جامعة مولاي إسماعيل
UNIVERSITÉ MOULAY ISMAÏL
كلية العلوم والتقنيات
FACULTÉ DES SCIENCES ET TECHNIQUES

Faculté des Sciences et Techniques

Errachidia

Département de Mathématiques

COURS D'ALGÈBRE 1

Pr. JAWAD SALHI

Année Universitaire 2021/2022

Table des matières

AVANT DE COMMENCER

Ce document a été écrit pour accompagner le cours dispensé à partir de 2020.

Ce polycopié est long : il est conçu pour être un document de travail plutôt qu'une transcription du cours dispensé en amphi. Il n'a pas vocation à se substituer à ce dernier : en particulier, il comporte des passages qui ne seront pas traités en amphi (et ne seront pas au programme de l'examen). À l'inverse, certains éléments ou exemples développés en cours ne figureront pas dans le polycopié.

Une particularité de ce document, qui explique en partie sa longueur, est la place qu'y occupent des exercices de manipulation. Leur but est de vous permettre de travailler sur des situations proches des exemples détaillés dans le polycopié, afin de consolider votre compréhension des notions du cours. Ils ne sont donc pas conçus pour être traités en TD, mais pour vous accompagner dans votre travail quotidien sur le cours.

Dans la version actuelle, il reste inévitablement des fautes de frappe et de mathématiques. N'hésitez pas à me les signaler.

CHAPITRE 1

Éléments de logique-Ensembles-Applications

Sommaire

1.1 Raisonnement

1.1.1 Implication

Définition 1.1.1 (Implication)**.** Si P et Q sont deux assertions, alors (NON P) ou Q se note $P \Rightarrow Q$. Le connecteur $\Rightarrow$ est appelé implication, et l'assertion $P \Rightarrow Q$ se lit alors P implique Q.

Remarque 1.1.2. Par définition du "ou", il est alors immédiat que l'assertion $P \Rightarrow Q$ est fausse lorsque P est vraie et Q fausse, et uniquement dans ce cas ; donc :
- si P est fausse alors $P \Rightarrow Q$ est vraie ;
- si $P \Rightarrow Q$ est vraie et si P est vraie, alors Q est vraie.

Remarque 1.1.3. Soient P et Q deux assertions. Pour démontrer $P \Rightarrow Q$:
- si P est fausse, alors il n'y a rien à faire ;
- si P est vraie, alors on est dans l'obligation de prouver que Q est vraie.

Remarque 1.1.4. Soit P et Q deux prédicats de la variable x. Pour démontrer l'assertion "$\forall x \in E, P(x) \Rightarrow Q(x)$" qui commence par "$\forall$", il suffit de prendre un élément x quelconque de E puis de prouver l'implication $P(x) \Rightarrow Q(x)$, ce qui amène à supposer que $P(x)$ est vraie. C'est pourquoi le début d'une telle démonstration doit être : Soit $x \in E$ tel que $P(x)$.

1.1.2 Négation de l'implication

En utilisant la définition de l'implication et la règle de négation d'un "ou", on a immédiatement le résultat suivant, très utile pour nier automatiquement une implication.

Définition 1.1.5 (Négation d'une implication)**.** Si P et Q sont deux assertions, la négation de $P \Rightarrow Q$ s'écrit :

$$P \text{ et } (\text{NON } Q).$$

En utilisant la règle de négation d'une assertion commençant par un "$\forall$" on en déduit la proposition suivante.

Proposition 1.1.6. *Soit $P(x)$ et $Q(x)$ deux prédicats de la variable x définis sur un ensemble E. La négation de l'assertion "$\forall x \in E, P(x) \Rightarrow Q(x)$" est :*

$$\exists x \in E, P(x) \text{ et } (\text{NON } Q(x)).$$

1.1.3 Équivalence

Définition 1.1.7 (Équivalence)**.** Si P et Q sont deux assertions, l'assertion $P \Leftrightarrow Q$ est l'abréviation de :

$$(P \Rightarrow Q) \textbf{ ET } (Q \Rightarrow P).$$

Le connecteur "$\Leftrightarrow$" est appelé équivalence, et l'assertion $P \Leftrightarrow Q$ se lit "P équivaut à Q" ou encore "P est équivalente à Q".

1.1.4 Contraposée, implication réciproque

Définition 1.1.8.
- $(\text{NON}\, Q) \Rightarrow (\text{NON}\, P)$ est la **contraposée** de l'implication $P \Rightarrow Q$.
- $Q \Rightarrow P$ est l'implication **réciproque** de l'implication $P \Rightarrow Q$.

Proposition 1.1.9. *L'assertion $P \Rightarrow Q$ est vraie si, et seulement si, sa contraposée est vraie.*

Démonstration. Par définition, $(\text{NON}\, Q) \Rightarrow (\text{NON}\, P)$ s'écrit

$$(\text{NON}(\text{NON}\, Q)) \text{ OU } (\text{NON}\, P) \quad \text{ou encore} \quad Q \text{ OU } (\text{NON}\, P),$$

ce qui est équivalent à $(\text{NON}\, P)$ OU Q, et donc à $P \Rightarrow Q$. □

1.2 Opérations sur les ensembles

1.2.1 Premières notions ensemblistes

Définition 1.2.1 (Définition en extension)**.** Pour décrire un ensemble, on peut donner la liste de ses éléments. Par exemple :

$$A = \{1, -8, \pi\}$$

est un ensemble comportant trois éléments.

Dans une telle description, l'ordre d'écriture "ne compte pas" : on a aussi $A = \{\pi, -8, 1\}$.

Définition 1.2.2 (Définition en compréhension)**.** On peut aussi décrire un ensemble en spécifiant quelles sont les propriétés qui distinguent les objets qui appartiennent à E des objets qui n'y appartiennent pas. Par exemple, si nous définissons

$$B = \{n \in \mathbb{N} \mid \exists p \in \mathbb{N} : n = p^2\}$$

on a $9 \in B$, mais $3 \notin B$.

Explication : L'écriture $E = \{x \in A \mid P(x)\}$ signifie " l'ensemble des x de A vérifiant la propriété $P(x)$".

Naturellement, un ensemble donné peut admettre plusieurs descriptions différentes : si

$$C = \{n \in \mathbb{N} \mid n < 5\},$$

on a bien sûr :

$$C = \{0, 1, 2, 3, 4\}.$$

Ici, une description est **en compréhension**, l'autre **en extension**, mais elles définissent le même ensemble.

1.2.2 Inclusion

Définition 1.2.3 (Relation d'inclusion)**.** Soient E et F deux ensembles. L'assertion

$$\forall x \in F \quad x \in E$$

se lit "F est inclus dans E" ou "F est une partie de E", et se note $F \subset E$.

Définition 1.2.4 (Ensemble vide)**.**
— On admet qu'il existe un unique ensemble, appelé **ensemble vide** et noté $\emptyset$, qui ne contient aucun élément et qui est donc tel que, pour tout prédicat $P(x)$ de la variable x, l'assertion "$\exists x \in \emptyset \quad P(x)$" est fausse.
— En appliquant l'item précédent à NON P, on déduit que, pour tout prédicat $P(x)$ de la variable x, l'assertion "$\forall x \in \emptyset \quad P(x)$" est vraie.

Proposition 1.2.5. *Pour tout ensemble E, on a : $\emptyset \subset E$.*

Démonstration. D'après le second item ci-dessus, l'assertion :

$$\forall x \in \emptyset \quad x \in E$$

est vraie, ce qui entraîne que $\emptyset \subset E$. □

1.2.3 Égalité d'ensembles

Définition 1.2.6 (Égalité d'ensembles)**.** Par convention, deux ensembles E et F sont égaux si, et seulement si, tout élément de l'un est élément de l'autre, ou encore :

$$(\forall x \in E \quad x \in F) \quad \text{ET} \quad (\forall x \in F \quad x \in E).$$

En termes d'inclusion, cette équivalence devient :

$$E = F \Leftrightarrow (E \subset F \text{ ET } F \subset E).$$

Remarque 1.2.7. La principale méthode pour montrer que deux ensembles E et F sont égaux est d'établir la double inclusion $E \subset F$ et $F \subset E$.

Exercice 1.2.8. Écrire une assertion permettant d'exprimer que $E \neq F$.

Solution. Pour exprimer $E \neq F$ il suffit de nier la relation $E = F$, qui s'écrit :

$$(\forall x \in E \quad x \in F) \quad \text{ET} \quad (\forall x \in F \quad x \in E).$$

— La règle de négation du "ET" nous dit que la négation de ce qui précède est :

$$\big(\text{NON}(\forall x \in E \quad x \in F)\big) \quad \text{OU} \quad \big(\text{NON}(\forall x \in F \quad x \in E)\big).$$

— En appliquant la règle de négation du quantificateur "$\forall$", cela devient :

$$(\exists x \in E \quad x \notin F) \quad \text{OU} \quad (\exists x \in F \quad x \notin E).$$

1.2.4 Intersection, réunion

Définition 1.2.9 (Intersection, réunion)**.** Soit E et F deux ensembles.

— La **réunion** de E et F est l'ensemble, noté $E \cup F$ (lire "E union F"), des éléments qui sont dans E ou dans F ; on a donc :

$$x \in E \cup F \Leftrightarrow (x \in E \quad \text{OU} \quad x \in F).$$

— L'**intersection** de E et F est l'ensemble, noté $E \cap F$ (lire "E inter F"), des éléments qui sont à la fois dans E et dans F ; on a donc :

$$x \in E \cap F \Leftrightarrow (x \in E \quad \text{ET} \quad x \in F).$$

Définition 1.2.10 (Ensembles disjoints)**.** Deux ensembles E et F sont disjoints si $E \cap F = \emptyset$.

Notation : Si E et F sont disjoints, l'union $E \cup F$ peut être notée $E \sqcup F$.

1.2.5 Différence, complémentaire

Définition 1.2.11 (Différence, complémentaire)**.** Soit E et F deux ensembles.

— On appelle **différence ensembliste** E moins F, l'ensemble, noté $E \setminus F$ (lire "E privé de F"), des éléments qui sont dans E mais pas dans F ; on a donc :

$$x \in E \setminus F \Leftrightarrow (x \in E \quad \text{ET} \quad x \notin F).$$

— Lorsque $F \subset E$, l'ensemble $E \setminus F$ s'appelle **complémentaire** de F dans E, et il se note alors $\complement_E F$ ou F^c ou encore $\bar{F}$.

Exercice 1.2.12. Pour $A \subset E$ et $B \subset E$, on pose $A \Delta B = (A \setminus B) \cup (B \setminus A)$. Montrer que

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

1.2.6 Ensemble des parties d'un ensemble

Définition 1.2.13. Soit E un ensemble. L'ensemble des parties de E se note $\mathcal{P}(E)$, et vérifie :

$$F \in \mathcal{P}(E) \Leftrightarrow F \subset E.$$

Exercice 1.2.14. Soient E et F deux ensembles. Montrer que :

$$\mathcal{P}(E) = \mathcal{P}(F) \Leftrightarrow E = F.$$

Solution. Si $E = F$, alors $\mathcal{P}(E) = \mathcal{P}(F)$.

Réciproquement, supposons que $\mathcal{P}(E) = \mathcal{P}(F)$. F est un élément de $\mathcal{P}(F)$ et donc F est un élément $\mathcal{P}(E)$. Mais alors $F \subset E$. En échangeant les rôles de E et F on a aussi $E \subset F$ et finalement $E = F$.

Exercice 1.2.15 (Paradoxe de Russell). Le but de cette question est de montrer que la notion d' "ensemble de tous les ensembles" est problématique.

Supposons que cette notion ne soit pas problématique et notons E l'ensemble de tous les ensembles. On définit alors l'ensemble

$$R = \{A \in E \mid A \notin A\}.$$

Montrer qu'on a à la fois $R \in R$ et $R \notin R$. Conclure.

Remarque 1.2.16. La théories des ensembles, oeuvre des mathématiciens allemands Georg Cantor et Richard Dedekind, apparaît à la fin du XIX^e siècle. Dans cette première approche, que l'on qualifie maintenant de "naîve", on appelle ensemble n'importe quelle collection d'objets, ce qui conduit à des paradoxes. Pour remédier à de tels paradoxes, une théorie axiomatique a été élaborée. C'est la théorie des ensembles Zermelo-Fraenkel, en abrégé ZF, conçue par Ernst Zermelo en 1908 et modifiée par Abraham Fraenkel en 1921 et 1922. La théorie ZF sert depuis cette époque de fondements aux mathématiques, dont elle permet une construction rigoureuse, même si sa consistance, c'est-à-dire l'absence des paradoxes, ne pourra jamais être prouvée, comme le montre Kurt Godel en 1931. Les mathématiciens d'aujourd'hui font le pari de cette consistance et fondent les mathématiques sur la théorie ZF.

Théorème 1.2.17 (Théorème de Knaster-Tarski)**.** *Soit E un ensemble, $\mathcal{P}(E)$ l'ensemble de ses parties et $\varphi : \mathcal{P}(E) \longrightarrow \mathcal{P}(E)$ une application croissante, c'est-à-dire telle que $A \subset B \Rightarrow \varphi(A) \subset \varphi(B)$. Alors φ admet un point fixe, c'est-à-dire qu'il existe une partie M de E telle que $\varphi(M) = M$.*

Démonstration. Voir TD. □

1.2.7 Couple, produit cartésien

Notations

— Á partir de deux éléments x et y, on peut construire le **couple** (x, y) avec, si x_1, x_2, y_1 et y_2 sont des éléments, la propriété fondamentale :

$$(x_1, y_1) = (x_2, y_2) \Leftrightarrow (x_1 = x_2 \quad \text{ET} \quad y_1 = y_2).$$

— Soit E et F deux ensembles. On appelle **produit cartésien** de E et F, l'ensemble, noté $E \times F$, des couples (x, y) avec $x \in E$ et $y \in F$. On a donc :

$$E \times F = \{z \mid \exists x \in E \, \exists y \in F \quad z = (x, y)\}.$$

Remarque 1.2.18. Avec ces notations, on écrit souvent :

$$E \times F = \{(x, y); x \in E \quad \text{ET} \quad y \in F\}.$$

Attention : Le couple (x, y) s'écrit avec des parenthèses, pas des accolades !

— Avec accolades, $\{x, y\}$ désigne un ensemble et, si $x = y$, alors $\{x, y\} = \{x\}$.

— Si $x \neq y$, alors $(x, y) \neq (y, x)$ mais $\{x, y\} = \{y, x\}$.

1.3 Grands types de démonstrations

1.3.1 Par contraposée

Pour prouver une implication du type $P \Rightarrow Q$, il peut être intéressant de prouver sa contraposée, l'implication $\text{NON}\, Q \Rightarrow \text{NON}\, P$, qui lui est équivalente. Cela n'est évidemment pertinent que si cette contraposée est plus facile à prouver !

Exemple 1.3.1. Fixons un entier naturel n. Supposons qu'on veuille démontrer l'assertion suivante :

$$\text{si} \quad n^2 \quad \text{est impair, alors} \quad n \quad \text{est nécessairement impair.}$$

On peut utiliser le fait que cette assertion est équivalente à sa contraposée, à savoir :

$$\text{Si} \quad n \quad \text{est pair, alors} \quad n^2 \quad \text{est pair.}$$

Cette dernière est peut-être plus facile à démontrer : si n est pair, alors on peut écrire $n = 2k$ avec $k \in \mathbb{N}$, et alors $n^2 = (2k)^2 = 2(2k^2)$, donc n^2 peut s'écrire sous la forme $2m$ où $m = 2k^2$ est entier, ainsi n^2 est pair.

Exercice 1.3.2. Soit $a \in \mathbb{R}$. Montrer que :

$$\big(\forall \varepsilon > 0, \quad |a| \leq \varepsilon\big) \Rightarrow a = 0.$$

Solution. Raisonnons par contraposée et montrons que :

$$a \neq 0 \Rightarrow \big(\exists \varepsilon > 0, \quad |a| > \varepsilon\big).$$

Supposons que $a \neq 0$. Posons $\varepsilon = \dfrac{|a|}{2}$. Alors $\varepsilon > 0$ et on a bien $|a| > \varepsilon$.

Exercice 1.3.3. On rappelle que tout nombre rationnel non nul peut s'écrire sous la forme $\dfrac{p}{q}$, où p et q sont des entiers relatifs premiers entre eux. Un nombre réel est dit irrationnel s'il n'appartient pas à $\mathbb{Q}$.

1. Soit n un entier naturel. Démontrer que si $\sqrt{n}$ n'est pas entier, alors il est irrationnel.
2. En déduire que si p désingne un nombre premier, alors $\sqrt{p}$ est irrationnel.

Solution.

1. $\sqrt{0} = 0$ et $\sqrt{1} = 1$ sont entiers. Soit n un entier supérieur ou égal à 2. Supposons que $\sqrt{n}$ soit rationnel. Il existe deux entiers naturels non nuls a et b tels que $\sqrt{n} = \dfrac{a}{b}$ ou encore tels que $n = \dfrac{a^2}{b^2}$. Si $b = 1$, alors $\sqrt{n} = a$ est un entier. Si $b > 2$, tout facteur premier de a^2 ou de b^2 apparaît à un exposant pair dans la décomposition primaire de a^2 ou de b^2. Il en est de même pour tout facteur premier de $n = \dfrac{a^2}{b^2}$ ce qui signifie que n est un carré parfait ou encore que $\sqrt{n}$ est un entier. On a montré que si $\sqrt{n}$ est rationnel, alors $\sqrt{n}$ est entier. Par contraposition, si $\sqrt{n}$ n'est pas entier, alors $\sqrt{n}$ est irrationnel.
2. Soit p un nombre premier. p est en particulier un entier supérieur ou égal à 2. Montrons que $\sqrt{p}$ n'est pas entier. Dans le cas contraire, il existe un entier naturel $n > 2$ tel que $\sqrt{p} = n$ ou encore tel que $n^2 = p$. Cette égalité est impossible par unicité de la décomposition en facteurs premier car le nombre premier p apparaît à un exposant dans le premier membre de cette égalité et à un exposant impair dans le second. Donc $\sqrt{p}$ n'est pas entier puis $\sqrt{p}$ est irrationnel d'après la question précédente.

Exercice 1.3.4. Soit $n \in \mathbb{N}^*$. Montrer que si l'entier $(n^2 - 1)$ n'est pas divisible par 8, alors l'entier n est pair.

1.3.2 Par l'absurde

Pour commencer, on appelle contradiction toute proposition de la forme : Q ET (NON Q).

Le principe du raisonnement par l'absurde s'énonce alors ainsi : Si, en supposant que NON P est vraie, on peut exhiber une assertion Q telle que Q ainsi que NON Q soient vraies, alors on en déduit que P est vraie.

Exercice 1.3.5 (Exemples de nombres irrationnels). Dans cet exercice, on se propose de démontrer l'irrationalité de quelques nombres réels.

1. Montrer que $\sqrt{2} \notin \mathbb{Q}$.
2. Montrer que $\sqrt{3} \notin \mathbb{Q}$.
3. Montrer que $\dfrac{\ln 2}{\ln 3} \notin \mathbb{Q}$.
4. On rappelle que $\boldsymbol{e} = \sum_{k=0}^{+\infty} \dfrac{1}{k!}$. On se propose de démontrer que le nombre $\boldsymbol{e}$ est un nombre irrationnel. Pour cela, on fait l'hypothèse qu'il existe p et q, entiers naturels non nuls, tels que $\boldsymbol{e} = \dfrac{p}{q}$ et on démontre que cette hypothèse conduit à une contradiction.
   Pour tout entier naturel n, on pose :
   $$u_n = \sum_{k=0}^{n} \frac{1}{k!} \quad \text{et} \quad v_n = u_n + \frac{1}{n \times n!}.$$
   4.1. Démontrer que les suites $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ sont adjacentes, puis montrer que :
   $$u_q < \boldsymbol{e} < v_q.$$
   4.2. Aboutir à une contradiction en multipliant les termes de cet encadrement par $q! \times q$.

Solution. Les deux premiers points sont laissés au lecteur.

3. $\dfrac{\ln 2}{\ln 3}$ est un réel strictement positif. Raisonnons par l'absurde et supposons que $\dfrac{\ln 2}{\ln 3}$ soit un rationnel strictement positif. Alors, il existe deux entiers naturels non nuls a et b tels que $\dfrac{\ln 2}{\ln 3} = \dfrac{a}{b}$ ou encore tels que $b \ln 2 = a \ln 3$ ou encore $e^{b \ln 2} = e^{a \ln 3}$ ou enfin tels que $2^b = 3^a$. Cette égalité est impossible par unicité de la décomposition en facteurs premiers car 2 et 3 sont des nombres premiers et car $a > 0$ et $b > 0$. Donc $\dfrac{\ln 2}{\ln 3}$ est irrationnel.

4.1. Pour tout entier naturel non nul n, $u_{n+1} - u_n = \frac{1}{(n+1)!} > 0$. Donc la suite $(u_n)_{n\in\mathbb{N}^*}$ est strictement croissante.
De plus, pour tout entier naturel non nul n, on a :
$$\begin{aligned} v_{n+1} - v_n &= \frac{1}{(n+1)!} + \frac{1}{(n+1)\times(n+1)!} - \frac{1}{n\times n!} \\ &= \frac{n(n+1) + n - (n+1)^2}{n(n+1)\times(n+1)!} \\ &= -\frac{1}{n(n+1)\times(n+1)!} < 0. \end{aligned}$$

Donc la suite $(v_n)_{n\in\mathbb{N}^*}$, est strictement décroissante.

Enfin, $\lim\limits_{n\to+\infty}(v_n-u_n)=\lim\limits_{n\to+\infty}\dfrac{1}{n\times n!}=0$. Donc les suites $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ sont adjacentes. La suite $(u_n)_{n\in\mathbb{N}}$ tend vers $\boldsymbol{e}$ en croissant strictement et donc pour tout entier naturel non nul n, $u_n<\boldsymbol{e}$. La suite $(v_n)_{n\in\mathbb{N}}$ a même limite que la suite $(u_n)_{n\in N}$ et donc la suite $(v_n)_{n\in\mathbb{N}}$ tend vers $\boldsymbol{e}$ en décroissant strictement. On en déduit que pour tout entier naturel non nul n, $v_n>\boldsymbol{e}$. On a montré que

$$\forall n\in\mathbb{N}^*, u_n<\boldsymbol{e}<v_n$$

En particulier, $u_q<\boldsymbol{e}<v_q$.

4.2. D'après la question précédente,

$$q!\times q\times u_q<q!\times q\times\boldsymbol{e}<q!\times q\times v_q,$$

ce qui s'écrit encore

$$q\sum_{k=0}^{q}\frac{q!}{k!}<p\times q!<1+q\sum_{k=0}^{q}\frac{q!}{k!}.$$

Pour tout entier $k\in[\![0,q]\!]$, $\dfrac{q!}{k!}$ est un entier et donc $q\sum_{k=0}^{q}\dfrac{q!}{k!}$ est un entier. Ainsi, l'entier $p\times q!$ est strictement compris entre deux entiers consécutifs. Ceci est une contradiction et il était donc absurde de supposer $\boldsymbol{e}$ rationnel. On a donc montré que $\boldsymbol{e}$ est irrationnel.

Exercice 1.3.6. 1. Soit $n\in\mathbb{N}^*$ et soient $p_1,\cdots,p_n$ des nombres premiers. Soit $N=1+p_1\cdot p_2\cdots p_n$. Montrer que, pour tout $k\in[\![1,n]\!]$, p_k ne divise pas N.

2. Montrer que l'ensemble des nombres premiers est infini.

Solution. 1. Raisonnons par l'absurde et supposons qu'il existe $k\in[\![1,n]\!]$ tel que p_k divise N. Alors, on aurait :
— p_k divise $p_1\cdot p_2\cdots p_n$.
— p_k divise $1+p_1\cdot p_2\cdots p_n$.
Ainsi p_k divise $1+p_1\cdot p_2\cdots p_n-p_1\cdot p_2\cdots p_n=1$ (On a utilisé le fait que si $d\mid a$ et $d\mid b$ alors $d\mid(a-b)$). Donc $p_k=1$ ou -1. Absurde car on avait supposé p_k premier (donc supérieure ou égale à 2). On cnclut que : $\forall k\in[\![1,n]\!]$, p_k ne divise pas N.

2. On raisonne une autre fois par l'absurde. Supposons que l'ensemble des nombres premiers soit fini. On pourrait alors tous les énumérer : $p_1,p_2,\cdots,p_n$. Considérons $N=1+p_1\cdot p_2\cdots p_n$. Comme $N\geq 2$, alors d'après l'exercice 1.3.20, N possède un diviseur premier p qui ne peut être l'un des p_k. L'hypothèse que l'ensemble des nombres premiers soit fini est donc contradictoire, ce qui prouve que cet ensemble est infini.

Exercice 1.3.7. Soit $n\in\mathbb{N}^*$. Soient $x_1,\ldots,x_{n+1}$ des points de l'intervalle $[0,1]$. Montrer qu'il existe $(i,j)\in[\![1,n+1]\!]^2$ tel que $i\neq j$ et $|x_i-x_j|\leqslant\frac{1}{n}$.

Solution. Deux pistes :
— Par l'absurde, en renumérotant les x_i de sorte à ce qu'ils soient rangés dans l'ordre croissant (pour une gestion plus facile des valeurs absolues). Sommer les $|x_{i+1}-x_i|$.

— Découper l'intervalle $[0,1]$ en n intervalles de longueur $\frac{1}{n}$.

Exercice 1.3.8. Soit $f : I \longrightarrow \mathbb{R}$ une fonction, où I est un intervalle de $\mathbb{R}$, continue et ne prenant qu'un nombre fini de valeurs. Montrer que f est constante.

Solution. Raisonnons par l'absurde et supposons qu'il existe $a < b$ dans I tel que $f(a) \neq f(b)$. Par le théorème des valeurs intermédiaire, toute valeur $[f(a), f(b)]$ (ou $[f(b), f(a)]$) est prise par f dans $[a, b]$. Comme cet intervalle contient un nombre infini de points, on aboutit à une contradiction. Donc, pour tous $a, b \in I$, on a : $f(a) = f(b)$ et f est constante.

Exercice 1.3.9. Montrer que :

$$\forall x \in \mathbb{R}_+^* \quad \forall y \in \mathbb{R} \quad \exists n \in \mathbb{N} \quad nx \geq y.$$

1.3.3 Par analyse-synthèse

Quand on veut déterminer l'ensemble des éléments d'un ensemble E qui satisfont une propriété $\mathcal{P}$, on raisonne souvent par analyse-synthèse de la manière suivante.

— Dans l'analyse, on part d'un élément quelconque de E et on montre que s'il satisfait la propriété $\mathcal{P}$, il a forcément telle ou telle tête et non telle autre. En résumé, **DANS L'ANALYSE, ON RESTREINT LE CHAMP DES SOLUTIONS POSSIBLES**.

— Dans la synthèse, on vérifie que les possibilités obtenues dans l'analyse sont plus que des possibilités, qu'elles sont bel et bien solutions du problème étudié, i.e. des éléments de E qui satisfont la propriété $\mathcal{P}$.

Exemple 1.3.10. Montrer qu'il existe une unique fonction $f : [-1,1] \to \mathbb{R}$ vérifiant :

$$\forall x \in \mathbb{R}, f(\cos(x)) = \cos(2x). \tag{1.1}$$

— **Analyse :** Considérons une fonction $f : [-1,1] \to \mathbb{R}$ vérifiant la contrainte (1.1), et voyons si nous pouvons comprendre qui est $f(a)$ lorsque $a \in [-1,1]$.
Soit a un élément de $[-1,1]$. Compte tenu des propriétés de la fonction cos, nous savons qu'il existe un réel x vérifiant $a = \cos(x)$. On a alors

$$f(a) = f(\cos(x)) = \cos(2x).$$

On rappelle maintenant une formule de trigonométrie : $\cos(2x) = 1 - 2\cos^2(x)$. On a donc nécessairement :

$$f(a) = 1 - 2a^2.$$

Mais alors, il n'y a qu'une possibilité pour f : c'est la fonction

$$\begin{aligned} \varphi : \mathbb{R} &\to \mathbb{R} \\ a &\mapsto 1 - 2a^2. \end{aligned} \tag{1.2}$$

— **Synthèse :** Montrons à présent que la fonction φ obtenue en (1.2) est effectivement solution du problème. Pour tout réel x, on a

$$\varphi(\cos(x)) = 1 - 2\cos(x)^2 = \cos(2x)$$

la fonction φ vérifie donc bien la condition espérée.

Exercice 1.3.11. Soit f une application de $\mathbb{R}$ dans $\mathbb{R}$. Montrer qu'il existe un unique couple (f_1, f_2) tel que l'on ait $f = f_1 + f_2$ avec f_1 (resp. f_2) fonction impaire (resp. paire) de $\mathbb{R}$ dans $\mathbb{R}$.

Solution. Soit f une application de $\mathbb{R}$ dans $\mathbb{R}$.

— **Analyse :** Supposons qu'il existe $f_1 : \mathbb{R} \longrightarrow \mathbb{R}$, impaire, et $f_2 : \mathbb{R} \longrightarrow \mathbb{R}$, paire, telles que $f = f_1 + f_2$ c'est-à-dire

$$\forall x \in \mathbb{R}, f(x) = f_1(x) + f_2(x).$$

On en déduit immédiatement :

$$\forall x \in \mathbb{R}, f(-x) = -f_1(x) + f_2(x),$$

ce qui donne :

$$\forall x \in \mathbb{R}, f_1(x) = \frac{f(x) - f(-x)}{2} \quad \text{et} \quad f_1(x) = \frac{f(x) + f(-x)}{2}.$$

Par suite il existe au plus un couple (f_1, f_2) répondant au problème.

— **Synthèse :** Pour tout $x \in \mathbb{R}$, posons :

$$f_1(x) = \frac{f(x) - f(-x)}{2} \quad \text{et} \quad f_1(x) = \frac{f(x) + f(-x)}{2}.$$

Il est alors immédiat de vérifier que, pour tout $x \in \mathbb{R}$:

$$f(x) = f_1(x) + f_2(x), f_1(-x) = -f_1(x) \text{ et } f_2(-x) = f_2(x),$$

ce qui prouve que le couple (f_1, f_2) répond au problème.

Exercice 1.3.12. On cherche toutes les isométries de $\mathbb{R}$, i.e. toutes les fonctions $f : \mathbb{R} \to \mathbb{R}$ telles que :

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

1. **Analyse :** Soit f une isométrie. On note δ la fonction $x \longmapsto f(x) - f(0)$ sur $\mathbb{R}$.
   a) Montrer, en étudiant la quantité $(f(x)-f(y))^2$, que pour tous $x, y \in \mathbb{R}$: $\delta(x)\delta(y) = xy$.
   b) En déduire la forme de f.
2. **Synthèse :** Conclure.

1.3.4 Par récurrence

Le théorème suivant permet de faire des raisonnements par récurrence.

Théorème 1.3.13. *Soit P un prédicat défini sur $\mathbb{N}$. Si $P(0)$ est vraie et si :*

$$\forall n \in \mathbb{N} \quad P(n) \Rightarrow P(n+1),$$

alors la propriété $P(n)$ est vraie pour tout entier naturel n.

Remarque 1.3.14. Pour démontrer une propriété $P(n)$ par récurrence, on doit commencer par établir $P(0)$ mais, ensuite

— soit on suppose $P(n)$ vraie pour un $n \in \mathbb{N}$ et on démontre $P(n+1)$;
— soit on suppose $P(n-1)$ vraie pour un $n \in \mathbb{N}^*$ et on démontre $P(n)$.

Un bon choix permet parfois des économies d'écriture (voir exercice suivant).

Exercice 1.3.15. Démontrer par récurrence l'identité :

$$\sum_{k=0}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}.$$

Solution. Pour $n \in \mathbb{N}$, soit $P(n)$ la relation : $\sum_{k=0}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$.

— **Initialisation :** La propriété $P(0)$ est vraie de façon évidente.
— **Hérédité :** Soit $n \in \mathbb{N}^*$. Supposons $P(n-1)$ vraie. On a alors :

$$\begin{aligned}
\sum_{k=0}^{n} k^2 &= \sum_{k=0}^{n-1} k^2 + n^2 \\
&= \frac{(n-1)n(2n-1)}{6} + n^2 \quad \text{d'après} \quad P(n-1) \\
&= \frac{n}{6}\Big((n-1)(2n-1) + 6n\Big) \\
&= \frac{n(2n^2+3n+1)}{6} \\
&= \frac{n(n+1)(2n+1)}{6}.
\end{aligned}$$

Ainsi $P(n)$ est vraie, ce qui termine la démonstration de la récurrence.

On peut aussi montrer par récurrence les identités classiques :

$$\sum_{k=0}^{n} k = \frac{n(n+1)}{2} \quad \text{et} \quad \sum_{k=0}^{n} k^3 = \frac{n^2(n+1)^2}{4}.$$

C'est un exercice facile laissé au lecteur.

Exercice 1.3.16. Démontrer, en utilisant un raisonnement par récurrence sur $n \in \mathbb{N}$, que tout ensemble fini à n éléments vérifie $\operatorname{card} \mathcal{P}(E) = 2^n$.

Solution. Raisonnons par récurrence sur $n = \operatorname{card} E$.

— **Initialisation :** Pour $n = 0$, $E = \emptyset$ et $\mathcal{P}(E) = \{\emptyset\}$. Donc $\operatorname{card} \mathcal{P}(E) = 1 = 2^0$.
— **Hérédité :** Soit $n \in \mathbb{N}$. Supposons le résultat acquis pour n et considérons un ensemble E de cardinal $n+1$. Soit $x \in E$ fixé. On peut distinguer deux catégories de parties de E :
  — celles qui contiennent x : $C_x = \{A \subset E; x \in A\}$.
  — celles qui ne contiennent pas x : $\mathcal{P}(E) \setminus C_x = \mathcal{P}(E \setminus \{x\})$.
  Comme $\operatorname{card} E \setminus \{x\} = n$, alors d'après l'hypothèse de récurrence : $\operatorname{card} \mathcal{P}(E \setminus \{x\}) = 2^n$. Par ailleurs, l'application

$$\begin{aligned}
\alpha : C_x &\longrightarrow \mathcal{P}(E \setminus \{x\}) \\
A &\longmapsto A \setminus \{x\}.
\end{aligned}$$

  est bijective de réciproque

$$\begin{aligned}
\beta : \mathcal{P}(E \setminus \{x\}) &\longrightarrow C_x \\
B &\longmapsto B \cup \{x\}.
\end{aligned}$$

(On vérifie facilement que $\beta \circ \alpha = Id_{C_x}$ et $\alpha \circ \beta = Id_{\mathcal{P}(E\setminus\{x\})}$).
D'où $\operatorname{card} C_x = 2^n$. Finalement, comme

$$\mathcal{P}(E) = C_x \sqcup \mathcal{P}(E \setminus \{x\}) \quad \text{(réunion disjointe)}$$

alors :

$$\begin{aligned}\operatorname{card} \mathcal{P}(E) &= \operatorname{card} C_x + \operatorname{card} \mathcal{P}(E \setminus \{x\}) \\ &= 2^n + 2^n = 2^{n+1}.\end{aligned}$$

1.3.5 Récurrence double

Si P est un prédicat défini sur $\mathbb{N}$, il arrive parfois que la justification de $P(n)$ nécessite l'utilisation de $P(n-1)$ et de $P(n-2)$. On fait alors ce quel'on appelle une récurrence d'ordre 2, ou encore récurrence à deux pas, qui est fondée sur le résultat suivant.

Corollaire 1.3.17 (Récurrence double)**.** *Soit P une propriété définie sur $\mathbb{N}$ avec $P(0)$ et $P(1)$ vraies ainsi que :*

$$\forall n \in \mathbb{N} \quad (P(n) \text{ et } P(n+1)) \Rightarrow P(n+2).$$

Alors, la propriété $P(n)$ est vraie pour tout $n \in \mathbb{N}$.

Exercice 1.3.18. On note $(u_n)_{n\in\mathbb{N}}$ la suite réelle définie par : $u_0 = 4, \quad u_1 = 5$ et pour tout $n \in \mathbb{N}$: $\quad u_{n+2} = 3u_{n+1} - 2u_n$. Montrer que pour tout $n \in \mathbb{N}$: $u_n = 2^n + 3$.

1.3.6 Récurrence forte

Si P est un prédicat défini sur $\mathbb{N}$, il arrive parfois que la justification de $P(n)$ nécessite l'utilisation de tous les $P(k)$ pour $k \in [\![0, n-1]\!]$. On fait alors ce que l'on appelle une récurrence forte qui est fondée sur le résultat suivant.

Corollaire 1.3.19 (Récurrence forte)**.** *Soit P une propriété définie sur $\mathbb{N}$ avec $P(0)$ vraie ainsi que :*

$$\forall n \in \mathbb{N}^* \quad (P(0) \text{ et } P(1) \cdots \text{ et } P(n-1)) \Rightarrow P(n).$$

Alors, la propriété $P(n)$ est vraie pour tout $n \in \mathbb{N}$.

Exercice 1.3.20. Montrer par récurrence que, pour tout entier $n \geq 2$ est divisible par au moins un nombre premier.

Solution. Montrons par récurrence que : $\forall n \geq 2$, n est divisible par au moins un nombre premier.

- **Initialisation :** 2 est divisible par 2 qui est un nombre premier. La propriété à démontrer est donc vraie quand $n = 2$.
- **Hérédité :** Soit $n \in \mathbb{N}$, $n \geq 2$ et supposons que pour tout $k \in [\![2, n]\!]$, il existe au moins un nombre premier qui divise k. Montrons que le nombre $n+1$ admet au moins un diviseur premier. Il y a deux cas possibles :
  1. Soit $n+1$ est premier et donc $(n+1) \mid (n+1)$ et $n+1$ possède un diviseur premier.

2. Soit $n+1$ n'est pas premier. Dans ce cas, il existe un diviseur d de $n+1$, $1 < d < n+1$. Ainsi $d \in [\![2, n]\!]$. Par hypothèse de récurrence, il existe au moins un nombre premier p qui divise d. On a donc $p \mid d$ et $d \mid (n+1)$. Par la propriété de transitivité de la divisibilité, on en déduit que $p \mid (n+1)$ et $n+1$ possède un diviseur premier.

— **Conclusion :** Grâce au principe de récurrence forte, on peut conclure que la propriété est vraie pour tout $n \geq 2$.

Exercice 1.3.21. Soit $(u_n)_{n\in\mathbb{N}}$ une suite réelle. On suppose que $u_0 \geqslant 0$ et que pour tout $n \in \mathbb{N}$: $u_{n+1} \leqslant \sum_{k=0}^{n} u_k$. Montrer que pour tout $n \in \mathbb{N}$: $u_n \leqslant 2^n u_0$.

1.4 Applications

Dans toute cette section E, F, G et H désignent des ensembles quelconques. De plus, on appelle graphe de E vers F une partie quelconque du produit cartésien $E \times F$.

Définition 1.4.1 (Application)**.** Une **application**, ou **fonction**, est un triplet $u = (E, F, \Gamma)$ où Γ est un graphe de E vers F tel que pour tout $x \in E$, il existe un unique $y \in F$ vérifiant $(x, y) \in \Gamma$, ce qui s'écrit encore :

$$\forall x \in E \quad \exists! y \in F \quad (x, y) \in \Gamma.$$

On dit aussi que u est une application de E dans F ou de E vers F.

Avec les notations de la définition précédente :
— E est appelé l'**ensemble de départ** ou ensemble de définition de u ;
— F est l'**ensemble d'arrivée** de u ;
— pour $x \in E$, l'unique élément $y \in F$ tel que $(x, y) \in \Gamma$ s'appelle **image de** x **par** u et se note $u(x)$;
— quand on a $y = u(x)$, on dit aussi que x est un **antécédent** de y ;
— l'ensemble :

$$\{y \in F \mid \exists x \in E \quad y = u(x)\} = \{u(x); x \in E\}$$

est l'**ensemble image** de u, c'est un sous-ensemble de F ;
— Γ, le graphe de u, est égal à $\{(x, u(x)); x \in E\}$;

Comme conséquence de la définition, on déduit que l'égalité de deux applications u et v signifie :
— l'égalité des ensembles de départ de u et de v,
— l'égalité des ensembles d'arrivée de u et de v,
— l'égalité $u(x) = v(x)$ pour tout x de l'ensemble de départ commun.

Notation L'ensemble des applications, ou des fonctions, de E dans F se note $\mathcal{F}(E, F)$ ou encore F^E. Cette dernière notation est justifiée par le fait que, pour E et F finis, on a $\operatorname{card}(F^E) = (\operatorname{card} F)^{\operatorname{card} E}$.

Définition 1.4.2 (Restriction, prolongement, corestriction)**.** Soit u une application de E dans F.
— Si E' est une partie de E, la **restriction** de u à E', notée $u_{|E'}$, est l'application de E' dans F définie par

$$\forall x \in E' \quad u_{|E'}(x) = u(x).$$

— On appelle **prolongement** de u toute application v définie sur un ensemble E_1 contenant E, et vérifiant

$$\forall x \in E \quad v(x) = u(x).$$

— Si F' est une partie de F, la **corestriction** de u à F', notée $u^{|F'}$, est l'application de E dans F' non définie en x tel que $u(x) \notin F'$, et telle que

$$u^{|F'}(x) = u(x) \quad \text{si} \quad u(x) \in F'.$$

1.4.1 Injectivité, surjectivité, bijectivité

Définition 1.4.3. Soit $u \in \mathcal{F}(E, F)$. On dit qu'elle est **injective**, ou que c'est une **injection**, si elle vérifie l'une des trois propriétés équivalentes suivantes.

— Tout élément de F a au plus un antécédent par u.
— Pour tout $y \in F$, l'équation $u(x) = y$ possède au plus une solution.
— On a : $\forall x_1 \in E \quad \forall x_2 \in E \quad u(x_1) = u(x_2) \Rightarrow x_1 = x_2$.

Exemples 1.4.4.
— L'identité de E est évidemment injective.
— Si X est une partie quelconque de $\mathbb{R}$ et si $u : X \longrightarrow \mathbb{R}$ est strictement croissante, alors elle est injective. En effet si $x \neq y$ alors on a par exemple $x < y$, et la stricte croissance de f nous donne $u(x) < u(y)$ et donc $u(x) \neq u(y)$.
Il en est de même si f est strictement décroissante.

Remarque 1.4.5.
— En général, c'est le 3 point que l'on utilise pour prouver l'injectivité.
— Pour $u : X \longrightarrow \mathbb{R}$, fonction réelle d'une variable réelle, l'injectivité se justifie souvent en prouvant que u est strictement monotone.
— Pour prouver que u n'est pas injective, il suffit d'exhiber $x_1 \in E$ et $x_2 \in E$ tels que $x_1 \neq x_2$ et $u(x_1) = u(x_2)$.

Exercice 1.4.6. Soit $A \in \mathcal{P}(E)$ et $u_A \mid \begin{array}{ccc} \mathcal{P}(E) & \longrightarrow & \mathcal{P}(E) \\ X & \longmapsto & X \cap A. \end{array}$

Montrer que si $A \neq E$, alors l'application u_A n'est pas injective.

Solution. Supposons $A \neq E$. Comme $u_A(A) = A = u_A(E)$, il est clair que u_A n'est pas injective.

Définition 1.4.7. Soit $u \in \mathcal{F}(E, F)$. On dit qu'elle est **surjective**, ou que c'est une **surjection**, si elle vérifie l'une des trois propriétés équivalentes suivantes.

— Tout élément de F a au moins un antécédent par u.
— Pour tout $y \in F$, l'équation $u(x) = y$ possède au moins une solution.
— On a : $\forall y \in F \quad \exists x \in E \quad y = u(x)$.

Remarque 1.4.8. Pour prouver que u n'est pas surjective, on utilise, la négation des propriétés précédente : il suffit donc d'exhiber un $y \in F$ qui n'a pas d'antécédent.

Exemple 1.4.9. La fonction sin : $\mathbb{R} \longrightarrow \mathbb{R}$ n'est pas surjective car 2 n'a pas d'antécédent.

Exercice 1.4.10. Soit $A \in \mathcal{P}(E)$ et $u_A \mid \begin{array}{ccc} \mathcal{P}(E) & \longrightarrow & \mathcal{P}(E) \\ X & \longmapsto & X \cap A. \end{array}$

Montrer que l'application u_A est surjective si, et seulement si, $A = E$.

Solution. — Supposons $A = E$. Alors u_A est l'identité de $\mathcal{P}(E)$; elle est donc surjective.
— Supposons $A \neq E$. Alors, pour tout $X \in \mathcal{P}(E)$, on a $u(X) \subset A$ et donc $u(X) \neq E$. Ainsi E n'a pas d'antécédent par u_A, et cette application n'est pas surjective.

Définition 1.4.11. Soit $u \in \mathcal{F}(E, F)$. On dit qu'elle est **bijective**, ou que c'est une **bijection**, si elle vérifie l'une des quatre propriétés équivalentes suivantes.
— L'application u est injective et surjective.
— Tout élément de F a un et un seul antécédent par u.
— Pour tout $y \in F$, l'équation $u(x) = y$ possède une unique solution.
— On a : $\forall y \in F \quad \exists! x \in E \quad y = u(x)$.

Remarque 1.4.12. Pour prouver qu'une application est bijective, le plus élémentaire est de prouver qu'elle est injective et qu'elle est surjective.

1.4.2 Composition d'applications

Définition 1.4.13. Si $u \in \mathcal{F}(E, F)$ et $v \in \mathcal{F}(F, G)$, l'application $x \longmapsto v(u(x))$, de E dans G est appelée **composée des applications** v et u ; on la note $v \circ u$.

Proposition 1.4.14. *Soit trois applications $u \in \mathcal{F}(E, F)$, $v \in \mathcal{F}(F, G)$ et $w \in \mathcal{F}(G, H)$. Alors les applications $w \circ (v \circ u)$ et $(w \circ v) \circ u$ sont égales.*

Remarque 1.4.15. On se réfère couramment à la propriété précédente en parlant de l'**associativité de la composition** des applications.

Exercice 1.4.16. Soit E un ensemble contenant au moins deux éléments. Construire deux applications $u \in E^E$ et $v \in E^E$ tels que $u \circ v \neq v \circ u$.

Proposition 1.4.17. *Soit $u \in \mathcal{F}(E, F)$ et $v \in \mathcal{F}(F, G)$.*

1. *Si u et v sont injectives, alors $v \circ u$ est injective.*
2. *Si u et v sont surjectives, alors $v \circ u$ est surjective.*
3. *Si u et v sont bijectives, alors $v \circ u$ est bijective.*

Démonstration.
1. Supposons u et v injectives. Soit $x, x' \in E$ tel que $v(u(x)) = v(u(x'))$. Grâce à l'injectivité de v on a $u(x) = u(x')$ et l'injectivité de u donne alors $x = x'$, ce qui prouve l'injectivité de $v \circ u$.
2. Supposons u et v surjectives. Soit $z \in G$. Grâce à la surjectivité de v on sait qu'il existe $y \in F$ tel que $z = v(y)$. Comme u est surjective, on peut aussi considérer un $x \in E$ tel que $y = u(x)$. On a alors $z = v(u(x)) = (v \circ u)(x)$, ce qui prouve la surjectivité de $v \circ u$.
3. Conséquence immédiate des précédents.

□

Exercice 1.4.18. Soient E, F, G trois ensembles ; on considère deux applications $f : E \to F$ et $g : F \to G$.

1. Montrer que si $g \circ f$ est injective et si f est surjective, alors g est injective.
2. Montrer que si $g \circ f$ est surjective et si g est injective, alors f est surjective.

Solution. 1. Soient y, y' dans F tels que $g(y) = g(y')$. Comme f est surjective, il existe x, x' dans E tels que $y = f(x)$ et $y' = f(x')$, ce qui donne $g \circ f(x) = g \circ f(x')$ et $x = x'$ puisque $g \circ f$ est injective, donc $y = y'$.

2. Soit $y \in F$. Comme $g \circ f$ est surjective, il existe $x \in E$ tel que $g(y) = (g \circ f)(x) = g(f(x))$. Comme g est injective alors : $y = f(x)$. En conséquence, f est surjective.

Exercice 1.4.19. Soit E un ensemble quelconque. Soit $f : E \longrightarrow E$ une application telle que $f \circ f \circ f = f$. Montrer que f est injective si et seulement si f est surjective.

Solution. — Supposons f est injective. Soit $y \in E$. Posons $w = f(y)$ et $z = f \circ f(y)$. On a : $f(z) = f \circ f \circ f(y) = f(y)$. Or f est injective. Donc $z = y$ et donc : $f \circ f(y) = y$. Ainsi : $f(w) = y$. Donc f est surjective.
— Supposons f est surjective. Soit $x, y \in E$ tel que $f(x) = f(y)$. Puisque f est surjective, il existe $x', y' \in E$ tel que $x = f(x')$ et $y = f(y')$. On a : $f \circ f(x) = f \circ f \circ f(x') = f(x') = x$ et $f \circ f(y) = f \circ f \circ f(y') = f(y') = y$. Or puisque $f(x) = f(y)$ alors $f \circ f(x) = f \circ f(y)$ et donc $x = y$. D'où f est injective.

Exercice 1.4.20. Soit E un ensemble, et $p : E \longrightarrow E$ telle que $p \circ p = p$. Montrer que si p est surjective ou injective, alors $p = Id_E$.

Solution. — Supposons p injective. Soit $x \in E$. Par hypothèse, on a $p(p(x)) = p(x)$. Puisque p est injective, on en déduit que $p(x) = x$. Ainsi, pour tout $x \in E$, $p(x) = x$ et donc $p = Id_E$.
— Supposons p surjective. On se ramène au cas précédent en montrant que p est injective. Soit x et x' deux éléments de E tels que $p(x) = p(x')$. Puisque p est surjective, il existe deux éléments y et y' de E tels que $x = p(y)$ et $x' = p(y')$. Ainsi, $p \circ p(y) = p \circ p(y')$ et par définition de p, on obtient $p(y) = p(y')$. Par suite $x = y$ et p est injective. Donc : $p = Id_E$.

1.4.3 Application réciproque

Définition 1.4.21. Soit u une application bijective de E dans F. Alors l'application de F dans E qui, à tout $y \in F$, associe l'unique $x \in E$ tel que $y = u(x)$, s'appelle application réciproque de u et se note u^{-1}.

Proposition 1.4.22. *Si u est une application bijective de E dans F, alors on a :*

$$u^{-1} \circ u = Id_E \quad et \quad u \circ u^{-1} = Id_F.$$

Démonstration. — Il est clair que $u^{-1} \circ u$ est une application de E dans E. De plus, pour tout $x \in E$, on a $u^{-1}(u(x)) = x$ car $u(x)$ possède, par u, un unique antécédent qui est évidemment x. On en déduit $u^{-1} \circ u = Id_E$.
— De même $u \circ u^{-1} \in F^F$ et, pour tout $y \in F$, on a $u(u^{-1}(y)) = y$ car $u^{-1}(y)$ est par définition l'antécédent de y. On a donc $u \circ u^{-1} = Id_F$.
□

Proposition 1.4.23. *Si $u \in F^E$ et $v \in E^F$ vérifient $u \circ v = Id_F$ et $v \circ u = Id_E$, alors elles sont toutes deux bijectives et réciproques l'une de l'autre.*

Démonstration. — Supposons qu'il existe une application $v \in \mathcal{F}(F, E)$ telle que $v \circ u = Id_E$, et montrons que u est injective. Soit donc $x, x' \in E$ tels que $u(x) = u(x')$; alors on a $v(u(x)) = v(u(x'))$. Comme $v \circ u = Id_E$, on en déduit $x = x'$, ce qui prouve l'injectivité de u.

— L'existence d'une application $v \in \mathcal{F}(F,E)$ telle que $u \circ v = Id_F$ entraîne que u est surjective. En effet, pour tout $y \in F$, si l'on pose $x = v(y)$, alors on a $u(x) = y$ car $u \circ v = Id_F$, ce qui prouve que u est surjective.

Par suite, l'application u est bijective. On peut alors écrire :

$$v = Id_E \circ v = u^{-1} \circ u \circ v = u^{-1} \circ Id_F = u^{-1}.$$

Par symétrie, on en déduit que v est bijective, et que $v^{-1} = u$. □

Remarque 1.4.24 (Méthode pour montrer que u est bijective et donner u^{-1}). Soit $u : E \longrightarrow F$.

— Si l'on exhibe une application $v : F \longrightarrow E$ telle que $u \circ v = Id_F$ et $v \circ u = Id_E$, alors on prouve que u est bijective et que $u^{-1} = v$.

— Souvent, la démonstration de la surjectivité de u mène à expliciter une solution de l'équation $u(x) = y$, de la forme $x = v(y)$. Une méthode efficace de rédaction consiste alors à exhiber directement v puis à prouver $u \circ v = Id_F$ et $v \circ u = Id_E$.

Exercice 1.4.25. Soit E un ensemble et A une partie de E. On pose $B = \complement_E A$. Montrer que $u \mid \begin{array}{ccc} \mathcal{P}(E) & \longrightarrow & \mathcal{P}(A) \times \mathcal{P}(B) \\ X & \longmapsto & (X \cap A, X \cap B) \end{array}$ est bijective.

Solution. Considérons $v : \begin{array}{c} \mathcal{P}(A) \times \mathcal{P}(B) \longrightarrow \mathcal{P}(E) \\ (Y,Z) \longmapsto Y \cup Z. \end{array}$

— Pour tout $X \in \mathcal{P}(E)$, on a :

$$v(u(X)) = v(X \cap A, X \cap B) = (X \cap A) \cup (X \cap B) = X \cap (A \cup B) = X.$$

Comme $v \circ u$ est une application de $\mathcal{P}(E)$ dans lui-même, on a donc $v \circ u = \mathrm{Id}_{\mathcal{P}(E)}$.

— Montrons $u \circ v = \mathrm{Id}_{\mathcal{P}(A)\times\mathcal{P}(B)}$ c'est-à-dire :

$$\forall (Y,Z) \in \mathcal{P}(A) \times \mathcal{P}(B) \quad u(v(Y,Z)) = (Y,Z).$$

Soit donc $(Y,Z) \in \mathcal{P}(A) \times \mathcal{P}(B)$. On a alors :

$$v(Y,Z) \cap A = (Y \cup Z) \cap A = (Y \cap A) \cup (Z \cap A).$$

⋆ Comme $Z \subset B$, on a $Z \cap A \subset (B \cap A) = \varnothing$.

⋆ Comme $Y \subset A$, on a $Y \cap A = Y$.

Ainsi $v(Y,Z) \cap A = Y$. On prouve de même $v(Y,Z) \cap B = Z$, ce qui donne :

$$u(v(Y,Z)) = (v(Y,Z) \cap A, v(Y,Z) \cap B) = (Y,Z).$$

Comme $u \circ v$ est une application de $\mathcal{P}(A) \times \mathcal{P}(B)$ dans lui-même, on en déduit $u \circ v = \mathrm{Id}_{\mathcal{P}(A)\times\mathcal{P}(B)}$. Par suite u est bijective, et v est son application réciproque.

Corollaire 1.4.26.

1. *Si $u \in F^E$ est bijective, alors u^{-1} est bijective et $(u^{-1})^{-1} = u$.*
2. *Si $u \in F^E$ et $v \in G^F$ sont deux applications bijectives, alors $v \circ u$ est une application bijective et $(v \circ u)^{-1} = u^{-1} \circ v^{-1}$.*

Démonstration. 1. Conséquence de la proposition précédente puisque u et u^{-1} vérifient :

$$u^{-1} \circ u = Id_E \quad \text{et} \quad u \circ u^{-1} = Id_F.$$

2. L'associativité de la composition nous donne :

$$(u^{-1}\circ v^{-1})\circ(v\circ u)=u^{-1}\circ(v^{-1}\circ v)\circ u=u^{-1}\circ Id_F\circ u=u^{-1}\circ u=Id_E,$$
$$(v\circ u)\circ(u^{-1}\circ v^{-1})=v\circ(u\circ u^{-1})\circ v^{-1}=v\circ Id_F\circ v^{-1}=v\circ v^{-1}=Id_G.$$

La proposition précédente nous permet alors de conclure.

□

Remarque 1.4.27.
— Soit u une application de E dans E. On dit qu'elle est **involutive**, ou que c'est une d'involution si elle vérifie $u\circ u=Id_E$.
— D'après la proposition précédente, une application involutive est bijective, et elle est sa propre réciproque.

Exemple 1.4.28. L'application $\begin{array}{ccc}\mathcal{P}(E) & \longrightarrow & \mathcal{P}(E)\\ X & \longmapsto & \complement_E X\end{array}$ est involutive et donc bijective.

1.4.4 Images directes, images réciproques

Définition 1.4.29. Soit $u\in F^E$. Si $A\subset E$ et $B\subset F$, on appelle :
— **image (directe)** de A par u, l'ensemble :

$$u(A)=\{y\in F\mid \exists x\in A\quad y=u(x)\}=\{u(x);x\in A\},$$

— **image réciproque** de B par u, l'ensemble :

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

— L'application $u\in\mathcal{F}(E,F)$ est surjective si et seulement si $u(E)=F$.
— L'application $u\in\mathcal{F}(E,F)$ est :
1. injective si, et seulement si, pour tout $y\in F$, l'ensemble $u^{-1}(\{y\})$ a au plus un élément ;
2. surjective si, et seulement si, pour tout$y\in F$, l'ensemble $u^{-1}(\{y\})$ a au moins un élément ;
3. bijective si, et seulement si, pour tout $y\in F$, l'ensemble $u^{-1}(\{y\})$ a exactement un élément.

Proposition 1.4.30. *Si $u\in\mathcal{F}(E,F)$, $B\subset F$ et $B'\subset F$, alors on a :*
1. $B\subset B'\Rightarrow u^{-1}(B)\subset u^{-1}(B')$.
2. $u^{-1}(B\cup B')=u^{-1}(B)\cup u^{-1}(B')$.
3. $u^{-1}(B\cap B')=u^{-1}(B)\cap u^{-1}(B')$.
4. $u^{-1}\left(\complement_F B\right)=\complement_E u^{-1}(B)$.

Démonstration. 1. Supposons $B\subset B'$ et prouvons $u^{-1}(B)\subset u^{-1}(B')$. Soit donc $x\in u^{-1}(B)$. Alors, on a $u(x)\in B$ et donc $u(x)\in B'$, ce qui prouve $x\in u^{-1}(B')$.

2. Établissons $u^{-1}(B\cup B')=u^{-1}(B)\cup u^{-1}(B')$ par double inclusion.
— Supposons $x\in u^{-1}(B\cup B')$. On a donc $u(x)\in B\cup B'$.
⋆ Si $u(x)\in B$, alors $x\in u^{-1}(B)$ et donc $x\in u^{-1}(B)\cup u^{-1}(B')$.
⋆ Si $u(x)\in B'$, alors $x\in u^{-1}(B')$ et donc $x\in u^{-1}(B)\cup u^{-1}(B')$.

On en déduit $x \in u^{-1}(B) \cup u^{-1}(B')$ et donc $u^{-1}(B \cup B') \subset u^{-1}(B) \cup u^{-1}(B')$.

— Réciproquement, on a $B \subset B \cup B'$ et donc $u^{-1}(B) \subset u^{-1}(B \cup B')$ d'après (1.) ; on a de même $u^{-1}(B') \subset u^{-1}(B \cup B')$, et donc : $u^{-1}(B) \cup u^{-1}(B') \subset u^{-1}(B \cup B')$.

3. Démonstration par double inclusion de $u^{-1}(B \cap B') = u^{-1}(B) \cap u^{-1}(B')$.
   — Supposons $x \in u^{-1}(B) \cap u^{-1}(B')$. On a donc $x \in u^{-1}(B)$ et $x \in u^{-1}(B')$, ce qui entraîne $u(x) \in B$ et $u(x) \in B'$, et donc $u(x) \in B \cap B'$ et $x \in u^{-1}(B \cap B')$. On a donc prouvé $u^{-1}(B) \cap u^{-1}(B') \subset u^{-1}(B \cap B')$.
   — Réciproquement, on a $B \cap B' \subset B$ donc, d'après (1.), on a $u^{-1}(B \cap B') \subset u^{-1}(B)$; on a de même $u^{-1}(B \cap B') \subset u^{-1}(B')$; donc $u^{-1}(B \cap B') \subset u^{-1}(B) \cap u^{-1}(B')$.
4. Démonstration similaire par double inclusion pour $u^{-1}\left(\complement_F B\right) = \complement_E u^{-1}(B)$. □

Il existe des résultats analogues (mais pas identiques) pour les images directes. Ces résultats, font l'objet de l'exercice suivant.

Exercice 1.4.31. Soit $u \in \mathcal{F}(E, F)$ ainsi que $A \subset E$ et $A' \subset E$.

1. Montrer que l'on a :
$$A \subset A' \Rightarrow u(A) \subset u(A') \quad \text{et} \quad u(A \cup A') = u(A) \cup u(A').$$
2. Établir $u(A \cap A') \subset u(A) \cap u(A')$.
3. On prend ici
$$\begin{aligned} u : \quad & \mathbb{R} \longrightarrow \mathbb{R} \\ & x \longmapsto x^2. \end{aligned}$$
Exhiber $A \subset \mathbb{R}$ et $A' \subset \mathbb{R}$ tels que $u(A \cap A') \neq u(A) \cap u(A')$.

Solution. 1. Supposons $A \subset A'$ et montrons $u(A) \subset u(A')$. Soit donc $y \in u(A)$. On peut alors trouver $x \in A$ tel que $y = u(x)$. Comme $A \subset A'$, on en déduit $x \in A'$ et donc $y \in u(A')$. Ainsi : $u(A) \subset u(A')$. Montrons $u(A \cup A') = u(A) \cup u(A')$.
   — Comme $A \subset A \cup A'$, la question précédente entraine que $u(A) \subset u(A \cup A')$. On a de même $u(A') \subset u(A \cup A')$; on en déduit $u(A) \cup u(A') \subset u(A \cup A')$.
   — Pour prouver $u(A \cup A') \subset u(A) \cup u(A')$, prenons $y \in u(A \cup A')$. On peut alors trouver $x \in A \cup A'$ tel que $y = u(x)$.
     ⋆ Si $x \in A$, alors $y \in u(A) \subset u(A) \cup u(A')$.
     ⋆ Si $x \in A'$, alors $y \in u(A') \subset u(A) \cup u(A')$. Par suite, on a $y \in u(A) \cup u(A')$, ce qui prouve l'inclusion souhaitée.

   On en déduit l'égalité par double inclusion.
2. Comme $A \cap A' \subset A$ et $A \cap A' \subset A'$, la première question nous donne :
$$u(A \cap A') \subset u(A) \quad \text{et} \quad u(A \cap A') \subset u(A').$$
On en déduit $u(A \cap A') \subset u(A) \cap u(A')$.

3. Posons $A = \mathbb{R}_+^*$ et $A' = \mathbb{R}_-^*$. On a alors :
   — $u(A \cap A') = u(\varnothing) = \varnothing$,
   — $u(A) = \mathbb{R}_+^* = u(A')$ et donc $u(A) \cap u(A') = \mathbb{R}_+^* \neq u(A \cap A')$.
   Ce contre-exemple montre que, dans le cas général, on ne peut pas améliorer l'inclusion de la question précédente.

Exercice 1.4.32. Soit $f : X \to Y$ une application, montrer que, pour tout $A \in \mathcal{P}(X)$, $A \subset f^{-1}(f(A))$ et que, pour tout $B \in \mathcal{P}(Y)$, $f\left(f^{-1}(B)\right) \subset B$.

Solution. — Démontrons l'inclusion $A \subset f^{-1}(f(A))$. Soit $x \in A$, il s'agit de vérifier que $x \in f^{-1}(f(A))$; or x appartient à A, donc son image $f(x)$ par f appartient à l'image $f(A)$ de A et ceci signifie précisément que x appartient à $f^{-1}(f(A))$, ce qui prouve le résultat voulu.
— Vérifions de même l'inclusion $f\left(f^{-1}(B)\right) \subset B$. Soit $y \in f\left(f^{-1}(B)\right)$, il s'agit de vérifier que $y \in B$. Il existe $x \in f^{-1}(B)$ tel que $y = f(x)$; dire que x appartient à $f^{-1}(B)$ signifie que $f(x)$ appartient à B et ceci prouve que $y = f(x) \in B$.

Exercice 1.4.33. Soit $f : X \to Y$ une application, montrer que :

1. f est injective $\iff \forall A \in \mathcal{P}(X), f^{-1}(f(A)) = A$.
2. f est surjective $\iff \forall B \in \mathcal{P}(Y), f\left(f^{-1}(B)\right) = B$.

Solution. 1. Supposons f injective et soit $A \in \mathcal{P}(X)$. D'après l'exercice 1.4.32, on sait que $A \subset f^{-1}(f(A))$. Montrons que $f^{-1}(f(A)) \subset A$. Soit $x \in f^{-1}(f(A))$, c'est-à-dire $f(x) \in f(A)$, il existe donc $y \in A$ tel que $f(x) = f(y)$; d'après l'injectivité de f, on a nécessairement $x = y$, d'où $x \in A$ et ceci prouve le résultat voulu.

   Réciproquement, supposons, que pour tout $A \in \mathcal{P}(X), A = f^{-1}(f(A))$ et montrons que f est injective. Soient $x, y \in X$ tels que $f(x) = f(y)$. Prenons $A = \{x\}$, puis $A = \{y\}$; on obtient

$$\{x\} = f^{-1}(f(\{x\})) \text{ et } \{y\} = f^{-1}(f(\{y\}))$$

   où $f(\{x\}) = \{f(x)\} = \{f(y)\} = f(\{y\})$, d'où $\{x\} = \{y\}$, c'est-à-dire $x = y$ et f est donc injective.

2. Supposons f surjective et soit $B \in \mathcal{P}(Y)$. D'après l'exercice 1.4.32, on a $f\left(f^{-1}(B)\right) \subset B$. Montrons l'inclusion opposée $B \subset f\left(f^{-1}(B)\right)$. Soit $y \in B$, f étant surjective il existe $x \in X$ tel que $y = f(x)$; alors $x \in f^{-1}(B)$, d'où $y = f(x) \in f\left(f^{-1}(B)\right)$ et le résultat voulu.

   Réciproquement, supposons que, pour tout $B \in \mathcal{P}(Y), f\left(f^{-1}(B)\right) = B$; soit $y \in Y$, prenons $B = \{y\}$, alors $f\left(f^{-1}(\{y\})\right) = \{y\}$ et ceci prouve que $f^{-1}(\{y\})$ est non vide, donc f est surjective.

1.4.5 Fonction indicatrice (ou caractéristique)

Définition 1.4.34 (Relation binaire)**.** Soit $A \subset E$. La **fonction indicatrice** de A, ou encore **fonction caractéristique** de A, est la fonction de E dans $\{0, 1\}$, notée 1_A et définie par :

$$1_A(x) = 1 \quad \text{si} \quad x \in A \quad \text{et} \quad 1_A(x) = 0 \quad \text{si} \quad x \notin A.$$

Exercice 1.4.35. Soit E est un ensemble non vide et A une partie donnée de E.

1. Préciser 1_A quand $A = \emptyset$ ou $A = E$.
2. Pour $A \in \mathcal{P}(E)$, exprimer $1_{\bar{A}}$ en fonction de 1_A.
3. Pour $(A, B) \in \mathcal{P}(E)^2$, exprimer $1_{A\cap B}$ en fonction de 1_A et 1_B.
4. Pour $(A, B) \in \mathcal{P}(E)^2$, exprimer $1_{A\cup B}$ et $A_{A\setminus B}$ en fonction de 1_A et 1_B.

Solution.
1. $1_\emptyset = 0$ et $1_E = 1$. où 0 désigne la fonction nulle, c'est-à-dire la fonction qui, à tout élément de E associe le nombre 0 et 1 désigne la fonction qui, à tout élément de E associe le nombre 1.
2. Pour tout élément x de E, on a : $1_A(x)+1_{\bar{A}}(x) = 1$. Par suite, $1_{\bar{A}} = 1-1_A$. (où 1 désigne toujours la fonction qui à tout élément associe 1).
3. Soit $x \in E$. Si $x \in A$ et $x \in B$, alors $1_A(x)1_B(x) = 1 = 1_{A\cap B}(x)$. Si $x \notin A$ ou $x \notin B$, $1_A(x)1_B(x) = 0 = 1_{A\cap B}(x)$. Finalement, $\forall x \in E$, $1_{A\cap B} = 1_A \times 1_B$.
4. On a :

$$1_{A\cup B} = 1 - 1_{\overline{A\cup B}} = 1 - 1_{\bar{A}\cap\bar{B}} = 1 - (1 - 1_A)(1 - 1_B) = 1_A + 1_B - 1_A 1_B.$$

D'autre part,

$$1_{A\setminus B} = \chi_{A\cap\bar{B}} = 1_A \times 1_{\bar{B}} = 1_A(1 - 1_B) = 1_A - 1_A 1_B.$$

Remarque 1.4.36. Si E ne possède qu'un nombre fini d'éléments, alors $\sum_{x\in E} 1_A(x)$ est égal au nombre d'éléments de A, noté $\mathrm{card}(A)$ ou $\mathrm{card}\, A$.

Proposition 1.4.37. *L'application $u : A \longmapsto 1_A$ est une bijection de $\mathcal{P}(E)$ sur l'ensemble $\mathcal{F}\Big(E, \{0,1\}\Big)$.*

Démonstration. Soit

$$\begin{aligned} v : \mathcal{F}(E,\{0,1\}) &\longrightarrow \mathcal{P}(E) \\ f &\longmapsto f^{-1}(\{1\}). \end{aligned}$$

— Par construction, on a : $\forall A \in \mathcal{P}(E) \quad (v\circ u)(A) = v(1_A) = 1_A^{-1}(\{1\}) = A$. On en déduit $v \circ u = \mathrm{Id}_{\mathcal{P}(E)}$.
— Soit $f \in \mathcal{F}(E,\{0,1\})$. Posons $A = v(f) = f^{-1}(\{1\})$. Alors :

$$\forall x \in E \quad f(x) = 1 \Longleftrightarrow x \in A.$$

Comme f ne prend que deux valeurs, on en déduit $u(A) = f$, et donc $u \circ v = \mathrm{Id}_{\mathcal{F}(E\{0,1\})}$.

Les deux relations $v \circ u = \mathrm{Id}_{\mathcal{P}(E)}$ et $u \circ v = \mathrm{Id}_{\mathcal{F}(E,\{0,1\})}$, entraînent que u et bijective. □

1.5 Relations binaires

Définition 1.5.1 (Relation binaire)**.** On appelle **relation binaire** sur un ensemble E toute partie de $E \times E$.

Si $\mathcal{R}$ est une telle relation, la proposition : $(x, y) \in \mathcal{R}$ sera notée de préférence : $x\mathcal{R}y$ pour tous $x, y \in E$ et lue : "x est en relation avec y par $\mathcal{R}$".

Exemples 1.5.2. Vous connaissez certaines relations binaires :

— la relation d'égalité $=$ sur E,
— les relations $\leq$ et $<$ sur $\mathbb{R}$,
— la relation d'inclusion $\subset$ sur $\mathcal{P}(E)$,
— pour tout $\alpha \in \mathbb{R}$, la relation $\equiv [\alpha]$ de congruence modulo α sur $\mathbb{R}$, définie pour tous $x, y \in \mathbb{R}$ par :

$$x \equiv y[\alpha] \quad \Longleftrightarrow \quad \exists k \in \mathbb{Z}, \quad x = y + k\alpha$$

1.5.1 Relation d'équivalence

Définition 1.5.3 (Relation d'équivalence)**.** Une relation binaire $\mathcal{R}$ sur E est une relation d'équivalence sur E si elle est :
— **réflexive** : $\forall x \in E, x\mathcal{R}x$.
— **symétrique** : $\forall (x, y) \in E^2, x\mathcal{R}y \Rightarrow y\mathcal{R}x$.
— **transitive** : $\forall (x, y, z) \in E^3, \Big(x\mathcal{R}y \quad \text{et} \quad y\mathcal{R}z\Big) \Rightarrow x\mathcal{R}z$.

Exemples 1.5.4.
1. Sur tout ensemble E, l'égalité est évidemment une relation d'équivalence.
2. Pour tout $\alpha \in \mathbb{R}$, la relation $\equiv [\alpha]$ de congruence modulo α sur $\mathbb{R}$ est une relation d'équivalence.
   De manière analogue, pour tout $n \in \mathbb{N}$, la relation $\equiv [n]$ de congruence modulo n sur $\mathbb{Z}$ est une relation d'équivalence.

Définition 1.5.5 (Classes d'équivalence, ensemble quotient)**.** Soit $\mathcal{R}$ une relation d'équivalence sur un ensemble E.
— Pour tout $x \in E$, l'ensemble $\{y \in E \mid x\mathcal{R}y\}$ est appelé **classe d'équivalence de** x pour $\mathcal{R}$; cet ensemble est noté $\mathrm{cl}(x)$ voire $\bar{x}$ ou $\dot{x}$.
— Une partie X de E est une **classe d'équivalence** s'il existe un $x \in E$ tel que $X = \mathrm{cl}(x)$; un tel x est alors appelé un **représentant** de X.
— L'ensemble des classes d'équivalences de E pour $\mathcal{R}$ est appelé l'**ensemble quotient** de E par $\mathcal{R}$ et souvent noté $E/\mathcal{R}$.

Remarque 1.5.6. Pour tout élément x de E, on a $x \in \mathrm{cl}(x)$.

Proposition 1.5.7. *Étant donné une relation d'équivalence $\mathcal{R}$ sur un ensemble E, ainsi que deux éléments x et y de E, les propriétés suivantes sont équivalentes :*

(i) $x\mathcal{R}y$.

(ii) $y \in \mathrm{cl}(x)$.

(iii) $\mathrm{cl}(x) = \mathrm{cl}(y)$.

Démonstration.
— Par définition des classes d'équivalence, les propriétés (i) et (ii) sont équivalentes.
— Supposons (i) et prouvons (iii).
  ⋆ Soit $z \in cl(x)$; alors on a $x\mathcal{R}z$ et, comme $x\mathcal{R}y$, par symétrie et transitivité, on en déduit $z\mathcal{R}y$; par suite, on a $\mathrm{ll}(x) \subset \mathrm{cl}(y)$.
  ⋆ Comme x et y jouent des rôles symétriques, on a aussi $\mathrm{cl}(y) \subset \mathrm{cl}(x)$. On en déduit $cl(y) = cl(x)$.
— Réciproquement, supposons (iii) c'est-à-dire $\mathrm{cl}(x) = \mathrm{cl}(y)$. Alors, comme $y \in \mathrm{cl}(y)$, on a $y \in \mathrm{cl}(x)$, ce qui montre (ii).

□

Définition 1.5.8 (Partition d'un ensemble)**.** Une **partition** d'un ensemble E est un ensemble de parties de E, toutes non vides, disjointes deux à deux, et dont la réunion est égale à E ; autrement dit, c'est une partie $\mathcal{U}$ de $\mathbb{P}(E)$ telle que :
- $\forall A \in \mathcal{U},\ A \neq \emptyset$.
- $\forall A \in \mathcal{U}\ \forall B \in \mathcal{U},\ A \neq B \Rightarrow A \cap B = \emptyset$.
- $\cup_{A\in\mathcal{U}} A = E$.

Théorème 1.5.9. *Si $\mathcal{R}$ est une relation d'équivalence sur un ensemble E, alors ses classes d'équivalence forment une partition de E.*

Démonstration. Soit $\mathcal{R}$ une relation d'équivalence sur E. Pour tout $x \in E$, notons $\mathrm{cl}(x)$ la classe d'équivalence de x pour $\mathcal{R}$. Pour tout $x \in E$: $x\mathcal{R}x$ par réflexivité, donc $x \in \mathrm{cl}(x)$, donc $\mathrm{cl}(x)$ est non vide. Rappelons à cette occasion que cette condition différencie les partitions des recouvrements disjoints.
- Clairement : $E = \bigcup_{x\in E} \mathrm{cl}(x)$ car $x \in \mathrm{cl}(x)$ pour tout $x \in E$.
- Soient $x, y \in E$. Pour montrer que $\mathrm{cl}(x)$ et $\mathrm{cl}(y)$ sont égales ou disjointes, supposons-les NON disjointes et montrons qu'elles sont égales. Par hypothèse, nous pouvons nous donner un élément z commun à $\mathrm{cl}(x)$ et $\mathrm{cl}(y)$. Par symétrie des rôles de x et y, il nous suffit de montrer que $\mathrm{cl}(x) \subset \mathrm{cl}(y)$. Soit $t \in \mathrm{cl}(x)$.
  - ⋆ D'abord : $z \in \mathrm{cl}(y)$, donc $y\mathcal{R}z$.
  - ⋆ Ensuite : $z \in \mathrm{cl}(x)$, donc $x\mathcal{R}z$, puis $z\mathcal{R}x$.
  - ⋆ Enfin : $t \in \mathrm{cl}(x)$, donc $x\mathcal{R}t$.

  Conclusion : $y\mathcal{R}t$ par transitivité.

Par suite, les classes d'équivalence forment une partition de E. □

Exercice 1.5.10. On définit sur $\mathbb{R}$ la relation suivante :

$$x\mathcal{R}y \Leftrightarrow x^2 - y^2 = x - y.$$

1. Montrer que $\mathcal{R}$ est une relation d'équivalence.
2. Calculer la classe d'équivalence d'un élément $x \in \mathbb{R}$. Combien y-a-t-il d'éléments dans cette classe ?

Solution. 1. Il suffit de remarquer que $x\mathcal{R}y \Leftrightarrow x^2 - x = y^2 - y \Leftrightarrow f(x) = f(y)$ avec $f : x \mapsto x^2 - x$. II est alors aisé de vérifier en appliquant la définition que $\mathcal{R}$ est une relation d'équivalence, c'est-à-dire qu'elle est réflexive, symétrique et transitive.

2. Soit $x \in \mathbb{R}$. On cherche les éléments y de $\mathbb{R}$ tels que $y\mathcal{R}x$. On doit donc résoudre l'équation (en y) $y^2 - x^2 = y - x$. Elle se factorise en

$$(y - x)(y + x) - (y - x) = 0 \Longleftrightarrow (y - x) \times (y + x - 1) = 0.$$

Ses solutions sont $y = x$ et $y = 1 - x$. La classe de x est donc égale à $\{x, 1-x\}$. Elle est constituée de deux éléments, sauf si $x = 1-x \Longleftrightarrow x = 1/2$. Dans ce cas, elle est égale à $\{1/2\}$.

Exercice 1.5.11. On définit sur $\mathbb{R}$ la relation suivante :

$$x\mathcal{R}y \Leftrightarrow x^3 - y^3 = 3(x - y).$$

1. Montrer que $\mathcal{R}$ est une relation d'équivalence.
2. Pour tout $x \in \mathbb{R}$, déterminer le nombre d'éléments de la classe de x.

1.5.2 Relation d'ordre

Définition 1.5.12 (Relation d'ordre)**.** Une relation binaire $\mathcal{R}$ sur E est une **relation d'ordre** sur E si elle est :

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

Le couple $(E, \mathcal{R})$ est alors appelé **ensemble ordonné**.

Exemple 1.5.13. Les relations $\leq$ sur $\mathbb{R}$ et $\mathbb{R}^{\mathbb{R}}$ sont des relations d'ordre, ainsi que la relation d'inclusion $\subset$ sur $\mathcal{P}(E)$.

Exercice 1.5.14. Montrer que la relation de divisibilité définie dans $\mathbb{N}$ par :

$$x \mid y \Longleftrightarrow \exists k \in \mathbb{N} \quad y = kx$$

est une relation d'ordre.

Solution.
- Réflexivité : Pour tout $n \in \mathbb{N}$: $n \mid n$ car $n = 1 \times n$.
- Transitivité : Soient $n, n', n'' \in \mathbb{N}$ des entiers pour lesquels $n \mid n'$ et $n' \mid n''$. Aussitôt $n' = kn$ et $n'' = k'n'$ pour certains $k, k' \in \mathbb{N}$, donc : $n'' = k'n' = kk'n$ et $kk' \in \mathbb{N}$ donc $n \mid n''$.
- Antisymétrie : Soient $n, n' \in \mathbb{N}$ des entiers pour lesquels $n \mid n'$ et $n' \mid n$. Aussitôt $n' = kn$ et $n = k'n'$ pour certains $k, k' \in \mathbb{N}$, donc $n = k'n' = kk'n$.
  - Si $n = 0$: $n' = kn = 0$, donc $n = n'$.
  - Si $n \neq 0$: $kk' = 1$, or k et k' sont des ENTIERS NATURELS, donc $k = k' = 1$, donc $n = k'n' = n'$.

Soit $(E, \leq)$ un ensemble ordonné. On dit que deux éléments x et y de E sont comparables pour la relation $\leq$ si l'on a $x \leq y$ ou $y \leq x$.

Exemple 1.5.15.
- Dans $\mathbb{R}$ muni de son ordre usuel deux éléments quelconques sont comparables.
- Dans $\mathbb{N}$ muni de la divisibilité, les éléments 2 et 3 ne sont pas comparables.

Définition 1.5.16. La relation d'ordre $\leq$ définit un ordre total sur E si deux éléments quelconques de E sont comparables pour $\leq$, c'est-à-dire si :

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

Dans le cas contraire on dit que c'est un ordre partiel.

Théorème 1.5.17. *Toute partie finie et non vide d'un ensemble totalement ordonné admet un plus grand et un plus petit élément.*

Théorème 1.5.18 (Deux propriétés de $\mathbb{N}$)**.**
1. *Toute partie non vide et majorée de $\mathbb{N}$ admet un plus grand élément.*
2. *Toute partie non vide de $\mathbb{N}$ admet un plus petit élément.*

1.6 Ensembles finis

Définition 1.6.1. Un ensemble E est **fini** s'il existe $n \in \mathbb{N}$ et une bijection de $[\![1, n]\!]$ sur E. Nous admettrons qu'un tel entier n est alors unique. Cet entier n est appelé **cardinal** de E, ou encore nombre d'éléments de E. On le note le plus souvent $\text{card}(E)$, mais aussi $|E|$ ou encore $\#E$.

Remarque 1.6.2.
1. Lorsque $n = 0$, l'intervalle $[\![1, n]\!]$ est vide, et il en est donc de même de E. Ainsi, l'ensemble vide est le seul ensemble de cardinal 0.
2. Si E est un ensemble fini de cardinal $n > 1$, alors une bijection $i \longmapsto a_i$ de $[\![1, n]\!]$ sur E permet de numéroter les éléments de E et d'écrire :
$$E = \{a_1, a_2, \cdots, a_n\}.$$
3. On appelle **singleton**, tout ensemble de cardinal 1.
4. Un ensemble est dit **infini** s'il n'est pas fini.

Proposition 1.6.3. *Soient E et F des ensembles finis non vides. Il existe une bijection de E sur F si et seulement si on a* $\operatorname{card} E = \operatorname{card} F$.

Démonstration. Posons $n = \operatorname{card} E$ et $k = \operatorname{card} F$. Par définition, il existe des bijections $g : E \longrightarrow \{1, \cdots, n\}$ et $h : F \longrightarrow \{1, \cdots, k\}$. Supposons qu'il existe une bijection $f : E \longrightarrow F$. L'application $h \circ f$ est alors une bijection de E sur $\{1, \cdots, k\}$. On en déduit que $n = k$. Réciproquement, si $n = k$, alors $h^{-1} \circ g$ est une bijection de E sur F. □

Maintenant, on donne les propriétés essentielles du cardinal d'un ensemble fini.

Proposition 1.6.4.
1. *Si E, F sont deux ensembles finis disjoints, alors $E \cup F$ est fini et :*
$$\operatorname{card}(E \cup F) = \operatorname{card} E + \operatorname{card} F.$$
2. *Si F est une partie d'un ensemble fini E, alors :*
$$\operatorname{card}(E \setminus F) = \operatorname{card} E - \operatorname{card} F.$$
3. *Toute partie F d'un ensemble fini E est finie et* $\operatorname{card} F \leq \operatorname{card} E$ *: L'égalité est réalisée si, et seulement si, $F = E$.*
4. *Si E, F sont deux ensembles finis, alors $E \cup F$ est fini et :*
$$\operatorname{card}(E \cup F) = \operatorname{card} E + \operatorname{card} F - \operatorname{card}(E \cap F).$$

Démonstration.
1. On désigne par n le cardinal de E et par m celui de F. On dispose donc d'une bijection f de E sur $E_n = \{1, \cdots, n\}$ et d'une bijection g de F sur $E_m = \{1, \cdots, m\}$. L'application h définie sur $E \cup F$ par :
$$h(x) = \begin{cases} f(x) \text{ si } x \in E \\ n + g(x) \text{ si } x \in F \end{cases}$$
réalise alors une bijection de $E \cup F$ sur $E_{n+m} = \{1, \cdots, n + m\}$. En effet, elle est bien définie puisque E et F sont disjoints et pour tout $k \in E_{n+m}$ il existe un unique $x \in E \cup F$ tel que $k = h(x)$, cet élément étant $x = f^{-1}(k)$ si $1 \leq k \leq n$ ou $x = g^{-1}(k - n)$ si $n + 1 \leq k \leq m$. L'ensemble $E \cup F$ est donc fini de cardinal $n + m = \operatorname{card} E + \operatorname{card} F$.
2. Avec la partition $E = (E \setminus F) \cup F$, on déduit que $\operatorname{card} E = \operatorname{card} F + \operatorname{card}(E \setminus F)$.

3. De l'égalité précédente, on déduit que $\operatorname{card} F \leq \operatorname{card} E$. Supposons que $\operatorname{crad} E = \operatorname{card} F$. Si $F \neq E$, il existe $x \in E \setminus F$ et de l'inclusion $F \cup \{x\} \subset E$ avec $F \cap \{x\} = \emptyset$, on déduit $\operatorname{card} F + 1 \leq \operatorname{card} E$, ce qui contredit l'égalité $\operatorname{card} E = \operatorname{card} F$. On a donc $F = E$. La réciproque est évidente.
4. Des partitions :

$$E \cup F = (E \setminus F) \cup F \text{ et } E = (E \setminus F) \cup (E \cap F)$$

on déduit que :

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

et :

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

Ce qui donne

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

□

Remarque 1.6.5. Pour démontrer l'égalité entre deux ensembles finis, il suffit de montrer une inclusion et l'égalité des cardinaux.

Exemple 1.6.6. On peut montrer par l'absurde que l'ensemble $\mathbb{N}$ est infini. Supposons-le fini et appelons n son cardinal. Comme $[\![1, n+1]\!] \subset \mathbb{N}$, d'après la proposition précédente, on a $n + 1 \leq n$, ce qui est impossible. Ainsi $\mathbb{N}$ est infini.

Remarque 1.6.7.
- Si E est un ensemble **fini** et si A est une partie de E telle que $A \neq E$, alors d'après ce qui précéde on a $\operatorname{card} A < \operatorname{card} E$. Puisque deux ensembles en bijection ont même cardinal, on en déduit qu'il n'existe pas de bijection de E sur A.
- Il n'en va pas de même lorsque l'ensemble E est infini. $\mathbb{N}^*$ est un sous-ensemble strict de $\mathbb{N}$. Pourtant, l'application $f : \mathbb{N} \longrightarrow \mathbb{N}^*$ qui à n associe $n + 1$ est bijective.

1.6.1 Applications entre ensembles finis

Proposition 1.6.8. *Soit E un ensemble fini non vide, F un ensemble quelconque et $u \in F^E$.*

1. *L'ensemble $u(E)$ est fini et $\operatorname{card} u(E) \leq \operatorname{card}(E)$.*
2. *On a $\operatorname{card} u(E) = \operatorname{card}(E)$ si, et seulement si, u est injective.*

Démonstration. Commencons par prouver que si u est injective, alors $\operatorname{card} u(E) = \operatorname{card}(E)$. Supposons u injective. Alors, l'application

$$\begin{aligned} u : E &\longrightarrow u(E) \\ x &\longmapsto u(x) \end{aligned}$$

est bijective, car
- injective comme restriction d'une injection,
- surjective puisque l'on a restreint l'ensemble d'arrivée à l'image.

Par suite, si u est injective, alors l'ensemble $u(E)$ est fini et $\operatorname{card} u(E) = \operatorname{card}(E)$.

Montrons par récurrence sur $n > 1$ la propriété H_n : "si E est un ensemble de cardinal n et u une application définie sur E, alors $u(E)$ est fini et $\operatorname{card} u(E) \leq \operatorname{card} E$; et s'il y égalité, alors u est injective".

— **Initialisation :** Si $n = 1$, alors il existe a tel que $E = \{a\}$. On alors $u(E) = \{u(a)\}$, donc $\operatorname{card} u(E) = 1 = \operatorname{card} E$ et u est injective, ce qui prouve que H_1 est vraie.

— **Hérédité :** Soit $n > 2$ tel que H_{n-1} soit vraie. Considérons un ensemble E de cardinal n.

(i) Si u est injective, alors d'après le premier point, on a $\operatorname{card} u(E) = \operatorname{card} E$.

(ii) Si u n'est pas injective, alors il existe deux éléments a et b de E distincts tels que $u(a) = u(b)$. Ainsi, on a $u(E) = u(E \setminus \{b\})$ avec $\operatorname{card}(E \setminus \{b\}) = n - 1$. D'après l'hypothèse de récurrence, $u(E \setminus \{b\})$ est fini et $\operatorname{card} u(E \setminus \{b\}) \leq n - 1$. Ainsi $u(E)$ est fini et $\operatorname{card} u(E) \leq n - 1 < \operatorname{card} E$. On a donc $\operatorname{card} u(E) \leq \operatorname{card} E$, avec égalité si, et seulement si, u est injective. La propriété est vraie au rang n, ce qui termine la démonstration par récurrence.

□

Remarque 1.6.9. De cette proposition, on déduit les résultats suivants.

1. S'il existe une application u surjective d'un ensemble fini E dans un ensemble fini F, alors on a $\operatorname{card} F \leq \operatorname{card} E$ puisqu'alors $F = u(E)$.
2. S'il existe une application injective u d'un ensemble fini E dans un ensemble fini F, on a alors $\operatorname{card} E \leq \operatorname{card} F$. En effet, dans ce cas, $\operatorname{card} E = \operatorname{card} u(E) \leq \operatorname{card} F$.
3. Il existe une bijection de E sur F si et seulement si on a $\operatorname{card} E = \operatorname{card} F$.

Théorème 1.6.10. *Soit E et F deux ensembles finis non vides, de même cardinal, ainsi que u une application de E dans F. Il y équivalence entre :*

(i) *u est injective.*

(ii) *u est surjective.*

(iii) *u est bijective.*

Démonstration. Il est clair que $(iii) \Longrightarrow (i)$ et $(iii) \Longrightarrow (ii)$.

— Montrons $(i) \Longrightarrow (iii)$. Supposons u injective. D'après la proposition 1.6.8, on a $\operatorname{card}(u(E)) = \operatorname{card}(E)$ et donc $\operatorname{card}(u(E)) = \operatorname{card}(F)$. Comme $u(E) \subset F$, d'après la proposition 1.6.4, on a $u(E) = F$ et l'application u est donc surjective.

— Montrons $(ii) \Longrightarrow (iii)$. Supposons u surjective. Alors on a $\operatorname{card}(u(E)) = \operatorname{card} F$ et donc $\operatorname{card}(u(E)) = \operatorname{card}(E)$. D'après la proposition 1.6.8, cela implique que u est injective.

□

Corollaire 1.6.11. *Si E est un ensemble fini et u une application de E dans E, alors les trois propriétés : u est bijective, u est injective et u est surjective sont équivalentes.*

1.7 Compléments

Définition 1.7.1 (Équipotence)**.** On dit que F est équipotent à E s'il existe une bijection de E sur F.

Explication : L'existence d'une bijection de E sur F nous garantit qu'on peut faire se correspondre parfaitement les éléments de E et les éléments de F, associer à tout élément de E un et un seul élément de F et vice versa. Dire que F est équipotent à E revient ainsi à dire que F a exactement le même nombre d'éléments que E.

Définition 1.7.2 (Notion d'ensemble dénombrable)**.** Un ensemble E est dit dénombrable s'il est en bijection avec $\mathbb{N}$ (ou, ce qui revient au même, s'il est équipotent à $\mathbb{N}$).

Passons à un résultat qui nous permettra de remplaçer si besoin l'hypothèse "il existe une injection de E dans F" par "il existe une surjection de F dans E".

Proposition 1.7.3. *Soient E et F des ensembles non vides. Il existe une injection de E dans F si et seulement s'il existe une surjection de F sur E.*

Démonstration. Supposons qu'il existe une injection f de E dans F. Chaque élément de $f(E)$ a donc exactement un antécédent par f. Soit $x_0 \in E$. Soit $g : F \longrightarrow E$ l'application définie ainsi : si $y \in f(E)$, alors $g(y)$ est l'unique antécédent de y par f ; si $y \notin f(E)$, $g(y) = x_0$. L'application g est bien définie. Montrons qu'elle est surjective. Soit $x \in E$. Soit y son image par f. Par définition, $y \in f(E)$, donc $g(y)$ est l'unique antécédent de y par f, donc $g(y) = x$. Donc x a au moins un antécédent par g, donc g est surjective. Donc il existe une surjection de F vers E.

Réciproquement, supposons qu'il existe une surjection g de F sur E. Chaque élément de E a donc au moins un antécédent par g. Pour chaque élément x de E, choisissons un de ses antécédents par g (peu importe lequel), et appellons-le $f(x)$. Cela définit une application $f : E \longrightarrow F$. Montrons que cette application est injective. Soit $x \in E$. Par définition, $f(x)$ est un antécédent de x par g, donc $g(f(x)) = x$, donc $g \circ f = Id_E$, donc $g \circ f$ est bijective, donc injective. Donc f est injective. Donc il existe une injection de E dans F. □

Voici un théorème qui va aider à prouver que deux ensembles sont équipotents.

Théorème 1.7.4 (Cantor-Bernstein)**.** *S'il existe une injection de E dans F et une injection de F dans E, E et F sont équipotents.*

Ce théorème revêt une grande importance en mathématiques. On l'utilise par exemple pour prouver qu'il y a autant d'éléments dans $\mathbb{N}$ que dans $\mathbb{Q}$.

Exercice 1.7.5. On souhaite démontrer le théorème de Théorème de Cantor-Bernstein. Soient E et F deux ensembles. On suppose qu'il existe une application f injective de E dans F, et une application g injective de F dans E.

1. On définit $\varphi : \mathcal{P}(E) \longrightarrow \mathcal{P}(E)$ par :

$$\varphi(A) = \complement_E\Big(g\big(\complement_F f(A)\big)\Big),$$

À l'aide du Théorème de point fixe de Knaster-Tarski, montrer que φ admet un point fixe $M \in \mathcal{P}(E)$, qu'on se donne pour la suite de cet exercice.

2. Montrer que f définit par restriction et corestriction une application $f_1 : M \longrightarrow f(M)$, et que f_1 est bijective.
3. Soit $N = \complement_F f(M)$.
   (a) Décrire g(N).
   (b) Montrer que g définit par restriction et corestriction une application $g_1 : N \longrightarrow \complement_E M$, et que g_1 est une bijection.
4. Construire à l'aide de f_1 et g_1 une bijection $h : E \longrightarrow F$.

On donne ci-dessous des exemples d'équipotence dont certains sont vraiment surprenants au premier abord. Nous allons notamment voir que deux ensembles peuvent être équipotents alors que l'un d'entre eux est inclus **STRICTEMENT** dans l'autre.

Exemple 1.7.6. $\mathbb{R}$ et $]-\frac{\pi}{2}, \frac{\pi}{2}[$ sont équipotents alors qu'ils n'ont pas la même longueur. Longueur et nombre de points n'ont donc aucun rapport ! En effet, tout simplement, la fonction tangente est bijective de $]-\frac{\pi}{2}, \frac{\pi}{2}[$ sur $\mathbb{R}$.

Exemple 1.7.7. $\mathbb{N}$ et $\mathbb{Z}$ sont équipotents. En effet, l'application f de $\mathbb{N}$ dans $\mathbb{Z}$ définie par

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

est bijective.

Exemple 1.7.8. $\mathbb{N}$ et $\mathbb{N}^2$ sont équipotents. Pour cela, il suffit de montrer que l'application $g : (p, q) \longmapsto 2^p(2q + 1)$ est bijective de $\mathbb{N}^2$ sur $\mathbb{N}^*$. À condition de composer ensuite par la bijection $n \longmapsto n - 1$ de $\mathbb{N}^*$ sur $\mathbb{N}$, on aura bien obtenu une bijection de $\mathbb{N}^2$ sur $\mathbb{N}$.

Remarque 1.7.9. On peut aussi raisonner de la manière suivante : L'application $n \longmapsto (n, 0)$ est une injection de $\mathbb{N}$ dans $\mathbb{N}^2$ et, 2 et 3 étant des nombres premiers, l'application $(k, \ell) \longmapsto 2^k 3^\ell$ est une injection de $\mathbb{N}^2$ dans $\mathbb{N}$. Donc on déduit du théorème de Cantor-Bernstein que $\mathbb{N}$ et $\mathbb{N}^2$ sont équipotents. Un raisonnement analogue, utilisant n nombres premiers $p_1, \cdots, p_n$ deux à deux distincts, permet de prouver que, pour tout entier $n \geq 2$, les ensembles $\mathbb{N}$ et $\mathbb{N}^n$ sont équipotents.

Exemple 1.7.10. $\mathbb{N}$ et $\mathbb{Q}$ sont équipotents. Nous conservons dans cet exemple les notations f et g des deux exemples précédents. Comme l'application $n \longmapsto n$ est injective de $\mathbb{N}$ dans $\mathbb{Q}$, alors, en vertu du théorème de Cantor-Bernstein, il nous suffit dès lors d'exhiber une injection de $\mathbb{Q}$ dans $\mathbb{N}$ pour montrer que $\mathbb{N}$ et $\mathbb{Q}$ sont équipotents. Tout d'abord, on rappelle que tout rationnel r s'écrit de façon unique comme fraction réduite $r = p/q$ où $q \geq 1$ et $p \wedge q = 1$. Ainsi, l'application qui à $r = \frac{p}{q} \in \mathbb{Q}$ associe $(p, q) \in \mathbb{Z} \times \mathbb{N}^*$ est bien définie et injective. Nous disposons donc d'une injection h de $\mathbb{Q}$ dans $\mathbb{Z} \times \mathbb{N}^*$. De plus, l'application $n \longmapsto n - 1$ est très clairement une bijection de $\mathbb{N}^*$ sur $\mathbb{N}$. Comme par ailleurs f est bijective de $\mathbb{N}$ sur $\mathbb{Z}$, l'application produit $(m, n) \longmapsto (f^{-1}(m), n - 1)$ est une bijection de $\mathbb{Z} \times \mathbb{N}^*$ sur $\mathbb{N}^2$ que nous noterons i. Finalement, l'application $h \circ i \circ g$ est injective de $\mathbb{Q}$ dans $\mathbb{N}$ par composition.

Exercice 1.7.11 (d'approfondissement). Montrer que $\mathbb{R}$ et $\mathbb{Q}$ ne sont pas équipotents.

▷ Il y a donc infiniment plus d'éléments dans $\mathbb{R}$ que dans $\mathbb{Q}$, donc a fortiori infiniment plus d'irrationnels que de rationnels.

Solution. Parce que $\mathbb{N}$ et $\mathbb{Q}$ sont équipotents, montrer que $\mathbb{R}$ et $\mathbb{Q}$ ne le sont pas revient à montrer que $\mathbb{R}$ et $\mathbb{N}$ ne le sont pas non plus. Et pour montrer qu'il n'existe pas de bijection de $\mathbb{N}$ sur $\mathbb{R}$, il suffit de prouver qu'aucune application de $\mathbb{N}$ dans $\mathbb{R}$ ne peut être surjective.

Nous terminerons ce chapitre par un résultat d'une portée épistémologique et historique considérable.

Théorème 1.7.12 (Cantor)**.** *Il n'existe pas de surjection de E sur $\mathcal{P}(E)$.*

Dans la mesure où l'application $x \longmapsto \{x\}$ est injective de E dans $\mathcal{P}(E)$, le théorème de Cantor montre au fond que E est toujours strictement plus petit que $\mathcal{P}(E)$ en termes d'équipotence. Il en découle un procédé de construction simple d'infinis de tailles **DIFFÉRENTES** toujours plus grandes :

$$\mathbb{N}, \mathcal{P}(\mathbb{N}), \mathcal{P}\Big(\mathcal{P}(\mathbb{N})\Big), \mathcal{P}\left(\mathcal{P}\Big(\mathcal{P}(\mathbb{N})\Big)\right) \cdots .$$

Il est d'ailleurs possible de prouver que $\mathcal{P}(\mathbb{N})$ et $\mathbb{R}$ sont équipotents.

— À la fin du XIXème siècle, Cantor se demande s'il existe ou non entre $\mathbb{N}$ et $\mathbb{R}$ (ou $\mathcal{P}(\mathbb{N})$) un infini de taille intermédiaire mais n'obtient aucun résultat ni dans un sens ni dans l'autre. L'énoncé selon lequel il **N**'y a **PAS** de tel infini intermédiaire s'appelle depuis l'**hypothèse du continu**.

— En 1963, Paul Cohen montre que l'hypothèse du continu n'est pas démontrable dans la théorie usuelle dite ZFC (pour Zermelo-Frenkel + axiome du choix), qui est le cadre admis par la plupart des mathématiciens. L'hypothèse du continu est donc un de ces énoncés qu'on dit **indécidables**, impossible à prouver, impossible à réfuter.

CHAPITRE 2

Polynômes et fractions rationnelles

Sommaire

2.1 Polynômes à une indéterminée sur le corps $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$

2.1.1 Construction des polynômes

Dans tout ce qui suit, $\mathbb{K}$ désigne le corps $\mathbb{R}$ ou $\mathbb{C}$.

Définition 2.1.1 (Polynôme à une indéterminée a coefficients dans $\mathbb{K}$)**.** On appelle polynôme (à une indéterminée) à coefficients dans $\mathbb{K}$ toute suite presque nulle d'éléments de $\mathbb{K}$, i.e. toute suite $(a_k)_{k\in\mathbb{N}}$ d'éléments de $\mathbb{K}$ dont tous les éléments sont nuls à partir d'un certain rang. Pour $k \in \mathbb{N}$, le coefficient a_k est appelé le coefficient de degré k du polynôme.
L'ensemble des polynômes à coefficients dans $\mathbb{K}$ est noté $\mathbb{K}[X]$ si on choisit de noter X l'indéterminée.

Un polynôme est donc une **suite** de la forme $(a_0, a_1, \cdots, a_n, \cdots)$ à coefficients dans $\mathbb{K}$.
► Avec cette présentation des polynômes, on peut immédiatement énoncer le résultat suivant :

Théorème 2.1.2. *Deux polynômes sont égaux si et seulement si ils ont les mêmes coefficients.*

Définition 2.1.3 (Polynôme constant, polynôme nul)**.** On appelle polynôme constant de $\mathbb{K}[X]$ tout polynôme $(\lambda, 0, 0, \cdots)$ avec $\lambda \in \mathbb{K}$. Un tel polynôme sera simplement noté λ.
Avec cette notation, le polynôme 0 est le polynôme nul.

Définition 2.1.4 (Degré d'un polynôme, coefficient dominant, polynôme unitaire)**.**
- Soit $P = (a_k)_{k\in\mathbb{N}} \in \mathbb{K}[X]$ un polynôme **non nul**. Le plus grand indice k pour lequel : $a_k \neq 0$ est appelé degré de P et noté $\deg(P)$, i.e. $\deg(P) = \max\{k \in \mathbb{N} | a_k \neq 0\}$.
  Le coefficient de degré $\deg(P)$ de P est appelé son coefficient dominant et est noté $\mathrm{dom}(P)$. S'il est égale à 1, on dit que P est **unitaire**.
- Par convention, le polynôme nul est de degré $-\infty$: $\deg(0) = -\infty$.

Remarque 2.1.5. On définit aussi la valuation de P comme étant le degré minimum dans P : $\mathrm{val}(P) = \min\{k \in \mathbb{N} |\, a_k \neq 0\}$.
Par convention la valuation du polynôme nul est $+\infty$.

Notation : L'ensemble des polynômes à coefficients dans $\mathbb{K}$, de degré inférieur ou égal à n, se note $\mathbb{K}_n[X]$:

$$\mathbb{K}_n[X] = \{\sum_{k=0}^{n} a_k X^k,\ (a_0, \cdots, a_n) \in \mathbb{K}^{n+1}\}.$$

Ainsi, $\mathbb{R}_2[X] = \{aX^2 + bX + c,\ (a, b, c) \in \mathbb{R}^3\}$. Un élément $aX^2 + bX + c$ de $\mathbb{R}_2[X]$ est de degré 2 si $a \neq 0$, de degré 1 si $a = 0$ et $b \neq 0$, de degré 0 si $a = b = 0$ et $c \neq 0$ et de degré $-\infty$ si $a = b = c = 0$. $\mathbb{K}_0[X]$ est l'ensemble des polynômes constants. Il est constitué des polynômes constants non nuls qui sont les polynômes de degré 0 et du polynôme nul qui est de degré $-\infty$.

On définit maintenant dans $\mathbb{K}[X]$ deux lois internes d'addition et de produit. Nous aimerons pouvoir écrire ceci :

$$\Big(\sum_{k=0}^{n} a_k X^k\Big) + \Big(\sum_{k=0}^{n} b_k X^k\Big) = \sum_{k=0}^{n} (a_k + b_k) X^k$$

et

$$\begin{aligned}\Big(\sum_{i=0}^{n} a_i X^i\Big) \times \Big(\sum_{j=0}^{n} b_j X^j\Big) &= \sum_{0\le i,j\le n} a_i b_j X^{i+j} \\ &= \sum_{k=0}^{2n} \overbrace{\sum_{\substack{0\le i,j\le n \\ i+j=k}} a_i b_j X^k}^{\text{On regroupe les termes de même degré } k} \\ &= \sum_{k=0}^{2n} \overbrace{\sum_{i=0}^{k} a_i b_{k-i} X^k}^{\text{On élimine } j \text{ via la relation } j=k-i}.\end{aligned}$$

Définition 2.1.6. Soient $P = (a_k)_{k\in\mathbb{N}}, Q = (b_k)_{k\in\mathbb{N}} \in \mathbb{K}[X]$.

— On appelle somme de P et Q la suite $(a_k + b_k)_{k\in\mathbb{N}}$, notée $P + Q$. Il s'agit bien d'un polynôme.

— On appelle produit de P et Q la suite $\Big(\sum_{i=0}^{k} a_i b_{k-i}\Big)_{k\in\mathbb{N}}$, notée $P \times Q$ ou PQ. Il s'agit bien d'un polynôme.
En particulier, pour tout $\lambda \in \mathbb{K}$, λP est le polynôme $(\lambda a_k)_{k\in\mathbb{N}}$.

Le triplet $(\mathbb{K}[X], +, \times)$ est alors un anneau commutatif d'élément neutre le polynôme nul 0 pour $+$ et le polynôme constant 1 pour $\times$.

Démonstration. Fixons une fois pour toutes $P = (a_k)_{k\in\mathbb{N}}, Q = (b_k)_{k\in\mathbb{N}}, R = (c_k)_{k\in\mathbb{N}} \in \mathbb{K}[X]$.

— **Lois internes** : Il s'agit de vérifier que la somme et le produit de deux polynômes sont bien des polynômes, i.e. des suites **presques nulles**. Notons N un rang à partir duquel : $a_k = b_k = 0$. Alors : $a_k + b_k = 0$ pour $k \ge N$, donc $P + Q$ est bien un polynôme. D'autre part, pour tout $k \ge 2N$, on a :

$$\sum_{i=0}^{k} a_i b_{k-i} = \sum_{i=0}^{N-1} a_i \underbrace{b_{k-i}}_{=0 \text{ car } k-i>k-N\ge N} + \sum_{i=N}^{k} \underbrace{a_i}_{=0} b_{k-i} = 0.$$

D'où PQ est un polynôme.

— **Multiplication par un scalaire :** Soit $\lambda \in \mathbb{K}$. Pour tout $k \in \mathbb{N}$, le coefficient de degré k de λP vaut : $0.a_0 + \cdots + 0.a_{k-1} + \lambda.a_k = \lambda a_k$, donc : $\lambda P = (\lambda a_k)_{k\in\mathbb{N}}$. En particulier : $1 \times P = P$.

— **Propriétés de** $+$: Il est facile de vérifier que $(\mathbb{K}[X], +)$ est groupe commutatif d'élément neutre 0. L'inverse pour $+$ d'un polynôme $P = (a_k)_{k\in\mathbb{N}}$ est le polynôme $(-a_k)_{k\in\mathbb{N}}$ noté $-P$.

— **Commutativité de** $\times$: Pour tout $k \in \mathbb{N}$, on a : $\sum_{i=0}^{k} a_i b_{k-i} \overbrace{=}^{j=k-i} \sum_{j=0}^{k} b_j a_{k-j}$ et donc $PQ = QP$.

— **Associativité de** $\times$: Pour tout $k \in \mathbb{N}$, le coefficient de degré k de $(PQ)R$ est :

$$\sum_{i=0}^{k}\Big(\sum_{j=0}^{i} a_j b_{i-j}\Big)c_{k-i} = \sum_{0\le j\le i\le k} a_j b_{i-j} c_{k-i} = \sum_{j=0}^{k} a_j\Big(\sum_{i=j}^{k} b_{i-j}c_{k-i}\Big) \overset{l=i-j}{=} \sum_{j=0}^{k} a_j\Big(\sum_{l=0}^{k-j} b_l c_{(k-j)-l}\Big),$$

donc est égal au coeffcient de degré k de $P(QR)$. Ainsi $(PQ)R = P(QR)$.

— **Distributivité de** $\times$ **sur** $+$: Pour tout $k \in \mathbb{N}$, le coefficient de degré k de $P(Q+R)$ est :

$$\sum_{i=0}^{k} a_i(b_{k-i} + c_{k-i}) = \sum_{i=0}^{k} a_i b_{k-i} + \sum_{i=0}^{k} a_i c_{k-i},$$

donc est égal au coefficient de degré k de $(PQ)+(PR)$. Ainsi : $P(Q+R) = (PQ) + (PR)$.

□

On va maintenant se diriger vers une notation définitive des polynômes (une notation de la forme $P = a_n X^n + a_{n-1}X^{n-1} + \cdots + a_1 X + a_0$) et abandonner la notation $P = (a_n)_{n\in\mathbb{N}}$.

Théorème 2.1.7 (Notation polynomiale)**.** *Dans $\mathbb{K}[X]$, on choisit de noter X le polynôme $(0, 1, 0, 0, \cdots)$.*

— *pour tout $k \in \mathbb{N}$: $X^k = (0, \cdots, 0, 1, 0, 0, \cdots)$, polynôme dans lequel le 1 est en position degré k.*

$$1 = (1, 0, 0, \cdots), \quad X = (0, 1, 0, 0, \cdots)$$
$$X^2 = (0, 0, 1, 0, 0, \cdots), \quad X^3 = (0, 0, 0, 1, 0, 0, \cdots)\cdots$$

— *Pour tout polynôme non nul $P = (a_k)_{k\in\mathbb{N}}$ de degré n : $P = \sum_{k=0}^{n} a_k X^k$. On peut aussi écrire que : $P = \sum_{k=0}^{\infty} a_k X^k$ et cette écriture est unique. Une telle somme est **finie** contrairement aux apparences car la suite $(a_k)_{k\in\mathbb{N}}$ est presque nulle. Cette notation "infinie" rend de précieux services de rédaction.*

Démonstration. Une récurrence simple permet de montrer que : $\forall k \in \mathbb{N}^*$, $X^k = (0, \cdots, 0, \underbrace{1}_{k^e}, 0, 0, \cdots)$. □

Attention : L'indéterminée X n'est pas un nombre ! Elle n'a pas de valeur. Elle représente la suite presque nulle $(0, 1, 0, 0, \cdots)$.

▶ Dès que le contexte fait intervenir l'anneau $\mathbb{K}[X]$, la lettre majuscule X est une notation réservée (au même titre que la lettre i des nombres complexes). On n'écrira donc **JAMAIS** des phrases du type "en posant $X = \cdots$" qui n'ont aucun sens ! Autrement dit, il ne faut pas confondre l'indéterminée X avec l'inconnue x.

Explication : En fait, **les polynômes ne sont pas des fonctions**. Notons par exemple P le polynôme $3X^2 + 4X + 1$. Calculer $P(5)$ c'est transformer 5 en un

autre nombre conformément à certaines opérations (puissances, multiplication par un réel, et addition). Or il y a un tas de mondes mathématiques dans lesquels on sait calculer les puissances, multiplier par un réel et additionner les objets :

- le corps $\mathbb{R}$ bien sûr, d'où la possibilité de calculer $P(5)$
- l'anneau $\mathcal{M}_n(\mathbb{R})$, d'où la possibilité de calculer $P(A)$ pour tout $A \in \mathcal{M}_n(\mathbb{R})$.
- l'anneau $\mathbb{R}^{\mathbb{R}}$ des fonctions de $\mathbb{R}$ dans $\mathbb{R}$, d'où la possibilité de noter $P(\exp)$ la fonction $x \mapsto 3e^{2x} + 4e^x + 1$.

Finalement, on ne sait toujours pas ce qu'est le polynôme $P = 3X^2 + 4X + 1$, mais ce n'est pas la gentille fonction $x \mapsto 3x^2 + 4x + 1$ en tout cas.

2.1.2 Propriétés des degrés

Théorème 2.1.8 (Degré d'une somme)**.** *Pour tous $P, Q \in \mathbb{K}[X]$, on a :*

$$\deg(P+Q) \leq \max\{\deg(P), \deg(Q)\}.$$

De plus, si $\deg(P) \neq \deg(Q)$, alors :

$$\deg(P+Q) = \max\{\deg(P), \deg(Q)\}.$$

Démonstration.

— Le résultat est évident lorsque P ou Q est nul. En effet, si par exemple $P = 0$, alors $\deg(P) = -\infty$ et donc

$$\max\{\deg(P), \deg(Q)\} = \deg(Q).$$

D'autre part, $P + Q = Q$ et donc $\deg(P+Q) = \deg(Q)$. Ainsi :

$$\deg(P+Q) = \max\{\deg(P), \deg(Q)\} \leq \max\{\deg(P), \deg(Q)\}.$$

— Supposons $P \neq 0$ et $Q \neq 0$ et notons m le degré de P et n celui de Q, ainsi que $P = (a_k)_{k\in\mathbb{N}}$ et $Q = (b_k)_{k\in\mathbb{N}}$. Pour tout $k > \max\{m, n\}$: $a_k + b_k = 0$, donc : $\deg(P+Q) \leq \max\{m, n\} = \max\{\deg(P), \deg(Q)\}$.

□

Théorème 2.1.9 (Degré d'un produit)**.** *Pour tous $P, Q \in \mathbb{K}[X]$, on a :*

$$\deg(PQ) = \deg(P) + \deg(Q).$$

En particulier, si $\lambda \neq 0$:

$$\deg(\lambda P) = \deg(P).$$

Démonstration. C'est immédiat lorsque P ou Q est nul. Supposons-les donc tous deux non nuls et notons m le degré de P et n celui de Q, ainsi que $P = (a_k)_{k\in\mathbb{N}}$, $Q = (b_k)_{k\in\mathbb{N}}$ et $PQ = (c_k)_{k\in\mathbb{N}}$. On a :

$$\begin{aligned} c_{m+n} &= \sum_{i=0}^{m+n} a_i b_{m+n-i} \\ &= \sum_{i=0}^{m-1} a_i \overbrace{b_{m+n-i}}^{=0 \text{ car } m+n-i>n} + a_m b_n + \sum_{i=m+1}^{m+n} \overbrace{a_i}^{=0} b_{m+n-i} \\ &= a_m b_n. \end{aligned}$$

Comme, $a_m \neq 0$ et $b_n \neq 0$, alors $c_{m+n} \neq 0$ et donc : $\deg(PQ) \geq m+n$.
Inversement, pour tout $k > m+n$:

$$c_k = \sum_{i=0}^{k} a_i b_{k-i} = \sum_{i=0}^{m} a_i \overbrace{b_{k-i}}^{=0 \text{ car } k-i>n} + \sum_{i=m+1}^{k} \overbrace{a_i}^{=0} b_{k-i} = 0,$$

donc $\deg(PQ) \leq m+n$. D'où le résultat. □

2.1.3 Intégrité de $\mathbb{K}[X]$

Théorème 2.1.10. *$\mathbb{K}[X]$ est intègre :*

$$\forall P, Q \in \mathbb{K}[X],\ \Big(PQ = 0 \ \Rightarrow\ P = 0 \, ou \, Q = 0\Big).$$

Démonstration. Pout tous $P, Q \in \mathbb{K}[X]$ tels que $PQ = 0$: $\deg(P) + \deg(Q) = -\infty$, donc nécessairement : $\deg(P) = -\infty$ ou $\deg(Q) = -\infty$, i.e. : $P = 0$ ou $Q = 0$.

Autre méthode : Raisonnons par contraposition et montrons que

$$\forall P, Q \in \mathbb{K}[X],\ \Big(P \neq 0 \, \text{et} \, Q \neq 0 \ \Rightarrow\ PQ \neq 0\Big).$$

Soit $P, Q \in \mathbb{K}[X]$. Supposons que $P \neq 0$ et $Q \neq 0$. Alors : $\deg(P), \deg(Q) \in \mathbb{N}$. Donc : $\deg(PQ) = \deg(P) + \deg(Q) \in \mathbb{N}$. En particulier, $PQ \neq 0$. D'où le résultat. □

Exemple 2.1.11. Soit $n \in \mathbb{N}^*$. Déterminer le degré du polynôme

$$\prod_{k=1}^{n} (X - 2k + 1)^k.$$

► Le polynôme $P = \prod_{k=1}^{n} (X-2k+1)^k$ se présente comme le produit des n polynômes $P_k = (X-2k+1)^k$, avec $1 \leq k \leq n$. Comme $\deg P_k = k$, alors : $\deg P = \sum_{k=1}^{n} k = \dfrac{n(n+1)}{2}$.

On déduit du théorème précédent les polynômes qui sont simplifiables pour la multiplication : ce sont les polynômes non nuls.

Théorème 2.1.12.

$$\forall (P, Q, R) \in (\mathbb{K}[X])^3,\ (P \times Q = P \times R \, et \, P \neq 0 \Rightarrow Q = R)\,.$$

Démonstration. Soit $(P, Q, R) \in (\mathbb{K}[X])^3$ tel que $P \neq 0$. Par intégrité de l'anneau $\mathbb{K}[X]$:

$$PQ = PR \Rightarrow P(Q-R) = 0 \Rightarrow Q - R = 0 \Rightarrow Q = R.$$

□

Explication : Ainsi, par exemple, $(X-1)P = 0 \Rightarrow P = 0$, car $X - 1$ n'est pas le polynôme nul (X n'est pas un nombre mais X est un polynôme et le polynôme X n'est pas le polynôme 1).

2.1.4 Inversibles de l'anneau $\mathbb{K}[X]$

Théorème 2.1.13. *Les inversibles de l'anneau $\mathbb{K}[X]$ sont les constantes non nulles. Ainsi $\mathbb{K}[X]^\times = \mathbb{K}^*$.*

Démonstration. Soit $P \in \mathbb{K}[X]$. On suppose que P est inversible. Donc, il existe un polynôme $Q \in \mathbb{K}[X]$ tel que $P \times Q = 1$. Ceci impose déjà $P \neq 0$ et $Q \neq 0$ et donc P et Q ont des degrés qui sont des entiers naturels. Ensuite,

$$PQ = 1 \Rightarrow \deg(PQ) = \deg(1) \Rightarrow \deg(P) + \deg(Q) = 0.$$

Donc, nécessairement, $\deg(P) = 0$ ou encore, P est une constante non nulle.

Réciproquement, si $P = a_0 \neq 0$, alors, en posant $Q = \frac{1}{a_0} \in \mathbb{K}[X]$, on a $P \times Q = 1$ et donc P est inversible dans l'anneau $\mathbb{K}[X]$. □

2.1.5 L'évaluation et les fonctions polynomiales

Définition 2.1.14 (Evaluation, fonction polynomiale)**.** Soient $P \in \mathbb{K}[X]$ et $\lambda \in \mathbb{K}$. Si $P = 0$, on pose $P(\lambda) = 0$. Si $P \neq 0$, en notant n son degré et $P = \sum_{k=0}^{n} a_k X^k$, on pose $P(\lambda) = \sum_{k=0}^{n} a_k \lambda^k$. L'application $\widetilde{P} : \mathbb{K} \to \mathbb{K}$ définie par $\widetilde{P}(x) = P(x)$ est appelée fonction polynomiale associée au polynôme P. On la note $\widetilde{P}$ quand on veut la distinguer proprement du polynôme P, mais aussi souvent P.

Remarque 2.1.15. Pour tout polynôme $P = \sum_{k=0}^{+\infty} a_k X^k$ avec $(a_k)_{k\in\mathbb{N}}$ presque nulle, on a clairement : $P(\lambda) = \sum_{k=0}^{+\infty} a_k \lambda^k$, cette dernière somme étant finie.
▶ On dit également que le nombre $P(\lambda)$ est la valeur de P en λ. Évaluer P en λ signifie calculer le nombre $P(\lambda)$.

2.1.6 Composition des polynômes

Définition 2.1.16 (Composition des polynômes)**.** Soient P et Q deux polynômes de $\mathbb{K}[X]$. Si $P = 0$, on pose $P \circ Q = 0$. Si $P \neq 0$, on note $n \geq 0$ son degré et l'on écrit P sous la forme $P = \sum_{k=0}^{n} a_k X^k$. On pose alors $P \circ Q = \sum_{k=0}^{n} a_k Q^k$, avec la convention $Q^0 = 1$. Le polynôme $P \circ Q$ est également noté $P(Q)$.

La notation usuelle P=P(X) : Lorsque $Q = X$, on a $P(Q) = P \circ Q = \sum_{k=0}^{n} a_k X^k = P$ et on peut donc écrire que $P(X) = P$.

Théorème 2.1.17 (Degré d'une composée)**.** *Si Q n'est pas constant :*

$$\deg(P \circ Q) = \deg(P) \times \deg(Q).$$

Démonstration. On suppose Q non constant et on pose : $m = \deg(P)$. Par produit, on a : $\deg(Q^k) = k \deg(Q)$ pour tout $k \in \{0, \cdots, m\}$. Donc, comme

$\deg(Q) \geq 1$, la suite $(\deg(Q^k))_{0\leq k\leq m}$ est strictement croissante. Finalement, par somme :

$$\deg(P\circ Q) = \deg(\sum_{k=0}^{m} a_k Q^k) \overset{a_m\neq 0}{=} \deg(Q^m) = m\deg(Q) = \deg(P)\times\deg(Q).$$

□

2.1.7 Dérivation des polynômes

Définition 2.1.18 (Dérivation des polynômes)**.** Soit $P \in \mathbb{K}[X]$. On définit le polynôme dérivé du polynôme P, noté P', de la manière suivante :

— Si $\deg(P) \leq 0$, on pose $P' = 0$.

— Si $n = \deg(P) \geq 1$ et si $P = \sum_{k=0}^{n} a_k X^k$, on pose $P' = \sum_{k=1}^{n} k a_k X^{k-1}$.

On définit ensuite pour tout $n \in \mathbb{N}$ le $n^{\text{ème}}$ polynôme dérivé de P, noté $P^{(n)}$. Pour commencer : $P^{(0)} = P$ et pour tout $n \in \mathbb{N}$: $P^{(n+1)} = (P^{(n)})'$. Pour $n = 1$ et $n = 2$, on préfère les notations P' et P'' aux notations $P^{(1)}$ et $P^{(2)}$.

► Autrement dit, pour tout polynôme $P = \sum_{k=0}^{+\infty} a_k X^k \in \mathbb{K}[X]$ avec $(a_k)_{k\in\mathbb{N}}$ presque nulle, on a : $P' = \sum_{k=1}^{+\infty} k a_k X^{k-1} = \sum_{\ell=0}^{+\infty}(\ell+1)a_{\ell+1}X^{\ell}$.

Théorème 2.1.19. *Soient $P, Q \in \mathbb{K}[X]$ et $n \in \mathbb{N}$.*

1. ***Degré :*** $\begin{cases} \deg(P^{(n)}) = \deg(P) - n & \text{si : } n \leq \deg(P) \\ P^{(n)} = 0, & \text{sinon.} \end{cases}$
2. ***Somme :*** $(P+Q)^{(n)} = P^{(n)} + Q^{(n)}$.
3. ***Produit :*** $(PQ)' = P'Q + PQ'$. *Plus généralement :* $(PQ)^{(n)} = \sum_{k=0}^{n} C_n^k P^{(k)} Q^{(n-k)}$ *(formule de Leibniz).*
4. ***Composition :*** $(P\circ Q)' = Q' \times P'\circ Q$.

Démonstration. Introduisant les coefficients de P et Q :

$$P = \sum_{k=0}^{+\infty} a_k X^k \text{ et } Q = \sum_{k=0}^{+\infty} b_k X^k.$$

— **Degré :** Posons : $d = \deg(P)$. Si $d \leq 0$: $P' = 0$. Si au contraire $d \geq 1$: $P' = \sum_{k=0}^{+\infty} k a_k X^{k-1}$ avec $d a_d \neq 0$, donc : $\deg(P') = d - 1$. On généralise par récurrence aux cas des dérivées successives.

— **Somme :** Evident.

— **Produit :** Montrons d'abord la formule : $(PQ)' = P'Q + PQ'$. D'après la définition du produit sur $\mathbb{K}[X]$, on a : $PQ = \Big(\sum_{k=0}^{+\infty} a_k X^k\Big)\Big(\sum_{k=0}^{+\infty} b_k X^k\Big) = \sum_{n=0}^{+\infty}\Big(\sum_{k=0}^{n} a_k b_{n-k}\Big)X^n$. Ainsi : $(PQ)' = \sum_{n=1}^{+\infty} n\big(\sum_{k=0}^{n} a_k b_{n-k}\big)X^{n-1}$ et donc

$$
\begin{aligned}
(PQ)' &= \sum_{n=1}^{+\infty} n a_0 b_n X^{n-1} + \sum_{k=1}^{+\infty} \Big(\sum_{n=k}^{+\infty} n a_k b_{n-k} X^{n-1} \Big) \\
&= a_0 Q' + \sum_{k=1}^{+\infty} a_k \Big(\sum_{n=k}^{+\infty} (n-k+k) b_{n-k} X^{n-1} \Big) \\
&= a_0 Q' + \sum_{k=1}^{+\infty} a_k \Big(\sum_{n=k}^{+\infty} (n-k) b_{n-k} X^{n-1} \Big) + \sum_{k=1}^{+\infty} k a_k \Big(\sum_{n=k}^{+\infty} b_{n-k} X^{n-1} \Big) \\
&= a_0 Q' + \sum_{k=1}^{+\infty} a_k \Big(\sum_{\ell=0}^{+\infty} \ell b_\ell X^{\ell-1+k} \Big) + \sum_{k=1}^{+\infty} k a_k \Big(\sum_{\ell=0}^{+\infty} b_\ell X^{\ell+k-1} \Big) \\
&= a_0 Q' + \sum_{k=1}^{+\infty} a_k X^k \Big(\sum_{\ell=1}^{+\infty} \ell b_\ell X^{\ell-1} \Big) + \sum_{k=1}^{+\infty} k a_k X^{k-1} \Big(\sum_{\ell=0}^{+\infty} b_\ell X^{\ell} \Big) \\
&= a_0 Q' + \sum_{k=1}^{+\infty} a_k X^k Q' + \sum_{k=1}^{+\infty} k a_k X^{k-1} Q = PQ' + P'Q.
\end{aligned}
$$

La formule de Leibniz s'en déduit par récurrence sur n.

— **Initialisation :** Pour $n = 0$, rien à faire !

— **Hérédité :** Soit $n \in \mathbb{N}$. Faisons l'hypothèse que la formule de Leibniz : $(PQ)^{(n)} = \ldots$ est vraie pour tous $P, Q \in \mathbb{K}[X]$. Alors pour tous $P, Q \in \mathbb{K}[X]$:

$$
\begin{aligned}
(PQ)^{(n+1)} &= ((PQ)')^{(n)} = (P'Q + PQ')^{(n)} = (P'Q)^{(n)} + (PQ')^{(n)} \\
&= \sum_{k=0}^{n} C_n^k (P')^{(k)} Q^{(n-k)} + \sum_{\ell=0}^{n} C_n^\ell P^{(\ell)} (Q')^{(n-\ell)} \\
&= \sum_{k=0}^{n} C_n^k P^{(k+1)} Q^{(n-k)} + \sum_{\ell=0}^{n} C_n^\ell P^{(\ell)} Q^{(n+1-\ell)} \\
&= \sum_{k=1}^{n+1} C_n^{k-1} P^{(k)} Q^{(n+1-k)} + \sum_{\ell=0}^{n} C_n^\ell P^{(\ell)} Q^{(n+1-\ell)}
\end{aligned}
$$

En décomposant les sommes précédentes de la façon suivante :

$$
\begin{aligned}
&(PQ)^{(n+1)} \\
&= P^{(n+1)} Q^{(0)} + \sum_{k=1}^{n} C_n^{k-1} P^{(k)} Q^{(n+1-k)} + \sum_{k=1}^{n} C_n^k P^{(k)} Q^{(n+1-k)} \\
&\quad + P^{(0)} Q^{(n+1)} \\
&= P^{(n+1)} Q^{(0)} + \sum_{k=1}^{n} \Big(C_n^{k-1} + C_n^k \Big) P^{(k)} Q^{(n+1-k)} + P^{(0)} Q^{(n+1)}
\end{aligned}
$$

et en utilisant la formule de Pascal $C_n^{k-1} + C_n^k = C_{n+1}^k$, on obtient :

$$
\begin{aligned}
(PQ)^{(n+1)} &= P^{(n+1)} Q^{(0)} + \sum_{k=1}^{n} C_{n+1}^k P^{(k)} Q^{(n+1-k)} + P^{(0)} Q^{(n+1)} \\
&= \sum_{k=0}^{n+1} C_{n+1}^k P^{(k)} Q^{(n+1-k)}.
\end{aligned}
$$

D'où le résultat.

— **Composition :** Tout d'abord, montrons par récurrence que $\forall P \in \mathbb{K}[X], \forall n \in \mathbb{N}^* : (P^n)' = nP'P^{n-1}$.

— **Initialisation :** Pour $n = 1$, rien à faire !

— **Hérédité :** Soit $n \in \mathbb{N}^*$. Supposons la propriété est vraie à l'ordre n et montrons-là à l'ordre $n+1$. On a : $(P^{n+1})' = (P^n \times P)' = (P^n)'P + P^nP' = (nP'P^{n-1}) \times P + P^nP' = nP'P^n + P'P^n = (n+1)P'P^n$. D'où le résultat.

Passons maintenant à la preuve du fait que $(P \circ Q)' = Q' \times P' \circ Q$.

Par définition : $P \circ Q = \sum_{k=0}^{+\infty} a_k Q^k$, donc : $(P \circ Q)' = \sum_{k=0}^{+\infty} a_k (Q^k)'$. Or : $(Q^k)' = kQ'Q^{k-1}$ pour tout $k \in \mathbb{N}^*$, donc :

$$(P \circ Q)' = \sum_{k=0}^{+\infty} a_k k Q' Q^{k-1} = Q' \times P' \circ Q.$$

□

Proposition 2.1.20 (Formule de Taylor pour les polynômes)**.** *Pour tous $n \in \mathbb{N}$, $P \in \mathbb{K}_n[X]$ et $\lambda \in \mathbb{K}$, on a :*

$$P(X) = \sum_{k=0}^{n} \frac{P^{(k)}(\lambda)}{k!}(X-\lambda)^k.$$

Démonstration. Un polynôme P de $\mathbb{K}_n[X]$ s'écrit $P = \sum_{i=0}^{n} a_i X^i$. Par linéarité de la dérivation, pour tout k entre 0 et n, on a $P^{(k)} = \sum_{i=0}^{n} a_i (X^i)^{(k)} = \sum_{i=k}^{n} a_i \frac{i!}{(i-k)!} X^{i-k}$.

D'où après évaluation en λ : $P^{(k)}(\lambda) = \sum_{i=k}^{n} a_i \frac{i!}{(i-k)!}\lambda^{i-k}$ et donc

$$\begin{aligned}
\sum_{k=0}^{n} \frac{P^{(k)}(\lambda)}{k!}(X-\lambda)^k &= \sum_{k=0}^{n} \frac{1}{k!}\Big(\sum_{i=k}^{n} a_i \frac{i!}{(i-k)!}\lambda^{i-k}\Big)(X-\lambda)^k \\
&= \sum_{i=0}^{n} \Big(\sum_{k=0}^{i} a_i \frac{i!}{k!(i-k)!}\lambda^{i-k}(X-\lambda)^k\Big) \\
&= \sum_{i=0}^{n} a_i \Big(\sum_{k=0}^{i} C_i^k \lambda^{i-k}(X-\lambda)^k\Big) = \sum_{i=0}^{n} a_i \Big(X - \lambda + \lambda\Big)^i \\
&= \sum_{i=0}^{n} a_i X^i = P(X),
\end{aligned}$$

où l'on a appliqué la formule du binôme à l'avant-dernière ligne. □

2.1.8 Relation de divisibilité

Définition 2.1.21 (Divisibilité, diviseur, multiple)**.** Soient $A, B \in \mathbb{K}[X]$. On dit que A divise B, ou que A est un diviseur de B, ou que B est divisible par A, ou que B est un multiple de A, s'il existe $P \in \mathbb{K}[X]$ pour lequel : $B = AP$. Cette relation se note $A|B$.

Exemple 2.1.22. Pour tout $n \in \mathbb{N}$, $X^{2n+1} + 1$ est divisible par $X + 1$.
▶ En effet,

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

Théorème 2.1.23 (Propriétés de la relation de divisibilité)**.** *Soient* $A, B, C, D \in \mathbb{K}[X]$.

- *La relation de divisibilité* $|$ *est réflexive et transitive dans* $\mathbb{K}[X]$*, c'est même une relation d'ordre sur l'ensemble des polynômes* ***unitaires ou nuls*** *de* $\mathbb{K}[X]$*. Elle est en revanche seulement réflexive et transitive sur* $\mathbb{K}[X]$ *car pour tous* $A, B \in \mathbb{K}[X]$ *:*

$$A|B \text{ et } B|A \Leftrightarrow \exists \lambda \in \mathbb{K}^* \mid A = \lambda B.$$

  On dit alors que A *et* B *sont* ***associés*** *(sur* $\mathbb{K}$*).*
- *Si :* $D|A$ *et* $D|B$*, alors :* $D|(AU + BV)$ *pout tous* $U, V \in \mathbb{K}[X]$.
- *Si :* $A|B$ *et* $C|D$*, alors :* $AC|BD$*. En particulier, si :* $A|B$*, alors :* $A^k|B^k$ *pour tout* $k \in \mathbb{N}$.

Remarque : La relation de divisibilité restreinte à l'ensemble des polynômes **unitaires** est une relation d'ordre. En effet, deux polynômes unitaires associés sont égaux, ce qui prouve l'antisymétrie.

Démonstration. La relation de divisibilité est réflexive et transitive. En effet :

- Soit $A \in \mathbb{K}[X]$. On a : $A = 1 \times A$ avec $1 \in \mathbb{K}[X]$. Donc, $A|A$.
- Soit $(A, B, C) \in (\mathbb{K}[X])^3$ tel que $A|B$ et $B|C$. Il existe $(Q_1, Q_2) \in (\mathbb{K}[X])^2$ tel que $B = Q_1 A$ et $C = Q_2 B$. Ainsi, $C = Q_2 Q_1 A$ avec $Q_2 Q_1 \in \mathbb{K}[X]$ et donc $A|C$.

Par contre, la relation de divisibilité n'est pas anti-symétrique. En effet, si : $A = \lambda B$ avec $\lambda \in \mathbb{K}^*$, on a aussi : $B = \frac{1}{\lambda} A$, donc $A|B$ et $B|A$. Réciproquement, supposons que : $A|B$ et $B|A$. Il existe alors $P, Q \in \mathbb{K}[X]$ pour lesquels : $A = PB$ et $B = QA$, donc : $A = PQA$. Deux cas se présentent :

- Si $A = 0$: $B = QA = 0$, donc $A = \lambda B$ pour $\lambda = 1$.
- Si au contraire $A \neq 0$: $PQ = 1$ par intégrité de $\mathbb{K}[X]$. Donc, les polynômes P et Q sont des inversibles de l'anneau $\mathbb{K}[X]$ et donc : P et Q sont des constantes non nulles. Ainsi, il existe $\lambda \in \mathbb{K}^*$ tel que $A = \lambda B$.

□

Théorème 2.1.24. *Soient* A *et* B *deux polynômes non nuls tels que* $B|A$*. Alors,* $\deg(B) \leq \deg(A)$.

Démonstration. Comme $B|A$, alors il existe $Q \in \mathbb{K}[x]$ tel que $A = QB$. Puisque $A \neq 0$, on a nécessairement $Q \neq 0$. Ainsi, $\deg(A), \deg(B), \deg(Q)$ sont des entiers naturels. Donc : $\deg(A) = \deg(QB) = \deg(Q) + \deg(B) \geq \deg(B)$. D'où le résultat. □

2.1.9 Division euclidienne

Théorème 2.1.25 (Division euclidienne)**.** *Soient $A, B \in \mathbb{K}[X]$ avec $B \neq 0$. Il existe un et un seul couple de polynômes $(Q, R) \in (\mathbb{K}[\mathbb{X}])^2$ pour lequel : $A = BQ+R$ et $\deg(R) < \deg(B)$. On appelle A le dividende de la division euclidienne de A par B, B son diviseur, Q son quotient et R son reste.*

Remarque 2.1.26. L'unicité dans la division euclidienne sur $\mathbb{K}[X]$ permet d'affirmer qu'un polynôme $B \neq 0$ divise A si et seulement si le reste dans la division euclidienne de A par B est nul.

Démonstration. Raisonnons en deux temps.

- **Unicité de (Q, R) :** soient (Q_1, R_1) et (Q_2, R_2) deux couples de polynômes de $\mathbb{K}[X]$ tels que $A = BQ_1 + R_1$ et $A = BQ_2 + R_2$, avec $\deg(R_1) < \deg(B)$ et $\deg(R_2) < \deg(B)$. On a donc $BQ_1 + R_1 = BQ_2 + R_2$, d'où l'égalité $B(Q_1 - Q_2) = R_2 - R_1$ avec $\deg(R_2 - R_1) \leqslant \max(\deg(R_1), \deg(R_2)) < \deg(B)$. Raisonnons par l'absurde et supposons que $Q_1 \neq Q_2$. Alors $\deg(Q_1 - Q_2) \geqslant 0$. Ce qui entraine que $\deg(R_2 - R_1) = \deg(B(Q_1 - Q_2)) = \deg(B) + \deg(Q_1 - Q_2) \geqslant \deg(B)$. Contradiction. Ainsi $Q_1 = Q_2$, puis $R_1 = A - BQ_1 = A - BQ_2 = R_2$.
- **Existence de (Q, R) :** Raisonnons par récurrence sur le degré de A. Notons $m = \deg(B) \geqslant 0$ et $B = b_m X^m + \cdots + b_0$ avec $b_m \neq 0$. Soient n un entier naturel et $HR(n)$ la propriété suivante : pour tout polynôme A de degré inférieur ou égal à n, il existe un couple (Q, R) de polynômes à coefficients dans $\mathbb{K}$ tels que $A = BQ+R$ et $\deg(R) < \deg(B)$. L'hypothèse $HR(0)$ est vraie : en effet, $A = a_0 \in \mathbb{K}$. Si $m \geqslant 1$, on pose $Q = 0$ et $R = A$; si $m = 0$, on pose $Q = a_0/b_0$ et $R = 0$. On a bien dans tous les cas $A = BQ + R$ et $\deg(R) < \deg(B)$. Soit $n \geqslant 0$ Supposons $HR(n)$ vérifiée. Soit A un polynôme de degré $n+1$: $A = a_{n+1}X^{n+1} + \cdots + a_1X + a_0$ Si $n+1 < m = \deg(B)$, il suffit de poser $R = A$ et $Q = 0$. Si $n+1 \geqslant m = \deg(B)$, posons $A_1 = A - (a_{n+1}/b_m) X^{n+1-m} B$. On a $\deg(A_1) \leqslant n$, donc d'après $HR(n)$, il existe un couple de polynômes (Q_1, R_1) tels que $A_1 = Q_1 B + R_1$ avec $\deg(R_1) < \deg(B)$. Posons $R = R_1$ et $Q = Q_1 + (a_{n+1}/b_m) X^{n+1-m}$. On a bien $A = BQ + R$ et $\deg(R) < \deg(B)$. D'où le résultat.

□

Parmi les exemples fondamentaux de division euclidienne, on retiendra l'exemple de la division d'un polynôme P par $X - \lambda$, où λ est un élément de $\mathbb{K}$.

Théorème 2.1.27 (Division par $X - \lambda$)**.** *Soient $\lambda \in \mathbb{K}$ et $P \in \mathbb{K}[X]$. Le reste de la division euclidienne de P par $X - \lambda$ est $P(\lambda)$.*

Démonstration. La division euclidienne de P par $X - \lambda$ s'écrit :

$$P = (X - \lambda)Q + R,$$

pour certains $(Q, R) \in \mathbb{K}[X]^2$ avec : $\deg(R) < 1$, donc en fait R est un polynôme constant. Evaluons en λ : $P(\lambda) = (\lambda - \lambda)Q(\lambda) + R(\lambda) = R$. □

Exemple 2.1.28. Pour tout $n \in \mathbb{N}$, le reste de division euclidienne de X^n par $X^2 - 3X + 2$ vaut : $(2^n - 1)X - (2^n - 2)$.

▶ En effet, soit $n \in \mathbb{N}$. La division euclidienne de X^n par $X^2 - 3X + 2$ s'écrit : $X^n = (X - 1)(X - 2)Q + aX + b$ pour certain $Q \in \mathbb{R}[X]$ et $a, b \in \mathbb{R}$. Evaluons

en 1 : $1 = a + b$, puis en 2 : $2^n = 2a + b$. Après calcul, on obtient : $a = 2^n - 1$ et $b = 2 - 2^n$.

Cette méthode peut être adaptée au calcul du reste dans la division euclidienne d'un polynôme P par $(X-a)^2$. Il faudra dans ce cadre faire appel à la dérivation.

▶ Déterminons de deux manières le reste R dans la division euclidienne d'un polynôme $P \in \mathbb{K}[X]$ par $B = (X-a)^2$ où $a \in \mathbb{K}$.

— Il existe $Q \in \mathbb{K}[X]$ et $R = \alpha X + \beta \in \mathbb{K}_1[X]$ tels que $P = (X-a)^2 Q + \alpha X + \beta$. On en déduit, après évaluation en a, que $P(a) = \alpha a + \beta$. De plus, $P' = 2(X-a)Q + (X-a)^2 Q' + \alpha$ et donc, après évaluation en a, on obtient : $\alpha = P'(a)$. D'où $R = P'(a)X + P(a) - aP'(a) = P(a) + (X-a)P'(a)$.

— D'après la formule de Taylor en a, pour $n = \deg(P)$ (on peut supposer $n \geqslant 2$ car sinon $Q = 0$ et $R = P$) $P = \sum_{k=0}^{n} \frac{P^{(k)}(a)}{k!}(X-a)^k = P(a) + P'(a)(X-a) + (X-a)^2 \sum_{k=2}^{n} \frac{P^{(k)}(a)}{k!}(X-a)^{k-2}$. On en déduit, par unicité dans la division euclidienne, que $R = P(a) + P'(a)(X-a)$.

Exercice 2.1.29. Pour $n \in \mathbb{N}$, on pose $P_n = X^n - 1$. Effectuer la division euclidienne de P_n par P_m.

Solution. Soit $(m, n) \in \mathbb{N}^2$. La division euclidienne de n par m s'écrit $n = qm + r$ où $(q, r) \in \mathbb{N}^2$ et $0 \leq r < m$. On écrit $X^n - 1 = X^{qm+r} - 1$ et donc

$$\begin{aligned} X^n - 1 = X^{qm+r} - X^r + X^r - 1 &= X^r\Big(\big(X^m\big)^q - 1\Big) + X^r - 1 \\ &= X^r\Big(1 + X^m + (X^m)^2 + \cdots + (X^m)^{q-1}\Big)(X^m - 1) + X^r - 1. \end{aligned}$$

Puisque $\deg(X^r - 1) \leq r < m$ (si $r \geq 1$ $\deg(X^r - 1) = r$ et si $r = 0$, $\deg(X^r - 1) = -\infty$), la division euclidienne est achevée. Le quotient est $Q = X^r\Big(1 + X^m + (X^m)^2 + \cdots + (X^m)^{q-1}\Big)$ et le reste est $R = X^r - 1$.

2.1.10 PGCD et PPCM

Définition 2.1.30 (Diviseur/multiple commun)**.** Soient $A_1, \cdots, A_r \in \mathbb{K}[X]$.

— On appelle divieur commun de $A_1, \cdots, A_r$ tout polynôme de $\mathbb{K}[X]$ qui divise à la fois $A_1, \cdots, A_r$.

— On appelle multiple commun de $A_1, \cdots, A_r$ tout polynôme de $\mathbb{K}[X]$ qui divisible à la fois par $A_1, \cdots, A_r$.

Théorème 2.1.31 (PGCD de deux polynômes)**.**

— *Soit $A, B \in \mathbb{K}[X]$ avec $A \neq 0$ et $B \neq 0$. On appelle plus grand commun diviseur ou (PGCD) de A et B tout diviseur commun de A et B de degré maximal.*

— *On convient que 0 est le seul PGCD de 0 et 0.*

Démonstration. Justifions l'existence d'un PGCD dans le cas où : $A \neq 0$. Or l'ensemble des **degrés** des diviseurs commun non nuls de A et B contient 0 (car A et B sont divisibles par 1) et il est majoré par $\deg(A)$. Donc, cet ensemble possède un plus grand élément car partie non vide et majorée de $\mathbb{N}$. □

Exemple 2.1.32. Pour tout $A \in \mathbb{K}[X]$, les PGCD de A et 0 sont exactement les associés de A.

▶ En effet, si $A \neq 0$, les diviseurs communs de A et 0 sont exactement les

diviseurs de A et les diviseurs de A de degré maximal sont exactement ses associés.

Théorème 2.1.33 (Idée fondamentale de l'algorithme d'Euclide)**.** *Pour tous $A, B, K \in \mathbb{K}[X]$, $A+BK$ et B ont les mêmes diviseurs communs que A et B, et donc aussi les mêmes PGCD.*

Explication : En particulier, pour tous $A, B \in \mathbb{K}[X]$ avec $B \neq 0$, en notant R le reste de la division euclidienne de A par B, B et R ont les mêmes diviseurs communs que A et B.

Démonstration. Tout diviseur commun de A et B divise aussi $A + BK$ et B, et inversement, tout diviseur commun de $A + BK$ et B divise aussi $A = (A + BK) - BK$ et B. □

Théorème 2.1.34 (Unicité du PGCD de deux polynômes)**.** *Soient $A, B \in \mathbb{K}[X]$.*

- *Les PGCD de A et B sont associés. Un seul d'entre eux est donc unitaire (ou nul si $A = B = 0$) et on l'appelle **LE** PGCD de A et B et on le note $A \wedge B$.*
- *Les diviseurs communs de A et B sont exactement les diviseurs de $A \wedge B$.*

Démonstration. Nous allons metre en oeuvre dans cette preuve un algorithme de calcul du PGCD qu'on appelle *l'algorithme d'Euclide*. Soient $A, B \in \mathbb{K}[X]$. On peut supposer que : $\deg(B) \leq \deg(A)$ sans perte de généralité. On définit une suite de polynômes $R_0, R_1, R_2, \cdots$ de la manière suivante :

- Au départ, on pose : $R_0 = A$ et $R_1 = B$.
- Ensuite, pour $k \in \mathbb{N}$, **TANT QUE :** $R_{k+1} \neq 0$, on note R_{k+2} le reste de la division euclidienne de R_k par R_{k+1}, et en particulier, on a : $\deg(R_{k+2}) < \deg(R_{k+1})$.

A l'issue de cette construction : $\deg(R_0) > \deg(R_1) > \deg(R_2) > \cdots$, et comme il n'existe qu'un nombre **FINI** d'entiers naturels entre 0 et $\deg(R_0)$, on obtient forcément $\deg(R_N) = -\infty$ pour un certain $N \in \mathbb{N}^*$ i.e. $R_N = 0$ et l'algorithme se termine. Or, en vertu de l'idée fondamentale de l'algorithme d'Euclide, $A = R_0$ et $B = R_1$ ont les mêmes diviseurs communs et le mêmes PGCD que R_1 et R_2, puis que R_2 et $R_3 \cdots$ et enfin que R_{N-1} et $R_N = 0$. Les PGCD de R_{N-1} et 0 étant exactement les associés de R_{N-1}, les diviseurs communs de A et B sont ainsi exactement les diviseurs de R_{N-1} et leurs PGCD sont exactement les associés de R_{N-1}. En particulier, les PGCD de A et B sont associés. □

Algorithme d'Euclide Comme on vient de le voir, l'algorithme d'Euclide est un algorithme de calcul du PGCD de deux polynômes. Il a été montré en particulier que : $A \wedge B$ est associé au dernier reste non nul dans l'algorithme d'Euclide.

Exemple 2.1.35. Calculons $(X^5 - X^4 + X^3 - X^2 + X - 1) \wedge (X^2 - 1)$.

$X^5 - X^4 + X^3 - X^2 + X - 1$	$X^2 - 1$
$-\left(X^5 - X^3\right)$	$X^3 - X^2 + 2X - 2$
$-X^4 + 2X^3 - X^2 + X - 1$ $-\left(-X^4 + X^2\right)$	
$2X^3 - 2X^2 + X - 1$ $-\left(2X^3 - 2X\right)$	
$-2X^2 + 3X - 1$ $-\left(-2X^2 + 2\right)$	
$3X - 3$	

D'où $X^5 - X^4 + X^3 - X^2 + X - 1 = (X^3 - X^2 + 2X - 2)(X^2 - 1) + 3X - 3$ puis, on a $X^2 - 1 = \frac{1}{3}(3X - 3)(X + 1) + 0$. Le pgcd est le dernier reste non nul, donc $(X^5 - X^4 + X^3 - X^2 + X - 1) \wedge (X^2 - 1) = X - 1$.

Remarque 2.1.36.
- $A \wedge B$ est un polynôme unitaire.
- Si $A|B$ et $A \neq 0$, $A \wedge B = \frac{1}{\text{dom}(A)} A$, où $\text{dom}(A)$ est le coefficient dominant de A.

Théorème 2.1.37 (Relations de Bézout pour deux polynômes)**.** *Soient $A, B \in \mathbb{K}[X]$. Il existe des polynômes $U, V \in \mathbb{K}[X]$ pour lesquels : $A \wedge B = AU + BV$. Une telle relation est appelé **une** relation de Bézout de A et B.*

Attention : Les polynômes U et V ne sont pas uniques. En fait, Si A et B sont non constants, on a de plus unicité des polynômes U et V lorsque l'on impose la condition supplémentaire $\deg(U) < \deg(B)$ et $\deg(V) < \deg(A)$.

Démonstration. Si B divise A, un PGCD de A et B est B ou encore $A \wedge B = \frac{1}{\text{dom}(B)} B = 0 \times A + \frac{1}{\text{dom}(B)} B$. Dans ce cas, les polynômes $U = 0$ et $V = \frac{1}{\text{dom}(B)}$ conviennent.
Sinon, avec les notations de l'algorithme d'Euclide exposé plus haut, on a $A \wedge B = \frac{1}{\text{dom}(R_{N-1})} R_{N-1}$. A partir de l'égalité, $R_{N-3} = Q_{N-3} R_{N-2} + R_{N-1}$, on obtient une égalité de la forme : $A \wedge B = \frac{1}{\text{dom}(R_{N-1})} R_{N-1} = R_{N-3} U_{N-3} + R_{N-2} V_{N-3}$, où $U_{N-3} = \frac{1}{\text{dom}(R_{N-1})}$ et $V_{N-3} = -\frac{1}{\text{dom}(R_{N-1})} Q_{N-3}$ sont des polynômes. Puis en remontant dans l'algorithme, par récurrence, on peut écrire $A \wedge B$ sous la forme $A \wedge B = R_k U_k + R_{k+1} V_k$ pour tout $k \in \{0, \cdots, N-3\}$, où U_k et V_k sont des polynômes. En particulier, pour $k = 0$, il existe deux polynômes U et V tels que

$$A \wedge B = R_0 U + R_1 V = AU + BV.$$

□

Exercice 2.1.38. Déterminer deux polynômes U et V tels que

$$U \times (X^2 + X + 1) + V \times (X^2 - X + 1) = 1.$$

Solution. On a :

$$X^2 + X + 1 = 1 \times (X^2 - X + 1) + 2X$$

puis

$$X^2 - X + 1 = (\frac{X}{2} - \frac{1}{2}) \times (2X) + 1.$$

Donc,

$$\begin{aligned} 1 &= (X^2 - X + 1) - (\frac{X}{2} - \frac{1}{2}) \times (2X) \\ &= (X^2 - X + 1) - (\frac{X}{2} - \frac{1}{2})\big[(X^2 + X + 1) - (X^2 - X + 1)\big] \\ &= (-\frac{X}{2} + \frac{1}{2})(X^2 + X + 1) + (\frac{X}{2} + \frac{1}{2})(X^2 - X + 1). \end{aligned}$$

Les polynômes $U = -\frac{X}{2} + \frac{1}{2}$ et $V = \frac{X}{2} + \frac{1}{2}$.

Théorème 2.1.39 (Factorisation par un diviseur commun)**.** *Soient $A, B, C \in \mathbb{K}[X] \setminus \{0\}$ avec C unitaire. Alors, $(CA) \wedge (CB) = C(A \wedge B)$.*

Démonstration. C divise CA et CB et donc C divise $(CA) \wedge (CB)$. Il existe un polynôme non nul Q tel que $(CA) \wedge (CB) = CQ$. CQ divise CA et donc Q divise A ($CA = Q_1CQ$ puis $A = Q_1Q$ car $C \neq 0$). De même, Q divise B et donc Q divise $A \wedge B$.

Inversement, $C(A \wedge B)$ divise CA et CB puis $C(A \wedge B)$ divise $(CA) \wedge (CB) = CQ$ et finalement $A \wedge B$ divise Q.

En résumé, Q divise $A \wedge B$ et $A \wedge B$ divise Q. Donc, il existe $\lambda \in \mathbb{K}^*$ tel que $Q = \lambda(A \wedge B)$. Enfin, CQ est unitaire car égal à $(CA) \wedge (CB)$ et C est unitaire. Donc Q est unitaire puis $\lambda = 1$ puis $Q = A \wedge B$ et finalement $(CA) \wedge (CB) = C(A \wedge B)$. □

Théorème 2.1.40. *Soient $A, B \in \mathbb{K}[X] \setminus \{0\}$. En posant $D = A \wedge B$, il existe $A_1, B_1 \in \mathbb{K}[X] \setminus \{0\}$ tel que $A = DA_1$ et $B = DB_1$ et $A_1 \wedge B_1 = 1$.*

Démonstration. Puisque D divise A et B, il existe deux polynômes A_1 et B_1 tels que $A = DA_1$ et $B = DB_1$. De plus, $D = A \wedge B = (DA_1) \wedge (DB_1) = D(A_1 \wedge B_1)$ puis $A_1 \wedge B_1 = 1$ après simplification par le polynôme non nul D. □

Définition 2.1.41 (PPCM de deux polynômes)**.** Soient $A, B \in \mathbb{K}[X]$. On appelle plus petit commun multiple (ou PPCM) de A et B tout polynôme $M \in \mathbb{K}[X]$ satisfaisant les deux assertions :

— M est multiple commun de A et B,

— M divise tout multiple commun de A et B.

Théorème 2.1.42 (Existence et unicité du PPCM)**.** *Soient $A, B \in \mathbb{K}[X]$. Les polynômes A et B possèdent un unique PPCM UNITAIRE OU NUL appelé LE PPCM de A et B et noté $A \vee B$, et les autres PPCM sont les associés de $A \vee B$, i.e. les polynômes $\lambda(A \vee B)$, λ décrivant $\mathbb{K}^*$.*

Proposition 2.1.43. *Soient $A, B \in \mathbb{K}[X]$. Si $A = PA_1$ et $B = PB_1$ avec $P, A_1, B_1 \in \mathbb{K}[X]$ et P unitaire, alors : $A \vee B = (A_1 \vee B_1)P$.*

Démonstration. Soient $M = A \vee B$ et $M_1 = A_1 \vee B_1$. Comme $A_1 | M_1$ et $B_1 | M_1$, on a $A = PA_1 | PM_1$ et $B = PB_1 | PM_1$. Donc, PM_1 est un multiple commun à A et B, et par suite $M | PM_1$.
Réciproquement, puisque M est un multiple de $A = PA_1$, on peut écrire $M = PQ$. On a alors $PA_1 | PQ$ et $PB_1 | PQ$ ce qui entraine, puisque $P \neq 0$, $A_1 | Q$ et $B_1 | Q$. Donc, $M_1 | Q$. Ce qui prouve que $PM_1 | PQ = M$.
En conclusion $M = PM_1$, puisque ce sont deux polynômes unitaires (ou nuls) associés.

□

2.1.11 Polynômes premiers entre eux

Définition 2.1.44 (Polynômes premiers entre eux)**.** Soient $A, B \in \mathbb{K}[X]$. On dit que A et B sont premiers entre eux lorsque leurs seuls diviseurs communs sont les constantes non nulles, autrement dit lorsque $A \wedge B = 1$.

Exemple 2.1.45. Pour tous a et b dans $\mathbb{K}$, $(X-a)\wedge(X-b)=1$ si et seulement si $a\neq b$.
▶ Prouvons ($\Rightarrow$) par contraposition : si $a=b$, on a $(X-a)\wedge(X-b)=X-a\neq 1$. Prouvons l'implication ($\Leftarrow$). Soit D un diviseur commun à $X-a$ et $X-b$. Comme D divise aussi $(X-b)-(X-a)=a-b\neq 0$, D est une constante non nulle. Ainsi $(X-a)\wedge(X-b)=1$.

Théorème 2.1.46 (Théorème de Bezout)**.** *Soient $A,B\in\mathbb{K}[X]$. Alors $A\wedge B=1$ si et seulement si $\exists U,V\in\mathbb{K}[X]$, $AU+BV=1$.*

Démonstration. Posons $D=A\wedge B$.
Si $D=1$, alors l'algorithme d'Euclide et plus spécialement de sa remontée, assure l'existence de deux polynômes U et V tels que $AU+BV=1$.
Réciproquement, supposons qu'il existe deux polynômes U et V tels que $AU+BV=1$. Ainsi, un diviseur unitaire commun à A et B divise encore $AU+BV=1$. Ceci montre que $D=1$. □

▶ On précisera ce théorème dans l'exercice suivant : on montrera que si $\deg(A)>1$ et $\deg(B)>1$, on peut imposer au couple (U,V) la condition supplémentaire $\deg(U)<\deg(B)$ et $\deg(V)<\deg(A)$ et qu'un tel couple (U,V) est unique.

Exercice 2.1.47. Étant donné des polynômes non constants A et B premiers entre eux, montrer qu'il existe un unique couple de polynômes (U_0,V_0) tel que :

$$AU_0+BV_0=1 \quad \text{avec} \quad \deg(U_0)<\deg(B) \quad \text{et} \quad \deg(V_0)<\deg(A).$$

Solution.
— **Unicité :** Si (U_1,V_1) et (U_2,V_2) sont deux tels couples, on a :

$$(U_1-U_2)A=(V_2-V_1)B.$$

Le polynôme A divise donc $(V_2-V-1)B$, et comme $A\wedge B=1$, le théorème de Gauss entraîne $A\mid(V_2-V_1)$. Or $\deg(V_2-V_1)<\deg A$. Donc $V_2-V_1=0$. Ainsi $V_1=V_2$ et par suite $U_1=U_2$ puisque $A\neq 0$.
— **Existence :** Soit (U,V) un couple de coefficients de Bézout pour A et B. Notons Q le quotient de la division euclidienne de U par B. L'égalité $AU+BV=1$ nous donne $A(U-QB)+B(V+QA)=1$, donc

$$AU_0+BV_0=1 \quad \text{avec} \quad U_0=(U-QB) \quad \text{et} \quad V_0=(V+QA).$$

Comme U_0 est le reste de la division euclidienne de U par B, on a $\deg U_0<\deg B$. D'autre part, puisque B n'est pas constant, le polynôme U_0 ne peut pas être nul (sinon on aurait $BV_0=1$) et donc $\deg(AU_0)\geq 1$. Alors :

$$\deg(BV_0)=\deg(1-AU_0)=\deg(AU_0),$$

ce qui donne $\deg V_0=\deg A+\deg U_0-\deg B<\deg A$.

Théorème 2.1.48 (Théorème de Gauss)**.** *Soient $A,B,C\in\mathbb{K}[X]$. Alors :*

$$\Big(A\wedge B=1 \quad \textit{et} \quad A|BC\Big)\Rightarrow\Big(A|C\Big).$$

Démonstration. Puisque $A|BC$, il existe $A_1\in\mathbb{K}[X]$ tel que $BC=AA_1$. Comme $A\wedge B=1$, il existe $(U,V)\in\mathbb{K}[X]^2$ tel que $AU+BV=1$. On multiple les deux membres de la dernière égalité par C et on obtient $C=ACU+BCV=ACU+AA_1V=A(CU+A_1V)$ et donc $A|C$. □

Lemme 2.1.49. *Soient A_1, A_2 et B dans $\mathbb{K}[X]$ tels que $A_1 \wedge A_2 = 1$, $A_1|B$ et $A_2|B$. Alors $A_1A_2|B$.*

Démonstration. Comme $A_1|B$, il existe $Q_1 \in \mathbb{K}[X]$ tel que $B = A_1Q_1$. Puisque $A_1 \wedge A_2 = 1$ et $A_2|B$, on déduit du théorème de Gauss que $A_2|Q_1$ et donc que $A_1A_2|B$. □

On en déduit le résultat suivant par une récurrence évidente sur l'entier $m \in \mathbb{N}^*$.

Proposition 2.1.50. *Soient $A_1, A_2, \cdots, A_m$ et B dans $\mathbb{K}[X]$ tels que, pour tout $i, j \in \{1, \cdots, m\}$ vérifiant $i \neq j$, $A_i \wedge A_j = 1$ et $A_i|B$. Alors $A_1A_2 \cdots A_m|B$.*

Lemme 2.1.51. *Si A et B sont deux polynômes premiers entre eux, alors $A \vee B$ et AB sont associés.*

Démonstration. Posons $P = A \vee B$. Il est évident que $P|AB$.
Réciproquement, on a $A|P$ et $B|P$. Il existe donc $Q_1 \in \mathbb{K}[X]$ tel que $P = BQ_1$. Comme $A|P$ et $A \wedge B = 1$, on en déduit d'après le théorème de Gauss que $A|Q_1$. Il existe donc un polynôme $Q_2 \in \mathbb{K}[X]$ tel que $Q_1 = AQ_2$, ce qui donne $P = ABQ_2$. D'où $AB|P$. Ce qui prouve que P et AB sont associés. □

Théorème 2.1.52 (Lien entre PPCM et PGCD)**.** *Soient $A, B \in \mathbb{K}[X]$. Les polynômes AB et $(A \wedge B)(A \vee B)$ sont associés.*

Démonstration. Le résultat étant évident si $AB = 0$, on peut supposer A et B non nuls, et unitaires quitte à les diviser par leurs coefficients dominants.
Soit $D = A \wedge B$. Prenons A_1 et B_1 tels que $A = DA_1$ et $B = DB_1$ et $A_1 \wedge B_1 = 1$. Comme A_1 et B_1 sont unitaires et premiers entre eux, on a d'après le lemme précédent : $A_1 \vee B_1 = A_1B_1$. Alors, $A \vee B = (DA_1) \vee (DB_1) = D(A_1 \vee B_1) = DA_1B_1$ et par suite $D(A \vee B) = AB$. □

Théorème 2.1.53. *Soient $A, B, C \in \mathbb{K}[X]$. On a :*

$$\Big(A \wedge B = 1 \quad et \quad A \wedge C = 1\Big) \Leftrightarrow A \wedge (BC) = 1.$$

Démonstration. ($\Rightarrow$) Supposons $A \wedge B = A \wedge C = 1$. Il existe donc des polynômes U_1, V_1, U_2, V_2 tels que : $AU_1 + BV_1 = 1$ et $AU_2 + CV_2 = 1$. En multipliant ces deux égalités, on obtient une relation de type $AU + (BC)V = 1$, ce qui prouve que $A \wedge (BC) = 1$.
($\Leftarrow$) La réciproque est évidente, puisque $A \wedge B$ et $A \wedge C$ sont des diviseurs communs a A et BC. □

2.1.12 Racines d'un polynôme

Définition 2.1.54 (Racine d'un polynôme)**.** Soient $P \in \mathbb{K}[X]$ et $\lambda \in \mathbb{K}$. On dit que λ est une racine de P (dans $\mathbb{K}$) lorsque $P(\lambda) = 0$, autrement dit lorsque λ est un zéro de la fonction $\widetilde{P} : \mathbb{K} \to \mathbb{K}$.

Théorème 2.1.55. *Soient $P \in \mathbb{K}[X]$ et $\lambda \in \mathbb{K}$. λ est une racine de P si et seulement si $X - \lambda|P$.*

Démonstration. Effectuons la division euclidienne de P par $X-\lambda \neq 0$: on a vu qu'elle s'écrit $P=(X-\lambda)Q+P(\lambda)$. Comme $X-\lambda|P$ si et seulement si le reste R dans la division euclidienne de P par $X-\lambda$ est nul, on voit directement que $X-\lambda|P$ si et seulement si $P(\lambda)=0$, c'est-à-dire λ est racine de P. □

Le résultat précédent se généralise au cas de $m \geq 1$ racines par des arguments d'arithmétique.

Proposition 2.1.56. *Si $a_1, \cdots, a_m$ sont des racines deux à deux distinctes de $P \in \mathbb{K}[X]$, alors $(X-a_1)\cdots(X-a_m)|P$.*

Démonstration. Puisque $a_i \neq a_j$ pour $i \neq j$, les polynômes $X-a_i$ sont deux à deux premiers entre eux et divisent P, on en déduit que : $(X-a_1)\cdots(X-a_m)|P$. □

Nous sommes maintenant en mesure d'établir le résultat central de ce paragraphe : le degré d'un polynôme non nul P de $\mathbb{K}[X]$ est un majorant du nombre de ses racines dans $\mathbb{K}$.

Théorème 2.1.57. *Tout polynôme P non nul à coefficients dans $\mathbb{K}$ admet au plus $\deg(P)$ racines dans $\mathbb{K}$.*

Démonstration. Si P n'admet aucune racine dans $\mathbb{K}$, le résultat est clair. Dans le cas contraire, notons $a_1, \cdots, a_m$ les racines (deux à deux distinctes) dans $\mathbb{K}$ du polynôme P. D'après la proposition précédente, il existe $Q \in \mathbb{K}[X]$ tel que $P=(X-a_1)\cdots(X-a_m)Q$, ainsi $\deg(P)=m+\deg(Q)$ et comme $(P\neq 0)\Rightarrow(Q\neq 0)$, on a $\deg(Q)\geq 0$, d'où l'inégalité $m\leq \deg(P)$. □

▶ On en déduit immédiatement que le seul polynôme admettant une infinité de racines dans $\mathbb{K}$ est le polynôme nul.

Remarque 2.1.58. On utilise souvent le théorème précédent pour démontrer qu'un polynôme P est nul :
- soit en exhibant $n+1$ racines lorsque l'on sait que $\deg(P)\leq n$,
- soit, plus radicalement, en exhibant une infinité de racines de P.

Exemple 2.1.59. Déterminons les polynômes P de $\mathbb{K}[X]$ vérifiant $P(X+1)=P(X)$.
▶ Raisonnons par analyse-synthèse. Soit P un tel polynôme et posons $Q=P-P(0)$. On prouve sans peine que $Q(X+1)=Q(X)$ et $Q(0)=0$. On en déduit par une récurrence immédiate que $\forall n\in\mathbb{N}, Q(n)=0$. Le polynôme Q admet donc une infinité de racines : $Q=0$ et donc P est constant. Réciproquement, il est clair qu'un polynôme P constant vérifie $P(X+1)=P(X)$.

▶ Nous avons maintenant les outils pour mieux comprendre les liens entre polynômes et fonctions polynomiales dans le cas où le corps de base est $\mathbb{K}=\mathbb{R}$ ou $\mathbb{C}$.

Théorème 2.1.60 (Identification polynôme/fonction polynômiale)**.** *Pour tous $P,Q\in\mathbb{K}[X]$, si les fonctions polynômiales $\widetilde{P}$ et $\widetilde{Q}$ sont égales, alors les polynômes P et Q eux mêmes le sont.*

On rappelle que deux polynômes sont égaux si et seulement si ils ont les mêmes coefficients.

Démonstration. Si $\widetilde{P} = \widetilde{Q}$, la fonction $\widetilde{P-Q}$ est nulle sur $\mathbb{K}$, donc tout élément de $\mathbb{K}$ est racine de $P-Q$. Comme $\mathbb{K}$ ($=\mathbb{R}$ ou $\mathbb{C}$) est infini, $P-Q$ possède ainsi une **infinité** de racines, donc $P-Q=0$, i.e. : $P=Q$. □

Remarque 2.1.61. Autrement dit, deux polynômes à coefficients dans $\mathbb{R}$ ou $\mathbb{C}$ sont égaux si et seulement si leurs fonctions polynomiales associées sont égales. Le lecteur aura remarqué que le caractère infini du corps de base $\mathbb{K}$ joue un rôle central dans cette démonstration. En effet, si l'on travaille par exemple sur $\mathbb{K}=\mathbb{Z}/2\mathbb{Z}=\{\bar{0},\bar{1}\}$, les polynômes $P_1=X^3+X$ et $P_2=X^2+X$ sont distincts bien que $\widetilde{P}_1=\widetilde{P}_2$.

Définition 2.1.62 (Multiplicité d'une racine)**.** Soient $\lambda\in\mathbb{K}$ et $P\in\mathbb{K}[X]$ **non nul**.

- L'ensemble $\{k\in\mathbb{N}|(X-\lambda)^k \text{ divise } P\}$ possède un plus grand élément m appelé la multiplicité de λ dans P. On dit souvent pour résumer que m est la plus grande puissance de $X-\lambda$ qui divise P. En particulier, dire que λ n'est pas racine de P, c'est dire que λ a pour multiplicité 0 dans P. Une racine est dite *simple* si elle est de multiplicité 1, *double* si elle est de multiplicité 2.
- Plus concrètement, m est caractérisé par les deux propositions suivantes, équivalentes :
  - P est divisible par $(X-\lambda)^m$ mais **PAS** par $(X-\lambda)^{m+1}$.
  - Il existe $Q\in\mathbb{K}[X]$ pour lequel : $P=(X-\lambda)^mQ$ et $Q(\lambda)\neq 0$.

Démonstration. Pour montrer qe l'ensemble $\mathcal{M}=\{k\in\mathbb{N}|(X-\lambda)^k \text{ divise } P\}$ possède un plus grand élément, nous allons montrer que c'est une partie non vide majorée de $\mathbb{N}$. Or déjà, $\mathcal{M}$ contient 0. Montrons ensuite que $\deg(P)$ majore $\mathcal{M}$. Pour tout $k\in\mathcal{M}$: $P=(X-\lambda)^kQ$ pour un certain $Q\in\mathbb{K}[X]$ avec $Q\neq 0$ car $P\neq 0$. En particulier : $\deg(Q)\geq 0$, donc $k\leq\deg(Q)+k=\deg(P)$. □

Théorème 2.1.63 (Utilisation des dérivées successives pour le calcul d'une multiplicité)**.** *Soient $P\in\mathbb{K}[X]$, $\lambda\in\mathbb{K}$ et $m\in\mathbb{N}$.*
λ est de multiplicité m dans P si et seulement si, $P^{(i)}(\lambda)=0$ pour tout $i\in\{0,\cdots,m-1\}$ ***MAIS :*** *$P^{(m)}(\lambda)\neq 0$.*

Démonstration. $\Rightarrow$ | Supposons λ de multiplicité m dans P. Dans ce cas : $P=(X-\lambda)^mQ$ pour un certain $Q\in\mathbb{K}[X]$ avec : $Q(\lambda)\neq 0$ (première division euclidienne de P par $(X-\lambda)^m$). Mais par ailleurs, Taylor donne (deuxième division euclidienne) :

$$P=\sum_{i=0}^{+\infty}\frac{P^{(i)}(\lambda)}{i!}(X-\lambda)^i=(X-\lambda)^m\sum_{i=m}^{+\infty}\frac{P^{(i)}(\lambda)}{i!}(X-\lambda)^{i-m}+\underbrace{\sum_{i=0}^{m-1}\frac{P^{(i)}(\lambda)}{i!}(X-\lambda)^i}_{\text{degré}<m}.$$

Par unicité de la division euclidienne :

$$\begin{cases}\sum_{i=0}^{m-1}\frac{P^{(i)}(\lambda)}{i!}(X-\lambda)^i=0 & (\clubsuit)\\ Q=\sum_{i=m}^{+\infty}\frac{P^{(i)}(\lambda)}{i!}(X-\lambda)^{i-m} & (\spadesuit)\end{cases}$$

Composons $\clubsuit$ à droite par $X+\lambda$: $\sum_{i=0}^{m-1}\frac{P^{(i)}(\lambda)}{i!}X^i=0$ et donc par identification : $P^{(i)}(\lambda)=0$ pour tout $i\in\{0,\cdots,m-1\}$.
Evaluons $\spadesuit$ en λ : $Q(\lambda)=\frac{P^{(m)}(\lambda)}{m!}$, donc : $P^{(m)}(\lambda)\neq 0$, car $Q(\lambda)\neq 0$.

$\Leftarrow$ | Supposons réciproquement que : $P(\lambda) = P'(\lambda) = \cdots = P^{(m-1)}(\lambda) = 0$ mais $P^{(m)}(\lambda) \neq 0$. D'après la formule de Taylor, on a : $P = \sum_{i=0}^{+\infty} \frac{P^{(i)}(\lambda)}{i!}(X-\lambda)^i = \sum_{i=m}^{+\infty} \frac{P^{(i)}(\lambda)}{i!}(X-\lambda)^i = (X-\lambda)^m \sum_{i=0}^{+\infty} \frac{P^{(i)}(\lambda)}{i!}(X-\lambda)^{i-m} = (X-\lambda)^m Q$, où l'on a posé $Q = \sum_{i=0}^{+\infty} \frac{P^{(i)}(\lambda)}{i!}(X-\lambda)^{i-m}$. Comme $Q(\lambda) = \frac{P^{(m)}(\lambda)}{m!} \neq 0$ par hypothèse, λ est bien de multiplicité m dans P.

□

Les racines appartenant à $\mathbb{C} \setminus \mathbb{R}$ d'un polynôme à coefficients réels vont toujours par deux : on peut les regrouper par paires de racines conjugués. Ce résultat repose sur le lemme suivant.

Lemme 2.1.64. *Pour tous $Q \in \mathbb{R}[X]$ et $\lambda \in \mathbb{C}$, $\overline{Q(\lambda)} = Q(\bar{\lambda})$.*

Démonstration. Écrivons $Q = \sum_{k=0}^{n} a_k X^k$ où les a_k sont réels. On a : $\overline{Q(\lambda)} = \sum_{k=0}^{n} \overline{a_k \lambda^k} = \sum_{k=0}^{n} a_k \overline{\lambda^k} = \sum_{k=0}^{n} a_k \bar{\lambda}^k = Q(\bar{\lambda})$, car les a_k sont réels. □

Théorème 2.1.65 (Racines complexes d'un polynôme réel)**.** *Soient $P \in \mathbb{R}[X]$, $n \in \mathbb{N}^*$ et $\lambda \in \mathbb{C}$. Le nombre λ est une racine de P de multiplicité n si et seulement si le nombre $\bar{\lambda}$ est une racine de P de multiplicité n.*

▶ Ce résultat va vous faire économiser la moitié de vos calculs. En effet, si vous parvenez à trouver une racine complexe (et non réelle) d'un polynôme réel, alors, automatiquement et sans calcul, vous en avez une autre : sa conjugué.

Démonstration. On déduit du lemme précédent que, pour tout entier naturel k, $P^{(k)}(\bar{\lambda}) = \overline{P^{(k)}(\lambda)}$. Ainsi, $P^{(k)}(\lambda) = 0$ si et seulement si $P^{(k)}(\bar{\lambda}) = 0$. D'où le résultat. □

Théorème 2.1.66 (Factorisation "par les racines")**.** *Soient $P \in \mathbb{K}[X]$* ***NON NUL*** *et $\lambda_1, \cdots, \lambda_r$ des racines distinctes de P de multiplicité respectives $m_1, \cdots, m_r$. Alors $(X-\lambda_1)^{m_1} \cdots (X-\lambda_r)^{m_r}$ divise P. En particulier :*

$$\sum_{k=1}^{r} m_k \leq \deg(P).$$

Démonstration. Puisque $\lambda_i \neq \lambda_j$ pour $i \neq j$, les polynômes $X - \lambda_i$ sont deux à deux premiers entre eux, et il en est donc de même des polynômes $(X-\lambda_i)^{m_i}$ qui divisent P. On en déduit que $(X-\lambda_1)^{m_1} \cdots (X-\lambda_r)^{m_r} | P$. □

▶ Lorsque l'on dénombre les racines d'un polynôme, on peut :
— soit compter le nombre de racines distinctes,
— soit compter chaque racine avec (c'est-à-dire autant de fois que) son ordre de multiplicité ; dans ce cas, une racine d'ordre r compte comme r racines.

Exemple 2.1.67. Le polynôme $(X-1)(X+1)^2(X-2)^3$ possède :
— 3 racines distinctes : -1,1,2.
— 6 racines comptées avec leur ordre de multiplicité : -1,-1,1,2,2,2.

Conséquence : Il découle du théorème précédent que : Un polynôme **NON NUL** de degré n possède au plus n racines comptées avec leur ordre de multiplicité.

Exemple 2.1.68. Déterminer les racines du polynôme $P = X^6 + X^5 + 3X^4 + 2X^3 + 3X^2 + X + 1$ sachant que i en est une racine multiple.
▶ On a $P(i) = P'(i) = 0$. mais $P''(i) = -8i \neq 0$. Le nombre i est donc une racine de P de multiplicité deux. Puisque P est à coefficients réels, $-i$ est également une racine de P de multiplicité deux. P est donc divisible par $(X-i)^2(X+i)^2 = (X^2+1)^2$. En posant la division euclidienne, on trouve $P = (X^2+1)^2(X^2+X+1)$. Comme $X^2+X+1 = (X-j)(X-j^2)$, les racines de P sont $i, -i$ (racines d'ordre deux) et j, j^2 (racines simples).

Remarque 2.1.69. Pour montrer qu'un polynôme P est divisible par un polynôme Q, il suffit de montrer que toutes las racines de Q dans $\mathbb{C}$ sont des racines de P avec un ordre de multiplicité au moins égal.

Exercice 2.1.70. A quelle condition nécessaire et suffisante sur n le polynôme X^2+1 divise-t-il X^n+1 ?

Solution. Pour tout $n \in \mathbb{N}$:

$$\begin{aligned}
&X^2+1 \,\text{divise}\, X^n+1 \Leftrightarrow i \,\text{et}\, -i \,\text{sont racines de}\, X^n+1\\
&\Leftrightarrow i \,\text{est racine de}\, X^n+1 \qquad (\text{car}\, X^n+1 \in \mathbb{R}[X])\\
&\Leftrightarrow i^n+1=0 \Leftrightarrow e^{\frac{in\pi}{2}} = e^{i\pi}\\
&\Leftrightarrow \frac{n\pi}{2} \equiv \pi[2\pi] \Leftrightarrow n \equiv 2[4].
\end{aligned}$$

Exercice 2.1.71. Soit $n \in \mathbb{N}$. Montrer que le polynôme

$$(X-1)^{n+2} + X^{2n+1}$$

est divisible par $X^2 - X + 1$.

Solution. Le polynôme $X^2 - X + 1$ admet deux racines simples distinctes $-j$ et $-j^2$. Il suffit donc de montrer que $-j$ et $-j^2$ sont racines de $(X-1)^{n+2} + X^{2n+1}$. Or, si α désigne l'une des racines de $X^2 - X + 1$, on a

$$(\alpha-1)^{n+2} + \alpha^{2n+1} = \alpha^{2n+4} + \alpha^{2n+1} = \alpha^{2n+1}(\alpha^3+1) = 0,$$

puisque $\alpha - 1 = \alpha^3$. Ce qui montre que $X^2 - X + 1$ divise $(X-1)^{n+2} + X^{2n+1}$.

Remarque 2.1.72. Pour montrer qu'un polynôme P est divisible par $(X-a)^m$, on montre que a est une racine d'ordre de multiplicité au moins m de P, et pour cela il suffit de prouver que : $P(a) = P'(a) = \cdots = P^{(m-1)}(a) = 0$.

Exercice 2.1.73. Soit $n \in \mathbb{N}$, $n \geq 2$. On considère le polynôme P défini par $P = aX^{n+1} + bX^n + 1$. Déterminer les réels a et b pour que P soit divisible par $(X-1)^2$.

Solution.

$$\begin{aligned}
P \,\text{est divisible par}\, (X-1)^2 &\Leftrightarrow 1 \,\text{est une racine d'ordre au moins}\; 2 \,\text{de}\, P\\
&\Leftrightarrow P(1) = P'(1) = 0\\
&\Leftrightarrow \begin{cases} a+b+1=0 \\ (n+1)a+nb=0 \end{cases}\\
&\Leftrightarrow \begin{cases} a=n \\ b=-1-n. \end{cases}
\end{aligned}$$

Exercice 2.1.74.
1. Quel est le reste de la division euclidienne du polynôme $A = X^n + X + b$, $n \in \mathbb{N}^*$ par $B = (X-a)^2$ où $a, b \in \mathbb{R}$.
2. Trouver a et b réels tel que le polynôme $P = X^3 + aX + b$ admette le nombre $z = 1 + i$ comme racine.
3. Montrer sans faire la division que $A = 2X^2 + X^3 - 6X^2 - X + 4$ est divisible par $B = X^2 - 1$.
4. Déterminer $n \in \mathbb{N}$ pour que le polynôme $(X+1)^{2n+1} + X^{2n+2}$ soit divisible par le polynôme $X^2 + X + 1$.
5. Soient m, n et $p \in \mathbb{N}$. Montrer que le polynôme $X^{3m} + X^{3n+1} + X^{3p+2}$ divisible par $X^2 + X + 1$.

2.1.13 Polynômes scindés et Théorème de d'Alembert-Gauss

Nous admettrons le "**théorème fondamental de l'algèbre**" :

Théorème 2.1.75 (d'Alembert-Gauss)**.** *Tout polynôme non constant de* $\mathbb{C}[X]$ *possède au moins une racine dans* $\mathbb{C}$*. On dit que* $\mathbb{C}$ *est un corps* ***algébriquement clos****.*

Corollaire 2.1.76. *Soit* $P \in \mathbb{C}[X]$ *non constant de degré* n*. Il existe des nombres complexes* $\lambda_1, \cdots, \lambda_n$*,* A *tels que :* $P = A\prod_{k=1}^{n}(X - \lambda_k)$*.*

Démonstration. Prouvons la propriété par récurrence sur $n \geq 1$. Le résultat est clair au rang 1. Supposons le résultat acquis au rang $n \geq 1$ et considérons un polynôme P de degré $n+1$. D'après le théorème de d'Alembert-Gauss, P admet une racine $\lambda_{n+1} \in \mathbb{C}$. Ainsi, il existe $Q \in \mathbb{C}[X]$ tel que $P = (X - \lambda_{n+1})Q$. Comme $\deg(Q) = n$, on déduit de l'hypothèse au rang n l'existence de $\lambda_1, \cdots, \lambda_n$ et A dans $\mathbb{C}$ tels que : $Q = A\prod_{k=1}^{n}(X - \lambda_k)$. D'où : $P = A\prod_{k=1}^{n+1}(X - \lambda_k)$.
L'hypothèse est donc vraie au rang $n+1$. On déduit du principe de récurrence que la propriété est vraie pour tout entier naturel non nul n. □

Définition 2.1.77 (Polynôme scindé)**.** Soit $P \in \mathbb{K}[X]$. On dit que P est scindé (sur $\mathbb{K}$) s'il **N'est PAS CONSTANT** et possède exactement $\deg(P)$ racines (dans $\mathbb{K}$) comptées avec multiplicité.
Dire que P est scindé sur $\mathbb{K}$ revient donc à dire que P est de la forme : $P = A\prod_{k=1}^{r}(X - \lambda_k)^{m_k}$, où $\lambda_1, \cdots, \lambda_r$ sont les racines distinctes de P dans $\mathbb{K}$, de multiplicité respectives $m_1, \cdots, m_r$, et où A est son coefficient dominant.

► Autrement dit, un polynôme est scindé sur $\mathbb{K}$ si et seulement si il est le produit de polynômes de $\mathbb{K}[X]$ de degré un. On peut donc reformuler ainsi le théorème de d'Alembert-Gauss : tout polynôme non constant de $\mathbb{C}[X]$ est scindé sur $\mathbb{C}$.

Remarque 2.1.78. Deux polynômes P et Q de $\mathbb{K}[X]$ scindés sur $\mathbb{K}$ sont premiers entre eux si et seulement s'ils n'ont aucune racine commune.

2.2 Factorisation irréductible sur $\mathbb{R}$ ou $\mathbb{C}$

L'objectif de cette partie est de définir une classe de polynômes qui joueraient le même rôle que les nombres premiers sur $\mathbb{Z}$, puis d'en déduire l'analogue polynomial du théorème fondamental de l'arithmétique des nombres entiers.

Définition 2.2.1 (Polynôme irréductible)**.** Un polynôme $P \in \mathbb{K}[X]$ est dit irréductible sur $\mathbb{K}$ si P **n'est PAS CONSTANT** et ses seuls diviseurs dans $\mathbb{K}[X]$ sont les polynômes constants et ceux de la forme λP, où $\lambda \in \mathbb{K}^*$.

Explication : Un polynôme P est irréductible s'il vérifie :
— $\deg(P) \geq 1$,
— les seuls diviseurs de P sont les éléments de $\mathbb{K}^*$ et les associés de P,
c-à-d tel que P soit non constant et que pour tout $A, B \in \mathbb{K}[X]$, on ait :

$$P = AB \Rightarrow \Big(\deg(A) = 0 \text{ ou } \deg(B) = 0\Big).$$

Avant de décrire tous les irréductibles sur $\mathbb{R}$ et $\mathbb{C}$, nous passerons en revue quelques exemples.

Exemple 2.2.2. Tout polynôme de degré un est irréductible sur $\mathbb{K}$.
▶ Soit P un polynôme de degré 1 et Q un diviseur de P. Il existe $Q_1 \in \mathbb{K}[X]$ tel que P $= QQ_1$. Donc $\deg(Q) \leq \deg(P)$. Ainsi Q est de degré 0 ou 1.
— Si $\deg(Q) = 0$, Q est constant non nul.
— Si $\deg(Q) = 1$, alors : $\deg(Q_1) = 0$, i.e. Q_1 est constant non nul a. Donc Q s'écrit : $Q = \frac{1}{a}P$.
P n'est pas constant et ses seuls diviseurs sont les constantes non nulles et les polynômes de la forme λP, avec $\lambda \in \mathbb{K}^*$: P est irréductible sur $\mathbb{K}$.

Exemple 2.2.3. Un polynôme de degré 2 est irréductible sur $\mathbb{K}$ si et seulement si il n'a pas de racine dans $\mathbb{K}$.
▶ Soit $P \in \mathbb{K}[X]$ de degré 2.
($\Rightarrow$) Raisonnons par contraposée et supposons que P admet une racine a dans $\mathbb{K}$. Alors $X - a | P$ et donc P n'est pas irréductible.
($\Leftarrow$) Raisonnons par contraposée et supposons que P n'est pas irréductible. Il admet donc un diviseur unitaire Q qui est de degré 1. Ainsi, il existe $a \in \mathbb{K}$ tel que $Q = X - a | P$ d'où $P(a) = 0$. Le polynôme P admet donc une racine dans $\mathbb{K}$.

Remarque 2.2.4. Un polynôme qui n'admet pas de racine dans $\mathbb{K}$ n'est pas nécessairement irréductible dans $\mathbb{K}[X]$, comme le prouve l'exemple de $(X^2+1)^2$ dans $\mathbb{R}[X]$.

Proposition 2.2.5. *Les polynômes irréductibles de* $\mathbb{C}[X]$ *sont les polynômes de degré* 1.

Démonstration. On sait que tout polynôme de degré 1 est irréductible.
Réciproquement, soit $P \in \mathbb{C}[X]$ irréductible. Etant **NON CONSTANT**, d'après le théorème de d'Alembert-Gauss, P possède une racine $a \in \mathbb{C}$ et donc $X - a$ divise P. L'irréductibilité de P sur $\mathbb{C}$ montre alors que P et $X - a$ sont associés, donc que P est de degré 1. □

Proposition 2.2.6. *Les polynômes irréductibles de* $\mathbb{R}[X]$ *sont :*
— *les polynômes de degré* 1*,*
— *les polynômes de degré* 2 *à discriminant strictement négatif, i.e. sans racine réelle.*

Démonstration. Soit $P \in \mathbb{R}[X]$ irréductible. Etant **NON CONSTANT**, d'après le théorème de d'Alembert-Gaus, P possède une racine a *COMPLEXE*. De plus, P étant à coefficients réels, alors $\bar{a}$ est aussi une racine de P.

— Si $a \in \mathbb{R}$, $X - a$ divise P dans $\mathbb{R}[X]$. Or, P est irréductible sur $\mathbb{R}$, donc P et $X - a$ sont associés et P est de degré 1.
— Si $a \notin \mathbb{R}$, alors $\bar{a} \neq a$, donc : $P = (X - a)(X - \bar{a})Q$ pour un certain $Q \in \mathbb{C}[X]$. En développant, on obtient : $P = (X^2 - 2\,\mathrm{Re}(a)X + |a|^2)Q$, donc en réalité Q est à coefficients **RÉELS** par unicité de la division euclidienne dans $\mathbb{C}[X]$. De là, P étant irréductible sur $\mathbb{R}$, P et $X^2 - 2\,\mathrm{Re}(a)X + |a|^2$ sont associés et P est de degré 2. De plus, P est bien sans racine réelle.

□

Lemme 2.2.7 (Euclide). *Soit $P \in \mathbb{K}[X]$ un polynôme irréductible. Si P divise un produit $\prod_{k=1}^{r} Q_k$ de $r \geq 2$ polynômes non nuls, alors il divise l'un des Q_k.*

Démonstration. On procède par récurrence sur $r \geq 2$.
— **Initialisation :** Soit P irréductible dans $\mathbb{K}[X]$. Supposons que $P | Q_1 Q_2$, où Q_1 et Q_2 sont deux polynômes non nuls. Quitte à diviser par le coefficient dominant, on peut supposer que P est unitaire. Soit $D = P \wedge Q_1$. D étant un diviseur de P, il est égal à 1 ou P, puisque P est irréductible. Si $D = P$ alors $P | Q_1$. Si $D = 1$, alors $\exists U, V \in \mathbb{K}[X]$ tels que $UP + VQ_1 = 1$ et donc $P | Q_2 = UPQ_2 + VQ_1Q_2$.
— **Hérédité :** Supposons le résultat acquis pour $r - 1 \geq 2$ et soit P irréductible qui divise $\prod_{k=1}^{r} Q_k$. Si $P | Q_r$, c'est terminé, sinon il divise $\prod_{k=1}^{r-1} Q_k$ (cas $r = 2$) et l'hypothèse de récurrence permet de conclure.

□

Théorème 2.2.8 (Existence et unicité de la factorisation irréductible). *Tout polynôme non constant $P \in \mathbb{K}[X]$ est produit de polynômes irréductibles et une telle décomposition est unique à l'ordre près des facteurs, ce qui signifie qu'il existe une constante $A \in \mathbb{K}^*$, un entier $r \geq 1$, des polynômes unitaires $P_1, \cdots, P_r$ deux à deux distincts et irréductibles et des entiers naturels non nuls $m_1, \cdots, m_r$ tels que $P = A \prod_{k=1}^{r} P_k^{m_k}$.*
L'unicité signifie que si on a une autre décomposition de même type $P = B \prod_{k=1}^{s} Q_k^{n_k}$, on a alors $B = A$, $s = r$ et il existe une permutation $\sigma \in \mathcal{S}_r$ telle que $Q_k = P_{\sigma(k)}$ et $n_k = m_{\sigma(k)}$ pour tout k compris entre 1 et r.

Démonstration. **Existence :** On procède par récurrence sur le degré $n \geq 1$ de P.
— Initialisation : Pour $n = 1$, on a vu que le polynôme P est irréductible.
— Hérédité : Supposons le résultat acquis pour tous les polynômes de degré k compris entre 1 et $n - 1 \geq 1$ et soit $P \in \mathbb{K}[X]$ de degré n. Si P est irréductible, c'est terminé. Sinon, P s'écrit $P = RS$ avec R, S non constants et degré compris entre 1 et $n - 1$. Il suffit d'utiliser l'hypothèse de récurrence pour R et S.

Unicité : On procède également par récurrence sur le degré $n \geq 1$ de P.
— Initialisation : Pour $n = 1$, on a $P = A(X - \lambda)$ avec $A \in \mathbb{K}^*$ et $X - \lambda$ irréductible uniquement déterminés.
— Hérédité : Supposons le résultat acquis pour tous les polynômes de degré k compris entre 1 et $n - 1 \geq 1$. Soit $P \in \mathbb{K}[X]$ de degré n ayant deux décompositions $P = A \prod_{k=1}^{r} P_k^{m_k} = B \prod_{k=1}^{s} Q_k^{n_k}$.

L'identification des coefficients dominants nous donne $A = B$. Comme Q_1 est irréductible et divise $\prod_{k=1}^{r} P_k^{m_k}$, le lemme d'Euclide nous dit qu'il divise l'un des P_j, il est donc égal à ce P_j puisque ces deux polynômes sont irréductibles.

Supposons que $n_1 \leq m_j$. En divisant par $Q_1^{n_1}$, on a $Q_1^{m_j-n_1}\prod_{\substack{k=1\\k\neq j}}^{r} P_k^{m_k} = \prod_{k=2}^{s} Q_k^{n_k}$ et nécessairement $m_j = n_1$ (sinon le lemme d'Euclide nous dit Q_1 est égal à l'un des Q_k avec $k \geq 2$, ce qui n'est pas). On note $j = \sigma(1)$ et on a $\prod_{\substack{k=1\\k\neq j}}^{r} = \prod_{k=2}^{s} Q_k^{n_k}$. L'hypothèse de récurrence permet alors de conclure. □

► La factorisation irréductible d'un polynôme non constant de $\mathbb{C}[X]$ coincide avec sa forme scindée.

Le théorème de décomposition en facteurs irréductibles dans $\mathbb{R}[X]$ prend la forme suivante :

Théorème 2.2.9. *Tout polynôme non constant $P \in \mathbb{R}[X]$ est produit de polynômes irréductibles de degré 1 ou 2, c-à-d qu'il existe une constante $A \in \mathbb{R}^*$, deux entiers naturels r et s, des réels $a_1, \cdots, a_r$ deux à deux distincts, des entiers naturels non nuls $n_1, \cdots, n_r$, des couples réels $(b_1, c_1), \cdots, (b_s, c_s)$ deux à deux distincts tels que $b_k^2 - 4c_k < 0$ pour tout k et des entiers naturels non nuls $m_1, \cdots, m_s$ tels que $P = A\prod_{k=1}^{r}(X-a_k)^{n_k}\prod_{k=1}^{s}(X^2+b_kX+c_k)^{m_k}$, une telle décomposition est unique à l'ordre près des facteurs.*

En pratique : Pour factoriser un polynôme P dans $\mathbb{R}[X]$, on peut, si l'on ne trouve pas de méthode plus simple, factoriser d'abord ce polynôme dans $\mathbb{C}[X]$. On sait que si b est une racine de P dans $\mathbb{C}$, alors $\bar{b}$ est aussi une racine de P dans $\mathbb{C}$.
Soit $a_1, \cdots, a_r$ les racines réelles de P, $b_1, \bar{b}_1, \cdots, b_s, \bar{b}_s$ les racines comlexes (non réelles) de P, alors :
dans $\mathbb{C}[X]$: $P = A\prod_{k=1}^{r}(X-a_k)\prod_{k=1}^{s}(X-b_k)(X-\bar{b}_k)$, donc
dans $\mathbb{R}[X]$: $P = A\prod_{k=1}^{r}(X-a_k)\prod_{k=1}^{s}(X^2-2Re(b_k)X+|b_k|^2)$.
► On peut retenir que :
— e^{it} et $e^{it'}$ sont des complexes conjugués $\Leftrightarrow \exists k \in \mathbb{Z} | t + t' = 2k\pi$.
— $(X-b)(X-\bar{b}) = X^2 - 2Re(b)X + |b|^2 \in \mathbb{R}[X]$.

Exemple 2.2.10. Décomposons en produit de facteurs irréductibles sur $\mathbb{R}$ et sur $\mathbb{C}$ le polynôme $P = X^n - 1$.

► Dans $\mathbb{C}[X]$, le polynôme $X^n - 1$ est unitaire et admet n racines simples, les nombres $z_k = e^{\frac{2ik\pi}{n}}, 0 \leqslant k \leqslant n-1$. Donc,

$$X^n - 1 = \prod_{k=0}^{n-1}\left(X - e^{\frac{2ik\pi}{n}}\right).$$

► Dans $\mathbb{R}[X]$, il y a deux cas.
— Si $n = 2p, p \in \mathbb{N}^*$, $X^n - 1 = X^{2p} - 1$ admet 2 racines réelles : 1 et -1.

En regroupant les facteurs conjugués, on obtient :

$$\begin{aligned}
X^{2p}-1 &= \prod_{k=0}^{2p-1}\left(X-e^{\frac{2ik\pi}{2p}}\right)\\
&= (X-1)\times\prod_{k=1}^{p-1}\left(X-e^{\frac{ik\pi}{p}}\right)\times(X+1)\times\prod_{k=p+1}^{2p-1}\left(X-e^{\frac{ik\pi}{p}}\right)\\
&= (X-1)(X+1)\prod_{k=1}^{p-1}\left(X-e^{\frac{ik\pi}{P}}\right)\prod_{\ell=1}^{p-1}\left(X-e^{\frac{i(2p-\ell)\pi}{p}}\right)\\
&= (X-1)(X+1)\prod_{k=1}^{p-1}\left(X-e^{\frac{ik\pi}{p}}\right)\left(X-e^{-\frac{ik\pi}{p}}\right)\\
&= (X-1)(X+1)\prod_{k=1}^{p-1}\left(X^2-2X\cos(\frac{k\pi}{p})+1\right).
\end{aligned}$$

— Si $n=2p+1, p\in\mathbb{N}^*, X^n-1=X^{2p+1}-1$ admet une seule racine réelle à savoir 1. En regroupant les facteurs conjugués, on obtient

$$\begin{aligned}
X^{2p+1}-1 &= \sum_{k=0}^{2p}\left(X-e^{\frac{2ik\pi}{2p+1}}\right)\\
&= (X-1)\times\prod_{k=1}^{p}\left(X-e^{\frac{2ik\pi}{2p+1}}\right)\times\prod_{k=p+1}^{2p}\left(X-e^{\frac{2ik\pi}{2p+1}}\right)\\
&= (X-1)\prod_{k=1}^{p}\left(X-e^{\frac{2ik\pi}{2p+1}}\right)\left(X-e^{-\frac{2ik\pi}{2p+1}}\right)\\
&= (X-1)\prod_{k=1}^{p}\left(X^2-2X\cos\left(\frac{2k\pi}{2p+1}\right)+1\right).
\end{aligned}$$

Exemple 2.2.11. Décomposons en produit de facteurs irréductibles sur $\mathbb{C}$ le polynôme $P=X^{n-1}+X^{n-2}+\cdots+X+1$.

▶ Dans $\mathbb{C}[X]$, on a

$$X^n-1=\prod_{k=0}^{n-1}\left(X-e^{\frac{2ik\pi}{n}}\right)=(X-1)\prod_{k=1}^{n-1}\left(X-e^{\frac{2ik\pi}{n}}\right)$$

mais aussi

$$X^n-1=(X-1)\left(X^{n-1}+\ldots+X+1\right).$$

Après simplification par le polynôme non nul $X-1$, on obtient :

$$X^{n-1}+\ldots+X+1=\prod_{k=1}^{n-1}\left(X-e^{\frac{2ik\pi}{n}}\right).$$

Exercice 2.2.12. Déterminer la décomposition en produit de facteurs irréductibles sur $\mathbb{R}$ du polynôme $P=(X+1)^7-X^7-1$ sachant que j est une racine multiple de P.

Solution. N'oublions pas la périodicité des puissances de j :

$$\Big((j_n)_{n\in\mathbb{N}} = (1, j, j^2, 1, j, j^2, \cdots)\Big)$$

ni la relation bien connue $1 + j + j^2 = 0$.
On a $P(j) = 0$. De même, on vérifie que $P'(j) = 0$. En revanche, $P''(j) = 42 \neq 0$. Ainsi, le nombre complexe j est une racine de P de multiplicité 2. Les nombres 0 et -1 sont des racines évidentes de P. Puisque $P \in \mathbb{R}[X]$, $\bar{j} = j^2$ en est également une racine de multiplicité 2. Le polynôme P est donc divisible par $Q = X(X+1)(X-j)^2(X-j^2)^2 = X(X+1)(X^2+X+1)^2$. D'autre part, en appliquant la formule du Binôme, on obtient : $P = \sum_{k=0}^{7} C_7^k X^k - X^7 - 1 = \sum_{k=1}^{6} C_7^k X^k$. P est donc de degré 6, de coefficient dominant égal à $C_7^6 = 7$. Puisque Q est unitaire de degré 6 et qu'il divisise P, on a $P = 7X(X+1)(X^2+X+1)^2$.

Exercice 2.2.13. On considère le polynôme $P = X^5 - 1$.

1. Décomposer P dans $\mathbb{C}[X]$.
2. En déduire les racines de $Q = X^4 + X^3 + X^2 + X + 1$ dans $\mathbb{C}$.
3. Décomoser Q dans $\mathbb{R}[X]$.
4. En déduire la valeur de $r_1 = \cos(\frac{2\pi}{5})$ et $r_2 = \cos(\frac{4\pi}{5})$.

2.3 Fractions rationnelles

2.3.1 Construction

Soit $\mathbb{K}$ un corps commutatif. Nous savons que l'ensemble $\mathbb{K}[X]$ des polynômes à une indéterminée à coefficients dans $\mathbb{K}$ est un anneau commutatif intègre dans lequel les seuls éléments inversibles sont les polynômes de degré 0. $\mathbb{K}[X]$ n' est donc pas un corps, mais on peut construire le corps des fractions de $\mathbb{K}[X]$. Ce corps s'appelle le **corps des fractions rationnelles** à une indéterminée à coefficients dans $\mathbb{K}$. On le note $\mathbb{K}(X)$.

On pose comme d'habitude $\mathbb{K}[X]^* = \mathbb{K}[X] \setminus \{0\}$.

— $\mathbb{K}(X)$ est l'ensemble quotient de $\mathbb{K}[X] \times \mathbb{K}[X]^*$ par la relation d'équivalence $\mathcal{R} : (P, Q)\mathcal{R}(P_1, Q_1) \Leftrightarrow PQ_1 = P_1Q$. Une fraction rationnelle F de $\mathbb{K}(X)$ est donc une classe d'équivalence représentée par un couple (P, Q) d'élément de $\mathbb{K}[X]$ dans lequel $Q \neq 0$; un autre couple (P_1, Q_1) représente la même fraction rationnelle F si, et seulement si $PQ_1 = P_1Q$.
Si (P, Q) est un représentant quelconque de F, on convient d'écrire $F = \frac{P}{Q}$; on dit que P est le numérateur et que Q est le dénominateur de la fraction rationnelle F.

— Dans $\mathbb{K}[X] \times \mathbb{K}[X]^*$ on définit l'addition et la multiplication en posant : $(P, Q)+(P_1, Q_1) = (PQ_1+P_1Q, QQ_1)$ et $(P, Q)\cdot(P_1, Q_1) = (PP_1, QQ_1)$.

Les deux lois de $\mathbb{K}(X)$ sont les lois quotients et on les note aussi, $+$ et $\cdot$; alors le triplet $(\mathbb{K}(X), +, \cdot)$ est un corps commutatif. L'élément neutre pour l'addition est la fraction rationnelle nulle 0 qui est la classe des couples $(0, Q)$ tels que $Q \neq 0$. L'élément neutre pour la multiplication, appelée fraction rationnelle unité, et notée 1, est la classe des couples (Q, Q) avec $Q \neq 0$. Pour la multiplication, l'inverse de la fraction rationnelle non nulle F, classe de (P, Q) avec $P \neq 0$ et $Q \neq 0$ est la fraction rationnelle notée $\frac{1}{F}$.

L'application qui, à $P \in \mathbb{K}[X]$, associe la fraction rationnelle dont un représentant est le couple $(P,1)$, est un morphisme injectif de l'anneau $\mathbb{K}[X]$ dans l'anneau $\mathbb{K}(X)$. On peut donc identifier $\mathbb{K}[X]$ au sous-anneau de $\mathbb{K}(X)$ constitué par les fractions rationnelles dont les représentants sont de la forme $(P,1)$.

2.3.2 Définition, règles de calcul

On a les règles de calcul suivantes (dans lesquelles Q, Q_1 et R sont des polynômes non nuls) :

- $\frac{P}{Q} + \frac{P_1}{Q_1} = \frac{PQ_1+P_1Q}{QQ_1}$.
- $\frac{P}{Q}\frac{P_1}{Q_1} = \frac{PP_1}{QQ_1}$.
- $\frac{P}{Q} = \frac{P_1}{Q_1} \Leftrightarrow PQ_1 = P_1Q$.
- $\frac{PR}{QR} = \frac{P}{Q}$.
- Si $P \neq 0$, $\left(\frac{P}{Q}\right)^{-1} = \frac{Q}{P}$.
- Conjuguée d'une fraction rationnelle à coefficients complexes :
  Si (P,Q) et (P_1,Q_1) sont deux représentants d'une fraction rationnelle F à coefficients dans $\mathbb{C}$, alors $\frac{\overline{P}}{\overline{Q}} = \frac{\overline{P_1}}{\overline{Q_1}}$ puisque l'on a :

$$PQ_1 = P_1Q \Leftrightarrow \overline{P}\,\overline{Q_1} = \overline{P_1}\,\overline{Q}.$$

Cette fraction rationnelle est appelée conjuguée de F et notée $\overline{F}$.
Les propriétés sur les polynômes nous donnent immédiatement, pour $(F,G) \in \mathbb{C}(X)^2$:

$$\overline{F+G} = \overline{F} + \overline{G} \quad \text{et} \quad \overline{FG} = \overline{F}\,\overline{G}$$

2.3.3 Représentant irréductible

Définition 2.3.1.

- On appelle *représentant irréductible* d'une fraction rationnelle F tout représentant (P,Q) de F où P et Q sont premiers entre eux.
- On appelle *représentant irréductible unitaire* d'une fraction rationnelle F tout représentant irréductible (P,Q) de F tel que Q soit un polynôme unitaire.

Proposition 2.3.2.

- *Si $\frac{P}{Q}$ est une forme irréductible d'une fraction rationnelle $F = \frac{P_1}{Q_1}$, alors :* $\exists R \in \mathbb{K}[X] : (P_1 = RP \quad et \quad Q_1 = RQ)$.
- *Si $\frac{P}{Q}$ et $\frac{P_1}{Q_1}$ sont deux formes irréductibles d'une fraction F, alors :* $\exists \lambda \in \mathbb{K}^* : (P_1 = \lambda P \quad et \quad Q_1 = \lambda Q)$.
- *Toute fraction rationnelle admet un représentant irréductible unitaire et un seul.*

Démonstration.

- On a $PQ_1 = QP_1$ et donc Q divise PQ_1. Comme P et Q sont premiers entre eux, on en déduit d'après le théorème de Gauss que Q divise Q_1. On peut donc trouver un polynôme $R \in \mathbb{K}[X]$ tel que $Q_1 = RQ$, et l'on a : $P_1 = \frac{PQ_1}{Q} = RP$.
- Soient (P,Q) et (P_1,Q_1) deux représentants irréductibles de F. Alors, d'après le point précédent, $Q|Q_1$ et $Q_1|Q$. Les polynômes Q et Q_1 sont donc associés et par conséquent, le polynôme R est une constante non nulle.

— L'unicité est évidente d'après ce qui précède. Pour l'existence, il suffit de prendre un représentant quelconque (P, Q), de diviser P et Q par leur PGCD et de diviser le numérateur et le dénominateur par le coefficient dominant de ce dernier (qui est non nul).

□

2.3.4 Degré d'une fraction rationnelle

Définition 2.3.3. Si F est une fraction rationnelle, la quantité $\deg(P)-\deg(Q) \in \mathbb{Z} \cup \{-\infty\}$ ne dépend pas du représentant (P, Q) choisi pour la fraction F. On l'appelle degré de F et on le note $\deg(F)$ ou $\deg F$. En particulier, on a $\deg 0 = -\infty$.

▶ On peut bien définir ainsi le degré de F car :
— la quantité $\deg(P)-\deg(Q)$ est toujours définie puisque Q étant non nul, on a $\deg(Q) \neq -\infty$.
— la quantité $\deg(P)-\deg(Q)$ ne dépend pas du représentant (P, Q) choisi pour la fraction rationnelle F, car si (P_1, Q_1) est un autre représentant, on a $PQ_1 = QP_1$ et donc : $\deg(P)+\deg(Q_1) = \deg(PQ_1) = \deg(P_1Q) = \deg(P_1)+\deg(Q)$ c'est-à-dire, puisque $\deg(Q)$ et $\deg(Q_1)$ sont des entiers naturels : $\deg(P) - \deg(Q) = \deg(P_1) - \deg(Q_1)$.

Remarque 2.3.4. Un polynôme P est égal à la fraction $\frac{P}{1}$ dont le degré est $\deg(P)$. Donc la définition du degré sur $\mathbb{K}(X)$ prolonge celle du degré défini sur $\mathbb{K}[X]$.

Proposition 2.3.5. *Étant données deux fractions rationnelles F_1 et F_2 de $\mathbb{K}(X)$, on a :*

1. $\deg(F_1 + F_2) \leq \max(\deg(F_1), \deg(F_2))$.
2. $\deg(F_1F_2) = \deg(F_1) + \deg(F_2)$.

2.3.5 Racines, pôles

Définition 2.3.6. Soit F une fraction rationnelle de forme irréductible $\frac{P}{Q}$.
— On appelle racine de F toute racine de P.
— On appelle pôle de F toute racine de Q.
— Si a est une racine (respectivement un pôle) de $F \neq 0$, l'ordre de multiplicité de a est l'ordre de multiplicité de a en tant que racine du polynôme P (respectivement Q).

Remarque 2.3.7. Un élément a de $\mathbb{K}$ ne peut pas être à la fois racine et pôle d'une fraction rationnelle F. Sinon, en prenant une forme irréductible $F = \frac{P}{Q}$, on aurait $P(a) = Q(a) = 0$, et donc les polynômes P et Q seraient divisibles par $X - a$, ce qui contredirait le caractère irréductible de $\frac{P}{Q}$.

Attention : Les racines (respectivement les pôles) d'une fraction rationnelle F ne peuvent être obtenues qu'à partir d'une forme irréductible de F. Par exemple, $F = \frac{X^3-1}{X^2-1}$ n'admet 1 ni comme racine ni comme pôle, car $F = \frac{X^2+X+1}{X+1}$.

Définition 2.3.8. Soit F une fraction rationnelle, de forme irréductible $\frac{P}{Q}$, dont on désigne par A l'ensemble des pôles.
— Pour $\alpha \in \mathbb{K} \setminus A$, on définit $F(\alpha) = \frac{P(\alpha)}{Q(\alpha)}$.

— La fonction définie sur $\mathbb{K}\backslash A$ par $x \mapsto F(x)$ est appelée fonction rationnelle associée à la fraction rationnelle F.

2.3.6 Composition

Si $F = \frac{P}{Q}$ est une fraction rationnelle et R un polynôme non constant, alors le polynôme $Q \circ R$ est non nul, puisqu'il est de degré $\deg(Q)\deg(R)$. De plus, si (P_1, Q_1) est un autre représentant de la fraction F, l'égalité $PQ_1 = P_1Q$ entraîne $(P\circ R)(Q_1 \circ R) = (P_1 \circ R)(Q \circ R)$. Le quotient $\frac{P\circ R}{Q\circ R}$ ne dépend donc pas du représentant (P, Q) choisi pour la fraction rationnelle F. C'est une fraction rationnelle dont le degré vaut $\deg(F)\deg(R)$, que l'on l'appelle composée de F par R et que l'on note $F(R)$.

2.3.7 Décomposition en éléments simples

Proposition 2.3.9 (Partie entière)**.** *Toute fraction rationnelle F s'écrit de façon unique comme la somme d'un polynôme, appelé* ***partie entière*** *de F, et d'une fraction rationnelle de degré strictement négatif.*

Démonstration. **Unicité :** Supposons : $F = E_1 + F_1 = E_2 + F_2$ avec $(E_1, E_2) \in \mathbb{K}[X]^2$, $(F_1, F_2) \in \mathbb{K}(X)^2$, $\deg(F_1) < 0$ et $\deg(F_2) < 0$. Alors $E_1 - E_2$ est un polynôme, et comme il est égal à $F_2 - F_1$, son degré est strictement négatif. Donc $E_1 - E_2 = 0$ c'est-à-dire $E_1 = E_2$ et par suite $F_1 = F_2$.
Existence : Soit $F = \frac{P}{Q}$ une fraction rationnelle, avec $Q \neq 0$. Si l'on appelle E le quotient et R le reste de la division euclidienne de P par Q, on obtient : $F = E + \frac{R}{Q}$ avec $\deg(R) < \deg(Q)$, ce qui constitue l'écriture cherchée. □

Méthode 2.3.10. D'après la démonstration précédente, la partie entière d'une fraction rationnelle $F = \frac{P}{Q}$ est le quotient de la division euclidienne du numérateur P par le dénominateur Q.

Exemple 2.3.11.
— Si $\deg(F) < 0$, alors sa partie entière est nulle.
— Si F est de degré 0, alors sa partie entière est le polynôme constant quotient du coefficient dominant du numérateur par celui du dénominateur.
— La partie entière de la fraction $\frac{X^5}{(X^2+X+1)^2}$ est $X - 2$.

Proposition 2.3.12 (Partie polaire)**.** *Si F est une fraction rationnelle admettant a pour pôle d'ordre n, il existe un unique n-uplet de scalaires $(\lambda_p)_{1\leq p\leq n}$ et une unique fraction F_0 n'admettant pas a pour pôle tels que :*

$$F = \sum_{p=1}^{n} \frac{\lambda_p}{(X-a)^p} + F_0.$$

La quantité $\sum_{p=1}^{n} \frac{\lambda_p}{(X-a)^p}$ s'appelle la ***partie polaire*** *de F relative au pôle a.*

Méthode 2.3.13. Soit F une fraction rationnelle admettant a pour pôle.

— Si a est pôle d'ordre 1 de F, on peut écrire $F = \frac{P}{(X-a)Q_1}$ où Q_1 est un polynôme n'admettant pas a pour racine. On cherche le scalaire λ_1 tel que : $F = \frac{\lambda_1}{X-a} + F_0$ où F_0 est une fraction rationnelle qui n'admet pas a pour pôle. En multipliant cette égalité par $X-a$, on obtient : $\frac{P}{Q_1} = \lambda_1 + (X-a)F_0$ ce qui, en substituant a à X, donne $\lambda_1 = \frac{P(a)}{Q_1(a)}$. La partie polaire relative au pôle a est donc : $\frac{\lambda_1}{X-a}$ avec $\lambda_1 = \frac{P(a)}{Q_1(a)} = \frac{P(a)}{Q'(a)}$ la dernière égalité venant de la relation $Q' = (X-a)Q_1' + Q_1$.

— Si a est pôle d'ordre 2 de F, on peut écrire $F = \frac{P}{(X-a)^2Q_2}$ où Q_2 est un polynôme n'admettant pas a pour racine. On cherche les scalaire λ_1 et λ_2 tels que : $F = \frac{\lambda_2}{(X-a)^2} + \frac{\lambda_1}{X-a} + F_0$ où F_0 est une fraction rationnelle qui n'admet pas a pour pôle. En multipliant cette égalité par $(X-a)^2$, on obtient : $\frac{P}{Q_2} = \lambda_2 + \lambda_1(X-a) + (X-a)^2F_0$ ce qui, en substituant a à X, donne $\lambda_2 = \frac{P(a)}{Q_2(a)}$. La partie polaire relative au pôle a est donc : $\frac{\lambda_2}{(X-a)^2} + \frac{\lambda_1}{X-a}$ avec $\lambda_2 = \frac{P(a)}{Q_2(a)}$. Pour trouver λ_1, on peut alors retrancher $\frac{\lambda_2}{(X-a)^2}$ pour obtenir une fraction dont a n'est pas pôle, ou est pôle simple ce qui ramène au cas précédent.

— Dans le cas général, si a est pôle d'ordre n de F, alors $F = \frac{P}{(X-a)^nQ_n}$ avec $Q_n(a) \neq 0$ et la partie polaire relative au pôle a est : $\sum_{p=1}^{n} \frac{\lambda_p}{(X-a)^p}$ avec $\lambda_n = \frac{P(a)}{Q_n(a)}$, comme on le voit en multipliant l'égalité : $F = \sum_{p=1}^{n} \frac{\lambda_p}{(X-a)^p} + F_0$ par $(X-a)^n$ et en substituant a à X. Comme de plus, $Q = (X-a)^nQ_n$, alors on a $Q_n(a) = \frac{Q^{(n)}(a)}{n!}$ (par exemple en utilisant la formule de Leibniz ou la formule de Taylor).

Exemple 2.3.14.

1. Soit $F = \frac{X^5+1}{X(X-1)^2}$.
   — La partie polaire associée au pôle 0 est $\frac{\lambda}{X}$ avec $\lambda = 1$.
   — La partie polaire associée au pôle 1 est $\frac{\lambda}{X-1} + \frac{\mu}{(X-1)^2}$ avec $\mu = 2$. Comme : $F - \frac{2}{(X-1)^2} = \frac{X^4+X^3+X^2+X-1}{X(1-X)}$ on en déduit $\lambda = 3$.
2. Soit $F = \frac{1}{X^5-1}$. Si ω est un pôle de F, c'est-à-dire une racine cinquième de l'unité, alors la partie polaire associée à ω est $\frac{\lambda}{X-\omega}$ avec : $\lambda = \frac{1}{5\omega^4} = \frac{\omega}{5\omega^5} = \frac{\omega}{5}$.

Théorème 2.3.15 (Décomposition en éléments simples sur $\mathbb{C}$)**.** *Étant donnée une fraction rationnelle $F \in \mathbb{C}(X)$ dont les pôles sont $a_1, a_2, \cdots, a_n$ distincts deux à deux et d'ordres de multiplicité respectifs $r_1, r_2, \cdots, r_n$, il existe un unique polynôme $E \in \mathbb{C}[X]$ et une unique famille de scalaires $(\lambda_{i,j})_{\substack{1\leq i\leq n \\ 1\leq j\leq r_i}}$ tels que :*

$$F = E + \sum_{i=1}^{n} \left(\underbrace{\sum_{j=1}^{r_i} \frac{\lambda_{i,j}}{(X-a_i)^j}}_{\text{Partie polaire associé au pôle } a_i} \right).$$

Autrement dit, toute fraction rationnelle de $\mathbb{C}(X)$ est la somme de sa partie entière et des parties polaires associées à chacun de ses pôles. Cette décomposition s'appelle la décomposition en éléments simples dans $\mathbb{C}(X)$ de la fraction F.

Méthode 2.3.16. Soit $F = \frac{P}{Q}$ une fraction rationnelle à coefficients complexes dont les pôles sont $a_1, a_2, \cdots, a_n$ d'ordres de multiplicité respectifs $r_1, r_2, \cdots, r_n$. Sa décomposition en éléments simples est de la forme :

$$F = E + \sum_{i=1}^{n} \left(\sum_{j=1}^{r_i} \frac{\lambda_{i,j}}{(X - a_i)^j} \right).$$

- La détermination de la partie entière E se fait à l'aide de la division euclidienne de P par Q, division limitée puisque seul le quotient nous intéresse.
- Les coefficients λ_{i,r_i} se calculent immédiatement à l'aide des formules : $\lambda_{i,r_i} = \frac{P(a_i)}{Q_i(a_i)} = \frac{r_i! P(a_i)}{Q^{(r_i)}(a_i)}$ avec $Q = (X - a_i)^{r_i} Q_i$.
- Si tous les pôles sont simples, on a ainsi la décomposition en éléments simples de F. Sinon, on peut retrancher à F chaque fraction $\frac{\lambda_{i,r_i}}{(X-a_i)^{r_i}}$ et recommencer avec la fraction ainsi obtenue. Mais il est souvent beaucoup plus rapide de déterminer les derniers coefficients en utilisant certaines des méthodes qui suivent :

Si la fraction est à coefficients réels

Si F est une fraction rationnelle à coefficients réels et si a est un pôle non réel de F d'ordre r, alors $\bar{a}$ est aussi un pôle d'ordre r et les coefficients des parties polaires associées à a et $\bar{a}$ sont conjugués deux à deux. En effet :

- Puisque le dénominateur Q de F est réel, si a est une racine non réelle de Q, alors $\bar{a}$ est aussi racine de Q au même ordre de multiplicité.
- Si $F = \sum_{i=1}^{r} \frac{\lambda_i}{(X-a)^i} + F_1$ où a n'est pas pôle de la fraction rationnelle F_1, alors $F = \overline{F} = \sum_{i=1}^{r} \frac{\overline{\lambda_i}}{(X-\bar{a})^i} + \overline{F_1}$ et la fraction rationnelle $\overline{F_1}$ n'admet pas $\bar{a}$ pour pôle. Donc, la partie polaire associée au pôle $\bar{a}$ est : $\sum_{i=1}^{r} \frac{\overline{\lambda_i}}{(X-\bar{a})^i}$ (unicité de la partie polaire).

Si la fraction est paire ou impaire

Si la fraction rationnelle F est paire ou impaire et que a est un pôle de F d'ordre n, alors $-a$ est aussi pôle de F d'ordre n et la comparaison des décompositions en éléments simples de $F(X)$ et $F(-X) = {}^{+}_{-}F(X)$ donne des relations entre les coefficients.

Exemple : La fraction $F = \frac{4}{(X^2-1)^2}$ se décompose en éléments simples : $F = \frac{a}{X-1} + \frac{b}{X+1} + \frac{c}{(X-1)^2} + \frac{d}{(X+1)^2}$. On a : $F(X) = F(-X) = \frac{a}{-X-1} + \frac{b}{-X+1} + \frac{c}{(-X-1)^2} + \frac{d}{(-X+1)^2}$. L'unicité de la décomposition en éléments simples nous donne alors $a = -b$ et $c = d$.

- On a immédiatement $c = \frac{4}{(1+1)^2} = 1$, donc $c = d = 1$.
- Pour déterminer a et b, il suffit de substituer 0 à X, ce qui donne : $4 = -a + b + c + d = 2 - 2a$, donc $a = -1$ et $b = 1$.

On a donc : $F = \frac{-1}{X-1} + \frac{1}{X+1} + \frac{1}{(X-1)^2} + \frac{1}{(X+1)^2}$.

Si la fraction est de degré strictement négatif

Soit F une fraction rationnelle de degré strictement négatif. Si f est la restriction à $\mathbb{R}$ de sa fonction rationnelle associée, alors la fonction $x \mapsto xf(x)$ a une limite finie en l'infini : on peut ainsi trouver des relations entre les coefficients des termes en $\frac{1}{X-a_i}$ de la décomposition en éléments simples de F.

Exemple : Soit la fraction rationnelle $F = \frac{4X^3}{(X^2-1)^2}$. L'imparité de F nous dit

que sa décomposition en éléments simples est du type : $F = \frac{a}{X-1} + \frac{a}{X+1} + \frac{b}{(X-1)^2} - \frac{b}{(X+1)^2}$. Alors $b = 1$ et puisque $\lim\limits_{x\to+\infty} xf(x) = 4$, on a $2a = 4$. Donc : $F = \frac{2}{X-1} + \frac{2}{X+1} + \frac{1}{(X-1)^2} - \frac{1}{(X+1)^2}$.

S'il ne reste qu'un ou deux coefficients à calculer

Lorsqu'il ne reste plus qu'un ou deux coefficients à déterminer, on peut substituer à X une ou deux valeurs simples.

Exemple : Soit $F = \frac{X^4+1}{(X+1)^2(X^2+1)} = 1 + \frac{a}{(X+1)^2} + \frac{b}{X+1} + \frac{c}{X-i} - \frac{\bar{c}}{X+i}$. On trouve d'abord : $c = \frac{i^4+1}{(i+1)^2(i+i)} = \frac{2}{(2i)^2} = -\frac{1}{2}$ et $a = \frac{1+1}{1+1} = 1$. En substituant 0 à X, on obtient de plus : $1 = 1 + a + b - \frac{c}{i} + \frac{\bar{c}}{i} = 2 + b$ et donc $b = -1$. D'où : $F = \frac{X^4+1}{(X+1)^2(X^2+1)} = 1 + \frac{1}{(X+1)^2} - \frac{1}{X+1} - \frac{1}{2(X-i)} - \frac{1}{2(X+i)}$.

Exercice 2.3.17. Décomposer en éléments simples dans $\mathbb{C}(X)$ les fractions rationnelles F suivantes :

1. $\frac{X^5}{X^4-1}$.
2. $\frac{X}{(X^2+X+1)(X+1)^3}$.

Solution. 1. — **Recherche de la partie entière** : Commençons par chercher la partie entière de $\frac{X^5}{X^4-1}$, ce qui revient à chercher le quotient de la division euclidienne de X^5 par $X^4 - 1$. Cette division s'écrit : $X^5 = X(X^4 - 1) + X$, donc X est la partie entière cherchée.

— **Forme de la décomposition en éléments simples** : Pour obtenir la forme de la DES de la fraction F dans $\mathbb{C}(X)$, on commence par factoriser le dénominateur $X^4 - 1$ en produit de facteurs irréductibles dans $\mathbb{C}[X]$. Or, dans $\mathbb{C}[X]$, on a :

$$X^4 - 1 = (X - 1)(X + 1)(X - i)(X + i).$$

Du coup, la décomposition en éléments simples de $\frac{X^5}{X^4-1}$ est :

$$\begin{aligned}\frac{X^5}{X^4-1} &= \frac{X^5}{(X-1)(X+1)(X-i)(X+i)} \\ &= X + \frac{a}{X-1} + \frac{b}{X+1} + \frac{c}{X-i} + \frac{d}{X+i}, \quad \bigstar\end{aligned}$$

où $a, b, c, d \in \mathbb{C}$ sont à déterminer.

— **Prise en compte de l'imparité** : On a :

$$\begin{aligned}F(X) = -F(-X) &= -\Big[-X + \frac{a}{-X-1} + \frac{b}{-X+1} + \frac{c}{-X-i} + \frac{d}{-X+i}\Big] \\ &= X + \frac{a}{X+1} + \frac{b}{X-1} + \frac{c}{X+i} + \frac{d}{X-i}.\end{aligned}$$

Ceci est une "nouvelle" décomposition de F en éléments simples. Une telle décomposition étant unique, nous obtenons par identification des coefficients : $b = a$ et $c = d$.

— **Calcul de** a : Multiplions $\bigstar$ par $X - 1$:

$$\frac{X^5}{(X+1)(X^2+1)} = X(X-1) + a + (X-1)\Big[\frac{b}{X+1} + \frac{c}{X-i} + \frac{d}{X+i}\Big],$$

puis évaluons en 1, pour obtenir : $a = \frac{1}{4}$.

— **Calcul de** c : Multiplions ★ par $X-i$:

$$\frac{X^5}{(X^2-1)(X+i)} = X(X-i)+c+(X-i)\Big[\frac{a}{X-1}+\frac{b}{X+1}+\frac{d}{X+i}\Big],$$

puis évaluons en i, pour obtenir : $c=-\frac{1}{4}$.

— **Conclusion** : Finalement, on obtient :

$$\frac{X^5}{X^4-1} = X+\frac{1}{4}\Big[\frac{1}{X-1}+\frac{1}{X+1}-\frac{1}{X-i}-\frac{1}{X+i}\Big].$$

2. — **Recherche de la partie entière** : Comme $\deg F<0$, la partie entière est nulle.

— **Forme de la décomposition en éléments simples** : Pour obtenir la forme de la DES de la fraction F dans $\mathbb{C}(X)$, on commence par factoriser le dénominateur $(X^2+X+1)(X+1)^3$ en produit de facteurs irréductibles dans $\mathbb{C}[X]$. Or, dans $\mathbb{C}[X]$, on a :

$$(X^2+X+1)(X+1)^3 = (X-j)(X-j^2)(X+1)^3.$$

Du coup, la décomposition cherchée s'écrit donc :

$$\begin{aligned}\frac{X}{(X^2+X+1)(X+1)^3} &= \frac{X}{(X-j)(X-j^2)(X+1)^3}\\ &= \frac{a}{X-j}+\frac{b}{X-j^2}+\frac{c}{(X+1)^3}+\frac{d}{(X+1)^2}+\frac{e}{X+1}, \quad ★\end{aligned}$$

où $a,b,c,d,e\in\mathbb{C}$ sont à déterminer.

— **Utilisation de la conjugaison** : Conjuguons ★ :

$$F(X)=\overline{F(X)} = \frac{\bar a}{X-j^2}+\frac{\bar b}{X-j}+\frac{\bar c}{(X+1)^3}+\frac{\bar d}{(X+1)^2}+\frac{\bar e}{X+1}$$

Ceci est une "nouvelle" décomposition de F en éléments simples. Une telle décomposition étant unique, nous obtenons par identification des coefficients : $b=\bar a$.

— **Calcul de** a : Multiplions ★ par $X-j$, puis évaluons en j pour obtenir : $a=\frac{j}{(j-j^2)(j+1)^3}=\frac{ij}{\sqrt{3}}$.

— **Calcul de** c : Multiplions ★ par $(X+1)^3$, puis évaluons en -1, pour obtenir : $c=-1$.

— **Utilisation du comportement en** ∞ : Pour calculer d et e, nous ne pouvons plus multiplier par $(X+1)^2$ et $(X+1)$ respectivement puis évaluer en -1, car le membre de gauche aura alors un pôle en -1. Pour trouver des équations faisant intervenir d et e, une solution consiste à multiplier par une puissance de X, à évaluer ensuite en $x\in\mathbb{R}$ quelconque, puis à faire tendre x vers ∞.
Ici, multiplions ★ par X :

$$\frac{X^2}{(X^2+X+1)(X+1)^3} = \frac{aX}{X-j}+\frac{bX}{X-j^2}+\frac{cX}{(X+1)^3}+\frac{dX}{(X+1)^2}+\frac{eX}{X+1},$$

puis évaluons en $x\in\mathbb{R}\setminus\{-1\}$:

$$\frac{x^2}{(x^2+x+1)(x+1)^3} = \frac{ax}{x-j}+\frac{bx}{x-j^2}+\frac{cx}{(x+1)^3}+\frac{dx}{(x+1)^2}+\frac{ex}{x+1},$$

et enfin faisons tendre x vers $+\infty$ pour obtenir : $0 = a + b + e$. Par conséquent : $e = -a - b = -a - \bar{a} = -2\,\mathrm{Re}(a) = 1$.

► **Remarque** : Il n'est pas correct de "faire tendre X vers ∞" car X est un polynôme et non un nombre.

— **Calcul de** d : Nous ne pouvons pas utiliser la technique précédente pour calculer d, car nous devrions pour cela multiplier par X^2, et certains termes à droite auraient alors une limite infinie. Evaluons simplement ★ en 0, pour obtenir : $0 = \frac{a}{-j} + \frac{b}{-j^2} + c + d + e$. Ainsi, sachant que $a = \frac{ij}{\sqrt{3}}$: $d = (\frac{a}{j} + \frac{\bar{a}}{j^2}) - c - e = -c - e = 0$.

— **Conclusion** : Finalement, on obtient :

$$\frac{X}{(X^2+X+1)(X+1)^3} = \frac{i}{\sqrt{3}}\Big(\frac{j}{X-j} - \frac{j^2}{X-j^2}\Big) - \frac{1}{(X+1)^3} + \frac{1}{X+1}.$$

Théorème 2.3.18 (Décomposition en éléments simples sur $\mathbb{R}$)**.** *Soit $F = \frac{P}{Q} \in \mathbb{R}(X)$* ***irréductible*** *de partie entière E. On introduit la factorisation irréductible de Q : $Q = B\prod_{i=1}^{r}(X-a_i)^{n_i}\prod_{j=1}^{s}(X^2+b_jX+c_j)^{m_j}$. Il existe des familles uniques $(\lambda_{i,k})_{\substack{1\le i\le r\\ 1\le k\le n_i}}$, $(u_{j,k})_{\substack{1\le j\le s\\ 1\le k\le m_j}}$ et $(v_{j,k})_{\substack{1\le j\le s\\ 1\le k\le m_j}}$ de réels telles que : $F = E + \sum_{i=1}^{r}\sum_{k=1}^{n_i}\frac{\lambda_{i,k}}{(X-a_i)^k} + \sum_{j=1}^{s}\sum_{k=1}^{m_j}\frac{u_{j,k}X+v_{j,k}}{(X^2+b_jX+c_j)^k}$.*

Définition 2.3.19. Dans la décomposition en éléments simples d'une fraction rationnelle dans $\mathbb{R}(X)$, une fraction de la forme $\frac{\lambda_{i,k}}{(X-a_i)^k}$ s'appelle **un élément simple de première espèce**, une fraction de la forme $\frac{u_{j,k}X+v_{j,k}}{(X^2+b_jX+c_j)^k}$ s'appelle **un élément simple de deuxième espèce**.

En pratique : Quand les pôles **non réels** d'une fraction **réelle** sont **simples**, on peut obtenir la décomposition en éléments simples sur $\mathbb{R}$ facilement à partir de la décomposition en éléments simples sur $\mathbb{C}$ par simple regroupement des parties polaires conjuguées.

Exercice 2.3.20. Soit $F = \frac{1}{(X^2+1)(X^2+X+1)}$.

1. Montrer qu'il existe $(\alpha, \beta) \in \mathbb{C}^2$ tels que :
$$\frac{1}{(X^2+1)(X^2+X+1)} = \frac{\alpha}{X-i} + \frac{\bar{\alpha}}{X+i} + \frac{\beta}{X-j} + \frac{\bar{\beta}}{X-j^2}$$
2. Déterminer les valeurs de α et β.
3. Donner alors la décomposition en éléments simples sur $\mathbb{R}$ de F.
4. Retrouver directement cette décomposition sur $\mathbb{R}$ (sans passer par celle sur $\mathbb{C}$).
   Indication : on pourra multiplier par X^2+1 et substituer i à X, puis multiplier par X^2+X+1 et substituer $j = \exp(2i\pi/3)$ à X.

Solution. 1. Remarquons que $F \in \mathbb{R}[X]$, que la partie entière vaut 0 et que les pôles sont : i, $-i$, j et j^2, tous d'ordre de multiplicité 1. On rappelle que $1 + j + j^2 = 0, j^3 = 1$ et $j^2 = \bar{j}$. En vertu de ce qui précède, il existe alors des complexes α et β tels que :
$$\frac{1}{(X^2+1)(X^2+X+1)} = \frac{\alpha}{X-i} + \frac{\bar{\alpha}}{X+i} + \frac{\beta}{X-j} + \frac{\bar{\beta}}{X-j^2}.$$

2. On évalue $(X-i)F(X) = \frac{1}{(X^2+X+1)(X+i)}$ en $X=i$ et on trouve $\alpha = -\frac{1}{2}$. Pour β, on évalue $(X-j)F(X) = \frac{1}{(X^2+1)(X-j^2)}$ en $X=j$. Le dénominateur est alors :

$$\left(j^2+1\right)\left(j-j^2\right) = (-j)j(1-j) = j^2(j-1) = 1-j^2.$$

On en déduit que $\beta = \frac{1}{1-j^2} = \frac{1-j}{3}$.

3. On a donc $F = -\frac{1}{2(X-i)} - \frac{1}{2(X+i)} + \frac{1-j}{3(X-j)} + \frac{1-j^2}{3(X-j^2)}$. En regroupant les termes deux à deux conjugués, on obtient :

$$F = -\frac{X}{X^2+1} + \frac{X+1}{X^2+X+1}.$$

4. La décomposition en éléments simples sur $\mathbb{R}$ de F s'écrit :

$$F = \frac{aX+b}{X^2+1} + \frac{cX+d}{X^2+X+1} \quad \text{avec} \quad (a,b,c,d) \in \mathbb{R}^4.$$

En multipliant par X^2+1, on obtient :

$$\frac{1}{X^2+X+1} = aX+b+\left(X^2+1\right)\frac{cX+d}{X^2+X+1},$$

ce qui, en remplaçant X par i, donne $ai+b = \frac{1}{i} = -i$ et donc $a=-1$ et $b=0$ puisque a et b sont réels. De même, en multipliant par X^2+X+1, on obtient :

$$\frac{1}{X^2+1} = cX+d+\left(X^2+X+1\right)\frac{aX+b}{X^2+1}$$

et en remplaçant X par j, on obtient :

$$cj+d = \frac{1}{j^2+1} = \frac{1}{-j} = -j^2 = j+1.$$

Puisque j n'est pas réel, on en déduit $c=1$ et $d=1$.

Exercice 2.3.21. Décomposer en éléments simples dans $\mathbb{R}(X)$ les fractions rationnelles F suivantes :

1. $\frac{X+3}{(X+1)^2(X+2)}$.
2. $\frac{X^4}{(X+3)(X^2+X+3)}$.
3. $\frac{1}{(X-1)^2(X^2+4)}$.

Solution. 1. — Forme de la décomposition en éléments simples : La partie entière est nulle, donc pour certains $a,b,c \in \mathbb{R}$:

$$\frac{X+3}{(X+1)^2(X+2)} = \frac{a}{(X+1)^2} + \frac{b}{X+1} + \frac{c}{X+2} \quad \bigstar.$$

— Calcul de a : On multiplie $\bigstar$ par $(X+1)^2$ puis on évalue en -1 pour obtenir : $a=2$.
— Calcul de c : On recommence. On multiplie $\bigstar$ par $X+2$ puis on évalue en -2 pour obtenir : $c=1$.

— Calcul de b : On ne peut malheureusement pas reproduire le raisonnement précédent pour calculer b. Multiplier ★ par $X+1$ puis évaluer en -1 nous conduirait en effet à diviser par 0 à cause du terme $(X+1)^2$. Cependant, plusieurs approches sont envisageables, **AU CHOIX** :
  — On peut multiplier ★ par X puis passer à la limite en $+\infty$ pour obtenir : $0 = 0+b+c$, donc : $b = -c = -1$. On obtient généralement ainsi une équation simple et agréable.
  — On peut évaluer ★ en un point, par exemple en 0 pour obtenir : $\frac{3}{2} = a+b+\frac{c}{2}$, ce qui donne aussi : $b = -1$. Les équations qu'on obtient en évaluant en un point sont souvent un peu plus compliquées que celles qu'on obtient en passant à la limite en $+\infty$.
— Conclusion : Finalement, on obtient :

$$\frac{X+3}{(X+1)^2(X+2)} = \frac{2}{(X+1)^2} - \frac{1}{X+1} + \frac{1}{X+2}.$$

2. — Partie entière : La division euclidienne de X^4 par $(X+3)(X^2+X+3)$ s'écrit :

$$X^4 = (X+3)(X^2+X+3)(X-4) + 10X^2 + 15X + 36,$$

donc la partie entière cherchée vaut $X-4$.
— Forme de la décomposition en éléments simples : Pour certains $a, b, c \in \mathbb{R}$:

$$\frac{X^4}{(X+3)(X^2+X+3)} = X - 4 + \frac{a}{X+3} + \frac{bX+c}{X^2+X+3}.$$

En tenant compte de la division euclidienne calculée juste avant, on peut aussi dire que :

$$\frac{10X^2+15X+36}{(X+3)(X^2+X+3)} = \frac{a}{X+3} + \frac{bX+c}{X^2+X+3} \quad ★.$$

▶ Il est toujours plus facile de calculer les coefficients d'une décomposition en éléments simples quand la partie entière est nulle.
— Calcul de a : On multiplie ★ par $X+3$ puis on évalue en -3 pour obtenir : $a = 9$.
— Calcul de b : On multiplie ★ par X puis on passe à la limite en $+\infty$ pour obtenir : $10 = a+b$ et donc $b = 1$.
— Calcul de c : On peut évaluer ★ par exemple en 0 pour obtenir : $4 = \frac{a}{3} + \frac{c}{3}$ et donc : $c = 12 - a = 3$.
— Conclusion : Finalement, on obtient :

$$\frac{X^4}{(X+3)(X^2+X+3)} = X - 4 + \frac{9}{X+3} + \frac{X+3}{X^2+X+3}.$$

3. — Forme de la décomposition en éléments simples : La partie entière est nulle, donc pour certains $a, b, c \in \mathbb{R}$:

$$\frac{1}{(X-1)^2(X^2+4)} = \frac{a}{(X-1)^2} + \frac{b}{X-1} + \frac{cX+d}{X^2+4} \quad ★.$$

— Calcul de a : On multiplie $\bigstar$ par $(X-1)^2$ puis on évalue en 1 pour obtenir : $a = \frac{1}{5}$.
— Calcul de c et d : Le polynôme X^2+4 admet $2i$ et $-2i$ pour racines. On multiplie $\bigstar$ par X^2+4 puis on évalue en $2i$ pour obtenir : $2ic + d = \frac{1}{(2i-1)^2} = \frac{1}{-3-4i} = \frac{-3+4i}{25}$. Or c et d sont des RÉELS, donc par identification des parties réelles et imaginaires : $c = \frac{2}{25}$ et $d = -\frac{3}{25}$.
— Calcul de b : On multiplie $\bigstar$ par X puis on passe à la limite en $+\infty$ pour obtenir : $0 = b + c$ et donc $b = -c = -\frac{2}{25}$.
— Conclusion : Finalement, on obtient :

$$\frac{1}{(X-1)^2(X^2+4)} = \frac{1}{5(X-1)^2} - \frac{2}{25(X-1)} + \frac{2X-3}{25(X^2+4)}.$$

CHAPITRE 3

Espaces vectoriels et applications linéaires

Sommaire

3.1 Structure d'espace vectoriel

Dans ce chapitre, $\mathbb{K}$ est l'un des corps $\mathbb{R}$ ou $\mathbb{C}$ et I est un ensemble non vide quelconque.

La notion d'espace vectoriel introduite dans ce chapitre est un nouvel exemple fondamental de structure algébrique, après les groupes, les anneaux et les corps. Comme nous le verrons, la notion d'espace vectoriel généralise, comme son nom l'indique, les notions de vecteurs du plan et de l'espace introduites au lycée. La théorie mathématique des espaces vectoriels s'appelle l'algèbre linéaire.

3.1.1 Espace vectoriel et combinaisons linéaires

Définition 3.1.1 (Espace vectoriel)**.** On appelle $\mathbb{K}$-espace vectoriel ou espace vectoriel sur $\mathbb{K}$ tout triplet $(E, +, \cdot)$ vérifiant les propriétés suivantes :

- $(E, +)$ est un groupe commutatif dont l'élément neutre est noté 0_E ou 0 et appelé le vecteur nul de E,
- $\cdot$ est une application de $\mathbb{K} \times E$ dans E. À partir d'un élément λ de $\mathbb{K}$ et d'un élément x de E, $\cdot$ fournit un élément de E noté $\lambda \cdot x$ ou plus simplement λx. Par définition, cette application $\cdot$ doit satisfaire les propriétés suivantes :
  - pour tout $x \in E$: $1 \cdot x = x$,
  - pour tous $x, y \in E$ et $\lambda \in \mathbb{K}$: $\lambda \cdot (x + y) = (\lambda \cdot x) + (\lambda \cdot y)$,
  - pour tous $x \in E$ et $\lambda, \mu \in \mathbb{K}$: $(\lambda + \mu) \cdot x = (\lambda \cdot x) + (\mu \cdot x)$,
  - pour tous $x \in E$ et $\lambda, \mu \in \mathbb{K}$: $\lambda \cdot (\mu \cdot x) = (\lambda\mu) \cdot x$.

Les éléments de E sont appelés des vecteurs, ceux de $\mathbb{K}$ des scalaires. La loi $\cdot$, qui n'est pas une loi de composition interne sur E puisqu'à travers elle des éléments de $\mathbb{K}$ agissent sur des vecteurs, est qualifiée de loi externe. La loi $+$ est appelée addition et la loi $\cdot$ est appelée multiplication par un scalaire. Le corps $\mathbb{K}$ est qualifié de corps de base pour E.

▶ Sachez que les mathématiciens ne mettent jamais de flèches au-dessus de leurs vecteurs, sauf quand ils font de la géométrie dans le plan et dans l'espace comme vous en avez fait jusqu'ici.

▶ Notons qu'un espace vectoriel contient toujours au moins le vecteur nul, il ne peut être vide. Si un ensemble ne contient pas de vecteur nul (élément neutre pour la loi $+$), ce n'est pas un espace vectoriel.

Exemples 3.1.2. Quelques exemples classiques d'espaces vectoriels (munis des lois usuelles) :

1. $\mathbb{R}$, $\mathbb{C}$ et plus généralement $\mathbb{K}^n$;
2. $\mathbb{K}[X]$, l'ensemble des polynômes à coefficients dans $\mathbb{K}$;
3. $\mathcal{M}_{n,p}(\mathbb{K})$ l'ensemble des matrices à coefficients dans $\mathbb{K}$ de taille $n \times p$;
4. $\mathcal{F}(\mathbb{R}, \mathbb{R})$ l'ensemble des fonctions définies sur $\mathbb{R}$ et à valeurs dans $\mathbb{R}$;
5. l'ensemble des suites à valeurs réelles ou complexes (que l'on pourrait noter $\mathcal{F}(\mathbb{N}, \mathbb{K})$ ou bien $\mathbb{K}^{\mathbb{N}}$).

▶ Pour montrer que ces différents ensembles possèdent une structure d'espace vectoriel, on revient à la définition précédente. Il s'agit d'un résultat classique du cours qui peut être réutilisé sans démonstration le jour d'examen.

Théorème 3.1.3 (Règles de calcul dans un espace vectoriel)**.** *Soit E un $\mathbb{K}$-espace vectoriel.*

1. *Pour tous $x \in E$ et $\lambda \in \mathbb{K}$: $\lambda \cdot x = 0_E \Leftrightarrow \lambda = 0$ ou $x = 0_E$.*
2. *Pour tout $x \in E$: $-x = (-1) \cdot x$, où $-x$ est l'opposé de x dans E et -1 l'opposé de 1 dans $\mathbb{K}$.*

Démonstration. 1. Trois étapes. Soient $x \in E$ et $\lambda \in \mathbb{K}$.

- Comme : $0 \cdot x = (0+0) \cdot x = 0 \cdot x + 0 \cdot x$, alors après simplification dans le groupe $(E, +)$: $0 \cdot x = 0_E$.

— Comme : $\lambda \cdot 0_E = \lambda \cdot (0_E + 0_E) = \lambda \cdot 0_E + \lambda \cdot 0_E$, alors après simplification : $\lambda \cdot 0_E = 0_E$.
— Si : $\lambda \cdot x = 0_E$ et si : $\lambda \neq 0$, alors :

$$x = 1 \cdot x = (\frac{1}{\lambda} \times \lambda) \cdot x = \frac{1}{\lambda} \cdot (\lambda \cdot x) = \frac{1}{\lambda} \cdot 0_E = 0_E.$$

2. Pour tout $x \in E$: $x + (-1) \cdot x = 1 \cdot x + (-1) \cdot x = (1-1) \cdot x = 0 \cdot x = 0_E$, donc : $-x = (-1) \cdot x$.

□

Définition 3.1.4 (Combinaisons linéaires d'un nombre fini de vecteurs)**.** Soient E un $\mathbb{K}$-espace vectoriel et $x_1, \cdots, x_n \in E$. On appelle combinaison linéaire de $x_1, \cdots, x_n$ tout vecteur de E de la forme : $\sum_{k=1}^{n} \lambda_k x_k$ pour certains $\lambda_1, \cdots, \lambda_n \in \mathbb{K}$.

Attention : Une égalité de la forme

$$\sum_{k=1}^{n} \lambda_k x_k = \sum_{k=1}^{n} \mu_k x_k$$

n'implique pas l'égalité :

$$\lambda_k = \mu_k, \forall k \in \{1, \cdots, n\}.$$

En d'autres termes, on ne peut pas pratiquer d'identification systématiquement quand deux combinaisons linéaires d'une même famille de vecteurs sont égales.

Par exemple :

$$(1,1) + 2(0,1) + 2(1,0) = (3,3) = 2(1,1) + (0,1) + (1,0).$$

Définition 3.1.5 (Famille presque nulle de scalaires)**.** On dit qu'une famille $X = (x_i)_{i \in I} \in \mathbb{K}^I$ est presque nulle (ou à support fini) si l'ensemble $\text{supp}(X) = \{i \in I | x_i \neq 0\}$ est de cardinal fini. Autrement dit, une famille d'éléments de $\mathbb{K}$ indexée par I est presque nulle si tous ses éléments sont nuls **SAUF UN NOMBRE FINI D'ENTRE EUX**.

▶ Dans le cas où I est un ensemble **FINI**, la précision "presque nulle" est évidemment sans intérêt.

Définition 3.1.6 (c.l. d'un nombre infini de vecteurs)**.** Soient E un espace vectoriel et $(x_i)_{i \in I}$ une famille de vecteurs de E. On appelle combinaison linéaire de $(x_i)_{i \in I}$ tout vecteur de E de la forme : $\sum_{i \in I} \lambda_i x_i$ où $(\lambda_i)_{i \in I}$ est une famille **PRESQUE NULLE** d'éléments de $\mathbb{K}$.

Attention : Pour un nombre **FINI** de vecteurs, pas besoin de familles **PRESQUE NULLES** de scalaires !

▶ Nous pourrons maintenant parler des combinaisons linéaires d'un nombre **INFINI** de vecteurs, mais chacune de ces combinaisons linéaires reste fondamentalement une somme **FINIE**. Les vraies sommes infinies n'ont aucun sens sans une notion de passage à la limite adéquat.

Exemple 3.1.7. $\mathbb{K}[X]$ est l'ensemble des combinaisons linéaires de la famille $(X^k)_{k \in \mathbb{N}}$.
▶ Rappelons à ce sujet que la notation $\sum_{k=0}^{+\infty} a_k X^k$ des polynômes, très pratique, désigne en fait une somme **FINIE**.

3.1.2 Sous-espace vectoriel

Définition 3.1.8 (Sous-espace vectoriel)**.** Soient E un espace vectoriel et F une partie de E. On dit que F est un sous-espace vectoriel de E si :
- F est stable par addition : $\forall x, y \in F,\ x + y \in F$;
- F est stable par multiplication par un scalaire : $\forall \lambda \in \mathbb{K}, \forall x \in F,\ \lambda \cdot x \in F$;
- F est un $\mathbb{K}$-espace vectoriel pour les lois de E.

▶ Si F un sous-espace vectoriel de E, F est un sous-groupe additif de E, donc : $0_F = 0_E \in F$.

Exemple 3.1.9. Si E est un $\mathbb{K}$-espace vectoriel, $\{0_E\}$ et E sont deux sous-espaces vectoriels de E.

▶ Pour montrer qu'un ensemble muni d'une addition et d'une multiplication par un scalaire est un espace vectoriel, il suffit souvent de montrer qu'il est **SOUS**-espace d'un autre espace vectoriel connu. On se ramènera notamment aux exemples fondamentaux présentés précédemment.

Théorème 3.1.10 (Caractérisation des sous-espaces vectoriels)**.** *Soient E un $\mathbb{K}$-espace vectoriel et F un sous-ensemble (ou partie) de E. Les assertions suivantes sont équivalentes :*

1. *F est un sous-espace vectoriel de E.*
2. $\begin{cases} 0_E \in F; \\ \forall x, y \in F, \forall \lambda \in \mathbb{K}, \lambda x + y \in F. \end{cases}$

Démonstration. ($\Rightarrow$) Si F est un sous-espace vectoriel de E, on a vu que : $0_E = 0_F \in F$. De plus, pour tous $x, y \in E$ et $\lambda \in \mathbb{K}$, λx et y sont éléments de F car F est stable par multiplication par un scalaire, et enfin : $\lambda x + y \in F$ car F est stable par addition.
($\Leftarrow$) Si l'assertion (2) est vraie, F est stable par différence (pour $\lambda = -1$) donc est un sous-groupe additif de E. Les autres axiomes de la définition des espaces vectoriels ne requièrent aucune vérification particulière car une relation vraie sur E tout entier l'est aussi sur F. □

▶ C'est **TOUJOURS** le résultat précédent qu'il faut utiliser pour montrer qu'une partie d'un espace vectoriel en est un sous-espace vectoriel. Si on utilisait la DÉFINITION des sous-espaces vectoriels, on serait obligé de vérifier beaucoup d'axiomes dont la CARACTÉRISATION fait l'économie.

S'en suit toute une série d'exemples qu'il convient de maîtriser parfaitement.

Exemples 3.1.11. Pour tout $n \in \mathbb{N}$, l'ensemble des polynômes de degré **INFÉRIEUR OU ÉGAL** à n, noté $\mathbb{K}_n[X]$, est un sous-espace vectoriel de $\mathbb{K}[X]$.
▶ Soit $n \in \mathbb{N}$.

1. Pour commencer : $\mathbb{K}_n[X] \subset \mathbb{K}[X]$. Ensuite : $0 \in \mathbb{K}_n[X]$ car : $\deg(0) = -\infty \leq n$.
2. Montrons enfin que $\mathbb{K}_n[X]$ est stable par combinaison linéaire. Soient $P, Q \in \mathbb{K}_n[X]$ et $\lambda \in \mathbb{K}$. Nous savons qu'alors :

$$\deg(\lambda P + Q) \leq \max(\deg(P), \deg(Q)) \leq n,$$

donc : $\lambda P + Q \in \mathbb{K}_n[X]$.

Attention : L'ensemble des polynômes de degré **ÉGAL** à n **N'est PAS** un sous-espace vectoriel de $\mathbb{K}[X]$, il ne contient même pas le polynôme nul !

Exemple 3.1.12. L'ensemble des suites convergentes à valeurs réelles est un sous-espace vectoriel de l'ensemble des suites à valeurs réelles :

- la suite nulle converge (vers 0) ;
- si $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ sont deux suites qui convergent respectivement vers ℓ et ℓ', λ un réel, alors la suite $(\lambda u_n + v_n)_{n\in\mathbb{N}}$ converge (vers $\lambda\ell + \ell'$). Il y a bien stabilité par combinaison linéaire.

Exemple 3.1.13. $F = \{(x, y, z) \in \mathbb{R}^3 \mid x + 2y - 3z = 0\}$ est un sous-espace vectoriel de $\mathbb{R}^3$:

- le vecteur nul appartient bien à F car $0 + 2 \cdot 0 - 3 \cdot 0 = 0$;
- si $u = (x, y, z), v = (x', y', z') \in F$, $\lambda \in \mathbb{R}$ alors $\lambda u + v \in F$. En effet, $\lambda u + v = (\lambda x + x', \lambda y + y', \lambda z + z')$ et :

$$\begin{aligned}&(\lambda x + x') + 2(\lambda y + y') - 3(\lambda z + z')\\ &= \lambda(x + 2y - 3z) + (x' + 2y' - 3z')\\ &= \lambda \cdot 0 + 0 = 0.\end{aligned}$$

Exemple 3.1.14. Soient $a, b, c \in \mathbb{R}$ tels que $(a, b) \neq (0, 0)$. Notons D l'ensemble $\{(x, y) \in \mathbb{R}^2 \mid ax + by + c = 0\}$.

- Si $c \neq 0$, alors D n'est pas un sous-espace vectoriel de $\mathbb{R}^2$ car il ne contient pas le vecteur nul $(0, 0)$.
- Si au contraire $c = 0$, D est un sous-espace vectoriel de $\mathbb{R}^2$, en l'occurence la droite d'équation $ax + by + c = 0$. En effet :
  - Pour commencer, D est une partie de $\mathbb{R}^2$ et il est clair que $(0, 0) \in D$.
  - Soient $(x, y), (x', y') \in D$ et $\lambda \in \mathbb{R}$. Nous devons montrer que $\lambda(x, y) + (x', y')$, égal à $(\lambda x + x', \lambda y + y')$ par définition, est un élément de D. On a $a(\lambda x + x') + b(\lambda y + y') = \lambda(ax + by) + (ax' + by') = \lambda \cdot 0 + 0 = 0$. D'où le résultat.

Théorème 3.1.15 (Intersection de sous-espaces vectoriels)**.** *Soit E un $\mathbb{K}$-espace vectoriel. Toute intersection de sous-espaces vectoriels de E est encore un sous-espace vectoriel de E.*

Démonstration. Soit $(F_i)_{i\in I}$ une famille de sous-espaces vectoriels de E. Nous voulons montrer que $F = \bigcap_{i\in I} F_i$ est un sous-espace vectoriel de E.

- Pour commencer : $F \subset E$. Ensuite : $0_E \in F_i$ pour tout $i \in I$ puisque F_i est un sous-espace vectoriel de E, donc : $0_E \in F$.
- Montrons enfin que F est stable par combinaison linéaire. Soient $x, y \in F$ et $\lambda \in \mathbb{K}$. Pour tout $i \in I$: $\lambda x + y \in F_i$ car F_i est un sous-espace vectoriel de E et : $x, y \in F_i$, donc : $\lambda x + y \in F$.

□

Attention : Ce théorème est faux pour la réunion ! Par exemple dans $\mathbb{R}^2$, les ensembles F et G définis par $F = \left\{ \begin{pmatrix} x \\ 0 \end{pmatrix} \mid x \in \mathbb{R} \right\}$ (l'axe des x) et $G = \left\{ \begin{pmatrix} 0 \\ y \end{pmatrix} \mid x \in \mathbb{R} \right\}$ (l'axe des y) sont des sous-espaces vectoriels, mais pas $F \cup G$ puisque $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ et $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ sont dans cette réunion, mais pas leur somme $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Exercice 3.1.16. Soient F et G deux s.e.v d'un $\mathbb{K}$-espace vectoriel E. Si $F \cup G$ est un s.e.v de E, montrer que $F \subset G$ ou $G \subset F$.

Solution. Raisonnons par l'absurde et supposons que $F \nsubseteq G$ et $G \nsubseteq F$, de sorte qu'il existe $x \in F$, $x \notin G$, et il existe $y \in G$, $y \notin F$. Le vecteur $x+y$ est dans le s.e.v $F \cup G$, donc $x+y \in F$ ou $x+y \in G$. Supposons par exemple $x+y \in F$. Comme F est un s.e.v et que $x \in F$, on a $(x+y)-x \in F$, c'est-à-dire $y \in F$, ce qui est absurde. D'où le résultat.

Définition 3.1.17 (Sous-espace vectoriel engendré par une partie)**.** Soient E un $\mathbb{K}$-espace vectoriel et X une partie de E.

1. L'intersection de tous les sous-espaces vectoriels de E contenant X est appelée le sous-espace vectoriel (de E) engendré par X et notée $\text{Vect}(X)$. À ce titre, $\text{Vect}(X)$ est le plus petit sous-espace vectoriel de E contenant X. Il est donné par la formule $\text{Vect}(X) = \bigcap_{\substack{X \subset F \subset E \\ F \text{ s.e.v de } E}} F$.
   En particulier, tout sous-espace vectoriel de E qui contient X contient aussi $\text{Vect}(X)$.
2. Si : $X = \{x_i | i \in I\}$, $\text{Vect}(X)$ est aussi l'ensemble des combinaisons linéaires de $(x_i)_{i\in I}$ et noté $\text{Vect}(x_i)_{i\in I}$. Autrement dit :

$$\text{Vect}(X) = \{\sum_{i\in I} \lambda_i x_i \mid (\lambda_i)_{i\in I} \in \mathbb{K}^I \text{ presque nulle}\}.$$

Démonstration. 1. En tant qu'intersection de sous-espaces vectoriels de E contenant X, $\text{Vect}(X)$ est lui-même un sous-espace vectoriel de E contenant X, et il est inclus par définition dans tous les sous-espaces vectoriels de E contenant X, ce qui montre que tout sous-espace vectoriel de E qui contient X contient en réalité $\text{Vect}(X)$ tout entier.

2. Notons $V = \{\sum_{i\in I} \lambda_i x_i \mid (\lambda_i)_{i\in I} \in \mathbb{K}^I \text{ presque nulle}\}$ l'ensemble des combinaisons linéaires de $(x_i)_{i\in I}$. Nous voulons montrer que : $V = \text{Vect}(X)$. Or $\text{Vect}(X)$ est stable par combinaison linéaire en tant que sous-espace vectoriel de E et contient X, donc contient toute combinaison linéaire de $(x_i)_{i\in I}$, autrement dit : $V \subset \text{Vect}(X)$. Pour l'inclusion réciproque, d'après le premier point, il nous suffit de montrer que V est un sous-espace vectoriel de E contenant X.
   — Pour commencer : $V \subset E$. Ensuite, V contient 0_E car : $0_E = \sum_{i\in I} 0 \cdot x_i$, mais aussi X car pour tout $j \in I$: $x_j = 1 \cdot x_j + \sum_{i\in I, i\neq j} 0 \cdot x_i$.
   — Pour la stabilité par combinaison linéaire, soient $v, w \in V$ et $\alpha \in \mathbb{K}$. Pour certaines familles presque nulles $(\lambda_i)_{i\in I}$ et $(\mu_i)_{i\in I}$ d'éléments de $\mathbb{K}$: $v = \sum_{i\in I} \lambda_i x_i$ et $w = \sum_{i\in I} \mu_i x_i$. La famille $(\alpha\lambda_i + \mu_i)_{i\in I}$ est alors elle aussi presque nulle et : $\alpha v + w = \sum_{i\in I} (\alpha\lambda_i + \mu_i) x_i$, donc : $\alpha v + w \in V$. □

Remarque 3.1.18 (s.e.v engendré par une famille finie de vecteurs)*.* Soient E un $\mathbb{K}$-espace vectoriel. Si $(x_1, \cdots, x_n)$ est une famille d'un nombre fini de vecteurs de E, on a :

$$\text{Vect}(x_1, \cdots, x_n) = \{\sum_{i=1}^{n} \alpha_i x_i | (\alpha_1, \cdots, \alpha_n) \in \mathbb{K}^n\}$$

et $\text{Vect}(x_1, \cdots, x_n)$ est le plus petit sous-espace vectoriel de E contenant $x_1, x_2, \cdots, x_n$; en d'autres termes, tout sous-espace vectoriel de E contenant $x_1, x_2, \cdots, x_n$ contient aussi $\text{Vect}(x_1, \cdots, x_n)$.

► Tout ensemble qui s'écrit sous la forme $\text{Vect}(\cdots)$ est donc un espace vectoriel. C'est une nouvelle façon de montrer qu'un ensemble possède une structure d'espace vectoriel.

Exemples 3.1.19.
1. $\text{Vect}(0_E) = \{0_E\}$.
2. $\text{Vect}(u)$ est l'ensemble des vecteurs qui s'écrivent sous la forme λu, c'est-à-dire l'ensemble des vecteurs colinéaires à u. Si u est non nul, on dit alors que $\text{Vect}(u)$ est la droite vectorielle dirigée (ou engendrée) par u.
3. $\text{Vect}(u, v)$ est l'ensemble des vecteurs qui s'écrivent sous la forme $\lambda u + \mu v$. Si u et v ne sont pas colinéaires, on dit alors que $\text{Vect}(u, v)$ est le plan vectoriel dirigé (ou engendré) par u et v.

Exemples 3.1.20.
1. Soit $(a, b) \in \mathbb{R}^2$ non nul. Dans $\mathbb{R}^2$, le sous-espace vectoriel $\text{Vect}\big((a, b)\big)$ est la droite passant par $(0, 0)$ dirigée par (a, b).
2. $\mathbb{K}[X] = \text{Vect}\left(X^k\right)_{k\in\mathbb{N}}$ et pour tout $n \in \mathbb{N}$: $\mathbb{K}_n[X] = \text{Vect}(1, X, \cdots, X^n)$.
3. Si le corps de base est $\mathbb{R}$: $\text{Vect}(1) = \{a \times 1 \mid a \in \mathbb{R}\} = \mathbb{R}$ et $\text{Vect}(1, i) = \{a + ib \mid a, b \in \mathbb{R}\} = \mathbb{C}$.
   Si par contre le corps de base est $\mathbb{C}$: $\text{Vect}(1) = \{a \times 1 | a \in \mathbb{C}\} = \mathbb{C}$.

Théorème 3.1.21 (Propriétés des Vect)**.** *Soient E un $\mathbb{K}$-espace vectoriel, X et Y deux parties de E et $x, a, b \in E$.*

1. ***Inclusion** : Si : $X \subset Y$, alors :* $\text{Vect}(X) \subset \text{Vect}(Y)$.
2. ***Ôter un vecteur** : Si : $x \in X$ est combinaison linéaire de $X \setminus \{x\}$:* $\text{Vect}(X) = \text{Vect}\Big(X \setminus \{x\}\Big)$.
3. ***Remplacer un vecteur** : Si b est combinaison linéaire de $X \cup \{a\}$ avec un coefficient **NON NUL** sur a :* $\text{Vect}\Big(X \cup \{a\}\Big) = \text{Vect}\Big(X \cup \{b\}\Big)$.

Démonstration.
1. Toute combinaison linéaire de X est bien sûr une combinaison linéaire de Y !
2. Soit $x \in X$ combinaison linéaire de $X \setminus \{x\}$. D'après (1) : $\text{Vect}\Big(X \setminus \{x\}\Big) \subset \text{Vect}(X)$. Réciproquement, $\text{Vect}\Big(X \setminus \{x\}\Big)$ est un sous-espace vectoriel contenant $X \setminus \{x\}$, or il contient aussi x par hypothèse, donc X tout entier, donc enfin $\text{Vect}(X)$.
3. Dans un sens, $\text{Vect}\Big(X \cup \{a\}\Big)$ contient X et a, donc aussi b par hypothèse, donc $X \cup \{b\}$, donc enfin $\text{Vect}\Big(X \cup \{b\}\Big)$. Dans l'autre sens, remarquons que par hypothèse : $b = \lambda a + x$ avec $\lambda \in \mathbb{K}$ **NON NUL** et où x est une combinaison linéaire de X. Ainsi : $a = \frac{b-x}{\lambda}$ donc a est combinaison linéaire de $X \cup \{b\}$ avec un coefficient **NON NUL** sur b. Comme a et b jouent des rôles symétriques, on obtient alors : $\text{Vect}\Big(X \cup \{a\}\Big) \subset \text{Vect}\Big(X \cup \{b\}\Big)$.

□

Théorème 3.1.22 (Opération sur les Vect)**.** *Soient E un $\mathbb{K}$-espace vectoriel et $x_1, x_2, \cdots, x_n \in E$. Le sous-espace vectoriel $\text{Vect}(x_1, x_2, \cdots, x_n)$ n'est pas modifié*

1. *lorsqu'on permute* $x_1, x_2, \cdots, x_n$;
2. *lorsqu'on retire tous les* x_k *nuls,* $k \in \{1, \cdots, n\}$;
3. *lorsqu'à* $k \in \{1, \cdots, n\}$ *fixé, on remplace* x_k *par une combinaison linéaire de* $x_1, x_2, \cdots, x_n$*, à condition toutefois d'affecter* x_k *lui-même d'un* ***coefficient non nul****.*

Démonstration. 1. Utiliser la commutativité de $+$; pour (2), le fait que 0_E est neutre.

2. Évident.

3. Fixons donc $\lambda_1, \lambda_2, \cdots, \lambda_n \in \mathbb{K}$ tels que $\lambda_k \neq 0$ et posons $V = \text{Vect}(x_1, x_2, \cdots, x_n)$ et $V' = \text{Vect}(x_1, x_2, \cdots, x_{k-1}, \sum_{i=1}^{n} \lambda_i x_i, x_{k+1}, \cdots, x_n)$.
   — Montrons que $V \subset V'$. Commençons par remarquer que $x_k \in V'$ car, puisque $\lambda_k \neq 0$:
   $$x_k = \frac{1}{\lambda_k} \sum_{i=1}^{n} \lambda_i x_i - \sum_{\substack{i=1 \\ i \neq k}}^{n} \frac{\lambda_i}{\lambda_k} x_i.$$
   Or $x_1, x_2, \cdots, x_{k-1}, x_{k+1}, \cdots, x_n \in V'$ également, donc V' contient $x_1, x_2, \cdots, x_n$. Enfin V' contient V car V est le plus petit sous-espace vectoriel de E contenant ces vecteurs.
   — Montrons que $V' \subset V$. Bien sûr :
   $$x_1, x_2, \cdots, x_{k-1}, \sum_{i=1}^{n} \lambda_i x_i, x_{k+1}, \cdots, x_n \in V.$$
   Par conséquent V contient V' car V' est le plus petit sous-espace vectoriel de E contenant ces vecteurs.

□

3.1.3 Famille de vecteurs

Définition 3.1.23 (Partie/famille génératrice)**.** Soient E un $\mathbb{K}$-espace vectoriel et X une partie de E. On note $\text{Vect}(X) = \text{Vect}(x)_{x \in X}$. On dit que X est une partie génératrice de E (ou $(x)_{x \in X}$ une famille génératrice de E) si tout élément de E est combinaison linéaire de X, i.e. si : $E = \text{Vect}(X)$.
Si : $X = (x_i)_{i \in I}$, on dit aussi que la famille $(x_i)_{i \in I}$ est génératrice de E ou engendre E.

Remarque 3.1.24. Soient E un $\mathbb{K}$-espace vectoriel et $x_1, x_2, \cdots, x_n \in E$. On dit que la famille $(x_1, x_2, \cdots, x_n)$ est une famille génératrice de E ou qu'elle engendre E si tout élément de E est combinaison linéaire de $x_1, x_2, \cdots, x_n$, autrement dit si $E = \text{Vect}(x_1, x_2, \cdots, x_n)$.

Exemples 3.1.25. 1. $(X^k)_{k \in \mathbb{N}}$ engendre $\mathbb{K}[X]$ et $(1, X, \cdots, X^n)$ engendre $\mathbb{K}_n[X]$ pour tout $n \in \mathbb{N}$.

2. Pour tout $(x, y) \in \mathbb{R}^2 : (x, y) = x(1, 0) + y(0, 1)$, donc la famille $\Big((1, 0), (0, 1)\Big)$ engendre $\mathbb{R}^2$.
   Plus généralement, posons pour tout $n \in \mathbb{N}^* : e_1 = (1, 0, \cdots, 0)$, $e_2 = (0, 1, 0, \cdots, 0), \cdots, e_n = (0, \cdots, 0, 1)$. La famille $(e_1, \cdots, e_n)$ engendre $\mathbb{K}^n$ car pour tout $(x_1, \cdots, x_n) \in \mathbb{K}^n : (x_1, \cdots, x_n) = \sum_{i=1}^{n} x_i e_i$.

3. La famille $(1, i)$ engendre le $\mathbb{R}$-espace vectoriel $\mathbb{C}$, mais (1) suffit à engendrer le $\mathbb{C}$-espace vectoriel $\mathbb{C}$.

Attention : Il peut exister plusieurs familles génératrices distinctes. Elles peuvent même ne pas avoir le même cardinal !

En pratique : Trouver une partie génératrice d'un sous-espace vectoriel, c'est l'écrire comme un Vect.

Exemple 3.1.26. L'ensemble

$$E = \{(x, y, z, t) \in \mathbb{R}^4 \mid x + 2y - z = 0 \,\text{et}\, x - y + t = 0\}$$

est un sous-espace vectoriel de $\mathbb{R}^4$ engendré par la famille $\Big((1, 0, 1, -1), (0, 1, 2, 1)\Big)$.

▶ En effet, pour tout $(x, y, z, t) \in \mathbb{R}^4$: $(x, y, z, t) \in E \Leftrightarrow \begin{cases} z &= x + 2y \\ t &= -x + y, \end{cases}$ donc :

$$\begin{aligned} E &= \{(x, y, x + 2y, -x + y) \mid x, y \in \mathbb{R}\} \\ &= \{x(1, 0, 1, -1) + y(0, 1, 2, 1) \mid x, y \in \mathbb{R}\} \\ &= \text{Vect}\Big((1, 0, 1, -1), (0, 1, 2, 1)\Big). \end{aligned}$$

Ceci montre **À LA FOIS** que E est un sous-espace vectoriel de $\mathbb{R}^4$ et que $\Big((1, 0, 1, -1), (0, 1, 2, 1)\Big)$ engendre E.

Exemple 3.1.27. L'ensemble $F = \{P \in \mathbb{R}_3[X] \mid 2P(X + 1) = XP'\}$ est un sous-espace vectoriel de $\mathbb{R}_3[X]$ engendré par $X^2 - 4X + 3$.

▶ En effet, pour tout $P = aX^3 + bX^2 + cX + d \in \mathbb{R}_3[X]$:

$$\begin{aligned} &P \in F \Leftrightarrow 2P(X + 1) = XP' \\ &\Leftrightarrow 2a(X + 1)^3 + 2b(X + 1)^2 + 2c(X + 1) + 2d = X(3aX^2 + 2bX + c) \\ &\Leftrightarrow \begin{cases} a &= 0 \\ c &= -4b \\ d &= 3b. \end{cases} \end{aligned}$$

Conclusion :

$$\begin{aligned} F &= \big\{bX^2 - 4bX + 3b \mid \quad b \in \mathbb{R}\big\} \\ &= \big\{b\left(X^2 - 4X + 3\right) \mid \quad b \in \mathbb{R}\big\} \\ &= \text{Vect}\left(X^2 - 4X + 3\right). \end{aligned}$$

Ceci montre **À LA FOIS** que F est un sous-espace vectoriel de $\mathbb{R}_3[X]$ et que $\left(X^2 - 4X + 3\right)$ en est une famille génératrice.

Le résultat qui suit n'est qu'une simple reformulation du théorème "Opération sur les Vect".

Proposition 3.1.28 (Propriétés des familles génératrices)**.** *Soit $x_1, \cdots, x_n \in E$. On pose $F = \text{Vect}(x_1, \cdots, x_n)$.*
On ne change pas l'espace vectoriel engendré F si :

1. *on permute plusieurs vecteurs dans la famille $(x_1, \cdots, x_n)$.*

2. *on ajoute à la famille un vecteur combinaison linéaire des* $x_1, \cdots, x_n$.
3. *on multiplie l'un des vecteurs par un scalaire non nul.*
4. *on retranche à la famille un vecteur combinaison linéaire des autres.*

Démonstration. 1. Évident.

2. Notons x_{n+1} le vecteur que l'on rajoute. Posons $F = \text{Vect}\,(x_1, \ldots, x_n)$ et $G = \text{Vect}\,(x_1, \ldots, x_n, x_{n+1})$. On chercher à prouver que $F = G$.
   — $F \subset G$ car tout vecteur combinaison linéaire des $x_1, \ldots, x_n$ est combinaison linéaire des $x_1, \ldots, x_n, x_{n+1}$:
   $$u = \alpha_1 x_1 + \cdots + \alpha_n x_n = \alpha_1 x_1 + \cdots + \alpha_n x_n + 0 x_{n+1}.$$
   — Réciproquement, montrons que $G \subset F$. Soit $u \in G$. Il existe donc $\alpha_1, \ldots, \alpha_n, \alpha_{n+1}$ tels que $u = \alpha_1 x_1 + \cdots + \alpha_n x_n + \alpha_{n+1} x_{n+1}$. Or, par hypothèse, $x_{n+1} = \sum_{i=1}^{n} \lambda_i x_i$ donc :
   $$\begin{aligned} u &= \alpha_1 x_1 + \cdots + \alpha_n x_n + \alpha_{n+1} \sum_{i=1}^{n} \lambda_i x_i \\ &= (\alpha_1 + \alpha_{n+1}\lambda_1)\, x_1 + \cdots + (\alpha_n + \alpha_{n+1}\lambda_n)\, x_n. \end{aligned}$$
   Donc $u \in F$.
3. Évident.
4. Soient E un $\mathbb{K}$-espace vectoriel et $(x_1, x_2, \cdots, x_n)$ une famille génératrice E. Si par exemple x_n est combinaison linéaire de $x_1, x_2, \cdots, x_{n-1}$, alors on peut écrire aussitôt, en vertu des opérations sur les Vect, que $E = \text{Vect}(x_1, x_2, \cdots, x_n) = \text{Vect}(x_1, x_2, \cdots, x_{n-1}, 0_E) = \text{Vect}(x_1, x_2, \cdots, x_{n-1})$; dans ce cas $(x_1, x_2, \cdots, x_{n-1})$ est une famille génératrice de E, comme voulu.

□

► En d'autres termes, si on rajoute à une famille génératrice de E des vecteurs de E, cette nouvelle famille reste génératrice. Si une famille génératrice de E contient des vecteurs combinaisons linéaires des autres, on peut les retirer, la nouvelle famille reste génératrice. Si l'on retire des vecteurs qui ne sont pas combinaisons linéaires des autres, la nouvelle famille ne sera en revanche plus génératrice !

Définition 3.1.29 (famille libre d'un nombre fini de vecteurs)**.** Soient E un $\mathbb{K}$-espace vectoriel et $x_1, \cdots, x_n \in E$. On dit que la famille $(x_1, \cdots, x_n)$ est libre ou que les vecteurs $x_1, \cdots, x_n$ sont linéairement indépendants si :

$$\forall (\lambda_1, \cdots, \lambda_n) \in \mathbb{K}^n, \Big(\sum_{i=1}^{n} \lambda_i x_i = 0_E \Rightarrow \lambda_1 = \cdots = \lambda_n = 0 \Big).$$

Quitte à remplacer λ_i par $\lambda_i - \mu_i$ dans la définition de la liberté, on peut aussi dire que $(x_1, \cdots, x_n)$ est libre si et seulement si : $\forall (\lambda_1, \cdots, \lambda_n), (\mu_1, \cdots, \mu_n) \in \mathbb{K}^n$,
$\Big(\sum_{i=1}^{n} \lambda_i x_i = \sum_{i=1}^{n} \mu_i x_i \Rightarrow \forall i \in \{1, \cdots, n\},\ \lambda_i = \mu_i \Big)$, ce qui n'est finalement rien de plus qu'un **PRINCIPE D'IDENTIFICATION**.

En résumé :

— **FAMILLE GÉNÉRATRICE = EXISTENCE** pour **TOUT** vecteur d'une décomposition comme combinaison linéaire.
— **FAMILLE LIBRE = UNICITÉ** des coefficients dans les combinaisons linéaires, donc possibilité de pratiquer des **IDENTIFICATIONS**.

Définition 3.1.30 (famille liée d'un nombre fini de vecteurs, couple de vecteurs colinéaires)**.** Soient E un $\mathbb{K}$-espace vectoriel et $x_1, \cdots, x_n, x, y \in E$.
— On dit que la famille $(x_1, \cdots, x_n)$ est liée ou que les vecteurs $x_1, \cdots, x_n$ sont linéairement dépendants si la famille $(x_1, \cdots, x_n)$ **N**'est **PAS** libre. Cela revient à dire que **L'UN AU MOINS** des vecteurs $x_1, \cdots, x_n$ est combinaison linéaire des autres.
— On dit que x et y sont colinéaires si la famille (x, y) est liée, i.e. si x ou y est un multiple de l'autre.

Dire que $(x_1, \cdots, x_n)$ est liée, c'est dire que pour certains $\lambda_1, \cdots, \lambda_n \in \mathbb{K}$: $\Big(\sum_{i=1}^{n} \lambda_i x_i = 0_E \quad \text{et} \quad \exists i_0 \in \{1, \cdots, n\},\ \lambda_{i_0} \neq 0\Big)$, auquel cas : $x_{i_0} = -\frac{1}{\lambda_{i_0}} \sum_{\substack{1 \leq i \leq n \\ i \neq i_0}} \lambda_i x_i$, i.e. x_{i_0} est "combinaison linéaire des autres".

Proposition 3.1.31. *1. Une famille est liée dès qu'elle contient le vecteur nul.*
2. Une famille composée d'un seul vecteur est libre si et seulement si ce vecteur n'est pas nul.
3. Une famille composée de deux vecteurs est libre si et seulement s'ils ne sont pas colinéaires.

Démonstration.
1. La famille $(x_1, \cdots, x_n, 0_E)$ est liée car $0 \cdot x_1 + \cdots + 0 \cdot x_n + 1 \cdot 0_E = 0_E$.
2. Si u est non nul, la famille (u) est libre car : $\lambda u = 0_E \Rightarrow \lambda = 0$.
3. Si (u, v) est liée alors il existe $\alpha, \beta \in \mathbb{K}$ tels que $\alpha u + \beta v = 0_E$ avec $(\alpha, \beta) \neq (0, 0)$. Si α est non nul, $u = -\frac{\beta}{\alpha} v$. Si $\alpha = 0$, β est nécessairement non nul et alors $v = -\frac{\alpha}{\beta} u$. Dans les deux cas, u et v sont colinéaires. Réciproquement, si u et v sont colinéaires, on aura, par exemple, $u = \lambda v$ donc $1 \cdot u + (-\lambda) \cdot v = 0_E$, ce qui montre que la famille est liée.

□

Exemple 3.1.32. La famille $(\sin, \cos)$ du $\mathbb{R}$-espace vectoriel $\mathbb{R}^{\mathbb{R}}$ est libre.
▶ Soient $\lambda, \mu \in \mathbb{R}$ tels que $\lambda \sin + \mu \cos = 0$, i.e. : $\forall x \in \mathbb{R}, \lambda \sin x + \mu \cos x = 0$. Alors en particulier, pour $x = 0$: $\lambda \sin 0 + \mu \cos 0 = 0$ et donc $\mu = 0$. De plus, pour $x = \frac{\pi}{2}$, on a : $\lambda \sin \frac{\pi}{2} + \mu \cos \frac{\pi}{2} = 0$ et donc $\lambda = 0$, comme voulu.

Définition 3.1.33 (famille libre/liée d'un nombre quelconque de vecteurs)**.** Soient E un $\mathbb{K}$-espace vectoriel et $(x_i)_{i \in I}$ une famille de vecteurs de E.
— On dit que la famille $(x_i)_{i \in I}$ est libre ou que les vecteurs x_i, i décrivant I, sont linéairement indépendants si :

$$\forall (\lambda_i)_{i \in I} \in \mathbb{K}^I \text{ presque nulle}, \Big(\sum_{i \in I} \lambda_i x_i = 0_E \Rightarrow \forall i \in I, \lambda_i = 0\Big).$$

— On dit que la famille $(x_i)_{i \in I}$ est liée ou que les vecteurs x_i, i décrivant I, sont linéairement dépendants si $(x_i)_{i \in I}$ **N**'est **PAS** libre. Cela revient à dire que **L'UN AU MOINS** d'entre eux est combinaison linéaire des autres.

▶ L'expression "presque nulle" nous ramène toujours à un nombre **FINI** de vecteurs utiles, donc la liberté d'une famille **INFINIE** de vecteurs est équivalente à la liberté de **TOUTES** ses sous-familles **FINIES**. Pour montrer qu'une famille $(x_n)_{n\in\mathbb{N}}$ de vecteurs est libre, il suffit même de montrer que la famille $(x_0, \cdots, x_n)$ est libre pour tout $n \in \mathbb{N}$.

Définition 3.1.34. On dit qu'une famille de polynômes non nuls $(P_k)_{k\in\mathbb{N}}$ est :
— étagée en degrés si on a $\deg(P_k) < \deg(P_{k+1})$ pour tout $k \in \mathbb{N}$.
— échelonnée en degrés si on a $\deg(P_k) = k$ pour tout $k \in \mathbb{N}$.

▶ Une famille de polynômes échelonnée en degrés est étagée en degrés.

Le théorème suivant est **TRÈS IMPORTANT !**

Théorème 3.1.35. *Une famille $(P_k)_{k\in\mathbb{N}}$ de polynômes non nuls étagée en degrés est libre dans $\mathbb{K}[X]$.*

Conséquence : Une famille de polynômes **NON NULS** échelonnés en degrés est libre.

Démonstration. Soit $(P_k)_{k\in\mathbb{N}}$ une famille de polynômes étagée en degrés. On vérifie par récurrence sur $n \geq 0$, que chaque famille $(P_k)_{0\leq k\leq n}$ est libre.
— Initialisation : Pour $n = 0$, P_0 est non nul, donc (P_0) est libre.
— Hérédité : Supposons le résultat acquis pour $n-1 \geq 0$ et soit $(\lambda_k)_{0\leq k\leq n}$ des scalaires tels que $\sum_{k=0}^{n} \lambda_k P_k = 0$. On a : $\lambda_n P_n = -\sum_{k=0}^{n-1} \lambda_k P_k$. Si $\lambda_n \neq 0$, alors : $\deg(P_n) = \deg(\sum_{k=0}^{n-1} \lambda_k P_k)$, ce qui est en contradiction avec $\deg(P_k) < \deg(P_n)$ pour tout k compris entre 0 et $n-1$. On a donc $\lambda_n = 0$ et $\sum_{k=0}^{n-1} \lambda_k P_k = 0$, ce qui impose $\lambda_k = 0$ pour tout k compris entre 0 et $n-1$ puisque $(\lambda_k)_{0\leq k\leq n-1}$ est libre par hypothèse de récurrence.

□

Théorème 3.1.36 (Propriétés des familles libres/liées)**.**
1. *Toute famille contenant une famille liée est liée.*
2. *Toute sous-famille d'une famille libre est libre.*
3. *Si on ajoute à une famille libre un vecteur **NON** combinaison linéaire des vecteurs de cette famille, alors la famille obtenue est encore libre.*

▶ Dire qu'une famille est libre, c'est dire qu'aucun de ses vecteurs n'est combinaison linéaire des autres, donc si on veut que l'ajout d'un vecteur conserve la liberté d'une famille libre, on doit faire attention de ne pas introduire de dépendance entre ses vecteurs (on ne peut donc ajouter qu'un vecteur linéairement indépendant des autres).

Démonstration.
1. Si une famille est liée, alors l'un de ses vecteurs est combinaison linéaire des autres. Quand on ajoute de nouveaux vecteurs à cette famille, ce fait ne s'en trouve pas modifié et donc la sur-famille obtenue est toujours liée.
2. Soient E un $\mathbb{K}$-espace vectoriel, $(x_1, x_2, \cdots, x_n)$ une famille libre de E et $p \leq n$. Montrons que la famille $(x_1, x_2, \cdots, x_p)$ est libre elle aussi. Soient $\lambda_1, \cdots, \lambda_p \in \mathbb{K}$ tels que $\sum_{k=1}^{p} \lambda_k x_k = 0_E$. Alors $\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_p x_p + 0 \cdot x_{p+1} + \cdots + 0 \cdot x_n = 0_E$. Or la famille $(x_1, x_2, \cdots, x_n)$ est libre. Par conséquent $\lambda_1 = \lambda_2 = \cdots = \lambda_p = 0$ comme voulu.

3. Soient E un $\mathbb{K}$-espace vectoriel, $(x_1, x_2, \cdots, x_n)$ une famille libre de E et $x_{n+1} \in E$ **non** combinaison linéaire de $x_1, x_2, \cdots, x_n$. Montrons que $(x_1, x_2, \cdots, x_n, x_{n+1})$ est libre elle aussi.

Soient $\lambda_1, \lambda_2, \cdots, \lambda_n, \lambda_{n+1}$ tels que $\sum_{k=1}^{n+1} \lambda_k x_k = 0_E$. Si λ_{n+1} était non nul, on aurait $x_{n+1} = -\frac{1}{\lambda_{n+1}} \sum_{k=1}^{n} \lambda_k x_k$; le vecteur x_{n+1} serait alors combinaison linéaire de $x_1, x_2, \cdots, x_n$, ce qui est faux par hypothèse. Par conséquent $\lambda_{n+1} = 0$. Du coup, on récupère l'égalité $\sum_{k=1}^{n} \lambda_k x_k = 0_E$. La liberté de $(x_1, x_2, \cdots, x_n)$ montre aussitôt que $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$. Nous avons bien prouvé que $\lambda_1 = \lambda_2 = \cdots = \lambda_n = \lambda_{n+1} = 0$. □

Définition 3.1.37 (Bases)**.** Soient E un $\mathbb{K}$-espace vectoriel et $\mathcal{B} = (e_i)_{i \in I}$ une famille de vecteurs de E.

— On dit que $\mathcal{B}$ est une base de E si $\mathcal{B}$ est à la fois libre et génératrice de E, i.e. si et seulement si tout vecteur de E est **D'UNE ET UNE SEULE MANIÈRE** combinaison linéaire de $\mathcal{B}$.

— Dans ce cas, pour tout $x \in E$, l'unique famille presque nulle $(x_i)_{i \in I} \in \mathbb{K}^I$ pour laquelle : $x = \sum_{i \in I} x_i e_i$ est appelée la famille des coordonnées de x dans $\mathcal{B}$.

La définition suivante est une synthèse des exemples précédents.

Définition 3.1.38 (Bases canoniques)**.**

— **Familles de scalaires** : Pour tout $n \in \mathbb{N}^*$, si on pose : $e_1 = (1, 0, \cdots, 0)$, $e_2 = (0, 1, 0, \cdots, 0)$, $\cdots$, $e_n = (0, \cdots, 0, 1)$, la famille $(e_1, \cdots, e_n)$ est une base de $\mathbb{K}^n$ appelée sa base canonique.

— **Polynômes** : La famille $(X^k)_{k \in \mathbb{N}}$ est une base de $\mathbb{K}[X]$ appelée sa base canonique et pour tout $n \in \mathbb{N}$, la famille $(1, X, \cdots, X^n)$ est une base de $\mathbb{K}_n[X]$ appelée sa base canonique.

Remarque 3.1.39.

— Que signifie "canonique" ? Réponse : "le plus naturel". De fait, les bases exhibées ci-dessus sont les plus naturelles, les plus simples, les plus faciles d'emploi auxquelles on peut penser dans $\mathbb{K}^n$, $\mathbb{K}[X]$ et $\mathbb{K}_n[X]$.

— On convient que la famille vide $\emptyset$ est une base de l'espace nul $E = \{0_E\}$.

Méthode 3.1.40. Pour trouver une base d'un espace vectoriel, on en cherche d'abord une famille génératrice en l'écrivant comme un Vect, puis on essaie de montrer que la famille ainsi obtenue est libre.

Exemple 3.1.41. L'ensemble $F = \{(x, y, z) \in \mathbb{R}^3 \mid x + y - 4z = 0\}$ est un sous-espace vectoriel de $\mathbb{R}^3$ et la famille $\mathcal{B} = \Big((4, 0, 1), (0, 4, 1)\Big)$ en est une base.

▶ Dans cet exemple, tâchons de trouver nous-mêmes une base de F, sans aide extérieure.

— Commençons par chercher une famille génératrice de F. On a :

$$\begin{aligned} F &= \{(x, y, \frac{x+y}{4}) \mid x, y \in \mathbb{R}\} \\ &= \text{Vect}\Big((1, 0, \frac{1}{4}), (0, 1, \frac{1}{4})\Big) \\ &= \text{Vect}\Big((4, 0, 1), (0, 4, 1)\Big). \end{aligned}$$

Ceci montre À LA FOIS que F est un sous-espace vectoriel de $\mathbb{R}^3$ et que la famille $\mathcal{B}$ est bien génératrice de F.

— Montrons que la famille $\mathcal{B}$ est libre. Soient $\lambda, \mu \in \mathbb{R}$ tels que $\lambda(4,0,1) + \mu(0,4,1) = (0,0,0)$. Alors : $4\lambda = 4\mu = 0$, donc $\lambda = \mu = 0$ comme voulu. C'est fini.

Exercice 3.1.42. Montrer que pour tous $n \in \mathbb{N}$ et $\lambda \in \mathbb{K}$, la famille $\Big(1, X - \lambda, (X-\lambda)^2, \cdots, (X-\lambda)^n\Big)$ est une base de $\mathbb{K}_n[X]$ et que les coordonnées d'un polynôme $P \in \mathbb{K}_n[X]$ dans cette base sont $\Big(P(\lambda), P'(\lambda), \frac{P''(\lambda)}{2} \cdots, \frac{P^{(n)}(\lambda)}{n!}\Big)$.

Solution. — Pour la liberté, soient $a_0, \cdots, a_n \in \mathbb{K}$. Faisons l'hypothèse que : $\sum_{i=0}^{n} a_i (X-\lambda)^i = 0$. Par composition à droite par $X+\lambda$: $\sum_{i=0}^{n} a_i X^i = 0$, donc aussitôt : $a_0 = \cdots = a_n = 0$.

— D'après la formule de Taylor : $P = \sum_{i=0}^{n} \frac{P^{(i)}(\lambda)}{i!}(X-\lambda)^i$ pour tout $P \in \mathbb{K}_n[X]$, donc $\Big(1, X-\lambda, (X-\lambda)^2, \cdots, (X-\lambda)^n\Big)$ engendre $\mathbb{K}_n[X]$ et la formule donne aussi les coordonnées de P.

3.1.4 Dimension finie

Définition 3.1.43 (Espace vectoriel de dimension finie)**.** Soit E un $\mathbb{K}$-espace vectoriel. On dit que E est de dimension finie s'il possède une partie génératrice **FINIE**, et de dimension infinie sinon.

Exemples 3.1.44.
1. $\mathbb{K}^n$ et $\mathbb{K}_n[X]$ sont des espaces vectoriels de dimension finie car ils possèdent des familles génératrices finies (notamment les bases canoniques).
2. $\mathbb{K}[X]$ est de dimension infinie. En effet, Pour toute famille **FINIE** $(P_1, \cdots, P_n)$ de polynômes non nuls, si nous posons :

$$d = \max_{1 \le i \le n} \deg(Pi),$$

alors : $\text{Vect}(P_1, \cdots, P_n) \subset \mathbb{K}_d[X] \neq \mathbb{K}[X]$, donc $(P_1, \cdots, P_n)$ n'engendre pas $\mathbb{K}[X]$. Conclusion : aucune famille **FINIE** de $\mathbb{K}[X]$ n'engendre $\mathbb{K}[X]$.

Passons maintenant à un théorème qui est à la base de la théorie sur la dimension des espaces vectoriels.

Théorème 3.1.45 (Nombre maximal de vecteurs linéairement indépendants)**.** *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie engendré par n éléments. Alors toute partie libre de E possède au plus n éléments.*

Démonstration. On procède par récurrence sur n.

Soit $\mathcal{B} = (e_i)_{1 \le i \le n}$ une famille génératrice de E. Si $n = 1$, alors pour tout couple (x, y) de vecteurs non nuls de E on peut trouver deux réels non nuls λ et μ tels que $x = \lambda e_1$ et $y = \mu e_1$ et on a la combinaison linéaire nulle $\mu x - \lambda y = 0$. avec μ et $-\lambda$ non nuls, ce qui signifie que le système (x, y) est lié. Il ne peut donc exister de famille libre à 2 éléments dans E et a fortiori il ne peut en exister à plus de 2 éléments.

Supposons le résultat acquis au rang $n-1 \ge 1$, c'est-à-dire que dans tout espace vectoriel F admettant un système générateur de $n-1$ vecteurs une famille de plus de n vecteurs est liée. Supposons que E soit un espace vectoriel admettant une famille génératrice à n éléments. Supposons qu'il existe une famille libre ayant $m \ge n+1$ éléments. On peut extraire de cette famille une famille

libre à $n+1$ éléments puisque toute sous-famille d'une famille libre est libre. Soit $\mathcal{L} = (f_i)_{1\leq i\leq n+1}$ une telle famille. Comme $\mathcal{B} = (e_i)_{1\leq i\leq n}$ est génératrice, il existe des réels a_{ij} tels que :

$$\begin{cases} f_1 = a_{11}e_1 + \cdots + a_{1n}e_n \\ \vdots \\ f_{n+1} = a_{n+1,1}e_1 + \cdots + a_{n+1,n}e_n. \end{cases}$$

Si tous les $a_{i,n}$ sont nuls alors les f_i sont dans l'espace vectoriel F engendré par les $n-1$ vecteurs $e_1, \cdots, e_{n-1}$ et en conséquence liés (hypothèse de récurrence), en contradiction avec $\mathcal{L}$ libre. Il existe donc un indice i compris entre 1 et $n+1$ tel que $a_{i,n} \neq 0$ et en changeant au besoin la numérotation des éléments de $\mathcal{L}$ on peut supposer que $i = n+1$. Les n vecteurs :

$$\begin{cases} g_1 = a_{n+1,n}f_1 - a_{1n}f_{n+1} \\ \vdots \\ g_n = a_{n+1,n}f_1 - a_{nn}f_n, \end{cases}$$

sont dans l'espace vectoriel F engendré par les $n-1$ vecteurs $e_1, \cdots, e_{n-1}$ (on a annulé les composantes en e_n) et en conséquence liés (hypothèse de récurrence), c'est-à-dire qu'il existe des réels $\lambda_1, \cdots, \lambda_n$ non tous nuls tels que :

$$\lambda_1 g_1 + \cdots + \lambda_n g_n = 0,$$

ce qui entraine :

$$a_{n+1,n}\left(\lambda_1 f_1 + \cdots + \lambda_n f_n\right) - \left(\lambda_1 a_{1n} + \cdots + \lambda_n a_{nn}\right) f_{n+1} = 0$$

les réels $a_{n+1,n}\lambda_1, \cdots, a_{n+1,n}\lambda_n$ n'étant pas tous nuls. Ce qui nous dit encore que les f_i sont liés et est en contradiction avec $\mathcal{L}$ libre. Il est donc impossible de trouver un tel système $\mathcal{L}$ libre. □

Théorème 3.1.46 (Théorème de la base extraite)**.** *Soit E un $\mathbb{K}$-espace vectoriel admettant une famille génératrice finie $\mathcal{F}$. Alors $\mathcal{F}$ contient une sous-famille qui est une base de E.*

▶ Un espace de dimension finie admet donc une base finie.

Démonstration. Raisonnons par récurrence. Soient E un $\mathbb{K}$-espace vectoriel de dimension finie et $\mathcal{P}_k$ la proposition "Toute famille génératrice de k vecteurs de E contient une sous-famille qui est une base de E".

— Initialisation : Si $k = 0$, E est engendré par 0 vecteur : $E = \{0_E\}$. La famille vide est une base de E.

— Hérédité : Supposons la propriété vraie au rang k et considérons une famille génératrice $\mathcal{F} = (x_1, \cdots, x_{k+1})$ de E. Si $\mathcal{F}$ est libre, c'est une base de E.
Sinon $\mathcal{F}$ est liée et l'un des vecteurs de $\mathcal{F}$ est combinaison linéaire des autres. Nous pouvons retirer ce vecteur et noter $\mathcal{F}'$ la sous-famille obtenue. D'après ce qui précède, celle-ci est toujours génératrice. Comme $\mathcal{F}'$ contient k vecteurs, elle contient par hypothèse de récurrence une sous-famille $\mathcal{F}''$ qui est une base de E. $\mathcal{F}''$ étant elle-même une sous-famille de $\mathcal{F}$, la propriété est démontrée au rang $k+1$.

□

▶ Si on connaît une famille génératrice de E, on peut toujours enlever des vecteurs combinaisons linéaires des autres jusqu'à obtenir une famille libre donc une base de E.

Remarque 3.1.47.
— Intuitivement, on a bien envie de définir la dimension d'un espace vectoriel comme le nombre de vecteurs qu'on trouve dans ses bases, mais $\cdots$ qui nous dit que toutes les bases ont le même nombre de vecteurs ? Pourquoi un espace vectoriel ne pourrait-il pas être à la fois de dimension 2 et de dimension 3 ?
— En principe, la notion de cardinal concerne les ensembles et les ensembles seulement, mais par abus de langage, une famille $(x_1, \cdots, x_n)$ de n objets est souvent appelée une famille de cardinal n, c'est très commode.

Définition 3.1.48 (Dimension)**.** Soit E un $\mathbb{K}$-espace vectoriel de dimension finie.
— Si : $E \neq \{0_E\}$, toutes les bases de E sont finies de même cardinal. Ce cardinal unique est appelé la dimension de E et notée $\dim E$.
— Si : $E = \{0_E\}$, on décrète par convention que : $\dim E = 0$.
Si : $\dim E = 1$, on dit que E est une droite (vectorielle), et si : $\dim E = 2$, que E est un plan (vectoriel).

Démonstration. Supposons : $E \neq \{0_E\}$. Comme E est de dimension finie, nous avons vu plus haut que toutes les familles libres de E sont finies, donc en particulier ses bases aussi. Soient alors $\mathcal{B}$ une base de E à n éléments et $\mathcal{B}'$ une base de E à n' éléments. Comme $\mathcal{B}$ engendre E et comme $\mathcal{B}'$ est libre, nous avons vu déjà que : $n' \leq n$. De plus, par symétrie des rôles de $\mathcal{B}$ et $\mathcal{B}'$, on a aussi : $n \leq n'$. Conclusion : $n = n'$. □

Théorème 3.1.49.
— *Pour tout* $n \in \mathbb{N}^*$ *:* $\dim \mathbb{K}^n = n$.
— *Pour tout* $n \in \mathbb{N}$ *:* $\dim \mathbb{K}_n[X] = n + 1$.

▶ À retenir : pour déterminer la dimension d'un espace vectoriel de dimension finie, il suffit d'exhiber une base de cet espace et de compter le nombre de vecteurs obtenus.

Exercice 3.1.50. On note $E = \mathbb{R}^{\mathbb{R}}$, $f, g, h : \mathbb{R} \to \mathbb{R}$ les applications définies par : $f(x) = 1, g(x) = e^x, h(x) = e^{-x}$ et $F = \text{Vect}(f, g, h)$. Déterminer $\dim F$.

Solution. Par définition de F, la famille (f, g, h) engendre F. Montrons que (f, g, h) est libre. Soit $(a, b, c) \in \mathbb{R}^3$ tel que $af + bg + ch = 0$. On a :

$$\forall x \in \mathbb{R}, a + be^x + ce^{-x} = 0.$$

Par le changement de variable $t = e^x$, on déduit :

$$\forall t \in]0, +\infty[, a + bt + c\frac{1}{t} = 0,$$

c-à-d : $\forall t \in]0, +\infty[, bt^2 + at + c = 0$.
Le polynôme $bX^2 + aX + c$ s'annule donc en une infinité de réels (les réels > 0), donc c'est le polynôme nul, d'où : $a = b = c = 0$. Ainsi, (f, g, h) est libre. Enfin, puisque (f, g, h) est libre et engendre F, (f, g, h) est une base de F et on conclut que : $\dim F = 3$.

Théorème 3.1.51 (Dimension et cardinal d'une famille génératrice)**.** *Soient E un $\mathbb{K}$-espace vectoriel de dimension finie n et $\mathcal{F}$ une famille génératrice de E. Alors* $\text{Card}(\mathcal{F}) \geq n$*, et si* $\text{Card}(\mathcal{F}) = n$*, c'est une base de E.*

Démonstration. Tout repose sur le théorème de la base extraite.
- $\mathcal{F}$ possède une sous-famille $\mathcal{F}'$ base de E qui comporte donc n vecteurs. Ainsi, $\mathcal{F}$ contient plus de n vecteurs.
- Si $\mathcal{F}$ possède n vecteurs, $\mathcal{F} = \mathcal{F}'$, ce qui fait de $\mathcal{F}$ une base de E.

□

▶ Dans la pratique, ce résultat a une importance capitale : il suffit qu'une famille soit génératrice et comporte autant de vecteurs que la dimension de E pour que celle-ci soit une base de E.

Voici maintenant l'équivalent du théorème de la base extraite pour les familles libres :

Théorème 3.1.52 (Théorème de la base incomplète)**.** *Soient E un $\mathbb{K}$-espace vectoriel de dimension finie n et $\mathcal{L}$ une famille libre de E. On peut compléter $\mathcal{L}$ en une base de E.*

Démonstration. Soient $\mathcal{L} = (u_1, \cdots, u_k)$ une famille libre et $\mathcal{B} = (e_1, \cdots, e_n)$ une base de E.
- Si $\mathcal{L}$ est génératrice, c'est une base de E. Supposons désormais que $\mathcal{L}$ n'est pas génératrice.
- Il existe au moins un vecteur de $\mathcal{B}$ qui n'est pas une combinaison linéaire des vecteurs u_i. En effet, par l'absurde, si tous les vecteurs e_j étaient combinaisons linéaires des vecteurs u_i, on aurait

$$E = \text{Vect}(e_1, \cdots, e_n) \subset \text{Vect}(u_1, \cdots, u_k).$$

  Et comme $\text{Vect}(u_1, \cdots, u_k) \subset E$ (tous les vecteurs u_i sont dans E), on aurait $E = \text{Vect}(u_1, \cdots, u_k)$, ce qui ferait de $\mathcal{L}$ une famille génératrice de E, absurde par hypothèse.
- Soit e_{j_0} un vecteur non combinaison linéaire des vecteurs u_i. On pose $\mathcal{L}' = (u_1, \cdots, u_k, e_{j_0})$. Montrons que la famille $\mathcal{L}'$ reste libre. Pour cela, supposons que $\lambda_1 u_1 + \cdots + \lambda_k u_k + \lambda_{k+1} u_{j_0} = 0_E$. Si $\lambda_{k+1} \neq 0$, $e_{j_0} = -\frac{1}{\lambda_{k+1}}(\lambda_1 u_1 + \cdots + \lambda_k u_k)$ est une combinaison linéaire des vecteurs u_i. Absurde. Donc $\lambda_{k+1} = 0$. Ainsi, $\lambda_1 u_1 + \cdots + \lambda_k u_k = 0_E$. Comme $\mathcal{L}$ est libre, $\lambda_1 = \cdots = \lambda_k = 0$. On a bien montré que $\lambda_1 = \cdots = \lambda_k = \lambda_{k+1} = 0$.
- On peut poursuivre le raisonnement et ajouter à $\mathcal{L}'$ des vecteurs non combinaison linéaire jusqu'à obtenir une famille génératrice, qui restera libre (comme $\mathcal{L}'$). On aura ainsi construit une base de E contenant $\mathcal{L}$.

□

Théorème 3.1.53 (Dimension et cardinal d'une famille libre)**.** *Soient E un $\mathbb{K}$-espace vectoriel de dimension finie n et $\mathcal{L}$ une famille libre de E. Alors* $\text{Card}(\mathcal{L}) \leq n$*, et si* $\text{Card}(\mathcal{L}) = n$*, c'est une base de E.*

Démonstration. Tout repose sur le théorème de la base incomplète.
- $\mathcal{L}$ possède une sur-famille $\mathcal{L}'$ base de E qui comporte donc n vecteurs. Ainsi, $\mathcal{L}$ contient moins de n vecteurs.

— Si $\mathcal{L}$ possède n vecteurs, $\mathcal{L} = \mathcal{L}'$, ce qui fait de $\mathcal{L}$ une base de E.

□

▶ Si on connaît une famille libre d'un espace vectoriel E qui contient $n = \dim(E)$ vecteurs, c'est une base !

En résumé : Dans un $\mathbb{K}$-espace vectoriel de dimension finie n, toute famille libre possède au plus n éléments et toute famille génératrice en possède au moins n.

Théorème 3.1.54 (Dimension d'un s.e.v.)**.** *Soient E un $\mathbb{K}$-espace vectoriel de dimension finie et 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 : $F = E$.*

▶ Pour montrer que deux espaces vectoriels E et F de dimension finie sont égaux, il suffit donc de prouver que $F \subset E$ puis que $\dim F = \dim E$. La double inclusion ne sera alors pas nécessaire.

Démonstration. Si : $F = \{0_E\}$, nous n'avons rien à démontrer. Si au contraire : $F \neq \{0_E\}$, l'ensemble $\mathcal{N}$ des nombres d'éléments des familles libres de F est non vide. Cet ensemble est par ailleurs majoré par $\dim E$ car toute famille libre de F est une famille libre de E, donc constituée d'au plus $\dim E$ vecteurs. Conclusion : $\mathcal{N}$ possède un plus grand élément n inférieur ou égal à $\dim E$.
Donnons-nous alors une famille libre $\mathcal{L}$ de F à n éléments. Pour tout $x \in F$, la famille $\mathcal{L}$ augmentée de x est liée par maximalité de n dans $\mathcal{N}$, donc comme $\mathcal{L}$ est libre, x est forcément combinaison linéaire de $\mathcal{L}$. Conclusion : libre et génératrice, $\mathcal{L}$ est une base de F, donc enfin F est de dimension finie et : $\dim F = n \leq \dim E$.
Et si : $\dim E = \dim F = n$? Dans ce cas $\mathcal{L}$ est une famille libre de E à $n = \dim E$ éléments, donc c'est déjà une base de E et : $E = \text{Vect}(\mathcal{L}) = F$. □

Exemple 3.1.55. Dans $\mathbb{R}^3$, on note : $u = (1, 1, 0)$, $v = (1, 0, 1)$, $x = (3, 1, 2)$ et $y = (1, 3, -2)$. Soient F et G les sous-espaces vectoriels de $\mathbb{R}^3$ suivants :

$$F = \text{Vect}(u, v) \quad \text{et} \quad G = \text{Vect}(x, y).$$

Montrons que $F = G$.
▶ On remarque que : $x = u + 2v$ et $y = 3u - 2v$. Ainsi : $x, y \in F$ et par suite $G \subset F$. De plus, il est clair que (u, v) est libre et que (x, y) est libre, donc : $\dim G = \dim F = 2$. On conclut que : $G = F$.

Théorème 3.1.56 (Dimension d'un espace vectoriel produit)**.** *Soient E et F deux $\mathbb{K}$-espaces vectoriels de dimension finie. Alors $E \times F = \{(x, y) \mid x \in E, y \in F\}$ est de dimension finie et :* $\dim(E \times F) = \dim E + \dim F$.
Le résultat se généralise au cas d'un nombre fini quelconque d'espaces vectoriels.

Démonstration. Supposons E et F de dimensions non nulles et donnons-nous $(e_1, \cdots, e_m)$ une base de E et $(f_1, \cdots, f_n)$ une base de F. Nous allons montrer que la famille $\mathcal{B} = \Big((e_1, 0_F), \cdots, (e_m, 0_F), (0_E, f_1), \cdots, (0_E, f_n)\Big)$ est une base de $E \times F$. Cela montrera bien que $E \times F$ est de dimension finie $m + n = \dim E + \dim F$.

— Montrons que $\mathcal{B}$ engendre $E \times F$. Pour tout $(x, y) \in E \times F$, si nous notons $(x_1, \cdots, x_m)$ les coordonnées de x dans $(e_1, \cdots, e_m)$ et $(y_1, \cdots, y_n)$ celles de y dans $(f_1, \cdots, f_n)$, alors :

$$(x, y) = \Big(\sum_{i=1}^{m} x_i e_i, \sum_{j=1}^{n} y_j f_j\Big) = \sum_{i=1}^{m} x_i (e_i, 0_F) + \sum_{j=1}^{n} y_j (0_E, f_j).$$

— Pour la liberté, soient $\lambda_1, \cdots, \lambda_m, \mu_1, \cdots, \mu_n \in \mathbb{K}$ pour lesquels :

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

Alors :

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

donc : $\sum_{i=1}^{m} \lambda_i e_i = 0_E$ et $\sum_{j=1}^{n} \mu_j f_j = 0_F$. Ainsi, par liberté de $(e_i)_{1 \le i \le m}$ et $(f_j)_{1 \le j \le n}$: $\lambda_1 = \cdots = \lambda_m = \mu_1 = \cdots = \mu_n = 0$.

□

3.1.5 Somme de deux s.e.v

Définition 3.1.57 (Somme de deux s.e.v.)**.** Soient E un $\mathbb{K}$-espace vectoriel et F et G deux sous-espaces vectoriels de E.

— L'ensemble $F + G = \{x_1 + x_2 \mid x_1 \in F \text{ et } x_2 \in G\}$ est un sous-espace vectoriel de E appelé la somme de F et G.

— Cette somme $F + G$ est également le plus petit sous-espace vectoriel de E contenant F et G. Cela signifie que tout sous-espace vectoriel de E contenant F et G contient aussi $F + G$.

Attention : Ne confondez pas **SOMME** et **RÉUNION** ! La somme est un sous-espace vectoriel, mais pas la réunion en général.

Démonstration. Montrons d'abord que $F + G$ est un sous-espace vectoriel de E.

— $0_E \in F + G$ car $0_E = \underbrace{0_E}_{\in F} + \underbrace{0_E}_{\in G}$.

— Soient $x, y \in F + G$ et $\lambda \in \mathbb{K}$. Il existe $x_1, y_1 \in F$ et $x_2, y_2 \in G$ tels que $x = x_1 + x_2$ et $y = y_1 + y_2$. Donc : $\lambda x + y = \lambda(x_1 + x_2) + (y_1 + y_2) = \underbrace{(\lambda x_1 + y_1)}_{\in F} + \underbrace{(\lambda x_2 + y_2)}_{\in G}$ car F et G sont des sous-espaces vectoriels, donc stables par combinaison linéaire. Ainsi, on a bien $\lambda x + y \in F + G$.

$F + G$ est donc un sous-espace vectoriel de E.
De plus, $F \subset F + G$ car tout vecteur $x \in F$ peut s'écrire sous la forme $x = \underbrace{x}_{\in F} + \underbrace{0_E}_{\in G}$. Idem pour G. □

On peut aussi définir $F + G$ comme le sous-espace vectoriel de E engendré par la réunion $F \cup G$ (qui en général n'est pas un espace vectoriel).

Théorème 3.1.58. *Si F, G sont deux sous-espaces vectoriels de E, alors la somme $F + G$ est le sous-espace vectoriel de E engendré par $F \cup G$.*

Démonstration. Dire que $x \in \text{Vect}(F \cup G)$ équivaut à dire qu'il s'écrit $x = \sum_{k=1}^{p} \lambda_k x_k$ où les x_k sont des éléments de $F \cup G$ et les λ_k des réels. En séparant les x_k qui sont dans F deux ceux qui sont dans G, cette somme peut s'écrire $x = y + z$ avec $y \in F$ et $z \in G$, ce qui signifie que $x \in F + G$.

Réciproquement $x \in F + G$ s'écrit $x = y + z$ avec $(y, z) \in F \times G$ et il est bien dans $\text{Vect}(F \cup G)$.

□

Théorème 3.1.59 (Parties génératrices d'une somme de deux s.e.v)**.** *Soient E un $\mathbb{K}$-espace vectoriel et F et G deux sous-espaces vectoriels de E. Si F est engendré par $(f_1, f_2, \cdots, f_m)$ et si G est engendré par $(g_1, g_2, \cdots, g_n)$, alors $F + G$ est engendré par $(f_1, f_2, \cdots, f_m, g_1, g_2, \cdots, g_n)$.*

Démonstration. Soit $x \in F + G$. Il existe alors $f \in F$ et $g \in G$ tels que $x = f + g$. Or $(f_1, f_2, \cdots, f_m)$ engendre F, donc il existe $\lambda_1, \lambda_2, \cdots, \lambda_m \in \mathbb{K}$ tels que $f = \lambda_1 f_1 + \lambda_2 f_2 + \cdots + \lambda_m f_m$; de même $(g_1, g_2, \cdots, g_n)$ engendre G, donc il existe $\mu_1, \mu_2, \cdots, \mu_n \in \mathbb{K}$ tels que $g = \mu_1 g_1 + \mu_2 g_2 + \cdots + \mu_n g_n$. Finalement on obtient $x = \lambda_1 f_1 + \lambda_2 f_2 + \cdots + \lambda_m f_m + \mu_1 g_1 + \mu_2 g_2 + \cdots + \mu_n g_n$. C'est le résultat voulu. □

Théorème 3.1.60 (Formule de Grassmann)**.** *Soient E un $\mathbb{K}$-espace vectoriel pas nécessairement de dimension finie et F et G deux sous-espaces vectoriels de dimension finie de E. La somme $F + G$ est alors elle aussi de dimension finie, et plus précisément :* $\dim(F + G) = \dim F + \dim G - \dim(F \cap G)$.

Démonstration. Pour ne pas alourdir la preuve par l'étude de cas particuliers sans intérêt, on peut supposer que $F \cap G = \{0_E\}$, $F \cap G \neq F$ et $F \cap G \neq G$.
Toute base $(e_1, \cdots, e_p)$ de $F \cap G$ peut être complétée en une base $(e_1, \cdots, e_p, f_1, \cdots, f_q)$ de F et en une base $(e_1, \cdots, e_p, g_1, \cdots, g_r)$ de G. Par concaténation, la famille $(e_1, \cdots, e_p, f_1, \cdots, f_q, g_1, \cdots, g_r)$ est alors une famille génératrice de $F + G$, donc en particulier, $F + G$ est de dimension finie.

Nous allons montrer qu'en réalité la famille $(e_1, \cdots, e_p, f_1, \cdots, f_q, g_1, \cdots, g_r)$ est aussi libre, et donc que c'est une base de $F + G$. La formule de Grassmann en découlera aussitôt : $\dim(F + G) = p + q + r = (p + q) + (p + r) - p = \dim F + \dim G - \dim(F \cap G)$.
Soient donc $\lambda_1, \cdots, \lambda_p, \mu_1, \cdots, \mu_q, \nu_1, \cdots, \nu_r \in \mathbb{K}^{p+q+r}$ pour lesquels :

$$\underbrace{\lambda_1 e_1 + \cdots + \lambda_p e_p}_{\in F \cap G} + \underbrace{\mu_1 f_1 + \cdots, \mu_q f_q}_{\in F} + \underbrace{\nu_1 g_1 + \cdots + \nu_r g_r}_{\in G} = 0_E.$$

- Pour commencer, on a $\nu_1 g_1 + \cdots + \nu_r g_r$ appartient à $F \cap G$ donc est combinaison linéaire de $e_1, \cdots, e_p$, ce qui n'est possible que si $\nu_1 = \cdots = \nu_r = 0$ par liberté de $(e_1, \cdots, e_p, g_1, \cdots, g_r)$.
- De même, $\mu_1 f_1 + \cdots + \mu_q f_q$ appartient à $F \cap G$ donc est combinaison linéaire de $e_1, \cdots, e_p$, ce qui n'est possible que si $\mu_1 = \cdots = \mu_q = 0$ par liberté de $(e_1, \cdots, e_p, f_1, \cdots, f_q)$.
- Finalement : $\lambda_1 e_1 + \cdots + \lambda_p e_p = 0_E$, donc $\lambda_1 = \cdots = \lambda_p = 0$ par liberté de $(e_1, \cdots, e_p)$.

□

Définition 3.1.61 (Somme directe de deux s.e.v)**.** Soient E un $\mathbb{K}$-espace vectoriel et F et G deux sous-espaces vectoriels de E. On dit que F et G sont en somme

directe si tout vecteur de $F+G$ s'écrit **de manière unique** comme somme d'un élément de F et d'un élément de G. On note alors souvent $F \oplus G$ la somme $F+G$, pour indiquer qu'il y a somme directe.

▶ Le petit rond qu'on ajoute à la notation "F + G" pour indiquer que la somme est directe ne fait pas de $F+G$ et $F \oplus G$ des ensembles différents, la notation "$F \oplus G$" contient juste une INFORMATION d'unicité en plus.

Théorème 3.1.62 (Caractérisation de la somme directe)**.** *Soient E un $\mathbb{K}$-espace vectoriel et F et G deux sous-espaces vectoriels de E. Les assertions suivantes sont équivalentes :*

1. *F et G sont en somme directe.*
2. *$F \cap G = \{0_E\}$.*

Dans ce cas, par ailleurs, si F et G sont de dimension finie : $\dim(F+G) = \dim F + \dim G$.

Démonstration. — $(1 \Rightarrow 2)$ Supposons F et G en somme directe et montrons que $F \cap G = \{0_E\}$, ce qui revient à montrer que $F \cap G \subset \{0_E\}$ puisque 0_E appartient au sous-espace vectoriel $F \cap G$ de toute façon. Soit $x \in F \cap G$. Alors $x = \overbrace{x}^{\in F} + \overbrace{0_E}^{\in G} = \overbrace{0_E}^{\in F} + \overbrace{x}^{\in G}$, donc : $x = 0_E$ par définition de la somme directe.

— $(2 \Rightarrow 1)$ Supposons que : $F \cap G = \{0_E\}$. Soient $x_1, x_2 \in F$ et $y_1, y_2 \in G$ pour lesquels : $x_1 + y_1 = x_2 + y_2$. Comme : $x_1 - x_2 = y_2 - y_1 \in F \cap G = \{0_E\}$, alors $x_1 = x_2$ et $y_1 = y_2$.

La formule : $\dim(F+G) = \dim F + \dim G$ n'est enfin qu'un cas particulier de la formule de Grassmann. □

Exemple 3.1.63. $F = \{(x,y,z) \in \mathbb{R}^3 \mid x+y+z=0\}$ et $G = \{(x,y,z) \in \mathbb{R}^3 \mid x=y=z\}$ sont en somme directe.

▶ On a : $F = \{(x,y,-x-y) \mid x,y \in \mathbb{R}\} = \text{Vect}\Big((1,0,-1),(0,1,-1)\Big)$ et $G = \text{Vect}\Big((1,1,1)\Big)$ et donc F et G sont bien des s.e.v de $\mathbb{R}^3$. Il nous reste donc de montrer que $F \cap G = \{0_{\mathbb{R}^3}\}$, et même que $F \cap G \subset \{0_{\mathbb{R}^3}\}$. Soit $(x,y,z) \in F \cap G$. Alors $x+y+z=0$ et $x=y=z$. Du coup : $x+x+x=0$, donc $x=0$ et enfin $x=y=z=0$, i.e. $(x,y,z) = 0_{\mathbb{R}^3}$.

Théorème 3.1.64 (Bases d'une somme directe de deux s.e.v)**.** *Soient E un $\mathbb{K}$-espace vectoriel et F et G deux sous-espaces vectoriels de E en somme directe. Si $(f_1, f_2, \cdots, f_m)$ est une base de F et $(g_1, g_2, \cdots, g_n)$ une base de G, alors $(f_1, f_2, \cdots, f_m, g_1, g_2, \cdots, g_n)$ est une base de $F \oplus G$.*
Une telle base dont les premiers vecteurs forment une base de F et les suivants une base de G est dite adaptée à la somme directe $F \oplus G$.

Attention : Ce résultat est faux si on ne suppose pas la somme directe. Pour les sommes en général, on dispose seulement du résultat vu précédemment sur les familles génératrices.

Démonstration. Nous avons déjà vu que $(f_1, f_2, \cdots, f_m, g_1, g_2, \cdots, g_n)$ est une famille génératrice de $F \oplus G$. Il ne nous reste donc qu'à montrer que cette famille est libre.

Soient donc $\lambda_1, \lambda_2, \cdots, \lambda_m, \mu_1, \mu_2, \cdots, \mu_n \in \mathbb{K}$ tels que $\sum_{k=1}^{m} \lambda_k f_k + \sum_{\ell=1}^{n} \mu_\ell g_\ell = 0_E$. Cette somme est une décomposition de 0_E comme somme d'un élément de F et d'un élément de G. Or la somme est directe donc il y a unicité d'une telle décomposition ; bref $\sum_{k=1}^{m} \lambda_k f_k = \sum_{\ell=1}^{n} \mu_\ell g_\ell = 0_E$. Enfin, les familles $(f_1, f_2, \cdots, f_m)$ et $(g_1, g_2, \cdots, g_n)$ étant libres, on en déduit comme voulu que $\lambda_1 = \lambda_2 = \cdots = \lambda_m = \mu_1 = \mu_2 = \cdots = \mu_n = 0$. □

Définition 3.1.65 (Sous-espaces vectoriels supplémentaires)**.** Soient E un $\mathbb{K}$-espace vectoriel et F et G deux sous-espaces vectoriels de E. Les assertions suivantes sont équivalentes :

1. Tout vecteur de E est d'une et une seule manière la somme d'un élément de F et d'un élément de G : $\forall x \in E, \exists!(f, g) \in F \times G, x = f + g$.
2. E est la somme directe de F et G : $E = F \oplus G$, ce qui revient à dire que : $E = F + G$ **ET** que F et G sont en somme directe.

On dit dans ces conditions que F et G sont supplémentaires dans E. On dit aussi que F est **UN** supplémentaire de G dans E et que G est **UN** supplémentaire de F dans E.

Démonstration.
— Supposons que, pour tout $x \in E$, il existe un unique couple $(f, g) \in F \times G$ tel que $x = f + g$ et montrons que $E = F \oplus G$. Il est clair que $E = F + G$. Soit $x \in F \cap G$. On peut alors écrire $x = x + 0_E$ avec $x \in F$ et $0_E \in G$, mais également $x = 0_E + x$ avec $0_E \in F$ et $x \in G$. L'unicité de la décomposition impose que $x = 0_E$. Donc $E = F \oplus G$.
— Réciproquement, supposons $E = F \oplus G$. Soit $x \in E$. Comme $E = F + G$, il existe $f \in F$ et $g \in G$ tels que $x = f + g$. Montrons que cette décomposition est unique. Soient $f' \in F$ et $g' \in G$ tels que $x = f' + g'$. On a alors $f + g = f' + g'$, c'est-à-dire $f - f' = g' - g$. Or $f - f' \in F$ et $g' - g \in G$. On en déduit que $f - f' \in F \cap G = \{0_E\}$. Donc $f = f'$ et, de même, $g = g'$.

□

Théorème 3.1.66 (Base adaptée)**.** *Soient E un $\mathbb{K}$-espace vectoriel et F et G deux sous-espaces vectoriels de E de bases respectives $(f_1, \cdots, f_m)$ et $(g_1, \cdots, g_n)$. Alors, $E = F \oplus G$ si et seulement si la famille $(f_1, f_2, \cdots, f_m, g_1, g_2, \cdots, g_n)$ est une base de E. Dans ce cas, la famille $(f_1, f_2, \cdots, f_m, g_1, g_2, \cdots, g_n)$ est qualifiée de base adaptée.*

Démonstration. Nous avons déjà vu que si F et G sont en somme directe alors la famille $(f_1, f_2, \cdots, f_m, g_1, g_2, \cdots, g_n)$ est une base de $F \oplus G = E$. Réciproquement, supposons que $(f_1, f_2, \cdots, f_m, g_1, g_2, \cdots, g_n)$ est une base de E. Alors tout $x \in E$ s'écrit d'une manière unique :

$$x = \underbrace{\lambda_1 f_1 + \cdots + \lambda_m f_m}_{\in F} + \underbrace{\mu_1 g_1 + \cdots + \mu_n g_n}_{\in G}$$

c'est-à-dire tout $x \in E$ s'écrit d'une manière unique $x = x_1 + x_2$ avec $x_1 \in F$ et $x_2 \in G$; donc $E = F \oplus G$. □

▶ Nous observerons l'importance des bases adaptées dans la prochaine partie mais remarquons déjà que ce théorème nous fournit une nouvelle méthode pour justifier que deux espaces sont supplémentaires : il suffit de montrer que la

concaténation de deux bases de F et G constitue une base de E pour que ceux-ci soient supplémentaires dans E.

Théorème 3.1.67 (Existence de supplémentaires en dimension finie)**.** *Soient E un $\mathbb{K}$-espace vectoriel de dimension finie et F un sous-espace vectoriel de E. Alors F possède un supplémentaire dans E.*
Le supplémentaire de F n'est pas unique, mais, les supplémentaires de F dans E ont tous pour dimension : $\dim E - \dim F$.

Démonstration. Soit $(f_1, \cdots, f_m)$ une base de F et soit $n = \dim E$. D'après le théorème de la base incomplète, il existe $f_{m+1}, \cdots, f_n$ vecteurs de E tels que $(f_1, \cdots, f_m, f_{m+1}, \cdots, f_n)$ est une base de E. En posant $G = \text{Vect}(f_{m+1}, \cdots, f_n)$ on obtient, d'après le théorème précédent, un supplémentaire de F dans E. Puisque le choix de $f_{m+1}, \cdots, f_n$ n'est pas unique, le supplémentaire de F n'est pas unique ; cependant, tous les supplémentaires de F sont de dimension $n-m$, m étant la dimension de F. □

Méthode 3.1.68. Pour montrer que deux s.e.v F, G d'un e.v E de dimension finie sont supplémentaires dans E, on peut essayer de montrer qu'il existe une base $\mathcal{F}$ de F et une base $\mathcal{G}$ de G telles que $\mathcal{F} \cup \mathcal{G}$, obtenue par concaténation des deux bases $\mathcal{F}$ et $\mathcal{G}$, soit une base de E.

Exercice 3.1.69. Soit : $F = \{P \in \mathbb{R}_3[X] \mid P(X+1) = P(1-X)\}$. Montrer que $\text{Vect}(X, X^3)$ est un supplémentaire de F dans $\mathbb{R}_3[X]$.

Solution. — Nous avons besoin d'abord d'une base de F. Soit $P = aX^3 + bX^2 + cX + d \in \mathbb{R}_3[X]$. On a :

$$\begin{aligned} P \in F &\Leftrightarrow a(X+1)^3 + b(X+1)^2 + c(X+1) + d \\ &\quad = a(1-X)^3 + b(1-X)^2 + c(1-X) + d \\ &\Leftrightarrow \begin{cases} a & = -a \\ 3a+b & = 3a+b \\ 3a+2b+c & = -3a-2b-c \\ a+b+c+d & = a+b+c+d \end{cases} \Leftrightarrow a = 0 \text{ et } 2b+c = 0. \end{aligned}$$

Ce calcul montre que $(1, X^2 - 2X)$ engendre F. Cette famille étant libre, c'est une base de F.

— Complétons-la en une base $(1, X^2 - 2X, X, X^3)$ de $\mathbb{R}_3[X]$. Conclusion : $\text{Vect}(X, X^3)$ est UN supplémentaire de F dans $\mathbb{R}_3[X]$.

Le résultat suivant est aussi d'un usage fréquent :

Théorème 3.1.70 (Caractérisation de la supplémentarité en dimension finie)**.** *Soient E un espace de dimension finie et F, G deux sous-espaces vectoriels de E. F et G sont supplémentaires dans E si et seulement si l'une des propositions suivantes est vérifiée :*

1. $E = F + G$ *et* $F \cap G = \{0_E\}$.
2. $E = F + G$ *et* $\dim(E) = \dim(F) + \dim(G)$.
3. $F \cap G = \{0_E\}$ *et* $\dim(E) = \dim(F) + \dim(G)$.

► On a souvent recours à la dernière caractérisation. On commence par montrer que l'intersection des deux espaces est réduite à $\{0_E\}$ puis on prouve l'égalité des dimensions.

Démonstration. Le point (1) correspond à la définition.

— Montrons que (1) $\Leftrightarrow$ (2). On suppose que $E = F + G$. On a :

$$\begin{aligned} F \cap G = \{0_E\} \Leftrightarrow \dim(F \cap G) = 0 &\Leftrightarrow \dim(F+G) = \dim F + \dim G \\ &\Leftrightarrow \dim E = \dim F + \dim G. \end{aligned}$$

— Montrons que (2) $\Leftrightarrow$ (3). On suppose que $\dim E = \dim F + \dim G$. Remarquons que $F + G \subset E$ donc pour avoir $F + G = E$, il faut et il suffit que $\dim(F + G) = \dim(E)$. On a :

$$\begin{aligned} F \cap G = \{0_E\} \Leftrightarrow \dim(F \cap G) = 0 &\Leftrightarrow \dim(F+G) = \dim F + \dim G \\ &\Leftrightarrow \dim(F+G) = \dim E. \end{aligned}$$

□

Exercice 3.1.71. On considère les deux ensembles suivants :

$$\begin{aligned} A &= \{(x+y, x-y, 2y) \mid (x,y) \in \mathbb{R}^2\}; \\ B &= \{(x,y,z) \in \mathbb{R}^3 \mid 2x = y \,\text{et}\, y = 3z\}. \end{aligned}$$

1. Montrer que A et B sont des sous-espaces vectoriels de $\mathbb{R}^3$.
2. Montrer que $\mathbb{R}^3 = A \oplus B$.

Solution. 1. On a :

— $A = \text{Vect}\Big((1,1,0),(1,-1,2)\Big)$ est un plan vectoriel.

— $B = \text{Vect}\Big((\frac{1}{2},1,\frac{1}{3})\Big) = \text{Vect}\Big((3,6,2)\Big)$ est une droite vectorielle.

2. Montrons maintenant que A et B sont supplémentaires :

— $A \cap B = \{0_{\mathbb{R}^3}\}$ car si on considère $u \in A \cap B$, alors,

$$u \in A \Leftrightarrow \exists (x,y) \in \mathbb{R}^2 : u = (x+y, x-y, 2y)$$

$$u \in B \Leftrightarrow \left\{ \begin{array}{ll} 2(x+y) & = x-y, \\ x-y & = 6y \end{array} \right. \Leftrightarrow \left\{ \begin{array}{ll} x+3y & = 0, \\ x-7y & = 0 \end{array} \right. \Leftrightarrow x = y = 0$$

Donc $u = (0,0,0)$.

— $\dim A + \dim B = 2 + 1 = 3 = \dim \mathbb{R}^3$.

Nous avons ainsi démontré que $\mathbb{R}^3 = A \oplus B$.

Exercice 3.1.72. $E = \mathbb{R}_4[X]$ et $F = \{P \in E \mid P(0) = P'(0) = P'(1) = 0\}$.

1. Montrer que F est un espace vectoriel, déterminer une base de F et préciser sa dimension.
2. Montrer que $G = \text{Vect}(1, X, 1+X+X^2)$ est un supplémentaire de F dans E.

Exercice 3.1.73. Soit $E = \mathcal{F}(\mathbb{R},\mathbb{R})$ l'espace vectoriel des fonctions de $\mathbb{R}$ dans $\mathbb{R}$. On note F le sous-espace vectoriel des fonctions paires (i.e. $f(-x) = f(x)$ pour tout $x \in \mathbb{R}$) et G le sous-espace vectoriel des fonctions impaires (i.e. $f(-x) = -f(x)$ pour tout $x \in \mathbb{R}$). Montrer que F et G sont supplémentaires.

3.2 Applications linéaires

La structure d'espace vectoriel ne devient vraiment intéressante que si l'on introduit la notion d'application linéaire. Il s'agit des applications entre espaces vectoriels qui, dans un sens que nous allons préciser, "conservent la structure d'espace vectoriel".

Dans cette partie, nous allons donner essentiellement les définitions et les résultats élémentaires de base.

3.2.1 Définitions et premières propriétés

Définition 3.2.1 (Application linéaire)**.** Soient E et F deux $\mathbb{K}$-espaces vectoriels. On appelle application linéaire de E dans F toute application $f : E \to F$ qui préserve les combinaisons linéaires :

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

L'ensemble des applications linéaires de E dans F est noté $\mathcal{L}(E, F)$.

- **Cas particulier où $E = F$:** Une application linéaire de E dans E est aussi appelée un endomorphisme de E. L'ensemble des endomorphismes de E est noté $\mathcal{L}(E)$.
- **Cas particulier où $F = \mathbb{K}$:** Une application linéaire de E dans $\mathbb{K}$ est aussi appelée une forme linéaire de E.

Explication : Une application linéaire n'est rien d'autre qu'un "morphisme d'espaces vectoriels", à ceci près qu'on n'emploie jamais cette expression. Bref, c'est une application qui conserve les combinaisons linéaires.

Remarque 3.2.2.
- Clairement : $f(0_E) = f(0_E + 0_E) = f(0_E) + f(0_E)$, donc après simplification : $f(0_E) = 0_F$.
- Ensuite, si A est un sous-espace vectoriel de E, alors $f_{|A}$ est aussi linéaire (mais sur A). En effet, s'il est vrai pour tous $x, y \in E$ et $\lambda, \mu \in \mathbb{K}$ que : $f(\lambda x + \mu y) = \lambda f(x) + \mu f(y)$, c'est a fortiori vrai pour tous $x, y \in A$.
- Enfin, pour vérifier que f est linéaire, il est suffisant de vérifier que : $f(\lambda x + y) = \lambda f(x) + f(y)$ pour tous $x, y \in E$ et $\lambda \in \mathbb{K}$ avec UN SEUL SCALAIRE. Dans ce cas, pour tous $x, y \in E$ et $\lambda, \mu \in \mathbb{K}$: $f(\lambda x) = f(\lambda x + 0_E) = \lambda f(x) + f(0_E) = \lambda f(x) + 0_F = \lambda f(x)$, puis de même : $f(\mu y) = \mu f(y)$, donc enfin : $f(\lambda x + \mu y) = \lambda f(x) + \mu f(y)$.

Définition 3.2.3 (Homothétie)**.** Soient E un $\mathbb{K}$-espace vectoriel et $\lambda \in \mathbb{K}$. On appelle homothétie de E de rapport λ l'application λId_E, i.e. : $\begin{cases} E & \to E \\ x & \mapsto \lambda x. \end{cases}$
Cette application est un endomorphisme de E. En particulier : $Id_E \in \mathcal{L}(E)$.

Démonstration. Pour tous $x, y \in E$ et $\alpha \in \mathbb{K}$:

$$(\lambda Id_E)(\alpha x + y) = \lambda(\alpha x + y) = \alpha(\lambda Id_E)(x) + (\lambda Id_E)(y).$$

□

Exemple 3.2.4. L'application $(x, y) \overset{f}{\mapsto} (x, x + y, x - 2y)$ est linéaire de $\mathbb{R}^2$ dans $\mathbb{R}^3$.

▶ Pour tous $(x,y),(x',y') \in \mathbb{R}^2$ et $\lambda \in \mathbb{R}$:

$$\begin{aligned} f(\lambda(x,y)+(x',y')) &= f(\lambda x+x', \lambda y+y') \\ &= \Big(\lambda x+x', (\lambda x+x')+(\lambda y+y'), (\lambda x+x')-2(\lambda y+y')\Big) \\ &= \lambda(x, x+y, x-2y)+(x', x'+y', x'-2y') \\ &= \lambda f(x,y)+f(x',y'). \end{aligned}$$

Exemple 3.2.5. — Soit $E = E_1 \oplus E_2$; alors tout vecteur $x \in E$ s'écrit d'une manière unique $x = x_1 + x_2$ où $x_1 \in E_1$ et $x_2 \in E_2$.
L'application :

$$\begin{aligned} pr_1 : \quad E &\longrightarrow E \\ x_1+x_2 &\longmapsto x_1 \end{aligned}$$

est une application linéaire dite projecteur sur E_1 parallèlement à E_2.
— Soit $v_0 \neq 0$ un vecteur de E ; l'application translation définie par

$$\begin{aligned} tr : \quad E &\longrightarrow E \\ v &\longmapsto v+v_0 \end{aligned}$$

n'est pas linéaire (noter, par exemple, que : $tr(0) = v_0 \neq 0$).

Théorème 3.2.6 (Opérations sur les applications linéaires)**.** *1. Soient E et F deux $\mathbb{K}$-espaces vectoriels. Alors $\mathcal{L}(E,F)$ est un sous-espace vectoriel de F^E, donc un $\mathbb{K}$-espace vectoriel.*

2. *Soient E, F et G trois $\mathbb{K}$-espaces vectoriels. Pour tous $f \in \mathcal{L}(E,F)$ et $g \in \mathcal{L}(F,G)$: $g \circ f \in \mathcal{L}(E,G)$.*
3. *Soient E, F et G trois $\mathbb{K}$-espaces vectoriels et $f, f' \in \mathcal{L}(E,F)$ et $g, g' \in \mathcal{L}(F,G)$. Alors : $g\circ(f+f') = (g\circ f)+(g\circ f')$ et $(g+g')\circ f = (g\circ f)+(g'\circ f)$.*
4. *Soit E un $\mathbb{K}$-espace vectoriel. Alors $(\mathcal{L}(E), +, \circ)$ est un anneau, non commutatif en général, avec $1_{\mathcal{L}(E)} = Id_E$.*

Démonstration. 1. Pour commencer, on a $\mathcal{L}(E,F) \subset F^E$ et l'application nulle $x \mapsto 0_F$ de E dans F est linéaire. Soient $f, g \in \mathcal{L}(E,F)$ et $\lambda \in \mathbb{K}$. Alors, pour tous $x, y \in E$ et $\alpha \in \mathbb{K}$:

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

2. Pour tous $x, y \in E$ et $\lambda \in \mathbb{K}$: $\big(g\circ f\big)(\lambda x+y) = g\big(f(\lambda x+y)\big) = g\big(\lambda f(x)+f(y)\big) = \lambda g\big(f(x)\big)+g\big(f(y)\big) = \lambda\big(g\circ f\big)(x)+\big(g\circ f\big)(y)$.
3. Facile.
4. L'assertion (1) montre en particulier que $\big(\mathcal{L}(E),+\big)$ est un groupe commutatif. On sait par ailleurs que la composition est une loi de composition interne sur $\mathcal{L}(E)$ via (2), qu'elle est associative, que $\mathrm{Id}_E \in \mathcal{L}(E)$ en est l'élément neutre, et enfin que la composition est distributive sur l'addition via (3).

□

Deux formules à connaître : Soient E un $\mathbb{K}$-espace vectoriel, $f, g \in \mathcal{L}(E)$ et $n \in \mathbb{N}$. On suppose que f et g commutent. Alors :

$$(f+g)^n = \sum_{k=0}^{n} C_n^k f^k \circ g^{n-k}$$

et

$$f^n - g^n = (f-g) \circ \sum_{k=0}^{n-1} f^k \circ g^{n-k-1}$$

(attention, les puissances désignent des compositions).

▶ Ces deux formules se démontrent ici exactement comme elles se démontrent dans $\mathbb{C}$.

Attention : Dans ces formules, il est essentiel que f et g commutent.

Théorème 3.2.7 (Images directe et réciproque d'un s.e.v par une application linéaire)**.** *Soient E et F deux $\mathbb{K}$-espaces vectoriels, $f \in \mathcal{L}(E,F)$, A un sous-espace vectoriel de E et B un sous-espace vectoriel de F.*

1. *L'image (directe) $f(A)$ de A par f est un sous-espace vectoriel de F.*
2. *L'image réciproque $f^{-1}(B)$ de B par f est un sous-espace vectoriel de E.*

Démonstration.

1. Montrons que $f(A)$ est un sous-espace vectoriel de F. Déjà, on a $f(A) \subset F$ et de plus $0_F = f(0_E) \in f(A)$. Soient $y, y' \in f(A)$ et $\lambda \in \mathbb{K}$. Puisque $y, y' \in f(A)$, il existe $a, a' \in A$ tels que $y = f(a)$ et $y' = f(a')$. Par linéarité de f : $\lambda y + y' = \lambda f(a) + f(a') = f(\lambda a + a')$, et par ailleurs : $\lambda a + a' \in A$ car A est un sous-espace vectoriel de E, donc comme voulu : $\lambda y + y' \in f(A)$.
2. Montrons que $f^{-1}(B)$ est un sous-espace vectoriel de E. Déjà, on a $f^{-1}(B) \subset E$ et de plus $0_E \in f^{-1}(B)$ car $f(0_E) = 0_F \in B$. Soient $x, x' \in f^{-1}(B)$ et $\lambda \in \mathbb{K}$. Par linéarité de f, on a : $f(\lambda x + x') = \lambda f(x) + f(x')$. Or par définition de x et x', $f(x), f(x') \in B$ et B est un sous-espace vectoriel de E. Donc $f(\lambda x + x') \in B$, ce qui montre que $\lambda x + x' \in f^{-1}(B)$.

□

3.2.2 Noyau et image d'une application linéaire

Définition 3.2.8 (Noyau et image d'une application linéaire)**.** Soient E et F deux $\mathbb{K}$-espaces vectoriels et $f \in \mathcal{L}(E,F)$.

- On appelle noyau de f, noté $\operatorname{Ker} f$, l'ensemble $f^{-1}\Big(\{0_F\}\Big) = \{x \in E \mid f(x) = 0_F\}$.
- On appelle image de f, noté $\operatorname{Im} f$, l'ensemble $f(E) = \{y \in F \mid \exists x \in E : y = f(x)\} = \{f(x), x \in E\}$.

Exercice 3.2.9. Soit E un $\mathbb{K}$-espace vectoriel, et f et g deux applications linéaires de E vers $\mathbb{K}$, telles que :

$$\forall x \in E \quad f(x)g(x) = 0.$$

Montrer que $f = 0$ ou $g = 0$.

Solution. Supposons les applications linéaires f et g non nulles. Soit $(x, y) \in E^2$ tels que $f(x) \neq 0$ et $g(y) \neq 0$. Puisque $f(x)g(x) = f(y)g(y) = 0$, on a $f(y) = 0$ et $g(x) = 0$. Or, $f(x + y) = f(x) + f(y)$ donc $f(x + y) = f(x) \neq 0$. De même, $g(x + y) \neq 0$. Ainsi $f(x + y)g(x + y) \neq 0$, contradiction.

Exercice 3.2.10. Soit E un $\mathbb{K}$-espace vectoriel et f un endomorphisme de E. Montrer que :

$$\operatorname{Ker} f^2 = \operatorname{Ker} f \Longleftrightarrow \operatorname{Im} f \cap \operatorname{Ker} f = 0.$$

Solution. Remarquons que l'on a toujours $\operatorname{Ker} f \subset \operatorname{Ker} f^2$. En effet, si $x \in \operatorname{Ker} f$, alors $f(x) = 0$, donc, par linéarité de f, on a $f(f(x)) = 0$ et donc $x \in \operatorname{Ker} f^2$.

— Supposons $\operatorname{Ker} f^2 \subset \operatorname{Ker} f$, et montrons $\operatorname{Im} f \cap \operatorname{Ker} f = \{0\}$. Soit $y \in \operatorname{Im} f \cap \operatorname{Ker} f$: il existe alors $x \in E$ tel que $y = f(x)$. Ainsi :

$$f^2(x) = f(y) = 0,$$

puisque $y \in \operatorname{Ker} f$; donc $x \in \operatorname{Ker} f^2$. Or, par hypothèse, $\operatorname{Ker} f^2 \subset \operatorname{Ker} f$, donc $x \in \operatorname{Ker} f$. Finalement :

$$y = f(x) = 0.$$

Par conséquent, on a $\operatorname{Im} f \cap \operatorname{Ker} f = \{0\}$.

— Supposons $\operatorname{Im} f \cap \operatorname{Ker} f = \{0\}$ et montrons $\operatorname{Ker} f^2 \subset \operatorname{Ker} f$. Soit $x \in \operatorname{Ker} f^2$. Alors, puisque $f(f(x)) = 0$, on a $f(x) \in \operatorname{Ker} f$. De plus, on a par définition $f(x) \in \operatorname{Im} f$. On a donc :

$$f(x) \in \operatorname{Im} f \cap \operatorname{Ker} f.$$

D'où $f(x) = 0$. Ainsi $x \in \operatorname{Ker} f$.

Théorème 3.2.11. *Soient E et F deux $\mathbb{K}$-espaces vectoriels et $f \in \mathcal{L}(E, F)$. Alors, $\operatorname{Ker} f$ est un sous-espace vectoriel de E.*

Démonstration.

— Tout d'abord on a : $0_E \in \operatorname{Ker} f$, car $f(0_E) = 0_F$.

— Soient $x, y \in \operatorname{Ker} f$ et $\lambda \in \mathbb{K}$. On a : $f(\lambda x + y) \overset{f \in \mathcal{L}(E,F)}{=} \lambda f(x) + f(y) = \lambda 0_F + 0_F = 0_F$ donc $\lambda x + y \in \operatorname{Ker} f$.

□

Théorème 3.2.12. *Soient E et F deux $\mathbb{K}$-espaces vectoriels et $f \in \mathcal{L}(E, F)$. L'application linéaire f est injective si et seulement si $\operatorname{Ker} f = \{0_E\}$.*

▶ Pour montrer qu'une application linéaire $f : E \to F$ est injective, il suffit de montrer que $\operatorname{Ker} f = \{0_E\}$ ou plus simplement (puisque $\{0_E\} \subset \operatorname{Ker} f$) :

$$\forall x \in E, \Big(f(x) = 0 \Rightarrow x = 0_E\Big).$$

Démonstration. Démontrons ce résultat par double implication.

— ($\Rightarrow$) Supposons f injective, c'est-à-dire que : $\forall (x, y) \in E^2, f(x) = f(y) \Rightarrow x = y$ et montrons qu'alors : $\operatorname{Ker} f = \{0_E\}$ ou encore : $\operatorname{Ker} f \subset \{0_E\}$. Or pour tout $x \in \operatorname{Ker} f$: $f(x) = 0_F = f(0_E)$, donc comme f est injective : $x = 0_E$.

— ($\Leftarrow$) Supposons maintenant que $\operatorname{Ker} f = \{0_E\}$ et montrons que f est injective. Soient $x, x' \in E$ tels que : $f(x) = f(x')$. Alors : $f(x - x') = f(x) - f(x') = 0_F$ par linéarité, donc : $x - x' \in \operatorname{Ker} f = \{0_E\}$, donc : $x - x' = 0_E$, i.e. : $x = x'$.

□

Exemple 3.2.13. Soit f l'endomorphisme de $\mathbb{R}^3$ défini par $f(x,y,z) = (2x - y, y + z, z - x)$. Étudions l'injectivité de f.
▶ Déterminons pour cela son noyau. Soit $(x,y,z) \in \mathbb{R}^3$. On a :

$$(x,y,z) \in \operatorname{Ker} f \Leftrightarrow f(x,y,z) = 0_{\mathbb{R}^3} \Leftrightarrow (2x - y, y + z, z - x) = 0_{\mathbb{R}^3}.$$

Ce qui conduit à la résolution suivante :

$$\begin{cases} 2x - y &= 0 \\ y + z &= 0 \\ z - x &= 0 \end{cases} \Leftrightarrow x = y = z = 0.$$

Ainsi : $\operatorname{Ker} f = 0_{\mathbb{R}^3}$ et f est injective.

Exercice 3.2.14. Soit $n \in \mathbb{N}^*$ ainsi que $x_1, \cdots, x_n$ des éléments deux à deux distincts de $\mathbb{K}$. On pose

$$\begin{aligned} \varphi : \quad \mathbb{K}[X] &\longrightarrow \mathbb{K}^n \\ P &\longmapsto (P(x_1), P(x_2), \ldots, P(x_n)). \end{aligned}$$

1. Vérifier que φ est une application linéaire.
2. Montrer que sa restriction φ_n à $\mathbb{K}_{n-1}[X]$ est injective.
3. Préciser le noyau de φ.

Solution.
1. La linéarité de φ est immédiate, puisque pour $P \in \mathbb{K}[X], Q \in \mathbb{K}[X]$, $\alpha \in \mathbb{K}, \beta \in \mathbb{K}$ et $a \in \mathbb{K}$, on a $(\alpha P + \beta Q)(a) = \alpha P(a) + \beta Q(a)$.
2. La restriction φ_n de φ à $\mathbb{K}_{n-1}[X]$ reste une application linéaire. Soit $P \in \mathbb{K}_{n-1}[X]$ un élément de $\operatorname{Ker} \varphi_n$. Chacun des x_i est donc racine de P ; ces nombres étant deux à deux distincts, le polynôme P possède n racines. Comme $\deg P \leqslant n - 1$, on a $P = 0$ et donc $\operatorname{Ker} \varphi_n = \{0\}$. Ainsi φ_n est injective.
3. Un polynôme P appartient au noyau de φ si, et seulement s'il admet chaque x_i pour racine. Comme ce sont des scalaires distincts deux à deux, cela revient à dire que P est divisible par $A = \prod_{i=1}^{n} (X - x_i)$. Ainsi $\operatorname{Ker} \varphi = \{AQ; Q \in \mathbb{K}[X]\}$.

Théorème 3.2.15. *Soient E et F deux $\mathbb{K}$-espaces vectoriels et $f \in \mathcal{L}(E,F)$. Alors, $\operatorname{Im} f$ est un sous-espace vectoriel de F.*

Démonstration.
— Comme $0_F = f(0_E)$, on a bien : $0_F \in \operatorname{Im} f$.
— De plus, si $x, y \in \operatorname{Im} f$ et $\lambda \in \mathbb{K}$, il existe $x', y' \in E$ tels que $x = f(x')$ et $y = f(y')$. D'où, $\lambda x + y = \lambda f(x') + f(y') \overset{f \in \mathcal{L}(E,F)}{=} f(\lambda x' + y') \in \operatorname{Im} f$.
□

Théorème 3.2.16 (Familles génératrices de l'image d'une application linéaire)**.** *Soient E et F deux $\mathbb{K}$-espaces vectoriels et $f \in \mathcal{L}(E,F)$. On suppose que E possède une famille génératrice $(x_1, x_2, \cdots, x_n)$. Alors $\Big(f(x_1), f(x_2), \cdots, f(x_n)\Big)$ est une famille génératrice de $\operatorname{Im} f$. En résumé :*

$$\operatorname{Im} f = \operatorname{Vect}\Big(f(x_1), f(x_2), \cdots, f(x_n)\Big).$$

Démonstration. Soit $y \in \operatorname{Im} f$. Alors il existe $x \in E$ tel que $y = f(x)$. Or $(x_1, x_2, \cdots, x_n)$ engendre E par hypothèse, donc $x = \sum_{k=1}^{n} \lambda_k x_k$ pour certains $\lambda_1, \lambda_2, \cdots, \lambda_n \in \mathbb{K}$. Par suite : $y = f(x) = f\Big(\sum_{k=1}^{n} \lambda_k x_k\Big) = \sum_{k=1}^{n} \lambda_k f(x_k)$, ce qui montre bien que la famille $\Big(f(x_1), f(x_2), \cdots, f(x_n)\Big)$ engendre $\operatorname{Im} f$. □

Exercice 3.2.17. Soit E un $\mathbb{K}$-espaces vectoriel. Montrer qu'une forme linéaire φ sur E non identiquement nulle est surjective.

Solution. Dire que $\varphi \neq 0$ signifie qu'il existe un vecteur $x_0 \in E$ tel que $\lambda = \varphi(x_0) \neq 0$. Pour tout réel y, on peut alors écrire :

$$y = \frac{y}{\lambda}\lambda = \frac{y}{\lambda}\varphi(x_0) = \varphi\big(\frac{y}{\lambda}x_0\big).$$

Ainsi $y = \varphi(x)$ avec $x = \dfrac{y}{\lambda}x_0 \in E$, ce qui signifie que φ est surjective.

Exercice 3.2.18. Montrer que si φ est une forme linéaire non nulle sur E, alors il existe un vecteur non nul a dans E tel que :

$$E = \ker(\varphi) \oplus \mathbb{R}a.$$

Solution. La forme linéaire φ étant non nulle, on peut trouver un vecteur a dans E tel que $\varphi(a) \neq 0$. Ce vecteur a est nécessairement non nul. Pour tout vecteur x dans E, le vecteur $h = x - \frac{\varphi(x)}{\varphi(a)}a$ est dans le noyau de φ et en écrivant que $x = h + \frac{\varphi(x)}{\varphi(a)}a$ on déduit que $E = \ker(\varphi) + \mathbb{R}a$. Si x est dans $\ker(\varphi) \cap \mathbb{R}a$ on a alors $x = \lambda a$ et $\lambda\varphi(a) = \varphi(x) = 0$ avec $\varphi(a) \neq 0$ ce qui entraîne $\lambda = 0$ et $x = 0$. On a donc $\ker(\varphi) \cap \mathbb{R}a = \{0\}$ et $E = \ker(\varphi) \oplus \mathbb{R}a$.

3.2.3 Isomorphisme et e. v. isomorphes

Définition 3.2.19 (Isomorphisme, espaces vectoriels isomorphes)**.** Soient E et F deux $\mathbb{K}$-espaces vectoriels.

- On appelle isomorphisme de E sur F toute application linéaire bijective de E sur F.
  Cas particulier où $E = F$: Un isomorphisme de E sur E est aussi appelée un automorphisme de E. L'ensemble des automorphismes de E est noté $\mathrm{GL}(E)$ et appelé le groupe linéaire de E.
- On dit que F est isomorphe à E s'il existe un isomorphisme de E sur F.

▶ Le fait que deux espaces vectoriels soient isomorphes signifie intuitivement que ces deux espaces sont "identiques" d'un strict point de vue vectoriel.

Théorème 3.2.20 (Composition d'isomorphismes, réciproque d'un isomorphisme)**.** *Soient E, F et G trois $\mathbb{K}$-espaces vectoriels.*

1. *Si f est un isomorphisme de E sur F et g un isomorphisme de F sur G, $g \circ f$ est un isomorphisme de E sur G.*
2. *Si f est un isomorphisme de E sur F, alors f^{-1} est un isomorphisme de F sur E.*

▶ En d'autres termes, la relation d'isomorphisme entre espaces vectoriels est une relation d'équivalence (pour la réflexivité, remarquer simplement que Id_E est un isomorphisme de E pour tout $\mathbb{K}$-espace vectoriel E).

Démonstration. 1. La composée de deux applications bijectives (resp. linéaires) est bijective (resp. linéaire).

2. Nous savons que f^{-1} est bijective de F sur E. Montrons que f^{-1} est linéaire. Soient $x', y' \in F$ et $\lambda \in \mathbb{K}$. Posons alors $x = f^{-1}(x')$ et $y = f^{-1}(y')$. On a alors $x' = f(x)$ et $y' = f(y)$ et,

$$\begin{aligned} f^{-1}(\lambda x' + y') &= f^{-1}\Big(\lambda f(x) + f(y)\Big) \\ &= f^{-1}\Big(f(\lambda x + y)\Big) \\ &= \lambda x + y, \quad \text{car} \quad f^{-1} \circ f = Id_E \\ &= \lambda f^{-1}(x') + f^{-1}(y'). \end{aligned}$$

□

Le théorème qui suit montre que les $\mathbb{K}$-espaces vectoriels de dimension finie sont tous "identiques" à un $\mathbb{K}^n$.

Théorème 3.2.21 (Effet d'un isomorphisme sur la dimension)**.** — *Soient E et F deux $\mathbb{K}$-espaces vectoriels. Si E est de dimension finie et si F est isomorphe à E, alors F est de dimension finie et :* $\dim E = \dim F$.

— *Réciproquement, deux $\mathbb{K}$-espaces vectoriels de **MÊMES** dimensions finies sont isomorphes. En particulier, tout $\mathbb{K}$-espace vectoriel de dimension finie $n \neq 0$ est isomorphe à $\mathbb{K}^n$.*

Démonstration. — On peut supposer $E \neq \{0_E\}$ et noter f un isomorphisme de E sur F. De dimension finie $n \neq 0$, E possède une base $(e_1, \cdots, e_n)$. En particulier, par surjectivité de f :

$$F = \operatorname{Im} f = \operatorname{Vect}\Big(f(e_1), \cdots, f(e_n)\Big),$$

donc F est engendré par la famille $\Big(f(e_1), \cdots, f(e_n)\Big)$ donc est de dimension finie. Nous allons en fait montrer que cette famille est libre. Il en découlera que c'est une base de F, et donc que : $\dim F = n = \dim E$. Soient $\lambda_1, \cdots, \lambda_n \in \mathbb{K}$ tels que : $\sum_{i=1}^n \lambda_i f(e_i) = 0_F$. Alors : $f\Big(\sum_{i=1}^n \lambda_i e_i\Big) = 0_F$, donc : $\sum_{i=1}^n \lambda_i e_i \in \operatorname{Ker} f$. Or f est injective et donc : $\sum_{i=1}^n \lambda_i e_i = 0_E$. Finalement comme voulu, la famille $(e_1, \cdots, e_n)$ étant libre : $\lambda_1 = \cdots = \lambda_n$.

— Soient E et F deux $\mathbb{K}$-espaces vectoriels de même dimension finie n. On peut supposer $n \neq 0$. Nous allons montrer que E est isomorphe à $\mathbb{K}^n$ et E et F seront alors isomorphes tout court.
Comme $n \neq 0$, E possède une base $(e_1, \cdots, e_n)$. Soit alors φ l'application $(x_1, \cdots, x_n) \in \mathbb{K}^n \mapsto \sum_{i=1}^n x_i e_i$ de $\mathbb{K}^n$ dans E. On a, φ est linéaire. En effet, pour tous $x = (x_1, \cdots, x_n), x = (x'_1, \cdots, x'_n) \in \mathbb{K}^n$ et $\lambda \in \mathbb{K}$:

$$\varphi(\lambda x + x') = \sum_{i=1}^n (\lambda x_i + x'_i) e_i = \lambda \sum_{i=1}^n x_i e_i + \sum_{i=1}^n x'_i e_i = \lambda \varphi(x) + \varphi(x').$$

De plus, φ est bijective de $\mathbb{K}^n$ sur E, car $(e_1, \cdots, e_n)$ étant une base de E alors : $\forall x \in E, \exists! (x_1, \cdots, x_n) \in \mathbb{K}^n, x = \sum_{i=1}^n x_i e_i = \varphi(x_1, \cdots, x_n)$. D'où φ est un isomorphisme de $\mathbb{K}^n$ sur E.

□

Remarque 3.2.22. L'application φ de la preuve précédente mérite qu'on s'y attarde un instant. Soient E un $\mathbb{K}$-espace vectoriel et $(e_1, \cdots, e_n)$ une famille **QUELCONQUE** de E (plus forcément une base de E). L'application $(x_1, \cdots, x_n) \in \mathbb{K}^n \overset{\varphi}{\mapsto} \sum_{i=1}^n x_i e_i$ est toujours linéaire de $\mathbb{K}^n$ dans E.

- φ surjective si et seulement si : $\forall x \in E, \exists (x_1, \cdots, x_n) \in \mathbb{K}^n, x = \sum_{i=1}^n x_i e_i$, i.e. si et seulement si la famille $(e_1, \cdots, e_n)$ engendre E.
- φ est injective si et seulement si : $\operatorname{Ker} \varphi = \{0_E\}$, i.e. si et seulement si : $\forall (x_1, \cdots, x_n) \in \mathbb{K}^n, \sum_{i=1}^n x_i e_i = 0_E \Rightarrow x_1 = \cdots = x_n = 0$, i.e. si et seulement si la famille $(e_1, \cdots, e_n)$ est libre.

Le théorème qui suit caractérise l'injectivité/surjectivité d'une application linéaire par l'image d'une base.

Théorème 3.2.23. *Soient E et F deux $\mathbb{K}$-espaces vectoriels et $f \in \mathcal{L}(E, F)$. On suppose que E possède une base $(e_i)_{i \in I}$.*

1. *f est surjective de E sur F si et seulement si $\Big(f(e_i)\Big)_{i \in I}$ engendre F.*
2. *f est injective sur E si seulement si $\Big(f(e_i)\Big)_{i \in I}$ est libre.*
3. *f est un isomorphisme de E sur F si et seulement si $\Big(f(e_i)\Big)_{i \in I}$ est une base de F.*

Démonstration. **1** On sait que : $\operatorname{Im} f = \operatorname{Vect} \Big(f(e_i)\Big)_{i \in I}$. Ainsi : $\operatorname{Im} f = F$ si et seulement si $\Big(f(e_i)\Big)_{i \in I}$ engendre F.

2 Supposons f injective et montrons que $\Big(f(e_i)\Big)_{i \in I}$ est libre. Soit $(\lambda_i)_{i \in I} \in \mathbb{K}^I$ presque nulle telle que : $\sum_{i \in I} \lambda_i f(e_i) = 0_F$. Alors : $f\Big(\sum_{i \in I} \lambda_i e_i\Big) = 0_F$ et donc : $\sum_{i \in I} \lambda_i e_i \in \operatorname{Ker} f$. Comme f est injective, alors $\sum_{i \in I} \lambda_i e_i = 0_E$ et donc : $\lambda_i = 0$ pour tout $i \in I$.
Réciproquement, supposons $\Big(f(e_i)\Big)_{i \in I}$ est libre et montrons que f est injective, i.e. que : $\operatorname{Ker} f \subset \{0_E\}$. Soient $x \in \operatorname{Ker} f$ de coordonnées $(x_i)_{i \in I}$ dans $(e_i)_{i \in I}$. Alors : $0_F = f(x) = f\Big(\sum_{i \in I} \lambda_i e_i\Big) = \sum_{i \in I} x_i f(e_i)$, donc par hypothèse pour tout $i \in I$: $x_i = 0$, et donc a fortiori : $x = 0_E$.

□

3.2.4 Notion de rang

Définition 3.2.24 (Application linéaire de rang fini, rang)**.** Soient E et F deux $\mathbb{K}$-espaces vectoriels pas nécessairement de dimension finie et $f \in \mathcal{L}(E, F)$.

- On dit que f est de rang fini si $\operatorname{Im} f$ est de dimension finie, et de rang infini sinon.
- Si f est de rang fini, on appelle rang de f , noté $\operatorname{rg}(f)$, la dimension de $\operatorname{Im} f$.

Théorème 3.2.25 (Inégalités sur le rang et cas d'égalité)**.** *Soient E et F deux $\mathbb{K}$-espaces vectoriels et $f \in \mathcal{L}(E, F)$.*

1. *Si F est de dimension finie, f est de rang fini et :* $\mathrm{rg}(f) \leq \dim F$, *avec égalité si et seulement si f est surjective.*
2. *Si E est de dimension finie, f est de rang fini et :* $\mathrm{rg}(f) \leq \dim E$, *avec égalité si et seulement si f est injective.*

Démonstration.
1. Comme : $\mathrm{Im}\, f \subset F$, alors $\mathrm{Im}\, f$ est de dimension finie et : $\mathrm{rg}(f) = \dim \mathrm{Im}\, f \leq \dim F$, avec égalité si et seulement si : $\mathrm{Im}\, f = F$, i.e. si et seulement si f est surjective.
2. Supposons $E \neq \{0_E\}$ et donnons-nous une base $(e_1, \cdots, e_n)$ de E. Par suite : $\mathrm{Im}\, f = \mathrm{Vect}\Big(f(e_1), \cdots, f(e_n)\Big)$, donc $\mathrm{Im}\, f$ est de dimension finie et : $\mathrm{rg}(f) = \dim \mathrm{Im}\, f \leq n = \dim E$, avec égalité si et seulement si $\Big(f(e_1), \cdots, f(e_n)\Big)$ est libre, i.e. si et seulement si f est injective.

□

Théorème 3.2.26 (Applications linéaires entre e. v. de mêmes dimensions finies)**.** *Soient E et F deux $\mathbb{K}$-espaces vectoriels de dimensions finies* ***ÉGALES*** *et $f \in \mathcal{L}(E, F)$.*

$$f \quad \textit{est bijective} \quad \Leftrightarrow f \quad \textit{est injective} \quad \Leftrightarrow f \quad \textit{est surjective.}$$

Attention : Le théorème ne dit pas que : bijectif = surjectif = injectif en algèbre linéaire ! Elle affirme seulement que c'est vrai lorsque les espaces vectoriels de départ et d'arrivée ont **MÊME DIMENSION FINIE**.

Démonstration. Par hypothèse : $\dim E = \dim F$. Du coup, f est injective si et seulement si : $\mathrm{rg}(f) = \dim E$, i.e. si et seulement si : $\mathrm{rg}(f) = \dim F$, i.e. si et seulement si f est surjective. □

Remarque 3.2.27. Ce résultat est faux en dimension infinie. En voici un contre-exemple d'un endomorphisme surjectif qui n'est pas injectif :

$$\begin{aligned} D : \quad \mathbb{R}[X] &\longrightarrow \mathbb{R}[X] \\ P &\longmapsto D(P) = P'. \end{aligned}$$

En effet :
- Le polynôme $X^0 = 1$ appartient à $\mathrm{Ker}(D)$ et il n'est pas nul, donc l'endomorphisme D n'est pas injectif.
- Soit $Q \in \mathbb{R}[X]$. Il existe $n \in \mathbb{N}$, $a_0, \cdots, a_n \in \mathbb{R}$ tels que : $Q = \sum_{k=0}^{n} a_k X^k$. En notant $P = \sum_{k=0}^{n} \frac{a_k}{k+1} X^{k+1}$, on a $P \in \mathbb{R}[X]$ et $D(P) = P' = Q$. Ceci montre que D est surjectif.

Exemple 3.2.28. En voici un autre contre-exemple d'un endomorphisme injectif qui n'est pas surjectif. :

$$\begin{aligned} f : \quad \mathbb{R}[X] &\longrightarrow \mathbb{R}[X] \\ P &\longmapsto f(P) = XP. \end{aligned}$$

En effet :
- Soit $P \in \mathbb{R}[X]$ tel que $f(P) = 0$. Alors : $XP = 0$. Or, comme l'anneau $\mathbb{R}[X]$ est intègre et $X \neq 0$, alors : $P = 0$. D'où $\mathrm{Ker}\, f = \{0_{\mathbb{R}[X]}\}$ et donc l'endomorphisme f est injectif.

— On a : $\operatorname{Im} f = \{f(P), P \in \mathbb{R}[X]\} = \{XP, P \in \mathbb{R}[X]\} = \{Q \in \mathbb{R}[X] \mid Q(0) = 0\}$. Donc par exemple $X^0 = 1 \notin \operatorname{Im} f$, ce qui montre que $\operatorname{Im} f \subsetneq E$. Il en résulte que l'endomorphisme f n'est pas une surjection de E sur E.

Exercice 3.2.29. Soit $n \in \mathbb{N}^*$. On note $E = \mathbb{R}_n[X]$ et

$$\begin{aligned} f : \quad & E \longrightarrow E \\ & P \longmapsto f(P) = XP' + P. \end{aligned}$$

Vérifier que $f \in \mathcal{L}(E)$ et montrer que f est bijectif.

Solution. — Pour tout $P \in E$, on a $\deg P \leq n$, donc :

$$\deg P' \leq n - 1 \text{ et } \deg(XP') \leq n.$$

Ainsi : $\deg(XP'+P) \leq n$ et donc $f(P) \in E$. D'autre part, soient $P, Q \in E$ et $\lambda \in \mathbb{R}$. On a :

$$\begin{aligned} f(\lambda P + Q) &= X(\lambda P + Q)' + (\lambda P + Q) = \lambda(XP' + P) + (XQ' + Q) \\ &= \lambda f(P) + f(Q). \end{aligned}$$

Donc f est linéaire. D'où : $f \in \mathcal{L}(E)$.

— Puisque E est de dimension finie ($\dim(E) = n + 1$), pour montrer que f est bijectif, il suffit de montrer que, par exemple, que f est injectif. Soit $P \in \operatorname{Ker} f$. Supposons $P \neq 0$ et notons $d = \deg P \leq n$. Il existe $a_0, \cdots, a_d \in \mathbb{R}$ avec $a_d \neq 0$, tels que $P = \sum_{k=0}^{d} a_k X^k$. Le coefficient du terme de degré d de $f(P)$ est $da_d + a_d = (d+1)a_d \neq 0$. Ce qui contredit le fait que $f(P) = 0$. Ceci montre que $\operatorname{Ker} f = \{0_E\}$ et donc f est injectif. On conclut que f est bijectif.

3.2.5 Le théorème du rang

Théorème 3.2.30 (Forme géométrique du théorème du rang)**.** *Soient E et F deux $\mathbb{K}$-espaces vectoriels pas nécessairement de dimension finie et $f \in \mathcal{L}(E, F)$.*
Si $\operatorname{Ker} f$ *possède un supplémentaire I dans E, alors $f_{|I}$ est un isomorphisme de I sur* $\operatorname{Im} f$.

Explication : Dans l'égalité : $E = I \oplus \operatorname{Ker} f$, $\operatorname{Ker} f$ est l'ensemble des éléments de E que f ne voit pas, donc f ne voit passer que I, et comme I ne touche $\operatorname{Ker} f$ que du bout de son zéro, f est injective sur I, donc envoie bijectivement I sur $\operatorname{Im} f$. De manière moins imagée, $\left(f_{|I}\right)^{-1}$ est l'application qui, à tout élément de $\operatorname{Im} f$, associe son unique antécédent dans I.

Démonstration. Par restriction, $f_{|I}$ est linéaire de I dans $\operatorname{Im} f$.

— Pour l'injectivité, sachant que I et $\operatorname{Ker} f$ sont en somme directe : $\operatorname{Ker} f_{|I} = I \cap \operatorname{Ker} f = \{0_E\}$.

— Pour la surjectivité, soit $y \in \operatorname{Im} f$, disons : $y = f(x)$ pour un certain $x \in E$. Comme $E = I + \operatorname{Ker} f$, alors : $x = i + k$ pour certains $i \in I$ et $k \in \operatorname{Ker} f$, donc : $y = f(x) = f(i) + f(k) = f(i) + 0_F = f_{|I}(i)$.

□

Théorème 3.2.31 (Théorème du rang)**.** *Soient E et F deux $\mathbb{K}$-espaces vectoriels et $f \in \mathcal{L}(E,F)$.*
Si E est de dimension finie : $\dim E = \dim \operatorname{Ker} f + \operatorname{rg}(f)$.

Morale de l'histoire : Si on connait le noyau, on connait un peu l'image (et vice versa).
▶ L'hypothèse selon laquelle E est de dimension finie garantit en particulier que $\operatorname{Ker} f$ et $\operatorname{Im} f$ le sont aussi.

Démonstration. Comme E est de dimension finie, $\operatorname{Ker} f$ possède un supplémentaire I dans E. Or, d'après la forme géométrique du théorème du rang, $f_{|I}$ est un isomorphisme de I sur $\operatorname{Im} f$ et donc : $\dim I = \dim \operatorname{Im} f$. Par supplémentarité de I et $\operatorname{Ker} f$ dans E, on obtient : $\operatorname{rg}(f) = \dim \operatorname{Im} f = \dim I = \dim E - \dim \operatorname{Ker} f$. □

Exemple 3.2.32. Soient E un $\mathbb{K}$-espace vectoriel de dimension finie et $f \in \mathcal{L}(E)$. Si : $f^3 = 0_{\mathcal{L}(E)}$, alors : $\operatorname{rg}(f) + \operatorname{rg}(f^2) \leq \dim E$.
▶ En effet, l'égalité : $f^3 = 0_{\mathcal{L}(E)}$ montre que pour tout $x \in E$: $f\left(f^2(x)\right) = 0_E$, i.e. que : $\operatorname{Im}(f^2) \subset \operatorname{Ker} f$ et donc en particulier que : $\operatorname{rg}(f^2) \leq \dim \operatorname{Ker} f$. Appliquons alors le théorème du rang à f, ce qui est possible car E est de dimension finie pour obtenir : $\dim E = \dim \operatorname{Ker} f + \operatorname{rg}(f) \geq \operatorname{rg}(f) + \operatorname{rg}(f^2)$.

Exercice 3.2.33. Soit E un espace vectoriel réel de dimension finie et $u \in \mathcal{L}(E)$.

1. Montrer que :
$$\operatorname{Im}(u) = \operatorname{Im}\left(u^2\right) \Leftrightarrow \ker(u) = \ker\left(u^2\right)$$
où $u^2 = u \circ u$.
2. Montrer que :
$$\operatorname{Im}(u) = \operatorname{Im}\left(u^2\right) \Leftrightarrow E = \ker(u) \oplus \operatorname{Im}(u) \Leftrightarrow \ker(u) = \ker\left(u^2\right).$$

Solution. 1. On a toujours :
$$\operatorname{Im}\left(u^2\right) \subset \operatorname{Im}(u) \quad \text{et} \quad \ker(u) \subset \ker\left(u^2\right).$$
Donc :
$$\operatorname{Im}(u) = \operatorname{Im}\left(u^2\right) \Leftrightarrow \operatorname{rg}(u) = \operatorname{rg}\left(u^2\right)$$
et :
$$\ker(u) = \ker\left(u^2\right) \Leftrightarrow \dim(\ker(u)) = \dim\left(\ker\left(u^2\right)\right)$$
D'autre part, d'après le théorème du rang on a :
$$\dim(E) = \dim(\ker(u)) + \operatorname{rg}(u) = \dim\left(\ker\left(u^2\right)\right) + \operatorname{rg}\left(u^2\right),$$
ce qui permet de déduire que :
$$\operatorname{Im}(u) = \operatorname{Im}\left(u^2\right) \Leftrightarrow \ker(u) = \ker\left(u^2\right).$$

2. Il suffit de montrer que :
$$\operatorname{Im}(u) = \operatorname{Im}\left(u^2\right) \Leftrightarrow E = \ker(u) \oplus \operatorname{Im}(u).$$

Si $\mathrm{Im}(u) = \mathrm{Im}\left(u^2\right)$, alors pour tout x dans E, il existe y dans E tel que $u(x) = u^2(y)$, donc $x - u(y) \in \ker(u)$ et $x = (x - u(y)) + u(y) \in \ker(u) + \mathrm{Im}(u)$. On a donc $E = \ker(u) + \mathrm{Im}(u)$ et avec le théorème du rang, on déduit que $E = \ker(u) \oplus \mathrm{Im}(u)$. Réciproquement, si $E = \ker(u) \oplus \mathrm{Im}(u)$ alors tout $x \in \ker\left(u^2\right)$ s'écrit $x = x_1 + u(x_2)$ avec $u(x_1) = 0$ et $0 = u^2(x) = u^3(x_2)$ entraine que $u^2(x_2) \in \ker(u) \cap \mathrm{Im}(u) = \{0\}$, donc $u(x_2) \in \ker(u) \cap \mathrm{Im}(u) = \{0\}$ et $x = x_1 \in \ker(u)$. On a donc $\ker\left(u^2\right) \subset \ker(u)$ et $\ker(u) = \ker\left(u^2\right)$, ce qui équivaut à $\mathrm{Im}(u) = \mathrm{Im}\left(u^2\right)$.

3.2.6 Existence d'applications linéaires

Pour connaitre une application en général, on n'a pas trop d'autre choix que de connaître l'ensemble de ses valeurs point par point. En revanche, pour une application linéaire, ce lot considérable d'informations peut être résumé par un nombre restreint de valeurs stratégiques. On connaît par exemple parfaitement l'application $(x, y, z) \mapsto (2x + y + z, 3x - z)$ de $\mathbb{R}^3$ dans $\mathbb{R}^2$ **SI ON SAIT QU'ELLE EST LINÉAIRE** et si on sait que : $f(1,0,0) = (2,3)$, $f(0,1,0) = (1,0)$ et $f(0,0,1) = (1,-1)$.

▶ En effet, pour tout $(x, y, z) \in \mathbb{R}^3$:

$$\begin{aligned} f(x,y,z) &= f\Big(x(1,0,0) + y(0,1,0) + z(0,0,1)\Big) \\ &\overset{\text{Linéarité}}{=} xf(1,0,0) + yf(0,1,0) + zf(0,0,1) \\ &= x(2,3) + y(1,0) + z(1,-1) = (2x + y + z, 3x - z). \end{aligned}$$

Le théorème qui suit, fondamental, généralise ce principe. Il signifie que, pour se donner une application linéaire complètement, on peut se contenter de donner les valeurs qu'elle prend sur une base quelconque fixée de l'espace vectoriel de départ.

Théorème 3.2.34 (Détermination d'une application linéaire par l'image d'une base)**.** *Soient E et F deux $\mathbb{K}$-espaces vectoriels. On suppose que E possède une base $(e_1, e_2, \cdots, e_n)$. Pour toute famille $(f_1, f_2, \cdots, f_n)$ de vecteurs de F, il existe une et une seule application linéaire f de E dans F telle que : $\forall k \in \{1, \cdots, n\}, f(e_k) = f_k$.*

Explication : Pour connaître/définir une application linéaire complètement, il suffit de connaître/définir les valeurs qu'elle prend sur une base de l'espace de départ.

Démonstration. Soit $(f_1, f_2, \cdots, f_n)$ une famille fixée de vecteurs de F.

— **Unicité :** Soient $f, g \in \mathcal{L}(E,F)$ telles que : $\forall k \in \{1, \cdots, n\}, f(e_k) = g(e_k) = f_k$. Montrons que $f = g$. Pour tout $x \in E$ de coordonnées $(x_1, x_2, \cdots, x_n)$ dans $(e_1, e_2, \cdots, e_n)$, on a :

$$\begin{aligned} f(x) = f\Big(\sum_{k=1}^{n} x_k e_k\Big) &= \sum_{k=1}^{n} x_k f(e_k) = \sum_{k=1}^{n} x_k g(e_k) \\ &= g\Big(\sum_{k=1}^{n} x_k e_k\Big) = g(x), \end{aligned}$$

et donc $f = g$.

— **Existence :** Soit f l'application de E dans F qui associe, à tout vecteur x de E de coordonnées $(x_1, x_2, \cdots, x_n)$ dans $(e_1, e_2, \cdots, e_n)$, le vecteur $f(x) = \sum_{k=1}^{n} x_k f_k$ de F. Montrons que f est linéaire. Soient $x, y \in E$ et $\lambda \in \mathbb{K}$. Si $(x_k)_{1\leq k\leq n}$ et $(y_k)_{1\leq k\leq n}$ sont les coordonnées respectives de x et y dans $(e_1, e_2, \cdots, e_n)$, les coordonnées de $\lambda x+y$ sont $(\lambda x_k+y_k)_{1\leq k\leq n}$. Du coup : $f(\lambda x+y) = \sum_{k=1}^{n}(\lambda x_k+y_k)f_k = \lambda\sum_{k=1}^{n} x_k f_k + \sum_{k=1}^{n} y_k f_k = \lambda f(x) + f(y)$, donc f est bien linéaire. Enfin, soit $k \in \{1, \cdots, n\}$. Alors les coordonnées de e_k sont la famille $(0, \cdots, 0, 1, 0, .., 0)$ dans laquelle le 1 est placé en $k^{ème}$ position, donc aussitôt $f(e_k) = f_k$.

□

3.2.7 Espace vectoriel d'applications linéaires

Théorème 3.2.35 (Dimension d'un espace vectoriel d'applications linéaires)**.** *Soient E et F deux $\mathbb{K}$-espaces vectoriels de dimension finie. Alors $\mathcal{L}(E, F)$ est de dimension finie et :* $\dim \mathcal{L}(E, F) = \dim E \times \dim F$.

Démonstration.

— Si : $\dim E = 0$, la seule application linéaire de E dans F est l'application nulle qui envoie 0_E sur 0_F, donc : $\mathcal{L}(E, F) = \{0_{\mathcal{L}(E,F)}\}$ et $\dim \mathcal{L}(E, F) = 0 = \dim E \times \dim F$.

— Si : $\dim E \neq 0$, nous pouvons nous donner une base $(e_1, \cdots, e_n)$ de E. Il n'est pas trop dur de vérifier que l'application $u \overset{\varphi}{\mapsto} \Big(u(e_1), \cdots, u(e_n)\Big)$ est linéaire de $\mathcal{L}(E, F)$ dans F^n, et d'après le théorème précédent :

$$\forall (f_1, \cdots, f_n) \in F^n, \exists! u \in \mathcal{L}(E, F), \varphi(u) = (f_1, \cdots, f_n).$$

Conclusion : linéaire bijective, φ est un isomorphisme de $\mathcal{L}(E, F)$ sur F^n. Or F est de dimension finie, donc F^n aussi par produit, puis $\mathcal{L}(E, F)$ par isomorphisme. Finalement :

$$\dim \mathcal{L}(E, F) = \dim F^n = n \times \dim F = \dim E \times \dim F.$$

□

CHAPITRE 4

Matrices et systèmes linéaires

Sommaire

4.1 Systèmes linéaires

4.1.1 Définitions et vocabulaire

Définition 4.1.1 (Équation linéaire)**.** On appelle équation linéaire à p inconnues une équation de la forme $a_1x_1 + a_2x_2 + \cdots + a_px_p = b$ où $a_1, a_2, \cdots, a_p$, b sont $p+1$ éléments de $\mathbb{K}$ donnés.

— $a_1, \cdots, a_p$ sont appelés coefficients de l'équation, b est appelé second membre de l'équation ; $x_1, \cdots, x_p$ sont les inconnues de l'équation.
— Tout p-uplet $(x_1, \cdots, x_p) \in \mathbb{K}^p$ vérifiant cette équation est appelé solution de l'équation.

Définition 4.1.2 (Système d'équations linéaires)**.** On appelle système d'équations linéaires à n équations et à p inconnues la donnée de n équations de la forme :

$$(\Sigma) \begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1p}x_p = b_1 \\ \vdots \\ a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{ip}x_p = b_i \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{np}x_p = b_n \end{cases}$$

— Les scalaires $a_{ij} (i \in [\![1, n]\!], j \in [\![1, p]\!])$ sont les coefficients du système.
— Les scalaires $b_1, \cdots, b_n$ forment le second membre du système.
— Les scalaires $x_1, \cdots, x_p$ sont les inconnues que l'on cherche à déterminer.
— On note généralement L_i la i^e équation.

▶ Une solution de (Σ) est un p-uplet $(x_1, \cdots, x_p) \in \mathbb{K}^p$ vérifiant les équation $L_1, \cdots, L_n$.

Vocabulaire :

— Un système qui admet au moins une solution est dit compatible. S'il ne l'est pas, il est dit incompatible.
— Le système est dit homogène (ou sans second membre) si $b_1 = b_2 = \cdots = b_n = 0$.
On appelle système homogène associé à (Σ) le système (Σ) dans lequel on remplace les b_i par 0.
Un système homogène admet toujours $(0, \cdots, 0)$ comme solution.
— Un système est dit carré si $n = p$, c'est-à-dire s'il admet autant d'équations que d'inconnues.
— Un système carré est dit triangulaire supérieur si $a_{ij} = 0$ pour $i > j$.

4.1.2 Opérations élémentaires sur les lignes

Définition 4.1.3 (Opérations élémentaires sur les lignes)**.** On appelle opération élémentaire (de base) sur les lignes d'un système linéaire (Σ) de la forme :

$$(\Sigma) \begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1p}x_p = b_1 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{np}x_p = b_n \end{cases}$$

les opérations suivantes :

— Échange de la ligne L_i et de la ligne L_j, noté $L_i \leftrightarrow L_j$.
— Multiplication d'une ligne par un scalaire non nul, notée $L_i \leftarrow \lambda L_j$.
— Ajout d'une ligne à une autre ligne multipliée par un scalaire, noté $L_i \leftarrow L_i + \lambda L_j$.
Cette dernière notation se lit "L_i reçoit $L_i + \lambda L_j$".

Remarque 4.1.4. Les opérations élémentaires préservent les équivalences, donc ne modifient pas l'ensemble des solutions d'un système linéaire, car on peut

toujours les défaire. L'opération : $L_i \leftrightarrow L_j$ se défait elle-même par exemple, l'opération : $L_i \leftarrow \lambda L_i$ est défaite par l'opération : $L_i \leftarrow \frac{1}{\lambda} L_i$ et l'opération : $L_i \leftarrow L_i + \lambda L_j$ par l'opération : $L_i \leftarrow L_i - \lambda L_j$.

► Transformer un système linéaire donné par des opérations élémentaires **SUR LES LIGNES** ne modifie pas l'ensemble de ses solutions.
► On peut définir d'autres opérations élémentaires comme composées des opérations élémentaires de base : changer l'ordre des lignes, additionner des lignes, ajouter à une ligne une combinaison linéaire des autres lignes, etc.

Définition 4.1.5. Deux systèmes (Σ) et (Σ') sont dits équivalents si l'un se déduit de l'autre par une succession d'opérations élémentaires.

► Les opérations élémentaires étant inversibles, on peut bien passer d'un système à l'autre et vice-versa.

Théorème 4.1.6. *Deux systèmes équivalents ont même ensemble de solutions.*

► Si deux systèmes (Σ) et (Σ') sont équivalents, on pourra écrire :

$$(\Sigma) \Leftrightarrow (\Sigma').$$

4.1.3 La méthode de Gauss

Dans cette section nous expliquons la méthode du pivot (ou méthode d'élimination de Gauss) qui fournit un algorithme simple et pratique pour résoudre les systèmes d'équations linéaires. Pour cela, nous aurons besoin de la définition suivante :

Définition 4.1.7 (systèmes échelonnés)**.** Un système est dit échelonné si les lignes commencent par un nombre de zéros strictement croissant à mesure que l'indice augmente (c'est-à-dire, par exemple, la ligne L_3 commence par un nombre de zéros strictement plus grand que la ligne L_2 et celle-ci par un nombre de zéros strictement plus grand que la ligne L_1).

Idée fondamentale : L'idée essentielle est de résoudre un système d'équations linéaires en se ramenant à un système échelonné équivalent par une succession d'opérations élémentaires.

Quelques exemples suffiront pour comprendre comment exploiter cette idée, avant d'expliquer d'une manière plus précise la technique.

Étape 1 On sélectionne dans la première colonne un coefficient non nul appelé pivot et on échange la première ligne et la ligne correspondante :

$$(\Sigma) \begin{cases} 3y + 2z + t = 1 \\ 2x + 4y + 6z = 2 \\ x - y + 2z - t = 3 \\ 5x + y + 9z - 3t = 4 \end{cases} \overset{L_3 \leftrightarrow L_1}{\Longleftrightarrow} \begin{cases} x - y + 2z - t = 3 \\ 3y + 2z + t = 1 \\ 2x + 4y + 6z = 2 \\ 5x + y + 9z - 3t = 4 \end{cases}$$

Étape 2 À l'aide de ce premier pivot, on annule les coefficients de la première inconnue dans les autres équations :

$$\begin{cases} x - y + 2z - t = 3 \\ 3y + 2z + t = 1 \\ 2x + 4y + 6z = 2 \\ 5x + y + 9z - 3t = 4 \end{cases} \overset{\substack{L_3 \leftarrow L_3 - 2L_1 \\ L_4 \leftarrow L_4 - 5L_1}}{\Longleftrightarrow} \begin{cases} x - y + 2z - t = 3 \\ 3y + 2z + t = 1 \\ 6y + 2z + 2t = -4 \\ 6y - z + 2t = -9 \end{cases}$$

Étape 3 On fixe ensuite la première ligne et on réitère avec la seconde inconnue, puis avec la troisième, et ainsi de suite. On aboutit de proche en proche à un système échelonné.

$$\left\{\begin{array}{r} x-y+2z-t=3 \\ 3y+2z+t=1 \\ 6y+2z+2t=-4 \\ 6y-z+2t=-9 \end{array}\right. \overset{\substack{L_3\leftarrow L_3-2L_2 \\ L_4\leftarrow L_4-2L_2}}{\Longleftrightarrow} \left\{\begin{array}{r} x-y+2z-t=3 \\ 3y+2z+t=1 \\ -2z=-6 \\ -5z=-11 \end{array}\right.$$

$$\overset{L_4\leftarrow L_4-\frac{5}{2}L_3}{\Longleftrightarrow} \left\{\begin{array}{r} x-y+2z-t=3 \\ 3y+2z+t=1 \\ -2z=-6 \\ 0=4 \end{array}\right.$$

Conclusion Le système est donc incompatible. L'échelonnement du système a été mené à terme mais on aurait pu s'arrêter à l'étape précédente car les deux dernières équations du système précédent donnent $z=3$ et $z=\frac{11}{5}$. (Σ) n'admet donc pas de solution.

En-voici un autre exemple :

$$\left\{\begin{array}{r} x+2y-z=1 \\ 2x+3y+z=2 \\ x+4y-6z=2 \end{array}\right. \overset{\substack{L_2\leftarrow L_2-2L_1 \\ L_3\leftarrow L_3-L_1}}{\Longleftrightarrow} \left\{\begin{array}{r} x+2y-z=1 \\ -y+3z=0 \\ 2y-5z=1 \end{array}\right.$$

$$\overset{L_3\leftarrow L_3+2L_2}{\Longleftrightarrow} \left\{\begin{array}{r} x+2y-z=1 \\ -y+3z=0 \\ z=1 \end{array}\right.$$

On a donc $z=1$. En reportant dans l'équation L_2 on obtient la valeur de y : $y=3$. En remontant maintenant dans l'équation L_1 on trouve x : $x=-2y+z+1=-6+1+1=4$.
Le système admet donc la solution unique $(x,y,z)=(-4,3,1)$.

▶ Comme on le voit, la méthode consiste à mettre le système "sous forme échelonnée" de manière à pouvoir, en partant de la solution de la dernière équation et en remontant, résoudre toutes les équations.

Un dernier exemple :

$$\left\{\begin{array}{r} x-3y+4z-2w=5 \\ x-y+9z-w=7 \\ x-2y+7z-2w=9 \end{array}\right. \overset{\substack{L_2\leftarrow L_2-L_1 \\ L_3\leftarrow L_3-L_1}}{\Longleftrightarrow} \left\{\begin{array}{r} x-3y+4z-2w=5 \\ 2y+5z+w=2 \\ y+3z=4 \end{array}\right.$$

$$\overset{L_3\leftarrow 2L_3-L_2}{\Longleftrightarrow} \left\{\begin{array}{r} x-3y+4z-2w=5 \\ 2y+5z+w=2 \\ z-w=6 \end{array}\right.$$

Le système est sous forme échelonnée. En donnant à w une valeur arbitraire λ, on trouve : $w=\lambda, z=6+\lambda, y=-14-3\lambda, x=-61-11\lambda$.
Le système admet donc une infinité à un paramètre de solutions.

▷ Les exemples que nous venons de traiter illustrent suffisamment la méthode pour résoudre un système d'équations linéaires.

Remarque 4.1.8. Trois cas de figure se présentent :

— Si tous les coefficients d'une ligne sont nuls et que le second membre b_i correspondant est non nul, le système n'admet pas de solution, il est incompatible.

— Si lors de la remontée, il reste des inconnues dont la valeur n'est pas imposée, le système admet une infinité de solutions.
— Sinon, le système admet une unique solution.

On peut ainsi montrer que :

Théorème 4.1.9. *1. Tout système linéaire est équivalent à un système échelonné.*

2. Tout système linéaire admet 0, 1 *ou une infinité de solutions.*

► On peut démontrer que quelles que soient les opérations effectuées, le nombre de pivots dans le système échelonné obtenu est constant. Ce théorème fondamental dont on admet la démonstration conduit à la définition suivante :

Définition 4.1.10 (Rang d'un système)**.** On appelle rang d'un système le nombre de pivots obtenus après échelonnement.

Exemple 4.1.11. Le rang du système

$$(\Sigma)\ \begin{cases} 3y+2z+t=1 \\ 2x+4y+6z=2 \\ x-y+2z-t=3 \\ 5x+y+9z-3t=4 \end{cases}$$

vaut 3 car d'après les calculs précédents, il est équivalent au système échelonné :

$$\begin{cases} 1x-y+2z-t=3 \\ 3y+2z+t=1 \\ 1z=3 \\ 0=4. \end{cases} \tag{4.1}$$

Un système carré est dit de **Cramer** s'il admet une unique solution.

Théorème 4.1.12. *Un système carré à n équations et à n inconnues est de Cramer si et seulement si son rang est égal à n.*

Exercice 4.1.13. On considère le système suivant :

$$(\mathcal{S})\ \begin{cases} x+y+mz=m \\ x+my-z=1 \\ x+y-z=1 \end{cases} \quad (m\in\mathbb{R})$$

Préciser pour quelle(s) valeur(s) du réel m le système précédent est de Cramer. Déterminer alors son unique solution en fonction de m. On pourra utiliser les formules de Cramer.

Solution. Procédons tout d'abord sans utiliser les formules de Cramer.

$$\begin{cases} x+y+mz=m \\ x+my-z=1 \\ x+y-z=1 \end{cases} \overset{\substack{L_2\leftarrow L_2-L_1 \\ L_3\leftarrow L_3-L_1}}{\Longleftrightarrow} \begin{cases} x+y+mz=m \\ (m-1)y-(m+1)z=1-m \\ -(m+1)z=1-m \end{cases}$$

Le dernier système étant échelonné, on peut facilement discuter son nombre de solutions.

⋆ Si $m = -1$, la dernière ligne montre que $(\mathcal{S})$ n'admet pas de solution.

⋆ Si $m \neq -1$, on a $z = \frac{m-1}{m+1}$ et $(m-1)y = 0$. Il faut donc distinguer deux cas.
- Si $m \neq 1, y = 0$ puis $x = m(1-z) = \frac{2m}{m+1}$. Le système est de Cramer et admet pour unique solution le triplet $\left(\frac{2m}{m+1}, 0, \frac{m-1}{m+1}\right)$
- Si $m = 1$, on trouve $x = 1 - y$ et $z = 0$. Le système admet une infinité de solutions de la forme $(1-y, y, 0)$.

Revenons à la méthode avec les formules de Cramer. Le système $(\mathcal{S})$ est de Cramer si et seulement si sont déterminant est non nul,

$$\begin{vmatrix} 1 & 1 & m \\ 1 & m & -1 \\ 1 & 1 & -1 \end{vmatrix} \neq 0 \Longleftrightarrow 1 - m^2 \neq 0 \Longleftrightarrow m \neq \pm 1.$$

Pour $m \neq \pm 1$, on peut donc utiliser les formules de Cramer :

$$x = \frac{\begin{vmatrix} m & 1 & m \\ 1 & m & -1 \\ 1 & 1 & -1 \end{vmatrix}}{1-m^2} = \frac{2m(1-m)}{1-m^2} = \frac{2m}{1+m},$$

$$y = \frac{\begin{vmatrix} 1 & m & m \\ 1 & 1 & -1 \\ 1 & 1 & -1 \end{vmatrix}}{1-m^2} = 0 \quad \text{et} \quad z = \frac{\begin{vmatrix} 1 & 1 & m \\ 1 & m & 1 \\ 1 & 1 & 1 \end{vmatrix}}{1-m^2} = \frac{-(m-1)^2}{1-m^2} = \frac{m-1}{1+m}.$$

4.2 Matrices

4.2.1 Opérations sur les matrices

Définition 4.2.1 (Matrice, coefficients, lignes, colonnes)**.** — On appelle matrice de taille $n \times p$ à coefficients dans $\mathbb{K}$ toute famille A de np éléments de $\mathbb{K}$ présentée sous la forme d'un tableau : $\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{pmatrix}$,
noté aussi : $(a_{ij})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant p}}$, où $a_{ij} \in \mathbb{K}$ pour tout $(i,j) \in [\![1,n]\!] \times [\![1,p]\!]$.

— Lensemble des matrices de taille $n \times p$ à coefficients dans $\mathbb{K}$ est noté $\mathcal{M}_{n,p}(\mathbb{K})$.
- Pour $n = p$, on parle de matrices carrées de taille n et la notation simplifiée $\mathcal{M}_n(\mathbb{K})$ est alors préférée. La famille $(a_{11}, a_{22}, \ldots, a_{nn})$ est alors appelée diagonale de A.
- Pour $p = 1$, on parle de matrices colonnes de taille n, et pour $n = 1$, de matrices lignes de taille p.

Remarque 4.2.2. À vrai dire, une matrice M de taille $n \times p$ à coefficients dans $\mathbb{K}$ n'est jamais qu'un élément de $\mathbb{K}^{[\![1,n]\!] \times [\![1,p]\!]}$, i.e. une famille $(m_{ij})_{(i,j) \in [\![1,n]\!] \times [\![1,p]\!]}$ d'éléments de $\mathbb{K}$ indexée par $[\![1,n]\!] \times [\![1,p]\!]$, c'est-à-dire encore une application $(i,j) \longmapsto m_{ij}$ de $[\![1,n]\!] \times [\![1,p]\!]$ dans $\mathbb{K}$.
En résumé : $\mathcal{M}_{n,p}(\mathbb{K}) = \mathbb{K}^{[\![1,n]\!] \times [\![1,p]\!]}$

Définition 4.2.3 (Matrice nulle)**.** La matrice de taille $n \times p$ dont tous les coefficients sont nuls est appelée la matrice nulle de $\mathcal{M}_{n,p}(\mathbb{K})$ et notée 0 ou $0_{n,p}$ quand on veut être précis.

Définition 4.2.4 (Addition matricielle et multiplication par un scalaire)**.** Pour tous $A, B \in \mathcal{M}_{n,p}(\mathbb{K})$ et $\lambda, \mu \in \mathbb{K}$, on note $\lambda A + \mu B$ la matrice :

$$\begin{pmatrix} \lambda a_{11} + \mu b_{11} & \cdots & \lambda a_{1p} + \mu b_{1p} \\ \vdots & & \vdots \\ \lambda a_{n1} + \mu b_{n1} & \cdots & \lambda a_{np} + \mu b_{np} \end{pmatrix},$$

appelée une combinaison linéaire de A et B.

Exercice 4.2.5. Montrer que

$$E = \{M(a,b) = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \mid (a,b) \in \mathbb{R}^2\}$$

est une $\mathbb{R}$-espace vectoriel et en déterminer une base et la dimension.

Solution. On a : $E = \{a \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}_{\text{notée } I} + b \underbrace{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}}_{\text{notée } J} \mid a, b \in \mathbb{R}\}$ et donc $E = \text{Vect}(I, J)$. De plus :

$$\alpha I + \beta J = 0 \Longleftrightarrow \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \Longleftrightarrow \alpha = \beta = 0.$$

Donc (I, J) est libre.
On conclut que : E est un $\mathbb{R} - \text{ev}$, (I, J) est une base de E et $\dim(E) = 2$.

4.2.2 Produit matriciel

Définition 4.2.6 (Produit matriciel)**.** Pour tous $A \in \mathcal{M}_{p,q}(\mathbb{K})$ et $B \in \mathcal{M}_{q,r}(\mathbb{K})$, on note $A \times B$ la matrice $\left(\sum_{k=1}^{q} a_{ik} b_{kj}\right)_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant r}}$ de taille $p \times r$.

Attention :

— Le produit de deux matrices n'est pas défini en général s'il n'y a pas, comme on dit, compatibilité des formats :

$$\overbrace{p \times q}^{\text{Matrice de taille}} \quad \times \quad \overbrace{q \times r}^{\text{Matrice de taille}} \quad = \quad \overbrace{p \times r}^{\text{Matrice de taille}}$$

— Le produit matriciel n'est pas commutatif ! Par exemple :

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

— Un produit de matrices peut être nul sans qu'aucune d'entre elles le soit. Exemple :

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Théorème 4.2.7 (Propriétés du produit matriciel, matrice identité)**.** *On a les propriétés suivantes :*

- ***Associativité :*** *Pour tous $A \in \mathcal{M}_{p,q}(\mathbb{K})$, $B \in \mathcal{M}_{q,r}(\mathbb{K})$, $C \in \mathcal{M}_{r,s}(\mathbb{K})$:* $(AB)C = A(BC)$.
- ***Bilinéarité :*** *Pour tous $A, B \in \mathcal{M}_{p,q}(\mathbb{K})$, $C, D \in \mathcal{M}_{q,r}(\mathbb{K})$ et $\lambda, \mu \in \mathbb{K}$:*

$$(\lambda A + \mu B)C = \lambda AC + \mu BC \quad et \quad B(\lambda C + \mu D) = \lambda BC + \mu BD.$$

- ***Élément neutre :*** *On appelle matrice identité (de taille n) la matrice carrée de taille n :*

$$I_n = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}$$

Pour toute matrice $A \in \mathcal{M}_{n,p}(\mathbb{K})$: $I_n A = A I_p = A$.

Exercice 4.2.8. On note $A = \begin{pmatrix} 8 & 0 \\ 0 & -1 \end{pmatrix} \in \mathcal{M}_2(\mathbb{R})$.
Résoudre l'équation : $M^3 = A$, d'inconnue $M \in \mathcal{M}_2(\mathbb{R})$, en remarquant que, si $M^3 = A$, alors $AM = MA$.

Solution. Raisonnons par analyse-synthèse. Soit $M \in \mathcal{M}_2(\mathbb{R})$ telle que $M^3 = A$. On a alors : $AM = M^3M = M^4 = MM^3 = MA$. En notant $M = \begin{pmatrix} x & y \\ z & t \end{pmatrix}, (x, y, z, t) \in \mathbb{R}^4$, on a :

$$AM = MA \Leftrightarrow \begin{pmatrix} 8 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} x & y \\ z & t \end{pmatrix} = \begin{pmatrix} x & y \\ z & t \end{pmatrix}\begin{pmatrix} 8 & 0 \\ 0 & -1 \end{pmatrix}$$
$$\Leftrightarrow \begin{pmatrix} 8x & 8y \\ -z & -t \end{pmatrix} = \begin{pmatrix} 8x & -y \\ 8z & -t \end{pmatrix} \Leftrightarrow y = z = 0.$$

On a donc : $M = \begin{pmatrix} x & 0 \\ 0 & t \end{pmatrix}$. Réciproquement, on a :

$$M^3 = A \Leftrightarrow \begin{pmatrix} x^3 & 0 \\ 0 & t^3 \end{pmatrix} = \begin{pmatrix} 8 & 0 \\ 0 & -1 \end{pmatrix} \Leftrightarrow \begin{cases} x^3 = 8 \\ t^3 = -1 \end{cases} \Leftrightarrow \begin{cases} x = 2 \\ t = -1. \end{cases}$$

On conclut : $\mathcal{S} = \left\{ \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix} \right\}$.

► À présent, pour une simple raison de compatibilité des formats, une matrice ne peut être multipliée avec elle-même que si elle est carrée. Pour une telle matrice $A \in \mathcal{M}_n(\mathbb{K})$, on peut ainsi parler des puissances de A :

$$A^k = \overbrace{A \times A \times \cdots \times A}^{k \text{ fois}} \quad \text{pour tout } k \in \mathbb{N}^* \quad \text{et} \quad A^0 = I_n.$$

Définition 4.2.9 (Matrice nilpotente)**.** Soit $A \in \mathcal{M}_n(\mathbb{K})$. On dit que A est nilpotente si pour un certain $p \in \mathbb{N}^*$: $A^p = 0$.
Le plus petit entier p pour laquelle cette identité est vraie est appelé l'indice de nilpotence de A.

Théorème 4.2.10 (Deux formules à connaître)**.** *Soient $A, B \in \mathcal{M}_n(\mathbb{K})$. On suppose que A et B* ***COMMUTENT****, i.e. que : $AB = BA$. Alors :*
$(A+B)^k = \sum_{i=0}^{k} C_k^i A^i B^{k-i}$ *formule du binôme*
et $A^k - B^k = (A-B)\sum_{i=0}^{k-1} A^i B^{k-i-1}$

Attention : L'hypothèse selon laquelle A et B commutent est essentielle, c'est déjà très clair pour $k = 2$:

$$(A+B)^2 = (A+B)(A+B) = A^2 + AB + BA + B^2 \overset{AB=BA}{=} A^2 + 2AB + B^2$$

et $(A-B)(A+B) = A^2 + AB - BA - B^2 \overset{AB=BA}{=} A^2 - B^2$.

Exemple 4.2.11. On pose : $A = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Pour tout $k \in \mathbb{N}$: $A^k = \begin{pmatrix} 1 & k & 2k \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

▶ En effet, écrivons : $A = I_3 + N$ avec : $N = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. Les matrices I_3 et N commutent car I_3 commute avec toute matrice carrée de taille 3. En outre, N est nilpotente car : $N^2 = 0$, donc : $N^i = 0$ pour tout $i \geqslant 2$. Finalement, pour tout $k \in \mathbb{N}$: $A^k = \sum_{i=0}^{k} C_k^i I_3^{k-i} N^i = I_3 + kN = \begin{pmatrix} 1 & k & 2k \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

Exercice 4.2.12.

4.2.3 Transposition

Définition 4.2.13 (Transposée)**.** Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$. On appelle transposée de A la matrice $(a_{ji})_{\substack{1 \leqslant j \leqslant p \\ 1 \leqslant i \leqslant n}}$ de $\mathcal{M}_{p,n}(\mathbb{K})$, notée A^T ou ${}^tA$.

Exemple 4.2.14. On a :

$$\begin{pmatrix} 3 & 0 & 1 \\ 5 & 2 & 7 \end{pmatrix}^\top = \begin{pmatrix} 3 & 5 \\ 0 & 2 \\ 1 & 7 \end{pmatrix}$$

et pour tous $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$: $\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}^\top = \begin{pmatrix} \lambda_1 & \cdots & \lambda_n \end{pmatrix}$

Théorème 4.2.15 (Propriétés de la transposition)**.**
- — ***Linéarité :*** *Pour tous $A, B \in \mathcal{M}_{n,p}(\mathbb{K})$ et $\lambda, \mu \in \mathbb{K}$: $(\lambda A + \mu B)^T = \lambda A^T + \mu B^T$.*
- — ***Involutivité :*** *Pour tous $A \in \mathcal{M}_{n,p}(\mathbb{K})$: $(A^T)^T = A$.*
- — ***Effet sur un produit :*** *Pour tous $A \in \mathcal{M}_{p,q}(\mathbb{K})$ et $B \in \mathcal{M}_{q,r}(\mathbb{K})$: $(AB)^T = B^T A^T$.*

Démonstration. Seule l'assertion sur le produit mérite une preuve. Pour tout $(i,j) \in [\![1,r]\!] \times [\![1,p]\!]$, on a : $\left((AB)^\top\right)_{ij} = (AB)_{ji} = \sum_{k=1}^{q} a_{jk} b_{ki} = \sum_{k=1}^{q} b_{ki} a_{jk} = \sum_{k=1}^{q} \left(B^\top\right)_{ik} \left(A^\top\right)_{kj} = \left(B^\top A^\top\right)_{ij}$. □

Définition 4.2.16 (Matrice symétrique/antisymétrique)**.** Soit $A \in \mathcal{M}_n(\mathbb{K})$. On dit que A est symétrique si : $A^T = A$, et antisymétrique si : $A^T = -A$.

— Une matrice symétrique est donc de la forme :

$$S = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{12} & a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1n} \\ a_{1n} & \dots & a_{n-1n} & a_{nn} \end{pmatrix}$$

On note $\mathrm{S}_n(\mathbb{K})$ l'ensemble des matrices symétriques d'ordre n à coefficients dans $\mathbb{K}$. Il s'agit d'un espace vectoriel.

— Une matrice antisymétrique est de la forme :

$$A = \begin{pmatrix} 0 & a_{12} & \dots & a_{1n} \\ -a_{12} & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1n} \\ -a_{1n} & \dots & -a_{n-1n} & 0 \end{pmatrix}$$

On note $\mathrm{A}_n(\mathbb{K})$ l'ensemble des matrices antisymétriques d'ordre n à coefficients dans $\mathbb{K}$. Il s'agit d'un espace vectoriel.

Exercice 4.2.17. Montrer que l'ensemble $\mathrm{S}_n(\mathbb{K})$ des matrices symétriques et l'ensemble $\mathrm{A}_n(\mathbb{K})$ des matrices antisymétriques de $\mathcal{M}_n(\mathbb{K})$ sont supplémentaires dans $\mathcal{M}_n(\mathbb{K})$.

Solution. Il est d'abord clair que ces deux ensembles sont des sous-espaces vectoriels. Soit $M \in \mathcal{M}_n(\mathbb{K})$. Nous voulons montrer ceci :

$$\exists!(S, A) \in \mathrm{S}_n(\mathbb{K}) \times \mathrm{A}_n(\mathbb{K}), \quad M = S + A.$$

Pour cela nous raisonnons par analyse-synthèse.

— **Analyse :** Soient $S \in \mathrm{S}_n(\mathbb{K})$ et $A \in \mathrm{A}_n(\mathbb{K})$. On suppose que : $M = S + A$. Alors : $M^\top = S^\top + A^\top = S - A$, donc par demi-somme et demi-différence : $S = \frac{M+M^\top}{2}$ et $A = \frac{M-M^\top}{2}$.

— **Synthèse :** Posons : $S = \frac{M+M^\top}{2}$ et $A = \frac{M-M^\top}{2}$. Alors : $M = S+A$, et S est symétrique et A antisymétrique car : $S^\top = \left(\frac{M+M^\top}{2}\right)^\top = \frac{M^\top+M}{2} = S$ et $A^\top = \left(\frac{M-M^\top}{2}\right)^\top = \frac{M^\top-M}{2} = -A$.

4.2.4 Matrices diagonales et triangulaires

Définition 4.2.18 (Matrice diagonale, matrice scalaire, matrice triangulaire)**.** —

Une matrice carrée est dite diagonale si tous ses coefficients non diagonaux sont nuls.

En particulier, les matrices λI_n, λ décrivant $\mathbb{K}$, sont appelées matrices scalaires.

— Une matrice carrée est dite triangulaire supérieure (resp. inférieure) si ses coefficients situés strictement au-dessous (resp. strictement au-dessus) de la diagonale sont nuls.

Explication :

Matrice triangulaire supérieure
$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & \cdots & a_{2n} \\ & & \ddots & \vdots \\ & & & a_{nn} \end{pmatrix}$$

$$\begin{pmatrix} a_{11} & & & \\ a_{21} & a_{22} & & \\ \vdots & \vdots & \ddots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$ Matrice triangulaire inférieure

Remarque 4.2.19. Les matrices diagonales sont exactement les matrices **À LA FOIS** triangulaires supérieures **ET** triangulaires inférieures.

Notation : Pour tous $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$, on note généralement $\operatorname{diag}(\alpha_1, \ldots, \alpha_n)$ la matrice diagonale : $\begin{pmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_n \end{pmatrix}$. Par exemple : $I_n = \operatorname{diag}(1, \ldots, 1)$.
Avec cette notation, pour tous $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n, \lambda \in \mathbb{K}$, on a :

$$\lambda \operatorname{diag}(\alpha_1, \ldots, \alpha_n) + \operatorname{diag}(\beta_1, \ldots, \beta_n) = \operatorname{diag}(\lambda\alpha_1 + \beta_1, \ldots, \lambda\alpha_n + \beta_n)$$
$$\text{et } \operatorname{diag}(\alpha_1, \ldots, \alpha_n) \times \operatorname{diag}(\beta_1, \ldots, \beta_n) = \operatorname{diag}(\alpha_1\beta_1, \ldots, \alpha_n\beta_n).$$

Théorème 4.2.20 (Combinaisons linéaires et produit de matrices triangulaires)**.** *Soient $A, B \in \mathcal{M}_n(\mathbb{K})$ triangulaires supérieures (resp. inférieures) et $\lambda, \mu \in \mathbb{K}$. Les matrices $\lambda A + \mu B$ et AB sont alors triangulaires supérieures (resp. inférieures). En outre, pour tout $i \in [\![1, n]\!]$, ona : $(AB)_{ii} = a_{ii}b_{ii}$.*

Démonstration. Supposons A et B triangulaires supérieures (le cas inférieur s'en déduit par transposition). Les matrices $\lambda A + \mu B$ et AB sont triangulaires supérieures car pour tous $i, j \in [\![1, n]\!]$, si : $i > j$, alors :

$$(\lambda A + \mu B)_{ij} = \lambda \underbrace{a_{ij}}_{=0} + \mu \underbrace{b_{ij}}_{=0} = 0 \quad \text{et}$$

$$(AB)_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} = \sum_{k=1}^{i-1} \underbrace{a_{ik}}_{=0} b_{kj} + \sum_{k=i}^{n} a_{ik} \underbrace{b_{kj}}_{=0} = 0.$$

Enfin, pour tout $i \in [\![1, n]\!]$, d'après le même calcul : $(AB)_{ii} = \sum_{k=1}^{i-1} \underbrace{a_{ik}}_{=0} b_{ki} + a_{ii}b_{ii} + \sum_{k=i+1}^{n} a_{ik} \underbrace{b_{ki}}_{=0} = a_{ii}b_{ii}$. □

4.2.5 Trace d'une matrice carrée

Définition 4.2.21 (Trace d'une matrice carrée)**.** Soit $A \in \mathcal{M}_n(\mathbb{K})$. On appelle trace de A et on note $\operatorname{tr}(A)$ ou $\operatorname{Tr}(A)$ la somme des éléments diagonaux de A.

Théorème 4.2.22. *1. **Linéarité :** Pour tous $A, B \in \mathcal{M}_n(\mathbb{K})$ et $\lambda, \mu \in \mathbb{K}$, on a : $\operatorname{tr}(\lambda A + \mu B) = \lambda \operatorname{Tr}(A) + \mu \operatorname{Tr}(B)$.*

2. ***Effet sur un produit :*** *Pour tous* $A \in \mathcal{M}_{np}(\mathbb{K})$ *et* $B \in \mathcal{M}_{pn}(\mathbb{K})$ *:* $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.

Démonstration. 1.

$$\begin{aligned}\operatorname{tr}(\lambda A + \mu B) &= \sum_{k=1}^{n} (\lambda a_{kk} + \mu b_{kk}) = \lambda \sum_{k=1}^{n} a_{kk} + \mu \sum_{k=1}^{n} b_{kk} \\ &= \lambda \operatorname{tr}(A) + \mu \operatorname{tr}(B).\end{aligned}$$

2. La matrice AB est carrée de taille n alors que BA est carrée de taille p, mais ces matrices ont même trace car : $\operatorname{tr}(AB) = \sum_{k=1}^{n}(AB)_{kk} = \sum_{k=1}^{n}\sum_{\ell=1}^{p} a_{k\ell}b_{\ell k} = \sum_{\ell=1}^{p}\sum_{k=1}^{n} b_{\ell k}a_{k\ell} = \sum_{\ell=1}^{p}(BA)_{\ell\ell} = \operatorname{tr}(BA)$.

□

Exemple 4.2.23. L'équation matricielle : $AB - BA = I_n$ d'inconnue $(A, B) \in \mathcal{M}_n(\mathbb{K})^2$ n'a pas de solution car pour toutes matrices $A, B \in \mathcal{M}_n(\mathbb{K})$: $\operatorname{tr}(AB - BA) = \operatorname{tr}(AB) - \operatorname{tr}(BA) = 0 \neq n = \operatorname{tr}(I_n)$.

4.2.6 Matrice inversible

Définition 4.2.24 (Matrice inversible, inverse, groupe linéaire)**.** Soit $A \in \mathcal{M}_n(\mathbb{K})$. On dit que A est inversible s'il existe une matrice $B \in \mathcal{M}_n(\mathbb{K})$ pour laquelle :

$$AB = BA = I_n.$$

SI ELLE EXISTE, une telle matrice B est unique, notée A^{-1} et appelée l'inverse de A.
L'ensemble des matrices inversibles de $\mathcal{M}_n(\mathbb{K})$ est noté $GL_n(\mathbb{K})$ et appelé le groupe linéaire de degré n sur $\mathbb{K}$.

Démonstration. Si B et B' sont deux inverses de A : $B = BI_n = B(AB') = (BA)B' = I_nB' = B'$. □

Remarque 4.2.25. $GL_n(\mathbb{K})$ n'est pas un espace vectoriel ! En effet, la matrice nulle n'est pas inversible. Si c'était le cas, il existerait B telle que $0_n \times B = B \times 0_n = I_n$, absurde !

Remarque 4.2.26. Pourquoi oblige-t-on par définition les matrices inversibles à être carrées ? Qu'est-ce qui empêche une matrice $A \in \mathcal{M}_{n,p}(\mathbb{K})$ avec : $n \neq p$ de posséder un inverse, i.e. une matrice $B \in \mathcal{M}_{p,n}(\mathbb{K})$ pour laquelle : $AB = I_n$ et $BA = I_p$? Nous y reviendrons plus tard, mais une première réponse simple peut être donnée tout de suite : $n = \operatorname{tr}(I_n) = \operatorname{tr}(AB) = \operatorname{tr}(BA) = \operatorname{tr}(I_p) = p$.

Théorème 4.2.27 (Opérations sur les matrices inversibles)**.** *Soient* $A, B \in GL_n(\mathbb{K})$*, donc INVERSIBLES.*

1. *Inversibilité de l'inverse :* A^{-1} *est inversible et :* $(A^{-1})^{-1} = A$.
2. *Inversibilité d'un produit :* AB *est inversible et :* $(AB)^{-1} = B^{-1}A^{-1}$.
3. *Inversibilité d'une puissance : Pour tout* $k \in \mathbb{Z}$*,* A^k *est inversible et :* $(A^K)^{-1} = (A^{-1})^k$.
4. *Inversibilité de la transposée :* A^T *est inversible et :* $(A^T)^{-1} = (A^{-1})^T$.

Démonstration. 1. On a : $A \times A^{-1} = A^{-1} \times A = I_n$, donc par définition de l'inversibilité, A^{-1} est inversible d'inverse A.

2. On a : $AB \times B^{-1}A^{-1} = A \times BB^{-1} \times A^{-1} = A \times I_n \times A^{-1} = AA^{-1} = I_n$ et $B^{-1}A^{-1} \times AB = I_n$, donc AB est inversible d'inverse $B^{-1}A^{-1}$.
3. Par récurrence à partir de l'assertion (2).
4. On a : $A^\top \left(A^{-1}\right)^\top = \left(A^{-1}A\right)^\top = I_n^\top = I_n$ et $\left(A^{-1}\right)^\top A^\top = I_n$, donc $A^\top$ est inversible d'inverse $\left(A^{-1}\right)^\top$.

□

Remarque 4.2.28. Il suffit en fait que $AB = I_n$ pour avoir $BA = I_n$:

Théorème 4.2.29. *Si deux matrices $A, B \in \mathcal{M}_n(\mathbb{K})$ vérifient $AB = I_n$ alors $BA = I_n$.*

Démonstration. Nous aurons, pour cette démonstration, recours aux applications linéaires et à leurs propriétés générales.

Soient $A, B \in \mathcal{M}_n(\mathbb{K})$ vérifiant $AB = I_n$. On considère alors l'application φ définie sur l'espace vectoriel $\mathcal{M}_n(\mathbb{K})$ par $\varphi : M \longmapsto MA$.

— φ est une application linéaire car pour tout $M_1, M_2 \in \mathcal{M}_n(\mathbb{K})$ et $\lambda \in \mathbb{K}$,

$$\varphi\left(\lambda M_1 + M_2\right) = \left(\lambda M_1 + M_2\right) A = \lambda M_1 A + M_2 A = \lambda\varphi\left(M_1\right) + \varphi\left(M_2\right).$$

— $\varphi : \mathcal{M}_n(\mathbb{K}) \to \mathcal{M}_n(\mathbb{K})$ donc $\varphi \in \mathscr{L}\left(\mathcal{M}_n(\mathbb{K})\right)$.
— φ est injective car $\mathrm{Ker}(\varphi) = \{0_n\}$:

$$M \in \mathrm{Ker}(\varphi) \Longrightarrow MA = 0_n \Longrightarrow MAB = 0_n \Longrightarrow MI_n = 0_n \Longrightarrow M = 0_n.$$

— $\mathcal{M}_n(\mathbb{K})$ étant de dimension finie, l'endomorphisme φ est surjectif.
— Comme $I_n \in \mathrm{Im}(\varphi)$, il existe $C \in \mathcal{M}_n(\mathbb{K})$ telle que $\varphi(C) = I_n$, c'est-à-dire, telle que $CA = I_n$.
— Il ne reste plus qu'à montrer que $B = C$, ce qui est bien le cas car on peut multiplier l'égalité $CA = I_n$ à droite par B pour obtenir $CAB = C = B$. On a donc bien $BA = I_n$.

□

Théorème 4.2.30 (Une condition suffisante de non-inversibilité)**.** *Soit $A \in \mathcal{M}_n(\mathbb{K})$. Si l'une des colonnes (resp. lignes) de A est combinaison linéaire de ses autres colonnes (resp. lignes), alors A **N**'est **PAS** inversible.*

Démonstration. Notons $L_1, \cdots, L_n$ les lignes de A et faisons par exemple l'hypothèse, pour y voir plus clair, que L_1 est combinaison linéaire des lignes $L_2, \cdots, L_n$: $L_1 = \lambda_2 L_2 + \cdots + \lambda_n L_n$. Dans ces conditions :

$$\begin{pmatrix} 1 & -\lambda_2 & \cdots & -\lambda_n \end{pmatrix} A = L_1 - \lambda_2 L_2 - \cdots - \lambda_n L_n = 0 = \begin{pmatrix} 0 & \cdots & 0 \end{pmatrix}.$$

Or si A est inversible, on peut multiplier cette égalité par A^{-1} à droite : $\begin{pmatrix} 1 & -\lambda_2 & \cdots & -\lambda_n \end{pmatrix} = \begin{pmatrix} 0 & \cdots & 0 \end{pmatrix}$, mais il en découle que : $1 = 0$. Conclusion : A n'est pas inversible. □

Exemple 4.2.31. La matrice $\begin{pmatrix} 1 & 1 & 0 \\ 2 & 1 & 0 \\ 3 & 2 & 0 \end{pmatrix}$ possède une colonne nulle, donc n'est pas inversible. La matrice $\begin{pmatrix} 0 & 1 & 1 \\ 1 & 4 & 3 \\ 1 & 3 & 2 \end{pmatrix}$ n'est pas inversible car sa deuxième ligne est égale à la somme des deux autres.

Attention ! Dans $\mathbb{R}$ et $\mathbb{C}$, 0 est le seul objet par lequel la division est impossible. Dans $\mathcal{M}_n(\mathbb{K})$ au contraire, l'exemple qui précède montre qu'une matrice non nulle peut ne pas être inversible.

Théorème 4.2.32 (Caractérisations diverses de l'inversibilité)**.** *Soit $A \in \mathcal{M}_n(\mathbb{K})$. Les assertions suivantes sont équivalentes :*

1. *A est inversible.*
2. *A est inversible à droite : $\exists B \in \mathcal{M}_n(\mathbb{K}), AB = I_n$.*
3. *A est inversible à gauche : $\exists B \in \mathcal{M}_n(\mathbb{K}), BA = I_n$.*
4. *Pour tout second membre $Y \in \mathbb{K}^n$, le système linéaire : $Y = AX$ d'inconnue $X \in \mathbb{K}^n$ possède **AU MOINS UNE** solution.*
5. *Le système linéaire homogène : $AX = 0$ d'inconnue $X \in \mathbb{K}^n$ admet 0 pour **UNIQUE** solution :*

$$\forall X \in \mathbb{K}^n, AX = 0 \Longrightarrow X = 0.$$

Démonstration. Notons $C_1, \cdots, C_n$ les colonnes de A. La matrice A étant CARRÉE, la famille $(C_1, \cdots, C_n)$ est une base de $\mathbb{K}^n$ si et seulement si elle est libre (ou génératrice).

— Les implications (1) $\Longrightarrow$ (2) et (1) $\Rightarrow$ (3) sont triviales.

— Supposons (2) est vraie. Alors pour tout $Y \in \mathbb{K}^n$: $A(BY) = (AB)Y = I_nY = Y$, donc le système : $Y = AX$ d'inconnue $X \in \mathbb{K}^n$ possède au moins une solution. Ainsi : (2) $\Longrightarrow$ (4).

— Supposons (3) vraie. Pour tout $X \in \mathbb{K}^n$, si : $AX = 0$, alors : $X = I_nX = (BA)X = B(AX) = B \times 0 = 0$, donc 0 est la seule solution du système homogène : $AX = 0$ d'inconnue $X \in \mathbb{K}^n$. Ainsi : (3) $\Longrightarrow$ (5).

— Ensuite, on a :

$$\begin{array}{rl} (4) & \Longleftrightarrow \quad \forall Y \in \mathbb{K}^n, \quad \exists X \in \mathbb{K}^n, \quad Y = AX \\ & \Longleftrightarrow \forall Y \in \mathbb{K}^n, \quad \exists\, (x_1, \ldots, x_n) \in \mathbb{K}^n, \quad Y = \sum_{k=1}^n x_k C_k \\ & \Longleftrightarrow (C_1, \ldots, C_n) \text{ engendre } \mathbb{K}^n \\ & \Longleftrightarrow (C_1, \ldots, C_n) \text{ est une base de } \mathbb{K}^n \quad \Longleftrightarrow (1). \end{array}$$

— De même, on a :

$$\begin{array}{rl} (5) & \Longleftrightarrow \quad \forall\, (x_1, \ldots, x_n) \in \mathbb{K}^n, \\ & \qquad (\sum_{k=1}^n x_k C_k = 0 \quad \Longrightarrow \quad x_1 = \ldots = x_n = 0) \\ & \Longleftrightarrow \quad (C_1, \ldots, C_n) \text{ est libre} \\ & \Longleftrightarrow \quad (C_1, \ldots, C_n) \text{ est une base de } \mathbb{K}^n \quad \Longleftrightarrow (1). \end{array}$$

□

4.2.7 Matrices diagonales inversibles

Théorème 4.2.33. *Soient* $\alpha_1, \cdots, \alpha_n \in \mathbb{K}$. $\text{diag}(\alpha_1, \cdots, \alpha_n)$ *est inversible si et seulement si :* $\alpha_k \neq 0$ *pour tout* $k \in [\![1, n]\!]$, *et dans ce cas :* $\text{diag}(\alpha_1, \cdots, \alpha_n)^{-1} = \text{diag}(\alpha_1^{-1}, \cdots, \alpha_n^{-1})$.

Démonstration. Si $\alpha_1, \cdots, \alpha_n$ sont tous non nuls, alors :

$$\text{diag}(\alpha_1, \cdots, \alpha_n) \times \text{diag}(\alpha_1^{-1}, \cdots, \alpha_n^{-1}) = \text{diag}(1, \cdots, 1) = I_n$$

et

$$\text{diag}(\alpha_1^{-1}, \cdots, \alpha_n^{-1}) \times \text{diag}(\alpha_1, \cdots, \alpha_n) = I_n,$$

donc $\text{diag}(\alpha_1, \cdots, \alpha_n)$ est inversible d'inverse $\text{diag}(\alpha_1^{-1}, \cdots, \alpha_n^{-1})$. Réciproquement, si l'un des nombres $\alpha_1, \cdots, \alpha_n$ est nul, la matrice $\text{diag}(\alpha_1, \cdots, \alpha_n)$ n'est pas inversible car l'une de ses colonnes est nulle. □

4.2.8 Matrices triangulaires inversibles

Théorème 4.2.34. *Une matrice triangulaire* A *est inversible si et seulement si ses coefficients diagonaux sont tous non nuls. Dans ce cas,* A^{-1} *est elle aussi triangulaire de même type et ses coefficients diagonaux sont exactement les inverses des coefficients diagonaux de* A.

Démonstration. Laissé au lecteur. □

4.2.9 Rang d'une matrice

Définition 4.2.35 (Rang d'une matrice)**.** On appelle rang d'une matrice A le rang du système $AX = 0$ associé.

En pratique : Pour calculer le rang d'une matrice, on pourra donc mettre sous forme échelonnée le système correspondant puis compter le nombre de pivots obtenus non nuls. Noter que l'on peut pratiquer directement la méthode du pivot de Gauss sur la matrice étudiée.

Exemple 4.2.36. Déterminer le rang de $A = \begin{pmatrix} 1 & 4 & 3 \\ 2 & 5 & 3 \\ 3 & 6 & 3 \end{pmatrix}$.

▶ On effectue pour cela une succession d'opérations élémentaires sur les lignes :

$$\text{rg}\begin{pmatrix} 1 & 4 & 3 \\ 2 & 5 & 3 \\ 3 & 6 & 3 \end{pmatrix} \overset{\substack{L_2 \leftarrow L_2 - 2L_1 \\ L_3 \leftarrow L_3 - 3L_1}}{=} \text{rg}\begin{pmatrix} 1 & 4 & 3 \\ 0 & -3 & -3 \\ 0 & -6 & -6 \end{pmatrix}$$

$$\overset{L_3 \leftarrow L_3 - 2L_2}{=} \text{rg}\begin{pmatrix} 1 & 4 & 3 \\ 0 & -3 & -3 \\ 0 & 0 & 0 \end{pmatrix} = 2.$$

Exercice 4.2.37. Déterminer le rang des matrices suivantes :

$$A = \begin{pmatrix} 0 & 3 \\ 2 & 1 \end{pmatrix}, B = \begin{pmatrix} 2 & 4 & 6 \\ 3 & 5 & 8 \end{pmatrix}, C = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 2 & 1 \\ 1 & 2 & 1 \end{pmatrix}.$$

Solution. — $\operatorname{rg}(A) = \operatorname{rg}\begin{pmatrix} 0 & 3 \\ 2 & 1 \end{pmatrix} \overset{L_1 \leftrightarrow L_2}{=} \operatorname{rg}\begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix} = 2.$

—

$$\begin{aligned} \operatorname{rg}(B) &= \operatorname{rg}\begin{pmatrix} 2 & 4 & 6 \\ 3 & 5 & 8 \end{pmatrix} \\ &\overset{L_1 \leftarrow \frac{1}{2}L_1}{=} \operatorname{rg}\begin{pmatrix} 1 & 2 & 3 \\ 3 & 5 & 8 \end{pmatrix} \\ &\overset{L_2 \leftarrow L_2 - 3L_1}{=} \operatorname{rg}\begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -1 \end{pmatrix} = 2. \end{aligned}$$

— $\operatorname{rg}(C) = \operatorname{rg}\begin{pmatrix} 1 & 2 & 1 \\ 1 & 2 & 1 \\ 1 & 2 & 1 \end{pmatrix} \overset{\substack{L_2 \leftarrow L_2 - L_1 \\ L_3 \leftarrow L_3 - L_1}}{=} \operatorname{rg}\begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = 1.$

4.2.10 Inversion de matrices et systèmes d'équations linéaires

Lemme 4.2.38. *Soient $A, B \in \mathcal{M}_n(\mathbb{K})$ telles que pour toute matrice colonne X de $\mathcal{M}_{n,1}(\mathbb{K})$ on ait $AX = BX$. Alors $A = B$.*

Remarque 4.2.39. Il ne s'agit en aucun cas d'une "simplification par X". En termes d'applications linéaires, l'analogie est la suivante :

$$\forall x \in E \quad f(x) = g(x) \quad \Longrightarrow \quad f = g.$$

Démonstration. Soient $j \in [\![1, n]\!]$ et X la colonne $\begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \leftarrow j$, c'est-à-dire que $X_k = \begin{cases} 0 & \text{si } k \neq j \\ 1 & \text{si } k = j \end{cases}$ Comme $AX = BX$ (égalité entre deux vecteurs colonnes),

$$\forall i \in [\![1, n]\!] \quad (AX)_{i,1} = (BX)_{i,1} \text{ c'est-à-dire } \sum_{k=1}^{n} A_{i,k} X_{k,1} = \sum_{k=1}^{n} B_{i,k} X_{k,1}.$$

Les coefficients du vecteur X étant nuls sauf pour $k = j$, on obtient :

$$\forall i \in [\![1, n]\!] \quad A_{i,j} = B_{i,j}.$$

L'égalité étant valable pour j quelconque, on a l'égalité des matrices. □

Théorème 4.2.40. *Soit $A \in \mathcal{M}_n(\mathbb{K})$. La matrice A est inversible si et seulement si le système associé à l'égalité $AX = Y$ avec Y quelconque est un système de Cramer.*

Démonstration. $\Rightarrow$ | On suppose A inversible. $AX = Y$ donc $X = A^{-1}Y$. Ainsi, pour tout Y, l'équation $AX = Y$ admet une unique solution $X = A^{-1}Y$. Le système est donc de Cramer.

$\Leftarrow$ | On suppose que $AX = Y$ est un système de Cramer. À l'aide de la méthode du pivot, on peut exprimer les inconnues x_i comme combinaisons linéaires des paramètres y_j, c'est-à-dire :

$$\begin{cases} x_1 = b_{11}y_1 + b_{12}y_2 + \ldots + b_{1n}y_n \\ \vdots \\ x_i = b_{i1}y_1 + b_{i2}y_2 + \ldots + b_{in}y_n \\ \vdots \\ x_n = b_{n1}y_1 + b_{n2}y_2 + \ldots + b_{nn}y_n \end{cases}$$

Notons alors B la matrice $(b_{ij})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant n}}$ associée. On a $X = BY$. D'où $Y = AX = ABY$, c'est-à-dire $(AB)Y = I_nY$. Y étant quelconque, d'après le lemme précédent, $AB = I_n$. La matrice A est donc inversible. □

4.2.11 Détermination pratique de l'inverse d'une matrice

La démonstration du théorème précédent nous fournit une méthode pratique pour inverser une matrice ! En effet, la matrice B qui apparaît lors de la résolution du système $AX = Y$ n'est rien d'autre que l'inverse de A. Ainsi, pour inverser une matrice $A \in GL_n(\mathbb{K})$, on pourra poser $X, Y \in \mathcal{M}_{n,1}(\mathbb{K})$ et résoudre le système $AX = Y$ à l'aide de la méthode du pivot de Gauss.

Exemple 4.2.41. Inversons la matrice $A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$. On résout pour cela

$$AX = Y \quad \text{avec} \quad X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad \text{et} \quad Y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} :$$

$$\begin{cases} x_1 + x_2 = y_1 \\ x_1 + 2x_2 = y_2 \end{cases} \overset{L_2 \leftarrow L_2 - L_1}{\Longleftrightarrow} \begin{cases} x_1 + x_2 = y_1 \\ x_2 = y_2 - y_1 \end{cases} \Leftrightarrow \begin{cases} x_1 = 2y_1 + y_2 \\ x_2 = -y_1 + y_2 \end{cases}$$

On a $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$. A est donc inversible et $A^{-1} = \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}$. On n'oubliera pas, lors de la dernière étape, de bien ordonner les termes y_1 et y_2 pour ne pas inverser les coefficients de la matrice A^{-1}.

Evidemment, toute matrice n'est pas inversible. L'équivalence du théorème précédent montre que la résolution du système $AX = Y$ avec $A \notin \mathrm{GL}_n(\mathrm{K})$ conduira à l'apparition d'un système échelonné incompatible.

Exemple 4.2.42. Essayons d'inverser la matrice $B = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$. Nous savons d'avance que c'est peine perdue car les deux colonnes sont identiques donc proportionnelles.

$$\begin{cases} x_1 + x_2 = y_1 \\ 2x_1 + 2x_2 = y_2 \end{cases} \overset{L_2 \leftarrow L_2 - 2L_1}{\Longleftrightarrow} \begin{cases} x_1 + x_2 = y_1 \\ 0 = y_2 - 2y_1 \end{cases}$$

Le système obtenu est bien incompatible : il admet 0 solution ou une infinité (selon la valeur de $y_2 - 2y_1$). La matrice B n'est donc pas inversible.

En pratique, si on s'intéresse seulement à l'inversibilité d'une matrice, on calculera son déterminant. Si l'énoncé demande de calculer l'inverse, on pourra se lancer dans la résolution d'un système. Il ne s'agit pas de la seule méthode ! On peut aussi recourir à l'approche suivante, présentée sur un exemple :

Exemple 4.2.43. Soit $A = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$. Montrer que $A^2 - 4A - I_2 = 0_2$, en déduire que A est inversible.

▷ On a :

$$A^2 = \begin{pmatrix} 5 & 8 \\ 8 & 13 \end{pmatrix}$$

donc :

$$A^2 - 4A - I_2 = \begin{pmatrix} 5 & 8 \\ 8 & 13 \end{pmatrix} - \begin{pmatrix} 4 & 8 \\ 8 & 12 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Ainsi, $A^2 - 4A = I_2$ donc : $A(A - 4I_2) = I_2$. Ainsi, A est inversible, d'inverse $A^{-1} = A - 4I_2 = \begin{pmatrix} -3 & 2 \\ 2 & -1 \end{pmatrix}$.

Exercice 4.2.44. Pour $a \in \mathbb{R}$, soient $A = \begin{pmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & a \end{pmatrix}, N = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ et $B = A + N$.

1. Vérifier que $AN = NA$ et $N^3 = 0$.
2. Calculer B^n pour tout entier $n \geqslant 2$.
3. Pour quelles valeurs de a la matrice B est-elle inversible ? Calculer B^{-1} pour ces valeurs de a.

Solution. 1. $AN = NA = \begin{pmatrix} 0 & a & 0 \\ 0 & 0 & a \\ 0 & 0 & 0 \end{pmatrix}$ et $N^3 = 0$.

2. Comme les matrices A et N commutent, on peut utiliser la formule du binome de Newton et on obtient :

$$\begin{aligned} B^n = (A+N)^n &= \sum_{k=0}^{n} \begin{pmatrix} n \\ k \end{pmatrix} N^k A^{n-k} \\ &= \begin{pmatrix} n \\ 0 \end{pmatrix} A^n + \begin{pmatrix} n \\ 1 \end{pmatrix} NA^{n-1} + \begin{pmatrix} n \\ 2 \end{pmatrix} N^2 A^{n-2} \end{aligned}$$

puisque $n \geqslant 2$ (sinon, il y aurait moins de termes dans la somme). Comme $A^k = \begin{pmatrix} a^k & 0 & 0 \\ 0 & a^k & 0 \\ 0 & 0 & a^k \end{pmatrix}$ et $N^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, on trouve pour $n \geqslant 2$,

$$\begin{aligned} B^n &= \begin{pmatrix} a^n & 0 & 0 \\ 0 & a^n & 0 \\ 0 & 0 & a^n \end{pmatrix} + n \begin{pmatrix} 0 & a^{n-1} & 0 \\ 0 & 0 & a^{n-1} \\ 0 & 0 & 0 \end{pmatrix} + \frac{n(n-1)}{2} \begin{pmatrix} 0 & 0 & a^{n-2} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= a^{n-2} \begin{pmatrix} a^2 & na & \frac{n(n-1)}{2} \\ 0 & a^2 & na \\ 0 & 0 & a^2 \end{pmatrix} \end{aligned}$$

3. La matrice B est triangulaire supérieure, elle est inversible si et seulement si tous ses coefficients diagonaux sont nont nuls (le déterminant est ici le

produit des coefficients diagonaux), c'est-à-dire si et seulement si a est non nul. Pour inverser B, on procède par pivot de Gauss :

$$BX = Y \Leftrightarrow \begin{pmatrix} a & 1 & 0 \\ 0 & a & 1 \\ 0 & 0 & a \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \Longleftrightarrow \begin{cases} ax_1 + x_2 = y_1 \\ ax_2 + x_3 = y_2 \\ ax_3 = y_3 \end{cases}$$

$$\Leftrightarrow \begin{cases} x_1 = \frac{1}{a}y_1 - \frac{1}{a^2}y_2 + \frac{1}{a^3}y_3 \\ x_2 = \frac{1}{a}y_2 - \frac{1}{a^2}y_3 \\ x_3 = \frac{1}{a}y_3 \end{cases}$$

$$\Leftrightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{a} & -\frac{1}{a^2} & \frac{1}{a} \\ 0 & \frac{1}{a} & -\frac{1}{a^2} \\ 0 & 0 & \frac{1}{a} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \Longleftrightarrow X = B^{-1}Y.$$

Exercice 4.2.45. Soient les matrices

$$M = \frac{1}{12}\begin{pmatrix} 6 & 8 & 6 \\ 3 & 4 & 3 \\ 3 & 0 & 3 \end{pmatrix} \quad \text{et} \quad D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

1. **a)** Trouver l'unique vecteur colonne X_1 dont la première coordonnée vaut 1 tel que $MX_1 = 0$.

   b) Trouver l'unique vecteur colonne X_2 dont la deuxième coordonnée vaut 1 tel que $MX_2 = \frac{1}{12}X_2$.

   c) Trouver l'unique vecteur colonne X_3 dont la dernière coordonnée vaut 2 tel que $MX_3 = X_3$.
2. On note P la matrice dont les vecteurs colonnes sont (dans l'ordre) X_1, X_2 et X_3. Montrer que la matrice P est inversible et calculer sa matrice inverse.
3. Montrer que la matrice $P^{-1}MP$ est diagonale et égale à D. Déterminer D^n pour tout entier $n \geqslant 1$.
4. Montrer par récurrence que, pour tout entier $n \geqslant 1$: $M^n = PD^nP^{-1}$.
5. En déduire que, pour tout entier naturel non nul n :

$$M^n = \frac{1}{11}\begin{pmatrix} 6 - 6 \times \left(\frac{1}{12}\right)^n & 6 + 16 \times \left(\frac{1}{12}\right)^n & 6 - 6 \times \left(\frac{1}{12}\right)^n \\ 3 - 3 \times \left(\frac{1}{12}\right)^n & 3 + 8 \times \left(\frac{1}{12}\right)^n & 3 - 3 \times \left(\frac{1}{12}\right)^n \\ 2 + 9 \times \left(\frac{1}{12}\right)^n & 2 - 24 \times \left(\frac{1}{12}\right)^n & 2 + 9 \times \left(\frac{1}{12}\right)^n \end{pmatrix}.$$

Solution. 1. On trouve $X_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, X_2 = \begin{pmatrix} 2 \\ 1 \\ -3 \end{pmatrix}$ et $X_3 = \begin{pmatrix} 6 \\ 3 \\ 2 \end{pmatrix}$.

2. On pourrait calculer le déterminant de P mais le fait qu'on puisse inverser P à l'aide de la méthode du pivot suffit à justifier l'inversibilité. On trouve $P^{-1} = \frac{1}{11}\begin{pmatrix} 11 & -22 & 0 \\ -3 & 8 & -3 \\ 1 & 1 & 1 \end{pmatrix}$.
3. On a bien $D = P^{-1}MP$ et $D^n = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{12^n} & 0 \\ 0 & 0 & 1 \end{pmatrix}$ pour $n \geqslant 1$.

4. Classique, c'est du cours !
5. Il suffit de calculer explicitement PD^nP^{-1} avec les expressions de P^{-1} et D^n données précédemment.

Exercice 4.2.46. Soit $A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$.

1. Déterminer A^n pour tout $n \in \mathbb{N}^*$. On pourra commencer par calculer $A^2, A^3, \ldots$
2. Soient $(x_n)_{n\in\mathbb{N}}$ et $(y_n)_{n\in\mathbb{N}}$ les suites définies par : $\begin{cases} x_0, y_0 \in \mathbb{R} \\ x_{n+1} = x_n - y_n \\ y_{n+1} = -x_n + y_n \end{cases}$
   a) On pose $X_n = \begin{pmatrix} x_n \\ y_n \end{pmatrix}$. Établir une relation entre X_{n+1}, A et X_n.
   b) Montrer alors que $X_n = A^n X_0$.
   c) En déduire une expression de x_n et y_n en fonction de x_0, y_0 et n pour tout $n \in \mathbb{N}^*$.

Solution. 1. On s'aperçoit que $A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, A^2 = \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}, A^3 = \begin{pmatrix} 4 & -4 \\ -4 & 4 \end{pmatrix}$ et $A^4 = \begin{pmatrix} 8 & -8 \\ -8 & 8 \end{pmatrix}$. On peut dès lors conjecturer que :

$$\forall n \geqslant 1 \quad A^n = 2^{n-1} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = 2^{n-1} A$$

Démontrons ce résultat par récurrence, sachant qu'il est évidemment vrai pour $n = 1$.

$$A^{n+1} = A^n A = 2^{n-1} A \times A = 2^{n-1} A^2 = 2^n A.$$

D'après le principe de récurrence, le résultat est vrai quel que soit $n \geqslant 1$.

2. **a)** On a $X_{n+1} = AX_n$.
   b) Par récurrence, $X_n = AX_{n-1} = A \times AX_{n-2} = A^2 X^{n-2} = \ldots = A^n X_0$.
   c) Ainsi,

$$\begin{aligned} X_n = \begin{pmatrix} x_n \\ y_n \end{pmatrix} &= A^n X_0 \\ &= 2^{n-1} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \\ &= \begin{pmatrix} 2^{n-1}x_0 - 2^{n-1}y_0 \\ -2^{n-1}x_0 + 2^{n-1}y_0 \end{pmatrix}. \end{aligned}$$

Par identification, pour tout $n \geqslant 1$,

$$x_n = 2^{n-1}x_0 - 2^{n-1}y_0 \quad \text{et} \quad y_n = -2^{n-1}x_0 + 2^{n-1}y_0.$$

4.2.12 Déterminant d'une matrice carrée

Pour les matrices carrées de taille 2, il est facile de trouver une condition nécessaire et suffisante simple d'inversibilité ainsi qu'une formule pour le calcul de l'inverse le cas échéant.

Définition 4.2.47 (Déterminant, inversibilité et inverse d'une matrice carrée de taille 2)**.** Soient $a, b, c, d \in \mathbb{K}$. On appelle déterminant de $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$, noté : $\det \begin{pmatrix} a & c \\ b & d \end{pmatrix}$ ou $\begin{vmatrix} a & c \\ b & d \end{vmatrix}$, le scalaire $ad-bc$. La matrice $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$ est inversible si et seulement si : $\begin{vmatrix} a & c \\ b & d \end{vmatrix} \neq 0$. Dans ce cas : $\begin{pmatrix} a & c \\ b & d \end{pmatrix}^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$.

Démonstration. Par un simple calcul :

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} d & -c \\ -b & a \end{pmatrix} = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix} \begin{pmatrix} a & c \\ b & d \end{pmatrix} = (ad-bc) I_2.$$

— Si : $ad - bc \neq 0$, $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$ est inversible par définition de l'inversibilité d'inverse $\frac{1}{ad-bc} \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$.

— Pour la réciproque, supposons par l'absurde que : $ad - bc = 0$ ET que $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$ est inversible. Alors :

$$\begin{aligned} \begin{pmatrix} d & -c \\ -b & a \end{pmatrix} = I_2 \times \begin{pmatrix} d & -c \\ -b & a \end{pmatrix} &= \begin{pmatrix} a & c \\ b & d \end{pmatrix}^{-1} \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} d & -c \\ -b & a \end{pmatrix} \\ &= \begin{pmatrix} a & c \\ b & d \end{pmatrix}^{-1} \times \underbrace{(ad-bc)}_{=0} I_2 \\ &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \end{aligned}$$

donc : $a = b = c = d = 0$. La matrice $\begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ est ainsi NON inversible. contradiction.

□

▷ Pour une matrice d'ordre 3, on a :

$$\begin{aligned} \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} &= a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - d \begin{vmatrix} b & c \\ h & i \end{vmatrix} + g \begin{vmatrix} b & c \\ e & f \end{vmatrix} \text{ (dév./ première colonne)} \\ &= a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - b \begin{vmatrix} d & f \\ g & i \end{vmatrix} + c \begin{vmatrix} d & e \\ g & h \end{vmatrix} \text{ (dév./ première ligne) .} \end{aligned}$$

Exercice 4.2.48. Soient $\alpha_1, \alpha_2, \alpha_3$ des réels. Calculer le déterminant de la matrice :

$$V(\alpha_1, \alpha_2, \alpha_3) = \begin{pmatrix} 1 & 1 & 1 \\ \alpha_1 & \alpha_2 & \alpha_3 \\ \alpha_1^2 & \alpha_2^2 & \alpha_3^2 \end{pmatrix}$$

Une telle matrice est dite de Vandermonde.

Solution. On a :

$$\begin{aligned}\det\left(V\left(\alpha_1,\alpha_2,\alpha_3\right)\right) &= \begin{vmatrix}\alpha_2 & \alpha_3\\ \alpha_2^2 & \alpha_3^2\end{vmatrix} - \alpha_1\begin{vmatrix}1 & 1\\ \alpha_2^2 & \alpha_3^2\end{vmatrix} + \alpha_1^2\begin{vmatrix}1 & 1\\ \alpha_2 & \alpha_3\end{vmatrix}\\ &= \alpha_2\alpha_3^2 - \alpha_2^2\alpha_3 - \alpha_1\left(\alpha_3^2-\alpha_2^2\right) + \alpha_1^2\left(\alpha_3-\alpha_2\right)\\ &= \alpha_2\alpha_3\left(\alpha_3-\alpha_2\right) - \alpha_1\left(\alpha_3-\alpha_2\right)\left(\alpha_3+\alpha_2\right) + \alpha_1^2\left(\alpha_3-\alpha_2\right)\\ &= \left(\alpha_3-\alpha_2\right)\left(\alpha_2\alpha_3 - \alpha_1\left(\alpha_3+\alpha_2\right) + \alpha_1^2\right)\\ &= \left(\alpha_3-\alpha_2\right)\left(\alpha_2\left(\alpha_3-\alpha_1\right) - \alpha_1\left(\alpha_3-\alpha_1\right)\right)\\ &= \left(\alpha_3-\alpha_2\right)\left(\alpha_3-\alpha_1\right)\left(\alpha_2-\alpha_1\right).\end{aligned}$$

Définition 4.2.49 (Mineurs, cofacteurs, comatrice)**.** Soient $A \in \mathcal{M}_n(\mathbb{K})$ et $i,j \in [\![1,n]\!]$.

- On appelle mineur de A de position $(\ i,j\ )$ le déterminant de la matrice extraite de A par suppression de la $i^{\text{ème}}$ ligne et de la $j^{\text{ème}}$ colonne. Nous le noterons $\Delta_{ij}(A)$ dans ce cours mais la notation n'est pas universelle.
- On appelle cofacteur de A de position (i,j) le scalaire : $(-1)^{i+j}\Delta_{ij}(A)$.
- On appelle comatrice de A, notée $\mathrm{com}(A)$, la matrice des cofacteurs de A : $\mathrm{com}(A) = \left((-1)^{i+j}\Delta_{ij}(A)\right)_{1\leqslant i,j\leqslant n}$.

Théorème 4.2.50 (Développement par rapport à une ligne ou une colonne)**.** *Soit $A \in \mathcal{M}_n(\mathbb{K})$.*

- *Développement par rapport à une ligne : Pour tout $i \in [\![1,n]\!]$:* $\det(A) = \sum_{j=1}^{n}(-1)^{i+j}a_{ij}\Delta_{ij}(A)$.
- *Développement par rapport à une colonne : Pour tout $j \in [\![1,n]\!]$:* $\det(A) = \sum_{i=1}^{n}(-1)^{i+j}a_{ij}\Delta_{ij}(A)$.

En pratique : Le développement par rapport à une ligne/colonne est souvent utile, mais dans la mesure du possible, il faut choisir de développer par rapport à une ligne/colonne contenant beaucoup de zéros. En règle générale, je vous conseille de privilégier la méthode du pivot, qui fournit davantage des résultats sous forme factorisée (car après tout, ce que l'on veut savoir d'un déterminant, c'est souvent s'il est nul ou non).

Le déterminant d'une matrice sert de critère d'inversibilité.

Théorème 4.2.51. *$A \in \mathcal{M}_n(\mathbb{K})$ est inversible si et seulement* si $\det(A) \neq 0$. *Dans ce cas,* $\det\left(A^{-1}\right) = \frac{1}{\det A}$.

Démonstration. Montrons seulement l'implication. Si A est inversible, $AA^{-1} = I_n$ donc : $\det(A)\det\left(A^{-1}\right) = \det\left(I_n\right) = 1$. Le produit étant égal à $1, \det(A) \neq 0$ et on a : $\det\left(A^{-1}\right) = \frac{1}{\det(A)}$. □

Théorème 4.2.52 (Formule d'inversion)**.** *Soit $A \in \mathcal{M}_n(\mathbb{K})$. Alors : $A\,\mathrm{com}(A)^\top = \mathrm{com}(A)^\top A = \det(A)I_n$. En particulier, si A est inversible :*

$$A^{-1} = \frac{1}{\det(A)}\mathrm{com}(A)^\top.$$

4.2.13 Calcul de déterminant par la méthode du pivot

Théorème 4.2.53 (Déterminant d'une matrice et opérations élémentaires)**.** *Soient $i,j \in [\![1,n]\!]$ avec $i \neq j$ et $\lambda \in \mathbb{K}$.*

- *Les opérations élémentaires de la forme* $L_i \leftarrow L_i + \lambda L_j$ *et* $C_j \leftarrow C_j + \lambda C_i$ *ne modifient pas les déterminants.*

— *Les opérations élémentaires* $L_i \leftarrow \lambda L_i$ *et* $C_j \leftarrow \lambda C_j$ *multiplient les déterminants par* λ.
— *Les opérations élémentaires* $L_i \leftrightarrow L_j$ *et* $C_j \leftrightarrow C_i$ *multiplient les déterminants par* -1.

Exercice 4.2.54. Calculer le déterminant de la matrice :

$$A = \begin{pmatrix} 5 & 4 & 2 & 1 \\ 2 & 3 & 1 & -2 \\ -5 & -7 & -3 & 9 \\ 1 & -2 & -1 & 4 \end{pmatrix}$$

en effectuant des opérations élémentaires.

Solution. Les opérations $L_2 \leftarrow L_2 - \frac{2}{5}L_1, L_3 \leftarrow L_3 + L_1, L_4 \leftarrow L_4 - \frac{1}{5}L_1$ donnent :

$$\begin{aligned} \det(A) &= \begin{vmatrix} 5 & 4 & 2 & 1 \\ 0 & \frac{7}{5} & \frac{1}{5} & -\frac{12}{5} \\ 0 & -3 & -1 & 10 \\ 0 & -\frac{14}{5} & -\frac{7}{5} & \frac{19}{5} \end{vmatrix} = 5 \begin{vmatrix} \frac{7}{5} & \frac{1}{5} & -\frac{12}{5} \\ -3 & -1 & 10 \\ -\frac{14}{5} & -\frac{7}{5} & \frac{19}{5} \end{vmatrix} \\ &= 5\frac{1}{5}\frac{1}{5} \begin{vmatrix} 7 & 1 & -12 \\ -3 & -1 & 10 \\ -14 & -7 & 19 \end{vmatrix} = \frac{1}{5} \begin{vmatrix} 7 & 1 & -12 \\ -3 & -1 & 10 \\ -14 & -7 & 19 \end{vmatrix} \end{aligned}$$

Puis les opérations $L_2 \leftarrow L_2 + \frac{3}{7}L_1, L_3 \leftarrow L_3 + \frac{14}{7}L_1 = L_3 + 2L_1$ donnent :

$$\begin{aligned} \det(A) &= \frac{1}{5} \begin{vmatrix} 7 & 1 & -12 \\ 0 & -\frac{4}{7} & \frac{34}{7} \\ 0 & -5 & -5 \end{vmatrix} = \frac{1}{5} \cdot 7 \cdot \frac{2}{7} \cdot 5 \begin{vmatrix} -2 & 17 \\ -1 & -1 \end{vmatrix} \\ &= 2 \cdot 19 = 38. \end{aligned}$$

Exercice 4.2.55. Soient $n \geq 2$ un entier et $\alpha_1, \alpha_2, \cdots, \alpha_n$ des réels.

1. Calculer le déterminant $\Delta(\alpha_1, \cdots, \alpha_n)$ de la matrice :

$$V(\alpha_1, \cdots, \alpha_n) = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{n-1} & \alpha_2^{n-1} & \cdots & \alpha_n^{n-1} \end{pmatrix}$$

   Une telle matrice est dite de Vandermonde.

2. A quelle condition une telle matrice est-elle inversible ?

Solution. Pour $n = 2$, on a $\Delta(\alpha_1, \alpha_2) = \alpha_2 - \alpha_1$ et pour $n = 3$, on a fait le calcul avec l'exercice 4.2.48.

1. Le calcul de $\Delta(\alpha_1, \cdots, \alpha_n)$ se fait par récurrence sur $n \geq 2$. En retranchant, pour $i = n, n-1, \cdots, 2$ à la ligne i la ligne $i-1$ multipliée par α_1,

on obtient :

$$\Delta(\alpha_1,\cdots,\alpha_n)=\begin{vmatrix} 1 & 1 & \cdots & 1 \\ 0 & \alpha_2-\alpha_1 & \cdots & \alpha_n-\alpha_1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \alpha_2^{n-2}(\alpha_2-\alpha_1) & \cdots & \alpha_n^{n-2}(\alpha_n-\alpha_1) \end{vmatrix}$$

$$=\begin{vmatrix} \alpha_2-\alpha_1 & \alpha_3-\alpha_1 & \cdots & \alpha_n-\alpha_1 \\ \alpha_2(\alpha_2-\alpha_1) & \alpha_3(\alpha_3-\alpha_1) & \cdots & \alpha_n(\alpha_n-\alpha_1) \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_2^{n-2}(\alpha_2-\alpha_1) & \alpha_3^{n-2}(\alpha_3-\alpha_1) & \cdots & \alpha_n^{n-2}(\alpha_n-\alpha_1) \end{vmatrix}$$

soit :

$$\Delta(\alpha_1,\cdots,\alpha_n)=\left(\prod_{k=2}^{n}(\alpha_k-\alpha_1)\right)\begin{vmatrix} 1 & \cdots & 1 \\ \alpha_2 & \cdots & \alpha_n \\ \vdots & \ddots & \vdots \\ \alpha_2^{n-2} & \cdots & \alpha_n^{n-2} \end{vmatrix}$$

$$=\left(\prod_{k=2}^{n}(\alpha_k-\alpha_1)\right)\Delta(\alpha_2,\cdots,\alpha_n)$$

et par récurrence :

$$\Delta(\alpha_1,\cdots,\alpha_n)=\prod_{k=2}^{n}(\alpha_k-\alpha_1)\prod_{2\leq i<j\leq n}(\alpha_j-\alpha_i)$$

$$=\prod_{1\leq i<j\leq n}(\alpha_j-\alpha_i).$$

2. Cette matrice est inversible si, et seulement si, les α_i sont deux à deux distincts.

Conclusion

Le lecteur curieux et courageux est renvoyé aux ouvrages suivant [AF87, DM18, Gri11, ML07, LM03, Tau05] qui contiennent des précisions sur les notions abordées.

Bibliographie

[AF87] Jean-Marie Arnaudiès and Henri Fraysse. *Cours de mathématiques 1, Algèbre*. Dunod, 1987.

[DM18] C. Deschamps and F. Moulin. *Mathématiques tout-en-un, 1re année Cours et exercices corrigés*. Dunod, 2018.

[Gri11] Joseph Grifone. *Algèbre linéaire*. CÉPADUÈS-ÉDITIONS, 2011.

[LM03] François Liret and Dominique Martinais. *algèbre 1e année*. Dunod, 2003.

[ML07] Jean-Pierre Marco and Laurent Lazzarini. *Mathématiques L1 Cours complet avec 1000 tests et exercices corrigés*. Pearson Education, 2007.

[Tau05] Patrice Tauvel. *algèbre*. Dunod, 2005.