

Chapitre 1

Logique et ensembles

Sommaire

1.1 Rudiments de logique	2
1.1.1 Propositions, démonstrations, etc.	2
1.1.2 Ensembles, éléments	3
1.1.3 Propriétés portant sur les éléments d'un ensemble	3
1.1.4 Opérations sur les propositions	4
1.1.5 Quantificateurs	4
1.1.6 Quelques synonymies classiques	5
1.1.7 Conditions nécessaires et/ou suffisantes	6
1.2 Raisonnements classiques	6
1.2.1 Conseils appuyés pour bien rédiger	6
1.2.2 Quelques figures usuelles du raisonnement	7
1.2.3 L'axiome de récurrence	8
1.2.4 Raisonnement par récurrence	8
1.2.5 Raisonnement par analyse-synthèse	10
1.2.6 Résolutions d'équations ou d'inéquations	10
1.2.7 Équations ou inéquations à un paramètre	11
1.3 Ensembles	12
1.3.1 Opérations sur les ensembles	12
1.3.2 Ensemble des parties d'un ensemble	13
1.3.3 Opérations sur les parties d'un ensemble	13
1.3.4 Produit cartésien d'un nombre fini d'ensembles	14
1.4 Applications	14
1.4.1 Applications entre ensembles non vides	14
1.4.2 Famille indexée par un ensemble non vide	15
1.4.3 Fonction indicatrice d'une partie	16
1.4.4 Restriction et prolongement	17
1.4.5 Image directe d'une partie par une application	17
1.4.6 Image réciproque d'une partie par une application	18
1.4.7 Composition d'applications	19
1.5 Injections, surjections, bijections	19
1.5.1 Applications injectives	19
1.5.2 Applications surjectives	20
1.5.3 Applications bijectives	20

1.6 Relations binaires	21
1.6.1 Généralités	21
1.6.2 Propriétés éventuelles des relations binaires	22
1.6.3 Relations d'ordre	22
1.6.4 Relations d'équivalence, classes d'équivalence	24
1.6.5 Congruence modulo un réel strictement positif	24
1.6.6 Division euclidienne et congruence modulo un entier	25

1.1 Rudiments de logique

1.1.1 Propositions, démonstrations, etc.

Définition 1.1.1

Une *proposition* est un énoncé dont on doit pouvoir dire qu'il est « vrai » ou « faux ».

On notera V et F (ou encore 1 et 0) les deux valeurs logiques possibles d'une proposition.

Exemples :

- « l'entier 2011 est premier » est une proposition vraie
- « l'entier 2012 est premier » est une proposition fausse
- « l'entier 2014 est une somme de deux carrés » est une proposition fausse (prouvez-le)
- « l'entier 2017 est somme de deux carrés » est une proposition vraie (en effet $2017 = 9^2 + 44^2$)

Définition 1.1.2

Certaines propositions sont déclarées vraies à priori : ce sont les *axiomes*. Sinon la véracité (ou la fausseté) d'une proposition doit résulter d'une *démonstration* (d'une *preuve*).

Remarque : dans le cadre d'un cours de mathématiques, quand on énonce une proposition, c'est pour affirmer qu'elle est vraie (et qu'on va la démontrer)!

Définition 1.1.3

Un *théorème* est une proposition vraie particulièrement importante.

Un *lemme* est une proposition vraie, utile à la démonstration d'une proposition plus importante.

Un *corollaire* est une proposition vraie, conséquence immédiate d'une autre proposition vraie.

Une *conjecture* est une proposition qu'on pense généralement vraie, sans en avoir de preuve.

Exemples :

- « l'axiome de récurrence » dans \mathbb{N} , « l'axiome de la borne supérieure » dans \mathbb{R} .
- « le théorème de Pythagore », le « théorème de Rolle », « le théorème de Bolzano-Weierstrass ».
- le « lemme des bergers », le « lemme de Gauss », le « lemme des noyaux ».
- la « conjecture de Syracuse », la « conjecture de Goldbach »
- la « conjecture de Fermat » est devenue le « grand théorème de Fermat » en 1994.

1.1.2 Ensembles, éléments

On ne se risque pas à donner une définition précise de ces notions premières.

Définition 1.1.4

On dit qu'un *ensemble* E est constitué d'*éléments* et qu'un élément a *appartient* à E (on écrit : $a \in E$) ou n'appartient pas à E (on écrit : $a \notin E$).

Deux ensembles E, F sont dits *égaux* (on note $E = F$) s'ils sont constitués des mêmes éléments.

Par convention l'*ensemble vide*, noté \emptyset , est l'ensemble ne contenant aucun élément.

Quelques remarques

- Un même objet mathématique peut, selon les circonstances, être vu comme un ensemble, ou comme un élément d'un ensemble. Par exemple, l'intervalle $[0, 1]$ est un *ensemble* de nombres réels, mais c'est également un *élément* de l'ensemble des intervalles de \mathbb{R} .
- Un ensemble *fini* peut être défini en *extension*, c'est-à-dire par la liste (non ordonnée) de ses éléments. C'est le cas par exemple de l'ensemble $E = \{1, 3, 6, 10, 15, 21, 28, 36, 45, 55\}$. Dans une écriture comme $\{a, b, c, \dots\}$ les éléments a, b, c , etc. sont à priori supposés distincts, et l'ordre dans lequel ils sont donnés n'a aucune importance.
- Un ensemble E peut être défini en *compréhension* (par une propriété caractérisant ses éléments). Ainsi $E = \left\{ \frac{n(n+1)}{2}, n \in \mathbb{N}, 1 \leq n \leq 10 \right\}$ est une autre définition de $\{1, 3, 6, 10, 15, 21, 28, 36, 45, 55\}$.
- Il y a bien d'autres conventions pour définir ou nommer des ensembles. Par exemple :
 - Si a, b sont deux réels, $[a, b[$ est l'ensemble des réels x qui vérifient $a \leq x < b$.
 - Si E est un ensemble, $\mathcal{P}(E)$ est l'ensemble des parties de E .
 - Certains ensembles ont des noms consacrés par l'usage : $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}, \dots$

Définition 1.1.5

Par convention l'*ensemble vide*, noté \emptyset , est l'ensemble ne contenant aucun élément.

Un ensemble $\{a\}$, formé d'un seul élément, est appelé un *singleton*.

Un ensemble $\{a, b\}$, formé de deux éléments *distincts*, est appelé une *paire*.

1.1.3 Propriétés portant sur les éléments d'un ensemble

On sait que la proposition « *l'entier 2017 est somme de deux carrés* » est vraie ($2017 = 9^2 + 44^2$).

Mais l'énoncé « *l'entier n est une somme de deux carrés* », qui contient la variable *libre* n (c'est-à-dire à laquelle on n'a pas donné de contenu particulier) ne peut pas être considéré comme une proposition, car on ne peut lui attribuer la valeur « Vrai » ou « Faux » tant qu'on ne donne pas une valeur à n .

On parlera plutôt ici d'une propriété \mathcal{P} portant sur les entiers naturels n .

Avec cette définition, la proposition $\mathcal{P}(2017)$ est vraie (ou encore : l'entier 2017 vérifie la propriété \mathcal{P}) et la proposition $\mathcal{P}(2014)$ est fausse (l'entier 2014 ne vérifie pas la propriété \mathcal{P}).

Soit \mathcal{P} une propriété portant sur les éléments d'un ensemble E . Pour affirmer qu'un élément x de E vérifie (ou satisfait à) la propriété \mathcal{P} , on écrira simplement « $\mathcal{P}(x)$ » plutôt que « $\mathcal{P}(x)$ est vraie ».

Il nous arrivera de considérer des énoncés contenant plusieurs variables libres.

Par exemple, la propriété \mathcal{P} suivante : « *l'entier n est la somme de k carrés* »

1.1.4 Opérations sur les propositions

On peut créer de nouvelles propositions à l'aide d'« opérations » sur des propositions existantes.

Définition 1.1.6

À partir des propositions \mathcal{A} , \mathcal{B} , on définit ainsi :

- la *négation* « non \mathcal{A} » (ou encore $\overline{\mathcal{A}}$, ou encore $\neg \mathcal{A}$).
- la *disjonction* « \mathcal{A} ou \mathcal{B} » (notée également $\mathcal{A} \vee \mathcal{B}$).
- la *conjonction* « \mathcal{A} et \mathcal{B} » (notée également $\mathcal{A} \wedge \mathcal{B}$).
- l'*implication* : la proposition « \mathcal{A} implique \mathcal{B} » (notée $\mathcal{A} \Rightarrow \mathcal{B}$)
- l'*équivalence* : la proposition « \mathcal{A} équivaut à \mathcal{B} » (notée $\mathcal{A} \Leftrightarrow \mathcal{B}$)

La valeur des propositions précédentes est résumée dans les *tableaux de vérité* ci-dessous.

\mathcal{A}	$\overline{\mathcal{A}}$
V	F
F	V

\mathcal{A}	\mathcal{B}	\mathcal{A} ou \mathcal{B}
V	V	V
V	F	V
F	V	V
F	F	F

\mathcal{A}	\mathcal{B}	\mathcal{A} et \mathcal{B}
V	V	V
V	F	F
F	V	F
F	F	F

\mathcal{A}	\mathcal{B}	$\mathcal{A} \Rightarrow \mathcal{B}$
V	V	V
V	F	F
F	V	V
F	F	V

\mathcal{A}	\mathcal{B}	$\mathcal{A} \Leftrightarrow \mathcal{B}$
V	V	V
V	F	F
F	V	F
F	F	V

Au vu des tableaux de vérités précédents, on peut donc dire, d'une façon moins formelle :

- La proposition $\overline{\mathcal{A}}$ est vraie quand \mathcal{A} est fausse, et fausse quand \mathcal{A} est vraie.
- La proposition « \mathcal{A} ou \mathcal{B} » est vraie quand l'une au moins des deux propositions \mathcal{A} , \mathcal{B} est vraie.
- La proposition « \mathcal{A} et \mathcal{B} » est vraie quand les deux propositions \mathcal{A} , \mathcal{B} sont vraies.
- « $\mathcal{A} \Rightarrow \mathcal{B}$ » est vraie quand \mathcal{A} est fausse (le « faux implique n'importe quoi ») ou quand \mathcal{B} est vraie.
- « $\mathcal{A} \Leftrightarrow \mathcal{B}$ » est vraie quand \mathcal{A} , \mathcal{B} sont toutes les deux vraies ou toutes les deux fausses.

Les opérations précédentes peuvent être répétées pour former des propositions $\mathcal{P}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \dots)$ dépendant de propositions initiales \mathcal{A} , \mathcal{B} , \mathcal{C} , etc.

Deux telles propositions $\mathcal{P}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \dots)$ et $\mathcal{Q}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \dots)$ sont dites *synonymes* si elles ont le même tableau de vérité, c'est-à-dire si l'équivalence $\mathcal{P}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \dots) \Leftrightarrow \mathcal{Q}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \dots)$ est *toujours* vraie (c'est-à-dire quelle que soit la valeur de vérité des propositions \mathcal{A} , \mathcal{B} , \mathcal{C} , ...).

1.1.5 Quantificateurs

Définition 1.1.7

Soit \mathcal{P} une propriété portant sur les éléments d'un ensemble E .

- La proposition « $\exists x \in E, \mathcal{P}(x)$ » dit qu'au moins un élément x de E vérifie la propriété \mathcal{P} .
On dit que « \exists » est le *quantificateur existentiel*.
- La proposition « $\forall x \in E, \mathcal{P}(x)$ » exprime que tout élément x de E vérifie la propriété \mathcal{P} .
On dit que « \forall » est le *quantificateur universel*.
- La proposition « $\exists! x \in E, \mathcal{P}(x)$ » exprime qu'un et un seul élément x de E vérifie la propriété \mathcal{P} .

De l'importance de la chronologie dans une phrase à plusieurs quantificateurs

On peut former des propositions « à tiroirs » avec plusieurs quantificateurs, notamment sur des énoncés $\mathcal{P}(x, y, \dots)$ à plusieurs variables. Dans ce cas, on prendra garde à l'ordre des quantificateurs, notamment dans les alternances de « \forall » et de « \exists ».

Ainsi les propositions $\begin{cases} \mathcal{A} : \forall x \in E, \exists y \in F, \mathcal{P}(x, y) \\ \mathcal{B} : \exists y \in F, \forall x \in E, \mathcal{P}(x, y) \end{cases}$ ont souvent des significations très différentes.

Le mieux est de penser à la traduction de ces deux propositions « en français ».

En effet, la proposition \mathcal{A} exprime que pour tout x de E , il existe un élément de F (dont la valeur dépend a priori du x qu'on vient d'évoquer) tel que la proposition $\mathcal{P}(x, y)$ soit vraie.

Quant à la proposition \mathcal{B} , elle exprime qu'il existe un certain élément y de F tel que, pour tout x de E , la proposition $\mathcal{P}(x, y)$ soit vraie.

Par exemple, la proposition « $\forall x \in \mathbb{N}, \exists y \in \mathbb{N}, x \leq y$ » (qui exprime que pour tout x de \mathbb{N} , il existe un y dans \mathbb{N} qui lui est supérieur ou égal) est évidemment vraie (prendre $y = x$, par exemple).

En revanche, la proposition « $\exists y \in \mathbb{N}, \forall x \in \mathbb{N}, x \leq y$ » (qui exprime qu'il existe un entier naturel y supérieur ou égal à tous les entiers naturels) est manifestement fausse.

Bien distinguer le « mode texte » et « mode mathématique »

Considérons la proposition suivante : « tout nombre réel est la puissance troisième d'un nombre réel ».

On pourra tout aussi bien écrire : « pour tout réel x , il existe un réel y tel que $y^3 = x$ ».

En revanche, on n'écrira pas :

- « pour tout $x \in \mathbb{R}$, il existe $y \in \mathbb{R}$ tq $y^3 = x$ »
- « $\forall x \in \mathbb{R}$, il existe $y \in \mathbb{R}$ tel que $y^3 = x$ »

Le programme le dit explicitement : « l'emploi de quantificateurs en guise d'abréviations est exclu »

Le mieux est de bien distinguer dans quel mode on écrit : en « mode texte » (c'est-à-dire en bon français, sans utiliser de quantificateurs) ou en « mode mathématique ».

Proposition 1.1.1 (négation d'une proposition avec quantificateurs)

Soit \mathcal{P} une propriété portant sur les éléments d'un ensemble E .

La négation de la proposition « $\exists x \in E, \mathcal{P}(x)$ » est « $\forall x \in E, \text{non } \mathcal{P}(x)$ »

La négation de la proposition « $\forall x \in E, \mathcal{P}(x)$ » est « $\exists x \in E, \text{non } \mathcal{P}(x)$ »

1.1.6 Quelques synonymies classiques

Soit $\mathcal{A}, \mathcal{B}, \mathcal{C}$, etc. des propositions.

Les équivalences suivantes sont toujours vraies :

$$\text{Double négation : } \text{non}(\text{non } \mathcal{A}) \Leftrightarrow \mathcal{A} \qquad \text{Idempotence : } \begin{cases} (\mathcal{A} \text{ et } \mathcal{A}) \Leftrightarrow \mathcal{A} \\ (\mathcal{A} \text{ ou } \mathcal{A}) \Leftrightarrow \mathcal{A} \end{cases}$$

$$\text{Commutativité : } \begin{cases} (\mathcal{A} \text{ et } \mathcal{B}) \Leftrightarrow (\mathcal{B} \text{ et } \mathcal{A}) \\ (\mathcal{A} \text{ ou } \mathcal{B}) \Leftrightarrow (\mathcal{B} \text{ ou } \mathcal{A}) \end{cases} \qquad \text{Associativité : } \begin{cases} ((\mathcal{A} \text{ et } \mathcal{B}) \text{ et } \mathcal{C}) \Leftrightarrow (\mathcal{A} \text{ et } (\mathcal{B} \text{ et } \mathcal{C})) \\ ((\mathcal{A} \text{ ou } \mathcal{B}) \text{ ou } \mathcal{C}) \Leftrightarrow (\mathcal{A} \text{ ou } (\mathcal{B} \text{ ou } \mathcal{C})) \end{cases}$$

$$\text{Dualité ou encore Lois de De Morgan : } \begin{cases} \text{non}(\mathcal{A} \text{ et } \mathcal{B}) \Leftrightarrow ((\text{non } \mathcal{A}) \text{ ou } (\text{non } \mathcal{B})) \\ \text{non}(\mathcal{A} \text{ ou } \mathcal{B}) \Leftrightarrow ((\text{non } \mathcal{A}) \text{ et } (\text{non } \mathcal{B})) \end{cases}$$

Double implication : $(\mathcal{A} \Leftrightarrow \mathcal{B}) \Leftrightarrow ((\mathcal{A} \Rightarrow \mathcal{B}) \text{ et } (\mathcal{B} \Rightarrow \mathcal{A}))$

Distributivité :
$$\begin{cases} (\mathcal{A} \text{ ou } (\mathcal{B} \text{ et } \mathcal{C})) \Leftrightarrow ((\mathcal{A} \text{ ou } \mathcal{B}) \text{ et } (\mathcal{A} \text{ ou } \mathcal{C})) \\ (\mathcal{A} \text{ et } (\mathcal{B} \text{ ou } \mathcal{C})) \Leftrightarrow ((\mathcal{A} \text{ et } \mathcal{B}) \text{ ou } (\mathcal{A} \text{ et } \mathcal{C})) \end{cases}$$

1.1.7 Conditions nécessaires et/ou suffisantes

On considère deux propositions \mathcal{A} et \mathcal{B} .

On suppose que « $\mathcal{A} \Rightarrow \mathcal{B}$ » est vraie.

Le tableau ci-contre illustre les trois cas possibles :

\mathcal{A}	\mathcal{B}	$\mathcal{A} \Rightarrow \mathcal{B}$
V	V	V
F	V	V
F	F	V

Définition 1.1.8

Soit \mathcal{A} et \mathcal{B} deux propositions.

Pour exprimer que la proposition « $\mathcal{A} \Rightarrow \mathcal{B}$ » est vraie, on dit indifféremment :

- La proposition \mathcal{A} est une *condition suffisante* de la proposition \mathcal{B} .
- La proposition \mathcal{B} est une *condition nécessaire* de la proposition \mathcal{A} .
- Pour que \mathcal{A} (soit vraie) il *faut* que \mathcal{B} (soit vraie).
- Pour que \mathcal{B} (soit vraie), il *suffit* que \mathcal{A} (soit vraie).
- \mathcal{B} (est vraie) *si* \mathcal{A} (est vraie).
- \mathcal{A} (est vraie) *seulement si* \mathcal{B} (est vraie).

Définition 1.1.9

Soit \mathcal{A} et \mathcal{B} deux propositions.

Pour exprimer que « $\mathcal{A} \Leftrightarrow \mathcal{B}$ » est vraie, on dit indifféremment :

- \mathcal{A} est une condition *nécessaire et suffisante* (CNS) de \mathcal{B} .
- \mathcal{A} (est vraie) *si et seulement si* \mathcal{B} (est vraie).
- Pour que \mathcal{A} (soit vraie), il *faut et il suffit* que \mathcal{B} (soit vraie).

Dans ces énoncés on peut bien sûr échanger le rôle de \mathcal{A} et \mathcal{B} .

1.2 Raisonnements classiques

1.2.1 Conseils appuyés pour bien rédiger

Dans le raisonnement logique, la syntaxe est primordiale. Elle va de pair avec la clarté du style.

Quelques bonnes habitudes doivent être prises :

- Indiquer clairement les hypothèses de la démonstration, et quel résultat on veut obtenir.
Si par exemple, on doit prouver que l'implication $\mathcal{A} \Rightarrow \mathcal{B}$ est vraie, on commencera en général par dire : « supposons \mathcal{A} , et montrons \mathcal{B} ». À partir de l'hypothèse sur \mathcal{A} , on procédera ensuite par conditions nécessaires, pour finalement prouver que \mathcal{B} est vraie.
Attention : on prouve ici que \mathcal{B} est vraie *si* \mathcal{A} est vrai, et pas que \mathcal{B} est vrai dans l'absolu.
À ce sujet, voir plus loin la règle de raisonnement dite du « modus ponens ».

- Mettre en évidence les liens logiques entre les phases successives de la démonstration.
Le symbole « \Rightarrow » n'est pas innocent. Son emploi doit être justifié.
- Ne pas mélanger les symboles « \Rightarrow » et « \Leftrightarrow ».
- Dans une proposition « à tiroirs », utiliser des parenthèses pour lever toute ambiguïté.
Ainsi la proposition $(\mathcal{A} \Rightarrow \mathcal{B}) \Rightarrow \mathcal{C}$ n'est pas synonyme de $\mathcal{A} \Rightarrow (\mathcal{B} \Rightarrow \mathcal{C})$.
- Éviter d'utiliser exclusivement le langage de la logique formelle (propositions, quantificateurs) là où on peut s'exprimer « en français ». En particulier, on ne mélangera pas les deux styles. On évitera d'écrire, par exemple : « pour tout $x \in E, \dots$ »).
- Varier le style, pour éviter toute sécheresse. Le mot « donc », par exemple, possède plusieurs synonymes : « on en déduit », « il s'ensuit », « par conséquent », etc.
- Soit \mathcal{P} une propriété portant sur les éléments d'un ensemble E .
Si on veut montrer la proposition « $\forall x \in E, \mathcal{P}(x)$ », on commencera souvent par écrire : « Soit x un élément de E , montrons que x vérifie la propriété \mathcal{P} ». Le fait de fixer un élément de x n'est ici pas restrictif, tant que cet élément est quelconque dans E .
Si on sait que la proposition « $\exists x \in E, \mathcal{P}(x)$ » est vraie, et si on veut en déduire quelque chose, on commencera en général par écrire « Soit x dans E , vérifiant la propriété \mathcal{P} ».
Le « Soit x » est un important car il installe dans la démonstration. Il crée et nomme un objet précis de E , sur lequel on va pouvoir travailler (le « $\exists x \in E$ » ne crée rien, lui).
- Soit \mathcal{P} et \mathcal{Q} deux propriétés portant sur les éléments d'un ensemble E .
La proposition $(\exists x \in E, \mathcal{P}(x))$ et $(\exists x \in E, \mathcal{Q}(x))$ n'implique pas, en général $\exists x \in E, (\mathcal{P}(x) \text{ et } \mathcal{Q}(x))$
La raison en est que le quantificateur \exists est *mutificateur*. En d'autres termes, la variable x dont il est question dans $(\exists x \in E, \mathcal{P}(x))$ est *muette* : elle ne désigne aucun élément particulier de E .
Il n'y a donc aucune raison de croire que les deux x de $(\exists x \in E, \mathcal{P}(x))$ et $(\exists x \in E, \mathcal{Q}(x))$ se réfèrent à un même élément de E .
Si on doit déduire quelque chose de $(\exists x \in E, \mathcal{P}(x))$ et $(\exists x \in E, \mathcal{Q}(x))$, on pourra commencer par écrire : soit x dans E tel que $\mathcal{P}(x)$, et soit x' dans E tel que $\mathcal{Q}(x')$.

1.2.2 Quelques figures usuelles du raisonnement

Proposition 1.2.1 (la règle du « modus ponens »)

L'implication suivante est toujours vraie : $((\mathcal{A} \Rightarrow \mathcal{B}) \text{ et } \mathcal{A}) \Rightarrow \mathcal{B}$

Proposition 1.2.2 (le raisonnement par contraposition)

L'équivalence suivante est toujours vraie : $(\mathcal{A} \Rightarrow \mathcal{B}) \Leftrightarrow ((\text{non } \mathcal{B}) \Rightarrow (\text{non } \mathcal{A}))$.

Proposition 1.2.3 (le raisonnement par l'absurde)

L'équivalence suivante est toujours vraie : $(\mathcal{A} \Rightarrow \mathcal{B}) \Leftrightarrow \text{non } (\mathcal{A} \text{ et non } \mathcal{B})$.

Proposition 1.2.4 (le syllogisme)

Soit \mathcal{A} , \mathcal{B} et \mathcal{C} trois propositions.

L'implication suivante est toujours vraie : $((\mathcal{A} \Rightarrow \mathcal{B}) \text{ et } (\mathcal{B} \Rightarrow \mathcal{C})) \Rightarrow (\mathcal{A} \Rightarrow \mathcal{C})$.

Proposition 1.2.5 (la disjonction des cas)

Soit A , B et C trois propositions.

L'implication suivante est vraie : $((A \Rightarrow C) \text{ et } (B \Rightarrow C)) \Rightarrow ((A \text{ ou } B) \Rightarrow C)$

1.2.3 L'axiome de récurrence

On admet l'existence et les propriétés de l'ensemble totalement ordonné \mathbb{N} , et notamment :

Axiome

Toute partie non vide de \mathbb{N} possède un plus petit élément.

Proposition 1.2.6

Toute partie non vide majorée de \mathbb{N} possède un plus grand élément.

Remarques

- Soit n dans \mathbb{N} . L'ensemble $A = \{m \in \mathbb{N}, m > n\}$ est non vide.
Le plus petit élément de A est bien sûr $n + 1$ (c'est le *successeur* de n).
Autrement dit, pour tout n de \mathbb{N} , on a : $m > n \Leftrightarrow m \geq n + 1$.
- Soit n dans \mathbb{N}^* . L'ensemble A des m de \mathbb{N} tel que $m < n$ est non vide (il contient 0).
Le plus grand élément de cet ensemble est bien sûr $n - 1$ (c'est le *prédécesseur* de n).
Autrement dit, pour tout n de \mathbb{N} , on a : $m < n \Leftrightarrow m \leq n - 1$.

Proposition 1.2.7 (principe de récurrence)

Soit A une partie de \mathbb{N} , contenant 0.

On suppose que si A contient un entier n , alors A contient $n + 1$. Dans ces conditions, $A = \mathbb{N}$.

Remarque

On peut modifier légèrement l'énoncé précédent, en « démarrant » en un entier n_0 plutôt qu'en 0.

Soit A une partie de \mathbb{N} , contenant n_0 .

On suppose que si A contient un entier $n \geq n_0$, alors A contient $n + 1$.

Dans ces conditions, A contient $[n_0, +\infty[$.

1.2.4 Raisonnement par récurrence**Raisonnement par récurrence simple****Proposition 1.2.8** (récurrence simple)

Soit \mathcal{P} une propriété portant sur les entiers naturels.

Soit n_0 un entier naturel. On suppose que $\mathcal{P}(n_0)$ est vraie.

On suppose également, pour tout entier $n \geq n_0$, que l'implication $\mathcal{P}(n) \Rightarrow \mathcal{P}(n + 1)$ est vraie.

Alors la proposition $\mathcal{P}(n)$ est vraie pour tout entier $n \geq n_0$.

Voici donc comment montrer qu'une propriété $\mathcal{P}(n)$ est vraie pour tous les entiers naturels :

- On vérifie que l'entier n_0 satisfait à la propriété : c'est le *pas initial* de la récurrence.
La plupart du temps, on a bien sûr $n_0 = 0$, ou $n_0 = 1$...
- On se **donne** ensuite un entier $n \geq n_0$, et on suppose que $\mathcal{P}(n)$ est vraie (*hypothèse de récurrence*).
On démontre alors que $\mathcal{P}(n+1)$ est vraie (c'est le « *passage du rang n au rang $n+1$* »).
On exprime l'implication $\mathcal{P}(n) \Rightarrow \mathcal{P}(n+1)$ en disant que la propriété \mathcal{P} est *héréditaire*.
- On conclut en annonçant que, par récurrence, la propriété est vraie pour tout entier $n \geq n_0$.

Variantes du raisonnement par récurrence

Une variante réside dans la manière d'avancer dans la récurrence. Il arrive en effet que l'hypothèse $\mathcal{P}(n)$ seule soit insuffisante pour démontrer $\mathcal{P}(n+1)$.

Le cas le plus fréquent est celui de la *récurrence double*, où le pas initial et l'hypothèse de récurrence portent sur deux entiers consécutifs.

Proposition 1.2.9 (récurrence de pas 2)

Soit \mathcal{P} une propriété portant sur les entiers naturels.

Soit n_0 un entier naturel. On suppose que $\mathcal{P}(n_0)$ et $\mathcal{P}(n_0+1)$ sont vraies.

On suppose que pour tout $n \geq n_0$, l'implication « $(\mathcal{P}(n) \text{ et } \mathcal{P}(n+1)) \Rightarrow \mathcal{P}(n+2)$ » est vraie.

Alors la proposition $\mathcal{P}(n)$ est vraie pour tout $n \geq n_0$.

Il reste à voir une variante importante : le raisonnement par récurrence forte.

Pour prouver $\mathcal{P}(n+1)$, on peut en effet utiliser tout ou partie des hypothèses $\mathcal{P}(n_0), \mathcal{P}(n_0+1), \dots, \mathcal{P}(n)$.

Proposition 1.2.10 (récurrence forte)

Soit \mathcal{P} une propriété portant sur les entiers naturels.

Soit n_0 un entier naturel. On suppose que $\mathcal{P}(n_0)$ est vraie.

On suppose que pour tout $n \geq n_0$, l'implication $(\mathcal{P}(n_0) \text{ et } \mathcal{P}(n_0+1) \cdots \text{ et } \mathcal{P}(n)) \Rightarrow \mathcal{P}(n+1)$ est vraie.

Alors la proposition $\mathcal{P}(n)$ est vraie pour tout $n \geq n_0$.

Pratique du raisonnement par récurrence

Voici enfin quelques conseils pour « réussir » un raisonnement par récurrence :

- Ne pas oublier le « pas initial » (la propriété est souvent triviale, mais on **doit** la prouver).
- Ne **jamais** écrire : « Supposons que pour **tout** n , $\mathcal{P}(n)$. Montrons $\mathcal{P}(n+1)$ ».
Il faut en fait écrire : « **Soit** n un entier naturel ; on suppose $\mathcal{P}(n)$. Montrons $\mathcal{P}(n+1)$ ».
- Bien articuler le pas initial et l'hypothèse de récurrence.
Si le pas initial est n_0 , et si on veut démontrer $\mathcal{P}(n) \Rightarrow \mathcal{P}(n+1)$, alors on doit fixer $n \geq n_0$.
On peut tout à fait prouver $\mathcal{P}(n-1) \Rightarrow \mathcal{P}(n)$, mais dans ce cas on doit fixer $n > n_0$.
- Bien séparer le « passage du rang n au rang $n+1$ », où l'entier n est **fixé**, et la conclusion finale (qui est obligatoire, et qui doit porter sur **tous** les entiers $n \geq n_0$).

Une phrase finale rituelle est « *ce qui prouve la propriété au rang $n+1$ et achève la récurrence* ».

1.2.5 Raisonnement par analyse-synthèse

Le *raisonnement par analyse-synthèse* est souvent utilisé quand on demande de prouver l'existence (ou la non existence) de solutions à un problème donné (et d'identifier la ou les solutions éventuelles).

Ce raisonnement se déroule alors en deux étapes (trois étapes avec la conclusion) :

- L'analyse :
on suppose que le problème admet une solution, et on procède par **conditions nécessaires**, en accumulant suffisamment de renseignements pour dégager un ou plusieurs « candidats-solutions ».
- La synthèse :
on vérifie si les « candidats » sont *effectivement* solutions du problème.
- La conclusion :
on donne la ou les solutions du problème (s'il y en a).

1.2.6 Résolutions d'équations ou d'inéquations

Commençons par une définition banale :

Définition 1.2.1

Soit f une fonction définie sur une partie \mathcal{D} de \mathbb{R} , à valeurs réelles.

Trouver l'ensemble \mathcal{S} (éventuellement vide) des réels x de \mathcal{D} tels que $f(x)$ soit nul (resp. positif ou nul, resp. strictement positif), c'est « résoudre » l'équation $f(x) = 0$ (resp. l'inéquation $f(x) \geq 0$, resp. l'inéquation $f(x) > 0$).

Les éléments x de \mathcal{S} sont appelés les *solutions* de l'équation (resp. de l'inéquation).

Généralisations

- Bien sûr, les équations $f(x) = \lambda$, ou $f(x) \geq \lambda$, ou $f(x) > \lambda$ (avec λ réel), et plus généralement les équations $f(x) = g(x)$, ou les inéquations $f(x) \geq g(x)$, $f(x) > g(x)$, $f(x) \leq g(x)$ ou $f(x) < g(x)$ se ramènent instantanément à l'une des situations évoquées dans la définition précédente.
- Soient f et g sont deux fonctions à valeurs réelles, et définies respectivement sur \mathcal{D}_f et \mathcal{D}_g .
Résoudre le système $\begin{cases} f(x) = 0 \\ g(x) = 0 \end{cases}$ c'est trouver les solutions communes à $f(x) = 0$ et $g(x) = 0$.
On généralise à la notion de système de n équations (ou inéquations, ou un mélange des deux).
- Plus généralement encore, et en introduisant des « fonctions de plusieurs variables », on peut être amené à considérer des systèmes d'équations (ou d'inéquations) à plusieurs inconnues x, y, z etc.
- Autre point de vue possible : on sera amené à résoudre des équations dont la ou les inconnues sont supposées appartenir à \mathbb{N} , ou à \mathbb{Z} , ou à \mathbb{Q} , ou à \mathbb{C} .
- Tout ça recouvre donc un grand nombre de situations différentes, et ça peut devenir très compliqué !
Dans certains cas, par exemple, on prouvera qu'une équation possède une solution sans pouvoir la calculer explicitement (considérer par exemple l'équation $x^3 + x + 1 = 0$, avec x cherché dans \mathbb{R}).

Deux grandes méthodes de résolution

Il est impossible de dresser la liste des « méthodes », tant les situations peuvent être différentes.

Contentons-nous d'imaginer une « équation » (E) (en donnant un sens très général à ce terme).

On peut chercher à résoudre (E) « par équivalences », ou « par conditions nécessaires ».

– **Résoudre (E) par équivalences :**

C'est former une succession d'équations (E) , puis (E_2) , (E_3) , etc. possédant exactement le même ensemble de solutions (ce qu'il est bien sûr primordial et obligatoire de justifier, sauf pour les équivalences vraiment évidentes).

On arrive alors à une dernière équation (E_n) dont l'ensemble des solutions est clair : on a ainsi obtenu l'ensemble des solutions de l'équation initiale (E) .

– **Résoudre (E) par implications (ou encore par « conditions nécessaires ») :**

C'est former des équations (E) , (E_2) , (E_3) , etc. telles qu'à chaque étape « (E_k) implique (E_{k+1}) ».

En termes plus précis, il s'agit de justifier que l'ensemble des solutions \mathcal{S}_k de (E_k) est inclus dans l'ensemble des solutions \mathcal{S}_{k+1} de (E_{k+1}) .

On arrive alors à une dernière équation (E_n) dont l'ensemble des solutions \mathcal{S}_n est clair.

Tout ce qu'on peut dire à ce stade est que l'ensemble \mathcal{S} des solutions de (E) (c'est lui qui nous intéresse) est *inclus* dans l'ensemble \mathcal{S}_n . Il importe alors de procéder à une *réciproque*, c'est-à-dire de vérifier, parmi les éléments de \mathcal{S}_n , qui est *effectivement* élément de \mathcal{S} .

Le prix consenti à établir la réciproque est compensé par le fait qu'il est souvent beaucoup facile de prouver les implications « $(E_k) \Rightarrow (E_{k+1})$ » que les équivalences « $(E_k) \Leftrightarrow (E_{k+1})$ ».

Si l'une des étapes contient une implication $(E_k) \Rightarrow (E_{k+1})$ qui n'est pas une équivalence, il ne sert à rien d'écrire des équivalences ici ou là (autant n'écrire que des implications, c'est plus sûr).

– **La clarté du style :**

Ce qui a été dit à la section 1.2.1 (conseils pour bien rédiger) s'applique particulièrement à la résolution d'équations. Il est absolument primordial d'éviter les successions interminables d'implications ou d'équivalences. Il **faut**, aussi souvent que possible, aérer le style et faire des pauses dans les implications (construire des phrases terminées par un point).

On enchaînera notamment les implications par des formules du type « il en découle », « on en déduit ». Pour les équivalences, on évitera l'utilisation abusive et répétée du symbole \Leftrightarrow , et on utilisera régulièrement des expressions comme « c'est-à-dire », ou « ce qui équivaut à », etc.

1.2.7 Équations ou inéquations à un paramètre

Définition 1.2.2

Soit \mathcal{M} une partie de \mathbb{R} , appelée ici ensemble des valeurs du paramètre m .

Pour tout m de \mathcal{M} , soit f_m une fonction définie sur une partie \mathcal{D}_m de \mathbb{R} , à valeurs réelles.

Trouver l'ensemble \mathcal{S}_m (éventuellement vide) des réels x de \mathcal{D}_m tels que $f_m(x)$ soit nul, c'est « résoudre » l'équation $f_m(x) = 0$. Les éléments x de \mathcal{S}_m sont appelés les *solutions* de l'équation (resp. de l'inéquation), pour la valeur m du paramètre.

Remarques

- On peut bien sûr généraliser aux inéquations à un paramètre, ou aux systèmes d'équations/inéquations, ou même à des systèmes à plusieurs paramètres...
- Évidemment, le nom donné au paramètre n'est pas toujours m : on trouve souvent λ par exemple !
- L'important est de comprendre que le paramètre (appelons-le m) n'est pas une *inconnue* du problème. Ce paramètre possède en fait une valeur, que nous ne connaissons pas, et il est probable que cette valeur conditionne la résolution de l'équation $(E_m) : f(x) = 0$ de l'inconnue x .

- Il est donc possible que l'équation (E_m) (dont l'énoncé dépend de la valeur attribuée à m) ne possède pas de solutions en x pour certaines valeurs de m , et qu'elle en possède pour d'autres (dans ce cas, il est probable que cette ou ces solutions dépendent de la valeur attribuée à m).
- Résoudre l'équation (E_m) , d'inconnue x et de paramètre m , c'est donc déterminer les valeurs (ou les intervalles de valeurs de m) pour lesquelles l'équation (E_m) possède des solutions (et lesquelles). Cette résolution prend alors la forme d'une « discussion suivant les valeurs du paramètre m ».
- Cette discussion permet souvent de dégager des valeurs particulières de m , et des situations plus générales (toujours pour m). Il est recommandé de traiter les cas particuliers en premier. Il convient aussi de ne pas « inventer » de cas particuliers (qui n'apparaissent que parce qu'on a effectué tel calcul plutôt que tel autre) et qui ne sont pas des cas particuliers « intrinsèques » au problème posé.

1.3 Ensembles

1.3.1 Opérations sur les ensembles

À partir de deux ensembles E et F , on peut en construire d'autres :

Définition 1.3.1 (intersection et réunion)

Soit E et F deux ensembles.

$E \cap F$ est l'ensemble formé des éléments qui sont à la fois dans E et dans F .

$E \cup F$ est l'ensemble formé des éléments qui sont dans l'un au moins des ensembles E et F .

Définition 1.3.2 (ensembles disjoints)

On dit que E, F sont *disjoints* si $E \cap F$ est vide.

Dans ce cas, on dit que $E \cup F$ est une *d'union disjointe*.

On ne confondra pas *distincts* et *disjoints* :

- Dire que E et F sont *distincts*, c'est dire : $(\exists x \in E, x \notin F)$ ou $(\exists x \in F, x \notin E)$.
- Dire que E et F sont *disjoints*, c'est dire : $(\forall x \in E, x \notin F)$ et $(\forall x \in F, x \notin E)$.

Définition 1.3.3 (différence et différence symétrique)

Soit E et F deux ensembles.

- *Différence* : l'ensemble $E \setminus F$ est formé des éléments qui sont dans E mais pas dans F .
- *Différence symétrique* : on note $E \Delta F$ l'ensemble $(E \cup F) \setminus (E \cap F)$.
C'est l'ensemble des éléments qui sont dans un et un seul des deux ensembles E et F .

Remarques :

- On dit encore que $E \setminus F$ est le *complémentaire* de F dans E et on peut le noter $\complement_E F$.
- Une définition équivalente de la différence symétrique est :
 $E \Delta F = (E \setminus F) \cup (F \setminus E)$ (c'est une union disjointe).

1.3.2 Ensemble des parties d'un ensemble

Définition 1.3.4

On dit qu'un ensemble F est *inclus* dans un ensemble E , et on note $F \subset E$, pour exprimer que tout élément de F est également élément de E .

On dit encore que E *contient* F , ou que F est une *partie* (ou un *sous-ensemble*) de E .

Définition 1.3.5

Soit E un ensemble. On note $\mathcal{P}(E)$ l'ensemble des parties de E .

Remarques

– Évidemment, si $E \subset F$ et $F \subset G$, alors $E \subset G$.

– On a l'équivalence $A \in \mathcal{P}(E) \Leftrightarrow A \subset E$.

De même : $a \in E \Leftrightarrow \{a\} \subset E \Leftrightarrow \{a\} \in \mathcal{P}(E)$.

Les ensembles E et \emptyset sont toujours des *éléments* de $\mathcal{P}(E)$.

– La réunion, l'intersection, la différence symétrique sont des *opérations binaires* sur $\mathcal{P}(E)$, en ce sens qu'à deux éléments de $\mathcal{P}(E)$ elles associent un élément de $\mathcal{P}(E)$.

– Si aucune confusion n'est à craindre sur l'ensemble E , on notera \bar{A} le complémentaire d'une partie A de E : c'est encore une partie de E .

Le passage au complémentaire est donc une opération *unaire* sur $\mathcal{P}(E)$.

1.3.3 Opérations sur les parties d'un ensemble

On observe que si $\mathcal{A}(x)$ et $\mathcal{B}(x)$ sont deux prédicats basés sur E , alors :

– On a l'égalité $\{x \in E, \mathcal{A}(x)\} \cup \{x \in E, \mathcal{B}(x)\} = \{x \in E, \mathcal{A}(x) \text{ ou } \mathcal{B}(x)\}$.

– On a l'égalité $\{x \in E, \mathcal{A}(x)\} \cap \{x \in E, \mathcal{B}(x)\} = \{x \in E, \mathcal{A}(x) \text{ et } \mathcal{B}(x)\}$.

– Le complémentaire dans E de $\{x \in E, \mathcal{A}(x)\}$ est $\{x \in E, \text{non } \mathcal{A}(x)\}$.

De cette remarque et des propriétés des opérations sur les propositions, on déduit les propriétés suivantes des opérations sur $\mathcal{P}(E)$. A, B, C désignent ici trois parties quelconques de E .

– *Double passage au complémentaire* : $\overline{\bar{A}} = A$. – *Idempotence* : $\begin{cases} A \cap A = A \\ A \cup A = A \end{cases}$

– *Commutativité* : $\begin{cases} A \cap B = B \cap A \\ A \cup B = B \cup A \end{cases}$ – *Associativité* : $\begin{cases} (A \cap B) \cap C = A \cap (B \cap C) \\ (A \cup B) \cup C = A \cup (B \cup C) \end{cases}$

– *Distributivité* : $\begin{cases} A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \\ A \cup (B \cap C) = (A \cup B) \cap (A \cup C) \end{cases}$ – *Dualité* : $\begin{cases} \overline{A \cap B} = \bar{A} \cup \bar{B} \\ \overline{A \cup B} = \bar{A} \cap \bar{B} \end{cases}$

– *Partie vide et partie pleine* : $\begin{cases} A \cup \emptyset = A & A \cap \emptyset = \emptyset \\ A \cap B = \emptyset \Leftrightarrow A \subset \bar{B} \end{cases} \quad \begin{cases} A \cap E = A & A \cup E = E \\ A \cup B = E \Leftrightarrow \bar{B} \subset A \end{cases}$

Proposition 1.3.1 (propriétés de l'inclusion)

Soit A, B, C trois parties quelconques d'un ensemble E . Les équivalences suivantes sont vraies :

$$\begin{aligned} A = B &\Leftrightarrow (A \subset B) \text{ et } (B \subset A) & A \subset B &\Leftrightarrow \overline{B} \subset \overline{A} \\ (A \cup B) \subset C &\Leftrightarrow (A \subset C) \text{ et } (B \subset C) & A \subset (B \cap C) &\Leftrightarrow (A \subset B) \text{ et } (A \subset C) \end{aligned}$$

Proposition 1.3.2 (propriétés de la différence symétrique)

Pour toutes parties A, B, C d'un ensemble E , on a les égalités suivantes :

$$\begin{aligned} A \Delta B &= B \Delta A & A \Delta A &= \emptyset & A \Delta E &= \overline{A} & A \Delta \emptyset &= A \\ A \Delta (B \Delta C) &= (A \Delta B) \Delta C & A \cap (B \Delta C) &= (A \cap B) \Delta (A \cap C) \end{aligned}$$

1.3.4 Produit cartésien d'un nombre fini d'ensembles**Définition 1.3.6** (n -uplets et produit cartésien)

Soit E_1, E_2, \dots, E_n n ensembles (non nécessairement distincts deux à deux), avec $n \geq 2$.

- Pour tout entier k (compris entre 1 et n), soit x_k un élément de l'ensemble E_k .
 (x_1, x_2, \dots, x_n) est appelé un n -uplet de composantes x_1, x_2, \dots, x_n (dans cet ordre).
- On appelle *produit cartésien* de E_1, E_2, \dots, E_n , et on note $E_1 \times E_2 \times \dots \times E_n$, l'ensemble des n -uplets (x_1, x_2, \dots, x_n) . Par exemple, $E \times F = \{(a, b), a \in E, b \in F\}$.

Remarques :

- Un n -uplet est donc le moyen de regrouper n éléments dans un ordre bien défini.
- On parle de *couple* si $n = 2$, de *triplet* si $n = 3$, de *quadruplet* si $n = 4$, etc.
- On ne confondra pas (par exemple) la *paire* $\{a, b\}$ avec le *couple* (a, b) :
 - Si a et b sont différents, les couples (a, b) et (b, a) désignent en effet deux objets différents, alors que $\{a, b\}$ et $\{b, a\}$ désignent le même ensemble.
 - De même si $a = b$: l'ensemble $\{a, b\}$ se réduit au singleton $\{a\}$, alors que (a, a) est toujours un couple (mais dont les deux composantes sont égales).
- Si E_1, E_2, \dots, E_n sont égaux à un même ensemble E , on note E^n plutôt que $E \times E \times \dots \times E$.
- Par définition, la *diagonale* de E^2 est l'ensemble $\Delta = \{(x, x), x \in E\}$.

1.4 Applications**1.4.1 Applications entres ensembles non vides****Définition 1.4.1**

Soient E et F deux ensembles non vides.

Une *application* (on dit aussi une *fonction*) f de E vers F est le moyen d'associer, à chaque élément x de E , un unique élément y de F . On note alors $y = f(x)$.

On exprime l'égalité $y = f(x)$ est l'*image* de x par f , ou que x est un *antécédent* de y par f .

On dit que E est l'*ensemble de départ* et que F l'*ensemble d'arrivée* de f .

Définition 1.4.2

Soit E et F deux ensembles.

On note $\mathcal{F}(E, F)$, ou encore F^E l'ensemble de toutes les applications de E vers F .

Si les deux ensembles E et F sont égaux, on note plus simplement $\mathcal{F}(E)$ (ou E^E).

Notations et remarques :

- Une application f de E vers F est souvent notée $E \xrightarrow{f} F$ ou $f : E \longrightarrow F$
 $x \mapsto y = f(x)$
- Si E est fini, on peut représenter $f : E \rightarrow F$ par la donnée de chacune des images par f .
On définit par exemple une application f de $E = \{a, b, c, d, e\}$ vers $F = \{t, u, v, w\}$ par :

$$f(a) = t, f(b) = t, f(c) = v, f(d) = w, f(e) = w$$

Comme on le voit sur cet exemple, tout élément de E possède une image et une seule (c'est la définition même d'une application) mais un élément donné de F peut très bien :

- ne posséder aucun antécédent (l'élément u n'en a pas)
- ou posséder un seul antécédent (celui de v est c)
- ou en posséder plusieurs (t et w en ont chacun deux) ou une infinité (pas ici...!)
- Deux applications f et g sont *égales* si :
 - elles ont le même ensemble de départ E et le même ensemble d'arrivée F .
 - pour tous x de E , on a $f(x) = g(x)$.

Définition 1.4.3 (application Identité)

Soit E un ensemble.

On définit l'application *identité* de E dans E , notée Id_E , par : $\forall x \in E, \text{Id}_E(x) = x$.

Définition 1.4.4 (applications constantes)

Une application f de E dans F est dite *constante* s'il existe un élément α de F , tel que, pour tout x de E , $f(x)$ soit égal à α .

1.4.2 Famille indexée par un ensemble non vide**Définition 1.4.5** (familles d'éléments d'un ensemble)

Soit I un ensemble non vide dont les éléments seront appelés *indices*. Soit E un ensemble.

Toute application x de I vers E est appelée une *famille d'éléments* de E , *indicée* par I .

On note x_i plutôt que $x(i)$, et on écrit $(x_i)_{i \in I}$ cette famille d'éléments.

On note E^I l'ensemble des familles d'éléments de E indicées par I : c'est donc une autre manière de désigner l'ensemble $\mathcal{F}(I, E)$.

Remarques et notations :

- La famille $(x_i)_{i \in I}$ est dite *finie* ou *infinie*, selon que l'ensemble I est fini ou infini.
- On ne confondra pas la famille $(x_i)_{i \in I}$, c'est-à-dire l'application x , avec l'ensemble $\{x_i, i \in I\}$ c'est-à-dire l'ensemble image de cette application.
En particulier, une famille infinie peut n'être constituée que d'un nombre fini d'éléments (et par exemple un seul si l'application x est constante !)

- Si I est un intervalle d'entiers $[m, n]$, où $m \leq n$, la famille est notée $(x_i)_{m \leq i \leq n}$.
- Si les x_i sont dans \mathbb{R} ou dans \mathbb{C} , leur somme et leur produit sont notés respectivement $\sum_{i=m}^n x_i$ et $\prod_{i=m}^n x_i$.

Définition 1.4.6 (famille de parties d'un ensemble)

Toute application de I vers $\mathcal{P}(E)$ est appelée une *famille de parties* de E , indicée par I .

Une telle famille s'écrit par exemple $(E_i)_{i \in I}$ (si on appelle E_i l'image de i par l'application).

Soit $(E_i)_{i \in I}$ une famille de parties de l'ensemble E .

La réunion et l'intersection de cette famille sont les parties de E définies par :

- $\bigcup_{i \in I} E_i = \{x \in E, \exists i \in I, x \in E_i\}$ (éléments de E appartenant à au moins un E_i)
- $\bigcap_{i \in I} E_i = \{x \in E, \forall i \in I, x \in E_i\}$ (éléments de E appartenant à tous les E_i).

Si I est un intervalle d'entiers $[m, n]$, avec $m \leq n$, on notera plutôt $\bigcup_{i=m}^n E_i$ et $\bigcap_{i=m}^n E_i$.

Définition 1.4.7 (partitions d'un ensemble)

On dit qu'une famille $(E_i)_{i \in I}$ de parties de E constitue une *partition* de E si :

- Aucun des ensembles E_i n'est vide : $\forall i \in I, E_i \neq \emptyset$.
- Les E_i sont disjoints deux à deux : $\forall (i, j) \in I^2$, avec $i \neq j$, $E_i \cap E_j = \emptyset$.
- La réunion des ensembles E_i est égale à E tout entier : $\bigcup_{i \in I} E_i = E$.

Dans ces conditions, tout élément x de E appartient à un sous-ensemble E_i unique.

1.4.3 Fonction indicatrice d'une partie

Définition 1.4.8

Soit A une partie d'un ensemble E . On appelle *fonction indicatrice* (ou *fonction caractéristique*) de A ,

et on note $\mathbb{1}_A$, la fonction définie sur E par :

$$\begin{cases} \mathbb{1}_A(x) = 1 & \text{si } x \in A \\ \mathbb{1}_A(x) = 0 & \text{si } x \notin A \end{cases}$$

- Opérations sur les applications numériques :

Soit f, g deux applications définies sur un ensemble E , à valeurs réelles.

Soit α un nombre réel.

On définit les applications $f + g$, fg et αf de la manière suivante :

- pour tout x de E , $(f + g)(x) = f(x) + g(x)$
- pour tout x de E , $(fg)(x) = f(x)g(x)$
- pour tout x de E , $(\alpha f)(x) = \alpha f(x)$

Par abus de langage, on note encore α l'application constante prenant la valeur α .

En particulier $\mathbb{1}_E = 1$ et $\mathbb{1}_\emptyset = 0$.

- Propriétés :

L'application $\mathbb{1}$, qui envoie A sur $\mathbb{1}_A$, est une bijection de $\mathcal{P}(E)$ sur $\mathcal{F}(E, \{0, 1\})$.

Toute application f de E vers $\{0, 1\}$ est en effet la fonction indicatrice d'une unique partie A de E , définie par : $A = \{x \in E, f(x) = 1\}$.

En particulier, deux parties A et B de E sont égales si et seulement si leurs fonctions indicatrices $\mathbb{1}_A$ et $\mathbb{1}_B$ sont identiques : c'est parfois un bon moyen de démontrer l'égalité entre deux parties d'un ensemble.

Proposition 1.4.1

Soit A et B deux parties de E . On a les égalités suivantes :

$$\mathbb{1}_{\overline{A}} = 1 - \mathbb{1}_A \qquad \mathbb{1}_{A \cap B} = \mathbb{1}_A \mathbb{1}_B$$

$$\mathbb{1}_{A \cup B} = \mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_A \mathbb{1}_B \qquad \mathbb{1}_{A \setminus B} = \mathbb{1}_A(1 - \mathbb{1}_B)$$

Les parties A et B sont disjointes si et seulement si $\mathbb{1}_{A \cup B} = \mathbb{1}_A + \mathbb{1}_B$.

Enfin, on a $\mathbb{1}_{A \Delta B} = \mathbb{1}_A + \mathbb{1}_B - 2\mathbb{1}_A \mathbb{1}_B = (\mathbb{1}_A - \mathbb{1}_B)^2 = |\mathbb{1}_A - \mathbb{1}_B|$

1.4.4 Restriction et prolongement

Soit f une application d'un ensemble E vers un ensemble F .

Il y a plusieurs moyens de créer de nouvelles applications en ne modifiant que l'ensemble de départ ou l'ensemble d'arrivée de f .

– Extension de l'ensemble d'arrivée :

Pour tout ensemble F' contenant F , on peut encore dire que f est une application de E vers F' (ce n'est pas vrai si F' est une partie de F , sauf si F' contient l'image de f , c'est-à-dire tous les éléments de F qui ont au moins un antécédent).

– Restriction d'une application :

Soit E' une partie de E .

On définit une application g , de E' vers F , en posant : $\forall x \in E', g(x) = f(x)$.

On dit que g est *la restriction* de f à E' .

– Prolongements d'une application :

Soit E'' un ensemble contenant E .

On dit qu'une application h de E'' vers F est *un prolongement* de f si f est la restriction de h à E , c'est-à-dire si : $\forall x \in E, h(x) = f(x)$.

On notera qu'on parle de *la* restriction et d'*un* prolongement.

Par exemple, si f est l'identité de \mathbb{R}^+ dans lui-même, elle possède une infinité de prolongements à \mathbb{R} tout entier, parmi lesquels :

– L'application identité de \mathbb{R} .

– L'application « valeur absolue », de \mathbb{R} dans lui-même.

– L'application h définie par : $h(x) = \frac{1}{2}(x + |x|)$, et qui est identiquement nulle sur \mathbb{R}^- .

1.4.5 Image directe d'une partie par une application

Soit E, F deux ensembles, et f une application de E vers F .

Définition 1.4.9

Soit $A \subset E$. On appelle *image* de A par f le sous-ensemble $f(A) = \{f(a), a \in A\}$ de F .

$f(A)$ est donc l'ensemble des images par f des éléments de A .

On peut écrire : $y \in f(A) \Leftrightarrow \exists x \in A, f(x) = y$.

On définit ainsi une nouvelle application de $\mathcal{P}(E)$ vers $\mathcal{P}(F)$: à toute partie de E on associe en effet une partie de F .

Cette application est encore notée f . On ne confondra pas ces deux significations de f !

Cas particuliers et exemples :

- L'image de l'ensemble vide par f est l'ensemble vide : $f(\emptyset) = \emptyset$.
Pour tout a de E , $f(\{a\}) = \{f(a)\}$ (on a là les deux acceptations de f).
- $f(E)$ est l'ensemble de toutes les images par f . On dit que c'est l'ensemble image de f .
Si on reprend l'exemple vu en 1.C.1 $f(\{a, b, c\}) = \{t, v\}$ et $f(\{d, e\}) = \{w\}$.

Proposition 1.4.2

Soit f une application d'un ensemble E vers un ensemble F .

Soit A et B deux parties quelconques de E . On a

- ◇ $A \subset B \Rightarrow f(A) \subset f(B)$
- ◇ $f(A \cup B) = f(A) \cup f(B)$
- ◇ $f(A \cap B) \subset f(A) \cap f(B)$

En revanche, il n'y a pas de règle générale permettant de comparer $f(\overline{A})$ et $\overline{f(A)}$.

1.4.6 Image réciproque d'une partie par une application

Définition 1.4.10

Soit f une application d'un ensemble E vers un ensemble F .

Soit B une partie de F . L'image réciproque de B par f , notée $f^{-1}(B)$, est l'ensemble des éléments de E dont l'image est dans B : Autrement dit : $f^{-1}(B) = \{x \in E, f(x) \in B\}$.

C'est la partie de E formée par les antécédents des éléments de B .

On peut donc écrire : $x \in f^{-1}(B) \Leftrightarrow f(x) \in B$.

Remarques :

- On définit ainsi une application f^{-1} de $\mathcal{P}(F)$ vers $\mathcal{P}(E)$: à toute partie de F on associe en effet une partie de E . Mais cette application ne doit pas être confondue (quand f est bijective) avec la bijection réciproque f^{-1} (voir plus loin).
- Soit f une application d'un ensemble E vers un ensemble F . Alors on a :
 $f^{-1}(\emptyset) = \emptyset$; $f^{-1}(F) = E$; $\forall b \in F, f^{-1}(\{b\}) = \{x \in E, f(x) = b\}$.
- Si on reprend l'exemple donné en 1.3.1,
 $f^{-1}(\{u\}) = \emptyset$, $f^{-1}(\{v\}) = \{c\}$, $f^{-1}(\{w\}) = \{d, e\}$ et $f^{-1}(\{t, w\}) = \{a, b, d, e\}$.

Proposition 1.4.3

Soit f une application d'un ensemble E vers un ensemble F .

Pour deux parties quelconques A et B de F , on a les égalités :

$$f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B), \quad f^{-1}(\overline{A}) = \overline{f^{-1}(A)}, \quad f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B)$$

1.4.7 Composition d'applications

Définition 1.4.11

Soit E, F, G trois ensembles. Soit f dans $\mathcal{F}(E, F)$ et g dans $\mathcal{F}(F, G)$.

La composée de f par g , notée $g \circ f$, est l'application de E vers G , définie par :

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

Proposition 1.4.4

Soit E, F, G, H quatre ensembles.

Soit f dans $\mathcal{F}(E, F)$, g dans $\mathcal{F}(F, G)$ et h dans $\mathcal{F}(G, H)$.

Alors on a l'égalité $h \circ (g \circ f) = (h \circ g) \circ f$.

Remarques et propriétés :

– On exprime cette propriété (qui ne devra jamais être redémontrée) en disant que *la loi de composition des applications est associative*.

De cette propriété il découle qu'une composition répétée, comme $f_n \circ f_{n-1} \circ \dots \circ f_2 \circ f_1$, ne nécessite pas de parenthèses.

– Soit f dans $\mathcal{F}(E, F)$. Alors $\text{Id}_F \circ f = f$ et $f \circ \text{Id}_E = f$.

– Si f appartient à $\mathcal{F}(E)$, on pose par convention $f^0 = \text{Id}_E$.

On définit alors : $\forall n \in \mathbb{N}^*, f^n = f^{n-1} \circ f = f \circ f \circ \dots \circ f \circ f$ (n fois).

– Soit f dans $\mathcal{F}(E, F)$ et g dans $\mathcal{F}(F, E)$.

On peut donc former à la fois $g \circ f$ (de E dans E) et $f \circ g$ (de F dans F).

Ces deux applications sont en général distinctes (notamment si $E \neq F$!).

– Si f et g appartiennent à $\mathcal{F}(E)$, on dit que f et g *commutent* si $g \circ f = f \circ g$.

On note que Id_E commute avec toute application de E dans E .

De même, les « puissances » f^n d'une application f commutent entre elles (on vérifie en effet par une récurrence évidente sur l'entier naturel n qu'on a toujours l'égalité : $f^n \circ f^p = f^p \circ f^n = f^{n+p}$)

1.5 Injections, surjections, bijections

1.5.1 Applications injectives

Définition 1.5.1

Soit f une application de E dans F . On dit que f est une application *injective* (ou encore une *injection*) si tout élément y de F possède *au plus un antécédent* par f .

Remarques :

– Une définition équivalente de l'injectivité de f est : $\forall (x, x') \in E^2, x \neq x' \Rightarrow f(x) \neq f(x')$.

Autrement dit, une application est injective si elle « conserve les différences ».

– Une autre définition équivalente (très utile) est : $\forall (x, x') \in E^2, f(x) = f(x') \Rightarrow x = x'$

Proposition 1.5.1

Soit f dans $\mathcal{F}(E, F)$ et g dans $\mathcal{F}(F, G)$.

- Si f et g sont injectives, alors $g \circ f$ est injective.
- Si $g \circ f$ est injective, alors f est injective.

1.5.2 Applications surjectives**Définition 1.5.2**

Soit f une application de E dans F .

On dit que f est une application *surjective* (ou encore une *surjection*) si tout élément y de F possède au moins un antécédent par f , autrement dit si : $\forall y \in F, \exists x \in E, f(x) = y$.

Une définition équivalente de la surjectivité de $f : E \rightarrow F$ est : $f(E) = F$.

On dit souvent que f est une surjection de E sur F (plutôt que dans F).

Proposition 1.5.2

Soit f dans $\mathcal{F}(E, F)$ et g dans $\mathcal{F}(F, G)$.

- Si f et g sont surjectives, alors $g \circ f$ est surjective.
- Si $g \circ f$ est surjective, alors g est surjective.

1.5.3 Applications bijectives**Définition 1.5.3**

Soit f une application de E dans F . On dit que f est une application *bijective* (ou encore une *bijection*) si f est à la fois injective et surjective, c'est-à-dire si tout élément y de l'ensemble F possède un antécédent x et un seul dans E : $\forall y \in F, \exists ! x \in E, f(x) = y$.

Proposition 1.5.3 (bijection réciproque)

Soit f une bijection de l'ensemble E sur l'ensemble F .

On définit une application de F vers E en associant à tout y de F son seul antécédent x .

Cette application, notée f^{-1} , vérifie donc : $\forall x \in E, \forall y \in F, x = f^{-1}(y) \Leftrightarrow y = f(x)$.

On a alors les égalités : $f^{-1} \circ f = \text{Id}_E$, et $f \circ f^{-1} = \text{Id}_F$.

L'application f^{-1} est également bijective : on l'appelle la bijection réciproque de f .

On a enfin l'égalité $(f^{-1})^{-1} = f$: l'application f est la bijection réciproque de f^{-1} .

Rappel : on ne confondra pas l'application f^{-1} (bijection réciproque de f), de F vers E , avec l'application $\overset{-1}{f}$ de $\mathcal{P}(F)$ vers $\mathcal{P}(E)$, qui existe même quand f n'est pas bijective.

Proposition 1.5.4

Soit f dans $\mathcal{F}(E, F)$. Pour montrer que l'application f est bijective, il suffit de trouver g dans $\mathcal{F}(F, E)$ telle que : $g \circ f = \text{Id}_E$, et $f \circ g = \text{Id}_F$. On a alors $g = f^{-1}$.

Proposition 1.5.5

Soit f dans $\mathcal{F}(E, F)$ et g dans $\mathcal{F}(F, G)$.

Si f et g sont bijectives, alors $g \circ f$ est bijective, et $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

Définition 1.5.4 (applications involutives)

Soit f dans $\mathcal{F}(E)$. On dit que f est *involutive* si $f \circ f = \text{Id}_E$.

Cela équivaut à dire que f est bijective et que $f^{-1} = f$.

Exemples

- L'application Id_E est involutive, ou encore les symétries dans un espace affine.
- L'application $x \mapsto \frac{1}{x}$ est une involution de \mathbb{R}^* .
- L'application $A \mapsto \overline{A}$ est une involution de $\mathcal{P}(E)$.
- L'application $z \mapsto \bar{z}$ est une involution de \mathbb{C} .

1.6 Relations binaires

1.6.1 Généralités

Soit E et F deux ensembles.

Définition 1.6.1

On appelle *relation* \mathcal{R} de E vers F la donnée d'une partie R du produit cartésien $E \times F$.

La partie R est appelée le *graphe* de la relation \mathcal{R} .

On dit qu'un élément x de E est *en relation* avec un élément y de F , pour la relation \mathcal{R} , si le couple (x, y) appartient au graphe R .

On exprime cette situation en écrivant $x \mathcal{R} y$.

Si $E = F$ (cas fréquent) on dit que \mathcal{R} est une relation *sur* E .

Exemples :

- La relation d'inclusion dans l'ensemble des parties de E : $A \mathcal{R} B \Leftrightarrow A \subset B$.
- La relation de divisibilité sur les entiers relatifs : $m \mathcal{R} n \Leftrightarrow m$ divise n .
- Dans \mathbb{Z} , et si a non nul, on définit la relation de *congruence modulo* a :
 $m \mathcal{R} n \Leftrightarrow m - n$ est divisible par a . On note le plus souvent $m \equiv n \pmod{a}$.
- Sur tout ensemble E , on peut définir la *relation égalité* : $x \mathcal{R} y \Leftrightarrow x = y$.
Le graphe de cette relation est la diagonale $\Delta(E) = \{(x, x), x \in E\}$.
- Soit E et F deux ensembles quelconques.
On définit la *relation universelle* de E vers F par : $\forall (x, y) \in E \times F, x \mathcal{R} y$.
Le graphe de cette relation est l'ensemble $E \times F$ tout entier.
- Soit E un ensemble *fini* et \mathcal{R} une relation sur E .
On peut représenter \mathcal{R} en donnant la liste de tous les couples (x, y) de E^2 tels que $x \mathcal{R} y$.
Voici par exemple la description d'une relation sur $E = \{a, b, c, d, e, f\}$.

	a	b	c	d	e	f
a	*					
b	*					*
c			*			
d		*		*	*	*
e	*				*	
f				*		

ce qui équivaut à

$$\left\{ \begin{array}{l} a \mathcal{R} a \\ b \mathcal{R} a, \quad b \mathcal{R} f \\ c \mathcal{R} c \\ d \mathcal{R} b, \quad d \mathcal{R} d, \quad d \mathcal{R} e, \quad d \mathcal{R} f \\ e \mathcal{R} a, \quad e \mathcal{R} e \\ f \mathcal{R} d \end{array} \right.$$

– Restriction d'une relation :

Si \mathcal{R} est une relation binaire sur E , et si F est une partie de E , alors on peut définir la restriction \mathcal{S} de \mathcal{R} à l'ensemble F .

Le graphe de \mathcal{S} est l'intersection de $F \times F$ et du graphe de \mathcal{R} .

Par exemple, Si P est l'ensemble des entiers naturels premiers, la restriction à P de la relation de divisibilité n'est autre que la relation d'égalité sur P .

1.6.2 Propriétés éventuelles des relations binaires

Définition 1.6.2

Soit \mathcal{R} une relation binaire sur l'ensemble E .

- \mathcal{R} est dite *réflexive* si : $\forall x \in E, x \mathcal{R} x$.
- \mathcal{R} est dite *symétrique* si : $\forall (x, y) \in E^2, x \mathcal{R} y \Rightarrow y \mathcal{R} x$.
- \mathcal{R} est dite *transitive* si : $\forall (x, y, z) \in E^3, (x \mathcal{R} y \text{ et } y \mathcal{R} z) \Rightarrow x \mathcal{R} z$.
- \mathcal{R} est dite *antisymétrique* si : $\forall (x, y) \in E^2, (x \mathcal{R} y \text{ et } y \mathcal{R} x) \Rightarrow x = y$.

Remarques :

- L'antisymétrie s'écrit aussi : $\forall (x, y) \in E^2, (x \mathcal{R} y \text{ et } x \neq y) \Rightarrow \text{non}(y \mathcal{R} x)$
- Si \mathcal{R} est symétrique, on dira de deux éléments x, y de E qu'ils sont ou qu'ils ne sont pas en relation (x et y ayant alors des rôles identiques).
- L'antisymétrie n'est pas le contraire de la symétrie : la relation « égalité », par exemple, possède les deux propriétés, alors que la relation donnée en exemple sur $\{a, b, c, d, e, f\}$ n'est ni symétrique (e est en relation avec a mais a n'est pas en relation avec e) ni antisymétrique (les éléments d et f sont en relation l'un avec l'autre, bien qu'ils soient distincts).
- A titre d'exercice, on pourra vérifier que si une relation est à la fois symétrique, antisymétrique, et réflexive, alors c'est la relation d'égalité.

1.6.3 Relations d'ordre

Définition 1.6.3

On dit qu'une relation \mathcal{R} sur un ensemble E est une *relation d'ordre* si \mathcal{R} est à la fois réflexive, antisymétrique et transitive. On note souvent \leq une telle relation.

On dit alors que (E, \leq) est un *ensemble ordonné*.

Une relation d'ordre sur E est donc caractérisée par :

- $\forall x \in E, x \leq x$
- $\forall x, y, z \in E, (x \leq y \text{ et } y \leq z) \Rightarrow x \leq z$
- $\forall x, y \in E, (x \leq y) \text{ et } (y \leq x) \Rightarrow x = y$

Définition 1.6.4 (ordre total ou ordre partiel)

Soit \mathcal{R} une relation d'ordre sur E .

- Deux éléments x et y de E sont dits *comparables* (pour \leq) si $x \leq y$ ou si $y \leq x$.
- Si deux éléments quelconques x et y sont toujours comparables, on dit que \leq est une relation d'*ordre total* : l'ensemble E est dit *totalement ordonné* par \leq .
- Sinon (c'est-à-dire s'il existe au moins deux éléments non comparables x et y) on dit que \leq est une relation d'*ordre partiel* (l'ensemble E est dit *partiellement ordonné* par \leq).

Exemple et remarques

- Sur les ensembles $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$, on dispose d'une relation d'ordre total, notée \leq .
Si m et n sont deux entiers relatifs, avec $m \leq n$, on note souvent $\llbracket m, n \rrbracket = \{p \in \mathbb{Z}, m \leq p \leq n\}$.
- *Relation d'ordre inverse*
Si \leq est une relation d'ordre sur E , on définit encore une relation d'ordre sur E , notée \geq , en posant : $\forall (x, y) \in E^2, x \geq y \Leftrightarrow y \leq x$.
- Si on pose : $x < y \Leftrightarrow (x \leq y) \text{ et } x \neq y$, on définit une nouvelle relation (appelée souvent ordre strict) qui en fait n'est pas une relation d'ordre (elle n'est pas réflexive).
- Soit (E, \leq) un ensemble ordonné.
La négation de la proposition $x \leq y$ est : « (x et y ne sont pas comparables) ou ($y < x$) ».
Bien sûr, si E est totalement ordonné, la négation de $x \leq y$ est $y < x$.
- On définit une relation d'ordre sur $\mathcal{P}(E)$ en posant : $A \leq B \Leftrightarrow A \subset B$.
Il s'agit d'un ordre partiel, sauf si E est vide ou se réduit à un élément.
Par exemple, si a et b sont distincts dans E , $\{a\}$ et $\{b\}$ ne sont pas comparables.
- Sur \mathbb{R}^2 , on peut poser : $(x, y) \leq (x', y') \Leftrightarrow (x \leq x') \text{ et } (y \leq y')$.
C'est un ordre partiel : les couples $(1, 2)$ et $(4, 0)$, par exemple, ne sont pas comparables.

Relation de divisibilité dans \mathbb{N} **Définition 1.6.5**

On dit que n *divise* m (ou que m est un *multiple* de n) si : $\exists q \in \mathbb{N}, m = nq$.

On note alors $n \mid m$. On définit ainsi une relation d'ordre partiel sur \mathbb{N} .

Pour cette relation, 1 est le minimum de \mathbb{N} .

Remarque : la divisibilité ne définit pas une relation d'ordre sur \mathbb{Z} ; en effet si m est un entier relatif non nul, les entiers m et $-m$ se divisent mutuellement (le quotient vaut -1) bien qu'ils soient distincts : la relation n'est donc pas antisymétrique.

Ordre lexicographique sur \mathbb{R}^2

Sur l'ensemble \mathbb{R}^2 , on peut poser : $(x, y) \leq (x', y')$ si $\begin{cases} x < x' \text{ ou} \\ (x = x') \text{ et } (y \leq y') \end{cases}$

on définit ainsi une relation d'ordre total sur \mathbb{R}^2 .

Cette relation (dont le modèle se généralise facilement à \mathbb{R}^n) est appelée *ordre lexicographique*, en référence à la relation d'ordre total qui permet de classer les mots du dictionnaire.

1.6.4 Relations d'équivalence, classes d'équivalence

Définition 1.6.6

On dit qu'une relation \mathcal{R} sur un ensemble E est une *relation d'équivalence* si \mathcal{R} est à la fois réflexive, symétrique et transitive.

Pour tout x de E , l'ensemble des éléments en relation avec x est appelé la *classe d'équivalence* de x , et notée $\mathcal{C}(x)$ (ou \dot{x} , ou \bar{x}) : $\mathcal{C}(x) = \{y \in E, x \mathcal{R} y\}$.

Conséquences de la définition :

- La relation \mathcal{R} étant réflexive, tout élément x de E appartient à sa propre classe d'équivalence. En particulier, aucune classe d'équivalence n'est vide.
- Soit x et y dans E : si $x \mathcal{R} y$, alors $\mathcal{C}(x) = \mathcal{C}(y)$. Sinon $\mathcal{C}(x) \cap \mathcal{C}(y) = \emptyset$.
Deux classes d'équivalence sont donc ou bien identiques ou bien disjointes.
- Une classe d'équivalence \mathcal{C} est entièrement déterminée par la donnée de l'un quelconque de ses éléments x (on dira que x est un *représentant* de \mathcal{C}).

Proposition 1.6.1 (partitions et relations d'équivalence)

Soit E un ensemble.

- Soit \mathcal{R} une relation d'équivalence sur E .
Soit Ω un sous-ensemble de E obtenu en choisissant dans chaque classe d'équivalence \mathcal{C} un représentant x unique : la famille $(\mathcal{C}(x))_{x \in \Omega}$ est une partition de l'ensemble E . Autrement dit, les différentes classes d'équivalence pour la relation \mathcal{R} définissent une partition de E .
- Réciproquement, soit $(E_i)_{i \in I}$ une partition de E .
Soit \mathcal{R} la relation définie sur E par : $\forall (x, y) \in E^2, x \mathcal{R} y \Leftrightarrow \exists i \in I, \{x, y\} \subset E_i$
Alors \mathcal{R} est une relation d'équivalence sur E , et les sous-ensembles E_i sont les classes d'équivalence de E pour \mathcal{R} .

Exemples de relations d'équivalence

- Soit $f : E \rightarrow F$ une application.
On définit une relation d'équivalence sur E en posant : $x \mathcal{R} y \Leftrightarrow f(x) = f(y)$.
La classe d'équivalence de x est l'image réciproque par f du singleton $\{f(x)\}$.
- Par exemple, si Ω est un point fixé du plan, on définit une relation sur ce plan en posant $M \mathcal{R} N \Leftrightarrow d(\Omega, M) = d(\Omega, N)$: l'ensemble des classes d'équivalence est l'ensemble des cercles de centre Ω .
- La relation définie sur l'ensemble \mathcal{D} des droites du plan affine par : $D \mathcal{R} \Delta \Leftrightarrow D // \Delta$.
Les classes d'équivalence correspondent aux différentes "directions" de droites.

1.6.5 Congruence modulo un réel strictement positif

Définition 1.6.7 (congruence modulo un réel $a > 0$)

Soit a un réel strictement positif. Les réels x et y sont dits *congrus* modulo a , et on note $x \equiv y (a)$, s'il existe un entier relatif q tel que $x - y = qa$.

Remarques

- La relation de congruence modulo a est une relation d'équivalence sur \mathbb{R} .
- Chaque classe a un représentant unique dans $[0, a[$ ou encore dans $\left[-\frac{a}{2}, \frac{a}{2}\right[$.
- Pour tout réel λ , on a : $x \equiv y (a) \Leftrightarrow x + \lambda \equiv y + \lambda (a)$
- Pour tout réel non nul λ , on a : $x \equiv y (a) \Leftrightarrow \lambda x \equiv \lambda y (\lambda a)$
- On a les équivalences classiques suivantes :

$$\cos y = \cos x \Leftrightarrow \begin{cases} y = x (2\pi) \\ \text{ou} \\ y = -x (2\pi) \end{cases} \quad \sin y = \sin x \Leftrightarrow \begin{cases} y = x (2\pi) \\ \text{ou} \\ y = \pi - x (2\pi) \end{cases}$$

Par exemple : $\cos x = 1 \Leftrightarrow x \equiv 0 (2\pi)$; $\sin(2x) = 0 \Leftrightarrow x \equiv 0 (\pi/2)$

1.6.6 Division euclidienne et congruence modulo un entier**Définition 1.6.8**

Soit (a, b) dans $\mathbb{Z} \times \mathbb{N}^*$.

Il existe un unique couple (q, r) de $\mathbb{Z} \times \mathbb{N}$ tel que : $(a = bq + r)$ et $(r \leq b - 1)$

Le passage du couple (a, b) au couple (q, r) s'appelle *division euclidienne* de a par b .

Dans cette division, a est le *dividende*, b le *diviseur*, q le *quotient*, et r le *reste*.

Remarque (divisibilité et division euclidienne) :

Soit a et b deux entiers relatifs. La proposition « b divise a » équivaut à dire que « $a = b = 0$ ou le reste dans la division de a par $|b|$ est nul ».

Définition 1.6.9 (congruence modulo un entier n de \mathbb{N}^*)

Soit n un entier naturel strictement positif. Les entiers relatifs a et b sont dits *congrus* modulo n , et on note $a \equiv b (n)$, s'il existe un entier relatif q tel que $a - b = qn$.

Remarques

- La relation de congruence modulo l'entier n est une relation d'équivalence sur \mathbb{Z} .
- Dire que les deux entiers a et b sont congrus modulo n , c'est dire que a et b ont le même reste dans la division euclidienne par n .
- Comme il y a n restes possibles dans la division par n , il y a exactement n classes d'équivalence. Les entiers $0, 1, 2, \dots, n - 1$ constituent une famille de représentants de chaque classe. L'ensemble des différentes classes d'équivalence s'écrit donc : $\{\overline{0}, \overline{1}, \overline{2}, \dots, \overline{n-1}\}$.
- On a les propriétés suivantes (a, b, c sont dans \mathbb{Z} , m, n sont dans \mathbb{N}^*)

$$\begin{cases} a \equiv b (n) \Leftrightarrow a + c \equiv b + c (n) \\ a \equiv b (n) \Rightarrow ma \equiv mb (mn) \Rightarrow ma \equiv mb (n) \\ a \equiv b (n) \Rightarrow a^m \equiv b^m (n) \end{cases}$$

Chapitre 2

Calculs algébriques

Sommaire

2.1	Les ensembles de nombres \mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}	27
2.1.1	L'existence admise des ensembles de nombres	27
2.1.2	Propriétés relatives à l'addition dans \mathbb{C}	27
2.1.3	Propriétés relatives au produit dans \mathbb{C}	28
2.2	Sommes et produits	29
2.2.1	Sommes et produits finis	29
2.2.2	Changements d'indice	30
2.2.3	Sommes et produits « télescopiques »	30
2.2.4	Quelques résultats classiques	31
2.3	Factorielles et coefficients binomiaux	32
2.3.1	Factorielle d'un entier	32
2.3.2	Coefficients binomiaux	32
2.3.3	Relations entre coefficients binomiaux	33
2.4	Sommes doubles, interversions	34
2.4.1	Somme sur un domaine rectangulaire	34
2.4.2	Somme sur un domaine triangulaire	35
2.4.3	Sommation par partition	35
2.4.4	Produits de sommes	36
2.5	Systèmes linéaires	37
2.5.1	Système linéaire de n équations à p inconnues	37
2.5.2	Exemples et définitions complémentaires	37
2.5.3	Interprétation géométrique (deux ou trois variables)	38
2.5.4	Système homogène associé	39
2.5.5	Structure de la solution générale d'un système linéaire quelconque	40
2.5.6	Systèmes de Cramer triangulaires	41
2.5.7	Opérations élémentaires sur les lignes d'un système	41
2.5.8	Méthode « du pivot de Gauss »	42
2.5.9	Trois exemples	44

2.1 Les ensembles de nombres \mathbb{N} , \mathbb{Z} , \mathbb{Q} , \mathbb{R} , \mathbb{C}

2.1.1 L'existence admise des ensembles de nombres

Conformément au programme, on admet l'existence et les principales propriétés des ensembles de nombres suivants :

- l'ensemble \mathbb{N} des entiers naturels.
- l'ensemble \mathbb{Z} des entiers relatifs.
- l'ensemble \mathbb{Q} des nombres rationnels.
- l'ensemble \mathbb{R} des nombres réels.
- l'ensemble \mathbb{C} des nombres complexes.

On a bien sûr les inclusions strictes : $\mathbb{N} \subsetneq \mathbb{Z} \subsetneq \mathbb{Q} \subsetneq \mathbb{R} \subsetneq \mathbb{C}$.

On sera également amené à considérer les ensembles suivants (dont les noms ne sont pas fixés par l'usage, mais dont la signification est bien connue) : l'ensemble \mathbb{P} des nombres premiers, l'ensemble \mathbb{D} des nombres décimaux, l'ensemble $\mathbb{R} \setminus \mathbb{Q}$ des nombres irrationnels.

2.1.2 Propriétés relatives à l'addition dans \mathbb{C}

Proposition 2.1.1 (propriétés de l'addition dans \mathbb{C})

L'ensemble \mathbb{C} est muni d'une opération (ou loi de composition), notée $+$, avec les propriétés suivantes :

- (1a) La loi $+$ est commutative : $\forall (x, y) \in \mathbb{C}^2, x + y = y + x$
- (1b) La loi $+$ est associative : $\forall (x, y, z) \in \mathbb{C}^3, x + (y + z) = (x + y) + z$
- (1c) Pour la loi $+$, l'entier naturel 0 est élément neutre : $\forall x \in \mathbb{C}, x + 0 = x$
- (1d) Pour la loi $+$, tout x de \mathbb{C} possède un opposé (noté $-x$), donc tel que $x + (-x) = 0$;
La notation $x - y$ doit être comprise comme une contraction de $x + (-y)$.

Remarques diverses

- On note $\mathbb{N}^*, \mathbb{Z}^*, \mathbb{Q}^*, \mathbb{R}^*, \mathbb{C}^*$ les ensembles $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ privés de l'entier 0.
- Chacun des ensembles $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ est *stable* pour la loi $+$ (ce qui signifie que la somme de deux éléments de \mathbb{Q} , par exemple, est encore un élément de \mathbb{Q}). On peut donc considérer (1a), (1b), et (1c) comme des propriétés de la loi $+$ *restreinte* à \mathbb{N} , à \mathbb{Z} , à \mathbb{Q} , ou à \mathbb{R} .
- De même, chacun des ensembles $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ est *stable par passage à l'opposé*, ce qui signifie que l'opposé d'un rationnel x , par exemple, est encore un rationnel. On peut donc considérer (1d) comme une propriété de la loi $+$ *restreinte* à \mathbb{Z} , à \mathbb{Q} , ou à \mathbb{R} . Pour ces trois ensembles, on peut donc également parler de *stabilité par différence* (si x et y sont dans \mathbb{Q} , par exemple, $x - y$ est dans \mathbb{Q}). Ici l'ensemble \mathbb{N} fait exception, car seul 0 possède un opposé dans \mathbb{N} (lui-même!). Enfin, on a bien sûr les identités : $-(-x) = x$, $-(x + y) = -x - y$, et $-(x - y) = -x + y$.
- On pourrait considérer l'opération *soustraction* sur \mathbb{C} (définie par $(x, y) \mapsto x - y$) mais cette loi ne présente que peu d'intérêt : elle n'est ni commutative, ni associative, et il n'y a pas d'élément neutre.
- Une conséquence de (1b), (1c) et (1d) est que tout élément de \mathbb{C} est *simplifiable pour l'addition*. En d'autres termes : $\forall (x, y, z) \in \mathbb{C}^3, x + y = x + z \Rightarrow y = z$ (ajouter $-x$ de part et d'autre). Pour des raisons analogues : on a l'équivalence $x + y = z \Leftrightarrow y = z - x$.

2.1.3 Propriétés relatives au produit dans \mathbb{C}

Proposition 2.1.2 (propriétés du produit dans \mathbb{C})

L'ensemble \mathbb{C} est muni d'une multiplication (notée \times , ou par juxtaposition : le produit de x par y est alors noté xy), avec les propriétés suivantes :

- (2a) La loi \times est commutative : $\forall (x, y) \in \mathbb{C}^2, xy = yx$
- (2b) La loi \times est associative : $\forall (x, y, z) \in \mathbb{C}^3, x(yz) = (xy)z$
- (2c) Pour la loi \times , l'entier naturel 1 est élément neutre : $\forall x \in \mathbb{C}, 1x = x$
- (2d) La loi \times est distributive par rapport à l'addition : $\forall (x, y, z) \in \mathbb{C}^3, x(y + z) = xy + xz$
- (2e) Pour la loi \times , tout élément x **non nul** de \mathbb{C} possède un inverse (noté x^{-1}), donc tel que $xx^{-1} = 1$

Remarques diverses

- On note $\frac{1}{x}$ plutôt que x^{-1} . La notation $\frac{y}{x}$ doit être comprise comme une contraction de yx^{-1} .
Cette dernière notation est rendue possible car le produit est une opération commutative.
Avec ces notations, l'ensemble \mathbb{Q} des nombres rationnels s'écrit $\mathbb{Q} = \left\{ \frac{a}{b}, a \in \mathbb{Z}, b \in \mathbb{Z}^* \right\}$.
- Chacun des ensembles $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ est *stable* pour la loi \times (si x et y sont dans \mathbb{Q} , par exemple, xy est encore un élément de \mathbb{Q}). On peut donc considérer (2a), (2b), (2c) et (2d) comme des propriétés de la loi \times *restreinte* à \mathbb{N} , à \mathbb{Z} , à \mathbb{Q} , ou à \mathbb{R} .
- Les identités $x0 = 0$ et $x(-y) = (-x)y = -(xy)$ semblent évidentes mais elles se démontrent.
Il en est de même pour l'équivalence $xy = 0 \Leftrightarrow (x = 0 \text{ ou } y = 0)$.
- Pour la propriété (2e), les ensembles \mathbb{N} et \mathbb{Z} font exception. En effet seul 1 possède un inverse dans \mathbb{N} (lui-même) et seuls 1 et -1 possèdent un inverse dans \mathbb{Z} (eux-mêmes!).
Dans \mathbb{N} , on a $xy = 1 \Leftrightarrow (x = 1 \text{ et } y = 1)$ et dans \mathbb{Z} on a : $xy = 1 \Leftrightarrow ((x = y = 1) \text{ ou } (x = y = -1))$.
- Bien sûr, l'inverse d'un rationnel (resp. d'un réel) non nul est encore un rationnel (resp. un réel). On parle dans ce cas de *stabilité par passage à l'inverse*.
- Une conséquence de (2b), (2c) et (2d) est que tout élément de \mathbb{C}^* est *simplifiable pour le produit*.
En d'autres termes : $\forall (x, y, z) \in \mathbb{C}^* \times \mathbb{C}^2, xy = xz \Rightarrow y = z$ (multiplier par x^{-1} de part et d'autre).
Pour des raisons analogues : si $x \neq 0$, on a l'équivalence $xy = z \Leftrightarrow y = zx^{-1}$.
Dans ces questions, il est essentiel de s'assurer que x est non nul avant de simplifier par x .
- Si x et y sont non nuls, on a bien sûr les identités : $(x^{-1})^{-1} = x, (-x)^{-1} = -(x^{-1}), (xy)^{-1} = x^{-1}y^{-1}$.
- On pourrait considérer l'opération *quotient* sur \mathbb{C} (définie par $(x, y) \mapsto x/y$) mais cette loi n'a que des défauts : elle n'est pas partout définie (y doit être non nul), elle n'est ni commutative ni associative, et il n'y a pas d'élément neutre.

2.2 Sommes et produits

2.2.1 Sommes et produits finis

Définition 2.2.1

Soit I un ensemble fini, et $(x_i)_{i \in I}$ une famille de nombres complexes.

On note $\sum_{i \in I} x_i$ la somme des x_i , et on note $\prod_{i \in I} x_i$ leur produit.

Dans le cas où l'ensemble I est vide, on convient que $\sum_{i \in I} x_i = 0$, et que $\prod_{i \in I} x_i = 1$.

Remarques

– Dans les notations précédentes, il faut bien comprendre que chaque indice i apparaît une fois et une seule, dans la somme ou dans le produit. La commutativité des opérations fait qu'il n'est pas nécessaire de préciser dans quel ordre sont effectués cette somme ou ce produit.

– Supposons que I soit la réunion $I = J \cup K$ de deux ensembles disjoints J et K .

Il est clair que $\sum_{i \in I} x_i = \sum_{i \in J} x_i + \sum_{i \in K} x_i$ et que $\prod_{i \in I} x_i = \left(\prod_{i \in J} x_i \right) \left(\prod_{i \in K} x_i \right)$.

Cette remarque est à rapprocher des conventions $\sum_{i \in \emptyset} x_i = 0$ et $\prod_{i \in \emptyset} x_i = 1$ (et elle aide à les comprendre).

– Un cas fréquent est celui où I est un intervalle d'entiers.

Supposons par exemple $I = \llbracket m, n \rrbracket$ (notation usuelle pour désigner $\{k \in \mathbb{N}, m \leq k \leq n\}$).

Dans ce cas, on notera $\sum_{i=m}^n x_i$ ou $\sum_{m \leq i \leq n} x_i$ pour la somme, et $\prod_{i=m}^n x_i$ ou $\prod_{m \leq i \leq n} x_i$ pour le produit.

Attention : l'utilisation de ces notations suggère qu'on effectue la somme (ou le produit) dans le sens des indices croissants. On considèrera donc que la somme est vide (donc vaut 0) et que le produit est vide (donc vaut 1) si $m > n$.

– La lettre utilisée pour décrire l'ensemble I est sans importance, dans la mesure où elle ne prête pas à confusion (on parle d'indice muet).

La même somme pourra donc être notée indifféremment $\sum_{i \in I} x_i$, $\sum_{j \in I} x_j$, ou $\sum_{k \in I} x_k$.

Par exemple, la somme $\sum_{j=1}^n x^j$ pourra être notée $\sum_{i=1}^n x^i$, mais certainement pas $\sum_{x=1}^n x^x$ ou $\sum_{x=1}^n x^n$.

Mise en facteur de termes constants

– Si λ ne dépend pas de l'indice i , alors $\sum_{i \in I} (\lambda x_i) = \lambda \sum_{i \in I} x_i$.

Bien sûr $\sum_{i \in I} \lambda = \lambda \text{Card}(I)$ (où $\text{Card}(I)$ désigne le cardinal, c'est-à-dire le nombre d'éléments de I)

Plus généralement, si λ et μ sont constants : $\sum_{i \in I} (\lambda x_i + \mu y_i) = \lambda \sum_{i \in I} x_i + \mu \sum_{i \in I} y_i$.

– Si λ ne dépend pas de i , et si $n = \text{Card}(I) \geq 1$, alors $\prod_{i \in I} (\lambda x_i) = \lambda^n \prod_{i \in I} x_i$ et notamment $\prod_{i \in I} \lambda = \lambda^n$.

On a les égalités si $\prod_{i \in I} (x_i y_i) = \left(\prod_{i \in I} x_i \right) \left(\prod_{i \in I} y_i \right)$ et plus généralement : $\prod_{i \in I} x_i^p y_i^q = \left(\prod_{i \in I} x_i \right)^p \left(\prod_{i \in I} y_i \right)^q$

– Attention surtout à ne pas faire l'erreur d'écrire $\sum_{i=1}^n x_i y_i = \sum_{i=1}^n x_i \sum_{i=1}^n y_i$

2.2.2 Changements d'indice

Soit I un ensemble fini, et $(x_i)_{i \in I}$ une famille de nombres complexes.

Soit J un ensemble fini, et φ une bijection de J sur I .

Ainsi à chaque indice i de I correspond un unique j de J tel que $i = \varphi(j)$.

Effectuer le changement d'indice $i = \varphi(j)$, c'est écrire (en posant $y_j = x_{\varphi(j)}$, pour tout j de J) :

– dans une somme : $S = \sum_{i \in I} x_i = \sum_{j \in J} x_{\varphi(j)} = \sum_{j \in J} y_j$ (que rien n'empêche alors d'écrire $S = \sum_{i \in J} y_i$).

– dans un produit : $P = \prod_{i \in I} x_i = \prod_{j \in J} x_{\varphi(j)} = \prod_{j \in J} y_j$ (que rien n'empêche alors d'écrire $P = \prod_{i \in J} y_i$).

Dans un tel calcul il faut distinguer le temps du changement d'indice $i = \varphi(j)$, pendant lequel i et j ne doivent pas être confondus, et le temps qui vient après (où l'indice muet j peut être renommé i).

La situation la plus fréquente est celle où I est un intervalle $\llbracket m, n \rrbracket$ d'entiers, et où le changement d'indice est une translation $i = j + p$ (le plus souvent $i = j + 1$ ou $i = j - 1$).

On écrira : $S = \sum_{i=m}^n x_i = \sum_{j=m+1}^{n+1} x_{j-1} = \sum_{k=m-1}^{n-1} x_{k+1}$ ou directement $S = \sum_{i=m}^n x_i = \sum_{i=m+1}^{n+1} x_{i-1} = \sum_{i=m-1}^{n-1} x_{i+1}$.

Autre situation classique, un changement d'indice suggérant que la somme ou le produit sont parcourus en sens contraire du sens initial :

Par exemple, avec « $k \leftarrow n - k$ » on écrira : $\sum_{k=0}^n x_k = \sum_{k=0}^n x_{n-k}$, ou $\prod_{k=0}^n x_k = \prod_{k=0}^n x_{n-k}$

2.2.3 Sommes et produits « télescopiques »

Pour comprendre cette notion, considérons l'exemple de la somme $u_n = \sum_{k=1}^n \frac{1}{k(k+1)}$.

Par exemple $u_6 = \frac{1}{2} + \frac{1}{6} + \frac{1}{12} + \frac{1}{20} + \frac{1}{30} + \frac{1}{42}$.

Pour calculer u_6 (ou plus généralement u_n), tout s'éclaire quand on réalise que $\frac{1}{k(k+1)} = \frac{1}{k} - \frac{1}{k+1}$.

Ainsi, $u_6 = \left(\frac{1}{1} - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \left(\frac{1}{4} - \frac{1}{5}\right) + \left(\frac{1}{5} - \frac{1}{6}\right) + \left(\frac{1}{6} - \frac{1}{7}\right) = 1 - \frac{1}{7} = \frac{6}{7}$.

Plus généralement, soit à calculer une somme $S = \sum_{i=m}^n x_i$.

Il peut être judicieux d'écrire $x_i = y_i - y_{i+1}$ (mais trouver y_i n'est pas forcément facile!).

Dans ce cas, on peut écrire toutes les étapes :

$$\begin{aligned} S &= \sum_{i=m}^n x_i = \sum_{i=m}^n (y_i - y_{i+1}) = \sum_{i=m}^n y_i - \sum_{i=m}^n y_{i+1} = \sum_{i=m}^n y_i - \sum_{i=m+1}^{n+1} y_i \\ &= \left(y_m + \sum_{i=m+1}^n y_i\right) - \left(\sum_{i=m+1}^n y_i + y_{n+1}\right) = y_m - y_{n+1} \end{aligned}$$

L'idéal est d'écrire directement : $\sum_{i=m}^n (y_i - y_{i+1}) = y_m - y_{n+1}$ (on parle de « somme télescopique »).

On est en présence d'un « produit télescopique » $P = \prod_{i=m}^n x_i$ quand x_i s'écrit $x_i = \frac{y_i}{y_{i+1}}$.

Dans ces conditions, on écrira directement (en justifiant) : $P = \prod_{i=m}^n \frac{y_i}{y_{i+1}} = \frac{y_m}{y_{n+1}}$.

2.2.4 Quelques résultats classiques

Somme d'une suite arithmétique :

Proposition 2.2.1 (somme des entiers de 1 à n)

Pour tout n de \mathbb{N} , on a $\sum_{k=1}^n k = 1 + 2 + \dots + n = \frac{n(n+1)}{2}$.

Plus généralement, si on cherche la somme S de n termes successifs a_k d'une progression arithmétique de raison r (c'est-à-dire $a_{k+1} = a_k + r$ pour tout k), et si les termes de début et de fin de cette somme sont d et f alors la formule à retenir est : $\sum a_k = n \frac{d+f}{2}$.

Un bon moyen de retrouver S est d'écrire (en notant a_k^+ et a_k^- les a_k « parcourus » de gauche à droite et de droite à gauche) : $2S = \sum a_k^+ + \sum a_k^- = \sum (a_k^+ + a_k^-) = \sum (d+f) = n(d+f)$.

Proposition 2.2.2 (somme des carrés, ou des cubes, des entiers de 1 à n)

Pour tout n de \mathbb{N} , on a $\sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6}$, et $\sum_{k=1}^n k^3 = \frac{n^2(n+1)^2}{4}$.

On notera (ça aide à la mémorisation) que la somme $\sum_{k=1}^n k^3$ est égale au carré de la somme $\sum_{k=1}^n k$.

Proposition 2.2.3 (somme d'une suite géométrique)

Pour tout $x \neq 1$, et tout n de \mathbb{N}^* : $S_n(x) = 1 + x + x^2 + \dots + x^{n-1} = \sum_{k=0}^{n-1} x^k = \frac{x^n - 1}{x - 1}$.

On a bien sûr $S_n(1) = n$.

Par exemple, pour tout n de \mathbb{N}^* : $1 + \frac{1}{2} + \frac{1}{4} + \dots + \frac{1}{2^{n-1}} = \sum_{k=0}^{n-1} \frac{1}{2^k} = 2 \left(1 - \frac{1}{2^n}\right)$

Plus généralement, si on cherche la somme S de n termes successifs a_k d'une progression géométrique de raison q ($a_{k+1} = qa_k$ pour tout k), et si les termes de début et de fin de cette somme sont d et f alors la formule à retenir est : $\sum a_k = \frac{qf - d}{q - 1}$ (si $q \neq 1$, car si $q = 1$ la somme vaut na_0).

On peut calculer $T_n(x) = 1 + 2x + 3x^2 + \dots + nx^{n-1}$ par dérivation de la somme $S_{n+1}(x)$.

On obtient $T_n(x) = \sum_{k=1}^n kx^{k-1} = S'_{n+1}(x) = \left(\frac{x^{n+1} - 1}{x - 1}\right)' = \frac{nx^{n+1} - (n+1)x^n + 1}{(x-1)^2}$ (si $x \neq 1$).

Le calcul précédent est valable pour x dans \mathbb{R} , mais l'expression finale de $T_n(x)$ reste vraie dans \mathbb{C} .

Voici un résultat analogue au précédent (mais davantage du point de vue de la factorisation) :

Proposition 2.2.4

Pour tous x, y dans \mathbb{C} , et pour tout n de \mathbb{N}^* , on a la factorisation :

$$x^n - y^n = (x - y) \sum_{k=0}^{n-1} x^{n-1-k} y^k = (x - y)(x^{n-1} + x^{n-2}y + \dots + xy^{n-2} + y^{n-1}).$$

Si l'entier n est impair (en appliquant ce qui précède à $-y$ plutôt qu'à y) :

$$x^n + y^n = (x + y) \sum_{k=0}^{n-1} (-1)^k x^{n-1-k} y^k = (x + y)(x^{n-1} - x^{n-2}y + \dots + (-1)^k x^{n-1-k} y^k + \dots - xy^{n-2} + y^{n-1})$$

On trouve en particulier les factorisations :

$$\begin{aligned}x^2 - y^2 &= (x - y)(x + y) & x^3 - y^3 &= (x - y)(x^2 + xy + y^2) & x^3 + y^3 &= (x + y)(x^2 - xy + y^2) \\x^4 - y^4 &= (x - y)(x^3 + x^2y + xy^2 + y^3) & x^5 - y^5 &= (x - y)(x^4 + x^3y + x^2y^2 + xy^3 + y^4) \\x^5 + y^5 &= (x + y)(x^4 - x^3y + x^2y^2 - xy^3 + y^4)\end{aligned}$$

De même, en choisissant $y = 1$, on a les factorisations :

$$\begin{aligned}x^2 - 1 &= (x - 1)(x + 1) & x^3 - 1 &= (x - 1)(x^2 + x + 1) & x^3 + 1 &= (x + 1)(x^2 - x + 1) \\x^4 - 1 &= (x - 1)(x^3 + x^2 + x + 1) & x^5 - 1 &= (x - 1)(x^4 + x^3 + x^2 + x + 1) \\x^5 + 1 &= (x + 1)(x^4 - x^3 + x^2 - x + 1)\end{aligned}$$

Et plus généralement :

Pour tout entier n : $x^n - 1 = (x - 1) \sum_{k=0}^{n-1} x^k$, et pour tout n impair $x^n + 1 = (x + 1) \sum_{k=0}^{n-1} (-1)^k x^{n-1-k}$

2.3 Factorielles et coefficients binomiaux

2.3.1 Factorielle d'un entier

Définition 2.3.1 (factorielle d'un entier)

On pose $0! = 1$, et pour tout n de \mathbb{N}^* , on note $n! = n(n-1)!$ (le symbole $n!$ se lit « factorielle n »)

Remarques

– L'énoncé précédent est un exemple de définition *réursive*.

– Pour tout n de \mathbb{N}^* , l'entier $n!$ est le produit des entiers de 1 à n , c'est-à-dire $n! = \prod_{k=1}^n k$

– On retiendra les valeurs :

$$0! = 1, 1! = 1, 2! = 2, 3! = 6, 4! = 24, 5! = 120, 6! = 720, 7! = 5040.$$

– L'entier $n!$ désigne le nombre de *permutations* d'un ensemble à n éléments (c'est-à-dire de bijections de cet ensemble sur lui-même).

Par exemple les $3! = 6$ permutations des lettres du mot *abc* sont : *abc, acb, bac, bca, cab* et *cba*.

– Pour les grandes valeurs de n , on verra plus tard la *formule de Stirling* : $n! \sim n^n e^{-n} \sqrt{2\pi n}$.

La signification de \sim est que le quotient des deux expressions tend vers 1 quand n tend vers $+\infty$.

Par exemple, pour $n = 20$, on a $n! = 2432902008176640000$ et $n^n e^{-n} \sqrt{2\pi n} \approx 2.42278684676 \cdot 10^{18}$.

2.3.2 Coefficients binomiaux

Définition 2.3.2 (combinaisons de p éléments parmi n)

Soient n et p deux entiers, avec $0 \leq p \leq n$. Soit E un ensemble fini possédant n éléments.

On note $\binom{n}{p}$ le nombre de sous-ensemble de E possédant p éléments.

Remarques

- L'ensemble E dont il est question ici est évidemment sans importance.
Par exemple, dans l'ensemble $E = \{a, b, c, d, e\}$, il y a 10 parties à trois éléments, qui sont $\{a, b, c\}, \{a, b, d\}, \{a, b, e\}, \{a, c, d\}, \{a, c, e\}, \{a, d, e\}, \{b, c, d\}, \{b, c, e\}, \{b, d, e\}, \{c, d, e\}$
- Le coefficient $\binom{n}{p}$ se lit « p parmi n ».
- Si E est un ensemble à n éléments (avec $n \geq 0$), il y a dans E :
une seule partie vide, donc $\binom{n}{0} = 1$, et une seule partie à n éléments (E lui-même), donc $\binom{n}{n} = 1$.
- On étend la définition en posant $\binom{n}{p} = 0$ si $p < 0$ ou si $p > n$ (ce qui est finalement assez logique).

Proposition 2.3.1

Soient n et p deux entiers, avec $0 \leq p \leq n$. On a l'égalité $\binom{n}{p} = \frac{n!}{p!(n-p)!}$

2.3.3 Relations entre coefficients binomiaux**Proposition 2.3.2**

On a les identités $\binom{n}{p} = \binom{n}{n-p}$, et $\binom{n}{p} = \binom{n-1}{p} + \binom{n-1}{p-1}$.

Cette dernière formule, avec $\binom{n}{0} = \binom{n}{n} = 1$, permet de calculer les $\binom{n}{p}$ de proche en proche.

On place souvent les $\binom{n}{p}$ dans un tableau triangulaire, de lignes et colonnes numérotées à partir de 0. Le coefficient $\binom{n}{p}$ vient alors se placer à l'intersection de la ligne d'indice n et de la colonne d'indice p . Le tableau obtenu est connu sous le nom de « triangle de Pascal » :

	$p = 0$	$p = 1$	$p = 2$	$p = 3$	$p = 4$	$p = 5$	$p = 6$	\dots
$n = 0$	1							
$n = 1$	1	1						
$n = 2$	1	2	1					
$n = 3$	1	3	3	1				
$n = 4$	1	4	6	4	1			
$n = 5$	1	5	10	10	5	1		
$n = 6$	1	6	15	20	15	6	1	
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\ddots	\ddots
n	$\binom{n}{0}$	$\binom{n}{1}$	$\binom{n}{2}$	$\binom{n}{3}$	$\binom{n}{4}$	$\binom{n}{5}$	$\binom{n}{6}$	\ddots
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\ddots

Proposition 2.3.3

on a les égalités : $\binom{n}{p+1} = \frac{n-p}{p+1} \binom{n}{p}$, $\binom{n}{p} = \frac{n}{p} \binom{n-1}{p-1}$, $\binom{n}{p} = \frac{n}{n-p} \binom{n-1}{p}$

Proposition 2.3.4 (Formule du binôme)

Pour tous x, y de \mathbb{C} , et pour tout n de \mathbb{N} : $(x + y)^n = \sum_{k=0}^n \binom{n}{k} x^k y^{n-k}$.

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

En particulier, pour tous x et y de \mathbb{C} :

$$\begin{aligned} (x + y)^2 &= x^2 + 2xy + y^2 & (x - y)^2 &= x^2 - 2xy + y^2 \\ (x + y)^3 &= x^3 + 3x^2y + 3xy^2 + y^3 & (x - y)^3 &= x^3 - 3x^2y + 3xy^2 - y^3 \\ (x + y)^4 &= x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + y^4 & (x - y)^4 &= x^4 - 4x^3y + 6x^2y^2 - 4xy^3 + y^4 \end{aligned}$$

2.4 Sommes doubles, interversions

Dans cette section, on désigne par Ω une partie finie de \mathbb{N}^2 .

On identifie tout couple (i, j) de Ω au point du plan d'abscisse i et d'ordonnée j .

On considère une application définie sur Ω , à valeurs dans \mathbb{C} , et on note $x_{i,j}$ l'image du couple (i, j) .

Soit $S = \sum_{(i,j) \in \Omega} x_{i,j}$ la somme des $x_{i,j}$, quand (i, j) parcourt Ω . On dit que S est une *somme double*.

Une méthode habituelle de calcul de S consiste à voir cette somme double comme l'enchaînement de deux sommes simples consécutives.

2.4.1 Somme sur un domaine rectangulaire

Ici Ω est le produit cartésien $I \times J$ de deux intervalles $I = \llbracket m, n \rrbracket$ et $J = \llbracket p, q \rrbracket$, avec $m \leq n$ et $p \leq q$.

L'ensemble Ω s'identifie alors aux points d'intersection d'une grille rectangulaire du plan.

Dans ces conditions $S = \sum_{(i,j) \in \Omega} x_{i,j} = \sum_{i=m}^n \left(\sum_{j=p}^q x_{i,j} \right) = \sum_{j=p}^q \left(\sum_{i=m}^n x_{i,j} \right)$.

La première expression évoque un parcours de Ω en colonnes : pour chaque i en abscisse ($m \leq i \leq n$), on forme la somme « verticale » $V_i = \sum_{j=p}^q x_{i,j}$ et on termine en calculant $\sum_{i=m}^n V_i$.

La deuxième expression évoque un parcours de Ω en lignes : pour chaque j en ordonnée ($p \leq j \leq q$), on forme la somme « horizontale » $H_j = \sum_{i=m}^n x_{i,j}$ et on termine en calculant $\sum_{j=p}^q H_j$.

Il y a donc une « somme interne » et une « somme externe ». Le caractère rectangulaire de Ω fait que les bornes de la somme interne ne dépendent pas de l'« indice courant » dans la somme externe, et qu'on peut librement intervertir les deux sommes.

Dans la pratique, on écrira indifféremment : $S = \sum_{\substack{m \leq i \leq n \\ p \leq j \leq q}} x_{i,j} = \sum_{i=m}^n \sum_{j=p}^q x_{i,j} = \sum_{j=p}^q \sum_{i=m}^n x_{i,j}$

Dans le cas particulier où $I = J = \llbracket m, n \rrbracket$, on pourra écrire $S = \sum_{m \leq i, j \leq n} x_{i,j}$

Mise en facteur de termes indépendants de l'indice de sommation

Considérons la somme $S = \sum_{\substack{m \leq i \leq n \\ p \leq j \leq q}} x_{i,j}$, où $x_{i,j}$ peut s'écrire $x_{i,j} = \lambda_i \mu_j y_{i,j}$.

On peut alors factoriser λ_i dans une somme interne sur j , et μ_j dans une somme interne sur i .

$$\text{Plus précisément, on peut écrire : } S = \sum_{i=m}^n \left(\sum_{j=p}^q \lambda_i \mu_j y_{i,j} \right) = \sum_{i=m}^n \left(\lambda_i \sum_{j=p}^q \mu_j y_{i,j} \right) = \sum_{j=p}^q \left(\mu_j \sum_{i=m}^n \lambda_i x_{i,j} \right)$$

Dans le cas où $x_{i,j}$ s'écrit $x_{i,j} = \lambda_i \mu_j$, la factorisation est encore plus prononcée :

$$\text{Dans ce cas on écrira : } S = \sum_{i=m}^n \left(\sum_{j=p}^q \lambda_i \mu_j \right) = \sum_{i=m}^n \left(\lambda_i \sum_{j=p}^q \mu_j \right) = \left(\sum_{i=m}^n \lambda_i \right) \left(\sum_{j=p}^q \mu_j \right)$$

2.4.2 Somme sur un domaine triangulaire

Un cas particulier important est celui où Ω est un domaine triangulaire du plan.

Pour fixer les idées, supposons que $\Omega = \{(i, j) \in \mathbb{N}^2, 0 \leq i \leq j \leq n\}$.

$$\text{Dans ces conditions } S = \sum_{(i,j) \in \Omega} x_{i,j} = \sum_{0 \leq i \leq j \leq n} x_{i,j} = \sum_{i=0}^n \left(\sum_{j=i}^n x_{i,j} \right) = \sum_{j=0}^n \left(\sum_{i=0}^j x_{i,j} \right).$$

Là aussi, on choisit un parcours de Ω en lignes ou en colonnes, mais il faut prendre garde au fait que les bornes de la somme interne dépendent de l'indice courant dans la somme externe : l'intervention nécessite donc un peu de réflexion.

Dans le même ordre d'idée, on écrira les interversions suivantes (liste non exhaustive) :

$$\sum_{0 \leq j \leq i \leq n} x_{i,j} = \sum_{j=0}^n \sum_{i=j}^n x_{i,j} = \sum_{i=0}^n \sum_{j=0}^i x_{i,j}, \quad \text{ou encore} \quad \sum_{0 \leq i < j \leq n} x_{i,j} = \sum_{i=0}^n \sum_{j=i+1}^n x_{i,j} = \sum_{j=1}^n \sum_{i=0}^{j-1} x_{i,j}$$

$$\sum_{\substack{0 \leq i,j \leq n \\ 0 \leq i+j \leq n}} x_{i,j} = \sum_{j=0}^n \sum_{i=0}^{n-j} x_{i,j} = \sum_{i=0}^n \sum_{j=0}^{n-i} x_{i,j} \quad (\text{ici, il est très recommandé de faire un dessin !})$$

2.4.3 Sommation par partition

Si Ω est la réunion **disjointe** de deux ou de plusieurs parties $\Omega_1, \Omega_2, \Omega_3, \dots$ du plan, on pourra écrire :

$$S = \sum_{(i,j) \in \Omega} x_{i,j} = \sum_{(i,j) \in \Omega_1} x_{i,j} + \sum_{(i,j) \in \Omega_2} x_{i,j} + \sum_{(i,j) \in \Omega_3} x_{i,j} \dots$$

Cette méthode est utile lorsque l'expression de $x_{i,j}$ varie en fonction de la position respectives de deux indices i et j (par exemple selon que $i < j$, $i = j$, ou $i > j$).

On pourra par exemple écrire, si on sait que $\Omega = \llbracket m, n \rrbracket^2$:

$$\sum_{i,j} x_{i,j} = \sum_{i \leq j} x_{i,j} + \sum_{i > j} x_{i,j} = \sum_{i=m}^n x_{i,i} + \sum_{i < j} x_{i,j} + \sum_{i > j} x_{i,j} = \sum_{i=m}^n x_{i,i} + \sum_{i \neq j} x_{i,j}$$

Il est possible aussi de décrire une zone triangulaire par des parallèles aux diagonales :

$$\text{Par exemple (en faisant un dessin!) : } \sum_{0 \leq j \leq i \leq n} x_{i,j} = \sum_{d=0}^n \left(\sum_{\substack{0 \leq j \leq i \leq n \\ i-j=d}} x_{i,j} \right) = \sum_{d=0}^n \left(\sum_{i=d}^n x_{i,i-d} \right) = \sum_{d=0}^n \left(\sum_{j=0}^{n-d} x_{j+d,j} \right)$$

Ou encore (et là aussi, on s'aide obligatoirement d'une représentation graphique) :

$$\sum_{0 \leq i \leq n} \left(\sum_{0 \leq j \leq n-i} x_{i,j} \right) = \sum_{0 \leq j \leq n} \left(\sum_{0 \leq i \leq n-j} x_{i,j} \right) = \sum_{d=0}^n \left(\sum_{\substack{0 \leq i,j \leq n \\ i+j=d}} x_{i,j} \right) = \sum_{d=0}^n \left(\sum_{i=0}^d x_{i,d-i} \right) = \sum_{d=0}^n \left(\sum_{j=0}^d x_{d-j,j} \right)$$

2.4.4 Produits de sommes

Soit $X = \sum_{i \in I} x_i$ et $Y = \sum_{j \in J} y_j$ deux sommes finies.

On développe leur produit en écrivant $XY = \sum_{i \in I} x_i \sum_{j \in J} y_j = \sum_{(i,j) \in I \times J} x_i y_j$.

Pour le développement de XY , il faut utiliser deux indices différents !

On n'écrira donc jamais : $\sum_{i=0}^n x_i \sum_{i=0}^n y_i = \dots$, le risque étant l'**erreur fatale** « $\sum_{i=0}^n x_i \sum_{i=0}^n y_i = \sum_{i=0}^n x_i y_i$ »

On peut généraliser ce qui précède au développement du produit de plusieurs sommes :

Par exemple, le développement $\sum_{i \in I} x_i \sum_{j \in J} y_j \sum_{k \in K} z_k = \sum_{(i,j,k) \in I \times J \times K} x_i y_j z_k$ donne une *somme triple*.

Enfin, les produits de produits ne se comportent pas comme les produits de sommes !

Par exemple : $\prod_{i \in I} x_i \prod_{i \in I} y_i = \prod_{i \in I} (x_i y_i)$, ou encore $\left(\prod_{i \in I} x_i \right)^2 = \prod_{i \in I} x_i^2$ (facile, si on réfléchit d'abord).

Développement du carré ou du cube d'une somme

Soit $I = \llbracket m, n \rrbracket$ un intervalle d'entiers.

On peut écrire de plusieurs manières différentes le développement du carré de la somme $\left(\sum_{i=m}^n x_i \right)^2$:

$$\left(\sum_{i=m}^n x_i \right)^2 = \sum_{m \leq i, j \leq n} x_i x_j = \sum_{i=m}^n \sum_{j=m}^n x_i x_j = \sum_{i=m}^n x_i^2 + \sum_{\substack{m \leq i, j \leq n \\ i \neq j}} x_i x_j = \sum_{i=m}^n x_i^2 + 2 \sum_{m \leq i < j \leq n} x_i x_j$$

« Le carré de la somme » c'est donc « la somme des carrés augmentée de tous les doubles produits ».

De la même manière, on peut développer le cube d'une somme en distinguant (parmi les $x_i x_j x_k$ du développement), ceux qui sont des cubes x_i^3 (donc obtenus quand $i = j = k$), ceux qui contiennent deux fois (et deux seulement) un même x_i et ceux pour lesquels les trois indices i, j, k sont distincts.

Avec un peu de réflexion, on obtient le développement suivant (on a simplifié l'écriture des sommes, sachant très bien que les indices i, j, k sont ici compris entre m et n) :

$$\left(\sum_{i=m}^n x_i \right)^3 = \sum_{m \leq i, j, k \leq n} x_i x_j x_k = \sum_{i=m}^n \sum_{j=m}^n \sum_{k=m}^n x_i x_j x_k = \sum_{i=m}^n x_i^3 + 3 \sum_{i \neq j} x_i^2 x_j + 6 \sum_{i < j < k} x_i x_j x_k$$

$$\text{Exemple } \begin{cases} 2x + 5y - 8z = 8 \\ 4x + 3y - 9z = 9 \\ 2x + 3y - 5z = 7 \\ x + 8y - 7z = 12 \end{cases} \quad (\text{sur-déterminé}) \quad \text{et} \quad \begin{cases} x + y + z + t + u = 7 \\ 3x + 2y + z + t - 3u = -2 \\ y + 2z + 2t + 6u = 23 \\ 5x + 4y + 3z + 3t - u = 12 \end{cases} \quad (\text{sous-déterminé})$$

– Certains systèmes sont dits *triangulaires* (ou « en escaliers », ou « en cascades », ou « échelonnés »).

$$\text{Par exemple, les systèmes } \begin{cases} x - 3y - z + t = 5 \\ 2y + z - 3t = 2 \\ z + t = 1 \end{cases} \quad \text{et} \quad \begin{cases} z = 5 \\ y + z = 1 \\ x + y + 3z = 1 \end{cases} \quad \text{sont échelonnés.}$$

– Dans certains cas, l'écriture d'un système linéaire (S) peut comporter un (ou plusieurs !) paramètre(s). Il faut alors résoudre en discutant suivant les valeurs du ou des paramètres.

$$\text{Exemples : } \begin{cases} x + y - mz = m + 2 \\ mx - y + 2z = 1 \\ x - my + z = m - 1 \end{cases} \quad (\text{paramètre } m) \quad \text{ou} \quad \begin{cases} ax + by + z = 1 \\ x + aby + z = b \\ x + by + az = 1 \end{cases} \quad (\text{paramètres } a, b)$$

Dans la résolution d'un système à paramètre m , il est important de savoir si m est à valeurs réelles ou complexes, car les cas particuliers de la discussion ne sont alors pas forcément les mêmes.

2.5.3 Interprétation géométrique (deux ou trois variables)

Dans cette sous-section, les coefficients et variables sont supposés appartenir à \mathbb{R} .

On identifie ici \mathbb{R}^2 (resp. \mathbb{R}^3) au plan (resp. à l'espace) muni d'un repère.

- Si $(a, b) \neq (0, 0)$, l'ensemble des (x, y) de \mathbb{R}^2 tels que $ax + by = c$ est une droite du plan.
- Si $(a, b, c) \neq (0, 0, 0)$, l'ensemble des (x, y, z) de \mathbb{R}^3 tels que $ax + by + cz = d$ est un plan de l'espace.

Résoudre $\begin{cases} ax + by = c \\ a'x + b'y = c' \\ \dots \end{cases}$ c'est donc chercher les points communs (éventuels) à des droites du plan.

Et résoudre $\begin{cases} ax + by + cz = d \\ a'x + b'y + c'z = d' \\ \dots \end{cases}$ c'est chercher les points communs (éventuels) à des plans de l'espace.

Intersection de deux droites du plan

On utilise la notation $\begin{vmatrix} \alpha & \gamma \\ \beta & \delta \end{vmatrix} = \alpha\delta - \beta\gamma$ (appelée « déterminant 2×2 »). On note que $\begin{vmatrix} \alpha & \gamma \\ \beta & \delta \end{vmatrix} = \begin{vmatrix} \alpha & \beta \\ \gamma & \delta \end{vmatrix}$.

Les deux vecteurs (α, β) et (γ, δ) sont proportionnels si et seulement si ce déterminant est nul.

On considère la droite \mathcal{D} d'équation $ax + by = c$, et la droite \mathcal{D}' d'équation $a'x + b'y = c'$.

Un vecteur directeur de \mathcal{D} est $u = (b, -a)$, et un vecteur directeur de \mathcal{D}' est $u' = (b', -a')$.

Un vecteur orthogonal à \mathcal{D} est $v = (a, b)$, et un vecteur orthogonal à \mathcal{D}' est $v' = (a', b')$.

Soit $(S) : \begin{cases} ax + by = c \\ a'x + b'y = c' \end{cases}$ le système précisant $\mathcal{D} \cap \mathcal{D}'$.

D'après le cours de TS, le système (S) s'écrit $AX = B$, avec $A = \begin{pmatrix} a & b \\ a' & b' \end{pmatrix}$, $X = \begin{pmatrix} x \\ y \end{pmatrix}$ et $B = \begin{pmatrix} c \\ c' \end{pmatrix}$.

On dit que A est la matrice du système (S) , et que B est la colonne des seconds membres.

On note $\begin{vmatrix} a & b \\ a' & b' \end{vmatrix} = ab' - ba'$ le *déterminant* de (S) (ou encore : le déterminant de la matrice A).

Proposition 2.5.1

Les conditions suivantes sont équivalentes :

- Les droites $(\mathcal{D}) : ax + by = c$ et $(\mathcal{D}') : a'x + b'y = c'$ ne sont pas parallèles.
- Les vecteurs $u = (b, -a)$ et $u' = (b', -a')$ ne sont pas proportionnels.
- Les vecteurs $v = (a, b)$ et $v' = (a', b')$ ne sont pas proportionnels.
- Le déterminant $\begin{vmatrix} a & b \\ a' & b' \end{vmatrix} = ab' - ba'$ n'est pas nul.
- La matrice $A = \begin{pmatrix} a & b \\ a' & b' \end{pmatrix}$ est inversible.

Si les deux droites \mathcal{D} et \mathcal{D}' ne sont pas parallèles, elles ont un unique point en commun.

Proposition 2.5.2

On suppose que le déterminant $\Delta = \begin{vmatrix} a & b \\ a' & b' \end{vmatrix} = ab' - ba'$ est non nul.

Le système $(S) : \begin{cases} ax + by = c \\ a'x + b'y = c' \end{cases}$ possède alors une solution unique (x, y) .

Elle est donnée par les « formules de Cramer » : $x = \frac{\Delta_x}{\Delta}$ et $y = \frac{\Delta_y}{\Delta}$, où $\Delta_x = \begin{vmatrix} c & b \\ c' & b' \end{vmatrix}$ et $\Delta_y = \begin{vmatrix} a & c \\ a' & c' \end{vmatrix}$

Avec les notations de TS, si $A = \begin{pmatrix} a & b \\ a' & b' \end{pmatrix}$ est inversible, alors sa matrice inverse est $A^{-1} = \frac{1}{\Delta} \begin{pmatrix} b' & -b \\ -a' & a \end{pmatrix}$.

Dans ce cas : $AX = B \Leftrightarrow X = A^{-1}B \Leftrightarrow \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{\Delta} \begin{pmatrix} b' & -b \\ -a' & a \end{pmatrix} \begin{pmatrix} c \\ c' \end{pmatrix} = \frac{1}{\Delta} \begin{pmatrix} cb' - c'b \\ ac' - a'c \end{pmatrix}$

On retrouve ainsi les formules de Cramer.

Cas où les droites \mathcal{D} et \mathcal{D}' sont parallèles

On suppose que $\mathcal{D} : ax + by = c$ et $\mathcal{D}' : a'x + b'y = c'$ sont parallèles.

Autrement dit, on suppose qu'il existe λ dans \mathbb{R} tel que $(a', b') = \lambda(a, b)$.

Les droites $\mathcal{D}, \mathcal{D}'$ sont alors confondues et si seulement si $(a', b', c') = \lambda(a, b, c)$.

Dans ce cas le système $(S) : \begin{cases} ax + by = c \\ a'x + b'y = c' \end{cases}$ se réduit à la seule équation $ax + by = c$ de \mathcal{D} .

Si \mathcal{D} et \mathcal{D}' sont parallèles mais pas confondues, alors $(S) : \begin{cases} ax + by = c \\ a'x + b'y = c' \end{cases}$ n'a pas de solution.

2.5.4 Système homogène associé

Définition 2.5.2

Un système linéaire est dit *homogène* si ses seconds membres b_i sont nuls.

À un système linéaire (S) on *associe* donc un système homogène (H) en annulant les seconds membres.

$$\text{Par exemple, } (H) : \begin{cases} 2x + 7y + 3z + 4t = 0 \\ x + 3y + 5z - 2t = 0 \\ x + 5y - 9z + 8t = 0 \\ 5x + 3y + 4z + 5t = 0 \end{cases} \text{ est homogène associé à } (S) : \begin{cases} 2x + 7y + 3z + 4t = 5 \\ x + 3y + 5z - 2t = 3 \\ x + 5y - 9z + 8t = 1 \\ 5x + 3y + 4z + 5t = 2 \end{cases}$$

Structure de la solution générale d'un système homogène

Un système homogène a toujours au moins la solution $(0, \dots, 0)$ (souvent notée $\vec{0}$) dite *solution triviale*.

Si $u = (x_1, x_2, \dots, x_p)$ et $v = (x'_1, x'_2, \dots, x'_p)$ sont deux solutions du système homogène (H) , alors tout p -uplet $w = \lambda u + \mu v = (\lambda x_1 + \mu x'_1, \lambda x_2 + \mu x'_2, \dots, \lambda x_p + \mu x'_p)$ est encore solution de (H) .

On exprime cela en disant que la solution générale de (H) est « stable par combinaisons linéaires ».

Avec les notations de TS, le système homogène (H) s'écrit $AX = 0$ (où 0 est ici une matrice colonne de coefficients tous nuls). La règle de calcul $A(\lambda X + \mu X') = \lambda AX + \mu AX'$ confirme que si les colonnes X, X' sont solutions de (H) , il en est de même de leurs combinaisons linéaires $\lambda X + \mu X'$.

Si la solution générale de (H) n'est pas *réduite* à $\vec{0}$, on montre qu'elle s'écrit comme l'ensemble des combinaisons linéaires $w = \lambda_1 u_1 + \lambda_2 u_2 + \dots + \lambda_r u_r$ d'un certain nombre r de solutions non nulles *indépendantes* (ces détails seront précisés ultérieurement). On dit alors que u_1, u_2, \dots, u_r constituent une *base de l'espace* des solutions de (H) .

2.5.5 Structure de la solution générale d'un système linéaire quelconque

Le résultat suivant donne la structure de l'ensemble des solutions d'un système linéaire quelconque.

Proposition 2.5.3

Soit (S) un système linéaire, et soit (H) le système homogène associé. La solution générale de (S) , si elle est non vide, s'obtient en ajoutant à la solution générale de (H) une solution particulière de (S) .

La proposition précédente se démontre facilement avec les notations de TS.

En effet, soit X_0 une colonne solution particulière du système (S) .

Dans ces conditions, la colonne X est solution de (S) si et seulement si $AX = B$ c'est-à-dire $AX = AX_0$, ou encore $A(X - X_0) = 0$, c'est-à-dire si et seulement si $X - X_0$ est une solution Z de (H) .

La solution générale de (S) s'écrit donc bien $X = X_0 + Z$, où Z est une solution quelconque de (H) .

Attention ! L'ensemble des solutions de (S) peut être vide. C'est le cas par exemple de $\begin{cases} x + 2y = 0 \\ x + 2y = 1 \end{cases}$

Le résultat précédent indique donc que la solution générale de (S) , si elle est non vide, est réduite à une seule solution (quand (H) n'a que la solution nulle) ou alors qu'elle est infinie.

- Par exemple, le système $(S) : \begin{cases} x - y = 0 \\ x + y = 2 \end{cases}$ possède l'unique solution $(x, y) = (1, 1)$.
- En revanche, l'ensemble des solutions de $(S) : \begin{cases} x - y = 1 \\ 2y + z = 1 \end{cases}$ est formé des triplets $(1 + \lambda, \lambda, 1 - 2\lambda)$.

On l'obtient en ajoutant à la solution particulière $(1, 0, 1)$ la solution générale $\lambda(1, 1, -2)$ de (H) .

2.5.6 Systèmes de Cramer triangulaires

Un *système de Cramer triangulaire* est un système linéaire à n équations et n inconnues, qui se présente sous forme triangulaire (supérieurement ou inférieurement), avec des coefficients diagonaux non nuls.

▷ **Systèmes de Cramer triangulaires supérieurs**

Dans ce cas, (S) s'écrit :

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1j}x_j + \cdots + a_{1n}x_n = b_1 \\ \phantom{a_{11}x_1} a_{22}x_2 + \cdots + a_{2j}x_j + \cdots + a_{2n}x_n = b_2 \\ \phantom{a_{11}x_1} \phantom{a_{22}x_2} \vdots \phantom{a_{2j}x_j} \phantom{a_{2n}x_n} \\ \phantom{a_{11}x_1} \phantom{a_{22}x_2} \phantom{a_{2j}x_j} \phantom{a_{2n}x_n} a_{n-1,n-1}x_{n-1} + a_{n-1,n}x_n = b_{n-1} \\ \phantom{a_{11}x_1} \phantom{a_{22}x_2} \phantom{a_{2j}x_j} \phantom{a_{2n}x_n} a_{nn}x_n = b_n \end{cases}$$

On trouve alors $x_n = \frac{b_n}{a_{nn}}$, puis $x_{n-1} = \frac{1}{a_{n-1,n-1}}(b_{n-1} - a_{n-1,n}x_n)$,

Une fois connus $x_n, x_{n-1}, \dots, x_{i+1}$, on trouve $x_i = \frac{1}{a_{ii}}(b_i - a_{i,i+1}x_{i+1} - \cdots - a_{ij}x_j - \cdots - a_{in}x_n)$

▷ **Systèmes de Cramer triangulaires inférieurs**

Le système (S) s'écrit ici :

$$\begin{cases} a_{11}x_1 & & & & = b_1 \\ a_{21}x_1 + a_{22}x_2 & & & & = b_2 \\ \vdots & \vdots & \vdots & \vdots & \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nj}x_j + \cdots + a_{nn}x_n & = & b_n \end{cases}$$

Dans ce cas, on trouve $x_1 = \frac{b_1}{a_{11}}$, puis $x_2 = \frac{1}{a_{22}}(b_2 - a_{21}x_1)$,

Une fois connus x_1, x_2, \dots, x_{i-1} , on trouve : $x_i = \frac{1}{a_{ii}}(b_i - a_{i1}x_1 - \cdots - a_{ij}x_j - \cdots - a_{i,i-1}x_{i-1})$

▷ **Systèmes de Cramer diagonaux**

La matrice du système étant diagonale, on obtient immédiatement : $\forall i \in \{1, \dots, n\}, x_i = \frac{b_i}{a_{ii}}$.

2.5.7 Opérations élémentaires sur les lignes d'un système

Définition 2.5.3

Soit (S) un système linéaire de n équations, à p inconnues et à coefficients dans \mathbb{K} .

Notons E_1, E_2, \dots, E_n les équations successives de (S).

On appelle *opération élémentaire* sur les lignes de (S) l'une des opérations suivantes :

- Multiplier une équation E_i par un scalaire *non nul* α . On note : $E_i \leftarrow \alpha E_i$.
- Ajouter à l'une des équations E_i un multiple d'une *autre* équation E_j . On note : $E_i \leftarrow E_i + \beta E_j$.
- Échanger deux équations E_i et E_j . On note : $E_i \leftrightarrow E_j$.

Proposition 2.5.4

Une opération élémentaire sur les lignes de (S) transforme le système (S) en un système (Σ) équivalent, c'est-à-dire ayant exactement les mêmes solutions que (S).

On peut enrichir la panoplie des opérations élémentaires avec les opérations suivantes :

- Remplacer l'équation E_i par $\alpha E_i + \beta E_j$, avec $\alpha \neq 0$ et $j \neq i$.

Il s'agit en fait de la composée des deux opérations $E_i \leftarrow \alpha E_i$ puis $E_i \leftarrow E_i + \beta E_j$.

- Ajouter à l'équation E_i une combinaison linéaire des autres équations du système.

Une telle opération peut s'écrire $E_i \leftarrow E_i + \sum_{j \neq i} \beta_j E_j$.

- On peut supprimer de (S) toute équation E_i qui est combinaison linéaire des autres équations de (S).

En effet si $E_i = \sum_{j \neq i} \beta_j E_j$, l'opération $E_i \leftarrow E_i - \sum_{j \neq i} \beta_j E_j$ remplace E_i par l'équation $0 = 0$, qui peut bien sûr être éliminée du système (S).

- Même si c'est moins fréquent, on peut adjoindre au système (S) une nouvelle équation obtenue par combinaison linéaire des équations initiales.

On peut interpréter cette modification de (S) en disant qu'on adjoint à (S) la nouvelle équation $0 = 0$ (on ne modifie pas l'ensemble des solutions) puis qu'on ajoute à celle-ci une combinaison linéaire des équations initiales.

2.5.8 Méthode « du pivot de Gauss »

▷ Principe de la méthode

Le principe est le suivant : par une suite d'opérations élémentaires, on transforme le système (S) en un système (Σ) équivalent et dont la matrice est échelonnée supérieurement. La résolution de (Σ) donne alors les solutions de (S).

▷ Mise en œuvre de la méthode

$$\text{Considérons le système : (S)} \left\{ \begin{array}{llll} a_{11} x_1 + a_{12} x_2 + \cdots + a_{1j} x_j + \cdots + a_{1p} x_p & = & b_1 \\ a_{21} x_1 + a_{22} x_2 + \cdots + a_{2j} x_j + \cdots + a_{2p} x_p & = & b_2 \\ \vdots & & \vdots & = & \vdots \\ a_{i1} x_1 + a_{i2} x_2 + \cdots + a_{ij} x_j + \cdots + a_{ip} x_p & = & b_i \\ \vdots & & \vdots & = & \vdots \\ a_{n1} x_1 + a_{n2} x_2 + \cdots + a_{nj} x_j + \cdots + a_{np} x_p & = & b_n \end{array} \right.$$

Supposons dans un premier temps que a_{11} est non nul.

On effectue alors des opérations élémentaires, avec a_{11} comme *pivot*, pour annuler les coefficients de x_1 dans les équations E_2, E_3, \dots, E_n :

$$\text{Avec } \left\{ \begin{array}{l} E_2 \leftarrow a_{11} E_2 - a_{21} E_1 \\ \vdots \\ E_i \leftarrow a_{11} E_i - a_{i1} E_1 \\ \vdots \\ E_n \leftarrow a_{11} E_n - a_{n1} E_1 \end{array} \right. \quad \text{(S) devient : (S')} \left\{ \begin{array}{llll} a_{11} x_1 + a_{12} x_2 + \cdots + a_{1j} x_j + \cdots + a_{1p} x_p & = & b_1 \\ & a'_{22} x_2 + \cdots + a'_{2j} x_j + \cdots + a'_{2p} x_p & = & b'_2 \\ & \vdots & \vdots & = & \vdots \\ & a'_{i2} x_2 + \cdots + a'_{ij} x_j + \cdots + a'_{ip} x_p & = & b'_i \\ & \vdots & \vdots & = & \vdots \\ & a'_{n2} x_2 + \cdots + a'_{nj} x_j + \cdots + a'_{np} x_p & = & b'_n \end{array} \right.$$

2.5.9 Trois exemples

Premier exemple

$$\text{Résoudre le système (S) : } \begin{cases} x + 2y + 3z + 4t = 11 \\ 2x + 3y + 4z + t = 12 \\ 3x + 4y + z + 2t = 13 \\ 4x + y + 2z + 3t = 14 \end{cases}$$

On applique la méthode du pivot, et on procède par équivalences :

$$\begin{aligned} & \begin{cases} x + 2y + 3z + 4t = 11 \\ 2x + 3y + 4z + t = 12 \\ 3x + 4y + z + 2t = 13 \\ 4x + y + 2z + 3t = 14 \end{cases} \iff \begin{pmatrix} E_2 \leftarrow 2E_1 - E_2 \\ E_3 \leftarrow 3E_1 - E_3 \\ E_4 \leftarrow 4E_1 - E_4 \end{pmatrix} \begin{cases} x + 2y + 3z + 4t = 11 \\ y + 2z + 7t = 10 \\ 2y + 8z + 10t = 20 \\ 7y + 10z + 13t = 30 \end{cases} \\ & \iff \begin{pmatrix} E_3 \leftarrow E_3 - 2E_2 \\ E_4 \leftarrow 7E_2 - E_4 \end{pmatrix} \begin{cases} x + 2y + 3z + 4t = 11 \\ y + 2z + 7t = 10 \\ 4z - 4t = 0 \\ 4z + 36t = 40 \end{cases} \\ & \iff \begin{pmatrix} E_4 \leftarrow E_4 - E_3 \end{pmatrix} \begin{cases} x + 2y + 3z + 4t = 11 \\ y + 2z + 7t = 10 \\ 4z - 4t = 0 \\ 40t = 40 \end{cases} \iff \begin{cases} t = 1 \\ z = t = 1 \\ y = -2z - 7t + 10 = 1 \\ x = 11 - 2y + 3z + 4t = 2 \end{cases} \end{aligned}$$

Le système (S) possède donc l'unique solution $(x, y, z, t) = (2, 1, 1, 1)$.

Deuxième exemple

$$\text{Résoudre le système (S) } \begin{cases} x + 3y + 5z - 2t - 7u = 3 \\ 3x + y + z - 2t - u = 1 \\ 2x - y - 3z + 7t + 5u = 2 \\ 3x - 2y - 5z + 7t + 8u = \lambda \end{cases} \quad (\text{où } \lambda \text{ un paramètre réel})$$

On applique la méthode du pivot, et on procède par équivalences :

$$\begin{aligned} (S) \quad & \iff \begin{pmatrix} E_2 \leftarrow (3E_1 - E_2)/2 \\ E_3 \leftarrow 2E_1 - E_3 \\ E_4 \leftarrow 3E_1 - E_4 \end{pmatrix} \begin{cases} x + 3y + 5z - 2t - 7u = 3 \\ 4y + 7z - 2t - 10u = 4 \\ 7y + 13z - 11t - 19u = 4 \\ 11y + 20z - 13t - 29u = 9 - \lambda \end{cases} \\ & \iff \begin{pmatrix} E_3 \leftarrow (4E_3 - 7E_2)/3 \\ E_4 \leftarrow 4E_4 - 11E_2 \end{pmatrix} \begin{cases} x + 3y + 5z - 2t - 7u = 3 \\ 4y + 7z - 2t - 10u = 4 \\ z - 10t - 2u = -4 \\ 3z - 30t - 6u = -4\lambda - 8 \end{cases} \\ & \iff \begin{pmatrix} E_4 \leftarrow E_4 - 3E_3 \end{pmatrix} \begin{cases} x + 3y + 5z - 2t - 7u = 3 \\ 4y + 7z - 2t - 10u = 4 \\ z - 10t - 2u = -4 \\ 0 = 4 - 4\lambda \end{cases} \end{aligned}$$

On constate que si $\lambda \neq 1$, alors (S) n'a pas de solution.

Supposons donc $\lambda = 1$. Le système (S) se réduit au trois premières équations ci-dessus.

On constate qu'on obtient un système de Cramer par rapport aux inconnues x, y, z , à condition de traiter les variables t et u comme des paramètres arbitraires.

Voici comment se termine la résolution de (S) :

$$(S) \iff \begin{cases} x + 3y + 5z = 3 + 2t + 7u \\ 4y + 7z = 4 + 2t + 10u \\ z = -4 + 10t + 2u \end{cases}$$

$$\begin{array}{l} E_1 \leftarrow E_1 - 5E_3 \\ E_2 \leftarrow (E_2 - 7E_3)/4 \\ \iff \end{array} \begin{cases} x + 3y = 23 - 48t - 3u \\ y = 8 - 17t - u \\ z = -4 + 10t + 2u \end{cases} \iff \begin{array}{l} E_1 \leftarrow E_1 - 3E_2 \\ \iff \end{array} \begin{cases} x = -1 + 3t \\ y = 8 - 17t - u \\ z = -4 + 10t + 2u \end{cases}$$

Les solutions de (S) sont les 5-uplets $A = (x, y, z, t, u)$ qui s'écrivent :

$$\begin{aligned} (x, y, z, t, u) &= (-1 + 3t, 8 - 17t - u, -4 + 10t + 2u, t, u) \\ &= (-1, 8, -4, 0, 0) + t(3, -17, 10, 1, 0) + u(0, -1, 2, 0, 1) \end{aligned}$$

où t et u ont des valeurs arbitraires.

On voit bien ici la structure de l'ensemble des solutions : à la solution particulière $A_0 = (-1, 8, -4, 0, 0)$ (obtenue pour $t = u = 0$), on ajoute la solution générale du système homogène (H) associé à (S) , et la solution générale de (H) est le plan engendré par $(3, -17, 10, 1, 0)$ et $(0, -1, 2, 0, 1)$.

Troisième exemple

$$\text{Résoudre le système } (S) : \begin{cases} \lambda x + y + z + t = 1 \\ x + \lambda y + z + t = -2 \\ x + y + \lambda z + t = 0 \\ x + y + z + \lambda t = 3 \end{cases}$$

Il est ici rentable d'adjoindre à (S) une nouvelle équation combinaison linéaire des équations initiales. Plus précisément on va lui adjoindre l'équation $E_1 + E_2 + E_3 + E_4$.

$$\text{On obtient le système équivalent : } \begin{cases} \lambda x + y + z + t = 1 \\ x + \lambda y + z + t = -2 \\ x + y + \lambda z + t = 0 \\ x + y + z + \lambda t = 3 \\ (\lambda + 3)(x + y + z + t) = 2 \end{cases}$$

On voit tout de suite que le système n'a pas de solution si $\lambda = -3$.

Supposons donc $\lambda \neq -3$.

On peut alors remplacer E_5 par $E_5 = E_5/(\lambda + 3)$.

On obtient

$$\begin{cases} \lambda x + y + z + t = 1 \\ x + \lambda y + z + t = -2 \\ x + y + \lambda z + t = 0 \\ x + y + z + \lambda t = 3 \\ x + y + z + t = \frac{2}{\lambda + 3} \end{cases} \iff \begin{array}{l} E_1 \leftarrow E_1 - E_5 \\ E_2 \leftarrow E_2 - E_5 \\ E_3 \leftarrow E_3 - E_5 \\ E_4 \leftarrow E_4 - E_5 \\ \iff \end{array} \begin{cases} (\lambda - 1)x = \frac{\lambda + 1}{\lambda + 3} \\ (\lambda - 1)y = -2\frac{\lambda + 4}{\lambda + 3} \\ (\lambda - 1)z = \frac{-2}{\lambda + 3} \\ (\lambda - 1)t = \frac{3\lambda + 7}{\lambda + 3} \end{cases}$$

On voit que si $\lambda = 1$ le système (S) n'a pas de solution.

On suppose donc que λ n'est ni égal à -3 ni égal à 1 .

Le système (S) possède alors une solution unique (x, y, z, t) donnée par :

$$(x, y, z, t) = \frac{1}{(\lambda - 1)(\lambda + 3)} (\lambda + 1, -2\lambda - 8, -2, 3\lambda + 7)$$

Chapitre 3

Nombres complexes et trigonométrie

Sommaire

3.1	Notation cartésienne, plan complexe	47
3.1.1	Notation cartésienne, partie réelle, partie imaginaire	47
3.1.2	Plan complexe. Affixe d'un point, d'un vecteur	48
3.1.3	Conjugué d'un nombre complexe	49
3.1.4	Notion de transformations du plan complexe	50
3.2	Module et distance dans le plan complexe	51
3.2.1	Module d'un nombre complexe	51
3.2.2	Distance dans le plan complexe	52
3.2.3	Nombres complexes de module 1	53
3.3	Trigonométrie circulaire	54
3.3.1	Une « définition » des fonctions $t \mapsto e^{it}$, $t \mapsto \cos t$ et $t \mapsto \sin t$	54
3.3.2	Propriétés de l'application e^{it}	55
3.3.3	Premières propriétés des fonctions $x \mapsto \sin x$ et $x \mapsto \cos x$	55
3.3.4	Formules d'Euler, linéarisation	58
3.3.5	Utilisation de la formule de De Moivre	59
3.3.6	Deux sommes trigonométriques classiques	60
3.3.7	La fonction tangente $x \mapsto \tan x$	60
3.4	Forme trigonométrique (polaire)	62
3.4.1	Module et argument d'un nombre complexe non nul	62
3.4.2	Forme polaire et opérations dans \mathbb{C}	63
3.4.3	Interprétation géométrique du produit	64
3.5	Équation du second degré dans \mathbb{C}	66
3.5.1	Racines carrées d'un nombre complexe	66
3.5.2	Équations du second degré dans \mathbb{C}	66
3.6	Racines n-ièmes	67
3.6.1	Racines n -ièmes de l'unité	67
3.6.2	Racines n -ièmes d'un nombre complexe	69
3.6.3	Généralisation (admise) aux racines des polynômes	70
3.7	Exponentielle complexe	70
3.7.1	Définition de e^z pour z dans \mathbb{C}	70
3.7.2	Propriétés de la fonction exponentielle	71
3.7.3	Résolution de l'équation $e^z = a$ dans \mathbb{C}	71
3.8	Interprétations géométriques	72

3.8.1	Module et argument de $(z - b)/(z - a)$	72
3.8.2	Similitudes directes	74
3.8.3	Symétries et projections orthogonales	75

3.1 Notation cartésienne, plan complexe

3.1.1 Notation cartésienne, partie réelle, partie imaginaire

Proposition 3.1.1 (opérations sur l'ensemble \mathbb{R}^2)

On munit l'ensemble \mathbb{R}^2 des deux lois suivantes :

$$\forall (x, y, x', y') \in \mathbb{R}^4, \begin{cases} (x, y) + (x', y') = (x + x', y + y') \\ (x, y)(x', y') = (xx' - yy', xy' + yx') \end{cases}$$

Muni de ces deux lois, \mathbb{R}^2 possède une « structure de corps commutatif ». Plus précisément :

- Le neutre pour la loi $+$ est $(0, 0)$, et l'opposé de (x, y) est $(-x, -y)$.
- Le neutre pour le produit est $(1, 0)$.
- Pour tout $z = (x, y)$ non nul, l'inverse de z est : $\frac{1}{z} = \left(\frac{x}{x^2 + y^2}, \frac{-y}{x^2 + y^2} \right)$.

Définition 3.1.1 (nombres complexes, notation provisoire)

On note \mathbb{C} l'ensemble \mathbb{R}^2 quand il est muni des deux lois précédentes.

Ses éléments $z = (x, y)$ sont appelés *nombres complexes*.

Proposition 3.1.2 (\mathbb{R} considéré comme une partie de \mathbb{C})

L'application $\varphi : x \mapsto (x, 0)$ est bijective de \mathbb{R} sur $\mathbb{K} = \{(x, 0), x \in \mathbb{R}\}$ et $\begin{cases} \varphi(x + x') = \varphi(x) + \varphi(x') \\ \varphi(xx') = \varphi(x)\varphi(x') \end{cases}$

Cette bijection permet donc d'identifier (algébriquement) le couple $(x, 0)$ avec le réel x .

De cette manière, on peut donc considérer que \mathbb{R} est une partie de \mathbb{C} .

Définition 3.1.2 (le nombre i , notation cartésienne des nombres complexes)

On pose $i = (0, 1)$. On rappelle que pour tout x de \mathbb{R} , on identifie $(x, 0)$ et x .

Tout z de \mathbb{C} s'écrit de façon unique $z = (x, y)$, avec x et y dans \mathbb{R} .

Mais l'égalité $z = (x, y)$ équivaut à $z = (x, 0) + (0, 1)(y, 0)$, c'est-à-dire $z = x + iy$.

On a ainsi obtenu la notation *cartésienne* (ou *algébrique*) des nombres complexes.

Le réel x est appelé *partie réelle* de z et est noté $\operatorname{Re}(z)$.

Le réel y est appelé *partie imaginaire* de z et est noté $\operatorname{Im}(z)$.

Définition 3.1.3 (nombres complexes réels ou imaginaires purs)

Dire que z est *réel*, c'est dire que sa partie imaginaire $\operatorname{Im}(z)$ est nulle.

On dit que z est *imaginaire pur* si $\operatorname{Re}(z) = 0$, c'est-à-dire si $z = iy$, avec y dans \mathbb{R} .

Attention : dire que le complexe z n'est pas réel ne signifie pas qu'il est imaginaire pur !

Identifications entre parties réelles et parties imaginaires

Soient $\begin{cases} z = x + iy \\ z' = x' + iy' \end{cases}$ deux nombres complexes, avec $(x, y, x', y') \in \mathbb{R}^4$.

On sait que $z = z' \Leftrightarrow \begin{cases} x = x' \\ y = y' \end{cases}$ (on dit qu'on *identifie* les parties réelles et les parties imaginaires).

En particulier : $z = 0 \Leftrightarrow x = y = 0$ (attention à vérifier que x et y sont réels!).

Plus généralement, soit ω un nombre complexe non réel.

Alors tout z de \mathbb{C} s'écrit encore façon unique $z = a + b\omega$, avec a, b dans \mathbb{R} .

On peut alors encore identifier : $\forall (x, y, x', y') \in \mathbb{R}^4 : x + \omega y = x' + \omega y' \Leftrightarrow x = x'$ et $y = y'$.

Nouvelle écriture des opérations sur \mathbb{C}

Avec les notations précédentes, le nombre i vérifie $i^2 = -1$.

Soient $\begin{cases} z = x + iy \\ z' = x' + iy' \end{cases}$ deux nombres complexes, avec $(x, y, x', y') \in \mathbb{R}^4$.

Les opérations sur \mathbb{C} s'écrivent maintenant : $\begin{cases} z + z' = (x + x') + i(y + y') \\ zz' = (xx' - yy') + i(xy' + yx') \end{cases}$

Si $z = x + iy$ est non nul (c'est-à-dire $x \neq 0$ ou $y \neq 0$), l'inverse de z est $\frac{1}{z} = \frac{x - iy}{x^2 + y^2}$.

Puissances du nombre i

On constate que $i^2 = -1$. Donc $\frac{1}{i} = -i$. En fait, $z^2 = -1 \Leftrightarrow z \in \{i, -i\}$.

Plus généralement $i^3 = -i$, et $i^4 = 1$ (la suite $n \mapsto i^n$ est périodique de période quatre).

3.1.2 Plan complexe. Affixe d'un point, d'un vecteur

On considère le plan \mathcal{P} muni d'un repère orthonormé direct $(0, e_1, e_2)$.

Chaque point M de \mathcal{P} est donc désigné par un unique couple de coordonnées (x, y) dans ce repère.

Pour cette raison, on parlera souvent du plan Oxy , du point $M(x, y)$, des axes Ox et Oy , etc.

Définition 3.1.4

L'application qui à $z = x + iy$ (x, y réels) associe $M(x, y)$ est une bijection de \mathbb{C} sur le plan Oxy .

On dit que M est le *point image* de z , ou encore que z est l'*affixe* de M .

On note $M(z)$ pour désigner simultanément M et son affixe z .

Le plan Oxy , muni de cette correspondance, est appelé le *plan complexe*.

Le vecteur $\overrightarrow{OM} = xe_1 + ye_2$ est appelé *vecteur image* de $z = x + iy$ (et on dit que z est l'affixe de \overrightarrow{OM}).

Bien sûr, à la partie réelle et à la partie imaginaire d'un nombre complexe z correspondent donc l'abscisse et l'ordonnée du point M dans le repère $(0, e_1, e_2)$.

Remarque : on dit « une affixe » et non « un affixe ». Il est prudent de dire « l'affixe ».

L'identification entre \mathbb{C} et le plan muni d'un repère orthonormé permet :

- d'interpréter les propriétés de \mathbb{C} dans un langage géométrique (jusqu'à résoudre avec les outils de la géométrie des problèmes posés initialement en termes de nombres complexes).

- de traduire des notions de géométrie du plan dans le langage algébrique des nombres complexes (jusqu'à résoudre par le calcul des problèmes énoncés en termes purement géométriques).

Par exemple :

- L'axe Ox (resp. Oy) est l'ensemble des points images des nombres réels (resp. des imaginaires purs).
- Si A, B ont pour affixes a, b alors \overrightarrow{AB} a pour affixe $b - a$, et le milieu Ω de $[A; B]$ a pour affixe $\frac{a+b}{2}$.
- Si A, B, C, D sont quatre points du plan Oxy , d'affixes respectives a, b, c, d , alors $ABCD$ est un parallélogramme si et seulement si $\overrightarrow{AB} = \overrightarrow{DC}$, c'est-à-dire $b - a = c - d$, ou encore $a + c = b + d$.
- Si A, B sont deux points distincts du plan, d'affixes respectives a et b , la droite (AB) est l'ensemble des points $M(z)$, où $z = a + \lambda(b - a) = (1 - \lambda)a + \lambda b$, avec λ réel.

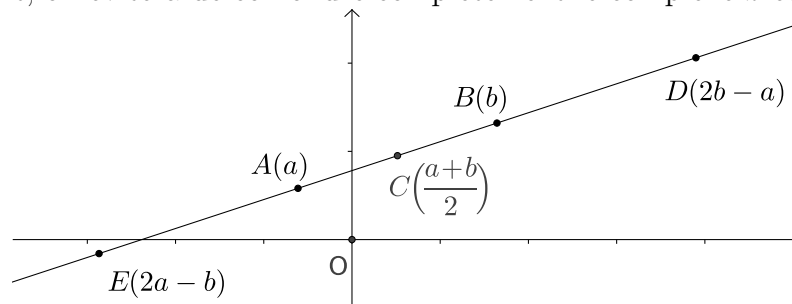
De même le segment $[A; B]$ est l'ensemble des points $M(z)$, où $z = a + \lambda(b - a)$, avec $0 \leq \lambda \leq 1$.

Soit u un élément de \mathbb{C}^* , et soit \vec{u} son vecteur image.

La droite \mathcal{D} passant par $A(a)$ et dirigée par \vec{u} est l'ensemble des $M(z = a + \lambda u)$, avec λ dans \mathbb{R} .

- L'isobarycentre des points $M_k(z_k)$ ($1 \leq k \leq n$) est le point G d'affixe $g = \frac{1}{n} \sum_{k=1}^n z_k$.

Même si c'est tentant, on évitera de confondre complètement le complexe z et le point M d'affixe z .



On a représenté ici deux points $A(a)$ et $B(b)$ et quelques points de la droite (AB) : le milieu C du segment $[A; B]$, le symétrique D de A par rapport à B , et le symétrique E de B par rapport à A .

3.1.3 Conjugué d'un nombre complexe

Définition 3.1.5

Soit $z = x + iy$ (x et y réels) un nombre complexe quelconque.

Le nombre complexe $\bar{z} = x - iy$ est appelé le *conjugué* de z .

On nomme *conjugaison* l'application de \mathbb{C} dans \mathbb{C} , définie par $z \rightarrow \bar{z}$.

Proposition 3.1.3

La conjugaison vérifie les propriétés suivantes :

Pour tout z de \mathbb{C} , on a : $\overline{\bar{z}} = z$ (on dit que la conjugaison est une opération involutive)

Pour tous z_1, z_2 de \mathbb{C} , on a : $\begin{cases} \overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2 \\ \overline{z_1 z_2} = \bar{z}_1 \bar{z}_2 \end{cases}$ (la conjugaison est compatible avec les opérations)

Remarques

- Les propriétés précédentes se généralisent à une somme ou à un produit fini.

Ainsi, pour tous nombres complexes z_1, \dots, z_n , $\sum_{k=1}^n z_k = \sum_{k=1}^n \bar{z}_k$ et $\prod_{k=1}^n z_k = \prod_{k=1}^n \bar{z}_k$

- Pour tout z complexe, on note les égalités $\operatorname{Re}(z) = \frac{z + \bar{z}}{2}$ et $\operatorname{Im}(z) = \frac{z - \bar{z}}{2i}$.

On en déduit que z est réel si et seulement si $\bar{z} = z$.

De même, z est imaginaire pur si et seulement si $\bar{z} = -z$.

3.1.4 Notion de transformations du plan complexe

Définition 3.1.6

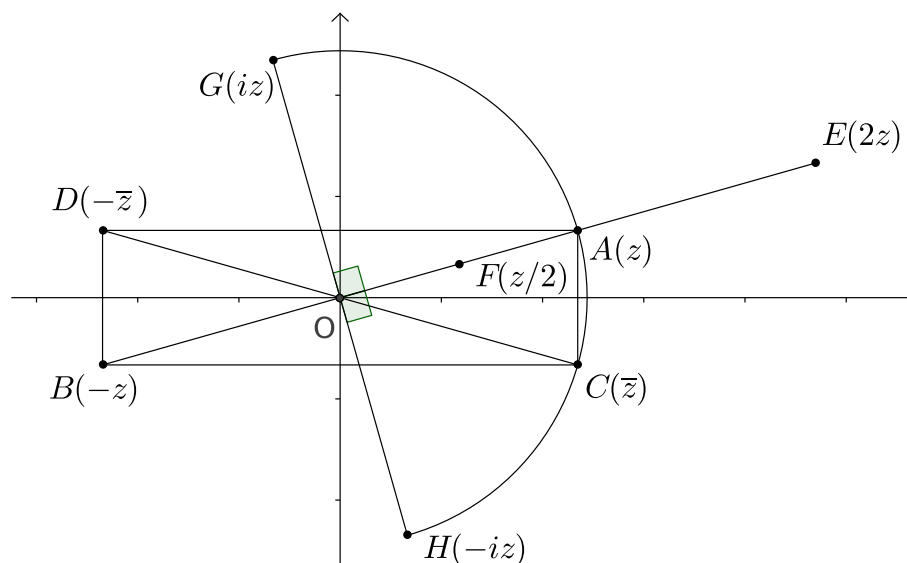
Soit g une application de \mathbb{C} dans \mathbb{C} (définie éventuellement sur une partie de \mathbb{C} .)

Il lui correspond de façon unique l'application f du plan dans lui-même, qui au point m d'affixe z associe le point M d'affixe $Z = g(z)$.

L'application $f : m(z) \mapsto M(Z)$ est appelée *transformation du plan complexe*.

Cas particuliers simples

- L'application $f : m(z) \mapsto M(Z = z + a)$ ($a \in \mathbb{C}$) est la translation de vecteur le vecteur image de a .
- L'application $f : m(z) \mapsto M(Z = -z)$ est la symétrie par rapport au point O .
Plus généralement la symétrie de centre $A(a)$ est donnée par $f : m(z) \mapsto M(Z = 2a - z)$.
- L'application $f : m(z) \mapsto M(Z = \bar{z})$ est la symétrie orthogonale par rapport à l'axe Ox .
L'application $f : m(z) \mapsto M(Z = -\bar{z})$ est la symétrie orthogonale par rapport à l'axe Oy .
- L'application $f : m(z) \mapsto M(Z = \lambda z)$, avec λ réel, est l'homothétie de centre O et de rapport λ .
L'homothétie de centre $A(a)$ et de rapport λ réel est donnée par $f : m(z) \mapsto Z = a + \lambda(z - a)$.
- L'application $f : m(z) \mapsto M(Z = iz)$ est la rotation de centre O et d'angle $\pi/2$.
L'application $f : m(z) \mapsto M(Z = -iz)$ est la rotation de centre O et d'angle $-\pi/2$.



On a représenté ci-dessus un point $A(z)$, et un certain nombre de points qui lui sont liés : le point B d'affixe $-z$, le point C d'affixe \bar{z} , et les points $D(-\bar{z})$, $E(2z)$, $F(z/2)$, $G(iz)$ et $H(-iz)$.

3.2 Module et distance dans le plan complexe

3.2.1 Module d'un nombre complexe

Définition 3.2.1

Soit $z = x + iy$ (x et y réels) un nombre complexe. La quantité $|z| = \sqrt{x^2 + y^2}$ est appelée *module* de z .

Relation entre le module et le conjugué

On constate que $z\bar{z} = |z|^2$. Cette égalité est (utile pour se « débarrasser » du module.

En particulier, si z est non nul, l'inverse de z s'écrit $\frac{1}{z} = \frac{\bar{z}}{|z|^2}$.

Si z est réel, le module de z est égal à sa valeur absolue. Les notations $||$ (valeur absolue dans \mathbb{R} et module dans \mathbb{C}) sont donc « compatibles ».

Module d'un produit, d'un quotient

Pour tous z et z' de \mathbb{C} , on a : $|z| \geq 0$; $|z| = 0 \Leftrightarrow z = 0$; $|zz'| = |z| |z'|$

Plus généralement, on a z_1, \dots, z_n , on a : $\left| \prod_{k=1}^n z_k \right| = \prod_{k=1}^n |z_k|$, et notamment : $\forall n \in \mathbb{N}, |z^n| = |z|^n$.

Si $z \neq 0$, alors : $\left| \frac{1}{z} \right| = \frac{1}{|z|}$, et $\left| \frac{z'}{z} \right| = \frac{|z'|}{|z|}$.

Inégalité triangulaire

Pour tous z, z' de \mathbb{C} , on a : $|z + z'| \leq |z| + |z'|$ (avec égalité $\Leftrightarrow \exists \lambda \in \mathbb{R}^+$ tel que $z' = \lambda z$ ou $z = \lambda z'$).

Cette inégalité se complète en : $||z| - |z'|| \leq |z \pm z'|$.

On peut donc écrire l'encadrement : $||z| - |z'|| \leq |z \pm z'| \leq |z| + |z'|$.

Conséquence : si $|z| \leq k < 1$, alors $1 - k \leq |1 + z| \leq 1 + k$.

Pour tout z de \mathbb{C} , on a aussi : $\max(|\operatorname{Re}(z)|, |\operatorname{Im}(z)|) \leq |z| \leq |\operatorname{Re}(z)| + |\operatorname{Im}(z)|$.

Généralisation au module d'une somme de n nombres complexes

Soient z_1, z_2, \dots, z_n des nombres complexes quelconques. Alors $\left| \sum_{k=1}^n z_k \right| \leq \sum_{k=1}^n |z_k|$

L'inégalité précédente est une égalité si et seulement si les z_k sont les produits de l'un d'entre eux par des réels positifs (c'est-à-dire, géométriquement, si les $M_k(z_k)$ sont sur une même demi-droite issue de O).

Module du carré d'une somme

Voici comment on peut développer le carré du module d'une somme (ou d'une différence) :

Pour tout u et v de \mathbb{C} :
$$\begin{cases} |u + v|^2 = |u|^2 + 2 \operatorname{Re}(u\bar{v}) + |v|^2 \\ |u - v|^2 = |u|^2 - 2 \operatorname{Re}(u\bar{v}) + |v|^2 \end{cases}$$

En ajoutant ces deux égalités, on obtient : $|u + v|^2 + |u - v|^2 = 2(|u|^2 + |v|^2)$.

Tout ça se généralise : on a en effet $\left| \sum_{k=1}^n z_k \right|^2 = \sum_{k=1}^n |z_k|^2 + 2 \sum_{1 \leq j < k \leq n} \operatorname{Re}(z_j \bar{z}_k)$.

3.2.2 Distance dans le plan complexe

Proposition 3.2.1 (Distance dans \mathbb{C})

Soit d l'application $\mathbb{C} \times \mathbb{C}$ vers \mathbb{R} , définie par : $\forall (z, z') \in \mathbb{C}^2, d(z, z') = |z - z'|$.

L'application d est une distance sur \mathbb{C} , ce qui signifie qu'elle vérifie les propriétés suivantes :

Pour tous nombres complexes u, v, w :

$$d(u, v) \geq 0; \quad d(u, v) = 0 \Leftrightarrow u = v; \quad d(u, v) = d(v, u).$$

$$d(u, v) \leq d(u, w) + d(w, v) \text{ (inégalité triangulaire)}$$

Interprétation géométrique (dans le plan Oxy)

Bien sûr $d(z, z') = |z' - z|$ est la distance, dans le plan, entre $M(z)$ et $M'(z')$.

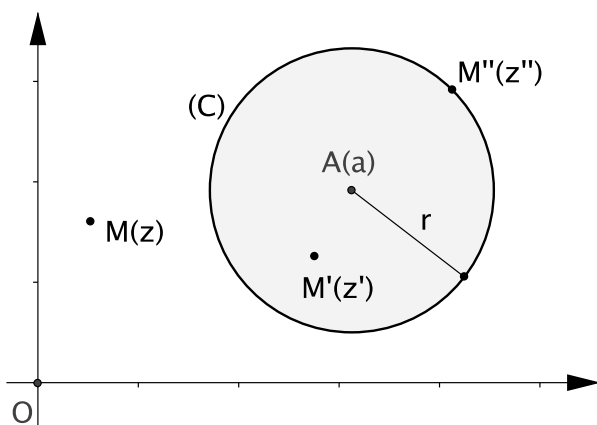
Les propriétés précédentes se comprennent alors aisément en termes géométriques.

Par exemple, si on note $U(u)$, $V(v)$, et $W(w)$, on a l'égalité $|w - u| = |w - v| + |v - u|$ si et seulement si $UV = UV + VW$, c'est-à-dire si et seulement si W est un élément du segment $[U; W]$. Cela équivaut à dire qu'il existe un réel λ dans $[0, 1]$ tel que $v = \lambda u + (1 - \lambda)w$.

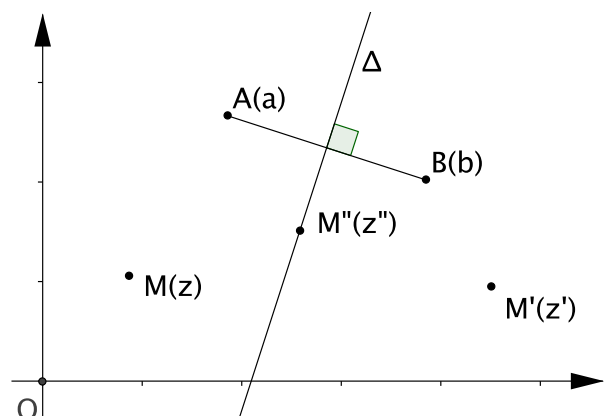
Propriétés géométriques liées au module

On note ici A et B deux points du plan, d'affixes respectives a et b . Soit r un réel positif ou nul.

- L'égalité $|z - a| = r$ caractérise les points $M(z)$ du cercle de centre A et de rayon r .
- L'inégalité large $|z - a| \leq r$ caractérise les points $M(z)$ du disque fermé de centre A et de rayon r .
L'inégalité stricte $|z - a| < r$ caractérise les points $M(z)$ du disque ouvert de centre A et de rayon r .
- Bien sûr les inégalités $|z - a| \geq r$ (ou $|z - a| > r$) caractérisent les points $M(z)$ qui sont à l'extérieur (largement, ou strictement) du cercle de centre A et de rayon r .
- L'égalité $|z - a| = |z - b|$ caractérise les points $M(z)$ de la médiatrice Δ du segment $[A; B]$.
Enfin $|z - a| < |z - b|$ caractérise les $M(z)$ du demi-plan ouvert délimité par Δ et contenant A .



(C) est le cercle de centre $A(a)$ et de rayon r .
 $M''(z'')$ est sur (C) car $|z'' - a| = r$.
 $M'(z')$ est intérieur à (C) car $|z' - a| < r$.
 $M(z)$ est extérieur à (C) car $|z - a| > r$.



Δ est la médiatrice du segment $[A; B]$.
 $M''(z'')$ est sur Δ car $|z'' - a| = |z'' - b|$.
 $M(z)$ et $M'(z')$ sont de part et d'autre de Δ .
 $|z - a| < |z - b|$ (ou encore $AM < BM$)
 $|z' - a| > |z' - b|$ (ou encore $AM' > BM'$)

3.2.3 Nombres complexes de module 1

Définition 3.2.2

On note \mathcal{U} l'ensemble des nombres complexes de module 1.

À cette définition algébrique correspondent des définitions géométriques :

Définition 3.2.3

On appelle « cercle unité » (ou encore « cercle trigonométrique ») le cercle de centre O et de rayon 1, c'est-à-dire l'ensemble des points $M(z)$ tels que $|z| = 1$ (c'est-à-dire avec z dans \mathcal{U}).

On appelle « disque unité ouvert » l'ensemble des points $M(z)$ tels que $|z| < 1$ (c'est l'intérieur, au sens strict, du cercle unité).

On appelle « disque unité fermé » l'ensemble des points $M(z)$ tels que $|z| \leq 1$ (c'est l'intérieur, au sens large, du cercle unité).

Stabilité de \mathcal{U} par produit

- Pour tous z de \mathbb{C}^* , on a $|z| = 1$ (c'est-à-dire z est dans \mathcal{U}) si et seulement si $\bar{z} = \frac{1}{z}$.
- Si z et z' sont dans \mathcal{U} , il en est de même de \bar{z} , de zz' et de $\frac{z}{z'}$, et de z^n , pour tout n de \mathbb{Z} .

Une construction intéressante

Supposons $|z| > 1$, c'est-à-dire supposons que $M(z)$ est extérieur au disque unité fermé.

Dans ces conditions, les deux points $N\left(\frac{1}{z}\right)$ et $P\left(\frac{1}{\bar{z}}\right)$ appartiennent au disque unité ouvert.

$$\text{On a : } \frac{1}{\bar{z}} = \frac{z}{|z|^2} \Rightarrow \overrightarrow{OP} = \frac{\overrightarrow{OM}}{OM^2} \Rightarrow OP \cdot OM = 1$$

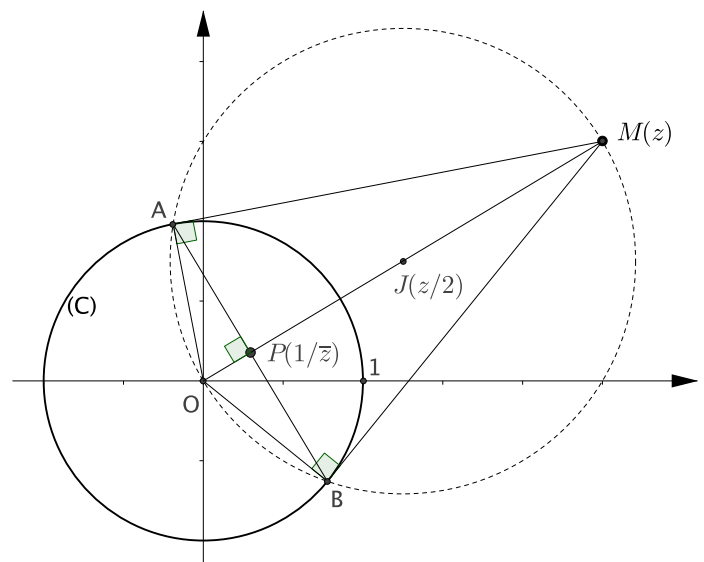
M, P sont donc sur une demi-droite issue de O .

Plus précisément, on a la construction suivante qui permet de passer de M à P , et inversement.

Le cercle de diamètre $[O; M]$ coupe le cercle unité (C) en deux points A, B (les points d'appuis des tangentes à (C) issues de M).

Le segment $[A; B]$ recoupe le segment $[O; M]$ orthogonalement au point P d'affixe $1/\bar{z}$.

Bien sûr le point d'affixe $1/z$ est le symétrique de P par rapport à l'axe Ox .



Pour justifier cette construction, on utilise des relations dans des triangles rectangles :

$$OM^2 = 1 + AM^2 = 1 + AP^2 + PM^2 = 2 - OP^2 + (OM - OP)^2 = OM^2 - 2(OP \cdot OM - 1)$$

Il en résulte l'égalité $OP \cdot OM = 1$. Donc si z est l'affixe de M , celle de P est $\frac{1}{\bar{z}}$.

3.3 Trigonométrie circulaire

3.3.1 Une « définition » des fonctions $t \mapsto e^{it}$, $t \mapsto \cos t$ et $t \mapsto \sin t$

Paradoxalement, il faut attendre la deuxième année de classe préparatoire pour une définition « rigoureuse » des fonctions $t \mapsto \cos t$ et $t \mapsto \sin t$.

On se contentera d'une définition intuitive qui consiste à « enrouler » sur le cercle unité l'axe vertical Δ passant par $A(1)$ et orienté vers le haut, comme indiqué ici.

On note e^{it} (et c'est une *définition*) l'affixe du point M' sur lequel vient s'enrouler un point $M(1 + it)$ de Δ .

On note π le demi-périmètre du cercle unité (considérons ça comme la *définition* de π !). Le point $C(1 + i\pi)$ s'enroule alors sur le point C' d'affixe -1 .

On a ainsi la remarquable *égalité d'Euler* $e^{i\pi} + 1 = 0$

De même on observe que :

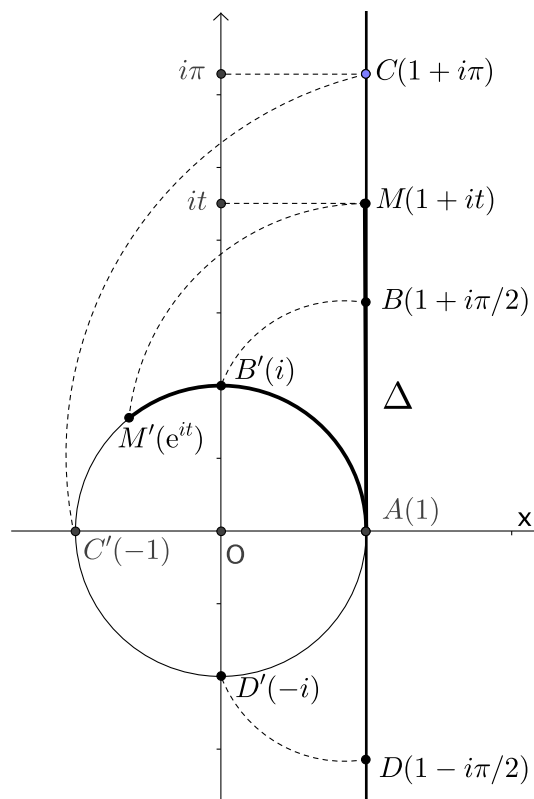
$e^{i\pi/2} = i$ (le point B vient s'enrouler en B')

$e^{-i\pi/2} = -i$ (le point D vient s'enrouler en D')

$e^{2i\pi} = 1$ (après un tour complet, le point d'ordonnée 2π sur Δ vient s'enrouler en A).

Pour tout réel t , on pose $e^{it} = \cos t + i \sin t$

Cela constitue une *définition* des fonctions \cos et \sin .



$$\forall (x, y) \in \mathbb{R}^2, e^{ix} e^{iy} = e^{i(x+y)}$$

La fonction $t \mapsto e^{it}$ vérifie la relation fondamentale :

Le figure ci-contre illustre cette propriété.

La translation de hauteur x , de A à M sur Δ , se traduit après enroulement en une rotation de centre O et d'angle x sur le cercle unité (x est en radians), et cette rotation envoie M sur M' (d'affixe e^{ix}).

Algébriquement, cette opération se traduit par la multiplication par e^{ix} (qui amène de 1 à e^{ix}).

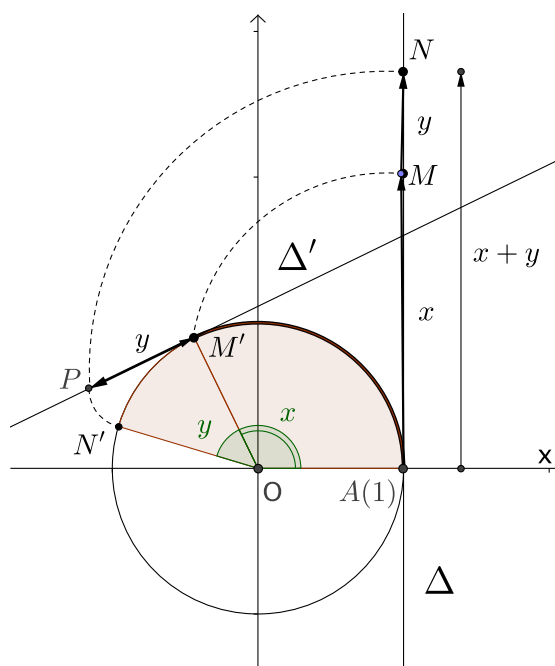
La translation de hauteur $x + y$, de A à N sur Δ , se traduit en la rotation d'angle $x + y$ sur le cercle unité (de A à N'), qui traduit algébriquement par le produit par $e^{i(x+y)}$ (qui amène de 1 à $e^{i(x+y)}$).

On peut également enrouler Δ' à partir de M' .

La translation (sur Δ') qui envoie M' sur P se traduit par une rotation d'angle y qui envoie M' sur N' (ou encore, par la multiplication par e^{iy}).

Les deux rotations, d'angle x puis y , équivalent à la seule rotation d'angle $x + y$. Les deux produits, par e^{ix} puis e^{iy} , équivalent donc au seul produit par $e^{i(x+y)}$.

L'égalité $e^{i(x+y)} = e^{ix} e^{iy}$ en est la traduction.



3.3.2 Propriétés de l'application e^{it}

Les propriétés suivantes sont immédiates et traduisent la présentation qui vient d'être donnée de $t \mapsto e^{it}$:

Proposition 3.3.1

Pour tout réel x $|e^{ix}| = 1$, c'est-à-dire : e^{ix} est élément de \mathcal{U} .

Pour tout réel x : $e^{ix} = 1$ si et seulement si x est congru à 0 modulo 2π .

En d'autres termes : $e^{ix} = 1 \Leftrightarrow x \equiv 0 [2\pi] \Leftrightarrow \exists k \in \mathbb{Z}, x = 2k\pi$.

Proposition 3.3.2 (périodicité de l'application $x \mapsto e^{ix}$)

Pour tous réels x et y , on a l'égalité fonctionnelle : $e^{ix}e^{iy} = e^{i(x+y)}$.

L'application $x \mapsto e^{ix}$ est 2π -périodique : $\forall x \in \mathbb{R}, e^{i(x+2\pi)} = e^{ix}$.

Plus précisément, pour tous réels x et y : $e^{ix} = e^{iy} \Leftrightarrow x \equiv y [2\pi] \Leftrightarrow \exists k \in \mathbb{Z}, y - x = 2k\pi$

Pour tout réel x , on a : $\frac{1}{e^{ix}} = e^{-ix} = \overline{e^{ix}}$

Valeurs particulières :

$$e^{i\pi/2} = i, \quad e^{i\pi} = -1, \quad e^{i3\pi/2} = -i, \quad e^{i\pi/4} = \frac{1}{\sqrt{2}}(1+i), \quad e^{3i\pi/4} = \frac{1}{\sqrt{2}}(-1+i), \quad e^{i2\pi/3} = j = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$$

Le résultat suivant est très important, et il sera illustré géométriquement en 3.4.2.

Proposition 3.3.3 (factorisation de $e^{ix} + 1$ et de $e^{ix} - 1$)

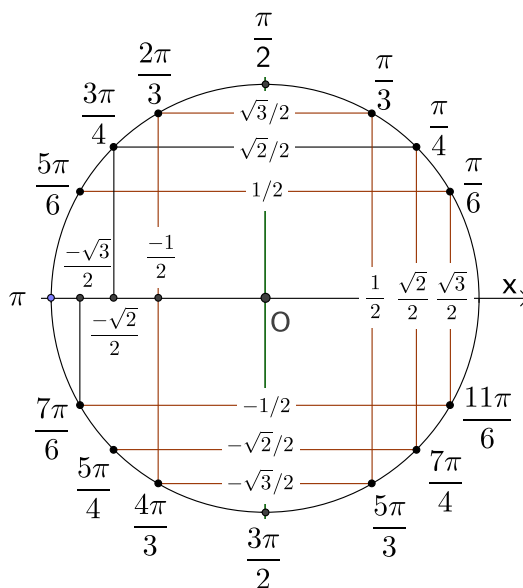
Pour tout réel x , on a : $e^{ix} + 1 = 2 \cos\left(\frac{x}{2}\right)e^{ix/2}$, et $e^{ix} - 1 = 2i \sin\left(\frac{x}{2}\right)e^{ix/2}$

3.3.3 Premières propriétés des fonctions $x \mapsto \sin x$ et $x \mapsto \cos x$

On rappelle qu'on a défini $x \mapsto \cos x$ et $x \mapsto \sin x$ par l'égalité $e^{ix} = \cos x + i \sin x$ (ça n'est pas rigoureux à 100%, mais on s'en contentera). Les propriétés des applications $x \mapsto \cos x$ et $x \mapsto \sin x$ découlent de ce qui précède (beaucoup sont assez évidentes, d'autres moins, mais on les tiendra toutes pour acquises!).

Valeurs particulières :

x	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$	$\frac{2\pi}{3}$	$\frac{3\pi}{4}$	$\frac{5\pi}{6}$	π
$\cos x$	1	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{1}{2}$	0	$-\frac{1}{2}$	$-\frac{\sqrt{2}}{2}$	$-\frac{\sqrt{3}}{2}$	-1
$\sin x$	0	$\frac{1}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{3}}{2}$	1	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{1}{2}$	0



Parité et périodicité

Les applications $x \mapsto \sin x$ et $x \mapsto \cos x$ sont définies sur \mathbb{R} , et elles sont 2π -périodiques.

L'application $x \mapsto \sin x$ est impaire et l'application $x \mapsto \cos x$ est paire.

Autrement dit, pour tout réel x :

$$\begin{cases} \cos(x + 2\pi) = \cos x \\ \sin(x + 2\pi) = \sin x \end{cases} \quad \text{et} \quad \begin{cases} \cos(-x) = \cos x \\ \sin(-x) = -\sin x \end{cases}$$

Égalités $\cos x = \cos \alpha$ et $\sin x = \sin \alpha$

Pour tout x de \mathbb{R} , on a : $\cos^2 x + \sin^2 x = 1$, $|\cos x| \leq 1$, $|\sin x| \leq 1$.

$$\forall (a, b) \in \mathbb{R}^2, a^2 + b^2 = 1 \Leftrightarrow \exists \alpha \in \mathbb{R}, \begin{cases} a = \cos \alpha \\ b = \sin \alpha \end{cases} \quad (\alpha \text{ étant unique à } 2\pi \text{ près})$$

Dans les notations suivantes, k est un entier relatif quelconque :

$$\cos x = \cos \alpha \Leftrightarrow \begin{cases} x = \alpha + 2k\pi \text{ ou} \\ x = -\alpha + 2k\pi \end{cases} \quad \sin x = \sin \alpha \Leftrightarrow \begin{cases} x = \alpha + 2k\pi \text{ ou} \\ x = \pi - \alpha + 2k\pi \end{cases}$$

On a en particulier les équivalences suivantes :

$$\begin{cases} \cos x = 0 \Leftrightarrow x = \frac{\pi}{2} + k\pi \\ \cos x = 1 \Leftrightarrow x = 2k\pi \\ \cos x = -1 \Leftrightarrow x = \pi + 2k\pi \end{cases} \quad \begin{cases} \sin x = 0 \Leftrightarrow x = k\pi \\ \sin x = 1 \Leftrightarrow x = \frac{\pi}{2} + 2k\pi \\ \sin x = -1 \Leftrightarrow x = -\frac{\pi}{2} + 2k\pi \end{cases}$$

Dérivées successives

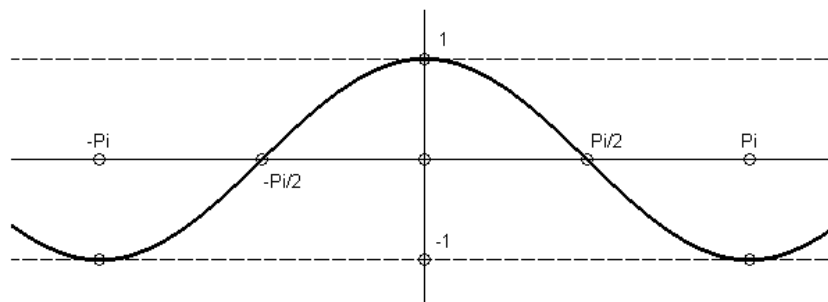
Les applications $\begin{cases} x \mapsto \sin x \\ x \mapsto \cos x \end{cases}$ sont indéfiniment dérivables sur \mathbb{R} .

Pour tout x réel, et tout n de \mathbb{N} , on a :

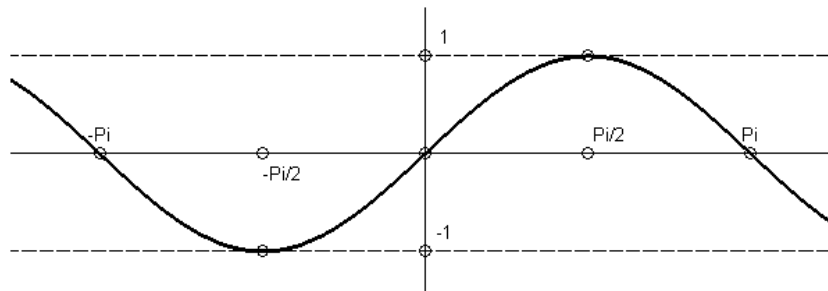
$$\begin{cases} \cos' x = -\sin x \\ \sin' x = \cos x \end{cases} \quad \begin{cases} \cos'' x = -\cos x \\ \sin'' x = -\sin x \end{cases} \quad \cos^{(n)} x = \cos\left(x + n\frac{\pi}{2}\right) \quad \text{et} \quad \sin^{(n)} x = \sin\left(x + n\frac{\pi}{2}\right)$$

Courbes représentatives

Courbe $y = \cos x$:



Courbe $y = \sin x$:

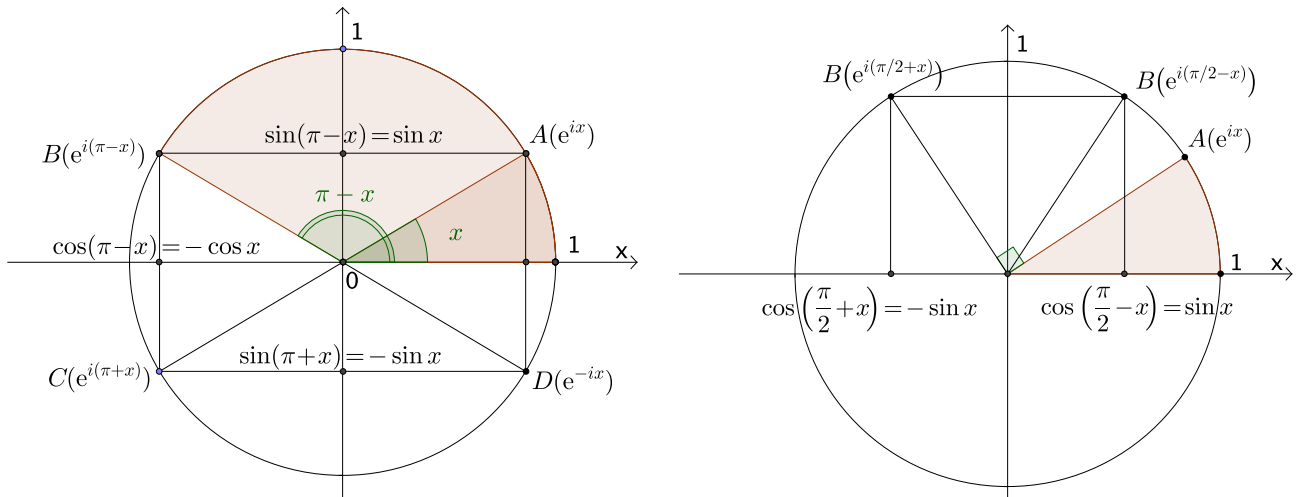


Passage de x à $\pi \pm x$ et à $\frac{\pi}{2} \pm x$

Pour tout réel x , on les égalités :

$$\begin{aligned} \sin(\pi + x) &= -\sin x & \cos(\pi + x) &= -\cos x & \sin(\pi - x) &= \sin x & \cos(\pi - x) &= -\cos x \\ \sin\left(\frac{\pi}{2} + x\right) &= \cos x & \cos\left(\frac{\pi}{2} + x\right) &= -\sin x & \sin\left(\frac{\pi}{2} - x\right) &= \cos x & \cos\left(\frac{\pi}{2} - x\right) &= \sin x \end{aligned}$$

Les relations précédentes se voient très bien sur le cercle unité (c'est comme ça qu'on les retrouve) :



Cosinus et sinus d'une somme ou d'une différence

Pour tous réels x et y :

$$\begin{cases} \cos(x + y) = \cos x \cos y - \sin x \sin y \\ \cos(x - y) = \cos x \cos y + \sin x \sin y \\ \sin(x + y) = \sin x \cos y + \cos x \sin y \\ \sin(x - y) = \sin x \cos y - \cos x \sin y \end{cases} \quad \begin{cases} \cos 2x = 2 \cos^2 x - 1 = 1 - 2 \sin^2 x \\ \sin 2x = 2 \sin x \cos x \end{cases}$$

Transformations de produits en sommes, et de sommes en produits

Pour tous réels x, y, p , et q :

$$\begin{cases} \cos x \cos y = \frac{1}{2}(\cos(x + y) + \cos(x - y)) \\ \sin x \sin y = \frac{1}{2}(\cos(x - y) - \cos(x + y)) \\ \sin x \cos y = \frac{1}{2}(\sin(x + y) + \sin(x - y)) \\ \cos^2 x = \frac{1}{2}(1 + \cos 2x) & \sin^2 x = \frac{1}{2}(1 - \cos 2x) \end{cases} \quad \begin{cases} \cos p + \cos q = 2 \cos \frac{p+q}{2} \cos \frac{p-q}{2} \\ \cos p - \cos q = -2 \sin \frac{p+q}{2} \sin \frac{p-q}{2} \\ \sin p + \sin q = 2 \sin \frac{p+q}{2} \cos \frac{p-q}{2} \\ \sin p - \sin q = 2 \sin \frac{p-q}{2} \cos \frac{p+q}{2} \end{cases}$$

3.3.4 Formules d'Euler, linéarisation

Proposition 3.3.4 (Formules d'Euler)

Pour tout réel x , on a les égalités : $\cos x = \frac{e^{ix} + e^{-ix}}{2}$, et $\sin x = \frac{e^{ix} - e^{-ix}}{2i}$

Ces formules permettent de calculer les puissances de $\cos x$ et de $\sin x$ en fonction de quantités du type $\cos(px)$ et/ou $\sin(px)$. Cette opération est appelée *linéarisation*.

Pour cela on développe $\cos^n x = \left(\frac{e^{ix} + e^{-ix}}{2}\right)^n$, $\sin^n x = \left(\frac{e^{ix} - e^{-ix}}{2i}\right)^n$, par la formule du binôme.

On termine en regroupant les termes équidistants des extrémités.

On réutilise alors les formules d'Euler pour retrouver des $\cos(px)$ et/ou des $\sin(px)$.

Quelques exemples

$$\cos^3 x = \left(\frac{e^{ix} + e^{-ix}}{2}\right)^3 = \frac{1}{8} (e^{3ix} + 3e^{ix} + 3e^{-ix} + e^{-3ix}) = \frac{1}{4} (\cos 3x + 3 \cos x)$$

$$\sin^3 x = \left(\frac{e^{ix} - e^{-ix}}{2i}\right)^3 = -\frac{1}{4} \frac{1}{2i} (e^{3ix} - 3e^{ix} + 3e^{-ix} - e^{-3ix}) = -\frac{1}{4} (\sin 3x - 3 \sin x)$$

$$\cos^4 x = \left(\frac{e^{ix} + e^{-ix}}{2}\right)^4 = \frac{1}{16} (e^{4ix} + 4e^{2ix} + 6 + 4e^{-2ix} + e^{-4ix}) = \frac{1}{8} (\cos 4x + 4 \cos 2x + 3)$$

$$\sin^4 x = \left(\frac{e^{ix} - e^{-ix}}{2i}\right)^4 = \frac{1}{16} (e^{4ix} - 4e^{2ix} + 6 - 4e^{-2ix} + e^{-4ix}) = \frac{1}{8} (\cos 4x - 4 \cos 2x + 3)$$

$$\begin{aligned} \cos^5 x &= \left(\frac{e^{ix} + e^{-ix}}{2}\right)^5 = \frac{1}{32} (e^{5ix} + 5e^{3ix} + 10e^{ix} + 10e^{-ix} + 5e^{-3ix} + e^{-5ix}) \\ &= \frac{1}{16} (\cos 5x + 5 \cos 3x + 10 \cos x) \end{aligned}$$

$$\begin{aligned} \sin^5 x &= \left(\frac{e^{ix} - e^{-ix}}{2i}\right)^5 = \frac{1}{16} \frac{1}{2i} (e^{5ix} - 5e^{3ix} + 10e^{ix} - 10e^{-ix} + 5e^{-3ix} - e^{-5ix}) \\ &= \frac{1}{16} (\sin 5x - 5 \sin 3x + 10 \sin x) \end{aligned}$$

$$\begin{aligned} \cos^6 x &= \left(\frac{e^{ix} + e^{-ix}}{2}\right)^6 = \frac{1}{64} (e^{6ix} + 6e^{4ix} + 15e^{2ix} + 20 + 15e^{-2ix} + 6e^{-4ix} + e^{-6ix}) \\ &= \frac{1}{32} (\cos 6x + 6 \cos 4x + 15 \cos 2x + 10) \end{aligned}$$

$$\begin{aligned} \sin^6 x &= \left(\frac{e^{ix} - e^{-ix}}{2i}\right)^6 = -\frac{1}{64} (e^{6ix} - 6e^{4ix} + 15e^{2ix} - 20 + 15e^{-2ix} - 6e^{-4ix} + e^{-6ix}) \\ &= -\frac{1}{32} (\cos 6x - 6 \cos 4x + 15 \cos 2x - 10) \end{aligned}$$

3.3.5 Utilisation de la formule de De Moivre

Proposition 3.3.5 (Formule de De Moivre)

Pour tout réel x , et pour tout entier n : $(e^{ix})^n = e^{inx}$.

Autrement dit : Pour tout réel x , et pour tout entier n : $(\cos x + i \sin x)^n = \cos nx + i \sin nx$.

Ce résultat permet, en développant $(\cos x + i \sin x)^n$ et en identifiant les parties réelles et imaginaires, d'exprimer $\cos nx$ et $\sin nx$ en fonction de puissances de $\cos x$ et/ou $\sin x$. Pour obtenir un résultat où figurent surtout des puissances de $\cos x$ (resp. de $\sin x$) on remplacera les puissances paires de $\sin x$ (resp. de $\cos x$) par des puissances de $(1 - \cos^2 x)$ (resp. de $(1 - \sin^2 x)$) puis de développer.

Quelques exemples

$$(\cos x + i \sin x)^3 = \cos^3 x + 3i \cos^2 x \sin x - 3 \cos x \sin^2 x - i \sin^3 x$$

$$\Rightarrow \begin{cases} \cos 3x = \cos^3 x - 3 \cos x \sin^2 x \\ \sin 3x = 3 \cos^2 x \sin x - \sin^3 x \end{cases}$$

$$\Rightarrow \begin{cases} \cos 3x = \cos^3 x - 3 \cos x(1 - \cos^2 x) \\ \sin 3x = 3(1 - \sin^2 x) \sin x - \sin^3 x \end{cases} \Rightarrow \begin{cases} \cos 3x = 4 \cos^3 x - 3 \cos x \\ \sin 3x = 3 \sin x - 4 \sin^3 x \end{cases}$$

$$(\cos x + i \sin x)^4 = \cos^4 x + 4i \cos^3 x \sin x - 6 \cos^2 x \sin^2 x - 4i \cos x \sin^3 x + \sin^4 x$$

$$\Rightarrow \begin{cases} \cos 4x = \cos^4 x - 6 \cos^2 x \sin^2 x + \sin^4 x \\ \sin 4x = 4 \cos^3 x \sin x - 4 \cos x \sin^3 x \end{cases}$$

$$\Rightarrow \begin{cases} \cos 4x = \cos^4 x - 6 \cos^2 x(1 - \cos^2 x) + (1 - \cos^2 x)^2 \\ \sin 4x = 4 \cos x((1 - \sin^2 x) \sin x - \sin^3 x) \end{cases}$$

$$\Rightarrow \begin{cases} \cos 4x = 8 \cos^4 x - 8 \cos^2 x + 1 \\ \sin 4x = 4 \cos x(-2 \sin^3 x + \sin x) \end{cases}$$

$$(\cos x + i \sin x)^5 = \cos^5 x + 5i \cos^4 x \sin x - 10 \cos^3 x \sin^2 x - 10i \cos^2 x \sin^3 x + 5 \cos x \sin^4 x + i \sin^5 x$$

$$\Rightarrow \begin{cases} \cos 5x = \cos^5 x - 10 \cos^3 x \sin^2 x + 5 \cos x \sin^4 x \\ \sin 5x = 5 \cos^4 x \sin x - 10 \cos^2 x \sin^3 x + \sin^5 x \end{cases}$$

$$\Rightarrow \begin{cases} \cos 5x = \cos^5 x - 10 \cos^3 x(1 - \cos^2 x) + 5 \cos x(1 - \cos^2 x)^2 \\ \sin 5x = 5(1 - \sin^2 x)^2 \sin x - 10(1 - \sin^2 x) \sin^3 x + \sin^5 x \end{cases}$$

$$\Rightarrow \begin{cases} \cos 5x = 16 \cos^5 x - 20 \cos^3 x + 5 \cos x \\ \sin 5x = 16 \sin^5 x - 20 \sin^3 x + 5 \sin x \end{cases}$$

Dans ce dernier cas, la formule donnant $\sin 5x$ se déduit facilement de celle donnant $\cos 5x$.

En effet, en posant $x = \frac{\pi}{2} - y$, on trouve :

$$\begin{aligned} \sin 5x &= \sin\left(\frac{5\pi}{2} - 5y\right) = \sin\left(\frac{\pi}{2} - 5y\right) \\ &= \cos 5y = 16 \cos^5 y - 20 \cos^3 y + 5 \cos y = 16 \sin^5 x - 20 \sin^3 x + 5 \sin x \end{aligned}$$

3.3.6 Deux sommes trigonométriques classiques

Pour x dans \mathbb{R} , et n dans \mathbb{N} , on pose : $C_n(x) = \sum_{k=0}^n \cos(kx)$, $S_n(x) = \sum_{k=0}^n \sin(kx)$, et $Z_n(x) = \sum_{k=0}^n e^{ikx}$.

On note bien sûr que $Z_n(x) = C_n(x) + iS_n(x)$, ou encore $C_n(x) = \operatorname{Re}(Z_n(x))$ et $S_n = \operatorname{Im}(Z_n(x))$.

D'autre part $Z_n(x) = \sum_{k=0}^n e^{ikx}$ s'écrit $Z_n(x) = \sum_{k=0}^n q^k$, avec $q = e^{ix}$ (somme géométrique).

Cas particulier : si $x \equiv 0 [2\pi]$, c'est-à-dire si $q = 1$, alors $Z_n(x) = n + 1$ donc $\begin{cases} C_n(x) = n + 1 \\ S_n(x) = 0 \end{cases}$

Cas général : on suppose $x \not\equiv 0 [2\pi]$, donc $q \neq 1$.

$$\text{Ainsi : } Z_n(x) = \frac{q^{n+1} - 1}{q - 1} = \frac{e^{i(n+1)x} - 1}{e^{ix} - 1} = \frac{e^{i(n+1)x/2} 2i \sin((n+1)x/2)}{e^{ix/2} 2i \sin(x/2)} = e^{inx/2} \frac{\sin((n+1)x/2)}{\sin(x/2)}$$

$$\text{On en tire : } C_n(x) = \frac{1}{\sin(x/2)} \left(\sin \frac{(n+1)x}{2} \cos \frac{nx}{2} \right) \text{ et } S_n(x) = \frac{1}{\sin(x/2)} \left(\sin \frac{(n+1)x}{2} \sin \frac{nx}{2} \right)$$

Autre méthode, avec une somme télescopique

$$\begin{aligned} \left(\sin \frac{x}{2} \right) C_n(x) &= \sin \frac{x}{2} \sum_{k=0}^n \cos(kx) = \sum_{k=0}^n \left(\sin \frac{x}{2} \cos(kx) \right) = \frac{1}{2} \sum_{k=0}^n \left(\sin \left(k + \frac{1}{2} \right) x - \sin \left(k - \frac{1}{2} \right) x \right) \\ &= \frac{1}{2} \sum_{k=1}^{n+1} \sin \left(k - \frac{1}{2} \right) x - \frac{1}{2} \sum_{k=0}^n \sin \left(k - \frac{1}{2} \right) x = \frac{1}{2} \left(\sin \left(n + \frac{1}{2} \right) x + \sin \frac{x}{2} \right) \end{aligned}$$

On en déduit : $\left(\sin \frac{x}{2} \right) C_n(x) = \sin \frac{(n+1)x}{2} \cos \frac{nx}{2}$, et on retrouve l'expression de $C_n(x)$.

$$\begin{aligned} \left(\sin \frac{x}{2} \right) S_n(x) &= \sin \frac{x}{2} \sum_{k=0}^n \sin(kx) = \sum_{k=0}^n \left(\sin \frac{x}{2} \sin(kx) \right) = \frac{1}{2} \sum_{k=0}^n \left(\cos \left(k - \frac{1}{2} \right) x - \cos \left(k + \frac{1}{2} \right) x \right) \\ &= \frac{1}{2} \sum_{k=0}^n \cos \left(k - \frac{1}{2} \right) x - \frac{1}{2} \sum_{k=1}^{n+1} \cos \left(k - \frac{1}{2} \right) x = \frac{1}{2} \left(\cos \frac{x}{2} - \cos \left(n + \frac{1}{2} \right) x \right) \end{aligned}$$

On en déduit : $\left(\sin \frac{x}{2} \right) S_n(x) = \sin \frac{(n+1)x}{2} \sin \frac{nx}{2}$, et on retrouve l'expression de $S_n(x)$.

3.3.7 La fonction tangente $x \mapsto \tan x$

Définition 3.3.1 (définition de la fonction tangente)

Pour tout réel x tel que $\cos x \neq 0$, c'est-à-dire tel que $x \neq \frac{\pi}{2} [\pi]$, on pose : $\tan x = \frac{\sin x}{\cos x}$.

Imparité et périodicité

Sur son domaine, l'application $x \mapsto \tan x$ est impaire et π -périodique : $\begin{cases} \tan(-x) = -\tan x \\ \tan(x + \pi) = \tan x \end{cases}$

Valeurs particulières, et égalités $\tan x = \tan \alpha$

On note les trois valeurs particulières : $\tan \frac{\pi}{6} = \frac{\sqrt{3}}{3}$, $\tan \frac{\pi}{4} = 1$, $\tan \frac{\pi}{3} = \sqrt{3}$.

Pour tout réel $\alpha \neq \frac{\pi}{2} (\pi)$, on a l'équivalence : $\tan x = \tan \alpha \Leftrightarrow x = \alpha (\pi)$.

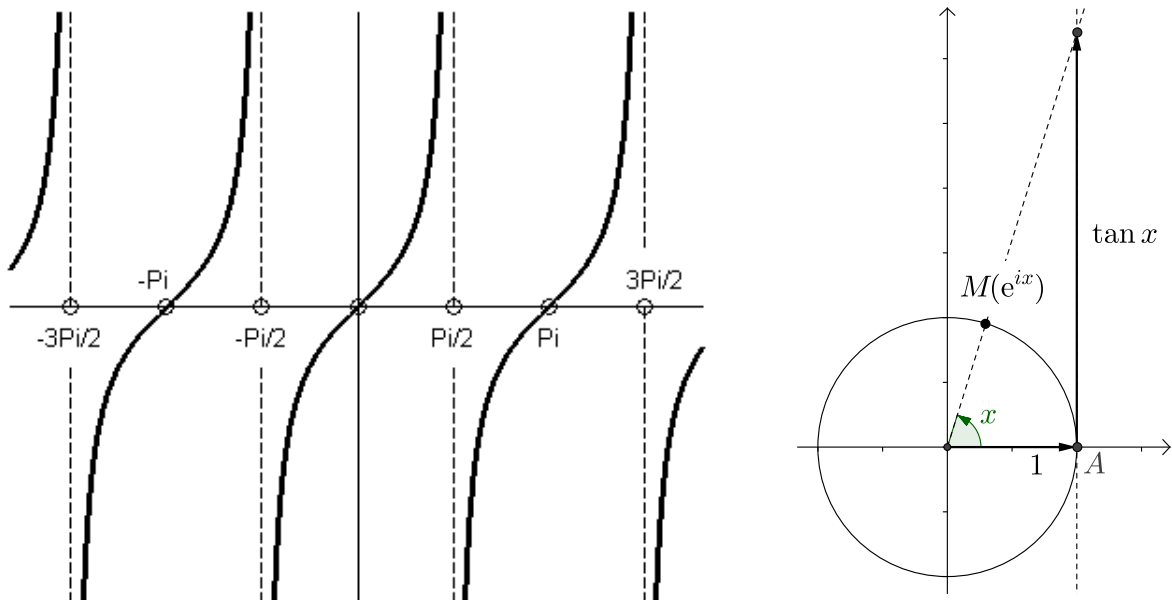
En particulier : $\tan x = 0 \Leftrightarrow x = 0 [\pi]$, $\tan x = 1 \Leftrightarrow x = \frac{\pi}{4} [\pi]$, $\tan x = -1 \Leftrightarrow x = -\frac{\pi}{4} [\pi]$

Dérivée et sens de variation de $x \mapsto \tan x$

Pour tout $x \neq \frac{\pi}{2} [\pi]$, on a : $\tan' x = 1 + \tan^2 x = \frac{1}{\cos^2 x}$.

L'application $x \mapsto \tan x$ est donc strictement croissante sur chaque intervalle de son domaine.

Plus généralement, l'application $x \mapsto \tan x$ est indéfiniment dérivable sur son domaine.

Représentation graphique, et interprétation sur le cercle unité**Passage de x à $\pi - x$ ou à $\frac{\pi}{2} \pm x$**

Les égalités suivantes se retrouvent très vite, notamment grâce à $\sin\left(x \pm \frac{\pi}{2}\right)$ et $\cos\left(x \pm \frac{\pi}{2}\right)$.

On retiendra : $\tan(\pi - x) = -\tan x$, $\tan\left(\frac{\pi}{2} + x\right) = -\frac{1}{\tan x}$, $\tan\left(\frac{\pi}{2} - x\right) = \frac{1}{\tan x}$

Tangente d'une somme ou d'une différence

On retrouve les égalités suivantes à partir de celles sur $\sin(x \pm y)$ et $\cos(x \pm y)$:

$$\tan(x + y) = \frac{\tan x + \tan y}{1 - \tan x \tan y}, \quad \tan(x - y) = \frac{\tan x - \tan y}{1 + \tan x \tan y}, \quad \tan 2x = \frac{2 \tan x}{1 - \tan^2 x}$$

Les identités relatives à la fonction tangente nécessitent une étude préalable du domaine de définition.

L'expression de $\tan(x + y)$, par exemple, suppose que x , y et $x + y$ soient différents de $\frac{\pi}{2}$ modulo π .

Utilisation du changement de variable $t = \tan \frac{x}{2}$

Le changement de variable $t = \tan \frac{x}{2}$ est utile en calcul intégral.

Par exemple, en posant $\theta = \frac{x}{2}$: $\cos x = \cos^2 \theta - \sin^2 \theta = \frac{\cos^2 \theta - \sin^2 \theta}{\cos^2 \theta + \sin^2 \theta} = \frac{1 - \tan^2 \theta}{1 + \tan^2 \theta} = \frac{1 - t^2}{1 + t^2}$

On retiendra qu'avec $t = \tan \frac{x}{2}$, on a : $\cos x = \frac{1 - t^2}{1 + t^2}$, $\sin x = \frac{2t}{1 + t^2}$, $\tan x = \frac{2t}{1 - t^2}$

3.4 Forme trigonométrique (polaire)

3.4.1 Module et argument d'un nombre complexe non nul

Proposition 3.4.1 (forme trigonométrique d'un nombre complexe non nul)

Soit z un nombre complexe non nul.

Il existe un unique réel $\rho > 0$ et une unique classe de réels θ modulo 2π , telle que $z = \rho e^{i\theta}$.

On dit que cette écriture de z est sa « forme trigonométrique », ou encore sa « forme polaire ».

Cette classe de réels modulo 2π est appelée l'argument de z .

Chacun des réels θ de cette classe est appelée une détermination de l'argument de z (ou, par abus de langage, un argument de z), et on note : $\arg z = \theta [2\pi]$.

Interprétation dans le plan complexe

L'interprétation de l'écriture $z = \rho e^{i\theta}$ est claire :

- Le réel $\rho > 0$ est le module de z
- θ est une mesure de l'angle orienté (\vec{Ox}, \vec{OM})

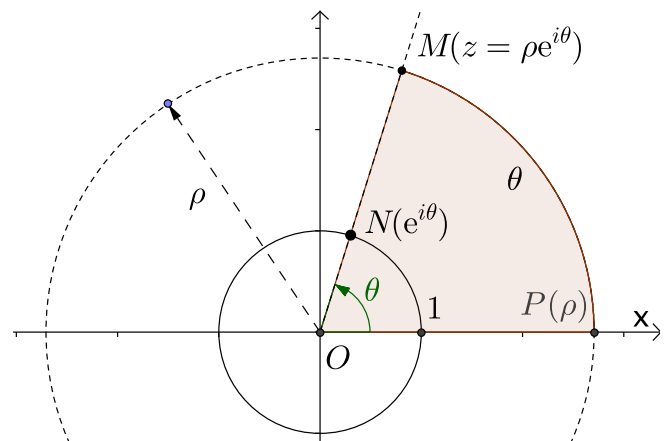
Pour tout $z \neq 0$, il y a une unique détermination de l'argument dans tout intervalle $]\alpha, \alpha + 2\pi]$, et en particulier dans $]-\pi, \pi]$ (cette dernière étant appelé *détermination principale* de l'argument de z).

On a $\rho e^{i\theta}$, avec $\rho = 0$ et pour tout réel θ .

Parler de l'argument de $z = 0$ n'a donc aucun sens.

La seule écriture $z = \rho e^{i\theta}$ ne caractérise pas la forme polaire, car il faut imposer $\rho > 0$.

Si $\rho < 0$, la forme polaire de $\rho e^{i\theta}$ est $(-\rho)e^{i(\theta+\pi)}$



Forme polaire et forme cartésienne

Soit z un nombre complexe non nul, écrit sous les deux formes $z = x + iy = \rho e^{i\theta}$ ($\rho > 0$). Alors :

Dans un sens $\begin{cases} x = \rho \cos \theta \\ y = \rho \sin \theta \end{cases}$ et dans l'autre $\begin{cases} \rho = |z| = \sqrt{x^2 + y^2} \\ \cos \theta = \frac{x}{\rho}, \quad \sin \theta = \frac{y}{\rho} \end{cases}$ (ce qui détermine ρ , et $\theta [2\pi]$)

On note que si $x \neq 0$ (donc si z n'est pas imaginaire pur), alors $\tan \theta = \frac{y}{x}$ (ce qui détermine $\theta [\pi]$)

Si z n'est pas un réel négatif, alors $\tan \frac{\theta}{2} = \frac{y}{x + \rho}$ (ce qui détermine θ modulo 2π .)

Si $z \neq 0$, mais si on n'est pas certain du signe du réel ρ :

$$z = \rho e^{i\theta} \Leftrightarrow \left(\rho = |z| \text{ et } \arg z = \theta [2\pi] \right) \text{ ou } \left(\rho = -|z| \text{ et } \arg z = \theta + \pi [2\pi] \right)$$

Cas particuliers

Soit z un nombre complexe non nul.

$$\text{On a les équivalences : } \begin{cases} z \text{ est réel} & \Leftrightarrow \arg z = 0 [\pi] \\ z \text{ est imaginaire pur} & \Leftrightarrow \arg z = \frac{\pi}{2} [\pi] \end{cases} \quad \text{et} \quad \begin{cases} z \in \mathbb{R}^{+*} & \Leftrightarrow \arg z = 0 [2\pi] \\ z \in \mathbb{R}^{-*} & \Leftrightarrow \arg z = \pi [2\pi] \end{cases}$$

3.4.2 Forme polaire et opérations dans \mathbb{C}

Proposition 3.4.2 (module et argument d'un produit)

Soient u, v deux complexes non nuls, écrits sous forme polaire $u = \rho e^{i\theta}$ et $v = r e^{i\varphi}$ ($\rho > 0, r > 0$).

$$\text{On a l'égalité : } uv = \rho r e^{i(\theta+\varphi)}, \text{ qui s'écrit aussi } \begin{cases} |uv| = |u| |v| \\ \arg(uv) = \arg u + \arg v [2\pi] \end{cases}$$

Cas particuliers

Avec les notations précédentes : $\frac{1}{u} = \frac{1}{\rho} e^{-i\theta}$, $\bar{u} = \rho e^{-i\theta}$, et $\frac{u}{v} = \frac{\rho}{r} e^{i(\theta-\varphi)}$

En termes d'arguments, on obtient donc : $\arg \frac{1}{u} = \arg \bar{u} = -\arg u [2\pi]$, et $\arg \frac{u}{v} = \arg u - \arg v [2\pi]$.

Pour tout n de \mathbb{Z} , on a $u^n = \rho^n e^{in\theta}$ et il en résulte : $\arg u^n = n \arg u [2\pi]$.

On a bien sûr : $\forall \lambda \in \mathbb{R}^{+*}, \arg \lambda u = \arg u [2\pi]$ et $\forall \lambda \in \mathbb{R}^{-*}, \arg \lambda u = \arg u + \pi [2\pi]$.

Le cas d'égalité dans l'inégalité triangulaire devient : $|u + v| = |u| + |v| \Leftrightarrow \arg u = \arg v [2\pi]$.

Illustration des factorisations de $e^{ix} \pm 1$

On rappelle que pour tout réel x , on a : $e^{ix} + 1 = 2 \cos\left(\frac{x}{2}\right) e^{ix/2}$, et $e^{ix} - 1 = 2i \sin\left(\frac{x}{2}\right) e^{ix/2}$.

Voici un schéma pour bien comprendre la signification géométrique des deux résultats précédents.

– On note M le point d'affixe e^{ix} sur le cercle unité.

On suppose ici que x est strictement compris entre 0 et π .

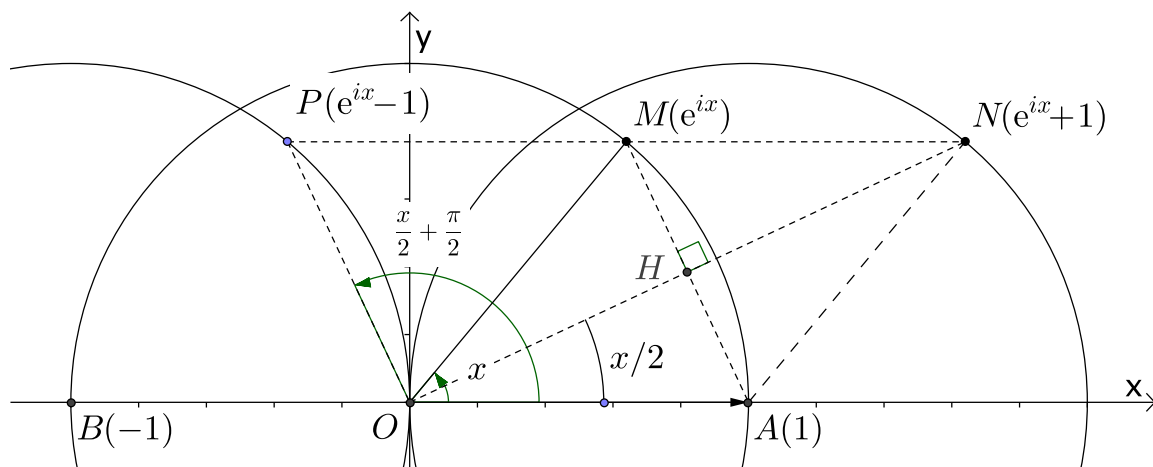
– On applique à ce cercle (donc à M) la translation de vecteur $(1, 0)$ et la translation de vecteur $(-1, 0)$.

– Quand M décrit le cercle unité, son image N dans la translation de vecteur $(1, 0)$ a pour affixe $e^{ix} + 1$ et parcourt le cercle de centre $A(1)$ et de rayon 1. L'image P de M dans la translation de vecteur $(-1, 0)$ a pour affixe $e^{ix} - 1$ et parcourt le cercle de centre $B(-1)$ et de rayon 1.

– On note que $(\widehat{Ox, \overrightarrow{ON}}) = \frac{x}{2}$ et $ON = 2OH = 2 \cos \frac{x}{2}$. Ainsi $|e^{ix} + 1| = 2 \cos \frac{x}{2}$ et $\arg(e^{ix} + 1) = \frac{x}{2}$.

De même, on voit bien que $(\widehat{Ox, \overrightarrow{OP}}) = \frac{x}{2} + \frac{\pi}{2}$ et $OP = 2HM = 2 \sin \frac{x}{2}$.

On en déduit $|e^{ix} - 1| = 2 \sin \frac{x}{2}$ et $\arg(e^{ix} - 1) = \frac{x}{2} + \frac{\pi}{2}$.



Transformation de $a \cos x + b \sin x$ en $A \cos(x - \varphi)$

On se donne (a, b) dans \mathbb{R}^2 , différent du couple nul, et x dans \mathbb{R} . Soit $A = \sqrt{a^2 + b^2}$ (donc $A > 0$).

Il existe un réel φ (défini de façon unique modulo 2π) tel que : $\frac{a}{A} = \cos \varphi$ et $\frac{b}{A} = \sin \varphi$.

Avec ces notations :

$$a \cos x + b \sin x = A \left(\frac{a}{A} \cos x + \frac{b}{A} \sin x \right) = A(\cos x \cos \varphi + \sin x \sin \varphi) = A \cos(x - \varphi)$$

Une autre approche est d'introduire $z = e^{ix} = \cos x + i \sin x$ et $\omega = a + ib = A e^{i\varphi}$.

On a en effet simultanément : $z \bar{\omega} = (\cos x + i \sin x)(a - ib)$ et $z \bar{\omega} = A e^{i(x-\varphi)}$.

On en déduit, en prenant les parties réelles : $\operatorname{Re}(z \bar{\omega}) = a \cos x + b \sin x = A \cos(x - \varphi)$.

Il y a une troisième approche, plus géométrique.

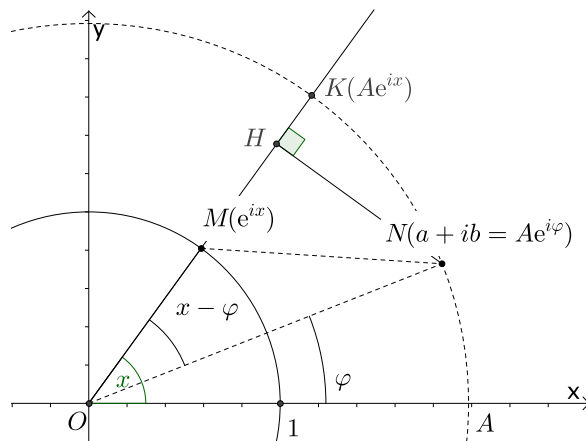
On peut en effet considérer que $a \cos x + b \sin x$ est le produit scalaire de $\overrightarrow{ON}(a, b)$ et $\overrightarrow{OM} = (\cos x, \sin x)$.

Le point N se projette en H sur le vecteur unitaire \overrightarrow{OM} .

$$\text{Ainsi } \overrightarrow{ON} \cdot \overrightarrow{OM} = \overline{OH} = \|\overrightarrow{OM}\| \|\overrightarrow{ON}\| \cos(\overrightarrow{ON}, \overrightarrow{OM})$$

c'est-à-dire : $a \cos x + b \sin x = A \cos(x - \varphi)$

Si on fixe x (c'est-à-dire M) et qu'on fait varier N sur le cercle de centre O et de rayon A , la projection H de N sur (OM) décrit le segment de centre O et dont une extrémité est le point K d'affixe Ae^{ix} .



3.4.3 Interprétation géométrique du produit

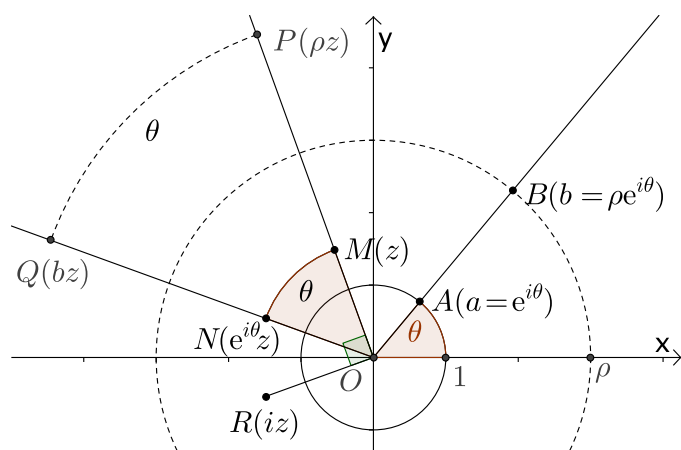
Soit $M(z)$ un point quelconque du plan, d'affixe z .

Soit $a = e^{i\theta}$ et $b = \rho e^{i\theta}$, avec $(\rho > 0)$. On définit les points $A(z)$ et $B(b)$.

On passe de $M(z)$ à $P(\rho z)$ par l'homothétie h de centre O de rapport ρ , et de $M(z)$ à $N(e^{i\theta} z)$ par la rotation r de centre O et d'angle θ .

On passe de $M(z)$ à $Q(bz)$ par la composée $f = h \circ r = r \circ h$ (on dit que f est la *similitude directe* de centre O , de rapport ρ , d'angle θ).

En particulier, $R(iz)$ se déduit de $M(z)$ par la rotation de centre O et d'angle $\frac{\pi}{2}$.



3.5 Équation du second degré dans \mathbb{C}

3.5.1 Racines carrées d'un nombre complexe

Proposition 3.5.1

Tout nombre complexe non nul Z admet exactement deux racines carrées, et elles sont opposées.

Méthode algébrique :

On pose $Z = A + iB$, et on cherche $z = x + iy$, avec A, B, x, y dans \mathbb{R} .

$$z^2 = Z \Leftrightarrow \begin{cases} x^2 - y^2 = A \\ 2xy = B \end{cases} \Leftrightarrow \begin{cases} x^2 - y^2 = A \\ x^2 + y^2 = |Z| \\ 2xy = B \end{cases} \Leftrightarrow \begin{cases} x^2 = \frac{|Z| + A}{2} \\ y^2 = \frac{|Z| - A}{2} \\ 2xy = B \end{cases}$$

$$\Leftrightarrow \begin{cases} x = \varepsilon \sqrt{\frac{|Z| + A}{2}} \\ y = \varepsilon' \sqrt{\frac{|Z| - A}{2}} \end{cases} \text{ avec } \varepsilon, \varepsilon' \in \{-1, 1\}, \text{ et } \varepsilon\varepsilon' \text{ du signe de } B$$

Par exemple :

Si $Z = 7 - 24i$ (donc $|Z| = \sqrt{49 + 576} = \sqrt{625} = 25$), et en posant $z = x + iy$:

$$z^2 = Z \Leftrightarrow \begin{cases} x^2 - y^2 = 7 \\ 2xy = -24 \end{cases} \Leftrightarrow \begin{cases} x^2 - y^2 = 7 \\ x^2 + y^2 = 25 \\ xy = -12 \end{cases} \Leftrightarrow \begin{cases} x^2 = 16 \\ y^2 = 9 \\ xy = -12 \end{cases} \Leftrightarrow \begin{cases} (x = 4 \text{ et } y = -3) \\ \text{ou } (x = -4 \text{ et } y = 3) \end{cases}$$

Les deux racines carrées de $Z = 7 - 24i$ sont donc $z_1 = 4 - 3i$ et $z_2 = -z_1 = -4 + 3i$.

Méthode trigonométrique :

On pose $Z = Re^{i\varphi}$ et $z = \rho e^{i\theta}$, avec $R > 0, \rho > 0$, et φ, θ dans \mathbb{R} .

Par exemple :

Si $Z = 2\sqrt{3} + 2i = 4e^{i\pi/6}$, et en posant $z = \rho e^{i\theta}$:

$$z^2 = Z \Leftrightarrow \rho^2 e^{2i\theta} = 4e^{i\pi/6} \Leftrightarrow \left(\rho = 2 \text{ et } \theta = \frac{\pi}{12} [\pi] \right) \Leftrightarrow \left(z = 2e^{i\pi/12} \text{ ou } z = 2e^{7i\pi/12} \right)$$

3.5.2 Équations du second degré dans \mathbb{C}

On considère l'équation $(E) : az^2 + bz + c = 0$, d'inconnue z , avec $(a, b, c) \in \mathbb{C}^3$, et $a \neq 0$.

On appelle *discriminant* de (E) le nombre complexe : $\Delta = b^2 - 4ac$.

Mise sous forme canonique :

$$(E) : az^2 + bz + c = 0 \Leftrightarrow z^2 + \frac{b}{a}z + \frac{c}{a} = 0 \Leftrightarrow \left(z + \frac{b}{2a} \right)^2 = \frac{b^2}{4a^2} - \frac{c}{a} \Leftrightarrow \left(z + \frac{b}{2a} \right)^2 = \frac{\Delta}{4a^2}$$

Cas où le discriminant est nul :

Si $\Delta = 0$, alors $(E) \Leftrightarrow \left(z + \frac{b}{2a}\right)^2 = 0 \Leftrightarrow z = -\frac{b}{2a}$ (on parle de « racine double »).

Cas où le discriminant est non nul :

Si $\Delta \neq 0$, soit δ l'une des deux racines carrées (distinctes et opposées) de Δ (ne pas noter $\sqrt{\Delta}$!!)

Alors $(E) \Leftrightarrow \left(z + \frac{b}{2a}\right)^2 - \left(\frac{\delta}{2a}\right)^2 = 0 \Leftrightarrow \left(z + \frac{b}{2a} - \frac{\delta}{2a}\right)\left(z + \frac{b}{2a} + \frac{\delta}{2a}\right) = 0$

Dans ce cas (E) a donc deux racines *distinctes* dans \mathbb{C} : $z = \frac{-b + \delta}{2a}$ et $z = \frac{-b - \delta}{2a}$.

Somme et produit des racines :

Dans tous les cas, la somme des racines est $-\frac{b}{a}$ et leur produit est $\frac{c}{a}$.

Inversement, si u, v sont cherchés dans \mathbb{C} , et si S, P sont donnés dans \mathbb{C} ,

On a l'équivalence : $\begin{cases} u + v = S \\ uv = P \end{cases} \Leftrightarrow \{u, v\}$ est l'ensemble des racines de $z^2 - Sz + P = 0$.

Utilisation du *discriminant réduit*

Si le coefficient b s'écrit manifestement $b = 2b'$, on peut utiliser le $\Delta' = b'^2 - ac$.

Les solutions de (E) s'écrivent alors : $z = \frac{-b' - \delta'}{a}$ et $z = \frac{-b' + \delta'}{a}$ où $\delta'^2 = \Delta'$.

Cas de l'équation du second degré à coefficients réels

Si (a, b, c) sont réels, le cas $\Delta \neq 0$ se subdivise en $\Delta > 0$ et $\Delta < 0$

Si $\Delta > 0$, les deux solutions de (E) sont réelles et s'écrivent : $z = \frac{-b \pm \sqrt{\Delta}}{2a}$.

Si $\Delta < 0$, elles sont non réelles, conjuguées l'une de l'autre et s'écrivent : $z = \frac{-b \pm i\sqrt{-\Delta}}{2a}$.

3.6 Racines n -ièmes

3.6.1 Racines n -ièmes de l'unité

Dans toute cette sous-section, on désigne par n un entier strictement positif.

Définition 3.6.1

On appelle racines n -ièmes de l'unité les solutions dans \mathbb{C} de l'équation $z^n = 1$.

On note \mathcal{U}_n l'ensemble des racines n -ièmes de l'unité.

Proposition 3.6.1

L'ensemble \mathcal{U}_n des racines n -ièmes de l'unité est formé de n nombres complexes distincts.

Les éléments de \mathcal{U}_n sont donnés par $\omega_k = e^{2ik\pi/n}$, avec $0 \leq k \leq n-1$.

Si on note $\omega = \omega_1 = e^{2i\pi/n}$, alors pour tout k : $\omega_k = \omega^k$ (en particulier $\omega_0 = 1$).

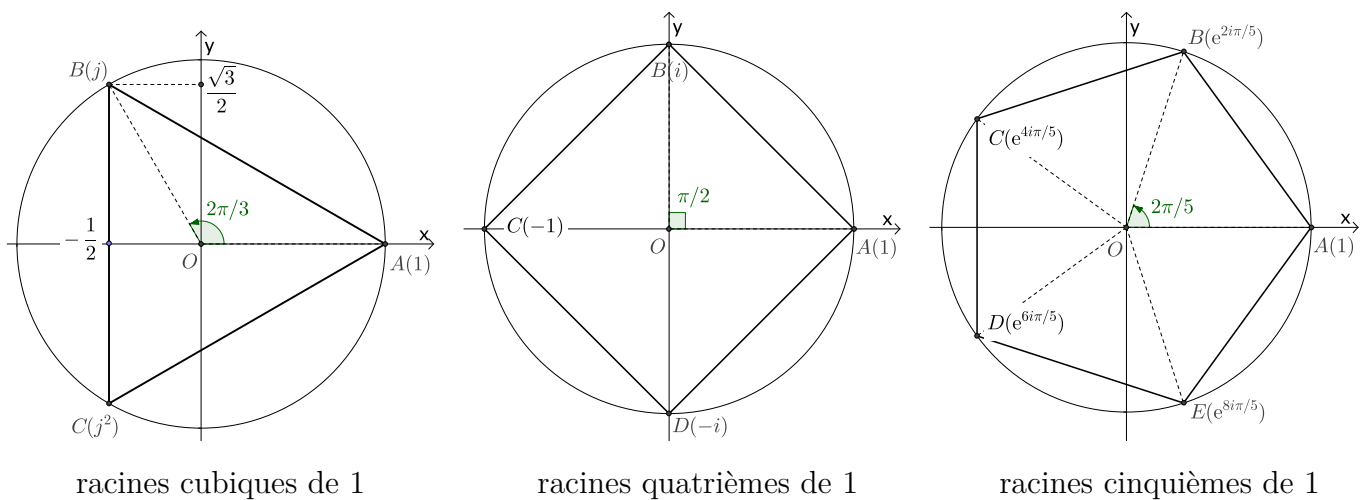
Autrement dit, avec ces notations : $\mathcal{U}_n = \{1, \omega, \omega^2, \dots, \omega^{n-1}\}$.

Cas particuliers

- La seule racine « une-ième » de l'unité est $z = 1$ (donc $\mathcal{U}_1 = \{1\}$).
- Les deux racines carrées (ou secondes, ou deuxièmes) de l'unité sont 1 et -1 (donc $\mathcal{U}_2 = \{1, -1\}$).
- Les trois racines cubiques (ou troisièmes) de l'unité sont $1, j, j^2$ avec $j = e^{2i\pi/3}$ (donc $\mathcal{U}_3 = \{1, j, j^2\}$).
- Les racines quatrièmes de l'unité sont : $1, i, -1, -i$ (donc $\mathcal{U}_4 = \{1, i, -1, -i\}$).
- Les racines cinquièmes de l'unité sont $1, \omega, \omega^2, \omega^3, \omega^4$, avec $\omega = e^{2i\pi/5}$ (donc $\mathcal{U}_5 = \{1, \omega, \omega^2, \omega^3, \omega^4\}$).
- Les racines sixièmes de l'unité sont : $1, e^{i\pi/3} = -j^2, j, -1, j^2$, et $-j$.

Proposition 3.6.2 (disposition dans le plan complexe)

Les points images des racines n -ièmes de l'unité forment les n sommets d'un polygone régulier convexe inscrit dans le cercle unité, l'un de ces sommets étant le point d'affixe 1.



L'ensemble des points-images des racines sixièmes de l'unité forme bien sûr un hexagone régulier de sommet 0, dont un sommet est $A(1)$, et on fera soi-même le dessin !

Remarques

- Le nombre $z = -1$ est une racine n -ième de l'unité si et seulement si n est pair
- Les racines n -ièmes de l'unité apparaissent dans la factorisation : $z^n - 1 = \prod_{k=0}^{n-1} (z - \omega_k)$.
- Si $n \geq 2$, la somme des racines n -ièmes de l'unité est nulle.
- Considérons l'équation $(E) : z^{n-1} + z^{n-2} + \dots + 1 = 0$, d'inconnue z dans \mathbb{C} .
Les solutions de (E) sont les $n - 1$ racines racines n -ièmes de l'unité distinctes de 1.
- En particulier : les solutions de $z^2 + z + 1 = 0$ sont j et j^2 .
Concernant le nombre $j = e^{2i\pi/3} = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$, on retiendra $j^2 + j + 1 = 0$, $j^2 = \frac{1}{j} = \bar{j}$.

Le groupe des racines n -ièmes de l'unité

Proposition 3.6.3

Si z et z' sont dans \mathcal{U}_n , il en est de même de zz' et de $\bar{z} = \frac{1}{z}$.

On exprime ces propriétés (stabilité par le produit et par le passage à l'inverse) en disant de \mathcal{U}_n qu'il est le « groupe des racines n -ièmes de l'unité ».

On sait que $\mathcal{U}_n = \{1, \omega, \omega^2, \dots, \omega^{n-1}\}$, avec $\omega = e^{2i\pi/n}$.

On exprime cette propriété en disant que ω « engendre » le groupe des racines n -ièmes de l'unité.

On peut vérifier que les autres « générateurs » de \mathcal{U}_n sont les ω_k , quand k est premier avec n .

3.6.2 Racines n -ièmes d'un nombre complexe

Définition 3.6.2

Soit Z un nombre complexe non nul, et n un entier strictement positif.

On appelle racine n -ième de Z tout nombre complexe z tel que $z^n = Z$.

Proposition 3.6.4

Soit $Z = \rho e^{i\theta}$ la forme trigonométrique de Z (avec $\rho > 0$).

Z possède n racines n -ièmes, données par : $z_k = \sqrt[n]{\rho} e^{i\theta_k}$, où $\theta_k = \frac{\theta}{n} + 2k\frac{\pi}{n}$, avec $0 \leq k \leq n-1$.

La méthode est la suivante, en cherchant z sous la forme $z = re^{i\varphi}$ ($r > 0$) :

$$z^n = Z \Leftrightarrow r^n e^{in\varphi} = \rho e^{i\theta} \Leftrightarrow \begin{cases} r^n = \rho \\ n\varphi \equiv \theta \pmod{2\pi} \end{cases} \Leftrightarrow \begin{cases} r = \sqrt[n]{\rho} \\ \varphi \equiv \frac{\theta}{n} \pmod{\frac{2\pi}{n}} \end{cases} \Leftrightarrow \begin{cases} r = \sqrt[n]{\rho} \\ \varphi = \frac{\theta + 2k\pi}{n} \\ 0 \leq k \leq n-1 \end{cases}$$

Remarques :

– Rappelons que $\mathcal{U}_n = \{1, \omega, \omega^2, \dots, \omega^{n-1}\}$, avec $\omega = e^{2i\pi/n}$, est l'ensemble des racines n -ièmes de l'unité. Il s'agit donc d'un cas particulier du problème des racines n -ièmes d'un élément Z de \mathbb{C}^* .

Mais inversement, si z_0 est une racine n -ième de Z , on les obtient toutes par $z_k = \omega_k z_0$ ($0 \leq k \leq n-1$).

On obtient donc les racines n -ièmes de Z en multipliant l'une d'elles par les racines n -ièmes de 1.

– Les points images M_k de ces n racines n -ièmes sont les sommets d'un polygone régulier convexe inscrit dans le cercle de centre O et de rayon $\rho^{1/n}$.

– Les n racines n -ièmes z_k de Z apparaissent dans la factorisation : $z^n - Z = \prod_{k=0}^{n-1} (z - z_k)$.

– La somme des n racines n -ièmes z_k de Z est nulle (si $n \geq 2$)

Proposition 3.6.5 (disposition dans le plan complexe)

Soit Z un nombre complexe non nul, et n un entier (avec $n \geq 2$). Les points images des racines n -ièmes de Z forment les n sommets d'un polygone régulier convexe de centre O .

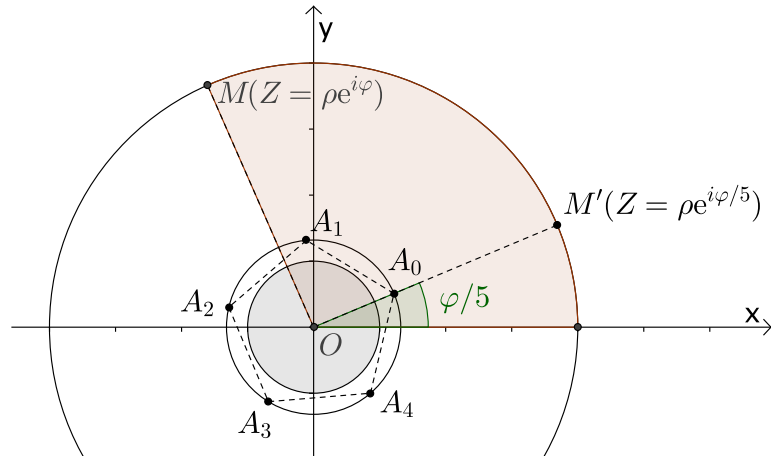
On illustre ici le cas d'un complexe Z de module $|Z| > 1$ (en fait $|Z| = 1$).

On a représenté les points A_0, \dots, A_4 images des racines cinquièmes de Z .

On a posé $\arg Z = \varphi [2\pi]$.

On a ici $(\widehat{\vec{Ox}, \vec{OA}_0}) = \frac{\varphi}{5} [2\pi]$.

Le pentagone $A_0A_1A_2A_3A_4$ se déduit, par rotation de centre O d'angle $\varphi/5$, et homothétie de centre O de rapport $\sqrt[5]{|Z|}$, du pentagone (non dessiné) des points-images des racines cinquièmes de 1.



3.6.3 Généralisation (admise) aux racines des polynômes

Dans cette sous-section, dont les résultats sont admis, on note $P(z) = a_n z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0$ une fonction polynomiale de degré $n \geq 1$ (donc ici $a_n \neq 0$), dont les coefficients a_k sont dans \mathbb{C} .

Les résultats suivants généralisent (considérablement) ce qui a été vu pour les racines n -ièmes d'un nombre complexe a (en effet, il suffit de considérer la fonction polynomiale $z \mapsto z^n - a$).

- La fonction polynomiale $P(z)$ se factorise en le produit de n fonctions polynomiales de degré 1.
- L'équation $P(z) = 0$ admet donc exactement n solutions, chacune étant répétée « autant de fois que sa multiplicité » (il peut y avoir des racines simples, des racines doubles, etc.)
- Dans le cas où les coefficients a_k sont réels, on a $P(\bar{z}) = \overline{P(z)}$ pour tout z . Il en résulte que si α est une racine de P , avec la multiplicité m , alors $\bar{\alpha}$ est également racine de P , et avec la même multiplicité.

3.7 Exponentielle complexe

3.7.1 Définition de e^z pour z dans \mathbb{C}

Définition 3.7.1

Soit $z = x + iy$ (avec x, y dans \mathbb{R}) un nombre complexe.

On pose $\exp(z) = e^x e^{iy}$, quantité également notée e^z .

On définit ainsi une application $z \mapsto \exp(z)$ de \mathbb{C} dans \mathbb{C} , appelée *exponentielle complexe*.

Premières propriétés :

- La restriction à \mathbb{R} de la fonction $z \rightarrow \exp(z)$ est l'exponentielle réelle déjà connue. De même, sa restriction aux imaginaires purs est l'application $i\theta \rightarrow e^{i\theta}$ définie précédemment.
- Pour tout complexe $z = x + iy$ (avec $x, y \in \mathbb{R}$) on a, par définition :
$$\begin{cases} |\exp(z)| = e^x \\ \arg(\exp(z)) = y [2\pi] \end{cases}$$
 En particulier, $\exp(z)$ n'est jamais nul. On note également que, pour tout z de \mathbb{C} , on a $\exp(\bar{z}) = \overline{\exp(z)}$.
- On constate l'équivalence : $\exp(z) = 1 \Leftrightarrow \exists k \in \mathbb{Z}$ tel que $z = 2ik\pi$.

3.7.2 Propriétés de la fonction exponentielle

Proposition 3.7.1 (relation fonctionnelle fondamentale)

Pour tous nombres complexes z et z' , on a l'égalité : $\exp(z + z') = \exp(z) \exp(z')$.

On en déduit que, pour tout z de \mathbb{C} : $\exp(z)$ est non nul et $\frac{1}{\exp(z)} = \exp(-z)$.

Proposition 3.7.2 (périodicité de la fonction exponentielle complexe)

On a l'équivalence : $\exp(z) = \exp(z') \Leftrightarrow (\exists k \in \mathbb{Z}, z = z' + 2ik\pi) \Leftrightarrow z \equiv z' (2i\pi)$.

L'application exponentielle $z \mapsto \exp(z)$ est donc périodique de période $2i\pi$.

Images d'une droite par la fonction exponentielle

Pour illustrer les propriétés suivantes, on est prié de faire un dessin !

- Quand $M(z = iy)$ décrit l'axe des imaginaires purs (la droite verticale $x = 0$), dans le sens des ordonnées y croissantes, le point $M(\exp(z))$ décrit le cercle trigonométrique (dans le sens trigonométrique).
- Plus généralement, quand $M(z = x_0 + iy)$ décrit la droite verticale $x = x_0$ (suivant les y croissants), le point $M(\exp(z))$ décrit le cercle de centre 0 et de rayon $R = e^{x_0}$ (dans le sens trigonométrique).
Si $x_0 < 0$ (resp $x_0 > 0$), on est à l'intérieur (resp. à l'extérieur) du cercle trigonométrique.
- Quand $M(z = x)$ décrit l'axe réel (la droite $y = 0$), dans le sens des x croissants, le point $M(\exp(z))$ décrit la demi-droite des réels strictement positifs (dans le sens des abscisses croissantes).
- Plus généralement, quand $M(z = x + iy_0)$ décrit la droite horizontale $y = y_0$ (selon les x croissants), le point $M(\exp(z))$ décrit la demi-droite issue de O et d'angle polaire y_0 (en s'éloignant de O).

3.7.3 Résolution de l'équation $e^z = a$ dans \mathbb{C}

Proposition 3.7.3

Soit $a = \rho e^{i\theta}$ un nombre complexe non nul, écrit sous forme trigonométrique ($\rho = |a| > 0$).

Soit $z = x + iy$ (avec x, y dans \mathbb{R}).

On a les équivalences : $\exp(z) = a \Leftrightarrow \begin{cases} x = \ln(\rho) \\ y \equiv \theta [2\pi] \end{cases} \Leftrightarrow \exists k \in \mathbb{Z}, z = \ln(\rho) + i(\theta + 2k\pi)$.

L'équation $\exp(z) = a$ possède donc une infinité de solutions.

Toutes se déduisent de l'une d'entre elles par ajout d'un multiple entier de $2i\pi$.

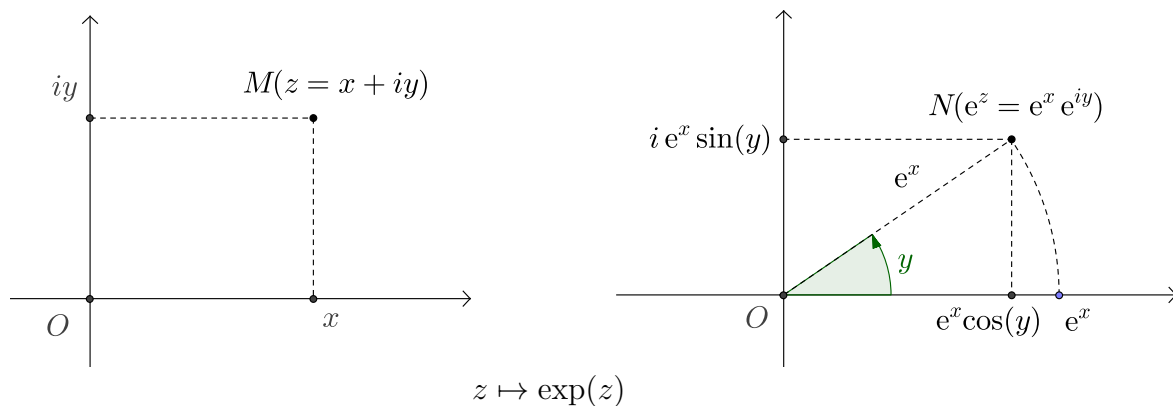
Remarques :

- L'équation $\exp(z) = a$ (a dans \mathbb{C}^* , et z cherché sous la forme $x + iy$) possède une solution unique si on se limite à y dans $] \alpha, \alpha + 2\pi]$ (par exemple y dans $] -\pi, \pi]$).
- On n'écrira **jamais** l'expression $\ln(a)$, quand a est un nombre complexe non réel strictement positif !
La raison principale est qu'il y a une infinité de solutions à l'équation $e^z = a$, et qu'on n'a pas de raison objective et sûre de choisir l'une plutôt que l'autre.
- Par exemple, on a bien $e^z = -1$ avec $z = i\pi$, mais ça n'autorise pas à écrire : $\ln(-1) = i\pi$.
En effet, les solutions de $e^z = -1$ sont les $z = (2k + 1)i\pi$, avec k dans \mathbb{Z} .

De même, on a bien $e^z = i$ avec $z = i\frac{\pi}{2}$, mais ça n'autorise pas à écrire : $\ln(i) = i\frac{\pi}{2}$.

En effet, les solutions de $e^z = -1$ sont les $z = i\frac{\pi}{2} + 2ik\pi$, avec k dans \mathbb{Z} .

– La figure suivante illustre le passage de $M(z)$ à $N(e^z)$ (on n'a pas indiqué les unités sur les axes) :



3.8 Interprétations géométriques

3.8.1 Module et argument de $(z - b)/(z - a)$

Dans cette sous-section, on désigne par A et B deux points distincts, d'affixes respectives a et b .

On note M un point quelconque du plan, distinct de A et B , d'affixe z .

On s'intéresse à l'application $\varphi : z \mapsto \varphi(z) = \frac{z - a}{z - b}$, définie sur $\mathbb{C} \setminus \{b\}$, et à valeurs dans $\mathbb{C} \setminus \{1\}$.

Par commodité, on identifie φ avec la transformation $M(z) \mapsto N(\varphi(z))$ du plan complexe.

Proposition 3.8.1

Soient A, B, M trois points distincts du plan complexe, d'affixes respectives a, b, z .

Avec ces notations : $\left| \frac{z - a}{z - b} \right| = \frac{AM}{BM}$ et $\arg\left(\frac{z - a}{z - b}\right) = (\widehat{BM, AM}) [2\pi]$

En particulier : M est aligné avec A et B si et seulement si $\frac{z - a}{z - b}$ est un nombre réel.

De même : les droites (AM) et (BM) sont orthogonales si et seulement si $\frac{z - a}{z - b}$ est imaginaire pur.

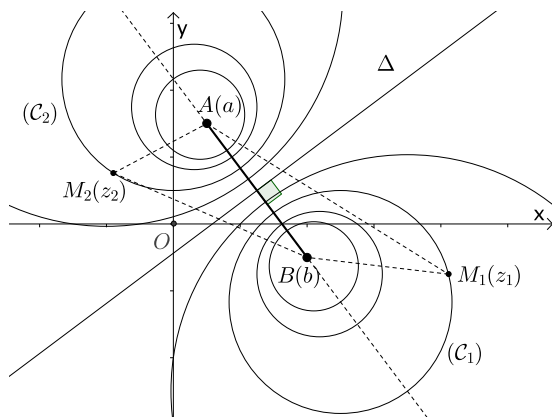
Ensemble des points $M(z)$ tels que $\left| \frac{z - a}{z - b} \right| = k > 0$

- Si $k = 1$, cet ensemble est la médiatrice Δ du segment $[A; B]$.
- Si $0 < k < 1$, c'est un cercle centré sur (AB) et contenu dans le demi-plan défini par Δ et A .
- Si $k > 1$, c'est un cercle centré sur (AB) et contenu dans le demi-plan défini par Δ et B .

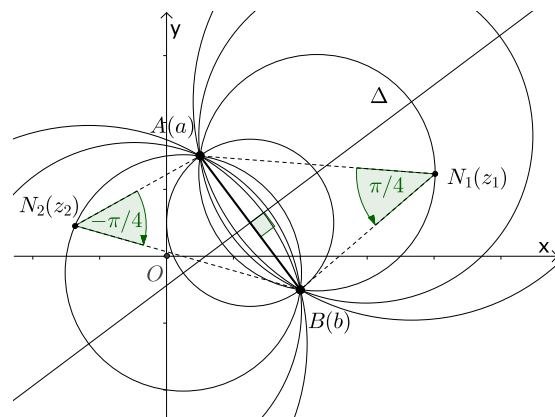
Ensemble des points $M(z)$ tels que $\arg\left(\frac{z - a}{z - b}\right) = \theta [\pi]$

- Si $\theta = 0 [\pi]$, cet ensemble est la droite (AB) (privée de A et B).
- Si $\theta \neq 0 [\pi]$, c'est un cercle centré sur la médiatrice du segment $[A; B]$, privé de A et B .

- Si $\theta = \frac{\pi}{2} [\pi]$, c'est le cercle de diamètre $[A; B]$, privé de A et B .
- Les cercles définis par θ et $-\theta$ sont symétriques l'un de l'autre par rapport à la droite (AB) .



Cercles définis par $\left| \frac{z-a}{z-b} \right| = k > 0$



Cercles définis par $\arg \left(\frac{z-a}{z-b} \right) = \theta [2\pi]$

Équation de droite, condition d'alignement

On a les équivalences :

$$\begin{aligned} M \in (AB) &\Leftrightarrow \frac{z-a}{z-b} \in \mathbb{R} \Leftrightarrow \frac{z-a}{z-b} = \frac{\bar{z}-\bar{a}}{\bar{z}-\bar{b}} \Leftrightarrow (z-a)(\bar{z}-\bar{b}) = (z-b)(\bar{z}-\bar{a}) \\ &\Leftrightarrow z(\bar{a}-\bar{b}) - \bar{z}(a-b) = a\bar{b} - b\bar{a} \end{aligned}$$

Plus symétriquement : $A(a)$, $B(b)$ et $C(c)$ sont alignés si et seulement si $a\bar{b} + b\bar{c} + c\bar{a}$ est réel.

Équation du cercle circonscrit au triangle ABC

On considère trois points distincts A, B, C , d'affixes respectives a, b, c .

Soit (Γ) le cercle circonscrit au triangle ABC .

On a les équivalences :

$$\begin{aligned} M \in (\Gamma) &\Leftrightarrow \arg \left(\frac{z-a}{z-b} \right) = \arg \left(\frac{c-a}{c-b} \right) [\pi] \Leftrightarrow \arg \frac{(z-a)(c-b)}{(z-b)(c-a)} = 0 [\pi] \\ &\Leftrightarrow \frac{(z-a)(c-b)}{(z-b)(c-a)} \in \mathbb{R} \Leftrightarrow (z-a)(\bar{z}-\bar{b})(c-b)(\bar{c}-\bar{a}) = (z-b)(\bar{z}-\bar{a})(\bar{c}-\bar{b})(c-a) \end{aligned}$$

Triangles équilatéraux

Soient $A(a)$, $B(b)$, $C(z)$ trois points du plan complexe.

On considère le triangle ABC du plan complexe. On a alors les équivalences :

$$ABC \text{ est équilatéral} \Leftrightarrow (a + jb + j^2c = 0 \text{ ou } a + jc + j^2b = 0) \Leftrightarrow a^2 + b^2 + c^2 = ab + ac + bc.$$

3.8.2 Similitudes directes

Proposition 3.8.2

Soient a et b deux nombres complexes, a étant non nul.

Soit f la transformation du plan complexe définie par $m(z) \mapsto M(Z = az + b)$.

Si $a = 1$, f est la translation dont le vecteur est le vecteur image de b .

Si $a \neq 1$, l'application f possède un point invariant unique Ω , dont l'affixe est $\omega = \frac{b}{1-a}$.

Si $a \neq 1$, l'application f est la composée commutative $f = r \circ h = h \circ r$ de :

– la rotation r de centre Ω et d'angle $\arg(a) [2\pi]$

– l'homothétie h de centre Ω et de rapport $|a|$

On dit alors que f est la similitude directe de centre Ω , de rapport $|a|$, et d'angle $\arg(a) [2\pi]$.

NB : une translation est une similitude directe de rapport 1 et d'angle 0 $[2\pi]$ (pas de centre).

Cas particuliers de similitudes directes

Soit l'application $f : m(z) \mapsto M(Z = az + b)$. On rappelle que si $a = 1$, f est une translation.

On suppose donc $a \neq 1$, et soit $\Omega(\omega)$ le point fixe de f , avec $\omega = \frac{b}{1-a}$.

– Si a est réel, alors f est l'homothétie de centre $\Omega(\omega)$ et de rapport a .

Réciproquement, si $\Omega(\omega)$ est un point du plan, et si λ est un nombre réel, l'homothétie de centre Ω et de rapport λ est définie par : $m(z) \mapsto M(Z) = \omega + \lambda(z - \omega)$.

– Si $a = e^{i\theta}$ (donc $|a| = 1$), f est la rotation de centre Ω et d'angle $\theta [2\pi]$.

Réciproquement, si $\Omega(\omega)$ est un point du plan, et si θ est un nombre réel, la rotation de centre Ω et d'angle $\theta [2\pi]$ est définie par : $m(z) \mapsto M(Z) = \omega + e^{i\theta}(z - \omega)$.

– Considérons par exemple l'application $m(z) \mapsto M(z' = f(z) = 2iz + 2 + i)$.

C'est une similitude directe du plan complexe. On a $\omega = f(\omega) \Leftrightarrow \omega = i$.

L'application f est la composée de la rotation r de centre $\Omega(i)$ et d'angle $\arg(2i) = \frac{\pi}{2} [2\pi]$ et de l'homothétie h de centre $\Omega(i)$ et de rapport $|2i| = 2$.

Multiplication des distances, conservation des angles

Soit f une similitude de rapport ρ .

Pour tous points $A' = f(A)$ et $B' = f(B)$, on a : $d(A', B') = \rho d(A, B)$.

L'application f multiplie donc les distances par le facteur $\rho = |a|$.

Pour tous points $A' = f(A)$, $B' = f(B)$, et $C' = f(C)$, on a : $(\widehat{A'B'C'}) = (\widehat{ABC}) [2\pi]$

Ainsi les similitudes directes « conservent les mesures d'angles ».

Inverse et composées de similitudes directes

L'application $f : z \mapsto Z = az + b$ est bijective, et $f^{-1}(Z) = (Z - b)/a$. L'inverse d'une similitude directe (rapport ρ , angle θ) est donc une similitude directe (rapport $1/\rho$, angle $-\theta$).

Si f et g sont deux similitudes directes de rapport ρ et ρ' (et d'angles θ, θ'), alors $g \circ f$ est une similitude directe de rapport $\rho\rho'$ et d'angle $\theta + \theta'$.

Similitude directe définie par l'image d'un segment

Proposition 3.8.3 (Similitude directe définie par l'image d'un segment)

Soient A, B, A', B' , quatre points du plan (avec $A \neq A'$ et $B \neq B'$) d'affixes respectives a, b, a', b' .

Il existe une unique similitude directe f telle que $f(A) = A'$ et $f(B) = B'$.

L'image du segment $[A; B]$ (resp. la droite (AB)) est alors le segment $[A'; B']$ (resp. la droite $(A'B')$).

L'angle de la similitude directe f est égale à $(\widehat{AB, A'B'}) [2\pi]$, c'est-à-dire $\arg \frac{b' - a'}{b - a} [2\pi]$.

Le rapport de la similitude f est $\frac{d(A', B')}{d(A, B)}$, c'est-à-dire $\left| \frac{b' - a'}{b - a} \right|$.

3.8.3 Symétries et projections orthogonales

Applications conservant l'alignement

On montre que les applications du plan complexe dans lui-même et qui « conservent l'alignement » (c'est-à-dire qui sont telles que, pour tous points alignés A, B, C les points $f(A), f(B), f(C)$ sont alignés) sont les applications $m(z) \mapsto M(Z) = uz + v\bar{z} + w$, avec (u, v, w) dans \mathbb{C}^3 .

Bien sûr les similitudes directes (et parmi elles les translations et les homothéties), c'est-à-dire les applications $f : m(z) \mapsto M(Z = uz + w)$, en font partie. Mais il y en a d'autres, comme les symétries (ou les projections) orthogonales sur (par rapport à) une droite.

Symétrie orthogonale par rapport à une droite

On sait que l'application $f : m(z) \mapsto M(Z = \bar{z})$ est la symétrie orthogonale par rapport à l'axe réel.

De même, $f : m(z) \mapsto M(Z = -\bar{z})$ est la symétrie orthogonale par rapport à l'axe des imaginaires purs.

Plus généralement, soit \mathcal{D} la droite passant par O et d'angle polaire $\theta [\pi]$.

Alors la symétrie orthogonale par rapport à \mathcal{D} est donnée par $f : m(z) \mapsto M(Z) = e^{2i\theta} \bar{z}$.

Plus généralement encore, soit \mathcal{D}' la droite passant par $\Omega(\omega)$, d'angle polaire $\theta [\pi]$.

Alors, la symétrie orthogonale par rapport à \mathcal{D}' s'écrit $z \mapsto Z = \omega + e^{2i\theta} \overline{z - \omega} = e^{2i\theta} \bar{z} + \omega - e^{2i\theta} \bar{\omega}$.

Projection orthogonale sur une droite

L'application $f : m(z) \mapsto M\left(Z = \operatorname{Re}(z) = \frac{z + \bar{z}}{2}\right)$ est la projection orthogonale sur l'axe Ox .

De même, l'application $f : m(z) \mapsto M\left(Z = \operatorname{Im}(z) = \frac{z - \bar{z}}{2i}\right)$ est la projection orthogonale sur Oy .

Plus généralement, soit \mathcal{D} la droite passant par O et d'angle polaire $\theta [\pi]$.

Alors la projection orthogonale sur \mathcal{D} est donnée par $f : m(z) \mapsto M(Z) = \frac{1}{2}(z + e^{2i\theta} \bar{z})$.

Plus généralement encore, soit \mathcal{D}' la droite passant par $\Omega(\omega)$, d'angle polaire $\theta [\pi]$.

Alors, la projection orthogonale sur \mathcal{D}' s'écrit $m(z) \mapsto (Z) = \frac{1}{2}(z + e^{2i\theta} \bar{z}) + \frac{1}{2}(\omega - e^{2i\theta} \bar{\omega})$.

Chapitre 4

Techniques d'analyse (dérivation)

Sommaire

4.1 Propriétés de la relation d'ordre dans \mathbb{R}	77
4.1.1 Existence d'une relation d'ordre total sur \mathbb{R}	77
4.1.2 Relation d'ordre et opérations dans \mathbb{R}	78
4.1.3 Valeur absolue, inégalité triangulaire	79
4.1.4 Intervalles de \mathbb{R}	80
4.1.5 Parties majorées, minorées, bornées	80
4.1.6 Borne supérieure et borne inférieure dans \mathbb{R}	82
4.1.7 Partie entière d'un nombre réel	83
4.1.8 Densité de \mathbb{Q} et de $\mathbb{R} \setminus \mathbb{Q}$ dans \mathbb{R}	84
4.1.9 Droite achevée $\overline{\mathbb{R}}$	85
4.2 Puissances à exposants entiers ou rationnels	85
4.2.1 Exposants entiers relatifs	85
4.2.2 Racine n -ième d'un réel positif	86
4.2.3 Exposants rationnels	87
4.3 Généralités sur les fonctions numériques	87
4.3.1 Représentations graphiques	88
4.3.2 Opérations sur les fonctions numériques	90
4.3.3 Fonctions paires ou impaires	91
4.3.4 Axes et centres de symétrie	93
4.3.5 Applications périodiques	93
4.3.6 Monotonie des fonctions numériques	94
4.3.7 Fonctions majorées, minorées, bornées	98
4.4 Dérivation des fonctions numériques	99
4.4.1 Notion de fonction continue	99
4.4.2 Équation de la tangente en un point	99
4.4.3 Opérations sur les fonctions dérivables	100
4.4.4 Dérivabilité et sens de variation	101
4.4.5 Dérivation de la bijection réciproque	101
4.4.6 Dérivée seconde, concavité, inflexions	102
4.4.7 Dérivées d'ordre supérieur	103
4.5 Fonctions usuelles	105
4.5.1 Exponentielle, logarithme népérien	105
4.5.2 Fonctions exponentielles de base quelconque	106

4.5.3	Fonctions puissances $x \mapsto x^\alpha$, avec α réel	107
4.5.4	Dérivée logarithmique	109
4.5.5	Fonctions circulaires réciproques	110
4.5.6	Fonctions hyperboliques	113
4.5.7	Trigonométrie hyperbolique	115
4.6	Études de fonctions, inégalités	117
4.6.1	Plan d'étude d'une fonction numérique	117
4.6.2	Dérivabilité sur le domaine de définition	117
4.6.3	Réduction du domaine d'étude	117
4.6.4	Tableau des variations	118
4.6.5	Études locales et tracé du graphe	118
4.6.6	Quelques inégalités utiles	120

4.1 Propriétés de la relation d'ordre dans \mathbb{R}

4.1.1 Existence d'une relation d'ordre total sur \mathbb{R}

Proposition 4.1.1 (propriétés de la relation d'ordre sur \mathbb{R})

L'ensemble \mathbb{R} est muni d'une relation d'ordre notée \leq , et vérifiant les propriétés :

- ordre total : pour tous x, y de \mathbb{R} , on a : $x \leq y$ ou $y \leq x$.
- compatibilité avec l'addition : $\forall (x, y, z) \in \mathbb{R}^3$, $x \leq y \Rightarrow x + z \leq y + z$.
- compatibilité avec le produit par un réel positif : $\forall (x, y, z) \in \mathbb{R}^3$, $(x \leq y \text{ et } 0 \leq z) \Rightarrow xz \leq yz$.

Relation d'ordre contraire, inégalités strictes

- On a bien sûr la relation d'ordre contraire de la précédente, définie par $x \geq y$ (qui équivaut à $y \leq x$). On utilise plus souvent \leq que \geq dans les calculs, mais essentiellement \leq dans les définitions et propriétés (sachant qu'à toute propriété relative à \leq correspondant une propriété relative à \geq).

- On définit les inégalités strictes :
$$\begin{cases} x < y & \text{équivaut à } (x \leq y \text{ et } x \neq y) \\ x > y & \text{équivaut à } y < x \end{cases}$$

On rappelle que les relations $<$ et $>$ ne définissent pas à proprement parler des relations d'ordre, car elles ne sont pas *réflexives* : on n'a en effet pas l'inégalité $x < x$.

Cas des ensembles \mathbb{N} , \mathbb{Z} et \mathbb{Q}

On peut *restreindre* la relation d'ordre de \mathbb{R} à chacun des ensembles \mathbb{N} , \mathbb{Z} , ou \mathbb{Q} (et considérer alors qu'il s'agit d'une relation d'ordre total sur chacun de ces trois ensembles).

Pour ce qui est de \mathbb{N} et \mathbb{Z} , on dispose de propriétés importantes :

- Toute partie non vide de \mathbb{N} possède un élément minimum (en particulier 0 est le minimum de \mathbb{N}).
- Toute partie minorée non vide de \mathbb{Z} possède un élément minimum.
- Toute partie non vide majorée de \mathbb{N} , ou de \mathbb{Z} , possède un élément maximum.
- Si x et y sont dans \mathbb{Z} , on a l'équivalence : $x \leq y \Leftrightarrow (\exists n \in \mathbb{N}, y = x + n)$.

Les propriétés précédentes sont fausses pour l'ensemble \mathbb{Q} .

Définition 4.1.1 (notations $\mathbb{R}^{+*}, \mathbb{R}^{-*}$ etc.)

On pose $\mathbb{R}^{+*} = \{x \in \mathbb{R}, x > 0\}$ et $\mathbb{R}^+ = \mathbb{R}^{+*} \cup \{0\} = \{x \in \mathbb{R}, x \geq 0\}$.

On définit de même les ensembles $\mathbb{Z}^{+*}, \mathbb{Z}^+, \mathbb{Q}^{+*}$, et \mathbb{Q}^+ .

On pose $\mathbb{R}^{-*} = \{x \in \mathbb{R}, x < 0\}$ et $\mathbb{R}^- = \mathbb{R}^{-*} \cup \{0\} = \{x \in \mathbb{R}, x \leq 0\}$.

On définit de même les ensembles $\mathbb{Z}^{-*}, \mathbb{Z}^-, \mathbb{Q}^{-*}$, et \mathbb{Q}^- .

Ne pas généraliser aux nombres complexes !

Soit x un réel : si $x \geq 0$ alors $x^2 \geq 0$ (comptabilité pour le produit par un réel positif).

Mais si $x \leq 0$, alors $-x \geq 0$ et là encore on a $(-x)^2 \geq 0$ c'est-à-dire $x^2 \geq 0$.

Retenons donc que pour tout x réel, on a $x^2 \geq 0$ (et bien sûr $x^2 > 0$ pour tout réel x non nul).

De cette remarque très simple, il découle qu'il n'existe pas de relation d'ordre dans \mathbb{C} qui puisse étendre celle de \mathbb{R} (en gardant notamment la comptabilité pour le produit par un nombre positif). En effet, si tel était le cas, on trouverait $i^2 > 0$, c'est-à-dire $-1 > 0$.

Il existe des relations d'ordre total sur \mathbb{C} , mais qui ne généralisent pas les propriétés de celle de \mathbb{R} .

Pour éviter les ennuis, on évitera donc d'écrire $z \leq z'$, quand z et z' sont dans \mathbb{C} .

4.1.2 Relation d'ordre et opérations dans \mathbb{R}

Signe de la somme, du produit

Le tableau ci-après résume les *règles des signes*, c'est-à-dire ce qu'on peut affirmer sur $x + y$ ou sur xy quand on connaît la position de x et de y par rapport à 0 (les points d'interrogation signifient qu'il n'y a pas de réponse générale dans tel ou tel cas).

x	≥ 0	≤ 0	≥ 0	> 0	< 0	> 0	> 0	> 0	< 0	< 0
y	≥ 0	≤ 0	≤ 0	> 0	< 0	< 0	≥ 0	≤ 0	≥ 0	≤ 0
$x + y$	≥ 0	≤ 0	?	> 0	< 0	?	> 0	?	?	< 0
xy	≥ 0	≥ 0	≤ 0	> 0	> 0	< 0	≥ 0	≤ 0	≤ 0	≥ 0

On démontre également les équivalences suivantes, valables pour tous réels x, y, z :

$$\left\{ \begin{array}{l} x + z \leq y + z \Leftrightarrow x \leq y \\ x \leq y \Leftrightarrow -y \leq -x \\ x \leq 0 \Leftrightarrow -x \geq 0 \end{array} \right. \quad \left\{ \begin{array}{l} x + z < y + z \Leftrightarrow x < y \\ x < y \Leftrightarrow -y < -x \\ x < 0 \Leftrightarrow -x > 0 \end{array} \right. \quad \left\{ \begin{array}{l} (x \leq y \text{ et } z \leq 0) \Rightarrow xz \geq yz \\ (x < y \text{ et } z > 0) \Rightarrow xz < yz \\ (x < y \text{ et } z < 0) \Rightarrow xz > yz \end{array} \right.$$

Relation d'ordre et passage à l'inverse

Pour tout réel x non nul, on a les équivalences évidentes : $x > 0 \Leftrightarrow \frac{1}{x} > 0$, et $x < 0 \Leftrightarrow \frac{1}{x} < 0$

Pour tout réel x non nul, on a aussi les équivalences suivantes :

$x > 1 \Leftrightarrow 0 < \frac{1}{x} < 1$	$x < -1 \Leftrightarrow -1 < \frac{1}{x} < 0$	$0 < x < 1 \Leftrightarrow \frac{1}{x} > 1$	$-1 < x < 0 \Leftrightarrow \frac{1}{x} < -1$
---	---	---	---

Plus généralement, pour tous réels x et y , on a :

$0 < x < y \Leftrightarrow 0 < \frac{1}{y} < \frac{1}{x}$	$x < y < 0 \Leftrightarrow \frac{1}{y} < \frac{1}{x} < 0$	$x < 0 < y \Leftrightarrow \frac{1}{x} < 0 < \frac{1}{y}$
---	---	---

4.1.3 Valeur absolue, inégalité triangulaire

Définition 4.1.2

Pour tout réel x , on pose $|x| = \max(-x, x)$. Cette quantité est appelée *valeur absolue* de x .

On vérifie immédiatement les propriétés suivantes :

- Pour tout réel x :
$$\begin{cases} |x| \geq 0, & |x| = 0 \Leftrightarrow x = 0 \\ |x| = x \Leftrightarrow x \geq 0, & |x| = -x \Leftrightarrow x \leq 0 \end{cases}$$
- Plus généralement, pour tout réel x , et pour tout réel positif ou nul α :
$$|x| = \alpha \Leftrightarrow x \in \{-\alpha, \alpha\}, \quad \begin{cases} |x| \leq \alpha \Leftrightarrow -\alpha \leq x \leq \alpha \\ |x| < \alpha \Leftrightarrow -\alpha < x < \alpha \end{cases} \quad \begin{cases} |x| \geq \alpha \Leftrightarrow (x \leq -\alpha \text{ ou } x \geq \alpha) \\ |x| > \alpha \Leftrightarrow (x < -\alpha \text{ ou } x > \alpha) \end{cases}$$
- Pour tous réels x et y :
$$\begin{cases} x^2 = y^2 \Leftrightarrow |x| = |y| \\ x^2 \leq y^2 \Leftrightarrow |x| \leq |y| \end{cases} \quad |xy| = |x||y| \quad (\text{et } |x^n| = |x|^n \text{ pour tout entier } n)$$
- Pour tous réels x et y , on a $\max(x, y) = \frac{1}{2}(x + y + |x - y|)$ et $\min(x, y) = \frac{1}{2}(x + y - |x - y|)$

Proposition 4.1.2 (inégalité triangulaire)

Pour tous réels x et y , on a : $|x + y| \leq |x| + |y|$.

On a l'égalité $|x + y| = |x| + |y|$ si et seulement si x et y ont le même signe.

Proposition 4.1.3 (valeur absolue d'une somme, ou d'un produit, de n nombres réels)

Pour tous réels x_1, x_2, \dots, x_n , on a : $\left| \prod_{k=1}^n x_k \right| = \prod_{k=1}^n |x_k|$ et $\left| \sum_{k=1}^n x_k \right| \leq \sum_{k=1}^n |x_k|$

On a l'égalité $\left| \sum_{k=1}^n x_k \right| = \sum_{k=1}^n |x_k|$ si et seulement si tous les x_k ont le même signe.

Définition 4.1.3

Pour tout réel x , on note $\begin{cases} x^+ = \max(x, 0) \\ x^- = \max(-x, 0) \end{cases}$ Ainsi : $x^+ = \begin{cases} x & \text{si } x \geq 0 \\ 0 & \text{sinon} \end{cases}$ et $x^- = \begin{cases} -x & \text{si } x \leq 0 \\ 0 & \text{sinon} \end{cases}$

On dit que x^+ (resp. x^-) est la partie positive (resp. négative) du réel x .

Avec ces notations, pour tout réel x :
$$\begin{cases} x^+ \geq 0 \\ x^- \geq 0 \end{cases} \quad \begin{cases} x = x^+ - x^- \\ |x| = x^+ + x^- \end{cases} \quad \begin{cases} x^+ = (x + |x|)/2 \\ x^- = (|x| - x)/2 \end{cases}$$

Définition 4.1.4

Pour tous réels x, y , la quantité $d(x, y) = |y - x|$ est appelée *distance* de x et de y .

Elle vérifie : $\forall (x, y, z) \in \mathbb{R}^3, \begin{cases} d(x, y) \geq 0, & d(x, y) = 0 \Leftrightarrow x = y \\ d(x, y) \leq d(x, z) + d(z, y) \end{cases}$

Remarques

- Pour tous réels x et y , on a $d(|x|, |y|) \leq d(x, y)$, c'est-à-dire $||x| - |y|| \leq |x - y|$.
Ce résultat complète donc l'inégalité triangulaire.
- Si a est dans \mathbb{R} et b dans \mathbb{R}^+ , on a l'équivalence $|x - a| \leq b \Leftrightarrow a - b \leq x \leq a + b$.

4.1.4 Intervalles de \mathbb{R}

Définition 4.1.5

Pour tous réels a et b , on définit les ensembles suivants, dits *intervalles* de \mathbb{R} .

$$\left\{ \begin{array}{ll} [a, b] = \{x \in \mathbb{R}, a \leq x \leq b\} &]a, b[= \{x \in \mathbb{R}, a < x < b\} \\]a, b] = \{x \in \mathbb{R}, a < x \leq b\} &]a, b[= \{x \in \mathbb{R}, a < x < b\} \\]-\infty, +\infty[= \mathbb{R} & \\ [a, +\infty[= \{x \in \mathbb{R}, a \leq x\} &]a, +\infty[= \{x \in \mathbb{R}, a < x\} \\]-\infty, b] = \{x \in \mathbb{R}, x \leq b\} &]-\infty, b[= \{x \in \mathbb{R}, x < b\} \end{array} \right.$$

En particulier : $\mathbb{R}^+ = [0, +\infty[$, $\mathbb{R}^{+*} =]0, +\infty[$, $\mathbb{R}^- =]-\infty, 0]$, $\mathbb{R}^{-*} =]-\infty, 0[$

Remarques et définitions

- On dit que $[a, b]$ (avec $a \leq b$) est le *segment* d'origine a et d'extrémité b .
- Les intervalles $]a, b[$, $]a, +\infty[$, $] -\infty, b[$ et $] -\infty, +\infty[$ sont dits *ouverts*.
- Les intervalles $[a, b]$, $[a, +\infty[$, $] -\infty, b]$ et $] -\infty, +\infty[$ sont dits *fermés*.
- Les intervalles $]a, b]$ et $[a, b[$ sont dits *semi-ouverts* (ou *semi-fermés* !).
- Le segment $[a, a]$ se réduit à $\{a\}$; L'intervalle $]a, a[$ est vide.
- Les segments sont les intervalles fermés bornés.

La proposition suivante est admise pour l'instant. Elle sera démontrée en 4.1.6

Proposition 4.1.4 (caractérisation des intervalles de \mathbb{R})

Soit I une partie non vide de \mathbb{R} . L'ensemble I est un intervalle si et seulement si I est convexe c'est-à-dire : $\forall (x, y) \in I \times I, [x, y] \subset I$

Autrement dit : I est un intervalle si et seulement si, dès que I contient deux réels x et y , alors I contient le segment qui les joint.

4.1.5 Parties majorées, minorées, bornées

Dans les définitions suivantes, A désigne une partie de \mathbb{R} (éventuellement vide).

Définition 4.1.6 (partie majorée)

On dit qu'un réel M *majoré* A (ou est un *majorant* de A) si $M \geq a$ pour tout a de A .

On dit que A est *majorée* si l'ensemble de ses majorants est non vide.

Définition 4.1.7 (partie minorée)

On dit qu'un réel m *mineur* A (ou est un *minorant* de A) si $m \leq a$ pour tout a de A .

On dit que A est *minorée* si l'ensemble de ses minorants est non vide.

Définition 4.1.8 (partie bornée)

On dit que A est *bornée* si A est à la fois minorée et majorée. Ainsi A est bornée si et seulement si il existe m et M dans \mathbb{R} tels que $m \leq a \leq M$ pour tout a de A .

Remarques

- La partie vide est bornée ! Toute partie finie de \mathbb{R} est bornée.
- Dire que A est majorée, c'est dire qu'il existe M dans \mathbb{R} tel que : $A \subset]-\infty, M]$.
Dire que A est minorée, c'est dire qu'il existe m dans \mathbb{R} tel que : $A \subset [m, +\infty[$.
Dire que A est bornée, c'est dire qu'il existe m et M dans \mathbb{R} tels que : $A \subset [m, M]$.
- Si M est un majorant de A , alors tout réel $M' \geq M$ est a fortiori un majorant de A .
Si m est un minorant de A , alors tout réel $m' \leq m$ est a fortiori un minorant de A .
- Parmi les intervalles de \mathbb{R} , seuls ceux qui sont de la forme $[a, b]$, $]a, b]$, $[a, b[$ ou $]a, b[$ sont bornés.
- Dire que A est bornée, c'est dire qu'il existe M dans \mathbb{R} tel que : $\forall x \in A, |x| \leq M$. On exprime cette propriété en disant que A est bornée si et seulement si elle est « majorée en valeur absolue ».
- Posons $-A = \{-a, a \in A\}$.
Avec ces notations : A est majorée (resp. minorée) si et seulement si $-A$ est minorée (resp. majorée).

Définition 4.1.9 (élément maximum)

Soit α un élément de A .

On dit que α est un *maximum* (ou un *plus grand élément*) de A si $\alpha \geq x$ pour tout x de A .

Si un tel élément existe, il est unique. On le note alors $\alpha = \max(A)$.

Définition 4.1.10 (élément minimum)

Soit α un élément de A .

On dit que α est un *minimum* (ou un *plus petit élément*) de A si $\alpha \leq x$ pour tout x de A .

Si un tel élément existe, il est unique. On le note alors $\alpha = \min(A)$.

Remarques et exemples

- Dire que A possède un maximum, c'est dire que l'un des éléments de A est un majorant de A .
Dire que A possède un minimum, c'est dire que l'un des éléments de A est un minorant de A .
- L'intervalle $A =]-\infty, 1[$ est majoré, mais non minoré (donc non borné).
L'ensemble de ses majorants est $[1, +\infty[$. Le maximum de A n'existe pas.
- L'intervalle $B = [-2, +\infty[$ est minoré mais non majoré.
L'ensemble de ses minorants est $] -\infty, -2]$, et on a $\min(B) = -2$.
- L'intervalle $C =]0, 1[$ est borné, mais $\min(C)$ et $\max(C)$ n'existent pas.
- L'intervalle $D = [0, 1]$ est borné. De plus $\min(D) = 0$ et $\max(D) = 1$.
- Si $A = \left\{ \frac{n}{n+1}, n \in \mathbb{N} \right\}$, $\min A = 0$, et $\max A$ n'existe pas. L'ensemble des majorants est $[1, +\infty[$.

4.1.6 Borne supérieure et borne inférieure dans \mathbb{R}

L'axiome suivant, très important, constitue une spécificité de \mathbb{R} par rapport à \mathbb{Q} .

Axiome (axiome de la borne supérieure)

Soit A une partie non vide et majorée de \mathbb{R} .

Alors l'ensemble des majorants de A possède un élément minimum.

Cet élément est appelé borne supérieure de A , et on le note $\sup(A)$.

Remarques

Soit A une partie non vide et majorée de \mathbb{R} .

– Le réel $\beta = \sup(A)$ est **caractérisé** par les deux propriétés $\begin{cases} \forall x \in A, x \leq \beta \\ \forall \varepsilon > 0, \exists a \in A, \beta - \varepsilon < a \leq \beta \end{cases}$

Autrement dit, ce qui caractérise β c'est que :

$\begin{cases} \text{d'une part, } \beta \text{ est un majorant de } A. \\ \text{d'autre part, tout réel strictement inférieur à } \beta \text{ n'est pas un majorant de } A. \end{cases}$

– L'ensemble des majorants de A est alors l'intervalle $[\beta, +\infty[$.

– On retiendra donc : *toute partie non vide majorée de \mathbb{R} possède une borne supérieure.*

Et d'une façon plus familière : *la borne supérieure de A , c'est le plus petit des majorants des A .*

L'axiome de la borne supérieure étant admis, on peut démontrer le résultat suivant :

Proposition 4.1.5 (propriété de la borne inférieure)

Soit A une partie non vide et minorée de \mathbb{R} .

Alors l'ensemble des minorants de A possède un élément maximum.

Cet élément est appelé borne inférieure de A , et on le note $\inf(A)$.

Remarques

Soit A une partie non vide et minorée de \mathbb{R} .

– Le réel $\alpha = \inf(A)$ est **caractérisé** par les deux propriétés $\begin{cases} \forall x \in A, x \geq \alpha \\ \forall \varepsilon > 0, \exists a \in A, \alpha \leq a < \alpha + \varepsilon \end{cases}$

Autrement dit, ce qui caractérise α c'est que :

$\begin{cases} \text{d'une part, } \alpha \text{ est un minorant de } A. \\ \text{d'autre part, tout réel strictement supérieur à } \alpha \text{ n'est pas un minorant de } A. \end{cases}$

– L'ensemble des minorants de A est alors l'intervalle $] -\infty, \alpha]$.

– On retiendra donc : *toute partie non vide minorée de \mathbb{R} possède une borne inférieure.*

Et d'une façon plus familière : *la borne inférieure de A , c'est le plus grand des minorants de A .*

Quelques propriétés de la borne Sup et la borne Inf

– Soit A une partie non vide et majorée de \mathbb{R} (donc telle que $\sup(A)$ existe).

Dire que x majore A (donc écrire que $x \geq a$ pour tout a de A) c'est dire que $x \geq \sup(A)$.

Dire que $\max(A)$ existe, c'est dire que $\sup(A)$ est élément de A . Dans ce cas, $\sup(A) = \max(A)$.

- Soit A une partie non vide et minorée de \mathbb{R} (donc telle que $\inf(A)$ existe).
Dire que x minore A (donc écrire que $x \leq a$ pour tout a de A) équivaut à écrire $x \leq \inf(A)$.
Dire que $\min(A)$ existe, c'est dire que $\inf(A)$ est élément de A . Dans ce cas, $\inf(A) = \min(A)$.
- Soit A, B deux parties non vides de \mathbb{R} .
Si B est majorée et si $A \subset B$, alors A est majorée et $\sup(A) \leq \sup(B)$.
Si B est minorée et si $A \subset B$, alors A est minorée et $\inf(B) \leq \inf(A)$.
- Soit A une partie non vide de \mathbb{R} . On rappelle la notation $-A = \{-a, a \in A\}$.
Si A est majorée, alors $-A$ est minorée et : $\inf(-A) = -\sup(A)$.
Si A est minorée, alors $-A$ est majorée et : $\sup(-A) = -\inf(A)$.
- Soit A, B deux parties non vides de \mathbb{R} . On rappelle la notation $A + B = \{a + b, a \in A, b \in B\}$.
Si A et B sont majorées, alors $A + B$ est majorée et : $\sup(A + B) = \sup(A) + \sup(B)$.
Si A et B sont minorées, alors $A + B$ est minorée et : $\inf(A + B) = \inf(A) + \inf(B)$.
- Enfin les résultats suivants sont évidents, pour tous réels a et b , avec $a < b$:

$$\begin{cases} \sup([a, b]) = \sup([a, b[) = \sup(]a, b]) = \sup(]a, b[) = \sup(]-\infty, b]) = \sup(]-\infty, b[) = b \\ \inf([a, b]) = \inf([a, b[) = \inf(]a, b]) = \inf(]a, b[) = \inf([a, +\infty[) = \inf(]a, +\infty[) = a \end{cases}$$

Caractérisation des intervalles

La propriété suivante est une caractérisation commode des intervalles de \mathbb{R} (ce n'en est pas une définition, car l'énoncé suppose connue celle des segments $[x, y]$).

Proposition 4.1.6 (caractérisation des intervalles)

Soit I une partie de \mathbb{R} (éventuellement vide).

Alors I est un intervalle si et seulement si : $\forall (x, y) \in I \times I, [x, y] \subset I$.

Autrement dit, les intervalles sont les parties de \mathbb{R} qui, dès qu'elles contiennent deux points, contiennent le segment qui les joint. On l'exprime en disant que les intervalles sont les parties convexes de \mathbb{R} .

4.1.7 Partie entière d'un nombre réel

On commence par démontrer un résultat qui semble évident, mais qui est une conséquence de l'axiome de la borne supérieure.

Proposition 4.1.7 (\mathbb{R} est archimédien)

Soit x un réel, et a un réel strictement positif. Alors il existe un entier n tel que $na > x$.

On exprime cette propriété en disant que \mathbb{R} est archimédien.

Proposition 4.1.8

Soit x un réel, et a un réel strictement positif.

Alors il existe un couple unique (n, y) de $\mathbb{Z} \times [0, a[$ tel que $x = na + y$.

Il suffit d'appliquer le résultat précédent avec $a = 1$ pour obtenir la notion de partie entière.

Définition 4.1.11 (partie entière)

Soit x un réel. Il existe un entier relatif unique m tel que $m \leq x < m + 1$.

On l'appelle *partie entière* de x et on le note $\lfloor x \rfloor$.

Quelques propriétés de la partie entière

- Pour tout réel x , on a : $\lfloor x \rfloor \leq x < \lfloor x \rfloor + 1$, ou encore : $x - 1 < \lfloor x \rfloor \leq x$
- Pour tout x de \mathbb{R} , et tout m de \mathbb{Z} , on a : $\lfloor x \rfloor = m \Leftrightarrow x \in [m, m + 1[$
- Pour tout x de \mathbb{R} , on a : $\lfloor x \rfloor = x \Leftrightarrow x \in \mathbb{Z}$
- Si x est réel non entier, alors : $\lfloor -x \rfloor = -\lfloor x \rfloor - 1$
- Pour tout x de \mathbb{R} , et tout m de \mathbb{Z} , on a : $\lfloor x + m \rfloor = \lfloor x \rfloor + m$
- Pour tous réels x et y : $\lfloor x + y \rfloor \in \{\lfloor x \rfloor + \lfloor y \rfloor, \lfloor x \rfloor + \lfloor y \rfloor + 1\}$

4.1.8 Densité de \mathbb{Q} et de $\mathbb{R} \setminus \mathbb{Q}$ dans \mathbb{R} **Proposition 4.1.9** (densité de \mathbb{Q} dans \mathbb{R})

Soient x et y deux réels, avec $x < y$. L'intervalle $]x, y[$ contient une infinité de nombres rationnels.

On exprime cette situation en disant que \mathbb{Q} est dense dans \mathbb{R} .

L'ensemble $\mathbb{R} \setminus \mathbb{Q}$ des nombres irrationnels est également dense dans \mathbb{R} .

On rappelle la définition des nombres décimaux.

Définition 4.1.12 (nombres décimaux)

On dit qu'un réel x est un nombre décimal s'il existe k dans \mathbb{Z} et n dans \mathbb{N} tel que $x = k 10^{-n}$.

Remarques

- Si on note \mathbb{D} l'ensemble des décimaux, on a $\mathbb{Z} \subsetneq \mathbb{D} \subsetneq \mathbb{Q}$.
- Les décimaux sont les réels qui peuvent s'écrire « avec un nombre fini de chiffres après la virgule ».
- Soit x un nombre rationnel, écrit sous forme irréductible $x = \frac{a}{b}$, avec a dans \mathbb{Z} , b dans \mathbb{N}^* , et $a \wedge b = 1$. Alors x est décimal si et seulement si les seuls diviseurs premiers éventuels de b sont 2 et 5.
- L'ensemble des décimaux est stable pour l'addition et pour le produit.
En revanche si x est un décimal non nul, son inverse $\frac{1}{x}$ n'est en général pas un décimal.
En fait les $x \neq 0$ tels que x et $\frac{1}{x}$ soient décimaux sont les $x = \pm 2^n 5^m$, avec (n, m) dans \mathbb{Z}^2 .

Définition 4.1.13 (approximations décimales d'un réel)

Soit x un nombre réel et n un entier naturel.

Il existe un unique entier relatif m tel que $m 10^{-n} \leq x < (m + 1) 10^{-n}$.

Le réel $\alpha_n = m 10^{-n}$ est appelé *valeur approchée* de x à 10^{-n} près *par défaut*. On a $\alpha_n = 10^{-n} \lfloor 10^n x \rfloor$.

Le réel $\beta_n = (m + 1) 10^{-n}$ est appelé *valeur approchée* de x à 10^{-n} près *par excès*.

Densité de l'ensemble des décimaux dans \mathbb{R}

Les propriétés ci-dessous montrent que l'ensemble des décimaux est dense dans \mathbb{R} .

Soit x un nombre réel. Pour tout n de \mathbb{N} , posons $\alpha_n = 10^{-n} \lfloor 10^n x \rfloor$ et $\beta_n = \alpha_n + 10^{-n}$.

- La suite (α_n) est une suite croissante de nombres décimaux.
- La suite (β_n) est une suite décroissante de nombres décimaux.
- Les deux suites (α_n) et (β_n) convergent vers x .

4.1.9 Droite achevée $\overline{\mathbb{R}}$

Définition 4.1.14

On note $\overline{\mathbb{R}}$ l'ensemble $\mathbb{R} \cup \{-\infty, +\infty\}$. Cet ensemble est appelé *droite numérique achevée*.

Relation d'ordre sur $\overline{\mathbb{R}}$

On étend l'ordre total de \mathbb{R} à $\overline{\mathbb{R}}$ en posant : $\forall x \in \mathbb{R}, -\infty \leq x \leq +\infty$ (en fait $-\infty < x < +\infty$).

Opérations sur $\overline{\mathbb{R}}$

De même, on étend (de façon toujours commutative) les lois $+$ et \times de \mathbb{R} en posant :

$$\text{Pour l'addition : } \begin{cases} (+\infty) + (+\infty) = +\infty & (-\infty) + (-\infty) = (-\infty) \\ \forall x \in \mathbb{R}, & x + (-\infty) = -\infty & x + (+\infty) = +\infty \end{cases}$$

$$\text{Pour le produit : } \begin{cases} (+\infty)(+\infty) = +\infty & (-\infty)(-\infty) = +\infty & (-\infty)(+\infty) = -\infty \\ \forall x \in \mathbb{R}^{+*}, & x(-\infty) = -\infty & x(+\infty) = +\infty \\ \forall x \in \mathbb{R}^{-*}, & x(-\infty) = +\infty & x(+\infty) = -\infty \end{cases}$$

Formes indéterminées

On ne donne pas de valeur aux expressions suivantes : $(+\infty) + (-\infty)$, $0(+\infty)$ et $0(-\infty)$.

Ces expressions sont appelées *formes indéterminées*.

Utiliser $\overline{\mathbb{R}}$ permet de simplifier les énoncés du genre :
$$\begin{cases} \lim_{n \rightarrow +\infty} u_n = \lambda \\ \lim_{n \rightarrow +\infty} v_n = \mu \end{cases} \Rightarrow \lim_{n \rightarrow +\infty} (u_n + v_n) = \lambda + \mu$$

Ce résultat est en effet vrai pour tous λ, μ de $\overline{\mathbb{R}}$ à l'exception des formes indéterminées pour lesquelles on devra faire une étude plus poussée (on devra donc *lever* la forme indéterminée).

4.2 Puissances à exposants entiers ou rationnels

4.2.1 Exposants entiers relatifs

Pour tout x de \mathbb{C} et tout n de \mathbb{N} , on définit par récurrence les puissances x^n :
$$\begin{cases} x^0 = 1 \\ \forall n \in \mathbb{N}, x^{n+1} = x^n x \end{cases}$$

Ainsi : $\forall n \in \mathbb{N}, 1^n = 1$, et $\forall n \in \mathbb{N}^*, 0^n = 0$.

Pour tout x non nul, et pour tout entier $n < 0$, on pose $x^n = (x^{-n})^{-1}$.

On connaît donc maintenant le sens de x^n , pour tout x de \mathbb{C} et tout n de \mathbb{Z} (avec $x \neq 0$ si $n < 0$).

Proposition 4.2.1 (propriétés des exposants relatifs)

Pour tous x, y de \mathbb{C} , pour tous n, p de \mathbb{Z} (et sous réserve d'existence en cas d'exposants négatifs), on a :

$$(xy)^n = x^n y^n, \quad x^n x^p = x^{n+p}, \quad (x^n)^p = x^{np}, \quad \frac{1}{x^n} = x^{-n}, \quad \frac{x^n}{x^p} = x^{n-p}$$

On en vient aux variations de $x \mapsto x^m$ sur \mathbb{R} (si m est dans \mathbb{N}) ou sur \mathbb{R}^* (si m est dans \mathbb{Z}^{-*}).

L'application $x \rightarrow x^m$ est paire si m est pair, et impaire si m est impair.

Sur \mathbb{R}^{+*} , elle est $\begin{cases} \text{strictement croissante si } m > 0, \\ \text{strictement décroissante si } m < 0, \\ \text{constante (valeur 1) si } m = 0 \end{cases}$

Le tableau ci-après indique ce que devient l'inégalité $x < y$ par élévation à la puissance m -ième.

	$m > 0$, pair	$m > 0$, impair	$m < 0$, pair	$m < 0$, impair
$0 < x < y \Rightarrow$	$0 < x^m < y^m$	$0 < x^m < y^m$	$0 < y^m < x^m$	$0 < y^m < x^m$
$x < y < 0 \Rightarrow$	$0 < y^m < x^m$	$x^m < y^m < 0$	$0 < x^m < y^m$	$y^m < x^m < 0$

4.2.2 Racine n -ième d'un réel positif**Proposition 4.2.2** (racine n -ième réelle positive, avec n dans \mathbb{N}^* , d'un réel positif)

Soit n un entier strictement positif. Soit a un réel positif (ou nul).

Il existe un unique réel positif ou nul x tel que $x^n = a$. On le note $\sqrt[n]{a}$, ou encore $a^{1/n}$.

Ce réel x est appelée est la racine n -ième positive du réel positif a .

Remarques

– Si $n = 2$, on note \sqrt{a} (plutôt que $\sqrt[2]{a}$) l'unique x de \mathbb{R}^+ tel que $x^2 = a$.

– Par exemple, l'unique réel x tel que $x^3 = 8$ est $x = 2$. Donc $\sqrt[3]{8} = 8^{1/3} = 2$.

L'unique réel x tel que $x^3 = -8$ est $x = -2$, mais on n'écrira pas $\sqrt[3]{-8} = -2$ ou $(-8)^{1/3} = -2$.

Citons le programme de la classe de MPSI : les fonctions puissances sont définies sur \mathbb{R}^{+*} et prolongées en 0 le cas échéant. Seules les fonctions puissances entières sont en outre définies sur \mathbb{R}^{-*} .

Attention ! (cas des nombres complexes)

Les remarques suivantes peuvent être passées en première lecture.

– On n'a parlé ici que de $x = \sqrt[n]{a}$, ou $x = a^{1/n}$ (n dans \mathbb{N}^*), avec a et x dans \mathbb{R}^+ .

Le problème précédent (trouver x tel que $x^n = a$) est très différent dans \mathbb{C} (voir chapitre 3).

– Par exemple, si a est dans \mathbb{C} (et $a \neq 0$), l'équation $x^2 = a$ possède exactement deux solutions distinctes et opposées dans \mathbb{C} , qu'on appelle les deux racines carrées complexes de a .

Les deux racines carrées complexes de 5 sont $x = \sqrt{5}$ et $x = -\sqrt{5}$ (elles sont réelles!).

De même, les deux racines carrées complexes de -1 sont $x = i$ et $x = -i$.

Dernier exemple : les deux racines carrées complexes de $2i$ sont $x = 1 + i$ et $x = -1 - i$.

Comme il n'y a aucun moyen objectif et sûr de choisir entre les deux racines carrées de a (dans le cas général) on n'écrira jamais \sqrt{a} (sauf si a est dans \mathbb{R}^+ , où \sqrt{a} désigne la racine carrée positive de a).

Par exemple, les notations $\sqrt{-1}$, ou encore $\sqrt{2i}$, sont interdites!

- Plus généralement l'équation $x^n = a$ (où a est un nombre complexe non nul) possède exactement n solutions dans \mathbb{C} (distinctes et non nulles).

Par exemple les trois solutions de $x^3 = 8$ sont $x = 2$, $x = -1 + i\sqrt{3}$ et $x = -1 - i\sqrt{3}$.

Les trois solutions de $x^3 = 8i$ sont $x = -2i$, $x = \sqrt{3} + i$ et $x = \sqrt{3} - i$.

La notation $\sqrt[3]{8}$ désigne 2, mais il est hors de question d'écrire $\sqrt[3]{8i}$.

4.2.3 Exposants rationnels

Définition 4.2.1 (puissances rationnelles d'un nombre réel)

Soit r un nombre rationnel, écrit sous la forme $r = p/q$ (avec p dans \mathbb{Z} et q dans \mathbb{N}^*).

On pose $x^r = (x^{1/q})^p$. Le domaine de définition de $x \mapsto x^r$ est : \mathbb{R}^+ si $p \geq 0$, et \mathbb{R}^{+*} si $p < 0$.

Remarque : la valeur x^r est indépendante de l'écriture choisie pour $r = \frac{p}{q}$, dans la mesure où $q > 0$.

Propriétés

- Les relations sur les fonctions puissances sont encore valables avec des exposants rationnels.

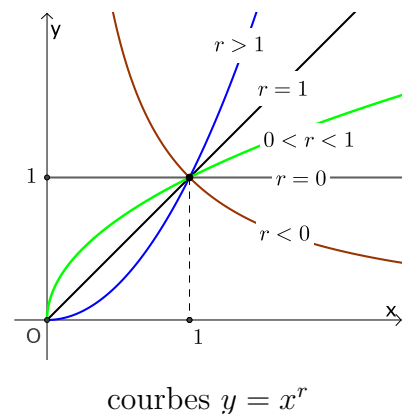
Ainsi, pour r, s dans \mathbb{Q} , on a les identités :

$$(xy)^r = x^r y^r, \quad x^r x^s = x^{r+s}, \quad (x^r)^s = x^{rs}, \quad \frac{1}{x^r} = x^{-r}, \quad \frac{x^r}{x^s} = x^{r-s}$$

- Ci-contre, on voit les différentes représentations graphiques.

On retiendra notamment que x^r et x^s sont dans le même ordre que r et s quand $x > 1$, et dans l'ordre contraire si $0 < x < 1$.

- La généralisation « ultime » de la notation x^α est $x^\alpha = e^{\alpha \ln x}$, mais elle suppose connue les fonctions $t \mapsto e^t$ et $t \mapsto \ln t$.



4.3 Généralités sur les fonctions numériques

Dans la suite, on considère des fonctions f à valeurs réelles, définies sur une partie \mathcal{D} de \mathbb{R} .

On note par exemple $f : \begin{matrix} \mathcal{D} & \rightarrow & \mathbb{R} \\ x & \mapsto & f(x) \end{matrix}$, et dit que f est une *fonction numérique*.

Avec ces notations, on dit que \mathcal{D} est le *domaine de définition* de f .

Le plus souvent ; \mathcal{D} est un intervalle ou une réunion disjointe d'intervalles.

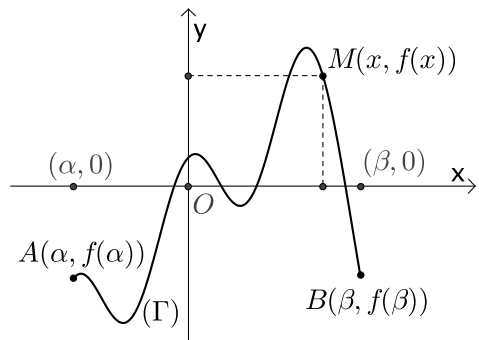
4.3.1 Représentations graphiques

On munit le plan d'un repère orthogonal (et même, le plus souvent, orthonormé).

L'ensemble (Γ) des points M de coordonnées $(x, f(x))$ est appelé *courbe représentative* (ou *graphe*) de f .

Chaque *droite verticale*, d'équation $x = x_0$, où x_0 est dans \mathcal{D} , rencontre (Γ) en l'unique point $(x_0, f(x_0))$.

On voit ici la courbe représentative d'une fonction définie sur un segment $[\alpha, \beta]$ de \mathbb{R} .



Soit $f : \mathcal{D} \rightarrow \mathbb{R}$ une fonction, et soit (Γ) sa courbe représentative.

À partir de f et d'un réel a , on va voir plusieurs définitions d'une fonction f_a .

À chaque fois, on va déterminer son domaine \mathcal{D}_a , et trouver la relation géométrique entre les courbes représentatives (Γ) de f et (Γ_a) de f_a .

Représentation graphique de $x \mapsto f_a(x) = f(x) + a$

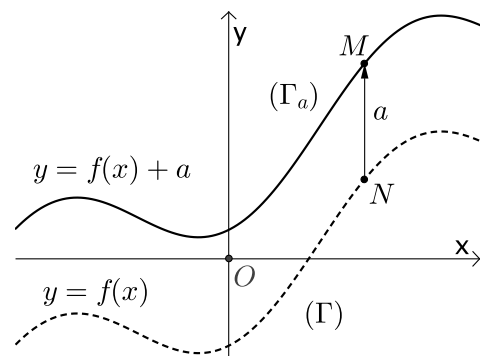
Posons par exemple $f_a(x) = f(x) + a$.

Le domaine \mathcal{D}_a de f_a égal au domaine \mathcal{D} de f .

(Γ_a) se déduit de (Γ) par la translation de vecteur $(0, a)$.

En effet, on a les équivalences :

$$\begin{aligned} M(x, y) \in \Gamma_a &\Leftrightarrow y = f_a(x) = f(x) + a \\ &\Leftrightarrow y - a = f(x) \\ &\Leftrightarrow N(x, y - a) \in \Gamma \end{aligned}$$



Représentation graphique de $x \mapsto f_a(x) = f(x + a)$

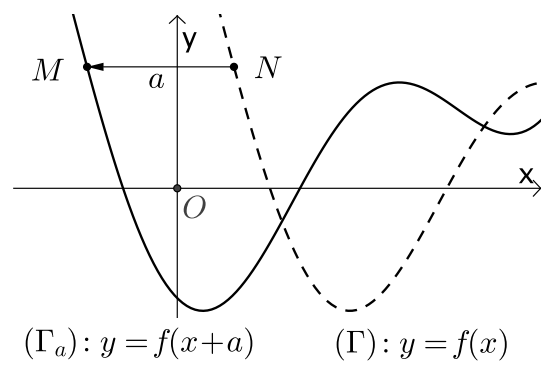
Si $f_a(x) = f(x + a)$, alors $\mathcal{D}_a = \{x - a, x \in \mathcal{D}\}$.

Ainsi (\mathcal{D}_a) se déduit de (\mathcal{D}) par la translation $x \mapsto x - a$.

De même, (Γ_a) se déduit de (Γ) par la translation de vecteur $(-a, 0)$.

En effet, on a les équivalences :

$$\begin{aligned} M(x, y) \in \Gamma_a &\Leftrightarrow y = f_a(x) = f(x + a) \\ &\Leftrightarrow N(x + a, y) \in \Gamma \end{aligned}$$



Représentation graphique de $x \mapsto f_a(x) = f(a - x)$

Si $f_a(x) = f(a - x)$, alors $\mathcal{D}_a = \{a - x, x \in \mathcal{D}\}$.

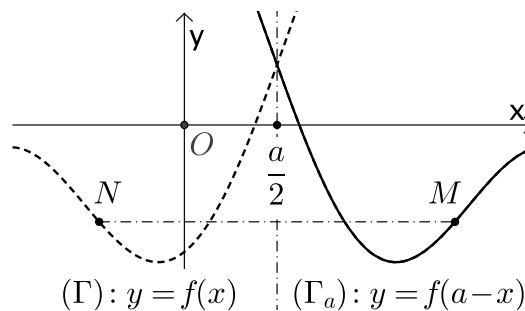
Ainsi (\mathcal{D}_a) se déduit de (\mathcal{D}) par la symétrie $x \mapsto a - x$ par rapport à la valeur $a/2$.

De même, la courbe (Γ_a) se déduit de (Γ) par la symétrie par rapport à l'axe vertical d'équation $x = a/2$.

En effet, on a les équivalences :

$$M(x, y) \in \Gamma_a \Leftrightarrow y = f(a - x) \Leftrightarrow N(a - x, y) \in \Gamma$$

Cas particulier : la courbe représentative de $x \mapsto f(-x)$ est symétrique de (Γ) par rapport à Oy .



Représentation graphique de $x \mapsto f_a(x) = f(ax)$

Si $f_a(x) = f(ax)$, alors $\mathcal{D}_a = \{x/a, x \in \mathcal{D}\}$.

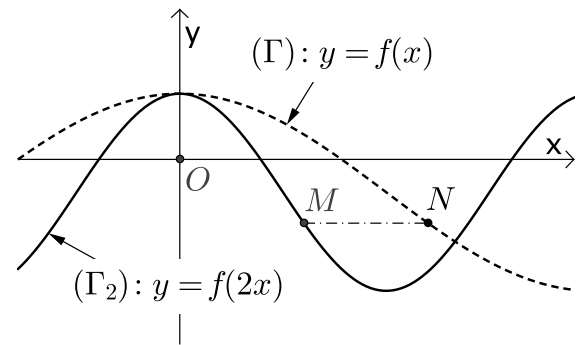
Ainsi (\mathcal{D}_a) se déduit de (\mathcal{D}) par l'homothétie $x \mapsto x/a$.

De même, (Γ_a) se déduit de (Γ) par une homothétie de rapport $1/a$ selon les abscisses uniquement (ce qui produit un effet en accordéon horizontal).

En effet, on a les équivalences :

$$M(x, y) \in \Gamma_a \Leftrightarrow y = f(ax) \Leftrightarrow N(ax, y) \in \Gamma$$

Sur l'exemple ci-contre, on a choisi $a = 2$.

**Représentation graphique de $x \mapsto f_a(x) = af(x)$**

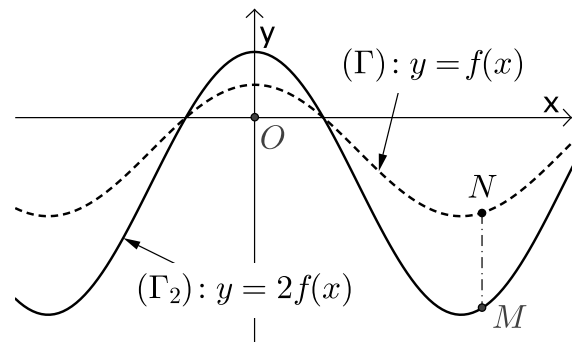
Si $f_a(x) = af(x)$, alors $\mathcal{D}_a = \mathcal{D}$.

Ici (Γ_a) se déduit de (Γ) par une homothétie de rapport a selon les ordonnées uniquement (ce qui produit un effet en accordéon vertical).

En effet, on a les équivalences :

$$M(x, y) \in \Gamma_a \Leftrightarrow y = af(x) \Leftrightarrow N(x, y/a) \in \Gamma$$

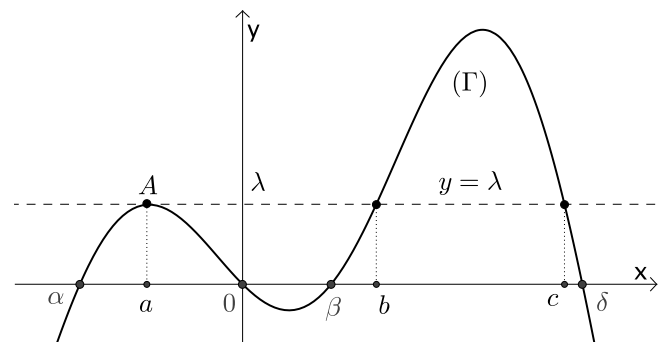
Sur l'exemple ci-contre, on a choisi $a = 2$.

**Interprétation graphique d'égalités et d'inégalités**

Des solutions de $f(x) = \lambda$ peuvent être représentées par les abscisses des points d'intersection de la courbe (Γ) et de la droite horizontale $y = \lambda$ (l'axe Ox dans le cas de l'équation $f(x) = 0$).

Sur cet exemple, on voit quatre solutions distinctes $\alpha < 0 < \beta < \delta$ de l'équation $f(x) = 0$. On voit aussi trois solutions $a < b < c$ de $f(x) = \lambda$ (la solution a pouvant être qualifiée de « double » ou « multiple »).

On interprète facilement les solutions des inéquations $f(x) < \lambda$, ou $f(x) \leq \lambda$, ou $f(x) \geq \lambda$, ou $f(x) > \lambda$.

**4.3.2 Opérations sur les fonctions numériques**

Soit f une fonction à valeurs réelles, définie sur une partie \mathcal{D} de \mathbb{R} (son *domaine de définition*).

L'ensemble \mathcal{D} consiste le plus souvent en une réunion d'intervalles.

L'étude de f (continuité, monotonie, extrémums, etc.) s'effectue cependant *intervalle par intervalle*.

C'est pourquoi on se limitera souvent à des fonctions réelles, définies sur un intervalle I de \mathbb{R} (évidemment l'intervalle I ne doit être ni vide, ni réduit à un point).

On note $\mathcal{F}(\mathcal{D}, \mathbb{R})$, ou $\mathbb{R}^{\mathcal{D}}$, l'ensemble des fonctions $f : \mathcal{D} \rightarrow \mathbb{R}$.

On rappelle que si f et g appartiennent à $\mathcal{F}(\mathcal{D}, \mathbb{R})$: $f = g \Leftrightarrow \forall x \in \mathcal{D}, f(x) = g(x)$.

Définition 4.3.1 (somme et produit de deux fonctions numériques)

Soit f et g deux éléments de $\mathcal{F}(\mathcal{D}, \mathbb{R})$.

On définit les fonctions $f + g$ et fg par :

$$\begin{cases} \forall x \in \mathcal{D}, (f + g)(x) = f(x) + g(x) \\ \forall x \in \mathcal{D}, (fg)(x) = f(x)g(x) \end{cases}$$
Propriétés immédiates

La fonction constante $x \mapsto 0$ (resp $x \mapsto 1$) est neutre pour l'addition (resp. le produit).

L'opposée d'une fonction f est la fonction notée $-f$ définie par : $\forall x \in I, (-f)(x) = -f(x)$.

On peut parler de $\frac{1}{f}$ si f ne s'annule pas sur \mathcal{D} . Dans ces conditions : $\forall x \in \mathcal{D}, \frac{1}{f}(x) = \frac{1}{f(x)}$.

Soit λ un réel. On note encore λ la fonction constante $x \mapsto \lambda$.

La notation λf désigne la fonction définie par : $\forall x \in \mathcal{D}, (\lambda f)(x) = \lambda f(x)$.

On peut ainsi former des *combinaisons linéaires* $g = \sum_{k=1}^n \lambda_k f_k$ de fonctions définies sur \mathcal{D} .

Si les fonctions $f : \mathcal{D} \rightarrow \mathbb{R}$ et $g : \mathcal{D}' \rightarrow \mathbb{R}$ ont des domaines différents, celui de $f + g$ et de fg est $\mathcal{D} \cap \mathcal{D}'$.

Définition 4.3.2 (composée de deux fonctions numériques)

Soit f et g deux fonctions numériques, de domaines respectifs \mathcal{D}_f et \mathcal{D}_g .

On définit une fonction numérique, notée $g \circ f$, en posant $(g \circ f)(x) = g(f(x))$.

Le domaine de la fonction $g \circ f$ est $\mathcal{D} = \{x \in \mathcal{D}_f, f(x) \in \mathcal{D}_g\}$.

L'ensemble de définition de $g \circ f$ est donc l'ensemble des réels x pour lesquels d'une part $f(x)$ existe, et d'autre part $f(x)$ est dans le domaine de définition de g .

Dans les questions théoriques portant sur la composition (monotonie par exemple), on considérera souvent des fonctions numériques $f : I \rightarrow \mathbb{R}$ et $g : J \rightarrow \mathbb{R}$ (définies respectivement sur les intervalles I et J) et on ajoutera la condition $f(I) \subset J$ pour s'assurer que la fonction $g \circ f$ est définie sur I .

On est souvent amené à étudier des fonctions numériques f obtenues par somme, produit, quotient, composition de fonctions usuelles f_k (dont il est bien sûr indispensable de connaître le domaine). On déterminera alors le domaine de définition de f en appliquant les règles indiquées précédemment, et en prêtant attention à la chronologie des compositions de fonctions.

4.3.3 Fonctions paires ou impaires

On considère ici des fonctions numériques définies sur une partie \mathcal{D} de \mathbb{R} .

On suppose que l'ensemble \mathcal{D} est symétrique par rapport à l'origine (donc : $x \in \mathcal{D} \Leftrightarrow -x \in \mathcal{D}$).

Le cas le plus courant est celui d'un intervalle de centre 0, et notamment $\mathcal{D} = \mathbb{R}$.

Définition 4.3.3 (parité, imparité)

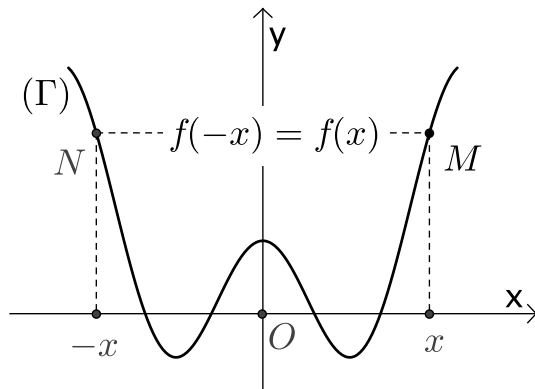
On dit que f est *paire* si : $\forall x \in \mathcal{D}, f(-x) = f(x)$.

On dit que f est *impaire* si : $\forall x \in \mathcal{D}, f(-x) = -f(x)$.

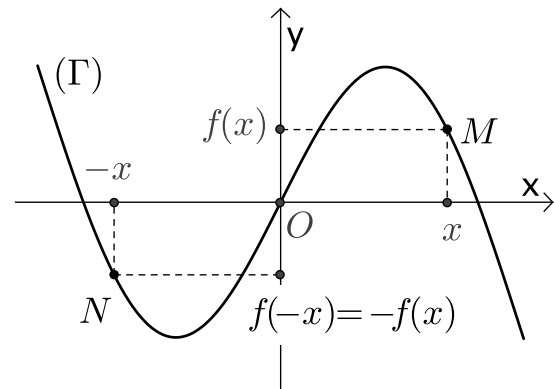
La seule fonction à la fois paire et impaire est la fonction nulle.

Évoquer la parité ou imparité de $f : \mathcal{D} \rightarrow \mathbb{R}$ n'a de sens que si \mathcal{D} est symétrique par rapport à 0.

L'interprétation graphique est évidente. La fonction f est paire (resp. impaire) si et seulement si sa représentation graphique (Γ) est symétrique par rapport à l'axe Oy (resp. l'origine O).



Graph (Γ) d'une fonction paire



Graph (Γ) d'une fonction impaire

Proposition 4.3.1 (Parties paire et impaire d'une fonction)

Soit f une fonction de \mathcal{D} dans \mathbb{R} (où le domaine \mathcal{D} est symétrique par rapport à 0).

Alors f s'écrit de manière unique $f = f_p + f_i$, où f_p est paire et f_i est impaire.

Les fonctions f_p et f_i sont définies par : $\forall x \in \mathcal{D}$, $f_p(x) = \frac{f(x) + f(-x)}{2}$ et $f_i(x) = \frac{f(x) - f(-x)}{2}$

On dit que f_p est la partie paire de f et que f_i en est la partie impaire.

Remarques :

On considère les fonctions ch et sh définies sur \mathbb{R} par $\text{ch}(x) = \frac{e^x + e^{-x}}{2}$ et $\text{sh}(x) = \frac{e^x - e^{-x}}{2}$.

Les fonctions ch et sh sont appelées *cosinus hyperbolique* et *sinus hyperbolique*.

Elles sont donc respectivement la partie paire et la partie impaire de $x \mapsto e^x$.

Proposition 4.3.2 (produits et sommes de fonctions paires ou impaires)

Soit f et g deux fonctions de \mathcal{D} dans \mathbb{R} , chacune d'elles étant paire ou impaire.

Si f et g ont même parité, alors fg est paire, sinon fg est impaire.

Soit (α, β) dans \mathbb{R}^2 : si f et g ont même parité, alors $h = \alpha f + \beta g$ a même parité que f et g .

Proposition 4.3.3 (opérations entre fonctions paires ou impaires)

Soit $f : \mathcal{D} \rightarrow \mathbb{R}$ et $g : \mathcal{D}' \rightarrow \mathbb{R}$ deux fonctions numériques.

Si f est paire, alors $g \circ f$ est paire (quelle que soit la fonction g , en fait).

Si f est impaire, et si g est paire ou impaire, alors $g \circ f$ a la même parité que g .

Quelques remarques faciles pour terminer

Soit n dans \mathbb{N} . Si f est paire, alors f^n est paire. Si f est impaire, alors f^n a la parité de n .

Si f est paire ou impaire alors $|f|$ est paire.

Si f est paire ou impaire et ne s'annule pas, alors $1/f$ a la parité de f .

Si f est bijective de \mathcal{D} sur \mathcal{D}' et impaire, alors sa bijection réciproque f^{-1} est impaire.

4.3.4 Axes et centres de symétrie

Proposition 4.3.4 (existence d'un axe de symétrie vertical)

Soit $f : \mathcal{D} \rightarrow \mathbb{R}$, le domaine \mathcal{D} étant symétrique par rapport au réel a .

Les conditions suivantes sont équivalentes :

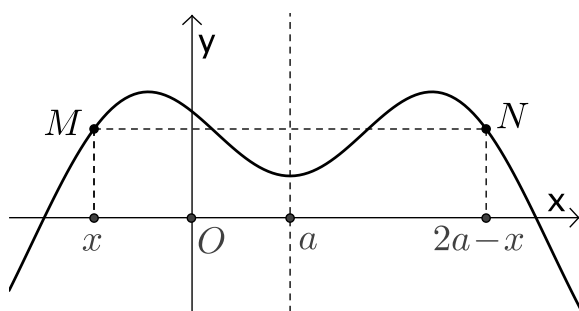
- la droite verticale $x = a$ est axe de symétrie du graphe Γ de f .
- pour tout x de \mathcal{D} , $f(2a - x) = f(x)$.
- pour tout h tel que $a \pm h$ appartienne à \mathcal{D} , $f(a + h) = f(a - h)$.
- la fonction g définie par $g(x) = f(a + x)$ est paire.

Proposition 4.3.5 (existence d'un centre de symétrie)

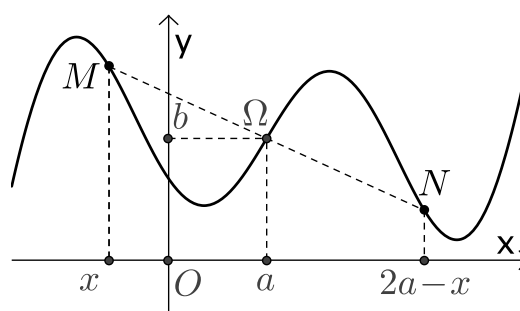
Soit $f : \mathcal{D} \rightarrow \mathbb{R}$, le domaine \mathcal{D} étant symétrique par rapport au réel a .

Les conditions suivantes sont équivalentes :

- le point $\Omega(a, b)$ est centre de symétrie du graphe Γ de f .
- pour tout x de \mathcal{D} , $f(2a - x) = 2b - f(x)$.
- pour tout h tel que $a \pm h$ appartienne à \mathcal{D} , $f(a + h) - b = b - f(a - h)$.
- la fonction g définie par $g(x) = f(a + x) - b$ est impaire.



Axe de symétrie $x = a$



Centre de symétrie $\Omega(a, b)$

4.3.5 Applications périodiques

Définition 4.3.4 (fonction T -périodique)

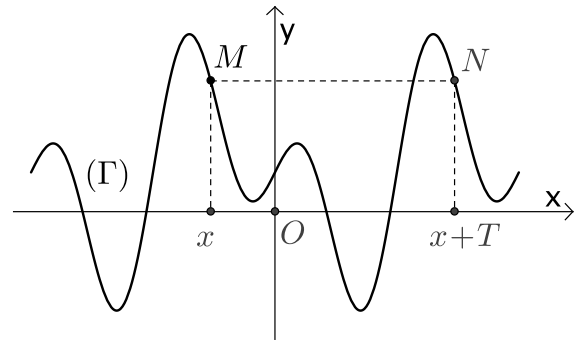
Soit \mathcal{D} une partie non vide de \mathbb{R} , et soit $f : \mathcal{D} \rightarrow \mathbb{R}$. Soit T un réel strictement positif.

La fonction f est dite T -périodique si : $\forall x \in \mathcal{D}$, $x + T \in \mathcal{D}$ et $f(x + T) = f(x)$

On a alors, pour tout x de \mathcal{D} et tout k de \mathbb{Z} : $f(x + kT) = f(x)$.

Une condition nécessaire pour qu'on puisse évoquer la T -périodicité de $f : \mathcal{D} \rightarrow \mathbb{R}$ est que le domaine \mathcal{D} soit lui-même *stable* par la translation $x \mapsto x + T$ (on le supposera dans ce qui suit).

L'interprétation graphique est facile : dire que la fonction f est T -périodique c'est dire que son graphe Γ est stable dans la translation de vecteur $(T, 0)$ (parallèle à Ox).



Notion de plus petite période positive

Si la fonction f est T -périodique, alors pour tout n de \mathbb{N}^* , elle est nT -périodique.

On cherchera donc, si possible, la plus petite période positive (qu'on appellera alors **la** période de f).

Dans certains cas très spécifiques, cette plus petite période n'existe pas : il en est ainsi par exemple de la fonction indicatrice de \mathbb{Q} , qui admet tout rationnel positif comme période.

Opérations entre fonctions T -périodiques

Si f et g sont T -périodiques, alors $\alpha f + \beta g$ et fg sont T -périodiques.

Si f est périodique et ne s'annule pas, alors $\frac{1}{f}$ est T -périodique.

Si f est T -périodique, alors pour toute fonction g , la fonction $g \circ f$ est T -périodique.

Quand on combine des fonctions de période T , on peut obtenir des fonctions de période inférieure à T .

Par exemple $\begin{cases} x \mapsto \sin x \\ x \mapsto \cos x \end{cases}$ sont 2π -périodiques, mais $\begin{cases} x \mapsto \sin x \cos x & \text{est } \pi\text{-périodique} \\ x \mapsto (\sin x \cos x)^2 & \text{est } \pi/2\text{-périodique} \end{cases}$

Cas de deux fonctions ayant des périodes distinctes

Soit f une fonction de période T_1 , et soit g une fonction de période T_2 .

On suppose que le rapport $\frac{T_1}{T_2}$ est rationnel. Alors $f + g$ et fg sont encore périodiques.

Par exemple : si $T_1 = \frac{3\pi}{4}$ et $T_2 = \frac{\pi}{2}$, alors $f + g$ et fg sont périodiques de période $\frac{3\pi}{2}$.

Mais si le rapport $\frac{T_1}{T_2}$ est irrationnel, alors $f + g$ et fg ne sont en général pas périodiques.

Si f est de période T , et si α est non nul, alors la fonction $g : x \mapsto f(\alpha x + \beta)$ est de période $\frac{T}{|\alpha|}$.

4.3.6 Monotonie des fonctions numériques

Définition 4.3.5

Soit $f : \mathcal{D} \rightarrow \mathbb{R}$ une fonction numérique. On dit que f est :

- *croissante* (au sens large) si : $\forall (x, y) \in \mathcal{D}^2, x \leq y \Rightarrow f(x) \leq f(y)$.
- *décroissante* (au sens large) si : $\forall (x, y) \in \mathcal{D}^2, x \leq y \Rightarrow f(x) \geq f(y)$.

- *strictement croissante* si : $\forall (x, y) \in \mathcal{D}^2, x < y \Rightarrow f(x) < f(y)$.
- *strictement décroissante* si : $\forall (x, y) \in \mathcal{D}^2, x < y \Rightarrow f(x) > f(y)$.
- *monotone* (au sens large) si f est croissante ou décroissante.
- *strictement monotone* si f est strictement croissante ou strictement décroissante.

Remarques

Quand on parle de monotonie, c'est par défaut de monotonie au sens large.

f est croissante (resp. décroissante) $\iff f$ conserve (resp. inverse) les inégalités larges.

f est strictement croissante (resp. décroissante) $\iff f$ conserve (resp. inverse) les inégalités strictes.

Seules les fonctions constantes sont à la fois croissantes et décroissantes.

Dire que $f : \mathcal{D} \rightarrow \mathbb{R}$ n'est pas monotone, c'est dire : $\exists (x, y, z) \in \mathcal{D}^3, z \in [x, y], f(z) \notin [f(x), f(y)]$.

Soit $f : I \rightarrow \mathbb{R}$ une fonction monotone sur un intervalle I . Dire que f n'est pas strictement monotone signifie qu'il existe un segment $[a, b]$ inclus dans I (avec $a < b$) sur lequel f garde une valeur constante.

Monotonie et taux d'accroissement

Pour tous a, b distincts de \mathcal{D} , on note $T_f(a, b) = \frac{f(b) - f(a)}{b - a}$ le taux d'accroissement de f entre a et b .

La quantité $T_f(a, b) = T_f(b, a)$ mesure le coefficient directeur de la corde entre $M(a, f(a))$ et $N(b, f(b))$.

Avec ces notations :

- La fonction f est croissante (au sens large) \iff tous ses taux d'accroissement sont positifs ou nuls.
- La fonction f est décroissante (au sens large) \iff tous ses taux d'accroissement sont négatifs ou nuls.
- La fonction f est strictement croissante \iff tous ses taux d'accroissement sont strictement positifs.
- La fonction f est strictement décroissante \iff tous ses taux d'accroissement sont strictement négatifs.

Opérations sur les taux d'accroissement

Soit f et g deux fonctions numériques, définies sur \mathcal{D} . Soit λ un nombre réel.

On a les égalités suivantes, en termes de taux d'accroissement :

$$T_{\lambda f}(a, b) = \frac{(\lambda f)(b) - (\lambda f)(a)}{b - a} = \lambda \frac{f(b) - f(a)}{b - a} = \lambda T_f(a, b)$$

$$T_{f+g}(a, b) = \frac{(f+g)(b) - (f+g)(a)}{b - a} = \frac{f(b) - f(a)}{b - a} + \frac{g(b) - g(a)}{b - a} = T_f(a, b) + T_g(a, b)$$

$$T_{fg}(a, b) = \frac{(fg)(b) - (fg)(a)}{b - a} = g(b) \frac{f(b) - f(a)}{b - a} + f(a) \frac{g(b) - g(a)}{b - a} = g(b)T_f(a, b) + f(a)T_g(a, b)$$

Voici deux propositions qui résultent directement des calculs précédents.

Proposition 4.3.6 (sommées de fonctions monotones)

Soit f et g deux fonctions monotones sur leur domaine \mathcal{D} .

Si f et g ont même monotonie, alors $f + g$ est monotone sur \mathcal{D} de même monotonie que f et g .

Si de plus f ou g est strictement monotone, alors $f + g$ est strictement monotone.

Il est clair que si $f : \mathcal{D} \rightarrow \mathbb{R}$ est monotone, et que si λ est un réel :

- si $\lambda \geq 0$, alors la fonction λf a la même monotonie que f
- si $\lambda \leq 0$, alors la fonction λf a la monotonie contraire de celle de f

En particulier, la fonction $-f$ a la monotonie contraire de celle de f .

Il y a davantage de cas particuliers à envisager pour le produit de deux fonctions monotones :

Proposition 4.3.7 (produits de fonctions monotones)

Soit f et g deux fonctions monotones sur leur domaine \mathcal{D} .

- Si f et g sont positives croissantes, alors fg est positive croissante.
- Si f et g sont positives décroissantes, alors fg est positive décroissante.
- Si f et g sont négatives croissantes, fg est positive décroissante.
- Si f et g sont négatives décroissantes, fg est positive croissante.
- Si f est positive croissante et g négative décroissante, fg est négative décroissante.
- Si f est positive décroissante et g négative croissante, fg est négative croissante.

Remarque :

Dans les cas autres que ceux énumérés ci-dessus, on ne peut rien dire.

Quand f et g sont de monotonies contraires, il n'y a donc pas de résultat général concernant $f + g$.

De même, on ne peut rien dire a priori de la monotonie de fg si, par exemple, f est positive croissante tandis que g est positive décroissante.

Taux d'accroissement de $1/f$ si f ne s'annule pas

Soit $f : \mathcal{D} \rightarrow \mathbb{R}$, ne s'annulant pas. Pour tous a, b de \mathcal{D} (avec $a \neq b$) :

$$T_{1/f}(a, b) = \frac{(1/f)(b) - (1/f)(a)}{b - a} = \frac{1}{b - a} \left(\frac{1}{f(b)} - \frac{1}{f(a)} \right) = \frac{-1}{f(a)f(b)} \frac{f(b) - f(a)}{b - a} = \frac{-1}{f(a)f(b)} T_f(a, b)$$

On en déduit alors facilement le résultat suivant :

Proposition 4.3.8 (monotonie de $1/f$ si f est monotone)

Soit f une fonction monotone de \mathcal{D} dans \mathbb{R} , strictement positive (ou strictement négative) sur \mathcal{D} .

Alors $\frac{1}{f}$ est monotone sur \mathcal{D} , de monotonie contraire à celle de f .

Remarque :

La condition que f ne change pas de signe est importante.

Par exemple la fonction $x \mapsto x$ est croissante sur \mathbb{R}^* , mais $x \mapsto 1/x$ n'est pas monotone sur \mathbb{R}^* (en revanche elle est strictement décroissante sur chacun des intervalles \mathbb{R}^{-*} et \mathbb{R}^{+*}).

Taux d'accroissement de $g \circ f$

Soit $f : \mathcal{D} \rightarrow \mathbb{R}$ et $g : \mathcal{D}' \rightarrow \mathbb{R}$ deux fonctions numériques, de domaines \mathcal{D} et \mathcal{D}' .

On suppose que f est injective sur \mathcal{D} . Pour tous a, b de \mathcal{D} (avec $a \neq b$) :

$$T_{g \circ f}(a, b) = \frac{g(f(b)) - g(f(a))}{b - a} = \frac{f(b) - f(a)}{b - a} \frac{g(f(b)) - g(f(a))}{f(b) - f(a)} = T_f(a, b) T_g(f(a), f(b))$$

Proposition 4.3.9 (composition de fonctions monotones)

Soit $f : \mathcal{D} \rightarrow \mathbb{R}$ et $g : \mathcal{D}' \rightarrow \mathbb{R}$ deux fonctions numériques, de domaines \mathcal{D} et \mathcal{D}' .

- Si f et g ont la même monotonie, alors $g \circ f$ est croissante sur son domaine.
- Si f et g sont de monotonies contraires, alors $g \circ f$ est décroissante sur son domaine.

Les deux propriétés précédentes restent vraies pour des monotonies strictes.

4.3.7 Fonctions majorées, minorées, bornées

Définition 4.3.6

On dit que f est *majorée* sur \mathcal{D} s'il existe un réel β tel que : $\forall x \in \mathcal{D}, f(x) \leq \beta$.

On dit que f est *minorée* sur \mathcal{D} s'il existe un réel α tel que : $\forall x \in \mathcal{D}, \alpha \leq f(x)$.

On dit que f est *bornée* sur \mathcal{D} si elle y est majorée et minorée.

Ainsi, dire que f est bornée sur \mathcal{D} , c'est dire qu'il existe α, β tels que : $\forall x \in \mathcal{D}, \alpha \leq f(x) \leq \beta$.

Remarques

Soit $f : \mathcal{D} \rightarrow \mathbb{R}$ une fonction numérique. Notons $f(\mathcal{D}) = \{f(x), x \in \mathcal{D}\}$ son *ensemble image*.

Dire que la fonction f est majorée (resp. minorée, bornée) sur \mathcal{D} , c'est dire que l'ensemble $f(\mathcal{D})$ est une partie majorée (resp. minorée, bornée) de \mathbb{R} .

Si f est majorée sur \mathcal{D} , on note $\sup_{\mathcal{D}} f$, ou encore $\sup_{x \in \mathcal{D}} f(x)$, la borne supérieure de l'ensemble $f(\mathcal{D})$.

Si f est minorée sur \mathcal{D} , on note $\inf_{\mathcal{D}} f$, ou encore $\inf_{x \in \mathcal{D}} f(x)$, la borne inférieure de l'ensemble $f(\mathcal{D})$.

Dire que la fonction f est bornée sur \mathcal{D} , c'est dire que la fonction $|f|$ est majorée sur \mathcal{D} .

S'il existe x_0 dans \mathcal{D} tel que $f(x_0) = \sup_{\mathcal{D}} f$, on note $f(x_0) = \max_{\mathcal{D}} f$ (on a un maximum global en x_0).

De même, s'il existe x_1 dans \mathcal{D} tel que $f(x_1) = \inf_{\mathcal{D}} f$, on note $f(x_1) = \min_{\mathcal{D}} f$ (minimum global en x_1).

Quelques propriétés faciles à vérifier

Si f et g sont majorées sur \mathcal{D} , alors $f + g$ est majorée sur \mathcal{D} , et $\sup_{\mathcal{D}}(f + g) \leq \sup_{\mathcal{D}} f + \sup_{\mathcal{D}} g$.

Si f et g sont minorées sur \mathcal{D} , alors $f + g$ est minorée sur \mathcal{D} , et $\inf_{\mathcal{D}}(f + g) \geq \inf_{\mathcal{D}} f + \inf_{\mathcal{D}} g$.

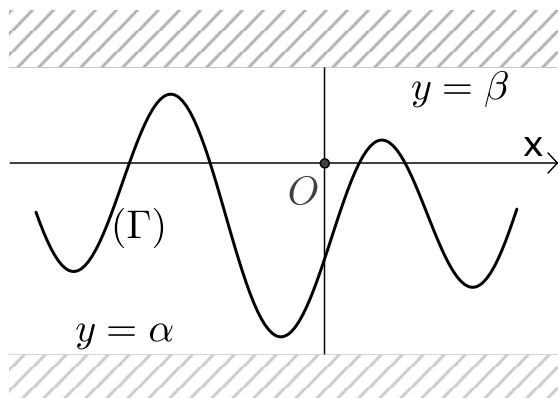
On a les égalités : $\inf_{\mathcal{D}}(-f) = -\sup_{\mathcal{D}} f$, et $\sup_{\mathcal{D}}(-f) = -\inf_{\mathcal{D}} f$.

Si $\alpha > 0$, alors $\sup_{\mathcal{D}}(\alpha f) = \alpha \sup_{\mathcal{D}} f$ et $\inf_{\mathcal{D}}(\alpha f) = \alpha \inf_{\mathcal{D}} f$ (mais si $\alpha < 0$?).

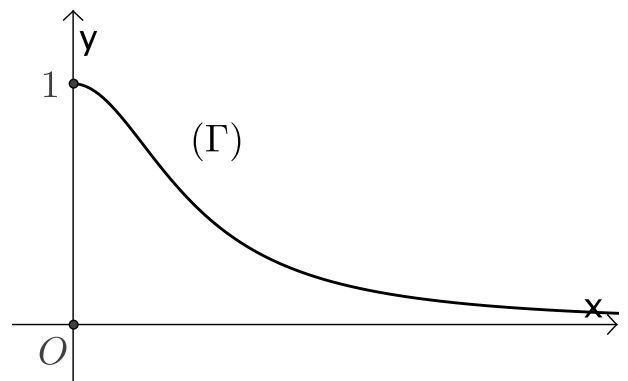
Si f et g sont bornées sur \mathcal{D} , alors fg et $\alpha f + \beta g$ sont bornées sur \mathcal{D} .

Question : si f est bornée et ne s'annule pas sur \mathcal{D} , peut-on affirmer que $\frac{1}{f}$ est bornée sur \mathcal{D} ?

L'interprétation graphique de la notion de fonction majorée et/ou minorée est facile :



$$\forall x \in \mathcal{D}, \alpha \leq f(x) \leq \beta$$



$$f \text{ décroissante sur } \mathbb{R}^+ : \max_{\mathbb{R}^+} f = 1, \inf_{\mathbb{R}^+} f = 0$$

4.4 Dérivation des fonctions numériques

4.4.1 Notion de fonction continue

Dans la suite de ce chapitre, on considère essentiellement des fonctions numériques définies sur un intervalle I non vide et non réduit à un point (on dit aussi « intervalle d'intérieur non vide »).

Dans le cas fréquent d'une fonction définie sur une réunion d'intervalles disjoints, on se ramène au cas précédent en étudiant la restriction de f à chacun de ces intervalles.

Dans ce chapitre, la notion de « fonction continue » (en un point, et plus généralement sur un intervalle) est supposée connue (on se référera au cours de Terminale S).

On se contentera donc pour l'instant d'une approche intuitive de la continuité (la faculté de pouvoir tracer la courbe représentative « sans lever le crayon »).

Les fonctions usuelles (fonctions puissances, trigonométriques, exponentielle, logarithme, etc.) sont continues sur leur domaine (la fonction « partie entière » est une exception notable).

Les deux énoncés suivants permettent donc d'affirmer la continuité d'une fonction numérique définie comme un « cocktail » de fonctions usuelles.

Proposition 4.4.1 (sommes, produits et quotients de fonctions continues)

Soit f et g deux fonctions continues sur l'intervalle I .

Alors les fonctions $\alpha f + \beta g$ et fg sont continues sur I .

Si g ne s'annule pas sur I , alors les fonctions $\frac{1}{g}$ et $\frac{f}{g}$ sont continues sur I .

Proposition 4.4.2 (composée de deux fonctions continues)

Soit $f : I \rightarrow \mathbb{R}$ et $g : J \rightarrow \mathbb{R}$ deux fonctions continues, avec $f(I) \subset J$.

Alors la fonction $g \circ f$ est continue sur I .

4.4.2 Équation de la tangente en un point

La notion de « fonction dérivable » est supposée connue (cf cours de 1ère S et de Terminale S).

On se bornera à dire, de façon très informelle, que $f : I \rightarrow \mathbb{R}$ est dérivable en un point a de I si sa courbe représentative présente, au point $A(a, f(a))$, une tangente **non verticale** Δ_a .

On dit que la fonction f est dérivable sur l'intervalle I si elle est dérivable en tout point de I .

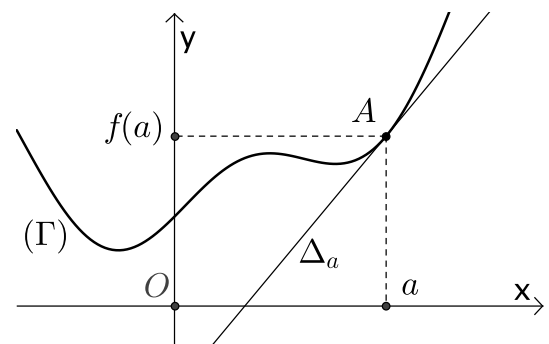
On note alors f' la fonction qui à toute valeur a de I associe le coefficient directeur de la tangente Δ_a .

L'équation de Δ_a est donc : $y = f(a) + (x - a)f'(a)$.

La fonction f' est parfois noté $D(f)$, ou $\frac{df}{dx}$.

On généralise à la notion de dérivée à droite (ou à gauche) en un point (notamment aux bornes de I).

On parle alors de demi-tangente à droite (ou à gauche) au point correspondant de la courbe (Γ) .



4.4.3 Opérations sur les fonctions dérivables

À ce stade, les résultats de cette sous-section sont admis.

Proposition 4.4.3 (dérivée d'une combinaison linéaire)

Soit f et g deux fonctions dérivables sur l'intervalle I .

Pour tous α, β dans \mathbb{R} , $h = \alpha f + \beta g$ est dérivable sur I et $h' = \alpha f' + \beta g'$.

Proposition 4.4.4 (dérivée d'un produit)

Soit f et g deux fonctions dérivables sur l'intervalle I .

La fonction fg est dérivable sur I et $(fg)' = f'g + fg'$.

Proposition 4.4.5 (dérivée de l'inverse, du quotient)

Soit g une fonction dérivable sur I , et ne s'annulant pas sur I .

Alors la fonction $\frac{1}{g}$ est dérivable sur I et $\left(\frac{1}{g}\right)' = -\frac{g'}{g^2}$.

Si de plus f est dérivable sur I , alors $\frac{f}{g}$ est dérivable sur I et $\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}$

Proposition 4.4.6 (dérivée d'une composée)

Soit $f : I \rightarrow \mathbb{R}$ et $g : J \rightarrow \mathbb{R}$ deux fonctions dérivables, avec $f(I) \subset J$.

Alors $g \circ f$ est dérivable sur I et $(g \circ f)' = f' \cdot (g' \circ f)$

Les fonctions usuelles (fonctions puissances, trigonométriques, exponentielle, logarithme, etc.) sont en général dérivables sur leur domaine, mais il y a (au moins) deux exceptions notables :

- la fonction $x \mapsto |x|$ est continue sur \mathbb{R} , mais n'est pas dérivable en 0 (demi-tangentes non colinéaires).
- la fonction $x \mapsto \sqrt{x}$ est continue sur \mathbb{R}^+ , mais n'est pas dérivable en 0 (demi-tangente verticale).

Les énoncés sur les opérations entre fonctions dérivables permettent donc de juger de la dérivabilité d'une fonction numérique définie comme un « cocktail » de fonctions usuelles.

Dans les cas particuliers où ces résultats généraux ne s'appliquent pas (notamment si on considère $x \mapsto |f(x)|$ ou $x \mapsto \sqrt{f(x)}$ en un point où f s'annule) on sera amené à effectuer une « étude locale ».

Dérivée d'un produit de plusieurs fonctions

Si f_1, f_2, \dots, f_n sont dérivables sur I , il en est de même de leur produit $g = f_1 f_2 \cdots f_n$.

Plus précisément, on a : $g' = f_1' f_2 \cdots f_n + f_1 f_2' f_3 \cdots f_n + f_1 f_2 \cdots f_{n-1}' f_n$.

Pour dériver le produit des f_k , on dérive donc chaque f_k « à son tour » et on fait la somme des résultats.

Un cas particulier est la dérivée de f^n , où n est dans \mathbb{N}^* . On trouve : $(f^n)' = n f' f^{n-1}$.

Cela se généralise à la formule $(f^\alpha)' = \alpha f' f^{\alpha-1}$, où α est un réel (à condition que f^α soit défini).

Si les f_k ne s'annulent pas, la dérivée de leur produit g se retrouve assez bien en considérant :

$$|g| = \prod_{k=1}^n |f_k| \Rightarrow \ln |g| = \sum_{k=1}^n \ln |f_k| \Rightarrow \frac{g'}{g} = \sum_{k=1}^n \frac{f_k'}{f_k} \Rightarrow g' = g \sum_{k=1}^n \frac{f_k'}{f_k} = \sum_{k=1}^n \frac{g}{f_k} f_k'$$

4.4.4 Dérivabilité et sens de variation

À ce stade, les résultats de cette sous-section sont admis. On rappelle que I est un *intervalle*.

Proposition 4.4.7 (caractérisation des fonctions constantes ou monotones)

Soit $f : I \rightarrow \mathbb{R}$ une fonction dérivable. On les équivalences :

La fonction f est constante sur I si et seulement si : $\forall x \in I, f'(x) = 0$.

La fonction f est croissante (au sens large) sur I si et seulement si : $\forall x \in I, f'(x) \geq 0$.

La fonction f est décroissante (au sens large) sur I si et seulement si : $\forall x \in I, f'(x) \leq 0$.

Proposition 4.4.8 (caractérisation des fonctions strictement monotones)

Soit $f : I \rightarrow \mathbb{R}$ une fonction dérivable et monotone. Alors f est strictement monotone sur I si et seulement si sa dérivée f' n'est identiquement nulle sur aucun sous-intervalle de I d'intérieur non vide (c'est-à-dire si et seulement si f' ne s'annule qu'en des points isolés de I).

Remarques

- Si $f'(x) > 0$ (resp. $f'(x) < 0$) sur I alors f est strictement croissante (resp. décroissante) sur I .
Mais l'exemple de $f : x \mapsto x + \cos(x)$ (strictement croissante sur \mathbb{R} , mais de dérivée nulle en tous les $x = \pi/2 + k\pi$ avec k dans \mathbb{Z}) confirme que la réciproque n'est pas vraie (la dérivée peut s'annuler, sans que ça remette en cause la stricte monotonie, du moment que c'est en des points isolés).
- Si on sait que f est constante (ou monotone) sur l'intérieur de I (sur I privé de ses « extrémités ») et si on sait que f est continue aux extrémités de I , alors le caractère constant ou monotone de f s'étend à l'intervalle I tout entier.
- Attention : les résultats énoncés ici sont valables sur un intervalle, pas sur une réunion d'intervalles.
Par exemple, si $f(x) = \frac{1}{x}$ alors $f'(x) = -\frac{1}{x^2} < 0$ sur \mathbb{R}^* , mais f n'est pas monotone sur \mathbb{R}^* .

4.4.5 Dérivation de la bijection réciproque

À ce stade, les résultats sont admis.

Proposition 4.4.9 (réciproque d'une fonction continue strictement monotone)

Soit f une fonction numérique, continue et strictement monotone sur I .

Alors f réalise une bijection de I sur $J = f(I)$ (qui est un intervalle).

La bijection réciproque f^{-1} est continue et strictement monotone (avec la même monotonie que f).

On peut en dire un peu plus dans le cas où la fonction f est dérivable de dérivée partout strictement positive (ou partout strictement négative) sur I .

Proposition 4.4.10

Soit f une fonction numérique, dérivable sur l'intervalle I .

On suppose que $f'(x) > 0$ pour tout x de I , ou que $f'(x) < 0$ pour tout x de I .

Le résultat précédent s'applique. De plus f^{-1} est dérivable sur $J = f(I)$ et $(f^{-1})' = \frac{1}{f' \circ f^{-1}}$.

Autrement dit, pour tout a de I , et si on note $b = f(a)$, alors $(f^{-1})'(b) = \frac{1}{f'(a)}$.

Interprétation graphique

On suppose ici que le repère est orthonormé.

On note (Γ) le graphe de f , et (Γ') celui de f^{-1} .

(Γ) et (Γ') sont symétriques par rapport à $y = x$.

Soit $A(a, b = f(a))$ sur (Γ) , et $B(b, a = f^{-1}(b))$ sur (Γ') .

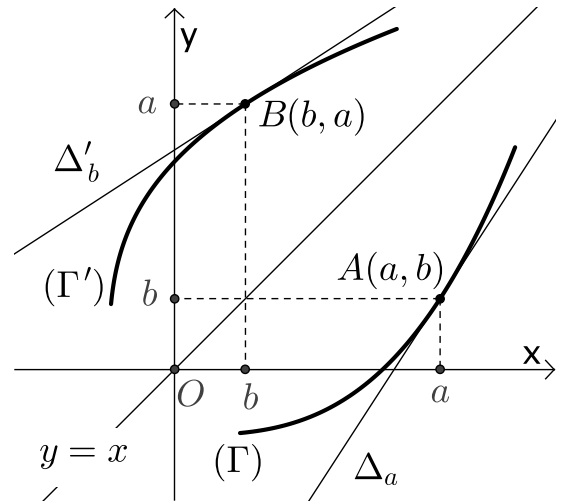
Soit Δ_a la tangente en A à (Γ) , Δ'_b la tangente en B à (Γ') .

La droite Δ_a a pour coefficient directeur $f'(a)$.

La droite Δ'_b a pour coefficient directeur $(f^{-1})'(b)$.

Δ_a et Δ'_b sont symétriques par rapport à $y = x$.

L'égalité $(f^{-1})'(b) = \frac{1}{f'(a)}$ traduit cette symétrie.



Une situation particulière

Avec les notations ci-dessus, supposons que f soit dérivable en a , mais qu'exceptionnellement $f'(a) = 0$. Alors la tangente Δ_a au point A de (Γ) est horizontale.

Par symétrie, il en découle que la courbe (Γ') présente une tangente *verticale* au point B .

La fonction f^{-1} n'est donc pas dérivable en b .

Exemples de fonctions réciproques

Les exemples avec $x \mapsto \sin x$, $x \mapsto \cos x$ et $x \mapsto \tan x$ seront étudiés dans la partie « trigonométrie ».

- La fonction $x \mapsto \exp(x)$ est bijective de \mathbb{R} sur \mathbb{R}^{+*} . Sa bijection réciproque est $x \mapsto \ln(x)$.
- Pour tout $\alpha > 0$, les fonctions $x \mapsto x^\alpha$ et $x \mapsto x^{1/\alpha}$ sont des bijections réciproques de \mathbb{R}^{+*} sur \mathbb{R}^{+*} .
- La restriction de $x \mapsto \sin(x)$ à $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ est bijective de $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ sur $[-1, 1]$.

La bijection réciproque, de $[-1, 1]$ sur $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, est notée $x \mapsto \arcsin(x)$ (*arc sinus de x*).

- La restriction de $x \mapsto \cos(x)$ à $[0, \pi]$ est bijective de $[0, \pi]$ sur $[-1, 1]$.

La bijection réciproque, de $[-1, 1]$ sur $[0, \pi]$, est notée $x \mapsto \arccos(x)$ (*arc cosinus de x*).

- La restriction de $x \mapsto \tan(x)$ à $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ est bijective de $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ sur \mathbb{R} .

La bijection réciproque, de \mathbb{R} sur $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$, est notée $x \mapsto \arctan(x)$ (*arc tangente de x*).

4.4.6 Dérivée seconde, concavité, inflexions

Définition 4.4.1

Soit $f : I \rightarrow \mathbb{R}$ une fonction numérique dérivable sur un intervalle I .

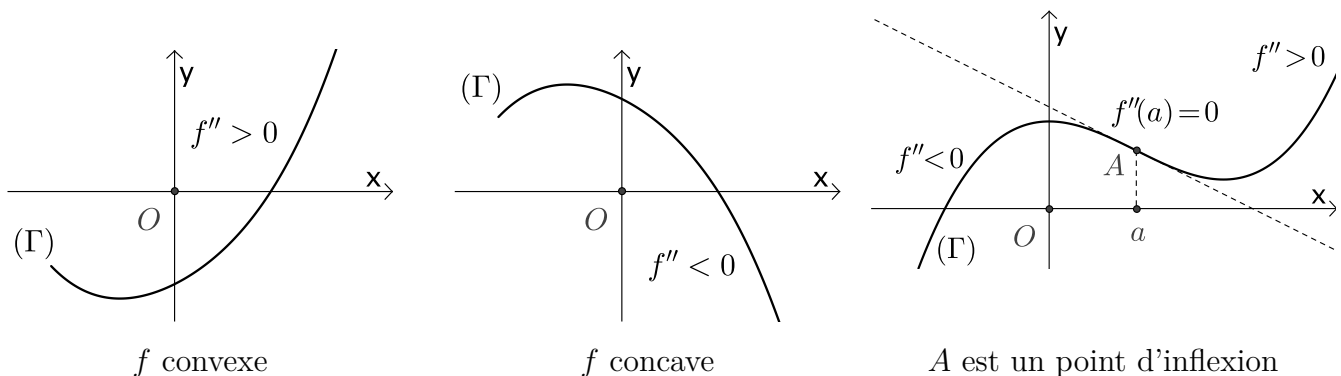
Si f' est dérivable sur I , on dit que f est *deux fois dérivable* sur I , et on note f'' plutôt que $(f')'$.

On dit que f'' est la *fonction dérivée seconde* de f .

Concavité

Soit $f : I \rightarrow \mathbb{R}$ une fonction numérique deux fois dérivable.

- Dire que $f'' \geq 0$ sur I , c'est dire que f' est croissante sur I .
Pour la courbe (Γ) cela se traduit par « une concavité tournée vers le haut ».
On exprimera cette situation en disant que f est une *fonction convexe* sur I .
- Dire que $f'' \leq 0$ sur I , c'est dire que f' est décroissante sur I .
Pour la courbe (Γ) cela se traduit par « une concavité tournée vers le bas ».
On exprimera cette situation en disant que f est une *fonction concave* sur I .
- Si la fonction f'' change de signe en un point a intérieur à l'intervalle I , cela se traduit par un « changement de concavité » au point $A(a, f(a))$ pour la courbe (Γ) .
En ce point la courbe (Γ) *traverse* sa tangente.
On exprime cette situation en disant que $A(a, f(a))$ est un *point d'inflexion* de (Γ) .



4.4.7 Dérivées d'ordre supérieur

Soit $f : I \rightarrow \mathbb{R}$ une fonction numérique, définie sur un intervalle I (si f est définie sur une réunion d'intervalles, on se limite à la restriction de f à l'un de ces intervalles).

Si f est dérivable sur I , la fonction f' est appelée *dérivée première* de f .

Si f' est à son tour dérivable sur I , la fonction $f'' = (f')'$ est appelée *dérivée seconde* de f .

Plus généralement, on est amené (si les propriétés de f le permettent) à définir les *fonctions dérivées successives* de f . Plus précisément :

Définition 4.4.2 (fonction n fois dérivable sur un intervalle)

Soit f une fonction numérique définie sur un intervalle de I .

Par convention, on pose $f^{(0)} = f$.

Soit k dans \mathbb{N} . On suppose que la fonction $f^{(k)} : I \rightarrow \mathbb{R}$ existe et qu'elle est dérivable sur I .

On note alors $f^{(k+1)}$ la fonction dérivée de $f^{(k)}$, c'est-à-dire : $f^{(k+1)} = (f^{(k)})'$.

Soit n dans \mathbb{N} . Si la fonction $f^{(n)} : I \rightarrow \mathbb{R}$ existe, on dit que f est n fois dérivable sur I .

On dit aussi que $f^{(n)}$ est la *fonction dérivée n -ième* de f sur I .

Quelques remarques

- On retiendra que $f^{(0)} = f$ (la dérivée « zéro-ième » de f) c'est f , et que $f^{(1)} = f'$.
On note souvent f'' (plutôt que $f^{(2)}$) et f''' (plutôt que $f^{(3)}$) les dérivées seconde et troisième de f .
À partir de la dérivée quatrième (donc si $k \geq 4$), on utilise obligatoirement la notation $f^{(k)}$.
- Si la fonction $f^{(n)} : I \rightarrow \mathbb{R}$ existe pour tout n , on dit que f est *indéfiniment dérivable* sur I .
- La fonction $f^{(n)}$ peut également être notée $D^n f$ ou encore $\frac{d^n f}{dx^n}$.
- Si f est n fois dérivable et si $f^{(n)}$ est p fois dérivable, alors f est $n+p$ fois dérivable et $(f^{(n)})^{(p)} = f^{(n+p)}$.

Proposition 4.4.11 (sommes de fonctions n fois dérivables)

Soit $f : I \rightarrow \mathbb{R}$ et $g : I \rightarrow \mathbb{R}$ deux fonctions n fois dérivables. Soit α et β deux réels.

Alors $\alpha f + \beta g$ est n fois dérivable sur I et : $(\alpha f + \beta g)^{(n)} = \alpha f^{(n)} + \beta g^{(n)}$.

Proposition 4.4.12 (produit de deux fonctions n fois dérivables)

Soit $f : I \rightarrow \mathbb{R}$ et $g : I \rightarrow \mathbb{R}$ deux fonctions n fois dérivables.

Alors la fonction produit fg est n fois dérivable sur I et on a : $(fg)^{(n)} = \sum_{k=0}^n \binom{n}{k} f^{(k)} g^{(n-k)}$.

Le résultat précédent (connu sous le nom de « formule de Leibniz ») dit par exemple que :

- La dérivée seconde de fg est : $(fg)'' = f''g + 2f'g' + fg''$
- La dérivée troisième de fg est : $(fg)''' = f'''g + 3f''g' + 3f'g'' + fg'''$
- La dérivée quatrième de fg est : $(fg)^{(4)} = f^{(4)}g + 4f'''g' + 6f''g'' + 4f'g''' + fg^{(4)}$

Proposition 4.4.13 (quotient de fonctions n fois dérivables)

Soit $f : I \rightarrow \mathbb{R}$ et $g : I \rightarrow \mathbb{R}$, n fois dérivables.

On suppose que g ne s'annule pas sur I . Alors les fonctions $\frac{1}{g}$ et $\frac{f}{g}$ sont n fois dérivables sur I .

Après $\left(\frac{1}{g}\right)' = -\frac{g'}{g^2}$, ça se complique : $\left(\frac{1}{g}\right)'' = -\left(\frac{g'}{g^2}\right)' = -(g'g^{-2})' = \frac{2(g')^2 - g''g}{g^3}$, etc.

Proposition 4.4.14 (composée de fonctions n fois dérivables)

Soit $f : I \rightarrow \mathbb{R}$ et $g : J \rightarrow \mathbb{R}$, deux fonctions n fois dérivables, avec $f(I) \subset J$.

Alors la fonction $g \circ f$ est n fois dérivable sur I .

Après $(g \circ f)' = f' \cdot (g' \circ f)$, ça se complique : $(g \circ f)'' = f'' \cdot (g' \circ f) + (f')^2 \cdot (g'' \circ f)$, etc.

Proposition 4.4.15 (réciproque d'une fonction n fois dérivable)

Soit $f : I \rightarrow \mathbb{R}$ une fonction numérique n fois dérivable.

On suppose que $f'(x) > 0$ pour tout x de I , ou que $f'(x) < 0$ pour tout x de I .

Alors la fonction réciproque f^{-1} est n fois dérivable sur l'intervalle $J = f(I)$.

Après $(f^{-1})' = \frac{1}{f' \circ f^{-1}}$, le calcul de $(f^{-1})''$ est plus compliqué (essayez).

4.5 Fonctions usuelles

4.5.1 Exponentielle, logarithme népérien

Définition 4.5.1 (fonction exponentielle)

Il existe une fonction unique $x \mapsto y(x)$, dérivable sur \mathbb{R} , et telle que :
$$\begin{cases} \forall x \in \mathbb{R}, y'(x) = y(x) \\ y(0) = 1 \end{cases}$$

On l'appelle la *fonction exponentielle*, et on la note $x \mapsto \exp(x)$.

Définition 4.5.2

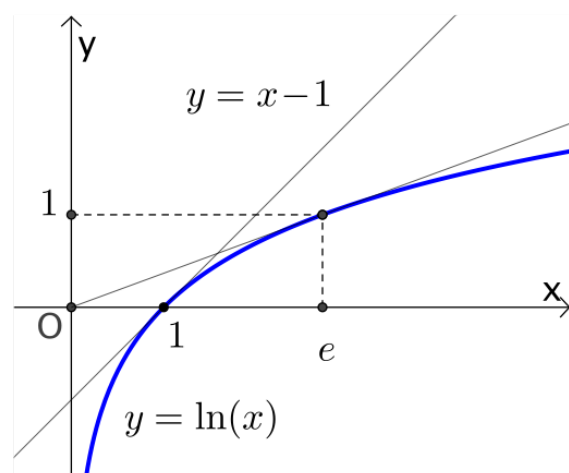
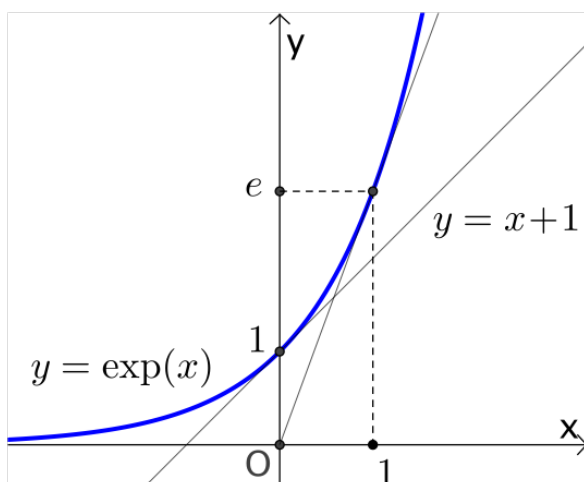
La fonction $x \mapsto \exp(x)$ est une bijection strictement croissante de \mathbb{R} sur \mathbb{R}^{+*} .

Sa bijection réciproque, de \mathbb{R}^{+*} sur \mathbb{R} , appelée *fonction logarithme népérien*, est notée $x \mapsto \ln x$.

Propriétés

- Par définition, on a l'équivalence :
$$\begin{cases} y = \exp(x) \\ x \in \mathbb{R} \end{cases} \Leftrightarrow \begin{cases} x = \ln(y) \\ y > 0 \end{cases}$$
- La fonction $x \mapsto \exp(x)$ est indéfiniment dérivable sur \mathbb{R} et : $\forall n \in \mathbb{N}, \exp^{(n)} = \exp$.
- La fonction $x \mapsto \ln(x)$ est définie sur \mathbb{R}^{+*} par : $\forall x > 0, \ln'(x) = \frac{1}{x}$ et $\ln(1) = 0$.
- La fonction $x \mapsto \ln(x)$ est strictement croissante et indéfiniment dérivable sur \mathbb{R}^{+*} .
- On note e l'unique réel strictement positif tel que $\ln(e) = 1$. On a : $e \approx 2.718281828$.
- Le logarithme népérien est l'unique primitive sur \mathbb{R}^{+*} , s'annulant en $x = 1$, de la fonction $x \mapsto \frac{1}{x}$.
En d'autres termes : $\forall x > 0, \ln(x) = \int_1^x \frac{dt}{t}$.
- La fonction $x \mapsto \exp(x)$ est convexe sur \mathbb{R} (sa dérivée seconde est $\exp(x) > 0$).
Pour tout x de \mathbb{R} , on a l'inégalité $\exp(x) \geq 1 + x$ (égalité $\Leftrightarrow x = 0$).
- La fonction $x \mapsto \ln(x)$ est concave (sa dérivée seconde est $-\frac{1}{x^2} < 0$).
Pour tout $x > 0$, on a l'inégalité $\ln(x) \leq x - 1$ (égalité $\Leftrightarrow x = 1$).

Courbes représentatives des fonctions $x \mapsto e^x$ et $x \mapsto \ln(x)$:



Propriétés fonctionnelles :

Pour tous x, y de \mathbb{R} : $\exp(x + y) = \exp(x) \exp(y)$, $\exp(-x) = \frac{1}{\exp(x)}$, $\exp(x - y) = \frac{\exp(x)}{\exp(y)}$.

Pour tous x, y de \mathbb{R}^{+*} : $\ln(xy) = \ln(x) + \ln(y)$, $\ln\left(\frac{1}{x}\right) = -\ln(x)$, $\ln\left(\frac{x}{y}\right) = \ln(x) - \ln(y)$.

Notation $x \mapsto e^x$

Pour tout n de \mathbb{N} , on a $\exp(n) = \exp(1)^n = e^n$.

Cette propriété se généralise aux exposants rationnels.

On décide d'étendre encore cette définition en posant : $\forall x \in \mathbb{R}$, $e^x = \exp(x)$.

On définit ainsi les puissances de e avec exposant réel quelconque.

Toutes les propriétés de la fonction exponentielle peuvent alors se réécrire en utilisant cette notation.

Limites usuelles :

Pour l'exponentielle :
$$\left\{ \begin{array}{lll} \lim_{x \rightarrow -\infty} \exp(x) = 0^+, & \lim_{x \rightarrow +\infty} \exp(x) = +\infty, & \lim_{x \rightarrow +\infty} \frac{\exp(x)}{x} = +\infty \\ \lim_{x \rightarrow -\infty} x \exp(x) = 0, & \lim_{x \rightarrow 0} \frac{\exp(x) - 1}{x} = 1 & \end{array} \right.$$

Pour le logarithme népérien :
$$\left\{ \begin{array}{lll} \lim_{x \rightarrow 0^+} \ln(x) = -\infty, & \lim_{+\infty} \ln(x) = +\infty, & \lim_{x \rightarrow 0^+} x \ln(x) = 0^- \\ \lim_{x \rightarrow +\infty} \frac{\ln(x)}{x} = 0^+, & \lim_{x \rightarrow 0} \frac{\ln(1+x)}{x} = 1, & \lim_{x \rightarrow 1} \frac{\ln(x)}{x-1} = 1 \end{array} \right.$$

4.5.2 Fonctions exponentielles de base quelconque

Définition 4.5.3 (fonctions exponentielles de base quelconque)

Pour tout réel $a > 0$, et pour tout réel x , on pose $a^x = \exp(x \ln(a))$.

La fonction $x \mapsto a^x$ est appelée *fonction exponentielle de base a* .

Propriétés des fonctions exponentielles de base quelconque

– Pour $a = e$, on retrouve la fonction $x \mapsto \exp(x)$, déjà notée $x \mapsto e^x$.

La fonction $x \mapsto \exp(x) = e^x$ est donc la fonction exponentielle de base e .

– La notation a^x étend la définition de a^r pour r rationnel.

– Pour tout réel $a > 0$, la fonction $x \mapsto a^x$ est définie et continue sur \mathbb{R} .

Elle est même indéfiniment dérivable : $\forall x \in \mathbb{R}$, $(a^x)' = (\ln(a))a^x$.

– La fonction $x \mapsto a^x$ est
$$\left\{ \begin{array}{l} \text{strictement croissante si } a > 1 \\ \text{strictement décroissante si } 0 < a < 1 \\ \text{constante égale à } 1 \text{ si } a = 1 \end{array} \right.$$

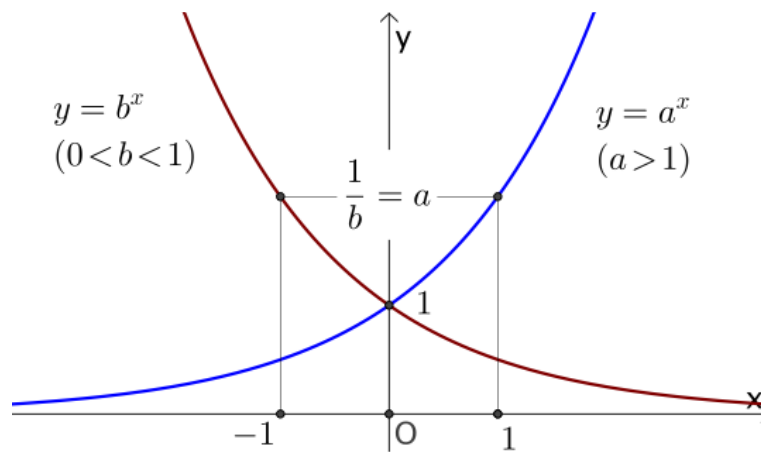
– Si $a \neq 1$, la fonction $x \mapsto a^x$ réalise une bijection de \mathbb{R} sur \mathbb{R}^{+*} .

La bijection réciproque est $x \mapsto \log_a x = \frac{\ln(x)}{\ln(a)}$ appelée fonction logarithme de base a .

Ainsi la fonction logarithme de base 10 est définie sur \mathbb{R}^{+*} par $\log_{10} x = \log x = \frac{\ln(x)}{\ln(10)}$ et elle est la bijection réciproque de la fonction $x \mapsto 10^x$.

– Pour tout x de \mathbb{R} et tout $a > 0$, on a $\left(\frac{1}{a}\right)^x = a^{-x}$. Les courbes représentatives de $x \mapsto a^x$ et $x \mapsto \left(\frac{1}{a}\right)^x$ sont donc symétriques l'une de l'autre par rapport à l'axe des ordonnées.

– Courbes représentatives des fonctions $x \mapsto a^x$ et $x \mapsto b^x$, avec $a > 1$ et $b = \frac{1}{a}$:



Propriétés fonctionnelles

Pour tous x, y de \mathbb{R} , pour tout a, b de \mathbb{R}^{+*} , on a :

$$\begin{cases} a^{x+y} = a^x a^y, & a^{-x} = \frac{1}{a^x}, & a^{x-y} = \frac{a^x}{a^y} \\ (a^x)^y = a^{xy}, & a^x b^x = (ab)^x, & \frac{a^x}{b^x} = \left(\frac{a}{b}\right)^x \end{cases}$$

Limites usuelles : $\lim_{x \rightarrow -\infty} a^x = \begin{cases} 0 & \text{si } a > 1 \\ +\infty & \text{si } 0 < a < 1 \end{cases}$ $\lim_{x \rightarrow +\infty} a^x = \begin{cases} +\infty & \text{si } a > 1 \\ 0 & \text{si } 0 < a < 1 \end{cases}$

4.5.3 Fonctions puissances $x \mapsto x^\alpha$, avec α réel

Définition 4.5.4 (fonctions puissances)

Soit α un nombre réel quelconque.

On appelle *fonction puissance d'exposant α* la fonction définie sur \mathbb{R}^{+*} par $x \mapsto x^\alpha = \exp(\alpha \ln(x))$.

Propriétés

– Quand l'exposant α est entier ou rationnel, cette définition de la fonction $x \mapsto x^\alpha$ est compatible avec celle qu'on connaissait déjà (sur un domaine parfois plus large que \mathbb{R}^{+*}).

– La dérivée de $x \mapsto x^\alpha$ est $x \mapsto \alpha x^{\alpha-1}$.

Sur son domaine \mathbb{R}^{+*} , la fonction $x \mapsto x^\alpha$ est $\begin{cases} \text{strictement croissante si } \alpha > 0 \\ \text{strictement décroissante si } \alpha < 0 \\ \text{constante en 1 si } \alpha = 0 \end{cases}$

- Si $\alpha \neq 0$, la fonction $x \mapsto x^\alpha$ est une bijection de \mathbb{R}^{+*} sur lui-même.
Sa bijection réciproque est la fonction $x \mapsto x^{1/\alpha}$
- Si $\alpha > 0$, $x \mapsto x^\alpha$ est prolongeable par continuité à l'origine en lui donnant la valeur 0.
En $(0, 0)$, la courbe présente alors une tangente horizontale si $\alpha > 1$ et verticale si $0 < \alpha < 1$.
Toutes les courbes représentatives des fonctions $x \mapsto x^\alpha$ passent par le point $(1, 1)$.
- Le placement des différentes courbes est le suivant :

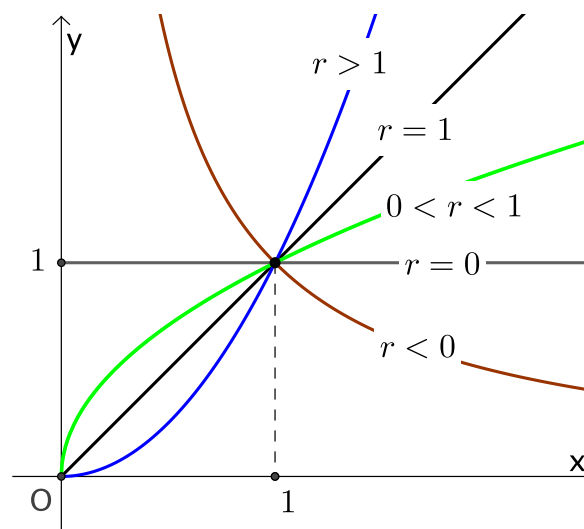
$$\forall x > 0, \forall (\alpha, \beta) \in \mathbb{R}^2, \text{ avec } \alpha < \beta : \begin{cases} \text{Si } 0 < x < 1 \text{ alors } x^\alpha > x^\beta \\ \text{Si } x > 1 \text{ alors } x^\alpha < x^\beta \end{cases}$$

Ainsi, les puissances sont dans l'ordre des exposants pour $x > 1$, et dans l'ordre contraire avant.

Propriétés fonctionnelles

$$\text{Pour tous } x, y \text{ de } \mathbb{R}^{+*}, \text{ pour tout } \alpha, \beta \text{ de } \mathbb{R}, \text{ on a : } \begin{cases} x^{\alpha+\beta} = x^\alpha x^\beta, & x^{-\alpha} = \frac{1}{x^\alpha}, & x^{\alpha-\beta} = \frac{x^\alpha}{x^\beta} \\ (x^\alpha)^\beta = x^{\alpha\beta}, & x^\alpha y^\alpha = (xy)^\alpha, & \frac{x^\alpha}{y^\alpha} = \left(\frac{x}{y}\right)^\alpha \end{cases}$$

Courbes représentatives des fonctions $x \mapsto x^r$



Croissances comparées

Les limites suivantes sont importantes à connaître. Elles constituent une « échelle de comparaison » entre fonctions puissances, fonction exponentielle, et fonction logarithme :

$$\begin{cases} \forall \alpha \in \mathbb{R}, \forall \beta > 0 : & \lim_{x \rightarrow -\infty} |x|^\alpha \exp^\beta(x) = 0 & \lim_{x \rightarrow +\infty} \frac{\exp^\beta(x)}{x^\alpha} = +\infty \\ \forall \alpha > 0, \forall \beta > 0 : & \lim_{x \rightarrow 0^+} x^\alpha |\ln(x)|^\beta = 0 & \lim_{x \rightarrow +\infty} \frac{\ln^\beta(x)}{x^\alpha} = 0 \end{cases}$$

On peut compléter les résultats précédents avec les fonctions exponentielles de base $a > 0$.

Tout dépend en fait de la position de a par rapport à 1.

$$\begin{cases} \forall m > 0, \forall a > 1 : & \lim_{x \rightarrow -\infty} |x|^m a^x = 0 & \lim_{x \rightarrow +\infty} \frac{a^x}{x^m} = +\infty \\ \forall m > 0, \forall a \in]0, 1[: & \lim_{x \rightarrow -\infty} \frac{a^x}{|x|^m} = +\infty & \lim_{x \rightarrow +\infty} x^m a^x = 0 \end{cases}$$

4.5.4 Dérivée logarithmique

Si x, y sont deux réels non nuls et de même signe, alors $\ln(xy) = \ln|x| + \ln|y|$.

En particulier, pour tout $x \neq 0$, on a : $\ln(x^2) = 2 \ln(|x|)$.

La fonction $x \mapsto \ln|x|$ est définie sur \mathbb{R}^* et sa dérivée est $x \mapsto \frac{1}{x}$.

Définition 4.5.5

Soit f une fonction dérivable sur un intervalle I , à valeurs dans \mathbb{R}^* .

On appelle *dérivée logarithmique* de f la dérivée $\frac{f'}{f}$ de la fonction $\ln|f|$.

Proposition 4.5.1

Soient f_1, f_2, \dots, f_n des fonctions dérivables et strictement positives sur l'intervalle I .

Soient $\alpha_1, \alpha_2, \dots, \alpha_n$ des réels, et $g = f_1^{\alpha_1} f_2^{\alpha_2} \dots f_n^{\alpha_n}$.

Alors la dérivée logarithmique de g est $\frac{g'}{g} = \alpha_1 \frac{f_1'}{f_1} + \alpha_2 \frac{f_2'}{f_2} + \dots + \alpha_n \frac{f_n'}{f_n}$.

La dérivée logarithmique s'avère donc un moyen commode de calculer la dérivée d'une fonction qui s'exprime essentiellement à l'aide de quotients, de produits ou de puissances.

Exemple d'utilisation de la dérivée logarithmique

Soit par exemple $f : x \mapsto \sqrt{|x(x+2)|} \exp\left(\frac{1}{x}\right)$.

Les théorèmes généraux ne permettent pas de se prononcer pour les valeurs $x = -2$ et $x = 0$.

En revanche, ils permettent de dire que f est dérivable sur (au moins) $\mathbb{R} \setminus \{-2, 0\}$.

Pour tout x de $\mathbb{R} \setminus \{-2, 0\}$, on a : $\ln(f(x)) = \frac{1}{2} \ln|x(x+2)| + \frac{1}{x}$.

En dérivant, on obtient : $\frac{f'(x)}{f(x)} = \frac{x+1}{x(x+2)} - \frac{1}{x^2} = \frac{x^2-2}{x^2(x+2)}$. Ainsi $f' = \frac{x^2-2}{x^2(x+2)} f$.

En dérivant à nouveau sur $\mathbb{R} \setminus \{-2, 0\}$, on trouve l'expression de f'' :

$$\begin{aligned} f''(x) &= \frac{x^2-2}{x^2(x+2)} f'(x) + \frac{-x^4+6x^2+8x}{x^4(x+2)^2} f(x) \\ &= \frac{(x^2-2)^2 + (-x^4+6x^2+8x)}{x^4(x+2)^2} f(x) = \frac{2(x^2+4x+2)}{x^4(x+2)^2} f(x) \end{aligned}$$

On pourra comparer ce calcul de f'' avec celui obtenu par les méthodes habituelles de dérivation (où la présence d'une valeur absolue n'arrange rien).

4.5.5 Fonctions circulaires réciproques

Les fonctions sinus, cosinus, tangente, ont été vues dans “Nombres complexes et trigonométrie”.

Définition 4.5.6 (fonction arcsin)

La restriction à $I = \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ de $x \mapsto \sin(x)$ est une bijection de I sur $J = [-1, 1]$.

La bijection réciproque est notée $x \mapsto \arcsin(x)$ (fonction “arc sinus”).

Propriétés

– La fonction $x \mapsto \arcsin(x)$ est une bijection de $[-1, 1]$ sur $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$.

Elle est continue, strictement croissante, et impaire.

– Pour tout x de $[-1, 1]$, $\arcsin(x)$ est l'angle compris entre $-\frac{\pi}{2}$ et $\frac{\pi}{2}$ dont le sinus est égal à x :

$$\begin{cases} y = \arcsin(x) \\ x \in [-1, 1] \end{cases} \Leftrightarrow \begin{cases} x = \sin(y) \\ y \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \end{cases}$$

– Quelques valeurs particulières :

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

– Pour tout x de $[-1, 1]$, $\sin(\arcsin(x)) = x$.

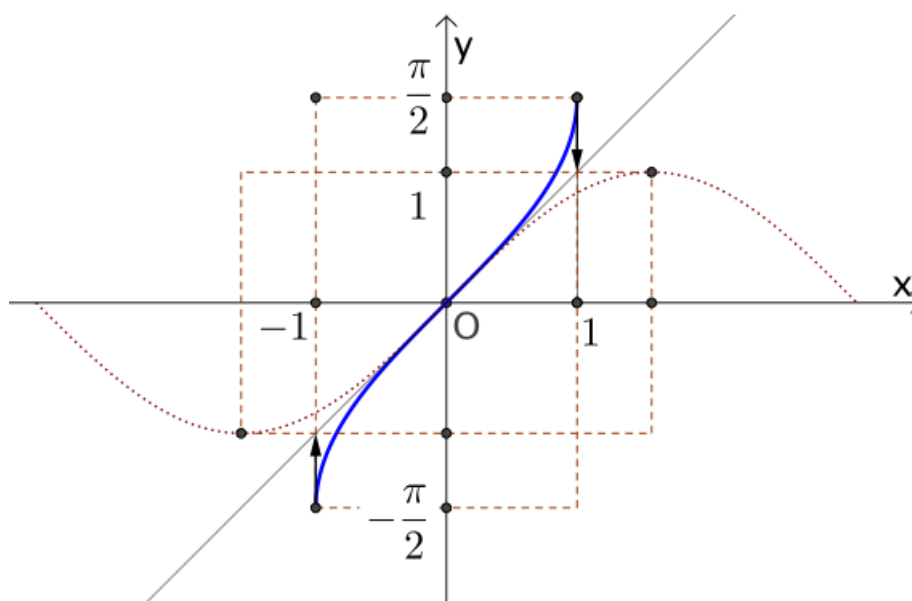
Pour tout x de $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, $\arcsin(\sin(x)) = x$ (attention au domaine!)

Pour tout x de $[-1, 1]$, $\cos(\arcsin(x)) = \sqrt{1-x^2}$.

Pour tout x de $] -1, 1[$, $\tan(\arcsin(x)) = \frac{x}{\sqrt{1-x^2}}$.

– Dérivée : pour tout x de $] -1, 1[$, $\arcsin'(x) = \frac{1}{\sqrt{1-x^2}}$.

Courbe représentative de la fonction $x \mapsto \arcsin(x)$:



Définition 4.5.7 (fonction arccos)

La restriction à $I = [0, \pi]$ de $x \mapsto \cos(x)$ est une bijection de I sur $J = [-1, 1]$.

La bijection réciproque est notée $x \mapsto \arccos(x)$ (fonction "arc cosinus").

Propriétés

– La fonction $x \mapsto \arccos(x)$ est une bijection de $[-1, 1]$ sur $[0, \pi]$.

Elle est continue et strictement décroissante.

– Pour tout x de $[-1, 1]$, $\arccos(x)$ est l'angle compris entre 0 et π dont le cosinus est égal à x :

$$\begin{cases} y = \arccos(x) \\ x \in [-1, 1] \end{cases} \Leftrightarrow \begin{cases} x = \cos(y) \\ y \in [0, \pi] \end{cases}$$

– Quelques valeurs particulières :

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

– Pour tout x de $[-1, 1]$, $\cos(\arccos(x)) = x$.

Pour tout x de $[0, \pi]$, $\arccos(\cos(x)) = x$ (attention au domaine!)

Pour tout x de $[-1, 1]$, $\sin(\arccos(x)) = \sqrt{1-x^2}$.

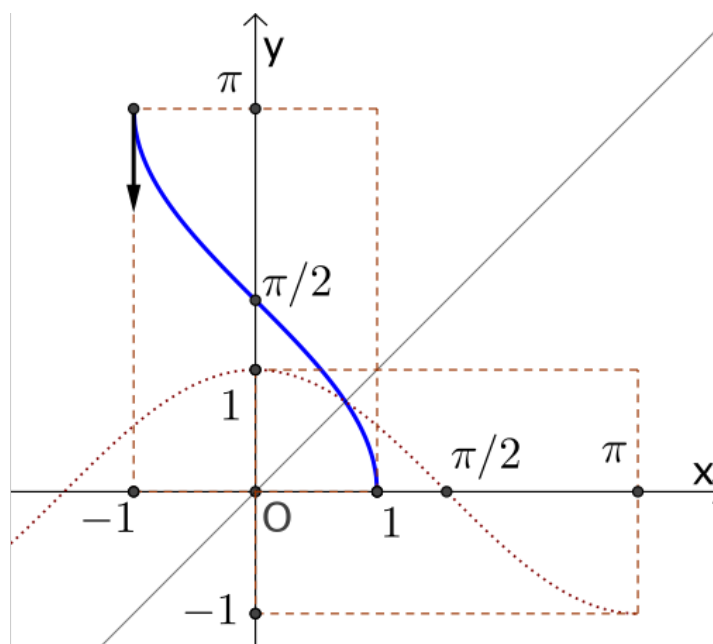
Pour tout x de $[-1, 0[\cup]0, 1]$, $\tan(\arccos(x)) = \frac{\sqrt{1-x^2}}{x}$.

– Pour tout x de $[-1, 1]$, $\arccos(-x) + \arccos(x) = \pi$.

Pour tout x de $[-1, 1]$, $\arcsin(x) + \arccos(x) = \frac{\pi}{2}$.

– Dérivée : pour tout x de $] -1, 1[$, $\arccos'(x) = -\frac{1}{\sqrt{1-x^2}}$.

Courbe représentative de la fonction $x \mapsto \arccos(x)$:



Définition 4.5.8 (fonction arctan)

La restriction à $I = \left] -\frac{\pi}{2}, \frac{\pi}{2} \right[$ de $x \mapsto \tan(x)$ est une bijection de I sur \mathbb{R} .

La bijection réciproque est notée $x \mapsto \arctan(x)$ (fonction “arc tangente”).

Propriétés

– La fonction $x \mapsto \arctan(x)$ est une bijection de \mathbb{R} sur $\left] -\frac{\pi}{2}, \frac{\pi}{2} \right[$.

Elle est continue, strictement croissante, et impaire.

– Pour tout x réel, $\arctan(x)$ est l'angle de $\left] -\frac{\pi}{2}, \frac{\pi}{2} \right[$ dont la tangente est égale à x :

$$\begin{cases} y = \arctan(x) \\ x \in \mathbb{R} \end{cases} \Leftrightarrow \begin{cases} x = \tan(y) \\ y \in \left] -\frac{\pi}{2}, \frac{\pi}{2} \right[\end{cases}$$

– Quelques valeurs particulières :

x	0	$\frac{\sqrt{3}}{3}$	1	$\sqrt{3}$
$\arctan(x)$	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$

– Pour tout x de \mathbb{R} , $\tan(\arctan(x)) = x$.

Pour tout x de $\left] -\frac{\pi}{2}, \frac{\pi}{2} \right[$, $\arctan(\tan(x)) = x$ (attention au domaine!).

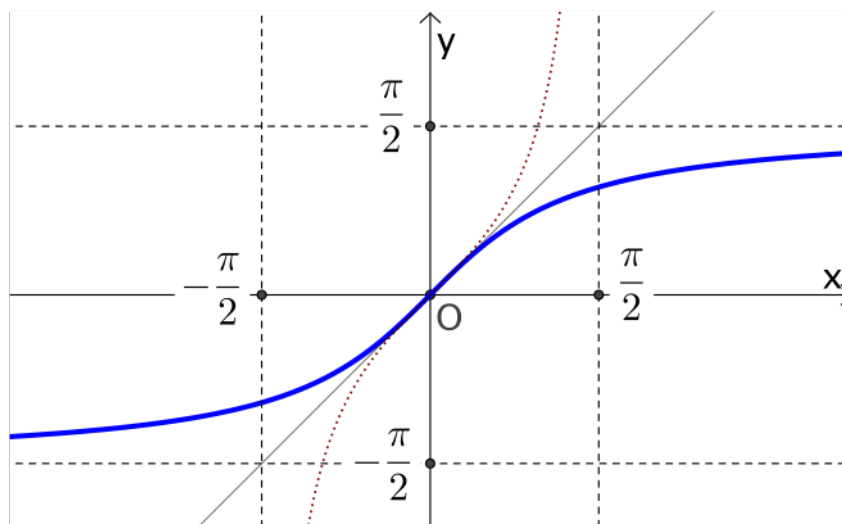
Pour tout x de \mathbb{R} , $\cos(\arctan(x)) = \frac{1}{\sqrt{1+x^2}}$.

Pour tout x de \mathbb{R} , $\sin(\arctan(x)) = \frac{x}{\sqrt{1+x^2}}$.

– Pour tout x de \mathbb{R}^* , $\arctan(x) + \arctan\left(\frac{1}{x}\right) = \varepsilon \frac{\pi}{2}$, avec $\varepsilon = \begin{cases} -1 & \text{si } x < 0 \\ +1 & \text{si } x > 0 \end{cases}$

– Dérivée : pour tout x de \mathbb{R} , $\arctan'(x) = \frac{1}{1+x^2}$.

Courbe représentative de la fonction $x \mapsto \arctan(x)$:



4.5.6 Fonctions hyperboliques

Définition 4.5.9 (fonctions $x \mapsto \text{sh}(x)$ et $x \mapsto \text{ch}(x)$)

Pour tout x de \mathbb{R} , on pose $\text{ch}(x) = \frac{e^x + e^{-x}}{2}$ (fonction “cosinus hyperbolique”)

Pour tout x de \mathbb{R} , on pose $\text{sh}(x) = \frac{e^x - e^{-x}}{2}$ (fonction “sinus hyperbolique”)

Propriétés

– Les fonctions $x \mapsto \text{ch}(x)$ et $x \mapsto \text{sh}(x)$ sont indéfiniment dérivables sur \mathbb{R} .

Pour tout x de \mathbb{R} , on a $\text{sh}'(x) = \text{ch}(x)$ et $\text{ch}'(x) = \text{sh}(x)$.

Les deux fonctions $x \mapsto y = \text{ch}(x)$ et $x \mapsto y = \text{sh}(x)$ sont donc solutions de $y'' = y$.

La fonction $x \mapsto \text{ch}(x)$ est paire, et la fonction $x \mapsto \text{sh}(x)$ est impaire.

$$- \forall x \in \mathbb{R}, \begin{cases} \text{ch}(x) + \text{sh}(x) = e^x \\ \text{ch}(x) - \text{sh}(x) = e^{-x} \end{cases}, \begin{cases} \text{ch}(x) \geq 1 \\ \text{ch}^2(x) - \text{sh}^2(x) = 1 \end{cases}$$

$$- \forall x \geq 0, \forall y \in \mathbb{R}, x^2 - y^2 = 1 \Leftrightarrow \begin{cases} \exists t \in \mathbb{R} \\ \text{ch}(t) = x, \text{sh}(t) = y \end{cases}$$

La fonction $t \mapsto (\text{ch } t, \text{sh } t)$ est un paramétrage de l'arc d'hyperbole $\begin{cases} x^2 - y^2 = 1 \\ x \geq 0 \end{cases}$

– Au voisinage de l'origine, on a : $\text{sh}(x) \sim x$ (la droite $y = x$ est tangente d'inflexion).

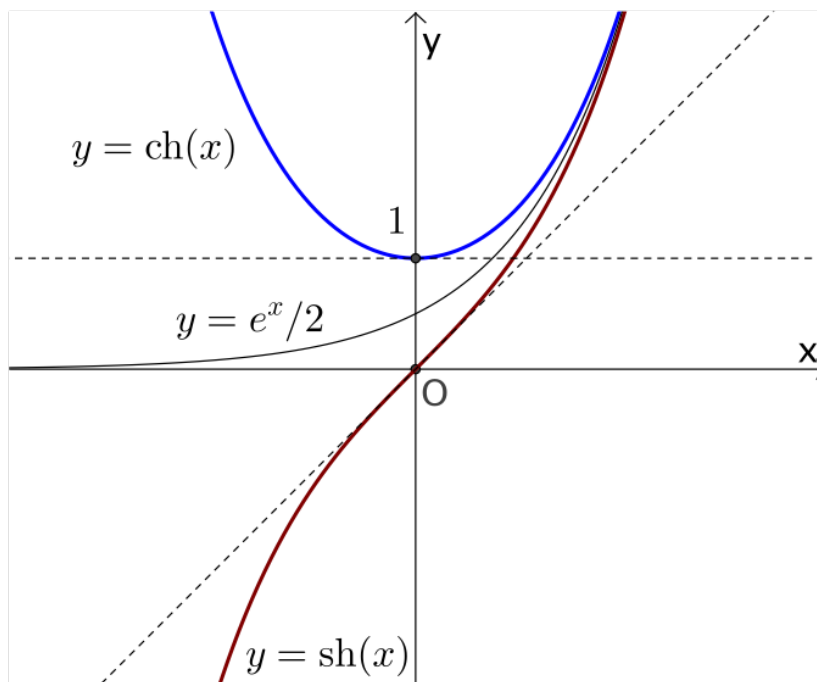
Toujours au voisinage de 0, on a l'équivalent : $\text{ch}(x) - 1 \sim \frac{x^2}{2}$.

– Au voisinage de $+\infty$, on a : $\text{ch}(x) \sim \text{sh}(x) \sim \frac{e^x}{2}$.

Les deux courbes $y = \text{ch}(x)$ et $y = \text{sh}(x)$ sont asymptotes à $y = \frac{e^x}{2}$ (avec $\text{sh}(x) < \frac{e^x}{2} < \text{ch}(x)$)

– Au voisinage de $-\infty$, on a : $\text{ch}(x) \sim \frac{e^{-x}}{2}$ et $\text{sh}(x) \sim -\frac{e^{-x}}{2}$.

Courbes représentatives des fonctions $x \mapsto \operatorname{ch}(x)$ et $x \mapsto \operatorname{sh}(x)$:



Définition 4.5.10 (fonction $x \mapsto \operatorname{th}(x)$)

Pour tout x de \mathbb{R} , on pose $\operatorname{th}(x) = \frac{\operatorname{sh}(x)}{\operatorname{ch}(x)}$ (fonction “tangente hyperbolique”)

Propriétés

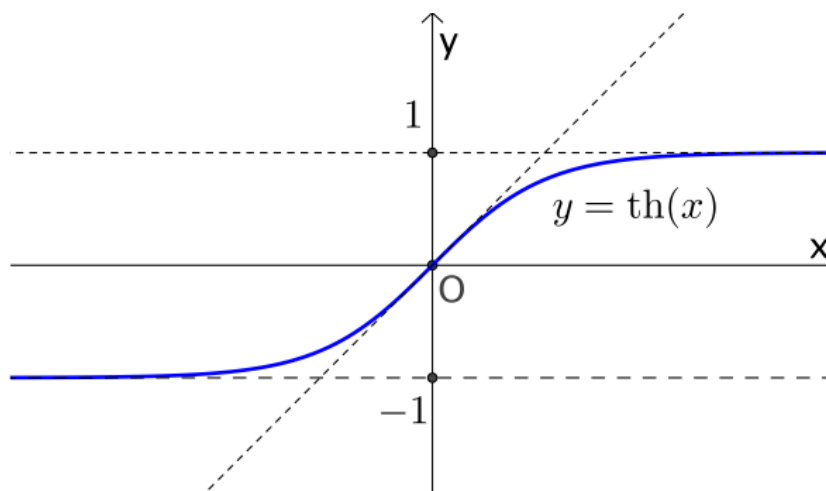
– La fonction $x \mapsto \operatorname{th}(x)$ est impaire. $\lim_{x \rightarrow +\infty} \operatorname{th}(x) = 1$.

Au voisinage de 0, on a : $\operatorname{th}(x) \sim x$ (la droite $y = x$ est tangente d’inflexion).

– La fonction $x \mapsto \operatorname{th}(x)$ est indéfiniment dérivable : $\forall x \in \mathbb{R}, \operatorname{th}'(x) = 1 - \operatorname{th}^2(x) = \frac{1}{\operatorname{ch}^2(x)}$.

– Pour tout x de \mathbb{R} , on a : $\operatorname{th}(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} = \frac{e^{2x} - 1}{e^{2x} + 1} = \frac{1 - e^{-2x}}{1 + e^{-2x}}$, et $|\operatorname{th}(x)| \leq 1$.

Courbe représentative de la fonction $x \mapsto \operatorname{th}(x)$:



4.5.7 Trigonométrie hyperbolique

– ch, sh et th d'une somme ou d'une différence :

$$\begin{cases} \text{ch}(x+y) = \text{ch}(x)\text{ch}(y) + \text{sh}(x)\text{sh}(y) \\ \text{ch}(x-y) = \text{ch}(x)\text{ch}(y) - \text{sh}(x)\text{sh}(y) \\ \text{sh}(x+y) = \text{sh}(x)\text{ch}(y) + \text{ch}(x)\text{sh}(y) \\ \text{sh}(x-y) = \text{sh}(x)\text{ch}(y) - \text{ch}(x)\text{sh}(y) \end{cases} \quad \begin{cases} \text{ch}(2x) = 2\text{ch}^2(x) - 1 = 1 + 2\text{sh}^2(x) \\ \text{sh}(2x) = 2\text{sh}(x)\text{ch}(x) \end{cases}$$

$$\text{th}(x+y) = \frac{\text{th}(x) + \text{th}(y)}{1 + \text{th}(x)\text{th}(y)}, \quad \text{th}(x-y) = \frac{\text{th}(x) - \text{th}(y)}{1 - \text{th}(x)\text{th}(y)}, \quad \text{th}2x = \frac{2\text{th}(x)}{1 + \text{th}^2(x)}$$

– Transformations de produits en sommes et de sommes en produits.

$$\begin{cases} \text{ch}(x)\text{ch}(y) = \frac{1}{2}(\text{ch}(x+y) + \text{ch}(x-y)) \\ \text{sh}(x)\text{sh}(y) = \frac{1}{2}(\text{ch}(x+y) - \text{ch}(x-y)) \\ \text{sh}(x)\text{ch}(y) = \frac{1}{2}(\text{sh}(x+y) + \text{sh}(x-y)) \\ \text{ch}^2(x) = \frac{1}{2}(1 + \text{ch}(2x)) \quad \text{et} \quad \text{sh}^2(x) = \frac{1}{2}(\text{ch}(2x) - 1) \end{cases} \quad \begin{cases} \text{ch}p + \text{ch}q = 2\text{ch}\left(\frac{p+q}{2}\right)\text{ch}\left(\frac{p-q}{2}\right) \\ \text{ch}p - \text{ch}q = 2\text{sh}\left(\frac{p+q}{2}\right)\text{sh}\left(\frac{p-q}{2}\right) \\ \text{sh}p + \text{sh}q = 2\text{sh}\left(\frac{p+q}{2}\right)\text{ch}\left(\frac{p-q}{2}\right) \\ \text{sh}p - \text{sh}q = 2\text{sh}\left(\frac{p-q}{2}\right)\text{ch}\left(\frac{p+q}{2}\right) \end{cases}$$

– Changement de variable $t = \text{th}\left(\frac{x}{2}\right)$: $\text{ch}(x) = \frac{1+t^2}{1-t^2}$, $\text{sh}(x) = \frac{2t}{1-t^2}$, $\text{th}(x) = \frac{2t}{1+t^2}$

– Changement de variable $u = e^x$: $\text{ch}(x) = \frac{u^2+1}{2u}$, $\text{sh}(x) = \frac{u^2-1}{2u}$, $\text{th}(x) = \frac{u^2-1}{u^2+1}$

– Linéarisation.

On écrit $\text{ch}^n(x) = \left(\frac{e^x + e^{-x}}{2}\right)^n$ et $\text{sh}^n(x) = \left(\frac{e^x - e^{-x}}{2}\right)^n$.

On développe (formule du binôme), on groupe les termes équidistants des extrémités, et on réutilise les définitions pour retrouver des $\text{ch}(px)$ et/ou des $\text{sh}(px)$. Par exemple :

$$\text{sh}^4(x) = \left(\frac{e^x - e^{-x}}{2}\right)^4 = \frac{1}{16} (e^{4x} - 4e^{2x} + 6 - 4e^{-2x} + e^{-4x}) = \frac{1}{8} (\text{ch}(4x) - 4\text{ch}(2x) + 3)$$

$$\text{sh}^5(x) = \left(\frac{e^x - e^{-x}}{2}\right)^5 = \frac{1}{32} (e^{5x} - 5e^{3x} + 10e^x - 10e^{-x} + 5e^{-3x} - e^{-5x})$$

$$= \frac{1}{16} (\text{sh}(5x) - 5\text{sh}(3x) + 10\text{sh}(x))$$

– Opération inverse de la linéarisation.

$$\forall x \in \mathbb{R}, \forall n \in \mathbb{N}, (\text{ch}(x) + \text{sh}(x))^n = (e^x)^n = e^{nx} = \text{ch}(nx) + \text{sh}(nx)$$

On peut ainsi exprimer $\text{ch}(nx)$, $\text{sh}(nx)$ en fonction de puissances de $\text{ch}(x)$ et/ou de $\text{sh}(x)$.

Pour cela on développe $(\text{ch}(x) + \text{sh}(x))^n$ par la formule du binôme.

La partie paire (resp. impaire) du résultat est alors égale à $\text{ch}(nx)$ (resp. $\text{sh}(nx)$).

$$(\text{ch}(x) + \text{sh}(x))^4 = \text{ch}^4(x) + 4\text{ch}^3(x)\text{sh}(x) + 6\text{ch}^2(x)\text{sh}^2(x) + 4\text{ch}(x)\text{sh}^3(x) + \text{sh}^4(x)$$

$$\Rightarrow \begin{cases} \text{ch}(4x) = \text{ch}^4(x) + 6\text{ch}^2(x)\text{sh}^2(x) + \text{sh}^4(x) \\ \text{sh}(4x) = 4\text{ch}^3(x)\text{sh}(x) + 4\text{ch}(x)\text{sh}^3(x) \end{cases}$$

$$\Rightarrow \begin{cases} \text{ch}(4x) = \text{ch}^4(x) + 6\text{ch}^2(x)(\text{ch}^2(x) - 1) + (\text{ch}^2(x) - 1)^2 \\ \text{sh}(4x) = 4\text{ch}(x)((1 + \text{sh}^2(x))\text{sh}(x) + \text{sh}^3(x)) \end{cases} \Rightarrow \begin{cases} \text{ch}(4x) = 8\text{ch}^4(x) - 8\text{ch}^2(x) + 1 \\ \text{sh}(4x) = 4\text{ch}(x)(2\text{sh}^3(x) + \text{sh}(x)) \end{cases}$$

– Liens entre la trigonométrie hyperbolique et la trigonométrie circulaire.

Les formules de la trigonométrie hyperbolique peuvent être retrouvées à partir de celles de la trigonométrie circulaire, avec : $\cos(ix) = \operatorname{ch}(x)$, $\sin(ix) = i \operatorname{sh}(x)$, $\tan(ix) = i \operatorname{th}(x)$.

Par exemple :

$$\begin{aligned}\sin^3(x) &= \frac{1}{4}(-\sin(3x) + 3\sin(x)) \Rightarrow \sin^3(ix) = \frac{1}{4}(-\sin(3ix) + 3\sin(ix)) \\ &\Rightarrow -i \operatorname{sh}^3(x) = \frac{1}{4}(-i \operatorname{sh}(3x) + 3i \operatorname{sh}(x)) \Rightarrow \operatorname{sh}^3(x) = \frac{1}{4}(\operatorname{sh}(3x) - 3\operatorname{sh}(x))\end{aligned}$$

4.6 Études de fonctions, inégalités

4.6.1 Plan d'étude d'une fonction numérique

L'« étude » d'une fonction numérique f consiste en général en les étapes suivantes :

- Préciser le domaine de définition \mathcal{D} et la dérivabilité de f (sur \mathcal{D} , sauf peut-être en des points isolés).
- Si f est périodique, ou si on devine un axe ou un centre de symétrie (du fait notamment de la parité ou de l'imparité de f), on en profite pour réduire « le domaine d'étude ».
Si on ne devine rien, il est prudent de dire « pas de réduction évidente du domaine d'étude ».
- Étudier le sens de variation de f , et dresser le *tableau de variations*.
On pourra compléter ce tableau par les « limites aux bornes du domaine ».
- Effectuer les études locales pour une compréhension fine du comportement de f en certains points : ceux par exemple où la dérivabilité de f pose problème, ou encore les *branches infinies*.
- Tracer **soigneusement** la courbe représentative.

Nous allons maintenant revenir en détail sur quelques-unes de ces étapes.

4.6.2 Dérivabilité sur le domaine de définition

On commence toujours par préciser le domaine de définition \mathcal{D} de f .

Il s'agit en général d'un intervalle, ou d'une réunion d'intervalles.

On indique ensuite sur quelle partie de ce domaine on peut appliquer les résultats généraux portant sur les opérations entre fonctions usuelles, et donc conclure à la continuité et/ou à la dérivabilité de f .

On est d'ailleurs souvent amené à affirmer directement que f est indéfiniment dérivable sur son domaine (ou sur une partie de celui-ci) en vertu de ces mêmes résultats.

À ce stade, il est possible que certains points isolés « posent problème » (on ne peut appliquer les résultats généraux). Cela ne veut pas dire, pour autant, que f ne sera pas dérivable en ces points, et il conviendra de répondre (plus tard) à cette question par des « études locales ».

4.6.3 Réduction du domaine d'étude

Connaissant le domaine de définition \mathcal{D} de f , on détermine le *domaine d'étude* \mathcal{D}_e de f , c'est-à-dire la partie de \mathcal{D} sur laquelle il suffit d'étudier f pour connaître son comportement global.

C'est ici le moment d'indiquer si la fonction f est T -périodique (ce qui permet de réduire l'étude à un intervalle de longueur T), et si elle est paire ou impaire (pour se limiter alors à $x \geq 0$).

Plus exceptionnellement, il se peut que f vérifie des égalités $f(2a - x) = f(x)$ (axe de symétrie $x = a$) ou $f(2a - x) = -f(x)$ (centre de symétrie $(a, 0)$) ou $f(2a - x) = 2b - f(x)$ (centre de symétrie (a, b)). Dans ce cas, on limite l'étude à $x \geq a$ (ou à $x \leq a$).

Il arrive parfois que f soit paire (ou impaire) et T -périodique, auquel cas on pourra se contenter de l'étude de f sur une *demi-période* (souvent l'intervalle $[0, T/2]$).

Si T est une période de f , on vérifiera qu'elle est la « plus petite période » (essayer $T/2$ par exemple).

Il arrive parfois que si T est une période f , on ait $f(x + T/2) = -f(x)$, ou $f(T - x) = f(x)$. L'étude peut alors être réduite à un intervalle de longueur $T/2$.

4.6.4 Tableau des variations

Connaissant le domaine d'étude \mathcal{D}_e de f , on examine le sens de variations de f .

Cela consiste à dire, intervalle par intervalle, si f est croissante ou décroissante, et à signaler les extrêmes *relatifs* (on dit aussi *locaux*) ou *absolus* (on dit aussi *globaux*).

L'étude du sens de variation de f passe souvent par le signe de f' , mais il arrive parfois qu'on puisse conclure par des résultats généraux sur les opérations entre fonctions monotones usuelles.

Il arrive aussi que l'étude du signe de f' ne soit pas particulièrement facile : dans certains cas, on est amené à étudier les variations de f'' , ou à étudier une fonction auxiliaire.

Le tableau des variations, soigné, doit indiquer clairement les intervalles formant le domaine d'étude.

On apportera un intérêt particulier aux frontières de ces intervalles, ainsi qu'aux points qui présentent un intérêt certain (notamment ceux où f possède un maximum ou un minimum local). On signalera par exemple les intersections du graphe (Γ) de f avec les axes (et en particulier avec Ox : il s'agit alors d'indiquer pour quelles valeurs de x la fonction f s'annule).

Le tableau de variation contient en général une ligne pour le signe de f' , mais on peut être amené à utiliser une ligne supplémentaire pour le signe de f'' : c'est l'occasion de préciser la concavité de f , et ses éventuels points d'inflexion.

4.6.5 Études locales et tracé du graphe

Après l'étude globale (réduction éventuelle du domaine d'étude, sens de variation) et avant de procéder au tracé de la courbe représentative de f , il peut être nécessaire de préciser quelques études locales.

Prolongement par continuité

On étudie les réels a éventuels où la continuité ou la dérivabilité de f pose problème (c'est-à-dire ne découle pas des résultats généraux sur les opérations entre fonctions continues ou dérivables).

On examine s'il y a une limite finie $b = \lim_{x \rightarrow a} f(x)$ (des deux cotés de a , ou d'un seul).

Si c'est le cas, on peut effectuer un *prolongement par continuité* en posant $f(a) = b$.

Demi-tangente horizontale ou oblique

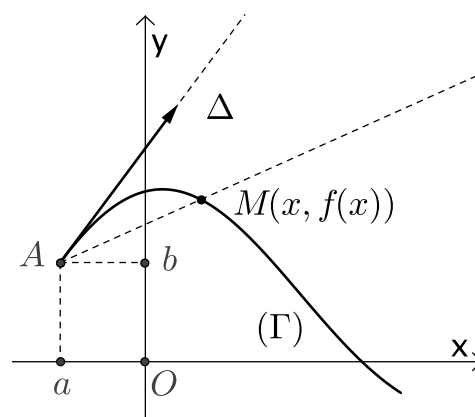
Pour examiner l'existence d'une tangente au point $A(a, b)$,

on calcule la limite $\lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a}$.

On cherche ainsi la position limite de la corde de $A(a, f(a))$ à $M(x, f(x))$ quand « M tend vers A ».

Supposons par exemple $\lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} = \delta$ ($\delta \in \mathbb{R}$).

On conclut alors à la présence en $A(a, f(a))$ d'une tangente (ou d'une demi-tangente) Δ à la courbe (Γ) , de coefficient directeur δ (donc horizontale si $\delta = 0$, et oblique si $\delta \neq 0$).



L'équation de cette (demi-)tangente Δ est $y = f(a) + (x - a)\delta$.

Il est recommandé de préciser le placement de (Γ) par rapport à cette demi-tangente, donc d'étudier le signe de la différence $f(x) - f(a) - \delta(x - a)$ quand x tend vers a .

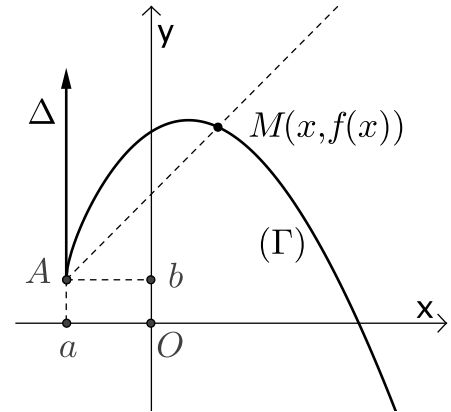
Demi-tangente verticale

Supposons au contraire $\lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} = \pm\infty$.

On conclut dans ce cas à la présence d'une tangente (ou demi-tangente) verticale en $A(a, b = f(a))$.

On illustrera par un dessin (pour une demi-tangente, on indiquera dans quel sens elle est dirigée).

Bien sûr cela demande parfois des connaissances et des techniques assez fines en analyse (par exemple l'utilisation des développements limités), qui ne sont pas forcément acquises à ce stade de l'année.



Asymptotes verticales ou horizontales

On note les réels a éventuels tels que $\lim_{x \rightarrow a} f(x) = \pm\infty$

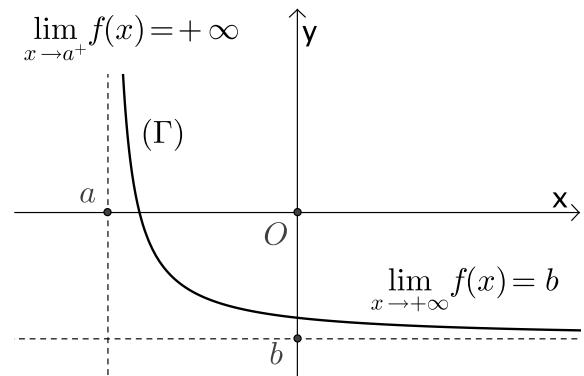
À un tel a correspond l'asymptote verticale $x = a$ (faire un dessin pour préciser l'allure).

Autre étude locale importante : quand $x \rightarrow \pm\infty$.

Supposons par exemple $\lim_{x \rightarrow \pm\infty} f(x) = b$ ($b \in \mathbb{R}$).

On a alors l'*asymptote horizontale* d'équation $y = b$.

On placera (Γ) par rapport à cette asymptote (signe de $f(x) - b$, ou sens de variations de f).



Asymptotes obliques

On suppose ici que $\lim_{x \rightarrow \pm\infty} f(x) = \pm\infty$.

Il est alors *possible* que le graphe (Γ) de f présente une *asymptote oblique* d'équation $y = ax + b$.

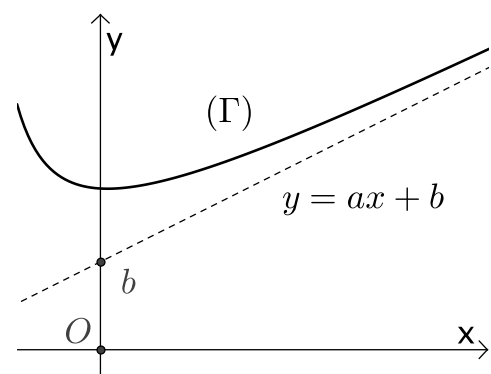
On pourra conclure à l'existence d'une telle asymptote de l'une des manières suivantes :

– On écrit $y = ax + b + \varphi(x)$, avec $\lim_{x \rightarrow \pm\infty} \varphi(x) = 0$.

Dans ce cas, le signe de φ indique la position de (Γ) par rapport à l'asymptote.

– On trouve $\lim_{x \rightarrow \pm\infty} \frac{f(x)}{x} = a$ puis $\lim_{x \rightarrow \pm\infty} (f(x) - ax) = b$.

Le signe de $f(x) - ax - b$ donne le placement.



Là encore, les « développements limités », non encore abordés, peuvent être très utiles.

Branches paraboliques

Il y a deux autres cas fréquents, toujours sous l'hypothèse $\lim_{x \pm \infty} f(x) = \infty$:

- Si $\lim_{x \pm \infty} \frac{f(x)}{x} = \infty$, on dit que (Γ) présente une branche parabolique dans la direction Oy .

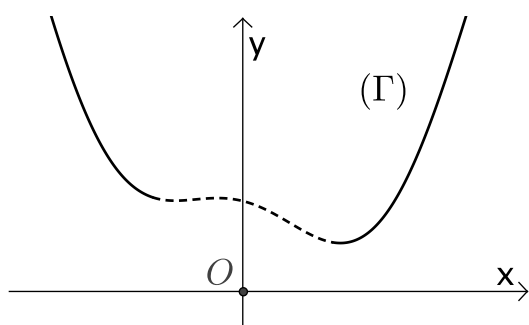
On précisera si la branche parabolique est tournée vers les $y > 0$ ou vers les $y < 0$.

Cette terminologie est à rapprocher de l'allure du graphe de $x \mapsto x^2$ quand $x \rightarrow \pm\infty$.

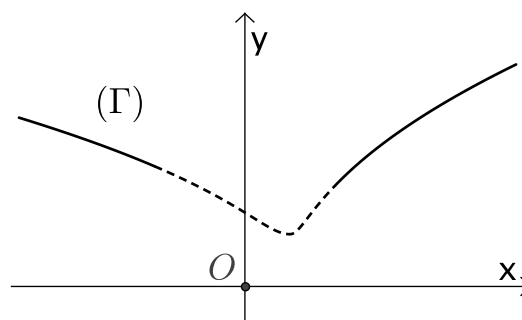
- Si $\lim_{x \pm \infty} \frac{f(x)}{x} = 0$, on dit que (Γ) présente une branche parabolique dans la direction Ox .

Cette terminologie est à rapprocher de l'allure du graphe de $x \mapsto \sqrt{x}$ quand $x \rightarrow +\infty$.

Dans ce cas, comme dans tous les précédents, penser à illustrer par un dessin !



$$\lim_{x \rightarrow -\infty} \frac{f(x)}{x} = -\infty \text{ et } \lim_{x \rightarrow +\infty} \frac{f(x)}{x} = +\infty$$



$$\lim_{x \rightarrow -\infty} \frac{f(x)}{x} = 0^- \text{ et } \lim_{x \rightarrow +\infty} \frac{f(x)}{x} = 0^+$$

4.6.6 Quelques inégalités utiles

Il est très souvent nécessaire de procéder à des majorations, des minorations, des encadrements.

Bien des techniques sont possibles, et notamment purement algébriques dans les cas les plus simples.

Souvent les inégalités seront obtenues par l'étude du signe d'une expression $f(x)$, ce signe résultant lui-même du sens de variations de f .

▷ Inégalités de convexité

Proposition 4.6.1

Soit $f : I \rightarrow \mathbb{R}$, deux fois dérivable et convexe (donc telle que $f''(x) \geq 0$ pour tout x de I).

Soit a un élément de I . Pour tout x de I , on a l'inégalité : $f(x) \geq f(a) + f'(a)(x - a)$.

Proposition 4.6.2

Soit $f : I \rightarrow \mathbb{R}$, deux fois dérivable et convexe (donc telle que $f''(x) \geq 0$ pour tout x de I).

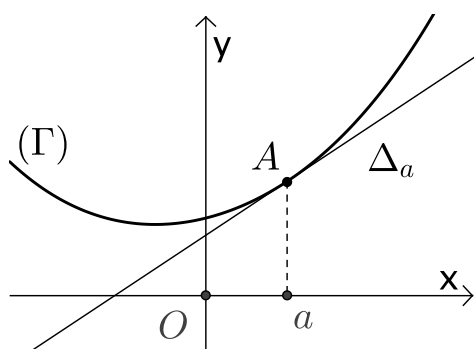
Soit a, b deux éléments de I , avec $a \leq b$. Alors on a l'inégalité : $f\left(\frac{a+b}{2}\right) \leq \frac{f(a) + f(b)}{2}$.

Plus généralement, pour tout λ de $[0, 1]$, on a : $f(\lambda a + (1 - \lambda)b) \leq \lambda f(a) + (1 - \lambda)f(b)$.

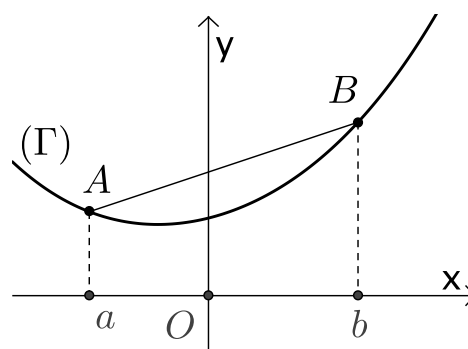
Remarque : dans le cas d'une application concave ($f'' \leq 0$), les inégalités sont inversées.

Interprétations graphiques, en notant (Γ) le graphe d'une fonction numérique convexe f :

- La courbe (Γ) est « partout au-dessus » de sa tangente Δ_a au point $A(a, f(a))$.
- L'arc de (Γ) situé entre $A(a, f(a))$ et $B(b, f(b))$ est situé « en-dessous » de la corde $[A; B]$



f convexe :
courbe (Γ) au-dessus de la tangente Δ_a



f convexe :
arc (AB) en-dessous de la corde $[A; B]$

Voici deux cas particuliers très importants :

Proposition 4.6.3

Pour tout x de \mathbb{R} , on a : $\exp(x) \geq 1 + x$, avec égalité si et seulement si $x = 0$.

Pour tout $x > -1$, on a : $\ln(1 + x) \leq x$, avec égalité si et seulement si $x = 0$.

▷ Moyenne géométrique et moyenne arithmétique

Proposition 4.6.4

Soit x, y deux réels positifs ou nuls.

Alors on a l'inégalité $\sqrt{ab} \leq \frac{1}{2}(a + b)$, avec égalité si et seulement si $a = b$.

L'inégalité classique suivante peut être vue comme un cas particulier :

Proposition 4.6.5

Pour tout x de $[0, 1]$, on a $0 \leq x(1 - x) \leq \frac{1}{4}$, avec égalité si et seulement si $x = \frac{1}{2}$.

On peut au contraire généraliser à une famille de n réels positifs ou nuls.

Proposition 4.6.6

Soit $(x_k)_{1 \leq k \leq n}$ une famille de n réels positifs ou nuls.

Soit $G = \left(\prod_{k=1}^n x_k\right)^{1/n}$ la moyenne géométrique des x_k , et $A = \frac{1}{n} \sum_{k=1}^n x_k$ leur moyenne arithmétique.

Alors on a l'inégalité $A \geq G$ (avec égalité si et seulement si les x_k sont tous égaux)

▷ Inégalités trigonométriques

Proposition 4.6.7

Pour tout réel x on a : $|\sin(x)| \leq |x|$, et $|1 - \cos(x)| \leq \frac{x^2}{2}$.

▷ Inégalité de Bernoulli

Proposition 4.6.8

Pour tout $n \geq 2$, et tout $x > -1$, on a : $(1+x)^n \geq 1+nx$ (égalité si et seulement si $x = 0$).

▷ Inégalité de Cauchy-Schwarz

Proposition 4.6.9 (inégalité de Cauchy-Schwarz)

Pour tous réels x_1, \dots, x_n et y_1, \dots, y_n , on a l'inégalité : $\left(\sum_{k=1}^n x_k y_k\right)^2 \leq \left(\sum_{k=1}^n x_k^2\right) \left(\sum_{k=1}^n y_k^2\right)$

Il y a égalité si et seulement si les n -uplets $u = (x_1, \dots, x_n)$ et $v = (y_1, \dots, y_n)$ sont proportionnels.

Chapitre 5

Techniques d'analyse (intégration)

Sommaire

5.1	Calculs de primitives	124
5.1.1	Primitives d'une fonction numérique	124
5.1.2	Primitives usuelles	125
5.1.3	Reconnaître la dérivée d'une composée	125
5.1.4	Primitivation de $x \mapsto \frac{px + q}{ax^2 + bx + c}$	126
5.2	Intégration sur un segment	126
5.2.1	Intégrale d'une fonction continue	126
5.2.2	Intégration et fonctions de classe \mathcal{C}^1	128
5.2.3	Intégration par parties	128
5.2.4	Changement de variable	129
5.2.5	Utilisation de la parité ou de la périodicité	130
5.3	Compléments sur les primitives	130
5.3.1	Primitives de $\sin^p(x) \cos^q(x)$	130
5.3.2	Primitives de $P(x)e^{ax}$ (et associées)	131
5.3.3	Utilisation de récurrences	131
5.3.4	Primitives des fractions rationnelles	132
5.3.5	Fractions trigonométriques	133
5.3.6	Primitives avec radicaux	134
5.4	Extension aux fonctions à valeurs complexes	135
5.4.1	Fonctions à valeurs complexes	135
5.4.2	Dérivée et intégrale des fonctions complexes	136
5.4.3	Extension des résultats relatifs aux fonctions réelles	136
5.4.4	Cas de la fonction exponentielle complexe	138
5.5	Équations différentielles $y' + a(x)y = b(x)$	138
5.5.1	Position du problème	138
5.5.2	Résolution de l'équation homogène $y' + a(x)y = 0$	139
5.5.3	Résolution de l'équation $y' + a(x)y = b(x)$	139
5.5.4	Principe de superposition	140
5.5.5	Méthode de variation de la constante	141
5.5.6	Problème de Cauchy	141
5.6	Équations différentielles du 2nd ordre	142
5.6.1	Position du problème	143
5.6.2	Résolution de l'équation homogène	143

5.6.3	Forme des solutions de l'équation complète	144
5.6.4	Problème de Cauchy	145
5.6.5	Quelques exemples	146

5.1 Calculs de primitives

5.1.1 Primitives d'une fonction numérique

Dans cette section, I est un intervalle de \mathbb{R} , d'intérieur non vide.

Définition 5.1.1 (primitive sur un intervalle)

Soit f une fonction de I dans \mathbb{R} . On dit qu'une fonction $F : I \rightarrow \mathbb{R}$ est une *primitive* de f sur I si F est dérivable sur I et si, pour tout x de I , on a : $F'(x) = f(x)$.

Proposition 5.1.1 (relation entre les primitives d'une même fonction)

Soit $f : I \rightarrow \mathbb{R}$ une fonction numérique, et soit F une primitive de f sur I .

Les primitives de f sur I sont les fonctions $x \mapsto G(x) = F(x) + \lambda$, avec λ dans \mathbb{R} .

Pour tout a de I , et tout y_0 dans \mathbb{R} , il existe une unique primitive G de f telle que $G(a) = y_0$.

Remarques

– Les primitives de f sur I sont définies « à une constante additive près ».

On note souvent $\int f(x) dx = F(x) + \lambda$ pour désigner l'ensemble des primitives de f sur I .

On dit alors communément que λ est la « constante d'intégration ».

Par exemple $\int \cos(x) dx = \sin(x) + \lambda$ désigne l'ensemble des primitives de $x \mapsto \cos(x)$ sur \mathbb{R} .

Dans cette notation, x joue le rôle de « variable muette ». Le symbole choisi n'a pas d'importance dans la mesure où il ne crée pas d'ambiguïté.

– Soit $f : I \rightarrow \mathbb{R}$ une fonction numérique. Soit F et G deux primitives de f sur I .

Si on souhaite déterminer la constante λ telle que $G = F + \lambda$, il suffit de calculer $G(a) - F(a)$ en un point de I (ou de calculer la différence des limites de F et G en une extrémité de I).

– Le calcul de primitives s'effectue toujours **sur un intervalle**.

Par exemple, parler des primitives de $x \mapsto \frac{1}{x}$ sur \mathbb{R}^* n'a aucun sens.

Supposons par exemple que f soit définie sur la réunion $\mathcal{D} = I \cup J$ de deux intervalles disjoints.

Supposons également que F et G soient dérivables sur \mathcal{D} et que : $\forall x \in \mathcal{D}, F'(x) = G'(x) = f(x)$.

D'une part : $\exists \lambda \in \mathbb{R}, \forall x \in I, G(x) = F(x) + \lambda$. D'autre part : $\exists \mu \in \mathbb{R}, \forall x \in J, G(x) = F(x) + \mu$.

Mais en aucun cas, on ne peut affirmer que les constantes λ et μ sont égales.

5.1.2 Primitives usuelles

Voici un memento des primitives $x \mapsto F(x)$ d'une fonction numérique $x \mapsto f(x)$, dans les cas « usuels ». Les résultats qui figurent dans le tableau suivant doivent donc être connus « par cœur ».

$f(x)$	$F(x)$	sur	$f(x)$	$F(x)$	sur
$x^\alpha, (\alpha \neq -1)$	$\frac{x^{\alpha+1}}{\alpha+1}$	\mathbb{R}^{+*}	$\frac{1}{1+x^2}$	$\arctan x$	\mathbb{R}
$\frac{1}{x}$	$\ln x $	$\mathbb{R}^{-*}, \mathbb{R}^{+*}$	$\frac{1}{1-x^2}$	$\frac{1}{2} \ln \left \frac{1+x}{1-x} \right $	$x \neq \pm 1$
e^x	e^x	\mathbb{R}	$\frac{1}{\sqrt{1+x^2}}$	$\ln(x + \sqrt{1+x^2})$	\mathbb{R}
$a^x (a \neq 1)$	$\frac{a^x}{\ln a}$	\mathbb{R}	$\frac{1}{\sqrt{1-x^2}}$	$\arcsin(x)$	$] -1, 1 [$
$\sin(x)$	$-\cos(x)$	\mathbb{R}	$\frac{1}{\sqrt{x^2-1}}$	$\ln x + \sqrt{x^2-1} $	$ x > 1$
$\cos(x)$	$\sin(x)$	\mathbb{R}	$\tan(x)$	$-\ln \cos(x) $	$x \neq \frac{\pi}{2} + k\pi$
$\operatorname{sh}(x)$	$\operatorname{ch}(x)$	\mathbb{R}	$\frac{1}{\sin(x)}$	$\ln \left \tan\left(\frac{x}{2}\right) \right $	$x \neq k\pi$
$\operatorname{ch}(x)$	$\operatorname{sh}(x)$	\mathbb{R}	$\frac{1}{\cos(x)}$	$\ln \left \tan\left(\frac{x}{2} + \frac{\pi}{4}\right) \right $	$x \neq \frac{\pi}{2} + k\pi$
$\frac{1}{\cos^2(x)}$	$\tan(x)$	$x \neq \frac{\pi}{2} + k\pi$	$\frac{1}{\operatorname{ch}^2(x)}$	$\operatorname{th}(x)$	\mathbb{R}
$\frac{1}{\sin^2(x)}$	$-\frac{1}{\tan(x)}$	$x \neq k\pi$	$\frac{1}{\operatorname{sh}^2(x)}$	$-\frac{1}{\operatorname{th}(x)}$	$\mathbb{R}^{+*}, \mathbb{R}^{-*}$

Remarque : si $\int f(x) dx = F(x) + \lambda$, alors $\int f(ax+b) dx = \frac{1}{a}F(ax+b) + \lambda$.

$$\text{Par exemple : } \int (ax+b)^\alpha dx = \frac{1}{a} \frac{(ax+b)^{\alpha+1}}{\alpha+1} + \lambda \quad \int \frac{dx}{ax+b} = \frac{1}{a} \ln|ax+b| + \lambda$$

$$\int \cos(ax+b) dx = \frac{1}{a} \sin(ax+b) + \lambda \quad \int e^{ax} dx = \frac{1}{a} e^{ax} + \lambda$$

5.1.3 Reconnaître la dérivée d'une composée

Soit $f : I \rightarrow \mathbb{R}$ une fonction numérique, et soit $\varphi : J \rightarrow \mathbb{R}$ une fonction dérivable, telle que $\varphi(J) \subset I$. Soit F une primitive de f sur I . Alors $F \circ \varphi$ est une primitive de $(f \circ \varphi)\varphi'$ sur J .

On peut donc écrire directement : $\int f(\varphi(x))\varphi'(x) dx = (F \circ \varphi)(x) + \lambda$

Voici trois situations classiques :

$$\int \varphi'(x) e^{\varphi(x)} dx = e^{\varphi(x)} + \lambda \quad \int \frac{\varphi'(x)}{\varphi(x)} dx = \ln|\varphi(x)| + \lambda \quad \int \varphi'(x) \varphi^r(x) dx = \frac{\varphi^{r+1}(x)}{r+1} + \lambda$$

5.1.4 Primitivation de $x \mapsto \frac{px + q}{ax^2 + bx + c}$

Les deux résultats ci-dessous sont utiles à connaître :

$$\int \frac{dx}{x^2 + a^2} = \frac{1}{a} \arctan \frac{x}{a} + \lambda \qquad \int \frac{dx}{x^2 - a^2} = \frac{1}{2a} \ln \left| \frac{x - a}{x + a} \right| + \lambda$$

Plus généralement, soit la fonction numérique $f : x \mapsto \frac{px + q}{ax^2 + bx + c}$, où p, q, a, b, c sont réels (et $a \neq 0$).

On veut calculer $\int f(x) dx$, sur un intervalle I où le dénominateur de f ne s'annule pas.

La première idée est d'écrire : $f(x) = \frac{p}{2a} \left(\frac{2ax + b}{ax^2 + bx + c} \right) + \left(q - \frac{pb}{2a} \right) \frac{1}{ax^2 + bx + c}$.

Cette décomposition permet d'utiliser $\int \frac{2ax + b}{ax^2 + bx + c} dx = \ln |ax^2 + bx + c| + \lambda$.

Il reste donc à calculer $\int \frac{dx}{ax^2 + bx + c}$.

Pour cela, on utilise la « forme canonique », et le discriminant $\Delta = b^2 - 4ac$.

$$ax^2 + bx + c = a \left(x^2 + \frac{b}{a}x + \frac{c}{a} \right) = a \left(\left(x + \frac{b}{2a} \right)^2 - \frac{b^2}{4a^2} + \frac{c}{a} \right) = a \left(\left(x + \frac{b}{2a} \right)^2 - \frac{\Delta}{4a^2} \right)$$

Suivant le signe de Δ , on est donc ramené à l'une des trois situations suivantes :

- Si $\Delta > 0$: $\int \frac{dx}{(x + \alpha)^2 - \beta^2} = \frac{1}{2\beta} \ln \left| \frac{x + \alpha - \beta}{x + \alpha + \beta} \right| + \lambda$
- Si $\Delta = 0$: $\int \frac{dx}{(x + \alpha)^2} = -\frac{1}{x + \alpha} + \lambda$
- Si $\Delta < 0$: $\int \frac{dx}{(x + \alpha)^2 + \beta^2} = \frac{1}{\beta} \arctan \frac{x + \alpha}{\beta} + \lambda$

5.2 Intégration sur un segment

5.2.1 Intégrale d'une fonction continue

Nous admettrons le résultat suivant :

Proposition 5.2.1

Si $f : I \rightarrow \mathbb{R}$ est une fonction continue, elle admet des primitives sur I .

Remarque : la réciproque de la propriété précédente est fautive (il existe des fonctions numériques qui admettent des primitives sur un intervalle I sans être continues en tout point de I) mais la question est relativement difficile et elle est hors-programme.

Définition 5.2.1 (intégrale d'une fonction continue sur un segment)

Soit $f : I \rightarrow \mathbb{R}$ une fonction numérique continue. Soit a, b deux éléments de I .

On note $\int_a^b f(x) dx$ la quantité $F(b) - F(a)$, où F est une primitive *quelconque* de f sur I .

Cette quantité est appelée *intégrale* de f sur le segment $[a, b]$ (ou « entre a et b »).

Remarques

– Si f vaut constamment λ sur I , alors on a : $\int_a^b f(x) dx = \lambda(b - a)$.

– La quantité $F(b) - F(a)$, notée $[F(x)]_a^b$, ne dépend pas de la primitive F choisie pour f .

– Soit $f : I \rightarrow \mathbb{R}$ une fonction numérique continue. Soit a un élément de I .

Alors la fonction $F : x \mapsto F(x) = \int_a^x f(t) dt$ est la primitive de f sur I qui s'annule en a .

Le nom de la « variable d'intégration » (ici t) doit être différent de celui de variable (ici x) de F .

– Quand on calcule $\int_a^b f(x) dx$, on dit qu'on « intègre » f sur le segment $[a, b]$.

Même si les deux notions sont très liées, on ne confondra pas la *primitivation* de f sur I (qui est l'art de chercher les primitives de f sur I , donc les fonctions dont la dérivée est f) avec l'*intégration* de f sur un segment $[a, b]$ de I (le résultat est dans ce cas un réel).

Proposition 5.2.2 (linéarité de l'intégrale)

Soit f et g deux fonctions continues sur I . Soit a et b deux éléments de I .

Pour tous réels α, β , on a : $\int_a^b (\alpha f(x) + \beta g(x)) dx = \alpha \int_a^b f(x) dx + \beta \int_a^b g(x) dx$.

Proposition 5.2.3 (positivité et croissance de l'intégrale)

Soit f et g deux fonctions continues sur le segment $[a, b]$, avec $a < b$.

Si $f \geq 0$ sur $[a, b]$, alors on a : $\int_a^b f(x) dx \geq 0$ (avec égalité $\Leftrightarrow f(x) = 0$ pour tout x de $[a, b]$).

Si $f \leq g$ sur $[a, b]$, alors $\int_a^b f(x) dx \leq \int_a^b g(x) dx$ (avec égalité $\Leftrightarrow f(x) = g(x)$ pour tout x de $[a, b]$).

Remarque : l'hypothèse $a < b$ est ici essentielle.

Proposition 5.2.4 (relation de Chasles)

Soit $f : I \rightarrow \mathbb{R}$, continue. Soit a, b, c dans I . Alors $\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx$.

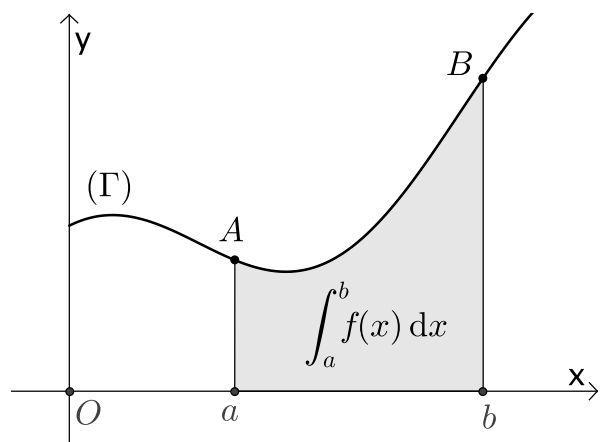
Interprétation de l'intégrale, en termes d'aire

Historiquement, la notion d'intégrale est liée au calcul d'aire de domaines du plan.

Notre définition de l'intégrale possède l'avantage d'être rapidement opérationnelle, mais elle recèle une difficulté qui ne peut pas être levée à ce stade de l'année.

Contentons-nous d'admettre, sans plus de précision, que si f est continue et positive ou nulle sur $[a, b]$, avec $a \leq b$, alors l'intégrale de f sur $[a, b]$ est une mesure de l'aire du domaine défini par $a \leq x \leq b$ et $0 \leq y \leq f(x)$.

L'aire est exprimée en « unités d'aires », l'unité étant l'aire du rectangle délimité par $(0, 0)$ et $(1, 1)$.



On peut étendre cette interprétation au cas d'une fonction ne gardant pas un signe constant, à condition de « compter positivement » les parties où $f \geq 0$ et négativement celles où $f \leq 0$.

5.2.2 Intégration et fonctions de classe \mathcal{C}^1

Pour tout ce qui touche à l'intégration, il est souvent utile de disposer de propriétés un peu plus fortes que la continuité ou même la dérivabilité, ce qui conduit à la définition suivante :

Définition 5.2.2 (fonctions de classe \mathcal{C}^1 sur un intervalle)

Soit $f : I \rightarrow \mathbb{R}$ une fonction numérique.

On dit que f est de classe \mathcal{C}^1 sur I si f est dérivable sur I et si f' est continue sur I .

Bien sûr, si f est de classe \mathcal{C}^1 , l'intégrale de f' sur $[a, b]$ est la différence des valeurs de f entre a et b :

Proposition 5.2.5 (intégrale de la dérivée d'une fonction de classe \mathcal{C}^1)

Soit $f : I \rightarrow \mathbb{R}$, de classe \mathcal{C}^1 . Pour tous a et b dans I , on a : $\int_a^b f'(t) dt = f(b) - f(a)$

Si f est continue sur I , et si $F(x) = \int_a^x f(t) dt$, on sait que $F'(x) = f(x)$ pour tout x .

On exprime la définition de F en disant qu'elle est une « intégrale fonction de sa borne supérieure ».

Le résultat suivant apporte une généralisation aux « intégrales fonctions de leurs bornes » :

Proposition 5.2.6 (intégrale fonction de ses bornes)

Soit f une fonction continue sur un intervalle I .

Soit u et v deux fonctions de classe \mathcal{C}^1 , de J dans \mathbb{R} , telles que $u(J) \subset I$ et $v(J) \subset I$.

La fonction G , définie sur J par $G(x) = \int_{u(x)}^{v(x)} f(t) dt$, est de classe \mathcal{C}^1 sur J .

Sa dérivée est donnée par : $\forall x \in J, G'(x) = v'(x)f(v(x)) - u'(x)f(u(x))$.

Si l'intégrale cherchée ne peut pas être obtenue immédiatement par utilisation d'une primitive usuelle, on cherche souvent à transformer l'intégrale initiale en une ou plusieurs autres que l'on sait calculer.

5.2.3 Intégration par parties

Proposition 5.2.7 (intégration par parties)

Soit f et g deux fonctions de classe \mathcal{C}^1 sur un intervalle I .

Pour tous a, b dans l'intervalle I , on a : $\int_a^b f(x)g'(x) dx = [f(x)g(x)]_a^b - \int_a^b f'(x)g(x) dx$.

Il y a aussi une méthode de « primitivation par parties » :

Si f et g sont de classe \mathcal{C}^1 sur I , alors $\int f(x)g'(x) dx = f(x)g(x) - \int f'(x)g(x) dx$.

Exemples d'intégrations par parties

– Pour tout n de \mathbb{N}^* , et pour tout m de \mathbb{N} , on pose $I_{n,m} = \int_0^1 x^n \ln^m(x) dx$.

Si $m \geq 1$, on a : $I_{n,m} = \left[\frac{x^{n+1}}{n+1} \ln^m(x) \right]_0^1 - \frac{m}{n+1} \int_0^1 x^n \ln^{m-1}(x) dx = -\frac{m}{n+1} I_{n,m-1}$.

Une récurrence facile donne alors : $I_{n,m} = \frac{(-1)^m m!}{(n+1)^m} I_{n,0} = \frac{(-1)^m m!}{(n+1)^m} \int_0^1 x^n dx = \frac{(-1)^m m!}{(n+1)^{m+1}}$.

– On peut calculer une primitive de $\frac{1}{(1+x^2)^2}$ en primitivant par partie $\frac{1}{1+x^2}$.

$$\begin{aligned} \text{En effet : } \int \frac{dx}{1+x^2} &= \frac{x}{1+x^2} + 2 \int \frac{x^2 dx}{(1+x^2)^2} = \frac{x}{1+x^2} + 2 \int \frac{(x^2+1)-1}{(1+x^2)^2} dx \\ &= \frac{x}{1+x^2} + 2 \int \frac{dx}{1+x^2} - 2 \int \frac{dx}{(1+x^2)^2} \end{aligned}$$

$$\text{On en déduit : } \int \frac{dx}{(1+x^2)^2} = \frac{1}{2} \left(\frac{x}{1+x^2} + \int \frac{dx}{1+x^2} \right) = \frac{1}{2} \left(\frac{x}{1+x^2} + \arctan(x) \right) + \lambda$$

– Soit (α, β) dans $\mathbb{R}^2 \setminus \{(0,0)\}$. Posons $C_{\alpha,\beta} = \int e^{\alpha x} \cos(\beta x) dx$ et $S_{\alpha,\beta} = \int e^{\alpha x} \sin(\beta x) dx$

$$\text{On trouve : } \alpha C_{\alpha,\beta} = e^{\alpha x} \cos(\beta x) + \beta \int e^{\alpha x} \sin(\beta x) dx = e^{\alpha x} \cos(\beta x) + \beta S_{\alpha,\beta}.$$

$$\text{De la même manière : } \alpha S_{\alpha,\beta} = e^{\alpha x} \sin(\beta x) - \beta \int e^{\alpha x} \cos(\beta x) dx = e^{\alpha x} \sin(\beta x) - \beta C_{\alpha,\beta}$$

$$\text{En combinant ces deux égalités : } \begin{cases} \alpha^2 C_{\alpha,\beta} = \alpha e^{\alpha x} \cos(\beta x) + \beta (e^{\alpha x} \sin(\beta x) - \beta C_{\alpha,\beta}) \\ \alpha^2 S_{\alpha,\beta} = \alpha e^{\alpha x} \sin(\beta x) - \beta (e^{\alpha x} \cos(\beta x) + \beta S_{\alpha,\beta}) \end{cases}$$

$$\text{Et finalement : } \begin{cases} C_{\alpha,\beta} = \int e^{\alpha x} \cos(\beta x) dx = \frac{e^{\alpha x}}{\alpha^2 + \beta^2} (\alpha \cos(\beta x) + \beta \sin(\beta x)) + \lambda \\ S_{\alpha,\beta} = \int e^{\alpha x} \sin(\beta x) dx = \frac{e^{\alpha x}}{\alpha^2 + \beta^2} (\alpha \sin(\beta x) - \beta \cos(\beta x)) + \lambda \end{cases}$$

5.2.4 Changement de variable

Proposition 5.2.8 (changement de variable dans une intégrale)

Soit $f : I \rightarrow \mathbb{R}$ une fonction continue.

Soit $\varphi : J \rightarrow \mathbb{R}$, de classe \mathcal{C}^1 sur J , telle que $\varphi(J) \subset I$.

Alors, pour tous a, b de J , on a $\int_a^b f(\varphi(t)) \varphi'(t) dt = \int_c^d f(x) dx$, où $c = \varphi(a)$, et $d = \varphi(b)$

Pratique du changement de variable :

Cette égalité peut être utilisée dans un sens ou dans l'autre selon les cas :

– Dans le sens $\int_a^b \varphi'(t) f(\varphi(t)) dt \Rightarrow \int_c^d f(x) dx$.

On veut calculer $\int_a^b g(t) dt$, et on constate que $g(t)$ se met sous la forme $g(t) = f(\varphi(t)) \varphi'(t)$.

On pose alors $x = \varphi(t)$, et on note que lorsque $t = a$ ou $t = b$, alors $x = c$ ou $x = d$.

On écrit $dx = \varphi'(t) dt$ puis $\int_a^b g(t) dt = \int_a^b \varphi'(t) f(\varphi(t)) dt = \int_c^d f(x) dx$.

– Dans le sens $\int_c^d f(x) dx \Rightarrow \int_a^b \varphi'(t) f(\varphi(t)) dt$.

On part donc de $\int_c^d f(x) dx$ et on pose (indication, intuition, expérience, etc.) $x = \varphi(t)$.

Dans ce cas, il faut trouver a et b tels que $\varphi(a) = c$ et $\varphi(b) = d$.

Il est préférable de choisir φ et l'intervalle sur lequel cette fonction est définie de manière à ce que φ soit bijective : on a alors $a = \varphi^{-1}(c)$ et $b = \varphi^{-1}(d)$.

Par exemple, pour calculer $\int_0^1 \sqrt{1-x^2} dx$, on pose $x = \varphi(t) = \sin t$, avec $t \in \left[0, \frac{\pi}{2}\right]$.

On a alors $\sqrt{1-x^2} = |\cos(t)| = \cos(t)$, et $dx = \cos(t) dt$.

D'autre part, quand $x = 0$ alors $t = 0$, et quand $x = 1$ alors $t = \frac{\pi}{2}$.

On en déduit : $\int_0^1 \sqrt{1-x^2} dx = \int_0^{\pi/2} \cos^2(t) dt = \frac{1}{2} \int_0^{\pi/2} (1 + \cos(2t)) dt = \frac{\pi}{4}$.

Remarque : il y a un moyen encore plus simple de calculer $\int_0^1 \sqrt{1-x^2} dx$ (lequel?)

5.2.5 Utilisation de la parité ou de la périodicité

Quelques changements de variables très simples permettent d'exploiter la parité, ou l'imparité, ou la périodicité, de la fonction à intégrer.

– Une simple translation permet d'écrire : $\int_a^b f(x) dx = \int_{a+\alpha}^{b+\alpha} f(x-\alpha) dx$.

– Si f est paire, alors $\int_{-a}^a f(x) dx = 2 \int_0^a f(x) dx$.

Si f est impaire, alors $\int_{-a}^a f(x) dx = 0$ (pas la peine de perdre du temps à calculer l'intégrale!)

– On suppose maintenant que f est T -périodique sur \mathbb{R} .

On a $\int_{a+T}^{b+T} f(x) dx = \int_a^b f(x) dx$, et plus généralement $\int_{a+kT}^{b+kT} f(x) dx = \int_a^b f(x) dx$ pour tout k de \mathbb{Z} .

Pour tous réels a et b , on a l'égalité $\int_a^{a+T} f(x) dx = \int_b^{b+T} f(x) dx$.

5.3 Compléments sur les primitives

Le calcul d'une intégrale se ramène souvent au calcul d'une primitive.

Dans ce paragraphe, on va passer en revue quelques situations courantes. Les méthodes décrites ici doivent être considérées comme des « compléments utiles » du cours.

On note $\int f(x) dx = F(x) + \lambda$ l'ensemble des primitives d'une application f .

5.3.1 Primitives de $\sin^p(x) \cos^q(x)$

Si on veut calculer $\int \sin^p(x) \cos^q(x) dx$, avec p et q dans \mathbb{N} , tout dépend de la parité de p et q .

– Si p est impair, on peut poser $t = \cos(x)$ (donc $dt = -\sin(x) dx$) :

$$\int \sin^3(x) \cos^4(x) dx = \int (t^2 - 1) t^4 dt = \frac{t^7}{7} - \frac{t^5}{5} + \lambda = \frac{\cos^7(x)}{7} - \frac{\cos^5(x)}{5} + \lambda$$

– Si q est impair, on peut poser $t = \sin(x)$ (donc $dt = \cos(x) dx$) :

$$\int \cos^5(x) dx = \int (1 - t^2)^2 dt = t - \frac{2t^3}{3} + \frac{t^5}{5} + \lambda = \sin(x) - \frac{2 \sin^3(x)}{3} + \frac{\sin^5(x)}{5} + \lambda$$

– Si p et q sont pairs, on linéarise :

$$\cos^4(x) dx = \frac{1}{8} \int (\cos(4x) + 4 \cos(2x) + 3) dx = \frac{\sin(4x)}{32} + \frac{\sin(2x)}{4} + \frac{3x}{8} + \lambda$$

5.3.2 Primitives de $P(x)e^{ax}$ (et associées)

On veut calculer $\int P(x)e^{ax} dx$, où P est un polynôme et a un scalaire.

On peut procéder par intégrations par parties successives, si $\deg P$ n'est pas trop grand.

Il est souvent préférable d'utiliser une méthode de coefficients indéterminés, et de chercher une primitive de $P(x)e^{ax}$ sous la forme $Q(x)e^{ax}$, avec $\deg Q = \deg P$.

Par exemple, on écrira $\int (x^3 - 2x + 1)e^{-x} dx = Q(x)e^{-x} + \lambda$, avec $Q(x) = \alpha x^3 + \beta x^2 + \gamma x + \delta$.

Par dérivation et identification, on obtient :

$$\begin{aligned} (Q(x)e^{-x})' &= (Q'(x) - Q(x))e^{-x} = (-\alpha x^3 + (3\alpha - \beta)x^2 + (2\beta - \gamma)x + \gamma - \delta)e^{-x} \\ &= (x^3 - 2x + 1)e^{-x} \quad \text{donc } \alpha = -1, \beta = -3, \gamma = -4 \text{ et } \delta = -5. \end{aligned}$$

Ainsi $\int (x^3 - 2x + 1)e^{-x} dx = -(x^3 + 3x^2 + 4x + 5)e^{-x} + \lambda$.

Primitives de $P(x)\sin(ax)$, ou $P(x)\cos(ax)$, ou $P(x)\operatorname{sh}(ax)$ ou $P(x)\operatorname{ch}(ax)$.

On est ramené au cas précédent en utilisant les formules d'Euler (quitte à aller vers des fonctions à valeurs complexes : ce sujet est traité un peu plus loin dans ce chapitre).

– On veut par exemple calculer $I = \int (x^3 - 2x + 1)\operatorname{ch}(x) dx$. On écrit $\operatorname{ch}(x) = \frac{e^x + e^{-x}}{2}$.

On obtient : $I = \frac{J + K}{2}$, avec $J = \int (x^3 - 2x + 1)e^x dx$ et $K = \int (x^3 - 2x + 1)e^{-x} dx$.

On sait déjà que $J = -(x^3 + 3x^2 + 4x + 5)e^{-x} + \lambda$ (voir exemple précédent).

Une méthode analogue donne $K = (x^3 - 3x^2 + 4x - 3)e^x + \lambda$.

On en déduit : $I = \frac{1}{2}e^x(x^3 - 3x^2 + 4x - 3) - \frac{1}{2}e^{-x}(x^3 + 3x^2 + 4x + 5) + \lambda$.

Dans le résultat, on peut remplacer e^x par $\operatorname{ch}(x) + \operatorname{sh}(x)$ et e^{-x} par $\operatorname{ch}(x) - \operatorname{sh}(x)$.

Tout calcul fait, on trouve : $I = -(3x^2 + 4)\operatorname{ch}(x) + (x^3 + 4x + 1)\operatorname{sh}(x) + \lambda$.

– Si on veut calculer par exemple $J = \int x^4 \cos(x) dx$, on écrit $J = \operatorname{Re}\left(\int x^4 e^{ix} dx\right)$.

Par identification on obtient $\int x^4 e^{ix} dx = -e^{ix}(ix^4 - 4x^3 - 12ix^2 + 24x + 24i) + \lambda$.

$$\begin{aligned} \text{Ainsi : } J &= \int x^4 \cos(x) dx = x^4 \sin(x) + 4x^3 \cos(x) - 12x^2 \sin(x) - 24x \cos(x) + 24 \sin(x) + \lambda \\ &= (x^4 - 12x^2 + 24) \sin(x) + 4(x^3 - 6x) \cos(x) + \lambda \end{aligned}$$

5.3.3 Utilisation de récurrences

Pour calculer $I_n = \int f_n(x) dx$, avec n dans \mathbb{N} , on peut chercher une relation de récurrence.

Cela passe souvent par une intégration par parties.

– **Premier exemple :**

On veut calculer de $I_n = \int \sin^n(x) dx$ ou $J_n = \int \cos^n(x) dx$ (« intégrales de Wallis »).

On suppose $n \geq 2$ et on intègre par partie $\sin(x) \sin^{n-1}(x)$ en dérivant $\sin^{n-1}(x)$.

$$\begin{aligned}
I_n &= \int \sin^n(x) dx = -\cos(x) \sin^{n-1}(x) + (n-1) \int \cos^2(x) \sin^{n-2}(x) dx \\
&= -\cos(x) \sin^{n-1}(x) + (n-1) \int (1 - \sin^2(x)) \sin^{n-2}(x) dx \\
&= -\cos(x) \sin^{n-1}(x) + (n-1)(I_{n-2} - I_n)
\end{aligned}$$

On en déduit la relation : $I_n = \frac{1}{n}(-\cos(x) \sin^{n-1}(x) + (n-1)I_{n-2})$

Connaissant $\begin{cases} I_0 = x + \lambda \\ I_1 = -\cos(x) + \lambda \end{cases}$, on peut donc trouver les I_n de proche en proche.

– **Deuxième exemple** : on veut calculer $I_n = \int \frac{dx}{(a^2 + x^2)^n}$.

Première méthode : on intègre par parties dans I_n , en primitivant 1 et en dérivant $\frac{1}{(a^2 + x^2)^n}$:

$$\begin{aligned}
I_n &= \frac{x}{(a^2 + x^2)^n} + 2n \int \frac{x^2}{(a^2 + x^2)^{n+1}} dx = \frac{x}{(a^2 + x^2)^n} + 2n \int \frac{(a^2 + x^2) - a^2}{(a^2 + x^2)^{n+1}} dx \\
&= \frac{x}{(a^2 + x^2)^n} + 2n(I_n - a^2 I_{n+1}) \implies I_{n+1} = \frac{1}{2na^2} \left[\frac{x}{(a^2 + x^2)^n} + (2n-1)I_n \right]
\end{aligned}$$

Connaissant $I_1 = \frac{1}{a} \arctan\left(\frac{x}{a}\right) + \lambda$, on en déduit I_n pour tout n de \mathbb{N}^* .

Il y a une autre méthode pour calculer $I_n = \int \frac{dx}{(a^2 + x^2)^n}$ (et qu'on recommandera si $n \leq 3$).

On effectue le changement de variable $x = a \tan t$. Ainsi $dx = a(1 + \tan^2 t) dt$.

Ainsi : $I_n = \int \frac{dt}{a^{2n-1}(1+t^2)^{n-1}} = \frac{1}{a^{2n-1}} \int \cos^{2n-2}(t) dt$, ce qui ramène aux intégrales de Wallis.

5.3.4 Primitives des fractions rationnelles

On décompose la fraction rationnelle *en éléments simples*.

La théorie sera abordée plus tard (et en se limitant à des situations assez simples).

Le problème est la primitivation de $f(x) = \frac{\lambda x + \mu}{(x^2 + bx + c)^n}$, où $b^2 - 4c < 0$.

On écrit $f(x) = \frac{\lambda(2x+b)}{2(x^2+bx+c)^n} + \frac{2\mu-\lambda b}{2(x^2+bx+c)^n}$.

La fonction $g(x) = \frac{2x+b}{2(x^2+bx+c)^n}$ s'intègre facilement car elle est du type $\frac{u'(x)}{u(x)}$.

Il reste à intégrer $h(x) = \frac{1}{(x^2+bx+c)^n}$. Or $x^2+bx+c = \left(x + \frac{b}{2}\right)^2 + \alpha^2$, avec $\alpha = \frac{1}{2}\sqrt{4c-b^2}$.

Le changement de variable $x = t - \frac{b}{2}$ donne : $\int \frac{dx}{(x^2+bx+c)^n} = \int \frac{dt}{(t^2+\alpha^2)^n}$.

On est ainsi ramené à une intégrale qu'on sait calculer (exemple précédent).

– **Un exemple très simple** : on veut calculer $I = \int \frac{dx}{x(x^2 + 2x + 5)}$

$$\begin{aligned} \frac{1}{x(x^2 + 2x + 5)} &= \frac{1}{5x} - \frac{x + 2}{5(x^2 + 2x + 5)} = \frac{1}{5x} - \frac{2x + 2}{10(x^2 + 2x + 5)} - \frac{1}{5(x^2 + 2x + 5)} \\ &= \frac{1}{5x} - \frac{2x + 2}{10(x^2 + 2x + 5)} - \frac{1}{5((x + 1)^2 + 2^2)} \end{aligned}$$

Conclusion : $\int \frac{dx}{x(x^2 + 2x + 5)} = \frac{1}{5} \ln|x| - \frac{1}{10} \ln(x^2 + 2x + 5) - \frac{1}{10} \arctan \frac{x + 1}{2} + \lambda.$

– **Cas des fractions rationnelles impaires** :

Poser $t = x^2$ permet d'abaisser le degré pratiquement de moitié. Par exemple :

$$\begin{aligned} \int \frac{dx}{x(x^2 + 1)^2} &= \int \frac{dt}{2t(t + 1)^2} = \int \left(\frac{1}{2t} - \frac{1}{2(t + 1)^2} - \frac{1}{2(t + 1)} \right) dt \\ &= \frac{1}{2} \ln t + \frac{1}{2(t + 1)} - \frac{1}{2} \ln(t + 1) + \lambda \\ &= \ln|x| + \frac{1}{2(x^2 + 1)} - \frac{1}{2} \ln(x^2 + 1) + \lambda \end{aligned}$$

Dans certains, cas, on peut abaisser le degré de manière plus spectaculaire.

Par exemple, avec le changement de variable $t = x^5$:

$$\begin{aligned} \int \frac{dx}{x(x^5 + 1)^2} &= \int \frac{dt}{5t(t + 1)^2} = \frac{1}{5} \ln|t| + \frac{1}{5(t + 1)} - \frac{1}{5} \ln(t + 1) + \lambda \\ &= \ln|x| + \frac{1}{5(x^5 + 1)} - \frac{1}{5} \ln(x^5 + 1) + \lambda \end{aligned}$$

5.3.5 Fractions trigonométriques

Règles de Bioche

Attention : les « règles de Bioche » sont des méthodes empiriques, hors-programme mais utiles.

On considère ici une expression $f(x)$ formée par sommes, produits, quotients et puissances entières de fonctions trigonométriques. Les règles de Bioche consistent à proposer un changement de variable quand l'expression $f(x) dx$ (donc y compris dx) est invariante dans une certaine transformation.

On est alors conduit à une fraction rationnelle (et on a des méthodes pour ça).

– On dit que $\int f(x) dx$ présente « l'invariant du cosinus » si $f(x) dx$ est inchangé dans $x \mapsto -x$.

Dans ce cas, on peut faire le changement de variable $t = \cos(x)$.

Exemple : $\int \frac{dx}{\sin(x)} = \int \frac{dt}{t^2 - 1} = \frac{1}{2} \ln \left| \frac{1 - t}{1 + t} \right| + \lambda = \frac{1}{2} \ln \left| \frac{1 - \cos(x)}{1 + \cos(x)} \right| + \lambda = \ln \left| \tan \left(\frac{x}{2} \right) \right| + \lambda$

– On dit que $\int f(x) dx$ présente « l'invariant du sinus » si $f(x) dx$ est inchangé dans $x \mapsto \pi - x$.

Dans ce cas, on peut faire le changement de variable $t = \sin(x)$.

Exemple : $\int \frac{2 \cos(x) dx}{3 - \cos(2x)} = \int \frac{\cos(x) dx}{1 + \sin^2(x)} = \int \frac{dt}{1 + t^2} = \arctan(t) + \lambda = \arctan \sin(x) + \lambda.$

– On dit que $\int f(x) dx$ présente « l'invariant de la tangente » si $f(x) dx$ est inchangé dans $x \mapsto x + \pi$.

Dans ce cas, on peut faire le changement de variable $t = \tan(x)$.

Exemple :
$$\int \frac{dx}{\sin(x)\cos(x)} = \int \frac{dt}{t} = \ln|t| + \lambda = \ln|\tan(x)| + \lambda.$$

S'il n'y a pas d'invariant

On peut poser $t = \tan\left(\frac{x}{2}\right)$, qui ramène à une fraction rationnelle, mais dont le degré est doublé.

Les règles de Bioche sont donc prioritaires si elles sont applicables.

On rappelle que $\sin(x) = \frac{2t}{1+t^2}$, $\cos(x) = \frac{1-t^2}{1+t^2}$, et $\tan(x) = \frac{2t}{1-t^2}$.

D'autre part, $t = \tan\frac{x}{2} \Rightarrow dt = \frac{1}{2}\left(1 + \tan^2\left(\frac{x}{2}\right)\right) dx = \frac{1}{2}(1+t^2) dx \Rightarrow dx = \frac{2 dt}{1+t^2}$.

Exemple :
$$\int \frac{dx}{1+\sin(x)} = \int \left(\frac{2 dt}{1+t^2} \frac{1}{1 + \frac{2t}{1+t^2}} \right) = \int \frac{2 dt}{(1+t)^2} = -\frac{2}{1+\tan\frac{x}{2}} + \lambda$$

Fractions trigonométriques en $\text{sh}(x)$, $\text{ch}(x)$ ou $\text{th}(x)$

On peut s'inspirer des règles de Bioche : on imagine de remplacer les fonctions hyperboliques par les fonctions circulaires correspondantes, et s'il y a par exemple l'invariant du sinus alors on effectue le changement de variable $t = \text{sh}(x)$ dans l'intégrale initiale.

On peut aussi poser $t = \text{th}\frac{x}{2}$, ou $u = e^x$ (doublement du degré).

Rappel : $\text{sh}(x) = \frac{u^2 - 1}{2u}$, $\text{ch}(x) = \frac{u^2 + 1}{2u}$, $\text{th}(x) = \frac{u^2 - 1}{u^2 + 1}$, et $\left(u = e^x \Rightarrow du = u dx \Rightarrow dx = \frac{du}{u}\right)$.

5.3.6 Primitives avec radicaux

– En présence de $\sqrt{\frac{ax+b}{cx+d}}$, on essaiera le changement de variable $y = \sqrt{\frac{ax+b}{cx+d}}$.

Exemple très simple : on veut calculer $\int \frac{x dx}{\sqrt{x+1}}$.

On pose $t = \sqrt{x+1}$, donc $x+1 = t^2$ et $dx = 2t dt$.

On en déduit :
$$\int \frac{x dx}{\sqrt{x+1}} = \int \frac{2t(t^2-1)}{t} dt = \frac{2}{3}t^3 - 2t + \lambda = \frac{2}{3}(x+1)^{3/2} - 2\sqrt{x+1} + \lambda.$$

– En présence de $\sqrt{ax^2+bx+c}$, on met ax^2+bx+c sous forme canonique :

Si ax^2+bx+c s'écrit $\alpha^2((x+\lambda)^2 + \mu^2)$, on pose $x+\lambda = \mu \text{sh}(t)$.

Si ax^2+bx+c s'écrit $\alpha^2((x+\lambda)^2 - \mu^2)$, on pose $x+\lambda = \pm \mu \text{ch}(t)$.

Si ax^2+bx+c s'écrit $\alpha^2(\mu^2 - (x+\lambda)^2)$, on pose $x+\lambda = \mu \sin t$.

On retiendra surtout que la présence de $\sqrt{1+x^2}$, $\sqrt{1-x^2}$ ou $\sqrt{x^2-1}$ incite aux changements de variables définis respectivement par $x = \text{sh}(t)$, $x = \sin(t)$ ou $x = \pm \text{ch}(t)$.

Exemple : on veut calculer $\int \frac{dx}{(x(2-x))^{3/2}}$.

On pose $y = \sqrt{x(2-x)}$ (mais attention ce n'est pas le changement de variable).

On obtient alors $y^2 = x(2-x) = 2x - x^2 = 1 - (x-1)^2$.

On pose alors $x-1 = \sin(t)$, avec t dans $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$.

Ainsi $dx = \cos(t) dt$ et $y = \sqrt{\cos^2(t)} = \cos(t)$.

On en déduit : $\int \frac{dx}{(x(2-x))^{3/2}} = \int \frac{dx}{y^3} = \int \frac{dt}{\cos^2 t} = \tan t + \lambda = \frac{x-1}{\sqrt{x(2-x)}} + \lambda$.

5.4 Extension aux fonctions à valeurs complexes

5.4.1 Fonctions à valeurs complexes

Définition 5.4.1

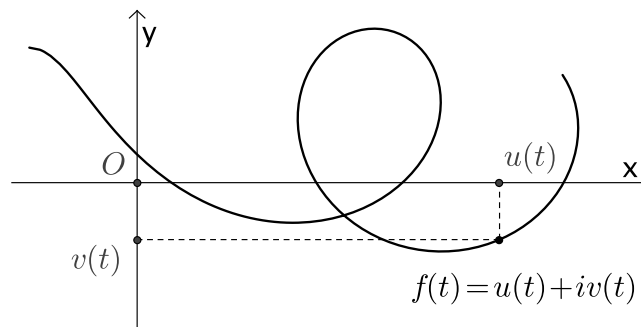
On note $\mathcal{F}(\mathcal{D}, \mathbb{C})$ l'ensemble des fonctions définies sur une partie \mathcal{D} de \mathbb{R} , à valeurs *complexes*.

Une telle fonction numérique *complexe* f est définie de façon unique par la donnée de deux fonctions numériques *réelles* u et v de la manière suivante : $\forall t \in \mathcal{D}, f(t) = u(t) + i v(t)$

Il revient au même d'écrire : $\forall t \in \mathcal{D}, u(t) = \operatorname{Re}(f(t))$ et $v(t) = \operatorname{Im}(f(t))$.

On dit que u est la *partie réelle* de f , et v sa *partie imaginaire*. On note $u = \operatorname{Re}(f)$ et $v = \operatorname{Im}(f)$.

Si on doit représenter graphiquement une telle fonction f , le mieux est d'imaginer un *arc paramétré* du plan, trajectoire du point $M(t)$ d'affixe $f(t) = u(t) + i v(t)$ quand le paramètre t parcourt l'intervalle I .



Si f, g sont dans $\mathcal{F}(\mathcal{D}, \mathbb{C})$, et si α, β sont dans \mathbb{C} , on peut former les fonctions suivantes :

- La fonction $\alpha f + \beta g$ définie sur \mathcal{D} par : $\forall t \in \mathcal{D}, (\alpha f + \beta g)(t) = \alpha f(t) + \beta g(t)$
- La fonction \bar{f} définie sur \mathcal{D} par : $\forall t \in \mathcal{D}, (\bar{f})(t) = \overline{f(t)}$
- La fonction fg définie sur \mathcal{D} par : $\forall t \in \mathcal{D}, (fg)(t) = f(t)g(t)$

5.4.2 Dérivée et intégrale des fonctions complexes

Définition 5.4.2 (continuité d'une fonction à valeurs dans \mathbb{C})

Soit $f : \mathcal{D} \rightarrow \mathbb{C}$ une fonction à valeurs complexes. Soit $u = \operatorname{Re}(f)$ et $v = \operatorname{Im}(f)$.

On dit que f est continue sur I si et seulement si les fonctions u et v sont continues sur I .

Définition 5.4.3 (dérivabilité d'une fonction à valeurs dans \mathbb{C})

Soit $f : \mathcal{D} \rightarrow \mathbb{C}$ une fonction à valeurs complexes. Soit $u = \operatorname{Re}(f)$ et $v = \operatorname{Im}(f)$.

On dit que f est dérivable sur I si et seulement si les fonctions u et v sont dérivables sur I .

On note alors $f' = u' + i v'$ et on dit que f' est la fonction dérivée première de f .

Définition 5.4.4 (intégrale sur un segment d'une fonction complexe)

Soit $f : \mathcal{D} \rightarrow \mathbb{C}$ une fonction à valeurs complexes. Soit $u = \operatorname{Re}(f)$ et $v = \operatorname{Im}(f)$.

Pour tous a, b de I , on pose : $\int_a^b f(x) dx = \int_a^b u(x) dx + i \int_a^b v(x) dx$

Remarques

On retiendra de ce qui précède que, concernant les fonctions à valeurs complexes, tout revient à procéder (continuité, dérivées, intégrales) séparément sur la partie réelle et sur la partie imaginaire.

– On dit que $f : I \rightarrow \mathbb{C}$ est n fois dérivable si u, v sont n fois dérivables, et on écrit $f^{(n)} = u^{(n)} + i v^{(n)}$.
En d'autres termes, on peut écrire $\operatorname{Re}(f^{(n)}) = (\operatorname{Re} f)^{(n)}$ et $\operatorname{Im}(f^{(n)}) = (\operatorname{Im} f)^{(n)}$.

– On définit de façon évidente les primitives d'une fonction $f : \mathcal{D} \rightarrow \mathbb{C}$:

$$\int f(x) dx = \int u(x) dx + i \int v(x) dx \quad (\text{à une constante additive complexe } \lambda \text{ quelconque}).$$

– De même, par définition de l'intégrale de f , on a les égalités :

$$\operatorname{Re}\left(\int_a^b f(x) dx\right) = \int_a^b \operatorname{Re}(f(x)) dx \quad \operatorname{Im}\left(\int_a^b f(x) dx\right) = \int_a^b \operatorname{Im}(f(x)) dx \quad \overline{\int_a^b f(x) dx} = \int_a^b \overline{f(x)} dx$$

5.4.3 Extension des résultats relatifs aux fonctions réelles

Des définitions précédentes, il découle que de nombreuses propriétés établies pour des fonctions à valeurs réelles s'étendent sans difficulté au cas des fonctions à valeurs complexes.

En revanche, et c'est important :

Tout ce qui a un rapport avec la relation d'ordre dans l'ensemble d'arrivée \mathbb{R} n'a plus de sens pour une fonction f à valeurs complexes. On ne parlera donc **jamais** de la monotonie ou des extremums de f (ça n'existe pas).

De même, si f est continue sur l'intervalle I et à valeurs dans \mathbb{C} , l'image $f(I)$ est un arc du plan complexe (mais certainement pas *un intervalle*). Le *signe* de $f : I \rightarrow \mathbb{C}$, ou le théorème des valeurs intermédiaires, ou le *théorème de la bijection réciproque* n'ont ici plus aucune signification.

Voici **ce qui reste vrai pour la dérivabilité**, dans le cas des fonctions à valeurs complexes :

– les égalités : $(\alpha f + \beta g)' = \alpha f' + \beta g' \quad (fg)' = f'g + fg' \quad \left(\frac{1}{g}\right)' = -\frac{g'}{g^2} \quad \left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}$

- avec $f : I \rightarrow \mathbb{R}$ et $g : J \rightarrow \mathbb{C}$ et $f(I) \subset J$, on a encore $(g \circ f)' = f'(g' \circ f)$
- la fonction dérivable $f : I \rightarrow \mathbb{C}$ est constante si et seulement si f' est identiquement nulle.
- la fonction $f - g$ est constante sur I si et seulement si on a $f' = g'$ sur I .

Voici **ce qui reste vrai pour l'intégration**, dans le cas des fonctions à valeurs complexes :

- l'égalité $\int_a^b f(x) dx = F(b) - F(a)$, où F est une primitive *quelconque* de f sur I .
- la linéarité de l'intégrale, la relation de Chasles, l'intégration par parties.

En revanche, dans le cas des fonctions à valeurs dans \mathbb{C} , il n'est évidemment plus question de positivité et de croissance de l'intégrale.

Un exemple à connaître, et à méditer !

Soit ω un nombre complexe *non réel*.

Considérons la fonction complexe définie sur \mathbb{R} par $f(x) = \frac{1}{x - \omega}$.

La fonction f est continue sur \mathbb{R} et on se propose d'en calculer les primitives.

Une erreur (catastrophique) consiste à écrire $\int \frac{dx}{x - \omega} = \ln|x - \omega| + \lambda$, ou $\int \frac{dx}{x - \omega} = \ln(x - \omega) + \lambda$.

La deuxième expression n'a aucun sens car la fonction $t \mapsto \ln(t)$ n'est définie, pour nous, que sur \mathbb{R}^{+*} .

En revanche (et avant de savoir ce qu'il convient réellement de faire), on peut s'interroger sur la première expression, car l'application $\varphi : x \mapsto \ln|x - \omega|$ est définie sur \mathbb{R} .

Posons $\omega = a + ib$, avec (a, b) dans $\mathbb{R} \times \mathbb{R}^*$.

Pour tout réel x , on a $\varphi(x) = \ln \sqrt{(x - a)^2 + b^2} = \frac{1}{2} \ln((x - a)^2 + b^2)$ donc $\varphi'(x) = \frac{x - a}{(x - a)^2 + b^2}$.

Mais $f(x) = \frac{1}{x - \omega} = \frac{x - \bar{\omega}}{|x - \omega|^2} = \frac{x - a + ib}{(x - a)^2 + b^2}$, et on constate que $\varphi'(x)$ n'est *jamais* égal à $f(x)$.

Pour calculer correctement les primitives de f , voici la méthode :

$$\begin{aligned} \int f(x) dx &= \int \frac{dx}{x - \omega} = \int \frac{x - \bar{\omega}}{|x - \omega|^2} dx = \int \frac{x - a + ib}{(x - a)^2 + b^2} dx = \int \frac{x - a}{(x - a)^2 + b^2} dx + ib \int \frac{dx}{(x - a)^2 + b^2} \\ &= \frac{1}{2} \ln((x - a)^2 + b^2) + i \arctan \frac{x - a}{b} + \lambda \quad (\text{avec } \lambda \text{ quelconque dans } \mathbb{C}) \end{aligned}$$

Si on veut *vraiment* exprimer les primitives de f à l'aide de ω (plutôt que $a = \operatorname{Re}\omega$ et $b = \operatorname{Im}\omega$), et également si on a le goût du risque, on peut observer que le résultat précédent s'écrit :

$$\int \frac{dx}{x - \omega} = \ln|x - \omega| + i \arg(x - \omega) + \lambda \quad \text{avec } \lambda \text{ quelconque dans } \mathbb{C}$$

(on choisit par exemple d'utiliser la détermination principale de l'argument)

Si pour z dans \mathbb{C}^* on note $\psi(z) = \ln|z| + i \arg(z)$, on sait que $\exp(\psi(z)) = z$.

En conclusion, l'évocation prématurée du logarithme dans le calcul de $\int \frac{dx}{x - \omega}$ reste une erreur flagrante,

mais on a tout de même $\int \frac{dx}{x - \omega} = \psi(x - \omega) + \lambda$, avec $\exp(\psi(x - \omega)) = x - \omega$ pour tout x de \mathbb{R} .

5.4.4 Cas de la fonction exponentielle complexe

Proposition 5.4.1

Soit $\varphi : I \rightarrow \mathbb{C}$ une fonction dérivable. Soit $f = \exp(\varphi)$, définie sur I par $f(x) = \exp(\varphi(x))$. Alors f est dérivable sur I et, pour tout x de I : $f'(x) = \varphi'(x) \exp(\varphi(x))$.

Applications

- Un cas particulier du résultat précédent est : si $f(x) = e^{\omega x}$, alors $f'(x) = \omega e^{\omega x}$.
- Si ω est dans \mathbb{C}^* , on peut écrire : $\int e^{\omega x} dx = \frac{1}{\omega} e^{\omega x} + \lambda$ (avec λ quelconque dans \mathbb{C}).
- On va utiliser cette idée pour reprendre un exemple déjà vu.

Soit (α, β) dans $\mathbb{R}^2 \setminus \{(0, 0)\}$. Posons $C_{\alpha, \beta} = \int e^{\alpha x} \cos(\beta x) dx$ et $S_{\alpha, \beta} = \int e^{\alpha x} \sin(\beta x) dx$

On introduit $\omega = \alpha + i\beta$, puis $Z_{\alpha, \beta} = C_{\alpha, \beta} + iS_{\alpha, \beta} = \int e^{\alpha x} (\cos(\beta x) + i \sin(\beta x)) dx = \int e^{\omega x} dx$

On trouve : $Z_{\alpha, \beta} = \frac{1}{\omega} e^{\omega x} + \lambda = \frac{\bar{\omega}}{|\omega|^2} e^{(\alpha+i\beta)x} + \lambda = \frac{\alpha - i\beta}{\alpha^2 + \beta^2} e^{\alpha x} (\cos(\beta x) + i \sin(\beta x)) + \lambda$.

En prenant la partie réelle et la partie imaginaire, on retrouve (avec λ quelconque dans \mathbb{R}) :

$$\begin{cases} C_{\alpha, \beta} = \operatorname{Re}(Z_{\alpha, \beta}) = \int e^{\alpha x} \cos(\beta x) dx = \frac{e^{\alpha x}}{\alpha^2 + \beta^2} (\alpha \cos(\beta x) + \beta \sin(\beta x)) + \lambda \\ S_{\alpha, \beta} = \operatorname{Im}(Z_{\alpha, \beta}) = \int e^{\alpha x} \sin(\beta x) dx = \frac{e^{\alpha x}}{\alpha^2 + \beta^2} (\alpha \sin(\beta x) - \beta \cos(\beta x)) + \lambda \end{cases}$$

5.5 Équations différentielles $y' + a(x)y = b(x)$

Dans cette section, il sera question de fonctions à valeurs réelles ou complexes.

Ces fonctions seront définies sur un intervalle ouvert non vide I de \mathbb{R} .

On notera \mathbb{K} pour désigner indifféremment \mathbb{R} et \mathbb{C} , et on parlera de fonctions à valeurs dans \mathbb{K} .

5.5.1 Position du problème

Définition 5.5.1

Soit I un intervalle ouvert non vide.

Soit $x \mapsto a(x)$ et $x \mapsto b(x)$ deux fonctions continues sur I , à valeurs dans \mathbb{K} .

On dit que (E) : $y' + a(x)y = b(x)$ est une *équation différentielle linéaire du premier ordre*.

On note (H) l'équation différentielle : $y' + a(x)y = 0$.

On dit que (H) est l'*équation différentielle homogène* associée à (E).

Compléments sur la définition

Résoudre (on dit aussi « intégrer ») l'équation (E) (resp. l'équation (H)) c'est trouver toutes les fonctions dérivables $y : I \rightarrow \mathbb{K}$ qui vérifient l'égalité (E) (resp. l'égalité (H)) sur I .

On pourra noter \mathcal{S}_E (resp. \mathcal{S}_H) l'ensemble des solutions de (E) (resp. de (H)) sur I .

À la fois pour (E) et pour (H) , la « solution générale » désigne une expression (en fonction d'une constante d'intégration, comme on le verra) de toutes les « solutions particulières ».

Cas particulier où a est une constante

Proposition 5.5.1 (solution générale de $y' + ay = 0$, où a est une constante)

On considère l'équation homogène $(H) : y' + ay = 0$ où a est un élément de \mathbb{K} (\mathbb{R} ou \mathbb{C}).

La solution générale de (H) sur \mathbb{R} est donnée par $y(x) = \lambda e^{-ax}$, où λ est quelconque dans \mathbb{K} .

Remarque importante sur l'intervalle de résolution

Les résultats de cette section peuvent être étendus aux équations différentielles qui se présentent sous la forme $u(x)y'(x) + v(x)y(x) = w(x)$, où u, v, w sont des fonctions numériques continues.

Mais il faut alors obligatoirement se placer sur un intervalle sur lequel la fonction $x \mapsto u(x)$ ne s'annule pas (de manière à revenir, en divisant par $u(x)$, à la forme précédente, dite « normalisée »).

Par exemple, pour résoudre $x(1-x)y'(x) + v(x)y(x) = w(x)$, il faudra absolument commencer par dire qu'on se place sur l'intervalle $I =]-\infty, 0[$, ou $I =]0, 1[$, ou $I =]1, +\infty[$.

5.5.2 Résolution de l'équation homogène $y' + a(x)y = 0$

Proposition 5.5.2 (solution générale de $(H) : y' + a(x)y = 0$)

On considère l'équation $(H) : y' + a(x)y = 0$, sur l'intervalle I .

Soit A une primitive particulière de $x \mapsto b(x)$ sur I .

La solution générale de (H) sur I s'écrit $y(x) = \lambda e^{-A(x)}$, où λ est quelconque dans \mathbb{K} .

Remarques et exemples

Le résultat précédent montre que l'ensemble \mathcal{S}_H des solutions y de (H) sur I est égal à l'ensemble des multiples d'une solution particulière y_0 non nulle de (H) sur I . On exprime cette situation en disant que \mathcal{S}_H est une droite vectorielle. On voit d'ailleurs qu'une solution de (H) sur I , si elle n'est pas la solution nulle, ne s'annule jamais sur I .

Si on « devine » une solution non nulle de (H) sur I , alors on connaît la solution générale (sans avoir à passer par la formule précédente).

Par exemple, on constate que la fonction $x \mapsto \frac{1}{x}$ est solution de $(H) : y' + \frac{y}{x} = 0$ sur \mathbb{R}^{-*} et sur \mathbb{R}^{+*} .

Sur $I = \mathbb{R}^{-*}$ ou sur $I = \mathbb{R}^{+*}$, la solution générale de (H) est donc l'ensemble des $x \mapsto \frac{\lambda}{x}$, ($\lambda \in \mathbb{K}$).

5.5.3 Résolution de l'équation $y' + a(x)y = b(x)$

Proposition 5.5.3 (solution générale de $(E) : y' + a(x)y = b(x)$)

On considère l'équation différentielle $(E) : y' + a(x)y = b(x)$, sur un intervalle I .

La solution générale de (E) sur I est donnée par $y_\lambda(x) = (B(x) + \lambda)e^{-A(x)}$, où :

– la fonction $x \mapsto A(x)$ est une primitive particulière de $x \mapsto a(x)$ sur I

- le scalaire λ est quelconque dans \mathbb{K} .
- la fonction $x \mapsto B(x)$ est une primitive particulière de $x \mapsto b(x)e^{A(x)}$ sur I

Proposition 5.5.4 (structure de l'ensemble des solutions de (E))

On considère les équations $(E) : y' + a(x)y = b(x)$ et $(H) : y' + a(x)y = 0$.

La solution générale de (E) est la somme de celle de (H) et d'une solution particulière de (E) .

Deux exemples

- Considérons par exemple l'équation différentielle $(E) : xy' + y = 2x$.

À cause du coefficient de y' , il faut obligatoirement se placer sur $I = \mathbb{R}^{-*}$ ou $I = \mathbb{R}^{+*}$.

On « remarque » qu'une solution particulière de $(E) : xy' + y = 2x$ est la fonction $x \mapsto x$.

On sait que la solution générale de (H) sur $I = \mathbb{R}^{+*}$ ou \mathbb{R}^{-*} s'écrit $y(x) = \frac{\lambda}{x}$, avec λ dans \mathbb{K} .

La solution générale de (E) sur $I = \mathbb{R}^{+*}$ ou \mathbb{R}^{-*} est donc $y(x) = x + \frac{\lambda}{x}$, avec λ dans \mathbb{K} .

- Considérons l'équation différentielle $(E) : \cos(x)y' + \sin(x)y = 1$, sur $I = \left] -\frac{\pi}{2}, \frac{\pi}{2} \right[$.

On « remarque » qu'une solution particulière de (E) sur I est $x \mapsto \sin(x)$.

On « remarque » qu'une solution particulière (non nulle) de (H) sur I est $x \mapsto \cos(x)$.

La solution générale de (E) sur I s'écrit donc $y(x) = \sin(x) + \lambda \cos(x)$, avec λ dans \mathbb{K} .

5.5.4 Principe de superposition**Proposition 5.5.5**

On considère l'équation différentielle $(E) : y' + a(x)y = b(x)$, sur un intervalle I .

On suppose que la fonction $x \mapsto b(x)$ s'écrit $b(x) = \sum_{k=1}^n \alpha_k b_k(x)$ (les α_k étant des scalaires).

Soit $x \mapsto y_k(x)$ une solution particulière sur I de l'équation différentielle $(E_k) : y' + a(x)y = b_k(x)$.

Alors une solution particulière sur I de (E) est : $x \mapsto y(x) = \sum_{k=1}^n \alpha_k y_k(x)$

Remarques et exemples

- Considérons par exemple l'équation différentielle $(E) : xy' + y = 2x + 1 + \ln(x)$, sur $I = \mathbb{R}^{+*}$.

Une solution particulière de $(E_1) : xy' + y = 2x$ sur I est $x \mapsto x$.

Une solution particulière de $(E_2) : xy' + y = 1 + \ln(x)$ sur I est $x \mapsto \ln(x)$.

Une solution particulière de (E) sur I est donc $x \mapsto x + \ln(x)$.

- Soit $a : I \rightarrow \mathbb{R}$ (donc à valeurs réelles) et $b : I \rightarrow \mathbb{C}$ (donc à valeurs complexes), continues sur I .

Soit $x \mapsto \omega(x)$ une solution particulière de $(E) : y'(x) + a(x)y(x) = b(x)$.

Alors $x \mapsto \operatorname{Re}(\omega(x))$ est solution particulière sur I de l'équation $y'(x) + a(x)y(x) = \operatorname{Re}(b(x))$.

De même, $x \mapsto \operatorname{Im}(\omega(x))$ est solution particulière sur I de l'équation $y'(x) + a(x)y(x) = \operatorname{Im}(b(x))$.

Cette idée sert quand $b(x) = P(x) \cos(\beta x)$, ou $P(x) \sin(\beta x)$ (P un polynôme à coefficients réels).

Les calculs sont en effet plus simples en écrivant le second membre sous la forme $P(x)e^{i\beta x}$.

Il suffit alors de prendre la partie réelle (ou imaginaire) de la solution particulière obtenue.

5.5.5 Méthode de variation de la constante

On considère les équations $(H) : y'(x) + a(x)y(x) = 0$ et $(E) : y'(x) + a(x)y(x) = b(x)$.

On rappelle que les fonctions a et b sont continues, à valeurs dans \mathbb{K} .

Soit $x \mapsto h(x)$ une solution de (H) sur I , non nulle (donc ne s'annulant pas sur I).

On sait que la solution générale de (H) sur I s'écrit $y : x \mapsto \lambda h(x)$, avec $\lambda \in \mathbb{K}$.

Pour résoudre complètement (E) sur I , il suffit d'en connaître une solution particulière.

On cherche une telle solution sous la forme $y : x \mapsto \lambda(x) h(x)$, où λ est maintenant une fonction dérivable sur I à valeurs dans \mathbb{K} (on fait donc « varier la constante »).

Avec ces notations, et pour tout x de l'intervalle I :

$$\begin{aligned} y'(x) + a(x)y(x) = b(x) &\Leftrightarrow \lambda'(x)h(x) + \lambda(x)h'(x) + a(x)\lambda(x)h(x) = b(x) \\ &\Leftrightarrow \lambda'(x)h(x) + \lambda(x)(h'(x) + a(x)h(x)) = b(x) \\ &\Leftrightarrow \lambda'(x)h(x) = b(x) \quad (\text{car } h'(x) + a(x)h(x) \equiv 0) \\ &\Leftrightarrow \lambda'(x) = \frac{b(x)}{h(x)} \quad (\text{ce qui détermine la fonction } x \mapsto \lambda(x) \text{ à une constante } \mu \text{ près}) \end{aligned}$$

Si Γ est une primitive de $x \mapsto \frac{b(x)}{h(x)}$ sur I , on a ainsi obtenu $x \mapsto y(x) = \Gamma(x)h(x) + \mu h(x)$, avec $\mu \in \mathbb{K}$.

Cette méthode donne donc l'ensemble des solutions de (E) sur I .

Remarque importante : il n'est pas nécessaire de calculer explicitement $h'(x)$ quand on applique la méthode, car cette dérivée se simplifie toujours !

5.5.6 Problème de Cauchy

Proposition 5.5.6 (existence et unicité de la solution d'un problème de Cauchy)

On considère l'équation différentielle $(E) : y' + b(x)y = c(x)$, sur l'intervalle I .

Soit x_0 un point I , et soit y_0 un élément quelconque de \mathbb{K} .

Il existe une unique solution de (E) sur I satisfaisant à la condition initiale $y(x_0) = y_0$.

Trouver cette solution, c'est résoudre le problème de Cauchy relatif à ces conditions initiales.

Supposons $\mathbb{K} = \mathbb{R}$ (toutes les fonctions sont donc à valeurs réelles).

Les courbes représentatives des solutions de (E) sont appelées *courbes intégrales* de (E) .

Ce que dit le résultat précédent, c'est que par tout point $M(x_0, y_0)$ de $I \times \mathbb{R}$, il passe une et une seule courbe intégrale de (E) .

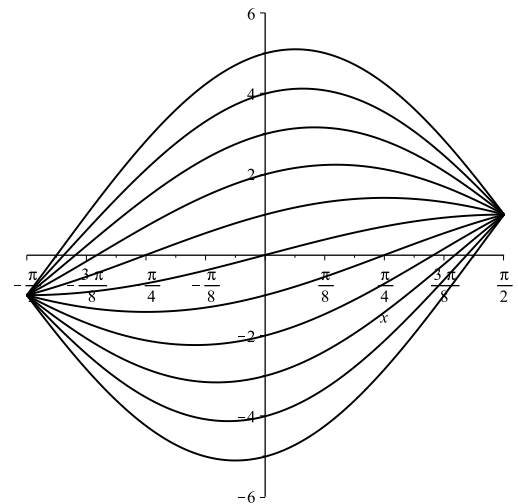
Bien sûr cela n'est valable que sur l'intervalle ouvert sur lequel on résout l'équation (E) , et cesse d'être vrai aux extrémités de cet intervalle.

On voit ici quelques solutions de l'équation :

$$(E) : \cos(x) y' + \sin(x) y = 1, \text{ sur } I = \left] -\frac{\pi}{2}, \frac{\pi}{2} \right[$$

Ce sont les $y(x) = \sin(x) + \lambda \cos(x)$ ($\lambda \in \mathbb{R}$)

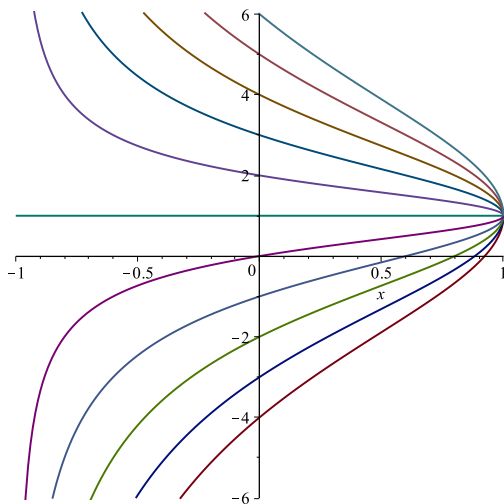
Sur cet exemple, on voit que les solutions se prolongent par continuité de la même manière aux extrémités de I .



$$\cos(x) y' + \sin(x) y = 1$$

$$\text{solutions : } y(x) = \sin(x) + \lambda \cos(x)$$

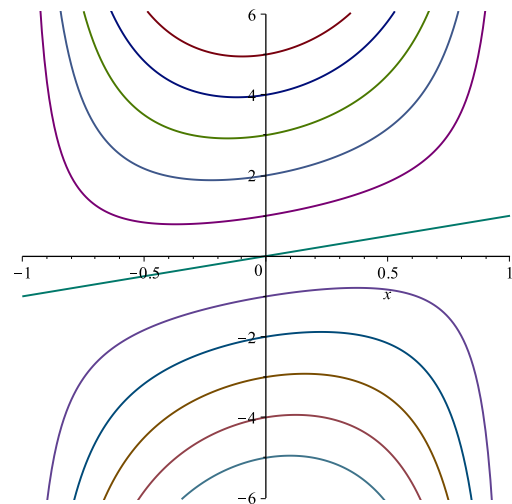
Voici deux autres exemples, avec des comportements différents aux bornes de I (ici $I =]-1, 1[$)



$$(1 - x^2)y' + y = 1$$

$$\text{solution générale } y(x) = 1 + \lambda \sqrt{\frac{1-x}{1+x}}$$

$$\text{solution particulière } y(x) = 1$$



$$(1 - x^2)y' - 2xy = 1 - 3x^2$$

$$\text{solution générale } y(x) = x + \frac{\lambda}{1-x^2}$$

$$\text{solution particulière } y(x) = x$$

On retiendra que les « courbes intégrales » d'une équation différentielle $(E) : y'(x) + a(x) = b(x)$ ne se « croisent » jamais sur l'intervalle ouvert I où on résout l'équation. L'ensemble de ces courbes constitue donc une partition de la bande verticale délimitée par l'intervalle I (on suppose ici $\mathbb{K} = \mathbb{R}$).

5.6 Équations différentielles du 2nd ordre

Dans cette section, il sera question de fonctions à valeurs réelles ou complexes.

Ces fonctions seront définies sur un intervalle ouvert non vide I de \mathbb{R} .

On notera \mathbb{K} pour désigner indifféremment \mathbb{R} et \mathbb{C} , et on parlera de fonctions à valeurs dans \mathbb{K} .

5.6.1 Position du problème

Définition 5.6.1

Soit I un intervalle ouvert non vide. Soit a, b deux éléments de \mathbb{K} .

Soit $x \mapsto f(x)$ une fonction continue sur I , à valeurs réelles ou complexes.

On considère l'équation (E) : $y'' + ay' + by = f(x)$.

On dit que (E) est une *équation différentielle linéaire du second ordre à coefficients constants*.

On note (H) l'équation différentielle : $y'' + ay' + by = 0$.

On dit que (H) est l'*équation différentielle homogène* associée à (E).

Compléments sur la définition

Résoudre (ou « intégrer ») l'équation (E) (resp. (H)) c'est trouver toutes les fonctions deux fois dérivables $y : I \rightarrow \mathbb{K}$ (resp. $y : \mathbb{R} \rightarrow \mathbb{K}$) qui vérifient l'égalité (E) sur I (resp. (H) sur \mathbb{R}).

On pourra noter \mathcal{S}_E (resp. \mathcal{S}_H) l'ensemble des solutions de (E) de I (resp. de (H) sur \mathbb{R}).

À la fois pour (E) et pour (H), la « solution générale » désigne une expression (en fonction de deux constantes d'intégration, comme on le verra) de toutes les « solutions particulières ».

5.6.2 Résolution de l'équation homogène

Proposition 5.6.1 (équation caractéristique)

Soit a, b dans \mathbb{K} , et soit (H) l'équation différentielle : $y'' + ay' + by = 0$.

On dit que (C) : $r^2 + ar + b = 0$ est l'équation caractéristique de (H).

L'application $y : x \mapsto e^{rx}$ est solution de (H) sur \mathbb{R} et si seulement si r est solution de (C).

Proposition 5.6.2 (solution générale de (H) dans le cas complexe)

Soit a, b dans \mathbb{C} , et soit (H) l'équation différentielle : $y'' + ay' + by = 0$.

Soit $\Delta = a^2 - 4b$, le discriminant de l'équation caractéristique (C) : $r^2 + ar + b = 0$

– Si $\Delta \neq 0$, l'équation (C) possède deux solutions complexes distinctes r et s .

La solution générale de (H) sur \mathbb{R} s'écrit alors : $y(x) = \lambda e^{rx} + \mu e^{sx}$, avec λ et μ dans \mathbb{C} .

– Si $\Delta = 0$, l'équation (C) possède une solution double r dans \mathbb{C} .

La solution générale de (H) sur \mathbb{R} s'écrit alors : $y(x) = (\lambda x + \mu) e^{rx}$, avec λ et μ dans \mathbb{C} .

Proposition 5.6.3 (solution générale de (H) dans le cas réel)

Soit a, b dans \mathbb{R} , et soit (H) l'équation différentielle : $y'' + ay' + by = 0$.

Soit $\Delta = a^2 - 4b$, le discriminant de l'équation caractéristique (C) : $r^2 + ar + b = 0$

– Si $\Delta > 0$, l'équation (C) possède deux solutions réelles distinctes r et s .

La solution générale de (H) sur \mathbb{R} s'écrit : $y(x) = \lambda e^{rx} + \mu e^{sx}$, avec λ et μ dans \mathbb{R} .

– Si $\Delta = 0$, l'équation (C) possède une solution double r dans \mathbb{R} .

La solution générale de (H) sur \mathbb{R} s'écrit : $y(x) = (\lambda x + \mu) e^{rx}$, avec λ et μ dans \mathbb{R} .

- Si $\Delta < 0$, l'équation (C) possède deux solutions complexes conjuguées distinctes r et \bar{r} .
Posons $r = \alpha + i\beta$, avec $(\alpha, \beta) \in \mathbb{R} \times \mathbb{R}^*$.
La solution générale de (H) sur \mathbb{R} est $y(x) = e^{\alpha x}(\lambda \cos(\beta x) + \mu \sin(\beta x))$, avec λ et μ dans \mathbb{R} .

Structure de l'ensemble des solutions de (H)

Dans tous les cas, la solution générale y de (H) sur \mathbb{R} s'écrit $x \mapsto y(x) = \lambda h(x) + \mu k(x)$, où $x \mapsto h(x)$ et $x \mapsto k(x)$ sont deux solutions particulières de (H) non nulles et non proportionnelles.

On exprime cette situation en disant que la solution générale de (H) sur \mathbb{R} est un *plan vectoriel*, dont une *base* est constituée des fonctions $x \mapsto h(x)$ et $x \mapsto k(x)$.

Deux cas particuliers importants

On suppose ici $\mathbb{K} = \mathbb{R}$. Soit ω un réel strictement positif.

- La solution générale de $y'' + \omega^2 y = 0$ s'écrit $y(x) = \lambda \cos(\omega x) + \mu \sin(\omega x)$, avec λ et μ dans \mathbb{R} .
Elle s'écrit aussi : $y(x) = A \cos(\omega(x - x_0))$, avec A et x_0 dans \mathbb{R} .
- La solution générale de $y'' - \omega^2 y = 0$ s'écrit $y(x) = \lambda e^{\omega x} + \mu e^{-\omega x}$, avec λ et μ dans \mathbb{R} .
Elle s'écrit aussi $y(x) = \lambda \operatorname{ch}(\omega x) + \mu \operatorname{sh}(\omega x)$, avec λ et μ dans \mathbb{R} .

5.6.3 Forme des solutions de l'équation complète

Proposition 5.6.4 (structure de la solution générale de (E))

Soit $f : I \rightarrow \mathbb{K}$ une fonction continue sur I . Soit a, b deux éléments de \mathbb{K} .

L'équation (E) : $y'' + ay' + by = f(x)$ possède des solutions sur \mathbb{R} . Soit $y_0 : I \rightarrow \mathbb{K}$ l'une d'elles.

Alors $y : I \rightarrow \mathbb{K}$ est solution de (E) si et seulement si $y - y_0$ est solution de (H).

En d'autres termes, la solution générale de (E) sur I s'obtient en ajoutant à une solution particulière de (E) la solution générale de (H) sur \mathbb{R} .

Remarque importante

On sait que la solution générale de (H) sur \mathbb{R} s'écrit $x \mapsto \lambda h(x) + \mu k(x)$, où h et k sont deux solutions particulières de (H) non nulles et non proportionnelles.

La solution générale de (E) sur I s'écrit donc $x \mapsto y(x) = y_0(x) + \lambda h(x) + \mu k(x)$. En particulier l'expression de cette solution générale fait toujours apparaître deux constantes λ et μ arbitraires.

Cas où on « devine » une solution particulière de (E)

On possède une « formule » pour la solution générale de (H), et tout le problème est de trouver une solution particulière de (E). Le cas le plus favorable est celui où on « devine » une telle solution.

Par exemple, une solution particulière de (E) : $y'' + y = 1$ est évidemment $y = 1$.

La solution générale de (E) est donc donnée par $y(x) = \alpha \cos(x) + \beta \sin(x) + 1$.

La proposition suivante est utile quand le second membre est combinaison linéaire de fonctions simples :

Proposition 5.6.5 (principe de superposition)

On considère l'équation (E) : $y'' + ay' + by = f(x)$, où $f : I \rightarrow \mathbb{K}$ est continue.

On suppose $f(x) = \sum_{k=1}^n \alpha_k f_k(x)$ (où les α_k sont scalaires, et les $f_k : I \rightarrow \mathbb{K}$ sont continues sur I).

Soit $x \mapsto y_k(x)$ une solution particulière sur I de l'équation $(E_k) : y'' + ay' + by = f_k(x)$.

Alors une solution particulière sur I de (E) est : $x \mapsto y(x) = \sum_{k=1}^n \alpha_k y_k(x)$.

5.6.4 Problème de Cauchy**Proposition 5.6.6** (Problème de Cauchy)

On considère l'équation (E) : $y'' + ay' + by = f(x)$, où $f : I \rightarrow \mathbb{K}$ est continue.

Soit x_0 un élément de I , et (y_0, m) un élément quelconque de \mathbb{K}^2 .

Alors il existe une unique solution de (E) sur I satisfaisant aux conditions initiales $\begin{cases} y(x_0) = y_0 \\ y'(x_0) = m \end{cases}$

La trouver c'est résoudre le problème de Cauchy relatif à ces conditions initiales.

Interprétation graphique

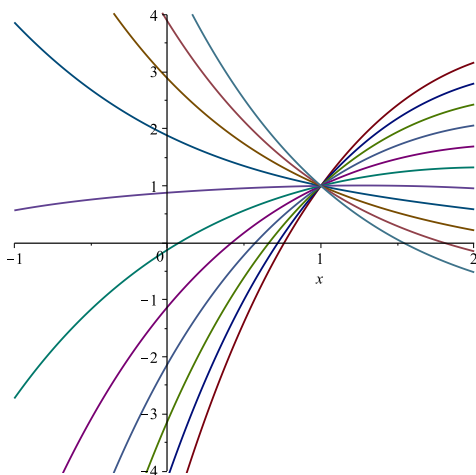
Supposons $\mathbb{K} = \mathbb{R}$ (toutes les fonctions sont donc ici à valeurs réelles).

Par tout point $M(x_0, y_0)$, avec x_0 dans I et y_0 dans \mathbb{R} , il passe une courbe intégrale unique ayant une pente donnée m . Il faut donc « deux conditions » au même point x_0 (l'une portant sur la valeur de la solution, l'autre sur sa dérivée) pour s'assurer de l'unicité d'une solution.

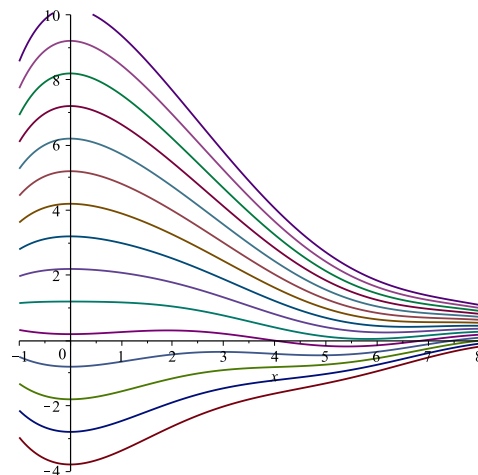
Si on ne fixe qu'une des deux conditions (soit $y(x_0) = y_0$, soit $y'(x_0) = m$), il reste encore une infinité de solutions dont les représentations graphiques forment :

- dans le premier cas : un faisceau de courbes passant par $M(x_0, y_0)$,
- dans le second cas, un faisceau de courbes dont les tangentes au point d'abscisse x_0 sont parallèles.

Voici deux exemples, avec l'équation différentielle $4y'' + 4y' + y = \cos(x)$.



$4y'' + 4y' + y = \cos(x)$
avec la condition $y(1) = 1$



$4y'' + 4y' + y = \cos(x)$
avec la condition $y'(0) = 0$

5.6.5 Quelques exemples

Conjuguée d'une solution complexe

On considère $(E) : y'' + ay' + by = f(x)$, avec a, b réels mais $f : I \rightarrow \mathbb{C}$ continue à valeurs complexes.

Soit $x \mapsto \varphi(x)$ une solution particulière de (E) sur l'intervalle I .

Alors l'application conjugquée $x \mapsto \overline{\varphi}(x)$ est une solution de $(\overline{E}) : y'' + ay' + by = \overline{f}(x)$ sur I .

Par superposition des deux seconds membres $x \mapsto f(x)$ et $x \mapsto \overline{f}(x)$, on en déduit que :

- la fonction $x \mapsto \operatorname{Re}(\varphi(x))$ est solution de $y'' + ay' + by = \operatorname{Re}(f(x))$ sur I .
- la fonction $x \mapsto \operatorname{Im}(\varphi(x))$ est solution de $y'' + ay' + by = \operatorname{Im}(f(x))$ sur I .

Par exemple, pour trouver une solution de l'équation $y'' + y' + 2y = x \cos(x)$, il suffit de trouver une solution φ de l'équation $y'' + y' + 2y = xe^{ix}$ et d'en prendre la partie réelle.

Seconds membres particuliers

On considère l'équation $(E) : y'' + ay' + by = f(x)$, avec a, b dans \mathbb{K} et $f : I \rightarrow \mathbb{K}$ continue.

Soit $(C) : r^2 + br + c = 0$ son équation caractéristique.

On suppose ici que $f(x) = P(x)e^{mx}$, où P est un polynôme à coefficients dans \mathbb{K} , et où m est dans \mathbb{K} .

Alors on peut chercher une solution particulière de (E) :

- sous la forme $y(x) = Q(x)e^{mx}$, avec $\deg(Q) = \deg(P)$, si m n'est pas racine de (C) .
- sous la forme $y(x) = xQ(x)e^{mx}$, avec $\deg(Q) = \deg(P)$, si m est racine simple de (C) .
- sous la forme $y(x) = x^2Q(x)e^{mx}$, avec $\deg(Q) = \deg(P)$, si m est racine double de (C) .

Une fois qu'on a compris sous quelle forme chercher cette solution particulière de l'équation (E) , il suffit d'écrire le polynôme Q avec des coefficients indéterminés, d'injecter cette expression de $y(x)$ dans (E) et de procéder à une identification pour déterminer les coefficients de Q .

Si le second membre est de la forme $f(x) = P(x) \cos(\omega x) + Q(x) \sin(\omega x)$, on se ramène à ce qui précède en écrivant $\cos(\omega x)$ et $\sin(\omega x)$ en fonction de $e^{i\omega x}$ et en utilisant le principe de superposition. On peut également chercher une solution particulière sous la forme $f(x) = R(x) \cos(\omega x) + S(x) \sin(\omega x)$ (en attendant l'identification pour déterminer le degré de R et S).

Un exemple

On considère l'équation différentielle $(E) : y'' - 3y' + 2y = (2x + 3)e^x + (4x - 5)e^{-x}$.

L'équation caractéristique est $(C) : r^2 - 3r + 2 = 0$, de racines $r = 1$ et $r = 2$.

La solution générale de (H) sur \mathbb{R} s'écrit : $y(x) = \lambda e^x + \mu e^{2x}$, avec λ, μ dans \mathbb{K} (disons dans \mathbb{R}).

On écrit le second membre $f(x) = g(x) + h(x)$, avec $g(x) = (2x + 3)e^x$ et $h(x) = (4x - 5)e^{-x}$.

On considère donc séparément les équations
$$\begin{cases} (E_1) : y'' - 3y' + 2y = (2x + 3)e^x \\ (E_2) : y'' - 3y' + 2y = (4x - 5)e^{-x} \end{cases}$$

Tout d'abord $r = 1$ est une racine simple de (C) .

On cherche donc une solution $(E_1) : y'' - 3y' + 2y = (2x + 3)e^x$ sous la forme $y_1(x) = x(ax + b)e^x$.

Ensuite $r = -1$ n'est pas racine de (C) .

On cherche une solution $(E_2) : y'' - 3y' + 2y = (4x - 5)e^{-x}$ sous la forme $y_1(x) = (cx + d)e^{-x}$.

Chapitre 6

Suites numériques

Sommaire

6.1	Généralités sur les suites réelles	148
6.1.1	Suites d'un ensemble quelconque	148
6.1.2	Suites majorées, minorées, bornées	149
6.1.3	Suites réelles monotones	149
6.2	Limite d'une suite réelle	150
6.2.1	Limite finie ou infinie	150
6.2.2	Suites convergentes ou divergentes	151
6.2.3	Opérations sur les limites	151
6.2.4	Passage à la limite et inégalités	152
6.3	Limites des suites monotones	153
6.3.1	Théorème de la suite monotone	153
6.3.2	Suites adjacentes	154
6.4	Suites extraites	154
6.4.1	Notion de suite extraite	154
6.4.2	Limites et suites extraites	155
6.4.3	Théorème de Bolzano-Weierstrass	155
6.4.4	Partie dense dans \mathbb{R}	155
6.5	Extension aux suites complexes	156
6.5.1	Limite d'une suite complexe	156
6.5.2	Suites complexes bornées	157
6.6	Suites particulières	158
6.6.1	Suites arithmétiques	158
6.6.2	Suites géométriques	158
6.6.3	Suites arithmético-géométriques	159
6.6.4	Suites récurrentes linéaires d'ordre 2	160
6.6.5	Suites définies par une relation $u_{n+1} = f(u_n)$	161
6.6.6	Exemples de suites définies par $u_{n+1} = f(u_n)$	163

6.1 Généralités sur les suites réelles

6.1.1 Suites d'un ensemble quelconque

Définition 6.1.1 (suites d'éléments d'un ensemble E)

Une *suite* d'éléments d'un ensemble E est une fonction (une application) de \mathbb{N} dans E .

Il revient au même de dire que u est une famille d'éléments de E indicée par \mathbb{N} .

Plutôt que de noter $u(n)$ l'image d'un entier n , on note en général u_n .

La suite u est elle-même notée $(u_n)_{n \in \mathbb{N}}$, ou $(u_n)_{n \geq 0}$, ou simplement (u_n) .

On dit que u_n est le *terme d'indice n* (ou *terme général*) de la suite u , et que u_0 en est le *terme initial*.

Suites numériques

On parle de suite *numérique* si $E = \mathbb{R}$ ou \mathbb{C} , *réelle* si $E = \mathbb{R}$, *complexe* si $E = \mathbb{C}$.

La donnée d'une suite complexe $(z_n)_{n \geq 0}$ équivaut à celle de deux suites réelles $(u_n)_{n \geq 0}$ et $(v_n)_{n \geq 0}$ définies par : $\forall n \in \mathbb{N}, z_n = u_n + iv_n$, c'est-à-dire $u_n = \operatorname{Re}(z_n)$ et $v_n = \operatorname{Im}(z_n)$.

Conformément au programme, l'essentiel de ce chapitre est consacré aux suites réelles.

Remarques importantes

Deux suites $(u_n)_{n \geq 0}$ et $(v_n)_{n \geq 0}$ sont égales si et seulement si $u_n = v_n$ pour tout n . Il suffit donc qu'il existe au moins un entier n tel que $u_n \neq v_n$ pour que les deux suites soient considérées comme distinctes.

On ne confondra pas une suite u (c'est-à-dire une fonction définie sur \mathbb{N}) avec l'ensemble des valeurs que prend cette fonction. Par exemple, les suites de termes généraux $u_n = (-1)^n$ et $v_n = (-1)^{n+1}$ sont distinctes (mieux que ça, on a $u_n \neq v_n$ pour tout n), mais elles ont le même ensemble de valeurs $\{-1, 1\}$.

Enfin, on ne confondra **jamais** la suite $(u_n)_{n \geq 0}$ avec son terme général u_n .

Ceci est particulièrement important si on choisit la notation (u_n) pour désigner la suite u .

Suites périodiques, stationnaires

On dit qu'une suite $(u_n)_{n \geq 0}$ est *constante* s'il existe un élément a tel que : $\forall n \in \mathbb{N}, u_n = a$.

On dit qu'elle est *stationnaire* s'il existe un élément a et un entier n_0 tel que : $\forall n \geq n_0, u_n = a$.

On dit qu'elle est *p -périodique* s'il existe un entier strictement positif p tel que : $\forall n \in \mathbb{N}, u_{n+p} = u_n$.

Le plus petit p de \mathbb{N}^* ayant cette propriété est appelé la période de la suite u .

Suites définies « à partir d'un certain rang »

On est souvent amené à considérer des suites définies non sur \mathbb{N} , mais sur $\llbracket n_0, +\infty[$ (n_0 dans \mathbb{N}).

Dans ce cas, le terme initial de la suite est bien sûr u_{n_0} , et la suite elle-même est notée $(u_n)_{n \geq n_0}$.

Cela n'a pas vraiment d'importance dans la mesure où les propriétés essentielles des suites numériques, notamment en ce qui concerne les limites, sont encore vraies si les hypothèses (majoration, monotonie par exemple) sont vraies « à partir d'un certain rang ».

C'est pour cette raison que la notation (u_n) est encore acceptable pour désigner une telle suite u .

Une suite stationnaire est une suite constante à partir d'un certain rang.

Dans la suite, on considérera des suites $(u_n)_{n \geq 0}$ (une suite $(v_n)_{n \geq n_0}$ s'y ramène en posant $u_n = v_{n-n_0}$).

6.1.2 Suites majorées, minorées, bornées

Définition 6.1.2 (suites réelles majorées ou minorées)

Soit $(u_n)_{n \geq 0}$ une suite de nombres réels.

On dit que la suite u est *majorée* s'il existe M dans \mathbb{R} tel que, pour tout n de \mathbb{N} , $u_n \leq M$.

Elle est dite *minorée* s'il existe m dans \mathbb{R} tel que, pour tout n de \mathbb{N} , $u_n \geq m$.

Elle est dite *bornée* si elle est à la fois minorée et majorée.

Dire qu'une suite réelle $(u_n)_{n \in \mathbb{N}}$ est *majorée* (resp. *minorée*) équivaut à dire que l'ensemble $\{u_n, n \in \mathbb{N}\}$ des valeurs prises par cette suite est une partie majorée (resp. minorée) de \mathbb{R} .

Dire qu'une suite **réelle** est bornée, c'est dire qu'elle est à la fois minorée et majorée. Cela revient aussi à dire qu'elle est *majorée en valeur absolue*.

6.1.3 Suites réelles monotones

Définition 6.1.3 (suites réelles monotones)

Soit $(u_n)_{n \geq 0}$ une suite de nombres réels.

La suite u est dite *croissante* si, pour tout n de \mathbb{N} , on a : $u_n \leq u_{n+1}$.

Elle est dite *décroissante* si, pour tout n de \mathbb{N} , on a : $u_{n+1} \leq u_n$.

Elle est dite *monotone* si elle est croissante ou décroissante.

Définition 6.1.4 (suites réelles strictement monotones)

Soit $(u_n)_{n \geq 0}$ une suite de nombres réels.

La suite u est dite *strictement croissante* si, pour tout n de \mathbb{N} , on a : $u_n < u_{n+1}$.

Elle est dite *strictement décroissante* si, pour tout n de \mathbb{N} , on a : $u_{n+1} < u_n$.

Elle est dite *strictement monotone* si elle est strictement croissante ou strictement décroissante.

Remarques

- Quand une suite est dite monotone, sans plus de précision, il s'agit de monotonie « au sens large ».
- Soit $u = (u_n)_{n \geq 0}$ une suite de nombres réels.
 - Dire que u est *croissante* c'est dire que, pour tous m, n de \mathbb{N} , on a : $m \leq n \Rightarrow u_m \leq u_n$.
 - Dire que u est *décroissante* c'est dire que, pour tous m, n de \mathbb{N} , on a : $m \leq n \Rightarrow u_m \geq u_n$.
- La monotonie d'une suite $(u_n)_{n \geq 0}$ se mesure en comparant deux termes consécutifs u_n et u_{n+1} . C'est une différence importante avec les fonctions définies sur un intervalle de \mathbb{R} . En effet, pour qu'une telle fonction soit monotone, il ne suffit pas de comparer $f(x+1)$ avec $f(x)$. Penser par exemple à la fonction f définie sur \mathbb{R} par $f(x) = x + \cos(2\pi x)$.
- Soit $u = (u_n)_{n \geq 0}$ une suite réelle. Notons $-u$ la suite de terme général $-u_n$. Si l'une des deux suites u ou $-u$ est minorée (resp. majorée), alors l'autre est majorée (resp. minorée). Si l'une des deux est croissante (resp. décroissante), alors l'autre est décroissante (resp. croissante). Si l'une est strictement monotone, l'autre est strictement monotone mais de monotonie contraire. Ces remarques montrent que, quitte à remplacer la suite u par la suite $-u$, on peut toujours se ramener à une suite croissante, ou à une suite majorée.

Le résultat suivant sera utile quand on parlera de « suites extraites » :

Proposition 6.1.1 (suites strictement croissantes d'entiers naturels)

Soit $n \mapsto \varphi(n)$ une suite strictement croissante d'entiers naturels.

Alors, on a $\varphi(n) \geq n$ pour tout n de \mathbb{N} .

Attention : l'hypothèse selon laquelle les $\varphi(n)$ sont dans \mathbb{N} est indispensable.

6.2 Limite d'une suite réelle

6.2.1 Limite finie ou infinie

Définition 6.2.1

Soit $u = (u_n)_{n \geq 0}$ une suite de nombres réels.

- On dit que la suite u tend vers $+\infty$ si : $\forall A \in \mathbb{R}, \exists N \in \mathbb{N}, n \geq N \Rightarrow u_n \geq A$.
- On dit que la suite u tend vers $-\infty$ si : $\forall A \in \mathbb{R}, \exists N \in \mathbb{N}, n \geq N \Rightarrow u_n \leq A$.
- Soit ℓ un nombre réel.
On dit que la suite u tend vers ℓ si : $\forall \varepsilon > 0, \exists N \in \mathbb{N}, n \geq N \Rightarrow |u_n - \ell| \leq \varepsilon$.

On a ainsi donné un sens à la phrase « la suite u tend vers ℓ », avec ℓ dans $\overline{\mathbb{R}}$.

On pourrait dire « la suite u tend vers ℓ quand n tend vers $+\infty$ », mais ce n'est pas vraiment utile, s'il n'y a pas d'ambiguïté possible sur le rôle de n dans cette définition.

On pourra comparer les définitions précédentes avec celles données en classe Terminale S :

- La suite u tend vers ℓ réel si tout intervalle ouvert contenant ℓ contient tous les termes de la suite à partir d'un certain rang
- La suite u tend vers $+\infty$ si tout intervalle ouvert de la forme $]A, +\infty[$ contient tous les termes de la suite à partir d'un certain rang

Définition 6.2.2

Soit ℓ un élément de $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$.

Si la suite u tend vers ℓ on note simplement $u_n \rightarrow \ell$.

Proposition 6.2.1 (unicité de la limite si existence)

Soit $u = (u_n)_{n \geq 0}$ une suite de nombres réels, tendant vers ℓ , avec ℓ dans $\overline{\mathbb{R}}$.

Alors ℓ est le seul élément de $\overline{\mathbb{R}}$ à posséder cette propriété.

On l'appelle la limite de la suite u , et on note $\ell = \lim u_n$.

Remarque : on peut noter $\lim_{n \rightarrow \infty} u_n$, ou $\lim_{n \rightarrow +\infty} u_n$, notamment en présence d'autres variables.

Par exemple, pour lever toute ambiguïté, on écrira : $\lim_{n \rightarrow \infty} \frac{2n + 5m}{3n + 7m} = \frac{2}{3}$ et $\lim_{m \rightarrow \infty} \frac{2n + 5m}{3n + 7m} = \frac{5}{7}$.

6.2.2 Suites convergentes ou divergentes

Définition 6.2.3

Soit $u = (u_n)_{n \geq 0}$ une suite de nombres réels.

On dit que la suite u est *convergente* si elle admet une limite *finie*, c'est-à-dire une limite ℓ dans \mathbb{R} .

Dans le cas contraire, c'est-à-dire si la suite u n'a pas de limite, ou si elle tend vers $+\infty$ ou vers $-\infty$, on dit qu'elle est *divergente*.

Proposition 6.2.2 (toute suite convergente est bornée)

Si une suite numérique $(u_n)_{n \geq 0}$ est convergente, alors elle est bornée.

La réciproque est fautive. Par exemple, la suite $n \mapsto (-1)^n$ est bornée, mais n'a pas de limite.

Toute suite de réels qui tend vers $+\infty$ ou $-\infty$ est non bornée. Mais là encore la réciproque est fautive : la suite $n \mapsto (-1)^n n$ n'est pas bornée, mais ne tend ni vers $+\infty$ ni vers $-\infty$ (elle n'a pas de limite).

Il ne faut pas croire qu'une suite tendant vers $+\infty$ (resp. $-\infty$) est nécessairement croissante (resp. décroissante) au moins à partir d'un certain rang. Considérer par exemple $n \mapsto u_n = n + (-1)^n$.

6.2.3 Opérations sur les limites

Dans cette sous-section, on note $(u_n)_{n \geq 0}$ et $(v_n)_{n \geq 0}$ deux suites de nombres réels.

Proposition 6.2.3 (limite et valeur absolue)

Si $\lim u_n = \ell$, alors $\lim |u_n| = |\ell|$

Ce résultat est encore valable si $\ell = -\infty$ ou $\ell = +\infty$, à condition de noter $|\infty| = |-\infty| = +\infty$.

L'existence de $\lim |u_n|$ n'implique pas celle de $\lim u_n$, comme on le voit avec $u_n = (-1)^n$.

En revanche, on a l'équivalence : $\lim u_n = 0 \Leftrightarrow \lim |u_n| = 0$.

Si ℓ est un réel, on a les équivalences : $\lim u_n = \ell \Leftrightarrow \lim(u_n - \ell) = 0 \Leftrightarrow \lim |u_n - \ell| = 0$.

Proposition 6.2.4 (limites et combinaisons linéaires)

Soit α et β deux réels. Si $\lim u_n = \ell$ et si $\lim v_n = \ell'$, alors $\lim(\alpha u_n + \beta v_n) = \alpha \ell + \beta \ell'$

Ce résultat s'étend si ℓ ou ℓ' sont dans $\{-\infty, +\infty\}$, à condition que $\alpha \ell + \beta \ell'$ ait un sens dans $\overline{\mathbb{R}}$.

Ainsi, on ne peut rien dire en général pour $\lim(u_n + v_n)$ si $\lim u_n = +\infty$ et si $\lim v_n = -\infty$.

On dit dans ce cas qu'on est en présence de la forme indéterminée « $\infty - \infty$ » .

Il faut alors faire une étude spécifique et « lever » cette indétermination.

Proposition 6.2.5 (limites et produits)

Si $\lim u_n = \ell$ et $\lim v_n = \ell'$, alors $\lim(u_n v_n) = \ell \ell'$.

Ce résultat s'étend si ℓ ou ℓ' sont dans $\{-\infty, +\infty\}$, à condition que $\ell \ell'$ ait un sens dans $\overline{\mathbb{R}}$.

Ainsi, on ne peut rien dire en général pour $\lim(u_n v_n)$ si $\lim u_n = 0$ et si $\lim v_n \in \{-\infty, +\infty\}$.

On dit dans ce cas qu'on est en présence de la forme indéterminée « 0∞ » .

Il faut alors faire une étude spécifique et « lever » cette indétermination.

Proposition 6.2.6 (produit d'une suite qui tend vers 0 et d'une suite bornée)

Si $\lim u_n = 0$ et si la suite $(v_n)_{n \geq 0}$ est bornée, alors $\lim(u_n v_n) = 0$.

Proposition 6.2.7

Si $\lim u_n = \ell$, avec ℓ dans \mathbb{R}^* , alors il existe n_0 dans \mathbb{N} tel que : $n \geq n_0 \Rightarrow |u_n| \geq \frac{1}{2} |\ell|$.

En particulier, pour tout $n \geq n_0$, on a la majoration : $\frac{1}{|u_n|} \leq \frac{2}{|\ell|}$.

Proposition 6.2.8 (limites et inverses)

Si $\lim u_n = \ell$, avec ℓ dans \mathbb{R}^* , il existe un entier à partir duquel $u_n \neq 0$. On a alors $\lim \frac{1}{u_n} = \frac{1}{\ell}$.

Si $\lim u_n = 0$ et si les u_n sont strictement positifs, alors $\lim \frac{1}{u_n} = +\infty$.

De même, si $\lim u_n = 0$ et si les u_n sont strictement négatifs, alors $\lim \frac{1}{u_n} = -\infty$.

Si $\lim u_n = -\infty$ ou $\lim u_n = +\infty$, alors $\lim \frac{1}{u_n} = 0$.

Ce qui précède permet de conclure dans le calcul de $\lim \frac{u_n}{v_n}$, sauf dans les cas suivants :

- Si $\lim u_n = 0$ et $\lim v_n = 0$, on parle de la forme indéterminée « $\frac{0}{0}$ »
- Si $\lim u_n = \pm\infty$ et $\lim v_n = \pm\infty$, on parle de la forme indéterminée « $\frac{\infty}{\infty}$ »

Pour $\lim u_n^{v_n}$, il y a trois formes indéterminées, se ramenant à « 0∞ » car $u_n^{v_n} = e^{v_n \ln(u_n)}$

Ces trois formes indéterminées sont :
$$\begin{cases} \text{« } 1^\infty \text{ » si } \lim u_n = 1 \text{ et } \lim v_n \neq 0 \\ \text{« } \infty^0 \text{ » si } \lim u_n = +\infty \text{ et } \lim v_n = 0 \\ \text{« } 0^0 \text{ » si } \lim u_n = 0^+ \text{ et } \lim v_n = 0 \end{cases}$$

Avec une forme indéterminée, tout est possible : il faut faire une étude spécifique pour chaque cas.

6.2.4 Passage à la limite et inégalités

Proposition 6.2.9 (conservation des inégalités larges)

Soit $(u_n)_{n \geq 0}$ et $(v_n)_{n \geq 0}$ deux suites réelles, de limites respectives ℓ et ℓ' dans $\overline{\mathbb{R}}$.

S'il existe un entier n_0 tel que $u_n \leq v_n$ pour tout $n \geq n_0$, alors $\ell \leq \ell'$.

Cas particuliers

- On exprime cette propriété en disant que le passage à la limite « conserve les inégalités larges ».

En revanche, si $u_n < v_n$ pour $n \geq n_0$, alors on ne peut (là encore) affirmer que $\ell \leq \ell'$.

- Cas particuliers : Soit λ un réel (le cas le plus utile étant $\lambda = 0$), et n_0 un entier naturel.

Si on a l'inégalité $u_n \geq \lambda$ pour tout entier $n \geq n_0$, alors $\ell \geq \lambda$.

Si on a l'inégalité $u_n \leq \lambda$ pour tout entier $n \geq n_0$, alors $\ell \leq \lambda$.

Proposition 6.2.10

Soit $(u_n)_{n \geq 0}$ une suite réelle, de limite ℓ dans $\overline{\mathbb{R}}$.

Si $\ell < \lambda$, alors il existe un entier n_0 tel que : $\forall n \geq n_0, u_n < \lambda$.

Si $\ell > \lambda$, alors il existe un entier n_0 tel que : $\forall n \geq n_0, u_n > \lambda$.

Le résultat précédent est souvent utilisé avec $\lambda = 0$: par exemple, si $\lim u_n = \ell > 0$, alors il existe un rang à partir duquel les u_n sont tous strictement positifs.

Convergence par encadrement**Proposition 6.2.11** (« principe des gendarmes »)

Soit $(u_n)_{n \geq 0}$, $(v_n)_{n \geq 0}$, $(w_n)_{n \geq 0}$ trois suites réelles.

On suppose que $\lim u_n = \lim v_n = \ell$, où ℓ est dans \mathbb{R} .

S'il existe un entier n_0 tel que : $u_n \leq w_n \leq v_n$ pour tout $n \geq n_0$, alors $\lim w_n = \ell$.

Cas particulier : si $\lim u_n = 0$, et s'il existe n_0 tel que $|v_n| \leq |u_n|$ pour $n \geq n_0$, alors $\lim v_n = 0$.

Divergence par minoration ou majoration**Proposition 6.2.12** (autres propriétés liées à la relation d'ordre)

Si $\lim u_n = +\infty$ s'il existe un entier n_0 tel que $v_n \geq u_n$ pour tout $n \geq n_0$, alors $\lim v_n = +\infty$.

Si $\lim u_n = -\infty$ s'il existe un entier n_0 tel que $v_n \leq u_n$ pour tout $n \geq n_0$, alors $\lim v_n = -\infty$.

Proposition 6.2.13 (convergence ou divergence par comparaison de quotients)

Soit u et v deux suites à valeurs dans \mathbb{R}^{+*} , et telles que : $\frac{u_{n+1}}{u_n} \leq \frac{v_{n+1}}{v_n}$ pour $n \geq n_0$.

Dans ces conditions, si $\lim v_n = 0$ alors $\lim u_n = 0$.

De même, si $\lim u_n = +\infty$, alors $\lim v_n = +\infty$.

Quelques limites utiles

Soit a un réel strictement supérieur à 1, et soit k un réel strictement positif.

Alors on a les limites : $\lim_{n \rightarrow \infty} \frac{a^n}{n^k} = +\infty$ $\lim_{n \rightarrow \infty} \frac{n!}{a^n} = +\infty$ $\lim_{n \rightarrow \infty} \frac{n^n}{n!} = +\infty$.

6.3 Limites des suites monotones**6.3.1 Théorème de la suite monotone****Proposition 6.3.1** (limite d'une suite réelle croissante)

Soit $(u_n)_{n \geq 0}$ une suite croissante de nombres réels.

Si cette suite est majorée, alors elle est convergente. Plus précisément, $\lim u_n = \sup\{u_n, n \geq 0\}$.

Si au contraire cette suite n'est pas majorée, alors $\lim u_n = +\infty$.

On retiendra le « *théorème de la suite monotone* » : toute suite réelle monotone possède une limite.

En considérant la suite de terme général $(-u_n)_{n \geq 0}$, on en déduit le résultat suivant :

Proposition 6.3.2 (limite d'une suite réelle décroissante)

Soit $(u_n)_{n \geq 0}$ une suite décroissante de nombres réels.

Si cette suite est minorée, alors elle est convergente. Plus précisément, $\lim u_n = \inf\{u_n, n \geq 0\}$.

Si au contraire cette suite n'est pas minorée, alors $\lim u_n = -\infty$.

Le résultat suivant constitue une sorte de réciproque des deux propositions précédentes :

Proposition 6.3.3

Soit X une partie non vide majorée de \mathbb{R} .

Alors il existe une suite u d'éléments de X telle que $\lim u_n = \sup X$.

On peut même faire en sorte que la suite u soit croissante.

De même, si X est non vide minorée, il existe une suite décroissante u de X telle que $\lim u_n = \inf X$.

Penser à étudier la monotonie

L'étude d'une suite réelle passe très souvent par celle de sa monotonie.

C'est donc un réflexe utile que de vérifier si la suite étudiée est croissante ou décroissante.

En général, on étudiera pour cela le signe de la différence $u_{n+1} - u_n$. Mais si le terme général u_n s'exprime sous forme de produits, de puissances ou de factorielles, il pourra être plus simple de comparer le rapport u_{n+1}/u_n avec la valeur 1 (mais attention cependant au signe de u_n avant de conclure!).

6.3.2 Suites adjacentes

Définition 6.3.1 (suites adjacentes)

On dit que deux suites réelles $(u_n)_{n \geq 0}$ et $(v_n)_{n \geq 0}$ sont *adjacentes* si l'une d'elles est croissante, si l'autre est décroissante, et si $\lim(v_n - u_n) = 0$.

Proposition 6.3.4 (théorème des suites adjacentes)

Soient $(u_n)_{n \geq 0}$ et $(v_n)_{n \geq 0}$ deux suites réelles adjacentes.

Alors ces deux suites sont convergentes et elles ont la même limite.

6.4 Suites extraites

6.4.1 Notion de suite extraite

Définition 6.4.1

Soit $(u_n)_{n \geq 0}$ une suite d'un ensemble E quelconque. On appelle *suite extraite* de u toute suite v de terme général $v_n = u_{\varphi(n)}$, où φ est strictement croissante de \mathbb{N} dans lui-même.

Rappel : avec les notations de l'énoncé, et pour tout entier n , on a l'inégalité $\varphi(n) \geq n$.

On considère souvent $\begin{cases} \text{la suite } (u_{2n})_{n \geq 0} \text{ des termes d'indices pairs : } \varphi(n) = 2n, \\ \text{la suite } (u_{2n+1})_{n \geq 0} \text{ des termes d'indices impairs : } \varphi(n) = 2n + 1. \end{cases}$

On utilise aussi l'expression « sous-suite de u » pour désigner une suite extraite de u .

Sous-suite d'une sous-suite

Soit $u = (u_n)_{n \geq 0}$ une suite d'un ensemble E quelconque.

Soit $v = (v_n)_{n \geq 0}$ une suite extraite de la suite u . Soit $w = (w_n)_{n \geq 0}$ une suite extraite de la suite v .

Alors w est une suite extraite de la suite u .

Ce résultat peut sembler évident, mais il y a quand même une subtilité.

Plus précisément, si φ et ψ sont deux applications strictement croissantes de \mathbb{N} dans lui-même telles que $v_m = u_{\varphi(m)}$ et $w_n = v_{\psi(n)}$ pour tous entiers naturels m, n , alors (en posant $m = \psi(n)$) :

$$\forall n \in \mathbb{N} : w_n = v_m = u_{\varphi(m)} = u_{\varphi(\psi(n))} = u_{\theta(n)}, \text{ avec } \theta = \varphi \circ \psi \text{ (et non pas } \theta = \psi \circ \varphi \text{ !!)}$$

6.4.2 Limites et suites extraites

Proposition 6.4.1 (limite des suites extraites)

Si la suite $u = (u_n)_{n \geq 0}$ a pour limite ℓ , alors toute suite extraite de u admet encore ℓ pour limite.

Le fait qu'une suite extraite de u possède une limite ℓ ne signifie pas que la suite u tende vers ℓ , ni même possède une limite (mais tout de même, ℓ devient la « seule limite possible » de u).

Si deux suites extraites de u ont des limites différentes, on est cependant certain que u n'a pas de limite. Cette remarque est souvent utilisée pour montrer qu'une suite u est divergente.

Proposition 6.4.2

Soit $(u_n)_{n \geq 0}$ une suite de nombres réels.

On suppose que les suites (u_{2n}) et (u_{2n+1}) ont une même limite ℓ . Alors $\lim u_n = \ell$.

6.4.3 Théorème de Bolzano-Weierstrass

Ce résultat possède des conséquences théoriques importantes :

Proposition 6.4.3 (théorème de Bolzano-Weierstrass)

De toute suite bornée de \mathbb{R} , on peut extraire une suite convergente.

Le théorème ne dit pas comment extraire une telle suite convergente, il dit simplement qu'elle existe.

6.4.4 Partie dense dans \mathbb{R}

Définition 6.4.2 (définition de la densité)

Soit A une partie de \mathbb{R} .

On dit que A est dense si, pour tout intervalle ouvert non vide I , l'intersection $A \cap I$ est non vide.

Remarque : il est clair que si A est dense dans \mathbb{R} , et si $A \subset B$, alors B est dense dans \mathbb{R} .

Proposition 6.4.4

Soit A une partie dense de \mathbb{R} , et soit I intervalle ouvert non vide.

Alors l'intersection $A \cap I$ est un ensemble infini.

Proposition 6.4.5 (caractérisation séquentielle de la densité)

Soit A une partie de \mathbb{R} .

Alors A est dense si et seulement si, pour tout réel α , il existe une suite $(u_n)_{n \geq 0}$ de A telle que $\lim u_n = \alpha$.

Proposition 6.4.6 (exemples de parties denses de \mathbb{R})

L'ensemble des nombres décimaux est une partie dense de \mathbb{R} .

L'ensemble des rationnels, et l'ensemble des irrationnels, sont des parties denses de \mathbb{R} .

6.5 Extension aux suites complexes

6.5.1 Limite d'une suite complexe

Définition 6.5.1 (définition d'une suite complexe)

Une suite *complexe* $z = (z_n)_{n \geq 0}$ est une fonction de \mathbb{N} dans \mathbb{C} .

Elle est définie de façon unique par deux suites réelles x, y en notant : $\forall n \in \mathbb{N}, z_n = x_n + iy_n$.

Il revient au même d'écrire : $\forall n \in \mathbb{N}, x_n = \operatorname{Re}(z_n)$ et $y_n = \operatorname{Im}(z_n)$.

On dit que $(x_n)_{n \geq 0}$ est la *partie réelle* de la suite $(z_n)_{n \geq 0}$, et $(y_n)_{n \geq 0}$ sa *partie imaginaire*.

Définition 6.5.2 (limite d'une suite complexe)

Soit $(z_n)_{n \geq 0}$ une suite de nombres complexes. Soit ℓ un nombre complexe.

On dit que la suite z tend vers ℓ si : $\forall \varepsilon > 0, \exists N \in \mathbb{N}, n \geq N \Rightarrow |z_n - \ell| \leq \varepsilon$.

On exprime aussi cette situation en disant que la suite $(z_n)_{n \geq 0}$ est convergente vers ℓ .

Remarques

- La définition précédente est la même que pour les suites réelles convergentes, mais il s'agit ici du module (qui généralise la notion de valeur absolue dans \mathbb{R}).
- La définition précédente peut aussi s'énoncer : *la suite z tend vers ℓ si tout disque fermé centré en ℓ (et de rayon strictement positif) contient tous les termes de la suite à partir d'un certain rang.*
- On a encore la propriété d'*unicité de la limite* (si existence), ce qui permet d'écrire $\ell = \lim z_n$.
- Si $(z_n)_{n \geq 0}$ une suite complexe, ça n'a aucun sens de dire qu'elle tend vers $+\infty$ ou $-\infty$.
Si $(z_n)_{n \geq 0}$ n'admet pas de limite (c'est-à-dire si elle n'est pas convergente), elle est dite divergente.

Proposition 6.5.1 (caractérisation en termes de partie réelle et partie imaginaire)

Soit $(z_n)_{n \geq 0}$ une suite complexe. Soit $\ell = a + ib$ un nombre complexe, avec a, b dans \mathbb{R} .

Pour tout n de \mathbb{N} , posons $z_n = x_n + iy_n$, avec x_n et y_n dans \mathbb{R} .

Alors on a l'équivalence : $\lim z_n = \ell \Leftrightarrow \begin{cases} \lim x_n = a \\ \lim y_n = b \end{cases}$

Dans ce cas, on a donc : $\lim z_n = (\lim x_n) + i(\lim y_n)$.

Voici une extension aux suites complexes des propriétés sur les limites et les opérations algébriques :

Proposition 6.5.2 (limites et combinaisons linéaires)

On suppose que $\lim z_n = \ell$ et $\lim z'_n = \ell'$, avec ℓ et ℓ' dans \mathbb{C} .

Pour tous α et β dans \mathbb{C} , on a : $\lim(\alpha z_n + \beta z'_n) = \alpha\ell + \beta\ell'$

Pour le produit, on a : $\lim(z_n z'_n) = \ell\ell'$. Si de plus $\ell' \neq 0$, alors on a : $\lim \frac{z_n}{z'_n} = \frac{\ell}{\ell'}$.

6.5.2 Suites complexes bornées

Définition 6.5.3 (suites complexes bornées)

Soit $(z_n)_{n \geq 0}$ une suite de nombres complexes.

On dit que la suite z est *bornée* s'il existe M dans \mathbb{R}^+ tel que, pour tout n de \mathbb{N} , $|z_n| \leq M$.

Proposition 6.5.3

Soit $(z_n)_{n \geq 0}$ une suite complexe. Posons $z_n = x_n + iy_n$, avec x_n et y_n dans \mathbb{R} .

Alors la suite $(z_n)_{n \geq 0}$ est bornée si et seulement si les suites $(x_n)_{n \geq 0}$ et $(y_n)_{n \geq 0}$ sont bornées.

Proposition 6.5.4 (convergence et module)

Si une suite complexe $(z_n)_{n \geq 0}$ est convergente, elle est bornée (réciproque fausse).

Si $\lim z_n = \ell$, alors $\lim |z_n| = |\ell|$.

On a l'équivalence : $\lim z_n = \ell \Leftrightarrow \lim(z_n - \ell) = 0 \Leftrightarrow \lim |z_n - \ell| = 0$.

Remarques

- Pour une suite à valeurs dans \mathbb{C} , ça n'a aucun sens de dire si elle est majorée ou minorée. De même, on ne parlera **jamais** de la monotonie d'une suite complexe (ça n'existe pas).
- Dire qu'une suite complexe est bornée, c'est dire qu'elle est *majorée en module*.
- Si $\lim z_n = 0$ et si la suite $(z'_n)_{n \geq 0}$ est bornée, alors $\lim(z_n z'_n) = 0$.
- Soit $(z_n)_{n \geq 0}$ une suite complexe. On suppose $\lim z_n = \ell$, avec ℓ dans \mathbb{C}^* . Alors il existe n_0 dans \mathbb{N} tel que : $n \geq n_0 \Rightarrow |z_n| \geq \frac{1}{2} |\ell|$. En particulier, $\frac{1}{|z_n|} \leq \frac{2}{|\ell|}$ pour $n \geq n_0$.

Proposition 6.5.5 (théorème de Bolzano-Weierstrass pour les suites complexes)

De toute suite bornée de \mathbb{C} , on peut extraire une suite convergente.

6.6 Suites particulières

6.6.1 Suites arithmétiques

On note toujours $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} .

Définition 6.6.1 (définition des suites arithmétiques)

Une suite $(u_n)_{n \geq 0}$ est dite *arithmétique* s'il existe un scalaire r tel que : $\forall n \in \mathbb{N}, u_{n+1} = u_n + r$.

Le scalaire r est appelé *raison* de la suite arithmétique. Il est défini de façon unique.

Pour tout n de \mathbb{N} , on a : $u_n = u_0 + nr$.

Quelques propriétés des suites arithmétiques

- Soit $(u_n)_{n \geq 0}$ une suite arithmétique de raison r .
Si $\mathbb{K} = \mathbb{R}$ et si $r > 0$ (resp. $r < 0$), elle est strictement croissante (resp. strictement décroissante).
Pour tous n, p de \mathbb{N} , on a : $u_n = u_p + (n-p)r$.
- Une suite arithmétique u n'est convergente que si sa raison r est nulle (et alors u est constante).
- Réciproquement, on suppose que le terme général d'une suite $(u_n)_{n \geq 0}$ s'écrit $u_n = a + nb$.
Alors la suite $(u_n)_{n \geq 0}$ est arithmétique de premier terme $u_0 = a$ et de raison b .
- Soit S la somme de n termes consécutifs d'une suite arithmétique.
Soit d le terme débutant, et f le terme finissant cette progression : alors $S = n \frac{d+f}{2}$.

Proposition 6.6.1 (caractérisation des suites arithmétiques)

Une suite $(u_n)_{n \geq 0}$ est arithmétique si et seulement si : $\forall n \in \mathbb{N}, u_n + u_{n+2} = 2u_{n+1}$.

On dit que a, b, c sont en *progression arithmétique* s'ils sont consécutifs dans une suite arithmétique. Cela équivaut à dire que $a + c = 2b$.

6.6.2 Suites géométriques

Définition 6.6.2 (définition des suites géométriques)

Une suite $(u_n)_{n \geq 0}$ est dite *géométrique* s'il existe un scalaire q tel que : $\forall n \in \mathbb{N}, u_{n+1} = q u_n$.

Le scalaire q est appelé *raison* de la suite géométrique.

Pour tout n de \mathbb{N} , on a : $u_n = q^n u_0$.

Remarques générales

- Soit $(u_n)_{n \geq 0}$ une suite *géométrique* de raison q .
Si $u_0 = 0$, la suite $(u_n)_{n \geq 0}$ est identiquement nulle, et q est quelconque.
Si $u_0 \neq 0$, le scalaire q est défini de façon unique par $q = u_1/u_0$.
Si $q = 1$, la suite u est constante. Si $q = 0$ elle est stationnaire en 0 (à partir de $n = 1$).
Enfin si $u_0 \neq 0$ et $q \neq 0$, aucun terme de la suite u n'est nul.

- Soit $(u_n)_{n \geq 0}$ une suite *géométrique* de raison q .
Pour tout n de \mathbb{N} , on a : $u_n = q^n u_0$, et pour $p \leq n$ dans \mathbb{N} , on a : $u_n = u_p q^{n-p}$.
- Réciproquement, si le terme général d'une suite $(u_n)_{n \geq 0}$ s'écrit $u_n = aq^n$, alors $(u_n)_{n \geq 0}$ est la suite géométrique de premier terme $u_0 = a$ et de raison q .

Monotonie des suites géométriques réelles

On suppose ici $\mathbb{K} = \mathbb{R}$. Soit $(u_n)_{n \geq 0}$ une suite géométrique de raison q .

- Si $q > 0$, la suite u garde un signe constant et est monotone. Plus précisément :
 - . si $u_0 > 0$ et $q > 1$, la suite u est positive strictement croissante.
 - . si $u_0 > 0$ et $0 < q < 1$, la suite u est positive strictement décroissante.
 - . si $u_0 < 0$ et $q > 1$, la suite u est négative strictement décroissante.
 - . si $u_0 < 0$ et $0 < q < 1$, la suite u est négative strictement croissante.
- Si $q < 0$, alors deux termes consécutifs u_n et u_{n+1} sont toujours de signes contraires.
La suite u n'est donc pas monotone.

Proposition 6.6.2 (caractérisation des suites géométriques)

Une suite $(u_n)_{n \geq 0}$ est géométrique si et seulement si, pour tout entier n : $u_n u_{n+2} = u_{n+1}^2$.

On dit que trois scalaires a, b, c sont en *progression géométrique* s'ils sont des termes successifs d'une suite géométrique : cela équivaut à dire que $ac = b^2$.

Soit S la somme de n termes consécutifs d'une progression géométrique de raison $q \neq 1$.

Soit u_0 le premier terme de cette progression : alors $S = u_0 \frac{q^n - 1}{q - 1}$.

Proposition 6.6.3 (convergence des suites géométriques)

Soit $(u_n)_{n \geq 0}$ une suite géométrique de raison q , et de premier terme $u_0 \neq 0$.

La suite $(u_n)_{n \geq 0}$ est convergente si et seulement si $q = 1$ ou $|q| < 1$.

Si $q = 1$, la suite u est constante. Si $|q| < 1$, la suite u converge vers 0.

6.6.3 Suites arithmético-géométriques

Définition 6.6.3 (suites arithmético-géométriques)

Une suite $(u_n)_{n \geq 0}$ est *arithmético-géométrique* si : $\exists (a, b) \in \mathbb{K}^2, \forall n \in \mathbb{N}, u_{n+1} = au_n + b$

Expression du terme général

On se donne une suite arithmético-géométrique : $\forall n \in \mathbb{N}, u_{n+1} = au_n + b$.

Si $b = 0$ (resp. $a = 1$) c'est une suite géométrique (resp. arithmétique).

Supposons $a \neq 1$: soit λ l'unique scalaire vérifiant $\lambda = a\lambda + b$. Donc $\lambda = \frac{b}{1-a}$.

Si on soustrait les égalités $\begin{cases} u_{n+1} = au_n + b \\ \lambda = a\lambda + b \end{cases}$ on obtient $u_{n+1} - \lambda = a(u_n - \lambda)$.

La suite $n \mapsto u_n - \lambda$ est donc géométrique de raison a : $\forall n \in \mathbb{N}, u_{n+1} - \lambda = a(u_n - \lambda)$.

On en déduit l'expression de u_n : $\forall n \in \mathbb{N}, u_n - \lambda = (u_0 - \lambda)a^n$, donc $u_n = (u_0 - \lambda)a^n + \lambda$.

Convergence des suites arithmético-géométriques

On se donne une suite arithmético-géométrique : $\forall n \in \mathbb{N}, u_{n+1} = au_n + b$.

On suppose $a \neq 1$ (donc la suite u n'est pas arithmétique).

On pose $\lambda = \frac{b}{1-a}$. Si $u_0 = \lambda$, alors la suite u est constante en λ .

En revanche, si $u_0 \neq \lambda$, la suite u converge si et seulement si $|a| < 1$. Dans ce cas $\lim u_n = \lambda$.

Monotonie des suites arithmético-géométriques réelles

On se donne une suite arithmético-géométrique : $\forall n \in \mathbb{N}, u_{n+1} = au_n + b$ (avec a et b dans \mathbb{R}).

On suppose $a \neq 1$ (la suite u n'est pas arithmétique) et $a \neq 0$ (sinon la suite u est constante en b).

On pose $\lambda = \frac{b}{1-a}$, et on suppose $u_0 \neq \lambda$ (sinon la suite u est constante en λ).

On a $u_{n+1} - u_n = a(u_n - u_{n-1})$ donc $u_{n+1} - u_n = a^n(u_1 - u_0) = a^n((a-1)u_0 + b)$.

Ainsi u est monotone $\Leftrightarrow a > 0$ (strictement croissante si $u_1 > u_0$, strictement décroissante sinon).

6.6.4 Suites récurrentes linéaires d'ordre 2

Dans tout ce qui suit, $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} .

Définition 6.6.4 (récurrences linéaires d'ordre 2)

On dit qu'une $(u_n)_{n \geq 0}$ de \mathbb{K} satisfait à une *récurrence linéaire d'ordre 2* s'il existe a, b, c dans \mathbb{K} , avec $a \neq 0$ et $c \neq 0$, tels que : $\forall n \in \mathbb{N}, au_{n+2} + bu_{n+1} + cu_n = 0$ (E).

L'équation (C) : $at^2 + bt + c = 0$ (d'inconnue t dans \mathbb{K}) est dite *équation caractéristique* de (E).

Remarques et exemple

- On a supposé $a \neq 0$ et $b \neq 0$ pour ne pas retomber dans le cas des suites géométriques.
La relation (E) fait donc réellement apparaître u_{n+2} et u_n .
- On sait que les suites arithmétiques sont caractérisées par la relation : $\forall n \in \mathbb{N}, u_n + u_{n+2} = 2u_{n+1}$.
Cette relation s'écrit : $\forall n \in \mathbb{N}, u_{n+2} - 2u_{n+1} + u_n = 0$.
Dans ce cas, l'équation caractéristique est $t^2 - 2t + 1 = 0$, de racine double $t = 1$.
- La « suite de Fibonacci » $(F_n)_{n \geq 0}$ est définie par $F_0 = 0, F_1 = 1$ et : $\forall n \in \mathbb{N}, F_{n+2} = F_{n+1} + F_n$.
Dans ce cas, l'équation caractéristique est (C) : $t^2 - t - 1 = 0$.
Elle possède les deux racines distinctes : $\Phi = \frac{1 + \sqrt{5}}{2}$ (le « nombre d'or ») et $\Phi' = \frac{1 - \sqrt{5}}{2} = -\frac{1}{\Phi}$.

Proposition 6.6.4 (suites géométriques solutions)

Soit \mathcal{S}_E l'ensemble des suites de \mathbb{K} qui vérifient la relation (E) : $\forall n \in \mathbb{N}, au_{n+2} + bu_{n+1} + cu_n = 0$.

- Un élément u de \mathcal{S}_E est défini de façon unique par la donnée de ses deux premiers termes.
L'ensemble \mathcal{S}_E est donc une famille de suites à deux paramètres.
- Soit q un élément de \mathbb{K}^* . La suite géométrique $n \mapsto q^n$ est élément de \mathcal{S}_E si et seulement si q est racine de l'équation caractéristique (C) : $at^2 + bt + c = 0$

Proposition 6.6.5 (solution générale dans le cas complexe)

Soit a, b, c trois nombres complexes, avec $a \neq 0$ et $b \neq 0$.

Soit \mathcal{S}_E l'ensemble des suites de \mathbb{C} qui vérifient la relation (E) : $\forall n \in \mathbb{N}, au_{n+2} + bu_{n+1} + cu_n = 0$.

Soit $\Delta = b^2 - 4ac$ le discriminant de l'équation caractéristique (C) : $at^2 + bt + c = 0$

– Si $\Delta \neq 0$, soit r et s les deux racines distinctes de l'équation caractéristique (C).

Alors \mathcal{S}_E est l'ensemble des suites $n \mapsto \lambda r^n + \mu s^n$, avec λ, μ dans \mathbb{C} .

– Si $\Delta = 0$, soit r la racine double de l'équation caractéristique (C).

Alors \mathcal{S}_E est l'ensemble des suites $n \mapsto (\lambda n + \mu)r^n$, avec λ, μ dans \mathbb{C} .

Proposition 6.6.6 (solution générale dans le cas réel)

Soit a, b, c trois nombres réels, avec $a \neq 0$ et $b \neq 0$.

Soit \mathcal{S}_E l'ensemble des suites de \mathbb{R} qui vérifient la relation (E) : $\forall n \in \mathbb{N}, au_{n+2} + bu_{n+1} + cu_n = 0$.

Soit $\Delta = b^2 - 4ac$ le discriminant de l'équation caractéristique (C) : $at^2 + bt + c = 0$

– Si $\Delta > 0$, soit r et s les deux racines réelles distinctes de l'équation caractéristique (C).

Alors \mathcal{S}_E est l'ensemble des suites $n \mapsto \lambda r^n + \mu s^n$, avec λ, μ dans \mathbb{R} .

– Si $\Delta = 0$, soit r la racine réelle double de l'équation caractéristique (C).

Alors \mathcal{S}_E est l'ensemble des suites $n \mapsto (\lambda n + \mu)r^n$, avec λ, μ dans \mathbb{R} .

– Si $\Delta < 0$, soit $r = \rho e^{i\theta}$ et $\bar{r} = \rho e^{-i\theta}$ les racines conjuguées distinctes de (C), avec $\theta \neq 0 \pmod{\pi}$.

Alors \mathcal{S}_E est l'ensemble des suites $n \mapsto (\lambda \cos(n\theta) + \mu \sin(n\theta))\rho^n$, avec λ, μ dans \mathbb{R} .

6.6.5 Suites définies par une relation $u_{n+1} = f(u_n)$ **Définition 6.6.5** (suite définie par récurrence)

Soit f une fonction définie sur une partie \mathcal{D} de \mathbb{K} , à valeurs dans \mathbb{K} .

Soit a un élément de \mathcal{D} . On peut définir une suite $(u_n)_{n \geq 0}$ de \mathbb{K} par :

– la donnée de son terme initial $u_0 = a$ dans \mathcal{D} .

– la relation de récurrence : $\forall n \in \mathbb{N}, u_{n+1} = f(u_n)$.

On dit alors que la suite u est définie par récurrence.

Remarque sur le domaine de définition

Avec les notations précédentes, il faut s'assurer de l'existence de la suite u .

On doit donc vérifier que les u_n sont dans \mathcal{D} pour tout n de \mathbb{N} .

C'est évidemment très simple si on sait que $f(\mathcal{D}) \subset \mathcal{D}$.

Prenons l'exemple de la suite réelle $(u_n)_{n \geq 0}$ par : u_0 dans \mathbb{R} et $\forall n \in \mathbb{N}, u_{n+1} = \sqrt{1 - u_n}$.

Pour que cette suite ait un sens, il faut en particulier que u_1 existe, c'est-à-dire $u_0 \leq 1$.

Mais pour que u_2 existe il faut $u_1 = \sqrt{1 - u_0} \leq 1$, c'est-à-dire $u_0 \geq 0$.

Enfin, la condition $0 \leq u_0 \leq 1$ est suffisante car $[0, 1]$ est stable par $f : x \mapsto \sqrt{1 - x}$.

Proposition 6.6.7 (limites « éventuelles » d'une suite définie par $u_{n+1} = f(u_n)$)

Soit $f : \mathcal{D} \rightarrow \mathbb{K}$ une fonction continue. On suppose que $f(\mathcal{D}) \subset \mathcal{D}$.

Soit $(u_n)_{n \geq 0}$ la suite définie par le choix de u_0 dans \mathcal{D} et par la relation de récurrence $u_{n+1} = f(u_n)$.

Si la suite u converge vers un élément ℓ de \mathcal{D} , alors la limite vérifie $f(\ell) = \ell$.

Résoudre l'équation $f(x) = x$ donne donc les limites éventuelles de la suite u dans l'ensemble \mathcal{D} .

Limites éventuelles et intervalles stables

Pour une suite définie par une récurrence $u_{n+1} = f(u_n)$, et si f est continue, on trouvera les limites éventuelles en cherchant les « points fixes » de f , c'est-à-dire en résolvant l'équation $f(x) = x$.

Il est recommandé d'étudier le signe de $f(x) - x$, et d'identifier des intervalles stables par f (souvent un intervalle séparant deux points fixes successifs de f).

Voici par exemple une situation typique :

- Supposons que α et β soient les seules solutions de $f(x) = x$.
- Supposons en outre que $\alpha < x < \beta \Rightarrow \alpha < f(x) < x < \beta$.
- Si $\alpha < u_0 < \beta$, alors par une récurrence évidente : $\forall n \in \mathbb{N}, \alpha < u_{n+1} < u_n < \beta$
- On conclut que la suite u , décroissante minorée, converge vers α (seule limite possible ici).

Utilisation de représentations graphiques

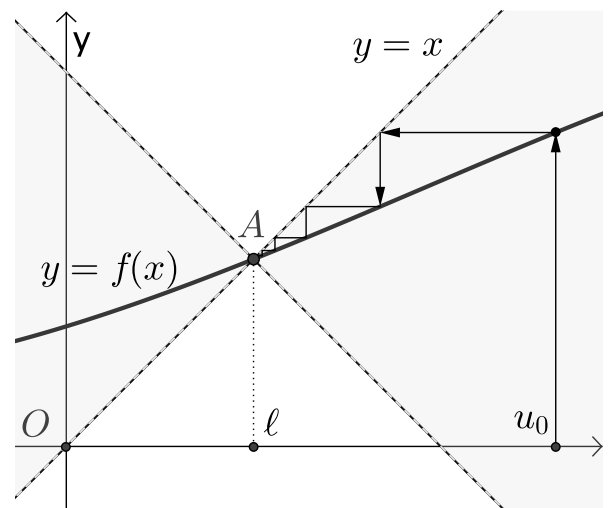
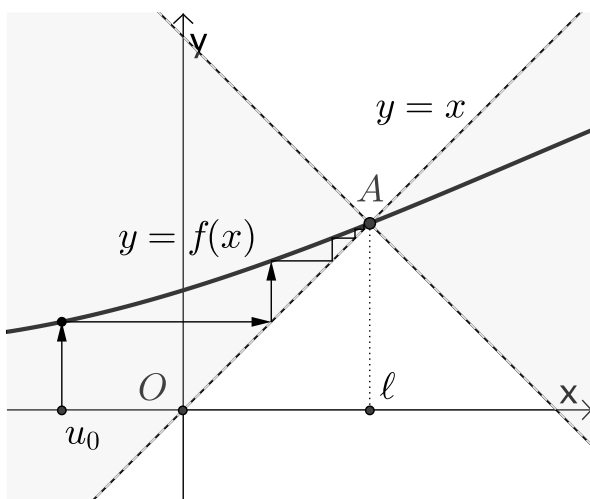
Pour illustrer le comportement d'une suite $(u_n)_{n \geq 0}$ définie par une relation $u_{n+1} = f(u_n)$, on utilise souvent des représentations graphiques « en escaliers » ou « en spirale » (le terme utilisé dépend de la monotonie de la suite au voisinage de sa limite).

On voit ci-dessous l'illustration de la convergence d'une telle suite $(u_n)_{n \geq 0}$ vers sa limite ℓ .

On a tracé la courbe $y = f(x)$, la droite $y = x$, ainsi que la perpendiculaire à cette droite qui passe par le « point-limite ». La partie grisée est la « zone de convergence » : si la courbe $y = f(x)$ reste dans cette zone au voisinage de $A(\ell, \ell)$, la suite (u_n) converge vers ℓ .

La convergence est d'autant plus rapide que la tangente à la courbe en A est proche de l'horizontale.

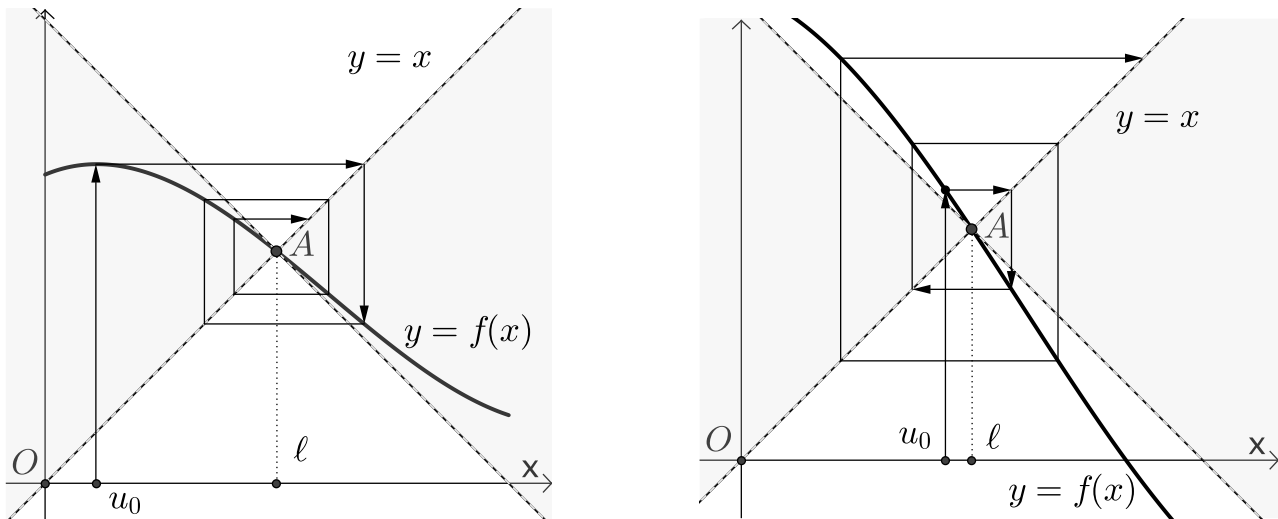
On voit sur ces deux exemples que le fait que f soit croissante implique seulement que (u_n) est monotone : elle est croissante si $u_1 > u_0$ (figure de gauche), et elle est décroissante si $u_1 < u_0$ (figure de droite).



Dans les deux illustrations ci-dessous, on considère le comportement de la suite $(u_n)_{n \geq 0}$ au voisinage d'un point fixe ℓ de f . Dans les deux cas, f est décroissante au voisinage de ℓ , et il en résulte que la suite $(u_n)_{n \geq 0}$ n'est pas monotone (en revanche les suites (u_{2n}) et (u_{2n+1}) sont monotones).

Dans l'exemple de gauche, on a $|f'(\ell)| < 1$, et la courbe $y = f(x)$ reste dans la zone de convergence au voisinage de A (on exprime cette situation en disant que ℓ est un point fixe « attractif » de f). Il en résulte que la suite $(u_n)_{n \geq 0}$ converge effectivement vers ℓ , et de façon alternée (et assez lentement sur notre exemple car la dérivée de f en a est proche de la valeur -1).

Dans l'exemple de droite, on a $|f'(\ell)| > 1$, et la courbe $y = f(x)$ sort de la zone de convergence au voisinage de A (on dit que ℓ est un point fixe « répulsif » de f). Il en résulte que la suite $(u_n)_{n \geq 0}$ ne converge pas vers ℓ (même si u_0 est proche de ℓ , les u_n s'en éloignent progressivement).



6.6.6 Exemples de suites définies par $u_{n+1} = f(u_n)$

▷ Exemple n°1

Étudier la suite $(u_n)_{n \geq 0}$ définie par $u_0 > 0$ et par : $\forall n \in \mathbb{N}, u_{n+1} = \frac{u_n}{1 + u_n}$

Une récurrence immédiate montre que la suite $(u_n)_{n \geq 0}$ est « bien définie » et que $u_n > 0$ pour tout n .

Si on pose $v_n = \frac{1}{u_n}$, on observe que : $\forall n \in \mathbb{N}, v_{n+1} = \frac{1}{u_{n+1}} = \frac{1 + u_n}{u_n} = \frac{1}{u_n} + 1 = v_n + 1$.

Autrement dit, la suite $(v_n)_{n \geq 0}$ est arithmétique de raison 1.

Pour tout n de \mathbb{N} , on en déduit : $v_n = v_0 + n = \frac{1}{u_0} + n = \frac{nu_0 + 1}{u_0}$ donc $u_n = \frac{1}{v_n} = \frac{u_0}{nu_0 + 1}$.

Conclusion : la suite $(u_n)_{n \geq 0}$ est décroissante de limite 0.

Ce premier exemple est quand même très particulier. Il est en effet très rare qu'on puisse trouver comme on l'a fait ici l'expression générale de u_n .

▷ Exemple n°2

Étudier la suite (u_n) définie par u_0 réel et $u_{n+1} = u_n(1 + u_n)$.

Pour tout n de \mathbb{N} , on a $u_{n+1} - u_n = u_n^2 \geq 0$. La suite u est donc croissante.

La seule limite finie possible, donnée par $\ell = \ell(1 + \ell)$, est $\ell = 0$.

– Si $u_0 > 0$, il est clair que $u_n > 0$ pour tout n .

On en déduit $\lim u_n = +\infty$ (en effet la suite est croissante et ne peut donc pas converger vers 0).

– Si $u_0 < -1$, on a $u_1 = u_0(1 + u_0) > 0$ et on est ramené au cas précédent : $\lim u_n = +\infty$.

– Si $u_0 = 0$, la suite u est constante : $\forall n \geq 0, u_n = 0$.

– Si $u_0 = -1$, on a $u_1 = 0$ ce qui ramène au cas précédent : $\forall n \geq 1, u_n = 0$.

– Si $-1 < u_0 < 0$:

Pour tout x de $] -1, 0[$, on a $-1 < x < f(x) < 0$.

L'intervalle $] -1, 0[$ est donc stable par f . On en déduit : $\forall n \geq 0, -1 < u_n < 0$.

La suite u , croissante et majorée, est convergente. Nécessairement $\lim u_n = 0$.

Conclusion : si $-1 \leq u_0 \leq 0$ alors $\lim u_n = 0$, sinon $\lim u_n = +\infty$.

▷ Exemple n°3

Étudier la suite (u_n) définie par $u_0 \neq 1$ et la relation $u_{n+1} = \frac{u_n^2 + 1}{u_n - 1}$.

La suite u est définie par $u_{n+1} = f(u_n)$, avec $f(x) = \frac{x^2 + 1}{x - 1}$.

On commence par étudier la fonction f , définie pour $x \neq 1$.

On constate que pour tout $x \neq 1$, $f(x) - 1 = \frac{x^2 - x + 2}{x - 1}$ est non nul.

Puisque $u_0 \neq 1$, on en déduit que la suite u est parfaitement définie.

On a $f(x) - x = \frac{x + 1}{x - 1}$ et $f(x) = x \Leftrightarrow x = -1$.

La seule limite finie possible de la suite u est donc $\ell = -1$.

Pour tout $x \neq 1$, on a :

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

Dans toute la suite, on note $\begin{cases} \alpha = 1 - \sqrt{2} \\ \beta = 1 + \sqrt{2} \end{cases}$

Avec ces notations, on peut écrire :

$$\forall x \neq 1, f'(x) = \frac{(x - \alpha)(x - \beta)}{(x - 1)^2}.$$

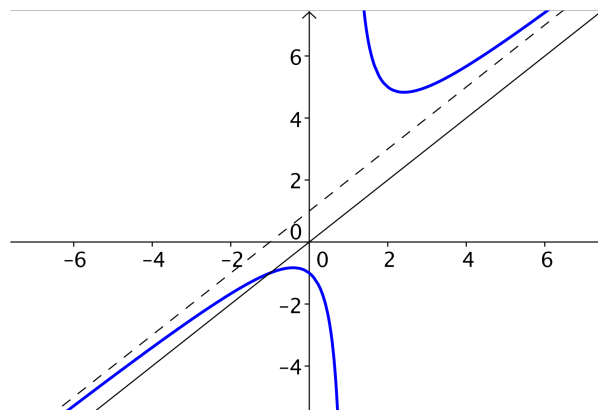
On en déduit le tableau de variations de f .

On y ajoute le signe de $f(x) - x = \frac{x + 1}{x - 1}$.

	$-\infty$	-1	α	0	-1	β	$+\infty$	
$f'(x)$		+	0	-		-	0	+
$f(x)$								
$f(x) - x$		+	0		-			+

Diagramme de variation de $f(x)$:
 - Pour $x < -1$, $f(x)$ est croissante de $-\infty$ vers 2α .
 - À $x = -1$, $f(x) = -1$.
 - Pour $-1 < x < \alpha$, $f(x)$ est décroissante de $-\infty$ vers 2α .
 - À $x = \alpha$, $f(x) = 2\alpha$.
 - Pour $\alpha < x < 0$, $f(x)$ est croissante de 2α vers $+\infty$.
 - À $x = 0$, $f(x) = -1$.
 - Pour $0 < x < \beta$, $f(x)$ est décroissante de $+\infty$ vers 2β .
 - À $x = \beta$, $f(x) = 2\beta$.
 - Pour $x > \beta$, $f(x)$ est croissante de 2β vers $+\infty$.

Voici la courbe représentative de f (avec l'asymptote $y = x + 1$ et la bissectrice $y = x$)



On peut maintenant étudier la suite u , suivant les valeurs de u_0 .

– Si $u_0 < -1$: pour tout $x < -1$, on a : $x < f(x) < -1$. En particulier $u_0 < u_1 < -1$ et $u_1 < u_2 < -1$.

Par une récurrence évidente, on en déduit : $\forall n \geq 0, u_n < u_{n+1} < -1$.

La suite u , croissante et majorée, converge vers sa seule limite finie possible $\ell = -1$.

– Si $u_0 = -1$: dans ce cas la suite u est constante : $\forall n \geq 0, u_n = -1$.

– Si $-1 < u_0 < 0$: pour tout x de $] -1, 0[$, on a : $-1 < f(x) < x < 0$.

En particulier $-1 < u_1 < u_0 < 0$ et $-1 < u_2 < u_1 < 0$.

Par une récurrence évidente, on en déduit : $\forall n \geq 0, -1 < u_{n+1} < u_n < 0$.

La suite u , décroissante et minorée, converge vers sa seule limite finie possible $\ell = -1$.

– Si $u_0 = 0$: on constate que $u_1 = f(0) = -1$, et donc : $\forall n \geq 1, u_n = -1$.

– Si $0 < u_0 < 1$: on a $u_1 = f(u_0) < -1$, et on est ramené au premier cas : $\lim u_n = -1$.

– Si $u_0 > 1$:

Pour tout $x > 1$ on a : $1 < x < f(x)$. En particulier $1 < u_0 < u_1$ et $1 < u_1 < u_2$.

Par une récurrence évidente, on en déduit : $\forall n \geq 0, 1 < u_n < u_{n+1}$.

La suite u est croissante ne peut pas converger vers -1 (sa seule limite finie possible).

On en déduit $u_n = +\infty$.

Conclusion : si $u_0 < 1$ on a $\lim u_n = -1$, et si $u_0 > 1$ on a $\lim u_n = +\infty$.

▷ Exemple n°4

On considère la suite $(u_n)_{n \geq 0}$ définie par $u_0 = 1$ et par : $\forall n \in \mathbb{N}, u_{n+1} = \sqrt{24 - 5u_n}$

La fonction $f : x \mapsto \sqrt{24 - 5x}$, définie sur $\mathcal{D} = \left] -\infty, \frac{24}{5} \right]$, est strictement décroissante.

L'équation $\ell = f(\ell)$ équivaut à $\ell^2 + 5\ell - 24 = 0$ et $\ell \geq 0$.

Or $\ell^2 + 5\ell - 24 = (\ell - 3)(\ell + 8)$. La seule limite finie possible de la suite u est donc $\ell = 3$.

On trouve $u_1 = \sqrt{19} \approx 4.359$, $u_2 = \sqrt{25 - 4\sqrt{19}} \approx 1.485$ et $u_3 = \sqrt{25 - 4\sqrt{25 - 4\sqrt{19}}} \approx 4.071$.

De même $u_4 = \sqrt{24 - 5u_3} \approx 1.908949125$, puis $u_5 = \sqrt{24 - 5u_4} \approx 3.802$, puis $u_6 = \sqrt{24 - 5u_5} \approx 2.234$.

Ces valeurs numériques indiquent que : $u_0 < u_2 < u_4 < u_6 < 3 < u_5 < u_3 < u_1$.

La fonction $g = f \circ f$ est croissante, et on a $u_{2(n+1)} = g(u_{2n})$ et $u_{2n+3} = g(u_{2n+1})$.

Une récurrence facile donne alors : $u_{2n} < u_{2n+2} < 3 < u_{2n+3} < u_{2n+1}$ pour tout n de \mathbb{N} .

Ainsi les suites de terme général $v_n = u_{2n}$ et $w_n = u_{2n+1}$ sont l'une croissante majorée, l'autre décroissante minorée, donc convergentes. On va montrer qu'elles convergent toutes les deux vers 3.

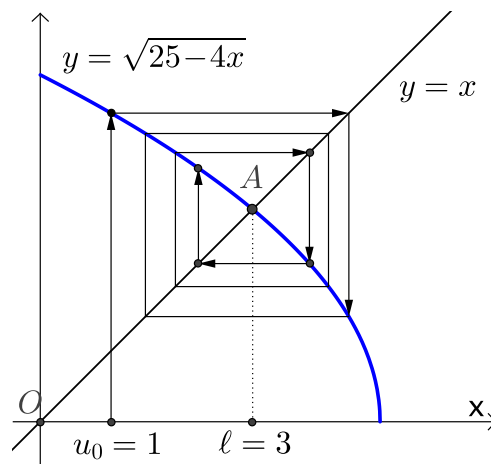
Pour tout n de \mathbb{N} , on a $|u_{n+1} - 3| = \left| \sqrt{24 - 5u_n} - 3 \right| = \frac{5}{3 + u_{n+1}} |u_n - 3|$.

On sait que $u_6 \leq u_n$ pour tout $n \geq 6$.

On en déduit $|u_{n+1} - 3| \leq \lambda |u_n - 3|$ pour tout $n \geq 6$, avec $\lambda = \frac{5}{3 + u_6} \approx 0.955 < 1$.

Il en résulte $|u_n - 3| \leq \lambda^{n-6} |u_6 - 3|$ pour tout $n \geq 6$, donc $\lim u_n = 3$.

La figure ci-dessous illustre la convergence (lente) de la suite $(u_n)_{n \geq 0}$ vers la valeur 3.



Chapitre 7

Limites, continuité

Sommaire

7.1	Limite en un point	168
7.1.1	Propriétés vraies « au voisinage d'un point »	168
7.1.2	Limite d'une fonction f en un point a de $\overline{\mathbb{R}}$	168
7.1.3	Limite d'une fonction f en $+\infty$ ou en $-\infty$	170
7.1.4	Unicité de la limite	171
7.1.5	Extensions de la notion de limite	171
7.1.6	Importance des limites à l'origine	173
7.1.7	Caractérisation séquentielle de la limite	173
7.2	Propriétés des limites	175
7.2.1	Opérations sur les limites	175
7.2.2	Limites et inégalités	176
7.2.3	Théorème de la limite monotone	177
7.3	Continuité	178
7.3.1	Continuité en un point	178
7.3.2	Caractérisation séquentielle de la continuité	180
7.3.3	Opérations sur les fonctions continues	180
7.4	Continuité sur un intervalle	180
7.4.1	Fonctions continues sur un intervalle	180
7.4.2	Théorème des valeurs intermédiaires	181
7.4.3	Approximation d'un zéro par dichotomie	181
7.4.4	Application continue sur un segment	182
7.4.5	Continuité et stricte monotonie sur un intervalle	183
7.5	Cas des fonctions continues complexes	184
7.5.1	Limite d'une fonction à valeurs complexes	185
7.5.2	Continuité en un point d'une fonction complexe	186
7.5.3	Continuité d'une fonction complexe sur un intervalle	186

7.1 Limite en un point

7.1.1 Propriétés vraies « au voisinage d'un point »

Dans ce chapitre, on sera amené à étudier la *limite* d'une fonction en un point a (éventuellement $a = +\infty$ ou $a = -\infty$). On sera aussi conduit à comparer des fonctions f et g « au voisinage de a ».

Par exemple pour écrire que si $f \leq g$ alors la limite de f en a est inférieure ou égale à celle de g en a :

- il *faut* que f et g soient toutes deux définies sur un même voisinage de a .
- il *suffit* que l'inégalité $f(x) \leq g(x)$ soit vraie au voisinage de a .

L'expression « au voisinage de a » peut s'entendre comme « suffisamment près de a », mais cela est un peu trop familier, et ça ne traduit pas le cas où a est égal à $+\infty$ ou à $-\infty$. On donne donc une définition plus précise de l'expression « une propriété vraie au voisinage d'un point a ».

Définition 7.1.1

Soit I un intervalle de \mathbb{R} , d'intérieur non vide.

Soit a un élément de I ou une « extrémité » de I (éventuellement $a = \pm\infty$).

Soit $\mathcal{P}(x)$ une proposition, vraie ou fausse selon les valeurs d'un élément x de I .

On dit que \mathcal{P} est vraie *au voisinage de a* si l'une des situations suivantes est réalisée :

- a est réel et il existe $\delta > 0$ tel que : $\forall x \in I \cap]a - \delta, a + \delta[, \mathcal{P}(x)$ est vraie.
- $a = +\infty$ et il existe un réel A tel que : $\forall x \geq A, \mathcal{P}(x)$ est vraie.
- $a = -\infty$ et il existe un réel A tel que : $\forall x \leq A, \mathcal{P}(x)$ est vraie.

Dans le premier cas, la clause « $x \in I$ » n'est utile que si a est une *extrémité* de I .

En effet si a est *intérieure* à I , alors pour tout δ assez petit, $]a - \delta, a + \delta[$ est inclus dans I .

Si f et g sont des fonctions définies sur I , on pourra par exemple écrire des propositions du genre : *si $f(x) \leq g(x)$ au voisinage de a , alors ...*

7.1.2 Limite d'une fonction f en un point a de $\overline{\mathbb{R}}$

Définition 7.1.2 (limite finie en un point a de \mathbb{R})

Soit $f : I \rightarrow \mathbb{R}$ une fonction à valeurs réelles.

Soit a un réel, élément ou extrémité de I . Soit ℓ un réel.

On dit que ℓ est limite de f en a si : $\forall \varepsilon > 0, \exists \delta > 0, (x \in I \text{ et } |x - a| \leq \delta) \Rightarrow |f(x) - \ell| \leq \varepsilon$.

Définition 7.1.3 (limite infinie en un point a de \mathbb{R})

Soit $f : I \rightarrow \mathbb{R}$ une fonction à valeurs réelles.

Soit a un réel, élément ou extrémité de I .

On dit que $+\infty$ est limite de f en a si : $\forall M \in \mathbb{R}, \exists \delta > 0, (x \in I \text{ et } |x - a| \leq \delta) \Rightarrow f(x) \geq M$.

On dit que $-\infty$ est limite de f en a si : $\forall M \in \mathbb{R}, \exists \delta > 0, (x \in I \text{ et } |x - a| \leq \delta) \Rightarrow f(x) \leq M$.

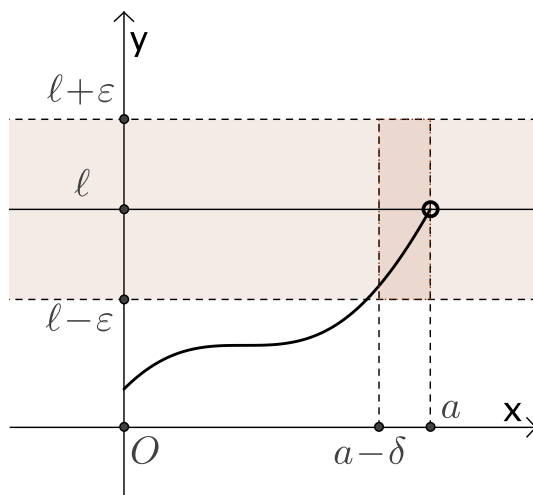
Remarques

- On vient de définir le sens de l'expression « ℓ est limite de f en a », avec a dans \mathbb{R} et ℓ dans $\overline{\mathbb{R}}$.
On dit aussi que $f(x)$ tend vers ℓ quand x tend vers a , et on note $f(x) \xrightarrow{x \rightarrow a} \ell$.
- Dans les trois cas, la clause « $x \in I$ » n'est pas nécessaire si a est intérieur à I .
Si a est dans l'intervalle de définition I de f , la seule limite possible de f en a est le réel $f(a)$.
Très souvent, les calculs de limites se font aux extrémités de l'intervalle de définition.
- Considérons la phrase : $\forall \varepsilon > 0, \exists \delta > 0, (x \in I \text{ et } |x - a| \leq \delta) \Rightarrow |f(x) - \ell| \leq \varepsilon$.
Cette proposition doit se lire de la façon suivante :
 - pour tout réel strictement positif ε (sous entendu : « aussi petit soit-il »),
 - il existe un réel strictement positif δ (qui à ce stade dépend donc certainement de ε),
 - tel que, pour tout x de I , si x est à moins de δ de a , alors $f(x)$ est à moins de ε de ℓ .
- Pour réviser un peu la logique, on considère la phrase, obtenue par interversion des quantificateurs :
 $\exists \delta > 0, \forall \varepsilon > 0, (x \in I \text{ et } |x - a| \leq \delta) \Rightarrow |f(x) - \ell| \leq \varepsilon$
On demande de la traduire en français, et de comprendre en quoi elle diffère de la phrase précédente.
- Considérons la phrase : $\forall M \in \mathbb{R}, \exists \delta > 0, (x \in I \text{ et } |x - a| \leq \delta) \Rightarrow f(x) \geq M$.
Cette proposition doit se lire de la façon suivante :
 - pour tout réel M (sous entendu : « aussi grand soit-il du côté de $+\infty$ »),
 - il existe un réel strictement positif δ (qui à ce stade dépend donc certainement de M),
 - tel que, pour tout x de I , si x est à moins de δ de a , alors $f(x)$ est supérieur ou égal à M .
 Dans cette phrase, on peut se limiter à $M > 0$ sans changer la portée de la définition.
- De même dans « $\forall M \in \mathbb{R}, \exists \delta > 0, (x \in I \text{ et } |x - a| \leq \delta) \Rightarrow f(x) \leq M$ » on peut supposer $M < 0$ (sans importance car la propriété doit être vraie pour M aussi grand qu'on veut du côté de $-\infty$).

On illustre ici le fait que ℓ (réel) est limite de f en a (réel). On suppose que f est définie sur un intervalle I dont a est l'extrémité droite.

On a choisi ici une valeur particulière de ε (un peu exagérée pour la clarté du dessin), et une valeur de δ qui « convient » pour un tel ε .

On devine que si δ convient pour ε , alors tout δ' , avec $0 < \delta' < \delta$ convient aussi. Il ne sert à rien de chercher le « meilleur » δ possible (c'est-à-dire le plus grand possible) pour un ε donné. L'existence d'un certain δ (dépendant de ε) est suffisante.



7.1.3 Limite d'une fonction f en $+\infty$ ou en $-\infty$

Définition 7.1.4 (limite finie en $+\infty$)

Soit $f : I \rightarrow \mathbb{R}$ une fonction, où I est un intervalle non majoré. Soit ℓ un nombre réel. On dit que ℓ est limite de f en $+\infty$ si : $\forall \varepsilon > 0, \exists A \in \mathbb{R}, (x \geq A \Rightarrow |f(x) - \ell| \leq \varepsilon)$.

La phrase : $\forall \varepsilon > 0, \exists A \in \mathbb{R}, (x \geq A \Rightarrow |f(x) - \ell| \leq \varepsilon)$ doit se lire de la façon suivante :

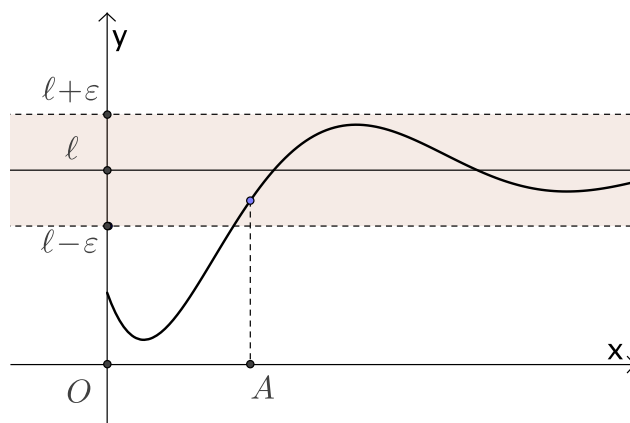
- pour tout réel strictement positif ε (sous entendu : « aussi petit soit-il »),
- il existe un réel A (qui à ce stade dépend donc certainement de ε),
- tel que, pour $x \geq A$ (c'est-à-dire « au-delà de A »), alors $f(x)$ est à moins de ε de ℓ .

On illustre le fait que ℓ (réel) est limite de f en $+\infty$. On suppose que f est définie sur un intervalle I non majoré (donc « qui va jusqu'à $+\infty$ »).

On a choisi ici une valeur particulière de ε , puis une valeur de A qui « convient » pour un tel ε .

Si A convient pour ε , tout $A' > A$ convient aussi.

Il ne sert à rien de chercher le « meilleur » A possible (le plus petit possible) pour un ε donné. L'existence d'un certain A (dépendant de ε) est suffisante.



On définit ensuite la notion de limite infinie, « quand x tend vers $+\infty$ » :

Définition 7.1.5 (limite infinie en $+\infty$)

Soit $f : I \rightarrow \mathbb{R}$ une fonction, où I est un intervalle non majoré.

On dit que $+\infty$ est limite de f en $+\infty$ si : $\forall M \in \mathbb{R}, \exists A \in \mathbb{R}, (x \geq A \Rightarrow f(x) \geq M)$.

On dit que $-\infty$ est limite de f en $+\infty$ si : $\forall M \in \mathbb{R}, \exists A \in \mathbb{R}, (x \geq A \Rightarrow f(x) \leq M)$.

On examine finalement la notion de limite « quand x tend vers $-\infty$ ».

Définition 7.1.6 (limite finie en $-\infty$)

Soit $f : I \rightarrow \mathbb{R}$ une fonction, où I est un intervalle non minoré. Soit ℓ un nombre réel.

On dit que ℓ est limite de f en $-\infty$ si : $\forall \varepsilon > 0, \exists A \in \mathbb{R}, (x \leq A \Rightarrow |f(x) - \ell| \leq \varepsilon)$.

Définition 7.1.7 (limite infinie en $-\infty$)

Soit $f : I \rightarrow \mathbb{R}$ une fonction, où I est un intervalle non minoré.

On dit que $+\infty$ est limite de f en $-\infty$ si : $\forall M \in \mathbb{R}, \exists A \in \mathbb{R}, (x \leq A \Rightarrow f(x) \geq M)$.

On dit que $-\infty$ est limite de f en $-\infty$ si : $\forall M \in \mathbb{R}, \exists A \in \mathbb{R}, (x \leq A \Rightarrow f(x) \leq M)$.

En définitive, on a donné un sens à l'expression « ℓ est limite de f en a », avec a et ℓ dans $\overline{\mathbb{R}}$.

Pour exprimer cette situation, on dit que $f(x)$ tend vers ℓ quand x tend vers a , et on note $f(x) \xrightarrow{x \rightarrow a} \ell$.

7.1.4 Unicité de la limite

Proposition 7.1.1 (Unicité de la limite)

Soit $f : I \rightarrow \mathbb{R}$ une fonction à valeurs réelles.

Soit a dans $\overline{\mathbb{R}}$, élément ou borne de l'intervalle I . Soit ℓ un élément de $\overline{\mathbb{R}}$.

On suppose que ℓ est limite de f en a , au sens de l'une des définitions précédentes.

Alors ℓ est le seul élément de $\overline{\mathbb{R}}$ à posséder cette propriété.

On l'appelle la limite de f en a , et on note $\lim_{x \rightarrow a} f(x) = \ell$, ou $\lim_a f = \ell$ ou $f(x) \xrightarrow{x \rightarrow a} \ell$.

Remarques

Il se peut qu'une fonction ne possède pas de limite en un point.

Par exemple, la fonction $x \mapsto \cos x$ n'a pas de limite en $-\infty$ ou en $+\infty$.

De même, la fonction $x \mapsto \lfloor x \rfloor$ (partie entière de x) n'a pas de limite en un élément k de \mathbb{Z} .

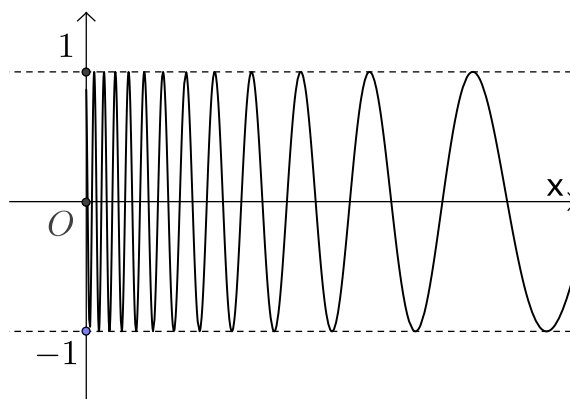
Soit f définie sur \mathbb{R}^{+*} par $f(x) = \cos\left(\frac{1}{x}\right)$.

On trace ici une partie du graphe de f .

Les oscillations de f s'accroissent quand $x \rightarrow 0$.

Cette fonction ne possède pas de limite en 0.

En fait, sur chaque intervalle $]0, a[$, avec $a > 0$, la fonction f prend une infinité de fois chacune des valeurs du segment $[-1, 1]$.



Le résultat suivant est une conséquence immédiate des définitions sur les limites :

Proposition 7.1.2 (limite de f en a si f est définie en a)

Si la fonction f est définie en le réel a , sa seule limite **possible** en a est $f(a)$.

Autrement dit : si $f(a)$ et la limite de f en a existent, alors nécessairement : $\lim_{x \rightarrow a} f(x) = f(a)$.

7.1.5 Extensions de la notion de limite

Définition 7.1.8 (limite par valeurs inférieures ou par valeurs supérieures)

On suppose que la limite de f en a (élément de $\overline{\mathbb{R}}$) est le réel ℓ .

Si $f(x) \geq \ell$ au voisinage de a , on dit que $f(x)$ tend en a vers ℓ « par valeurs supérieures ».

Si $f(x) \leq \ell$ au voisinage de a , on dit que $f(x)$ tend en a vers ℓ « par valeurs inférieures ».

On peut alors noter $\lim_{x \rightarrow a} f(x) = \ell^+$ (respectivement $\lim_{x \rightarrow a} f(x) = \ell^-$).

Remarque

Si $\lim_{x \rightarrow a} f(x)$ est un réel ℓ , il n'y a aucune raison a priori pour ℓ soit « obtenu » par valeurs supérieures ou « inférieures » (c'est-à-dire pour que $f(x) - \ell$ garde un signe constant au voisinage de a).

Par exemple, $\lim_{x \rightarrow +\infty} \frac{\sin(x)}{x} = 0$, mais le signe de $\frac{\sin(x)}{x}$ n'est constant sur aucun intervalle $[A, +\infty[$.

Définition 7.1.9 (limite à gauche)

Soit $f : I \rightarrow \mathbb{R}$ une fonction réelle définie sur un intervalle I .

On suppose que le réel a est intérieur à l'intervalle I . Soit ℓ un élément de $\overline{\mathbb{R}}$.

Soit g la restriction de f à l'intervalle $J = I \cap]-\infty, a[$.

Le réel a est donc l'extrémité droite de J , et n'appartient pas à J .

On dit que f admet ℓ pour limite en a à gauche si g admet ℓ pour limite en a .

On note alors $\lim_{x \rightarrow a^-} f(x) = \ell$, ou $\lim_{a^-} f = \ell$, ou $f(x) \xrightarrow{x \rightarrow a^-} \ell$

Définitions équivalentes : en notant ℓ un nombre réel, on a les définitions équivalentes

$$\begin{cases} \lim_{x \rightarrow a^-} f(x) = \ell & \Leftrightarrow \forall \varepsilon > 0, \exists \delta > 0, (a - \delta \leq x < a) \Rightarrow |f(x) - \ell| \leq \varepsilon \\ \lim_{x \rightarrow a^-} f(x) = +\infty & \Leftrightarrow \forall M \in \mathbb{R}, \exists \delta > 0, (a - \delta \leq x < a) \Rightarrow f(x) \geq M \\ \lim_{x \rightarrow a^-} f(x) = -\infty & \Leftrightarrow \forall M \in \mathbb{R}, \exists \delta > 0, (a - \delta \leq x < a) \Rightarrow f(x) \leq M \end{cases}$$

Définition 7.1.10 (limite à droite)

Soit $f : I \rightarrow \mathbb{R}$ une fonction réelle définie sur un intervalle I .

On suppose que le réel a est intérieur à l'intervalle I . Soit ℓ un élément de $\overline{\mathbb{R}}$.

Soit g la restriction de f à l'intervalle $J = I \cap]a, +\infty[$.

Le réel a est donc l'extrémité gauche de J , et n'appartient pas à J .

On dit que f admet ℓ pour limite en a à droite si g admet ℓ pour limite en a .

On note alors $\lim_{x \rightarrow a^+} f(x) = \ell$, ou $\lim_{a^+} f = \ell$, ou $f(x) \xrightarrow{x \rightarrow a^+} \ell$

Définitions équivalentes : en notant ℓ un nombre réel, on a les définitions équivalentes

$$\begin{cases} \lim_{x \rightarrow a^+} f(x) = \ell & \Leftrightarrow \forall \varepsilon > 0, \exists \delta > 0, (a < x \leq a + \delta) \Rightarrow |f(x) - \ell| \leq \varepsilon \\ \lim_{x \rightarrow a^+} f(x) = +\infty & \Leftrightarrow \forall M \in \mathbb{R}, \exists \delta > 0, (a < x \leq a + \delta) \Rightarrow f(x) \geq M \\ \lim_{x \rightarrow a^+} f(x) = -\infty & \Leftrightarrow \forall M \in \mathbb{R}, \exists \delta > 0, (a < x \leq a + \delta) \Rightarrow f(x) \leq M \end{cases}$$

Remarque : la limite de f en a , à gauche ou à droite, si elle existe, est unique. De même, la plupart des propriétés vraies pour les limites le sont encore s'il s'agit de limites à gauche ou à droite.

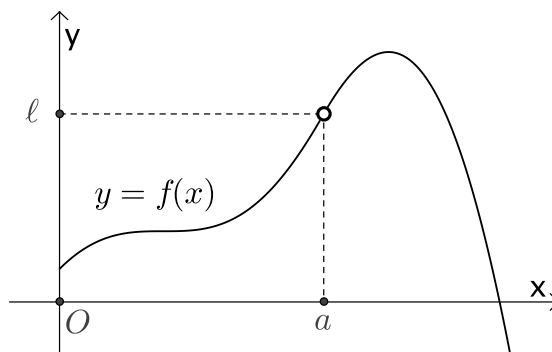
Extension de la notion de limite si f est définie sur $I \setminus \{a\}$

On suppose ici que f est définie sur $I \setminus \{a\}$, où a est un point intérieur à l'intervalle I (autrement dit la fonction f est définie sur I sauf en a).

On suppose également qu'il existe un réel ℓ tel que $\lim_{x \rightarrow a^-} f(x) = \lim_{x \rightarrow a^+} f(x) = \ell$.

Dans ces conditions, on dira encore que ℓ est la limite de f en a (même si f n'est pas définie en a).

Le plus souvent d'ailleurs, on prolongera f au point a en posant $f(a) = \ell$.



7.1.6 Importance des limites à l'origine

Le résultat suivant explique qu'on peut toujours se ramener à des limites calculées à l'origine (donc « quand x tend vers 0 »), et même (dans le cas de limites finies) à une limite égale à 0 :

Proposition 7.1.3 (importance des limites nulles ou des limites en 0)

Si a est un réel, et ℓ un élément de $\overline{\mathbb{R}}$, on a l'équivalence : $\lim_{x \rightarrow a} f(x) = \ell \Leftrightarrow \lim_{x \rightarrow 0} f(a+x) = \ell$.

Si ℓ est un réel, et a un élément de $\overline{\mathbb{R}}$, on a l'équivalence : $\lim_{x \rightarrow a} f(x) = \ell \Leftrightarrow \lim_{x \rightarrow a} (f(x) - \ell) = 0$.

Les « limites usuelles » sont pour la plupart connues à l'origine (en 0). Pour une limite en un réel a non nul, on se ramènera donc souvent au cas usuel en posant $x = a + h$ avec « h tendant vers 0 ».

Proposition 7.1.4 (limite finie et caractère borné de f)

Soit a un élément de $\overline{\mathbb{R}}$. Si $\lim_{x \rightarrow a} f(x)$ existe et est finie, alors f est bornée au voisinage de a .

7.1.7 Caractérisation séquentielle de la limite

Proposition 7.1.5 (caractérisation séquentielle des limites)

Soit f une fonction définie sur l'intervalle I , à valeurs réelles.

Soit a un élément de $\overline{\mathbb{R}}$ (élément de I ou extrémité de I). Soit ℓ un élément de $\overline{\mathbb{R}}$.

Alors $\lim_{x \rightarrow a} f(x) = \ell$ si et seulement si, pour toute suite (u_n) de I tendant vers a , on a : $\lim f(u_n) = \ell$.

Cette propriété est importante pour deux raisons :

– D'une part, elle permet d'utiliser les résultats connus sur les limites de suites pour en déduire des résultats sur les limites de fonctions (voir plus loin).

– D'autre part, elle est utile pour montrer qu'une fonction ne possède pas de limite en un point a .

Il suffit en effet de mettre en évidence deux suites (u_n) et (v_n) de l'intervalle I , tendant toutes deux vers a , et telles que les suites $(f(u_n))$ et $(f(v_n))$ n'aient pas la même limite.

Considérons par exemple la fonction f définie sur \mathbb{R}^{+*} par $f(x) = \cos\left(\frac{1}{x}\right)$.

Posons $u_n = \frac{1}{2n\pi}$ et $v_n = \frac{1}{(2n+1)\pi}$. Pour tout n , on a $f(u_n) = 1$ et $f(v_n) = -1$.

Ainsi $\lim u_n = \lim v_n = 0$, mais les suites $(f(u_n))$ et $(f(v_n))$ n'ont pas la même limite.

Il en résulte que la fonction f n'a pas de limite en 0.

On notera bien que la proposition précédente est valable pour tous les cas particuliers de limites (en un point a de \mathbb{R} , ou quand x tend vers $\pm\infty$), que la limite soit elle-même un nombre réel ou $\pm\infty$.

Enfin, cette propriété se généralise aux limites à gauche et à droite.

7.2 Propriétés des limites

7.2.1 Opérations sur les limites

Proposition 7.2.1 (limite et valeur absolue)

Si $\lim_{x \rightarrow a} f(x) = \ell$, avec ℓ dans \mathbb{R} , alors $\lim_{x \rightarrow a} |f(x)| = |\ell|$.

Ce résultat est encore valable si $\ell = -\infty$ ou $\ell = +\infty$, à condition de noter $|\infty| = |-\infty| = +\infty$.

L'existence de $\lim_{x \rightarrow a} |f(x)|$ n'implique pas celle de $\lim_{x \rightarrow a} f(x)$, si on ne sait rien du signe de f .

En revanche, on a l'équivalence : $\lim_{x \rightarrow a} f(x) = 0 \Leftrightarrow \lim_{x \rightarrow a} |f(x)| = 0$.

Si ℓ est un réel, on a les équivalences : $\lim_{x \rightarrow a} f(x) = \ell \Leftrightarrow \lim_{x \rightarrow a} (f(x) - \ell) = 0 \Leftrightarrow \lim_{x \rightarrow a} |f(x) - \ell| = 0$.

Proposition 7.2.2 (limites et combinaisons linéaires)

On suppose que $\lim_{x \rightarrow a} f(x) = \ell$ et que $\lim_{x \rightarrow a} g(x) = \ell'$, avec ℓ et ℓ' dans \mathbb{R} .

Alors $\lim_{x \rightarrow a} (f + g)(x) = \ell + \ell'$.

Plus généralement, pour tout (α, β) de \mathbb{R}^2 , on a : $\lim_{x \rightarrow a} (\alpha f + \beta g)(x) = \alpha \ell + \beta \ell'$.

Ce résultat s'étend à ℓ ou ℓ' dans $\{-\infty, +\infty\}$, à condition que $\alpha \ell + \beta \ell'$ ait un sens dans $\overline{\mathbb{R}}$.

Ainsi, on ne peut rien dire en général pour $\lim_{x \rightarrow a} (f + g)(x)$ si $\lim_{x \rightarrow a} f(x) = +\infty$ et si $\lim_{x \rightarrow a} g(x) = -\infty$.

On dit dans ce cas qu'on est en présence de la forme indéterminée « $\infty - \infty$ ».

Il faut alors faire une étude spécifique et « lever » cette indétermination.

Proposition 7.2.3 (limites et produits)

On suppose que $\lim_{x \rightarrow a} f(x) = \ell$ et $\lim_{x \rightarrow a} g(x) = \ell'$, avec ℓ et ℓ' dans \mathbb{R} . Alors $\lim_{x \rightarrow a} (fg)(x) = \ell \ell'$.

Ce résultat s'étend à ℓ ou ℓ' dans $\{-\infty, +\infty\}$, à condition que $\ell \ell'$ ait un sens dans $\overline{\mathbb{R}}$.

Ainsi, on ne peut rien dire en général pour $\lim_{x \rightarrow a} (fg)(x)$ si $\lim_{x \rightarrow a} f(x) = 0$ et si $\lim_{x \rightarrow a} g(x) \in \{-\infty, +\infty\}$.

On dit dans ce cas qu'on est en présence de la forme indéterminée « 0∞ ».

Il faut alors faire une étude spécifique et « lever » cette indétermination.

Proposition 7.2.4 (majoration ou minoration de l'inverse)

Si $\lim_{x \rightarrow a} f(x) = \ell$, avec ℓ dans \mathbb{R}^* , alors au voisinage de a : $|f(x)| \geq \frac{1}{2} |\ell|$, donc $\frac{1}{|f(x)|} \leq \frac{2}{|\ell|}$.

Si $\lim_{x \rightarrow a} f(x) = \ell > 0$ (resp. $\ell < 0$), alors au voisinage de a : $f(x) > \frac{\ell}{2} > 0$ (resp. $f(x) < \frac{\ell}{2} < 0$).

Proposition 7.2.5 (limites et passage à l'inverse)

Si $\lim_{x \rightarrow a} f(x) = \ell$, avec ℓ dans \mathbb{R}^* , alors $\lim_{x \rightarrow a} \frac{1}{f(x)} = \frac{1}{\ell}$.

Si $\lim_{x \rightarrow a} f(x) = 0$ et $f(x) > 0$ (resp. $f(x) < 0$) au vois. de a , alors $\lim_{x \rightarrow a} \frac{1}{f(x)} = +\infty$ (resp. $\lim_{x \rightarrow a} \frac{1}{f(x)} = -\infty$).

Si $\lim_{x \rightarrow a} f(x) = -\infty$ ou $\lim_{x \rightarrow a} f(x) = +\infty$, alors $\lim_{x \rightarrow a} \frac{1}{f(x)} = 0$.

Ce qui précède permet de conclure dans le calcul de $\lim_{x \rightarrow a} \frac{f(x)}{g(x)}$, sauf dans les cas suivants :

- Si $\lim_{x \rightarrow a} f(x) = 0$ et $\lim_{x \rightarrow a} g(x) = 0$, on parle de la forme indéterminée « $\frac{0}{0}$ »
- Si $\lim_{x \rightarrow a} f(x) = \pm\infty$ et $\lim_{x \rightarrow a} g(x) = \pm\infty$, on parle de la forme indéterminée « $\frac{\infty}{\infty}$ »

Pour $\lim_{x \rightarrow a} f(x)^{g(x)}$, trois formes indéterminées se ramènent à « 0∞ » car $f(x)^{g(x)} = e^{g(x)\ln(f(x))}$:

Ces trois formes indéterminées sont

$$\begin{cases} \text{« } 1^\infty \text{ » si } \lim_{x \rightarrow a} f(x) = 1 \text{ et } \lim_{x \rightarrow a} g(x) = \pm\infty \\ \text{« } \infty^0 \text{ » si } \lim_{x \rightarrow a} f(x) = +\infty \text{ et } \lim_{x \rightarrow a} g(x) = 0 \\ \text{« } 0^0 \text{ » si } \lim_{x \rightarrow a} f(x) = 0^+ \text{ et } \lim_{x \rightarrow a} g(x) = 0 \end{cases}$$

Avec une forme indéterminée, tout est possible : il faut faire une étude spécifique pour chaque cas.

Proposition 7.2.6 (composition des limites)

On suppose que la fonction $g \circ f$ est définie au voisinage de a .

Si $\lim_{x \rightarrow a} f(x) = b$ et $\lim_{x \rightarrow b} g(x) = \ell$, alors $\lim_{x \rightarrow a} (g \circ f)(x) = \ell$.

7.2.2 Limites et inégalités

Proposition 7.2.7 (conservation des inégalités large par passage à la limite)

On suppose que $\lim_{x \rightarrow a} f(x) = \ell$ et $\lim_{x \rightarrow a} g(x) = \ell'$, avec ℓ et ℓ' dans $\overline{\mathbb{R}}$.

Si on a $f(x) \leq g(x)$ au voisinage de a , alors $\ell \leq \ell'$.

- En particulier, si $\lim_{x \rightarrow a} f(x) = \ell$ et si λ est un nombre réel :
Si $f(x) \leq \lambda$ au voisinage de a , alors $\ell \leq \lambda$. Si $f(x) \geq \lambda$ au voisinage de a , alors $\ell \geq \lambda$.
- Si $f(x) < g(x)$ au voisinage de a , alors on peut seulement affirmer que $\lim_{x \rightarrow a} f(x) \leq \lim_{x \rightarrow a} g(x)$.
Ainsi, par passage à la limite, les inégalités strictes « deviennent » des inégalités larges.

Proposition 7.2.8

On suppose que $\lim_{x \rightarrow a} f(x) = \ell$ et $\lim_{x \rightarrow a} g(x) = \ell'$, avec a, ℓ, ℓ' dans $\overline{\mathbb{R}}$.

Si $\ell < \ell'$, alors l'inégalité $f(x) < g(x)$ est vraie au voisinage de a .

En particulier, si λ est un nombre réel :

Si $\ell < \lambda$, alors $f(x) < \lambda$ au voisinage de a ; si $\ell > \lambda$, alors $f(x) > \lambda$ au voisinage de a .

On retrouve le « théorème des gendarmes » :

Proposition 7.2.9 (« passage à la limite par encadrement »)

On suppose que $\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} g(x) = \ell$, avec a et ℓ dans $\overline{\mathbb{R}}$.

Si on a l'encadrement $f(x) \leq h(x) \leq g(x)$ au voisinage de a , alors $\lim_{x \rightarrow a} h(x) = \ell$.

Cas particuliers importants

- Si $|f(x)| \leq g(x)$ au voisinage de a et si $\lim_{x \rightarrow a} g(x) = 0$, alors $\lim_{x \rightarrow a} f(x) = 0$.
- Si f est bornée au voisinage de a et si $\lim_{x \rightarrow a} g(x) = 0$, alors $\lim_{x \rightarrow a} (fg)(x) = 0$.
- Supposons $f(x) \leq g(x)$ au voisinage de a :
$$\begin{cases} \text{si } \lim_{x \rightarrow a} f(x) = +\infty, \text{ alors } \lim_{x \rightarrow a} g(x) = +\infty. \\ \text{si } \lim_{x \rightarrow a} g(x) = -\infty, \text{ alors } \lim_{x \rightarrow a} f(x) = -\infty. \end{cases}$$

7.2.3 Théorème de la limite monotone

Proposition 7.2.10 (limite aux bornes, pour une fonction monotone)

Soit f une fonction monotone de $]a, b[$ dans \mathbb{R} , avec a et b dans $\overline{\mathbb{R}}$.

Alors la limite ℓ de f en a et la limite ℓ' de f en b existent dans $\overline{\mathbb{R}}$.

Plus précisément, si f est croissante : on a $\ell = \inf_{a < x < b} f(x)$ et $\ell' = \sup_{a < x < b} f(x)$.

Et si f est décroissante, alors : $\ell = \sup_{a < x < b} f(x)$ et $\ell' = \inf_{a < x < b} f(x)$.

Les différents cas possibles

- Supposons f croissante :
$$\begin{cases} \text{si elle est majorée, } \ell' \text{ est un réel, sinon } \ell' = +\infty \\ \text{si elle est minorée, } \ell \text{ est un réel, sinon } \ell = -\infty \end{cases}$$
- Supposons f décroissante :
$$\begin{cases} \text{si elle est minorée, } \ell' \text{ est un réel, sinon } \ell' = -\infty \\ \text{si elle est majorée, } \ell \text{ est un réel, sinon } \ell = +\infty \end{cases}$$

Proposition 7.2.11 (théorème de la limite monotone)

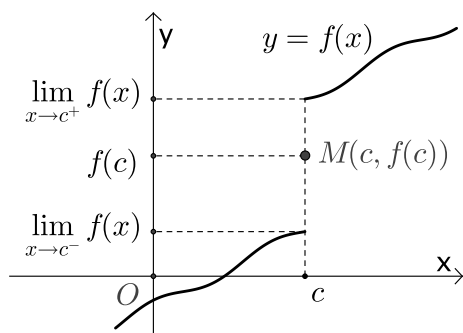
Soit f une fonction monotone de $]a, b[$ dans \mathbb{R} , avec a et b dans $\overline{\mathbb{R}}$. Soit c dans $]a, b[$.

La fonction f admet en c une limite à gauche et une limite à droite, toutes deux finies.

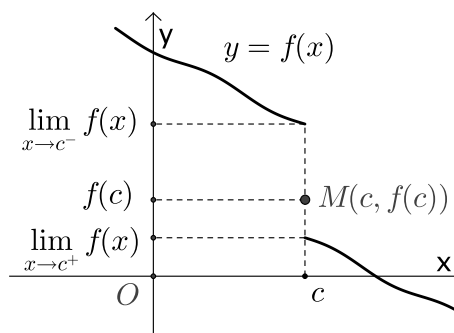
Si f est croissante, alors : $\lim_{x \rightarrow c^-} f(x) \leq f(c) \leq \lim_{x \rightarrow c^+} f(x)$.

Si f est décroissante, alors : $\lim_{x \rightarrow c^-} f(x) \geq f(c) \geq \lim_{x \rightarrow c^+} f(x)$.

Voici une illustration du résultat précédent, dans le cas (un peu rare tout de même) où les limites à gauche et à droite de f en c sont toutes deux distinctes de la valeur $f(c)$.



Limites à gauche et à droite en c avec f croissante au voisinage de c



Limites à gauche et à droite en c avec f décroissante au voisinage de c

7.3 Continuité

7.3.1 Continuité en un point

Dans toute la suite I est un intervalle de \mathbb{R} , d'intérieur non vide.

Définition 7.3.1 (définition de la continuité en un point)

Soit $f : I \rightarrow \mathbb{R}$ une fonction numérique, et soit a un élément de I .

On dit que f est *continue* en a si la limite de f en a existe.

Puisque f est définie en a , cela équivaut à dire : $\lim_{x \rightarrow a} f(x) = f(a)$.

Donc f est continue en a si et seulement si : $\forall \varepsilon > 0, \exists \delta > 0, (x \in I \text{ et } |x - a| \leq \delta) \Rightarrow |f(x) - f(a)| \leq \varepsilon$

Définition 7.3.2 (prolongement par continuité)

Soit I un intervalle d'intérieur non vide. Soit a un élément de I .

Soit f une fonction numérique réelle définie sur $I \setminus \{a\}$.

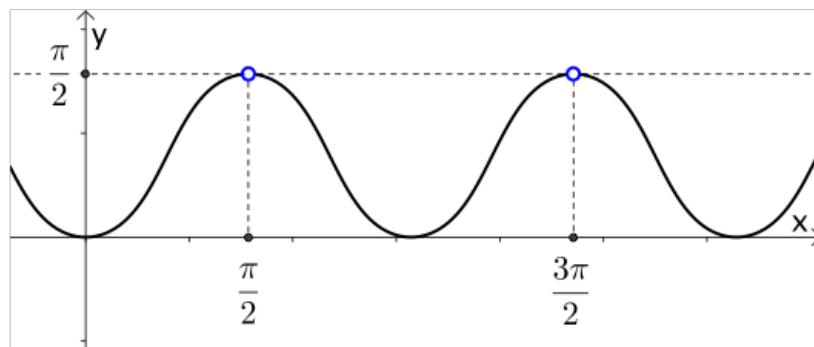
On dit que f est prolongeable par continuité en a si la limite $\ell = \lim_{x \rightarrow a} f(x)$ existe et est finie.

Si on pose $f(a) = \ell$, la fonction f ainsi prolongée devient continue en a .

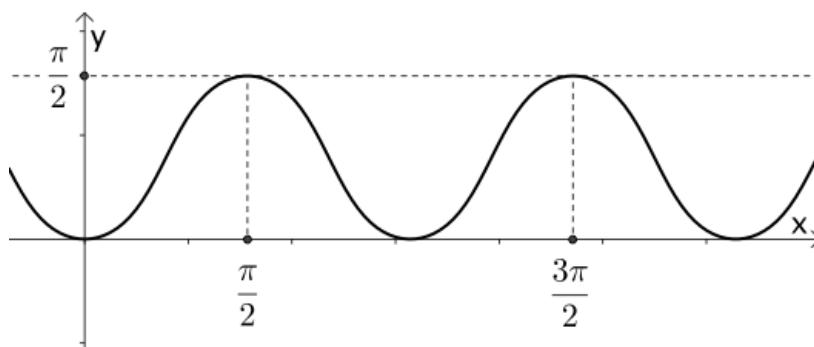
On dit qu'on a effectué le *prolongement par continuité* de f au point a .

Exemple : on a représenté une partie de la courbe représentative de $f : x \mapsto \arctan(\tan^2(x))$.

Cette fonction est définie sauf aux $a_k = \frac{\pi}{2} + k\pi$, mais $\lim_{x \rightarrow a_k^-} f(x) = \lim_{x \rightarrow a_k^+} f(x) = \frac{\pi}{2}$.



Si on pose $f\left(\frac{\pi}{2} + k\pi\right) = \frac{\pi}{2}$, la fonction f devient continue en tout point de \mathbb{R} (elle est π -périodique).



Définition 7.3.3 (continuité à gauche en un point)

Soit $f : I \rightarrow \mathbb{R}$. Soit a un élément de I , qui n'en soit pas l'extrémité gauche.

Posons $J = I \cap]-\infty, a]$: l'intervalle J est d'intérieur non vide, et a est son extrémité droite.

Soit g la restriction de f à J . On dit que f est *continue à gauche* en a si g est continue en a .

Cela équivaut à $\lim_{x \rightarrow a^-} f(x) = f(a)$ ou encore à : $\forall \varepsilon > 0, \exists \delta > 0, (a - \delta \leq x \leq a) \Rightarrow |f(x) - f(a)| \leq \varepsilon$

Définition 7.3.4 (continuité à droite en un point)

Soit $f : I \rightarrow \mathbb{R}$. Soit a un élément de I , qui n'en soit pas l'extrémité droite.

Posons $J = I \cap [a, +\infty[$: l'intervalle J est d'intérieur non vide, et a est son extrémité gauche.

Soit g la restriction de f à J . On dit que f est *continue à droite* en a si g est continue en a .

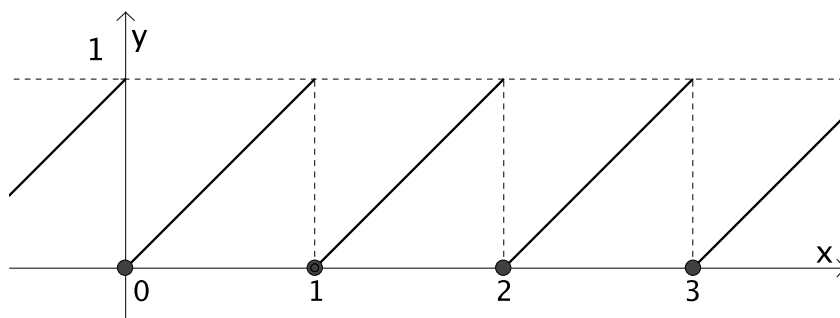
Cela équivaut à : $\lim_{x \rightarrow a^+} f(x) = f(a)$ ou encore : $\forall \varepsilon > 0, \exists \delta > 0, (a \leq x \leq a + \delta) \Rightarrow |f(x) - f(a)| \leq \varepsilon$

Soit a un point intérieur à l'intervalle I . Soit f une fonction de I dans \mathbb{R} . Alors f est continue en a si et seulement si f est continue à droite et à gauche en a .

On a représenté ci-dessous la fonction $f : x \mapsto x - \lfloor x \rfloor$.

En tout point de \mathbb{Z} , la fonction f est continue à droite mais pas à gauche.

Pour tout a de \mathbb{Z} , on a en effet : $\lim_{x \rightarrow a^-} f(x) = 1$ et $\lim_{x \rightarrow a^+} f(x) = 0 = f(a)$.

**Définition 7.3.5** (discontinuité de première espèce)

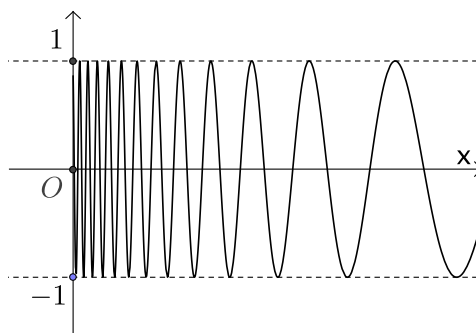
Soit f une fonction numérique réelle, définie sur l'intervalle I . Soit a un point de I .

Si f n'est pas continue en a , on dit que f est *discontinue* en ce point.

Si a est intérieur à I , si f est discontinue en a , mais si les limites à gauche et à droite en a existent et sont finies, on dit que f présente en a une *discontinuité de première espèce*.

Si f est discontinue en a , et si l'une au moins des deux limites (à gauche ou à droite) en a n'existe pas (ou est infinie), on dit que la discontinuité de f en a est de seconde espèce.

C'est le cas notamment de la discontinuité de la fonction $x \mapsto \cos(1/x)$ à l'origine (quelle que soit la valeur qu'on serait tenté de donner à $f(0)$).



7.3.2 Caractérisation séquentielle de la continuité

Proposition 7.3.1 (caractérisation séquentielle de la continuité en un point)

Soit f une fonction numérique réelle, définie sur l'intervalle I . Soit a un point de I .

La fonction f est continue en a si et seulement si, pour toute suite (u_n) de I convergeant vers a , la suite de terme général $f(u_n)$ converge vers $f(a)$.

Le résultat précédent est utile pour montrer que f n'est pas continue en un point a : il suffit en effet de construire une suite (u_n) convergeant vers a , mais telle que la suite $(f(u_n))$ ne converge pas vers $f(a)$.

7.3.3 Opérations sur les fonctions continues

Proposition 7.3.2 (combinaisons linéaires, produits, quotients)

Soit f et g deux fonctions numériques réelles, définies sur l'intervalle I . Soit a un élément de I .

On suppose que f et g sont continues en a .

Alors il en est de même pour $\alpha f + \beta g$ (avec α, β dans \mathbb{R}), pour fg , et (si $g(a) \neq 0$) pour $\frac{1}{g}$ et $\frac{f}{g}$.

Proposition 7.3.3 (composition de fonctions continues en un point)

Soit $f : I \rightarrow \mathbb{R}$ et $g : J \rightarrow \mathbb{R}$ deux fonctions numériques, avec $f(I) \subset J$.

Si f est continue en a et si g est continue en $b = f(a)$, alors $g \circ f$ est continue en a .

7.4 Continuité sur un intervalle

7.4.1 Fonctions continues sur un intervalle

Définition 7.4.1

Soit f une fonction numérique réelle, définie sur l'intervalle I .

On dit que f est continue sur I si f est continue en tout point de I .

On note $\mathcal{C}(I, \mathbb{R})$ l'ensemble des fonctions continues sur I , à valeurs réelles.

Applications continues usuelles

- Les fonctions constantes, les fonctions $x \mapsto x$ et $x \mapsto |x|$, sont continues sur \mathbb{R} .
- Les fonctions polynomiales sont continues sur \mathbb{R} . Les fonctions rationnelles (quotient de fonctions polynomiales) sont continues sur chaque intervalle de leur domaine de définition.
- Les fonctions usuelles $x \mapsto \sin(x)$, $x \mapsto \cos(x)$, $x \mapsto \tan(x)$, $x \mapsto \exp(x)$, $x \mapsto \ln(x)$ et $x \mapsto x^\alpha$ sont continues sur chaque intervalle de leur domaine.

Proposition 7.4.1 (opérations entre fonctions continues sur un intervalle)

Soit f et g deux fonctions continues sur I .

Alors $\alpha f + \beta g$ (α, β dans \mathbb{R}), fg , $\inf(f, g)$ et $\sup(f, g)$ sont continues sur I .

Proposition 7.4.2 (compositions de fonctions continues sur un intervalle)

Si $f: I \rightarrow \mathbb{R}$ et $g: J \rightarrow \mathbb{R}$ sont continues, avec $f(I) \subset J$, alors $g \circ f$ est continue sur I .

Remarques

Pour démontrer qu'une fonction est continue sur un intervalle I , on ne revient pratiquement jamais à la définition « epsilonlesque ». Le plus souvent, la fonction à étudier est en effet un « cocktail » de fonctions continues usuelles et les propriétés précédentes permettent de conclure.

La continuité, même sur un intervalle, reste une *propriété locale*, ce qui signifie qu'elle n'est que le bilan de la continuité de f en chacun des points de I .

7.4.2 Théorème des valeurs intermédiaires**Proposition 7.4.3** (propriété des valeurs intermédiaires)

Soit f une fonction continue sur l'intervalle I .

Soit a et b deux éléments de I , et soit β un réel compris entre $f(a)$ et $f(b)$.

Alors il existe un réel α , compris entre a et b , tel que $f(\alpha) = \beta$.

Proposition 7.4.4 (un énoncé équivalent au précédent)

Soit f une fonction continue sur l'intervalle I , à valeurs réelles.

On suppose qu'il existe a et b dans I tels que $f(a) \leq 0$ et $f(b) \geq 0$.

Alors il existe c dans I , compris entre a et b , tel que $f(c) = 0$.

Proposition 7.4.5 (un énoncé équivalent aux deux précédents)

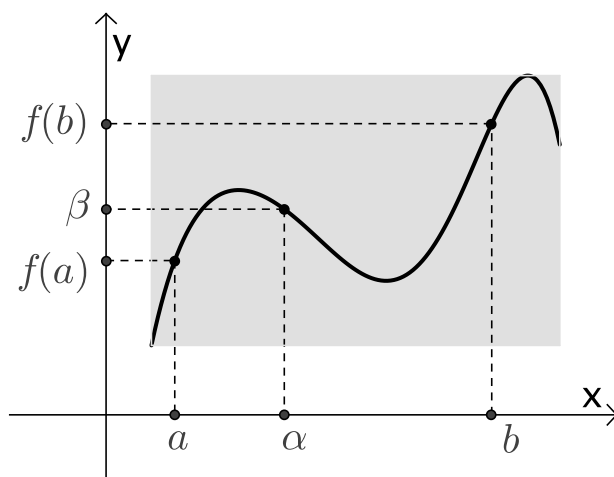
Soit f une fonction continue sur l'intervalle I , à valeurs réelles. Alors $f(I)$ est un intervalle.

On représente ici une situation typique, la fonction f étant continue sur un intervalle I .

On a choisi deux abscisses a et b de I , et une ordonnée β comprise entre $f(a)$ et $f(b)$.

On a grisé le rectangle délimitant le graphe de f : sa projection sur Ox est l'intervalle I , et sa projection sur Oy est l'image $J = f(I)$.

On a marqué une abscisse α comprise entre a et b , telle que $f(\alpha) = \beta$ (en fait il y avait ici trois solutions α possibles, ce qui est manifestement une conséquence de la non monotonie de f).

**7.4.3 Approximation d'un zéro par dichotomie**

Soit f une fonction définie et continue sur un segment $[a, b]$ (avec $a < b$) à valeurs réelles.

On suppose que $f(a)f(b) \leq 0$.

D'après le théorème des valeurs intermédiaires, il existe α dans $[a, b]$ tel que $f(\alpha) = 0$.

La méthode d'approximation de α par dichotomie consiste en les étapes suivantes :

- Initialisation : on pose $a_0 = a$ et $b_0 = b$.
- Soit n dans \mathbb{N} . On suppose a_n et b_n connus, avec $f(a_n)f(b_n) \leq 0$. On pose $c_n = \frac{a_n + b_n}{2}$.
Si $f(a_n)f(c_n) \leq 0$, on pose $\begin{cases} a_{n+1} = a_n \\ b_{n+1} = c_n \end{cases}$ sinon on pose $\begin{cases} a_{n+1} = c_n \\ b_{n+1} = b_n \end{cases}$
- On définit ainsi deux suites adjacentes $(a_n)_{n \geq 0}, (b_n)_{n \geq 0}$.
Ces deux suites convergent vers une solution α de l'équation $f(x) = 0$.
- Dans la pratique, on se donne $\varepsilon > 0$, et on s'arrête dès que $0 \leq b_n - a_n \leq \varepsilon$.
On a alors obtenu un encadrement de α avec une erreur absolue majorée par ε .

On va écrire une fonction Python qui renvoie l'intervalle final de cette dichotomie.

Pour ε , on utilise une valeur par défaut : $\varepsilon = 10^{-8}$.

Voici le listing, et un exemple d'utilisation.

On obtient ainsi un encadrement de chacune des trois racines de $f(x) = x^3 - 2x^2 - x + 1$.

```
def dichotomie(f, a, b, eps=1e-8):
    while abs(b-a) > eps:
        c = (a+b)/2
        if f(a)*f(c) <= 0:
            b=c
        else:
            a=c
    return [a,b]
```

```
>>> def f(x): return x*x*x-2*x*x-x+1
>>> dichotomie(f, -1, 0)
[-0.8019377365708351, -0.8019377291202545]
>>> dichotomie(f, 0, 1)
[0.5549581274390221, 0.5549581348896027]
>>> dichotomie(f, 2, 3)
[2.2469796016812325, 2.246979609131813]
```

7.4.4 Application continue sur un segment

Les deux énoncés qui suivent sont équivalents :

Proposition 7.4.6 (présence d'un maximum et d'un minimum)

Soit f une fonction continue sur le segment $[a, b]$, à valeurs réelles.

Alors f est bornée sur $[a, b]$, et elle atteint ses bornes.

Il existe donc x_0 dans $[a, b]$ tel que $f(x_0) = \min\{f(x), a \leq x \leq b\}$.

De même, il existe x_1 dans $[a, b]$ tel que $f(x_1) = \max\{f(x), a \leq x \leq b\}$.

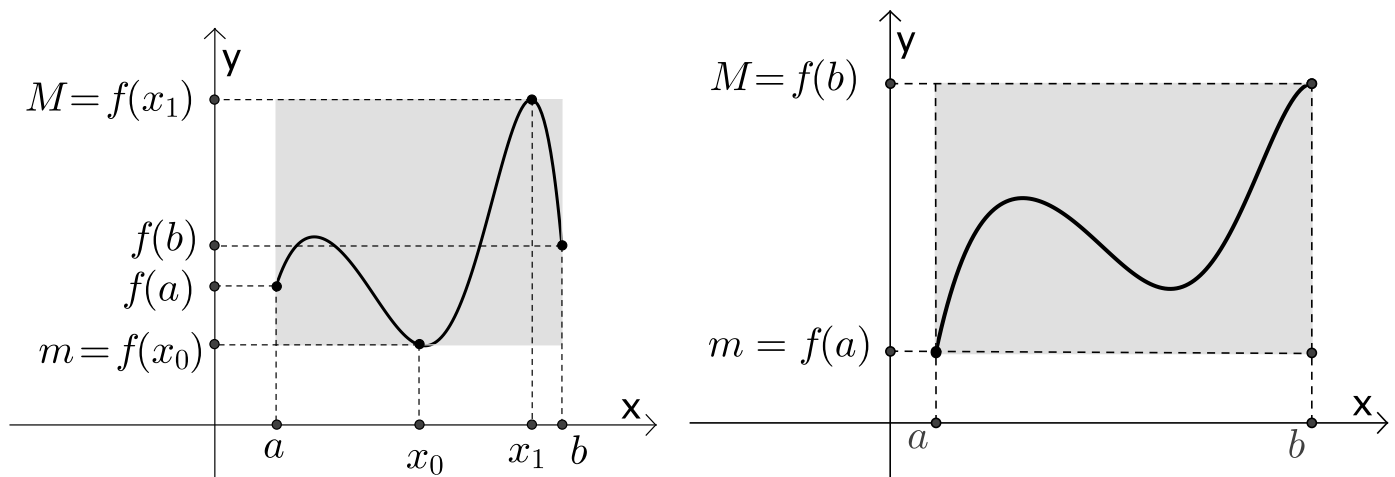
Proposition 7.4.7 (image d'un segment par une fonction continue)

Soit f une fonction continue sur le segment $[a, b]$, à valeurs réelles.

Alors $f([a, b])$ est un segment $[m, M]$ de \mathbb{R} .

On a représenté ici deux situations typiques. La fonction choisie f n'est pas monotone sur $[a, b]$.

Les abscisses où f atteint ses bornes ne sont pas nécessairement les extrémités de $[a, b]$ (figure de gauche), même si ça reste possible (indépendamment de la non monotonie de f , voir figure de droite).



7.4.5 Continuité et stricte monotonie sur un intervalle

Proposition 7.4.8 (continuité et injectivité \Rightarrow stricte monotonie)

Soit $f : I \rightarrow \mathbb{R}$, continue et injective sur I . Alors f est strictement monotone sur I .

Proposition 7.4.9 (une condition suffisante de continuité sur un intervalle)

Soit I un intervalle, et soit $f : I \rightarrow \mathbb{R}$ une fonction monotone.

On suppose que $f(I)$ est un intervalle. Alors f est continue sur l'intervalle I .

Proposition 7.4.10 (théorème de la bijection réciproque)

Soit $f : I \rightarrow \mathbb{R}$, continue et strictement monotone sur I .

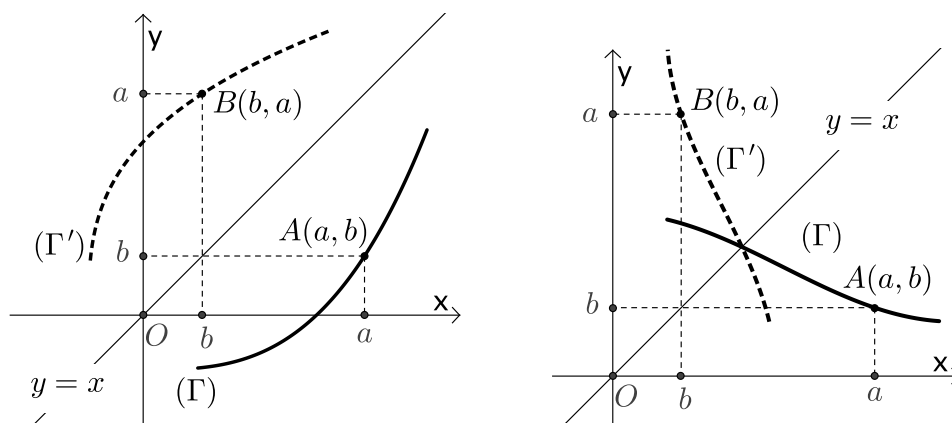
Alors f réalise une bijection de I sur l'intervalle image $J = f(I)$.

La bijection réciproque $f^{-1} : J \rightarrow I$ est continue et strictement monotone (de même monotonie que f).

Symétrie des courbes représentatives

Les courbes représentatives (Γ) de f et (Γ') de f^{-1} sont symétriques l'une de l'autre dans la symétrie par rapport à la droite $y = x$, parallèlement à la droite $y = -x$ (si le repère est orthonormé, il s'agit de la symétrie orthogonale par rapport à la droite $y = x$).

On a représenté ici deux situations possibles (le graphe (Γ') de f^{-1} est dessiné en traits pointillés). Dans le premier cas f et f^{-1} sont strictement croissantes. Dans le deuxième cas, elles sont strictement décroissantes (et on a fait figurer une valeur x telle que $f(x) = x$ donc telle que $f^{-1}(x) = x$: les deux graphes traversent alors simultanément la première bissectrice).



Unicité de la racine à une équation $f(x) = 0$

Le théorème des valeurs intermédiaires montre l'existence d'une solution pour $f(x) = 0$.

Le théorème de la bijection réciproque assure l'unicité de cette solution.

Conservation des caractéristiques de l'intervalle

Si f est continue sur I , alors $J = f(I)$ n'a pas nécessairement les mêmes caractéristiques que I (caractère ouvert ou fermé, borné ou non borné).

En revanche, on sait que si I est un segment, alors $f(I)$ est un segment.

On peut ajouter que si f est strictement monotone, alors le caractère ouvert, semi-ouvert, ou fermé est conservé quand on passe de l'intervalle I à l'intervalle J .

Rappel : exemples de réciproques d'applications continues

- L'application $x \mapsto \exp(x)$ est une bijection de \mathbb{R} sur \mathbb{R}^{+*} .
La bijection réciproque est $x \mapsto \ln(x)$.
- Pour tout α de \mathbb{R}^{+*} , les applications $x \mapsto x^\alpha$ et $x \mapsto x^{1/\alpha}$ sont deux bijections de \mathbb{R}^{+*} sur lui-même, réciproques l'une de l'autre.
- L'application $x \mapsto \sin(x)$ réalise une bijection de $[-\frac{\pi}{2}, \frac{\pi}{2}]$ sur $[-1, 1]$.
La bijection réciproque est notée $x \mapsto \arcsin(x)$ (*arc sinus de x*).
- L'application $x \mapsto \cos(x)$ réalise une bijection de $[0, \pi]$ sur $[-1, 1]$.
La bijection réciproque est notée $x \mapsto \arccos(x)$ (*arc cosinus de x*).
- L'application $x \mapsto \tan(x)$ réalise une bijection de $]-\frac{\pi}{2}, \frac{\pi}{2}[$ sur \mathbb{R} .
La bijection réciproque est notée $x \mapsto \arctan(x)$ (*arc tangente de x*).

7.5 Cas des fonctions continues complexes

Soit f une fonction complexe $f : I \rightarrow \mathbb{C}$, définie sur l'intervalle I .

On sait que f est caractérisée par $g : I \rightarrow \mathbb{R}$ et $h : I \rightarrow \mathbb{R}$ telles : $\forall x \in I, f(x) = g(x) + ih(x)$.

Ces deux fonctions réelles sont notées $g = \operatorname{Re}(f)$ et $h = \operatorname{Im}(f)$.

7.5.1 Limite d'une fonction à valeurs complexes

Voici les trois définitions qu'on peut donner pour exprimer qu'une fonction à valeurs complexes possède une limite en un point a (avec a réel), ou en $\pm\infty$.

Les trois définitions suivantes sont calquées sur celles qui ont été données pour des fonctions à valeurs réelles. La seule différence est que la distance $|f(x) - \ell|$ est mesurée dans \mathbb{C} , avec un module et non plus une valeur absolue.

Définition 7.5.1 (limite finie en un point a de \mathbb{R})

Soit I un intervalle, et soit $f : I \rightarrow \mathbb{C}$ une fonction à valeurs complexes.

Soit a un réel, élément ou extrémité de I . Soit ℓ un nombre complexe.

On dit que ℓ est limite de f en a si : $\forall \varepsilon > 0, \exists \delta > 0, (x \in I \text{ et } |x - a| \leq \delta) \Rightarrow |f(x) - \ell| \leq \varepsilon$.

Définition 7.5.2 (limite finie en $+\infty$)

Soit $f : I \rightarrow \mathbb{C}$ une fonction complexe, où I est un intervalle non majoré. Soit ℓ un nombre complexe.

On dit que ℓ est limite de f en $+\infty$ si : $\forall \varepsilon > 0, \exists A \in \mathbb{R}, (x \geq A \Rightarrow |f(x) - \ell| \leq \varepsilon)$.

Définition 7.5.3 (limite finie en $-\infty$)

Soit $f : I \rightarrow \mathbb{C}$ une fonction complexe, où I est un intervalle non minoré. Soit ℓ un nombre complexe.

On dit que ℓ est limite de f en $-\infty$ si : $\forall \varepsilon > 0, \exists A \in \mathbb{R}, (x \leq A \Rightarrow |f(x) - \ell| \leq \varepsilon)$.

Les définitions portant sur les fonctions réelles ne sont pas toutes transposables aux fonctions complexes.

Par exemple, si $f : I \rightarrow \mathbb{C}$, cela n'a aucun sens d'écrire que f tend vers $+\infty$ ou $-\infty$.

De même, on ne peut plus parler de limite « par valeurs supérieure » (ou inférieure).

En revanche, on peut encore parler de la limite à gauche et de la limite à droite en a .

Proposition 7.5.1 (caractérisation en termes de partie réelle et partie imaginaire)

Soit $f : I \rightarrow \mathbb{C}$ une fonction complexe. Soit $g = \operatorname{Re}(f)$ et $h = \operatorname{Im}(f)$.

Soit a un réel, élément ou extrémité de I .

Soit $\ell = u + iv$ un nombre complexe, avec u, v dans \mathbb{R} .

Alors on a l'équivalence : $\lim_{x \rightarrow a} f(x) = \ell \Leftrightarrow \left(\lim_{x \rightarrow a} g(x) = u \text{ et } \lim_{x \rightarrow a} h(x) = v \right)$

Dans ce cas, on peut donc écrire : $\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} g(x) + i \lim_{x \rightarrow a} h(x)$.

Le résultat précédent s'étend immédiatement à $\lim_{x \rightarrow +\infty} f(x) = \ell$ et $\lim_{x \rightarrow -\infty} f(x) = \ell$, avec ℓ dans \mathbb{C} .

Résultats qui s'étendent aux limites de fonctions complexes

Un certain nombre de résultats concernant les limites de fonctions à valeurs réelles s'étendent sans difficulté au cas des fonctions à valeurs complexes. Citons entre autres :

- L'unicité de la limite (cf proposition 7.1.1)
- Le fait que si f est définie en a , sa seule limite possible en a est $f(a)$ (cf proposition 7.1.2)
- La caractérisation séquentielle des limites (cf proposition 7.1.5)
- Les opérations sur les limites (cf les propositions 7.2.2, 7.2.3, 7.2.6)

En revanche, on ne peut pas généraliser aux fonctions complexes les propriétés des limites réelles quand elles ont un rapport avec les inégalités et la monotonie (cf sous-sections 7.2.2 et 7.2.3)

7.5.2 Continuité en un point d'une fonction complexe

La définition suivante reprend mot pour mot celle qui a été donnée pour des fonctions à valeurs réelles. La seule différence est que la distance $|f(x) - f(a)|$ utilise maintenant le module dans \mathbb{C} .

Définition 7.5.4 (définition de la continuité d'une fonction complexe en un point)

Soient $f : I \rightarrow \mathbb{C}$ une fonction complexe, et soit a un élément de I .

On dit que f est *continue* en a si la limite de f en a existe.

Puisque f est définie en a , cela équivaut à dire : $\lim_{x \rightarrow a} f(x) = f(a)$

Donc f est continue en a si et seulement si : $\forall \varepsilon > 0, \exists \delta > 0, (x \in I \text{ et } |x - a| \leq \delta) \Rightarrow |f(x) - f(a)| \leq \varepsilon$

On a immédiatement la caractérisation par la partie réelle et la partie imaginaire :

Proposition 7.5.2 (caractérisation en termes de partie réelle et partie imaginaire)

Soit $f : I \rightarrow \mathbb{C}$ une fonction complexe. Soit $g = \operatorname{Re}(f)$ et $h = \operatorname{Im}(f)$.

Soit a un élément de I . Alors f est continue en a si et seulement si g et h sont continues en a .

Un certain nombre de résultats concernant la continuité de fonctions à valeurs réelles s'étendent sans difficulté au cas des fonctions à valeurs complexes. Citons entre autres :

- La définition de la continuité à gauche ou à droite (cf définitions 7.3.3 et 7.3.4).
- La caractérisation séquentielle de la continuité en un point (cf proposition 7.3.1)
- Les opérations sur les fonctions continues (cf propositions 7.3.2 et 7.3.3).

7.5.3 Continuité d'une fonction complexe sur un intervalle

Sans surprise, on a la définition suivante, puis une caractérisation avec les parties réelle et imaginaire.

Définition 7.5.5

Soit f une fonction numérique à valeurs complexes, définie sur l'intervalle I .

On dit que f est continue sur I si f est continue en tout point de I .

On note $\mathcal{C}(I, \mathbb{C})$ l'ensemble des fonctions continues sur I , à valeurs complexes.

Proposition 7.5.3 (caractérisation en termes de partie réelle et partie imaginaire)

Soit $f : I \rightarrow \mathbb{C}$ une fonction complexe. Soit $g = \operatorname{Re}(f)$ et $h = \operatorname{Im}(f)$.

Alors f est continue sur I si et seulement si g et h sont continues sur I .

Voici des résultats concernant la continuité sur un intervalle, et qui généralisent le cas des fonctions à valeur réelles au cas des fonctions à valeurs complexes :

- Opérations entre fonctions continues sur un intervalle (cf proposition 7.4.1).
- Compositions de fonctions continues sur un intervalle (cf proposition 7.4.2).

Si $f : I \rightarrow \mathbb{R}$ et $g : J \rightarrow \mathbb{C}$ sont continues, avec $f(I) \subset J$, alors $g \circ f$ est continue sur I .

En revanche, on notera bien les différences suivantes.

Pour des fonctions continues à valeurs complexes :

- Il n'y a plus de théorème des valeurs intermédiaires.

Considérons par exemple la fonction $f : x \mapsto e^{i\pi x}$.

Elle est continue sur $[0, 1]$, on a $f(0) = 1$, $f(1) = -1$, mais f ne s'annule jamais (faire un dessin)

- Il n'y a plus de théorème de la bijection réciproque (tout simplement parce que parler de la monotonie de $f : I \rightarrow \mathbb{C}$ n'a pas de signification).

On notera tout de même le résultat suivant, dans le cas où f est continue sur un *segment* :

Proposition 7.5.4

Soit f une fonction continue sur le segment $[a, b]$, à valeurs complexes.

Alors f est bornée sur $[a, b]$, et son module atteint ses bornes.

Il existe donc x_0 dans $[a, b]$ tel que $|f(x_0)| = \min\{|f(x)|, a \leq x \leq b\}$.

De même, il existe x_1 dans $[a, b]$ tel que $|f(x_1)| = \max\{|f(x)|, a \leq x \leq b\}$.

Ce résultat signifie que l'arc $t \mapsto f(t) = u(t) + iv(t)$, avec $0 \leq t \leq b$ est tout entier compris dans la couronne circulaire définie par $m \leq |f(z)| \leq M$, où m et M sont respectivement le minimum et le maximum de $|f|$.

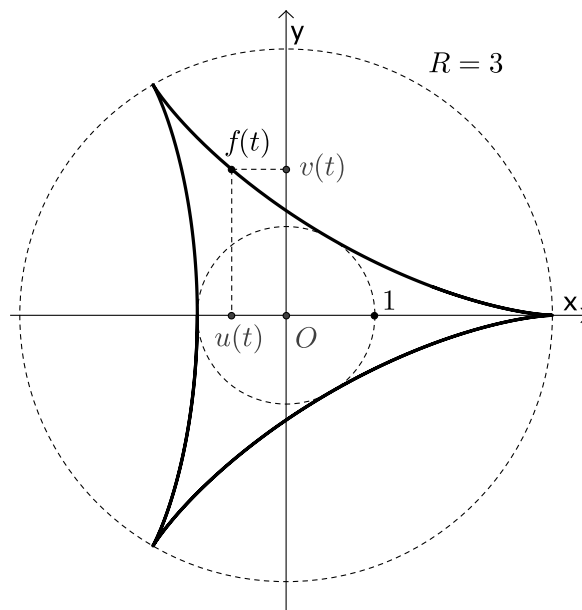
Le fait que ces bornes soient atteintes signifie que l'arc représentant f rencontre les deux cercles qui délimitent cette couronne.

On a représenté ici l'arc défini par $f(t) = 2e^{it} + e^{-2it}$.

$$\text{On a ici } \begin{cases} u(t) = \operatorname{Re}(f(t)) = 2 \cos(t) + \cos(2t) \\ v(t) = \operatorname{Im}(f(t)) = 2 \sin(t) - \sin(2t) \end{cases}$$

Le mouvement de $f(t)$ est périodique de période 2π .

L'arc est ici tout entier inclus dans la couronne circulaire de centre 0 défini par $1 \leq |z| \leq 3$. Sur chaque période, il y a trois points de contact avec le cercle intérieur $|z| = 1$, et trois points de contact avec le cercle extérieur $|z| = 3$.



Chapitre 8

Dérivabilité

Sommaire

8.1	Nombre dérivé, fonction dérivée	189
8.1.1	Dérivabilité en un point, nombre dérivé	189
8.1.2	Notion de développement limité d'ordre 1	190
8.1.3	Dérivabilité à gauche, à droite	191
8.1.4	Dérivabilité sur un intervalle	192
8.1.5	Opérations sur les fonctions dérivables	192
8.1.6	Approximation d'un zéro par la méthode de Newton	194
8.2	Rolle et accroissements finis	194
8.2.1	Extremum local et point critique	194
8.2.2	Théorème de Rolle	195
8.2.3	Égalité des accroissements finis	196
8.2.4	Inégalité des accroissements finis	197
8.2.5	Dérivabilité et sens de variation	199
8.2.6	Théorème de la limite de la dérivée	200
8.3	Fonctions de classe \mathcal{C}^k	201
8.3.1	Fonctions de classe \mathcal{C}^k sur un intervalle	201
8.3.2	Opérations sur les fonctions de classe \mathcal{C}^k	201
8.3.3	Théorème de classe \mathcal{C}^k par prolongement	202
8.4	Extension aux fonctions complexes	203
8.4.1	Fonctions de classe \mathcal{C}^k à valeurs dans \mathbb{C}	203
8.4.2	Extension des résultats	203

8.1 Nombre dérivé, fonction dérivée

8.1.1 Dérivabilité en un point, nombre dérivé

Dans tout ce chapitre, on considère des fonctions qui sont définies sur un intervalle I de \mathbb{R} non réduit à un point, et qui sont à valeurs dans \mathbb{R} .

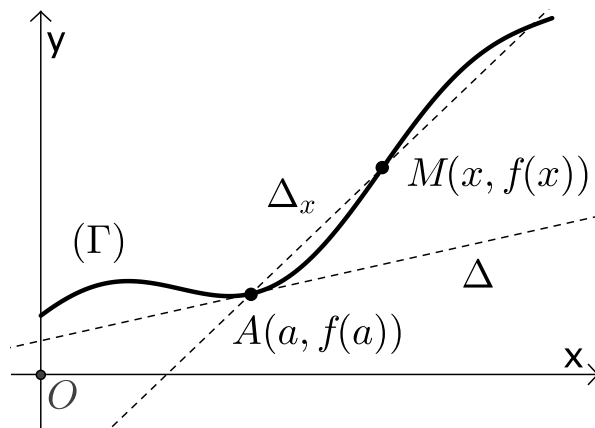
On se place au voisinage d'un point $A(a, f(a))$ de la courbe représentative (Γ) de f .

Soit $M(x, f(x))$ un point mobile sur (Γ) , avec $x \neq a$.
Soit Δ_x la droite passant par A et M .

Le coefficient directeur de Δ_x est $\delta_x = \frac{f(x) - f(a)}{x - a}$.

Quand x tend vers a (donc ici quand M se rapproche de A sur (Γ)), on examine si la droite Δ_x (qui pivote autour de A) possède une position limite Δ (c'est-à-dire si δ_x possède une valeur limite).

Si tel est le cas, on dit que la droite Δ est la tangente en A à la représentation graphique (Γ) .



Définition 8.1.1 (nombre dérivé en un point)

Soit $f : I \rightarrow \mathbb{R}$ une fonction numérique réelle. Soit a un élément de I .

On dit que f est *dérivable* en a si $\lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a}$ existe dans \mathbb{R} .

Cette limite est appelée *nombre dérivé* de f en a et est notée $f'(a)$, ou $D(f)(a)$, ou $\frac{df}{dx}(a)$.

Interprétation géométrique

Dire que f est dérivable en a , c'est dire que la courbe représentative (Γ) de f présente au point $A(a, f(a))$ une tangente Δ **non verticale**.

Par définition, le coefficient directeur de Δ est $f'(a)$.

L'équation de Δ est donc $y = f(a) + (x - a)f'(a)$.

Proposition 8.1.1 (une autre définition de la dérivabilité)

Soit $f : I \rightarrow \mathbb{R}$ une fonction numérique réelle. Soit a un élément de I .

f est dérivable en a si et seulement si il existe :

- un réel ℓ
- une fonction $x \mapsto \varepsilon(x)$ de I dans \mathbb{R} , vérifiant $\lim_{x \rightarrow a} \varepsilon(x) = 0$ et $\varepsilon(a) = 0$

tels que : $\forall x \in I, f(x) = f(a) + (x - a)\ell + (x - a)\varepsilon(x)$

Au sens de la définition précédente, le réel ℓ est le nombre dérivé $f'(a)$.

Représentons la courbe (Γ) de f au voisinage de $A(a, f(a))$, et la tangente Δ (non verticale) en A .

Plaçons également deux points M_0 et M_1 sur (Γ) d'abscisses respectives x_0 et x_1 .

On sait que l'équation de la tangente en A à (Γ) est : $y = f(a) + (x - a)f'(a)$.

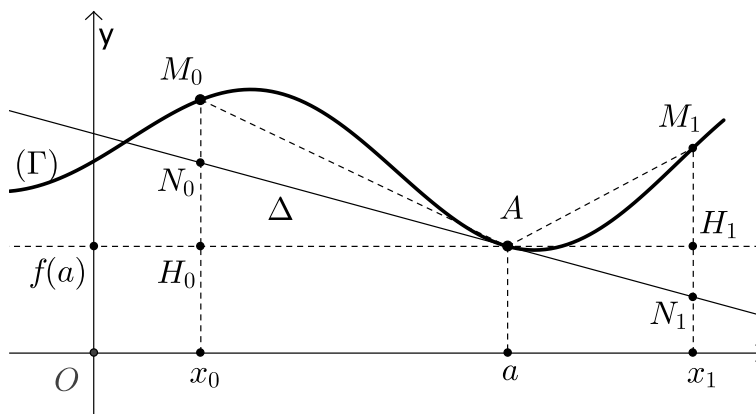
Dans l'illustration ci-dessous, on a les coordonnées :

- d'une part : $H_0(x_0, f(a))$, $N_0(x_0, f(a) + (x_0 - a)f'(a))$ et $M_0(x_0, f(x_0))$.
- d'autre part : $H_1(x_1, f(a))$, $N_1(x_1, f(a) + (x_1 - a)f'(a))$ et $M_1(x_1, f(x_1))$.

Si on parle en termes de mesures algébriques sur un axe vertical (dirigé vers le haut), on a :

$$\begin{cases} \overline{H_0N_0} = (x_0 - a)f'(a) \\ \overline{H_1N_1} = (x_1 - a)f'(a) \end{cases} \quad \begin{cases} \overline{N_0M_0} = f(x_0) - f(a) - (x_0 - a)f'(a) \\ \overline{N_1M_1} = f(x_1) - f(a) - (x_1 - a)f'(a) \end{cases} \quad \text{et} \quad \begin{cases} \overline{H_0M_0} = f(x_0) - f(a) \\ \overline{H_1M_1} = f(x_1) - f(a) \end{cases}$$

Avec les notations de la proposition 8.1.1, on a donc $\begin{cases} \overline{N_0M_0} = (x_0 - a)\varepsilon(x_0) \\ \overline{N_1M_1} = (x_1 - a)\varepsilon(x_1) \end{cases}$



Ce que « dit » la dérivabilité de f en a , c'est qu'au voisinage « immédiat » de A une mesure algébrique comme $\overline{N_0M_0}$ est « négligeable » devant $\overline{H_0N_0}$ (ça ne saute pas aux yeux sur l'illustration car on s'est placé finalement très loin de a).

En termes imagés, la dérivabilité de f en a nous dit que la différence $f(x) - f(a)$ (donc l'accroissement de f entre a et x) est approché par $f'(a)(x - a)$ (donc proportionnellement à l'accroissement $x - a$ de la variable), avec un terme d'erreur qui est lui-même négligeable devant $x - a$.

En termes plus géométriques (mais pas très précis) la tangente Δ à (Γ) en $A(a, f(a))$ est une « bonne approximation » de la courbe au voisinage de A .

Proposition 8.1.2

Soit $f : I \rightarrow \mathbb{R}$ une fonction numérique, et soit a un point de I .

Si f est dérivable en a , alors f est continue en a .

La réciproque est fautive : penser à $x \mapsto |x|$ en 0.

8.1.2 Notion de développement limité d'ordre 1

Soit $f : I \rightarrow \mathbb{R}$ une fonction numérique, et soit a un point de I .

Si f est dérivable en a alors : $f(x) = f(a) + f'(a)(x - a) + (x - a)\varepsilon(x)$, avec $\lim_{x \rightarrow a} \varepsilon(x) = 0$.

Réciproquement, on suppose que a est une extrémité de l'intervalle de définition de f .

On suppose également qu'on peut écrire $f(x) = \lambda_0 + \lambda_1(x - a) + (x - a)\varepsilon(x)$, avec $\lim_{x \rightarrow a} \varepsilon(x) = 0$.

On trouve $\lim_{x \rightarrow a} f(x) = \lambda_0$, ce qui prouve que f est prolongeable par continuité en a , en posant $f(a) = \lambda_0$.

On trouve ensuite $\frac{f(x) - f(a)}{x - a} = \lambda_1 + \varepsilon(x)$, donc $\lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} = \lambda_1$.

Cela prouve que f est dérivable en a , avec $f'(a) = \lambda_1$.

On peut donc énoncer le résultat suivant, qui sera amélioré ultérieurement :

Proposition 8.1.3

Soit f une fonction numérique, définie sur $I \setminus \{a\}$.

On suppose que f possède un « développement limité d'ordre 1 en a ».

Autrement dit, on suppose qu'on peut écrire : $f(x) = \lambda_0 + \lambda_1(x - a) + (x - a)\varepsilon(x)$, avec $\lim_{x \rightarrow a} \varepsilon(x) = 0$.

Alors f est prolongeable de façon dérivable en a , en posant $f(a) = \lambda_0$ et $f'(a) = \lambda_1$.

Variante de notation

Le développement limité $f(x) = \lambda_0 + \lambda_1(x - a) + (x - a)\varepsilon(x)$ s'écrit : $f(x) = \lambda_0 + \lambda_1(x - a) + o(x - a)$.

Dans cette écriture, $o(x - a)$ est une notation générique désignant toute fonction « négligeable devant » $x - a$ quand x tend vers a , c'est-à-dire qui s'écrit $(x - a)\varepsilon(x)$, avec $\lim_{x \rightarrow a} \varepsilon(x) = 0$.

Le changement de variable $x = a + h$ permet de se ramener à une étude quand h tend vers 0.

On écrira donc $f(a + h) = \lambda_0 + \lambda_1 h + o(h)$ pour exprimer le développement limité d'ordre 1 de f en a .

8.1.3 Dérivabilité à gauche, à droite

On complète les définitions précédentes avec la notion de nombre dérivé à gauche ou à droite.

Définition 8.1.2 (nombre dérivé à gauche)

Soit $f : I \rightarrow \mathbb{R}$ une fonction numérique, et soit a un point de I (distinct de l'extrémité gauche de I).

On dit que f est dérivable à gauche en a si $\lim_{x \rightarrow a^-} \frac{f(x) - f(a)}{x - a}$ existe dans \mathbb{R} .

Cette limite est appelée *nombre dérivé à gauche* de f en a , et est notée $f'_g(a)$.

Définition 8.1.3 (nombre dérivé à droite)

Soit $f : I \rightarrow \mathbb{R}$ une fonction numérique, et soit a un point de I (distinct de l'extrémité droite de I).

On dit que f est dérivable à droite en a si $\lim_{x \rightarrow a^+} \frac{f(x) - f(a)}{x - a}$ existe dans \mathbb{R} .

Cette limite est appelée *nombre dérivé à droite* de f en a , et est notée $f'_d(a)$.

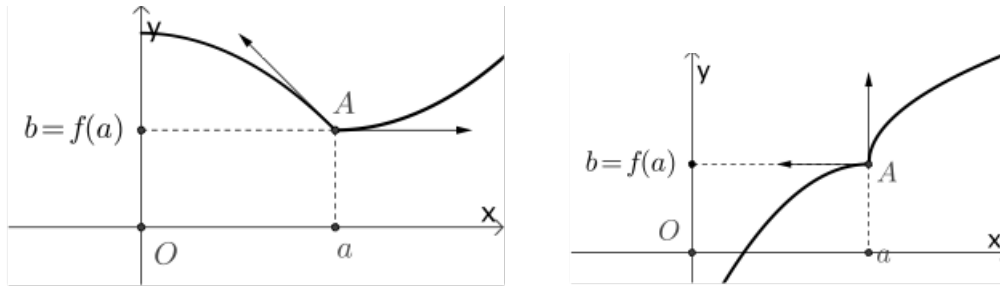
Interprétation géométrique

Dire que f est dérivable à droite (resp. à gauche) en a , c'est dire que la courbe (Γ) de f admet au point $A(a, f(a))$ une demi-tangente à droite (resp. à gauche) **non verticale**.

Le coefficient directeur de cette demi-tangente est $f'_d(a)$ (resp. $f'_g(a)$).

Sur l'exemple de gauche, f est dérivable à gauche et à droite en a , avec $f'_g(a) = -1$ (demi-tangente oblique, parallèle à $y = -x$) et $f'_d(a) = 0$ (demi-tangente horizontale.)

Sur l'exemple de droite, on a $f'_g(a) = 0$ (demi-tangente horizontale), mais f n'est pas dérivable à droite en a (il y a bien une demi-tangente mais elle est verticale).

**Remarques :**

— Soit a un point de I qui n'est pas une extrémité de I .

Alors f est dérivable en a si seulement si $\begin{cases} \text{elle est dérivable à gauche et à droite en } a \\ \text{on a l'égalité } f'_g(a) = f'_d(a) \end{cases}$
On a alors $f'(a) = f'_g(a) = f'_d(a)$.

— Si f est dérivable à gauche (resp. à droite) en a , elle y est continue à gauche (resp. à droite).

8.1.4 Dérivabilité sur un intervalle**Définition 8.1.4** (fonctions dérivables, ou de classe \mathcal{C}^1 , sur I)

On dit que f est dérivable sur l'intervalle I si f est dérivable en tout point de I .

La fonction $f' : I \rightarrow \mathbb{R}$ qui à tout a associe $f'(a)$ est appelée *fonction dérivée* de f .

Cette fonction est également notée $D(f)$ ou $\frac{df}{dx}$.

Si de plus f' est continue sur I , on dit que f est de classe \mathcal{C}^1 sur I .

On note $\mathcal{D}(I, \mathbb{R})$ (resp. $\mathcal{C}^1(I, \mathbb{R})$) l'ensemble des fonctions dérivables (resp. de classe \mathcal{C}^1) de I dans \mathbb{R} .

Remarque : on se souviendra que la dérivabilité (même étendue à un intervalle) reste une *notion locale* : la dérivabilité de f sur I , ça n'est que la dérivabilité de f en chacun des points de I .

8.1.5 Opérations sur les fonctions dérivables**Proposition 8.1.4** (linéarité de la dérivation en un point)

Soient f et g deux fonctions dérivables au point a .

Pour tous scalaires α, β , la fonction $h = \alpha f + \beta g$ est dérivable en a , et $h'(a) = \alpha f'(a) + \beta g'(a)$.

Proposition 8.1.5 (produit de fonctions dérivables en un point)

Soient f et g deux fonctions dérivables en un point a .

Alors la fonction $h = fg$ est dérivable en a , et $h'(a) = f'(a)g(a) + f(a)g'(a)$.

Proposition 8.1.6 (dérivée de l'inverse)

Si g est dérivable en a , avec $g(a) \neq 0$, alors $h = \frac{1}{g}$ est dérivable en a , et $h'(a) = -\frac{g'(a)}{g^2(a)}$.

Supposons en outre que f soit dérivable en a .

Alors le quotient $\frac{f}{g}$ est dérivable au point a , et : $\left(\frac{f}{g}\right)'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{g^2(a)}$.

Proposition 8.1.7 (composition et dérivation)

Soit $f : I \rightarrow \mathbb{R}$, dérivable en a . Soit $g : J \rightarrow \mathbb{R}$, dérivable en $b = f(a)$.

Alors $g \circ f$ est dérivable en a et $(g \circ f)'(a) = f'(a)(g' \circ f)(a)$.

Remarque : dans la proposition précédente, il faut supposer que $g \circ f$ est définie au voisinage de a (par exemple, si $f(I) \subset J$, alors $g \circ f$ est définie sur I).

Proposition 8.1.8 (dérivation et bijection réciproque)

Soit $f : I \rightarrow \mathbb{R}$ une fonction dérivable, strictement monotone.

f est donc bijective de I sur un intervalle J . Soit a dans I tel que $f'(a) \neq 0$.

Alors $g = f^{-1}$ est dérivable en $b = f(a)$ et $g'(b) = \frac{1}{f'(a)} = \frac{1}{f' \circ f^{-1}(b)}$.

Généralisation aux fonctions dérivables sur un intervalle

Les propriétés précédentes se généralisent à la dérivabilité sur un intervalle. On peut même d'ailleurs énoncer les mêmes résultats en remplaçant « dérivable » par « de classe \mathcal{C}^1 »).

– Soient f et g deux fonctions dérivables sur l'intervalle I .

Alors la combinaison linéaire $\alpha f + \beta g$ et le produit fg sont dérivables sur I et $\begin{cases} h' = \alpha f' + \beta g' \\ (fg)' = f'g + fg' \end{cases}$

Si g ne s'annule pas sur I , alors $\left(\frac{1}{g}\right)' = -\frac{g'}{g^2}$, et $\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}$.

– Soit $f : I \rightarrow \mathbb{R}$ et $g : J \rightarrow \mathbb{R}$ deux fonctions dérivables, avec $f(I) \subset J$.

Alors $g \circ f$ est dérivable sur I et $(g \circ f)' = f' \cdot (g' \circ f)$

– Soit $f : I \rightarrow \mathbb{R}$ une fonction dérivable, strictement monotone.

La fonction f réalise donc une bijection de I sur un intervalle J .

Si f' ne s'annule pas sur I , alors $g = f^{-1}$ est dérivable sur J et $g' = \frac{1}{f' \circ f^{-1}}$.

Cas usuels de dérivation composée

La dérivée de la fonction puissance $x \mapsto x^\alpha$ est $x \mapsto \alpha x^{\alpha-1}$ (c'est vrai sur \mathbb{R}^{+*} pour tous les α ; c'est prolongeable en 0 si $\alpha > 1$; c'est prolongeable sur \mathbb{R}^{-*} pour les exposants entiers).

Dans les égalités suivantes, f est dérivable, et on suppose qu'on a réglé les problèmes de définition.

$$(f^\alpha)' = \alpha f' f^{\alpha-1} \quad (\alpha \text{ constant !}), \quad (\sqrt{f})' = \frac{f'}{2\sqrt{f}}, \quad (\ln|f|)' = \frac{f'}{f}, \quad (e^f)' = f' e^f$$

Si f, g sont dérivables : $(fg)' = (e^{g \ln(f)})' = (g \ln(f))' f^g = \left(g' \ln(f) + \frac{gf'}{f}\right) f^g = g' \ln(f) f^g + g f' f^{g-1}$

Rappel sur la dérivation des fonctions trigonométriques inverses

La dérivée de $x \rightarrow \sin x$ sur $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ est $x \rightarrow \cos x$, nulle en $\pm \frac{\pi}{2}$. On en déduit :

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

La dérivée de $x \rightarrow \cos x$ sur $[0, \pi]$ est $x \rightarrow -\sin x$, nulle en $x = 0$ et $x = \pi$. On en déduit :

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

On rappelle que les courbes représentatives des fonctions $x \mapsto \arcsin x$ et $x \mapsto \arccos x$ présentent une demi-tangente verticale aux points d'abscisse -1 et 1 , ce qui résulte du théorème de la limite de la dérivée (voir en 8.2.6).

La dérivée de $x \rightarrow \tan x$ sur $]-\frac{\pi}{2}, \frac{\pi}{2}[$ est $x \rightarrow 1 + \tan^2 x$, toujours non nulle. On en déduit :

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

8.1.6 Approximation d'un zéro par la méthode de Newton

Soit f une fonction de classe \mathcal{C}^1 sur un intervalle I .

On suppose que f' ne s'annule pas, mais que f s'annule en un point α de I .

La méthode d'approximation de α par la « méthode de Newton », consiste en les étapes suivantes :

- Initialisation : on part d'une valeur x_0 (plutôt proche de α).
- Si x_n est connu dans I , on mène la tangente Δ_n au point d'abscisse x_n de la courbe de f .
La droite Δ_n , d'équation $y = f(x_n) + (x - x_n)f'(x_n)$, recoupe Ox au point d'abscisse $x_n - \frac{f(x_n)}{f'(x_n)}$.
Si cette abscisse est dans I , on la nomme x_{n+1} et on continue à partir de x_{n+1} .
- Dans la pratique, on se donne $\varepsilon > 0$, et on s'arrête dès que $|x_{n+1} - x_n| \leq \varepsilon$.

Voici une fonction Python `newton(f, x0, eps, h)`, avec $\varepsilon = 10^{-8}$ par défaut.

On note h un « petit réel positif » (par défaut $h = 10^{-3}$) qui approche $f'(x)$ par $\frac{f(x+h) - f(x)}{h}$.

On a ajouté un exemple d'utilisation.

On obtient ici une valeur approchée des trois racines de $f(x) = x^3 - 2x^2 - x + 1$.

```
def newton(f, x0, eps=1e-8, h=1e-3):
    (xn, xn1) = (x0, x0+2*eps)
    while abs(xn1-xn) > eps:
        xn = xn1
        d = (f(xn+h)-f(xn))/h
        xn1 = xn - f(xn)/d
    return xn
```

```
>>> def f(x): return x*x*x-2*x*x-x+1
>>> newton(f, -0.8)
-0.801937729356303
>>> newton(f, 0.5)
0.5549581324591675
>>> newton(f, 2)
2.2469796051199724
```

8.2 Rolle et accroissements finis

8.2.1 Extremum local et point critique

Définition 8.2.1 (extremum local)

Soit $f : I \rightarrow \mathbb{R}$ une fonction numérique définie sur l'intervalle I . Soit a un élément de I .

On dit que f présente un *maximum local* en a si : $\exists \varepsilon > 0, \forall x \in I, (|x - a| \leq \varepsilon \Rightarrow f(x) \leq f(a))$

On dit que f présente un *minimum local* en a si : $\exists \varepsilon > 0, \forall x \in I, (|x - a| \leq \varepsilon \Rightarrow f(x) \geq f(a))$

On dit que f présente un *extremum local* en a si f présente un maximum ou un minimum local en a .

Si $f(x) \leq f(a)$ pour tout x de I , on dit que f présente un *maximum absolu* en a .

On définit de même la notion de *minimum absolu* et d'*extremum absolu*.

Définition 8.2.2 (point critique d'une fonction dérivable)

Soit $f : I \rightarrow \mathbb{R}$ une fonction dérivable. Soit a un élément de I .

On dit que a est un *point critique* de f si la dérivée de f est nulle en a .

Proposition 8.2.1

Soit $f : I \rightarrow \mathbb{R}$ une fonction dérivable. Soit a un point intérieur à I .

Si f possède un *extrémum local* en a , alors $f'(a) = 0$ (a est donc un *point critique* de f).

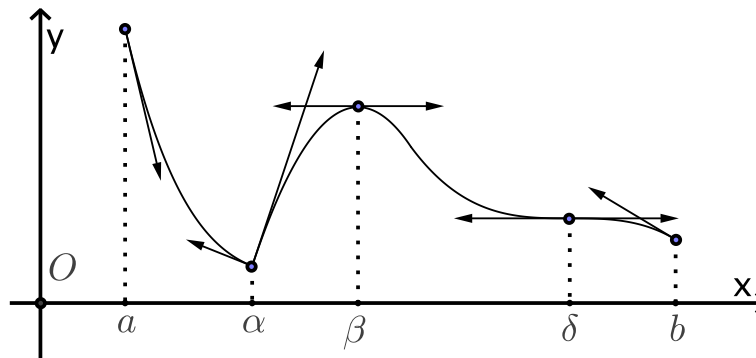
Remarques :

La réciproque est fautive : un point critique n'est pas nécessairement le signe d'un extrémum relatif.

Par exemple, avec $f(x) = x^3$, on a $f'(0) = 0$ mais f n'a pas d'extrémum en 0.

En fait, les extrémums locaux de f sur I sont à chercher parmi les points où f n'est pas dérivable, parmi les extrémités de I , et parmi les points critiques intérieurs à I . Le graphe ci-dessous illustre quelques cas possibles. On y voit une fonction définie sur le segment $[a, b]$, avec les propriétés suivantes :

- en a , la dérivée n'est pas nulle, mais f présente en ce point un maximum absolu.
- en α , il y a un minimum absolu, et en ce point f n'est pas dérivable.
- on voit que β est un point critique, et que ça correspond à un maximum relatif.
- on voit que δ est un point critique, mais ça ne correspond à aucun extrémum relatif.
- en b , la dérivée n'est pas nulle, mais f présente en ce point un minimum relatif.



8.2.2 Théorème de Rolle

Proposition 8.2.2 (Théorème de Rolle)

Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction définie sur le segment $[a, b]$, avec $a < b$, à valeurs réelles.

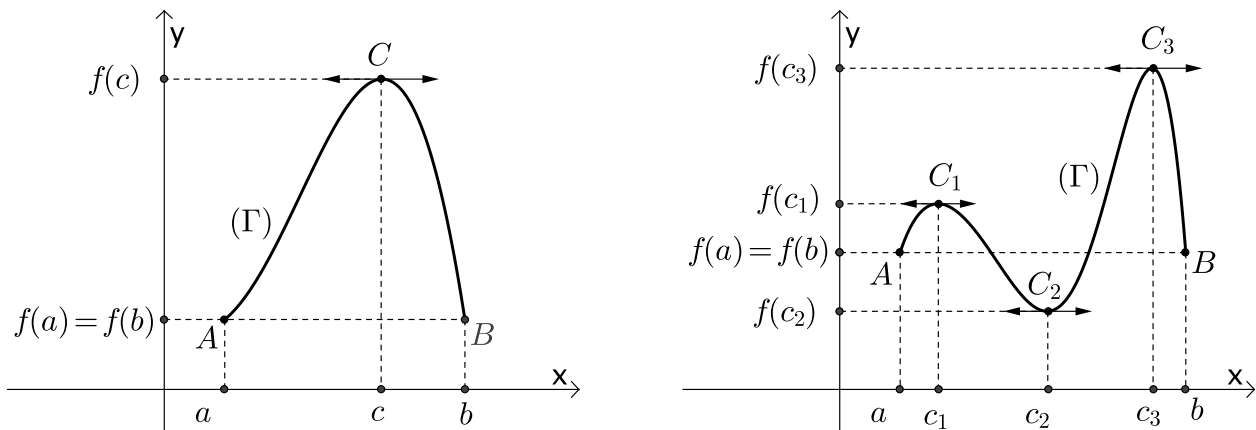
On suppose que f est continue sur $[a, b]$, dérivable sur $]a, b[$, et que $f(a) = f(b)$.

Alors il existe c dans $]a, b[$ tel que $f'(c) = 0$.

Interprétation géométrique

Soit (Γ) la courbe de f . Soit A, B les points d'abscisse a, b de (Γ) . Avec les hypothèses du théorème de Rolle, il y a un point C de (Γ) , distinct de A et B , en lequel la tangente à (Γ) est horizontale.

Comme on le voit sur l'illustration ci-dessous (à droite), il est possible qu'il existe plusieurs points strictement compris entre A et B et en lesquels la tangente à (Γ) est horizontale (donc plusieurs valeurs de c dans $]a, b[$ pour lesquelles $f'(c) = 0$).



Remarques

– Les hypothèses du théorème de Rolle sont extrêmement importantes (continuité sur le segment, dérivabilité à l'intérieur du segment, valeurs identiques aux extrémités).

Si l'une de ces hypothèses disparaît, alors le résultat ne tient plus (trouver des contre-exemples!).

Quand on utilise le théorème de Rolle, il faut le citer avec les hypothèses ci-dessus (et qui sont minimales), même si dans la pratique les propriétés de f sont souvent plus « confortables ».

– Le théorème de Rolle est souvent utilisé de manière répétée.

Supposons par exemple que f soit deux fois dérivable sur un intervalle I , et qu'il existe trois points $a < b < c$ de I tels que $f(a) = f(b) = f(c)$.

Avec « Rolle » sur $[a, b]$ et sur $[b, c]$, il existe α dans $]a, b[$ et β dans $]b, c[$ tels que $f'(\alpha) = f'(\beta) = 0$.

En appliquant Rolle à f' sur $[\alpha, \beta]$, on en déduit qu'il existe γ dans $]a, b[$ tel que $f''(\gamma) = 0$.

Plus généralement, si f est n fois dérivable sur I , et si elle s'annule en $n + 1$ points distincts, on montre (par une application répétée du théorème de Rolle) qu'il existe un élément c de I en lequel la fonction dérivée n -ième $f^{(n)}$ s'annule.

8.2.3 Égalité des accroissements finis

Proposition 8.2.3 (égalité des accroissements finis)

Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction définie sur le segment $[a, b]$, avec $a < b$, à valeurs réelles.

On suppose que f est continue sur $[a, b]$, dérivable sur $]a, b[$.

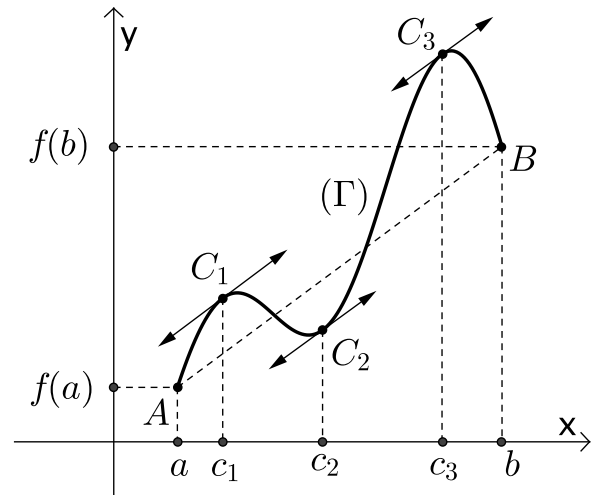
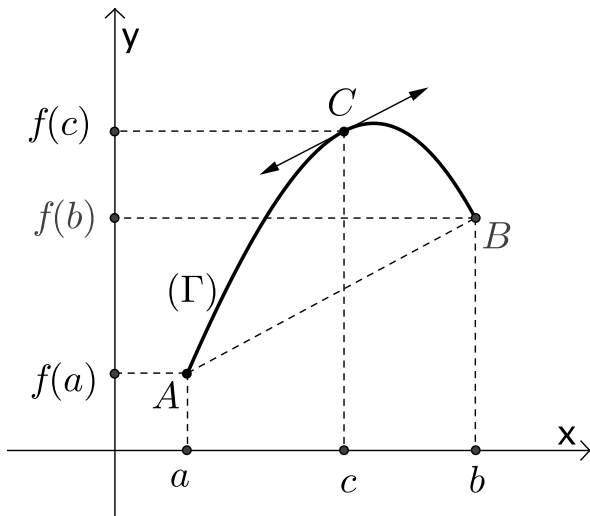
Alors il existe c dans $]a, b[$ tel que $f(b) - f(a) = (b - a)f'(c)$, c'est-à-dire : $\frac{f(b) - f(a)}{b - a} = f'(c)$.

Interprétation géométrique

Soit (Γ) la courbe de f . Soit A, B les points d'abscisse a, b de (Γ) .

Avec les hypothèses de l'égalité des accroissements finis, il y a un point (C) de (Γ) , distinct de A et B , en lequel la tangente à (Γ) est parallèle à la corde (AB) .

Comme on le voit sur l'illustration ci-dessous (à droite), il est possible qu'il existe plusieurs points strictement compris entre A et B et en lesquels la tangente à (Γ) est parallèle à la corde (AB) (donc plusieurs valeurs de c dans $]a, b[$ pour lesquelles $f(b) - f(a) = (b - a)f'(c)$).



Remarques

- Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction continue sur $[a, b]$, dérivable sur $]a, b[$ ($a < b$).
On suppose que : $\forall x \in]a, b[, m \leq f'(x) \leq M$. Alors $m(b - a) \leq f(b) - f(a) \leq M(b - a)$.
- L'égalité $f(b) - f(a) = (b - a)f'(c)$ est invariante par échange de a et b . Il n'est donc pas nécessaire d'imposer $a < b$. Il suffit d'indiquer que c est strictement compris entre a et b .
- Interprétation cinématique : si un point mobile se déplace sur une droite, pendant un intervalle de temps $a \leq t \leq b$, avec une vitesse variable $v(t)$, alors il existe au moins un instant t_0 en lequel sa vitesse instantanée est exactement égale à sa vitesse moyenne sur tout l'intervalle de temps.

L'égalité des accroissements finis est souvent écrite de la manière suivante :

Proposition 8.2.4 (une autre version de l'égalité des accroissements finis)

Soit f une fonction continue sur $[a, a + h]$ et dérivable sur $]a, a + h[$.
Alors il existe θ dans $]0, 1[$ tel que : $f(a + h) = f(a) + hf'(a + \theta h)$.

Dans la version précédente de l'égalité des accroissements finis, le signe de h est quelconque.
On peut donc reformuler le résultat de la façon suivante :

Proposition 8.2.5 (une version locale de l'égalité des accroissements finis)

Soit f une fonction dérivable au voisinage de a (sur un intervalle $]a - \delta, a + \delta[$, avec $\delta > 0$).
Alors, pour tout h vérifiant $|h| < \delta$, il existe θ dans $]0, 1[$ tel que : $f(a + h) = f(a) + hf'(a + \theta h)$.

Remarque : dans les deux énoncés précédents, θ dépend de h (le mieux serait donc de le noter θ_h).

8.2.4 Inégalité des accroissements finis

Définition 8.2.3 (fonction lipschitzienne)

Soit $f : I \rightarrow \mathbb{R}$ une fonction numérique définie sur un intervalle I . Soit K un réel strictement positif.
On dit que f est K -lipschitzienne sur I si : $\forall (x, y) \in I^2, |f(y) - f(x)| \leq K|y - x|$.

Remarques

- Si f est lipschitzienne sur l'intervalle I , elle est continue sur cet intervalle. La réciproque est fautive comme on le voit avec $x \mapsto \sqrt{x}$ sur $[0, 1]$.
- Le caractère lipschitzien est une propriété globale, alors que la continuité est une notion locale. Cela n'a donc aucun sens de parler de fonction lipschitzienne en un point.
- Quand une fonction est K -lipschitzienne, avec $K < 1$, on dit qu'elle est *contractante*.

Le résultat ci-dessous est appelé « inégalité des accroissements finis ».

Il constitue une condition suffisante pour qu'une fonction soit lipschitzienne sur un intervalle.

Proposition 8.2.6 (inégalité des accroissements finis)

Soit $f : I \rightarrow \mathbb{R}$ une fonction dérivable. On suppose que $|f'(x)| \leq K$ pour tout x de I .

Alors la fonction f est K -lipschitzienne sur I .

Autrement dit, pour tous x, y de I , on a : $|f(y) - f(x)| \leq K|y - x|$.

Une situation usuelle où les hypothèses précédentes sont réalisées est : f est de classe \mathcal{C}^1 sur $[a, b]$.

En effet f' est alors continue donc bornée sur $[a, b]$, et on peut utiliser $K = \sup_{x \in [a, b]} |f'(x)|$.

Utilisation pour les suites définies par une récurrence $u_{n+1} = f(u_n)$

On se donne une suite $(u_n)_{n \geq 0}$, définie par une relation de récurrence $u_{n+1} = f(u_n)$.

Soit ℓ un point fixe de f (c'est-à-dire $f(\ell) = \ell$, donc ℓ est une limite éventuelle de la suite $(u_n)_{n \geq 0}$).

On suppose que f est de classe \mathcal{C}^1 sur un intervalle I contenant ℓ .

On suppose également que $|f'(\ell)| < 1$.

Par continuité, il existe un segment J contenant ℓ et sur lequel on a $|f'(x)| \leq K < 1$.

On en déduit que f est K -lipschitzienne (donc *contractante*) sur le segment J .

Supposons que le terme initial u_0 soit dans le segment J .

Alors, pour tout n de \mathbb{N} , le terme u_n est encore dans J (le segment J est donc *stable* par f).

Cela résulte d'une récurrence facile et de : $|u_{n+1} - \ell| = |f(u_n) - f(\ell)| \leq K|u_n - \ell| \leq |u_n - \ell|$.

Toujours par récurrence évidente, on trouve $|u_n - \ell| \leq K^n |u_0 - \ell|$ pour tout n de \mathbb{N} .

Il en résulte $\lim u_n = \ell$.

On exprime cette situation (le fait que $|f'(\ell)| < 1$) en disant que ℓ est un point fixe *attractif* de f .

On constate que la convergence de la suite $(u_n)_{n \geq 0}$ est d'autant plus rapide que K peut être choisi proche de 0 (et donc que $|f'(\ell)|$ est lui-même proche de 0, c'est-à-dire que la tangente au point fixe est proche de l'horizontale).

Remarque (cas d'un point fixe *répulsif*)

Supposons $|f'(\ell)| > 1$, les autres hypothèses étant inchangées : $f(I) \subset I$, f est \mathcal{C}^1 sur I , et $f(\ell) = \ell$.

On dit alors que ℓ est un point fixe *répulsif* de f . La suite $(u_n)_{n \geq 0}$ ne peut alors pas converger vers ℓ (sauf s'il existe un entier n_0 tel que $u_{n_0} = \ell$, car alors la suite *stationne* en ℓ).

Pour ces questions, on se reportera au contenu de la section 6.6.5

8.2.5 Dérivabilité et sens de variation

Dans les propositions suivantes, I est un intervalle de \mathbb{R} , d'intérieur non vide.

Pour assurer des résultats de monotonie de f sur l'intervalle I (donc y compris aux extrémités) il suffit de vérifier des hypothèses sur la dérivée de f à l'intérieur de I .

Pour que la monotonie reste vraie « extrémités incluses » on utilise la continuité de f en ces extrémités.

On note $\overset{\circ}{I}$ « l'intérieur » de I , c'est-à-dire I privé de ses extrémités éventuelles.

Proposition 8.2.7 (caractérisation des fonctions constantes)

Toute fonction constante f de I dans \mathbb{R} est dérivable sur I et : $\forall x \in I, f'(x) = 0$.

Réciproquement, soit $f : I \rightarrow \mathbb{R}$, une fonction continue sur I , dérivable sur l'intérieur $\overset{\circ}{I}$ de I .

Si on a $f'(x) = 0$ pour tout x de $\overset{\circ}{I}$, alors f est constante sur I .

Proposition 8.2.8 (caractérisation des fonctions monotones)

Soit $f : I \rightarrow \mathbb{R}$ une fonction continue sur I , dérivable sur l'intérieur $\overset{\circ}{I}$ de I .

- la fonction f est croissante sur I si et seulement si : $\forall x \in \overset{\circ}{I}, f'(x) \geq 0$.
- la fonction f est décroissante sur I si et seulement si : $\forall x \in \overset{\circ}{I}, f'(x) \leq 0$.

Proposition 8.2.9 (caractérisation des fonctions strictement monotones)

Soit $f : I \rightarrow \mathbb{R}$ une fonction continue sur I , dérivable sur l'intérieur $\overset{\circ}{I}$ de I .

La fonction f est strictement monotone sur I si et seulement si sa dérivée f' garde un signe constant sur $\overset{\circ}{I}$, ne s'annulant sur aucun sous-intervalle de $\overset{\circ}{I}$ d'intérieur non vide (c'est-à-dire ne s'annulant éventuellement qu'en des points isolés de $\overset{\circ}{I}$).

Exemple : la fonction $f : x \mapsto x - \sin(x)$ a pour dérivée $f'(x) = 1 - \cos(x)$. Cette dérivée reste positive ou nulle sur \mathbb{R} , et ne s'annule qu'aux points isolés $x_k = 2k\pi$ (avec k dans \mathbb{Z}). La fonction f est donc strictement croissante sur \mathbb{R} .

Proposition 8.2.10 (fonctions ayant la même dérivée)

Soit f et g deux fonctions continues sur I , dérivables sur l'intérieur $\overset{\circ}{I}$ de I .

Les deux conditions suivantes sont équivalentes :

- pour tout x de I , on a $f'(x) = g'(x)$.
- il existe une constante λ telle que : $\forall x \in I, g(x) = f(x) + \lambda$.

Remarques

Les résultats précédents découlent du théorème de Rolle.

Comme lui, il ne sont valables que sur un intervalle, et pas sur une réunion d'intervalles.

Par exemple, si $f(x) = \frac{1}{x}$ alors $f'(x) = -\frac{1}{x^2} < 0$ sur \mathbb{R}^* , mais f n'est pas monotone sur \mathbb{R}^* .

De même, si deux fonctions dérivables sur \mathbb{R}^* vérifient $f' = g'$ sur \mathbb{R}^* , alors elles diffèrent d'une constante λ sur \mathbb{R}^{-*} et d'une constante μ (a priori sans rapport avec λ) sur \mathbb{R}^{+*} .

8.2.6 Théorème de la limite de la dérivée

Proposition 8.2.11 (théorème de la limite de la dérivée)

Soit f une fonction continue sur un intervalle I . Soit a un point de I .

On suppose que f est dérivable sur $I \setminus \{a\}$, et que $\lim_{x \rightarrow a} f'(x) = \ell$, avec ℓ dans $\overline{\mathbb{R}}$.

Alors, sous ces hypothèses, on a $\lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} = \ell$.

Cas particuliers et interprétations géométriques

Avec les notations de la proposition précédente :

– Si $\ell = +\infty$ ou $\ell = -\infty$, la courbe représentative de f présente en $A(a, f(a))$ une tangente verticale. C'est le cas notamment pour $x \mapsto \arcsin x$ et $x \mapsto \arccos x$ en $x = -1$ et en $x = 1$.

– Si ℓ est réel, alors f est dérivable en a et $f'(a) = \ell$. Ainsi prolongée, f' est continue en a .

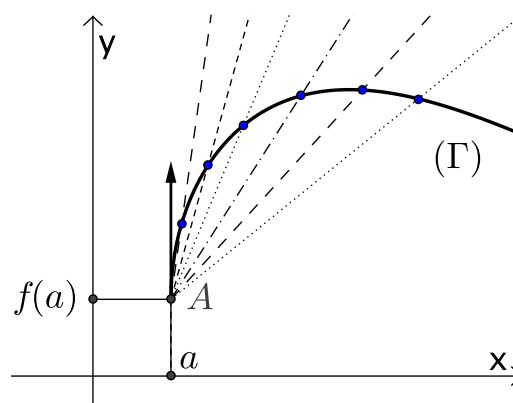
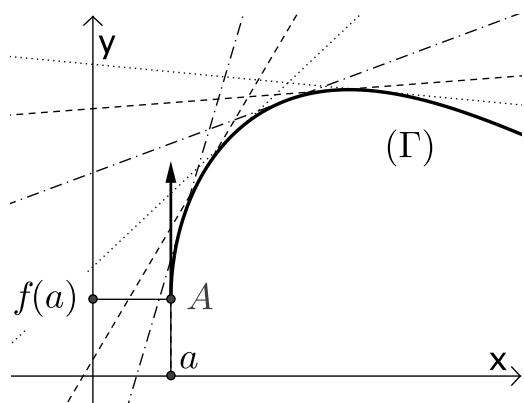
En particulier, si on suppose la continuité de f' sur $I \setminus \{a\}$, la fonction f est maintenant de classe \mathcal{C}^1 sur I tout entier (on peut alors parler de « théorème de classe \mathcal{C}^1 par prolongement »).

– Le résultat précédent permet d'affirmer la dérivabilité d'une fonction en un « point à problème ».

Par rapport à la méthode traditionnelle (limite du taux d'accroissement entre $A(a, f(a))$ et $M(x, f(x))$ quand x tend vers a , donc position limite de la corde (AM)), on utilise ici la limite de $f'(x)$ quand x tend vers a (donc la position limite de la tangente en M « quand M tend vers A »).

Les calculs sont souvent plus compliqués avec $\lim_{x \rightarrow a} f'(x)$, mais on y gagne la continuité de f' en a .

– On illustre ici le cas d'une tangente verticale en $A(a, f(a))$: à gauche avec le théorème de la limite de la dérivée ($\lim_{x \rightarrow a} f'(x) = +\infty$) et à droite avec la position limite des cordes.



Un contre-exemple utile

On considère la fonction f définie sur \mathbb{R}^* par $f(x) = x^2 \sin\left(\frac{1}{x}\right)$ et par $f(0) = 0$.

La fonction f est continue sur \mathbb{R}^* (théorèmes généraux) et en 0 (inégalité $|f(x)| \leq x^2$).

La fonction f est de classe \mathcal{C}^1 sur \mathbb{R}^* : sa dérivée est alors $f'(x) = 2x \sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right)$.

Pour $x \neq 0$, on a $\left| \frac{f(x) - f(0)}{x} \right| = \left| x \sin\left(\frac{1}{x}\right) \right| \leq |x|$ donc $\lim_{x \rightarrow 0} \frac{f(x) - f(0)}{x} = 0$.

Ainsi f est dérivable en 0 et $f'(0) = 0$.

En revanche, la limite de $f'(x) = 2x \sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right)$ n'existe pas quand x tend vers 0.

8.3 Fonctions de classe \mathcal{C}^k

8.3.1 Fonctions de classe \mathcal{C}^k sur un intervalle

On rappelle que I désigne un intervalle de \mathbb{R} d'intérieur non vide.

Définition 8.3.1 (fonctions k fois dérivables sur un intervalle)

Soit $f : I \rightarrow \mathbb{R}$ une fonction numérique réelle définie sur l'intervalle I . On pose $f^{(0)} = f$.

Si la fonction $f^{(k)}$, avec k dans \mathbb{N} , est définie et dérivable sur I , on pose $f^{(k+1)} = (f^{(k)})'$.

Si $f^{(k)}$ est définie sur I , on dit que f est k fois dérivable sur I .

La fonction $f^{(k)}$ est appelée *dérivée k -ième* de f sur I . On peut la noter $D^k(f)$ ou $\frac{d^k f}{dx^k}$.

Évidemment, on note souvent f'' et f''' , plutôt que $f^{(2)}$ et $f^{(3)}$, les dérivées seconde et troisième de f . Évidemment toujours, on ne confondra jamais les notations $f^{(k)}$ et f^k .

Définition 8.3.2 (fonction de classe \mathcal{C}^k sur un intervalle)

Soit $f : I \rightarrow \mathbb{R}$ une fonction numérique réelle définie sur l'intervalle I . Soit k un entier naturel.

Si $f : I \rightarrow \mathbb{R}$ est k fois dérivable, et si $f^{(k)}$ est continue sur I , on dit que f est de classe \mathcal{C}^k sur I .

On dit que f est de classe \mathcal{C}^∞ sur I si f est indéfiniment dérivable sur I .

On note $\mathcal{C}^k(I, \mathbb{R})$ l'ensemble des fonctions de classe \mathcal{C}^k de I dans \mathbb{R} (avec k dans $\mathbb{N} \cup \{+\infty\}$).

Remarques :

- Si f est de classe \mathcal{C}^k sur I , avec $k \geq 1$, toutes les fonctions dérivées j -ièmes, avec $0 \leq j < k$ sont évidemment continues sur I puisqu'elles y sont (au moins une fois) dérivables.
- On dit souvent d'une fonction de classe \mathcal{C}^k qu'elle est « k fois continûment dérivable ».
- Dire que f est de classe \mathcal{C}^0 , c'est dire que f est continue.
Pour tout k de \mathbb{N} , on a bien sûr $\mathcal{C}^{k+1}(I, \mathbb{R}) \subset \mathcal{C}^k(I, \mathbb{R})$. Par ailleurs $\mathcal{C}^\infty(I, \mathbb{R}) = \bigcap_{k \in \mathbb{N}} \mathcal{C}^k(I, \mathbb{R})$.
- On dit parfois « f est \mathcal{C}^k sur I » plutôt que « f est de classe \mathcal{C}^k sur I ».
- On a $f^{(k)} \equiv 0$ sur I si et seulement si f est une fonction polynomiale, avec $\deg(f) < k$.

8.3.2 Opérations sur les fonctions de classe \mathcal{C}^k

Dans les énoncés suivants, k est un élément de $\mathbb{N} \cup \{+\infty\}$.

Proposition 8.3.1 (combinaisons linéaires de fonctions de classe \mathcal{C}^k)

Soit f et g deux fonctions de classe \mathcal{C}^k de I dans \mathbb{R} . Soit α, β deux réels.

Alors $\alpha f + \beta g$ est de classe \mathcal{C}^k sur I et : $(\alpha f + \beta g)^{(k)} = \alpha f^{(k)} + \beta g^{(k)}$.

Proposition 8.3.2 (formule de Leibniz)

Soit k un élément de $\mathbb{N} \cup \{+\infty\}$.

Soit f et g deux fonctions de classe \mathcal{C}^k de I dans \mathbb{R} .

Alors la fonction fg est de classe \mathcal{C}^k sur I et on a : $(fg)^{(k)} = \sum_{j=0}^k \binom{k}{j} f^{(j)} g^{(k-j)}$.

Proposition 8.3.3 (inverse d'une fonction de classe \mathcal{C}^k)

Si $f : I \rightarrow \mathbb{R}$ est de classe \mathcal{C}^k sur I et ne s'annule pas, alors $\frac{1}{f}$ est de classe \mathcal{C}^k sur I .

Proposition 8.3.4 (composition de fonctions de classe \mathcal{C}^k)

Soit $f : I \rightarrow \mathbb{R}$, et $g : J \rightarrow \mathbb{R}$, toutes deux de classe \mathcal{C}^k , avec $f(I) \subset J$.

Alors la fonction $g \circ f$ est de classe \mathcal{C}^k de I dans \mathbb{R} .

Proposition 8.3.5 (bijection réciproque d'une fonction de classe \mathcal{C}^k)

Soit f une fonction de classe \mathcal{C}^k de I dans \mathbb{R} .

On suppose que $f'(x) > 0$ pour tout x de I , ou que $f'(x) < 0$ pour tout x de I .

La fonction f réalise donc une bijection de I sur un intervalle J .

Dans ces conditions, la bijection réciproque f^{-1} est également de classe \mathcal{C}^k .

Exemples de fonctions de classe \mathcal{C}^∞

– Les fonctions polynomiales sont de classe \mathcal{C}^∞ sur \mathbb{R} .

Il en est de même des fonctions rationnelles (sur leur domaine de définition).

– Les fonctions $x \mapsto x^\alpha$ et $x \mapsto \ln x$ sont de classe \mathcal{C}^∞ sur \mathbb{R}^{+*} .

Les fonctions $x \mapsto \exp x$, $x \mapsto \operatorname{ch} x$, $x \mapsto \operatorname{sh} x$ et $x \mapsto \operatorname{th} x$ sont de classe \mathcal{C}^∞ sur \mathbb{R} .

– Les fonctions $x \mapsto \sin x$ et $x \mapsto \cos x$ sont de classe \mathcal{C}^∞ sur \mathbb{R} .

La fonction $x \mapsto \tan x$ est de classe \mathcal{C}^∞ sur son domaine. La fonction $x \mapsto \arctan x$ est \mathcal{C}^∞ sur \mathbb{R} .

Les fonctions $x \mapsto \arcsin x$, et $x \mapsto \arccos x$ sont de classe \mathcal{C}^∞ sur $] -1, 1[$.

– Les fonctions qui se déduisent des précédentes par somme, produit, quotient, puissance et composition sont de classe \mathcal{C}^∞ sur leur domaine de définition.

8.3.3 Théorème de classe \mathcal{C}^k par prolongement

Proposition 8.3.6

Soit I un intervalle d'intérieur non vide, et soit a un élément de I .

Soit k un entier naturel. On se donne une fonction f , définie et de classe \mathcal{C}^k sur $I \setminus \{a\}$.

On suppose que, pour tout j de $\{0, \dots, k\}$, la limite $\ell_j = \lim_{x \rightarrow a} f^{(j)}(x)$ existe et est finie.

En particulier, on peut prolonger f en a , par continuité, en posant $f(a) = \ell_0$.

Ainsi prolongée, f est de classe \mathcal{C}^k sur I et : $\forall j \in \{0, \dots, k\}$, $f^{(j)}(a) = \ell_j$.

Le résultat précédent s'étend à $k = +\infty$, en remplaçant « $\forall j \in \{0, \dots, k\}$ » par « $\forall j \in \mathbb{N}$ ».

Le cas de la fonction définie sur \mathbb{R}^{+*} par $f(x) = \exp\left(-\frac{1}{x}\right)$ est classique.

Il est clair en effet que f est de classe \mathcal{C}^∞ sur \mathbb{R}^{+*} (opérations entre fonctions de classe \mathcal{C}^∞).

On montre par récurrence que $f^{(n)}(x) = \frac{P_n(x)}{x^{2n}} \exp\left(-\frac{1}{x}\right)$, où P_n est un polynôme tel que $P_n(0) = 1$.

On en déduit : $\lim_{x \rightarrow 0^+} f^{(n)}(x) = \lim_{x \rightarrow 0^+} \frac{1}{x^{2n}} \exp\left(-\frac{1}{x}\right) = \lim_{X \rightarrow +\infty} X^{2n} \exp(-X) = 0$.

Ainsi la fonction f , prolongée par $f(0) = 0$, est de classe \mathcal{C}^∞ sur \mathbb{R}^+ .

8.4 Extension aux fonctions complexes

8.4.1 Fonctions de classe \mathcal{C}^k à valeurs dans \mathbb{C}

Définition 8.4.1 (fonctions de classe \mathcal{C}^k à valeurs complexes)

Soit $f : I \rightarrow \mathbb{C}$ une fonction à valeurs complexes. Soit $u = \operatorname{Re}(f)$ et $v = \operatorname{Im}(f)$.

On dit que f est de classe \mathcal{C}^k sur I si et seulement si les fonctions u et v sont de classe \mathcal{C}^k sur I .

On note alors $f^{(k)} = u^{(k)} + iv^{(k)}$ et on dit que $f^{(k)}$ est la fonction dérivée k -ième de f .

Important : s'il est possible de considérer des fonctions dérivables sur un intervalle I de \mathbb{R} et à valeurs complexes, on ne parlera **jamais** de la dérivée d'une fonction qui serait définie sur une partie de \mathbb{C} .

On retiendra qu'on ne dérive **jamais** (autant le répéter) par rapport à une variable complexe.

Cela rejoint une faute à ne **jamais** commettre : on ne dérive **jamais** par rapport à une variable entière !

Par exemple, pour étudier la monotonie d'une suite réelle, **ça n'a aucun sens** de dériver l'expression de u_n par rapport à l'entier n . Voilà c'est dit.

8.4.2 Extension des résultats

Un certain nombre de résultats s'étendent aux fonctions complexes, mais pas tous :

- Les résultats sur les opérations entre fonctions de classe \mathcal{C}^k (combinaison linéaire, produit, quotient, et les calculs qui vont avec) sont encore valables pour les fonctions à valeurs dans \mathbb{C} .

La formule de Leibniz est encore valable.

Le cas de la composition est un peu à part. Pour pouvoir affirmer que $g \circ f$ est de classe \mathcal{C}^k , il faut que f et g soient de classe \mathcal{C}^k mais que f soit à valeurs réelles (alors que g peut être à valeurs complexes).

- La fonction $f - g$ est constante sur I si et seulement si on a $f' = g'$ sur I .

En revanche, il n'est plus question de lier le « signe de la dérivée » avec le « sens de variations », car aucune de ces deux notions n'a de sens pour les fonctions à valeurs dans \mathbb{C} .

On pourrait toujours parler de « point critique » (un point où la dérivée s'annule), mais cette notion ne présente plus beaucoup d'intérêt (puisque l'on a perdu la notion d'extremum relatif).

- Il n'y a plus de théorème de la bijection réciproque.

- **Important** : il n'y a pas de théorème de Rolle pour les fonctions à valeurs dans \mathbb{C} .

On retiendra l'exemple classique de la fonction $f : t \mapsto e^{it}$, dont la dérivée est $f'(t) = ie^{it}$.

On a $f(0) = f(2\pi)$ (valeur commune 1), mais la dérivée de f ne s'annule jamais !

- Conséquence de la perte du théorème de Rolle, **on perd l'égalité des accroissements finis**, mais attention, il reste toujours **l'inégalité des accroissements finis** (voir plus loin).

- Le théorème de classe \mathcal{C}^k par prolongement est encore valable, bien que rarement utilisé.

- On peut encore définir les fonctions K -lipschitziennes (dans $|f(y) - f(x)| \leq K|y - x|$, c'est un module à gauche, une valeur absolue à droite).

Proposition 8.4.1 (inégalité des accroissements finis pour une fonction de classe \mathcal{C}^1)

Soit $f : I \rightarrow \mathbb{C}$ une fonction de classe \mathcal{C}^1 . On suppose que $|f'(x)| \leq K$ pour tout x de I .

Alors la fonction f est K -lipschitzienne sur I .

Autrement dit, pour tous x, y de I , on a : $|f(y) - f(x)| \leq K|y - x|$.

Chapitre 9

Analyse asymptotique

Sommaire

9.1	Rappels de quelques limites usuelles	205
9.2	Comparaison des suites	205
9.2.1	Domination, négligeabilité, équivalence	206
9.2.2	Utilisation des notations de Landau	207
9.2.3	Traduction des croissances comparées	208
9.2.4	Opérations sur les équivalents	208
9.2.5	Limites usuelles et équivalents de suites	209
9.3	Comparaison des fonctions	209
9.3.1	Domination, négligeabilité, équivalence	209
9.3.2	Propriétés des relations de comparaison	210
9.3.3	Conseils pour utiliser les équivalents	211
9.3.4	Comparaisons usuelles	211
9.4	Développements limités	212
9.4.1	DL, unicité des coefficients, troncature	212
9.4.2	Développement limité en 0 et parité	213
9.4.3	Développements limités et dérivabilité	214
9.4.4	Développements limités usuels	215
9.5	Opérations sur les développements limités	216
9.5.1	Utilisation de combinaisons linéaires	216
9.5.2	Produit de deux développements limités	216
9.5.3	Composition de deux DL	217
9.5.4	Inverse d'un développement limité	218
9.5.5	Quotient de deux développements limités	218
9.5.6	Primitivation d'un DL	220
9.5.7	Dérivation d'un DL	220
9.5.8	Pratique des compositions de DL	220
9.6	Application des développements limités	222
9.6.1	Équivalents et DL	222
9.6.2	Position par rapport à une tangente	222
9.6.3	Branches infinies	225
9.7	Exemples de développements asymptotiques	226
9.7.1	Position du problème	226
9.7.2	Quelques exemples simples	227

9.7.3 Étude d’une suite définie implicitement 228
 9.7.4 Formule de Stirling 229

9.1 Rappels de quelques limites usuelles

Logarithme, exponentielle, puissances : croissances comparées

Les limites qui suivent (cf chapitre 4 : « Techniques d’analyse (dérivation) », constituent une échelle de comparaison entre fonctions puissances, fonction exponentielle, et fonction logarithme :

Pour tous réels strictement positifs α, β, δ , on a :

$$\lim_{x \rightarrow -\infty} |x|^\alpha e^{\delta x} = 0$$

$$\lim_{x \rightarrow +\infty} \frac{e^{\delta x}}{x^\alpha} = +\infty$$

$$\lim_{x \rightarrow 0^+} x^\alpha |\ln(x)|^\beta = 0$$

$$\lim_{x \rightarrow +\infty} \frac{\ln^\beta(x)}{x^\alpha} = 0$$

Dérivées usuelles à l’origine

Par ailleurs, soit f une fonction définie et dérivable en 0.

On sait que $\lim_{x \rightarrow 0} \frac{f(x) - f(0)}{x} = f'(0)$ (c’est même la définition de la dérivée en 0).

En particulier : $\lim_{x \rightarrow 0} \frac{e^x - 1}{x} = 1$ $\lim_{x \rightarrow 0} \frac{\ln(1+x)}{x} = 1$ et $\lim_{x \rightarrow 0} \frac{(1+x)^\alpha - 1}{x} = \alpha$

Une simple translation de la variable permet d’écrire : $\lim_{x \rightarrow 1} \frac{\ln(x)}{x-1} = 1$ et $\lim_{x \rightarrow 1} \frac{x^m - 1}{x-1} = m$

Dans de nombreux cas usuels, on a $f(0) = 0$ et $f'(0) = 1$ (ce qui traduit le fait que la première bissectrice $y = x$ est la tangente au point $(0, 0)$ de la courbe représentative). On a alors $\lim_{x \rightarrow 0} \frac{f(x)}{x} = 1$.

Dans cette catégorie, on trouve : $\lim_{x \rightarrow 0} \frac{\sin(x)}{x} = \lim_{x \rightarrow 0} \frac{\tan(x)}{x} = \lim_{x \rightarrow 0} \frac{\text{sh}(x)}{x} = \lim_{x \rightarrow 0} \frac{\text{th}(x)}{x} = 1$

Voici deux limites utiles : $\lim_{x \rightarrow 0} \frac{\cos(x) - 1}{x^2} = -\frac{1}{2}$ et $\lim_{x \rightarrow 0} \frac{\text{ch}(x) - 1}{x^2} = \frac{1}{2}$ (attention au signe !)

Les limites précédentes ont été encadrées car... il faut les connaître !

9.2 Comparaison des suites

Dans toute cette section, on considère des suites à valeurs réelles ou complexes.

Quand on *compare* une suite (u_n) à une suite (v_n) , on suppose que v_n est non nul pour « n assez grand ».

Cette hypothèse (qui sera toujours sous-entendue) permet d’évoquer $\lim \frac{u_n}{v_n}$.

Les comparaisons de suites sont toujours effectuées « quand n tend vers $+\infty$ ».

9.2.1 Domination, négligeabilité, équivalence

Définition 9.2.1 (suite dominée par une autre)

Soit (u_n) et (v_n) deux suites à valeurs réelles ou complexes.

On dit que la suite (u_n) est *dominée* par la suite (v_n) si la suite de terme général $\frac{u_n}{v_n}$ est bornée.

On note alors $u_n = O(v_n)$, et on prononce : « u_n est un grand O de v_n ».

Remarques

Dire que $u_n = O(v_n)$, c'est dire qu'il existe M dans \mathbb{R}^+ et n_0 dans \mathbb{N} tels que $|u_n| \leq M |v_n|$ pour $n \geq n_0$.

Si $0 < m \leq \left| \frac{u_n}{v_n} \right| \leq M$ pour n assez grand, chacune des deux suites (u_n) et (v_n) est dominée par l'autre.

Il est clair que si $u_n = O(v_n)$ et $v_n = O(w_n)$, alors $u_n = O(w_n)$ (transitivité de la domination).

Définition 9.2.2 (suite négligeable devant une autre)

Soit (u_n) et (v_n) deux suites à valeurs réelles ou complexes.

On dit que la suite (u_n) est *négligeable* devant la suite (v_n) si $\lim \frac{u_n}{v_n} = 0$.

On note alors $u_n = o(v_n)$ (et on prononce : u_n est un « petit o » de v_n).

Remarques

Dire que $u_n = o(v_n)$, c'est dire que : $\forall \varepsilon > 0, \exists n_0 \in \mathbb{N}, \forall n \geq n_0, |u_n| \leq \varepsilon |v_n|$.

Il est manifeste que si $u_n = o(v_n)$, alors $u_n = O(v_n)$, mais la réciproque est fautive.

Il est clair que si $\begin{cases} u_n = O(v_n) \\ v_n = o(w_n) \end{cases}$ ou $\begin{cases} u_n = o(v_n) \\ v_n = O(w_n) \end{cases}$ ou $\begin{cases} u_n = o(v_n) \\ v_n = o(w_n) \end{cases}$ alors $u_n = o(w_n)$

Définition 9.2.3 (suite équivalente à une autre)

Soit (u_n) et (v_n) deux suites à valeurs réelles ou complexes.

On dit que la suite (u_n) est *équivalente* à la suite (v_n) si $\lim \frac{u_n}{v_n} = 1$. On note alors $u_n \sim v_n$.

Remarques

– Rappelons qu'on considère ici des suites réelles ou complexes dont le terme général est non nul au voisinage de $+\infty$ (c'est-à-dire au-delà d'un certain entier n_0).

Sur l'ensemble de ces suites, la relation $u_n \sim v_n$ est une relation d'équivalence.

Pour toutes suites (u_n) , (v_n) et (w_n) , on a en effet les propriétés suivantes :

$$u_n \sim u_n \text{ (réflexivité)}, u_n \sim v_n \Leftrightarrow v_n \sim u_n \text{ (symétrie)}, \begin{cases} u_n \sim v_n \\ v_n \sim w_n \end{cases} \Rightarrow u_n \sim w_n \text{ (transitivité)}.$$

– La symétrie de la relation \sim permet de parler de *suites équivalentes* (sans préciser dans quel ordre).

– Si λ est un scalaire **non nul**, on a : $u_n \sim \lambda \Leftrightarrow \lim u_n = \lambda$.

Attention!!! on n'écrira jamais $u_n \sim 0$ (car ça n'a aucun sens).

Proposition 9.2.1

Soit (u_n) et (v_n) deux suites à valeurs réelles ou complexes.

Alors il revient au même de dire $u_n \sim v_n$ ou $u_n - v_n = o(v_n)$.

9.2.2 Utilisation des notations de Landau

Les notations o et O , dans $u_n = o(v_n)$ et $u_n = O(v_n)$ sont appelées « notations de Landau ».

Elles permettent de renseigner sur l'ordre de grandeur du terme général d'une suite (quand c'est la seule information importante). Plutôt que des égalités, les propositions $u_n = o(v_n)$ (resp. $u_n = O(v_n)$) doivent être interprétées comme l'appartenance de (u_n) à la catégorie de suites qui sont négligeables devant (resp. dominées par) la suite (v_n) .

Dans les écritures $u_n = o(v_n)$ et $u_n = O(v_n)$, la suite (v_n) est le plus souvent une *suite de référence*.

Par exemple :

- écrire $u_n = O(1)$, c'est dire que la suite (u_n) est bornée.
- écrire $u_n = o(1)$, c'est dire que la suite (u_n) converge vers 0.
- écrire $u_n = O(n^2)$, c'est dire qu'il existe $K \geq 0$ tel que $|u_n| \leq Kn^2$ pour n assez grand.
- écrire $u_n = o(n^2)$, c'est dire que $\lim \frac{u_n}{n^2} = 0$.
- écrire $u_n = o\left(\frac{1}{n^2}\right)$, c'est dire que $\lim n^2 u_n = 0$.

La relation $u_n \sim v_n$, synonyme de $u_n - v_n = o(v_n)$, sera souvent écrite $u_n = v_n + o(v_n)$: sous cette forme, elle signifie que u_n est la somme de v_n et d'une suite elle-même négligeable devant la suite (v_n) .

Inversement, si à l'issue d'un calcul on arrive à $u_n = v_n + o(v_n)$, on pourra écrire $u_n \sim v_n$.

La relation $u_n = v_n + o(v_n)$ est parfois plus fiable que la notation $u_n \sim v_n$, notamment dans les opérations de sommation (voir plus loin). Mais on utilise par ailleurs très souvent les notations de Landau dans des calculs algébriques (notamment dans la partie « développements limités » de ce chapitre).

Il suffit de connaître quelques règles assez intuitives, notamment :

Proposition 9.2.2

Soit (u_n) et (v_n) deux suites à valeurs réelles ou complexes. Soit λ un scalaire non nul.

Si $u_n = O(v_n)$, alors $\lambda u_n = O(v_n)$. Si $u_n = o(v_n)$, alors $\lambda u_n = o(v_n)$.

Le résultat précédent signifie que dans les relations de domination et de négligeabilité, les constantes numériques multiplicatives sont simplement « cachées ».

Par exemple, on n'écrira jamais $u_n = O(2n^2)$ ou $u_n = O(-n^2)$, mais plus simplement $u_n = O(n^2)$.

Le résultat suivant dit que l'ensemble des suites qui sont dominées par (resp. négligeables devant) une suite (w_n) donnée est stable par combinaisons linéaires :

Proposition 9.2.3

Soit (u_n) , (v_n) et (w_n) trois suites à valeurs réelles ou complexes.

Si $u_n = O(w_n)$ et si $v_n = O(w_n)$, alors pour tous scalaires α et β , on a : $\alpha u_n + \beta v_n = O(w_n)$.

Si $u_n = o(w_n)$ et si $v_n = o(w_n)$, alors pour tous scalaires α et β , on a : $\alpha u_n + \beta v_n = o(w_n)$.

Plus généralement, considérons une suite (u_n) de terme général $u_n = v_n + \alpha a_n + \beta b_n + \gamma c_n + \dots$, où les suites (a_n) , (b_n) , (c_n) , etc. sont négligeables devant la suite (v_n) .

Alors on peut écrire $\alpha a_n + \beta b_n + \gamma c_n + \dots = o(v_n)$, donc $u_n = v_n + o(v_n)$, donc $u_n \sim v_n$.

9.2.3 Traduction des croissances comparées

Soit α, β, δ trois réels strictement positifs quelconques.

– On connaît la limite $\lim_{n \rightarrow +\infty} n^\alpha e^{-\delta n} = 0$, qui peut maintenant s'écrire : $e^{-\delta n} = o\left(\frac{1}{n^\alpha}\right)$

Ainsi $e^{-\delta n}$ (qui tend vers 0 quand n tend vers $+\infty$) est négligeable devant toute puissance de $1/n$.

On dit aussi que lorsque n tend vers $+\infty$, $e^{-\delta n}$ (qui tend alors vers 0) est un « infiniment petit d'ordre supérieur à tout ordre donné par rapport à l'infiniment petit principal $1/n$ ».

– On sait que $\lim_{n \rightarrow +\infty} \frac{e^{\delta n}}{n^\alpha} = +\infty$, ou encore $\lim_{n \rightarrow +\infty} \frac{n^\alpha}{e^{\delta n}} = 0$, ce qui peut s'écrire : $n^\alpha = o(e^{\delta n})$

Ainsi toute puissance de n est négligeable devant $e^{\delta n}$ quand n tend vers $+\infty$.

– On sait que $\lim_{n \rightarrow +\infty} \frac{\ln^\beta(n)}{n^\alpha} = 0$, ce qui s'écrit maintenant $\ln^\beta(n) = o(n^\alpha)$

Ainsi toute puissance de $\ln(n)$ est négligeable devant toute puissance de n quand n tend vers $+\infty$.

– Pour résumer, de façon très informelle, on peut retenir : $\ln^\beta(n) \ll n^\alpha \ll e^{\delta n}$ (avec α, β, δ dans \mathbb{R}^{+*})

9.2.4 Opérations sur les équivalents

Proposition 9.2.4 (conservation du signe)

Soit (u_n) et (v_n) deux suites à valeurs réelles.

Si $u_n \sim v_n$, alors u_n et v_n gardent le même signe au voisinage de $+\infty$.

Proposition 9.2.5 (conservation de la limite)

Soit (u_n) et (v_n) deux suites à valeurs réelles ou complexes.

Si $u_n \sim v_n$, et si $\lim v_n = \ell$ (avec ℓ dans $\overline{\mathbb{R}}$ ou dans \mathbb{C}), alors $\lim u_n = \ell$.

Réciproquement : si $\lim u_n = \lim v_n = \ell$ (où ℓ est dans \mathbb{C}^*), alors $u_n \sim v_n$.

Attention !! Si $\lim u_n = \lim v_n = \ell$, avec $\ell \in \{-\infty, 0, +\infty\}$, rien ne permet d'affirmer que $u_n \sim v_n$.

Attention !! même si $\lim u_n = 0$, on n'écrira jamais $u_n \sim 0$ (on l'a déjà dit!).

Proposition 9.2.6 (équivalences dans un produit ou dans un quotient)

Soit (u_n) , (v_n) et (a_n) trois suites à valeurs réelles ou complexes.

Si $u_n \sim v_n$ alors $u_n a_n \sim v_n a_n$ et $\frac{a_n}{u_n} \sim \frac{a_n}{v_n}$, et plus généralement : $a_n u_n^\alpha \sim a_n v_n^\alpha$.

Dans un produit ou un quotient de suites, on peut donc remplacer l'une des suites par une suite équivalente. L'expression initiale est alors équivalente à l'expression finale. En particulier, il y a conservation de la limite éventuelle quand n tend vers $+\infty$.

Un cas particulier très simple est le produit par une constante non nulle λ : si $u_n \sim v_n$ alors $\lambda u_n \sim \lambda v_n$.

Attention aux constantes multiplicatives !

On sait qu'elles sont « cachées » dans les relations du type $u_n = O(v_n)$ ou $u_n = o(v_n)$, mais il en va tout autrement dans les relations du type $u_n \sim v_n$.

Pour prendre un exemple, on sait que $u_n = o(2n^3) \Leftrightarrow u_n = o(-n^3) \Leftrightarrow u_n = o(n^3)$ (et cette dernière formulation est recommandée).

En revanche, les propositions $u_n \sim 2n^3$, $u_n \sim -n^3$, et $u_n \sim n^3$ n'ont pas la même signification !

Proposition 9.2.7 (jamais d'équivalents dans une somme!!!)

Attention !! Si $u_n \sim v_n$, alors on n'a pas, en général, $u_n + a_n \sim v_n + a_n$.

On retiendra qu'on ne doit **JAMAIS** utiliser d'équivalents dans une somme.

Posons par exemple $u_n = 1 + \frac{1}{n}$, $v_n = 1 + \frac{1}{n^2}$, et $a_n = -1$.

On a : $u_n \sim v_n$ car ces deux suites tendent vers la même limite non nulle $\ell = 1$.

En revanche $u_n + a_n = \frac{1}{n}$ n'est pas équivalent à $v_n + a_n = \frac{1}{n^2}$.

9.2.5 Limites usuelles et équivalents de suites

Les limites usuelles rappelées au début du chapitre peuvent servir aux calculs d'équivalents de suites.

L'idée générale est que si $\lim_{x \rightarrow a} f(x) = \ell \in \mathbb{R}^*$ et $\lim u_n = a$, alors $\lim f(u_n) = \ell$, ou encore $f(u_n) \sim \ell$.

Par exemple : si $\lim u_n = 0$, alors $\frac{e^{u_n} - 1}{u_n} \sim 1$, ou encore $e^{u_n} - 1 \sim u_n$.

On trouve donc facilement :

- si $\lim u_n = 0$, alors : $e^{u_n} - 1 \sim u_n$, $\ln(1 + u_n) \sim u_n$, $(1 + u_n)^\alpha - 1 \sim \alpha u_n$
en particulier, toujours si $\lim u_n = 0$, on a $\sqrt{1 + u_n} - 1 \sim \frac{u_n}{2}$
- si $\lim u_n = 1$, alors : $\ln(u_n) \sim u_n - 1$, $u_n^\alpha - 1 \sim \alpha(u_n - 1)$
- si $\lim u_n = 0$, alors : $\sin(u_n) \sim u_n$, $\tan(u_n) \sim u_n$, $\text{sh}(u_n) \sim u_n$, $\text{th}(u_n) \sim u_n$
- si $\lim u_n = 0$, alors : $\cos(u_n) - 1 \sim -\frac{u_n^2}{2}$, $\text{ch}(u_n) - 1 \sim \frac{u_n^2}{2}$

9.3 Comparaison des fonctions

9.3.1 Domination, négligeabilité, équivalence

Dans ce paragraphe, on considère un intervalle I de \mathbb{R} , d'intérieur non vide.

On désigne par a un élément ou une extrémité de I (a est dans $\overline{\mathbb{R}}$).

On se limite ici à des fonctions définies sur $I \setminus \{a\}$ et à valeurs réelles ou complexes (à valeurs réelles dans la plupart des cas), et qui ne s'annulent pas au voisinage de a (elles peuvent être définies et nulles au point a lui-même, mais ça n'intervient pas dans les définitions).

Les comparaisons de fonctions ont lieu au voisinage du point a , avec a dans \mathbb{R} , ou $a = +\infty$ ou $a = -\infty$ (alors que pour les suites, les comparaisons ont lieu uniquement quand n tend vers $+\infty$).

Les définitions suivantes constituent des *comparaisons locales* de fonctions quand x tend vers a (on dit aussi, de façon plus informelle mais moins précise, que les comparaisons sont effectuées « en a »).

Les définitions et les propriétés qui en découlent sont très proches de celles vues pour les comparaisons des suites numériques. On se contente donc d'une simple adaptation de ce qui a été dit précédemment.

Définition 9.3.1 (fonction dominée par une autre en un point)

Soient f, g deux fonctions de $I \setminus \{a\}$, à valeurs dans \mathbb{R} (ou \mathbb{C}).

On dit que f est *dominée* par g quand x tend vers a (ou en a) si $\frac{f}{g}$ est bornée au voisinage de a .

On note alors $f = O(g)$, ou éventuellement $f = O_a(g)$.

Définition équivalente : il existe $M \geq 0$ tel que $|f(x)| \leq M|g(x)|$ au voisinage de a .

Définition 9.3.2 (fonction négligeable devant une autre en un point)

Soient f, g deux fonctions de $I \setminus \{a\}$ à valeurs dans \mathbb{R} (ou \mathbb{C}).

On dit que f est *négligeable* devant g quand x tend vers a (ou en a) si $\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = 0$.

On note alors $f = o(g)$, ou éventuellement $f = o_a(g)$.

Définition équivalente : pour tout $\varepsilon > 0$, il existe un voisinage de a (dépendant bien sûr de ε) sur lequel on a l'inégalité $|f(x)| \leq \varepsilon|g(x)|$.

Définition 9.3.3 (fonction équivalente à une autre en un point)

Soient f, g deux fonctions de $I \setminus \{a\}$ à valeurs dans \mathbb{R} (ou \mathbb{C}).

On dit que f est équivalente à g au voisinage de a (ou en a) si $\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = 1$.

On note alors $f \sim g$, ou éventuellement $f \sim_a g$.

Définition équivalente : dire que $f \sim_a g$, c'est dire que $f - g$ est négligeable devant g en a .

9.3.2 Propriétés des relations de comparaison

Dans les résultats suivants, on désigne par f, g, h des fonctions numériques définies sur $I \setminus \{a\}$ et qui ne s'annulent pas au voisinage de a . Les relations de comparaison sont établies quand x tend vers a .

- Si $f = o(g)$ alors $f = O(g)$.
Si $f = o(g)$ et $g = O(h)$ alors $f = o(h)$ (même résultat si $f = O(g)$ et $g = o(h)$).
- Si $f = O(h)$ et si $g = O(h)$, alors $\alpha f + \beta g = O(h)$ (même résultat avec des « o »).
Si $\alpha \neq 0$ et si $f = O(g)$ (resp. $f = o(g)$, $f \sim g$), alors $f = O(\alpha g)$ (resp. $f = o(g)$, $\alpha f \sim \alpha g$).
- Si $f \sim g$ (et sont à valeurs réelles), alors f et g gardent le même signe au voisinage de $+\infty$.
Si $f \sim g$, et si $\lim g = \ell$ (avec ℓ dans $\overline{\mathbb{R}}$ ou dans \mathbb{C}), alors $\lim f = \ell$.
Réciproquement : si $\lim f = \lim g = \ell$ (où ℓ est un scalaire non nul), alors $f \sim g$.
- Si $f \sim g$ alors $fh \sim gh$ et $\frac{h}{f} \sim \frac{h}{g}$, et plus généralement : $h f^\alpha \sim h g^\alpha$.
- Attention !! Si $f \sim g$ alors on n'a pas, en général, $f + h \sim g + h$.
- Si f_2, f_3, \dots, f_n sont des $o(f_1)$ alors $f_1 + f_2 + \dots + f_n \sim f_1$.

Proposition 9.3.1 (changement de variable dans un équivalent)

Soient f, g deux fonctions de $I \setminus \{a\}$ à valeurs dans \mathbb{R} (ou \mathbb{C}).

Soit φ une fonction de J dans I , qui tend vers a quand x tend vers b dans J .

Si f est dominée par g en a , alors $f \circ \varphi$ est dominée par $g \circ \varphi$ en b .

Si f est négligeable devant g en a , alors $f \circ \varphi$ est négligeable devant $g \circ \varphi$ en b .

Si f et g sont équivalentes en a , alors $f \circ \varphi$ et $g \circ \varphi$ sont équivalentes en b .

C'est surtout cette dernière propriété qui est utilisée.

Par exemple, du fait que $\sin(x) \underset{0}{\sim} x$ en 0, on peut écrire : $\sin(x^2) \underset{0}{\sim} x^2$.

Toujours grâce à $\sin(x) \underset{0}{\sim} x$, on trouve : $\sin\left(\frac{1}{x}\right) \underset{\infty}{\sim} \frac{1}{x}$ ou encore $\sin(\ln(x)) \underset{1}{\sim} \ln(x) \underset{1}{\sim} (x-1)$.

9.3.3 Conseils pour utiliser les équivalents

– Les équivalents servent essentiellement aux calculs de limites :

On transforme une expression $f(x)$, dont on cherche la limite ℓ en un point a , en une expression équivalente $g(x)$ dont la limite en ce point est évidente (si ℓ est un scalaire non nul, il est courant qu'on aboutisse à $g(x) = \ell$).

Les outils essentiels sont les équivalents classiques (voir plus loin) et la possibilité qu'on a de remplacer les facteurs d'un produit ou d'un quotient par des équivalents.

– L'erreur la plus fréquente consiste à utiliser les équivalents dans des sommes. La seule propriété concernant les équivalents et les sommes peut s'écrire : $g = o(f) \Rightarrow f + g \sim f$.

– On évitera d'utiliser un équivalent d'une fonction f sous la forme $f \sim g + h$, avec $h = o(g)$, et surtout de donner un rôle à h : on se contentera d'écrire $f \sim g$.

Écrire par exemple $\cos(x) \underset{0}{\sim} 1 - \frac{x^2}{2}$ n'est pas faux mais c'est dangereux si on donne un rôle à $-\frac{x^2}{2}$.

En effet, on a aussi : $\cos(x) \underset{0}{\sim} 1 + x^2 \underset{0}{\sim} 1 - 36x^2 \dots$

Pour cet exemple, la solution est sans doute d'écrire : $1 - \cos(x) \underset{0}{\sim} \frac{x^2}{2}$.

– Soit $\lim_{x \rightarrow a} f(x) = \ell$, avec $\ell \neq 0$, alors $f(x) \underset{a}{\sim} \ell$. Mais si $\ell = 0$, on n'écrira pas $f(x) \sim 0$!

– Si $f \sim g$ (f et g étant positives et tendant vers une limite ℓ différente de 1) alors $\ln(f) \sim \ln(g)$.

C'est faux si f et g tendent vers 1.

Par exemple, $(1+x) \underset{0}{\sim} (1+x^2)$, mais en ce point $\begin{cases} \ln(1+x) \sim x \\ \ln(1+x^2) \sim x^2 \end{cases}$ et $x \not\sim x^2$

– On évitera surtout de prendre des « exponentielles » d'équivalents :

En effet $e^f \sim e^g \Leftrightarrow f - g = o(1)$, ce qui n'équivaut pas du tout à $f \sim g$.

Exemples : considérer x et x^2 en 0, ou encore x et $x+1$ en $+\infty$.

9.3.4 Comparaisons usuelles

Exponentielles, puissances et logarithmes

Soient α , β et δ des réels strictement positifs.

On sait que : $\lim_{x \rightarrow +\infty} x^\alpha e^{-\delta x} = 0$, $\lim_{x \rightarrow -\infty} |x|^\alpha e^{\delta x} = 0$, $\lim_{x \rightarrow +\infty} x^{-\alpha} \ln^\beta(x) = 0$, $\lim_{x \rightarrow 0+} x^\alpha |\ln^\beta(x)| = 0$.

Autrement dit : $x^\alpha \underset{+\infty}{=} o(e^{\beta x})$, $e^{\delta x} \underset{-\infty}{=} o(|x|^{-\alpha})$, $\ln(x)^\beta \underset{+\infty}{=} o(x^\alpha)$, $|\ln(x)|^\beta \underset{0}{=} o(x^{-\alpha})$.

Équivalents classiques

Si f est dérivable en 0 et si de plus $f(0) = 0$ et $f'(0) = 1$, alors $f(x) \sim x$ en 0.

En particulier, on a les équivalents : $\sin(x) \underset{0}{\sim} x$, $\tan(x) \underset{0}{\sim} x$, $\ln(1+x) \underset{0}{\sim} x$, et $e^x - 1 \underset{0}{\sim} x$.

Toujours à l'origine : $(1+x)^m - 1 \underset{0}{\sim} mx$, et $1 - \cos(x) \underset{0}{\sim} \frac{x^2}{2}$.

On peut aussi écrire, au voisinage de $x = 1$: $x^m - 1 \underset{1}{\sim} m(x-1)$ et $\ln(x) \underset{1}{\sim} x-1$.

Polynômes et fractions rationnelles

Soit $P(x) = a_mx^m + a_{m+1}x^{m+1} + \dots + a_{n-1}x^{n-1} + a_nx^n$ un polynôme ($m < n$, $a_m \neq 0$, $a_n \neq 0$).

Au voisinage de l'origine : $P(x) \underset{0}{\sim} a_mx^m$ (monôme de plus bas degré).

Au voisinage de $\pm\infty$: $P(x) \underset{\infty}{\sim} a_nx^n$ (monôme de plus haut degré).

Soit $f(x) = \frac{P(x)}{Q(x)}$ une fraction rationnelle (P et Q deux polynômes).

Au voisinage de 0, $f(x)$ est équivalente au quotient des monômes de plus bas degré.

Au voisinage de $\pm\infty$, $f(x)$ est équivalente au quotient des monômes de plus haut degré.

9.4 Développements limités

9.4.1 DL, unicité des coefficients, troncature

Définition 9.4.1 (développement limité d'ordre n en un point)

Soit $f : I \rightarrow \mathbb{R}$ une fonction, et soit x_0 un réel élément ou extrémité de I .

Soit n un entier naturel. On dit que f admet un *développement limité* (en abrégé un DL) à l'ordre n en x_0 s'il existe des réels a_0, a_1, \dots, a_n et une fonction $x \mapsto \varepsilon(x)$ tels que :

$$\forall x \in I, f(x) = \sum_{k=0}^n a_k(x-x_0)^k + (x-x_0)^n \varepsilon(x), \text{ avec } \lim_{x \rightarrow x_0} \varepsilon(x) = 0.$$

Utilisation des notations « o » ou « O »

Avec les notations de Landau, on écrira plutôt $f(x) = \sum_{k=0}^n a_k(x-x_0)^k + o((x-x_0)^n)$.

Il arrive qu'on utilise les notations "O" de Landau dans un développement limité.

Par exemple, si $f(x) = 1 + 2x^2 + x^3 - x^4 + o(x^4)$, alors $f(x) = 1 + 2x^2 + x^3 + O(x^4)$.

Cette dernière écriture contient un peu plus d'informations que $f(x) = 1 + 2x^2 + x^3 + o(x^3)$.

Troncature d'un développement limité

Supposons que f admette un DL d'ordre n en x_0 . Soit p un entier naturel, avec $p \leq n$.

Alors f admet un DL d'ordre p en x_0 , obtenu par *troncature*.

Plus précisément : $f(x) = \sum_{k=0}^n a_k(x-x_0)^k + o((x-x_0)^n) \Rightarrow f(x) = \sum_{k=0}^p a_k(x-x_0)^k + o((x-x_0)^p)$.

Par exemple, si $f(x) = 1 - x + 2x^3 + x^4 + o(x^4)$, alors $f(x) = 1 - x + 2x^3 + o(x^3)$.

Proposition 9.4.1 (unicité du développement limité)

Soit f une fonction admettant un DL d'ordre n au point x_0 : $f(x) = \sum_{k=0}^n a_k(x-x_0)^k + o((x-x_0)^n)$.
Alors les coefficients a_0, a_1, \dots, a_n sont définis de façon unique.

Le polynôme $P(x) = \sum_{k=0}^n a_k(x-x_0)^k$ est appelé partie régulière du développement limité.

Importance des développements à l'origine

Soit $f : I \rightarrow \mathbb{R}$ une fonction, et soit x_0 un réel élément ou extrémité de I .

Considérons la fonction g définie (au voisinage de $h = 0$) par $g(h) = f(x_0 + h)$.

Alors f possède un DL d'ordre n en x_0 si et seulement si g possède un DL d'ordre n en 0 .

Plus précisément, on a l'équivalence (pour f le DL est en $x = x_0$, et pour g il est en $h = 0$) :

$$f(x) = \sum_{k=0}^n a_k(x-x_0)^k + o((x-x_0)^n) \Leftrightarrow g(h) = f(x_0+h) = \sum_{k=0}^n a_k h^k + o(h^n)$$

Forme normalisée d'un développement limité

Soit f une fonction à valeurs réelles, admettant un DL d'ordre n en x_0 .

On sait que ce développement : $f(x) = a_0 + a_1(x-x_0) + a_2(x-x_0)^2 + \dots + a_n(x-x_0)^n + o((x-x_0)^n)$

peut aussi s'écrire (en posant $x = x_0 + h$) : $f(x_0+h) = a_0 + a_1 h + a_2 h^2 + \dots + a_n h^n + o(h^n)$.

Sous cette forme, on note p le plus petit indice tel que $a_p \neq 0$.

On obtient : $f(x_0+h) = h^p(a_p + a_{p+1}h + \dots + a_n h^{n-p} + o(h^{n-p}))$

Cette écriture est appelée *forme normalisée* du développement limité de f en x_0 .

Avec cette écriture (et puisque $a_p \neq 0$), on trouve $f(x_0+h) \underset{h \rightarrow 0}{\sim} a_p h^p$.

Cet équivalent donne le signe de f au voisinage de $x = x_0$:

- Si p est pair, alors $f(x)$ garde le signe de a_p au voisinage de x_0 .
- Si p est impair, alors $f(x)$ change de signe en x_0 :
Plus précisément, $f(x)$ est du signe contraire de a_p si $x < x_0$, puis du signe de a_p si $x > x_0$.

9.4.2 Développement limité en 0 et parité

Soit f une fonction définie sur un intervalle symétrique par rapport à 0 .

On suppose que f admet un développement limité d'ordre n à l'origine : $f(x) = \sum_{k=0}^n a_k x^k + o(x^n)$.

Le changement de x en $-x$ donne alors le DL : $f(-x) = \sum_{k=0}^n (-1)^k a_k x^k + o(x^n)$.

- Si f est paire, la partie régulière de son DL en 0 est un polynôme pair.
Autrement dit les coefficients d'indice impair a_{2k+1} sont nuls.
Le développement limité de f se réduit donc à : $f(x) = a_0 + a_2 x^2 + \dots + a_{2k} x^{2k} + \dots$
- Si f est impaire, la partie régulière de son DL en 0 est un polynôme impair.
Autrement dit les coefficients d'indice pair a_{2k} sont nuls.
Le développement limité de f se réduit donc à : $f(x) = a_1 x + a_3 x^3 + \dots + a_{2k+1} x^{2k+1} + \dots$

Si on forme le DL d'une fonction dont on sait qu'elle paire ou impaire, il pourra être utile d'utiliser la notation "O" pour améliorer à peu de frais la précision du développement.

Supposons par exemple que f soit paire : le développement $f(x) = a_0 + a_2x^2 + a_4x^4 + O(x^6)$ est plus précis que $f(x) = a_0 + a_2x^2 + a_4x^4 + o(x^5)$, lui-même plus précis que $f(x) = a_0 + a_2x^2 + a_4x^4 + o(x^4)$.

De même, si f est impaire, le développement $f(x) = a_1x + a_3x^3 + a_5x^5 + O(x^7)$ est plus précis que $f(x) = a_1x + a_3x^3 + a_5x^5 + o(x^6)$, lui-même plus précis que $f(x) = a_1x + a_3x^3 + a_5x^5 + o(x^5)$.

9.4.3 Développements limités et dérivabilité

Développement limité d'ordre 0 et continuité, DL et dérivabilité

Un DL de f d'ordre 0 en x_0 s'écrit $f(x) = a_0 + o(1)$, où $o(1)$ est une fonction tendant vers 0 en x_0 .

Dire que f a un tel DL en x_0 , c'est dire que f est continue (ou prolongeable par continuité) en x_0 .

Après prolongement éventuel, ce développement s'écrit : $f(x) = f(x_0) + o(1)$.

Développement limité d'ordre 1 et dérivabilité

Un DL de f d'ordre 1 en x_0 s'écrit $f(x) = a_0 + a_1(x - x_0) + o(x - x_0)$.

Dire que f admet un tel DL en x_0 , c'est dire que f est dérivable (après prolongement éventuel en x_0).

Après un tel prolongement, ce DL s'écrit $f(x) = f(x_0) + f'(x_0)(x - x_0) + o(x - x_0)$.

Développement limité d'ordre $n \geq 2$ et dérivabilité

Attention ! l'existence d'un DL d'ordre ≥ 2 en x_0 n'implique pas que f soit deux fois dérivable en x_0 .

Un contre-exemple est donné par $f(x) = x^3 \sin(1/x)$ en 0.

En effet, on a $f(x) = x^2(x \sin(1/x))$ et $\lim_{x \rightarrow 0} x \sin(1/x) = 0$, donc $f(x) = o(x^2)$.

On a donc un DL d'ordre 2 à l'origine : $f(x) = a_0 + a_1x + a_2x^2 + o(x^2)$, avec $a_0 = a_1 = a_2 = 0$.

Par troncature à l'ordre 1, on voit que f est dérivable en 0, avec $f(0) = f'(0) = 0$.

Mais f n'est pas deux fois dérivable en 0.

En effet, pour $x \neq 0$ on a $f'(x) = 3x^2 \sin\left(\frac{1}{x}\right) + x \cos\left(\frac{1}{x}\right)$, et $\frac{f'(x) - f'(0)}{x}$ n'a pas de limite en 0.

Obtention d'un DL par la formule de Taylor-Young

Proposition 9.4.2 (DL obtenu par la formule de Taylor-Young)

Si f est de classe \mathcal{C}^n de I dans \mathbb{R} , et si x_0 appartient à I , alors f possède un DL d'ordre n en x_0 .

Ce développement est obtenu par la formule de Taylor-Young :

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \cdots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + o((x - x_0)^n)$$

En particulier, si f est de classe \mathcal{C}^n au voisinage de 0, on a le développement :

$$f(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \cdots + \frac{f^{(n)}(0)}{n!}x^n + o(x^n)$$

9.4.4 Développements limités usuels

Tous les développements ci-dessous sont valables à l'origine, et peuvent être obtenus par la formule de Taylor-Young (ou par d'autres méthodes qui seront exposées plus loin).

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} + o(x^n) = \sum_{k=0}^n \frac{x^k}{k!} + o(x^n)$$

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + o(x^{2n+2}) = \sum_{k=0}^n (-1)^k \frac{x^{2k+1}}{(2k+1)!} + o(x^{2n+2})$$

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots + (-1)^n \frac{x^{2n}}{(2n)!} + o(x^{2n+1}) = \sum_{k=0}^n (-1)^k \frac{x^{2k}}{(2k)!} + o(x^{2n+1})$$

$$\operatorname{sh}(x) = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots + \frac{x^{2n+1}}{(2n+1)!} + o(x^{2n+2}) = \sum_{k=0}^n \frac{x^{2k+1}}{(2k+1)!} + o(x^{2n+2})$$

$$\operatorname{ch}(x) = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots + \frac{x^{2n}}{(2n)!} + o(x^{2n+1}) = \sum_{k=0}^n \frac{x^{2k}}{(2k)!} + o(x^{2n+1})$$

$$\tan(x) = x + \frac{x^3}{3} + \frac{2x^5}{15} + o(x^6) \quad \operatorname{th}(x) = x - \frac{x^3}{3} + \frac{2x^5}{15} + o(x^6)$$

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + \cdots + (-1)^n x^n + o(x^n) = \sum_{k=0}^n (-1)^k x^k + o(x^n)$$

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \cdots + x^n + o(x^n) = \sum_{k=0}^n x^k + o(x^n)$$

$$(1+x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!} x^2 + \cdots + \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!} x^n + o(x^n)$$

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} + \cdots + (-1)^{n+1} \frac{x^n}{n} + o(x^n) = \sum_{k=1}^n (-1)^{k+1} \frac{x^k}{k} + o(x^n)$$

$$\ln(1-x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \cdots - \frac{x^n}{n} + o(x^n) = -\sum_{k=1}^n \frac{x^k}{k} + o(x^n)$$

$$\arctan(x) = x - \frac{x^3}{3} + \frac{x^5}{5} + \cdots + (-1)^n \frac{x^{2n+1}}{2n+1} + o(x^{2n+2}) = \sum_{k=0}^n (-1)^k \frac{x^{2k+1}}{2k+1} + o(x^{2n+2})$$

$$\arcsin(x) = x + \frac{1}{2} \frac{x^3}{3} + \frac{1 \cdot 3}{2 \cdot 4} \frac{x^5}{5} + \frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6} \frac{x^7}{7} + \cdots + o(x^{2n+2}) \quad \arccos(x) = \frac{\pi}{2} - \arcsin(x) = \cdots$$

9.5 Opérations sur les développements limités

Pour simplifier, les résultats sont énoncés pour des développements limités à l'origine, mais on peut facilement les adapter à des développements en un autre point.

Pour chaque méthode, on donne un exemple d'utilisation (qui utilise éventuellement certains des développements qui seront présentés un peu plus loin).

9.5.1 Utilisation de combinaisons linéaires

Proposition 9.5.1 (combinaison linéaire de deux développements limités)

Soit $f, g : I \rightarrow \mathbb{R}$ telles que $f(x) = \sum_{k=0}^n a_k x^k + o(x^n)$ et $g(x) = \sum_{k=0}^n b_k x^k + o(x^n)$.

Alors, pour tous scalaires α, β , on a : $(\alpha f + \beta g)(x) = \sum_{k=0}^n (\alpha a_k + \beta b_k) x^k + o(x^n)$.

Exemple 1 : $\sin\left(x + \frac{\pi}{4}\right) = \frac{\sqrt{2}}{2} (\sin(x) + \cos(x)) = \frac{\sqrt{2}}{2} \left(1 + x - \frac{x^2}{2!} - \frac{x^3}{3!} + \frac{x^4}{4!} + o(x^4)\right)$.

Exemple 2 : $\frac{1}{2} \ln\left(\frac{1+x}{1-x}\right) = \frac{1}{2} (\ln(1+x) - \ln(1-x)) = x + \frac{x^3}{3} + \frac{x^5}{5} + \dots + \frac{x^{2n+1}}{2n+1} + o(x^{2n+2})$.

Exemple 3 : on sait que $e^x = \sum_{k=0}^{2n} \frac{x^k}{k!} + o(x^{2n})$ donc $e^{-x} = \sum_{k=0}^{2n} (-1)^k \frac{x^k}{k!} + o(x^{2n})$ pour tout n .

On en déduit : $\operatorname{ch}(x) = \frac{e^x + e^{-x}}{2} = \sum_{k=0}^n \frac{x^{2k}}{(2k)!} + o(x^{2n})$ et $\operatorname{sh}(x) = \frac{e^x - e^{-x}}{2} = \sum_{k=0}^{n-1} \frac{x^{2k+1}}{(2k+1)!} + o(x^{2n})$

9.5.2 Produit de deux développements limités

Proposition 9.5.2 (produit de deux développements limités)

Soit $f, g : I \rightarrow \mathbb{R}$ telles que $f(x) = \sum_{k=0}^n a_k x^k + o(x^n)$ et $g(x) = \sum_{k=0}^n b_k x^k + o(x^n)$.

Alors $(fg)(x) = \sum_{k=0}^n c_k x^k + o(x^n)$, avec $c_k = \sum_{i+j=k} a_i b_j$.

Exemple 1 :

On sait que $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + o(x^4)$ et $\frac{1}{1-x} = 1 + x + x^2 + x^3 + x^4 + o(x^4)$.

On en déduit $\frac{e^x}{1-x} = 1 + 2x + \frac{5}{2}x^2 + \frac{8}{3}x^3 + \frac{65}{24}x^4 + o(x^4)$

Exemple 2 :

On sait que : $\frac{1}{1-x} = 1 + x + x^2 + \dots + x^n + o(x^n)$

Par élévation au carré, on en déduit : $\frac{1}{(1-x)^2} = 1 + 2x + 3x^2 + \dots + (n+1)x^n + o(x^n)$

Utilisation des formes normalisées :

Soit $f, g : I \rightarrow \mathbb{R}$ deux fonctions admettant un DL en 0.

La forme normalisée de ces deux DL aide à comprendre quel est l'ordre du développement produit.

Posons par exemple
$$\begin{cases} f(x) = x^p (a_0 + a_1x + \dots + a_nx^n + o(x^n)) \\ g(x) = x^q (b_0 + b_1x + \dots + b_mx^m + o(x^m)) \end{cases} \quad \text{avec } a_0 \neq 0 \text{ et } b_0 \neq 0.$$

Ainsi $f(x)g(x) = x^{p+q} (a_0 + a_1x + \dots + a_nx^n + o(x^n)) (b_0 + b_1x + \dots + b_mx^m + o(x^m))$.

Les deux DL entre parenthèses doivent être tronqués à l'ordre minimum (supposons par exemple $n < m$).

On obtient alors le développement limité de fg en 0 :

$$\begin{aligned} f(x)g(x) &= x^{p+q} (a_0 + a_1x + \dots + a_nx^n + o(x^n)) (b_0 + b_1x + \dots + b_nx^n + o(x^n)) \\ &= x^{p+q} (c_0 + c_1x + \dots + c_nx^n + o(x^n)) \quad \text{où } c_k = \sum_{i+j=k} a_i b_j \end{aligned}$$

Par exemple, pour calculer le DL de $(1 - \cos(x))(\sin(x) - x)$ en 0 à l'ordre 8 :

On écrit le développement normalisé : $1 - \cos(x) = \frac{x^2}{2!} - \frac{x^4}{4!} + o(x^5) = x^2 \left(\frac{1}{2} - \frac{x^2}{24} + o(x^3) \right)$

On écrit celui de : $\sin(x) - x = -\frac{x^3}{3!} + \frac{x^5}{5!} + o(x^6) = x^3 \left(-\frac{1}{6} + \frac{x^2}{120} + o(x^3) \right)$.

Ainsi : $(1 - \cos(x))(\sin(x) - x) = x^5 \left(\frac{1}{2} - \frac{x^2}{24} + o(x^3) \right) \left(-\frac{1}{6} + \frac{x^2}{120} + o(x^3) \right) = x^5 \left(-\frac{1}{12} + \frac{x^2}{90} + o(x^3) \right)$

9.5.3 Composition de deux DL**Proposition 9.5.3** (composition de deux développements limités)

On suppose $f(x) = P(x) + o(x^n)$, avec $P(x) = \sum_{k=1}^n a_k x^k$ et $g(y) = Q(y) + o(y^n)$, avec $Q(y) = \sum_{k=0}^n b_k y^k$.

On note l'hypothèse $P(0) = a_0 = 0$, donc $\lim_{x \rightarrow 0} f(x) = 0$, qui permet de considérer $g \circ f$ quand $x \rightarrow 0$.

Dans ces conditions, la fonction $g \circ f$ admet un DL d'ordre n en 0, dont la partie régulière s'obtient en conservant les termes de degrés inférieurs ou égaux à n dans $Q(y)$ quand on a posé $y = P(x)$.

Dans la pratique, on pose $g(y) = \sum_{k=0}^n b_k y^k + o(y^n)$ et on remplace y par le DL de $f(x)$.

On calcule de proche en proche les DL des puissances successives $y^k = f(x)^k$, en ne gardant à chaque étape que les puissances x^m avec $m \leq n$.

Exemple :

Supposons $f(x) = x - x^2 + 2x^3 + x^4 + o(x^4)$ et $g(y) = 1 + y + 3y^2 - y^3 - y^4 + o(y^4)$.

Posons $y = f(x) = x - x^2 + 2x^3 + x^4 + o(x^4)$.

On trouve $y^2 = x^2 - 2x^3 + 5x^4 + o(x^4)$, puis $y^3 = x^3 - 3x^4 + o(x^4)$ et $y^4 = x^4 + o(x^4)$.

On en déduit le développement limité de $g \circ f$ à l'ordre 4 à l'origine :

$$(g \circ f)(x) = 1 + y + 3y^2 - y^3 - y^4 + o(y^4) = 1 + x + 2x^2 - 5x^3 + 18x^4 + o(x^4)$$

Les calculs précédents peuvent avantageusement prendre place dans un tableau comme indiqué ci-contre. Un tel tableau est particulièrement indiqué quand aucun des deux DL à composer n'est pair ou impair.

					coeff
y	x	$-x^2$	$2x^3$	x^4	1
y^2		x^2	$-2x^3$	$5x^4$	3
y^3			x^3	$-3x^4$	-1
y^4				x^4	-1
	x	$2x^2$	$-5x^3$	$18x^4$	

Une remarque importante

La forme « triangulaire supérieure » du tableau précédent est plus que normale !

Elle traduit le fait que $y = f(x)$ est un infiniment petit quand x tend vers 0, c'est-à-dire $\lim_{x \rightarrow 0} f(x) = 0$.

Cette condition est indispensable pour que la composition du DL de f par celui de g soit possible.

9.5.4 Inverse d'un développement limité

Proposition 9.5.4 (inverse d'un développement limité)

Soit $f : I \rightarrow \mathbb{R}$ admettant un DL d'ordre n en 0 : $f(x) = \sum_{k=0}^n a_k x^k + o(x^n)$, avec $a_0 \neq 0$.

Dans ces conditions la fonction $x \mapsto \frac{1}{f(x)}$ possède un DL d'ordre n en 0.

Pour cela on écrit $\frac{1}{f(x)} = \frac{1}{a_0(1+g(x))}$ où $g(x) = \frac{1}{a_0} (a_1 x + a_2 x^2 + \dots + a_n x^n + o(x^n))$.

On compose ensuite le développement de $g(x)$ par celui de la fonction $y \mapsto \frac{1}{1+y}$.

Exemple :

On veut calculer le développement limité de $x \mapsto \frac{1}{\cos(x)}$ à l'origine, à l'ordre 7.

On sait que $\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + o(x^7) = 1 - \frac{x^2}{2} + \frac{x^4}{24} - \frac{x^6}{720} + o(x^7)$.

On pose donc $\frac{1}{\cos(x)} = \frac{1}{1+g(x)}$, avec $g(x) = -\frac{x^2}{2} + \frac{x^4}{24} - \frac{x^6}{720} + o(x^7)$.

On injecte ensuite $y = g(x)$ dans le développement $\frac{1}{1+y} = 1 - y + y^2 - y^3 + O(y^4)$.

On trouve $y^2 = \frac{x^4}{4} - \frac{x^6}{24} + o(x^7)$, $y^3 = -\frac{x^6}{8} + o(x^7)$, et finalement $\frac{1}{\cos(x)} = 1 + \frac{x^2}{2} + \frac{5x^4}{24} + \frac{61x^6}{720} + o(x^7)$.

9.5.5 Quotient de deux développements limités

Proposition 9.5.5

Soit $f, g : I \rightarrow \mathbb{R}$ telles que $f(x) = \sum_{k=0}^n a_k x^k + o(x^n)$ et $g(x) = \sum_{k=0}^n b_k x^k + o(x^n)$, avec $b_0 \neq 0$.

On suppose donc que g ne tend vers 0 à l'origine. Alors $\frac{f}{g}$ admet un DL en 0 à l'ordre n .

Ce développement est obtenu en effectuant le produit de celui de f par celui de $\frac{1}{g}$.

Exemple : on peut obtenir le développement limité de $\tan(x)$ en 0 par quotient.

On sait que $\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + o(x^8)$, et on a vu que $\frac{1}{\cos(x)} = 1 + \frac{x^2}{2} + \frac{5x^4}{24} + \frac{61x^6}{720} + o(x^7)$.

$$\begin{aligned} \text{Ainsi : } \tan(x) &= \frac{\sin(x)}{\cos(x)} = x \left(1 - \frac{x^2}{6} + \frac{x^4}{120} - \frac{x^6}{5040} + o(x^7) \right) \left(1 + \frac{x^2}{2} + \frac{5x^4}{24} + \frac{61x^6}{720} + o(x^7) \right) \\ &= x + \frac{x^3}{3} + \frac{2x^5}{15} + \frac{17x^7}{315} + o(x^8) \end{aligned}$$

9.5.6 Primitivation d'un DL

Proposition 9.5.6 (primitivation d'un développement limité)

Soit $f : I \rightarrow \mathbb{R}$ admettant un DL d'ordre n en 0 : $f(x) = \sum_{k=0}^n a_k x^k + o(x^n)$.

Soit F une primitive de f sur l'intervalle I (donc une fonction dérivable telle que $F' = f$).

Alors F possède en 0 un DL d'ordre $n + 1$ obtenu par intégration terme à terme de celui de f .

Plus précisément : $F(x) = F(0) + \sum_{k=0}^n \frac{a_k}{k+1} x^{k+1} + o(x^{n+1})$ (ne pas oublier $F(0)$...)

Exemple 1 : Si $f(x) = \ln(\cos(x))$, alors $f'(x) = -\tan(x) = -x - \frac{x^3}{3} - \frac{2x^5}{15} + o(x^6)$.

On en déduit $f(x) = -\frac{x^2}{2} - \frac{x^4}{12} - \frac{x^6}{45} + o(x^7)$.

Exemple 2 : Si $f(x) = \arctan\left(\frac{x+2}{1-2x}\right)$, alors $f'(x) = \frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + o(x^7)$.

On en déduit $f(x) = \arctan(2) + x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + o(x^8)$.

9.5.7 Dérivation d'un DL

Proposition 9.5.7 (dérivation d'un développement limité)

Soit f une fonction de classe \mathcal{C}^∞ au voisinage de 0 .

Alors le développement limité de f' en 0 à l'ordre n s'obtient en dérivant terme à terme le développement limité de f en 0 à l'ordre $n + 1$ (ces deux DL résultent de la formule de Taylor-Young).

Exemple : On sait que $\frac{1}{1-x} = 1 + x + x^2 + \dots + x^n + o(x^n)$.

Par dérivation, on en déduit : $\frac{1}{(1-x)^2} = 1 + 2x + 3x^2 + 4x^3 + \dots + (n+1)x^n + o(x^n)$.

Après une nouvelle dérivation, on obtient : $\frac{1}{(1-x)^3} = 1 + 3x + 6x^2 + \dots + \frac{(n+2)(n+1)}{2}x^n + o(x^n)$.

9.5.8 Pratique des compositions de DL

Quand on doit calculer le développement limité, à un ordre déterminé, d'une fonction f qui s'exprime elle-même à partir d'autres fonctions g, h, \dots il faut prendre le temps de comprendre à quel ordre les développements de g, h, \dots doivent être calculés.

Il y a en effet deux risques : partir avec des développements « trop longs » et donc faire des calculs inutiles, ou au contraire partir avec des développements « trop courts » ce qui oblige à tout recommencer.

Exemple 1 : on demande le DL de $f(x) = \ln(1+x - \arctan(x))$ en 0 à l'ordre 6.

Il est clair qu'il faut composer le développement de $g(x) = x - \arctan(x)$ par celui de $\ln(1+y)$.

D'une part, $g(x) = x - \arctan(x) = \frac{x^3}{3} - \frac{x^5}{5} + o(x^6)$ (ici, il faut pousser le DL à l'ordre 6 : pourquoi ?)

D'autre part, $\ln(1+y) = y - \frac{y^2}{2} + O(y^3)$ (ici un développement en $O(y^3)$ suffit : pourquoi ?).

On trouve $y^2 = \frac{x^6}{18} + o(x^6)$ puis finalement : $\ln(1 + x - \arctan(x)) = \frac{x^3}{3} - \frac{x^5}{5} - \frac{x^6}{18} + o(x^6)$.

Exemple 2 : On demande le DL en 0 de $f(x) = \frac{1}{x} \ln(\cos(\sqrt{x}))$ à l'ordre 2.

On écrit $\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + O(x^8)$ puis $\cos(\sqrt{x}) = 1 - \frac{x}{2} + \frac{x^2}{24} - \frac{x^3}{720} + o(x^3)$.

On pose $y = -\frac{x}{2} + \frac{x^2}{24} - \frac{x^3}{720} + o(x^3)$ et on compose par $\ln(1 + y) = y - \frac{y^2}{2} + \frac{y^3}{3} + o(y^3)$.

Après calcul : $\ln(\cos(\sqrt{x})) = -\frac{x}{2} - \frac{x^2}{12} - \frac{x^3}{45} + o(x^3)$, donc $\frac{1}{x} \ln(\cos(\sqrt{x})) = -\frac{1}{2} - \frac{x}{12} - \frac{x^2}{45} + o(x^2)$.

Pour obtenir un DL à l'ordre 2, il a donc fallu développer $x \mapsto \cos(x)$ à l'ordre 6 (exemple instructif).

Cas particuliers de composition

Quand on veut calculer le DL de $g \circ f(x)$ en 0 en composant les développements de $y = f(x)$ et de $g(y)$ à l'origine, il faut veiller à ce que $y = f(x)$ soit bien un infiniment petit lorsque x tend vers 0, afin que la substitution de y par $f(x)$ soit justifiée dans le développement de $g(y)$.

Si ce n'est pas le cas, on peut parfois s'y ramener, comme dans les exemples suivants :

▷ Si $g(x) = \exp(f(x)) = \exp(a_0 + a_1x + a_2x^2 + \dots) = \exp(a_0) \exp(a_1x + a_2x^2 + \dots)$.

On pose $y = a_1x + a_2x^2 + \dots$ et on utilise $\exp(y) = 1 + y + \frac{y^2}{2!} + \dots$

▷ Si $g(x) = \ln(f(x)) = \ln(a_0 + a_1x + a_2x^2 + \dots) = \ln(a_0) + \ln\left(1 + \frac{a_1}{a_0}x + \frac{a_2}{a_0}x^2 + \dots\right)$.

On pose $y = \frac{a_1}{a_0}x + \frac{a_2}{a_0}x^2 + \dots$ et on utilise $\ln(1 + y) = y - \frac{y^2}{2} + \dots$

▷ Si $g(x) = f(x)^\alpha = (a_0 + a_1x + a_2x^2 + \dots)^\alpha = a_0^\alpha \left(1 + \frac{a_1}{a_0}x + \frac{a_2}{a_0}x^2 + \dots\right)^\alpha$.

On pose $y = \frac{a_1}{a_0}x + \frac{a_2}{a_0}x^2 + \dots$ et on utilise $(1 + y)^\alpha = 1 + \alpha y + \frac{\alpha(\alpha - 1)}{2}y^2 + \dots$

Remarque sur le développement de $(1 + x)^\alpha$

On sait que $(1 + x)^\alpha = 1 + a_1x + a_2x^2 + \dots + a_nx^n + o(x^n)$, où $a_k = \frac{\alpha(\alpha - 1) \dots (\alpha - k + 1)}{k!}$.

On remarque la relation $a_{k+1} = a_k \frac{\alpha - k}{k + 1}$, qui permet de calculer les a_k de proche en proche.

Si on doit former le DL de $(1 + x)^\alpha$ avec une valeur particulière de α , et plutôt que d'utiliser la formule donnant a_k , il est préférable d'exploiter cette récurrence dans un tableau comme indiqué ci-dessous :

multiplicateur	1	α	$(\alpha - 1)\frac{1}{2}$	$(\alpha - 2)\frac{1}{3}$	$(\alpha - 3)\frac{1}{4}$	$(\alpha - 4)\frac{1}{5}$
coefficient	$a_0 = 1$	a_1	a_2	a_3	a_4	a_5

Par exemple, pour développer $f(x) = \sqrt{1 + x}$:

multiplicateur	1	$\frac{1}{2}$	$-\frac{1}{2} \times \frac{1}{2}$	$-\frac{3}{2} \times \frac{1}{3}$	$-\frac{5}{2} \times \frac{1}{4}$	$-\frac{7}{2} \times \frac{1}{5}$
coefficient	$a_0 = 1$	$a_1 = \frac{1}{2}$	$a_2 = \frac{-1}{8}$	$a_3 = \frac{1}{16}$	$a_4 = \frac{-5}{128}$	$a_5 = \frac{7}{256}$

On en déduit : $\sqrt{1+x} = 1 + \frac{x}{2} - \frac{x^2}{8} + \frac{x^3}{16} - \frac{5x^4}{128} + \frac{7x^5}{256} + o(x^5)$

9.6 Application des développements limités

9.6.1 Équivalents et DL

On utilise principalement les équivalents dans les recherches de limites, mais on se tourne vers les développements limités si on a besoin de davantage de précision (par exemple non seulement l'existence d'une demi-tangente mais encore la position de la courbe par rapport à celle-ci) ou quand il est difficile d'utiliser des équivalents (notamment dans les sommes).

– Il arrive qu'on ait besoin de développements limités pour trouver un simple équivalent d'une somme.

Par exemple, pour un équivalent de $\sin(\operatorname{sh}(x)) - \operatorname{sh}(\sin(x))$ en 0, il faut développer $\sin(x)$ et $\operatorname{sh}(x)$ à l'ordre 7 (pour atteindre les premiers coefficients qui ne se simplifient pas) :

On trouve $\sin(\operatorname{sh}(x)) = x - \frac{x^5}{15} - \frac{x^7}{90} + o(x^7)$, et $\operatorname{sh}(\sin(x)) = x - \frac{x^5}{15} + \frac{x^7}{90} + o(x^7)$.

On obtient finalement : $\sin(\operatorname{sh}(x)) - \operatorname{sh}(\sin(x)) = -\frac{x^7}{45} + o(x^7) \underset{0}{\sim} -\frac{x^7}{45}$.

– Plus généralement, considérons le développement $f(x) = \sum_{k=0}^n a_k(x-x_0)^k + o((x-x_0)^n)$.

Si tous les a_k sont nuls, alors $f(x)$ est négligeable devant $(x-x_0)^n$ au voisinage de x_0 .

Sinon, et si m est l'indice minimum tel que $a_m \neq 0$, alors $f(x) \sim a_m(x-x_0)^m$ en x_0 .

Inversement, si $f(x) \sim a_m(x-x_0)^m$ en x_0 , avec $m \in \mathbb{N}$, alors $f(x) = a_m(x-x_0)^m + o((x-x_0)^m)$.

Plus généralement, considérons par exemple le DL usuel : $\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} + o(x^4)$.

Il permet d'écrire les équivalents : $\cos x - 1 \underset{0}{\sim} -\frac{x^2}{2!}$, ou encore $\cos x - 1 + \frac{x^2}{2!} \underset{0}{\sim} \frac{x^4}{4!}$

9.6.2 Position par rapport à une tangente

On suppose que f admet un développement limité d'ordre $n \geq 3$ au point x_0 :

$$f(x) = a_0 + a_1(x-x_0) + \dots + a_n(x-x_0)^n + o((x-x_0)^n).$$

On sait que cela implique la dérivabilité de f en x_0 , avec $f(x_0) = a_0$ et $f'(x_0) = a_1$.

L'équation de la tangente Δ à la courbe $y = f(x)$ en $x = x_0$ est donc $y = a_0 + a_1(x-x_0)$.

Remarque : si le DL n'est valable qu'à gauche ou à droite de x_0 , il s'agit d'une demi-tangente.

Soit m l'indice minimum tel que $m \geq 2$ et $a_m \neq 0$.

Alors $f(x) - a_0 - a_1(x-x_0) \sim a_m(x-x_0)^m$ au voisinage de x_0 .

On en déduit le placement local de la courbe $y = f(x)$ par rapport à la (demi-)tangente Δ .

– Si m est pair, le placement de $y = f(x)$ par rapport à Δ est donné par le signe de a_m .

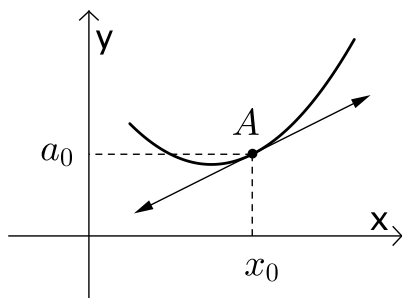
Si $a_m > 0$, la courbe est localement "au-dessus" de sa tangente.

Si $a_m < 0$, la courbe est localement "en-dessous" de sa tangente.

– Si m est impair, la courbe $y = f(x)$ "traverse" Δ au voisinage de M_0 .

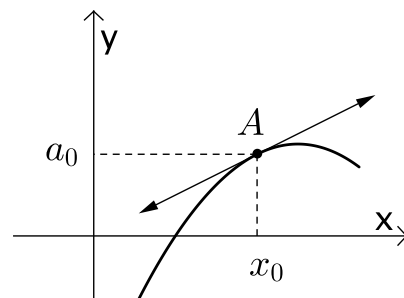
Dans ce cas, la droite Δ est donc une tangente d'inflexion.

Voici les allures possibles au voisinage de $A(x_0, a_0)$, quand $a_1 = f'(x_0)$ est strictement positif.



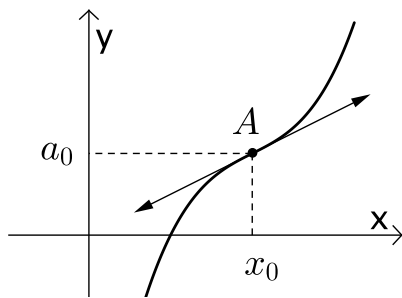
$$f(x_0 + h) = a_0 + a_1h + a_mh^m + o(h^m)$$

$$a_1 > 0, m \text{ pair}, a_m > 0$$



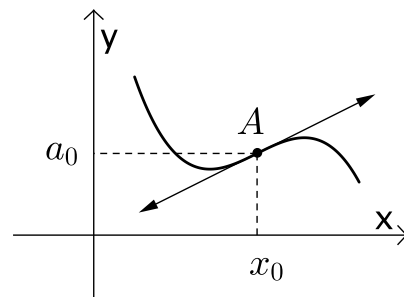
$$f(x_0 + h) = a_0 + a_1h + a_mh^m + o(h^m)$$

$$a_1 > 0, m \text{ pair}, a_m < 0$$



$$f(x_0 + h) = a_0 + a_1h + a_mh^m + o(h^m)$$

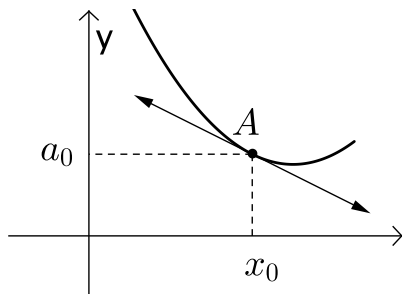
$$a_1 > 0, m \text{ impair}, a_m > 0$$



$$f(x_0 + h) = a_0 + a_1h + a_mh^m + o(h^m)$$

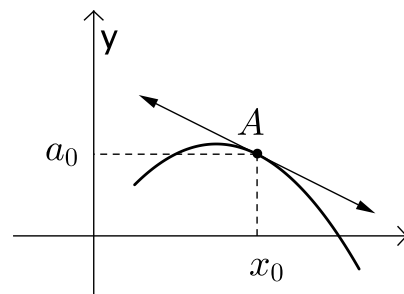
$$a_1 > 0, m \text{ impair}, a_m < 0$$

Voici les allures possibles au voisinage de $A(x_0, a_0)$, quand $a_1 = f'(x_0)$ est strictement négatif.



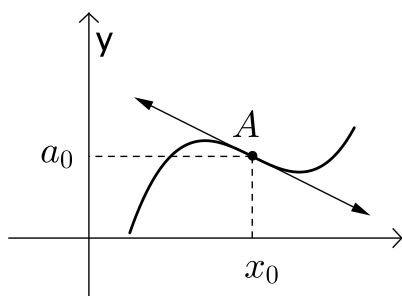
$$f(x_0 + h) = a_0 + a_1h + a_mh^m + o(h^m)$$

$$a_1 < 0, m \text{ pair}, a_m > 0$$



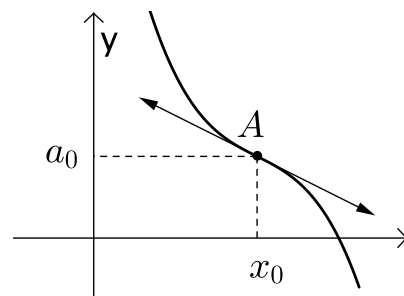
$$f(x_0 + h) = a_0 + a_1h + a_mh^m + o(h^m)$$

$$a_1 < 0, m \text{ pair}, a_m < 0$$



$$f(x_0 + h) = a_0 + a_1h + a_mh^m + o(h^m)$$

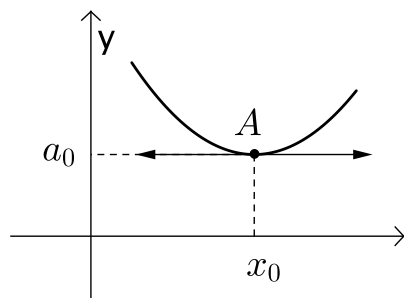
$$a_1 < 0, m \text{ impair}, a_m > 0$$



$$f(x_0 + h) = a_0 + a_1h + a_mh^m + o(h^m)$$

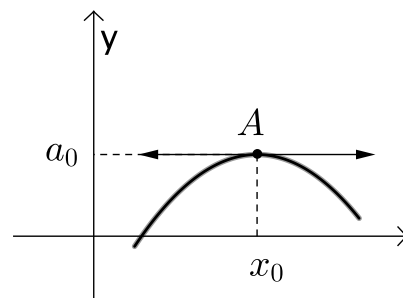
$$a_1 < 0, m \text{ impair}, a_m < 0$$

Enfin, voici les allures possibles au voisinage de $A(x_0, a_0)$, quand $a_1 = f'(x_0)$ est nul.



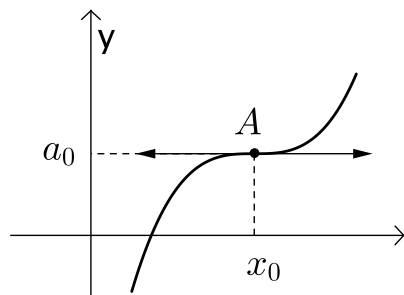
$$f(x_0 + h) = a_0 + a_m h^m + o(h^m)$$

$$m \text{ pair, } a_m > 0$$



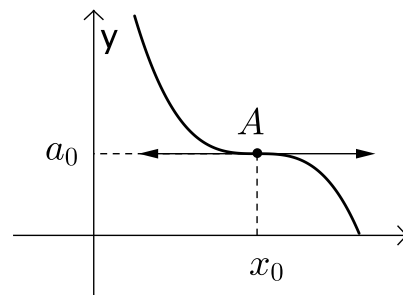
$$f(x_0 + h) = a_0 + a_m h^m + o(h^m)$$

$$m \text{ pair, } a_m < 0$$



$$f(x_0 + h) = a_0 + a_m h^m + o(h^m)$$

$$m \text{ impair, } a_m > 0$$



$$f(x_0 + h) = a_0 + a_m h^m + o(h^m)$$

$$m \text{ impair, } a_m < 0$$

Au vu des douze cas précédents, on comprend qu'une condition *nécessaire* pour que f admette un extrémum local en x_0 est que le coefficient de degré 1 dans le développement de $f(x_0 + h)$ soit nul.

Ce n'est pas une condition *suffisante* comme on le voit avec les deux derniers cas ci-dessus.

En revanche, si $f(x_0 + h) = a_0 + a_m h^m + o(h^m)$, avec $m \geq 2$ et pair (en général $m = 2$) et $a_m \neq 0$, alors la fonction f présente en x_0 un extrémum local (un minimum si $a_m > 0$, un maximum si $a_m < 0$).

Une remarque sur les DL en dehors de l'origine

On considère le DL : $f(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \dots + a_n(x - x_0)^n + o((x - x_0)^n)$, où $x_0 \neq 0$.

Dans un tel développement, la différence $x - x_0$ constitue « l'infiniment principal », et la présence des puissances successives de $(x - x_0)$ illustre une approximation locale de $f(x)$ (de plus en plus précise au fur et à mesure de l'avancement dans ce développement).

Dans une telle écriture, il est important de conserver l'ordre des puissances croissantes de $(x - x_0)$.

Surtout il ne faut **jamais** développer les $(x - x_0)^k$ quand $k \geq 2$.

En revanche, on rappelle que $y = a_0 + a_1(x - x_0) = a_1x + (a_0 - a_1x_0)$ représente l'équation de la tangente à la courbe représentative de f au point $A(x_0, a_0 = f(x_0))$ à la courbe $y = f(x)$. On peut réordonner cette partie du développement sans risque.

9.6.3 Branches infinies

Définition 9.6.1 (développement limité au voisinage de $\pm\infty$)

Soit $f : I \rightarrow \mathbb{R}$ une fonction définie au voisinage de $+\infty$ (ou de $-\infty$). Soit n dans \mathbb{N} .

On dit que f a un développement limité (un DL) à l'ordre n en $+\infty$ (resp. en $-\infty$) s'il existe des réels a_0, a_1, \dots, a_n et une fonction $x \mapsto \varepsilon(x)$ tels que : $\forall x \in I, f(x) = \sum_{k=0}^n \frac{a_k}{x^k} + \frac{\varepsilon(x)}{x^n}$, avec $\lim_{x \rightarrow \infty} \varepsilon(x) = 0$.

Remarques

- Avec les notations de Landau, le développement précédent s'écrit : $f(x) = \sum_{k=0}^n \frac{a_k}{x^k} + o\left(\frac{1}{x^n}\right)$.
- Les DL en ∞ ont des propriétés analogues à celles des DL en x_0 (unicité, troncature, etc.)
- D'ailleurs un DL en ∞ peut toujours se ramener à un DL en 0 en posant $y = \frac{1}{x}$.

En effet, si g est définie au voisinage de $y = 0$ par $g(y) = f\left(\frac{1}{y}\right)$, on a l'équivalence :

$$f(x) = \sum_{k=0}^n \frac{a_k}{x^k} + o\left(\frac{1}{x^n}\right) \Leftrightarrow g(y) = \sum_{k=0}^n a_k y^k + o(y^n)$$

Placement par rapport à une asymptote

On suppose qu'au voisinage de $\pm\infty$ on a le développement : $\frac{f(x)}{x} = a_0 + \frac{a_1}{x} + \dots + \frac{a_n}{x^n} + o\left(\frac{1}{x^n}\right)$.

Alors $f(x) = a_0 x + a_1 + \frac{a_2}{x} + \dots + \frac{a_n}{x^{n-1}} + o\left(\frac{1}{x^{n-1}}\right)$

En particulier, on observe que $\lim_{x \rightarrow \pm\infty} (f(x) - a_0 x - a_1) = 0$.

La droite Δ d'équation $y = a_0 x + a_1$ est donc asymptote à la courbe $y = f(x)$ quand $x \rightarrow \pm\infty$.

Soit m l'indice minimum tel que $m \geq 2$ et $a_m \neq 0$. Alors $f(x) - a_0 x - a_1 \sim \frac{a_m}{x^{m-1}}$.

On en déduit le placement de la courbe $y = f(x)$ par rapport à Δ au voisinage de $\pm\infty$.

Un exemple de recherche d'asymptote

On considère la fonction f_λ définie sur \mathbb{R}^* par $f_\lambda(x) = (2x - \lambda)e^{1/x}$, où λ est un paramètre réel.

On effectue le changement de variable $x = \frac{1}{y}$ pour se ramener au voisinage de $y = 0$. On trouve :

$$f_\lambda(x) = \left(\frac{2}{y} - \lambda\right)e^y = \frac{1}{y}(2 - \lambda y) \left(1 + y + \frac{y^2}{2} + \frac{y^3}{6} + o(y^3)\right) = \frac{1}{y} \left(2 + (2 - \lambda)y + (1 - \lambda)y^2 + \left(\frac{1}{3} - \frac{\lambda}{2}\right)y^3 + o(y^3)\right)$$

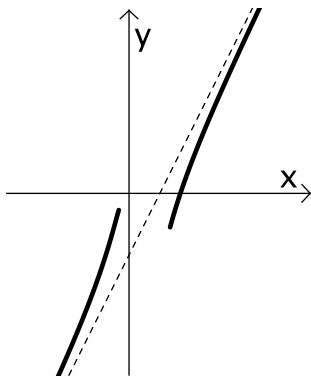
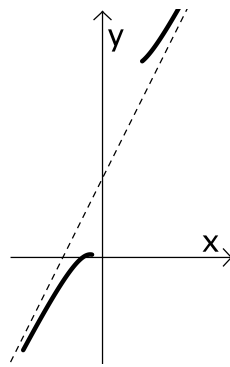
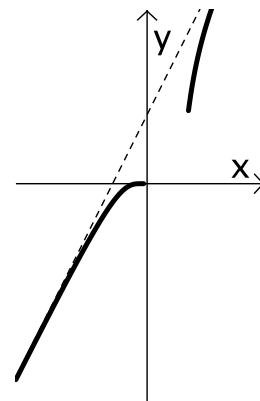
Ainsi, de retour à la variable x , on obtient : $f_\lambda(x) = 2x + 2 - \lambda + \frac{1 - \lambda}{x} + \left(\frac{1}{3} - \frac{\lambda}{2}\right)\frac{1}{x^2} + o\left(\frac{1}{x^2}\right)$

On en déduit la présence de l'asymptote Δ_λ d'équation $y = 2x + 2 - \lambda$ quand x tend vers $\pm\infty$.

Le placement est donné par le signe de $\frac{1 - \lambda}{x}$ si $\lambda \neq 1$, et par celui de $\left(\frac{1}{3} - \frac{\lambda}{2}\right)\frac{1}{x^2} = -\frac{1}{6x^2}$ si $\lambda = 1$.

Remarque : à cause de $\lambda = 1$, on ne pouvait contenter de $f_\lambda(x) = 2x + 2 - \lambda + \frac{1 - \lambda}{x} + o\left(\frac{1}{x}\right)$.

Voici l'allure de la courbe $y = f(x)$ dans les trois cas $\lambda > 1$, $\lambda < 1$, et $\lambda = 1$:

allure quand $\lambda > 1$ allure quand $\lambda < 1$ allure quand $\lambda = 1$

9.7 Exemples de développements asymptotiques

9.7.1 Position du problème

Considérons le développement limité usuel, quand x tend vers 0 : $\sin(x) = x - \frac{x^3}{6} + \frac{x^5}{120} + o(x^6)$.

Si on remplace x par \sqrt{x} , on obtient : $\sin(\sqrt{x}) = \sqrt{x} - \frac{x\sqrt{x}}{6} + \frac{x^2\sqrt{x}}{24} + o(x^3)$

Le résultat obtenu n'est pas un développement limité au sens qu'on a donné à ce terme (car sa partie régulière n'est pas un polynôme en x). On parle plutôt ici de « développement asymptotique ».

Plus généralement, un développement asymptotique en un point x_0 nécessite ce qu'on appelle une « échelle de comparaison ». Celle-ci est formée d'une suite de fonctions $x \mapsto \varphi_n(x)$, avec la particularité que, pour tout entier n , $\varphi_{n+1}(x)$ est négligeable devant $\varphi_n(x)$ quand x tend vers x_0 .

Un développement asymptotique de f en x_0 , à la précision $o(\varphi_n(x))$, est alors une écriture :

$$f(x) = a_0\varphi_0(x) + a_1\varphi_1(x) + a_2\varphi_2(x) + \cdots + a_n\varphi_n(x) + o(\varphi_n(x))$$

On note que les développements usuels en x_0 sont un cas particulier de cette situation, quand on utilise l'échelle de comparaison en x_0 définie par les $\varphi_n(x) = (x - x_0)^n$.

Il est également possible de définir des développements asymptotiques au voisinage de $+\infty$ ou de $-\infty$.

Les échelles de comparaisons les plus utilisées en 0 ou en ∞ sont :

- les fonctions puissances de x ou de $\frac{1}{x}$ à exposant rationnel
- les fonctions puissances de $\ln x$ ou des fonctions du type $x \mapsto x^p \ln^q(x)$, avec p, q dans \mathbb{Z} .

Le programme de mathématiques en MPSI précise que les échelles de comparaison sont hors-programme, et qu'on doit se contenter de quelques exemples de développements asymptotiques « simples ».

9.7.2 Quelques exemples simples

Développement asymptotique, quand $x \rightarrow +\infty$, à la précision $o\left(\frac{1}{x^4}\right)$, de $f(x) = \frac{1}{x + \ln(x)}$.

On écrit $f(x) = \frac{1}{x} \left(\frac{1}{1 + y(x)} \right)$, avec $y(x) = \frac{\ln(x)}{x}$. On sait que $\lim_{x \rightarrow +\infty} y(x) = 0$.

On peut donc utiliser le développement : $\frac{1}{1 + y} = 1 - y + y^2 - y^3 + O(y^4)$.

Ainsi $f(x) = \frac{1}{x} \left(1 - \frac{\ln(x)}{x} + \frac{\ln^2(x)}{x^2} - \frac{\ln^3(x)}{x^3} + O\left(\frac{\ln^4(x)}{x^4}\right) \right) = \frac{1}{x} - \frac{\ln(x)}{x^2} + \frac{\ln^2(x)}{x^3} - \frac{\ln^3(x)}{x^4} + O\left(\frac{\ln^4(x)}{x^5}\right)$

Enfin, $\frac{\ln^4(x)}{x^5} = \frac{1}{x^4} \frac{\ln^4(x)}{x} = o\left(\frac{1}{x^4}\right)$ quand $x \rightarrow +\infty$.

Finalement : $f(x) = \frac{1}{x} - \frac{\ln(x)}{x^2} + \frac{\ln^2(x)}{x^3} - \frac{\ln^3(x)}{x^4} + o\left(\frac{1}{x^4}\right)$.

Il est important de noter que, dans le développement précédent, chacun des termes qui figurent dans l'échelle de comparaison est négligeable (au voisinage de $+\infty$) devant celui qui le précède (dans une lecture de gauche à droite, bien sûr).

Développement asymptotique, quand $x \rightarrow 0$, à la précision $o(x^3)$, de $f(x) = \frac{1}{x + \ln(x)}$.

On écrit $f(x) = \frac{1}{\ln(x)} \left(\frac{1}{1 + y(x)} \right)$, avec $y(x) = \frac{x}{\ln(x)}$. On sait que $\lim_{x \rightarrow 0} y(x) = 0$.

On peut donc utiliser le développement : $\frac{1}{1 + y} = 1 - y + y^2 + O(y^3)$.

Ainsi $f(x) = \frac{1}{\ln(x)} \left(1 - \frac{x}{\ln(x)} + \frac{x^2}{\ln^2(x)} + O\left(\frac{x^3}{\ln^3(x)}\right) \right) = \frac{1}{\ln(x)} - \frac{x}{\ln^2(x)} + \frac{x^2}{\ln^3(x)} + O\left(\frac{x^3}{\ln^4(x)}\right)$.

Enfin, $\frac{x^3}{\ln^4(x)} = o(x^3)$ quand $x \rightarrow 0$. Finalement : $f(x) = \frac{1}{\ln(x)} - \frac{x}{\ln^2(x)} + \frac{x^2}{\ln^3(x)} + o(x^3)$.

Développement asymptotique, quand $x \rightarrow 0$, à la précision $o(x^3)$, de $f(x) = \sqrt{x + \sqrt{x}}$.

On écrit $f(x) = x^{1/4} \sqrt{1 + \sqrt{x}}$, et on utilise $\sqrt{1 + y} = 1 + \frac{y}{2} - \frac{y^2}{8} + \frac{y^3}{16} - \frac{5y^4}{128} + \frac{7y^5}{256} + O(y^6)$.

Il en résulte : $\sqrt{1 + \sqrt{x}} = 1 + \frac{x^{1/2}}{2} - \frac{x}{8} + \frac{x^{3/2}}{16} - \frac{5x^2}{128} + \frac{7x^{5/2}}{256} + O(x^3)$.

Finalement : $f(x) = x^{1/4} + \frac{x^{3/4}}{2} - \frac{x^{5/4}}{8} + \frac{x^{7/4}}{16} - \frac{5x^{9/4}}{128} + \frac{7x^{11/4}}{256} + o(x^3)$.

Développement asymptotique, quand $x \rightarrow +\infty$, à la précision $o\left(\frac{1}{x^2}\right)$, de $f(x) = \sqrt{x + \sqrt{x}}$.

On écrit $f(x) = \sqrt{x} \left(1 + \frac{1}{\sqrt{x}} \right)^{1/2}$, et on utilise $\sqrt{1 + y} = 1 + \frac{y}{2} - \frac{y^2}{8} + \frac{y^3}{16} - \frac{5y^4}{128} + \frac{7y^5}{256} + o(y^5)$.

Il en résulte : $\left(1 + \frac{1}{\sqrt{x}} \right)^{1/2} = 1 + \frac{1}{2\sqrt{x}} - \frac{1}{8x} + \frac{1}{16x\sqrt{x}} - \frac{5}{128x^2} + \frac{7}{256x^2\sqrt{x}} + o\left(\frac{1}{x^2\sqrt{x}}\right)$.

Finalement : $f(x) = \sqrt{x} + \frac{1}{2} - \frac{1}{8\sqrt{x}} + \frac{1}{16x} - \frac{5}{128x\sqrt{x}} + \frac{7}{256x^2} + o\left(\frac{1}{x^2}\right)$.

Déterminer un développement, quand $x \rightarrow 0$, à la précision $o(x^2)$, de $f(x) = x^{\ln(1+x)}$.

On écrit : $f(x) = \exp(\ln(x) \ln(1+x)) = \exp(y(x))$, avec $y(x) = x \ln(x) \left(1 - \frac{x}{2} + O(x^2)\right)$.

On a : $y(x) \underset{0}{\sim} x \ln(x)$, et en particulier : $\lim_{x \rightarrow 0} y(x) = 0$.

On peut donc utiliser le développement usuel : $e^y = 1 + y + \frac{y^2}{2} + O(y^3)$ quand $y \rightarrow 0$.

On trouve $y^2(x) = x^2 \ln^2(x) \left(1 - \frac{x}{2} + O(x^2)\right)^2 = x^2 \ln^2(x) \left(1 - x + O(x^2)\right)$

En utilisant $\lim_{x \rightarrow 0} x \ln^2(x) = 0$, le développement de $y^2(x)$ s'écrit : $y^2(x) = x^2 \ln^2(x) + o(x^2)$.

Par ailleurs : $y(x) = x \ln(x) - \frac{x^2}{2} \ln(x) + O(x^3 \ln(x)) = x \ln(x) - \frac{x^2}{2} \ln(x) + o(x^2)$.

Enfin, et puisque $y(x) \underset{0}{\sim} x \ln(x)$, un $O(y^3)$ est un $O(x^3 \ln^3(x))$ donc un $o(x^2)$.

Finalement, on trouve : $f(x) = 1 + x \ln(x) + \frac{x^2}{2} \ln^2(x) - \frac{x^2}{2} \ln(x) + o(x^2)$.

9.7.3 Étude d'une suite définie implicitement

On se propose de résoudre l'exercice suivant :

Montrer que l'équation $\ln(x) + x = n$ possède une unique solution, notée x_n , pour tout n dans \mathbb{N} .

Déterminer a, b, c tels que : $x_n = a n + b \ln(n) + c \frac{\ln(n)}{n} + o\left(\frac{\ln(n)}{n}\right)$ quand $n \rightarrow +\infty$.

– Tout d'abord, l'application $x \mapsto f(x) = x + \ln(x)$ est continue et strictement croissante sur \mathbb{R}^{+*} .

Puisque $\lim_{x \rightarrow 0^+} f(x) = -\infty$ et $\lim_{x \rightarrow +\infty} f(x) = +\infty$, l'application f est une bijection de \mathbb{R}^{+*} sur \mathbb{R} .

Pour tout n dans \mathbb{N} , il existe donc un unique $x_n > 0$ tel que $f(x_n) = n$.

Retenons que pour tout entier n , on a l'égalité (E_1) : $\ln(x_n) + x_n = n$.

– L'application réciproque f^{-1} est une bijection strictement croissante de \mathbb{R} sur \mathbb{R}^{+*} .

Pour tout n , on a $x_n = f^{-1}(n)$. On en déduit $\lim_{n \rightarrow +\infty} x_n = +\infty$.

– Pour tout entier n , l'égalité (E_1) s'écrit aussi (E_2) : $\frac{n}{x_n} = \frac{\ln(x_n)}{x_n} + 1$.

Puisque $\lim_{n \rightarrow +\infty} x_n = +\infty$, un passage à la limite dans (E_2) donne $\lim_{n \rightarrow +\infty} \frac{n}{x_n} = 1$ donc $x_n \sim n$.

– Posons $x_n = n + y_n$, donc $y_n = x_n - n$.

Puisque $x_n \sim n$, cette définition nous dit que $y_n = o(n)$ en $+\infty$.

L'égalité (E_1) devient (E_3) : $\ln(n + y_n) + y_n = 0$, donc $y_n = -\ln(n + o(n)) \underset{+\infty}{\sim} -\ln(n)$.

– Posons alors $y_n = -\ln(n) + z_n$, donc $x_n = n - \ln(n) + z_n$.

Avec cette définition, et sachant que $y_n \underset{+\infty}{\sim} -\ln(n)$, on a $z_n = o(\ln(n))$.

On rappelle que $\ln(x_n) + x_n = n$, pour tout n de \mathbb{N} .

On effectue un développement de $\ln(x_n)$ quand $n \rightarrow +\infty$:

$$\ln(x_n) = \ln(n - \ln(n) + o(\ln(n))) = \ln(n) + \ln\left(1 - \frac{\ln(n)}{n} + o\left(\frac{\ln(n)}{n}\right)\right) = \ln(n) - \frac{\ln(n)}{n} + o\left(\frac{\ln(n)}{n}\right)$$

Ainsi : $x_n = n - \ln(x_n) = n - \ln(n) + \frac{\ln(n)}{n} + o\left(\frac{\ln(n)}{n}\right)$.

– On a donc obtenu le résultat souhaité avec $a = 1$, $b = -1$ et $c = 1$.

9.7.4 Formule de Stirling

Proposition 9.7.1 (formule de Stirling)

Quand l'entier n tend vers $+\infty$, on a l'équivalent : $n! \sim n^n e^{-n} \sqrt{2\pi n}$.

Le résultat précédent est à connaître, mais pas sa démonstration (qui constitue un problème classique).

Puisque : $n! = (n^n e^{-n} \sqrt{2\pi n})(1 + o(1))$, on a le développement asymptotique, quand $n \rightarrow +\infty$:

$$\ln(n!) = n \ln(n) - n + \frac{\ln(n)}{2} + \frac{\ln(2\pi)}{2} + o(1)$$

Et ce développement donne les équivalents : $\ln(n!) \underset{+\infty}{\sim} n \ln(n)$, ou encore $\ln\left(\frac{n^n}{n!}\right) \underset{+\infty}{\sim} n$.

Chapitre 10

Arithmétique dans \mathbb{Z}

Sommaire

10.1	Divisibilité et division euclidienne	231
10.1.1	Divisibilité dans \mathbb{Z} , diviseurs, multiples	231
10.1.2	Théorème de la division euclidienne	232
10.2	Pgcd et algorithme d'Euclide	233
10.2.1	Pgcd de a, b dans \mathbb{N} (avec $a \neq 0$ ou $b \neq 0$)	233
10.2.2	Algorithme d'Euclide	235
10.2.3	Un peu de programmation Python	236
10.2.4	Quelques propriétés du pgcd	237
10.2.5	Relation de Bézout	238
10.2.6	Ppcm de deux entiers	240
10.3	Entiers premiers entre eux	242
10.3.1	Couples d'entiers premiers entre eux	242
10.3.2	Le théorème de Bézout et ses conséquences	242
10.3.3	Pgcd de plusieurs entiers	243
10.3.4	Entiers premiers entre eux dans leur ensemble	244
10.4	Nombres premiers	245
10.4.1	Définition et « premières » propriétés	245
10.4.2	Décomposition en produit de facteurs premiers	247
10.4.3	Un peu de programmation Python	249
10.5	Congruences	252
10.5.1	Congruence modulo un entier sur \mathbb{Z}	252
10.5.2	Opérations sur les congruences	252

10.1 Divisibilité et division euclidienne

10.1.1 Divisibilité dans \mathbb{Z} , diviseurs, multiples

Définition 10.1.1

Soient a et b deux entiers relatifs.

On dit que b *divise* a , ou encore b est un *diviseur* de a , ou encore a est *divisible* par b , ou encore a est un *multiple* de b , s'il existe q dans \mathbb{Z} tel que $a = qb$. On note alors $b \mid a$.

Définition 10.1.2

Pour tout n de \mathbb{Z} , on note $n\mathbb{Z} = \{qn, q \in \mathbb{Z}\}$ l'ensemble des multiples de n .

On note $\mathcal{D}(n)$ l'ensemble des diviseurs de n dans \mathbb{Z} .

L'ensemble $\mathcal{D}(n) \cap \mathbb{N}$ est donc l'ensemble des diviseurs positifs ou nuls de n .

Pour tous a, b dans \mathbb{Z} , on a les équivalences : $b \mid a \Leftrightarrow a \in b\mathbb{Z} \Leftrightarrow b \in \mathcal{D}(a)$.

Remarques sur la relation de divisibilité

En écrivant $a \mid b$, on définit une relation binaire sur l'ensemble \mathbb{Z} .

Cette relation est réflexive (on a toujours $a \mid a$) et transitive (si $a \mid b$ et $b \mid c$, alors $a \mid c$).

Mais ce n'est pas une relation d'ordre car elle n'est pas antisymétrique.

Plus précisément, on a l'équivalence : $(a \mid b \text{ et } b \mid a) \Leftrightarrow a = \pm b$.

En revanche, la restriction de la relation de divisibilité à \mathbb{N} est une relation d'ordre (partiel).

On dit de deux entiers relatifs qui se divisent mutuellement (c'est-à-dire : qui sont égaux ou opposés, ou encore : qui ont la même valeur absolue) qu'ils sont *associés*.

On retiendra que deux entiers associés ont exactement les mêmes propriétés par rapport à la relation de divisibilité (et c'est la raison pour laquelle on choisira souvent celui des deux qui est positif).

Quelques cas particuliers

– La notation $2\mathbb{Z}$ désigne l'ensemble des entiers pairs, et on a $\mathcal{D}(2) = \{-2, -1, 1, 2\}$.

Deux exemples : $\mathcal{D}(12) = \{-12, -6, -4, -3, -2, -1, 1, 2, 3, 4, 6, 12\}$, et $\mathcal{D}(13) = \{-13, -1, 1, 13\}$

De même : $12\mathbb{Z} = \{12k, k \in \mathbb{Z}\} = \{\dots, -36, -24, -12, 0, 12, 24, 36, \dots\}$.

– L'entier 0 est multiple de tout entier b (car $0 = 0b$), ou encore : tout entier b divise 0.

En revanche, 0 ne divise que lui-même (car $a = q0 \Rightarrow a = 0$), ou encore : le seul multiple de 0 est 0.

On peut donc écrire : $\mathcal{D}(0) = \mathbb{Z}$, et $0\mathbb{Z} = \{0\}$.

– Les entiers 1 et -1 divisent tout entier a (considérer $a = a1 = (-a)(-1)$).

Ainsi tout entier est multiple de 1 et -1 .

On peut donc écrire : $1\mathbb{Z} = (-1)\mathbb{Z} = \mathbb{Z}$.

En revanche les seuls diviseurs de -1 et 1 sont -1 et 1 eux-mêmes.

Ainsi $\mathcal{D}(1) = \mathcal{D}(-1) = \{-1, 1\}$.

Propriétés diverses sur les ensembles $\mathcal{D}(n)$ et $n\mathbb{Z}$.

- Il nous arrivera souvent de considérer $\mathcal{D}(n) \cap \mathbb{N}$, ensemble des diviseurs positifs ou nuls de n .
- Les ensembles $\mathcal{D}(n)$ et $n\mathbb{Z}$ ne sont jamais vides, car ils contiennent n .

Les ensembles $\mathcal{D}(n)$ et $n\mathbb{Z}$ sont tous deux symétriques par rapport à l'origine.

Si $n \neq 0$: l'ensemble $\mathcal{D}(n)$ est inclus dans $\llbracket -n, n \rrbracket$ donc est fini ; en revanche $n\mathbb{Z}$ est infini.

- Pour tous entiers relatifs a et b , on a les équivalences : $a\mathbb{Z} \subset b\mathbb{Z} \Leftrightarrow b \mid a \Leftrightarrow \mathcal{D}(b) \subset \mathcal{D}(a)$.

On en déduit : $a\mathbb{Z} = b\mathbb{Z} \Leftrightarrow |a| = |b| \Leftrightarrow \mathcal{D}(a) = \mathcal{D}(b)$.

Les égalités $(-n)\mathbb{Z} = n\mathbb{Z}$ et $\mathcal{D}(-n) = \mathcal{D}(n)$ font qu'on se limite souvent à $n \geq 0$.

- Si a, b sont dans $n\mathbb{Z}$, et si u, v sont dans \mathbb{Z} , alors $au + bv$ est dans $n\mathbb{Z}$.

Une telle propriété est fautive pour $\mathcal{D}(n)$: par exemple 3 et 5 sont dans $\mathcal{D}(15)$ mais $3 + 5$ n'y est pas.

Comparaisons dans \mathbb{N} entre ordre naturel et ordre pour la divisibilité.

- On sait que \mathbb{N} est muni de sa relation d'ordre naturelle, notée \leq .

C'est un ordre total pour lequel $\min(\mathbb{N}) = 0$, mais $\max(\mathbb{N})$ n'existe pas.

On sait que toute partie non vide de \mathbb{N} possède un plus petit élément.

On sait également que toute partie finie non vide de \mathbb{N} possède un plus grand élément.

- Passons maintenant à la relation de divisibilité sur \mathbb{N} .

C'est une relation d'ordre partiel (il y a des entiers non comparables, par exemple 2 et 3).

Pour la relation de divisibilité, le minimum de \mathbb{N} est 1 (car 1 divise tous les éléments de \mathbb{N}).

Pour la relation de divisibilité, le maximum de \mathbb{N} est 0 (car 0 est un multiple de tout n de \mathbb{N}).

Une partie non vide de \mathbb{N} peut très bien ne pas posséder de plus petit élément (au sens de la division!).

Par exemple :

- Si $A = \{6, 9, 12, 15, 18\}$: pour la divisibilité, il n'y a ni minimum ni maximum.
- Si $A = \{3, 6, 9, 12, 15, 18\}$: pour la divisibilité, l'entier 3 est minimum mais il n'y a pas de maximum.
- Si $A = \{6, 9, 12, 15, 18, 180\}$: pour la divisibilité, il n'y a pas de minimum, et 180 est maximum.
- Si $A = \{3, 6, 9, 12, 15, 18, 180\}$: pour la divisibilité, 3 est minimum, et 180 est maximum.
- Si $A = \{0, 1, 3, 6, 9, 12, 15, 18, 180\}$: pour la divisibilité, 1 est minimum, et 0 est maximum!

- On n'est donc pas à l'abri d'une ambiguïté quand on parle de divisibilité dans \mathbb{N} et qu'on utilise les expressions « plus petit que », « plus grand que », « maximum », « minimum ».

Remarquons tout de même que, pour (a, b) dans $\mathbb{N}^* \times \mathbb{N}^*$, on a : $a \mid b \Rightarrow a \leq b$.

Mais dans \mathbb{N}^2 , l'implication $a \mid b \Rightarrow a \leq b$ est fautive à cause du cas $a \geq 1$ et $b = 0$.

10.1.2 Théorème de la division euclidienne**Définition 10.1.3** (division euclidienne dans \mathbb{Z})

Soit (a, b) dans $\mathbb{Z} \times \mathbb{N}^*$.

Il existe un unique couple (q, r) de $\mathbb{Z} \times \mathbb{N}$ tel que $a = qb + r$ et $0 \leq r < b$

Le passage du couple (a, b) au couple (q, r) s'appelle *division euclidienne* de a par b .

Dans cette division, a est le *dividende*, b le *diviseur*, q le *quotient*, et r le *reste*.

Rapport entre divisibilité et division euclidienne

Attention à ne pas confondre la relation de divisibilité avec la division euclidienne.

On rappelle que $a \mid b$ est une expression booléenne (vraie ou fausse) quels que soient a et b dans \mathbb{Z} .

En revanche, effectuer la division euclidienne de a par b suppose que b est strictement positif.

Plus précisément, soient a et b deux entiers, avec $b > 0$.

Dire que b divise a , c'est dire que le reste dans la division de a par b est nul.

Division euclidienne de a ou de $-a$ par b

Soit $a = qb + r$ la division euclidienne de a par b .

Si a est multiple de b (c'est-à-dire $r = 0$), la division euclidienne de $-a$ par b s'écrit : $-a = (-q)b$.

Sinon (donc si $1 \leq r \leq b - 1$), elle s'écrit : $-a = q'b + r'$, avec $q' = -q - 1$ et $r' = b - r$.

Par exemple, avec $a = 2014$ et $b = 27$, on a : $2014 = 74 \cdot 27 + 16$ et $-2014 = -74 \cdot 27 - 16 = (-75) \cdot 27 + 11$.

On vérifie tous ces résultats avec la fonction `divmod` du langage Python :

```
>>> [divmod(2014,27), divmod(-2014,27)]
[(74, 16), (-75, 11)]
```

Le quotient entier est la partie entière du quotient

Soit q le quotient dans la division euclidienne de a par b (souvent appelé quotient entier de a par b).

Alors q est la partie entière du rationnel $x = \frac{a}{b}$.

En effet l'égalité $a = bq + r$ équivaut à $x = q + \frac{r}{b}$ et on a $0 \leq \frac{r}{b} < 1$, donc $q = \left\lfloor \frac{a}{b} \right\rfloor$.

On rappelle que l'opérateur `//` de Python fournit ce quotient entier, mais que l'opérateur `/` renvoie un flottant (confusion classique, d'autant que la syntaxe de Python a changé entre les versions 2 et 3).

```
>>> [2014//27, -2014//27, 2014/27]
[74, -75, 74.5925925925926]
```

10.2 Pgcd et algorithme d'Euclide

10.2.1 Pgcd de a, b dans \mathbb{N} (avec $a \neq 0$ ou $b \neq 0$)

Rappelons que $\mathcal{D}(0) = \mathbb{Z}$ donc $\mathcal{D}(0) \cap \mathbb{N} = \mathbb{N}$ (pas de maximum pour l'ordre naturel, mais 0 est le maximum pour l'ordre défini par la divisibilité).

En revanche si $n > 0$, alors $\mathcal{D}(n) \cap \mathbb{N}$ est fini non vide, et son maximum (à la fois pour l'ordre naturel et pour l'ordre défini par la divisibilité) est n .

Dans la suite, on considère deux entiers a et b de \mathbb{N} dont l'un au moins est non nul.

Il en résulte que $\mathcal{D}(a) \cap \mathcal{D}(b) \cap \mathbb{N}$ est une partie finie non vide de \mathbb{N}^* (elle contient l'entier 1).

Définition 10.2.1 (pgcd de deux entiers naturels dont l'un au moins est non nul)

Soit a et b deux entiers naturels, dont l'un au moins est non nul.

Soit $\mathcal{D}(a) \cap \mathcal{D}(b) \cap \mathbb{N}$ l'ensemble des diviseurs communs de a et b dans \mathbb{N} .

Cet ensemble, fini et inclus dans \mathbb{N}^* , possède un maximum $n \geq 1$ (pour l'ordre naturel).

Cet entier strictement positif n est appelé pgcd de a et de b . Il est noté $n = \text{pgcd}(a, b)$, ou $n = a \wedge b$.

Choix d'une notation

Ainsi $a \wedge b = \max(\mathcal{D}(a) \cap \mathcal{D}(b) \cap \mathbb{N})$, où le « max » s'entend au sens de l'ordre naturel.

Mais on verra bientôt (avec l'algorithme d'Euclide) que $a \wedge b$ est aussi le maximum de $\mathcal{D}(a) \cap \mathcal{D}(b) \cap \mathbb{N}$ au sens de l'ordre défini par la divisibilité (en d'autres termes, non seulement $d \leq a \wedge b$ pour tout diviseur positif d de a et b , mais mieux : $d \mid a \wedge b$).

Le nom « pgcd » est bien sûr une abréviation de « plus grand commun diviseur ».

En anglais, la dénomination est « gcd » (*greatest common divisor*).

Il faut mieux utiliser la notation $a \wedge b$ qui a le mérite d'être universelle.

Le langage Python possède une fonction intégrée `gcd`, mais elle fait partie du module `fractions`.

De toutes façons, on verra qu'il est très simple de programmer soi-même cette fonction.

```
>>> from fractions import gcd
>>> gcd(1155,910)
35
```

Premières propriétés

Il est clair que $a \wedge b = b \wedge a$. De même, si $a > 0$, la définition donne $a \wedge 0 = a$.

Pour tout a de \mathbb{N} , on a : $\mathcal{D}(a) \cap \mathcal{D}(1) \cap \mathbb{N} = \{1\}$ donc $a \wedge 1 = 1$.

Par définition, l'entier $a \wedge b$ est un diviseur commun aux entiers a et b .

Les diviseurs de $a \wedge b$ sont donc également des diviseurs communs à a et b .

On verra que la réciproque est vraie : les diviseurs communs à a et b sont aussi des diviseurs de $a \wedge b$.

On a l'égalité $a \wedge b = a$ si et seulement si a est un diviseur de b .

Prolongement aux entiers relatifs

La définition utilisée ici, $a \wedge b = \max(\mathcal{D}(a) \cap \mathcal{D}(b) \cap \mathbb{N})$, est celle du programme de la classe de MPSI.

On pose $0 \wedge 0 = 0$, ce qui prolonge l'égalité $a \wedge 0 = a$ quand $a \geq 1$.

Attention : poser $0 \wedge 0 = 0$ ne cadre pas avec la définition $a \wedge b = \max(\mathcal{D}(a) \cap \mathcal{D}(b) \cap \mathbb{N})$ si on considère le max au sens de l'ordre naturel, mais ça marche encore si le max est celui de l'ordre défini par la divisibilité dans \mathbb{N} (en effet $\mathcal{D}(0) \cap \mathcal{D}(0) \cap \mathbb{N} = \mathbb{N}$ dont le maximum pour la divisibilité est 0).

Par ailleurs, on rappelle que $\mathcal{D}(a) = \mathcal{D}(-a)$ pour tout a de \mathbb{Z} .

On peut donc étendre la définition du pgcd aux entiers relatifs en posant $a \wedge b = |a| \wedge |b|$.

Dans la mesure où les propriétés relatives à la divisibilité (donc aux calculs de pgcd) ne dépendent pas du signe des entiers a et b , on supposera essentiellement que a et b sont des entiers naturels.

10.2.2 Algorithme d'Euclide

Proposition 10.2.1

Soit a, b et q trois entiers relatifs. Alors $\mathcal{D}(a) \cap \mathcal{D}(b) = \mathcal{D}(b) \cap \mathcal{D}(a - qb)$

Autrement dit, les diviseurs communs de a et b sont exactement ceux de b et $a - qb$.

Ce résultat a une conséquence immédiate qui est la clef de l'algorithme d'Euclide :

Proposition 10.2.2 (pgcd et division euclidienne)

Soit a, b deux entiers naturels, avec $b > 0$.

Soit $a = qb + r$ la division euclidienne de a par b . Alors $\mathcal{D}(a) \cap \mathcal{D}(b) = \mathcal{D}(b) \cap \mathcal{D}(r)$

Un application répétée de ce principe conduit au célèbre algorithme d'Euclide.

Proposition 10.2.3 (algorithme d'Euclide)

Soit a et b deux entiers naturels strictement positifs. On veut calculer $a \wedge b$.

On forme une suite finie d'entiers r_k , à commencer par $r_0 = a$ et $r_1 = b$.

Soit $k \geq 1$. On suppose que r_{k-1} et r_k sont connus.

Si $r_k > 0$, on note $r_{k-1} = q_k r_k + r_{k+1}$ la division euclidienne de r_{k-1} par r_k .

Sous l'hypothèse $r_k > 0$, on a donc défini r_{k+1} , avec $0 \leq r_{k+1} < r_k$.

La suite d'entiers naturels $(r_k)_{k \geq 1}$ est finie car elle est strictement décroissante.

Il existe donc un entier naturel n tel que $r_n > 0$ et $r_{n+1} = 0$. Avec ces notations, on a : $a \wedge b = r_n$.

Ainsi $a \wedge b$ est le dernier reste non nul dans cette succession de divisions.

Exemples

– Voici les huit divisions successives qui aboutissent à $14938 \wedge 9471 = 77$:

$14938 = 1 \cdot 9471 + 5467$	$9471 = 1 \cdot 5467 + 4004$	$5467 = 1 \cdot 4004 + 1463$	$4004 = 2 \cdot 1463 + 1078$
$1463 = 1 \cdot 1078 + 385$	$1078 = 2 \cdot 385 + 308$	$385 = 1 \cdot 308 + 77$	$308 = 4 \cdot 77$

– Le nombre de divisions est parfois bien moindre : tout dépend bien sûr des quotients successifs : la méthode est d'autant plus rapide que les quotients sont très supérieurs à 1...

Voici par exemple le pgcd de $a = 267914296$ et de $b = 317811$ en seulement trois divisions :

$267914296 = 842 \cdot 317811 + 317434$	$317811 = 1 \cdot 317434 + 377$	$317434 = 842 \cdot 377 + 0$
---	---------------------------------	------------------------------

– Dernier exemple, plus théorique : montrons que $(n^a - 1) \wedge (n^b - 1) = n^{a \wedge b} - 1$.

Soit $a = qb + r$ la division euclidienne de a par b (avec a, b dans \mathbb{N}^*).

Pour tout n de \mathbb{N}^* , on a $n^a - 1 = n^{qb+r} - 1 = (n^b)^q n^r - 1 = ((n^b)^q - 1)n^r + n^r - 1$

Mais l'entier $(n^b)^q - 1$ est factorisable (donc divisible) par $n^b - 1$.

On peut donc écrire $n^a - 1 = q'(n^b - 1) + r'$, avec $0 \leq r' = n^r - 1 < n^b - 1$.

Autrement dit le reste dans la division de $n^a - 1$ par $n^b - 1$ est $n^r - 1$.

Il en découle facilement (algorithmes d'Euclide en parallèle), que : $(n^a - 1) \wedge (n^b - 1) = n^{a \wedge b} - 1$.

Par exemple : $(2^{30} - 1) \wedge (2^{24} - 1) = 2^{30 \wedge 24} - 1 = 2^6 - 1 = 63$.

10.2.3 Un peu de programmation Python

On écrit une fonction Python très simple qui calcule le pgcd par l'algorithme d'Euclide :

```
def pgcd1(a, b):
    while b: (a, b) = (b, a % b) # tant que b non nul, remplacer (a,b) par (b,r)
    return a # renvoyer le dernier reste non nul, donc le pgcd
```

On vérifie que ça marche sur les trois exemples précédents :

```
>>> pgcd1(14938,9471)
77
>>> pgcd1(267914296,317811)
377
>>> pgcd1(2**30-1,2**24-1)
63
```

Avec un petit effort de programmation, on modifie la fonction précédente de façon à ce qu'elle renvoie non seulement le pgcd, mais aussi la description des étapes de l'algorithme d'Euclide :

```
def pgcd2(a, b):
    e = []
    while b:
        (q, r) = (a // b, a % b)
        e.append('{0}={1}*{2}+{3}'.format(a, q, b, r))
        (a, b) = (b, r)
    return (a, e)
```

Voici ce que renvoie la fonction `pgcd2`, toujours sur les exemples précédents :

```
>>> pgcd2(14938,9471)
(77, ['14938=1*9471+5467', '9471=1*5467+4004', '5467=1*4004+1463', '4004=2*1463+1078',
      '1463=1*1078+385', '1078=2*385+308', '385=1*308+77', '308=4*77+0'])
>>> pgcd2(267914296,317811)
(377, ['267914296=842*317811+317434', '317811=1*317434+377', '317434=842*377+0'])
>>> pgcd2(2**30-1,2**24-1)
(63, ['1073741823=64*16777215+63', '16777215=266305*63+0'])
```

L'algorithme d'Euclide est par nature *récuratif* (calculer le pgcd de a et b , c'est calculer celui de b et r , à moins que b soit nul auquel cas c'est fini). Voici une fonction Python qui utilise la récursivité :

```
def pgcd3(a, b):
    if b: return pgcd3(b, a % b)
    else: return a
```

On peut même écrire une version encore plus compacte (avec `lambda` et le « if ternaire ») :

```
pgcd4 = lambda a, b : pgcd4(b, a % b) if b else a
```

Voici maintenant une fonction Python qui renvoie l'ensemble des diviseurs d'un entier $n > 0$.

On initialise l'ensemble $d = \mathcal{D}(n)$ à la valeur $d = \emptyset$.

On utilise un compteur k qu'on incrémente à partir de 1, sans jamais atteindre \sqrt{n} .

À chaque diviseur k de n , on ajoute à la fois k et n/k à l'ensemble d (on a $1 \leq k < \sqrt{n} < n/k \leq n$). À la sortie de la boucle, il faut vérifier si la dernière valeur de k vaut exactement \sqrt{n} (ce qui se produit si n est un carré parfait) car alors n/k et k sont égaux (et il suffit d'ajouter k à l'ensemble d).

```
def diviseurs(n):
    d = set(); k = 1
    while k*k < n:
        if n % k == 0: d |= {k, n//k}
        k += 1
    if k*k == n: d.add(k)
    return sorted(d)
```

Voici par exemple la liste des diviseurs de l'entier 210

```
>>> diviseurs(210)
[1, 2, 3, 5, 6, 7, 10, 14, 15, 21, 30, 35, 42, 70, 105, 210]
```

10.2.4 Quelques propriétés du pgcd

On peut considérer que l'algorithme d'Euclide constitue une nouvelle définition du pgcd (c'est le dernier reste non nul dans la succession des divisions).

Les propriétés suivantes sont alors des conséquences de cette définition algorithmique du pgcd.

On sait par exemple (depuis la toute première définition) que l'entier $a \wedge b$ divise a et b .

Mais l'algorithme d'Euclide donne la réciproque : si un entier divise a et b , alors il divise leur pgcd.

Ceci constitue une caractérisation de l'entier strictement positif $n = a \wedge b$.

Proposition 10.2.4 (caractérisation du pgcd)

Soit a et b deux entiers naturels, dont l'un au moins est non nul. Soit $n = a \wedge b$.

L'entier strictement positif $n = a \wedge b$ est caractérisé par l'égalité $\mathcal{D}(n) = \mathcal{D}(a) \cap \mathcal{D}(b)$.

Interprétation et exemple

Soit a et b deux entiers naturels, dont l'un au moins est non nul. Soit $n = a \wedge b$.

- d'une part n est le maximum de $\mathcal{D}(a) \cap \mathcal{D}(b)$ au sens de la relation \leq sur \mathbb{N} .
- d'autre part n est le maximum de $\mathcal{D}(a) \cap \mathcal{D}(b)$ au sens de la relation de divisibilité $|$ sur \mathbb{N} .

Prenons l'exemple des deux entiers $a = 210$ et $b = 330$.

Avec la fonction `diviseurs` définie précédemment, on forme la liste des diviseurs de a puis celle de b .

On vérifie que le pgcd de a et b vaut 30.

On forme l'ensemble des diviseurs communs aux deux entiers a et b .

On vérifie qu'on obtient la même chose en demandant directement la liste des diviseurs de 30.

Dans cette liste, l'entier 30 est à la fois le maximum pour l'ordre habituel dans \mathbb{N} , mais il est aussi le maximum en terme de divisibilité car il est multiple de chacun des éléments de la liste.

```

>>> diviseurs(210)
[1, 2, 3, 5, 6, 7, 10, 14, 15, 21, 30, 35, 42, 70, 105, 210]
>>> diviseurs(330)
[1, 2, 3, 5, 6, 10, 11, 15, 22, 30, 33, 55, 66, 110, 165, 330]
>>> pgcd1(210,330)
30
>>> set(diviseurs(210)) & set(diviseurs(330))
{1, 2, 3, 5, 6, 10, 15, 30}
>>> diviseurs(30)
[1, 2, 3, 5, 6, 10, 15, 30]

```

La proposition suivante montre qu'il est possible d'effectuer une mise en facteur (ou de diviser par un facteur commun) dans un calcul de pgcd.

Proposition 10.2.5 (mise en facteur dans un calcul de pgcd)

Soit a et b deux entiers naturels, dont l'un au moins est non nul.

Pour tout entier strictement positif k , on a l'égalité : $(ka) \wedge (kb) = k(a \wedge b)$.

De même, si k est un diviseur commun à a et b , on a : $\frac{a}{k} \wedge \frac{b}{k} = \frac{a \wedge b}{k}$

10.2.5 Relation de Bézout

Proposition 10.2.6 (relation de Bézout entre deux entiers)

Soit a et b deux entiers naturels, dont l'un au moins est non nul.

Alors il existe des entiers relatifs u_0 et v_0 tels que $a \wedge b = au_0 + bv_0$.

Une telle égalité s'appelle une relation de Bézout entre a et b .

On dit que le couple (u_0, v_0) constitue un couple de coefficients de Bézout de a et b .

Proposition 10.2.7 (combinaisons linéaires de deux entiers)

Soit a et b deux entiers naturels, dont l'un au moins est non nul.

L'ensemble $\{au + bv, u \in \mathbb{Z}, v \in \mathbb{Z}\}$ est exactement égal à l'ensemble des multiples de $a \wedge b$.

Recherche d'un couple de coefficients de Bézout

Pour trouver un couple de coefficients de Bézout de deux entiers strictement positifs a et b , il suffit de « remonter les calculs » dans l'algorithme d'Euclide.

Voici comment procéder avec $a = 14938$ et $b = 9471$. Les calculs de la colonne 1 donnent $a \wedge b = 77$, puis ceux de la colonne 2 donnent $26a - 41b = 77$.

14938 = 1 · 9471 + 5467	77 = 385 - 308
9471 = 1 · 5467 + 4004	77 = 385 - (1078 - 2 · 385) = 3 · 385 - 1078
5467 = 1 · 4004 + 1463	77 = 3(1463 - 1078) - 1078 = -4 · 1078 + 3 · 1463
4004 = 2 · 1463 + 1078	77 = -4 · (4004 - 2 · 1463) + 3 · 1463 = 11 · 1463 - 4 · 4004
1463 = 1 · 1078 + 385	77 = 11 · (5467 - 4004) - 4 · 4004 = -15 · 4004 + 11 · 5467
1078 = 2 · 385 + 308	77 = -15 · (9471 - 5467) + 11 · 5467 = 26 · 5467 - 15 · 9471
385 = 1 · 308 + 77	77 = 26 · (14938 - 9471) - 15 · 9471 = 26 · 14938 - 41 · 9471
308 = 4 · 77	

Programmation d'une méthode récursive

Soit $a = bq + r$ la division euclidienne de a par b . On sait que $a \wedge b = b \wedge r$.

Supposons connu un couple (x, y) de coefficients de Bézout de b et r .

Autrement dit $bx + ry = b \wedge r$, et il en résulte $a \wedge b = b \wedge r = bx + (a - bq)y = ay + b(x - qy)$.

Ainsi $(x' = y, y' = x - qy)$ est un couple de coefficients de Bézout de a et b .

Le problème se résout donc par des appels récursifs, qui finissent par aboutir au cas de base $b = 0$ pour lequel une solution (x, y) de l'équation $ax + by = a \wedge b = a$ est $(1, 0)$:

```
def recbezout(a,b):
    if b == 0: return (1,0)          # si b=0, alors (x,y)=(1,0)
    else:
        (q, r) = divmod(a, b)       # division euclidienne a=bq+r
        (x, y) = recbezout(b, r)    # coeffs de Bézout de (b,r)
    return (y, x-q*y)               # on en déduit des coeffs de Bézout de (a,b)
```

On vérifie que ça marche, en reprenant l'exemple $a = 14938$, $b = 9471$.

```
>>> recbezout(14938, 9471)
(26, -41)
```

Programmation d'une méthode itérative

Notons (E_d) l'équation $ax + by = d$, où d est dans \mathbb{Z} .

Soit α et β dans \mathbb{Z} , avec $\beta \neq 0$, et soit $\alpha = q\beta + r$ la division euclidienne de α par β .

Soit (x_1, y_1) une solution de (E_α) et (x_2, y_2) une solution de (E_β) .

Les égalités $\begin{cases} ax_1 + by_1 = \alpha \\ ax_2 + by_2 = \beta \end{cases}$ impliquent $a(x_1 - qx_2) + b(y_1 - qy_2) = \alpha - q\beta = r$.

Autrement dit le couple $(x_3 = x_1 - qx_2, y_3 = y_1 - qy_2)$ est solution de E_r .

On remarque que $(1, 0)$ est une solution de (E_a) et que $(0, 1)$ est une solution de (E_b) .

Si on applique cette idée tout au long de l'algorithme d'Euclide de (a, b) , on forme, pour chacun des restes successifs r_k , une solution (x_k, y_k) de (E_{r_k}) . À la fin, avec $r_n = a \wedge b$, on obtient une solution (x_n, y_n) de l'équation (E_{r_n}) , c'est-à-dire de l'équation $ax + by = a \wedge b$.

Voici une fonction Python utilisant cette méthode :

```
def iterbezout(a,b):
    (x1, y1, x2, y2) = (1, 0, 0, 1)
    while b:
        (q, r) = divmod(a, b); (a, b) = (b, r)
        (x1, y1, x2, y2) = (x2, y2, x1-q*x2, y1-q*y2)
    return (x1,y1)
```

On vérifie, en reprenant l'exemple $a = 14938$ et $b = 9471$:

```

>>> iterbezout(14938,9471)
(26, -41)

```

10.2.6 Ppcm de deux entiers

Dans la suite, on considère deux entiers strictement positifs a et b .

L'intersection $a\mathbb{Z} \cap b\mathbb{Z} \cap \mathbb{N}^*$ désigne l'ensemble des multiples communs strictement positifs de a et b .

Cet ensemble est non vide, car il contient le produit ab .

Il possède donc un plus petit entier strictement positif, ce qui conduit à la définition suivante :

Définition 10.2.2 (ppcm de deux entiers strictement positifs)

Soit a et b deux entiers naturels strictement positifs.

Soit n le minimum de l'ensemble $a\mathbb{Z} \cap b\mathbb{Z} \cap \mathbb{N}^*$ des multiples communs strictement positifs de a et b .

L'entier strictement positif n est appelé ppcm de a et de b . Il est noté $n = \text{ppcm}(a, b)$, ou $n = a \vee b$.

Remarque sur le caractère minimum du ppcm

Le nom « ppcm » est bien sûr une abréviation de « plus petit commun multiple ».

En anglais, la dénomination est « lcm » (*least common multiple*). On utilisera la notation $a \vee b$.

On va voir que l'entier $a \vee b$ est non seulement le minimum de $a\mathbb{Z} \cap b\mathbb{Z} \cap \mathbb{N}^*$ au sens de l'ordre naturel, mais qu'il en est aussi le minimum au sens de l'ordre défini par la divisibilité.

En d'autres termes, non seulement $a \vee b \leq m$ pour tout multiple commun strictement positif m des entiers a et b , mais on a mieux : $a \vee b \mid m$.

Premières propriétés

Il est clair qu'on a les égalités $a \vee b = b \vee a$, et $a \vee 1 = a$.

Par définition, l'entier $a \vee b$ est un multiple commun aux entiers a et b .

On verra que la réciproque est vraie : les multiples communs à a et b sont aussi des multiples de $a \vee b$.

On a l'égalité $a \vee b = a$ si et seulement si a est un multiple de b .

Un exemple

On forme la liste m_{12} des cinquante premiers multiples de l'entier 12 dans \mathbb{N}^* .

On crée ensuite la liste m_{15} des cinquante premiers multiples de l'entier 15 dans \mathbb{N}^* .

On forme enfin l'intersection ordonnée de ces deux listes. On obtient alors une liste de dix entiers strictement positifs, qui sont donc les dix premiers multiples communs de 12 et 15 dans \mathbb{N}^* .

On constate que le plus petit d'entre eux est l'entier 60 (donc $12 \vee 15 = 60$), et qu'il est effectivement un diviseur de tous les entiers de cette liste.


```
>>> m12 = list(12*k for k in range(1,51)); m12
[12, 24, 36, 48, 60, 72, 84, 96, 108, 120, 132, 144, 156, 168, 180, 192, 204, 216,
 228, 240, 252, 264, 276, 288, 300, 312, 324, 336, 348, 360, 372, 384, 396, 408,
 420, 432, 444, 456, 468, 480, 492, 504, 516, 528, 540, 552, 564, 576, 588, 600]
>>> m15 = list(15*k for k in range(1,51)); m15
[15, 30, 45, 60, 75, 90, 105, 120, 135, 150, 165, 180, 195, 210, 225, 240, 255, 270,
 285, 300, 315, 330, 345, 360, 375, 390, 405, 420, 435, 450, 465, 480, 495, 510,
 525, 540, 555, 570, 585, 600, 615, 630, 645, 660, 675, 690, 705, 720, 735, 750]
>>> sorted(set(m12) & set(m15))
[60, 120, 180, 240, 300, 360, 420, 480, 540, 600]
```

On en vient à une caractérisation du ppcm de deux entiers strictement positifs.

Proposition 10.2.8 (caractérisation du ppcm de deux entiers)

Soit a, b dans \mathbb{N}^* . Le ppcm de a et b est l'unique entier strictement positif n tel que $a\mathbb{Z} \cap b\mathbb{Z} = n\mathbb{Z}$.
Autrement dit, les multiples communs de a et b sont exactement ceux de leur ppcm.

Extension aux entiers relatifs

Si $a = 0$ ou si $b = 0$, on pose $a \vee b = 0$, ce qui est compatible avec la caractérisation $a\mathbb{Z} \cap b\mathbb{Z} = (a \vee b)\mathbb{Z}$, mais ne cadre pas avec la première définition (car 0 est le seul multiple commun de a et b).

Enfin, si a et b sont dans \mathbb{Z} , on peut encore appeler ppcm de a et b celui de $|a|$ et de $|b|$, c'est-à-dire l'unique entier naturel n tel que $a\mathbb{Z} \cap b\mathbb{Z} = (a \vee b)\mathbb{Z}$.

Dans la pratique, on ne perd pas grand-chose à se limiter aux cas $a > 0$ et $b > 0$.

Quelques propriétés du ppcm

Proposition 10.2.9 (mise en facteur ou simplification dans un calcul de ppcm)

Soit a, b, k trois entiers naturels strictement positifs.

Alors on a l'égalité : $(ka) \vee (kb) = k(a \vee b)$.

De même, si k est un diviseur commun de a et b , on a l'égalité : $\frac{a}{k} \vee \frac{b}{k} = \frac{a \vee b}{k}$.

La proposition suivante est admise pour l'instant. Elle sera démontrée après la proposition 10.3.3.

Proposition 10.2.10 (lien entre le pgcd et le ppcm)

Soit a, b deux entiers naturels strictement positifs. Alors on a l'égalité : $(a \wedge b)(a \vee b) = ab$

Il suffit donc de calculer $a \wedge b$ (par l'algorithme d'Euclide) pour obtenir le ppcm par : $a \vee b = \frac{ab}{a \wedge b}$.

Voici comment faire avec Python :

```
def pgcd(a, b):
    return pgcd(b, a % b) if b else a
```

```
def ppcm(a, b):
    return a*b // pgcd(a, b)
```

Et un exemple d'utilisation :

```

|| >>> ppcm(240, 350)
|| 8400

```

10.3 Entiers premiers entre eux

10.3.1 Couples d'entiers premiers entre eux

Définition 10.3.1

Soit a et b deux entiers relatifs.

On dit que a et b sont *premiers entre eux* (ou encore *étrangers*) si $a \wedge b = 1$.

Il revient au même d'écrire $\mathcal{D}(a) \cap \mathcal{D}(b) = \{-1, 1\}$.

Cela équivaut aussi à dire que le seul diviseur commun strictement positif de a et b est 1.

Dans le cas où $a = b = 0$, on rappelle que $a \wedge b = 0$, donc le problème ne pose pas.

Seuls les entiers 1 et -1 sont premiers avec eux-mêmes.

Comme on a toujours $a \wedge b = |a| \wedge |b|$, on peut se ramener au cas de deux entiers naturels dont l'un au moins est non nul. Et dire alors que a et b sont premiers entre eux, c'est dire que le dernier reste non nul dans leur algorithme d'Euclide est égal à 1.

On ne confondra pas cette notion avec celle de « nombre premier » (voir plus loin).

Proposition 10.3.1 (les quotients de deux entiers par leur pgcd sont premiers entre eux)

Soit a et b deux entiers relatifs (non tous deux nuls), et d leur pgcd (donc $d > 0$).

Les deux entiers a' et b' tels que $a = da'$ et $b = db'$ sont premiers entre eux.

Réciproquement, si $u \wedge v = 1$, et pour tout δ dans \mathbb{N}^* , le pgcd de δu et δv est égal à δ .

Avec les notations précédentes, le rationnel $r = \frac{a}{b}$ admet la forme dite *irréductible* : $r = \frac{a'}{b'}$.

Cette forme irréductible (simplifiée) est unique si on impose un dénominateur strictement positif.

10.3.2 Le théorème de Bézout et ses conséquences

Proposition 10.3.2 (théorème de Bezout)

Soit a et b deux entiers relatifs. Les deux propositions suivantes sont équivalentes :

- les entiers a et b sont premiers entre eux.
- il existe deux entiers relatifs u et v tels que $au + bv = 1$.

Le théorème de Bézout a des conséquences fondamentales en arithmétique des entiers :

Proposition 10.3.3 (lemme de Gauss)

Soit a, b, c trois entiers relatifs.

Si a divise bc et si a est premier avec b , alors a divise c .

Proposition 10.3.4 (résolution de $ax + by = 1$ quand a et b sont premiers entre eux)

Soit a et b deux entiers relatifs premiers entre eux.

Alors il existe une infinité de couples (x, y) de \mathbb{Z}^2 tels que $ax + by = 1$.

Si (x_0, y_0) est l'un d'eux, les autres sont donnés par
$$\begin{cases} x = x_0 + kb \\ y = y_0 - ka \end{cases} \text{ avec } k \text{ dans } \mathbb{Z}.$$

Proposition 10.3.5 (entier premier avec un produit)

Soit a, b, c trois entiers relatifs. Les deux propositions suivantes sont équivalentes :

- l'entier a est premier avec l'entier b et avec l'entier c .
- l'entier a est premier avec le produit bc .

Plus généralement, soient a_1, a_2, \dots, a_m et b_1, b_2, \dots, b_n deux familles d'entiers relatifs.

Les deux propositions suivantes sont équivalentes :

- chacun des entiers a_1, a_2, \dots, a_m est premier avec chacun des entiers b_1, b_2, \dots, b_n .
- le produit $a_1 a_2 \cdots a_m$ est premier avec le produit $b_1 b_2 \cdots b_n$.

Par exemple, si $a \wedge b = 1$, alors $a^m \wedge b^n = 1$.

Proposition 10.3.6 (divisibilité par deux entiers premiers entre eux)

Soit a, b, c trois entiers relatifs, les entiers a et b étant supposés premiers entre eux.

Si l'entier c est divisible par a et par b , alors il est divisible par leur produit ab .

Plus généralement si les entiers a_1, a_2, \dots, a_n sont premiers entre eux deux à deux, et si l'entier c est divisible par chacun des a_k , alors il est divisible par leur produit.

10.3.3 Pgcd de plusieurs entiers

Proposition 10.3.7 (associativité du pgcd et du ppcm)

Pour tous entiers relatifs a, b, c , on a les égalités
$$\begin{cases} a \wedge (b \wedge c) = (a \wedge b) \wedge c \\ a \vee (b \vee c) = (a \vee b) \vee c \end{cases}$$

L'associativité (et la commutativité) du pgcd et du ppcm ont pour conséquence que si on se donne n entiers relatifs a_1, a_2, \dots, a_n , alors les notations $a_1 \wedge a_2 \wedge \cdots \wedge a_n$ et $a_1 \vee a_2 \vee \cdots \vee a_n$ ont un sens, indépendamment de l'ordre des facteurs a_k et de celui dans lequel on effectue les calculs.

On peut alors poser la définition suivante :

Définition 10.3.2

Soit a_1, a_2, \dots, a_n , une famille de n entiers relatifs, avec $n \geq 2$.

L'entier $a_1 \wedge a_2 \wedge \cdots \wedge a_n$ est appelé le pgcd des entiers a_1, a_2, \dots, a_n .

L'entier $a_1 \vee a_2 \vee \cdots \vee a_n$ est appelé le ppcm des entiers a_1, a_2, \dots, a_n .

Caractérisation du pgcd et du ppcm

- L'entier $d = a_1 \wedge a_2 \wedge \cdots \wedge a_n$ est l'unique entier naturel tel que $\mathcal{D}(a_1) \cap \mathcal{D}(a_2) \cap \cdots \cap \mathcal{D}(a_n) = \mathcal{D}(d)$.

Autrement dit, les diviseurs communs aux entiers a_1, a_2, \dots, a_n sont les diviseurs de leur pgcd.

– L'entier $m = a_1 \vee a_2 \vee \dots \vee a_n$ est l'unique entier naturel tel que $a_1\mathbb{Z} \cap a_2\mathbb{Z} \cap \dots \cap a_n\mathbb{Z} = m\mathbb{Z}$.

Autrement dit, les multiples communs aux entiers a_1, a_2, \dots, a_n sont les multiples de leur ppcm.

Proposition 10.3.8 (factorisations et simplifications dans un calcul de pgcd/ppcm)

Soit a_1, \dots, a_n des entiers relatifs. Soit k un entier strictement positif.

On les égalités :
$$\begin{cases} (ka_1) \wedge (ka_2) \wedge \dots \wedge (ka_n) = k(a_1 \wedge a_2 \wedge \dots \wedge a_n) \\ (ka_1) \vee (ka_2) \vee \dots \vee (ka_n) = k(a_1 \vee a_2 \vee \dots \vee a_n) \end{cases}$$

Si k est un diviseur commun des a_k , alors

$$\frac{a_1}{k} \wedge \frac{a_2}{k} \wedge \dots \wedge \frac{a_n}{k} = \frac{a_1 \wedge a_2 \wedge \dots \wedge a_n}{k} \quad \text{et} \quad \frac{a_1}{k} \vee \frac{a_2}{k} \vee \dots \vee \frac{a_n}{k} = \frac{a_1 \vee a_2 \vee \dots \vee a_n}{k}.$$

Proposition 10.3.9 (relation de Bézout)

Soit a_1, a_2, \dots, a_n une famille de n entiers relatifs, avec $n \geq 2$. Soit d leur pgcd.

Alors il existe des entiers relatifs u_1, u_2, \dots, u_n tels que $a_1u_1 + a_2u_2 + \dots + a_nu_n = d$.

10.3.4 Entiers premiers entre eux dans leur ensemble

Définition 10.3.3 (entiers premiers entre eux dans leur ensemble)

Soit a_1, a_2, \dots, a_n une famille de n entiers relatifs, avec $n \geq 2$.

On dit que ces n entiers sont *premiers entre eux dans leur ensemble* si leur pgcd est égal à 1.

Cela équivaut à dire que le seul diviseur positif commun à a_1, a_2, \dots, a_n est 1.

Proposition 10.3.10 (caractérisation par une identité de Bézout)

Soit a_1, a_2, \dots, a_n une famille de n entiers relatifs, avec $n \geq 2$.

Les deux propositions suivantes sont équivalentes :

- les entiers a_1, a_2, \dots, a_n sont premiers entre eux dans leur ensemble.
- il existe n entiers relatifs u_1, u_2, \dots, u_n tels que $a_1u_1 + a_2u_2 + \dots + a_nu_n = 1$.

Remarques

– On ne confondra pas « premiers entre eux deux à deux » et « premiers entre eux dans leur ensemble ».

Plus précisément : si deux au moins des entiers a_1, \dots, a_n sont premiers entre eux, et a fortiori si les entiers a_1, \dots, a_n sont premiers entre eux deux à deux, alors ils le sont *dans leur ensemble*.

Dès que $n \geq 3$, la réciproque est fautive comme le montre l'exemple de 6, 10, 15.

On a en effet $6 \wedge 10 \wedge 15 = 1$ (car $6 + 10 - 15 = 1$), mais $6 \wedge 10 = 2$, $6 \wedge 15 = 3$, $10 \wedge 15 = 5$.

– Soit d un diviseur strictement positif commun aux entiers a_1, \dots, a_n .

Alors $d = a_1 \wedge a_2 \wedge \dots \wedge a_n$ si et seulement si les entiers $\frac{a_k}{d}$ sont premiers entre eux dans leur ensemble.

– Si a_1, \dots, a_n sont premiers entre eux deux à deux, alors $a_1 \vee a_2 \vee \dots \vee a_n = a_1 a_2 \dots a_n$.

– Attention : l'égalité $(a \wedge b)(a \vee b) = ab$ ne se généralise pas à plus de deux entiers.

Par exemple, avec $a = 6$, $b = 10$, $c = 15$, on a : $a \wedge b \wedge c = 1$, $a \vee b \vee c = 30$ et $abc = 900$.

Un peu de programmation Python

On utilise ici les fonctions `pgcd` et `ppcm` définies précédemment (sous-section 10.2.6)

Pour calculer le `pgcd` ou le `ppcm` des entiers d'une liste, on importe la fonction `reduce` du module `functools`, qui offre un moyen d'itérer une même fonction de deux arguments.

```
def mpgcd(seq):
    from functools import reduce
    return reduce(pgcd, seq)
```

```
def mppcm(seq):
    from functools import reduce
    return reduce(ppcm, seq)
```

Voici deux exemples très simples d'utilisation de ces deux fonctions :

```
>>> mpgcd([378, 630, 882, 945])
63
```

```
>>> mppcm([6, 10, 14, 15, 21])
210
```

10.4 Nombres premiers

10.4.1 Définition et « premières » propriétés

Définition 10.4.1

Soit p un entier naturel.

On dit que p est premier si $p \geq 2$ et si ses seuls diviseurs dans \mathbb{N} sont 1 et p .

On remarque que 1 n'est pas considéré comme un nombre premier.

On peut aussi adopter la définition suivante : un entier naturel p est dit premier s'il possède exactement deux diviseurs distincts dans \mathbb{N} (ce qui exclut les entiers 0 et 1)

Les dix plus petits nombres premiers sont 2, 3, 5, 7, 11, 13, 17, 19, 23, 29.

A l'exception de 2, tous les nombres premiers sont impairs.

Dans la phrase « a est premier avec b », il n'y a souvent pas de nombre premier...

Proposition 10.4.1

Soit p un nombre premier, et a un entier relatif.

Si p ne divise pas a , alors p est premier avec a .

En particulier, un entier premier p est premier avec tous les entiers de $\{1, \dots, p-1\}$.

Une autre conséquence est que deux nombres premiers distincts sont premiers entre eux.

Proposition 10.4.2 (entier premier divisant un produit)

Soit p un nombre premier, et soit a_1, a_2, \dots, a_n une famille d'entiers relatifs.

Si p divise le produit $a_1 a_2 \dots a_n$, alors p divise l'un au moins des entiers a_k .

Proposition 10.4.3 (existence d'un diviseur premier)

Tout entier naturel $n \geq 2$ est divisible par au moins un nombre premier.

Proposition 10.4.4

L'ensemble des nombres premiers est infini.

Crible d'Erathosthène

Une méthode artisanale bien connue.

Ici on se restreint à l'intervalle $\llbracket 1, 100 \rrbracket$.

On barre l'entier 1, qui n'est pas premier.

Ensuite on barre les entiers pairs (sauf 2 lui-même).

Puis on barre les multiples de 3 qui ne sont pas déjà barrés (donc les multiples impairs de 3), sauf 3 lui-même, puis les multiples non barrés de 5, puis ceux de 7.

L'entier premier 11 vérifie $11^2 > 100$ donc c'est fini.

À ce stade, il ne reste plus que des entiers premiers.

	1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20	
21	22	23	24	25	26	27	28	29	30	
31	32	33	34	35	36	37	38	39	40	
41	42	43	44	45	46	47	48	49	50	
51	52	53	54	55	56	57	58	59	60	
61	62	63	64	65	66	67	68	69	70	
71	72	73	74	75	76	77	78	79	80	
81	82	83	84	85	86	87	88	89	90	
91	92	93	94	95	96	97	98	99	100	

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100

(après avoir barré les entiers pairs > 2)

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100

(on a barré les multiples impairs de 3)

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100

(après avoir barré les multiples de 5)

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100

(puis ceux de 7)

Voici une fonction Python qui renvoie la liste des tous les entiers premiers strictement inférieurs à n^2 , où n est l'entier passé en argument :

```
def crible(n):
    return [p for p in range(2, n*n)
            if p not in
            set(j for i in range(2, n) for j in range(i*i, n*n, i))]
```

On retrouve la liste des entiers premiers strictement inférieurs à 10^2 :

```
>>> crible(10)
[2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41,
 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97]
```

La fonction `crible` n'est pas très facile à comprendre au premier abord.

Disons que l'ensemble `set(j for i in range(2, n) for j in range(i*i, n*n, i))` est celui de tous les entiers qui sont « grisés » à un moment ou à un autre de la méthode :

```
>>> set(j for i in range(2, 10) for j in range(i*i, 10*10, i))
{4, 6, 8, 9, 10, 12, 14, 15, 16, 18, 20, 21, 22, 24, 25, 26, 27, 28, 30, 32, 33, 34,
 35, 36, 38, 39, 40, 42, 44, 45, 46, 48, 49, 50, 51, 52, 54, 55, 56, 57, 58, 60, 62,
 63, 64, 65, 66, 68, 69, 70, 72, 74, 75, 76, 77, 78, 80, 81, 82, 84, 85, 86, 87, 88,
 90, 91, 92, 93, 94, 95, 96, 98, 99}
```

Sinon, voici une autre solution : la fonction `premiers` renvoie le tableau des entiers premiers compris entre 2 et n . L'intérêt repose dans l'utilisation d'un tableau Numpy de booléens (8 fois moins de taille que des int) et surtout dans le « fancy indexing » qui permet de barrer tous les multiples d'un nombre premier sans créer pour cela une boucle (on utilise une coupe du tableau pour l'affectation : on sait que ça ne nécessite aucune mémoire supplémentaire, car les calculs se font sur place).

```
def premiers(n):
    import numpy as np
    a = np.ones(n, 'bool')          # n fois 1
    a[0] = a[1] = False            # on barre 0 et 1
    for i in range(2, int(n**0.5)): # de i=2 à sqrt(n)
        if a[i]: a[i*i: n: i] = False # si i non barré, barrer ses suivants
    return np.nonzero(a)[0]        # renvoyer les entiers premiers
```

On obtient par exemple (en moins de 50ms) :

```
>>> premiers(1000)
array([ 2,  3,  5,  7, 11, 13, 17, 19, 23, 29, 31, 37, 41,
        43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97, 101,
        103, 107, 109, 113, 127, 131, 137, 139, 149, 151, 157, 163, 167,
        173, 179, 181, 191, 193, 197, 199, 211, 223, 227, 229, 233, 239,
        241, 251, 257, 263, 269, 271, 277, 281, 283, 293, 307, 311, 313,
        317, 331, 337, 347, 349, 353, 359, 367, 373, 379, 383, 389, 397,
        401, 409, 419, 421, 431, 433, 439, 443, 449, 457, 461, 463, 467,
        479, 487, 491, 499, 503, 509, 521, 523, 541, 547, 557, 563, 569,
        571, 577, 587, 593, 599, 601, 607, 613, 617, 619, 631, 641, 643,
        647, 653, 659, 661, 673, 677, 683, 691, 701, 709, 719, 727, 733,
        739, 743, 751, 757, 761, 769, 773, 787, 797, 809, 811, 821, 823,
        827, 829, 839, 853, 857, 859, 863, 877, 881, 883, 887, 907, 911,
        919, 929, 937, 941, 947, 953, 961, 967, 971, 977, 983, 991, 997])
```

10.4.2 Décomposition en produit de facteurs premiers

Dans la suite de ce chapitre, on note \mathbb{P} l'ensemble des nombres premiers (notation non standard).

Définition 10.4.2 (valuation p -adique d'un entier non nul)

Soit n un entier relatif non nul, et soit p un entier premier.

Il existe un plus grand entier naturel k tel que p^k divise n .

Cet entier est noté $v_p(n)$ et est appelé *valuation p -adique de n* .

Dire que $v_p(n)$ est nul, c'est dire que n n'est pas divisible avec p (donc est premier avec p).

Dire que $v_p(n)$ vaut 1, c'est dire que n est divisible par p , mais pas par p^2 .

Pour tout n de \mathbb{Z}^* , on a $v_p(n) = v_p(-n)$, ce qui permet de se limiter à $n > 0$.

Proposition 10.4.5 (une propriété de la valuation p -adique)

Pour tous entiers non nuls m et n , et pour tout p de \mathbb{P} , on a : $v_p(mn) = v_p(m) + v_p(n)$.

Proposition 10.4.6 (décomposition en produit de facteurs premiers)

Soit n un entier strictement positif.

Alors $n = \prod_{p \in \mathbb{P}} p^{\alpha_p}$ où les α_p sont dans \mathbb{N} et tous nuls sauf un nombre fini d'entre eux.

Dans cette écriture, et pour chaque p de \mathbb{P} , l'entier α_p est égal à la valuation p -adique $v_p(n)$ de n .

Une telle écriture de n est donc unique (à l'ordre près des facteurs).

On l'appelle décomposition de n en produits de facteurs premiers.

Si $n = \prod_{p \in \mathbb{P}} p^{\alpha_p}$, alors on peut dire que $-n = - \prod_{p \in \mathbb{P}} p^{\alpha_p}$ est la factorisation de $-n$ (peu utilisé en fait).

Le produit $\prod_{p \in \mathbb{P}} p^{\alpha_p}$ a un sens parce qu'il contient un nombre fini de facteurs distincts de 1.

Dans cette écriture de n , les entiers p pour lesquels $\alpha_p \geq 1$ sont les diviseurs premiers de n .

Pour l'entier $n = 1$, tous les exposants α_p sont nuls.

Pour prendre un exemple, la décomposition de $n = 4200$ est $n = 2^3 \cdot 3 \cdot 5^2 \cdot 7$.

Proposition 10.4.7 (caractérisation de la divisibilité en termes de valuations p -adiques)

Soit n et m deux entiers strictement positifs.

Soit $n = \prod_{p \in \mathbb{P}} p^{\alpha_p}$ et $m = \prod_{p \in \mathbb{P}} p^{\beta_p}$ leurs décompositions en produits de facteurs premiers.

Alors m divise n si et seulement si $\beta_p \leq \alpha_p$ pour tout p de \mathbb{P} .

Autrement dit : $m \mid n \Leftrightarrow (\forall p \in \mathbb{P}, v_p(m) \leq v_p(n))$.

Une conséquence est que le nombre de diviseurs distincts positifs de n est $\prod_{p \in \mathbb{P}} (\alpha_p + 1)$.

Par exemple l'entier $4200 = 2^3 \cdot 3 \cdot 5^2 \cdot 7$ possède $4 \cdot 2 \cdot 3 \cdot 2 = 48$ diviseurs distincts dans \mathbb{N} .

D'ailleurs les voici (on utilise la fonction `diviseurs` de la section 10.2.3) :

```
>>> diviseurs(4200)
[1, 2, 3, 4, 5, 6, 7, 8, 10, 12, 14, 15, 20, 21, 24, 25, 28, 30, 35, 40, 42, 50,
 56, 60, 70, 75, 84, 100, 105, 120, 140, 150, 168, 175, 200, 210, 280, 300, 350,
 420, 525, 600, 700, 840, 1050, 1400, 2100, 4200]
>>> len(_)
48
```


Proposition 10.4.8 (expression du pgcd et du ppcm à l'aide des valuations p -adiques)

Soit n et m deux entiers strictement positifs.

Soit $n = \prod_{p \in \mathbb{P}} p^{\alpha_p}$ et $m = \prod_{p \in \mathbb{P}} p^{\beta_p}$ leurs décompositions en produits de facteurs premiers.

Alors $m \wedge n = \prod_{p \in \mathbb{P}} p^{\gamma_p}$ et $m \vee n = \prod_{p \in \mathbb{P}} p^{\delta_p}$, avec $\begin{cases} \gamma_p = \min(\alpha_p, \beta_p) \\ \delta_p = \max(\alpha_p, \beta_p) \end{cases}$ pour tout p de \mathbb{P} .

Ce résultat permet de retrouver $(m \wedge n)(m \vee n) = mn$. En effet, $\gamma_p + \delta_p = \alpha_p + \beta_p$ pour tout p de \mathbb{P} .

10.4.3 Un peu de programmation Python

La fonction `factor` forme la liste (ordonnée suivant les valeurs croissantes) de tous les diviseurs premiers p de n (avec répétitions éventuelles, selon la valuation de p dans n) :

```
def factor(n):
    fp = []                # la liste des facteurs premiers de n, vide au départ
    pfp = 1               # le produit des éléments de cette liste, 1 au départ
    x = n                 # fait une copie de l'entier n qu'on veut factoriser
    d = 2                 # le plus petit diviseur potentiel de n
    while pfp != n:      # tant qu'on n'a pas complètement factorisé n
        while x % d == 0: # tant que d est un diviseur de x
            fp.append(d)  # on ajoute le diviseur (premier) d à la liste fp
            pfp *= d      # on actualise le produit des facteurs premiers
            x //= d       # on actualise l'entier x restant à factoriser
        d += 1           # on passe au diviseur potentiel suivant
    return fp            # renvoie la factorisation
```

Voici la factorisation de l'entier 1037400 :

```
>>> factor(1037400)
[2, 2, 2, 3, 5, 5, 7, 13, 19]
```

On écrit maintenant une deuxième fonction de factorisation. C'est juste une variation sur l'idée précédente. Mais on utilise un dictionnaire plutôt qu'une liste : dans ce dictionnaire les clefs sont les diviseurs premiers de n , et les valeurs associées à ces clefs sont les valuations p -adiques correspondantes.

La toute dernière instruction sert à renvoyer un résultat sous forme d'une liste de couples $(p, v_p(n))$, triée suivant les valeurs croissantes de p .

```
def factor2(n):
    # donne la liste des couples (p, vp(n))
    vp = {}
    # le dictionnaire des valuations p-adiques
    pf = 1
    # le produit des facteurs déjà obtenus
    x = n
    # fait une copie de l'entier n qu'on veut factoriser
    d = 2
    # le plus petit diviseur potentiel de n
    while pf != n:
        # tant qu'on n'a pas complètement factorisé n
        while x % d == 0:
            # tant que d est un diviseur de x
            if d in vp:
                # si le facteur premier d avait déjà été rencontré
                vp[d] += 1
                # on augmente la valuation d'une unité
            else: vp[d] = 1
                # sinon on pose une valuation égale à 1
            pf *= d
            # on actualise le produit des facteurs
            x //= d
            # on actualise l'entier x restant à factoriser
            d += 1
            # on passe au diviseur potentiel suivant
    return [(p, vp[p]) for p in sorted(vp)] # liste triée des (p, vp(n))
```

Voici à nouveau la factorisation de l'entier 1037400 :

```
>>> factor2(1037400)
[(2, 3), (3, 1), (5, 2), (7, 1), (13, 1), (19, 1)]
```

Voici maintenant une fonction (très compacte) renvoyant le n -ème nombre premier :

```
def nthprime(n):
    c = 1          # un compteur de nombres premiers
    p = 1          # initialise le futur n-ième nombre premier
    while c < n:   # tant qu'on n'a pas atteint n nombres premiers
        p += 2     # on passe au candidat suivant
        j = 3      # diviseur potentiel de p
        ok = True  # initialise condition de primalité
        while ok and j*j <= p: # tant qu'on ne dépasse pas sqrt(p)
            ok = p%j # on continue si j ne divise pas p
            j += 2   # et on passe au diviseur potentiel suivant
        if ok: c += 1 # si on est sorti avec ok=True, p était premier
    return max(p, 2) # max utile uniquement si n=1 donc si p=1
```

On obtient le 10000-ième nombre premier en une demi-seconde environ.

```
>>> nthprime(10000)
104729
```

Enfin, la fonction suivante renvoie la liste ordonnées de n plus petits entiers premiers :

```
def nprimes(N):
    primes = [2]          # initialise la liste des nombres premiers
    def primable(k):
        for j in primes: # pour chaque nombre premier de la liste
            if k % j == 0: # si j divise k
                return 0 # alors k n'est pas premier!
            if j*j > k: # si on dépasse sqrt(k)
                return 1 # alors k est "primable"
    n = 1 # initialise le compteur de nombres premiers
    k = 3 # 3 est le plus petit entier premier impair
    while n < N: # tant qu'on n'a pas obtenu N nombres premiers
        if primable(k): # si k est "primable" (donc premier!)
            n += 1 # alors on incrémente le compteur n
            primes.append(k) # et on ajoute k à la liste "primes"
        k += 2 # on passe à l'entier impair suivant
    return primes # renvoie la liste des n plus petits premiers
```

Voici par exemple la liste des 100 plus petits nombres premiers :

```
>>> nprimes(100)
[2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79,
83, 89, 97, 101, 103, 107, 109, 113, 127, 131, 137, 139, 149, 151, 157, 163, 167,
173, 179, 181, 191, 193, 197, 199, 211, 223, 227, 229, 233, 239, 241, 251, 257,
263, 269, 271, 277, 281, 283, 293, 307, 311, 313, 317, 331, 337, 347, 349, 353,
359, 367, 373, 379, 383, 389, 397, 401, 409, 419, 421, 431, 433, 439, 443, 449,
457, 461, 463, 467, 479, 487, 491, 499, 503, 509, 521, 523, 541]
```

10.5 Congruences

10.5.1 Congruence modulo un entier sur \mathbb{Z}

Définition 10.5.1

Soit n un entier strictement positif. Soit a, b deux entiers relatifs quelconques.

On note $a \equiv b [n]$ si la différence $b - a$ est dans $n\mathbb{Z}$ (c'est-à-dire est un multiple de n).

On exprime cette situation en disant que a et b sont congrus modulo n .

On définit ainsi une relation sur \mathbb{Z} , appelée *relation de congruence modulo n* .

Définitions équivalentes

– On a l'équivalence : $a \equiv b [n] \Leftrightarrow (\exists k \in \mathbb{Z}, a = b + kn)$.

– De même : $a \equiv b [n]$ équivaut à « a et b ont le même reste dans la division par n ».

Proposition 10.5.1

Soit n un entier strictement positif.

La relation de congruence modulo n est une relation d'équivalence sur \mathbb{Z} .

Classes d'équivalences

Soit n un entier strictement positif fixé.

On note souvent \bar{a} la classe d'équivalence de a pour la relation de congruence modulo n , c'est-à-dire l'ensemble des b de \mathbb{Z} tels que $b \equiv a [n]$, c'est-à-dire l'ensemble $\{a + kn, k \in \mathbb{Z}\}$.

Avec ces notations, $\bar{0} = \bar{n}$ est l'ensemble de tous les multiples de n dans \mathbb{Z} .

Tout entier relatif a est congru, modulo n , à un unique entier r de $\{0, \dots, n - 1\}$, l'entier r étant le reste dans la division de a par n .

Il y a donc exactement n classes d'équivalences, et on note souvent $\mathbb{Z}/n\mathbb{Z} = \{\bar{0}, \bar{1}, \dots, \overline{n-1}\}$.

10.5.2 Opérations sur les congruences

Proposition 10.5.2

Soit n un entier strictement positif.

Pour tous entiers relatifs a, b, c, d , on a les implications
$$\begin{cases} a \equiv b [n] \\ a' \equiv b' [n] \end{cases} \Rightarrow \begin{cases} a + a' \equiv b + b' [n] \\ aa' \equiv bb' [n] \end{cases}$$

Pour tout entier naturel k , on a l'implication : $a \equiv b [kn] \Rightarrow a \equiv b [n]$.

Pour tout entier naturel k , on a l'implication : $a \equiv b [n] \Rightarrow a^k \equiv b^k [n]$.

Si q est un entier strictement positif, on a l'équivalence : $a \equiv b [n] \Leftrightarrow qa \equiv qb [qn]$.

Si les entiers q et n sont premiers entre eux, alors : $qa \equiv qb [n] \Rightarrow a \equiv b [n]$.

Proposition 10.5.3 (petit théorème de Fermat)

Soit p un entier premier. Pour tout entier relatif a , on a : $a^p \equiv a [p]$.

En particulier, si a n'est pas divisible par p , on a : $a^{p-1} \equiv 1 [p]$.

Chapitre 11

Structures algébriques

Sommaire

11.1 Lois de composition	253
11.1.1 Loi de composition interne	253
11.1.2 Exemples de lois usuelles	254
11.1.3 Élément neutre et inversibilité	256
11.1.4 Distributivité	258
11.1.5 Partie stable pour une loi	258
11.2 Groupes et sous-groupes	259
11.2.1 Structure de groupe	259
11.2.2 Sous-groupe : définition, caractérisation	261
11.3 Structures d'anneau et de corps	262
11.3.1 Structure d'anneau	262
11.3.2 Calculs dans un anneau	262
11.3.3 Structure de corps	263

11.1 Lois de composition

11.1.1 Loi de composition interne

Définition 11.1.1 (loi de composition interne sur un ensemble)

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

Plutôt que de noter par exemple $f(u, v)$ (notation *préfixée*) l'image d'un couple (u, v) , on la note $u * v$, $u \top v$, $u + v$, etc. (notation *infixée*) et on parle alors des lois $*$, \top , $+$, etc.

On note souvent $(E, *)$ pour désigner un ensemble E muni d'une loi de composition $*$.

On retiendra qu'une *loi de composition* $*$ sur E est un mécanisme permettant, à partir de deux éléments quelconques x et y de E , de former un élément z de E , noté $z = x * y$ et qu'on pourra appeler *composé de x par y pour la loi $*$* . Il est important qu'une loi soit *partout définie* : le résultat $x * y$ doit donc avoir un sens, quels que soient les éléments x et y de E .

Il arrive souvent qu'on utilise plusieurs fois le mécanisme précédent dans un même calcul. Il faut alors préciser, au moyen de parenthèses, dans quel ordre on a effectué les compositions.

Par exemple l'expression $x * y * z$ est a priori dépourvue de signification, et il faudrait écrire :

- soit $(x * y) * z$ si on a d'abord calculé $a = x * y$ avant de calculer $a * z$.
- soit $x * (y * z)$ si on a d'abord calculé $b = y * z$ avant de calculer $x * b$.

Plus compliqué, une expression comme $x * y * z * t$ possède les cinq parenthésages possibles suivants, qui indiquent chacune une chronologie particulière dans les compositions par la loi $*$:

$$(x * y) * (z * t), \quad ((x * y) * z) * t, \quad (x * (y * z)) * t, \quad x * ((y * z) * t) \quad \text{et} \quad x * (y * (z * t))$$

On appréciera donc qu'une loi de composition possède la propriété suivante :

Définition 11.1.2 (associativité d'une loi de composition)

Soit $*$ une loi de composition sur un ensemble E .

On dit que la loi $*$ est *associative* si, pour tous x, y, z de E , on a : $(x * y) * z = x * (y * z)$.

Quand une loi de composition $*$ est associative, une expression comme $a * b * \dots * x * y * z$ est définie sans ambiguïté : les parenthèses qui indiquent dans quel ordre on combine les éléments deux à deux sont en effet inutiles. En revanche, l'ordre dans lequel les éléments apparaissent, de gauche à droite, reste important, à moins que...

Définition 11.1.3 (éléments qui commutent pour une loi)

Soit $*$ une loi de composition sur un ensemble E .

Soit x et y deux éléments de E . On dit que x et y *commutent* (pour la loi $*$) si $x * y = y * x$.

Définition 11.1.4 (commutativité d'une loi de composition)

Soit $*$ une loi de composition sur un ensemble E .

On dit que la loi $*$ est *commutative* si, pour tous x et y de E , on a : $x * y = y * x$.

Autrement dit : une loi de composition sur E est commutative si tous les éléments de E commutent deux à deux pour cette loi. Quand une loi $*$ est associative et commutative, non seulement une expression comme $a * b * \dots * x * y * z$ est définie sans ambiguïté, mais on peut aussi changer l'ordre des termes et notamment regrouper ceux d'entre eux qui sont identiques.

On pourra ainsi noter $x * y * x * y * z * y * x * y = x^3 * y^4 * z$ à condition, pour tout entier strictement positif n , de poser $a^n = a * a * \dots * a$ (l'élément a apparaissant n fois).

11.1.2 Exemples de lois usuelles

Somme et produit sur les ensembles de nombres :

Les lois $+$ et \times usuelles sur \mathbb{N} , \mathbb{Z} , \mathbb{Q} , \mathbb{R} , et \mathbb{C} sont associatives et commutatives.

La loi produit \times est le plus souvent notée par simple *juxtaposition* : xy plutôt que $x \times y$.

On peut considérer l'opération « différence » sur \mathbb{Z} , \mathbb{Q} , \mathbb{R} , et \mathbb{C} (mais pas sur \mathbb{N}).

Cette loi, définie par $(x, y) \mapsto x - y$, n'est ni associative ni commutative (donc sans intérêt!).

La loi de composition des applications

Soit E un ensemble, et soit $\mathcal{F}(E)$ l'ensemble des applications de E dans E .

On définit la loi \circ (loi de composition) sur $\mathcal{F}(E)$ par $(f, g) \mapsto f \circ g$.

Cette loi est associative, mais elle n'est pas commutative (sauf si E est réduit à un singleton).

Même si la loi \circ n'est pas commutative, il existe des applications f et g qui commutent entre elles, c'est-à-dire qui vérifient $g \circ f = f \circ g$.

Par exemple, en géométrie du plan, les rotations de même centre commutent deux à deux.

Lois sur l'ensemble des parties d'un ensemble

Soit E un ensemble. Les lois \cup (union), \cap (intersection) et Δ (différence symétrique) sur $\mathcal{P}(E)$ sont toutes trois commutatives et associatives.

En revanche, la « différence ensembliste » définie sur $\mathcal{P}(E)$ par $A \setminus B = A \cap \overline{B} = \{x \in A, x \notin B\}$, n'est ni associative ni commutative.

Maximum et minimum sur un ensemble totalement ordonné

Soit E un ensemble muni d'une relation d'ordre total noté \leq .

Les lois \min et \max (minimum et maximum) sont notées de façon préfixée : $\min(x, y)$ et $\max(x, y)$.

Ces deux lois sont associatives et commutatives.

Pgcd et ppcm sur les entiers

Les lois pgcd et ppcm sur \mathbb{N} ou \mathbb{Z} sont commutatives et associatives.

Elles sont notées de façon tantôt préfixe (pgcd(a, b) et ppcm(a, b)), tantôt infixes ($a \wedge b$ et $a \vee b$).

C'est l'associativité qui permet de noter, sans ambiguïté : $\begin{cases} a \wedge b \wedge c \wedge \dots \\ a \vee b \vee c \vee \dots \end{cases}$ pour tous entiers a, b, c, \dots

Addition et produit « modulo n »

Si n est un entier strictement positif, notons $\mathbb{N}_n = \{0, 1, 2, \dots, n-1\}$.

\mathbb{N}_n est donc l'ensemble des restes possibles dans la division euclidienne par n .

On peut définir deux lois sur \mathbb{N}_n à partir des lois $+$ et \times de \mathbb{N} , en calculant le résultat modulo n .

Ces deux lois sont associatives et commutatives.

Par exemple, dans \mathbb{N}_{15} , on a $11 + 23 = 4$ (car $34 \equiv 4 [15]$) et $11 \cdot 23 = 13$ (car $11 \cdot 23 = 253 \equiv 13 [15]$).

Voici la table des lois $+$ et \times dans l'ensemble $\mathbb{N}_{10} = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ (la première ligne et la première colonne de chaque tableau ne sont pas considérées comme faisant partie de la table : elles sont simplement un rappel de la valeur des éléments de \mathbb{N}_{10}).

+	0	1	2	3	4	5	6	7	8	9
0	0	1	2	3	4	5	6	7	8	9
1	1	2	3	4	5	6	7	8	9	0
2	2	3	4	5	6	7	8	9	0	1
3	3	4	5	6	7	8	9	0	1	2
4	4	5	6	7	8	9	0	1	2	3
5	5	6	7	8	9	0	1	2	3	4
6	6	7	8	9	0	1	2	3	4	5
7	7	8	9	0	1	2	3	4	5	6
8	8	9	0	1	2	3	4	5	6	7
9	9	0	1	2	3	4	5	6	7	8

Table de l'addition
dans $\mathbb{N}_{10} = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$

\times	0	1	2	3	4	5	6	7	8	9
0	0	0	0	0	0	0	0	0	0	0
1	0	1	2	3	4	5	6	7	8	9
2	0	2	4	6	8	0	2	4	6	8
3	0	3	6	9	2	5	8	1	4	7
4	0	4	8	2	6	0	4	8	2	6
5	0	5	0	5	0	5	0	5	0	5
6	0	6	2	8	4	0	6	2	8	4
7	0	7	4	1	8	5	2	9	6	3
8	0	8	6	4	2	0	8	6	4	2
9	0	9	8	7	6	5	4	3	2	1

Table du produit
dans $\mathbb{N}_{10} = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$

Loi sur l'ensemble des fonctions à valeurs dans $(E, *)$

Soit E un ensemble muni d'une loi $*$, et soit X un ensemble quelconque.

On définit une loi, encore notée $*$, sur l'ensemble $\mathcal{F}(X, E)$ des applications de X vers E .

On pose pour cela : $\forall (f, g) \in \mathcal{F}(X, E)^2, \forall x \in X, (f * g)(x) = f(x) * g(x)$.

On vérifie que si la loi $*$ sur E est associative (resp. commutative), alors la loi $*$ sur $\mathcal{F}(X, E)$ est encore associative (resp. commutative).

On définit par exemple les lois $+$ et \times sur l'ensemble des applications de X vers \mathbb{R} (ou $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{C}$).

Autre exemple : si $X = \mathbb{N}$ et $E = \mathbb{R}$, on définit la somme $s = u + v$ et le produit $p = uv$ de deux suites $(u_n)_{n \geq 0}$ et $(v_n)_{n \geq 0}$ de nombres réels en posant, pour tout n de \mathbb{N} : $s_n = u_n + v_n$ et $p_n = u_n v_n$.

11.1.3 Élément neutre et inversibilité

Définition 11.1.5

Soit E un ensemble muni d'une loi de composition $*$. Soit e un élément de E .

On dit que e est *élément neutre* pour la loi $*$ si, pour tout élément x de E , on a : $x * e = e * x = x$.

Remarque : si on sait que la loi $*$ est commutative, l'égalité $x * e = e * x$ est automatiquement réalisée.

Proposition 11.1.1 (unicité de l'élément neutre)

L'élément neutre de l'ensemble E pour la loi $*$, s'il existe, est unique.

Conventions de vocabulaire

Il est beaucoup plus juste de dire que c'est E qui *possède* un élément neutre e pour la loi $*$, plutôt que de dire que c'est la loi $*$ qui possède l'élément neutre e .

La notation $+$ peut être employée en dehors des ensembles $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$: elle doit cependant être réservée aux lois commutatives. Dans ce cas, l'élément neutre, s'il existe, sera souvent noté 0 .

De même, pour une loi noté multiplicativement (ou par juxtaposition), on pourra noter 1 l'élément neutre éventuel (s'il n'y a pas de risque d'ambiguïté).

Quelques exemples

– Dans $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ et \mathbb{C} : 0 est neutre pour la loi $+$ et 1 est neutre pour la loi \times .

– Dans $\mathcal{F}(E)$: l'application Identité Id_E est neutre pour la loi \circ (composition).

– Soit X un ensemble quelconque, et soit E un ensemble muni d'une loi $*$ avec un neutre e .

On munit $\mathcal{F}(X, E)$ de la loi $*$, définie par : $\forall (f, g) \in \mathcal{F}(X, E)^2, \forall x \in X, (f * g)(x) = f(x) * g(x)$.

Alors l'application constante, qui à tout x de E associe e , est neutre pour cette loi.

Ainsi, sur l'ensemble $\mathcal{F}(\mathbb{N}, \mathbb{K})$ des suites (à valeurs dans $\mathbb{K} = \mathbb{R}$ ou \mathbb{C}), la suite constante 0 est neutre pour l'addition, et la suite constante 1 est neutre pour le produit.

– Dans $\mathcal{P}(E)$: \emptyset est neutre pour la loi \cup (et pour la loi Δ), et E est neutre pour la loi \cap .

– Dans \mathbb{Z}, \mathbb{Q} et \mathbb{R} : les lois \min et \max n'ont pas d'élément neutre.

Définition 11.1.6 (inversibilité d'un élément)

Soit E un ensemble muni d'une loi associative $*$.

On suppose qu'il existe un élément neutre e .

On dit qu'un élément x est *inversible* (pour la loi $*$) s'il existe x' dans E tel que $x * x' = x' * x = e$.

Si un tel élément x' existe, il est unique.

On le note en général x^{-1} , et on l'appelle l'*inverse* (ou le *symétrique*) de x pour la loi $*$.

Proposition 11.1.2 (inversibilité du produit de deux éléments inversibles)

Soit E un ensemble muni d'une loi associative $*$.

On suppose qu'il existe un élément neutre e .

Soit x et y deux éléments de E , inversibles pour la loi $*$, d'inverses respectifs x^{-1} et y^{-1} .

Alors $x * y$ est inversible, et son inverse est $(x * y)^{-1} = y^{-1} * x^{-1}$.

Remarque : attention à la permutation dans la formule précédente, si la loi $*$ n'est pas commutative.

Notation additive

Dans le cas d'une loi $+$ (nécessairement commutative, d'élément neutre 0), on ne parle pas d'inverse ou de symétrique, mais d'*opposé*, et l'élément en question n'est pas noté x^{-1} ou x' mais $-x$.

L'opposé de x est donc l'unique élément de E tel que $x + (-x) = 0$.

Pour tous éléments x et y , et si x possède un opposé, on note $y - x$ plutôt que $y + (-x)$.

En notation additive, la propriété $(x * y)^{-1} = y^{-1} * x^{-1}$ devient $-(x + y) = -y - x = -x - y$.

Propriétés et remarques

- S'il n'y a pas de neutre dans E pour la loi $*$, alors la notion d'élément inversible n'a aucun sens.
- On a demandé à la loi $*$ d'être associative pour garantir l'*unicité du symétrique* si existence.
- L'élément neutre e de E pour la loi $*$ est inversible et il est son propre inverse (car $e * e = e$).

Exemples

- Dans $(\mathbb{N}, +)$ seul l'entier 0 possède un opposé.
 Bien sûr, tous les éléments de $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$ et $(\mathbb{C}, +)$ possèdent un opposé.
 Les éléments inversibles de (\mathbb{R}, \times) sont les éléments non nuls (idem avec (\mathbb{Q}, \times) et (\mathbb{C}, \times)).
 Le seul élément inversible de (\mathbb{N}, \times) est 1 . Les seuls éléments inversibles de (\mathbb{Z}, \times) sont -1 et 1 .
- Dans $(\mathcal{F}(E), \circ)$, une application est inversible si et seulement si elle est bijective.
 Son inverse est alors sa bijection réciproque f^{-1} . La notation f^{-1} ne prête donc pas à confusion.
- Dans le cas des applications de \mathbb{R} dans \mathbb{R} , on ne confondra le produit fg et la composition $f \circ g$.
 On sait que le neutre pour la loi \circ est Id_E , et que f est inversible pour cette loi si f est bijective.
 En revanche, le neutre pour la loi produit est l'application constante $x \mapsto 1$, et f est inversible pour cette loi (c'est-à-dire il existe une application g telle que $fg = 1$) si et seulement si f ne s'annule jamais. L'inverse de f (pour le produit !) est alors l'application $1/f$.

11.1.4 Distributivité

Définition 11.1.7

Soit E un ensemble muni de deux lois $*$ et \top .

On dit que la loi $*$ est *distributive* par rapport à la loi \top si, pour tous x, y, z de E :

- d’une part $x * (y \top z) = (x * y) \top (x * z)$ (distributivité à gauche)
- d’autre part $(x \top y) * z = (x * z) \top (y * z)$ (distributivité à droite)

Exemples et remarques

- Si la loi $*$ est commutative, chacune des deux distributivités (à gauche ou à droite) implique l’autre.
- Dans $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$, on a toujours $x(y + z) = xy + xz$: la loi \times est distributive par rapport à la loi $+$.
En revanche la loi $+$ n’est pas distributive par rapport à la loi \times .
- Dans $\mathcal{P}(E)$, les lois \cup et \cap sont distributives l’une par rapport à l’autre.

On a en effet toujours les égalités :
$$\begin{cases} A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \\ A \cup (B \cap C) = (A \cup B) \cap (A \cup C) \end{cases}$$

Toujours dans $\mathcal{P}(E)$, la loi \cap est distributive par rapport à la loi Δ .

Mais réciproquement, la loi Δ n’est pas distributive par rapport à la loi \cap .

- Les propriétés de distributivité, d’associativité, de commutativité sont utiles aux développements.

Il est par exemple clair que $(a + b)(c + d) = ac + ad + bc + bd$ pour tous réels a, b, c, d .

Mais si A, B, C, D sont quatre parties d’un ensemble E , les mêmes propriétés permettent d’écrire :

d’une part : $(A \cup B) \cap (C \cup D) = (A \cap C) \cup (A \cap D) \cup (B \cap C) \cup (B \cap D)$

d’autre part : $(A \cap B) \cup (C \cap D) = (A \cup C) \cap (A \cup D) \cap (B \cup C) \cap (B \cup D)$

11.1.5 Partie stable pour une loi

Définition 11.1.8

Soit E un ensemble muni d’une loi $*$, et soit F une partie de E .

On dit que F est *stable* pour la loi $*$ si : $\forall (x, y) \in F \times F, x * y \in F$.

La restriction à $F \times F$ de la loi $*$ définit alors une loi de composition sur F , qu’on appelle *loi induite sur F par celle de E* , et qu’en général on note encore $*$.

Quelques exemples dans $(\mathbb{R}, +)$ ou (\mathbb{R}, \times)

L’intervalle $\mathbb{R}^+ = [0, +\infty[$ est stable, à la fois pour l’addition et pour le produit.

L’intervalle $[-1, 1]$ est stable pour le produit, mais pas pour l’addition.

L’intervalle $[-2, 2]$ n’est stable ni pour le produit, ni pour l’addition.

L’intervalle $\mathbb{R}^- =]-\infty, 0]$ est stable pour l’addition, mais pas pour le produit.

L’ensemble des rationnels est une partie stable de \mathbb{R} , pour les deux lois $+$ et \times .

Remarques en cas de loi induite

Soit F une partie stable de E pour la loi $*$.

Si on munit F de la loi induite (toujours notée $*$), on dispose donc à la fois de $(E, *)$ et de $(F, *)$.

– Si la loi $*$ sur E est commutative (resp. associative), il en est de même de la loi induite $*$ sur F .

– Si e est neutre dans $(E, *)$, et si e est dans F , alors bien sûr e est encore neutre dans $(F, *)$.

Mais attention, il est possible qu'il y ait un neutre e' dans F et qu'il n'y ait pas de neutre dans E . Il est possible aussi qu'il y ait un neutre e dans E , mais que e n'appartienne pas à F (dans ces conditions, il est possible que F possède lui-même son propre neutre, ou qu'il n'en possède pas!).

– Supposons qu'un élément e de F soit neutre dans $(E, *)$, donc neutre dans $(F, *)$.

Soit x un élément de F . Si on examine l'inversibilité de x pour la loi $*$, il faut savoir sans ambiguïté si on parle d'inversibilité dans F (pour la loi $*$ induite) ou dans E (pour la loi $*$ initiale).

Par exemple 2 est inversible dans (\mathbb{R}, \times) , mais pas dans (\mathbb{Z}, \times) (car $1/2$ n'existe pas dans \mathbb{Z} !).

11.2 Groupes et sous-groupes

11.2.1 Structure de groupe

Définition 11.2.1

Soit G un ensemble muni d'une loi de composition $*$.

On dit que $(G, *)$ est un *groupe* si :

- la loi $*$ est associative, et il y a un élément neutre e .
- tout élément de G possède un inverse.

Si de plus la loi $*$ est commutative, on dit que $(G, *)$ est un groupe *commutatif* (ou encore *abélien*).

Premières remarques

- Par définition un groupe est toujours non vide (puisqu'il y a au moins l'élément neutre).
- Si la loi est notée $+$, on dit que $(G, +)$ est un *groupe additif*. Le neutre est noté 0 . On rappelle qu'une loi $+$ est toujours supposée commutative, et qu'on note $-x$ l'opposé (plutôt que l'inverse) de x .
- En cas de loi produit (notation \times ou par juxtaposition), on dit que (G, \times) est un *groupe multiplicatif*.

Exemples usuels

- Les ensembles $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$ et $(\mathbb{C}, +)$ sont des groupes additifs.
- Les ensembles (\mathbb{Q}^*, \times) , $(\mathbb{Q}^{+*}, \times)$, (\mathbb{R}^*, \times) , $(\mathbb{R}^{+*}, \times)$ et (\mathbb{C}^*, \times) sont des groupes multiplicatifs.
- L'ensemble $\mathcal{U} = \{z \in \mathbb{C}, |z| = 1\}$ est un groupe multiplicatif.
Il en est de même de l'ensemble $\mathcal{U}_n = \{z \in \mathbb{C}, z^n = 1\}$ des racines n -ièmes de l'unité.

Définition 11.2.2 (groupe des permutations d'un ensemble E)

Soit E un ensemble. On note \mathcal{S}_E l'ensemble des bijections de E dans E (on dit les *permutations* de E). Alors \mathcal{S}_E est un groupe pour la loi de composition des applications.

Remarque : dès que l'ensemble E possède au moins trois éléments, le groupe \mathcal{S}_E est non commutatif.

Définition 11.2.3 (puissances entières d'un élément)

Soit $(G, *)$ un groupe multiplicatif, d'élément neutre e , et soit x un élément de G .

On définit les puissances entières x^n ($n \in \mathbb{Z}$) de x de la manière suivante :

- on pose $x^0 = e$.
- pour tout n de \mathbb{N}^* , on pose $x^n = x * x^{n-1}$, c'est-à-dire $x^n = x x \cdots x$ (n fois).
- pour tout n de \mathbb{N}^* , on pose $x^{-n} = (x^n)^{-1}$, ou ce qui revient au même : $x^{-n} = (x^{-1})^n$.

Avec ces notations, on a : $x^n x^m = x^{n+m}$, et $(x^n)^m = x^{nm}$ pour tout x de G et tous m, n de \mathbb{Z} .

Si les deux éléments x et y commutent, alors $(x * y)^n = x^n * y^n$ (attention, c'est faux si $x * y \neq y * x$).

En notation additive, on ne note plus x^n mais nx , pour tout n de \mathbb{Z} .

Dans un groupe additif $(G, +)$, on considérera donc $3x$, par exemple, non pas comme le produit de 3 par x , mais comme un raccourci pour désigner $x + x + x$.

Proposition 11.2.1 (dans un groupe tout élément est simplifiable)

Soit G un groupe pour la loi $*$.

Pour tous éléments x, y, z de G , on a les implications $\begin{cases} (x * y = x * z) \Rightarrow y = z \\ (y * x = z * x) \Rightarrow y = z \end{cases}$

On exprime cette propriété en disant que dans un groupe tout élément est simplifiable.

Attention, cette propriété cesse d'être vraie si on n'est pas dans un groupe (il peut exister des éléments « non simplifiables »). Par exemple, dans (\mathbb{R}, \times) , on n'a pas l'implication $0x = 0y \Rightarrow x = y$ (mais en revanche tout réel non nul est simplifiable pour le produit).

Dans tout ensemble E muni d'une loi $*$, les implications $\begin{cases} y = z \Rightarrow (x * y = x * z) \\ y = z \Rightarrow (y * x = z * x) \end{cases}$ sont toujours vraies (et elles sont mêmes évidentes, et sans grand intérêt).

Proposition 11.2.2

Soit G un groupe pour la loi $*$. Soit a un élément de G .

L'application $g_a : x \mapsto a * x$ (dite « multiplication à gauche par a ») est bijective.

Sa bijection réciproque est $g_{a^{-1}} : x \mapsto a^{-1} * x$ (c'est-à-dire la multiplication à gauche par a^{-1}).

L'application $d_a : x \mapsto x * a$ (dite « multiplication à droite par a ») est bijective.

Sa bijection réciproque est $d_{a^{-1}} : x \mapsto x * a^{-1}$ (c'est-à-dire la multiplication à droite par a^{-1}).

On peut réécrire le résultat précédent en termes de résolutions d'équations dans un groupe.

Proposition 11.2.3 (équations $a * x = b$ et $x * a = b$ dans un groupe)

Soit G un groupe pour la loi $*$. Soit a, b deux éléments de G .

L'équation $a * x = b$ possède une solution unique, à savoir $x = a^{-1} * b$.

L'équation $x * a = b$ possède une solution unique, à savoir $x = b * a^{-1}$.

11.2.2 Sous-groupe : définition, caractérisation

Définition 11.2.4 (sous-groupe d'un groupe)

Soit $(G, *)$ un groupe et soit H une partie non vide de G .

On dit que H est un *sous-groupe* de $(G, *)$ si :

- l'ensemble H est stable pour la loi $*$: $\forall (x, y) \in H^2, x * y \in H$.
- l'ensemble H est « stable pour le passage à l'inverse » : $\forall x \in H, x^{-1} \in H$.

Remarque : on n'oubliera pas la condition disant que H est une partie *non vide* de G .

La proposition suivante dit qu'un sous-groupe, c'est un groupe à part entière (le mot « sous » n'a donc rien de péjoratif : il se réfère simplement à l'inclusion des ensembles).

Proposition 11.2.4

Soit H un sous-groupe de $(G, *)$. On munit H de la loi induite.

Alors $(H, *)$ est lui-même un groupe.

Les deux groupes $(G, *)$ et $(H, *)$ ont le même neutre (qui est donc élément de H).

Si x est élément de H , l'inverse x^{-1} de x est le même (du point de vue de G ou de celui de H).

Proposition 11.2.5 (caractérisation des sous-groupes)

Soit $(G, *)$ un groupe et soit H une partie non vide de G .

H est un sous-groupe de $(G, *)$ si et seulement si : $\forall (x, y) \in H^2, x * y^{-1} \in H$.

En notation additive : H est un sous-groupe de $(G, +)$ si et seulement si : $\forall (x, y) \in H^2, x - y \in H$.

Exemples

- Soit $(G, *)$ un groupe de neutre e . Alors $\{e\}$ et G en sont deux sous-groupes (dits *triviaux*).
- Dans $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$, $(\mathbb{C}, +)$, chacun est un sous-groupe du suivant.
- C'est la même chose avec $(\{-1, 1\}, \times)$, (\mathbb{Q}^*, \times) , (\mathbb{R}^*, \times) , (\mathbb{C}^*, \times) .
- De même, $(\mathbb{R}^{+*}, \times)$ est un sous-groupe de (\mathbb{R}^*, \times) .
- L'ensemble \mathcal{U} des nombres complexes de module 1 est un sous-groupe de (\mathbb{C}^*, \times) .
- L'ensemble \mathcal{U}_n des racines n -ièmes de l'unité est un sous-groupe de (\mathcal{U}, \times) .

Proposition 11.2.6 (intersection de sous-groupes)

Une intersection quelconque de sous-groupes de G est encore un sous-groupe de G .

Remarque : c'est faux pour la réunion ! Plus précisément, si H et K sont deux sous-groupes de G , $H \cup K$ est un sous-groupe de G si et seulement si $H \subset K$ ou $K \subset H$ (et alors $H \cup K = K$ ou $H \cup K = H$).

Proposition 11.2.7 (sous-groupes de $(\mathbb{Z}, +)$)

On rappelle que si n est un élément de \mathbb{N} , on note $n\mathbb{Z} = \{kn, k \in \mathbb{Z}\}$.

Les sous-groupes de $(\mathbb{Z}, +)$ sont les $n\mathbb{Z}$, avec n dans \mathbb{N} .

11.3 Structures d'anneau et de corps

11.3.1 Structure d'anneau

Définition 11.3.1 (structure d'anneau)

Soit A un ensemble muni de deux lois de composition, notées $+$ et \times .

On dit que $(A, +, \times)$ est un *anneau* si :

- $(A, +)$ est un groupe commutatif (son neutre est en général noté 0).
- La loi \times est associative et distributive par rapport à l'addition.
- Il existe un élément neutre pour le produit \times (en général noté 1).

Si de plus la loi \times est commutative, on dit que $(A, +, \times)$ est un *anneau commutatif*.

Exemples

- $(\mathbb{Z}, +, \times)$, $(\mathbb{Q}, +, \times)$, $(\mathbb{R}, +, \times)$ et $(\mathbb{C}, +, \times)$ sont des anneaux commutatifs.
- Soit $(A, +, \times)$ un anneau de neutres 0 (pour la loi $+$) et 1 (pour la loi \times).
Il est possible que les deux éléments 0 et 1 de A soient identiques!
Mais dans ce cas A se réduit à $\{0\}$ (anneau nul, sans grand intérêt).

Proposition 11.3.1 (groupe des éléments inversibles d'un anneau)

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

L'ensemble des éléments de A qui sont inversibles pour le produit est un groupe pour la loi \times .

Exemples

Le groupe des inversibles de l'anneau $(\mathbb{Z}, +, \times)$ se réduit à la paire $\{-1, 1\}$.

Le groupe des inversibles de l'anneau $(\mathbb{R}, +, \times)$ est l'ensemble de tous les réels non nuls.

11.3.2 Calculs dans un anneau

Soit $(A, +, \times)$ un anneau (on note 0 le neutre pour $+$, et 1 le neutre pour \times).

Rappelons qu'on note $a - b$ plutôt que $a + (-b)$.

On rappelle également que, pour tout n de \mathbb{N}^* , la notation na désigne $a + a + \dots + a$ (n fois).

Pour tout (a, b, c) de A^3 , et tout entier relatif m , on a :

$$a0 = 0a = 0, \quad \begin{cases} (-a)b = a(-b) = -(ab) \\ (-a)(-b) = ab \end{cases}, \quad \begin{cases} a(b - c) = ab - ac \\ (a - b)c = ac - bc \end{cases}$$

Sommes et produits, développements

Pour tous a_m, a_{m+1}, \dots, a_n de A , on écrit : $a_m + a_{m+1} + \dots + a_n = \sum_{k=m}^n a_k$ et $a_m a_{m+1} \dots a_n = \prod_{k=m}^n a_k$

Si $m > n$, on pose $\sum_{k=m}^n a_k = 0$ et $\prod_{k=m}^n a_k = 1$ (somme et produit « vides »).

Pour tout b de A , on a : $b \left(\sum_{k=m}^n a_k \right) = \sum_{k=m}^n (ba_k)$, et $\left(\sum_{k=m}^n a_k \right) b = \sum_{k=m}^n (a_k b)$.

Plus généralement :
$$\left(\sum_{j=m}^n a_j\right)\left(\sum_{k=p}^q b_k\right) = \sum_{j=m}^n \left(a_j \sum_{k=p}^q b_k\right) = \sum_{j=m}^n \sum_{k=p}^q a_j b_k.$$

Proposition 11.3.2 (factorisation de $a^n - b^n$ dans un anneau)

Soit $(A, +, \times)$ un anneau, et soit a, b deux éléments de A . On suppose que $ab = ba$.

Alors, pour tout n de \mathbb{N}^* , on a l'égalité :
$$a^n - b^n = (a - b) \left(\sum_{k=0}^{n-1} a^{n-1-k} b^k\right).$$

Cas particuliers (toujours si $ab = ba$) : $a^2 - b^2 = (a - b)(a + b)$, et $a^3 - b^3 = (a - b)(a^2 + ab + b^2)$.

Puisque 1 commute avec tous les éléments de A , on a toujours la factorisation :

$$\forall a \in A, \forall n \in \mathbb{N}^*, 1 - a^n = (1 - a) \sum_{k=0}^{n-1} a^k = (1 - a)(1 + a + a^2 + \dots + a^{n-1})$$

Proposition 11.3.3 (formule du binôme dans un anneau)

Soit $(A, +, \times)$ un anneau, et soit a, b deux éléments de A . On suppose que $ab = ba$.

Alors, pour tout entier naturel n , on a la « formule du binôme » :
$$(a + b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}.$$

Remarque sur l'importance de l'hypothèse $ab = ba$

Dans les deux propositions précédentes, l'hypothèse selon laquelle a et b commutent est essentielle.

Évidemment, cela est automatiquement vérifié dans un anneau commutatif.

Si les deux éléments a et b ne commutent pas, le développement de $(a + b)^n$ est beaucoup plus compliqué.

On trouve par exemple :
$$\begin{cases} (a + b)^2 = a^2 + ab + ba + b^2 \\ (a + b)^3 = a^3 + a^2b + aba + ab^2 + ba^2 + bab + b^2a + b^3 \end{cases}$$

De même, si $ab \neq ba$, le produit $(a + b)(a - b)$ se développe en $a^2 - ab + ba - b^2$.

11.3.3 Structure de corps

Définition 11.3.2

Soit K un ensemble muni de deux lois $+$ et \times .

On dit que $(K, +, \times)$ est un *corps* si $(K, +, \times)$ est un anneau *commutatif* non réduit à $\{0\}$, et si tous les éléments non nuls de K sont inversibles pour le produit.

Exemples

$(\mathbb{Q}, +, \times)$, $(\mathbb{R}, +, \times)$ et $(\mathbb{C}, +, \times)$ sont des corps, mais pas $(\mathbb{Z}, +, \times)$.

L'ensemble $\{r + s\sqrt{2}, (r, s) \in \mathbb{Q}^2\}$ est un corps, contenant strictement \mathbb{Q} , contenu strictement dans \mathbb{R} .

Chapitre 12

Polynômes, fractions rationnelles

Sommaire

12.1 Anneau des polynômes	265
12.1.1 Suites à support fini de \mathbb{K}	265
12.1.2 L'anneau des polynômes $\mathbb{K}[X]$	266
12.1.3 Degré d'un polynôme	267
12.1.4 Degré d'une somme ou d'un produit	268
12.1.5 Composition de deux polynômes	269
12.1.6 Divisibilité dans $\mathbb{K}[X]$, diviseurs, multiples	269
12.1.7 Polynômes associés	270
12.1.8 Théorème de la division euclidienne	271
12.2 Fonctions polynomiales, racines	271
12.2.1 Fonction polynomiale associée	271
12.2.2 Racines (ou zéros) d'un polynôme	272
12.2.3 Nombre maximum de racines d'un polynôme	273
12.2.4 Polynômes scindés	274
12.2.5 Identification entre polynômes et fonctions polynomiales	274
12.2.6 Relations entre coefficients et racines	275
12.3 Dérivation des polynômes	276
12.3.1 Dérivée formelle d'un polynôme	276
12.3.2 Formule de Taylor polynomiale	277
12.4 Arithmétique dans $\mathbb{K}[X]$	279
12.4.1 Pgcd de deux polynômes A et B	279
12.4.2 Algorithme d'Euclide	279
12.4.3 Relation de Bézout	280
12.4.4 Ppcm de deux polynômes	281
12.4.5 Couples de polynômes premiers entre eux	282
12.4.6 Pgcd et ppcm de plusieurs polynômes	283
12.4.7 Polynômes premiers entre eux dans leur ensemble	284
12.5 Polynômes irréductibles et factorisations	284
12.5.1 Polynômes irréductibles dans $\mathbb{K}[X]$	284
12.5.2 Décomposition en facteurs irréductibles dans $\mathbb{C}[X]$	286
12.5.3 Décomposition en facteurs irréductibles dans $\mathbb{R}[X]$	287
12.6 Interpolation de Lagrange	289
12.7 Fractions rationnelles	291

12.7.1	La construction du corps $\mathbb{K}(X)$	291
12.7.2	Degré, partie entière	293
12.7.3	Zéros et pôles, multiplicités	293
12.8	Décomposition en éléments simples	294
12.8.1	Décomposition en éléments simples dans $\mathbb{C}(X)$	294
12.8.2	Décomposition en éléments simples dans $\mathbb{R}(X)$	295
12.8.3	Cas d'un pôle simple	296
12.8.4	Décomposition en éléments simples de P'/P	297
12.8.5	Pratique de la décomposition en éléments simples	298
12.8.6	Compléments sur quelques exemples	299

12.1 Anneau des polynômes

12.1.1 Suites à support fini de \mathbb{K}

Dans tout le chapitre, \mathbb{K} désigne le corps \mathbb{R} ou le corps \mathbb{C} .

Définition 12.1.1 (suites à support fini dans \mathbb{K})

On dit qu'une suite $(a_n)_{n \geq 0}$ de \mathbb{K} est à *support fini* s'il existe n_0 dans \mathbb{N} tel que : $\forall n \geq n_0, a_n = 0$.

Cela équivaut à dire que l'ensemble $\{n \in \mathbb{N}, a_n \neq 0\}$ est fini (éventuellement vide).

On note $\mathbb{K}^{(\mathbb{N})}$ l'ensemble des suites de \mathbb{K} qui sont à support fini.

Soit la suite $(a_n)_{n \geq 0}$ définie par $a_0 = 1, a_1 = -3, a_2 = 0, a_3 = 5, a_4 = -1$ et $a_n = 0$ pour tout $n \geq 5$.

Cette suite est à support fini, et on pourrait la noter, sans ambiguïté : $a = (1, -3, 0, 5, -1, 0\dots)$

Un cas très particulier est celui de la suite identiquement nulle $(0\dots)$.

Définition 12.1.2 (somme de suites à support fini, et produit par un scalaire)

Soit $a = (a_n)_{n \geq 0}$ et $b = (b_n)_{n \geq 0}$ deux suites à support fini de \mathbb{K} . Soit λ un élément de \mathbb{K} .

On note $a + b$ la suite de terme général $a_n + b_n$, et on note λa la suite de terme général λa_n .

Avec ces notations, les suites $a + b$ et λa sont encore à support fini.

Structure de groupe commutatif

L'ensemble $(\mathbb{K}^{(\mathbb{N})}, +)$ a une structure de groupe commutatif.

Plus précisément : l'élément neutre est la suite nulle, et la suite opposée à $(a_n)_{n \geq 0}$ est $(-a_n)_{n \geq 0}$.

Combinaisons linéaires

L'opération qui à un scalaire λ et à une suite $(a_n)_{n \geq 0}$ associe la suite λa est appelée « loi externe ».

Avec cette opération et la loi $+$, on peut définir des *combinaisons linéaires* dans $\mathbb{K}^{(\mathbb{N})}$.

Par exemple :

- soit $a = (a_n)_{n \geq 0}, b = (b_n)_{n \geq 0}$ et $c = (c_n)_{n \geq 0}$ sont trois éléments de $\mathbb{K}^{(\mathbb{N})}$
- soit α, β, γ trois scalaires (c'est-à-dire trois éléments de \mathbb{K})
- alors $d = \alpha a + \beta b + \gamma c$ désigne la suite (à support fini) dont le terme général est $d_n = \alpha a_n + \beta b_n + \gamma c_n$; on dit que d est une combinaison linéaire de a, b, c , avec les coefficients α, β, γ .

Base canonique

Pour tout m dans \mathbb{N} , notons $e_m = (a_n)_{n \geq 0}$ l'élément de $\mathbb{K}^{(\mathbb{N})}$ défini par
$$\begin{cases} a_m = 1 \\ a_n = 0 \text{ si } n \neq m \end{cases}$$

Ainsi $e_0 = (1, 0, \dots)$, $e_1 = (0, 1, 0, \dots)$, $e_2 = (0, 0, 1, 0, \dots)$.

On remarque par exemple que $a = (3, 0, 1, -2, 0, 0, 4, 0, \dots)$ s'écrit $a = 3e_0 + e_2 - 2e_3 + 4e_6$.

Plus généralement, si $a = (a_0, a_1, \dots, a_m, 0, \dots)$, alors $a = a_0e_0 + a_1e_1 + \dots + a_me_m = \sum_{n=0}^m a_n e_n$.

On notera aussi $a = \sum_{n \geq 0} a_n e_n$ ou $a = \sum_{n=0}^{+\infty} a_n e_n$, en se souvenant que cette somme est finie.

L'écriture de a comme *combinaison linéaire* des e_n est unique, à l'ordre près. On exprime cette propriété en disant que les suites e_n forment une *base* de $\mathbb{K}^{(\mathbb{N})}$ (dite « base canonique »).

Définition 12.1.3 (définition d'un produit sur les suites à support fini)

Soit $a = (a_n)_{n \geq 0}$ et $b = (b_n)_{n \geq 0}$ deux suites à support fini de \mathbb{K} .

On définit la suite $c = ab$ (dite suite produit de a et b) de la manière suivante :

$$c_0 = a_0 b_0, \quad c_1 = a_1 b_0 + a_0 b_1, \quad c_2 = a_2 b_0 + a_1 b_1 + a_0 b_2, \quad c_3 = a_3 b_0 + a_2 b_1 + a_1 b_2 + a_0 b_3$$

et plus généralement : $\forall n \in \mathbb{N}, c_n = \sum_{k=0}^n a_{n-k} b_k = \sum_{k=0}^n a_k b_{n-k} = \sum_{j+k=n} a_j b_k$.

Avec ces notations, la suite $c = ab$ est encore à support fini.

Remarques

- La loi produit sur $\mathbb{K}^{(\mathbb{N})}$ est commutative, associative, distributive par rapport à l'addition. La suite $e_0 = (1, 0, \dots)$ est élément neutre pour ce produit.
- Pour tous indices j et k , on remarque que $e_j e_k = e_{j+k}$.
On en déduit que pour tout n de \mathbb{N} , on a $e_1^n = e_n$ (en posant $e_1^0 = e_0$).

Proposition 12.1.1

Muni des deux lois $+$ et \times , l'ensemble $\mathbb{K}^{(\mathbb{N})}$ est un anneau commutatif.

Les éléments de cet anneau sont appelés polynômes à coefficients dans \mathbb{K} .

12.1.2 L'anneau des polynômes $\mathbb{K}[X]$ **Les scalaires sont des polynômes particuliers**

Pour tout scalaire λ , considérons la suite $(\lambda, 0, \dots)$, c'est-à-dire la suite λe_0 .

Pour toute suite $a = (a_n)_{n \geq 0}$ de $\mathbb{K}^{(\mathbb{N})}$, on a : $(\lambda e_0)a = \lambda a$.

Cette égalité permet d'identifier la suite λe_0 avec le scalaire λ (et donc e_0 avec 1)

Notation définitive des polynômes

On pose $X = e_1 = (0, 1, 0, \dots)$. Alors $X^2 = (0, 0, 1, 0, \dots) = e_2$, et $X^3 = (0, 0, 0, 1, 0, \dots) = e_3$.

Plus généralement, on a $e_n = X^n$ pour tout n de \mathbb{N}^* (et on complète par $X^0 = e_0 = 1$).

$P = (a_0, a_1, \dots, a_n, \dots) = (a_k)_{k \geq 0}$ s'écrit donc $P = a_0 + a_1 X + \dots + a_n X^n + \dots = \sum_{k \geq 0} a_k X^k = \sum_{k=0}^{+\infty} a_k X^k$

Une telle somme, toujours finie, représente P de façon unique (à l'ordre près).

Pour tout entier n tel que ($k > n \Rightarrow a_k = 0$), on peut bien sûr noter $P = \sum_{k=0}^n a_k X^k$.

On dit que les a_k sont les *coefficients* du polynôme P .

Unicité des coefficients et identification

L'unicité de l'écriture $P = \sum_{k \geq 0} a_k X^k$ permet de procéder à des identifications.

En particulier, P est le polynôme nul si et seulement si tous ses coefficients sont nuls.

De même, on a l'équivalence : $\sum_{k \geq 0} a_k X^k = \sum_{k \geq 0} b_k X^k \Leftrightarrow \forall k \in \mathbb{N}, a_k = b_k$.

Notation définitive des opérations sur $\mathbb{K}[X]$

On notera maintenant $\mathbb{K}[X]$ l'anneau des polynômes à coefficients dans \mathbb{K} .

Les lois sur $\mathbb{K}[X]$ sont donc définies par :

- le produit d'un polynôme par un scalaire : $\lambda \sum_{k \geq 0} a_k X^k = \sum_{k \geq 0} (\lambda a_k) X^k$
- la somme de deux polynômes : $\sum_{k \geq 0} a_k X^k + \sum_{k \geq 0} b_k X^k = \sum_{k \geq 0} (a_k + b_k) X^k$
- le produit de deux polynômes : $\left(\sum_{k \geq 0} a_k X^k \right) \left(\sum_{k \geq 0} b_k X^k \right) = \sum_{k \geq 0} \left(\sum_{i+j=k} a_i b_j \right) X^k$

Deux cas particuliers

On dit qu'un élément P de $\mathbb{K}[X]$ est un *monôme* s'il s'écrit $P = \lambda X^n$, avec λ dans \mathbb{K} et n dans \mathbb{N} .

On dit que P est un *polynôme constant* si $P = \lambda$, avec λ dans \mathbb{K} .

L'indéterminée

Dans les notations précédentes, le polynôme très particulier X est appelée *l'indéterminée*, et $\mathbb{K}[X]$ devient l'ensemble des polynômes à une indéterminée X . Le nom X est consacré par l'usage.

On pourrait généraliser et considérer des polynômes à plusieurs indéterminées (par exemple le polynôme $P = 1 + X + 2Y - XY + X^2Y$ est un élément de $\mathbb{K}[X, Y]$) mais cela dépasse le cadre du programme.

12.1.3 Degré d'un polynôme

Définition 12.1.4 (degré d'un polynôme)

Soit $P = \sum_{n \geq 0} a_n X^n$ un polynôme non nul de $\mathbb{K}[X]$.

On appelle *degré* de P , et on note $\deg(P)$, l'entier n maximum tel que $a_n \neq 0$.

Par convention, on pose $\deg(0) = -\infty$.

Exemples et remarques

- Le polynôme $P_\lambda = 1 + 3X^2 + X^3 - 2X^5 + \lambda X^7$ est de degré 7 si λ est non nul, et de degré 5 sinon.
- Si $P = \sum_{k \geq 0} a_k X^k$, on dit que a_k est le coefficient du terme ou du monôme de degré k dans P .
On dit que a_0 est le coefficient *constant* de P .

- Les polynômes sont en général écrits suivant les degrés croissants de l'indéterminée X ($P = 1 + 3X^2 + X^3 - 2X^5 + X^7$) ou suivant les degrés décroissants ($P = X^7 - 2X^5 + X^3 + 3X^2 + 1$). Ce choix n'est souvent qu'une question de confort (identification de coefficients, division de deux polynômes, etc.)
- Écrire que $\deg(P)$ appartient à \mathbb{N} (donc $\deg(P) \neq -\infty$), c'est écrire que P est non nul. Écrire que $\deg(P) \geq 1$, c'est écrire que P n'est pas un polynôme constant.

Définition 12.1.5 (polynômes unitaires, ou normalisés)

Soit $P = \sum_{n \geq 0} a_n X^n$ un polynôme non nul, et soit $d = \deg(P) \geq 0$.

On dit que a_d est le coefficient *dominant* de P (ou coefficient du terme de plus haut degré de P).

Si $a_d = 1$, on dit que le polynôme P est *normalisé* (ou encore *unitaire*).

Définition 12.1.6 (ensemble des polynômes de degré inférieur ou égal à n)

Soit n un entier naturel. On note $\mathbb{K}_n[X]$ l'ensemble des polynômes P tels que $\deg(P) \leq n$.

Par exemple : $\mathbb{K}_0[X]$ est l'ensemble des polynômes constants (donc $\mathbb{K}_0[X] = \mathbb{K}$).

De même : $\mathbb{K}_1[X] = \{aX + b, (a, b) \in \mathbb{K}^2\}$ et $\mathbb{K}_2[X] = \{aX^2 + bX + c, (a, b, c) \in \mathbb{K}^3\}$.

Un élément $P = \sum_{k=0}^n a_k X^k$ de $\mathbb{K}_n[X]$ est caractérisé par les $n + 1$ coefficients a_0, a_1, \dots, a_n .

12.1.4 Degré d'une somme ou d'un produit

On rappelle que, par convention, le degré du polynôme nul vaut $-\infty$.

Proposition 12.1.2 (degré d'une somme ou d'un produit)

Soit A et B deux éléments de $\mathbb{K}[X]$.

- on a $\deg(A + B) \leq \max(\deg(A), \deg(B))$, avec égalité si $\deg(A) \neq \deg(B)$.
- on a $\deg(AB) = \deg(A) + \deg(B)$.

– Ces résultats sont encore vrais si $A = 0$ ou $B = 0$, toujours avec la convention $\deg(0) = -\infty$.

– Si $\deg(A) = \deg(B) = n$, il est possible qu'on ait $\deg(A + B) < n$.

Il suffit pour cela que les termes de plus haut degré de A et B se « neutralisent ».

Par exemple, si $\begin{cases} A = X^3 + X \\ B = -X^3 + 1 \end{cases}$, alors $A + B = X + 1$, donc $\deg(A + B) = 1 < 3$.

On peut aussi prendre l'exemple extrême $B = -A$, car alors $\deg(A + B) = -\infty$.

– Pour tout A dans $\mathbb{K}[X]$ et pour tout λ dans \mathbb{K} , on a $\begin{cases} \deg(\lambda A) = \deg(A) & \text{si } \lambda \neq 0 \\ \deg(\lambda A) = -\infty & \text{si } \lambda = 0 \end{cases}$

– On a $\deg(A_1 A_2 \cdots A_m) = \sum_{k=1}^m \deg(A_k)$. En particulier $\deg(A^m) = m \deg(A)$.

Pour tous scalaires λ_k , on a $\deg\left(\sum_{k=1}^m \lambda_k A_k\right) \leq \max_{1 \leq k \leq m} (\deg(A_k))$.

Dans ce dernier résultat, si l'un des A_k est de degré strictement supérieur aux autres et si le coefficient λ_k correspondant est non nul, alors il y a égalité.

Proposition 12.1.3 (le produit de deux polynômes non nuls est non nul)

L'égalité $\deg(AB) = \deg(A) + \deg(B)$ montre que si $A \neq 0$ et $B \neq 0$ alors $AB \neq 0$.

Autrement dit : $AB = 0 \Rightarrow (A = 0 \text{ ou } B = 0)$.

Plus généralement, si $A_1 A_2 \dots A_n = 0$, alors l'un au moins des A_k est nul.

Une conséquence importante est que tout polynôme non nul est *simplifiable pour le produit*.

En effet, si A, B, C sont des polynômes et si $A \neq 0$, alors : $AB = AC \Rightarrow A(B - C) = 0 \Rightarrow B = C$.

Proposition 12.1.4 (polynômes inversibles pour le produit)

Soit A et B deux éléments de $\mathbb{K}[X]$.

On a $AB = 1$ si et seulement si A et B sont des constantes inverses l'une de l'autre.

Autrement dit, les seuls éléments inversibles pour le produit dans l'anneau $\mathbb{K}[X]$ sont les polynômes de degré 0, c'est-à-dire les polynômes constants non nuls.

12.1.5 Composition de deux polynômes

Définition 12.1.7 (composition de deux polynômes)

Soit $A = \sum_{n \geq 0} a_n X^n$ un élément de $\mathbb{K}[X]$. Pour tout polynôme B , on pose $A(B) = \sum_{n \geq 0} a_n B^n$.

On dit que $A(B)$ est le *composé* du polynôme B par le polynôme A .

Par exemple, posons $A = X^3 + X + 1$ et $B = X^2 - 1$.

Alors $A(B) = B^3 + B + 1 = (X^2 - 1)^3 + (X^2 - 1) + 1 = X^6 - 3X^4 + 4X^2 - 1$.

De même $B(A) = A^2 - 1 = (X^3 + X + 1)^2 - 1 = X^6 + 2X^4 + 2X^3 + X^2 + 2X$.

Proposition 12.1.5 (degré du composé de deux polynômes)

Soit A et B deux polynômes non nuls.

Soit $A(B)$ le polynôme composé de B par A . Alors $\deg(A(B)) = \deg(A) \deg(B)$.

Remarques et propriétés

Si $B = X$, alors $A(B) = A$. Ceci justifie qu'on note souvent $A(X)$ un polynôme de $\mathbb{K}[X]$.

Un cas classique est le calcul des polynômes $A(X + h)$, appelés *translatés* du polynôme A .

Par exemple : $A = aX^2 + bX + 1 \Rightarrow A(X + 1) = aX^2 + (2a + b)X + a + b + 1$.

Pour tous polynômes A, B, C , et pour tous scalaires λ, μ , on a
$$\begin{cases} (\lambda A + \mu B)(C) = \lambda A(C) + \mu B(C) \\ (AB)(C) = A(C)B(C) \end{cases}$$

12.1.6 Divisibilité dans $\mathbb{K}[X]$, diviseurs, multiples

Définition 12.1.8 (multiples et diviseurs)

Soit A et B deux éléments de $\mathbb{K}[X]$. On dit que B est un *diviseur* de A , ou encore que A est un *multiple* de B , et on note $B \mid A$, s'il existe un polynôme Q tel que $A = BQ$.

On note $\mathcal{D}(A)$ l'ensemble des diviseurs du polynôme A , et $A\mathbb{K}[X]$ l'ensemble de ses multiples.

Unicité du « quotient exact » quand il y a divisibilité

Si $A = BQ$ avec $B \neq 0$, alors Q (le *quotient exact* de A par B) est défini de façon unique.

Étant donnés deux polynômes A, B , avec $B \neq 0$, il est exceptionnel que B divise A .

Si cela se produit, on évitera de noter $\frac{A}{B}$ leur quotient exact.

Cas particuliers des polynômes 0 et 1

Pour tout polynôme B , on a $0 = QB$ avec $Q = 0$. Le polynôme nul est donc un multiple de tout polynôme B (ou encore : tout polynôme B de $\mathbb{K}[X]$ divise le polynôme nul). Ainsi $\mathcal{D}(0) = \mathbb{K}[X]$.

En revanche le polynôme nul ne divise que lui-même (car $A = Q0 \Rightarrow A = 0$), donc $0\mathbb{K}[X] = \{0\}$.

L'égalité évidente $B = 1A$ dit que le polynôme constant 1 divise tout polynôme B , ou encore que tout polynôme de $\mathbb{K}[X]$ est multiple du polynôme 1. Ainsi $1\mathbb{K}[X] = \mathbb{K}[X]$.

Mais seuls les polynômes constants non nuls divisent le polynôme 1 (car $B \mid 1$ signifie que B est inversible pour le produit, ou encore : $BQ = 1 \Rightarrow \deg(B) = 0$). Ainsi $\mathcal{D}(1) = \mathbb{K}^*$.

12.1.7 Polynômes associés

En posant $A \mid B$, on définit une relation binaire sur $\mathbb{K}[X]$.

Cette relation est réflexive (on a toujours $A \mid A$) et transitive (si $A \mid B$ et $B \mid C$ alors $A \mid C$).

Cette relation est loin d'être symétrique (si $A \mid B$, on n'a généralement pas $B \mid A$).

Elle n'est pas antisymétrique car $(A \mid B \text{ et } B \mid A) \not\Rightarrow A = B$, et la définition suivante précise ce point :

Définition 12.1.9 (polynômes associés)

On dit que deux polynômes A et B sont *associés* si A divise B et si B divise A .

Proposition 12.1.6 (caractérisation des couples de polynômes associés)

Soit (A, B) un couple de polynômes de $\mathbb{K}[X]$.

Alors A et B sont associés si et seulement si il existe un scalaire non nul λ tel que $B = \lambda A$.

Définition 12.1.10 (normalisé d'un polynôme non nul)

Soit A un polynôme non nul, et soit a_d son coefficient de plus haut degré.

Le polynôme unitaire $A^* = \frac{1}{a_d}A$ est appelé le *normalisé* du polynôme A .

Quelques propriétés

- Dire que deux polynômes non nuls sont associés, c'est dire qu'ils ont le même normalisé.
- Si deux polynômes unitaires A et B se divisent mutuellement, alors ils sont égaux.
- Pour tous polynômes A, B , on a les équivalences : $A\mathbb{K}[X] \subset B\mathbb{K}[X] \Leftrightarrow B \mid A \Leftrightarrow \mathcal{D}(B) \subset \mathcal{D}(A)$.
On en déduit, pour A, B non nuls : $A\mathbb{K}[X] = B\mathbb{K}[X] \Leftrightarrow A^* = B^* \Leftrightarrow \mathcal{D}(A) = \mathcal{D}(B)$.
- Soit A un polynôme non nul, et soit A^* son normalisé.
Alors A^* est l'unique polynôme unitaire tel que $\mathcal{D}(A^*) = \mathcal{D}(A)$.
De même A^* est l'unique polynôme unitaire tel que $A\mathbb{K}[X] = A^*\mathbb{K}[X]$.

12.1.8 Théorème de la division euclidienne

Proposition 12.1.7 (division euclidienne des polynômes)

Soit A et B deux éléments de $\mathbb{K}[X]$, avec $B \neq 0$.

Il existe un unique couple (Q, R) de polynômes tels que
$$\begin{cases} A = QB + R \\ \deg(R) < \deg B \end{cases}$$

Le passage du couple (A, B) au couple (Q, R) s'appelle division euclidienne de A par B .

Dans cette division, A est le dividende, B le diviseur, Q le quotient et R le reste.

Remarques

Il ne faut jamais oublier de mentionner la condition $\deg(R) < \deg(B)$.

Si $\deg(A) < \deg(B)$, la division euclidienne de A par B s'écrit $A = 0B + A$.

Si $B \neq 0$, dire que B divise A , c'est dire que le reste dans la division de A par B est nul.

Un exemple de division euclidienne

On divise ici le polynôme $A = X^5 + 2X^3 - X^2 - 4X + 3$ par le polynôme $B = X^2 + 3X + 1$.

$$\begin{array}{r|l} X^5 & + 2X^3 - X^2 - 4X + 3 \\ - 3X^4 & + X^3 - X^2 - 4X + 3 \\ & 10X^3 + 2X^2 - 4X + 3 \\ & - 28X^2 - 14X + 3 \\ & 70X + 31 \end{array} \quad \begin{array}{l} X^2 + 3X + 1 \\ \hline X^3 - 3X^2 + 10X - 28 \end{array}$$

Ainsi $A = BQ + R$ avec $Q = X^3 - 3X^2 + 10X - 28$ et $R = 70X + 31$.

Division euclidienne dans $\mathbb{R}[X]$ puis dans $\mathbb{C}[X]$

Soit A, B deux éléments de $\mathbb{R}[X]$, le polynôme B étant non nul.

Soit $A = BQ + R$ la division euclidienne de A par B dans $\mathbb{R}[X]$.

Par unicité, cette égalité représente aussi la division euclidienne de A par B dans $\mathbb{C}[X]$.

Cette propriété est souvent utilisée de la manière suivante : on part d'une division euclidienne dans $\mathbb{R}[X]$, et on la considère momentanément comme une division dans $\mathbb{C}[X]$, ce qui permet de substituer à X des valeurs complexes (notamment des racines complexes de B , voir plus loin).

12.2 Fonctions polynomiales, racines

12.2.1 Fonction polynomiale associée

Définition 12.2.1 (fonction polynomiale associée à un polynôme)

Soit $A = \sum_{k \geq 0} a_k X^k$ un élément de $\mathbb{K}[X]$. Pour tout x de \mathbb{K} , on pose $A(x) = \sum_{k \geq 0} a_k x^k$.

On dit que $A(x)$ est la *valeur* du polynôme A en x .

On dit que la fonction $x \mapsto A(x)$ est la *fonction polynomiale* associée au polynôme A .

On note souvent \tilde{A} cette fonction de \mathbb{K} dans \mathbb{K} , pour la distinguer du polynôme A lui-même.

Différence entre polynôme et fonction polynomiale

Un polynôme A de $\mathbb{K}[X]$ tel que nous l'avons défini, est un objet « formel ».

La fonction polynomiale \tilde{A} qui lui est associée est une fonction de \mathbb{K} dans \mathbb{K} .

On ne doit donc pas confondre A et \tilde{A} , même si l'égalité $A(\lambda) = \tilde{A}(\lambda)$ recèle une ambiguïté.

Quand on envisage $A(\lambda)$, il ne faut pas dire qu'on donne à X la valeur λ (ou qu'on pose $X = \lambda$) car ça n'a pas de sens : X est un polynôme de degré 1 et il ne saurait être égal à la constante λ .

En fait, il s'agit d'une simple substitution : on se contente de remplacer X par λ .

Pour éviter toute ambiguïté, il est d'usage d'utiliser le nom X quand on parle de polynômes et la variable x quand on parle de fonctions polynomiales.

Remarques et propriétés

– Si A est un polynôme, alors $A(0)$ est le coefficient constant de A .

De même, $A(1)$ représente la somme des coefficients de A .

– La fonction polynomiale associée au polynôme constant λ est la fonction constante $x \mapsto \lambda$.

La fonction polynomiale associée au polynôme X est la fonction identité $x \mapsto x$ de \mathbb{K} dans \mathbb{K} .

– On rappelle qu'on note \tilde{A} la fonction polynomiale associée à un polynôme A .

Avec ces notations, et pour tous A, B dans $\mathbb{K}[X]$, on a $\widetilde{A+B} = \tilde{A} + \tilde{B}$ et $\widetilde{AB} = \tilde{A} \tilde{B}$.

De même, pour tous polynômes A et B , on vérifie que $\widetilde{A(B)} = \tilde{A} \circ \tilde{B}$.

Valeurs en un point de \mathbb{C} d'un polynôme réel

Soit A dans $\mathbb{R}[X]$. On peut considérer A comme un élément particulier de $\mathbb{C}[X]$.

Alors pour tout nombre complexe z , on a l'égalité : $A(\bar{z}) = \overline{A(z)}$.

Plus généralement, soit $A = \sum_{n \geq 0} a_n X^n$ un polynôme de $\mathbb{C}[X]$. Posons $\bar{A} = \sum_{n \geq 0} \bar{a}_n X^n$.

Pour tous A, B dans $\mathbb{C}[X]$ et α, β dans \mathbb{C} , on a alors : $\overline{\alpha A + \beta B} = \bar{\alpha} \bar{A} + \bar{\beta} \bar{B}$ et $\overline{AB} = \bar{A} \bar{B}$.

12.2.2 Racines (ou zéros) d'un polynôme

Proposition 12.2.1 (division euclidienne par $X - \alpha$)

Soit A un élément de $\mathbb{K}[X]$ et soit α un scalaire.

Le reste dans la division de A par le polynôme $X - \alpha$ est le scalaire $A(\alpha)$.

Plus généralement si $A = QB + R$ et si $B(\alpha) = 0$ alors $A(\alpha) = R(\alpha)$.

Supposons par exemple qu'on veuille calculer $A(\alpha)$ avec $\deg(A) \geq 2$ et $\alpha = \frac{-1 + \sqrt{13}}{2}$.

Il est sans doute plus commode de diviser A par $B = X^2 + X - 3$ car $B(\alpha) = 0$.

Le reste R dans cette division s'écrit en effet $R = aX + b$ et on a alors $A(\alpha) = a\alpha + b$.

Définition 12.2.2 (racine (ou zéro) d'un polynôme)

Soit A un élément de $\mathbb{K}[X]$ et α un élément de \mathbb{K} .

On dit que α est une *racine* (ou encore un *zéro*) du polynôme A si $A(\alpha) = 0$.

Cela équivaut à dire que A est divisible par le polynôme $X - \alpha$.

Définition 12.2.3 (racines simples, racines multiples)

Soit A un polynôme non nul et soit α une racine de A dans \mathbb{K} .

On appelle *multiplicité* de α (comme racine de A) l'entier $m \geq 1$ tel que
$$\begin{cases} (X - \alpha)^m \mid A \\ (X - \alpha)^{m+1} \nmid A \end{cases}$$

– si $m = 1$, c'est-à-dire si $(X - \alpha) \mid A$ mais $(X - \alpha)^2 \nmid A$ on dit que α est une *racine simple* de A .

– si $m > 1$, c'est-à-dire si $(X - \alpha)^2 \mid A$, on dit que α est une *racine multiple* de A .

– si $m = 2$ (resp. $m = 3$) on dit que α est une *racine double* (resp. une *racine triple*) de A .

Proposition 12.2.2 (une caractérisation de la multiplicité)

Soit A dans $\mathbb{K}[X]$, soit α dans \mathbb{K} , et m dans \mathbb{N} .

Alors α est racine de A avec la multiplicité $m \Leftrightarrow \exists B \in \mathbb{K}[X], A = (X - \alpha)^m B$, avec $B(\alpha) \neq 0$

Remarques

Si α n'est pas racine de A , il est commode de dire que α est « racine de A avec la multiplicité 0 ».

Si α est une racine de multiplicité m de A , alors nécessairement $m \leq \deg(A)$.

Dire que A est divisible par $(X - \alpha)^m$, c'est dire que α est racine de A avec une multiplicité $\geq m$.

Dépendance par rapport au corps des coefficients

Il faut toujours préciser dans quel « corps de coefficients » on cherche les racines d'un polynôme.

Par exemple, si on considère $A = (X^2 - 2)(X^2 + 1)$: le polynôme A n'a pas de racine dans \mathbb{Q} , ses racines dans \mathbb{R} sont $-\sqrt{2}$ et $\sqrt{2}$, et ses racines dans \mathbb{C} sont $-\sqrt{2}$, $\sqrt{2}$, $-i$, et i .

Proposition 12.2.3 (racines complexes d'un polynôme à coefficients réels)

Soit A un polynôme de $\mathbb{R}[X]$. Soit α une racine de A dans \mathbb{C} de multiplicité m .

Alors $\bar{\alpha}$ est une racine de A dans \mathbb{C} , avec la même multiplicité m .

12.2.3 Nombre maximum de racines d'un polynôme

Quand on dénombre les racines d'un polynôme A , on peut soit considérer les racines distinctes de A (indépendamment de leur multiplicité) soit au contraire compter chaque racine autant de fois que sa multiplicité. Ainsi une racine double sera comptée deux fois, une racine triple trois fois, etc.

Proposition 12.2.4 (racines, multiplicités et factorisations)

Soit A un polynôme non nul de $\mathbb{K}[X]$.

Soit $\alpha_1, \alpha_2, \dots, \alpha_p$ dans \mathbb{K} , distincts deux à deux. Soit m_1, m_2, \dots, m_p dans \mathbb{N} .

Les deux propositions suivantes sont équivalentes :

– pour tout k de $\{1, \dots, p\}$, α_k est racine de A avec la multiplicité m_k .

– il existe Q dans $\mathbb{K}[X]$ tel que $A = Q \prod_{k=1}^p (X - \alpha_k)^{m_k}$, avec pour tout k de $\{1, \dots, p\}$, $Q(\alpha_k) \neq 0$.

Proposition 12.2.5 (nombre maximum de racines d'un polynôme)

Soit A un polynôme de $\mathbb{K}[X]$, de degré $n \geq 0$.

Le polynôme A admet au plus n racines dans \mathbb{K} , chacune étant comptée autant de fois que sa multiplicité.
En particulier, A admet au plus n racines distinctes dans \mathbb{K} .

12.2.4 Polynômes scindés**Proposition 12.2.6** (polynômes scindés)

Soit A dans $\mathbb{K}[X]$, de degré $n \geq 1$.

Soit $\alpha_1, \alpha_2, \dots, \alpha_p$ ses racines distinctes dans \mathbb{K} .

Soit m_1, m_2, \dots, m_p leurs multiplicités respectives (les $m_k \geq 1$).

On suppose que $m_1 + m_2 + \dots + m_p = n$ (autrement dit on suppose que A admet n racines dans \mathbb{K} , chacune étant comptée autant de fois que sa multiplicité).

Alors $A = \lambda \prod_{k=1}^p (X - \alpha_k)^{m_k}$, où λ est le coefficient dominant de A .

On exprime cette situation en disant que le polynôme A est scindé dans \mathbb{K} .

Dire que A est scindé dans \mathbb{K} c'est dire que A s'écrit sous la forme d'un produit $A = \lambda \prod_{k=1}^n (X - \beta_k)$ de facteurs de degré 1, où $n = \deg(A)$ et où β_1, \dots, β_n sont les racines (distinctes ou non) de A dans \mathbb{K} .

Proposition 12.2.7 (théorème de d'Alembert-Gauss (admis))

Tout polynôme de $\mathbb{C}[X]$, de degré $n \geq 1$, admet au moins une racine dans \mathbb{C} .

On exprime cette propriété en disant que le corps \mathbb{C} est algébriquement clos.

Proposition 12.2.8 (nombre de racines d'un polynôme à coefficients complexes)

Dans $\mathbb{C}[X]$, tout polynôme non constant est scindé.

Ainsi tout polynôme de degré $n \geq 1$ dans $\mathbb{C}[X]$ admet exactement n racines dans \mathbb{C} , chacune étant comptée autant de fois que sa multiplicité.

12.2.5 Identification entre polynômes et fonctions polynomiales**Proposition 12.2.9**

Soit A un polynôme de $\mathbb{K}[X]$, de degré inférieur ou égal à n .

On suppose que A s'annule en au moins $n + 1$ points distincts de \mathbb{K} .

Alors A est le polynôme nul.

Proposition 12.2.10

Soit A, B deux polynômes de $\mathbb{K}[X]$, de degré inférieur ou égal à n .

On suppose que A et B prennent la même valeur en au moins $n + 1$ points distincts de \mathbb{K} .

Alors les polynômes A et B sont identiques (autrement dit : ils ont les mêmes coefficients.)

Conséquence importante

Soit A et B deux polynômes à coefficients dans \mathbb{K} .

Si les fonctions polynomiales \tilde{A} et \tilde{B} sont égales, alors les polynômes A et B sont égaux.

Il suffit d'ailleurs que l'égalité $\tilde{A}(x) = \tilde{B}(x)$ soit vraie sur une partie infinie de l'ensemble \mathbb{K} pour qu'on ait l'égalité *formelle* des deux polynômes A et B (c'est-à-dire l'égalité de leurs coefficients).

En d'autres termes, l'application de $\mathbb{K}[X]$ dans $\mathcal{F}(\mathbb{K}, \mathbb{K})$ qui à un polynôme A associe la fonction polynomiale \tilde{A} est injective (elle réalise donc une bijection de $\mathbb{K}[X]$ sur son image, c'est-à-dire sur l'ensemble de toutes les fonctions polynomiales).

La fonction polynomiale \tilde{A} est par définition connue dès qu'on connaît A . Les remarques précédentes disent que réciproquement, un polynôme A (c'est-à-dire ses *coefficients*) est entièrement déterminé par la fonction polynomiale associée \tilde{A} (c'est-à-dire par l'ensemble des *valeurs* de A).

12.2.6 Relations entre coefficients et racines

Définition 12.2.4 (fonctions symétriques élémentaires)

Soit x_1, x_2, \dots, x_n une famille de n éléments de \mathbb{K} .

Pour tout k de $\{1, \dots, n\}$, on note $\sigma_k(x_1, x_2, \dots, x_n) = \sum_{i_1 < i_2 < \dots < i_k} x_{i_1} x_{i_2} \cdots x_{i_k}$ (somme des produits k à k)

En particulier : $\sigma_1 = \sum_{i=1}^n x_i$, $\sigma_2 = \sum_{i < j} x_i x_j$ (somme des produits deux à deux) et $\sigma_n = \prod_{i=1}^n x_i$

Les quantités précédentes sont appelées *fonctions symétriques élémentaires* de x_1, x_2, \dots, x_n .

Exemples et remarques

Chaque expression σ_k est une somme de $\binom{n}{k}$ termes.

Les fonctions symétriques élémentaires de a, b sont $\sigma_1 = a + b$ et $\sigma_2 = ab$.

Celles de a, b, c sont $\sigma_1 = a + b + c$, $\sigma_2 = ab + ac + bc$ et $\sigma_3 = abc$.

Celles de a, b, c, d sont :
$$\begin{cases} \sigma_1 = a + b + c + d, & \sigma_2 = ab + ac + ad + bc + bd + cd \\ \sigma_3 = abc + abd + acd + bcd, & \sigma_4 = abcd \end{cases}$$

Fonctions symétriques, élémentaires ou non

Soit s une permutation de $\{1, 2, \dots, n\}$ et soit k un élément de $\{1, 2, \dots, n\}$.

On a $\sigma_k(x_1, \dots, x_n) = \sigma_k(x_{s(1)}, \dots, x_{s(n)})$, ce qui justifie le nom de « fonctions symétriques ».

Elles sont dites « élémentaires » parce que de nombreuses autres fonctions symétriques de x_1, x_2, \dots, x_n peuvent s'écrire en fonction de $\sigma_1, \sigma_2, \dots, \sigma_n$.

Ainsi $S_2 = x_1^2 + x_2^2 + \dots + x_n^2 = \sigma_1^2 - 2\sigma_2$, ou encore $\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_n} = \frac{\sigma_{n-1}}{\sigma_n}$.

Proposition 12.2.11 (relations coefficients-racines, pour un polynôme scindé)

Soit $A = a_n X^n + a_{n-1} X^{n-1} + \dots + a_1 X + a_0$ un polynôme de $\mathbb{K}[X]$, avec $\deg(A) = n \geq 1$.

On écrit $A = a_n \prod_{k=1}^n (X - \alpha_k)$ où $\alpha_1, \dots, \alpha_n$ sont les racines (non nécessairement distinctes) de A .

Soit $\sigma_1, \sigma_2, \dots, \sigma_n$ les fonctions symétriques élémentaires de $\alpha_1, \alpha_2, \dots, \alpha_n$.

Alors on a les égalités : $\forall k \in \{1, \dots, n\}, \sigma_k = (-1)^k \frac{a_{n-k}}{a_n}$.

Exemples

– si $P = aX^2 + bX + c = a(X - \alpha)(X - \beta)$, alors $\sigma_1 = \alpha + \beta = -\frac{b}{a}$ et $\sigma_2 = \alpha\beta = \frac{c}{a}$.

– si $P = aX^3 + bX^2 + cX + d = a(X - \alpha)(X - \beta)(X - \gamma)$, alors :

$$\sigma_1 = \alpha + \beta + \gamma = -\frac{b}{a}, \quad \sigma_2 = \alpha\beta + \alpha\gamma + \beta\gamma = \frac{c}{a}, \quad \sigma_3 = \alpha\beta\gamma = -\frac{d}{a}$$

Proposition 12.2.12 (utilisation des valeurs des fonctions symétriques élémentaires)

Soit $\sigma_1, \dots, \sigma_n$ les fonctions symétriques élémentaires de n scalaires x_1, \dots, x_n .

L'ensemble $\{x_1, x_2, \dots, x_n\}$ est égal à l'ensemble des solutions de l'équation :

$$x^n - \sigma_1 x^{n-1} + \dots + (-1)^{n-1} \sigma_{n-1} x + (-1)^n \sigma_n = 0.$$

Exemples

– Soit (E) l'équation $t^2 - St + P = 0$, d'inconnue t dans \mathbb{K} .

Soit \mathcal{S} l'ensemble de ses solutions dans \mathbb{K} . On a : $\{x, y\} = \mathcal{S} \Leftrightarrow \begin{cases} x + y = S \\ xy = P \end{cases}$

– Soit (E) l'équation $t^3 - St^2 + \lambda t - P = 0$, d'inconnue t dans \mathbb{K} .

Soit \mathcal{S} l'ensemble de ses solutions dans \mathbb{K} .

On a $\{x, y, z\} = \mathcal{S}$ si et seulement si le triplet (x, y, z) est solution du système $\begin{cases} x + y + z = S \\ xy + xz + yz = \lambda \\ xyz = P \end{cases}$

Un exemple d'utilisation des fonctions symétriques élémentaires

On considère le polynôme $A = X^3 + X^2 + 3X + 2$. Soit α, β, γ ses racines dans \mathbb{C} .

On se propose de calculer les sommes $S_n = \alpha^n + \beta^n + \gamma^n$, pour n dans \mathbb{N} .

On a tout d'abord $S_0 = 3$, $S_1 = \sigma_1 = -1$, $S_2 = \sigma_1^2 - 2\sigma_2 = -5$.

$$\text{On a } \begin{cases} \alpha^3 + \alpha^2 + 3\alpha + 2 = 0 \\ \beta^3 + \beta^2 + 3\beta + 2 = 0 \\ \gamma^3 + \gamma^2 + 3\gamma + 2 = 0 \end{cases} \text{ puis, pour tout } n \geq 3 : \begin{cases} \alpha^n + \alpha^{n-1} + 3\alpha^{n-2} + 2\alpha^{n-3} = 0 \\ \beta^n + \beta^{n-1} + 3\beta^{n-2} + 2\beta^{n-3} = 0 \\ \gamma^n + \gamma^{n-1} + 3\gamma^{n-2} + 2\gamma^{n-3} = 0 \end{cases}$$

On en déduit, après addition terme à terme : $\forall n \geq 3$, $S_n = -S_{n-1} - 3S_{n-2} - 2S_{n-3}$.

Ainsi $S_3 = -S_2 - 3S_1 - 2S_0 = 2$, $S_4 = -S_3 - 3S_2 - 2S_1 = 15$, etc.

12.3 Dérivation des polynômes

12.3.1 Dérivée formelle d'un polynôme

On rappelle que \mathbb{K} désigne indifféremment \mathbb{R} ou \mathbb{C} .

Définition 12.3.1 (polynôme dérivé formel)

Soit $A = \sum_{k \geq 0} a_k X^k$ dans $\mathbb{K}[X]$. Le polynôme $A' = \sum_{k \geq 0} (k+1)a_{k+1} X^k$ est appelé *polynôme dérivé* de A .

Il s'agit ici d'une dérivée « formelle », c'est-à-dire purement symbolique.

Mais si $\mathbb{K} = \mathbb{R}$, alors la fonction polynomiale associée au polynôme A' est bien la dérivée (au sens habituel donné à ce nom) de la fonction polynomiale associée à A .

En revanche si $\mathbb{K} = \mathbb{C}$, ça n'a aucun sens de parler de fonction dérivée.

Remarques et propriétés

- Si $A = aX^3 + bX^2 + cX + d$ alors $A' = 3aX^2 + 2bX + c$.
- Si $A = \sum_{k=0}^n a_k X^k$, alors $A' = \sum_{k=0}^{n-1} (k+1)a_{k+1}X^k$, ou encore $A' = \sum_{k=1}^n k a_k X^{k-1}$.
- Si $\deg(A) \geq 1$ on a $\deg(A') = \deg(A) - 1$.
Pour cette raison, A' est le polynôme nul si et seulement si A est un polynôme constant.
Plus généralement : $\forall (A, B) \in \mathbb{K}[X]^2$, $A' = B' \Leftrightarrow \exists \lambda \in \mathbb{K}, A = B + \lambda$.

Définition 12.3.2 (polynômes dérivés successifs)

Soit A dans $\mathbb{K}[X]$. On définit les polynômes dérivés successifs de A en posant $\begin{cases} A^{(0)} = A \\ \forall m \in \mathbb{N}, A^{(m+1)} = (A^{(m)})' \end{cases}$
On dit que $A^{(m)}$ est le polynôme dérivé m -ième de A .

On note bien sûr A' et A'' plutôt que $A^{(1)}$ et $A^{(2)}$.

Proposition 12.3.1 (polynômes dérivés d'une combinaison linéaire)

Soit A, B dans $\mathbb{K}[X]$, et α, β dans \mathbb{K} . Pour tout m de \mathbb{N} , on a $(\alpha A + \beta B)^{(m)} = \alpha A^{(m)} + \beta B^{(m)}$.

Si $m \leq k$, on a $(X^k)^{(m)} = k(k-1) \cdots (k-m+1)X^{k-m} = \frac{k!}{(k-m)!}X^{k-m}$.

En particulier, $(X^m)^{(m)} = m!$, et $(X^k)^{(m)} = 0$ si $m > k$.

Plus généralement, si $A = \sum_{k \geq 0} a_k X^k$, alors $A^{(m)} = \sum_{k \geq m} \frac{k!}{(k-m)!} a_k X^{k-m} = \sum_{k \geq 0} \frac{(k+m)!}{k!} a_{k+m} X^k$.

Dérivation et degré

Si $m \leq \deg(A)$, alors $\deg(A^{(m)}) = \deg(A) - m$, mais si $m > \deg A$ alors $A^{(m)} = 0$.

On a donc $A^{(m)} = 0$ si et seulement si A est un polynôme de degré strictement inférieur à m .

Si $A = \sum_{k \geq 0} a_k X^k$ est de degré n , alors $A^{(n)}$ est le polynôme constant $n! a_n$.

Proposition 12.3.2 (formule de Leibniz pour les polynômes)

Soit A, B dans $\mathbb{K}[X]$, et m dans \mathbb{N} . On a $(AB)^{(m)} = \sum_{k=0}^m \binom{m}{k} A^{(k)} B^{(m-k)}$.

On retrouve $(AB)' = A'B + AB'$, mais on a aussi $\begin{cases} (AB)'' = A''B + 2A'B' + AB'' \\ (AB)''' = A'''B + 3A''B' + 3A'B'' + AB''' \end{cases}$

On se méfiera de l'analogie entre $(AB)^{(n)}$ dans $\mathbb{K}[X]$ et $(a+b)^n$ dans \mathbb{K} .

En effet on a $A^{(0)} = A$ et $B^{(0)} = B$ « aux extrémités », alors que dans \mathbb{K} on a $a^0 = b^0 = 1$.

12.3.2 Formule de Taylor polynomiale**Proposition 12.3.3** (formule de Taylor en un point)

Soit $A = \sum_{k \geq 0} a_k X^k$ un élément de $\mathbb{K}[X]$, et soit λ un élément de \mathbb{K} .

On a l'égalité : $A = A(\lambda) + A'(\lambda)(X - \lambda) + \frac{A''(\lambda)}{2!}(X - \lambda)^2 + \cdots = \sum_{k \geq 0} \frac{A^{(k)}(\lambda)}{k!}(X - \lambda)^k$

Remarques

La somme précédente est finie, et si $\deg(A) = n$ son dernier terme est $\frac{A^{(n)}(\lambda)}{n!}(X - \lambda)^n$.

La formule de Taylor montre qu'un polynôme est déterminé par ses dérivées successives en un point.

Le cas particulier $\lambda = 0$ est connu sous le nom de *formule de Mac Laurin*.

Si $A = \sum_{k \geq 0} a_k X^k$, cette formule s'écrit $A = \sum_{k \geq 0} \frac{A^{(k)}(0)}{k!} X^k$. Ainsi : $\forall k \geq 0, a_k = \frac{A^{(k)}(0)}{k!}$.

On a l'équivalence : $A(X) = \sum_{k \geq 0} \frac{A^{(k)}(\lambda)}{k!} (X - \lambda)^k \Leftrightarrow A(X + \lambda) = \sum_{k \geq 0} \frac{A^{(k)}(\lambda)}{k!} X^k$.

Un exemple

On considère le polynôme $A = X^4 + 2X^3 - X + 1$ et le scalaire $\lambda = 2$:

On trouve $A' = 4X^3 + 6X^2 - 1$, puis $A'' = 12X^2 + 12X$, $A^{(3)} = 24X + 12$, $A^{(4)} = 24$ et $A^{(5)} = 0$.

Ainsi $A(2) = 31$, $A'(2) = 55$, $A''(2) = 72$, $A^{(3)}(2) = 60$, $A^{(4)}(2) = 24$ et $A^{(n)}(2) = 0$ pour $n \geq 5$.

La formule de Taylor de A au point 2 s'écrit donc :

$$\begin{aligned} A(X) &= A(2) + A'(2)(X - 2) + \frac{A''(2)}{2!}(X - 2)^2 + \frac{A^{(3)}(2)}{3!}(X - 2)^3 + \frac{A^{(4)}(2)}{4!}(X - 2)^4 \\ &= 31 + 55(X - 2) + 36(X - 2)^2 + 10(X - 2)^3 + (X - 2)^4 \end{aligned}$$

On aurait tout aussi bien pu écrire :

$$\begin{aligned} A(X + 2) &= (X + 2)^4 + 2(X + 2)^3 - (X + 2) + 1 \\ &= (X^4 + 8X^3 + 24X^2 + 32X + 16) + 2(X^3 + 6X^2 + 12X + 8) - (X + 2) + 1 \\ &= X^4 + 10X^3 + 36X^2 + 55X + 31 \end{aligned}$$

et on retrouve : $A(X) = 31 + 55(X - 2) + 36(X - 2)^2 + 10(X - 2)^3 + (X - 2)^4$.

Proposition 12.3.4 (caractérisation de la multiplicité par les dérivées successives)

Soit A dans $\mathbb{K}[X]$, α dans \mathbb{K} , et m dans \mathbb{N} .

Alors α est racine de A avec la multiplicité $m \Leftrightarrow \begin{cases} A(\alpha) = A'(\alpha) = \dots = A^{(m-1)}(\alpha) = 0 \\ A^{(m)}(\alpha) \neq 0 \end{cases}$

La multiplicité d'une racine α de A est donc l'ordre de la première dérivée non nulle en α .

Exemples

$$\alpha \text{ est racine simple} \Leftrightarrow \begin{cases} A(\alpha) = 0 \\ A'(\alpha) \neq 0 \end{cases}$$

$$\alpha \text{ est racine multiple} \Leftrightarrow \begin{cases} A(\alpha) = 0 \\ A'(\alpha) = 0 \end{cases}$$

$$\alpha \text{ est racine double} \Leftrightarrow \begin{cases} A(\alpha) = A'(\alpha) = 0 \\ A''(\alpha) \neq 0 \end{cases}$$

$$\alpha \text{ est racine triple} \Leftrightarrow \begin{cases} A(\alpha) = A'(\alpha) = A''(\alpha) = 0 \\ A'''(\alpha) \neq 0 \end{cases}$$

Proposition 12.3.5 (racines de A et racines de A')

Soit A dans $\mathbb{K}[X]$, α dans \mathbb{K} , et m dans \mathbb{N} .

Si α est racine de A de multiplicité $m \geq 1$, alors α est racine de A' de multiplicité $m - 1$.

Si α est racine de A' de multiplicité m , et si $A(\alpha) = 0$, alors α est racine de A de multiplicité $m + 1$.

12.4 Arithmétique dans $\mathbb{K}[X]$

12.4.1 Pgcd de deux polynômes A et B

Rappelons que l'ensemble des diviseurs du polynôme nul est $\mathcal{D}(0) = \mathbb{K}[X]$.

En revanche, si $A \neq 0$, tous les polynômes de $\mathcal{D}(A)$ sont de degré inférieur ou égal à $\deg(A)$.

Dans la suite, on considère deux polynômes A et B de $\mathbb{K}[X]$ dont l'un au moins est non nul.

Il en résulte que $\mathcal{D}(A) \cap \mathcal{D}(B)$ contient des polynômes non nuls (par exemple le polynôme 1), mais que ceux-ci sont tous de degré inférieur ou égal à $\min(\deg(A), \deg(B))$.

Définition 12.4.1 (pgcd de A et B , avec $A \neq 0$ ou $B \neq 0$)

Soit A et B deux polynômes de $\mathbb{K}[X]$, dont l'un au moins est non nul.

L'ensemble des diviseurs communs à A et B possède des polynômes de degré maximum $d \geq 0$.

Tout polynôme D de $\mathcal{D}(A) \cap \mathcal{D}(B)$ et de degré d est appelé *un* pgcd de A et de B .

Contrairement au pgcd des entiers, le pgcd de deux polynômes n'est pas défini de façon unique.

Plus précisément, on a le résultat suivant :

Proposition 12.4.1 (le pgcd est défini « à association près »)

Soit A et B deux polynômes de $\mathbb{K}[X]$, dont l'un au moins est non nul.

Soit D un pgcd de A et B . Soit Δ un polynôme de $\mathbb{K}[X]$.

Alors Δ est un pgcd de A et B si et seulement si D et Δ sont associés.

En particulier, A et B possèdent un pgcd unitaire et un seul. Celui est noté $A \wedge B$.

Premières propriétés

Rappelons qu'on note A^* le normalisé d'un polynôme A non nul.

Pour tous polynômes A et B , on a : $A \wedge B = B \wedge A$, $A \wedge 0 = A^*$ (si $A \neq 0$), et $A \wedge 1 = 1$.

On a l'égalité $A \wedge B = A^*$ si et seulement si A est un diviseur de B .

Le polynôme $A \wedge B$, et ses diviseurs, sont des diviseurs communs à A et B .

On verra que la réciproque est vraie : les diviseurs communs à A et B sont aussi des diviseurs de $A \wedge B$.

On peut (éventuellement) poser $0 \wedge 0 = 0$.

12.4.2 Algorithme d'Euclide

Proposition 12.4.2 (pgcd et division euclidienne)

Soit A , B et Q trois polynômes.

Alors $\mathcal{D}(A) \cap \mathcal{D}(B) = \mathcal{D}(B) \cap \mathcal{D}(A - QB)$

Autrement dit, les diviseurs communs de A et B sont exactement ceux de B et $A - QB$.

En particulier, si R est le reste dans la division de A par B , on a $\mathcal{D}(A) \cap \mathcal{D}(B) = \mathcal{D}(B) \cap \mathcal{D}(R)$.

Un application répétée de ce principe conduit à l'algorithme d'Euclide :

Proposition 12.4.3 (algorithme d'Euclide)

Soit A et B deux polynômes non nuls. On veut calculer $A \wedge B$.

On forme une suite finie de polynômes R_k , à commencer par $R_0 = A$ et $R_1 = B$.

Soit $k \geq 1$. On suppose que R_{k-1} et R_k sont connus.

Si R_k est non nul, on note $R_{k-1} = Q_k R_k + R_{k+1}$ la division euclidienne de R_{k-1} par R_k .

Sous l'hypothèse $R_k \neq 0$, on a donc défini R_{k+1} , avec $0 \leq \deg(R_{k+1}) < \deg(R_k)$.

La suite de polynômes $(R_k)_{k \geq 1}$ est finie car la suite de leurs degrés est strictement décroissante.

Il existe donc un n de \mathbb{N} tel que $R_n \neq 0$ et $R_{n+1} = 0$.

Avec ces notations, on a : $A \wedge B = R_n^*$.

Ainsi $A \wedge B$ est le normalisé du dernier reste non nul dans cette succession de divisions.

On peut considérer l'algorithme d'Euclide comme une nouvelle définition du pgcd (c'est le normalisé du dernier reste non nul dans la succession des divisions).

Les propriétés suivantes sont alors des conséquences de cette définition algorithmique.

On sait par exemple (depuis la toute première définition) que le polynôme $A \wedge B$ divise A et B .

Mais l'algorithme d'Euclide donne la réciproque : si un polynôme divise A et B , alors il divise leur pgcd.

On aboutit ainsi à la caractérisation suivante du polynôme unitaire $D = A \wedge B$:

Proposition 12.4.4 (caractérisation du pgcd)

Soit A et B deux polynômes de $\mathbb{K}[X]$, dont l'un au moins est non nul. Soit $D = A \wedge B$.

Le polynôme D est le seul polynôme unitaire tel que $\mathcal{D}(D) = \mathcal{D}(A) \cap \mathcal{D}(B)$.

La proposition suivante montre qu'on peut effectuer une mise en facteur (ou une simplification par un facteur commun) dans un calcul de pgcd.

Proposition 12.4.5 (mise en facteur dans un calcul de pgcd)

Soit A et B deux polynômes de $\mathbb{K}[X]$, dont l'un au moins est non nul.

Pour tout polynôme non nul C , on a l'égalité : $(CA) \wedge (CB) = C^* (A \wedge B)$.

12.4.3 Relation de Bézout

Proposition 12.4.6 (relation de Bézout entre deux polynômes)

Soit A et B deux polynômes de $\mathbb{K}[X]$, dont l'un au moins est non nul.

Alors il existe des polynômes U_0 et V_0 tels que $A \wedge B = AU_0 + BV_0$.

Une telle égalité s'appelle une relation de Bézout entre A et B .

On dit que le couple (U_0, V_0) constitue un couple de coefficients de Bézout de A et B .

Proposition 12.4.7 (combinaisons $AU + BV$ de deux polynômes A et B)

Soit A et B deux polynômes de $\mathbb{K}[X]$, dont l'un au moins est non nul.

L'ensemble $\{AU + BV, U \in \mathbb{K}[X], V \in \mathbb{K}[X]\}$ est exactement égal à l'ensemble des multiples de $A \wedge B$.

Recherche d'un couple de coefficients de Bézout

Pour trouver un couple de coefficients de Bézout de deux polynômes non nuls A et B , il suffit de « remonter les calculs » dans l'algorithme d'Euclide (s'inspirer de ce qui a été fait avec les entiers).

12.4.4 Ppcm de deux polynômes

Dans la suite, on considère deux polynômes non nuls A et B .

L'intersection $A\mathbb{K}[X] \cap B\mathbb{K}[X]$ désigne l'ensemble des multiples communs à A et B .

Cet ensemble contient des polynômes non nuls, et par exemple le produit AB .

L'ensemble des degrés de ces multiples communs non nuls possède donc un plus petit élément, ce qui conduit à la définition suivante :

Définition 12.4.2 (ppcm de deux polynômes non nuls)

Soit A et B deux polynômes non nuls de $\mathbb{K}[X]$.

Soit m le degré minimum d'un polynôme parmi tous les multiples communs non nuls de A et B .

Tout multiple commun de A et B , de degré m , est appelé un ppcm de A et B .

Tout comme le pgcd, le ppcm de deux polynômes n'est pas défini de façon unique.

Plus précisément, on a le résultat suivant :

Proposition 12.4.8 (le ppcm est défini « à association près »)

Soit A et B deux polynômes non nuls de $\mathbb{K}[X]$.

Soit M un ppcm de A et B . Soit N un polynôme de $\mathbb{K}[X]$.

Alors N est un ppcm de A et B si et seulement si M et N sont associés.

En particulier, A et B possèdent un ppcm unitaire et un seul. Celui est noté $A \vee B$.

Premières propriétés

Il est clair qu'on a les égalités $A \vee B = B \vee A$, et $A \vee 1 = A^*$ (normalisé de A).

On a l'égalité $A \vee B = A^*$ si et seulement si A est un multiple de B .

Le polynôme $A \vee B$ et ses multiples sont des multiples communs aux polynômes A et B .

La proposition suivante dit que la réciproque est vraie :

Proposition 12.4.9 (caractérisation du ppcm de deux polynômes)

Soit A et B deux polynômes non nuls de $\mathbb{K}[X]$.

Le polynôme $A \vee B$ est l'unique polynôme normalisé M tel que $A\mathbb{K}[X] \cap B\mathbb{K}[X] = M\mathbb{K}[X]$.

Autrement dit, les multiples communs de A et B sont exactement ceux de $A \vee B$.

Si $A = 0$ ou $B = 0$, on peut poser $A \vee B = 0$, ce qui est compatible avec la caractérisation précédente.

Proposition 12.4.10 (mise en facteur dans un calcul de ppcm)

Soit A, B, C trois polynômes non nuls de $\mathbb{K}[X]$.

Alors on a l'égalité : $(CA) \vee (CB) = C^*(A \vee B)$.

Proposition 12.4.11 (lien entre le pgcd et le ppcm)

Soit A, B deux polynômes non nuls de $\mathbb{K}[X]$. Alors on a l'égalité : $(A \wedge B)(A \vee B) = (AB)^*$.

Conséquence : le polynôme $A \vee B$ est le normalisé du quotient de AB par $A \wedge B$.

12.4.5 Couples de polynômes premiers entre eux

Définition 12.4.3 (polynômes premiers entre eux)

Soit A et B deux polynômes. On dit que A et B sont *premiers entre eux* si $A \wedge B = 1$.

Cela revient à dire que les seuls diviseurs communs à A et B sont les polynômes constants non nuls.

Dire que A et B sont premiers entre eux, c'est dire que le dernier reste non nul dans leur algorithme d'Euclide est un polynôme de degré 0 (c'est-à-dire une constante non nulle).

Proposition 12.4.12 (théorème de Bézout)

Soit A et B deux polynômes. Les deux propositions suivantes sont équivalentes :

- les polynômes A et B sont premiers entre eux.
- il existe deux polynômes U et V tels que $AU + BV = 1$.

Le théorème de Bézout a des conséquences fondamentales en arithmétique des polynômes :

Proposition 12.4.13 (lemme de Gauss)

Soit A, B, C trois polynômes. Si A divise BC et si A est premier avec B , alors A divise C .

Proposition 12.4.14 (résolution de $AU + BV = 1$ quand A et B sont premiers entre eux)

Soit A et B deux polynômes premiers entre eux.

Alors il existe une infinité de couples (U, V) de $\mathbb{K}[X]^2$ tels que $AU + BV = 1$.

Si (U_0, V_0) est l'un d'eux, les autres sont donnés par
$$\begin{cases} U = U_0 + QB \\ V = V_0 - QA \end{cases} \text{ avec } Q \text{ dans } \mathbb{K}[X].$$

Proposition 12.4.15 (polynôme premier avec un produit)

Soit A, B, C trois polynômes. Les deux propositions suivantes sont équivalentes :

- le polynôme A est premier avec le polynôme B et avec le polynôme C .
- le polynôme A est premier avec le produit BC .

Plus généralement, soit A_1, A_2, \dots, A_m et B_1, B_2, \dots, B_n deux familles de polynômes.

Les deux propositions suivantes sont équivalentes :

- chacun des polynômes A_1, A_2, \dots, A_m est premier avec chacun des polynômes B_1, B_2, \dots, B_n .
- le produit $A_1 A_2 \cdots A_m$ est premier avec le produit $B_1 B_2 \cdots B_n$.

Par exemple, si $A \wedge B = 1$, alors $A^m \wedge B^n = 1$ pour tous entiers naturels m et n .

Proposition 12.4.16 (divisibilité par deux polynômes premiers entre eux)

Soit A, B, C trois polynômes, A et B étant supposés premiers entre eux.

Le polynôme C est divisible par A et par B si et seulement si il est divisible par leur produit.

Plus généralement si les polynômes A_1, A_2, \dots, A_n sont premiers entre eux deux à deux, le polynôme C est divisible par chacun des A_k si et seulement si il est divisible par leur produit.

Quotient de deux polynômes par leur pgcd

Soit A et B deux polynômes (non tous deux nuls).

Soit \hat{A} et \hat{B} les polynômes tels que $A = (A \wedge B)\hat{A}$ et $B = (A \wedge B)\hat{B}$. Alors $\hat{A} \wedge \hat{B} = 1$.

Cas particulier des polynômes $(X - \lambda)^m$

Si λ et μ sont deux scalaires distincts, alors $(X - \lambda) \wedge (X - \mu) = 1$.

Plus généralement, soit $\lambda_1, \lambda_2, \dots, \lambda_n, \mu_1, \mu_2, \dots, \mu_p$ des scalaires distincts deux à deux.

Soit $\alpha_1, \alpha_2, \dots, \alpha_n, \beta_1, \beta_2, \dots, \beta_p$ des entiers naturels.

Alors les polynômes $\begin{cases} A = (X - \lambda_1)^{\alpha_1} (X - \lambda_2)^{\alpha_2} \dots (X - \lambda_n)^{\alpha_n} \\ B = (X - \mu_1)^{\beta_1} (X - \mu_2)^{\beta_2} \dots (X - \mu_p)^{\beta_p} \end{cases}$ sont premiers entre eux.

12.4.6 Pgcd et ppcm de plusieurs polynômes

Proposition 12.4.17 (associativité du pgcd et du ppcm)

Pour tous polynômes A, B, C , on a les égalités $\begin{cases} A \wedge (B \wedge C) = (A \wedge B) \wedge C \\ A \vee (B \vee C) = (A \vee B) \vee C \end{cases}$

L'associativité (et la commutativité) du pgcd et du ppcm ont pour conséquence que si on se donne n polynômes A_1, A_2, \dots, A_n , alors les expressions $A_1 \wedge A_2 \wedge \dots \wedge A_n$ et $A_1 \vee A_2 \vee \dots \vee A_n$ ont un sens, indépendamment de l'ordre des facteurs A_k et de celui dans lequel on effectue les calculs.

On peut alors poser la définition suivante :

Définition 12.4.4

Soit A_1, A_2, \dots, A_n une famille de n polynômes, avec $n \geq 2$.

Le polynôme $A_1 \wedge A_2 \wedge \dots \wedge A_n$ est appelé le pgcd des polynômes A_1, A_2, \dots, A_n .

Le polynôme $A_1 \vee A_2 \vee \dots \vee A_n$ est appelé le ppcm des polynômes A_1, A_2, \dots, A_n .

Caractérisation du pgcd et du ppcm

$D = A_1 \wedge A_2 \wedge \dots \wedge A_n$ est l'unique polynôme unitaire tel que $\mathcal{D}(A_1) \cap \mathcal{D}(A_2) \cap \dots \cap \mathcal{D}(A_n) = \mathcal{D}(D)$. Autrement dit, les diviseurs communs aux polynômes A_1, A_2, \dots, A_n sont les diviseurs de leur pgcd.

$M = A_1 \vee A_2 \vee \dots \vee A_n$ est l'unique polynôme unitaire tel que $A_1\mathbb{Z} \cap A_2\mathbb{Z} \cap \dots \cap A_n\mathbb{Z} = M\mathbb{Z}$.

Autrement dit, les multiples communs aux polynômes A_1, A_2, \dots, A_n sont les multiples de leur ppcm.

Proposition 12.4.18 (factorisation dans un calcul de pgcd/ppcm)

Soit A_1, \dots, A_n des polynômes. Soit Q un polynôme non nul.

On les égalités : $\begin{cases} (QA_1) \wedge (QA_2) \wedge \dots \wedge (QA_n) = Q^*(A_1 \wedge A_2 \wedge \dots \wedge A_n) \\ (QA_1) \vee (QA_2) \vee \dots \vee (QA_n) = Q^*(A_1 \vee A_2 \vee \dots \vee A_n) \end{cases}$

Proposition 12.4.19 (relation de Bézout)

Soit A_1, A_2, \dots, A_n une famille de n polynômes, avec $n \geq 2$. Soit D leur pgcd.

Alors il existe des polynômes U_1, U_2, \dots, U_n tels que : $A_1U_1 + A_2U_2 + \dots + A_nU_n = D$.

12.4.7 Polynômes premiers entre eux dans leur ensemble**Définition 12.4.5** (polynômes premiers entre eux dans leur ensemble)

Soit A_1, A_2, \dots, A_n une famille de n polynômes, avec $n \geq 2$.

On dit que ces n polynômes sont *premiers entre eux dans leur ensemble* si leur pgcd est égal à 1.

Cela équivaut à dire que leurs seuls diviseurs communs sont les polynômes constants non nuls.

Proposition 12.4.20 (caractérisation par une identité de Bézout)

Soit A_1, A_2, \dots, A_n une famille de n polynômes, avec $n \geq 2$.

Les propositions suivantes sont équivalentes :

- les polynômes A_1, A_2, \dots, A_n sont premiers entre eux dans leur ensemble
- il existe n polynômes U_1, U_2, \dots, U_n tels que : $A_1U_1 + A_2U_2 + \dots + A_nU_n = 1$

Premiers entre eux : « deux à deux » ou « dans leur ensemble » ?

On ne confondra pas « premiers entre eux deux à deux » et « premiers entre eux dans leur ensemble ».

Si deux au moins des polynômes A_1, \dots, A_n sont premiers entre eux, et a fortiori si les polynômes A_1, \dots, A_n sont premiers entre eux *deux à deux*, alors ils le sont *dans leur ensemble*.

Mais la réciproque est fautive, comme on le voit avec $A = X^2 - X$, $B = X^2 + X$ et $C = X^2 - 1$.

Les polynômes A, B, C sont en effet premiers entre eux dans leur ensemble (on a $(A + B)/2 - C = 1$).

Pourtant $A \wedge B = X(X - 1) \wedge X(X + 1) = X$.

De même : $A \wedge C = X(X - 1) \wedge (X + 1)(X - 1) = X - 1$, et $B \wedge C = X(X + 1) \wedge (X + 1)(X - 1) = X + 1$.

Remarques

Attention : l'égalité $(A \wedge B)(A \vee B) = AB$ ne se généralise pas à plus de deux polynômes.

On reprend l'exemple de $A = X^2 - X$, $B = X^2 + X$ et $C = X^2 - 1$.

Avec ces trois polynômes, on a $A \wedge B \wedge C = 1$ et $A \vee B \vee C = X^3 - X$, mais ABC est de degré 6.

En revanche, si A_1, \dots, A_n sont premiers entre eux deux à deux, alors $A_1 \vee A_2 \vee \dots \vee A_n = A_1 A_2 \dots A_n$.

12.5 Polynômes irréductibles et factorisations**12.5.1 Polynômes irréductibles dans $\mathbb{K}[X]$**

On rappelle qu'un polynôme P quelconque de $\mathbb{K}[X]$ est divisible par tous les polynômes constants non nuls, et que si P est non nul il est divisible par tous ses associés, c'est-à-dire les λP avec λ dans \mathbb{K}^* .

Définition 12.5.1 (polynômes irréductibles)

Soit P un polynôme non constant de $\mathbb{K}[X]$.

On dit que P est irréductible dans $\mathbb{K}[X]$ si ses seuls diviseurs dans $\mathbb{K}[X]$ sont :

- les polynômes constants non nuls.
- les polynômes associés à P , c'est-à-dire les λP , avec λ dans \mathbb{K}^* .

A contrario, un polynôme non constant P n'est pas irréductible (donc est « réductible ») dans $\mathbb{K}[X]$ s'il possède une « factorisation stricte », c'est-à-dire s'il existe A et B dans $\mathbb{K}[X]$ tels que $P = AB$ avec $\deg(A) < \deg(P)$ et $\deg(B) < \deg(P)$.

La description des polynômes irréductibles est facile dans $\mathbb{C}[X]$, assez facile dans $\mathbb{R}[X]$, mais difficile dans $\mathbb{Q}[X]$ (polynômes à coefficients rationnels). On se contente ici d'un aperçu très sommaire du cas général $\mathbb{K}[X]$, sachant que le programme de la classe de MPSI se limite à $\mathbb{R}[X]$ et $\mathbb{C}[X]$.

On a vu à quel point l'arithmétique des polynômes est proche de celle des entiers. Cette proximité s'explique par l'existence d'une division euclidienne, à la fois dans \mathbb{Z} et dans $\mathbb{K}[X]$ (et c'est cette division qui conduit, via l'algorithme d'Euclide, à la notion de Pgcd). La notion de polynôme irréductible est dans la continuité de cette proximité, les polynômes irréductibles jouant le rôle qu'ont joué les nombres premiers dans l'arithmétique des entiers.

Remarques et propriétés

- Tout polynôme de degré 1 est irréductible.

Si $\deg(P) \geq 2$ et si P admet une racine dans \mathbb{K} , alors P n'est *pas* irréductible dans $\mathbb{K}[X]$.

Si $\deg(P) \in \{2, 3\}$ et si P n'a pas de racine dans \mathbb{K} alors il *est* irréductible dans $\mathbb{K}[X]$.

Cette propriété cesse d'être vraie si $\deg(P) \geq 4$ (exemple : $P = (X^2 + 1)^2$ dans $\mathbb{R}[X]$).

- La notion de polynôme irréductible dépend du corps \mathbb{K} .

Ainsi $A = X^2 - 2$ est irréductible dans $\mathbb{Q}[X]$, mais pas dans $\mathbb{R}[X]$ car $A = (X - \sqrt{2})(X + \sqrt{2})$.

De même $A = X^2 + 1$ est irréductible dans $\mathbb{R}[X]$, mais pas dans $\mathbb{C}[X]$ car $A = (X - i)(X + i)$.

- Si un polynôme irréductible P ne divise pas un polynôme A , alors il est premier avec A .
En particulier, P est premier avec les polynômes de degré strictement inférieur à $\deg(P)$.
- Soit P un polynôme irréductible, et soit A_1, A_2, \dots, A_n une famille de polynômes.
Alors P divise le produit $A_1 A_2 \dots A_n$ si et seulement si P divise l'un au moins des A_k .
- Si un polynôme P est irréductible, ses associés le sont aussi.
Si deux polynômes irréductibles ne sont pas associés, ils sont premiers entre eux.
Deux polynômes irréductibles unitaires distincts sont premiers entre eux.

Proposition 12.5.1 (existence d'un diviseur irréductible)

Tout polynôme non constant est divisible par au moins un polynôme irréductible.

Proposition 12.5.2 (décomposition en produit de polynômes irréductibles)

Tout polynôme A non constant s'écrit $A = \lambda P_1^{\alpha_1} P_2^{\alpha_2} \dots P_m^{\alpha_m}$, où :

- m est un entier strictement positif.
- P_1, P_2, \dots, P_m sont des polynômes irréductibles unitaires distincts deux à deux.
- $\alpha_1, \alpha_2, \dots, \alpha_m$ sont des entiers strictement positifs.

Une telle écriture de A est unique à l'ordre près des facteurs.

On l'appelle décomposition de A en produit de facteurs irréductibles.

Dans cette écriture de A , les P_k sont les diviseurs irréductibles unitaires du polynôme A .

12.5.2 Décomposition en facteurs irréductibles dans $\mathbb{C}[X]$

On rappelle le résultat fondamental suivant (dont la démonstration est admise!)

Proposition 12.5.3 (théorème de d'Alembert-Gauss)

Tout polynôme de $\mathbb{C}[X]$, de degré $n \geq 1$, admet au moins une racine dans \mathbb{C} .

On exprime cette propriété en disant que le corps \mathbb{C} est algébriquement clos.

Le théorème de d'Alembert-Gauss a une conséquence immédiate :

Proposition 12.5.4

Les polynômes irréductibles dans $\mathbb{C}[X]$ sont les polynômes de degré 1.

Ce qui conduit à la forme générale de la factorisation en polynômes irréductibles dans $\mathbb{C}[X]$:

Proposition 12.5.5 (factorisation dans $\mathbb{C}[X]$)

Soit A un polynôme non constant de $\mathbb{C}[X]$.

Alors A s'écrit de manière unique (à l'ordre près des facteurs) $A = \lambda \prod_{k=1}^p (X - \alpha_k)^{m_k}$ où :

- le scalaire non nul λ est le coefficient dominant de A .
- les scalaires $\alpha_1, \alpha_2, \dots, \alpha_p$ sont les racines distinctes de A dans \mathbb{C} .
- les entiers strictement positifs m_1, m_2, \dots, m_p sont leurs multiplicités respectives.

Un exemple important

Les racines de $X^n - 1$ dans \mathbb{C} sont les racines n -ièmes de l'unité $\omega_0, \omega_1, \dots, \omega_{n-1}$.

Ce sont toutes des racines de multiplicité 1.

La factorisation dans $\mathbb{C}[X]$ est donc $X^n - 1 = \prod_{k=0}^{n-1} (X - \omega_k) = \prod_{k=0}^{n-1} \left(X - \exp \frac{2ik\pi}{n} \right)$

Avec les relations coefficients-racines, on trouve immédiatement $\begin{cases} \omega_0 + \omega_1 + \dots + \omega_{n-1} = 0 \\ \omega_0 \omega_1 \dots \omega_{n-1} = (-1)^{n-1} \end{cases}$

Par exemple, si $n = 3$: $X^3 - 1 = (X - 1)(X - j)(X - j^2)$.

Autre exemple, si $n = 6$: $X^6 - 1 = (X - 1)(X + 1)(X - j)(X - j^2)(X + j)(X + j^2)$.

Dans le cas particulier où A est un polynôme à coefficients réels, et si on veut le factoriser dans $\mathbb{C}[X]$, la décomposition fait apparaître les (éventuelles) racines réelles d'une part, et les (éventuelles) racines non réelles d'autre part. On sait que si α est une racine non réelle de A , avec la multiplicité $m \geq 1$, alors $\bar{\alpha}$ est également une racine de A , avec la même multiplicité :

Proposition 12.5.6 (factorisation dans $\mathbb{C}[X]$ d'un polynôme à coefficients réels)

–
–
–

On peut proposer une autre forme de la décomposition en facteurs irréductibles de A dans $\mathbb{C}[X]$, en y faisant figurer tous les polynômes irréductibles unitaires $X - \alpha$, mais avec des exposants positifs ou nuls (exposants strictement positifs uniquement pour les racines effectives du polynôme A) :

Proposition 12.5.7 (autre forme de la factorisation dans $\mathbb{C}[X]$)

Soit A un polynôme non nul de $\mathbb{C}[X]$.

Alors A s'écrit de manière unique (à l'ordre près des facteurs) $A = \lambda \prod_{\alpha \in \mathbb{C}} (X - \alpha)^{m_\alpha}$ où :

- le scalaire λ est le coefficient dominant de A .
- l'entier $m_\alpha \geq 0$ est la multiplicité de α comme racine de A .

La forme précédente permet de caractériser la divisibilité dans $\mathbb{C}[X]$.

Proposition 12.5.8

Soit A et B deux polynômes non nuls de $\mathbb{C}[X]$.

Soit $A = \lambda \prod_{\alpha \in \mathbb{C}} (X - \alpha)^{m_\alpha}$ et $B = \mu \prod_{\alpha \in \mathbb{C}} (X - \alpha)^{n_\alpha}$ leurs factorisations dans $\mathbb{C}[X]$.

Alors A divise B si et seulement si, pour tout α de \mathbb{C} , on a $m_\alpha \leq n_\alpha$.

On en déduit une forme du pgcd et du ppcm, à partir des factorisations :

Proposition 12.5.9 (expression factorisée du pgcd et du ppcm)

Soit A et B deux polynômes non nuls de $\mathbb{C}[X]$.

Soit $A = \lambda \prod_{\alpha \in \mathbb{C}} (X - \alpha)^{m_\alpha}$ et $B = \mu \prod_{\alpha \in \mathbb{C}} (X - \alpha)^{n_\alpha}$ leurs factorisations dans $\mathbb{C}[X]$.

Alors $A \wedge B = \prod_{\alpha \in \mathbb{C}} (X - \alpha)^{p_\alpha}$ et $A \vee B = \prod_{\alpha \in \mathbb{C}} (X - \alpha)^{q_\alpha}$ avec pour tout α de \mathbb{C} :
$$\begin{cases} p_\alpha = \min(m_\alpha, n_\alpha) \\ q_\alpha = \max(m_\alpha, n_\alpha) \end{cases}$$

12.5.3 Décomposition en facteurs irréductibles dans $\mathbb{R}[X]$

Proposition 12.5.10

Les polynômes irréductibles dans $\mathbb{R}[X]$ sont :

- les polynômes de degré 1.
- les polynômes $P = aX^2 + bX + c$ de degré 2 et de discriminant $\Delta = b^2 - 4ac < 0$.

Proposition 12.5.11 (factorisation dans $\mathbb{R}[X]$)

Soit A un polynôme non constant de $\mathbb{R}[X]$.

Alors A s'écrit de façon unique (à l'ordre près) : $A = \lambda \prod_{k=1}^p (X - \alpha_k)^{r_k} \prod_{k=1}^q (X^2 + b_k X + c_k)^{s_k}$ où :

– Le réel non nul λ est le coefficient dominant de A .

– Les réels $\alpha_1, \alpha_2, \dots, \alpha_p$ sont les racines distinctes de A dans \mathbb{R} .

Les entiers strictement positifs r_1, r_2, \dots, r_p sont leurs multiplicités respectives.

– Les polynômes $X^2 + b_k X + c_k$, distincts deux à deux, sont à discriminant négatif.

Dans $\mathbb{C}[X]$, chaque $X^2 + b_k X + c_k$ se factoriserait en $(X - \beta_k)(X - \overline{\beta_k})$,

où β_k est l'une des racines complexes non réelles de A .

Les entiers strictement positifs s_1, \dots, s_q sont les multiplicités respectives de β_1, \dots, β_q dans $\mathbb{C}[X]$.

Dans l'écriture précédente de A , il est possible que q soit nul (si A est scindé dans \mathbb{R}), ou que p soit nul (si A ne possède que des racines complexes non réelles).

Dans ce cas, le produit “vide” $\prod_{k=1}^p \dots$ ou $\prod_{k=1}^q \dots$ prend la valeur 1.

Pour factoriser dans $\mathbb{R}[X]$ un polynôme A ayant des racines non réelles, on peut le factoriser dans $\mathbb{C}[X]$ puis regrouper les facteurs correspondant à des racines complexes non réelles.

Polynômes bicarrés

Il arrive souvent qu'on ait à factoriser dans $\mathbb{R}[X]$ des polynômes “bicarrés” à discriminant négatif, c'est-à-dire du type $A = X^4 + bX^2 + c$, avec $\Delta = b^2 - 4c < 0$.

Dans ce cas, on considère les termes X^4 et c comme provenant du développement d'un carré.

Plus précisément, on écrit : $X^4 + bX^2 + c = (X^2 + \sqrt{c})^2 - (2\sqrt{c} - b)X^2$.

Or le coefficient $2\sqrt{c} - b$ est strictement positif. On peut donc l'écrire comme un carré β^2 .

On arrive alors à $X^4 + bX^2 + c = (X^2 + \sqrt{c})^2 - \beta^2 X^2 = (X^2 + \beta X + \sqrt{c})(X^2 - \beta X + \sqrt{c})$.

À titre d'exemple, voici la factorisation dans $\mathbb{R}[X]$ du polynôme « quadricarré » $A = X^8 + X^4 + 1$.

On a tout d'abord $A = (X^4 + 1)^2 - X^4 = (X^4 - X^2 + 1)(X^4 + X^2 + 1)$.

Ensuite
$$\begin{cases} X^4 - X^2 + 1 = (X^2 + 1)^2 - 3X^2 = (X^2 + \sqrt{3}X + 1)(X^2 - \sqrt{3}X + 1) \\ X^4 + X^2 + 1 = (X^2 + 1)^2 - X^2 = (X^2 + X + 1)(X^2 - X + 1) \end{cases}$$

Conclusion : $X^8 + X^4 + 1 = (X^2 + X + 1)(X^2 - X + 1)(X^2 + \sqrt{3}X + 1)(X^2 - \sqrt{3}X + 1)$

12.6 Interpolation de Lagrange

Dans cette section, comme dans le reste du chapitre, \mathbb{K} désigne \mathbb{R} ou \mathbb{C} .

On considère n scalaires distincts $x_1, x_2, \dots, x_{n-1}, x_n$ dans \mathbb{K} .

On se donne également y_1, y_2, \dots, y_n dans \mathbb{K} , mais pas nécessairement distincts deux à deux.

Le problème est : existe-t-il des polynômes P (et lesquels ?) tels que : $\forall k \in \{1, \dots, n\}, P(x_k) = y_k$.

En d'autres termes, et si $\mathbb{K} = \mathbb{R}$, peut-on faire « passer » la courbe représentative d'un polynôme P par n points $A_k(x_k, y_k)$ donnés (dont les abscisses x_k sont deux à deux distinctes). Et quel est le degré minimum d'un tel polynôme ?

Si $f : I \rightarrow \mathbb{R}$ est une fonction dont seules sont connues les valeurs $y_k = f(x_k)$ en n abscisses x_1, x_2, \dots, x_n de I , la question devient alors : trouver un polynôme P tel que $P(x_k) = f(x_k)$ pour $1 \leq k \leq n$.

Une fois ce polynôme connu, et de degré minimum si possible, on peut « interpoler » la fonction f entre deux abscisses successives x_k et x_{k+1} en posant $f(x) = P(x)$ sur $[x_k, x_{k+1}]$. La question est alors de mesurer la qualité de cette approximation, et de savoir si le choix des abscisses x_k est important.

Il y a plusieurs approches possibles pour former le « polynôme d'interpolation » P , et notamment la méthode de Lagrange. L'idée consiste à former d'abord, pour une famille d'abscisses $x_1, x_2, \dots, x_{n-1}, x_n$ donnée, des polynômes prenant des valeurs simples sur les x_k .

Définition 12.6.1 (polynômes interpolateurs de Lagrange)

On se donne une famille de n scalaires $x_1, x_2, \dots, x_{n-1}, x_n$, distincts deux à deux.

Pour tout entier k de $\{1, \dots, n\}$ on définit le polynôme $L_k(X) = \prod_{\substack{j=1 \\ j \neq k}}^n \frac{X - x_j}{x_k - x_j}$.

Proposition 12.6.1

Avec les notations précédentes, et pour tout k de $\{1, \dots, n\}$: $\deg(L_k) = n - 1$ et $\begin{cases} L_k(x_k) = 1 \\ \forall j \neq k, L_k(x_j) = 0 \end{cases}$

Par exemple, si $n = 4$:

$$\begin{aligned} L_1(X) &= \frac{(X - x_2)(X - x_3)(X - x_4)}{(x_1 - x_2)(x_1 - x_3)(x_1 - x_4)} & L_2(X) &= \frac{(X - x_1)(X - x_3)(X - x_4)}{(x_2 - x_1)(x_2 - x_3)(x_2 - x_4)} \\ L_3(X) &= \frac{(X - x_1)(X - x_2)(X - x_4)}{(x_3 - x_1)(x_3 - x_2)(x_3 - x_4)} & L_4(X) &= \frac{(X - x_1)(X - x_2)(X - x_3)}{(x_4 - x_1)(x_4 - x_2)(x_4 - x_3)} \end{aligned}$$

On voit que L_1, L_2, L_3, L_4 sont des polynômes de degré 3, et qu'ils vérifient :

$$\begin{aligned} L_1(x_2) = L_1(x_3) = L_1(x_4) = 0 \text{ et } L_1(x_1) = 1 & \quad L_2(x_1) = L_2(x_3) = L_2(x_4) = 0 \text{ et } L_2(x_2) = 1 \\ L_3(x_1) = L_3(x_2) = L_3(x_4) = 0 \text{ et } L_3(x_3) = 1 & \quad L_4(x_1) = L_4(x_2) = L_4(x_3) = 0 \text{ et } L_4(x_4) = 1 \end{aligned}$$

Par un principe de « superposition », et dans le cas de n ordonnées y_1, \dots, y_n quelconques, on peut former un polynôme P de degré $n - 1$ tel que $P(x_k) = y_k$ pour tout k de $\{1, \dots, n\}$:

Proposition 12.6.2 (existence et unicité du polynôme d'interpolation de degré minimum)

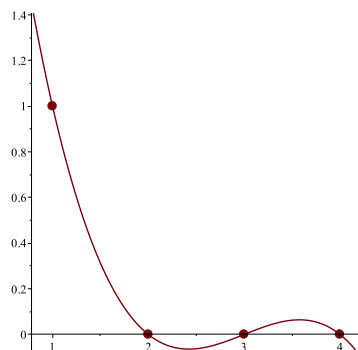
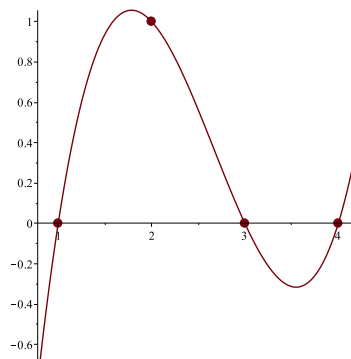
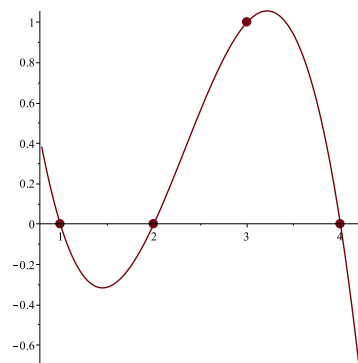
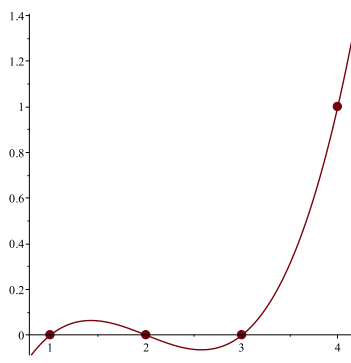
On se donne une famille de n éléments (x_k, y_k) de \mathbb{K}^2 , les x_k étant distincts deux à deux.

Il existe un unique polynôme P de $\mathbb{K}[X]$, avec $\deg(P) \leq n - 1$, tel que : $\forall k \in \{1, \dots, n\}, P(x_k) = y_k$.

Ce polynôme s'écrit $P(X) = \sum_{k=1}^n y_k L_k(X)$, où on rappelle que : $\forall k \in \{1, \dots, n\}, L_k(X) = \prod_{\substack{j=1 \\ j \neq k}}^n \frac{X - x_j}{x_k - x_j}$.

Pour prendre un exemple, on suppose $n = 4$ et on pose $x_1 = 1, x_2 = 2, x_3 = 3, x_4 = 4$.

On représente ici les quatre polynômes L_1, L_2, L_3, L_4 qui correspondent à ces abscisses : chaque L_k est de degré 3, il vaut 1 en x_k et s'annule sur les trois autres abscisses.

Le polynôme L_1 Le polynôme L_2 Le polynôme L_3 Le polynôme L_4

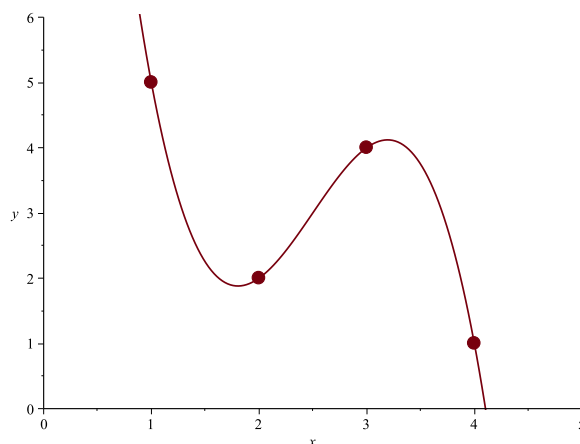
Pour former un polynôme de degré inférieur ou égal à 3 et qui prend les valeurs y_1, y_2, y_3, y_4 sur les abscisses x_1, x_2, x_3, x_4 , il suffit de poser :

$$P(X) = y_1 L_1(X) + y_2 L_2(X) + y_3 L_3(X) + y_4 L_4(X)$$

En reprenant l'exemple précédent, on trace ici la courbe représentative du polynôme P de degré inférieur ou égal à 3 et qui vérifie les quatre conditions :

$$P(1) = 5, P(2) = 2, P(3) = 4, P(4) = 1.$$

$$\text{On trouverait : } P(X) = -\frac{5}{3}X^3 + \frac{25}{2}X^2 - \frac{173}{6}X + 23$$



Proposition 12.6.3 (expression de tous les polynômes vérifiant les égalités $P(x_k) = y_k$)

On reprend ici les notations de la proposition 12.6.2.

Les polynômes P de $\mathbb{K}[X]$ qui vérifient $\forall k \in \{1, \dots, n\}, P(x_k) = y_k$ sont donnés par :

$$P(X) = P_I(X) + Q(X) \prod_{k=1}^n (X - x_k), \text{ où : } \begin{cases} P_I \text{ est le « polynôme interpolateur » de la proposition 12.6.2} \\ Q(X) \text{ est un polynôme quelconque de } \mathbb{K}[X] \end{cases}$$

12.7 Fractions rationnelles

Comme d'habitude, on pose $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} .

12.7.1 La construction du corps $\mathbb{K}(X)$

Note : dans le programme de la classe de MPSI, cette construction n'est pas exigible.

Définition 12.7.1 (une relation d'équivalence sur $\mathbb{K}[X] \times (\mathbb{K}[X] \setminus \{0\})$)

On définit une relation sur $\mathbb{K}[X] \times (\mathbb{K}[X] \setminus \{0\})$ de la façon suivante :

Pour tous couples de polynômes (A, B) et (C, D) , avec $B \neq 0$ et $D \neq 0$: $(A, B) \mathcal{R} (C, D) \Leftrightarrow AD = BC$.

Proposition 12.7.1 (définition des fractions rationnelles)

Avec les définitions précédentes, \mathcal{R} est une relation d'équivalence sur $\mathbb{K}[X] \times (\mathbb{K}[X] \setminus \{0\})$.

Les classes d'équivalence pour cette relation sont appelées fractions rationnelles à coefficients dans \mathbb{K} .

On note $\mathbb{K}(X)$ l'ensemble des fractions rationnelles à coefficients dans \mathbb{K} .

La classe d'équivalence R du couple (A, B) , avec $B \neq 0$, est notée $F = \frac{A}{B}$.

Ainsi pour A, B, C, D dans $\mathbb{K}[X]$, avec $B \neq 0$ et $D \neq 0$, on a l'équivalence : $\frac{A}{B} = \frac{C}{D} \Leftrightarrow AD = BC$.

Remarques

On ne confondra pas les notations $\mathbb{K}[X]$ (polynômes) et $\mathbb{K}(X)$ (fractions rationnelles).

Quand on dit « soit $F = \frac{A}{B}$ une fraction rationnelle », on sous-entend « A, B sont dans $\mathbb{K}[X]$ et $B \neq 0$ ».

Dans l'écriture $F = \frac{A}{B}$, on dit bien sûr que A est le *numérateur* et que B est le *dénominateur*.

Forme irréductible d'une fraction rationnelle

Pour tous polynômes A, B, Q , avec $B \neq 0$ et $Q \neq 0$, on a $\frac{AQ}{BQ} = \frac{A}{B}$.

Soit $F = \frac{A}{B}$ un élément de $\mathbb{K}(X)$. Soit $\Delta = A \wedge B$.

Il existe deux polynômes C et D tels que $\begin{cases} A = \Delta C \\ B = \Delta D \end{cases}$ donc tels que $F = \frac{C}{D}$, avec $C \wedge D = 1$.

Sous cette dernière forme, on dit que F est écrite sous forme *irréductible*, ou *simplifiée*.

Réciproquement, si $F = \frac{A}{B} = \frac{C}{D}$, avec $C \wedge D = 1$, alors il existe Q dans $\mathbb{K}[X]$ tel que $\begin{cases} A = QC \\ B = QD \end{cases}$

Si on suppose de plus $A \wedge B = 1$, alors il existe λ dans \mathbb{K}^* tel que $A = \lambda C$ et $B = \lambda D$.

On en déduit que la forme irréductible d'une fraction rationnelle non nulle est unique si on impose au dénominateur d'être un polynôme unitaire.

Par exemple, la forme irréductible de $F = \frac{X^6 - 1}{X^4 - 1} = \frac{(X^2 - 1)(X^4 + X^2 + 1)}{(X^2 - 1)(X^2 + 1)}$ est $F = \frac{X^4 + X^2 + 1}{X^2 + 1}$

Les polynômes sont des fractions rationnelles particulières

Dans toute la suite, on identifie la fraction rationnelle $F = \frac{A}{1}$ avec le polynôme A .

Cette identification permet de considérer les polynômes comme des fractions rationnelles particulières.

Le polynôme nul notamment s'identifie à $F = \frac{0}{1}$ (avec d'ailleurs $F = \frac{0}{B}$ pour tout $B \neq 0$).

Dire qu'une fraction rationnelle $F = \frac{A}{B}$ est non nulle, c'est dire que son numérateur A est non nul.

Opérations sur les fractions rationnelles

Proposition 12.7.2 (corps des fractions rationnelles à coefficients dans \mathbb{K})

Soit $F = \frac{A}{B}$ et $G = \frac{C}{D}$ dans $\mathbb{K}(X)$. On pose $\frac{A}{B} + \frac{C}{D} = \frac{AD + BC}{BD}$, et $\frac{A}{B} \frac{C}{D} = \frac{AC}{BD}$.

Ces définitions ne dépendent pas des couples (A, B) et (C, D) utilisés pour représenter F et G .

Munis de ces deux opérations, l'ensemble $\mathbb{K}(X)$ possède une structure de corps.

Le neutre dans $\mathbb{K}(X)$ pour la loi $+$ est la fraction rationnelle (le polynôme) $F = 0$.

L'opposé de la fraction rationnelle $F = \frac{A}{B}$ est la fraction rationnelle $-F = \frac{-A}{B}$.

Le neutre dans $\mathbb{K}(X)$ pour la loi \times est la fraction rationnelle (le polynôme) $F = 1$.

Si $F = \frac{A}{B} \neq 0$ (c'est-à-dire $A \neq 0$) l'inverse de $F = \frac{A}{B}$ pour la loi \times est la fraction rationnelle $\frac{1}{F} = \frac{B}{A}$.

Définition 12.7.2 (fonction rationnelle associée à une fraction rationnelle)

Soit $F = \frac{A}{B}$ une fraction rationnelle, supposée irréductible (c'est-à-dire telle que $A \wedge B = 1$).

Soit \tilde{A}, \tilde{B} les fonctions polynomiales associées aux polynômes A et B .

On appelle fonction rationnelle \tilde{F} , associée à fraction rationnelle F , la fonction $x \mapsto \tilde{F}(x) = \frac{\tilde{A}(x)}{\tilde{B}(x)}$.

Son domaine de définition est \mathbb{K} privé de l'ensemble des racines de B .

L'application qui à une fraction rationnelle F associe la fonction \tilde{F} est injective. Si les fonctions \tilde{F} et \tilde{G} sont égales en un nombre infini de points, alors les fractions rationnelles F et G sont égales. On peut donc sans danger identifier une fraction rationnelle et sa fonction rationnelle associée.

Remarque : soit F dans $\mathbb{R}[X]$, et soit α dans \mathbb{C} . Alors $\tilde{F}(\bar{\alpha})$ est le conjugué de $\tilde{F}(\alpha)$.

Parité ou imparité d'une fraction rationnelle

Si une fraction rationnelle est paire (invariante par $X \rightarrow -X$), on peut l'écrire $F(X) = \frac{A(X^2)}{B(X^2)}$.

Si elle est impaire, on peut l'écrire $F(X) = X \frac{A(X^2)}{B(X^2)}$.

Par exemple $G = \frac{X^4 - 1}{X^6 + 3X^2 + 1}$ est paire : $G = \frac{A(X^2)}{B(X^2)}$ avec $A = X^2 - 1$ et $B = X^3 + 3X + 1$.

De même, $G = \frac{X^4 + 1}{X(X^2 + 1)}$ est impaire : $G = X \frac{A(X^2)}{B(X^2)}$ avec $A = X^2 + 1$ et $B = X(X + 1)$.

12.7.2 Degré, partie entière

Définition 12.7.3 (degré d'une fraction rationnelle)

Soit $F = \frac{A}{B}$ une fraction rationnelle. On dit que $\deg(F) = \deg(A) - \deg(B)$ est le *degré* de F .

Cette définition ne dépend pas du couple (A, B) utilisé pour représenter F .

Remarques

Le degré d'une fraction rationnelle non nulle est un entier relatif, et on a encore $\deg(0) = -\infty$.

Si la fraction F est un polynôme, son degré est le même, qu'on se place dans $\mathbb{K}(X)$ ou dans $\mathbb{K}[X]$.

Comme avec les polynômes, on a les relations :
$$\begin{cases} \deg(F + G) \leq \max(\deg(F), \deg(G)) \\ \deg(FG) = \deg(F) + \deg(G) \end{cases}$$

Proposition 12.7.3 (partie entière d'une fraction rationnelle)

Soit $F = \frac{A}{B}$ une fraction rationnelle à coefficients dans \mathbb{K} .

Alors F s'écrit de manière unique $F = E_F + G$, avec E_F dans $\mathbb{K}[X]$, G dans $\mathbb{K}(X)$, et $\deg(G) < 0$.

On dit que le polynôme E_F est la partie entière de la fraction rationnelle F .

En fait, le polynôme E_F n'est autre que le quotient dans la division euclidienne de A par B .

En effet, si cette division est $A = BQ + R$, on trouve : $F = \frac{BQ + R}{B} = Q + \frac{R}{B}$ avec $\deg(R) < \deg(B)$.

Remarques et propriétés

On a $E_F = 0$ si et seulement si $\deg(A) < \deg(B)$. Sinon $\deg(E_F) = \deg(A) - \deg(B)$.

Si $\deg(A) = \deg(B)$, alors E_F est le quotient des coefficients dominants de A et B .

Soit F, G dans $\mathbb{K}(X)$ et λ, μ dans \mathbb{K} . Alors $E_{\lambda F + \mu G} = \lambda E_F + \mu E_G$.

Si F est paire (resp. impaire), alors E_F est un polynôme pair (resp. impair).

12.7.3 Zéros et pôles, multiplicités

Définition 12.7.4 (pôles d'une fraction rationnelle)

Soit $F = \frac{A}{B}$ une fraction rationnelle écrite sous forme irréductible.

Les deux polynômes A et B , premiers entre eux, n'ont donc pas de racine commune dans \mathbb{K} .

Soit α un élément de \mathbb{K} , et m un élément de \mathbb{N}^* .

On dit que α est un zéro de F , avec la multiplicité m , si α est un zéro de A , avec la multiplicité m .

On dit que α est un pôle de F , avec la multiplicité m , si α est un zéro de B , avec la multiplicité m .

Remarques :

– Dire que α est un zéro de F avec la multiplicité m , c'est dire que $F = \frac{(X - \alpha)^m Q}{B}$, avec $\begin{cases} Q(\alpha) \neq 0 \\ B(\alpha) \neq 0 \end{cases}$

De même, α est un pôle de F de multiplicité $m \Leftrightarrow F = \frac{A}{(X - \alpha)^m Q}$, avec $\begin{cases} A(\alpha) \neq 0 \\ Q(\alpha) \neq 0 \end{cases}$

– On parle de pôle simple si $m = 1$, et de pôle multiple si $m > 1$.

On parle de pôle double si $m = 2$, triple si $m = 3$, etc.

– Dire que α est pôle simple de $F = \frac{A}{B}$, c'est dire que $A(\alpha) \neq 0$, $B(\alpha) = 0$, et $B'(\alpha) \neq 0$.

De même, α est un pôle double de $F = \frac{A}{B}$ si et seulement si $\begin{cases} A(\alpha) \neq 0 \\ B(\alpha) = B'(\alpha) = 0, B''(\alpha) \neq 0 \end{cases}$

– Soit F une fraction rationnelle à coefficients réels.

Soit α un pôle complexe non réel de F de multiplicité m .

Alors $\bar{\alpha}$ est un pôle de F de multiplicité m .

12.8 Décomposition en éléments simples

12.8.1 Décomposition en éléments simples dans $\mathbb{C}(X)$

Proposition 12.8.1 (forme de la décomposition en éléments simples dans $\mathbb{C}(X)$)

Soit $F = \frac{A}{B}$ une fraction rationnelle à coefficients complexes, écrite sous forme irréductible.

Soit $\alpha_1, \dots, \alpha_p$ les pôles distincts de F , avec les multiplicités respectives r_1, \dots, r_p .

Alors F s'écrit de manière unique $F = E_F + \sum_{k=1}^p \left(\sum_{j=1}^{r_k} \frac{\lambda_{k,j}}{(X - \alpha_k)^j} \right)$

où E_F est la partie entière de F et où les $\lambda_{k,j}$ sont des éléments de \mathbb{C} .

Cette écriture est appelée décomposition en éléments simples de F dans $\mathbb{C}(X)$.

Par exemple, considérons $F = \frac{X^{12} + 1}{X^4(X - i)^2(X + 1)(X - j)^3}$, dont le degré est : $12 - (4 + 2 + 1 + 3) = 2$.

Les racines du dénominateur sont $0, i, -1, j$ et aucune d'elles n'est racine du numérateur.

La fraction rationnelle est donc écrite sous forme irréductible.

Ses pôles sont 0 (quadruple), i (double), -1 (simple) et j (triple).

Sa décomposition en éléments simples dans $\mathbb{C}(X)$ s'écrit donc sous la forme :

$$\begin{aligned} F(X) &= aX^2 + bX + c && \text{(partie entière de } F, \text{ ici un polynôme de degré 2)} \\ &+ \frac{\alpha_4}{X^4} + \frac{\alpha_3}{X^3} + \frac{\alpha_2}{X^2} + \frac{\alpha_1}{X} && \text{(« partie polaire » relative au pôle quadruple 0)} \\ &+ \frac{\beta_2}{(X - i)^2} + \frac{\beta_1}{X - i} && \text{(partie polaire relative au pôle double } i) \\ &+ \frac{\gamma_1}{X + 1} && \text{(partie polaire relative au pôle simple } -1) \\ &+ \frac{\delta_3}{(X - j)^3} + \frac{\delta_2}{(X - j)^2} + \frac{\delta_1}{X - j} && \text{(partie polaire relative au pôle triple } j) \end{aligned}$$

Le calcul de tous ces coefficients demande une technicité qui n'est pas dans l'esprit du programme.

On se contentera donc, dans la suite de cette section, de situations assez élémentaires.

Décomposition dans $\mathbb{R}(X)$ quand le dénominateur est scindé

Ce qui précède s'applique aux fractions rationnelles réelles dont le dénominateur est scindé dans $\mathbb{R}[X]$. Dans ce cas, les coefficients qui apparaissent dans la décomposition sont tous des nombres réels.

Par exemple, considérons la fraction rationnelle irréductible $F(X) = \frac{X^3 + X + 1}{X^5(X-1)^2}$.

Le degré de F est -4 : il est strictement négatif donc la partie entière E_F est nulle.

Ses pôles sont 0 (multiplicité 5), et 1 (multiplicité 2).

Sa décomposition en éléments simples dans $\mathbb{R}(X)$ s'écrit donc sous la forme :

$$F(X) = \frac{\alpha_5}{X^5} + \frac{\alpha_4}{X^4} + \frac{\alpha_3}{X^3} + \frac{\alpha_2}{X^2} + \frac{\alpha_1}{X} + \frac{\beta_2}{(X-1)^2} + \frac{\beta_1}{X-1} \quad (\text{où les } \alpha_k \text{ et } \beta_k \text{ sont réels})$$

12.8.2 Décomposition en éléments simples dans $\mathbb{R}(X)$

Proposition 12.8.2 (forme de la décomposition en éléments simples dans $\mathbb{R}(X)$)

Soit $F = \frac{A}{B}$ une fraction rationnelle à coefficients réels, écrite sous forme irréductible.

On note $B = \lambda \prod_{k=1}^p (X - \alpha_k)^{r_k} \prod_{k=1}^q (X^2 + b_k X + c_k)^{s_k}$ la factorisation de B dans $\mathbb{R}[X]$.

Alors F s'écrit de manière unique : $F = E_F + \sum_{k=1}^p \left(\sum_{j=1}^{r_k} \frac{\lambda_{k,j}}{(X - \alpha_k)^j} \right) + \sum_{k=1}^q \left(\sum_{j=1}^{s_k} \frac{c_{k,j} X + d_{k,j}}{(X^2 + b_k X + c_k)^j} \right)$

où E_F est la partie entière de F et où les $\lambda_{k,j}$, $c_{k,j}$, $d_{k,j}$ sont des réels.

Cette écriture est appelée décomposition en éléments simples de F dans $\mathbb{R}(X)$.

Dans la décomposition de la fraction rationnelle F dans $\mathbb{R}(X)$:

- les fractions $\frac{\lambda_{k,j}}{(X - \alpha_k)^j}$ sont appelées *éléments simples de première espèce*.
- les fractions $\frac{c_{k,j} X + d_{k,j}}{(X^2 + b_k X + c_k)^j}$ sont appelées *éléments simples de seconde espèce*.

Par exemple, soit la fraction rationnelle irréductible $F(X) = \frac{64 X^{13}}{(X+1)^2(X^2+1)^2(X^2+X+1)(X-1)^3}$.

Le degré de F est $13 - (2 + 4 + 2 + 3) = 2$: il y a donc une partie entière de degré 2.

Ses pôles réels sont -1 (multiplicité 2), et 1 (multiplicité 3).

Les deux facteurs $X^2 + 1$ et $X^2 + X + 1$ sont irréductibles dans $\mathbb{R}[X]$.

La décomposition de F en éléments simples dans $\mathbb{R}(X)$ s'écrit donc sous la forme :

$$\begin{aligned} F(X) &= aX^2 + bX + c && \text{(partie entière de } F, \text{ ici un polynôme de degré 2)} \\ &+ \frac{\alpha_2}{(X+1)^2} + \frac{\alpha_1}{X+1} && \text{(partie polaire relative au pôle double } -1) \\ &+ \frac{\beta_2 X + \gamma_2}{(X^2+1)^2} + \frac{\beta_1 X + \gamma_1}{X^2+1} && \text{(partie relative au facteur } (X^2+1)^2) \\ &+ \frac{\delta_1 X + \lambda_1}{X^2+X+1} && \text{(partie relative au facteur } X^2+X+1) \\ &+ \frac{\mu_3}{(X-1)^3} + \frac{\mu_2}{(X-1)^2} + \frac{\mu_1}{X-1} && \text{(partie polaire relative au pôle réel triple 1)} \end{aligned}$$

On trouve en fait :

$$F(X) = 576X^2 + \frac{108}{(X-1)^2} + \frac{64(X+2)}{X^2+X+1} + \frac{36(7X+11)}{X^2+1} + \frac{18}{(X+1)^2} + \frac{413}{X-1} + \frac{12}{(X-1)^3} - \frac{72(X+1)}{(X^2+1)^2} - \frac{153}{8(X+1)}$$

Autre exemple, considérons la fraction rationnelle $F = \frac{1}{(X^2+1)^4(X^2-X+1)^3}$

Sa décomposition en éléments simples dans $\mathbb{R}(X)$ est :

$$F(X) = -\frac{X}{(X^2+1)^4} - \frac{2X+3}{(X^2+1)^3} + \frac{3X-6}{(X^2+1)^2} + \frac{14X+1}{X^2+1} + \frac{X-1}{(X^2-X+1)^3} - \frac{2X+3}{(X^2-X+1)^2} - \frac{14X+13}{X^2-X+1}$$

12.8.3 Cas d'un pôle simple

Soit $F(X) = \frac{A(X)}{B(X)}$ un élément de $\mathbb{K}(X)$ admettant α comme pôle simple.

On a donc $A(\alpha) \neq 0$, et on peut écrire $F(X)$ sous la forme $F(X) = \frac{A(X)}{(X-\alpha)Q(X)}$ avec $Q(\alpha) \neq 0$.

La décomposition de F s'écrit $F(X) = \frac{\lambda}{X-\alpha} + G(X)$, où α n'est pas un pôle de G .

Après produit par $X-\alpha$, on trouve $\frac{A(X)}{Q(X)} = \lambda + (X-\alpha)G(X)$, donc $\lambda = \frac{A(\alpha)}{Q(\alpha)}$.

Dans la pratique, on multiplie F par $X-\alpha$, et après simplification on substitue α à X .

Cette méthode est très adaptée au cas (fréquent) où B est factorisé.

On remarque également que $B'(X) = (X-\alpha)Q'(X) + Q(X)$, et il en résulte $B'(\alpha) = Q(\alpha)$.

Le coefficient λ peut donc également s'écrire : $\lambda = \frac{A(\alpha)}{B'(\alpha)}$.

Cette dernière méthode est adaptée au cas où B n'est pas factorisé (cf deuxième exemple ci-dessous).

Un premier exemple

On se propose de décomposer $F = \frac{1}{X(X+1)(X+2)\cdots(X+n)}$ dans $\mathbb{R}(X)$.

Les pôles sont $0, -1, -2, \dots, -n$, tous de multiplicité 1.

La forme de la décomposition est donc : $F = \sum_{k=0}^n \frac{\lambda_k}{X+k}$.

Le scalaire λ_k s'obtient en multipliant F par $X+k$ et en substituant $-k$ à X .

On trouve : $\lambda_k = \prod_{j \neq k} \frac{1}{-k+j} = \prod_{j=0}^{k-1} \frac{1}{j-k} \prod_{j=k+1}^n \frac{1}{j-k} = (-1)^k \prod_{i=1}^k \frac{1}{i} \prod_{i=1}^{n-k} \frac{1}{i} = \frac{(-1)^k}{k!(n-k)!} = \frac{(-1)^k}{n!} \binom{n}{k}$.

Conclusion : $F = \frac{1}{X(X+1)(X+2)\cdots(X+n)} = \frac{1}{n!} \sum_{k=0}^n \binom{n}{k} \frac{(-1)^k}{X+k}$.

Par exemple, si $n=3$:

$$\frac{1}{X(X+1)(X+2)(X+3)} = \frac{1}{24X} - \frac{1}{6(X+1)} + \frac{1}{4(X+2)} - \frac{1}{6(X+3)} + \frac{1}{24(X+4)}$$

Un deuxième exemple

On se propose de décomposer $F = \frac{1}{X^n - 1}$ dans $\mathbb{C}(X)$.

Ici $F = \frac{A}{B}$, avec $\begin{cases} A = 1 \\ B = X^n - 1 \end{cases}$ et les pôles (simples) sont les racines n -ièmes ω_k de l'unité.

Pour tout k de $\{0, \dots, n-1\}$, on a $\frac{A(\omega_k)}{B'(\omega_k)} = \frac{1}{n\omega_k^{n-1}} = \frac{\omega_k}{n}$.

La décomposition en éléments simples de F s'écrit donc $\frac{1}{X^n - 1} = \frac{1}{n} \sum_{k=0}^{n-1} \frac{\omega_k}{X - \omega_k}$.

Par exemple : $\frac{1}{X^4 - 1} = \frac{1}{4(X-1)} + \frac{i}{4(X-i)} - \frac{1}{4(X+1)} - \frac{i}{4(X+i)}$.

12.8.4 Décomposition en éléments simples de P'/P

Soit P un polynôme scindé dans $\mathbb{K}[X]$.

On se propose de décomposer $F = \frac{P'}{P}$ dans $\mathbb{K}(X)$.

Soit α une racine de P , avec la multiplicité $m \geq 1$.

On peut donc écrire $P(X) = (X - \alpha)^m Q(X)$, avec $Q(\alpha) \neq 0$.

On en déduit $P'(X) = (X - \alpha)^{m-1}(mQ(X) + (X - \alpha)Q'(X))$, donc $\frac{P'}{P} = \frac{m}{X - \alpha} + \frac{Q'}{Q}$.

Ainsi $\frac{m}{X - \alpha}$ représente la partie polaire de F pour le pôle α .

On peut conclure de la façon suivante :

Proposition 12.8.3 (décomposition en éléments simples de P'/P)

Soit P un polynôme scindé dans $\mathbb{K}[X]$.

Soit $\alpha_1, \alpha_2, \dots, \alpha_r$ les racines distinctes de P dans \mathbb{K} , de multiplicités respectives m_1, \dots, m_r .

Alors la décomposition de $\frac{P'}{P}$ dans $\mathbb{K}(X)$ est : $\frac{P'}{P} = \sum_{k=1}^r \frac{m_k}{X - \alpha_k}$.

Par exemple, si $P = X^4(X-1)^2(X+1)^3(X-3)$, alors $\frac{P'}{P} = \frac{4}{X} + \frac{2}{X-1} + \frac{3}{X+1} + \frac{1}{X-3}$.

Remarques

Soit $P = \lambda \prod_{k=1}^r (X - \alpha_k)^{m_k}$ un polynôme scindé dans $\mathbb{K}[X]$.

– Les pôles de $\frac{P'}{P}$ sont les racines de P , et ce sont tous des pôles simples.

– Le pgcd de P et de P' est : $P \wedge P' = \prod_{k=1}^r (X - \alpha_k)^{m_k - 1}$.

12.8.5 Pratique de la décomposition en éléments simples

Les méthodes précédentes permettent de décomposer en éléments simples dans $\mathbb{R}(X)$ ou dans $\mathbb{C}(X)$, dans les cas usuels (le programme de la classe de MPSI demande expressément de « limiter la technicité des exercices »). On va tout de même décrire ici un certain nombre d'idées permettant de faciliter la recherche de telles décompositions dans des situations plus techniques.

Décomposition dans $\mathbb{C}(X)$ d'une fraction à coefficients réels

Soit $F = \frac{A}{B}$ un élément de $\mathbb{R}(X)$.

On peut considérer F comme un élément de $\mathbb{C}(X)$ et la décomposer en tant que telle.

Cette décomposition dans $\mathbb{C}(X)$ est nécessairement invariante par conjugaison.

Il en résulte par exemple que la partie entière de F est un polynôme réel.

Il en résulte également que les parties polaires sont conjuguées deux à deux.

Plus précisément, si α et $\bar{\alpha}$ sont deux pôles conjugués non réels de F , de multiplicité m , les parties polaires s'écrivent respectivement : $\sum_{k=1}^m \frac{\lambda_k}{(X - \alpha)^k}$ et $\sum_{k=1}^m \frac{\bar{\lambda}_k}{(X - \bar{\alpha})^k}$.

Cette idée permet donc de diminuer de moitié environ le nombre d'inconnues.

Il est également possible d'utiliser la décomposition dans $\mathbb{C}(X)$ d'une fraction rationnelle réelle F pour obtenir sa décomposition dans \mathbb{R} après regroupement des termes conjugués. Cette méthode n'est envisageable que si les pôles non réels sont de multiplicité 1.

Utilisation de la parité ou de l'imparité

Si une fraction rationnelle est paire ou impaire, sa décomposition doit refléter cette propriété.

Si on exprime cette invariance par les transformations $X \mapsto F(-X)$ ou $X \mapsto -F(-X)$, on en déduit des relations sur les coefficients (le nombre d'inconnues diminue environ de moitié).

Il est possible que la fraction rationnelle F soit invariante (ou changée en son opposée) par une autre transformation « simple », comme : $X \mapsto \lambda - X$. La décomposition de F doit refléter la même invariance. Exprimer cette invariance donne là encore des relations sur les coefficients inconnus.

Injection de valeurs particulières

Quand il reste peu de coefficients à calculer, il peut être intéressant d'injecter, dans l'égalité entre F et sa décomposition, une ou plusieurs valeurs qui ne sont pas des pôles de F .

Si F est dans $\mathbb{R}(X)$, on peut injecter une valeur complexe (comme i ou j), l'identification donnant alors deux relations entre les coefficients réels inconnus.

Utilisation de la méthode $\lim_{x \rightarrow \infty} xF(x)$

On suppose ici que le degré de F est strictement négatif (la partie entière est donc nulle).

La décomposition de F fait apparaître des termes du type $\frac{\lambda_k}{X - \alpha_k}$ ou $\frac{a_k X + b_k}{X^2 + \beta_k X + \gamma_k}$.

Le calcul de $\lim_{x \rightarrow \infty} xF(x)$ donne alors une relation liant les coefficients λ_k et a_k .

Cette méthode est intéressante quand il ne reste plus qu'un ou deux coefficients à calculer.

12.8.6 Compléments sur quelques exemples

Décomposition de $F = \frac{1}{X^{2n} - 1}$ dans $\mathbb{R}(X)$

On utilise la décomposition dans $\mathbb{C}(X)$ et on regroupe deux à deux les termes conjugués.

Les racines $2n$ -ièmes de l'unité sont les $\omega_k = \exp \frac{ik\pi}{n}$, avec $0 \leq k \leq 2n - 1$.

Pour tout k de $\{1, \dots, n - 1\}$, on a $\omega_{2n-k} = \overline{\omega_k}$. Ainsi :

$$\begin{aligned} \frac{1}{X^{2n} - 1} &= \frac{1}{2n(X - 1)} - \frac{1}{2n(X + 1)} + \frac{1}{2n} \sum_{k=1}^{n-1} \left(\frac{\omega_k}{X - \omega_k} + \frac{\overline{\omega_k}}{X - \overline{\omega_k}} \right) \\ &= \frac{1}{2n(X - 1)} - \frac{1}{2n(X + 1)} + \frac{1}{n} \sum_{k=1}^{n-1} \frac{X \cos \frac{k\pi}{n} - 1}{X^2 - 2X \cos \frac{k\pi}{n} + 1} \end{aligned}$$

Décomposition de $F = \frac{X^5}{(X - 1)^4}$ dans $\mathbb{R}(X)$

On écrit $X^5 = (X - 1 + 1)^5 = (X - 1)^5 + 5(X - 1)^4 + 10(X - 1)^3 + 10(X - 1)^2 + 5(X - 1) + 1$.

On en déduit : $F = X + 4 + \frac{10}{(X - 1)} + \frac{10}{(X - 1)^2} + \frac{5}{(X - 1)^3} + \frac{1}{(X - 1)^4}$.

Décomposition de $F = \frac{1}{(X^n - 1)^2}$ dans $\mathbb{C}(X)$

Les pôles, tous doubles, sont les racines n -ièmes de l'unité.

Le mieux est de procéder par « dérivation » de la décomposition de $G = \frac{1}{X^n - 1}$.

$$\begin{aligned} \frac{1}{X^n - 1} &= \frac{1}{n} \sum_{k=0}^{n-1} \frac{\omega_k}{X - \omega_k} \Rightarrow \frac{nX^{n-1}}{(X^n - 1)^2} = \frac{1}{n} \sum_{k=0}^{n-1} \frac{\omega_k}{(X - \omega_k)^2} \\ \Rightarrow \frac{nX^n}{(X^n - 1)^2} &= \frac{1}{n} \sum_{k=0}^{n-1} \frac{\omega_k X}{(X - \omega_k)^2} \Rightarrow \frac{n(X^n - 1 + 1)}{(X^n - 1)^2} = \frac{1}{n} \sum_{k=0}^{n-1} \frac{\omega_k(X - \omega_k + \omega_k)}{(X - \omega_k)^2} \\ \Rightarrow \frac{1}{X^n - 1} + \frac{1}{(X^n - 1)^2} &= \frac{1}{n^2} \sum_{k=0}^{n-1} \frac{\omega_k}{X - \omega_k} + \frac{1}{n^2} \sum_{k=0}^{n-1} \frac{\omega_k^2}{(X - \omega_k)^2} \\ \Rightarrow \frac{1}{(X^n - 1)^2} &= \frac{1 - n}{n^2} \sum_{k=0}^{n-1} \frac{\omega_k}{X - \omega_k} + \frac{1}{n^2} \sum_{k=0}^{n-1} \frac{\omega_k^2}{(X - \omega_k)^2} \end{aligned}$$

Décomposition de $F = \frac{1}{X^3(X^2 - 1)}$ dans $\mathbb{R}(X)$

Les pôles sont 0 (triple), 1 (simple) et -1 (simple). La partie entière est nulle.

Posons $F = \frac{a}{X^3} + \frac{b}{X^2} + \frac{c}{X} + \frac{d}{X - 1} + \frac{e}{X + 1}$. L'imparité de F donne $b = 0$ et $e = d$.

On trouve a en multipliant par X^3 et en substituant 0 à X . Donc $a = -1$.

On trouve d en multipliant par $(X - 1)$ et en substituant 1 à X . Donc $d = \frac{1}{2}$.

Si on utilise la méthode $\lim_{x \rightarrow \infty} xF(x)$, on trouve $0 = c + 1$.

Finalement : $F = \frac{1}{X^3(X^2 - 1)} = -\frac{1}{X^3} - \frac{1}{X} + \frac{1}{2(X - 1)} + \frac{1}{2(X + 1)}$.

Décomposition de $F = \frac{1}{X(X^2 + 1)(X^2 + X + 1)(X^2 - X + 1)}$ **dans** $\mathbb{R}(X)$

La décomposition est de la forme : $F = \frac{a}{X} + \frac{cX + d}{X^2 + 1} + \frac{\alpha X + \beta}{X^2 + X + 1} + \frac{\lambda X + \mu}{X^2 - X + 1}$.

L'imparité de F donne immédiatement : $d = 0$, $\mu = -\beta$ et $\lambda = \alpha$.

On cherche donc a, c, α, β tels que : $F = \frac{a}{X} + \frac{cX}{X^2 + 1} + \frac{\alpha X + \beta}{X^2 + X + 1} + \frac{\alpha X - \beta}{X^2 - X + 1}$.

On trouve a en multipliant par X et en substituant 0 à X . Donc $a = 1$.

On trouve c en multipliant par $X^2 + 1$ et en substituant i à X . Donc $c = -1$.

On trouve α, β en multipliant par $X^2 + X + 1$ et en substituant j à X :

$$\begin{aligned} (X^2 + X + 1)F &= \frac{1}{X(X^2 + 1)(X^2 - X + 1)} = \alpha X + \beta + (X^2 + X + 1)(\dots) \\ &\Rightarrow \alpha j + \beta = \frac{1}{j(j^2 + 1)(j^2 - j + 1)} = \frac{1}{2} \Rightarrow \alpha = 0, \beta = \frac{1}{2} \end{aligned}$$

Finalement : $F = \frac{1}{X} - \frac{X}{X^2 + 1} + \frac{1}{2(X^2 + X + 1)} - \frac{1}{2(X^2 - X + 1)}$.

Décomposition de $F = \frac{X^8}{(X^2 - X + 1)^3}$ **dans** $\mathbb{R}(X)$

On procède à des divisions successives par $B = X^2 - X + 1$.

- on trouve $X^8 = Q_1 B + R_1$, avec $Q_1 = X^6 + X^5 - X^3 - X^2 + 1$ et $R_1 = X - 1$.
- on trouve $Q_1 = Q_2 B + R_2$, avec $Q_2 = X^4 + 2X^3 + X^2 - 2X - 4$ et $R_2 = -2X + 5$.
- on trouve $Q_2 = Q_3 B + R_3$, avec $Q_3 = X^2 + 3X + 3$ et $R_3 = -2X - 7$.

Ainsi : $X^8 = R_1 + R_2 B + R_3 B^2 + Q_3 B^3$ puis $F = \frac{X^8}{B^3} = Q_3 + \frac{R_3}{B} + \frac{R_2}{B^2} + \frac{R_1}{B^3}$.

Conclusion : $F = X^2 + 3X + 3 - \frac{2X + 7}{X^2 - X + 1} - \frac{2X - 5}{(X^2 - X + 1)^2} + \frac{X - 1}{(X^2 - X + 1)^3}$

Chapitre 13

Espaces vectoriels

Sommaire

13.1 Généralités	302
13.1.1 Espace vectoriel sur \mathbb{K} , avec $\mathbb{K} = \mathbb{R}$ ou \mathbb{C}	302
13.1.2 Exemples d'espaces vectoriels	302
13.1.3 Combinaisons linéaires	303
13.2 Sous-espaces vectoriels	303
13.2.1 Notion de sous-espace vectoriel	303
13.2.2 Droites vectorielles et plans vectoriels	304
13.2.3 Intersections de sous-espaces vectoriels	307
13.3 Familles génératrices, libres. Bases	308
13.3.1 Familles génératrices, familles libres	308
13.3.2 Bases et coordonnées	309
13.4 Somme de sous-espaces vectoriels	310
13.4.1 Somme de deux sous-espaces vectoriels	310
13.4.2 Couples de sous-espaces supplémentaires	311
13.4.3 Somme d'un nombre fini de sous-espaces vectoriels	311
13.5 Espaces de dimension finie	312
13.5.1 Existence de bases en dimension finie	312
13.5.2 Dimension d'un espace de dimension finie	313
13.5.3 Rang d'une famille de vecteurs	314
13.5.4 Exemples d'espaces vectoriels de dimension finie	315
13.6 Sous-espaces et dimension	317
13.6.1 Dimension d'un sous-espace d'un espace de dimension finie	317
13.6.2 Supplémentaires en dimension finie	318
13.6.3 Dimension d'une somme de sous-espaces	318

13.1 Généralités

13.1.1 Espace vectoriel sur \mathbb{K} , avec $\mathbb{K} = \mathbb{R}$ ou \mathbb{C}

Définition 13.1.1 (structure d'espace vectoriel sur le corps \mathbb{K} , avec $\mathbb{K} = \mathbb{R}$ ou \mathbb{C})

On dit que l'ensemble E est un *espace vectoriel* sur \mathbb{K} (ou encore un \mathbb{K} -espace vectoriel) si :

- l'ensemble E est muni d'une loi interne $+$ pour laquelle il a une structure de groupe commutatif.
- Il existe une application $(\alpha, u) \rightarrow \alpha u$ de $\mathbb{K} \times E$ dans E , dite *loi externe*, telle que :

$$\begin{cases} \text{pour tous scalaires } \alpha \text{ et } \beta & \left\{ \begin{array}{l} (\alpha + \beta)u = \alpha u + \beta u, \quad \alpha(u + v) = \alpha u + \alpha v \\ \text{pour tous éléments } u, v \text{ de } E \end{array} \right. \\ & \left\{ \begin{array}{l} \alpha(\beta u) = (\alpha\beta)u, \quad 1u = u \end{array} \right. \end{cases}$$

Les éléments d'un \mathbb{K} -espace vectoriel E sont appelés *vecteurs* et les éléments de \mathbb{K} sont appelés *scalaires*.

Le neutre du groupe $(E, +)$ est noté $\vec{0}$ (parfois $\vec{0}_E$) et est appelé *vecteur nul*.

L'espace vectoriel E est parfois noté $(E, +, \cdot)$ pour rappeler l'existence des deux lois.

Proposition 13.1.1 (règles de calcul dans un espace vectoriel)

Soit E un espace vectoriel sur \mathbb{K} . Pour tout scalaire α et pour tous vecteurs u et v :

- on a l'équivalence : $\alpha u = \vec{0} \Leftrightarrow (\alpha = 0 \text{ ou } u = \vec{0})$.
- on a les égalités : $\alpha(-u) = (-\alpha)u = -(\alpha u)$ et $\alpha(u - v) = \alpha u - \alpha v$.

13.1.2 Exemples d'espaces vectoriels

- L'ensemble $\mathbb{K}[X]$ des polynômes à coefficients dans \mathbb{K} est un espace vectoriel sur \mathbb{K} .
- L'ensemble \mathbb{K} est un espace vectoriel sur lui-même, la loi externe étant ici le produit de \mathbb{K} .

Définition 13.1.2 (espace vectoriel produit)

Soit E_1, E_2, \dots, E_n des espaces vectoriels sur \mathbb{K} , et soit $E = E_1 \times E_2 \times \dots \times E_n$ leur produit cartésien.

Alors l'ensemble E est muni d'une structure d'espace vectoriel sur \mathbb{K} quand on pose,

$$\text{pour tous } \begin{cases} u = (u_1, u_2, \dots, u_n) \\ v = (v_1, v_2, \dots, v_n) \end{cases} \text{ de } E, \text{ et pour tout scalaire } \lambda : \begin{cases} u + v = (u_1 + v_1, u_2 + v_2, \dots, u_n + v_n) \\ \lambda u = (\lambda u_1, \lambda u_2, \dots, \lambda u_n) \end{cases}$$

Si E est un \mathbb{K} -espace vectoriel, E^n est un \mathbb{K} -espace vectoriel.

L'ensemble $\mathbb{K}^n = \{(x_1, x_2, \dots, x_n), \text{ les } x_i \in \mathbb{K}\}$ est un \mathbb{K} -espace vectoriel (et ici $\vec{0} = (0, 0, \dots, 0)$).

Proposition 13.1.2 (espace des fonctions d'un ensemble non vide dans un espace vectoriel)

Soit Ω un ensemble non vide quelconque, et soit E un espace vectoriel sur \mathbb{K} .

On note $\mathcal{F}(\Omega, E)$ (ou E^Ω) l'ensemble des applications de Ω dans E .

Pour f, g dans $\mathcal{F}(\Omega, E)$, et λ dans \mathbb{K} , on définit $f + g$ et λf par : $\forall x \in \Omega, \begin{cases} (f + g)(x) = f(x) + g(x) \\ (\lambda f)(x) = \lambda f(x) \end{cases}$

Avec ces opérations, $\mathcal{F}(\Omega, E)$ est un espace vectoriel sur \mathbb{K} .

Le vecteur nul de $\mathcal{F}(\Omega, E)$ est « l'application nulle », qui à tout x de Ω associe le vecteur nul $\vec{0}$ de E .

Pour $\Omega = \mathbb{N}$ et $E = \mathbb{K}$, on obtient le \mathbb{K} -espace vectoriel $\mathbb{K}^{\mathbb{N}}$ des suites d'éléments de \mathbb{K} .

Un \mathbb{C} -espace vectoriel est aussi un \mathbb{R} -espace vectoriel, mais ils doivent être considérés comme différents.

13.1.3 Combinaisons linéaires

Définition 13.1.3 (combinaison linéaire d'un nombre fini de vecteurs)

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

Soit u_1, u_2, \dots, u_n une famille de n vecteurs de E . Soit $\lambda_1, \lambda_2, \dots, \lambda_n$ une famille de n scalaires.

On dit que $v = \lambda_1 u_1 + \lambda_2 u_2 + \dots + \lambda_n u_n$ est une *combinaison linéaire* des vecteurs u_1, u_2, \dots, u_n .

Définition 13.1.4 (famille presque nulle (ou à support fini) de scalaires)

Soit I un ensemble quelconque d'indices.

On dit qu'une famille $(\lambda_i)_{i \in I}$ d'éléments de \mathbb{K} est à *support fini*, ou encore « presque nulle » s'il n'y a qu'un nombre fini d'indices i pour lesquels λ_i est différent de 0.

Définition 13.1.5 (combinaisons linéaires quelconques)

Soit E un espace vectoriel sur \mathbb{K} . Soit I un ensemble quelconque d'indices.

Soit $(u_i)_{i \in I}$ une famille de vecteurs de E . Soit $(\lambda_i)_{i \in I}$ une famille de scalaires, à support fini.

Même si I est infini, la somme $\sum_{i \in I} \lambda_i u_i$ a un sens car elle n'est formée que d'un nombre fini de termes.

La somme $\sum_{i \in I} \lambda_i u_i$ est appelée *combinaison linéaire* des vecteurs u_i avec les coefficients λ_i .

13.2 Sous-espaces vectoriels

13.2.1 Notion de sous-espace vectoriel

Définition 13.2.1 (sous-espace vectoriel)

Soit E un espace vectoriel sur \mathbb{K} . Soit F une partie non vide de E .

On dit que F est un *sous-espace vectoriel* de E si : $\forall (u, v) \in F^2, \forall \lambda \in \mathbb{K}, \begin{cases} u + v \in F \\ \lambda u \in F \end{cases}$

Il revient au même d'écrire : $\forall (u, v) \in F^2, \forall (\lambda, \mu) \in \mathbb{K}^2, \lambda u + \mu v \in F$.

Proposition 13.2.1

Soit E un espace vectoriel sur \mathbb{K} , et soit F un sous-espace vectoriel de E .

Alors l'ensemble F , muni des « lois induites », est lui-même un espace vectoriel sur \mathbb{K} .

La réciproque de la propriété est vraie. On comprendra donc l'expression « sous-espace vectoriel » comme une traduction de l'inclusion d'un espace vectoriel dans un autre.

Remarques

On dit souvent *sous-espace* plutôt que sous-espace vectoriel.

Tous les sous-espaces vectoriels de E contiennent au moins le vecteur nul de E .

Pour montrer que F est un sous-espace vectoriel de E , on n'oubliera pas la condition $F \neq \emptyset$.

On vérifiera par exemple que le vecteur nul $\vec{0}$ de E appartient à F .

Définition 13.2.2 (stabilité par combinaisons linéaires)

Soit F un sous-espace vectoriel de E .

Pour toute famille $(u_i)_{i \in I}$ de vecteurs de F , et pour toute famille $(\lambda_i)_{i \in I}$ de \mathbb{K} (à support fini), la combinaison linéaire $\sum_{i \in I} \lambda_i u_i$ est encore un élément de F .

On exprime cette propriété en disant que F est *stable par combinaisons linéaires*.

Exemples usuels

- Pour tout espace vectoriel E , le singleton $\{\vec{0}\}$ et E lui-même sont des sous-espaces vectoriels de E .
On dit que $\{\vec{0}\}$ est le « sous-espace nul » de E .
- L'ensemble $\mathbb{K}_n[X]$ des polynômes de degré inférieur ou égal à n est un sous-espace vectoriel de $\mathbb{K}[X]$.
- Soit I un intervalle de \mathbb{R} , et soit $\mathcal{F}(I, \mathbb{K})$ l'espace vectoriel de toutes les fonctions de I dans \mathbb{K} .
L'ensemble $\mathcal{C}(I, \mathbb{K})$ des fonctions continues de I dans \mathbb{K} est un sous-espace vectoriel de $\mathcal{F}(I, \mathbb{K})$.
L'ensemble $\mathcal{C}^k(I, \mathbb{K})$ des fonctions de classe \mathcal{C}^k est un sous-espace vectoriel de $\mathcal{C}(I, \mathbb{K})$ donc de $\mathcal{F}(I, \mathbb{K})$.

13.2.2 Droites vectorielles et plans vectoriels**Définition 13.2.3** (vecteurs colinéaires)

Soit E un espace vectoriel sur \mathbb{K} . Soit u et v deux vecteurs de E .

On dit que u et v sont colinéaires (ou proportionnels) s'il existe α dans \mathbb{K} tel que $v = \alpha u$ ou $u = \alpha v$.

Le vecteur nul $\vec{0}$ est colinéaire à tout vecteur u de E (on a en effet $\vec{0} = \alpha u$ avec $\alpha = 0$).

Si u est non nul, les vecteurs colinéaires à u forment l'ensemble $\mathbb{K}u = \{\lambda u, \lambda \in \mathbb{K}\}$.

Définition 13.2.4 (droite vectorielle engendrée par un vecteur)

Soit u un vecteur non nul d'un espace vectoriel E sur \mathbb{K} .

L'ensemble $\mathbb{K}u = \{\lambda u, \lambda \in \mathbb{K}\}$ est appelé *la droite vectorielle engendrée par u* .

Si u et v sont deux vecteurs non nuls et colinéaires, ils engendrent la même droite vectorielle.

Soit u un vecteur non nul d'un espace vectoriel E .

La droite vectorielle engendrée par u , c'est-à-dire $\mathbb{K}u = \{\lambda u, \lambda \in \mathbb{K}\}$ est un sous-espace vectoriel de E .

Réciproquement, si un sous-espace F de E contient u , alors il contient la droite engendrée par u .

Définition 13.2.5 (plan vectoriel engendré par deux vecteurs non colinéaires)

Soit u et v deux vecteurs non colinéaires (donc non nuls) d'un espace vectoriel E sur \mathbb{K} .

On note $\mathbb{K}u + \mathbb{K}v = \{\lambda u + \mu v, (\lambda, \mu) \in \mathbb{K}^2\}$ l'ensemble des combinaisons linéaires de u et v .

On dit que $\mathbb{K}u + \mathbb{K}v$ est le *plan vectoriel engendré* par les vecteurs u et v .

Avec les notations précédentes (donc u et v non colinéaires dans E).

- $\mathbb{K}u + \mathbb{K}v$ est un sous-espace de E qui contient strictement les droites vectorielles $\mathbb{K}u$ et $\mathbb{K}v$.
- Tout vecteur w de $F = \mathbb{K}u + \mathbb{K}v$ s'écrit de manière *unique* sous la forme $w = \lambda u + \mu v$, et on dit que les scalaires λ et μ sont les *coordonnées* de w dans la *base* (u, v) du plan F .

Proposition 13.2.2 (condition de non colinéarité de deux vecteurs d'un plan)

Soit u et v deux vecteurs non colinéaires (donc non nuls) d'un espace vectoriel E sur \mathbb{K} .

Soit $w = \lambda u + \mu v$ et $w' = \lambda' u + \mu' v$ deux vecteurs du plan $F = \mathbb{K}u + \mathbb{K}v$ engendré par u et v .

Alors w et w' sont non colinéaires si et seulement si $\lambda\mu' \neq \lambda'\mu$.

Dans ce cas, les deux vecteurs w et w' à leur tour constituent une base du plan F .

▷ **Droites vectorielles de \mathbb{K}^2**

On se place dans $\mathbb{K}^2 = \{(x, y), x \in \mathbb{K}, y \in \mathbb{K}\}$, muni de sa structure d'espace vectoriel sur \mathbb{K} .

Proposition 13.2.3 (droites vectorielles de \mathbb{K}^2)

Soit F un sous-espace de \mathbb{K}^2 , avec $F \neq \{\vec{0}\}$ et $F \neq \mathbb{K}^2$.

Soit u non nul dans F . Alors F est la droite vectorielle engendrée par u : $F = \mathbb{K}u = \{\lambda u, \lambda \in \mathbb{K}\}$.

Posons $u = (a, b)$, avec $u \neq \vec{0}$. Soit $w = (x, y)$ un vecteur quelconque de \mathbb{K}^2 .

Alors le vecteur w est dans la droite $\mathbb{K}u$ si et seulement si il existe un scalaire λ tel que $\begin{cases} x = \lambda a \\ y = \lambda b \end{cases}$

On dit que le système $\begin{cases} x = \lambda a \\ y = \lambda b \end{cases}$ est une *représentation paramétrique* de la droite $\mathbb{K}u$.

De même, le vecteur $v = (x, y)$ est dans $\mathbb{K}u$ si et seulement si $bx - ay = 0$.

On dit que l'égalité $bx - ay = 0$ est une *équation cartésienne* de la droite $\mathbb{K}u$.

Par exemple : la droite engendrée par $u = (2, -3)$ a pour équation $\frac{x}{2} = \frac{y}{-3}$ c'est-à-dire $3x + 2y = 0$.

▷ **Plans vectoriels de \mathbb{K}^3**

On se place dans $\mathbb{K}^3 = \{(x, y, z), x \in \mathbb{K}, y \in \mathbb{K}, z \in \mathbb{K}\}$, muni de sa structure d'espace vectoriel sur \mathbb{K} .

Proposition 13.2.4 (colinéarité ou non de deux vecteurs de \mathbb{K}^3)

Soit $u = (a, b, c)$ et $v = (a', b', c')$ deux vecteurs de \mathbb{K}^3 . On note $u \wedge v = (bc' - cb', ca' - ac', ab' - ba')$.

Alors u et v sont non colinéaires si et seulement si $u \wedge v \neq \vec{0}$.

On note souvent $u \wedge v = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \wedge \begin{pmatrix} a' \\ b' \\ c' \end{pmatrix} = \begin{pmatrix} bc' - cb' \\ ca' - ac' \\ ab' - ba' \end{pmatrix}$ (c'est le « produit vectoriel » de u et v).

Proposition 13.2.5 (plans vectoriels de \mathbb{K}^3)

Soit F un sous-espace de \mathbb{K}^3 , avec $F \neq \{\vec{0}\}$ et $F \neq \mathbb{K}^3$.

On suppose que F contient deux vecteurs non colinéaires u et v .

Alors F est le plan vectoriel engendré par u et v : $F = \mathbb{K}u + \mathbb{K}v = \{\lambda u + \mu v, (\lambda, \mu) \in \mathbb{K}^2\}$.

Soit $u = (a, b, c)$ et $v = (a', b', c')$ deux vecteurs non colinéaires de \mathbb{K}^3 .

Soit $w = (x, y, z)$ un vecteur quelconque de \mathbb{K}^3 .

Alors w est dans le plan $F = \mathbb{K}u + \mathbb{K}v$ si et seulement si il existe (λ, μ) tel que $\begin{cases} x = \lambda a + \mu a' \\ y = \lambda b + \mu b' \\ z = \lambda c + \mu c' \end{cases}$

Ce système est appelé *représentation paramétrique* du plan $F = \mathbb{K}u + \mathbb{K}v$.

De même, $w = (x, y, z)$ est dans $\mathbb{K}u + \mathbb{K}v$ si et seulement si $(bc' - cb')x + (ca' - ac')y + (ab' - ba')z = 0$.
On dit qu'une telle égalité est une *équation cartésienne* du plan $F = \mathbb{K}u + \mathbb{K}v$.

Prenons un exemple :

Les deux vecteurs $u = (1, 1, 2)$ et $v = (3, 1, 1)$ ne sont pas proportionnels dans \mathbb{R}^3 .

Ils engendrent donc dans \mathbb{R}^3 un plan vectoriel F ayant pour représentation paramétrique :

$$\begin{cases} x = \lambda + 3\mu \\ y = \lambda + \mu \\ z = 2\lambda + \mu \end{cases}$$

On résout deux équations par rapport à λ et μ et on reporte dans l'autre.

On trouve $\begin{cases} \lambda = z - y \\ \mu = 2y - z \end{cases}$ puis $x = (z - y) + 3(2y - z)$ donc $x - 5y + 2z = 0$.

On a ainsi obtenu une équation cartésienne du plan F .

On aurait pu trouver les coefficients de cette équation en formant $v \wedge u = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} \wedge \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ -5 \\ 2 \end{pmatrix}$.

▷ Droites vectorielles de \mathbb{K}^3

Proposition 13.2.6 (droites vectorielles de \mathbb{K}^3)

Soit F un sous-espace de \mathbb{K}^3 , avec $F \neq \{\vec{0}\}$ et $F \neq \mathbb{K}^3$. On suppose que F n'est pas un plan vectoriel. Alors F est une droite vectorielle. Il existe donc u non nul dans \mathbb{K}^3 tel que $F = \mathbb{K}u = \{\lambda u, \lambda \in \mathbb{K}\}$.

Avec les notations précédentes, posons $u = (a, b, c)$ et $v = (x, y, z)$.

Alors le vecteur v est dans $F = \mathbb{K}u$ si et seulement si il existe un scalaire λ tel que $\begin{cases} x = \lambda a \\ y = \lambda b \\ z = \lambda c \end{cases}$

Ce système est appelé *représentation paramétrique* de la droite $F = \mathbb{K}u$ dans \mathbb{K}^3 .

De même, $w = (x, y, z)$ est dans $\mathbb{K}u$ si et seulement si : $\frac{x}{a} = \frac{y}{b} = \frac{z}{c}$.

On dit que ces deux égalités forment *un système d'équations cartésiennes* de $F = \mathbb{K}u$ dans \mathbb{K}^3 .

Remarque : dans ce système, si par exemple $a = 0$, on remplace $\frac{x}{a} = \dots$ par l'équation $x = 0$.

Prenons un exemple :

La droite F engendrée par $u = (2, -3, 1)$ a pour équations $\frac{x}{2} = \frac{y}{-3} = \frac{z}{1}$ c'est-à-dire $\begin{cases} 3x + 2y = 0 \\ x - 2z = 0 \end{cases}$

Mais ce n'est pas le seul système d'équations possible pour F .

On pourrait écrire aussi $\begin{cases} 3x + 2y = 0 \\ y + 3z = 0 \end{cases}$ ou encore $\begin{cases} x - 2z = 0 \\ y + 3z = 0 \end{cases}$

Ces différents systèmes sont autant de façons d'interpréter la droite F de \mathbb{R}^3 comme l'intersection de deux plans de \mathbb{R}^3 (il y a d'ailleurs une infinité de couples de plans qui « s'intersectent » en F).

▷ Intersection de deux plans vectoriels dans \mathbb{K}^3

Proposition 13.2.7 (intersection de deux plans vectoriels de \mathbb{K}^3)

Soit F et G deux plans vectoriels de \mathbb{K}^3 .

Si F et G sont distincts, leur intersection est une droite vectorielle.

Plus précisément, soit $ax + by + cz = 0$ et $a'x + b'y + c'z = 0$ des équations cartésiennes de F et G .

Alors la droite vectorielle $F \cap G$ est engendrée par le vecteur $(bc' - cb', ca' - ac', ab' - ba')$.

Prenons un exemple :

Soit F et G les plans de \mathbb{R}^3 d'équations respectives $3x + 2y + 5z = 0$ et $x + y + 4z = 0$.

Ces deux plans ne sont pas confondus : par exemple $(1, -1, 0)$ est dans G mais pas dans F .

La droite $D = F \cap G$ s'obtient en résolvant
$$\begin{cases} 3x + 2y + 5z = 0 \\ x + y + 4z = 0 \end{cases} \quad \text{et on trouve } \begin{cases} x = 3z \\ y = -7z \end{cases}$$

Ainsi les vecteurs de D sont les $u = (x, y, z) = (3z, -7z, z) = z(3, -7, 1)$, avec z dans \mathbb{R} .

Autrement dit, $D = F \cap G$ est la droite vectorielle engendrée par $w = (3, -7, 1)$.

On aurait pu trouver directement ce résultat en écrivant : $v \wedge u = \begin{pmatrix} 3 \\ 2 \\ 5 \end{pmatrix} \wedge \begin{pmatrix} 1 \\ 1 \\ 4 \end{pmatrix} = \begin{pmatrix} 3 \\ -7 \\ 1 \end{pmatrix}$.

13.2.3 Intersections de sous-espaces vectoriels

Proposition 13.2.8 (intersections de sous-espaces vectoriels)

Soit E un espace vectoriel sur \mathbb{K} . Soit I un ensemble quelconque d'indices.

Soit $(F_i)_{i \in I}$ une famille de sous-espaces vectoriels de E .

Alors leur intersection $F = \bigcap_{i \in I} F_i$ est encore un sous-espace vectoriel de E .

Définition 13.2.6 (ensemble des combinaisons linéaires d'une famille de vecteurs)

Soit E un espace vectoriel sur \mathbb{K} . Soit I un ensemble quelconque d'indices.

Soit $(u_i)_{i \in I}$ une famille d'éléments de E . Soit $X = \{u_i, i \in I\}$.

On note $\text{Vect}(X)$, ou encore $\text{Vect}(u_i)_{i \in I}$, l'ensemble des combinaisons linéaires d'éléments de X .

Autrement dit, le vecteur v est dans $\text{Vect}(X)$ si et seulement si il existe une famille $(\lambda_i)_{i \in I}$ à support fini, telle que $v = \sum_{i \in I} \lambda_i u_i$.

Si X est un ensemble fini $\{u_i, 1 \leq i \leq n\}$, alors $\text{Vect}(X) = \left\{ v = \sum_{i=1}^n \lambda_i u_i, \lambda_i \in \mathbb{K} \right\}$.

Rien n'indique que l'écriture d'un élément v de $\text{Vect}(u_i)_{i \in I}$ en fonction des u_i soit unique (à suivre...)

Proposition 13.2.9 (sous-espace engendré par une partie)

Soit E un espace vectoriel sur \mathbb{K} . Soit I un ensemble quelconque d'indices.

Soit $(u_i)_{i \in I}$ une famille d'éléments de E . Soit $X = \{u_i, i \in I\}$.

Alors $\text{Vect}(u_i)_{i \in I}$ est un sous-espace vectoriel de E appelé sous-espace engendré par X .

Au sens de l'inclusion, $\text{Vect}(u_i)_{i \in I}$ est le plus petit sous-espace de E qui contient X .

13.3 Familles génératrices, libres. Bases

On rappelle que la lettre \mathbb{K} désigne indifféremment \mathbb{R} ou \mathbb{C} .

Dans ce qui suit, les familles de scalaires $(\lambda_i)_{i \in I}$ sont toujours supposées à support fini.

13.3.1 Familles génératrices, familles libres

Définition 13.3.1 (familles génératrices)

Soit E un espace vectoriel sur \mathbb{K} . Soit $(u_i)_{i \in I}$ une famille d'éléments de E .

On dit que la famille $(u_i)_{i \in I}$ est *génératrice* dans E si $\text{Vect}(u_i)_{i \in I} = E$.

Cela signifie que tout vecteur u de E est, *au moins d'une manière*, combinaison linéaire des u_i .

Définition 13.3.2 (familles libres)

Soit E un espace vectoriel sur \mathbb{K} . Soit $(u_i)_{i \in I}$ une famille d'éléments de E .

On dit que la famille $(u_i)_{i \in I}$ est *libre* si on a l'implication : $\sum_{i \in I} \lambda_i u_i = \vec{0} \Rightarrow (\forall i \in I, \lambda_i = 0)$.

On exprime aussi cette situation en disant que les vecteurs u_i sont *linéairement indépendants*.

Sinon, donc s'il existe une famille $(\lambda_i)_{i \in I}$ de scalaires *non tous nuls* telle que $\sum_{i \in I} \lambda_i u_i = \vec{0}$, on dit que la famille $(u_i)_{i \in I}$ est *liée*, ou encore que les u_i sont *linéairement dépendants*.

Dans la proposition précédente, on ne doit pas confondre « non tous nuls » et « tous non nuls ».

Cas d'une famille finie de vecteurs

On se donne n vecteurs u_1, u_2, \dots, u_n de l'espace vectoriel E .

Dire que ces vecteurs « engendrent » E , c'est dire que : $\forall v \in E, \exists (\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n, v = \sum_{i=1}^n \lambda_i u_i$.

Dire qu'ils sont libres, c'est dire que : $\forall (\lambda_1, \lambda_2, \dots, \lambda_n) \in \mathbb{K}^n, \left(\sum_{i=1}^n \lambda_i u_i = \vec{0} \Rightarrow \lambda_1 = \dots = \lambda_n = 0 \right)$.

Une famille réduite à un seul vecteur u est libre si et seulement si ce vecteur est non nul.

Extension au cas d'une « famille vide »

Si l'ensemble I des indices est vide, on peut encore parler de la famille $(u_i)_{i \in I}$ (c'est une famille vide, il n'y a personne dedans!), et dire que cette famille est libre.

Le sous-espace engendré par la famille vide est le singleton $\{\vec{0}\}$, ce qui est cohérent avec le fait qu'une somme vide est toujours considérée comme ayant une valeur nulle (ici $\vec{0}$).

Proposition 13.3.1 (une caractérisation des familles liées)

Une famille de vecteurs de E est liée si et seulement si l'un (au moins) des vecteurs qui la compose peut s'écrire comme une combinaison linéaire des autres.

Une famille de deux vecteurs u et v est donc liée si et seulement si u et v sont *colinéaires*, c'est-à-dire s'il existe un scalaire λ tel que $u = \lambda v$ ou $v = \lambda u$.

Attention à ne pas dire (si on ne sait rien de u et v) que u et v sont liés si et seulement si il existe un scalaire λ tel que $u = \lambda v$, car c'est faux si $v = \vec{0}$ et $u \neq \vec{0}$ (en revanche c'est vrai si $v \neq \vec{0}$).

Remarques

- Toute sous-famille d'une famille libre est, a fortiori, une famille libre.
Toute « sur-famille » d'une famille génératrice de E est, a fortiori, une famille génératrice.
En particulier toute famille contenant $\vec{0}$, ou deux vecteurs colinéaires, est liée.
- Soit F un sous-espace vectoriel *strict* de E .
Soit $(u_i)_{i \in I}$ une famille de vecteurs de F , donc de E . Le caractère libre ou non de cette famille ne dépend pas de l'espace vectoriel, F ou E , auxquels ils sont censés appartenir.
En revanche, si cette famille est génératrice dans F , elle ne l'est pas dans E . Quand il y a un risque d'ambiguïté, on précisera donc de quel espace vectoriel une famille de vecteurs est génératrice.

Proposition 13.3.2 (famille de polynômes à degrés distincts deux à deux)

On se place dans l'espace vectoriel $\mathbb{K}[X]$ des polynômes à coefficients dans \mathbb{K} .

Toute famille de polynômes non nuls dont les degrés sont différents deux à deux est une famille libre.

Cette propriété est vraie en particulier pour $(P_n)_{n \in \mathbb{N}}$ si : $\deg P_0 < \deg P_1 < \dots < \deg P_n < \dots$.

On exprime cette situation en disant que la famille des polynômes P_k est à *degrés échelonnés*.

Attention à ne pas annoncer de réciproque !

Par exemple, les trois polynômes X^2 , $X^2 + X$ et $X^2 + X + 1$ sont libres, alors qu'ils ont le même degré.

13.3.2 Bases et coordonnées

Définition 13.3.3 (bases d'un espace vectoriel)

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

On dit qu'une famille $(u_i)_{i \in I}$ d'éléments de E est une *base* si elle est à la fois libre et génératrice.

Proposition 13.3.3 (existence de bases (théorème admis))

Dans tout espace vectoriel E , il y a des bases.

Ce théorème sera démontré pour les « espaces vectoriels de dimension finie ».

Cas très particulier : on peut considérer que la « famille vide » est une base d'un espace réduit à $\{\vec{0}\}$.

Proposition 13.3.4 (caractérisation des bases)

Soit E un espace vectoriel sur \mathbb{K} . Soit $(u_i)_{i \in I}$ une famille d'éléments de E .

La famille $(u_i)_{i \in I}$ est une base si et seulement si tout vecteur v de E peut s'écrire, et de manière unique, comme une combinaison linéaire $v = \sum_{i \in I} \lambda_i u_i$ des vecteurs u_i .

Dans cette écriture, les coefficients λ_i sont appelés les coordonnées de v dans la base $(u_i)_{i \in I}$.

Cas d'une famille finie de vecteurs

On se donne n vecteurs u_1, u_2, \dots, u_n de l'espace vectoriel E .

Dire que ces vecteurs forment une base de E , c'est dire que : $\forall v \in E, \exists ! (\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n, v = \sum_{i=1}^n \lambda_i u_i$.

Importance de l'ordre dans lequel sont donnés les vecteurs d'une base

Si par exemple u_1, u_2, u_3 (dans cet ordre) forment une base de E , et si les coordonnées d'un vecteur v dans cette base sont a, b, c , (c'est-à-dire si $v = au_1 + bu_2 + cu_3$), alors u_2, u_3, u_1 (dans cet ordre) forment une base de E dans laquelle les coordonnées de v sont b, c, a . Conclusion : deux bases se déduisant l'une de l'autre par modification de l'ordre des vecteurs doivent être considérées comme *distinctes*.

Exemples usuels

– Soit n dans \mathbb{N} . On note $e_1 = (1, 0, \dots, 0)$, $e_2 = (0, 1, 0, \dots, 0)$, etc., $e_n = (0, \dots, 0, 1)$.

La famille $(e_k)_{1 \leq k \leq n}$ est une base de \mathbb{K}^n , dite « base canonique ».

Les coordonnées de $v = (x_1, \dots, x_n)$ dans cette base sont les coefficients x_1, \dots, x_n , car $v = \sum_{k=1}^n x_k e_k$.

– Soit n dans \mathbb{N} . La famille finie $1, X, X^2, \dots, X^n$ est une base de $\mathbb{K}_n[X]$, dite « base canonique ».

Plus généralement, la famille infinie $(X^k)_{k \in \mathbb{N}}$ est une base (dite *base canonique*) de $\mathbb{K}[X]$.

Les coordonnées d'un polynôme $A = \sum_{k=0}^{+\infty} a_k X^k$ dans cette base sont les coefficients a_k de A .

13.4 Somme de sous-espaces vectoriels

13.4.1 Somme de deux sous-espaces vectoriels

Soit F et G deux sous-espaces de E . On sait que $F \cap G$ est un sous-espace de E .

En revanche $H = F \cup G$ n'est pas un sous-espace de E , sauf si $F \subset G$ ou $G \subset F$.

Définition 13.4.1 (somme de deux sous-espaces vectoriels)

Soit E un espace vectoriel sur \mathbb{K} . Soit F et G deux sous-espaces de E .

On pose $F + G = \{u + v, u \in F, v \in G\}$.

On dit que $F + G$ est la somme des deux sous-espaces F et G .

Proposition 13.4.1 (une caractérisation de la somme de deux sous-espaces vectoriels)

Avec les notations précédentes, $F + G$ est un sous-espace vectoriel de E .

Plus précisément, $F + G = \text{Vect}(F \cup G)$, le plus petit sous-espace de E contenant à la fois F et G .

Définition 13.4.2 (somme directe de deux sous-espaces vectoriels)

Soit E un espace vectoriel sur \mathbb{K} . Soit F et G deux sous-espaces de E .

Les conditions suivantes sont équivalentes :

- tout vecteur w de $F + G$ s'écrit de façon *unique* sous la forme $w = u + v$, avec u dans F et v dans G .
- pour tout vecteur u de F et tout vecteur v dans G , on a l'implication : $u + v = \vec{0} \Rightarrow u = v = \vec{0}$.
- l'intersection $F \cap G$ est réduite à $\{\vec{0}\}$.

Si elles sont réalisées, on dit que F et G sont en *somme directe* et la somme $F + G$ est notée $F \oplus G$.

13.4.2 Couples de sous-espaces supplémentaires

Définition 13.4.3 (sous-espaces supplémentaires)

Soit F et G deux sous-espaces de E .

On dit que F et G sont *supplémentaires* si $\begin{cases} E = F + G \\ F \cap G = \{\vec{0}\} \end{cases}$ c'est-à-dire si $E = F \oplus G$.

Cela signifie que tout u de E s'écrit d'une manière unique $u = v + w$, avec v dans F et w dans G .

Proposition 13.4.2 (existence d'un supplémentaire (théorème admis))

Soit F un sous-espace vectoriel de E .

Alors F possède au moins un supplémentaire G dans E .

Remarques

Le résultat précédent sera démontré dans le cas particulier des « espaces vectoriels de dimension finie ».

Un même sous-espace F de E possède en général une infinité de supplémentaires dans E .

Il y a cependant un cas d'unicité : le seul supplémentaire de E (resp. de $\{\vec{0}\}$) est $\{\vec{0}\}$ (resp. E).

On ne confondra pas *supplémentaire* et *complémentaire* ! Le complémentaire d'un sous-espace F de E est sans grand intérêt : ce n'est pas un sous-espace vectoriel car il ne contient pas le vecteur nul.

Une bêtise classique est de dire que $F + G$ est directe si et seulement si $F \cap G$ est vide ! Cette intersection n'est en effet *jamais* vide car elle contient $\vec{0}$. Il faut en fait vérifier que $F \cap G$ se réduit à $\{\vec{0}\}$.

Un exemple classique de sous-espaces supplémentaires

Dans l'espace vectoriel $\mathcal{F}(\mathbb{R}, \mathbb{R})$ des fonctions de \mathbb{R} dans \mathbb{R} , les sous-espaces $\mathcal{P}(\mathbb{R}, \mathbb{R})$ et $\mathcal{A}(\mathbb{R}, \mathbb{R})$ formés respectivement des fonctions paires et des fonctions impaires sont supplémentaires.

Plus précisément, soit f une fonction de \mathbb{R} dans \mathbb{R} .

Sa décomposition en la somme $f = p + i$ d'une fonction paire p et d'une fonction impaire i s'écrit :

$$\forall x \in \mathbb{R}, p(x) = \frac{1}{2}(f(x) + f(-x)) \quad \text{et} \quad \forall x \in \mathbb{R}, i(x) = \frac{1}{2}(f(x) - f(-x))$$

13.4.3 Somme d'un nombre fini de sous-espaces vectoriels

Définition 13.4.4

Soit $(F_i)_{1 \leq i \leq n}$ une famille de n sous-espaces vectoriels de E .

On note $\sum_{i=1}^n F_i$ l'ensemble des sommes $v = \sum_{i=1}^n u_i$, où pour tout i de $\{1, \dots, n\}$, u_i est dans F_i .

Proposition 13.4.3 (une caractérisation de la somme de n sous-espaces vectoriels)

Soit $(F_i)_{1 \leq i \leq n}$ une famille de n sous-espaces vectoriels de E .

Alors $\sum_{i=1}^n F_i$ est un sous-espace vectoriel de E , appelé somme des F_i .

Plus précisément, $\sum_{i=1}^n F_i = \text{Vect}\left(\bigcup_{i=1}^n F_i\right)$, c'est-à-dire le plus petit sous-espace de E contenant tous les F_i .

Proposition 13.4.4 (somme directe de n sous-espaces vectoriels)

Soit $(F_i)_{1 \leq i \leq n}$ une famille de n sous-espaces de E . Les conditions suivantes sont équivalentes :

- tout vecteur v de $\sum_{i=1}^n F_i$ s'écrit de façon unique sous la forme $\sum_{i=1}^n u_i$, avec u_i dans F_i pour tout i .
- pour tout choix de vecteurs u_i dans F_i , on a l'implication : $\sum_{i=1}^n u_i = \vec{0} \Rightarrow (\forall i \in \{1, \dots, n\}, u_i = \vec{0})$

Si elles sont réalisées, on dit que la somme $\sum_{i=1}^n F_i$ est directe et on la note $\bigoplus_{i=1}^n F_i$.

Remarques

Si la somme $\sum_{i=1}^n F_i$ est directe, et si J est une partie de I , alors la somme $\sum_{i \in J} F_i$ est encore directe.

En particulier, pour tous indices distincts i et j , $F_i \cap F_j = \{\vec{0}\}$.

La réciproque est fautive ! Pour montrer que F_1, F_2, \dots, F_n sont en somme directe, avec $n \geq 3$, il ne suffit pas de vérifier que pour tous indices distincts i et j , $F_i \cap F_j = \{\vec{0}\}$.

Ce serait encore pire de se contenter de vérifier que $F_1 \cap F_2 \cap \dots \cap F_n = \{\vec{0}\}$.

13.5 Espaces de dimension finie

13.5.1 Existence de bases en dimension finie

Définition 13.5.1 (notion d'espace vectoriel de dimension finie)

Soit E un \mathbb{K} -espace vectoriel.

On dit que E est de dimension finie si E possède une famille génératrice finie.

Avec cette définition, l'espace réduit à $\{\vec{0}\}$ est de dimension finie.

Si un espace vectoriel n'est pas de dimension finie, il est dit... de dimension infinie.

C'est le cas de l'espace vectoriel $\mathbb{K}[X]$ des polynômes à coefficients dans \mathbb{K} .

Remarque : à ce stade, on sait ce qu'est un espace vectoriel de dimension finie E , mais on ne sait pas encore ce qu'est (donc ce que vaut) la dimension de E .

Proposition 13.5.1 (existence de bases dans un espace de dimension finie)

Soit E un espace vectoriel de dimension finie sur \mathbb{K} . Soit $(u_i)_{i \in I}$ une famille génératrice finie de E .

Soit J une partie de I pour laquelle la famille $(u_j)_{j \in J}$ est libre.

Alors il existe une partie K telle que $J \subset K \subset I$ et pour laquelle la famille $(u_k)_{k \in K}$ est une base de E .

Conséquence : dans tout espace vectoriel de dimension finie, il existe des bases.

Proposition 13.5.2 (théorèmes de la base extraite et de la base incomplète)

Soit E un espace vectoriel de dimension finie sur \mathbb{K} .

Théorème « de la base extraite » : de toute famille génératrice de E on peut extraire une base.

Théorème « de la base incomplète » : toute famille libre de E peut être complétée en une base.

13.5.2 Dimension d'un espace de dimension finie

Proposition 13.5.3 (une propriété essentielle pour la suite!)

Soit E un espace vectoriel de dimension finie sur \mathbb{K} .

On suppose que E est engendré par une famille $(u_i)_{1 \leq i \leq n}$ de n vecteurs.

Alors toute famille d'au moins $n + 1$ vecteurs de E est une famille liée.

Proposition 13.5.4 (dimension d'un espace vectoriel)

Soit E un espace vectoriel de dimension finie sur \mathbb{K} .

Toutes les bases de E sont finies et ont exactement le même nombre d'éléments.

Ce nombre est appelé la dimension de E et il est noté $\dim(E)$.

Avec cette définition, on a $\dim\{\vec{0}\} = 0$, car la seule base de l'espace $\{\vec{0}\}$ est la famille vide.

Proposition 13.5.5 (condition suffisante pour qu'une famille forme une base)

Soit E un \mathbb{K} -espace vectoriel de dimension finie $n \geq 1$.

Soit $(u) = u_1, u_2, \dots, u_n$ une famille de n éléments de E (donc autant que la dimension de E).

Si cette famille est génératrice, ou si elle est libre, alors c'est une base de E .

Une interprétation de la dimension

Soit E un \mathbb{K} -espace vectoriel, de dimension n .

Toute famille libre de E possède au plus n vecteurs (toute famille de plus de n vecteurs est liée).

Toute famille génératrice de E possède au moins n vecteurs.

La dimension n de l'espace vectoriel E est donc à la fois le nombre minimum d'éléments d'une famille génératrice, et le nombre maximum d'éléments d'une famille libre.

Droites vectorielles

On appelle *droite vectorielle* tout espace de dimension 1, c'est-à-dire l'ensemble $E = \{\lambda u, \lambda \in \mathbb{K}\}$ des multiples d'un certain vecteur non nul u .

En fait, tout vecteur v non nul de E en constitue une base, donc $E = \{\lambda v, \lambda \in \mathbb{K}\}$.

Plans vectoriels

On appelle *plan vectoriel* tout espace E de dimension 2, donc tout ensemble $E = \{\lambda u + \mu v, (\lambda, \mu) \in \mathbb{K}^2\}$ des combinaisons linéaires de deux vecteurs non colinéaires u et v .

Deux vecteurs u' et v' non proportionnels de E forment à leur tour une base de E .

Dépendance de la dimension par rapport au corps des scalaires

Si E est muni d'une structure d'espace vectoriel sur \mathbb{C} , on sait qu'on peut le munir d'une structure d'espace vectoriel sur \mathbb{R} (il suffit de restreindre l'ensemble des coefficients de \mathbb{C} à \mathbb{R}), mais que ces deux espaces vectoriels ne doivent pas être confondus.

Plus précisément, si E est un espace de dimension n sur \mathbb{C} , c'est un espace de dimension $2n$ sur \mathbb{R} .

Par exemple, \mathbb{C} est une droite vectorielle sur \mathbb{C} , mais c'est un plan vectoriel sur \mathbb{R} .

Autre exemple :

– Si on considère $E = \mathbb{C}^2$ comme un espace vectoriel sur \mathbb{C} , il est de dimension 2, et les vecteurs $e_1 = (1, 0)$ et $e_2 = (0, 1)$ en constituent la base canonique.

En effet tout élément $z = (z_1, z_2)$ de $E = \mathbb{C}^2$ s'écrit de façon unique $z = z_1(1, 0) + z_2(0, 1)$ (les coordonnées z_1 et z_2 sont des complexes).

– Mais si on considère $E = \mathbb{C}^2$ comme un espace sur \mathbb{R} , il est alors de dimension 4, et une base est formée des vecteurs $\varepsilon_1 = (1, 0)$, $\varepsilon_2 = (i, 0)$, $\varepsilon_3 = (0, 1)$, $\varepsilon_4 = (0, i)$.

En effet tout élément $z = (z_1, z_2)$ de $E = \mathbb{C}^2$ s'écrit de façon unique :

$$z = (x_1 + iy_1, x_2 + iy_2) = x_1(1, 0) + y_1(i, 0) + x_2(0, 1) + y_2(0, i) \text{ (les coordonnées sont des réels).}$$

Pour éviter toute ambiguïté, on peut noter $\dim_{\mathbb{K}}(E)$ la dimension de E comme espace vectoriel sur \mathbb{K} .

Proposition 13.5.6 (produit d'espaces vectoriels de dimension finie)

Soit E et F deux espaces vectoriels de dimension finie sur \mathbb{K} .

Alors $E \times F$ est de dimension finie, et $\dim(E \times F) = \dim E + \dim F$.

Plus généralement : $\dim(E_1 \times E_2 \times \cdots \times E_m) = \sum_{i=1}^m \dim(E_i)$, et $\dim(E^m) = m \dim(E)$.

13.5.3 Rang d'une famille de vecteurs

Définition 13.5.2 (rang d'une famille finie de vecteurs)

Soit $(u) = u_1, u_2, \dots, u_p$ une famille de p vecteurs d'un espace vectoriel E sur \mathbb{K} .

On appelle *rang* de la famille (u) la dimension du sous-espace de E engendré par cette famille.

Autrement dit : $\text{rg}(u_1, u_2, \dots, u_p) = \dim(\text{Vect}\{u_1, u_2, \dots, u_p\})$.

On a $\text{rg}(u_1, u_2, \dots, u_p) \leq p$, avec égalité si et seulement si la famille (u) est libre.

Si $\dim(E) = n$, alors $\text{rg}(u_1, u_2, \dots, u_p) \leq n$, avec égalité si et seulement si la famille (u) engendre E .

Définition 13.5.3 (opérations élémentaires sur les vecteurs d'une famille)

Soit $(u) = u_1, u_2, \dots, u_p$ une famille de p vecteurs d'un espace vectoriel E .

On appelle *opération élémentaire* sur les vecteurs de cette famille l'une des opérations suivantes :

- multiplier un des vecteurs par un scalaire non nul.
- ajouter à l'un des vecteurs un multiple d'un autre vecteur de la famille.
- échanger deux vecteurs de la famille.

Proposition 13.5.7

Soit (u') la famille obtenue en appliquant une opération élémentaire à une famille de vecteurs (u) .

Alors les deux familles (u) et (u') ont le même rang.

On ne modifie donc pas le rang d'une famille de vecteurs en lui appliquant une *succession* d'opérations élémentaires. Il en est ainsi quand on ajoute à l'un des vecteurs une combinaison linéaire des autres vecteurs de la famille.

Un exemple de calcul de rang d'une famille de vecteurs

Dans \mathbb{R}^4 , on se donne les vecteurs $a = (1, 2, 2, 1)$, $b = (4, 3, 10, 5)$, $c = (-1, -3, 4, 0)$, $d = (0, 4, -3, -1)$. Le système $\alpha a + \beta b + \gamma c + \delta d = \vec{0}$ équivaut à :

$$\begin{cases} \alpha + 4\beta - \gamma = 0 \\ 2\alpha + 3\beta - 3\gamma + 4\delta = 0 \\ 2\alpha + 10\beta + 4\gamma - 3\delta = 0 \\ \alpha + 5\beta - \delta = 0 \end{cases} \Leftrightarrow \begin{cases} \alpha = -4\beta + \gamma \\ 5\beta + \gamma - 4\delta = 0 \\ 2\beta + 6\gamma - 3\delta = 0 \\ \beta + \gamma - \delta = 0 \end{cases} \Leftrightarrow \begin{cases} \alpha = -4\beta + \gamma \\ \beta - 3\gamma = 0 \\ \beta - 3\gamma = 0 \\ \delta = \beta + \gamma \end{cases} \Leftrightarrow \begin{cases} \alpha = -11\gamma \\ \beta = 3\gamma \\ \delta = 4\gamma \end{cases}$$

On constate que la famille a, b, c, d est liée (donc de rang inférieur ou égal à 3).

Avec $\gamma = 0$, c'est-à-dire si on résout $\alpha a + \beta b + \delta d = \vec{0}$, on trouve $\alpha = \beta = \delta = 0$.

La famille a, b, d est donc libre : c'est une base de $\text{Vect}(a, b, c, d)$. Ainsi $\text{rg}(a, b, c, d) = 3$.

13.5.4 Exemples d'espaces vectoriels de dimension finie

▷ L'espace vectoriel \mathbb{K}^n

L'ensemble \mathbb{K}^n est un espace vectoriel de dimension n .

Une base (dite « canonique ») de \mathbb{K}^n est la famille $(e_k)_{1 \leq k \leq n}$, où, pour tout entier k de $\{1, \dots, n\}$: $e_1 = (1, 0, \dots, 0)$, $e_2 = (0, 1, 0, \dots, 0)$, ..., $e_n = (0, \dots, 0, 1)$.

Les coordonnées de $u = (x_1, \dots, x_n)$ dans cette base sont x_1, \dots, x_n , car $u = \sum_{k=1}^n x_k e_k$.

▷ L'espace $\mathbb{K}_n[X]$ des polynômes de degré inférieur ou égal à n

L'ensemble $\mathbb{K}_n[X]$ des polynômes de degré inférieur ou égal à n est un espace de dimension $n + 1$.

Une base (dite « canonique ») de $\mathbb{K}_n[X]$ est : $1, X, X^2, \dots, X^n$.

Les coordonnées de $A = \sum_{k=0}^n a_k X^k$ dans cette base sont les coefficients a_0, a_1, \dots, a_n de A .

Il est bien sûr possible de former de nombreuses bases de $\mathbb{K}_n[X]$.

Par exemple, toute famille P_0, P_1, \dots, P_n , avec $\deg(P_k) = k$ (degrés échelonnés), est une base de $\mathbb{K}_n[X]$.

Autre idée : soit a, b deux éléments distincts de \mathbb{K} .

Considérons les $n + 1$ polynômes $P_k = (X - a)^k (X - b)^{n-k}$, avec $0 \leq k \leq n$, tous de degré n .

Alors ces $n + 1$ polynômes forment une base de $\mathbb{K}_n[X]$ (il suffit de vérifier qu'ils forment une famille libre, ce qui est plus simple que de prouver qu'ils forment une famille génératrice).

▷ Solutions d'équations différentielles linéaires homogènes d'ordre 1

On rappelle ici un résultat établi dans la section 5.5.2

Proposition 13.5.8 (solution générale de (H))

Soit I un intervalle ouvert non vide. Soit $x \mapsto a(x)$ une fonction continue sur I , à valeurs dans \mathbb{K} .

On considère l'équation (H) : $y' + a(x)y = 0$, sur l'intervalle I .

Soit A une primitive particulière de $x \mapsto a(x)$ sur I .

La solution générale de (H) sur I s'écrit $y(x) = \lambda e^{-A(x)}$, où λ est quelconque dans \mathbb{K} .

En d'autres termes l'ensemble des solutions \mathcal{S}_H est une droite vectorielle sur \mathbb{K} .

▷ Solutions d'équations différentielles linéaires homogènes d'ordre 2

On rappelle ici un résultat établi dans la section 5.6.2

Proposition 13.5.9 (solution générale de (H) dans le cas complexe)

Soit a, b dans \mathbb{C} , et soit (H) l'équation différentielle : $y'' + ay' + by = 0$.

Soit $\Delta = a^2 - 4b$, le discriminant de l'équation caractéristique (C) : $r^2 + ar + b = 0$

– Si $\Delta \neq 0$, l'équation (C) possède deux solutions complexes distinctes r et s .

La solution générale de (H) s'écrit alors : $y(x) = \lambda e^{rx} + \mu e^{sx}$, avec λ et μ dans \mathbb{C} .

– Si $\Delta = 0$, l'équation (C) possède une solution double r dans \mathbb{C} .

La solution générale de (H) s'écrit alors : $y(x) = (\lambda x + \mu) e^{rx}$, avec λ et μ dans \mathbb{C} .

Proposition 13.5.10 (solution générale de (H) dans le cas réel)

Soit a, b dans \mathbb{R} , et soit (H) l'équation différentielle : $y'' + ay' + by = 0$.

Soit $\Delta = a^2 - 4b$, le discriminant de l'équation caractéristique (C) : $r^2 + ar + b = 0$

– Si $\Delta > 0$, l'équation (C) possède deux solutions réelles distinctes r et s .

La solution générale de (H) s'écrit : $y(x) = \lambda e^{rx} + \mu e^{sx}$, avec λ et μ dans \mathbb{R} .

– Si $\Delta = 0$, l'équation (C) possède une solution double r dans \mathbb{R} .

La solution générale de (H) s'écrit : $y(x) = (\lambda x + \mu) e^{rx}$, avec λ et μ dans \mathbb{R} .

– Si $\Delta < 0$, l'équation (C) possède deux solutions complexes conjuguées distinctes r et \bar{r} .

Posons $r = \alpha + i\beta$, avec $(\alpha, \beta) \in \mathbb{R} \times \mathbb{R}^*$.

La solution générale de (H) est $y(x) = e^{\alpha x}(\lambda \cos(\beta x) + \mu \sin(\beta x))$, avec λ et μ dans \mathbb{R} .

En d'autres termes, l'ensemble des solutions \mathcal{S}_H est toujours un plan vectoriel sur \mathbb{K} .

Par exemple, considérons l'équation différentielle (E) : $y'' - 3y' + 2y = 0$.

L'équation caractéristique est (C) : $r^2 - 3r + 2 = 0$, de racines $r = 1$ et $r = 2$.

La solution générale de (H) sur \mathbb{R} s'écrit : $y(x) = \lambda e^x + \mu e^{2x}$, avec λ, μ dans \mathbb{R} .

C'est effectivement un plan vectoriel, dont une base est formée des fonctions $x \mapsto e^x$ et $x \mapsto e^{2x}$.

▷ Suites satisfaisant à une relation de récurrence linéaire d'ordre 2

On rappelle ici un résultat établi dans la section 6.6.4

Proposition 13.5.11 (solution générale dans le cas complexe)

Soit a, b, c trois nombres complexes, avec $a \neq 0$ et $b \neq 0$.

Soit \mathcal{S}_E l'ensemble des suites de \mathbb{C} qui vérifient la relation (E) : $\forall n \in \mathbb{N}, au_{n+2} + bu_{n+1} + cu_n = 0$.

Soit $\Delta = b^2 - 4ac$ le discriminant de l'équation caractéristique (C) : $at^2 + bt + c = 0$

– Si $\Delta \neq 0$, soit r et s les deux racines distinctes de l'équation caractéristique (C) .

Alors \mathcal{S}_E est l'ensemble des suites $n \mapsto \lambda r^n + \mu s^n$, avec λ, μ dans \mathbb{C} .

– Si $\Delta = 0$, soit r la racine double de l'équation caractéristique (C) .

Alors \mathcal{S}_E est l'ensemble des suites $n \mapsto (\lambda n + \mu) r^n$, avec λ, μ dans \mathbb{C} .

Proposition 13.5.12 (solution générale dans le cas réel)

Soit a, b, c trois nombres réels, avec $a \neq 0$ et $b \neq 0$.

Soit \mathcal{S}_E l'ensemble des suites de \mathbb{R} qui vérifient la relation (E) : $\forall n \in \mathbb{N}, au_{n+2} + bu_{n+1} + cu_n = 0$.

Soit $\Delta = b^2 - 4ac$ le discriminant de l'équation caractéristique (C) : $at^2 + bt + c = 0$

– Si $\Delta > 0$, soit r et s les deux racines réelles distinctes de l'équation caractéristique (C).

Alors \mathcal{S}_E est l'ensemble des suites $n \mapsto \lambda r^n + \mu s^n$, avec λ, μ dans \mathbb{R} .

– Si $\Delta = 0$, soit r la racine réelle double de l'équation caractéristique (C).

Alors \mathcal{S}_E est l'ensemble des suites $n \mapsto (\lambda n + \mu)r^n$, avec λ, μ dans \mathbb{R} .

– Si $\Delta < 0$, soit $r = \rho e^{i\theta}$ et $\bar{r} = \rho e^{-i\theta}$ les racines complexes conjuguées distinctes de (C), avec $\theta \neq 0 [\pi]$.

Alors \mathcal{S}_E est l'ensemble des suites $n \mapsto (\lambda \cos(n\theta) + \mu \sin(n\theta))\rho^n$, avec λ, μ dans \mathbb{R} .

Autrement dit, l'ensemble des suites solutions de la récurrence est toujours un plan vectoriel sur \mathbb{K} .

Par exemple, les suites arithmétiques, caractérisées par $u_{n+2} - 2u_{n+1} + u_n = 0$, sont les suites de terme général $u_n = a + bn$, avec a, b dans \mathbb{K} . Elles forment un plan, dont une base est $n \mapsto 1$ et $n \mapsto n$.

Pour les suites « à la Fibonacci » ($u_{n+2} = u_{n+1} + u_n$) l'équation caractéristique est (C) : $t^2 - t - 1 = 0$.

Elle possède les racines distinctes : $\Phi = \frac{1 + \sqrt{5}}{2}$ (le « nombre d'or ») et $\Psi = \frac{1 - \sqrt{5}}{2} = -\frac{1}{\Phi}$.

L'ensemble des solutions est le plan des suites $n \mapsto u_n = \lambda\Phi^n + \mu\Psi^n$.

13.6 Sous-espaces et dimension

13.6.1 Dimension d'un sous-espace d'un espace de dimension finie

Proposition 13.6.1 (sous-espaces d'un espace vectoriel de dimension finie)

Soit E un \mathbb{K} -espace vectoriel de dimension n . Soit F un sous-espace vectoriel de E .

Alors F est de dimension finie et $\dim(F) \leq \dim(E)$, avec égalité si et seulement si $F = E$.

Sous-espaces de \mathbb{R}^2

Les sous-espaces vectoriels de \mathbb{R}^2 sont $\{\vec{0}\}$, \mathbb{R}^2 lui-même, et les droites vectorielles $\mathbb{R}u$ avec $u \neq \vec{0}$.

Si $u = (a, b) \neq \vec{0}$, alors : $v(x, y) \in \mathbb{R}u \Leftrightarrow (\exists \lambda \in \mathbb{R}, \begin{cases} x = \lambda a \\ y = \lambda b \end{cases}) \Leftrightarrow bx - ay = 0$.

Sous-espaces de \mathbb{R}^3

Les sous-espaces de \mathbb{R}^3 sont $\{\vec{0}\}$, \mathbb{R}^3 lui-même, les droites $\mathbb{R}u$ et les plans $\mathbb{R}u \oplus \mathbb{R}v$.

– Soit $u = (a, b, c)$ non nul dans \mathbb{R}^3 .

Alors $v(x, y, z) \in \mathbb{R}u \Leftrightarrow (\exists \lambda \in \mathbb{R}, \begin{cases} x = \lambda a \\ y = \lambda b \\ z = \lambda c \end{cases}) \Leftrightarrow \frac{x}{a} = \frac{y}{b} = \frac{z}{c}$ (si $a = 0$, remplacer par $x = 0$).

– Soit $u = (a, b, c)$ et $v = (a', b', c')$ deux vecteurs non colinéaires.

Posons $h = (a'', b'', c'')$, avec $a'' = bc' - cb'$, $b'' = ca' - ac'$, $c'' = ab' - ba'$ (vecteur non nul de \mathbb{R}^3).

Alors : $w(x, y, z) \in \mathbb{R}u \oplus \mathbb{R}v \Leftrightarrow (\exists (\lambda, \mu) \in \mathbb{R}^2, \begin{cases} x = \lambda a + \mu a' \\ y = \lambda b + \mu b' \\ z = \lambda c + \mu c' \end{cases}) \Leftrightarrow a''x + b''y + c''z = 0$.

13.6.2 Supplémentaires en dimension finie

Proposition 13.6.2 (existence d'un supplémentaire en dimension finie)

Soit E un \mathbb{K} -espace vectoriel de dimension finie, et soit F un sous-espace vectoriel de E .
Alors F possède des supplémentaires dans E .

Proposition 13.6.3 (base adaptée à une somme directe)

Soit E un espace vectoriel de dimension finie. Soit F et G deux sous-espaces supplémentaires de E .
Soit (f) une base de F , et soit (g) une base de G .
Soit (e) la famille de vecteurs obtenue en « juxtaposant » les familles (f) et (g) .
Alors (e) est une base de E , dite « adaptée à la somme directe » $E = F \oplus G$.

Conséquence importante :

Si $\dim(E) = n$, et si $\dim(F) = p$, tous les supplémentaires de F dans E sont de dimension $n - p$.

Généralisation au cas d'une somme directe de plusieurs sous-espaces

Soit E un espace vectoriel quelconque.

Soit F_1, F_2, \dots, F_p une famille de p sous-espaces de dimension finie de E , en somme directe.

Pour tout j de $\{1, \dots, p\}$, soit $(e)_j$ une base de F_j .

On forme la famille $(e) = (e)_1 \cup (e)_2 \cup \dots \cup (e)_p$ en « juxtaposant » les familles $(e)_j$.

Alors la famille (e) est une base de l'espace $\bigoplus_{j=1}^p F_j$, dite adaptée à cette somme directe.

13.6.3 Dimension d'une somme de sous-espaces

Proposition 13.6.4 (dimension de la somme de deux sous-espaces vectoriels)

Soit E un espace vectoriel quelconque (donc pas nécessairement de dimension finie).

Soit F et G deux sous-espaces vectoriels de dimension finie de E .

Alors $F + G$ est un sous-espace de dimension finie de E .

De plus on a la « formule de Grassmann » : $\dim(F + G) = \dim(F) + \dim(G) - \dim(F \cap G)$.

Cas particulier important :

Soit F et G deux sous-espaces vectoriels de dimension finie de E .

Alors, on a l'équivalence : $\dim(F + G) = \dim(F) + \dim(G) \Leftrightarrow F \cap G = \{\vec{0}\}$.

Proposition 13.6.5 (généralisation à la somme d'un nombre fini de sous-espaces)

Soit E un espace vectoriel quelconque (donc pas nécessairement de dimension finie).

Soit F_1, F_2, \dots, F_p une famille de p sous-espaces vectoriels de dimension finie de E .

On a $\dim\left(\sum_{i=1}^p F_i\right) \leq \sum_{i=1}^p \dim(F_i)$, avec égalité si et seulement si la somme $\sum_{i=1}^p F_i$ est directe.

Chapitre 14

Applications linéaires

Sommaire

14.1 Généralités	320
14.1.1 Notion d'application linéaire	320
14.1.2 Premiers exemples d'applications linéaires	320
14.1.3 Opérations sur les applications linéaires	321
14.1.4 Applications linéaires et sous-espaces vectoriels	321
14.2 Endomorphismes	322
14.2.1 Structure d'anneau de $(\mathcal{L}(E), +, \circ)$	322
14.2.2 Le groupe linéaire $(GL(E), \circ)$	323
14.2.3 Projections et symétries vectorielles	323
14.3 Détermination des applications linéaires	324
14.3.1 Applications linéaires et familles de vecteurs	324
14.3.2 Restrictions aux sous-espaces d'une somme directe	325
14.3.3 Isomorphismes et dimensions	325
14.3.4 Applications linéaires de rang fini	326
14.3.5 Le théorème du rang et ses conséquences	327
14.4 Formes linéaires, hyperplans	328
14.4.1 Formes linéaires	328
14.4.2 Hyperplans vectoriels	329
14.4.3 Systèmes d'équations de sous-espaces vectoriels	330

14.1 Généralités

14.1.1 Notion d'application linéaire

Définition 14.1.1 (applications linéaires d'un espace vectoriel dans un autre)

Soit E et F deux espaces vectoriels sur \mathbb{K} .

Une application f de E dans F est dite *linéaire* si : $\forall (u, v) \in E^2, \forall \lambda \in \mathbb{K}, \begin{cases} f(u + v) = f(u) + f(v) \\ f(\lambda u) = \lambda f(u) \end{cases}$

On dit aussi que f est un *morphisme* d'espaces vectoriels.

Premières propriétés

$f : E \rightarrow F$ est linéaire si et seulement si :

$$\forall (u, v) \in E^2, \forall (\alpha, \beta) \in \mathbb{K}^2, f(\alpha u + \beta v) = \alpha f(u) + \beta f(v)$$

Si f est linéaire, alors $f\left(\sum_{i \in I} \lambda_i u_i\right) = \sum_{i \in I} \lambda_i f(u_i)$ pour toute combinaison linéaire.

On exprime cette propriété en disant qu'une application linéaire « conserve les combinaisons linéaires ».

Si $f : E \rightarrow F$ est linéaire, alors $f(\vec{0}_E) = \vec{0}_F$ (utile pour montrer la non-linéarité).

Notations et terminologie

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

Un *endomorphisme* de E est une application linéaire de E dans lui-même.

On note $\mathcal{L}(E)$ (plutôt que $\mathcal{L}(E, E)$) l'ensemble des endomorphismes de E .

Un *isomorphisme* est une application linéaire bijective.

On dit qu'un isomorphisme de E dans lui-même est un *automorphisme* de E .

On dit qu'une application linéaire de E dans \mathbb{K} est une *forme linéaire*.

14.1.2 Premiers exemples d'applications linéaires

– Soit E et F deux espaces vectoriels sur \mathbb{K} . L'application nulle de E dans F est linéaire.

– Soit $E = \mathcal{C}([a, b], \mathbb{K})$ l'espace vectoriel des applications continues de $[a, b]$ dans \mathbb{K} .

L'application $f \mapsto \varphi(f) = \int_a^b f(t) dt$ est une forme linéaire sur l'espace $E = \mathcal{C}([a, b], \mathbb{K})$.

– L'application $f \mapsto f'$ est un endomorphisme de $E = \mathcal{C}^\infty(I, \mathbb{R})$.

– Les formes linéaires sur \mathbb{K}^n sont les applications $f : (x_1, x_2, \dots, x_n) \mapsto \sum_{k=1}^n \lambda_k x_k$.

– Soit f_1, \dots, f_p des applications linéaires de E dans F .

Alors l'application $u \mapsto f(u) = (f_1(u), \dots, f_p(u))$ est linéaire de E dans F^p .

Par exemple, l'application $(x, y, z) \mapsto (2x - y + z, x - 3y + 2z)$ est linéaire de \mathbb{R}^3 dans \mathbb{R}^2 .

Plus généralement, tout application $f : (x_1, \dots, x_n) \mapsto \left(\sum_{k=1}^n a_k x_k, \sum_{k=1}^n b_k x_k, \sum_{k=1}^n c_k x_k, \text{etc.}\right)$ est linéaire.

14.1.3 Opérations sur les applications linéaires

Proposition 14.1.1 (combinaisons linéaires d'applications linéaires)

Soit E et F deux espaces vectoriels sur \mathbb{K} .

Soit f et g deux applications linéaires de E dans F , et soit α, β deux scalaires.

Alors l'application $\alpha f + \beta g$ est linéaire de E dans F .

Il en résulte que $\mathcal{L}(E, F)$ est un espace vectoriel sur \mathbb{K} .

Proposition 14.1.2 (composition d'applications linéaires)

Soit E, F et G trois espaces vectoriels sur \mathbb{K} .

Si $f : E \rightarrow F$ et $g : F \rightarrow G$ sont linéaires, alors $g \circ f$ est linéaire de E dans G .

Proposition 14.1.3 (isomorphisme réciproque)

Soit f un isomorphisme de E sur F . Sa bijection réciproque f^{-1} est un isomorphisme de F sur E .

Bilinéarité de la composition des applications linéaires

Soit f, g, h des applications linéaires, et soit α, β des scalaires.

L'égalité $(\alpha f + \beta g) \circ h = \alpha f \circ h + \beta g \circ h$ est toujours vraie (et cela n'utilise pas la linéarité).

L'égalité $h \circ (\alpha f + \beta g) = \alpha h \circ f + \beta h \circ g$ est vraie (et cela utilise la linéarité de h).

On retiendra qu'entre applications linéaires, la composition $(f, g) \mapsto f \circ g$ est « bilinéaire ».

14.1.4 Applications linéaires et sous-espaces vectoriels

Proposition 14.1.4 (image d'un sous-espace par une application linéaire)

Soit E et F deux espaces vectoriels sur \mathbb{K} . Soit f une application linéaire de E dans F .

Si E' est un sous-espace vectoriel de E , alors $f(E')$ est un sous-espace vectoriel de F .

Il y a un cas particulier important :

Définition 14.1.2 (image d'une application linéaire)

Soit E et F deux espaces vectoriels sur \mathbb{K} . Soit f une application linéaire de E dans F .

Le sous-espace $f(E) = \{f(u), u \in E\}$ de F est appelé image de f , et il est noté $\text{Im}(f)$.

Proposition 14.1.5 (image réciproque d'un sous-espace par une application linéaire)

Soit E et F deux espaces vectoriels sur \mathbb{K} . Soit f une application linéaire de E dans F .

Si F' est un sous-espace vectoriel de F , alors $f^{-1}(F')$ est un sous-espace vectoriel de E .

Il y a un cas particulier important :

Définition 14.1.3 (noyau d'une application linéaire)

Soit E et F deux espaces vectoriels sur \mathbb{K} . Soit f une application linéaire de E dans F .

Le sous-espace $f^{-1}(\{\vec{0}_F\}) = \{u \in E, f(u) = \vec{0}_F\}$ de E est appelé noyau de f , et il est noté $\text{Ker}(f)$.

On peut parfois montrer qu'une partie d'un espace vectoriel en est un sous-espace vectoriel en l'interprétant comme le noyau ou l'image d'une application linéaire.

Proposition 14.1.6 (caractérisation de l'injectivité par le noyau)

Soit E et F deux espaces vectoriels sur \mathbb{K} . Soit f une application linéaire de E dans F .

f est injective si et seulement si son noyau $\text{Ker}(f)$ se réduit à $\{\vec{0}_E\}$.

Autrement dit, f est injective si et seulement si : $\forall u \in E, f(u) = \vec{0}_F \Rightarrow u = \vec{0}_E$.

Sous-espace $E_\lambda(f)$ des vecteurs u tels que $f(u) = \lambda u$

Soit f un endomorphisme de l'espace vectoriel E , et soit λ un scalaire.

Notons $E_\lambda(f)$ l'ensemble des vecteurs u de E tels que $f(u) = \lambda u$.

On a les équivalences : $f(u) = \lambda u \Leftrightarrow (f - \lambda \text{Id})(u) = \vec{0} \Leftrightarrow u \in \text{Ker}(f - \lambda \text{Id})$.

Ainsi $E_\lambda(f) = \text{Ker}(f - \lambda \text{Id})$, et il en résulte que $E_\lambda(f)$ est un sous-espace vectoriel de E .

Il y a deux cas particuliers importants :

- les vecteurs *invariants* par f : $\text{Inv}(f) = E_1 = \text{Ker}(f - \text{Id}) = \{u \in E, f(u) = u\}$
- les vecteurs changés en leur opposé par f : $\text{Opp}(f) = E_{-1} = \text{Ker}(f + \text{Id}) = \{u \in E, f(u) = -u\}$

14.2 Endomorphismes

14.2.1 Structure d'anneau de $(\mathcal{L}(E), +, \circ)$

L'application *identité*, notée Id_E , définie par $\text{Id}_E(u) = u$ pour tout u , est un automorphisme de E .

Si f et g sont deux endomorphismes de E , alors $g \circ f$ est un endomorphisme de E .

Il en résulte que $(\mathcal{L}(E), +, \circ)$ est muni d'une structure d'anneau.

On note souvent gf la composée des endomorphismes g et f , plutôt que $g \circ f$.

Si f est un endomorphisme de E , on note $f^n = f \circ f \circ \dots \circ f$ (n fois), avec par convention $f^0 = \text{Id}_E$.

Homothéties d'un espace vectoriel

Pour tout scalaire λ , l'application $h_\lambda : u \rightarrow \lambda u$ est un endomorphisme de E .

On a toujours $h_\lambda \circ h_\mu = h_{\lambda\mu}$.

L'application h_λ un automorphisme si $\lambda \neq 0$, et alors $h_\lambda^{-1} = h_{1/\lambda}$.

Si $\lambda \neq 0$, on dit que h_λ est l'*homothétie* de rapport λ .

Endomorphismes qui commutent

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

Si E est réduit à $\{\vec{0}\}$, alors $\mathcal{L}(E)$ est réduit à l'application nulle (sans grand intérêt).

Si E est une droite vectorielle, alors $\mathcal{L}(E)$ se réduit à l'application nulle et aux homothéties de E (et la loi \circ est commutative).

Dans tous les autres cas, donc dès que $\dim(E) \geq 2$, l'anneau $(\mathcal{L}(E), +, \circ)$ est non commutatif (autrement dit : il existe des endomorphismes f et g tels que $f \circ g \neq g \circ f$).

Dans tous les cas, il existe des applications linéaires qui commutent (donc qui vérifient $f \circ g = g \circ f$). Les homothéties $h_\lambda : u \mapsto \lambda u$ de E commutent avec tous les endomorphismes de E .

Si f et g commutent dans $\mathcal{L}(E)$, on peut utiliser la formule du binôme : $(f + g)^n = \sum_{k=0}^n \binom{n}{k} f^k \circ g^{n-k}$.

Ainsi, pour tout f de $\mathcal{L}(E)$ et tout λ de \mathbb{K} , on a : $(f + \lambda \text{Id}_E)^n = \sum_{k=0}^n \binom{n}{k} \lambda^{n-k} f^k$.

14.2.2 Le groupe linéaire $(GL(E), \circ)$

Si f et g sont des automorphismes de E , alors f^{-1} et $g \circ f$ sont encore des automorphismes de E .

On note $GL(E)$ l'ensemble des automorphismes de E .

C'est un groupe pour la loi de composition, appelé le « groupe linéaire de E ».

Le groupe $GL(E)$ est le groupe des éléments inversibles de l'anneau $(\mathcal{L}(E), +, \circ)$.

Ce groupe est non commutatif dès que $\dim(E) \geq 2$.

14.2.3 Projections et symétries vectorielles

Définition 14.2.1 (projections et symétries vectorielles)

Soit F et G deux sous-espaces vectoriels supplémentaires de E .

On sait que pour tout u de E , il existe un unique v de F et un unique w de G tels que $u = v + w$.

L'application $p : u \rightarrow p(u) = v$ est appelée *projection* sur F , parallèlement à G .

l'application $s : u \rightarrow s(u) = v - w$ est appelée *symétrie* par rapport à F , parallèlement à G .

Propriétés

On garde ici les notations de la définition précédente.

L'application p est un endomorphisme de E , et vérifie $p \circ p = p$.

On a $F = \text{Im}(p) = \text{Inv}(p)$ et $G = \text{Ker}(p)$.

L'application s est un automorphisme de E , et vérifie $s \circ s = \text{Id}_E$, donc $s^{-1} = s$ (application *involutive*).

On a la relation $s = 2p - \text{Id}_E$, qui s'écrit encore $p = \frac{1}{2}(s + \text{Id}_E)$.

On a $F = \text{Inv}(f)$ (vecteurs invariants) et $G = \text{Opp}(s)$ (vecteurs changés en leur opposé).

Soit p' la projection sur G parallèlement à F , et s' la symétrie par rapport à G parallèlement à F .

Alors on a les relations :
$$\begin{cases} p + p' = \text{Id}_E, & p \circ p' = p' \circ p = 0 \\ s + s' = 0, & s \circ s' = s' \circ s = -\text{Id}_E \end{cases}$$

Quelques cas très particuliers

On considère la somme directe $E = E \oplus \{\vec{0}\}$.

La projection sur E parallèlement à $\{\vec{0}\}$ est l'application Id_E .

C'est le seul cas où une projection vectorielle est injective.

La projection sur $\{\vec{0}\}$ parallèlement à E est l'application nulle.

La symétrie par rapport à E parallèlement à $\{\vec{0}\}$ est l'application Id_E .

La symétrie par rapport à $\{\vec{0}\}$ parallèlement à E est l'application $-\text{Id}_E$.

Définition 14.2.2 (projecteur d'un espace vectoriel)

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

On appelle *projecteur* de E tout endomorphisme p de E tel que $p \circ p = p$.

On sait déjà qu'une projection vectorielle est un projecteur, mais la réciproque est vraie :

Proposition 14.2.1 (équivalence entre projecteur et projection vectorielle)

Si p est un projecteur de l'espace vectoriel E , on a : $E = \text{Ker}(p) \oplus \text{Im}(p)$.

L'application p est alors la projection vectorielle sur $\text{Im}(p)$ parallèlement à $\text{Ker}(p)$.

On sait déjà qu'une symétrie vectorielle est involutive, mais la réciproque est vraie :

Proposition 14.2.2 (équivalence entre endomorphisme involutif et symétrie vectorielle)

Si s est un endomorphisme involutif de E , donc si $s \circ s = \text{Id}_E$, on a : $E = \text{Inv}(s) \oplus \text{Opp}(s)$.

L'application s est alors la symétrie vectorielle par rapport à $\text{Inv}(s)$ parallèlement à $\text{Opp}(s)$.

Remarque importante

Si f est un endomorphisme de E , beaucoup de choses sont possibles entre $\text{Im}(f)$ et $\text{Ker}(f)$.

Par exemple l'inclusion $\text{Im}(f) \subset \text{Ker}(f)$ équivaut à $f \circ f = 0$.

En particulier, on n'interprétera pas abusivement la propriété $E = \text{Ker}(f) \oplus \text{Im}(f)$.

On montre en fait que : $E = \text{Ker}(f) \oplus \text{Im}(f) \Leftrightarrow \begin{cases} \text{Ker}(f^2) = \text{Ker}(f) \\ \text{Im}(f^2) = \text{Im}(f) \end{cases}$ (ce qui n'équivaut pas à $f^2 = f$)

14.3 Détermination des applications linéaires

14.3.1 Applications linéaires et familles de vecteurs

Proposition 14.3.1 (applications linéaires et familles génératrices)

Soit E et F deux espaces vectoriels sur \mathbb{K} et f une application linéaire de E dans F .

Soit $(u_i)_{i \in I}$ une famille génératrice de E .

Alors la famille $(f(u_i))_{i \in I}$ est génératrice dans $\text{Im}(f)$.

En particulier, si f est surjective, alors la famille $(f(u_i))_{i \in I}$ est génératrice dans F .

Une application linéaire *surjective* transforme donc une famille génératrice en une famille génératrice.

Proposition 14.3.2 (applications linéaires et familles libres)

Soit E et F deux espaces vectoriels sur \mathbb{K} et soit f une application linéaire de E dans F .

Soit $(u_i)_{i \in I}$ une famille d'éléments de E .

Si la famille $(u_i)_{i \in I}$ est liée, alors la famille $(f(u_i))_{i \in I}$ est liée.

Si la famille $(u_i)_{i \in I}$ est libre et si f est injective, alors la famille $(f(u_i))_{i \in I}$ est libre.

Une application linéaire *injective* transforme donc une famille libre en une famille libre.

Proposition 14.3.3 (image d'une base par un isomorphisme)

Soit E et F deux espaces vectoriels sur \mathbb{K} et soit f un isomorphisme de E sur F .

Soit $(u_i)_{i \in I}$ une base de E . Alors la famille $(f(u_i))_{i \in I}$ est une base de F .

Une application linéaire *bijective* transforme donc une base en une base.

On peut finalement énoncer un résultat plus général, qui indique qu'une application linéaire f est *déterminée de manière unique par les images des vecteurs d'une base*. De plus les propriétés de la famille image renseignent sur les propriétés de f (injectivité, surjectivité) :

Proposition 14.3.4 (applications linéaires et bases)

Soit E et F deux espaces vectoriels sur \mathbb{K} , E étant muni d'une base $(e_i)_{i \in I}$.

Soit $(v_i)_{i \in I}$ une famille quelconque de vecteurs de F .

Alors il existe une unique application linéaire f de E dans F telle que : $\forall i \in I, f(e_i) = v_i$.

- l'application f est injective si et seulement si la famille $(v_i)_{i \in I}$ est libre dans F .
- l'application f est surjective si et seulement si la famille $(v_i)_{i \in I}$ est génératrice dans F .
- l'application f est bijective si et seulement si la famille $(v_i)_{i \in I}$ est une base de F .

On retiendra en particulier :

Soit $f = E \rightarrow F$ une application linéaire.

L'application f est un isomorphisme si et seulement si f transforme *une* base de E en une base de F .

Elle transforme alors *toute* base de E en une base de F .

14.3.2 Restrictions aux sous-espaces d'une somme directe

Soit E un espace vectoriel et $(E_i)_{1 \leq i \leq n}$ une famille de sous-espaces de E telle que $E = \bigoplus_{i=1}^n E_i$.

Soit f une application linéaire de E dans un espace vectoriel F .

Pour tout i de $\{1, \dots, n\}$, soit f_i la restriction de f à E_i (c'est un élément de $\mathcal{L}(E_i, F)$).

Si $u = \sum_{i=1}^n u_i$ (avec u_i dans E_i), on a $f(u) = f\left(\sum_{i=1}^n u_i\right) = \sum_{i=1}^n f(u_i)$, c'est-à-dire $f(u) = \sum_{i=1}^n f_i(u_i)$.

Ainsi, connaître la restriction de f à chaque E_i permet de « reconstruire » f de manière unique.

Réciproquement, pour tout i de $\{1, \dots, n\}$ on se donne φ_i linéaire de E_i dans F .

Alors il existe une et une seule application linéaire de E dans F telle que $f|_{E_i} = \varphi_i$ pour tout i .

Toujours avec la notation $u = \sum_{i=1}^n u_i$, cette application f est définie par $f(u) = \sum_{i=1}^n \varphi_i(u_i)$.

14.3.3 Isomorphismes et dimensions

En dimension finie, les résultats précédents ont une conséquence immédiate et importante :

Proposition 14.3.5 (conservation de la dimension par isomorphisme)

Soit E et F deux espaces vectoriels sur \mathbb{K} et soit f un isomorphisme de E sur F .

Si E est de dimension finie, alors F est de dimension finie et on a $\dim(F) = \dim(E)$.

Plus généralement, la proposition suivante indique que la notion de dimension permet d'effectuer une véritable classification des espaces de dimension finie.

Proposition 14.3.6 (espaces isomorphes à un espace de dimension finie)

Soit E et F deux espaces vectoriels sur \mathbb{K} , E étant de dimension finie.

On dit que E et F sont isomorphes s'il existe un isomorphisme de E sur F .

Alors E et F sont isomorphes si et seulement si F est de dimension finie et $\dim(F) = \dim(E)$.

Isomorphisme défini par les images des vecteurs d'une base

Soit E et F deux espaces vectoriels de même dimension $n \geq 1$.

Soit $(e_i)_{1 \leq i \leq n}$ une base de E , et soit $(\varepsilon_i)_{1 \leq i \leq n}$ une base de F .

L'isomorphisme f de E sur F qui transforme chaque e_i en ε_i est défini par : $f\left(\sum_{i=1}^n x_i e_i\right) = \sum_{i=1}^n x_i \varepsilon_i$.

Autrement dit, f est l'application qui envoie un vecteur u de E sur le vecteur v de F qui a (dans la base $(\varepsilon_i)_{1 \leq i \leq n}$) les mêmes coordonnées que u dans la base $(e_i)_{1 \leq i \leq n}$.

Importance de l'espace vectoriel \mathbb{K}^n

Tout \mathbb{K} -espace vectoriel E de dimension $n \geq 1$, est isomorphe à \mathbb{K}^n .

Si $(u_i)_{1 \leq i \leq n}$ est une base de E , alors $\varphi : (x_1, \dots, x_n) \mapsto \sum_{k=1}^n x_k u_k$ est un isomorphisme de \mathbb{K}^n sur E .

Ainsi \mathbb{K}^n est l'exemple-type du \mathbb{K} -espace de dimension n , sa base canonique étant la plus « naturelle ».

La dimension de l'espace $\mathcal{L}(E, F)$

Soit E un espace vectoriel de dimension n et soit $(e_i)_{1 \leq i \leq n}$ une base de E .

Soit F un espace vectoriel quelconque (donc pas forcément de dimension finie).

Soit φ l'application définie sur $\mathcal{L}(E, F)$ par $\varphi(f) = (f(e_1), f(e_2), \dots, f(e_n))$.

Alors φ est un isomorphisme de $\mathcal{L}(E, F)$ sur F^n .

Dans le cas où F est lui-même de dimension finie, on en déduit le résultat suivant :

Proposition 14.3.7 (dimension de l'espace $\mathcal{L}(E, F)$)

Soit E et F deux espaces vectoriels de dimension finie.

Alors l'espace $\mathcal{L}(E, F)$ est de dimension finie. Plus précisément : $\dim(\mathcal{L}(E, F)) = \dim(E) \dim(F)$.

14.3.4 Applications linéaires de rang fini

Définition 14.3.1 (application linéaire de rang fini)

Soit f une application linéaire de E dans F .

On dit que f est de rang fini si $\text{Im}(f)$ est un sous-espace de dimension finie de F .

Si cette condition est réalisée, on dit que $\dim(\text{Im}(f))$ est le rang de f (et on le note $\text{rg}(f)$).

La condition précédente est automatiquement réalisée si l'espace d'arrivée F est de dimension finie p .

En effet $\text{Im}(f)$ est un sous-espace F , donc $\text{rg}(f) = \dim(\text{Im}(f)) \leq p$.

On a $\text{rg}(f) = \dim(F)$ si et seulement si $\text{Im}(f) = F$, c'est-à-dire si et seulement si f est surjective.

Invariance du rang par composition par un isomorphisme

Soit E_1, E_2, E_3, E_4 des espaces vectoriels sur \mathbb{K} .

Soit φ un isomorphisme de E_1 sur E_2 , et soit Ψ un isomorphisme de E_3 sur E_4 .

Soit f une application linéaire de E_2 dans E_3 , de rang fini.

Alors $f \circ \varphi$ et $\psi \circ f$ sont également de rang fini, et : $\text{rg}(f \circ \varphi) = \text{rg}(\psi \circ f) = \text{rg}(f)$.

14.3.5 Le théorème du rang et ses conséquences

Le résultat suivant est très (très!) important :

Proposition 14.3.8 (théorème du rang)

Soit f une application linéaire de E dans F . On suppose que E est de dimension finie.

Soit S un supplémentaire de $\text{Ker}(f)$ dans E .

Alors f induit un isomorphisme de S sur $\text{Im}(f)$.

Il en résulte que $\text{Im}(f)$ est un sous-espace de dimension finie de F (ou encore : f est de rang fini).

Mais plus précisément, on a l'égalité : $\dim(E) = \text{rg}(f) + \dim(\text{Ker}(f))$.

On voit que $\text{rg}(f) \leq \dim(E)$ avec égalité si et seulement si f est injective.

On pourra interpréter ce résultat en disant qu'une application linéaire « ne crée pas de dimension ». Tout au plus, elle « conserve la dimension » si elle est injective.

Les notions de rang d'une application linéaire et de rang d'une famille de vecteurs se rejoignent.

En effet, si $f : E \rightarrow F$ est linéaire, si e_1, e_2, \dots, e_n est une base de E , alors : $\text{rg}(f) = \text{rg}(f(e_1), \dots, f(e_n))$.

Proposition 14.3.9 (caractérisation des isomorphismes en cas de dimensions égales)

Soit E et F deux espaces vectoriels de même dimension n .

Soit f une application linéaire de E dans F .

Alors on a les équivalences : f est un isomorphisme $\Leftrightarrow f$ est injective $\Leftrightarrow f$ est surjective.

On sait que pour montrer qu'une application linéaire f est bijective, une possibilité est d'exhiber un isomorphisme réciproque, c'est-à-dire une application linéaire g telle que $g \circ f = f \circ g = \text{Id}$.

On sait également que la composition des applications linéaires n'est pas commutative, donc qu'il faut a priori vérifier les deux égalités $g \circ f = \text{Id}$ et $f \circ g = \text{Id}$.

On va voir qu'une seule de ces deux égalités est suffisante si E et F sont de même dimension finie, et notamment si f est un endomorphisme d'un espace de dimension finie.

Définition 14.3.2 (inversibilité à gauche ou à droite)

Soit E, F deux espaces vectoriels quelconques, et soit f une application linéaire de E dans F .

On dit que f est « inversible à gauche » s'il existe $g : F \rightarrow E$, linéaire, telle que $g \circ f = \text{Id}_E$.

On dit que f est « inversible à droite » s'il existe $h : F \rightarrow E$, linéaire, telle que $f \circ h = \text{Id}_F$.

Il est clair que si f est inversible à gauche (resp. à droite) elle est injective (resp. surjective).

Si f est inversible à gauche et à droite, alors f est un isomorphisme de E sur F , et (avec les notations précédentes) on a les égalités $g = h = f^{-1}$.

Proposition 14.3.10 (équivalence entre inversibilité à droite et à gauche en même dimension)

Soit E, F deux espaces vectoriels de même dimension finie. Soit $f : E \rightarrow F$, linéaire.

Alors f est inversible à gauche si et seulement si elle est inversible à droite.

On retiendra que pour montrer qu'un endomorphisme f d'un espace E de dimension finie est un automorphisme de E , il suffit de vérifier l'une des deux égalités $g \circ f = \text{Id}_E$ ou $f \circ g = \text{Id}_E$ (l'autre étant alors automatiquement vérifiée, et on a alors $g = f^{-1}$).

14.4 Formes linéaires, hyperplans

14.4.1 Formes linéaires

Définition 14.4.1 (formes linéaires sur un espace vectoriel)

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

On appelle *forme linéaire* sur E toute application linéaire de E dans \mathbb{K} .

Remarques et exemples

- Si une forme linéaire f n'est pas identiquement nulle, alors elle est surjective (et donc de rang 1).
- L'application de \mathbb{R}^3 dans \mathbb{R} définie par $f(x, y, z) = 2x - 3y + 5z$ est une forme linéaire sur \mathbb{R}^3 .
- L'application $\varphi : f \mapsto \int_a^b f(t) dt$ est une forme linéaire sur $\mathcal{C}([a, b], \mathbb{K})$.
- Soit $E = \mathcal{F}(I, \mathbb{K})$ l'espace des applications définies sur l'intervalle I et à valeurs dans \mathbb{K} . Soit x_0 un élément de I . L'application $\varphi : f \mapsto f(x_0)$ est une forme linéaire sur E .

Définition 14.4.2 (formes linéaires coordonnées)

Soit E un espace vectoriel sur \mathbb{K} . Soit $(e_i)_{i \in I}$ une base de E .

On sait que tout vecteur u de E s'écrit de façon unique $u = \sum_{i \in I} x_i e_i$ (somme à support fini).

Pour tout i de I , on note e_i^* l'application qui au vecteur u associe sa coordonnée x_i sur le vecteur e_i .

Les applications $(e_i^*)_{i \in I}$ sont des formes linéaires sur E .

On les appelle les « formes linéaires coordonnées » sur E relativement à la base $(e_i)_{i \in I}$.

Proposition 14.4.1 (en dimension finie, les formes linéaires coordonnées forment une base)

Soit E un espace vectoriel sur \mathbb{K} , de dimension finie n . Soit $(e_i)_{1 \leq i \leq n}$ une base de E .

Alors la famille $(e_i^*)_{1 \leq i \leq n}$ des formes linéaires coordonnées est une base de $\mathcal{L}(E, \mathbb{K})$.

Soit E un espace de dimension finie n , muni d'une base $(e_i)_{1 \leq i \leq n}$.

Soit f une forme linéaire sur E .

Alors f s'écrit, de manière unique : $f = \sum_{i=1}^n a_i e_i^*$. Dans cette écriture on a $a_i = f(e_i)$.

Autrement dit, f est définie de façon unique sous la forme : $f\left(\sum_{i=1}^n x_i e_i\right) = \sum_{k=1}^n a_k x_k$.

Considérons par exemple \mathbb{R}^3 muni de sa base canonique $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 = (0, 0, 1)$.

Les formes linéaires coordonnées sont définies par : $\forall (x, y, z) \in \mathbb{R}^3$,
$$\begin{cases} e_1^*(x, y, z) = x \\ e_2^*(x, y, z) = y \\ e_3^*(x, y, z) = z \end{cases}$$

Les formes linéaires sur \mathbb{R}^3 sont les applications $f = ae_1^* + be_2^* + ce_3^*$, où (a, b, c) est dans \mathbb{R}^3 .

Ou encore : les formes linéaires sur \mathbb{R}^3 sont les applications $(x, y, z) \mapsto ax + by + cz$.

Par exemple si f est définie par $f(x, y, z) = 2x - 3y + 5z$, on a $f(e_1) = 2$, $f(e_2) = -3$ et $f(e_3) = 5$.

14.4.2 Hyperplans vectoriels

Définition 14.4.3 (hyperplans d'un espace vectoriel)

Soit E un espace vectoriel sur \mathbb{K} . Soit H un sous-espace vectoriel de E .

On dit que H est un *hyperplan* de E s'il existe une forme linéaire non nulle f telle que $H = \text{Ker } f$.

Proposition 14.4.2 (caractérisations des hyperplans)

Soit E un espace vectoriel sur \mathbb{K} . Soit H un sous-espace vectoriel de E .

Les conditions suivantes sont équivalentes :

- la sous-espace H est un hyperplan de E .
- pour toute droite vectorielle D non incluse dans H , on a $E = H \oplus D$.
- il existe une droite vectorielle D non incluse dans H , telle que $E = H \oplus D$.

On dispose finalement de trois caractérisations des hyperplans de E :

- les hyperplans de E sont les sous-espaces de E qui sont supplémentaires d'une droite vectorielle.
- les hyperplans de E sont les noyaux des formes linéaires non nulles sur E .
- si $\dim(E) = n \geq 1$, les hyperplans de E sont les sous-espaces de dimension $n - 1$.

Proposition 14.4.3 (formes linéaires non nulles ayant le même noyau)

Soit E un espace vectoriel sur \mathbb{K} , et soit f, g deux formes linéaires non nulles sur E .

Alors f et g ont le même hyperplan noyau si et seulement si f et g sont proportionnelles.

Équations d'un hyperplan

Si H est un hyperplan de E et si f est une forme linéaire non nulle telle que $H = \text{Ker } f$, alors l'égalité $f(x) = 0$ (où x est quelconque dans E) est appelée une *équation* de l'hyperplan H .

La propriété précédente dit que l'équation d'un hyperplan est unique à un facteur multiplicatif près :

Proposition 14.4.4 (équation d'un hyperplan en dimension finie)

Soit E un espace de dimension $n \geq 1$, muni d'une base $(e_i)_{1 \leq i \leq n}$. Soit H un hyperplan de E .

L'équation de H s'écrit $\sum_{i=1}^n a_i x_i = 0$, où les x_i sont les coordonnées d'un vecteur quelconque u .

Cette équation est unique à un facteur multiplicatif non nul près.

Hyperplans de \mathbb{R}^2 et de \mathbb{R}^3

- Les hyperplans de \mathbb{R}^2 sont les droites vectorielles de \mathbb{R}^2 .

Si \mathbb{R}^2 est muni d'une base e_1, e_2 , et si on note (x, y) les coordonnées d'un vecteur quelconque dans cette base, l'équation d'une droite D de \mathbb{R}^2 s'écrit d'une manière unique (à un facteur multiplicatif non nul près) sous la forme $ax + by = 0$, avec $(a, b) \neq (0, 0)$.

Sous cette forme, un vecteur directeur de D est le vecteur $u = -be_1 + ae_2$.

- Les hyperplans de \mathbb{R}^3 sont les plans vectoriels de \mathbb{R}^3 .

Si \mathbb{R}^3 est muni d'une base e_1, e_2, e_3 , et si on note (x, y, z) les coordonnées d'un vecteur u quelconque dans cette base, l'équation d'un plan P de \mathbb{R}^3 s'écrit d'une manière unique (à un facteur multiplicatif non nul près) sous la forme $ax + by + cz = 0$, avec $(a, b, c) \neq (0, 0, 0)$.

Sous cette forme, et si par exemple $a \neq 0$, une base de ce plan est formée des vecteurs $\begin{cases} u_1 = -be_1 + ae_2 \\ u_2 = -ce_1 + ae_3 \end{cases}$

14.4.3 Systèmes d'équations de sous-espaces vectoriels

Proposition 14.4.5 (intersection de deux hyperplans dans un espace de dimension n)

Soit E un \mathbb{K} -espace vectoriel de dimension $n \geq 1$.

Soit H et K deux hyperplans distincts de E . Alors $\dim(H \cap K) = n - 2$.

Par exemple, l'intersection de deux plans distincts de \mathbb{R}^3 est une droite vectorielle.

Proposition 14.4.6 (intersection de m hyperplans dans un espace de dimension n)

Soit E un \mathbb{K} -espace vectoriel de dimension $n \geq 1$.

On se donne m hyperplans H_1, H_2, \dots, H_m de E . Alors $\dim\left(\bigcap_{i=1}^m H_i\right) \geq n - m$.

Un exemple

On se place dans \mathbb{R}^5 muni de sa base canonique.

Soit H_1, H_2, H_3 les hyperplans d'équations respectives :
$$\begin{cases} x_1 - 2x_2 + 3x_3 - x_4 + x_5 = 0 \\ 2x_1 - x_2 + x_3 + x_4 - x_5 = 0 \\ x_1 + 2x_2 - 2x_3 - x_4 + 3x_5 = 0 \end{cases}$$

Soit $F = H_1 \cap H_2 \cap H_3$. D'après la proposition précédente $\dim(F) \geq 5 - 3 = 2$.

Pour déterminer F , on applique la méthode du pivot au système (S) formé par ces trois équations :

En appliquant $\begin{cases} L_2 \leftarrow L_2 - 2L_1 \\ L_3 \leftarrow L_3 - L_1 \end{cases}$ on trouve $(S) \Leftrightarrow \begin{cases} x_1 - 2x_2 + 3x_3 - x_4 + x_5 = 0 \\ 3x_2 - 5x_3 + 3x_4 - 3x_5 = 0 \\ 4x_2 - 5x_3 + 2x_5 = 0 \end{cases}$

En appliquant $\begin{cases} L_1 \leftarrow 3L_1 + 2L_2 \\ L_3 \leftarrow 3L_3 - 4L_2 \end{cases}$, on a $(S) \Leftrightarrow \begin{cases} 3x_1 - x_3 + 3x_4 - 3x_5 = 0 \\ 3x_2 - 5x_3 + 3x_4 - 3x_5 = 0 \\ 5x_3 - 12x_4 + 18x_5 = 0 \end{cases}$

Après $\begin{cases} L_1 \leftarrow 5L_1 + L_3 \\ L_2 \leftarrow L_2 + L_3 \end{cases}$, on trouve $(S) \Leftrightarrow \begin{cases} 15x_1 + 3x_4 + 3x_5 = 0 \\ 3x_2 - 9x_4 + 15x_5 = 0 \\ 5x_3 - 12x_4 + 18x_5 = 0 \end{cases}$

Ainsi $u = (x_1, x_2, x_3, x_4, x_5) \in E \Leftrightarrow \left(x_1 = -\frac{1}{5}(x_4 + x_5), x_2 = 3x_4 - 5x_5, x_3 = \frac{1}{5}(12x_4 - 18x_5)\right)$

On encore : $u \in E \Leftrightarrow u = (x_1, x_2, x_3, x_4, x_5) = \frac{x_4}{5}(-1, 15, 12, 5, 0) - \frac{x_5}{5}(1, 25, 18, 0, -5)$.

Dans cette expression, les composantes x_4 et x_5 sont arbitraires.

Ainsi F est le plan de \mathbb{R}^5 engendré par les deux vecteurs indépendants $\begin{cases} u_1 = (-1, 15, 12, 5, 0) \\ u_2 = (1, 25, 18, 0, -5) \end{cases}$

Voici un résultat qui énonce la réciproque de la proposition précédente.

Proposition 14.4.7

Soit E un \mathbb{K} -espace vectoriel de dimension $n \geq 1$.

Soit F un sous-espace de E , de dimension m , avec $0 \leq m \leq n - 1$.

Alors F peut s'écrire comme l'intersection de $n - m$ hyperplans de E .

Système d'équations cartésiennes d'un sous-espace vectoriel

Si $\dim(F) = m \leq n - 2$, c'est-à-dire si F n'est pas lui-même un hyperplan de E , l'écriture de F comme intersection de $n - m$ hyperplans est loin d'être unique.

Soit $F = \bigcap_{i=1}^n H_i$ une telle écriture du sous-espace F .

Si on munit E d'une base $(e_i)_{1 \leq i \leq n}$, chaque hyperplan H_i est défini par une équation cartésienne (E_i) définie elle-même de façon unique à un facteur multiplicatif non nul près.

Le système formé par ces $n - m$ équations est appelé un « système d'équations » de F .

Un exemple

On se place dans \mathbb{R}^4 muni de sa base canonique.

Soit F le plan de \mathbb{R}^4 engendré par les deux vecteurs $\begin{cases} v_1 = (1, -1, 2, 1) \\ v_2 = (1, 1, 3, -2) \end{cases}$

On va écrire F comme l'intersection de $n - m$ hyperplans, avec ici $n - m = 4 - \dim(F) = 4 - 2 = 2$.

Soit $u = (x, y, z, t)$ un vecteur quelconque de \mathbb{R}^4 .

Dire que u est dans F , c'est dire que la famille (v_1, v_2, u) est de rang 2.

On procède par opérations élémentaires (elles ne modifient pas le rang d'une famille de vecteurs).

$$\left(\begin{array}{cc|c} 1 & 1 & x \\ -1 & 1 & y \\ 2 & 3 & z \\ 1 & -2 & t \end{array} \right) \begin{array}{l} \\ L_2 \leftarrow L_2 + L_1 \\ L_3 \leftarrow L_3 - 2L_1 \\ L_4 \leftarrow L_4 - L_1 \end{array} \Rightarrow \left(\begin{array}{cc|c} 1 & 1 & x \\ 0 & 2 & x + y \\ 0 & 1 & -2x + z \\ 0 & -3 & -x + t \end{array} \right) \begin{array}{l} \\ \\ L_3 \leftarrow 2L_3 - L_1 \\ L_4 \leftarrow 2L_4 + 3L_1 \end{array} \Rightarrow \left(\begin{array}{cc|c} 1 & 1 & x \\ 0 & 2 & x + y \\ 0 & 0 & -5x - y + 2z \\ 0 & 0 & x + 3y + 2t \end{array} \right)$$

Les deux premiers vecteurs obtenus sont indépendants et leurs coefficients sont « échelonnés ».

Dire que la famille obtenue est de rang 2, c'est dire qu'on a $\begin{cases} -5x - y + 2z = 0 \\ x + 3y + 2t = 0 \end{cases}$

On a ainsi obtenu un système d'équations de F .

Chapitre 15

Sous-espaces affines

Sommaire

15.1 Points et vecteurs, translations	333
15.2 Sous-espaces affines	334
15.2.1 Translaté d'un sous-espace vectoriel	334
15.2.2 Dimension d'un sous-espace affine	334
15.2.3 Points alignés, points coplanaires	335
15.2.4 Exemples de sous-espaces affines	335
15.3 Paramétrage d'un sous-espace affine	336
15.4 Parallélisme et intersection	337
15.4.1 Parallélisme de sous-espaces affines	337
15.4.2 Intersection de sous-espaces affines	338
15.5 Équation cartésienne d'un hyperplan	339
15.5.1 Équation cartésienne d'un hyperplan de \mathbb{R}^n	339
15.5.2 Droites affines de \mathbb{R}^2	340
15.5.3 Plans affines de \mathbb{R}^3	340
15.6 Systèmes d'équations d'un sous-espace affine	342
15.6.1 Intersection de p hyperplans affines de \mathbb{R}^n	342
15.6.2 Système d'équations d'un sous-espace de \mathbb{R}^n	343
15.7 Barycentres et repères affines	345
15.7.1 Barycentres	345
15.7.2 Barycentres et sous-espaces affines	346

15.1 Points et vecteurs, translations

Dans tout ce chapitre, E est un espace vectoriel sur \mathbb{R} .

Les éléments de E , selon le rôle qu'on leur fait jouer, sont appelés *points* ou *vecteurs*.

Conventions de notation

Pour limiter les ambiguïtés, on utilisera quelques conventions typographiques :

– Les *points* seront notés A, B, \dots, M, N, \dots

Le vecteur nul $\vec{0}$, considéré comme un point de E , pourra être noté O .

– Les vecteurs seront notés a, b, u, v, \dots , parfois $\vec{a}, \vec{b}, \dots, \vec{u}, \vec{v} \dots$

On choisira la notation \vec{u} par exemple, plutôt que u , pour insister sur le rôle « vectoriel » de cet élément. La distinction entre les deux notations est une simple affaire de commodité.

– Les sous-espaces vectoriels de E seront notés F, G, H, \dots

On définira plus loin les *sous-espaces affines* de E , et on les notera $\mathcal{F}, \mathcal{G}, \mathcal{H}, \dots$

Définition 15.1.1 (translations)

Soit u un vecteur de E .

L'application $t_u : E \rightarrow E$ définie par $t_u(M) = M + u$ est appelée *translation* de vecteur u .

On a bien sûr $t_u(M) = M + u = t_M(u)$, mais l'idée est de considérer u comme un vecteur et M comme un point (auquel on ajoute le vecteur u), même si M et u sont l'un comme l'autre des éléments de E .

Pour tous vecteurs u, v et tout point A , on a : $(A + u) + v = A + (u + v) = (A + v) + u$.

Autrement dit on a les égalités : $t_v \circ t_u = t_{u+v} = t_u \circ t_v$.

On a bien sûr $t_{\vec{0}} = \text{Id}_E$. D'autre part toute translation t_u est bijective et $(t_u)^{-1} = t_{-u}$.

Seule la translation $t_{\vec{0}} = \text{Id}_E$ est linéaire. En effet, si $u \neq \vec{0}$, alors $t_u(\vec{0}) = u \neq \vec{0}$.

Définition 15.1.2 (utilisation de la notation \overrightarrow{AB})

Soit A, B deux points de E . Il existe un unique vecteur u de E tel que $B = t_u(A) = A + u$.

Ce vecteur, égal à $B - A$, est noté \overrightarrow{AB} .

En particulier, on a l'équivalence des notations : $\overrightarrow{AB} = u \Leftrightarrow B = A + u$.

De manière évidente, on a : $\overrightarrow{AB} = \vec{0} \Leftrightarrow A = B$, et $\overrightarrow{BA} = -\overrightarrow{AB}$.

Pour tous points A, B, C de E , on a la « relation de Chasles » : $\overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC}$.

Segment délimitant deux points

Soit A, B deux points de E . On note $[A, B] = \{M \in E, \overrightarrow{AM} = \lambda \overrightarrow{AB}, \lambda \in [0, 1]\}$.

On dit que $[A, B]$ est le *segment* d'origine A et d'extrémité B , et on a $[A, B] = [B, A]$.

On note que $\overrightarrow{AM} = \lambda \overrightarrow{AB} \Leftrightarrow M - A = \lambda(B - A) \Leftrightarrow M = (1 - \lambda)A + \lambda B$.

En particulier, le point I défini par $\overrightarrow{AI} = \frac{1}{2} \overrightarrow{AB}$, c'est-à-dire $I = \frac{A + B}{2}$, est le *milieu* du segment $[A, B]$.

15.2 Sous-espaces affines

15.2.1 Translaté d'un sous-espace vectoriel

Définition 15.2.1 (translaté d'un sous-espace vectoriel)

Soit A un point de E , et F un sous-espace vectoriel de E . On note $A + F = \{A + u, u \in F\}$. Autrement dit, $A + F$ est l'image de F par la translation $u \mapsto A + u$.

Proposition 15.2.1 (égalité de deux translatés de sous-espaces vectoriels)

Soit A, B deux points de E , et soit F, G deux sous-espaces vectoriels de E .

On a l'équivalence $A + F = B + G \Leftrightarrow (F = G \text{ et } \overrightarrow{AB} \in F)$

Définition 15.2.2 (sous-espace affine d'un espace vectoriel)

Soit \mathcal{F} une partie de l'espace vectoriel E . On dit que \mathcal{F} est un *sous-espace affine* de E s'il existe un point A et un sous-espace vectoriel F tel que $\mathcal{F} = A + F$.

On voit que A est dans \mathcal{F} (prendre $\vec{0}$ dans F).

On sait également (proposition précédente) que F est défini de façon unique par l'égalité $\mathcal{F} = A + F$.

On dit alors que $\mathcal{F} = A + F$ est le sous-espace affine de E « passant par A et de direction F ».

Remarques et propriétés

- Un sous-espace affine est toujours non vide.
Les singletons de E sont les sous-espaces affines de direction $\{\vec{0}\}$.
L'ensemble E est un sous-espace affine de lui-même, de direction E .
- Par définition, les sous-espaces affines de E sont les translatés des sous-espaces vectoriels de E .
Les sous-espaces vectoriels F de E sont les sous-espaces affines qui passent par l'origine.
Si F est un sous-espace vectoriel de E , il est sa propre direction.
- Soit \mathcal{F} un sous-espace affine de E , de direction F . Pour tout point B de \mathcal{F} , on a $\mathcal{F} = B + F$.
Un sous-espace affine est donc défini par sa direction et par l'un *quelconque* de ses points.
- Deux sous-espaces affines sont égaux \Leftrightarrow (ils ont la même direction et un point en commun).

15.2.2 Dimension d'un sous-espace affine

Définition 15.2.3 (sous-espaces affines de dimension finie)

Soit F un sous-espace vectoriel de E , et \mathcal{F} un sous-espace affine de E de direction F .

On dit que \mathcal{F} est de dimension finie si F est lui-même de dimension finie.

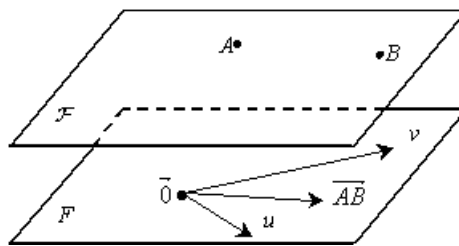
Dans ce cas on note $\dim \mathcal{F} = \dim F$. En particulier :

- les singletons de E sont les sous-espaces affines de dimension 0.
- on appelle *droites affines* les sous-espaces affines de dimension 1.
- on appelle *plans affines* les sous-espaces affines de dimension 2.
- si F est un hyperplan de E , on dit que \mathcal{F} est un *hyperplan affine* de E .

On se place ici dans un espace de dimension 3.

On voit un plan affine \mathcal{F} passant par un point A et de direction un plan vectoriel F de base (u, v) .

Dire que B est dans \mathcal{F} , c'est dire que le vecteur \overrightarrow{AB} est dans F , ou encore est combinaison linéaire de u et v .



15.2.3 Points alignés, points coplanaires

Définition 15.2.4 (points alignés, points coplanaires)

On dit que des points de E sont *alignés* s'ils appartiennent à une même droite affine.

On dit qu'ils sont *coplanaires* s'ils appartiennent à un même plan affine.

– Deux points A, B sont toujours alignés.

S'ils sont distincts, ils appartiennent à la seule droite $\mathcal{D} = (A, \overrightarrow{AB})$.

Trois points A, B, C sont alignés si et seulement si les vecteurs \overrightarrow{AB} et \overrightarrow{AC} sont liés.

Si les trois points A, B, C ne sont pas alignés, on dit parfois qu'ils forment un *vrai triangle*.

– Trois points A, B, C sont toujours coplanaires.

Supposons qu'ils ne soient pas alignés (donc que les vecteurs \overrightarrow{AB} et \overrightarrow{AC} soient libres).

Alors ils appartiennent au seul plan $\mathcal{P} = (A, \overrightarrow{AB}, \overrightarrow{AC})$.

Ce plan peut tout aussi bien être noté $(B, \overrightarrow{BA}, \overrightarrow{BC})$ ou $(C, \overrightarrow{CA}, \overrightarrow{CB})$.

Quatre points A, B, C, D sont coplanaires si et seulement si les vecteurs $\overrightarrow{AB}, \overrightarrow{AC}, \overrightarrow{AD}$ sont liés.

15.2.4 Exemples de sous-espaces affines

Proposition 15.2.2 (image réciproque d'un point par une application linéaire)

Soit E, F deux \mathbb{R} -espaces vectoriels, et soit f dans $\mathcal{L}(E, F)$. Soit b un vecteur de F .

Supposons que $\mathcal{S} = \{u \in E, f(u) = b\}$ soit non vide, c'est-à-dire que b appartienne à $\text{Im}(f)$.

Alors \mathcal{S} est un sous-espace affine de E , de direction $\text{Ker}(f)$.

Solution générale d'un système linéaire

Considérons un système linéaire (S) à p inconnues.

La solution générale de (S) , si elle n'est pas vide, est un sous-espace affine de \mathbb{R}^p .

Solution générale d'une équation différentielle linéaire d'ordre 2

Considérons l'espace $E = \mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ des fonctions indéfiniment dérivables de \mathbb{R} dans \mathbb{R} .

Considérons l'équation différentielle $(\star) : y'' - 5y' + 6y = 6$ (équation caractéristique $(r - 2)(r - 3) = 0$).

L'application f qui à une fonction y de E associe la fonction $y'' - 5y' + 6y$ est un endomorphisme de E .

Chercher le noyau de f , c'est résoudre l'équation homogène associée à (\star) .

Résoudre (\star) , c'est trouver l'image réciproque de la fonction $x \mapsto 6$ par l'endomorphisme f .

La solution \mathcal{S}_H de l'équation homogène, donc $\text{Ker}(f)$, est $x \mapsto y(x) = \lambda e^{2x} + \mu e^{3x}$, avec (λ, μ) dans \mathbb{R}^2 . On constate que la fonction constante $x \mapsto 1$ est solution de l'équation (E).

La solution générale \mathcal{S}_\star de (\star) , c'est-à-dire $f^{-1}(1)$, s'écrit : $x \mapsto 1 + \lambda e^{2x} + \mu e^{3x}$, avec (λ, μ) dans \mathbb{R}^2 .

Ici \mathcal{S}_\star est un sous-espace affine de $\mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$, de direction \mathcal{S}_H .

Solution générale d'une récurrence linéaire d'ordre 2

Considérons l'espace $E = \mathbb{R}^{\mathbb{N}}$ des suites à valeurs réelles.

Considérons la récurrence $(\star) : u_{n+2} - 5u_{n+1} + 6u_n = 2$ (équation caractéristique $(r-2)(r-3) = 0$).

Soit $(H) : u_{n+2} - 5u_{n+1} + 6u_n = 0$ la récurrence linéaire homogène associée.

L'application f qui à une suite $(u_n)_{n \geq 0}$ associe la suite $n \mapsto u_{n+2} - 5u_{n+1} + 6u_n$ est dans $\mathcal{L}(E)$.

Chercher le noyau de f , c'est résoudre la récurrence linéaire homogène (H).

Résoudre (\star) , c'est trouver l'image réciproque de la suite $n \mapsto 1$ par l'endomorphisme f .

La solution générale \mathcal{S}_H de (H) s'écrit $n \mapsto u_n = \lambda 2^n + \mu 3^n$, avec (λ, μ) dans \mathbb{R}^2 .

On constate que la suite constante $n \mapsto 1$ est solution de l'équation (\star) .

La solution générale \mathcal{S}_\star de (\star) s'écrit donc $x \mapsto 1 + \lambda 2^n + \mu 3^n$, avec (λ, μ) dans \mathbb{R}^2 .

Ici \mathcal{S}_\star est un sous-espace affine de $\mathbb{R}^{\mathbb{N}}$, de direction \mathcal{S}_H .

Retour à l'interpolation polynomiale

On se donne une famille de n éléments $(x_k, y_k)_{1 \leq k \leq n}$ de \mathbb{R}^2 , les x_k étant distincts deux à deux.

L'application $f: \mathbb{R}[X] \rightarrow \mathbb{R}^n$ qui à un polynôme P associe $(P(x_1), P(x_2), \dots, P(x_n))$ est linéaire.

Le noyau de f est l'ensemble des multiples de $M = \prod_{k=1}^n (X - x_k)$.

Soit $Y = (y_1, y_2, \dots, y_n)$. On a $\{P \in \mathbb{R}[X], \forall k \in \{1, \dots, n\}, P(x_k) = y_k\} = f^{-1}(Y)$.

On sait qu'il existe un unique A de $\mathbb{R}_{n-1}[X]$, tel que $f(A) = Y$.

En d'autres termes, la restriction f_n de f à $\mathbb{R}_{n-1}[X]$ est un isomorphisme de $\mathbb{R}_{n-1}[X]$ sur \mathbb{R}^n .

L'ensemble des polynômes qui « interpolent » la famille des (x_k, y_k) est l'ensemble des $P = A + QM$, avec Q quelconque dans $\mathbb{R}[X]$. On obtient donc un sous-espace affine de $\mathbb{R}[X]$, passant par le polynôme A , et dont la direction est l'ensemble des multiples du polynôme M .

15.3 Paramétrage d'un sous-espace affine

Soit Ω un point de E , et F un sous-espace vectoriel de E , de base u_1, u_2, \dots, u_p .

Soit \mathcal{F} le sous-espace affine de E , passant par Ω et de direction F .

On dit que $\mathcal{R} = \{\Omega, (u_1, u_2, \dots, u_p)\}$ est un *repère cartésien* de \mathcal{F} .

Un point M est dans \mathcal{F} si et seulement si : $(\star) \exists (\lambda_1, \dots, \lambda_p) \in \mathbb{R}^p, M = \Omega + \sum_{k=1}^p \lambda_k u_k$.

Le repère cartésien \mathcal{R} nous a donc permis d'obtenir une « représentation paramétrique » de \mathcal{F} .

Réciproquement, (\star) définit le sous-espace affine passant par Ω et de direction $F = \text{Vect}(u_1, \dots, u_p)$.

On dit souvent, par abus de langage, que le sous-espace affine \mathcal{F} est dirigé par les vecteurs u_1, \dots, u_p .

Cas particulier des droites affines

Une partie \mathcal{D} de E est une droite affine si et seulement si il existe un point Ω de E et un vecteur non nul u de E tels que : $M \in \mathcal{D} \Leftrightarrow (\exists \lambda \in \mathbb{R}, M = \Omega + \lambda u)$.

On peut alors noter $\mathcal{D} = \mathcal{D}(\Omega, u)$ et on dit que u est un *vecteur directeur* de \mathcal{D} .

Si $M = \Omega + \lambda u$, on dit aussi que λ est l'*abscisse* de M sur l'axe $\mathcal{D}(\Omega, u)$.

Si $M = \Omega + \lambda u$ et $N = \Omega + \mu u$, alors la quantité $\overline{MN} = \mu - \lambda$ est appelée *mesure algébrique* du vecteur \overrightarrow{MN} sur l'axe $\mathcal{D}(\Omega, u)$ (cette mesure ne dépend pas du choix du point Ω de \mathcal{D}).

Cas particulier des plans affines

Une partie \mathcal{P} de E est un plan affine si et seulement si il existe un point Ω de E et deux vecteurs indépendants u, v de E tels que : $M \in \mathcal{P} \Leftrightarrow (\exists (\lambda, \mu) \in \mathbb{R}^2, M = \Omega + \lambda u + \mu v)$.

On peut alors noter $\mathcal{P} = \mathcal{P}(\Omega, u, v)$.

Exemple

On se place dans \mathbb{R}^3 muni de sa base canonique.

Soit \mathcal{P} le plan (l'hyperplan) passant par $\Omega = (1, 2, 3)$ et dirigé par les vecteurs $\begin{cases} u = (3, 1, 2) \\ v = (4, 0, 5) \end{cases}$

On a l'équivalence : $M \in \mathcal{P} \Leftrightarrow \exists (\lambda, \mu) \in \mathbb{R}^2, \begin{cases} x = 1 + 3\lambda + 4\mu \\ y = 2 + \lambda \\ z = 3 + 2\lambda + 5\mu \end{cases}$ ce qui constitue un paramétrage de \mathcal{P} .

Réciproquement, on voit bien que ce système caractérise un plan passant par $\Omega = (1, 2, 3)$ (poser $\lambda = \mu = 0$) et dirigé par $u = (3, 1, 2)$ (coefficients de λ) et $v = (4, 0, 5)$ (coefficients de μ).

Demi-droites, demi-plans

Soit A un point de E et u un vecteur non nul.

On dit que $\{M = A + \lambda u, \lambda \in \mathbb{R}^+\}$ est la *demi-droite* d'origine A et de vecteur directeur u .

Soit A un point de E et u, v deux vecteurs indépendants.

Considérons l'ensemble \mathcal{P}^+ défini par $\mathcal{P}^+ = \{M = A + \lambda u + \mu v, \lambda \in \mathbb{R}, \mu \in \mathbb{R}^+\}$.

On dit que \mathcal{P}^+ est le *demi-plan* défini par la droite (A, u) et le vecteur v .

15.4 Parallélisme et intersection

15.4.1 Parallélisme de sous-espaces affines

Définition 15.4.1 (sous espaces affines parallèles)

Soit \mathcal{F} et \mathcal{G} deux sous-espaces affines de E , de directions respectives F et G .

On dit que \mathcal{F} est *faiblement parallèle* à \mathcal{G} si on a l'inclusion $F \subset G$.

On dit que \mathcal{F} est *parallèle* à \mathcal{G} (ou encore que \mathcal{F} et \mathcal{G} sont parallèles) si on a l'égalité $F = G$.

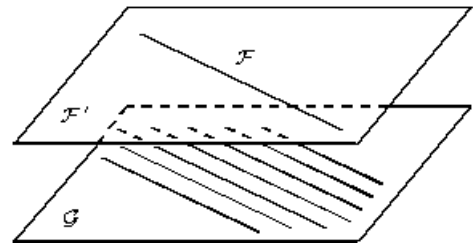
Remarques et exemples

- Si \mathcal{F} et \mathcal{G} ont même dimension, les deux notions précédentes sont équivalentes.
- Un singleton est faiblement parallèle à n'importe quel sous-espace affine.
Une droite peut être faiblement parallèle à un plan, mais l'inverse est impossible.
Si deux sous-espaces affines de dimension finie sont parallèles, ils ont même dimension.
- Deux droites affines sont parallèles si et seulement si elles ont un vecteur directeur commun.
- On pourra noter $\mathcal{F} \parallel \mathcal{G}$ pour exprimer que \mathcal{F} et \mathcal{G} sont parallèles.
On définit ainsi une relation d'équivalence sur l'ensemble des sous-espaces affines de E .
- \mathcal{F} et \mathcal{G} sont parallèles \Leftrightarrow il existe u dans E tel que $t_u(\mathcal{F}) = \mathcal{G}$.
Plus précisément, si $\mathcal{F} \parallel \mathcal{G}$ on a $\mathcal{G} = t_u(\mathcal{F})$ pour tout vecteur $u = \overrightarrow{AB}$ tel que $A \in \mathcal{F}$ et $B \in \mathcal{G}$.
- Soit \mathcal{F} un sous-espace affine de E , et A un point de E .
Par le point A , il passe un unique sous-espace affine parallèle à \mathcal{F} .
- Un sous-espace affine \mathcal{F} est parallèle à sa propre direction F .
Celle-ci est d'ailleurs l'unique sous-espace affine parallèle à \mathcal{F} et passant par O .
- Soit \mathcal{F}, \mathcal{G} deux sous-espaces affines de E . On suppose que \mathcal{F} est parallèle à \mathcal{G} .
Alors il existe un unique sous-espace affine \mathcal{F}' contenant \mathcal{F} et tel que \mathcal{F}' et \mathcal{G} soient parallèles.

Sur cette figure, on voit une droite \mathcal{F} faiblement parallèle au plan \mathcal{G} .

Il existe un plan unique \mathcal{F}' contenant \mathcal{F} et parallèle à \mathcal{G} .

En revanche, \mathcal{G} contient une infinité de droites parallèles à \mathcal{F} .



15.4.2 Intersection de sous-espaces affines

Proposition 15.4.1 (parallélisme et intersection de sous-espaces affines)

Soit \mathcal{F} et \mathcal{G} deux sous-espaces affines de E .

Si \mathcal{F} est faiblement parallèle à \mathcal{G} , alors ou bien $\mathcal{F} \cap \mathcal{G} = \emptyset$ ou bien $\mathcal{F} \subset \mathcal{G}$.

Si \mathcal{F} et \mathcal{G} sont parallèles, alors ils sont disjoints ou confondus.

Proposition 15.4.2 (intersection de sous-espaces affines)

Soit \mathcal{F}, \mathcal{G} deux sous-espaces affines de E , de directions respectives F et G .

L'intersection $\mathcal{F} \cap \mathcal{G}$, si elle n'est pas vide, est un sous-espace affine de direction $F \cap G$.

Si $E = F + G$, alors on est certain que l'intersection $\mathcal{F} \cap \mathcal{G}$ n'est pas vide.

Si $E = F \oplus G$, alors l'intersection $\mathcal{F} \cap \mathcal{G}$ se réduit à un singleton.

Cas particuliers

- Soit \mathcal{H} un hyperplan affine de E , et soit \mathcal{D} une droite affine, non faiblement parallèle à \mathcal{H} .
Alors la droite \mathcal{D} « coupe » l'hyperplan \mathcal{H} en un point et un seul.

- En particulier, si $\dim E = 2$, deux droites non parallèles ont un unique point en commun.
Si $\dim E = 3$, une droite \mathcal{D} non parallèle à un plan \mathcal{P} rencontre ce plan en un point unique.
- Soit \mathcal{P}_1 et \mathcal{P}_2 deux plans affines de E , avec $\dim E = 3$.
Si \mathcal{P}_1 et \mathcal{P}_2 ne sont pas parallèles, alors leur intersection est une droite.
- Si $\dim E \geq 4$, il est possible que deux plans \mathcal{P}_1 et \mathcal{P}_2 de E soient disjoints.

Droites coplanaires

Supposons $\dim E \geq 3$, et soient \mathcal{D}_1 et \mathcal{D}_2 deux droites affines de E .

- On dit que \mathcal{D}_1 et \mathcal{D}_2 sont coplanaires si elles sont incluses dans un même plan affine \mathcal{P} .
Cela équivaut à dire que \mathcal{D}_1 et \mathcal{D}_2 sont parallèles ou concourantes.
Si $\mathcal{D}_1 = (A, u)$ et $\mathcal{D}_2 = (B, v)$, cela équivaut à dire que $\text{rg}(\overrightarrow{AB}, u, v) \leq 2$.
Dans ce cas, et si $\mathcal{D}_1 \neq \mathcal{D}_2$, le plan \mathcal{P} est défini de manière unique par \mathcal{D}_1 et \mathcal{D}_2 .
- Si les droites \mathcal{D}_1 et \mathcal{D}_2 ne sont pas coplanaires, leur intersection est vide.

15.5 Équation cartésienne d'un hyperplan

Proposition 15.5.1 (caractérisation et équation d'un hyperplan affine)

Soit \mathcal{H} une partie de l'espace vectoriel E . Les conditions suivantes sont équivalentes :

- L'ensemble \mathcal{H} est un hyperplan affine de E .
- Il existe une forme linéaire f non nulle et un scalaire α tels que : $M \in \mathcal{H} \Leftrightarrow f(M) = \alpha$.
Une telle caractérisation est appelée une équation de l'hyperplan \mathcal{H} .

Remarque : l'équation $f(M) = \alpha$ de l'hyperplan \mathcal{H} est unique à un facteur multiplicatif non nul près.

Avec les notations précédentes :

Les équations $f(M) = \beta$ sont celles des hyperplans parallèles à \mathcal{H} .

Par exemple $f(M) = f(M_0)$ est l'équation de l'hyperplan parallèle à \mathcal{H} et passant par M_0 .

L'équation $f(M) = 0$ est celle de la direction H de \mathcal{H} .

15.5.1 Équation cartésienne d'un hyperplan de \mathbb{R}^n

On se place dans \mathbb{R}^n , muni de son repère canonique.

Soit (x_1, x_2, \dots, x_n) les coordonnées d'un point M quelconque de \mathbb{R}^n dans ce repère.

L'équation d'un hyperplan \mathcal{H} s'écrit : $\sum_{k=1}^n a_k x_k = \alpha$, où α est un scalaire et où les a_k sont non tous nuls.

Quand on fait varier β , on obtient les équations $\sum_{k=1}^n a_k x_k = \beta$ des hyperplans affines parallèles à \mathcal{H} .

L'équation $\sum_{k=1}^n a_k x_k = 0$ est celle de l'hyperplan vectoriel H direction de \mathcal{H} .

Soit Ω un point de \mathbb{R}^n , de coordonnées $(\omega_1, \omega_2, \dots, \omega_n)$.

L'hyperplan de direction H et passant par Ω a pour équation : $\sum_{k=1}^n a_k (x_k - \omega_k) = 0$.

Parallélisme de deux hyperplans de \mathbb{R}^n

Considérons les équations
$$\begin{cases} a_1x_1 + a_2x_2 + \dots + a_nx_n = \alpha \\ b_1x_1 + b_2x_2 + \dots + b_nx_n = \beta \end{cases}$$
 de deux hyperplans \mathcal{H}_1 et \mathcal{H}_2 de \mathbb{R}^n .

Les deux hyperplans sont parallèles si et seulement si on a les égalités : $\frac{a_1}{b_1} = \frac{a_2}{b_2} = \dots = \frac{a_n}{b_n}$.

Ils sont confondus si et seulement si : $\frac{a_1}{b_1} = \frac{a_2}{b_2} = \dots = \frac{a_n}{b_n} = \frac{\alpha}{\beta}$

(par convention, si un dénominateur est nul, le numérateur correspondant l'est également).

15.5.2 Droites affines de \mathbb{R}^2

On se place dans \mathbb{R}^2 , muni de son repère canonique d'origine O et de base $\vec{e}_1 = (1, 0)$, $\vec{e}_2 = (0, 1)$.

Une droite affine \mathcal{D} de \mathbb{R}^2 admet une équation du type $ax + by = c$, avec $(a, b) \neq (0, 0)$.

Avec ces notations, un vecteur directeur de \mathcal{D} est $\vec{u} = (b, -a)$.

– les droites parallèles à l'axe (Ω, \vec{e}_2) ont une équation du type $x = \alpha$.

– les droites parallèles à l'axe (Ω, \vec{e}_1) ont une équation du type $y = \beta$.

– si \mathcal{D} est non parallèle à (Ω, \vec{e}_2) son équation s'écrit $y = \alpha x + \beta$ ($\alpha =$ coefficient directeur de \mathcal{D}).

Les droites $\begin{cases} \mathcal{D}: ax + by = c \\ \mathcal{D}': a'x + b'y = c' \end{cases}$ sont parallèles $\Leftrightarrow \frac{a}{a'} = \frac{b}{b'}$. Elles sont confondues $\Leftrightarrow \frac{a}{a'} = \frac{b}{b'} = \frac{c}{c'}$.

Si $\begin{cases} \mathcal{D}: ax + by = c \\ \mathcal{D}': a'x + b'y = c' \end{cases}$ sont sécantes, alors $\mathcal{D} \cap \mathcal{D}' = \{\Omega\}$, avec $x_\Omega = \frac{cb' - bc'}{ab' - ba'}$ et $y_\Omega = \frac{ac' - ca'}{ab' - ba'}$.

La droite passant par $A(x_0, y_0)$ et dirigée par $u(\alpha, \beta)$ a pour équation : $(x - x_0)\beta - (y - y_0)\alpha = 0$.

La droite passant par $A(a, 0)$ et $B(0, b)$, avec $a \neq 0$ et $b \neq 0$, a pour équation $\frac{x}{a} + \frac{y}{b} = 1$.

Complément : faisceaux de droites dans le plan

Soit \mathcal{D} et \mathcal{D}' deux droites distinctes du plan, d'équations respectives (E) et (E') .

On forme une famille d'équations de droites en écrivant $\lambda(E) + \mu(E')$, avec $(\lambda, \mu) \neq (0, 0)$.

– si \mathcal{D} et \mathcal{D}' sont parallèles, on obtient ainsi toutes les droites qui leur sont parallèles.

– sinon, on obtient toutes les droites passant par le point d'intersection I de \mathcal{D} et \mathcal{D}' .

Dans tous les cas, on dit que l'ensemble obtenu est le *faisceau de droites* engendré par $\mathcal{D}, \mathcal{D}'$.

Les équations $(E) + \mu(E')$, avec μ dans \mathbb{R} donnent toutes les droites du faisceau sauf la droite \mathcal{D}' .

Trois droites sont parallèles ou concourantes si et seulement si elles appartiennent à un même faisceau, c'est-à-dire si et seulement si leurs équations sont « liées ».

15.5.3 Plans affines de \mathbb{R}^3

On se place dans \mathbb{R}^3 , avec repère canonique d'origine O et de base
$$\begin{cases} \vec{e}_1 = (1, 0, 0) \\ \vec{e}_2 = (0, 1, 0) \\ \vec{e}_3 = (0, 0, 1) \end{cases}$$

Soit \mathcal{P} une partie de \mathbb{R}^3 .

L'ensemble \mathcal{P} est un plan affine s'il a une équation du type $ax + by + cz = d$, avec $(a, b, c) \neq (0, 0, 0)$.

Si $a \neq 0$, les vecteurs $u = (b, -a, 0)$ et $v = (c, 0, -a)$ forment une base de la direction de \mathcal{P} .

Cas particuliers

Dans certains cas, l'équation du plan \mathcal{P} prend une forme plus simple :

- le plan \mathcal{P} est parallèle au plan (O, e_2, e_3) (appelé « plan Oyz ») \Leftrightarrow son équation s'écrit $x = \alpha$.
- le plan \mathcal{P} est parallèle au plan (O, e_1, e_3) (appelé « plan Oxz ») \Leftrightarrow son équation s'écrit $y = \beta$.
- le plan \mathcal{P} est parallèle au plan (O, e_1, e_2) (appelé « plan Oxy ») \Leftrightarrow son équation s'écrit $z = \gamma$.
- la droite (Ω, e_1) (dite « droite Ox ») est parallèle à \mathcal{P} \Leftrightarrow l'équation de \mathcal{P} s'écrit $by + cz = d$.
- la droite (Ω, e_2) (dite « droite Oy ») est parallèle à \mathcal{P} \Leftrightarrow l'équation de \mathcal{P} s'écrit $ax + cz = d$.
- la droite (Ω, e_3) (dite « droite Oz ») est parallèle à \mathcal{P} \Leftrightarrow l'équation de \mathcal{P} s'écrit $ax + by = d$.
- L'équation d'un plan \mathcal{P} non parallèle à Oxy peut s'écrire $z = \alpha x + \beta y + \gamma$.
- On considère les points $A(a, 0, 0)$, $B(0, b, 0)$, $C(0, 0, c)$, avec $a \neq 0$, $b \neq 0$, $c \neq 0$.

Le plan passant par A, B, C a pour équation $\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1$.

Parallélisme de deux plans de \mathbb{R}^3

Considérons deux plans d'équations respectives $ax + by + cz = d$ pour \mathcal{P} , et $a'x + b'y + c'z = d'$ pour \mathcal{P}' .

Les plans \mathcal{P} et \mathcal{P}' sont parallèles si et seulement si on a les égalités $\frac{a}{a'} = \frac{b}{b'} = \frac{c}{c'}$.

Ils sont confondus si et seulement si $\frac{a}{a'} = \frac{b}{b'} = \frac{c}{c'} = \frac{d}{d'}$.

Si les plans $\begin{cases} \mathcal{P}: z = \alpha x + \beta y + \gamma \\ \mathcal{P}': z = \alpha' x + \beta' y + \gamma' \end{cases}$ ne sont pas parallèles à Oxy , alors : $\mathcal{P} \parallel \mathcal{P}' \Leftrightarrow (\alpha = \alpha' \text{ et } \beta = \beta')$.

Si $\begin{cases} \mathcal{P}: ax + by + cz = d \\ \mathcal{P}': a'x + b'y + c'z = d' \end{cases}$ sont non parallèles, $\mathcal{D} = \mathcal{P} \cap \mathcal{P}'$ est dirigée par $\begin{pmatrix} a \\ b \\ c \end{pmatrix} \wedge \begin{pmatrix} a' \\ b' \\ c' \end{pmatrix} = \begin{pmatrix} bc' - cb' \\ ca' - ac' \\ ab' - ba' \end{pmatrix}$

Exemple de recherche d'équation cartésienne d'un plan de \mathbb{R}^3

On se place ici dans \mathbb{R}^3 , muni de son repère canonique.

Soit \mathcal{P} le plan (l'hyperplan) passant par $A = (1, 2, 3)$ et dirigé par $\begin{cases} u = (3, 1, 2) \\ v = (4, 0, 5) \end{cases}$

On forme le vecteur $w = u \wedge v = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} \wedge \begin{pmatrix} 4 \\ 0 \\ 5 \end{pmatrix} = \begin{pmatrix} 5 \\ -7 \\ -4 \end{pmatrix}$

L'équation de \mathcal{P} s'écrit : $5(x - 1) - 7(y - 2) - 4(z - 3) = 0$ c'est-à-dire $5x - 7y - 4z = -21$.

On peut aussi trouver l'équation cartésienne de \mathcal{P} à partir d'une représentation paramétrique.

En effet, on a l'équivalence : $M(x, y, z) \in \mathcal{P} \Leftrightarrow \exists (\lambda, \mu) \in \mathbb{R}^2, \begin{cases} x = 1 + 3\lambda + 4\mu \\ y = 2 + \lambda \\ z = 3 + 2\lambda + 5\mu \end{cases}$

On résout ce système par rapport aux inconnues (λ, μ) , et l'équation cartésienne cherchée est la condition sur les paramètres x, y, z pour que ce système admette une solution (λ, μ) .

On trouve successivement :

$$\exists (\lambda, \mu) \in \mathbb{R}^2, \begin{cases} x = 1 + 3\lambda + 4\mu \\ y = 2 + \lambda \\ z = 3 + 2\lambda + 5\mu \end{cases} \Leftrightarrow \exists (\lambda, \mu) \in \mathbb{R}^2, \begin{cases} \lambda = y - 2 \\ x = 1 + 3(y - 2) + 4\mu \\ z = 3 + 2(y - 2) + 5\mu \end{cases}$$

$$\exists (\lambda, \mu) \in \mathbb{R}^2, \begin{cases} \lambda = y - 2 \\ 4\mu = x - 3y + 5 \\ 5\mu = 1 - 2y + z \end{cases} \Leftrightarrow \begin{cases} 5(x - 3y + 5) = 4(1 - 2y + z) \\ 5x - 7y - 4z = -21 \end{cases}$$

De l'équation cartésienne à une représentation paramétrique

Réciproquement, on passe facilement d'une équation cartésienne de plan à un paramétrage.

Par exemple, soit \mathcal{P} le plan (donc l'hyperplan) d'équation $2x + 3y - 5z = 8$.

Une base de la direction vectorielle P de \mathcal{P} (d'équation $2x + 3y - 5z = 0$) est $\begin{cases} u = (3, -2, 0) \\ v = (5, 0, 2) \end{cases}$

Le point $\Omega(4, 0, 0)$ appartient à \mathcal{P} .

Une représentation paramétrique de \mathcal{P} est donc $\begin{cases} x = 4 + 3\lambda + 5\mu \\ y = -2\lambda \\ z = 2\mu \end{cases}$ avec (λ, μ) dans \mathbb{R}^2 .

Exemple de recherche d'intersection d'une droite et d'un plan dans \mathbb{R}^3

On se place dans \mathbb{R}^3 muni de son repère canonique.

Soit \mathcal{D} la droite passant par le point $A(1, 0, 4)$ et dirigée par le vecteur $u = (1, -1, -1)$.

Soit \mathcal{P} le plan d'équation $2x + 3y - 2z = 5$.

Le vecteur u n'est pas dans la direction P de \mathcal{P} (qui a pour équation $2x + y - 2z = 0$.)

La droite \mathcal{D} rencontre donc \mathcal{P} en un point $M_\lambda = A + \lambda u$ unique.

On reporte les coordonnées $x_\lambda = 1 + \lambda, y_\lambda = -\lambda, z_\lambda = 4 - \lambda$ de M_λ dans $2x + 3y - 2z = 5$.

On trouve $2(1 + \lambda) - 3\lambda - 2(4 - \lambda) = 5$ c'est-à-dire : $\lambda = 11$.

Ainsi le point d'intersection de \mathcal{D} et de \mathcal{P} est $M = A + 11u = (12, -11, -7)$.

Complément : faisceaux de plans de \mathbb{R}^3

Soit \mathcal{P} et \mathcal{P}' deux plans distincts de \mathbb{R}^3 , d'équations respectives (E) et (E') .

On forme une famille d'équations de plans en écrivant $\lambda(E) + \mu(E')$, avec $(\lambda, \mu) \neq (0, 0)$.

– si \mathcal{P} et \mathcal{P}' sont parallèles, on obtient ainsi tous les plans qui leur sont parallèles.

– sinon, on obtient tous les plans contenant la droite $\mathcal{D} = \mathcal{P} \cap \mathcal{P}'$.

Dans tous les cas, on dit que l'ensemble obtenu est le *faisceau de plans* engendré par $\mathcal{P}, \mathcal{P}'$.

Les équations $(E) + \mu(E')$, avec μ dans \mathbb{R} , donnent tous les plans du faisceau sauf \mathcal{P}' .

On considère trois plans, d'équations $ax + by + cz = d, a'x + b'y + c'z = d'$ et $a''x + b''y + c''z = d''$.

Ces trois plans sont parallèles ou passent par une même droite si et seulement si ils appartiennent à un même faisceau, c'est-à-dire si et seulement si leurs équations sont « liées ».

15.6 Systèmes d'équations d'un sous-espace affine

Pour simplifier, on se place ici dans \mathbb{R}^n , muni de son repère canonique.

15.6.1 Intersection de p hyperplans affines de \mathbb{R}^n

Proposition 15.6.1 (intersection de p hyperplans affines)

Soit $(H_i)_{1 \leq i \leq p}$ une famille de p hyperplans affines de \mathbb{R}^n , avec $1 \leq p \leq n$. Soit $\mathcal{F} = \bigcap_{i=1}^p H_i$.

Si \mathcal{F} n'est pas vide, alors \mathcal{F} est un sous-espace affine de \mathbb{R}^n , de dimension supérieure ou égale à $n - p$.

Considérons par exemple trois plans affines $\mathcal{P}_1, \mathcal{P}_2$ et \mathcal{P}_3 de \mathbb{R}^3 .

La proposition précédente nous dit que si $\mathcal{P}_1 \cap \mathcal{P}_2 \cap \mathcal{P}_3$ n'est pas vide, alors cette intersection est un sous-espace affine de dimension supérieure ou égale à $n - p = 0$, ce qui est un renseignement assez vague, mais on s'aperçoit rapidement que tous les cas sont possibles :

– L'intersection de $\mathcal{P}_1, \mathcal{P}_2$ et \mathcal{P}_3 peut être vide.

C'est le cas par exemple si \mathcal{P}_1 et \mathcal{P}_2 sont parallèles distincts (car $\mathcal{P}_1 \cap \mathcal{P}_2 = \emptyset$ donc $\mathcal{P}_1 \cap \mathcal{P}_2 \cap \mathcal{P}_3 = \emptyset$).

C'est aussi le cas si \mathcal{P}_1 et \mathcal{P}_2 ne sont pas parallèles (leur intersection est alors une droite \mathcal{D}), mais si la droite \mathcal{D} est faiblement parallèle à \mathcal{P}_3 sans être incluse dans \mathcal{P}_3 .

– L'intersection de $\mathcal{P}_1, \mathcal{P}_2$ et \mathcal{P}_3 peut être réduite à un point Ω .

C'est le cas si \mathcal{P}_1 et \mathcal{P}_2 ne sont pas parallèles (leur intersection est alors une droite \mathcal{D}) et si \mathcal{D} n'est pas faiblement parallèle à \mathcal{P}_3 (donc si elle coupe \mathcal{P}_3 en un point Ω).

– L'intersection $\mathcal{P}_1 \cap \mathcal{P}_2 \cap \mathcal{P}_3$ est une droite \mathcal{D} si $\mathcal{P}_1, \mathcal{P}_2$ et \mathcal{P}_3 appartiennent au « faisceau » dirigé par la droite \mathcal{D} (sans être tous les trois confondus).

– L'intersection de $\mathcal{P}_1, \mathcal{P}_2$ et \mathcal{P}_3 peut être un plan (si $\mathcal{P}_1 = \mathcal{P}_2 = \mathcal{P}_3$, bien sûr).

Un exemple

On considère les quatre hyperplans $\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3$, et \mathcal{H}_4 de \mathbb{R}^5 , dont les équations respectives sont :

$$\begin{aligned} \mathcal{H}_1: & x + 3y + 5z - 2t - 7u = 3 \\ \mathcal{H}_2: & 3x + y + z - 2t - u = 1 \\ \mathcal{H}_3: & 2x - y - 3z + 7t + 5u = 2 \\ \mathcal{H}_4: & 3x - 2y - 5z + 7t + 8u = \lambda \end{aligned} \quad (\text{où } \lambda \text{ désigne un paramètre réel})$$

Soit $\mathcal{F} = \mathcal{H}_1 \cap \mathcal{H}_2 \cap \mathcal{H}_3 \cap \mathcal{H}_4$ (ici il y a $p = 4$ hyperplans, et on est en dimension $n = 5$).

La proposition nous dit que si $\mathcal{F} \neq \emptyset$, alors c'est un sous-espace affine de \mathbb{R}^5 , de dimension $\geq 5 - 4 = 1$.

Pour déterminer \mathcal{F} , on résout le système (S) formé par les équations des quatre hyperplans.

$$\text{On montre que } (S) \Leftrightarrow \begin{cases} x = -1 + 3t \\ y = 8 - 17t - u \\ z = -4 + 10t + 2u \\ \lambda = 1 \end{cases} \Leftrightarrow (\lambda = 1) \text{ et } \begin{pmatrix} x \\ y \\ z \\ t \\ u \end{pmatrix} = \begin{pmatrix} -1 \\ 8 \\ -4 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} 3 \\ -17 \\ 10 \\ 1 \\ 0 \end{pmatrix} + u \begin{pmatrix} 0 \\ -1 \\ 2 \\ 0 \\ 1 \end{pmatrix}$$

Ici λ est ici un paramètre, dont la valeur est supposée fixée (mais inconnue).

Le résultat ci-dessus nous dit que si $\lambda \neq 1$, alors le système (S) n'a pas de solutions.

Autrement dit, si $\lambda \neq 1$, l'intersection $\mathcal{F} = \mathcal{H}_1 \cap \mathcal{H}_2 \cap \mathcal{H}_3 \cap \mathcal{H}_4$ est vide.

Mais si $\lambda = 1$, alors \mathcal{F} est le plan passant par $\Omega = (-1, 8, -4, 0, 0)$ et dirigé par $\begin{cases} (3, -17, 10, 1, 0) \\ (0, -1, 2, 0, 1) \end{cases}$

15.6.2 Système d'équations d'un sous-espace de \mathbb{R}^n

Proposition 15.6.2 (système d'équations d'un sous-espace affine)

Soit \mathcal{F} un sous-espace affine de \mathbb{R}^n , de dimension $n - p$, avec $1 \leq p \leq n$.

Alors \mathcal{F} peut s'écrire comme l'intersection de p hyperplans affines.

Le système formé par les équations de ces hyperplans est appelé un système d'équations de \mathcal{F} .

Sauf si F est lui-même un hyperplan de \mathbb{R}^n (c'est-à-dire si $p = 1$), il n'y a pas (loin de là) unicité de la famille d'hyperplans d'intersection F (donc unicité du système d'équations).

Un exemple

On se place ici dans \mathbb{R}^5 , muni de son repère canonique.

On note (x, y, z, t, u) les coordonnées d'un point M de E .

Soit \mathcal{F} le sous-espace passant par $A(1, 0, 1, 2, -1)$ et dirigé par les vecteurs

$$\begin{cases} w_1 = (1, -1, 2, 1, 1) \\ w_2 = (1, 1, 3, 0, 1) \\ w_3 = (1, 2, 0, 1, 1) \end{cases}$$

Dire que M est dans \mathcal{F} , c'est dire que \overrightarrow{AM} est dans $\text{Vect}(w_1, w_2, w_3)$.

On utilise une méthode la méthode du pivot :

$$\begin{array}{l} \left(\begin{array}{ccc|c} w_1 & w_2 & w_3 & \overrightarrow{AM} \\ 1 & 1 & 1 & x-1 \\ -1 & 1 & 2 & y \\ 2 & 3 & 0 & z-1 \\ 1 & 0 & 1 & t-2 \\ 1 & 1 & 1 & u+1 \end{array} \right) \begin{array}{l} \implies \\ L_2 \leftarrow L_2 + L_1 \\ L_3 \leftarrow L_3 - 2L_1 \\ L_4 \leftarrow L_4 - L_1 \\ L_5 \leftarrow L_5 - L_1 \end{array} \left(\begin{array}{cccc} 1 & 1 & 1 & x-1 \\ 0 & 2 & 3 & x+y-1 \\ 0 & 1 & -2 & -2x+z+1 \\ 0 & -1 & 0 & -x+t-1 \\ 0 & 0 & 0 & -x+u+2 \end{array} \right) \begin{array}{l} \implies \\ L_3 \leftarrow 2L_3 - L_2 \\ L_4 \leftarrow 2L_4 + L_2 \end{array} \\ \\ \left(\begin{array}{cccc} 1 & 1 & 1 & x-1 \\ 0 & 2 & 3 & x+y-1 \\ 0 & 0 & -7 & -5x-y+2z+3 \\ 0 & 0 & 3 & -x+y+2t-3 \\ 0 & 0 & 0 & u-x+2 \end{array} \right) L_4 \leftarrow 7L_4 + 3L_3 \left(\begin{array}{ccc|c} w'_1 & w'_2 & w'_3 & w'_4 \\ 1 & 1 & 1 & x-1 \\ 0 & 2 & 3 & x+y-1 \\ 0 & 0 & -7 & -5x-y+2z-1 \\ 0 & 0 & 0 & -22x+4y+6z+14t-12 \\ 0 & 0 & 0 & u-x+2 \end{array} \right) \end{array}$$

La méthode du pivot ne modifie pas le rang de la famille des vecteurs à laquelle est s'applique.

Ici on a transformé les vecteurs $w_1, w_2, w_3, \overrightarrow{AM}$ en les vecteurs w'_1, w'_2, w'_3, w'_4 .

Il est clair que w'_1, w'_2, w'_3 sont libres (coefficients « échelonnés »).

Cela confirme que w_1, w_2, w_3 sont libres, donc que \mathcal{F} est bien de dimension 3.

De même, dire que \overrightarrow{AM} est lié à w_1, w_2, w_3 , c'est dire que w'_4 est lié à w'_1, w'_2, w'_3 .

Mais le résultat précédent montre que cela équivaut à
$$\begin{cases} 11x - 2y - 3z - 7t = -6 \\ x - u = 2 \end{cases}$$

On a donc obtenu un système d'équations de \mathcal{F} , qui est ici l'intersection de deux hyperplans.

Droites dans \mathbb{R}^3

Soit \mathcal{D} une droite affine de \mathbb{R}^3 .

Un système d'équations de \mathcal{D} est
$$\begin{cases} ax + by + cz = d \\ a'x + b'y + c'z = d' \end{cases}$$
 où (a, b, c) et (a', b', c') sont non proportionnels.

Ainsi définie, $\mathcal{D} = \mathcal{P} \cap \mathcal{P}'$ où
$$\begin{cases} \mathcal{P} \text{ est le plan d'équation } ax + by + cz = d \\ \mathcal{P}' \text{ est le plan d'équation } a'x + b'y + c'z = d' \end{cases}$$

Un vecteur directeur de la droite \mathcal{D} est
$$u = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \wedge \begin{pmatrix} a' \\ b' \\ c' \end{pmatrix} = \begin{pmatrix} bc' - cb' \\ ca' - ac' \\ ab' - ba' \end{pmatrix}.$$

Le système
$$\begin{cases} ax + by + cz = 0 \\ a'x + b'y + c'z = 0 \end{cases}$$
 caractérise la direction D de \mathcal{D} .

Les droites parallèles à \mathcal{D} ont pour système d'équations :
$$\begin{cases} ax + by + cz = \alpha \\ a'x + b'y + c'z = \beta \end{cases} \quad \text{avec } (\alpha, \beta) \text{ dans } \mathbb{R}^2.$$

Les équations $\lambda(ax + by + cz - d) + \mu(a'x + b'y + c'z - d') = 0$ (avec $(\lambda, \mu) \neq (0, 0)$) sont celles des plans contenant la droite \mathcal{D} (c'est le *faisceau* de plans dirigé par \mathcal{D}).

15.7 Barycentres et repères affines

15.7.1 Barycentres

E désigne un espace vectoriel sur \mathbb{R} .

Définition 15.7.1 (points pondérés)

On appelle *point pondéré* le couple (A, λ) formé d'un point A de E et d'un réel λ .

On dit que le réel λ est le *poids* du point pondéré (A, λ) .

Soit $(A_1, \lambda_1), (A_2, \lambda_2), \dots, (A_p, \lambda_p)$ une famille de p points pondérés.

La quantité $m = \sum_{k=1}^p \lambda_k$ est appelée *poids total* de ce système de points.

Définition 15.7.2 (barycentre d'une famille de points pondérés)

Soit $(A_1, \lambda_1), (A_2, \lambda_2), \dots, (A_p, \lambda_p)$ une famille de p points pondérés.

On suppose que le poids total m de cette famille est non nul.

Il existe alors un unique point G de E tel que : $\sum_{k=1}^p \lambda_k \overrightarrow{GA_k} = \vec{0}$. Il est défini par : $G = \frac{1}{m} \sum_{k=1}^p \lambda_k A_k$.

On dit que G est le *barycentre* des points pondérées (A_k, λ_k) .

On ne modifie pas le barycentre G en multipliant les poids λ_k par un même coefficient non nul μ .

En particulier, en choisissant $\mu = \frac{1}{m}$, on peut toujours se ramener à $\sum_{k=1}^p \lambda_k = 1$.

Dans ce cas le barycentre G est défini par $G = \sum_{k=1}^p \lambda_k A_k$.

Isobarycentre

Si les poids λ_k sont constants, on obtient l'*isobarycentre* (ou *éuibarycentre*) de A_1, A_2, \dots, A_p .

On peut bien sûr choisir $\lambda = \frac{1}{p}$. L'isobarycentre G est alors défini par l'égalité $G = \frac{1}{p} \sum_{k=1}^p A_k$.

L'isobarycentre de A, B est le milieu $G = \frac{1}{2}(A + B)$ du segment $[A, B]$.

L'isobarycentre de A, B, C est le centre de gravité $G = \frac{1}{3}(A + B + C)$ du triangle ABC .

Parallélogramme

Soit A, B, C, D quatre points de E . On a les équivalences suivantes :

$$\overrightarrow{AB} = \overrightarrow{DC} \Leftrightarrow \overrightarrow{AD} = \overrightarrow{BC} \Leftrightarrow \exists u \in E, \begin{cases} t_u(A) = D \\ t_u(B) = C \end{cases} \Leftrightarrow [A, C] \text{ et } [B, D] \text{ ont même milieu}$$

On exprime ces conditions en disant que A, B, C, D (dans cet ordre) forment un *parallélogramme*. Il en est alors de même pour les quadruplets B, C, D, A , ou C, D, A, B .

Le milieu commun I de $[A, C]$ et $[B, D]$ vérifie $I = \frac{1}{2}(A + C) = \frac{1}{2}(B + D) = \frac{1}{4}(A + B + C + D)$.

Autrement dit, I est l'isobarycentre des quatre points A, B, C, D .

Associativité du barycentre

– Lorsqu'on cherche un barycentre G , on peut remplacer une sous-famille de points (de poids total m_k non nul) par le barycentre G_k de cette sous-famille affecté lui-même du poids m_k .

– L'isobarycentre G des points A, B, C est aussi le barycentre de $(A, 1), (I, 2)$, où $I = \frac{1}{2}(B + C)$.

Autrement dit $\overrightarrow{AG} = \frac{2}{3}\overrightarrow{AI}$. Les médianes d'un triangle sont donc concourantes en son centre de gravité G , qui est au deux-tiers de chaque médiane en partant du sommet.

– De même, on se donne quatre points A, B, C, D non coplanaires : ils forment un vrai *tétraèdre*. Soit G l'isobarycentre de A, B, C, D , et soit I celui de B, C, D .

Par associativité, G est le barycentre des points pondérés $(A, 1)$ et $(I, 3)$.

Autrement dit $\overrightarrow{AG} = \frac{3}{4}\overrightarrow{AI}$. Ainsi, dans un tétraèdre, les segments joignant un sommet au centre de gravité de la face opposée sont concourants en le centre de gravité du tétraèdre. Celui-ci est aux $\frac{3}{4}$ de chacun de ces segments (en partant du sommet du tétraèdre.)

15.7.2 Barycentres et sous-espaces affines

Proposition 15.7.1 (stabilité d'un sous-espace affine par barycentration)

Soit \mathcal{F} un sous-espace affine de E . Soit $(A_1, \lambda_1), \dots, (A_p, \lambda_p)$ des points pondérés de \mathcal{F} .

On suppose que $m = \sum_{k=1}^p \lambda_k \neq 0$. Soit G le barycentre des (A_k, λ_k) .

Alors le point G appartient au sous-espace affine \mathcal{F} .

On peut se demander réciproquement si tout point de \mathcal{F} peut être considéré comme le barycentre d'une famille *fixée* de points de \mathcal{F} .

La réponse fait l'objet de la définition et de la proposition suivantes.

Définition 15.7.3 (repère affine d'un sous-espace affine de dimension p)

Soit \mathcal{F} un sous-espace affine de dimension $p \geq 1$ de E .

Soit A_0, A_1, \dots, A_p une famille de $p + 1$ points de \mathcal{F} .

On dit que cette famille est un *repère affine* de \mathcal{F} si les vecteurs $\overrightarrow{A_0A_1}, \overrightarrow{A_0A_2}, \dots, \overrightarrow{A_0A_p}$ sont libres, c'est-à-dire s'ils forment une base de la direction F de \mathcal{F} .

Proposition 15.7.2 (coordonnées dans un repère affine)

Soit \mathcal{F} un sous-espace affine de dimension $p \geq 1$ de E . Soit A_0, A_1, \dots, A_p un repère affine de \mathcal{F} .

Alors tout point M de \mathcal{F} est, d'une façon unique, le barycentre des (A_k, λ_k) , avec $\sum_{k=0}^p \lambda_k = 1$.

Les λ_k sont appelés *coordonnées (barycentriques) de M dans ce repère affine*.

Repères affines d'une droite ou d'un plan

- Deux points distincts d'une droite en forment un repère affine.

Si A, B sont distincts, la droite (AB) est l'ensemble des $M = \alpha A + \beta B$ avec $\alpha + \beta = 1$.

- Trois points non alignés d'un plan en forment un repère affine.

Soient A, B, C trois points non alignés de E .

Le plan (ABC) est l'ensemble des $M = \alpha A + \beta B + \gamma C$ avec $\alpha + \beta + \gamma = 1$.

Chapitre 16

Calcul matriciel

Sommaire

16.1	Espaces de matrices	349
16.1.1	Matrices à n lignes et p colonnes	349
16.1.2	Structure d'espace vectoriel de $\mathcal{M}_{n,p}(\mathbb{K})$	351
16.1.3	Matrices carrées	352
16.2	Produit matriciel	354
16.2.1	Produit des matrices	354
16.2.2	La non-commutativité du produit	355
16.2.3	Propriétés du produit matriciel	356
16.2.4	Une justification du produit matriciel	356
16.2.5	Produits, lignes et colonnes	358
16.2.6	Produits de matrices de la base canonique	359
16.3	Calculs sur les matrices carrées	360
16.3.1	L'anneau $(\mathcal{M}_n(\mathbb{K}), +, \times)$	360
16.3.2	La formule du binôme	361
16.3.3	Matrices nilpotentes ou diviseurs de zéro	362
16.3.4	Calcul des puissances d'une matrice carrée	363
16.3.5	Matrices inversibles	365
16.3.6	Calcul de l'inverse d'une matrice inversible	369
16.4	Transposition	370
16.4.1	Propriétés de la transposition	370
16.4.2	Matrices symétriques ou antisymétriques	371
16.5	Calculs par blocs	372
16.5.1	Décompositions en blocs	372
16.5.2	Opérations sur les décompositions en blocs	374

16.1 Espaces de matrices

16.1.1 Matrices à n lignes et p colonnes

Définition 16.1.1 (matrices à n lignes, à p colonnes, et à coefficients dans \mathbb{K})

Une *matrice* A de type (n, p) (avec n, p dans \mathbb{N}^*) est une application de $\{1, \dots, n\} \times \{1, \dots, p\}$ dans \mathbb{K} .

On note souvent $a_{i,j}$ (« coefficient d'indice i, j ») l'image du couple (i, j) par l'application A .

On écrit alors $A = (a_{i,j})_{\substack{i=1..n \\ j=1..p}}$, ou plus simplement $A = (a_{ij})$.

On note $\mathcal{M}_{n,p}(\mathbb{K})$ l'ensemble des matrices de type (n, p) à coefficients dans \mathbb{K} .

Pour décrire un élément A de $\mathcal{M}_{n,p}(\mathbb{K})$, on dispose les coefficients dans un tableau à n lignes et p colonnes, le coefficient $a_{i,j}$ venant se placer à l'intersection de la i -ème ligne et de la j -ème colonne.

Par exemple, la matrice A de $\mathcal{M}_{2,3}(\mathbb{K})$ définie par : $\begin{cases} a_{1,1} = 5, & a_{1,2} = 3, & a_{1,3} = 2 \\ a_{2,1} = 7, & a_{2,2} = 4, & a_{2,3} = 1 \end{cases}$ est $A = \begin{pmatrix} 5 & 3 & 2 \\ 7 & 4 & 1 \end{pmatrix}$.

Finalement, c'est ce tableau lui-même qu'on appelle une matrice. On note souvent $[A]_{i,j} = a_{i,j}$.

On dit donc qu'un élément A de $\mathcal{M}_{n,p}(\mathbb{K})$ est une « matrice à n lignes et à p colonnes ».

Voici comment on peut former cette matrice dans Python, après avoir chargé le module `numpy` :

```
>>> import numpy as np          # charge le module numpy et le renomme np
>>> A = np.array([[5,3,2],[7,4,1]]); A  # forme et affiche une matrice d'entiers
array([[5, 3, 2],
       [7, 4, 1]])
```

On peut convertir les coefficients entiers de la matrice A précédente au format 'float' ou 'complex' :

```
>>> A.astype(float)
array([[ 5.,  3.,  2.],
       [ 7.,  4.,  1.]])
>>> A.astype(complex)
array([[ 5.+0.j,  3.+0.j,  2.+0.j],
       [ 7.+0.j,  4.+0.j,  1.+0.j]])
```

Attention!! En Python, les indices de ligne et de colonne commencent à 0.

```
>>> A[1,0]          # la ligne d'indice 1 est la deuxième ligne
7                  # la colonne d'indice 0 est la première colonne
```

Important Dans tous les exemples de ce chapitre, on supposera que le module `numpy` a été importé.

Avec Python, il est facile de créer une matrice dont le terme général répond à une formule donnée.

On peut par exemple créer la liste des listes de ces valeurs (en utilisant une syntaxe « en compréhension ») et convertir le résultat en un « array » :

```
>>> A = np.array([[10*i+j for j in range(10)] for i in range(5)])
>>> print(A)      # la matrice de terme général 10i+j, avec 0<=i<=4, et 0<=j<=9
[[ 0  1  2  3  4  5  6  7  8  9]
 [10 11 12 13 14 15 16 17 18 19]
 [20 21 22 23 24 25 26 27 28 29]
 [30 31 32 33 34 35 36 37 38 39]
 [40 41 42 43 44 45 46 47 48 49]]
```

Voici une autre méthode pour former la matrice précédente.

Ici les coefficients sont des flottants : si on veut des entiers, il suffit de rajouter l'option `dtype='int'`.

```
>>> B = np.fromfunction(lambda x,y: 10*x+y, (5,10))
>>> print(B)
[[ 0.  1.  2.  3.  4.  5.  6.  7.  8.  9.]
 [10. 11. 12. 13. 14. 15. 16. 17. 18. 19.]
 [20. 21. 22. 23. 24. 25. 26. 27. 28. 29.]
 [30. 31. 32. 33. 34. 35. 36. 37. 38. 39.]
 [40. 41. 42. 43. 44. 45. 46. 47. 48. 49.]]
```

On peut aussi créer des matrices aléatoires :

```
>>> A = np.random.rand(4,5) # quatre lignes, cinq colonnes, coeffs dans [0,1]
>>> print(A)
[[ 0.4672486  0.42452394  0.55652697  0.50122566  0.45039697]
 [ 0.10245706  0.39642783  0.18395313  0.43191611  0.59490124]
 [ 0.97133547  0.06577487  0.33274442  0.39404957  0.34601463]
 [ 0.60141568  0.01085386  0.69123586  0.33062893  0.68429281]]
```

Matrices constantes, matrice nulle

Une *matrice constante* $A = (a_{i,j})$ de $\mathcal{M}_{n,p}(\mathbb{K})$ est définie par $a_{i,j} = \lambda$ pour tous i, j .

En particulier, la *matrice nulle* $A = (a_{i,j})$ de $\mathcal{M}_{n,p}(\mathbb{K})$ est définie par $a_{i,j} = 0$ pour tous i, j .

Voici quelques façons de créer des matrices constantes avec Python :

```
>>> np.zeros([3, 4])
array([[ 0.,  0.,  0.,  0.],
       [ 0.,  0.,  0.,  0.],
       [ 0.,  0.,  0.,  0.]])
```

```
>>> np.ones([3, 4])
array([[ 1.,  1.,  1.,  1.],
       [ 1.,  1.,  1.,  1.],
       [ 1.,  1.,  1.,  1.]])
```

Ici la somme (ou le produit) d'une matrice et d'une constante s'effectue terme à terme :

```
>>> np.zeros([3, 4]) + 8
array([[ 8.,  8.,  8.,  8.],
       [ 8.,  8.,  8.,  8.],
       [ 8.,  8.,  8.,  8.]])
```

```
>>> np.ones([3, 4]) * 8
array([[ 8.,  8.,  8.,  8.],
       [ 8.,  8.,  8.,  8.],
       [ 8.,  8.,  8.,  8.]])
```

« Matrices-ligne » et « matrices-colonne »

Les éléments de $\mathcal{M}_{1,p}(\mathbb{K})$ sont dits *matrices-ligne*.

Ceux de $\mathcal{M}_{p,1}(\mathbb{K})$ sont dits *matrices-colonne*.

On peut identifier (en restant prudent) :

- le vecteur (a_1, a_2, \dots, a_p) de \mathbb{K}^p
- la matrice-ligne $(a_1 \ a_2 \ \dots \ a_p)$ de $\mathcal{M}_{1,p}(\mathbb{K})$

- la matrice-colonne $\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{pmatrix}$ de $\mathcal{M}_{p,1}(\mathbb{K})$

```
>>> # on définit un vecteur
>>> np.array([1,3,5])
array([1, 3, 5])

>>> # on définit une matrice-ligne
>>> np.array([[1,3,5]])
array([[1, 3, 5]])

>>> # on définit une matrice-colonne
>>> np.array([[1],[3],[5]])
array([[1],
       [3],
       [5]])
```

16.1.2 Structure d'espace vectoriel de $\mathcal{M}_{n,p}(\mathbb{K})$

Définition 16.1.2 (somme de deux matrices, produit par un scalaire)

Soit $A = (a_{ij})$ et $B = (b_{ij})$ deux matrices de $\mathcal{M}_{n,p}(\mathbb{K})$, et λ un scalaire.

On définit les matrices $C = A + B$ et $D = \lambda A$ dans $\mathcal{M}_{n,p}(\mathbb{K})$ de la manière suivante :

Pour tous indices i et j : $c_{ij} = a_{ij} + b_{ij}$ et $d_{ij} = \lambda a_{ij}$.

Proposition 16.1.1 (structure d'espace vectoriel de $\mathcal{M}_{n,p}(\mathbb{K})$)

L'ensemble $\mathcal{M}_{n,p}(\mathbb{K})$ est un espace vectoriel sur \mathbb{K} , de dimension np .

L'élément neutre est la matrice nulle, notée habituellement 0 .

L'opposé de la matrice $A = (a_{ij})$ est la matrice notée $-A$, de terme général $-a_{ij}$.

Base canonique

On fixe n et p dans \mathbb{N}^* , et on se place dans l'espace vectoriel $\mathcal{M}_{n,p}(\mathbb{K})$.

Soit i dans $\{1, \dots, n\}$ et j dans $\{1, \dots, p\}$.

On note $E_{i,j}$ la matrice dont tous les coefficients sont nuls, sauf celui d'indice (i, j) qui vaut 1.

La famille des np matrices $(E_{i,j})$, pour $1 \leq i \leq n$ et $1 \leq j \leq p$ est une base de $\mathcal{M}_{n,p}(\mathbb{K})$.

On l'appelle la « base canonique » de $\mathcal{M}_{n,p}(\mathbb{K})$.

Pour toute matrice $M = (a_{ij})$ de $\mathcal{M}_{n,p}(\mathbb{K})$, on a $M = \sum_{i=1}^n \sum_{j=1}^p a_{ij} E_{ij}$.

Voici par exemple les six matrices de la base canonique de $M_{2,3}(\mathbb{K})$:

$$\begin{aligned} E_{1,1} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} & E_{1,2} &= \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} & E_{1,3} &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \\ E_{2,1} &= \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} & E_{2,2} &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} & E_{2,3} &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

Pour toute matrice $A = \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}$ de $M_{2,3}(\mathbb{K})$, on a effectivement (et de façon unique) :

$$A = a \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + b \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + c \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} + d \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} + e \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} + f \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Voici une fonction Python qui renvoie une matrice de la base canonique de $\mathcal{M}_{n,p}(\mathbb{K})$ (attention encore une fois à la numérotation des lignes et des colonnes qui commence à 0 dans les tableaux Python) :

```
>>> def E(n, p, i, j, format='int'):
    a = np.zeros([n,p],format); a[i-1,j-1] = 1; return a
```

On calcule ici deux fois $E_{2,4}$ dans $\mathcal{M}_{3,5}(\mathbb{K})$ (d'abord en format 'int', puis en format 'float') :

```
>>> E(3, 5, 2, 4)
array([[0, 0, 0, 0, 0],
       [0, 0, 0, 1, 0],
       [0, 0, 0, 0, 0]])
```

```
>>> E(3, 5, 2, 4, 'float')
array([[0., 0., 0., 0., 0.],
       [0., 0., 0., 1., 0.],
       [0., 0., 0., 0., 0.]])
```

Avec Python, l'addition des matrices et le produit par des scalaires s'effectuent de façon très naturelle :

```
>>> A = np.random.randint(10, size=(3,4)) # coeffs aléatoires entiers dans [0,10[
>>> B = np.random.randint(10, size=(3,4)) # une autre matrice aléatoire
```

```
>>> print(A)
[[1 3 4 2]
 [0 5 5 1]
 [2 1 9 7]]
```

```
>>> print(B)
[[3 5 4 7]
 [1 2 3 1]
 [3 8 8 0]]
```

```
>>> print(A+B)
[[ 4  8  8  9]
 [ 1  7  8  2]
 [ 5  9 17  7]]
```

```
>>> print(10*A+B)
[[13 35 44 27]
 [ 1 52 53 11]
 [23 18 98 70]]
```

16.1.3 Matrices carrées

Définition 16.1.3 (matrices carrées d'ordre n)

On appelle *matrice carrée* d'ordre n (ou de taille n) toute matrice de type (n, n) .

On note $\mathcal{M}_n(\mathbb{K})$ (plutôt que $\mathcal{M}_{n,n}(\mathbb{K})$) l'ensemble des matrices carrées d'ordre n à coefficients dans \mathbb{K} .

Diagonale d'une matrice carrée

Soit $A = (a_{i,j})$ une matrice carrée d'ordre n .

Les coefficients $a_{i,i}$ (appelés *coefficients diagonaux*) de A forment la *diagonale* de A .

Les coefficients $a_{i,j}$ tels que $j < i$ sont donc *en dessous* de cette diagonale.

Les coefficients $a_{i,j}$ tels que $i < j$ sont *au dessus* de celle-ci.

On forme ici une matrice A de $\mathcal{M}_5(\mathbb{R})$, à coefficients entiers choisis aléatoirement dans $[0, 9[$.

On voit que la fonction `diagonal` du module `numpy` permet de récupérer la diagonale de A .

Remarque : il y a une deuxième diagonale dans une matrice carrée, mais seule celle qui débute « en haut en gauche » et se termine « en bas à droite » est importante (on le verra plus loin).

```
>>> A = np.random.randint(10, size=(4,4))
>>> print(A)
[[6 3 4 3]
 [4 1 3 9]
 [7 0 9 6]
 [4 8 6 4]]
>>> A.diagonal() # ou encore np.diag(A)
array([6, 1, 9, 4])
```

Matrices diagonales

Une matrice $A = (a_{i,j})$ de $\mathcal{M}_n(\mathbb{K})$ est dite *diagonale* si on a $a_{i,j} = 0$ pour tous indices distincts i et j .

Ainsi A est diagonale si les seuls coefficients non nuls *éventuels* de A sont ses coefficients diagonaux.

La matrice nulle de $\mathcal{M}_n(\mathbb{K})$ est donc un cas (très) particulier de matrice diagonale.

Matrice identité

La *matrice identité d'ordre n* est la matrice diagonale de $\mathcal{M}_n(\mathbb{K})$ de coefficients diagonaux égaux à 1.

Cette matrice est notée I_n .

On remarque que pour tous indices i et j , on a : $[I_n]_{i,j} = \delta_{i,j}$ (notation de Kronecker).

```
>>> np.diag([1,4,7,2])
array([[1, 0, 0, 0],
       [0, 4, 0, 0],
       [0, 0, 7, 0],
       [0, 0, 0, 2]])
```

```
>>> np.eye(4)
array([[ 1.,  0.,  0.,  0.],
       [ 0.,  1.,  0.,  0.],
       [ 0.,  0.,  1.,  0.],
       [ 0.,  0.,  0.,  1.]])
```

```
>>> np.eye(4, dtype='int')
array([[1, 0, 0, 0],
       [0, 1, 0, 0],
       [0, 0, 1, 0],
       [0, 0, 0, 1]])
```


Matrices scalaires

Les matrices carrées de la forme $A = \lambda I_n$, c'est-à-dire les matrices diagonales dont tous les coefficients diagonaux sont égaux, sont dites *matrices scalaires*.

```
>>> (5+2j)*np.eye(3)
array([[ 5.+2.j,  0.+0.j,  0.+0.j],
       [ 0.+0.j,  5.+2.j,  0.+0.j],
       [ 0.+0.j,  0.+0.j,  5.+2.j]])
```

Matrices triangulaires ou strictement triangulaires

Soit $A = (a_{i,j})$ une matrice de $\mathcal{M}_n(\mathbb{K})$.

On dit que A est *triangulaire supérieure* si les coefficients situés *sous* la diagonale sont nuls.

On dit que A est *triangulaire inférieure* si les coefficients situés *au-dessus* de la diagonale sont nuls.

Ainsi A est triangulaire inférieure si $(i < j \Rightarrow a_{i,j} = 0)$, et *triangulaire supérieure* si $(j < i \Rightarrow a_{i,j} = 0)$.

La matrice A est dite strictement triangulaire si elle est triangulaire et si de plus sa diagonale est nulle.

Avec Python, on forme ici une matrice aléatoire A , carrée d'ordre 4, à coefficients dans $\llbracket 1, 9 \rrbracket$.

Les fonctions `triu` et `tril` (avec `u` pour « up » et `l` pour « low ») permettent d'extraire de A une matrice triangulaire supérieure ou (éventuellement strictement, avec un argument optionnel).

```
>>> A=np.random.randint(1,10,size=(4,4))
>>> print(A)
[[4 6 2 7]
 [9 7 2 7]
 [3 4 5 1]
 [9 8 7 7]]
```

```
>>> np.triu(A)
array([[4, 6, 2, 7],
       [0, 7, 2, 7],
       [0, 0, 5, 1],
       [0, 0, 0, 7]])
```

Matrice triangulaire
supérieure

```
>>> np.triu(A,1)
array([[0, 6, 2, 7],
       [0, 0, 2, 7],
       [0, 0, 0, 1],
       [0, 0, 0, 0]])
```

Matrice strictement
triangulaire supérieure

```
>>> np.tril(A)
array([[4, 0, 0, 0],
       [9, 7, 0, 0],
       [3, 4, 5, 0],
       [9, 8, 7, 7]])
```

Matrice triangulaire
inférieure

```
>>> np.tril(A,-1)
array([[0, 0, 0, 0],
       [9, 0, 0, 0],
       [3, 4, 0, 0],
       [9, 8, 7, 0]])
```

Matrice strictement
triangulaire inférieure

Toujours en Python, la fonction `tri` permet de construire des matrices (strictement) triangulaires inférieures dont les coefficients sous-diagonaux valent 1 :

```
>>> np.tri(4,4,dtype='int')
array([[1, 0, 0, 0],
       [1, 1, 0, 0],
       [1, 1, 1, 0],
       [1, 1, 1, 1]])
```

```
>>> np.tri(4,4,-1,dtype='int')
array([[0, 0, 0, 0],
       [1, 0, 0, 0],
       [1, 1, 0, 0],
       [1, 1, 1, 0]])
```

16.2 Produit matriciel

16.2.1 Produit des matrices

Définition 16.2.1 (définition du produit de deux matrices)

Soit $A = (a_{ik})$ dans $\mathcal{M}_{n,p}(\mathbb{K})$ et $B = (b_{kj})$ dans $\mathcal{M}_{p,q}(\mathbb{K})$.

On définit $M = AB$, dans $\mathcal{M}_{n,q}(\mathbb{K})$, par : $\forall i \in \{1, \dots, n\}, \forall j \in \{1, \dots, q\}, m_{ij} = \sum_{k=1}^p a_{ik} b_{kj}$.

Interprétation et visualisation

Le terme de la i -ième ligne et de la j -ième colonne de $M = AB$ est donc obtenu en sommant les produits des termes de même rang dans la i -ième ligne L_i de A et dans la j -ième colonne C_j de B , selon le schéma ci-après (on a encadré les coefficients de A et de B utiles au calcul du coefficient c_{ij}).

$$L_i \rightarrow \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1k} & \dots & a_{1p} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ik} & \dots & a_{ip} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nk} & \dots & a_{np} \end{pmatrix} \begin{pmatrix} b_{11} & \dots & b_{1j} & \dots & b_{1q} \\ b_{21} & \dots & b_{2j} & \dots & b_{2q} \\ \vdots & & \vdots & & \vdots \\ b_{k1} & & b_{kj} & & b_{kq} \\ \vdots & & \vdots & & \vdots \\ b_{p1} & \dots & b_{pj} & \dots & b_{pq} \end{pmatrix} = \begin{pmatrix} m_{11} & \dots & m_{1j} & \dots & m_{1q} \\ \vdots & & \vdots & & \vdots \\ m_{i1} & \dots & m_{ij} & \dots & m_{iq} \\ \vdots & & \vdots & & \vdots \\ m_{n1} & \dots & m_{nj} & \dots & m_{nq} \end{pmatrix}$$

Pour former le coefficient $m_{i,j}$, on met en regard la ligne L_i de A et la colonne C_j de B , et on somme les produits deux à deux des a_{ik} et b_{kj} de même position k , avec $1 \leq k \leq p$

On verra plus loin les raisons pour lesquelles on choisit de définir de cette façon le produit matriciel.

Condition impérative de compatibilité

Le produit AB n'est possible que si le nombre de colonnes de A est égal au nombre de lignes de B . On obtient alors une matrice ayant autant de lignes que A , et autant de colonnes que B .

En résumé : $\boxed{\text{(matrice de type } (n, p)) \times \text{(matrice de type } (p, q)) = \text{(matrice de type } (n, q))}$

Avec Python on forme ici deux matrices aléatoires, A dans $\mathcal{M}_{4,3}(\mathbb{R})$, l'autre dans $\mathcal{M}_{3,5}(\mathbb{R})$:

```
>>> A=np.random.randint(10,size=(4,3))
>>> print(A)
[[2 2 7]
 [6 3 8]
 [4 3 3]
 [7 7 0]]
```

```
>>> B=np.random.randint(10,size=(3,5))
>>> print(B)
[[6 8 1 2 6]
 [5 6 2 8 7]
 [5 6 8 4 9]]
```

On voit qu'on peut calculer AB et le résultat est une matrice de type $(4, 5)$.

Mais le produit BA est impossible à calculer et conduirait à l'erreur de format suivante :

ValueError: objects are not aligned

```
>>> print(A.dot(B)) # ou np.dot(A,B)
[[ 57  70  62  48  89]
 [ 91 114  76  68 129]
 [ 54  68  34  44  72]
 [ 77  98  21  70  91]]
```

En Python, attention à l'opération `*` qui désigne en fait le produit « terme à terme » des tableaux. Dans l'exemple ci-après, on effectue le produit terme à terme de deux matrices A et B de taille 4×3 .

Le résultat est encore une matrice 4×3 dont le terme général est $c_{i,j} = a_{i,j}b_{i,j}$. Ce produit (qui peut avoir son utilité) n'a rien à voir avec le produit matriciel dont nous parlons dans ce chapitre!

```

>>> print(A)
[[9 2 4]
 [4 3 7]
 [9 3 5]
 [6 1 9]]
>>> print(B)
[[3 2 2]
 [1 8 1]
 [6 8 6]
 [4 5 2]]
>>> print(A*B)
[[27 4 8]
 [ 4 24 7]
 [54 24 30]
 [24 5 18]]

```

Attention, ce n'est pas le « vrai » produit matriciel!

16.2.2 La non-commutativité du produit

Les produits AB et BA ne sont simultanément possibles que si A est de type (n, p) et B de type (p, n) . La matrice AB est alors carrée d'ordre n , tandis que BA est carrée d'ordre p .

Si $n \neq p$, les matrices AB et BA , de formats différents, sont évidemment distinctes!

Si A et B sont toutes deux carrées d'ordre n , alors AB et BA sont encore carrées d'ordre n , mais là encore on a en général $AB \neq BA$. Dans le cas contraire (très rare), on dit que les matrices carrées A et B *commutent* (on en reparle plus loin).

Exemple avec Python : on utilise ici la fonction `arange` (qui forme le vecteur des éléments d'une suite arithmétique, ici la suite $0, 1, \dots, 5$) et la fonction `reshape` (qui redéfinit la taille du tableau, du moment que le nombre de coefficients est le même).

```

>>> A = np.arange(6).reshape(3,2); print(A)
[[0 1]
 [2 3]
 [4 5]]
>>> B = np.arange(6).reshape(2,3)
>>> print(B)
[[0 1 2]
 [3 4 5]]

```

On a donc formé une matrice A de type $(3, 2)$, et une matrice B de type $(2, 3)$.

Les deux produits AB et BA sont possibles, mais AB est carrée d'ordre 3 et BA est carrée d'ordre 2 :

```

>>> print(A.dot(B))
[[ 3  4  5]
 [ 9 14 19]
 [15 24 33]]
>>> print(B.dot(A))
[[10 13]
 [28 40]]
>>>

```

Autre exemple, on forme ici deux matrices A et B carrées d'ordre 3 :

```

>>> A = np.random.rand(3,3); print(A)
[[ 0.75596295  0.96396382  0.45077587]
 [ 0.67323071  0.04498405  0.88736757]
 [ 0.86576073  0.03023288  0.42663428]]
>>> B = np.random.rand(3,3); print(B)
[[ 0.68788001  0.35191549  0.98365263]
 [ 0.38822915  0.80922831  0.75737939]
 [ 0.92474908  0.11545089  0.3735433 ]]

```

On constate effectivement que les produits AB et BA sont différents (c'est le cas général!) :

```

>>> print(np.dot(A,B))
[[ 1.31110522  1.09814436  1.64207559]
 [ 1.30115841  0.37577006  1.02776536]
 [ 1.00180644  0.37839522  1.03387196]]
>>> print(np.dot(B,A))
[[ 1.60853994  0.70866068  1.04201804]
 [ 1.49399354  0.43353898  1.2162113 ]
 [ 1.10020024  0.90791139  0.67866832]]

```

16.2.3 Propriétés du produit matriciel

Proposition 16.2.1 (bilinéarité du produit matriciel)

Soit A, B dans $\mathcal{M}_{n,p}(\mathbb{K})$, soit C, D dans $\mathcal{M}_{p,q}(\mathbb{K})$, et soit λ, μ deux scalaires.

Alors on a les égalités $A(\lambda C + \mu D) = \lambda AC + \mu AD$ et $(\lambda A + \mu B)C = \lambda AC + \mu BC$.

Cette propriété peut s'énoncer de la façon suivante :

- Si on fixe A dans $\mathcal{M}_{n,p}(\mathbb{K})$, l'application $M \mapsto AM$ est linéaire de $\mathcal{M}_{p,q}(\mathbb{K})$ dans $\mathcal{M}_{n,q}(\mathbb{K})$.
- Si on fixe A dans $\mathcal{M}_{p,q}(\mathbb{K})$, l'application $M \mapsto MA$ est linéaire de $\mathcal{M}_{n,p}(\mathbb{K})$ dans $\mathcal{M}_{n,q}(\mathbb{K})$.
- On peut résumer : l'application $(M, N) \mapsto MN$ est *bilinéaire* de $\mathcal{M}_{n,p}(\mathbb{K}) \times \mathcal{M}_{p,q}(\mathbb{K})$ dans $\mathcal{M}_{n,q}(\mathbb{K})$.

Proposition 16.2.2 (associativité du produit matriciel)

Soit A dans $\mathcal{M}_{n,p}(\mathbb{K})$, soit B dans $\mathcal{M}_{p,q}(\mathbb{K})$, et soit C dans $\mathcal{M}_{q,r}(\mathbb{K})$.

Alors, on a l'égalité $A(BC) = (AB)C$ dans $\mathcal{M}_{n,r}(\mathbb{K})$.

On définit ici trois matrices A, B, C , de coefficients aléatoires dans $\{-1, 0, 1\}$.

```
>>> A = np.random.randint(-1,2,size=(3,4))
>>> B = np.random.randint(-1,2,size=(4,3))
>>> C = np.random.randint(-1,2,size=(3,2))
```

On a donc choisi A dans $\mathcal{M}_{3,4}(\mathbb{R})$,
 B dans $\mathcal{M}_{4,3}(\mathbb{R})$, et C dans $\mathcal{M}_{3,2}(\mathbb{R})$.

Voici les trois matrices A, B, C .

```
>>> print(A)
[[ 0  1  1  0]
 [ 0 -1  1  0]
 [ 0  1  1 -1]]
>>>
>>> print(B)
[[ 0  1  0]
 [ 0  1  1]
 [ 1  0  0]
 [ 1 -1  0]]
>>>
>>> print(C)
[[ 0  1]
 [ 0  1]
 [-1 -1]]
>>>
```

On affiche alors AB (qui est dans $\mathcal{M}_{3,3}(\mathbb{R})$), puis $(AB)C$ (qui est dans $\mathcal{M}_{3,2}(\mathbb{R})$).

Ensuite on affiche alors BC (qui est dans $\mathcal{M}_{4,2}(\mathbb{R})$), puis $A(BC)$ (qui est dans $\mathcal{M}_{3,2}(\mathbb{R})$).

On constate que les deux matrices $A(BC)$ et $(AB)C$ sont égales dans $\mathcal{M}_{3,2}(\mathbb{R})$:

```
>>> print(A.dot(B))
[[ 1  1  1]
 [ 1 -1 -1]
 [ 0  2  1]]
>>> print((A.dot(B)).dot(C))
[[-1  1]
 [ 1  1]
 [-1  1]]
>>>
```

```
>>> print(B.dot(C))
[[ 0  1]
 [-1  0]
 [ 0  1]
 [ 0  0]]
>>> print(A.dot(B.dot(C)))
[[-1  1]
 [ 1  1]
 [-1  1]]
>>>
```

16.2.4 Une justification du produit matriciel

Une situation particulière mais importante

L'application qui à $u = (x_1, x_2, \dots, x_p)$ associe $U = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$ est un isomorphisme de \mathbb{K}^p sur $\mathcal{M}_{p,1}(\mathbb{K})$. Cet isomorphisme permet d'identifier u et U .

Soit A un élément de $\mathcal{M}_{n,p}(\mathbb{K})$. L'application $U \mapsto V = AU$ est linéaire de $\mathcal{M}_{p,1}(\mathbb{K})$ dans $\mathcal{M}_{n,1}(\mathbb{K})$.

À un isomorphisme près, il s'agit donc d'une application linéaire de \mathbb{K}^p dans \mathbb{K}^n .

Considérons par exemple la matrice $A = \begin{pmatrix} 2 & 1 & 0 & 3 \\ 0 & 4 & 1 & 5 \\ 1 & 6 & 2 & 0 \end{pmatrix}$, élément de $\mathcal{M}_{3,4}(\mathbb{R})$.

Si $U = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$ est dans $\mathcal{M}_{4,1}(\mathbb{R})$ (disons dans \mathbb{R}^4), $V = AU = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$ est dans $\mathcal{M}_{3,1}(\mathbb{R})$ (disons dans \mathbb{R}^3).

Le produit matriciel donne : $\begin{pmatrix} 2 & 1 & 0 & 3 \\ 0 & 4 & 1 & 5 \\ 1 & 6 & 2 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 2x_1 + x_2 + 3x_4 \\ 4x_2 + x_3 + 5x_4 \\ x_1 + 6x_2 + 2x_3 \end{pmatrix}$ donc $\begin{cases} y_1 = 2x_1 + x_2 + 3x_4 \\ y_2 = 4x_2 + x_3 + 5x_4 \\ y_3 = x_1 + 6x_2 + 2x_3 \end{cases}$

À isomorphisme près, on a donc défini $f : u = (x_1, x_2, x_3, x_4) \mapsto v = (y_1, y_2, y_3)$ linéaire de \mathbb{R}^4 dans \mathbb{R}^3 .

On dira que la matrice A est *canoniquement associée* à l'application linéaire f .

Dans le cas général : on peut légitimement identifier une application linéaire de \mathbb{K}^p dans \mathbb{K}^n avec une matrice A de $\mathcal{M}_{n,p}(\mathbb{K})$. Si on note U, V les matrices colonnes représentant respectivement u dans \mathbb{K}^p et v dans \mathbb{K}^n , cela nous permet d'identifier les égalités $v = f(u)$ et $V = AU$.

Identification d'une matrice-ligne et d'une forme linéaire

Un élément L de $\mathcal{M}_{1,n}(\mathbb{K})$ (une matrice-ligne de largeur n) s'identifie à une forme linéaire sur \mathbb{K}^n .

Par exemple, considérons la matrice ligne $A = (a \ b \ c \ d)$.

Pour toute matrice colonne $U = \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix}$, on a : $AX = (a \ b \ c \ d) \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = ax + by + cz + dt$.

La matrice ligne A s'identifie donc à la forme linéaire f définie sur \mathbb{K}^4 par $f(x, y, z, t) = ax + by + cz + dt$.

Justification du produit matriciel

Considérons deux applications linéaires $f : \mathbb{K}^n \rightarrow \mathbb{K}^p$ et $g : \mathbb{K}^p \rightarrow \mathbb{K}^q$.

L'application $h = g \circ f$ est linéaire de \mathbb{K}^n dans \mathbb{K}^q .

– à l'application f est canoniquement associée une matrice A de $\mathcal{M}_{p,n}(\mathbb{K})$.

– à l'application g est canoniquement associée une matrice B de $\mathcal{M}_{q,p}(\mathbb{K})$.

– à l'application h est canoniquement associée une matrice C de $\mathcal{M}_{q,n}(\mathbb{K})$.

Pour tous vecteurs u dans \mathbb{K}^n , v dans \mathbb{K}^p et w dans \mathbb{K}^q , on a l'équivalence : $w = h(u) \Leftrightarrow w = g(f(u))$.

Matriciellement, $w = h(u)$ s'écrit $W = CU$, et $w = g(f(u))$ s'écrit $W = B(AU)$.

Précisément, la définition du produit matriciel est choisie pour qu'on ait l'égalité $BA = C$.

Autrement dit, si on attache aux matrices A et B deux applications linéaires f et g , le produit BA est la matrice attachée à l'application linéaire $g \circ f$.

On résume ça en imaginant que le diagramme : $\mathbb{K}^n \xrightarrow[A]{f} \mathbb{K}^p \xrightarrow[B]{g} \mathbb{K}^q$ se « contracte » en $\mathbb{K}^n \xrightarrow[BA]{g \circ f} \mathbb{K}^q$.

Prenons un exemple simple, en restant en dimension 2 :

$\begin{cases} f : (x, y) \rightarrow (x + y, 2x - y) \\ g : (x, y) \rightarrow (x + 3y, x - y) \end{cases}$ sont attachées à $A = \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & 3 \\ 1 & -1 \end{pmatrix}$. On a $BA = \begin{pmatrix} 7 & -2 \\ -1 & 2 \end{pmatrix}$.

On vérifie que $g \circ f : (x, y) \rightarrow (7x - 2y, -x + 2y)$ est attachée à $C = \begin{pmatrix} 7 & -2 \\ -1 & 2 \end{pmatrix}$ et qu'on a bien $C = BA$.

16.2.5 Produits, lignes et colonnes

Dans le calcul du produit AB de deux matrices, on met en rapport les lignes de A et les colonnes de B .

Il est alors commode de se représenter une matrice quelconque sous la forme :

- de la superposition de n matrices-ligne L_1, L_2, \dots, L_n de même largeur.
- de la juxtaposition de p matrices-colonne C_1, C_2, \dots, C_p de même hauteur.

On pourra alors noter symboliquement $\left(\begin{array}{c} L_1 \\ L_2 \\ \vdots \\ L_n \end{array} \right)$ et $\left(C_1 \mid C_2 \mid \dots \mid C_p \right)$

Le terme d'indice (i, j) de $\left(\begin{array}{c} L_1 \\ L_2 \\ \vdots \\ L_n \end{array} \right) \left(C_1 \mid C_2 \mid \dots \mid C_p \right)$ est le « produit scalaire » de L_i par C_j .

Avec Python, on forme ici les matrices A dans $\mathcal{M}_{4,5}(\mathbb{R})$ et B dans $\mathcal{M}_{5,4}(\mathbb{K})$.

```
>>> A=np.random.randint(-5,5,size=(4,5))
>>> print(A)
[[ 1 -4 -1 -3  4]
 [-1  1  4  4 -5]
 [ 3 -3 -1 -1  1]
 [-1 -1  2  0  1]]
>>>

>>> B=np.random.randint(-5,5,size=(5,4))
>>> print(B)
[[-1 -4 -3 -3]
 [-1  1  1  0]
 [-2  1 -2 -5]
 [ 2 -5  2  2]
 [-1 -5 -3 -3]]
```

On extrait la ligne d'indice 3 de A (attention encore une fois à la numérotation Python qui commence à 0 : il s'agit donc ici, avec nos notations, de la ligne L_4).

On extrait également la colonne d'indice 2 de B (donc c'est la colonne C_3 !).

On forme ensuite le produit scalaire de cette ligne et de cette colonne.

Le résultat, égal à -5 , est effectivement le coefficient de position $(3, 2)$ de $C = AB$ (donc c'est $c_{4,3}$!) :

```
>>> A[3,:] # quatrième ligne !!!
array([-1, -1,  2,  0,  1])
>>> B[:,2] # troisième colonne !!!
array([-3,  1, -2,  2, -3])
>>> np.vdot(A[3,:],B[:,2])
-5

>>> A.dot(B) # ou np.dot(A,B)
array([[ -5, -14, -23, -16],
 [  5,  14,  19,   6],
 [ -1, -16, -15,  -9],
 [ -3,   0,  -5, -10]])
>>>
```

Si A est un élément de $\mathcal{M}_{n,p}(\mathbb{K})$ et si les x_k sont scalaires, on peut écrire :

$$\underbrace{\left(C_1 \mid C_2 \mid \dots \mid C_p \right)}_A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix} = \sum_{j=1}^p x_j C_j \quad \text{et} \quad \begin{pmatrix} x_1 & x_2 & \dots & x_n \end{pmatrix} \underbrace{\left(\begin{array}{c} L_1 \\ L_2 \\ \vdots \\ L_n \end{array} \right)}_A = \sum_{i=1}^n x_i L_i$$

De même, les produits matriciels MA et AN peuvent s'écrire :

$$M \underbrace{\left(C_1 \mid C_2 \mid \dots \mid C_p \right)}_A = \left(MC_1 \mid MC_2 \mid \dots \mid MC_p \right) \quad \text{et} \quad \underbrace{\left(\begin{array}{c} L_1 \\ L_2 \\ \vdots \\ L_n \end{array} \right)}_A N = \left(\begin{array}{c} L_1 N \\ L_2 N \\ \vdots \\ L_n N \end{array} \right)$$

Ces idées peuvent être mises à profit pour optimiser le calcul d'un produit AB de deux matrices :

- si la matrice A est « simple », on privilégiera un calcul de AB par lignes.
- si la matrice B est « simple », on privilégiera un calcul de AB par colonnes.

Posons par exemple $A = \begin{pmatrix} a & b & c & d \\ a' & b' & c' & d' \\ a'' & b'' & c'' & d'' \end{pmatrix}$ et $B = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix}$.

On a ici intérêt à effectuer le calcul de AB colonne par colonne :

$$\begin{aligned} AB &= \begin{pmatrix} a & b & c & d \\ a' & b' & c' & d' \\ a'' & b'' & c'' & d'' \end{pmatrix} \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix} = \left(\begin{array}{c|c|c|c} C_1 & C_2 & C_3 & C_4 \end{array} \right) \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix} \\ &= \left(\begin{array}{c|c|c} C_1 + C_3 + C_4 & 2C_1 + C_2 & -C_2 + C_4 \end{array} \right) = \begin{pmatrix} a + c + d & 2a + b & -b + d \\ a' + c' + d' & 2a' + b' & -b' + d' \\ a'' + c'' + d'' & 2a'' + b'' & -b'' + d'' \end{pmatrix} \end{aligned}$$

De même, on a intérêt à effectuer le produit BA ligne par ligne :

$$\begin{aligned} BA &= \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} a & b & c & d \\ a' & b' & c' & d' \\ a'' & b'' & c'' & d'' \end{pmatrix} = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} L_1 \\ L_2 \\ L_3 \end{pmatrix} \\ &= \begin{pmatrix} L_1 + 2L_2 \\ L_2 - L_3 \\ L_1 \\ L_1 + L_3 \end{pmatrix} = \begin{pmatrix} a + 2a' & b + 2b' & c + 2c' & d + 2d' \\ a' - a'' & b' - b'' & c' - c'' & d' - d'' \\ a & b & c & d \\ a + a'' & b + b'' & c + c'' & d + d'' \end{pmatrix} \end{aligned}$$

16.2.6 Produits de matrices de la base canonique

Proposition 16.2.3 (produits de matrices de la base canonique)

On se donne les ensembles $\mathcal{M}_{n,p}(\mathbb{K})$, $\mathcal{M}_{p,q}(\mathbb{K})$ et $\mathcal{M}_{n,q}(\mathbb{K})$, avec n, p, q dans \mathbb{N}^* .

Soit $E_{i,j}$ (resp. $E'_{j,k}$, $E''_{i,k}$) les matrices de la base canonique de $\mathcal{M}_{n,p}(\mathbb{K})$ (resp. $\mathcal{M}_{p,q}(\mathbb{K})$, $\mathcal{M}_{n,q}(\mathbb{K})$).

Si les indices j et j' sont distincts, alors le produit $E_{i,j} E'_{j',k}$ est égal à la matrice nulle de $\mathcal{M}_{n,q}(\mathbb{K})$.

Si $j' = j$, alors le produit $E_{i,j} E'_{j,k}$ est égal à $E''_{i,k}$.

Plus généralement, quand on effectue le produit $A E_{i,j}$ le résultat est une matrice nulle à l'exception de sa j -ème colonne qui est la copie de la i -ème colonne de A .

De même $E_{i,j} A$ est nulle à l'exception de sa i -ème ligne qui est la copie de la j -ème ligne de A .

Voici un exemple, où on réutilise la fonction E de la sous-section 16.1.2 :

```
>>> print(A)
[[-3  2  0  3  1]
 [ 0  1 -5 -4  4]
 [ 3  2  2  1 -3]
 [-4  3 -1  1 -3]
 [-4 -1  3  0  4]]
```

Une matrice de $\mathcal{M}_{5,5}(\mathbb{R})$

```
>>> print(A.dot(E(5,5,2,4)))
[[ 0  0  0  2  0]
 [ 0  0  0  1  0]
 [ 0  0  0  2  0]
 [ 0  0  0  3  0]
 [ 0  0  0 -1  0]]
```

Produit par $E_{2,4}$ à droite
 C_2 a été copiée en C_4

```
>>> print(E(5,5,2,4).dot(A))
[[ 0  0  0  0  0]
 [-4  3 -1  1 -3]
 [ 0  0  0  0  0]
 [ 0  0  0  0  0]
 [ 0  0  0  0  0]]
```

Produit par $E_{2,4}$ à gauche
 L_4 a été copiée en L_2

16.3 Calculs sur les matrices carrées

16.3.1 L'anneau $(\mathcal{M}_n(\mathbb{K}), +, \times)$

Proposition 16.3.1 (l'anneau $\mathcal{M}_n(\mathbb{K})$)

Muni de la somme et du produit des matrices, l'ensemble $\mathcal{M}_n(\mathbb{K})$ possède une structure d'anneau. L'anneau $(\mathcal{M}_n(\mathbb{K}), +, \times)$ est non commutatif dès que $n \geq 2$.

– Le neutre additif de l'anneau $\mathcal{M}_n(\mathbb{K})$ est la matrice nulle, notée 0 (ou éventuellement 0_n).
Le neutre multiplicatif de l'anneau $\mathcal{M}_n(\mathbb{K})$ est la matrice identité I_n .

– Si $n = 1$, on identifie une matrice $A = (a)$ avec l'unique scalaire qu'elle contient.
Dans ce cas, $\mathcal{M}_1(\mathbb{K})$ s'identifie donc au corps (commutatif) \mathbb{K} .

– Il est facile de former des matrices qui ne commutent pas dans $\mathcal{M}_2(\mathbb{K})$.

Par exemple si $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ et $B = \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix}$, alors $AB = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ et $BA = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$.

Si $n \geq 2$, l'anneau $\mathcal{M}_n(\mathbb{K})$ est non commutatif, comme on le voit en « étendant » l'exemple précédent :

$$\text{si } A = \begin{pmatrix} 0 & 1 & 0 & \cdots \\ 1 & 0 & 0 & \cdots \\ 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \text{ et } B = \begin{pmatrix} 0 & 2 & 0 & \cdots \\ 1 & 0 & 0 & \cdots \\ 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \text{ alors } AB = \begin{pmatrix} 1 & 0 & 0 & \cdots \\ 0 & 2 & 0 & \cdots \\ 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \text{ et } BA = \begin{pmatrix} 2 & 0 & 0 & \cdots \\ 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Proposition 16.3.2 (produits de matrices diagonales ou triangulaires)

Soit A et B deux matrices de $\mathcal{M}_n(\mathbb{K})$.

On suppose que A et B sont diagonales (resp. triangulaires supérieures, triangulaires inférieures).

Alors $C = AB$ est encore diagonale (resp. triangulaire supérieure, triangulaire inférieure).

De plus les coefficients diagonaux c_{ii} de C vérifient $c_{ii} = a_{ii}b_{ii}$.

Propriétés diverses

– Les matrices triangulaires supérieures forment un sous-espace de $\mathcal{M}_n(\mathbb{R})$, de dimension $\frac{n(n+1)}{2}$.

Le résultat précédent nous dit que c'est aussi un sous-anneau de $\mathcal{M}_n(\mathbb{R})$.

Il en est bien sûr de même pour les matrices triangulaires inférieures.

Les matrices diagonales forment un sous-espace de $\mathcal{M}_n(\mathbb{R})$ (et un sous-anneau) de dimension n .

– La proposition précédente s'étend (récurrence facile) à un nombre quelconque de matrices carrées (toutes diagonales, ou toutes triangulaires supérieures, ou toutes triangulaires inférieures).

Cela s'applique donc aux puissances A^p (avec p dans \mathbb{N}^*) d'une matrice diagonale ou triangulaire.

– Si deux matrices A et B sont triangulaires (disons supérieurement) et si l'une d'elle est strictement triangulaire, alors leur produit AB est strictement triangulaire.

– On ne peut rien dire du produit de deux matrices triangulaires, l'une inférieure et l'autre supérieure.

Par exemple, si $A = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$ et $B = \begin{pmatrix} 4 & 0 \\ 5 & 6 \end{pmatrix}$, alors $AB = \begin{pmatrix} 14 & 12 \\ 15 & 18 \end{pmatrix}$ et $BA = \begin{pmatrix} 4 & 8 \\ 5 & 28 \end{pmatrix}$.

En Python, on voit ici une matrice A de $\mathcal{M}_4(\mathbb{R})$, ainsi que son carré A^2 et son cube A^3 : si un calcul est nécessaire pour les coefficients « sur-diagonaux », les coefficients diagonaux sont immédiats :

```
>>> print(A)
[[ 2  3 -2  4]
 [ 0 -1  1  0]
 [ 0  0  1  5]
 [ 0  0  0  3]]

>>> A2 = A.dot(A); print(A2)
[[ 4  3 -3 10]
 [ 0  1  0  5]
 [ 0  0  1 20]
 [ 0  0  0  9]]

>>> A3 = A.dot(A2); print(A3)
[[ 8  9 -8 31]
 [ 0 -1  1 15]
 [ 0  0  1 65]
 [ 0  0  0 27]]
```

Une matrice triangulaire très particulière

On note ici T_n la matrice strictement triangulaire dont tous les coefficients sont nuls, sauf ceux situés immédiatement au-dessus de la diagonale et qui valent 1.

Voici une fonction Python pour définir une telle matrice :

```
def T(n):
    return np.array([[1 if j==i+1 else 0 for j in range(n)] for i in range(n)])
```

Voici par exemple la matrice $J = T_4$, ainsi que J^2 et J^3 . On voit très bien comment la « sur-diagonale » de coefficients égaux à 1 « remonte » quand l'exposant augmente (la prochaine étape est $J^4 = 0$).

Remarque : en mode Python interactif, on désigne par `_` le résultat du calcul précédent.

```
>>> J = T(4); J
array([[0, 1, 0, 0],
       [0, 0, 1, 0],
       [0, 0, 0, 1],
       [0, 0, 0, 0]])

>>> J.dot(_)
array([[0, 0, 1, 0],
       [0, 0, 0, 1],
       [0, 0, 0, 0],
       [0, 0, 0, 0]])

>>> J.dot(_)
array([[0, 0, 0, 1],
       [0, 0, 0, 0],
       [0, 0, 0, 0],
       [0, 0, 0, 0]])
```

La matrice $J = T_4$

Le carré J^2 de J

Le cube J^3 de J

Les matrices triangulaires $J = T_n$ reviennent souvent.

Le passage de A au produit JA revient à « décaler vers le haut » les lignes de A .

Le passage de A au produit AJ revient à « décaler vers la droite » les parallèles à la diagonale de A .

Pour l'exemple ci-dessous, on a fabriqué la matrice A de la manière suivante :

```
A = np.arange(16).reshape(4,4)
```

On multiplie ensuite A sur sa gauche, puis sur sa droite, par la matrice J de l'exemple précédent :

```
>>> A
array([[ 1,  2,  3,  4],
       [ 5,  6,  7,  8],
       [ 9, 10, 11, 12],
       [13, 14, 15, 16]])

>>> J.dot(A)
array([[ 5,  6,  7,  8],
       [ 9, 10, 11, 12],
       [13, 14, 15, 16],
       [ 0,  0,  0,  0]])

>>> A.dot(J)
array([[ 0,  1,  2,  3],
       [ 0,  5,  6,  7],
       [ 0,  9, 10, 11],
       [ 0, 13, 14, 15]])
```

La matrice A

Le produit par J à gauche
remonte les lignes vers le haut

Le produit par J à droite
décale les colonnes vers la droite

16.3.2 La formule du binôme

Le résultat suivant est un cas particulier d'une situation rencontrée dans 11.3.3 :

Proposition 16.3.3 (formule du binôme pour deux matrices carrées qui commutent)

Soit A et B deux matrices de $\mathcal{M}_n(\mathbb{K})$. On suppose que A et B commutent.

Alors, pour tout entier naturel p , on a : $(A + B)^p = \sum_{k=0}^p \binom{p}{k} A^k B^{p-k} = \sum_{k=0}^p \binom{p}{k} A^{p-k} B^k$.

Remarques

– L’hypothèse selon laquelle $AB = BA$ est essentielle. Si elle n’est pas réalisée, alors le développement de $(A + B)^p$ est totalement différent (et beaucoup plus compliqué).

Par exemple, si $AB \neq BA$, on a :
$$\begin{cases} (A + B)^2 = A^2 + AB + BA + B^2 \\ (A + B)^3 = A^3 + A^2B + ABA + AB^2 + BA^2 + BAB + B^2A + B^3 \end{cases}$$

Toujours si $AB \neq BA$ on a : $(A + B)(A - B) = A^2 - AB + BA - B^2$, donc $(A + B)(A - B) \neq A^2 - B^2$.

– Les matrices scalaires λI_n commutent avec toutes les matrices de $\mathcal{M}_n(\mathbb{K})$.

Pour tout A de $\mathcal{M}_n(\mathbb{K})$, on peut donc écrire : $(A + \lambda I_n)^p = \sum_{k=0}^p \binom{p}{k} \lambda^{p-k} A^k = \sum_{k=0}^p \binom{p}{k} \lambda^k A^{p-k}$.

– Si A_1, \dots, A_m commutent deux à deux dans $\mathcal{M}_n(\mathbb{K})$, alors $\left(\sum_{k=1}^m A_k\right)^2 = \sum_{k=1}^m A_k^2 + 2 \sum_{1 \leq i < j \leq m} A_i A_j$

– Si A et B commutent et si $p \geq 1$, alors $(AB)^p = A^p B^p$, sinon $(AB)^p = ABAB \cdots AB$ (p fois).

16.3.3 Matrices nilpotentes ou diviseurs de zéro

Si $J = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$, on a $J^2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$, $J^3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$, et $J^p = 0$ pour tout $p \geq 4$.

On exprime cette propriété en disant que J est nilpotente, et que son indice de nilpotence est 4.

Définition 16.3.1 (matrices nilpotentes)

Soit A une matrice de $\mathcal{M}_n(\mathbb{K})$.

On dit que la matrice A est *nilpotente* s’il existe un entier $p \geq 1$ tel que $A^p = 0$.

Le plus petit entier $p \geq 1$ tel que $A^p = 0$ est appelé *indice de nilpotence* de A .

Avec cette définition, on convient que la matrice nulle est nilpotente d’indice 1.

Proposition 16.3.4 (condition suffisante pour qu’une matrice soit nilpotente)

Toute matrice strictement triangulaire est nilpotente (la réciproque est fausse).

L’exemple de $A = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$, qui vérifie $A^2 = 0$, montre effectivement que la réciproque est fausse.

En revanche : si A est triangulaire, elle est nilpotente si et seulement si elle est *strictement* triangulaire.

Proposition 16.3.5 (condition nécessaire sur l’indice de nilpotence)

Soit A une matrice nilpotente de $\mathcal{M}_n(\mathbb{K})$.

Alors l’indice de nilpotence de A est inférieur ou égal à n . On est donc certain que $A^n = 0$.

Par la contraposée : si A est dans $\mathcal{M}_n(\mathbb{K})$ et si $A^n \neq 0$, il est certain que A n'est pas nilpotente !

Proposition 16.3.6 (somme ou produit de deux matrices nilpotentes)

Soit A et B deux matrices nilpotentes de $\mathcal{M}_n(\mathbb{K})$.

Si A et B commutent, alors AB et $A + B$ sont nilpotentes.

L'hypothèse selon laquelle $AB = BA$ est essentielle.

Par exemple si $A = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$ et $B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ alors $AB = \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix}$ et $BA = \begin{pmatrix} -1 & -1 \\ 0 & 0 \end{pmatrix}$.

On trouve $(AB)^2 = \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix} = AB$ et $A + B = \begin{pmatrix} 1 & 0 \\ -1 & -1 \end{pmatrix}$ (donc ni AB ni $A + B$ ne sont nilpotentes).

Définition 16.3.2 (matrices diviseurs de zéro)

Soit A une matrice de $\mathcal{M}_n(\mathbb{K})$. On dit que A est un *diviseur de zéro* s'il existe une matrice B non nulle dans $\mathcal{M}_n(\mathbb{K})$ telle que $AB = 0$ ou $BA = 0$.

Avec la définition précédente, on convient que la matrice nulle est un diviseur de 0.

Par exemple, si $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}$ et $C = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$, alors $BA = AC = 0$.

La matrice A est donc un diviseur de zéro (l'une des deux égalités $BA = 0$ ou $AC = 0$ suffisait).

En revanche, on notera que $BA = 0$ mais que $AB = \begin{pmatrix} 2 & -2 \\ 2 & -2 \end{pmatrix}$ n'est pas la matrice nulle.

On montrera que s'il existe $B \neq 0$ telle que $AB = 0$ alors, il existe $C \neq 0$ telle que $CA = 0$ (et réciproquement). Il n'y a donc pas de matrice qui soit diviseur de zéro « d'un seul côté ».

Proposition 16.3.7 (nilpotent \Rightarrow diviseur de zéro)

Soit A une matrice de $\mathcal{M}_n(\mathbb{K})$.

Si A est nilpotente, alors A est un diviseur de zéro. La réciproque est fausse.

Bien sûr si A est nilpotente d'indice $p \geq 1$, il suffit d'écrire $0 = A^p = AA^{p-1}$ (avec $A^{p-1} \neq 0$) pour réaliser que A est un diviseur de zéro.

La matrice $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ est un diviseur de zéro mais elle n'est pas nilpotente.

Non simplifiabilité pour le produit, pour une matrice diviseur de zéro

Si A est un diviseur de zéro, les implications $\begin{cases} AM = AN \Rightarrow M = N \\ MA = NA \Rightarrow M = N \end{cases}$ sont fausses.

En effet, si $AB = 0$ avec $B \neq 0$, et si $N = M + B$, alors $AM = AN$ (bien que M et N soient distinctes).

On exprime cela en disant qu'une matrice diviseur de zéro n'est pas simplifiable pour le produit.

16.3.4 Calcul des puissances d'une matrice carrée

On présente ici quelques méthodes usuelles de calcul de A^p avec A dans $\mathcal{M}_n(\mathbb{K})$ et p dans \mathbb{N} .

▷ Utilisation de la formule du binôme

Pour calculer A^p , il est parfois possible d'écrire $A = B + C$, et d'utiliser la formule du binôme. La condition indispensable est bien sûr l'égalité $BC = CB$.

Il faut que le calcul des puissances de B et C soit beaucoup plus facile que celui de A !

Un cas fréquent est celui où $B = \lambda I_n$ (matrice scalaire) et où C est nilpotente.

Si $C^m = 0$, alors, pour tout p : $A^p = \sum_{k=0}^p \binom{p}{k} C^k B^{p-k} = \sum_{k=0}^p \binom{p}{k} \lambda^{p-k} C^k = \sum_{k=0}^{m-1} \binom{p}{k} \lambda^{p-k} C^k$.

Par exemple si $C^3 = 0$, alors pour $p \geq 3$ on a : $(I+C)^p = I + pC + \frac{p(p-1)}{2}C^2 + \frac{p(p-1)(p-2)}{6}C^3$.

NB : cette formule reste vraie si $0 \leq p \leq 2$ (car les termes « excédentaires » sont nuls).

Premier exemple

Soit $A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$. La matrice A s'écrit $A = I_4 + T$ avec $T = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$.

On a facilement : $T^2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$, $T^3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ puis $T^k = 0$ si $k \geq 4$.

On en déduit : $A^p = I + pT + \frac{p(p-1)}{2}T^2 + \frac{p(p-1)(p-2)}{6}T^3 = \begin{pmatrix} 1 & p & \frac{p(p-1)}{2} & \frac{p(p-1)(p-2)}{6} \\ 0 & 1 & p & \frac{p(p-1)}{2} \\ 0 & 0 & 1 & p \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

Deuxième exemple

Voici un autre d'exemple d'utilisation de la formule du binôme.

La matrice $A = \begin{pmatrix} 2 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 2 \end{pmatrix}$ s'écrit $A = I_4 + J$ avec $J = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$.

On trouve facilement $J^2 = 4J$, donc $J^k = 4^{k-1}J$ pour $k \geq 1$ (mais attention, c'est faux si $k = 0$).

On obtient finalement (bien vérifier toutes les étapes du calcul) :

$$A^p = \sum_{k=0}^p \binom{p}{k} T^k = I_4 + \frac{1}{4} \left(\sum_{k=1}^p \binom{p}{k} 4^k \right) J = I_4 + \frac{5^p - 1}{4} J = \frac{1}{4} \begin{pmatrix} 5^p + 3 & 5^p - 1 & 5^p - 1 & 5^p - 1 \\ 5^p - 1 & 5^p + 3 & 5^p - 1 & 5^p - 1 \\ 5^p - 1 & 5^p - 1 & 5^p + 3 & 5^p - 1 \\ 5^p - 1 & 5^p - 1 & 5^p - 1 & 5^p + 3 \end{pmatrix}$$

▷ Utilisation d'une récurrence

Il arrive que les coefficients des premières puissances de A satisfassent à une formule simple.

Il reste à établir si cette formule est vraie pour toutes les puissances de A , à l'aide d'une récurrence.

Par exemple, si $R(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$ (rotation d'angle θ), on trouve $R(\theta)R(\varphi) = R(\theta + \varphi)$.

Une récurrence évidente donne alors $R(\theta)^p = R(p\theta) = \begin{pmatrix} \cos(p\theta) & -\sin(p\theta) \\ \sin(p\theta) & \cos(p\theta) \end{pmatrix}$ pour tout p .

▷ Utilisation d'un polynôme annulateur

Supposons par exemple qu'une matrice A vérifie $A^3 - 2A^2 - A + 2I_n = 0$ (1)

On dit que $P = X^3 - 2X^2 - X + 2$ est un *polynôme annulateur* de A , l'égalité (1) s'écrivant $P(A) = 0$. Mais le polynôme P se factorise en $P = (X - 1)(X + 1)(X - 2)$.

Pour calculer A^p , on écrit la division euclidienne $X^p = Q(X)P(X) + a_pX^2 + b_pX + c_p$ (2).

Dans cette division, on remplace X par 1, -1 et 2 et on trouve

$$\begin{cases} a_p + b_p + c_p = 1 \\ a_p - b_p + c_p = (-1)^p \\ 4a_p + 2b_p + c_p = 2^p \end{cases}$$

On trouve alors : $a_p = -\frac{1}{2} + \frac{(-1)^p}{6} + \frac{2^p}{3}$, $b_p = \frac{1}{2} - \frac{(-1)^p}{2}$ et $c_p = 1 + \frac{(-1)^p}{3} - \frac{2^p}{3}$.

Dans la division (2), on remplace X par A et on trouve : $A^p = Q(A)P(A) + a_pA^2 + b_pA + c_pI_n$.

Du fait que $P(A) = 0$, cela se simplifie en $A^p = a_pA^2 + b_pA + c_pI_n$.

On a ainsi obtenu A^p en fonction de A^2 , de A et de la matrice identité.

▷ Puissances entières avec Python

La fonction (rudimentaire) suivante calcule les puissances A^k d'une matrice A , avec $k \geq 1$.

```
def powmat(A, k):
    B = A
    for i in range(1, k): B = B.dot(A)
    return B
```

```
>>> A = np.tri(4, dtype='int')
>>> A
array([[1, 0, 0, 0],
       [1, 1, 0, 0],
       [1, 1, 1, 0],
       [1, 1, 1, 1]])
```

```
>>> powmat(A, 2)
array([[1, 0, 0, 0],
       [2, 1, 0, 0],
       [3, 2, 1, 0],
       [4, 3, 2, 1]])
>>>
```

```
>>> powmat(A, 10)
array([[ 1,  0,  0,  0],
       [10,  1,  0,  0],
       [55, 10,  1,  0],
       [220, 55, 10,  1]])
>>>
```

16.3.5 Matrices inversibles

Définition 16.3.3 (le groupe des matrices carrées inversibles d'ordre n)

On note $GL_n(\mathbb{K})$ l'ensemble des matrices de $\mathcal{M}_n(\mathbb{K})$ qui sont inversibles pour le produit.

L'ensemble $GL_n(\mathbb{K})$ est un groupe (non commutatif si $n \geq 2$) pour le produit des matrices.

En fait $GL_n(\mathbb{K})$ est le groupe des inversibles de l'anneau $\mathcal{M}_n(\mathbb{K})$ (voir 11.3.1).

On l'appelle souvent « le groupe linéaire d'indice n ».

Rappel : si A et B sont inversibles, alors AB est inversible et $(AB)^{-1} = B^{-1}A^{-1}$ (l'ordre est important!).

Rappelons que dire que A est inversible dans $\mathcal{M}_n(\mathbb{K})$, c'est dire qu'il existe B telle que $AB = BA = I_n$.

Le produit de $\mathcal{M}_n(\mathbb{K})$ n'étant pas commutatif, il semble indispensable de vérifier $AB = I_n$ et $BA = I_n$.

La proposition 16.3.9 va nous montrer que l'une de ces deux égalités implique l'autre.

Une formule pour l'inverse d'une matrice carrée d'ordre 2

Il n'est pas toujours facile de voir au premier coup d'œil si une matrice carrée A est inversible ou non. En revanche, c'est facile pour les matrices carrées d'ordre 2 :

Proposition 16.3.8 (inversibilité d'une matrice carrée d'ordre 2)

Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ dans $\mathcal{M}_2(\mathbb{K})$. Soit $\Delta = ad - bc$ le « déterminant » de A .

Alors A est inversible si et seulement si Δ est non nul. Dans ce cas, $A^{-1} = \frac{1}{\Delta} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

Par exemple, avec $A = \begin{pmatrix} 1 & 4 \\ 2 & 10 \end{pmatrix}$, on a $\Delta = 2$, donc $A^{-1} = \frac{1}{2} \begin{pmatrix} 10 & -4 \\ -2 & 1 \end{pmatrix}$.

On reprend cet exemple avec Python, le temps de mettre en garde contre une erreur facile à commettre ! L'expression Python $1/A$ désigne en effet non pas l'inverse A^{-1} de A , mais la matrice dont les coefficients sont les inverses, terme à terme, de ceux de A .

Pour calculer l'inverse de A avec Python, il faut utiliser la syntaxe `np.linalg.inv(A)`.

```
>>> A = np.array([[1,4],[2,10]])
>>> print(A)
[[ 1  4]
 [ 2 10]]
```

```
>>> print(1/A)
[[ 1.   0.25]
 [ 0.5  0.1 ]]
```

```
>>> B = np.linalg.inv(A)
>>> print(B)
[[ 5.  -2. ]
 [-1.  0.5]]
```

Avec la matrice B ainsi obtenue, on vérifie qu'on a bien les égalités $AB = BA = I_2$. La matrice B et les produits AB et BA sont obtenus en mode 'float'.

```
>>> A.dot(B)
array([[ 1.,  0.],
       [ 0.,  1.]])
```

```
>>> B.dot(A)
array([[ 1.,  0.],
       [ 0.,  1.]])
```

Inverser une matrice à coefficients rationnels avec Python

Si une matrice carrée inversible A est à coefficients entiers (ou plus généralement rationnels), alors son inverse est à coefficients rationnels.

Avec Python, le calcul de l'inverse de A provoque un passage en mode float.

Il est alors assez délicat de retrouver l'expression exacte (sous forme de rationnels) de la matrice A^{-1} . C'est tout de même possible, en utilisant la fonction suivante dont le rôle est de convertir les coefficients flottants d'une matrice en leur équivalent rationnel (avec une précision donnée) :

```
>>> def tofrac(M, d=1000): # ici d est le dénominateur maximum
    from fractions import Fraction
    return [[Fraction(y).limit_denominator(d) for y in x] for x in M]
```

Par exemple, on définit une matrice A d'ordre 3, on calcule son inverse, et on convertit le résultat avec notre fonction `tofrac` :

```

>>> A = np.array([[1,3,2],[4,1,6],[3,2,5]]); A      # une matrice A d'ordre 3
array([[1, 3, 2],
       [4, 1, 6],
       [3, 2, 5]])

>>> B = np.linalg.inv(A); B                        # son inverse, en mode float
array([[ 2.33333333,  3.66666667, -5.33333333],
       [ 0.66666667,  0.33333333, -0.66666667],
       [-1.66666667, -2.33333333,  3.66666667]])

>>> tofrac(B)                                     # convertit le résultat en rationnels
[[Fraction(7, 3), Fraction(11, 3), Fraction(-16, 3)],
 [Fraction(2, 3), Fraction(1, 3), Fraction(-2, 3)],
 [Fraction(-5, 3), Fraction(-7, 3), Fraction(11, 3)]]

```

Le résultat précédent nous dit donc que si $A = \begin{pmatrix} 1 & 3 & 2 \\ 4 & 1 & 6 \\ 3 & 2 & 5 \end{pmatrix}$, alors $A^{-1} = \frac{1}{3} \begin{pmatrix} 7 & 11 & -16 \\ 2 & 1 & -2 \\ -5 & -7 & 11 \end{pmatrix}$.

On en vient maintenant à plusieurs caractérisations équivalentes de l'inversibilité d'une matrice carrée :

Proposition 16.3.9 (caractérisation des matrices inversibles)

Soit A une matrice de $\mathcal{M}_n(\mathbb{K})$, avec $n \geq 1$.

Les conditions suivantes sont équivalentes :

- la matrice A est inversible dans $\mathcal{M}_n(\mathbb{K})$
- il existe une matrice B de $\mathcal{M}_n(\mathbb{K})$ telle que $AB = I_n$ (on a alors $B = A^{-1}$)
- il existe une matrice B de $\mathcal{M}_n(\mathbb{K})$ telle que $BA = I_n$ (on a alors $B = A^{-1}$)
- la seule matrice-colonne X telle que $AX = 0$ est la colonne nulle
- la seule matrice-ligne X telle que $XA = 0$ est la ligne nulle
- la matrice A n'est pas un diviseur de zéro

Proposition 16.3.10 (inversibilité des matrices triangulaires)

Soit A une matrice triangulaire supérieure (resp. inférieure).

Alors est inversible si et seulement si ses coefficients diagonaux a_{ii} sont non nuls.

Son inverse A^{-1} est alors triangulaire supérieure (resp. inférieure).

De plus les coefficients diagonaux de A^{-1} sont les inverses des a_{ii} .

- Cas particulier : une matrice diagonale D est inversible si et seulement si ses coefficients diagonaux d_{ii} sont tous non nuls. La matrice D^{-1} est alors la matrice diagonale dont les coefficients diagonaux sont les inverses respectifs des d_{ii} .
- Si la matrice A est triangulaire inversible, alors la propriété disant que A^p est triangulaire (et du même côté que A) avec pour coefficients diagonaux les a_{ii}^p est encore valable pour les exposants négatifs.

Soit A dans $\mathcal{M}_n(\mathbb{K})$. Notons C_1, C_2, \dots, C_n les colonnes de A , et L_1, L_2, \dots, L_n les lignes de A .

$$\text{Avec ces notations, on peut écrire : } A = \left(\begin{array}{c|c|c|c} L_1 & & & \\ \hline L_2 & & & \\ \vdots & & & \\ \hline L_n & & & \end{array} \right) = \left(C_1 \mid C_2 \mid \dots \mid C_n \right)$$

On sait (voir sous-section 16.2.5) que si les λ_k sont des scalaires, on peut écrire :

$$\left(C_1 \mid C_2 \mid \dots \mid C_n \right) \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} = \sum_{j=1}^n \lambda_j C_j \quad \text{et} \quad (\lambda_1 \ \lambda_2 \ \dots \ \lambda_n) \begin{pmatrix} L_1 \\ L_2 \\ \vdots \\ L_n \end{pmatrix} = \sum_{i=1}^n \lambda_i L_i$$

On peut alors écrire de nouvelles caractérisations de l'inversibilité d'une matrice carrée :

Proposition 16.3.11 (caractérisation de l'inversibilité d'une matrice carrée)

Soit A une matrice de $\mathcal{M}_n(\mathbb{K})$, avec $n \geq 1$. Les conditions suivantes sont équivalentes :

- la matrice A est inversible dans $\mathcal{M}_n(\mathbb{K})$
- les lignes de A sont linéairement indépendantes
- les colonnes de A sont linéairement indépendantes

La proposition précédente permet donc de caractériser la non inversibilité !

Proposition 16.3.12 (caractérisation de la non inversibilité d'une matrice carrée)

Soit A une matrice de $\mathcal{M}_n(\mathbb{K})$, avec $n \geq 1$. Les conditions suivantes sont équivalentes :

- la matrice A n'est pas inversible dans $\mathcal{M}_n(\mathbb{K})$
- les lignes de A forment une famille liée
- les colonnes de A forment une famille liée

Voici donc des conditions *suffisantes* pour dire, au premier coup d'œil, qu'une matrice est non inversible :

- si elle contient une ligne nulle, ou une colonne nulle
- si elle contient deux lignes (ou deux colonnes) proportionnelles

Considérons par exemple $A = \begin{pmatrix} 1 & 3 & 4 \\ 2 & 1 & 3 \\ 3 & 2 & 5 \end{pmatrix}$. On « voit » que ses colonnes sont liées car $C_3 = C_1 + C_2$.

La matrice A n'est donc pas inversible. Plus précisément $AX = 0$, avec $X = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$.

Puisque A n'est pas inversible, ses lignes L_1, L_2, L_3 doivent être liées.

Par le calcul, on trouve $5L_3 = L_1 + 7L_2$, ce qui était difficile à voir « au premier coup d'œil » !

16.3.6 Calcul de l'inverse d'une matrice inversible

▷ Inversion d'une matrice par résolution d'un système

Soit A un élément de $\mathcal{M}_n(\mathbb{K})$, et soit X, B deux matrices-colonne de hauteur n .

Si le système $AX = B$, d'inconnue X , possède une solution unique alors A est inversible.

La résolution de ce système fournit d'ailleurs A^{-1} car $AX = B \Leftrightarrow X = A^{-1}B$.

Considérons par exemple la matrice $A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$.

On a les équivalences :

$$A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \Leftrightarrow \begin{cases} y + z = a \\ x + y + z = b \\ x + z = c \end{cases} \Leftrightarrow \begin{cases} x = -a + b \\ y = b - c \\ z = a - b + c \end{cases} \Leftrightarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -1 & 1 & 0 \\ 0 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

Ce calcul prouve que A est inversible et que $A^{-1} = \begin{pmatrix} -1 & 1 & 0 \\ 0 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix}$.

▷ Utilisation d'un polynôme annulateur

Supposons par exemple qu'une matrice A vérifie $A^3 - 2A^2 - A + 2I_n = 0$ (1)

On dit que $P = X^3 - 2X^2 - X + 2$ est un *polynôme annulateur* de A , l'égalité (1) s'écrivant $P(A) = 0$.

Mais (1) s'écrit $A(A^2 - 2A + 2I_n) = -2I_n$ et prouve que A est inversible avec $A^{-1} = -\frac{1}{2}(A^2 - 2A + 2I_n)$.

Autre exemple : avec $A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & -1 \end{pmatrix}$, on a $A^2 = \begin{pmatrix} 4 & 0 & 0 & -2 \\ 0 & 4 & 0 & 2 \\ 0 & 0 & 4 & 2 \\ -2 & 2 & 2 & 4 \end{pmatrix}$ et $A^3 = \begin{pmatrix} 2 & 6 & 6 & 6 \\ 6 & 2 & -6 & -6 \\ 6 & -6 & 2 & -6 \\ 6 & -6 & -6 & -10 \end{pmatrix}$.

On remarque que $A^3 = 6A - 4I_4$. Ainsi $A(6I_4 - A^2) = 4I_4$.

On en déduit que A est inversible et que $A^{-1} = \frac{1}{4}(6I_4 - A^2) = \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$.

▷ Exposants négatifs

Supposons qu'on ait trouvé une formule donnant $A^p = \varphi(p)$ en fonction de l'entier $p \geq 0$. On peut chercher à prouver que cette formule est encore valable pour les exposants négatifs, à condition que A soit inversible. Il suffit alors de prouver que pour tout p de \mathbb{N} , $\varphi(p)\varphi(-p) = I_n$.

Soit par exemple $A = \begin{pmatrix} 1 & 2\lambda & 0 \\ 0 & 1 & 2\lambda \\ 0 & 0 & 1 \end{pmatrix} = I + 2\lambda T$, avec $T = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$, $T^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ et $T^3 = 0$.

La formule du binôme donne $A^p = \varphi(p) = I + 2p\lambda T + 2p(p-1)\lambda^2 T^2 = \begin{pmatrix} 1 & 2p\lambda & 2p(p-1)\lambda^2 \\ 0 & 1 & 2p\lambda \\ 0 & 0 & 1 \end{pmatrix}$

Mais $\varphi(p)\varphi(-p) = (I + 2p\lambda T + 2p(p-1)\lambda^2 T^2)(I - 2p\lambda T + 2p(p+1)\lambda^2 T^2) = I$ (grâce à $T^3 = 0$).

Ainsi $\varphi(-p) = (\varphi(p))^{-1}$, donc la formule $A^p = \varphi(p)$ est encore valable pour les exposants négatifs.

16.4 Transposition

16.4.1 Propriétés de la transposition

Définition 16.4.1 (transposée d'une matrice)

Soit A une matrice de $\mathcal{M}_{n,p}(\mathbb{K})$, de terme général $a_{i,j}$.

On appelle *transposée* de A et on note A^\top la matrice B de $\mathcal{M}_{p,n}(\mathbb{K})$ de terme général $b_{i,j} = a_{j,i}$.

Remarque : on peut utiliser les deux notations A^\top ou A^t .

$$\text{Si par exemple } A = \begin{pmatrix} 1 & 4 & 2 & 3 \\ 8 & 4 & 3 & 6 \\ 7 & 1 & 0 & 5 \end{pmatrix}, \text{ alors } A^\top = \begin{pmatrix} 1 & 8 & 7 \\ 4 & 4 & 1 \\ 2 & 3 & 0 \\ 3 & 6 & 5 \end{pmatrix}.$$

Voici comment on peut transposer une matrice avec Python. On notera les deux syntaxes possibles.

```
>>> A = np.array([[1,4,2,3],[8,4,3,6],[7,1,0,5]]);
>>> A
array([[1, 4, 2, 3],
       [8, 4, 3, 6],
       [7, 1, 0, 5]])

>>> A.transpose() # ou A.T
array([[1, 8, 7],
       [4, 4, 1],
       [2, 3, 0],
       [3, 6, 5]])
```

Proposition 16.4.1 (linéarité et bijectivité de la transposition)

La transposition des matrices induit un isomorphisme de $\mathcal{M}_{n,p}(\mathbb{K})$ sur $\mathcal{M}_{p,n}(\mathbb{K})$.

Plus précisément, si $n = p$, la transposition est un automorphisme involutif de $\mathcal{M}_n(\mathbb{K})$.

On retiendra que si A, B ont le même format, et pour tous λ, μ de \mathbb{K} , on a :

$$\begin{cases} (\lambda A + \mu B)^\top = \lambda A^\top + \mu B^\top \\ (A^\top)^\top = A \end{cases}$$

Proposition 16.4.2 (transposée d'un produit de deux matrices)

Pour A dans $\mathcal{M}_{n,p}(\mathbb{K})$, et B dans $\mathcal{M}_{p,q}(\mathbb{K})$, on a l'égalité : $(AB)^\top = B^\top A^\top$ (Attention à l'ordre !)

On peut généraliser à un produit de k matrices : $({}^\top A_1 A_2 \cdots A_k) = A_k^\top A_{k-1}^\top \cdots A_2^\top A_1^\top$.

Proposition 16.4.3 (transposition de l'inverse d'une matrice carrée)

Si A est une matrice carrée inversible, alors la matrice A^\top est inversible et $(A^\top)^{-1} = (A^{-1})^\top$.

On peut retenir que l'inverse de la transposée, c'est la transposée de l'inverse.

Proposition 16.4.4

Pour toute matrice A de $\mathcal{M}_n(\mathbb{K})$ et tout entier naturel k , on a : $(A^k)^\top = (A^\top)^k$.

Si A est inversible, cette égalité s'étend aux entiers strictement négatifs.

16.4.2 Matrices symétriques ou antisymétriques

Définition 16.4.2 (matrices carrées symétriques)

Une matrice $A = (a_{i,j})$ de $\mathcal{M}_n(\mathbb{K})$ est dite *symétrique* si elle vérifie $A^\top = A$.

Cela équivaut donc à dire que $a_{j,i} = a_{i,j}$ pour tous indices i et j .

Dire que A est symétrique, c'est donc dire qu'elle est « symétrique par rapport à sa diagonale ».

Une matrice symétrique A est définie de façon unique par la donnée de ses coefficients $a_{i,j}$ avec $j \geq i$ (c'est-à-dire qui sont « sur » ou « au-dessus » de la diagonale).

Définition 16.4.3 (matrices carrées antisymétriques)

Une matrice $A = (a_{i,j})$ de $\mathcal{M}_n(\mathbb{K})$ est dite *antisymétrique* si $A^\top = -A$.

Cela équivaut donc à dire que $a_{j,i} = -a_{i,j}$ pour tous indices i et j .

Si A est antisymétrique, alors ses coefficients diagonaux sont nuls.

Une matrice antisymétrique A est définie de façon unique par la donnée de ses coefficients $a_{i,j}$ avec $j > i$ (c'est-à-dire qui sont « strictement au-dessus » de la diagonale).

Par exemple, $A = \begin{pmatrix} 4 & 1 & 0 & 8 \\ 1 & 3 & 2 & 5 \\ 0 & 2 & 7 & 6 \\ 8 & 5 & 6 & 3 \end{pmatrix}$ est symétrique, et $B = \begin{pmatrix} 0 & -1 & 3 & 6 \\ 1 & 0 & 2 & -4 \\ -3 & -2 & 0 & 7 \\ -6 & 4 & -7 & 0 \end{pmatrix}$ est antisymétrique.

Cela n'a pas de sens de parler de symétrie ou d'antisymétrie pour une matrice qui n'est pas carrée!

Quelques propriétés

- si A est inversible et symétrique alors A^{-1} est symétrique.
- si A est inversible et antisymétrique alors A^{-1} est antisymétrique.
- si A est symétrique alors A^k est symétrique pour tout k de \mathbb{N} (et k dans \mathbb{Z} si A est inversible).
- si A est antisymétrique, alors A^k est symétrique si k est pair et antisymétrique si k est impair.

Notation

On note $\mathcal{S}_n(\mathbb{K})$ l'ensemble des matrices de $\mathcal{M}_n(\mathbb{K})$ qui sont symétriques.

On note $\mathcal{A}_n(\mathbb{K})$ l'ensemble des matrices de $\mathcal{M}_n(\mathbb{K})$ qui sont antisymétriques.

Proposition 16.4.5 (sous-espaces des matrices symétriques ou antisymétriques)

Les ensembles $\mathcal{S}_n(\mathbb{K})$ et $\mathcal{A}_n(\mathbb{K})$ sont deux sous-espaces supplémentaires de $\mathcal{M}_n(\mathbb{K})$.

La dimension de $\mathcal{S}_n(\mathbb{K})$ est $\frac{1}{2}n(n+1)$ et celle de $\mathcal{A}_n(\mathbb{K})$ est $\frac{1}{2}n(n-1)$.

Toute matrice M de $\mathcal{M}_n(\mathbb{K})$ s'écrit de manière unique $M = S + A$, avec S dans $\mathcal{S}_n(\mathbb{K})$ et A dans $\mathcal{A}_n(\mathbb{K})$.

Les matrices S et A sont données par : $S = \frac{1}{2}(M + M^\top)$ et $A = \frac{1}{2}(M - M^\top)$.

Les fonctions `sym` et `asym` renvoient les composantes symétrique et antisymétrique d'une matrice M :

```
|| >>> def sym(M): return (M+M.T)/2
```

```
|| >>> def asym(M): return (M-M.T)/2
```

```

>>> M=np.array([[1,3,4],[2,7,6],[4,1,2]])
>>> print(M)
[[1 3 4]
 [2 7 6]
 [4 1 2]]

>>> S = sym(M)
>>> print(S)
[[ 1.  2.5  4. ]
 [ 2.5  7.  3.5]
 [ 4.  3.5  2. ]]

>>> A = asym(M)
>>> print(A)
[[ 0.  0.5  0. ]
 [-0.5  0.  2.5]
 [ 0. -2.5  0. ]]

```

Les deux fonctions suivantes (`randsym` et `randasym`) permettent de fabriquer des matrices symétriques ou antisymétriques aléatoires d'ordre n (avec $n = 4$ par défaut). Elles forment en fait, respectivement, la composante symétrique ou antisymétrique d'une matrice aléatoire d'ordre n à coefficients entiers dans un intervalle donné $[a, b]$ (avec $[a, b] = [-5, 5]$ par défaut).

Ici les matrices renvoyées sont dans le type `int`.

Attention à la « division par 2 » qu'il faut forcer dans le type `float` (on utilise donc `/` et non pas `//`) avant de revenir dans le type `int` avec la fonction `astype` :

```

>>> def randsym(n=4, a=-5, b=5):
    m = np.random.randint(a, b+1, size=(n,n))
    return ((m+m.T)/2).astype('int')

>>> def randasym(n=4, a=-5, b=5):
    m = np.random.randint(a, b+1, size=(n,n))
    return ((m-m.T)/2).astype('int')

```

Voici quelques exemples :

```

>>> randsym()
array([[ -1,  -2,   0,   0],
       [ -2,  -5,   4,  -1],
       [  0,   4,   5,   1],
       [  0,  -1,   1,  -1]])

>>> randasym()
array([[ 0,  2,  3,  1],
       [-2,  0,  1,  0],
       [-3, -1,  0,  3],
       [-1,  0, -3,  0]])

>>> randsym(a=0,b=1)
array([[0, 0, 1, 0],
       [0, 1, 1, 0],
       [1, 1, 0, 1],
       [0, 0, 1, 0]])

```

```

>>> randsym(6,-9,9)
array([[ 7,  2,  5,  5, -1,  0],
       [ 2,  1, -3, -1,  0, -4],
       [ 5, -3,  6, -2, -6, -1],
       [ 5, -1, -2, -3, -6, -5],
       [-1,  0, -6, -6, -9, -1],
       [ 0, -4, -1, -5, -1, -2]])

>>> randasym(6,-99,99)
array([[ 0,  4, -62, -5, -45, 27],
       [-4,  0, -30, -70, 24, 38],
       [62, 30,  0, -33, 15, -61],
       [ 5, 70, 33,  0,  8, -60],
       [45, -24, -15, -8,  0, 28],
       [-27, -38, 61, 60, -28,  0]])

```

16.5 Calculs par blocs

16.5.1 Décompositions en blocs

▷ Commençons par quelques exemples

La matrice $A = \begin{pmatrix} 2 & 5 & 3 & 1 & 6 & 7 \\ 4 & 1 & 5 & 6 & 8 & 9 \\ 1 & 0 & 4 & 5 & 1 & 3 \\ 7 & 9 & 6 & 2 & 1 & 5 \\ 1 & 5 & 8 & 2 & 7 & 4 \end{pmatrix}$ est « décomposable en blocs » de plusieurs manières, par exemple :

$$A = \left(\begin{array}{ccc|ccc} 2 & 5 & 3 & 1 & 6 & 7 \\ 4 & 1 & 5 & 6 & 8 & 9 \\ 1 & 0 & 4 & 5 & 1 & 3 \\ 7 & 9 & 6 & 2 & 1 & 5 \\ 1 & 5 & 8 & 2 & 7 & 4 \end{array} \right) = \begin{pmatrix} M & N \\ P & Q \end{pmatrix}, \text{ où } M = \begin{pmatrix} 2 & 5 & 3 \\ 4 & 1 & 5 \end{pmatrix}, N = \begin{pmatrix} 1 & 6 & 7 \\ 6 & 8 & 9 \end{pmatrix}, P = \begin{pmatrix} 1 & 0 & 4 \\ 7 & 9 & 6 \\ 1 & 5 & 8 \end{pmatrix}, Q = \begin{pmatrix} 5 & 1 & 3 \\ 2 & 1 & 5 \\ 2 & 7 & 4 \end{pmatrix}$$

$$\text{On pourrait aussi décomposer en } A = \left(\begin{array}{cccc|cc} 2 & 5 & 3 & 1 & 6 & 7 \\ 4 & 1 & 5 & 6 & 8 & 9 \\ 1 & 0 & 4 & 5 & 1 & 3 \\ 7 & 9 & 6 & 2 & 1 & 5 \\ 1 & 5 & 8 & 2 & 7 & 4 \end{array} \right), \text{ ou même en } A = \left(\begin{array}{cc|ccc|c} 2 & 5 & 3 & 1 & 6 & 7 \\ 4 & 1 & 5 & 6 & 8 & 9 \\ 1 & 0 & 4 & 5 & 1 & 3 \\ 7 & 9 & 6 & 2 & 1 & 5 \\ 1 & 5 & 8 & 2 & 7 & 4 \end{array} \right)$$

▷ Décompositions en blocs avec Python

On reprend la matrice initiale A , et on forme les matrices M, P, N, Q comme indiqué ci-dessus.

Rappelons le principe du « slicing » qui dit que l'expression $A[a:b;c:d]$ renvoie le bloc formé des lignes dont l'indice est dans $[a, b[$ et des colonnes dont l'indice est dans $[c, d[$.

Dans cette syntaxe, le premier indice vaut 0 par défaut, et le deuxième vaut $+\infty$ par défaut.

```
>>> A
array([[2, 5, 3, 1, 6, 7],
       [4, 1, 5, 6, 8, 9],
       [1, 0, 4, 5, 1, 3],
       [7, 9, 6, 2, 1, 5],
       [1, 5, 8, 2, 7, 4]])
>>>
```

```
>>> M = A[:2, :3]; M
array([[2, 5, 3],
       [4, 1, 5]])
>>> P = A[2:, :3]; P
array([[1, 0, 4],
       [7, 9, 6],
       [1, 5, 8]])
```

```
>>> N = A[:2, 3:]; N
array([[1, 6, 7],
       [6, 8, 9]])
>>> Q = A[2:, 3:]; Q
array([[5, 1, 3],
       [2, 1, 5],
       [2, 7, 4]])
```

Voici ensuite deux façons de reconstituer A à partir des blocs M, N, P, Q :

```
>>> np.bmat([[M,N], [P,Q]])
matrix([[2, 5, 3, 1, 6, 7],
        [4, 1, 5, 6, 8, 9],
        [1, 0, 4, 5, 1, 3],
        [7, 9, 6, 2, 1, 5],
        [1, 5, 8, 2, 7, 4]])
```

```
>>> np.bmat('M N; P Q')
matrix([[2, 5, 3, 1, 6, 7],
        [4, 1, 5, 6, 8, 9],
        [1, 0, 4, 5, 1, 3],
        [7, 9, 6, 2, 1, 5],
        [1, 5, 8, 2, 7, 4]])
```

Attention : les résultats sont au format « matrix », une sous-classe de « array ». Il y a quelques différences de syntaxe (notamment pour le produit matriciel), et il est préférable de s'en tenir au format `array`.

On pouvait par exemple écrire : `np.asarray(np.bmat('M N; P Q'))`

▷ Forme générale d'une décomposition en blocs

Plus généralement, soit A une matrice quelconque de $\mathcal{M}_{n,p}(\mathbb{K})$.

$$\text{On peut écrire } A = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1k} & \dots & A_{1q} \\ \vdots & \vdots & & \vdots & & \vdots \\ A_{i1} & A_{i2} & \dots & A_{ik} & \dots & A_{iq} \\ \vdots & \vdots & & \vdots & & \vdots \\ A_{m1} & A_{m2} & \dots & A_{mk} & \dots & A_{mq} \end{pmatrix} \text{ où les } A_{i,j} \text{ sont elles-mêmes des matrices.}$$

On a ainsi décomposé A en mq blocs.

La contrainte est que chaque bloc $A_{i,j}$ soit de taille (n_i, p_j) , avec $\sum_{i=1}^m n_i = n$, et $\sum_{j=1}^q p_j = p$.

Certaines décompositions par blocs reviennent plus souvent que d'autres.

Une matrice A peut par exemple être décomposée en quatre blocs : $A = \begin{pmatrix} M & N \\ P & Q \end{pmatrix}$

▷ Décompositions en lignes ou en colonnes

Soit A dans $\mathcal{M}_{n,p}(\mathbb{K})$. Notons C_1, C_2, \dots, C_p les colonnes de A , et L_1, L_2, \dots, L_n les lignes de A .

Les écritures $A = \begin{pmatrix} \text{---} L_1 \text{---} \\ \text{---} L_2 \text{---} \\ \vdots \\ \text{---} L_n \text{---} \end{pmatrix} = \left(C_1 \mid C_2 \mid \cdots \mid C_p \right)$ sont des décompositions particulières de A .

▷ Matrices diagonales ou triangulaires par blocs

Désignons par D_1, D_2, \dots, D_m une succession de m matrices carrées (pas nécessairement de même taille, certaines d'entre elles pouvant même être réduites à un seul scalaire).

On note également $*$ des blocs matriciels quelconques.

Si $A = \begin{pmatrix} D_1 & 0 & \cdots & 0 \\ 0 & D_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & D_m \end{pmatrix}$ et $B = \begin{pmatrix} D_1 & * & \cdots & * \\ 0 & D_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & D_m \end{pmatrix}$ on dit que $\begin{cases} A \text{ est « diagonale par blocs »} \\ B \text{ est « triangulaire par blocs »} \end{cases}$

16.5.2 Opérations sur les décompositions en blocs

▷ Addition par blocs

Soit A et B dans $\mathcal{M}_{n,p}(\mathbb{K})$, décomposées en blocs : $A = \begin{pmatrix} M & N \\ P & Q \end{pmatrix}$ et $B = \begin{pmatrix} M' & N' \\ P' & Q' \end{pmatrix}$

On demande ici que ces deux décompositions soient « compatibles pour la somme », c'est-à-dire que M et M' aient le même format (il en est alors de même pour N et N' , pour P et P' , pour Q et Q').

Si ces conditions sont réunies, alors $\lambda A + \mu B = \begin{pmatrix} \lambda M + \mu M' & \lambda N + \mu N' \\ \lambda P + \mu P' & \lambda Q + \mu Q' \end{pmatrix}$.

Cette propriété se généralise à des décompositions en blocs quelconques, sous réserve bien sûr que les blocs qui se correspondent aient exactement la même taille.

▷ Produit par blocs

Soit A dans $\mathcal{M}_{n,p}(\mathbb{K})$, B dans $\mathcal{M}_{p,q}(\mathbb{K})$, décomposées en blocs : $A = \begin{pmatrix} M & N \\ P & Q \end{pmatrix}$ et $B = \begin{pmatrix} M' & N' \\ P' & Q' \end{pmatrix}$

On souhaite effectuer le produit AB en exploitant cette décomposition.

On voudrait en fait pouvoir écrire : $AB = \begin{pmatrix} M & N \\ P & Q \end{pmatrix} \begin{pmatrix} M' & N' \\ P' & Q' \end{pmatrix} = \begin{pmatrix} MM' + NP' & MN' + NQ' \\ PM' + QP' & PN' + QQ' \end{pmatrix}$

Pour cela, il faut déjà pouvoir effectuer le produit MM' des blocs M et M' .

Supposons par exemple que le bloc M soit de type (m, r) , avec les inégalités $\begin{cases} 1 \leq m \leq n \\ 1 \leq r \leq p \end{cases}$

Supposons également que M' soit de type (r, s) , avec $\begin{cases} 1 \leq r \leq p \\ 1 \leq s \leq q \end{cases}$

Avec ces hypothèses (qui permettent effectivement de calculer MM'), on a :

$$\begin{cases} M \in \mathcal{M}_{m,r}(\mathbb{K}) \\ M' \in \mathcal{M}_{r,s}(\mathbb{K}) \end{cases} \quad \begin{cases} N \in \mathcal{M}_{m,p-r}(\mathbb{K}) \\ P' \in \mathcal{M}_{p-r,s}(\mathbb{K}) \end{cases} \quad \text{donc} \quad \begin{cases} MM' \in \mathcal{M}_{m,s}(\mathbb{K}) \\ NP' \in \mathcal{M}_{m,s}(\mathbb{K}) \end{cases} \quad \text{donc} \quad MM' + NP' \in \mathcal{M}_{m,s}(\mathbb{K}).$$

$$\begin{cases} M \in \mathcal{M}_{m,r}(\mathbb{K}) \\ N' \in \mathcal{M}_{r,q-s}(\mathbb{K}) \end{cases} \quad \begin{cases} N \in \mathcal{M}_{m,p-r}(\mathbb{K}) \\ Q' \in \mathcal{M}_{p-r,q-s}(\mathbb{K}) \end{cases} \quad \text{donc} \quad \begin{cases} MN' \in \mathcal{M}_{m,q-s}(\mathbb{K}) \\ NQ' \in \mathcal{M}_{m,q-s}(\mathbb{K}) \end{cases} \quad \text{donc} \quad MN' + NQ' \in \mathcal{M}_{m,q-s}(\mathbb{K}).$$

$$\begin{cases} P \in \mathcal{M}_{n-m,r}(\mathbb{K}) \\ M' \in \mathcal{M}_{r,s}(\mathbb{K}) \end{cases} \quad \begin{cases} Q \in \mathcal{M}_{n-m,p-r}(\mathbb{K}) \\ P' \in \mathcal{M}_{p-r,s}(\mathbb{K}) \end{cases} \quad \text{donc} \quad \begin{cases} PM' \in \mathcal{M}_{n-m,s}(\mathbb{K}) \\ QP' \in \mathcal{M}_{n-m,s}(\mathbb{K}) \end{cases} \quad \text{donc} \quad PM' + QP' \in \mathcal{M}_{n-m,s}(\mathbb{K}).$$

$$\begin{cases} P \in \mathcal{M}_{n-m,r}(\mathbb{K}) \\ N' \in \mathcal{M}_{r,q-s}(\mathbb{K}) \end{cases} \quad \begin{cases} Q \in \mathcal{M}_{n-m,p-r}(\mathbb{K}) \\ Q' \in \mathcal{M}_{p-r,q-s}(\mathbb{K}) \end{cases} \quad \text{donc} \quad \begin{cases} PN' \in \mathcal{M}_{n-m,q-s}(\mathbb{K}) \\ QQ' \in \mathcal{M}_{n-m,q-s}(\mathbb{K}) \end{cases} \quad \text{donc} \quad PN' + QQ' \in \mathcal{M}_{n-m,q-s}(\mathbb{K}).$$

$$\text{On vérifie alors qu'effectivement : } AB = \begin{pmatrix} M & N \\ P & Q \end{pmatrix} \begin{pmatrix} M' & N' \\ P' & Q' \end{pmatrix} = \begin{pmatrix} MM' + NP' & MN' + NQ' \\ PM' + QP' & PN' + QQ' \end{pmatrix}$$

Cette propriété se généralise à des décompositions en un nombre quelconque de blocs, sous réserve que tous les produits de blocs qui sont nécessaires à la construction du résultat soient possibles :

Proposition 16.5.1 (théorème du produit par blocs)

Soit A dans $\mathcal{M}_{n,p}(\mathbb{K})$, et B dans $\mathcal{M}_{p,q}(\mathbb{K})$.

On suppose que A et B sont décomposées en blocs : $A = (A_{i,k})_{\substack{1 \leq i \leq r \\ 1 \leq k \leq s}}$ et $B = (B_{k,j})_{\substack{1 \leq k \leq s \\ 1 \leq j \leq t}}$.

Alors $M = AB$ se décompose en blocs $M = (M_{i,j})_{\substack{1 \leq i \leq r \\ 1 \leq j \leq t}}$, avec $\begin{cases} \forall i \in \{1, \dots, r\} \\ \forall j \in \{1, \dots, t\} \end{cases} M_{ij} = \sum_{k=1}^s A_{ik} B_{kj}$

▷ Cas des matrices triangulaires par blocs

Soit A et B deux matrices carrées, triangulaires par blocs.

$$\text{On peut donc écrire } A = \begin{pmatrix} D_1 & * & \dots & * \\ 0 & D_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & D_m \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} \Delta_1 & * & \dots & * \\ 0 & \Delta_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & \Delta_m \end{pmatrix}$$

On suppose ici que D_i et Δ_i ont la même taille. Les blocs marqués $*$ sont sans importance.

Avec ces notations, on montre que :

$$AB = \begin{pmatrix} D_1 \Delta_1 & * & \dots & * \\ 0 & D_2 \Delta_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & D_m \Delta_m \end{pmatrix}, \quad A^p = \begin{pmatrix} D_1^p & * & \dots & * \\ 0 & D_2^p & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & D_m^p \end{pmatrix}, \quad A^{-1} = \begin{pmatrix} D_1^{-1} & * & \dots & * \\ 0 & D_2^{-1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & D_m^{-1} \end{pmatrix}$$

La dernière égalité doit être précisée : elle exprime que A est inversible si et seulement si chacun des blocs diagonaux D_i est inversible. L'inverse de A est alors triangulaire par blocs, les blocs diagonaux de A^{-1} étant, respectivement, les inverses des blocs diagonaux de A .

Chapitre 17

Matrices et applications linéaires

Sommaire

17.1 Matrices et applications linéaires	377
17.1.1 Matrice d'une famille de vecteurs dans une base	377
17.1.2 Matrice d'une app ⁿ linéaire dans un couple de bases	377
17.1.3 Coordonnées de l'image d'un vecteur par une app ⁿ linéaire	378
17.1.4 Propriétés opératoires	379
17.2 Changements de bases	381
17.2.1 Matrice de passage d'une base à une autre	381
17.2.2 Effet d'un changement de base(s)	381
17.2.3 Matrices équivalentes et matrices semblables	382
17.2.4 Réduction de la matrice de f à une forme J_r	383
17.2.5 Matrices semblables	384
17.3 Trace d'une matrice, d'un endomorphisme	385
17.3.1 Trace d'une matrice carrée	385
17.3.2 Trace d'un endomorphisme	386
17.4 Noyau, image et rang d'une matrice	386
17.4.1 Application linéaire canoniquement associée	386
17.4.2 Noyau, image et rang d'une matrice	387
17.4.3 Matrices équivalentes et rang	390
17.4.4 Rang et matrices extraites	390
17.5 Calcul effectif du rang	391
17.5.1 Matrices échelonnées	391
17.5.2 Opérations élémentaires	392
17.5.3 Calcul du rang par la méthode du pivot	393
17.5.4 Calcul de l'inverse par la méthode du pivot	394
17.6 Systèmes linéaires	398
17.6.1 Généralités et définitions	398
17.6.2 Interprétations d'un système linéaire	398
17.6.3 Structure de l'ensemble des solutions	400
17.6.4 Systèmes de Cramer	401
17.6.5 Résolution par la méthode du pivot de Gauss	402

17.1 Matrices et applications linéaires

17.1.1 Matrice d'une famille de vecteurs dans une base

Définition 17.1.1 (matrice d'une famille de vecteurs dans une base)

Soit E un \mathbb{K} -espace vectoriel, de $\dim n \geq 1$, muni d'une base $\varepsilon = (\varepsilon_i)_{1 \leq i \leq n}$.

Soit $v = (v_j)_{1 \leq j \leq p}$ une famille de p vecteurs de E . Pour tout j de $\{1, \dots, p\}$, posons $v_j = \sum_{i=1}^n a_{ij} \varepsilon_i$.

La matrice de $\mathcal{M}_{n,p}(\mathbb{K})$ de terme général a_{ij} est appelée *matrice de la famille v dans la base (ε)* .

On pourra noter cette matrice $A = \text{Mat}_\varepsilon(v)$.

La j -ième colonne de A est donc formée des composantes de v_j dans la base ε .

Par exemple, si $n = 3$, $j = 2$, et si $\begin{cases} v_1 = 3\varepsilon_1 + 5\varepsilon_2 + \varepsilon_3 \\ v_2 = 2\varepsilon_1 + 4\varepsilon_2 + 7\varepsilon_3 \end{cases}$ on a $A = \text{Mat}_\varepsilon(v) = \begin{pmatrix} 3 & 2 \\ 5 & 4 \\ 1 & 7 \end{pmatrix}$

On notera $[v]_\varepsilon$ la matrice-colonne des coordonnées d'un vecteur v de E dans la base (ε) .

Si $v = (v_j)_{1 \leq j \leq p}$, on peut donc écrire : $A = M_\varepsilon(v) = \left(\begin{array}{c|c|c|c} [v_1]_\varepsilon & [v_2]_\varepsilon & \cdots & [v_p]_\varepsilon \end{array} \right)$

17.1.2 Matrice d'une appⁿ linéaire dans un couple de bases

Définition 17.1.2 (matrice d'une application linéaire)

Soit E et F deux espaces vectoriels sur \mathbb{K} .

On suppose que $\dim(E) = p \geq 1$, et que E est muni d'une base $(e_j)_{1 \leq j \leq p}$.

On suppose que $\dim(F) = n \geq 1$, et que F est muni d'une base $(\varepsilon_i)_{1 \leq i \leq n}$.

Soit f une application linéaire de E dans F .

On appelle matrice de f dans les bases e et ε la matrice A de la famille $(f(e_j))_{1 \leq j \leq p}$ dans la base ε .

Cette matrice, élément de $\mathcal{M}_{n,p}(\mathbb{K})$, est notée $\text{Mat}_{e,\varepsilon}(f)$.

Pour $1 \leq j \leq p$, la j -ième colonne de $\text{Mat}_{e,\varepsilon}(f)$ est donc formée des composantes de $f(e_j)$ dans ε .

Supposons par exemple qu'une base de E soit $e = (e_1, e_2, e_3)$, et qu'une base de F soit $\varepsilon = (\varepsilon_1, \varepsilon_2)$.

Si $f : E \rightarrow F$ est l'application linéaire définie par $\begin{cases} f(e_1) = \varepsilon_1 + 2\varepsilon_2 \\ f(e_2) = 7\varepsilon_1 + 5\varepsilon_2 \\ f(e_3) = 3\varepsilon_1 \end{cases}$ alors $\text{Mat}_{e,\varepsilon}(f) = \begin{pmatrix} 1 & 7 & 3 \\ 2 & 5 & 0 \end{pmatrix}$

Remarques

- Une application linéaire est complètement déterminée par sa matrice dans un couple de bases donné. Si $\dim(E) = p \geq 1$ et $\dim(F) = n \geq 1$, l'application de $\mathcal{L}(E, F)$ dans $\mathcal{M}_{n,p}(\mathbb{K})$ qui à f associe sa matrice dans un couple de bases donné est donc une bijection.
- Si $f : E \rightarrow F$ est linéaire et si change la base de E ou de F , la matrice de f est, en général, modifiée. On analysera plus loin cette dépendance en fonction du couple de bases. En revanche, la matrice de l'application nulle est la matrice nulle, quel que soit le couple de bases.

- On retiendra que si A est la matrice d'une application linéaire, le nombre de colonnes de A est la dimension de l'espace de départ, et le nombre de lignes de A est la dimension de l'espace d'arrivée.

Cas particulier : matrice d'un endomorphisme dans *une* base

Soit f un endomorphisme de E , où $\dim(E) = n \geq 1$.

On munit souvent E de la même base $e = (e_i)_{1 \leq i \leq n}$ au départ et à l'arrivée.

Plutôt que de noter $\text{Mat}_{e,e}(f)$, on note $\text{Mat}_e(f)$ et on parle de *matrice de f dans la base e* .

Cette matrice est bien sûr carrée d'ordre n .

Par exemple, soit f l'endomorphisme de $\mathbb{R}_3[X]$ défini par $f(P(X)) = (X+1)P'(X) + P(X)$.

On munit $\mathbb{R}_3[X]$ de la base canonique $e = 1, X, X^2, X^3$ (dans cet ordre!).

$$\text{On constate que } \begin{cases} f(1) = 1 \\ f(X) = 1 + 2X \\ f(X^2) = 2X + 3X^2 \\ f(X^3) = 3X^2 + 4X^3 \end{cases} \quad . \text{ On en déduit que } \text{Mat}_e(f) = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 2 & 2 & 0 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 4 \end{pmatrix}.$$

17.1.3 Coordonnées de l'image d'un vecteur par une appⁿ linéaire

Proposition 17.1.1 (interprétation matricielle de l'égalité $v = f(u)$)

Soit E et F deux espaces vectoriels sur \mathbb{K} .

On suppose que $\dim(E) = p \geq 1$, et que E est muni d'une base $e = (e_j)_{1 \leq j \leq p}$.

On suppose que $\dim(F) = n \geq 1$, et que F est muni d'une base $\varepsilon = (\varepsilon_i)_{1 \leq i \leq n}$.

Soit f une application linéaire de E dans F , de matrice A dans les bases e et ε .

Pour tout u de E , l'égalité vectorielle $v = f(u)$ équivaut à l'égalité matricielle $[f(u)]_\varepsilon = A[u]_e$.

Réciproquement, si pour une application $f: E \rightarrow F$ il existe une matrice A telle que $[f(u)]_\varepsilon = A[u]_e$ pour tout vecteur u de E , alors f est linéaire et $A = \text{Mat}_{e,\varepsilon}(f)$.

Exemples

Supposons qu'une base de E soit $e = (e_1, e_2, e_3)$, et qu'une base de F soit $\varepsilon = (\varepsilon_1, \varepsilon_2)$.

$$\text{Si } f: E \rightarrow F \text{ est l'application linéaire définie par } \begin{cases} f(e_1) = \varepsilon_1 + 2\varepsilon_2 \\ f(e_2) = 7\varepsilon_1 + 5\varepsilon_2 \\ f(e_3) = 3\varepsilon_1 \end{cases} \quad \text{alors } \text{Mat}_{e,\varepsilon}(f) = \begin{pmatrix} 1 & 7 & 3 \\ 2 & 5 & 0 \end{pmatrix}$$

Si u a pour coordonnées (x, y, z) dans e , et si v a pour coordonnées (x', y') dans ε , alors :

$$v = f(u) \Leftrightarrow [v]_\varepsilon = A[u]_e \Leftrightarrow \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & 7 & 3 \\ 2 & 5 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \Leftrightarrow \begin{cases} x' = x + 7y + 3z \\ y' = 2x + 5y \end{cases}$$

Réciproquement, soit $g: E \rightarrow F$ envoyant $u = xe_1 + ye_2 + ze_3$ sur $v = x'\varepsilon_1 + y'\varepsilon_2$ avec $\begin{cases} x' = x + 2y + 3z \\ y' = 9x + 8y + 7z \end{cases}$

Ce système s'écrit $\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 9 & 8 & 7 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, et on a par exemple $g(e_1) = \varepsilon_1 + 9\varepsilon_2$.

Ainsi g est linéaire et sa matrice dans les bases e et ε est $\text{Mat}_{e,\varepsilon}(g) = \begin{pmatrix} 1 & 2 & 3 \\ 9 & 8 & 7 \end{pmatrix}$

Application identité et matrice identité

La matrice de Id_E dans une base e de E est la matrice identité I_n , quelle que soit la base e .

Mais attention ! la matrice de Id_E n'est pas I_n si dans E on utilise une certaine base e *au départ* et une autre base ε *à l'arrivée* (voir plus loin la question des « matrices de passage »).

Application linéaire canoniquement associée à une matrice

On sait qu'à toute application linéaire $f : E \rightarrow F$ (avec $\dim(E) = p \geq 1$ et $\dim(F) = n \geq 1$) correspond une matrice unique de $\mathcal{M}_{n,p}(\mathbb{K})$ dans un couple de bases donné.

Inversement, une matrice A de $\mathcal{M}_{n,p}(\mathbb{K})$ peut « représenter » une infinité d'applications linéaires :

- on a en effet d'abord le choix des espaces E (de dimension p) et F (de dimension n).
- on a ensuite le choix d'une base de E et d'une base de F .

Si rien n'est imposé, on prend souvent $E = \mathbb{K}^p$ et $F = \mathbb{K}^n$, munis de leurs bases canoniques. L'application linéaire $f : \mathbb{K}^p \rightarrow \mathbb{K}^n$ ainsi obtenue est dite « canoniquement associée » à la matrice A .

Par exemple, soit $A = \begin{pmatrix} 1 & 0 & -1 & 3 & 2 \\ 3 & 2 & -1 & 0 & 4 \\ 2 & -1 & 5 & -2 & 0 \end{pmatrix}$ dans $\mathcal{M}_{3,5}(\mathbb{R})$.

Elle est canoniquement associée à une application linéaire $f : \mathbb{K}^5 \rightarrow \mathbb{K}^3$.

Celle-ci est définie par $\begin{cases} x' = x - z + 3t + 2u \\ y' = 3x + 2y - z + 4u \\ z' = 2x - y + 5z - 2t \end{cases}$ c'est-à-dire par $\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & -1 & 3 & 2 \\ 3 & 2 & -1 & 0 & 4 \\ 2 & -1 & 5 & -2 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \\ u \end{pmatrix}$.

Forme linéaire canoniquement associée à une matrice-ligne

Une matrice ligne $A = (a_1 \ a_2 \ \dots \ a_p)$ est canoniquement associée à une forme linéaire f sur \mathbb{K}^p .

Plus précisément, f est définie par $f(x_1, x_2, \dots, x_p) = (a_1 \ a_2 \ \dots \ a_p) \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} = \sum_{k=1}^p a_k x_k$.

17.1.4 Propriétés opératoires

▷ Matrice d'une combinaison linéaire

Proposition 17.1.2 (matrice de $\lambda f + \mu g$)

Soit E et F deux espaces vectoriels sur \mathbb{K} , de dimensions respectives $p \geq 1$ et $n \geq 1$.

On suppose que E est muni d'une base $e = (e_j)_{1 \leq j \leq p}$, et que F est muni d'une base $\varepsilon = (\varepsilon_i)_{1 \leq i \leq n}$.

L'application qui à f associe $\text{Mat}_{e,\varepsilon}(f)$ est un isomorphisme de $\mathcal{L}(E, F)$ sur $\mathcal{M}_{n,p}(\mathbb{K})$.

Ainsi, $\begin{cases} \text{pour tous } f, g \text{ dans } \mathcal{L}(E, F) \\ \text{pour tous } \lambda, \mu \text{ de } \mathbb{K} \end{cases} : \text{Mat}_{e,\varepsilon}(\lambda f + \mu g) = \lambda \text{Mat}_{e,\varepsilon}(f) + \mu \text{Mat}_{e,\varepsilon}(g).$

Cet isomorphisme implique $\dim \mathcal{L}(E, F) = \dim \mathcal{M}_{n,p}(\mathbb{K}) = np = \dim(E) \dim(F)$.

Cet isomorphisme n'est pas « canonique », en ce sens qu'il dépend des bases e, ε choisies dans E, F .

Complément : une base de $\mathcal{L}(E, F)$

On considère toujours l'isomorphisme de $\mathcal{L}(E, F)$ sur $\mathcal{M}_{n,p}(\mathbb{K})$ qui à f associe $\text{Mat}_{e,\varepsilon}(f)$.

On obtient une base de $\mathcal{L}(E, F)$ en prenant l'image réciproque de la base canonique de $\mathcal{M}_{n,p}(\mathbb{K})$.

Une base de $\mathcal{L}(E, F)$ est donc formée des $(f_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$, avec
$$\begin{cases} \forall k \in \{1, \dots, p\} \setminus \{j\}, f_{i,j}(e_k) = \vec{0} \\ f_{i,j}(e_j) = \varepsilon_i \end{cases}$$

En d'autres termes, $f_{i,j}$ est définie par $f_{i,j}\left(\sum_{k=1}^p x_k e_k\right) = x_j \varepsilon_i$.

Cette base de $\mathcal{L}(E, F)$ n'est pas « canonique » car elle dépend du choix des bases e et ε .

▷ Matrice d'une composée**Proposition 17.1.3** (matrice de la composée $g \circ f$)

Soit E, F, G trois espaces vectoriels sur \mathbb{K} , tous les trois de dimension finie non nulle.

On suppose que E, F, G sont munis de bases notées respectivement α, β et γ .

Soit f une application linéaire de E dans F , et soit g une application linéaire de F dans G .

On note $A = \text{Mat}_{\alpha,\beta}(f)$ et $B = \text{Mat}_{\beta,\gamma}(g)$.

Alors la matrice de l'application linéaire $g \circ f$, dans les bases α de E et γ de G , est $C = BA$.

Autrement dit : $\text{Mat}_{\alpha,\gamma}(g \circ f) = \text{Mat}_{\beta,\gamma}(g) \text{Mat}_{\alpha,\beta}(f)$.

▷ Matrice d'une application nilpotente

Soit E un \mathbb{K} -espace vectoriel de dimension $n \geq 1$, muni d'une base e .

Soit f un endomorphisme de E , de matrice A dans la base e .

Alors l'application f est nilpotente si et seulement si la matrice A est nilpotente.

Si f est nilpotente, il existe une base (ε) de E telle que $\text{Mat}_{\varepsilon}(f)$ est strictement triangulaire supérieure.

▷ Matrice d'un isomorphisme**Proposition 17.1.4** (matrice de f^{-1})

Soit E et F deux \mathbb{K} -espaces vectoriels de même dimension $n \geq 1$.

On suppose que E est muni d'une base e , et que F est muni d'une base ε .

Soit f une application linéaire de E dans F , de matrice A dans les bases e et ε .

Alors f est un isomorphisme si et seulement si la matrice A est inversible.

Dans ce cas, la matrice de f^{-1} dans les bases ε et e est A^{-1} , ou encore : $\text{Mat}_{\varepsilon,e}(f^{-1}) = (\text{Mat}_{e,\varepsilon}(f))^{-1}$

Cas particulier : matrice d'un automorphisme dans une base

Soit E un espace vectoriel sur \mathbb{K} , de dimension $n \geq 1$, muni d'une base e .

Soit f un endomorphisme de E , de matrice A dans la base e .

Alors f est un automorphisme de E si et seulement si A est inversible dans $\mathcal{M}_n(\mathbb{K})$.

En d'autres termes, on a l'équivalence : $f \in \text{GL}(E) \Leftrightarrow A \in \text{GL}_n(\mathbb{K})$.

En cas d'inversibilité, la matrice de f^{-1} dans la base e est A^{-1} , ou encore : $\text{Mat}_e(f^{-1}) = (\text{Mat}_e(f))^{-1}$.

▷ **Matrice des puissances d'une application linéaire****Proposition 17.1.5** (matrice de f^n)

Soit E un espace vectoriel sur \mathbb{K} , de dimension $n \geq 1$, muni d'une base e .

Soit f un endomorphisme de E , de matrice A dans la base e .

Pour tout entier naturel k , la matrice de f^k dans la base e est A^k .

Cette propriété s'étend aux k négatifs si f est un automorphisme de E , c'est-à-dire si A est inversible.

17.2 Changements de bases

17.2.1 Matrice de passage d'une base à une autre

Proposition 17.2.1 (inversibilité de la matrice d'une famille de vecteurs)

Soit E un espace vectoriel sur \mathbb{K} , de dimension $n \geq 1$, muni d'une base e .

Soit $v = (v_j)_{1 \leq j \leq n}$ une famille de n vecteurs de E (autant donc que la dimension de E).

Soit A la matrice de la famille v dans la base e . C'est un élément de $\mathcal{M}_n(\mathbb{K})$.

Alors la famille v est une base de E si et seulement si la matrice A est inversible.

Définition 17.2.1 (matrice de passage d'une base à une autre)

Soit E un espace vectoriel sur \mathbb{K} , de dimension $n \geq 1$, muni de deux bases $e = (e_i)_{1 \leq i \leq n}$ et $\varepsilon = (\varepsilon_j)_{1 \leq j \leq n}$.

La matrice de la famille ε dans la base e est appelée *matrice de passage* de la base e à la base ε .

Cette matrice est notée P_e^ε . D'après ce qui précède, elle est inversible.

Deux interprétations d'une matrice de passage

Avec les notations précédentes, la matrice de passage de e à ε est à la fois :

- la matrice de l'identité, de E muni de la base ε , vers E muni de la base e : $P_e^\varepsilon = \text{Mat}_{\varepsilon, e}(\text{Id}_E)$.
- la matrice dans la base e de l'automorphisme f de E défini par : $\forall j \in \{1, \dots, n\}, f(e_j) = \varepsilon_j$.

Opérations entre matrices de passage

- l'inverse de la matrice de passage P_e^ε de e à ε est la matrice de passage P_ε^e de ε à e .
- si α, β et γ sont trois bases de E , alors on a la relation : $P_\alpha^\gamma = P_\alpha^\beta P_\beta^\gamma$.

17.2.2 Effet d'un changement de base(s)

Proposition 17.2.2 (changement de base pour les coordonnées d'un vecteur)

Soit E un espace vectoriel sur \mathbb{K} , de dimension $n \geq 1$, muni de deux bases e et ε .

Soit P la matrice de passage de e à ε . Pour tout u de E : $\boxed{[u]_e = P [u]_\varepsilon}$

On retiendra le paradoxe : la matrice de passage de l'ancienne base à la nouvelle donne les anciennes coordonnées en fonction des nouvelles.

Proposition 17.2.3 (changement de base pour la matrice d'une application linéaire)

Soit E et F deux espaces vectoriels sur \mathbb{K} , de dimensions respectives $p \geq 1$ et $n \geq 1$.

On suppose que E est muni d'une ancienne base e et d'une nouvelle base e' .

De même soit ε l'ancienne base de F , et soit ε' la nouvelle base de F .

Soit P la matrice de passage de e à e' . Soit Q la matrice de passage de ε à ε' .

Soit f une application linéaire de E dans F .

Soit A la matrice de f dans les bases e et ε (ancienne matrice de f).

Soit B la matrice de f dans les bases e' et ε' (nouvelle matrice de f).

Alors on a l'égalité : $\boxed{B = Q^{-1} A P}$.

Cas particulier important

Soit E un espace vectoriel sur \mathbb{K} , de dimension $n \geq 1$.

Soit e l'ancienne base de E et e' la nouvelle base de E .

Soit f un endomorphisme de E , de matrice A dans la base e , et de matrice B dans la base e' .

Soit P la matrice de passage de e à e' . Alors on a l'égalité : $\boxed{B = P^{-1} A P}$.

17.2.3 Matrices équivalentes et matrices semblables

Définition 17.2.2 (matrices équivalentes)

Soit A et B deux matrices de $\mathcal{M}_{n,p}(\mathbb{K})$.

On dit que A et B sont équivalentes s'il existe G dans $\text{GL}_p(\mathbb{K})$ et D dans $\text{GL}_n(\mathbb{K})$, telles que $B = GAD$.

Remarques

On pourra noter $A \equiv B$ pour dire « A est équivalente à B ».

On définit ainsi une relation ... d'équivalence dans $\mathcal{M}_{np}(\mathbb{K})$:

- réflexivité : toute matrice A est équivalente à elle-même (écrire $A = I_n A I_p$).
- symétrie : si $A \equiv B$, alors $B \equiv A$. En effet $B = GAD \Rightarrow A = G^{-1} B D^{-1}$.
- transitivité : si $\begin{cases} A \equiv B \\ B \equiv C \end{cases}$ alors $A \equiv C$. En effet $\begin{cases} B = GAD \\ C = G' B D' \end{cases} \Rightarrow C = (G' G) A (D D')$.

Caractérisation de l'équivalence des matrices

– Soit E et F deux espaces vectoriels sur \mathbb{K} , de dimensions respectives $p \geq 1$ et $n \geq 1$.

Soit f une application linéaire de E dans F .

Soit A la matrice de f dans un certain couple de bases (e dans E et ε dans F).

Soit B la matrice de f dans un autre couple de bases (e' dans E et ε' dans F).

On sait que $B = Q^{-1} A P$ où P et Q sont les deux matrices de passages.

Les matrices A et B sont donc équivalentes dans $\mathcal{M}_{n,p}(\mathbb{K})$.

- Réciproquement, soit A et B deux matrices équivalentes dans $\mathcal{M}_{n,p}(\mathbb{K})$.
Soit E et F deux espaces vectoriels sur \mathbb{K} , de dimensions respectives $p \geq 1$ et $n \geq 1$.
On se donne un couple de bases, e dans E et ε dans F .
On sait que dans ce couple de bases, A représente une unique application linéaire $f : E \rightarrow F$.
Alors il existe un couple de bases, e' dans E et ε' dans F , tel que $B = \text{Mat}_{e',\varepsilon'}(f)$.
- On peut résumer de la façon suivante :

Proposition 17.2.4 (une caractérisation de l'équivalence des matrices)

Deux matrices A et B sont équivalentes dans $\mathcal{M}_{n,p}(\mathbb{K})$ si et seulement si elles sont susceptibles de représenter une même application linéaire f (chacune dans un certain couple de bases).

17.2.4 Réduction de la matrice de f à une forme J_r

Définition 17.2.3 (matrices J_r)

On se place dans l'espace vectoriel $\mathcal{M}_{n,p}(\mathbb{K})$, avec $n \geq 1$ et $p \geq 1$.

Soit r un entier vérifiant $0 \leq r \leq \min(n, p)$.

On note J_r la matrice de $\mathcal{M}_{n,p}(\mathbb{K})$, de coefficients a_{ij} définis par $\begin{cases} a_{i,i} = 1 \text{ pour } 1 \leq i \leq r \\ a_{i,j} = 0 \text{ dans tous les autres cas} \end{cases}$

Par exemple, dans $\mathcal{M}_{4,5}(\mathbb{K})$: $J_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ et $J_3 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$

On devrait noter $J_r(n, p)$, mais en général les valeurs de n et p sont connues sans ambiguïté.

La matrice J_0 est la matrice nulle de $\mathcal{M}_{n,p}(\mathbb{K})$.

Si $n = p$, la matrice J_n est la matrice identité I_n .

Voici une fonction Python pour former les matrices J_r . Les arguments sont les entiers r , n et p (ces deux derniers étant facultatifs avec la valeur 5 par défaut) :

```
>>> def J(r, n=5, p=5):
    return np.array([[i==j<r for j in range(p)] for i in range(n)]).astype(int)
```

Voici deux exemples d'utilisation de la fonction J :

```
>>> J(3) # ici r=3, et n=p=5 par défaut
array([[1, 0, 0, 0, 0],
       [0, 1, 0, 0, 0],
       [0, 0, 1, 0, 0],
       [0, 0, 0, 0, 0],
       [0, 0, 0, 0, 0]])
```

```
>>> J(2,4,6) # ici r=2, n=4, p=6
array([[1, 0, 0, 0, 0, 0],
       [0, 1, 0, 0, 0, 0],
       [0, 0, 0, 0, 0, 0],
       [0, 0, 0, 0, 0, 0]])
>>>
```

Proposition 17.2.5 (réduction de la matrice de f à une forme J_r)

Soient E et F deux espaces vectoriels sur \mathbb{K} , de dimension finie non nulle.

Soit f une application linéaire de E dans F , de rang r .

Alors il existe une base e de E et une base ε de F telles que $\text{Mat}_{e,\varepsilon}(f) = J_r$.

Attention !

Soit E un espace vectoriel de dimension finie non nulle. Soit f un endomorphisme de E , de rang r . La proposition précédente **ne dit pas** qu'il existe *une* base e de E dans laquelle $\text{Mat}_e(f) = J_r$. Donc même si f est un endomorphisme d'un espace vectoriel E , la proposition précédente énonce l'existence **de deux bases**, une base e « au départ » et une base ε « à l'arrivée ».

17.2.5 Matrices semblables**Définition 17.2.4** (matrices semblables)

Soit A et B deux matrices de $\mathcal{M}_n(\mathbb{K})$.

On dit que A et B sont semblables s'il existe P dans $\text{GL}_n(\mathbb{K})$ telle que $B = P^{-1}AP$.

Remarques

Cette notion ne concerne que les matrices carrées.

Deux matrices semblables sont équivalentes, mais la réciproque est fautive.

La seule matrice semblable à la matrice nulle de $\mathcal{M}_n(\mathbb{K})$ est la matrice nulle.

Pour tout scalaire λ , la seule matrice semblable à λI_n est la matrice λI_n .

Caractérisation de la similitude des matrices

- Soit E un espace vectoriel sur \mathbb{K} , de dimension $n \geq 1$. Soit f un endomorphisme de E . Soit A la matrice de f dans une base e de E . Soit B la matrice de f dans une autre base e' de E . On sait que $B = P^{-1}AP$ où P est la matrice de passage de e à e' . Les matrices A et B sont donc semblables dans $\mathcal{M}_n(\mathbb{K})$.
- Réciproquement, soit A et B deux matrices semblables dans $\mathcal{M}_n(\mathbb{K})$. Soit E un espace vectoriel sur \mathbb{K} , de dimension n , muni d'une base e . On sait que, dans cette base e , la matrice A représente un unique endomorphisme f de E . Dans ces conditions, si e' est la base de E définie par $P = P_{e,e'}$, on a $B = \text{Mat}_{e'}(f)$.
- On peut résumer de la façon suivante :

Proposition 17.2.6 (une caractérisation de la similitude des matrices)

Deux matrices A, B de $\mathcal{M}_n(\mathbb{K})$ sont semblables si et seulement si elles sont susceptibles de représenter un même endomorphisme f d'un espace vectoriel E de dimension n (chacune dans une certaine base).

Puissances de matrices semblables

Si $B = P^{-1}AP$, alors pour tout entier naturel n on a : $B^n = P^{-1}A^nP$.

Cette relation s'étend aux exposants négatifs si A (et donc B) est inversible.

On peut donc calculer B^n si A^n est plus facile à obtenir, notamment si A est diagonale.

17.3 Trace d'une matrice, d'un endomorphisme

17.3.1 Trace d'une matrice carrée

Définition 17.3.1 (trace d'une matrice carrée)

Soit A une matrice carrée d'ordre n , à coefficients dans \mathbb{K} , de terme général a_{ij} .

On appelle *trace* de A , et on note $\text{tr}(A)$ ou $\text{Tr}(A)$, la somme $\sum_{i=1}^n a_{ii}$ des coefficients diagonaux de A .

Proposition 17.3.1 (linéarité de la trace)

L'application « trace » est une forme linéaire sur l'espace vectoriel $\mathcal{M}_n(\mathbb{K})$.

Ainsi pour toutes matrices A et B de $\mathcal{M}_n(\mathbb{K})$, et tous α, β de \mathbb{K} , on a : $\text{tr}(\alpha A + \beta B) = \alpha \text{tr}(A) + \beta \text{tr}(B)$.

Proposition 17.3.2 (trace d'un produit de deux matrices)

Soient A une matrice de $\mathcal{M}_{n,p}(\mathbb{K})$ et B une matrice de $\mathcal{M}_{p,n}(\mathbb{K})$.

La matrice AB est donc carrée d'ordre n , tandis que BA est carrée d'ordre p .

Les matrices carrées AB et BA ont la même trace : $\text{tr}(AB) = \text{tr}(BA)$.

L'égalité $\text{tr}(AB) = \text{tr}(BA)$ est vraie en particulier pour toutes matrices A, B de $\mathcal{M}_n(\mathbb{K})$.

On ne doit pas généraliser abusivement à des produits de plus de deux matrices.

On peut par exemple écrire : $\text{tr}(ABC) = \text{tr}(A(BC)) = \text{tr}((BC)A) = \text{tr}(BCA)$.

On peut également écrire $\text{tr}(ABC) = \text{tr}(CAB)$ (en échangeant AB et C).

Mais en général on a : $\text{tr}(ABC) \neq \text{tr}(ACB)$.

Proposition 17.3.3 (invariance de la trace par similitude)

Soit A et B deux matrices de $\mathcal{M}_n(\mathbb{K})$.

Si A et B sont semblables, alors elles ont la même trace.

Pour démontrer ce résultat, on écrit $\text{tr}(P^{-1}AP) = \text{tr}((P^{-1}A)P) = \text{tr}(P(P^{-1}A)) = \text{tr}(A)$.

Mais on évitera d'écrire : $\text{tr}(P^{-1}AP) = \text{tr}(P^{-1}PA) = \text{tr}(A)$!

La réciproque de la proposition précédents est évidemment fausse.

Attention à ne pas confondre...

On ne confondra pas la *trace* $\text{tr}(A)$ d'une matrice carrée A avec sa *transposée* A^\top .

– la trace n'est définie que pour les matrices carrées, et la trace de A est un *scalaire*.

– la transposée d'une matrice A de type (n, p) est une *matrice* de type (p, n) .

– on a la relation $\text{tr}(AB) = \text{tr}(BA)$, mais on a la relation $(^\top AB) = B^\top A^\top$.

17.3.2 Trace d'un endomorphisme

Définition 17.3.2 (trace d'un endomorphisme en dimension finie)

Soit E un espace vectoriel sur \mathbb{K} , de dimension finie $n \geq 1$. Soit f un endomorphisme de E .

La *trace* de f , notée $\text{tr}(f)$ ou $\text{Tr}(f)$, est la trace de la matrice de f dans une base *quelconque* de E .

Cette définition est légitime : bien que la matrice de f dépende en général de la base choisie dans l'espace vectoriel E , la trace de cette matrice n'en dépend pas (en effet toutes les matrices susceptibles de représenter l'application f dans une base e de E sont semblables entre elles).

Les deux propriétés suivantes sont des conséquences immédiates de la définition :

- si f et g sont deux endomorphismes de E , on a : $\text{tr}(\lambda f + \mu g) = \lambda \text{tr}(f) + \mu \text{tr}(g)$.
- si f et g sont dans $\mathcal{L}(E)$, on a : $\text{tr}(gf) = \text{tr}(fg)$ (rappelons qu'on note gf pour $g \circ f$).

Trace d'une projection vectorielle (d'un projecteur)

Si p est la projection vectorielle de E sur un sous-espace F de dimension r , alors $\text{tr}(p) = \text{rg}(p) = r$.

Pour s'en persuader, il suffit de choisir un supplémentaire G de F dans E et de se placer dans une base adaptée à la somme directe $E = F \oplus G$: la matrice de p dans cette base est diagonale, les r premiers coefficients diagonaux valant 1 et les $n - r$ derniers valant 0.

Trace d'une rotation vectorielle en dimension 3

On se place dans l'espace vectoriel euclidien orienté \mathbb{R}^3 .

Soit r la rotation vectorielle d'angle θ autour d'un vecteur unitaire w . Alors $\text{tr}(r) = 1 + 2 \cos \theta$.

Il suffit en effet de se placer dans une base orthonormée directe u, v, w .

La matrice de r dans cette base est $A = \begin{pmatrix} \cos \theta & -\sin \theta & 1 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$ et $\text{tr}(A) = 1 + 2 \cos \theta$.

Si on connaît la matrice B de la rotation r dans une base quelconque (même non orthonormée), on sait qu'on a l'égalité $\text{tr}(B) = 1 + 2 \cos \theta$. On en déduit immédiatement le cosinus de l'angle de la rotation r . Pour calculer son sinus il faut *orienter* l'axe de la rotation.

17.4 Noyau, image et rang d'une matrice

17.4.1 Application linéaire canoniquement associée

On revient ici sur des remarques déjà formulées dans la section 16.2.4.

Considérons l'application qui à $u = (x_1, x_2, \dots, x_p)$ associe la matrice colonne $U = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$.

C'est un isomorphisme « canonique » de \mathbb{K}^p sur $\mathcal{M}_{p,1}(\mathbb{K})$.

Il permet d'identifier un p -uplet u avec la matrice colonne U correspondante, de hauteur p .

Soit A un élément de $\mathcal{M}_{n,p}(\mathbb{K})$. L'application $\varphi: U \mapsto V = AU$ est linéaire de $\mathcal{M}_{p,1}(\mathbb{K})$ dans $\mathcal{M}_{n,1}(\mathbb{K})$.

À isomorphisme près, il s'agit donc d'une application linéaire de \mathbb{K}^p dans \mathbb{K}^n .

Si par exemple $A = \begin{pmatrix} 2 & 1 & 0 & 3 \\ 0 & 4 & 1 & 5 \\ 1 & 6 & 2 & 0 \end{pmatrix}$, et $U = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$, alors $V = AU = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$, où $\begin{cases} y_1 = 2x_1 + x_2 + 3x_4 \\ y_2 = 4x_2 + x_3 + 5x_4 \\ y_3 = x_1 + 6x_2 + 2x_3 \end{cases}$

Dans cet exemple, on peut donc identifier :

- l'application linéaire $\varphi: U \mapsto V = AU$ de $\mathcal{M}_{4,1}(\mathbb{K})$ dans $\mathcal{M}_{3,1}(\mathbb{K})$.
- l'application linéaire $f: u = (x_1, x_2, x_3, x_4) \mapsto v = (y_1, y_2, y_3)$ de \mathbb{K}^4 dans \mathbb{K}^3 .

Définition 17.4.1 (application linéaire canoniquement associée à une matrice)

Soit A un élément de $\mathcal{M}_{n,p}(\mathbb{K})$.

- l'application $\varphi: U \mapsto V = AU$ est linéaire de $\mathcal{M}_{p,1}(\mathbb{K})$ dans $\mathcal{M}_{n,1}(\mathbb{K})$.
- si on munit \mathbb{K}^p et \mathbb{K}^n de leurs bases canoniques, A est la matrice d'un unique f de $\mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$.

L'application linéaire φ (ou l'application linéaire f) est dite canoniquement associée à la matrice A .

Compte-tenu de ce qui a été dit plus haut, on peut donc quasiment considérer que f et φ sont une seule et même application linéaire.

Dans la suite, on utilisera indifféremment l'une ou l'autre des deux terminologies.

17.4.2 Noyau, image et rang d'une matrice

Définition 17.4.2 (image et noyau d'une matrice)

Soit A une matrice, élément de $\mathcal{M}_{n,p}(\mathbb{K})$.

Soit $f: \mathbb{K}^p \rightarrow \mathbb{K}^n$ l'application linéaire canoniquement associée à la matrice A .

On appelle *image* de A , et on note $\text{Im}(A)$, l'image $\text{Im}(f)$ de l'application linéaire f .

On appelle *noyau* de A , et on note $\text{Ker}(A)$, l'image $\text{Ker}(f)$ de l'application linéaire f .

On appelle *rang* de A , et on note $\text{rg}(A)$, le rang $\text{rg}(f)$ de l'application linéaire f .

Le *théorème de la dimension* nous assure que $\text{rg}(A) + \dim(\text{Ker}(A)) = p$.

▷ Premières remarques

- Si A est dans $\mathcal{M}_{n,p}(\mathbb{K})$, alors $\text{rg}(A) \leq n$ et $\text{rg}(A) \leq p$.
- Le rang d'une matrice A est nul si et seulement si A est elle-même la matrice nulle.
- Le rang d'une matrice A est égal à 1 si et seulement si A possède une colonne non nulle et si les autres colonnes de A lui sont proportionnelles (voir ci-après).
- Le rang de A est celui de toute application linéaire susceptible d'être représentée par A .

▷ Image d'une matrice, et vecteur colonnes

Soit A un élément de $\mathcal{M}_{n,p}(\mathbb{K})$. Notons C_1, C_2, \dots, C_p les p colonnes de A .

Pour toute colonne $U = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}$ de $\mathcal{M}_{p,1}(\mathbb{K})$, on peut écrire $AU = \left(C_1 \mid C_2 \mid \dots \mid C_p \right) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix} = \sum_{j=1}^p x_j C_j$.

Ainsi l'image de A , c'est-à-dire l'ensemble des AU , est le sous-espace engendré par C_1, C_2, \dots, C_p .

On retiendra donc : *les colonnes d'une matrice A forment une famille génératrice de $\text{Im}(A)$.*

▷ **Noyau d'une matrice, et vecteurs-ligne**

Soit $A = (a_{i,j})$ un élément de $\mathcal{M}_{n,p}(\mathbb{K})$. Notons L_1, L_2, \dots, L_n les n lignes de A .

On peut donc écrire $A = \begin{pmatrix} L_1 \\ L_2 \\ \vdots \\ L_n \end{pmatrix}$. On se donne ensuite $U = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}$ quelconque dans $\mathcal{M}_{p,1}(\mathbb{K})$.

Avec ces notations, on peut caractériser les vecteurs U de $\text{Ker}(A)$:

$$U \in \text{Ker}(A) \Leftrightarrow \begin{pmatrix} L_1 \\ L_2 \\ \vdots \\ L_n \end{pmatrix} U = \vec{0} \Leftrightarrow \begin{pmatrix} L_1 U \\ L_2 U \\ \vdots \\ L_n U \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \Leftrightarrow (\forall i \in \{1, \dots, n\}, L_i U = 0)$$

Mais $L_i = (a_{i,1} \ a_{i,2} \ \dots \ a_{i,p})$ donc l'égalité $L_i U = 0$ s'écrit $\sum_{j=1}^p a_{ij} x_j = 0$.

Ainsi l'égalité $AU = \vec{0}$ se traduit par un système de n équations linéaires impliquant les lignes de A .

On retiendra donc : *les lignes d'une matrice A permettent de former un système d'équations de $\text{Ker}(A)$.*

▷ **Illustration par un exemple**

On demande de déterminer le noyau, l'image, et le rang de la matrice $A = \begin{pmatrix} 2 & -1 & 1 & 5 \\ -1 & 2 & 3 & -4 \\ 3 & 0 & 5 & 6 \end{pmatrix}$

On obtient $\text{Ker}(A)$ en résolvant $AU = 0$, où U est dans $\mathcal{M}_{4,1}(\mathbb{K})$.

$$AU = 0 \Leftrightarrow A \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} 2x - y + z + 5t = 0 & (E_1) \\ -x + 2y + 3z - 4t = 0 & (E_2) \\ 3x + 5z + 6t = 0 & (E_3) \end{cases} \Leftrightarrow \begin{cases} 2x - y + z + 5t = 0 \\ -x + 2y + 3z - 4t = 0 \end{cases}$$

On voit bien comment les trois lignes de A ont permis d'écrire un système d'équations du noyau.

On s'est ramené à un système de deux équations car $(E_3) = 2(E_1) + (E_2)$.

Le dernier système obtenu équivaut à $\begin{cases} 2x - y = -z - 5t \\ x - 2y = 3z - 4t \end{cases}$ c'est-à-dire $x = -\frac{5}{3}z - 2t$ et $y = -\frac{7}{3}z + t$.

Cela équivaut à $(x, y, z, t) = \frac{z}{3}(-5, -7, 1, 0) + t(-2, 1, 0, 1)$, avec $(z, t) \in \mathbb{R}^2$.

Ainsi : $U = \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \text{Ker}(A) \Leftrightarrow X = \frac{z}{3} \begin{pmatrix} -5 \\ -7 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} -2 \\ 1 \\ 0 \\ 1 \end{pmatrix}$ où z et t sont deux scalaires quelconques.

Finalement, on peut dire que $\text{Ker}(A)$ est le plan engendré par $V = \begin{pmatrix} -5 \\ -7 \\ 1 \\ 0 \end{pmatrix}$ et $W = \begin{pmatrix} -2 \\ 1 \\ 0 \\ 1 \end{pmatrix}$.

On peut aussi invoquer l'application linéaire $f : \mathbb{K}^4 \rightarrow \mathbb{K}^3$ canoniquement associée à A . La différence (très mince) consiste à considérer $u = (x, y, z, t)$ plutôt que la matrice colonne U , et à dire que le noyau est le plan engendré par $v = (-5, -7, 1, 0)$ et $w = (-2, 1, 0, 1)$ dans \mathbb{K}^4 .

Le théorème de la dimension donne $\dim \mathbb{K}^4 = \dim \text{Im}(f) + \dim \text{Ker}(f)$.

Or $\dim \text{Ker}(f) = 2$. On en déduit $\dim \text{Im}(f) = 2$, c'est-à-dire $\dim \text{Im}(A) = 2$.

Ainsi l'application linéaire f , et donc la matrice A , sont de rang 2.

Les vecteurs $\begin{cases} f(e_1) = (2, -1, 3) \\ f(e_2) = (-1, 2, 0) \end{cases}$ sont dans le plan $\text{Im}(f)$ et sont libres. Ils en forment donc une base.

En termes quasiment identiques : les colonnes $\begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}$ et $\begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}$ forment une base de $\text{Im}(A)$.

Le sous-espace $\text{Im}(f)$ est un plan de \mathbb{R}^3 , donc le noyau d'une forme linéaire non nulle sur \mathbb{R}^3 .

Il existe donc $(a, b, c) \neq \vec{0}$ dans \mathbb{R}^3 tel que : $(x, y, z) \in \text{Im}(f) \Leftrightarrow ax + by + cz = 0$.

En utilisant $\begin{cases} f(e_1) = (2, -1, 3) \\ f(e_2) = (-1, 2, 0) \end{cases}$ on trouve $\begin{cases} 2a - b + 3c = 0 \\ -a + 2b = 0 \end{cases}$ donc $\begin{cases} a = 2b \\ c = -b \end{cases}$

Les coefficients de l'équation $ax + by + cz = 0$ sont uniques à un facteur multiplicatif non nul près.

En choisissant $b = 1$, on voit que l'équation de $\text{Im}(f)$ dans la base canonique est : $2x + y - z = 0$.

Mais à y regarder de plus près, on a déjà rencontré cette relation !

En effet, on a vu que les lignes $(E_1), (E_2), (E_3)$ du système $AU = 0$ étaient liées par : $(E_3) = 2(E_1) + (E_2)$.

Cette relation signifie que toutes les colonnes $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ de $\text{Im}(A)$ vérifient $z = 2x + y$.

Le même exemple, avec Python

Attention : la fonction `matrix_rank` renvoie le rang d'une matrice (au sens que nous avons donné à ce terme). Il ne faut pas confondre avec la fonction `rank` qui donne le nombre d'indices utilisés pour décrire le tableau A (ici `rank(A)` donne 2 car il faut deux indices pour décrire A).

```
>>> print(a)
[[ 2 -1  1  5]
 [-1  2  3 -4]
 [ 3  0  5  6]]
```

Voici la **bonne** fonction !

```
>>> np.linalg.matrix_rank(a)
2
```

Voici la **mauvaise** fonction !

```
>>> np.rank(a)
2
```

▷ Inversibilité et noyau, pour une matrice carrée

Proposition 17.4.1

Soit A une matrice carrée. Alors A est inversible si et seulement si son noyau est réduit à $\vec{0}$.

Pour prouver que A est inversible dans $\mathcal{M}_n(\mathbb{K})$, il suffit donc de prouver que $AX = \vec{0} \Rightarrow X = \vec{0}$. Évidemment, cela ne donne pas A^{-1} (il faudrait pour cela résoudre $AX = B$ en $X = A^{-1}B$).

▷ Effet du produit par une matrice inversible

Proposition 17.4.2 (invariance du noyau par multiplication à gauche)

Soit A une matrice de $\mathcal{M}_{n,p}(\mathbb{K})$, et soit Q dans $\text{GL}_n(\mathbb{K})$ (donc Q est carrée inversible d'ordre n). Alors les matrices A et QA ont le même noyau. En particulier elles ont le même rang.

Proposition 17.4.3 (invariance de l'image par multiplication à droite)

Soit A une matrice de $\mathcal{M}_{n,p}(\mathbb{K})$, et soit P dans $\text{GL}_p(\mathbb{K})$ (donc P est carrée inversible d'ordre p). Alors les matrices A et AP ont la même image. En particulier elles ont le même rang.

17.4.3 Matrices équivalentes et rang

On se place dans l'espace vectoriel $\mathcal{M}_{n,p}(\mathbb{K})$, avec $n \geq 1$ et $p \geq 1$.

On rappelle qu'on note J_r la matrice de coefficients a_{ij} définie par
$$\begin{cases} a_{i,i} = 1 \text{ pour } 1 \leq i \leq r \\ a_{i,j} = 0 \text{ dans tous les autres cas} \end{cases}$$

Par exemple, dans $\mathcal{M}_{4,5}(\mathbb{K})$, on a $J_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ et $J_3 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$.

Proposition 17.4.4 (caractérisation du rang avec les matrices J_r)

Soit A une matrice de $\mathcal{M}_{n,p}(\mathbb{K})$.

Alors A est de rang r si et seulement si elle est équivalente à la matrice J_r .

Proposition 17.4.5 (rapport entre équivalence et rang des matrices)

Soit A et B deux matrices de $\mathcal{M}_{n,p}(\mathbb{K})$.

Alors A et B sont équivalentes si et seulement si elles ont le même rang.

On sait qu'en écrivant « A est équivalente à B », on définit une relation d'équivalence sur $\mathcal{M}_{n,p}(\mathbb{K})$.

La proposition précédente dit que dans $\mathcal{M}_{n,p}(\mathbb{K})$, il y a exactement $1 + \min(n, p)$ classes d'équivalence pour cette relation : chaque classe d'équivalence est en effet formée des matrices A dont le rang a une valeur donnée r , avec $0 \leq r \leq \min(n, p)$.

En particulier, dans $\mathcal{M}_n(\mathbb{K})$ on a l'inégalité $\text{rg}(A) \leq n$ et l'équivalence ($\text{rg}(A) = n \Leftrightarrow A$ est inversible).

Proposition 17.4.6 (invariance du rang par transposition)

Le rang d'une matrice A est égal au rang de sa matrice transposée A^\top .

Proposition 17.4.7 (interprétations du rang d'une matrice)

Soit A dans $\mathcal{M}_{n,p}(\mathbb{K})$. Le rang de A est égal au nombre maximum de colonnes libres dans A .

Le rang de A est aussi égal au nombre maximum de lignes libres dans A .

17.4.4 Rang et matrices extraites

Définition 17.4.3 (matrices extraites)

Soit A un élément de $\mathcal{M}_{n,p}(\mathbb{K})$.

Soit I une famille strictement croissante d'indices de ligne, c'est-à-dire d'éléments de $\{1, \dots, n\}$.

Soit J une famille strictement croissante d'indices de colonne, c'est-à-dire d'éléments de $\{1, \dots, p\}$.

Soit B la matrice formée par les intersections des lignes L_i et C_j de A , avec i dans I et j dans J .

On dit que B est une *matrice extraite* de A .

Par exemple, si $\begin{cases} I = (1, 3, 4) \\ J = (2, 3, 5, 6) \end{cases}$ et $A = \begin{pmatrix} 2 & \boxed{5} & \boxed{3} & 1 & \boxed{6} & \boxed{7} \\ 4 & 1 & 5 & 6 & 8 & 9 \\ 1 & \boxed{0} & \boxed{4} & 5 & \boxed{1} & \boxed{3} \\ 7 & \boxed{9} & \boxed{6} & 2 & \boxed{1} & \boxed{5} \\ 1 & 5 & 8 & 2 & 7 & 4 \end{pmatrix}$ alors $B = \begin{pmatrix} 5 & 3 & 6 & 7 \\ 0 & 4 & 1 & 3 \\ 9 & 6 & 1 & 5 \end{pmatrix}$

Voici une fonction `sousmatrice` permettant d'extraire d'une matrice A une sous-matrice B en précisant un ensemble I d'indices de ligne, puis un ensemble J d'indices de colonne.

```
>>> def sousmatrice(a, I, J):
    return np.array([[a[i-1, j-1] for j in J] for i in I])
```

Et voici une illustration de cette fonction, en reprenant l'exemple précédent :

```
>>> a
array([[2, 5, 3, 1, 6, 7],
       [4, 1, 5, 6, 8, 9],
       [1, 0, 4, 5, 1, 3],
       [7, 9, 6, 2, 1, 5],
       [1, 5, 8, 2, 7, 4]])

>>> sousmatrice(a, [1,3,4], [2,3,5,6])
array([[5, 3, 6, 7],
       [0, 4, 1, 3],
       [9, 6, 1, 5]])
>>>
```

Proposition 17.4.8 (rang d'une matrice extraite)

Soit A un élément de $\mathcal{M}_{n,p}(\mathbb{K})$, et soit B une matrice extraite de A . Alors $\text{rg}(B) \leq \text{rg}(A)$.

Conséquence importante : si de A on peut extraire une matrice de rang r , alors $\text{rg}(A) \geq r$.

Proposition 17.4.9 (caractérisation du rang par les matrices extraites inversibles)

Soit A un élément de $\mathcal{M}_{n,p}(\mathbb{K})$.

Alors $\text{rg}(A)$ est l'ordre maximum d'une matrice carrée inversible extraite de A .

17.5 Calcul effectif du rang

17.5.1 Matrices échelonnées

Soit $A = (a_{ij})$ une matrice de $\mathcal{M}_{n,p}(\mathbb{K})$. On note L_1, \dots, L_n les lignes successives de A .

Pour chaque ligne L_i de A , soit $d(i)$ le plus petit indice j , s'il existe, tel que $a_{ij} \neq 0$.

On dit que A est *échelonnée* supérieurement s'il existe un entier r de $\{0, \dots, n\}$ tel que :

- pour tout indice i inférieur ou égal à r , la ligne L_i est non nulle.
- pour tout indice i strictement supérieur à r , la ligne L_i est nulle.
- la suite $d(1), d(2), \dots, d(r)$ est strictement croissante.

Proposition 17.5.1 (rang d'une matrice échelonnée)

Avec les notations précédentes, la matrice échelonnée A est de rang r .

Les r coefficients non nuls situés aux positions $(i, d(i))$ sont appelés les *pivots* de A .

Par exemple, $A = \begin{pmatrix} 0 & \boxed{1} & 3 & 4 & 0 & 1 & 5 \\ 0 & 0 & \boxed{3} & 5 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \boxed{4} & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & \boxed{9} & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ est échelonnée, avec quatre pivots, donc $\text{rg}(A) = 4$.

La matrice nulle de $\mathcal{M}_{n,p}(\mathbb{K})$ est un cas particulier de matrice échelonnée (avec zéro pivot!).

On définit comme précédemment les matrices « échelonnées inférieurement ». En fait une matrice A est échelonnée inférieurement si et seulement si sa transposée est échelonnée supérieurement.

17.5.2 Opérations élémentaires

Définition 17.5.1 (opérations élémentaires sur les lignes d'une matrice)

Soit A une matrice de $\mathcal{M}_{n,p}(\mathbb{K})$. Notons L_1, L_2, \dots, L_n les lignes de A .

On appelle *opération élémentaire* sur les lignes de A l'une des opérations suivantes :

- multiplier une ligne L_i par un scalaire **non nul** α : on note $L_i \leftarrow \alpha L_i$.
- ajouter à l'une des lignes L_i un multiple d'une **autre** ligne L_j : on note $L_i \leftarrow L_i + \beta L_j$.
- échanger deux lignes L_i et L_j : on note $L_i \leftrightarrow L_j$.

On définit de même les opérations élémentaires sur les colonnes de la matrice A .

Ces opérations sont notées : $C_i \leftarrow \alpha C_i$ (avec $\alpha \neq 0$), $C_i \leftarrow C_i + \beta C_j$ (avec $j \neq i$), et $C_i \leftrightarrow C_j$.

▷ Nécessiter d'utiliser un pivot non nul

Dans les opérations $L_i \leftarrow \alpha L_i$ et $C_i \leftarrow \alpha C_i$, on dit que α joue le rôle de « pivot ».

Dans ces deux opérations il est absolument indispensable que α soit non nul.

On fera notamment attention au cas où α dépend d'un paramètre : pour les valeurs de ce paramètre qui annuleraient α , l'opération se traduirait par $L_i \leftarrow 0$ ou $C_i \leftarrow 0$ et serait « illégale ».

On note souvent $L_i \leftarrow \alpha L_i + \beta L_j$ la composée de $L_i \leftarrow \alpha L_i$ (avec $\alpha \neq 0$) puis de $L_i \leftarrow L_i + \beta L_j$.

▷ Interprétation matricielle des opérations élémentaires sur les lignes

Proposition 17.5.2 (interprétation matricielle d'une opération élémentaire sur les lignes)

On se place dans $\mathcal{M}_{n,p}(\mathbb{K})$ et on considère une opération élémentaire particulière sur les lignes.

Notons φ l'application qui à toute matrice A de $\mathcal{M}_{n,p}(\mathbb{K})$ associe sa transformée par cette opération.

Alors il existe une matrice inversible Q d'ordre n telle que : $\forall A \in \mathcal{M}_{n,p}(\mathbb{K}), \varphi(A) = QA$.

- Dans $\mathcal{M}_{n,p}(\mathbb{K})$, toute opération élémentaire sur les *lignes* équivaut donc à une multiplication à *gauche* par une certaine matrice inversible Q d'ordre n . Pour savoir quelle est cette matrice inversible Q , il suffit d'appliquer cette opération élémentaire à la matrice identité I_n .
- Dans $\mathcal{M}_{n,p}(\mathbb{K})$, on considère une succession donnée d'opérations élémentaires sur les lignes. Soit φ l'application qui à A dans $\mathcal{M}_{n,p}(\mathbb{K})$ associe sa transformée par cette succession d'opérations. Alors il existe une matrice inversible Q d'ordre n telle que : $\forall A \in \mathcal{M}_n(\mathbb{K}), \varphi(A) = QA$.
- En particulier, si on applique à A une opération élémentaire sur les lignes (ou une succession de telles opérations), on transforme A en une matrice B telle que $\text{Ker}(B) = \text{Ker}(A)$.
D'une façon plus familière, les opérations élémentaires sur les lignes « conservent le noyau ».

▷ Interprétation matricielle des opérations élémentaires sur les colonnes

Proposition 17.5.3 (interprétation matricielle d'une opération élémentaire sur les colonnes)

On se place dans $\mathcal{M}_{n,p}(\mathbb{K})$ et on considère une opération élémentaire particulière sur les colonnes.

Notons φ l'application qui à toute matrice A de $\mathcal{M}_{n,p}(\mathbb{K})$ associe sa transformée par cette opération.

Alors il existe une matrice inversible P d'ordre p telle que : $\forall A \in \mathcal{M}_{n,p}(\mathbb{K}), \varphi(A) = AP$.

- Dans $\mathcal{M}_{n,p}(\mathbb{K})$, toute opération élémentaire sur les *colonnes* équivaut donc à une multiplication à *droite* par une certaine matrice inversible P d'ordre p . Pour savoir quelle est cette matrice inversible P , il suffit d'appliquer cette opération élémentaire à la matrice identité I_p .
- Dans $\mathcal{M}_{n,p}(\mathbb{K})$, on considère une succession donnée d'opérations élémentaires sur les colonnes. Soit φ l'application qui à A dans $\mathcal{M}_{n,p}(\mathbb{K})$ associe sa transformée par cette succession d'opérations. Alors il existe une matrice inversible P d'ordre p telle que : $\forall A \in \mathcal{M}_{n,p}(\mathbb{K}), \varphi(A) = AP$.
- En particulier, si on applique à A une opération élémentaire sur les colonnes (ou une succession de telles opérations), on transforme A en une matrice B telle que $\text{Im}(B) = \text{Im}(A)$.
D'une façon plus familière, les opérations élémentaires sur les colonnes « conservent l'image ».

Proposition 17.5.4 (conservation du rang par des opérations élémentaires)

Toute opération élémentaire (ou toute suite d'opérations élémentaires), sur les lignes et/ou sur les colonnes, transforme une matrice A en une matrice de même rang.

On peut donc dire que les opérations élémentaires (sur les colonnes, sur les lignes) « conservent le rang ».

17.5.3 Calcul du rang par la méthode du pivot

Rappelons les trois acceptions du mot « rang » :

- Soit E, F deux espaces vectoriels sur \mathbb{K} , avec $\dim(E) = p \geq 1$, et $\dim(F) = n \geq 1$.
Soit $f : E \rightarrow F$ une application linéaire. Le rang de f est la dimension de $\text{Im}(f)$.
- Soit A une matrice de $\mathcal{M}_{n,p}(\mathbb{K})$.
Alors $\text{rg}(A) = \text{rg}(f)$, où $f : \mathbb{K}^p \rightarrow \mathbb{K}^n$ est l'application linéaire canoniquement associée à A .
- Soit E un \mathbb{K} espace vectoriel, avec $\dim(E) = n$. Soit $(v_j)_{1 \leq j \leq p}$ une famille de p vecteurs de E .
Soit $F = \text{Vect}(v)$ le sous-espace de E engendré par les vecteurs v_j .
Le rang de la famille $v = (v_j)_{1 \leq j \leq p}$ est la dimension du sous-espace $F = \text{Vect}(v)$.

En fait, chacune de ces trois définitions se rattache immédiatement aux deux autres :

- Le rang d'une matrice A est celui de la famille de ses vecteurs-colonne, et celui de la famille de ses vecteurs-ligne, et celui de toute application linéaire susceptible d'être représentée par A .
- Le rang d'une famille de vecteurs de E est celui de la matrice des coordonnées de ces vecteurs dans une base e quelconque de E .
- Le rang d'une application linéaire $f : E \rightarrow F$ est le rang de la matrice de f dans un couple de bases quelconques (e dans E et ε dans F). C'est aussi le rang de la famille des images des vecteurs d'une base quelconque e de E .

Finalement tout peut se ramener au calcul du rang d'une matrice.

Proposition 17.5.5 (passage à une forme échelonnée supérieure)

Soit A une matrice de $\mathcal{M}_{n,p}(\mathbb{K})$. Par une succession bien choisie d'opérations élémentaires sur les lignes, on peut transformer A en une matrice échelonnée supérieurement.

Les opérations élémentaires ne modifiant pas le rang de la matrice initiale, on peut ainsi calculer le rang de A : c'est celui de la matrice échelonnée finale, c'est-à-dire le nombre de ses pivots non nuls.

Exemple de calcul de rang

On demande de calculer le rang de $A = \begin{pmatrix} 1 & 5 & 9 & 13 & 17 \\ 3 & 7 & 11 & 15 & 19 \\ 2 & 6 & 1 & 0 & 11 \\ 1 & 3 & 14 & 21 & 16 \end{pmatrix}$

Au moyen des opérations $\begin{cases} L_2 \leftarrow L_2 - 3L_1 \\ L_3 \leftarrow L_3 - 2L_1 \\ L_4 \leftarrow L_4 - L_1 \end{cases}$ on passe de A à $A_1 = \begin{pmatrix} 1 & 5 & 9 & 13 & 17 \\ 0 & -8 & -16 & -24 & -32 \\ 0 & -4 & -17 & -26 & -23 \\ 0 & -2 & 5 & 8 & -1 \end{pmatrix}$

Au moyen des opérations $\begin{cases} L_3 \leftarrow 2L_3 - L_2 \\ L_4 \leftarrow 4L_4 - L_2 \end{cases}$ on passe de A_1 à $A_2 = \begin{pmatrix} 1 & 5 & 9 & 13 & 17 \\ 0 & -8 & -16 & -24 & -32 \\ 0 & 0 & 18 & 28 & 14 \\ 0 & 0 & -36 & -56 & -28 \end{pmatrix}$

Au moyen de l'opération $L_4 \leftarrow L_4 + 2L_3$, on passe de A_2 à $A_3 = \begin{pmatrix} \boxed{1} & 5 & 9 & 13 & 17 \\ 0 & \boxed{-8} & -16 & -24 & -32 \\ 0 & 0 & \boxed{18} & 28 & 14 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$

La matrice finale, donc la matrice initiale A , est de rang 3.

17.5.4 Calcul de l'inverse par la méthode du pivot

Proposition 17.5.6 (inversion d'une matrice par opérations élémentaires)

Soit A une matrice de $\mathcal{M}_n(\mathbb{K})$. Alors A est inversible si et seulement si il est possible, par une succession bien choisie d'opérations élémentaires sur les lignes, de passer de la matrice A à la matrice I_n .

La même succession d'opérations, dans le même ordre, transforme alors I_n en la matrice A^{-1} .

▷ Principe de la méthode

Soit A la matrice de $\mathcal{M}_n(\mathbb{K})$ dont on veut calculer l'inverse.

On place A et I_n côte à côte dans un tableau $(A \mid I_n)$ à n lignes et $2n$ colonnes.

On procède ensuite à une succession d'opérations élémentaires sur les lignes de ce tableau.

Cette suite d'opérations correspond à la multiplication à gauche par une matrice inversible P .

On passe donc du tableau $(A \mid I_n)$ au tableau $P(A \mid I_n) = (PA \mid PI_n)$.

On choisit la succession d'opérations sur les lignes de A de manière à transformer A en $PA = I_n$.

Cela signifie que $P = A^{-1}$, et le tableau final est donc $(I_n \mid A^{-1})$.

▷ Dans la pratique

– On part du tableau $(A \mid I_n) = \left(\begin{array}{cccc|cccc} a_{11} & a_{12} & \dots & a_{1p} & 1 & 0 & \dots & 0 & 0 \\ a_{21} & a_{22} & \dots & a_{2p} & 0 & 1 & \ddots & \ddots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \vdots & 0 & \ddots & \ddots & 1 & 0 \\ a_{n1} & a_{n2} & \dots & a_{n,p} & 0 & 0 & \dots & 0 & 1 \end{array} \right)$

– Il y a au moins un coefficient non nul sur la première colonne sinon A ne serait pas inversible.

Au besoin après un échange de deux lignes, on peut supposer $a_{11} \neq 0$.

- On applique les opérations élémentaires : $L_i \leftarrow a_{11}L_i - a_{i1}L_1$, pour $i = 2, \dots, n$.

Dans cette succession d'opérations, le coefficient non nul a_{11} est appelé le *pivot*.

Le tableau prend alors la forme suivante :

$$\left(\begin{array}{cccc|cccc} a_{11} & a_{12} & \dots & a_{1p} & 1 & 0 & \dots & 0 & 0 \\ 0 & b_{22} & \dots & b_{2p} & -a_{21} & a_{11} & \ddots & \ddots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & a_{11} & 0 \\ 0 & b_{n2} & \dots & b_{n,p} & -a_{n1} & 0 & \dots & 0 & a_{11} \end{array} \right)$$

- La partie gauche du tableau est une matrice qui a le même rang que A , et qui est donc inversible. Il y a donc au moins un coefficient non nul dans la deuxième colonne et à partir de la deuxième ligne (sinon les deux premières colonnes seraient proportionnelles et A ne serait pas inversible). Au besoin après un échange de L_2 avec une ligne L_i d'indice $i \geq 3$, on peut donc supposer $b_{22} \neq 0$. Avec le pivot b_{22} , on annule tous les coefficients de la deuxième colonne (sans toucher à la ligne L_2).
- On continue ainsi jusqu'à obtenir à la place de A une matrice diagonale. On termine en divisant chaque ligne par le coefficient diagonal obtenu. À la place qu'occupait I_n se trouve maintenant A^{-1} .
- Il n'est pas utile (au contraire) de rendre égaux à 1 les coefficients diagonaux de la moitié gauche du tableau avant d'y avoir obtenu une matrice diagonale, pour éviter d'introduire des coefficients fractionnaires difficiles à manipuler : puisqu'on demande souvent d'inverser des matrices à coefficients entiers, autant garder les coefficients entiers le plus longtemps possible !

▷ Un exemple assez technique

Soit $A = \begin{pmatrix} 1 & 1/2 & 1/3 & 1/4 \\ 1/2 & 1/3 & 1/4 & 1/5 \\ 1/3 & 1/4 & 1/5 & 1/6 \\ 1/4 & 1/5 & 1/6 & 1/7 \end{pmatrix}$. Pour calculer A^{-1} , on forme $\left(\begin{array}{cccc|cccc} 1 & 1/2 & 1/3 & 1/4 & 1 & 0 & 0 & 0 \\ 1/2 & 1/3 & 1/4 & 1/5 & 0 & 1 & 0 & 0 \\ 1/3 & 1/4 & 1/5 & 1/6 & 0 & 0 & 1 & 0 \\ 1/4 & 1/5 & 1/6 & 1/7 & 0 & 0 & 0 & 1 \end{array} \right)$

Pour limiter la technicité de l'exercice, il est préférable de travailler sur des coefficients entiers.

Les opérations $\begin{cases} L_1 \leftarrow 12L_1 \\ L_2 \leftarrow 60L_2 \\ L_3 \leftarrow 60L_3 \\ L_4 \leftarrow 420L_4 \end{cases}$ conduisent au tableau $\left(\begin{array}{cccc|cccc} 12 & 6 & 4 & 3 & 12 & 0 & 0 & 0 \\ 30 & 20 & 15 & 12 & 0 & 60 & 0 & 0 \\ 20 & 15 & 12 & 10 & 0 & 0 & 60 & 0 \\ 105 & 84 & 70 & 60 & 0 & 0 & 0 & 420 \end{array} \right)$

On va utiliser le pivot 12 en position (1, 1).

Les opérations $\begin{cases} L_2 \leftarrow 2L_2 - 5L_1 \\ L_3 \leftarrow 3L_3 - 5L_1 \\ L_4 \leftarrow 4L_4 - 35L_1 \end{cases}$ conduisent à $\left(\begin{array}{cccc|cccc} 12 & 6 & 4 & 3 & 12 & 0 & 0 & 0 \\ 0 & 10 & 10 & 9 & -60 & 120 & 0 & 0 \\ 0 & 15 & 16 & 15 & -60 & 0 & 180 & 0 \\ 0 & 126 & 140 & 135 & -420 & 0 & 0 & 1680 \end{array} \right)$

On va utiliser le pivot 10 en position (2, 2).

Les opérations $\begin{cases} L_1 \leftarrow 5L_1 - 3L_2 \\ L_3 \leftarrow 2L_3 - 3L_2 \\ L_4 \leftarrow 5L_4 - 63L_2 \end{cases}$ conduisent à $\left(\begin{array}{cccc|cccc} 60 & 0 & -10 & -12 & 240 & -360 & 0 & 0 \\ 0 & 10 & 10 & 9 & -60 & 120 & 0 & 0 \\ 0 & 0 & 2 & 3 & 60 & -360 & 360 & 0 \\ 0 & 0 & 70 & 108 & 1680 & -7560 & 0 & 8400 \end{array} \right)$

On va utiliser le pivot 2 en position (3, 3).

$$\text{Les opérations } \begin{cases} L_1 \leftarrow L_1 + 5L_3 \\ L_2 \leftarrow L_2 - 5L_3 \\ L_4 \leftarrow L_4 - 35L_3 \end{cases} \text{ conduisent à } \left(\begin{array}{cccc|cccc} 60 & 0 & 0 & 3 & 540 & -2160 & 1800 & 0 \\ 0 & 10 & 0 & -6 & -360 & 1920 & -1800 & 0 \\ 0 & 0 & 2 & 3 & 60 & -360 & 360 & 0 \\ 0 & 0 & 0 & 3 & -420 & 5040 & -12600 & 8400 \end{array} \right)$$

On va utiliser le pivot 3 en position (4, 4).

$$\text{Les opérations } \begin{cases} L_1 \leftarrow L_1 - L_4 \\ L_2 \leftarrow L_2 + 2L_4 \\ L_3 \leftarrow L_3 - L_4 \end{cases} \text{ conduisent à } \left(\begin{array}{cccc|cccc} 60 & 0 & 0 & 0 & 960 & -7200 & 14400 & -8400 \\ 0 & 10 & 0 & 0 & -1200 & 12000 & -27000 & 16800 \\ 0 & 0 & 2 & 0 & 480 & -5400 & 12960 & -8400 \\ 0 & 0 & 0 & 3 & -420 & 5040 & -12600 & 8400 \end{array} \right)$$

On a transformé la partie gauche du tableau en une matrice diagonale dont les coefficients diagonaux sont non nuls (cette matrice est donc inversible, ce qui prouve que la matrice initiale A est inversible).

On peut maintenant transformer la partie gauche du tableau en la matrice identité.

$$\text{Les opérations } \begin{cases} L_1 \leftarrow L_1/60 \\ L_2 \leftarrow L_2/10 \\ L_3 \leftarrow L_3/2 \\ L_4 \leftarrow L_4/3 \end{cases} \text{ conduisent à } \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 16 & -120 & 240 & -140 \\ 0 & 1 & 0 & 0 & -120 & 1200 & -2700 & 1680 \\ 0 & 0 & 1 & 0 & 240 & -2700 & 6480 & -4200 \\ 0 & 0 & 0 & 1 & -140 & 1680 & -4200 & 2800 \end{array} \right)$$

$$\text{On a ainsi montré que } A \text{ est inversible et que } A^{-1} = \begin{pmatrix} 16 & -120 & 240 & -140 \\ -120 & 1200 & -2700 & 1680 \\ 240 & -2700 & 6480 & -4200 \\ -140 & 1680 & -4200 & 2800 \end{pmatrix}$$

▷ Le même exemple en Python

On importe le module `numpy` (renommé comme d'habitude en `np`), et on fabrique la matrice A :

```
>>> import numpy as np
>>> A = np.array([[1/(i+j+1) for j in range(4)] for i in range(4)]); A
[[ 1.          0.5          0.33333333  0.25        ]
 [ 0.5         0.33333333  0.25         0.2         ]
 [ 0.33333333  0.25         0.2          0.16666667 ]
 [ 0.25        0.2          0.16666667  0.14285714 ]]
```

Par sécurité, on fait une sauvegarde de A dans la variable $A_$.

On borde ensuite A par la matrice identité d'ordre 4 :

```
>>> A_ = A.copy(); A = np.hstack((A,np.eye(4))); print(A)
[[ 1.          0.5          0.33333333  0.25         1.         0.         0.         0.        ]
 [ 0.5         0.33333333  0.25         0.2         0.         1.         0.         0.        ]
 [ 0.33333333  0.25         0.2          0.16666667  0.         0.         1.         0.        ]
 [ 0.25        0.2          0.16666667  0.14285714  0.         0.         0.         1.        ]]
```

On multiplie chaque ligne par un coefficient bien choisi, ce qui permet de passer en coefficients entiers :

```
>>> A[0]=12*A[0]; A[1]=60*A[1]; A[2]=60*A[2]; A[3]=420*A[3]
>>> A = np.asarray(A, dtype='int'); print(A)
[[ 12   6   4   3  12   0   0   0]
 [ 30  20  15  12   0  60   0   0]
 [ 20  15  12  10   0   0  60   0]
 [105  84  70  60   0   0   0 420]]]
```

On effectue les opérations élémentaires avec le pivot qui est en position (1, 1) :

```
>>> A[1]=2*A[1]-5*A[0]; A[2]=3*A[2]-5*A[0]; A[3]=4*A[3]-35*A[0]; print(A)
[[ 12   6   4   3  12   0   0   0]
 [  0  10  10   9 -60 120   0   0]
 [  0  15  16  15 -60   0 180   0]
 [  0 126 140 135 -420   0   0 1680]]
```

On effectue les opérations élémentaires avec le pivot qui est en position (2, 2) :

```
>>> A[0]=5*A[0]-3*A[1]; A[2]=2*A[2]-3*A[1]; A[3]=5*A[3]-63*A[1]; print(A)
[[ 60   0 -10 -12 240 -360   0   0]
 [  0  10  10   9 -60 120   0   0]
 [  0   0   2   3  60 -360 360   0]
 [  0   0  70 108 1680 -7560   0 8400]]
```

On effectue les opérations élémentaires avec le pivot qui est en position (3, 3) :

```
>>> A[0]=A[0]+5*A[2]; A[1]=A[1]-5*A[2]; A[3]=A[3]-35*A[2]; print(A)
[[ 60   0   0   3 540 -2160 1800   0]
 [  0  10   0  -6 -360 1920 -1800   0]
 [  0   0   2   3   60 -360 360   0]
 [  0   0   0   3 -420 5040 -12600 8400]]
```

On effectue les opérations élémentaires avec le pivot qui est en position (4, 4) :

```
>>> A[0]=A[0]-A[3]; A[1]=A[1]+2*A[3]; A[2]=A[2]-A[3]; print(A)
[[ 60   0   0   0 960 -7200 14400 -8400]
 [  0  10   0   0 -1200 12000 -27000 16800]
 [  0   0   2   0 480 -5400 12960 -8400]
 [  0   0   0   3 -420 5040 -12600 8400]]
```

On divise chaque ligne par un coefficient bien choisi, pour obtenir l'identité à gauche :

```
>>> A[0] = A[0]/60; A[1] = A[1]/10; A[2] = A[2]/2; A[3] = A[3]/3; print(A)
[[ 1   0   0   0 16 -120 240 -140]
 [ 0   1   0   0 -120 1200 -2700 1680]
 [ 0   0   1   0 240 -2700 6480 -4200]
 [ 0   0   0   1 -140 1680 -4200 2800]]
```

On extrait le bloc de droite : c'est la matrice $B = A^{-1}$.

On vérifie le résultat en effectuant le produit AB avec la copie de sauvegarde de A . On obtient bien la matrice identité (on a arrondi ici à douze décimales, donc c'est très précis) :

```
>>> B = A[:,4:]; print(B)
[[ 16 -120 240 -140]
 [-120 1200 -2700 1680]
 [ 240 -2700 6480 -4200]
 [-140 1680 -4200 2800]]

>>> (A_.dot(B)).round(12)
array([[ 1.,  0.,  0.,  0.],
       [ 0.,  1.,  0.,  0.],
       [ 0.,  0.,  1., -0.],
       [ 0.,  0.,  0.,  1.]])
```


Le système (S) s'écrit
$$\begin{pmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1p} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \dots & a_{ij} & \dots & a_{ip} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & a_{nj} & \dots & a_{np} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_j \\ \vdots \\ x_p \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_i \\ \vdots \\ b_n \end{pmatrix},$$
 c'est-à-dire $AX = B$.

Si A est carrée inversible (système dit « de Cramer »), (S) possède une solution unique $X = A^{-1}B$.

Le système homogène (H) associé à (S) s'écrit : $AX = 0$.

▷ Interprétation en termes d'applications linéaires

Soit f l'application linéaire de \mathbb{K}^p vers \mathbb{K}^n , de matrice A dans les bases canoniques.

Soit $x = (x_1, x_2, \dots, x_p)$ dans \mathbb{K}^p et $b = (b_1, b_2, \dots, b_n)$ dans \mathbb{K}^n .

Le système (S) équivaut alors à l'égalité vectorielle $f(x) = b$.

Résoudre (S), c'est donc chercher l'image réciproque du vecteur b de \mathbb{K}^n par l'application linéaire f .

Le système (S) admet au moins une solution si et seulement si b est dans l'image de f .

Le système homogène associé (H) équivaut à $f(x) = \vec{0}$.

Résoudre (H), c'est donc trouver $\text{Ker}(f)$.

▷ Interprétation en termes de combinaisons linéaires

Pour $1 \leq j \leq p$, soit $C_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix}$ la j -ième colonne de A , et $B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$ la colonne des seconds membres.

Le système (S) s'écrit :
$$x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix} + \dots + x_j \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix} + \dots + x_p \begin{pmatrix} a_{1p} \\ a_{2p} \\ \vdots \\ a_{np} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

Autrement dit, (S) s'écrit : $x_1 C_1 + \dots + x_j C_j + \dots + x_p C_p = B$.

Résoudre (S), c'est trouver toutes les façons d'écrire B comme combinaison linéaire de C_1, C_2, \dots, C_p .

Le système (S) a donc au moins une solution si et seulement si B est dans $\text{Vect}(C_1, C_2, \dots, C_p)$.

Résoudre (H), c'est trouver tous les p -uplets (x_1, \dots, x_p) tels que $\sum_{k=1}^p x_k C_k = \vec{0}$.

Le système (H) a d'autres solutions que la solution triviale si et seulement si C_1, C_2, \dots, C_p sont liés.

▷ Interprétation en termes d'intersections d'hyperplans affines

Revenons à la définition initiale du système (S).

Considérons la i -ème équation $(E_i) : a_{i1} x_1 + a_{i2} x_2 + \dots + a_{ij} x_j + \dots + a_{ip} x_p = b_i$.

On commence par éliminer le cas particulier où tous les coefficients $a_{i,j}$ de cette ligne sont nuls : si $b_i = 0$, alors l'équation (E_i) se réduit à $0 = 0$ et elle peut évidemment être supprimée, mais si $b_i \neq 0$ l'équation (E_i) (et donc le système (S)) n'a aucune solution.

On suppose donc que chacune des p -uplets $(a_{i1}, a_{i2}, \dots, a_{ip})$, avec $1 \leq i \leq n$ est non identiquement nul.

Chaque application $\varphi_i: \mathbb{K}^p \rightarrow \mathbb{K}$ définie par $\varphi_i(x_1, x_2, \dots, x_p) = a_{i1}x_1 + a_{i2}x_2 + \dots + a_{ij}x_j + \dots + a_{ip}x_p$ est alors une forme linéaire non nulle sur \mathbb{K}^p .

L'équation (E_i) s'écrit $\varphi_i(x) = b$, et elle définit un hyperplan affine \mathcal{H}_i de \mathbb{K}^p .

L'équation correspondante du système homogène associé (H) s'écrit $\varphi_i(x) = 0$, et elle définit l'hyperplan vectoriel H_i de \mathbb{K}^p qui est la direction de \mathcal{H}_i .

Avec cette interprétation, résoudre (S) c'est déterminer l'intersection éventuelle des hyperplans affines \mathcal{H}_i , et résoudre (H) c'est déterminer l'intersection des hyperplans vectoriels H_i .

17.6.3 Structure de l'ensemble des solutions

On reprend le système (S) défini dans les deux sous-sections précédentes.

▷ Structure de l'ensemble des solutions du système homogène

Utilisons l'interprétation de (H) en termes d'applications linéaires : $x = (x_1, x_2, \dots, x_p)$ est solution du système homogène (H) si et seulement si $f(x) = \vec{0}$ c'est-à-dire si et seulement si x est dans $\text{Ker}(f)$.

Or le théorème du rang nous dit que $\text{Ker}(f)$ est un sous-espace vectoriel de \mathbb{K}^p de dimension $p - \text{rg}(f)$, où $\text{rg}(f)$ est le rang de l'application linéaire f c'est-à-dire le rang de la matrice A .

On peut donc énoncer :

Proposition 17.6.1 (structure de l'ensemble des solutions du système linéaire homogène (H))
*Soit (H) un système linéaire homogène de n équations, à p inconnues et à coefficients dans \mathbb{K} .
 Soit $AX = 0$ l'écriture matricielle de ce système, avec A dans $\mathcal{M}_{n,p}(\mathbb{K})$ et X dans $\mathcal{M}_{p,1}(\mathbb{K})$.
 Soit r le rang du système, c'est-à-dire le rang de la matrice A .
 Alors l'ensemble des solutions X de (H) est un espace vectoriel de dimension $p - r$.*

▷ Structure de l'ensemble des solutions du système (S)

Supposons que (S) possède au moins une solution x_0 . Soit x un vecteur de \mathbb{K}^p .

On a les équivalences : $(x \text{ est solution de (S)}) \Leftrightarrow f(x) = b \Leftrightarrow f(x) = f(x_0) \Leftrightarrow f(x - x_0) = \vec{0}$.

Ainsi les solutions de (S) sont les $x = x_0 + h$ où h est une solution quelconque de (H).

L'ensemble des solutions \mathcal{S}_S de (S) est alors un sous-espace affine dont la direction est l'ensemble des solutions \mathcal{S}_H de (H) (et dont la dimension est donc $p - r$).

Attention : le calcul précédent est fait sous réserve que le système (S) admette au moins une solution !

Définition 17.6.1 (système linéaire compatible)

Soit (S) un système linéaire homogène de n équations, à p inconnues et à coefficients dans \mathbb{K} .

Soit $AX = B$ l'écriture matricielle de (S), avec A dans $\mathcal{M}_{n,p}(\mathbb{K})$, B dans $\mathcal{M}_{n,1}(\mathbb{K})$, X dans $\mathcal{M}_{p,1}(\mathbb{K})$.

On dit que le système (S) est *compatible* si l'ensemble de ses solutions est non vide.

Cela équivaut, avec l'interprétation matricielle, à dire que B est dans l'image de A .

On peut donc énoncer :

Proposition 17.6.2 (structure de l'ensemble des solutions d'un système linéaire compatible)

Soit (S) un système linéaire homogène de n équations, à p inconnues et à coefficients dans \mathbb{K} .

Soit $AX = B$ l'écriture matricielle de (S) , avec A dans $\mathcal{M}_{n,p}(\mathbb{K})$, B dans $\mathcal{M}_{n,1}(\mathbb{K})$, X dans $\mathcal{M}_{p,1}(\mathbb{K})$.

Soit r le rang du système, c'est-à-dire le rang de la matrice A .

On suppose que le système (S) est compatible.

Alors l'ensemble des solutions X de (S) est un sous-espace affine de dimension $p - r$.

Il s'obtient en ajoutant à une solution particulière de (S) la solution générale de (H) .

▷ Remarques

On garde bien sûr les notations précédentes, le système (S) s'écrivant indifféremment $f(x) = b$ (avec f dans $\mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$) ou $AX = B$ (avec A dans $\mathcal{M}_{n,p}(\mathbb{K})$).

Le système (S) est compatible si et seulement si b appartient à $\text{Im}(f)$.

Cela dépend de b , mais si f est surjective (c'est-à-dire $\text{rg}(A) = n$) alors (S) est toujours compatible.

Si f est injective (c'est-à-dire $\text{rg}(A) = p$) alors (S) possède au plus une solution, quelque soit le second membre b . Le système homogène (H) , quant à lui, possède alors uniquement la solution triviale.

L'ensemble des solutions d'un système linéaire quelconque (S) est donc ou bien vide (système incompatible) ou bien réduit à un seul élément (cas où f est bijective), ou bien infini (en tant que sous-espace affine de dimension supérieure ou égale à 1).

17.6.4 Systèmes de Cramer

On se place ici dans le cas d'un système carré (n équations et n inconnues).

Matriciellement, le système (S) s'écrit donc $AX = B$, avec A dans $\mathcal{M}_n(\mathbb{K})$, et X, B dans $\mathcal{M}_{n,1}(\mathbb{K})$.

Proposition 17.6.3 (solution unique d'un système de Cramer)

Soit (S) un système linéaire de n équations, à n inconnues et à coefficients dans \mathbb{K} .

Soit $AX = B$ son écriture matricielle, avec A dans $\mathcal{M}_n(\mathbb{K})$, B dans $\mathcal{M}_{n,1}(\mathbb{K})$, X cherché dans $\mathcal{M}_{n,1}(\mathbb{K})$.

Soit r le rang du système, c'est-à-dire le rang de la matrice A .

Le système (S) possède une solution unique si et seulement si $r = n$, c'est-à-dire si et seulement si la matrice A est inversible. Cette solution est $X = A^{-1}B$.

On dit alors que le système (S) est un « système de Cramer »

Avec l'interprétation en termes d'applications linéaires : $f(x) = b \Leftrightarrow x = f^{-1}(b)$.

Si on utilise l'interprétation de (S) en termes de combinaisons linéaires, l'unique solution (x_1, x_2, \dots, x_n) de (S) est le n -uplet des coordonnées du vecteur B (colonne des seconds membres) dans la base de \mathbb{K}^n formée par les colonnes C_1, C_2, \dots, C_n de la matrice A .

Résolution d'un système de Cramer avec Python

On utilise la fonction `solve` dans `numpy.linalg` :

```
>>> import numpy as np                # toujours s'assurer que numpy est importé!
>>> a = np.array([[1,3,2],[4,1,5],[6,2,7]]); a # matrice A carrée d'ordre 3
array([[1, 3, 2],
       [4, 1, 5],
       [6, 2, 7]])
>>> np.linalg.det(a)                  # déterminant non nul, matrice inversible
7.00000000000000071
>>> b = np.array([1,4,2]); b          # le vecteur B des seconds membres
array([1, 4, 2])
>>> x = np.linalg.solve(a, b); x      # la solution X du système
array([-6.42857143, -1.71428571,  6.28571429])
>>> np.dot(a,x)                       # on vérifie qu'on a bien AX=B
array([ 1.,  4.,  2.])
```

Voici ce qui se passe si la matrice carrée A n'est pas inversible.

Dans l'exemple ci-dessous la matrice A est de rang 2 (on a en effet $L_3 = L_2 - 2L_1$).

```
>>> a = np.array([[1,3,2],[4,1,5],[2,-5,1]]); a
array([[ 1,  3,  2],
       [ 4,  1,  5],
       [ 2, -5,  1]])
>>> x = np.linalg.solve(a, b);        # avec les mêmes seconds membres
[...]
numpy.linalg.linalg.LinAlgError: Singular matrix
>>> np.linalg.det(a)                  # effectivement le déterminant est nul
0.0
```

17.6.5 Résolution par la méthode du pivot de Gauss

Les aspects techniques de la résolution des systèmes linéaires par la méthode du pivot de Gauss ont été traités dans le chapitre 2 (« Calculs algébriques ») de ce cours (voir 2.5.8).

Chapitre 18

Déterminants

Sommaire

18.1 Le groupe symétrique	404
18.1.1 Permutations de l'ensemble $E_n = \{1, \dots, n\}$	404
18.1.2 Cycles	404
18.1.3 Transpositions	406
18.1.4 Décomposition en produit de cycles à supports disjoints	406
18.1.5 Décomposition en produit de transpositions	407
18.1.6 Signature d'une permutation	407
18.2 Formes n-linéaires alternées	409
18.2.1 Applications multilinéaires	409
18.2.2 Formes multilinéaires alternées	409
18.2.3 Application « déterminant dans une base »	410
18.2.4 Relation entre applications « déterminant dans une base »	412
18.3 Déterminant d'un endomorphisme, d'une matrice	412
18.3.1 Déterminant d'un endomorphisme	412
18.3.2 Déterminant d'une matrice carrée	413
18.4 Calcul des déterminants	415
18.4.1 Déterminants et opérations élémentaires	415
18.4.2 Développement d'un déterminant, comatrice	416
18.4.3 Quelques déterminants particuliers	417
18.5 Déterminants et orientation	419
18.5.1 Orientation d'un espace réel de dimension finie	419
18.5.2 Si $n = 2$, le déterminant est une aire orientée	419
18.5.3 Si $n = 3$, le déterminant est un volume orienté	421

18.1 Le groupe symétrique

18.1.1 Permutations de l'ensemble $E_n = \{1, \dots, n\}$

Définition 18.1.1 (groupe symétrique d'indice n)

Pour tout entier $n \geq 1$, on note $E_n = \{1, \dots, n\}$.

On appelle *permutation* de E_n , toute bijection de E_n sur lui-même.

On note \mathcal{S}_n l'ensemble de toutes les permutations de E_n .

L'ensemble \mathcal{S}_n est un groupe pour la loi de composition, appelé *groupe symétrique* d'indice n .

Le groupe \mathcal{S}_n est d'ordre $n!$, et il est non commutatif dès que $n \geq 3$.

Remarque : on écrira $\sigma_2\sigma_1$ (plutôt que $\sigma_2 \circ \sigma_1$) pour désigner la composée de σ_1 par σ_2 .

Un élément σ de \mathcal{S}_n est représenté par $\sigma = \begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix}$.

En particulier l'application identité, neutre du groupe \mathcal{S}_n , se note $\text{Id} = \begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix}$.

Pour prendre un exemple, $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 5 & 1 & 4 & 6 & 2 \end{pmatrix}$ représente l'élément σ de \mathcal{S}_6 défini par :

$$\sigma(1) = 3, \quad \sigma(2) = 5, \quad \sigma(3) = 1, \quad \sigma(4) = 4, \quad \sigma(5) = 6, \quad \sigma(6) = 2$$

On voit facilement que $\sigma^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 6 & 1 & 4 & 2 & 5 \end{pmatrix}$ (lire le tableau σ à partir de sa deuxième ligne).

Si $n = 1$, le groupe \mathcal{S}_1 se réduit à l'application identité de $E_1 = \{1\}$ dans lui-même.

Si $n = 2$, $\mathcal{S}_2 = \{\text{Id}, \sigma\}$, où σ est définie par : $\sigma(1) = 2$ et $\sigma(2) = 1$.

Si $n = 3$, \mathcal{S}_3 est formé de six éléments :

$$\begin{aligned} \sigma_0 = \text{Id} &= \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, & \sigma_1 &= \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, & \sigma_2 &= \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \\ \sigma_3 &= \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, & \sigma_4 &= \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, & \sigma_5 &= \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = \sigma_4^2 \end{aligned}$$

Le groupe (\mathcal{S}_3, \circ) n'est pas commutatif : on vérifie par exemple que $\begin{cases} \sigma_1 \sigma_3 = \sigma_5 \\ \sigma_3 \sigma_1 = \sigma_4 \end{cases}$

18.1.2 Cycles

Définition 18.1.2 (les cycles sont des permutations particulières)

Soit σ un élément de \mathcal{S}_n , avec $n \geq 2$. Soit p un entier de $\{2, \dots, n\}$.

On dit que σ est un *cycle* de longueur p s'il existe a_1, a_2, \dots, a_p distincts dans $\{1, \dots, n\}$ tels que :

- pour tout k de $\{1, \dots, p-1\}$, $\sigma(a_k) = a_{k+1}$, et de plus $\sigma(a_p) = a_1$.
- pour tout élément b de $E_n \setminus \{a_1, \dots, a_p\}$ on a $\sigma(b) = b$.

On dit alors que l'ensemble $\{a_1, \dots, a_p\}$ est le *support* du cycle σ .

Remarques

Le support d'un cycle σ est l'ensemble des éléments qui ne sont pas invariants par σ .

En général, on représente un cycle σ en écrivant $\sigma = (a_1 \ a_2 \ \dots \ a_p)$.

Dans \mathcal{S}_n , un cycle de longueur n est appelé une *permutation circulaire*.

Exemples

Dans \mathcal{S}_7 , la permutation $\sigma_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 5 & 6 & 3 & 1 & 2 & 4 & 7 \end{pmatrix}$ est le cycle $(1 \ 5 \ 2 \ 6 \ 4)$.

Cette dernière notation ne dit pas que σ_1 est dans \mathcal{S}_7 , mais qu'elle est dans \mathcal{S}_n pour tout $n \geq 6$ (en principe le contexte est clair, mais de toutes façons c'est sans grande importance).

Le support de σ_1 est $\{1, 2, 4, 5, 6\}$. Les éléments 3 et 7 sont fixes par σ_1 .

On remarque qu'on peut aussi écrire $\sigma_1 = (5 \ 2 \ 6 \ 4 \ 1)$, ou $\sigma_1 = (2 \ 6 \ 4 \ 1 \ 5) \dots$

En revanche la permutation $\sigma_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 6 & 5 & 3 & 8 & 7 & 4 & 2 & 1 \end{pmatrix}$ n'est pas un cycle.

Cependant on a visiblement $\sigma = st = ts$, où $s = (1 \ 6 \ 4 \ 8)$ et $t = (2 \ 5 \ 7)$.

Dans \mathcal{S}_7 , la permutation $\sigma_3 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 7 & 5 & 6 & 1 & 4 & 2 & 3 \end{pmatrix}$ est la permutation circulaire $(1 \ 7 \ 3 \ 6 \ 2 \ 5 \ 4)$.

Proposition 18.1.1 (inverse d'un cycle)

Si σ est le cycle $(a_1 \ a_2 \ \dots \ a_p)$ alors σ^{-1} est le cycle $(a_p \ a_{p-1} \ \dots \ a_1)$.

Attention, les puissances d'un cycle ne sont pas toujours des cycles.

Si par exemple $\sigma = (1 \ 2 \ 3 \ 4 \ 5 \ 6)$, alors $\sigma^2 = (1 \ 3 \ 5) (2 \ 4 \ 6)$ et $\sigma^3 = (1 \ 4) (2 \ 5) (3 \ 6)$.

En revanche σ^5 est le cycle $(1 \ 6 \ 5 \ 4 \ 3 \ 2)$.

On montre que si σ est un cycle de longueur p , alors σ^k est un cycle si et seulement si $k \wedge p = 1$.

Proposition 18.1.2 (calcul des puissances d'un cycle)

Soit σ un cycle de longueur $p \geq 2$. Alors $\sigma^p = \text{Id}$ et, pour tout k de $\{1, \dots, p-1\}$, $\sigma^k \neq \text{Id}$.

Pour tout entier relatif m , si $m = qp + r$ est la division euclidienne de m par p , alors $\sigma^m = \sigma^r$.

Proposition 18.1.3 (deux cycles à supports disjoints commutent)

Soit σ_1 et σ_2 deux cycles de \mathcal{S}_n .

On suppose que les supports de σ_1 et σ_2 sont disjoints. Alors $\sigma_2 \sigma_1 = \sigma_1 \sigma_2$.

Plus généralement, soit $\sigma_1, \dots, \sigma_m$ une famille de m cycles de \mathcal{S}_n .

Si leurs supports sont disjoints deux à deux, alors les cycles σ_k commutent deux à deux.

Dans ce cas, et pour tout entier k , on peut alors écrire : $(\sigma_1 \cdots \sigma_m)^k = \sigma_1^k \cdots \sigma_m^k$.

18.1.3 Transpositions

Définition 18.1.3 (les transpositions sont les cycles de longueur 2)

Soit n un entier supérieur ou égal à 2. Soit σ un élément de \mathcal{S}_n .

On dit que σ de \mathcal{S}_n est une *transposition* si σ est un cycle $(a \ b)$ de longueur 2.

Cela signifie qu'il existe a et b distincts dans E_n tels que
$$\begin{cases} \sigma(a) = b \text{ et } \sigma(b) = a \\ \forall c \notin \{a, b\}, \sigma(c) = c \end{cases}$$

Remarques

Une transposition est donc une permutation qui se contente d'échanger deux éléments.

On ne confondra pas les mots « permutation » et « transposition ».

Si $\tau = (a \ b)$, on a bien sûr : $\tau = (b \ a)$, $\tau^2 = \text{Id}$, $\tau^{-1} = \tau$.

Soit $\tau_1 = (a \ b)$ et $\tau_2 = (c \ d)$ deux transpositions : on a $\tau_1 \tau_2 = \tau_2 \tau_1 \Leftrightarrow \begin{cases} \{a, b\} \cap \{c, d\} = \emptyset \\ \text{ou } \{a, b\} = \{c, d\} \end{cases}$

Dans \mathcal{S}_n , il y a $\binom{n}{2} = \frac{n(n-1)}{2}$ transpositions (autant que de parties à deux éléments dans E_n).

18.1.4 Décomposition en produit de cycles à supports disjoints

Proposition 18.1.4 (décomposition en produit de cycles)

Toute permutation de \mathcal{S}_n (avec $n \geq 2$) se décompose en un produit de cycles à supports deux à deux disjoints. Cette décomposition est unique à l'ordre près des facteurs.

Cette décomposition est dite commutative, car les cycles qui la composent commutent deux à deux.

Étude détaillée d'un exemple

Soit la permutation $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 \\ 10 & 5 & 9 & 4 & 14 & 3 & 1 & 11 & 12 & 7 & 13 & 6 & 2 & 8 \end{pmatrix}$.

On a $\sigma(1) = 10$, $\sigma(10) = 7$, et $\sigma(7) = 1$. Il apparaît donc le cycle $\sigma_1 = (1 \ 10 \ 7)$.

En considérant les images successives de 2, on trouve le cycle $\sigma_2 = (2 \ 5 \ 14 \ 8 \ 11 \ 13)$.

En considérant celles de 3 (qui n'est pas apparu dans σ_1 et σ_2) on trouve $\sigma_3 = (3 \ 9 \ 12 \ 6)$.

On voit que $\sigma(4) = 4$, et les éléments restant dans $\{1, \dots, 14\}$ sont tous apparus dans σ_1 , σ_2 , ou σ_3 .

On peut donc écrire $\sigma = \sigma_1 \sigma_2 \sigma_3$.

Les supports de ces cycles sont respectivement $\{1, 7, 10\}$, $\{2, 5, 8, 11, 13, 14\}$ et $\{3, 6, 9, 12\}$.

Ils sont disjoints deux à deux : les cycles $\sigma_1, \sigma_2, \sigma_3$ commutent entre eux.

On pourrait donc aussi écrire : $\sigma = \sigma_3 \sigma_2 \sigma_1 = \sigma_1 \sigma_3 \sigma_2 = \sigma_2 \sigma_3 \sigma_1 = \dots$

On en déduit également le calcul des puissances de σ : $\sigma^m = (\sigma_1 \sigma_2 \sigma_3)^m = \sigma_1^m \sigma_2^m \sigma_3^m$.

Compte tenu des longueurs des cycles $\sigma_1, \sigma_2, \sigma_3$, on a : $\sigma_1^3 = \text{Id}$, $\sigma_2^6 = \text{Id}$, $\sigma_3^4 = \text{Id}$.

Le ppcm de 3, 6, 4 est 12. On a donc $\sigma^{12} = (\sigma_1^3)^4 (\sigma_2^6)^2 (\sigma_3^4)^3 = \text{Id}$.

On vérifie que pour tout entier k compris entre 1 et 12 on a $\sigma^k \neq \text{Id}$.

Si on veut calculer σ^m , on calcule les restes de m dans les divisions euclidiennes par 3, 6, 4.

On observe par exemple que : $2014 \equiv 1 \pmod{3}$, $2014 \equiv 4 \pmod{6}$ et $2014 \equiv 2 \pmod{4}$.

On en déduit $\sigma^{2014} = \sigma_1 \sigma_2^4 \sigma_3^2$.

Or $\sigma_2^4 = (2 \ 11 \ 14) (5 \ 13 \ 8)$, et $\sigma_3^2 = (3 \ 12) (9 \ 6)$.

On trouve donc : $\sigma^{2014} = (1 \ 10 \ 7) (2 \ 11 \ 14) (5 \ 13 \ 8) (3 \ 12) (9 \ 6)$.

Finalement : $\sigma^{2014} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 \\ 10 & 11 & 12 & 4 & 13 & 9 & 1 & 5 & 6 & 7 & 14 & 3 & 8 & 2 \end{pmatrix}$.

Enfin le calcul de σ^{-1} peut s'effectuer en écrivant :

$\sigma^{-1} = \sigma_1^{-1} \sigma_2^{-1} \sigma_3^{-1} = (1 \ 7 \ 10) (2 \ 13 \ 11 \ 8 \ 14 \ 5) (3 \ 6 \ 12 \ 9)$.

On pouvait aussi trouver σ^{-1} directement (en lisant dans σ à partir de la ligne du bas) :

$$\sigma^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 \\ 7 & 13 & 6 & 4 & 2 & 12 & 10 & 14 & 3 & 1 & 8 & 9 & 11 & 5 \end{pmatrix}$$

18.1.5 Décomposition en produit de transpositions

Proposition 18.1.5 (décomposition d'une permutation en produit de transpositions)

Tout cycle de \mathcal{S}_n peut s'écrire comme un produit de transpositions.

Il en découle que toute permutation de \mathcal{S}_n peut s'écrire comme un produit de transpositions.

Remarques

Il n'y a pas unicité de la décomposition d'une permutation en un produit de transpositions.

Par exemple : $\sigma = (1 \ 2 \ 3) = (1 \ 2) (2 \ 3) = (1 \ 2) (1 \ 3) (1 \ 2) (1 \ 3) = (2 \ 3) (1 \ 3)$.

Il y a une façon très simple de décomposer un cycle en produit de transpositions.

En effet : $(a_1 \ a_2 \ a_3 \ \dots \ a_p) = (a_1 \ a_2) (a_2 \ a_3) \cdots (a_k \ a_{k+1}) \cdots (a_{p-1} \ a_p)$.

Pour décomposer une permutation quelconque σ en un produit de transpositions, il est judicieux d'écrire la permutation σ comme un produit de cycles σ_k à supports disjoints, puis d'écrire chaque σ_k comme un produit de transpositions. Par exemple :

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 4 & 7 & 3 & 8 & 6 & 2 & 5 & 1 \end{pmatrix} = (1 \ 4 \ 8) (2 \ 7 \ 5 \ 6) = (1 \ 4) (4 \ 8) (2 \ 7) (7 \ 5) (5 \ 6)$$

18.1.6 Signature d'une permutation

Définition 18.1.4 (signature d'une permutation)

Il existe une et une seule application ε de \mathcal{S}_n dans $\{-1, 1\}$ telle que :

- pour toute transposition τ , on a : $\varepsilon(\tau) = -1$.
- pour toutes permutations σ et σ' , on a : $\varepsilon(\sigma\sigma') = \varepsilon(\sigma)\varepsilon(\sigma')$.

L'application ε est appelée *signature*.

De par la définition (et parce que toute permutation peut s'écrire au moins d'une manière comme un produit de transpositions), il est clair que $\varepsilon(\sigma)$ est toujours égal 1 ou à -1 .

Définition 18.1.5 (permutations paires ou impaires)

Soit σ un élément de \mathcal{S}_n , avec $n \geq 2$.

On dit que σ est une *permutation paire* si $\varepsilon(\sigma) = 1$.

On dit que σ est une *permutation impaire* si $\varepsilon(\sigma) = -1$.

Propriétés immédiates

- Par définition de la signature, les transpositions sont des permutations impaires.
- L'application identité est une permutation paire (elle est le carré d'une transposition quelconque).
- Une permutation σ et son inverse σ^{-1} ont la même signature.
- La composée de deux permutations de même parité est une permutation paire.
La composée de deux permutations de parités opposées est une permutation impaire.
- Si σ est une permutation paire, alors pour tout p de \mathbb{Z} la permutation σ^p est paire.
Si σ est une permutation impaire, alors la permutation σ^p a la parité de l'entier relatif p .

Proposition 18.1.6 (signature et décompositions en produits de transpositions)

Soit σ une permutation de l'ensemble $E_n = \{1, 2, \dots, n\}$.

La parité de σ est celle du nombre de facteurs dans toute décomposition de σ en produit de transpositions.

Dire qu'une permutation σ est paire (resp. impaire), c'est donc dire que les décompositions de σ en produits de transpositions comportent un nombre pair (resp. impair) de facteurs.

Proposition 18.1.7 (nombre de permutations paires, ou impaires)

Dans \mathcal{S}_n , il y a autant de permutations paires que de permutations impaires, c'est-à-dire $\frac{1}{2} n!$.

Proposition 18.1.8 (signature d'un cycle)

La signature d'un cycle σ de longueur p est $\varepsilon(\sigma) = (-1)^{p-1}$. Autrement dit :

- un cycle de longueur paire est une permutation impaire.
- un cycle de longueur impaire est une permutation paire.

Pour calculer la signature d'un élément σ de \mathcal{S}_n , le plus simple est souvent de décomposer σ en cycles à supports disjoints $\sigma = \sigma_1 \circ \sigma_2 \circ \dots \circ \sigma_p$ et d'écrire $\varepsilon(\sigma) = \varepsilon(\sigma_1) \varepsilon(\sigma_2) \dots \varepsilon(\sigma_p)$.

Par exemple : $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 4 & 7 & 3 & 8 & 6 & 2 & 5 & 1 \end{pmatrix} = \sigma_1 \circ \sigma_2$ avec $\sigma_1 = (1 \ 4 \ 8)$ et $\sigma_2 = (2 \ 7 \ 5 \ 6)$.

On a $\varepsilon(\sigma_1) = 1$ et $\varepsilon(\sigma_2) = -1$, donc $\varepsilon(\sigma) = -1$: la permutation σ est impaire.

18.2 Formes n -linéaires alternées

18.2.1 Applications multilinéaires

Définition 18.2.1 (applications multilinéaires)

Soit E et F deux espaces vectoriels sur \mathbb{K} . Soit f une application de E^n dans \mathbb{K} , avec $n \geq 1$.

On dit que f est n -linéaire si elle vérifie la propriété suivante :

Pour tout i de $\{1, \dots, n\}$ et pour tout choix des vecteurs $u_1, \dots, u_{i-1}, u_{i+1}, \dots, u_n$ dans E , l'application f_i de E dans \mathbb{K} définie par $f_i(v) = f(u_1, \dots, u_{i-1}, v, u_{i+1}, \dots, u_n)$ est linéaire.

Une application $f : E^n \rightarrow \mathbb{K}$ est donc n -linéaire (on dit aussi « multilinéaire ») si elle est « linéaire par rapport à chacune de ses variables quand on fixe toutes les autres ».

Remarques et propriétés

– Pour $n = 1$, la n -linéarité se confond avec la linéarité.

Si $n = 2$, on parle d'application *bilinéaire*. Si $n = 3$, on parle d'application *trilinéaire*.

Si $F = \mathbb{K}$, on parle de *forme* n -linéaire.

– Si $f : E^n \rightarrow \mathbb{K}$ est n -linéaire et si l'un des u_i est nul, alors $f(u_1, u_2, \dots, u_n) = \vec{0}$.

Cela résulte en effet de la linéarité par rapport à la i -ième composante.

– Une application f de E^2 dans \mathbb{K} est bilinéaire si et seulement si :

pour tous vecteurs u, u', v, v' de E , et pour tous scalaires $\alpha, \beta, \gamma, \delta$:

$$\begin{aligned} f(\alpha u + \beta u', \gamma v + \delta v') &= \alpha f(u, \gamma v + \delta v') + \beta f(u', \gamma v + \delta v') \\ &= \alpha \gamma f(u, v) + \alpha \delta f(u, v') + \beta \gamma f(u', v) + \beta \delta f(u', v') \end{aligned}$$

– Si $n \geq 2$, on ne confondra pas linéarité et n -linéarité.

Par exemple $\begin{cases} \text{si } f \text{ est linéaire, } f(\lambda u_1, \lambda u_2, \dots, \lambda u_n) = \lambda f(u_1, u_2, \dots, u_n) \\ \text{si } f \text{ est } n\text{-linéaire, } f(\lambda u_1, \lambda u_2, \dots, \lambda u_n) = \lambda^n f(u_1, u_2, \dots, u_n) \end{cases}$

De même, si $n = 2$ $\begin{cases} \text{si } f \text{ linéaire, } f(u + u', v + v') = f(u, v) + f(u', v) = f(u, v') + f(u', v') \\ \text{si } f \text{ est bilinéaire, } f(u + u', v + v') = f(u, v) + f(u, v') + f(u', v) + f(u', v') \end{cases}$

18.2.2 Formes multilinéaires alternées

Définition 18.2.2 (formes multilinéaires alternées sur un espace de dimension n)

Soit E un \mathbb{K} espace vectoriel de dimension $n \geq 1$.

Soit $f : E^n \rightarrow \mathbb{K}$ une forme n -linéaire. On dit que f est *alternée* (ou encore *antisymétrique*) si :

$$\begin{cases} \forall (u_1, u_2, \dots, u_n) \in E^n \\ \forall (i, j) \in \{1, \dots, n\}^2 \text{ avec } i \neq j \end{cases} \quad f(u_1, \dots, u_i, \dots, u_j, \dots, u_n) = -f(u_1, \dots, u_j, \dots, u_i, \dots, u_n).$$

Autrement dit : l'application $f : E^n \rightarrow \mathbb{K}$ est dite alternée si l'échange de deux vecteurs quelconques dans la liste (u_1, \dots, u_n) change le scalaire $f(u_1, \dots, u_n)$ en son opposé.

On note $\mathcal{A}_n(E, \mathbb{K})$ l'ensemble des « formes n -linéaires alternées sur E ».

Attention à la terminologie : on parle de formes n -linéaires alternées « sur E », et on note $\mathcal{A}_n(E, \mathbb{K})$, mais il s'agit bien d'applications qui sont définies sur E^n et qui sont à valeurs dans \mathbb{K} .

Il faut bien noter que l'entier n est présent deux fois dans la définition : d'une part il représente la dimension de E , et d'autre part l'exposant dans le domaine E^n de f .

On vérifie facilement que $\mathcal{A}_n(E, \mathbb{K})$ est un espace vectoriel sur \mathbb{K} (à suivre!).

Proposition 18.2.1

Soit E un \mathbb{K} -espace vectoriel de dimension $n \geq 1$. Soit $f : E^n \rightarrow \mathbb{K}$ une forme n -linéaire alternée.

- si deux des vecteurs u_1, \dots, u_n sont égaux, alors $f(u_1, \dots, u_n) = 0$
- on ne modifie pas $f(u_1, u_2, \dots, u_n)$ en ajoutant à l'un des u_i une combinaison linéaire des **autres**
- si les vecteurs u_1, u_2, \dots, u_n sont liés, alors $f(u_1, u_2, \dots, u_n) = 0$

Proposition 18.2.2 (caractère alterné et permutation sur l'ordre des vecteurs)

Soit E un espace vectoriel sur \mathbb{K} , de dimension $n \geq 1$.

Soit $f : E^n \rightarrow \mathbb{K}$ une forme n -linéaire alternée.

Soit σ une permutation de $\{1, 2, \dots, n\}$, et soit $\varepsilon(\sigma)$ sa signature.

Pour tous vecteurs u_1, u_2, \dots, u_n de E , on a : $f(u_{\sigma(1)}, u_{\sigma(2)}, \dots, u_{\sigma(n)}) = \varepsilon(\sigma)f(u_1, u_2, \dots, u_n)$.

18.2.3 Application « déterminant dans une base »

Proposition 18.2.3 (forme nécessaire d'une forme n -linéaire alternée)

Soit E un \mathbb{K} -espace vectoriel de dimension $n \geq 1$, muni d'une base $e = (e_i)_{1 \leq i \leq n}$.

Soit $f : E^n \rightarrow \mathbb{K}$ une forme n -linéaire alternée sur E . Posons $\lambda = f(e_1, e_2, \dots, e_n)$.

Alors l'application f est déterminée de manière unique par le scalaire λ , en effet :

- soit $u = (u_j)_{1 \leq j \leq n}$ une famille de n vecteurs quelconques de E , de matrice $A = (a_{i,j})$ dans e .
- alors $f(u_1, u_2, \dots, u_n) = \lambda \left(\sum_{\sigma} \varepsilon(\sigma) \prod_{j=1}^n a_{\sigma(j)j} \right)$ (somme sur les $n!$ permutations σ de $\{1, 2, \dots, n\}$)

Définition 18.2.3 (application déterminant dans une base)

Soit E un \mathbb{K} -espace vectoriel de dimension $n \geq 1$, muni d'une base $e = (e_i)_{1 \leq i \leq n}$.

Soit $u = (u_j)_{1 \leq j \leq n}$ une famille de n vecteurs de E , et $A = (a_{i,j})$ la matrice de cette famille dans e .

On pose $\det_e(u_1, u_2, \dots, u_n) = \sum_{\sigma} \varepsilon(\sigma) \prod_{j=1}^n a_{\sigma(j)j}$ (somme étendue aux $n!$ permutations σ de $\{1, 2, \dots, n\}$)

L'application $\det_e : E^n \rightarrow \mathbb{K}$ ainsi définie est appelée « application déterminant dans la base e ».

Cas particulier des petites dimensions

- Si $n = 1$ (E est une droite vectorielle), soit e un vecteur non nul de E .

Pour tout vecteur $u = ae$ de E , on a : $\det_e(u) = a$.

- Si $n = 2$: Soit $e = (e_1, e_2)$ une base du plan E .

Soit $u = ae_1 + a'e_2$ et $v = be_1 + b'e_2$, donc $A = \begin{pmatrix} a & b \\ a' & b' \end{pmatrix}$. Alors $\det_e(u, v) = ab' - a'b$.

– Si $n = 3$: Soit $e = (e_1, e_2, e_3)$ une base de E .

Soit $u = ae_1 + a'e_2 + a''e_3$, $v = be_1 + b'e_2 + b''e_3$, et $w = ce_1 + c'e_2 + c''e_3$ donc $A = \begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix}$.

Alors $\det_e(u, v, w) = a b' c'' - a b'' c' - a' b c'' + a' b'' c + a'' b c' - a'' b' c$.

Proposition 18.2.4 (l'application \det_e est n -linéaire alternée, et non nulle)

Soit E un \mathbb{K} -espace vectoriel de dimension $n \geq 1$. Soit $e = (e_i)_{1 \leq i \leq n}$ une base de E .

Alors l'application $\det_e : E^n \rightarrow \mathbb{K}$ est n -linéaire alternée sur E , et elle vérifie $\det_e(e_1, e_2, \dots, e_n) = 1$

Proposition 18.2.5 (la droite des formes n -linéaires alternées sur un espace de dimension n)

Soit E un \mathbb{K} -espace vectoriel de dimension $n \geq 1$. Soit $e = (e_i)_{1 \leq i \leq n}$ une base de E .

Pour toute forme n -linéaire alternée f sur E , on a : $f = \lambda \det_e$, avec $\lambda = f(e_1, e_2, \dots, e_n)$.

L'espace $\mathcal{A}_n(E, \mathbb{K})$ est une droite vectorielle, et l'application \det_e est une base de cette droite.

L'application \det_e est en fait l'unique élément f de $\mathcal{A}_n(E, \mathbb{K})$ tel que $f(e_1, e_2, \dots, e_n) = 1$.

18.2.4 Relation entre applications « déterminant dans une base »

Proposition 18.2.6 (relation entre deux applications « déterminant dans une base »)

Soit E un \mathbb{K} -espace vectoriel de dimension $n \geq 1$.

Soit $e = (e_i)_{1 \leq i \leq n}$ et $\varepsilon = (\varepsilon_i)_{1 \leq i \leq n}$ deux bases de E .

Pour tous vecteurs u_1, u_2, \dots, u_n de E , on a :

$$\det_\varepsilon(u_1, u_2, \dots, u_n) = \det_\varepsilon(e_1, e_2, \dots, e_n) \det_e(u_1, u_2, \dots, u_n).$$

D'une façon plus compacte, le résultat précédent peut s'écrire : $\det_\varepsilon(u) = \det_\varepsilon(e) \det_e(u)$.

Proposition 18.2.7 (caractérisation des bases)

Soit E un \mathbb{K} -espace vectoriel de dimension $n \geq 1$, muni d'une base $e = (e_i)_{1 \leq i \leq n}$.

Soit $u = (u_i)_{1 \leq i \leq n}$ une famille de n vecteurs de E .

La famille (u) est une base de E si et seulement si $\det_e(u_1, u_2, \dots, u_n) \neq 0$.

18.3 Déterminant d'un endomorphisme, d'une matrice

18.3.1 Déterminant d'un endomorphisme

Proposition 18.3.1

Soit E un \mathbb{K} -espace vectoriel de dimension $n \geq 1$. Soit f un endomorphisme de E .

Le scalaire $\det_e(f(e_1), f(e_2), \dots, f(e_n))$ est indépendant de la base e choisie dans E .

On l'appelle le déterminant de l'endomorphisme f , et on le note $\det(f)$.

Propriétés immédiates

Par définition, le déterminant d'un endomorphisme f est égal au déterminant dans la base e des images par f des vecteurs de e , et ceci pour toute base de E .

En particulier, le déterminant de l'application Id vaut 1.

En effet ce déterminant est égal à $\det_e(e_1, e_2, \dots, e_n)$, pour une base e quelconque.

Pour tout endomorphisme f , tous vecteurs u_1, u_2, \dots, u_n , et toute base e , on a :

$$\det_e(f(u_1), f(u_2), \dots, f(u_n)) = \det(f) \det_e(u_1, u_2, \dots, u_n).$$

Proposition 18.3.2 (déterminant du composé de deux endomorphismes)

Soit E un \mathbb{K} -espace vectoriel de dimension $n \geq 1$.

Soit f et g deux endomorphismes de E . Alors $\det(gf) = \det(g) \det(f)$.

Rappel : on note gf plutôt que $g \circ f$.

Proposition 18.3.3 (déterminant d'un automorphisme)

Soit E un \mathbb{K} -espace vectoriel de dimension $n \geq 1$.

Soit f un endomorphisme de E .

Alors f est un automorphisme si et seulement $\det(f) \neq 0$. Dans ce cas : $\det(f^{-1}) = \frac{1}{\det(f)}$.

Proposition 18.3.4 (déterminant des puissances d'un endomorphisme)

Soit f un endomorphisme de E et soit p un entier naturel. Alors $\det(f^p) = (\det f)^p$.

Ce résultat se généralise aux exposants négatifs si f est un automorphisme.

18.3.2 Déterminant d'une matrice carrée

Définition 18.3.1 (déterminant d'une matrice carrée)

Soit $A = (a_{ij})$ un élément de $\mathcal{M}_n(\mathbb{K})$.

On pose $\det(A) = \sum_{\sigma} \varepsilon(\sigma) \prod_{j=1}^n a_{\sigma(j)j}$, où la somme est étendue aux $n!$ permutations σ de $\{1, 2, \dots, n\}$.

Cette définition, qui exprime $\det(A)$ comme une expression développée des coefficients, est « neutre ».

En variant les points de vue, on aboutit en fait à plusieurs définitions équivalentes possibles...

Définitions équivalentes du déterminant d'une matrice carrée

Soit $A = (a_{ij})$ une matrice carrée d'ordre n à coefficients dans \mathbb{K} .

- Soit C_1, C_2, \dots, C_n les vecteurs-colonne de A , considérés comme des éléments de $\mathcal{M}_{n,1}(\mathbb{K})$.
Alors $\det(A)$ est le déterminant de la famille $(C_j)_{1 \leq j \leq n}$ dans la base canonique de $\mathcal{M}_{n,1}(\mathbb{K})$.
- Soit v_1, v_2, \dots, v_n les vecteurs-colonne de A , considérés comme des éléments de \mathbb{K}^n .
Alors $\det(A)$ est le déterminant de la famille $(v_j)_{1 \leq j \leq n}$ dans la base canonique de \mathbb{K}^n .
- Soit E un espace vectoriel quelconque sur \mathbb{K} , de dimension n , muni d'une base $e = (e_i)_{1 \leq i \leq n}$.
Soit $v = (v_j)_{1 \leq j \leq n}$ la famille de E telle que $A = \text{Mat}_e(v)$: alors $\det(A) = \det_e(v)$.
- Soit $f: \mathbb{K}^n \rightarrow \mathbb{K}^n$, linéaire, de matrice A dans la base canonique de \mathbb{K}^n : alors $\det(A) = \det(f)$.
(ici f est l'application linéaire canoniquement associée à A).

En Python, on dispose de la fonction `det`, dans le module `np.linalg`.

Dans l'exemple ci-contre, on forme une matrice aléatoire d'ordre 3, à coefficients entiers dans $\llbracket 0, 9 \rrbracket$, et on calcule son déterminant.

Le résultat est renvoyé au format `float`, mais c'est ici $\det(A) = 42$ qu'il faut comprendre.

```
>>> import numpy as np
>>> a = np.random.randint(10, size=(3, 3))
>>> print(a)
[[8 3 3]
 [0 3 1]
 [1 9 5]]
>>> np.linalg.det(a)
41.999999999999986
```

Déterminants et matrices de passage

Soit E un \mathbb{K} -espace vectoriel de dimension $n \geq 1$.

Soit e et ε deux bases de E , et soit P la matrice de passage de e à ε .

Alors on a l'égalité : $\det P = \det_e(\varepsilon_1, \dots, \varepsilon_n)$.

Les applications « déterminant dans (e) » et « déterminant dans (ε) » sont reliées par : $\det_e = \det P \det_\varepsilon$.

Ce résultat rappelle l'égalité $[u]_e = P[u]_\varepsilon$ reliant les coordonnées dans e et ε d'un vecteur u de E .

Notation des déterminants

Soit $A = (a_{ij})$ dans $\mathcal{M}_n(\mathbb{K})$. Le déterminant Δ de A est noté $\Delta =$

Plus généralement, c'est un tel « tableau » qu'on appellera déterminant d'ordre n (avec la signification qu'on lui a donné)

sans qu'il soit nécessaire de préciser son « origine »

(c'est-à-dire : matrice, famille de vecteurs, endomorphisme).

On ne confondra surtout pas Δ (une valeur numérique) avec la matrice $A = (a_{ij})$ elle-même.

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix}$$

Déterminants d'ordre 1, 2, ou 3

Pour tout scalaire a , on a bien sûr $|a| = a$ (ne pas confondre avec la valeur absolue...).

Pour tous scalaires a, b, c, d : $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$.

Pour tous $a, a', a'', b, b', b'', c, c', c''$, on a : $\begin{vmatrix} a & a' & a'' \\ b & b' & b'' \\ c & c' & c'' \end{vmatrix} = ab'c'' + bc'a'' + ca'b'' - cb'a'' - ac'b'' - ba'c''$.

Proposition 18.3.5 (déterminant du produit de deux matrices carrées)

Soit A et B deux éléments de $\mathcal{M}_n(\mathbb{K})$. Alors $\det(AB) = (\det A)(\det B)$.

Proposition 18.3.6 (déterminant d'une matrice carrée inversible)

Soit A dans $\mathcal{M}_n(\mathbb{K})$.

Alors A est inversible si et seulement $\det(A) \neq 0$. Dans ce cas : $\det(A^{-1}) = \frac{1}{\det(A)}$.

Conséquence : si les matrices carrées A et B sont semblables, alors elles ont le même déterminant.

Proposition 18.3.7 (déterminant des puissances d'une matrice carrée)

Soit A dans $\mathcal{M}_n(\mathbb{K})$ et soit p un entier naturel. Alors $\det(A^p) = (\det A)^p$.

Ce résultat se généralise aux exposants négatifs si A est une matrice inversible.

Proposition 18.3.8 (déterminant de la transposée d'une matrice carrée)

Soit A une matrice de $\mathcal{M}_n(\mathbb{K})$. Alors $\det(A) = \det(A^\top)$.

Conséquence importante : toutes les propriétés des déterminants qui s'expriment en termes de colonnes peuvent également s'exprimer en termes de lignes.

18.4 Calcul des déterminants

18.4.1 Déterminants et opérations élémentaires

Les propriétés des déterminants découlent ce qu'ils représentent des fonctions multilinéaires alternées de leurs colonnes (et aussi de leurs lignes).

Dans l'énoncé des propriétés suivantes (exprimées en termes de colonnes, mais qui pourraient l'être en termes de lignes), on note Δ un déterminant d'ordre n .

On convient de confondre « la valeur » Δ et « le tableau » Δ . Avec cette convention :

- La valeur d'un déterminant Δ dépend linéairement de chacune de ses colonnes (de ses lignes).
Si on multiplie une colonne (une ligne) par λ , la valeur de Δ est multipliée par λ .
En particulier, pour toute matrice A de $\mathcal{M}_n(\mathbb{K})$, et tout λ de \mathbb{K} , on a : $\det(\lambda A) = \lambda^n \det(A)$.
- Si on échange deux colonnes (deux lignes) de Δ , la valeur de Δ est changée en son opposé.
Plus généralement, si on effectue une permutation sur les colonnes (sur les lignes) de Δ , la valeur de Δ est inchangée (resp. changée en son opposé) selon que cette permutation est paire ou impaire, c'est-à-dire selon qu'elle se décompose en un nombre pair ou impair d'échanges.
- On ne modifie pas la valeur de Δ en ajoutant à l'une de ses colonnes (de ses lignes) une combinaison linéaire des autres colonnes (des autres lignes) de Δ .
- La valeur du déterminant Δ est nulle si et seulement si ses colonnes (ses lignes) sont liées.
En particulier, si Δ contient une colonne (ou une ligne) nulle, alors la valeur de Δ est nulle.

On va résumer en termes d'opérations élémentaires sur les lignes (ou colonnes) :

Proposition 18.4.1 (effet d'une opération élémentaire sur un déterminant)

Soit A dans $\mathcal{M}_n(\mathbb{K})$, et A' obtenue en appliquant à A une opération élémentaire φ .

Soit $\Delta = \det(A)$ et $\Delta' = \det(A')$.

- si φ est l'opération $C_i \leftarrow \alpha C_i$ (ou $L_i \leftarrow \alpha L_i$), avec $\alpha \neq 0$, alors $\Delta' = \alpha \Delta$.
- si φ est l'opération $C_i \leftrightarrow C_j$ (ou $L_i \leftrightarrow L_j$), alors $\Delta' = -\Delta$.
- si φ est l'opération $C_i \leftarrow C_i + \beta C_j$ (ou $L_i \leftarrow L_i + \beta L_j$), avec $j \neq i$, alors $\Delta' = \Delta$.

18.4.2 Développement d'un déterminant, comatrice

Définition 18.4.1 (mineur, cofacteur, comatrice)

Soit A une matrice de $\mathcal{M}_n(\mathbb{K})$, avec $n \geq 2$, de terme général a_{ij} .

Pour tout couple d'indices (i, j) , on appelle *mineur* de a_{ij} dans A (ou dans Δ), le déterminant Δ_{ij} d'ordre $n - 1$ obtenu en supprimant dans Δ la ligne et la colonne de a_{ij} .

La quantité $A_{ij} = (-1)^{i+j} \Delta_{ij}$ est appelée *cofacteur* du coefficient a_{ij} .

On appelle *comatrice* de A et on note $\text{com}(A)$ la matrice de $\mathcal{M}_n(\mathbb{K})$ de terme général A_{ij} .

Exemple dans le cas $n = 3$

$$\text{Si } A = \begin{pmatrix} a & a' & a'' \\ b & b' & b'' \\ c & c' & c'' \end{pmatrix}, \text{ alors } \text{com}(A) = \begin{pmatrix} \begin{vmatrix} b' & b'' \\ c' & c'' \end{vmatrix} & - \begin{vmatrix} b & b'' \\ c & c'' \end{vmatrix} & \begin{vmatrix} b & b' \\ c & c' \end{vmatrix} \\ - \begin{vmatrix} a' & a'' \\ c' & c'' \end{vmatrix} & \begin{vmatrix} a & a'' \\ c & c'' \end{vmatrix} & - \begin{vmatrix} a & a' \\ c & c' \end{vmatrix} \\ \begin{vmatrix} a' & a'' \\ b' & b'' \end{vmatrix} & - \begin{vmatrix} a & a'' \\ b & b'' \end{vmatrix} & \begin{vmatrix} a & a' \\ b & b' \end{vmatrix} \end{pmatrix}$$

Proposition 18.4.2 (développement d'un déterminant suivant une ligne)

Soit A une matrice de $\mathcal{M}_n(\mathbb{K})$, avec $n \geq 2$, de terme général a_{ij} .

Pour tout indice i de $\{1, \dots, n\}$, on a : $\Delta = \sum_{j=1}^n (-1)^{i+j} a_{ij} \Delta_{ij} = \sum_{j=1}^n a_{ij} A_{ij}$.

Cette égalité est appelée *développement de Δ suivant sa i -ème ligne*.

Proposition 18.4.3 (développement d'un déterminant suivant une colonne)

Soit A une matrice de $\mathcal{M}_n(\mathbb{K})$, avec $n \geq 2$, de terme général a_{ij} .

Pour tout indice j de $\{1, \dots, n\}$, on a : $\Delta = \sum_{i=1}^n (-1)^{i+j} a_{ij} \Delta_{ij} = \sum_{i=1}^n a_{ij} A_{ij}$.

Cette égalité est appelée *développement de Δ suivant sa j -ème colonne*.

Proposition 18.4.4 (expression de l'inverse à l'aide de la comatrice)

Soit A une matrice de $\mathcal{M}_n(\mathbb{K})$, avec $n \geq 2$. Alors $A \text{com}(A)^\top = \text{com}(A)^\top A = (\det A) I_n$.

En particulier, si A est inversible, alors $A^{-1} = \frac{1}{\det A} \text{com}(A)^\top$.

La formule précédente possède surtout un intérêt théorique.

Cas très particulier : si $n = 2$ et si $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ est inversible, alors $A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

Voici une fonction Python pour calculer les cofacteurs et la comatrice (on suppose comme d'habitude que le module `numpy` a été préalablement importé) :

```
def cofacteur(a, i, j):
    return (-1)**(i+j)*np.linalg.det(np.delete(np.delete(a, j, 1), i, 0))
def comatrice(a):
    n, p = np.shape(a)
    return np.array([[cofacteur(a, i, j) for j in range(n)] for i in range(p)])
```


On forme ici une matrice aléatoire A d'ordre 5 (à coefficients entiers dans $\llbracket 0, 9 \rrbracket$).

```
>>> a = np.random.randint(10, size=(6, 6)); print(a)
[[6 7 3 3 1 2]
 [2 4 1 6 0 9]
 [6 9 0 6 0 2]
 [0 3 7 7 0 5]
 [7 8 0 1 6 1]
 [2 5 1 1 3 5]]
```

On forme ensuite sa comatrice B (il est prudent de rester en mode `float`, même si on sait que les coefficients de B , tout comme ceux de A , sont entiers) :

```
b = comatrice(a); print(b)
[[ 6117. -1320.  3945. -4800. -4908.  1989.]
 [ 4697. -4916. -1649.  1733.   314.  2832.]
 [-4367.  4739. -2452.  1981. -1190. -2184.]
 [-2225.   104.  1868.  2557.  2275. -1464.]
 [ 2717. -3854. -1010.  3416.  5570. -1056.]
 [-7473.  8148.   705. -5232. -1515.  1449.]]
```

On calcule AB^T , et on obtient une matrice scalaire (il faut arrondir raisonnablement le résultat, pour annuler les coefficients non diagonaux, qui sans cela refléteraient les erreurs d'arrondis) :

```
>>> print(np.dot(a,b.T).round(8))
array([[ 23967.,    0.,    0.,    0.,   -0.,   -0.],
       [   -0.,  23967.,   -0.,    0.,   -0.,   -0.],
       [   -0.,    0.,  23967.,    0.,   -0.,   -0.],
       [   -0.,    0.,   -0.,  23967.,   -0.,    0.],
       [   -0.,   -0.,    0.,    0.,  23967.,   -0.],
       [   -0.,    0.,   -0.,    0.,   -0.,  23967.]])
```

On vérifie enfin que le coefficient sur la diagonale de AB^T est le déterminant de A :

```
>>> np.linalg.det(a)
23967.000000000004
```

18.4.3 Quelques déterminants particuliers

Déterminants triangulaires

Proposition 18.4.5 (déterminants triangulaires)

Soit A une matrice de $\mathcal{M}_n(\mathbb{K})$, triangulaire (supérieure ou inférieure).

Alors $\det(A)$ est égal au produit $\prod_{i=1}^n a_{ii}$ des coefficients diagonaux de A .

C'est le cas en particulier si la matrice est diagonale !

Proposition 18.4.6 (déterminants triangulaires par blocs)

Soit A une matrice carrée triangulaire (supérieure ou inférieure) « par blocs ».

Alors le déterminant de A est égal au produit des déterminants des blocs diagonaux.

$$\text{Par exemple : } \begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} \\ 0 & 0 & 0 & a_{44} & a_{45} & a_{46} \\ 0 & 0 & 0 & 0 & a_{55} & a_{56} \\ 0 & 0 & 0 & 0 & a_{65} & a_{66} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} a_{44} \begin{vmatrix} a_{55} & a_{56} \\ a_{65} & a_{66} \end{vmatrix}$$

Déterminant de Vandermonde

Proposition 18.4.7 (déterminant de Vandermonde)

Soit $x = (x_i)_{1 \leq i \leq n}$ une famille de n scalaires.

Soit A la matrice carrée d'ordre n , de terme général $a_{ij} = x_i^{j-1}$.

Le déterminant de A (et de sa transposée) est appelé « déterminant de Vandermonde » de x_1, x_2, \dots, x_n .

Sa valeur est $\det A = \prod_{i < j} (x_j - x_i)$.

La matrice évoquée dans la définition précédente est $A = \begin{pmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_i & x_i^2 & \dots & x_i^{n-1} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{pmatrix}$

Exemple : $\begin{vmatrix} 1 & x_1 & x_1^2 & x_1^3 \\ 1 & x_2 & x_2^2 & x_2^3 \\ 1 & x_3 & x_3^2 & x_3^3 \\ 1 & x_4 & x_4^2 & x_4^3 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 & 1 \\ x_1 & x_2 & x_3 & x_4 \\ x_1^2 & x_2^2 & x_3^2 & x_4^2 \\ x_1^3 & x_2^3 & x_3^3 & x_4^3 \end{vmatrix} = (x_4 - x_3)(x_4 - x_2)(x_4 - x_1)(x_3 - x_2)(x_3 - x_1)(x_2 - x_1)$

Dans le module `numpy` de Python, on dispose d'une fonction `vander`.

Mais attention, l'indexation n'est pas conforme à notre définition :

```
>>> a = np.vander([1,10,100]); print(a)
[[ 1  1  1]
 [100 10 1]
 [10000 100 1]]
>>> np.linalg.det(a)
-80190.000000000102
```

```
>>> a = np.vander(range(1,6)); print(a)
[[ 1  1  1  1  1]
 [16  8  4  2  1]
 [81 27  9  3  1]
 [256 64 16  4  1]
 [625 125 25  5  1]]
```

Voici comment écrire notre propre fonction `vandermonde` :

```
>>> def vandermonde(x): # ici on attend une liste ou un intervalle
    import numpy as np
    n = len(x); x = np.array(x)
    return np.vstack([x**i for i in range(n)])
```

```
>>> a = vandermonde([1,10,100]); print(a)
[[ 1  1  1]
 [ 1 10 100]
 [ 1 100 10000]]
>>> np.linalg.det(a)
80189.999999999971
```

```
>>> vandermonde(range(1,6))
array([[ 1,  1,  1,  1,  1],
       [ 1,  2,  4,  8, 16],
       [ 1,  3,  9, 27, 81],
       [ 1,  4, 16, 64, 256],
       [ 1,  5, 25, 125, 625]])
```

18.5 Déterminants et orientation

18.5.1 Orientation d'un espace réel de dimension finie

Proposition 18.5.1 (orientation d'un espace réel de dimension finie)

Soit E un espace vectoriel sur \mathbb{R} , de dimension $n \geq 1$.

Soit \mathcal{B} et \mathcal{B}' deux bases de E . Soit P la matrice de passage de \mathcal{B} à \mathcal{B}' .

Si $\det P > 0$, on dit que la base \mathcal{B}' a la même orientation que la base \mathcal{B} .

On définit ainsi une relation d'équivalence sur l'ensemble des bases de E .

Pour cette relation, il y a exactement deux classes d'équivalence.

Orienter E , c'est choisir l'une de ces deux classes.

- les bases de la classe d'équivalence choisie sont dites directes.
- les bases de l'autre classe d'équivalence sont dites indirectes.

Effet d'une permutation des vecteurs de base

Supposons que la base \mathcal{B}' se déduise de \mathcal{B} par une permutation σ sur les vecteurs de \mathcal{B} .

- si σ est une transposition, alors \mathcal{B} et \mathcal{B}' sont d'orientation contraire.
- si σ est une permutation paire, les bases \mathcal{B} et \mathcal{B}' sont de même orientation.
- si σ est une permutation impaire, alors elles sont d'orientation contraire.

Effet de l'opération $u \mapsto -u$ sur l'orientation d'une base

Si on passe de \mathcal{B} à \mathcal{B}' en changeant un vecteur en son opposé alors $\mathcal{B}, \mathcal{B}'$ sont d'orientation contraire.

Par exemple, si $\dim(E) = 2$, supposons que (u, v) soit une base directe de E .

- les bases $(-u, v)$, $(u, -v)$, (v, u) et $(-v, -u)$ sont indirectes.
- les bases (u, v) , $(-u, -v)$, $(v, -u)$, et $(-v, u)$ sont directes.

De même, si $\dim(E) = 3$, supposons que (u, v, w) soit une base directe de E .

- les bases $(-u, v, w)$, $(u, -v, w)$, $(u, v, -w)$ et $(-u, -v, -w)$ sont indirectes.
- les bases (v, u, w) , (w, v, u) , (u, w, v) sont indirectes, etc.
- les bases (u, v, w) , $(u, -v, -w)$, $(-u, v, -w)$, et $(-u, -v, w)$ sont directes.
- les bases (v, w, u) , (w, u, v) sont directes, etc.

En résumé : il y a toujours deux orientations possibles sur un \mathbb{R} -espace vectoriel de dimension finie.

Le choix de la classe des bases dites positives est arbitraire.

Néanmoins on orientera toujours \mathbb{R}^n en décrétant que sa base canonique est directe.

18.5.2 Si $n = 2$, le déterminant est une aire orientée

On se place ici dans \mathbb{R}^2 , muni de son orientation canonique.

Définition 18.5.1 (parallélogramme dans \mathbb{R}^2)

Soit u, v deux vecteurs de \mathbb{R}^2 .

On appelle *parallélogramme* formé sur u, v l'ensemble des vecteurs $\alpha u + \beta v$, avec $0 \leq \alpha \leq 1$ et $0 \leq \beta \leq 1$.

Si on note $\mathcal{P}_{u,v}$ l'ensemble précédent, on a bien sûr $\mathcal{P}_{u,v} = \mathcal{P}_{v,u}$.

On pourra appeler « parallélogramme unité » le parallélogramme formé sur les vecteurs $(1, 0)$ et $(0, 1)$ de la base canonique, c'est-à-dire les couples (α, β) , avec $0 \leq \alpha \leq 1$ et $0 \leq \beta \leq 1$.

Définition 18.5.2 (aire orientée d'un parallélogramme dans \mathbb{R}^2)

On se place dans \mathbb{R}^2 , muni de sa base canonique e , et de son orientation canonique (la base e est directe).

Soit u, v deux vecteurs de \mathbb{R}^2 , et soit $\mathcal{P}_{u,v}$ le parallélogramme formé sur u et v .

La quantité $\det_e(u, v)$ est appelée « aire orientée » du parallélogramme $\mathcal{P}_{u,v}$.

Avec cette définition, l'aire du « parallélogramme unité » est égale à 1.

L'aire orientée du parallélogramme $\mathcal{P}_{v,u}$ est l'opposée de celle du parallélogramme $\mathcal{P}_{u,v}$.

L'aire orientée de $\mathcal{P}_{u,v}$ est nulle si et seulement si u et v sont liés (« parallélogramme plat »).

Lien avec l'aire au sens usuel

On suppose ici que les vecteurs u et v sont libres (donc forment une base du plan \mathbb{R}^2).

Dans ces conditions, $\det_e(u, v)$ est strictement positif si la base (u, v) est directe, strictement négatif sinon. Le signe de l'aire orientée du parallélogramme $\mathcal{P}_{u,v}$ reflète donc l'orientation de la base u, v .

Supposons par exemple que la base u, v est directe.

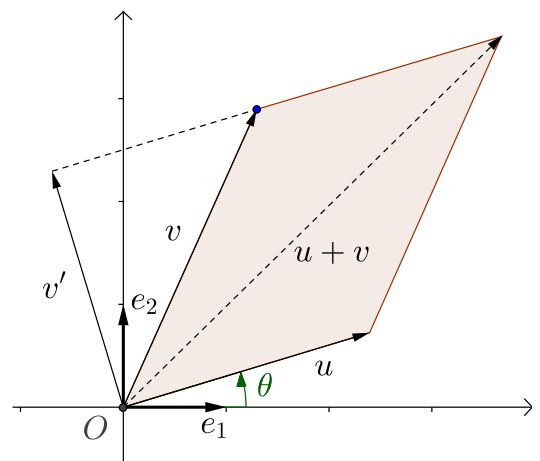
Soit v' la projection de v sur la droite orthogonale à u .

Soit $\varepsilon_1 = (\cos \theta, \sin \theta)$ et $\varepsilon_2 = (-\sin \theta, \cos \theta)$ les vecteurs unitaires normalisés respectifs de u et v' .

Avec ces notations, on a :

$$\begin{aligned} \det_e(u, v) &= \|u\| \|v'\| \det_e(\varepsilon_1, \varepsilon_2) \\ &= \|u\| \|v'\| \begin{vmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{vmatrix} = \|u\| \|v'\| \end{aligned}$$

Enfin, la quantité $\|u\| \|v'\|$ est bien l'aire, au sens usuel, du parallélogramme formé sur u et v (le produit de la « base » $\|u\|$, par la « hauteur » $\|v'\|$).



18.5.3 Si $n = 3$, le déterminant est un volume orienté

Définition 18.5.3 (parallélépipède dans \mathbb{R}^3)

Soit u, v, w trois vecteurs de \mathbb{R}^3 .

On appelle *parallélépipède* formé sur u, v, w l'ensemble des vecteurs $\alpha u + \beta v + \delta w$, avec
$$\begin{cases} 0 \leq \alpha \leq 1 \\ 0 \leq \beta \leq 1 \\ 0 \leq \delta \leq 1 \end{cases}$$

Si on note $\mathcal{P}_{u,v,w}$ l'ensemble précédent, on a bien sûr $\mathcal{P}_{u,v,w} = \mathcal{P}_{v,u,w} = \mathcal{P}_{w,u,v} = \dots$

On pourra appeler « parallélépipède unité » le parallélépipède formé sur les vecteurs $(1, 0, 0)$, $(0, 1, 0)$ et $(0, 0, 1)$ de la base canonique, c'est-à-dire les couples (α, β, δ) , avec $0 \leq \alpha \leq 1$, $0 \leq \beta \leq 1$ et $0 \leq \delta \leq 1$.

Définition 18.5.4 (volume orienté d'un parallélépipède dans \mathbb{R}^3)

On se place dans \mathbb{R}^3 , muni de sa base canonique e , et de son orientation canonique (la base e est directe).

Soit u, v, w trois vecteurs de \mathbb{R}^3 , et soit $\mathcal{P}_{u,v,w}$ le parallélépipède formé sur u, v, w .

La quantité $\det_e(u, v, w)$ est appelée « volume orienté » du parallélépipède $\mathcal{P}_{u,v,w}$.

Avec cette définition, le volume du « parallélépipède unité » est égal à 1.

Si on échange deux vecteurs dans la famille u, v, w , le volume orienté est changé en son opposé.

Le volume orienté de $\mathcal{P}_{u,v,w}$ est nulle si et seulement si u, v, w sont liés (« parallélépipède plat »).

Lien avec le volume au sens usuel

On suppose ici que les vecteurs u, v, w sont libres (donc forment une base de \mathbb{R}^3).

Alors $\det_e(u, v, w)$ est strictement positif si la base (u, v, w) est directe, strictement négatif sinon.

Le signe du volume orienté du parallélépipède $\mathcal{P}_{u,v,w}$ reflète donc l'orientation de la base u, v, w .

Supposons pour fixer les idées, que la base u, v, w est directe.

Dans le plan $\mathbb{R}u \oplus \mathbb{R}v$, soit v' la projection de v sur la droite vectorielle orthogonale à u .

Soit w' la projection de w sur la droite orthogonale au plan $\mathbb{R}u \oplus \mathbb{R}v = \mathbb{R}u \oplus \mathbb{R}v'$ (faire un dessin).

Soit $\varepsilon_1, \varepsilon_2$ et ε_3 les vecteurs unitaires normalisés respectifs de u, v', w' .

La base $\varepsilon_1, \varepsilon_2, \varepsilon_3$ est orthonormale directe, donc $\det_e(\varepsilon_1, \varepsilon_2, \varepsilon_3) = 1$.

Alors $\det_e(u, v, w) = \det_e(u, v, w') = \det_e(u, v', w') = \|u\| \|v'\| \|w'\| \det_e(\varepsilon_1, \varepsilon_2, \varepsilon_3) = \|u\| \|v'\| \|w'\|$.

Enfin, $\|u\| \|v'\| \|w'\|$ est bien le volume, au sens usuel, du parallélépipède formé sur les vecteurs u, v, w (c'est-à-dire le produit de l'aire de la « base » $\|u\| \|v'\|$ par la « hauteur » $\|w'\|$).

Chapitre 19

Espaces préhilbertiens réels

Sommaire

19.1	Produit scalaire, norme et distance	423
19.1.1	Produit scalaire sur un \mathbb{R} espace vectoriel	423
19.1.2	Norme et distance associée	425
19.2	Orthogonalité	426
19.2.1	Vecteurs orthogonaux	426
19.2.2	Orthogonal d'une partie	428
19.2.3	Algorithme d'orthonormalisation de Schmidt	429
19.2.4	Calculs dans une base orthonormale	430
19.3	Produit mixte, produit vectoriel	431
19.3.1	Produit mixte dans un espace euclidien orienté	431
19.3.2	Produit vectoriel	433
19.4	Projections orthogonales	434
19.4.1	Supplémentaire orthogonal	434
19.4.2	Projection orthogonale	434
19.4.3	Distance à un sous-espace	436
19.5	Hyperplans affines d'un espace euclidien	436
19.5.1	Vecteur normal à un hyperplan d'un espace euclidien	436
19.5.2	Équations d'un hyperplan dans une base orthonormale	438
19.5.3	Calcul de la distance à un hyperplan affine	439
19.5.4	Orientation d'un hyperplan par un vecteur normal	440
19.6	Isométries vectorielles	440
19.6.1	Isométries vectorielles	440
19.6.2	Symétries vectorielles orthogonales	441
19.7	Matrices orthogonales	442
19.7.1	Matrices orthogonales	442
19.7.2	Matrices orthogonales positives ou négatives	443
19.7.3	Isométries positives, négatives	444
19.8	Isométries en dimension 2	445
19.8.1	Matrices orthogonales de taille 2	445
19.8.2	Angle de rotations et de vecteurs du plan	445
19.8.3	Classification des isométries d'un plan euclidien orienté	446

19.1 Produit scalaire, norme et distance

Dans tout le chapitre, E désigne un espace vectoriel sur \mathbb{R} .

19.1.1 Produit scalaire sur un \mathbb{R} espace vectoriel

Définition 19.1.1 (produit scalaire sur un \mathbb{R} espace vectoriel)

Soit E un espace vectoriel sur \mathbb{R} . Soit f une application de $E \times E$ dans \mathbb{R} .

On dit que $f : E \times E \rightarrow \mathbb{R}$ est un *produit scalaire* sur E si elle vérifie les propriétés suivantes :

- l'application f est *bilinéaire*.
- pour tous vecteurs u et v de E , on a $f(u, v) = f(v, u)$ (on dit que f est *symétrique*).
- pour tout vecteur u de E , on a : $f(u, u) \geq 0$ (on dit que f est *positive*).
- pour tout vecteur u de E , on a : $f(u, u) = 0 \Leftrightarrow u = \vec{0}$ (on dit que f est *définie*).

Rappelons que la bilinéarité s'écrit :
$$\begin{cases} \forall (u, u', v, v') \in E^4 \\ \forall (\alpha, \beta) \in \mathbb{R}^2 \end{cases} \begin{cases} f(\alpha u + \beta u', v) = \alpha f(u, v) + \beta f(u', v) \\ f(u, \alpha v + \beta v') = \alpha f(u, v) + \beta f(u, v') \end{cases}$$

Si le caractère symétrique de f est établi, la « linéarité à droite » équivaut à la « linéarité à gauche ».

Un produit scalaire sur E est donc une « forme bilinéaire symétrique définie positive ».

Définition 19.1.2 (espace préhilbertien réel, espace euclidien)

Un \mathbb{R} espace vectoriel E muni d'un produit scalaire est dit *préhilbertien réel*.

Un espace *euclidien* est un espace préhilbertien réel de dimension finie.

Notations

Plutôt que de noter $f(u, v)$, on note souvent $\langle u, v \rangle$, ou $u \cdot v$, ou $(u | v)$.

Avec la notation $(\cdot | \cdot)$, que nous utiliserons, la définition d'un produit scalaire devient :

$$\forall (u, u', v, v') \in E^4, \forall (\alpha, \beta) \in \mathbb{R}^2 : \begin{cases} (\alpha u + \beta u' | v) = \alpha (u | v) + \beta (u' | v) \\ (u | \alpha v + \beta v') = \alpha (u | v) + \beta (u | v') \\ (u | v) = (v | u) \quad (u | u) \geq 0 \quad \text{et} \quad (u | u) = 0 \Leftrightarrow u = \vec{0} \end{cases}$$

Proposition 19.1.1 (produit scalaire canonique sur \mathbb{R}^n)

Soit $u = (x_1, \dots, x_n)$ et $v = (y_1, \dots, y_n)$ deux éléments quelconques de \mathbb{R}^n .

En posant $(u | v) = \sum_{k=1}^n x_k y_k$, on définit un produit scalaire sur \mathbb{R}^n .

On l'appelle le produit scalaire canonique de \mathbb{R}^n .

Notation matricielle :

Si on note $[u]$ la matrice-colonne associée à tout vecteur u de \mathbb{R}^n , alors $(u | v) = [u]^\top [v]$.

Proposition 19.1.2 (un produit scalaire sur $\mathcal{C}([a, b], \mathbb{R})$)

Soit $E = \mathcal{C}([a, b], \mathbb{R})$ l'espace vectoriel des applications continues de $[a, b]$ dans \mathbb{R} , avec $a < b$.

En posant $(f | g) = \int_a^b f(t) g(t) dt$, on définit un produit scalaire sur E .

19.1.2 Norme et distance associée

Proposition 19.1.3 (norme et distance associée à un produit scalaire)

Soit E un espace préhilbertien réel.

Pour tout vecteur u de E , on appelle norme de u la quantité $\|u\| = \sqrt{(u | u)}$.

Pour tous vecteurs u, v on appelle distance de u à v la quantité $d(u, v) = \|u - v\|$.

Les applications « norme » et « distance » sont dites associées au produit scalaire sur E .

Définition 19.1.3 (vecteurs unitaires)

Soit E un espace préhilbertien réel. Un vecteur u de E est dit *unitaire* (ou encore *normé*) si $\|u\| = 1$.

Proposition 19.1.4 (inégalité de Cauchy-Schwarz)

Soit E un espace préhilbertien réel.

Pour tous vecteurs u, v de E , on a l'inégalité dite « de Cauchy-Schwarz » : $|(u | v)| \leq \|u\| \|v\|$.

Il y a égalité dans ce résultat si et seulement si u et v sont liés.

Les deux exemples classiques

– On se place dans \mathbb{R}^n , muni de son produit scalaire canonique.

Pour tout vecteur $u = (x_1, \dots, x_n)$, on a $\|u\| = \left(\sum_{k=1}^n x_k^2\right)^{1/2}$.

Pour tous $\begin{cases} u = (x_1, \dots, x_n) \\ v = (y_1, \dots, y_n) \end{cases}$ l'inégalité de Cauchy-Schwarz s'écrit : $\left(\sum_{k=1}^n x_k y_k\right)^2 \leq \sum_{k=1}^n x_k^2 \sum_{k=1}^n y_k^2$

– On se place dans $\mathcal{C}([a, b], \mathbb{R})$ muni de $(f | g) = \int_a^b f(t) g(t) dt$. Alors $\|f\| = \sqrt{\int_a^b f(t)^2 dt}$.

Dans ce cas, l'inégalité de Cauchy-Schwarz s'écrit $\left(\int_a^b f(t) g(t) dt\right)^2 \leq \int_a^b f(t)^2 dt \int_a^b g(t)^2 dt$

Proposition 19.1.5 (propriétés de la norme associée à un produit scalaire)

Soit E un espace préhilbertien réel.

– pour tout vecteur u de E , on a l'inégalité $\|u\| \geq 0$, et l'équivalence $\|u\| = 0 \Leftrightarrow u = \vec{0}$.

– pour tout vecteur u de E , et pour tout réel λ , on a : $\|\lambda u\| = |\lambda| \|u\|$.

– pour tous vecteurs u, v de E , on a l'inégalité triangulaire : $\|u + v\| \leq \|u\| + \|v\|$.

cette inégalité est une égalité si et seulement si u et v sont « positivement liés ».

Remarques

– l'expression « positivement liés » signifie l'existence de λ dans \mathbb{R}^+ tel que $v = \lambda u$ ou $u = \lambda v$.

– pour tous vecteurs u et v de E , on a l'encadrement : $\left| \|u\| - \|v\| \right| \leq \|u \pm v\| \leq \|u\| + \|v\|$.

– Si u est non nul, les vecteurs $\pm \frac{u}{\|u\|}$ sont les seuls vecteurs unitaires de la droite $\mathbb{R}u$.

Proposition 19.1.6 (propriétés de la distance associée à un produit scalaire)

Soit E un espace préhilbertien réel. On note $d(u, v)$ la distance associée.

- pour tous vecteurs u et v de E , on a $d(u, v) = d(v, u)$.
- pour tous vecteurs u et v de E , on a l'inégalité $d(u, v) \geq 0$, et l'équivalence ($d(u, v) = 0 \Leftrightarrow u = v$).
- pour tous vecteurs u, v, w , on a $d(u, v) = d(u + w, v + w)$ (la distance est invariante par translation).
- pour tous vecteurs u, v, w de E , on a l'inégalité triangulaire : $d(u, v) \leq d(u, w) + d(w, v)$;
il y a égalité dans ce résultat si et seulement si il existe λ dans $[0, 1]$ tel que $w = \lambda u + (1 - \lambda)v$.

Remarques

La notion de distance est surtout utilisée dans le cadre de la géométrie affine.

On parle alors de la distance $d(A, B) = \|\overrightarrow{AB}\|$ entre deux points A et B .

Avec ces notations, $d(A, B) \leq d(A, C) + d(C, B)$ (égalité si et seulement si C est sur le segment $[A; B]$)

Proposition 19.1.7 (un développement usuel)

Soit E un espace préhilbertien réel.

Pour tous u, v de E , et tous réels α, β on a : $\|\alpha u + \beta v\|^2 = \alpha^2 \|u\|^2 + 2\alpha\beta (u | v) + \beta^2 \|v\|^2$.

En particulier,
$$\begin{cases} \|u + v\|^2 = \|u\|^2 + 2(u | v) + \|v\|^2 \\ \|u - v\|^2 = \|u\|^2 - 2(u | v) + \|v\|^2 \end{cases}$$

Par addition, on en déduit : $\forall (u, v) \in E^2, \|u + v\|^2 + \|u - v\|^2 = 2(\|u\|^2 + \|v\|^2)$.

Cette égalité est connue sous le nom d'*identité du parallélogramme*.

Proposition 19.1.8 (formule de polarisation)

Soit E un espace préhilbertien réel.

Pour tous vecteurs u, v de E , on a : $(u | v) = \frac{1}{2} (\|u + v\|^2 - \|u\|^2 - \|v\|^2)$.

19.2 Orthogonalité

19.2.1 Vecteurs orthogonaux

Définition 19.2.1 (orthogonalité de deux vecteurs)

Soit E un espace préhilbertien réel.

Deux vecteurs u et v de E sont dits *orthogonaux* s'ils vérifient $(u | v) = 0$.

Remarques

- ces notions dépendent évidemment du produit scalaire utilisé sur E . si on en change, les vecteurs qui étaient orthogonaux ne le sont donc plus nécessairement.
- la définition de l'orthogonalité est symétrique car $(u | v) = (v | u)$.

- le seul vecteur u qui est orthogonal à lui-même est le vecteur nul.
a fortiori, le seul vecteur u qui est orthogonal à tous les vecteurs de E est $u = \vec{0}$.

Définition 19.2.2 (familles orthogonales ou orthonormales)

Soit E un espace préhilbertien réel.

On dit qu'une famille $(u_i)_{i \in I}$ de vecteurs de E est *orthogonale* si les u_i sont orthogonaux deux à deux. Si de plus ils sont unitaires, alors la famille est dite *orthonormale* (ou orthonormée).

La famille $(u_i)_{i \in I}$ est orthonormale $\Leftrightarrow \forall (i, j) \in I^2, (u_i | u_j) = \delta_{ij}$ (notations de Kronecker).

Deux exemples classiques

– La base canonique de \mathbb{R}^n est orthonormale pour le produit scalaire canonique.

– On se place dans $\mathcal{C}([0, 2\pi], \mathbb{R})$ muni du produit scalaire $(f | g) = \int_0^{2\pi} f(t)g(t) dt$.

La famille des $f_n : x \mapsto \cos(nx)$, avec n dans \mathbb{N} , est orthogonale pour ce produit scalaire.

Proposition 19.2.1 (une famille orthogonale de vecteurs non nuls est libre)

Soit E un espace préhilbertien réel.

Si une famille $(u_i)_{i \in I}$ est orthogonale et formée de vecteurs non nuls, alors c'est une famille libre.

C'est le cas notamment d'une famille $(u_i)_{i \in I}$ orthonormale.

En particulier, si $\dim E = n \geq 1$, une famille orthonormale de n vecteurs est une base orthonormale.

Proposition 19.2.2 (théorème de Pythagore)

Soit E un espace préhilbertien réel.

Si la famille $(u_k)_{1 \leq k \leq p}$ est orthogonale, alors $\left\| \sum_{k=1}^p u_k \right\|^2 = \sum_{k=1}^p \|u_k\|^2$ (relation de Pythagore).

Attention, la réciproque n'est vraie que si $p = 2$.

Ainsi : $(u | v) = 0 \Leftrightarrow \|u + v\|^2 = \|u\|^2 + \|v\|^2$.

19.2.2 Orthogonal d'une partie**Définition 19.2.3** (orthogonal d'une partie d'un espace préhilbertien)

Soit E un espace préhilbertien réel.

Soit X une partie non vide de E .

L'*orthogonal* de X , noté X^\perp , est l'ensemble des vecteurs orthogonaux à tous les éléments de X .

Définition 19.2.4

Soit E un espace préhilbertien réel.

Soit X, Y deux parties non vides de E .

On dit que les parties X et Y sont orthogonales si : $\forall u \in X, \forall v \in Y, (u | v) = 0$.

Cela équivaut à l'inclusion $Y \subset X^\perp$ (ou bien sûr à l'inclusion $X \subset Y^\perp$).

Propriétés

- De manière évidente, on a $\{\vec{0}\}^\perp = E$, et $E^\perp = \{\vec{0}\}$. Si $X \subset Y$, alors $Y^\perp \subset X^\perp$.
- L'orthogonal X^\perp de X est *toujours* un sous-espace vectoriel de E , même si X n'en est pas un.
- Si X est une partie non vide de E , alors $X^\perp = \text{Vect}(X)^\perp$.
En particulier, si $X = \text{Vect}\{u_j, j \in J\}$, alors : $v \in X^\perp \Leftrightarrow \forall j \in J, (v | u_j) = 0$.
- Pour toute partie non vide de E , on a l'inclusion $X \subset X^{\perp\perp}$ (X est inclus dans son *double orthogonal*). Cette inclusion peut être stricte, notamment si X n'est pas un sous-espace vectoriel de E .
Par exemple, dans \mathbb{R}^3 muni de son produit scalaire canonique (et en notant e_1, e_2, e_3 les vecteurs de la base canonique) : si $X = \{e_1, e_2\}$, alors $X^\perp = \mathbb{R}e_3$, puis $X^{\perp\perp} = \mathbb{R}e_1 \oplus \mathbb{R}e_2$.
- Si F est un sous-espace vectoriel de E , alors $F \cap F^\perp = \{\vec{0}\}$: la somme $F + F^\perp$ est donc directe.
La proposition suivante généralise ce résultat :

Proposition 19.2.3 (sommées directes orthogonales)

Soit E un espace préhilbertien réel.

Soit $(F_j)_{j \in J}$ une famille de sous-espaces de E , orthogonaux deux à deux.

Alors la somme $G = \sum F_j$ est directe, et on notera $G = \bigoplus^\perp F_j$ (on parle de somme directe orthogonale).

19.2.3 Algorithme d'orthonormalisation de Schmidt**Proposition 19.2.4** (principe de l'algorithme)

Soit E un espace préhilbertien réel. Soit $(e_k)_{1 \leq k \leq n}$ une famille libre de E .

Pour tout k de $\{1, \dots, n\}$, on pose $\varepsilon_k = \frac{1}{\|u_k\|} u_k$, où $u_k = e_k - \sum_{j=1}^{k-1} (\varepsilon_j | e_k) \varepsilon_j$

La première étape de l'algorithme consiste bien sûr à poser $\varepsilon_1 = \frac{1}{\|e_1\|} e_1$

Proposition 19.2.5 (preuve de l'algorithme)

Avec les notations précédentes, l'algorithme de Schmidt se termine.

Pour tout k de $\{1, \dots, n\}$, on a $\text{Vect}\{\varepsilon_1, \dots, \varepsilon_k\} = \text{Vect}\{e_1, \dots, e_k\}$, et $(\varepsilon_k | e_k) > 0$.

La famille $(\varepsilon_k)_{1 \leq k \leq n}$ ainsi construite est donc une famille orthonormée.

Illustration du procédé

On illustre le passage d'une famille libre $e = (e_1, e_2, e_3)$ à une famille orthonormale $\varepsilon = (\varepsilon_1, \varepsilon_2, \varepsilon_3)$.

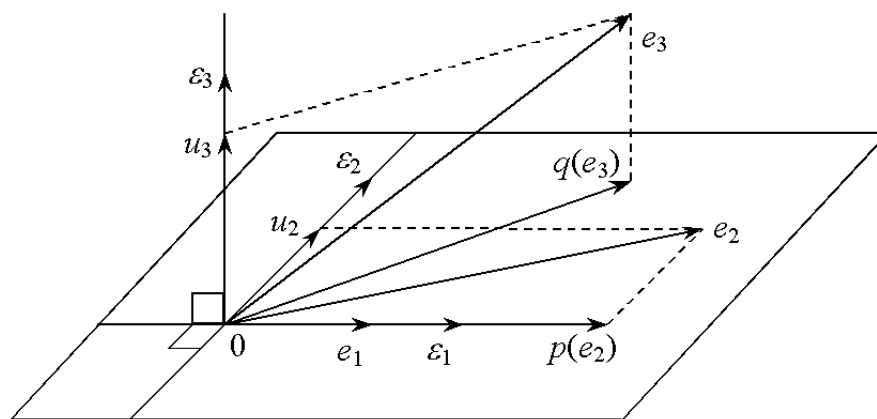
On a conservé les notations de la proposition en ce qui concerne les vecteurs u_2 et u_3 .

On a cependant noté $p(e_2) = (\varepsilon_1 | e_2) \varepsilon_1$, donc $u_2 = e_2 - p(e_2)$.

De même, on a noté $q(e_3) = (\varepsilon_1 | e_3) \varepsilon_1 + (\varepsilon_2 | e_3) \varepsilon_2$, donc $u_3 = e_3 - q(e_3)$.

On voit bien, ce qui sera repris plus tard, que :

- le vecteur $p(e_2)$ est la « projection orthogonale » de e_2 sur la droite engendrée par ε_1 (donc par e_1)
- le vecteur $q(e_3)$ est la projection orthogonale de e_3 sur le plan engendré par $\varepsilon_1, \varepsilon_2$ (donc par e_1, e_2)



Si on se place dans un espace E de dimension finie, le résultat ci-dessous découle immédiatement de l'algorithme d'orthonormalisation appliqué à une base quelconque e de E .

Proposition 19.2.6 (existence de bases orthonormales dans un espace euclidien)

Soit E un espace euclidien (c'est-à-dire un espace préhilbertien réel de dimension finie).

Alors, dans l'espace vectoriel E , il existe des bases orthonormales.

Proposition 19.2.7 (théorème de la base orthonormée incomplète)

Soit E un espace euclidien de dimension $n \geq 1$.

Soit $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p)$ une famille orthonormée de E , avec $p < n$ (donc non génératrice).

Alors il est possible de compléter ε en une base orthonormée $\varepsilon' = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p, \varepsilon_{p+1}, \dots, \varepsilon_n)$ de E .

19.2.4 Calculs dans une base orthonormale

Proposition 19.2.8 (expressions des coordonnées dans une base orthonormale)

Soit E un espace euclidien de dimension $n \geq 1$, et soit $\varepsilon = (\varepsilon_k)_{1 \leq k \leq n}$ une base orthonormale de E .

Pour tout vecteur u de E , on a : $u = \sum_{k=1}^n (u | \varepsilon_k) \varepsilon_k$.

Tout vecteur u de E est donc entièrement déterminé par ses produits scalaires sur les vecteurs de ε .

Les applications coordonnées dans la base ε sont les applications $u \mapsto (u | \varepsilon_k)$.

Proposition 19.2.9 (expressions du produit scalaire et de la norme dans une base orthonormale)

Soit E un espace euclidien de dimension $n \geq 1$, et soit $\varepsilon = (\varepsilon_k)_{1 \leq k \leq n}$ une base orthonormale de E .

Pour tous vecteurs $x = \sum_{k=1}^n x_k \varepsilon_k$ et $y = \sum_{k=1}^n y_k \varepsilon_k$, on a : $(x | y) = \sum_{k=1}^n x_k y_k$ et $\|x\|^2 = \sum_{k=1}^n x_k^2$.

Expression matricielle du produit scalaire dans une base orthonormale

Soit $\varepsilon = (\varepsilon_k)_{1 \leq k \leq n}$ une base orthonormale de E .

On peut écrire $(x | y) = X^T Y$, en notant X, Y les matrices-colonne des coordonnées de x, y dans ε .

Expression matricielle du produit scalaire dans une base quelconque

Si la base ε de E est *quelconque*, alors on a seulement $(x | y) = \sum_{j=1}^n \sum_{k=1}^n x_j y_k (\varepsilon_j | \varepsilon_k)$.

On peut alors écrire : $(x | y) = X^T G Y$, où G est la matrice des $(\varepsilon_i | \varepsilon_j)$, pour $1 \leq i, j \leq n$.

Matrice d'une famille de vecteurs dans une base orthonormale

Soit $\varepsilon = (\varepsilon_i)_{1 \leq i \leq n}$ une base orthormée de E , et soit $v = (v_j)_{1 \leq j \leq n}$ une famille de n vecteurs de E .

Soit $A = (a_{i,j})$ la matrice de la famille v dans la base ε . Alors $a_{i,j} = (\varepsilon_i | v_j)$ pour tous i, j .

Si on note ε la base orthonormée obtenue à partir d'une base quelconque e de E par l'algorithme de Schmidt, alors les matrices de passage $P_{e,\varepsilon}$ (de e à ε) et $P_{\varepsilon,e}$ (de ε à e , inverse de la matrice précédente) sont triangulaires supérieures avec coefficients diagonaux strictement positifs.

19.3 Produit mixte, produit vectoriel

19.3.1 Produit mixte dans un espace euclidien orienté

Dans cette section on se place dans un espace vectoriel euclidien orienté E .

Il y a dans E des bases orthonormales directes et des bases orthonormales indirectes. En effet, si $\mathcal{B} = (e_1, e_2, \dots, e_n)$ est orthonormale, $\mathcal{B}' = (-e_1, e_2, \dots, e_n)$ est orthonormale d'orientation contraire.

Proposition 19.3.1 (produit mixte)

Soit u_1, \dots, u_n une famille de n vecteurs d'un espace euclidien orienté E de dimension n .

Le déterminant $\det_e(u_1, u_2, \dots, u_n)$ est le même dans toute base orthonormale directe e .

Ce déterminant est appelé produit mixte de u_1, u_2, \dots, u_n et il est noté $[u_1, u_2, \dots, u_n]$.

Produit mixte et orientation

L'application « produit mixte » est une forme n -linéaire alternée sur E .

Si e_1, e_2, \dots, e_n forment une base orthonormale directe, alors $[e_1, e_2, \dots, e_n] = 1$.

Si e_1, e_2, \dots, e_n forment une base orthonormale indirecte, alors $[e_1, e_2, \dots, e_n] = -1$.

Soit u_1, u_2, \dots, u_n une famille de n vecteurs de E .

On a évidemment $[u_1, u_2, \dots, u_n] \neq 0$ si et seulement si les u_k forment une base de E .

Si la base u_1, u_2, \dots, u_n est directe (resp. indirecte) alors $[u_1, u_2, \dots, u_n] > 0$ (resp. < 0)

Produit mixte et produit scalaire

Soit u, v deux vecteurs d'un plan euclidien orienté E . Alors $(u | v)^2 + [u, v]^2 = \|u\|^2 \|v\|^2$.

Dans E euclidien orienté de dimension n , on a : $|[u_1, u_2, \dots, u_n]| \leq \|u_1\| \|u_2\| \cdots \|u_n\|$.

Si les u_k sont libres, ce résultat est une égalité si et seulement si les u_k sont orthogonaux deux à deux.

Produit mixte et applications linéaires

Soit u_1, u_2, \dots, u_n une famille de n vecteurs de l'espace euclidien orienté E .

Pour tout endomorphisme f de E , on a : $[f(u_1), f(u_2), \dots, f(u_n)] = (\det f)[u_1, u_2, \dots, u_n]$.

En particulier, si $\det(f) = 1$, on a $[f(u_1), f(u_2), \dots, f(u_n)] = [u_1, u_2, \dots, u_n]$.

On peut donc dire que les applications linéaires de déterminant 1 *conservent le produit mixte*.

Interprétation du produit mixte dans un plan orienté

Soit u, v deux vecteurs d'un plan euclidien orienté E_2 .

Alors $[u, v]$ est l'aire orientée du parallélogramme construit sur les vecteurs u et v .

L'aire orientée du triangle formé sur u et v est $\frac{1}{2}[u, v]$.

Interprétation du produit mixte en dimension 3

On se place dans un espace euclidien orienté E_3 de dimension 3.

On identifie ici les éléments de E_3 avec des points de l'espace.

On se donne un parallélépipède dont les arêtes issues de A sont AB , AC , et AD .

Son volume orienté est $[\vec{AB}, \vec{AC}, \vec{AD}]$. Celui du tétraèdre $ABCD$ est $\frac{1}{6}[\vec{AB}, \vec{AC}, \vec{AD}]$.

On a représenté ci-dessous le parallélépipède.

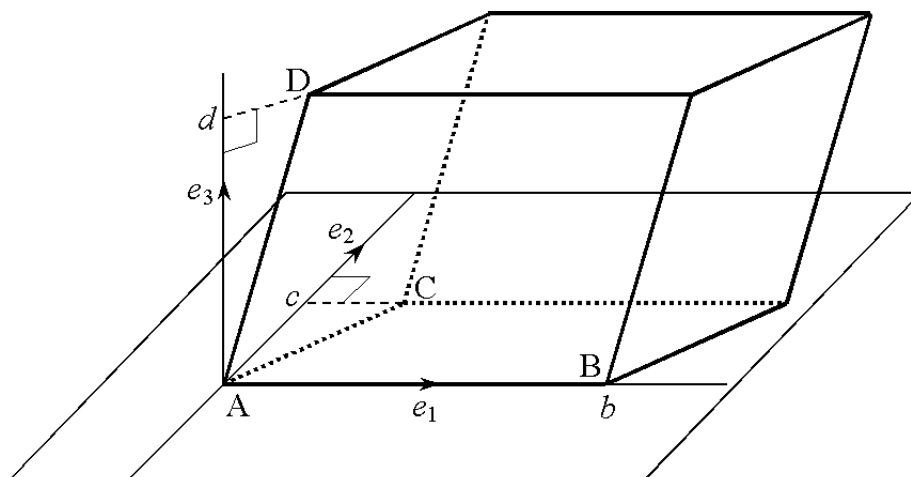
Ici la base $\vec{AB}, \vec{AC}, \vec{AD}$ est directe, donc le produit mixte $[\vec{AB}, \vec{AC}, \vec{AD}]$ est positif.

Le procédé de Schmidt transforme $\vec{AB}, \vec{AC}, \vec{AD}$ en une base orthonormale directe e_1, e_2, e_3 .

On peut alors écrire $\vec{AB} = be_1$, $\vec{AC} = c'e_1 + ce_2$, $\vec{AD} = d''e_1 + d'e_2 + de_3$.

Alors $[\vec{AB}, \vec{AC}, \vec{AD}] = \det_e(\vec{AB}, \vec{AC}, \vec{AD}) = bcd$: c'est bien le volume du parallélépipède.

En effet, bc est l'aire du parallélogramme de base, et d est la hauteur du parallélépipède.



19.3.2 Produit vectoriel

Proposition 19.3.2 (produit vectoriel dans un espace euclidien orienté de dimension 3)

Soit u, v deux vecteurs d'un espace euclidien orienté E de dimension 3.

Il existe un unique vecteur a de E tel que : $\forall w \in E, [u, v, w] = (a | w)$.

Ce vecteur a est appelé produit vectoriel de u par v , et il est noté $u \wedge v$.

On a donc l'égalité, pour tous vecteurs u, v, w de E : $[u, v, w] = ((u \wedge v) | w)$.

Remarques et propriétés

– Alors que le produit mixte existe en dimension n , le produit vectoriel n'existe qu'en dimension 3.

– L'application $(u, v) \mapsto u \wedge v$ est bilinéaire alternée : $\forall (u, v) \in E^2, u \wedge v = -v \wedge u$.

Pour tous vecteurs u, v, w , on peut écrire : $[u, v, w] = ((u \wedge v) | w) = (u | (v \wedge w))$.

– Le vecteur $u \wedge v$ est orthogonal à u et à v .

On a : $u \wedge v = \vec{0} \Leftrightarrow u, v$ sont liés. Si u, v sont libres, alors $u, v, u \wedge v$ forment une base directe.

– Si u, v sont unitaires et orthogonaux, alors $u, v, u \wedge v$ est une base orthonormale directe.

Si i, j, k est orthonormale directe on a :
$$\begin{cases} i \wedge j = k & j \wedge k = i & k \wedge i = j \\ j \wedge i = -k & k \wedge j = -i & i \wedge k = -j \end{cases}$$

– On suppose que E est muni d'une base orthonormale directe i, j, k .

Soit $u = xi + yj + zk$ et $v = x'i + y'j + z'k$.

Alors le produit vectoriel $u \wedge v$ se calcule en écrivant :
$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \wedge \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} yz' - zy' \\ zx' - xz' \\ xy' - yx' \end{pmatrix}.$$

– Soit u, v deux vecteurs de E .

On a l'égalité : $(u | v)^2 + \|u \wedge v\|^2 = \|u\|^2 \|v\|^2$.

En particulier $\|u \wedge v\| \leq \|u\| \|v\|$, avec égalité si et seulement si $(u | v) = 0$.

– L'aire du parallélogramme $ABDC$ est $\|\vec{AB} \wedge \vec{AC}\|$, celle du triangle ABC est $\frac{1}{2} \|\vec{AB} \wedge \vec{AC}\|$.

– Distance d'un point à une droite

Soit \mathcal{D} la droite passant Ω et dirigée par u . La distance de M à \mathcal{D} est $d(M, \mathcal{D}) = \frac{\|\vec{\Omega M} \wedge u\|}{\|u\|}$.

Proposition 19.3.3 (formule du double produit vectoriel)

Pour tous vecteurs u, v, w , on a : $u \wedge (v \wedge w) = (u | w)v - (u | v)w$.

Proposition 19.3.4 (problème de la division vectorielle)

Soit a, b dans E , avec a non nul ; on cherche les vecteurs u de E tels que $a \wedge u = b$.

Si $(a | b) \neq 0$, il n'y a pas de solution, sinon on obtient les $u = u_0 + \lambda a$, avec $u_0 = \frac{1}{\|a\|^2} b \wedge a$.

On illustre ici le problème de la division vectorielle.

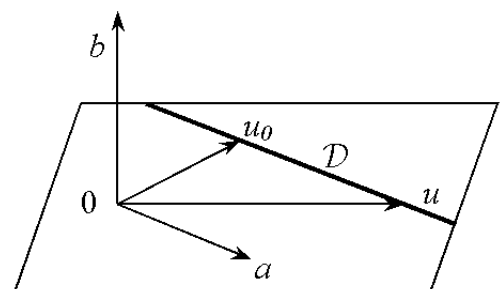
On voit deux vecteurs a et b qui sont orthogonaux.

On cherche les vecteurs u tels que $a \wedge u = b$.

Les solutions u sont forcément orthogonales à b .

Le vecteur u_0 est la seule solution qui soit orthogonale à a .

Les autres solutions forment la droite affine \mathcal{D} passant par u_0 et dirigée par a .



19.4 Projections orthogonales

19.4.1 Supplémentaire orthogonal

Proposition 19.4.1 (supplémentaire orthogonal d'un sous-espace de dimension finie)

Soit E un espace préhilbertien réel. Soit F un sous-espace vectoriel de dimension finie de E .

Alors $E = F \oplus F^\perp$. Le sous-espace F^\perp est appelé le supplémentaire orthogonal de F .

Le cadre du résultat précédent est : F sous-espace de dimension finie de E (qui lui est de dimension quelconque). Évidemment, il n'y a pas d'hypothèse à faire sur F si E est lui-même de dimension finie.

Proposition 19.4.2 (dimension du supplémentaire orthogonal en dimension finie)

Soit E un espace euclidien, donc de dimension finie n .

Soit F un sous-espace vectoriel de E . Alors $\dim(F^\perp) = n - \dim(F)$.

On a l'égalité $F = F^{\perp\perp}$. Ainsi F est lui-même le supplémentaire orthogonal de F^\perp .

Si on est en dimension finie, on pourra donc dire des sous-espaces F et F^\perp qu'ils sont supplémentaires orthogonaux l'un de l'autre.

Supplémentaire orthogonal et bases orthonormées

Soit F un sous-espace vectoriel de l'espace euclidien E .

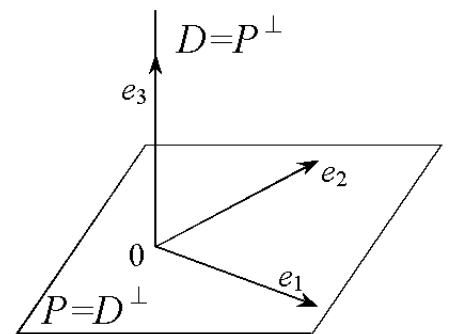
Si e est une base orthonormale de F et si e' est une base orthonormale de F^\perp , alors $e \cup e'$ (obtenue par juxtaposition) est une base orthonormale de E .

Réciproquement, si on complète une base orthonormale e_1, \dots, e_p de F en une base orthonormale $e_1, \dots, e_p, e_{p+1}, \dots, e_n$ de E , alors e_{p+1}, \dots, e_n est une base orthonormale de F^\perp .

Plaçons nous dans un espace euclidien E de dimension 3.

Ici, le plan vectoriel P et la droite vectorielle D sont supplémentaires orthogonaux l'un de l'autre.

Si (e_1, e_2) est une base de P et si e_3 est une base de D , alors la famille (e_1, e_2, e_3) est une base orthonormale de E si et seulement si (e_1, e_2) est une base orthonormale de P et e_3 est unitaire.



19.4.2 Projection orthogonale

Proposition 19.4.3 (projection orthogonale sur un sous-espace de dimension finie)

Soit E un espace préhilbertien réel. Soit F un sous-espace vectoriel de dimension finie de E .

La projection p_F sur F parallèlement à F^\perp est appelée projection orthogonale sur F .

Soit $\varepsilon = \varepsilon_1, \dots, \varepsilon_p$ une base orthonormale de F .

Alors pour tout vecteur u de E , on a l'égalité : $p_F(u) = \sum_{k=1}^p (u | \varepsilon_k) \varepsilon_k$.

Si p est la projection orthogonale sur F , alors $\text{Id} - p$ est la projection orthogonale sur F^\perp .

Projection orthogonale sur une droite ou un hyperplan

Soit a un vecteur non nul de l'espace préhilbertien réel E .

La projection orthogonale sur la droite $\mathbb{R}a$ est définie par $p_a(u) = \frac{(a | u)}{\|a\|^2} a$.

Bien sûr, si a est unitaire : $p_a(u) = (a | u) a$.

On suppose E de dimension finie, et on considère l'hyperplan $H = (\mathbb{R}a)^\perp$.

La projection orthogonale sur $H = (\mathbb{R}a)^\perp$ est définie par $p_H : u \mapsto p_H(u) = u - \frac{(a | u)}{\|a\|^2} a$.

Illustration en dimension 3

On suppose ici $\dim E = 3$.

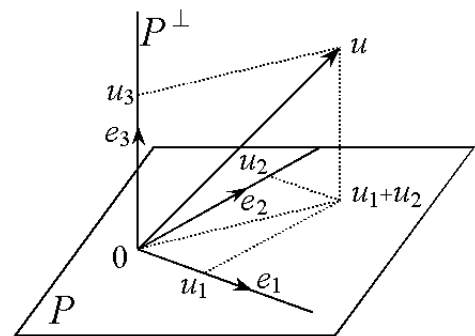
Soit e_1, e_2 une base orthonormale du plan P .

Soit e_3 un vecteur unitaire de la droite P^\perp .

On s'est donné un vecteur u de E .

Chaque $u_k = (u | e_k) e_k$ est la projection orthogonale de u sur la droite engendrée par e_k .

Le vecteur $u_1 + u_2 = (u | e_1) e_1 + (u | e_2) e_2$ est la projection orthogonale de u sur le plan P .



Retour au procédé de Schmidt

Soit E un espace préhilbertien réel. Soit $e = (e_k)_{1 \leq k \leq n}$ une famille libre de E .

Le procédé de Schmidt transforme la famille e en une famille orthonormale $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p)$.

La formation du vecteur ε_k peut être interprétée de la manière suivante :

– Soit $F_{k-1} = \text{Vect}(e_1, \dots, e_{k-1}) = \text{Vect}(\varepsilon_1, \dots, \varepsilon_{k-1})$.

La projection orthogonale v_k de e_k sur F_{k-1} est donnée par $v_k = \sum_{j=1}^{k-1} (e_k | \varepsilon_j) \varepsilon_j$.

– On en déduit $u_k = e_k - v_k$, orthogonal à F_{k-1} et non nul.

Il suffit alors de normer le vecteur u_k pour obtenir le vecteur ε_k .

Proposition 19.4.4 (caractérisations des projections orthogonales)

Soit p une projection vectorielle de l'espace euclidien E .

Les conditions suivantes sont équivalentes :

- la projection p est une projection orthogonale.
- pour tous vecteurs u, v de E , on a l'égalité $(p(u) | v) = (u | p(v))$.
- la matrice de p dans toute base orthonormale est symétrique.
- la matrice de p dans une base orthonormale est symétrique.

Proposition 19.4.5 (une autre caractérisation des projections orthogonales)

Soit p une projection vectorielle de l'espace euclidien E .

L'application p est une projection orthogonale si et seulement si : $\forall u \in E, \|p(u)\| \leq \|u\|$.

19.4.3 Distance à un sous-espace

Définition 19.4.1 (distance d'un vecteur à un sous-espace vectoriel)

Soit E un espace préhilbertien réel. Soit F un sous-espace vectoriel de E .

Pour tout vecteur u de E , on appelle distance de u à F la quantité $d(u, F) = \inf\{\|u - w\|, w \in F\}$.

Proposition 19.4.6 (caractérisation du projeté orthogonal)

Soit E un espace préhilbertien réel. Soit F un sous-espace vectoriel de dimension finie de E .

Soit p la projection orthogonale de E sur F .

Alors $p(u)$ vérifie $\|u - p(u)\| = d(u, F)$, et il est le seul dans F à posséder cette propriété.

Interprétation

Parmi tous les vecteurs de F , le projeté orthogonal de u est celui qui est « plus proche » de u .

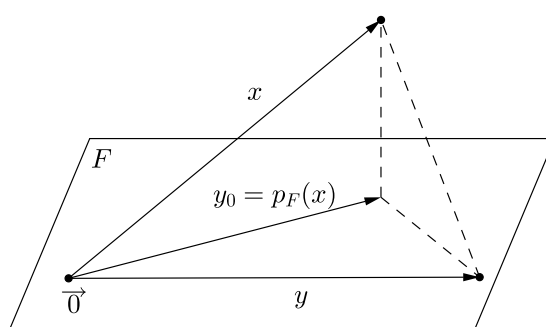
On illustre ici la projection orthogonale $p(u)$ de u sur F .

Pour tout vecteur w de F , on a :

$$\begin{aligned}\|u - w\|^2 &= \|u - p(u) + p(u) - w\|^2 \\ &= \|u - p(u)\|^2 + \|p(u) - w\|^2 \geq \|u - p(u)\|^2\end{aligned}$$

(avec égalité si et seulement si $w = p(u)$).

De plus : $d(u, F)^2 = \|u - p(u)\|^2 = \|u\|^2 - \|p(u)\|^2$.



19.5 Hyperplans affines d'un espace euclidien

19.5.1 Vecteur normal à un hyperplan d'un espace euclidien

Définition 19.5.1 (vecteur normal à un hyperplan affine)

Soit \mathcal{H} un hyperplan affine d'un espace euclidien E , de direction un hyperplan vectoriel H .

On appelle vecteur normal à \mathcal{H} tout vecteur non nul de la droite vectorielle $D = H^\perp$.

Proposition 19.5.1 (caractérisation d'un hyperplan par un point et un vecteur normal)

Soit \mathcal{H} un hyperplan affine d'un espace euclidien E . Soit \vec{n} un vecteur normal à \mathcal{H} .

Soit A un point de \mathcal{H} . Alors on a l'équivalence : $M \in \mathcal{H} \Leftrightarrow (\overrightarrow{AM} \mid \vec{n}) = 0$.

Un hyperplan affine \mathcal{H} est donc déterminé par la donnée d'un point Ω et d'un vecteur normal \vec{n} .

Proposition 19.5.2 (lignes de niveau de $M \mapsto (\overrightarrow{AM} \mid \vec{n})$)

Soit \vec{n} un vecteur non nul d'un espace euclidien E . Soit A un point quelconque de E .

On considère l'application f définie sur E par $f(M) = (\overrightarrow{AM} \mid \vec{n})$.

On appelle lignes de niveau de f les ensembles $\mathcal{H}_\lambda = \{M \in E, f(M) = \lambda\}$.

Les lignes de niveau de f sont les hyperplans affines de vecteur normal \vec{n} .

En particulier \mathcal{H}_0 est l'hyperplan de vecteur normal \vec{n} et qui passe par A .

19.5.2 Équations d'un hyperplan dans une base orthonormale

Proposition 19.5.3 (normale et équations d'un hyperplan affine)

Soit E un espace euclidien muni d'une base orthonormée e .

Soit \mathcal{H} un hyperplan affine de E , de direction H , et soit $\vec{n} = \sum_{k=1}^n \alpha_k e_k$ un vecteur non nul de E .

Les conditions suivantes sont équivalentes :

- le vecteur \vec{n} est normal à l'hyperplan vectoriel H (à l'hyperplan affine \mathcal{H}).
- une équation de H est $(a | u) = 0$, c'est-à-dire $\sum_{k=1}^n a_k x_k = 0$.
- une équation de \mathcal{H} est $(a | u) = \lambda$, avec λ réel, c'est-à-dire $\sum_{k=1}^n a_k x_k = \lambda$.

Exemples dans \mathbb{R}^2

On se place dans l'espace euclidien \mathbb{R}^2 , avec son produit scalaire canonique.

- La normale à la droite vectorielle d'équation $2x + 5y = 0$ est dirigée par $\vec{n} = (2, 5)$.

Soit \mathcal{D} la droite affine orthogonale au vecteur $\vec{n} = (2, 5)$ et passant par $\Omega(4, 1)$.

La droite \mathcal{D} a pour équation $2(x - 4) + 5(y - 1) = 0$, donc $2x + 5y = 13$.

- Soit \mathcal{D} et \mathcal{D}' deux droites affines de \mathbb{R}^2 , de directions respectives D et D' .

On dit que \mathcal{D} et \mathcal{D}' sont *perpendiculaires* si $D^\perp = D'$, c'est-à-dire si la direction de chaque droite affine est le supplémentaire orthogonal de la direction de l'autre.

Cela équivaut aussi à dire que les vecteurs normaux à \mathcal{D} et \mathcal{D}' sont orthogonaux.

Supposons que les équations de \mathcal{D} et \mathcal{D}' soient
$$\begin{cases} ax + by = \lambda \\ a'x + b'y = \mu \end{cases}$$

Alors les droites affines \mathcal{D} et \mathcal{D}' sont perpendiculaires si et seulement si $aa' + bb' = 0$.

Exemples dans \mathbb{R}^3

On se place dans l'espace euclidien \mathbb{R}^3 , avec son produit scalaire canonique.

- La normale au plan vectoriel d'équation $2x + 5y - 3z = 0$ est dirigée par $\vec{n} = (2, 5, -3)$.

Soit \mathcal{P} le plan affine orthogonal au vecteur $\vec{n} = (2, 5, -3)$ et passant par $\Omega(4, -5, -7)$.

Le plan \mathcal{P} a pour équation $(x - 4) + 3(y + 5) - 2(z + 7) = 0$, donc $x + 3y - 2z = 3$.

- Soit \mathcal{P} et \mathcal{P}' deux plans affines de \mathbb{R}^3 , de directions P et P' .

On dit que \mathcal{P} et \mathcal{P}' sont *perpendiculaires* si $P^\perp \subset P'$, c'est-à-dire si $(P')^\perp \subset P$.

Cela signifie que la direction de chacun des deux plans contient un vecteur normal à l'autre.

Cela équivaut aussi à dire que leurs vecteurs normaux sont orthogonaux.

Supposons que les équations de \mathcal{P} et \mathcal{P}' soient
$$\begin{cases} ax + by + cz = \lambda \\ a'x + b'y + c'z = \lambda' \end{cases}$$

Alors \mathcal{P} et \mathcal{P}' sont perpendiculaires si et seulement si $aa' + bb' + cc' = 0$.

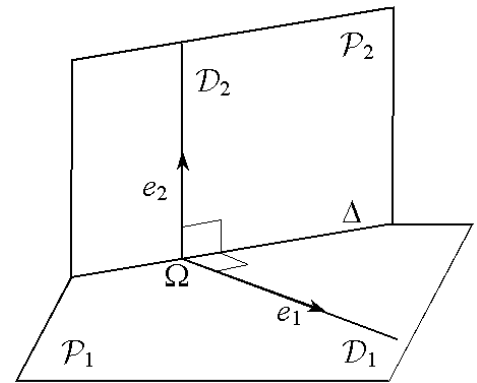
On a représenté ici deux plans affines \mathcal{P}_1 et \mathcal{P}_2 .

Ces plans se coupent suivant une droite affine Δ .

Soit Ω un point de Δ .

Puisque la droite \mathcal{D}_1 passant par Ω et orthogonale à \mathcal{P}_2 est dans \mathcal{P}_1 , Les plans \mathcal{P}_1 et \mathcal{P}_2 sont perpendiculaires.

De même, la droite \mathcal{D}_2 passant par le point Ω et orthogonale au plan affine \mathcal{P}_1 est dans le plan affine \mathcal{P}_2 .



19.5.3 Calcul de la distance à un hyperplan affine

Proposition 19.5.4 (distance d'un point à un hyperplan affine)

Soit \mathcal{H} un hyperplan affine d'un espace euclidien E .

Soit A un point de \mathcal{H} , et soit \vec{n} un vecteur normal unitaire à \mathcal{H} .

Pour tout point M de E , on a : $d(M, \mathcal{H}) = |(\overrightarrow{AM} | \vec{n})|$.

Si le vecteur \vec{n} n'est pas unitaire, alors la distance de M à \mathcal{H} s'écrit : $d(M, \mathcal{H}) = \frac{1}{\|\vec{n}\|} |(\overrightarrow{AM} | \vec{n})|$

Cas particulier : distance à une droite affine dans \mathbb{R}^2

On se place dans l'espace euclidien \mathbb{R}^2 , avec son produit scalaire canonique.

Soit \mathcal{D} une droite affine d'équation $ax + by = h$. Soit $M(x_0, y_0)$ un point quelconque.

Alors la distance du point M à la droite \mathcal{D} est donnée par : $d(M, \mathcal{D}) = \frac{|ax_0 + by_0 - h|}{\sqrt{a^2 + b^2}}$.

Cas particulier : distance à un plan affine dans \mathbb{R}^3

On se place dans l'espace euclidien \mathbb{R}^3 , avec son produit scalaire canonique.

Soit \mathcal{P} un plan affine d'équation $ax + by + cz = h$. Soit $M(x_0, y_0, z_0)$ un point quelconque.

Alors la distance du point M au plan \mathcal{P} est donnée par : $d(M, \mathcal{P}) = \frac{|ax_0 + by_0 + cz_0 - h|}{\sqrt{a^2 + b^2 + c^2}}$.

Distance à une droite affine dans \mathbb{R}^3

On se place dans \mathbb{R}^3 , avec son produit scalaire et son orientation canoniques.

Soit \mathcal{D} une droite affine, passant par un point A et dirigée par un vecteur unitaire u .

Alors la distance du point M à la droite \mathcal{D} est donnée par : $d(M, \mathcal{D}) = \|\overrightarrow{AM} \wedge u\|$.

Si \mathcal{D} est l'intersection de plans perpendiculaires $\mathcal{P}_1, \mathcal{P}_2$, alors $d(M, \mathcal{D})^2 = d(M, \mathcal{P}_1)^2 + d(M, \mathcal{P}_2)^2$.

19.5.4 Orientation d'un hyperplan par un vecteur normal

Proposition 19.5.5 (orientation d'un hyperplan affine par un vecteur normal)

Soit E un espace euclidien orienté de dimension n .

Soit \mathcal{H} un hyperplan affine, de direction H , et soit \vec{n} un vecteur normal à \mathcal{H} .

On oriente la droite $D = H^\perp$ par la donnée du vecteur \vec{n} .

On en déduit une orientation de H (donc de \mathcal{H}) de la manière suivante : une base orthonormale $e = (e_1, \dots, e_{n-1})$ de H est dite directe si $(e_1, \dots, e_{n-1}, e_n)$ est une base orthonormale directe de E .

Si on inverse l'orientation de D (en choisissant $-\vec{n}$ plutôt que \vec{n}), celle de \mathcal{H} s'en trouve inversée.

Illustration en dimension 3

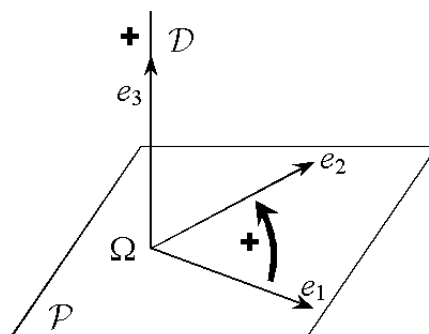
On suppose ici $\dim E = 3$.

La droite \mathcal{D} est la normale en Ω au plan affine \mathcal{P} .

On oriente \mathcal{D} par le choix du vecteur unitaire e_3 .

Il en découle une orientation positive du plan \mathcal{P} .

Le repère orthonormal (Ω, e_1, e_2) de \mathcal{P} est direct dans le plan \mathcal{P} si et seulement si le repère orthonormal (Ω, e_1, e_2, e_3) est direct dans l'espace E .



19.6 Isométries vectorielles

19.6.1 Isométries vectorielles

Définition 19.6.1 (isométries vectorielles, automorphismes orthogonaux)

Soit E un espace euclidien. Soit f une application de E dans lui-même.

On dit que f est une isométrie vectorielle (ou encore : un automorphisme orthogonal) si f est linéaire et si elle « conserve la norme », c'est-à-dire si : $\forall u \in E, \|f(u)\| = \|u\|$.

Les expressions « automorphisme orthogonal » et « isométrie vectorielle » sont donc synonymes.

Toute isométrie vectorielle f de E est effectivement un automorphisme de E .

Soit f un automorphisme orthogonal et soit u un vecteur non nul de E : s'il existe un réel λ tel que $f(u) = \lambda u$, alors nécessairement λ est dans $\{-1, 1\}$.

Proposition 19.6.1 (caractérisation des isométries vectorielles)

Soit f un endomorphisme d'un espace vectoriel euclidien E .

Les propriétés suivantes sont équivalentes :

- l'application f est une isométrie vectorielle (c'est-à-dire elle conserve la norme).
- l'application f conserve le produit scalaire : $\forall (x, y) \in E^2, (f(x) | f(y)) = (x | y)$.
- l'application f transforme toute base orthonormale de E en une base orthonormale de E .
- l'application f transforme une base orthonormale de E en une base orthonormale de E .

Le groupe orthogonal

Les applications Id et $-\text{Id}$ sont des automorphismes orthogonaux de E .

Le composé de deux automorphismes orthogonaux de E est un automorphisme orthogonal de E .

Enfin, si f est un automorphisme orthogonal alors f^{-1} est un automorphisme orthogonal.

On peut résumer ces propriétés de la façon suivante :

Proposition 19.6.2 (groupe orthogonal d'un espace euclidien)

Soit E un espace euclidien. On note $O(E)$ l'ensemble des automorphismes orthogonaux de E .

Alors $O(E)$ est un groupe pour la composition des applications, appelé groupe orthogonal de E .

19.6.2 Symétries vectorielles orthogonales

Définition 19.6.2 (symétries vectorielles orthogonales d'un espace euclidien)

Soit E un espace euclidien, et soit F un sous-espace vectoriel de E .

La symétrie par rapport à F parallèlement à F^\perp est appelée *symétrie orthogonale* par rapport à F .

Si F est un hyperplan, on dit que f est la *réflexion* par rapport à F .

Si F est une droite, on dit que f est le *demi-tour* (ou *retournement*) d'axe F .

Si s est la symétrie vectorielle orthogonale par rapport à F , $-s$ est celle par rapport à F^\perp .

Subtilités de terminologie

Soit E un espace euclidien. Soit s une application linéaire involutive de E , c'est-à-dire telle que $s^2 = \text{Id}$.

On sait que s est la symétrie vectorielle de E par rapport à $F = \text{Inv}(s)$ parallèlement à $G = \text{Opp}(s)$.

Alors s est une symétrie vectorielle *orthogonale* (c'est-à-dire $G = F^\perp$) si et seulement si s est un *automorphisme orthogonal* (c'est-à-dire conserve la norme).

En revanche, une projection orthogonale p n'est *pas* un automorphisme orthogonal, sauf si $p = \text{Id}_E$.

Illustration en dimension 3

On se place ici dans un espace euclidien de dimension 3.

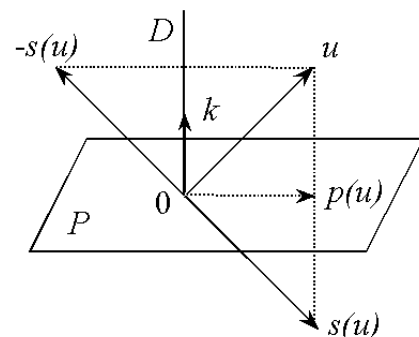
La droite D et le plan P sont orthogonaux.

La droite D est dirigée par le vecteur unitaire k .

On a représenté la réflexion s par rapport au plan vectoriel P et la projection orthogonale p sur P .

On a $p(u) = u - (u | k)k$, et $s(u) = u - 2(u | k)k$.

L'application $-s$ est la symétrie orthogonale par rapport à D .



Proposition 19.6.3 (matrice d'une symétrie vectorielle orthogonale dans une base orthonormée)

Soit E un espace euclidien, et soit s une symétrie vectorielle orthogonale de E .

Alors on a l'égalité $(s(u) | v) = (u | s(v))$ pour tous vecteurs u et v .

Il en résulte que la matrice de s dans toute base orthonormale est symétrique.

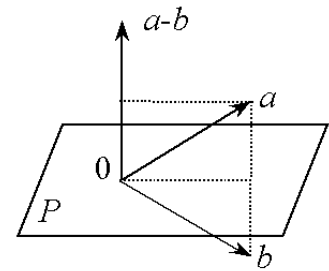
Réflexion échangeant deux vecteurs de même norme

Soit E un espace vectoriel euclidien.

Soit a, b deux vecteurs distincts non nuls, tels que $\|a\| = \|b\|$.

Alors il existe une unique réflexion vectorielle qui échange a et b .

Cette application est la réflexion par rapport à l'hyperplan vectoriel P orthogonal au vecteur $a - b$.



19.7 Matrices orthogonales

19.7.1 Matrices orthogonales

Remarque préliminaire :

Soit M dans $\mathcal{M}_n(\mathbb{R})$, de colonnes C_1, \dots, C_n . Alors le terme général de $A = M^\top M$ est $a_{ij} = C_i^\top C_j$.

Définition 19.7.1 (matrices orthogonales)

Soit M une matrice de $\mathcal{M}_n(\mathbb{R})$. Les conditions suivantes sont équivalentes :

- la matrice M vérifie $M^\top M = I_n$.
- la matrice M est inversible et $M^{-1} = M^\top$.
- les vecteurs-colonne de M forment une famille orthonormale.

Si ces conditions sont réalisées, on dit que M est une *matrice orthogonale*.

Remarques et exemples

Si M est une matrice orthogonale, il en est de même de M^\top (car $M^\top = M^{-1}$).

Une matrice M est donc orthogonale si et seulement si ses lignes forment une famille orthonormale.

Les matrices $R(\theta) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$ et $S(\theta) = \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix}$ sont orthogonales.

On verra plus loin que ce sont les seules les matrices orthogonales d'ordre 2.

Les matrices $M = \frac{1}{3} \begin{pmatrix} 2 & 2 & 1 \\ 1 & -2 & 2 \\ 2 & -1 & -2 \end{pmatrix}$ et $M = \begin{pmatrix} \cos \theta \cos \varphi & -\sin \theta & \cos \theta \sin \varphi \\ \sin \theta \cos \varphi & \cos \theta & \sin \theta \sin \varphi \\ \sin \varphi & 0 & -\cos \varphi \end{pmatrix}$ sont orthogonales.

Proposition 19.7.1 (le groupe orthogonal $O(n)$)

On note $O(n)$ ou $O_n(\mathbb{R})$ l'ensemble des matrices orthogonales d'ordre n .

C'est un groupe pour le produit des matrices (donc un sous-groupe de $GL(n, \mathbb{R})$).

On l'appelle le groupe orthogonal d'indice n .

Proposition 19.7.2 (lien entre base orthonormale, isométrie et matrice orthogonale)

Soit M la matrice d'un endomorphisme f dans une base orthonormale de l'espace euclidien E .

Les conditions suivantes sont équivalentes :

- l'application f est un automorphisme orthogonal de E (c'est-à-dire un élément du groupe $O(E)$).
- la matrice M est une matrice orthogonale (c'est-à-dire un élément du groupe $O(n)$).

On peut interpréter la proposition précédente en disant que les matrices orthogonales sont les matrices des automorphismes orthogonaux dans les bases orthonormales.

Si on se place dans \mathbb{R}^n (avec son produit scalaire canonique), une matrice M de $\mathcal{M}_n(\mathbb{R})$ est orthogonale si et seulement si l'endomorphisme f canoniquement associé à M est une isométrie vectorielle.

Proposition 19.7.3 (matrices de passage entre bases orthonormales)

Soit E un espace euclidien, muni d'une base orthonormale $e = (e_i)_{1 \leq i \leq n}$.

Soit $\varepsilon = (\varepsilon_j)_{1 \leq j \leq n}$ une famille de n vecteurs, et soit M la matrice de la famille ε dans la base e .

Alors la famille ε est une base orthonormale de E si et seulement si la matrice M est orthogonale.

On peut donc dire que les matrices orthogonales sont les matrices de passage entre bases orthonormales.

19.7.2 Matrices orthogonales positives ou négatives

Proposition 19.7.4 (déterminant d'une matrice orthogonale)

Si M est une matrice orthogonale, alors $\det M$ est égal à 1 ou à -1 .

Attention la **réciproque** est **fausse**!! Considérer par exemple la matrice $M = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

Définition 19.7.2 (matrices orthogonales positives ou négatives)

Soit M une matrice orthogonale d'ordre n .

Si $\det(M) = 1$, on dit que M est une matrice orthogonale *positive*.

Si $\det(M) = -1$, on dit que M est une matrice orthogonale *négative*.

Remarques

Si on échange deux colonnes (ou deux lignes) d'une matrice orthogonale positive, on obtient une matrice orthogonale négative (et réciproquement).

C'est la même chose si on remplace une colonne (ou une ligne) par son opposée.

Il existe des matrices orthogonales positives (considérer par exemple I_n).

Il existe des matrices orthogonales négatives (changer un coefficient diagonal de I_n en -1).

Pour M dans $\mathcal{M}_n(\mathbb{R})$, on sait que $\det(-M) = (-1)^n \det(M)$: il en résulte que si M est orthogonale, les matrices M et $-M$ ont la même « orientation » si n est pair, et sont d'orientation contraire sinon.

Comatrice d'une matrice orthogonale

Si M appartient à $O(n)$ alors $\text{Com}(M) = \varepsilon M$, avec $\varepsilon = \det(M) = \pm 1$.

En comparant un coefficient non nul de M avec son cofacteur, on peut donc déterminer ε .

Considérons par exemple la matrice $M = \frac{1}{3} \begin{pmatrix} 2 & 2 & 1 \\ 1 & -2 & 2 \\ 2 & -1 & -2 \end{pmatrix}$, qui est élément de $O(3)$.

Le coefficient $m_{11} = \frac{2}{3}$ et son cofacteur égaux. On est donc certain que $\det(M) = 1$.

Proposition 19.7.5 (le groupe spécial orthogonal $SO(n)$)

On note $SO(n)$, ou $SO_n(\mathbb{R})$, l'ensemble des matrices orthogonales positives d'ordre n .

L'ensemble $SO(n)$ est un sous-groupe de $O(n)$, appelé groupe spécial orthogonal d'indice n .

L'ensemble des matrices orthogonales *négatives* (le complémentaire de $SO(n)$ dans $O(n)$) n'est pas un groupe : non seulement il ne contient pas le neutre I_n , mais il n'est pas stable : en effet si M et N sont orthogonales négatives, alors MN est orthogonale positive.

En revanche, l'inverse d'une matrice orthogonale négative est encore orthogonale négative.

La matrice I_n est orthogonale positive, alors que la matrice $-I_n$ n'est dans $SO(n)$ que si n est pair.

19.7.3 Isométries positives, négatives**Définition 19.7.3** (isométries positives ou négatives)

Si f est une isométrie vectorielle d'un espace euclidien E , alors $\det(f)$ est égal à 1 ou à -1 .

Si $\det(f) = 1$, on dit que f est une isométrie *positive*.

Si $\det(f) = -1$, on dit que f est une isométrie *négative*.

Proposition 19.7.6 (le groupe spécial orthogonal d'un espace euclidien)

Soit E un espace euclidien. On note $SO(E)$ l'ensemble des isométries positives de E .

L'ensemble $SO(E)$ est un sous-groupe de $O(E)$, appelé groupe spécial orthogonal de E .

Remarques

L'application Id est dans $SO(E)$. Mais $-\text{Id}$ est dans $SO(E)$ si et seulement si $\dim(E)$ est paire.

La réflexion s par rapport à un hyperplan vectoriel est toujours un automorphisme orthogonal négatif.

Un demi-tour de E est une isométrie positive si $\dim(E)$ est impaire, et négative sinon.

Proposition 19.7.7 (isométries positives ou négatives dans un espace orienté)

Soit f une isométrie vectorielle d'un espace euclidien orienté E .

Soit e une base orthonormée de E . On sait que f transforme e en une base orthonormée ε .

Si f est une isométrie positive, alors les bases e et ε ont la même orientation.

Si f est une isométrie négative, alors les bases e et ε sont d'orientation contraire.

Ainsi les isométries positives *conservent l'orientation*, alors que les isométries négatives l'inversent.

Proposition 19.7.8 (conservation du produit mixte par une isométrie positive)

Soit f une isométrie vectorielle positive d'un espace euclidien orienté E .

Alors pour tous vecteurs u_1, u_2, \dots, u_n de E , on a : $[f(u_1), f(u_2), \dots, f(u_n)] = [u_1, u_2, \dots, u_n]$.

On peut donc dire que les isométries vectorielles positives *conservent le produit mixte* (donc l'aire orientée en dimension 2, et le volume orienté en dimension 3).

19.8 Isométries en dimension 2

19.8.1 Matrices orthogonales de taille 2

Proposition 19.8.1 (matrices orthogonales d'ordre 2)

Les matrices orthogonales positives d'ordre 2 sont les matrices $R(\theta) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$, avec $\theta \in \mathbb{R}$.

Les matrices orthogonales négatives d'ordre 2 sont les matrices $S(\theta) = \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix}$, avec $\theta \in \mathbb{R}$.

Propriétés des matrices orthogonales positives d'ordre 2

Pour tous réels θ et φ , on a : $R(\theta)R(\varphi) = R(\varphi)R(\theta) = R(\theta + \varphi)$.

Il en découle que le groupe $SO(2)$ est commutatif (c'est faux pour $SO(n)$ si $n \geq 3$).

On a $R(0) = I_2$ et, pour tout réel θ : $R(\theta)^{-1} = R(-\theta)$.

Pour tous réels θ et φ , on a : $R(\theta) = R(\varphi) \Leftrightarrow \theta \equiv \varphi [2\pi]$.

Propriétés des matrices orthogonales négatives d'ordre 2

Pour tous réels θ et φ , on a : $S(\theta)S(\varphi) = R(\theta - \varphi)$.

Pour tout réel θ , on a $S(\theta)^{-1} = S(\theta)^\top = S(\theta)$.

Toutes les matrices orthogonales négatives d'ordre 2 sont donc des matrices de symétrie.

Les automorphismes orthogonaux négatifs d'un plan sont des symétries orthogonales (à suivre...).

19.8.2 Angle de rotations et de vecteurs du plan

Proposition 19.8.2 (angle d'une rotation dans le plan euclidien orienté)

Soit E_2 un plan euclidien orienté, et soit r une isométrie vectorielle positive de E_2 .

Il existe un réel θ (défini modulo 2π) vérifiant la propriété suivante :

La matrice de r dans toute base orthonormale directe est égale à $R(\theta) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$.

On dit alors que r est la rotation d'angle θ , et on note $r = r(\theta)$.

On a en particulier $r(0) = \text{Id}$ et $r(\pi) = -\text{Id}$ (et ce sont les seules rotations involutives).

Dans E_2 , les expressions « isométries vectorielles positives » et « rotations vectorielles » sont synonymes.

Remarques et propriétés

– La rotation inverse de $r(\theta)$ est $r(-\theta)$.

Pour tous réels θ et φ , on a : $r(\theta)r(\varphi) = r(\varphi)r(\theta) = r(\theta + \varphi)$. Le groupe $SO(E_2)$ est donc commutatif.

– La matrice de la rotation $r(\theta)$ dans toute base orthonormale indirecte est $R(-\theta)$.

Si on inverse l'orientation de E_2 , la mesure d'une rotation est donc changée en son opposée.

– Si $r = r(\theta)$, avec $\theta \neq 0 [2\pi]$, alors le seul vecteur invariant de r est $\vec{0}$.

– Si $r = r\left(\frac{\pi}{2}\right)$, et si e_1, e_2 forment une base orthonormale directe, alors $r(e_1) = e_2$ et $r(e_2) = -e_1$.

Proposition 19.8.3 (mesure de l'angle orienté de deux vecteurs unitaires)

Soit E_2 un plan vectoriel euclidien orienté, et soit u, v deux vecteurs unitaires de E_2 .

Il existe une et une seule rotation $r = r(\theta)$ telle que $r(u) = v$.

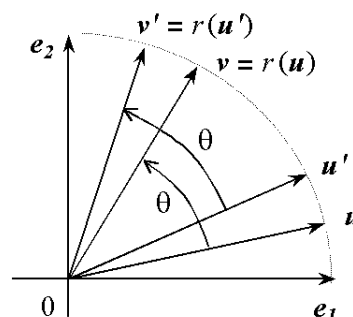
On note $\widehat{(u, v)} = \theta [2\pi]$, et on dit que θ est une mesure (modulo 2π) de l'angle orienté $\widehat{(u, v)}$.

La figure ci-dessous illustre cette propriété. La base e_1, e_2 , orthonormale directe, n'est là que pour visualiser l'orientation positive choisie dans le plan. Tous les vecteurs considérés ici sont unitaires.

Il existe bien une *unique* rotation vectorielle $r = r(\theta)$ telle que $v = r(u)$. On commet souvent l'erreur de croire qu'il y a deux rotations transformant u en v (l'une « tournant » dans un sens, la deuxième tournant dans l'autre) ou même une infinité (selon le « nombre de tours effectués »).

La rotation r n'est qu'une application : seul compte où se trouve l'image v d'un vecteur u , et pas la manière dont on « passe » de u à v . L'erreur évoquée vient de la confusion entre la rotation r et les différentes mesures de son angle.

On se rend compte de l'unicité de la rotation r transformant u en v en se donnant un autre vecteur unitaire u' . La condition $v = r(u)$ détermine en effet le vecteur $v' = r(u')$ de manière unique.

**Définition 19.8.1** (mesure de l'angle orienté de deux vecteurs non nuls)

Dans un plan euclidien orienté E_2 , soit u, v deux vecteurs non nuls.

On appelle mesure de l'angle orienté $\widehat{(u, v)}$ la mesure θ de l'angle orienté $\left(\frac{u}{\|u\|}, \frac{v}{\|v\|}\right)$.

On note alors $\widehat{(u, v)} = \theta [2\pi]$.

Proposition 19.8.4 (calcul d'une mesure de l'angle orienté de deux vecteurs non nuls)

Soit u, v deux vecteurs non nuls dans un plan euclidien orienté E_2 .

Une mesure θ de l'angle $\widehat{(u, v)}$ est donnée par : $\cos \theta = \frac{(u | v)}{\|u\| \|v\|}$, et $\sin \theta = \frac{[u, v]}{\|u\| \|v\|}$.

Rappel : $[u, v]$ est le produit mixte de u et v (leur déterminant dans toute base orthonormale directe).

19.8.3 Classification des isométries d'un plan euclidien orienté**Proposition 19.8.5** (automorphismes orthogonaux négatifs de E_2)

On se place dans un plan euclidien orienté E_2 , rapporté une base orthonormale directe $e = (e_1, e_2)$.

Soit s une isométrie négative de E_2 . Soit $S(\theta) = \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix}$ la matrice de s dans la base e .

Alors s est la réflexion par rapport à la droite dirigée par le vecteur $u = \cos\left(\frac{\theta}{2}\right)e_1 + \sin\left(\frac{\theta}{2}\right)e_2$.

Classification des isométries d'un plan euclidien orienté E_2

Les isométries positives de E_2 sont les rotations vectorielles.

Les isométries négatives de E_2 sont les réflexions par rapport à des droites vectorielles.

Le fait que E_2 soit orienté n'intervient pas dans cette classification, mais dans la possibilité de mesurer l'angle d'une rotation et l'angle polaire de l'axe d'une réflexion.

Chapitre 20

Intégration

Sommaire

20.1	Continuité uniforme	449
20.2	Continuité par morceaux	450
20.2.1	Fonctions en escaliers	450
20.2.2	Fonctions continues par morceaux	451
20.2.3	Intégrale des fonctions en escaliers	452
20.3	Intégrale sur un segment	453
20.3.1	Définition de l'intégrale sur un segment	453
20.3.2	Propriétés de l'intégrale	454
20.3.3	Extension de la définition et nouvelle notation	455
20.3.4	Intégrale d'une fonction continue de signe constant	456
20.3.5	Sommes de Riemann de f sur $[a, b]$	456
20.3.6	Méthode des trapèzes	457
20.4	Intégrale et primitives	457
20.4.1	Dérivée de la fonction $x \mapsto \int_0^x f(t) dt$	457
20.4.2	Méthodes d'intégration	458
20.4.3	Calcul de primitives	458
20.5	Formules de Taylor	459

20.1 Continuité uniforme

Dans tout le chapitre I est un intervalle de \mathbb{R} non vide et non réduit à un point.

La lettre \mathbb{K} désigne indifféremment le corps \mathbb{R} des nombres réels ou le corps \mathbb{C} des nombres complexes.

On rappelle que $\mathcal{F}(I, \mathbb{K})$ est l'ensemble des fonctions définies sur I et à valeurs dans \mathbb{K} .

Définition 20.1.1 (uniforme continuité)

Soit f un élément de $\mathcal{F}(I, \mathbb{K})$. On dit que f est *uniformément continue* sur I si :

$$\forall \varepsilon > 0, \exists \delta > 0, \forall (x, y) \in I \times I, (|x - y| \leq \delta \Rightarrow |f(x) - f(y)| \leq \varepsilon).$$

La définition précédente doit se lire de la manière suivante :

« Pour tout réel strictement positif ε (sous-entendu : aussi petit soit-il), il existe un réel strictement positif δ (dépendant a priori de ε) tel que, pour tous éléments x et y de I distants de moins de δ , alors leurs images $f(x)$ et $f(y)$ sont distantes de moins de ε ».

Pour montrer qu'une fonction n'est pas uniformément continue

Soit $f : I \rightarrow \mathbb{K}$ une fonction dont on souhaite prouver qu'elle n'est pas uniformément continue sur I .

Si on prend la négation de la définition, on doit prouver l'existence de $\varepsilon > 0$ tel que, pour tout $\delta > 0$, on peut trouver x et y dans I tels que $|x - y| < \delta$ mais cependant tels que $|f(x) - f(y)| \geq \varepsilon$.

Il revient au même (et il est plus simple) de trouver deux suites (x_n) et (y_n) de l'intervalle I telles que $\lim_{n \rightarrow \infty} (y_n - x_n) = 0$, mais telles qu'on n'ait pas $\lim_{n \rightarrow \infty} (f(y_n) - f(x_n)) = 0$

Continuité et continuité uniforme

Rappelons la définition de la continuité de f en un point a de l'intervalle I :

$$\forall \varepsilon > 0, \exists \delta > 0, \forall x \in I, (|x - a| \leq \delta \Rightarrow |f(x) - f(a)| \leq \varepsilon)$$

Dans cette définition, le réel δ dépend a priori de ε et du point a .

La continuité uniforme exprime l'existence d'un réel δ ne dépendant que de ε , et donc pas du point a .

Si f est uniformément continue sur l'intervalle I , alors elle est continue sur I .

Mais la réciproque est fautive, comme le montrent les deux exemples suivants :

– la fonction f définie sur $]0, 1]$ par $f(x) = \frac{1}{x}$: considérer $x_n = \frac{1}{n}$ et $y_n = \frac{1}{n+1}$.

– la fonction f définie sur \mathbb{R} par $f(x) = \cos(x^2)$: considérer $x_n = \sqrt{2n\pi}$ et $y_n = \sqrt{(2n+1)\pi}$.

L'uniforme continuité est une notion globale, et non locale

Quand on dit qu'une fonction est continue sur un intervalle I , c'est pour exprimer qu'elle est continue en chacun des points de I . En ce sens, on dit que la continuité est une *notion locale*.

En revanche, l'uniforme continuité d'une fonction f n'a de sens que relativement à un intervalle I . Cela ne signifie donc rien d'énoncer que f est uniformément continue en un point ! On exprime cela en disant que l'uniforme continuité est une *notion globale*.

Le théorème de Heine

Le théorème suivant dit que « la continuité sur un segment implique la continuité uniforme ».

Proposition 20.1.1 (théorème de Heine)

Soit f une fonction continue sur un segment $[a, b]$ de \mathbb{R} , à valeurs dans \mathbb{K} .

Alors la fonction f est uniformément continue sur $[a, b]$.

20.2 Continuité par morceaux

Dans ce chapitre, $[a, b]$ désigne un segment de \mathbb{R} , avec $a < b$.

20.2.1 Fonctions en escaliers

Définition 20.2.1 (subdivisions d'un segment)

On appelle *subdivision* de $[a, b]$ toute suite finie $(x_0 = a < x_1 < \dots < x_{n-1} < x_n = b)$.

L'ensemble $\{a = x_0, \dots, x_k, \dots, x_n = b\}$ est appelé le *support* de la subdivision.

La quantité $h = \max(x_{k+1} - x_k)$ est appelée le *pas* de la subdivision.

Finesse d'une subdivision

Soit σ et σ' deux subdivisions de $[a, b]$.

On dit que σ est *plus fine* que σ' si le support de σ contient celui de σ' .

La subdivision notée $\sigma \cup \sigma'$ et dont le support est la réunion de ceux de σ et de σ' est plus fine que chacune des subdivisions σ et σ' . Réciproquement si une subdivision de $[a, b]$ est plus fine que σ et σ' , alors elle est plus fine que $\sigma \cup \sigma'$.

Définition 20.2.2 (fonctions en escaliers sur un segment)

Soit φ une fonction définie sur $[a, b]$, à valeurs dans \mathbb{K} .

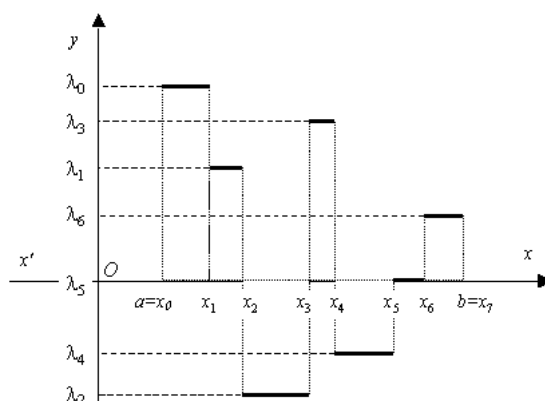
On dit que φ est *en escaliers* sur le segment $[a, b]$ s'il existe une subdivision $\sigma = (x_k)_{0 \leq k \leq n}$ de $[a, b]$ et s'il existe n scalaires $\lambda_0, \lambda_1, \dots, \lambda_{n-1}$ tels que : $\forall k = 0, \dots, n-1, \forall t \in]x_k, x_{k+1}[, \varphi(t) = \lambda_k$.

On dit alors que la subdivision σ est *adaptée* (ou encore *subordonnée*) à la fonction φ .

On note $\mathcal{E}([a, b], \mathbb{K})$ l'ensemble des fonctions en escaliers sur $[a, b]$ et à valeurs dans \mathbb{K} .

La figure ci-contre représente une fonction en escaliers φ sur le segment $[a, b]$, à valeurs réelles.

On n'a pas représenté les valeurs de φ aux points x_k , car ces valeurs sont sans importance.



Remarques et propriétés

- Si σ est une subdivision adaptée à φ , toute subdivision plus fine que σ est adaptée à φ .
- Les fonctions constantes sur $[a, b]$ sont des cas particuliers de fonctions en escaliers.
- Si φ et ψ sont en escaliers sur $[a, b]$, alors : $\forall (\alpha, \beta) \in \mathbb{K}^2$, $\alpha\varphi + \beta\psi$ est en escaliers sur $[a, b]$.
Plus généralement, toute combinaison linéaire de fonctions en escaliers est encore en escaliers.
De même le produit $\varphi\psi$ est en escaliers sur $[a, b]$.

Définition 20.2.3 (fonctions en escaliers sur un intervalle quelconque)

Soit I un intervalle de \mathbb{R} d'intérieur non vide. Soit φ une fonction de I dans \mathbb{K} .

On dit que φ est en escaliers sur I si φ est en escaliers sur tout segment de I .

Par exemple, l'application « partie entière » $x \mapsto [x]$ est en escaliers sur \mathbb{R} .

20.2.2 Fonctions continues par morceaux**Définition 20.2.4** (fonction continue par morceaux sur un segment)

Soit f une fonction définie sur le segment $[a, b]$, à valeurs dans \mathbb{K} .

On dit que f est *continue par morceaux* sur $[a, b]$ s'il existe une subdivision $\sigma = (x_k)_{0 \leq k \leq n}$ de $[a, b]$ (dite adaptée à f , ou encore subordonnée à f) telle que, pour tout k de $\{0, \dots, n-1\}$:

- la restriction f_k de f à chaque intervalle ouvert $]x_k, x_{k+1}[$ est continue.
- cette restriction est prolongeable par continuité aux points x_k et x_{k+1} .

On note $\mathcal{C}_m([a, b], \mathbb{K})$ l'ensemble des fonctions continues par morceaux sur $[a, b]$ et à valeurs dans \mathbb{K} .

Remarques et propriétés

- Si σ est une subdivision adaptée à f , toute subdivision plus fine que σ est encore adaptée à f .
- Dire que f est dans $\mathcal{C}_m([a, b], \mathbb{K})$, c'est dire qu'elle n'a qu'un nombre fini de discontinuités, toutes de *première espèce* : en chaque discontinuité, il y a une limite à gauche et une limite à droite finies.
- Toute fonction continue par morceaux sur $[a, b]$ est bornée sur $[a, b]$.

Définition 20.2.5 (fonctions continues par morceaux sur un intervalle quelconque)

Soit I un intervalle de \mathbb{R} d'intérieur non vide. Soit f une fonction de I dans \mathbb{K} .

On dit que f est continue par morceaux sur I si elle l'est sur tout segment de I .

On note $\mathcal{C}_m(I, \mathbb{K})$ l'ensemble des fonctions continues par morceaux sur I et à valeurs dans \mathbb{K} .

Toute fonction continue sur I est continue par morceaux sur I (il y a un seul « morceau »!).

Toute fonction en escaliers sur I est continue par morceaux sur I .

Opérations sur les fonctions continues par morceaux

Soit f et g dans $\mathcal{C}_m(I, \mathbb{K})$. Pour tous α, β de \mathbb{K} , $\alpha f + \beta g$ est dans $\mathcal{C}_m(I, \mathbb{K})$.

De même, la fonction fg est continue par morceaux sur I .

Cas des fonctions à valeurs complexes

Soit $f : I \rightarrow \mathbb{C}$ une fonction à valeurs complexes. Soit $u = \operatorname{Re}(f)$ et $v = \operatorname{Im}(f)$.

La fonction f est dans $\mathcal{C}_m(I, \mathbb{C})$ si et seulement si u et v sont dans $\mathcal{C}_m(I, \mathbb{R})$.

20.2.3 Intégrale des fonctions en escaliers

Définition 20.2.6

Soit $\varphi : [a, b] \rightarrow \mathbb{K}$ une fonction en escaliers et $\sigma = (x_k)_{0 \leq k \leq n}$ une subdivision adaptée.

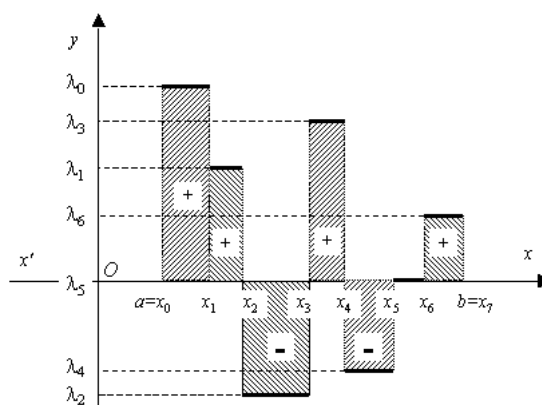
On suppose que : $\forall k \in \{0, \dots, n-1\}, \forall t \in]x_k, x_{k+1}[, \varphi(t) = \lambda_k$.

Le réel $\sum_{k=0}^{n-1} (x_{k+1} - x_k) \lambda_k$ est appelé *intégrale* de φ et est noté $\int_{[a,b]} \varphi$.

Interprétation graphique (cas réel) :

Comme le montre la figure ci-contre, l'intégrale de φ est égale à la somme des « aires algébriques » (comptées positivement ou négativement selon le signe des réels λ_k) des rectangles définis par le graphe de φ .

L'intégrale de φ ne dépend pas de la subdivision σ choisie (dans la mesure bien sûr où elle est adaptée à φ).



Linéarité de l'intégrale

Soit φ, ψ deux fonctions en escaliers sur $[a, b]$, et soit α, β deux scalaires.

Alors on a l'égalité $\int_{[a,b]} (\alpha\varphi + \beta\psi) = \alpha \int_{[a,b]} \varphi + \beta \int_{[a,b]} \psi$.

La fonction qui à φ associe $\int_{[a,b]} \varphi$ est donc une forme *linéaire* sur $\mathcal{E}([a, b], \mathbb{K})$.

Positivité et croissance

On rappelle que $[a, b]$ désigne un segment de \mathbb{R} , avec $a < b$.

Soit φ et ψ dans $\mathcal{E}([a, b], \mathbb{R})$. Si $\varphi \geq 0$ alors $\int_{[a,b]} \varphi \geq 0$. Si $\varphi \leq \psi$ alors : $\int_{[a,b]} \varphi \leq \int_{[a,b]} \psi$.

Si $\varphi \in \mathcal{E}([a, b], \mathbb{K})$, alors $|\varphi| : t \mapsto |\varphi(t)| \in \mathcal{E}([a, b], \mathbb{R})$ et $\left| \int_{[a,b]} \varphi \right| \leq \int_{[a,b]} |\varphi|$.

Si φ est en escaliers de $[a, b]$ dans \mathbb{R} , alors $\left| \int_{[a,b]} \varphi \right| \leq (b-a) \sup_{t \in [a,b]} |\varphi(t)|$.

Remarques et propriétés

Si φ est constante et égale à λ sur $[a, b]$, alors $\int_{[a,b]} \varphi = (b-a)\lambda$.

Soit φ une fonction nulle sur $[a, b]$, sauf peut-être en un nombre fini de points.

Alors φ est en escaliers sur $[a, b]$ et $\int_{[a,b]} \varphi = 0$.

Soit φ en escaliers sur $[a, b]$, et soit ψ ne différant de φ qu'en un nombre fini de points.

Alors ψ est en escaliers sur $[a, b]$ et $\int_{[a,b]} \psi = \int_{[a,b]} \varphi$.

Soit φ un élément de $\mathcal{E}([a, b], \mathbb{K})$, et soit c un élément de $]a, b[$.

Alors les restrictions de φ à $[a, c]$ et $[c, b]$ sont en escaliers et : $\int_{[a,b]} \varphi = \int_{[a,c]} \varphi + \int_{[c,b]} \varphi$.

20.3 Intégrale sur un segment

20.3.1 Définition de l'intégrale sur un segment

Proposition 20.3.1 (approximation par des fonctions en escaliers)

Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction continue par morceaux, à valeurs réelles.

Pour tout $\varepsilon > 0$, il existe deux fonctions en escaliers φ et ψ sur $[a, b]$ telles que :

- pour tout x de $[a, b]$, $\varphi(x) \leq f(x) \leq \psi(x)$.
- pour tout x de $[a, b]$, $0 \leq \psi(x) - \varphi(x) \leq \varepsilon$.

Proposition 20.3.2 (intégrale des fonctions continues par morceaux à valeurs réelles)

Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction continue par morceaux à valeurs réelles.

On considère les deux quantités :

d'une part $I_-(f) = \sup_{\varphi \leq f} \int_{[a,b]} \varphi$, borne supérieure des intégrales des φ en escaliers telles que $\varphi \leq f$.

d'autre part $I^+(f) = \inf_{\psi \geq f} \int_{[a,b]} \psi$, borne inférieure des intégrales des ψ en escaliers telles que $\psi \geq f$.

Alors $I_-(f)$ et $I^+(f)$ sont des réels égaux.

Leur valeur commune est appelée intégrale de f sur $[a, b]$, et elle est notée $\int_{[a,b]} f$.

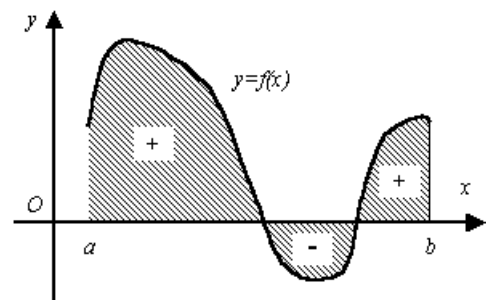
Si f est en escaliers, donc continue par morceaux, les deux significations de $\int_{[a,b]} f$ coïncident.

Interprétation en terme d'aire

Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction continue par morceaux.

L'intégrale de f sur $[a, b]$ représente l'aire algébrique du domaine situé entre la courbe $y = f(x)$ et l'axe Ox , cette « aire » étant comptée positivement sur les intervalles où $f \geq 0$ et négativement sur les intervalles où $f \leq 0$.

Numériquement, le résultat est exprimé en *unités d'aire* (ua).



Extension aux fonctions à valeurs complexes

Soit $f : I \rightarrow \mathbb{C}$ une fonction à valeurs complexes, continue par morceaux.

Soit $u = \operatorname{Re}(f)$ et $v = \operatorname{Im}(f)$. On pose $\int_{[a,b]} f = \int_{[a,b]} u + i \int_{[a,b]} v$.

Invariance de l'intégrale par translation

Soit $f : [a, b] \rightarrow \mathbb{K}$, continue par morceaux. Soit α un nombre réel.

On définit la fonction g de $J = [a + \alpha, b + \alpha]$ dans \mathbb{K} par $g(t) = f(t - \alpha)$.

Alors g est continue par morceaux sur J et $\int_J g = \int_I f$.

20.3.2 Propriétés de l'intégrale

Pour les fonctions à valeurs réelles, les propriétés de l'intégrale découlent des propriétés analogues sur $\mathcal{E}([a, b], \mathbb{R})$, par passage à borne supérieure. Pour les fonctions à valeurs complexes, elles découlent immédiatement de la définition $\int_{[a, b]} f = \int_{[a, b]} \operatorname{Re}(f) + i \int_{[a, b]} \operatorname{Im}(f)$.

Proposition 20.3.3 (linéarité de l'intégrale)

Soit f et g dans $\mathcal{C}_m([a, b], \mathbb{K})$, et soit α, β dans \mathbb{K} . Alors $\int_{[a, b]} (\alpha f + \beta g) = \alpha \int_{[a, b]} f + \beta \int_{[a, b]} g$.

Ainsi la fonction qui à f associe $\int_{[a, b]} f$ est une forme linéaire sur $\mathcal{C}_m([a, b], \mathbb{K})$.

Proposition 20.3.4 (positivité et croissance, pour les fonctions à valeurs réelles)

Soit f et g deux fonctions continues par morceaux sur $[a, b]$, à valeurs réelles :

- si f est positive ou nulle sur $[a, b]$, alors $\int_{[a, b]} f \geq 0$ (positivité de l'intégrale).
- si $f \leq g$ sur $[a, b]$, alors $\int_{[a, b]} f \leq \int_{[a, b]} g$ (croissance de l'intégrale).

Proposition 20.3.5 (inégalité de la moyenne)

Soit $f : [a, b] \rightarrow \mathbb{K}$, continue par morceaux. Alors $\left| \int_{[a, b]} f \right| \leq \int_{[a, b]} |f| \leq (b - a) \sup_{[a, b]} |f(x)|$.

Définition 20.3.1 (valeur moyenne d'une fonction)

Soit f dans $\mathcal{C}_m([a, b], \mathbb{K})$. La quantité $\frac{1}{b - a} \int_{[a, b]} f$ est appelée *valeur moyenne* de f sur $[a, b]$.

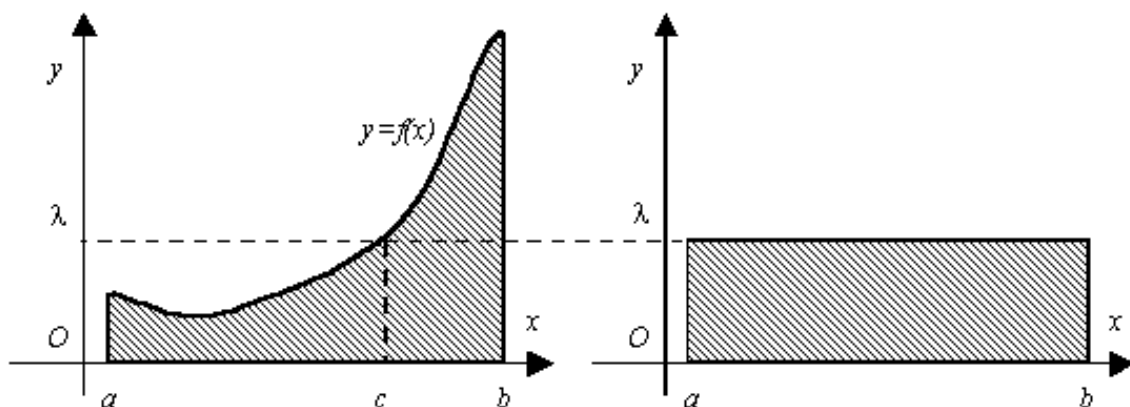
Si f est à valeurs réelles, on a $\inf_{[a, b]} f \leq \frac{1}{b - a} \int_{[a, b]} f \leq \sup_{[a, b]} f$.

En particulier, si $f : [a, b] \rightarrow \mathbb{R}$ est continue, il existe c dans $[a, b]$ tel que $\int_{[a, b]} f = (b - a)f(c)$.

La valeur moyenne λ de f vérifie l'égalité $\int_{[a, b]} f = \int_{[a, b]} \lambda$.

C'est donc le scalaire par lequel on peut remplacer f sans changer son intégrale sur $[a, b]$.

Sur la figure ci-dessous (où f est à valeurs réelles !) les deux aires hachurées sont donc égales.



Proposition 20.3.6 (relation de Chasles)

Soit $f : [a, b] \rightarrow \mathbb{K}$ une fonction continue par morceaux, et soit c un point de $]a, b[$.

Alors f est continue par morceaux sur $[a, c]$ et sur $[c, b]$, et on a : $\int_{[a,b]} f = \int_{[a,c]} f + \int_{[c,b]} f$.

20.3.3 Extension de la définition et nouvelle notation

Dans cette section, I désigne un intervalle de \mathbb{R} , d'intérieur non vide.

Définition 20.3.2 (notation de l'intégrale entre deux points)

Soit $f : I \rightarrow \mathbb{K}$, continue par morceaux. Soit a, b deux éléments quelconques de I .

Si $a < b$, on note $\int_a^b f = \int_{[a,b]} f$; si $a > b$, on note $\int_a^b f = -\int_{[b,a]} f$; si $a = b$, on note $\int_a^b f = 0$.

Si f est continue par morceaux sur I , on a donc défini $\int_a^b f$ pour deux points quelconques de I .

Il est clair que $\int_a^b f = -\int_b^a f$ pour tous a et b de I .

Plutôt que de noter $\int_a^b f$, on note souvent $\int_a^b f(t) dt$ (où t est une variable *muette*).

Cette notation s'avère pratique dans le calcul des intégrales par changement de variable.

Attention à la position des bornes

On est passé de la notation $\int_{[a,b]} f$ (avec $a < b$) à la notation $\int_a^b f(t) dt$ (a, b quelconques).

Avec cette nouvelle notation, les propriétés relatives à la linéarité restent valables.

Mais attention : les propriétés relatives à la positivité et à la croissance (dans le cas de fonctions à valeurs réelles, bien sûr) dépendent de la position respective des bornes de l'intégrale (le mieux est de vérifier que ces bornes sont « dans le bon sens »)

Par exemple, l'inégalité de la moyenne devient : $\forall (a, b) \in I^2, \left| \int_a^b f \right| \leq |b - a| \sup_{[a,b]} |f|$.

Relation de Chasles

Soit $f : I \rightarrow \mathbb{K}$, continue par morceaux, et soit a, b, c trois points quelconques de I .

On a toujours la relation de Chasles : $\int_a^b f = \int_a^c f + \int_c^b f$.

On peut généraliser à une suite finie c_1, \dots, c_n de points de I : $\int_{c_1}^{c_n} f = \sum_{k=1}^{n-1} \int_{c_k}^{c_{k+1}} f$.

Utilisation d'une translation, de la parité, de la périodicité

L'invariance de l'intégrale par translation s'écrit : $\int_a^b f(t) dt = \int_{a+\alpha}^{b+\alpha} f(t - \alpha) dt$.

Si f est paire, alors $\int_{-a}^a f(t) dt = 2 \int_0^a f(t) dt$. Si f est impaire, alors $\int_{-a}^a f(t) dt = 0$.

Si f est T -périodique, alors on a $\int_{a+kT}^{b+kT} f = \int_a^b f$ (avec k dans \mathbb{Z}) et $\int_a^{a+T} f = \int_b^{b+T} f$.

20.3.4 Intégrale d'une fonction continue de signe constant

Proposition 20.3.7 (intégrale d'une fonction continue de signe constant)

Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction continue, gardant un signe constant sur $[a, b]$ (avec $a \neq b$).

Alors on a l'équivalence : $\int_{[a,b]} f(t) dt = 0 \Leftrightarrow \forall t \in [a, b], f(t) = 0$.

Conséquences immédiates :

- Si $f : [a, b] \rightarrow \mathbb{R}^+$ est continue, avec $a < b$, et si f n'est pas identiquement nulle, alors $\int_{[a,b]} f > 0$.
- Si f, g sont continues sur $[a, b]$, avec $a < b$, et si $f \leq g$ sur $[a, b]$ mais $f \neq g$, alors $\int_{[a,b]} f < \int_{[a,b]} g$.

Proposition 20.3.8 (inégalité de Cauchy-Schwarz)

Soit $f, g : [a, b] \rightarrow \mathbb{R}$, continues par morceaux. Alors on a $\left(\int_a^b f(t)g(t) dt\right)^2 \leq \int_a^b f^2(t) dt \int_a^b g^2(t) dt$.

Quand f, g sont continues, il y a égalité si et seulement si f et g sont proportionnelles.

NB : l'inégalité de Cauchy-Schwarz est valable quelle que soit la position relative de a et b .

20.3.5 Sommes de Riemann de f sur $[a, b]$

Définition 20.3.3 (sommées de Riemann d'une fonction continue sur un segment)

Soit $f : [a, b] \rightarrow \mathbb{K}$ une fonction continue. Soit n un entier strictement positif.

La quantité $R_n(f) = \frac{b-a}{n} \sum_{k=0}^{n-1} f\left(a + k \frac{b-a}{n}\right)$ est appelée *somme de Riemann* d'indice n de f sur $[a, b]$.

Avec ces notations, on a $\lim_{n \rightarrow +\infty} R_n(f) = \int_a^b f(t) dt$.

Dans ce résultat, on peut remplacer $\sum_{k=0}^{n-1}$ par $\sum_{k=0}^n$ ou par $\sum_{k=1}^n$: cela ne change rien.

Interprétation géométrique

Si f est à valeurs réelles, la quantité $R_n(f)$ est la somme des aires algébriques des rectangles de base $[x_k, x_{k+1}]$ de hauteur $f(x_k)$. Il est « clair » que lorsque n tend vers $+\infty$, la somme $R_n(f)$ tend vers l'aire algébrique comprise entre l'axe Ox et la courbe $y = f(x)$, c'est-à-dire vers l'intégrale de f de a à b .

Utilisation pour calculer des limites de suites

La proposition précédente permet de calculer la limite d'une suite dont le terme général u_n peut être interprété comme une somme de Riemann. Il est alors recommandé de comparer u_n avec la forme générale d'une somme de Riemann, pour éviter toute erreur sur le segment $[a, b]$ et sur la fonction f .

Posons par exemple : $u_n = \frac{1}{n+1} + \frac{1}{n+2} + \dots + \frac{1}{2n}$.

On constate que $u_n = \frac{1}{n} \sum_{k=1}^n f\left(a + k \frac{b-a}{n}\right)$, avec $a = 0$, $b = 1$ et $f(x) = \frac{1}{1+x}$.

On en déduit $\lim_{n \rightarrow \infty} u_n = \frac{1}{b-a} \int_a^b f(x) dx = \int_0^1 \frac{dx}{1+x} = \ln 2$.

20.3.6 Méthode des trapèzes

Proposition 20.3.9 (formule à un seul trapèze)

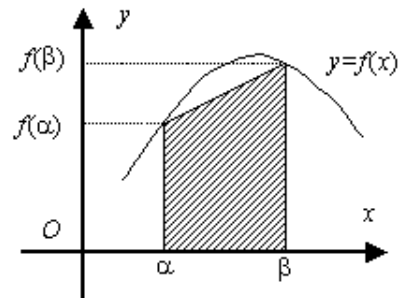
Soit $f : [\alpha, \beta] \rightarrow \mathbb{K}$ une fonction de classe \mathcal{C}^2 .

Soit $M_2(f) = \sup_{[\alpha, \beta]} |f''|$. Alors on a la majoration $\left| \int_{\alpha}^{\beta} f(t) dt - \frac{f(\alpha) + f(\beta)}{2} (\beta - \alpha) \right| \leq \frac{(\beta - \alpha)^3}{12} M_2(f)$.

Sur le schéma ci-contre, on voit comment on approche l'intégrale de f sur $[\alpha, \beta]$ par celle d'une fonction affine φ .

Graphiquement, on approche l'aire du domaine situé entre l'axe Ox et la courbe $y = f(x)$ par celle du trapèze construit sur les points $(\alpha, 0)$, $(\beta, 0)$, $(\alpha, f(\alpha))$, $(\beta, f(\beta))$.

La proposition précédente indique un majorant de l'erreur commise dans cette approximation.



Proposition 20.3.10 (formule à n trapèzes)

Soit $f : [a, b] \rightarrow \mathbb{K}$ une fonction de classe \mathcal{C}^2 . Soit $M_2(f) = \sup_{[a, b]} |f''|$.

Soit n un entier strictement positif. Pour tout k de $\{0, 1, \dots, n-1\}$, on pose $x_k = a + k \frac{b-a}{n}$.

Alors on a la majoration $\left| \int_a^b f(t) dt - \frac{b-a}{2n} \sum_{k=0}^{n-1} (f(x_k) + f(x_{k+1})) \right| \leq \frac{(b-a)^3}{12n^2} M_2(f)$.

La « méthode des trapèzes » consiste à décomposer $[a, b]$ en n sous-segments égaux $[x_k, x_{k+1}]$ et à écrire

$$\int_{x_k}^{x_{k+1}} f(t) dt \approx \frac{f(x_k) + f(x_{k+1})}{2} (x_{k+1} - x_k) \text{ c'est-à-dire } \int_{x_k}^{x_{k+1}} f(t) dt \approx \frac{b-a}{2n} (f(x_k) + f(x_{k+1}))$$

Après sommation, on en déduit l'approximation suivante de l'intégrale de f sur $[a, b]$:

$$\int_a^b f(t) dt \approx I_n(f) \text{ avec } I_n(f) = \frac{b-a}{2n} \sum_{k=0}^{n-1} (f(x_k) + f(x_{k+1})) = \frac{b-a}{n} \left(\frac{f(a) + f(b)}{2} + \sum_{k=1}^{n-2} f(x_k) \right)$$

Si $f : [a, b] \rightarrow \mathbb{R}$ est concave alors $I_n(f) \leq \int_a^b f(t) dt$, et si f est convexe alors $\int_a^b f(t) dt \leq I_n(f)$.

20.4 Intégrale et primitives

20.4.1 Dérivée de la fonction $x \mapsto \int_0^x f(t) dt$

Pour les définitions relatives aux primitives, on se reportera à la section 5.1.1.

Définition 20.4.1 (dérivée d'une intégrale fonction de sa borne supérieure)

Soit $f : I \rightarrow \mathbb{K}$ une fonction continue. Soit a, b deux éléments de I .

Alors la fonction $F : x \mapsto F(x) = \int_a^x f(t) dt$ est la primitive de f sur I qui s'annule au point a .

En particulier, toute fonction continue sur un intervalle y possède des primitives.

Le résultat précédent peut s'écrire, de façon « décontractée » : $\left(\int_a^x f(t) dt \right)' = f(x)$.

Il exprime que l'intégration est en quelque sorte l'opération inverse de la dérivation.

Proposition 20.4.1 (expression d'une intégrale à l'aide d'une primitive quelconque)

Soit $f : I \rightarrow \mathbb{K}$ une application continue.

Pour toute primitive F de f , on a : $\forall (a, b) \in I^2, \int_a^b f(t) dt = [F]_a^b = F(b) - F(a)$.

20.4.2 Méthodes d'intégration

Proposition 20.4.2 (intégration par parties)

Soit f et g deux fonctions de classe \mathcal{C}^1 sur un intervalle I , à valeurs dans \mathbb{K} .

Pour tous a, b dans l'intervalle I , on a : $\int_a^b f(x)g'(x) dx = [f(x)g(x)]_a^b - \int_a^b f'(x)g(x) dx$.

Remarque : le thème de l'intégration par parties a été abordé dans la section 5.2.3

Proposition 20.4.3 (intégrations par parties répétées)

Soit f et g deux fonctions de classe \mathcal{C}^n sur le segment $[a, b]$, à valeurs dans \mathbb{K} .

On a alors l'égalité : $\int_a^b f(t)g^{(n)}(t) dt = \left[\sum_{k=0}^{n-1} (-1)^k f^{(k)}(t)g^{(n-1-k)}(t) \right]_{t=a}^{t=b} + (-1)^n \int_a^b f^{(n)}(t)g(t) dt$.

Si $n = 2$, on obtient : $\int_a^b f(t)g''(t) dt = [f(t)g'(t) - f'(t)g(t)]_{t=a}^{t=b} + \int_a^b f''(t)g(t) dt$.

De même, si $n = 3$, on obtient : $\int_a^b f(t)g'''(t) dt = [f(t)g''(t) - f'(t)g'(t) + f''(t)g(t)]_{t=a}^{t=b} - \int_a^b f'''(t)g(t) dt$.

Proposition 20.4.4 (intégration par changement de variable)

Soit $f : I \rightarrow \mathbb{K}$ une fonction continue.

Soit $\varphi : J \rightarrow \mathbb{R}$, de classe \mathcal{C}^1 sur J , telle que $\varphi(J) \subset I$.

Alors, pour tous a, b de J , on a $\int_a^b f(\varphi(t)) \varphi'(t) dt = \int_c^d f(x) dx$, où $c = \varphi(a)$, et $d = \varphi(b)$.

Remarque : le thème de l'intégration par changement de variable a été abordé dans la section 5.2.4

Utilisation d'un changement de variable affine

Il est possible de transformer une intégrale sur $[a, b]$ en une intégrale sur $[0, 1]$ ou $[-1, 1]$.

Il suffit pour cela de poser $x = a + t(b - a)$: quand t parcourt $[0, 1]$, x parcourt $[a, b]$.

On obtient alors : $\int_a^b f(x) dx = (b - a) \int_0^1 g(t) dt$, avec $g(t) = f(a + t(b - a))$.

De même, en posant $x = \frac{a+b}{2} + t \frac{b-a}{2}$: quand t parcourt $[-1, 1]$, x parcourt $[a, b]$.

On en déduit l'égalité : $\int_a^b f(x) dx = \frac{b-a}{2} \int_{-1}^1 h(t) dt$, avec $h(t) = f\left(\frac{a+b}{2} + t \frac{b-a}{2}\right)$.

20.4.3 Calcul de primitives

Les méthodes de primitivation ont été vues dans le chapitre « Techniques d'analyse (intégration) ».

On se reportera notamment à la section 5.3

20.5 Formules de Taylor

Proposition 20.5.1 (formule de Taylor avec reste intégral, au point a , à l'ordre n)

Soit $f : [a, b] \rightarrow \mathbb{K}$ une application de classe \mathcal{C}^{n+1} .

$$\text{Alors on a l'égalité : } f(b) = \sum_{k=0}^n \frac{(b-a)^k}{k!} f^{(k)}(a) + \overbrace{\int_a^b \frac{(b-t)^n}{n!} f^{(n+1)}(t) dt}^{R_n}.$$

La quantité R_n est appelée reste intégral de la formule de Taylor de f à l'ordre n en a .

Proposition 20.5.2 (inégalité de Taylor-Lagrange pour une fonction de classe \mathcal{C}^{n+1})

Soit $f : I \rightarrow \mathbb{K}$ une application de classe \mathcal{C}^{n+1} . Soit a et b deux points de I .

$$\text{Alors on a la majoration : } \left| f(b) - \sum_{k=0}^n \frac{(b-a)^k}{k!} f^{(k)}(a) \right| \leq M \frac{|b-a|^{n+1}}{(n+1)!}, \text{ où } M = \sup_{[a,b]} |f^{(n+1)}|.$$

$$\text{En posant } h = b - a, \text{ cela s'écrit : } \left| f(a+h) - \sum_{k=0}^n \frac{h^k}{k!} f^{(k)}(a) \right| \leq M \frac{|h|^{n+1}}{(n+1)!}, \text{ où } M = \sup_{[a, a+h]} |f^{(n+1)}|.$$

Cas particuliers $n = 0$ et $n = 1$

Pour $n = 0$, c'est l'inégalité des accroissements finis : $|f(b) - f(a)| \leq M_1 |b - a|$ où $M_1 = \sup_{[a,b]} |f'|$.

Pour $n = 1$, on trouve : $|f(b) - f(a) - (b-a)f'(a)| \leq M_2 \frac{(b-a)^2}{2!}$ où $M_2 = \sup_{[a,b]} |f''|$.

Exemples d'utilisation

On applique ici l'inégalité de Taylor-Lagrange à $t \mapsto \sin(t)$ et $t \mapsto \cos(t)$ sur $[0, x]$:

$$\begin{array}{lll} \left| \sin x - x \right| \leq \frac{|x|^3}{3!} & \left| \sin x - x + \frac{x^3}{3!} \right| \leq \frac{|x|^5}{5!} & \left| \sin x - x + \frac{x^3}{3!} - \frac{x^5}{5!} \right| \leq \frac{|x|^7}{7!} \\ \left| \cos x - 1 \right| \leq \frac{x^2}{2!} & \left| \cos x - 1 + \frac{x^2}{2!} \right| \leq \frac{x^4}{4!} & \left| \cos x - 1 + \frac{x^2}{2!} - \frac{x^4}{4!} \right| \leq \frac{x^6}{6!} \end{array}$$

Proposition 20.5.3 (formule de Taylor-Young pour une fonction \mathcal{C}^n au voisinage d'un point)

Soit $f : I \rightarrow \mathbb{K}$. On suppose que f est de classe \mathcal{C}^n au voisinage d'un point a de I .

$$\text{Alors, au voisinage de } a, \text{ on a l'égalité } f(x) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-a)^k + o((x-a)^n)$$

En particulier, si f est de classe \mathcal{C}^n au voisinage de 0 , on a le développement :

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(0)}{k!} x^k + o(x^n) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \dots + \frac{f^{(n)}(0)}{n!}x^n + o(x^n).$$

Différence de nature entre les formules de Taylor

La formule de Taylor-Young n'a qu'un rôle « local ». Elle permet d'obtenir une approximation de f (d'autant meilleure que n est élevé) au voisinage immédiat d'un point donné a de I .

L'intérêt essentiel de la formule de Taylor-Young est donc qu'elle fournit un développement limité d'une fonction en un point (on se reportera au chapitre « Analyse asymptotique », notamment 9.4.3).

La formule de Taylor avec reste intégral, et l'inégalité de Taylor-Lagrange, ont une nature « globale » : elles nécessitent des hypothèses sur $[a, b]$, et elles fournissent un résultat lui aussi valable sur $[a, b]$.

Chapitre 21

Séries numériques

Sommaire

21.1 Séries convergentes ou divergentes	512
21.1.1 Définitions de base	512
21.1.2 Premiers exemples	513
21.1.3 Propriétés des séries convergentes	513
21.2 Séries à termes positifs	515
21.2.1 Convergence par utilisation de comparaisons	515
21.2.2 Utilisation des séries de référence	516
21.2.3 Convergence absolue	516
21.3 Représentation décimale des réels	516

21.1 Séries convergentes ou divergentes

21.1.1 Définitions de base

Définition 21.1.1 (sommations partielles d'une série)

Soit $(u_n)_{n \geq 0}$ une suite d'éléments de \mathbb{K} . Soit N un entier naturel.

La quantité $S_N = \sum_{n=0}^N u_n$ est appelée *somme partielle d'indice N de la série* $\sum u_n$.

Avec les notations précédentes, on a $u_0 = S_0$ et, $\forall n \in \mathbb{N}^*$, $u_n = S_n - S_{n-1}$.

La suite $(u_n)_{n \geq 0}$ est donc à son tour déterminée par la donnée des sommes partielles $(S_n)_{n \geq 0}$.

Définition 21.1.2 (convergence ou divergence d'une série)

Soit $(u_n)_{n \geq 0}$ une suite de \mathbb{K} .

On dit que la *série* $\sum u_n$ est *convergente* si la suite (S_N) de ses sommes partielles est convergente.

Dans le cas contraire, on dit que la série $\sum_{n \geq 0} u_n$ est *divergente*.

Définition 21.1.3 (somme d'une série convergente)

Soit $\sum u_n$ une série convergente.

La quantité $\lim_{N \rightarrow \infty} S_N$ est notée $\sum_{n=0}^{\infty} u_n$ et est appelée *somme* de la série $\sum u_n$.

L'unicité de la limite implique l'unicité de la somme d'une série convergente.

Ne pas confondre « nature » et « somme » d'une série

Déterminer la *nature* d'une série, c'est dire si elle est convergente ou divergente. C'est ensuite un autre problème que de calculer la somme en cas de convergence.

Parfois les deux problèmes peuvent être traités simultanément, mais l'énoncé pourra demander de prouver d'abord la convergence, *puis* de calculer la somme.

Enfin il est fréquent qu'on puisse prouver la convergence d'une série sans pouvoir en calculer la somme.

Influence de la modification d'un nombre fini de termes

On ne modifie pas la *nature* de la série $\sum u_n$ en changeant la valeur d'un nombre fini des u_n .

En revanche, en cas de convergence, on a toutes les chances de modifier la somme de la série.

Si la suite (u_n) n'est définie que pour $n \geq k$, on adapte facilement les définitions précédentes, et en cas de convergence, la somme de la série est notée $\sum_{n=k}^{\infty} u_n$.

Définition 21.1.4 (reste d'ordre N d'une série convergente)

Soit $\sum_{n \geq 0} u_n$ une série de \mathbb{K} , convergente, de somme S . Soit N un entier naturel.

On appelle *reste d'ordre N* de cette série, la quantité $R_N = S - S_N = \sum_{n=0}^{\infty} u_n - \sum_{n=0}^N u_n = \sum_{n=N+1}^{\infty} u_n$.

Par définition de la convergence d'une série, on a $\lim_{N \rightarrow \infty} R_N = 0$.

Mais attention : on ne doit pas dire qu'une série est convergente \Leftrightarrow son reste d'indice N tend vers 0, car l'existence même de ce reste suppose déjà que la série converge.

21.1.2 Premiers exemples

▷ **La série** $\sum \frac{1}{n(n+1)}$

Pour tout $n \geq 1$, posons $u_n = \frac{1}{n(n+1)}$. On a $u_n = \frac{1}{n} - \frac{1}{n+1}$ donc $S_N = \sum_{n=1}^N u_n = 1 - \frac{1}{N+1}$.

Il en résulte que $\sum \frac{1}{n(n+1)}$ converge et que $\sum_{n=1}^{+\infty} \frac{1}{n(n+1)} = 1$.

▷ **La série exponentielle**

Pour tout x de \mathbb{R} , on sait que $\lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{x^k}{k!} = e^x$ (utilisation de la formule de Taylor avec reste intégral).

La série $\sum \frac{x^n}{n!}$ (dite *série exponentielle*) est donc convergente et sa somme est $\sum_{n=0}^{+\infty} \frac{x^n}{n!} = e^x$.

▷ **La série harmonique**

Pour tout $n \geq 1$, on a $\frac{1}{n} \geq \int_n^{n+1} \frac{dt}{t} = \ln(n+1) - \ln(n)$. Ainsi $S_N = \sum_{n=1}^N \frac{1}{n} \geq \ln(N+1)$.

Il en résulte $\lim_{N \rightarrow +\infty} S_N = +\infty$. La série $\sum_{n \geq 1} \frac{1}{n}$ (dite *série harmonique*) est donc divergente.

▷ **La série harmonique alternée**

Pour tout $n \geq 1$, on a $\frac{(-1)^{n-1}}{n} = (-1)^{n-1} \int_0^1 t^{n-1} dt$.

Il en résulte $S_N = \sum_{n=1}^N \frac{(-1)^{n-1}}{n} = \int_0^1 \sum_{n=1}^N (-t)^{n-1} dt = \int_0^1 \frac{1 - (-t)^N}{1+t} dt = \ln(2) - \int_0^1 \frac{(-t)^N}{1+t} dt$.

Mais $\left| \int_0^1 \frac{(-t)^N}{1+t} dt \right| \leq \int_0^1 t^N dt = \frac{1}{N+1}$. Il en résulte $\lim_{N \rightarrow +\infty} S_N = \ln(2)$.

Ainsi la série $\sum_{n \geq 1} \frac{(-1)^{n-1}}{n}$ (dite *série harmonique alternée*) est convergente et sa somme est $\ln(2)$.

21.1.3 Propriétés des séries convergentes

Proposition 21.1.1 (linéarité de la somme)

Soit $(u_n)_{n \geq 0}$ et $(v_n)_{n \geq 0}$ deux suites de \mathbb{K} . Soit λ et μ dans \mathbb{K} .

Si les séries $\sum_{n \geq 0} u_n$ et $\sum_{n \geq 0} v_n$ sont convergentes, alors la série $\sum_{n \geq 0} (\lambda u_n + \mu v_n)$ est convergente.

Dans ce cas, on a l'égalité $\sum_{n=0}^{\infty} (\lambda u_n + \mu v_n) = \lambda \sum_{n=0}^{\infty} u_n + \mu \sum_{n=0}^{\infty} v_n$.

Si $\lambda \neq 0$, les séries $\sum_{n \geq 0} (\lambda u_n)$ et $\sum_{n \geq 0} u_n$ sont de même nature.

Remarque importante sur les sommes de deux séries

Si les séries $\sum u_n$ et $\sum v_n$ sont de natures différentes, alors la série $\sum(u_n + v_n)$ est divergente.

Il est possible que $\sum(u_n + v_n)$ soit convergente alors que ni $\sum u_n$ ni $\sum v_n$ ne le sont.

On ne développera donc pas $\sum_{n=0}^{\infty} (u_n + v_n)$ sans vérifier préalablement l'existence de $\sum_{n=0}^{\infty} u_n$ et $\sum_{n=0}^{\infty} v_n$.

Proposition 21.1.2 (une condition nécessaire, mais non suffisante, de convergence)

Si la série $\sum_{n \geq 0} u_n$ est convergente, alors $\lim_{n \rightarrow \infty} u_n = 0$. Attention la réciproque est fautive !

Si $\lim_{n \rightarrow \infty} u_n$ n'existe pas ou est non nulle, la série $\sum_{n \geq 0} u_n$ est dite grossièrement divergente.

Définition 21.1.5 (séries géométriques)

Soit a un élément de \mathbb{K} . La série $\sum a^n$ est appelée série géométrique de raison a .

Si $a \neq 1$, on sait que, pour tout N de \mathbb{N} : $S_N = \sum_{n=0}^N a^n = \frac{1 - a^{N+1}}{1 - a}$

Proposition 21.1.3 (condition nécessaire et suffisante de convergence de la série géométrique)

La série géométrique $\sum a^n$ est convergente si et seulement si $|a| < 1$.

Dans ce cas, la somme est $S = \sum_{n=0}^{\infty} a^n = \frac{1}{1 - a}$, et plus généralement : $R_N = \sum_{n=N+1}^{\infty} a^n = \frac{a^{N+1}}{1 - a}$.

Proposition 21.1.4 (lien suite-série)

Soit $(u_n)_{n \geq 0}$ une suite de \mathbb{K} .

La suite (u_n) et la série $\sum(u_{n+1} - u_n)$ sont de même nature.

En cas de convergence, on $\sum_{n=0}^{+\infty} (u_{n+1} - u_n) = (\lim_{n \rightarrow +\infty} u_n) - u_0$.

Cette propriété ramène l'étude de la nature d'une suite (u_n) à celle de la série $\sum_{n \geq 0} (u_{n+1} - u_n)$.

Elle permet aussi d'étudier $\sum_{n \geq 0} v_n$ si on sait écrire v_n sous la forme $v_n = u_{n+1} - u_n$.

Proposition 21.1.5 (convergence des séries à termes complexes)

Soit $(z_n)_{n \geq 0}$ une suite de nombres complexes.

La série $\sum_{n \geq 0} z_n$ est convergente si et seulement si les séries réelles $\sum_{n \geq 0} \operatorname{Re}(z_n)$ et $\sum_{n \geq 0} \operatorname{Im}(z_n)$ le sont.

En cas de convergence, on a : $\sum_{n=0}^{\infty} z_n = \sum_{n=0}^{\infty} \operatorname{Re}(z_n) + i \sum_{n=0}^{\infty} \operatorname{Im}(z_n)$.

21.2 Séries à termes positifs

21.2.1 Convergence par utilisation de comparaisons

Proposition 21.2.1 (convergence par majoration des sommes partielles)

Soit $\sum u_n$ une série à termes réels positifs ou nuls.

Il est clair que la suite (S_N) des sommes partielles de la série $\sum u_n$ est croissante.

Dans ces conditions, la série $\sum u_n$ converge si et seulement si la suite (S_N) est majorée.

Si l'hypothèse $u_n \geq 0$ n'est vraie qu'à partir d'un certain rang n_0 , le résultat reste valable.

Sachant que $\sum u_n$ et $\sum(-u_n)$ sont de même nature, l'énoncé se généralise (avec des modifications évidentes) aux séries réelles dont le terme général est de signe constant à partir d'un certain rang.

Proposition 21.2.2 (convergence par domination)

Soit $\sum u_n$ et $\sum v_n$ deux séries, telles que $0 \leq u_n \leq v_n$ pour tout $n \geq n_0$.

— si on sait que la série $\sum_{n \geq 0} u_n$ diverge, alors la série $\sum_{n \geq 0} v_n$ diverge.

— si la série $\sum_{n \geq 0} v_n$ converge, la série $\sum_{n \geq 0} u_n$ converge, et pour tout $N \geq n_0$, on a : $\sum_{n=N}^{\infty} u_n \leq \sum_{n=N}^{\infty} v_n$.

Proposition 21.2.3 (convergence par équivalence)

Soit $\sum u_n$ et $\sum v_n$ deux séries à termes réels positifs (au moins à partir d'un certain rang).

Si $u_n \sim v_n$, alors les séries $\sum u_n$ et $\sum v_n$ sont de même nature.

La proposition précédente vaut aussi pour des séries à termes négatifs à partir d'un certain rang.

L'hypothèse selon laquelle les u_n et v_n gardent un signe constant est essentielle.

En effet, si $u_n = \frac{(-1)^n}{\sqrt{n}}$ et $v_n = \ln(1 + u_n)$, on a $u_n \sim v_n$ mais $\sum u_n$ converge alors que $\sum v_n$ diverge.

Proposition 21.2.4 (comparaison série-intégrale dans le cas monotone)

Soit $f : [n_0, +\infty[\rightarrow \mathbb{R}^+$ une fonction, continue par morceaux, décroissante et à valeurs positives.

Pour tout entier $n > n_0$, on a l'encadrement : $\int_n^{n+1} f(t) dt \leq f(n) \leq \int_{n-1}^n f(t) dt$.

Pour tout $N > n_0$, on en déduit : $\int_{n_0+1}^{N+1} f(t) dt \leq \sum_{n=n_0+1}^N f(n) \leq \int_{n_0}^N f(t) dt$

Il en résulte que la série $\sum u_n$ converge si et seulement si $\lim_{n \rightarrow +\infty} \int_{n_0}^n f(t) dt$ est un nombre réel.

Proposition 21.2.5 (séries de référence)

Séries de Riemann : la série $\sum_{n \geq 1} \frac{1}{n^\alpha}$ est convergente si et seulement si $\alpha > 1$.

Séries de Bertrand : la série $\sum_{n \geq 2} \frac{1}{n^\alpha \ln^\beta(n)}$ converge si et seulement si $\alpha > 1$ ou ($\alpha = 1$ et $\beta > 1$)

Par exemple, la série $\sum \frac{1}{n^2}$ est convergente, et on montre que $\sum_{n=1}^{+\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$.

21.2.2 Utilisation des séries de référence

Proposition 21.2.6 (utilisation des séries de référence de Riemann)

S'il existe $\alpha > 1$ et $M \geq 0$ tels que $0 \leq n^\alpha u_n \leq M$ pour $n \geq n_0$, alors la série $\sum u_n$ converge.

C'est notamment le cas si $\lim_{n \rightarrow \infty} n^\alpha u_n = 0$, c'est-à-dire si $u_n = o\left(\frac{1}{n^\alpha}\right)$.

S'il existe $M > 0$ tel que $u_n \geq \frac{M}{n}$ pour $n \geq n_0$, alors $\sum_{n \geq 0} u_n$ diverge. C'est le cas si $u_n \sim \frac{\lambda}{n}$, où $\lambda > 0$.

Proposition 21.2.7 (règles de d'Alembert)

Soit (u_n) une suite de \mathbb{R}^{+*} . On suppose que $\lim_{n \rightarrow \infty} \frac{u_{n+1}}{u_n} = \alpha$.

Si $0 \leq \alpha < 1$, la série $\sum_{n \geq 0} u_n$ est convergente. Si $\alpha > 1$, la série $\sum_{n \geq 0} u_n$ est divergente.

Si $\alpha = 1$ on ne peut rien dire : c'est le cas douteux de la règle de d'Alembert.

Toutefois, si $\lim_{n \rightarrow \infty} \frac{u_{n+1}}{u_n} = 1^+$, alors la série $\sum_{n \geq 0} u_n$ est divergente.

21.2.3 Convergence absolue

Définition 21.2.1 (convergence absolue d'une série numérique)

On dit que la série $\sum u_n$ est *absolument convergente* si la série $\sum |u_n|$ est convergente.

Proposition 21.2.8 (la convergence absolue implique la convergence)

Si la série $\sum u_n$ est absolument convergente, alors elle est convergente et $\left| \sum_{n=0}^{\infty} u_n \right| \leq \sum_{n=0}^{\infty} |u_n|$.

La réciproque est fautive, comme le montre l'exemple de la série harmonique alternée $\sum \frac{(-1)^n}{n}$.

Proposition 21.2.9 (critère suffisant de convergence absolue par domination)

Soit (z_n) une suite complexe, et soit (v_n) une suite d'éléments de \mathbb{R}^+ .

Si $z_n = O(v_n)$, et si $\sum v_n$ converge, alors $\sum z_n$ est absolument convergente, donc convergente.

21.3 Représentation décimale des réels

Proposition 21.3.1 (développement décimal propre d'un réel de l'intervalle $[0, 1[$)

Pour tout réel x de $[0, 1[$, il existe une unique suite $(a_n)_{n \geq 1}$, vérifiant les propriétés suivantes :

– pour tout $n \geq 1$, a_n est un entier de $\{0, 1, \dots, 9\}$, et : $\forall n_0 \in \mathbb{N}, \exists n \geq n_0, a_n \neq 9$.

– on a l'égalité $x = \sum_{n=1}^{+\infty} a_n 10^{-n}$, appelée *développement décimal propre* du réel x .

Ce résultat s'étend aux réels positifs en écrivant : $x = [x] + (x - [x]) = \sum_{n=-\infty}^{+\infty} a_n 10^{-n}$.

Ici $\sum_{n=-\infty}^0 a_n 10^{-n} = a_{-m} 10^m + \dots + a_{-2} 10^2 + a_{-1} 10 + a_0$ est l'écriture décimale finie de l'entier $[x]$.

Chapitre 22

Ensembles finis, probabilités

Sommaire

22.1 Ensembles finis, dénombrements	520
22.1.1 Cardinal d'un ensemble fini	520
22.1.2 Calculs de quelques cardinaux usuels	521
22.1.3 p -listes et combinaisons	521
22.1.4 Rappels sur les coefficients binomiaux	523

22.1 Ensembles finis, dénombrements

22.1.1 Cardinal d'un ensemble fini

Pour tout entier naturel, on note $E_n = \llbracket 0, n-1 \rrbracket = \{m \in \mathbb{N}, 0 \leq m \leq n-1\}$.

Ainsi $E_0 = \emptyset$, $E_1 = \{0\}$, $E_2 = \{0, 1\}$, $E_3 = \{0, 1, 2\}$, etc.

La fonction Python `E` ci-contre renvoie la liste ordonnée des éléments de E_n :

```
>>> def E(n): return list(range(n))
>>> E(10)
[0, 1, 2, 3, 4, 5, 6, 7, 8, 9]
```

Proposition 22.1.1 (existence d'injections, de surjections, de bijections, entre E_n et E_p)

Il existe une injection de E_n dans E_p si et seulement si $n \leq p$.

Il existe une surjection de E_n sur E_p si et seulement si $n \geq p$.

Il existe une bijection de E_n sur E_p si et seulement si $n = p$.

Définition 22.1.1 (notion d'ensemble fini)

Un ensemble non vide A est dit *fini* s'il existe une bijection de E_n sur A , avec $n \geq 0$.

L'entier n , s'il existe, est unique et est appelé le *cardinal* de A . On note $n = \text{Card}(A)$.

En particulier $\text{Card}(\emptyset) = 0$. Un ensemble non fini est dit *infini*.

Notation : le cardinal d'un ensemble fini A est également noté $|A|$, ou $\#A$.

On peut dire que $E_n = \llbracket 0, n-1 \rrbracket$ est l'ensemble « modèle » de tout ensemble fini de cardinal n (un peu comme \mathbb{K}^n est le modèle des espaces vectoriels de dimension n).

La bijection $f : E_n \rightarrow A$ représente une « numérotation » des éléments de A .

Cette numérotation est exhaustive (surjectivité de f), et à deux indices différents i et j correspondent deux éléments différents a_i et a_j de A (c'est l'injectivité de f).

Intuitivement, l'entier $\text{Card}(A)$ représente le « nombre d'éléments » de A (conformément au programme de la classe de MPSI, on en reste d'ailleurs le plus souvent à cette formulation intuitive).

Si $m \leq n$, l'intervalle $A = \llbracket m, n \rrbracket$ est fini de cardinal $n - m + 1$.

En effet l'application f définie par $f(k) = m + k$ est bijective de $E_{n-m+1} = \llbracket 0, n-m \rrbracket$ sur A .

Proposition 22.1.2 (ensemble en bijection avec un ensemble fini)

Soit A un ensemble fini. S'il existe une bijection de A sur B , alors B est fini et $\text{Card}(B) = \text{Card}(A)$.

Réciproquement, si B est fini de même cardinal que A , il existe une bijection de A sur B .

On peut caractériser les parties finies de \mathbb{N} :

Proposition 22.1.3 (parties finies de \mathbb{N})

Une partie A non vide de \mathbb{N} est finie \Leftrightarrow elle est majorée. En particulier \mathbb{N} est infini.

On en déduit le résultat suivant :

Proposition 22.1.4 (cardinal d'une partie d'un ensemble fini)

Soit A un ensemble fini et B une partie de A . Alors B est un ensemble fini et $\text{Card}(B) \leq \text{Card}(A)$.

Plus précisément, on a $\text{Card}(B) = \text{Card}(A)$ si et seulement si $B = A$.

Proposition 22.1.5 (bijection entre deux ensembles finis de même cardinal)

Soit A, B deux ensembles finis de même cardinal. Soit f une application de A dans B .

Alors f est bijective si et seulement si elle est injective, et si et seulement si elle est surjective.

Si E est infini, il existe des applications de E dans E qui sont injectives et non surjectives, ou surjectives et non injectives. Par exemple, l'application $n \mapsto 2n$ est injective et non surjective de \mathbb{N} dans lui-même, et l'application $n \mapsto \lfloor n/2 \rfloor$ est surjective et non injective..

22.1.2 Calculs de quelques cardinaux usuels

Proposition 22.1.6 (produit cartésien de deux ensembles finis)

Soit A, B deux ensembles finis. Alors $A \times B$ est un ensemble fini et $\text{Card}(A \times B) = \text{Card}(A) \text{Card}(B)$.

Plus généralement, si les ensembles A_1, A_2, \dots, A_n sont finis, alors $\text{Card}\left(\prod_{k=1}^n A_k\right) = \prod_{k=1}^n \text{Card}(A_k)$.

Si A est fini, alors $\text{Card}(A^n) = \text{Card}(A)^n$ pour tout $n \geq 1$.

Proposition 22.1.7 (cardinal de l'union de deux ensembles finis)

Si A et B sont finis disjoints, alors $A \cup B$ est fini et $\text{Card}(A \cup B) = \text{Card}(A) + \text{Card}(B)$.

Dans le cas général, on a : $\text{Card}(A \cup B) = \text{Card}(A) + \text{Card}(B) - \text{Card}(A \cap B)$.

Si A_1, \dots, A_n sont finis, alors $\bigcup_{k=1}^n A_k$ est fini et $\text{Card}\left(\bigcup_{k=1}^n A_k\right) \leq \sum_{k=1}^n \text{Card}(A_k)$.

L'inégalité précédente est une égalité si et seulement si les A_k sont disjoints deux à deux.

Si A, B, C sont trois ensembles finis (en notant $|E|$ plutôt que $\text{Card}(E)$) :

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|$$

Proposition 22.1.8 (cardinal de l'ensemble des applications entre deux ensembles finis)

Soit A, B deux ensembles finis, avec $\text{Card}(A) = n$ et $\text{Card}(B) = p$.

Soit $\mathcal{F}(A, B)$ l'ensemble des applications de A dans B .

Alors $\mathcal{F}(A, B)$ est un ensemble fini, et $\text{Card}(\mathcal{F}(A, B)) = p^n$.

Ainsi $\text{Card}(\mathcal{F}(A, B)) = \text{Card}(B)^{\text{Card}(A)}$, ce qui justifie qu'on note parfois B^A plutôt que $\mathcal{F}(A, B)$.

Proposition 22.1.9 (cardinal de l'ensemble des parties d'un ensemble fini)

Soit A un ensemble fini, de cardinal n . Alors $\mathcal{P}(A)$ est un ensemble fini et $\text{Card}(\mathcal{P}(A)) = 2^n$.

22.1.3 p -listes et combinaisons

Définition 22.1.2 (p -listes d'éléments quelconques d'un ensemble de cardinal n)

Soit B un ensemble fini de cardinal $n \geq 1$. Soit p un entier, avec $p \geq 1$.

Les p -uplets (b_1, b_2, \dots, b_p) , où les b_k sont quelconques dans B , sont appelés p -listes d'éléments de B .

Le nombre de p -listes d'un ensemble B de cardinal n est $\text{Card}(B^p) = n^p$.

Il faut bien noter que dans la définition précédente, les b_k ne sont pas supposés distincts.

En revanche, l'ordre dans lequel les b_k apparaissent est important.

Ainsi, si a, b, c sont trois éléments distincts de A , les six 3-listes (a, b, c) , (a, c, b) , (b, a, c) , (b, c, a) , (c, a, b) et (c, b, a) sont distinctes.

De même, si a et b sont distincts, les huit 3-listes (a, a, a) , (a, a, b) , (a, b, a) , (a, b, b) , (b, a, a) , (b, a, b) , (b, b, a) et (b, b, b) sont distinctes.

Proposition 22.1.10 (p -listes d'éléments distincts d'un ensemble de cardinal n)

Soit B un ensemble fini de cardinal $n \geq 1$. Soit p un entier, avec $1 \leq p \leq n$.

Le nombre de p -listes (b_1, b_2, \dots, b_p) d'éléments distincts deux à deux dans B est $\frac{n!}{(n-p)!}$.

On a supposé $p \leq n$ car si $p > n$ il est impossible de former une p -liste d'éléments distincts de B .

Si par exemple $B = \{x, y, z, t\}$:

– le nombre de 2-listes d'éléments distincts de B est $\frac{4!}{2!} = 12$.

les voici : (x, y) , (x, z) , (x, t) , (y, x) , (y, z) , (y, t) , (z, x) , (z, y) , (z, t) , (t, x) , (t, y) et (t, z) .

– le nombre de 3-listes d'éléments distincts de B est $\frac{4!}{1!} = 24$.

les voici :

((x, y, z)	(x, y, t)	(x, z, y)	(x, z, t)	(x, t, y)	(x, t, z)
)	(y, x, z)	(y, x, t)	(y, z, x)	(y, z, t)	(y, t, x)	(y, t, z)
)	(z, x, y)	(z, x, t)	(z, y, x)	(z, y, t)	(z, t, x)	(z, t, y)
)	(t, x, y)	(t, x, z)	(t, y, x)	(t, y, z)	(t, z, x)	(t, z, y)

Proposition 22.1.11 (nombre d'injections entre deux ensembles finis)

Soit A et B deux ensembles finis, avec $\text{Card}(A) = p$, $\text{Card}(B) = n$, et $1 \leq p \leq n$.

Le nombre d'applications injectives de A dans B est $\frac{n!}{(n-p)!}$.

Proposition 22.1.12 (nombre de permutations d'un ensemble de cardinal n)

Soit A un ensemble fini de cardinal n . Le nombre de bijections de A dans lui-même est $n!$.

Ces applications sont appelées permutations de l'ensemble A .

Définition 22.1.3 (p -combinaisons d'éléments d'un ensemble de cardinal n)

Soit B un ensemble fini de cardinal $n \geq 1$. Soit p un entier, avec $p \geq 1$.

On dit que $\{b_1, b_2, \dots, b_p\}$, où les b_k sont distincts dans B , est une p -combinaison d'éléments de B .

Le nombre de p -combinaisons d'un ensemble B de cardinal n est $\binom{n}{p} = \frac{n!}{p!(n-p)!}$.

L'expression « p -combinaison d'éléments de B » est synonyme de « partie à p éléments de B ».

Quand on parle de p -combinaison, les p éléments concernés sont forcément distincts deux à deux (puisque seul compte l'ensemble formé par ces éléments).

Si par exemple $B = \{x, y, z, t\}$:

– il y a $\binom{4}{2} = 6$ « 2-combinaisons » : $\{x, y\}$, $\{x, z\}$, $\{x, t\}$, $\{y, z\}$, $\{y, t\}$, et $\{z, t\}$.

– il y a $\binom{4}{3} = 4$ « 3-combinaisons » : $\{x, y, z\}$, $\{x, y, t\}$, $\{x, z, t\}$ et $\{y, z, t\}$.

22.1.4 Rappels sur les coefficients binomiaux

Les coefficients $\binom{n}{p}$ sont appelés « coefficients binomiaux ».

Ils ont été vus au chapitre « Calculs algébriques » (cf 2.3.2). On se contente ici de quelques rappels :

Proposition 22.1.13 (les deux relations fondamentales sur les coefficients binomiaux)

On a les identités $\binom{n}{p} = \binom{n}{n-p}$, et $\binom{n}{p} = \binom{n-1}{p} + \binom{n-1}{p-1}$.

Proposition 22.1.14 (formule du binôme dans \mathbb{C})

Pour x, y de \mathbb{C} , et pour n de \mathbb{N} : $(x + y)^n = \sum_{k=0}^n \binom{n}{k} x^k y^{n-k}$. En particulier : $(1 + x)^n = \sum_{k=0}^n \binom{n}{k} x^k$.

Par exemple, pour tous x et y de \mathbb{C} :

$$\begin{aligned} (x + y)^2 &= x^2 + 2xy + y^2 & (x - y)^2 &= x^2 - 2xy + y^2 \\ (x + y)^3 &= x^3 + 3x^2y + 3xy^2 + y^3 & (x - y)^3 &= x^3 - 3x^2y + 3xy^2 - y^3 \\ (x + y)^4 &= x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + y^4 & (x - y)^4 &= x^4 - 4x^3y + 6x^2y^2 - 4xy^3 + y^4 \end{aligned}$$

Proposition 22.1.15 (propriétés de conversion des coefficients binomiaux)

On a les égalités : $\binom{n}{p+1} = \frac{n-p}{p+1} \binom{n}{p}$, $\binom{n}{p} = \frac{n}{p} \binom{n-1}{p-1}$, $\binom{n}{p} = \frac{n}{n-p} \binom{n-1}{p}$.

Les coefficients binomiaux peuvent être calculés de proche en proche, en les disposant dans un tableau triangulaire connu sous le nom de « triangle de Pascal ».

On écrit ici une fonction (rustique!) pour former le triangle de Pascal sous forme de liste de listes.

```
>>> def trpascal(N):
    from math import factorial as f
    return [[f(n)//f(p)//f(n-p) for p in range(n+1)] for n in range(N+1)]
```

On importe le module pprint pour obtenir un affichage agréable :

```
>>> from pprint import pprint
>>> pprint(trpascal(9))
[[1],
 [1, 1],
 [1, 2, 1],
 [1, 3, 3, 1],
 [1, 4, 6, 4, 1],
 [1, 5, 10, 10, 5, 1],
 [1, 6, 15, 20, 15, 6, 1],
 [1, 7, 21, 35, 35, 21, 7, 1],
 [1, 8, 28, 56, 70, 56, 28, 8, 1],
 [1, 9, 36, 84, 126, 126, 84, 36, 9, 1]]
```