frac{(-1)^{k-1}}{k} - \sum_{k=1}^{2n} \frac{(-1)^{k-1}}{k}$$

En simplifiant les termes communs de ces deux sommes,

$$S_{2(n+1)} - S_{2n} = \frac{1}{2n+1} - \frac{1}{2n+2} \geqslant 0$$

Ainsi la suite extraite (S_{2n}) est croissante.

$$S_{2(n+1)+1} - S_{2n+1} = S_{2n+3} - S_{2n+1} = -\frac{1}{2n+2} + \frac{1}{2n+3} \leqslant 0$$

Ainsi la suite extraite (S_{2n+1}) est décroissante.
Enfin

$$S_{2n+1} - S_{2n} = \frac{1}{2n+1} \to 0$$

donc les suites extraites (S_{2n}) et (S_{2n+1}) sont adjacentes.
Par suite elles converge vers une limite commune et on en déduit que (S_n) converge vers cette même limite.
On peut montrer que cette limite commune vaut $\ln(2)$ mais c'est une autre histoire.

Remarque Le théorème qui suit est essentiel pour la construction de la suite du cours d'analyse (il est notamment à la base du calcul intégral. . .)

Théorème

De toute suite réelle bornée on peut extraire une suite convergente.

dém. :
La démonstration suivante exploite le principe de dichotomie qui consiste à découper un objet en deux portions, en conserver une et la découper à nouveau de sorte de générer un processus récurrent.
L'idée essentielle ici est la suivante : Lorsqu'on découpe un ensemble infini en deux sous-ensembles,

nécessairement l'un d'entre eux (au moins) doit être infini. Ceci permettra d'exploiter le principe de dichotomie...
Passons aux faits :
Soit (u_n) une suite réelle bornée par un certain réel $M \in \mathbb{R}^+$.
Nous allons construire par un procédé dichotomique deux suites adjacentes (a_n) et (b_n) telles que, pour tout entier naturel n, l'ensemble

$$A_n = \{k \in \mathbb{N} / a_n \leqslant u_k \leqslant b_n\}$$

□
soit infini :
Etape initiale :
Pour $a_0 = -M, b_0 = M$ l'ensemble A_0 est infini car égal à $\mathbb{N}$ puisque tous les termes de la suite u sont comprise entre $-M$ et M.
Etape n :
Soient a_n et b_n tels que l'ensemble

$$A_n = \{k \in \mathbb{N} / a_n \leqslant u_k \leqslant b_n\}$$

soit infini.
Construisons a_{n+1} et b_{n+1}.
Posons

$$d = \frac{a_n + b_n}{2}$$

et considérons :

$$A_n^- = \{k \in \mathbb{N} / a_n \leqslant u_k \leqslant d\} \text{ et } A_n^+ = \{k \in \mathbb{N} / d \leqslant u_k \leqslant b_n\}$$

On a $A_n = A_n^- \cup A_n^+$. Puisque A_n est infini, au moins l'un des deux ensembles A_n^- ou A_n^+ l'est.
Si A_n^+ est infini, on pose $a_{n+1} = d$ et $b_{n+1} = b_n$, sinon pose $a_{n+1} = a_n$ et $b_{n+1} = d$.
Dans les deux cas, l'ensemble

$$A_{n+1} = \{k \in \mathbb{N} / a_{n+1} \leqslant u_k \leqslant b_{n+1}\}$$

est infini.
De plus, dans les deux cas

$$b_{n+1} - a_{n+1} = \frac{b_n - a_n}{2}$$

Par se schéma, nous avons défini deux suites (a_n) et (b_n) et on vérife aisément que (a_n) est croissante, (b_n) décroissante et $b_{n+1} - a_{n+1} = (b_n - a_n)/2$ de sorte que

$$b_n - a_n = \frac{1}{2^n}(b - a) \to 0$$

Ainsi les suites (a_n) et (b_n) sont adjacentes. Notons ℓ leur limite commune.
Nous allons maintenant pouvoir construire une suite extraite de (u_n) qui converge vers ℓ en définissant par récurrence une extractrice $\varphi : \mathbb{N} \to \mathbb{N}$.
Pour $n = 0$, on pose $\varphi(0) = 0$.
Lorsque $\varphi(n)$ est défini, on pose

$$\varphi(n+1) = \min\left(A_{n+1} \backslash \{0, 1, 2, ..., \varphi(n)\}\right)$$

Ce min existe car l'ensemble $A_{n+1} \setminus \{0, 1, 2, ..., \varphi(n)\}$ est une partie de $\mathbb{N}$ et cette partie est non vide puisqu'on a retiré l'ensemble infini A_{n+1} un nombre fini d'éléments.
On définit ainsi une application $\varphi : \mathbb{N} \to \mathbb{N}$ vérifiant $\varphi(n+1) > \varphi(n)$ pour tout $n \in \mathbb{N}$; cette application est strictement croissante.
Considérons maintenant la suite extraite $(u_{\varphi(n)})$.
Par construction de φ, on a pour tout $n \in \mathbb{N}$, $\varphi(n) \in A_n$ et donc

$$a_n \leqslant u_{\varphi(n)} \leqslant b_n$$

Comme les suites (a_n) et (b_n) convergent vers la même limite ℓ, par encadrement $(u_{\varphi(n)})$ converge vers ℓ.
Finalement, nous avons extrait de la suite (u_n) une suite convergente.

15.3 Extension aux suites complexes

Les notions qui vont suivrent prolongent celles vues pour les suites réelles au cadre des suites complexes.

15.3.1 Définition

Définition

On appelle suite complexe toute application $z : \mathbb{N} \to \mathbb{C}$. Une telle suite est notée z, (z_n) ou $(z_n)_{n\in\mathbb{N}}$.
L'ensemble des suites complexes est noté $\mathbb{C}^{\mathbb{N}}$ ou $\mathcal{F}(\mathbb{N}, \mathbb{C})$.

Exemple $z_n = \dfrac{n+i}{n-i}$ est le terme général d'une suite complexe.

Définition

Soit $z = (z_n)$ une suite complexe.
On appelle partie réelle de z la suite Re(z) de terme général Re(z_n).
On appelle partie imaginaire de z la suite Im(z) de terme général Im(z_n).
On appelle conjuguée de suite z la suite $\bar{z}$ de terme général $\overline{z_n}$.
On appelle module de suite z la suite $|z|$ de terme général $|z_n|$.

Remarque Dans le cadre des suites complexes :
- on retrouve les notions de suites constantes, stationnaires et d'opérations sur les suites ;
- on perd les notions de monotonie, de suite majoraée et de suite minorée.

Cependant, on préserve la notion de suite bornée.

Définition

Une suite complexe $z = (z_n)$ est dite bornée s'il existe $M \in \mathbb{R}^+$ vérifiant

$$\forall n \in \mathbb{N}, |z_n| \leqslant M$$

Exemple $z_n = \mathrm{e}^{in}$ est le terme général d'une suite complexe bornée.
En effet, pour tout $n \in \mathbb{N}$, $|z_n| \leqslant 1$.

Proposition

Une suite complexe z est bornée si, et seulement si, $\text{Re}(z)$ et $\text{Im}(z)$ le sont.

dém. :
Puisque $|\text{Re}(z)|, |\text{Im}(z)| \leqslant |z|$, si la suite z est bornée, les suites $\text{Re}(z)$ et $\text{Im}(z)$ le sont aussi.
Inversement, puisque $|z| \leqslant |\text{Re}(z)| + |\text{Im}(z)|$, si les suites $\text{Re}(z)$ et $\text{Im}(z)$ sont bornées, la suite z l'est encore.
□

15.3.2 Convergence

Définition

On dit qu'une suite complexe z tend vers $\ell \in \mathbb{C}$ si

$$|z_n - \ell| \xrightarrow[n\to+\infty]{} 0$$

On note alors $z \to \ell$, $z_n \to \ell$ ou $z_n \xrightarrow[n\to+\infty]{} \ell$.

Exemple Si z est une suite stationnaire égale à C alors $z_n \to C$.

Exemple On a $z_n = \dfrac{in}{n+i} \to i$.
En effet

$$|z_n - i| = \left|\frac{in}{n+i} - i\right| = \frac{1}{|n+i|} = \frac{1}{\sqrt{n^2+1}} \to 0$$

Définition

On dit qu'une suite complexe z converge s'il existe $\ell \in \mathbb{C}$ vérifiant $z_n \to \ell$.
Sinon, on dit que z diverge.

Théorème

Si une suite complexe z converge alors il existe un unique $\ell \in \mathbb{C}$ tel que $z_n \to \ell$.
Ce complexe ℓ est alors appelé limite de la suite z et on note $\ell = \lim z$, $\ell = \lim\limits_{n\to+\infty} z_n$ ou $\ell = \lim z_n$.

dém. :
Si $z_n \to \ell$ et $z_n \to \ell'$ alors $0 \leqslant |\ell - \ell'| \leqslant |\ell - z_n| + |z_n - \ell'|$ donne à la limite $|\ell - \ell'| = 0$ et donc $\ell = \ell'$.
□

Remarque Pour le suites complexes, il n'y a pas de notion de limite infinie, tout au plus peut-on dire $|z_n| \to +\infty$.

15.3.3 Théorèmes liés à la convergence

Proposition

Si $z_n \to \ell$ alors $\bar{z}_n \to \bar{\ell}$ et $|z_n| \to |\ell|$.

dém. :
Supposons $z_n \to \ell$.
On a alors $|z_n - \ell| \to 0$.
Or

$$\left|\bar{z}_n - \bar{\ell}\right| = \left|\overline{z_n - \ell}\right| = |z_n - \ell|$$

et

$$||z_n| - |\ell|| \leqslant |z_n - \ell|$$

donc

$$\left|\bar{z}_n - \bar{\ell}\right| \to 0 \text{ et } ||z_n| - |\ell|| \to 0$$

□

Théorème

Toute suite complexe convergente est bornée.

dém. :
Soit (z_n) une suite convergente de limite ℓ.
Puisque $|z_n| \to |\ell|$, la suite $(|z_n|)$ est bornée, donc majorée, ce qui signifie que z est majorée.
□

Théorème

Soient z et z' deux suites complexes.
Si $z_n \to \ell \in \mathbb{C}$ et $z'_n \to \ell' \in \mathbb{C}$ alors $z_n + z'_n \to \ell + \ell'$, $z_n z'_n \to \ell\ell'$
De plus si $z_n \to \ell \neq 0$ alors $1/z_n \to 1/\ell$.

dém. :
Suppsons $z_n \to \ell$ et $z'_n \to \ell'$.
Puisque

$$|(z_n + z'_n) - (\ell + \ell')| \leqslant |z_n - \ell| + |z'_n - \ell'|$$

on a

$$|(z_n + z'_n) - (\ell + \ell')| \to 0$$

Ainsi $z_n + z'_n \to \ell + \ell'$.
Puisque

$$|z_n z'_n - \ell\ell'| \leqslant |z_n|\,|z'_n - \ell'| + |\ell'|\,|z_n - \ell|$$

avec $(|z_n|)$ majorée, on a

$$|z_n z'_n - \ell\ell'| \to 0$$

Ainsi $z_n z'_n \to \ell\ell'$.
Enfin, si $z_n \to \ell \neq 0$, l'égalité

$$\left|\frac{1}{z_n} - \frac{1}{\ell}\right| = \frac{|\ell - z_n|}{|z_n|\,|\ell|}$$

entraîne

$$\left|\frac{1}{z_n} - \frac{1}{\ell}\right| \to 0$$

et donc $\dfrac{1}{z_n} \to \dfrac{1}{\ell}$.

□

Théorème

Soient z une suite complexe et $\ell \in \mathbb{C}$.
On a équivalence entre :
(i) $z_n \to \ell$;
(ii) $\text{Re}(z_n) \to \text{Re}(\ell)$ et $\text{Im}(z_n) \to \text{Im}(\ell)$.

dém. :
(i) $\Rightarrow$ (ii) Car $|\text{Re}(z_n) - \text{Re}(\ell)| \leqslant |z_n - \ell|$ et $|\text{Im}(z_n) - \text{Im}(\ell)| \leqslant |z_n - \ell|$.
(ii) $\Rightarrow$ (i) Car $z_n = \text{Re}(z_n) + i\text{Im}(z_n)$.

□

Exemple On a

$$\lim_{n\to+\infty} \mathrm{e}^{2i\pi/n} = 1$$

En effet

$$\mathrm{e}^{2i\pi/n} = \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n} \to 1$$

15.3.4 Nature de (q^n) avec q complexe

Proposition

Soit $q \in \mathbb{C}$.
Si $|q| < 1$ alors $q^n \to 0$.
Si $q = 1$ alors $q^n \to 1$.
Si $|q| \geqslant 1$ et $q \neq 1$ alors (q^n) diverge.

dém. :
Si $|q| < 1$ alors pour tout $n \in \mathbb{N}$,

$$\left|q^{n+1}\right| = |q^n|\,|q| \leqslant |q^n|$$

Ainsi la suite réelle $(|q^n|)$ est décroissante, elle est de plus minorée par 0 et donc elle converge. Posons ℓ sa limite. Puisque $\left|q^{n+1}\right| = |q|\,|q^n|$, on obtient à la limite $\ell = |q|\,\ell$ d'où $\ell = 0$ car $|q| \neq 1$.
Ainsi $|q^n| \to 0$ et donc $q^n \to 0$.
Si $q = 1$ alors $q^n = 1 \to 1$.
Si $|q| \geqslant 1$ et $q \neq 1$, montrons la divergence de la suite (q^n) en raisonnant par l'absurde :
Supposons que la suite (q^n) converge vers un certain $\ell \in \mathbb{C}$.
La relation $q^{n+1} = q.q^n$ donne à la limite $\ell = q.\ell$ d'où $\ell = 0$ (sachant $q \neq 1$).
Or $|q^n| = |q|^n \geqslant 1$ et donc $q^n \not\to 0$.
C'est absurde et par conséquent la suite (q^n) diverge.

□

Exemple La suite $((-1/2)^n)$ converge vers 0.
Pour $\theta \neq 0 \quad [2\pi]$, la suite complexe $(\mathrm{e}^{in\theta})$ diverge.

15.3.5 Suites extraites

La notion de suite extraite et les propriétés énoncées dans le cadre réel se prolongent immédiatement et en les mêmes termes au cadre complexe y compris le théorème de Bolzano-Weierstrass :

Théorème

De toutes suites complexes bornées, on peut extraire une suite convergente.

dém. :
Soit (z_n) une suite complexe bornée.
Notons (x_n) et (y_n) les suites réelles égales à la partie réelle et à la partie imaginaire de (z_n).
Puisque la suite (z_n) est bornée, les deux suites (x_n) et (y_n) le sont aussi.
Par le théorème de Bolzano-Weierstrass relatif aux suites réelles, on peut extraire de la suite bornée (x_n) une suite convergente $(x_{\varphi(n)})$.
Considérons alors la suite extraite $(y_{\varphi(n)})$. Celle-ci est bornée car (y_n) l'est et donc par le théorème de Bolzano-Weierstrass relatif aux suites réelles, on peut en extraire une suite convergente $(y_{\varphi(\psi(n))})$.
Puisque la suite $(x_{\varphi(n)})$ converge, la suite extraite $(x_{\varphi(\psi(n))})$ converge aussi et finalement la suite extraite $(z_{\varphi(\psi(n))})$ converge car ses parties réelles et imaginaires convergent.
□

15.3.6 Suites d'éléments de $\mathbb{R}^2$

15.3.6.1 Distance euclidienne

Définition

Soient $a = (x, y) \in \mathbb{R}^2$ et $b = (x', y') \in \mathbb{R}^2$.
On appelle distance euclidienne de a à b le réel

$$d(a, b) = \sqrt{(x' - x)^2 + (y' - y)^2}$$

Remarque Si l'on pose $z_a = x + i.y$ et $z_b = x' + i.y'$ alors $d(a, b) = |z_b - z_a|$.

Proposition

$\forall a, b, c \in \mathbb{R}^2$,

$$d(a, b) = 0 \Leftrightarrow a = b,\ d(a, b) = d(b, a) \text{ et } d(a, c) \leqslant d(a, b) + d(b, c)$$

dém. :
En exploitant $d(a, b) = |z_b - z_a|$, ces propriétés découlent immédiatement de propriétés connues du module.
□

15.3.6.2 Suite

Définition

On appelle suite d'éléments de $\mathbb{R}^2$ toute application $u : \mathbb{N} \to \mathbb{R}^2$.

Exemple $u_n = (2 + 3n, 1 + \mathrm{e}^n)$ est le terme général d'une suite d'éléments de $\mathbb{R}^2$.

Définition

Soit $u = (u_n)$ une suite d'éléments de $\mathbb{R}^2$.
Pour tout $n \in \mathbb{N}$, on peut écrire $u_n = (x_n, y_n)$.
Ceci introduit deux suites réelles (x_n) et (y_n) appelées suites coordonnées de (u_n).

15.3.6.3 Convergence

Soit $u = (u_n)$ une suite d'éléments de $\mathbb{R}^2$.

Définition

On dit que u tend vers $\ell \in \mathbb{R}^2$ si
$$d(u_n, \ell) \to 0$$
On note alors $u \to \ell$, $u_n \to \ell$ ou $u_n \xrightarrow[n\to+\infty]{} \ell$.

Définition

On dit que la suite u converge s'il existe $\ell \in \mathbb{R}^2$ tel que $u_n \to \ell$.

Proposition

Si u converge alors il existe un unique $\ell \in \mathbb{R}^2$ tel que $u_n \to \ell$.
ℓ est alors appelé limite de la suite u et on note $\ell = \lim u$, $\ell = \lim u_n$ ou $\ell = \lim\limits_{n\to+\infty} u_n$

dém. :
Si $u_n \to \ell$ et $u_n \to \ell'$ alors $0 \leqslant d(\ell, \ell') \leqslant d(\ell, u_n) + d(u_n, \ell')$ donne à la limite $d(\ell, \ell') = 0$ et donc $\ell = \ell'$.
□

Proposition

Soit u une suite d'éléments de $\mathbb{R}^2$ de suites coordonnées x et y.
On a équivalence entre :
(i) u converge vers $\ell = (a, b)$;
(ii) x converge vers a et y converge vers b.

dém. :
On a
$$d(u_n, \ell) = \sqrt{(x_n - a)^2 + (y_n - b)^2}$$
Par suite $|x_n - a|, |y_n - b| \leqslant d(u_n, \ell)$ et donc (i) $\Rightarrow$ (ii).
Aussi $d(u_n, \ell) \leqslant |x_n - a| + |y_n - b|$ et donc (ii) $\Rightarrow$ (i).
□

Exemple Soit
$$u_n = (\frac{1}{n}, n \sin\frac{1}{n}) \in \mathbb{R}^2$$
On a $u_n \to (0, 1)$ car $\dfrac{1}{n} \to 0$ et $n \sin\dfrac{1}{n} = \dfrac{\sin(1/n)}{1/n} \to 1$.

15.4 Comparaisons de suites numériques

Les suites considérées ici sont réelles ou complexes et définies à partir d'un certain rang.

15.4.1 Négligeabilité

Définition

On dit qu'une suite u est négligeable devant une suite α si pour n assez grand on peut écrire

$$u_n = \alpha_n \varepsilon_n$$

avec (ε_n) une suite de limite nulle.
On note alors $u_n = \underset{n\to+\infty}{o}(\alpha_n)$, $u_n = o(\alpha_n)$ ou $u_n << \alpha_n$ quand $n \to +\infty$.

Proposition

Si la suite α ne s'annule pas à partir d'un certain rang alors on a équivalence entre :
(i) $u_n = o(\alpha_n)$;
(ii) $u_n/\alpha_n \to 0$.

dém. :
(i) $\Rightarrow$ (ii) Si $u_n = o(\alpha_n)$ alors on peut écrire pour n assez grand $u_n = \alpha_n\varepsilon_n$ avec $\varepsilon_n \to 0$.
Or pour n assez grand $\alpha_n \neq 0$ et donc pour n assez grand, $u_n/\alpha_n = \varepsilon_n \to 0$.
(ii) $\Rightarrow$ (i) Si $u_n/\alpha_n \to 0$ alors on peut écrire $u_n = \alpha_n\varepsilon_n$ avec $\varepsilon_n = u_n/\alpha_n \to 0$ et donc $u_n = o(\alpha_n)$.
□

Exemple Quand $n \to +\infty$, $\ln n = o(n)$.
En effet $\dfrac{\ln n}{n} \to 0$ par limite de référence.
De même, quand $n \to +\infty$, on peut écrire $n = o(n^2)$, $1/n^2 = o(1/n)$,...

Exemple Ecrire $u_n = o(1)$ signifie que $u_n/1 \to 0$ i.e. $u_n \to 0$.
Il est fréquent d'écrire $u_n = o(1)$ pour signifier que la suite (u_n) est de limite nulle.
Si par exemple $u_n \to \ell \in \mathbb{R}$, on peut aussi écrire $u_n = \ell + o(1)$.

Attention : $u_n = o(\alpha_n)$ et $v_n = o(\alpha_n)$ n'impliquent pas $u_n = v_n$.
La relation $u_n = o(\alpha_n)$ se comprend :
« la suite u appartient à l'ensemble des suites négligeables devant α ».
En fait, lors de la manipulation de $o(\ldots)$, les égalités ne doivent plus être considérées comme possédant la propriété de symétrie : $a = b \Rightarrow b = a$!

Proposition

Si $u_n = o(\lambda\alpha_n)$ alors $u_n = o(\alpha_n)$.
Si $u_n = o(\alpha_n)$ et $v_n = o(\alpha_n)$ alors $u_n + v_n = o(\alpha_n)$.
Si $u_n = o(\alpha_n)$ alors $u_n v_n = o(\alpha_n v_n)$.

dém. :
Supposons $u_n = o(\lambda\alpha_n)$
A partir d'un certain rang, on peut écrire $u_n = \lambda\alpha_n\varepsilon_n$ avec $\varepsilon_n \to 0$ et donc $u_n = \alpha_n(\lambda\varepsilon_n)$ avec $\lambda\varepsilon_n \to 0$. Ainsi $u_n = o(\alpha_n)$.
Supposons $u_n = o(\alpha_n)$ et $v_n = o(\alpha_n)$.
A partir d'un certain rang, on peut écrire $u_n = \alpha_n\varepsilon_n$ et $v_n = \alpha_n\varepsilon'_n$ avec $\varepsilon_n, \varepsilon'_n \to 0$ et donc $u_n + v_n = \alpha_n(\varepsilon_n + \varepsilon'_n)$ avec $\varepsilon_n + \varepsilon'_n \to 0$. Ainsi $u_n + v_n = o(\alpha_n)$.

Supposons $u_n = o(\alpha_n)$.
A partir d'un certain rang, on peut écrire $u_n = \alpha_n \varepsilon_n$ avec $\varepsilon_n \to 0$ et donc $u_n v_n = \alpha_n v_n \varepsilon_n$ avec $\varepsilon_n \to 0$.
Ainsi $u_n v_n = o(\alpha_n v_n)$
□

Remarque Dans la pratique, on exploite ces formules en écrivant les égalités :
- $o(\lambda\alpha_n) = o(\alpha_n)$;
- $o(\alpha_n) + o(\alpha_n) = o(\alpha_n)$;
- $o(\alpha_n)v_n = o(\alpha_n v_n)$.

Attention : $o(\alpha_n) - o(\alpha_n) = o(\alpha_n)$ et non $o(\alpha_n) - o(\alpha_n) = 0$
On ne peut jamais simplifier les $o(\ldots)$.

Remarque Systématiquement, on écrit $o(\alpha_n)$ au lieu de $o(2\alpha_n)$.
Ainsi, on écrit $o(1)$ au lieu de $o(2)$.

Proposition

Si $u_n = o(\alpha_n)$ et $\alpha_n = o(\beta_n)$ alors $u_n = o(\alpha_n)$.

dém. :
Supposons $u_n = o(\alpha_n)$ et $\alpha_n = o(\beta_n)$.
Pour n assez grand, on peut écrire à la fois $u_n = \alpha_n\varepsilon_n$ et $\alpha_n = \beta_n\varepsilon'_n$ avec $\varepsilon_n, \varepsilon'_n \to 0$.
On a alors $u_n = \beta_n\varepsilon_n\varepsilon'_n$ avec $\varepsilon_n\varepsilon'_n \to 0$ et donc $u_n = o(\beta_n)$.
□

Remarque Dans la pratique on écrit $o(\alpha_n) = o(\beta_n)$.
Ici, il convient de souligner que l'égalité n'est pas symétrique mais se comprend comme une transformation du premier membre en le second.

Exemple Quand $n \to +\infty$, on peut écrire $o(n) - no(n) = o(n) - o(n^2) = o(n^2) - o(n^2) = o(n^2)$.

15.4.2 Prépondérance

Définition

On dit qu'une suite u est dominée par une suite α si pour n assez grand on peut écrire

$$u_n = \alpha_n M_n$$

avec (M_n) une suite bornée.
On note alors $u_n = \underset{n\to+\infty}{O}(\alpha_n)$ ou $u_n = O(\alpha_n)$ quand $n \to +\infty$.

Exemple $u_n = o(\alpha_n) \Rightarrow u_n = O(\alpha_n)$.
En effet, si une suite converge vers 0 elle est bornée

Exemple $u_n = O(1)$ signifie que la suite (u_n) est bornée.
Ainsi, on peut écrire $\sin n = O(1)$, $(-1)^n = O(1)$.

Attention : $u_n = O(\alpha_n)$ et $v_n = O(\alpha_n) \not\Rightarrow u_n = v_n$.
La relation $u_n = O(\alpha_n)$ se comprend :
« la suite (u_n) appartient à l'ensemble des suites dominées par (α_n) ».
Encore une fois l'égalité manipulée n'est pas symétrique.

Proposition

Si la suite (α_n) ne s'annule pas à partir d'un certain rang alors on a équivalence entre :
(i) $u_n = O(\alpha_n)$;
(ii) (u_n/α_n) est bornée.

dém. :
(i) $\Rightarrow$ (ii) Supposons $u_n = O(\alpha_n)$.
A partir d'un certain rang, on peut écrire $u_n = \alpha_n M_n$ avec (M_n) bornée et alors puisque pour n assez grand $u_n/\alpha_n = M_n$, la suite (u_n/α_n) est bornée à partir d'un certain rang donc bornée.
(ii) $\Rightarrow$ (i) Supposons (u_n/α_n) bornée.
Puisque pour n assez grand $u_n = \alpha_n M_n$ avec $M_n = u_n/\alpha_n$, on a $u_n = O(\alpha_n)$.
□

Proposition

Si $u_n = O(\lambda\alpha_n)$ alors $u_n = O(\alpha_n)$.
Si $u_n = O(\alpha_n)$ et $v_n = O(\alpha_n)$ alors $u_n + v_n = O(\alpha_n)$.
Si $u_n = O(\alpha_n)$ alors $u_n\beta_n = O(\alpha_n\beta_n)$.

dém. :
Supposons $u_n = O(\lambda\alpha_n)$.
A partir d'un certain rang, on peut écrire $u_n = \lambda\alpha_n M_n$ avec (M_n) bornée et donc $u_n = \alpha_n(\lambda M_n)$ avec (λM_n) bornée. Ainsi $u_n = O(\alpha_n)$.
Supposons $u_n = O(\alpha_n)$ et $v_n = O(\alpha_n)$.
A partir d'un certain rang, on peut écrire $u_n = \alpha_n M_n$ et $v_n = \alpha_n M_n'$ avec (M_n) et (M_n') bornées. On a donc $u_n + v_n = \alpha_n(M_n + M_n')$ avec $(M_n + M_n')$ bornée. Ainsi $u_n + v_n = O(\alpha_n)$.
Supposons $u_n = O(\alpha_n)$.
A partir d'un certain rang, on peut écrire $u_n = \alpha_n M_n$ avec (M_n) bornée et donc $u_n v_n = \alpha_n v_n M_n$ avec (M_n) bornée. Ainsi $u_n v_n = O(\alpha_n v_n)$
□

Remarque Dans la pratique, on exploite ces formules en écrivant :
- $O(\lambda\alpha_n) = O(\alpha_n)$;
- $O(\alpha_n) + O(\alpha_n) = O(\alpha_n)$;
- $O(\alpha_n)\beta_n = O(\alpha_n\beta_n)$.

Proposition

Si $u_n = o(\alpha_n)$ et $\alpha_n = O(\beta_n)$ alors $u_n = o(\beta_n)$.
Si $u_n = O(\alpha_n)$ et $\alpha_n = o(\beta_n)$ alors $u_n = o(\beta_n)$.
Si $u_n = O(\alpha_n)$ et $\alpha_n = O(\beta_n)$ alors $u_n = O(\beta_n)$.

dém. :
Supposons $u_n = o(\alpha_n)$ et $\alpha_n = O(\beta_n)$.
A partir d'un certain on peut écrire $u_n = \alpha_n \varepsilon_n$ et $\alpha_n = \beta_n M_n$ avec $\varepsilon_n \to 0$ et (M_n) bornée.
On a alors $u_n = \beta_n(M_n \varepsilon_n)$ avec $M_n \varepsilon_n \to 0$ et donc $u_n = o(\beta_n)$.
Les deux autres implications s'obtiennent de façon semblable.
□

15.4.3 Croissance comparée des suites de référence

Proposition

Pour $\alpha < \beta$, $n^\alpha = o(n^\beta)$ et $(\ln n)^\alpha = o((\ln n)^\beta)$.
Pour $0 < a < b$ on a $a^n = o(b^n)$.

dém. :
$\dfrac{n^\alpha}{n^\beta} = n^{\alpha-\beta} = \exp((\alpha - \beta)\ln n) \to 0$ car $\alpha - \beta < 0$.
$\dfrac{(\ln n)^\alpha}{(\ln n)^\beta} = \exp\left((\alpha - \beta)\ln(\ln n)\right) \to 0$ car $\alpha - \beta < 0$.
$\dfrac{a^n}{b^n} = \exp(n(\ln a - \ln b)) \to 0$ car $\ln a - \ln b < 0$.
□

Proposition

Pour $\alpha, \beta > 0$ et $a > 1$ on a
$(\ln n)^\alpha = o(n^\beta)$, $n^\alpha = o(a^n)$ et $a^n = o(n!)$.

dém. :
Jusitifions $(\ln n)^\alpha = o(n^\beta)$.

$$\frac{(\ln n)^\alpha}{n^\beta} = \exp\left(\alpha \ln(\ln n) - \beta \ln n\right)$$

or

$$\alpha \ln(\ln n) - \beta \ln n = \ln n \left(-\beta + \alpha \frac{\ln(\ln n)}{\ln n}\right)$$

On sait que $\lim\limits_{x\to+\infty} \dfrac{\ln x}{x} = 0$ donc par composition de limites $\dfrac{\ln(\ln n)}{\ln n} \to 0$ puis par opérations $\alpha \ln(\ln n) - \beta \ln n \to -\infty$ (car $\beta > 0$) et enfin

$$\frac{(\ln n)^\alpha}{n^\beta} \to 0$$

Jusitifions $n^\alpha = o(a^n)$.

$$\frac{n^\alpha}{a^n} = \exp\left(\alpha \ln n - n \ln a\right)$$

or

$$\alpha \ln n - n \ln a = n\left(-\ln a + \alpha \frac{\ln n}{n}\right)$$

Puisque $\dfrac{\ln n}{n} \to 0$, on obtient par opérations $\alpha \ln n - n \ln a \to -\infty$ (car $\ln a > 0$) puis

$$\frac{n^\alpha}{a^n} \to 0$$

Justifions $a^n = o(n!)$.
Posons $u_n = \dfrac{a^n}{n!}$. On a $u_{n+1} = \dfrac{a}{n+1} u_n$ et $u_n \geqslant 0$

Pour n assez grand $n+1 \geqslant a$ et donc $u_{n+1} \leqslant u_n$.
La suite (u_n) est décroissante et minorée donc convergente. Posons ℓ sa limite.
En passant la relation $u_{n+1} = \dfrac{a}{n+1} u_n$ à la limite, on obtient $\ell = 0 \times \ell$ et donc $\ell = 0$.
□

Lemme

Si u et v ne s'annulent pas à partir d'un certain rang alors

$$u_n = o(v_n) \Rightarrow \frac{1}{v_n} = o\left(\frac{1}{u_n}\right)$$

dém. :

$$\frac{u_n}{v_n} \to 0 \Rightarrow \frac{1/v_n}{1/u_n} \to 0$$

puisque

$$\frac{1/v_n}{1/u_n} = \frac{u_n}{v_n}$$

□

Proposition

Pour $\alpha, \beta > 0$ et $0 < a < 1$ on a :

$$\frac{1}{n!} = o(a^n), q^n = o\left(\frac{1}{n^\alpha}\right) \text{ et } \frac{1}{n^\alpha} = o\left(\frac{1}{(\ln n)^\beta}\right)$$

dém. :
Ces résultats découlent immédiatement du lemme précédent.
□

Remarque Par l'étude précédente, on peut écrire :

$$\mathrm{e}^{-n} << \frac{1}{n^2} << \frac{1}{n} << \frac{1}{\ln n} << 1 << \ln n << n << n^2 << \mathrm{e}^n << n!$$

A titre d'exercice, on peut positionner entre les précédents les termes suivants :

$$2^n, \frac{1}{3^n}, \frac{1}{(\ln n)^2}, \frac{1}{n \ln n}, \frac{\ln n}{n^2}, \frac{n}{\ln n}, \ln \ln n, \sqrt{n} \ln n, \frac{1}{\sqrt{n}} \ln n, ...$$

15.4.4 Equivalence

15.4.4.1 Définition

Définition

Une suite u est dite équivalente à une suite v si à partir d'un certain rang on peut écrire

$$u_n = v_n \theta_n$$

avec $\theta_n \to 1$
On note alors $u_n \underset{n \to +\infty}{\sim} v_n$ ou encore $u_n \sim v_n$ quand $n \to +\infty$.

Proposition

Si la suite (v_n) ne s'annule pas à partir d'un certain rang alors on a équivalence entre :
(i) $u_n \sim v_n$;
(ii) $u_n/v_n \to 1$.

dém. :
(i) $\Rightarrow$ (ii) Supposons $u_n \sim v_n$.
A partir d'un certain rang, on peut écrire $u_n = v_n\theta_n$ avec $\theta_n \to 1$ bornée et alors puisque pour n assez grand $u_n/v_n = \theta_n$, la suite (u_n/v_n) converge vers 1.
(ii) $\Rightarrow$ (i) Supposons $u_n/v_n \to 1$ bornée.
Puisque pour n assez grand $u_n = v_n\theta_n$ avec $\theta_n = u_n/v_n \to 1$, on a $u_n \sim v_n$.
□

Exemple Quand $n \to +\infty$,

$$n + \ln n + 2 \sim n$$
$$\sqrt{n^2 + n + 1} \sim n,$$
$$\frac{1}{n} + \frac{1}{n^2} \sim \frac{1}{n} \text{ et } \frac{n+1}{n} \sim 1$$

Attention : $u_n \sim 0$ signifie $u_n = 0$ à partir d'un certain rang (c'est assez rare d'être dans cette situation. . .)
Ecrire $u_n \sim +\infty$ n'a aucun sens !

Proposition

Si $u_n \sim v_n$ alors $v_n \sim u_n$.
Si $u_n \sim v_n$ et $v_n \sim w_n$ alors $u_n \sim w_n$.

dém. :
Supposons $u_n \sim v_n$.
On peut écrire à partir d'un certain rang $u_n = v_n\theta_n$ avec $\theta_n \to 1$.
Puisque $\theta_n \to 1$, pour n assez grand, on a $\theta_n \neq 0$.
Ainsi à partir d'un certain rang, on a $v_n = u_n\theta'_n$ avec $\theta'_n = 1/\theta_n \to 1$ et donc $v_n \sim u_n$.
Supposons $u_n \sim v_n$ et $v_n \sim w_n$.
A partir d'un certain rang on peut écrire $u_n = v_n\theta_n$ et $v_n = \theta'_n w_n$ avec $\theta_n, \theta'_n \to 1$ et donc $u_n = w_n\theta_n\theta'_n$ avec $\theta_n\theta'_n \to 1$. Ainsi $u_n \sim w_n$.
□

Proposition

On a équivalence entre :
(i) $u_n \sim v_n$;
(ii) $u_n = v_n + o(v_n)$.

dém. :
(i) $\Rightarrow$ (ii) Supposons $u_n \sim v_n$.
Pour n assez grand, on peut écrire $u_n = v_n\theta_n$ avec $\theta_n \to 1$.
On a alors $u_n = v_n + v_n\varepsilon_n$ avec $\varepsilon_n = \theta_n - 1 \to 0$ et donc $u_n = v_n + o(v_n)$.
(ii) $\Rightarrow$ (i) Supposons $u_n = v_n + o(v_n)$.
Pour n assez grand, on peut écrire $u_n = v_n + v_n\varepsilon_n$ avec $\varepsilon_n \to 0$.
On a alors $u_n = v_n\theta_n$ avec $\theta_n = 1 + \varepsilon_n \to 1$ et donc $u_n \sim v_n$.
□

Remarque Ce résultat est essentiel pour obtenit facilement un équivalent à une somme.

Exemple Quand $n \to +\infty$,

$$2^n + n! + n^{10} = n! + o(n!) \sim n!$$

$$\ln n + 2\sqrt{n} + 1 = 2\sqrt{n} + o(\sqrt{n}) \sim 2\sqrt{n}$$

$$\frac{1}{2^n} + \frac{1}{\sqrt{n}} + \frac{1}{\ln n} = \frac{1}{\ln n} + o\left(\frac{1}{\ln n}\right) \sim \frac{1}{\ln n}$$

Exemple Si $u_n \to \ell$ avec $\ell \neq 0$ alors $u_n = \ell + o(1) = \ell + o(\ell) \sim \ell$.
Cette manipulation ne peut pas se faire dans les cas $\ell = 0$ et $\ell = +\infty$.

15.4.4.2 Applications des équivalents

Théorème

Si $u_n \sim v_n$ et $v_n \to \ell \in \bar{\mathbb{R}}$ ou $\ell \in \mathbb{C}$ alors $u_n \to \ell$.

dém. :
Pour n assez grand $u_n = v_n \theta_n$ avec $\theta_n \to 1$ donc $u_n \to \ell$.
□
Exemple Quand $n \to +\infty$, $n^2 - n - \sqrt{n} = n^2 + o(n^2) \sim n^2 \to +\infty$.

Remarque Cette formalisation permet de lever efficacement des indéterminations.

Remarque Inversement :
Supposons $u_n \to \ell$ et $v_n \to \ell$.
Si $\ell \in \mathbb{R}^\star$ alors $u_n \sim v_n$ car $u_n \sim \ell$ et $v_n \sim \ell$.
Sinon, on ne peut rien dire !

Attention : $u_n \to 0$ et $v_n \to 0$ n'implique pas $u_n \sim v_n$.
Par exemple $u_n = 1/n$ et $v_n = 1/n^2$ définissent des suites de limite nulle non équivalente.

Attention : $u_n \to +\infty$ et $v_n \to +\infty$ n'implique pas $u_n \sim v_n$.
Par exemple $u_n = n$ et $v_n = n^2$ définissent des suites de limite $+\infty$ non équivalentes.

Théorème

Soient u et v deux suites réelles.
Si $u_n \sim v_n$ alors, à partir d'un certain rang, u_n et v_n ont le même signe.

dém. :
Pour n assez grand $u_n = v_n\theta_n$ avec $\theta_n \to 1$.
Puisque $\theta_n \to 1$, pour n assez grand $\theta_n > 0$ et alors u_n et v_n ont le même signe.
□

Exemple Quand $n \to +\infty$,

$$u_n = -\frac{1}{n} + \frac{1}{\sqrt{n}} - \frac{2}{n^2} \sim \frac{1}{\sqrt{n}} \to 0^+$$

donc $u_n \to 0^+$.

Proposition

Si $u_n \sim v_n$ alors $u_n = O(v_n)$ et $v_n = O(u_n)$.

dém. :
Supposons $u_n \sim v_n$.
Pour n assez grand, on peut écrire $u_n = v_n\theta_n$ avec $\theta_n \to 1$.
Puisque la suite (θ_n) est convergente, elle est bornée et donc $u_n = O(v_n)$.
Aussi $u_n \sim v_n$ entraîne $v_n \sim u_n$ et donc $v_n = O(u_n)$.
□

Attention : Il n'y a pas de réciproque.
On a par exemple $n = O(2n)$ et $2n = O(n)$ alors que n et $2n$ ne sont pas équivalents.

Proposition

Si $u_n \sim v_n$ et $v_n = o(w_n)$ alors $u_n = o(w_n)$.
Si $u_n \sim v_n$ et $v_n = O(w_n)$ alors $u_n = O(w_n)$.
Si $u_n = o(v_n)$ et $v_n \sim w_n$ alors $u_n = o(w_n)$.
Si $u_n = O(v_n)$ et $v_n \sim w_n$ alors $u_n = O(w_n)$.

dém. :
Supposons $u_n \sim v_n$ et $v_n = o(w_n)$.
On a alors $u_n = O(v_n)$ et $v_n = o(w_n)$ et donc $u_n = o(w_n)$.
Les trois autres implications s'obtiennent de façon analogue.
□

Remarque Par les deux dernières propriétés, il est usuel lors d'expression de négligeabilité ou de domination de ne comparer qu'à des suites « simples ».
Par exemple, on écrit $o(n)$ au lieu de $o(n+1)$ car $n \sim n+1$.

15.4.4.3 Obtention d'équivalents

Attention : On ne peut pas sommer avec les équivalents :
$u_n \sim v_n$ n'entraîne pas $u_n + w_n \sim v_n + w_n$.
Par exmple $n+1 \sim n-1$ alors que $(n+1)-n$ n'est pas équivalent à $(n-1)-n$.

Remarque Pour déterminer un équivalent à une somme, on écrit celle-ci sous la forme
$u_n + o(u_n) \sim u_n$.

Théorème

> Si $u_n \sim v_n$ et $w_n \sim t_n$ alors $u_n w_n \sim v_n t_n$ et $\dfrac{u_n}{w_n} \sim \dfrac{v_n}{t_n}$.
> Si $u_n \sim v_n$ alors $\forall p \in \mathbb{Z}, u_n^p \sim v_n^p$.

dém. :
Suppsons $u_n \sim v_n$ et $w_n \sim t_n$.
A partir d'un certain rang, on peut écrire $u_n = v_n\theta_n$ et $w_n = t_n\theta'_n$ avec $\theta_n, \theta'_n \to 1$.
On a alors alors $u_n w_n = v_n t_n(\theta_n\theta'_n)$ et $\dfrac{u_n}{w_n} = \dfrac{v_n}{t_n}\dfrac{\theta_n}{\theta'_n}$ avec $\theta_n\theta'_n \to 1$ et $\theta_n/\theta'_n \to 1$.
De plus $u_n^p = v_n^p\theta_n^p$ avec $\theta_n^p \to 1$.
□

Exemple Déterminons un équivalent simple de

$$u_n = (1+\frac{1}{n})(n+\ln n)^2$$

D'une part, $1+\dfrac{1}{n} \sim 1$ car $1+\dfrac{1}{n} \to 1$.
D'autre part $n+\ln n = n + o(n) \sim n$.
Par opérations multiplicatives

$$u_n \sim n^2$$

Exemple Déterminons un équivalent simple de

$$u_n = \frac{\ln n + 1}{n+\sqrt{n}}$$

D'une part $\ln n + 1 \sim \ln n$ car $\ln n \to +\infty$ et 1 est une constante.
D'autre part $n+\sqrt{n} \sim n$ car $n+\sqrt{n} = n+o(n)$.
Par opérations multiplicatives

$$u_n \sim \frac{\ln n}{n}$$

Théorème

> Soient (u_n) et (v_n) deux suites de réels strictement positifs.
> Si $u_n \sim v_n$ alors
> $$\forall \alpha \in \mathbb{R}, u_n^\alpha \sim v_n^\alpha$$

dém. :
Pour n assez grand, on peut écrire $u_n = v_n\theta_n$ avec $\theta_n \to 1$.
Par suite $u_n^\alpha = v_n^\alpha\theta_n^\alpha$ avec $\theta_n^\alpha \to 1^\alpha = 1$.
□

Exemple Déterminons un équivalent simple de

$$u_n = \frac{\sqrt{n^2+1}}{\sqrt[3]{n^3+1}}$$

$n^2+1 \sim n^2$ donc $\sqrt{n^2+1} \sim \sqrt{n^2} = |n| = n$ et $\sqrt[3]{n^3+1} \sim n$ donc

$$u_n \sim \frac{n}{n} = 1$$

Attention : Ce résultat ne vaut que pour un exposant α fixe.
En particulier $u_n \sim v_n$ n'entraîne pas $u_n^n \sim v_n^n$.

Proposition

Si $u_n \to 0$ alors $\sin u_n \sim u_n$ et $\ln(1+u_n) \sim u_n$.

dém. :
On exploite les limites connues $\frac{\sin x}{x} \xrightarrow[x\to 0]{} 1$ et $\frac{\ln(1+x)}{x} \xrightarrow[x\to 0]{} 1$
□

Exemple Déterminons un équivalent simple de

$$u_n = \ln(n+1) - \ln(n)$$

$\ln(n+1) - \ln n = \ln(1+1/n) \sim 1/n$ car $1/n \to 0$.

Exemple Déterminons un équivalent simple de

$$u_n = (n+1)\sin\frac{1}{\sqrt{n^2+1}}$$

Puisque $\frac{1}{\sqrt{n^2+1}} \to 0$, on a $\sin\frac{1}{\sqrt{n^2+1}} \sim \frac{1}{\sqrt{n^2+1}}$ puis

$$u_n \sim n \times \frac{1}{n} = 1$$

Exemple Déterminons un équivalent de

$$u_n = \tan\frac{1}{n} - \frac{1}{n^2}$$

On a

$$\tan\frac{1}{n} = \frac{\sin(1/n)}{\cos(1/n)} \sim \frac{1}{n}$$

donc

$$\tan\frac{1}{n} = \frac{1}{n} + o\left(\frac{1}{n}\right)$$

puis

$$u_n = \frac{1}{n} + o\left(\frac{1}{n}\right) - \frac{1}{n^2} = \frac{1}{n} + o\left(\frac{1}{n}\right) \sim \frac{1}{n}$$

Proposition

Soient u et v deux suites de réels strictement positifs.
Si $u_n \sim v_n$ et $v_n \to \ell$ avec $\ell = 0^+$, $\ell = +\infty$ ou $\ell \in \mathbb{R}^{+\star} \setminus \{1\}$
alors $\ln u_n \sim \ln v_n$.

dém. :
Pour n assez grand, on peut écrire $u_n = v_n\theta_n$ avec $\theta_n \to 1$.
On a alors $\ln u_n = \ln v_n + \ln\theta_n$. Or $\ln\theta_n \to 0$ et $\ln v_n \to -\infty, \ln\ell$ ou $+\infty$.
Dans chaque cas, $\ln\theta_n = o(\ln v_n)$ et donc $\ln u_n \sim \ln v_n$.
□

Exemple Déterminons un équivalent simple de

$$u_n = \ln(n^2+1)$$

Puisque $n^2 + 1 \sim n^2 \to +\infty \neq 1$, on a

$$u_n = \ln(n^2+1) \sim \ln n^2 = 2\ln n$$

Exemple Déterminons un équivalent simple de $u_n = \ln(\sin(1/n))$.
Puisque $\sin(1/n) \sim 1/n \to 0 \neq 1$, $u_n \sim \ln(1/n) = -\ln n$.

Attention : $u_n \sim v_n$ et $v_n \to 1$ n'implique par $\ln u_n \sim \ln v_n$.
Par exemple $1 + \dfrac{1}{n} \sim 1 - \dfrac{1}{n}$ alors que

$$\ln\left(1+\frac{1}{n}\right) \sim \frac{1}{n} \text{ et } \ln\left(1-\frac{1}{n}\right) \sim -\frac{1}{n}$$

Remarque Si $u_n \to 1$, pour déterminer un équivalent simple de $\ln u_n$, il suffit d'écrire $u_n = 1 + v_n$ puis d'exploiter $\ln(1+v_n) \sim v_n$ car $v_n \to 0$.

Exemple Déterminons un équivalent simple de

$$u_n = \ln \frac{n+1}{n-1}$$

On a $\dfrac{n+1}{n-1} \to 1$ ce qui permet d'écrire $\dfrac{n+1}{n-1} = 1 + \dfrac{2}{n-1}$ avec $\dfrac{2}{n-1} \to 0$.
On a alors

$$u_n = \ln\left(1 + \frac{2}{n-1}\right) \sim \frac{2}{n-1} \sim \frac{2}{n}$$

Attention : On ne peut pas passer les équivalents à l'exponentielle.
Pour $u_n = n+1$ et $v_n = n$ on a $u_n \sim v_n$ alors que e^{u_n} et e^{v_n} ne sont pas équivalent puisque le rapport ne tend pas vers 1.
Pour déterminer la limite d'une exponentielle, on peut étudier un équivalent de son contenu puis de passe à l'exponentielle la limite obtenue.

Exemple Déterminons

$$\lim_{n\to+\infty} \left(1 + \frac{1}{n}\right)^n$$

On a

$$\left(1 + \frac{1}{n}\right)^n = \exp\left(n \ln\left(1 + \frac{1}{n}\right)\right)$$

or

$$\ln\left(1 + \frac{1}{n}\right) \sim \frac{1}{n}$$

donc

$$n \ln\left(1 + \frac{1}{n}\right) \to 1$$

puis par composition de limites

$$\lim_{n\to+\infty} \left(1 + \frac{1}{n}\right)^n = \mathrm{e}$$

Chapitre 16

Fonctions numériques

$\mathcal{D}$ et $\mathcal{D}'$ désignent des parties de $\mathbb{R}$.

16.1 Fonctions réelles

16.1.1 Définitions générales

Définition

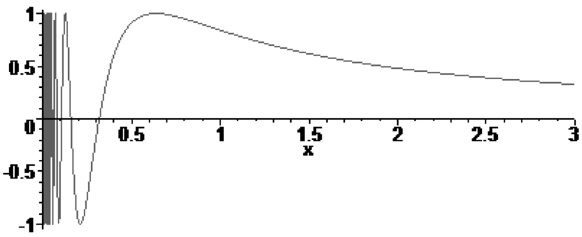
On appelle fonction réelle définie sur $\mathcal{D}$ toute application $f : \mathcal{D} \to \mathbb{R}$.
$\mathcal{D}$ est alors appelé ensemble de définition de f, on note parfois $\mathcal{D} = \mathcal{D}_f$.
On note $\mathcal{F}(\mathcal{D}, \mathbb{R})$ l'ensemble des applications de $\mathcal{D}$ vers $\mathbb{R}$.

Définition

On appelle graphe (ou courbe représentative) d'une fonction $f : \mathcal{D} \to \mathbb{R}$ l'ensemble

$$\Gamma_f = \{M(x, y)/x \in \mathcal{D} \text{ et } y = f(x)\}$$

Exemple $f : \mathbb{R}^{+\star} \to \mathbb{R}$ définie par $f(x) = \sin(1/x)$ est une fonction réelle.



Définition

Une fonction $f : \mathcal{D} \to \mathbb{R}$ est dite constante s'il existe $C \in \mathbb{R}$ tel que

$$\forall x \in \mathcal{D}, f(x) = C$$

Une telle fonction est alors dite constante égale à C et est parfois simplement notée C.

Exemple La fonction nulle est la fonction constante égale à 0 ; on la note 0, $\tilde{0}$ ou plus rarement θ.

Définition

On appelle valeur absolue d'une fonction $f : \mathcal{D} \to \mathbb{R}$ la fonction $|f| : \mathcal{D} \to \mathbb{R}$ définie par

$$\forall x \in \mathcal{D}, |f|(x) = |f(x)|$$

Proposition

Soit $f : \mathcal{D} \to \mathbb{R}$. On a équivalence entre :
(i) f est bornée ;
(ii) $\exists M \in \mathbb{R}, \forall x \in \mathcal{D}, |f(x)| \leqslant M$;
(iii) $|f|$ est majorée.

dém. :
(i) $\Rightarrow$ (ii) Supposons f bornée.
La fonction f est minorée et majorée et donc alors il existe $m_1, m_2 \in \mathbb{R}$ tels que

$$\forall x \in \mathcal{D}, m_1 \leqslant f(x) \leqslant m_2$$

Pour $M = \max\{-m_1, m_2\} \in \mathbb{R}$, on a

$$\forall x \in \mathcal{D}, |f(x)| \leqslant M$$

(ii) $\Rightarrow$ (i) Supposons qu'il existe $M \in \mathbb{R}$ tel que

$$\forall x \in \mathcal{D}, |f(x)| \leqslant M$$

On a alors pour tout $x \in \mathcal{D}$, $-M \leqslant f(x) \leqslant M$.
Ainsi f est minorée et majorée et donc bornée.
(ii) $\Leftrightarrow$ (iii) immédiat.
□

Exemple Considérons $f : \mathbb{R} \to \mathbb{R}$ définie par

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

La fonction f est bornée.
En effet, sachant

$$|ab| \leqslant \frac{1}{2}\left(a^2 + b^2\right)$$

on a

$$|x| \leqslant \frac{1}{2}\left(1 + x^2\right)$$

donc

$$|f(x)| \leqslant \frac{1}{2}$$

Définition

Pour $f, g : \mathcal{D} \to \mathbb{R}$, on définit les fonctions $\sup(f,g) : \mathcal{D} \to \mathbb{R}$ et $\inf(f,g) : \mathcal{D} \to \mathbb{R}$ par :

$$\forall x \in \mathcal{D},\ \sup(f,g)(x) = \max(f(x), g(x)) \text{ et } \inf(f,g)(x) = \min(f(x), g(x))$$

Proposition

$\inf(f,g) \leqslant f, g \leqslant \sup(f,g)$.

dém. :
Pour tout $x \in \mathcal{D}$, $\min(f(x), g(x)) \leqslant f(x), g(x) \leqslant \max(f(x), g(x))$.
□

Exemple Pour $f : \mathcal{D} \to \mathbb{R}$, on note $f^+ = \sup(f, 0)$ et $f^- = \sup(-f, 0)$ appelées partie positive et négative de f.

16.1.2 Opérations sur les fonctions numériques

Définition

Pour $f, g : \mathcal{D} \to \mathbb{R}$ et $\lambda \in \mathbb{R}$, on définit les fonction $\lambda.f$, $f + g$ et fg de $\mathcal{D}$ vers $\mathbb{R}$ par :

$$\forall x \in \mathcal{D}, (\lambda.f)(x) = \lambda f(x), (f+g)(x) = f(x) + g(x) \text{ et } (fg)(x) = f(x)g(x)$$

Si de plus g ne s'annule pas, on définit les fonctions $1/g$ et f/g de $\mathcal{D}$ vers $\mathbb{R}$ par :

$$\forall x \in \mathcal{D}, (1/g)(x) = 1/g(x) \text{ et } f/g = f \times 1/g$$

Exemple La fonction $(-1).f$ est encore notée $-f$.

Exemple En introduisant les parties positives et négatives de f, on peut écrire $f = f^+ - f^-$ et $|f| = f^+ + f^-$.

Proposition

Soient $f, g : \mathcal{D} \to \mathbb{R}$ et $\lambda \in \mathbb{R}$.
Si f et g sont bornées alors $\lambda.f$, $f + g$ et fg le sont aussi.

dém. :
Supposons f et g bornées.
Par hypothèse il existe $M, M' \in \mathbb{R}$ vérifiant

$$\forall x \in \mathcal{D},\ |f(x)| \leqslant M \text{ et } |g(x)| \leqslant M'$$

On a alors pour tout $x \in \mathcal{D}$, $|(\lambda.f)(x)| = |\lambda f(x)| = |\lambda|\,|f(x)| \leqslant |\lambda|\,M$,

$$|(f+g)(x)| = |f(x) + g(x)| \leqslant |f(x)| + |g(x)| \leqslant M + M'$$

et

$$|(fg)(x)| = |f(x)g(x)| = |f(x)|\,|g(x)| \leqslant MM'$$

□

16.1.3 Parité

Définition

Une fonction $f : \mathcal{D} \to \mathbb{R}$ est dite paire (resp. impaire) si :
1) le domaine $\mathcal{D}$ est symétrique par rapport à 0 i.e. : $\forall x \in \mathcal{D}, -x \in \mathcal{D}$;
2) $\forall x \in \mathcal{D}, f(-x) = f(x)$ (resp. $f(-x) = -f(x)$).

Exemple La fonction nulle est la seule fonction à la fois paire et impaire.

Proposition

Soient $f : \mathcal{D} \to \mathbb{R}$ et $\varphi : \mathcal{D}' \to \mathbb{R}$ telles que $f(\mathcal{D}) \subset \mathcal{D}'$.
Si f est paire alors $\varphi \circ f$ l'est aussi.
Si f est impaire et φ paire alors $\varphi \circ f$ est paire.
Si f est impaire et φ impaire alors $\varphi \circ f$ est impaire.

dém. :
$\varphi \circ f$ est définie sur $\mathcal{D}$ intervalle symétrique par rapport à 0.
Si f est paire alors pour tout $x \in \mathcal{D}$,

$$(\varphi \circ f)(-x) = \varphi(f(-x)) = \varphi(f(x)) = (\varphi \circ f)(x)$$

donc $\varphi \circ f$ est paire.
Si f est impaire alors pour tout $x \in \mathcal{D}$,

$$(\varphi \circ f)(-x) = \varphi(f(-x)) = \varphi(-f(x))$$

Si de plus φ est paire alors

$$(\varphi \circ f)(-x) = \varphi(f(x)) = (\varphi \circ f)(x)$$

et donc $\varphi \circ f$ est paire.
Si en revanche φ est impaire alors

$$(\varphi \circ f)(-x) = -\varphi(f(x)) = -(\varphi \circ f)(x)$$

et donc $\varphi \circ f$ est impaire.
□

16.1.4 Périodicité

Définition

Une fonction $f : \mathcal{D} \to \mathbb{R}$ est dite T périodique si :
1) le domaine $\mathcal{D}$ est T périodique i.e. $\forall x \in \mathcal{D}, x + T \in \mathcal{D}$;
2) $\forall x \in \mathcal{D}, f(x + T) = f(x)$.

Exemple Toute fonction est 0 périodique.

Définition

Une fonction f est dite périodique s'il existe $T \neq 0$ telle qu'elle soit T périodique.

Exemple $x \mapsto \cos x, x \mapsto \sin x$ sont 2π périodiques.
$x \mapsto x - E(x)$ est 1 périodique.

Exemple Soit $\alpha > 0$.
La fonction $x \mapsto \cos \dfrac{x}{\alpha}$ est $2\pi\alpha$ périodique.
La fonction $x \mapsto |\sin \alpha x|$ est $\dfrac{\pi}{\alpha}$ périodique.

Proposition

Soit $f : \mathbb{R} \to \mathbb{R}$.
Si f est T périodique alors

$$\forall k \in \mathbb{Z}, \forall x \in \mathbb{R}, f(x + kT) = f(x)$$

dém. :
Par récurrence sur $n \in \mathbb{N}$, on obtient facilement,

$$\forall n \in \mathbb{N}, \forall x \in \mathbb{R}, f(x + nT) = f(x)$$

Pour $k \in \mathbb{Z}^-$, on peut écrire $k = -n$ avec $n \in \mathbb{N}$. Pour $x \in \mathbb{R}$, $f(x + kT) = f(x - nT)$. Or par le résultat qui précède $f(x - nT) = f\left((x - nT) + nT\right)$ donc $f(x + kT) = f(x)$.
□

16.1.5 Lipschitzianité

Définition

On dit que $f : \mathcal{D} \to \mathbb{R}f$ est lipschitzienne s'il existe $k \in \mathbb{R}^+$ tel que

$$\forall x, y \in \mathcal{D}, |f(y) - f(x)| \leqslant k\,|y - x|$$

Plus précisément on dit alors que f est k-lipschitzienne.

Remarque Si f est k lipschitzienne alors pour tout $k' \geqslant k$, f est aussi k' lipschitzienne.

Exemple Les fonctions constantes sont les fonctions 0 lipschitziennes.

Exemple $x \mapsto x$ et $x \mapsto |x|$ sont des fonctions 1 lipschitziennes.

Exemple La fonction $x \mapsto \sin x$ est lipschitzienne sur $\mathbb{R}$.
En effet

$$|\sin y - \sin x| = 2\left|\sin\frac{y-x}{2}\right|\left|\cos\frac{y+x}{2}\right| \leqslant 2\left|\sin\frac{y-x}{2}\right|$$

Or pour tout $t \in \mathbb{R}$, $|\sin t| \leqslant |t|$ donc

$$|\sin y - \sin x| \leqslant 2\frac{|y-x|}{2} = |y-x|$$

Exemple La fonction $x \mapsto x^2$ n'est pas lipschitzienne sur $\mathbb{R}$.
En effet, supposons par l'absurde $x \mapsto x^2$ lipschitzienne sur $\mathbb{R}$.
Il existe alors $M \in \mathbb{R}$ tel que pour tout $x, y \in \mathbb{R}$,

$$\left|y^2 - x^2\right| \leqslant M\,|y-x|$$

Pour $y = x+1$, on obtient $|2x+1| \leqslant M$.
Or quand $x \to +\infty$, $|2x+1| \to +\infty$: c'est absurde.

Proposition

> Soient $f, g : \mathcal{D} \to \mathbb{R}$ et $\lambda \in \mathbb{R}$.
> Si f est k lipschitzienne alors $\lambda.f$ est $|\lambda|\,k$ lipschitzienne.
> Si f et g sont k et k' lipschitziennes alors $f+g$ est $k+k'$ lipschitzienne.

dém. :

$$|(\lambda.f)(y) - (\lambda.f)(x)| = |\lambda f(y) - \lambda f(x)| = |\lambda|\,|f(y)-f(x)| \leqslant |\lambda|\,k\,|y-x|$$

et

$$\begin{aligned}|(f+g)(y) - (f+g)(x)| &= |f(y)+g(y)-f(x)-g(x)| \\ &\leqslant |f(y)-f(x)| + |g(y)-g(x)| \leqslant (k+k')\,|y-x|\end{aligned}$$

□

Proposition

> Soit $f : \mathcal{D} \to \mathbb{R}$ et $\varphi : \mathcal{D}' \to \mathbb{K}$ telles que $f(\mathcal{D}) \subset \mathcal{D}'$.
> Si f et φ sont k et k' lipschitziennes alors $\varphi \circ f$ est kk' lipschitzienne.

dém. :

$$|(\varphi\circ f)(y) - (\varphi\circ f)(x)| = |\varphi(f(y)) - \varphi(f(x))| \leqslant k'\,|f(y)-f(x)| \leqslant kk'\,|y-x|$$

□

16.1.6 Propriété vraie sur un voisinage

Définition

> Soient une fonction $f : \mathcal{D} \to \mathbb{R}$ et Δ une partie de $\mathcal{D}$.
> On dit que f présente une certaine propriété sur Δ si la restriction $f_{\restriction\Delta}$ présente cette propriété.

Exemple $x \mapsto \sin x$ est croissante sur $[-\pi/2, \pi/2]$.

Exemple $x \mapsto \ln x$ est positive sur $[1, +\infty[$.

Exemple $x \mapsto E(x)$ est constante sur chaque $I_k = [k, k+1[$ avec $k \in \mathbb{Z}$.

Définition

Soit $a \in \mathbb{R}$, on appelle voisinage de a tout ensemble $V = [a - \alpha, a + \alpha]$ avec $\alpha > 0$.

Définition

On appelle voisinage de $+\infty$ (resp. $-\infty$) tout ensemble de la forme $V = [A, +\infty[$ (resp. $V =]-\infty, A]$) avec $A \in \mathbb{R}$.

Remarque L'intersection de deux voisinages de $a \in \bar{\mathbb{R}}$ est encore un voisinage de a.

Définition

On dit que $f : \mathcal{D} \to \mathbb{R}$ est définie au voisinage de $a \in \bar{\mathbb{R}}$ si pour tout voisinage V de a, on a $V \cap \mathcal{D} \neq \emptyset$.

Exemple La fonction $x \mapsto \ln x$ est définie au voisinage de $1, 0$ ou $+\infty$ mais pas au voisinage de -1.

Exemple Si I est un intervalle non vide de $\mathbb{R}$, une fonction $f : I \to \mathbb{R}$ est définie au voisinage de chaque point de I et aussi au voisinage des extrémités de I.

Exemple Si I est un intervalle non singulier de $\mathbb{R}$ et si a est un élément de I alors une fonction définie sur $I \setminus \{a\}$ est définie au voisinage de a.

Définition

Soit $f : \mathcal{D} \to \mathbb{R}$ une fonction définie au voisinage de $a \in \bar{\mathbb{R}}$.
On dit que f présente une propriété au voisinage de a s'il existe un voisnage V de a tel que f présente la propriété sur $V \cap \mathcal{D}$.

Exemple La fonction $x \mapsto \sin x$ est croissante au voisinage de 0,
La fonction $x \mapsto \sin x$ est croissante au voisinage de chaque point de $]-\pi/2, \pi/2[$.
La fonction $x \mapsto \sin x$ n'est pas croissante au voisinage de $\pi/2$.

Exemple La fonction $x \mapsto 1/x$ est bornée au voisinage de $+\infty$.

Exemple La fonction $x \mapsto \cos x$ n'est pas monotone au voisinage de $+\infty$.

Exemple $x \mapsto E(x)$ est constante au voisinage de tout $a \in \mathbb{R} \setminus \mathbb{Z}$.

16.2 Limites d'une fonction réelle

Soient $f, g, h : \mathcal{D} \to \mathbb{R}$ trois fonctions définies au voisinage de $a \in \bar{\mathbb{R}}$.

16.2.1 Limite finie

16.2.1.1 Définition

Définition

On dit que $f : \mathcal{D} \to \mathbb{R}$ tend vers $\ell \in \mathbb{R}$ en a si l'on a la propriété

$$\forall \varepsilon > 0, \exists V \text{ voisinage de } a, \forall x \in \mathcal{D}, x \in V \Rightarrow |f(x) - \ell| \leqslant \varepsilon$$

On note alors $f \underset{a}{\to} \ell$, $f(x) \xrightarrow[x \to a]{} \ell$ ou $f(x) \xrightarrow[x \to a, x \in \mathcal{D}]{} \ell$

Ainsi :

Cas $a \in \mathbb{R}$

$$f \underset{a}{\to} \ell \Leftrightarrow \forall \varepsilon > 0, \exists \alpha > 0, \forall x \in \mathcal{D}, |x - a| \leqslant \alpha \Rightarrow |f(x) - \ell| \leqslant \varepsilon$$

Cas $a = +\infty$

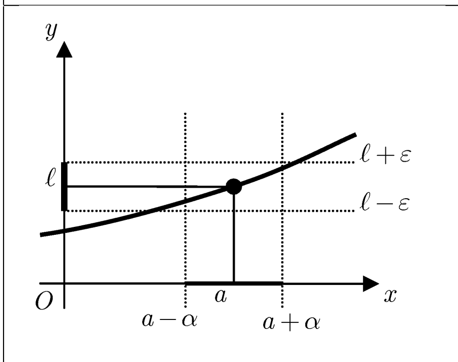
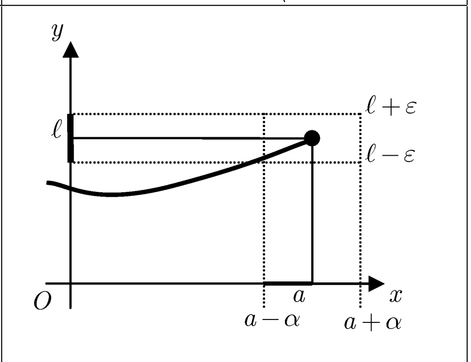
$$f \xrightarrow[+\infty]{} \ell \Leftrightarrow \forall \varepsilon > 0, \exists A \in \mathbb{R}, \forall x \in \mathcal{D}, x \geqslant A \Rightarrow |f(x) - \ell| \leqslant \varepsilon$$

Cas $a = -\infty$

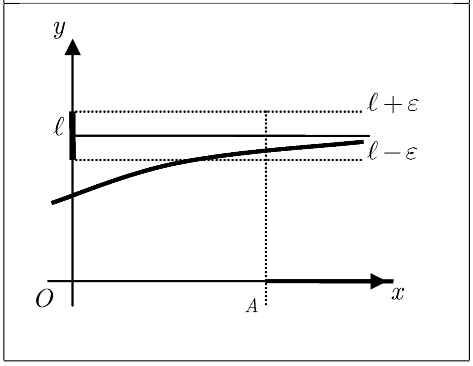
$$f \xrightarrow[-\infty]{} \ell \Leftrightarrow \forall \varepsilon > 0, \exists A \in \mathbb{R}, \forall x \in \mathcal{D}, x \leqslant A \Rightarrow |f(x) - \ell| \leqslant \varepsilon$$

Remarque Dire que f tend ℓ en a signifie que pour tout $\varepsilon > 0$, aussi petit soit-il, il existe un voisinage de a sur lequel :

$$\ell - \varepsilon \leqslant f(x) \leqslant \ell + \varepsilon$$

Cas $a \in \mathcal{D}$	Cas $a \in \mathbb{R} \backslash \mathcal{D}$
	

Cas $a = +\infty$



Remarque On obtient une définition équivalente en transformant « $\forall \varepsilon > 0$ »en « $\forall \varepsilon \in]0,1]$ ».
En effet un voisinage convenant pour $\varepsilon = 1$ conviendra a fortiori pour tout $\varepsilon \geqslant 1$.

Proposition

> Si f est définie en $a \in \mathbb{R}$ et si $f \underset{a}{\to} \ell$ alors nécessairement $\ell = f(a)$.
> On dit alors que f est continue en a (et ce concept est suffisamment important pour nécessiter un chapitre entier…)

dém. :
Puisqu'on suppose que $f \underset{a}{\to} \ell$, pour tout $\varepsilon > 0$, il existe $\alpha > 0$ vérifiant :

$$\forall x \in \mathcal{D}, |x - a| \leqslant \alpha \Rightarrow |f(x) - \ell| \leqslant \varepsilon$$

Or on suppose que a appartient à $\mathcal{D}$ et puisqu'on a $|a - a| = 0 \leqslant \alpha$, l'implication qui précède entraîne $|f(a) - \ell| \leqslant \varepsilon$. Puisque ceci vaut pour tout $\varepsilon > 0$, on peut conclure $f(a) = \ell$.
□

Exemple Soient $f : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = x$ et $a \in \mathbb{R}$.
On a $f(x) \xrightarrow[x \to a]{} a$ c'est-à-dire $x \xrightarrow[x \to a]{} a$.
En effet, soit $\varepsilon > 0$. Pour $\alpha = \varepsilon$, on a

$$\forall x \in \mathbb{R}, |x - a| \leqslant \alpha \Rightarrow |f(x) - a| \leqslant \varepsilon \text{ car } |f(x) - a| = |x - a|$$

Exemple Soit $f : \mathbb{R}^{+\star} \to \mathbb{R}$ définie par $f(x) = 1/x$.
On a $f(x) \xrightarrow[x \to +\infty]{} 0$.
En effet, soit $\varepsilon > 0$. Pour $A = 1/\varepsilon \in \mathbb{R}$, on a

$$\forall x \in \mathbb{R}^{\star}, x \geqslant A \Rightarrow |f(x) - 0| \leqslant \varepsilon \text{ car } f(x) - 0 = 1/x$$

Exemple Soient $f : \mathbb{R} \to \mathbb{R}$ la fonction constante égale à C et $a \in \bar{\mathbb{R}}$.
On montre aisément $f(x) \xrightarrow[x \to a]{} C$.

Proposition

Soit $\ell \in \mathbb{R}$. On a équivalence entre :
(i) $f(x) \xrightarrow[x\to a]{} \ell$;
(ii) $|f(x) - \ell| \xrightarrow[x\to a]{} 0$.

dém. :
Puisque $||f(x) - \ell| - 0| = |f(x) - \ell|$, les phrases quantifiées traduisant $f(x) \xrightarrow[x\to a]{} \ell$ et $|f(x) - \ell| \xrightarrow[x\to a]{} 0$ sont indentiques.
□

Exemple Si $f \xrightarrow[a]{} \ell$ alors $-f \xrightarrow[a]{} -\ell$.
En effet $|-f(x) - (-\ell)| = |\ell - f(x)| = |f(x) - \ell| \xrightarrow[x\to a]{} 0$.

Remarque Il est fréquent que pour établir $f(x) \xrightarrow[x\to a]{} \ell$ qu'on « se ramène à 0 »en étudiant si $|f(x) - \ell| \xrightarrow[x\to a]{} 0$.

16.2.1.2 Convergence et divergence

Définition

On dit que f converge en a s'il existe $\ell \in \mathbb{R}$ tel que $f \xrightarrow[a]{} \ell$.
Sinon, on dit que f diverge en a.
Déterminer la nature de f en a c'est savoir si cette fonction converge ou diverge en a.

Exemple $x \mapsto \cos x$ diverge en $+\infty$ et $x \mapsto \sin(1/x)$ diverge en 0.

Théorème

Si f converge en a alors il existe un unique $\ell \in \mathbb{R}$ tel que $f \xrightarrow[a]{} \ell$.
Ce réel ℓ est alors appelé limite de f en a et on note $\ell = \lim_a f$, $\ell = \lim_{x\to a} f(x)$ ou $\ell = \lim_{x\to a, x\in\mathcal{D}} f(x)$.

dém. :
Supposons $f \xrightarrow[a]{} \ell$ et $f \xrightarrow[a]{} \ell'$.
Soit $\varepsilon > 0$. Il existe un voisinage de a sur lequel $|f(x) - \ell| \leqslant \varepsilon/2$ et un autre voisinage sur lequel $|f(x) - \ell'| \leqslant \varepsilon/2$. Pour un élément x appartenant à l'intersection de ces deux voisinages (un tel élément existe car f est supposée définie au voisinage de a), on a à la fois $|f(x) - \ell| \leqslant \varepsilon/2$ et $|f(x) - \ell'| \leqslant \varepsilon/2$.
On en déduit

$$|\ell - \ell'| \leqslant |\ell - f(x)| + |f(x) - \ell'| \leqslant \varepsilon$$

Or ceci vaut pour tout $\varepsilon > 0$ donc $\ell = \ell'$.
□

Exemple On peut écrire $\lim_{x\to a} C = C$, $\lim_{x\to a} x = a$ et $\lim_{x\to +\infty} 1/x = 0$.

16.2.1.3 Limite finie et relation d'ordre

Théorème

Si f converge en a alors f est bornée au voisinage de a.

dém. :
Supposons $f \underset{a}{\to} \ell \in \mathbb{R}$.
Pour $\varepsilon = 1$, il existe un voisinage de a sur lequel $|f(x) - \ell| \leqslant 1$ et alors sur ce voisinage $|f(x)| = |\ell + f(x) - \ell| \leqslant |\ell| + |f(x) - \ell| \leqslant |\ell| + 1$.
Ainsi f est bornée au voisinage de a.
□

Attention : La réciproque fausse, la fonction $x \mapsto \sin(1/x)$ est bornée sur $\mathbb{R}$ et en particulier au voisinage de 0 alors que cette fonction diverge en 0.

Proposition

Soient $\alpha, \beta \in \mathbb{R}$ et $\ell \in \mathbb{R}$
Si $f \underset{a}{\to} \ell$ et si $\ell > \alpha$ alors $f(x) > \alpha$ au voisinage de a.
Si $f \underset{a}{\to} \ell$ et si $\ell < \beta$ alors $f(x) < \beta$ au voisinage de a.
Si $f \underset{a}{\to} \ell$ et si $\alpha < \ell < \beta$ alors $\alpha < f(x) < \beta$ au voisinage de a.

dém. :
Supposons $f \underset{a}{\to} \ell$ et $\ell > \alpha$.
Pour $\varepsilon = (\ell - \alpha)/2 > 0$, il existe un voisinage de a sur lequel $|f(x) - \ell| \leqslant \varepsilon$.
Sur ce voisinage, on a $\ell - \varepsilon \leqslant f(x) \leqslant \ell + \varepsilon$ et donc $f(x) \geqslant \dfrac{\ell + \alpha}{2} > \alpha$.
Le cas $f \underset{a}{\to} \ell$ avec $\ell < \beta$ se traite de façon analogue et le cas $\alpha < \ell < \beta$ se résout en considérant l'intersection des deux voisinages précédents.
□

Exemple Si $f \underset{a}{\to} \ell \neq 0$ alors la fonction $1/f$ est définie au voisinage de a.
En effet, selon que $\ell > 0$ ou $\ell < 0$, il existe un voisinage de a sur lequel la fonction f est de signe strict constant.

Théorème

Soient $f, g : \mathcal{D} \to \mathbb{R}$ telles $f \underset{a}{\to} \ell$ et $g \underset{a}{\to} \ell'$.
Si $f(x) \leqslant g(x)$ au voisinage de a alors $\ell \leqslant \ell'$

dém. :
Par contraposée, supposons $\ell > \ell'$ et considérons $\alpha = (\ell + \ell')/2$.
Puisque $f \underset{a}{\to} \ell > \alpha$, il existe un voisinage de a sur lequel $f(x) > \alpha$.
Aussi, puisque $g \underset{a}{\to} \ell' < \alpha$, il existe un voisinage de a sur lequel $g(x) < \alpha$.
Sur l'intersection de ces deux voisinages, on a $g(x) < \alpha < f(x)$ ce qui fait plus que contredire l'hypothèse $f(x) \leqslant g(x)$ au voisinage de a.
□

Attention : Ce résultat est faux en terme d'inégalité stricte ; on retient que, par passage à la limite, les inégalités strictes deviennent larges.

16.2.2 Limites infinies

16.2.2.1 Définition

Définition

On dit que f tend vers $+\infty$ en a si on a la propriété

$$\forall M \in \mathbb{R}, \exists V \text{ voisinage de } a, \forall x \in \mathcal{D}, x \in V \Rightarrow f(x) \geqslant M$$

On note alors $f \underset{a}{\rightarrow} +\infty$, $f(x) \xrightarrow[x \to a]{} +\infty$ ou $f(x) \xrightarrow[x \to a, x \in \mathcal{D}]{} +\infty$

Ainsi :

Cas $a \in \mathbb{R}$

$$f \underset{a}{\rightarrow} +\infty \Leftrightarrow \forall M \in \mathbb{R}, \exists \alpha > 0, \forall x \in \mathcal{D}, |x-a| \leqslant \alpha \Rightarrow f(x) \geqslant M$$

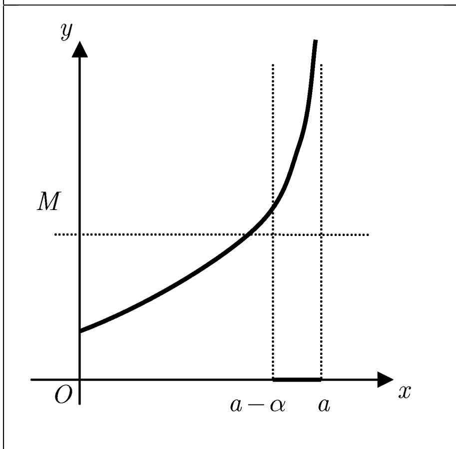
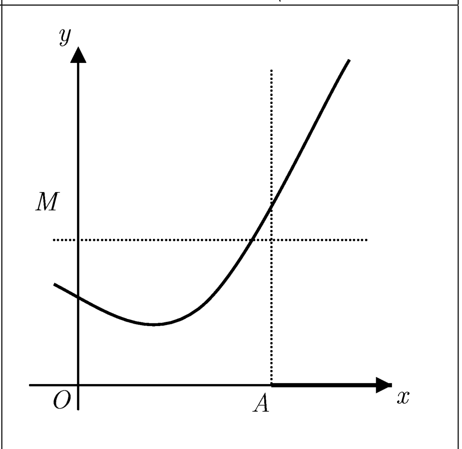
Cas $a = +\infty$

$$f \xrightarrow[+\infty]{} +\infty \Leftrightarrow \forall M \in \mathbb{R}, \exists A \in \mathbb{R}, \forall x \in \mathcal{D}, x \geqslant A \Rightarrow f(x) \geqslant M$$

Cas $a = -\infty$

$$f \xrightarrow[-\infty]{} +\infty \Leftrightarrow \forall M \in \mathbb{R}, \exists A \in \mathbb{R}, \forall x \in \mathcal{D}, x \leqslant A \Rightarrow f(x) \geqslant M$$

Remarque Dire que f tend vers $+\infty$ en a signifier que pour tout A, arbitrairement grand, il existe un voisinage de a sur lequel $f(x) \geqslant A$.

Cas $a = +\infty$	Cas $a \in \mathbb{R} \backslash \mathcal{D}$
	

Remarque On obtient une définition équivalente en remplaçant « $\forall M \in \mathbb{R}$ »par « $\forall M \in \mathbb{R}^+$ », « $\forall M \in [1, +\infty[$ »,...
En effet un voisinage convenant pour $M = 1$ conviendra a fortiori pour tout $M \leqslant 1$.

Remarque Si f est définie en $a \in \mathbb{R}$ alors on ne peut pas avoir $f(x) \xrightarrow[x \to a]{} +\infty$.
En effet, si f est définie en a alors on obtient l'absurdité $f(a) \geqslant M$ pour tout $M \in \mathbb{R}$ car au voisinage de a, $f(x) \geqslant M$.

Exemple Soit $f : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = x$.
On a $f(x) \xrightarrow[x \to +\infty]{} +\infty$ c'est-à-dire $x \xrightarrow[x \to +\infty]{} +\infty$.
En effet, quel que soit $M \in \mathbb{R}$, pour $A = M$, on a

$$\forall x \in \mathbb{R}, x \geqslant A \Rightarrow x \geqslant M$$

Exemple Soit $f : \mathbb{R}^{+\star} \to \mathbb{R}$ définie par $f(x) = 1/x$.
Montrons que $f(x) \xrightarrow[x \to 0]{} +\infty$.
Soit $M \in \mathbb{R}^{+\star}$. Pour $\alpha = 1/M > 0$, on a

$$\forall x \in \mathbb{R}^{+\star}, |x| \leqslant \alpha \Rightarrow f(x) \geqslant M$$

Définition

On dit que f tend vers $-\infty$ en a si on a la propriété

$$\forall M \in \mathbb{R}, \exists V \text{ voisinage de } a, \forall x \in \mathcal{D}, x \in V \Rightarrow f(x) \leqslant M$$

On note alors $f \underset{a}{\to} -\infty$ ou $f(x) \underset{x \to a}{\to} -\infty$.

Proposition

$f \underset{a}{\to} +\infty \Leftrightarrow -f \underset{a}{\to} -\infty$.

dém. :
($\Rightarrow$) Supposons $f \underset{a}{\to} +\infty$.
Pour $M \in \mathbb{R}$, puisque $-M \in \mathbb{R}$, il existe un voisinage de a sur lequel $f(x) \geqslant -M$ et alors $-f(x) \leqslant M$.
Ainsi $-f \underset{a}{\to} -\infty$.
($\Leftarrow$) Même principe.
□

16.2.2.2 Limites infinies et relation d'ordre

Proposition

Si $f \underset{a}{\to} +\infty$ alors f est minorée mais non majorée au voisinage de a.
Si $f \underset{a}{\to} -\infty$ alors f est majorée mais non minorée au voisinage de a.

dém. :
Suppsons $f \underset{a}{\to} +\infty$.
Pour $M = 0$, il existe un voisinage de a sur lequel $f(x) \geqslant 0$ et ainsi f est minorée au voisinage de a.
Pour $M \in \mathbb{R}$, en considérant $M + 1 \in \mathbb{R}$, il existe un voisinage de a sur lequel $f(x) \geqslant M + 1$ et par

suite f n'est pas majorée par M au voisinage de a.
Par passage à l'opposé, ce qui précède se transpose au cas $f \underset{a}{\rightarrow} -\infty$.

□

Remarque Si $f \underset{a}{\rightarrow} +\infty$ (resp. $f \underset{a}{\rightarrow} -\infty$) alors f n'est pas bornée au voisinage de a et donc f diverge en a.

Définition

> Si $f \underset{a}{\rightarrow} +\infty$ (resp. $f \underset{a}{\rightarrow} -\infty$) alors on dit que f diverge vers $+\infty$ en a (resp. $-\infty$) et on note $\lim\limits_{a} f = +\infty$ (resp. $\lim\limits_{a} f = -\infty$).

Exemple On peut écrire $\lim\limits_{x \to +\infty} x = +\infty$ et $\lim\limits_{x \to 0, x \in \mathbb{R}^{+\star}} \dfrac{1}{x} = +\infty$.

16.2.3 Limite à droite, limite à gauche en un point

Définition

> Soient $f : \mathcal{D} \to \mathbb{R}$, $a \in \mathbb{R}$ et $\mathcal{D}^+ = \mathcal{D} \cap \,]a, +\infty[$
> On dit que f est définie au voisinage à droite en a si $f_{\restriction \mathcal{D}^+}$ est définie au voisinage de a.
> L'éventuelle limite $\ell \in \bar{\mathbb{R}}$ de $f_{|\mathcal{D}^+}$ en a est alors appelée limite à droite de f en a.
> On note $\ell = \lim\limits_{a^-} f$ ou $\ell = \lim\limits_{x \to a^+} f(x)$.

Définition

> Soient $f : \mathcal{D} \to \mathbb{R}$, $a \in \mathbb{R}$ et $\mathcal{D}^- = \mathcal{D} \cap \,]-\infty, a[$
> On dit que f est définie au voisinage à gauche en a si $f_{\restriction \mathcal{D}^-}$ est définie au voisinage de a.
> L'éventuelle limite $\ell \in \bar{\mathbb{R}}$ de $f_{|\mathcal{D}^-}$ en a est alors appelée limite à gauche de f en a.
> On note alors $\ell = \lim\limits_{a^-} f$ ou $\lim\limits_{x \to a^-} f(x)$.

Exemple $\lim\limits_{x \to 0^+} \dfrac{1}{x} = +\infty$ et $\lim\limits_{x \to 0^-} \dfrac{1}{x} = -\infty$.

Exemple Pour $n \in \mathbb{Z}$, on a $\lim\limits_{x \in n^+} E(x) = n$ et $\lim\limits_{x \to n^-} E(x) = n - 1$.

Remarque Précisement

$$f \underset{a^+}{\longrightarrow} \ell \in \mathbb{R} \Leftrightarrow \forall \varepsilon > 0, \exists \alpha > 0, \forall x \in \mathcal{D}, a < x < a + \alpha \Rightarrow |f(x) - \ell| \leqslant \varepsilon$$

$$f \underset{a^-}{\longrightarrow} \ell \in \mathbb{R} \Leftrightarrow \forall \varepsilon > 0, \exists \alpha > 0, \forall x \in \mathcal{D}, a - \alpha < x < a \Rightarrow |f(x) - \ell| \leqslant \varepsilon$$

et de même on peut exprimer $f \underset{a^+}{\longrightarrow} +\infty, f \underset{a^+}{\longrightarrow} -\infty$ etc.
Notons qu'à chaque fois l'éventuelle valeur de f en a n'est pas prise en compte pour la détermination des limites $\lim\limits_{a+} f$ et $\lim\limits_{a-} f$.

Remarque Puisque les limites à droite et à gauche de f ne sont autres que des limites de restriction de f, les résultats qui suivent s'appliquent aux limites à droite et à gauche

16.2.4 Opérations sur les limites

16.2.4.1 Caractérisation séquentielle des limites

Théorème

Soient $f : \mathcal{D} \to \mathbb{R}$ définie au voisinage de $a \in \bar{\mathbb{R}}$ et $\ell \in \bar{\mathbb{R}}$.
On a équivalence entre :
(i) $f \underset{a}{\to} \ell$;
(ii) $\forall (u_n) \in \mathcal{D}^{\mathbb{N}}, u_n \to a \Rightarrow f(u_n) \to \ell$.

dém. :
Nous ne traiterons que le cas $a \in \mathbb{R}$ et $\ell \in \mathbb{R}$; les autres cas sont très semblables.
(i) $\Rightarrow$ (ii)
Supposons $f \underset{a}{\to} \ell$.
Soit $(u_n) \in \mathcal{D}^{\mathbb{N}}$ tel que $u_n \to a$.
Soit $\varepsilon > 0$.
Puisque $f \underset{a}{\to} \ell$, il existe $\alpha > 0$ vérifiant

$$\forall x \in \mathcal{D}, |x - a| \leqslant \alpha \Rightarrow |f(x) - \ell| \leqslant \varepsilon$$

Puisque $u_n \to a$, il existe $N \in \mathbb{N}$ vérifiant

$$\forall n \geqslant N, |u_n - a| \leqslant \alpha$$

On a alors

$$\forall n \geqslant N, |f(u_n) - \ell| \leqslant \varepsilon$$

Ainsi $f(u_n) \to \ell$.
(ii) $\Rightarrow$ (i) Par contraposée :
Supposons que f ne tend pas vers ℓ en a.
Il existe alors $\varepsilon > 0$ tel que pour tout $\alpha > 0$, il existe $x \in \mathcal{D}$ vérifiant $|x - a| \leqslant \alpha$ et $|f(x) - \ell| > \varepsilon$.
Pour $n \in \mathbb{N}$, prenons $\alpha = 1/(n+1)$ et notons u_n un élément de $\mathcal{D}$ tel que $|u_n - a| \leqslant \alpha$ et $|f(u_n) - \ell| > \varepsilon$.
En faisant varier n, ce qui précède définit une suite (u_n) d'éléments de $\mathcal{D}$ telle que $u_n \to a$ et $(f(u_n))$ ne tend pas vers ℓ.
□

Remarque (i) $\Rightarrow$ (ii) correspond au théorème de composition de limites vu dans le cadre des suites réelles.
(i) $\Rightarrow$ (ii) permet aussi d'établir qu'une fonction n'a pas de limite en a.

Exemple Montrer que la fonction $x \mapsto \cos x$ n'a pas de limite en $+\infty$.
Considérons $u_n = 2n\pi \to +\infty$ et $v_n = (2n+1)\pi \to +\infty$.
On a $\cos(u_n) = 1 \to 1$ et $\cos(v_n) = -1 \to -1 \neq 1$ donc la fonction $\cos$ ne peut avoir de limite en $+\infty$.
De même, on peut aussi montrer que la n $x \mapsto \sin(1/x)$ n'a pas de limite en 0.

Remarque (ii) $\Rightarrow$ (i) est utile pour construite rapidement la suite de la théorie :

16.2.4.2 Somme et produit

Théorème

Supposons $f \underset{a}{\to} \ell \in \bar{\mathbb{R}}$ et $g \underset{a}{\to} \ell' \in \bar{\mathbb{R}}$.
Si $\ell + \ell'$ est définie dans $\bar{\mathbb{R}}$ alors $f + g \underset{a}{\to} \ell + \ell'$.
Si $\ell\ell'$ est définie dans $\bar{\mathbb{R}}$ alors $fg \underset{a}{\to} \ell\ell'$.

dém. :
Supposons $\ell + \ell'$ défini dans $\bar{\mathbb{R}}$.
Soit $(u_n) \in \mathcal{D}^{\mathbb{N}}$.
Si $u_n \to a$ alors $f(u_n) \to \ell$ et $g(u_n) \to \ell'$.
Donc $(f+g)(u_n) = f(u_n) + g(u_n) \to \ell + \ell'$.
Par la caractérisation séquentielle de limites on peut alors affirmer que $f + g \underset{a}{\to} \ell + \ell'$.
L'étude pour le produit est évidemment analogue.
□

Remarque Rappelons que certaines formes sont indéterminées parmis lesquelles figurent
$(+\infty) + (-\infty)$ et $0 \times (+\infty)$.

16.2.4.3 Passage à l'inverse

Définition

Soit $\ell \in \mathbb{R}$.
On note $f \underset{a}{\to} \ell^+$ pour signifier que $f \underset{a}{\to} \ell$ et $f(x) > \ell$ au voisinage de a.
On note $f \underset{a}{\to} \ell^-$ pour signifier que $f \underset{a}{\to} \ell$ et $f(x) < \ell$ au voisinage de a.

Théorème

1) Si $f \underset{a}{\to} \ell \in \mathbb{R}^\star$ alors $1/f \underset{a}{\to} 1/\ell$.
2) Si $f \underset{a}{\to} 0^+$ alors $1/f \underset{a}{\to} +\infty$.
3) Si $f \underset{a}{\to} 0^-$ alors $1/f \underset{a}{\to} -\infty$.
4) Si $f \underset{a}{\to} +\infty$ alors $1/f \underset{a}{\to} 0^+$.
5) Si $f \underset{a}{\to} -\infty$ alors $1/f \underset{a}{\to} 0^-$.

dém. :
1) Supposons $f \underset{a}{\to} \ell \in \mathbb{R}^\star$.
La fonction f est alors de signe strict au voisinage de a et donc la fonction $1/f$ est définie au voisinage de a sur un certaine domaine $\mathcal{D}' = \mathcal{D} \backslash f^{-1}(\{0\})$.
Soit (u_n) une suite d'éléments de $\mathcal{D}'$.
Si $u_n \to a$ alors $f(u_n) \to \ell \neq 0$ puis, par passage à l'inverse d'une suite convergente, on obtient $1/f(u_n) \to 1/\ell$.
En vertu de la caractérisation séquentielle des limites, on peut affirmer que $1/f \underset{a}{\to} 1/\ell$.

2) 3) 4) et 5) s'obtiennent de façon analogue en observant à chaque fois que $1/f$ est définie au voisinage de a.

□

Corollaire

On en déduit les règles relatives au rapport de deux limites.

16.2.4.4 Composition

Théorème

Soient $f : \mathcal{D} \to \mathbb{R}$ et $\varphi : \mathcal{D}' \to \mathbb{R}$ telles que $f(\mathcal{D}) \subset \mathcal{D}'$.
Si $f \underset{a}{\to} b \in \bar{\mathbb{R}}$ et si $\varphi \underset{b}{\to} \ell \in \bar{\mathbb{R}}$ alors $\varphi \circ f \underset{a}{\to} \ell$.

dém. :
Soit (u_n) une suite d'éléments de $\mathcal{D}$.
Si $u_n \to a$ alors $f(u_n) \to b$ puis $\varphi(f(u_n)) \to \ell$.
Ainsi $(\varphi \circ f)(u_n) \to \ell$.
En vertu de la caractéristisation séquentielle des limites, on peut affirmer que $\varphi \circ f \underset{a}{\to} \ell$.

□

16.2.5 Etude pratique

16.2.5.1 Limite en l'infini

Rappelons quelques limites de référence résolvant des indéterminations classiques en $+\infty$:

$$\forall \alpha > 0,\ \lim_{x \to +\infty} \frac{\ln x}{x^\alpha} = 0 \text{ et } \lim_{x \to +\infty} x^\alpha \mathrm{e}^{-x} = 0$$

Exemple Etudions

$$\lim_{x \to +\infty} (x - \ln x)$$

Quand $x \mapsto +\infty$,
$x \to +\infty$ et $\ln x \to +\infty$, il y a une forme indéterminée additive.

$$x - \ln x = x(1 - \frac{\ln x}{x})$$

$x \to +\infty$ et par limite de référence $\dfrac{\ln x}{x} \to 0$ donc par opérations $x - \ln x \to +\infty$.
Ainsi

$$\lim_{x \to +\infty} (x - \ln x) = +\infty$$

Exemple Etudions

$$\lim_{x \to +\infty} (\sqrt{x^2+1} - \sqrt{x^2-1})$$

Quand $x \to +\infty$,
$\sqrt{x^2+1} \to +\infty$ et $\sqrt{x^2-1} \to +\infty$, il y a une forme indéterminée additive.

$$\sqrt{x^2+1} - \sqrt{x^2-1} = \frac{(x^2+1) - (x^2-1)}{\sqrt{x^2+1} + \sqrt{x^2-1}} = \frac{2}{\sqrt{x^2+1} + \sqrt{x^2-1}} \to 0$$

Ainsi

$$\lim_{x \to +\infty} (\sqrt{x^2+1} - \sqrt{x^2-1}) = 0$$

Exemple Etudions

$$\lim_{x\to+\infty} \frac{\sqrt{x^2+x+1}}{x}$$

Quand $x \to +\infty$,
$\sqrt{x^2+x+1} \to +\infty$ et $x \to +\infty$, il y a une forme indéterminée multiplicative.

$$\frac{\sqrt{x^2+x+1}}{x} = \frac{\sqrt{x^2}\sqrt{1+1/x+1/x^2}}{x} = \sqrt{1+\frac{1}{x}+\frac{1}{x^2}} \to 1$$

car on peut supposer $x \geqslant 0$ puisque $x \to +\infty$.
Ainsi

$$\lim_{x\to+\infty} \frac{\sqrt{x^2+x+1}}{x} = 1$$

Exemple Etudions

$$\lim_{x\to-\infty} \frac{\sqrt{x^2+x+1}}{x}$$

Quand $x \to -\infty$,

$$\frac{\sqrt{x^2+x+1}}{x} = \frac{\sqrt{x^2}\sqrt{1+1/x+1/x^2}}{x} = -\sqrt{1+\frac{1}{x}+\frac{1}{x^2}} \to -1$$

car on peut supposer $x \leqslant 0$ puisque $x \to -\infty$.
Ainsi

$$\lim_{x\to-\infty} \frac{\sqrt{x^2+x+1}}{x} = -1$$

Exemple Etudions

$$\lim_{x\to+\infty} x\mathrm{e}^{-\sqrt{x}}$$

Quand $x \to +\infty$,
$x \to +\infty$ et $\mathrm{e}^{-\sqrt{x}} \to 0$, il y a une forme indéterminée multiplicative.
Posons $y = \sqrt{x} \to +\infty$,
$x\mathrm{e}^{-\sqrt{x}} = y^2\mathrm{e}^{-y} \to 0$ par limite de référence.
Ainsi

$$\lim_{x\to+\infty} x\mathrm{e}^{-\sqrt{x}} = 0$$

16.2.5.2 Limite en 0

Rappelons la limite de référence

$$\forall\alpha > 0,\ \lim_{x\to0^+} x^\alpha \ln x = 0$$

Exemple Etudions de

$$\lim_{x\to0^+} \left(\frac{1}{\sqrt{x}} + \ln x\right)$$

Quand $x \to 0^+$,
$\dfrac{1}{\sqrt{x}} \to +\infty$ et $\ln x \to -\infty$, il y a une forme indéterminée additive.

$$\frac{1}{\sqrt{x}} + \ln x = \frac{1}{\sqrt{x}}(1 + \sqrt{x}\ln x)$$

Puisque par limite de référence $\sqrt{x}\ln x \to 0$, on obtient $\dfrac{1}{\sqrt{x}} + \ln x \to +\infty$
Ainsi

$$\lim_{x\to 0^+}\left(\frac{1}{\sqrt{x}} + \ln x\right) = +\infty$$

Exemple Etudions

$$\lim_{x\to 0^+} |\ln x|^{1/\ln x}$$

Quand $x \to 0^+$,
$|\ln x| \to +\infty$ et $1/\ln x \to 0$, il y a une forme indéterminée.

$$|\ln x|^{1/\ln x} = \exp\left(\frac{\ln|\ln x|}{\ln x}\right)$$

Or $|\ln x| \to +\infty$ et on sait que $\dfrac{\ln t}{t} \xrightarrow[t\to+\infty]{} 0$ donc $\dfrac{\ln|\ln x|}{\ln x} \to 0$ puis $|\ln x|^{1/\ln x} \to 1$.
Ainsi

$$\lim_{x\to 0^+} |\ln x|^{1/\ln x} = 1$$

16.2.5.3 Limite en $a \in \mathbb{R}$

Pour $a \in \mathbb{R}$ et $\ell \in \bar{\mathbb{R}}$, on démontre par composition de limite l'équivalence suivante

$$f(x) \xrightarrow[x\to a]{} \ell \Leftrightarrow f(a+h) \xrightarrow[h\to 0]{} \ell$$

Par suite, il est très fréquent de ramener en 0 la détermination d'une limite en $a \in \mathbb{R}$ via le changement de variable $x = a + h$ pour lequel $h = x - a$.

Exemple Etudions

$$\lim_{x\to 1}\left(\frac{\ln x}{x-1}\right)$$

Quand $x \to 1$.
On pose $x = 1 + h$ de sorte que $h = x - 1 \to 0$.

$$\frac{\ln x}{x-1} = \frac{\ln(1+h)}{h} = \frac{\ln(1+h) - \ln 1}{h} \to (\ln)'(1) = 1$$

Ainsi

$$\lim_{x\to 1}\left(\frac{\ln x}{x-1}\right) = 1$$

Exemple Etudions

$$\lim_{x\to\pi/2}\frac{\sin(2x)}{\pi-2x}$$

Quand $x\to\pi/2$.
On pose $x=\pi/2+h$ de sorte que $h=x-\pi/2\to 0$.

$$\frac{\sin(2x)}{\pi-2x}=\frac{\sin(\pi+2h)}{-2h}=\frac{\sin(2h)}{2h}\to 1$$

car

$$\frac{\sin t}{t}=\frac{\sin t-\sin 0}{t}\xrightarrow[t\to 0]{}(\sin)'(0)=1$$

Ainsi

$$\lim_{x\to\pi/2}\frac{\sin(2x)}{\pi-2x}=1$$

16.2.6 Etude de limite par de comparaison

16.2.6.1 Théorème d'encadrement

Théorème

On suppose qu'au voisinage de a,

$$g(x)\leqslant f(x)\leqslant h(x)$$

Si $g\underset{a}{\to}\ell\in\mathbb{R}$ et $h\underset{a}{\to}\ell$ alors $f\underset{a}{\to}\ell$.

dém. :
Soit (u_n) une suite d'éléments de $\mathcal{D}$
Supposons $u_n\to a$.
Pour n assez grand, u_n appartient au voisinage de a sur lequel $g(x)\leqslant f(x)\leqslant h(x)$.
On a alors $g(u_n)\leqslant f(u_n)\leqslant h(u_n)$.
Or $g(u_n)\to\ell$ et $h(u_n)\to\ell$ donc par encadrement $f(u_n)\to\ell$.
En vertu de la caractérisation séquentielle des limites, on a $f\underset{a}{\to}\ell$.
□

Exemple Etudions

$$\lim_{x\to+\infty}\frac{E(x)}{x}$$

Quand $x\to+\infty$.
On a $x-1\leqslant E(x)\leqslant x$ donc

$$1-\frac{1}{x}\leqslant\frac{E(x)}{x}\leqslant 1$$

Or $1-1/x\to 1$ donc par encadrement $E(x)/x\to 1$.
Ainsi

$$\lim_{x\to+\infty}\frac{E(x)}{x}=1$$

16.2.6.2 Obtention de convergence

Théorème

> Si $|f(x) - \ell| \leqslant g(x)$ au voisinage de a et si $g \underset{a}{\to} 0$ alors $f \underset{a}{\to} \ell$.

dém. :
Soit (u_n) une suite d'éléments de $\mathcal{D}$
Supposons $u_n \to a$.
Pour n assez grand, u_n appartient au voisinage de a sur lequel

$$|f(x) - \ell| \leqslant g(x)$$

On a alors

$$|f(u_n) - \ell| \leqslant g(u_n)$$

Or $g(u_n) \to 0$ donc par comparaion $f(u_n) \to \ell$.
En vertu de la caractérisation séquentielle des limites, on a $f \underset{a}{\to} \ell$.
□

Exemple Etudions

$$\lim_{x \to +\infty} \frac{x + \sin x}{x}$$

Quand $x \to +\infty$,

$$\left|\frac{x + \sin x}{x} - 1\right| = \frac{|\sin x|}{x} \leqslant \frac{1}{x}$$

or $\frac{1}{x} \to 0$ donc $\frac{x + \sin x}{x} \to 1$.
Ainsi

$$\lim_{x \to +\infty} \frac{x + \sin x}{x} = 1$$

Proposition

> Si f est bornée au voisinage de a et si $g \underset{a}{\to} 0$ alors $fg \underset{a}{\to} 0$.

dém. :
Si f est bornée par M au voisinage de a alors au voisinage de a,

$$|f(x)g(x) - 0| \leqslant M\,|g(x)|$$

avec $M\,|g| \underset{a}{\to} 0$.
□

Exemple Etudions

$$\lim_{x \to 0} x\cos(1/x)$$

Quand $x \to 0$,
$x \to 0$ et $\cos(1/x)$ est borné donc $x\cos(1/x) \to 0$.
Ainsi

$$\lim_{x \to 0} x\cos(1/x) = 0$$

16.2.6.3 Obtention de limite infinie

Théorème

Supposons que $f(x) \leqslant g(x)$ au voisinage de a.
Si $f \underset{a}{\rightarrow} +\infty$ alors $g \underset{a}{\rightarrow} +\infty$.
Si $g \underset{a}{\rightarrow} -\infty$ alors $f \underset{a}{\rightarrow} -\infty$.

dém. :
Supposons $f \underset{a}{\rightarrow} +\infty$.
Soit (u_n) une suite d'éléments de $\mathcal{D}$
Supposons $u_n \rightarrow a$.
Pour n assez grand, u_n appartient au voisinage de a sur lequel $f(x) \leqslant g(x)$.
On a alors $f(u_n) \leqslant g(u_n)$.
Or $f(u_n) \rightarrow +\infty$ donc par comparaion $g(u_n) \rightarrow +\infty$.
En vertu de la caractérisation séquentielle des limites, on a $g \underset{a}{\rightarrow} +\infty$.
La deuxième implication s'obtient de façon semblable.

□

Exemple Etudions

$$\lim_{x \rightarrow +\infty} (x^2 + x \sin x)$$

Quand $x \rightarrow +\infty$,

$$x^2 + x \sin x \geqslant x^2 - x = x(x-1)$$

or $x(x-1) \rightarrow +\infty$ donc par comparaison $x^2 + x \sin x \rightarrow +\infty$.
Ainsi

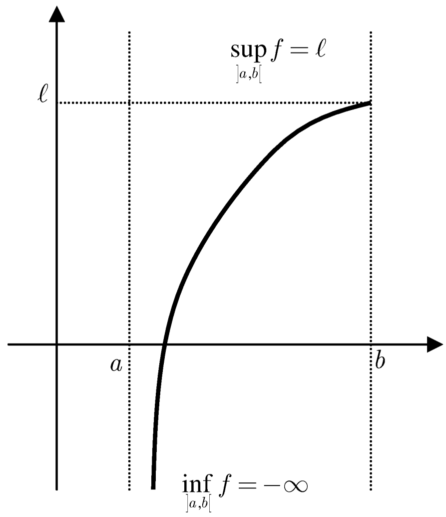
$$\lim_{x \rightarrow +\infty} (x^2 + x \sin x) = +\infty$$

16.2.7 Fonctions monotones sur un intervalle

Théorème

Soient $a, b \in \bar{\mathbb{R}}$ tels que $a < b$ et $f :]a,b[\to \mathbb{R}$.
Si f est croissante alors f admet des limites (éventuellement infinies) en a et b.
Plus précisément

$$\lim_a f = \inf_{]a,b[} f \text{ et } \lim_b f = \sup_{]a,b[} f$$



dém. :
Etude de la limite en b.
Cas f non majorée.
On a $\sup_{]a,b[} f = +\infty$.
Soit $M \in \mathbb{R}$, puisque f n'est pas majorée, il existe $x_0 \in]a,b[$ tel que $f(x_0) > M$.
Or f est croissante donc

$$\forall x \in [x_0, b[\,, f(x) > M$$

Ainsi, on peut affirmer qu'au voisinage de b, $f(x) \geqslant M$ ce qui établit $f \xrightarrow[b]{} +\infty$.
Cas f majorée.
osons $\ell = \sup_{]a,b[} f \in \mathbb{R}$.
Soit $\varepsilon > 0$, puisque $\ell - \varepsilon < \ell = \sup f$, $\ell - \varepsilon$ ne majore pas f et donc il existe $x_0 \in]a,b[$ tel que $f(x_0) \geqslant \ell - \varepsilon$.
Comme f est croissante, on a alors

$$\forall x \in [x_0, b[\,, f(x) \geqslant \ell - \varepsilon$$

D'autre part $f(x) \leqslant \ell$ car ℓ majore f et donc au voisinage de b, $|f(x) - \ell| \leqslant \varepsilon$.
Ainsi $f \xrightarrow[b]{} \ell$
L'étude en a est très similaire.
□

Corollaire

Si f est décroissante alors f admet des limites en a et b.
Plus précisément

$$\lim_a f = \sup_{]a,b[} f \text{ et } \lim_b f = \inf_{]a,b[} f$$

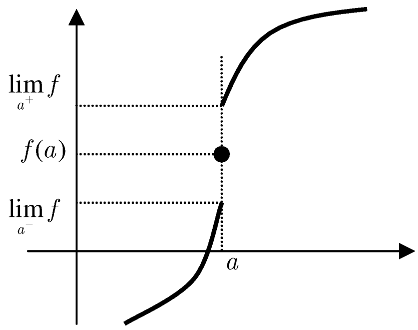
dém. :
Par passage à l'opposé.
□

Exemple $\sup_{x\in\mathbb{R}} \mathrm{e}^x = \lim_{x\to+\infty} \mathrm{e}^x = +\infty$ et $\inf_{x\in\mathbb{R}} \mathrm{e}^x = \lim_{x\to-\infty} \mathrm{e}^x = 0$.
Plus généralement, c'est ce résultat qui justifie que sup et inf se lisent sur un tableau de variation.

Corollaire

Soit I un intervalle non singulier et $a \in I$.
Si f est croissante alors f converge en a^- et en a^+ et on a

$$\lim_{a^-} f \leqslant f(a) \leqslant \lim_{a^+} f$$



dém. :
La restriction de f à $I \cap]-\infty, a[$ est croissante et majorée par $f(a)$ donc f converge en a^- et $\lim_{a^-} f \leqslant f(a)$
La restriction de f à $I \cap]a, +\infty[$ est croissante et minorée par $f(a)$ donc f converge en a^+ et $f(a) \leqslant \lim_{a^+} f$
□

16.3 Extension aux fonctions complexes

Les notions qui vont suivrent prolongent celles vues pour les fonctions réelles au cadre des fonctions complexes.

16.3.1 Définition

Définition

On appelle fonction complexe définie sur $\mathcal{D}$ toute application $f : \mathcal{D} \to \mathbb{C}$.
On note $\mathcal{F}(\mathcal{D}, \mathbb{C})$ l'ensemble de ces fonctions.

Exemple $t \mapsto \dfrac{t+i}{t-i}$ est une fonction complexe définie sur $\mathbb{R}$.

Exemple $t \mapsto \mathrm{e}^{it} = \cos t + i.\sin t$ est une fonction complexe définie sur $\mathbb{R}$.

Définition

Pour $f : \mathcal{D} \to \mathbb{C}$, on définit les fonctions $\mathrm{Re}(f) : \mathcal{D} \to \mathbb{R}$, $\mathrm{Im}(f) : \mathcal{D} \to \mathbb{R}$, $|f| : \mathcal{D} \to \mathbb{R}$ et $\bar{f} : \mathcal{D} \to \mathbb{C}$ par

$$\forall t \in \mathcal{D}, \mathrm{Re}(f)(t) = \mathrm{Re}(f(t)), \mathrm{Im}(f)(t) = \mathrm{Im}(f(t)),\ |f|(t) = |f(t)| \text{ et } \bar{f}(t) = \overline{f(t)}$$

Remarque En dehors des notions de minoration, de majoration et de monotonie, les notions relatives aux fonctions réelles s'étendent aux fonctions complexes.

Définition

Une fonction $f : \mathcal{D} \to \mathbb{C}$ est dite bornée s'il existe $M \in \mathbb{R}$ tel que $\forall t \in \mathcal{D}, |f(t)| \leqslant M$.

Exemple La fonction $t \mapsto \mathrm{e}^{it}$ est bornée par 1.

Proposition

$f : \mathcal{D} \to \mathbb{C}$ est bornée si, et seulement si, les fonctions $\mathrm{Re}(f)$ et $\mathrm{Im}(f)$ le sont.

dém. :
Puisque $|\mathrm{Re}(f)|, |\mathrm{Im} f| \leqslant |f|$, si f est bornée alors $\mathrm{Re}(f)$ et $\mathrm{Im}(f)$ le sont.
Inversement puisque $|f| \leqslant |\mathrm{Re}(f)| + |\mathrm{Im}(f)|$, si les fonctions $\mathrm{Re}(f)$ et $\mathrm{Im}(f)$ sont bornées alors f l'est aussi.
□

16.3.2 Limites

Soit $f : \mathcal{D} \to \mathbb{C}$ définie au voisinage de $a \in \bar{\mathbb{R}}$.

Définition

On dit que f tend vers $\ell \in \mathbb{C}$ en a si $|f(t) - \ell| \xrightarrow[t \to a]{} 0$.
On note alors $f \underset{a}{\to} \ell$ ou $f(t) \xrightarrow[t \to a]{} \ell$.

Exemple On a

$$\frac{t+i}{t-i} \xrightarrow[t \to +\infty]{} 1$$

En effet

$$\left|\frac{t+i}{t-i} - 1\right| = \frac{|2i|}{|t-i|} = \frac{2}{\sqrt{t^2+1}} \xrightarrow[t \to +\infty]{} 0$$

Définition

On dit que f converge en a s'il existe $\ell \in \mathbb{C}$ tel que $f \underset{a}{\to} \ell$.
Sinon, on dit que f diverge en a.

Théorème

Si f converge en a alors il existe un unique $\ell \in \mathbb{C}$ tel que $f \underset{a}{\to} \ell$.
Ce complexe ℓ est appelé limite de f en a et on note $\ell = \lim\limits_{t \to a} f(t)$.

dém. :
L'existence est acquise par définition de la convergence ; il reste à montrer l'unicité.
Supposons $f \underset{a}{\to} \ell$ et $f \underset{a}{\to} \ell'$.
On a alors

$$|\ell - \ell'| = |\ell - f(t) + f(t) - \ell'| \leqslant |\ell - f(t)| + |f(t) - \ell'|$$

En passant à la limite quand $t \to a$, on obtient $|\ell - \ell'| = 0$ et donc $\ell = \ell'$.
□

Remarque Pour les fonctions complexes, il n'y a pas de notion de limite infinie, tout au plus peut-on dire $|f(t)| \xrightarrow[t \to a]{} +\infty$.

16.3.3 Théorème liés à la convergence

Soient $f, g : \mathcal{D} \to \mathbb{C}$ définies au voisinage de $a \in \bar{\mathbb{R}}$.

Proposition

Si $f \underset{a}{\to} \ell \in \mathbb{C}$ alors $\bar{f} \underset{a}{\to} \bar{\ell}$ et $|f| \underset{a}{\to} |\ell|$.

dém. :
Supposons $f \underset{a}{\to} \ell$.
On a

$$|f(t) - \ell| \xrightarrow[t \to a]{} 0$$

Or

$$\left|\bar{f}(t) - \bar{\ell}\right| = \left|\overline{f(t) - \ell}\right| = |f(t) - \ell|$$

et

$$||f(t)| - |\ell|| \leqslant |f(t) - \ell|$$

donc $\left|\bar{f}(t) - \bar{\ell}\right| \underset{a}{\to} 0$ et $||f(t)| - |\ell|| \underset{a}{\to} 0$
□

Théorème

Si f converge en a alors f est bornée au voisinage de a.

dém. :
Supposons $f \underset{a}{\to} \ell \in \mathbb{C}$. On a alors $|f| \underset{a}{\to} |\ell|$, la fonction réelle $|f|$ est alors bornée au voisinage de a et donc majorée. Par suite f est bornée au voisinage de a.
□

Théorème

Si $f \underset{a}{\to} \ell \in \mathbb{C}$ et $g \underset{a}{\to} \ell' \in \mathbb{C}$ alors $f + g \underset{a}{\to} \ell + \ell'$, $fg \underset{a}{\to} \ell\ell'$ et si $\ell \neq 0$ alors $1/f \underset{a}{\to} 1/\ell$.

dém. :
Supposons $f \underset{a}{\to} \ell$ et $g \underset{a}{\to} \ell'$.
Puisque
$$|(f+g)(t) - (\ell + \ell')| \leqslant |f(t) - \ell| + |g(t) - \ell'|$$
on a
$$|(f+g)(t) - (\ell + \ell')| \underset{a}{\to} 0$$
Ainsi $f + g \underset{a}{\to} \ell + \ell'$.
Puisque
$$|(fg)(t) - \ell\ell'| = |f(t)|\,|g(t) - \ell'| + |\ell'|\,|f(t) - \ell|$$
avec $|f|$ majorée au voisinage de a, on a $|(fg)(t) - \ell\ell'| \underset{a}{\to} 0$. Ainsi $fg \underset{a}{\to} \ell\ell'$.
Enfin, si $\ell \neq 0$ alors au voisinage de a,
$$\left|\frac{1}{f(t)} - \frac{1}{\ell}\right| = \frac{|\ell - f(t)|}{|f(t)|\,|\ell|}$$
et donc
$$\left|\frac{1}{f(t)} - \frac{1}{\ell}\right| \underset{a}{\to} 0$$
Ainsi $1/f \underset{a}{\to} 1/\ell$.
□

Théorème

Soit $\ell \in \mathbb{C}$. On a équivalence entre :
(i) $f \underset{a}{\to} \ell$;
(ii) $\mathrm{Re}(f) \underset{a}{\to} \mathrm{Re}(\ell)$ et $\mathrm{Im}(f) \underset{a}{\to} \mathrm{Im}(\ell)$.

dém. :
(i) $\Rightarrow$ (ii) Car $|\mathrm{Re}(f)(t) - \mathrm{Re}(\ell)| \leqslant |f(t) - \ell|$ et $|\mathrm{Im}(f)(t) - \mathrm{Im}(\ell)| \leqslant |f(t) - \ell|$.
(ii) $\Rightarrow$ (i) Car $f(t) = \mathrm{Re}(f)(t) + i\mathrm{Im}(f)(t)$.
□

Exemple Etudions
$$\lim_{t \to +\infty} \frac{it}{t - i}$$
Quand $t \to +\infty$,
$$\frac{it}{t-i} = \frac{i}{1 - i/t} \to i$$
donc
$$\lim_{t \to +\infty} \frac{it}{t - i} = i$$

Exemple Etudions
$$\lim_{t \to +\infty} \frac{\mathrm{e}^{it} - 1}{t}$$
Quand $t \to +\infty$,
$$\left|\frac{\mathrm{e}^{it} - 1}{t}\right| \leqslant \frac{\left|\mathrm{e}^{it}\right| + 1}{t} = \frac{2}{t} \to 0$$

donc

$$\lim_{t\to+\infty}\frac{\mathrm{e}^{it}-1}{t}=0$$

Exemple Etudions

$$\lim_{t\to 0}\frac{\mathrm{e}^{it}-1}{t}$$

Quand $t\to 0$,

$$\frac{\mathrm{e}^{it}-1}{t}=\frac{\cos t-1}{t}+i\frac{\sin t-0}{t}\to(\cos)'(0)+i(\sin)'(0)=0+i\times 1=i$$

donc

$$\lim_{t\to 0}\frac{\mathrm{e}^{it}-1}{t}=i$$

16.3.4 Fonctions à valeurs dans $\mathbb{R}^2$

On note $d(a,b)$ la distance euclidienne entre deux couples éléments de $\mathbb{R}^2$.

Définition

On appelle fonction définie sur $\mathcal{D}$ à valeurs dans $\mathbb{R}^2$ toute application $f:\mathcal{D}\to\mathbb{R}^2$.
On note $\mathcal{F}(\mathcal{D},\mathbb{R}^2)$ l'ensemble de ces fonctions.

Exemple L'application $f:t\mapsto(\cos t,\dfrac{\sin t}{t})$ est une fonction à valeurs dans $\mathbb{R}^2$ définie sur $\mathbb{R}^\star$.

Définition

Soit $f:\mathcal{D}\to\mathbb{R}^2$. Pour tout $t\in\mathcal{D}$, on peut écrire $f(t)=(x(t),y(t))$.
Ceci introduit des fonctions réelles $x,y:\mathcal{D}\to\mathbb{R}$ appelées fonctions coordonnées de f.

Définition

Soit $f:\mathcal{D}\to\mathbb{R}^2$ définie au voisinage de $a\in\bar{\mathbb{R}}$
On dit que f tend vers $\ell\in\mathbb{R}^2$ en a si $d(f(t),\ell)\xrightarrow[t\to a]{}0$.
On dit alors que f converge vers ℓ en a et on montre aisément que cet élément ℓ est unique.
On l'appelle limite de f en a et on note $\ell=\lim\limits_a f$.

Proposition

Soient $f:\mathcal{D}\to\mathbb{R}^2$ définie au voisinage de $a\in\bar{\mathbb{R}}$ et $\ell=(\ell_x,\ell_y)\in\mathbb{R}^2$.
On a équivalence entre :
(i) $f\underset{a}{\to}\ell$;
(ii) $x(t)\xrightarrow[t\to a]{}\ell_x$ et $y(t)\xrightarrow[t\to a]{}\ell_y$.

dém. :
(i) $\Rightarrow$ (ii) Car $|x(t)-\ell_x|\leqslant d(f(t),\ell)$ et $|y(t)-\ell_y|\leqslant|d(f(t),\ell)|$.
(ii) $\Rightarrow$ (i) Car $|d(f(t),\ell)|\leqslant|x(t)-\ell_x|+|y(t)-\ell_y|$.
□

Exemple Considérons la fonction donnée par

$$f(t) = \left(\frac{\ln(1+t)}{t}, \frac{\sin t}{t}\right)$$

définie sur $]0, +\infty[$ et étudions ses limites en 0 et $+\infty$
Quand $t \to 0^+$,
D'une part

$$\frac{\ln(1+t)}{t} = \frac{\ln(1+t) - \ln 1}{t} \to (\ln)'(1) = 1$$

D'autre part

$$\frac{\sin(t)}{t} = \frac{\sin(t) - \sin(0)}{t} \to (\sin)'(0) = 1$$

Ainsi $f(t) \to (1, 1)$
Quand $t \to +\infty$,
D'une part

$$\frac{\ln(1+t)}{t} = \frac{1+t}{t}\frac{\ln(1+t)}{1+t}$$

or $\dfrac{1+t}{t} \to 1$ et $\dfrac{\ln y}{y} \xrightarrow[y \to +\infty]{} 0$ donc $\dfrac{\ln(1+t)}{t} \to 0$
D'autre part $\dfrac{\sin t}{t} \to 0$ car $\sin t$ est bornée et $\dfrac{1}{t} \to 0$.
Ainsi $f(t) \to (0, 0)$.
Finalement $\lim\limits_{0^+} f = (1, 1)$ et $\lim\limits_{+\infty} f = (0, 0)$.

16.4 Comparaison des fonctions numériques

Les fonctions considérées ici sont réelles ou complexes et supposées définies sur un même voisinage de $a \in \bar{\mathbb{R}}$.

16.4.1 Négligeabilité

Définition

Une fonction f est dite négligeable devant φ en a si on peut écrire

$$f(x) = \varphi(x)\varepsilon(x)$$

au voisinage de a avec $\varepsilon \underset{a}{\to} 0$.
On note alors

$$f = \underset{a}{o}(\varphi),\ f(x) = \underset{x \to a}{o}(\varphi(x)) \text{ ou } f(x) = o(\varphi(x)) \text{ quand } x \to a$$

On trouve aussi la notation $f(x) \ll \varphi(x)$ quand $x \to a$.

Proposition

Si la fonction φ ne s'annule pas au voisinage de a alors, on a équivalence entre
(i) $f(x) = o(\varphi(x))$ quand $x \to a$;
(ii) $f(x)/\varphi(x) \xrightarrow[x \to a]{} 0$.

dém. :
La fonction ε exprimant $f(x) = \varphi(x)\varepsilon(x)$ n'est autre que $f(x)/\varphi(x)$ au voisinage de a.
□

Remarque L'énoncé précédent reste vrai dans le cas où les fonctions f et φ sont définies en a et s'y annulent.

Exemple Quand $x \to +\infty$,

$$\ln x = o(x),\, x = o(x^2)\ x^2 = o(\mathrm{e}^x) \text{ et } \frac{1}{x^2} = o\left(\frac{1}{x}\right)$$

Exemple Quand $x \to 0$,

$$x^2 = o(x),\ \ln x = o\left(\frac{1}{x}\right) \text{ et } \frac{1}{x} = o\left(\frac{1}{x^2}\right)$$

Exemple Ecrire $f(x) = o(1)$ quand $x \to a \in \bar{\mathbb{R}}$ signifie que f converge vers 0 en a.
Il est fréquent d'écrire $f(x) = o(1)$ pour signifier que f est de limite nulle.
En particulier, si $f(x) \to \ell$ quand $x \to a$, alors on peut écrire $f(x) = \ell + o(1)$.

Attention : $f = \underset{a}{o}(h)$ et $g = \underset{a}{o}(h)$ n'impliquent pas $f = g$.
La relation $f = \underset{a}{o}(h)$ se comprend :
« la fonction f appartient à l'ensemble des fonctions négligeable devant h au voisinage de a ».
En fait, lors de la manipulation de $o(\ldots)$, les égalités ne doivent plus être considérées comme possédant la propriété de symétrie : $a = b \Rightarrow b = a$!

Proposition

Si $f = \underset{a}{o}(\lambda\varphi)$ alors $f = \underset{a}{o}(\varphi)$.
Si $f = \underset{a}{o}(\varphi)$ et $g = \underset{a}{o}(\varphi)$ alors $f + g = \underset{a}{o}(\varphi)$.
Si $f = \underset{a}{o}(\varphi)$ alors $fg = \underset{a}{o}(\varphi g)$.

dém. :
Supposons $f = \underset{a}{o}(\lambda\varphi)$.
Au voisinage de a, on peut écrire $f(x) = \lambda\varphi(x)\varepsilon(x)$ avec $\varepsilon \underset{a}{\to} 0$.
On a alors sur ce voisinage $f(x) = \varphi(x)\,(\lambda\varepsilon(x))$ avec $\lambda\varepsilon \underset{a}{\to} 0$. Ainsi $f = \underset{a}{o}(\varphi)$.
Supposons $f = \underset{a}{o}(\varphi)$ et $g = \underset{a}{o}(\varphi)$.
Au voisinage de a, on peut écrire $f(x) = \varphi(x)\varepsilon(x)$ avec $\varepsilon \underset{a}{\to} 0$ et $g(x) = \varphi(x)\eta(x)$ avec $\eta \underset{a}{\to} 0$.
Sur un voisinage commun, les deux relations précédentes sont valables et alors $(f + g)(x) = f(x) + g(x) = \varphi(x)(\varepsilon + \eta)(x)$ avec $\varepsilon + \eta \underset{a}{\to} 0$. Ainsi $f + g = \underset{a}{o}(\varphi)$.
Supposons $f = \underset{a}{o}(\varphi)$
Au voisinage de a, on peut écrire $f(x) = \varphi(x)\varepsilon(x)$ avec $\varepsilon \underset{a}{\to} 0$.
On a alors $(fg)(x) = f(x)g(x) = \varphi(x)g(x)\varepsilon(x)$ avec $\varepsilon \underset{a}{\to} 0$ donc $fg = \underset{a}{o}(\varphi g)$.
□

Remarque Dans la pratique, on exploite ces formules en écrivant les égalités :
- $\underset{a}{o}(\lambda\varphi) = \underset{a}{o}(\varphi)$;
- $\underset{a}{o}(\varphi) + \underset{a}{o}(\varphi) = \underset{a}{o}(\varphi)$;
- $g\,\underset{a}{o}(\varphi) = \underset{a}{o}(g\varphi)$.

Attention : $\underset{a}{o}(\varphi) - \underset{a}{o}(\varphi) = \underset{a}{o}(\varphi)$ et non $\underset{a}{o}(\varphi) - \underset{a}{o}(\varphi) = 0$.
On ne peut jamais simplifier les $o(\ldots)$.

Proposition

> En particulier, si $f = \underset{a}{o}(\varphi)$ et $\varphi = \underset{a}{o}(\psi)$ alors $f = \underset{a}{o}(\psi)$.

dém. :
Suppsons $f = \underset{a}{o}(\varphi)$ et $\varphi = \underset{a}{o}(\psi)$.
Au voisinage de a, on peut écrire $f(x) = \varphi(x)\varepsilon(x)$ avec $\varepsilon \underset{a}{\to} 0$ et $\varphi(x) = \psi(x)\eta(x)$ avec $\eta \underset{a}{\to} 0$.
Sur un voisinage commun, les deux relations précédentes sont valables et alors $f(x) = \psi(x)(\varepsilon\eta)(x)$ avec $\varepsilon\eta \underset{a}{\to} 0$. Ainsi $f = \underset{a}{o}(\psi)$.
□

Remarque Dans la pratique on écrit $\underset{a}{o}(\varphi) = \underset{a}{o}(\psi)$.
Ici, il convient de souligner que l'égalité n'est pas symétrique mais se comprend comme une transformation du premier membre en le second.

Exemple Quand $x \to 0$,
on peut écrire $o(x) - xo(x) = o(x) - o(x^2) = o(x) - o(x) = o(x)$.

16.4.2 Prépondérance

Définition

> On dit que f est dominée par φ en a s'il existe $M \in \mathbb{R}^+$ tel qu'au voisinage de a,
>
> $$|f(x)| \leqslant M\,|\varphi(x)|$$
>
> On note alors
>
> $$f = \underset{a}{O}(\varphi),\ f(x) = \underset{x\to a}{O}(\varphi(x)) \text{ ou } f(x) = O(\varphi(x)) \text{ quand } x \to a$$

Exemple $f = \underset{a}{o}(h) \Rightarrow f = \underset{a}{O}(h)$.
En effet si une fonction converge vers 0, elle est bornée au voisinage de 0.

Exemple $f = \underset{a}{O}(1)$ signifie que la fonction f est bornée au voisinage de a.
On peut par exemple écrire $\cos x = O(1)$ quand $x \to +\infty$ et $\sin(1/x) = O(1)$ quand $x \to 0$.

Proposition

Si la fonction φ ne s'annule pas au voisinage de a alors on a équivalence entre :
(i) $f(x) = O(\varphi(x))$ quand $x \to a$;
(ii) $f(x)/\varphi(x)$ est bornée quand $x \to a$.

dém. :
Si φ ne s'annule pas au voisinage de a alors l'inéquation $|f(x)| \leqslant M\,|\varphi(x)|$ équivaut à $|f(x)/\varphi(x)| \leqslant M$.
□

Remarque L'énoncé précédent reste vrai dans le cas où les fonctions f et φ sont définies en a et s'y annulent.

Proposition

Si $f = \underset{a}{O}(\lambda\varphi)$ alors $f = \underset{a}{O}(\varphi)$.
Si $f = \underset{a}{O}(\varphi)$ et $g = \underset{a}{O}(\varphi)$ alors $f + g = \underset{a}{O}(\varphi)$.
Si $f = \underset{a}{O}(\varphi)$ alors $fg = \underset{a}{O}(\varphi g)$.

dém. :
Supposons $f = \underset{a}{O}(\lambda\varphi)$.
Au voisinage de a, on peut écrire $|f(x)| \leqslant \lambda M\,|\varphi(x)|$ et alors $|f(x)| \leqslant M'\,|\varphi(x)|$ avec $M' = \lambda M$.
Supposons $f = \underset{a}{O}(\varphi)$ et $g = \underset{a}{O}(\varphi)$.
Au voisinage de a, on peut écrire $|f(x)| \leqslant M\,|\varphi(x)|$ et $|g(x)| \leqslant N\,|\varphi(x)|$.
Sur un voisinage commun, les deux relations précédentes sont valables et $|f(x) + g(x)| \leqslant (M + N)\,|\varphi(x)|$. Ainsi $f + g = \underset{a}{O}(\varphi)$.
Supposons $f = \underset{a}{O}(\varphi)$
Au voisinage de a, on peut écrire $|f(x)| \leqslant M\,|\varphi(x)|$.
On a alors $|(fg)(x)| = |f(x)|\,|g(x)| \leqslant M\,|\varphi(x)g(x)|$. Ainsi $fg = \underset{a}{O}(\varphi g)$.
□

Remarque Dans la pratique, on exploite ces formules en écrivant :
- $\underset{a}{O}(\lambda\varphi) = \underset{a}{O}(\varphi)$;
- $\underset{a}{O}(\varphi) + \underset{a}{O}(\varphi) = \underset{a}{O}(\varphi)$;
- $g\,\underset{a}{O}(\varphi) = \underset{a}{O}(g\varphi)$.

Proposition

Si $f = \underset{a}{o}(\varphi)$ et $\varphi = \underset{a}{O}(\psi)$ alors $f = \underset{a}{o}(\psi)$.
Si $f = \underset{a}{O}(\varphi)$ et $\varphi = \underset{a}{o}(\psi)$ alors $f = \underset{a}{o}(\psi)$.
Si $f = \underset{a}{O}(\varphi)$ et $\varphi = \underset{a}{O}(\psi)$ alors $f = \underset{a}{O}(\psi)$.

dém. :
Suppsons $f \underset{a}{=} o(\varphi)$ et $\varphi \underset{a}{=} O(\psi)$.
Au voisinage de a, on peut écrire $f(x) = \varphi(x)\varepsilon(x)$ avec $\varepsilon \underset{a}{\to} 0$ et $|\varphi(x)| \leqslant M\,|\psi(x)|$ au voisinage de a.
Posons

$$\eta(x) = \begin{cases} \varphi(x)\varepsilon(x)/\psi(x) & \text{si } \psi(x) \neq 0 \\ 0 & \text{si} \psi(x) = 0 \end{cases}$$

Au voisinage de a, on a $f(x) = \eta(x)\psi(x)$ que $\psi(x) \neq 0$ ou non car $\psi(x) = 0 \Rightarrow \varphi(x) = 0$.
Puisqu'au voisinage de a, $|\eta| \leqslant M\,|\varepsilon|$, on a $\eta \underset{a}{\to} 0$ et donc $f \underset{a}{=} o(\psi)$.
Supposons $f \underset{a}{=} O(\varphi)$ et $\varphi \underset{a}{=} o(\psi)$.
Au voisinage de a, on peut écrire $|f(x)| \leqslant M\,|\varphi(x)|$ et $\varphi(x) = \psi(x)\varepsilon(x)$ avec $\varepsilon \underset{a}{\to} 0$
Posons

$$\eta(x) = \begin{cases} f(x)\varepsilon(x)/\psi(x) & \text{si } \psi(x) \neq 0 \\ 0 & \text{si} \psi(x) = 0 \end{cases}$$

Au voisinage de a, on a $f(x) = \eta(x)\psi(x)$ que $\psi(x) \neq 0$ ou non car $\psi(x) = 0 \Rightarrow f(x) = 0$.
Puisqu'au voisinage de a, $|\eta| \leqslant M\,|\varepsilon|$, on a $\eta \underset{a}{\to} 0$ et donc $f \underset{a}{=} o(\psi)$.
Suppsons $f \underset{a}{=} O(\varphi)$ et $\varphi \underset{a}{=} O(\psi)$.
Au voisinage de a, on peut écrire $|f(x)| \leqslant M\,|\varphi(x)|$ et $|\varphi(x)| \leqslant M\,|\psi(x)|$.
Sur un voisinage commun, les deux relations précédentes sont valables et $|f(x)| \leqslant MN\,|\psi(x)|$ donc $f \underset{a}{=} O(\psi)$.
□

16.4.3 Croissance comparée des fonctions usuelles

16.4.3.1 Comparaison en $+\infty$

Proposition

> Quand $x \to +\infty$
> Pour $\alpha < \beta$, $x^\alpha = o(x^\beta)$ et $(\ln x)^\alpha = o((\ln x)^\beta)$.
> Pour $0 < a < b$, $a^x = o(b^x)$.

dém. :
Quand $x \to +\infty$, $\dfrac{x^\alpha}{x^\beta} = x^{\alpha-\beta} \to 0$ car $\alpha - \beta < 0$. Ainsi $x^\alpha = o(x^\beta)$.
De même, on obtient $(\ln x)^\alpha = o((\ln x)^\beta)$.
Quand $x \to +\infty$, $\dfrac{a^x}{b^x} = \mathrm{e}^{x(\ln a - \ln b)} \to 0$ car $\ln a - \ln b < 0$. Ainsi $a^x = o(b^x)$.
□

Proposition

> Quand $x \to +\infty$
> Pour $\alpha, \beta > 0$ et $a > 1$, $(\ln x)^\beta = o(x^\alpha)$ et $x^\alpha = o(a^x)$.

dém. :
Quand $x \to +\infty$,

$$\frac{(\ln x)^\beta}{x^\alpha} = \exp\left(\beta \ln(\ln x) - \alpha \ln x\right)$$

or

$$\beta \ln(\ln x) - \alpha \ln x = \ln x \left(-\alpha + \beta \frac{\ln(\ln x)}{\ln x}\right)$$

Puisque $\ln x \to +\infty$ et puisqu'on sait $\dfrac{\ln y}{y} \xrightarrow[y\to+\infty]{} 0$, on a $\dfrac{\ln(\ln x)}{\ln x} \to 0$ puis

$$\ln x\left(-\alpha+\beta\frac{\ln(\ln x)}{\ln x}\right)\to -\infty$$

et enfin $\dfrac{(\ln x)^\beta}{x^\alpha} \to 0$.
Ainsi $(\ln x)^\beta = o(x^\alpha)$.
Quand $x \to +\infty$,

$$\frac{x^\alpha}{a^x} = \exp\left(\alpha\ln x - x\ln a\right)$$

or

$$\alpha\ln x - x\ln a = x\left(-\ln a+\alpha\frac{\ln x}{x}\right)\to -\infty$$

et donc $\dfrac{x^\alpha}{a^x} \to 0$. Ainsi $x^\alpha = o(a^x)$.
□

Proposition

Quand $x \to +\infty$
Pour $\alpha,\beta > 0$ et $0 < a < 1$, $a^x = o\left(1/x^\alpha\right)$ et $1/x^\alpha = o(1/(\ln x)^\beta)$.

dém. :
Par passage à l'inverse, les comparaisons sont renversées !
□

Remarque Quand $x \to +\infty$, on peut écire

$$\mathrm{e}^{-x} \ll \frac{1}{x^2} \ll \frac{1}{x} \ll \frac{1}{\ln x} \ll 1 \ll \ln x \ll x \ll x^2 \ll \mathrm{e}^x$$

A titre d'exercice, on peut positionner entre les termes précédents les suivants :
$x\ln x, \dfrac{x}{\ln x}, \dfrac{\ln x}{x}, \dfrac{1}{x\ln x}, ...$

Remarque Via le changement de variable $y = -x$, on peut énoncer des règles semblable en $-\infty$.

16.4.3.2 Comparaison en 0

Proposition

Quand $x \to 0^+$, pour $\alpha < \beta$, $x^\beta = o(x^\alpha)$.
Quand $x \to 0^+$, pour $\alpha > 0$, $\ln x = o\left(\dfrac{1}{x^\alpha}\right)$ et $x^\alpha = o\left(\dfrac{1}{\ln x}\right)$.

dém. :
Quand $x \to 0$, $\dfrac{x^\beta}{x^\alpha} = x^{\beta-\alpha} \to 0$ car $\beta - \alpha > 0$. Ainsi $x^\beta = o(x^\alpha)$.
Quand $x \to 0^+$, $\dfrac{\ln x}{1/x^\alpha} = x^\alpha \ln x \to 0$ par limite de référence et donc $\ln x = o\left(\dfrac{1}{x^\alpha}\right)$.
De même $\dfrac{x^\alpha}{1/\ln x} = x^\alpha \ln x \to 0$ et donc $x^\alpha = o\left(\dfrac{1}{\ln x}\right)$.
□

Remarque Quand $x \to 0^+$, on peut écrire

$$x^2 \ll x \ll \sqrt{x} \ll \frac{1}{\ln x} \ll 1 \ll \ln x \ll \frac{1}{\sqrt{x}} \ll \frac{1}{x} \ll \frac{1}{x^2}$$

A titre d'exercice, on peut positionner entre les termes précédents les suivants : $\frac{x}{\ln x}$, $x \ln x$ et $\frac{\sqrt{x}}{\ln x}$.

Remarque Via le changement de variable $x = a + h$ on peut énoncer des règles similaires au voisinage de $a \in \mathbb{R}$.
Ainsi, $\forall \alpha < \beta, (x - a)^\beta = o((x - a)^\alpha)$ au voisinage de a.

16.4.4 Equivalence de fonctions

16.4.4.1 Définition

Définition

On dit que f est équivalente à g en a si on peut écrire au voisinage de a

$$f(x) = g(x)\theta(x)$$

avec $\theta \underset{a}{\to} 1$. On note alors

$$f \underset{a}{\sim} g,\ f(x) \underset{x \to a}{\sim} g(x) \text{ ou } f(x) \sim g(x) \text{ quand } x \to a$$

Proposition

Si la fonction g ne s'annule pas au voisinage de a alors on a équivalence entre :
(i) $f(x) \sim g(x)$ quand $x \to a$;
(ii) $f(x)/g(x) \xrightarrow[x \to a]{} 1$.

dém. :
La fonction θ exprimant $f(x) = g(x)\theta(x)$ n'est autre que $f(x)/g(x)$ au voisinage de a.
□

Exemple Quand $x \to +\infty$, $x^2 + x + 2\ln x \sim x^2$.
En effet

$$\frac{x^2 + x + 2\ln x}{x^2} = 1 + \frac{1}{x} + 2\frac{\ln x}{x^2} \to 1$$

Quand $x \to 0$, $x^2 + x + 2\ln x \sim 2\ln x$.
En effet

$$\frac{x^2 + x + 2\ln x}{2\ln x} = \frac{x^2}{2\ln x} + \frac{x}{2\ln x} + 1 \to 1$$

Exemple Quand $x \to +\infty$, $\sqrt{x + x^2} \sim x$,
Quand $x \to 0^+$, $\sqrt{x + x^2} \sim \sqrt{x}$
Quand $x \to -\infty$, $\sqrt{x + x^2} \sim -x$.

Exemple Si f converge vers $\ell \neq 0$ en a alors $f \underset{a}{\sim} \ell$

Attention : Ecrire $f \underset{a}{\sim} 0$ signifie que f est constante égale à la fonction nulle au voisinage de a.
Ecrire $f \underset{a}{\sim} +\infty$ n'a pas de sens.

Proposition

Si $f \underset{a}{\sim} g$ alors $g \underset{a}{\sim} f$.
Si $f \underset{a}{\sim} g$ et $g \underset{a}{\sim} h$ alors $f \underset{a}{\sim} h$.

dém. :
Supposons $f \underset{a}{\sim} g$.
On peut écrire $f(x) = g(x)\theta(x)$ au voisinage de a avec $\theta \underset{a}{\to} 1$.
La fonction θ n'est pas nulle au voisinage de a et donc on peut encore écrire $g(x) = f(x)\xi(x)$ avec $\xi = 1/\theta \underset{a}{\to} 1$. Ainsi $g \underset{a}{\sim} f$.
Supposons $f \underset{a}{\sim} g$ et $g \underset{a}{\sim} h$.
On peut écrire $f(x) = g(x)\theta(x)$ et $g(x) = h(x)\xi(x)$ au voisinage de a avec $\theta, \xi \underset{a}{\to} 1$.
Sur un voisinage commun, on peut alors écrire $f(x) = h(x)(\theta\xi)(x)$ avec $\theta\xi \underset{a}{\to} 1$.
Ainsi $f \underset{a}{\sim} h$.
□

Proposition

On a équivalence entre :
(i) $f \underset{a}{\sim} g$;
(ii) $f = g + \underset{a}{o}(g)$.

dém. :
(i) ⇒ (ii) Supposons $f \underset{a}{\sim} g$.
On peut écrire $f(x) = g(x)\theta(x)$ au voisinage de a avec $\theta \underset{a}{\to} 1$.
On a alors sur ce voisinage $f(x) = g(x) + g(x)(\theta(x) - 1)$ au voisinage de a et puisque $\theta(x) - 1 \xrightarrow[x \to a]{} 0$, la relation qui précède donne $f(x) = g(x) + o(g(x))$ quand $x \to a$.
(ii) ⇒ (i) Supposons $f = g + \underset{a}{o}(g)$.
On peut écrire $f(x) = g(x) + g(x)\varepsilon(x)$ au voisinage de a avec $\varepsilon \underset{a}{\to} 0$.
On a alors sur ce voisinage $f(x) = g(x)\theta(x)$ avec $\theta(x) = 1 + \varepsilon(x) \xrightarrow[x \to a]{} 1$.
Ainsi $f \underset{a}{\sim} g$.
□

Exemple Quand $x \to +\infty$, $\ln x + 2x = 2x + o(x) \sim 2x$.
Quand $x \to 0$, $\ln x + 2x = \ln x + o(x) \sim \ln x$.

Exemple Quand $x \to +\infty$, $\mathrm{e}^x + x^2 = \mathrm{e}^x + o(\mathrm{e}^x) \sim \mathrm{e}^x$.
Quand $x \to 0$, $\mathrm{e}^x + x^2 = 1 + o(1) \sim 1$.

16.4.4.2 Applications des équivalents

Théorème

> Si $f \underset{a}{\sim} g$ et si $g \underset{a}{\to} \ell \in \bar{\mathbb{R}}$ ou $\ell \in \mathbb{C}$ alors $f \underset{a}{\to} \ell$.

dém. :
On peut écrire $f(x) = g(x)\theta(x)$ au voisinage de a avec $\theta \underset{a}{\to} 1$.
Si $g \underset{a}{\to} \ell$ alors par produits de limites, $f \underset{a}{\to} \ell$.
□

Exemple $\lim\limits_{x\to+\infty}(x - x^2 + \ln x) = -\infty$ car $x - x^2 + \ln x = -x^2 + o(x^2) \sim -x^2 \to -\infty$.

Exemple $\lim\limits_{x\to 0+}\left(\ln x + \dfrac{1}{x}\right) = +\infty$ car $\ln x + \dfrac{1}{x} = \dfrac{1}{x} + o(\dfrac{1}{x}) \sim \dfrac{1}{x} \to +\infty$.

Remarque Inversement, supposons $f \underset{a}{\to} \ell$ et $g \underset{a}{\to} \ell$.
Si $\ell \in \mathbb{R}^\star$ ou $\mathbb{C}^\star$ alors $f \underset{a}{\sim} g$ car $f \underset{a}{\sim} \ell$ et $g \underset{a}{\sim} \ell$
Sinon, on ne peut rien dire !

Attention : $f \underset{a}{\to} 0$ et $g \underset{a}{\to} 0$ n'implique pas $f \underset{a}{\sim} g$.
$f \underset{a}{\to} +\infty$ et $g \underset{a}{\to} +\infty$ n'implique pas $f \underset{a}{\sim} g$.

Théorème

> Supposons que f et g soient des fonctions à valeurs réelles.
> Si $f \underset{a}{\sim} g$ alors $f(x)$ et $g(x)$ ont même signe au voisinage de a.

dém. :
On peut écrire $f(x) = g(x)\theta(x)$ au voisinage de a avec $\theta \underset{a}{\to} 1$.
Puisque $\theta \underset{a}{\to} 1 > 0$, au voisinage de a, on a $\theta(x) > 0$.
Sur un voisinage commun, $f(x) = g(x)\theta(x)$ avec $\theta(x) > 0$ donc $f(x)$ et $g(x)$ ont le même signe.
□

Exemple Quand $x \to 0^+$, $x - 2x^2 + x^5 = x + o(x) \sim x$ donc $x - 2x^2 + x^5 \to 0^+$.

Proposition

> Si $f \underset{a}{\sim} g$ alors $f = \underset{a}{O}(g)$ et $g = \underset{a}{O}(f)$.

dém. :
Supposons $f \underset{a}{\sim} g$.
On a $f = g + o(g) = O(g) + O(g) = O(g)$ et par symétrie $g = O(f)$.
□

Attention : La réciproque est fausse !

Proposition

Si $f \underset{a}{\sim} g$ et $g = \underset{a}{o}(h)$ alors $f = \underset{a}{o}(h)$.
Si $f \underset{a}{\sim} g$ et $g = \underset{a}{O}(h)$ alors $f = \underset{a}{O}(h)$.
Si $f = \underset{a}{o}(g)$ et $g \underset{a}{\sim} h$ alors $f = \underset{a}{o}(h)$.
Si $f = \underset{a}{O}(g)$ et $g \underset{a}{\sim} h$ alors $f = \underset{a}{O}(h)$.

dém. :
Supposons $f \underset{a}{\sim} g$ et $g = \underset{a}{o}(h)$.
On a alors $f = \underset{a}{O}(g)$ et $g = \underset{a}{o}(h)$ donc $f = \underset{a}{o}(h)$.
Les trois autres implications s'obtiennent de façon analogue.
□

Remarque Par les deux dernières propriétés, il est usuel lors d'expression de négligeabilité ou de domination de ne comparer qu'à des fonctions « simples ».
Par exemple, quand $x \to +\infty$, on écrit $o(x)$ au lieu de $o(x+1)$ car $x \sim x+1$.

16.4.4.3 Obtention d'équivalents

Attention : On ne peut pas sommer avec les équivalents :
$f \underset{a}{\sim} g$ n'entraîne pas $f + h \underset{a}{\sim} g + h$.
Pour déterminer l'équivalent d'une somme, on peut écrire celle-ci sous la forme $g + o(g) \sim g$.

Théorème

Si $f_1 \underset{a}{\sim} g_1$ et $f_2 \underset{a}{\sim} g_2$ alors

$$f_1 f_2 \underset{a}{\sim} g_1 g_2 \text{ et } f_1/f_2 \underset{a}{\sim} g_1/g_2$$

Si $f \underset{a}{\sim} g$ alors

$$\forall p \in \mathbb{Z}, f^p \underset{a}{\sim} g^p$$

dém. :
Supposons $f_1 \underset{a}{\sim} g_1$ et $f_2 \underset{a}{\sim} g_2$.
On peut écrire au voisinage de a, $f_1(x) = g_1(x)\theta_1(x)$ et $f_2(x) = g_2(x)\theta_2(x)$ avec $\theta_1, \theta_2 \underset{a}{\to} 1$.
On a alors $f_1(x)f_2(x) = g_1(x)g_2(x)\theta_1(x)\theta_2(x)$ et $\dfrac{f_1(x)}{f_2(x)} = \dfrac{g_1(x)}{g_2(x)}\dfrac{\theta_1(x)}{\theta_2(x)}$ avec $\theta_1\theta_2 \underset{a}{\to} 1$ et $\dfrac{\theta_1}{\theta_2} \underset{a}{\to} 1$.
Ainsi $f_1 f_2 \underset{a}{\sim} g_1 g_2$ et $f_1/f_2 \underset{a}{\sim} g_1/g_2$.
Supposons $f \underset{a}{\sim} g$.
On peut écrire au voisinage de a, $f(x) = g(x)\theta(x)$ avec $\theta \underset{a}{\to} 1$.
On a alors $[f(x)]^p = [g(x)]^p [\theta(x)]^p$ avec $\theta^p \underset{a}{\to} 1$ donc $f^p \underset{a}{\sim} g^p$
□

Théorème

On suppose que f et g sont à valeurs réelles strictement positives.
Si $f \underset{a}{\sim} g$ alors
$$\forall \alpha \in \mathbb{R}, f^\alpha \underset{a}{\sim} g^\alpha$$

dém. :
Supposons $f \underset{a}{\sim} g$.
On peut écrire au voisinage de a, $f(x) = g(x)\theta(x)$ avec $\theta \underset{a}{\to} 1$.
On a alors $[f(x)]^\alpha = [g(x)]^\alpha [\theta(x)]^\alpha$ avec $\theta^\alpha \underset{a}{\to} 1$ donc $f^\alpha \underset{a}{\sim} g^\alpha$
□

Attention : Ce résultat ne vaut que pour un exposant α constant (c'est-à-dire qui ne dépend pas de x).

Exemple Quand $x \to +\infty$, déterminons un équivalent simple de
$$\frac{\sqrt{x^3+x}}{\sqrt[3]{x^2+1}}$$
$x^3 + x = x^3 + o(x^3) \sim x^3$ donc $\sqrt{x^3+x} \sim x^{3/2}$.
$x^2 \to +\infty$ et 1 est constant donc $x^2 + 1 = x^2 + o(x^2) \sim x^2$ puis $\sqrt[3]{x^2+1} \sim x^{2/3}$.
Enfin
$$\frac{\sqrt{x^3+x}}{\sqrt[3]{x^2+1}} \sim \frac{x^{3/2}}{x^{2/3}} = x^{5/6}$$

Proposition

Si $f \underset{a}{\to} 0$ alors
$$\sin f \underset{a}{\sim} f \text{ et } \ln(1+f) \underset{a}{\sim} f$$

dém. :
Puisqu'on sait $\dfrac{\sin x}{x} = \dfrac{\sin x - \sin 0}{x} \xrightarrow[x \to 0, x \neq 0]{} (\sin)'(0) = 1$, on peut écrire, pour $x \neq 0$, $\sin x = x\theta(x)$ avec $\theta \xrightarrow[x\to 0]{} 1$. En posant $\theta(0) = 1$, l'égalité qui précède vaut encore pour $x = 0$.
Si $f \underset{a}{\to} 0$ alors $\sin(f(x)) = f(x)\xi(x)$ avec $\xi(x) = \theta(f(x)) \xrightarrow[x\to a]{} 1$. Ainsi $\sin f \underset{a}{\sim} f$.
De même, en exploitant $\dfrac{\ln(1+x)}{x} \xrightarrow[x\to 0]{} 1$, on obtient $\ln(1+f) \underset{a}{\sim} f$.
□

Exemple En particulier $\sin x \sim x$ et $\ln(1+x) \sim x$ quand $x \to 0$.

Exemple Déterminons un équivalent simple à $\sqrt{x^3+1}\sin\dfrac{1}{x}$ quand $x \to +\infty$.
D'une part $x^3 + 1 = x^3 + o(x^3) \sim x^3$ donc $\sqrt{x^3+1} \sim x^{3/2}$.

D'autre part $1/x \to 0$ donc $\sin(1/x) \sim 1/x$.
Par produit, on obtient

$$\sqrt{x^3+1}\sin\frac{1}{x} \sim \sqrt{x}$$

Exemple Déterminons un équivalent simple à $\sqrt{\ln(1+x^2)}$ quand $x \to 0$.
Puisque $x^2 \to 0$, $\ln(1+x^2) \sim x^2$ puis

$$\sqrt{\ln(1+x^2)} \sim \sqrt{x^2} = |x|$$

Exemple Déterminons un équivalent simple à $1-\cos x$ quand $x \to 0$.
Ecrire $\cos x \sim 1$ n'est pas pertinent ici, nous allons transformer l'expression pour ob
On sait

$$1-\cos x = 2\sin^2\frac{x}{2}$$

Puisque $x/2 \to 0$, on a $\sin(x/2) \sim x/2$ puis

$$1-\cos x \sim \frac{x^2}{2}$$

Exemple Déterminons un équivalent simple à $\ln(1+x)+x$ quand $x \to 0$.
Ici, il faut prendre garde à ne pas rédiger une somme d'équivalents, pour cela on va manipuler des $o(\ldots)$.
Puisque $x \to 0$, $\ln(1+x) \sim x$ et donc $\ln(1+x) = x + o(x)$.
Par suite

$$\ln(1+x)+x = 2x+o(x) = 2x+o(2x) \sim 2x$$

Remarque Pour étudier un équivalent à $\ln(1+x)-x$ quand $x \to 0$, on ne peut guère dire mieux que

$$\ln(1+x)-x = x+o(x)-x = o(x)$$

Dans un chapitre ultérieur, on pourra résoudre ce problème à l'aide de développements limités.

Proposition

> On suppose que f et g sont valeurs strictement positives.
> Si $f \underset{a}{\sim} g$ et si $g \underset{a}{\to} \ell$ avec $\ell = 0^+, \ell \in \mathbb{R}^{+\star} \setminus \{1\}$ ou $\ell = +\infty$ alors $\ln f \underset{a}{\sim} \ln g$.

dém. :
Puisque $f \underset{a}{\sim} g$, on peut écrire $f(x) = g(x)\theta(x)$ au voisinage de a avec $\theta \underset{a}{\to} 1$.
On a alors $\ln f(x) = \ln g(x) + \ln\theta(x)$. Or $\ln\theta(x) \xrightarrow[x\to a]{} 1$ et $\ln g(x) \xrightarrow[x\to a]{} -\infty, \ln\ell \neq 0$ ou $+\infty$.
Dans chaque cas, $\ln\theta(x) = o(\ln g(x))$ et donc $\ln f(x) \sim \ln g(x)$ quand $x \to a$.
□

Exemple Quand $x \to +\infty$,

$$\ln(1 + x + 2x^2) \sim 2 \ln x$$

car $1 + x + 2x^2 \sim 2x^2 \to +\infty \neq 1$ donc

$$\ln(1 + x + 2x^2) \sim \ln(2x^2) = 2 \ln x + \ln 2 \sim 2 \ln x$$

Remarque Pour déterminer un équivalent de $\ln f$ quand $f \underset{a}{\to} 1$, il suffit d'écrire $f = 1 + g$ car alors $g \underset{a}{\to} 0$ et donc $\ln f = \ln(1 + g) \underset{a}{\sim} g$.

Exemple Quand $x \to 0$,

$$\ln(1 + x + 2x^2) \sim x + 2x^2 \sim x$$

Attention : On ne peut pas passer les équivalents à l'exponentielle.

Chapitre 17

Continuité

17.1 Continuité des fonctions réelles

17.1.1 Définition

Définition

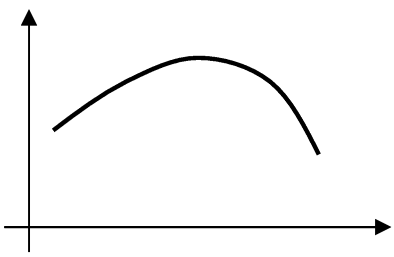
On dit qu'une fonction $f : \mathcal{D} \to \mathbb{R}$ est continue en un point $a \in \mathcal{D}$ si $f(x) \xrightarrow[x \to a]{} f(a)$.
Sinon, on dit que f est discontinue en a.

Définition

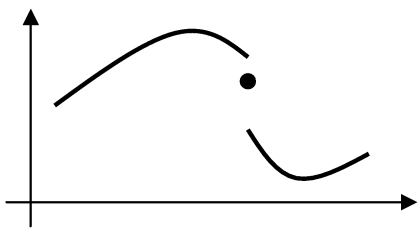
On dit que f est continue si f est continue en chaque $a \in \mathcal{D}$.
On note $\mathcal{C}(\mathcal{D}, \mathbb{R})$ l'ensemble des fonctions réelles définies et continues sur $\mathcal{D}$.

Remarque Graphiquement, une fonction continue est une fonction qui peut être représentée sans lever le crayon là où elle est définie.

Graphe d'une fonction continue



Graphe d'une fonction discontinue



Remarque Par définition

$$f : \mathcal{D} \to \mathbb{R} \text{ est continue } \Leftrightarrow \forall a \in \mathcal{D}, f(x) \xrightarrow[x \to a]{} f(a)$$

Cela signifie encore

$$\forall a \in \mathcal{D}, \forall \varepsilon > 0, \exists \alpha > 0, \forall x \in \mathcal{D}, |x - a| \leqslant \alpha \Rightarrow |f(x) - f(a)| \leqslant \varepsilon$$

Exemple Les fonctions constantes sont continues.

Exemple La fonction $x \mapsto x$ est continue sur $\mathbb{R}$, la fonction $x \mapsto |x|$ est continue sur $\mathbb{R}$ et plus généralement les fonctions usuelles : $\sin, \cos, \tan, \exp, \ln, x \mapsto \sqrt{x}, \ldots$ sont continues sur leur domaine de définition.

Exemple La fonction indicatrice de $\mathbb{Q}$ est totalement discontinue.

Attention : Nous verrons que dérivabilité entraîne la continuité. Cependant la continuité n'entraîne pas la dérivabilité
Par exemple, les fonction $x \mapsto \sqrt{x}$ et $x \mapsto |x|$ sont continues alors que non dérivables en 0 !

17.1.2 Opérations

Théorème

Soient $f, g : \mathcal{D} \to \mathbb{R}$ et $\lambda \in \mathbb{R}$.
Si f et g sont continues alors $\lambda.f$, $f + g$ et fg le sont aussi.
De plus, si g ne s'annule pas alors f/g est continue.

dém. :
Soit $a \in \mathcal{D}$.
Par continuité de f en a, on a $f(x) \xrightarrow[x \to a]{} f(a)$.
Par opérations sur les limites $(\lambda.f)(x) = \lambda f(x) \xrightarrow[x \to a]{} \lambda f(a) = (\lambda.f)(a)$.
Ainsi $\lambda.f$ est continue en a et puisque ceci vaut pour tout $a \in \mathcal{D}$, on peut affirmer que $\lambda.f$ est continue.
L'obtention des continuités de $f + g$, fg et f/g est semblable.
□

Exemple Par somme et produit de fonctions continues, les fonctions polynomiales sont continues sur $\mathbb{R}$.
Par rapport de fonctions continues, les fonctions rationnelles sont continues là où elles sont définies.

Théorème

Soient $f : \mathcal{D} \to \mathbb{R}$ et $\varphi : \mathcal{D}' \to \mathbb{R}$ telles que $f(\mathcal{D}) \subset \mathcal{D}'$.
Si f et φ sont continues alors $\varphi \circ f$ l'est aussi.

dém. :
Soit $a \in \mathcal{D}$.
Par continuité de f en a, on a $f(x) \xrightarrow[x \to a]{} f(a)$.
Or $f(a) \in \mathcal{D}'$ et φ est continue sur $\mathcal{D}'$ donc $\varphi(y) \xrightarrow[y \to f(a)]{} \varphi(f(a))$.
Par composition de limites, on obtient $(\varphi \circ f)(x) = \varphi(f(x)) \xrightarrow[x \to a]{} \varphi(f(a)) = (\varphi \circ f)(a)$.
□

Exemple La fonction
$$x \mapsto \frac{\mathrm{e}^{\sqrt{x^2+1}}}{\sin^2(x) + 1}$$
est continue par opérations sur les fonctions continues !

Proposition

Si $f, g : \mathcal{D} \to \mathbb{R}$ sont continues alors les fonctions $|f|$, $\sup(f, g)$ et $\inf(f, g)$ le sont aussi.

dém. :
La fonction $|f|$ est continue par composition de fonction continues.
Par définition $\forall x \in \mathcal{D}, \sup(f, g)(x) = \max(f(x), g(x))$.
Or il est remarquable que pour $a, b \in \mathbb{R}$, on a $\max(a, b) = \dfrac{1}{2}(a + b + |a - b|)$.
Par suite

$$\sup(f, g) = \frac{1}{2}\left(f + g + |f - g|\right)$$

et cette fonction est donc continue par opérations sur les fonctions continues.
De même

$$\inf(f, g) = \frac{1}{2}\left(f + g - |f - g|\right)$$

est continue par opérations sur les fonctions continues.
□

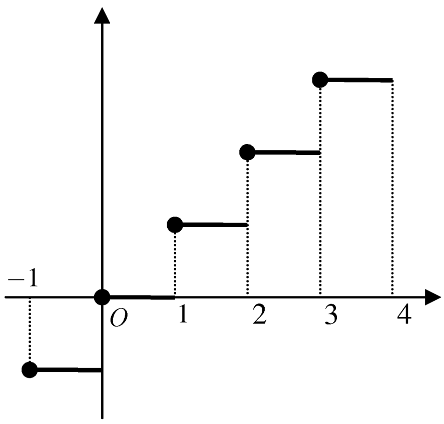
Exemple Si f est continue alors les fonctions $f^+ = \sup(f, 0)$ et $f^- = \sup(-f, 0)$ sont continues.

17.1.3 Continuité à droite, continuité à gauche

Définition

Soit $f : \mathcal{D} \to \mathbb{R}$ et $a \in \mathcal{D}$.
Si f est définie à droite de a, on dit que f est continue à droite en a si $f \xrightarrow[a^+]{} f(a)$.
Si f est définie à gauche de a, on dit que f est continue à gauche en a si $f \xrightarrow[a^-]{} f(a)$.

Exemple La fonction partie entière $x \mapsto E(x)$ est continue à droite en tout $a \in \mathbb{R}$.



Proposition

Soit $f : \mathcal{D} \to \mathbb{R}$ et a un point de $\mathcal{D}$ tel que f soit définie à droite et à gauche de a.
On a équivalence entre :
(i) f est continue en a ;
(ii) f est continue à droite et à gauche en a.

dém. :
(i) $\Rightarrow$ (ii) Immédiat.
(ii) $\Rightarrow$ (i) C'est un point immédiat que une limite à droite ou à gauche en a ne prend pas en compte la valeur de la fonction en a...
Supposons que $f(t) \xrightarrow[t \to a^+]{} f(a)$ et $f(t) \xrightarrow[t \to a^-]{} f(a)$.
Pour $\varepsilon > 0$, il existe $\alpha^+ > 0$ et $\alpha^- > 0$ tel que :

$$\forall t \in \mathcal{D} \cap \left]-\infty, a\right[, |t-a| \leqslant \alpha^- \Rightarrow |f(t) - f(a)| \leqslant \varepsilon$$

et

$$\forall t \in \mathcal{D} \cap \left]a, +\infty\right[, |t-a| \leqslant \alpha^+ \Rightarrow |f(t) - f(a)| \leqslant \varepsilon$$

En posant $\alpha = \min(\alpha^+, \alpha^-) > 0$, on a

$$\forall t \in \mathcal{D} \setminus \{a\}, |t-a| \leqslant \alpha \Rightarrow |f(t) - f(a)| \leqslant \varepsilon$$

puis immédiatement

$$\forall t \in \mathcal{D}, |t-a| \leqslant \alpha \Rightarrow |f(t) - f(a)| \leqslant \varepsilon$$

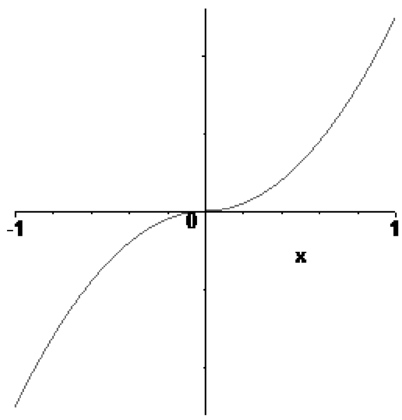
Ainsi f est continue en a.
□

Remarque Cet outil est utile pour obtenir la continuité d'une fonction définie par une alternative.

Exemple Soit $f : \mathbb{R} \to \mathbb{R}$ la fonction définie par

$$f(x) = \begin{cases} x^2 & \text{si } x \geqslant 0 \\ -x^2 & \text{sinon} \end{cases}$$

Montrer que f est continue sur $\mathbb{R}$.
1ère méthode :
Soit $a \in \mathbb{R}$
Cas $a > 0$ alors au voisinage de a, $f(x) = x^2$ et donc f est continue en a.
Cas $a < 0$ alors au voisinage de a, $f(x) = -x^2$ et à nouveau f est continue en a.
Il reste le cas $a = 0$.
Quand $x \to 0^+$, $f(x) = x^2 \to 0 = f(0)$ et quand $x \to 0^-$, $f(x) = -x^2 \to 0 = f(0)$.
Par continuité à droite et à gauche en 0, f est continue en 0.
2ème méthode :
On a $\forall x \geqslant 0, f(x) = x^2$ donc $f_{\restriction [0,+\infty[}$ est continue et par suite f est continue en chaque point de $\mathbb{R}^{+\star}$ et continue à droite en 0.
Aussi, $\forall x \leqslant 0, f(x) = -x^2$ donc $f_{\restriction]-\infty,0]}$ est continue et par suite f est continue en chaque point de $\mathbb{R}^{-\star}$ et continue à gauche en 0.
3ème méthode :
$\forall x \in \mathbb{R}, f(x) = x\,|x|$ donc f est continue par opération sur les fonctions continues.



17.1.4 Restrictions et prolongement de fonctions continues

Proposition

Soient $f : \mathcal{D} \to \mathbb{R}$ et $\Delta \subset \mathcal{D}$.
Si f est continue alors $f_{\restriction \Delta}$ l'est aussi.

dém. :
Soit $a \in \Delta$.
Par continuité de f en a, on a la propriété :

$$\forall \varepsilon > 0, \exists \alpha > 0, \forall x \in \mathcal{D}, |x - a| \leqslant \alpha \Rightarrow |f(x) - f(a)| \leqslant \varepsilon$$

Puisque $\Delta \subset \mathcal{D}$, on a encore

$$\forall \varepsilon > 0, \exists \alpha > 0, \forall x \in \Delta, |x - a| \leqslant \alpha \Rightarrow |f(x) - f(a)| \leqslant \varepsilon$$

Or pour $x \in \Delta$ on a $|f(x) - f(a)| = |f_{\restriction \Delta}(x) - f_{\restriction \Delta}(a)|$ et donc on obtient la continuité de $f_{\restriction \Delta}$ en a.
□

Définition

Soit $f : \mathcal{D} \to \mathbb{R}$ continue et $\tilde{\mathcal{D}}$ une partie de $\mathbb{R}$ contenant $\mathcal{D}$.
On appelle prolongement par continuité de f à $\tilde{\mathcal{D}}$ toute application $\tilde{f} : \tilde{\mathcal{D}} \to \mathbb{R}$ telle que :
1) $\tilde{f}$ prolonge f ;
2) $\tilde{f}$ est continue sur $\tilde{\mathcal{D}}$.

Proposition

Soit $f : \mathcal{D} \to \mathbb{R}$ continue et $a \in \mathbb{R} \backslash \mathcal{D}$ tel que f soit définie au voisinage de a.
Si $f \underset{a}{\to} \ell \in \mathbb{R}$ alors la fonction $\tilde{f} : \mathcal{D} \cup \{a\} \to \mathbb{R}$ définie par

$$f(x) = \begin{cases} f(x) & \text{si } x \neq a \\ \ell & \text{si } x = a \end{cases}$$

est l'unique prolongement par continuité de f au domaine $\mathcal{D} \cup \{a\}$.

dém. :
Unicité :
Si $\tilde{f}$ prolonge f par continuité à $\mathcal{D} \cup \{a\}$ alors pour tout $x \in \mathcal{D}$, $\tilde{f}(x) = f(x)$ et pour $x = a$,

$$\tilde{f}(a) = \lim_{x \to a} \tilde{f}(x) = \lim_{x \to a, x \in \mathcal{D}} \tilde{f}(x) = \lim_{x \to a} f(x) = \ell$$

Ainsi $\tilde{f}$ est déterminé de façon unique.
Existence :
La fonction proposée prolonge bien évidemment la fonction f à $\mathcal{D} \cup \{a\}$; il reste à voir qu'elle est continue.
Pour tout $x \in \mathcal{D}$, $\tilde{f}$ est continue en x car coïncide avec la fonction continue f au voisinage de x.
Pour $x = a$, on a $\tilde{f}(a) = \ell = \lim\limits_{x \to a} f(x) = \lim\limits_{x \to a, x \in \mathcal{D}} \tilde{f}(x)$ et puisque la valeur de $\tilde{f}$ en a est justement ℓ, on a encore $\tilde{f}(a) = \lim\limits_{x \to a, x \in \mathcal{D} \cup \{a\}} \tilde{f}(x)$ et donc $\tilde{f}$ est continue en a.
□

Exemple Soit $f :]0, +\infty[\to \mathbb{R}$ définie par

$$f(x) = x \ln x$$

f est continue sur $\mathbb{R}^{+\star}$ et $\lim\limits_{x \to 0^+} f(x) = 0$.
On prolonge f par continuité en 0 en posant $f(0) = 0$.

Exemple Soit $f : \mathbb{R}^\star \to \mathbb{R}$ définie par

$$f(x) = \frac{\sin x}{x}$$

f est continue sur $\mathbb{R}^\star$ et $\lim\limits_{x \to 0} f(x) = 1$.
On prolonge f par continuité en 0 en posant $f(0) = 1$.
La fonction obtenue est appelée fonction sinus cardinal.

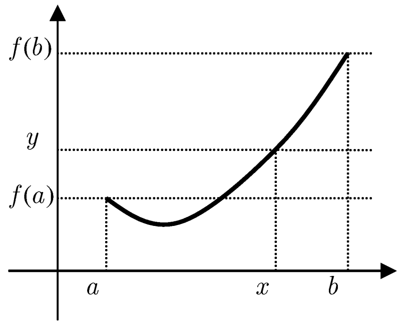
17.2 Fonctions continue sur un intervalle

Ici I désigne un intervalle de $\mathbb{R}$.

17.2.1 Théorème des valeurs intermédiaires

Théorème

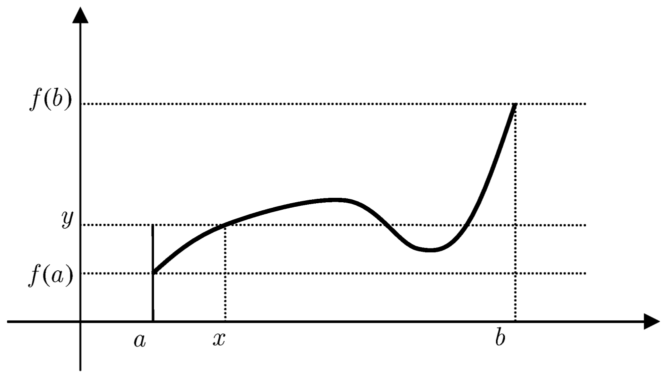
Soit $f : I \to \mathbb{R}$ et $a, b \in I$ tels que $a \leqslant b$.
Si f est continue alors f prend toutes les valeurs intermédiaires comprises entre $f(a)$ et $f(b)$.



dém. :
Quitte à considérer $-f$ et passer le problème à l'opposé on peut supposer $f(a) \leqslant f(b)$.
Soit $y \in [f(a), f(b)]$. On désire établir l'existence de $x \in [a, b]$ vérifiant $f(x) = y$.
1ère méthode : Par une borne supérieure...
Soit

$$A = \{c \in [a, b] \, / \forall t \in [a, c] \, , f(t) \leqslant y\}$$

La partie A est une partie de $\mathbb{R}$ non vide (car $a \in A$) et majorée (par b) ; elle admet donc une borne supérieure que nous notons x. Montrons $f(x) = y$.
Pour tout $t < x$, il existe $c \in A$ tel que $t < c$ car t n'est pas majorant de A. Par suite $t \in [a, c]$ avec $c \in A$ et donc $f(t) \leqslant y$. En passant cette relation à la limite quand $t \to x^-$, on obtient $f(x) \leqslant y$.
Supposons maintenant par l'absurde $f(x) < y$.
Par continuité de f en x, la fonction f est strictement inférieure à y au voisinage de x. Cet x est forcément strictement inférieur à b et il existe $\alpha > 0$ tel que $c = x + \alpha \in [a, b]$ et $\forall t \in [a, c] = [a, x] \cup [x, x + \alpha] \, , f(t) \leqslant y$.
Ainsi $x + \alpha \in A$ ce qui contredit la définition de x comme majorant de A.
Au final $f(x) = y$

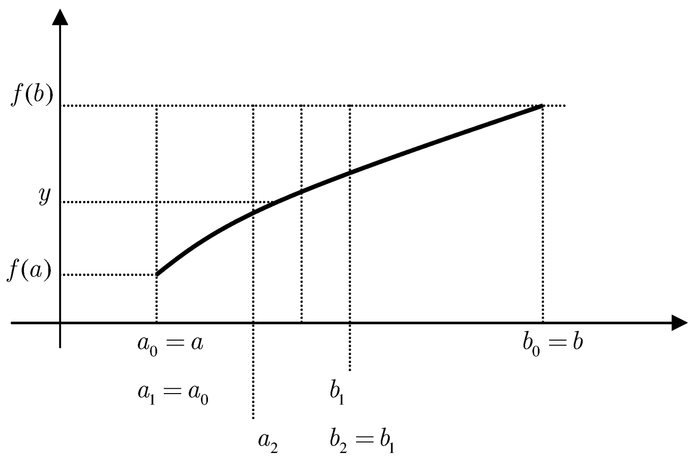


2ème méthode : Par dichotomie
On va construire par un procédé dichotomique deux suites adjacentes (a_n) et (b_n) dont la limite commune sera un x tel que $y = f(x)$.
Etape 0 :
On pose $a_0 = a$ et $b_0 = b$; on vérifie $f(a_0) \leqslant y \leqslant f(b_0)$.
Etape 1 :
On introduit $d = (a_0 + b_0)/2$.

Si $f(d) \leqslant y$ on pose $a_1 = d$ et $b_1 = b_0$.
Si $f(d) > y$ on pose $a_1 = a_0$ et $b_1 = d$.
Dans les deux cas, on vérifie : $f(a_1) \leqslant y \leqslant f(b_1)$
...
Etape $n+1$:
On introduit $d = (a_n + b_n)/2$.
Si $f(d) \leqslant y$ on pose $a_{n+1} = d$ et $b_{n+1} = b_n$.
Si $f(d) > y$ on pose $a_{n+1} = a_n$ et $b_{n+1} = d$.
Dans les deux cas, on vérifie : $f(a_{n+1}) \leqslant y \leqslant f(b_{n+1})$.
...
On définit ainsi deux suites (a_n) et (b_n) et par construction on vérifie :

$$\forall n \in \mathbb{N}, a_{n+1} \geqslant a_n, b_{n+1} \leqslant b_n \text{ et } b_{n+1} - a_{n+1} = \frac{b_n - a_n}{2}$$

La suite (a_n) est croissante, la suite (b_n) est décroissante et $b_n - a_n = \dfrac{b-a}{2^n} \to 0$ donc les suites (a_n) et (b_n) sont adjacentes. Notons x leur limite commune.
Puisque $a_n \to x$ et $b_n \to x$ on a $f(a_n) \to f(x)$ et $f(b_n) \to f(x)$.
Or pour tout $n \in \mathbb{N}$, $f(a_n) \leqslant y \leqslant f(b_n)$ donc en passant à la limite $f(x) \leqslant y \leqslant f(x)$ puis $y = f(x)$.



□

Remarque La méthode par dichotomie présentée dans la démonstration permet de déterminer numériquement une valeur approchée d'une solution x à l'équation $f(x) = y$.

Remarque Le théorème des valeurs intermédiaires assure l'existence de x mais n'évoque pas son unicité qui d'ailleurs n'est pas toujours vraie.
Un argument d'injectivité, par exemple par stricte monotonie, permet d'obtenir l'éventuelle unicité de x.

Corollaire

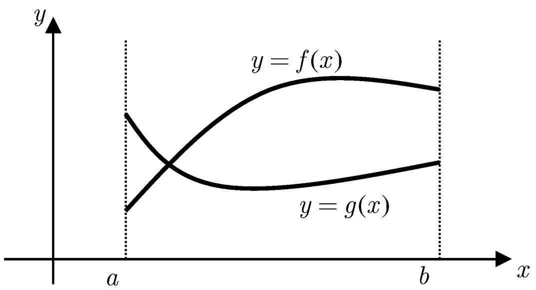
> Si une fonction continue $f : I \to \mathbb{R}$ prend une valeur positive et une valeur négative alors f s'annule.

dém. :
0 est valeur intermédiaire entre deux valeurs prises

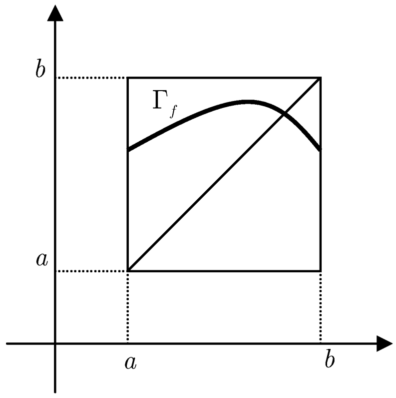
□

Exemple Soient $a, b, c \in \mathbb{R}$.
Montrer que l'équation $x^3 + ax^2 + bx + c = 0$ possède au moins une solution réelle.
La fonction $f : x \mapsto x^3 + ax^2 + bx + c$ est continue car polynomiale.
Puisque $\lim\limits_{+\infty} f = +\infty$ et $\lim\limits_{-\infty} f = -\infty$, la fonction f prend une valeur positive et une valeur négative et donc cette fonction s'annule au moins une fois.
Plus généralement, la démarche qui précède s'adaptent aux fonctions polynomiales de degrés impairs.

Exemple Soient $f, g : [a, b] \to \mathbb{R}$ continues.
On suppose $f(a) \leqslant g(a)$ et $f(b) \geqslant g(b)$.
Montrons qu'il existe $x \in [a, b]$ tel que $f(x) = g(x)$.
On introduit pour cela la fonction auxiliaire $\varphi : x \mapsto f(x) - g(x)$.
Cette fonction est continue par opérations sur les fonctions continues.
De plus $\varphi(a) \leqslant 0$ et $\varphi(b) \geqslant 0$ donc φ s'annule et cette annulation détermine un $x \in [a, b]$ vérifiant $f(x) = g(x)$.



Exemple Soit $f : [a, b] \to [a, b]$ continue.
Montrons qu'il existe $x \in [a, b]$ tel que $f(x) = x$.
On introduit pour cela la fonction auxiliaire $\varphi : x \mapsto f(x) - x$.
Cette fonction est continue par opérations sur les fonctions continues.
De plus $\varphi(a) = f(a) - a \geqslant 0$ et $\varphi(b) = f(b) - b \leqslant 0$ car $f(a), f(b) \in [a, b]$.
Par suite la fonction φ s'annule et cette annulation détermine un $x \in [a, b]$ vérifiant $f(x) = x$.



17.2.2 Image d'un intervalle

Rappel Soit J une partie de $\mathbb{R}$. On a équivalence entre :
(i) J est un intervalle de $\mathbb{R}$;
(ii) $\forall x, y \in J, x \leqslant y \Rightarrow [x, y] \subset J$.

Théorème

L'image d'un intervalle par une fonction continue est un intervalle.

dém. :
Soit $f : \mathcal{D} \to \mathbb{R}$ une fonction continue et I un intervalle inclus dans le domaine de définition $\mathcal{D}$.
Montrons que $J = f(I)$ est un intervalle de $\mathbb{R}$.
Soient $y, y' \in J$ tels que $y \leqslant y'$.
Il existe $x, x' \in I$ tels que $y = f(x)$ et $y' = f(x')$.
En appliquant le théorème des valeurs intermédiaires entre x et x', on obtient que pour tout $y'' \in [y, y']$, il existe x'' compris entre x et x' tel que $y'' = f(x'')$.
Puisque x'' est compris entre x et x' avec $x, x' \in I$, on a $x'' \in I$ puis $y'' = f(x'') \in f(I) = J$
Ainsi $[y, y'] \subset J$ et ceci valant pour tout $y \leqslant y' \in J$, la caractérisation des intervalles de $\mathbb{R}$ permet de conclure que J est un intervalle.
□

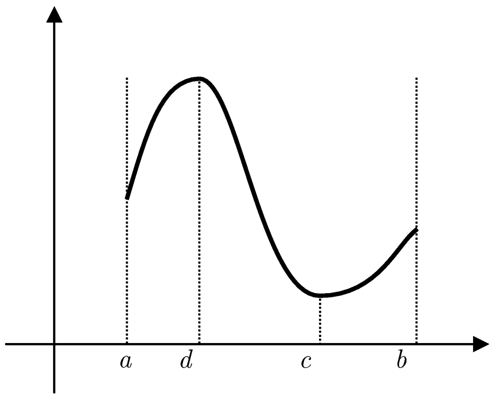
Remarque Les extrémités de l'intervalle $f(I)$ sont $\inf_I f$ et $\sup_I f$; on ne peut sans hypothèses supplémentaires affirmer si ces extrémités sont ouvertes ou fermées.

Exemple Soit $f : I \to \mathbb{R}$ continue. On suppose que f n'est ni majorée, ni minorée. Montrons que f est surjective.
Il suffit pour cela d'établir que $f(I) = \mathbb{R}$.
Or $f(I)$ est un intervalle de $\mathbb{R}$ et celui-ci n'est ni majoré, ni minoré, ce ne peut qu'être $\mathbb{R}$.

17.2.3 Image d'un segment

Théorème

Soient $a, b \in \mathbb{R}$ tels que $a \leqslant b$ et $f : [a, b] \to \mathbb{R}$
Si f est continue alors f admet un minimum et un maximum.
On dit que f est bornée et atteint ses bornes.
Ainsi
$$\exists c, d \in [a, b], \forall x \in [a, b], f(c) \leqslant f(x) \leqslant f(d)$$



dém. :
Posons $M = \sup\limits_{[a,b]} f \in \mathbb{R} \cup \{+\infty\}$.
Par la réalisation séquentielle d'une borne supérieure, il existe une suite (y_n) de valeurs prises par f sur $[a, b]$ vérifiant $y_n \to M$ et par conséquent il existe une suite (x_n) d'éléments de $[a, b]$ vérifiant $f(x_n) \to M$.
Puisque la suite (x_n) est bornée, le théorème de Bolzano-Weierstrass affirme l'existence d'une suite extraite $(x_{\varphi(n)})$ convergente. Posons d sa limite.
Puisque pour tout $n \in \mathbb{N}$, $a \leqslant x_{\varphi(n)} \leqslant b$, on obtient à la limite $d \in [a, b]$.
Mais par continuité de f en d, on a alors $f(x_{\varphi(n)}) \to f(d)$.
Or $(f(x_{\varphi(n)}))$ est une suite extraite de la suite $(f(x_n))$ qui tend vers M donc $f(x_{\varphi(n)}) \to M$.
Par unicité de la limite on obtient $M \in \mathbb{R}$ et $M = f(d)$.
Ainsi f admet un maximum en d.
De façon symétrique, on montre que f admet un minimum en un certain $c \in [a, b]$.
□

Remarque Puisque f admet un minimum et un maximum sur $[a, b]$, on peut écrire :
$$\inf_{[a,b]} f = \min_{[a,b]} f \text{ et } \sup_{[a,b]} f = \max_{[a,b]} f$$

Corollaire

L'image d'un segment par une fonction continue est un segment.
Plus précisément :

$$f([a,b]) = \left[\min_{[a,b]} f, \max_{[a,b]} f\right]$$

Attention : Cette propriété n'est vraie que pour les segments.

Exemple Soit $f : [a,b] \to \mathbb{R}$ continue telle que

$$\forall x \in [a,b], f(x) > 0$$

Montrons que : $\inf_{x\in[a,b]} f(x) > 0$.
Puisque f est continue sur le segment $[a,b]$ il y admet un minimum en un certain $c \in [a,b]$.
On a alors

$$\inf_{x\in[a,b]} f(x) = \min_{[a,b]} f(x) = f(c) > 0$$

Exemple Soit $f : \mathbb{R}^+ \to \mathbb{R}$ continue telle que f converge en $+\infty$. Montrons que f est bornée.
Puisque f converge en $+\infty$, f est bornée au voisinage de $+\infty$ et donc il existe $M \in \mathbb{R}$ et il existe $A \in \mathbb{R}^+$ tel que

$$\forall x \in [A, +\infty[, |f(x)| \leqslant M$$

De surcroît f est bornée par un certain M' sur $[0,A]$ car continue sur un segment.
Par conséquent f est bornée par $\max(M, M')$ sur $\mathbb{R}^+$.

17.2.4 Fonction continue strictement monotone

Théorème

Si $f : I \to \mathbb{R}$ est une fonction continue strictement croissante définie sur un intervalle I alors
1) $f(I)$ est un intervalle de même type que I et dont les extrémités sont les limites de f aux extrémités de I ;
2) f réalise une bijection de I vers $f(I)$;
3) son application réciproque $f^{-1} : f(I) \to I$ est continue, de même monotonie que f et les limites de f^{-1} aux extrémités de $f(I)$ sont les extrémités de I.

dém. :
1) Notons $a < b \in \bar{\mathbb{R}}$ les extrémités de I.
Cas $I =]a,b[$
Comme f est croissante, les limites de f en a et b existent dans $\bar{\mathbb{R}}$ et

$$\lim_a f = \inf_{]a,b[} f \text{ et } \lim_b f = \sup_{]a,b[} f$$

On a alors pour tout $x \in]a,b[$, $\lim_a f \leqslant f(x) \leqslant \lim_b f$ et de plus ces inégalités sont strictes car sinon x serait un extremum de f ce qui est impossible puisque f est strictement croissante. Ainsi

$$f(]a,b[) \subset \left]\lim_a f, \lim_b f\right[$$

Inversement.
Soit $y \in \left]\lim_a f, \lim_b f\right[$. Il existe $x_1, x_2 \in]a,b[$ tel que $f(x_1) < y < f(x_2)$
En appliquant le théorème des valeurs intermédiaires entre x_1 et x_2 il existe $x \in [x_1, x_2] \subset I$ tel que $y = f(x)$
Ainsi

$$\left]\lim_a f, \lim_b f\right[\subset f(]a,b[)$$

puis finalement

$$f(]a,b[) = \left]\lim_a f, \lim_b f\right[$$

Cas $I =]a,b]$
La continuité de f en b impose $\lim_{b^-} f = f(b)$.
Par suite

$$f(I) = f(]a,b[\cup \{b\}) = f(]a,b[) \cup \{f(b)\} = \left]\lim_a f, \lim_{b^-} f\right[\cup \{f(b)\} = \left]\lim_a f, f(b)\right]$$

De même, si $I = [a,b[$ ou $I = [a,b]$ on obtient respectivement : $f(I) = \left[f(a), \lim_b f\right[$ et $f(I) = [f(a), f(b)]$.
Puisque
2) f est strictement croissante donc injective et par suite f réalise une bijective de I vers $f(I)$.
3) f^{-1} est strictement croissante car application réciproque d'une fonction qu'il l'est.
Montrons maintenant que f^{-1} est continue.
Soit y_0 un élément de $f(I)$
Si y_0 n'est pas extrémité droite de $f(I)$ alors comme f^{-1} est croissante, f^{-1} admet une limite finie en y_0^+ qui est supérieure à $f(y_0)$
Quand $y \to y_0^+$.
D'une part $f(f^{-1}(y)) = y \to y_0$.
D'autre part $f^{-1}(y) \to \lim_{y_0^+} f^{-1}$, puis par continuité de f,

$$f(f^{-1}(y)) \to f(\lim_{y_0^+} f^{-1})$$

Par unicité de la limite, on obtient

$$f(\lim_{y_0^+} f^{-1}) = y_0 \text{ puis } \lim_{y_0^+} f^{-1} = f^{-1}(y_0)$$

Ainsi f^{-1} est continue à droite en y_0.
De même, on obtient f^{-1} continue à gauche en y_0 quand y_0 n'est pas extrémité gauche de $f(I)$.
Finalement f^{-1} est une fonction continue.
Enfin, puisque f^{-1} est continue et strictement croissante définie sur $f(I)$, le point 1) appliqué à f^{-1}

permet d'affirmer que f^{-1} réalise une bijection de $f(I)$ vers un intervalle dont les extrémités sont ses limites en les extrémités $f(I)$. Or f^{-1} réalise une bijection de $f(I)$ vers I, donc les limites de f^{-1} aux extrémités de $f(I)$ sont les extrémités de I.

□

Corollaire

Par passage à l'opposé, on peut énoncer un résultat semblable pour f continue et strictement décroissante sur un intervalle I.

Remarque Rappelons que les graphes de f et f^{-1} sont symétriques par rapport à Δ : $y = x$ (1ère bissectrice).

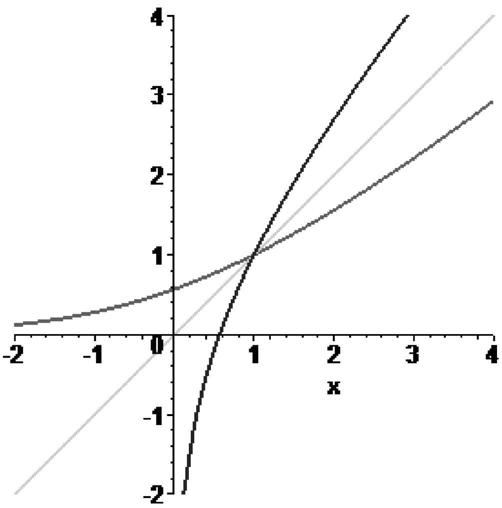
Remarque Par ce théorème, il est possible de déduire des bijections d'un tableau de variation. C'est ce théorème qui permet de définir bon nombre des fonctions usuelles comme application réciproque de fonction déjà connue.

Exemple Soit $f : \mathbb{R}^{+\star} \to \mathbb{R}$ définie par

$$f(x) = x + \ln x$$

La fonction f est continue et strictement croissante.
Ses limites en 0^+ et $+\infty$ sont respectivement $-\infty$ et $+\infty$.
Par le théorème de la bijection monotone, on peut affirmer que f réalise une bijection de $\mathbb{R}^{+\star}$ vers $\left]\lim_0 f, \lim_{+\infty} f\right[= \mathbb{R}$. De plus, son application réciproque $f^{-1} : \mathbb{R} \to \mathbb{R}^{+\star}$ est continue, strictement croissante et on a

$$\lim_{x\to-\infty} f^{-1}(x) = 0^+ \text{ et } \lim_{x\to+\infty} f^{-1}(x) = +\infty$$



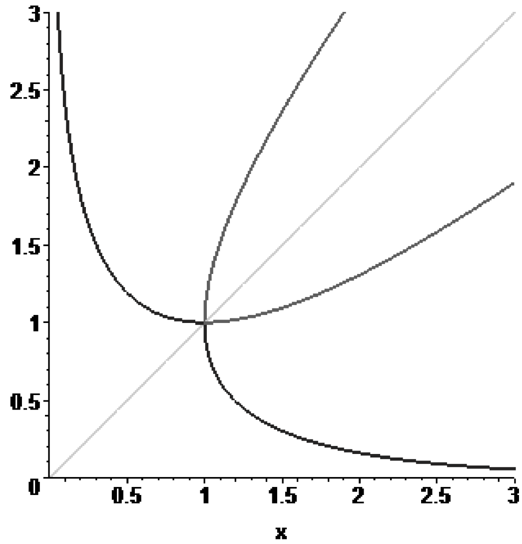
Exemple Soit $f : \mathbb{R}^{+\star} \to \mathbb{R}$ définie par

$$f(x) = x - \ln x$$

La fonction f est dérivable et $f'(x) = 1 - 1/x = (x-1)/x$ du signe de $x - 1$.

x	0		1		$+\infty$
$f'(x)$		$-$	0	$+$	
$f(x)$	$+\infty$	$\searrow$	1	$\nearrow$	$+\infty$

La fonction f est continue est strictement décroissante sur $]0,1]$ donc réalise une bijection de $]0,1]$ vers $[1,+\infty[$.
La fonction f est continue et strictement croissante sur $[1,+\infty[$ donc réalise une bijection de $[1,+\infty[$ vers $[1,+\infty[$



Proposition

> Si $f : I \to J$ est une bijection continue alors f est strictement monotone et f^{-1} continue.

dém. :
Soit $f : I \to \mathbb{R}$ une fonction continue et injective
Si I est réduit à un point, il n'y a rien à montrer.
Sinon, I possède au moins deux points que l'on notera a, b de sorte que $a < b$.
Puisque f est injective $f(a) \neq f(b)$.
Cas $f(a) < f(b)$
Montrons que f est strictement croissante.
Soient $x, y \in I$ tels que $x < y$.
Posons $\varphi : [0,1] \to \mathbb{R}$ définie par :

$$\varphi(t) = f\left((1-t)b + ty\right) - f\left((1-t)a + tx\right)$$

Par construction $\varphi(0) = f(b) - f(a) > 0$ et $\varphi(1) = f(y) - f(x)$.
Par opérations sur les fonctions continues φ est continue.

Supposons par l'absurde $f(y) \leqslant f(x)$. On a alors $\varphi(1) \leqslant 0$.
En appliquant le théorème des valeurs intermédiaires entre 0 et 1, on obtient l'existence de $t \in]0,1]$ tel que $\varphi(t) = 0$.
Posons alors $u = (1-t)a + tx$ et $v = (1-t)b + ty$.
$\varphi(t) = 0$ donne $f(u) = f(v)$ puis $u = v$ car f est supposée injective.
Or, $a \leqslant b$ et $(1-t) \geqslant 0$ donc $(1-t)a \leqslant (1-t)b$,
de plus $x < y$ et $t > 0$ donc $tx < ty$ et par suite $u < v$. C'est absurde.
Par suite $f(x) < f(y)$ et nous avons établi : $x < y \Rightarrow f(x) < f(y)$.
Ainsi f est strictement croissante.
Cas $f(b) < f(a)$.
On applique l'étude précédente à $g = -f$ et on obtient que g est strictement croissante puis f est strictement décroissante.
Sachant f strictement monotone, on peut alors appliquer à f le théorème de la bijection continue strictement monotone et conclure que son application réciproque est continue.
□

17.3 Extension aux fonctions complexes

Les notions qui suivent prolongent celles vue dans le cadre des fonctions à valeurs réelles.

17.3.1 Définition

Définition

On dit que $f : \mathcal{D} \to \mathbb{C}$ est continue en $a \in \mathcal{D}$ si $f(t) \xrightarrow[t \to a]{} f(a)$.
On dit que f est continue si f est continue en tout point de $\mathcal{D}$.
On note $\mathcal{C}(\mathcal{D}, \mathbb{C})$ l'ensemble des fonctions complexes continues sur $\mathcal{D}$.

Remarque On peut aussi introduire les notions de continuité à droite et à gauche en un point.

17.3.2 Opérations

Théorème

Soient $f, g : \mathcal{D} \to \mathbb{C}$ et $\lambda \in \mathbb{C}$.
Si f et g sont continues alors $\lambda.f, f+g, fg, \bar{f}$ et $|f|$ le sont aussi.
Si de plus si g ne s'annule pas alors la fonction f/g est continue.

dém. :
Par opérations sur les limites en chaque $a \in \mathcal{D}$.
□

Exemple Les fonctions polynomiales et les fonctions rationnelles sont continues.

Théorème

Soit $f : \mathcal{D} \to \mathbb{C}$. On a équivalence entre :
(i) f est continue ;
(ii) $\mathrm{Re}(f)$ et $\mathrm{Im}(f)$ le sont.

dém. :
(i) $\Rightarrow$ (ii) via $\mathrm{Re}(f) = \dfrac{1}{2}(f + \bar{f})$ et $\mathrm{Im}(f) = \dfrac{1}{2i}(f - \bar{f})$.
(ii) $\Rightarrow$ (i) via $f = \mathrm{Re}(f) + i.\mathrm{Im}(f)$.
□

Exemple Pour $\alpha \in \mathbb{C}$, la fonction $t \mapsto \mathrm{e}^{\alpha t}$ est continue sur $\mathbb{R}$.
En effet $\mathrm{e}^{\alpha t} = \cos(\omega t)\mathrm{e}^{at} + i\sin(\omega t)\mathrm{e}^{at}$ avec $a = \mathrm{Re}(\alpha)$ et $\omega = \mathrm{Im}(\alpha)$.

17.3.3 Propriétés

Attention : Pour les fonctions complexes le théorème des valeurs intermédiaires n'a pas de sens.
En effet on ne peut pas parler de valeurs intermédiaire entre deux nombres complexes
Par exemple la fonction continue $f(t) = \mathrm{e}^{it}$ prend la valeur 1 en 0 et la valeur -1 en π sans pour autant s'annuler.

Proposition

Si $f : [a, b] \to \mathbb{C}$ est continue alors f est bornée.

dém. :
La fonction $t \mapsto |f(t)|$ est une fonction à valeur réelle définie et continue sur un segment donc majorée.
□

17.3.4 Fonction à valeurs dans $\mathbb{R}^2$

Soit $f : \mathcal{D} \to \mathbb{R}^2$ une fonction à valeurs dans $\mathbb{R}^2$ définie sur $\mathcal{D}$.
Pour tout $t \in \mathcal{D}$, on peut écrire $f(t) = (x(t), y(t))$ ce qui introduit x et $y : \mathcal{D} \to \mathbb{R}$ appelées les fonctions coordonnées de f.

Définition

On dit que $f : \mathcal{D} \to \mathbb{R}^2$ est continue en $a \in \mathcal{D}$ si $f(t) \xrightarrow[t \to a]{} f(a)$.
On dit que f est continue si elle l'est en tout $a \in \mathcal{D}$.
On note $\mathcal{C}(\mathcal{D}, \mathbb{R}^2)$ l'ensemble des fonctions continues de $\mathcal{D}$ vers $\mathbb{R}^2$.

Proposition

$f : \mathcal{D} \to \mathbb{R}^2$ est continue si, et seulement si, ses fonctions coordonnées le sont.

dém. :
Notons x et y les fonctions coordonnées de f.
Soit $a \in \mathcal{D}$, par caractérisation des limites via les fonctions coordonnées, on a $f(t) \xrightarrow[t \to a]{} f(a)$ si, et seulement si, $x(t) \xrightarrow[t \to a]{} x(a)$ et $y(t) \xrightarrow[t \to a]{} y(a)$.
□

17.4 Uniforme continuité

17.4.1 Définition

Définition

Soit $f : \mathcal{D} \to \mathbb{K}$ avec $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.
On dit que f est uniformément continue si on a la propriété

$$\forall \varepsilon > 0, \exists \alpha > 0, \forall x, y \in \mathcal{D}, |y - x| \leqslant \alpha \Rightarrow |f(y) - f(x)| \leqslant \varepsilon$$

Remarque f continue signifie

$$\forall x \in \mathcal{D}, \forall \varepsilon > 0, \exists \alpha > 0, \forall y \in \mathcal{D}, |y - x| \leqslant \alpha \Rightarrow |f(y) - f(x)| \leqslant \varepsilon$$

La différence entre ces deux assertions est dans le positionnement du terme $\forall x \in \mathcal{D}$.
L'uniforme continuité est plus forte que la continuité car cette notion exige que α soit indépendant de x, (on dit que α est uniforme).

Proposition

Si $f : \mathcal{D} \to \mathbb{K}$ est uniformément continue alors f est continue.

dém. :
S'il existe un $\alpha > 0$ convenant pour tout $x \in \mathcal{D}$, il existe a fortiori pour chaque $x \in \mathcal{D}$ un $\alpha > 0$ convenable.
□

Exemple Les fonctions lipschitziennes sont uniformément continues.
Soit $f : \mathcal{D} \to \mathbb{K}$ une fonction lipschitzienne de rapport k
On a

$$\forall x, y \in \mathcal{D}, |f(y) - f(x)| \leqslant k\,|y - x|$$

Soit $\varepsilon > 0$.
Cherchons $\alpha > 0$ tel que

$$|y - x| \leqslant \alpha \Rightarrow |f(y) - f(x)| \leqslant \varepsilon$$

Sachant $|f(y) - f(x)| \leqslant k\,|y - x|$
Pour $\alpha = \varepsilon/(k + 1) > 0$, on a

$$|y - x| \leqslant \alpha \Rightarrow |f(y) - f(x)| \leqslant \frac{k}{k+1}\varepsilon$$

et donc

$$|y - x| \leqslant \alpha \Rightarrow |f(y) - f(x)| \leqslant \varepsilon$$

Ainsi la fonction f est uniformément continue.

17.4.2 Fonction continue sur un segment

Théorème

Toute fonction réelle ou complexe continue sur un segment y est uniformément continue.

dém. :
Soient $a \leqslant b \in \mathbb{R}$ et $f : [a, b] \to \mathbb{K}$ continue.
Supposons par l'absurde f ne soit pas uniformément continue.
Il existe alors un $\varepsilon > 0$ vérifiant

$$\forall \alpha > 0, \exists x, y \in [a, b]\,,\ |y - x| \leqslant \alpha \text{ et } |f(y) - f(x)| > \varepsilon$$

Pour tout $n \in \mathbb{N}$, en posant $\alpha = \dfrac{1}{n+1} > 0$, il existe deux éléments de $[a, b]$ que l'on note x_n et y_n tels que

$$|y_n - x_n| \leqslant \frac{1}{n+1} \text{ et } |f(y_n) - f(x_n)| > \varepsilon$$

En faisant varier n, cela détermine deux suites (x_n) et (y_n) d'éléments de $[a, b]$
La suite (x_n) est une suite réelle bornée. Par le théorème de Bolzano-Weierstrass, on peut en extraire une suite $(x_{\varphi(n)})$ convergente. Posons $c = \lim\limits_{n\infty} x_{\varphi(n)}$.
Puisque pour tout $n \in \mathbb{N}$, $a \leqslant x_{\varphi(n)} \leqslant b$, on obtient à la limite $a \leqslant c \leqslant b$.
De plus

$$\left|y_{\varphi(n)} - c\right| \leqslant \left|y_{\varphi(n)} - x_{\varphi(n)}\right| + \left|x_{\varphi(n)} - c\right| \leqslant \frac{1}{\varphi(n)+1} + \left|x_{\varphi(n)} - c\right| \to 0$$

donc $y_{\varphi(n)} \to c$.
La fonction f étant continue en $c \in [a, b]$ on obtient $f(x_{\varphi(n)}) \to f(c)$ et $f(y_{\varphi(n)}) \to f(c)$.
Or par construction $\left|f(y_{\varphi(n)}) - f(x_{\varphi(n)})\right| > \varepsilon$ donc à la limite $0 \geqslant \varepsilon$.
C'est absurde !
□

Remarque Ce résultat sera utile dans le chapitre d'intégration pour définir l'intégrale d'une fonction continue sur un segment.

Chapitre 18

Dérivation

Dans ce chapitre I et J désignent des intervalles non singuliers de $\mathbb{R}$ i.e. des intervalles non vides et non réduits à un point.

18.1 Dérivées d'une fonction réelle

18.1.1 Nombre dérivé

Soient $f : I \to \mathbb{R}$ et Γ_f le graphe de f.
Pour étudier la courbe Γ_f, on introduit le paramétrage

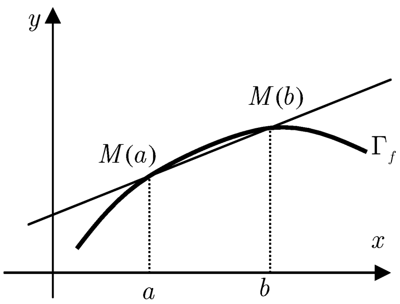
$$\begin{cases} x = t \\ y = f(t) \end{cases} \text{ avec } t \in I$$

de point courant $M(t)$

Définition

Pour $a, b \in I$ distincts, on appelle taux de variation de f entre a et b le réel :

$$\tau(a, b) = \frac{f(b) - f(a)}{b - a}$$



Remarque $\tau(a, b)$ est la pente (ou coefficient directeur) de la droite $(M(a)M(b))$.

Définition

On dit que f est dérivable en $a \in I$ si le taux de variation de f entre a et x converge quand $x \to a$ (avec $x \neq a$) autrement dit si

$$\lim_{x \to a, x \neq a} \frac{f(x) - f(a)}{x - a} = \lim_{h \to 0, h \neq 0} \frac{1}{h} \left(f(a+h) - f(a) \right)$$

existe et est finie.
Cette limite est alors appelée nombre dérivé de f en a et est notée $f'(a)$.

Remarque Ainsi et sous réserve d'existence :

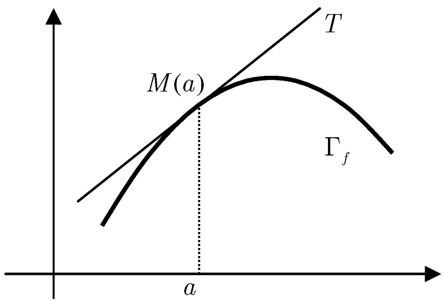
$$f'(a) = \lim_{x \to a, x \neq a} \frac{f(x) - f(a)}{x - a}$$

Proposition

Si f est dérivable en $a \in I$ alors la courbe Γ_f admet une tangente au point $M(a)$ qui est la droite T d'équation

$$y = f'(a)(x - a) + f(a)$$

Par abus, on dit que T est la tangente à f en a.



dém. :
Les droites $(M(a)M(a+h))$ pivotent autour du point $M(a)$ et on pour pente $\tau(a, a+h)$.
Or quand $h \to 0$, $\tau(a, a+h) \to f'(a)$, donc les droites $(M(a)M(a+h))$ tendent vers la droite passant par $M(a)$ de pente $f'(a)$.
□

Théorème

Si f est dérivable en $a \in I$ alors

$$f(x) = f(a) + f'(a)(x - a) + o(x - a)$$

quand $x \to a$.
Cette relation est appelée développement limité à l'ordre 1 de f en a.

dém. :
Quand $x \to a$ (avec $x \neq a$) on peut écrire

$$\frac{f(x) - f(a)}{x - a} = f'(a) + \varepsilon(x)$$

avec $\varepsilon(x) \xrightarrow[x \to a, x \neq a]{} 0$.
On en déduit

$$f(x) = f(a) + f'(a)(x-a) + (x-a)\varepsilon(x)$$

En posant $\varepsilon(a) = 0$, la relation qui précède reste valable pour $x = a$ et on a $\varepsilon(x) \xrightarrow[x \to a]{} 0$.
Par suite

$$f(x) = f(a) + f'(a)(x-a) + o(x-a)$$

quand $x \to a$.
□

Remarque Cette relation traduit que la tangente en a est la droite la plus proche de la courbe Γ_f au voisinage du point d'abscisse a.

Corollaire

Si $f : I \to \mathbb{R}$ est dérivable en a alors f est continue en a.

dém. :
Quand $x \to a$,

$$f(x) = f(a) + f'(a)(x-a) + o(x-a) \to f(a)$$

□

Exemple Soit $f : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = C$ une fonction constante.
Soit $a \in \mathbb{R}$.
Quand $x \to a$ (avec $x \neq a$),

$$\frac{f(x) - f(a)}{x - a} = 0 \xrightarrow[x \to a]{} 0$$

Par suite f est dérivable en $a \in \mathbb{R}$ et $f'(a) = 0$.

Exemple Soit $f : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = x$.
Soit $a \in \mathbb{R}$.
Quand $x \to a$ (avec $x \neq a$),

$$\frac{f(x) - f(a)}{x - a} = \frac{x - a}{x - a} = 1 \xrightarrow[x \to a]{} 1$$

Par suite f est dérivable en $a \in \mathbb{R}$ et $f'(a) = 1$.

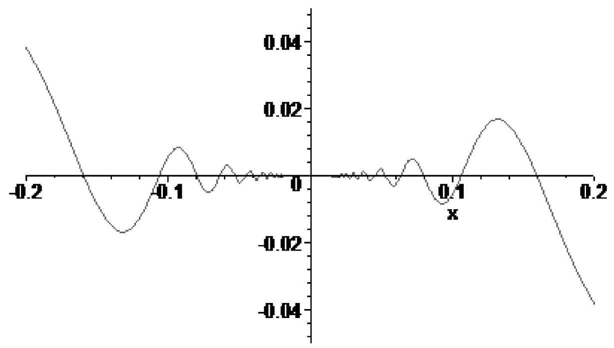
Exemple Soit $f : \mathbb{R} \to \mathbb{R}$ définie par

$$f(x) = \begin{cases} x^2 \sin 1/x & \text{si } x \neq 0 \\ 0 & \text{sinon} \end{cases}$$

Etudions la dérivabilité de f en 0.
Quand $x \to 0$ (avec $x \neq 0$),

$$\frac{f(x) - f(0)}{x} = x \sin \frac{1}{x} \to 0$$

car le produit d'une fonction bornée par une fonction de limite nulle est de limite nulle.
On en déduit que f est dérivable en 0 et $f'(0) = 0$.



Proposition

Si f est continue en $a \in I$ et si

$$\lim_{h\to 0, h\neq 0} \frac{1}{h}\left(f(a+h) - f(a)\right) = +\infty \text{ ou } -\infty$$

alors f n'est pas dérivable en a et Γ_f admet une tangente verticale en $M(a)$ qui est la droite d'équation $x = a$.
Par abus on dit que f admet une tangente verticale en a.

dém. :
Les droites $(M(a)M(a+h))$ pivotent autour de $M(a)$ et ont pour pente $\tau(a, a+h)$.
Or quand $h \to 0$ (avec $h \neq 0$), $\tau(a, a+h) \to +\infty$ (ou $-\infty$) et donc les droites $(M(a)M(a+h))$ tendent vers la droite verticale passant par $M(a)$.
□

Exemple Soit $f : \mathbb{R}^+ \to \mathbb{R}$ définie par $f(x) = \sqrt{x}$.
Quand $x \to 0^+$,

$$\frac{f(x) - f(0)}{x} = \frac{1}{\sqrt{x}} \to +\infty$$

La fonction $\sqrt{.}$ n'est pas dérivable en 0 mais son graphe présente une tangente verticale en 0.

Remarque Il existe des fonctions continues n'admettant pas de tangente en un point :

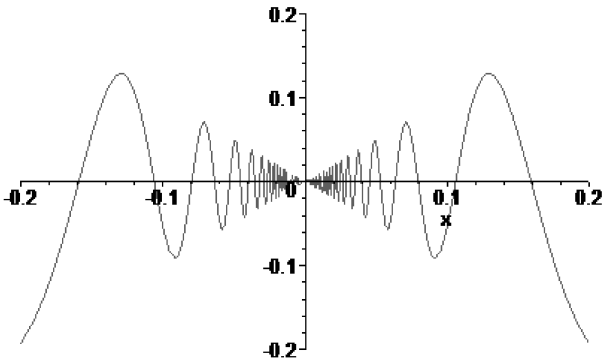
Exemple Soit $f : \mathbb{R} \to \mathbb{R}$ définie par

$$f(x) = \begin{cases} x \sin 1/x & \text{si } x \neq 0 \\ 0 & \text{sinon} \end{cases}$$

Quand $x \to 0$ (avec $x \neq 0$),

$$\frac{f(x) - f(0)}{x} = \sin\left(\frac{1}{x}\right)$$

diverge tout en étant borné.



18.1.2 Fonction dérivée

Définition

On dit que $f : I \to \mathbb{R}$ est dérivable si f est dérivable en tout point de I.
On note $\mathcal{D}(I, \mathbb{R})$ l'ensemble des fonctions dérivables de I vers $\mathbb{R}$.

Exemple Les fonctions constantes et la fonction $x \mapsto x$ sont dérivables.

Proposition

Si $f : I \to \mathbb{R}$ est dérivable alors f est continue.

dém. :
Si f est dérivable alors f est dérivable en chaque $a \in I$ et donc f est continue en chaque $a \in I$.

□

Attention : Il existe des fonctions continues non dérivables !
Parmi celle-ci soulignons les fonctions continues $\sqrt{.}$ et $|\,.\,|$ toutes deux non dérivables en 0.
Soulignons l'existence de la fonction de Bolzano qui est continue sur $[0, 1]$ et qui n'est dérivable en aucun point de $[0, 1]$! Voici son allure

Définition

Si $f : I \to \mathbb{R}$ est dérivable, on introduit sa fonction dérivée qui est l'application f' de I vers $\mathbb{R}$ qui à $x \in I$ associe le nombre dérivé de f en x.
Cette fonction est encore parfois notée $\mathrm{D}f$.

Exemple Rappelons quelques formules de dérivation bien à connaître :
$(\mathrm{e}^x)' = \mathrm{e}^x$ sur $\mathbb{R}$ et $(\ln x)' = 1/x$ sur $\mathbb{R}^{+\star}$
Pour $n \in \mathbb{Z}$, $(x^n)' = nx^{n-1}$ et en ainsi $\left(\dfrac{1}{x^n}\right) = -\dfrac{n}{x^{n+1}}$ sur $\mathbb{R}^{+\star}$ ou $\mathbb{R}^{-\star}$
Pour $\alpha \in \mathbb{R}$, $(x^\alpha)' = \alpha x^{\alpha-1}$ et en particulier $(\sqrt{x})' = \dfrac{1}{2\sqrt{x}}$ sur $\mathbb{R}^{+\star}$.
$(\cos x)' = -\sin x$, $(\sin x)' = \cos x$ sur $\mathbb{R}$.
$(\tan x) = 1 + \tan x = \dfrac{1}{\cos^2 x}$ sur $]-\pi/2, \pi/2[+ k\pi$ (avec $k \in \mathbb{Z}$)
$(\arccos x)' = \dfrac{-1}{\sqrt{1-x^2}}$, $(\arcsin x)' = \dfrac{1}{\sqrt{1-x^2}}$ sur $]-1, 1[$ et $(\arctan x)' = \dfrac{1}{1+x^2}$ sur $\mathbb{R}$.
$(\mathrm{ch}x)' = \mathrm{sh}x$, $(\mathrm{sh}x)' = \mathrm{ch}x$, $(\mathrm{th}x)' = 1 - \mathrm{th}^2 x$ sur $\mathbb{R}$.
$(\mathrm{argch}x)' = \dfrac{1}{\sqrt{x^2-1}}$ sur $]1, +\infty[$, $(\mathrm{argsh}x)' = \dfrac{1}{\sqrt{x^2+1}}$ sur $\mathbb{R}$ et $(\mathrm{argth}x)' = \dfrac{1}{1-x^2}$ sur $]-1, 1[$.

Remarque Ci-dessus, on a abusivement écrit $(f(x))'$ au lieu de $f'(x)$.
Cette écriture peut-être ambiguë si plusieurs variables interviennent ; on peut alors employer la notation $\dfrac{\mathrm{d}}{\mathrm{d}x}(f(x))$ pour désigner le résultat de la dérivation de l'expression $f(x)$ en la variable x.
Ainsi écrire $(x^t)'$ n'est pas nécessairement clair alors qu'écrire $\dfrac{\mathrm{d}}{\mathrm{d}x}(x^t)$ ou $\dfrac{\mathrm{d}}{\mathrm{d}t}(x^t)$ n'est pas ambigu.

18.1.3 Opérations sur les fonctions dérivables

18.1.3.1 Somme, produit et rapport

Théorème

Si f et $g : I \to \mathbb{R}$ sont dérivables alors pour tout $\lambda \in \mathbb{R}$ les fonctions $\lambda.f, f+g, fg$ le sont aussi.
De plus on a alors $(\lambda.f)' = \lambda.f'$, $(f+g)' = f'+g'$ et $(fg)' = f'g + fg'$.

dém. :
Soit $a \in I$.
Quand $x \to a$ (avec $x \neq a$).

$$\frac{(\lambda.f)(x) - (\lambda.f)(a)}{x-a} = \lambda \frac{f(x)-f(a)}{x-a} \to \lambda f'(a)$$

$$\frac{(f+g)(x) - (f+g)(a)}{x-a} = \frac{f(x)-f(a)}{x-a} + \frac{g(x)-g(a)}{x-a} \to f'(a) + g'(a)$$

$$\frac{(fg)(x) - (fg)(a)}{x-a} = \frac{f(x)-f(a)}{x-a} g(x) + f(a) \frac{g(x)-g(a)}{x-a} \to f'(a)g(a) + f(a)g'(a)$$

car g est continue en a.
□

Corollaire

Si $f_1, ..., f_n : I \to \mathbb{R}$ sont dérivables alors $f_1 + \dots + f_n$ et $f_1 \dots f_n$ le sont aussi et de plus

$$(f_1 + \dots + f_n)' = f_1' + \dots + f_n'$$

et

$$(f_1 \dots f_n)' = \sum_{i=1}^{n} f_1 \dots f_{i-1} (f_i)' f_{i+1} \dots f_n$$

dém. :
En raisonnant par récurrence sur $n \in \mathbb{N}^\star$.
□

Exemple En particulier

$$(fgh)' = f'gh + fg'h + fgh'$$

Exemple Pour $n \in \mathbb{N}^\star$, puisque $x^n = x \times x \times \ldots \times x$ (produit à n termes), la formule de dérivation d'un produit de n fonctions donne $(x^n)' = nx^{n-1}$.

Exemple Par opérations, les fonctions polynomiales sont dérivables.

Théorème

Soit $f : I \to \mathbb{R}$ ne s'annulant pas
Si f est dérivable alors sa fonction inverse $1/f$ l'est aussi et

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

dém. :
Soit $a \in \mathbb{R}$.
Quand $x \to a$ (avec $x \neq a$),

$$\frac{(1/f)(x) - (1/f)(a)}{x - a} = \frac{f(a) - f(x)}{x - a} \frac{1}{f(x)f(a)} \to -f'(a)\frac{1}{f(a)^2}$$

car f est continue en a.
□

Exemple $\forall n \in \mathbb{N}^\star$,

$$\left(\frac{1}{x^n}\right)' = -\frac{nx^{n-1}}{(x^n)^2} = -\frac{n}{x^{n+1}}$$

Corollaire

Soient $f, g : I \to \mathbb{R}$ avec g ne s'annulant pas.
Si f et g sont dérivables alors $\dfrac{f}{g}$ l'est aussi et

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

dém. :
Il suffit d'écrire $\dfrac{f}{g} = f \times \dfrac{1}{g}$.
□

Exemple Par opérations, les fonctions rationnelles sont dérivables.

18.1.3.2 Composition

Théorème

Soient $f : I \to \mathbb{R}$ et $\varphi : J \to \mathbb{R}$ vérifiant $f(I) \subset J$.
Si f et φ sont dérivables alors $\varphi \circ f$ l'est aussi et

$$(\varphi \circ f)' = f' \times \varphi' \circ f$$

dém. :
Soit $a \in I$.
Puisque f est dérivable en a, on peut écrire par développement limité à l'ordre 1

$$f(x) = f(a) + f'(a)(x - a) + (x - a)\varepsilon(x)$$

avec $\varepsilon \underset{a}{\to} 0$.
Aussi, puisque g est dérivable en $f(a)$, on peut écrire par développement limité à l'ordre 1

$$g(y) = g(f(a)) + g'(f(a))(y - f(a)) + (y - f(a))\tilde{\varepsilon}(y)$$

avec $\tilde{\varepsilon} \xrightarrow[f(a)]{} 0$.
On en déduit

$$g(f(x)) = g(f(a)) + g'(f(a))f'(a)(x-a) + (x-a)\hat{\varepsilon}(x)$$

avec

$$\hat{\varepsilon}(x) = g'(f(a))\varepsilon(x) + (f'(a) + \varepsilon(x))\tilde{\varepsilon}(f(x)) \xrightarrow[x\to a]{} 0$$

Ainsi

$$g(f(x)) = g(f(a)) + g'(f(a))f'(a)(x-a) + o(x-a)$$

quand $x \to a$ et donc $g \circ f$ est dérivable en a avec $(g \circ f)'(a) = g'(f(a)) \times f'(a)$.
□

Exemple $\forall \alpha \in \mathbb{R}$, $(x^\alpha)' = (\mathrm{e}^{\alpha \ln x})' = \alpha x^{\alpha-1}$ sur $\mathbb{R}^{+\star}$

Exemple Par dérivation de fonction composée et pour une fonction u dérivable convenable :
$(\mathrm{e}^u)' = u'\mathrm{e}^u$, $(\ln u)' = \dfrac{u'}{u}$, $(u^\alpha)' = \alpha u' u^{\alpha-1}$
$(\cos u)' = -u' \sin u$, $(\sin u)' = u' \cos u$,, $(\tan u)' = u'(1+\tan^2 u)$,
$(\arccos u)' = -\dfrac{u'}{\sqrt{1-u^2}}$, $(\arcsin u)' = \dfrac{u'}{\sqrt{1-u^2}}$, $(\arctan u)' = \dfrac{u'}{1+u^2}$,
$(\mathrm{ch} u)' = u'\mathrm{sh} u$, $(\mathrm{sh} u)' = u'\mathrm{ch} u$, $(\mathrm{th} u)' = u'(1-\mathrm{th}^2 u)$,
$(\arg\mathrm{ch} u)' = \dfrac{u'}{\sqrt{u^2-1}}$, $(\mathrm{argsh} u)' = \dfrac{u'}{\sqrt{u^2+1}}$ et $(\mathrm{argth} u)' = \dfrac{u'}{1-u^2}$.

Exemple Calculons quelques dérivées...

$$\left(\frac{1}{1+x^2}\right)' = \left(\frac{1}{u}\right)' = -\frac{u'}{u^2} = \frac{-2x}{(1+x^2)^2}$$

$$\left(\frac{1}{(1+x^2)^2}\right)' = \left(\frac{1}{u^2}\right)' = -\frac{2u'}{u^3} = -\frac{4x}{(1+x^2)^3}$$

$$\left(\frac{x}{(1+x)^n}\right)' = \left(u \times \frac{1}{v}\right)' = u'\frac{1}{v} + u\left(\frac{1}{v}\right)' = \frac{1}{(1+x)^n} - \frac{nx}{(1+x)^{n+1}} = \frac{1+(1-n)x}{(1+x)^{n+1}}$$

$$(a^x)' = \left(\mathrm{e}^{x\ln a}\right)' = \ln a . a^x \text{ et } (x^x)' = \left(\mathrm{e}^{x\ln x}\right)' = (1+\ln x)x^x$$

18.1.3.3 Application réciproque

Théorème

Soit $f : I \to J$ une bijection continue.
Si f est dérivable en $a \in I$ et si $f'(a) \neq 0$ alors $f^{-1} : J \to I$ est dérivable en $b = f(a)$ et

$$(f^{-1})'(b) = \frac{1}{f'(a)} = \frac{1}{f'(f^{-1}(b))}$$

dém. :
Puisque f est une bijection continue donc son application réciproque $f^{-1}: J \to I$ l'est aussi.
Soit $a \in I$ tel que $f'(a)$ existe et est non nul. Posons $b = f(a) \in J$.
Pour $h \neq 0$, en introduisant $x_h = f^{-1}(b+h)$, on a

$$f(x_h) - f(a) = b + h - b = h$$

donc

$$\frac{1}{h}\left(f^{-1}(b+h) - f^{-1}(b)\right) = \frac{x_h - a}{f(x_h) - f(a)}$$

Quand $h \to 0$ (avec $h \neq 0$), on a $x_h = f^{-1}(b+h) \to f^{-1}(b) = a$ car f^{-1} est continue en b et par composition de limite on a

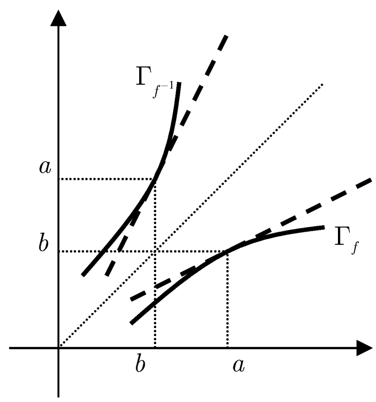
$$\frac{f(x_h) - f(a)}{x_h - a} \to f'(a)$$

puis

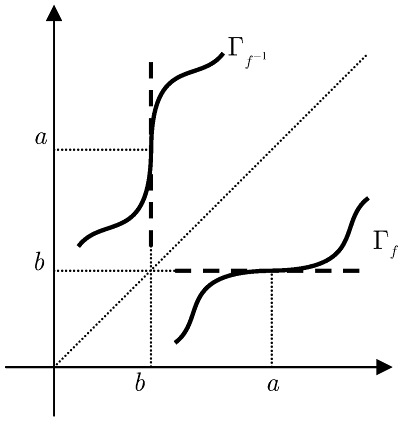
$$\frac{1}{h}\left(f^{-1}(b+h) - f^{-1}(b)\right) \to \frac{1}{f'(a)}$$

□

Remarque Géométriquement, les courbes représentant f et f^{-1} sont symétriques par rapport à la droite d'équation $y = x$. Cette symétrie transforme une tangente à f en a de pente α en une tangente à f^{-1} en $b = f(a)$ de pente $1/\alpha$.



Attention : La condition $f'(a) \neq 0$ est indispensable ; en effet si $f'(x) = 0$ alors f admet une tangente horizontale en a et par symétrie f^{-1} admet une tangente verticale en $b = f(a)$.



Corollaire

Si $f : I \to J$ est une bijection dérivable vérifiant $\forall x \in I, f'(x) \neq 0$ alors son application réciproque f^{-1} est dérivable et

$$(f^{-1})' = \frac{1}{f' \circ f^{-1}}$$

Remarque La formule de dérivation proposée peut être retrouvée en dérivant la relation $f \circ f^{-1} = \mathrm{Id}$.

Exemple On peut introduire la fonction logarithme népérien comme étant la primitive sur $\mathbb{R}^{+\star}$ de la fonction inverse s'annulant en 1. En montrant que cette fonction réalise une bijection de $\mathbb{R}^{+\star}$ vers $\mathbb{R}$, on peut introduire son application réciproque, la fonction exponentielle et la formule de dérivation précédente donne

$$(\mathrm{e}^x)' = \frac{1}{(\ln)'(\mathrm{e}^x)} = \frac{1}{1/\mathrm{e}^x} = \mathrm{e}^x$$

Exemple De façon semblable on introduit les fonctions : $\arcsin$, $\arccos$, $\arctan$ et argsh, argch, argth et on établit les formules de dérivation précédemment énoncées.

Exemple Soit $f : \mathbb{R}^+ \to \mathbb{R}$ définie par

$$f(x) = \sqrt{x} + x + 1$$

f réalise une bijection de $\mathbb{R}^+$ sur $[1, +\infty[$ car c'est une fonction continue, strictement croissante (par opération sur de telles fonctions) vérifiant $f(0) = 1$ et $\lim\limits_{+\infty} f = +\infty$.

La fonction f est dérivable sur $\mathbb{R}^{+\star}$ et

$$\forall x > 0, f'(x) = \frac{1}{2\sqrt{x}} + 1 \neq 0$$

Par le théorème précédent, on peut affirmer que son application réciproque f^{-1} est dérivable sur

$$f(\mathbb{R}^{+\star}) =]1, +\infty[$$

Etude de la dérivabilité en 1.
Quand $h \to 0$ (avec $h \neq 0$),

$$\frac{1}{h}\left(f^{-1}(1+h) - f^{-1}(1)\right) = \frac{1}{h}f^{-1}(1+h) \underset{x=f^{-1}(1+h)}{=} \frac{x}{f(x)-1} = \frac{x}{\sqrt{x}+x} \sim \frac{x}{\sqrt{x}} = \sqrt{x} \to 0$$

Ainsi f^{-1} est dérivable en 1 et $(f^{-1})'(1) = 0$.
Cela pouvait être attendu car la fonction f admet une tangente verticale en 0.

18.1.4 Nombre dérivé à droite, nombre dérivé à gauche

Définition

Soient $f : I \to \mathbb{R}$ et $a \in I$.
Si a n'est pas extrémité droite de I, on dit que f est dérivable à droite en a si le taux de variation $\frac{1}{h}(f(a+h) - f(a))$ converge quand $h \to 0^+$. On pose alors

$$f'_d(a) = \lim_{h \to 0^-} \frac{1}{h}(f(a+h) - f(a))$$

Si a n'est pas extrémité gauche de I, on dit que f est dérivable à gauche en a si le taux de variation $\frac{1}{h}(f(a+h) - f(a))$ converge quand $h \to 0^-$. On pose alors

$$f'_g(a) = \lim_{h \to 0^-} \frac{1}{h}(f(a+h) - f(a))$$

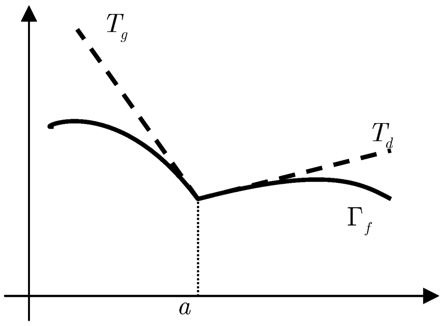
Exemple Pour $f : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = |x|$, on a $f'_d(0) = 1$ et $f'_g(0) = -1$.

Définition

Si $f : I \to \mathbb{R}$ est dérivable à droite en a alors la droite T_d d'équation

$$y = f'_d(a)(x-a) + f(a)$$

est appelée demi-tangente à droite à f en a.
De façon semblable, on définit aussi l'éventuelle demi-tangente à gauche.



Proposition

Soient $f : I \to \mathbb{R}$ et $a \in I$ qui n'est pas extrémité de I. On a équivalence entre :
(i) f est dérivable en a ;
(ii) f est dérivable à droite et à gauche en a et $f'_d(a) = f'_g(a)$.
De plus, si tel est le cas, $f'(a) = f'_d(a) = f'_g(a)$.

dém. :
Le taux de variation $\frac{1}{h}(f(a+h) - f(a))$ converge quand $h \to 0$ avec $h \neq 0$ si, et seulement si, ce taux de variation converge quand $h \to 0^+$ et quand $h \to 0^-$ vers une même limite. Cette limite commune est alors la limite de $\frac{1}{h}(f(a+h) - f(a))$ quand $h \to 0$ avec $h \neq 0$.
□

Exemple Etudions la dérivabilité de $f : \mathbb{R} \to \mathbb{R}$ définie par

$$f(x) = \begin{cases} x^2 & \text{si } x \geqslant 0 \\ -x^2 & \text{si } x < 0 \end{cases}$$

1ère méthode :
Soit $a \in \mathbb{R}$.
Cas $a > 0$
Au voisinage de a, on a $f(x) = x^2$ et donc f est dérivable en a et $f'(a) = 2a$.
Cas $a < 0$
Au voisinage de a, on a $f(x) = -x^2$ et donc f est dérivable en a et $f'(a) = -2a$.
Cas $a = 0$
Quand $h \to 0^+$,
$\frac{1}{h}(f(h) - f(0)) = \frac{h^2 - 0}{h} \to 0$ donc f est dérivable à droite en 0 et $f'_d(0) = 0$.
Quand $h \to 0^-$,
$\frac{1}{h}(f(h) - f(0)) = \frac{-h^2 - 0}{h} \to 0$ donc f est dérivable à gauche en 0 et $f'_g(0) = 0$.
Puisque f est dérivable à droite et à gauche en 0 avec $f'_d(0) = f'_g(0) = 0$, f est dérivable en 0 et $f'(0) = 0$.
Au final f est dérivable sur $\mathbb{R}$ et $f'(x) = |x|$.
2ème méthode :
Sur $\mathbb{R}^+$, on a $f(x) = x^2$ donc $f_{\restriction \mathbb{R}^+}$ est dérivable et $f'_{\restriction \mathbb{R}^+}(x) = 2x$.
Ainsi f est dérivable en tout $a > 0$ et est dérivable à droite en 0 avec $f'_d(0) = 0$.
Sur $\mathbb{R}^-$, on a $f(x) = -x^2$ donc $f_{\restriction \mathbb{R}^-}$ est dérivable et $f'_{\restriction \mathbb{R}^-}(x) = -2x$.

Ainsi f est dérivable en tout $a < 0$ et est dérivable à gauche en 0 avec $f'_g(0) = 0$.
On peut ensuite conclure comme ci-dessus.

18.1.5 Dérivées successives

18.1.5.1 Définition

Définition

Soit $f : I \to \mathbb{R}$.
On pose $f^{(0)} = f$ appelée dérivée d'ordre 0 de f.
Si f est dérivable, on note $f^{(1)} = f'$ appelée dérivée d'ordre 1 de f.
Si de plus f' est dérivable, on pose $f^{(2)} = f'' = (f')'$ appelée dérivée d'ordre 2 (ou dérivée seconde) de f.
Ainsi, de proche en proche, si f possède une dérivée d'ordre $n \in \mathbb{N}$, notée $f^{(n)}$, et si celle-ci est dérivable, on pose $f^{(n+1)} = (f^{(n)})'$ appelée dérivée d'ordre $n+1$ de f.

Remarque Lorsqu'elle existe, la dérivée d'ordre n de f est aussi notée $\mathrm{D}^n f$.

Définition

On dit que $f : I \to \mathbb{R}$ est n fois dérivable si sa dérivée d'ordre n existe.
On dit que f est indéfiniment dérivable si pour tout $n \in \mathbb{N}$, f est n fois dérivable.

Exemple La fonction $x \mapsto \mathrm{e}^x$ est indéfiniment dérivable et

$$\forall n \in \mathbb{N}, (\mathrm{e}^x)^{(n)} = \mathrm{e}^x$$

En effet, on montre facilement par récurrence sur $n \in \mathbb{N}$, la propriété

« $x \mapsto \mathrm{e}^x$ est n fois dérivable et $(\mathrm{e}^x)^{(n)} = \mathrm{e}^x$ »

Exemple Les fonctions $x \mapsto \sin x$ et $x \mapsto \cos x$ sont indéfiniment dérivables et

$$\forall n \in \mathbb{N}, (\sin x)^{(n)} = \begin{cases} \sin x & \text{si } n = 0 \quad [4] \\ \cos x & \text{si } n = 1 \quad [4] \\ -\sin x & \text{si } n = 2 \quad [4] \\ -\cos x & \text{si } n = 3 \quad [4] \end{cases} = \sin(x + n\frac{\pi}{2})$$

De même

$$(\cos x)^{(n)} = \cos(x + n\frac{\pi}{2})$$

Exemple Les fonctions $x \mapsto \mathrm{ch}x$ et $x \mapsto \mathrm{sh}x$ sont indéfiniment dérivables et

$$\forall n \in \mathbb{N}, (\mathrm{ch}x)^{(n)} = \begin{cases} \mathrm{ch}x & \text{si } n \text{ pair} \\ \mathrm{sh}x & \text{sinon} \end{cases} \text{ et } (\mathrm{sh}x)^{(n)} = \begin{cases} \mathrm{sh}x & \text{si } n \text{ pair} \\ \mathrm{ch}x & \text{sinon} \end{cases}$$

Exemple Les fonctions polynomiales sont indéfiniment dérivables.
En effet, on montre facilement par récurrence sur $n \in \mathbb{N}$, la propriété
« les fonctions polynomiales sont n fois dérivables et la dérivée n-ième d'une fonction polynomiale est encore une fonction polynomiale »

Exemple En Pour tout $n \in \mathbb{N}$, la fonction $x \mapsto x^n$ est indéfiniment dérivable et
$(x^n)' = nx^{n-1}$, $(x^n)'' = n(n-1)x^{n-2}$,...
$\forall 0 \leqslant k \leqslant n, (x^n)^{(k)} = n(n-1)...(n-k+1)x^{n-k} = \dfrac{n!}{(n-k)!}x^{n-k}$,
$(x^n)^{(n)} = n!$ et $\forall k > n, (x^n)^{(k)} = 0$.

18.1.5.2 Opérations

Soit $n \in \mathbb{N}$.

Théorème

Si $f, g : I \to \mathbb{R}$ sont n fois dérivables alors pour tout $\lambda \in \mathbb{R}$, $\lambda.f$ et $f + g$ le sont aussi et

$$(\lambda.f)^{(n)} = \lambda.f^{(n)} \text{ et } (f+g)^{(n)} = f^{(n)} + g^{(n)}$$

dém. :
Par récurrence sur $n \in \mathbb{N}$.
La propriété est immédiate pour $n = 0$.
Supposons la propriété vraie au rang $n \geqslant 0$.
Soient $f, g : I \to \mathbb{R}$ des fonctions $n+1$ fois dérivables.
Les fonctions f et g sont en particulier n fois dérivables et donc par hypothèse de récurrence les fonctions $\lambda.f$ et $f + g$ le sont aussi avec $(\lambda.f)^{(n)} = \lambda.f^{(n)}$ et $(f+g)^{(n)} = f^{(n)} + g^{(n)}$.
Puisque les fonctions $f^{(n)}$ et $g^{(n)}$ sont dérivables, on peut affirmer que les fonctions $(\lambda.f)^{(n)}$ et $(f+g)^{(n)}$ sont dérivables par opérations sur les fonctions dérivables.
Par suite les fonctions $\lambda.f$ et $f+g$ sont $n+1$ fois dérivables avec $(\lambda.f)^{(n+1)} = (\lambda.f^{(n)})' = \lambda.f^{(n+1)}$ et

$$(f+g)^{(n+1)} = \left(f^{(n)} + g^{(n)}\right)' = f^{(n+1)} + g^{(n+1)}$$

□

Théorème

Si $f, g : I \to \mathbb{R}$ sont n fois dérivables alors fg l'est aussi et

$$(fg)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} f^{(n-k)} g^{(k)}$$

dém. :
Par récurrence sur $n \in \mathbb{N}$.
La propriété est immédiate pour $n = 0$.
Supposons la propriété vraie au rang $n \geqslant 0$.

Soient $f, g : I \to \mathbb{R}$ des fonctions $n+1$ fois dérivables
Les fonctions f et g sont en particulier n fois dérivables et donc par hypothèse de récurrence la fonction fg l'est aussi avec

$$(fg)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} f^{(n-k)} g^{(k)}$$

Pour tout $k \in \{0, \dots, n\}$, les fonctions $f^{(n-k)}$ et $g^{(k)}$ sont dérivables donc par opération sur les fonctions dérivables, la fonction $(fg)^{(n)}$ est encore dérivable.
Ainsi la fonction fg est $n+1$ fois dérivable et

$$(fg)^{(n+1)} = \left[(fg)^{(n)}\right]' = \sum_{k=0}^{n} \binom{n}{k} \left[f^{(n+1-k)} g^{(k)} + f^{(n-k)} g^{(k+1)}\right]$$

En décomposant la somme en deux et en procédant à un décalage d'indice sur la deuxième somme

$$(fg)^{(n+1)} = \sum_{k=0}^{n} \binom{n}{k} f^{(n+1-k)} g^{(k)} + \sum_{k=1}^{n+1} \binom{n}{k-1} f^{(n+1-k)} g^{(k)}$$

En adjoignant un terme nul à chaque somme

$$(fg)^{(n+1)} = \sum_{k=0}^{n+1} \binom{n}{k} f^{(n+1-k)} g^{(k)} + \sum_{k=0}^{n+1} \binom{n}{k-1} f^{(n+1-k)} g^{(k)}$$

Enfin en combinant les deux sommes et en exploitant la formule du triangle de Pascal

$$(fg)^{(n+1)} = \sum_{k=0}^{n+1} \binom{n+1}{k} f^{(n+1-k)} g^{(k)}$$

Récurrence établie
□

Exemple La fonction $x \mapsto (x^2+x+1)\mathrm{e}^{-x}$ est indéfiniment dérivable.
Par la formule de Leibniz

$$\left((x^2+x+1)\mathrm{e}^{-x}\right)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} (x^2+x+1)^{(k)} (\mathrm{e}^{-x})^{(n-k)}$$

Or pour tout $k \geqslant 3$, $(x^2+x+1)^{(k)} = 0$ donc

$$\left((x^2+x+1)\mathrm{e}^{-x}\right)^{(n)} = \binom{n}{0} (x^2+x+1)(-1)^n \mathrm{e}^{-x} + \binom{n}{1} (2x+1)(-1)^{n-1} \mathrm{e}^{-x} + \binom{n}{2} 2(-1)^{n-2} \mathrm{e}^{-x}$$

Au final

$$\left((x^2+x+1)\mathrm{e}^{-x}\right)^{(n)} = (-1)^n (x^2 - (2n-1)x + (n-1)^2) \mathrm{e}^{-x}$$

Lemme

Soit $f : I \to \mathbb{R}$. On a équivalence entre :
(i) f est $n+1$ fois dérivable ;
(ii) f est dérivable et f' est n fois dérivable.
De plus, on a alors $f^{(n+1)} = (f')^{(n)}$

dém. :
Par récurrence sur $n \in \mathbb{N}$.
Pour $n = 0$, la propriété est immédiate.
Supposons la propriété établie au rang $n \geqslant 0$.
(i) $\Rightarrow$ (ii) Supposons que f est $n+2$ fois dérivable.
La fonction est en particulier $n+1$ fois dérivable et par hypothèse de récurrence, on peut affirmer que f est dérivable et sa dérivée f' est n fois dérivable avec $(f')^{(n)} = f^{(n+1)}$. Or la fonction $f^{(n+1)}$ est encore dérivable donc $(f')^{(n)}$ est dérivable et par conséquent f' est $n+1$ fois dérivable avec $(f')^{(n+1)} = (f^{(n+1)})' = f^{(n+2)}$.
(ii) $\Rightarrow$ (i) Supposons que f est dérivable et f' est $n+1$ fois dérivable.
Par hypothèse de récurrence, on peut déjà affirmer que f est $n+1$ fois dérivable et $f^{(n+1)} = (f')^{(n)}$. Or $(f')^{(n)}$ est dérivable donc $f^{(n+1)}$ aussi et ainsi f est $n+2$ fois dérivable.
Récurrence établie.
□

Théorème

> Soient $f, g : I \to \mathbb{R}$ telles que g ne s'annule pas.
> Si f et g sont n fois dérivables alors f/g l'est aussi.

dém. :
Par récurrence sur $n \in \mathbb{N}$
Pour $n = 0$, la propriété est immédiate.
Supposons la propriété établie au rang $n \geqslant 0$.
Soient $f, g : I \to \mathbb{R}$ des fonctions $n+1$ fois dérivables avec g ne s'annulant pas.
Les fonctions f et g sont dérivables donc le rapport f/g l'est encore et on a

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

Or les fonctions f' et g sont n fois dérivables donc le produit $f'g$ l'est aussi. De même les produits fg' et g^2 sont aussi n fois dérivables et par hypothèse de récurrence le rapport $(f'g - fg')/g^2$ est n fois dérivable.
En vertu du lemme précédent, on peut affirmer que le rapport f/g est $n+1$ fois dérivable.
Récurrence établie.
□

Exemple Les fonctions rationnelles sont indéfiniment dérivables.

Exemple La fonction $x \mapsto 1/x$ est indéfiniment dérivable et

$$\left(\frac{1}{x}\right)' = \frac{-1}{x^2}, \left(\frac{1}{x}\right)'' = \frac{2}{x^3}, \ldots, \left(\frac{1}{x}\right)^{(n)} = (-1)^{n-1}\frac{n!}{x^{n+1}}$$

Exemple La fonction $x \mapsto \ln x$ est indéfiniment dérivable car sa dérivée est la fonction inverse qui est indéfiniment dérivable.

Exemple Les fonctions $x \mapsto \tan x$ et $x \mapsto \mathrm{th} x$ sont indéfiniment dérivables par rapport de fonctions indéfiniment dérivables.

Théorème

Soient $f : I \to \mathbb{R}$ et $\varphi : J \to \mathbb{R}$ vérifiant $f(I) \subset J$.
Si f et φ sont n fois dérivables alors $\varphi \circ f$ l'est aussi.

dém. :
Par récurrence sur $n \in \mathbb{N}$
Pour $n = 0$, la propriété est immédiate.
Supposons la propriété établie au rang $n \geqslant 0$.
Soient $f : I \to \mathbb{R}$ et $\varphi : J \to \mathbb{R}$ vérifiant $f(I) \subset J$ des fonctions $n + 1$ fois dérivables.
Les fonctions f et φ sont dérivables donc la composée $\varphi \circ f$ l'est encore et on a

$$(\varphi \circ f)' = f' \times \varphi' \circ f$$

Or les fonctions φ' et f sont n fois dérivables donc par hypothèse de récurrence, la fonction $\varphi' \circ f$ est aussi n fois dérivable. De plus la fonction f' est n fois dérivable donc par produit la fonction $(\varphi \circ f)'$ est n fois dérivable.
En vertu du lemme précédent, on peut affirmer que la composée $\varphi \circ f$ est $n + 1$ fois dérivable.
Récurrence établie.
□

Exemple Pour tout $\alpha \in \mathbb{R}$, la fonction $x \mapsto x^\alpha = \mathrm{e}^{\alpha \ln x}$ est indéfiniment dérivable.

$$(x^\alpha)' = \alpha x^{\alpha-1}, \ldots, (x^\alpha)^{(n)} = \alpha(\alpha - 1)...(\alpha - n + 1)x^{\alpha-n}$$

Théorème

Soit $f : I \to J$ une bijection continue.
Si f est n fois dérivable (avec $n \geqslant 1$) et si dérivée première ne s'annule pas alors $f^{-1} : J \to I$ est n fois dérivable.

dém. :
Par récurrence sur $n \in \mathbb{N}^\star$
Pour $n = 1$, la propriété est déjà connue.
Supposons la propriété établie au rang $n \geqslant 1$.
Soit $f : I \to J$ une bijection continue $n + 1$ fois dérivable dont la dérivée ne s'annule pas.
Puisque f est en particulier n fois dérivable, on peut déjà affirmer par hypothèse de récurrence que f^{-1} est n fois dérivable. En particulier f^{-1} est dérivable et on sait

$$(f^{-1})' = \frac{1}{f' \circ f^{-1}}$$

Or f' et f^{-1} sont n fois dérivables donc par opérations $(f^{-1})'$ est n fois dérivable.
En vertu du lemme précédent, on peut affirmer que la bijection réciproque f^{-1} est $n + 1$ fois dérivable.
Récurrence établie.
□

Exemple Les fonctions $\arccos$ et $\arcsin$ sont indéfiniment dérivables sur $]-1, 1[$ car bijections réciproques des fonctions indéfiniment dérivables $\cos_{\restriction]0,\pi[}$ et $\sin_{\restriction]-\pi/2,\pi/2[}$ dont les dérivées ne s'annulent pas.

Exemple La fonction $\arctan$ est indéfiniment dérivable sur $\mathbb{R}$ car bijection réciproque de la fonction indéfiniment dérivable $\tan_{\restriction]-\pi/2,\pi/2[}$ dont la dérivée ne s'annule pas.

18.1.6 Classe d'une fonction

18.1.6.1 Définition

Définition

On dit qu'une fonction $f : I \to \mathbb{R}$ est de classe $\mathcal{C}^n$ (avec $n \in \mathbb{N}$) si sa dérivée d'ordre n existe et est continue.
On note $\mathcal{C}^n(I, \mathbb{R})$ l'ensemble des fonctions réelles de classe $\mathcal{C}^n$ définies sur I.

Exemple Dire que $f : I \to \mathbb{R}$ est de classe $\mathcal{C}^0$ signifie que f est continue.

Exemple Dire que $f : I \to \mathbb{R}$ est de classe $\mathcal{C}^1$ signifie que f est dérivable et que sa dérivée est continue.

Exemple Soit $f : \mathbb{R} \to \mathbb{R}$ définie par

$$f(x) = \begin{cases} x^2 \sin(1/x) & \text{si } x \neq 0 \\ 0 & \text{sinon} \end{cases}$$

On a déjà vu est f est dérivable sur $\mathbb{R}$ avec $f'(0) = 0$
Cependant pour $x \neq 0$, $f'(x) = 2x \sin(1/x) - \cos(1/x)$ diverge quand $x \to 0$.
Ainsi la dérivée de f n'est pas continue en 0 ; cette fonction n'est pas de classe $\mathcal{C}^1$.

Remarque Puisqu'une fonction dérivable est continue, une fonction $n + 1$ fois dérivable est de classe $\mathcal{C}^n$. En particulier les fonctions de classe $\mathcal{C}^{n+1}$ sont a fortiori de classe $\mathcal{C}^n$. Ainsi
$\mathcal{C}^{n+1}(I, \mathbb{R}) \subset \mathcal{C}^n(I, \mathbb{R})$.

Définition

On dit qu'une fonction $f : I \to \mathbb{R}$ est de classe $\mathcal{C}^\infty$ si elle est de classe $\mathcal{C}^n$ pour tout $n \in \mathbb{N}$.
On note $\mathcal{C}^\infty(I, \mathbb{R})$ l'ensemble de ces fonctions.
Ainsi

$$\mathcal{C}^\infty(I, \mathbb{R}) = \bigcap_{n \in \mathbb{N}} \mathcal{C}^n(I, \mathbb{R})$$

Remarque Les fonctions de classe $\mathcal{C}^\infty$ correspondent exactement aux fonctions indéfiniment dérivables.

Exemple Les fonctions polynomiales, les fonctions rationnelles, les fonctions usuelles sont de classe $\mathcal{C}^\infty$ là où elles sont dérivables et a fortiori elles y sont aussi de classe $\mathcal{C}^n$.

18.1.6.2 Opérations

Lemme

Soient $f : I \to \mathbb{R}$ et $n \in \mathbb{N}$.
On a équivalence entre :
(i) f est de classe $\mathcal{C}^{n+1}$;
(ii) f est dérivable et f' est de classe $\mathcal{C}^n$.

dém. :
(i) $\Rightarrow$ (ii) Si f est de classe $\mathcal{C}^{n+1}$ alors f est $n+1$ fois dérivable et $f^{(n+1)}$ est continue.
Par suite f est dérivable et f' est n fois dérivable avec $(f')^{(n)} = f^{(n+1)}$ continue.
Ainsi f' est de classe $\mathcal{C}^n$.
(ii) $\Rightarrow$ (i) Si f est dérivable et f' de classe $\mathcal{C}^n$ alors f' est n fois dérivable et donc f est $n+1$ fois dérivable. De plus, $f^{(n+1)} = (f')^{(n)}$ est continue et donc f est de classe $\mathcal{C}^{n+1}$.
Les théorèmes qui suivent se justifient par récurrence sur $n \in \mathbb{N}$ et s'étendent alors $n = \infty$:

□

Théorème

Si f et $g : I \to \mathbb{R}$ sont de classe $\mathcal{C}^n$ alors pour tout $\lambda \in \mathbb{R}$ les fonctions $\lambda.f, f+g, fg$ le sont aussi.
Si de plus g ne s'annule pas alors le rapport f/g est aussi $\mathcal{C}^n$.

Théorème

Soient $f : I \to \mathbb{R}$ et $\varphi : J \to \mathbb{R}$ vérifiant $f(I) \subset J$.
Si f et φ sont $\mathcal{C}^n$ alors la composée $\varphi \circ f$ l'est aussi.

Théorème

Si $f : I \to J$ est une bijection de classe $\mathcal{C}^n$ (avec $n \geqslant 1$) et si sa dérivée ne s'annule pas alors sa bijection réciproque $f^{-1} : J \to I$ est aussi de classe $\mathcal{C}^n$.

18.2 Théorème des accroissements finis

18.2.1 Extremum local

Définition

On dit qu'une fonction $f : I \to \mathbb{R}$ admet un minimum local en $a \in I$ s'il existe $\alpha > 0$ vérifiant

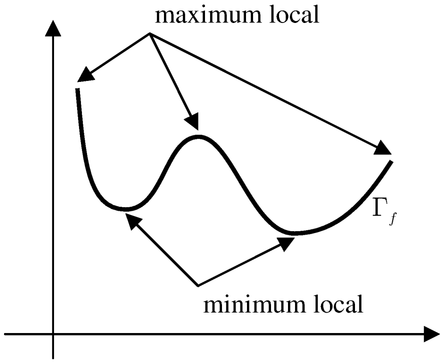
$$\forall x \in I \cap [a-\alpha, a+\alpha]\,, f(x) \geqslant f(a)$$

On dit qu'une fonction $f : I \to \mathbb{R}$ admet un maximum local en $a \in I$ s'il existe $\alpha > 0$ vérifiant

$$\forall x \in I \cap [a-\alpha, a+\alpha]\,, f(x) \leqslant f(a)$$

Dans les deux cas, on parle d'extremum local en a.

Exemple



Exemple Les extremum (globaux) de f sont a fortiori des extremum locaux.

Théorème

Si $f : I \to \mathbb{R}$ admet un extremum local en $a \in I$ qui n'est pas extrémité de I et si f est dérivable en a alors $f'(a) = 0$.

dém. :
Cas a minimum local
Il existe $\alpha > 0$ tel que

$$\forall x \in I \cap [a - \alpha, a + \alpha], f(x) \geqslant f(a)$$

Quitte à réduire α, on peut aussi supposer $[a - \alpha, a + \alpha] \subset I$ car a n'est pas extrémité de I.
Quand $h \to 0^+$,
Pour h assez petit, $h \leqslant \alpha$ et alors $f(a + h) \geqslant f(a)$ d'où

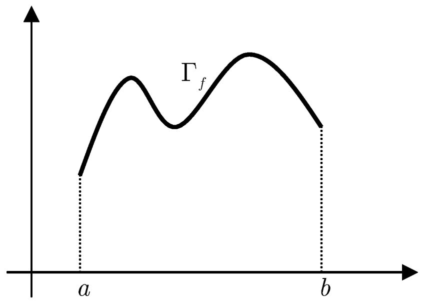
$$\frac{1}{h}(f(a+h) - f(a)) \geqslant 0$$

En passant à la limite, on obtient $f'(a) \geqslant 0$.
Quand $h \to 0^-$,
Pour h assez petit, $f(a + h) \geqslant f(a)$ puis

$$\frac{1}{h}(f(a+h) - f(a)) \leqslant 0$$

□
car ici $h < 0$.
En passant à la limite, on obtient $f'(a) \leqslant 0$.
Finalement $f'(a) = 0$.
Cas a maximum local
Il suffit d'exploiter le résultat qui précède en l'appliquant à la fonction $-f$.
(!)Pour appliquer ce résultat il ne faut pas que a soit extrémité de I.

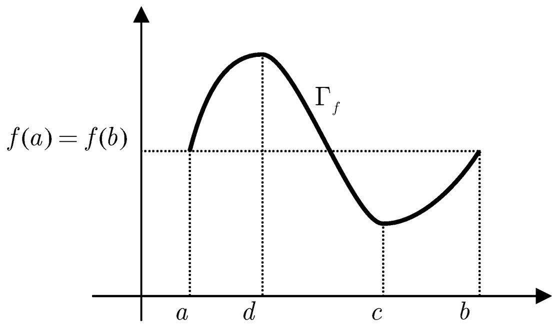
Remarque Si a est minimum local de f et est une extrémité gauche de I (resp. extrémité droite de I), le raisonnement qui précède donne $f'(a) \geqslant 0$ (resp. $f'(a) \leqslant 0$).



18.2.2 Théorème de Rolle

Théorème

Soient $a, b \in \mathbb{R}$ avec $a < b$.
Si $f : [a, b] \to \mathbb{R}$ est continue sur $[a, b]$, dérivable sur $]a, b[$ et si $f(a) = f(b)$ alors il existe $c \in\,]a, b[$ vérifiant $f'(c) = 0$.



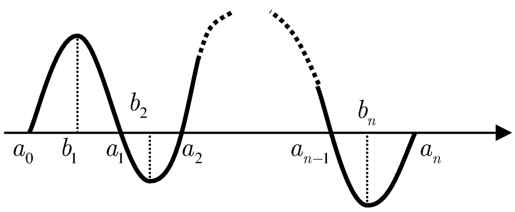
dém. :
Puisque la fonction f est continue sur le segment, elle y admet un minimum et un maximum.
Ainsi il existe $c, d \in [a, b]$ vérifiant

$$\forall x \in [a, b]\,, f(c) \leqslant f(x) \leqslant f(d)$$

Si c et d sont tous deux extrémités de I alors la condition $f(a) = f(b)$ entraîne que f est constante sur $[a, b]$. Par suite n'importe quel élément de $]a, b[$ annule la dérivée de f.
Sinon, l'un au moins de c ou de d n'est pas extrémité de l'intervalle $[a, b]$ et puisque f admet un extremum local en cet élément et y est dérivable, la dérivée de f s'y annule.
□

Exemple Soit $n \in \mathbb{N}$ et $f : I \to \mathbb{R}$ une fonction n fois dérivable.
Montrons que si f s'annule en $(n + 1)$ points alors sa dérivée d'ordre n s'annule au moins une fois.
Commençons par un résultat intermédiaire.
Supposons qu'une fonction dérivable $f : I \to \mathbb{R}$ s'annule aux points $a_0 < a_1 < ... < a_n$.
Pour chaque $i \in \{1, \ldots, n\}$, on peut appliquer le théorème de Rolle à f sur le segment $[a_{i-1}, a_i]$ car f est continue sur $[a_{i-1}, a_i]$, dérivable sur $]a_{i-1}, a_i[$ et vérifie $f(a_{i-1}) = 0 = f(a_i)$.
Ainsi, il existe $b_i \in\,]a_{i-1}, a_i[$ vérifiant $f'(b_i) = 0$.

Puisque les $b_1, \dots, b_n$ vérifient $a_0 < b_1 < a_1 < b_2 < \dots < a_{n-1} < b_n < a_n$, ces éléments sont deux à deux distincts.
Ainsi, nous venons de démontrer que si une fonction dérivable $f : I \to \mathbb{R}$ s'annule $n+1$ fois, sa dérivée s'annule au moins n fois.
Par suite, si f est n fois dérivable et s'annule $n+1$ fois, sa dérivée s'annule au moins n fois, sa dérivée seconde s'annule au moins $n-1$ et plus généralement, pour $k \in \{0, \dots, n\}$, $f^{(k)}$ s'annule au moins $n+1-k$ fois.
En particulier $f^{(n)}$ s'annule au moins une fois.



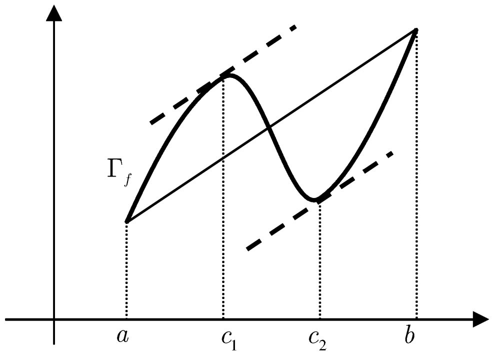
18.2.3 Théorème des accroissements finis

Théorème

> Soient $a, b \in \mathbb{R}$ avec $a < b$.
> Si $f : [a, b] \to \mathbb{R}$ est continue sur $[a, b]$ et dérivable sur $]a, b[$ alors il existe $c \in]a, b[$ vérifiant
> $$f(b) - f(a) = f'(c)(b - a)$$

dém. :
Posons
$$K = \frac{f(b) - f(a)}{b - a}$$
et considérons la fonction auxiliaire $\varphi : [a, b] \to \mathbb{R}$ définie par
$$\varphi(x) = f(x) - K(x - a)$$
La fonction φ est continue sur $[a, b]$ et dérivable sur $]a, b[$ par opérations sur de telles fonctions.
Par construction $\varphi(a) = f(a)$ et $\varphi(b) = f(b) - K(b - a) = f(a) = \varphi(a)$.
On peut donc appliquer le théorème de Rolle à la fonction φ sur $[a, b]$ et donc il existe $c \in]a, b[$ vérifiant $\varphi'(c) = 0$. Or $\varphi'(x) = f'(x) - K$ sur $]a, b[$ et donc $\varphi'(c) = 0$ donne $f'(c) = K$.
□

Remarque L'égalité
$$f'(c) = \frac{f(b) - f(a)}{b - a}$$
signifie que la tangente en c à f est parallèle à la droite joignant les points d'abscisses a et b de f.
D'un point de vue cinématique, la vitesse moyenne sur $[a, b]$ est égale à la vitesse instantanée pour un instant c bien choisi.



Exemple Montrons que pour tout $x > 0$, on a l'encadrement

$$\frac{x}{x+1} < \ln(1+x) < x$$

Pour cela appliquons le théorème des accroissements finis à la fonction $t \mapsto \ln(1+t)$ sur $[0, x]$.
Cette fonction est continue sur $[0, x]$ et dérivable sur $]0, x[$ car de classe $\mathcal{C}^\infty$ sur $]-1, +\infty[$.
Par le théorème des accroissements finis, il existe $c \in]0, x[$ vérifiant

$$\ln(1+x) - \ln(1) = \frac{x}{1+c}$$

Puisque $0 < c < x$, on a $\dfrac{1}{1+x} < \dfrac{1}{1+c} < 1$ puis $\dfrac{x}{x+1} < \ln(1+x) < x$.

18.2.4 Applications du TAF

18.2.4.1 Inégalités des accroissements finis

Théorème

Soit $f : I \to \mathbb{R}$ dérivable.
S'il existe $m, M \in \mathbb{R}$ vérifiant

$$\forall x \in I, m \leqslant f'(x) \leqslant M$$

alors pour tout $a, b \in I$ tels que $a \leqslant b$, on a

$$m(b-a) \leqslant f(b) - f(a) \leqslant M(b-a)$$

dém. :
Cas $a = b$
C'est immédiat
Cas $a < b$
La fonction f est continue sur $[a, b]$ et dérivable sur $]a, b[$ car dérivable sur I.

Par l'égalité des accroissements finie, il existe $c \in]a, b[$ vérifiant $f(b) - f(a) = f'(c)(b - a)$.
Puisque $m \leqslant f'(c) \leqslant M$ et $b - a \geqslant 0$, on en déduit $m(b - a) \leqslant f(b) - f(a) \leqslant M(b - a)$.
□

Remarque D'un point de vue cinématique, l'inégalité des accroissements finis sous la forme

$$m \leqslant \frac{f(b) - f(a)}{b - a} \leqslant M$$

signifie que la vitesse moyenne est comprise entre l'inf et le sup des vitesses instantanées.

Corollaire

Soit $f : I \to \mathbb{R}$ dérivable.
S'il existe $M \in \mathbb{R}^+$ vérifiant

$$\forall x \in I, |f'(x)| \leqslant M$$

alors f est M lipschitzienne i.e.

$$\forall x, y \in I, |f(y) - f(x)| \leqslant M |y - x|$$

dém. :
Puisque pour tout $x \in I$, $-M \leqslant f'(x) \leqslant M$, l'inégalité des accroissements finis donne que pour tout $x \leqslant y$,

$$-M(y - x) \leqslant f(y) - f(x) \leqslant M(y - x)$$

et donc

$$|f(y) - f(x)| \leqslant M |y - x|$$

Cette relation vaut aussi pour $y \leqslant x$, il suffit d'échanger x et y pour l'obtenir.
□

Exemple Les fonctions $\sin, \cos, \arctan$ sont des fonctions 1 lipschitziennes car leurs dérivées sont bornées par 1.

Exemple Soit $f : [a, b] \to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$. Montrer que f est lipschitzienne.
Puisque f est de classe $\mathcal{C}^1$, f est dérivable et f' est continue.
Puisque f' est continue sur le segment $[a, b]$, on peut affirmer que f' est bornée par un certain $M \in \mathbb{R}^+$ sur ce segment et alors f est M-lipschitzienne en vertu de l'inégalité des accroissements finis.

18.2.4.2 Variations d'une fonction dérivable

Théorème

Soit $f : I \to \mathbb{R}$ une fonction dérivable.
La fonction f est croissante si, et seulement si,

$$\forall x \in I, f'(x) \geqslant 0$$

La fonction f est décroissante si, et seulement si,

$$\forall x \in I, f'(x) \leqslant 0$$

La fonction f est constante si, et seulement si,

$$\forall x \in I, f'(x) = 0$$

dém. :
Cas de la croissance :
($\Rightarrow$) Supposons la fonction f croissante.
Pour tout $x \in I$, $f'(x)$ est la limite quand $h \to 0$ (avec $h \neq 0$) du taux de variation $\frac{1}{h}\left(f(x+h) - f(x)\right)$.
En vertu de la croissance de f, ce taux de variation est positif pour tout h et donc en passant à la limite on obtient $f'(x) \geqslant 0$.
($\Leftarrow$) Supposons

$$\forall x \in I, f'(x) \geqslant 0$$

Soient $a < b$ deux éléments de I.
La fonction f est continue sur $[a, b]$ et dérivable sur $]a, b[$ car dérivable sur I.
Par le théorème des accroissements finis, il existe $c \in]a, b[$ vérifiant $f(b) - f(a) = f'(c)(b - a)$.
Or $b - a \geqslant 0$ et $f'(c) \geqslant 0$ donc $f(b) \geqslant f(a)$.
Ainsi f est croissante.
Cas de la décroissance :
Il suffit d'appliquer ce qui précède en considérant la fonction $-f$.
Cas de la constante :
Il suffit d'observer que les fonctions constantes sur I sont les fonctions à la fois croissantes et décroissantes.
□

Exemple Montrons que pour tout $x \in \mathbb{R}$, on a

$$\cos x \geqslant 1 - \frac{x^2}{2}$$

Introduisons la fonction auxiliaire

$$\varphi : x \mapsto \cos x - 1 + x^2/2$$

La fonction φ est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$.
$\varphi'(x) = -\sin x + x$ et $\varphi''(x) = 1 - \cos x$.
Puisque $\varphi'' \geqslant 0$, la fonction φ' est croissante.
Puisque $\varphi'(0) = 0$, la fonction φ' est négative sur $]-\infty, 0]$ et positive sur $[0, +\infty[$.
Par suite la fonction φ est décroissante sur $]-\infty, 0]$ et croissante sur $[0, +\infty[$.
La fonction φ admet donc un minimum en 0 et puisque $\varphi(0) = 0$, on peut conclure que φ est positive.
Ainsi pour tout $x \in \mathbb{R}$, $\cos x - 1 + x^2/2 \geqslant 0$ puis $\cos x \geqslant 1 - x^2/2$.

Théorème

Soient $a < b$ les extrémités de l'intervalle I.
Si $f : I \to \mathbb{R}$ est continue sur I et dérivable sur l'intervalle ouvert $]a, b[$ avec

$$\forall x \in\,]a, b[\,,\, f'(x) > 0 \text{ (resp. } f'(x) < 0 \text{)}$$

alors f est strictement croissante (resp. strictement décroissante)

dém. :
Cas de la stricte croissance :
Soient $x < y$ deux éléments de I.
La fonction f est continue sur $[x, y]$ car continue sur I et dérivable sur $]x, y[$ car dérivable sur $]a, b[$.
Par le théorème des accroissements finis, il existe $c \in\,]x, y[$ vérifiant $f(y) - f(x) = f'(c)(y - x)$.
Or $y - x > 0$ et $f'(c) > 0$ donc $f(y) \geqslant f(x)$.
Cas de la stricte décroissance :
Il suffit d'appliquer ce qui précède en considérant la fonction $-f$.
□

Exemple La fonction

$$f : x \mapsto \sqrt{\mathrm{e}^x - 1}$$

est strictement croissante sur $\mathbb{R}^+$.
En effet la fonction f est continue sur $\mathbb{R}^+$, dérivable sur $]0, +\infty[$ et pour tout $x > 0$,

$$f'(x) = \frac{\mathrm{e}^x}{2\sqrt{\mathrm{e}^x - 1}} > 0$$

On peut aussi affirmer que f est strictement croissante par composition des fonctions strictement croissantes $x \mapsto \mathrm{e}^x - 1$ et $t \mapsto \sqrt{t}$.

Corollaire

Soit $f : I \to \mathbb{R}$ continue.
Si $f'(x) > 0$ (resp. $f'(x) < 0$) sur I sauf peut-être en un nombre fini de points où $f'(x)$ s'annule ou bien n'est pas défini, alors f est strictement croissante (resp. strictement décroissante).

dém. :
On découpe l'intervalle I en intervalles J contigus isolant les points où $f'(x)$ « pose problème »et on applique le théorème ci-dessus sur chacun des intervalles J avant de recoller les strictes monotonies ainsi obtenues.
□

Exemple La fonction $x \mapsto x^3$ est strictement croissante.
En effet cette fonction est continue sur $\mathbb{R}$, dérivable sur $\mathbb{R}$ et $f'(x) = 3x^2 > 0$ sauf pour $x = 0$.

18.2.4.3 Obtention de la dérivabilité par limite

Théorème

Soit $f : I \to \mathbb{R}$ et $a \in I$.
On suppose que f est dérivable sur $I \backslash \{a\}$ et est continue en a.
Si $\lim\limits_{x \to a} f'(x) = \ell$ alors f est dérivable en a et $f'(a) = \ell$.
Si $\lim\limits_{x \to a} f'(x) = +\infty$ (ou $-\infty$) alors f n'est pas dérivable en a mais présente une tangente verticale en a.

dém. :
Supposons $\lim\limits_{x \to a} f'(x) = \ell \in \bar{\mathbb{R}}$.
Soit $x \in I$ avec $x \neq a$. En appliquant le théorème des accroissements finis entre a et x, on obtient l'existence de c_x strictement compris entre a et x vérifiant $f(x) - f(a) = f'(c_x)(x - a)$.
On a alors

$$\frac{f(x) - f(a)}{x - a} = f'(c_x)$$

Quand $x \to a$ (avec $x \neq a$),
Par encadrement c_x tend vers a et par composition de limite

$$\frac{f(x) - f(a)}{x - a} \to \ell$$

Si $\ell \in \mathbb{R}$ alors f est dérivable en a et $f'(a) = \ell$.
Si $\ell = \pm\infty$ alors f n'est pas dérivable en a mais présente une tangente verticale en a.
□

Remarque Si la limite de $f'(x)$ quand $x \to a$ n'existe pas, on ne peut rien dire.

Attention : Pour exploiter ce résultat il faut préalablement observer la continuité de f en a.

Exemple Considérons la fonction $f : x \mapsto \sqrt{\mathrm{e}^x - 1}$ définie et continue sur $\mathbb{R}^+$.
La fonction f est dérivable sur $]0, +\infty[$ et

$$f'(x) = \frac{\mathrm{e}^x}{2\sqrt{\mathrm{e}^x - 1}} \xrightarrow[x \to 0^+]{} +\infty$$

Par suite f n'est pas dérivable en 0 mais y présente une tangente verticale.

Corollaire

Soit $f : [a, b] \to \mathbb{R}$ une fonction continue et de classe $\mathcal{C}^1$ sur $]a, b]$.
Si f' converge en a^+ alors f est de classe $\mathcal{C}^1$ sur $[a, b]$.

dém. :
Par le théorème qui précède, f est dérivable en a et $f'(a) = \lim\limits_{x \to a^+} f'(x)$ ce qui assure la continuité de f' en a.
□

Exemple Considérons la fonction $f :]0, +\infty[\to \mathbb{R}$ définie par $f(x) = x^2 \ln x$.
La fonction f est de classe $\mathcal{C}^1$ sur $]0, +\infty[$.
Quand $x \to 0^+$, $f(x) \to 0$.
On prolonge f par continuité en 0 en posant $f(0) = 0$.
Quand $x \to 0^+$, $f'(x) = 2x \ln x + x \to 0$.
Par le théorème du prolongement $\mathcal{C}^1$, f est de classe $\mathcal{C}^1$ sur $[0, +\infty[$ et $f'(0) = 0$.

Attention : Dans ce dernier raisonnement, on n'a pas prolongé f' par continuité en 0 car on ne peut pas poser la valeur d'une dérivée. En revanche, le théorème du prolongement $\mathcal{C}^1$ justifie la dérivabilité au point considéré et la continuité de la dérivée en ce point.

18.3 Extension aux fonctions complexes

Les définitions qui suivent prolongent celles vues dans le cadre réel.

18.3.1 Fonction dérivée

Définition

On dit que $f : I \to \mathbb{C}$ est dérivable en a si le taux de variation complexe

$$\frac{1}{h}\left(f(a+h) - f(a)\right)$$

converge quand $h \to 0$ (avec $h \neq 0$)
La limite obtenue est alors appelé nombre dérivé de f en a et est noté $f'(a)$.

Définition

On dit que $f : I \to \mathbb{C}$ est dérivable si f est dérivable en tout point de I et on peut alors introduire sa fonction dérivée $f' : I \to \mathbb{C}$ parfois encore notée Df.

Proposition

Si f est dérivable en $a \in I$ alors

$$f(t) = f(a) + f'(a)(t-a) + o(t-a)$$

quand $t \to a$.
En particulier f est continue en a.

dém. :
Analogue à celle présentée dans le cadre réel.
□

Théorème

Si f et $g : I \to \mathbb{C}$ sont dérivables alors pour tout $\lambda \in \mathbb{C}$, les fonctions $\lambda.f$, $f+g$, fg et $\bar{f}$ sont aussi dérivables et on a les relations

$$(\lambda.f)' = \lambda.f', (f+g)' = f'+g', (fg)' = f'g + fg' \text{ et } (\bar{f})' = \overline{f'}$$

Si de plus la fonction g ne s'annule pas alors le rapport f/g est dérivable et

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

dém. :
Analogue à celle présentée dans le cadre réel.
□

Exemple Les fonctions polynomiales et rationnelles à coefficients complexes sont dérivables.

Exemple La fonction

$$t \mapsto \frac{t+i}{t-i}$$

est une fonction définie et dérivable sur $\mathbb{R}$. De plus

$$\left(\frac{t+i}{t-i}\right)' = \left(\frac{u}{v}\right)' = \frac{u'v - uv'}{v^2} = \frac{2i}{(t-i)^2}$$

Corollaire

On a équivalence entre :
(i) la fonction $f : I \to \mathbb{C}$ est dérivable ;
(ii) les fonctions réelles $\text{Re}(f)$ et $\text{Im}(f)$ sont dérivables.
De plus, si tel est le cas,

$$\text{Re}(f') = (\text{Re}f)' \text{ et } \text{Im}(f') = (\text{Im}f)'$$

dém. :
(i) $\Rightarrow$ (ii) via

$$\text{Re}f = \frac{1}{2}(f+\bar{f}) \text{ et } \text{Im}f = \frac{1}{2i}(f+\bar{f})$$

(ii) $\Rightarrow$ (i) via

$$f = \text{Re}(f) + i.\text{Im}(f)$$

□

Théorème

Soit $f : I \to \mathbb{R}$ et $\varphi : J \to \mathbb{C}$ vérifiant $f(I) \subset J$.
Si f et φ sont dérivables alors $\varphi \circ f$ l'est aussi et

$$(\varphi \circ f)' = f' \times \varphi' \circ f$$

dém. :
Analogue à celle présentée dans le cadre réel.
□

Proposition

Soit $f : I \to \mathbb{C}$.
Si f est dérivable alors $t \mapsto \mathrm{e}^{f(t)}$ l'est aussi et

$$(\mathrm{e}^{f(t)})' = f'(t)\mathrm{e}^{f(t)}$$

dém. :
Posons $a(t) = \mathrm{Re}(f(t))$ et $b(t) = \mathrm{Im}(f(t))$. Les fonctions réelles a et b sont dérivables.
Puisque $\mathrm{e}^{f(t)} = \mathrm{e}^{a(t)} (\cos b(t) + i.\sin b(t))$, la fonction f est dérivable et

$$\left(\mathrm{e}^{f(t)}\right)' = a'(t)\mathrm{e}^{a(t)} (\cos b(t) + i.\sin b(t)) + b'(t)\mathrm{e}^{a(t)} (-\sin b(t) + i\cos b(t))$$

Ainsi

$$\left(\mathrm{e}^{f(t)}\right)' = (a'(t) + ib'(t))\, \mathrm{e}^{a(t)} (\cos b(t) + i.\sin b(t)) = f'(t)\mathrm{e}^{if(t)}$$

□

Exemple Pour $\alpha \in \mathbb{C}$, la fonction $t \mapsto \mathrm{e}^{\alpha t}$ est dérivable et $(\mathrm{e}^{\alpha t})' = \alpha \mathrm{e}^{\alpha t}$.

18.3.2 Classe d'une fonction complexe

Définition

Tout comme pour les fonctions réelles on définit, lorsqu'elle existe, la dérivée d'ordre $n \in \mathbb{N}$ d'une fonction complexe f. Celle-ci est notée $f^{(n)}$.

Proposition

On a équivalence entre :
(i) la fonction $f : I \to \mathbb{C}$ est n fois dérivable ;
(ii) les fonctions réelles $\mathrm{Re}(f)$ et $\mathrm{Im}(f)$ sont n fois dérivables.
De plus, si tel est le cas

$$\mathrm{Re}(f^{(n)}) = \mathrm{Re}(f)^{(n)} \text{ et } \mathrm{Im}(f^{(n)}) = \mathrm{Im}(f)^{(n)}$$

dém. :
Par récurrence sur $n \in \mathbb{N}$.
Pour $n = 0$, c'est immédiat car toute fonction f est 0 fois dérivable et $f^{(0)} = f$.
Supposons la propriété établie au rang.
(i) $\Rightarrow$ (ii) Si la fonction $f : I \to \mathbb{C}$ est $n + 1$ fois dérivable alors f est n dérivable. Par hypothèse de récurrence, les fonctions $\mathrm{Re}(f)$ et $\mathrm{Im}(f)$ sont alors n fois dérivables avec $\mathrm{Re}(f)^{(n)} = \mathrm{Re}(f^{(n)})$ et $\mathrm{Im}(f)^{(n)} = \mathrm{Im}(f^{(n)})$. Puisque f est $n + 1$ fois dérivable, la dérivée d'ordre n de f est encore dérivable et ainsi sa partie réelle l'est aussi. Puisque $\mathrm{Re}(f)^{(n)} = \mathrm{Re}(f^{(n)})$, la fonction réelle $\mathrm{Re}(f)$ est $n + 1$ fois

dérivable et de plus $(\mathrm{Re}(f))^{(n+1)} = \mathrm{Re}(f^{(n)})' = \mathrm{Re}((f^{(n)})') = \mathrm{Re}(f^{(n+1)})$. On procède de même pour la partie imaginaire de f.
(ii) $\Rightarrow$ (i) Supposons les fonctions réelles $\mathrm{Re}(f)$ et $\mathrm{Im}(f)$ $n+1$ fois dérivables. Ces fonctions sont a fortiori n fois dérivables et par hypothèse de récurrence f est n fois dérivable avec $\mathrm{Re}(f^{(n)}) = \mathrm{Re}(f)^{(n)}$ et $\mathrm{Im}(f^{(n)}) = \mathrm{Im}(f)^{(n)}$. Puisque $f^{(n)} = \mathrm{Re}(f)^{(n)} + i\mathrm{Im}(f)^{(n)}$ avec $\mathrm{Re}(f)^{(n)}$ et $\mathrm{Im}(f)^{(n)}$ dérivables, on obtient $f^{(n)}$ dérivable puis f $n+1$ fois dérivable.
Récurrence établie.
□

Définition

On dit que $f : I \to \mathbb{C}$ est de classe $\mathcal{C}^n$ (avec $n \in \mathbb{N}$) si f est n fois dérivable et si $f^{(n)}$ est continue.
On dit que $f : I \to \mathbb{C}$ est de classe $\mathcal{C}^\infty$ si f est $\mathcal{C}^n$ pour tout $n \in \mathbb{N}$.
Pour $n \in \mathbb{N} \cup \{\infty\}$, on note $\mathcal{C}^n(I, \mathbb{C})$ l'ensemble des fonctions complexes définies et de classe $\mathcal{C}^n$ sur I.

Proposition

On a équivalence entre :
(i) la fonction $f : I \to \mathbb{C}$ est de classe $\mathcal{C}^n$;
(ii) les fonctions réelles $\mathrm{Re}(f)$ et $\mathrm{Im}(f)$ sont de classe $\mathcal{C}^\infty$.

dém. :
C'est immédiate sachant $\mathrm{Re}(f^{(n)}) = \mathrm{Re}(f)^{(n)}$ et $\mathrm{Im}(f^{(n)}) = \mathrm{Im}(f)^{(n)}$.
□

Exemple Soit $f : I \to \mathbb{C}$ de classe $\mathcal{C}^n$.
Si ne s'annule pas alors la fonction $|f| : I \to \mathbb{R}$ est de classe $\mathcal{C}^n$.
En effet $|f| = \left(\mathrm{Re}(f)^2 + \mathrm{Im}(f)^2\right)^{1/2}$ et la fonction $\mathrm{Re}(f)^2 + \mathrm{Im}(f)^2$ de classe $\mathcal{C}^n$ à valeurs strictement positives et la fonction $\sqrt{.}$ est de classe $\mathcal{C}^\infty$ sur $]0, +\infty[$.

Les résultats qui suivent, énoncés pour $n \in \mathbb{N} \cup \{\infty\}$, peuvent être obtenus raisonnant via les parties réelles et imaginaires.

Théorème

Si f et $g : I \to \mathbb{C}$ sont de classe $\mathcal{C}^n$ alors pour tout $\lambda \in \mathbb{C}$, les fonctions $\lambda.f, f+g, fg$ et $\bar{f}$ le sont aussi.
Si de plus g ne s'annule pas alors le rapport f/g est aussi de classe $\mathcal{C}^n$.

Exemple Les fonctions polynomiales et les fonctions rationnelles sont $\mathcal{C}^\infty$.

Théorème

Soient $f : I \to \mathbb{R}$ et $\varphi : J \to \mathbb{C}$ telles que $f(I) \subset J$.
Si f et γ sont de classe $\mathcal{C}^n$ alors $\varphi \circ f$ l'est aussi.

Proposition

Si $f : I \to \mathbb{C}$ est de classe $\mathcal{C}^n$ alors la fonction $\mathrm{e}^f : t \mapsto \mathrm{e}^{f(t)}$ l'est aussi.

Exemple Calculons la dérivée n-ième de la fonction réelle $t \mapsto \cos(t)\mathrm{e}^t$.

$$(\cos(t)\mathrm{e}^t)^{(n)} = (\mathrm{Re}(\mathrm{e}^{(1+i)t}))^{(n)} = \mathrm{Re}((1+i)^n \mathrm{e}^{(1+i)t})$$

Or $(1+i)^n = 2^{n/2}\mathrm{e}^{in\pi/4}$ puis

$$(\cos(t)\mathrm{e}^t)^{(n)} = 2^{n/2}\mathrm{e}^t\left(\cos(t+n\pi/4) + i\sin(t+n\pi/4)\right)$$

18.3.3 Théorèmes de dérivation

18.3.3.1 Inégalité des accroissements finis

Attention : Il n'y a pas de théorème de Rolle ni a fortiori de théorème des accroissements finis pour les fonctions complexes.

Exemple Considérons la fonction complexe $f : t \mapsto \mathrm{e}^{it}$ définie et de classe $\mathcal{C}^\infty$ sur $[0, 2\pi]$.
On a $f(0) = f(2\pi)$ alors que $\forall t \in [0, 2\pi], f'(t) = i\mathrm{e}^{it} \neq 0$.

Théorème

Soit $f : I \to \mathbb{C}$ dérivable.
S'il existe $M \in \mathbb{R}^+$ vérifiant

$$\forall t \in I, |f'(t)| \leqslant M$$

alors la fonction f est M-lipschitzienne i.e.

$$\forall x, y \in I, |f(y) - f(x)| \leqslant M\,|y - x|$$

dém. :
Soient $x, y \in I$. On peut écrire $f(y) - f(x) = r\mathrm{e}^{i\theta}$ avec $r = |f(y) - f(x)| \in \mathbb{R}^+$ et $\theta \in \mathbb{R}$.
Considérons la fonction $g : I \to \mathbb{C}$ définie par $g(t) = f(t)\mathrm{e}^{-i\theta}$.
g est dérivable sur I et

$$g(y) - g(x) = (f(y) - f(x))\,\mathrm{e}^{-i\theta} = |f(y) - f(x)| \in \mathbb{R}^+$$

Par suite

$$g(y) - g(x) = |\mathrm{Re}\,(g(y) - g(x))| = |\mathrm{Re}(g(y)) - \mathrm{Re}(g(x))|$$

Appliquons l'inégalité des accroissements finis à la fonction réelle $t \mapsto \mathrm{Re}(g(t))$.
La fonction $\mathrm{Re}(g)$ est dérivable et

$$|\mathrm{Re}(g(t))'| = |\mathrm{Re}(g'(t))| \leqslant |g'(t)| = \left|f'(t)\mathrm{e}^{-i\theta}\right| \leqslant M$$

la fonction $\mathrm{Re}(g)$ est donc M-lipschitzienne et par suite

$$|\mathrm{Re}(g(y)) - \mathrm{Re}(g(x))| \leqslant M\,|y - x|$$

puis

$$|f(y) - f(x)| \leqslant M\,|y - x|$$

□

Corollaire

Soit $f : I \to \mathbb{C}$ dérivable.
La fonction f est constante si, et seulement si, $f' = 0$.

dém. :
($\Rightarrow$) Immédiat.
($\Leftarrow$) Il suffit d'appliquer l'inégalité des accroissements finis avec $M = 0$.
□

18.3.3.2 Prolongement de la dérivée

Théorème

Soit $f : I \to \mathbb{C}$ une fonction dérivable sur $I\backslash\{a\}$ et continue en a.
Si $f'(x)$ converge quand $x \to a$ alors f est dérivable en a et $f'(a) = \lim\limits_{x \to a} f'(x)$.

dém. :
Il suffit d'appliquer le théorème analogues aux fonctions parties réelles et parties imaginaires.
□

18.3.4 Fonctions à valeurs dans $\mathbb{R}^2$

Soit $f : I \to \mathbb{R}^2$ une fonction à valeurs dans $\mathbb{R}^2$ définie sur I.
Pour tout $t \in I$, on peut écrire $f(t) = (x(t), y(t))$ ce qui introduit x et $y : I \to \mathbb{R}$ appelées les fonctions coordonnées de f.

Définition

On dit qu'une fonction $f : I \to \mathbb{R}^2$ est dérivable en $a \in I$ si le taux de variation

$$\frac{1}{h}\left(f(a+h) - f(a)\right)$$

converge quand $h \to 0$ (avec $h \neq 0$)
Sa limite est alors appelée vecteur dérivé de f en a et est notée $f'(a)$.

Définition

On dit que $f : I \to \mathbb{R}^2$ est dérivable si f est dérivable en tout point de I et on peut alors introduire sa fonction dérivée $f' : I \to \mathbb{R}^2$ parfois encore notée Df.

Proposition

$f : I \to \mathbb{R}^2$ dérivable si, et seulement si, ses fonctions coordonnées x et y le sont.
De plus, si tel est le cas

$$\forall t \in I, f'(t) = (x'(t), y'(t))$$

dém. :
Le taux de variation

$$\frac{1}{h}\left(f(a+h) - f(a)\right) = \left(\frac{1}{h}\left(x(a+h) - x(a)\right), \frac{1}{h}\left(y(a+h) - y(a)\right)\right)$$

converge quand $h \to 0$ (avec $h \neq 0$) si, et seulement si, les taux de variations $(x(a+h) - x(a))/h$ et $(y(a+h) - y(a))/h$ convergent.
□

Exemple La fonction $f : t \mapsto (t^2, t^3 + t)$ est dérivable et $f'(t) = (2t, 3t^2 + 1)$

Définition

Tout comme pour les fonctions réelles et complexes, on définit la notion de dérivée d'ordre $n \in \mathbb{N}$ et de classe d'une fonction à valeurs dans $\mathbb{R}^2$.

Proposition

$f : I \to \mathbb{R}^2$ est de classe $\mathcal{C}^n$ (avec $n \in \mathbb{N} \cup \{\infty\}$) si, et seulement si, ses fonctions coordonnées le sont.

18.4 Convexité

On munit le plan géométrique $\mathcal{P}$ d'un repère $\mathcal{R} = (O; \vec{i}, \vec{j})$.

18.4.1 Paramétrage d'un segment

Soient $a, b \in \mathbb{R}$ et $\varphi : [0, 1] \to \mathbb{R}$ la fonction définie par

$$\varphi(\lambda) = (1 - \lambda)a + \lambda b$$

La fonction φ est continue, monotone et vérifie $\varphi(0) = a$ et $\varphi(1) = b$.
La fonction φ réalise une surjection de $[0, 1]$ sur le segment I d'extrémités a et b.

Définition

On dit φ réalise un paramétrage du segment I d'extrémités a et b.

Remarque En échangeant a et b, on observe que la fonction $\psi : [0, 1] \to \mathbb{R}$ définie par

$$\psi(\lambda) = \lambda a + (1 - \lambda)b$$

est encore un paramétrage du segment I.

Soient

$$A \left| \begin{matrix} x_A \\ y_A \end{matrix} \right. \text{ et } B \left| \begin{matrix} x_B \\ y_B \end{matrix} \right.$$

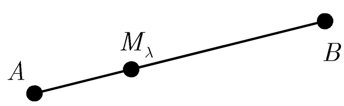
deux points du plan et $\varphi : [0, 1] \to \mathcal{P}$ la fonction définie par

$$\varphi(\lambda) = M_\lambda \left| \begin{matrix} (1 - \lambda)x_A + \lambda x_B \\ (1 - \lambda)y_A + \lambda y_B \end{matrix} \right.$$

Puisque

$$\overrightarrow{AM_\lambda} \left| \begin{matrix} \lambda(x_B - x_A) \\ \lambda(y_B - y_A) \end{matrix} \right.$$

on a $\overrightarrow{AM_\lambda} = \lambda \overrightarrow{AB}$ et ainsi φ réalise une surjection de $[0, 1]$ sur $[A, B]$.



Définition

On dit que φ est un paramétrage du segment $[A, B]$.

Remarque En échangeant A et B, on observe que la fonction $\psi : [0,1] \to \mathbb{R}^2$ définie par

$$\psi(\lambda) = M_\lambda \left| \begin{array}{l} \lambda x_A + (1-\lambda)x_B \\ \lambda y_A + (1-\lambda)y_B \end{array} \right.$$

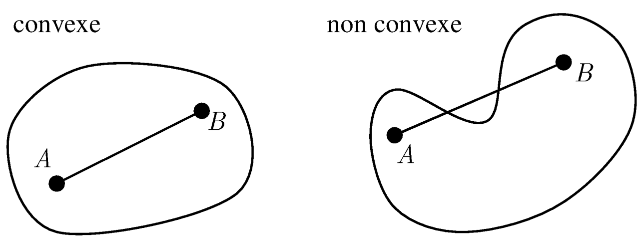
est encore un paramétrage de $[A, B]$.

18.4.2 Partie convexe du plan

Définition

Une partie $\mathcal{C}$ du plan $\mathcal{P}$ est dite convexe si elle vérifie $\forall A, B \in \mathcal{C}, [A, B] \subset \mathcal{C}$.

Exemple



Exemple $\emptyset$ et $\mathcal{P}$ sont des parties convexes.

Exemple Les segments, les droites et les demi-plans sont des parties convexes.

Remarque De manière semblable, on peut définir la notion de partie convexe de l'espace géométrique.

18.4.3 Fonction convexe, fonction concave

Définition

On dit qu'une fonction $f : I \to \mathbb{R}$ est convexe si elle vérifie

$$\forall a, b \in I, \forall \lambda \in [0,1], f(\lambda a + (1-\lambda)b) \leqslant \lambda f(a) + (1-\lambda) f(b)$$

Remarque Pour $a, b \in I$, notons

$$A \left| \begin{array}{l} a \\ f(a) \end{array} \right. \text{ et } B \left| \begin{array}{l} b \\ f(b) \end{array} \right.$$

les points du graphe de f d'abscisses a et b.
On appelle corde d'extrémités A et B le segment $[A, B]$.

On appelle arc d'extrémités A et B l'ensemble noté $\widehat{AB}$ formé des points de Γ_f d'abscisses comprises entre a et b.

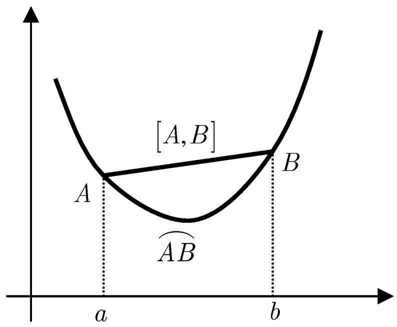
$$[A, B] = \left\{ M \left| \begin{array}{l} (1-\lambda)x_A + \lambda x_B \\ (1-\lambda)y_A + \lambda y_B \end{array} \right. / \lambda \in [0, 1] \right\}$$

et

$$\widehat{AB} = \left\{ N \left| \begin{array}{l} \lambda a + (1-\lambda)b \\ f(\lambda a + (1-\lambda)b) \end{array} \right. / \lambda \in [0, 1] \right\}$$

L'inégalité de convexité affirme que pour une abscisse x donnée, l'ordonnée du point d'abscisse x de la corde est supérieure à l'ordonnée du point d'abscisse x de l'arc.

Ainsi, pour une fonction convexe, l'arc $\widehat{AB}$ est en dessous de la corde $[A, B]$.



Exemple Les fonctions affines $x \mapsto \alpha x + \beta$ sont convexes ; en fait pour ces fonctions, l'inégalité de convexité est une égalité.

Exemple La fonction $x \mapsto |x|$ est convexe car

$$\forall a, b \in \mathbb{R}, |\lambda a + (1-\lambda)b| \leqslant |\lambda|\,|a| + |1-\lambda|\,|b| = \lambda\,|a| + (1-\lambda)\,|b|$$

Définition

On dit qu'une fonction $f : I \to \mathbb{R}$ est concave si elle vérifie

$$\forall a, b \in I, \forall \lambda \in [0, 1]\,, f(\lambda a + (1-\lambda)b) \geqslant \lambda f(a) + (1-\lambda) f(b)$$

Remarque Pour une fonction concave, l'arc est au dessus de la corde.

Exemple Les fonctions affines sont concaves.

Proposition

Pour $f : I \to \mathbb{R}$, on a équivalence entre :
(i) f est concave ;
(ii) $-f$ est convexe.

dém. :
Par passage à l'opposé l'inégalité de convexité est renversée.
□

Remarque Par passage à l'opposé et renversement d'inégalité, les résultats suivant présenté pour les fonctions convexes se transposent aux fonctions concaves.

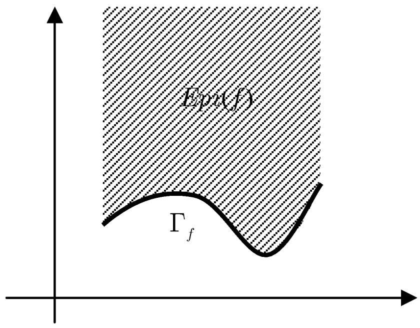
18.4.4 Caractérisation de la convexité

18.4.4.1 Épigraphe

Définition

On appelle épigraphe d'une $f : I \to \mathbb{R}$ l'ensemble

$$Epi(f) = \left\{ M \left| \begin{matrix} x \\ y \end{matrix} \right. \in \mathcal{P} / x \in I, f(x) \leqslant y \right\}$$



Théorème

Pour $f : I \to \mathbb{R}$, on a équivalence entre :
(i) la fonction f est convexe ;
(ii) l'épigraphe de f est convexe.

dém. :
(i) $\Rightarrow$ (ii) Supposons f convexe.
Soient

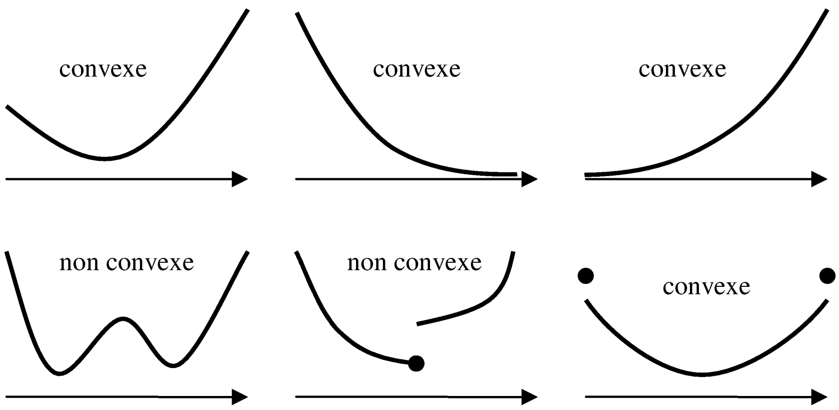
$$A \left| \begin{matrix} x_A \\ y_A \end{matrix} \right., B \left| \begin{matrix} x_B \\ y_B \end{matrix} \right.$$

des points de l'épigraphe de f et

$$A' \left| \begin{matrix} x_A \\ f(x_A) \end{matrix} \right. , B' \left| \begin{matrix} x_B \\ f(x_B) \end{matrix} \right.$$

les points correspondants du graphe de f. Par comparaison des ordonnées de points de même abscisse, on peut affirme que le segment $[A, B]$ est au dessus du segment $[A', B']$ et lui même au dessus de l'arc $\widehat{A'B'}$. On en déduit que le segment $[A, B]$ est inclus dans l'épigraphe de f.
(ii) $\Rightarrow$ (i) Supposons l'épigraphe de f convexe.
Soient $a, b \in I$ et A, B les points d'abscisses correspondantes du graphe de f.
Ces points sont aussi dans l'épigraphe de f et donc le segment $[A, B]$ est inclus dans l'épigraphe de f et par suite le segment $[A, B]$ est au dessus de la corde $\widehat{AB}$.
Ainsi f est convexe.
$\square$

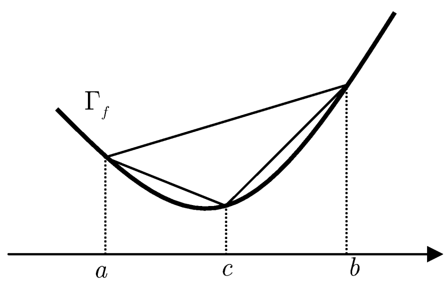
Exemple



18.4.4.2 Taux de variation

Lemme

Soit $f : I \to \mathbb{R}$. On a équivalence entre :
(i) f est convexe ;
(ii) $\forall a, b, c \in I$, $a < c < b \Rightarrow \tau(a, c) \leqslant \tau(a, b) \leqslant \tau(c, b)$;
(iii) $\forall a, b, c \in I$, $a < c < b \Rightarrow \tau(a, c) \leqslant \tau(c, b)$.



dém. :
(i) $\Rightarrow$ (ii) Supposons f convexe
Soient $a, b, c \in I$ tels que $a < c < b$. $c = \lambda a + (1-\lambda)b$ avec

$$\lambda = \frac{b-c}{b-a} \in]0,1[$$

Par convexité

$$f(c) = f(\lambda a + (1-\lambda)b) \leqslant \lambda f(a) + (1-\lambda)f(b)$$

donc

$$f(c) - f(a) \leqslant (1-\lambda)(f(b) - f(a)) = \frac{c-a}{b-a}(f(b) - f(a))$$

d'où $\tau(a,c) \leqslant \tau(a,b)$.
De plus :

$$f(b) - f(c) \geqslant \lambda(f(b) - f(a)) = \frac{b-c}{b-a}(f(b) - f(a))$$

d'où $\tau(a,b) \leqslant \tau(b,c)$.
(ii) $\Rightarrow$ (iii) ok
(iii) $\Rightarrow$ (i) Supposons (iii)
Soient $a, b \in I$ et $\lambda \in [0,1]$.
Montrons

$$f(\lambda a + (1-\lambda)b) \leqslant \lambda f(a) + (1-\lambda)f(b)$$

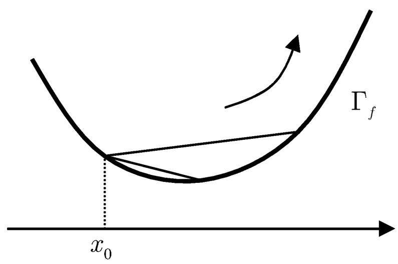
Si $a = b$: ok
Si $a \neq b$, quitte à échanger a et b d'une part, et λ et $1-\lambda$ d'autre part, on peut supposer $a < b$.
Si $\lambda = 0$ ou $\lambda = 1$: ok
Si $\lambda \in]0,1[$, posons $c = \lambda a + (1-\lambda)b$.
Puisque $a < c < b$, on a $\tau(a,c) \leqslant \tau(c,b)$ ce qui donne

$$f(c) - f(a) \leqslant \frac{c-a}{b-c}(f(b) - f(c)) = \frac{1-\lambda}{\lambda}(f(b) - f(c))$$

puis $f(c) \leqslant \lambda f(a) + (1-\lambda)f(b)$.
Ainsi f est convexe.
□

Théorème

Soit $f : I \to \mathbb{R}$. On a équivalence entre :
(i) f est convexe ;
(ii) $\forall x_0 \in I$ la fonction $\tau_{x_0} : I \setminus \{x_0\} \to \mathbb{R}$ définie par $\tau_{x_0}(x) = \tau(x_0, x)$ est croissante.



dém. :
(i) $\Rightarrow$ (ii) Supposons f convexe.
Soit $x_0 \in I$. Montrons que τ_{x_0} est croissante.
Soient $x, y \in I \setminus \{x_0\}$ tels que $x < y$.
Si $x_0 < x < y$ alors $\tau_{x_0}(x) \leqslant \tau_{x_0}(y)$ (en prenant $a = x_0, c = x$ et $b = y$),
Si $x < x_0 < y$ alors $\tau_{x_0}(x) \leqslant \tau_{x_0}(y)$ (en prenant $a = x, c = x_0$ et $b = y$),
Si $x < y < x_0$ alors $\tau_{x_0}(x) \leqslant \tau_{x_0}(y)$ (en prenant $a = x, c = y$ et $b = x_0$).
Dans tous les cas $\tau_{x_0}(x) \leqslant \tau_{x_0}(y)$
(ii) $\Rightarrow$ (i) Supposons (ii)
Soient $a, b, c \in I$ tels que $a < c < b$.
On a $\tau(a, c) = \tau_c(a) \leqslant \tau_c(b) = \tau(c, b)$ donc f est convexe en vertu du lemme précédent.
□

18.4.4.3 Dérivation

Théorème

> Soit $f : I \to \mathbb{R}$ dérivable. On a équivalence entre :
> (i) f est convexe ;
> (ii) f' est croissante.

dém. :
(i) $\Rightarrow$ (ii) Supposons f convexe.
Soient $a, b \in I$ tels que $a < b$ et $x \in]a, b[$.
On a

$$\tau(a, x) \leqslant \tau(a, b) \leqslant \tau(b, x)$$

Quand $x \to a^+$: $f'(a) \leqslant \tau(a, b)$.
Quand $x \to b^-$: $\tau(a, b) \leqslant f'(b)$.
Ainsi $f'(a) \leqslant f'(b)$ et f' est croissante.
(ii) $\Rightarrow$ (i) Supposons f' croissante.
Soient $a, b, c \in I$ tels que $a < c < b$.
Par le théorème des accroissements finis, il existe $\alpha \in]a, c[$ tel que $\tau(a, c) = f'(\alpha)$ et il existe $\beta \in]c, b[$ tel que $\tau(c, b) = f'(\beta)$. Puisque $\alpha \leqslant \beta$, on obtient $\tau(a, c) \leqslant \tau(c, b)$.
On peut alors conclure que f est convexe en vertu du précédent lemme.
□

Corollaire

> Soit $f : I \to \mathbb{R}$ deux fois dérivable.
> On a équivalence entre :
> (i) f est convexe ;
> (ii) $f'' \geqslant 0$.

dém. :
La monotonie de f' est donnée par le signe de f''.
□

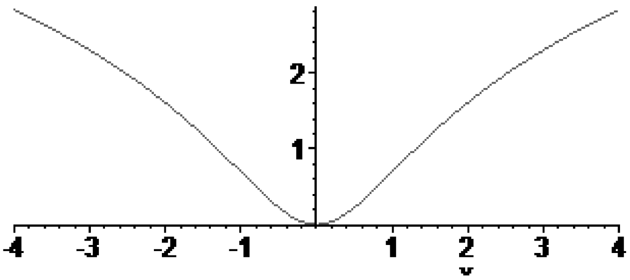
Exemple Les fonctions $x \mapsto x^2$, $x \mapsto \mathrm{e}^x$, $x \mapsto \mathrm{ch}x$ sont des fonctions convexes car de dérivées secondes positives

Exemple La fonction $x \mapsto \ln x$ est une fonction concave car de dérivée seconde négative.

Exemple Etudions la convexité de la fonction $f : x \mapsto \ln(1 + x^2)$ définie sur $\mathbb{R}$.
La fonction f est deux fois dérivable,

$$f'(x) = \frac{2x}{1 + x^2} \text{ et } f''(x) = 2\frac{1 - x^2}{(1 + x^2)^2}$$

du signe de $1 - x^2$
On en déduit que f est convexe sur $[-1, 1]$ et concave sur $]-\infty, -1]$ et sur $[1, +\infty[$.
Notons que nous ne dirons pas que f est concave sur la réunion $]-\infty, -1] \cup [1, +\infty[$ car la notion de convexité d'une fonction réelle n'a de sens que pour une fonction définie sur un intervalle.



18.4.5 Position relative d'une courbe et de sa tangente en un point

Théorème
> Si $f : I \to \mathbb{R}$ dérivable est convexe alors son graphe Γ_f est au dessus de chacune de ses tangentes.

dém. :
Soit $a \in I$.
L'équation de la tangente T en a est :

$$y = f'(a)(x - a) + f(a)$$

Considérons la fonction $g : I \to \mathbb{R}$ définie par

$$g(x) = f(x) - (f'(a)(x - a) + f(a))$$

Par opérations, la fonction g est dérivable et $g'(x) = f'(x) - f'(a)$.
La croissance de f' donne le signe de g' et on en déduit que g admet un minimum en a avec $g(a) = 0$.
Par suite, pour tout $x \in I$, $g(x) \geqslant 0$ puis l'inégalité

$$f(x) \geqslant f'(a)(x - a) + f(a)$$

□

Corollaire
> Si $f : I \to \mathbb{R}$ dérivable est concave alors son graphe Γ_f est en dessous de chacune de ses tangentes.

dém. :
Il suffit de considérer la fonction $-f$ qui est convexe.
□

Exemple Puisque la fonction $x \mapsto \mathrm{e}^x$ est convexe, en positionnant son graphe par rapport à sa tangente en 0, on obtient la propriété

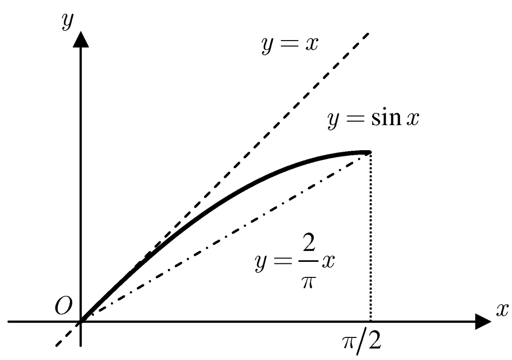
$$\forall x \in \mathbb{R}, \mathrm{e}^x \geqslant 1 + x$$

Puisque la fonction $x \mapsto \ln(1+x)$ est concave, en positionnant son graphe par rapport à sa tangente en 0, on obtient la propriété

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

Puisque la fonction $x \mapsto \sin x$ est concave sur $[0, \pi/2]$, en positionnant son graphe par rapport à sa tangente en 0 et par rapport à sa corde joignant les points d'abscisse 0 et $\pi/2$, on obtient la propriété

$$\forall x \in [0, \pi/2], \frac{2}{\pi}x \leqslant \sin x \leqslant x$$



18.4.6 Inégalités de convexité

Lemme

Soit I un intervalle et $n \in \mathbb{N}^\star$.

$$\forall a_1, \ldots, a_n \in I, \forall \lambda_1, \ldots, \lambda_n \in \mathbb{R}^+, \lambda_1 + \cdots + \lambda_n = 1 \Rightarrow x = \lambda_1 a_1 + \cdots + \lambda_n a_n \in I$$

dém. :
Soient $a_1, \ldots, a_n \in I$ et $\lambda_1, \ldots, \lambda_n \geqslant 0$ vérifiant $\lambda_1 + \cdots + \lambda_n = 1$.
Considérons $a = \min\limits_{1 \leqslant i \leqslant n} a_i$ et $b = \min\limits_{1 \leqslant i \leqslant n} a_i$.
Puisque a et b sont éléments de I, le segment $[a, b]$ est un inclus dans I.
Or $x = \lambda_1 a_1 + \cdots + \lambda_n a_n \geqslant (\lambda_1 + \cdots + \lambda_n)a = a$ et de même $x \leqslant b$ donc $x \in I$.
□

Théorème

Soit $f : I \to \mathbb{R}$ une fonction convexe et $n \in \mathbb{N}^\star$.

$$\forall a_1, \dots, a_n \in I, \forall \lambda_1, \dots, \lambda_n \in \mathbb{R}^+$$

$$\lambda_1 + \dots + \lambda_n = 1 \Rightarrow f(\lambda_1 a_1 + \dots + \lambda_n a_n) \leqslant \lambda_1 f(a_1) + \dots + \lambda_n f(a_n)$$

dém. :
Par récurrence sur $n \in \mathbb{N}^\star$.
Pour $n = 1$: ok
Supposons la propriété établie au rang $n \geqslant 1$.
Soient $a_1, \dots, a_n, a_{n+1} \in I$ et $\lambda_1, \dots, \lambda_n, \lambda_{n+1} \geqslant 0$ vérifiant $\lambda_1 + \dots + \lambda_n + \lambda_{n+1} = 1$.
Posons alors $\mu = \lambda_1 + \dots + \lambda_n = 1 - \lambda_{n+1}$.
Cas $\mu > 0$:
On introduit $\mu_1 = \lambda_1/\mu, \dots, \mu_n = \lambda_n/\mu$.
On vérifie $\mu_1, \dots, \mu_n \geqslant 0$ et $\mu_1 + \dots + \mu_n = 1$.
En introduisant $a = \mu_1 a_1 + \dots + \mu_n a_n$, l'hypothèse de récurrence donne

$$f(a) \leqslant \mu_1 f(a_1) + \dots + \mu_n f(a_n)$$

Puisque

$$\lambda_1 a_1 + \dots + \lambda_n a_n + \lambda_{n+1} a_{n+1} = \mu a + (1 - \mu) a_{n+1}$$

avec $\mu \in [0, 1]$, l'inégalité de convexité donne

$$f(\lambda_1 a_1 + \dots + \lambda_n a_n + \lambda_{n+1} a_{n+1}) \leqslant \mu f(a) + (1 - \mu) f(a_{n+1})$$

puis on obtient

$$f(\lambda_1 a_1 + \dots + \lambda_n a_n + \lambda_{n+1} a_{n+1}) \leqslant \lambda_1 f(a_1) + \dots + \lambda_n f(a_n) + \lambda_{n+1} f(a_{n+1})$$

Cas $\mu = 0$:
On a $\lambda_1 = \ldots = \lambda_n = 0$ et $\lambda_{n+1} = 1$ et l'inégalité voulue est immédiate.
Récurrence établie.
□

Corollaire

Pour $f : I \to \mathbb{R}$ convexe, on a

$$\forall a_1, \dots, a_n \in I, f\left(\frac{a_1 + \dots + a_n}{n}\right) \leqslant \frac{1}{n}\left(f(a_1) + \dots + f(a_n)\right)$$

dém. :
Il suffit de prendre $\lambda_1 = \ldots = \lambda_n = 1/n$
□

Exemple Montrons

$$\forall a_1, \dots, a_n \in \mathbb{R}^+, \sqrt[n]{a_1 \dots a_n} \leqslant \frac{a_1 + \dots + a_n}{n}$$

Si l'un des a_i est nul, c'est immédiat.
Sinon, exploitons la concavité de $x \mapsto \ln x$.

Pour tout $a_1, \dots, a_n > 0$,

$$\frac{1}{n}\left(\ln a_1 + \cdots + \ln a_n\right) \leqslant \ln\left(\frac{a_1 + \cdots + a_n}{n}\right)$$

donc

$$\ln \sqrt[n]{a_1 \dots a_n} \leqslant \ln\left(\frac{a_1 + \cdots + a_n}{n}\right)$$

puis en composant avec la fonction exponentielle qui est croissante on obtient l'inégalité voulue.

18.4.7 Musculation : dérivabilité et continuité des fonctions convexes

Théorème

Si $f : I \to \mathbb{R}$ est convexe alors en tout point $x_0 \in I$ qui n'est pas extrémité de I, f est dérivable à droite et à gauche et

$$f'_g(x_0) \leqslant f'_d(x_0)$$

dém. :
Soit $a \in I$ tel que $a < x_0$.
L'application restreinte τ_{x_0} : $]x_0, +\infty[\, \cap I$ est croissante et minorée par $\tau(a, x_0)$, cette application converge donc en x_0^+. Ainsi f est dérivable à droite en x_0 et $f'_d(x_0) \geqslant \tau(a, x_0)$.
L'application restreinte τ_{x_0} : $]-\infty, x_0[\, \cap I$ est croissante et majorée, en vertu de l'étude précédente, par $f'_d(x_0)$. Cette application converge donc en x_0^- et f est dérivable à gauche en x_0 avec $f'_g(x_0) \leqslant f'_d(x_0)$.
□

Corollaire

Si $f : I \to \mathbb{R}$ est convexe alors f est continue en tout point intérieur à I.

dém. :
Car continue à droite et à gauche par dérivabilité à droite et à gauche...
□

18.5 Etude graphique d'une fonction

Pour étudier le graphe Γ_f d'une fonction $f : I \to \mathbb{R}$, on introduit le paramétrage

$$\begin{cases} x = t \\ y = f(t) \end{cases} \text{ avec } t \in I$$

de point courant $M(t)$.

18.5.1 Réduction du domaine d'étude

Cas $I = \mathbb{R}$ et f est une fonction T périodique :
On peut limiter l'étude de f à un intervalle de longueur T car le point $M(t + T)$ est l'image du point $M(t)$ par la translation de vecteur $T.\vec{i}$.
Cas I est symétrique par rapport à t_0 et $f(2t_0 - t) = f(t)$:
On a $f(t_0 - h) = f(t_0 + h)$ et donc les points $M(t_0 + h)$ et $M(t_0 - h)$ sont symétriques l'un de l'autre par rapport à la droite $\Delta : x = t_0$.
Cas I est symétrique par rapport à t_0 et $f(2t_0 - t) = -f(t)$:

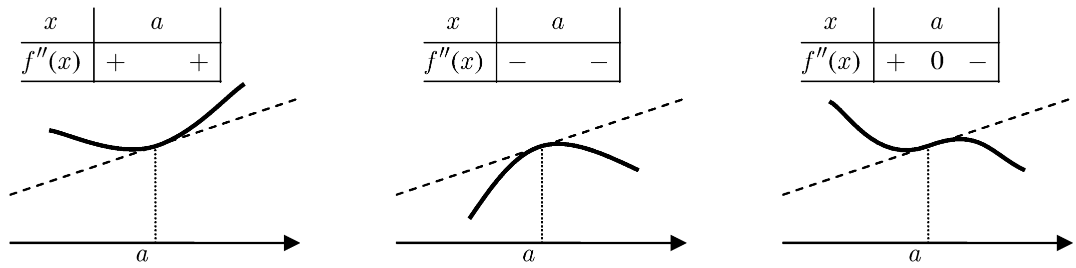
On a $f(t_0 - h) = -f(t_0 + h)$ et les points $M(t_0 + h)$ et $M(t_0 - h)$ sont symétriques l'un de l'autre par rapport au point $\Omega(t_0, 0)$.
Dans ces deux cas, on peut limiter l'étude à l'intervalle $I \cap [t_0, +\infty[$ ou à l'intervalle $I \cap]-\infty, t_0]$.

18.5.2 Etude locale

Si f est dérivable en a alors le graphe Γ_f admet une tangente en $M(a)$ qui est de pente $f'(a)$. C'est la droite d'équation

$$y = f'(a)(x - a) + f(a)$$

Supposons maintenant f deux fois dérivable au voisinage de a.
Si $f''(x) \geqslant 0$ (resp. $f''(x) \leqslant 0$) au voisinage de a, alors f est convexe (resp. concave) au voisinage de a et Γ_f est alors localement au dessus (resp. en dessous) de sa tangente en a.
Si $f''(x)$ s'annule en changeant de signe en a alors f est convexe d'un côté de a et concave de l'autre, on dit que f présente un point d'inflexion en a et que Γ_f traverse sa tangente en a.

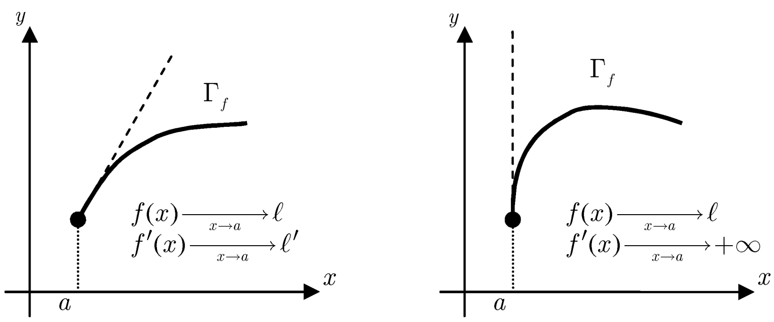


Remarque Pour localiser les éventuels points d'inflexion d'une fonction $f : I \to \mathbb{R}$ deux fois dérivable, on détermine les points où f'' s'annule en changeant de signe.

18.5.3 Etude aux extrémités ouvertes de l'intervalle de définition

18.5.3.1 Prolongement par continuité

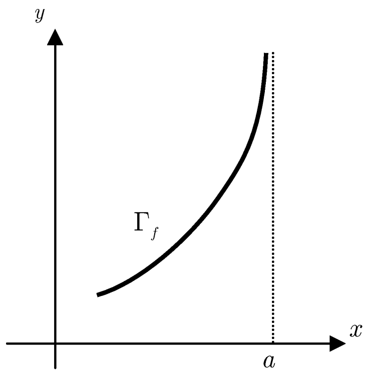
Soit a une extrémité ouverte finie de I.
Si $f(x) \xrightarrow[x \to a]{} \ell$ alors on peut prolonger f par continuité en a en posant $f(a) = \ell$.
Si de plus $f'(x) \xrightarrow[x \to a]{} \ell'$ alors f est dérivable en a et $f'(a) = \ell$.
En revanche si $f'(x) \xrightarrow[x \to a]{} \pm\infty$ alors f n'est pas dérivable en a mais son graphe y présente une tangente verticale.



18.5.3.2 Asymptote verticale

Soit a une extrémité ouverte finie de I.
Si $f(x) \xrightarrow[x\to a]{} \pm\infty$ alors la droite Δ d'équation $x = a$ est asymptote à Γ_f.
En effet

$$\frac{y(t)}{x(t)} = \frac{f(t)}{t} \xrightarrow[t\to a]{} \pm\infty \text{ et } x(t) \to a$$

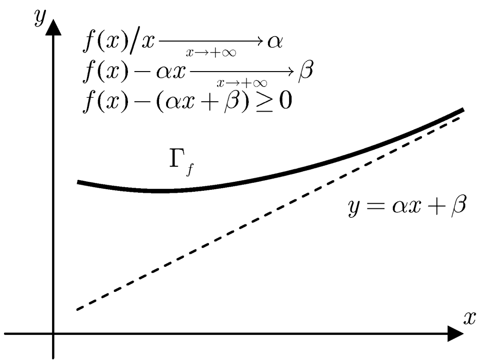


18.5.3.3 Asymptote oblique et branche parabolique

Supposons que $+\infty$ soit une extrémité de I ; l'étude est identique en $-\infty$.
Cas $f(x)/x \xrightarrow[x\to+\infty]{} \alpha \in \mathbb{R}$ et $f(x) - \alpha x \xrightarrow[x\to+\infty]{} \beta \in \mathbb{R}$:
La droite d'équation $y = \alpha x + \beta$ est asymptote à Γ_f en $+\infty$.
En effet

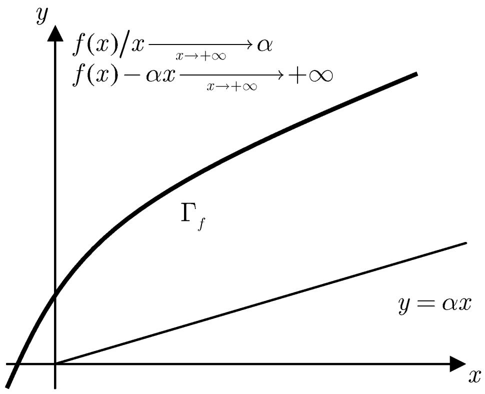
$$\frac{y(t)}{x(t)} \xrightarrow[t\to+\infty]{} \alpha \in \mathbb{R} \text{ et } y(t) - \alpha x(t) \to \beta \in \mathbb{R}$$

Il est alors usuel de positionner la courbe par rapport à son asymptote en étudiant le signe de $f(x) - (\alpha x + \beta)$ au voisinage de $+\infty$.

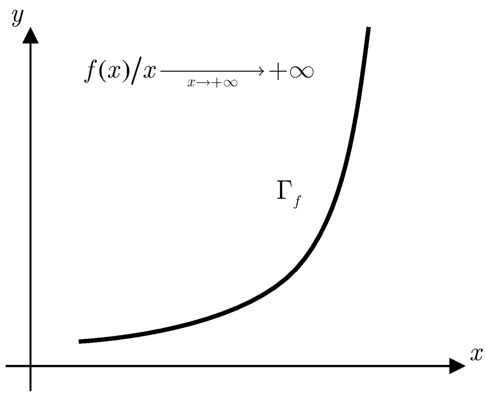


Cas $f(x)/x \xrightarrow[x\to+\infty]{} \alpha \in \mathbb{R}$ et $f(x) - \alpha x \xrightarrow[x\to+\infty]{} \pm\infty$:
Le graphe Γ_f présente une branche parabolique de direction $y = \alpha x$.
En effet

$$\frac{y(t)}{x(t)} \xrightarrow[t\to+\infty]{} \alpha \in \mathbb{R} \text{ et } y(t) - \alpha x(t) \to \pm\infty$$



Cas $f(x)/x \xrightarrow[x\to+\infty]{} \pm\infty$:
Le graphe Γ_f présente une branche parabolique verticale.



18.5.4 Exemples d'étude

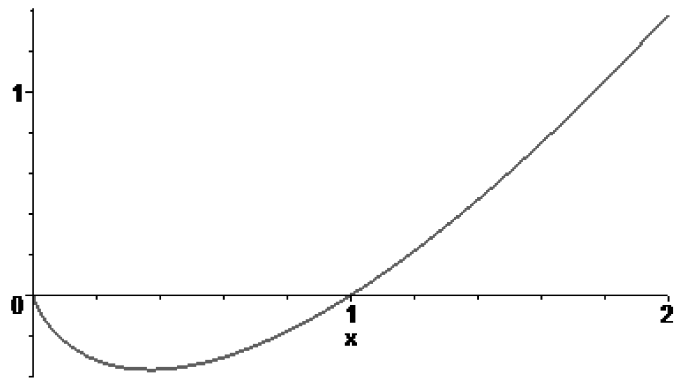
Exemple Etudions la fonction

$$f : x \mapsto x \ln x$$

La fonction f est définie et de classe $\mathcal{C}^\infty$ sur $]0,+\infty[$ et $f'(x) = 1 + \ln x$

x	0		$1/\mathrm{e}$		$+\infty$
$f'(x)$		$-$	0	$+$	
$f(x)$	0	$\searrow$	$-1/\mathrm{e}$	$\nearrow$	$+\infty$

$f''(x) = 1/x \geqslant 0$.
La fonction f est convexe donc au dessus de chacune de ses tangentes.
Quand $x \to 0^+$,
$f(x) \to 0$ donc on peut prolonger f par continuité en posant $f(0) = 0$.
$f'(x) \to -\infty$, le graphe de f présente une tangente verticale en 0.
Quand $x \to +\infty$,
$f(x)/x \to +\infty$, le graphe de f présente une branche parabolique verticale.



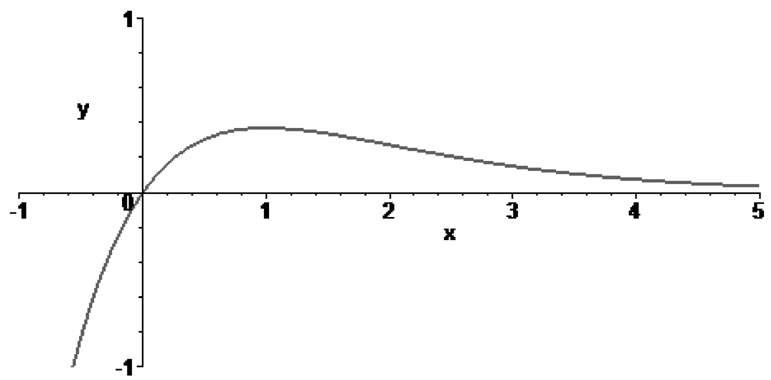
Exemple Etudions la fonction

$$f : x \mapsto x\mathrm{e}^{-x}$$

La fonction f est définie et de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$.
$f'(x) = (1-x)\mathrm{e}^{-x}$ est du signe de $1-x$.

x	$-\infty$		1		$+\infty$
$f'(x)$		$+$	0	$-$	
$f(x)$	$-\infty$	$\nearrow$	$1/\mathrm{e}$	$\searrow$	0

$f''(x) = (x-2)\mathrm{e}^{-x}$ est du signe de $x-2$.
Le graphe de f présente un point d'inflexion en 2.
Quand $x \to +\infty$,
$f(x) \to 0^+$, le graphe de f présente une asymptote horizontale d'équation $y = 0$, courbe au dessus.
Quand $x \to -\infty$,
$f(x)/x \to -\infty$, le graphe de f présente une branche parabolique verticale.



Exemple Etudions la fonction

$$f : x \mapsto \frac{x^2 + 2x}{x + 1}$$

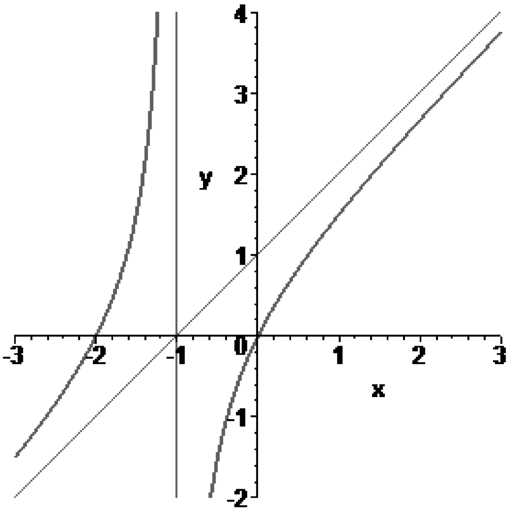
La fonction f est définie et $\mathcal{C}^\infty$ sur les intervalles $]-\infty, -1[$ et $]-1, +\infty[$.

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

x	$-\infty$			-1			$+\infty$
$f'(x)$		$+$		$\|\|$		$+$	
$f(x)$	$-\infty$	$\nearrow$	$+\infty$	$\|\|$	$-\infty$	$\nearrow$	$+\infty$

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

La fonction f est concave sur $]-1, +\infty[$ et convexe sur $]-\infty, -1[$.
Quand $x \to 1^+$,
$f(x) \to +\infty$.
La droite verticale d'équation $x = 1$ est asymptote au graphe de f.
Quand $x \to 1^-$,
$f(x) \to +\infty$.
La droite verticale d'équation $x = 1$ est asymptote au graphe de f.
Quand $x \to +\infty$,
$f(x)/x \to 1$ et $f(x) - x \to 1^-$.
La droite d'équation $y = x + 1$ est asymptote au graphe de f, courbe en dessous.
Quand $x \to -\infty$,
$f(x)/x \to 1$ et $f(x) - x \to 1^+$.
La droite d'équation $y = x + 1$ est asymptote au graphe de f, courbe au dessus.



18.6 Suites récurrentes réelles

18.6.1 Vocabulaire

$\mathcal{D}$ désigne une partie non vide de $\mathbb{R}$.

Définition

On appelle suite récurrente d'ordre 1 de fonction itératrice $f : \mathcal{D} \to \mathbb{R}$ toute suite $(u_n) \in \mathcal{D}^{\mathbb{N}}$ vérifiant

$$\forall n \in \mathbb{N}, u_{n+1} = f(u_n)$$

Exemple Les suites arithmético-géométriques sont les suites récurrentes réelles de fonction itératrices affines.

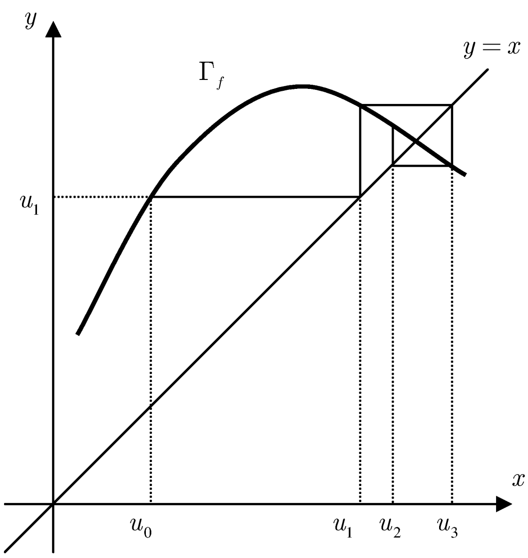
Exemple La suite (u_n) définie par

$$u_0 \in \mathbb{N}^\star \text{ et } \forall n \in \mathbb{N}, u_{n+1} = \begin{cases} u_n/2 & \text{si } u_n \text{ est pair} \\ 3u_n + 1 & \text{sinon} \end{cases}$$

est une suite récurrente réelle de fonction itératrice l'application $f : \mathbb{N}^\star \to \mathbb{N}^\star$ définie par

$$f(x) = \begin{cases} x/2 & \text{si } x \text{ est pair} \\ 3x + 1 & \text{sinon} \end{cases}$$

Remarque On peut visualiser l'évolution des termes de la suite (u_n) en représentant conjointement le graphe de la fonction itératrice f et la droite d'équation $y = x$.



Remarque Pour une telle suite, la connaissance de u_0 suffit à déterminer l'intégralité des termes de la suite puisque

$$u_1 = f(u_0), u_2 = f(f(u_0)), u_3 = (f(f(f(u_0))), ...$$

Le terme u_0 est appelé le germe de la suite récurrente (u_n).

Théorème

Si $f(\mathcal{D}) \subset \mathcal{D}$ (i.e. $\forall x \in \mathcal{D}, f(x) \in \mathcal{D}$) alors il existe une unique suite $(u_n) \in \mathcal{D}^{\mathbb{N}}$ vérifiant

$$u_0 = a \text{ et } \forall n \in \mathbb{N}, u_{n+1} = f(u_n)$$

dém. :
Unicité :
Soient (u_n) et (v_n) deux suites solutions.
Montrons par récurrence sur $n \in \mathbb{N}$ la propriété $u_n = v_n$.
Pour $n = 0$, $u_0 = a$ et $v_0 = a$ donc $u_0 = v_0$.
Supposons la propriété vraie au rang $n \geqslant 0$.
Par hypothèse de récurrence $u_n = v_n$.
Or $u_{n+1} = f(u_n)$ et $v_{n+1} = f(v_n)$ donc $u_{n+1} = v_{n+1}$.
Récurrence établie.
Existence :
Si l'unicité est facile à établir par récurrence, l'existence ne se démontre pas aussi simplement qu'on pourrait le croire. En effet, établir par récurrence la propriété « u_n existe »est ambigu car ne précise pas ce qu'est u_n...
Pour commencer montrons par récurrence sur $n \in \mathbb{N}$ la propriété :
Il existe une unique famille $(u_0, u_1, \ldots, u_n) \in \mathcal{D}^{n+1}$ vérifiant la propriété $\mathcal{P}_n$: « $u_0 = a$ et $\forall k \in \{0, \ldots, n-1\}, u_{k+1} = f(u_k)$. »
Pour $n = 0$, la propriété est immédiate.
Supposons la propriété vraie au rang $n \geqslant 0$.
Unicité :
Si la famille $(u_0, u_1, \ldots, u_{n+1}) \in \mathcal{D}^{n+2}$ vérifie la propriété $\mathcal{P}_{n+1}$ alors la sous-famille $(u_0, u_1, \ldots, u_n) \in \mathcal{D}^{n+1}$ vérifie la propriété $\mathcal{P}_n$ et donc cette dernière est déterminée de manière unique en vertu de l'hypothèse de récurrence.
De plus $u_{n+1} = f(u_n)$ et donc u_{n+1} est aussi déterminé de manière unique.
Existence :
Par hypothèse de récurrence, il existe une famille $(u_0, u_1, \ldots, u_n) \in \mathcal{D}^{n+1}$ vérifiant la propriété $\mathcal{P}_n$.
Posons alors $u_{n+1} = f(u_n)$ ce qui est possible car $u_n \in \mathcal{D}$. On obtient alors $u_{n+1} \in \mathcal{D}$ et la famille $(u_0, u_1, \ldots, u_{n+1})$ vérifie la propriété $\mathcal{P}_{n+1}$.
Récurrence établie.
Malheureusement, pour conclure il ne suffit pas de faire tendre n vers $+\infty$ (cela n'aurait pas de sens) ; nous allons mettre en place ce qu'on appelle un procédé diagonal de Cantor...
Pour tout $n \in \mathbb{N}$, notons v_n le dernier élément de la famille $(u_0, u_1, \ldots, u_n) \in \mathcal{D}^{n+1}$ vérifiant la propriété $\mathcal{P}_n$.
On vérifie aisément que (v_n) est une suite d'éléments de $\mathcal{D}$ et que $v_0 = a$. Il reste à montrer que pour tout $n \in \mathbb{N}$, $v_{n+1} = f(v_n)$ afin de pouvoir affirmer que (v_n) est solution et conclure.
Pour tout $n \in \mathbb{N}$, v_{n+1} est le dernier élément de la famille $(u_0, u_1, \ldots, u_{n+1})$ vérifiant la propriété $\mathcal{P}_{n+1}$.
Puisque alors la sous-famille $(u_0, u_1, \ldots, u_n)$ vérifie la propriété $\mathcal{P}_n$, on peut affirmer $v_n = u_n$ et puisque

$u_{n+1} = f(u_n)$, on obtient $v_{n+1} = f(v_n)$.
□

18.6.2 Exemples d'étude de suite récurrente réelle

Pour étudier une suite (u_n) donnée par $u_0 = a$ et $\forall n \in \mathbb{N}, u_{n+1} = f(u_n)$:
- on précise et on étudie la fonction itératrice f (domaine de définition, tableau de variation,. . .) ;
- on détermine $\mathcal{D}$ vérifiant $u_0 \in \mathcal{D}$ et $\forall x \in \mathcal{D}, f(x) \in \mathcal{D}$, ceci justifie l'existence de (u_n) et localise ses termes :

$$\forall n \in \mathbb{N}, u_n \in \mathcal{D}$$

- si cela est possible, on exprime le terme général de (u_n) et on poursuit l'étude à partir de celui-ci ;
- sinon, on analyse les limites finies possibles en passant la relation de récurrence $u_{n+1} = f(u_n)$ à la limite ;
- on peut ensuite, étudier la monotonie de (u_n), étudier $|u_n - \ell|$ avec ℓ limite finie,. . .

Remarque On pourra aussi s'appuyer sur une visualisation graphique de l'évolution de (u_n) pour orienter l'étude.

Exemple Soit (u_n) la suite déterminée par

$$u_0 = a \text{ et } \forall n \in \mathbb{N}, u_{n+1} = \sqrt{u_n^2 + 1}$$

Considérons l'application $f : x \mapsto \sqrt{x^2 + 1}$.
Puisque l'application f est bien définie de $\mathbb{R}$ vers $\mathbb{R}$, la suite (u_n) est bien définie.
Supposons que la suite (u_n) converge.
En posant $\ell \in \mathbb{R}$ la limite de (u_n), la relation $u_{n+1} = \sqrt{u_n^2 + 1}$ donne à la limite $\ell = \sqrt{\ell^2 + 1}$.
Cette équation n'ayant pas de solution réelle, on peut conclure que la suite (u_n) diverge.
De plus $u_{n+1} - u_n = \sqrt{u_n^2 + 1} - u_n \geqslant 0$ donc la suite (u_n) est croissante.
Finalement (u_n) diverge vers $+\infty$.

Exemple Soit (u_n) la suite définie par

$$u_0 = a \in \mathbb{R}^+ \text{ et } u_{n+1} = \ln(1 + u_n)$$

Considérons l'application $f : x \mapsto \ln(1 + x)$.
L'application f définie sur $\mathbb{R}^+$ et prend ses valeurs dans $\mathbb{R}^+$.
Puisque $u_0 = a \in \mathbb{R}^+$ et que $\forall x \in \mathbb{R}^+, f(x) \in \mathbb{R}^+$, la suite (u_n) est bien définie et tous ses termes sont éléments de $\mathbb{R}^+$.
Supposons que la suite (u_n) converge et notons $\ell \in \mathbb{R}$ sa limite.
Puisque pour tout $n \in \mathbb{N}$, $u_n \geqslant 0$, à la limite $\ell \geqslant 0$.
Puisque pour tout $n \in \mathbb{N}$, $u_{n+1} = \ln(1 + u_n)$, à la limite $\ell = \ln(1 + \ell)$
Pour résoudre cette équation, introduisons la fonction $\varphi : x \mapsto x - \ln(1 + x)$
La fonction φ est dérivable, $\varphi'(x) > 0$ sur $\mathbb{R}^{+\star}$ et $\varphi(0) = 0$.
On en déduit que $x = 0$ est la seule solution de l'équation $\varphi(x) = 0$
Ainsi, si la suite (u_n) converge, sa limite est 0.
Etudions maintenant la monotonie de (u_n).

$$u_{n+1} - u_n = \ln(1 + u_n) - u_n = -\varphi(u_n)$$

En vertu de l'étude qui précède, $u_{n+1} - u_n \leqslant 0$.
La suite (u_n) est décroissante, or elle est minorée par 0 donc elle converge et sa limite ne peut être que 0.
Finalement (u_n) converge vers 0.

Exemple Soit (u_n) la suite déterminée par

$$u_0 = a \in [0,6] \text{ et } \forall n \in \mathbb{N}, u_{n+1} = \sqrt{6 - u_n}$$

Considérons l'application $f : x \mapsto \sqrt{6-x}$ définie sur $]-\infty, 6]$
Une brève étude de la fonction f permet d'affirmer que

$$\forall x \in [0,6], f(x) \in [0,6]$$

Puisque $u_0 = a \in [0,6]$, la suite (u_n) est bien définie et tous ses termes appartiennent à l'intervalle $[0,6]$
Supposons que la suite (u_n) converge et notons $\ell \in \mathbb{R}$ sa limite.
Puisque pour tout $n \in \mathbb{N}$, $u_n \in [0,6]$, à la limite $\ell \in [0,6]$.
Puisque pour tout $n \in \mathbb{N}$, $u_{n+1} = \sqrt{6 - u_n}$, à la limite $\ell = \sqrt{6 - \ell}$.
On a alors $\ell^2 + \ell - 6 = 0$ d'où $\ell = 2$ ou $\ell = -3$ mais cette dernière solution est à exclure car $\ell \geqslant 0$.
Ainsi, si la suite (u_n) converge, sa limite est 2.
Pour montrer la convergence vers 2, étudions $|u_n - 2|$.
Pour tout $n \in \mathbb{N}$,

$$|u_{n+1} - 2| = \frac{|u_n - 2|}{2 + \sqrt{6 - u_n}} \leqslant \frac{1}{2} |u_n - 2|$$

Par une récurrence facile, on obtient

$$|u_n - 2| \leqslant \frac{1}{2^n} |a - 2| \to 0$$

Par suite, on peut conclure que la suite (u_n) converge vers 2.

18.6.3 Exploitation de l'IAF

Exemple Etudions l'équation $\cos x = x$ d'inconnue $x \in \mathbb{R}$.
Si l'équation possède une solution, celle-ci ne peut être que dans $[-1,1]$.
Considérons alors l'application $f : x \mapsto \cos(x) - x$ définie sur $[-1,1]$.
f est continue, strictement croissante, $f(-1) \geqslant 0$ et $f(1) \leqslant 0$ donc l'équation $f(x) = 0$ possède une unique solution dans $[-1,1]$, notons la α.
Pour obtenir une valeur numérique approchée de α, considérons la suite (u_n) définie par $u_0 \in [-1,1]$ et

$$\forall n \in \mathbb{N}, u_{n+1} = \cos(u_n)$$

La suite (u_n) est bien définie et $\forall n \in \mathbb{N}, u_n \in [-1,1]$
On a

$$|u_{n+1} - \alpha| = |\cos u_n - \cos \alpha|$$

La fonction cos est $\mathcal{C}^1$ et $|(\cos x)'| = |\sin x| \leqslant |\sin 1| = \rho$ avec $\rho \in [0,1[$.
Par l'inégalité des accroissements finis

$$|u_{n+1} - \alpha| \leqslant \rho |u_n - \alpha|$$

Par récurrence on obtient

$$|u_n - \alpha| \leqslant \rho^n |u_0 - \alpha|$$

Puisque $\rho^n \to 0$, on conclut $u_n \to \alpha$.

Remarque Derrière cette étude, se cache le résultat plus général suivant :

Théorème

Soit $f : [a,b] \to [a,b]$ de classe $\mathcal{C}^1$ vérifiant

$$\forall x \in [a,b]\,, |f'(x)| < 1$$

La fonction f admet un unique point fixe α et toute suite récurrente $(u_n) \in [a,b]^{\mathbb{N}}$ de fonction itératrice f converge vers α.

dém. :
Considérons la fonction auxiliaire $g : x \mapsto f(x) - x$ définie sur $[a,b]$.
La fonction g est dérivable et $g'(x) = f'(x) - 1 < 0$ donc g est strictement décroissante. Puisque $g(a) \geqslant 0$ et $g(b) \leqslant 0$, la fonction g s'annule une unique fois et donc l'équation $f(x) = x$ possède une unique solution α.
Considérons maintenant une suite (u_n) définie par $u_0 \in [a,b]$ et pour tout $n \in \mathbb{N}$, $u_{n+1} = f(u_n)$.
Une telle suite est bien définie et formée d'éléments du segment $[a,b]$.
On a

$$|u_{n+1} - \alpha| = |f(u_n) - f(\alpha)|$$

Puisque la fonction $x \mapsto |f'(x)|$ est continue sur le segment $[a,b]$ elle admet un maximum en un point $c \in [a,b]$. En posant ρ la valeur en ce maximum, on vérifie $\rho = |f'(c)| < 1$.
Par l'inégalité des accroissements finis

$$|u_{n+1} - \alpha| \leqslant \rho\, |u_n - \alpha|$$

Par une récurrence facile, on obtient

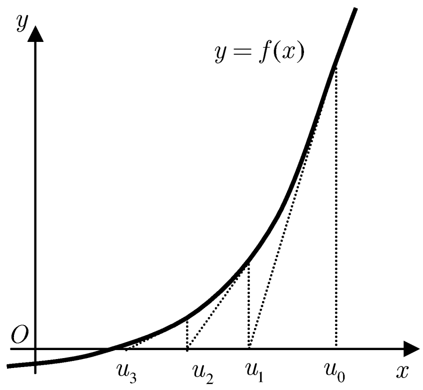
$$|u_n - \alpha| \leqslant \rho^n\, |u_0 - \alpha|$$

Sachant $\rho^n \to 0$, on conclut $u_n \to \alpha$.
□

Remarque On peut alors obtenir une valeur approchée de α grâce aux termes de la suite (u_n)
Concrètement, pour déterminer une valeur décimale de α à 10^{-p} près, on commence par déterminer un rang n vérifiant $\rho^n \leqslant 0,5.10^{-p}$. Pour ce rang, on calcule u_n et on propose à l'aide d'un arrondi une valeur décimale approchée β de u_n à $0,5.10^{-p}$ près. Par cumul d'erreur, on peut affirmer que β est une valeur décimale approchée de α à 10^{-p} près

18.6.4 Méthode de Newton

Soit $f : I \to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$ s'annulant. On désire déterminer une valeur numérique approchée d'une solution α de l'équation $f(x) = 0$ par une méthode plus efficace que la dichotomie.
Pour cela on construit une suite récurrente (u_n) de la manière suivante :
- on choisit $u_0 \in I$ « voisin »de α ;
- pour tout $n \in \mathbb{N}$, on définit u_{n+1} comme étant l'abscisse d'intersection de la tangente à f en u_n avec l'axe (Ox).



Puisque la tangente à f en u_n a pour équation

$$y = f'(u_n)(x - u_n) + f(u_n)$$

on peut exprimer u_{n+1} en fonction de u_n :

$$u_{n+1} = u_n - \frac{f(u_n)}{f'(u_n)}$$

Une telle suite n'est malheureusement pas toujours définie, car on peut avoir $f'(u_n) = 0$ ou encore $u_n \notin I$ ce qui empêche de calculer u_{n+1}. De plus, quand la suite est définie, elle n'est pas nécessaire convergente. Cependant, si elle est définie et qu'elle converge dans I, par passage à la limite de la relation de récurrence, on obverse que sa limite α vérifie l'équation

$$\alpha = \alpha - \frac{f(\alpha)}{f'(\alpha)}$$

et donc $f(\alpha) = 0$.
En pratique la méthode de Newton est satisfaisante, car au voisinage d'une solution, on est souvent dans le cadre du théorème suivant.

Théorème

Soit $f : I \to \mathbb{R}$ une fonction de classe $\mathcal{C}^2$ s'annulant en $\alpha \in I$ telle que f' et f'' ne s'annule pas.
Pour tout $a \in I$ tel que $f(a)f''(a) \geqslant 0$, la suite (u_n) définie par

$$u_0 = a \text{ et } \forall n \in \mathbb{N}, u_{n+1} = u_n - \frac{f(u_n)}{f'(u_n)}$$

est bien définie et converge vers α.
dém. :
Puisque f' est continue et ne s'annule pas, f est strictement monotone.
Puisque f'' est continue et ne s'annule pas, f'' est de signe constant.
Quitte à considérer $-f$ au lieu de f, ce qui ne modifie ni les hypothèses, ni la conclusion, on peut supposer

$$\forall x \in I, f''(x) > 0$$

Soit $a \in I$ tel que $f(a)f''(a) \geqslant 0$ (i.e. $f(a) \geqslant 0$)
Si f est strictement croissante alors $a \geqslant \alpha$ et f est positive sur $[\alpha, a]$.
Considérons alors la fonction itératrice $\varphi : x \mapsto x - f(x)/f'(x)$.
La fonction φ est dérivable, $\varphi'(x) \geqslant 0$, $\varphi(\alpha) = \alpha$ et $\varphi(a) \leqslant a$.
Par suite on a la propriété

$$\forall x \in [\alpha, a]\,, \varphi(x) \in [\alpha, a]$$

La suite (u_n) est donc bien définie et c'est une suite d'éléments de $[\alpha, a]$.
Puisque $u_{n+1} - u_n = -f(u_n)/f'(u_n) \leqslant 0$, la suite (u_n) est décroissante et finalement convergente.
Posons β sa limite. On a $\beta \in [\alpha, a]$ et en passant la relation de récurrence

$$u_{n+1} = u_n - \frac{f(u_n)}{f'(u_n)}$$

à la limite on obtient $f(\beta) = 0$ et donc $\beta = \alpha$.
Cas f strictement décroissante : semblable.

Exemple Considérons la fonction $f : x \mapsto x^2 - 2$ définie sur $[1, +\infty[$
Les hypothèses du théorème sont satisfaites, notamment pour $x_0 = 2$.
La suite (u_n) définie par

$$u_0 = 2 \text{ et } u_{n+1} = \frac{1}{2}\left(u_n + \frac{2}{u_n}\right)$$

converge donc vers $\sqrt{2}$.
Cette convergence est très rapide puisque

$$u_{n+1} - \sqrt{2} = \frac{(u_n - \sqrt{2})^2}{2u_n}$$

et donc

$$\frac{u_{n+1} - \sqrt{2}}{(u_n - \sqrt{2})^2} \to \frac{1}{2\sqrt{2}}$$

Ainsi : « plus on est proche, plus on se rapproche » ; plus précisément « on se rapproche d'une quantité proportionnelle au carrée de la distance ». On parle de convergence quadratique.
De façon générale c'est ce type de convergence que l'on obtient avec la méthode de Newton et c'est pour cela qu'elle est très efficace.

Chapitre 19

Intégration sur un segment

19.1 Fonctions continues par morceaux

a et b désignent deux réels avec $a < b$.

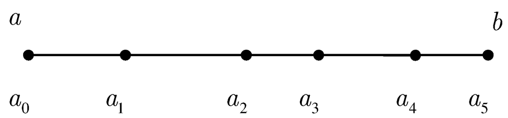
19.1.1 Subdivision d'un segment

Définition

On appelle subdivision d'un segment $[a, b]$, toute famille finie de réels $\sigma = (a_0, a_1, ..., a_n)$ telle que :

$$a = a_0 < a_1 < ... < a_n = b$$

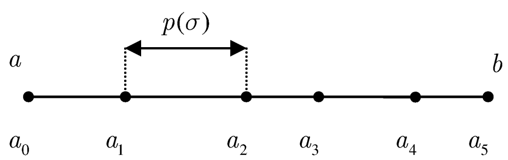
Les a_i sont alors appelés points de la subdivision σ et les intervalles $]a_{i-1}, a_i[$ sont appelés intervalles de la subdivision σ.



Définition

On appelle pas d'une subdivision $\sigma = (a_0, a_1, ..., a_n)$ de $[a, b]$ le réel

$$p(\sigma) = \max_{1 \leqslant i \leqslant n} (a_i - a_{i-1})$$

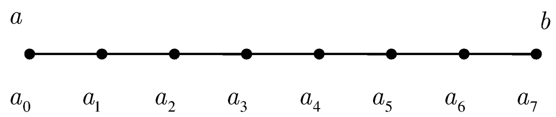


Exemple Pour $n \in \mathbb{N}^\star$, posons

$$a_i = a + i\frac{b-a}{n}$$

pour chaque $i \in \{0, 1, ..., n\}$.
$\sigma = (a_0, a_1, ..., a_n)$ est une subdivision de $[a, b]$ vérifiant

$$\forall 1 \leqslant i \leqslant n, a_i - a_{i-1} = \frac{b-a}{n}$$



Définition

On appelle support d'une subdivision $\sigma = (a_0, a_1, ..., a_n)$ de $[a, b]$ l'ensemble

$$\text{Supp}(\sigma) = \{a_0, a_1, \ldots, a_n\}$$

Proposition

Si S est un ensemble fini d'éléments de $[a, b]$ contenant a et b, il existe une unique subdivision σ du segment $[a, b]$ de support égal à S.

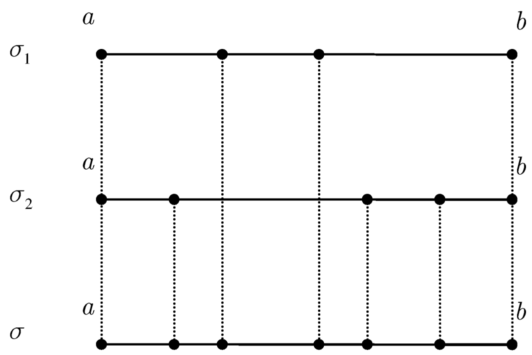
dém. :
Il suffit d'indexer les éléments de S par ordre croissant pour former une subdivision σ de support S.
□

Définition

Soient σ et σ' deux subdivisions d'un même segment $[a, b]$.
On dit que la subdivision σ est plus fine que σ' si $\text{Supp}(\sigma') \subset \text{Supp}(\sigma)$.

Définition

On appelle réunion de deux subdivisions σ_1 et σ_2 d'un même segment $[a, b]$ la subdivision σ de support
$\text{Supp}(\sigma_1) \cup \text{Supp}(\sigma_2)$.



Proposition

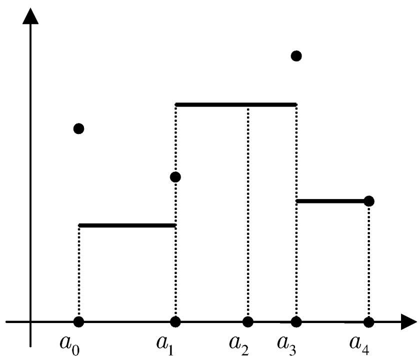
La réunion de deux subdivision est plus fine que chacune.

dém. :
En effet $\text{Supp}(\sigma_1) \subset \text{Supp}(\sigma_1) \cup \text{Supp}(\sigma_2)$ et $\text{Supp}(\sigma_2) \subset \text{Supp}(\sigma_1) \cup \text{Supp}(\sigma_2)$.
□

19.1.2 Fonction en escalier

Définition

Une fonction $f : [a, b] \to \mathbb{R}$ est dite en escalier s'il existe une subdivision $\sigma = (a_0, a_1, \ldots, a_n)$ de
$[a, b]$ telle que f est constante sur chaque intervalle de subdivision $]a_{i-1}, a_i[$.
Une telle subdivision est alors dite adaptée à f.
On note $\mathcal{E}([a, b], \mathbb{R})$ l'ensemble des fonctions en escalier de $[a, b]$ vers $\mathbb{R}$.



Remarque Les valeurs prises par une fonction en escalier aux points de subdivision n'a pas d'importance.

Remarque Si une subdivision σ est adaptée à une fonction en escalier f, toute subdivision plus fine que σ l'est aussi.

Exemple Les fonctions constantes sont des fonctions en escalier.

Proposition

Soient $f, g : [a, b] \to \mathbb{R}$ et $\lambda \in \mathbb{R}$.
Si f et g sont en escalier alors $\lambda.f, f + g, fg, |f|$ le sont aussi.

dém. :
Soit $\sigma = (a_0, \ldots, a_n)$ une subdivision adaptée à f.
Sur chaque intervalle de la subdivision σ, la fonction f est constante et donc les fonctions $\lambda.f$ et $|f|$ sont aussi constantes sur ces intervalles. Ainsi les fonctions $\lambda.f$ et $|f|$ sont en escalier.
Soient σ' une subdivision adaptée à g et σ'' la réunion des subdivisions σ et σ'.
La subdivision σ'' est adaptée à la fois à f et g et donc, sur chaque intervalle de cette subdivision, les

fonctions f et g sont constantes puis les fonctions $f+g$ et fg le sont aussi. Ainsi les fonctions $f+g$ et fg sont en escalier.
□

Proposition

Soit $f:[a,b]\to\mathbb{R}$ et $[\alpha,\beta]\subset[a,b]$ avec $\alpha<\beta$.
Si f est en escalier alors $f_{\restriction[\alpha,\beta]}$ l'est aussi.

dém. :
Soit $\sigma=(a_0,\dots,a_n)$ une subdivision adaptée à f.
Quitte à ajouter des points à la subdivision σ, on peut supposer que α et β sont des points de cette subdivision et il existe alors $k<\ell$ tel que $\alpha=a_k$ et $\beta=a_\ell$.
La famille $(a_k,a_{k+1},\dots,a_\ell)$ forme alors une subdivision de $[\alpha,\beta]$ telle que $f_{\restriction[\alpha,\beta]}$ est constante sur chaque intervalle de subdivision. On en déduit que $f_{\restriction[\alpha,\beta]}$ est en escalier.
□

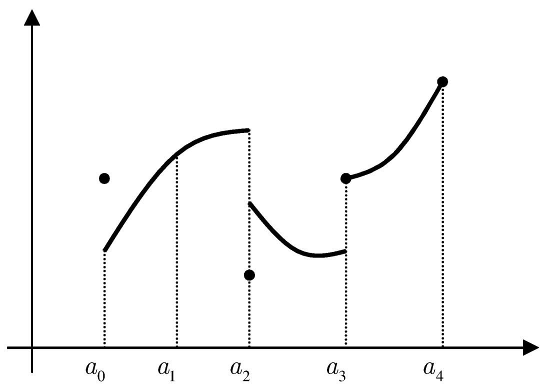
19.1.3 Fonction continue par morceaux

Définition

Une fonction $f:[a,b]\to\mathbb{R}$ est dite continue par morceaux s'il existe une subdivision $\sigma=(a_0,a_1,\dots,a_n)$ de $[a,b]$ vérifiant :

$$\forall 1\leqslant i\leqslant n,\ f \text{ est continue sur }]a_{i-1},a_i[\text{ et } \lim_{a_{i-1}^+} f, \lim_{a_i^-} f \text{ existent et sont finies}$$

Une telle subdivision σ est alors dite adaptée à la fonction f.
On note $\mathcal{C}^0_{pm}([a,b],\mathbb{R})$ l'ensemble des fonctions continue par morceaux de $[a,b]$ vers $\mathbb{R}$.



Exemple Les fonctions continues et les fonctions en escalier sont des fonctions continues par morceaux.

Remarque Les valeurs prises par une fonctions continue par morceaux aux points de subdivision n'importent pas.

Remarque Si une subdivision σ est adaptée à une fonction continue par morceaux alors toute subdivision plus fine que σ l'est aussi.

Proposition

Soit $f, g : [a, b] \to \mathbb{R}$ et $\lambda \in \mathbb{R}$.
Si f et g sont continues par morceaux alors $\lambda.f, f + g, fg, |f|$ le sont aussi.

dém. :
Une subdivision adaptée à f permet d'établir que les fonctions $\lambda.f$ et $|f|$ sont continues par morceaux.
La réunion d'une subdivision adaptée à f et d'une subdivision adaptée à g permet d'établir que les fonctions $f + g$ et fg sont continues par morceaux.

□

Proposition

Soit $f : [a, b] \to \mathbb{R}$ et $[\alpha, \beta] \subset [a, b]$ avec $\alpha < \beta$.
Si f est continue par morceaux alors $f_{\restriction[\alpha,\beta]}$ l'est aussi.

dém. :
Soit $\sigma = (a_0, \dots, a_n)$ une subdivision adaptée à f.
Quitte à ajouter des points à la subdivision σ, on peut supposer que α et β sont des points de cette subdivision et il existe alors $k < \ell$ tel que $\alpha = a_k$ et $\beta = a_\ell$.
La famille $(a_k, a_{k+1}, \dots, a_\ell)$ forme alors une subdivision de $[\alpha, \beta]$ permettant d'affirmer que $f_{\restriction[\alpha,\beta]}$ est continue par morceaux.

□

Proposition

Toute fonction continue par morceaux de $[a, b]$ vers $\mathbb{R}$ est bornée.

dém. :
Soient $f : [a, b] \to \mathbb{R}$ une fonction continue par morceaux et $\sigma = (a_0, a_1, ..., a_n)$ une subdivision adaptée à f.
Pour tout $i \in \{1, \dots, n\}$, considérons la restriction $f_i = f_{\restriction]a_{i-1},a_i[}$.
La fonction f_i est continue sur l'intervalle ouvert $]a_{i-1}, a_i[$ et peut-être prolongée par continuité en a_{i-1} et a_i en posant $f_i(a_{i-1}) = \lim\limits_{a_{i-1}^+} f$ et $f_i(a_i) = \lim\limits_{a_i^-} f$. La fonction ainsi obtenue est continue sur le segment $[a_{i-1}, a_i]$ et par suite y est bornée par un certain réel M_i. En particulier, on obtient $\forall t \in]a_{i-1}, a_i[\,, |f(t)| \leqslant M_i$.
En posant $M = \max(M_1, \dots, M_n, |f(a_0)|, \dots, |f(a_n)|)$, on a $\forall t \in [a, b]\,, |f(t)| \leqslant M$.

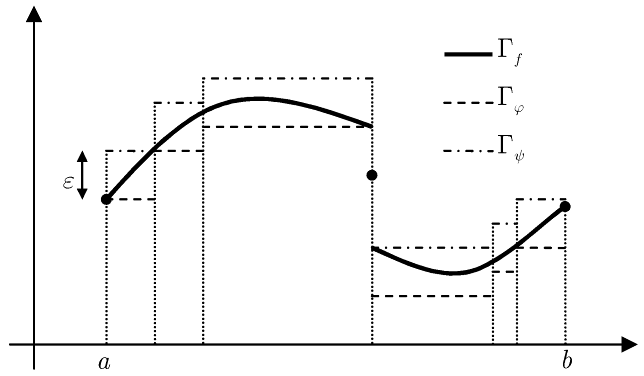
□

19.1.4 Approximation uniforme par des fonctions en escalier

Théorème

Soit $f : [a,b] \to \mathbb{R}$ une fonction continue par morceaux.
Pour tout $\varepsilon > 0$, il existe des fonction $\varphi, \psi : [a,b] \to \mathbb{R}$ en escalier vérifiant

$$\varphi \leqslant f \leqslant \psi \text{ et } 0 \leqslant \psi - \varphi \leqslant \varepsilon$$



dém. :
Commençons par le cas d'une fonction f définie et continue sur $[a,b]$.
Soit $\varepsilon > 0$. Par le théorème de Heine, on peut dire que f est uniformément continue sur $[a,b]$ et par conséquent il existe $\alpha > 0$ vérifiant

$$\forall x, y \in [a,b], |y - x| \leqslant \alpha \Rightarrow |f(x) - f(y)| \leqslant \varepsilon$$

Soit $n \in \mathbb{N}^\star$ suffisamment grand pour que $(b-a)/n \leqslant \alpha$ et $\sigma = (a_0, a_1, \dots, a_n)$ la subdivision à pas constant déterminée par

$$a_i = a + i.\frac{b-a}{n}$$

Pour tout $1 \leqslant i \leqslant n$, la fonction f est continue sur le segment $[a_{i-1}, a_i]$ et y admet donc un minimum et un maximum en des points $\alpha, \beta \in [a_{i-1}, a_i]$. Posons $m_i = f(\alpha)$ et $M_i = f(\beta)$.
Comme $|\beta - \alpha| \leqslant |a_i - a_{i-1}| \leqslant \alpha$ on a $|f(\beta) - f(\alpha)| \leqslant \varepsilon$ et on en déduit que $0 \leqslant M_i - m_i \leqslant \varepsilon$.
Définissons maintenant des fonctions φ et ψ qui vont être solutions de notre problème.
Pour $1 \leqslant i \leqslant n$, on pose φ constante égale à m_i et ψ constante égale à M_i sur $]a_{i-1}, a_i[$.
Pour $0 \leqslant i \leqslant n$, on pose $\varphi(a_i) = \psi(a_i) = f(a_i)$.
Les fonctions φ et ψ sont bien définies sur $[a,b]$ et ce sont évidemment des fonctions en escalier.
L'encadrement $\varphi(x) \leqslant f(x) \leqslant \psi(x)$ est vérifié sur chaque intervalle $]a_{i-1}, a_i[$ et aussi en les a_i.
Enfin, l'encadrement $0 \leqslant \psi(x) - \varphi(x) \leqslant \varepsilon$ est vérifié pour les mêmes raisons.
Il reste à généraliser le résultat aux fonctions continues par morceaux.
Soient f une fonction continue par morceaux sur $[a,b]$, $\sigma = (a_0, a_1, \dots, a_n)$ une subdivision adaptée à f et $\varepsilon > 0$. Pour $1 \leqslant i \leqslant n$, notons f_i la restriction de f au départ de $]a_{i-1}, a_i[$.
On peut prolonger f_i par continuité en a_i et en a_{i-1} en posant $f_i(a_i) = \lim\limits_{a_i^-} f$ et $f_i(a_{i-1}) = \lim\limits_{a_{i-1}^+} f$.
En appliquant le résultat précédent aux fonctions f_i continues sur les segments $[a_{i-1}, a_i]$, il existe des fonctions en escaliers φ_i et ψ_i définies sur $[a_{i-1}, a_i]$ vérifiant

$$\forall x \in [a_{i-1}, a_i], \varphi_i(x) \leqslant f(x) \leqslant \psi_i(x) \text{ et } 0 \leqslant \psi_\mathrm{i}(x) - \varphi_i(x) \leqslant \varepsilon$$

On peut alors définir deux fonctions en escalier φ et ψ qui vont résoudre notre problème.
Pour $1 \leqslant i \leqslant n$, on pose $\varphi(x) = \varphi_i(x)$ et $\psi(x) = \psi_i(x)$ sur $]a_{i-1}, a_i[$.
Pour $0 \leqslant i \leqslant n$, on pose $\varphi(a_i) = \psi(a_i) = f(a_i)$.
Les fonctions φ et ψ sont bien définies sur $[a, b]$ et ce sont évidemment des fonctions en escalier
Par construction, l'encadrement $\varphi(x) \leqslant f(x) \leqslant \psi(x)$ est vérifié sur chaque intervalle $]a_{i-1}, a_i[$ mais aussi en les a_i.
Enfin, l'encadrement $0 \leqslant \psi(x) - \varphi(x) \leqslant \varepsilon$ est aussi vérifié pour les mêmes raisons.
□

19.2 Construction de l'intégrale

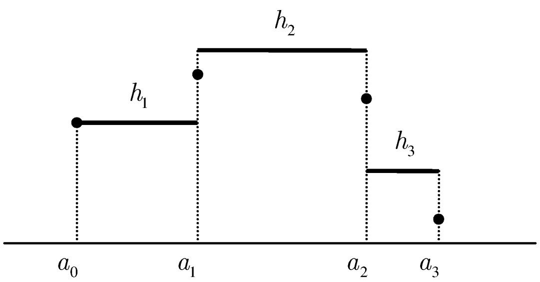
19.2.1 Intégrale d'une fonction en escalier

Soit $a < b \in \mathbb{R}$.

19.2.1.1 Définition

Soit $f : [a, b] \to \mathbb{R}$ une fonction en escalier et $\sigma = (a_0, \ldots, a_n)$ une subdivision adaptée à f.
Pour tout $i \in \{1, \ldots, n\}$ notons h_i la valeur de f sur l'intervalle $]a_{i-1}, a_i[$ et posons

$$I_\sigma(f) = \sum_{i=1}^{n} h_i(a_i - a_{i-1}).$$

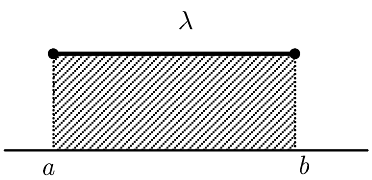


On peut montrer que $I_\sigma(f)$ est indépendante de la subdivision σ adaptée à f choisie. En effet si l'on forme une subdivision σ' en adjoignant un point à la subdivision σ, on montre facilement que $I_\sigma(f) = I_{\sigma'}(f)$. En raisonnant par récurrence, on montre que la propriété perdure si σ' est plus fine que σ. Enfin, en transitant par la réunion des deux subdivisions, on observe que la propriété est encore valable quand σ et σ' sont des subdivisions toutes deux adaptées à f.

Définition

La quantité $I_\sigma(f)$ est appelée intégrale de la fonction en escalier f sur le segment $[a, b]$.
On la note (temporairement) $I_{[a,b]}(f)$.

Exemple Si f est constante égale à λ alors $I_{[a,b]}(f) = \lambda(b - a)$.



Exemple Si $f \in \mathcal{E}([a,b],\mathbb{R})$ est nulle, sauf en un nombre fini de points, $I_{[a,b]}(f) = 0$.

19.2.1.2 Propriétés

Proposition

Pour $f : [a,b] \to \mathbb{R}$ en escalier et $\lambda \in \mathbb{R}$ on a

$$I_{[a,b]}(\lambda.f) = \lambda.I_{[a,b]}(f)$$

dém. :
Soit $\sigma = (a_0, \ldots, a_n)$ une subdivision adaptée à f.
Pour tout $i \in \{1, \ldots, n\}$, notons h_i la valeur de la fonction f sur $]a_{i-1}, a_i[$.
Puisque σ est adaptée à λf et que λh_i est la valeur de la fonction λf sur $]a_{i-1}, a_i[$, on a

$$I_{[a,b]}(\lambda.f) = \sum_{i=1}^{n} \lambda h_i(a_i - a_{i-1}) = \lambda \sum_{i=1}^{n} h_i(a_i - a_{i-1}) = \lambda I_{[a,b]}(f)$$

□

Proposition

Pour $f, g : [a,b] \to \mathbb{R}$ en escalier on a

$$I_{[a,b]}(f+g) = I_{[a,b]}(f) + I_{[a,b]}(g)$$

dém. :
Soit $\sigma = (a_0, \ldots, a_n)$ une subdivision adaptée à f et à g.
Pour tout $i \in \{1, \ldots, n\}$, notons h_i la valeur de la fonction f sur $]a_{i-1}, a_i[$ et k_i la valeur de g sur $]a_{i-1}, a_i[$.
Puisque σ est adaptée à $f+g$ et que $h_i + k_i$ est la valeur de $f+g$ sur $]a_{i-1}, a_i[$, on a

$$I_{[a,b]}(f+g) = \sum_{i=1}^{n} (h_i + k_i)(a_{i-1} - a_i) = \sum_{i=1}^{n} h_i(a_{i-1} - a_i) + \sum_{i=1}^{n} k_i(a_{i-1} - a_i) = I_{[a,b]}(f) + I_{[a,b]}(g)$$

□

Proposition

Soient $f, g : [a,b] \to \mathbb{R}$ en escalier.
Si $f \geqslant 0$ alors $I_{[a,b]}(f) \geqslant 0$.
Si $f \leqslant g$ alors $I_{[a,b]}(f) \leqslant I_{[a,b]}(g)$.

dém. :
Supposons $f \geqslant 0$.
Soit $\sigma = (a_0, \ldots, a_n)$ une subdivision adaptée à f.
Pour tout $i \in \{1, \ldots, n\}$, notons h_i la valeur de la fonction f sur $]a_{i-1}, a_i[$.
Puisque la fonction f est positive, on a $h_i \geqslant 0$ pour tout $i \in \{1, \ldots, n\}$ et donc

$$I_{[a,b]}(f) = \sum_{i=1}^{n} h_i(a_i - a_{i-1}) \geqslant 0$$

Supposons $f \leqslant g$.
Puisque la fonction $g - f$ est positive, l'étude précédente donne $I_{[a,b]}(g - f) \geqslant 0$.
Or $I_{[a,b]}(g - f) = I_{[a,b]}(g) - I_{[a,b]}(f)$ et donc $I_{[a,b]}(g) \geqslant I_{[a,b]}(f)$.
□

Proposition

Soient $f : [a, b] \to \mathbb{R}$ en escalier et $c \in]a, b[$.

$$I_{[a,b]}(f) = I_{[a,c]}(f) + I_{[c,b]}(f)$$

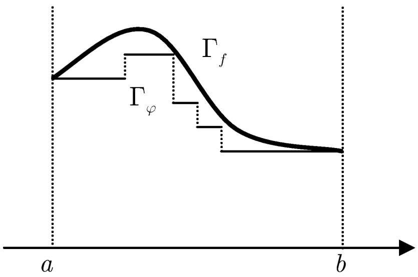
dém. :
Soit $\sigma = (a_0, \ldots, a_n)$ une subdivision adaptée à f.
Quitte à ajouter un point à la subdivision σ, on peut supposer qu'il existe $k \in \{1, \ldots, n - 1\}$ tel que $c = a_k$.
Pour tout $i \in \{1, \ldots, n\}$, notons h_i la valeur de la fonction f sur $]a_{i-1}, a_i[$.
Puisque $(a_0, \ldots, a_k)$ est une subdivision de $[a, c]$ adaptée à $f_{\restriction [a,c]}$, on a $I_{[a,c]}(f) = \sum_{i=1}^{k} h_i(a_i - a_{i-1})$.
Puisque $(a_k, \ldots, a_n)$ est une subdivision de $[c, b]$ adaptée à $f_{\restriction [c,b]}$, on a $I_{[c,b]}(f) = \sum_{i=k+1}^{n} h_i(a_i - a_{i-1})$.
On en déduit

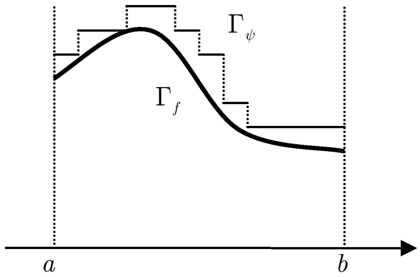
$$I_{[a,c]}(f) + I_{[c,b]}(f) = \sum_{i=1}^{n} h_i(a_i - a_{i-1}) = I_{[a,b]}(f)$$

□

19.2.2 Définition de l'intégrale d'une fonction continue par morceaux.

Soient $a < b \in \mathbb{R}$ et $f : [a, b] \to \mathbb{R}$ continue par morceaux.





Introduisons l'ensemble $\Phi = \{\varphi \in \mathcal{E}([a, b], \mathbb{R}) / \varphi \leqslant f\}$ des fonctions en escalier inférieures à f et celui $\Psi = \{\psi \in \mathcal{E}([a, b], \mathbb{R}) / f \leqslant \psi\}$ des fonctions en escalier supérieures à f.

Théorème

Les bornes $\alpha = \sup \left\{I_{[a,b]}(\varphi) / \varphi \in \Phi\right\}$ et $\beta = \inf \left\{I_{[a,b]}(\psi) / \psi \in \Psi\right\}$ existent et sont égales.

dém. :
Notons $I^- = \left\{I_{[a,b]}(\varphi) / \varphi \in \Phi\right\}$ et $I^+ = \left\{I_{[a,b]}(\psi) / \psi \in \Psi\right\}$ les ensembles formés des intégrales des éléments de Φ et Ψ.
Comme la fonction f est continue par morceaux, elle est bornée et donc il existe $m, M \in \mathbb{R}$ vérifiant $m \leqslant f \leqslant M$. On en déduit que l'ensemble Φ est non vide car $\varphi = m$ est élément de Φ et par suite

l'ensemble I^- est non vide. De même, avec $\psi = M$, on obtient $\Psi \neq \emptyset$ puis $I^+ \neq \emptyset$.
De plus, pour tout $\varphi \in \Phi$, on a $\varphi \leqslant f \leqslant M$ donc

$$I_{[a,b]}(\varphi) \leqslant I_{[a,b]}(M) = M(b-a)$$

Ainsi l'ensemble I^- est majorée par $M(b-a)$.
Finalement I^- est une partie de $\mathbb{R}$ non vide et majorée donc $\alpha = \sup I^-$ existe.
De même $\beta = \inf I^+$ existe car I^+ est une partie de $\mathbb{R}$ non vide et minorée par $m(b-a)$.
Il reste à montrer $\alpha = \beta$.
Pour tout $\varphi \in \Phi, \psi \in \Psi$, on a $\varphi \leqslant f \leqslant \psi$ donc $\varphi \leqslant \psi$ puis $I_{[a,b]}(\varphi) \leqslant I_{[a,b]}(\psi)$.
Par suite $I_{[a,b]}(\varphi)$ est un minorant de I^+ et donc $I_{[a,b]}(\varphi) \leqslant \beta$.
Ainsi β est un majorant de I^- et donc $\alpha \leqslant \beta$.
D'autre part, pour $\varepsilon > 0$, il existe $\varphi, \psi \in \mathcal{E}([a,b], \mathbb{R})$ telles que $\varphi \leqslant f \leqslant \psi$ et $\psi - \varphi \leqslant \varepsilon$.
Puisque $\varphi \in \Phi$ et $\psi \in \Psi$ on a $I_{[a,b]}(\varphi) \leqslant \alpha$ et $\beta \leqslant I_{[a,b]}(\psi)$.
De plus $\psi \leqslant \varphi + \varepsilon$ donc

$$I_{[a,b]}(\psi) \leqslant I_{[a,b]}(\varphi) + I_{[a,b]}(\varepsilon)$$

On en déduit $\beta \leqslant \alpha + \varepsilon(b-a)$.
Cette relation valant pour tout $\varepsilon > 0$, on obtient $\beta \leqslant \alpha$.
Finalement $\beta = \alpha$.
□

Définition

Cette valeur commune est appelée intégrale de f sur $[a,b]$. On la note

$$\int_{[a,b]} f \text{ ou } \int_{[a,b]} f(t)\,\mathrm{d}t$$

Remarque Par construction, la valeur de $\displaystyle\int_{[a,b]} f$ s'interprète comme étant l'aire comprise entre l'axe des abscisses et la courbe Γ_f, celle-ci étant comptée positivement au dessus de l'axe des abscisses et négativement en dessous.

Exemple Si $f : [a,b] \to \mathbb{R}$ est en escalier alors

$$\int_{[a,b]} f = I_{[a,b]}(f)$$

En effet $\displaystyle\int_{[a,b]} f = \max I^- = \min I^+$ puisque $f \in \Phi$ et $f \in \Psi$.

19.2.3 Propriétés de l'intégrale

Soient $a < b \in \mathbb{R}$.

19.2.3.1 Linéarité

Théorème

Soient $f, g : [a, b] \to \mathbb{R}$ continues par morceaux et $\lambda \in \mathbb{R}$

$$\int_{[a,b]} \lambda f = \lambda \int_{[a,b]} f \text{ et } \int_{[a,b]} f + g = \int_{[a,b]} f + \int_{[a,b]} g$$

dém. :
Cas $\lambda \geqslant 0$:
Pour toute fonction $\varphi : [a, b] \to \mathbb{R}$ en escalier inférieure à f, on a $\lambda\varphi \leqslant \lambda f$. Par définition de l'intégrale de la fonction $\lambda\varphi$, on a $I_{[a,b]}(\lambda\varphi) \leqslant \int_{[a,b]} \lambda f$ et donc $\lambda I_{[a,b]}(\varphi) \leqslant \int_{[a,b]} \lambda f$.
En passant cette inégalité en la borne supérieure en $\varphi \in \Phi$, on obtient $\lambda \int_{[a,b]} f \leqslant \int_{[a,b]} \lambda f$.
De même, en partant d'une fonction en escalier ψ supérieure à f, on parvient à $\int_{[a,b]} \lambda f \leqslant \lambda \int_{[a,b]} f$ et on peut conclure $\int_{[a,b]} \lambda f = \lambda \int_{[a,b]} f$.
Cas $\lambda \leqslant 0$:
Même principe mais la multiplication par un scalaire négatif renverse les inégalités.
Etude de la somme :
Soient φ_1 et $\varphi_2 : [a, b] \to \mathbb{R}$ des fonctions en escalier respectivement inférieures à f et g.
La fonction $\varphi_1 + \varphi_2$ est alors inférieure à $f + g$ et par définition de l'intégrale de la fonction $f + g$, on a

$$I_{[a,b]}(\varphi_1 + \varphi_2) = I_{[a,b]}(\varphi_1) + I_{[a,b]}(\varphi_2) \leqslant \int_{[a,b]} f + g$$

En passant cette inégalité en la borne supérieure en φ_1 et φ_2 fonctions en escalier inférieure à f et g, on obtient

$$\int_{[a,b]} f + \int_{[a,b]} g \leqslant \int_{[a,b]} f + g$$

De même, en partant de fonctions en escalier supérieure à f et g, on obtient $\int_{[a,b]} f + g \leqslant \int_{[a,b]} f + \int_{[a,b]} g$ et on peut conclure à l'égalité.
□

Corollaire

Si $f = g$ sauf en un nombre fini de points alors $\int_{[a,b]} f = \int_{[a,b]} g$.

dém. :
$f - g$ est une fonction en escalier à paliers nuls.
□

Remarque On ne modifie pas $\int_{[a,b]} f$ en changeant les valeurs de f en un nombre fini de points.

19.2.3.2 Croissance

Théorème

Soient $f, g : [a,b] \to \mathbb{R}$ continues par morceaux
Si $f \geqslant 0$ alors $\int_{[a,b]} f \geqslant 0$.
Si $f \leqslant g$ alors $\int_{[a,b]} f \leqslant \int_{[a,b]} g$.

dém. :
Si f est positive alors la fonction nulle est une fonction en escalier inférieure à f donc $\int_{[a,b]} f \geqslant I_{[a,b]}(\tilde{0}) = 0$.
Si $f \leqslant g$ alors $g - f \geqslant 0$ donc $\int_{[a,b]} g - f \geqslant 0$ puis par linéarité $\int_{[a,b]} g - \int_{[a,b]} f \geqslant 0$.
□

Remarque Ces résultats se prolongent au cas où les inégalités sont vraies sauf en un nombre fini de points.

Corollaire

Pour $f : [a,b] \to \mathbb{R}$ continue par morceaux, on a

$$\left|\int_{[a,b]} f\right| \leqslant \int_{[a,b]} |f|$$

dém. :
On a $-|f| \leqslant f \leqslant |f|$ donc par croissance

$$-\int_{[a,b]} |f| \leqslant \int_{[a,b]} f \leqslant \int_{[a,b]} |f|$$

□

19.2.3.3 Relation de Chasles

Théorème

Pour $f : [a,b] \to \mathbb{R}$ continue par morceaux et $c \in]a,b[$,

$$\int_{[a,b]} f = \int_{[a,c]} f + \int_{[c,b]} f$$

dém. :
Soient $\varphi_1 : [a,c] \to \mathbb{R}$ une fonction en escalier inférieure à $f_{\restriction [a,c]}$, $\varphi_2 : [c,b] \to \mathbb{R}$ une fonction en escalier inférieure à $f_{\restriction [c,b]}$ et $\varphi : [a,b] \to \mathbb{R}$ la fonction définie par $\varphi(t) = \varphi_1(t)$ si $t \in [a,c]$ et $\varphi(t) = \varphi_2(t)$ si $t \in]c,b]$.
La fonction φ est en escalier et est inférieure à f donc $I_{[a,b]}(\varphi) \leqslant \int_{[a,b]} f$.

Or

$$I_{[a,b]}(\varphi) = I_{[a,c]}(\varphi) + I_{[c,b]}(\varphi) = I_{[a,c]}(\varphi_1) + I_{[c,b]}(\varphi_2)$$

donc

$$I_{[a,c]}(\varphi_1) + I_{[c,b]}(\varphi_2) \leqslant \int_{[a,b]} f$$

En passant à la borne supérieure en φ_1 et φ_2, on obtient

$$\int_{[a,c]} f + \int_{[c,b]} f \leqslant \int_{[a,b]} f$$

De même, en partant de fonctions en escalier supérieure à $f_{\upharpoonright[a,c]}$ et à $f_{\upharpoonright[c,b]}$, on obtient

$$\int_{[a,b]} f \leqslant \int_{[a,c]} f + \int_{[c,b]} f$$

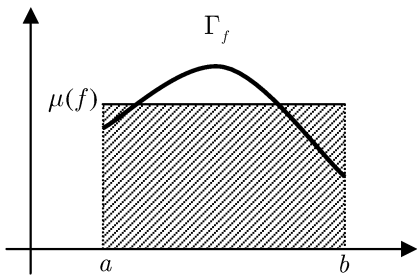
et on peut conclure à l'égalité.
□

19.2.3.4 Inégalité de la moyenne

Définition

On appelle valeur moyenne d'une fonction $f : [a,b] \to \mathbb{R}$ continue par morceaux le réel

$$\mu(f) = \frac{1}{b-a} \int_{[a,b]} f$$



Remarque L'aire du rectangle est égale à l'aire sous la courbe.

Théorème

Pour $f, g : [a,b] \to \mathbb{R}$ continues par morceaux,

$$\left|\int_{[a,b]} fg\right| \leqslant \sup_{[a,b]} |f| \int_{[a,b]} |g|$$

dém. :
Par l'inégalité triangulaire

$$\left|\int_{[a,b]} fg\right| \leqslant \int_{[ab]} |fg|$$

Or en notant $M = \sup\limits_{[a,b]} |f|$, on a $|fg| \leqslant M\,|g|$ donc par croissance de l'intégrale

$$\left|\int_{[a,b]} fg\right| \leqslant \int_{[ab]} M\,|g|$$

et enfin par linéarité

$$\left|\int_{[a,b]} fg\right| \leqslant M \int_{[ab]} |g|$$

□

19.2.3.5 Inégalité de Cauchy Schwarz

Théorème

Pour $f, g : [a,b] \to \mathbb{R}$ continues

$$\left(\int_{[a,b]} fg\right)^2 \leqslant \left(\int_{[a,b]} f^2\right)\left(\int_{[a,b]} g^2\right)$$

dém. :
Par linéarité

$$\int_{[a,b]} (\lambda f + g)^2 = \lambda^2 \int_{[a,b]} f^2 + 2\lambda \int_{[a,b]} f + \int_{[a,b]} g^2$$

Or $(\lambda f + g)^2 \geqslant 0$ donc par croissance $\displaystyle\int_{[a,b]} (\lambda f + g)^2 \geqslant 0$.

Si $\displaystyle\int_{[a,b]} f^2 \neq 0$ alors la fonction du second degré

$$\lambda \mapsto \lambda^2 \int_{[a,b]} f^2 + 2\lambda \int_{[a,b]} fg + \int_{[a,b]} g^2$$

étant de signe constant son discriminant Δ est négatif ce qui donne l'inégalité voulue.

Si $\displaystyle\int_{[a,b]} f^2 = 0$ alors la fonction affine

$$\lambda \mapsto 2\lambda \int_{[a,b]} fg + \int_{[a,b]} g^2$$

étant de signe constant le coefficient de λ dans son expression doit être nulle et donc $\displaystyle\int_{[a,b]} fg = 0$ et l'inégalité voulue est encore vraie.

□

19.2.4 Intégrale de deux bornes

Soit I un intervalle de $\mathbb{R}$.

Définition

Une fonction $f : I \to \mathbb{R}$ est dite continue par morceaux si pour tout $[a,b] \subset I$, $f_{\restriction [a,b]}$ est continue par morceaux. On note $\mathcal{C}^0_{pm}(I,\mathbb{R})$ l'ensemble des fonctions continue par morceaux de I vers $\mathbb{R}$.

Exemple La fonction partie entière $x \mapsto E(x)$ est continue par morceaux sur $\mathbb{R}$.

Exemple Les fonctions continue de I vers $\mathbb{R}$ sont continues par morceaux.

Définition

Soient $f : I \to \mathbb{R}$ continue par morceaux et $a, b \in I$.
On définit $\displaystyle\int_a^b f$ (ou encore $\displaystyle\int_a^b f(t)\,\mathrm{d}t$) par :

$$\int_a^b f = \begin{cases} \displaystyle\int_{[a,b]} f & \text{si } a < b \\ 0 & \text{si } a = b \\ \displaystyle -\int_{[b,a]} f & \text{si } a > b \end{cases}$$

Remarque Pour tout $a, b \in I$, $\displaystyle\int_b^a f = -\int_a^b f$.

Exemple $\forall a, b \in \mathbb{R}$, $\displaystyle\int_a^b 1\,\mathrm{d}t = b - a$.

Théorème

Soient $f, g : I \to \mathbb{R}$ continues par morceaux et $\lambda \in \mathbb{R}$. Pour tout $a, b \in I$ on a

$$\int_a^b \lambda f = \lambda \int_a^b f \text{ et } \int_a^b f + g = \int_a^b f + \int_a^b g$$

Théorème

Soient $f : I \to \mathbb{R}$ continue par morceaux. Pour tout $a, b, c \in I$,

$$\int_a^b f = \int_a^c f + \int_c^b f$$

dém. :
Il suffit de discuter selon les positions relatives de a, b, c dans I.
□
(!)Si $f \geqslant 0$ alors $\displaystyle\int_a^b f \geqslant 0$ lorsque $a \leqslant b$ et $\displaystyle\int_b^a f \leqslant 0$ lorsque $b \leqslant a$.
Lors d'intégration d'inégalités, il faut vérifier le bon ordre des bornes d'intégration.

19.3 Primitives et intégrales

I désigne un intervalle de $\mathbb{R}$ non singulier.

19.3.1 Primitives d'une fonction

Définition

On appelle primitive d'une fonction $f : I \to \mathbb{R}$, s'il en existe, toute fonction $F : I \to \mathbb{R}$ dérivable vérifiant $F' = f$.

Proposition

Si $f : I \to \mathbb{R}f$ admet une primitive F alors l'ensemble des primitives de f est constitué des fonctions de la forme $t \mapsto F(t) + C$ avec $C \in \mathbb{R}$

dém. :
Soit $G : I \to \mathbb{R}$ dérivable.
G est primitive de f si, et seulement si, $(G - F)' = 0$ c'est-à-dire si, et seulement si, $G - F$ est une fonction constante.

□

Définition

On note $\displaystyle\int f(t)\,\mathrm{d}t = F(t) + C^{te}$ pour signifier que F est une primitive de f et pour rappeler que les autres s'obtiennent par addition d'une constante.
Abusivement, on écrit parfois aussi $\displaystyle\int f(t)\,\mathrm{d}t = F(t)$

Exemple Pour $f : I \to \mathbb{R}$ dérivable, $\displaystyle\int f'(t)\,\mathrm{d}t = f(t) + C^{te}$.

Proposition

Soient $f, g : I \to \mathbb{R}$ et $\lambda \in \mathbb{R}$.
Si F et G sont primitives de f et g alors λF et $F + G$ sont respectivement primitives de λf et $f + g$.

dém. :
$(\lambda F)' = \lambda F' = \lambda f$ et $(F + G)' = F' + G' = f + g$.

□

19.3.2 Primitives de fonctions usuelles

Le tableau suivant exprime des primitives d'expressions fonctionnelles usuelles

$f(t)$	$\int f(t)\,\mathrm{d}t$	I
t^n avec $n \in \mathbb{N}$	$\dfrac{t^{n+1}}{n+1} + C^{te}$	$\mathbb{R}$
$\dfrac{1}{t^n}$ avec $n \in \mathbb{N}\setminus\{0,1\}$	$-\dfrac{1}{n-1}\dfrac{1}{t^{n-1}} + C^{te}$	$\mathbb{R}^{+\star}$ ou $\mathbb{R}^{-\star}$
t^α avec $\alpha \in \mathbb{R}\setminus\{-1\}$	$\dfrac{1}{\alpha+1}t^{\alpha+1} + C^{te}$	$\mathbb{R}^{+\star}$
$\dfrac{1}{t}$	$\ln\lvert t\rvert + C^{te}$	$\mathbb{R}^{+\star}$ ou $\mathbb{R}^{-\star}$
e^t	$\mathrm{e}^t + C^{te}$	$\mathbb{R}$
$\ln t$	$t\ln t - t + C^{te}$	$\mathbb{R}^{+\star}$
$\sin t$	$-\cos t + C^{te}$	$\mathbb{R}$
$\cos t$	$\sin t + C^{te}$	$\mathbb{R}$
$\mathrm{sh}t$	$\mathrm{ch}t + C^{te}$	$\mathbb{R}$
$\mathrm{ch}t$	$\mathrm{sh}t + C^{te}$	$\mathbb{R}$
$\dfrac{1}{1+t^2}$	$\arctan t + C^{te}$	$\mathbb{R}$
$\dfrac{1}{1-t^2}$	$\dfrac{1}{2}\ln\left\lvert\dfrac{1+t}{1-t}\right\rvert + C^{te}$	$]-\infty,-1[\,,\,]-1,1[$ et $]1,+\infty[$
$\dfrac{1}{\sqrt{1-t^2}}$	$\arcsin t + C^{te}$	$]-1,1[$
$\dfrac{1}{\sqrt{1+t^2}}$	$\ln(t+\sqrt{t^2+1}) + C^{te}$	$\mathbb{R}$
$\dfrac{1}{\sqrt{t^2-1}}$	$\ln\left\lvert t+\sqrt{t^2-1}\right\rvert + C^{te}$	$]-\infty,-1]$ et $[1,+\infty[$

19.3.3 Calcul de primitives par linéarisation

Pour déterminer une primitive d'une expression formée par une somme de produits de fonctions $\cos$ et $\sin$, on linéarise l'expression étudiée.

Exemple Détermination de $\displaystyle\int \cos^4 t \sin^2 t\,\mathrm{d}t$ sur $\mathbb{R}$.

$$\cos^4 t \sin^2 t = \left(\frac{\mathrm{e}^{it}+\mathrm{e}^{-it}}{2}\right)^4 \left(\frac{\mathrm{e}^{it}-\mathrm{e}^{-it}}{2i}\right)^2$$

En développant et en recombinant les exponentielles imaginaires, on obtient

$$\cos^4 t \sin^2 t = -\frac{1}{32}\left(\cos 6t + 2\cos 4t - \cos 2t - 2\right)$$

On en déduit

$$\int \cos^4 t \sin^2 t\,\mathrm{d}t = -\frac{1}{192}\sin 6t - \frac{1}{64}\sin 4t + \frac{1}{64}\sin 2t + \frac{t}{16} + C^{te}$$

Exemple Détermination $\displaystyle\int \cos^2 t\,\mathrm{d}t$ et $\displaystyle\int \sin^2 t\,\mathrm{d}t$

Via $\cos 2t = 2\cos^2 t - 1 = 1 - 2\sin^2 t$, on obtient

$$\int \cos^2 t\,\mathrm{d}t = \frac{1}{4}\sin 2t + \frac{1}{2}t + C^{te} \text{ et } \int \sin^2 t\,\mathrm{d}t = -\frac{1}{4}\sin 2t + \frac{t}{2} + C^{te}$$

Exemple Détermination de $\displaystyle\int \cos t \cos 2t \, \mathrm{d}t$ sur $\mathbb{R}$.
Via
$$\cos a \cos b = \frac{1}{2}\left(\cos(a+b) + \cos(a-b)\right)$$
on obtient
$$\cos t \cos 2t = \frac{1}{2}(\cos 3t + \cos t)$$
puis
$$\int \cos t \cos 2t \, \mathrm{d}t = \frac{1}{6}\sin 3t + \frac{1}{2}\sin t + C^{te}$$

Pour déterminer une primitive d'une expression formée par une somme de produits de fonctions ch et sh, on procède par linéarisation via les exponentielles réelles

Exemple Détermination de $\displaystyle\int \mathrm{ch}^2 t \mathrm{sh}^2 t \, \mathrm{d}t$ sur $\mathbb{R}$.
$$\mathrm{ch}^2 t \mathrm{sh}^2 t = \left(\frac{\mathrm{e}^t + \mathrm{e}^{-t}}{2}\right)^2 \left(\frac{\mathrm{e}^t - \mathrm{e}^{-t}}{2}\right)^2 = \left(\frac{\mathrm{e}^{2t} - \mathrm{e}^{-2t}}{4}\right)^2$$
En développant et en recombinant les exponentielles, on obtient
$$\mathrm{ch}^2 t \mathrm{sh}^2 t = \frac{1}{8}\mathrm{ch} 4t - \frac{1}{8}$$
puis
$$\int \mathrm{ch}^2 t \mathrm{sh}^2 t \, \mathrm{d}t = \frac{1}{32}\mathrm{sh} 4t - \frac{t}{8} + C^{te}$$

19.4 Intégrales et primitives

19.4.1 Théorème fondamental de l'intégration

Théorème

Si $f : I \to \mathbb{R}$ est une fonction continue alors pour tout $a \in I$, f possède une unique primitive qui s'annule en a, c'est la fonction
$$x \mapsto \int_a^x f(t) \, \mathrm{d}t$$

dém. :
Unicité :
Soient F et G deux fonctions solutions.
Puisque F et G sont primitives d'une même fonction, elles diffèrent d'une constante. Or elles prennent la même valeur en a, la constante est donc nulle et les fonctions F et G égales.
Existence :
Soit $F : I \to \mathbb{R}$ la fonction définie par
$$F(x) = \int_a^x f(t) \, \mathrm{d}t$$

On a immédiatement $F(a) = 0$.
Soit $x \in I$ qui n'est pas extrémité droite de I.

$$\frac{1}{h}\left(F(x+h) - F(x)\right) - f(x) = \frac{1}{h}\int_x^{x+h} f(t) - f(x)\,\mathrm{d}t$$

Pour $h > 0$,

$$\left|\frac{1}{h}\left(F(x+h) - F(x)\right) - f(x)\right| \leqslant \frac{1}{h}\int_x^{x+h} |f(t) - f(x)|\,\mathrm{d}t$$

Soit $\varepsilon > 0$. Puisque f est continue en x, il existe $\alpha > 0$ tel que

$$\forall t \in I, |t - x| \leqslant \alpha \Rightarrow |f(t) - f(x)| \leqslant \varepsilon$$

Pour $0 < h \leqslant \alpha$ on a

$$\forall t \in [x, x+h]\,,\ |f(t) - f(x)| \leqslant \varepsilon$$

puis

$$\left|\frac{1}{h}\left(F(x+h) - F(x)\right) - f(x)\right| \leqslant \frac{1}{h}\int_x^{x+h} \varepsilon\,\mathrm{d}t = \varepsilon$$

On en déduit

$$\lim_{h \to 0^+} \frac{1}{h}\left(F(x+h) - F(x)\right) = f(x)$$

De même, pour $x \in I$ qui n'est pas extrémité gauche de I, on a

$$\lim_{h \to 0^-} \frac{1}{h}\left(F(x+h) - F(x)\right) = f(x)$$

On en déduit que F est dérivable en chaque $x \in I$ et on a $F'(x) = f(x)$.
□

Corollaire

Toute fonction réelle continue sur un intervalle I y admet des primitives.

Exemple On peut définir la fonction $\ln$ en tant que primitive s'annulant en 1 de la fonction inverse sur l'intervalle $]0, +\infty[$.

Corollaire

Soient $f : I \to \mathbb{R}$ continue et F une primitive de f.

$$\forall a, b \in I, \int_a^b f(t)\,\mathrm{d}t = [F(t)]_a^b = F(b) - F(a)$$

dém. :
Considérons l'application $G : x \mapsto \displaystyle\int_a^x f(t)\,\mathrm{d}t$.
La fonction G est une primitive de f, donc il existe $C \in \mathbb{R}$ tel que $G = F + C$.
Puisque $G(a) = F(a) + C = 0$, on a $C = -F(a)$ puis $\displaystyle\int_a^b f(t)\,\mathrm{d}t = G(b) = F(b) + C = F(b) - F(a)$.
□

Exemple $\displaystyle\int_0^1 t^n \,\mathrm{d}t = \left[\frac{1}{n+1}t^{n+1}\right]_0^1 = \frac{1}{n+1}.$

Exemple $\displaystyle\int_0^1 \frac{\mathrm{d}t}{1+t^2} = [\arctan t]_0^1 = \frac{\pi}{4}.$

Exemple $\displaystyle\int_0^{\pi/2} \cos^2 t \,\mathrm{d}t = \left[\frac{1}{4}\sin(2t) + \frac{t}{2}\right]_0^{\pi/2} = \frac{\pi}{4}.$

19.4.2 Positivité de l'intégrale d'une fonction continue

Théorème

> Soient $a < b \in \mathbb{R}$ et $f : [a,b] \to \mathbb{R}$.
> Si f est continue, positive et si $\displaystyle\int_a^b f(t)\,\mathrm{d}t = 0$ alors $f = \tilde{0}$.

dém. :
Considérons la fonction $F : x \mapsto \displaystyle\int_a^x f(t)\,\mathrm{d}t$.
La fonction F est croissante car $F' = f \geqslant 0$.
De plus $F(b) = F(a)$ car $\displaystyle\int_a^b f(t)\,\mathrm{d}t = 0$.
On en déduit que F est constante puis $f = F' = \tilde{0}$.
□

Corollaire

> Soient $a < b$ et $f : [a,b] \to \mathbb{R}$ continue.
> Si $\displaystyle\int_a^b |f(t)|\,\mathrm{d}t = 0$ alors $f = 0$.
> Si $\displaystyle\int_a^b f^2(t)\,\mathrm{d}t = 0$ alors $f = 0$.

dém. :
Les fonctions $|f|$ et f^2 sont continues et positives.
□

Corollaire

> Soient $a < b \in \mathbb{R}$ et $f : [a,b] \to \mathbb{R}$ continue.
> Si $f \geqslant 0$ et $f \neq \tilde{0}$ alors $\displaystyle\int_a^b f(t)\,\mathrm{d}t > 0$.
> Si $\displaystyle\int \frac{\mathrm{d}t}{\sqrt{a^2 - t^2}} \underset{x=ta}{=} \arcsin\frac{t}{a} + C^{te}$ et $f \neq \tilde{0}$ alors $\displaystyle\int_a^b f(t)\,\mathrm{d}t < 0$.

dém. :
C'est une contraposition de l'implication énoncée dans le théorème.
□

Remarque Cet outil permet de justifier des inégalités strictes entre intégrales.

Corollaire

Soient $a < b \in \mathbb{R}$ et $f : [a,b] \to \mathbb{R}$ continue.
Si $\int_a^b f(t)\,\mathrm{d}t = 0$ alors la fonction f s'annule.

dém. :
Si f est de signe constant alors f est la fonction nulle.
Sinon f change de signe et donc s'annule en vertu du théorème des valeurs intermédiaires.
□

Exemple Soit $f : [a,b] \to \mathbb{R}$ continue. Montrons qu'il existe $c \in [a,b]$ tel que $\mu(f) = f(c)$.
Considérons la fonction continue $g : t \mapsto f(t) - \mu(f)$.
Puisque

$$\int_a^b g(t)\,\mathrm{d}t = \int_a^b f(t)\,\mathrm{d}t - (b-a)\mu(f) = 0$$

la fonction g s'annule et donc il existe $c \in [a,b]$ tel que $g(c) = 0$ ce qui donne $f(c) = \mu(f)$.
Ce résultat peut aussi être établi en appliquant le théorème des accroissements finis.

Exemple Soit $f : [0,1] \to \mathbb{R}$ continue telle que

$$\int_0^1 f(t)\,\mathrm{d}t = \int_0^1 tf(t)\,\mathrm{d}t = 0$$

Montrons que f s'annule au moins deux fois dans $[0,1]$.
Puisque $\int_0^1 f(t)\,\mathrm{d}t = 0$, il existe déjà $\alpha \in \left]0,1\right[$ tel que $f(\alpha) = 0$.
Par l'absurde, supposons qu'il n'y a pas d'autres annulations que α.
La fonction f ne peut pas être de signe constant car sinon c'est la fonction nulle ce qui est exclu.
Le tableau de signe de f est alors

$$\begin{array}{c|ccc} t & 0 & \alpha & 1 \\ \hline f(t) & + & 0 & - \end{array} \text{ ou } \begin{array}{c|ccc} t & 0 & \alpha & 1 \\ \hline f(t) & - & 0 & + \end{array}$$

Dans les deux cas, on a $\int_a^b (t-\alpha)f(t)\,\mathrm{d}t \neq 0$ car intégrale d'une fonction continue de signe constant non nulle.
Or

$$\int_a^b (t-\alpha)f(t)\,\mathrm{d}t = \int_a^b tf(t)\,\mathrm{d}t - \alpha\int_a^b f(t)\,\mathrm{d}t = 0$$

C'est absurde.

19.4.3 Suites d'intégrales

(!)Il n'est pas possible de passer directement les limites à l'intégrale :
$(\forall t \in [a,b], f_n(t) \to f(t))$ n'implique pas $\int_a^b f_n(t)\,\mathrm{d}t \to \int_a^b f(t)\,\mathrm{d}t$!

Exemple Soit $f_n : [0,1] \to \mathbb{R}$ définie par $f_n(t) = n^2t^n(1-t)$.

Pour tout $t \in [0,1[$ et pour $t = 1$, on vérifie $f_n(t) \xrightarrow[n\to+\infty]{} 0$.

Or $\displaystyle\int_0^1 n^2 t^n(1-t)\,\mathrm{d}t = n^2\int_0^1 t^n - t^{n+1}\,\mathrm{d}t = \frac{n^2}{(n+1)(n+2)} \xrightarrow[n\to+\infty]{} 1 \neq 0$

Remarque Pour étudier la limite d'une intégrale qu'on ne sait pas calculer, on peut néanmoins considérer ce qui précède pour « deviner »l'éventuelle limite puis justifier celle-ci en procédant par comparaison...

Exemple Etudions la limite de la suite de terme général

$$I_n = \int_0^1 t^n \mathrm{e}^{-t}\,\mathrm{d}t$$

Puisque la fonction intégrée tend « essentiellement »vers 0 quand $n \to +\infty$, on peut présumer que (I_n) tend vers 0... Montrons par comparaison $|I_n - 0| \xrightarrow[n\to+\infty]{} 0$.

$$|I_n - 0| = \int_0^1 t^n \mathrm{e}^t\,\mathrm{d}t \leqslant \mathrm{e}\int_0^1 t^n\,\mathrm{d}t = \frac{\mathrm{e}}{n+1} \to 0$$

Effectivement, la suite (I_n) tend vers 0.

Exemple Etudions la limite de la suite de terme général

$$I_n = \int_0^1 \frac{\mathrm{d}t}{1+t^n}$$

Puisque la fonction intégrée tend « essentiellement »vers 1 quand $n \to +\infty$, on peut présumer que (I_n) tend vers $\displaystyle\int_0^1 1\,\mathrm{d}t = 1$... Etudions $|I_n - 1|$.

$$|I_n - 1| = \left|\int_0^1 \frac{\mathrm{d}t}{1+t^n} - \int_0^1 1\,\mathrm{d}t\right| = \int_0^1 \frac{t^n}{1+t^n}\,\mathrm{d}t \leqslant \int_0^1 t^n\,\mathrm{d}t = \frac{1}{n+1} \to 0$$

19.4.4 Fonction définie par une intégrale dont les bornes dépendent de la variable

Soit $f : I \to \mathbb{R}$ continue et deux fonctions $u, v : J \to \mathbb{R}$ deux fonctions dérivables prenant leurs valeurs dans I.

Considérons la fonction $\varphi : x \mapsto \displaystyle\int_{u(x)}^{v(x)} f(t)\,\mathrm{d}t$.

Pour étudier la dérivée de φ, il suffit d'introduire une primitive F de la fonction f. En effet, on peut alors écrire $g(x) = F(v(x)) - F(u(x))$ ce qui permet d'affirmer que g est dérivable et d'obtenir

$$g'(x) = v'(x)f(v(x)) - u'(x)f(u(x))$$

Exemple Pour $x > 0$, on pose
$$\varphi(x) = \int_{1/x}^{x} \frac{t}{1+t+t^2+t^3}\,\mathrm{d}t$$
Puisque la fonction $f : t \mapsto \dfrac{t}{1+t+t^2+t^3}$ est définie et continue sur l'intervalle $\mathbb{R}^+$, pour tout $x > 0$, il est possible d'intégrer cette fonction entre les bornes $1/x$ et x.
Ainsi la fonction f est bien définie.
Soit F une primitive sur $\mathbb{R}^+$ de la fonction continue f.
Pour tout $x > 0$, on a $\varphi(x) = F(x) - F(1/x)$.
Puisque la fonction F est de classe $\mathcal{C}^1$, on peut affirmer que φ l'est aussi par opérations sur les fonctions de classe $\mathcal{C}^1$
De plus
$$\varphi'(x) = F'(x) - \frac{1}{x^2}F(1/x) = \frac{x}{1+x+x^2+x^3} + \frac{1}{x^3+x^2+x+1} = \frac{1}{x^2+1}$$
On en déduit qu'il existe une constante C telle que $\varphi(x) = \arctan x + C$.
Puisque $\varphi(1) = 0$, on a $C = -\pi/4$ puis $\varphi(x) = \arctan x - \pi/4$.

Exemple Soit $\varphi :]0,1[\to \mathbb{R}$ définie par
$$\varphi(x) = \int_{x}^{x^2} \frac{\mathrm{d}t}{\ln t}$$
La fonction $f : t \mapsto \dfrac{1}{\ln t}$ est définie et continue sur l'intervalle $]0,1[$ et pour tout $x \in]0,1[$, on a $x, x^2 \in]0,1[$ et donc il est possible d'intégrer f entre x et x^2.
Par suite la fonction φ est bien définie sur $]0,1[$.
Soit F une primitive sur $]0,1[$ de la fonction continue f.
Pour tout $x \in]0,1[$,
$$\varphi(x) = F(x^2) - F(x)$$
Puisque la fonction F est de classe $\mathcal{C}^1$, on peut affirmer que φ l'est aussi par opérations sur les fonctions de classe $\mathcal{C}^1$
De plus
$$\varphi'(x) = 2f(x^2) - f(x) = \frac{x-1}{\ln x}$$
Cela ne permet pas d'exprimer φ mais cependant nous avons amorcée l'étude de cette fonction...

19.5 Intégration par parties

Soit I un intervalle non singulier.

19.5.1 Calcul de primitive par parties

Soient u, v deux fonctions dérivables sur I.
La fonction uv est dérivable sur I et $(uv)' = u'v + uv'$.
On en déduit la formule
$$\int u'v = \int (uv)' - uv' = uv - \int uv'$$

Remarque Cette formule est utile pour déterminer des primitives de fonctions dont la dérivée est simple :

Exemple Sur $\mathbb{R}^{+\star}$,

$$\int \ln t \, \mathrm{d}t = \int 1 \times \ln t \, \mathrm{d}t = t \ln t - \int t \times \frac{1}{t} \, \mathrm{d}t = t \ln t - t + C^{te}$$

Exemple Sur $\mathbb{R}$,

$$\int \arctan t \, \mathrm{d}t = \int 1 \times \arctan t \, \mathrm{d}t = t \arctan t - \int \frac{t}{1+t^2} \, \mathrm{d}t = t \arctan t - \frac{1}{2} \ln(1+t^2) + C^{te}$$

Exemple Sur $]-1,1[$,

$$\int \arcsin t \, \mathrm{d}t = \int 1 \times \arcsin t \, \mathrm{d}t = t \arcsin t - \int \frac{t \, \mathrm{d}t}{\sqrt{1-t^2}} = t \arcsin t + \sqrt{1-t^2} + C^{te}$$

19.5.2 Détermination de $\int P(x)\mathrm{e}^{\alpha x} \, \mathrm{d}x$

Soit P une fonction polynomiale de degré $n \in \mathbb{N}$ et $\alpha \in \mathbb{R}^\star$.
Deux méthodes sont usuelles :
(1) Par calculs successifs de primitives par parties, on supprime le terme polynomial à force de dérivation :

$$\int P(x)\mathrm{e}^{\alpha x} \, \mathrm{d}x = \frac{1}{\alpha} P(x)\mathrm{e}^{\alpha x} - \frac{1}{\alpha} \int P'(x)\mathrm{e}^{\alpha x} \, \mathrm{d}x$$
$$= \left(\frac{1}{\alpha} P(x) - \frac{1}{\alpha^2} P'(x) + \cdots + \frac{(-1)^n}{\alpha^{n+1}} P^{(n)}(x) \right) \mathrm{e}^{\alpha x} + C^{te}$$

(2) En vertu de l'étude ci-dessus, on sait $\int P(x)\mathrm{e}^{\alpha x} \, \mathrm{d}x = Q(x)\mathrm{e}^{\alpha x} + C^{te}$ avec Q fonction polynomiale de degré n. On peut alors rechercher Q par coefficients inconnus de sorte que

$$(Q(x)\mathrm{e}^{\alpha x})' = (Q'(x) + \alpha Q(x))\mathrm{e}^{\alpha x} = P(x)\mathrm{e}^{\alpha x}$$

Remarque Si $\deg P \leqslant 2$, il est efficace d'utiliser (1)
Si $\deg P \geqslant 3$, il est efficace d'utiliser (2)

Exemple Déterminons $\int (x^2+1)\mathrm{e}^{2x} \, \mathrm{d}x$ sur $\mathbb{R}$.

$$\int (x^2+1)\mathrm{e}^{2x} \, \mathrm{d}x = \frac{(x^2+1)}{2} \mathrm{e}^{2x} - \int x \, \mathrm{e}^{2x} \, \mathrm{d}x = \frac{x^2-x+1}{2} \mathrm{e}^{2x} + \int \frac{1}{2} \mathrm{e}^{2x}$$
$$= \left(\frac{1}{2}x^2 - \frac{1}{2}x + \frac{3}{4} \right) \mathrm{e}^{2x} + C^{te}$$

Exemple Déterminons $\int (x^3 + x^2)\,\mathrm{e}^{\,-x}\,\mathrm{d}x$ sur $\mathbb{R}$.
La primitive cherchée est de la forme

$$\int (x^3 + x)\mathrm{e}^{-x}\,\mathrm{d}x = (ax^3 + bx^2 + cx + d)\mathrm{e}^{-x} + C^{te}$$

Or

$$\big((ax^3 + bx^2 + cx + d)\mathrm{e}^{-x}\big)' = (-ax^3 + (3a - b)x^2 + (2b - c)x + (c - d))\mathrm{e}^{-x}$$

On est donc ramené à résoudre le système

$$\begin{cases} -a = 1 \\ 3a - b = 1 \\ 2b - c = 0 \\ c - d = 0 \end{cases}$$

et on obtient pour solution

$$\begin{cases} a = -1 \\ b = -4 \\ c = -8 \\ d = -8 \end{cases}$$

19.5.3 Détermination de $\int P(x)\cos(\alpha x)\,\mathbf{d}x$ et $\int P(x)\sin(\alpha x)\,\mathbf{d}x$

Soit P une fonction polynomiale de degré $n \in \mathbb{N}$ et $\alpha \in \mathbb{R}^\star$.
Deux méthodes sont usuelles, pour déterminer $\int P(x)\cos(\alpha x)\,\mathrm{d}x$
(1) Par calculs successifs de primitives par parties, on supprime le terme polynomial à force de dérivation

$$\int P(x)\cos(\alpha x)\,\mathrm{d}x = \frac{1}{\alpha}P(x)\sin\alpha x - \frac{1}{\alpha}\int P'(x)\sin(\alpha x)\,\mathrm{d}x$$

puis

$$\int P(x)\cos(\alpha x)\,\mathrm{d}x = \frac{1}{\alpha}P(x)\sin\alpha x + \frac{1}{\alpha^2}P'(x)\cos\alpha x - \frac{1}{\alpha^2}\int P''(x)\cos(\alpha x)\,\mathrm{d}x = ...$$

(2) En vertu de l'étude ci-dessus, sait

$$\int P(x)\cos(\alpha x)\,\mathrm{d}x = A(x)\cos(\alpha x) + B(x)\sin(\alpha x) + C^{te}$$

avec A, B fonctions polynomiales de degrés inférieurs à n.
On peut alors chercher A et B par coefficients inconnues de sorte que

$$(A(x)\cos\alpha x + B(x)\sin\alpha x)' = P(x)\cos\alpha x$$

La problématique est bien entendu identique pour la détermination de $\int P(x)\sin(\alpha x)\,\mathrm{d}x$.

Remarque Si $\deg P \leqslant 2$ il est efficace d'utiliser (1)
Si $\deg P \geqslant 3$ il est efficace d'utiliser (2)

Exemple Déterminons $\int x^2 \sin x \cos x \,\mathrm{d}x$ sur $\mathbb{R}$.

$$\int x^2 \sin x \cos x \,\mathrm{d}x = \frac{1}{2}\int x^2 \sin 2x = -\frac{1}{4}x^2 \cos 2x + \frac{1}{2}\int x \cos 2x$$

puis

$$\int x^2 \sin x \cos x \,\mathrm{d}x = -\frac{1}{4}x^2 \cos 2x + \frac{1}{4}x \sin 2x - \frac{1}{4}\int \sin 2x$$

et enfin

$$\int x^2 \sin x \cos x \,\mathrm{d}x = (-\frac{1}{4}x^2 + \frac{1}{8}) \cos 2x + \frac{1}{4}x \sin 2x + C^{te}$$

Exemple Déterminons $\int (x^3 + x^2 + x + 1) \sin x \,\mathrm{d}x$ sur $\mathbb{R}$.

$$\int (x^3 + x^2 + x + 1) \sin x \,\mathrm{d}x = (ax^3 + bx^2 + cx + d) \cos x + (\alpha x^3 + \beta x^2 + \gamma x + \delta) \sin x + C^{te}$$

Après calculs, et un peu de courage, on obtient

$$a = -1, \alpha = 0, b = -1, \beta = 3, c = 5, \gamma = 2, d = 1, \delta = -5$$

Remarque En procédant à une linéarisation, la méthode qui précède permet le calcul de $\int P(x) \cos^n x \,\mathrm{d}x$

Remarque Les méthodes qui précèdent permettent aussi de déterminer $\int P(x)\mathrm{ch}(\alpha x)\,\mathrm{d}x$ et $\int P(x)\mathrm{sh}(\alpha x)\,\mathrm{d}x$.

Exemple Déterminons $\int (x+1)\mathrm{sh}x \,\mathrm{d}x$ sur $\mathbb{R}$.

$$\int (x+1)\mathrm{sh}x \,\mathrm{d}x = (x+1)\mathrm{ch}x - \int \mathrm{ch}x = (x+1)\mathrm{ch}x - \mathrm{sh}x$$

19.5.4 Intégration par parties

Théorème

Soient $u, v : I \to \mathbb{R}$ de classe $\mathcal{C}^1$.
Pour tout $a, b \in I$, on a

$$\int_a^b u'v = [uv]_a^b - \int_a^b uv'$$

dém. :
La fonction uv est primitive de la fonction continue $(uv)' = u'v + uv'$.
On en déduit $\displaystyle\int_a^b u'v + uv' = [uv]_a^b$ puis par linéarité la relation énoncée.
□

Remarque L'intégration par parties est utile pour calculer directement des intégrales ou pour former une relation sur les termes d'une suites d'intégrales.

Exemple Par intégration par parties

$$\int_0^1 t\arctan t\,\mathrm{d}t = \left[\frac{t^2\arctan t}{2}\right]_0^1 - \frac{1}{2}\int_0^1 \frac{t^2}{t^2+1}\,\mathrm{d}t = \frac{\pi}{8} - \frac{1}{2}\int_0^1 \frac{(t^2+1)-1}{t^2+1}\,\mathrm{d}t = \frac{\pi-2}{4}$$

Exemple Pour $p, q \in \mathbb{N}$, posons

$$I_{p,q} = \int_0^1 t^p(1-t)^q\,\mathrm{d}t$$

Pour $q > 0$, on a par intégration par parties

$$I_{p,q} = \int_0^1 t^p(1-t)^q\,\mathrm{d}t = \left[\frac{t^{p+1}(1-t)^q}{p+1}\right]_0^1 + \frac{q}{p+1}\int_0^1 t^{p+1}(1-t)^{q-1}\,\mathrm{d}t = \frac{q}{p+1}I_{p+1,q-1}$$

Ainsi

$$I_{p,q} = \frac{q}{p+1}I_{p+1,q-1} = \frac{q(q-1)}{(p+1)(p+2)}I_{p+2,q-2} = \ldots = \frac{q(q-1)\ldots 1}{(p+1)(p+2)\ldots(p+q)}I_{p+q,0}$$

puis

$$I_{p,q} = \frac{q(q-1)\ldots 1}{(p+1)(p+2)\ldots(p+q+1)} = \frac{q!p!}{(p+q+1)!}$$

Exemple On étudie

$$I_n = \int_0^{\pi/2} \sin^n(t)\,\mathrm{d}t$$

Pour $n \geqslant 2$,

$$I_n = \int_0^{\pi/2} \sin t.\sin^{n-1}(t)\,\mathrm{d}t$$

Par intégration par parties,

$$I_n = \left[-\cos t.\sin^{n-1} t\right]_0^{\pi/2} + (n-1)\int_0^{\pi/2} \cos^2(t)\sin^{n-2}(t)\,\mathrm{d}t$$

Or

$$\left[-\cos t.\sin^{n-1} t\right]_0^{\pi/2} = 0$$

et

$$\int_0^{\pi/2} \cos^2(t)\sin^{n-2}(t)\,\mathrm{d}t = \int_0^{\pi/2} (1-\sin^2(t))\sin^{n-2}(t)\,\mathrm{d}t = I_n - I_{n-2}$$

donc $I_n = (n-1)(I_n - I_{n-2})$ puis enfin

$$I_n = \frac{n-1}{n} I_{n-2}$$

Par cette relation de récurrence, il est possible d'exprimer I_n en fonction de I_1 ou de I_0 selon la parité de n.
Cas n impair : $n = 2p+1$.

$$I_{2p+1} = \frac{2p}{2p+1} I_{2p-1} = \frac{2p}{2p+1}\frac{2p-2}{2p-1} I_{2p-3}, \ldots I_{2p+1} = \frac{2p}{2p+1}\frac{2p-2}{2p-1}\cdots\frac{2}{3} I_1$$

Or

$$2p(2p-2)\ldots 2 = 2^p p!$$

et

$$(2p+1)(2p-1)\ldots 3 = \frac{(2p+1)!}{2^p p!}$$

Enfin $I_1 = \displaystyle\int_0^{\pi/2} \sin(t)\,\mathrm{d}t = 1$ donc

$$I_{2p+1} = \frac{(2^p p!)^2}{(2p+1)!}$$

Cas n pair : $n = 2p$
De façon analogue

$$I_{2p} = \frac{(2p)!}{(2^p p)!}\frac{\pi}{2}$$

Exemple Soit $f : [a,b] \to \mathbb{R}$ de classe $\mathcal{C}^1$. Montrons

$$\lim_{n\to+\infty} \int_a^b f(t)\sin(nt)\,\mathrm{d}t = 0$$

Par intégration par parties

$$\int_a^b f(t)\sin(nt)\,\mathrm{d}t = \left[-\frac{f(t)\cos(nt)}{n}\right]_a^b + \frac{1}{n}\int_a^b f'(t)\cos(nt)\,\mathrm{d}t$$

On en déduit

$$\left|\int_a^b f(t)\sin(nt)\,\mathrm{d}t\right| \leqslant \frac{|f(a)|+|f(b)|}{n} + \frac{1}{n}\int_a^b |f'(t)|\ \mathrm{d}t \to 0$$

puis

$$\int_a^b f(t)\sin(nt)\,\mathrm{d}t \to 0$$

Le même principe peut être repris pour obtenir

$$\int_a^b f(t)\cos(nt)\,\mathrm{d}t \to 0$$

19.6 Changement de variables

Soient I et J des intervalles non singuliers

19.6.1 Idée

Soient $u : I \to J$ et $F : J \to \mathbb{R}$ des fonctions dérivables.
La fonction $F \circ u$ est alors dérivable et $(F \circ u)' = u' \times F' \circ u$.
Plus légèrement, cette relation est notée $(F(u))' = u'F'(u)$ et ainsi on peut écrire

$$\int u'F'(u) = F(u) + C^{te}$$

Exemple Pour u fonction dérivable convenable, on a
$\displaystyle\int u'u^n = \frac{1}{n+1}u^{n+1} + C^{te}$ pour $n \in \mathbb{N}$;
$\displaystyle\int \frac{u'}{u^n} = -\frac{1}{n-1}\frac{1}{u^{n-1}} + C^{te}$ pour $n \in \mathbb{N}$, $n \geqslant 2$;
$\displaystyle\int u'u^\alpha \,\mathrm{d}u = \frac{1}{\alpha+1}u^{\alpha+1} + C^{te}$ pour $\alpha \in \mathbb{R}$
et en particulier $\displaystyle\int u'\sqrt{u}\,\mathrm{d}u = \frac{2}{3}u^{3/2} + C^{te}$ et $\displaystyle\int \frac{u'}{\sqrt{u}} = 2\sqrt{u} + C^{te}$;
$\displaystyle\int \frac{u'}{u} = \ln|u| + C^{te}$;
$\displaystyle\int u'\mathrm{e}^u = \mathrm{e}^u + C^{te}$;
$\displaystyle\int u' \sin u = -\cos u + C^{te}$;
$\displaystyle\int u' \cos u = \sin u + C^{te}$;
$\displaystyle\int u'\mathrm{sh}u = \mathrm{ch}u + C^{te}$;
$\displaystyle\int u'\mathrm{ch}u = \mathrm{sh}u + C^{te}$ et
$\displaystyle\int \frac{u'}{1+u^2} = \arctan u + C^{te}$.

Exemple $\displaystyle\int \frac{\ln t}{t}\,\mathrm{d}t = \int u'u = \frac{1}{2}(\ln t)^2 + C^{te}$.

Exemple $\displaystyle\int \frac{\mathrm{d}t}{t\ln t} = \int \frac{u'}{u} = \ln|\ln t| + C^{te}$.

Exemple $\displaystyle\int \tan t\,\mathrm{d}t = \int \frac{\sin t}{\cos t}\,\mathrm{d}t = \int -\frac{u'}{u} = -\ln|\cos t| + C^{te}$.

Exemple $\displaystyle\int \frac{t}{1+t^2}\,\mathrm{d}t = \int \frac{u'}{u} = \frac{1}{2}\ln(1+t^2) + C^{te}.$

Exemple $\displaystyle\int \frac{t}{1+t^4}\,\mathrm{d}t = \int \frac{1}{2}\frac{u'}{1+u^2} = \frac{1}{2}\arctan t^2 + C^{te}.$

Exemple $\displaystyle\int \frac{t}{\sqrt{1+t^2}}\,\mathrm{d}t = \int \frac{1}{2}\frac{u'}{\sqrt{u}} = \sqrt{1+t^2} + C^{te}.$

Exemple $\displaystyle\int \cos^3 t\,\mathrm{d}t = \int \cos t(1-\sin^2 t)\,\mathrm{d}t = \int u'(1-u^2) = \sin t - \frac{1}{3}\sin^3 t + C^{te}$

19.6.2 Détermination de primitives par changement de variables

Soient $u : I \to J$ dérivable et $f : J \to \mathbb{R}$ possédant une primitive F.
On a
$$\int u'(t)f(u(t))\,\mathrm{d}t = F(u(t)) + C^{te}$$
Pour exploiter cette formule, on écrit $x = u(t)$, $\mathrm{d}x = u'(t)\,\mathrm{d}t$
et ainsi
$$\int u'(t)f(u(t))\,\mathrm{d}t = \int f(x)\,\mathrm{d}x = F(x) + C^{te} = F(u(t)) + C^{te}$$
Lors de cette manipulation, on dit qu'on a réalisé le changement de variable défini par la relation $x = u(t)$.

Exemple Calculons
$$\int \cos t \sin^n t\,\mathrm{d}t$$
Procédons au changement de variable $x = \sin t$ pour lequel $\mathrm{d}x = \cos t\,\mathrm{d}t$.
On obtient
$$\int \cos t \sin^n t\,\mathrm{d}t \int x^n\,\mathrm{d}x = \frac{1}{n+1}x^{n+1} + C^{te} = \frac{1}{n+1}\sin^{n+1} t + C^{te}$$

Exemple Calculons
$$\int \frac{\mathrm{d}t}{\mathrm{ch}\,t} = \int \frac{2\,\mathrm{d}t}{\mathrm{e}^t + \mathrm{e}^{-t}}$$
Procédons au changement de variable $x = \mathrm{e}^t$ pour lequel $\mathrm{d}x = \mathrm{e}^t\,\mathrm{d}t = x\,\mathrm{d}t$
On obtient
$$\int \frac{\mathrm{d}t}{\mathrm{ch}\,t} = \int \frac{2\,\mathrm{d}x}{x(x+1/x)} = \arctan \mathrm{e}^t + C^{te}$$
Ici le changement de variable a révélé une forme $\dfrac{u'}{1+u^2}$ qu'on n'avait pas forcément entrevue...

Exemple Calculons
$$\int \frac{\mathrm{d}t}{\sqrt{t}+t}$$
Procédons au changement de variable $x=\sqrt{t}$ pour lequel $t=x^2$ et $\mathrm{d}t=2x\,\mathrm{d}x$.
$$\int \frac{\mathrm{d}t}{\sqrt{t}+t}=\int \frac{2x\,\mathrm{d}x}{x+x^2}=\ln(1+x)+C^{te}=2\ln(1+\sqrt{t})+C^{te}$$

Exemple Calculons
$$\int \frac{\mathrm{d}t}{\mathrm{e}^t+1}$$
Procédons au changement de car $x=\mathrm{e}^t$ pour lequel $\mathrm{d}x=\mathrm{e}^t\,\mathrm{d}t=x\,\mathrm{d}t$
$$\int \frac{\mathrm{d}t}{\mathrm{e}^t+1}\underset{x=\mathrm{e}^t}{=}\int \frac{\mathrm{d}x}{x(x+1)}=\int \frac{1}{x}-\frac{1}{x+1}\,\mathrm{d}x=\ln\frac{x}{x+1}+C^{te}=\ln\frac{\mathrm{e}^t}{1+\mathrm{e}^t}+C^{te}$$

19.6.3 Intégration par changement de variables

Théorème

Soient $u:I\to J$ de classe $\mathcal{C}^1$ et $f:J\to\mathbb{R}$ continue.
Pour tout $a,b\in I$,
$$\int_{t=a}^{b} f(u(t))u'(t)\,\mathrm{d}t=\int_{x=u(a)}^{u(b)} f(x)\,\mathrm{d}x$$

dém. :
Soit F une primitive de la fonction continue f.
La fonction $F\circ u$ est primitive de $u'\times f'\circ u$ et par suite
$$\int_a^b f(u(t))u'(t)\,\mathrm{d}t=[F(u(t))]_a^b=[F(x)]_{u(a)}^{u(b)}=\int_{u(a)}^{u(b)} f(x)\,\mathrm{d}x$$

□

Remarque Pour exploiter cette formule, on écrit :
$x=u(t)$, $\mathrm{d}x=u'(t)\,\mathrm{d}t$
pour $t=a$, $x=u(a)$,
pour $t=b$, $x=u(b)$.
Ceci permet de transformer formellement une intégrale en l'autre.
On dit alors qu'on a réalisé le changement de variable défini par la relation $x=u(t)$.

Exemple Calculons
$$\int_1^{\mathrm{e}} \frac{\mathrm{d}t}{t+t\ln t}$$
On procède au changement de variable $x=\ln t$ pour lequel $\mathrm{d}x=\mathrm{d}t/t$.

Pour $t = 1$, $x = 0$ et pour $t = \mathrm{e}$, $x = 1$.
On obtient
$$\int_{t=1}^{\mathrm{e}} \frac{\mathrm{d}t}{t + t\ln t} = \int_{x=0}^{1} \frac{\mathrm{d}x}{x+1} = \ln 2$$

Exemple Calculons
$$\int_1^3 \frac{\sqrt{t}\,\mathrm{d}t}{t+t^2}$$
On procède au changement de variable $x = \sqrt{t}$ pour lequel $t = x^2$ et $\mathrm{d}t = 2x\,\mathrm{d}x$.
Pour $t = 1$, $x = 1$ et pour $t = 3$, $x = \sqrt{3}$.
On obtient
$$\int_{t=1}^{3} \frac{\sqrt{t}\,\mathrm{d}t}{t+t^2} = \int_{x=1}^{\sqrt{3}} \frac{2x^2\,\mathrm{d}x}{x^2+x^4} = [2\arctan x]_1^{\sqrt{3}} = \frac{\pi}{6}$$

Exemple Calculons
$$\int_1^2 \frac{\mathrm{d}t}{t+t^3}$$
On procède au changement de variable $x = 1/t$ pour lequel $\mathrm{d}x = -\,\mathrm{d}t/t^2$.
Pour $t = 1$, $x = 1$ et pour $t = 2$, $x = 1/2$.
$$\int_{t=1}^{2} \frac{\mathrm{d}t}{t+t^3} = -\int_{t=1}^{1/2} \frac{x\,\mathrm{d}x}{x^2+1} = \int_{1/2}^{1} \frac{x\,\mathrm{d}x}{x^2+1} = \left[\frac{1}{2}\ln(1+x^2)\right]_{1/2}^{1} = \frac{1}{2}\ln\frac{8}{5}$$

Exemple Calculons
$$\int_0^1 \sqrt{1-t^2}\,\mathrm{d}t$$
On procède au changement de variable $t = \sin x$ pour lequel $\mathrm{d}t = \cos x\,\mathrm{d}x$.
Pour $x = 0$, $t = 0$ et pour $x = \pi/2$, $t = 1$.
Après linéarisation, on obtient
$$\int_{t=0}^{1} \sqrt{1-t^2}\,\mathrm{d}t = \int_{x=0}^{1} \cos^2 x\,\mathrm{d}x = \frac{\pi}{4}$$
Dans cet exemple, le changement de variable est réalisé dans le sens inverse des exemples précédents, c'est possible car l'égalité du théorème de changement de variable peut être utilisée dans les deux sens !

19.6.4 Changement de variables affines

19.6.4.1 Principe

Les changements de variables affines sont ceux de la forme $t = ax + b$ avec $a, b \in \mathbb{R}$ et $a \neq 0$.

Exemple Déterminons $\displaystyle\int \frac{\mathrm{d}t}{t^2+a^2}$ pour $a \neq 0$.

Par un changement de variable affine, nous allons transformer cette étude en $\displaystyle\int \frac{\mathrm{d}t}{t^2+1}$.
Posons le changement de variable $t = ax$, on obtient

$$\int \frac{\mathrm{d}t}{t^2+a^2} = \int \frac{a\,\mathrm{d}x}{a^2x^2+a^2} = \frac{1}{a}\int \frac{\mathrm{d}x}{x^2+1} = \frac{1}{a}\arctan\frac{t}{a} + C^{te}$$

Exemple Soient $a, b \in \mathbb{R}$ tels que $a < b$, calculons $J_{p,q} = \displaystyle\int_a^b (t-a)^p(b-t)^q\,\mathrm{d}t$ pour $p, q \in \mathbb{N}$.
Cette étude ressemble à celle, particulière, $\displaystyle\int_0^1 t^p(1-t)^q\,\mathrm{d}t$ déjà menée.
Par un changement de variable affine transformant le segment $[a, b]$ en $[0, 1]$, nous allons transformer la première en la seconde.
Posons le changement de variable $t = a + x(b-a)$, on obtient

$$J_{p,q} = \int_a^b (t-a)^p(b-t)^q\,\mathrm{d}t = \int_0^1 (b-a)^{p+q+1}t^p(1-t)^q\,\mathrm{d}t = (b-a)^{p+q+1}\frac{p!q!}{(p+q+1)!}$$

19.6.4.2 Propriétés géométriques

Par changement de variable affines on peut facilement établir quelques propriétés géométriques des intégrales...

Proposition

Soient $\tau \in \mathbb{R}$ et $f : \mathbb{R} \to \mathbb{R}$ continue.
Pour $a, b \in \mathbb{R}$, on a

$$\int_a^b f(t+\tau)\,\mathrm{d}t = \int_{a+\tau}^{b+\tau} f(t)\,\mathrm{d}t$$

dém. :
Par le changement de variable $x = t+\tau$ pour lequel $\mathrm{d}x = \mathrm{d}t$, on obtient $\displaystyle\int_a^b f(t+\tau)\,\mathrm{d}t = \int_{a+\tau}^{b+\tau} f(x)\,\mathrm{d}x$.
□

Proposition

Soient $a > 0$ et $f : [-a, a] \to \mathbb{R}$ continue.
Si f est paire alors

$$\int_{-a}^a f(t)\,\mathrm{d}t = 2\int_0^a f(t)\,\mathrm{d}t$$

Si f est impaire alors

$$\int_{-a}^a f(t)\,\mathrm{d}t = 0$$

dém. :
Par le changement de variable $t = -x$ pour lequel $\mathrm{d}t = -\mathrm{d}x$, on obtient

$$\int_{-a}^0 f(t)\,\mathrm{d}t = \int_0^a f(-x)\,\mathrm{d}x = \int_0^a f(-t)\,\mathrm{d}t$$

Si f est paire $\int_{-a}^{0} f(t)\,\mathrm{d}t = \int_{0}^{a} f(t)\,\mathrm{d}t$ puis $\int_{-a}^{a} f(t)\,\mathrm{d}t = 2\int_{0}^{a} f(t)\,\mathrm{d}t$

Si f est impaire $\int_{-a}^{0} f(t)\,\mathrm{d}t = -\int_{0}^{a} f(t)\,\mathrm{d}t$ puis $\int_{-a}^{a} f(t)\,\mathrm{d}t = 0$

□

Proposition

> Soit $f : \mathbb{R} \to \mathbb{R}$ continue et T périodique.
> Pour tout $a \in \mathbb{R}$, on a
> $$\int_{a}^{a+T} f(t)\,\mathrm{d}t = \int_{0}^{T} f(t)\,\mathrm{d}t$$

dém. :
Par la relation de Chasles

$$\int_{a}^{a+T} f(t)\,\mathrm{d}t = \int_{a}^{0} f(t)\,\mathrm{d}t + \int_{0}^{T} f(t)\,\mathrm{d}t + \int_{T}^{a+T} f(t)\,\mathrm{d}t$$

Par le changement de variable $t = x + T$, on obtient

$$\int_{T}^{a+T} f(t)\,\mathrm{d}t = \int_{0}^{a} f(x+T)\,\mathrm{d}x$$

Or

$$\int_{0}^{a} f(x+T)\,\mathrm{d}x = \int_{0}^{a} f(x)\,\mathrm{d}x$$

car f est T périodique et la relation initiale donne alors

$$\int_{a}^{a+T} f(t)\,\mathrm{d}t = \int_{0}^{T} f(t)\,\mathrm{d}t$$

□

Remarque On peut étendre les résultats qui précèdent aux fonctions continues par morceaux en procédant à un découpage aux points de discontinuité des intégrales de celles-ci.

Exemple Soient $f : \mathbb{R} \to \mathbb{R}$ une fonction continue, périodique et impaire et F une primitive de f.
Montrons que F est périodique.
Pour tout $x \in \mathbb{R}$, on a

$$\int_{0}^{x+T} f(t)\,\mathrm{d}t = \int_{0}^{x} f(t)\,\mathrm{d}t + \int_{x}^{x+T} f(t)\,\mathrm{d}t$$

donc

$$F(x+T) = F(x) + \int_{x}^{x+T} f(t)\,\mathrm{d}t$$

Or par périodicité de f,

$$\int_{x}^{x+T} f(t)\,\mathrm{d}t = \int_{0}^{T} f(t)\,\mathrm{d}t = \int_{-T/2}^{T/2} f(t)\,\mathrm{d}t = 0$$

car f est impaire.
Ainsi F est T-périodique.

19.7 Méthodes d'approximation d'intégrales

Soient $a < b \in \mathbb{R}$ et $f : [a, b] \to \mathbb{R}$ continue. Il n'est pas toujours possible d'expliciter $\int_a^b f(t)\,\mathrm{d}t$; notamment parce qu'il n'est pas toujours possible d'exprimer une primitive de f à l'aide des fonctions usuelles.
A défaut d'un calcul exact, on se propose dans cette partie de présenter quelques techniques usuelle permettant un calcul numérique approché de $\int_a^b f(t)\,\mathrm{d}t$.

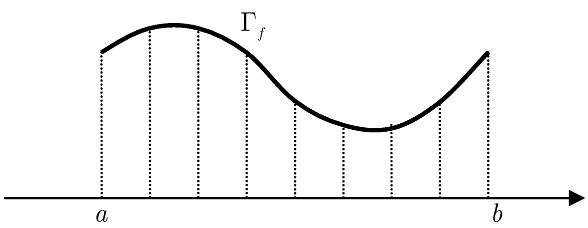
19.7.1 Les sommes de Riemann

Pour $n \in \mathbb{N}^\star$, considérons la subdivision $\sigma = (a_0, \dots, a_n)$ à pas constant déterminée par

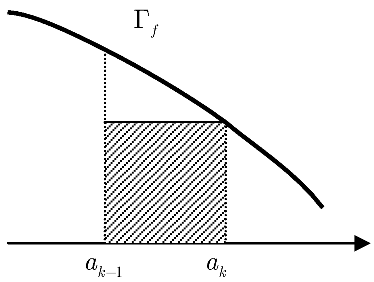
$$\forall 0 \leqslant k \leqslant n, a_k = a + k\frac{b-a}{n}$$

En vertu de la relation de Chasles, on a

$$\int_a^b f(t)\,\mathrm{d}t = \sum_{k=1}^{n} \int_{a_{k-1}}^{a_k} f(t)\,\mathrm{d}t$$



On approche chaque intégrale $\int_{a_{k-1}}^{a_k} f(t)\,\mathrm{d}t$ par l'aire du rectangle à droite suivant



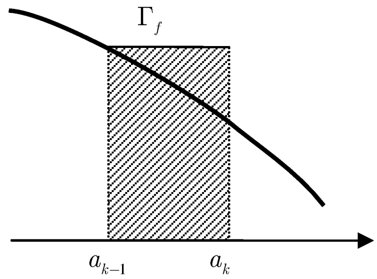
Ainsi, on approche $\int_{a_{k-1}}^{a_k} f(t)\,\mathrm{d}t$ par

$$\frac{1}{2}(a_k - a_{k-1})f(a_k) = \frac{b-a}{n}f(a_k)$$

En sommant ces approximations, on approche $\int_a^b f(t)\,\mathrm{d}t$ par la quantité

$$R_n^+ = \frac{b-a}{n}\sum_{k=1}^{n} f(a_k) = \frac{b-a}{n}\left(f(a_1) + \cdots + f(a_n)\right).$$

On peut aussi approcher l'intégrale $\displaystyle\int_{a_{k-1}}^{a_k} f(t)\,\mathrm{d}t$ par l'aire du rectangle à gauche suivant

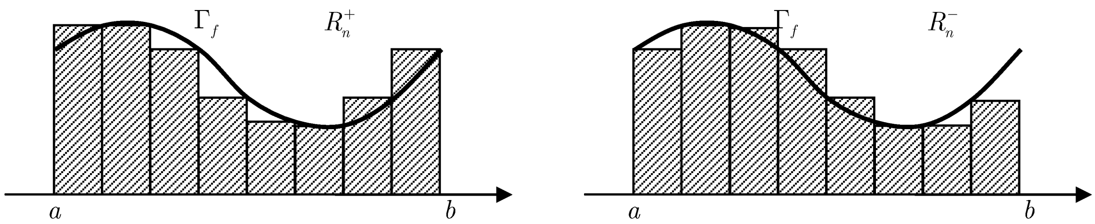


Ainsi on approche $\displaystyle\int_{a_{k-1}}^{a_k} f(t)\,\mathrm{d}t$ par

$$\frac{1}{2}(a_k - a_{k-1})f(a_{k-1}) = \frac{b-a}{n}f(a_{k-1})$$

puis on approche $\displaystyle\int_a^b f(t)\,\mathrm{d}t$ par la quantité

$$R_n^- = \frac{b-a}{n}\sum_{k=0}^{n-1} f(a_k) = \frac{b-a}{n}\left(f(a_0) + \cdots + f(a_{n-1})\right)$$



Définition

Les quantités R_n^+ et R_n^- sont appelées somme de Riemann de la fonction f associées à la méthodes des rectangles à droite et à gauche.

Théorème

Si $f : [a,b] \to \mathbb{R}$ est de classe $\mathcal{C}^1$ alors

$$\left|\int_a^b f(t)\,\mathrm{d}t - R_n^{+/-}\right| \leqslant \frac{(b-a)}{2n}M$$

avec $M = \max\limits_{[a,b]} |f'|$.
Ainsi les suites (R_n^+) et (R_n^-) convergent vers l'intégrale de f sur $[a,b]$.

dém. :

Par construction de la méthode des rectangles

$$\left|\int_a^b f(t)\,\mathrm{d}t - R_n^+\right| = \left|\sum_{k=1}^n \left[\int_{a_{k-1}}^{a_k} f(t)\,\mathrm{d}t - \frac{(b-a)}{n} f(a_k)\right]\right|$$
$$= \left|\sum_{k=1}^n \int_{a_{k-1}}^{a_k} \left[f(t) - f(a_k)\right]\,\mathrm{d}t\right| \leqslant \sum_{k=1}^n \int_{a_{k-1}}^{a_k} \left|f(t) - f(a_k)\right|\,\mathrm{d}t$$

Or en vertu de l'inégalité des accroissements finis, pour tout $t \in [a_{k-1}, a_k]$, $|f(t) - f(a_k)| \leqslant M(a_k - t)$ et donc

$$\left|\int_a^b f(t)\,\mathrm{d}t - R_n^+\right| \leqslant \sum_{k=1}^n \int_{a_{k-1}}^{a_k} M(a_k - t)\,\mathrm{d}t = \sum_{k=1}^n \frac{M(b-a)^2}{2n^2} = \frac{(b-a)^2}{2n} M$$

L'étude est identique pour R_n^-.
□

Remarque En pratique, la convergence en $O(1/n)$ est numériquement trop lente.

Remarque On peut aussi obtenir la convergence des suites (R_n^+) et (R_n^-) avec une hypothèse moindre.

Théorème

Si $f : [a, b] \to \mathbb{R}$ est continue alors $R_n^{+/-} \to \displaystyle\int_a^b f(t)\,\mathrm{d}t$.

dém. :
En reprenant les calculs ci-dessus

$$\left|\int_a^b f(t)\,\mathrm{d}t - R_n^+\right| \leqslant \sum_{k=1}^n \int_{a_{k-1}}^{a_k} \left|f(t) - f(a_k)\right|\,\mathrm{d}t$$

Soit $\varepsilon > 0$. Puisque la fonction f est continue sur le segment $[a, b]$, elle y est uniformément continue en vertu du théorème de Heine. Par suite il existe $\alpha > 0$ vérifiant

$$\forall x, y \in [a, b]\,, |y - x| \leqslant \alpha \Rightarrow |f(y) - f(x)| \leqslant \varepsilon$$

Pour n assez grand, $(b - a)/n \leqslant \alpha$ et alors pour tout $t \in [a_{k-1}, a_k]$, $|t - a_k| \leqslant \dfrac{b-a}{n} \leqslant \alpha$ donc $|f(t) - f(a_k)| \leqslant \varepsilon$.
On a alors

$$\sum_{k=1}^n \int_{a_{k-1}}^{a_k} \left|f(t) - f(a_k)\right|\,\mathrm{d}t \leqslant \sum_{k=1}^n \int_{a_{k-1}}^{a_k} \varepsilon\,\mathrm{d}t = \int_a^b \varepsilon\,\mathrm{d}t = \varepsilon(b-a)$$

Ainsi, pour tout $\varepsilon > 0$, à partir d'un certain rang

$$\left|\int_a^b f(t)\,\mathrm{d}t - R_n^+\right| \leqslant \varepsilon(b-a)$$

Finalement $R_n^+ \to \int_a^b f(t)\,\mathrm{d}t$ et de façon semblable, on obtient $R_n^- \to \int_a^b f(t)\,\mathrm{d}t$.
□

Corollaire

Si $f : [0,1] \to \mathbb{R}$ est continue alors

$$\frac{1}{n}\sum_{k=1}^{n} f\left(\frac{k}{n}\right) \to \int_0^1 f(t)\,dt \text{ et } \frac{1}{n}\sum_{k=0}^{n-1} f\left(\frac{k}{n}\right) \to \int_0^1 f(t)\,\mathrm{d}t$$

Exemple Etudions

$$\lim_{n\to\infty}\sum_{k=1}^{n}\frac{1}{n+k}$$

On peut écrire

$$\sum_{k=1}^{n}\frac{1}{n+k} = \frac{1}{n}\sum_{k=1}^{n}\frac{1}{1+\frac{k}{n}} = \frac{1}{n}\sum_{k=1}^{n} f\left(\frac{k}{n}\right)$$

avec $f : x \mapsto \dfrac{1}{1+x}$ continue sur $[0,1]$.
On en déduit

$$\sum_{k=1}^{n}\frac{1}{n+k} \to \int_0^1 \frac{\mathrm{d}t}{1+t} = \ln 2$$

Exemple Etudions

$$\lim_{n\to\infty}\sum_{k=n+1}^{2n}\frac{n}{k^2}$$

On peut écrire

$$\sum_{k=n+1}^{2n}\frac{n}{k^2} = \frac{1}{n}\sum_{k=1}^{n}\frac{n^2}{(n+k)^2} = \frac{1}{n}\sum_{k=1}^{n}\frac{1}{(1+k/n)^2} = \frac{1}{n}\sum_{k=1}^{n} f\left(\frac{k}{n}\right)$$

avec $f : x \mapsto \dfrac{1}{(1+x)^2}$ continue sur $[0,1]$.
On en déduit

$$\sum_{k=n+1}^{2n}\frac{n}{k^2} \to \int_0^1 \frac{\mathrm{d}t}{(1+t)^2} = \frac{1}{2}$$

Exemple Etudions

$$\lim_{n\to\infty}\frac{1}{n}\sqrt[n]{(n+1)(n+2)\dots(2n)}$$

On peut écrire

$$\ln\left(\frac{1}{n}\sqrt[n]{(n+1)(n+2)\dots(2n)}\right) = \ln\left(\sqrt[n]{\left(1+\frac{1}{n}\right)\left(1+\frac{2}{n}\right)\cdots\left(1+\frac{n}{n}\right)}\right)$$

donc

$$\ln\left(\frac{1}{n}\sqrt[n]{(n+1)(n+2)\dots(2n)}\right) = \frac{1}{n}\sum_{k=1}^{n}\ln\left(1+\frac{k}{n}\right) = \frac{1}{n}\sum_{k=1}^{n} f\left(\frac{k}{n}\right)$$

avec $f : x \mapsto \ln(1+x)$ continue sur $[0,1]$.
On en déduit

$$\ln\left(\frac{1}{n}\sqrt[n]{(n+1)(n+2)\dots(2n)}\right) \to \int_0^1 \ln(1+t)\,\mathrm{d}t = [(1+t)\ln(1+t)-t]_0^1 = 2\ln 2 - 1$$

puis

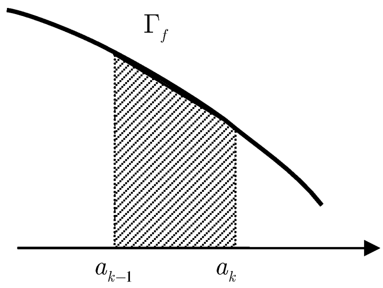
$$\frac{1}{n}\sqrt[n]{(n+1)(n+2)\dots(2n)} \to \frac{4}{\mathrm{e}}$$

19.7.2 Méthode des trapèzes

Comme ci-dessus, en introduisant $\sigma = (a_0,\dots,a_n)$ subdivision de $[a,b]$ à pas constante, on découpe l'intégrale $\int_a^b f(t)\,\mathrm{d}t$ pour écrire

$$\int_a^b f(t)\,\mathrm{d}t = \sum_{k=1}^{n}\int_{a_{k-1}}^{a_k} f(t)\,\mathrm{d}t$$

On approche maintenant chaque intégrale $\int_{a_{k-1}}^{a_k} f(t)\,\mathrm{d}t$ par l'aire du trapèze suivant

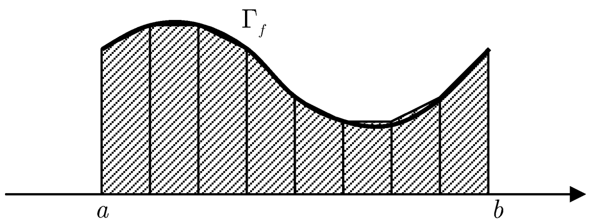


Ainsi, on approche $\int_{a_{k-1}}^{a_k} f(t)\,\mathrm{d}t$ par

$$\frac{1}{2}(a_k - a_{k-1})(f(a_{k-1}) + f(a_k)) = \frac{b-a}{2n}(f(a_{k-1}) + f(a_k))$$

En sommant ces approximations, on approche $\int_a^b f(t)\,\mathrm{d}t$ par la quantité

$$T_n = \sum_{k=1}^{n}\frac{b-a}{2n}(f(a_{k-1}) + f(a_k)) = \frac{b-a}{n}\left(\frac{f(a_0)}{2} + f(a_1) + \dots + f(a_{n-1}) + \frac{f(a_n)}{2}\right).$$



Théorème

Si $f:[a,b]\to\mathbb{R}$ est de classe $\mathcal{C}^2$ alors

$$\left|\int_a^b f(t)\,\mathrm{d}t - T_n\right| \leqslant \frac{M(b-a)}{12n^2}$$

avec $M=\max\limits_{[a,b]}|f''|$.

dém. :

Par construction de la méthode des trapèzes

$$\begin{aligned}\left|\int_a^b f(t)\,\mathrm{d}t - T_n\right| &= \left|\sum_{k=1}^n \int_{a_{k-1}}^{a_k} f(t)\,\mathrm{d}t - \frac{b-a}{2n}\sum_{k=1}^n \left(f(a_{k-1})+f(a_k)\right)\right| \\ &\leqslant \sum_{k=1}^n \left|\int_{a_{k-1}}^{a_k} f(t)\,\mathrm{d}t - \int_{a_{k-1}}^{a_k} \varphi_k(t)\,\mathrm{d}t\right|\end{aligned}$$

en notant φ_k la fonction affine déterminée par $\varphi_k(a_{k-1})=f(a_{k-1})$ et $\varphi_k(a_k)=f(a_k)$.

Pour chaque $k\in\{1,\dots,n\}$,

$$\left|\int_{a_{k-1}}^{a_k} f(t)\,\mathrm{d}t - \int_{a_{k-1}}^{a_k} \varphi_k(t)\,\mathrm{d}t\right| = \left|\int_{a_{k-1}}^{a_k} f(t)-\varphi_k(t)\,\mathrm{d}t\right| \leqslant \int_{a_{k-1}}^{a_k} |f(t)-\varphi_k(t)|\,\mathrm{d}t$$

Pour poursuivre, montrons que pour chaque $t\in[a_{k-1},a_k]$, il existe $u\in[a_{k-1},a_k]$ vérifiant

$$f(t)-\varphi_k(t) = \frac{(t-a_{k-1})(t-a_k)}{2}f''(u)$$

Lorsque $t=a_{k-1}$ ou $t=a_k$, la propriété est immédiate car n'importe quel $u\in[a_{k-1},a_k]$ convient.

Pour $t\in\,]a_{k-1},a_k[$, considérons la fonction

$$\psi : x\mapsto f(x)-\varphi_k(x)-\frac{(x-a_{k-1})(x-a_k)}{2}K_t$$

définie à partir d'une constante K_t choisie de sorte que $\psi(t)=0$.

La fonction ψ est de classe $\mathcal{C}^2$ et s'annule en au moins trois points a_{k-1},a_k et t. Par application du théorème de Rolle, on peut affirmer que la dérivée seconde de ψ s'annule et donc il existe $u\in[a_{k-1},a_k]$ vérifiant $\psi''(u)=0$. Or $\psi''(x)=f''(x)-K_t$ car la dérivée seconde d'une fonction affine est nulle. Ainsi $\psi''(u)=0$ donne $K_t=f''(u)$ et puisque la constante K_t a été choisie de sorte que $\psi(t)=0$ on obtient

$$f(t)-\varphi_k(t) = \frac{(t-a_{k-1})(t-a_k)}{2}f''(u)$$

On en déduit

$$|f(t) - \varphi_k(t)| \leqslant \frac{(t-a_{k-1})(a_k-t)}{2} M$$

puis

$$\int_{a_{k-1}}^{a_k} |f(t) - \varphi_k(t)| \,\mathrm{d}t \leqslant \int_{a_{k-1}}^{a_k} \frac{(t-a_{k-1})(a_k-t)}{2} M \,\mathrm{d}t$$

Par une intégration par parties, on peut facilement calculer cette dernière intégrale et obtenir

$$\int_{a_{k-1}}^{a_k} |f(t) - \varphi_k(t)| \,\mathrm{d}t \leqslant \frac{(b-a)^3}{12n^3} M$$

On en déduit

$$\left| \int_a^b f(t) \,\mathrm{d}t - T_n \right| \leqslant \sum_{k=1}^n \frac{(b-a)^3}{12n^3} M = \frac{(b-a)^3}{12n^2} M$$

□

Remarque L'erreur d'approximation est en $O(1/n^2)$, c'est honorable mais on peut faire mieux avec la méthode qui suit.

19.7.3 Méthode de Simpson

La méthode de Simpson, qui n'est pas au programme, consiste à approcher l'intégrale de f sur les segments $[a_{k-1}, a_k]$ par l'intégrale de la fonction polynomiale de degré $\leqslant 2$ prenant les mêmes valeurs que f en $a_{k-1}, (a_{k-1}+a_k)/2$ et a_k.

Au terme des calculs, cela revient à approcher $\displaystyle\int_a^b f(t) \,\mathrm{d}t$ par la quantité

$$S_n = \frac{b-a}{6n} \sum_{k=1}^n \left[f(a_{k-1}) + 4f\left(\frac{a_{k-1}+a_k}{2}\right) + f(a_k) \right].$$

On montre que lorsque f est de classe $\mathcal{C}^4$, l'erreur est un $O\left(1/n^4\right)$.
La méthode de Simpson est la méthode numérique la plus utilisée pour calculer des intégrales.

19.8 Extension aux fonctions complexes

I désigne un intervalle non singulier de $\mathbb{R}$.
Les définitions qui suivent prolongent celles vue précédemment.

19.8.1 Construction de l'intégrale d'une fonction complexe

Définition

> On dit qu'une fonction $f : I \to \mathbb{C}$ est continue par morceaux si les fonctions réelles $\mathrm{Re}(f)$ et $\mathrm{Im}(f)$ le sont.
> On note $\mathcal{C}^0_{pm}(I, \mathbb{C})$ l'ensemble des fonctions continues par morceaux de I vers $\mathbb{C}$.

Exemple Les fonctions complexes continues sont continues par morceaux.

Définition

Soient $f : I \to \mathbb{C}$ continue par morceaux et $a, b \in I$.
On appelle intégrale de la fonction f de a à b le complexe

$$\int_a^b f(t)\,\mathrm{d}t = \int_a^b \mathrm{Re}(f)(t)\,\mathrm{d}t + i\int_a^b \mathrm{Im}(f)(t)\,\mathrm{d}t$$

Exemple $\displaystyle\int_0^{2\pi} \mathrm{e}^{it}\,\mathrm{d}t = \int_0^{2\pi} \cos t\,\mathrm{d}t + i\int_0^{2\pi} \sin t\,\mathrm{d}t = 0 + i.0 = 0.$

Exemple $\displaystyle\int_0^1 \frac{\mathrm{d}t}{t+i} = \int_0^1 \frac{t-i}{t^2+1}\,\mathrm{d}t = \int_0^1 \frac{t}{t^2+1}\,\mathrm{d}t + i\int_0^1 \frac{\mathrm{d}t}{t^2+1} = \frac{1}{2}\ln 2 - i\frac{\pi}{4}.$

Théorème

Soient $f, g : I \to \mathbb{C}$ continues par morceaux, $\lambda \in \mathbb{C}$.
Pour tout $a, b \in I$,

$$\int_a^b \lambda f = \lambda\int_a^b f,\ \int_a^b f + g = \int_a^b f + \int_a^b g \text{ et } \int_a^b \bar{f} = \overline{\int_a^b f}$$

dém. :
Par étude des parties des parties réelles et imaginaires.
□

Théorème

Soit $f : I \to \mathbb{C}$ continue par morceaux.
Pour tout $a, b, c \in I$,

$$\int_a^b f = \int_a^c f + \int_c^b f$$

dém. :
Par étude des parties des parties réelles et imaginaires.
□

(!)Les propriétés de croissance et de positivité de l'intégrale disparaissent dans le cadre des fonctions complexes. Par exemple, l'intégrale $\displaystyle\int_0^{2\pi} \mathrm{e}^{it}\,\mathrm{d}t$ est nulle alors la fonction $t \mapsto \mathrm{e}^{it}$ ne s'annule pas.

Théorème

Soient $a \leqslant b \in \mathbb{R}$ et $f : [a, b] \to \mathbb{C}$ continue par morceaux. On a

$$\left|\int_a^b f(t)\,\mathrm{d}t\right| \leqslant \int_a^b |f(t)|\,\mathrm{d}t$$

dém. :

Posons $J = \displaystyle\int_a^b f(t)\,\mathrm{d}t$. On peut écrire $J = r\mathrm{e}^{i\theta}$ avec $r = \left|\displaystyle\int_a^b f(t)\,\mathrm{d}t\right|$.

Considérons alors la fonction continue par morceaux $g : I \to \mathbb{C}$ définie par $g(t) = f(t)\mathrm{e}^{-i\theta}$.

On a

$$\int_a^b g(t)\,\mathrm{d}t = \mathrm{e}^{-i\theta}\int_a^b f(t)\,\mathrm{d}t = \mathrm{e}^{-i\theta}J = r \in \mathbb{R}^+$$

donc

$$r = \left|\mathrm{Re}\left(\int_a^b g(t)\,\mathrm{d}t\right)\right|$$

Or

$$\left|\mathrm{Re}\left(\int_a^b g(t)\,\mathrm{d}t\right)\right| = \left|\int_a^b \mathrm{Re}(g)(t)\,\mathrm{d}t\right| \leqslant \int_a^b |\mathrm{Re}(g)(t)|\,\mathrm{d}t \leqslant \int_a^b |g(t)|\,\mathrm{d}t = \int_a^b |f(t)|\,\mathrm{d}t$$

donc

$$r = \left|\int_a^b f(t)\,\mathrm{d}t\right| \leqslant \int_a^b |f(t)|\,\mathrm{d}t$$

□

19.8.2 Intégration et dérivation

Définition

On appelle primitive de $f : I \to \mathbb{C}$, s'il en existe, toute fonction $F : I \to \mathbb{C}$ dérivable vérifiant $F' = f$.

Remarque Si f possède une primitive F alors les autres primitives de f sont les fonctions de la forme $t \mapsto F(t) + C$ avec $C \in \mathbb{C}$. On note alors $\displaystyle\int f(t)\,\mathrm{d}t = F(t) + C^{te}$ ou plus abusivement $\displaystyle\int f(t)\,\mathrm{d}t = F(t)$.

Exemple Pour $\alpha \in \mathbb{C}^\star$, $\displaystyle\int \mathrm{e}^{\alpha t}\,\mathrm{d}t = \frac{1}{\alpha}\mathrm{e}^{\alpha t} + C^{te}$ car $\left(\mathrm{e}^{\alpha t}\right)' = \alpha\mathrm{e}^{\alpha t}$.

Proposition

On a équivalence entre :
(i) F est primitive de $f : I \to \mathbb{C}$;
(ii) $\mathrm{Re}(F)$ et $\mathrm{Im}(F)$ sont primitives de $\mathrm{Re}(f)$ et $\mathrm{Im}(f)$.

dém. :

Pour F dérivable, on sait que $(\mathrm{Re}F)' = \mathrm{Re}(F')$ et $(\mathrm{Im}F)' = \mathrm{Im}(F')$.

□

Théorème

Soit $f : I \to \mathbb{C}$ continue.
Pour tout $a \in I$, f possède une unique primitive s'annulant en a, c'est la fonction

$$x \mapsto \int_a^x f(t)\,\mathrm{d}t$$

dém. :
Unicité :
Deux primitives d'une même fonction diffèrent d'une constante, si donc elle prennent la même valeur en a, elles sont égales.
Existence :
Les fonctions $x \mapsto \int_a^x \mathrm{Re}(f(t))\,\mathrm{d}t$ et $x \mapsto \int_a^x \mathrm{Im}(f(t))\,\mathrm{d}t$ sont primitives s'annulant en a des fonctions réelles $\mathrm{Re}(f)$ et $\mathrm{Im}(f)$ donc $x \mapsto \int_a^x f(t)\,\mathrm{d}t = \int_a^x \mathrm{Re}(f(t))\,\mathrm{d}t + i\int_a^x \mathrm{Im}(f(t))\,\mathrm{d}t$ est primitive s'annulant en a de la fonction $x \mapsto f(t)$.
□

Corollaire

Si F est une primitive d'une fonction continue $f : I \to \mathbb{C}$ alors pour tout $a, b \in I$, $\int_a^b f(t)\,\mathrm{d}t = [F(t)]_a^b$.

dém. :
Semblable à celle du cadre réel.
□

Exemple Soit $a \in \mathbb{R}$ déterminons $\int_0^{2\pi} \mathrm{e}^{at}\cos(t)\,\mathrm{d}t$ et $\int_0^{2\pi} \mathrm{e}^{at}\sin(t)\,\mathrm{d}t$.

$$\int_0^{2\pi} \mathrm{e}^{at}\cos(t)\,\mathrm{d}t + i\int_0^{2\pi} \mathrm{e}^{at}\sin(t)\,\mathrm{d}t = \int_0^{2\pi} \mathrm{e}^{(a+i)t}\,\mathrm{d}t$$

Or

$$\int_0^{2\pi} \mathrm{e}^{(a+i)t}\,\mathrm{d}t = \frac{1}{a+i}\left[\mathrm{e}^{(a+i)t}\right]_0^{2\pi} = \frac{\mathrm{e}^{2\pi a}-1}{a+i} = \frac{\mathrm{e}^{2\pi a}-1}{a^2+1}(a-i)$$

Par suite

$$\int_0^{2\pi} \mathrm{e}^{at}\cos(t)\,\mathrm{d}t = \frac{\mathrm{e}^{2a\pi}-1}{a^2+1} \text{ et } \int_0^{2\pi} \mathrm{e}^{at}\sin(t)\,\mathrm{d}t = \frac{1-\mathrm{e}^{2a\pi}}{a^2+1}$$

Remarque Les techniques d'intégration par parties et de changement de variables (réelles) se généralisent aux intégrales complexes.

Exemple Calculons $\int_0^{\pi} t\mathrm{e}^{it}\,\mathrm{d}t$.
Par intégration par parties

$$\int_0^{\pi} t\mathrm{e}^{it}\,\mathrm{d}t = \left[-it\mathrm{e}^{it}\right]_0^{\pi} + i\int_0^{\pi} \mathrm{e}^{it} = i\pi + \left[\mathrm{e}^{it}\right]_0^{\pi} = -2 + i\pi$$

19.9 Formules de Taylor

I désigne un intervalle non singulier de $\mathbb{R}$.

19.9.1 Formule de Taylor avec reste intégral

Pour une fonction $f : I \to \mathbb{C}$ de classe $\mathcal{C}^1$, on sait exprimer la fonction f à partir de sa dérivée f' en vertu de la formule

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

Nous allons ici généraliser cette formule.

Théorème

Soient $f : I \to \mathbb{C}$ une fonction de classe $\mathcal{C}^{n+1}$ et $a \in I$.
Pour tout $x \in I$ on a :

$$f(x) = f(a)+(x-a)f'(a)+\frac{(x-a)^2}{2!}f''(a)+\cdots+\frac{(x-a)^n}{n!}f^{(n)}(a)+\int_a^x \frac{(x-t)^n}{n!}f^{(n+1)}(t)\,\mathrm{d}t$$

ou plus succinctement

$$f(x) = \sum_{k=0}^n \frac{(x-a)^k}{k!}f^{(k)}(a) + \int_a^x \frac{(x-t)^n}{n!}f^{(n+1)}(t)\,\mathrm{d}t$$

Le terme polynomial $\displaystyle\sum_{k=0}^n \frac{(x-a)^k}{k!}f^{(k)}(a)$ est appelée partie régulière du développement de Taylor de f à l'ordre n en a.
Le terme $\displaystyle\int_a^x \frac{(x-t)^n}{n!}f^{(n+1)}(t)\,\mathrm{d}t$ est appelé reste intégral de ce développement.

dém. :
Par récurrence sur $n \in \mathbb{N}$.
Pour $n = 0$: on a la formule connue

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

Supposons la propriété vérifiée au rang $n \geqslant 0$.
Soit $f : I \to \mathbb{C}$ de classe $\mathcal{C}^{n+2}$.
La fonction f est de classe $\mathcal{C}^{n+1}$ donc par hypothèse de récurrence

$$f(x) = \sum_{k=0}^n \frac{(x-a)^k}{k!}f^{(k)}(a) + \int_a^x \frac{(x-t)^n}{n!}f^{(n+1)}(t)\,\mathrm{d}t$$

Puisque la fonction $f^{(n+1)}$ est de classe $\mathcal{C}^1$ on peut réaliser une intégration par parties sur le reste intégral et ainsi

$$\int_a^x \frac{(x-t)^n}{n!}f^{(n+1)}(t)\,\mathrm{d}t = \left[-\frac{(x-t)^{n+1}}{(n+1)!}f^{(n+1)}(t)\right]_a^x + \int_a^x \frac{(x-t)^{n+1}}{(n+1)!}f^{(n+2)}(t)\,\mathrm{d}t$$

et ainsi

$$\int_a^x \frac{(x-t)^n}{n!}f^{(n+1)}(t)\,\mathrm{d}t = \frac{(x-a)^{n+1}}{(n+1)!}f^{(n+1)}(a) + \int_a^x \frac{(x-t)^{n+1}}{(n+1)!}f^{(n+2)}(t)\,\mathrm{d}t$$

puis

$$f(x)=\sum_{k=0}^{n+1}\frac{(x-a)^k}{k!}f^{(k)}(a)+\int_a^x\frac{(x-t)^{n+1}}{(n+1)!}f^{(n+2)}(t)\,\mathrm{d}t$$

Récurrence établie.

□

19.9.2 Inégalité de Taylor Lagrange

Théorème

Soient $f:I\to\mathbb{C}$ une fonction de classe $\mathcal{C}^{n+1}$ telle que $f^{(n+1)}$ soit bornée et $a\in I$.
Pour tout $x\in I$, on a

$$\left|f(x)-\sum_{k=0}^{n}\frac{(x-a)^k}{k!}f^{(k)}(a)\right|\leqslant\frac{|x-a|^{n+1}}{(n+1)!}M$$

avec $M=\sup_I\left|f^{(n+1)}\right|$

dém. :
Par la formule de Taylor avec reste intégral on peut écrire

$$f(x)-\sum_{k=0}^{n}\frac{(x-a)^k}{k!}f^{(k)}(a)=\int_a^x\frac{(x-t)^n}{n!}f^{(n+1)}(t)\,\mathrm{d}t$$

Cas $x\geqslant a$

$$\left|\int_a^x\frac{(x-t)^n}{n!}f^{(n+1)}(t)\,\mathrm{d}t\right|\leqslant\int_a^x\frac{(x-t)^n}{n!}\left|f^{(n+1)}(t)\right|\,\mathrm{d}t\leqslant\int_a^x\frac{(x-t)^n}{n!}M\,\mathrm{d}t=\frac{(x-a)^{n+1}}{(n+1)!}M$$

Cas $x\leqslant a$

$$\left|\int_a^x\frac{(x-t)^n}{n!}f^{(n+1)}(t)\,\mathrm{d}t\right|\leqslant\int_x^a\frac{(t-x)^n}{n!}\left|f^{(n+1)}(t)\right|\,\mathrm{d}t\leqslant\int_x^a\frac{(t-x)^n}{n!}M\,\mathrm{d}t=\frac{(a-x)^{n+1}}{(n+1)!}M$$

□

Remarque Cette énoncé est une généralisation de l'inégalité des accroissements finis qui se voit alors comme le cas particulier $n=1$.

Remarque L'hypothèse « $f^{(n+1)}$ est bornée »est automatiquement vérifiée si l'intervalle I est un segment $[a,b]$.

19.9.3 Applications

Exemple Montrons

$$\lim_{n\to\infty}\sum_{k=0}^{n}\frac{1}{k!}=\mathrm{e}$$

Considérons la fonction $f : x \mapsto \mathrm{e}^x$ qui est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$.
La partie régulière de f à l'ordre n en 0 est $\displaystyle\sum_{k=0}^{n} \frac{x^k}{k!} f^{(k)}(0) = \sum_{k=0}^{n} \frac{x^k}{k!}$ car $f^{(k)}(0) = 1$ pour tout $k \in \mathbb{N}$.
Puisque $f^{(n+1)}(x) = \mathrm{e}^x$, on a $\left|f^{(n+1)}(x)\right| \leqslant \mathrm{e}$ sur $[0, 1]$
En appliquant l'inégalité de Taylor Lagrange sur le segment $[0, 1]$, on obtient que pour tout $x \in [0, 1]$,

$$\left|f(x) - \sum_{k=0}^{n} \frac{x^k}{k!}\right| \leqslant \frac{x^{n+1}}{(n+1)!}\mathrm{e}$$

En particulier pour $x = 1$, on obtient

$$\left|\mathrm{e} - \sum_{k=0}^{n} \frac{1}{k!}\right| \leqslant \frac{\mathrm{e}}{(n+1)!} \to 0$$

et donc

$$\sum_{k=0}^{n} \frac{1}{k!} \to \mathrm{e}$$

Cette limite permet en pratique de calculer efficacement le nombre e à la précision voulue.

Exemple Montrons

$$\forall x \in [0, \pi/2]\,, x - \frac{x^3}{6} \leqslant \sin x \leqslant x - \frac{x^3}{6} + \frac{x^5}{120}$$

Pour cela formons le développement de Taylor de la fonction sin à l'ordre 4 en 0.
On obtient

$$\sin x = x - \frac{x^3}{6} + \int_0^x \frac{(x-t)^4}{24} \cos t \,\mathrm{d}t$$

Pour $x \in [0, \pi/2]$, on a $\forall t \in [0, x]\,, 0 \leqslant \cos t \leqslant 1$
Par suite

$$0 \leqslant \int_0^x \frac{(x-t)^4}{24} \cos t \,\mathrm{d}t \leqslant \int_0^x \frac{(x-t)^4}{24} \,\mathrm{d}t = \frac{x^5}{120}$$

puis l'encadrement proposé.

Exemple Soit $g : [0, 1] \to \mathbb{R}$ une fonction continue.
Déterminons les fonctions $f : [0, 1] \to \mathbb{R}$ deux fois dérivables vérifiant $f(0) = f(1) = 0$ et $f'' = g$.
Si f est solution alors f est de classe $\mathcal{C}^2$ et la formule de Taylor avec reste intégrale permet d'écrire

$$f(x) = f(0) + f'(0)x + \int_0^x (x-t)f''(t) \,\mathrm{d}t \text{ d'où } f(x) = f'(0) + \int_0^x (x-t)g(t) \,\mathrm{d}t$$

Or $f(1) = 0$ donc $f'(0) = -\displaystyle\int_0^1 (1-t)g(t) \,\mathrm{d}t$ puis

$$f(x) = x\int_0^1 (t-1)g(t) \,\mathrm{d}t + \int_0^x (x-t)g(t) \,\mathrm{d}t$$

Ce qui détermine f de manière unique.
Inversement, pour

$$f : x \mapsto x \int_0^1 (t-1)g(t)\,\mathrm{d}t + \int_0^x (x-t)g(t)\,\mathrm{d}t$$

on vérifie aisément $f(0) = f(1) = 0$.
De plus, puisque

$$f(x) = x \int_0^1 (t-1)g(t)\,\mathrm{d}t + x \int_0^x g(t)\,\mathrm{d}t - \int_0^x tg(t)\,\mathrm{d}t$$

la fonction f est dérivable par opérations sur de telles fonctions et

$$f'(x) = \int_0^1 (t-1)g(t)\,\mathrm{d}t + \int_0^x g(t)\,\mathrm{d}t + xg(x) - xg(x) = \int_0^1 (t-1)g(t)\,\mathrm{d}t + \int_0^x g(t)\,\mathrm{d}t$$

La fonction f' est encore dérivable et donc f est deux fois dérivable avec $f''(x) = g(x)$.
Finalement f est solution du problème posé.

Chapitre 20

Calcul de primitives

20.1 Primitives de fonctions rationnelles

20.1.1 Détermination de $\displaystyle\int \frac{\mathrm{d}x}{x-a}$ avec $a \in \mathbb{C}$

Pour $a \neq 0$, rappelons la formule

$$\int \frac{\mathrm{d}u}{u^2+a^2} = \frac{1}{a} \arctan \frac{u}{a}$$

Soit $a \in \mathbb{C}$ déterminons $\displaystyle\int \frac{\mathrm{d}x}{x-a}$.
Si $a \in \mathbb{R}$ alors

$$\int \frac{\mathrm{d}x}{x-a} = \ln|x-a|$$

Si $a \in \mathbb{C}\backslash\mathbb{R}$ alors on peut écrire $a = \alpha + i\beta$ avec $\alpha, \beta \in \mathbb{R}$ et $\beta \neq 0$.
On a alors

$$\frac{1}{x-a} = \frac{x-\alpha+i\beta}{(x-\alpha)^2+\beta^2} = \frac{x-\alpha}{(x-\alpha)^2+\beta^2} + i\frac{\beta}{(x-\alpha)^2+\beta^2}$$

or

$$\int \frac{x-\alpha}{(x-\alpha)^2+\beta}\,\mathrm{d}x = \frac{1}{2}\ln\left((x-\alpha)^2+\beta^2\right) = \ln|x-a|$$

et

$$\int \frac{\beta}{(x-\alpha)^2+\beta^2}\,\mathrm{d}x \underset{u=x-\alpha}{=} \int \frac{\beta}{u^2+\beta^2}\,\mathrm{d}u \underset{a=\beta}{=} \arctan\frac{u}{\beta} = \arctan\frac{x-\alpha}{\beta}$$

On en déduit

$$\int \frac{1}{x-a}\,\mathrm{d}x = \ln|x-a| + i\arctan\frac{x-\alpha}{\beta}$$

Exemple Déterminons $\displaystyle\int \frac{\mathrm{d}x}{x-j}$ sur $\mathbb{R}$.
On a

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

Or

$$\int \frac{x+\frac{1}{2}}{\left(x+\frac{1}{2}\right)^2+\frac{3}{4}}\,\mathrm{d}x = \frac{1}{2}\ln\left(x^2+x+1\right)$$

et

$$\int \frac{1}{\left(x+\frac{1}{2}\right)^2+\frac{3}{4}} = \frac{2}{\sqrt{3}}\arctan\frac{2x+1}{\sqrt{3}}$$

Par suite

$$\int \frac{\mathrm{d}x}{x-j} = \frac{1}{2}\ln\left(x^2+x+1\right) + i\arctan\frac{2x+1}{\sqrt{3}}$$

20.1.2 Détermination de $\int \frac{\alpha x+\beta}{x^2+px+q}\,\mathrm{d}x$ sur $\mathbb{R}$

Soient $\alpha, \beta \in \mathbb{R}$ et $p, q \in \mathbb{R}$ tels que $p^2-4q<0$.
Notons que l'équation $x^2+px+q=0$ n'a pas de racines réelles.
On peut écrire

$$\frac{\alpha x+\beta}{x^2+px+q} = \frac{\frac{\alpha}{2}(2x+p)+\lambda}{x^2+px+q}$$

Par suite

$$\int \frac{\alpha x+\beta}{x^2+px+q}\,\mathrm{d}x = \frac{\alpha}{2}\int \frac{(2x+p)}{x^2+px+q}\,\mathrm{d}x + \lambda\int \frac{\mathrm{d}x}{x^2+px+q}$$

On a

$$\int \frac{2x+p}{x^2+px+q}\,\mathrm{d}x = \ln(x^2+px+q)$$

Il reste à déterminer $\int \frac{\mathrm{d}x}{x^2+px+q}$. Pour cela on écrit le trinôme du dénominateur sous forme canonique

$$x^2+px+q = \left(x+\frac{p}{2}\right)^2 + \frac{4q-p^2}{4} = \left(x+\frac{p}{2}\right)^2+\delta^2$$

On réalise le changement de variable : $u = x+p/2$

$$\int \frac{\mathrm{d}x}{x^2+px+q} = \int \frac{\mathrm{d}x}{\left(x+\frac{p}{2}\right)^2+\delta^2} = \int \frac{\mathrm{d}u}{u^2+\delta^2} = \frac{1}{\delta}\arctan\frac{2x+p}{2\delta}$$

Exemple Déterminons $\int \frac{x+1}{x^2+x+1}\,\mathrm{d}x$ sur $\mathbb{R}$.

$$\int \frac{x+1}{x^2+x+1}\,\mathrm{d}x = \int \frac{\frac{1}{2}(2x+1)+\frac{1}{2}}{x^2+x+1}\,\mathrm{d}x = \frac{1}{2}\int \frac{2x+1}{x^2+x+1}\,\mathrm{d}x + \frac{1}{2}\int \frac{\mathrm{d}x}{x^2+x+1}$$

avec

$$\int \frac{2x+1}{x^2+x+1}\,\mathrm{d}x = \ln(x^2+x+1)$$

et

$$\int \frac{\mathrm{d}x}{x^2+x+1} = \int \frac{\mathrm{d}x}{(x+\frac{1}{2})^2+\frac{3}{4}} \underset{t=x+\frac{1}{2}}{=} \int \frac{\mathrm{d}t}{t^3+\frac{3}{4}} \underset{a=\frac{\sqrt{3}}{2}}{=} \frac{2}{\sqrt{3}}\arctan\frac{2x+1}{\sqrt{3}}$$

Finalement

$$\int \frac{x+1}{x^2+x+1}\,\mathrm{d}x = \frac{1}{2}\ln(x^2+x+1) + \frac{1}{\sqrt{3}}\arctan\frac{2x+1}{\sqrt{3}}$$

20.1.3 Détermination d'une primitive d'une fonctions rationnelles

Soit $F \in \mathbb{K}(X)$, pour déterminer $\int F(x)\,\mathrm{d}x$ sur les intervalles où $x \mapsto F(x)$ est définie :
- on écrit F sous forme irréductible ;
- si on peut écrire $F(x) = x^{n-1}G(x^n)$ avec $G \in \mathbb{C}(X)$ et $n \geqslant 2$ alors on réalise le changement de variable $t = x^n$ car

$$\int F(x)\,\mathrm{d}x = \int x^{n-1}G(x^n)\,\mathrm{d}x = \frac{1}{n}\int G(t)\,\mathrm{d}t$$

ce qui réduit le problème ;
- on réalise la décomposition en éléments simple de F dans $\mathbb{C}(X)$.
- on détermine une primitive de chacun des termes de la décomposition précédente.

Exemple Déterminons $\displaystyle\int \frac{x^2}{1+x^3}\,\mathrm{d}x$ sur $]-\infty,-1[$ et $]-1,+\infty[$.

$$\int \frac{x^2}{1+x^3}\,\mathrm{d}x \underset{t=x^3}{=} \int \frac{1}{3}\frac{\mathrm{d}t}{1+t} = \frac{1}{3}\ln|1+t| = \frac{1}{3}\ln\left|1+x^3\right|$$

Exemple Déterminons $\displaystyle\int \frac{\mathrm{d}x}{1-x^2}$ sur $]-\infty,-1[\,,]-1,1[$ ou $]1,+\infty[$.
On a la décomposition en éléments simples

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

On en déduit

$$\int \frac{\mathrm{d}x}{1-x^2} = \frac{1}{2}\int \frac{\mathrm{d}x}{x+1} - \frac{1}{2}\int \frac{\mathrm{d}x}{x-1} = \frac{1}{2}\ln|x+1| - \frac{1}{2}\ln|x-1| = \frac{1}{2}\ln\left|\frac{x+1}{x-1}\right|$$

Remarque Pour une fonction rationnelle réelle, il est recommandé de regrouper les termes conjugués en $\dfrac{1}{x-a}$ et $\dfrac{1}{x-\bar{a}}$ avant de déterminer une primitive.

Exemple Déterminons $\displaystyle\int \frac{x}{x^3+1}\,\mathrm{d}x$ sur $]-\infty,-1[$ ou $]-1,+\infty[$.
Puisque $x^3+1 = (x+1)(x^2-x+1) = (x+1)(x+j)(x+j^2)$, on a la décomposition

$$\frac{x}{x^3+1} = \frac{a}{x+1} + \frac{b}{x+j} + \frac{\bar{b}}{x+j^2} = \frac{a}{x+1} + \frac{\alpha x+\beta}{x^2-x+1}$$

avec

$$a = \left.\frac{x}{x^2-x+1}\right|_{x=-1} = -\frac{1}{3}$$

$x \times \dfrac{x}{x^3+1} \xrightarrow[x\to+\infty]{} 0$ donne $a+\alpha = 0$ donc $\alpha = 1/3$.
Enfin, en évaluant la relation en 0, on obtient $a+\beta = 0$ donc $\beta = 1/3$.

Ainsi, on obtient la décomposition

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

D'une part

$$\int \frac{\mathrm{d}x}{x+1} = \ln|x+1|$$

D'autre part

$$\int \frac{x+1}{x^2-x+1}\,\mathrm{d}x = \frac{1}{2}\int \frac{2x-1}{x^2-x+1}\,\mathrm{d}x + \frac{3}{2}\int \frac{\mathrm{d}x}{x^2-x+1}$$

avec

$$\int \frac{2x-1}{x^2-x+1}\,\mathrm{d}x = \frac{1}{2}\ln\left(x^2-x+1\right)$$

et

$$\int \frac{\mathrm{d}x}{x^2-x+1} = \int \frac{\mathrm{d}x}{\left(x-\frac{1}{2}\right)^2+\frac{3}{4}} = \frac{2}{\sqrt{3}}\arctan\frac{2x-1}{\sqrt{3}}$$

Au final

$$\int \frac{x}{x^3+1}\,\mathrm{d}x = -\frac{1}{3}\ln|x+1| + \frac{1}{6}\ln\left|x^2-x+1\right| + \frac{1}{\sqrt{3}}\arctan\frac{2x-1}{\sqrt{3}}$$

Exemple Déterminons $\displaystyle\int \frac{\mathrm{d}x}{x(x^2+2x+3)}$ sur $]-\infty,0[$ ou $]0,+\infty[$.
En introduisant ω et $\bar{\omega}$ les racines de $x^2+2x+3=0$, on a la décomposition

$$\frac{1}{x(x^2+2x+3)} = \frac{a}{x} + \frac{\lambda}{x-\omega} + \frac{\bar{\lambda}}{x-\bar{\omega}} = \frac{a}{x} + \frac{bx+c}{x^2+2x+3}$$

avec

$$a = \left.\frac{1}{x^2+2x+3}\right|_{x=0} = \frac{1}{3}$$

$x\dfrac{1}{x(x^2+2x+3)} \xrightarrow[x\to+\infty]{} 0$ donne $a+b=0$ donc $b=-1/3$.
En évaluant en 1, on obtient $1/6 = a+(b+c)/6$ donc $c=-2/3$.

$$\int \frac{\mathrm{d}x}{x(x^2+2x+3)} = \frac{1}{3}\int\frac{\mathrm{d}x}{x} - \frac{1}{3}\int \frac{x+2}{x^2+2x+3}\,\mathrm{d}x$$

avec

$$\int \frac{x+2}{x^2+2x+3}\,\mathrm{d}x = \frac{1}{2}\int \frac{x+1}{x^2+2x+3}\,\mathrm{d}x + \int \frac{\mathrm{d}x}{(x+1)^2+2}$$

Au final

$$\int \frac{\mathrm{d}x}{x(x^2+2x+3)} = \frac{1}{3}\ln|x| - \frac{1}{6}\ln(x^2+2x+3) - \frac{1}{3\sqrt{2}}\arctan\frac{x+1}{\sqrt{2}} + C^{te}$$

Exemple Déterminons $\displaystyle\int \frac{x^2+1}{(x^2+x+1)^2}\,\mathrm{d}x$ sur $\mathbb{R}$
Puisque $(x^2+x+1)^2 = (x-j)^2(x-j^2)^2$, on a la décomposition

$$\frac{x^2+1}{(x^2+x+1)^2} = \frac{a}{(x-j)^2} + \frac{b}{x-j} + \frac{\bar{a}}{(x-j^2)^2} + \frac{\bar{b}}{x-j^2} = \frac{a}{(x-j)^2} + \frac{\bar{a}}{(x-j^2)^2} + \frac{\alpha x+\beta}{x^2+x+1}$$

avec

$$a = \left.\frac{x^2+1}{(x-j^2)^2}\right|_{x=j} = \frac{j}{3}$$

$x \times \dfrac{x^2+1}{(x^2+x+1)^2} \xrightarrow[x\to+\infty]{} 0$ donne $\alpha = 0$.
En évaluant en 0, on obtient $1 = \dfrac{1}{3j} + \dfrac{1}{3j^2} + \beta$ donc $\beta = \dfrac{4}{3}$.

$$\int \frac{a}{(x-j)^2}\,\mathrm{d}x = -\frac{1}{3}\frac{j}{x-j} = -\frac{1}{3}\frac{jx-1}{x^2+x+1}$$

Par conjugaison

$$\int \frac{\bar{a}}{(x-j^2)^2}\,\mathrm{d}x = -\frac{1}{3}\frac{j^2x-1}{x^2+x+1}$$

Enfin

$$\int \frac{\mathrm{d}x}{x^2+x+1} = \int \frac{\mathrm{d}x}{\left(x+\frac{1}{2}\right)^2+\frac{3}{4}} = \frac{2}{\sqrt{3}}\arctan\frac{2x+1}{\sqrt{3}}$$

Finalement :

$$\int \frac{x^2+1}{(x^2+x+1)^2}\,\mathrm{d}x = \frac{1}{3}\frac{x+2}{x^2+x+1} + \frac{8}{3\sqrt{3}}\arctan\frac{2x+1}{\sqrt{3}}$$

20.2 Détermination se ramenant à des fonctions rationnelles

20.2.1 Fonctions rationnelles en $\mathrm{e}^{\alpha x}$ avec $\alpha \in \mathbb{R}^\star$

Soit $F \in \mathbb{R}(X)$, on veut déterminer $\displaystyle\int F(\mathrm{e}^{\alpha x})\,\mathrm{d}x$ sur les intervalles où cela a un sens.
On réalise le changement de variable $u = \mathrm{e}^{\alpha x}$, $\mathrm{d}u = \alpha u\,\mathrm{d}x$ et alors

$$\int F(\mathrm{e}^{\alpha x})\,\mathrm{d}x = \frac{1}{\alpha}\int \frac{F(u)\,\mathrm{d}u}{u}$$

On sait ensuite poursuivre le calcul…

Exemple Déterminons $\displaystyle\int \frac{\mathrm{d}x}{(\mathrm{e}^x+1)^2}$ sur $\mathbb{R}$.
On pose $u = \mathrm{e}^x$ et on a

$$\int \frac{\mathrm{d}x}{(\mathrm{e}^x+1)^2} = \int \frac{\mathrm{d}u}{u(u+1)^2}$$

Par la décomposition

$$\frac{1}{u(u+1)^2} = \frac{1}{u} - \frac{1}{(u+1)^2} - \frac{1}{u+1}$$

on obtient

$$\int \frac{\mathrm{d}x}{(\mathrm{e}^x+1)^2} = \ln|u| - \ln|u+1| + \frac{1}{u+1} = x - \ln(1+\mathrm{e}^x) + \frac{1}{\mathrm{e}^x+1}$$

20.2.2 Fonctions rationnelles en cos x et sin x

On appelle polynôme réels en deux indéterminées X et Y toute expression de la forme

$$P(X,Y) = \sum_{p=0}^{N}\sum_{q=0}^{M} a_{p,q}X^pY^q \text{ avec } a_{p,q} \in \mathbb{R}$$

On note $\mathbb{R}[X,Y]$ l'ensemble de ces polynômes.
On appelle fraction rationnelle réelle en les indéterminées X et Y toute expression de la forme

$$R(X,Y) = \frac{P(X,Y)}{Q(X,Y)} \text{ avec } P,Q \in \mathbb{R}[X,Y], Q \neq 0$$

On note $\mathbb{R}(X,Y)$ l'ensemble de ces fractions.
Soit $R \in \mathbb{R}(X,Y)$, on veut déterminer $\int R(\cos x, \sin x)\,\mathrm{d}x$ sur tout intervalle où $x \mapsto R(\sin x, \cos x)$ est définie.
On procède par changement de variable.

20.2.2.1 Règles de Bioche

Théorème

Notons $f(x) = R(\cos x, \sin x)$
Si $f(-x)\,\mathrm{d}(-x) = f(x)\,\mathrm{d}x$ alors on pose $t = \cos x$
Si $f(\pi - x)\,\mathrm{d}(\pi - x) = f(x)\,\mathrm{d}x$ alors on pose $t = \sin x$
Si $f(x+\pi)\,\mathrm{d}(x+\pi) = f(x)\,\mathrm{d}x$ alors on pose $t = \tan x$
Dans chaque cas, le changement de variable proposé transforme la détermination de $\int f(x)\,\mathrm{d}x$ en la détermination d'une primitive d'une fonction rationnelle en t.

dém. :
Ecrivons $R(X,Y) = \dfrac{P(X,Y)}{Q(X,Y)}$ avec P,Q polynômes de $\mathbb{R}[X,Y]$.
Cas $f(-x)\,\mathrm{d}(-x) = f(x)\,\mathrm{d}x$:
En regroupant les puissances paires et les puissances impaires de Y au numérateur et au dénominateur, on obtient
□

$$R(X,Y) = \frac{P_1(X,Y^2) + YP_2(X,Y^2)}{Q_1(X,Y^2) + YQ_2(X,Y^2)}$$

En multipliant, en haut et en bas par la quantité $Q_1(X,Y^2) - YQ_2(X,Y^2)$, on parvient à

$$R(X,Y) = \frac{\left(P_1(X,Y^2) + YP_2(X,Y^2)\right)\left(Q_1(X,Y^2) - YQ_2(X,Y^2)\right)}{Q_1^2(X,Y^2) - Y^2Q_2^2(X,Y^2)}$$

Le dénominateur apparaît désormais comme n'étant constitué que de puissances paires de Y et en développant le numérateur, on parvient à écrire

$$R(X,Y) = R_1(X,Y^2) + YR_2(X,Y^2)$$

Ainsi

$$f(x) = R_1(\cos x, \sin^2 x) + \sin x R_2(\cos x, \sin^2 x)$$

Puisque $\sin^2 x = 1 - \cos^2 x$, on obtient

$$f(x) = f_1(\cos x) + \sin x.f_2(\cos x)$$

avec f_1 et f_2 fonctions rationnelles.
Puisque l'on suppose $f(-x)\,\mathrm{d}(-x) = f(x)\,\mathrm{d}x$, on a $f(-x) = -f(x)$ et donc

$$f_1(\cos x) - \sin x.f_2(\cos x) = -f_1(\cos x) - \sin x.f_2(\cos x)$$

Par suite $f_1(\cos x) = 0$ puis $f(x) = \sin x.f_2(\cos x)$.
Par le changement de variable $t = \cos x$,

$$\int f(x)\,\mathrm{d}x = -\int f_2(t)\,\mathrm{d}t$$

Cas $f(\pi - x)\,\mathrm{d}(\pi - x) = f(x)\,\mathrm{d}x$:
La démarche est semblable en écrivant

$$R(X, Y) = R_1(X^2, Y) + XR_2(X^2, Y)$$

puis

$$f(x) = f_1(\sin x) + \cos x.f_2(\sin x)$$

avec f_1 nulle par l'hypothèse $f(\pi - x) = -f(x)$.
Cas $f(x + \pi)\,\mathrm{d}(x + \pi) = f(x)\,\mathrm{d}x$:
En reprenant les démarches qui précèdent, on écrit

$$R(X, Y) = R_1(X^2, Y^2) + XR_2(X^2, Y) + YR_3(X^2, Y^2) + XYR_4(X^2, Y^2)$$

On peut alors écrire

$$f(x) = f_1(\cos^2 x) + \cos x.f_2(\cos^2 x) + \sin x.f_3(\cos^2 x) + \cos x.\sin x.f_4(\cos^2 x)$$

avec f_1, f_2, f_3, f_4 fonctions rationnelles.
L'hypothèse $f(x + \pi)\,\mathrm{d}(x + \pi) = f(x)\,\mathrm{d}x$ donne $f(x + \pi) = f(x)$ ce qui entraîne

$$\cos x.f_2(\cos^2 x) + \sin x.f_3(\cos^2 x) = 0$$

Par suite

$$f(x) = f_1(\cos^2 x) + \cos x.\sin x.f_4(\cos^2 x) = f_1(\cos^2 x) + \frac{\sin x}{\cos x}\cos^2 x.f_4(\cos^2 x)$$

Sachant $\cos^2 x = \dfrac{1}{1+\tan^2 x}$, on obtient $f(x) = g(\tan x)$ avec g fonction rationnelle.
Par suite en posant $t = \tan x$,

$$\int f(x)\,\mathrm{d}x = \int \frac{g(t)}{1+t^2}\,\mathrm{d}t$$

Exemple Déterminons $\displaystyle\int \frac{\sin^3 x}{1+\cos^2 x}\,\mathrm{d}x$ sur $\mathbb{R}$.
Ici $f(-x)\,\mathrm{d}(-x) = f(x)\,\mathrm{d}x$.
Posons $t = \cos x$.

$$\int \frac{\sin^3 x}{1+\cos^2 x}\,\mathrm{d}x = \int \frac{1-\cos^2 x}{1+\cos^2 x}\sin x\,\mathrm{d}x$$

donc

$$\int \frac{\sin^3 x}{1+\cos^2 x}\,\mathrm{d}x = \int \frac{t^2-1}{t^2+1}\,\mathrm{d}t = t - 2\arctan t = \cos x - 2\arctan(\cos x)$$

Exemple Déterminons $\displaystyle\int \frac{\mathrm{d}x}{\cos x}$ sur $I_k = \left]-\frac{\pi}{2}+k\pi, \frac{\pi}{2}+k\pi\right[$ (avec $k \in \mathbb{Z}$)
Ici $f(\pi - x)\,\mathrm{d}(\pi - x) = f(x)\,\mathrm{d}x$.
Posons $t = \sin x$.

$$\int \frac{\mathrm{d}x}{\cos x} = \int \frac{\cos x\,\mathrm{d}x}{\cos^2 x} = \int \frac{\mathrm{d}t}{1-t^2} = \frac{1}{2}\ln\frac{1+\sin x}{1-\sin x}$$

Exemple Détermination de $\displaystyle\int \frac{\mathrm{d}x}{1+\sin^2 x}$ sur $\mathbb{R}$.
Ici $f(x+\pi)\,\mathrm{d}(x+\pi) = f(x)\,\mathrm{d}x$.
Posons $t = \tan x$.
Or la fonction tangente n'est pas définie en les $x_k = \dfrac{\pi}{2} + k\pi$ avec $k \in \mathbb{Z}$.
On ne peut donc réaliser le changement de variable voulu que sur les intervalles
$I_k = \left]-\dfrac{\pi}{2}+k\pi, \dfrac{\pi}{2}+k\pi\right[=]x_{k-1}, x_k[$ (avec $k \in \mathbb{Z}$)
Commençons par déterminer $\displaystyle\int \frac{\mathrm{d}x}{1+\sin^2 x}$ sur chaque I_k.
$t = \tan x,\ \mathrm{d}t = (1+\tan^2 x)\,\mathrm{d}x = (1+t^2)\,\mathrm{d}x$ et

$$\sin^2 x = 1 - \cos^2 x = 1 - \frac{1}{1+t^2} = \frac{t^2}{1+t^2}$$

donc

$$\int \frac{\mathrm{d}x}{1+\sin^2 x} = \int \frac{\mathrm{d}t}{(1+t^2)(1+\frac{t^2}{1+t^2})} = \int \frac{\mathrm{d}t}{1+2t^2} = \frac{1}{2}\int \frac{\mathrm{d}t}{t^2+\frac{1}{2}} = \frac{1}{\sqrt{2}}\arctan(\sqrt{2}\tan x)$$

Déterminons maintenant $\displaystyle\int \frac{\mathrm{d}x}{1+\sin^2 x}$ sur $\mathbb{R}$.
Soit F une primitive de $x \mapsto \dfrac{1}{1+\sin^2 x}$ sur $\mathbb{R}$, il en existe car $x \mapsto \dfrac{1}{1+\sin^2 x}$ est continue sur $\mathbb{R}$.

Pour tout $k \in \mathbb{Z}$, F est primitive sur I_k donc il existe $C_k \in \mathbb{R}$ tel que

$$\forall x \in I_k, F(x) = \frac{1}{\sqrt{2}} \arctan(\sqrt{2} \tan x) + C_k$$

La continuité de F en $x_k = \frac{\pi}{2} + k\pi$ va donner une relation entre C_k et C_{k+1} :
D'une part

$$F(x_k) = \lim_{x \to x_k^-} F(x) = \frac{\pi}{2\sqrt{2}} + C_k$$

D'autre part

$$F(x_k) = \lim_{x \to x_k^+} F(x) = -\frac{\pi}{2\sqrt{2}} + C_{k+1}$$

On en déduit $C_{k+1} = \frac{\pi}{\sqrt{2}} + C_k$ puis $C_k = \frac{k\pi}{\sqrt{2}} + C_0$.
Finalement

$$F(x) = \begin{cases} \frac{1}{\sqrt{2}} \arctan(\sqrt{2} \tan x) + \frac{k\pi}{\sqrt{2}} + C_0 & \text{si } x \in I_k \\ \frac{(2k+1)\pi}{2\sqrt{2}} + C_0 & \text{si } x = x_k \end{cases}$$

Inversement
Etant assuré de l'existence de primitives sur $\mathbb{R}$ et sachant celles-ci déterminées à une constante près, les fonctions proposées sont bien solutions.
En application, calculons

$$\int_0^{2\pi} \frac{\mathrm{d}x}{1 + \sin^2 x}$$

On a

$$\int_0^{2\pi} \frac{\mathrm{d}x}{1 + \sin^2 x} = [F(x)]_0^{2\pi} = F(2\pi) - F(0)$$

$0 \in I_0$ donc $F(0) = C_0$ et $2\pi \in I_2$ donc $F(2\pi) = \sqrt{2}\pi + C_0$.
Ainsi

$$\int_0^{2\pi} \frac{\mathrm{d}x}{1 + \sin^2 x} = \sqrt{2}\pi$$

Attention : Si on calcule directement l'intégrale par le changement de variable $t = \tan x$,
$\mathrm{d}t = (1 + t^2)\,\mathrm{d}x$,
Pour $x = 0, t = 0$ et pour $x = 2\pi$, $t = 0$.
Le changement de variable donne $\displaystyle\int_0^{2\pi} \frac{\mathrm{d}x}{1 + \sin^2 x} = \int_0^0 (...) = 0$!
Ici le résultat est faux car le changement de variable $t = \tan x$ est incorrect car non défini aux points intermédiaires $\pi/2$ et $3\pi/2$.

20.2.2.2 Méthode systématique

Lorsque les règles de Bioche échouent, on peut néanmoins réaliser le changement de variable $t = \tan(x/2)$¨.
Pour celui-ci $\mathrm{d}t = \dfrac{1}{2}(1+t^2)\,\mathrm{d}x$ et sachant $\cos x = \dfrac{1-t^2}{1+t^2}$ et $\sin x = \dfrac{2t}{1+t^2}$ on obtient :

$$\int R(\cos x, \sin x)\,\mathrm{d}x = \int R\left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}\right)\frac{2}{1+t^2}\,\mathrm{d}t$$

On est ainsi ramené au calcul d'une primitive d'une fonction rationnelle.

Attention : Ce changement de variable pose problème aux points $x_k = \pi + 2k\pi$ et ne peut donc être réalisé que sur des intervalles inclus dans $I_k =]-\pi + 2k\pi, \pi + 2k\pi[$ (avec $k \in \mathbb{Z}$)
Pour déterminer une primitive sur un intervalle plus grand, il conviendra de procéder au recollement des primitives aux points de jonction.

Attention : Lors du changement de variable $t = \tan\dfrac{x}{2}$ on a $\mathrm{d}t = \dfrac{1}{2}(1+t^2)\,\mathrm{d}x$.
Il ne faut pas oublier le facteur $\dfrac{1}{2}$!

Exemple Déterminons $\displaystyle\int \frac{\mathrm{d}x}{1+\cos x}$ sur $I_k =]-\pi + 2k\pi, \pi + 2k\pi[$.
Ici, les règles de Bioche échouent.
Réalisons le changement de variable $t = \tan\dfrac{x}{2}$.

$$\int \frac{\mathrm{d}x}{1+\cos x} = \int \frac{2}{1+t^2}\frac{1}{1+\frac{1-t^2}{1+t^2}}\,\mathrm{d}t = \int \mathrm{d}t = t = \tan\frac{x}{2}$$

Exemple Calculons $\displaystyle\int_0^{\pi/2} \frac{\mathrm{d}x}{2+\sin x}$.
Les règles de Bioche échouent.
Réalisons le changement de variable $t = \tan\dfrac{x}{2}$.

$$\int_0^{\pi/2} \frac{\mathrm{d}x}{2+\sin x} = \int_0^1 \frac{2\,\mathrm{d}t}{1+t^2}\frac{1}{2+\frac{2t}{1+t^2}} = \int_0^1 \frac{\mathrm{d}t}{t^2+t+1} = \frac{\pi}{3\sqrt{3}}$$

20.2.3 Fonctions rationnelles en ch$x$ et sh$x$

Soit $R \in \mathbb{R}(X,Y)$, on veut déterminer $\displaystyle\int R(\mathrm{ch}x, \mathrm{sh}x)\,\mathrm{d}x$ sur tout intervalle où $x \mapsto R(\mathrm{ch}x, \mathrm{sh}x)$ est définie.

20.2.3.1 Règles de Bioche

Soit à calculer $\displaystyle\int R(\mathrm{ch}x, \mathrm{sh}x)\,\mathrm{d}x$.

Si les règles de Bioche suggèrent de calculer $\displaystyle\int R(\cos x, \sin x)\,\mathrm{d}x$ par le changement de variable $t =$

$\cos x, \sin x$ ou $\tan x$ alors il peut être pertinent de calculer $\int R(\mathrm{ch}x, \mathrm{sh}x)\,\mathrm{d}x$ par le changement de variable resp. $t = \mathrm{ch}x, \mathrm{sh}x$ ou $t = \mathrm{th}x$.

Exemple Déterminons $\int \frac{\mathrm{d}x}{\mathrm{ch}x}$ sur $\mathbb{R}$.
Les règles de Bioche invitent au changement de variable $t = \mathrm{sh}x$.

$$\int \frac{\mathrm{d}x}{\mathrm{ch}x} = \int \frac{\mathrm{ch}x\,\mathrm{d}x}{1+\mathrm{sh}^2 x} \underset{t=\mathrm{sh}x}{=} \int \frac{\mathrm{d}t}{1+t^2} = \arctan t = \arctan(\mathrm{sh}x)$$

20.2.3.2 Méthode systématique

Si les règles de Bioche échouent, on écrit $\mathrm{sh}x$ et $\mathrm{ch}x$ sous forme exponentielle et on réalise le changement de variable : $t = \mathrm{e}^x$.

Exemple Déterminons $\int \frac{\mathrm{d}x}{1+\mathrm{ch}x}$ sur $\mathbb{R}$.
Les règles de Bioche échouent.
Réalisons le changement de variable $t = \mathrm{e}^x$.

$$\int \frac{\mathrm{d}x}{1+\mathrm{ch}x} = \int \frac{\mathrm{d}t}{t(1+\frac{1}{2}(t+1/t))} = \int \frac{2\,\mathrm{d}t}{(t+1)^2} = -\frac{2}{t+1} = -\frac{2}{\mathrm{e}^x+1}$$

20.2.4 Fonctions rationnelles en x et $\sqrt[n]{\frac{ax+b}{cx+d}}$

Soient $n \in \mathbb{N}$ tel que $n \geqslant 2$ et $a, b, c, d \in \mathbb{R}$ tels que $ad - bc \neq 0$.
Pour $R \in \mathbb{R}(X, Y)$, on veut déterminer

$$\int R\left(x, \sqrt[n]{\frac{ax+b}{cx+d}}\right)\mathrm{d}x$$

sur les intervalles où cela est possible.
On effectue pour cela le changement de variable

$$t = \sqrt[n]{\frac{ax+b}{cx+d}}$$

et on commence par exprimer x en fonction de t.
$t^n = \frac{ax+b}{cx+d}$ donne $t^n(cx+d) = ax+b$ puis $x(ct^n - a) = b - dt^n$ et enfin $x = \frac{b-dt^n}{ct^n-a}$
En effet $ct^n - a \neq 0$ puisque $ct^n - a = 0 \Rightarrow b - dt^n = 0$ puis $bc - ad = 0$ ce qui est exclu.
On peut alors exprimer simplement $\mathrm{d}x$ en fonction de $\mathrm{d}t$

$$\mathrm{d}x = \left(\frac{b-dt^n}{ct^n-a}\right)'\mathrm{d}t = \frac{n(ad-bc)t^{n-1}}{(ct^n-a)^2}\,\mathrm{d}t$$

puis on est prêt pour procéder au changement de variable :

$$\int R\left(x, \sqrt[n]{\frac{ax+b}{cx+d}}\right)\mathrm{d}x = \int R\left(\frac{b-dt^n}{ct^n-a}, t\right)\frac{n(ad-bc)t^{n-1}}{(ct^n-a)^2}\,\mathrm{d}t$$

Exemple Déterminons $\displaystyle\int \frac{1}{x}\sqrt{\frac{1-x}{1+x}}\,\mathrm{d}x$ sur $]-1,0[$ ou $]0,1]$.
On pose $t = \sqrt{\dfrac{1-x}{1+x}}$.
On a $x = \dfrac{1-t^2}{1+t^2} = -1 + \dfrac{2}{1+t^2}$ et par suite $\mathrm{d}x = -\dfrac{4t}{(1+t^2)^2}\,\mathrm{d}t$.
Ainsi

$$\int \frac{1}{x}\sqrt{\frac{1-x}{1+x}}\,\mathrm{d}x = \int \frac{4t^2}{(t^2-1)(t^2+1)}\,\mathrm{d}t$$

Via une décomposition en éléments simples on peut écrire

$$\frac{4t^2}{(t^2-1)(t^2+1)} = \frac{1}{t-1} - \frac{1}{t+1} + \frac{2}{1+t^2}$$

Par suite

$$\int \frac{1}{x}\sqrt{\frac{1-x}{1+x}}\,\mathrm{d}x = \ln\left|\frac{t-1}{t+1}\right| + 2\arctan t = \ln\frac{\sqrt{1+x}-\sqrt{1-x}}{\sqrt{1+x}+\sqrt{1-x}} + 2\arctan\sqrt{\frac{1-x}{1+x}}$$

Remarque Cas $c = 0$ et $d = 1$:
On détermine aisément $\displaystyle\int R(x,\ \sqrt[n]{ax+b})\,\mathrm{d}x$ via le changement de variable $t = \sqrt[n]{ax+b}$.

Exemple Déterminons $\displaystyle\int \frac{(1+x)^{1/3}-1}{x(1+x)^{2/3}}\,\mathrm{d}x$ sur $]0,+\infty[$.
On pose $t = (1+x)^{1/3}$.
On a $x = t^3 - 1$ et par suite $\mathrm{d}x = 3t^2\,\mathrm{d}t$
Ainsi

$$\int \frac{(1+x)^{1/3}-1}{x(1+x)^{2/3}}\,\mathrm{d}x = 3\int \frac{\mathrm{d}t}{t^2+t+1}$$

Puisque

$$\int \frac{\mathrm{d}t}{t^2+t+1} = \int \frac{\mathrm{d}t}{\left(t+\frac{1}{2}\right)^2+\frac{3}{4}} = \frac{2}{\sqrt{3}}\arctan\frac{2t+1}{\sqrt{3}}$$

on obtient

$$\int \frac{(1+x)^{1/3}-1}{x(1+x)^{2/3}}\,\mathrm{d}x = 2\sqrt{3}\arctan\frac{2(1+x)^{1/3}+1}{\sqrt{3}}$$

Exemple Déterminons $\displaystyle\int \frac{\mathrm{d}x}{\sqrt{x}+\sqrt[3]{x}}$ sur $]0,+\infty[$.
On pose $t = \sqrt[6]{x}$, on a $x = t^6$ et $\mathrm{d}x = 6t^5\,\mathrm{d}t$.
Ainsi

$$\int \frac{\mathrm{d}x}{\sqrt{x}+\sqrt[3]{x}} = 6\int \frac{t^3\,\mathrm{d}t}{t+1}$$

Par décomposition en éléments simples on peut écrire

$$\frac{t^3}{t+1} = t^2 - t + 1 - \frac{1}{t+1}$$

On en déduit

$$\int \frac{\mathrm{d}x}{\sqrt{x} + \sqrt[3]{x}} = 2\sqrt{x} - 3\sqrt[3]{x} + 6\sqrt[6]{x} - 6\ln(1+\sqrt[6]{x})$$

20.2.5 Fonctions rationnelles en x et $\sqrt{ax^2+bx+c}$

Soient $a, b, c \in \mathbb{R}$ tels que $a \neq 0$ et $\Delta = b^2 - 4ac \neq 0$.
Pour $R \in \mathbb{R}(X,Y)$, on veut déterminer $\displaystyle\int R(x, \sqrt{ax^2+bx+c})\,\mathrm{d}x$ sur tout intervalle où cela est possible.
On commence par écrire le trinôme ax^2+bx+c sous forme canonique

$$ax^2+bx+c = a\left(\left(x+\frac{b}{2a}\right)^2 - \frac{\Delta}{4a^2}\right)$$

Par un changement de variable affine, il est alors possible de transformer $\sqrt{ax^2+bx+c}$ en l'une des formes $\sqrt{1-u^2}$, $\sqrt{u^2+1}$ ou $\sqrt{u^2-1}$. On réalise ensuite les changements de variable respectifs $u = \sin t$, $u = \mathrm{sh} t$, $u = \pm\mathrm{ch} t$ et on sait alors parvenir à une fonction rationnelle.

Exemple Déterminons $\displaystyle\int \sqrt{3+2x-x^2}\,\mathrm{d}x$.
On réécrit le trinôme $3+2x-x^2$ sous forme canonique
$3+2x-x^2 = 4-(x-1)^2$.
En posant $x-1 = 2u$, $\sqrt{3+2x-x^2}$ est transformé en $\sqrt{4-4u^2} = 2\sqrt{1-u^2}$
Procédons au changement de variable correspondant

$$\int \sqrt{3+2x-x^2}\,\mathrm{d}x = 4\int \sqrt{1-u^2}\,\mathrm{d}u$$

Posons ensuite $u = \sin t$ avec $t \in [-\pi/2, \pi/2]$ de sorte que $t = \arcsin x$.

$$\int \sqrt{1-u^2}\,\mathrm{d}u = \int \cos^2 t\,\mathrm{d}t = \frac{1}{2}\int \cos(2t)+1\,\mathrm{d}t = \frac{1}{4}\sin(2t) + \frac{1}{2}t$$

On en déduit

$$\int \sqrt{3+2x-x^2}\,\mathrm{d}x = \frac{1}{2}(x-1)\sqrt{3+2x-x^2} + 2\arcsin\frac{x-1}{2}$$

Exemple Déterminons $\displaystyle\int \frac{\mathrm{d}x}{\sqrt{x^2+x+1}}$ sur $\mathbb{R}$
On réécrit le trinôme x^2+x+1 sous forme canonique

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

En posant $x + \frac{1}{2} = \frac{\sqrt{3}}{2}u$, $\sqrt{x^2+x+1}$ est transformé en $\frac{\sqrt{3}}{2}\sqrt{u^2+1}$.
Procédons au changement de variable correspondant

$$\int \frac{\mathrm{d}x}{\sqrt{x^2+x+1}} = \int \frac{\mathrm{d}u}{\sqrt{u^2+1}}$$

Posons ensuite $u = \mathrm{sh}t$ soit encore $t = \mathrm{argsh}x$.

$$\int \frac{\mathrm{d}u}{\sqrt{u^2+1}} = \int \mathrm{d}t = t = \mathrm{argsh}u$$

Par l'expression logarithmique de la fonction argsh et en rappelant qu'une primitive n'est déterminée qu'à une constante près, on parvient à

$$\int \frac{\mathrm{d}x}{\sqrt{x^2+x+1}} = \ln(2x+1+2\sqrt{x^2+x+1})$$

Exemple Déterminons $\int \sqrt{x^2+x}\,\mathrm{d}x$ sur $\mathbb{R}^+$.
On réécrit le trinôme $x^2 + x$ sous forme canonique

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

En posant $x + \frac{1}{2} = \frac{1}{2}u$, $\sqrt{x^2+x}$ est transformé en $\frac{1}{2}\sqrt{u^2-1}$.
Procédons au changement de variable correspondant

$$\int \sqrt{x^2+x}\,\mathrm{d}x = \frac{1}{4}\int \sqrt{u^2-1}\,\mathrm{d}u$$

Pour $x \in \mathbb{R}^+$, $u = 2x+1 \in [1, +\infty[$, on peut donc écrire $u = \mathrm{ch}t$ avec $t = \arg\mathrm{ch}u$.

$$\int \sqrt{u^2-1}\,\mathrm{d}u = \int \mathrm{sh}^2 t\,\mathrm{d}t = \frac{1}{4}\int \mathrm{e}^{2t} - 2 + \mathrm{e}^{-2t} = \frac{1}{4}\mathrm{sh}(2t) - \frac{1}{2}t$$

Ainsi

$$\int \sqrt{x^2+x}\,\mathrm{d}x = \frac{1}{16}\mathrm{sh}(2\mathrm{argch}(2x+1)) - \frac{1}{8}\mathrm{argch}(2x+1)$$

Par l'expression logarithmique de la fonction $\arg\mathrm{ch}$, on a

$$\mathrm{argch}(2x+1) = \ln(2x+1+2\sqrt{x^2+x})$$

De plus sachant $\mathrm{ch}(\mathrm{argch}(2x+1)) = 2x+1$ et $\mathrm{sh}(\mathrm{argch}(2x+1)) = 2\sqrt{x^2+x}$, on obtient finalement

$$\int \sqrt{x^2+x}\,\mathrm{d}x = \frac{1}{4}(2x+1)\sqrt{x^2+x} - \frac{1}{8}\ln(2x+1+2\sqrt{x^2+x})$$

Pour déterminer un primitive sur $]-\infty, -1]$ de la fonction $x \mapsto \sqrt{x^2+x}$, on reprend les calculs qui précèdent mais on pose $u = -\mathrm{ch}t$ car dans le cas présent $u = 2x+1 \in]-\infty, -1]$.

Chapitre 21

Développements limités

I désigne un intervalle non singulier et n un entier naturel.
Les fonctions considérées ici sont à valeurs réelles ou complexes.

21.1 Développements limités

Soit a un point de I ou une extrémité finie de I et $\mathcal{D} = I$ ou $\mathcal{D} = I \setminus \{a\}$.

21.1.1 Définition

Définition

On dit que $f : \mathcal{D} \to \mathbb{C}$ admet un développement limité à l'ordre n en a (en abrégé $DL_n(a)$) s'il existe $a_0, a_1, \ldots, a_n \in \mathbb{C}$ tels que quand $x \to a$,

$$f(x) = a_0 + a_1(x-a) + \cdots + a_n(x-a)^n + o((x-a)^n)$$

La fonction polynomiale $x \mapsto a_0 + a_1(x-a) + \cdots + a_n(x-a)^n$ est alors appelée partie régulière du $DL_n(a)$ de f en a.

Remarque Dans l'écriture

$$f(x) = a_0 + a_1(x-a) + \cdots + a_n(x-a)^n + o((x-a)^n)$$

chaque terme de la somme est négligeable devant celui qui le précède.

Remarque Un $DL_n(a)$ donne une information sur le comportement de f en a et uniquement en a.

Exemple Un développement limité de f à l'ordre 0 en a est de la forme $f(x) = a_0 + o(1)$.
Par suite f admet un développement limité à l'ordre 0 en a si, et seulement si, f converge en a.
De plus si telle est le cas, ce développement est de la forme $f(x) = a_0 + o(1)$ avec $a_0 = \lim\limits_a f$

Exemple Si $f : I \to \mathbb{R}$ est dérivable en $a \in I$ alors on a

$$f(x) = f(a) + f'(a)(x-a) + o(x-a)$$

Cette relation est un développement limité à l'ordre 1 en a.

Exemple $DL_n(0)$ de $x \mapsto \dfrac{1}{1-x}$
On a
$$1 + x + \cdots + x^n = \frac{1 - x^{n+1}}{1-x}$$
donc
$$\frac{1}{1-x} = 1 + x + \cdots + x^n + \frac{x^{n+1}}{1-x}$$
Or
$$\frac{x^{n+1}}{1-x} = x^n \frac{x}{1-x} = o(x^n)$$
car $\dfrac{x}{1-x} \xrightarrow[x \to 0]{} 0$.
Ainsi
$$\frac{1}{1-x} = 1 + x + \cdots + x^n + o(x^n)$$
ce qui est développement limité à l'ordre n en 0.

Exemple Soit $P(x) = a_0 + a_1 x + \cdots + a_p x^p$ une fonction polynomiale.
Pour $n \leqslant p$, on a $P(x) = a_0 + a_1 x + \cdots + a_n x^n + o(x^n)$.
Pour $n > p$, on a $P(x) = a_0 + a_1 x + \cdots + a_p x^p + o(x^n)$.

Proposition

Si $f : \mathcal{D} \to \mathbb{C}$ admet un $DL_n(a)$ de la forme :
$$f(x) = a_0 + a_1(x-a) + \cdots + a_n(x-a)^n + o((x-a)^n)$$
alors, pour tout $m \leqslant n$, f admet un $DL_m(a)$ s'obtenant par troncature :
$$f(x) = a_0 + a_1(x-a) + \cdots + a_m(x-a)^m + o((x-a)^m)$$

dém. :
via
□

$$a_{m+1}(x-a)^{m+1} + \cdots + a_n(x-a)^n + o((x-a)^n) = o((x-a)^m)$$

21.1.2 Unicité

Théorème

Si $f : \mathcal{D} \to \mathbb{C}$ admet un $DL_n(a)$ alors celui-ci est unique.

dém. :
Par l'absurde, supposons que f admet deux développements limités à l'ordre n en a distincts :
□

$$f(x) = a_0 + a_1(x-a) + ... + a_n(x-a)^n + o((x-a)^n)$$

et

$$f(x) = b_0 + b_1(x-a) + ... + b_n(x-a)^n + o((x-a)^n)$$

Soit $m \in \mathbb{N}$ le plus petit entier tel que $a_m \neq b_m$.
Par troncature, on a les deux relations

$$f(x) = a_0 + a_1(x-a) + ... + a_m(x-a)^m + o((x-a)^m)$$

et

$$f(x) = b_0 + b_1(x-a) + ... + b_m(x-a)^m + o((x-a)^m)$$

En simplifiant, on obtient

$$(a_m - b_m)(x-a)^m = o((x-a)^m)$$

Ainsi

$$(a_m - b_m)(x-a)^m = (x-a)^m \varepsilon(x)$$

avec $\varepsilon \underset{a}{\to} 0$.
En simplifiant par $(x-a)^m$ pour $x \neq a$, on obtient $a_m - b_m = \varepsilon(x)$
En passant à la limite quand $x \to a$ (avec $x \neq a$), on conclut $a_m - b_m = 0$.
C'est absurde.

Corollaire

Supposons I symétrique par rapport à 0 et $\mathcal{D} = I$ ou $\mathcal{D} = I \setminus \{0\}$.
Si $f : \mathcal{D} \to \mathbb{C}$ admet un développement limité à l'ordre n en 0 et si f est une fonction paire (resp. impaire) alors la partie régulière de f ne contient que des termes d'exposants pairs (resp. impairs).

dém. :
Cas f paire :
Quand $x \to 0$, on a
□

$$f(x) = a_0 + a_1 x + \cdots + a_n x^n + o(x^n)$$

En substituant $-x$ à x, on obtient

$$f(-x) = a_0 - a_1 x + \cdots + (-1)^n a_n x^n + o(x^n)$$

Or f est paire et donc

$$f(x) = f(-x) = a_0 - a_1 x + \cdots + (-1)^n a_n x^n + o(x^n)$$

Par unicité du développement limité à l'ordre n de f en 0, on obtient

$$\forall k \in \{0, 1, \dots, n\}, a_k = (-1)^k a_k$$

En particulier pour les indices k impairs, on obtient $a_k = 0$.
Cas f impaire : semblable.

21.1.3 Existence

Théorème

Soient $f : I \to \mathbb{C}$ et $a \in I$.
Si f est de classe $\mathcal{C}^n$ alors f admet un développement limité à l'ordre n en a de la forme :

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + o((x-a)^n)$$

dém. :
Par la formule de Taylor avec reste-intégrale, on a pour tout $x \in I$
□

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!}(x-a)^k + \int_a^x \frac{(x-t)^{n-1}}{(n-1)!} f^{(n)}(t)\,\mathrm{d}t$$

La fonction $f^{(n)}$ étant continue, on peut écrire $f^{(n)}(t) = f^{(n)}(a) + \varepsilon(t)$ avec $\varepsilon \xrightarrow[a]{} 0$.
On a alors

$$\int_a^x \frac{(x-t)^{n-1}}{(n-1)!} f^{(n)}(t)\,\mathrm{d}t = \int_a^x \frac{(x-t)^{n-1}}{(n-1)!} f^{(n)}(a)\,\mathrm{d}t + \int_a^x \frac{(x-t)^{n-1}}{(n-1)!} \varepsilon(t)\,\mathrm{d}t$$

D'une part

$$\int_a^x \frac{(x-t)^{n-1}}{(n-1)!} f^{(n)}(a)\,\mathrm{d}t = \frac{(x-a)^n}{n!} f^{(n)}(a)$$

D'autre part

$$\left|\int_a^x \frac{(x-t)^{n-1}}{(n-1)!} \varepsilon(t)\,\mathrm{d}t\right| \leqslant \sup_{t \in [a,x]} |\varepsilon(t)| \left|\int_a^x \frac{(x-t)^{n-1}}{(n-1)!}\,\mathrm{d}t\right| = |x-a|^n \sup_{t \in [a,x]} |\varepsilon(t)|$$

avec $[a, x]$ désignant le segment d'extrémités a et x que a soit inférieur à x ou l'inverse.
Quand $x \to a$, $\sup\limits_{t \in [a,x]} |\varepsilon(t)| \to 0$ et donc

$$\int_a^x \frac{(x-t)^{n-1}}{(n-1)!} \varepsilon(t)\,\mathrm{d}t = o\left((x-a)^n\right)$$

Par suite

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + o((x-a)^n)$$

Corollaire

> Si f est de classe $\mathcal{C}^\infty$ sur I alors f admet un développement limité en tout point de I et à tout ordre.

(!)Ce résultat ne peut être utilisé que pour une fonction définie et régulière en a.

21.1.4 DL de référence

Exemple $DL_n(0)$ de $x \mapsto \dfrac{1}{1-x}$.
On a déjà vu la relation

$$\begin{aligned}\frac{1}{1-x} &= \sum_{k=0}^{n} x^k + o(x^n)\\ &= 1 + x + \cdots + x^n + o(x^n)\end{aligned}$$

Exemple $DL_n(0)$ de $x \mapsto \dfrac{1}{1+x}$.
En remplaçant x par $-x$, la relation précédente donne

$$\begin{aligned}\frac{1}{1+x} &= \sum_{k=0}^{n} (-1)^k x^k + o(x^n)\\ &= 1 - x + x^2 + \cdots + (-1)^n x^n + o(x^n)\end{aligned}$$

Exemple $DL_n(0)$ de $x \mapsto \mathrm{e}^x$.
La fonction $f : x \mapsto \mathrm{e}^x$ est $\mathcal{C}^\infty$ avec $f^{(k)}(x) = \mathrm{e}^x$ et donc $f^{(k)}(0) = 1$.
Par la formule de Taylor-Young

$$\begin{aligned}\mathrm{e}^x &= \sum_{k=0}^{n} \frac{1}{k!} x^k + o(x^n)\\ &= 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \cdots + \frac{1}{n!}x^n + o(x^n)\end{aligned}$$

En particulier

$$\mathrm{e}^x = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \frac{1}{24}x^4 + o(x^4)$$

Exemple $DL_n(0)$ de $x \mapsto \mathrm{e}^{\lambda x}$ avec $\lambda \in \mathbb{C}$.
En reprenant la même démarche avec $x \mapsto \mathrm{e}^{\lambda x}$, on obtient

$$\mathrm{e}^{\lambda x} = \sum_{k=0}^{n} \frac{\lambda^k}{k!} x^k + o(x^n)$$

Exemple $DL_{2n+1}(0)$ de $x \mapsto \cos x$.
Puisque $\cos x = \mathrm{Re}(\mathrm{e}^{ix})$ et $\mathrm{e}^{ix} = \sum_{k=0}^{2n+1} \frac{i^k}{k!} x^k + o(x^{2n+1})$, en séparant la somme en deux sommes formés respectivement des termes d'indices pairs et impairs, on obtient

$$\begin{aligned}\cos x &= \sum_{k=0}^{n} \frac{(-1)^k}{(2k)!} x^{2k} + o(x^{2n+1}) \\ &= 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 + \cdots + \frac{(-1)^n}{(2n)!}x^{2n} + o(x^{2n+1})\end{aligned}$$

En particulier

$$\cos x = 1 - \frac{1}{2}x^2 + \frac{1}{24}x^4 + o(x^4)$$

Exemple $DL_{2n+2}(0)$ de $x \mapsto \sin x$.
Puisque $\sin x = \mathrm{Re}(\mathrm{e}^{ix})$, la démarche ci-dessus donne

$$\begin{aligned}\sin x &= \sum_{k=0}^{n} \frac{(-1)^k}{(2k+1)!} x^{2k+1} + o(x^{2n+1}) \\ &= x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 + \cdots + \frac{(-1)^n}{(2n+1)!}x^{2n+1} + o(x^{2n+2})\end{aligned}$$

En particulier

$$\sin x = x - \frac{1}{6}x^3 + o(x^4)$$

Exemple $DL_{2n+1}(0)$ de $x \mapsto \mathrm{ch}x$.
Puisque $\mathrm{ch}x = \frac{1}{2}\left(\mathrm{e}^x + \mathrm{e}^{-x}\right)$, on obtient en sommant les développements de e^x et de e^{-x}

$$\begin{aligned}\mathrm{ch}x &= \sum_{k=0}^{n} \frac{1}{(2k)!} x^{2k} + o(x^{2n+1}) \\ &= 1 + \frac{1}{2!}x^2 + \frac{1}{4!}x^4 + \cdots + \frac{1}{(2n)!}x^{2n} + o(x^{2n+1})\end{aligned}$$

Exemple $DL_{2n+2}(0)$ de $x \mapsto \mathrm{sh}x$.
Puisque $\mathrm{sh}x = \frac{1}{2}\left(\mathrm{e}^x - \mathrm{e}^{-x}\right)$, en procédant comme ci-dessus :

$$\begin{aligned}\mathrm{sh}x &= \sum_{k=0}^{n} \frac{1}{(2k+1)!} x^{2k+1} + o(x^{2n+2}) \\ &= x + \frac{1}{3!}x^3 + \frac{1}{5!}x^5 + \cdots + \frac{1}{(2n+1)!}x^{2n+1} + o(x^{2n+2})\end{aligned}$$

Exemple $DL_n(0)$ de $x \mapsto \ln(1+x)$.
La fonction $f : x \mapsto \ln(1+x)$ est de classe $\mathcal{C}^\infty$ sur $]-1,+\infty[$ avec $f'(x) = \dfrac{1}{1+x}$,. . . ,

$$f^{(k)}(x) = \frac{(-1)^{k-1}(k-1)!}{(1+x)^k}$$

Puisque $f(0) = 0$ et pour tout $k \in \mathbb{N}^\star f^{(k)}(0) = (-1)^k(k-1)!$, la formule de Taylor-Young donne

$$\begin{aligned}\ln(1+x) &= \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k} x^k + o(x^n) \\ &= x - \frac{1}{2}x^2 + \cdots + \frac{(-1)^{n-1}}{n} x^n + o(x^n)\end{aligned}$$

En particulier $\ln(1+x) = x - \dfrac{1}{2}x^2 + \dfrac{1}{3}x^3 + o(x^3)$.

Exemple $DL_n(0)$ de $x \mapsto (1+x)^\alpha$ avec $\alpha \in \mathbb{R}$ fixé.
La fonction $f : x \mapsto (1+x)^\alpha$ est de classe $\mathcal{C}^\infty$ sur $]-1,+\infty[$ avec $f'(x) = \alpha(1+x)^{\alpha-1}$, $f''(x) = \alpha(\alpha-1)(1+x)^{\alpha-2}$,. . . ,

$$f^{(k)}(x) = \alpha(\alpha-1)...(\alpha-k+1)(1+x)^{\alpha-k}$$

On alors $f^{(k)}(0) = \alpha(\alpha-1)...(\alpha-k+1)$ pour tout $k \in \mathbb{N}$, et la formule de Taylor-Young donne

$$(1+x)^\alpha = 1 + \alpha.x + \frac{\alpha(\alpha-1)}{2!}x^2 + \cdots + \frac{\alpha(\alpha-1)...(\alpha-n+1)}{n!}x^n + o(x^n)$$

Remarque Pour $\alpha = p \in \mathbb{N}$:

$$(1+x)^p = \binom{p}{0} + \binom{p}{1}x + \binom{p}{2}x^2 + \cdots + \binom{p}{n}x^n + o(x^n)$$

La formule correspond à la troncature du développement du binôme à l'ordre n.

Remarque On note souvent :

$$\binom{\alpha}{k} = \frac{\alpha(\alpha-1)\dots(\alpha-k+1)}{k!}$$

pour $k \in \mathbb{N}$ et $\alpha \in \mathbb{R}$.
Ainsi

$$\begin{aligned}(1+x)^\alpha &= \sum_{k=0}^{n} \binom{\alpha}{k} x^k + o(x^n) \\ &= \binom{\alpha}{0} + \binom{\alpha}{1}x + \binom{\alpha}{2}x^2 + \cdots + \binom{\alpha}{n}x^n + o(x^n)\end{aligned}$$

Exemple Pour $\alpha = -1$, $\binom{\alpha}{k} = \dfrac{(-1)(-2)...(-k)}{k!} = (-1)^k$ et on retrouve

$$\frac{1}{1+x} = 1 - x + x^2 + \cdots + (-1)^n x^n + o(x^n)$$

Pour $\alpha = 1/2$, $(1+x)^\alpha = 1 + \alpha x + \dfrac{\alpha(\alpha-1)}{2}x^2 + o(x^2)$ donne

$$\sqrt{1+x} = 1 + \frac{1}{2}x - \frac{1}{8}x^2 + o(x^2)$$

Pour $\alpha = -1/2$, on obtient

$$\frac{1}{\sqrt{1+x}} = 1 - \frac{1}{2}x + \frac{3}{8}x^2 + \mathrm{o}(x^2)$$

(!)La formule donnant le développement limité de $(1+x)^\alpha$ ne peut s'appliquer que pour un exposant $\alpha \in \mathbb{R}$ fixé. En particulier, elle ne s'applique pas à une expression du type $(1+x)^x$.

21.2 Détermination de développements limités

21.2.1 Positionnement du problème en 0

Pour déterminer un développement limité en a d'une fonction $x \mapsto f(x)$, on relocalise le problème en 0 via le changement de variable $x = a + h$. On détermine alors un développement limité en 0 de la fonction $h \mapsto f(a+h)$ puis on transpose ce transpose en développement limité en a en remplaçant h par $x - a$.

Exemple $DL_2(1)$ de $x \mapsto \mathrm{e}^x$.
Quand $x \to 1$
$x = 1 + h$, $h = x - 1$ avec $h \to 0$.

$$\mathrm{e}^x = \mathrm{e}^{1+h} = \mathrm{e}.\mathrm{e}^h = \mathrm{e}(1 + h + \frac{1}{2}h^2 + o(h^2)) = \mathrm{e} + \mathrm{e}.h + \frac{\mathrm{e}}{2}h^2 + o(h^2)$$

puis

$$\mathrm{e}^x = \mathrm{e} + \mathrm{e}(x-1) + \frac{\mathrm{e}}{2}(x-1)^2 + o((x-1)^2)$$

(!)Ne pas développer cette expression car sinon on perd la visualisation des ordres de grandeurs au voisinage de 1.

Exemple $DL_3(\pi/3)$ de $x \mapsto \cos x$.
Quand $x \to \pi/3$
$x = \pi/3 + h$, $h = \pi/3 - x$ avec $h \to 0$.

$$\cos x = \cos\left(\frac{\pi}{3} + h\right) = \frac{1}{2}\cos h - \frac{\sqrt{3}}{2}\sin h$$

Or $\cos h = 1 - \dfrac{1}{2}h^2 + o(h^3)$ et $\sin h = h - \dfrac{1}{6}h^3 + o(h^3)$ donc

$$\cos x = \frac{1}{2} + \frac{\sqrt{3}}{2}h - \frac{1}{4}h^2 - \frac{\sqrt{3}}{12}h^3 + o(h^3)$$

puis

$$\cos x = \frac{1}{2} + \frac{\sqrt{3}}{2}\left(x - \frac{\pi}{3}\right) - \frac{1}{4}\left(x - \frac{\pi}{3}\right)^2 - \frac{\sqrt{3}}{12}\left(x - \frac{\pi}{3}\right)^3 + o\left(\left(x - \frac{\pi}{3}\right)^3\right)$$

Exemple $DL_2(1)$ de $x \mapsto \ln x$.
Quand $x \to 1$
$x = 1 + h$, $h = x - 1$ avec $h \to 0$.

$$\ln x = \ln(1+h) = h - \frac{1}{2}h^2 + o(h^2) = (x-1) - \frac{1}{2}(x-1)^2 + o((x-1)^2)$$

Exemple $DL_2(2)$ de $x \mapsto \sqrt{x}$.
Quand $x \to 2$
$x = 2 + h$, $h = x - 2$ avec $h \to 0$.
$\sqrt{x} = \sqrt{2+h} = \sqrt{2}\sqrt{1+h/2}$.
Or $\sqrt{1+u} = 1 + \frac{1}{2}u - \frac{1}{8}u^2 + o(u^2)$ quand $u \to 0$
donc pour $u = h/2 \to 0$ on a $\sqrt{1+h/2} = 1 + \frac{1}{4}h - \frac{1}{32}h^2 + o(h^2)$,
puis

$$\sqrt{x} = \sqrt{2} + \frac{\sqrt{2}}{4}(x-2) - \frac{\sqrt{2}}{32}(x-2)^2 + o((x-2)^2)$$

(!)Ici le changement de variable $x = 1 + h$ eut été inadapté, puisque quand $x \to 2$, on a $h \to 1$ et non $h \to 0$.

21.2.2 DL d'un produit

Supposons qu'au voisinage de 0
$f(x) = a_0 + a_1x + \cdots + a_nx^n + o(x^n)$ et $g(x) = b_0 + b_1x + \cdots + b_nx^n + o(x^n)$.
On a $f(x)g(x) = a_0b_0 + (a_0b_1 + a_1b_0)x + \cdots + (a_0b_n + \cdots + a_nb_0)x^n + o(x^n)$
ce qui détermine un $DL_n(0)$ de $x \mapsto f(x)g(x)$.

Exemple $DL_3(0)$ de $x \mapsto \dfrac{\mathrm{e}^x}{1-x}$.
Quand $x \to 0$

$$\mathrm{e}^x = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + o(x^3) \text{ et } \frac{1}{1-x} = 1 + x + x^2 + x^3 + o(x^3)$$

donc

$$\frac{\mathrm{e}^x}{1-x} = 1 + (1+1)x + \left(1 + 1 + \frac{1}{2}\right)x^2 + \left(1 + 1 + \frac{1}{2} + \frac{1}{6}\right)x^3 + o(x^3)$$

puis

$$\frac{\mathrm{e}^x}{1-x} = 1 + 2x + \frac{5}{2}x^2 + \frac{8}{3}x^3 + o(x^3)$$

Exemple $DL_4(0)$ de $x \mapsto \cos x \text{ch} x$.
Quand $x \to 0$

$$\cos x = 1 - \frac{1}{2}x^2 + \frac{1}{24}x^4 + o(x^4) \text{ et } \text{ch} x = 1 + \frac{1}{2}x^2 + \frac{1}{24}x^4 + o(x^4)$$

donc

$$\cos x \text{ch} x = 1 + \left(\frac{1}{2} - \frac{1}{2}\right)x^2 + \left(\frac{1}{24} - \frac{1}{4} + \frac{1}{24}\right)x^4 + o(x^4)$$

puis

$$\cos x \text{ch} x = 1 - \frac{1}{6}x^4 + o(x^4)$$

Exemple $DL_3(0)$ de $x \mapsto \ln(1+x)\mathrm{e}^x$.
Puisque le développement de $\ln(1+x)$ commence par le terme x, un développement à l'ordre 2 de e^x suffit à mener les calculs. En effet, par multiplication par $\ln(1+x)$, le terme $o(x^2)$ du développement de e^x devient $o(x^3)$.
Quand $x \to 0$

$$\ln(1+x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 + o(x^3) \text{ et } \mathrm{e}^x = 1 + x + \frac{1}{2}x^2 + o(x^2)$$

donc

$$\ln(1+x)\mathrm{e}^x = x + \left(1 - \frac{1}{2}\right)x^2 + \left(\frac{1}{2} - \frac{1}{2} + \frac{1}{3}\right)x^3 + o(x^3)$$

puis

$$\ln(1+x)\mathrm{e}^x = x + \frac{1}{2}x^2 + \frac{1}{3}x^3 + o(x^3)$$

Exemple $DL_4(0)$ de $x \mapsto \ln(1+x)(1-\cos x)$.
Les développements de $\ln(1+x)$ commençant par x, un développement à l'ordre 3 de $1-\cos x$ est suffisant pour mener les calculs.
Aussi, le développement de $1-\cos x$ commençant par un terme x^2, un développement à l'ordre 2 de $\ln(1+x)$ est suffisant pour mener les calculs.
Quand $x \to 0$

$$\ln(1+x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 + o(x^3) \text{ et } 1 - \cos x = \frac{1}{2}x^2 + o(x^3)$$

donc

$$\ln(1+x)(1-\cos x) = \frac{1}{2}x^3 - \frac{1}{4}x^4 + o(x^4)$$

21.2.3 DL d'une composée.

Supposons $f(x) \xrightarrow[x\to 0]{} 0$ et supposons que $g(u) = a_0 + a_1 u + \cdots + a_n u^n + o(u^n)$ quand $u \to 0$.
Puisqu'on peut écrire $o(u^n) = u^n \varepsilon(u)$ avec $\varepsilon \underset{0}{\to} 0$, on a si la composition est possible

$$g(f(x)) = a_0 + a_1 f(x) + \cdots + a_n (f(x))^n + (f(x))^n \varepsilon(f(x))$$

avec $\varepsilon(f(x)) \underset{x\to 0}{\to} 0$
Ceci permet alors d'écrire :

$$g(f(x)) = a_0 + a_1 f(x) + \cdots + a_n (f(x))^n + o((f(x))^n)$$

Ainsi on a pu substituer $f(x)$ à u dans le $DL_n(0)$ de $g(u)$ et cela a été possible car $f(x) \xrightarrow[x\to 0]{} 0$
Si l'on connaît alors un développement limité de f, on peut en déduire un développement limité de $g(f(x))$.

Exemple $DL_3(0)$ de $x \mapsto \mathrm{e}^{x+x^2}$.
Quand $x \to 0$,
$\mathrm{e}^{x+x^2} = \mathrm{e}^u$ avec
$u = x + x^2 \to 0$.
Nous sommes amenés à réaliser un $DL(0)$ de $u \mapsto \mathrm{e}^u$.
Commençons car calculer les développement limites de puissances de u à la précision $o(x^3)$.
$u^2 = x^2 + 2x^3 + o(x^3)$,
$u^3 = x^3 + o(x^3)$
et $o(u^3) = o(x^3)$.
Un développement limité à l'ordre 3 de e^u peut alors être transformé en développement limité à l'ordre 3 en x.

$$\mathrm{e}^u = 1 + u + \frac{1}{2}u^2 + \frac{1}{6}u^3 + o(u^3)$$

donc

$$\mathrm{e}^{x+x^2} = 1 + x + \left(1 + \frac{1}{2}\right)x^2 + \left(1 + \frac{1}{6}\right)x^3 + o(x^3)$$

puis

$$\mathrm{e}^{x+x^2} = 1 + x + \frac{3}{2}x^2 + \frac{7}{6}x^3 + o(x^3)$$

Exemple $DL_6(0)$ de $x \mapsto \ln(1 + x^2 + x^3)$
Quand $x \to 0$,
$\ln(1 + x^2 + x^3) = \ln(1 + u)$ avec
$u = x^2 + x^3 \to 0$,
$u^2 = x^4 + 2x^5 + x^6$,
$u^3 = x^6 + o(x^6)$
et $o(u^3) = o(x^6)$.
Un développement limité à l'ordre de $\ln(1 + u)$ convient.

$$\ln(1 + u) = u - \frac{1}{2}u^2 + \frac{1}{3}u^3 + o(u^3)$$

donc
$$\ln(1+x^2+x^3) = x^2+x^3-\frac{1}{2}x^4-x^5-\frac{1}{6}x^6+o(x^6)$$

Exemple $DL_3(0)$ de $x \mapsto \ln(1+\sin x)$.
Quand $x \to 0$,
$$\ln(1+\sin x) = \ln(1+x-\frac{1}{6}x^3+o(x^3)) = \ln(1+u)$$
avec
$u = x-\frac{1}{6}x^3+o(x^3) \to 0$,
$u^2 = x^2+o(x^3)$,
$u^3 = x^3+o(x^3)$
et $o(u^3) = o(x^3)$.
$$\ln(1+u) = u-\frac{1}{2}u^2+\frac{1}{3}u^3+o(u^3)$$
donc
$$\ln(1+\sin x) = x-\frac{1}{2}x^2+\frac{1}{6}x^3+o(x^3)$$

Exemple $DL_2(0)$ de $x \mapsto \sqrt{\cos x}$.
Quand $x \to 0$,
$$\sqrt{\cos x} = \sqrt{1-\frac{1}{2}x^2+o(x^2)} = \sqrt{1+u}$$
avec
$u = -\frac{1}{2}x^2+o(x^2) \to 0$ et $o(u) = o(x^2)$.
$$\sqrt{1+u} = 1+\frac{1}{2}u+o(u)$$
donc
$$\sqrt{\cos x} = 1-\frac{1}{4}x^2+o(x^2)$$

Exemple $DL_3(0)$ de $x \mapsto \mathrm{e}^{\frac{1}{1+x}}$.
Quand $x \to 0$,
$$\mathrm{e}^{\frac{1}{1+x}} = \mathrm{e}^{1-x+x^2-x^3+o(x^3)} = \mathrm{e}.\mathrm{e}^u$$
avec
$u = -x+x^2-x^3+o(x^3) \to 0$,
$u^2 = x^2-2x^3+o(x^3)$,
$u^3 = x^3+o(x^3)$
et $o(u^3) = o(x^3)$.
$$\mathrm{e}^u = 1+u+\frac{1}{2}u^2+\frac{1}{6}u^3+o(u^3)$$
donc
$$\mathrm{e}^{\frac{1}{1+x}} = \mathrm{e}-\mathrm{e}x+\frac{3\mathrm{e}}{2}x^2-\frac{13\mathrm{e}}{6}x^3+o(x^3)$$

Exemple $DL_2(0)$ de $x \mapsto \sqrt{1+\mathrm{e}^x}$.
Quand $x \to 0$,

$$\sqrt{1+\mathrm{e}^x} = \sqrt{2+x+\frac{1}{2}x^2+o(x^2)} = \sqrt{2}\sqrt{1+\frac{1}{2}x+\frac{1}{4}x^2+o(x^2)} = \sqrt{2}\sqrt{1+u}$$

avec
$u = \frac{1}{2}x+\frac{1}{4}x^2+o(x^2) \to 0$,
$u^2 = \frac{1}{4}x^2+o(x^2)$
et $o(u^2) = o(x^2)$.

$$\sqrt{1+u} = 1+\frac{1}{2}u-\frac{1}{8}u^2+o(u^2)$$

donc

$$\sqrt{1+\mathrm{e}^x} = \sqrt{2}+\frac{\sqrt{2}}{4}x+\frac{3\sqrt{2}}{32}x^2+o(x^2)$$

21.2.4 DL d'un inverse

Supposons : $f(x) = a_0+a_1x+...+a_nx^n+o(x^n)$ avec $a_0 \neq 0$ au voisinage de 0. En écrivant

$$\frac{1}{f(x)} = \frac{1}{a_0}.\frac{1}{1+\left(\frac{a_1}{a_0}x+...+\frac{a_n}{a_0}x^n+o(x^n)\right)} = \frac{1}{a_0}.\frac{1}{1+u}$$

on peut former un développement limité à l'ordre n de f en 0…

Exemple $DL_3(0)$ de $x \mapsto \frac{1}{1+\mathrm{e}^x}$
Quand $x \to 0$,

$$\frac{1}{1+\mathrm{e}^x} = \frac{1}{2+x+\frac{1}{2}x^2+\frac{1}{6}x^3+o(x^3)} = \frac{1}{2}\frac{1}{1+u}$$

avec
$u = \frac{1}{2}x+\frac{1}{4}x^2+\frac{1}{6}x^3+o(x^3) \to 0$,
$u^2 = \frac{1}{4}x^2+\frac{1}{4}x^3+o(x^3)$,
$u^3 = \frac{1}{8}x^3+o(x^3)$
et $o(u^3) = o(x^3)$.

$$\frac{1}{1+u} = 1-u+u^2-u^3+o(u^3)$$

donne

$$\frac{1}{1+\mathrm{e}^x} = \frac{1}{2}-\frac{1}{4}x+\frac{1}{48}x^3+o(x^3)$$

Exemple $DL_4(0)$ de $x \mapsto \frac{1}{\cos x}$
Quand $x \to 0$,

$$\frac{1}{\cos x} = \frac{1}{1-\frac{1}{2}x^2+\frac{1}{24}x^4+o(x^4)} = \frac{1}{1+u}$$

avec
$u = -\dfrac{1}{2}x^2 + \dfrac{1}{24}x^4 + o(x^4) \to 0$,
$u^2 = \dfrac{1}{4}x^4 + o(x^4)$ et $o(u^2) = o(x^4)$

$$\frac{1}{1+u} = 1 - u + u^2 + o(u^2)$$

donne

$$\frac{1}{\cos x} = 1 + \frac{1}{2}x^2 + \frac{5}{24}x^4 + o(x^4)$$

Exemple $DL_5(0)$ de $x \mapsto \tan x$

$$\tan x = \frac{\sin x}{\cos x} = \sin x \frac{1}{\cos x}$$

Puisque le développement limité de $\sin x$ commence par x, un développement limité à l'ordre 4 de $1/\cos x$ suffit à poursuivre les calculs.
Quand $x \to 0$,

$$\frac{1}{\cos x} = \frac{1}{1 - \frac{1}{2}x^2 + \frac{1}{24}x^4 + o(x^4)} = \frac{1}{1-u}$$

avec
$u = \dfrac{1}{2}x^2 - \dfrac{1}{24}x^4 + o(x^4) \to 0$.
$u^2 = \dfrac{1}{4}x^4 + o(x^4)$
et $o(u^2) = o(x^4)$.

$$\frac{1}{1-u} = 1 + u + u^2 + o(u^2)$$

donne

$$\frac{1}{\cos x} = 1 + \frac{1}{2}x^2 + \frac{5}{24}x^4 + o(x^4)$$

puis

$$\tan x = \left(x - \frac{1}{6}x^3 + \frac{1}{120}x^5 + o(x^5)\right)\left(1 + \frac{1}{2}x^2 + \frac{5}{24}x^4 + o(x^4)\right)$$

En développant, on obtient

$$\tan x = x + \frac{1}{3}x^3 + \frac{2}{15}x^5 + o(x^5)$$

21.2.5 DL délicats

Lors des calculs, des divisions peuvent réduire l'ordre d'un développement limité. En anticipant celles-ci, on peut éviter de devoir reprendre un calcul initié avec des développements trop courts.

Exemple $DL_3(0)$ de $x \mapsto (\cos x)^{1/x}$
$(\cos x)^{1/x} = \mathrm{e}^{\frac{1}{x}\ln(\cos x)}$
Pour former le développement limité voulu, on développe à l'ordre 3 l'expression $\dfrac{1}{x}\ln(\cos x)$.

Puisque la division par x, réduit l'ordre d'un développement limité, nous allons former un développement limité à l'ordre 4 de $\ln(\cos x)$.

$$\ln(\cos x) = \ln\left(1 - \frac{1}{2}x^2 + \frac{1}{24}x^4 + o(x^4)\right) = \ln(1+u)$$

avec

$$u = -\frac{1}{2}x^2 + \frac{1}{24}x^4 + o(x^4)$$

Par composition de développements limités

$$\ln(\cos x) = -\frac{1}{2}x^2 - \frac{1}{12}x^4 + o(x^4)$$

puis

$$\frac{1}{x}\ln(\cos x) = -\frac{1}{2}x - \frac{1}{12}x^3 + o(x^3)$$

$(\cos x)^{1/x} = \mathrm{e}^u$ avec $u = -\frac{1}{2}x - \frac{1}{12}x^3 + o(x^3)$.
Par composition de développements limités, on obtient

$$(\cos x)^{1/x} = 1 - \frac{1}{2}x + \frac{1}{8}x^2 - \frac{5}{48}x^3 + o(x^3)$$

Exemple $DL_2(0)$ de $x \mapsto \left(\frac{\sin x}{x}\right)^{3/x^2}$.

$$\left(\frac{\sin x}{x}\right)^{3/x^2} = \exp\left(\frac{3}{x^2}\ln\left(\frac{\sin x}{x}\right)\right)$$

A cause de la division par x^2, on forme un développement limité à l'ordre 4 du terme $\ln\frac{\sin x}{x}$.
A cause de la division par x, on part d'un développement limité à l'ordre 5 du terme $\sin x$.

$$\sin x = x - \frac{1}{6}x^3 + \frac{1}{120}x^5 + o(x^5)$$

donc

$$\ln\frac{\sin x}{x} = \ln\left(1 - \frac{1}{6}x^2 + \frac{1}{120}x^4 + o(x^4)\right)$$

Par composition

$$\ln\frac{\sin x}{x} = -\frac{1}{6}x^2 - \frac{1}{180}x^4 + o(x^4)$$

On en déduit

$$\frac{3}{x^2}\ln\left(\frac{\sin x}{x}\right) = -\frac{1}{2} - \frac{1}{60}x^2 + o(x^2)$$

puis

$$\left(\frac{\sin x}{x}\right)^{3/x^2} = \exp\left(-\frac{1}{2} - \frac{1}{60}x^2 + o(x^2)\right) = \frac{1}{\sqrt{2}}\mathrm{e}^{-\frac{1}{60}x^2 + o(x^2)} = \frac{1}{\sqrt{\mathrm{e}}} - \frac{1}{60\sqrt{\mathrm{e}}}x^2 + o(x^2)$$

Exemple $DL_2(0)$ de $x \mapsto \dfrac{\ln(1+x)-x}{\mathrm{ch}x-1}$.
Les développements limités de $\ln(1+x)-x$ et $\mathrm{ch}x-1$ commençant par un terme en x^2, une division par x^2 aura lieu durant les calculs. On initie donc ceci avec des développements limités à l'ordre 4.

$$\frac{\ln(1+x)-x}{\mathrm{ch}x-1} = \frac{-\frac{1}{2}x^2+\frac{1}{3}x^3-\frac{1}{4}x^4+o(x^4)}{\frac{1}{2}x^2+\frac{1}{24}x^4+o(x^4)} = \frac{-1+\frac{2}{3}x-\frac{1}{2}x^2+o(x^2)}{1+\frac{1}{12}x^2+o(x^2)}$$

puis

$$\frac{\ln(1+x)-x}{\mathrm{ch}x-1} = \left(-1+\frac{2}{3}x-\frac{1}{2}x^2+o(x^2)\right)\left(1-\frac{1}{12}x^2+o(x^2)\right)$$

et en développant

$$\frac{\ln(1+x)-x}{\mathrm{ch}x-1} = -1+\frac{2}{3}x-\frac{5}{12}x^2+o(x^2)$$

21.2.6 Intégration et non dérivation de DL

Théorème

Soit $f : I \to \mathbb{C}$ dérivable et $a \in I$.
Si f' admet un développement limité à l'ordre n en a de la forme :

$$f'(x) = a_0 + a_1(x-a) + \cdots + a_n(x-a)^n + o\left((x-a)^n\right)$$

alors f admet un développement limité à l'ordre $n+1$ en a de la forme :

$$f(x) = f(a) + a_0(x-a) + \cdots + \frac{a_n}{n+1}(x-a)^{n+1} + o\left((x-a)^{n+1}\right)$$

dém. :
Considérons la fonction $\varphi : I \to \mathbb{R}$ définie par
□

$$\varphi(x) = f(x) - \left(f(a) + a_0(x-a) + \cdots + \frac{a_n}{n+1}(x-a)^{n+1}\right)$$

La fonction φ est dérivable, $\varphi(a) = 0$ et $\varphi'(x) = o((x-a)^n) = (x-a)^n\varepsilon(x)$ avec $\varepsilon \xrightarrow[a]{} 0$.
Par l'inégalité des accroissements finis

$$|\varphi(x)| = |\varphi(x) - \varphi(a)| \leqslant |x-a| \sup_{t\in[a,x]} |\varphi'(t)|$$

en notant $[a, x]$ le segment d'extrémités a et x que a soit inférieur à x ou non.
Or

$$\sup_{t\in[a,x]} |\varphi'(t)| \leqslant |x-a|^n \sup_{t\in[a,x]} |\varepsilon(t)|$$

et puisque $\varepsilon \underset{a}{\to} 0$, on a $\sup_{t\in[a,x]} |\varepsilon(t)| \to 0$.
Ainsi $\varphi(x) = o\left((x-a)^{n+1}\right)$
puis l'on obtient la relation

$$f(x) = f(a) + a_0(x-a) + \cdots + \frac{a_n}{n+1}(x-a)^{n+1} + o\left((x-a)^{n+1}\right)$$

(!)Ne pas oublier le terme $f(a)$ lors de cette « intégration ».

Exemple On peut retrouver le développement limité de $x \mapsto \ln(1+x)$ par cet outil.
En effet $(\ln(1+x))' = \dfrac{1}{1+x}$
Or, quand $x \to 0$,

$$\frac{1}{1+x} = 1 - x + x^2 + \cdots + (-1)^n x^n + o(x^n)$$

donc par intégration de ce développement limité et sachant $\ln(1) = 0$, on obtient

$$\ln(1+x) = x - \frac{1}{2}x^2 + \cdots + \frac{(-1)^n}{n+1}x^{n+1} + o(x^{n+1})$$

Exemple $DL_{2n+1}(0)$ de $x \mapsto \arctan x$
La fonction $x \mapsto \arctan x$ est dérivable et $(\arctan x)' = \dfrac{1}{1+x^2}$.
Quand $x \to 0$,

$$\frac{1}{1+x^2} = 1 - x^2 + x^4 + \cdots + (-1)^n x^{2n} + o(x^{2n})$$

donc par intégration de ce développement limité et sachant $\arctan(0) = 0$, on obtient

$$\arctan x = x - \frac{1}{3}x^3 + \frac{1}{5}x^5 + \cdots + \frac{(-1)^n}{2n+1}x^{2n+1} + o(x^{2n+1}) = \sum_{k=0}^{n} \frac{(-1)^k}{2k+1}x^{2k+1} + o(x^{2n+1})$$

Exemple $DL_5(0)$ de $x \mapsto \arccos x$:
La fonction $x \mapsto \arccos x$ est dérivable et $(\arccos x)' = -\dfrac{1}{\sqrt{1-x^2}}$
Quand $x \to 0$,

$$-\frac{1}{\sqrt{1-x^2}} = -(1 + \frac{1}{2}x^2 + \frac{3}{8}x^4 + o(x^4))$$

donc par intégration de ce développement limité et sachant $\arccos(0) = \dfrac{\pi}{2}$, on obtient
puis

$$\arccos x = \frac{\pi}{2} - x - \frac{1}{6}x^3 - \frac{3}{40}x^5 + o(x^5)$$

(!)On peut intégrer un développement limité mais on ne peut pas dériver un développement limité. Plus précisément, il se peut que la fonction dérivée ne possède pas de développement limité.

Exemple Soient $n \in \mathbb{N}^\star$ et

$$f(x) = \begin{cases} 1 + x + x^2 \cos(1/x) & \text{si } x \neq 0 \\ 1 & \text{sinon} \end{cases}$$

Quand $x \to 0$, on a $f(x) = 1 + x + o(x)$ et ainsi f admet un développement limité à l'ordre 1 en 0.
f est dérivable sur $\mathbb{R}^\star$ et

$$f'(x) = 1 + 2x^2 \cos(1/x) + \sin(1/x)$$

Aussi, par limite de taux de variation, la fonction f est dérivable en 0 et $f'(0) = 1$.
Ainsi la fonction f est dérivable sur $\mathbb{R}$.
Cependant la dérivée de f n'admet pas de limite en 0 et donc a fortiori n'admet pas de développement limité en 0.
f' n'a pas de limite en 0 donc n'y admet pas de DL.

21.2.7 DL de solution d'équation différentielle

On peut parfois former des développements limités de solutions d'équations différentielle :
- on justifie l'existence du développement limité par le théorème de Taylor-Young ;
- on exprime ce développement limité avec des coefficients inconnus ;
- on détermine ceux-ci en exploitant l'équation différentielle.

Exemple Formons $DL_7(0)$ de la fonction $x \mapsto \tan x$ sachant qu'elle est solution de l'équation différentielle

$$y' = 1 + y^2$$

La fonction $\tan$ est de classe $\mathcal{C}^\infty$ et impaire donc admet un $DL_7(0)$ de la forme :

$$\tan x = ax + bx^3 + cx^5 + dx^7 + o(x^7)$$

Comme la fonction $(\tan)'$ est $\mathcal{C}^\infty$ et paire, $(\tan)'$ admet un $DL_6(0)$ de la forme :

$$(\tan x)' = \alpha + \beta x^2 + \gamma x^4 + \delta x^6 + o(x^6)$$

En intégrant celui-ci :

$$\tan x = \alpha x + \frac{\beta}{3}x^3 + \frac{\gamma}{5}x^5 + \frac{\delta}{7}x^7 + o(x^7)$$

Par unicité du DL : $\alpha = a$, $\beta = 3b$, $\gamma = 5c$ et $\delta = 7d$.
Ainsi

$$(\tan x)' = a + 3bx^2 + 5cx^4 + 7dx^6 + o(x^6)$$

La relation $(\tan)' = 1 + \tan^2$ donne alors

$$a + 3bx^2 + 5cx^4 + 7dx^6 + o(x^6) = a^2x^2 + 2abx^4 + (b^2 + 2ax)x^6 + o(x^6)$$

Par unicité des développements limité, on obtient un système en les inconnus a, b, c, d dont la solution est

$$a = 1, b = \frac{1}{3}, c = \frac{2}{15}, d = \frac{17}{315}$$

Remarque Ici pour obtenir le développement limité de y', on n'a pas procéder par dérivation mais par intégration connaissant l'existence du développement de y'.

21.2.8 DL d'application réciproque

On peut parfois former des développements limités de bijections réciproques :
- on justifie l'existence du développement limité par le théorème de Taylor-Young ;
- on exprime ce développement limité avec des coefficients inconnus ;
- on forme un développement limité de $f^{-1}(f(x)) = x$ (ou, selon les cas, de $f(f^{-1}(y)) = y$) ;
- par unicité des développements limités, on forme un système déterminant les coefficients inconnus de départ.

Exemple Soit $f : \mathbb{R} \to \mathbb{R}$ définie par

$$f(x) = x\text{ch}x$$

La fonction f est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$ et $f'(x) = x\text{sh}x + \text{ch}x \geqslant \text{ch}x \geqslant 1 > 0$.
La fonction f réalise donc une bijection de $\mathbb{R}$ vers $\mathbb{R}$ et son application réciproque f^{-1} est de classe $\mathcal{C}^\infty$.
Réalisons un développement limité à l'ordre 5 de f^{-1} en 0.
Puisque f^{-1} est de classe $\mathcal{C}^\infty$, elle admet un $DL_5(0)$.
De plus, f est impaire, donc f^{-1} aussi et le $DL_5(0)$ de f^{-1} est de la forme

$$f^{-1}(y) = ay + by^3 + cy^5 + o(y^5)$$

Pour déterminer les coefficients a, b, c, exploitons la relation $f^{-1}(f(x)) = x$
Quand $x \to 0$,
$f^{-1}(f(x)) = f^{-1}(x\text{ch}x) = f^{-1}(y)$ avec
$y = x + \frac{1}{2}x^3 + \frac{1}{24}x^5 + o(x^5) \to 0$.
$y^3 = x^3 + \frac{3}{2}x^5 + o(x^5)$,
$y^5 = x^5 + o(x^5)$ et $o(y^5) = o(x^5)$.
Par composition

$$f^{-1}(f(x)) = ax + \left(\frac{a}{2} + b\right) x^3 + \left(\frac{a}{24} + \frac{3}{2}b + c\right) x^5 + o(x^5) = x$$

Par unicité des développements limités on obtient le système

$$\begin{cases} a = 1 \\ a/2 + b = 0 \\ a/24 + 3b/2 + c = 0 \end{cases}$$

et on en déduit

$$a = 1, b = -\frac{1}{2} \text{ et } c = \frac{17}{24}$$

21.3 Notion de développements asymptotiques

21.3.1 Définition

Définition

Soient $a \in I$ ou une extrémité, éventuellement infinie de l'intervalle I et $f : \mathcal{D} \to \mathbb{C}$ avec $\mathcal{D} = I$ ou $\mathcal{D} = I \setminus \{a\}$.
On appelle développement asymptotique de f en a toute écriture

$$f(x) = a_0 g_0(x) + a_1 g_1(x) + \cdots + a_n g_n(x) + o(g_n(x)) \text{ quand } x \to a$$

avec $a_0, \ldots, a_n \in \mathbb{C}$ et $g_0, \ldots, g_n$ des fonctions « simples »telles que

$$g_n \ll g_{n-1} \ll \cdots \ll g_0$$

Un tel développement asymptotique est alors dit réalisé à la précision $g_n(x)$.

Remarque Dans un développement asymptotique, chaque terme de la somme est négligeable devant le précédent.

Remarque Un développement asymptotique de f en a donne des informations sur le comportement de f en a.
En particulier le premier terme non nul donne la limite de f en a.

Exemple Un $DL_n(a)$ est un développement asymptotique à la précision $(x-a)^n$.

21.3.2 Détermination

Pour forme un développement asymptotique, on exploite les techniques calculatoires des développements limité sans se limiter à l'obtention de terme de la forme $(x-a)^n$…

Exemple Développement asymptotique de $x \mapsto \ln(1+\sqrt{x})$ à la précision x^2 en 0.
On a $\ln(1+\sqrt{x}) = \ln(1+u)$ avec $u = \sqrt{x}$.
Quand $x \to 0$, on a $u \to 0$.
$u = \sqrt{x}$, $u^2 = x$, $u^3 = x\sqrt{x}$, $u^4 = x^2$ et $o(u^4) = o(x^2)$.
Or

$$\ln(1+u) = u - \frac{1}{2}u^2 + \frac{1}{3}u^3 - \frac{1}{4}u^4 + o(u^4)$$

donc par composition

$$\ln(1+\sqrt{x}) = \sqrt{x} - \frac{1}{2}x + \frac{1}{3}x\sqrt{x} - \frac{1}{4}x^2 + o(x^2)$$

Exemple Développement asymptotique de $x \mapsto \dfrac{1}{\mathrm{e}^x - 1}$ à 3 termes en 0.
Quand $x \to 0$,

$$\frac{1}{\mathrm{e}^x - 1} = \frac{1}{x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + o(x^3)} = \frac{1}{x}\frac{1}{1 + \frac{1}{2}x + \frac{1}{6}x^2 + o(x^2)}$$

Or

$$\frac{1}{1+\frac{1}{2}x+\frac{1}{6}x^2+o(x^2)} = \frac{1}{1+u} \text{ avec } u = \frac{1}{2}x+\frac{1}{6}x^2+o(x^2)$$

donc

$$\frac{1}{1+\frac{1}{2}x+\frac{1}{6}x^2+o(x^2)} = 1-\frac{1}{2}x+\frac{1}{12}x^2+o(x^2)$$

puis

$$\frac{1}{\mathrm{e}^x-1} = \frac{1}{x}-\frac{1}{2}+\frac{1}{12}x+o(x)$$

Exemple Développement asymptotique de $x \mapsto \ln(x+1)$ à 3 termes en $+\infty$.
Quand $x \to +\infty$,

$$\ln(x+1) = \ln x + \ln(1+u)$$

avec $u = 1/x \to 0$.
Or

$$\ln(1+u) = u - \frac{1}{2}u^2 + o(u^2)$$

donc

$$\ln(x+1) = \ln x + \frac{1}{x} - \frac{1}{2x^2} + o\left(\frac{1}{x^2}\right)$$

Exemple Développement asymptotique de $x \mapsto \sqrt{x+1}$ à 3 termes en $+\infty$.
Quand $x \to +\infty$,

$$\sqrt{x+1} = \sqrt{x}\sqrt{1+u}$$

avec $u = 1/x \to 0$.
Or

$$\sqrt{1+u} = 1+\frac{1}{2}u-\frac{1}{8}u^2+o(u^2)$$

donc

$$\sqrt{x+1} = \sqrt{x} + \frac{1}{2\sqrt{x}} - \frac{1}{8x^{3/2}} + o\left(\frac{1}{x^{3/2}}\right)$$

Exemple Développement asymptotique de $x \mapsto \dfrac{1}{\sqrt{x^2+x+1}}$ à 3 termes en $+\infty$.
Quand $x \to +\infty$,

$$\frac{1}{\sqrt{x^2+x+1}} = \frac{1}{x}\frac{1}{\sqrt{1+u}}$$

avec $u = \dfrac{1}{x}+\dfrac{1}{x^2} \to 0$.
Or

$$\frac{1}{\sqrt{1+u}} = 1-\frac{1}{2}u+\frac{3}{8}u^2+o(u^2)$$

donc

$$\frac{1}{\sqrt{x^2+x+1}} = \frac{1}{x}-\frac{1}{2}-\frac{1}{8x}+o\left(\frac{1}{x}\right)$$

Exemple Développement asymptotique à trois termes de $x \mapsto \arctan x$ en $+\infty$.
On sait que pour tout $x > 0$,

$$\arctan x + \arctan \frac{1}{x} = \frac{\pi}{2}$$

On en déduit

$$\arctan x = \frac{\pi}{2} - \arctan u$$

avec $u = \dfrac{1}{x} \to 0$.
Or

$$\arctan u = u - \frac{1}{3}u^3 + o(u^3)$$

donc

$$\arctan x = \frac{\pi}{2} - \arctan \frac{1}{x} = \frac{\pi}{2} - \frac{1}{x} + \frac{1}{3x^3} + o\left(\frac{1}{x^3}\right)$$

21.3.3 Développement asymptotique de suites

La notion de développement asymptotique s'étend aux suites, un développement asymptotique d'une suite est la décomposition de son terme général en somme de « termes simples »ordonnés par négligeabilité croissante.
L'obtention d'un développement asymptotique d'une suite à partir de son terme général se fait en exploitant les démarches calculatoires qui précèdent.

Exemple Développement asymptotique à trois termes de $\sqrt{n^2+1}$.

$$\sqrt{n^2+1} = n\sqrt{1+u}$$

avec $u = 1/n^2 \to 0$.
Puisque

$$\sqrt{1+u} = 1 + \frac{1}{2}u - \frac{1}{8}u^2 + o(u^2)$$

on obtient

$$\sqrt{n^2+1} = 1 + \frac{1}{2n^2} - \frac{1}{8n^4} + o\left(\frac{1}{n^4}\right)$$

Exemple Développement asymptotique à trois termes de $n\ln(n+1)$.

$$n\ln(n+1) = n\ln n + n\ln(1+u)$$

avec $u = 1/n \to 0$.
Puisque

$$\ln(1+u) = u - \frac{1}{2}u^2 + o(u^2)$$

on obtient

$$n\ln(n+1) = n\ln n + 1 - \frac{1}{2n} + o\left(\frac{1}{n}\right)$$

Exemple Développement asymptotique à trois termes de $\sqrt[n]{n}$.

$$\sqrt[n]{n} = n^{1/n} = \mathrm{e}^{\frac{1}{n}\ln n} = \mathrm{e}^{u}$$

avec $u = \dfrac{1}{n}\ln n \to 0$.
Puisque

$$\mathrm{e}^{u} = 1 + u + \frac{1}{2}u^2 + o(u^2)$$

on obtient

$$\sqrt[n]{n} = 1 + \frac{\ln n}{n} + \frac{1}{2}\left(\frac{\ln n}{n}\right)^2 + o\left(\left(\frac{\ln n}{n}\right)^2\right)$$

Remarque Si le terme général de la suite n'est pas explicite, les termes d'un développement asymptotique peuvent parfois s'obtenir les uns après les autres...

Exemple Développement asymptotique à la précision $\dfrac{1}{n}$ de $u_n = \displaystyle\int_0^1 \frac{x\,\mathrm{d}x}{1+x^n}$.

$$\left|u_n - \frac{1}{2}\right| = \left|u_n - \int_0^1 x\,\mathrm{d}x\right| \leqslant \int_0^1 \left|\frac{x^{n+1}}{1+x^n}\right|\,\mathrm{d}x \leqslant \int_0^1 x^{n+1}\,\mathrm{d}x = \frac{1}{n+2} \to 0$$

Ainsi $u_n \to \dfrac{1}{2}$ ce qui permet d'écrire $u_n = \dfrac{1}{2} + o(1)$, c'est un développement asymptotique à un terme.
Posons $v_n = u_n - \dfrac{1}{2} \to 0$ de sorte que $u_n = \dfrac{1}{2} + v_n$ et cherchons un équivalent simple de v_n.

$$\begin{aligned} v_n &= -\int_0^1 \frac{x^{n+1}}{1+x^n}\,\mathrm{d}x \\ &= -\frac{1}{n}\left[x^2\ln(1+x^n)\right]_0^1 + \frac{1}{n}\int_0^1 2x\ln(1+x^n)\,\mathrm{d}x \\ &= -\frac{1}{n}\ln 2 + \frac{1}{n}\int_0^1 2x\ln(1+x^n)\,\mathrm{d}x \end{aligned}$$

Puisque

$$\left|\int_0^1 2x\ln(1+x^n)\,\mathrm{d}x\right| = \int_0^1 2x\ln(1+x^n)\,\mathrm{d}x \leqslant \int_0^1 2x^{n+1}\,\mathrm{d}x = \frac{2}{n+2} \to 0$$

On a

$$\frac{1}{n}\int_0^1 2x\ln(1+x^n)\,\mathrm{d}x = o\left(\frac{1}{n}\right)$$

puis

$$v_n = -\frac{\ln 2}{n} + o\left(\frac{1}{n}\right)$$

et enfin

$$u_n = \frac{1}{2} - \frac{\ln 2}{n} + o\left(\frac{1}{n}\right)$$

Exemple Pour $n \in \mathbb{N}$, étudions l'équation $\tan x = x$ d'inconnue $x \in]-\pi/2 + n\pi, \pi/2 + n\pi[= I_n$.
Considérons la fonction $\varphi : x \mapsto \tan x - x$ définie sur I_n.
φ est dérivable et $\varphi'(x) = \tan^2 x > 0$ sauf pour $x = n\pi$.
Ainsi φ est strictement croissante et l'étude des limite de φ assure que φ réalise une bijection de I_n vers $\mathbb{R}$. Par suite, l'équation $\tan x = x$ admet une unique solution $x_n \in I_n$
Formons un développement asymptotique à trois termes de x_n.
On a

$$\underbrace{-\pi/2 + n\pi}_{\sim n\pi} < x_n < \underbrace{\pi/2 + n\pi}_{\sim n\pi}$$

donc $x_n \sim n\pi$.
Ainsi on peut écrire $x_n = n\pi + o(n)$.
Considérons maintenant

$$y_n = x_n - n\pi$$

On a $y_n \in]-\pi/2, \pi/2[$ et $\tan y_n = \tan x_n = x_n$ donc $y_n = \arctan x_n \to \pi/2$.
Ainsi

$$y_n = \frac{\pi}{2} + o(1)$$

puis

$$x_n = n\pi + \frac{\pi}{2} + o(1)$$

Considérons maintenant

$$z_n = x_n - n\pi - \pi/2 = y_n - \pi/2$$

On a

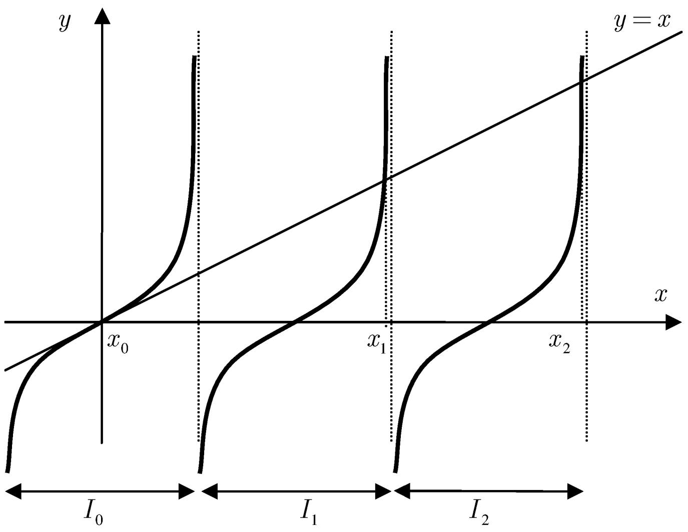
$$z_n = \frac{\pi}{2} - \arctan x_n = -\arctan \frac{1}{x_n}$$

donc $z_n \sim -\dfrac{1}{n\pi}$
Ainsi

$$z_n = -\frac{1}{n\pi} + o\left(\frac{1}{n}\right)$$

et enfin

$$x_n = n\pi + \frac{\pi}{2} - \frac{1}{n\pi} + o\left(\frac{1}{n}\right)$$



21.4 Applications des développements limités et asymptotiques

21.4.1 Détermination d'équivalents

Le premier terme non nul d'un développement limité ou d'un développement asymptotique fournit un équivalent simple de la fonction étudiée au point considéré.

Exemple A partir des développements limités des fonctions usuelles, on obtient les équivalents de référence suivants :
Quand $x \to 0$, $\sin x \sim x$, $\tan x \sim x$, $\ln(1+x) \sim x$, $\mathrm{e}^x - 1 \sim x$, $1 - \cos x \sim \dfrac{x^2}{2}$, $(1+x)^\alpha - 1 \sim \alpha x$.

Exemple Déterminons un équivalent simple de $\tan x - \arctan x$ quand $x \to 0$.
On a $\tan x = x + \dfrac{1}{3}x^3 + o(x^3)$ et $\arctan x = x - \dfrac{1}{3}x^3 + o(x^3)$ donc

$$\tan x - \arctan x = \frac{2}{3}x^3 + o(x^3) \underset{0}{\sim} \frac{2}{3}x^3$$

Exemple Déterminons un équivalent simple de $x^x - x$ quand $x \to 1$.
Il importe de commencer par se repositionner en 0.
On pose $x = 1 + h$, avec $h = x - 1 \to 0$.

$$x^x - x = (1+h)^{1+h} - (1+h) = \mathrm{e}^{(1+h)\ln(1+h)} - (1+h)$$

Or

$$\mathrm{e}^{(1+h)\ln(1+h)} = \mathrm{e}^{(1+h)\left(h - \frac{1}{2}h^2 + o(h^2)\right)} = \mathrm{e}^{h + \frac{1}{2}h^2 + o(h^2)} = 1 + h + h^2 + o(h^2)$$

donc

$$x^x - x = h^2 + o(h^2) \sim h^2 = (x-1)^2$$

Exemple Déterminons un équivalent simple de $\frac{1}{x} - \cot x$ quand $x \to 0$.

$$\cot x = \frac{\cos x}{\sin x} = \frac{1}{x}\frac{1 - \frac{1}{2}x^2 + o(x^2)}{1 - \frac{1}{6}x^2 + o(x^2)} = \frac{1}{x} - \frac{1}{3}x + o(x)$$

donc

$$\frac{1}{x} - \cot x \sim -\frac{1}{3}x$$

Exemple Déterminons un équivalent simple de $\frac{1}{\sqrt{x+1}} - \frac{1}{\sqrt{x-1}}$ quand $x \to +\infty$.

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

et

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

donc

$$\frac{1}{\sqrt{x+1}} - \frac{1}{\sqrt{x-1}} \sim -\frac{1}{x\sqrt{x}}$$

Exemple Déterminons un équivalent simple $\ln(n+1) - \ln(n-1)$.

$$\ln(n+1) = \ln n + \ln(1 + 1/n) = \ln n + \frac{1}{n} + o\left(\frac{1}{n}\right)$$

et

$$\ln(n-1) = \ln n - \frac{1}{n} + o\left(\frac{1}{n}\right)$$

donc

$$\ln(n+1) - \ln(n-1) \sim \frac{2}{n}$$

21.4.2 Détermination de limite

L'obtention d'un équivalent permet d'obtenir la limite de la fonction considérée.

Exemple Etudions

$$\lim_{x\to 0} \frac{\tan 2x - 2\tan x}{\sin 2x - 2\sin x}$$

Quand $x \to 0$,

$\tan x = x + \frac{1}{3}x^3 + o(x^3)$ donc $\tan 2x - 2\tan x \sim 2x^3$

$\sin x = x - \frac{1}{6}x^3 + o(x^3)$ donc $\sin 2x - 2\sin x \sim -x^3$

Par suite

$$\frac{\tan 2x - 2\tan x}{\sin 2x - 2\sin x} \sim \frac{2x^3}{-x^3} \to -2$$

Exemple Etudions

$$\lim_{x\to 0}\left(\frac{1}{x^2}-\frac{1}{\tan^2 x}\right)$$

Quand $x \to 0$,

$$\frac{1}{x^2}-\frac{1}{\tan^2 x}=\frac{\tan^2 x-x^2}{x^2\tan^2 x}=\frac{(\tan x-x)(\tan x+x)}{x^2\tan^2 x}\sim\frac{\frac{1}{3}x^3\times 2x}{x^4}\to\frac{2}{3}$$

Exemple Etudions

$$\lim_{x\to+\infty}\left(\cos\frac{1}{x}\right)^{x^2}$$

Quand $x \to +\infty$

$$\left(\cos\frac{1}{x}\right)^{x^2}=\exp\left(x^2\ln\cos\frac{1}{x}\right)$$

Or

$$\cos\frac{1}{x}=1-\frac{1}{2x^2}+o\left(\frac{1}{x^2}\right)$$

donc $\ln\cos\frac{1}{x}\sim-\frac{1}{2x^2}$ puis $x^2\ln\cos\frac{1}{x}\to-\frac{1}{2}$ et enfin

$$\left(\cos\frac{1}{x}\right)^{x^2}\to\frac{1}{\sqrt{\mathrm{e}}}$$

Exemple Etudions

$$\lim_{n\to+\infty}\left(\frac{\sqrt[n]{a}+\sqrt[n]{b}}{2}\right)^n \text{ pour } a,b>0$$

Quand $n \to +\infty$,

$$\sqrt[n]{a}=\mathrm{e}^{\frac{1}{n}\ln a}=1+\frac{1}{n}\ln a+o\left(\frac{1}{n}\right)$$

donc

$$\frac{\sqrt[n]{a}+\sqrt[n]{b}}{2}=1+\frac{1}{n}\ln\sqrt{ab}+o\left(\frac{1}{n}\right)$$

puis

$$\left(\frac{\sqrt[n]{a}+\sqrt[n]{b}}{2}\right)^n=\exp\left(n\ln\left(1+\frac{1}{n}\ln\sqrt{ab}+o\left(\frac{1}{n}\right)\right)\right)$$

Or

$$n\ln\left(1+\frac{1}{n}\ln\sqrt{ab}+o\left(\frac{1}{n}\right)\right)\to\ln\sqrt{ab}$$

donc

$$\left(\frac{\sqrt[n]{a}+\sqrt[n]{b}}{2}\right)^n\to\sqrt{ab}$$

21.4.3 Prolongement d'une fonction en un point

Proposition

Soient a un point de I ou une extrémité finie de I et $f : I \backslash \{a\} \to \mathbb{C}$.
Si f admet un $DL_1(a)$ de la forme

$$f(x) = a_0 + a_1(x-a) + o(x-a)$$

alors on peut prolonger f par continuité en a en posant $f(a) = a_0$ et ce prolongement est dérivable en a avec $f'(a) = a_1$.

dém. :
Quand $x \to a$, $f(x) = a_0 + o(1) \to a_0$ donc on prolonge f par continuité en a en posant $f(a) = a_0$.
De plus $\dfrac{f(x)-f(a)}{x-a} = a_1 + o(1) \to a_1$ et donc ce prolongement est dérivable avec $f'(a) = a_1$.
□
(!)En revanche, l'existence d'un développement limité à un ordre supérieur n'assure pas l'existence des dérivées d'ordre supérieur.

Exemple Montrons que la fonction $f : x \mapsto \dfrac{\sin x}{x}$ se prolonge en 0 en une fonction de classe $\mathcal{C}^1$ sur $\mathbb{R}$.
La fonction est évidemment définie et de classe $\mathcal{C}^1$ sur les intervalles $\mathbb{R}^\star$.
Quand $x \to 0$, $f(x) = 1 + o(x)$.
On peut donc prolonger f par continuité en 0 en posant $f(0) = 1$.
De plus, ce prolongement est dérivable et $f'(0) = 0$.
Etudions la continuité de f' en 0.
Pour $x \neq 0$,

$$f'(x) = \frac{x\cos x - \sin x}{x^2}$$

Quand $x \to 0$ (avec $x \neq 0$)

$$f'(x) = \frac{x + o(x^2) - x + o(x^2)}{x^2} = o(1) \to 0$$

Par suite f' est continue en 0.

21.4.4 Positionnement local d'une courbe et de sa tangente

Soient $f : I \to \mathbb{R}$ et $a \in I$.
On suppose que f admet un développement limité à l'ordre n en a de la forme :

$$f(x) = a_0 + a_1(x-a) + a_2(x-a)^2 + \cdots + a_n(x-a)^n + o((x-a)^n)$$

La fonction étant définie en a, on a nécessairement $a_0 = f(a)$.
Comme dans l'étude ci-dessus, on établit aussi que $a_1 = f'(a)$.
L'équation de la tangente T à f en a est alors $y = f(a) + f'(a)(x-a)$ ou encore

$$y = a_0 + a_1(x-a)$$

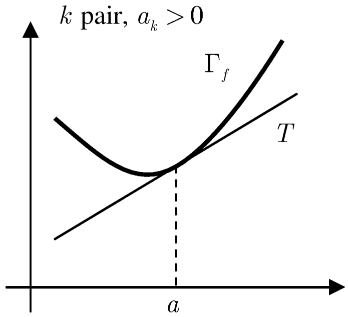
Ainsi les deux premiers termes du développement limité de f en a correspondent au second membre de l'équation de la tangente à f en a.
Etudions maintenant la position du graphe de f par rapport à T en a. Pour cela déterminons le signe de la quantité

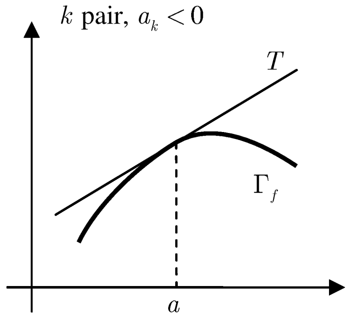
$$f(x) - (f(a) + f'(a)(x-a)) = a_2(x-a)^2 + ... + a_n(x-a)^n + o((x-a)^n)$$

quand $x \to a$.
Soit k le plus petit entier supérieur ou égal à 2 tel que $a_k \neq 0$.
On a

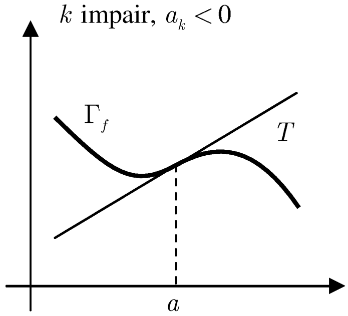
$$f(x) - (f(a) + f'(a)(x-a)) \sim a_k(x-a)^k$$

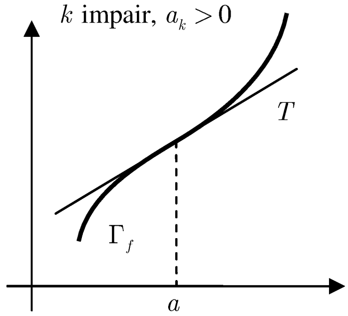
quand $x \to a$ et donc la quantité étudiée est du signe de $a_k(x-a)^k$ au voisinage de a.
Cas k pair
Le terme $(x-a)^k$ est toujours positif.
Si $a_k > 0$ alors le graphe de f est au dessus de T au voisinage de a.
Si $a_k < 0$ alors le graphe de f est en dessous de T au voisinage de a.





Cas k impair
Le terme $(x-a)^k$ s'annule en changeant de signe en a.
Dans ce cas le graphe de f traverse sa tangente T en a et le signe de a_k permet de préciser de quel côté de a, Γ_f est au dessus de T.





(!)Ces positionnements ne valent qu'au voisinage de a.

Exemple Soit $f : \mathbb{R}^\star \to \mathbb{R}$ définie par

$$f(x) = \frac{x}{\mathrm{e}^x - 1}$$

Quand $x \to 0$, on obtient

$$f(x) = 1 - \frac{1}{2}x + \frac{1}{12}x^2 + o(x^2)$$

On peut prolonger f par continuité en 0 en posant $f(0) = 1$.
Ce prolongement est dérivable et $f'(0) = -1/2$.
L'équation de la tangente à f en 0 est

$$y = 1 - \frac{1}{2}x$$

De plus $f(x) - (1 - \frac{1}{2}x) \sim \frac{1}{12}x^2 \geqslant 0$ quand $x \to 0$, donc la courbe est localement au dessus de sa tangente en 0.

Exemple Soit $f : \mathbb{R} \to \mathbb{R}$ définie par

$$f(x) = \ln(x^2 + 2x + 2)$$

Quand $x \to 0$, on obtient

$$f(x) = \ln 2 + x - \frac{1}{6}x^3 + o(x^3)$$

L'équation de la tangente à f en 0 est

$$y = \ln 2 + x$$

Puisque $f(x) - (\ln 2 + x) \sim -\frac{1}{6}x^3 + o(x^3)$ quand $x \to 0$, la courbe traverse sa tangente en 0 en étant d'abord au dessus puis en dessous.

21.4.5 Etude de droite asymptote

L'obtention d'un développement asymptotique en $+\infty$ de la forme $f(x) = ax + b + o(1)$ permet de conclure que la droite d'équation $y = ax + b$ est asymptote à f en $+\infty$.
De plus l'étude du signe du $o(1)$ permet de positionner la courbe par rapport à cette asymptote.

Exemple Etudions $f : x \mapsto \sqrt{x^2 + x + 1}$ en $+\infty$.
Après calcul, on obtient le développement asymptotique

$$\sqrt{x^2 + x + 1} = x + \frac{1}{2} + \frac{3}{8x} + o\left(\frac{1}{x}\right)$$

quand $x \to +\infty$.
La droite d'équation $y = x + \frac{1}{2}$ est asymptote à f en $+\infty$ et puisque $\frac{3}{8x} > 0$, la courbe est au dessus de celle-ci.

21.4.6 Etude locale des points d'une courbe paramétrée

On note $\mathcal{P}$ le plan géométrique muni d'un repère orthonormé $(O; \vec{i}, \vec{j})$.
Soit $\Gamma = (I, M)$ un arc de classe $\mathcal{C}^m$ avec m assez grand.
On étudie l'allure de l'arc Γ au point $M(t_0)$ avec $t_0 \in I$ fixé.
On note $x(t)$ et $y(t)$ les coordonnées du point $M(t)$.

21.4.6.1 Formule de Taylor-Young

Théorème

Pour tout $n \leqslant m$, on a

$$\overrightarrow{OM}(t_0+h) = \overrightarrow{OM}(t_0) + h\frac{\mathrm{d}\overrightarrow{OM}}{\mathrm{d}t}(t_0) + \cdots + \frac{h^n}{n!}\frac{\mathrm{d}^n\overrightarrow{OM}}{\mathrm{d}t^n}(t_0) + h^n\overrightarrow{\varepsilon(h)}$$

avec $\overrightarrow{\varepsilon(h)} \underset{0}{\to} \vec{o}$.

dém. :
Par les coordonnées des vecteurs.
□

Remarque Le terme $h^n\overrightarrow{\varepsilon(h)}$ est parfois aussi noté $\overrightarrow{o(h^n)}$.

21.4.6.2 Tangente

Théorème

S'il existe un plus petit entier $p \geqslant 1$ tel que

$$\frac{\mathrm{d}^p\overrightarrow{OM}}{\mathrm{d}t^p}(t_0) \neq \vec{0}$$

alors Γ admet une tangente en $M(t_0)$ dirigée par ce vecteur.

dém. :
Sous réserve d'existence, la tangente en $M(t_0)$ est la droite limite des droites pivotantes $(M(t_0)M(t))$ quand $t \to t_0$. Par la formule de Taylor-Young :
□

$$\overrightarrow{OM}(t_0+h) = \overrightarrow{OM_0} + \frac{h^p}{p!}\frac{\mathrm{d}^p\overrightarrow{OM}}{\mathrm{d}t^p}(t_0) + \overrightarrow{o\left(h^p\right)}$$

donc

$$\overrightarrow{M_0M}(t_0+h) = \frac{h^p}{p!}\frac{\mathrm{d}^p\overrightarrow{OM}}{\mathrm{d}t^p}(t_0) + \overrightarrow{o\left(h^p\right)}$$

La droite $(M_0M(t_0+h))$ est alors dirigée par

$$\vec{u}_t = \frac{p!}{h^p}\overrightarrow{M_0M}(t_0+h) = \frac{\mathrm{d}^p\overrightarrow{OM}}{\mathrm{d}t^p}(t_0) + \overrightarrow{o(1)} \to \frac{\mathrm{d}^p\overrightarrow{OM}}{\mathrm{d}t^p}(t_0) \neq \vec{0}$$

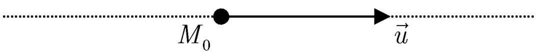
Remarque Si $\vec{v}(t_0) \neq \vec{0}$ (i.e. $M(t_0)$ régulier) alors $p = 1$ et on retrouve que la vitesse dirige la tangente lorsqu'elle n'est pas nulle.
Si $\vec{v}(t_0) = \vec{0}$ et $\vec{a}(t_0) \neq \vec{0}$ alors $p = 2$ et la tangente en un point stationnaire est dirigée par l'accélération lorsque cette dernière n'est pas nulle.

21.4.6.3 Position de la courbe par rapport à sa tangente

Soit $t_0 \in I$, on désire préciser l'allure de Γ au voisinage du point $M_0 = M(t_0)$.
Supposons qu'il existe un plus petit entier $p \geqslant 1$ tel que

$$\vec{u} = \frac{\mathrm{d}^p\overrightarrow{OM}}{\mathrm{d}t^p}(t_0) \neq \vec{0}$$

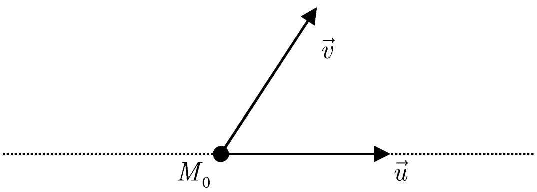
La tangente en M_0 existe et est dirigée par $\vec{u}$.



Exemple Si M_0 est régulier alors $p = 1$ et $\vec{u} = \vec{v}(t_0)$

Supposons de plus, qu'il existe un plus petit entier $q \geqslant p + 1$ tel que

$$\vec{v} = \frac{\mathrm{d}^q\overrightarrow{OM}}{\mathrm{d}t^q}(t_0) \text{ ne soit pas colinéaire à } \vec{u}$$



Exemple Si M_0 est birégulier $p = 1$, $q = 2$, $\vec{u} = \vec{v}(t_0)$ et $\vec{v} = \vec{a}(t_0)$.

Par la formule de Taylor-Young,

$$\overrightarrow{OM}(t_0 + h) = \overrightarrow{OM_0} + h\frac{\mathrm{d}\overrightarrow{OM}}{\mathrm{d}t}(t_0) + \cdots + \frac{h^q}{q!}\frac{\mathrm{d}^q\overrightarrow{OM}}{\mathrm{d}t^q}(t_0) + \overrightarrow{o(h^q)}$$

Par définition des entiers p et q on a
pour tout $1 \leqslant j < p$, $\dfrac{\mathrm{d}^j\overrightarrow{OM}}{\mathrm{d}t^j}(t_0) = \vec{0}$,
pour $j = p$, $\dfrac{\mathrm{d}^p\overrightarrow{OM}}{\mathrm{d}t^p}(t_0) = \vec{u}$,

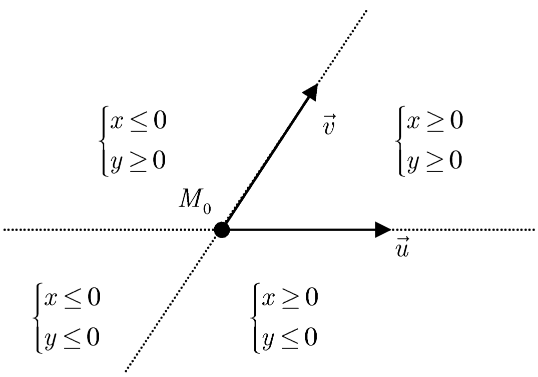
pour tout $p < j < q$, $\dfrac{\mathrm{d}^j\overrightarrow{OM}}{\mathrm{d}t^j}(t_0) = \lambda_j\vec{u}$,
et pour $j = q$, $\dfrac{\mathrm{d}^q\overrightarrow{OM}}{\mathrm{d}t^q}(t_0) = \vec{v}$.
On obtient alors

$$\overrightarrow{OM}(t_0+h) = \overrightarrow{OM_0} + \left(\frac{1}{p!}h^p + o(h^p)\right)\vec{u} + \frac{h^q}{q!}\vec{v} + \overrightarrow{o(h^q)}$$

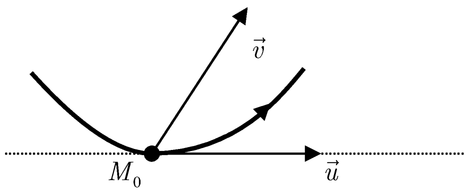
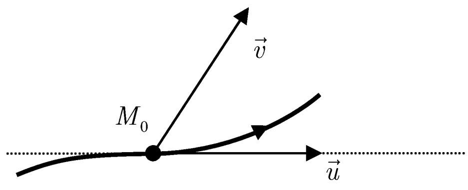
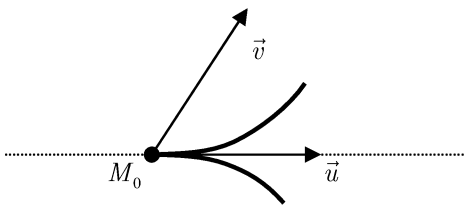
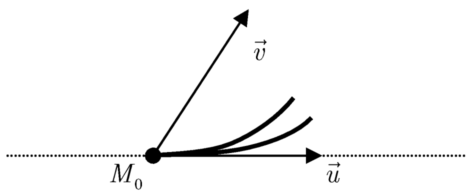
Or $\overrightarrow{o(h^q)} = o(h^q)\vec{u} + o(h^q)\vec{v}$ car les composantes de $\overrightarrow{o(h^q)}$ dans $\vec{u}, \vec{v}$ sont combinaisons linéaires des composantes dans la base $(\vec{i}, \vec{j})$ et ces dernières sont en $o(h^q)$.
On parvient alors à la relation

$$\overrightarrow{M_0M(t_0+h)} = \left(\frac{h^p}{p!} + o(h^p)\right)\vec{u} + \left(\frac{h^q}{q!} + o(h^q)\right)\vec{v}$$

L'étude du signe de $x = \dfrac{h^p}{p!} + o(h^p)$ et de $y = \dfrac{h^q}{q!} + o(h^q)$ permet de positionner le point $M(t_0+h)$ dans le repère oblique $(M_0; \vec{u}, \vec{v})$.



On peut alors affirmer les figures suivantes :

Cas p impair et q pair	Cas p impair et q impair
	
Point ordinaire	Point d'inflexion
Cas p pair et q impair	Cas p pair et q pair
	
Point de rebroussement de 1ère espèce	Point de rebroussement de 2nde espèce

Remarque Un point régulier est soit un point ordinaire, soit un point d'inflexion.
Un point birégulier est un point ordinaire.

Remarque Lorsque p et q sont pairs, il est impossible de préciser le sens de parcours sur les deux branches, il arrive même parfois que celles-ci se superposent.

21.4.6.4 Mise en pratique

1ère méthode
On calcule les vecteurs dérivés successifs

$$\frac{\mathrm{d}^n\overrightarrow{OM}}{\mathrm{d}t^n}(t_0)$$

jusqu'à détermination des entiers p et q.
2ème méthode
On réalise un développement limité de x et y en t_0 :

$$\begin{cases} x(t) = a_0 + a_1(t-t_0) + \cdots + a_n(t-t_0)^n + o((t-t_0)^n) \\ y(t) = b_0 + b_1(t-t_0) + \cdots + b_n(t-t_0)^n + o((t-t_0)^n) \end{cases}$$

Par la formule de Taylor-Young, les coefficients de ces développements limités déterminent les nombres dérivés successifs de x et y en t_0.
Ainsi $M(t_0)$ est le point de coordonnées (a_0, b_0) et pour $k \geqslant 1$,

$$a_k = \frac{x^{(k)}(t_0)}{k!} \text{ et } b_k = \frac{y^{(k)}(t_0)}{k!}$$

En notant $\vec{u}_1 \left| \begin{matrix} a_1 \\ b_1 \end{matrix} \right., \ldots, \vec{u}_n \left| \begin{matrix} a_n \\ b_n \end{matrix} \right.$ on a

$$\vec{u}_k = \frac{1}{k!} \frac{\mathrm{d}^k \overrightarrow{OM}}{\mathrm{d}t^k}(t_0)$$

On peut alors raisonner à partir des $\vec{u}_k$ au lieu des vecteurs dérivés successifs :
- le premier $\vec{u}_p$ non nul donne la direction de la tangente ;
- le premier $\vec{u}_q$ suivant non colinéaire à $\vec{u}_p$ donne la position de la courbe par rapport à sa tangente.

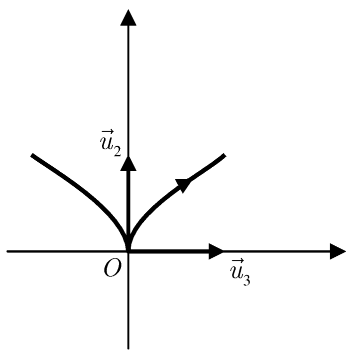
Exemple Etudions $\begin{cases} x = t - \sin t \\ y = 1 - \cos t \end{cases}$ en $t = 0$.

$$x(t) = \frac{1}{6}t^3 + o(t^3) \text{ et } y(t) = \frac{1}{2}t^2 + o(t^3) \text{ quand } t \to 0$$

On obtient

$$M(0) \left| \begin{matrix} 0 \\ 0 \end{matrix} \right., \vec{u}_1 \left| \begin{matrix} 0 \\ 0 \end{matrix} \right., \vec{u}_2 \left| \begin{matrix} 0 \\ 1/2 \end{matrix} \right., \vec{u}_3 \left| \begin{matrix} 1/6 \\ 0 \end{matrix} \right.$$

$p = 2, q = 3$, on a un point de rebroussement de première espèce.



Remarque Pour les figures, comme seuls la direction et le sens des vecteurs $\vec{u}_k$ importe, il est fréquent de représenter des vecteurs de longueur quelconque.

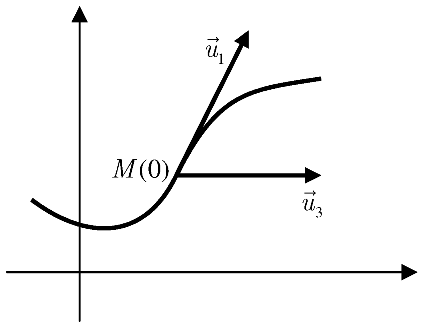
Exemple Etudions $\begin{cases} x = \mathrm{e}^t \\ y = (t+1)^2 \end{cases}$ en $t = 0$.

$$x(t) = 1 + t + \frac{1}{2}t^2 + \frac{1}{6}t^3 + o(t^3) \text{ et } y(t) = 1 + 2t + t^2 \text{ quand } t \to 0$$

On obtient

$$M(0)\left|\begin{matrix}1\\1\end{matrix}\right., \vec{u}_1\left|\begin{matrix}1\\2\end{matrix}\right., \vec{u}_2\left|\begin{matrix}1/2\\1\end{matrix}\right., \vec{u}_3\left|\begin{matrix}1/6\\0\end{matrix}\right.$$

$p = 1, q = 3$, on a un point d'inflexion



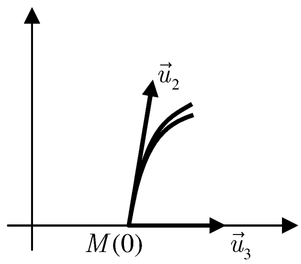
Exemple Etudions $\begin{cases} x = \mathrm{e}^t - t \\ y = t^2(t+3) \end{cases}$ en $t = 0$.

$$x(t) = 1 + \frac{1}{2}t^2 + \frac{1}{6}t^3 + \frac{1}{24}t^4 + o(t^4) \text{ et } y(t) = 3t^2 + t^3 + o(t^4) \text{ quand } t \to 0$$

On obtient

$$M(0)\left|\begin{matrix}1\\0\end{matrix}\right., \vec{u}_1\left|\begin{matrix}0\\0\end{matrix}\right., \vec{u}_2\left|\begin{matrix}1/2\\3\end{matrix}\right., \vec{u}_3\left|\begin{matrix}1/6\\1\end{matrix}\right., \vec{u}_4\left|\begin{matrix}1/24\\0\end{matrix}\right.$$

$p = 2, q = 4$, on a un point de rebroussement de seconde espèce.



Chapitre 22

Equations différentielles linéaires

$\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$.
I désigne un intervalle non singulier de $\mathbb{R}$.
Une équation différentielle d'ordre n en la fonction inconnue $x \mapsto y(x)$ est une égalité E exprimée à l'aide de $x, y, y', \dots, y^{(n)}$.
Une solution sur un intervalle I de cette équation différentielle est une fonction n fois dérivable $x \mapsto y(x)$ telle que, pour tout $x \in I$, l'égalité E soit vérifiée pour $x, y(x), y'(x), \dots, y^{(n)}(x)$.

Exemple $E : xy' = y(y+1) + x^2$ est une équation différentielle d'ordre 1.
La fonction $x \mapsto x \tan x$ est solution de celle-ci sur l'intervalle $]-\pi/2, \pi/2[$.

On ne sait résoudre qu'assez peu d'équations différentielles parmi lesquelles les suivantes...

22.1 Equation linéaire du premier ordre

22.1.1 Définition

Définition

On appelle équation différentielle linéaire d'ordre 1 définie sur I toute équation différentielle E de la forme

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

avec $a, b : I \to \mathbb{K}$ des fonctions continues.
On appelle équation homogène associée (ou équation sans second membre) l'équation

$$y' + a(x)y = 0$$

notée E_0 ou encore $ESSM$.
Lorsque la fonction a est constante, on parle d'équation à coefficient constant.

Exemple $E : y' + xy = x$ est une équation différentielle linéaire d'ordre 1 sur $\mathbb{R}$.
$E_0 : y' + xy = 0$ est l'équation homogène associée.
$y(x) = 1$ est une solution sur $\mathbb{R}$ de E.

Exemple $E:(x^2+1)y'+xy=1$ n'est pas à proprement parler une équation différentielle linéaire d'ordre 1 car il y a ici une fonction de x en facteur de y'. Cependant elle se ramène à une équation linéaire d'ordre 1 car l'équation différentielle E est équivalente à l'équation différentielle

$$y'+\frac{x}{1+x^2}y=\frac{1}{1+x^2}$$

Théorème

Soient $a,b:I\to\mathbb{K}$ continues, $x_0\in I$ et $y_0\in\mathbb{K}$.
Il existe une unique solution sur l'intervalle I à l'équation différentielle

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

vérifiant la condition initiale $y(x_0)=y_0$.

dém. :
Unicité :
Soit y une solution du problème posé.
Considérons A la primitive de a qui s'annule en x_0 ; celle-ci existe car la fonction a est continue sur l'intervalle I.
On a

$$\left(y(x)\mathrm{e}^{A(x)}\right)'=\left(y'(x)+a(x)y(x)\right)\mathrm{e}^{A(x)}=b(x)\mathrm{e}^{A(x)}$$

Considérons alors B la primitive de $x\mapsto b(x)\mathrm{e}^{A(x)}$ qui s'annule en x_0 ; celle-ci existe aussi car la fonction $x\mapsto b(x)\mathrm{e}^{A(x)}$ est continue sur l'intervalle I.
Puisque les fonctions $x\mapsto y(x)\mathrm{e}^{A(x)}$ et B sont primitives sur I de la même fonction $x\mapsto b(x)\mathrm{e}^{A(x)}$, celles-ci diffèrent ne différent que d'une constante et, connaissant les valeurs de chacune en x_0, on obtient $y(x)\mathrm{e}^{A(x)}=B(x)+y_0$ pour tout $x\in I$.
Ainsi $y(x)=(B(x)+y_0)\mathrm{e}^{-A(x)}$ pour tout $x\in I$ ce qui détermine la fonction y de façon unique.
Existence :
Considérons A la primitive de a qui s'annule en x_0 et B la primitive de $x\mapsto b(x)\mathrm{e}^{A(x)}$ qui s'annule en x_0.
Soit $y:I\to\mathbb{K}$ la fonction définie par $y(x)=(B(x)+y_0)\mathrm{e}^{-A(x)}$.
D'une part $y(x_0)=(B(x_0)+y_0)\,\mathrm{e}^{-A(x_0)}=y_0$.
D'autre part, y est dérivable et $y'(x)=-a(x)y(x)+B'(x)\mathrm{e}^{-A(x)}$ donc $y'(x)+a(x)y(x)=b(x)$.
Ainsi y est solution de l'équation différentielle étudiée et satisfait la condition initiale $y(x_0)=y_0$.
□

Exemple Pour $a\in\mathbb{C}$, il existe unique solution sur $\mathbb{R}$ à l'équation différentielle $y'=ay$ vérifiant $y(0)=1$, c'est la fonction $x\mapsto\mathrm{e}^{ax}$.

Remarque En pratique, la formule solution proposée par ce théorème n'est pas employée.

22.1.2 Démarche de résolution

Proposition

Soient $a, b : I \to \mathbb{K}$ continues et

$$E : y' + a(x)y = b(x)$$

Si y_1 désigne une solution particulière de l'équation E alors les solutions de E sont les fonctions de la forme $x \mapsto y_1(x) + y_0(x)$ avec y_0 solution de l'équation homogène E_0 : $y' + a(x)y = 0$.

dém. :
Supposons y_1 solution de E sur I.
Soit $y : I \to \mathbb{K}$ une fonction dérivable.
y est solution de E sur I si, et seulement si, pour tout $x \in I$, $y'(x) + a(x)y(x) = b(x)$ i.e. $y'(x) + a(x)y(x) = y_1'(x) + a(x)y_1(x)$ ce qui équivaut encore à $(y - y_1)'(x) + a(x)(y - y_1)(x) = 0$.
Ainsi, y est solution de E sur I si, et seulement si, la fonction $x \mapsto y(x) - y_1(x)$ est solution de l'équation homogène $E_0 : y' + a(x)y = 0$.
□

Remarque Pour résoudre une équation différentielle $E : y' + a(x)y = b(x)$:
- on présente le type d'équation différentielle ;
- on résout l'équation homogène associée : $y_0(x) = \dots$;
- on détermine une solution particulière : $y_1(x) = \dots$;
- on exprime la solution générale : $y(x) = y_0(x) + y_1(x)$.

Remarque Afin de trouver une solution particulière, il arrive parfois qu'on décompose le second membre en plusieurs fonctions plus simples ; on peut alors exploiter le résultat suivant :

Proposition

Soient $E : y' + a(x)y = b_1(x) + b_2(x)$ continues.
Si, pour $i \in \{1, 2\}$, y_i est solution de l'équation $y' + a(x)y = b_i(x)$
alors la fonction $y_1 + y_2$ est solution de $y' + a(x)y = b_1(x) + b_2(x)$.

dém. :
C'est immédiat par le calcul.
□

22.1.3 Cas des équations à coefficients constants

22.1.3.1 Résolution de l'équation homogène

Définition

On appelle équation caractéristique associée à l'équation différentielle linéaire à coefficient constant $E : y' + ay = 0$ (avec $a \in \mathbb{K}$) l'équation $r + a = 0$.

Proposition

Les solutions de l'équation $E : y' + ay = 0$ sont les fonctions $y : x \mapsto \lambda \mathrm{e}^{\alpha x}$ avec λ parcourant $\mathbb{K}$ et où α désigne la solution de l'équation caractéristique associée.

dém. :
Soit y une fonction dérivable.
La fonction $x \mapsto y(x)\mathrm{e}^{ax}$ est dérivable et

$$(y(x)\mathrm{e}^{ax})' = (y'(x) + ay(x))\,\mathrm{e}^{ax}$$

Par suite y est solution de l'équation E si, et seulement si, la fonction $x \mapsto y(x)\mathrm{e}^{ax}$ est constante ce qui donne y de la forme $x \mapsto \lambda\mathrm{e}^{-ax}$ avec $\lambda \in \mathbb{K}$.
□

Exemple Résolvons sur $\mathbb{R}$ l'équation $E : y' = y$.
E est une équation différentielle linéaire d'ordre 1 homogène à coefficient constant d'équation caractéristique $r = 1$. Sa solution générale est $y(x) = \lambda\mathrm{e}^{x}$ avec λ parcourant $\mathbb{K}$.

22.1.3.2 Résolution de l'équation entière

Soient $a \in \mathbb{K}$ et $b : I \to \mathbb{K}$ continue.
Pour achever de résoudre l'équation $E : y' + ay = b(x)$ il suffit de savoir déterminer une solution particulière.

Proposition

Si $b(x) = P(x)$ avec P fonction polynomiale alors on peut trouver une solution particulière à l'équation

$$E : y' + ay = b(x)$$

de la forme $y_1(x) = x^m Q(x)$ avec Q une fonction polynomiale de même degré que P et

$$m = \begin{cases} 0 & \text{si } a \neq 0 \\ 1 & \text{si } a = 0 \end{cases}$$

dém. :
Cas $a = 0$:
On chercher une solution particulière à l'équation $y' = P(x)$.
En déterminant une primitive de P s'annulant en 0, on obtient une solution particulière de la forme $x \mapsto xQ(x)$ avec Q fonction polynomiale de même degré que P.
Cas $a \neq 0$:
Posons $n = \deg P$ et écrivons $P(x) = a_0 + a_1 x + \cdots + a_n x^n$.
Considérons ensuite $Q(x) = b_0 + b_1 x + \cdots + b_n x^n$.

$$Q'(x) + aQ(x) = (b_1 + ab_0) + (2b_2 + ab_1)x + \cdots + (nb_n + ab_{n-1})x^{n-1} + ab_n x^n$$

En déterminant $b_0, b_1, \ldots, b_n$ tels que

$$\begin{cases} ab_n = a_n \\ nb_n + ab_{n-1} = a_{n-1} \\ \vdots \\ 2b_2 + ab_1 = a_1 \\ 2b_1 + ab_0 = a_0 \end{cases}$$

ce qui est possible car $a \neq 0$, on obtient une solution particulière $x \mapsto Q(x)$ de la forme annoncée.
□

Exemple Résolvons sur $\mathbb{R}$ l'équation $E : y' + y = x$.
E est une équation différentielle linéaire d'ordre 1 à coefficients constants.
Solution homogène $y(x) = \lambda \mathrm{e}^{x}$.
Solution particulière $y(x) = x - 1$.
Solution générale

$$y(x) = x - 1 + \lambda \mathrm{e}^{-x} \text{ avec } \lambda \text{ parcourant } \mathbb{K}$$

Exemple Résolvons sur $\mathbb{R}$ l'équation $E : y' + ay = b$ avec $a \neq 0$ et $b \in \mathbb{K}$.
E est une équation différentielle linéaire d'ordre 1 à coefficients constants.
Solution homogène $y_0(x) = \lambda \mathrm{e}^{-ax}$.
Solution particulière $y_1(x) = b/a$.
Solution générale

$$y(x) = b/a + \lambda \mathrm{e}^{-ax} \text{ avec } \lambda \text{ parcourant } \mathbb{K}$$

Proposition

Si $b(x) = P(x)\mathrm{e}^{\alpha x}$ avec $\alpha \in \mathbb{K}$ et P une fonction polynomiale alors on peut trouver une solution particulière à l'équation

$$E : y' + ay = b(x)$$

de la forme $y_1(x) = x^m Q(x)\mathrm{e}^{\alpha x}$ avec Q fonction polynomiale de même degré que P et

$$m = \begin{cases} 1 & \text{si } \alpha \text{ est solution de l'équation caractéristique} \\ 0 & \text{sinon} \end{cases}$$

dém. :
Soit y une fonction dérivable et $z : x \mapsto y(x)\mathrm{e}^{-\alpha x}$ une fonction dérivable telle que $y(x) = z(x)\mathrm{e}^{\alpha x}$.
On a

$$y'(x) + ay(x) = (z'(x) + (a + \alpha)z(x))\mathrm{e}^{\alpha x}$$

Par suite y est solution de l'équation E si, et seulement si, z est solution de l'équation $z' + (a + \alpha)z = P(x)$.
L'étude qui précède assure que cette dernière équation présente une solution particulière de la fonction $z_1(x) = x^m Q(x)$ avec

$$m = \begin{cases} 0 & \text{si } a + \alpha \neq 0 \\ 1 & \text{si } a + \alpha = 0 \end{cases}$$

et on obtient alors une solution particulière $y_1(x) = x^m Q(x)\mathrm{e}^{\alpha x}$ à l'équation E de la forme annoncée.
□

Exemple Résolvons sur $\mathbb{R}$ l'équation $E : y' + y = 2x\mathrm{ch}x$.
E est une équation différentielle linéaire d'ordre 1 à coefficients constants.
Solution homogène $y(x) = \lambda \mathrm{e}^{-x}$.
Solution particulière à l'équation $E_1 : y' + y = x\mathrm{e}^{x}$, $y(x) = \dfrac{2x - 1}{4}\mathrm{e}^{x}$.

Solution particulière à l'équation $E_2 : y' + y = x\,\mathrm{e}^{-x}$, $y(x) = \dfrac{x^2}{2}\mathrm{e}^{-x}$
En vertu du principe de superposition, la solution générale de E est

$$y(x) = \frac{2x-1}{2}\mathrm{e}^{x} + \frac{x^2}{2}\mathrm{e}^{-x} + \lambda\,\mathrm{e}^{-x} \text{ avec } \lambda \text{ parcourant } \mathbb{K}.$$

Proposition

Cas $\mathbb{K} = \mathbb{R}$
Si $b(x) = P(x)\cos(\omega x)$ (respectivement $P(x)\sin(\omega x)$) avec $\omega \in \mathbb{R}$ et P une fonction polynomiale réelle alors on peut trouver une solution particulière à l'équation E en considérant la partie réelle (respectivement imaginaire) d'une solution de l'équation différentielle complexe $z' + az = P(x)\mathrm{e}^{i\omega x}$.

dém. :
Si z est solution de l'équation $z' + az = P(x)\mathrm{e}^{i\omega x}$ alors en passant à la partie réelle et à la partie imaginaire la relation $z'(x) + az(x) = P(x)\mathrm{e}^{i\omega x}$, on obtient $\mathrm{Re}(z)'(x) + a\mathrm{Re}(z)(x) = P(x)\cos(\omega x)$ et $\mathrm{Im}(z)'(x) + a\mathrm{Im}(z)(x) = P(x)\sin(\omega x)$ car a et $P(x)$ sont réels.
□

Exemple Résolvons sur $\mathbb{R}$ l'équation $E : y' + y = x\cos x$.
E est une équation différentielle linéaire d'ordre 1 à coefficients constants.
Solution homogène $y(x) = \lambda\mathrm{e}^{-x}$
Solution particulière à l'équation $z' + z = x\,\mathrm{e}^{ix}$

$$z(x) = \left(\frac{x}{1+i} - \frac{1}{(1+i)^2}\right)\mathrm{e}^{ix} = \left(\frac{(1-i)x}{2} + \frac{i}{2}\right)\mathrm{e}^{ix}$$

Solution générale à l'équation $E : y' + y = x\cos x$

$$y(x) = \frac{1}{2}x\cos x + \frac{1}{2}(x-1)\sin x + \lambda\mathrm{e}^{-x} \text{ avec } \lambda \text{ parcourant } \mathbb{R}$$

22.1.4 Cas général

22.1.4.1 Résolution de l'équation homogène

Théorème

Soient $a : I \to \mathbb{K}$ continue et A une primitive de a.
L'ensemble des solutions sur I de l'équation $y' + a(x)y = 0$ est constitué des fonctions de la forme $x \mapsto \lambda\mathrm{e}^{-A(x)}$ avec λ parcourant $\mathbb{K}$.

dém. :
Soit y une fonction dérivable définie sur I.
La fonction $x \mapsto y(x)\mathrm{e}^{A(x)}$ est dérivable et

$$\left(y(x)\mathrm{e}^{A(x)}\right)' = \left(y'(x) + a(x)y(x)\right)\mathrm{e}^{A(x)}$$

Par suite la fonction y est solution de l'équation $y' + a(x)y = 0$ si, et seulement si, la fonction $x \mapsto y(x)\mathrm{e}^{A(x)}$ est constante ce qui donne y de la forme $x \mapsto \lambda\mathrm{e}^{-A(x)}$ avec $\lambda \in \mathbb{K}$.
□

Remarque La seule solution à l'équation $y' + a(x)y = 0$ qui s'annule est la fonction nulle.

Remarque Dans le cas où a fonction constante, on retrouve le résultat précédemment obtenu.

Remarque Pour résoudre $y' + a(x)y = 0$:
- on exprime y' en fonction de y ;
- on détermine une primitive du facteur de y ainsi obtenu ;
- on exprime la solution générale.

Cette méthodologie évite toute erreur de signe et permet une vérification rapide de la résolution.

Exemple Résolvons sur $\mathbb{R}$ l'équation $E : y' + xy = 0$.
E est une équation différentielle linéaire d'ordre 1 homogène

$$y' + xy = 0 \Leftrightarrow y' = -xy$$

et

$$\int -x\,\mathrm{d}x = -\frac{1}{2}x^2 + C^{te}$$

Solution générale de E sur $\mathbb{R}$

$$y(x) = \lambda \mathrm{e}^{-x^2/2} \text{ avec } \lambda \text{ parcourant } \mathbb{K}$$

Exemple Résolvons sur $\mathbb{R}$ l'équation $E : y' + \sin(x)y = 0$.
E est une équation différentielle linéaire d'ordre 1 homogène
$y' + \sin(x)y = 0 \Leftrightarrow y' = -\sin(x)y$ et $\displaystyle\int -\sin(x)\,\mathrm{d}x = \cos x$
Solution générale de E sur $\mathbb{R}$
$y(x) = \lambda \mathrm{e}^{\cos x}$ avec λ parcourant $\mathbb{K}$.

Exemple Résolvons sur $I = \mathbb{R}^{+\star}$ ou $\mathbb{R}^{-\star}$ l'équation

$$E : y' + \frac{1}{x}y = 0$$

E est une équation différentielle linéaire d'ordre 1 homogène

$$y' + \frac{1}{x}y = 0 \Leftrightarrow y' = -\frac{1}{x}y \text{ et } \int -\frac{\mathrm{d}x}{x} = \ln|x|$$

Solution générale de E sur $I = \mathbb{R}^{+\star}$ ou $\mathbb{R}^{-\star}$, $y(x) = \lambda/|x|$ avec λ parcourant $\mathbb{K}$.
Quitte à changer λ en $-\lambda$ dans le cas où $I = \mathbb{R}^{-\star}$, on peut affirmer avec plus de légèreté que la solution générale de E sur I est $y(x) = \lambda/x$ avec λ parcourant $\mathbb{K}$.

22.1.4.2 Résolution de l'équation entière

Soient $a, b : I \to \mathbb{K}$ continues et
$$E : y' + a(x)y = b(x)$$
La résolution de l'équation homogène $E_0 : y' + a(x)y = 0$ a donné $y(x) = \lambda \mathrm{e}^{-A(x)}$.
Pour achever de résoudre E, il suffit maintenant de déterminer une solution particulière $y(x)$.
Si celle-ci n'est pas apparente, on peut la rechercher par la méthode de la variation de la constante qui suit :
On cherche y de la forme $y(x) = \lambda(x)\mathrm{e}^{-A(x)}$ avec $x \mapsto \lambda(x)$ fonction dérivable.
$$y'(x) + a(x)y(x) = (\lambda'(x) - a(x)\lambda(x))\mathrm{e}^{-A(x)} + a(x)\lambda(x)\mathrm{e}^{-A(x)} = \lambda'(x)\mathrm{e}^{-A(x)}$$
Par suite y est solution de E si, et seulement si,
$$\lambda'(x) = b(x)\mathrm{e}^{A(x)}$$
Par détermination de primitive, ceci permet d'obtenir une fonction λ puis une fonction y convenables.

Exemple Résolvons sur $\mathbb{R}$ l'équation
$$E : y' + \frac{2x}{1+x^2}y = \frac{1}{1+x^2}$$
E est une équation différentielle linéaire d'ordre 1.
$$y' + \frac{2x}{1+x^2}y = 0 \Leftrightarrow y' = -\frac{2x}{1+x^2}y$$
et
$$\int \frac{2x\,\mathrm{d}x}{1+x^2} = \ln(1+x^2)$$
Solution homogène $y(x) = \dfrac{\lambda}{1+x^2}$
Solution particulière :
Cherchons celle-ci de la forme $y(x) = \dfrac{\lambda(x)}{1+x^2}$ avec $x \mapsto \lambda(x)$ dérivable.
$$y'(x) + \frac{2x}{1+x^2}y(x) = \frac{\lambda'(x)}{1+x^2} = \frac{1}{1+x^2} \Leftrightarrow \lambda'(x) = 1$$
$\lambda(x) = x$ convient et $y(x) = \dfrac{x}{1+x^2}$ est solution particulière de E.
Solution générale
$$y(x) = \frac{x+\lambda}{x^2+1} \text{ avec } \lambda \text{ parcourant } \mathbb{K}$$

Exemple Résolvons sur $]-1, 1[$ l'équation
$$E : (1-x^2)y' - xy = 1$$
E est une équation différentielle linéaire d'ordre 1.
$$(1-x^2)y' - xy = 0 \Leftrightarrow y' = \frac{x}{1-x^2}y$$

et

$$\int \frac{x\,\mathrm{d}x}{1-x^2} = -\frac{1}{2}\ln(1-x^2)$$

Solution homogène $y(x) = \dfrac{\lambda}{\sqrt{1-x^2}}$
Solution particulière :
Cherchons celle-ci de la former $y(x) = \dfrac{\lambda(x)}{\sqrt{1-x^2}}$ avec $x \mapsto \lambda(x)$ fonction dérivable.

$$(1-x^2)y'(x) - xy(x) = \sqrt{1-x^2}\lambda'(x) = 1 \Leftrightarrow \lambda'(x) = \frac{1}{\sqrt{1-x^2}}$$

$\lambda(x) = \arcsin x$ convient et $y(x) = \dfrac{\arcsin x}{\sqrt{1-x^2}}$ est solution particulière de E.
Solution générale

$$y(x) = \frac{\arcsin x + \lambda}{\sqrt{1-x^2}} \text{ avec } \lambda \text{ parcourant } \mathbb{K}$$

Exemple Résolvons sur $\mathbb{R}$ l'équation

$$E : y' - xy = x$$

E est une équation différentielle linéaire d'ordre 1.
Solution homogène

$$y' - xy = 0 \Leftrightarrow y' = xy$$

et

$$\int x\,\mathrm{d}x = \frac{1}{2}x^2$$

Solution homogène $y(x) = \lambda \mathrm{e}^{x^2/2}$
Solution particulière $y(x) = -1$ (celle-ci est apparente, il n'est pas nécessaire de mettre en place la méthode précédente)
Solution générale

$$y(x) = -1 + \lambda \mathrm{e}^{x^2/2} \text{ avec } \lambda \text{ parcourant } \mathbb{R}$$

22.2 Equation linéaire du second ordre à coefficients constants

22.2.1 Définition

Définition

On appelle équation différentielle linéaire d'ordre 2 à coefficients constants définie sur I toute équation différentielle E de la forme

$$y'' + ay' + by = c(x)$$

avec $a, b \in \mathbb{K}$ et $c : I \to \mathbb{K}$ une fonction continue.
On appelle équation homogène associée l'équation

$$y'' + ay' + b = 0$$

notée E_0 ou encore $ESSM$.
On appelle équation caractéristique associée l'équation

$$r^2 + ar + b = 0$$

d'inconnue $r \in \mathbb{C}$.

Exemple $y'' + y' - 2y = x$ est une équation différentielle linéaire d'ordre 2 à coefficients constants d'équation homogène $y'' + y' - 2y = 0$ et d'équation caractéristique $r^2 + r - 2 = 0$.

Théorème

Soient $a, b \in \mathbb{K}$, $c : I \to \mathbb{K}$ une fonction continue, $x_0 \in I$ et $y_0, y_1 \in \mathbb{K}$.
Il existe une unique solution sur I à l'équation différentielle

$$y'' + ay' + by = c(x)$$

vérifiant les conditions initiales $y(x_0) = y_0$ et $y'(x_0) = y_1$.

dém. :
Cas $\mathbb{K} = \mathbb{C}$.
Soient α, β les deux racines de l'équation caractéristique $r^2 + ar + b = 0$.
On sait $\alpha + \beta = -a$ et $\alpha\beta = b$ et l'équation différentielle étudiée se réécrit $y'' - (\alpha + \beta)y' + \alpha\beta y = c(x)$ ou encore $(y' - \alpha y)' - \beta(y' - \alpha y) = c(x)$.
Unicité :
Soient $y : I \to \mathbb{C}$ solution du problème posé et z la fonction donnée par $z = y' - \alpha y$.
La fonction z est dérivable et $z' - \beta z = y'' - (\alpha + \beta)y' + \alpha\beta y = c(x)$.
Ainsi z est solution de l'équation différentielle $z' - \beta z = c(x)$ linéaire d'ordre 1 .
De plus, z vérifie la condition initiale $z(x_0) = y_1 - \alpha y_0$ et donc la fonction z est déterminée de façon unique.
Puisque y peut se voir comme solution de l'équation différentielle $y' - \alpha y = z(x)$ linéaire d'ordre 1 et puisque y vérifie la condition initiale $y(x_0) = y_0$, la fonction y est elle aussi déterminée de manière unique.
Existence :
Soit z la solution de l'équation $z' - \beta z = c(x)$ vérifiant $z(x_0) = y_1 - \alpha y_0$ et soit y la solution de l'équation $y' - \alpha y = z(x)$ vérifiant $y(x_0) = y_0$.

Puisque la fonction z est dérivable et puisque $y' = z + \alpha y$, la fonction y est deux fois dérivable et

$$y'' - (\alpha + \beta)y' + \alpha\beta y = z' - \beta z = c(x)$$

Ainsi y est solution de l'équation

$$y'' + ay' + by = c(x)$$

De plus $y(x_0) = y_0$ et $y'(x_0) = z'(x_0) + \alpha y(x_0) = y_1$ donc y est solution du problème posé.
Cas $\mathbb{K} = \mathbb{R}$.
Ici les paramètres considérés sont tous réels et peut donc aussi se voir comme étant complexes.
Par l'étude qui précède, il existe une unique fonction $y : I \to \mathbb{C}$ solution du problème posé. Pour conclure il suffit de montrer que y est à valeurs dans $\mathbb{R}$.
Pour cela considérons la fonction $z : x \mapsto \overline{y(x)}$.
La fonction z est deux fois dérivable et on vérifie aisément que z est solution de l'équation différentielle étudiée vérifiant les conditions initiales proposées. Par unicité de la solution, on obtient $z = y$ et donc y est une fonction réelle.
□

22.2.2 Démarche de résolution

Proposition

Soient $a, b \in \mathbb{K}$ et $c : I \to \mathbb{K}$ une fonction continue.
Si y_1 désigne une solution particulière de l'équation $E : y'' + ay' + by = c(x)$ alors les solutions de E sont les fonctions de la forme $x \mapsto y_1(x) + y_0(x)$ avec y_0 solution de l'équation homogène $E_0 : y'' + ay' + by = 0$.

dém. :
Supposons y_1 solution de E sur I.
Soit $y : I \to \mathbb{K}$ une fonction deux fois dérivable.
y est solution de E sur I si, et seulement si, pour tout $x \in I$, $y''(x) + ay'(x) + by(x) = c(x)$ i.e.

$$y''(x) + ay'(x) + by(x) = y_1''(x) + ay_1'(x) + by_1(x)$$

ce qui équivaut encore à

$$(y - y_1)''(x) + a(y - y_1)'(x) + b(y - y_1)(x) = 0$$

Ainsi, y est solution de E sur I si, et seulement si, la fonction $x \mapsto y(x) - y_1(x)$ est solution de l'équation homogène $E_0 : y'' + ay' + by = 0$.
□

Remarque Pour résoudre l'équation différentielle $E : y'' + ay' + by = c(x)$:
- on présente le type de l'équation différentielle ;
- on résout l'équation homogène associée : $y_0(x) = \ldots$:
- on détermine une solution particulière : $y_1(x) = \ldots$;
- on exprime la solution générale : $y(x) = y_0(x) + y_1(x)$.

Remarque Afin de trouver une solution particulière, il arrive parfois qu'on décompose le second membre en plusieurs fonctions plus simples ; on peut alors exploiter le résultat suivant :

Proposition

Soient $a, b \in \mathbb{K}$ et $c_1, c_2 : I \to \mathbb{K}$ continues.
Si, pour $i \in \{1, 2\}$, y_i est solution de l'équation $y'' + ay' + by = c_i(x)$
alors la fonction $y_1 + y_2$ est solution de $y'' + ay' + bqy = c_1(x) + c_2(x)$.

dém. :
C'est immédiat par le calcul.
□

22.2.3 Résolution pratique de l'équation homogène

22.2.3.1 Cadre complexe

Théorème

Soient $a, b \in \mathbb{C}$ et $E_0 : y'' + ay' + by = 0$.
Si l'équation caractéristique $r^2 + ar + b = 0$ possède deux racines distinctes α et β alors les solutions sur $\mathbb{R}$ de E_0 sont les fonctions de la forme

$$y(x) = \lambda \mathrm{e}^{\alpha x} + \mu \mathrm{e}^{\beta x} \text{ avec } \lambda, \mu \text{ parcourant } \mathbb{C}$$

Si l'équation caractéristique $r^2 + ar + b = 0$ possède une racine double α
alors les solutions sur $\mathbb{R}$ de E_0 sont les fonctions de la forme

$$y(x) = (\lambda x + \mu)\mathrm{e}^{\alpha x} \text{ avec } \lambda, \mu \text{ parcourant } \mathbb{C}$$

dém. :
Notons α, β les deux racines de l'équation $r^2 + ar + b = 0$.
Soit $y : \mathbb{R} \to \mathbb{C}$ une fonction deux fois dérivable et $z : \mathbb{R} \to \mathbb{C}$ définie par $z(x) = y(x)\mathrm{e}^{-\alpha x}$.
z est deux fois dérivable et $y''(x) + ay'(x) + by(x) = (z''(x) + (\alpha - \beta)z'(x))\,\mathrm{e}^{\alpha x}$ car $a = \alpha + \beta$.
Ainsi y est solution de E sur $\mathbb{R}$ si, et seulement si, z est solution sur $\mathbb{R}$ de l'équation $z'' + (\alpha - \beta)z' = 0$.
Cas $\alpha \neq \beta$
z est solution de l'équation $z'' + (\alpha - \beta)z' = 0$ si, et seulement si, $z'(x) = \gamma \mathrm{e}^{(\beta - \alpha)x}$ soit

$$z(x) = \frac{\gamma}{\beta - \alpha}\mathrm{e}^{(\beta - \alpha)x} + \lambda \text{ avec } \lambda, \mu \text{ parcourant } \mathbb{C}$$

Par suite, on obtient que y est de la forme

$$y : x \mapsto \lambda \mathrm{e}^{\alpha x} + \mu \mathrm{e}^{\beta x} \text{ avec } \mu = \frac{\gamma}{\beta - \alpha} \text{ parcourant } \mathbb{C}$$

Cas $\alpha = \beta$
z est solution de l'équation $z'' = 0$ si, et seulement si, $z(x) = \lambda x + \mu$ avec $\lambda, \mu \in \mathbb{C}$.
□

Exemple Résolvons sur $\mathbb{R}$ l'équation

$$y'' + (1 + 2i)y' + (i - 1)y = 0$$

C'est une équation différentielle linéaire d'ordre 2 homogène d'équation caractéristique

$$r^2 + (1 + 2i)r + (i - 1) = 0$$

$\Delta = (1+2i)^2 - 4(i-1) = 1$.
Les racines de l'équation caractéristique sont $-i$ et $-(1+i)$.
La solution générale de l'équation étudiée est $y(x) = (\lambda + \mu \mathrm{e}^{-x})\mathrm{e}^{-ix}$ avec λ, μ parcourant $\mathbb{C}$.

22.2.3.2 Cadre réel

Théorème

Soient $a, b \in \mathbb{R}$ et $E_0 : y'' + ay' + by = 0$.
Si l'équation caractéristique $r^2 + ar + b = 0$ possède deux racines réelles distinctes α et β alors les solutions sur $\mathbb{R}$ de E_0 sont les fonctions de la forme

$$y(x) = \lambda \mathrm{e}^{\alpha x} + \mu \mathrm{e}^{\beta x} \text{ avec } \lambda, \mu \text{ parcourant } \mathbb{R}$$

Si l'équation caractéristique $r^2 + ar + b = 0$ possède une racine double α
alors les solutions sur $\mathbb{R}$ de E_0 sont les fonctions de la forme

$$y(x) = (\lambda x + \mu)\mathrm{e}^{\alpha x} \text{ avec } \lambda, \mu \text{ parcourant } \mathbb{R}$$

Si l'équation caractéristique $r^2 + ar + b = 0$ possède deux racines complexes conjuguées $\alpha \pm i\omega$ (avec $\alpha, \omega \in \mathbb{R}$, $\omega \neq 0$) alors les solutions sur $\mathbb{R}$ de E_0 sont les fonctions de la forme

$$y(x) = (\lambda \cos(\omega x) + \mu \sin(\omega x))\mathrm{e}^{\alpha x} \text{ avec } \lambda, \mu \text{ parcourant } \mathbb{R}$$

dém. :
Les deux premiers cas se traitent comme dans le cadre complexe.
Traitons le cas où l'équation caractéristique $r^2 + ar + b = 0$ possède deux racines complexes conjuguées $\alpha \pm i\omega$.
Les fonctions complexes solutions de E_0 sont les fonctions de la forme :

$$y(x) = \lambda \mathrm{e}^{(\alpha+i\omega)x} + \mu \mathrm{e}^{(\alpha-i\omega)x} = ((\lambda+\mu)\cos\omega x + i(\lambda-\mu)\sin\omega x)\,\mathrm{e}^{\alpha x}$$

avec $\lambda, \mu \in \mathbb{C}$.
Parmi celle-ci les fonctions réelles vérifient $y(0) \in \mathbb{R}$ et $y(2\pi/\omega) \in \mathbb{R}$ ce qui donne $\lambda + \mu \in \mathbb{R}$ et $\lambda - \mu \in i\mathbb{R}$.
Par suite $\mathrm{Im}(\lambda) + \mathrm{Im}(\mu) = 0$ et $\mathrm{Re}(\lambda) = \mathrm{Re}(\mu)$.
Posons alors $\gamma = \lambda + \mu \in \mathbb{R}$ et $\delta = i(\lambda - \mu) \in \mathbb{R}$.
La fonction y est alors de la forme

$$y(x) = (\gamma \cos(\omega x) + \delta \sin(\omega x))\mathrm{e}^{\alpha x}$$

avec $\gamma, \delta \in \mathbb{R}$.
Inversement, une telle fonction est solution de l'équation E_0.
□

Exemple Résolvons sur $\mathbb{R}$ l'équation

$$y'' + 4y' + 4y = 0$$

C'est une équation différentielle linéaire d'ordre 2 homogène à coefficients constants d'équation caractéristique $r^2 + 4r + 4 = 0$ de racine double $r = -2$.

Solution générale
$$y(x) = (\lambda x + \mu)\mathrm{e}^{-2x} \text{ avec } \lambda, \mu \text{ parcourant } \mathbb{R}$$

Exemple Résolvons sur $\mathbb{R}$ l'équation $y'' + 2y' + 2y = 0$.
C'est une équation différentielle linéaire d'ordre 2 homogène à coefficients constants d'équation caractéristique $r^2 + 2r + 2 = 0$ de racines complexes $-1 \pm i$.
Solution générale
$$y(x) = (\lambda \cos x + \mu \sin x)\mathrm{e}^{-x} \text{ avec } \lambda, \mu \text{ parcourant } \mathbb{R}$$

Exemple Soit $\omega \in \mathbb{R}^{+\star}$. Résolvons sur $\mathbb{R}$ l'équation $y'' + \omega^2 y = 0$.
C'est une équation différentielle linéaire d'ordre 2 homogène à coefficients constants d'équation caractéristique $r^2 + \omega^2 = 0$ de racines complexes $\pm i\omega$.
Solution générale
$$y(x) = \lambda \cos(\omega x) + \mu \sin(\omega x) \text{ avec } \lambda, \mu \text{ parcourant } \mathbb{R}$$

Exemple Soit $\omega \in \mathbb{R}^{+\star}$. Résolvons sur $\mathbb{R}$ l'équation $y'' - \omega^2 y = 0$.
C'est une équation différentielle linéaire d'ordre 2 homogène à coefficients constants d'équation caractéristique $r^2 - \omega^2 = 0$ de racines réelles $\pm\omega$.
Solution générale
$$y(x) = \lambda \mathrm{e}^{\omega x} + \mu \mathrm{e}^{-\omega x} \text{ avec } \lambda, \mu \text{ parcourant } \mathbb{R}$$

Par analogie avec la précédente équation, on peut aussi écrire la solution générale sous la forme
$$y(x) = \alpha \mathrm{ch}(\omega x) + \beta \mathrm{sh}(\omega x) \text{ avec } \alpha, \beta \text{ parcourant } \mathbb{R}$$

22.2.4 Résolution pratique de l'équation entière

Soient $a, b \in \mathbb{K}$ et $c : I \to \mathbb{K}$ continue.
Pour achever de résoudre l'équation $E : y'' + ay' + by = c(x)$, il suffit d'en déterminer une solution particulière.

Proposition

Si $c(x) = P(x)$ avec P fonction polynomiale alors on peut trouver une solution particulière à l'équation
$$E : y'' + ay' + by = P(x)$$
de la forme $y(x) = x^m Q(x)$ avec Q fonction polynomiale de même degré que P et
$$m = \begin{cases} 0 & \text{si } b \neq 0 \\ 1 & \text{si } b = 0, a \neq 0 \\ 2 & \text{si } b = 0, a = 0 \end{cases}$$

dém. :
Cas $a = b = 0$:
On cherche une solution particulière à l'équation $y'' = P(x)$.

En déterminant une primitive de P s'annulant en 0 et en déterminant une primitive cette dernière s'annulant encore en 0, on obtient une solution particulière de la forme $x \mapsto x^2 Q(x)$ avec Q fonction polynomiale de même degré que P.
Cas $a \neq 0$ et $b = 0$:
On cherche une solution particulière à l'équation $y'' + ay' = P(x)$.
Considérons alors l'équation $z' + az = P(x)$.
On peut trouver une solution particulière à celle-ci de la forme $z(x) = Q(x)$ avec Q fonction polynomiale de même degré que P.
En déterminant une primitive de Q s'annulant en 0, on obtient une solution particulière de la forme $x \mapsto xR(x)$ avec R fonction polynomiale de même degré que P.
Cas $b \neq 0$:
Posons $n = \deg P$ et écrivons $P(x) = a_0 + a_1 x + \cdots + a_n x^n$.
Considérons ensuite $Q(x) = b_0 + b_1 x + \cdots + b_n x^n$.
Par identification des coefficients polynomiaux, l'équation $Q''(x) + aQ'(x) + bQ(x) = P(x)$ conduit au système

$$\begin{cases} 2b_2 + ab_1 + bb_0 = a_0 \\ \vdots \\ n(n-1)b_n + b(n-1)b_{n-1} + ab_n = a_{n-2} \\ anb_n + bb_{n-1} = a_{n-1} \\ bb_n = a_n \end{cases}$$

Puisque $b \neq 0$, il est possible de déterminer une solution à ce système triangulaire.
□

Exemple Résolvons sur $\mathbb{R}$ l'équation

$$E : y'' + y' + y = x^2 + 1$$

C'est une équation différentielle linéaire d'ordre 2 à coefficients constant d'équation caractéristique $r^2 + r + 1 = 0$ de racines $-\dfrac{1}{2} \pm i\dfrac{\sqrt{3}}{2}$.
Solution générale de l'équation homogène

$$y(x) = \left(\lambda \cos\left(\frac{\sqrt{3}}{2}x\right) + \mu \sin\left(\frac{\sqrt{3}}{2}x\right)\right) \mathrm{e}^{-\frac{1}{2}x}$$

Solution particulière
On cherche celle-ci de la forme $y(x) = ax^2 + bx + c$.

$$y''(x) + y'(x) + y(x) = (2a + b + c) + (2a + b)x + ax^2$$

Pour $a = 1$, $b = -2$ et $c = 1$, $y(x) = x^2 - 2x + 1$ est solution particulière.
Solution générale

$$y(x) = (x-1)^2 + \left(\lambda \cos\frac{\sqrt{3}}{2}x + \mu \sin\frac{\sqrt{3}}{2}x\right) \mathrm{e}^{-\frac{1}{2}x} \text{ avec } \lambda, \mu \text{ parcourant } \mathbb{R}$$

Proposition

Si $c(x) = P(x)\mathrm{e}^{\alpha . x}$ avec P fonction polynomiale et $\alpha \in \mathbb{K}$ alors on peut trouver une solution à l'équation
$$E : y'' + ay' + by = P(x)\mathrm{e}^{\alpha x}$$
de la forme $y(x) = x^m Q(x)\mathrm{e}^{\alpha x}$ avec Q fonction polynomiale de même degré que P et
$$m = \begin{cases} 0 & \text{si } \alpha \text{ non racine de } r^2 + ar + b = 0 \\ 1 & \text{si } \alpha \text{ racine simple de } r^2 + ar + b = 0 \\ 2 & \text{si } \alpha \text{ racine double de } r^2 + ar + b = 0 \end{cases}$$

dém. :
En effet si $y(x) = z(x)\mathrm{e}^{\alpha x}$ avec z une fonction deux fois dérivable alors
$$y''(x) + ay'(x) + by(x) = \left(z''(x) + (2\alpha + a)z'(x) + (\alpha^2 + a\alpha + b)z(x)\right)\mathrm{e}^{\alpha x}$$
Si z est solution de l'équation
$$z'' + (2\alpha + a)z' + (\alpha^2 + a\alpha + b)z = P(x)$$
alors y est solution particulière de E.
Or on peut trouver une solution particulière $z(x) = x^m Q(x)$ avec Q fonction polynomiale de même degré que P et
$$m = \begin{cases} 0 & \text{si } \alpha^2 + a\alpha + b \neq 0 \text{ i.e. } \alpha \text{ racine de } ar^2 + br + c = 0 \\ 1 & \text{si } \alpha^2 + a\alpha + b = 0 \text{ et } 2\alpha + b \neq 0 \text{ i.e. } \alpha \text{ racine simple de } ar^2 + br + c = 0 \\ 2 & \text{si } \alpha^2 + a\alpha + b = 0 \text{ et } 2\alpha + b = 0 \text{ i.e. } \alpha \text{ racine double de } ar^2 + br + c = 0 \end{cases}$$
Ainsi, on peut trouver une solution y à l'équation E de la forme proposée.
□

Exemple Résolvons sur $\mathbb{R}$ l'équation
$$E : y'' - 2y' + y = x\mathrm{e}^{-x}$$
E est une équation différentielle linéaire d'ordre 2 à coefficients constants d'équation caractéristique $r^2 - 2r + 1 = 0$ de racine double 1.
Solution homogène $y(x) = (\lambda x + \mu)\mathrm{e}^{x}$.
Solution particulière :
On cherche celle-ci de la forme $y(x) = (ax + b)\mathrm{e}^{-x}$.
$$y''(x) - 2y'(x) + y(x) = (4ax + 4b - 4a)\mathrm{e}^{-x}$$
Pour $a = 1/4$ et $b = 1/4$, $y(x) = \dfrac{x+1}{4}\mathrm{e}^{-x}$ est solution particulière.
Solution générale
$$y(x) = \frac{1}{4}(x+1)\mathrm{e}^{-x} + (\lambda x + \mu)\mathrm{e}^{x} \text{ avec } \lambda, \mu \text{ parcourant } \mathbb{R}$$

Exemple Résolvons sur $\mathbb{R}$ l'équation

$$E : y'' - 3y' + 2y = x\mathrm{e}^x$$

E est une équation différentielle linéaire d'ordre 2 à coefficients constants d'équation caractéristique $r^2 - 3r + 2 = 0$ de racines 1 et 2.
Solution homogène $y(x) = \lambda \mathrm{e}^x + \mu \mathrm{e}^{2x}$.
Solution particulière :
On cherche celle-ci de la forme $y(x) = (ax^2 + bx)\mathrm{e}^x$.

$$y''(x) - 3y'(x) + 2y(x) = (-2ax + 2a - b)\mathrm{e}^x$$

Pour $a = -1/2$ et $b = -1$, $y(x) = -\frac{1}{2}(x^2 + 2x)\mathrm{e}^x$ est solution particulière.
Solution générale

$$y(x) = -\frac{1}{2}(x^2 + 2x)\mathrm{e}^x + \lambda \mathrm{e}^x + \mu \mathrm{e}^{2x} \text{ avec } \lambda, \mu \text{ parcourant } \mathbb{R}$$

Proposition

Cas $\mathbb{K} = \mathbb{R}$.
Si $c(x) = P(x)\cos(\omega x)$ (resp. $c(x) = P(x)\sin(\omega x)$) avec P fonction polynomiale et $\omega \in \mathbb{R}$ alors on peut trouver une solution particulière à l'équation $E : y'' + ay' + by = c(x)$ en considérant la partie réelle (resp. imaginaire) d'une solution de l'équation différentielle complexe $z'' + az' + bz = P(x)\mathrm{e}^{i\omega x}$.

dém. :
Sachant que $a, b \in \mathbb{R}$ et $P(x)$ réel, en passant à la partie réelle et la partie imaginaire la relation $z''(x) + az'(x) + bz(x) = P(x)\mathrm{e}^{i\omega x}$ on obtient que $\mathrm{Re}(z)$ et $\mathrm{Im}(z)$ sont solutions respectivement des équations

$$y'' + ay' + by = P(x)\cos(\omega x) \text{ et } y'' + ay' + by = P(x)\sin(\omega x)$$

□

Exemple Résolvons sur $\mathbb{R}$ l'équation $y'' + y = x\cos(2x)$.
E est une équation différentielle linéaire d'ordre 2 à coefficients constants d'équation caractéristique $r^2 + 1 = 0$ de racines complexes $\pm i$.
Solution homogène $y(x) = \lambda \cos x + \mu \sin x$.
Cherchons une solution particulière à l'équation $z'' + z = x\mathrm{e}^{2ix}$ de la forme $z(x) = (ax + b)\mathrm{e}^{2ix}$,

$$z''(x) + z(x) = (-3(ax + b) + 4ia)\mathrm{e}^{2ix}$$

Pour $a = -1/3$ et $b = -4i/9$, $z(x) = -\frac{3x + 4i}{9}\mathrm{e}^{2ix}$ est solution particulière.
On en déduit que

$$y(x) = \mathrm{Re}(z(x)) = -\frac{3x\cos(2x) - 4\sin(2x)}{9}$$

est solution particulière de E
Solution générale

$$y(x) = \frac{4\sin(2x) - 3x\cos(2x)}{9} + \lambda \cos x + \mu \sin x \text{ avec } \lambda, \mu \text{ parcourant } \mathbb{R}$$

Exemple Résolvons sur $\mathbb{R}$ l'équation $y'' + 2y' + 2y = \cos x + \sin x$.
E est une équation différentielle linéaire d'ordre 2 à coefficients constants d'équation caractéristique $r^2 + 2r + 2 = 0$ de racines complexes $-1 \pm i$.
Solution homogène $y(x) = (\lambda \cos x + \mu \sin x)\mathrm{e}^{-x}$.
Cherchons une solution particulière à l'équation $z'' + 2z' + 2z = \mathrm{e}^{ix}$.

$$z(x) = \frac{1}{1+2i}\mathrm{e}^{ix} = \frac{1-2i}{5}\mathrm{e}^{ix} \text{ convient}$$

Par le principe de superposition

$$y(x) = \frac{\cos x + 2\sin x}{5} + \frac{\sin x - 2\cos x}{5} = \frac{3\sin x - \cos x}{5}$$

est solution particulière.
Solutions générale

$$y(x) = \frac{3\sin x - \cos x}{5} + (\lambda \cos x + \mu \sin x)\mathrm{e}^{-x} \text{ avec } \lambda, \mu \text{ parcourant } \mathbb{R}$$

22.2.5 Oscillateurs linéaires libres

On se propose d'étudier l'équation différentielle

$$y'' + 2my' + \omega_0^2 y = 0$$

sur $\mathbb{R}^+$ avec $m \geqslant 0$ et $\omega_0 > 0$.
m se comprend comme un paramètre d'amortissement.
ω_0 se comprend comme un paramètre de pulsation propre.

Exemple La décharge d'un condensateur dans un circuit RLC série conduit à l'équation différentielle

$$q'' + \frac{R}{L}q' + \frac{1}{LC}q = 0$$

L'équation $y'' + 2my' + \omega_0^2 y = 0$ est une équation différentielle linéaire d'ordre 2 à coefficients constant d'équation caractéristique $r^2 + 2mr + \omega_0^2 = 0$ et de discriminant $\Delta = 4(m^2 - \omega_0^2)$.

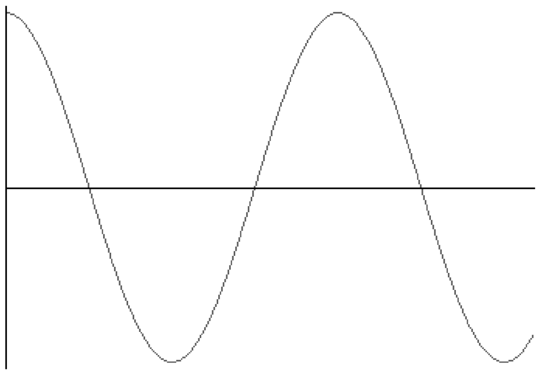
22.2.5.1 Cas $m = 0$, amortissement nul

$$\Delta = (2i\omega_0)^2$$

L'équation caractéristique a pour racines $\pm i\omega_0$.
La solution générale de l'équation différentielle est

$$y(t) = \lambda \cos(\omega_0 t) + \mu \sin(\omega_0 t) = A \cos(\omega_0 t - \varphi)$$

On obtient un mouvement périodique d'amplitude A et de période propre $T_0 = 2\pi/\omega_0$.
La solution correspondant aux conditions initiales $y(0) = y_0$ et $y'(0) = 0$ est $y(t) = y_0 \cos(\omega_0 t)$.



22.2.5.2 Cas $m > 0$ et $\Delta < 0$, amortissement faible

Si $m < \omega_0$ alors $\Delta < 0$

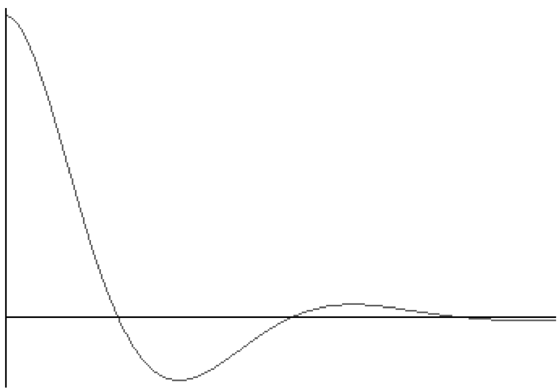
$$\Delta = \left(2i\sqrt{\omega_0^2 - m^2}\right)^2 = (2i\omega)^2$$

avec $\omega = \sqrt{\omega_0^2 - m^2}$.
L'équation caractéristique a pour racines $-m \pm i\omega$.
La solution générale de l'équation différentielle est

$$y(x) = (\lambda\cos(\omega t) + \mu\sin(\omega t))\,\mathrm{e}^{-mt} = A\cos(\omega t - \varphi)\mathrm{e}^{-mt}$$

On parle de mouvement pseudo périodique de pseudo-période $T = \dfrac{2\pi}{\omega}$ avec $T > T_0$.
La solution correspondant aux conditions initiales $y(0) = y_0$ et $y'(0) = 0$ est

$$y(t) = y_0\left(\cos(\omega t) + \frac{m}{\omega}\sin(\omega t)\right))\mathrm{e}^{-mt}$$



22.2.5.3 Cas $m > 0$ et $\Delta = 0$, amortissement critique

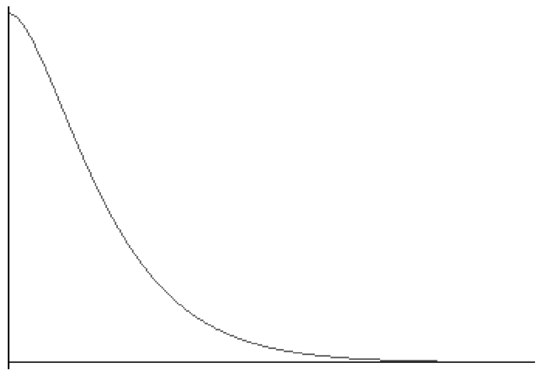
Si $m = \omega_0$ alors

$$\Delta = 0$$

L'équation caractéristique a pour racine double $-m$
La solution générale de l'équation différentielle est

$$y(t) = (\lambda t + \mu)\,\mathrm{e}^{-mt}$$

On parle de mouvement apériodique critique
La solution correspondant aux conditions initiales $y(0) = y_0$ et $y'(0) = 0$ est $y(t) = y_0(mt + 1)\mathrm{e}^{-mt}$.



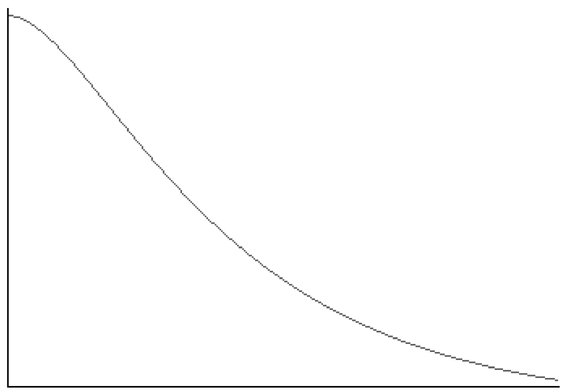
22.2.5.4 Cas $m > 0$ et $\Delta > 0$, amortissement fort

Si $m > \omega_0$ alors

$$\Delta > 0$$

L'équation caractéristique a pour racines $-m \pm \omega$ avec $\omega = \sqrt{\Delta}/2$.
La solution générale de l'équation différentielle est
$y(t) = (\lambda \mathrm{e}^{\omega t} + \mu \mathrm{e}^{-\omega t})\mathrm{e}^{-mt} = (\alpha \mathrm{ch}(\omega t) + \beta \mathrm{sh}(\omega t))\,\mathrm{e}^{-mt}$
On parle de mouvement apériodique.
La solution correspondant aux conditions initiales $y(0) = y_0$ et $y'(0) = 0$ est

$$y(t) = y_0 \left(\mathrm{ch}(\omega t) + \frac{\omega}{m} \mathrm{sh}(\omega t) \right) \mathrm{e}^{-mt}$$



22.2.6 Oscillateurs linéaires forcés

On se propose d'étudier l'équation différentielle

$$y'' + 2my' + \omega_0^2 y = A\cos(\omega t)$$

avec m, ω_0, A, ω des paramètres strictement positifs.
A est l'amplitude forcée et ω la pulsation forcée.

Exemple Cette équation s'obtient lors de l'étude d'un circuit RLC série en régime sinusoïdal forcé.

Cette équation est une équation différentielle linéaire d'ordre 2 dont l'équation homogène a été résolue ci-dessus. Cherchons maintenant une solution particulière et pour cela déterminons une solution particulière à l'équation complexe $z'' + 2mz' + \omega_0^2 z = A\mathrm{e}^{i\omega t}$ de la forme $z(t) = B\mathrm{e}^{i\omega t}$.
L'équation

$$z'' + 2mz' + \omega_0^2 z = A\mathrm{e}^{i\omega t}$$

est vérifiée si

$$(-\omega^2 + 2im\omega + \omega_0^2)B = A$$

i.e. pour

$$B = \frac{A}{\omega_0^2 - \omega^2 + 2im\omega}$$

On peut écrire $B = |B|\,\mathrm{e}^{-i\varphi}$ avec

$$|B| = \frac{A}{\sqrt{(\omega_0^2 - \omega^2)^2 + 4m^2\omega^2}}$$

et

$$\mathrm{e}^{i\varphi} = \frac{(\omega_0^2 - \omega^2) + 2im\omega}{\sqrt{(\omega_0^2 - \omega^2) + 4m^2\omega^2}}$$

On obtient alors

$$z(t) = \frac{A}{\sqrt{(\omega_0^2 - \omega)^2 + 4m^2\omega^2}}\mathrm{e}^{i(\omega t - \varphi)}$$

solution particulière de l'équation complexe puis

$$y(t) = \frac{A}{\sqrt{(\omega_0^2 - \omega)^2 + 4m^2\omega^2}}\cos(\omega t - \varphi)$$

solution particulière de l'équation initiale étudiée.

Remarque Lorsque $m > 0$, la solution homogène tend vers 0 quand $t \to +\infty$ et donc, pour t assez grand, c'est la solution particulière ci-dessus obtenue qui demeure ; celle-ci ne dépend pas des conditions initiales, on parle alors de régime permanent (ou forcé). Préalablement, on dit que le régime est transitoire.

Remarque L'amplitude du régime permanent

$$\frac{A}{\sqrt{(\omega_0^2 - \omega)^2 + 4m^2\omega^2}}$$

est maximale quand la quantité

$$(\omega_0^2 - \omega)^2 + 4m^2\omega^2 = \omega^4 + 2(2m^2 - \omega_0^2)\omega^2 + \omega_0^4$$

est minimale.
Si $2m^2 - \omega_0^2 \geqslant 0$ alors l'amplitude est maximale pour $\omega = 0$.
Si $2m^2 - \omega_0^2 < 0$ alors l'amplitude est maximale pour $\omega_r = \sqrt{\omega_0^2 - 2m^2}$; cette quantité est appelée pulsation de résonance.

22.3 Equation linéaire du second ordre

22.3.1 Définition

Soient $a, b, c : I \to \mathbb{K}$ trois fonctions continues.

Définition

On appelle équation différentielle linéaire d'ordre 2 définie sur I toute équation différentielle de la forme

$$y'' + a(x)y' + b(x)y = c(x)$$

avec $a, b, c : I \to \mathbb{K}$ fonctions continues.

Exemple $y'' + (x^2 + 1)y + (x - 1)y = 1$ est une équation différentielle linéaire d'ordre 2 définie sur $\mathbb{R}$.

Attention : On ne peut pas introduire d'équation caractéristique pour résoudre une telle équation et il n'y a pas de démarche générale à connaître.
Dans la pratique, une telle équation se résout par un changement de fonction inconnue ou un changement de variable ramenant le problème à une situation que l'on sait résoudre.

22.3.2 Changement de fonction inconnue

Exemple Résolvons sur $\mathbb{R}$ l'équation

$$E : (1 + x^2)y'' + 4xy' + (1 - x^2)y = 0$$

via le changement de fonction inconnue $z(x) = (1 + x^2)y(x)$.
Soient y une fonction deux fois dérivable et z la fonction définie par $z(x) = (1 + x^2)y(x)$.
La fonction z est deux fois dérivable et

$$y(x) = \frac{z(x)}{1 + x^2}, y'(x) = \frac{z'(x)}{1 + x^2} - \frac{2xz(x)}{(1 + x^2)^2}$$

et

$$y''(x) = \frac{z''(x)}{1 + x^2} - \frac{4xz'(x)}{(1 + x^2)^2} - \frac{2z(x)}{(1 + x^2)^2} + \frac{8x^2z(x)}{(1 + x^2)^3}$$

Par suite, y est solution de E sur $\mathbb{R}$ si, et seulement si,

$$\forall x \in \mathbb{R}, z''(x) - z(x) = 0$$

Résolvons l'équation $z'' - z = 0$.
C'est une équation différentielle linéaire d'ordre 2 homogène d'équation caractéristique $r^2 - 1 = 0$ et de solution générale : $z(x) = \lambda e^x + \mu e^{-x}$ avec λ, μ parcourant $\mathbb{K}$.
Par suite, la solution générale de E est

$$y(x) = \frac{\lambda e^x + \mu e^{-x}}{1 + x^2} \text{ avec } \lambda, \mu \text{ parcourant } \mathbb{K}$$

Exemple Résolvons sur $\mathbb{R}$ l'équation

$$E : (1 + e^x)y'' + (2e^x + 1)y' + e^x y = 0$$

via le changement de fonction inconnue $z = y' + y$.
Soient y une fonction deux fois dérivable sur $\mathbb{R}$ et z la fonction définie par $z = y' + y$.
La fonction z est dérivable et $z' = y'' + y'$.
Par suite la fonction y est solution de E si, et seulement si,

$$\forall x \in \mathbb{R}, (1 + e^x)z'(x) + e^x z(x) = 0$$

Résolvons l'équation $(1 + e^x)z' + e^x z = 0$.
C'est une équation différentielle linéaire d'ordre 1 homogène.

$$(1 + e^x)z' + e^x z = 0 \Leftrightarrow z' = -\frac{e^x}{1 + e^x} z$$

et

$$\int -\frac{e^x}{1 + e^x}\, dx = -\ln(1 + e^x)$$

La solution générale de l'équation $(1 + e^x)z' + e^x z = 0$ est

$$z(x) = \frac{\lambda}{1 + e^x} \text{ avec } \lambda \text{ parcourant } \mathbb{K}$$

La fonction y se soit alors comme solution d'une équation différentielle

$$y' + y = \frac{\lambda}{1 + e^x}$$

C'est une équation différentielle linéaire d'ordre 1.
Solution homogène $y(x) = \mu e^{-x}$.
Solution particulière :
On recherche celle-ci de la forme $y(x) = \mu(x)e^{-x}$ avec $x \mapsto \mu(x)$ fonction dérivable.

$$y'(x) + y(x) = \mu'(x)e^{-x} = \frac{\lambda}{1 + e^x} \Leftrightarrow \mu'(x) = \frac{\lambda e^x}{1 + e^x}$$

$\mu(x) = \lambda \ln(1 + e^x)$ convient et donc $y(x) = \lambda e^{-x} \ln(1 + e^x)$ est solution particulière.
Solution générale de $y' + y = \dfrac{\lambda}{1 + e^x}$

$$y(x) = (\lambda \ln(1 + e^x) + \mu)e^{-x} \text{ avec } \mu \text{ parcourant } \mathbb{R}$$

Solution générale de E

$$y(x) = (\lambda \ln(1 + e^x) + \mu)e^{-x} \text{ avec } \lambda, \mu \text{ parcourant } \mathbb{R}$$

22.3.3 Changement de variable

Exemple Résolvons sur $]0,+\infty[$ l'équation

$$E : x^2y'' - 2xy' + 2y = \ln x$$

via le changement de variable $t = \ln x$.
L'application $x \mapsto t = \ln x$ est une bijection de $]0,+\infty[$ vers $\mathbb{R}$ dont l'application réciproque est l'application $t \mapsto x = \mathrm{e}^t$.
Soient $y : x \to y(x)$ une fonction deux fois dérivable définie sur $]0,+\infty[$ et $z : t \mapsto z(t)$ définie de sorte que « $z(t) = y(x)$ »i.e. z la fonction définie par $z(t) = y(\mathrm{e}^t)$. La fonction z est deux fois dérivable.
Pour tout $x > 0$,

$$y(x) = z(\ln x), y'(x) = \frac{1}{x}z'(\ln x)$$

et

$$y''(x) = -\frac{1}{x^2}z'(\ln x) + \frac{1}{x^2}z''(\ln x)$$

Par suite, y est solution de E sur $]0,+\infty[$ si, et seulement si,

$$\forall x > 0, z''(\ln x) - 3z'(\ln x) + 2z(\ln x) = \ln x$$

i.e. si, et seulement si, z est solution de l'équation différentielle

$$z''(t) - 3z'(t) + 2z(t) = t$$

La résolution de cette équation différentielle linéaire d'ordre 2 à coefficients constants donne pour solution générale

$$z(t) = \frac{1}{2}(t+3) + \lambda\mathrm{e}^t + \mu\mathrm{e}^{2t}$$

On en déduit que la solution générale de E est

$$y(x) = \frac{1}{2}\left(\ln x + 3\right) + \lambda x + \mu x^2 \text{ avec } \lambda, \mu \text{ parcourant } \mathbb{K}$$

Attention : Lorsqu'on dérive les relations du type $y(x) = z(t)$ il faut prendre soin de dériver de part et d'autre par rapport à la même variable x ou t. L'emploi de notations $\frac{\mathrm{d}y}{\mathrm{d}x}$ et $\frac{\mathrm{d}z}{\mathrm{d}t}$ peut s'avérer pertinente.

Exemple Résolvons sur $\mathbb{R}$ l'équation

$$E : (1+x^2)^2y'' + 2x(1+x^2)y' + y = 0$$

via le changement de variable $t = \arctan x$.
L'application $x \mapsto t = \arctan x$ réalise une bijection de $\mathbb{R}$ vers $I =]-\pi/2, \pi/2[$ dont l'application réciproque est $t \mapsto x = \tan t$.
Soient $y : x \mapsto y(x)$ une fonction définie et deux fois dérivable sur $\mathbb{R}$ et $z : t \mapsto z(t)$ la fonction définie de sorte que « $z(t) = y(x)$ »i.e. z la fonction définie sur I par $z(t) = y(\tan t)$.
La fonction z est deux fois dérivable et pour tout $x \in \mathbb{R}$,

$$y(x) = z(\arctan x), y'(x) = \frac{z'(\arctan x)}{1+x^2}$$

et
$$y''(x) = -\frac{2xz'(\arctan x)}{(1+x^2)^2} + \frac{z''(\arctan x)}{(1+x^2)^2}$$
Par suite y est solution de E sur $\mathbb{R}$ si, et seulement si,
$$\forall x \in \mathbb{R}, z''(\arctan x) + z(\arctan x) = 0$$
i.e. si, et seulement si, z est solution sur I de l'équation différentielle
$$z''(t) + z(t) = 0$$
La résolution de cette équation différentielle linéaire d'ordre 2 à coefficients constants donne pour solution générale
$$z(t) = \lambda \cos t + \mu \sin t$$
On en déduit que la solution générale de E
$$y(x) = \frac{\lambda + \mu x}{\sqrt{1+x^2}} \text{ avec } \lambda, \mu \text{ parcourant } \mathbb{K}$$

Chapitre 23

Fonctions de deux variables réelles

Dans ce chapitre, on étudie les notions de continuité et de dérivabilité (partielle) des fonctions de deux variables réelles. La théorie et les résultats se généralisent aux fonctions de 3 voire n variables.
On note $(.\mid .)$ le produit scalaire canonique sur $\mathbb{R}^2$ et $\|\,.\,\|$ la norme euclidienne associée.
Les éléments de $\mathbb{R}^2$ seront, selon l'interprétation voulu, appelés points ou vecteurs ; par exemple le couple $(0,0)$ est appelé point O ou vecteur nul.

23.1 Limite et continuité des fonctions de deux variables réelles

23.1.1 Fonctions réelles de deux variables réelles

Définition

On appelle fonction réelle de deux variables définies sur une partie X de $\mathbb{R}^2$ toute application

$$f : \begin{cases} X \subset \mathbb{R}^2 \to \mathbb{R} \\ x = (x_1, x_2) \mapsto f(x) = f(x_1, x_2) \end{cases}$$

L'ensemble de ces fonctions est noté $\mathcal{F}(X, \mathbb{R})$ et est muni d'une structure d'anneau et de $\mathbb{R}$-espace vectoriel pour les lois usuelles.

Remarque Les fonctions de deux variables réelles peuvent se visualiser comme opérant sur deux variables réelles x_1, x_2 (c'est l'expression $f(x_1, x_2)$) ou bien comme opérant sur un couple x élément de $\mathbb{R}^2$ (c'est l'expression $f(x)$).

Remarque Les deux variables réelles sont fréquemment notées x_1, x_2 dans la théorie ou x, y dans la pratique. Bien évidemment d'autres notations sont tout à fait possibles et on saura s'adapter à ces dernières.

Exemple La fonction

$$f : (x, y) \mapsto \frac{x \sin(xy)}{x^2 + y^2}$$

est une fonction réelle de deux variables réelles définies sur $\mathbb{R}^2 \setminus \{(0,0)\}$.

Exemple Les applications $p_1, p_2 : \mathbb{R}^2 \to \mathbb{R}$ définies par $p_1(x_1, x_2) = x_1$ et $p_2(x_1, x_2) = x_2$ sont des fonctions de deux variables appelées première et deuxième projections canoniques.

Définition

On appelle fonction polynomiale réelle de deux variables définie sur une partie X de $\mathbb{R}^2$ toute application de la forme

$$p : \begin{cases} X \to \mathbb{R} \\ (x_1, x_2) \mapsto \displaystyle\sum_{p=0}^{m} \sum_{q=0}^{n} a_{p,q} x_1^p x_2^q \end{cases}$$

avec $m, n \in \mathbb{N}$ et $a_{p,q} \in \mathbb{R}$ pour tout p et q.

Exemple $(x, y) \mapsto x^2 - xy + y$ est une fonction polynomiale réelle définie sur $\mathbb{R}^2$.

Définition

On appelle fonction rationnelle réelle de deux variables définie sur une partie X de $\mathbb{R}^2$ toute application de la forme :

$$f : \begin{cases} X \to \mathbb{R} \\ (x_1, x_2) \mapsto \dfrac{p(x_1, x_2)}{q(x_1, x_2)} \end{cases}$$

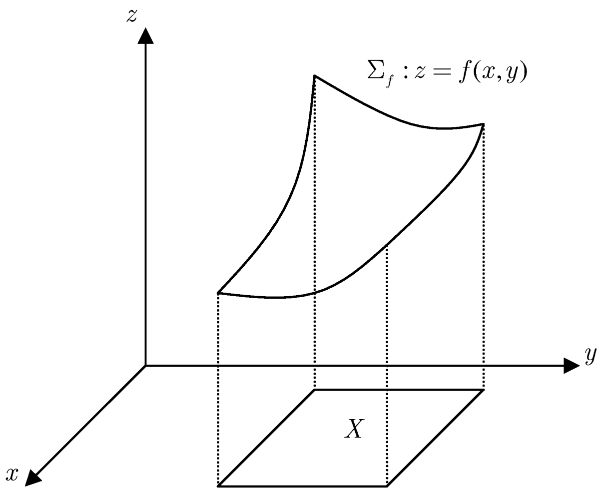
avec p, q des fonctions polynomiales telles que q ne s'annule pas sur X.

Exemple $(x, y) \mapsto \dfrac{x + y}{xy}$ est une fonction rationnelle réelle définie sur $\mathbb{R}^\star \times \mathbb{R}^\star$.

Définition

On appelle surface représentative d'une fonction $f : X \subset \mathbb{R}^2 \to \mathbb{R}$ l'ensemble

$$\Sigma_f = \{(x, y, z) \in X \times \mathbb{R} / z = f(x, y)\}$$



Remarque Tout comme pour les fonctions d'une variable réelles, on peut parler de fonction de deux variables réelles minorée, majorée, bornée et on peut aussi parler d'extremum d'une fonction de deux variables.
En revanche, il n'y a pas de notion de monotonie définie pour les fonctions de deux variables

Exemple Soit $f : \mathbb{R}^2 \setminus \{(0,0)\} \to \mathbb{R}$ la fonction rationnelle définie par

$$f(x, y) = \frac{xy}{x^2 + y^2}$$

La fonction f est bornée.
En effet, par l'inégalité bien connue $2|ab| \leqslant a^2 + b^2$, on obtient $|f(x, y)| \leqslant 1/2$ pour tout $(x, y) \in \mathbb{R}^2 \setminus \{(0,0)\}$

Exemple Soit $f : \mathbb{R}^2 \to \mathbb{R}$ la fonction polynomiale définie par

$$f(x, y) = x^2 + y^2$$

La fonction f admet un minimum en $(0, 0)$.
En effet $f(x, y) = x^2 + y^2 \geqslant 0 = f(0, 0)$ pour tout $(x, y) \in \mathbb{R}^2$.

23.1.2 Limite

Soit $a = (a_1, a_2) \in \mathbb{R}^2$.

23.1.2.1 Définition

Définition

Pour $\alpha > 0$, on note $D(a, \alpha)$ le disque de centre a et de rayon α,

$$D(a, \alpha) = \left\{x \in \mathbb{R}^2 / \, \|x - a\| \leqslant \alpha\right\}$$

Définition

On dit qu'une fonction $f : X \subset \mathbb{R}^2 \to \mathbb{R}$ est définie au voisinage de a si le domaine X de définition de f intercepte tous les disques centrés en a i.e.

$$\forall \alpha > 0, D(a, \alpha) \cap X \neq \emptyset$$

Exemple Une fonction $f : X \subset \mathbb{R}^2 \to \mathbb{R}$ est définie au voisinage de tout point élément de X.

Exemple Une fonction $f : \mathbb{R}^2 \setminus \{a\} \to \mathbb{R}$ est définie au voisinage de a.

Exemple Une fonction $f : \mathbb{R}^{+\star} \times]\alpha, \beta[\to \mathbb{R}$ est définie au voisinage de tout élément de $\mathbb{R}^+ \times [\alpha, \beta]$

Définition

On suppose que $f : X \subset \mathbb{R}^2 \to \mathbb{R}$ est une fonction définie au voisinage de a.
On dit que f tend vers $\ell \in \mathbb{R}$ en a si

$$\forall \varepsilon > 0, \exists \alpha > 0, \forall x \in X, \|x - a\| \leqslant \alpha \Rightarrow |f(x) - \ell| \leqslant \varepsilon$$

On dit que f tend vers $+\infty$ en a si

$$\forall M \in \mathbb{R}, \exists \alpha > 0, \forall x \in X, \|x - a\| \leqslant \alpha \Rightarrow f(x) \geqslant M$$

On dit que f tend vers $-\infty$ en a si

$$\forall M \in \mathbb{R}, \exists \alpha > 0, \forall x \in X, \|x - a\| \leqslant \alpha \Rightarrow f(x) \leqslant M$$

Lorsque f tend vers $\ell \in \bar{\mathbb{R}}$ en a, on dit que f admet une limite en a et on note :

$$f \underset{a}{\to} \ell, f(x) \xrightarrow[x \to a]{} \ell \text{ ou encore } f(x_1, x_2) \xrightarrow[(x_1, x_2) \to a]{} \ell$$

Remarque Ces définitions sont analogues à celles proposées pour définir les limites des fonctions réelles d'une variable réelle.

Exemple Soit $f : \mathbb{R}^2 \to \mathbb{R}$ fonction constante égale à C.
On a $f(x) \xrightarrow[x\to a]{} C$.
En effet, pour tout $\varepsilon > 0$, en prenant $\alpha = 123 > 0$, on a

$$\forall x \in \mathbb{R}^2, \|x - a\| \leqslant \alpha \Rightarrow |f(x) - C| = 0 \leqslant \varepsilon$$

Exemple Soit $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x) = \|x - a\|$.
On a $f(x) \xrightarrow[x\to a]{} 0$.
En effet, pour tout $\varepsilon > 0$, en prenant $\alpha = \varepsilon > 0$, on a

$$\forall x \in \mathbb{R}^2, \|x - a\| \leqslant \alpha \Rightarrow |f(x) - 0| = \|x - a\| \leqslant \varepsilon$$

Exemple Soit $f : \mathbb{R}^2 \backslash \{a\} \to \mathbb{R}$ définie par $f(x) = \dfrac{1}{\|x - a\|}$
On a $f(x) \xrightarrow[x\to a]{} +\infty$.
En effet, pour tout $M > 0$, en prenant $\alpha = 1/M > 0$, on a

$$\forall x \in \mathbb{R}^2 \backslash \{a\}, \|x - a\| \leqslant \alpha \Rightarrow f(x) = \frac{1}{\|x - a\|} \geqslant \frac{1}{\alpha} = M$$

Remarque En cas d'ambiguïté sur le domaine de définition X de la fonction f dont on considère la limite, on écrit $f(x) \xrightarrow[x\to a, x\in X]{} \ell$ au lieu de $f(x) \xrightarrow[x\to a]{} \ell$.

Proposition

> Soient $f : X = X' \cup X'' \subset \mathbb{R}^2 \to \mathbb{R}$ et $\ell \in \bar{\mathbb{R}}$.
> Si $f(x) \xrightarrow[x\to a, x\in X']{} \ell$ et $f(x) \xrightarrow[x\to a, x\in X'']{} \ell$ alors $f(x) \xrightarrow[x\to a, x\in X]{} \ell$.

dém. :
Soit $\varepsilon > 0$. Il existe $\alpha', \alpha'' > 0$ vérifiant

$$\forall x \in X', \|x - a\| \leqslant \alpha' \Rightarrow |f(x) - \ell| \leqslant \varepsilon \text{ et } \forall x \in X'', \|x - a\| \leqslant \alpha'' \Rightarrow |f(x) - \ell| \leqslant \varepsilon$$

Pour $\alpha = \min(\alpha', \alpha'') > 0$, on a alors

$$\forall x \in X = X' \cup X'', \|x - a\| \leqslant \alpha \Rightarrow |f(x) - \ell| \leqslant \varepsilon$$

□

23.1.2.2 Propriétés

Théorème

> Soit $f : X \subset \mathbb{R}^2 \to \mathbb{R}$ définie au voisinage de $a \in \mathbb{R}^2$.
> Si $f \underset{a}{\to} \ell \in \bar{\mathbb{R}}$ et si $f \underset{a}{\to} \ell' \in \bar{\mathbb{R}}$ alors $\ell = \ell'$.

dém. :
Cas $\ell, \ell' \in \mathbb{R}$
Soit $\varepsilon > 0$. Il existe $\alpha, \alpha' > 0$ tels que

$$\forall x \in X, \|x - a\| \leqslant \alpha \Rightarrow \|f(x) - \ell\| \leqslant \varepsilon \text{ et } \forall x \in X, \|x - a\| \leqslant \alpha' \Rightarrow \|f(x) - \ell'\| \leqslant \varepsilon$$

Considérons un $x \in X$ tel que $\|x - a\| \leqslant \min(\alpha, \alpha')$, il en existe car f est supposée définie au voisinage de a.
Pour cet x, on a à la fois

$$|f(x) - \ell| \leqslant \varepsilon \text{ et } |f(x) - \ell'| \leqslant \varepsilon$$

On en déduit $|\ell - \ell'| \leqslant 2\varepsilon$ et puisque ceci vaut pour tout $\varepsilon > 0$, on obtient $\ell = \ell'$.
Cas $\ell \in \mathbb{R}$ et $\ell' = +\infty$.
Soient $\varepsilon = 1/2$ et $M = |\ell| + 1$.
Il existe $\alpha, \alpha' > 0$ tels que

$$\forall x \in X, \|x - a\| \leqslant \alpha \Rightarrow |f(x) - \ell| \leqslant \varepsilon \text{ et } \forall x \in X, \|x - a\| \leqslant \alpha' \Rightarrow f(x) \geqslant M$$

Considérons un $x \in X$ tel que $\|x - a\| \leqslant \min(\alpha, \alpha')$, il en existe car f est supposée définie au voisinage de a.
Pour cet x, on a à la fois

$$|f(x) - \ell| \leqslant 1/2 \text{ et } f(x) \geqslant |\ell| + 1$$

ce qui est absurde car $|f(x) - \ell| \leqslant 1/2$ entraîne $|f(x)| \leqslant |\ell| + 1/2$.
Les cas restants sont semblables.
□

Définition

Si $f \underset{a}{\to} \ell \in \bar{\mathbb{R}}$ alors on dit que ℓ est la limite de f en a et on note

$$\ell = \lim_a f,\ \ell = \lim_{x \to a} f(x) \text{ ou encore } \ell = \lim_{(x_1,x_2)\to(a_1,a_2)} f(x_1, x_2)$$

Théorème

Soit $f : X \subset \mathbb{R}^2 \to \mathbb{R}$ définie au voisinage de $a \in \mathbb{R}^2$ et $\ell \in \bar{\mathbb{R}}$.
On a équivalence entre :
(i) $f \underset{a}{\to} \ell$;
(ii) $\forall (x_n) \in X^{\mathbb{N}}, x_n \to a \Rightarrow f(x_n) \to \ell$.

dém. :
Cas $\ell \in \mathbb{R}$.
(i) $\Rightarrow$ (ii) Supposons que f tend vers ℓ en a. Soit (x_n) une suite d'éléments de X convergeant vers a.
Pour tout $\varepsilon > 0$, il existe $\alpha > 0$ tel que

$$\forall x \in X, \|x - a\| \leqslant \alpha \Rightarrow |f(x) - \ell| \leqslant \varepsilon$$

Puisque la suite (x_n) converge vers a, il existe un rang $N \in \mathbb{N}$ tel que

$$\forall n \geqslant N, \|x_n - a\| \leqslant \alpha$$

et alors

$$\forall n \in \mathbb{N}, |f(x_n) - \ell| \leqslant \varepsilon$$

On peut alors affirmer que la suite $(f(x_n))$ converge vers ℓ.
(ii) $\Rightarrow$ (i) Par contraposée. Supposons que f ne tend vers ℓ en a. Il existe alors $\varepsilon > 0$ tel que

$$\forall \alpha > 0, \exists x \in X, \|x - a\| \leqslant \alpha \text{ et } |f(x) - \ell| > \varepsilon$$

Pour $n \in \mathbb{N}$, en prenant $\alpha = 1/(n+1) > 0$, on peut affirmer qu'il existe $x_n \in X$ vérifiant

$$\|x_n - a\| \leqslant 1/(n+1) \text{ et } |f(x_n) - \ell| \geqslant \varepsilon$$

En faisant varier n, ceci détermine une suite (x_n) d'éléments de X convergeant vers a mais dont l'image par f ne tend pas vers ℓ.
Les cas $\ell = +\infty$ et $\ell = -\infty$ sont analogues.
□

Proposition

Soient $f, g, h : X \subset \mathbb{R}^2 \to \mathbb{R}$ définies au voisinage de $a \in \mathbb{R}^2$ et $\ell \in \mathbb{R}$.
Si $|f(x) - \ell| \leqslant g(x)$ et $g \underset{a}{\to} 0$ alors $f \underset{a}{\to} \ell$.
Si $g \leqslant f \leqslant h$ et si $g \underset{a}{\to} \ell$ et $h \underset{a}{\to} \ell$ alors $f \underset{a}{\to} \ell$.
Si $f \leqslant g$ et $f \underset{a}{\to} +\infty$ alors $g \underset{a}{\to} +\infty$.
Si $f \leqslant g$ et $g \underset{a}{\to} -\infty$ alors $f \underset{a}{\to} -\infty$.

dém. :
On peut procéder de façon élémentaire ou encore exploiter la caractérisation séquentielle des limites et les résultats analogues relatifs aux suites réelles.
□

Proposition

Soient $f, g : X \subset \mathbb{R}^2 \to \mathbb{R}$ définie au voisinage de $a \in \mathbb{R}^2$ telles que $f \underset{a}{\to} \ell \in \bar{\mathbb{R}}$ et $g \underset{a}{\to} \ell' \in \bar{\mathbb{R}}$.
Si $\ell + \ell'$ est définie dans $\bar{\mathbb{R}}$ alors $f + g \underset{a}{\to} \ell + \ell'$.
Si $\ell\ell'$ est définie dans $\bar{\mathbb{R}}$ alors $fg \underset{a}{\to} \ell\ell'$. $\bar{\mathbb{R}}$

dém. :
On peut procéder de façon élémentaire ou encore exploiter la caractérisation séquentielle des limites et les résultats analogues relatifs aux suites réelles.
□

Proposition

Soient

$$f : \begin{cases} X \subset \mathbb{R}^2 \to \mathbb{R} \\ (x_1, x_2) \mapsto f(x_1, x_2) \end{cases} \text{ et } \varphi : \begin{cases} \mathcal{D} \subset \mathbb{R} \to \mathbb{R} \\ t \to \varphi(t) \end{cases}$$

telles que $f(X) \subset \mathcal{D}$
Si $f(x) \xrightarrow[x \to a]{} \ell \in \bar{\mathbb{R}}$ et si $\varphi(t) \xrightarrow[t \to \ell]{} \ell' \in \bar{\mathbb{R}}$ alors $(\varphi \circ f)(x) \xrightarrow[x \to a]{} \ell'$.

dém. :
On peut procéder de façon élémentaire ou encore exploiter la caractérisation séquentielle des limites et les résultats analogues relatifs aux suites réelles.
□

Proposition

Soient

$$\gamma : \begin{cases} \mathcal{D} \subset \mathbb{R} \to \mathbb{R}^2 \\ t \to (x_1(t), x_2(t)) \end{cases} \text{ et } f : \begin{cases} X \subset \mathbb{R}^2 \to \mathbb{R} \\ (x_1, x_2) \mapsto f(x_1, x_2) \end{cases}$$

telles que $\gamma(\mathcal{D}) \subset X$

Si $\gamma(t) \xrightarrow[t \to t_0]{} a$ et $f(x) \xrightarrow[x \to a]{} \ell \in \bar{\mathbb{R}}$ alors $(f \circ \gamma)(t) \xrightarrow[t \to t_0]{} \ell$.

dém. :

On peut procéder de façon élémentaire ou encore exploiter la caractérisation séquentielle des limites et les résultats analogues relatifs aux suites réelles.

□

23.1.2.3 Exemples d'étude

Soient $p_1 : (x_1, x_2) \mapsto x_1$ et $p_2 : (x_1, x_2) \mapsto x_2$ définies sur $\mathbb{R}^2$.

En tout $a = (a_1, a_2) \in \mathbb{R}^2$, on a $p_1 \underset{a}{\to} a_1$ et $p_2 \underset{a}{\to} a_2$ car

$$|p_1(x) - a_1| = |x_1 - a_1| \leqslant \|x - a\| \xrightarrow[x \to a]{} 0 \text{ et } |p_2(x) - a_2| = |x_2 - a_2| \leqslant \|x - a\| \xrightarrow[x \to a]{} 0$$

Autrement dit , on a

$$x_1 \xrightarrow[x \to a]{} a_1 \text{ et } x_2 \xrightarrow[x \to a]{} a_2$$

Exemple Etudions

$$\lim_{(x,y) \to (0,0)} \sqrt{x^2 + xy + y^2}$$

Quand $(x, y) \to (0, 0)$.

En vertu de ce qui précède, on peut écrire $x \to 0$ et $y \to 0$.

Par opérations algébriques

$$x^2 + xy + y^2 \to 0$$

Par composition de limites

$$\sqrt{x^2 + xy + y^2} \to 0$$

On en déduit $\lim\limits_{(x,y) \to (0,0)} \sqrt{x^2 + xy + y^2} = 0$

Attention : Cette démarche ne consiste pas à faire tendre x vers 0 puis y vers 0 ce qui serait incorrect. Ici, c'est le couple (x, y) qui tend vers $(0, 0)$.

Exemple Etudions

$$\lim_{(x,y) \to (0,0)} \frac{1}{x^2 + y^2}$$

Quand $(x, y) \to (0, 0)$ avec $(x, y) \neq (0, 0)$, on a $x \to 0$, $y \to 0$ et donc $x^2 + y^2 \to 0$ avec $x^2 + y^2 > 0$.

Or $1/t \xrightarrow[t \to 0^+]{} +\infty$ donc par composition de limite $\dfrac{1}{x^2 + y^2} \to +\infty$.

Ainsi $\lim\limits_{(x,y) \to (0,0)} \dfrac{1}{x^2 + y^2} = +\infty$.

Exemple Etudions
$$\lim_{(x,y)\to(0,0)} \frac{xy}{\sqrt{x^2+y^2}}$$

1ère méthode
Quand $(x,y)\to(0,0)$ avec $(x,y)\neq(0,0)$, on a
$$\left|\frac{xy}{\sqrt{x^2+y^2}}-0\right| \leqslant \frac{1}{2}\sqrt{x^2+y^2}\to 0$$

donc par comparaison $\lim\limits_{(x,y)\to(0,0)} \dfrac{xy}{\sqrt{x^2+y^2}}=0$
2ème méthode
Quand $(x,y)\to(0,0)$ avec $(x,y)\neq(0,0)$, on peut écrire en passant en coordonnées polaires $x=r\cos\theta$ et $y=r\sin\theta$ avec $r=\sqrt{x^2+y^2}\to 0^+$ et θ bien choisi.
On a alors
$$\frac{xy}{\sqrt{x^2+y^2}}=r\cos\theta\sin\theta\to 0$$

On retrouve $\lim\limits_{(x,y)\to(0,0)} \dfrac{xy}{\sqrt{x^2+y^2}}=0$

Exemple Etudions
$$\lim_{(x,y)\to(0,0)} \frac{x^2+y^2}{x^4+y^4}$$

1ère méthode
Quand $(x,y)\to(0,0)$ avec $(x,y)\neq(0,0)$, on a
$$\frac{x^2+y^2}{x^4+y^4}\geqslant\frac{x^2+y^2}{(x^2+y^2)^2}\geqslant\frac{1}{x^2+y^2}\to+\infty$$

donc par comparaison $\lim\limits_{(x,y)\to(0,0)} \dfrac{x^2+y^2}{x^4+y^4}=+\infty$
2ème méthode
Quand $(x,y)\to(0,0)$ avec $(x,y)\neq(0,0)$, on peut écrire en passant en coordonnées polaires $x=r\cos\theta$ et $y=r\sin\theta$ avec $r=\sqrt{x^2+y^2}\to 0^+$ et θ bien choisi.
On a alors
$$\frac{x^2+y^2}{x^4+y^4}=\frac{1}{r^2}\frac{1}{\cos^4\theta+\sin^4\theta}\geqslant\frac{1}{2r^2}\to+\infty$$

On retrouve $\lim\limits_{(x,y)\to(0,0)} \dfrac{x^2+y^2}{x^4+y^4}=+\infty$.

Exemple Etudions
$$\lim_{(x,y)\to(0,0)} \frac{xy}{x^2+y^2}$$

Quand $(x,y)\to(0,0)$ avec $(x,y)\neq(0,0)$, on peut écrire en passant en coordonnées polaires $x=r\cos\theta$ et $y=r\sin\theta$ avec $r=\sqrt{x^2+y^2}\to 0^+$ et θ bien choisi

On a alors
$$\frac{xy}{x^2+y^2} = \cos\theta\sin\theta$$
qui ne semble pas avoir de limite...
Posons $f(x,y) = \dfrac{xy}{x^2+y^2}$ définie sur $\mathbb{R}^2 - \{(0,0)\}$.
On a
$$f(0,1/n) = 0 \to 0 \text{ et } f(1/n,1/n) \to 1/2$$
En vertu de la caractérisation séquentielle des limites, si f admet une limite en $(0,0)$ celle-ci doit valoir 0 et $1/2$; c'est absurde. On peut donc affirmer l'inexistence de la limite étudiée.

Exemple Etudions
$$\lim_{(x,y)\to(0,0)} \frac{x^2-y^2}{x^2+y^2}$$
Quand $(x,y)\to(0,0)$ avec $(x,y)\neq(0,0)$, on peut écrire en passant en coordonnées polaires $x = r\cos\theta$ et $y = r\sin\theta$ avec $r = \sqrt{x^2+y^2}\to 0^+$ et θ bien choisi
On a alors
$$\frac{x^2-y^2}{x^2+y^2} = \cos 2\theta$$
qui ne semble pas avoir de limite...
Posons $f(x,y) = \dfrac{x^2-y^2}{x^2+y^2}$ définie sur $\mathbb{R}^2 - \{(0,0)\}$.
On a $f(1/n,0) = 1 \to 1$ et $f(0,1/n) = -1 \to -1$
On peut affirmer l'inexistence de la limite étudiée.

Remarque Si on étudie une limite quand $(x,y)\to(a,b)$, on ramène le problème en $(0,0)$ par translation des variables, $x = a+h, y = b+k$ avec $(h,k)\to(0,0)$. Comme dans les exemples précédents, on peut alors cerner les ordres de grandeur en passant en coordonnées polaires $h = r\cos\theta$, $k = r\sin\theta$ avec $r = \sqrt{h^2+k^2}\to 0$.

23.1.3 Continuité

23.1.3.1 Définition

Définition

> On dit que $f : X\subset\mathbb{R}^2\to\mathbb{R}$ est continue en un point $a\in X$ si $f \underset{a}{\to} f(a)$.
> On dit que f est continue sur X si f est continue en tout point $a\in X$.
> On note $\mathcal{C}(X,\mathbb{R})$ l'ensemble des fonctions réelles définies et continues sur X.

Exemple Les fonctions constantes sont continues.

Exemple Les projections canoniques p_1 et p_2 sont continues sur $\mathbb{R}^2$.

23.1.3.2 Opérations

Proposition

Soient $f, g : X \to \mathbb{R}$ et $\lambda \in \mathbb{R}$.
Si f et g sont continues alors $\lambda.f, f+g, fg$ le sont aussi.
Ainsi $\mathcal{C}(X, \mathbb{R})$ est un sous-espace vectoriel et un sous anneau de $\mathcal{F}(X, \mathbb{R})$.

dém. :
Par opérations sur les limites en tout point $a \in X$.
□

Exemple Les fonctions $\sigma : \mathbb{R}^2 \to \mathbb{R}$ et $\pi : \mathbb{R}^2 \to \mathbb{R}$ définie par

$$\sigma(x_1, x_2) = x_1 + x_2 \text{ et } \pi(x_1, x_2) = x_1 x_2$$

sont continues sur $\mathbb{R}^2$ car $\sigma = p_1 + p_2$ et $\pi = p_1 p_2$.

Exemple Plus généralement les fonctions polynomiales sont continues sur $\mathbb{R}^2$.

Proposition

Soient

$$f : \begin{cases} X \subset \mathbb{R}^2 \to \mathbb{R} \\ (x_1, x_2) \mapsto f(x_1, x_2) \end{cases} \text{ et } \varphi : \begin{cases} \mathcal{D} \subset \mathbb{R} \to \mathbb{R} \\ t \to \varphi(t) \end{cases}$$

telles que $f(X) \subset \mathcal{D}$
Si f et φ sont continues alors $\varphi \circ f : (x_1, x_2) \mapsto \varphi(f(x_1, x_2))$ est continue.

dém. :
Par opérations sur les limites en tout point $a \in X$.
□

Exemple Si $f : X \to \mathbb{R}$ est continue et ne s'annule pas sur X alors $1/f$ est continue.

Exemple Les fonctions rationnelles de deux variables sont continues.

Proposition

Soient

$$\gamma : \begin{cases} \mathcal{D} \subset \mathbb{R} \to \mathbb{R}^2 \\ t \to (x_1(t), x_2(t)) \end{cases} \text{ et } f : \begin{cases} X \subset \mathbb{R}^2 \to \mathbb{R} \\ (x_1, x_2) \mapsto f(x_1, x_2) \end{cases}$$

telles que $\gamma(\mathcal{D}) \subset X$
Si f et γ sont continues alors $f \circ \gamma : t \mapsto f(x_1(t), x_2(t))$ l'est aussi.

dém. :
Par opérations sur les limites en tout point $t \in \mathcal{D}$.
□

23.1.3.3 En pratique

Exemple L'application

$$(x, y) \mapsto \frac{xy + x \sin y}{x^2 + y^2 + 1}$$

est continue sur $\mathbb{R}^2$ par opérations sur les fonctions continues...

Exemple Montrons la continuité sur $\mathbb{R}^2$ de l'application

$$f : (x, y) \mapsto \begin{cases} (1 - x^2)(1 - y^2) & \text{si } (x, y) \in [-1, 1]^2 \\ 0 & \text{sinon} \end{cases}$$

Soit $a = (x_0, y_0) \in \mathbb{R}^2$.
Cas $a \in]-1, 1[^2$
Sur un disque centré en a de rayon $\alpha > 0$ suffisamment petit, on a $f(x, y) = (1 - x^2)(1 - y^2)$ ce qui permet d'affirmer que f est continue en a.
Cas $a \in \mathbb{R}^2 \setminus [-1, 1]^2$
Sur un disque centré en a de rayon $\alpha > 0$ suffisamment petit, on a $f(x, y) = 0$ ce qui permet à nouveau d'affirmer que f est continue en a.
Cas $a \in [-1, 1]^2 \setminus]-1, 1[^2$.
Quand $x \to a$ avec $a \in [-1, 1]^2$, on a

$$f(x, y) = (1 - x^2)(1 - y^2) \to (1 - x_0^2)(1 - y_0^2) = 0 = f(a)$$

et quand $x \to a$ avec $a \notin [-1, 1]^2$, on a

$$f(x, y) = 0 \to 0 = f(a)$$

On peut donc affirmer que $f(x)$ tend vers $f(a)$ quand $x \to a$ et donc encore une fois f est continue en a.

23.1.4 Applications partielles

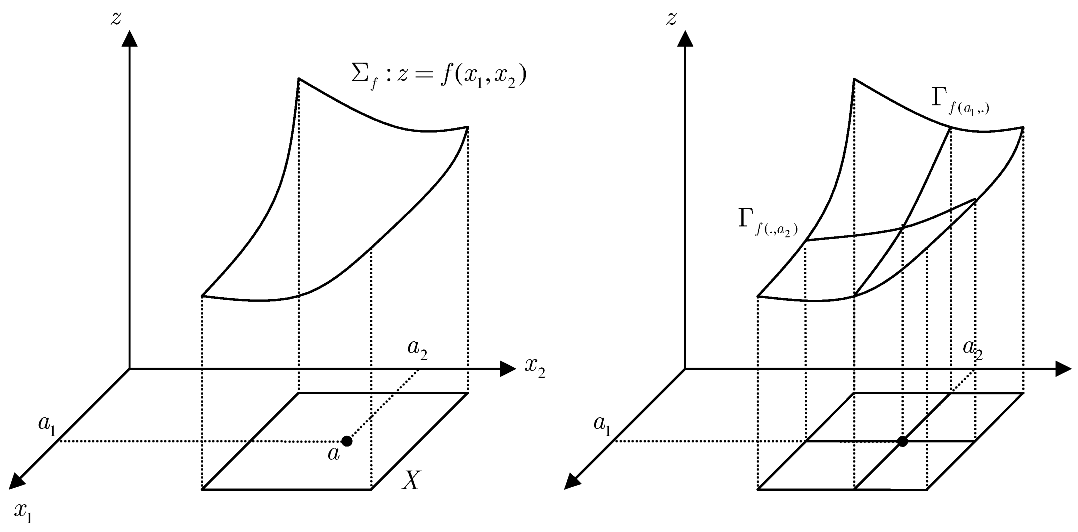
Définition

Soient $f : X \subset \mathbb{R}^2 \to \mathbb{R}$ et $a = (a_1, a_2) \in X$.
On appelle première application partielle de f au point a la fonction d'une variable réelle

$$f(., a_2) : x_1 \mapsto f(x_1, a_2)$$

On appelle deuxième application partielle de f au point a la fonction d'une variable réelle

$$f(a_1, .) : x_2 \mapsto f(a_1, x_2)$$



Remarque Sur cette figure, on perçoit que les applications partielles en un point ne déterminent par entièrement la fonction autour du point.

Proposition

Si $f : X \subset \mathbb{R}^2 \to \mathbb{R}$ est continue alors les applications partielles de f en tout point $a = (a_1, a_2) \in X$ sont continues.

dém. :
L'application partielle $f(., a_2)$ est continue par composition de l'application continue $t \mapsto (t, a_2)$ avec l'application f supposée continue. Il en est de même pour $f(a_1, .)$.
□

Attention : La réciproque est fausse ; la continuité des applications partielles n'assure pas la continuité de la fonction.

Exemple Considérons

$$f(x,y) = \begin{cases} \dfrac{(x+y)^2}{x^2+y^2} & \text{si } (x,y) \neq (0,0) \\ 1 & \text{sinon} \end{cases}$$

Toutes les applications partielles de f sont continues car

$$f(.,y) : x \mapsto \frac{(x+y)^2}{x^2+y^2} \text{ si } y \neq 0 \text{ et } f(.,0) : x \mapsto 1$$

$$f(x,.) : y \mapsto \frac{(x+y)^2}{x^2+y^2} \text{ si } x \neq 0 \text{ et } f(0,.) : y \mapsto 1$$

Néanmoins la fonction f n'est pas continue en $(0,0)$ car

$$f(1/n,1/n) = 2 \to 2 \neq 1 = f(0,0)$$

En fait, par la continuité des applications partielles de f, on connaît la continuité de f en tout point lorsqu'on s'en rapproche de façon horizontale ou verticale ; ceci ne permet d'affirmer la continuité au point lorsqu'on s'en rapproche de façon oblique.

23.1.5 Extension aux fonctions à valeurs dans R2

Soit

$$f : \begin{cases} X \subset \mathbb{R}^2 \to \mathbb{R}^2 \\ (x_1,x_2) \mapsto (f_1(x_1,x_2), f_2(x_1,x_2)) \end{cases}$$

Définition

Les fonctions réelles f_1, f_2 sont appelées fonctions coordonnées de f.

Exemple $f(x,y) = (x+y, x-y)$ définit une fonction de $\mathbb{R}^2$ vers $\mathbb{R}^2$ dont les fonctions coordonnées sont polynomiales.

Définition

Soit $a \in \mathbb{R}^2$ tel que f est définie au voisinage de a.
On dit que f converge vers $\ell \in \mathbb{R}^2$ en a et on note $f \underset{a}{\to} \ell$ si

$$\forall \varepsilon > 0, \exists \alpha > 0, \forall x \in X, \|x-a\| \leqslant \alpha \Rightarrow \|f(x) - \ell\| \leqslant \varepsilon$$

On démontre aisément qu'un tel ℓ est alors unique, on l'appelle limite de f en a.

Proposition

En écrivant $\ell = (\ell_1, \ell_2)$, on a équivalence entre :
(i) $f \underset{a}{\to} \ell$;
(ii) $\|f(x) - \ell\| \xrightarrow[x \to a]{} 0$;
(iii) $f_1 \underset{a}{\to} \ell_1$ et $f_2 \underset{a}{\to} \ell_2$.

dém. :
On a (i) $\Leftrightarrow$ (ii) car

$$\|f(x) - \ell\| \leqslant \varepsilon \Leftrightarrow |\|f(x) - \ell\| - 0| \leqslant \varepsilon$$

On a (ii) $\Rightarrow$ (iii) car

$$|f_1(x) - \ell_1| \leqslant \|f(x) - \ell\| \text{ et } |f_2(x) - \ell_2| \leqslant \|f(x) - \ell\|$$

et on a (iii) $\Rightarrow$ (ii) car

$$\|f(x) - \ell\| = \sqrt{(f_1(x) - \ell_1)^2 + (f_2(x) - \ell_2)^2}$$

□

Proposition

Soient

$$f : \begin{cases} X \subset \mathbb{R}^2 \to \mathbb{R}^2 \\ (x_1, x_2) \mapsto f(x_1, x_2) \end{cases}, g : \begin{cases} Y \subset \mathbb{R}^2 \to \mathbb{R} \\ (y_1, y_2) \mapsto g(y_1, y_2) \end{cases}$$

telles que $f(X) \subset Y$.
Si $f \underset{a}{\to} b$ et $g \underset{b}{\to} \ell \in \bar{\mathbb{R}}$ alors $g \circ f \underset{a}{\to} \ell$.

dém. :
Cas $\ell \in \mathbb{R}$
Soit $\varepsilon > 0$. Il existe $\beta > 0$ tel que

$$\forall y \in Y, \|y - b\| \leqslant \beta \Rightarrow |g(x) - \ell| \leqslant \varepsilon$$

et il existe $\alpha > 0$ tel que

$$\forall x \in X, \|x - a\| \leqslant \alpha \Rightarrow \|f(x) - b\| \leqslant \beta$$

On a alors

$$\forall x \in X, \|x - a\| \leqslant \alpha \Rightarrow\Rightarrow |g(f(x)) - \ell| \leqslant \varepsilon$$

Les cas $\ell = +\infty$ ou $\ell = -\infty$ sont analogues.
□

Remarque On peut aussi généralises ce résultat au cas où g serait à valeurs dans $\mathbb{R}^2$.

Définition

On dit que $f : X \subset \mathbb{R}^2 \to \mathbb{R}$ est continue en $a \in X$ si $f \underset{a}{\to} f(a)$.
On dit que f est continue si f l'est en tout point $a \in X$.
On note $\mathcal{C}(X, \mathbb{R}^2)$ l'ensemble des fonctions à valeurs dans $\mathbb{R}^2$ définies et continues sur X.

Proposition

La fonction $f : X \subset \mathbb{R}^2 \to \mathbb{R}^2$ est continue si ses fonctions coordonnées f_1 et f_2 le sont.

dém. :
Car $f(x) \underset{a}{\to} f(a)$ si, et seulement si, $f_1(x) \underset{a}{\to} f_1(a)$ et $f_2(x) \underset{a}{\to} f_2(a)$.
□

Exemple La fonction $f : (x, y) \mapsto (x + y, x - y)$ est continue sur $\mathbb{R}^2$ car ses fonctions coordonnées le sont.

Proposition

Soient

$$f : \begin{cases} X \subset \mathbb{R}^2 \to \mathbb{R}^2 \\ (x_1, x_2) \mapsto f(x_1, x_2) \end{cases} \text{ et } g : \begin{cases} B \subset \mathbb{R}^2 \to \mathbb{R} \\ (y_1, y_2) \mapsto g(y_1, y_2) \end{cases}$$

telles que $f(X) \subset B$
Si f et g sont continues alors $g \circ f$ l'est aussi.

dém. :
Par composition de limites.
□

Remarque Ce résultat se généralise évidemment au cas où g serait une fonction à valeurs dans $\mathbb{R}^2$.

23.2 Dérivées partielles

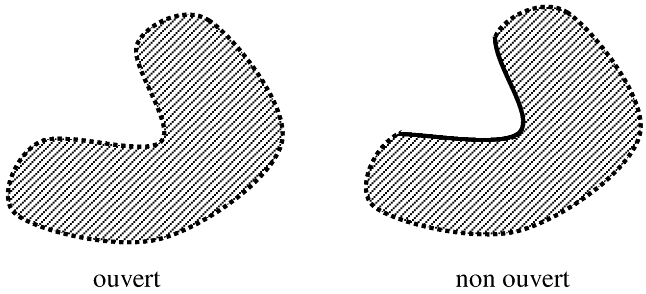
23.2.1 Ouvert de $\mathbb{R}^2$

Définition

Une partie U de $\mathbb{R}^2$ est dite ouverte si elle vérifie

$$\forall a \in U, \exists \alpha > 0, D(a, \alpha) \subset U$$

Exemple



Exemple Les parties $\mathbb{R}^2$ et $\emptyset$ sont ouvertes

Exemple Le produit cartésien de deux intervalles ouverts de $\mathbb{R}$ est une partie ouverte de $\mathbb{R}^2$.

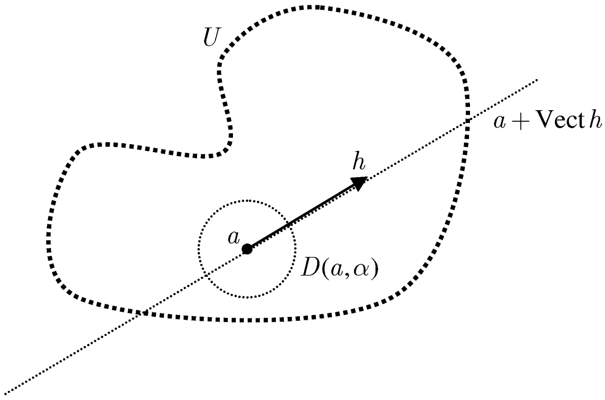
Désormais, U et V désignent des ouverts de $\mathbb{R}^2$ et I un intervalle non singulier de $\mathbb{R}$.

23.2.2 Dérivée selon un vecteur

Soient $a = (a_1, a_2) \in U$ et

$$f : \begin{cases} U \to \mathbb{R} \\ (x_1, x_2) \mapsto f(x_1, x_2) \end{cases} \text{ et } a = (a_1, a_2) \in U$$

Puisque U est ouvert, il existe $\alpha > 0$ tel que $D(a, \alpha) \subset U$.
Soit $h = (h_1, h_2)$ un vecteur de $\mathbb{R}^2$, l'application $t \mapsto f(a + t.h) = f(a_1 + th_1, a_2 + th_2)$ est définie au voisinage de 0.

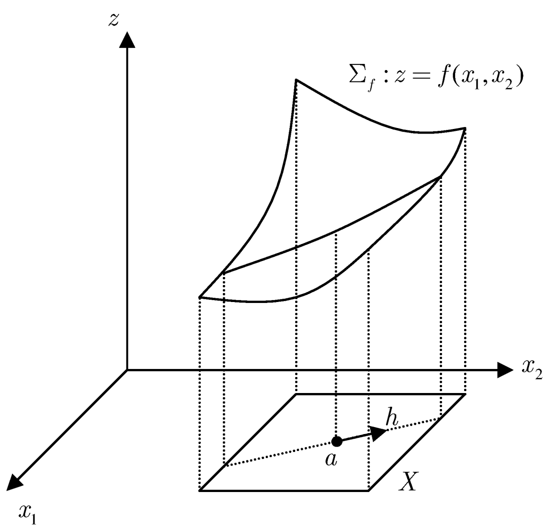


Définition

On dit que f est dérivable en a selon le vecteur h si l'application $t \mapsto f(a+t.h)$ est dérivable en 0.
Si tel est le cas, son nombre dérivé est appelé nombre dérivé de f en a selon le vecteur h et on le note $\mathrm{D}_h f(a)$. Ainsi et sous réserve d'existence

$$\mathrm{D}_h f(a) = \lim_{t \to 0} \frac{1}{t}\left(f(a+t.h) - f(a)\right)$$

Remarque Le calcul de $\mathrm{D}_h f(a)$ ne prend en compte que les valeurs de f sur la droite passant par a et dirigé par h.



Exemple Soient $f : (x_1, x_2) \mapsto x_1 x_2^2$, $a = (a_1, a_2)$ et $h = (h_1, h_2)$.

On a

$$\frac{1}{t}(f(a+t.h)-f(a)) = \frac{1}{t}\left((a_1+th_1)(a_2+th_2)^2 - a_1a_2^2\right) \xrightarrow[t\to 0]{} a_2^2h_1 + 2a_1a_2h_2$$

On en déduit que f est dérivable en a selon le vecteur h et

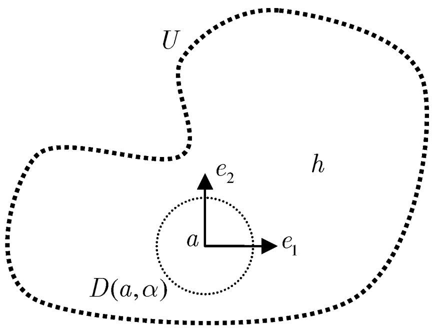
$$\mathrm{D}_h f(a) = a_2^2h_1 + 2a_1a_2h_2$$

23.2.3 Dérivées partielles

Soient $a = (a_1, a_2) \in U$ et

$$f : \begin{cases} U \to \mathbb{R} \\ (x_1, x_2) \mapsto f(x_1, x_2) \end{cases}$$

Notons $e_1 = (1,0)$ et $e_2 = (0,1)$ les vecteurs de la base canonique de $\mathbb{R}^2$.



Pour $i = 1$ ou 2.

Définition

Pour $i = 1$ ou 2 et sous réserve d'existence, on appelle nombre dérivé partiel de f en a selon sa i-ème variable, le nombre dérivé de f en a selon le vecteur e_i. Celui-ci est noté $\mathrm{D}_i f(a)$ au lieu de $\mathrm{D}_{e_i} f(a)$.

Remarque Ainsi et sous réserve d'existence, on a

$$\mathrm{D}_i f(a) = \lim_{t\to 0} \frac{1}{t}\left(f(a+t.e_i) - f(a)\right)$$

Définition

Sous réserve d'existence, on appelle dérivée partielle de f selon sa i-ème variable l'application

$$\mathrm{D}_i f : \begin{cases} U \to \mathbb{R} \\ a \mapsto \mathrm{D}_i f(a) \end{cases}$$

Attention : On ne parle pas de la dérivée d'une fonction de deux variables réelles mais de ses deux dérivées partielles.

Théorème

Les dérivées partielles de f sont les dérivées de ses applications partielles.

dém. :
Considérons les applications partielles de f au point (x_1, x_2)

$$f_1 = f(., x_2) : t \mapsto f(t, x_2) \text{ et } f_2 = f(x_1, .) : t \mapsto f(t, x_2)$$

Pour étudier $\mathrm{D}_1 f(x_1, x_2)$, on considère le taux de variation

$$\frac{1}{t}(f(x_1 + t, x_2) - f(x_1)) = \frac{1}{t}(f_1(x_1 + t) - f_1(x_1))$$

Par suite $\mathrm{D}_1 f(x_1, x_2)$ existe si, et seulement si, $f_1'(x_1)$ existe et si tel le cas

$$\mathrm{D}_1 f(x_1, x_2) = f_1'(x_1)$$

On obtient un résultat similaire pour l'étude de $\mathrm{D}_2 f(x_1, x_2)$.
□

Corollaire

On exprime que les dérivées partielles sont les dérivées des applications partielles en écrivant

$$\mathrm{D}_1 f(x_1, x_2) = \frac{\mathrm{d}}{\mathrm{d}x_1}(f(x_1, x_2)) \text{ et } \mathrm{D}_2 f(x_1, x_2) = \frac{\mathrm{d}}{\mathrm{d}x_2}(f(x_1, x_2))$$

Exemple Considérons la fonction définie sur $\mathbb{R}^2$ donnée par la relation $f(x, y) = xy^2$.
On a

$$\mathrm{D}_1 f(x, y) = (x \mapsto xy^2)' = \frac{\mathrm{d}}{\mathrm{d}x}(xy^2) = y^2$$

et

$$\mathrm{D}_2 f(x, y) = (y \mapsto xy^2)' = \frac{\mathrm{d}}{\mathrm{d}y}(xy^2) = 2xy$$

Remarque Si on convient de noter (x, y) le couple de variables réelles de f, il est alors fréquent de noter

$$\frac{\partial f}{\partial x} \text{ et } \frac{\partial f}{\partial y}$$

au lieu de $\mathrm{D}_1 f$ et $\mathrm{D}_2 f$ les deux dérivées partielles de f.
Ainsi

$$\frac{\partial f}{\partial x}(x, y) = \frac{\mathrm{d}}{\mathrm{d}x}(f(x, y)) = \lim_{t \to 0} \frac{1}{t}(f(x + t, y) - f(x, y))$$

et

$$\frac{\partial f}{\partial y}(x, y) = \frac{\mathrm{d}}{\mathrm{d}y}(f(x, y)) = \lim_{t \to 0} \frac{1}{t}(f(x, y + t) - f(x, y))$$

Remarque Si on convient plutôt de noter (u, v) le couple de variables réelles de f, les dérivées partielles de f sont notées

$$\frac{\partial f}{\partial u} \text{ et } \frac{\partial f}{\partial v}$$

Exemple Considérons la fonction définie sur $\mathbb{R}^2$ donnée par la relation $f(x, y) = x^2y + x$.
On a

$$\frac{\partial f}{\partial x}(x, y) = \frac{\mathrm{d}}{\mathrm{d}x}(x^2y + x) = 2xy + 1 \text{ et } \frac{\partial f}{\partial y}(x, y) = x^2$$

Exemple Considérons la fonction définie sur $\mathbb{R}^{+\star} \times \mathbb{R}\mathbb{R}^2$ donnée par la relation $f(x, t) = x^t$.
On a

$$\frac{\partial f}{\partial x}(x, t) = tx^{t-1} \text{ et } \frac{\partial f}{\partial t}(x, t) = \ln x.x^t$$

Exemple Considérons la fonction définie sur $\mathbb{R}^2$ donnée par la relation

$$f(x, y) = \begin{cases} \dfrac{xy}{x^2 + y^2} \text{ si } (x, y) \neq (0, 0) \\ 0 \text{ sinon} \end{cases}$$

Calculons les dérivées partielles de f en $(0, 0)$.
On a

$$\frac{1}{t}\left(f(t, 0) - f(0, 0)\right) = 0 \xrightarrow[t \to 0]{} 0$$

On en déduit l'existence de la première dérivée partielle de f en $(0, 0)$ et

$$\frac{\partial f}{\partial x}(0, 0) = 0$$

De même on obtient

$$\frac{\partial f}{\partial y}(0, 0) = 0$$

Remarque Dans l'exemple précédent, la fonction n'est pas continue en $(0, 0)$ car

$$f(1/n, 1/n) \xrightarrow[n \to \infty]{} 1/2 \neq f(0, 0)$$

et cependant la fonction f admet deux dérivées partielles en $(0, 0)$!
Contrairement aux fonctions d'une variable réelle, l'existence des dérivées partielles n'assure pas la continuité au point.

23.2.4 Opérations sur les dérivées partielles

Soit $i = 1$ ou 2.

Proposition

Soient $f, g : U \to \mathbb{R}$ et $\lambda \in \mathbb{R}$.
Si les dérivées partielles $\mathrm{D}_i f$ et $\mathrm{D}_i g$ existent alors les dérives partielles $\mathrm{D}_i(\lambda f), \mathrm{D}_i(f + g), \mathrm{D}_i(fg)$ existent et on a les identités

$$\mathrm{D}_i(\lambda f) = \lambda \mathrm{D}_i f,\ \mathrm{D}_i(f+g) = \mathrm{D}_i f + \mathrm{D}_i g \text{ et } \mathrm{D}_i(fg) = \mathrm{D}_i f \times g + f \times \mathrm{D}_i g$$

dém. :
C'est immédiat puisque les dérivées partielles sont les dérivées des applications partielles

□

Proposition

Soient

$$f : \begin{cases} U \subset \mathbb{R}^2 \to \mathbb{R} \\ (x_1, x_2) \mapsto f(x_1, x_2) \end{cases} \text{ et } \varphi : \begin{cases} I \subset \mathbb{R} \to \mathbb{R} \\ t \to \varphi(t) \end{cases}$$

telles que $f(U) \subset I$
Si $\mathrm{D}_i f$ et φ' existent alors $\mathrm{D}_i(\varphi \circ f)$ existe et

$$\mathrm{D}_i(\varphi \circ f) = \mathrm{D}_i f \times \varphi' \circ f$$

dém. :
C'est immédiat puisque les dérivées partielles sont les dérivées des applications partielles

□

Exemple Soient $g : \mathbb{R} \to \mathbb{R}$ dérivable et $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x, y) = g(xy)$.
Par composition, les dérivées partielles de f existent et

$$\frac{\partial f}{\partial x}(x, y) = yg'(xy) \text{ et } \frac{\partial f}{\partial y}(x, y) = xg'(xy)$$

23.2.5 Fonction à valeurs dans $\mathbb{R}^2$

Comme pour les fonctions à valeurs réelles, on définit la notion de dérivée selon un vecteur et de dérivées partielles pour les fonctions à valeurs dans $\mathbb{R}^2$. On peut étudier une fonction à valeurs dans $\mathbb{R}^2$ à partir de ses fonctions coordonnées en vertu du résultat suivant

Proposition

Soit

$$f : \begin{cases} U \to \mathbb{R}^2 \\ (x_1, x_2) \mapsto (f_1(x_1, x_2), f_2(x_1, x_2)) \end{cases}$$

La fonction f est dérivable en $a \in U$ selon $h \in \mathbb{R}^2$ si, et seulement si, ses fonctions coordonnées f_1 et f_2 le sont et alors

$$\mathrm{D}_h f(a) = (\mathrm{D}_h f_1(a), \mathrm{D}_h f(a))$$

dém. :
Pour étudier la dérivabilité de f au point a selon le vecteur h on étudie la convergence quand $h \to 0$ de

$$\frac{1}{h}\left[f(a+h) - f(a)\right] = \left(\frac{1}{h}\left[f_1(a+h) - f_1(a)\right], \frac{1}{h}\left[f_2(a+h) - f_2(a)\right]\right)$$

Cette convergence est liée aux convergences des fonctions coordonnées associées c'est-à-dire à la dérivabilité des fonctions coordonnées f_1 et f_2 au point a selon le vecteur h.
□

Remarque Sous réserve d'existence, on a pour $i = 1$ ou 2

$$\mathrm{D}_i f = (\mathrm{D}_i f_1, \mathrm{D}_i f_2)$$

puisque les dérivées partielles sont les dérivées selon les vecteurs de la base canonique.

23.3 Fonctions de classe $\mathcal{C}^1$

23.3.1 Définition

Définition

On dit qu'une fonction $f : U \subset \mathbb{R}^2 \to \mathbb{R}$ est de classe $\mathcal{C}^1$ si ses dérivées partielles existent et sont continues sur U. On note $\mathcal{C}^1(U, \mathbb{R})$ l'ensemble de ces fonctions.

Exemple Les fonctions polynomiales et les fonctions rationnelles sont de classe $\mathcal{C}^1$.
En particulier les projections canoniques p_1 et p_2 sont de classe $\mathcal{C}^1$.

Théorème

Soient $f : U \to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$ et $a \in U$.
Pour tout $h = (h_1, h_2)$ tel que $a + h \in U$, on peut écrire :

$$f(a+h) = f(a) + \mathrm{D}_1 f(a).h_1 + \mathrm{D}_2 f(a).h_2 + \|h\|\,\varepsilon(h)$$

avec ε une fonction réelle telle que $\varepsilon(h) \xrightarrow[h \to 0]{} 0$.
Cette relation est appelée développement limité à l'ordre 1 de la fonction f au point a.

dém. :
Posons

$$\Delta(h) = f(a+h) - (f(a) + \mathrm{D}_1 f(a) h_1 + \mathrm{D}_2 f(a) h_2)$$

Notre objectif est d'établir

$$\varepsilon(h) = \frac{\Delta(h)}{\|h\|} \xrightarrow[h \to (0,0)]{} 0$$

Puisque $a \in U$ et que U est ouvert, il existe $\alpha > 0$ tel que

$$[a_1 - \alpha, a_1 + \alpha] \times [a_2 - \alpha, a_2 + \alpha] \subset U$$

Considérons $h = (h_1, h_2) \in \mathbb{R}^2$ tel que $\|h\| \leqslant \alpha$ de sorte que $h_1 \in [a_1 - \alpha, a_1 + \alpha]$ et $h_2 \in [a_2 - \alpha, a_2 + \alpha]$.
On peut écrire

$$\Delta(h) = f(a_1 + h_1, a_2 + h_2) - f(a_1, a_2) - h_1 \mathrm{D}_1 f(a_1, a_2) - h_2 \mathrm{D}_2 f(a_1, a_2) = \Delta_1(h) + \Delta_2(h)$$

avec
□

$$\Delta_1(h) = f(a_1 + h_1, a_2 + h_2) - f(a_1, a_2 + h_2) - h_1 \mathrm{D}_1 f(a_1, a_2)$$

et

$$\Delta_2(h) = f(a_1, a_2 + h_2) - f(a_1, a_2) - h_2 \mathrm{D}_2 f(a_1, a_2)$$

Considérons l'application partielle $t \mapsto f(t, a_2 + h_2)$, celle-ci est définie au moins sur l'intervalle $[a_1 - \alpha, a_1 + \alpha]$ et a pour fonction dérivée $t \mapsto \mathrm{D}_1 f(t, a_2)$. En lui appliquant le théorème des accroissements finis entre a_1 et $a_1 + h_1$ on obtient l'existence d'un élément $c_1(h)$ compris entre a_1 et $a_1 + h_1$ vérifiant :

$$f(a_1 + h_1, a_2 + h_2) - f(a_1, a_2 + h_2) = h_1 \mathrm{D}_1 f(c_1(h), a_2 + h_2)$$

Cela permet d'écrire

$$\Delta_1(h) = h_1 \left(\mathrm{D}_1 f(c_1(h), a_2 + h_2) - \mathrm{D}_1 f(a_1, a_2) \right)$$

De même, en considérant l'application partielle $t \mapsto f(a_1, t)$, on obtient l'existence d'un élément $c_2(h)$ compris entre a_2 et $a_2 + h_2$ vérifiant

$$f(a_1, a_2 + h_2) - f(a_1, a_2) = h_2 D_2 f(a_1, c_2(h))$$

Cela permet d'écrire

$$\Delta_2(h) = h_2 \left(\mathrm{D}_2 f(a_1, c_2(h)) - \mathrm{D}_2 f(a_1, a_2) \right)$$

Sachant $\|h\| = \sqrt{h_1^2 + h_2^2}$, on a $|h_1|, |h_2| \leqslant \|h\|$ et donc

$$\frac{\Delta(h)}{\|h\|} \leqslant |\mathrm{D}_1 f(c_1(h), a_2 + h_2) - \mathrm{D}_1 f(a_1, a_2)| + |\mathrm{D}_2 f(a_1, c_2(h)) - \mathrm{D}_2 f(a_1, a_2)|$$

Quand $h \to (0,0)$, $h_1 \to 0$ et $h_2 \to 0$ donc, par encadrement

$$c_1(h) \to a_1 \text{ et } c_2(h) \to a_2$$

Par composition de limite, sachant que les fonctions $\mathrm{D}_1 f$ et $\mathrm{D}_2 f$ sont continues en a, on obtient :

$$\mathrm{D}_1 f(c_1(h), a_2 + h_2) \to \mathrm{D}_1 f(a_1, a_2) \text{ et } \mathrm{D}_2 f(a_1, c_2(h)) \to \mathrm{D}_2 f(a_1, a_2)$$

et donc on peut conclure que

$$\frac{\Delta(h)}{\|h\|} \to 0$$

Corollaire

> Si f est une fonction de classe $\mathcal{C}^1$ alors f est continue.

dém. :
Quand $h \to 0$,

$$f(a+h) = f(a) + \mathrm{D}_1 f(a).h_1 + \mathrm{D}_2 f(a).h_2 + \|h\| \, \varepsilon(h) \to f(a).$$

□

23.3.2 Gradient

Soient $f : U \to \mathbb{R}$ une fonction et un point $a \in U$.

Proposition

> Si f est de classe $\mathcal{C}^1$ alors f est dérivable en a selon tout vecteur $h = (h_1, h_2)$ de $\mathbb{R}^2$ et
>
> $$\mathrm{D}_h f(a) = \mathrm{D}_1 f(a).h_1 + \mathrm{D}_2 f(a).h_2$$

dém. :
Pour étudier $\mathrm{D}_h f(a)$, on étudie la convergence quand $t \to 0$ de la quantité

$$\frac{1}{t}\left(f(a+t.h) - f(a)\right)$$

Or, par développement limité à l'ordre 1, on a

$$\frac{1}{t}\left(f(a+t.h) - f(a)\right) = \mathrm{D}_1 f(a).h_1 + \mathrm{D}_2 f(a).h_2 + \frac{|t|}{t}\varepsilon(t.h)$$

avec une fonction ε qui converge vers 0 en $(0,0)$.
On en déduit

$$\frac{1}{t}\left(f(a+t.h) - f(a)\right) \xrightarrow[t \to 0]{} \mathrm{D}_1 f(a).h_1 + \mathrm{D}_2 f(a).h_2$$

□

Exemple Reprenons la fonction polynomiale définie sur $\mathbb{R}^2$ par la relation $f(x_1, x_2) = x_1 x_2^2$.
f est de classe $\mathcal{C}^1$ car polynomiale et par un calcul immédiat

$$\mathrm{D}_1 f(a) = a_2^2 \text{ et } \mathrm{D}_2 f(a) = 2a_1 a_2$$

On en déduit que pour tout $h = (h_1, h_2) \in \mathbb{R}^2$,

$$\mathrm{D}_h f(a) = a_2^2 h_1 + 2a_1 a_2 h_2$$

On retrouve ainsi, avec plus de simplicité, un résultat obtenu dans un exemple précédent.

Remarque Si f est de classe $\mathcal{C}^1$, on peut noter que l'application

$$\begin{cases} \mathbb{R}^2 \to \mathbb{R} \\ h \mapsto \mathrm{D}_h f(a) \end{cases}$$

est une forme linéaire sur $\mathbb{R}^2$.

Définition

Si f est de classe $\mathcal{C}^1$, on appelle gradient de f en a le vecteur

$$\mathrm{grad} f(a) = (\mathrm{D}_1 f(a), \mathrm{D}_2 f(a))$$

Celui-ci est défini de sorte que

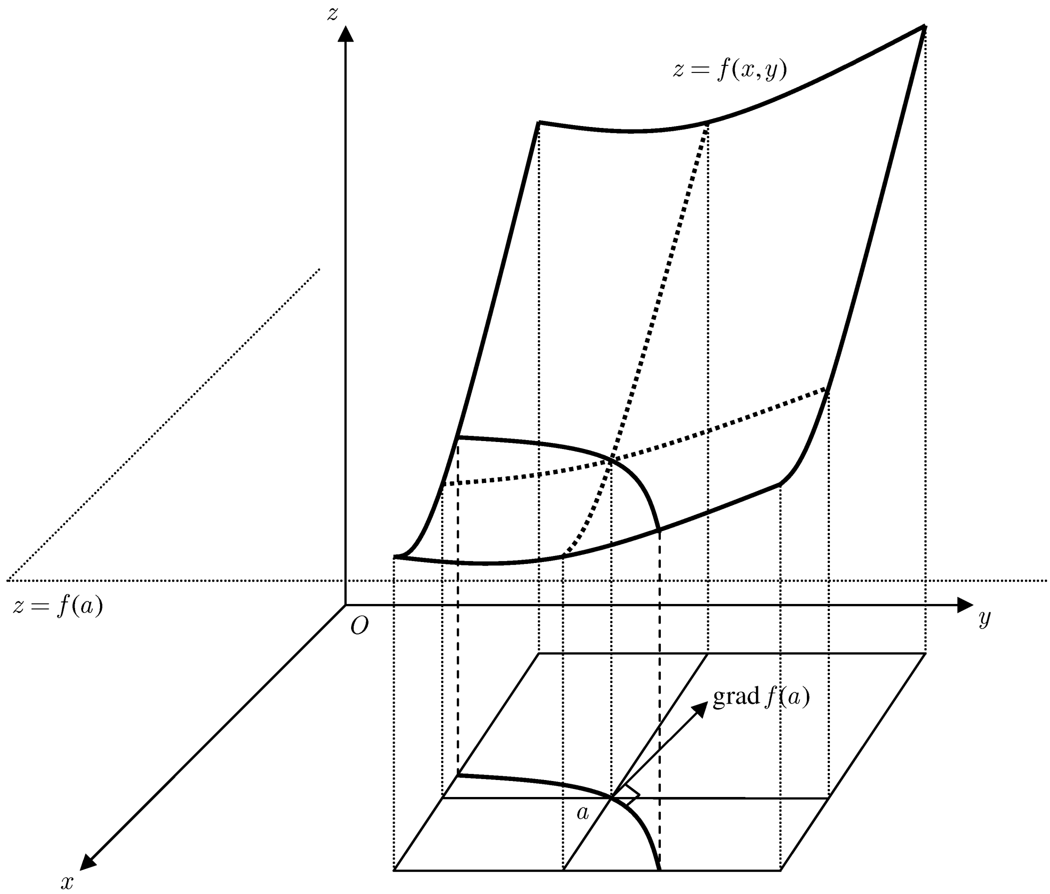
$$\forall h \in \mathbb{R}^2, \mathrm{D}_h f(a) = (\mathrm{grad}\, f(a) \mid h)$$

en notant $(. \mid .)$ le produit scalaire canonique sur $\mathbb{R}^2$.

Remarque Pour u un vecteur unitaire :

$$(\mathrm{grad} f(a) \mid u) = \mathrm{D}_u f(a) = \lim_{t \to 0} \frac{1}{t} \left(f(a + tu) - f(a) \right)$$

Cette quantité est maximale quand u a le sens et la direction du vecteur grad$f(a)$. Ainsi le vecteur gradient indique la direction de la plus grande pente, son sens donne le sens de progression croissante sur cette pente et sa norme donne la valeur de cette pente maximale.



23.3.3 Opérations sur les fonctions de classe $\mathcal{C}^1$

Proposition

Soient $f, g : U \to \mathbb{R}$ et $\lambda \in \mathbb{R}$.
Si f et g sont de classe $\mathcal{C}^1$ alors les fonctions $\lambda.f, f + g, fg$ le sont aussi.

dém. :
Les dérivées partielles à considérer existent et sont continues.

□

Remarque Par suite, l'ensemble $\mathcal{C}^1(U, \mathbb{R})$ est un sous-espace vectoriel et un sous anneau de $\mathcal{C}(U, \mathbb{R})$.

Proposition

Soient

$$f : \begin{cases} U \to \mathbb{R} \\ (x_1, x_2) \mapsto f(x_1, x_2) \end{cases} \text{ et } \varphi : \begin{cases} I \to \mathbb{R} \\ t \mapsto \varphi(t) \end{cases}$$

telles que $f(U) \subset I$

Si f et φ sont de classe $\mathcal{C}^1$ alors $\varphi \circ f : (x_1, x_2) \mapsto \varphi(f(x_1, x_2))$ l'est aussi.

dém. :

Les dérivées partielles à considérer existent et sont continues.

□

Exemple Si $f : U \subset \mathbb{R}^2 \to \mathbb{R}$ est de classe $\mathcal{C}^1$ et ne s'annule pas alors $1/f$ est $\mathcal{C}^1$.

Proposition

Soient

$$\varphi : \begin{cases} I \to \mathbb{R}^2 \\ t \mapsto (\varphi_1(t), \varphi_2(t)) \end{cases} \text{ et } f : \begin{cases} U \to \mathbb{R} \\ (x_1, x_2) \mapsto f(x_1, x_2) \end{cases}$$

telles que $\varphi(I) \subset U$

Si f et φ sont de classe $\mathcal{C}^1$ alors $f \circ \varphi : t \mapsto f(\varphi_1(t), \varphi_2(t))$ l'est aussi et

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(f(\varphi_1(t), \varphi_2(t)\right) = \varphi_1'(t)\mathrm{D}_1 f(\varphi(t)) + \varphi_2'(t)\mathrm{D}_2 f(\varphi(t))$$

dém. :

Pour étudier la dérivabilité de $f \circ \varphi$ en $t_0 \in I$, étudions la convergence quand $t \to 0$ du taux de variation

$$\frac{1}{t}\left((f \circ \varphi)(t_0 + t) - (f \circ \varphi)(t_0)\right) = \frac{1}{t}\left(f(a + h) - f(a)\right)$$

avec $a = \varphi(t_0)$ et $a + h = \varphi(t_0 + t)$ i.e. $h = \varphi(t_0 + t) - \varphi(t_0)$.

On a

$$a = (\varphi_1(t_0), \varphi_2(t_0))$$

et

$$h = (h_1, h_2) = (\varphi_1(t_0 + t) - \varphi_1(t_0), \varphi_2(t_0 + t) - \varphi_2(t_0))$$

Par développement limité à l'ordre 1 de la fonction f, on peut écrire

$$\frac{1}{t}(f(a + h) - f(a)) = \frac{1}{t}\left[\mathrm{D}_1 f(a) h_1 + \mathrm{D}_2 f(a).h_2 + \|h\| \varepsilon(h)\right]$$

ou encore

$$\frac{1}{t}\left(f(a + h) - f(a)\right) = \frac{\varphi_1(t_0 + t) - \varphi_1(t_0)}{t}\mathrm{D}_1 f(a) + \frac{\varphi_2(t_0 + t) - \varphi_2(t_0)}{t}\mathrm{D}_2 f(a) + \frac{\|h\|}{t}\varepsilon(h)$$

Quand $t \to 0$, on a $h \to (0, 0)$ par continuité de φ_1 et φ_2.

On a aussi

$$\frac{\varphi_i(t_0 + t) - \varphi_i(t_0)}{t} \to \varphi_i'(t_0)$$

et donc

$$\left|\frac{\|h\|}{t}\right| = \left\|\frac{h}{t}\right\| \to \|(\varphi_1'(t_0), \varphi_2'(t_0)\|$$

et on en déduit

$$\frac{\|h\|}{t}\varepsilon(h) \to 0$$

et finalement

$$\frac{1}{t}\left((f\circ\varphi)(t_0+t) - (f\circ\varphi)(t_0)\right) \to \varphi_1'(t_0)\mathrm{D}_1 f(\varphi(t_0)) + \varphi_2'(t_0)\mathrm{D}_2 f(\varphi(t_0))$$

Ainsi la fonction $f\circ\varphi$ est dérivable en t_0 et

$$(f\circ\varphi)'(t_0) = \mathrm{D}_1 f(\varphi(t_0))\varphi_1'(t_0) + \mathrm{D}_2 f(\varphi(t_0))\varphi_2'(t_0)$$

La fonction $f\circ\varphi$ est alors dérivable sur I et puisque sa dérivée est continue par opération sur des fonctions qui le sont, on peut affirmer que la fonction $f\circ\varphi$ est de classe $\mathcal{C}^1$.
□

Remarque En pratique, il convient de savoir mettre en place cette formule en utilisant les notations des dérivées partielles de f en fonction des notations convenues des variables réelles de f. Autrement dit, si f est vue en le couple de variable (x, y), on écrit $\dfrac{\partial f}{\partial x}$ et $\dfrac{\partial f}{\partial y}$ plutôt que $\mathrm{D}_1 f$ et $\mathrm{D}_2 f$.

Exemple Soient

$$f : \begin{cases} \mathbb{R}^2 \to \mathbb{R} \\ (x,y) \mapsto f(x,y) \end{cases}$$

une fonction de classe $\mathcal{C}^1$ et $g : \mathbb{R} \to \mathbb{R}$ la fonction définie par la relation $g(t) = f(\cos t, \sin t)$.
Par composition, la fonction g est de classe $\mathcal{C}^1$ et on a

$$g'(t) = \frac{\mathrm{d}}{\mathrm{d}t}\left(f(\cos t, \sin t)\right) = (-\sin t) \times \frac{\partial f}{\partial x}(\cos t, \sin t) + \cos t \times \frac{\partial f}{\partial y}(\cos t, \sin t)$$

Remarque Dans le contexte en cours, écrire $\dfrac{\partial f}{\partial t}$ n'a aucun sens car t ne désigne pas une variable de f et donc $\dfrac{\partial f}{\partial t}$ ne désigne par une dérivée partielle de f...

Exemple Soient $f : \begin{cases} \mathbb{R}^2 \to \mathbb{R} \\ (x,y) \mapsto f(x,y) \end{cases}$ une fonction de classe $\mathcal{C}^1$ et $g : \mathbb{R} \to \mathbb{R}$ la fonction définie par la relation $g(t) = f(2t, 1+t^2)$.
Par composition, la fonction g est de classe $\mathcal{C}^1$ et

$$g'(t) = 2\frac{\partial f}{\partial x}(2t, 1+t^2) + 2t\frac{\partial f}{\partial y}(2t, 1+t^2)$$

Remarque Si, pour changer, on avait convenait de noter (u, v) les variables de la fonction f, la relation précédente aurait été réécrite

$$g'(t) = 2\frac{\partial f}{\partial u}(2t, 1+t^2) + 2t\frac{\partial f}{\partial v}(2t, 1+t^2)$$

23.3.4 Fonctions à valeurs dans $\mathbb{R}^2$

Définition

On dit qu'une fonction $f : U \to \mathbb{R}^2$ est de classe $\mathcal{C}^1$ si ses dérivées partielles existent et sont continues.

Proposition

$f : U \to \mathbb{R}^2$ est $\mathcal{C}^1$ si, et seulement si, ses fonctions coordonnées le sont.

dém. :
Car les fonctions coordonnées des dérivées partielles de f, sont les dérivées partielles des fonctions coordonnées de f au contraire du frère de ma sœur qui n'est pas la sœur de mon frère. . .
□

Théorème

Soient

$$f : \begin{cases} U \to \mathbb{R}^2 \\ (x_1, x_2) \mapsto (f_1(x_1, x_2), f_2(x_1, x_2)) \end{cases} \text{ et } g : \begin{cases} V \to \mathbb{R} \\ (y_1, y_2) \mapsto g(y_1, y_2) \end{cases}$$

telles que $f(U) \subset V$.
Si les fonctions f et g sont de classe $\mathcal{C}^1$ alors la fonction composée $g \circ f$ l'est aussi et ses dérivées partielles sont

$$\mathrm{D}_1(g \circ f)(x_1, x_2) = \mathrm{D}_1 f_1(x_1, x_2)\mathrm{D}_1 g(f(x_1, x_2)) + \mathrm{D}_1 f_2(x_1, x_2)\mathrm{D}_2 g(f(x_1, x_2))$$

et

$$\mathrm{D}_2(g \circ f)(x_1, x_2) = \mathrm{D}_2 f_1(x_1, x_2)\mathrm{D}_1 g(f(x_1, x_2)) + \mathrm{D}_2 f_2(x_1, x_2)\mathrm{D}_2 g(f(x_1, x_2))$$

dém. :
Soit x_2 fixé. Considérons les applications partielles des fonctions f_1 et f_2 suivantes

$$\varphi_1 : x_1 \mapsto f_1(x_1, x_2) \text{ et } \varphi_2 : x_1 \mapsto f_2(x_1, x_2)$$

Sous réserve d'existence, la première dérivée partielle de $g \circ f$ en (x_1, x_2) est la dérivée de l'application partielle

$$x_1 \mapsto g(\varphi_1(x_1), \varphi_2(x_1)) = g(\varphi(x))$$

Or, par composition de fonction de classe $\mathcal{C}^1$, cette dernière fonction est dérivable et

$$\frac{\mathrm{d}}{\mathrm{d}x_1}\left(g(\varphi(x_1))\right) = \varphi_1'(x_1)\mathrm{D}_1 g(f(x)) + \varphi_2'(x_1)\mathrm{D}_2 g(f(x))$$

Sachant

$$\mathrm{D}_1(g \circ f)(x_1, x_2) = \frac{\mathrm{d}}{\mathrm{d}x_1}(g(\varphi_1(x_1), \varphi_2(x_1))) \text{ et } \mathrm{D}_i f(x_1, x_2) = \varphi_i'(x_1)$$

□

on obtient la relation proposée.
Rq Ce résultat se généralise aux fonctions g à valeurs dans $\mathbb{R}^2$.

Remarque En pratique, il convient de savoir adapter ce qui précède aux notations convenues des dérivées partielles des fonctions engagées…

Exemple Soient

$$f : \begin{cases} \mathbb{R}^2 \to \mathbb{R}^2 \\ (x, y) \mapsto (xy, x^2 + y^2) \end{cases} \text{ et } g : \begin{cases} \mathbb{R}^2 \to \mathbb{R} \\ (u, v) \mapsto g(u, v) \end{cases}$$

de classe $\mathcal{C}^1$.
L'application $g \circ f : \begin{cases} \mathbb{R}^2 \to \mathbb{R} \\ (x; y) \mapsto g(xy, x^2 + y^2) \end{cases}$ est de classe $\mathcal{C}^1$ par composition et on a :

$$\frac{\partial(g \circ f)}{\partial x}(x, y) = y\frac{\partial g}{\partial u}(x + y, x^2 + y^2) + 2x\frac{\partial g}{\partial v}(x + y, x^2 + y^2)$$

et

$$\frac{\partial(g \circ f)}{\partial y}(x, y) = x\frac{\partial g}{\partial u}(x + y, x^2 + y^2) + 2y\frac{\partial g}{\partial v}(x + y, x^2 + y^2)$$

Exemple Soit $f : \begin{cases} \mathbb{R}^2 \to \mathbb{R} \\ (x, y) \mapsto f(x, y) \end{cases}$ une fonction de classe $\mathcal{C}^1$ et $g : \mathbb{R}^2 \to \mathbb{R}$ définie par $g(u, v) = f(u + v, uv)$.
g est de classe $\mathcal{C}^1$ par la composition et de plus

$$\frac{\partial g}{\partial u}(u, v) = \frac{\mathrm{d}}{\mathrm{d}u}\left(f(u + v, uv)\right) = \frac{\partial f}{\partial x}(u + v, uv) + v\frac{\partial f}{\partial y}(u + v, uv)$$

et

$$\frac{\partial g}{\partial v}(u, v) = \frac{\mathrm{d}}{\mathrm{d}v}\left(f(u + v, uv)\right) = \frac{\partial f}{\partial x}(u + v, uv) + u\frac{\partial f}{\partial y}(u + v, uv)$$

23.4 Fonctions de classe $\mathcal{C}^2$

23.4.1 Dérivées partielles d'ordre 2

Soit une fonction

$$f : \begin{cases} U \to \mathbb{R} \\ (x_1, x_2) \mapsto f(x_1, x_2) \end{cases}$$

Définition

> Pour $i, j = 1$ ou 2, l'application $\mathrm{D}_j(\mathrm{D}_i f)$, si elle existe, est appelée dérivée partielle d'ordre 2 de f en sa i-ème puis j-ème variable. Celle-ci est notée $\mathrm{D}_{j,i} f$.

Remarque Lorsqu'on convient de noter x_1, x_2 les variables de f, on note

$$\frac{\partial^2 f}{\partial x_j \partial x_i} \text{ ou } \frac{\partial^2 f}{\partial x_i^2} \text{ quand } i = j$$

les dérivées partielles d'ordre 2 de f.

Exemple Calculons les dérivées partielles d'ordre 2 de $f(x, y) = x\mathrm{e}^{xy}$.
$\frac{\partial f}{\partial x}(x, y) = (1 + xy)\mathrm{e}^{xy}$ donc $\frac{\partial^2 f}{\partial x^2}(x, y) = (2y + xy^2)\mathrm{e}^{xy}$ et $\frac{\partial^2 f}{\partial y \partial x}(x, y) = (2x + x^2 y)\mathrm{e}^{xy}$.
$\frac{\partial f}{\partial y}(x, y) = x^2\mathrm{e}^{xy}$ donc $\frac{\partial^2 f}{\partial x \partial y}(x, y) = (2x + x^2 y)\mathrm{e}^{xy}$ et $\frac{\partial^2 f}{\partial y^2}(x, y) = x^3\mathrm{e}^{xy}$

23.4.2 Fonctions de classe $\mathcal{C}^2$

Soit une fonction

$$f : \begin{cases} U \to \mathbb{R} \\ (x_1, x_2) \mapsto f(x_1, x_2) \end{cases}$$

Définition

> On dit que f est de classe $\mathcal{C}^2$ sur U si toutes les dérivées partielles d'ordre 2 de f existent et sont continues sur U.
> On note $\mathcal{C}^2(U, \mathbb{R})$ l'ensemble des fonctions de U vers $\mathbb{R}$ de classe $\mathcal{C}^2$.

Exemple Les fonctions polynomiales et rationnelles sont de classe $\mathcal{C}^2$.
En particulier, les projections canoniques p_1, p_2 sont de classe $\mathcal{C}^2$.

Proposition

> f est de classe $\mathcal{C}^2$ si, et seulement si, ses dérivées partielles (d'ordre 1) existent et sont de classe $\mathcal{C}^1$.

dém. :
Les dérivées partielles d'ordre 2 de f sont les dérivées partielles des dérivées partielles de f.
Si f est de classe $\mathcal{C}^2$ alors les dérivées partielles (d'ordre 1) de f existent et admettent des dérivées partielles continues, elles sont donc de classe $\mathcal{C}^1$.
Inversement, si les dérivées partielles de f existent et sont de classe $\mathcal{C}^1$ alors f admet des dérivées partielles d'ordre 2 et celles-ci sont continues donc f est de classe $\mathcal{C}^2$.
□

Remarque On a les inclusions

$$\mathcal{C}^2(U, \mathbb{R}) \subset \mathcal{C}^1(U, \mathbb{R}) \subset \mathcal{C}(U, \mathbb{R})$$

Théorème

Soit une fonction
$$f : \begin{cases} U \to \mathbb{R} \\ (x_1, x_2) \mapsto f(x_1, x_2) \end{cases}$$
Si f est de classe $\mathcal{C}^2$ alors
$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial^2 f}{\partial x_2 \partial x_1}$$

dém. :
Soit $a = (a_1, a_2) \in U$.
Puisque $a \in U$ et que U est ouvert, il existe $\alpha > 0$ tel que $[a_1 - \alpha, a_1 + \alpha] \times [a_2 - \alpha, a_2 + \alpha] \subset U$.
Considérons $h = (h_1, h_2) \in \mathbb{R}^2$ tel que $\|h\| \leqslant \alpha$ de sorte que $h_1 \in [a_1 - \alpha, a_1 + \alpha]$ et $h_2 \in [a_2 - \alpha, a_2 + \alpha]$.
Posons
$$\Delta(h) = f(a_1 + h_1, a_2 + h_2) - f(a_1 + h_1, a_2) - f(a_1, a_2 + h_2) + f(a_1, a_2)$$
Nous allons établir l'égalité voulue en étudiant de deux manières la limite quand $h \to 0$ du rapport
$$\frac{\Delta(h)}{h_1 h_2}$$
Commençons en considérant l'application $t \mapsto f(t, a_2 + h_2) - f(t, a_2)$. Celle-ci est définie au moins sur l'intervalle $[a_1 - \alpha, a_1 + \alpha]$ et a pour fonction dérivée
$$t \mapsto \frac{\partial f}{\partial x_1}(t, a_2 + h_2) - \frac{\partial f}{\partial x_1}(t, a_2)$$
En lui appliquant le théorème des accroissements finis entre a_1 et a_1+h_1 qui sont éléments de $[a_1 - \alpha, a_1 + \alpha]$, on obtient l'existence d'un élément $c_1(h)$ compris entre a_1 et $a_1 + h_1$ tel que
$$\Delta(h) = h_1 \left(\frac{\partial f}{\partial x_1}(c_1(h), a_2 + h_2) - \frac{\partial f}{\partial x_1}(c_1(h), a_2) \right)$$
En appliquant à nouveau le théorème des accroissements finis, cette fois-ci à l'application $t \mapsto \dfrac{\partial f}{\partial x_1}(c_1(h), t)$ entre a_2 et $a_2 + h_2$, on obtient l'existence d'un élément $c_2(h)$ compris entre a_2 et $a_2 + h_2$ tel que
$$\Delta(h) = h_1 h_2 \frac{\partial^2 f}{\partial x_2 \partial x_1}(c_1(h), c_2(h))$$
Quand $h \to (0, 0)$, $h_1 \to 0$ et $h_2 \to 0$ donc par encadrement $c_1(h) \to a_1$ et $c_2(h) \to a_2$. Or $\dfrac{\partial^2 f}{\partial x_2 \partial x_1}$ est continue donc, par composition de limites
□
$$\frac{\Delta(h)}{h_1 h_2} \to \frac{\partial^2 f}{\partial x_2 \partial x_1}(a_1, a_2)$$
En procédant de même, mais en appliquant le théorème des accroissements finis à la fonction $t \mapsto f(a_1 + h_1, t) - f(a_1, t)$ entre a_2 et a_2+h_2, on obtient l'existence d'un élément $d_2(h)$ intermédiaire à a_2 et a_2+h_2 tel que
$$\Delta(h) = h_2 \left(\frac{\partial f}{\partial x_2}(a_1 + h_1, d_2(h)) - \frac{\partial f}{\partial x_2}(a_2 + d_2(h)) \right)$$

En appliquant une dernière fois le TAF, on obtient $d_1(h)$ intermédiaire à a_1 et $a_1 + h_1$ tel que

$$\Delta(h) = h_2 h_1 \frac{\partial^2 f}{\partial x_1 \partial x_2}(d_1(h), d_2(h))$$

Cette fois-ci, quand $h \to (0,0)$

$$\frac{\Delta(h)}{h_1 h_2} \to \frac{\partial^2 f}{\partial x_1 \partial x_2} f(a_1, a_2)$$

Par unicité de la limite, on conclut

$$\frac{\partial^2 f}{\partial x_1 \partial x_2}(a) = \frac{\partial^2 f}{\partial x_2 \partial x_1}(a)$$

Exemple Considérons la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ définie par

$$f(x,y) = \frac{xy^3}{x^2+y^2} \text{ si } (x,y) \neq (0,0) \text{ et } f(0,0) = 0$$

Quand $(x,y) \to (0,0)$ avec $(x,y) \neq (0,0)$, on peut écrire $x = r\cos\theta$ et $y = r\sin\theta$ avec $r = \sqrt{x^2+y^2} \to 0$ et alors

$$f(x,y) = r^2 \cos\theta \sin^3\theta \to 0 = f(0,0)$$

On en déduit que la fonction f est continue en $(0,0)$.
Par opérations, la fonction f est de classe $\mathcal{C}^1$ et même de classe $\mathcal{C}^2$ sur $\mathbb{R}^2 \setminus \{(0,0)\}$ avec

$$\frac{\partial f}{\partial x}(x,y) = \frac{(y^2-x^2)y^3}{(x^2+y^2)^2} \text{ et } \frac{\partial f}{\partial y}(x,y) = \frac{(3x^2+y^2)xy^2}{(x^2+y^2)^2}$$

De plus

$$\frac{\partial f}{\partial x}(0,0) = \lim_{t\to 0} \frac{1}{t}\left(f(t,0) - f(0,0)\right) = 0 \text{ et } \frac{\partial f}{\partial y}(0,0) = \lim_{t\to 0} \frac{1}{t}\left(f(0,t) - f(0,0)\right) = 0$$

et on vérifie comme ci-dessus que les dérivées partielles de f sont continues en $(0,0)$. Ainsi la fonction f est de classe $\mathcal{C}^1$.
Cependant

$$\frac{\partial^2 f}{\partial y \partial x}(0,0) = \lim_{t\to 0} \frac{1}{t}\left(\frac{\partial f}{\partial x}(0,t) - \frac{\partial f}{\partial x}(0,0)\right) = 1$$

et

$$\frac{\partial^2 f}{\partial x \partial y}(0,0) = \lim_{t\to 0} \frac{1}{t}\left(\frac{\partial f}{\partial y}(t,0) - \frac{\partial f}{\partial y}(0,0)\right) = 0$$

On en déduit que la fonction f n'est pas de classe $\mathcal{C}^2$.

23.4.3 Opérations sur les fonctions de classe $\mathcal{C}^2$

Proposition

Soient $f, g : U \to \mathbb{R}$ et $\lambda \in \mathbb{R}$.
Si f et g sont de classe $\mathcal{C}^2$ alors $\lambda.f, f+g, fg$ le sont aussi.

dém. :
Les dérivées partielles des fonctions $\lambda.f, f+g$ et fg existent et sont de classe $\mathcal{C}^1$ par opérations.
□

Remarque $C^2(U,\mathbb{R})$ est donc un sous-espace vectoriel et un sous-anneau de $\mathcal{C}^1(U,\mathbb{R})$.

Proposition

Soient

$$f : \begin{cases} U \to \mathbb{R} \\ (x_1,x_2) \mapsto f(x_1,x_2) \end{cases} \text{ et } \varphi : \begin{cases} I \to \mathbb{R} \\ t \mapsto \varphi(t) \end{cases}$$

telles que $f(U) \subset I$.
Si f et φ sont de classe $\mathcal{C}^2$ alors $\varphi \circ f : (x_1,x_2) \mapsto \varphi(f(x_1,x_2))$ l'est aussi.

dém. :
Les dérivées partielles de la fonctions $\varphi \circ f$ existent et sont de classe $\mathcal{C}^1$ par opérations.
□

Exemple Si $f : U \subset \mathbb{R}^2 \to \mathbb{R}$ est de classe $\mathcal{C}^2$ et ne s'annule pas alors la fonction $1/f$ est de classe $\mathcal{C}^2$.

Proposition

Soient

$$\gamma : \begin{cases} I \to \mathbb{R}^2 \\ t \mapsto (x_1(t),x_2(t)) \end{cases} \text{ et } f : \begin{cases} U \to \mathbb{R} \\ (x_1,x_2) \mapsto f(x_1,x_2) \end{cases}$$

telles que $\gamma(I) \subset U$.
Si f et γ sont de classe $\mathcal{C}^2$ alors $f \circ \varphi$ l'est aussi.

dém. :
La dérivée de $f \circ \gamma$ existe et est de classe $\mathcal{C}^1$ par opérations.
□

Exemple Soit $f : \mathbb{R}^2 \to \mathbb{R}$ de classe $\mathcal{C}^2$.
Exprimons la dérivée seconde de la fonction $g : t \mapsto f(2t, 1+t^2)$.
On a

$$g'(t) = 2\frac{\partial f}{\partial x}(2t,1+t^2) + 2t\frac{\partial f}{\partial y}(2t,1+t^2)$$

puis en dérivant et en simplifiant par le théorème de Schwarz

$$g''(t) = 4\frac{\partial^2 f}{\partial x^2}(2t,1+t^2) + 8t\frac{\partial^2 f}{\partial x \partial y}(2t,1+t^2) + \frac{\partial^2 f}{\partial y^2}(2t,1+t^2) + 2\frac{\partial f}{\partial y}(2t,1+t^2)$$

23.4.4 Fonctions à valeurs dans $\mathbb{R}^2$

Remarque Comme pour les fonctions à valeurs réelles, on définit les dérivées partielles d'ordre 2 et la notion de fonction de classe $\mathcal{C}^2$ dans le cadre des fonctions à valeurs dans $\mathbb{R}^2$. On obtient les résultats suivants :

Proposition

Une fonction $f : U \to \mathbb{R}^2$ est de classe $\mathcal{C}^2$ si, et seulement si, ses fonctions coordonnées le sont.

Proposition

Soient

$$f : \begin{cases} U \to \mathbb{R}^2 \\ (x_1, x_2) \mapsto (f_1(x_1, x_2), f_2(x_1, x_2)) \end{cases} \text{ et } g : \begin{cases} V \to \mathbb{R} \\ (y_1, y_2) \mapsto g(y_1, y_2) \end{cases}$$

telles que $f(U) \subset V$.
Si f et g sont $\mathcal{C}^2$ alors $g \circ f$ l'est aussi.

Exemple Soient une fonction $f : (x, y) \mapsto f(x, y)$ de classe $\mathcal{C}^2$ définie sur $\mathbb{R}^2$ et $g(u, v) = f(u + v, uv)$
La fonction g est de classe $\mathcal{C}^2$ par composition ; exprimons ses dérivées partielles d'ordre 2.
On a déjà

$$\frac{\partial g}{\partial u} = \frac{\partial f}{\partial x} + v\frac{\partial f}{\partial y} \text{ et } \frac{\partial g}{\partial v} = \frac{\partial f}{\partial x} + u\frac{\partial f}{\partial y}$$

On en déduit

$$\frac{\partial^2 g}{\partial u^2} = \frac{\partial^2 f}{\partial x^2} + 2v\frac{\partial^2 f}{\partial x \partial y} + v^2\frac{\partial^2 f}{\partial y^2}$$

$$\frac{\partial^2 g}{\partial u \partial v} = \frac{\partial^2 f}{\partial x^2} + (u + v)\frac{\partial^2 f}{\partial x \partial y} + uv\frac{\partial^2 f}{\partial y^2} + \frac{\partial f}{\partial y}$$

et

$$\frac{\partial^2 g}{\partial v^2} = \frac{\partial^2 f}{\partial x^2} + 2u\frac{\partial^2 f}{\partial x \partial y} + u^2\frac{\partial^2 f}{\partial y^2}$$

23.5 Manipulation de fonctions de deux variables

23.5.1 Recherche d'extremum

Définition

On dit que $f : A \subset \mathbb{R}^2 \to \mathbb{R}$ admet un minimum (resp. maximum) local en $a \in A$ si

$$\exists \alpha > 0, \forall x \in D(a, \alpha) \cap A, f(x) \geqslant f(a) \text{ (resp. } f(x) \leqslant f(a) \text{)}$$

Exemple Un extremum global est bien évidemment un extremum local.

Définition

Soit $f : U \to \mathbb{R}$ de classe $\mathcal{C}^1$. On dit que f admet un point critique en $a \in U$ si ses dérivées partielles s'y annulent.

Théorème

Soient U un ouvert non vide de $\mathbb{R}^2$ et $f : U \to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$.
Si f admet un extremum local en $a \in U$ alors c'est un point critique de f.

dém. :
Quitte à considérer $-f$, supposons que f admet un minimum local en $a = (a_1, a_2)$.
Il existe donc $\alpha > 0$ tel que pour tout $x \in D(a, \alpha)$, $f(x) \geqslant f(a)$.
Pour $i = 1$ ou 2, on a

$$\mathrm{D}_i f(a) = \lim_{t \to 0} \frac{1}{t}\left(f(a + t.e_i) - f(a)\right)$$

D'une part, pour $t > 0$,

$$\frac{1}{t}\left(f(a + t.e_i) - f(a)\right) \geqslant 0$$

donc à la limite quand $t \to 0^+$, $\mathrm{D}_i f(a) \geqslant 0$.
D'autre part, pour $t < 0$,

$$\frac{1}{t}\left(f(a + t.e_i) - f(a)\right) \leqslant 0$$

donc à la limite quand $t \to 0^-$, $\mathrm{D}_i f(a) \leqslant 0$.
On en déduit $\mathrm{D}_i f(a) = 0$.
□

Attention : La réciproque est fausse : un point critique peut ne pas être extremum local de la fonction.

Remarque Pour déterminer les extremum de $f : (x, y) \mapsto f(x, y)$:
- on recherche les points critiques de f en résolvant le système

$$\begin{cases} \dfrac{\partial f}{\partial x}(x, y) = 0 \\ \dfrac{\partial f}{\partial y}(x, y) = 0 \end{cases}$$

- on étudie chaque point critique (x_0, y_0) en se ramenant si besoin en $(0, 0)$ via translation des variables

$$\begin{cases} x = x_0 + u \\ y = y_0 + v \end{cases}$$

- enfin on passe éventuellement en coordonnées polaires pour mieux cerner les ordres de grandeur.

Exemple Etudions la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ définie par

$$f(x, y) = x^2 + y^2 + xy + 1$$

La fonction f est de classe $\mathcal{C}^1$ et ses points critiques s'obtiennent en résolvant le système

$$\begin{cases} 2x + y = 0 \\ x + 2y = 0 \end{cases}$$

On obtient $(0,0)$ seul point critique.
On a $f(0,0) = 1$. Etudions alors le signe de

$$g(x,y) = f(x,y) - f(0,0) = x^2 + y^2 + xy$$

(1) On écrit

$$\begin{cases} x = r\cos\theta \\ y = r\sin\theta \end{cases} \text{avec } r = \sqrt{x^2+y^2} \text{ et } \theta \in \mathbb{R}$$

On a alors

$$g(x,y) = r^2(1 + \cos\theta\sin\theta) \geqslant 0$$

Il y a donc un minimum en $(0,0)$.
(2) On a

$$g(x,y) = (x + \frac{1}{2}y)^2 + \frac{3}{4}y^2 \geqslant 0$$

(3) Sachant $|xy| \leqslant \dfrac{1}{2}(x^2+y^2)$ on a

$$g(x,y) \geqslant \frac{1}{2}(x^2+y^2) \geqslant 0$$

Exemple Etudions la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ définie par

$$f(x,y) = x^2 - y^2 + 1$$

La fonction f est de classe $\mathcal{C}^1$ et ses points critiques s'obtiennent en résolvant le système

$$\begin{cases} 2x = 0 \\ -2y = 0 \end{cases}$$

On obtient $(0,0)$ seul point critique.
On a $f(0,0) = 1$. Etudions alors le signe de

$$g(x,y) = f(x,y) - f(0,0) = x^2 - y^2$$

(1) On écrit

$$\begin{cases} x = r\cos\theta \\ y = r\sin\theta \end{cases} \text{avec } r = \sqrt{x^2+y^2} \text{ et } \theta \in \mathbb{R}$$

On a alors

$$g(x,y) = r^2(\cos^2\theta - \sin^2\theta)$$

Or cette expression n'est pas de signe constant au voisinage de $(0,0)$, on peut donc affirmer que $(0,0)$ n'est pas un extremum local de f.

(2) On a
$$g(1/n,0) = 1/n^2 > 0$$
donc au voisinage de $(0,0)$, il existe des valeurs prises par f strictement supérieure à celle en $(0,0)$.
On en déduit que $(0,0)$ n'est pas un maximum local.
On a aussi
$$g(0,1/n) = -1/n^2 < 0$$
et donc au voisinage de $(0,0)$, il existe des valeurs prises par f strictement inférieure à celle en $(0,0)$.
On en déduit que $(0,0)$ n'est pas un maximum local.

Exemple Etudions la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ définie par
$$f(x,y) = x^2 + y^2 + 4xy - 1$$
La fonction f est de classe $\mathcal{C}^1$ et ses points critiques s'obtiennent en résolvant le système
$$\begin{cases} 2x + 4y = 0 \\ 4x + 2y = 0 \end{cases}$$
On obtient $(0,0)$ seul point critique.
On a $f(0,0) = -1$. Etudions le signe de
$$g(x,y) = f(x,y) - f(0,0) = x^2 + y^2 + 4xy$$
(1) En passant en coordonnées polaires
$$g(x,y) = r^2(1 + 4\cos\theta\sin\theta) = r^2(1 + 2\sin 2\theta)$$
n'est pas de signe constant au voisinage de $(0,0)$ donc $(0,0)$ n'est pas extremum local.
(2) On a
$$g(x,y) = (x+2y)^2 - 3y^2$$
donc $g(1/n,0) = 1/n^2 > 0$ et $g(-2/n,1/n) = -3/n^2 < 0$.

Exemple Etudions la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ définie par
La fonction f est de classe $\mathcal{C}^1$ et ses points critiques s'obtiennent en résolvant le système
$$\begin{cases} 2x - 3x^2 = 0 \\ 2y = 0 \end{cases}$$
On obtient deux points critiques $(0,0)$ et $(2/3,0)$.
Etude en $(0,0)$:
$f(0,0) = 0$.
En passant en coordonnées polaires
$$f(x,y) = r^2(1 - r\cos^3\theta)$$
avec $1 - r\cos^3\theta \geqslant 0$ pour r suffisamment petit.

Par suite, au voisinage de $(0,0)$, $f(x,y) \geqslant 0$ et donc $(0,0)$ est un minimum local.
Etude en $(2/3,0)$:
$f(2/3,0) = 4/27$.
On se ramène en $(0,0)$ par translation

$$\begin{cases} x = 2/3 + u \\ y = v \end{cases}, \begin{cases} u = x - 2/3 \\ v = y \end{cases}$$

On a

$$f(x,y) - f(2/3,0) = v^2 - u^2 - u^3 =$$

En posant

$$\begin{cases} u = r\cos\theta \\ v = r\sin\theta \end{cases} \text{ avec } r = \sqrt{u^2+v^2} \text{ et } \theta \in \mathbb{R}$$

$$f(x,y) - f(2/3,0) = r^2(\sin^2\theta - \cos^2\theta - r\cos^3\theta)$$

Or $\sin^2\theta - \cos^2\theta - r\cos^3\theta$ prend des valeurs > 0 et < 0 même pour r petit donc f n'admet pas d'extremum local en $(0,0)$.

23.5.2 Equations aux dérivées partielles d'ordre 1

Définition

Résoudre sur U une équation aux dérivées partielles d'ordre 1 en l'inconnue $f : (x,y) \mapsto f(x,y)$, c'est déterminer toutes les fonctions réelles f de classe $\mathcal{C}^1$ sur U vérifiant une relation donnée entre

$$f, \frac{\partial f}{\partial x} \text{ et } \frac{\partial f}{\partial y}$$

Proposition

Soient I et J deux intervalles ouverts non vides de $\mathbb{R}$ et $U = I \times J$.
Les solutions sur U de l'équation

$$\frac{\partial f}{\partial x}(x,y) = 0$$

sont les fonctions $f : (x,y) \mapsto C(y)$ où $C \in \mathcal{C}^1(J,\mathbb{R})$.
Les solutions sur U de l'équation

$$\frac{\partial f}{\partial y}(x,y) = 0$$

sont les fonctions $f : (x,y) \mapsto C(x)$ où $C \in \mathcal{C}^1(I,\mathbb{R})$.

dém. :
Soit f une fonction de classe $\mathcal{C}^1$ solution sur U de l'équation

$$\frac{\partial f}{\partial x}(x,y) = 0$$

Soit $y \in J$ fixé.
L'application partielle $x \mapsto f(x,y)$ a pour dérivée $\dfrac{\partial f}{\partial x}(x,y) = 0$, c'est donc une fonction constante.

Ainsi il existe $C(y) \in \mathbb{R}$ tel que
$$\forall x \in I, f(x,y) = C(y)$$
On peut alors définir une application $C : J \to \mathbb{R}$ telle que
$$\forall (x,y) \in I \times J, f(x,y) = C(y)$$
Choisissons $a \in I$. On a
$$\forall y \in J, C(y) = f(a,y)$$
Puisque la fonction f est de classe $\mathcal{C}^1$, la fonction C et de classe $\mathcal{C}^1$ par composition de telles fonctions.
Inversement, une telle fonction f est évidemment solution de l'équation étudiée.
□

Exemple Résolvons sur $\mathbb{R}^2$ l'équation
$$\frac{\partial f}{\partial y}(x,y) = xy$$
Soit f une fonction solution.
Soit $x \in \mathbb{R}$ fixé.
L'application partielle $y \mapsto f(x,y)$ a pour dérivée
$$\frac{\partial f}{\partial y}(x,y) = xy$$
donc il existe $C(x) \in \mathbb{R}$ tel que
$$\forall (x,y) \in \mathbb{R}^2, f(x,y) = \frac{1}{2}xy^2 + C(x)$$
Ceci définit une fonction $C : \mathbb{R} \to \mathbb{R}$.
Soit $b \in \mathbb{R}$ fixé, $C : x \mapsto (x,b) \mapsto f(x,b) - \dfrac{1}{2}xb^2$ est de classe $\mathcal{C}^1$ par composition.
Inversement, une telle fonction f est solution.
On peut conclure
$$\mathcal{S} = \left\{(x,y) \mapsto \frac{1}{2}xy^2 + C(x) / C \in \mathcal{C}^1(\mathbb{R},\mathbb{R})\right\}$$

Remarque En pratique, on résout plus directement l'équation précédente en « intégrant en la variable y »ce qui « introduit une constante additive dépendant de x ».

Exemple Résolvons sur $\mathbb{R}^2$ l'équation
$$\frac{\partial f}{\partial x}(x,y) = yf(x,y)$$
Soit f une fonction solution.
Soit $y \in \mathbb{R}$ fixé.
L'application partielle $x \mapsto f(x,y)$ est solution de l'équation différentielle
$$z' = y.z$$
Les solution de cette équation différentielle sont les fonctions $z(x) = C\mathrm{e}^{xy}$.

Donc il existe $C(y) \in \mathbb{R}$ telle que

$$\forall (x,y) \in \mathbb{R}^2, f(x,y) = C(y)\mathrm{e}^{xy}$$

Soit $a \in \mathbb{R}$. On a alors $C(y) = f(a,y)\mathrm{e}^{-ay}$ donc C est une fonction de classe $\mathcal{C}^1$ par composition.
Inversement de telles fonctions f sont solutions.

$$\mathcal{S} = \left\{(x,y) \mapsto C(y)\mathrm{e}^{-xy} / C \in \mathcal{C}^1(\mathbb{R},\mathbb{R})\right\}$$

Exemple Résolvons sur $\mathbb{R}^2$ l'équation

$$2\frac{\partial f}{\partial x}(x,y) - \frac{\partial f}{\partial y}(x,y) = 0$$

en procédant au changement de variable

$$\begin{cases} u = x + y \\ v = x + 2y \end{cases}$$

Etudions ce changement de variables :

$$\begin{cases} u = x + y \\ v = x + 2y \end{cases}, \begin{cases} x = 2u - v \\ y = v - u \end{cases}, (x,y) \in \mathbb{R}^2 \text{ et } (u,v) \in \mathbb{R}^2$$

Soit $f : \mathbb{R}^2 \to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$.
Considérons $g : \mathbb{R}^2 \to \mathbb{R}$ définie par

$$g(u,v) =' f(x,y)' = f(2u - v, v - u)$$

g est une fonction de classe $\mathcal{C}^1$ définie sur $\mathbb{R}^2$ et

$$\frac{\partial g}{\partial u}(u,v) = 2\frac{\partial f}{\partial x}(2u - v, v - u) - \frac{\partial f}{\partial y}(2u - v, v - u)$$

Par suite la fonction f est solution de l'équation étudiée si, et seulement si, g est solution de l'équation

$$\frac{\partial g}{\partial u} = 0$$

i.e. g de la forme $g : (u,v) \mapsto C(v)$ avec $C \in \mathcal{C}^1(\mathbb{R},\mathbb{R})$ ce qui donne $f(x,y) = C(x + 2y)$.
Finalement

$$\mathcal{S} = \left\{(x,y) \mapsto C(x + 2y) / C \in \mathcal{C}^1(\mathbb{R},\mathbb{R})\right\}$$

23.5.3 Equations aux dérivées partielles d'ordre 2

Définition

Résoudre sur U une équation aux dérivées partielles d'ordre 2, c'est déterminer toutes les fonctions réelles $f : (x,y) \mapsto f(x,y)$ de classe $\mathcal{C}^2$ sur U vérifiant une relation donnée entre f,

$$\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial^2 f}{\partial x^2}, \frac{\partial^2 f}{\partial x \partial y} \text{ et } \frac{\partial^2 f}{\partial y^2}$$

Proposition

Soient I et J deux intervalles ouverts non vides de $\mathbb{R}$ et $U = I \times J$.
Les solutions sur U des équations

$$\frac{\partial^2 f}{\partial x^2}(x,y) = 0,\ \frac{\partial^2 f}{\partial y^2}(x,y) = 0 \text{ et } \frac{\partial^2 f}{\partial x \partial y}(x,y) = 0$$

sont respectivement les fonctions

$$f : (x,y) \mapsto xC(y) + D(y),\ f : (x,y) \mapsto yC(x) + D(x)$$

et

$$f : (x,y) \mapsto C(x) + D(y)$$

avec C et D fonctions réelles de classe $\mathcal{C}^2$ définies sur I ou J selon les cas
dém. :
Soit f une fonction de classe $\mathcal{C}^2$ solution sur U de l'équation

$$\frac{\partial^2 f}{\partial x^2}(x,y) = 0$$

Soit $y \in J$ fixé.
L'application partielle $x \mapsto f(x,y)$ a pour dérivée $\dfrac{\partial^2 f}{\partial x^2}(x,y) = 0$, c'est donc une fonction affine.
Ainsi il existe $C(y), D(y) \in \mathbb{R}$ tels que

$$\forall x \in I, f(x,y) = xC(y) + D(y)$$

Ceci définit des fonctions $C, D : J \to \mathbb{R}$ telles que

$$\forall (x,y) \in I \times J, f(x,y) = xC(y) + D(y)$$

Choisissons $a < b \in I$. Puisque

$$\begin{cases} f(a,y) = aC(y) + D(y) \\ f(b,y) = bC(y) + D(y) \end{cases}$$

on a

$$C(y) = \frac{f(b,y) - f(a,y)}{b-a} \text{ et } D(y) = \frac{bf(a,y) - af(b,y)}{b-a}$$

ce qui permet d'affirmer que les fonctions C et D sont de classe $\mathcal{C}^2$ par opérations sur les fonctions de classe $\mathcal{C}^2$.
Inversement, une telle fonction f est évidemment solution de l'équation étudiée.
Soit f une fonction de classe $\mathcal{C}^2$ solution sur U de l'équation

$$\frac{\partial^2 f}{\partial x \partial y}(x,y) = 0$$

La fonction $\dfrac{\partial f}{\partial y}$ est alors solution sur U de l'équation aux dérivées partielles

$$\frac{\partial g}{\partial x} = 0$$

et donc il existe une fonction $c : I \to \mathbb{R}$ de classe $\mathcal{C}^1$ telle que

$$\frac{\partial f}{\partial y}(x, y) = c(x)$$

En intégrant en la variable y on obtient

$$f(x, y) = C(x) + D(y)$$

avec C fonction de classe $\mathcal{C}^2$ car primitive de c et D une fonction de classe $\mathcal{C}^1$.
Or, en choisissant $a \in I$, on a

$$D(y) = f(a, y) - C(a)$$

et donc D est une fonction de classe $\mathcal{C}^2$ par opération sur les fonctions de classe $\mathcal{C}^2$.
Inversement, une telle fonction f est évidemment solution de l'équation étudiée.

Exemple Résolvons sur $\mathbb{R}^{+\star} \times \mathbb{R}$ l'équation

$$x^2\frac{\partial^2 f}{\partial x^2} + 2xy\frac{\partial^2 f}{\partial x \partial y} + y^2\frac{\partial^2 f}{\partial y^2} = xy$$

en passant aux coordonnées polaires.
Passer aux coordonnées polaires consiste à écrire

$$\begin{cases} x = r\cos\theta \\ y = r\sin\theta \end{cases}$$

Sachant $x > 0$ et $y \in \mathbb{R}$, on peut « inverser »ce système en proposant

$$\begin{cases} r = \sqrt{x^2 + y^2} \\ \theta = \arctan y/x \end{cases} \text{ avec } r \in \mathbb{R}^{+\star}, \theta \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$$

Soient $f : \mathbb{R}^{+\star} \times \mathbb{R} \to \mathbb{R}$ une fonction de classe $\mathcal{C}^2$ et g la fonction définie par

$$g(r, \theta) =' f(x, y)' = f(r\cos\theta, r\sin\theta)$$

La fonction g est de classe $\mathcal{C}^2$ et par dérivation de fonction composée

$$r^2\frac{\partial^2 g}{\partial r^2}(r, \theta) = \left[x^2\frac{\partial^2 f}{\partial x^2}(x, y) + 2xy\frac{\partial^2 f}{\partial x \partial y}(x, y) + y^2\frac{\partial^2 f}{\partial y^2}(x, y)\right]_{\substack{x = r\cos\theta \\ y = r\sin\theta}}$$

Par conséquent f est solution de l'équation étudiée si, et seulement si,

$$\forall (r, \theta) \in \mathbb{R}^{+\star} \times \left]-\pi/2, \pi/2\right[, r^2\frac{\partial g}{\partial r}(r, \theta) = r^2\cos\theta\sin\theta$$

soit encore

$$\forall (r, \theta) \in \mathbb{R}^{+\star} \times \left]-\pi/2, \pi/2\right[, \frac{\partial g}{\partial r}(r, \theta) = \cos\theta\sin\theta$$

Après résolution cela conduit à

$$g(r, \theta) = \frac{1}{2}r^2\cos\theta\sin\theta + rC(\theta) + D(\theta)$$

avec $C, D \in \mathcal{C}^2(]-\pi/2, \pi/2[\,, \mathbb{R})$
puis

$$f(x,y) = \frac{1}{2}xy + \sqrt{x^2+y^2}\tilde{C}\left(\frac{y}{x}\right) + \tilde{D}\left(\frac{y}{x}\right)$$

avec $\tilde{C}, \tilde{D} \in \mathcal{C}^2(\mathbb{R}, \mathbb{R})$ fonctions définie par composition à partir des fonctions C et D.

Exemple

$$\frac{\partial^2 y}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 y}{\partial t^2}$$

Modèle : on considère une corde tendue soumise à de petites oscillations, $y(x,t)$ est la hauteur de la corde au point d'abscisse x et à l'instant t, c est une vitesse de propagation, c'est une constante positive. Pour résoudre cette équation, on procède au changement de variable

$$\begin{cases} u = x + ct \\ v = x - ct \end{cases} \text{soit } \begin{cases} x = (u+v)/2 \\ t = (u-v)/2c \end{cases}$$

Soit $y : \mathbb{R}^2 \to \mathbb{R}$ une fonction de classe $\mathcal{C}^2$ et $z : \mathbb{R}^2 \to \mathbb{R}$ définie par

$$z(u,v) = y(x,t) = y\left(\frac{u+v}{2}, \frac{u-v}{2c}\right)$$

La fonction z est de classe $\mathcal{C}^2$ et on observe que y est solution de l'équation étudiée si, et seulement si,

$$\frac{\partial^2 z}{\partial u \partial v} = 0$$

ce qui donne

$$z(u,v) = C(u) + D(v)$$

avec C, D fonctions de classe $\mathcal{C}^2$ définie sur $\mathbb{R}$
puis

$$y(x,t) = C(x+ct) + D(x-ct)$$

Cette solution y se comprend comme la somme de deux ondes de propageant à la vitesse c et en sens inverses.

23.5.4 Problème de primitivation

Soient $P, Q : U \to \mathbb{R}$ deux fonctions continues.
On cherche s'il existe des fonctions $f : U \to \mathbb{R}$ de classe $\mathcal{C}^1$ vérifiant

$$\begin{cases} \dfrac{\partial f}{\partial x}(x,y) = P(x,y) \\ \dfrac{\partial f}{\partial y}(x,y) = Q(x,y) \end{cases}$$

Exemple Etudions le système
$$\begin{cases} \dfrac{\partial f}{\partial x}(x,y) = y^2 \ \ (1) \\ \dfrac{\partial f}{\partial y}(x,y) = 2xy + y \ \ (2) \end{cases}$$
Soit f une solution sur $\mathbb{R}^2$.
(1) donne $f(x,y) = xy^2 + C(y)$ avec C fonction de classe $\mathcal{C}^1$.
Injectée dans (2), on obtient $C'(y) = y$ d'où $C(y) = \dfrac{1}{2}y^2 + C$ avec $C \in \mathbb{R}$.
Finalement $f(x,y) = x^2y + \dfrac{1}{2}y^2 + C$ avec $C \in \mathbb{R}$.
Inversement, on peut affirmer qu'une telle fonction est solution du système étudié.

Exemple Etudions le système
$$\begin{cases} \dfrac{\partial f}{\partial x}(x,y) = 2xy \ \ (1) \\ \dfrac{\partial f}{\partial y}(x,y) = 1 + y \ \ (2) \end{cases}$$
Soit f une solution sur $\mathbb{R}^2$.
(1) donne $f(x,y) = x^2y + C(y)$ avec C fonction de classe $\mathcal{C}^1$.
Injectée dans (2), on obtient $C'(y) = 1 + y - x^2$. Ceci est impossible car $C'(y)$ est fonction de y et aucunement de x.
On en déduit que le système proposé n'a pas de solutions.

A quelle condition le système
$$\begin{cases} \dfrac{\partial f}{\partial x}(x,y) = P(x,y) \\ \dfrac{\partial f}{\partial y}(x,y) = Q(x,y) \end{cases}$$
possède-t-il une solution ?
Supposons que P et Q soient des fonctions classe $\mathcal{C}^1$ sur U. Si le système possède une solution f, celle-ci est de classe $\mathcal{C}^2$ et par le théorème de Schwarz
$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y}$$
ce qui entraîne
$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$$
Cette condition est donc une condition nécessaire pour que le système possède une solution. Voyons, que dans certaines situations, elle peut être suffisante.

Définition

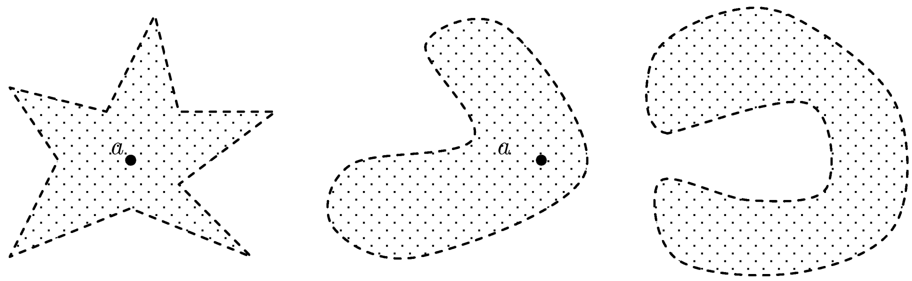
Soient $a, b \in \mathbb{R}^2$. On appelle segment d'extrémités a et b l'ensemble
$$[a,b] = \{a + \lambda(b-a)/\lambda \in [0,1]\}$$

Définition

On dit qu'un ouvert U de $\mathbb{R}^2$ est étoilé s'il existe $a \in U$ vérifiant

$$\forall x \in U, [a, x] \subset U$$

Exemple



étoilé étoilé non étoilé

Exemple $\mathbb{R}^2$ est étoilé.

Exemple $\mathbb{R}^2 \setminus \{(0,0)\}$ n'est pas étoilé.

Exemple Toute partie convexe est étoilée, en particulier le produit cartésien de deux intervalles réels est étoilé.

Théorème

Soient U un ouvert étoilé de $\mathbb{R}^2$ et $P, Q : U \to \mathbb{R}$ des fonctions de classe $\mathcal{C}^1$.
Le système

$$\begin{cases} \dfrac{\partial f}{\partial x}(x,y) = P(x,y) \\ \dfrac{\partial f}{\partial y}(x,y) = Q(x,y) \end{cases}$$

admet une solution sur U si, et seulement si,

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$$

De plus, si tel est le cas, les solutions se déduisent les une des autres par l'addition d'une constante.

Remarque Ce théorème sera démontré dans le cours de seconde année.

Remarque En sciences physiques, ce résultat permet de caractériser les forces du plan dérivant d'un potentiel.

23.6 Analyse vectorielle

$\mathcal{P}$ désigne le plan géométrique muni d'un repère orthonormé direct $\mathcal{R} = (O; \vec{i}, \vec{j})$.
$\mathcal{E}$ désigne l'espace géométrique muni d'un repère orthonormé direct $\mathcal{R} = (O; \vec{i}, \vec{j}, \vec{k})$.

23.6.1 Champ scalaire

23.6.1.1 Définition

Définition

On appelle champ scalaire du plan (resp. de l'espace) toute fonction $F : \mathcal{D} \to \mathbb{R}$ définie sur une partie $\mathcal{D}$ du plan (resp. de l'espace).

Exemple $F(M) = OM^2$ définit un champ scalaire.

23.6.1.2 Représentations dans le cadre du plan

Soit $F : \mathcal{D} \to \mathbb{R}$ un champ scalaire du plan.
Pour tout $(x, y) \in \mathbb{R}^2$ tel que le point M de coordonnées (x, y) est élément de $\mathcal{D}$, on pose

$$f(x, y) = F(M)$$

Définition

La fonction f est appelée représentation cartésienne du champ scalaire F (dans le repère $\mathcal{R}$).
On dit que champ scalaire F est continue (resp. de classe $\mathcal{C}^1$, $\mathcal{C}^2$) si la fonction f l'est.

Remarque Par les formules de changement de repère, on peut montrer que la notion ne dépend pas du choix du repère $\mathcal{R}$ du plan $\mathcal{P}$.

Exemple Pour $F(M) = OM^2$, $f(x, y) = x^2 + y^2$. F est un champ scalaire de classe $\mathcal{C}^2$.

Définition

Si l'on convient de noter (x, y) les coordonnées dans $\mathcal{R}$ d'un point générique M, on pose, sous réserve d'existence

$$\frac{\partial F}{\partial x}(M) = \frac{\partial f}{\partial x}(x, y) \text{ et } \frac{\partial F}{\partial y}(M) = \frac{\partial f}{\partial y}(x, y)$$

Exemple Pour le champ scalaire $F(M) = OM^2$ on obtient

$$\frac{\partial F}{\partial x}(M) = 2x \text{ et } \frac{\partial F}{\partial y}(M) = 2y$$

Pour tout couple $(r, \theta) \in \mathbb{R}^{+\star} \times \mathbb{R}$ tel que le point M de coordonnées polaires (r, θ) est élément de $\mathcal{D}$, on pose

$$\tilde{f}(r, \theta) = F(M)$$

Définition

La fonction $\tilde{f}$ est appelée représentation polaire du champ scalaire F (dans le repère $\mathcal{R}$).

Exemple Pour $F(M) = OM^2$, on obtient $\tilde{f}(r, \theta) = r^2$.

Définition

Si l'on convient de noter les coordonnées polaire dans $\mathcal{R}$ d'une point générique M, on pose, sous réserve d'existence

$$\frac{\partial F}{\partial r}(M) = \frac{\partial \tilde{f}}{\partial r}(r, \theta) \text{ et } \frac{\partial F}{\partial \theta}(M) = \frac{\partial \tilde{f}}{\partial \theta}(r, \theta)$$

Exemple Pour $F(M) = OM^2$ on obtient

$$\frac{\partial F}{\partial r}(M) = 2r \text{ et } \frac{\partial F}{\partial \theta}(M) = 0$$

Remarque Dans la pratique, il est très fréquent de ne pas distinguer les objets F, f et $\tilde{f}$.
Cela permet l'écriture abusive

$$F(M) = OM^2 = x^2 + y^2 = r^2$$

qui rend les manipulations qui précèdent immédiates.

23.6.1.3 Représentations dans le cadre de l'espace

Soit $F : \mathcal{D} \to \mathbb{R}$ un champ scalaire de l'espace.
Comme dans le plan on introduit sa représentation cartésienne dans un repère qui permet de définir, sous réserve d'existence,

$$\frac{\partial F}{\partial x}, \frac{\partial F}{\partial y} \text{ et } \frac{\partial F}{\partial z}$$

A partir des notions de coordonnées cylindriques (ρ, φ, z) et sphériques (r, θ, φ), on introduit aussi les notions de représentations cylindriques et sphériques de la fonction F. Ceci permet de donner un sens aux dérivées partielles :

$$\frac{\partial F}{\partial \rho}(M), \frac{\partial F}{\partial \varphi}(M), \frac{\partial F}{\partial z}(M)$$

et au dérivées partielles

$$\frac{\partial F}{\partial r}(M), \frac{\partial F}{\partial \theta}(M), \frac{\partial F}{\partial \varphi}(M)$$

Exemple Considérons le champ scalaire donné par

$$F(M) = OM$$

On a

$$F(M) = \sqrt{x^2 + y^2 + z^2}$$

donc

$$\frac{\partial F}{\partial x}(M) = \frac{x}{\sqrt{x^2+y^2+z^2}},\ \frac{\partial F}{\partial y}(M) = \frac{y}{\sqrt{x^2+y^2+z^2}} \text{ et } \frac{\partial F}{\partial z}(M) = \frac{z}{\sqrt{x^2+y^2+z^2}}$$

On a aussi

$$F(M) = \sqrt{\rho^2 + z^2}$$

donc

$$\frac{\partial F}{\partial \rho}(M) = \frac{\rho}{\sqrt{\rho^2+z^2}} \text{ et } \frac{\partial F}{\partial \varphi}(M) = 0$$

On a encore

$$F(M) = r$$

donc

$$\frac{\partial F}{\partial r}(M) = 1 \text{ et } \frac{\partial F}{\partial \theta}(M) = 0$$

23.6.2 Champ de vecteurs

Définition

On appelle champ de vecteurs du plan (resp. de l'espace) toute application $\vec{F} : \mathcal{D} \to P$ (resp. $\vec{F} : \mathcal{D} \to E$) définie sur une partie $\mathcal{D}$ du plan (resp. de l'espace) et à valeurs dans l'espace vectoriel P (resp. E) direction de $\mathcal{P}$ (resp. $\mathcal{E}$).

Exemple $\vec{F}(M) = \overrightarrow{OM}$ définit un champ de vecteurs du plan.

Soit $\vec{F}$ un champ de vecteurs du plan défini sur $\mathcal{D}$. Pour tout $M \in \mathcal{D}$, on peut écrire

$$\vec{F}(M) = P(M)\vec{i} + Q(M)\vec{j}$$

ce qui introduit des applications $P, Q : \mathcal{D} \to \mathbb{R}$.

Définition

Les champs scalaires P et Q sont appelés fonctions composantes du champ scalaire $\vec{F}$ (relatives au repère $\mathcal{R}$.

De même si $\vec{F}$ est un champ scalaire de l'espace on introduit ses fonctions composantes

$$P, Q, R : \mathcal{D} \to P$$

permettant d'écrire

$$\vec{F}(M) = P(M)\vec{i} + Q(M)\vec{j} + R(M)\vec{k}$$

Exemple Pour $\vec{F}(M) = \overrightarrow{OM}$ dans le plan, on obtient

$$P(M) = x \text{ et } Q(M) = y$$

Exemple Pour $\vec{F}(M) = \vec{\omega} \wedge \overrightarrow{OM}$ dans l'espace, on obtient

$$P(M) = \omega_y z - \omega_z y,\ Q(M) = \omega_z x - \omega_x z \text{ et } R(M) = \omega_x y - \omega_y x$$

Définition

Un champ de vecteurs est dit continue (resp. de classe $\mathcal{C}^1$, $\mathcal{C}^2$) si ses fonctions composantes le sont.

Remarque A partir de ses fonctions composantes, on peut définir les dérivées partielles d'un champ de vecteur $\vec{F}$.

23.6.3 Gradient d'un champ scalaire

Définition

On appelle gradient d'un champ scalaire F de classe $\mathcal{C}^1$ du plan le champ de vecteurs défini par

$$\overrightarrow{\text{grad}}\, F(M) = \frac{\partial F}{\partial x}(M)\vec{i} + \frac{\partial F}{\partial y}(M)\vec{j}$$

Remarque Respectivement, pour un champ scalaire de l'espace

$$\overrightarrow{\text{grad}}\, F(M) = \frac{\partial F}{\partial x}(M)\vec{i} + \frac{\partial F}{\partial y}(M)\vec{j} + \frac{\partial F}{\partial z}(M)\vec{k}$$

Remarque On peut montrer que levecteur gradient est indépendant du choix du repère orthonormé $\mathcal{R}$.

Remarque Ce vecteur indique la direction dans laquelle la progression du champ scalaire F est la plus importante.

Exemple Pour $F(M) = OM^2$, on a $\overrightarrow{\text{grad}}\, F = 2\overrightarrow{OM}$

Proposition

Soient F et G deux champs scalaires de classe $\mathcal{C}^1$ et $\lambda \in \mathbb{R}$. On a

$$\overrightarrow{\text{grad}}\, \lambda.F = \lambda.\overrightarrow{\text{grad}}\, F,\ \overrightarrow{\text{grad}}\,(F+G) = \overrightarrow{\text{grad}}\, F + \overrightarrow{\text{grad}}\, G$$

et

$$\overrightarrow{\text{grad}}\,(FG) = F.\overrightarrow{\text{grad}}\, G + G.\overrightarrow{\text{grad}}\, F$$

dém. :
Par opération sur les dérivées partielles...
□

Proposition

Soit F un champ scalaire de classe $\mathcal{C}^1$ du plan.
On a

$$\overrightarrow{\text{grad}}\, F(M) = \frac{\partial F}{\partial r}(M)\vec{u}_\theta + \frac{1}{r}\frac{\partial F}{\partial \theta}(M)\vec{v}_\theta$$

dém. :
Introduisons f et $\tilde{f}$ les représentations cartésienne et polaire de F. On a

$$f(M) = f(x, y) = \tilde{f}(r, \theta)$$

avec M de coordonnées cartésiennes (x, y) et de coordonnées polaires (r, θ) dans le repère $\mathcal{R}$.
Par définition

$$\overrightarrow{\text{grad}}\, F(M) = \frac{\partial F}{\partial x}(M)\vec{i} + \frac{\partial F}{\partial y}(M)\vec{j}$$

avec

$$\frac{\partial F}{\partial x}(M) = \frac{\partial f}{\partial x}(x, y), \frac{\partial F}{\partial y}(M) = \frac{\partial f}{\partial y}(x, y)$$

Puisque

$$\frac{\partial F}{\partial r}(M) = \frac{\partial \tilde{f}}{\partial r}(\rho, \theta), \frac{\partial F}{\partial \theta}(M) = \frac{\partial \tilde{f}}{\partial \theta}(\rho, \theta)$$

déterminons les relations qui existent entre les dérivées partielles de f et celle de $\tilde{f}$.
Puisque

$$\begin{cases} x = r\cos\theta \\ y = r\sin\theta \end{cases}$$

on a $\tilde{f}(r, \theta) = f(r\cos\theta, r\sin\theta)$ donc par dérivation de fonctions composées

$$\frac{\partial \tilde{f}}{\partial r}(r, \theta) = \cos\theta\frac{\partial f}{\partial x}(r\cos\theta, r\sin\theta) + \sin\theta\frac{\partial f}{\partial y}(r\cos\theta, r\sin\theta)$$

$$\frac{\partial \tilde{f}}{\partial \theta}(r, \theta) = r\sin\theta\frac{\partial f}{\partial x}(r\cos\theta, r\sin\theta) + -r\cos\theta\frac{\partial f}{\partial y}(r\cos\theta, r\sin\theta)$$

Ainsi

$$\frac{\partial F}{\partial r}(M) = \cos\theta\frac{\partial F}{\partial x}(M) + \sin\theta\frac{\partial F}{\partial y}(M) \text{ (1)}$$

et

$$\frac{\partial F}{\partial \theta}(M) = -\rho\sin\theta\frac{\partial F}{\partial x}(M) + r\cos\theta\frac{\partial F}{\partial y}(M) \text{ (2)}$$

Les combinaisons $\cos\theta \times (1) - \dfrac{1}{r}\sin\theta \times (2)$ et $\sin\theta \times (1) + \dfrac{1}{\rho}\cos\theta \times (2)$ donnent alors

$$\frac{\partial F}{\partial x}(M) = \cos\theta\frac{\partial F}{\partial r}(M) - \frac{1}{r}\sin\theta\frac{\partial F}{\partial \theta}(M)$$

et

$$\frac{\partial F}{\partial y}(M) = \sin\theta \frac{\partial F}{\partial r}(M) + \frac{1}{r}\cos\theta \frac{\partial F}{\partial \theta}(M)$$

et on en déduit

$$\overrightarrow{\text{grad}}\, F(M) = \frac{\partial F}{\partial r}(M)\vec{u}_\theta + \frac{1}{r}\frac{\partial F}{\partial \theta}(M)\vec{v}_\theta$$

□

23.6.4 Potentiel scalaire

23.6.4.1 Définition

Définition

On dit qu'un champ de vecteurs $\vec{F}$ dérive d'un potentiel scalaire s'il existe un champ scalaire V de classe $\mathcal{C}^1$ vérifian

$$\vec{F} = \overrightarrow{\text{grad}}\, V$$

Le champ scalaire V est alors appelé potentiel scalaire de $\vec{F}$.

Exemple Le champ de vecteur $\vec{F}(M) = \overrightarrow{OM}$ dérive du potentiel $V(M) = \frac{1}{2}OM^2$.

23.6.4.2 Dans le cadre du plan

Si $\vec{F}(M) = P(M)\vec{\imath} + Q(M)\vec{\jmath}$ alors $\vec{F} = \overrightarrow{\text{grad}}\ V$ si, et seulement si,

$$\begin{cases} \dfrac{\partial V}{\partial x}(M) = P(M) \\ \dfrac{\partial V}{\partial y}(M) = Q(M) \end{cases}$$

On peut alors transposer le théorème de Poincaré et ainsi

Théorème

Soit $\vec{F}$ un champ de vecteurs $\mathcal{C}^1$ du plan de fonctions composantes P et Q.
Pour que $\vec{F}$ dérive d'un champ scalaire il est nécessaire que

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$$

Inversement, si $\vec{F}$ est défini sur un ouvert étoilé, cette condition est suffisante.

23.6.4.3 Cadre de l'espace

Si $\vec{F}(M) = P(M)\vec{\imath} + Q(M)\vec{\jmath} + R(M)\vec{k}$ alors $\vec{F} = \overrightarrow{\text{grad}}\ V$ si, et seulement si,

$$\begin{cases} \dfrac{\partial V}{\partial x}(M) = P(M) \\ \dfrac{\partial V}{\partial y}(M) = Q(M) \\ \dfrac{\partial V}{\partial z}(M) = R(M) \end{cases}$$

On peut généraliser le théorème de Poincaré à la dimension supérieure

Théorème

Soit $\vec{F}$ un champ de vecteurs $\mathcal{C}^1$ de l'espace de fonctions composantes P, Q, R.
Pour que $\vec{F}$ dérive d'un champ scalaire il est nécessaire que :

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}, \frac{\partial Q}{\partial z} = \frac{\partial R}{\partial y} \text{ et } \frac{\partial R}{\partial x} = \frac{\partial P}{\partial z}$$

Inversement, si $\vec{F}$ est défini sur un ouvert étoilé, cette condition est suffisante.

23.6.5 Divergence d'un champ de vecteurs

Définition

Soit $\vec{F}$ un champ de vecteurs du plan de classe $\mathcal{C}^1$ et de fonctions composantes P et Q
On appelle divergence du champ de vecteurs $\vec{F}$ le champ scalaire défini par

$$\text{div}\vec{F}(M) = \frac{\partial P}{\partial x}(M) + \frac{\partial Q}{\partial y}(M)$$

Remarque Respectivement, pour champ de vecteurs $\vec{F}$ de l'espace, on pose

$$\text{div}\vec{F}(M) = \frac{\partial P}{\partial x}(M) + \frac{\partial Q}{\partial y}(M) + \frac{\partial R}{\partial z}(M)$$

Remarque On peut montrer que ce vecteur est indépendant du choix du repère orthonormé $\mathcal{R}$.

Remarque La divergence se comprend comme étant la « quantité de vecteurs »naissant au point considéré lorsqu'elle est positive ou y disparaissant lorsqu'elle est négative.

Exemple Dans le plan, pour $\vec{F}(M) = \overrightarrow{OM}$ on a div$\vec{F}(M) = 2$.

Exemple Dans le plan, pour $\vec{F}(M) = \dfrac{\overrightarrow{OM}}{OM^2}$ on a div$\vec{F}(M) = 0$.

Proposition

Soient $\vec{F}, \vec{G}$ champs de vecteurs de classe $\mathcal{C}^1$, F un champ scalaire de classe $\mathcal{C}^1$ et $\lambda \in \mathbb{R}$. On a

$$\text{div}(\lambda.\vec{F}) = \lambda.\text{div}\vec{F}, \text{div}(\vec{F} + \vec{G}) = \text{div}\vec{F} + \text{div}\vec{G}$$

et

$$\text{div}(F.\vec{G}) = F.\text{div}\vec{G} + \overrightarrow{\text{grad}}\, F.\vec{G}$$

dém. :
Par opérations sur les dérivées partielles. . .
□

23.6.6 Rotationnel d'un champ de vecteurs de l'espace

Définition

Soit $\vec{F}$ un champ de vecteurs de classe $\mathcal{C}^1$ de l'espace de fonctions composantes P, Q, R.
On appelle rotationnel de $\vec{F}$ le champ de vecteurs défini par :

$$\overrightarrow{\text{Rot}}\,\vec{F} = \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\vec{i} + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)\vec{j} + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\vec{k}$$

Remarque On peut montrer que ce vecteur est indépendant du choix du repère orthonormé direct $\mathcal{R}$.

Remarque Le rotationnel de $\vec{F}$ en un point caractérise, par un vecteur rotation, l'action de rotation de $\vec{F}$ en ce point.

Exemple Pour $\vec{F}(M) = \overrightarrow{OM}$ on a $\overrightarrow{\text{Rot}}\,\vec{F} = \vec{0}$.

Exemple Pour $\vec{F}(M) = \vec{\omega} \wedge \overrightarrow{OM}$ on a $\overrightarrow{\text{Rot}}\,\vec{F} = 2\vec{\omega}$.

Proposition

Soient $\vec{F}, \vec{G}$ deux champs de vecteurs de classe $\mathcal{C}^1$, F un champ scalaire de classe $\mathcal{C}^1$ et $\lambda \in \mathbb{R}$.
On a

$$\overrightarrow{\text{Rot}}\,(\lambda.\vec{F}) = \lambda.\overrightarrow{\text{Rot}}\,\vec{F},\ \overrightarrow{\text{Rot}}\,(\vec{F} + \vec{G}) = \overrightarrow{\text{Rot}}\,\vec{F} + \overrightarrow{\text{Rot}}\,\vec{G}$$

et

$$\overrightarrow{\text{Rot}}\,(F.\vec{G}) = F.\overrightarrow{\text{Rot}}\vec{G} + \overrightarrow{\text{grad}}\,F \wedge \vec{G}$$

dém. :
Par opérations sur les dérivées partielles...
□

Théorème

Si F est un champ scalaire de classe $\mathcal{C}^2$ alors

$$\overrightarrow{\text{Rot}}(\overrightarrow{\text{grad}}\,F) = \vec{0}$$

Si $\vec{F}$ est un champ de vecteurs de classe $\mathcal{C}^2$ alors

$$\text{div}(\overrightarrow{\text{Rot}}\,\vec{F}) = 0$$

dém. :
Par opérations sur les dérivées partielles et exploitation du théorème de Schwarz.
□

Remarque Si $\vec{F}$ dérive d'un champ scalaire alors $\overrightarrow{\text{Rot}}\,\vec{F} = \vec{0}$.
De plus, la réciproque est vraie lorsque le domaine de définition est étoilé.

23.6.7 Opérateur nabla

On introduit l'opérateur différentiel nabla

$$\vec{\nabla} = \frac{\partial}{\partial x}\vec{i} + \frac{\partial}{\partial y}\vec{j} \text{ (resp. } \vec{\nabla} = \frac{\partial}{\partial x}\vec{i} + \frac{\partial}{\partial y}\vec{j} + \frac{\partial}{\partial z}\vec{k}\text{).}$$

Formellement, on a :

$$\overrightarrow{\text{grad}}\,F = \vec{\nabla}(F), \text{div}\vec{F} = \vec{\nabla}.\vec{F} \text{ et } \overrightarrow{\text{Rot}}\,\vec{F} = \vec{\nabla} \wedge \vec{F}$$

La propriété $\overrightarrow{\text{Rot}}(\overrightarrow{\text{grad}}\,f) = \vec{0}$ se relit

$$\vec{\nabla} \wedge (\vec{\nabla} f) = \vec{0}$$

La propriété $\text{div}(\overrightarrow{\text{Rot}}\,\vec{F}) = 0$ se relit

$$\vec{\nabla}.(\vec{\nabla} \wedge \vec{F}) = 0$$

Table des matières

II Analyse 425