Dauphine | PSL
UNIVERSITÉ PARIS

Département MIDO, première année

Algèbre 1

Applications, nombres complexes, polynômes et matrices

Alexandre Afgoustidis

(version du 25 décembre 2019)

Alexandre Afgoustidis
CEREMADE, Université Paris-Dauphine, 75016 Paris, France.
E-mail : `afgoustidis@ceremade.dauphine.fr`

Ce polycopié a été créé avec LaTeX; pour la mise en forme, nous avons adapté des fichiers de style fournis par la Société Mathématique de France, notamment la classe `smfbook`.

ALGÈBRE 1

Alexandre Afgoustidis

TABLE DES MATIÈRES

AVANT DE COMMENCER

CHAPITRE 1

APPLICATIONS

Si E et F sont deux ensembles quelconques, la notion générale d'application de E dans F a été évoquée dans le cours « raisonnement » de la pré-rentrée, au chapitre 2. Nous y avons notamment défini les notions suivantes :

- Application $f : E \to F$;
- Composée de deux applications ;
- Ensemble image $f(A)$ d'une partie $A \subset E$ de l'ensemble de départ ;
- Ensemble image réciproque $f^{-1}(B)$ d'une partie $B \subset F$ de l'ensemble d'arrivée.

Dans ce chapitre, nous nous baserons sur ces notions et sur les résultats vus lors de la pré-rentrée.

1. Rappels ; notion de restriction et de prolongement

1.1. Restriction et co-restriction. —

Définition 1.1 – Restriction d'une application

Soient E et F deux ensembles et $f : E \to F$ une application.
Si A est une partie de l'ensemble de départ E, on appelle *restriction de f à A*, et on note $f|_A$, l'application

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

La restriction de f à A n'est donc rien d'autre qu'une version de f où on a « changé l'ensemble de départ ».

Exemple 1.2. — Si on considère

$$\begin{array}{rccc} f & : & \mathbb{R}^\star & \to & \mathbb{R} \\ & & x & \mapsto & \frac{5x}{|x|}, \end{array}$$

alors f n'est pas constante. Cependant, la restriction $f|_{\mathbb{R}_+^\star}$ est une fonction constante : il s'agit de la fonction

$$\begin{array}{rccc} f|_{\mathbb{R}_+^\star} & : & \mathbb{R}_+^\star & \to & \mathbb{R} \\ & & x & \mapsto & 5. \end{array}$$

Exemple 1.3. — Les fonctions

$$\begin{array}{rccc} f & : & \mathbb{R} & \to & \mathbb{R} \\ & & x & \mapsto & x^3 \end{array} \quad \text{et} \quad \begin{array}{rccc} g & : & \mathbb{R} & \to & \mathbb{R} \\ & & x & \mapsto & \sqrt{x^6} \end{array}$$

ne sont pas identiques, puisque $f(-1) \neq g(-1)$; cependant, on a $f|_{\mathbb{R}^+} = g|_{\mathbb{R}^+}$.

Définition 1.4 – Co-restriction d'une application

Soient E et F deux ensembles et $f : E \to F$ une application.
Soit B une partie de l'ensemble d'arrivée F vérifiant la propriété suivante : $\forall x \in E,\ f(x) \subset B$.
On appelle *co-restriction de f à B*, et on note $f|^B$, l'application

$$\begin{array}{rccc} f|^B & : & E & \to & B \\ & & x & \mapsto & f(x). \end{array}$$

Exemple 1.5. — Considérons la fonction

$$\begin{array}{rccc} f & : & \mathbb{R} & \to & \mathbb{R} \\ & & x & \mapsto & x^6. \end{array}$$

Comme $f(x)$ est positif pour tout $x \in \mathbb{R}$, on peut considérer la co-restriction $f|^{\mathbb{R}^+}$: il s'agit simplement de l'application

$$\begin{array}{rccc} g & : & \mathbb{R} & \to & \mathbb{R}^+ \\ & & x & \mapsto & x^6. \end{array}$$

Les fonctions f et $g = f|^{\mathbb{R}^+}$ ne sont pas identiques : par exemple, il existe des éléments de l'espace d'arrivée de f n'ayant aucun antécédent par f, alors que tout élément de l'espace d'arrivée de g admet au moins un antécédent par g.

Remarque 1.6. — On ne peut considérer la co-restriction de f à B que si l'ensemble $f(E) = \{f(x),\ x \in E\}$ est inclus dans B. Si l'on voulait définir la fonction « $f = \exp|^{[6,+\infty[}$ » comme

$$\begin{array}{ccc} \mathbb{R} & \to & [6,+\infty[\\ x & \mapsto & \exp(x), \end{array} \qquad \text{(Définition impossible !)}$$

on aurait bien du mal à ne pas avoir de problème pour $f(0)$.

1.2. Prolongement. —

Définition 1.7 – Prolongement d'une application

Soient E et F deux ensembles et $f : E \to F$ une application.
Soient Ω un ensemble contenant E et $g : \Omega \to F$ une application.
On dit que g est un *prolongement de f à Ω* lorsque $g|_E = f$.

Exemple 1.8. — Considérons l'application

$$\begin{array}{rccc} f & : & \mathbb{R}_+^\star & \to & \mathbb{R} \\ & & x & \mapsto & x^\pi = e^{\pi \ln(x)}. \end{array}$$

Cette fonction est définie sur $E = \mathbb{R}_+^\star$, et non sur $\Omega = \mathbb{R}$. Cependant, si l'on définit

$$\begin{array}{rccc} g & : & \mathbb{R} & \to & \mathbb{R} \\ & & x & \mapsto & \begin{cases} e^{\pi \ln(|x|)} & \text{si } x \neq 0, \\ 0 & \text{si } x = 0, \end{cases} \end{array}$$

alors on obtient une une application qui prolonge f à $\Omega = \mathbb{R}$.

***Exemple 1.9* (Prolongement du "sinus cardinal").** — Considérons l'application

$$\begin{array}{rccc} f & : & \mathbb{R}^\star & \to & \mathbb{R} \\ & & x & \mapsto & \frac{\sin(x)}{x}. \end{array}$$

Cette fonction n'est pas définie sur $\mathbb{R}$, puisque $f(0)$ n'a pas de sens. Cependant, si l'on définit

$$\begin{array}{rccc} g & : & \mathbb{R} & \to & \mathbb{R} \\ & & x & \mapsto & \begin{cases} \frac{\sin(x)}{x} & \text{si } x \neq 0, \\ 1 & \text{si } x = 0, \end{cases} \end{array}$$

alors on obtient une application qui est définie sur $\mathbb{R}$ tout entier, et qui prolonge f. On peut remarquer que $\lim\limits_{x\to 0} \frac{\sin(x)}{x} = 1$: la fonction g ci-dessus est donc continue en zéro (vous reparlerez en détail de la notion de continuité dans le cours « Analyse 1 »).

Il existe bien sûr d'autres prolongements de f à $\Omega = \mathbb{R}$: si l'on définit

$$\begin{array}{rccc} h & : & \mathbb{R} & \to & \mathbb{R} \\ & & x & \mapsto & \begin{cases} \frac{\sin(x)}{x} & \text{si } x \neq 0, \\ 927846 & \text{si } x = 0, \end{cases} \end{array}$$

alors h est elle aussi un prolongement de f, mais elle n'est pas continue en zéro.

***Remarque 1.10* (Il existe de nombreux prolongements).** — Soient E et F deux ensembles quelconques et $f : E \to F$ une application quelconque. Si Ω est un ensemble qui contient E, et si $\Omega \neq E$, alors il existe presque toujours de nombreuses applications $g : \Omega \to F$ qui prolongent f. En effet, pour définir une telle application g,

- les valeurs de $g(x)$ pour $x \in E$ sont prescrites par f,
- mais le fait que g prolonge f ne donne, si l'on n'impose pas de condition supplémentaire, aucune contrainte sur la valeur de $g(x)$ pour $x \in \Omega \setminus E$.

***Exemple 1.11* (Prolongement par zéro).** — Soient E un ensemble quelconque et $f : E \to \mathbb{R}$ une application définie sur E à valeurs dans $\mathbb{R}$. Si Ω est un ensemble qui contient E, alors il est toujours possible de définir une application $\tilde{f} : \Omega \to \mathbb{R}$ en définissant

$$\begin{array}{rccc} \tilde{f} & : & \Omega & \to & \mathbb{R} \\ & & x & \mapsto & \begin{cases} f(x) & \text{si } x \in E, \\ 0 & \text{si } x \in \Omega \setminus E. \end{cases} \end{array}$$

Cette application prolonge toujours f : on a $\tilde{f}|_E = f$.

2. Injectivité et surjectivité

2.1. Injectivité : définition et premières propriétés. —

Commençons par un rappel de vocabulaire. Considérons deux ensembles E et F, une application $f : E \to F$. Fixons un élément y est de l'espace d'arrivée F. Si x est un élément de E, alors on dit que x est un *antécédent* de y par f lorsque $f(x) = y$.

Un élément y de F peut n'avoir aucun antécédent par f, il peut en avoir un et un seul, il peut en avoir plusieurs...

2.1.1. *Définition et exemples.* —

Définition 1.12 – Injectivité d'une application

Soient E et F deux ensembles et $f : E \to F$ une application.
On dit que f est *injective* lorsque tout élément de F admet *au plus un antécédent* par f.

Exemple 1.13. — Si $f : \mathbb{R} \to \mathbb{R}$ est la fonction $x \mapsto e^x + 3$, alors f est injective. En effet, si y est un élément de l'ensemble d'arrivée $\mathbb{R}$, alors deux cas peuvent se présenter :

- Si $y \leq 3$, alors y n'admet aucun antécédent par f ;
- Si $y > 3$, alors y admet un et un seul antécédent par f : le nombre $x = \ln(y-3)$ vérifie $f(x) = y$, et c'est le seul réel à vérifier cette propriété.

Dans les deux cas, le nombre y admet au plus un antécédent par f.

Exemple 1.14. — La fonction $\sin : \mathbb{R} \to \mathbb{R}$ n'est pas injective, car $\sin(0) = \sin(2\pi)$ alors que $0 \neq 2\pi$.

Exemple 1.15. — Si $f : \mathbb{R} \to \mathbb{R}$ est la fonction $x \mapsto x^2$, alors f n'est pas injective. En effet, le nombre 5 admet deux antécédents distincts, à savoir $\sqrt{5}$ et $-\sqrt{5}$.

***Exemple 1.16* (De l'importance du domaine de définition).** — En revanche, si l'on reprend la fonction f de l'exemple précédent, la restriction $h = f|_{\mathbb{R}^+}$ est l'application

$$\begin{array}{rccc} h & : & \mathbb{R}^+ & \to & \mathbb{R} \\ & & x & \mapsto & x^2, \end{array}$$

et cette application est injective. En effet, pour tout élément y de l'ensemble d'arrivée $\mathbb{R}$, il existe au plus un élément $x \in \mathbb{R}^+$ vérifiant $h(x) = y$, puisque selon la valeur de y, deux cas peuvent se présenter :

- Si $y < 0$, alors y n'admet aucun antécédent par h ;
- Si $y \geq 0$, alors y admet un et un seul antécédent par h : il s'agit du nombre $x = \sqrt{y}$.

Exemple 1.17. — Si E est un ensemble quelconque, alors l'application $\mathrm{id}_E : E \to E$ est injective (rappelons que l'application id_E est définie par : $\forall x \in E,\ \mathrm{id}_E(x) = x$).

Exemple 1.18. — Si $f : \mathbb{R} \to \mathbb{R}$ est une fonction constante, alors f n'est pas injective.

Proposition 1.19 – Injectivité d'une application : en pratique

Soient E et F deux ensembles et $f : E \to F$ une application.
Pour que f soit injective, il faut et il suffit que la propriété suivante soit vérifiée :

$$\forall x \in E,\ \forall x' \in E,\quad [\ f(x) = f(x') \implies x = x'\ ]. \tag{$\star$}$$

Démonstration. —

- La négation de la propriété ($\star$) de l'énoncé est la suivante :

$$\exists x \in E,\ \exists x' \in E \quad : \quad x \neq x' \quad \text{mais} \quad f(x) = f(x'). \tag{2.1}$$

Si cette négation est vérifiée, et si x et x' sont deux éléments de E comme dans (2.1), alors l'élément $y = f(x)$ de F admet deux antécédents distincts par f, à savoir x et x'. Il est donc impossible que f soit injective.

Nous constatons donc que si f est injective, alors la propriété ($\star$) de l'énoncé est nécessairement vérifiée.

- Réciproquement, montrons que si la propriété ($\star$) est vérifiée, alors f est injective. Si ($\star$) est vérifiée et si y est un élément de F, alors de deux choses l'une :
  - soit y n'admet aucun antécédent par f ;
  - soit il en admet au moins un, disons x, et la propriété de l'énoncé indique qu'il ne peut en admettre aucun autre : en effet si x' est un élément de E vérifiant $f(x') = y$, alors par ($\star$), on a nécessairement $x = x'$.

  Tout élément y de F admet donc au plus un antécédent par f, et l'application f est donc injective.

□

La proposition 1.19 est extrêmement pratique pour vérifier l'injectivité d'une application, notamment dans des situations abstraites. Voici un exemple d'utilisation.

***Exemple 1.20* (Un exemple détaillé).** — Considérons l'application

$$\begin{array}{rccc} f & : & \mathbb{N}^\star & \to & \mathbb{R} \\ & & n & \mapsto & n^2 - n\sqrt{5}. \end{array}$$

Montrons que f est injective en utilisant la proposition ci-dessus.

Soient n et n' deux éléments de $\mathbb{N}$. Supposons vérifiée l'égalité $f(n) = f(n')$ et montrons qu'alors on a nécessairement $n = n'$.

Partons du fait que l'hypothèse $f(n) = f(n')$ signifie

$$n^2 - n\sqrt{5} = (n')^2 - (n')\sqrt{5}.$$

En réarrangeant cette égalité, on obtient

$$n^2 - (n')^2 = n\sqrt{5} - (n')\sqrt{5},$$

autrement dit :

$$(n - n')(n + n') = (n - n')\sqrt{5}. \tag{2.2}$$

Si $(n - n')$ n'était pas nul, on pourrait en déduire que $\sqrt{5} = n + n'$, donc que $\sqrt{5}$ est un entier, et ce n'est pas vrai. De l'égalité (2.2), on peut donc déduire que $n = n'$: c'est le but que nous devions atteindre.

2.1.2. *L'exemple des fonctions strictement monotones.* — Dans le cas où l'ensemble de départ et l'ensemble d'arrivée sont *tous les deux inclus dans* $\mathbb{R}$, notre prochain résultat fournit de nombreux exemples d'applications injectives.

Rappelons qu'une fonction f est dite *strictement monotone* si elle est soit *strictement croissante*, soit *strictement décroissante*.

N'oubliez pas que si $f : \mathbb{R} \to \mathbb{R}$ est une application, la situation où f n'est *ni croissante sur* $\mathbb{R}$, *ni décroissante sur* $\mathbb{R}$ se présente très fréquemment.

Dans le cas *très particulier* où f est strictement monotone, on dispose du résultat suivant.

Proposition 1.21 – Cas des applications strictement monotones

Soient E et F deux parties de $\mathbb{R}$ et $f : E \to F$ une application. Si f est strictement monotone, alors f est injective.

Démonstration. — Supposons que f soit strictement croissante et vérifions que f est injective en utilisant le critère de la Proposition 1.19. Nous devons vérifier la propriété ($\star$) qui y apparaît ; en s'appuyant sur l'équivalence entre une implication et sa contraposée, cela revient à vérifier l'énoncé suivant :

$$\forall x \in E,\ \forall x' \in E,\quad (x \neq x' \implies f(x) \neq f(x'))\,.$$

Soient alors x et x' deux éléments de E. Supposons $x \neq x'$ et vérifions qu'on a nécessairement $f(x) \neq f(x')$. Comme l'ensemble de définition E est inclus dans $\mathbb{R}$, on sait que x et x' sont deux nombres réels ; on peut donc distinguer deux cas :

- Si $x < x'$, alors en utilisant l'hypothèse « f strictement croissante », on a $f(x) < f(x')$;
- Si $x > x'$, alors en utilisant l'hypothèse « f strictement croissante », on a $f(x) > f(x')$;

Dans tous les cas on a bien $f(x) \neq f(x')$, et cela conclut notre démonstration dans le cas où f est strictement croissante. Le cas où f est strictement décroissante se traite de la même manière. □

Exemple 1.22. — Considérons l'application

$$\begin{array}{rccc} f & : & \mathbb{R} & \to & \mathbb{R} \\ & & x & \mapsto & 2x + \cos(x). \end{array}$$

Il s'agit d'une fonction dérivable sur $\mathbb{R}$ dont la dérivée f' vérifie : $\forall x \in \mathbb{R},\ f'(x) = 2 - \sin(x)$; comme sin prend ses valeurs dans $[-1, 1]$, la dérivée f' est partout strictement positive, d'où l'on déduit que f est strictement croissante. Le résultat ci-dessus montre donc que la fonction f est injective.

Attention. —
- Ce résultat n'est valable que si l'ensemble de départ et l'ensemble d'arrivée de f sont des parties de $\mathbb{R}$.
- Si f est une fonction de $\mathbb{R}$ dans $\mathbb{R}$, le résultat ci-dessus donne une condition *suffisante, mais pas nécessaire,* pour que f soit injective. Il existe des fonctions qui sont injectives sans être strictement monotones : c'est le cas, par exemple, de l'application

$$\begin{array}{rccc} f & : & \mathbb{R} & \to & \mathbb{R} \\ & & x & \mapsto & \begin{cases} \frac{1}{x} & \text{si } x \neq 0, \\ 0 & \text{si } x = 0. \end{cases} \end{array}$$

(attention, cette fonction est *strictement décroissante sur* $\mathbb{R}_+^\star$ et *strictement décroissante sur* $\mathbb{R}_-^\star$, mais elle n'est pas *strictement décroissante sur* $\mathbb{R}$).

2.1.3. *Composition et injectivité.* —

Proposition 1.23 – Composition et injectivité

Soient E, F et G trois ensembles ; considérons deux applications $f : E \to F$ et $g : F \to G$.

1. *Si f et g sont injectives, alors $g \circ f$ est injective.*
2. *Si $g \circ f$ est injective, alors f est injective (mais on ne peut rien en déduire pour g).*

Démonstration. —

1. Supposons que f et g soient injectives et montrons que l'application $(g \circ f) : E \to G$ est injective. Utilisons pour cela la Proposition 1.19 : fixons deux éléments x, x' de E et supposons qu'on a $(g \circ f)(x) = (g \circ f)(x')$. Nous devons montrer que $x = x'$.

   Mais notre hypothèse peut s'écrire $g(f(x)) = g(f(x'))$, et si nous notons $a = f(x)$ et $a' = f(x')$, on peut la réécrire comme : $g(a) = g(a')$. La fonction g étant injective, on peut en déduire que $a = a'$. Mais dire que $a = a'$, c'est dire que $f(x) = f(x')$; comme f est injective, on en déduit $x = x'$, comme espéré.
2. Supposons que $(g \circ f)$ soit injective et montrons que l'application $f : E \to F$ est injective. Fixons deux éléments x, x' de E et supposons qu'on a $f(x) = f(x')$; nous devons montrer que $x = x'$.

   Mais si $f(x)$ et $f(x')$ sont identiues, alors $g(f(x))$ et $g(f(x'))$ sont identiques aussi : on a donc $(g \circ f)(x) = (g \circ f)(x')$. Comme $(g \circ f)$ est supposée injective, on peut bien en déduire $x = x'$.

□

2.2. Surjectivité. —

2.2.1. *Définition et exemples.* —

Définition 1.24 – Surjectivité d'une application

Soient E et F deux ensembles et $f : E \to F$ une application.
On dit que f est *surjective* lorsque tout élément de F admet *au moins un antécédent* par f.

Remarque 1.25. — Il y a un lien simple entre la surjectivité et l'ensemble *image directe de E par f*, noté $f(E)$ dans le cours « raisonnement » de la pré-rentrée. Rappelons que $f(E)$ est l'ensemble des éléments de F qui sont atteints par f. Affirmer que f est surjective revient donc à affirmer l'égalité $f(E) = F$.

Exemple 1.26. —

- La fonction $\ln : \mathbb{R}_+^\star \to \mathbb{R}$ est surjective : en effet, pour tout $y \in \mathbb{R}$, le nombre $x = e^y$ appartient à l'espace de départ $\mathbb{R}_+^\star$ et vérifie $f(x) = y$.
- Si $f : \mathbb{R} \to \mathbb{R}$ est la fonction $x \mapsto e^x + 3$, alors f n'est pas surjective. En effet, le nombre 2 n'admet aucun antécédent par f.
- Si $g : \mathbb{R} \to \mathbb{R}$ est la fonction $x \mapsto x^2$, alors g n'est pas surjective. En effet, l'élément -3 de l'espace d'arrivée n'admet aucun antécédent par g.

***Exemple 1.27* (De l'importance de l'espace d'arrivée).** — En revanche, si l'on reprend la fonction g de l'exemple précédent, la co-restriction $h = f|^{\mathbb{R}^+}$ est l'application

$$\begin{array}{rccc} h & : & \mathbb{R} & \to & \mathbb{R}^+ \\ & & x & \mapsto & x^2, \end{array}$$

et cette application est surjective. En effet, pour tout élément y de l'ensemble d'arrivée $\mathbb{R}$, il existe au moins un élément x de $\mathbb{R}$ vérifiant $x = h(y)$: on peut choisir par exemple $x = \sqrt{y}$.

Exemple 1.28. — Si E est un ensemble quelconque, alors l'application $\mathrm{id}_E : E \to E$ est surjective.

Exemple 1.29. — Si $f : \mathbb{R} \to \mathbb{R}$ est une fonction constante, alors f n'est pas surjective.

***Remarque 1.30* (Remarque théorique).** — Si $f : E \to F$ est une application qui n'est pas surjective, il est toujours possible de « modifier l'espace d'arrivée » pour obtenir une application surjective. En effet, si nous considérons l'ensemble $B = f(E) = \{y \in F \ / \ \exists x \in E \ : \ y = f(x)\}$, et si nous formons la co-restriction

$$\begin{array}{rccc} f|^B & : & E & \to & B \\ & & x & \mapsto & f(x), \end{array}$$

alors on obtient une application surjective.

2.2.2. *Composition et surjectivité.* —

Proposition 1.31 – Composition et surjectivité

Soient E, F et G trois ensembles ; considérons deux applications $f : E \to F$ et $g : F \to G$.

1. *Si f et g sont surjectives, alors $g \circ f$ est surjective.*
2. *Si $g \circ f$ est surjective, alors g est surjective (mais on ne peut rien en déduire pour f).*

Démonstration. —

1. Supposons que f et g soient surjectives ; pour vérifier que $(g \circ f) : E \to G$ est surjective, nous devons montrer que tout élément de G admet au moins un antécédent par $g \circ f$.

   Soit z un élément de G. Comme $g : F \to G$ est surjective, il existe au moins un élément y de F vérifiant $g(y) = z$. Fixons un tel y. Comme $f : E \to F$ est surjective, il existe au moins un élément x de E vérifiant $f(x) = y$.

   Mais si nous fixons un tel x, alors $(g \circ f)(x) = g(f(x)) = g(y) = z$, donc x fournit un antécédent de z par $(g \circ f)$, comme espéré.
2. Supposons $g \circ f$ surjective, et montrons qu'alors g est surjective. Pour cela, considérons un élément z de G et vérifions qu'il admet au moins un antécédent par g.

   Comme $(g \circ f) : E \to G$ est supposée surjective, il existe au moins un élément x de E vérifiant $(g \circ f)(x) = z$. Mais alors $z = g(f(x))$, et si nous posons $y = f(x)$, nous obtenons un élément $y \in F$ vérifiant $g(y) = z$, comme espéré.

□

3. Bijectivité ; bijection réciproque

3.1. Définition et exemples. —

Définition 1.32 – Bijectivité d'une application

Soient E et F deux ensembles et $f : E \to F$ une application.
On dit que f est *bijective* lorsqu'elle est à la fois injective et surjective, autrement dit, lorsque tout élément de F admet *un et un seul antécédent* par f.

Exemple 1.33. —

- La fonction $f : \mathbb{R}_+^\star \to \mathbb{R}$ définie par : $\forall x \in \mathbb{R},\ f(x) = \ln\left(\sqrt{3x}\right)$. Si y est un élément de l'espace d'arrivée $\mathbb{R}$, alors il existe un et un seul élément x de l'espace de départ $\mathbb{R}_+^\star$ qui vérifie $f(x) = y$: il s'agit du nombre $x = \frac{1}{3}(e^y)^2$.
- La fonction $\exp : \mathbb{R} \to \mathbb{R}$ n'est pas bijective, puisqu'elle n'est pas surjective : le nombre (-3) n'admet aucun antécédent par f.

Définition 1.34 – Bijection réciproque d'une application bijective

Soit $f : E \to F$ une application bijective.
La *bijection réciproque* de f est l'application $f^{-1} : F \to E$ définie de la manière suivante :

$$\begin{array}{rccl} f^{-1} & : & F & \to & E \\ & & y & \mapsto & \text{l'unique élément } x \text{ de } E \text{ vérifiant } y = f(x). \end{array}$$

Exemple 1.35. — Si l'on considère la fonction

$$\begin{array}{rccl} f & : & \mathbb{R} & \to & \mathbb{R}_+^\star \\ & & x & \mapsto & 4e^{5x}, \end{array}$$

alors on constate que pour tout $y \in \mathbb{R}_\star^+$, il existe un unique $x \in \mathbb{R}$ vérifiant $f(x) = y$, à savoir $x = \frac{1}{4}\ln(\frac{y}{5})$. Ainsi, f est bijective ; de plus, sa bijection réciproque est l'application

$$\begin{array}{rccl} f^{-1} & : & \mathbb{R}_+^\star & \to & \mathbb{R} \\ & & a & \mapsto & \frac{1}{4}\ln(\frac{a}{5}). \end{array}$$

Remarque 1.36. — Si $f : E \to F$ est une application quelconque et si y est un élément de l'espace d'arrivée f, la recherche des antécédents de y par f peut être vue comme la résolution de l'équation $f(x) = y$, d'inconnue $x \in E$. Dire que f est bijective, c'est dire que cette équation admet toujours une et une seule solution, à savoir $f^{-1}(y)$. On pourra retenir l'idée suivante :

Si f est bijective, alors pour tous $x \in E$ et $y \in F$, on a : $\quad y = f(x) \iff x = f^{-1}(y)$.

***Attention* (notations : $f^{-1}(B)$ vs $f^{-1}(y)$).** — Dans ce cours, la notation « f^{-1} » a été utilisée dans deux contextes différents, qu'il ne faut pas confondre :

- Si y est un *élément* de F, la notation $f^{-1}(y)$ n'a de sens *que si f est bijective* ; elle désigne alors un *élément de E*.
- Si B est une *partie* de F, la notation $f^{-1}(B)$ a *toujours* un sens, même si f n'est pas bijective : elle désigne alors un *ensemble* inclus dans E.

***Remarque 1.37* (La réciproque de la réciproque).** — Si $f : E \to F$ est bijective, alors l'application $f^{-1} : F \to E$ est elle-même bijective : si a est un élément de E, les antécédents de a par f^{-1} sont les éléments y de F vérifiant $f^{-1}(y) = a$; par définition de la bijection réciproque, le seul y vérifiant cela est $y = f(a)$.

Si $f : E \to F$ est bijective, alors $f^{-1} : F \to E$ l'est aussi et on a $(f^{-1})^{-1} = f$.

***Remarque 1.38* (Remarque théorique).** — Si $f : E \to F$ est une application qui n'est pas surjective, mais qui est injective, alors il est toujours possible de « modifier l'espace d'arrivée » pour obtenir une application bijective. En effet, comme nous l'avons vu dans la remarque 1.30, la co-restriction

$$\begin{array}{rccc} f|^{f(E)} & : & E & \to & f(E) \\ & & x & \mapsto & f(x) \end{array}$$

est surjective, et elle est bien sûr injective.

Pour résumer la discussion ci-dessus, on dit souvent que lorsque $f : E \to F$ est injective (mais non nécessairement surjective), l'application f *induit une bijection de E sur $f(E)$*.

3.2. Bijectivité et composition. —

Commençons par une application très simple de la définition de la bijection réciproque :

Proposition 1.39 – Composée d'une bijection et de sa réciproque

Soient E et F deux ensembles et $f : E \to F$ une application.
Si f est bijective, alors la bijection réciproque $f^{-1} : F \to E$ vérifie :

$$f^{-1} \circ f = id_E \qquad et \qquad f \circ f^{-1} = id_F.$$

Ce dernier résultat fournit une manière d'appréhender la bijectivité en utilisant la composition des applications plutôt qu'en examinant les antécédents des éléments de l'espace d'arrivée :

Proposition 1.40 – Caractérisation de la bijectivité avec la composition

Soient E et F deux ensembles et $f : E \to F$ une application.
Pour que f soit bijective, il faut et il suffit qu'il existe une application $g : F \to E$ vérifiant :

$$g \circ f = id_E \qquad et \qquad f \circ g = id_F.$$

Lorsque c'est le cas, l'application g coïncide nécessairement avec la bijection réciproque f^{-1}.

Démonstration. — Soit f une application de E dans F.

- Si f est bijective, alors nous venons de voir que l'application $g = f^{-1}$ vérifie $g \circ f = \mathrm{id}_E$ et $f \circ g = \mathrm{id}_F$.
- Réciproquement, supposons qu'il existe une application $g : F \to E$ vérifiant $g \circ f = \mathrm{id}_E$ et $f \circ g = \mathrm{id}_F$; montrons que f est bijective et que $f^{-1} = g$.
  - *Surjectivité :* l'égalité $f \circ g = \mathrm{id}_F$ signifie que pour tout y de F, on a $f(g(y)) = y$. Un antécédent de y est donc donné par $g(y)$: en particulier, l'application f est surjective.
  - *Injectivité :* Soient x et x' deux éléments de E. Supposons $f(x) = f(x')$. Puisque $f(x)$ et $f(x')$ sont identiques, leurs images par g le sont également : ainsi, $g(f(x)) = g(f(x'))$. Mais grâce à l'égalité $g \circ f = \mathrm{id}_E$, on sait que $g(f(x)) = x$ et $g(f(x')) = x'$. C'est donc que $x = x'$.
  - *Réciproque :* Nous en concluons donc que f est bijective. De plus, si y est un élément de F, nous avons vu que l'unique antécédent de y par f est donné par $g(y)$: c'est donc que $g = f^{-1}$, comme espéré.

□

Exemple 1.41 **(Un exemple d'involution).** — Considérons l'application suivante :

$$\begin{array}{rccc} f & : & \mathbb{R}^2 & \to & \mathbb{R}^2 \\ & & (x,y) & \mapsto & (x,-y). \end{array}$$

On remarque que pour tout $(x, y) \in \mathbb{R}^2$, on a $f\,[f(x, y)] = (x, -(-y)) = (x, y)$.
Par conséquent, on a $f \circ f = \mathrm{id}_{\mathbb{R}^2}$. On en déduit deux choses :

- l'application f est bijective
- et de plus, la bijection réciproque f^{-1} n'est autre que f.

Lorsqu'une application $f : E \to E$ vérifie $f \circ f = \mathrm{id}_E$ (autrement dit, lorsqu'en appliquant f « deux fois de suite » à un élément x de E, on obtient simplement x), on dit que f est une *involution*.

Proposition 1.42 – Bijection réciproque d'une composée

Soient E, F et G trois ensembles ; considérons deux applications $f : E \to F$ et $g : F \to G$.
Si f et g sont bijectives, alors $g \circ f$ est bijective et sa bijection réciproque est donnée par la formule suivante :

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

Démonstration. — Nous devons montrer que l'application $(g \circ f) : E \to G$ est bijective et identifier sa bijection réciproque. Pour cela, utilisons la Proposition 1.40 : si nous définissons $\Phi : G \to E$ comme l'application $f^{-1} \circ g^{-1}$, alors la proposition ci-dessus montre que si nous vérifions que $\Phi \circ (g \circ f) = \mathrm{id}_E$ et $(g \circ f) \circ \Phi = \mathrm{id}_G$, nous aurons gagné.

Mais pour tout x de E, nous avons

$$\begin{aligned}[\Phi \circ (g \circ f)](x) &= \Phi\left(g(f(x))\right) \\ &= f^{-1}(g^{-1}(g(f(x)))) \\ &= f^{-1}(f(x)) \quad \text{(en appliquant } g^{-1}(g(y)) = y \text{ à } y = f(x)) \\ &= x,\end{aligned}$$

et on prouve de même que $(g \circ f) \circ \Phi(z) = z$ pour tout $z \in G$. Cela achève notre démonstration.

□

Remarque 1.43. — Si $f : E \to F$ et $g : F \to G$ sont deux applications et si on sait que $g \circ f$ est bijective, alors en combinant les propositions 1.23 et 1.31, on constate que f est nécessairement injective et que g est nécessairement surjective.

Mais en général, on ne peut pas en déduire plus : il est tout à fait possible que g ne soit pas injective et que f ne soit pas surjective, comme le montre l'exemple où

$$\begin{array}{ccccc} f & : & \{0\} & \to & \mathbb{R} \\ & & x & \mapsto & 5 \end{array} \quad \text{et} \quad \begin{array}{ccccc} g & : & \mathbb{R} & \to & \{0\} \\ & & x & \mapsto & 0 : \end{array}$$

dans ce cas, l'application $g \circ f : \{0\} \to \{0\}$ est l'application $\mathrm{id}_{\{0\}}$, qui est (pour des raisons peu pronfondes...) une bijection, alors que f n'est pas injective et que g n'est pas surjective.

Exercices du chapitre 1

Retour sur « antécédent, image directe, image réciproque ». —

Exercice 1.1. — ★☆☆

Soient E, F et G trois ensembles ; considérons une application $f : E \to F$ et une application $g : F \to G$.

1. Montrer que pour toute partie A de E, on a l'égalité $(g \circ f)(A) = g(f(A))$.
2. Montrer que pour toute partie B de F, on a l'égalité $(g \circ f)^{-1}(B) = f^{-1}(g^{-1}(B))$.

Exercice 1.2. — ★☆☆

Considérons l'application

$$\begin{array}{rccc} f & : & \mathbb{N}\times\mathbb{N} & \to & \mathbb{N} \\ & & (n,p) & \mapsto & n+p. \end{array}$$

1. Déterminer les ensembles $f^{-1}(\{3\})$, $f(\mathbb{N}\times\{2\})$.
2. Déterminer l'ensemble $f\left((2\mathbb{N})\times(3\mathbb{N})\right)$ où, pour $a \in \mathbb{N}$, la notation $a\mathbb{N}$ désigne l'ensemble $\{ak,\ k \in \mathbb{N}\}$.

Exercice 1.3. — ★★☆

1. Considérons l'application

$$\begin{array}{rccc} f & : & \mathbb{R}^2 & \to & \mathbb{R} \\ & & (x,y) & \mapsto & e^x + e^y. \end{array}$$

   Déterminer l'image directe $f(\mathbb{R}^2)$.
2. Considérons l'application

$$\begin{array}{rccc} g & : & \mathbb{R}^2 & \to & \mathbb{R} \\ & & (x,y) & \mapsto & e^x - e^y. \end{array}$$

   (a) Déterminer l'image réciproque $g^{-1}(\{0\})$.
   (b) Déterminer l'image directe $g(\mathbb{R}^2)$.
3. On considère à présent l'application

$$\begin{array}{rccc} h & : & \mathbb{R}^2 & \to & \mathbb{R}^2 \\ & & (x,y) & \mapsto & (e^x+e^y, e^x-e^y). \end{array}$$

   (a) Soit (X,Y) un élément fixé de $\mathbb{R}^2$. Déterminer une condition nécessaire et suffisante pour que (X,Y) soit atteint par h ; dans ce cas, déterminer tous les antécédents de (X,Y).
   (b) Déterminer l'image réciproque $h^{-1}\left([0,2]\times\{0\}\right)$.

Exercice 1.4. — ★☆☆

Soient E un ensemble et $f : E \to E$ une application. On suppose vérifiée l'égalité $f \circ f = f$.

Montrer l'équivalence suivante :

$$\forall x \in E, \quad \left(x \in f(E) \iff f(x) = x\right).$$

Injectivité et surjectivité : exemples concrets. —

Exercice 1.5. — ★☆☆

Les applications suivantes sont-elles injectives ? Surjectives ?

$$\begin{array}{rccc} f_1 & : & \mathbb{N} & \to & \mathbb{N} \\ & & n & \mapsto & n+1 \end{array} \qquad \begin{array}{rccc} f_2 & : & \mathbb{Z} & \to & \mathbb{Z} \\ & & n & \mapsto & -n \end{array} \qquad \begin{array}{rccc} f_3 & : & \mathbb{N} & \to & \mathbb{Z} \\ & & n & \mapsto & (-1)^n \end{array}$$

$$\begin{array}{rccc} f_4 & : & \mathbb{N} & \to & \mathbb{Z} \\ & & n & \mapsto & n\cdot(-1)^n \end{array} \qquad \begin{array}{rccc} f_5 & : & \mathbb{N} & \to & \mathbb{Z} \\ & & n & \mapsto & n^2\cdot(-1)^n. \end{array}$$

Exercice 1.6. — ★★☆

1. Existe-t-il une application $f : \mathbb{N} \to \mathbb{N}$ strictement décroissante ?
2. (a) Donner un exemple d'application $f : \mathbb{N} \to \mathbb{N}$ qui soit injective mais pas strictement croissante.
   (b) Donner un exemple d'application $f : \mathbb{N} \to \mathbb{N}$ vérifiant $f \circ f = \mathrm{id}_{\mathbb{N}}$ et $f \neq \mathrm{id}_{\mathbb{N}}$.
3. (a) Soit $f : \mathbb{N} \to \mathbb{N}$ strictement croissante. Montrer que pour tout $n \in \mathbb{N}$, on a $f(n) \geq n$.
   (b) Si f est injective, peut-on affirmer en général que $f(n) \geq n$ pour tout $n \in \mathbb{N}$?

Exercice 1.7. — ★☆☆

On considère l'application

$$\begin{array}{rccc} f & : & \mathbb{R} & \to & \mathbb{R} \\ & & x & \mapsto & xe^{-x}. \end{array}$$

1. Déterminer si f est injective et si f est surjective.
   (*On pourra utiliser librement des résultats d'analyse : tableau de variations, limites...).*
2. Déterminer les ensembles $f^{-1}(\{-e\})$, $f^{-1}(\{1\})$, $f(\mathbb{R}_+)$, $f^{-1}(\mathbb{R}_+)$.

Exercice 1.8. — ★★☆

Dans cet exercice, on fixe un entier $n \in \mathbb{N}^\star$ et on considère l'application

$$\begin{array}{rccc} f & : & \{1, 2, \ldots, (2n-1), 2n\} & \to & \{1, \ldots, n\} \\ & & k & \mapsto & \begin{cases} k/2 & \text{si } k \text{ est pair,} \\ (k+1)/2 & \text{si } k \text{ est impair.} \end{cases} \end{array}$$

1. Soit a un élément de $\{1, \ldots, n\}$; déterminer tous les antécédents de a par f.
   En déduire que f est surjective.
2. Le but de cette question est de montrer qu'il existe une fonction $g : \{1, \ldots, n\} \to \{1, \ldots, 2n\}$ qui vérifie $f \circ g = \mathrm{id}_{\{1,\ldots,n\}}$.
   (a) Considérons une fonction g solution du problème.
   Quelles sont les valeurs possibles pour $g(k)$ à $k \in \{1, \ldots, 2n\}$ fixé ?
   (b) En déduire explicitement une fonction g solution du problème.
   (c) Combien y a-t-il de fonctions g qui sont solution du problème ?

Exercice 1.9. — ★★☆

Dans cet exercice, on fixe un entier $n \in \mathbb{N}^\star$ et on considère l'application

$$\begin{array}{rccc} f & : & \{1, n\} & \to & \{1, \ldots, 2n\} \\ & & k & \mapsto & 2k. \end{array}$$

1. Montrer que f est injective.
2. Montrer qu'il existe une application $g : \{1, \ldots, 2n\} \to \{1, \ldots, n\}$ vérifiant $g \circ f = \mathrm{id}_{\{1,\ldots,n\}}$.
3. Combien y a-t-il d'applications $g : \{1, \ldots, 2n\} \to \{1, \ldots, n\}$ vérifiant $g \circ f = \mathrm{id}_{\{1,\ldots,n\}}$?

Injectivité et surjectivité : manipulations abstraites. —

Exercice 1.10. — ★☆☆

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 f est injective, alors g est surjective.

Exercice 1.11. — ★★☆

Soient $f : \mathbb{R} \to \mathbb{R}$ et $g : \mathbb{R} \to \mathbb{R}$ deux applications. On considère l'application

$$\begin{array}{rccc} h & : & \mathbb{R} & \to & \mathbb{R}^2 \\ & & x & \mapsto & (f(x), g(x)). \end{array}$$

1. Montrer que si l'une des deux applications f, g est injective, alors l'application h est injective.

2. On suppose que f et g sont surjectives. A-t-on nécessairement h surjective ?
3. Montrer que si h est surjective, alors les applications f et g sont nécessairement surjectives.
4. Donner un exemple où h est injective mais où aucune des deux applications f, g n'est injective.

Exercice 1.12. — ★★☆

Soient E et F deux ensembles et $f : E \to F$ une application.

1. (a) Montrer que pour toute partie A de E, on a $A \subset f^{-1}(f(A))$.
   (b) Montrer que si f est injective, alors pour toute partie A de E, on a $f^{-1}(f(A)) = A$.
   (c) Montrer réciproquement que si on a $f^{-1}(f(A)) \subset A$ pour toute partie A de E, alors f est injective.
2. (a) Montrer que pour toute partie B de F, on a $f(f^{-1}(B)) \subset B$.
   (b) Montrer que si f est surjective, alors pour toute partie B de F, on a $f(f^{-1}(B)) = B$.
   (c) Montrer réciproquement que si on a $f(f^{-1}(B)) = B$ pour toute partie B de F, alors f est surjective.

Exercice 1.13. — ★★★

Dans cet exercice, on fixe deux ensembles E et F et une application $f : E \to F$.

1. Dans cette question, on étudie l'application

$$\begin{array}{rccc} \mathfrak{Imdir} & : & \mathcal{P}(E) & \to & \mathcal{P}(F) \\ & & A & \mapsto & f(A). \end{array}$$

   (a) Montrer que f est injective si et seulement si $\mathfrak{Imdir}$ est injective.
   (b) Montrer que f est surjective si et seulement si $\mathfrak{Imdir}$ est surjective.
2. Dans cette question, on étudie l'application

$$\begin{array}{rccc} \mathfrak{Imrec} & : & \mathcal{P}(F) & \to & \mathcal{P}(E) \\ & & B & \mapsto & f^{-1}(B). \end{array}$$

   (a) Montrer que f est injective si et seulement si $\mathfrak{Imrec}$ est surjective.
   (b) Montrer que f est surjective si et seulement si $\mathfrak{Imrec}$ est injective.

Bijectivité : exemples concrets. —

Exercice 1.14. — ★☆☆ On considère les applications

$$\begin{array}{rccc} f & : & R_- & \to & \mathbb{R}_+ \\ & & x & \mapsto & x^2 \end{array} \quad \text{et} \quad \begin{array}{rccc} g & : & \mathbb{R}_- & \to & \mathbb{R}_+ \\ & & x & \mapsto & \sqrt{|x|}. \end{array}$$

1. Les applications $g \circ f$ et $f \circ g$ sont-elles bien définies ?
2. Montrer que f et g sont bijectives et déterminer leurs bijections réciproques.

Exercice 1.15. — ★☆☆ On considère les fonctions

$$\begin{array}{rccc} f & : &]1,+\infty[& \to &]0,+\infty[\\ & & x & \mapsto & \ln\left(\frac{x+1}{x-1}\right), \end{array} \qquad \begin{array}{rccc} g & : & \mathbb{R} & \to & \mathbb{R} \\ & & x & \mapsto & \sqrt[3]{1-x^3} \end{array}$$

et

$$\begin{array}{rccc} h & : & \mathbb{R} & \to & \mathbb{R} \\ & & x & \mapsto & \begin{cases} x^6 & \text{si } x \geq 0, \\ x^3 & \text{si } x \leq 0. \end{cases} \end{array}$$

Montrer que f, g et h sont bijectives et déterminer leurs bijections réciproques.

Bijectivité : manipulations abstraites. —

Exercice 1.16. — ★☆☆

On considère des ensembles E, F, G, H et des applications $f : E \to F$, $g : F \to G$ et $h : G \to H$.
Montrer que si $g \circ f$ et $h \circ g$ sont toutes les deux bijectives, alors f, g et h sont toutes les trois bijectives.

Exercice 1.17. — ★★☆

On considère un ensemble E et une application $f : E \to E$. On suppose vérifiée l'égalité $f \circ f \circ f = f$.
Montrer l'équivalence suivante : f est surjective si et seulement si f est injective.

Exercice 1.18. — ★★☆

On considère deux ensembles E et F et une application $f : E \to F$.

1. Montrer qu'il n'y a en général pas d'inclusion systématique entre $f(E \setminus A)$ et $F \setminus f(A)$.
2. Montrer l'équivalence suivante : (f est bijective) $\iff$ ($\forall A \in \mathcal{P}(E), \quad f(E \setminus A) = F \setminus f(A)$).

Exercice 1.19. — ★★☆

On considère un ensemble E et une application $f : E \to E$. On suppose vérifiée l'égalité $f \circ f = f$.
Montrer l'équivalence suivante : f est injective ou surjective si et seulement si $f = \mathrm{id}_E$.

Exercice 1.20. — ★★☆

Soit E un ensemble non vide. On fixe deux parties A et B de E et on considère l'application

$$\begin{array}{rccc} \Phi & : & \mathcal{P}(E) & \to & \mathcal{P}(A) \times \mathcal{P}(B) \\ & & X & \mapsto & (X \cap A, X \cap B). \end{array}$$

1. Dans cette question, on suppose $A \cup B = E$ et $A \cap B = \emptyset$.
   Montrer que Φ est bijective et déterminer la bijection réciproque Φ^{-1}.
2. Dans cette question, on suppose que Φ est bijective. Montrer que $A \cup B = E$ et $A \cap B = \emptyset$.

CHAPITRE 2

NOMBRES COMPLEXES

1. Généralités

1.1. Le corps des nombres complexes. —

Théorème 2.1 – Existence de $\mathbb{C}$

Il existe un ensemble $\mathbb{C}$, muni de deux opérations $+$ et $\times$, et vérifiant les propriétés suivantes :

(1) L'ensemble $\mathbb{C}$ contient $\mathbb{R}$ et contient un élément, noté i, qui vérifie $i \times i = (-1)$;

(2) pour tout élément z de $\mathbb{C}$, il existe un unique couple $(x, y) \in \mathbb{R}^2$ vérifiant $z = x + iy$;

(3) les opérations $+$ et $\times$ peuvent s'expliciter de la manière suivante :
pour tous z, z' dans $\mathbb{C}$, si l'on écrit $z = x + iy$ et $z' = x' + iy'$ avec x, y, x', y' réels, alors

$$z + z' = (x + x') + i(y + y') \qquad \text{et} \qquad zz' = (xx' - yy') + i(xy' + x'y). \tag{$\star$}$$

Vocabulaire : parties réelle et imaginaire, forme algébrique... —
Soit z un nombre complexe ; considérons l'unique couple (x, y) de nombres réels vérifiant $z = x + iy$.

- On dit que x est la *partie réelle* de z et que y est la *partie imaginaire* de z,
  et on note $x = \mathfrak{Re}(z)$ et $y = \mathfrak{Im}(z)$.
- Lorsque $\mathfrak{Im}(z) = 0$, on dit (bien sûr) que le complexe z est *réel*,
  et lorsque $\mathfrak{Re}(z) = 0$, on dit que z est *imaginaire pur*. On note parfois $i\mathbb{R}$ l'ensemble des nombres imaginaires purs.
- On dit que l'écriture $z = x + iy$ donne la *forme algébrique* du nombre complexe z.
- L'expression « si z est un nombre complexe et si $z = a + ib$ est son écriture sous forme algébrique, » signifiera dans ce cours : « si z est un nombre complexe et si l'on note $a = \mathfrak{Re}(z)$ et $b = \mathfrak{Im}(z)$, ».

Remarque 2.2 **(Propriétés de l'addition et de la multiplication).** — Les calculs sur les nombres complexes reposent sur les propriétés suivantes des opérations $+$ et $\times$ de $\mathbb{C}$, qui sont des conséquences des propriétés de $\mathbb{R}$ et des formules $(\star)$:

- Associativité de $+$ et $\times$: pour tous z_1, z_2, z_3 de $\mathbb{C}$, on a $(z_1 + z_2) + z_3 = z_1 + (z_2 + z_3)$ et $z_1(z_2 z_3) = (z_1 z_2)z_3$;
- Commutativité de $+$ et $\times$: si z_1, z_2 sont deux nombres complexes, alors $z_1 + z_2 = z_2 + z_1$ et $z_1 z_2 = z_2 z_1$;
- Distributivité de $\times$ sur $+$: si z_1, z_2, z_3 sont trois nombres complexes, alors $z_1(z_2 + z_3) = z_1 z_2 + z_2 z_3$;
- Rôle des complexes $0 = 0 + i0$ et $1 = 1 + i0$: pour tout $z \in \mathbb{C}$, on a $z + 0 = z$, $z0 = 0z = 0$, $1z = z$.
- Existence des opposés et des inverses :
  - pour tout $z \in \mathbb{C}$, il existe un unique $z' \in \mathbb{C}$ vérifiant $z + z' = 0$, et il s'agit de $z' = (-1) \times z$.
  - pour tout $z \in \mathbb{C}$ non nul, il existe un unique complexe z' vérifiant $zz' = 1,$; on note $z' = \frac{1}{z}$.

La seule propriété qui ne soit pas conséquence immédiate des propriétés de $\mathbb{R}$ est l'existence des inverses multiplicatifs : si z est un nombre complexe et si l'on considère l'unique couple $(x, y) \in \mathbb{R}^2$ vérifiant $z = x + iy$, alors

on vérifie aisément que le nombre $\frac{x-iy}{x^2+y^2}$, bien défini car $z \neq 0$, donne l'inverse de z mentionné dans la dernière propriété.

Comme dans le cas de $\mathbb{R}$, les propriétés formelles de $+$ et $\times$ ont des conséquences pratiques considérables pour la manipulation des nombres complexes. Mentionnons les deux suivantes :

- si z et z' sont deux nombres complexes vérifiant $zz' = 0$, alors $z = 0$ ou $z' = 0$;
- si z et z' sont deux nombres complexes, alors pour tout $n \in \mathbb{N}^\star$, on a $(z+z')^n = \sum_{k=0}^{n} \binom{n}{k} z^k (z')^{n-k}$.

1.2. Représentation géométrique : le plan d'Argand. — Si z est un nombre complexe et si $z = x+iy$ est son écriture sous forme algébrique, alors le couple $(x, y) \in \mathbb{R}^2$ fournit un moyen de *représenter z par un point du plan* : si l'on a fixé un repère permettant de décrire les points du plan par deux coordonnées "cartésiennes", alors on représente z par le point $M(z)$ d'abscisse x et d'ordonnée y.

Réciproquement, si M est un point du plan et si (x, y) est le couple de nombres réels donnant son abscisse et son ordonnée dans le repère donné, alors le nombre complexe $z = x + iy$ est appelé *l'affixe de M*.

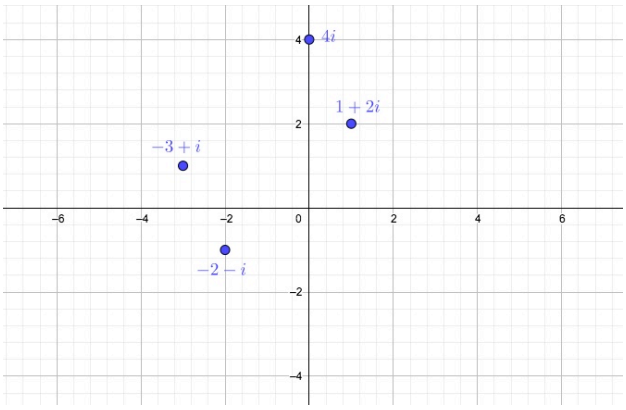

FIGURE 1. Si $z = x+iy$ est un nombre complexe, on peut le représenter par le point du plan dont les coordonnées cartésiennes sont données par le couple $(x, y) \in \mathbb{R}^2$.

L'addition des nombres complexes admet une interprétation géométrique simple : si z_1 et z_2 sont deux nombres complexes et si l'on note O, M_1, M_2 les points du plan d'affixes respectifs 0, z_1 et z_2, alors le point d'affixe $z_1 + z_2$ se trouve au quatrième sommet du parallélogramme dont les autres sommets sont O, M_1 et M_2.

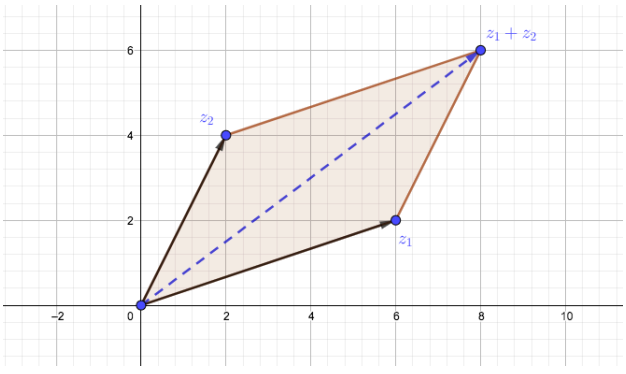

FIGURE 2. Somme de nombres complexes, version géométrique : le point M d'affixe $z_1 + z_2$ se trouve au quatrième sommet du parallélogramme dont les autres sommets sont O, $M_1 = M(z_1)$ et $M_2 = M(z_2)$.

1.3. Conjugué. —

Définition 2.3 – Conjugué d'un nombre complexe

Si z est un nombre complexe et si $z = a + ib$ est son écriture sous forme algébrique, on appelle conjugué de z, et on note $\overline{z}$, le nombre complexe

$$\overline{z} = a - ib.$$

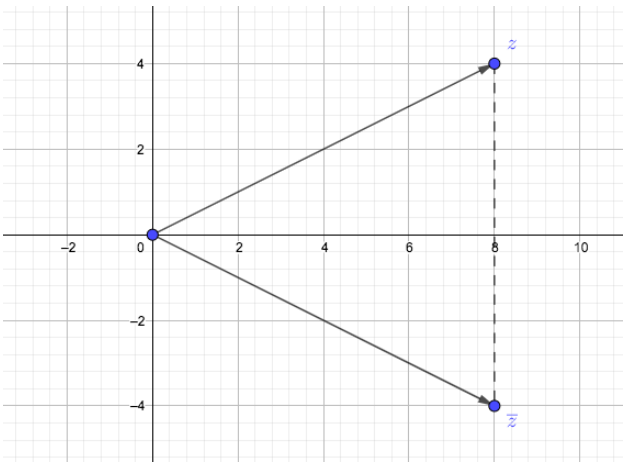

FIGURE 3. Si $z = x + iy$ est un nombre complexe, le point d'affixe $\bar{z}$ est le symétrique du point d'affixe z par rapport à l'axe des abscisses.

Relevons quelques propriétés élémentaires de la conjugaison :

Pour tout nombre complexe z, on a $\overline{\overline{z}} = z$.

Proposition 2.4 – Conjugué d'une somme et d'un produit

Pour tout nombre complexe z, on a les égalités suivantes :

$$\overline{z + z'} = \overline{z} + \overline{z'} \qquad \text{et} \qquad \overline{zz'} = \overline{z} \cdot \overline{z'}.$$

Démonstrations. — Si z et z' sont deux nombres complexes et si $z = x + iy$, $z = x' + iy'$ sont leurs écritures sous forme algébrique, alors on a les égalités suivantes :

$$\begin{aligned}
\overline{\overline{z}} &= \overline{x - iy} = x + i(-(-y)) = z, \\
\overline{z + z'} &= \overline{(x + x') + i(y + y')} = (x + x') - i(y + y') = (x - iy) + (x' - iy') = \overline{z} + \overline{z'} \\
\overline{zz'} &= \overline{(xx' - yy') + i(xy' + x'y)} = xx' - yy' - i(xy' + x'y) \\
&= (xx' - (-y)(-y')) + i(x(-y') + x'(-y)) = (x - iy)(x' - iy') = \overline{z}\overline{z'}.
\end{aligned}$$

Tout cela prouve les trois formules ci-dessus. □

Les remarques suivantes, qui rappellent comment reconstituer la partie réelle et la partie imaginaire d'un complexe z en fonction de z et de $\bar{z}$, sont très utiles en pratique :

Proposition 2.5 – Parties réelle et imaginaire et conjugué

Pour tout nombre complexe z, on a les égalités suivantes :

$$\mathfrak{Re}(z) = \frac{z+\bar{z}}{2} \qquad \textit{et} \qquad \mathfrak{Im}(z) = \frac{z-\bar{z}}{2i}.$$

Pour tout $z \in \mathbb{C}$, on a de plus les équivalences suivantes :

$$(z \text{ est un nombre réel} \iff \bar{z} = z) \qquad \text{et} \qquad (z \text{ est un imaginaire pur} \iff \bar{z} = -z)$$

Démonstration. — Si z est un nombre complexe et si $z = x + iy$ est son écriture sous forme algébrique, alors $\frac{z+\overline{z}}{2} = \frac{(x+iy)+(x-iy)}{2} = x = \mathfrak{Re}(z)$, tandis que $\frac{z-\overline{z}}{2i} = \frac{(x+iy)-(x-iy)}{2i} = y = \mathfrak{Im}(z)$. Cela prouve les deux égalités annoncées.

De plus, z est réel si et seulement si $\mathfrak{Im}(z)$=0, et d'après la formule précédente cela équivaut à $\frac{z-\overline{z}}{2i} = 0$, autrement dit à $z = \overline{z}$. De même, z est imaginaire pur si et seulement si $\mathfrak{Re}(z) = 0$, et cela équivaut à $\frac{z+\overline{z}}{2} = 0$, autrement dit $z = -\overline{z}$. comme annoncé. □

2. Module et argument

2.1. Le module d'un nombre complexe. —

Définition 2.6 – Module d'un nombre complexe

Soit z un nombre complexe. Écrivons $z = x + iy$ avec $(x, y) \in \mathbb{R}^2$.
On appelle *module de* z, et on note $|z|$, le nombre

$$|z| = \sqrt{x^2 + y^2}$$

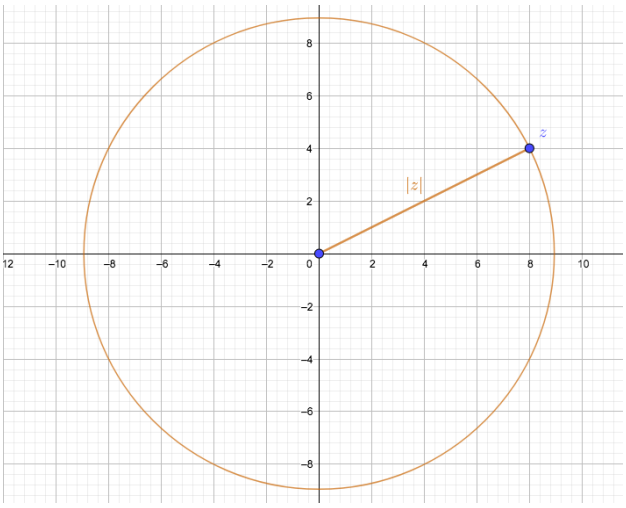

FIGURE 4. Si z est un nombre complexe, le réel positif $|z|$ donne la distance de z à l'origine.

Notation : le cercle unité $\mathbb{U}$. — *Compte tenu de l'interprétation du module , on appelle* cercle unité *de* $\mathbb{C}$ *l'ensemble des nombres complexes de module* 1, *et on le note souvent* $\mathbb{U}$ *:*

$$\mathbb{U} = \{z \in \mathbb{C} \ / \ |z| = 1\}.$$

Proposition 2.7 – Module et conjugué

1. *Pour tout $z \in \mathbb{C}$, on a $|\bar{z}| = |z|$.*
2. *Pour tout $z \in \mathbb{C}$, on a $z\bar{z} = |z|^2$.*
3. *Si z est un nombre complexe de module 1, alors $\bar{z} = \frac{1}{z}$.*

Démonstration. — La première propriété est claire d'après la définition du module et du conjugué. Pour la deuxième, si z est un nombre complexe et si $z = x + iy$ est son écriture sous forme algébrique, alors $z\bar{z} = (x+iy)(x-iy) = x^2 - (iy)^2 = x^2 - (-1)y^2 = x^2 + y^2 = |z|^2$. La troisième égalité est conséquence de la seconde et du fait que si $|z| = 1$, alors z ne peut pas être nul. □

Proposition 2.8 – Module et produit

1. *Pour tout $z \in \mathbb{C}$ et $z' \in \mathbb{C}$, on a $|zz'| = |z||z'|$.*
2. *Pour tout $z \in \mathbb{C}$, on a $|-z| = |z|$.*
3. *Si z est un nombre complexe non nul, on a $|1/z| = 1/|z|$.*

Démonstration. —

1. Si z et z' sont deux nombres complexes et si $z = a + ib$ et $z' = c + id$ sont leurs écritures sous forme algébrique, alors on peut calculer le module de zz' par la formule

$$\begin{aligned} |zz'| &= \sqrt{(ac-bd)^2 + (ad+bc)^2} = \sqrt{(a^2c^2 + b^2d^2 - 2abcd) + (a^2d^2 + b^2c^2 + 2abcd)} \\ &= \sqrt{a^2c^2 + b^2d^2 + a^2d^2b^2c^2} = \sqrt{(a^2+b^2)(c^2+d^2)} = \sqrt{|z|^2|z'|^2} = |z||z'|. \end{aligned}$$

2. En appliquant le point 1. à z quelconque et $z' = (-1)$, on trouve le deuxième point de la proposition.
3. Enfin, si z est un complexe non nul et si l'on applique le premier point à $z' = \frac{1}{z}$, on trouve $|z\frac{1}{z}| = |z||\frac{1}{z}|$; comme $z\frac{1}{z} = 1$ on a $|z\frac{1}{z}|$. On en déduit que les nombres réels $|z|$ et $|\frac{1}{z}|$ sont inverses l'un de l'autre, ce qui prouve le dernier point.

□

Proposition 2.9 – Inégalités entre module et partie réelle et imaginaire :

1. *Pour tout $z \in \mathbb{C}$, on a $|\mathfrak{Re}(z)| \leq |z|$, avec égalité si et seulement si z est un nombre réel.*
2. *Pour tout $z \in \mathbb{C}$, on a $|\mathfrak{Im}(z)| \leq |z|$, avec égalité si et seulement si z est imaginaire pur.*

Démonstration. — Si $z \in \mathbb{C}$ et si $z = x + iy$ est son écriture sous forme algébrique, alors on a $|z| = \sqrt{x^2+y^2} \geq \sqrt{x^2} = |x| = |\mathfrak{Re}(z)|$, avec égalité si et seulement si $y = 0$. Cela prouve le premier point. Le second se prouve de la même manière. □

Proposition 2.10 – Module et somme/différence

1. *Pour tout $z \in \mathbb{C}$ et $z' \in \mathbb{C}$, on a $|z + z'| \leq |z| + |z'|$,*
   avec égalité si et seulement s'il existe un nombre $\alpha \in \mathbb{R}$ vérifiant : α est positif et $z = \alpha z'$.
2. *Pour tout $z \in \mathbb{C}$ et $z' \in \mathbb{C}$, on a $||z| - |z'|| \leq |z - z'|$,*
   avec égalité si et seulement s'il existe un nombre $\alpha \in \mathbb{R}^+$ vérifiant $z = \alpha z'$.

Démonstration. — Soient z et z' deux nombres complexes.

1. • Commençons par remarquer que $s = |z+z'|$ et $t = |z|+|z'|$ sont deux nombres réels positifs ; on a donc l'équivalence $s \leq t \iff s^2 \leq t^2$.
Remarquons à présent que

$$\begin{aligned} s^2 &= (z+z')(\overline{z+z'}) = (z+z')(\overline{z}+\overline{z'}) = z\overline{z} + (z')(\overline{z'}) + \left[z\overline{z'} + z'\overline{z}\right] \\ &= |z|^2 + |z'|^2 + \left[z\overline{z'} + \overline{z\overline{z'}}\right]. \end{aligned}$$

Pour comparer s^2 et t^2, insérons deux remarques :
- On a $t^2 = |z|^2 + |z'|^2 + 2|zz'| = |z|^2 + |z'|^2 + 2|z||z'|$.
- On reconnaît dans l'expression entre crochets le nombre $2\mathfrak{Re}(z\bar{z'})$.

On constate donc l'égalité suivante :

$$s^2 = t^2 - 2|z||z'| + 2\mathfrak{Re}(z\bar{z'}),$$

qui peut se reformuler comme suit :

$$s^2 - t^2 = 2\left(\mathfrak{Re}(z\bar{z'}) - |z||\overline{z'}|\right). \tag{2.1}$$

D'après la proposition 2.9, cette quantité est négative, et vaut zéro si et seulement si $z\bar{z'}$ est un nombre réel.

• Étudions le cas d'égalité dans (2.1). Si $z = 0$ ou $z' = 0$, il y a bien sûr égalité. Si on suppose $z \neq 0$ et $z' \neq 0$, alors $z = \frac{z\overline{z'}}{\overline{z'}}$; d'après la proposition 2.7, on a $\frac{1}{\overline{z'}} = \frac{z'}{|z'|^2}$, ce qui prouve que $z = \left(\frac{z\overline{z'}}{|z'|^2}\right) z'$.
S'il y a égalité dans (2.1), on a donc bien $z = \alpha z'$ avec $\alpha = \left(\frac{z\overline{z'}}{|z'|^2}\right) \in \mathbb{R}$.

• Réciproquement, si $z = \alpha z'$ avec $\alpha \in \mathbb{R}$, alors on a $|z+z'| = |1+\alpha||z'|$, tandis que $|z|+|z'| = (1+|\alpha|)|z'|$. Ces deux quantités sont égales lorsque $\alpha \geq 0$, et ne sont pas égales lorsque $\alpha < 0$. L'égalité dans (2.1) a donc lieu si et seulement si $z = \alpha z'$ avec $\alpha \geq 0$.

2. • Si x et y sont deux nombres réels positifs et si l'on a $x \leq y$ et $-x \leq y$, alors $|x|y$. Pour prouver le point 2, il nous suffit donc de prouver les deux inégalités $(|z| - |z'|) \leq |z - z'|$ et $(|z'| - |z|) \leq |z - z'|$.

• Notons $a = z - z'$ et $b = z'$. D'après l'inégalité du point 1, on a $|a + b| \leq |a| + |b|$, autrement dit $|z| \leq |z - z'| + |z'|$. On a donc $(|z| - |z'|) \leq |z - z'|$.
En appliquant de même le point 1 à $a = z' - z$ et $b = z$, on obtient l'inégalité $(|z'| - |z|) \leq |z - z'|$.

• Pour les cas d'égalité :
  - On remarque que si l'inégalité $(|z| - |z'|) \leq |z - z'|$ est une égalité , alors d'après le point 1, il existe un réel $\alpha \in \mathbb{R}^+$ tel que $z - z' = \alpha z'$, mais en choisissant $\beta = 1 + \alpha$, on obtient $z = \beta z'$ où β est un nombre réel positif. On traite de même le cas où c'est l'inégalité $(|z'| - |z|) \leq |z - z'|$ qui est une égalité.
  - Enfin, si $z = \alpha z'$ avec $\alpha \in \mathbb{R}^+$, on a bien $||z| - |z'|| = |z - z'| = |1 - |\alpha|||z|$.

□

2.2. L'exponentielle complexe. — Dans ce cours, nous supposerons connues les fonction cos : $\mathbb{R} \to \mathbb{R}$ et sin : $\mathbb{R} \to \mathbb{R}$, ainsi que les formules de trigonométrie usuelles. Nous nous en servirons pour définir un nombre $e^{i\theta}$ si θ est un nombre réel, puis pour donner un sens l'expression e^z si z est un nombre complexe.

Définition 2.11 – Exponentielle complexe

Si θ est un nombre réel, on définit un nombre complexe $e^{i\theta}$ de module 1 par la formule

$$e^{i\theta} = \cos(\theta) + i\sin(\theta).$$

Si z est un nombre complexe et si $z = a + ib$ est son écriture sous forme algébrique, on définit un nombre complexe e^z par la formule

$$e^z = e^a e^{ib} = e^a\left(\cos(b) + i\sin(b)\right).$$

Proposition 2.12 – L'exponentielle complexe transforme les sommes en produits

Si z et z' sont deux nombres complexes, alors $e^{z+z'} = e^z e^{z'}$.

Démonstration. — Si b et b' sont deux nombres réels, le fait que $e^{i(b+b')} = e^{ib}e^{ib'}$ résulte des formules de trigonométrie pour $\cos(b+b')$ et $\sin(b+b')$. On a en effet

$$\begin{aligned} e^{ib}e^{ib'} &= (\cos(b) + i\sin(b))\,(\cos(b') + i\sin(b')) = [\cos(b)\cos(b') - \sin(b)\sin(b')] + i\,[\sin(b)\cos(b') + \cos(b)\sin(b')] \\ &= \cos(b+b') + i\sin(b+b') = e^{i(b+b')}. \end{aligned}$$

Si maintenant z et z' sont deux nombres complexes et si $z = a + ib$ et $x' = a' + ib'$ sont leurs écritures sous forme algébrique, alors

$$\begin{aligned} e^{z+z'} &= e^{(a+a')+i(b+b')} && \text{par définition de } z+z' \\ &= e^{a+a'}e^{i(b+b')} && \text{compte tenu de la définition 2.11} \\ &= e^a e^{a'} e^{ib} e^{ib'} && \text{puisque } e^{a+a'} = e^a e^{a'} \text{ (propriété de exp sur } \mathbb{R}) \text{ et que avons prouvé } e^{i(b+b')} = e^{ib}e^{ib'} \\ &= e^a e^{ib} e^{a'} e^{ib'} && \text{par commutativité du produit dans } \mathbb{C} \\ &= e^z e^{z'} && \text{compte tenu de la définition 2.11.} \end{aligned}$$

□

***Remarque 2.13* (Définir cos et sin à partir de exp, plutôt que l'inverse...)**

Dans ce cours, nous avons choisi de *supposer connues* cos *et* sin et leurs propriétés, y compris les mystérieuses formules de trigonométrie que vous avez apprises « par coeur » au lycée.

Mais comment, au juste, peut-on définir rigoureusement cos et sin, et d'où sortent ces formules ? Leur définition classique, si l'on ne veut pas s'appuyer sur des notions de géométrie (longueur, arc de cercle...), consiste précisément à renverser la logique du présent paragraphe :

- À l'aide d'outils que vous verrez en deuxième année, on peut d'abord *construire complètement* une fonction $\exp_{\mathbb{C}} : \mathbb{C} \to \mathbb{C}$ vérifiant les trois propriétés suivantes [1] :
  (i) $\exp_{\mathbb{C}}(0) = 1$
  (ii) $\exp_{\mathbb{C}}(z+z') = \exp_{\mathbb{C}}(z)\exp_{\mathbb{C}}(z')$ pour tout $(z,z') \in \mathbb{C}^2$,
  (iii) $\exp_{\mathbb{C}}(\overline{z}) = \overline{\exp_{\mathbb{C}}(z)}$ pour tout $z \in \mathbb{C}$.
- On peut alors *prouver* avec (i), (ii) et (iii) que pour tout $\theta \in \mathbb{R}$, $\exp_{\mathbb{C}}(i\theta)$ est un nombre complexe de module 1 : en effet, pour tout $z \in \mathbb{C}$, on a
$$|\exp_{\mathbb{C}}(i\theta)|^2 = \exp_{\mathbb{C}}(i\theta)\overline{\exp_{\mathbb{C}}(i\theta)} \overset{(iii)}{=} \exp_{\mathbb{C}}(i\theta)\exp_{\mathbb{C}}(-i\theta) \overset{(ii)}{=} \exp_{\mathbb{C}}(i\theta - i\theta) \overset{(i)}{=} \exp_{\mathbb{C}}(0) = 1.$$
- Pour $\theta \in \mathbb{R}$, on *définit* ensuite $\cos(\theta) = \mathfrak{Re}(e^{i\theta})$ et $\sin(\theta) = \mathfrak{Im}(e^{i\theta})$. C'est ainsi qu'on obtient les fonctions $\cos : \mathbb{R} \to \mathbb{R}$ et $\sin : \mathbb{R} \to \mathbb{R}$.
- Les propriétés de cos et sin sont alors des *conséquences* des théorèmes sur l'exponentielle complexe : par exemple, pour prouver les formules pour $\cos(b+b')$ et $\sin(b+b')$, il suffit d'utiliser le fait que résultent du fait que $e^{i(b+b')} = e^{ib}e^{ib'}$, alors que dans notre cours, nous avons présenté cette égalité sur l'exponentielle complexe comme une conséquence des formules pour cos et sin.

  Il y a beaucoup de travail (pour trouver les tableaux de variations et calculer les dérivées de cos et sin, pour définir π, pour prouver toutes les formules usuelles...), mais ce travail est possible.

1. La formule est la suivante : pour $z \in \mathbb{C}$, on définit $\exp_{\mathbb{C}}(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}$. Cette définition nécessite de comprendre la notion de « somme infinie, ce que vous ferez en deuxième année. De plus, prouver les propriétés (i), (ii) et (iii) à partir de cette formule n'est pas évident...

2.3. Arguments d'un complexe non nul ; notion d'argument principal. — L'étude de l'argument d'un nombre complexe, autrement dit de la notion d'angle dans le plan, repose sur l'étude du cercle unité $\mathbb{U}$ en utilisant la fonction

$$\begin{array}{rccc} \rho & : & \mathbb{R} & \to & \mathbb{U} \\ & & \theta & \mapsto & e^{i\theta}. \end{array}$$

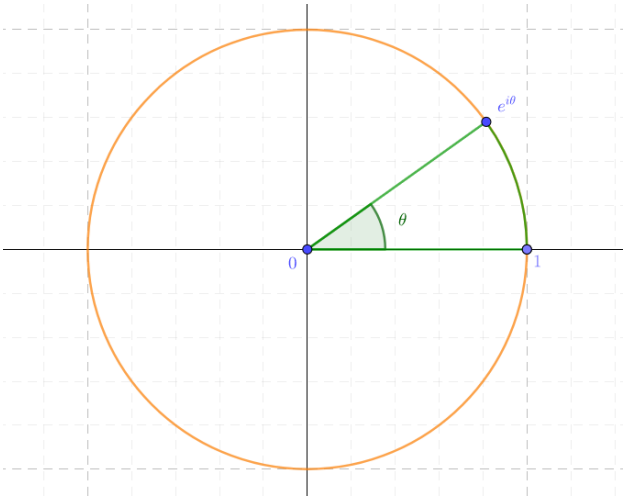

FIGURE 5. Si θ est un nombre réel, le point du plan d'affixe $e^{i\theta}$ est situé sur le cercle unité : il est obtenu en parcourant à partir du point d'affixe 1 un angle θ dans le sens direct.

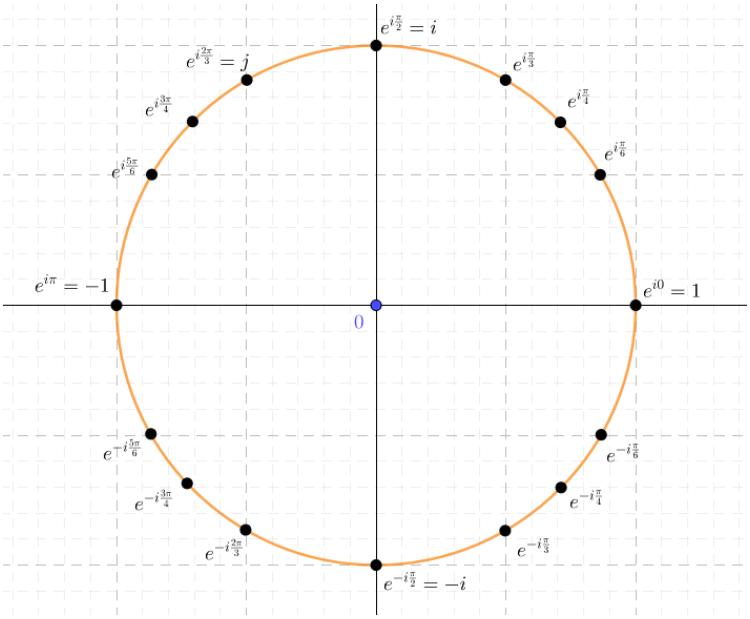

FIGURE 6. Quelques points remarquables du cercle unité.

Proposition 2.14 – Propriétés de l'application $\theta \mapsto e^{i\theta}$, partie I

1. *Si θ est un nombre réel, alors on a $e^{i\theta} = 1$ si et seulement s'il existe $k \in \mathbb{Z}$ tel que $\theta = 2k\pi$.*
2. *Si θ et θ' sont deux nombres réels, alors $e^{i\theta} = e^{i\theta'}$ si et seulement si $\theta - \theta' \equiv 0 \mod (2\pi)$.*

Démonstration. —

1. Si θ est un nombre réel, alors on a $e^{i\theta} = 1 = 1 + i0$ si et seulement si $\cos(\theta) = 1$ et $\sin(\theta) = 0$; les nombres θ vérifiant cette équation sont ceux qui sont de la forme $\theta = 2k\pi$, $k \in \mathbb{Z}$.
2. Si θ et θ' sont deux nombre réels, alors on a $e^{i\theta} = e^{i\theta'}$ si et seulement si $e^{i\theta}e^{-i\theta'} = 0$ (en effet, $e^{i\theta'}$ est non nul et de module 1, donc d'inverse $\overline{e^{i\theta'}} = e^{-i\theta'}$). Le résultat annoncé est donc conséquence du premier point.

□

Proposition 2.15 – Propriétés de l'application $\theta \mapsto e^{i\theta}$, partie II

1. L'application $\rho : \mathbb{R} \to \mathbb{U}$ est surjective : pour tout $u \in \mathbb{U}$, il existe un réel θ vérifiant $u = e^{i\theta}$.
2. L'application $\rho : \mathbb{R} \to \mathbb{U}$ n'est pas injective : tout élément de $\mathbb{U}$ admet une infinité d'antécédents.

Démonstration. —

1. Si u est un élément de $\mathbb{U}$, alors le module de u est 1, et d'après la proposition 2.9, sa partie réelle est comprise entre -1 et 1. On peut alors obtenir un antécédent à u en se rappelant que pour tout $x \in [-1, 1]$, on peut considérer le nombre $\arccos(x) \in [0, \pi]$, et que de plus $\arccos(x) \in [0, \frac{\pi}{2}]$ si $x \geq 0$, tandis que $\arccos(x) \in [-\frac{\pi}{2}, 0]$ si $x \leq 0$. Distinguons alors plusieurs cas selon les signes de $\mathfrak{Re}(u)$ et $\mathfrak{Im}(u)$:
   - Si $\mathfrak{Re}(u)$ et $\mathfrak{Im}(u)$ sont dans $[0, 1]$, choisissons $\theta = \arccos(\mathfrak{Re}(u))$; on constate alors que $\cos(\theta) = \mathfrak{Re}(u)$. De plus, on a $\theta \in [0, \frac{\pi}{2}]$, donc $\sin(\theta)$ est positif et $\sin(\theta) = \sqrt{1 - \cos^2(\theta)} = \mathfrak{Im}(u)$ puisque ce dernier nombre est positif. On a donc bien $e^{i\theta} = u$ pour ce choix de θ.
   - Si $\mathfrak{Re}(u) \in [0, 1]$ et $\mathfrak{Im}(u) \in [-1, 0]$ et , choisissons $\theta = -\arccos(\mathfrak{Re}(u))$; on constate alors que $\sin(\theta) < 0$ et $\cos(\theta) > 0$, ce qui prouve que $\sin(\theta) = -\sqrt{1 - \cos^2(\theta)} = -|\mathfrak{Im}(u)| = -\mathfrak{Im}(u)$ puisque ce dernier nombre est négatif.
   - Si $\mathfrak{Re}(u) \in [-1, 0]$ et $\mathfrak{Im}(u) \in [0, 1]$, en choisissant $\theta = \arccos(\mathfrak{Re}(u))$, on constate de même qu'on a $e^{i\theta} = u$.
   - Enfin, si $\mathfrak{Re}(u) \in [-1, 0]$ et $\mathfrak{Im}(u) \in [-1, 0]$, en choisissant $\theta = -\arccos(\mathfrak{Re}(u))$, on constate de même qu'on a $e^{i\theta} = u$.
2. Si u est un élément de $\mathbb{U}$ et si θ_0 est un antécédent de u par ρ (un tel antécédent existe d'après le premier point), alors tous les nombres de la forme $\theta_0 + 2k\pi$, $k \in \mathbb{Z}$, sont aussi des antécédents de $\mathbb{U}$.

□

Définition 2.16 – Arguments d'un nombre complexe non nul

Soient z un nombre complexe non nul et θ un nombre réel.
On dit que θ est un argument de z lorsque l'égalité suivante est vérifiée : $z = |z|e^{i\theta}$.

- Un nombre complexe non nul admet toujours une infinité d'arguments différents.
- Si $z \in \mathbb{C}^\star$ et si θ et θ' sont deux arguments de z, alors on a $\theta \equiv \theta' \mod (2\pi)$.

***Remarque 2.17* (Pourquoi 0 n'a pas d'argument).** — L'égalité $0 = |0|e^{i\theta}$ est vérifiée par *tous* les nombres réels θ ; il n'est donc pas raisonnable de parler d'argument du nombre complexe 0 (sinon, la cohérence imposerait que tout nombre réel soit un argument de 0).

Vocabulaire : forme trigonométrique. — Si z est un nombre complexe non nul et si θ est un argument de z, on dit que l'écriture $z = |z|e^{i\theta}$ est une écriture de z sous *forme trigonométrique*. Dans ce cours (et

dans les exercices qui vous attendent), l'expression « mettre z sous forme trigonométrique » signifie donc « trouver un réel $r > 0$ et un nombre $\theta \in \mathbb{R}$ vérifiant $z = re^{i\theta}$ ».

Exemple 2.18. — Considérons le nombre $z = 1 + i$. Le module de z est un nombre positif de carré $|z|^2 = 1^2 + 1^2 = 2$, donc $|z| = \sqrt{2}$. Pour trouver un argument de z, écrivons $z = \sqrt{2}\left(\frac{1}{\sqrt{2}} + \frac{i}{\sqrt{2}}\right) = \sqrt{2}\left(\frac{\sqrt{2}}{2} + i\frac{\sqrt{2}}{2}\right)$. Comme $\cos(\frac{\pi}{4}) = \sin(\frac{\pi}{4}) = \sqrt{2}2$, on reconnaît dans la parenthèse le nombre $e^{i\frac{\pi}{4}}$, donc une écriture de z sous forme trigonométrique est

$$z = \sqrt{2}e^{i\frac{\pi}{4}}.$$

***Exemple 2.19* (La "méthode de l'angle moitié").** — Soient α et β deux éléments de $[0, \frac{\pi}{2}]$. Déterminons le module et l'argument du nombre complexe $z = e^{i\alpha} + e^{i\beta}$.

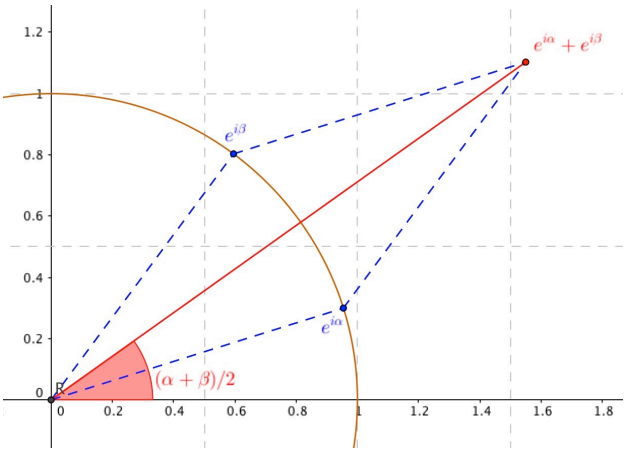

FIGURE 7. Illustration de la méthode de l'angle moitié : si A est le point d'affixe $e^{i\alpha}$ et B le point d'affixe $e^{i\beta}$, alors la droite qui passe par 0 et par $z = e^{i\alpha} + e^{i\beta}$ est la bissectrice de l'angle $\widehat{AOB}$.

Factorisons, dans l'écriture de z, par $e^{i\frac{\alpha+\beta}{2}}$. On obtient

$$\begin{aligned} z &= e^{i\frac{\alpha+\beta}{2}}\left[e^{i\left(\alpha-\frac{\alpha+\beta}{2}\right)} + e^{i\left(\beta-\frac{\alpha+\beta}{2}\right)}\right] \\ &= e^{i\frac{\alpha+\beta}{2}}\left[e^{i\left(\frac{\alpha-\beta}{2}\right)} + e^{i\left(\frac{\beta-\alpha}{2}\right)}\right] \\ &= e^{i\frac{\alpha+\beta}{2}}\left[2\cos\left(\frac{\alpha-\beta}{2}\right)\right]. \end{aligned}$$

A-t-on bien trouvé le module et un argument de z ? Pour le savoir, nous devons voir si $\cos\left(\frac{\alpha-\beta}{2}\right)$ est positif. Mais sous notre hypothèse sur α et β, on a $0 \leq \alpha \leq \frac{\pi}{2}$ et $-\frac{\pi}{2} \leq -\beta \leq 0$, donc $\frac{\alpha-\beta}{2} \in \left[-\frac{\pi}{4}, \frac{\pi}{4}\right]$, ce qui garantit bien $\cos\left(\frac{\alpha-\beta}{2}\right) \geq 0$.

Nous pouvons donc conclure : le module de z est égal à $2\cos\left(\frac{\alpha-\beta}{2}\right)$, et le nombre $\theta = \frac{\alpha+\beta}{2}$ donne un argument de z.

Théorème et définition 2.20 – Argument principal

Si z est un nombre complexe non nul, alors il existe un unique nombre réel θ vérifiant :

$$\theta \in]-\pi, \pi] \qquad \text{et} \qquad z = |z|e^{i\theta}.$$

Le nombre θ ci-dessus est appelé argument principal de z, et on note $\mathrm{Arg}(z)$. Pour $\alpha \in \mathbb{R}$, on a donc l'équivalence :

$$\alpha = \mathrm{Arg}(z) \iff \begin{cases} \alpha \in]-\pi, \pi] \\ z = |z|e^{i\alpha}. \end{cases}$$

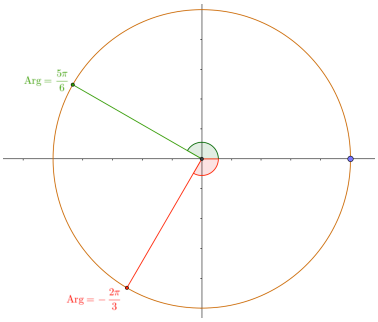

FIGURE 8. Illustration de la notion d'argument principal $\mathrm{Arg}(z)$: toujours compris entre $-\pi$ et π, il s'agit de l'angle nécessaire pour aller de 1 à z. On a $\mathrm{Arg}(z) \in [0, \pi]$ lorsque z est dans le demi-plan « haut », et $\mathrm{Arg}(z) \in]-\pi, 0]$ lorsque z est dans le demi-plan « bas ».

Proposition 2.21 – Propriétés de la fonction $\mathrm{Arg} : \mathbb{C}^\star \to]-\pi, \pi]$

Soient z et z' deux nombres complexes non nuls. On a les égalités suivantes :

- $\mathrm{Arg}(-z) \equiv \pi + \mathrm{Arg}(z) \mod (2\pi)$
- $\mathrm{Arg}(\overline{z}) \equiv -\mathrm{Arg}(z) \mod (2\pi)$
- $\mathrm{Arg}(\frac{1}{z}) \equiv -\mathrm{Arg}(z) \mod (2\pi)$
- $\mathrm{Arg}(zz') \equiv \mathrm{Arg}(z) + \mathrm{Arg}(z') \mod (2\pi)$

2.4. Formules de Moivre et d'Euler ; applications à la trigonométrie. —

Proposition 2.22 – Formules d'Euler : expression de $\cos(\theta)$ et $\sin(\theta)$ à l'aide de $e^{i\theta}$.

Pour tout réel θ, on a

$$\cos(\theta) = \frac{e^{i\theta} + e^{-i\theta}}{2} \qquad \textit{et} \qquad \sin(\theta) = \frac{e^{i\theta} - e^{-i\theta}}{2i}.$$

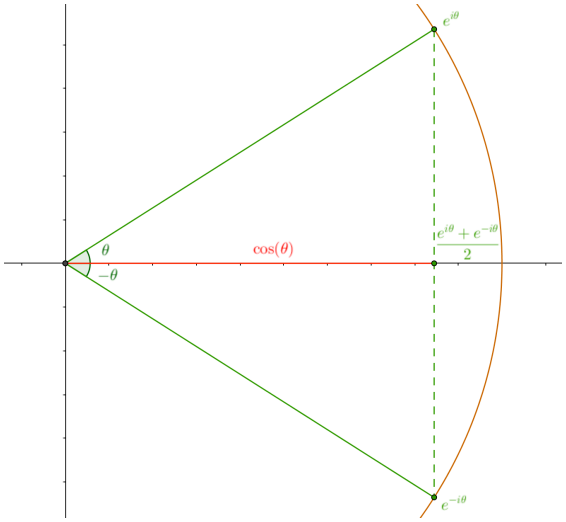

FIGURE 9. Illustration de la proposition 2.22 : le point d'affixe $\frac{e^{i\theta}+e^{-i\theta}}{2}$ est le milieu des points d'affixes $e^{i\theta}$ et $e^{-i\theta}$: il est situé sur l'axe des abscisses et son abscisse est exactement $\cos(\theta)$.

Exemple 2.23 **(Développement de $\cos(\theta)^3$ à l'aide du binôme.)** — Soit θ un nombre réel. En écrivant $\cos(\theta)^3 = \left(\frac{e^{i\theta}+e^{-i\theta}}{2}\right)^3$, on peut utiliser la formule du binôme pour obtenir une expression de $\cos(\theta)^3$:

$$\begin{aligned}\cos(\theta)^3 &= \frac{(e^{i\theta})^3 + 3(e^{i\theta})^2(e^{i-\theta}) + 3(e^{i\theta})(e^{-i\theta})^2 + (e^{-i\theta})^3}{2^3} \\ &= \frac{e^{3i\theta} + 3e^{i\theta} + 3e^{i\theta} + e^{-3i\theta}}{8} = \frac{(e^{3i\theta} + e^{-3i\theta}) + 3(e^{i\theta} + e^{-i\theta})}{8} \\ &= \frac{2\cos(3\theta) + 3\cdot(2\cos(\theta))}{8} \\ &= \frac{\cos(3\theta) + \cos(\theta)}{4}.\end{aligned}$$

Proposition 2.24 – Formules de Moivre pour exprimer $\cos(n\theta)$ et $\sin(n\theta)$.

Pour tout entier $n \in \mathbb{N}$ et pour tout réel θ, on a

$$\cos(n\theta) + i\sin(n\theta) = (\cos(\theta) + i\sin(\theta))^n .$$

En utilisant la formule du binôme et en séparant parties réelle et imaginaire, on peut alors trouver un lien entre $\cos(n\theta)$ et les puissances de $\cos(\theta)$ et $\sin(\theta)$

Exemple 2.25 **(Expression de $\cos(3\theta)$ à l'aide de la formule de Moivre)**

Soit θ un nombre réel. Remarquons que $\cos(3\theta)$ est la partie réelle de $e^{in\theta} = (e^{i\theta})^3 = (\cos(\theta) + i\sin(\theta))^3$, et développons avec la formule du binôme :

$$\begin{aligned}e^{3i\theta} &= \cos(\theta)^3 + 3\cos(\theta)^2 \times (i\sin(\theta)) + 3\cos(\theta) \times (i\sin(\theta))^2 + (i\sin(\theta))^3 \\ &= \cos(\theta)^3 + 3i\cos(\theta)^2\sin(\theta) - 3\cos(\theta)\sin(\theta)^2 - i\sin(\theta)^3 \\ &= \left[\cos(\theta)^3 - 3\cos(\theta)\sin(\theta)^2\right] + i\left[3\cos^2(\theta)\sin(\theta) - \sin^3(\theta)\right]\end{aligned}$$

En passant aux parties réelles, on obtient l'expression suivante :

$$\cos(3\theta) = \cos(\theta)^3 - 3\cos(\theta)\sin(\theta)^2.$$

3. Racines n-èmes

3.1. Racines n-èmes de l'unité. —

Définition 2.26 – Racine n-ème de l'unité

Soient n un élément de $\mathbb{N}^\star$ et z un nombre complexe.
On dit que z est une racine n-ème de l'unité lorsque

$$z^n = 1.$$

Exemple 2.27. —

- Les nombres 1 et (-1) sont des racines 2-èmes de l'unité.
- Le nombre i n'est pas une racine 2-ème de l'unité, puisque $i^2 = (-1) \neq 1$. En revanche, c'est une racine quatrième de l'unité, puisque $i^4 = (i^2)^2 = 1$.

***Exemple 2.28* (Racines deuxièmes de l'unité : 1 et (-1)).** —

- Les nombres 1 et -1 ont pour carré 1, donc ce sont des racines deuxièmes de l'unité.
- Ce sont les seules racines deuxièmes de l'unité. En effet, si z est un nombre complexe vérifiant $z^2 = 1$ et si l'on considère une écriture $z = |z|e^{i\theta}$ de z sous forme trigonométrique, alors on a $|z|^2 e^{2i\theta} = 1$, ce qui prouve que $|z|^2 = 1$ et que 2θ est multiple entier de 2π. Comme $|z|$ est positif, cela signifie que $|z| = 1$; de plus on constate que θ est de la forme $k\pi$ avec $k \in \mathbb{Z}$. On a donc $z = e^{ik\pi} = (e^{i\pi})^k = (-1)^k$, et nous avons donc bien $z = 1$ ou $z = (-1)$.

***Exemple 2.29* (Racines troisièmes de l'unité : 1, j et j^2).** —

- Considérons le nombre $j = e^{\frac{2i\pi}{3}}$. On constate que $j^3 = e^{3\cdot\frac{2i\pi}{3}} = e^{2i\pi} = 1$, donc j^3 est une racine troisième de l'unité.
  C'est également le cas du nombre $j^2 = e^{4\cdot\frac{2i\pi}{3}}$, puisque $(j^2)^3 = (j^3)^2 = 1^2 = 1$.
  Nous connaissons donc au moins trois racines 3-èmes de l'unité, les nombres 1, j et j^2.

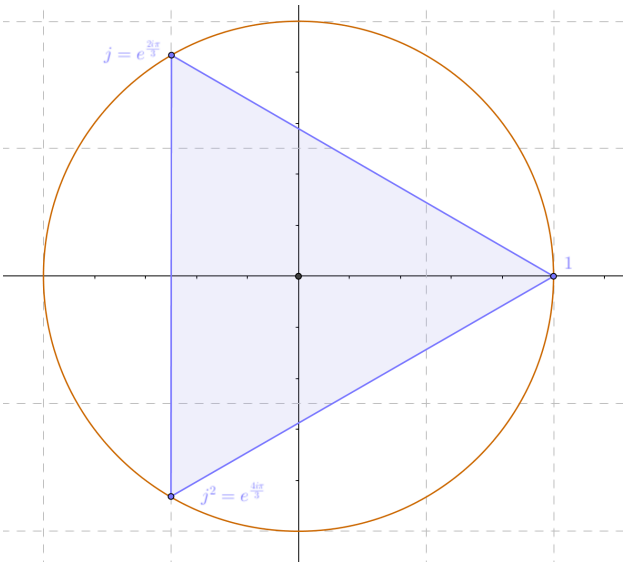

FIGURE 10. Les complexes 1, j et j^2 forment, sur le cercle unité, les trois sommets d'un triangle équilatéral.

- Il n'existe en fait *aucune autre racine troisième de l'unité que* 1, j *et* j^2. En effet, si z est un nombre complexe vérifiant $z^3 = 1$, alors si l'on écrit $z = re^{i\theta}$ avec $r > 0$ et $\theta \in \mathbb{R}$, on doit avoir $(re^{i\theta})^3 = 1$, autrement dit $r^3 e^{i3\theta} = 1$. Par unicité du module et de l'argument, on doit donc avoir

(a) $r^3 = 1$, d'où $r = 1$ (puisque r est un réel positif et que le seul réel positif de cube 1 est $r = 1$),
(b) $3\theta = 2k\pi$ pour un certain $k \in \mathbb{Z}$. Distinguons alors trois cas :
$\rightsquigarrow$ Si $k \equiv 0 \mod 3$, on peut alors écrire $k = 3\ell$ avec $\ell \in \mathbb{Z}$; on a alors $3\theta = 6\ell\pi$, d'où $z = re^{i\theta} = re^{2\ell i\pi} = 1$.
$\rightsquigarrow$ Si $k = 1 \mod 3$ on peut écrire $k = 3\ell + 1$ avec $\ell \in \mathbb{Z}$; on a alors $3\theta = 6\ell\pi + 2\pi$, d'où $z = re^{i\theta} = re^{\frac{2i\pi}{3}+2\ell i\pi} = j$.
$\rightsquigarrow$Enfin, si $k = 1 \mod 3$ on a alors on peut alors écrire $k = 3\ell + 2$ avec $\ell \in \mathbb{Z}$; on a alors $3\theta = 6\ell\pi + 4\pi$, d'où $z = re^{i\theta} = re^{\frac{4i\pi}{3}+2\ell i\pi} = j^2$.
Dans tous les cas, on constate que z est égal soit à 1, soit à j, soit à j^2.

Plus généralement, on constate qu'il y a toujours n racines n-èmes de l'unité, et que la liste n'en est pas difficile à retenir :

Proposition 2.30 – Liste des racines n-èmes de l'unité

L'ensemble des racines n-èmes de l'unité est

$$\mathbb{U}_n = \left\{1, \quad e^{\frac{2i\pi}{n}}, \quad e^{2\cdot\frac{2i\pi}{n}}, \quad e^{3\cdot\frac{2ik\pi}{n}}, \ldots, \quad e^{(n-1)\cdot\frac{2i\pi}{n}}\right\}$$

Il comporte exactement n éléments.

Si on note $\omega = e^{\frac{2i\pi}{n}}$, alors on a l'égalité suivante :

$$\mathbb{U}_n = \left\{\omega^k, \quad k \in [\![0, n-1]\!]\right\}.$$

***Exemple 2.31* (Racines quatrièmes et cinquièmes de l'unité).** — D'après la proposition ci-dessus :

- Il y a exactement 4 racines quatrièmes de l'unité : il s'agit des nombres 1, $e^{\frac{2i\pi}{4}}$, $e^{\frac{4i\pi}{4}}$ et $e^{\frac{6i\pi}{4}}$. On remarque que $e^{\frac{2i\pi}{4}} = e^{i\frac{\pi}{2}} = i$, tandis que les racines suivantes sont $e^{\frac{4i\pi}{4}} = (e^{\frac{2i\pi}{4}})^2 = i^2 = -1$ et $e^{\frac{6i\pi}{4}} = (e^{\frac{2i\pi}{4}})^4 = i^3 = -i$.
  En conclusion, les racines quatrièmes de l'unité sont 1, i, -1 et $-i$.
- Il y a exactement 5 racines cinquièmes de l'unité : il s'agit des nombres 1, $e^{\frac{2i\pi}{5}}$, $e^{\frac{4i\pi}{5}}$, $e^{\frac{6i\pi}{5}}$ et $e^{\frac{8i\pi}{4}}$. Si on note $\omega = e^{\frac{2i\pi}{5}}$, ces nombres sont égaux à 1, ω, ω^2, ω^3 et ω^3.

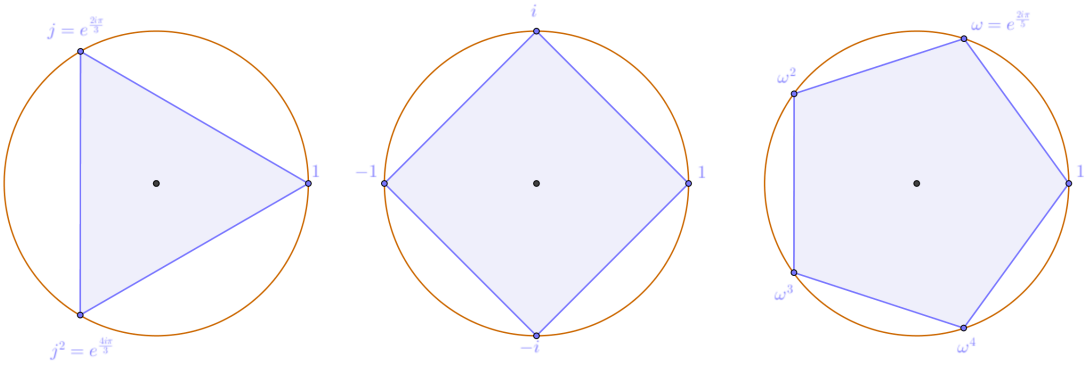

FIGURE 11. Si n est donné, les racines n-èmes de l'unité forment les sommets d'un polygône régulier à n côtés inscrit dans le cercle unité : un triangle pour $n = 3$, un carré pour $n = 4$, un pentagone pour $n = 5$, etc.

Proposition 2.32 – La somme des racines n-èmes de l'unité vaut 0

Fixons $n \in \mathbb{N}^\star$ et notons $\omega = e^{\frac{2i\pi}{n}}$. On a l'égalité $\sum_{k=0}^{n-1} \omega^k = 0$.

Démonstration. — La quantité $\sum_{k=0}^{n-1} \omega^k = 0$ est la somme des n premiers termes d'une suite géométrique de premier terme 1 et de raison ω : elle est donc égale à $\frac{1-\omega^n}{1-\omega}$. Or $\omega^n = 1$, puisque ω est une racine n-ème de l'unité. On trouve donc $\sum_{k=0}^{n-1} \omega^k = 0$. □

***Exemple 2.33* (Cas $n = 3$).** — Pour les racines troisièmes de l'unité, la proposition ci-dessus exprime l'égalité $1 + j + j^2 = 0$.

3.2. Racines n-èmes d'un nombre complexe quelconque. —

Définition 2.34 – Racine n-ème de a pour $a \neq 1$

Soient n un élément de $\mathbb{N}^\star$ et a un nombre complexe.
On dit qu'un nombre complexe z est une racine n-ème de a lorsque

$$z^n = a.$$

Exemple 2.35. — Si $a = i = e^{i\frac{\pi}{2}}$, alors le nombre $e^{i\frac{\pi}{8}}$ est une racine quatrième de a. C'est également le cas des nombres $-e^{i\frac{\pi}{8}}$ et $ie^{i\frac{\pi}{8}}$.

Attention aux notations ! — Si $n \geq 2$ et si $a \neq 0$, il y a toujours *plusieurs* racines n-èmes de a ; si a est un nombre complexe, l'écriture $\sqrt[n]{a}$ ou $a^{1/n}$ n'a donc pas de sens (y compris si $n = 2$: écrire $\sqrt{7-i}$ n'a pas de sens !). On parlera toujours d'*une* racine n-ème ou *des* racines n-èmes, mais le vocabulaire "la racine n-ème" est donc à proscrire quand il s'agit de nombres complexes.

La notation $\sqrt[n]{\cdot}$ que vous connaissez est spécifique au cas réel : si a est un nombre *réel positif*, nous verrons que parmi les racines n-èmes de a dans $\mathbb{C}$, il y en a une et une seule qui appartient à $\mathbb{R}^+$; c'est elle qu'on note habituellement $\sqrt[n]{a}$, mais cette notation est *spécifique au cas réel*.

Proposition 2.36 – Si on en connaît une, on connaît toutes les autres

Soit a un nombre complexe non nul. Soit $n \in \mathbb{N}^\star$.
Supposons connu un nombre complexe u_0 donnant une racine n-ème de a.
L'ensemble des racines n-èmes de a est alors

$$\left\{u_0 \cdot e^{\frac{2ik\pi}{n}}, \quad k \in [\![0, n-1]\!]\right\} = \left\{u_0, \quad u_0 e^{\frac{2i\pi}{n}}, \quad u_0 e^{\frac{4i\pi}{n}}, \ldots, \quad u_0 e^{\frac{2i(n-1)\pi}{n}}\right\}.$$

Il comporte exactement n éléments.

Démonstration. — Si u_0 est un nombre complexe vérifiant $u_0^n = a$ et si a est non-nul, on ne peut pas avoir $u_0 = 0$ (sinon on aurait $a = u_0^n = 0$).

En remarquant que pour $z \in \mathbb{C}$, on a $z^n = a$ si et seulement si on a $z^n = u_0^n$, et en divisant par u_0^n, on constate que $z^n = a$ équivaut à $\left(\frac{z}{u_0}\right)^n = 1$. Pour que cette équation soit vérifiée, il faut et il suffit que $\frac{z}{u_0}$ soit une racine n-ème de l'unité. La proposition 2.30 permet alors de conclure. □

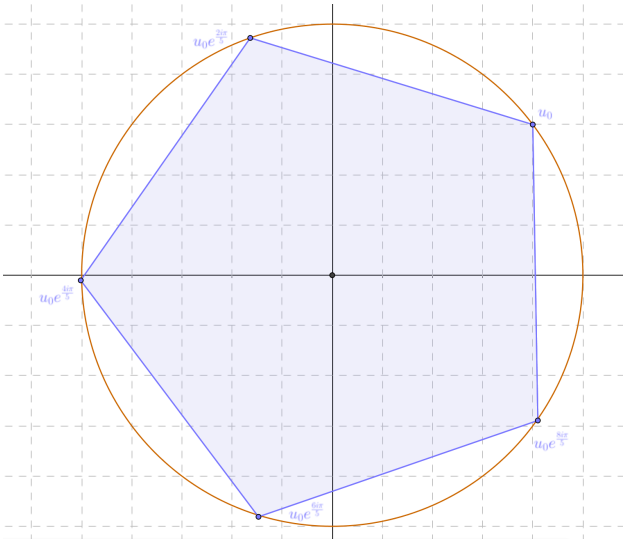

FIGURE 12. Si $a \neq 0$ et $n \geq 2$ sont donnés, les racines n-èmes de a forment les sommets d'un polygône régulier à n côtés inscrit dans le cercle de centre 0 et de rayon $|u_0|$. La figure représente les racines cinquièmes de a dans le cas où une racine particulière est $u_0 = 4 + 3i$

Proposition 2.37 – Trouver une racine n-ème de a à l'aide du module et de l'argument

Soit a un nombre complexe non nul. Si $a = re^{i\theta}$ est une écriture de a sous forme trigonométrique (avec $r > 0$ et $\theta \in \mathbb{R}$), alors le nombre

$$u_0 = r^{1/n} e^{i\theta/n}$$

est une racine n-ème de a.

Démonstration. — Nous avons bien $\left(r^{1/n} e^{i\theta/n}\right)^n = re^{i\theta} = a$. □

***Exemple 2.38* (racines n-èmes de -1).** — Fixons $n \in \mathbb{N}^\star$ et cherchons les nombres z vérifiant $z^n = (-1)$. Utilisons la proposition 2.37 et partons du fait que $-1 = e^{i\pi}$: on constate alors que le nombre $e^{i\frac{\pi}{n}}$ est une racine n-ème de -1. D'après la proposition 2.36, la liste des racines n-èmes de (-1) est donc

$$e^{\frac{i\pi}{n}}, \quad e^{i\frac{\pi}{n}+i\frac{2\pi}{n}}, \quad e^{i\frac{\pi}{n}+i\frac{4\pi}{n}}, \quad \ldots, \quad e^{i\frac{\pi}{n}+i\frac{2(n-1)\pi}{n}}.$$

3.3. Solutions d'une équation du second degré à coefficients complexes. —

Proposition 2.39 – Résolution d'une équation du second degré dans $\mathbb{C}$

Soient a, b, c trois nombres complexes avec $a \neq 0$.
Considérons le nombre complexe $\Delta = b^2 - 4ac$.

- *Si $\Delta = 0$, alors il existe un et un seul nombre complexe z vérifiant l'équation $az^2 + bz + c = 0$: il est donné par $z = \frac{-b}{2a}$.*
- *Si $\Delta \neq 0$, alors l'équation $az^2 + bz + c = 0$, d'inconnue $z \in \mathbb{C}$, admet exactement deux solutions : si l'on note ρ l'une des deux racines carrées du nombre complexe Δ, alors les deux solutions de $az^2 + bz + c = 0$ sont $\frac{-b-\rho}{2a}$ et $\frac{-b+\rho}{2a}$.*

Exemple 2.40. — L'équation $z^2 + 2iz + i$, d'inconnue $z \in \mathbb{C}$, est une équation du second degré de discriminant $\Delta = (2i)^2 - 4i = -4 - 4i = 4(-1 - i)$.

En constatant que $-1 - i = \sqrt{2}\left(\cos(\pi + \frac{\pi}{4}) + \sin(\pi + \frac{\pi}{4})\right)$, on obtient une forme trigonométrique du discriminant : on a $\Delta = 4\sqrt{2}e^{i\frac{5\pi}{4}}$.

L'équation initiale a donc deux solutions distinctes dans $\mathbb{C}$: il s'agit des nombres complexes $\frac{-2i-\sqrt{4\sqrt{2}e^{i\frac{5\pi}{8}}}}{2} = -i - 2^{\frac{5}{4}}e^{i\frac{5\pi}{8}}$ et $\frac{-2i+\sqrt{4\sqrt{2}e^{i\frac{5\pi}{8}}}}{2} = i + 2^{\frac{5}{4}}e^{i\frac{5\pi}{8}}$.

Attention : ne pas abuser des notations Δ, a, b, c ! — Rappelons qu'il n'est pas possible, dans un texte mathématique correctement rédigé, d'utiliser des symboles sans avoir introduit leur signification. Par exemple, si vous voulez résoudre l'équation $z^2 + 3z + 5$, d'inconnue $z \in \mathbb{C}$, et si vous écrivez « $\Delta = b^2 - 4ac = 9 - 20 = -11$, donc $z_1 = \frac{-3-i\sqrt{11}}{2}$ et $z_2 = \frac{-3+i\sqrt{11}}{2}$ », vous utilisez de nombreux symboles sans les introduire, et c'est mal. Si vous introduisez le symbole Δ, il faut expliquer qu'il s'agit du discriminant, par exemple en écrivant « le discriminant de cette équation est donné par $\Delta = \ldots$; de même, il est le plus souvent inutile de donner des noms aux coefficients et de parler de "a", "b", "c" si ce n'est pas nécessaire à la compréhension de votre texte.

4. Transformations du plan complexe

4.1. Interprétation géométrique des applications de $\mathbb{C}$ dans $\mathbb{C}$. — Considérons deux ensembles E et F et supposons que E et F soient tous les deux des parties de $\mathbb{C}$. Considérons une application $f : E \to F$.

Si z est un point de E, alors z peut être représenté géométriquement un point du plan. Si l'on observe $f(z)$, qui est un point de F, donc de $\mathbb{C}$, on obtient un autre point du plan. On peut donc voir f comme une *transformation géométrique*.

4.2. Translations, rotations, homothéties. —

Définition 2.41 – Translation dans $\mathbb{C}$

Soit $f : \mathbb{C} \to \mathbb{C}$ une application. On dit que f est une translation s'il existe un nombre complexe $v \in \mathbb{C}$ vérifiant

$$\forall z \in \mathbb{C},\ f(z) = z + v.$$

On dit alors que f est la translation de vecteur v.

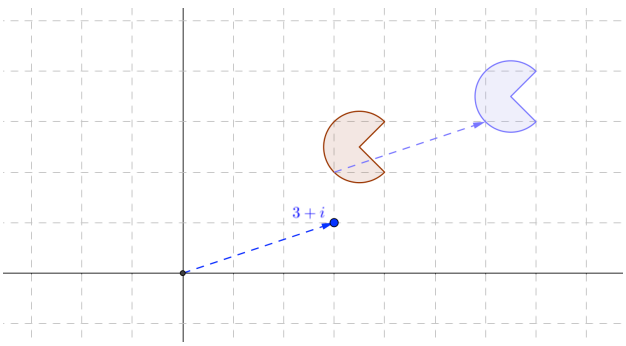

FIGURE 13. Un pac-man et son image par la translation de vecteur $3 + i$.

Définition 2.42 – Rotation autour de l'origine, puis rotation autour d'un point z_0

Soit $f : \mathbb{C} \to \mathbb{C}$ une application. Soit θ un nombre réel fixé.

1. Lorsque f est l'application $z \mapsto e^{i\theta} z$, on dit que f est la *rotation d'angle θ autour de l'origine.*
2. Lorsqu'il existe un nombre complexe z_0 tel que f soit l'application $z \mapsto e^{i\theta}(z - z_0) + z_0$, on dit que f est la *rotation d'angle θ autour de z_0.*

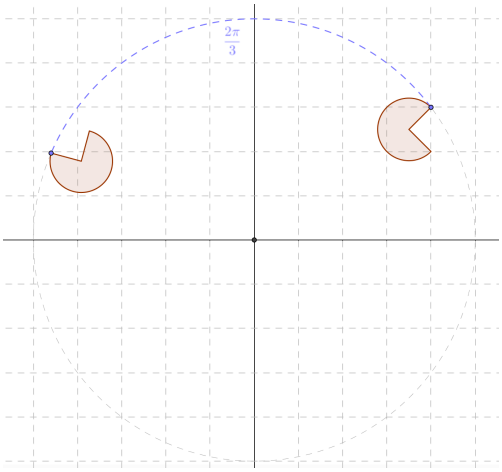

FIGURE 14. Le pac-man de la figure 13 et son image par la rotation d'angle $\frac{2\pi}{3}$ autour de l'origine.

Définition 2.43 – Homothétie de centre l'origine, puis homothétie de centre un point z_0

Soit $f : \mathbb{C} \to \mathbb{C}$ une application. Soit $k \in \mathbb{R}^\star$.

1. Lorsque f est l'application $z \mapsto kz$, on dit que f est l' *homothétie de centre 0 et de rapport k.*.
2. Lorsqu'il existe un nombre complexe z_0 tel que f soit l'application $z \mapsto k(z - z_0) + z_0$, on dit que f est l'*homothétie de centre z_0 et de rapport k.*

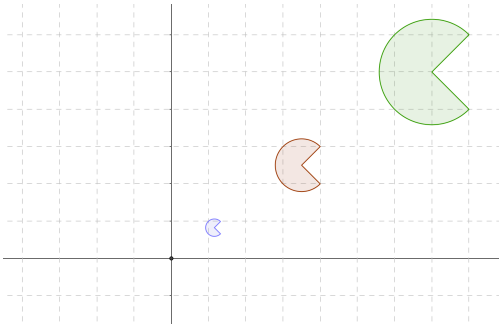

FIGURE 15. Le pac-man de la figure 13 et, en vert, son image par l'homothétie de centre 0 et de rapport 2, en bleu, son image par l'homothétie de centre 0 et de rapport $\frac{1}{3}$. .

4.3. Similitudes directes. —

Définition 2.44 – Similitude directe dans $\mathbb{C}$

Soit $f : \mathbb{C} \to \mathbb{C}$ une application. On dit que f est une *similitude directe* lorsqu'il existe un nombre complexe $a \neq 0$ et un nombre complexe b vérifiant :

$$\forall z \in \mathbb{C},\ f(z) = az + b.$$

Proposition 2.45 – Similitude directe dans $\mathbb{C}$

Soient a et b deux nombres complexes et $f : \mathbb{C} \to \mathbb{C}$ l'application $z \mapsto az + b$.

1. *Si $a = 1$, alors f est la translation de vecteur b.*
2. *Si $a \neq 1$, alors il existe un unique point $z_0 \in \mathbb{C}$ vérifiant $f(z_0) = z_0$. De plus, toujours sous l'hypothèse $a \neq 1$:*
   - *S'il existe $\theta \in (\mathbb{R} \setminus 2\pi\mathbb{Z})$ tel que $a = e^{i\theta}$, alors f est la rotation d'angle θ autour de z_0.*
   - *Si $a \in (\mathbb{R} \setminus \{0, 1\})$, alors f est l'homothétie de rapport a et de centre z_0.*

Exercices du chapitre 2

Forme algébrique et forme trigonométrique. —

Exercice 2.1. — ★☆☆

Trouver la forme algébrique de chacun des nombres complexes suivants :

$$z_1 = \frac{3+6i}{3-4i}, \qquad z_2 = \left(\frac{1+i}{2-i}\right)^2, \qquad z_3 = \frac{2+5i}{1-i} + \frac{2-5i}{1+i}.$$

Rappel : « trouver la forme algébrique de z » signifie « écrire z sous la forme $x+iy$ avec $x \in \mathbb{R}$ et $y \in \mathbb{R}$ ».

Exercice 2.2. — ★☆☆

1. Trouver une forme trigonométrique de chacun des nombres complexes suivants :

$$z_1 = 1+i, \qquad z_2 = 1+i\sqrt{3}, \qquad z_3 = \sqrt{3}+i, \qquad z_4 = \frac{1+i\sqrt{3}}{\sqrt{3}-i}.$$

   Rappel : « trouver une forme trigonométrique de z » signifie « écrire z sous la forme $re^{i\theta}$ avec $r \in \mathbb{R}_+^\star$ et $\theta \in \mathbb{R}$ ».
2. Soit $\alpha \in \mathbb{R}$. Trouver une forme trigonométrique de $e^{e^{i\alpha}}$.

Exercice 2.3. — ★☆☆

Déterminer une forme trigonométrique de $\frac{1+i}{1-i}$, puis calculer $\left(\frac{1+i}{1-i}\right)^{32}$.

Exercice 2.4. — ★★☆

1. Soit $\theta \in]0, \frac{\pi}{2}[$. Trouver une forme trigonométrique de $1+e^{i\theta}$.
2. Soit $\theta \in]-\pi, \pi[$. Trouver une forme trigonométrique de $1+e^{i\theta}$.
3. Soit θ un réel. Déterminer une forme trigonométrique de $1-\cos(\theta)+i\sin(\theta)$.
4. Soit θ un réel. Déterminer une forme trigonométrique de $e^{i\theta}+e^{2i\theta}$.
5. Soit θ un réel. Déterminer une forme trigonométrique de $\frac{\tan(\theta)-i}{\tan(\theta)+i}$.

Exercices divers sur : module, argument, conjugué. —

Exercice 2.5. — ★☆☆

Soit $z = 1+i(1+\sqrt{2})$. Calculer z^2, en déduire une forme trigonométrique de z^2 puis de z.

Exercice 2.6. — ★☆☆

Dans cet exercice, on fixe $\theta \in \mathbb{R}$, on rappelle la notation $j = e^{\frac{2i\pi}{3}}$ et on définit

$$a = \frac{1+i\sqrt{3}}{1+i}, \qquad b = 1+j \qquad \text{et} \qquad c = \frac{1+i\tan(\theta)}{1-i\tan(\theta)}.$$

Soit $n \in \mathbb{N}^\star$. Calculer a^n, b^n et c^n.

Exercice 2.7. — ★★☆

Soient a et b deux nombres complexes. On suppose que a et b sont de module 1 et que $ab \neq -1$. Démontrer que le nombre $\frac{a+b}{1+ab}$ est réel.

Exercice 2.8. — ★★☆

Soit z un nombre complexe de module 1. Montrer qu'il existe un réel t vérifiant

$$z = \frac{1+it}{1-it}.$$

Indication : commencer par faire l'exercice 2.5(c).

Exercice 2.9. — ★☆☆

Soient z et z' deux nombres complexes. Démontrer l'égalité du parallélogramme :

$$|z+z'|^2+|z-z'|^2=2\left(|z|+|z'|^2\right).$$

Pourquoi cette égalité est-elle ainsi nommée ?

Exercice 2.10. — ★★☆

Déterminer tous les nombres complexes $z\in\mathbb{C}^\star$ tels que les points d'affixes z, $\frac{1}{z}$ et $(1-z)$ soient situés sur un même cercle centré à l'origine.

Applications des nombres complexes à la trigonométrie. —

Exercice 2.11. — ★☆☆

1. Soit $\theta\in\mathbb{R}$. Exprimer $\cos(3\theta)$ et $\sin(3\theta)$ à l'aide des puissances de $\cos(\theta)$ et $\sin(\theta)$.
2. Soit $\theta\in\mathbb{R}$. Exprimer $\cos(\theta)^3$ à l'aide de $\sin(3\theta)$ et $\sin(\theta)$.

Exercice 2.12. — ★★☆

1. Soit $\theta\in\mathbb{R}$. Exprimer $\sin(6\theta)\cos(8\theta)$ à l'aide des puissances de $\cos(\theta)$ et $\sin(\theta)$.
2. Soit $\theta\in\mathbb{R}$. Exprimer $\sin(\theta)^4\cos(\theta)^5$ à l'aide des quantités $\cos(\theta),\cos(2\theta),\dots,\cos(5\theta),\sin(\theta),\dots,\sin(5\theta)$.

(*les résultats ne sont pas très jolis et les calculs peuvent être longs*).

Exercice 2.13. — ★★☆

1. Pour tout réel θ, exprimer $\cos(5\theta)$ à l'aide des puissances de $\cos(\theta)$ et $\sin(\theta)$.
2. En utilisant le fait que $\cos\left(5\frac{\pi}{10}\right)=0$, trouver la valeur de $\cos(\frac{\pi}{10})$.
3. Déterminer une forme trigonométrique du nombre complexe $\sqrt{10+2\sqrt{5}}+i(1-\sqrt{5})$.

Exercice 2.14. — ★★☆

Dans tout l'exercice, on fixe $\theta\in\mathbb{R}$ et $n\in\mathbb{N}^\star$.

1. Calculer $a=\sum_{k=0}^{n}\cos(k\theta)$ et $b=\sum_{k=0}^{n}\sin(k\theta)$ *(on calculera $a+ib$).*
2. Calculer $c=\sum_{k=-n}^{n}e^{i\theta}$.
3. Calculer $\sum_{k=0}^{n}\binom{n}{k}\cos(k\theta)$.

Exercice 2.15. — ★☆☆

Dans cet exercice, on fixe $\theta\in\mathbb{R}$ et $n\in\mathbb{N}^\star$. Calculer les deux quantités suivantes :

$$a=\sum_{k=1}^{n}\cos\left(\theta+\frac{2k\pi}{n}\right)\qquad\text{et}\qquad b=\sum_{k=1}^{n}\cos\left(\theta+\frac{2k\pi}{n}\right).$$

Racines n-èmes. —

Exercice 2.16. — ★☆☆

1. Calculer les racines carrées des nombres complexe $z=\frac{1}{2}+i\frac{\sqrt{3}}{2}$ et $z'=8(i+\sqrt{3})$.
2. Calculer les racines carrées des nombres complexe $z=-2+2i\sqrt{3}$ et $z'=9i$.

Exercice 2.17. — ★☆☆

On fixe deux réels a et b. Calculer les racines carrées du nombre complexe $z=4ab+2i\left(a^2-b^2\right)$.

Exercice 2.18. — ★☆☆

1. Calculer les racines troisièmes du nombre $1+i$.
2. Calculer les racines cinquièmes du nombre $z=1-i\sqrt{3}$.
3. Calculer les racines quatrièmes du nombre -4.

Exercice 2.19. — ★☆☆
Soit $n \in \mathbb{N}^\star$. Déterminer tous les nombres réels x qui vérifient l'équation

$$(x+i)^n = (x-i)^n.$$

Exercice 2.20. — ★★☆

1. Montrer qu'un nombre complexe z vérifie $|1+iz| = |1-iz|$ si et seulement si on a $z \in \mathbb{R}$.
2. On fixe un réel a et on considère dans $\mathbb{C}$
$$\left(\frac{1+iz}{1-iz}\right)^n = \frac{1+ia}{1-ia} \qquad (\star)$$
d'inconnue $z \in \mathbb{C}$.
   (a) Montrer, sans les calculer, que les solutions de cette équation sont réelles.
   (b) En écrivant $a = \tan(\alpha)$ et $z = \tan(\theta)$ avec α et θ réels, résoudre l'équation $(\star)$.
3. Calculer les racines cubiques de $\frac{\sqrt{3}+i}{\sqrt{3}-i}$

Autour des équations du second degré. —

Exercice 2.21. — ★★☆
On considère l'équation suivante, d'inconnue $z \in \mathbb{C}$:

$$z^2 + (1-i\sqrt{3})z - (1+i\sqrt{3}) = 0. \qquad (\star)$$

1. Montrer que $(\star)$ a deux solutions distinctes, qu'on exprimera en fonction de $a = \frac{\sqrt{3}+i}{2}$ et $b = \frac{-1+i\sqrt{3}}{2}$.
2. Trouver une forme trigonométrique de chacune des deux solutions de $(\star)$.
3. En déduire les valeurs de $\cos(\frac{5\pi}{12})$, $\sin(\frac{5\pi}{12})$, $\cos(\frac{11\pi}{12})$ et $\sin(\frac{11\pi}{12})$.

Exercice 2.22. — ★☆☆
Résoudre l'équation suivante, d'inconnue $z \in \mathbb{C}$:

$$z^2 - 2iz - 1 + 2i = 0.$$

Exercice 2.23. — ★☆☆
Résoudre l'équation suivante, d'inconnue $z \in \mathbb{C}$:

$$z^6 + z^3 + 1 = 0.$$

Exercice 2.24. — ★★☆
On fixe un réel θ. Résoudre l'équation suivante, d'inconnue $z \in \mathbb{C}$:

$$z^2 - 2e^{i\theta}z + 2i\sin(\theta)e^{i\theta} = 0.$$

Transformations du plan complexe. —

Exercice 2.25. — ★☆☆
On considère l'application

$$\begin{array}{rccc} f & : & \mathbb{C} & \to & \mathbb{C} \\ & & z & \mapsto & z^2. \end{array}$$

1. L'application f est-elle injective ? Surjective ?
2. On note $\mathbb{H}$ le demi-plan $\{z \in \mathbb{C} \ / \ \mathfrak{Im}(z) > 0\}$. Montrer que la restriction $f|_{\mathbb{H}}$ est injective, mais pas surjective.
3. Trouver un ensemble $A \subset \mathbb{C}$ tel que la restriction $f|_A : A \to \mathbb{C}$ soit bijective.

Exercice 2.26. — ★★☆
On considère l'application

$$\begin{array}{rccc} f & : & \mathbb{C}\setminus\{i\} & \to & \mathbb{C}\setminus\{1\} \\ & & z & \mapsto & \frac{z+i}{z-i}. \end{array}$$

1. Montrer que f est bijective et déterminer f^{-1}.
2. Déterminer les points fixes de f, c'est-à-dire les nombres complexes z vérifiant $f(z) = z$.
3. On note $\mathbb{U}$ l'ensemble des nombres complexes de module 1.
   (a) Montrer que $f(\mathbb{R}) \subset \mathbb{U}$.
   (b) Montrer que $f(\mathbb{R}) = \mathbb{U}$.

Exercice 2.27. — ★★☆
On considère l'application

$$\begin{array}{rccc} f & : & \mathbb{C}^\star & \to & \mathbb{C}^\star \\ & & z & \mapsto & \frac{2}{(\overline{z})}. \end{array}$$

1. Montrer que $f \circ f = \mathrm{id}_{\mathbb{C}^\star}$ (on dit que f est *involutive*).
2. L'application f est-elle bijective, et si oui, quelle est l'application f^{-1} ?.
3. Soit R un réel strictement positif et $\mathcal{C}$ le cercle $\{z \in \mathbb{C} \ / \ |z| = R\}$.
   Déterminer l'ensemble $f(\mathcal{C})$.
4. Quel est l'ensemble des points fixes de f ?

Exercice 2.28. — ★★☆
Dans tout l'exercice, on note $\mathbb{H}$ le demi-plan $\{z \in \mathbb{C} \ / \ \mathfrak{Im}(z) > 0\}$. On fixe des nombres *entiers relatifs* a, b, c et d vérifiant $ad - bc = 1$. On considère

$$\begin{array}{rccc} f & : & \mathbb{H} & \to & \mathbb{C} \\ & & z & \mapsto & \frac{az+b}{cz+d}. \end{array}$$

1. Montrer que $f(\mathbb{H}) \subset \mathbb{H}$.
2. Montrer que $f(\mathbb{H}) = \mathbb{H}$.
3. Montrer que f induit une bijection de $\mathbb{H}$ sur $\mathbb{H}$ dont on précisera la bijection réciproque.

Exercice 2.29. — ★★☆
Soit $f : \mathbb{C} \to \mathbb{C}$ l'application définie par la formule suivante : si z est un nombre et si $z = x + iy$ est son écriture sous forme algébrique, alors

$$f(z) = \frac{1}{2}\left(e^{-y}e^{ix} + e^{y}e^{-ix}\right).$$

1. Quelle est la fonction $f|_{\mathbb{R}}$?
2. Soit $z \in \mathbb{C}$. Montrer que $f(z + 2\pi) = f(z)$, que $f(-z) = f(z)$ et que $f(2z) = 2(f(z))^2 - 1$.
3. L'application f est-elle injective ?
4. Déterminer l'ensemble $f^{-1}(\{0\})$.

Exercice 2.30. — ★★☆
On considère l'application

$$\begin{array}{rccc} f & : & \mathbb{C}^\star & \to & \mathbb{C} \\ & & z & \mapsto & \frac{1}{2}\left(z + \frac{1}{z}\right). \end{array}$$

1. L'application f est-elle injective ? surjective ?
2. Déterminer l'ensemble $f^{-1}(\{i\})$.
3. Déterminer l'image directe par f du cercle unité $\mathbb{U}$.

4. On note $\mathbb{E}$ le complémentaire dans $\mathbb{C}$ du segment $[-1,1]$, et $\mathbb{D}$ l'ensemble $\{z \in \mathbb{C} \ / \ 0 < |z| < 1\}$.
   Montrer qu'on obtient une application bien définie en posant
$$\begin{array}{rccc} g & : & \mathbb{D} & \to & \mathbb{E} \\ & & z & \mapsto & f(z). \end{array}$$
5. Montrer que g est bijective (pour $a \in \mathbb{C}$ fixé, on pourra étudier dans $\mathbb{C}^\star$ l'équation $f(z) = a$, et remarquer qu'on connaît le produit de ses racines).

CHAPITRE 3

POLYNÔMES

1. Définitions élémentaires et degré

1.1. Notion abstraite de polynôme. —

Vous connaissez la notion de *fonction polynomiale de* $\mathbb{R}$ *dans* $\mathbb{R}$: si n est un entier naturel, si $a_0, a_1, \ldots, a_n$ sont des nombres réels fixés, et si l'on définit

$$\begin{array}{rccc} f & : & \mathbb{R} & \to & \mathbb{R} \\ & & x & \mapsto & a_n x^n + a_{n-1}x^{n-1} + \cdots + a_2 x^2 + a_1 x + a_0, \end{array}$$

alors on dit que f est une *fonction polynomiale*; si de plus on a $a_n \neq 0$, on dit que f est de *degré* n. Par exemple, la fonction

$$\begin{array}{rccc} g & : & \mathbb{R} & \to & \mathbb{R} \\ & & x & \mapsto & x^4 + 3x^2 + 7, \end{array}$$

est une fonction polynomiale de degré 4.

Partant de nombres $a_0, a_1, \ldots, a_n$, on peut également définir une fonction de $\mathbb{C}$ dans $\mathbb{C}$. Par exemple, en utilisant le même « jeu de coefficients » que ci-dessus, on peut considérer la fonction

$$\begin{array}{rccc} h & : & \mathbb{C} & \to & \mathbb{C} \\ & & z & \mapsto & z^4 + 3z^2 + 7. \end{array}$$

Il est important de comprendre que les fonctions g et h sont *différentes* : non seulement elles n'ont pas la même définition (puisque leurs ensembles de départ et d'arrivée sont différents), mais elles n'ont pas du tout les mêmes propriétés : g est croissante sur $\mathbb{R}$ alors que cela n'a pas de sens de se demander si h est croissante ; g n'est pas surjective (puisque $g(\mathbb{R}) \subset [7, +\infty[$) alors que h est surjective (savez-vous le montrer... ?).

Dans ce chapitre, nous définissons des *objets abstraits*, les « polynômes », qui ne sont pas des fonctions, mais qui peuvent donner lieu à des fonctions des deux types ci-dessus. Nous considérerons par exemple l'objet abstrait

$$P(X) = X^4 + 3X^2 + 7$$

qui n'est pas la même chose que g, qui n'est pas non plus la même chose que h, mais qui pourra « s'incarner », selon les besoins, en la fonction g ci-dessus ou en la fonction h ci-dessus.

Nous nous permettrons au passage d'augmenter un peu le niveau de généralité et d'inclure à ce chapitre l'étude des polynômes « à coefficients complexes », comme

$$Q(X) = X^5 + (3 - 7i)X^4 + (-8 + 2i)X^2 + (3 + 4i).$$

Théorème et définition 3.1 – Polynômes à coefficients complexes

Il existe un ensemble E, muni de deux opérations $+$ et $\times$, vérifiant les cinq propriétés suivantes :

(1) l'ensemble E contient $\mathbb{C}$ et contient de plus un élément noté X, appelé *l'indéterminée*,

(2) **Forme générale des éléments.**
Si P est un élément de E, alors P peut s'écrire sous la forme

$$a_0 + a_1X + a_2X + \cdots + a_nX^n \qquad (\star)$$

où n est un élément de $\mathbb{N}$ et où $a_0, \ldots, a_n$ sont des nombres complexes.

(3) **Unicité de l'écriture.**
Si P est un élément *non nul* de E, alors il existe une *unique* écriture de P sous la forme $(\star)$ si l'on impose $a_n \neq 0$.

Si $P = a_0 + a_1X + \cdots + a_nX^n$ et $Q = b_0 + b_1X + \cdots + a_pX^p$ sont deux éléments de E, alors

(4) **Définition de l'addition.**
L'élément $P + Q$ est défini par la formule suivante :

$$P + Q = (a_0 + b_0) + (a_1 + b_1)X + \cdots + (a_d + b_d)X^d$$

où $d = \max(n, p)$ et où on a défini $a_k = 0$ pour tout $k > n$, et $b_k = 0$ pour tout $k > p$.

(5) **Définition de la multiplication.**
L'élément PQ de E est défini par la formule suivante :

$$PQ = \sum_{k=0}^{n+p} \left(\sum_{j=0}^{k} a_j b_{k-j} \right) X^k$$

autrement dit,

$$PQ = (a_0b_0)+(a_0b_1+b_1a_0)X+\cdots+(a_0b_2+a_1b_1+a_2b_0)X^2+\cdots+(a_{n-1}b_p+a_nb_{p-1})X^{n+p-1}(a_nb_p)X^{n+p}.$$

Exemple 3.2. — La définition ci-dessus peut sembler très abstraite. Cette apparence est trompeuse et la manipulation des polynômes est tout aussi concrète que celle des expressions dont vous avez l'habitude.

- Les propriétés (1) et (2) permettent de donner un sens à des expressions comme :

$$\begin{gathered}\text{Notons } P \text{ le polynôme } X^3 - 1\\ \text{et notons } Q \text{ le polynôme } X^4 - X^3 + 7X^2.\end{gathered}$$

- Pour comprendre l'unicité de l'écriture (propriété (3)), relevons simplement, par exemple, que si l'on considère quatre réels a, b, c, d, alors on a l'équivalence

$$\left(X^3 - 1 = aX^4 + bX^3 + cX + d\right) \iff (a = 0,\ b = 1,\ c = 0 \text{ et } d = -1)\,.$$

- Quant aux définitions de l'addition et de la multiplication (propriétés (4) et (5)), malgré la lourdeur des notations générales, elles sont transparentes sur des exemples : ainsi,

$$\begin{gathered}P + Q \text{ est le polynôme } X^4 + 7X^2 - 1\\ PQ \text{ est le polynôme } X^7 - X^6 + 7X^5 - X^4 + X^3 - 7X^2.\end{gathered}$$

  Les propriétés (4) et (5) disent donc simplement qu'on peut faire la multiplication « comme d'habitude ».

1.2. Notations et vocabulaire de base. —

Si P est un élément de E, il sera parfois psychologiquement utile d'écrire $P(X)$ plutôt que P. C'est uniquement une question de notations, et il n'y a pas de différence entre P et $P(X)$. Par exemple, les

expressions « Si P est le polynôme X^2+1 » et « Si $P(X)=X^2+1$ » ont exactement la même signification.

Introduisons maintenant quelques éléments de terminologie.

- ***Coefficients d'un polynôme.*** Si P est un polynôme et si l'on a $P(X)=a_0+a_1X+\cdots+a_nX^n$ avec $n\in\mathbb{N}$ et $(a_0,a_1,\ldots,a_n)\in\mathbb{C}^{n+1}$, alors les nombres complexes $a_0,a_1,\ldots,a_n$ sont appelés les *coefficients* de P.

  On dit que P est à *coefficients réels* lorsque les coefficients $a_0,a_1,\ldots,a_n$ sont tous réels. Par exemple, le polynôme
  $$P(X)=7X^5+\sqrt{3}X^4-X^2+2$$
  est à coefficients réels, mais le polynôme $7X^5+\sqrt{3}X^4-(1+i)X^2+2$ ne l'est pas.

- ***Degré d'un polynôme.*** Si P est un polynôme non nul et si l'on a $P(X)=a_0+a_1X+\cdots+a_nX^n$ avec $n\in\mathbb{N}$ et $(a_0,a_1,\ldots,a_n)\in\mathbb{C}^{n+1}$, on appelle *degré* de P, et on note $\deg(P)$, le plus grand entier k tel que $a_k\neq 0$. Par exemple, on a l'égalité
  $$\deg(X^3-7X+2)=3.$$
  Autre exemple : si l'on fixe un nombre réel α et si l'on considère le polynôme
  $$P(X)=(\alpha^2-4)X^5+(\alpha-2)X^2+3X+1,$$
  alors P est de degré 5 lorsque α n'est égal ni à 2 ni à -2, il est de degré 2 lorsque $\alpha=-2$, et il est de degré 1 lorsque $\alpha=2$.

  On adopte de plus la convention suivante :

  Le degré du polynôme nul est $-\infty$.

- ***Coefficient constant.*** Si P est un polynôme et si l'on a $P(X)=a_0+a_1X+\cdots+a_nX^n$ avec $n\in\mathbb{N}$ et $(a_0,a_1,\ldots,a_n)\in\mathbb{C}^{n+1}$, alors le nombre a_0 est appelé le *coefficient constant de P*. Par exemple, le coefficient constant de X^2-7 est (-7).

- ***Coefficient dominant.*** Si P est un polynôme non nul et si l'on note $n=\deg(P)$ et si l'on écrit
  $$P(X)=a_nX^n+a_{n-1}X^{n-1}+\cdots+a_1X+a_0$$
  avec $(a_1,\ldots,a_n)\in\mathbb{C}^{n+1}$ et $a_n\neq 0$, alors le coefficient a_n est appelé *coefficient dominant de P*. Par exemple, le polynôme $7X^3+8X^2-1$ est de coefficient dominant 7, tandis que le polynôme α^2X^3-4X+4 a pour coefficient dominant α^2 lorsque $\alpha\neq 0$ et -4 lorsque $\alpha=0$.

- ***Ensembles $\mathbb{C}[X]$ et $\mathbb{R}[X]$.*** On note $\mathbb{C}[X]$ l'ensemble des polynômes à coefficients complexes et $\mathbb{R}[X]$ l'ensemble des polynômes à coefficients réels. On a par exemple :
  $$(3X^8-X^3+4)\in\mathbb{R}[X],\qquad (3X^8-(2+3i)X^2)\in\mathbb{C}[X],\qquad \text{et aussi}\quad (3X^8-X^3+4)\in\mathbb{C}[X].$$
  Pour la dernière affirmation, remarquons que les coefficients réels peuvent aussi être vus comme des coefficients complexes, puisque $\mathbb{R}\subset\mathbb{C}$. Il sera souvent commode de pouvoir le considérer au choix comme élément de $\mathbb{R}[X]$ ou comme élément de $\mathbb{C}[X]$.

- ***Polynôme unitaire.*** On dit qu'un polynôme (non nul) est *unitaire* si son coefficient dominant est égal à 1. Par exemple, le polynôme $X-3-2X+1$ est unitaire, mais le polynôme $3X^2-5$ n'est pas unitaire.

- ***Polynômes associés.*** Soient P et Q deux polynômes. On dit que P et Q sont *associés* s'il existe un nombre complexe non nul $\alpha\in\mathbb{C}^\star$ vérifiant $P(X)=\alpha Q(X)$. Par exemple, les polynômes $P(X)=$

$3X^6 - 9X^5 + 21X + 15$ et $Q(X) = X^6 - 3X^5 + 7X + 5$ sont associés, alors que les polynômes $P(X)$ et $R(X) = 3X^6 + 17X^5 + 21X + 15$ ne sont pas associés.

On remarquera que si P est un polynôme non nul et si a est son coefficient dominant, alors le polynôme $\frac{P(X)}{a}$ est unitaire et qu'il est associé à P.

Nombre $P(x)$ pour $x \in \mathbb{C}$. — Si P est un polynôme à coefficients réels ou complexes et si l'on écrit $P(X) = a_0 + a_1X + \cdots + a_nX^n$, alors dès que x est un nombre complexe, on peut obtenir le nombre $P(x) = a_0 + a_1x + \cdots + a_nx^n$. Dans ce cours, nous veillerons toujours à bien distinguer les deux notations :

$$P(X)$$

désigne l'objet abstrait P, où X est « l'indéterminée », tandis que

$$P(x), \text{ lorsque } x \in \mathbb{R} \text{ ou } \mathbb{C}$$

désigne le nombre réel ou complexe obtenu en « évaluant P en x ».

Notation « $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ » . — Si l'on considère les deux polynômes suivants :

$$P(X) = X^4 - (3 + 5i)X^3 + X + 1 \quad \text{et} \quad Q(X) = X^3 + X,$$

alors

- P est un polynôme à coefficients complexes, et pas à coefficients réels. Il appartient donc à $\mathbb{C}[X]$, mais pas à $\mathbb{R}[X]$.
- Q est à coefficients réels, mais si on le souhaite, on peut aussi le voir comme un polynôme à coefficients complexes : il appartient donc à la fois à $\mathbb{R}[X]$ et à $\mathbb{C}[X]$.

Il pourra arriver qu'on veuille étudier Q comme élément de $\mathbb{R}[X]$ (par exemple en étudiant les variations de la fonction $x \mapsto Q(x)$, de $\mathbb{R}$ dans $\mathbb{R}$, pour constater qu'elle ne s'annule qu'en zéro), ou comme élément de $\mathbb{C}[X]$ (par exemple pour écrire que $Q(i) = 0$). Certains théorèmes n'auront de sens que si l'on travaille dans $\mathbb{R}[X]$, d'autres n'auront de sens que si l'on travaille dans $\mathbb{C}[X]$.

Dans ce cours, on écrira parfois

$$\text{« Fixons } \mathbb{K} = \mathbb{R} \text{ ou } \mathbb{K} = \mathbb{C} \text{ et travaillons dans } \mathbb{K}[X] \text{ »}$$

pour fixer le cadre avant de travailler. Par exemple, considérons le texte suivant :

$$\text{Fixons } \mathbb{K} = \mathbb{R} \text{ ou } \mathbb{K} = \mathbb{C}.$$
$$\text{Si } P \in \mathbb{K}[X] \text{ et si } P(0) = 0, \text{ alors il existe un polynôme } Q \in \mathbb{K}[X] \text{ vérifiant } P(X) = XQ(X) \text{ »}$$

est à comprendre comme suit :

- si P est un polynôme à coefficients complexes vérifiant $P(0) = 0$, alors il existe un polynôme Q à coefficients complexes vérifiant $P(X) = XQ(X)$,
- et si de plus P est à coefficients réels, alors il existe en fait un polynôme Q à coefficients *réels* vérifiant $P(X) = XQ(X)$.

1.3. Remarques sur la notion de degré. —

Proposition 3.3 – Degré d'une somme et d'un produit

Soient P et Q deux polynômes à coefficients réels ou complexes.

1. *On a toujours* $\deg(P + Q) \leq \max(\deg(P), \deg(Q))$.
   De plus, le seul cas où il n'y a pas d'égalité est celui où les polynômes P et Q ont le même degré et des coefficients dominants opposés.
2. *Si P et Q sont non-nuls, on a toujours* $\deg(PQ) = \deg(P) + \deg(Q)$.

Exemple 3.4. — Si $P(X) = X^2 - 1$ et $Q(X) = X^6 + X^4$, alors $(P+Q)(X) = X^6 + X^4 + X^2 - 1$ a pour degré 6, qui est bien le « plus grand des deux degrés », tandis que $(PQ)(X) = X^8 + X^6 - X^6 - X^4 = X^8 - X^4$, qui est bien de degré $8 = \deg(P) + \deg(Q)$.

Exemple 3.5. — Si $P(X) = X^6 - 1$ et $Q(X) = -X^6 + X^4$, alors $(P+Q)(X) = X^4 - 1$, et il est de degré 4 alors que P et Q sont de degré 6 : cette « chute du degré » s'explique par le fait que les coefficients dominants de P et Q sont opposés, donc que les termes en X^6 « disparaissent » lorsqu'on effectue la sommation.

***Remarque 3.6* (Autour du degré $-\infty$).** — Dans l'énoncé ci-dessus, pour tenir compte du cas où l'un des deux polynômes est nul, on adopte pour convention que $-\infty + n = -\infty$ pour tout entier naturel n.

La propriété 2. ci-dessus est ainsi compatible avec le fait que si l'un des deux polynômes P, Q est nul, alors PQ est nul aussi.

***Remarque 3.7* (Intégrité).** — Si P et Q sont deux polynômes et si $PQ = 0$, alors on a $P = 0$ ou $Q = 0$.

En effet, puisque $\deg(PQ) = \deg(P) + \deg(Q)$, lorsque $PQ = 0$ on a nécessairement $\deg(P) + \deg(Q) = -\infty$, ce qui n'est possible dans notre contexte que si $\deg(P) = -\infty$ ou $\deg(Q) = -\infty$.

Cette propriété est moins évidente qu'il n'y paraît : elle permet par exemple de voir que

Si $(X-3)P(X)$ est le polynôme nul, alors $P(X)$ est le polynôme nul

sans avoir besoin de se préoccuper de « ce qui se passe pour $P(3)$ ».

2. Division euclidienne et arithmétique des polynômes

2.1. Diviseurs d'un polynôme. —

Définition 3.8

Soient A et B deux polynômes à coefficients réels ou complexes.
On dit que B divise A, et on note $B|A$, lorsque

- B n'est pas nul
- et il existe un polynôme Q vérifiant : $A(X) = B(X)Q(X)$.

Exemple 3.9. — Le polynôme $X^3 - 1$ est divisible par le polynôme $X - 1$, puisqu'on a $X^3 - 1 = (X-1)(X^2 + X + 1)$.

Par contre, le polynôme $X^7 - 2X + 3$ n'est pas divisible par $X^4 - X$: si c'était le cas, il existerait un polynôme $Q(X)$ vérifiant : $(X^7 - 2X + 3) = (X^4 - X)Q(X)$; or, le coefficient constant du polynôme $(X^4 - X)Q(X)$ est nécessairement égal à 0, il est donc impossible que ce polynôme soit égal à $X^7 - 2X + 3$ (de coefficient constant 3).

Attention. — Dans la notion de divisibilité, on prendra garde au fait que les multiplications peuvent faire intervenir des nombres réels arbitraires. Cela peut donner lieu à quelques étrangetés : par exemple, 5 divise $X + 1$, puisque $X + 1 = 5\left(\frac{1}{5}X + \frac{1}{5}\right)$! De même, si l'on voit les nombres 7 et 2 comme des polynômes et si l'on s'intéresse à la notion de divisibilité dans l'ensemble $\mathbb{R}[X]$ des polynômes, plutôt que dans l'ensemble $\mathbb{Z}$, alors il faut bien convenir que 7 divise 2 dans $\mathbb{R}[X]$...

Proposition 3.10 – Divisibilité et degré

1. *Si A et B sont deux polynômes non nuls et si A divise B , alors* $\deg(A) \leq \deg(B)$.
2. *Si $A|B$ et $B|A$, alors il existe un scalaire $\alpha \neq 0$ vérifiant : $A = \alpha B$.*
3. *Si $A|B$ et* $\deg(A) = \deg(B)$, *alors il existe un scalaire $\alpha \neq 0$ vérifiant : $A = \alpha B$.*

Démonstration. — 1. S'il existe un polynôme Q vérifiant : $A(X) = B(X)Q(X)$, alors d'après la proposition 3.3, on a $\deg(A) = \deg(B) + \deg(Q)$, et puisque $Q \neq 0$, le degré de Q est un entier positif ou nul ; on a donc bien $\deg(A) \leq \deg(B)$.

2. D'après le point précédent, si $A|B$ et $B|A$, alors on a nécessairement $\deg(A) = \deg(B)$; de plus, si on écrit $A(X) = B(X)Q(X)$ pour tenir compte du fait que B divise A, alors on a $\deg(A) = \deg(B) + \deg(Q)$; comme nous venons de voir que $\deg(A) = \deg(B)$, on a $\deg(Q) = 0$; on peut donc écrire $Q(X) = \alpha$ où α est un scalaire non nul. Ainsi $A(X) = \alpha B(X)$: c'est ce qu'il fallait démontrer. □

Exemple 3.11. — Fixons $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$. Si P est un polynôme non nul de $\mathbb{K}[X]$ et si l'on note n le degré de P, alors on a l'équivalence suivante :

$$P \text{ est divisible par } X^n \iff \text{on peut écrire } P(X) = \alpha X^n \text{ avec } \alpha \in \mathbb{K}^\star.$$

2.2. Division euclidienne des polynômes. —

Théorème 3.12 – Existence et unicité de la division euclidienne

Fixons $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$. Soient A et B deux polynômes de $\mathbb{K}[X]$. On suppose $B \neq 0$. Il existe un unique couple (Q, R) de polynômes de $\mathbb{K}[X]$ vérifiant

$$A(X) = B(X)Q(X) + R(X) \quad \textit{et} \quad \deg(R) < \deg(B).$$

***Exemple 3.13* (Un exemple concret).** — Considérons les polynômes

$$A(X) = X^3 + X^2 + 1 \qquad \text{et} \qquad B(X) = X - 1.$$

La division euclidienne doit être de la forme $A(X) = B(X)Q(X) + R(X)$ où $Q(X)$ est un polynôme et où $R(X)$ est un polynôme de degré $< \deg(B) = 1$, donc un polynôme constant. En jouant sur les « termes de plus haut degré », on constate que

$$\begin{aligned} X^3 + X^2 + 1 &= (X-1)(X^2) + 2X^2 + 1 \\ &= (X-1)X^2 + (X-1)(2X) + 2X + 1 \\ &= (X-1)(X^2 + 2X) + 2(X-1) + 3 \\ &= (X-1)(X^2 + 2X + 2) + 3. \end{aligned}$$

Compte tenu de l'unicité dans le théorème ci-dessus, la dernière égalité prouve que le quotient est $X^2 + 2X + 2$ et le reste est 3.

***Exemple 3.14* (Un exemple plus théorique).** — Soit P un polynôme non constant à coefficients réels ; déterminons le reste de la division euclidienne de P par $X(X-1)(X+1)$. On sait que si l'on écrit la division euclidienne sous la forme

$$P(X) = Q(X)X(X-1)(X+1) + R(X) \qquad \text{avec } \deg(R) < \deg\left[X(X-1)(X+1)\right]$$

alors le polynôme R est de degré < 3, donc de la forme $aX^2 + bX + c$ avec a, b, c réels. On cherche donc a, b et c pour avoir

$$P(X) = X(X-1)(X+1)Q(X) + aX^2 + bX + c.$$

En observant ce que donne cette égalité en $x = 0$, on constate alors que $P(0) = 0 + c$, donc $c = P(0)$. En prenant $x = 1$, on obtient aussi $P(1) = a + b + c$, et en prenant $x = -1$, on obtient aussi $P(-1) = a - b + c$.

Ces trois renseignements, combinés, permettent de trouver a, b et c : on a $c = P(0)$, $P(1) - P(-1) = 2b$ et $P(1) + P(-1) = 2(a + c)$, d'où

$$a = \frac{P(1) + P(-1)}{2} - P(0), \quad b = \frac{P(1) - P(-1)}{2} \quad \text{et} \quad c = P(0).$$

Démonstration du théorème 3.12. —

- *Unicité sous réserve d'existence.* Supposons qu'il existe deux couples (Q_1, R_1) et (Q_2, R_2) de polynômes vérifiant

$$A(X) = B(X)Q_1(X) + R_1(X) \quad \text{et} \quad \deg(R_1) < \deg(B),$$
$$A(X) = B(X)Q_1(X) + R_2(X) \quad \text{et} \quad \deg(R_2) < \deg(B).$$

  On a alors l'égalité $BQ_1 + R_1 = BQ_2 + R_2$; en la réarrangeant, on trouve

$$R_1 - R_2 = B(Q_2 - Q_1).$$

  Examinons les degrés des deux membres de cette égalité :
  - Les polynômes R_1 et R_2 étant de degrés $< \deg(B)$, on a $\deg(R_1 - R_2) < \deg(B)$ (voir la proposition 3.3).
  - Par ailleurs, le polynôme $U = B(Q_2 - Q_1)$ est divisible par B ; si $Q_2 - Q_1 \neq 0$, on en déduit (par la proposition 3.10) que $\deg(B) \leq \deg(U)$.

  Ces deux remarques mènent à une absurdité, sauf si l'on a $Q_2 = Q_1$; c'est donc que $Q_2 = Q_1$, et on en déduit que $BQ_1 + R_1 = BQ_1 + R_2$, d'où $R_1 = R_2$.
- *Existence.* On suppose que B ne divise pas A, sinon le résultat à montrer est immédiat. Considérons alors l'ensemble suivant :

$$E = \{\deg(A - BQ), \quad Q \in \mathbb{K}[X]\}.$$

  Comme B ne divise pas A, on n'a jamais $A - BQ = 0$, donc jamais $\deg(A - BQ) = -\infty$. L'ensemble E est donc formé d'entiers naturels. De plus, E est non vide (il contient par exemple $\deg(A-0) = \deg(A)$).

  L'ensemble E admet donc un plus petit élément. Si nous notons d le plus petit élément de E, nous savons alors qu'il existe un polynôme $Q \in \mathbb{K}[X]$ tel que le polynôme $R = A - BQ$ ait pour degré d.

  On a alors $A = BQ + R$, avec $\deg(R) = d$; si nous parvenons à montrer que $d < \deg(B)$, nous aurons gagné.

  Pour démontrer que $d < \deg(B)$, nous allons raisonner par l'absurde. Supposons donc $d \geq \deg(B)$. Montrons qu'en retirant à $R = A - BQ$ un polynôme de la forme BU, avec $U \in \mathbb{K}[X]$, on peut obtenir un polynôme $R' = A - BQ - BU$ dont le degré est $< d$; cela contredira le fait que d est le plus petit élément de E.

  Pour cela, isolons dans R le « monôme de plus haut degré » : écrivons $R(X)$ sous la forme

$$R(X) = \rho X^d + \tilde{R}$$

  où ρ est le coefficient dominant de R (donc un scalaire non nul) et où $\tilde{R}(X)$ est un polynôme de degré $< d$. De même, écrivons

$$B(X) = bX^{\deg(B)} + \tilde{B},$$

  où b est le coefficient dominant de B (donc un scalaire non nul) et où $\tilde{B}$ est un polynôme de degré $< \deg(B)$.

  Choisissons alors

$$U(X) = \frac{\rho}{b} X^{d - \deg(B)}.$$

  On a alors $B(X)U(X) = \rho X^d + \frac{\rho}{b}\tilde{B}X^{d-\deg(B)}$; on remarque que le degré de $\tilde{B}X^{d-\deg(B)}$ est égal à $\deg(\tilde{B}) + d - \deg(B) = d - (\deg(B) - \deg(\tilde{B})) < d$. Les polynômes R et $B(X)U(X)$ ont donc le même degré (égal à d), et le même coefficient dominant (égal à ρ). On en conclut que

$$R' = R - UB \text{ est de degré } < d,$$

  mais $R' = R - UB = (A - BQ) - BU = A - B(U + Q)$; en revenant à la définition de l'ensemble E, on constate que $\deg(R')$ est un élément de E, et qu'il est $< d$, alors que d est censé être le plus petit élément de E. On obtient ainsi la contradiction espérée, ce qui achève notre démonstration.

□

***Remarque 3.15* (Divisibilité sur $\mathbb{C}$ et divisibilité sur $\mathbb{R}$).** — Soient A et B deux polynômes non nuls à coefficients réels. On peut se demander si A divise B. Le plus naturel est d'imaginer que la question est la suivante :

$$\text{Existe-t-il un polynôme } U \text{ à coefficients } \textit{réels} \text{ tels que } A(X)U(X) = B(X) \,?$$

Mais si l'on voit A et B comme deux éléments de $\mathbb{C}[X]$, alors la question est :

$$\text{Existe-t-il un polynôme } U \text{ à coefficients } \textit{complexes} \text{ tels que } A(X)U(X) = B(X) \,?$$

On pourrait penser que la réponse peut être "oui" pour la deuxième question sans que ce soit le cas pour la première (c'est-à-dire qu'on puisse trouver U à coefficients complexes, mais pas réels, pour avoir $B = AU$). Cependant, grâce à l'unicité dans le théorème de division euclidienne, on constate que les deux questions sont en fait équivalentes :

> Si A et B sont à coefficients réels et si A divise B dans $\mathbb{C}[X]$, alors A divise B dans $\mathbb{R}[X]$.

En effet, si nous écrivons la division euclidienne de A par B dans $\mathbb{C}[X]$, le reste est nul et la division est de la forme $A(X) = B(X)U(X) + 0$ avec $U(X) \in \mathbb{C}[X]$. Si nous notons à présent $\bar{A}$, $\bar{B}$ et $\bar{U}$ les polynômes dont les coefficients sont les conjugués des coefficients de A, B et U, alors en passant au conjugué dans l'égalité $A = BU$, on obtient $\bar{A} = \bar{B}\bar{U}$; mais A et B étant supposés à coefficients réels, on a $\bar{A} = A$ et $\bar{B} = B$, d'où $A = B\bar{U}$. Comme on a aussi $A = BU$, on en déduit $BU = B\bar{U}$, puis $B(U - \bar{U}) = 0$. Comme le polynôme B est non-nul, on en déduit $U = \bar{U}$ (voir la remarque 3.7) ; on a donc en fait $U \in \mathbb{R}[X]$ et $A = BU$, donc B divise A dans $\mathbb{R}[X]$.

2.3. Notion de PGCD, algorithme d'Euclide, relation de Bézout... —

Théorème et définition 3.16 – PGCD de deux polynômes

Fixons $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$.
Soient A et B deux polynômes non nuls de $\mathbb{K}[X]$. Il existe un unique polynôme $P \in \mathbb{K}[X]$ vérifiant les trois conditions suivantes :

(i) P est un diviseur commun de A et B,
(ii) P est unitaire,
(iii) et pour tout $Q \in \mathbb{K}[X]$, si Q est unitaire et divise à la fois A et B, alors $\deg(Q) \leq \deg(P)$.

Ce polynôme est appelé plus grand diviseur commun de A et B, noté $\text{PGCD}(A, B)$.

Remarque 3.17. — Le fait d'imposer que le coefficient dominant du PGCD soit égal à 1 est important pour garantir l'unicité dans le théorème ci-dessus. C'est ce qui permet que la notation $\text{PGCD}(A, B)$ ne comporte pas d'ambiguité.

Démonstration de l'existence et de l'unicité du PGCD. — Considérons l'ensemble

$$E = \{\deg(P), \quad P \text{ est un polynôme unitaire qui est un diviseur commun de } A \text{ et } B\}\,.$$

L'ensemble E est une partie de $\mathbb{N}$ (puisque le polynôme 0 ne divise aucun polynôme non nul, donc le degré $-\infty$ n'est pas à prendre dans E). De plus, il est non vide puisqu'il contient le polynôme 1, et il est majoré par exemple par $\deg(A)$, puisqu'un diviseur de A ne peut pas avoir un degré $> \deg(A)$. Ainsi, E admet un plus grand élément. Notons d ce plus grand élément. Il existe alors un polynôme unitaire P, diviseur commun de A et B, tel que $\deg(P) = d$. Le polynôme P vérifie les conditions (i) et (ii) du théorème ; ce qu'il reste à prouver, c'est la propriété (iii). Mais si Q est un polynôme unitaire qui est un diviseur commun de A et B, le degré de Q fournit un élément de E, et par définition du plus petit élément de E, on a alors $\deg(Q) \leq d$. Or, $d = \deg(P)$; on obtient donc $\deg(Q) \leq \deg(P)$, comme espéré. Notre démonstration est donc achevée. □

À partir du théorème de division euclidienne et de la définition du PGCD, on peut formuler des théorèmes sur les polynômes qui sont les "décalques" directs des théorèmes montrés sur les entiers dans la pré-rentrée. Nous donnons les énoncés et des exemples, et reléguons les démonstrations (qui sont calquées sur celles pour les entiers) à la fin du paragraphe.

Proposition 3.18 – Relation de Bézout pour le PGCD

Fixons $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$. Soient A et B deux polynômes non nuls de $\mathbb{K}[X]$. Notons D le PGCD *de A et de B.*
Il existe deux polynômes U et V de $\mathbb{K}[X]$ vérifiant

$$D(X) = A(X)U(X) + B(X)V(X).$$

Corollaire 3.19 – Propriété universelle du PGCD

Soient A et B deux polynômes non nuls.

(i) Si Q est un polynôme qui divise à la fois A et B, alors Q est un diviseur du PGCD *de A et B.*
(ii) De plus, si D est un polynôme unitaire diviseur commun de A et B et si tout diviseur commun de A et B est nécessairement un diviseur de D, alors $D = \text{PGCD}(A, B)$.

Proposition 3.20 – Propriété des restes pour le PGCD

Soient A et B deux polynômes non nuls. Soit R le reste de A dans la division euclidienne par B. Si R n'est pas nul, alors on a l'égalité

$$\text{PGCD}(A, B) = \text{PGCD}(B, R).$$

Proposition 3.21 – Algorithme d'Euclide pour le calcul du PGCD

Soient A et B deux polynômes non nuls.
On travaille avec deux variables U et V et on procède comme suit.

- Initialisation : *on commence avec $U = A$ et $V = B$.*
- Boucle : *tant que $\beta \neq 0$, on effectue les deux substitutions*

$$U \leftarrow V$$
$$V \leftarrow \text{ le reste de la division euclidienne de } U \text{ par } V$$

- Conclusion : *on s'arrête quand $V = 0$.*

*Le contenu de U, **divisé par son coefficient dominant pour devenir unitaire**, donne alors le* PGCD *de A et B.*

Exemple 3.22. — Considérons les polynômes $A(X) = X^4 + 1$ et $B(X) = X^3 + 1$. En écrivant

$$\begin{aligned} X^4 + 1 &= (X^3 + 1)X - X + 1 \\ X^3 + 1 &= (-X + 1)(-X^2 - X - 1) + 2 \\ -X + 1 &= 2\left[\frac{1}{2}(-X + 1)\right] + 0, \end{aligned}$$

on constate que le dernier reste non nul est un polynôme constant. En le divisant par son coefficient dominant, on obtient A ; ainsi $\text{PGCD}(A, B) = 1$..

Les seules différences entre ces résultats et les théorèmes d'arithmétique des entiers viennent du fait que le PGCD doit être *unitaire* : c'est ce qui explique la subtilité énoncée en gras dans l'algorithme d'Euclide.

Démonstration de la relation de Bézout 3.18. — Imitons la démonstration effectuée pour les entiers et considérons l'ensemble de polynômes suivant :

$$\mathcal{I} = \{P \in \mathbb{K}[X] \ / \ P \text{ est non nul, unitaire, et il existe deux polynômes } U \text{ et } V \text{ de } \mathbb{K}[X] \text{ vérifiant } P = AU + BV\}.$$

Notre but est de montrer que le PGCD de A et B appartient à $\mathcal{I}$. Pour cela, nous adoptons une stratégie abstraite et considérons l'ensemble

$$E = \{\deg(P),\ P \in \mathcal{I}\}$$

des degrés des éléments de $\mathcal{I}$. Comme le polynôme nul n'est pas dans $\mathcal{I}$, l'ensemble E est une partie de $\mathbb{N}$; de plus, E est non vide (en effet, si l'on note α le coefficient dominant de A, le polynôme $P = \frac{1}{\alpha}A + 0B$ appartient à $\mathcal{I}$, et son degré est égal à $\deg(A)$).

C'est donc que E admet un plus petit élément ; notons-le d. Nous savons alors qu'il existe un polynôme unitaire P de degré d qui peut s'écrire sous la forme $AU + BV$, où U et V sont deux polynômes de $\mathbb{K}[X]$.

Notons D le PGCD de A et B. Quel est le rapport entre D et le polynôme P que nous venons d'introduire ?

- D'abord, on sait que $P = AU + BV$, et que A et B sont divisibles par D. On en déduit que D divise P ; en particulier $\deg(D) \leq \deg(P)$.
- Pour prouver que les degrés sont en fait égaux, effectuons la division euclidienne de A par P. On peut écrire $A = PQ + R$, où Q et R sont deux polynômes et $\deg(R) < \deg(P) = d$ On a alors

  $$R = A - PQ = A - Q(AU + BV) = A(1 - UQ) + BV;$$

  si R est non nul et si nous notons r le coefficient dominant de R, on constate alors que $\frac{R}{r}$ appartient à $\mathcal{I}$; mais c'est impossible puisque $\deg(R) < d$ et que d est le plus petit degré possible pour un élément de $\mathcal{I}$.

  De cela on conclut que $R = 0$, donc que A est multiple de P ; de même, on vérifie que B est multiple de P. Nous constatons donc que P est un diviseur commun de A et B ; on a donc $\deg(P) \leq \deg(D)$.

Pour résumer, nous avons nécessairement $\deg(P) = \deg(D)$, nous avons vu que D divise P ; par la proposition 3.10, c'est que D et P sont associés ; comme ils sont tous les deux unitaires, ils sont égaux. On a donc $D = AU + BV$, comme espéré.

□

Démonstration de la propriété universelle 3.19. — Notons Δ le PGCD de A et B.

- Si nous écrivons une relation de Bézout $\Delta = AU + BV$, on constate que si un polynôme C est un diviseur commun de A et B, on peut écrire $A = C\tilde{A}$ et $B = C\tilde{B}$ où $\tilde{A}$ et $\tilde{B}$ sont deux polynômes ; on a alors $\Delta = AU + BV = C\tilde{A}U + C\tilde{B}V = C(\tilde{A}U + \tilde{B}V)$, ce qui prouve que C divise Δ.
- Si D est un polynôme unitaire et si tout diviseur commun de A et B divise D, alors Δ divise D puisque Δ est un diviseur commun de A et B. De plus, par la première partie de la proposition, le fait que D soit un diviseur commun implique que D divise Δ. C'est donc que D et Δ sont associés ; comme ils sont tous les deux unitaires, ils sont égaux.

□

Démonstration de la propriété des restes 3.20 et justification de l'algorithme d'Euclide

Les démonstrations sont pour ainsi dire identiques à celles effectuées pour les entiers dans le polycopié de la pré-rentrée « raisonnement », page 63. □

Une fois énoncées ces propriétés, on peut formuler des analogues exacts de la plupart des théorèmes vus dans le chapitre d'arithmétique. Voici certains des énoncés correspondants, sans détailler les démonstrations qui sont identiques à celles des résultats déjà vus sur les entiers.

2.4. Polynômes premiers entre eux. —

Définition 3.23 – Polynômes premiers entre eux

Soient A et B deux polynômes non nuls. On dit que A et B sont premiers entre eux si le polynôme $\text{PGCD}(A, B)$ est égal à 1.

Proposition 3.24 – Théorème de Bézout pour deux polynômes premiers entre eux

Soient A et B deux polynômes non nuls. Les polynômes A et B sont premiers entre eux si et seulement s'il existe deux polynômes U et V tels que soit vérifiée la relation

$$A(X)U(X) + B(X)V(X) = 1.$$

Recherche d'une relation de Bézout. — Comme dans le cas des entiers, elle s'effectue en « remontant les calculs dans l'algorithme d'Euclide ». Par exemple, si l'on reprend $A(X) = X^4 + 1$ et $B(X) = X^3 + 1$, alors la recherche du PGCD s'effectue par l'algorithme d'Euclide : en écrivant

$$\begin{aligned} X^4 + 1 &= (X^3 + 1)X - X + 1 \\ X^3 + 1 &= (-X + 1)(-X^2 - X - 1) + 2 \\ -X + 1 &= 2\left[\frac{1}{2}(-X + 1)\right] + 0, \end{aligned}$$

on constate que le dernier reste non nul est un polynôme constant, donc que $\text{PGCD}(A, B) = 1$. Pour trouver deux polynômes U et V vérifiant $AU + BV = 1$, on remonte les calculs pour exprimer 1 comme suit :

$$\begin{aligned} 1 &= \frac{1}{2}\left(X^3 + 1 - (-X + 1)(-X^2 - X - 1)\right) \\ &= \frac{1}{2}\left(X^3 + 1 - \left[(X^4 + 1) - X(X^3 + 1)\right](-X^2 - X - 1)\right) \\ &= (X^4 + 1)(\tfrac{1}{2}X^2 + \tfrac{1}{2}X + \tfrac{1}{2}) + (X^3 + 1)(-\tfrac{1}{2}X^3 - \tfrac{1}{2}X^2 - \tfrac{1}{2}X + \tfrac{1}{2}). \end{aligned}$$

Proposition 3.25 – Lemme de Gauss

Soient A, B, C trois polynômes non nuls.
Si A divise (BC) et si A est premier avec B, alors A divise C.

Proposition 3.26 – Deux diviseurs premiers entre eux...

Soient A, B, C trois polynômes non nuls.
Si A et B divisent C et si A et B sont premiers entre eux, alors (AB) divise C.

Proposition 3.27 – Comment se ramener au cas de deux polynômes premiers entre eux

Soient A et B deux polynômes non nuls. Soit D le PGCD *de A et B. Il existe deux polynômes $\tilde{A}$ et $\tilde{B}$ vérifiant*

$$A = D\tilde{A}, \quad B = D\tilde{B} \quad \text{et} \quad \text{PGCD}(\tilde{A}, \tilde{B}) = 1.$$

2.5. Notion de PPCM. —

Théorème et définition 3.28 – PPCM de deux polynômes

Fixons $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$.
Soient A et B deux polynômes non nuls de $\mathbb{K}[X]$. Il existe un unique polynôme $P \in \mathbb{K}[X]$ vérifiant les trois conditions suivantes :

(i) P est un multiple commun de A et B,
(ii) P est unitaire,
(iii) et pour tout $Q \in \mathbb{K}[X]$, si Q est unitaire et divise à la fois A et B, alors $\deg(Q) \geq \deg(P)$.

Ce polynôme est noté $\text{PPCM}(A, B)$.

Proposition 3.29 – Propriété universelle du PPCM

Soient A et B deux polynômes non nuls.

(i) Si Q est un polynôme qui est multiple de A et de B, alors Q est un multiple du polynôme $\text{PPCM}(A, B)$.

(ii) De plus, si M est un polynôme unitaire et si tout multiple commun de A et B est nécessairement un diviseur de M, alors $M = \text{PGCD}(A, B)$.

Proposition 3.30 – Pour deux polynômes premiers entre eux, le ppcm, c'est le produit, à condition de normaliser

Soient A et B deux polynômes non nuls. Notons $dom(A)$ et $dom(B)$ leurs coefficients dominants. Si A et B sont premiers entre eux, alors on a l'égalité $\text{PPCM}(A, B) = \frac{1}{dom(A)dom(B)} AB$.

Corollaire 3.31 – Lien entre le PGCD et le PPCM

Soient A et B deux polynômes non nuls. Notons $dom(A)$ et $dom(B)$ leurs coefficients dominants. On a l'égalité suivante :

$$\text{PPCM}(A, B)\text{PGCD}(A, B) = \frac{1}{dom(A)dom(B)} AB.$$

3. Polynômes irréductibles ; théorème de factorisation

3.1. Définition et premières remarques. —

Si l'on se rappelle la définition, dans $\mathbb{Z}$, de « nombre premier », et si l'on cherche ce que peut être l'analogue pour les polynômes de la notion de nombre premier, on est conduit naturellement à la notion suivante.

Définition 3.32 – Polynôme irréductible dans $\mathbb{C}[X]$ ou $\mathbb{R}[X]$

Fixons $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$.
Soit P un polynôme. On dit que P est *irréductible dans* $\mathbb{K}[X]$ lorsque les deux conditions suivantes sont vérifiées :

- P n'est pas constant
- et les seuls diviseurs de P dans $\mathbb{K}[X]$ sont les polynômes de la forme $\alpha 1$ et αP, $\alpha \in \mathbb{K}^\star$.

On remarquera la subtilité liée aux difficultés que posent les constantes dans la notion de divisibilité (voir la remarque en haut de la page 44). À cause de cette subtilité, la définition impose *qu'aucun polynôme constant ne soit premier.*

Notre premier exemple de polynôme irréductible est le suivant :

Proposition 3.33 – Les polynômes de degré 1 sont irréductibles

Fixons $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$ et considérons deux éléments a et b de $\mathbb{K}$, avec $a \neq 0$.
Le polynôme $aX + b$ est irréductible dans $\mathbb{K}[X]$.

Démonstration. — Si Q est un diviseur de $P(X) = aX + b$ dans $\mathbb{K}[X]$, alors $\deg(Q)$ est un entier naturel vérifiant $\deg(Q) \leq \deg(P) = 1$, donc $\deg(Q) = 0$ ou $\deg(Q) = 1$.

Si $\deg(Q) = 0$, c'est que Q est un polynôme constant non nul, donc $Q(X) = \alpha 1$ où α est un élément de $\mathbb{K}^\star$.

Si $\deg(Q) = 1$, on a $Q|P$ et $\deg(Q) = \deg(P)$: d'après la troisième partie de la proposition 3.10, c'est que $Q = \alpha P$ avec $\alpha \in \mathbb{K}^\star$.

Nous constatons donc que les seuls diviseurs de P dans $\mathbb{K}[X]$ sont les diviseurs « triviaux », c'est-à-dire les polynômes de la forme $\alpha 1$ et αP, $\alpha \in \mathbb{K}^\star$. C'est bien que P est irréductible dans $\mathbb{K}[X]$. □

***Exemple 3.34* (Exemple fondamental : le cas de X^2+1).** — Considérons le polynôme $P(X) = X^2+1$.

- Voyons-le d'abord comme un élément de l'ensemble $\mathbb{R}[X]$ des polynômes à coefficients réels. Examinons alors les diviseurs de P dans $\mathbb{R}[X]$.

  Soit Q un diviseur de P dans $\mathbb{R}[X]$. Le degré de Q est un entier naturel inférieur ou égal à $\deg(P) = 2$, donc on a $\deg(Q) = 0$, 1 ou 2.

  Si $\deg(Q) = 0$, on a $Q(X) = \alpha 1$ avec $\alpha \in \mathbb{R}^\star$; si $\deg(Q) = 2$, en utilisant à nouveau la proposition 3.10.3, on a $Q = \alpha P$ avec $\alpha \in \mathbb{R}^\star$. Les seuls diviseurs de degré 0 ou 2 sont donc les diviseurs « triviaux ».

  Montrons à présent qu'il est impossible que Q soit de degré 1.

  Si c'était le cas, on pourrait écrire $Q(X) = aX + b$ avec a, b réels et $a \neq 0$. Par ailleurs, puisque Q est un diviseur de P, on aurait $P(X) = Q(X)U(X)$, et cela ne peut arriver que si U est aussi de

degré 1, donc de la forme $a'X + b'$ avec a', b' réels et $a' \neq 0$. Mais alors, on aurait

$$X^2 + 1 = (aX + b)(a'X + b') = (aa')X^2 + (ab' + a'b)X + bb'$$

d'où $aa' = 1$, $ab' + a'b = 0$ et $bb' = 1$.

On rappelle que a est non nul ; la première contrainte donnerait donc $a' = \frac{1}{a}$, tandis que la deuxième donnerait $b' = -\frac{a'b}{a} = -\frac{b}{a^2}$. On obtiendrait alors $bb' = -\frac{b^2}{a^2} \leq 0$, ce qui est impossible si $bb' = 1 > 0$.

On en conclut que $P(X) = X^2 + 1$ est irréductible dans $\mathbb{R}[X]$.

- Si nous voyons maintenant P comme un élément de l'ensemble $\mathbb{C}[X]$ des polynômes à coefficients complexes, alors l'égalité

$$X^2 + 1 = (X - i)(X + i)$$

montre que P n'est pas irréductible dans $\mathbb{C}[X]$.

En résumé, nous avons montré que

> Le polynôme $X^2 + 1$ est irréductible dans $\mathbb{R}[X]$, mais pas dans $\mathbb{C}[X]$.

Contrairement peut-être aux apparences issues de ces deux premiers exemples, la notion de polynôme irréductible de $\mathbb{R}[X]$ ou $\mathbb{C}[X]$ est *plus facile* à manier que la notion de nombre premier. La raison est que dans le cas des entiers, il est *difficile* de savoir si un nombre donné est premier, alors que dans le cas des polynômes, nous verrons plus loin dans ce cours que c'est très facile :

> La liste des polynômes irréductibles de $\mathbb{R}[X]$ et $\mathbb{C}[X]$ sera donnée au §4.4.

3.2. Théorème de factorisation, partie théorique. —

Dans ce paragraphe, on énonce un analogue pour les polynômes du théorème de « décomposition en produit de facteurs premiers » de $\mathbb{Z}$. Comme nous le verrons, une fois connue la liste des polynômes irréductibles de $\mathbb{R}[X]$ et $\mathbb{C}[X]$, le théorème deviendra très concret (et vous en verrez de nombreux exemples dans les exercices).

Proposition 3.35 – Tout polynôme de degré ≥ 1 admet au moins un diviseur irréductible

Fixons $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$. Si P est un polynôme de $\mathbb{K}[X]$ de degré supérieur ou égal à 1, alors il existe un polynôme irréductible de $\mathbb{K}[X]$ qui divise P dans $\mathbb{K}[X]$.

Démonstration. — On raisonne par récurrence forte sur le degré de P. Si P est de degré 1, il est lui-même irréductible d'après la proposition 3.33.

Fixons maintenant un entier $n \geq 1$, et supposons que pour tout $k \in \{1, \ldots, n\}$, les polynômes de degré k de $\mathbb{K}[X]$ admettent chacun au moins un diviseur irréductible dans $\mathbb{K}[X]$.

Soit maintenant P un polynôme de $\mathbb{K}[X]$ de degré $n + 1$. Distinguons deux cas :

- Si P est irréductible, alors il n'y a rien à prouver.
- Sinon, on peut écrire $P = P_1 P_2$ où P_1 et P_2 sont deux polynômes non constants de $\mathbb{K}[X]$. Dans ce cas, on a $\deg(P_1) \geq 1$ et $\deg(P_2) \geq 1$. Si nous notons $k = \deg(P_1)$, nous constatons donc que $k \in \{1, \ldots, n\}$. D'après l'hypothèse de récurrence, le polynôme P_1 admet au moins un diviseur irréductible dans $\mathbb{K}[X]$. Mais un tel diviseur est aussi un diviseur de P, ce qui établit le résultat à prouver.

□

Théorème 3.36 – Existence de la décomposition en produit de facteurs irréductibles

Fixons $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$. Si A est un polynôme de $\mathbb{K}[X]$ de degré supérieur ou égal à 1, alors il existe un entier $r \in \mathbb{N}^\star$, des polynômes $P_1, \ldots, P_r$ qui sont unitaires et irréductibles dans $\mathbb{K}[X]$, et des entiers $m_1, \ldots m_r$ de $\mathbb{N}^\star$, ainsi qu'une constante $c \in \mathbb{K}^\star$, vérifiant :

$$A(X) = cP_1(X)^{m_1}P_2(X)^{m_2}\cdots P_r(X)^{m_k}.$$

De plus, une telle écriture est unique à l'ordre près des facteurs.

Ce théorème est une conséquence du résultat précédent, appliqué plusieurs fois à partir d'un polynôme quelconque jusqu'à obtenir une factorisation comme dans le théorème. L'unicité demande un peu de soin pour être écrite, mais ne pose pas de difficulté conceptuelle majeure. Nous n'écrirons pas la démonstration.

Exemple 3.37. — Considérons le polynôme $A(X) = X^4 - 1$. En utilisant l'identité remarquable $X^4 - 1 = (X^2)^2 - 1^2 = (X^2 - 1)(X^2 + 1)$, on constate que

$$A(X) = (X - 1)(X + 1)(X^2 + 1).$$

Nous savons que les polynômes $X - 1$ et $X + 1$ sont irréductibles, puisqu'ils sont de degré 1 ; de plus, nous avons vu à l'exemple 3.34 que $X^2 + 1$ est irréductible dans $\mathbb{R}[X]$. L'écriture ci-dessus donne donc la décomposition en facteurs irréductibles de A dans $\mathbb{R}[X]$.

Par ailleurs, nous avons vu que $X^2 + 1$ n'est plus irréductible si l'on se place dans $\mathbb{C}[X]$; si l'on décide de travailler dans $\mathbb{C}[X]$, alors on peut poursuivre la factorisation et écrire

$$A(X) = (X - 1)(X + 1)(X - i)(X + i).$$

Cette écriture en produit de quatre facteurs de degré 1, donc irréductibles, donne la factorisation de A dans $\mathbb{C}[X]$.

On notera l'existence de deux factorisations pour P, l'une dans $\mathbb{R}[X]$ et l'autre dans $\mathbb{C}[X]$, et leur forme très particulière.

Dans la suite de ce cours, nous montrerons que le théorème de décomposition ci-dessus n'est, dans le cas le plus général, pas beaucoup plus abstrait que ce que nous avons rencontré sur cet exemple. Mais pour cela, nous devrons étudier soigneusement la notion de racine, dans $\mathbb{C}$ et dans $\mathbb{R}$.

À suivre au §4.4 : version concrète de la factorisation

4. Racines d'un polynôme

4.1. Définition et premières remarques. —

Définition 3.38

Soit $P(X)$ un polynôme de $\mathbb{K}[X]$. Soit a un élément de $\mathbb{K}$.
On dit que a est une racine de P lorsqu'on a $P(a) = 0$.

Exemple 3.39 **(Degré 1).** — Soit P un polynôme à coefficients réels ou complexes ; on peut donc écrire $P(X) = \alpha X + \beta$ où a et b sont deux nombres réels ou complexes et où $a \neq 0$.

Le nombre $\frac{-\beta}{\alpha}$ est une racine de P ; de plus, il s'agit de la seule racine de P, puisque l'équation $\alpha x + \beta$, d'inconnue $x \in \mathbb{C}$, admet $\frac{-\beta}{\alpha}$ pour unique solution.

***Exemple 3.40* (Cas de racines sur $\mathbb{C}$ mais pas sur $\mathbb{R}$).** — Le polynôme $P(X) = X^2 + 1$ n'a pas de racine réelle ; en revanche, $P(i) = P(-i) = 0$, donc les nombres i et $-i$ sont des racines de P dans $\mathbb{C}$.

***Remarque 3.41* (Coefficients réels et racines conjuguées).** — Considérons le polynôme

$$P(X) = X^5 + X^4 + X^3 + X^2 + X + 1.$$

Nous avons vu au chapitre sur les nombres complexes que la somme des racines cinquièmes de l'unité vaut 0 ; par conséquent, si $a = e^{i\frac{2\pi}{5}}$, alors $P(a) = 0$.

On peut remarquer que a est un nombre complexe, non réel, alors que P est à coefficients réels. Si l'on fait intervenir la conjugaison complexe, l'égalité

$$a^5 + a^4 + a^3 + a^2 + a + 1 = 0$$

a pour conséquence

$$\bar{a}^5 + \bar{a}^4 + \bar{a}^3 + \bar{a}^2 + \bar{a} + 1 = 0.$$

On connaît donc une autre racine de P : le nombre $\bar{a} = e^{-i\frac{2\pi}{5}}$.

Plus généralement, nous nous servirons souvent de la remarque suivante :

> Si P est à coefficients réels et si a est une racine complexe de P, alors $\bar{a}$ est aussi une racine de P.

***Remarque 3.42* (Diviseurs et racines).** — Soient P et Q deux polynômes. Supposons que Q divise P. On peut alors écrire $P(X) = Q(X)U(X)$, où U est un polynôme.

Si a est une racine de Q, alors $Q(a) = 0$, et on constate avec l'égalité $P = QU$ que $P(a) = Q(a)U(a) = 0$.

Nous constatons donc que si a est racine d'un polynôme qui divise P, alors a est en fait racine de P.

La réciproque est fausse (par exemple, le polynôme $X^3 - 1$ est divisible par $X^2 + X + 1$, puisque $X^3 - 1 = (X - 1)(X^2 + X + 1)$. Mais le nombre 1 est racine de $X^3 - 1$, mais pas du diviseur $X^2 + X + 1$.

Proposition 3.43 – Racine et divisibilité par $X - a$

Soit $P(X)$ un polynôme de $\mathbb{K}[X]$. Soit a un élément de $\mathbb{K}$. Les deux assertions suivantes sont équivalentes :

- *a est une racine de P*
- *le polynôme $(X - a)$ divise P.*

Démonstration. —

- Supposons que a soit une racine de P. Pour montrer que $(X - a)$ divise P dans $\mathbb{K}[X]$, écrivons la division euclidienne de P par $(X - a)$ et montrons que le reste est nul.

  Introduisons donc l'unique couple (Q, R) de polynômes de $\mathbb{K}[X]$ vérifiant :

  $$P(X) = (X - a)Q(X) + R(X) \text{ et } \deg(R) < \deg(X - a).$$

  Comme $X - a$ est de degré 1, le reste R doit être constant. Mais en observant ce que donne l'égalité ci-dessus en a, on constate que $P(a) = (a - a)Q(a) + R(a) = 0 + R(a)$; comme $P(a) = 0$, on en déduit que $R(a) = 0$; puisque R est constant, c'est que $R = 0$. On a donc $P(X) = (X - a)Q(X)$ où Q est un polynôme de $\mathbb{K}[X]$. Nous avons bien montré que $(X - a)$ divise P.
- Réciproquement, si $(X - a)$ divise P dans $\mathbb{K}[X]$, alors on peut écrire $P(X) = (X - a)Q(X)$ où Q est un polynôme de $\mathbb{K}[X]$. On a alors $P(a) = (a - a)Q(a) = 0$, donc a est racine de P.

□

Proposition 3.44 – Un polynôme de degré n admet au plus n racines

Soit P un polynôme non nul à coefficients réels ou complexes. Notons $n = \deg(P)$.
Le nombre de racines distinctes de P dans $\mathbb{C}$ est inférieur ou égal à n.

Exemple 3.45. — En observant le polynôme $X^7 + 5X^3 - X^2 + 8$, il n'est pas facile de savoir combien il a exactement de racines (que ce soit dans $\mathbb{R}$ ou dans $\mathbb{C}$). Mais on sait qu'il ne peut pas admettre plus de 7 racines dans $\mathbb{C}$ (et a fortiori pas plus de 7 racines dans $\mathbb{R}$)

Démonstration. —

- D'après la proposition précédente, si a est une racine de P, alors $(X - a)$ divise P.
- Remarquons maintenant que si a et b sont deux racines distinctes de P dans $\mathbb{C}$, alors le polynôme $(X - a)(X - b)$ divise P. Puisque nous savons déjà que $(X - a)$ et $(X - b)$ divisent P, il nous suffit de montrer que les polynômes $(X - a)$ et $(X - b)$ sont premiers entre eux : la conclusion sera alors garantie par la proposition **??**. Mais il est facile d'écrire une identité de Bézout : on a
$$\frac{1}{a+b}(X - a) + \frac{-1}{a+}(X - b) = 1.$$
Les polynômes $(X - a)$ et $(X - b)$ sont donc bien premiers entre eux.
- En poursuivant le raisonnement du point précédent, on constate que si P admet au moins k racines distinctes, et si $a_1, \ldots, a_k$ sont ces racines, alors le polynôme $(X - a_1)(X - a_2)\cdots(X - a_k)$ divise P. Comme ce diviseur de P est de degré k, et comme P est de degré n, on en déduit qu'on a nécessairement $k \leq n$. Il est donc impossible que P admette strictement plus de n racines distinctes.

□

Le résultat ci-dessus est crucial. L'une de ses conséquences immédiates est la suivante :

Corollaire 3.46 – Trop de racines, c'est le polynôme nul

1. *Si $P(X) = a_0 + a_1X + \cdots + a_nX^n$ et si P admet $n + 1$ racines distinctes, alors P est le polynôme nul.*
2. *Si P est un polynôme admettant une infinité de racines, alors $P = 0$.*

Relevons une variante utile du corollaire précédent :

Corollaire 3.47 – Deux polynômes de degré n qui coïncident en $n+1$ points...

Si P et Q sont deux polynômes de degré $n \in \mathbb{N}$ et si l'équation $P(x) = Q(x)$ admet au moins $n + 1$ solutions distinctes dans $\mathbb{R}$ ou $\mathbb{C}$, alors les polynômes P et Q sont identiques.

Démonstration. — Si P et Q sont comme dans l'énoncé du théorème, alors le polynôme $A = P - Q$ est de degré inférieur ou égal à n (voir la proposition 3.3), et l'équation $A(x) = 0$ équivaut à $P(x) - Q(x) = 0$. Sous l'hypothèse de l'énoncé, le polynôme A admet donc au moins $(n + 1)$ racines distinctes dans $\mathbb{C}$, alors que son degré est inférieur ou égal à n ; on a donc $A = 0$, autrement dit $P - Q = 0$, et on a bien $P = Q$. □

4.2. Notion de multiplicité d'une racine, version avec la divisibilité. —

Définition 3.48 – Multiplicité d'une racine

Soit $P(X)$ un polynôme de $\mathbb{K}[X]$. Soit a un élément de $\mathbb{K}$ vérifiant $P(a) = 0$.
On appelle *multiplicité de a comme racine de P* le plus grand entier k tel que $(X-a)^k$ divise P.

Ainsi, la racine a est de multiplicité m lorsqu'on a : $(X-a)^m | P$ mais $(X-a)^{m+1} \not| P$.

Exemple 3.49. —

- Dans le polynôme $P(X) = X^3 - 1 = (X-1)(X^2+X+1)$, le nombre 1 est racine simple.
- Dans le polynôme $P(X) = X^4 + 3X^2 = X^2(X^2+3)$, le nombre 0 est racine double, mais pas triple.

Proposition 3.50 – Total des racines avec multiplicité

Soit P un polynôme à coefficients réels ou complexes.
Écrivons la liste des racines (réelles ou complexes) de P sous la forme $\{a_1, \dots, a_k\}$. Pour chaque racine a_i, notons m_i l'ordre de multiplicité de a_i comme racine de P. On a la majoration suivante :

$$\sum_{i=1}^{k} m_i \leq n.$$

De plus, s'il y a égalité dans cette inégalité, alors on a $P(X) = dom(P)(X-a_1)^{m_1} \dots (X-a_k)^{m_k}$.

Démonstration. — Avec les notations de l'énoncé, et compte tenu de la définition de la notion de multiplicité, nous savons que $(X-a_1)^{m_1}$ divise P. Nous savons que $(X-a_2)^{m_2}$ divise P. Or, les polynômes $(X-a_1)^{m_1}$ et $(X-a_2)^{m_2}$ sont premiers entre eux : en effet, les diviseurs de $(X-a_1)^{m_1}$ qui sont unitaires sont des puissances de $X-a_1$, et les diviseurs de $(X-a_2)^{m_2}$ qui sont unitaires sont des puissances de $X-a_2$: il n'y a aucun diviseur unitaire en commun à part 1.

C'est donc que $(X-a_1)^{m_1}(X-a_2)^{m_2}$ divise P. On montre ainsi, de proche en proche, que $Q(X) = (X-a_1)^{m_1} \dots (X-a_k)^{m_k}$ divise P. Le degré de Q étant égal à $d = \sum_{i=1}^{k} m_i$, on a nécessairement $d \leq \deg(P)$, ce qui est l'inégalité $\sum_{i=1}^{k} m_i \leq n$ annoncée dans la proposition.

Si cette égalité est une égalité, alors P et Q ont le même degré et Q divise P ; d'après la proposition **??**, il existe $\alpha \in \mathbb{C}^\star$ tel que $P = \alpha Q$. Comme le coefficient dominant de Q est 1 et celui de P est $\mathrm{dom}(P)$, on a nécessairement $\alpha = \mathrm{dom}(P)$: c'est ce qu'il fallait démontrer. □

4.3. Le théorème de d'Alembert. —

Nous avons vu qu'il existe des polynômes n'ayant pas de racine réelle : c'est le cas pour X^2+1. L'un des piliers de l'algèbre est le fait qu'il n'existe *aucun* polynôme n'ayant pas de racine complexe, sauf les polynômes de la forme $P(X) = c$ avec $c \in \mathbb{C}$:

Théorème 3.51 – Théorème de d'Alembert-Gauss

Soit P un polynôme à coefficients réels ou complexes. Si P n'est pas constant, alors P admet au moins une racine dans $\mathbb{C}$.

Démontrer ce théorème est difficile et fait appel à des techniques fines d'analyse. Nous ne le ferons pas dans ce cours.

***Remarque 3.52* (Existence en théorie vs recherche en pratique)**

Le cas des polynômes de degré 2, avec la notion de discriminant, vous a donné l'habitude de formules existe pour trouver les racines complexes d'un polynôme de degré 2. Il existe des formules de ce type pour le degré 3 et pour le degré 4, mais il n'est pas utile de les connaître.

Mais cette idée s'effondre complètement au-delà du degré 5, et e fait, il est *impossible de donner une formule* permettant de trouver les racines d'un polynôme au-delà du degré 5. Une recherche rapide sur les termes « théorie de Galois » devrait vous donner des éléments à ce sujet.

Pour prendre un exemple concret, si nous considérons le polynôme

$$P(X) = X^7 - 83X^6 + 9X^5 + 21X^4 - X^2 + 1,$$

alors d'après le théorème de d'Alembert-Gauss, le polynôme P admet au moins une racine dans $\mathbb{C}$... Mais il est difficile de trouver concrètement une telle racine, et on sait qu'il n'existe pas de formule générale permettant de le faire.

Corollaire 3.53 – Tout polynôme à cofficients réels est scindé sur $\mathbb{C}$

Tout polynôme non constant à coefficients complexes peut s'écrire sous la forme

$$P(X) = c(X - a_1)^{m_1}(X - a_1)^{m_1}(X - a_1)^{m_1}$$

où $c \in \mathbb{C}$, $r \in \mathbb{N}^\star$, où $a_1, \dots, a_r$ sont des nombres complexes deux à deux distincts, et où $m_1, \dots, m_r$ sont des entiers naturels non nuls.

De plus, une telle écriture est unique et les données y sont nécessairement les suivantes :

- *le coefficient c est le coefficient dominant de P,*
- *l'entier r est le nombre de racines de P dans $\mathbb{C}$,*
- *les nombres $a_1, \dots, a_r$ donnent la liste des racines de P,*
- *pour chaque $i \in \{1, \dots, r\}$, l'entier m_i donne l'ordre de multiplicité de a_i comme racine de P.*

Vocabulaire : scindé, scindé à racines simples. —

- Soit P un polynôme de $\mathbb{K}[X]$. On dit que P est *scindé sur* $\mathbb{K}$ s'il existe un entier $r \in \mathbb{N}^\star$, des éléments $a_1, \dots, a_r$ de $\mathbb{K}$ et des entiers $m_1, \dots, m_r$ de $\mathbb{N}^\star$ vérifiant $P = \mathrm{dom}(P)(X - a_1)^{m_1}(X - a_2)^{m_2} \dots (X - a_r)^{m_r}$.
- Le théorème ci-dessus dit que tout polynôme à coefficients réels ou complexes est scindé sur $\mathbb{C}$. Par contre, l'exemple de $X^2 + 1$ montre qu'il existe des polynômes qui ne sont pas scindés sur $\mathbb{R}$.
- Si P est à coefficients réels, dire que P est scindé sur $\mathbb{R}$ revient à dire que si l'on observe l'écriture $P = \mathrm{dom}(P)(X - a_1)^{m_1}(X - a_2)^{m_2} \dots (X - a_r)^{m_r}$ de P dans $\mathbb{C}[X]$, alors les racines a_i sont toutes réelles.
- Si P est un polynôme de $\mathbb{K}[X]$, on dit que P est *scindé à racines simples* lorsque P est scindé et lorsque toutes les racines de P sont de multiplicité 1. Si P est de degré $n \geq 2$, dire que P est scindé à racines simples est équivalent à dire que P admet exactement n racines distinctes dans $\mathbb{C}$, ou encore que P puisse s'écrire sous la forme $P(X) = \mathrm{dom}(P)(X - a_1)\ (X - a_2) \dots (X - a_n)$ où $a_1, \dots, a_n$ sont des éléments distincts de $\mathbb{K}$.

Exemple 3.54. — Le polynôme

$$X^4 - 1 = (X^2 - 1)(X^2 + 1) = (X - 1)(X + 1)(X^2 + 1) = (X - 1)(X + 1)(X - i)(X + i)$$

est scindé sur $\mathbb{C}$ (comme tout polynôme non constant), et il est scindé à racines simples sur $\mathbb{C}$. Par contre, il n'est pas scindé sur $\mathbb{R}$, car il admet des racines complexes non réelles.

4.4. Conséquences sur les polynômes irréductibles et la factorisation. —

4.4.1. *Sur le corps des complexes.* —

Proposition 3.55 – Sur $\mathbb{C}$, les seuls polynômes irréductibles sont les $aX+b$, $a\neq 0$

Soit P un polynôme à coefficients complexes. Le polynôme P est irréductible dans $\mathbb{C}[X]$ si et seulement s'il est de la forme $P(X)=aX+b$ avec $(a,b)\in\mathbb{C}^\star\times\mathbb{C}$.

Proposition 3.56 – Sur $\mathbb{C}$, lien entre factorisation et racines

Soit P un polynôme non constant à coefficients complexes. La décomposition de P en produit de facteurs irréductibles est donnée par l'écriture

$$P(X)=dom(P)(X-a_1)^{m_1}(X-a_2)^{m_2}(X-a_k)^{m_k}$$

où $a_1,a_2,\dots,a_k$ sont les racines de P et où les entiers $m_1,\dots m_k$ sont leurs multiplicités comme racine de P.

4.4.2. *Sur le corps des réels.* —

Proposition 3.57 – Sur $\mathbb{R}$, liste des polynômes irréductibles

Les polynômes irréductibles de $\mathbb{R}[X]$ sont

- *les polynômes de degré 1, de la forme $aX+b$ avec $a\neq 0$,*
- *les polynômes de degré 2 sans racine réelle, de la forme aX^2+bX+c avec $a\neq 0$ et $b^2-4ac<0$.*

Démonstration. — Soit P un polynôme irréductible de $\mathbb{R}[X]$. Par définition de l'irréductibilité, il est impossible que P soit constant. Il existe donc au moins une racine a de P dans $\mathbb{C}$; notons a une telle racine complexe. Distinguons alors deux cas :

- Si $a\in\mathbb{R}$, alors a est une racine réelle de P ; le polynôme $X-a$ divise donc P, et il est dans $\mathbb{R}[X]$. Comme P est supposé irréductible les seuls diviseurs de P sont soit constants, soit associés à P ; c'est donc que $P(X)=c(X-a)$ avec $c\in\mathbb{R}^\star$, et donc P est de degré 1.
- Si $a\notin\mathbb{R}$, alors comme P est à coefficients réels, son conjugué $\bar{a}$ est aussi une racine complexe de P (voir la remarque 3.41). Les polynômes $(X-a)$ et $X-\bar{a})$ sont donc des diviseurs de P dans $\mathbb{C}[X]$. Ces polynômes sont à coefficients complexes et non à coefficients réels ; mais ils sont premiers entre eux (voir la démonstration de la proposition 3.44) ; on en déduit que $(X-a)(X-\bar{a})$ divise P. Or, $(X-a)(X-\bar{a})$ est un polynôme de degré 2 à coefficients réels : il est égal à $X^2-2\mathfrak{Re}(a)X+|a|^2$. Nous avons donc trouvé un polynôme de degré 2 qui divise P ; comme P est supposé irréductible, c'est que P est lui-même de degré 2. De plus, il n'a pas de racine réelle, sinon nous serions dans le cas précédent. Nous constatons donc que P est de la forme aX^2+bX+c avec $a\neq 0$ et $b^2-4ac<0$.

Tout cela montre que les polynômes listés dans la proposition sont les seuls qui peuvent être irréductibles. Il reste à montrer qu'ils le sont effectivement :

- Si P est de degré 1, nous avons vu que P est irréductible (proposition 3.33)
- Si P est de degré 2 sans racine réelle, aucun polynôme de la forme $aX+b$, avec a,b réels et $a\neq 0$, ne peut diviser P (puisqu'un tel diviseur donnerait pour P une racine en $\frac{-b}{a}$). Comme les seuls diviseurs possibles pour un polynôme de degré 2 sont soit constants, soit associés à P, soit de degré 1, cela montre que P est irréductible dans $\mathbb{R}[X]$.

□

Corollaire 3.58 – Sur $\mathbb{R}$, forme concrète du théorème de factorisation

Soit P un polynôme non constant à coefficients complexes. La décomposition de P en produit de facteurs irréductibles est de la forme

$$P(X) = dom(P)(X-a_1)^{m_1}\dots(X-a_k)^{m_2}(X^2+\alpha_1 X+\beta_1)^{n_1}\dots(X^2+\alpha_\ell X+\beta_\ell)^{n_\ell}$$

où

- *$a_1, a_2, \dots, a_k$ sont les racines réelles de P, et les entiers $m_1, \dots m_k$ sont leurs multiplicités,*
- *$\ell \in \mathbb{N}$, et pour tout $i \in \{1, \dots, \ell\}$, le polynôme $X^2+\alpha_i X+\beta_i$, $i \in \{1, \dots \ell\}$ est un polynôme de degré 2 sans racine réelle, et $n_i \in \mathbb{N}^\star$.*

Remarque 3.59. — Il est facile d'obtenir la factorisation sur $\mathbb{R}$ si l'on connaît la factorisation sur $\mathbb{C}$. En effet, si P est un polynôme non constant à coefficients réels, les racines complexes de P se séparent en deux catégories :

- celles qui appartiennent à $\mathbb{R}$; notons-les $a_1, \dots, a_r$
- celles qui n'appartiennent pas à $\mathbb{R}$. Mais nous savons que si b est une telle racine de P et si P est à coefficients réels, alors son conjugué $\bar{b}$ est aussi une racine de P. La liste des racines complexes non réelles de P est donc de la forme : $b_1, \bar{b_1}, b_2, \bar{b_2}, \dots, b_\ell, \bar{b_\ell}$, où les b_i, $i \in \{1, \dots, \ell\}$ sont des nombres complexes non réels.

La factorisation de P sur $\mathbb{C}$ est alors de la forme

$$P(X) = c(X-a_1)^{m_1}\dots(X-a_k)^{m_k}(X-b_1)^{n_1}(X-\bar{b_1})^{n'_1}\dots(X-b_\ell)^{n_\ell}(X-\bar{b_1})^{n'_\ell}$$

et on remarque que pour tout $i \in \{1, \dots, \ell\}$, les multiplicités n_i et n'_i sont égales (cela vient du fait que si une puissance $(X-b_1)^k$ divise P, alors comme P est à coefficients réels, le polynôme conjugué $(X-\bar{b_1})^k$) divise aussi P).

On en déduit que

$$\begin{aligned} P(X) &= c(X-a_1)^{m_1}\dots(X-a_k)^{m_k}(X-b_1)^{n_1}(X-\bar{b_1})^{n_1}\dots(X-b_\ell)^{n_\ell}(X-\bar{b_1})^{n_\ell} \\ &= c(X-a_1)^{m_1}\dots(X-a_k)^{m_k}\left[(X-b_1)(X-\bar{b_1})\right]^{n_1}\dots\left[(X-b_\ell)(X-\bar{b_\ell})\right]^{n_\ell} \\ &= c(X-a_1)^{m_1}\dots(X-a_k)^{m_k}\left[X^2+\beta_1 X+\gamma_1\right]^{n_1}\dots\left[X^2+\beta_\ell X+\gamma_\ell\right]^{n_\ell} \end{aligned}$$

où on a noté $X^2+\beta_i X+\gamma_i$ le produit $(X-b_i)(X-\bar{b_i})$, qui est un polynôme de degré 2 sans racine réelle.

C'est ainsi que l'on trouve la factorisation de P sur $\mathbb{R}$ à partir de la factorisation sur $\mathbb{C}$.

5. Polynôme dérivé ; lien avec la multiplicité des racines

5.1. Polynôme dérivé : définition et propriétés élémentaires. —

Nous étions partis, au début de ce chapitre, de la notion de *fonction polynomiale* $f : x \mapsto a_0 + a_1 x + \dots + a_n x^n$, puis avions considéré une notion abstraite, celle de polynôme $a_0 + a_1 X + \dots + a_n X^n$.

La notion de dérivée, déjà connue de vous pour les fonctions, n'existe pas encore dans le cadre plus abstrait qui est le nôtre depuis le début de ce chapitre. Cependant, il n'est pas difficile de l'y transposer :

Définition 3.60 – Polynôme dérivé

Fixons $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$. Soit P un élément de $\mathbb{K}[X]$.
Écrivons $P(X) = a_0 + a_1X + \cdots + a_nX^n = \sum_{k=0}^{n} a_kX^k$, où $a_0, \ldots, a_n$ sont des éléments de $\mathbb{K}$.
On appelle *polynôme dérivé* de P, et on note P', le polynôme de $\mathbb{K}[X]$ donné par la formule suivante :

$$P'(X) = a_1 + 2a_2X + \cdots + na_nX^{n-1} = \sum_{k=1}^{n} ka_kX^{k-1}.$$

Exemple 3.61. — Si $P(X) = X^4 - 5X^3 + X^2 + 1$, alors $P'(X) = 4X^3 - 15X^2 + 2X$.

Relevons le fait suivant :

Proposition 3.62 – Degré du polynôme dérivé

Soit P un polynôme non constant. On a l'égalité

$$\deg(P') = \deg(P) - 1.$$

Attention. — Dans l'affirmation ci-dessus, l'hypothèse « P non constant » est nécessaire : pour P constant, le degré de P' est $-\infty$. Dans les raisonnements faisant intervenir le degré de P', il ne faudra donc pas oublier ce *caveat*.

Proposition 3.63 – Dérivée d'une somme et d'un produit

Soient P et Q deux polynômes à coefficients réels ou complexes. On a les égalités suivantes

$$(P + Q)' = P' + Q'$$
$$(PQ)' = P'Q + QP'$$

Démonstration. — Seule l'assertion sur le produit présente une difficulté, c'est donc elle que nous démontrons.

- Montrons d'abord que le résultat est vrai si $P_1(X) = X^m$ et $P_2(X) = X^n$. Dans ce cas

$$(P_1P_2)'(X) = (X^{m+n})' = (m+n)X^{n+m-1}$$

tandis que

$$P_1'(X)P_2(X) + P_1(X)P_2'(X) = mX^{m-1}X^n + X^m(nX^{n-1}) = mX^{m+n-1} + nX^{m+n-1} = (m+n)X^{n+m-1}\ .$$

L'égalité voulue est bien vérifiée.

- Démontrons maintenant le résultat dans le cas général. Soient $P_1(X) = \sum_{k=0}^n a_k X^k$ et $P_2(X) = \sum_{i=0}^n b_i X^i$. Alors

$$\begin{aligned}
(P_1P_2)'(X) &= \left(\sum_{k=0}^n \sum_{i=0}^n a_k b_i X^{k+i}\right)' \\
&= \sum_{k=0}^n \sum_{i=0}^n a_k b_i (X^{k+i})' \\
&= \sum_{k=0}^n \sum_{i=0}^n a_k b_i \left((X^k)' X^i + X^k (X^i)'\right) \\
&= \left(\sum_{k=0}^n \sum_{i=0}^n a_k b_i (X^k)' X^i\right) + \left(\sum_{k=0}^n \sum_{i=0}^n a_k b_i X^k (X^i)'\right) \\
&= \left(\sum_{k=0}^n a_k (X^k)'\right)\left(\sum_{i=0}^n b_i X^i\right) + \left(\sum_{k=0}^n a_k X^k\right)\left(\sum_{i=0}^n b_i (X^i)'\right) \\
&= P_1'(X)P_2(X) + P_1(X)P_2'(X)
\end{aligned}$$

  ce qu'il fallait démontrer.

□

5.2. Dérivées d'ordre supérieur. —

Définition 3.64 – Dérivées d'ordre supérieur

Soit P un polynôme à coefficients réels ou complexes. Pour tout $k \geq 2$, on définit le polynôme $P^{(k)}$ (dérivée k-ème de P) par la formule récursive suivante : $P^{(0)} = P$, $P^{(1)} = P'$ et pour tout $k \geq 2$, $P^{(k)} = (P^{(k-1)})'$.

Ainsi, $P^{(2)} = (P')'$, souvent noté P'' ; de même, $P^{(5)} = (((((P')')')')')'$.

***Exemple 3.65* (Dérivées successives du polynôme X^n).** — Si on fixe un entier $n \geq 2$ et si P est le polynôme X^n, on a $P'(X) = nX^{n-1}$, $P''(X) = n(n-1)X^{n-2}$; en poursuivant, on constate que pour tout entier $k \leq n$, on a $P^{(k)} = n(n-1)\dots(n-k+1)X^{n-k}$. Comme la n-ème dérivée $P^{(n)}$ est un polynôme constant, on a ensuite $P^{(k)} = 0$ pour $k > n$.

Plus généralement, si P est un polynôme non constant et si n est son degré, alors pour tout $k \leq n$, on a $\deg(P^{(k)}) = n - k$, tandis que $P^{(k)} = 0$ pour $k > n$.

Les dérivées successives d'une somme de deux polynômes sont faciles à calculer : en effet, si P et Q sont deux polynômes à coefficients réels ou complexes, alors $(P+Q)' = P' + Q'$, et en dérivant à nouveau, on trouve $(P+Q)'' = P'' + Q''$, puis $(P+Q)^{(k)} = P^{(k)} + Q^{(k)}$ pour tout k.

Par exemple, si $P(X) = X^5 + X^4 + X^3 + X^2 + 1$ et $Q(X) = X^3 - 7X + 9$ et si l'on cherche à calculer la dérivée cinquième de $P + Q$, on sait que $Q^{(5)} = 0$ et $P^{(5)}$ est le polynôme constant $5 \times 4 \times 3 \times 2 \times 1 = 120$.

La formule pour la dérivée d'un produit, en revanche, est suffisamment alambiquée pour qu'on

$$\begin{aligned}
(PQ)' &= P'Q + PQ' \\
(PQ)'' &= P''Q + P'Q' + P'Q' + PQ'' \\
&= P''Q + 2P'Q' + PQ'' \\
(PQ)''' &= (P'''Q + P''Q') + 2(P''Q' + P'Q'') + (P'Q'' + PQ''') \\
&= P'''Q + 3P''Q' + 3P'Q'' + Q'''.
\end{aligned}$$

Nous voyons apparaître une structure qui rappelle la formule du binôme. Elle est complètement générale :

Proposition 3.66 – Formule de Leibniz pour les dérivées successives d'un produit

Soient P et Q deux polynômes à coefficients réels ou complexes. Pour tout entier $n \in \mathbb{N}^\star$, on a l'égalité suivante :

$$(PQ)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} P^{(k)} Q^{(n-k)}.$$

Démonstration. — Raisonnons par récurrence sur n.

Nous avons vu que la formule est vraie pour $n = 1$, pour $n = 2$, pour $n = 3$...

Fixons $n \in \mathbb{N}^\star$ et supposons vérifiée l'égalité $(PQ)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} P^{(k)} Q^{(n-k)}$. Dérivons une fois de plus pour trouver $(PQ)^{(n+1)}$:

$$\begin{aligned}(PQ)^{(n+1)} &= \sum_{k=0}^{n} \binom{n}{k} \left[P^{(k+1)} Q^{(n-k)} + P^{(k)} Q^{(n-k+1)}\right] \\ &= \sum_{k=0}^{n} \binom{n}{k} P^{(k+1)} Q^{(n-k)} + \sum_{k=0}^{n} \binom{n}{k} P^{(k)} Q^{(n-k+1)}.\end{aligned}$$

Pour poursuivre, faisons le changement d'indice $\ell = k + 1$ dans la première somme. On trouve

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

en isolant le dernier terme et en renommant l'indice k de la deuxième somme. En utilisant la formule $\binom{n}{\ell-1} + \binom{n}{\ell} = \binom{n+1}{\ell}$ et en rassemblant, on trouve

$$(PQ)^{(n+1)} = P^{(n+1)} Q^{(0)} + \sum_{\ell=0}^{n} \binom{n+1}{\ell} P^{(\ell)} Q^{(n+1-\ell)}$$

qui est la formule que nous voulions démontrer.

□

5.3. La formule de Taylor pour les polynômes. —

Soit P un polynôme à coefficients réels ou complexes. Écrivons $P(X) = a_0 + a_1 X + \cdots + a_n X^n$ où

Peut-on exprimer les coefficients $a_0, \ldots, a_n$ à l'aide de dérivées et de dérivées successives ?

La réponse est oui, si l'on utilise les dérivées successives en zéro : en effet, en dérivant plusieurs fois grâce aux formules pour les dérivées des puissances de X vues à l'exemple 3.65, on constate que

$$\begin{aligned} P(0) &= a_0 \\ P'(0) &= a_1 \\ P''(0) &= 2a_2 \\ P'''(0) &= 3 \cdot 2 \cdot a_3 = (3!)a_3 \\ P^{(4)}(0) &= 4 \cdot 3 \cdot 2 \cdot 1 \cdot a_4 = (4!)a_4 \\ &\text{etc...} \end{aligned}$$

On peut donc retrouver le coefficient a_k à partir de $P^{(k)}(0)$: la relation est $a_k = \frac{P^{(k)}(0)}{k!}$.

Nous constatons donc que la formule suivante est vraie pour tout polynôme non nul P de degré $n \in \mathbb{N}$:

$$P(X) = \sum_{k=0}^{n} \frac{P^{(k)}(0)}{k!} X^k. \tag{5.1}$$

Plus généralement, on a la relation suivante :

Théorème 3.67 – Formule de Taylor pour les polynômes

Soient P un polynôme non nul à coefficients réels ou complexes et a un nombre réel ou complexe. Notons n le degré de P.
On a l'égalité de polynômes suivantes :

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

Démonstration. — Définissons le polynôme $Q(X) = P(X+a)$. On a $Q(0) = P(a)$, et de même on constate que pour tout $k \in \mathbb{N}$ $Q^{(k)}(0) = P^{(k)}(a)$. En appliquant la formule (5.1) au polynôme Q, on obtient

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

En appliquant cette formule en X-a, on obtient

$$P(X-a+a) = \sum_{k=0}^{n} \frac{P^{(k)}(a)}{k!} (X-a)^k,$$

ce qui est l'égalité cherchée. □

Exemple 3.68. — Si P est un polynôme de degré 3 vérifiant $P(5) = 0$, $P'(5) = 0$ et $P''(5) = 0$, alors on a $P(X) = a(X-5)^3$, où a est le nombre $\frac{P'''(5)}{5!}$.

Exemple 3.69. — Si P est un polynôme de degré n vérifiant $P^{(k)}(5) = 0$ pour tout $k \leq n$, alors P est en fait le polynôme nul.

5.4. Multiplicité des racines et dérivée. — Rappelons que si P est un polynôme à coefficients réels ou complexes et si a est un nombre réel ou complexe, alors a est une racine de multiplicité 2 de P si $(X-a)^2$ divise P, mais $(X-a)^3$ ne divise pas P.

Observons ce que veut dire cette condition lorsque $a = 0$. Il s'agit de voir si P est divisible par X^2. Mais alors, P s'écrit sous la forme $P(X) = a_2X^2 + a_3X^3 + \cdots + a_nX^n$, où n est un entier supérieur ou égal à 2 et où $a_2, \ldots, a_n$ sont des nombres réels ou complexes.

On constate alors que $P(0) = 0$, ce qui n'est pas étonnant puisque 0 est racine de P. Mais on constate aussi que P' est divisible par X, donc que $P'(0) = 0$, alors que $P''(0) \neq 0$.

On peut donc lire sur les nombres $P(0)$, $P'(0)$ et $P''(0)$ le fait que 0 soit ou non une racine de multiplicité 2 de P.

Plus généralement, on a le résultat suivant.

Proposition 3.70 – Multiplicité d'une racine et annulation des dérivées successives

Soient P un polynôme à coefficients réels ou complexes et a un nombre réel ou complexe. Fixons un entier n_0 de $\mathbb{N}^\star$. Les deux assertions suivantes sont équivalentes :

- *le nombre a est racine de multiplicité n_0 de P ;*
- *on a $P^{(k)}(a) = 0$ pour tout $k \leq n_0 - 1$, mais on a $P^{(n_0)}(a) \neq 0$.*

Démonstration. —

- Si a est racine de multiplicité n_0 de P, on a $P(X) = (X-a)^{n_0}Q(X)$, où Q est un polynôme qui n'est pas divisible par $(X-a)$, donc qui vérifie $Q(a) \neq 0$.

  Montrons alors que $P^{(k)}(a) = 0$ pour tout $k \leq n_0 - 1$, mais on a $P^{(n_0)}(a) \neq 0$.

  Si nous utilisons la formule de Leibniz pour le calcul des dérivées successives, on obtient

$$P^{(k)}(X) = \sum_{i=0}^{k} \binom{k}{i} \left[n_0(n_0-1)\dots(n_0-i+1)(X-a)^{n_0-i}\right] Q^{(k-i)}(X).$$

  Or, pour tout $i < n_0$, le polynôme $(X-a)^{n_0-i}$ s'annule en a ; on en déduit que $P^{(k)}(a) = 0$ pour tout $k < n_0$. Pour $k = n_0$, dans

$$P^{(n_0)}(X) = \sum_{i=0}^{n_0} \binom{k}{i} \left[n_0(n_0-1)\dots(n_0-i+1)(X-a)^{n_0-i}\right] Q^{(k-i)}(X),$$

  tous les termes correspondant à des indices $i < n_0$ s'annulent en a pour la raison ci-dessus ; en revanche, le terme qui correspond à $i = n_0$ est égal à $(n_0!)Q^{(0)}(X)$. On en déduit que $P^{(n_0)}(a) = (n_0!)Q(a)$, qui est non-nul.

  Nous avons bien prouvé que $P^{(k)}(a) = 0$ pour tout $k < n_0$, mais $P^{(n_0)}(a) \neq 0$.

- L'autre implication est conséquence de la formule de Taylor pour les polynômes : si P est non nul (pour le polynôme nul il n'y a rien à prouver), on a

$$P(X) = \sum_{k=0}^{\deg(P)} \frac{P^{(k)}(a)}{k!}(X-a)^k$$

  et si toutes les dérivées $\frac{P^{(k)}(a)}{k!}$ sont nulles pour $k \leq n_0 - 1$, on peut réécrire cette somme comme

$$\begin{aligned} P(X) &= \sum_{k=n_0}^{\deg(P)} \frac{P^{(k)}(a)}{k!}(X-a)^k \\ &= (X-a)^{n_0} \left(\sum_{k=n_0}^{\deg(P)} \frac{P^{(k)}(a)}{k!}(X-a)^{k-n_0} \right) \end{aligned}$$

  ce qui prouve que $(X-a)^{n_0}$ divise P. De plus, si $(X-a)^{n_0+1}$ divisait P, d'après la première partie de notre démonstration, on aurait $P^{(n_0)}(a) = 0$, et nous avons supposé que ce n'était pas le cas. C'est donc bien que a est racine de multiplicité n_0 dans P.

□

Exercices du chapitre 3

Généralités, notion de degré. —

Exercice 3.1. — ★☆☆
Soient P et Q deux polynômes à coefficients réels. On suppose que P et Q ne sont pas constants.

1. Peut-on avoir $P+Q=PQ$?
2. Peut-on avoir $(P+Q)^2=(P-Q)^2$?
3. Peut-on avoir $(P-Q)^3=P^3-Q^3$? Et si P et Q n'ont pas le même degré ?

Exercice 3.2. — ★☆☆
Soient p et q deux entiers naturels non nuls. En développant de deux façons différentes le polynôme $P(X)=(X+1)^{p+q}=(X+1)^p(X+1)^q$, démontrer que l'identité suivante est vraie pour tout entier n vérifiant $n\leq p$ et $n\leq q$:

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

Arithmétique : division euclidienne, PGCD*...* —

Exercice 3.3. — ★★☆
Étant donné un couple $(a,b)\in\mathbb{R}^2$, on considère les polynômes

$$A(X)=X^5+X^4+aX^3+bX^2+5X-2 \text{ et}$$
$$B(X)=X^3-2X+1.$$

Est-il possible de choisir a et b de façon à ce que B divise A ?

Exercice 3.4. — ★☆☆
Trouver le quotient et le reste de la division euclidienne

(a) $\begin{cases}\text{de} & X^3+6X^2+2X+5\\ \text{par} & 2X^2+4.\end{cases}$

(b) $\begin{cases}\text{de} & X^7+2X^5+7X^3+15X+2\\ \text{par} & X^3+2X.\end{cases}$

(c) $\begin{cases}\text{de} & X^4+1\\ \text{par} & X^2+1.\end{cases}$

(d) $\begin{cases}\text{de} & 2X^3+17X^2-7X+2\\ \text{par} & 2X^5-1.\end{cases}$

Exercice 3.5. — ★☆☆
Dans tout l'exercice, on fixe deux réels a et b et on suppose $a\neq b$.

1. Exprimer le reste de la division euclidienne de P par $(X-a)(X-b)$ en fonction de a, b, $P(a)$ et $P(b)$.
2. Exprimer le reste de la division euclidienne de X^n+X+b par $X-a$ en fonction de a et b.

Exercice 3.6. — ★☆☆

1. Fixons un entier $n\geq 2$. Déterminer le reste de la division euclidienne de X^n+X+1 par $(X-1)^2$.
2. Fixons deux entiers naturels p et q avec $p>q$.
   Déterminer le reste de la division euclidienne de X^p+X^q+1 par X^2+X.

Exercice 3.7. — ★☆☆

1. Montrer que le polynôme $X+1$ divise les polynômes X^5+1 et X^3+1.
2. On note P_1 et P_2 les polynômes vérifiant $X^5+1=(X+1)P_1(X)$ et $X^3+1=(X+1)P_2(X)$. Écrire explicitement P_1 et P_2, puis montrer que P_1 et P_2 sont premiers entre eux.
3. En déduire le PGCD de X^5+1 et X^3+1, puis le PPCM de X^5+1 et X^3+1.

Notion de racine : lien avec la divisibilité. —

Exercice 3.8. — ★☆☆

Dans tout l'exercice, on fixe trois entiers n, m, p dans $\mathbb{N}^\star$.

1. Démontrer que $B(X)=X(X+1)(2X+1)$ divise $A(X)=(X+1)^{2n}-X^{2n}-2X-1$.
2. Démontrer que $B(X)=X^2+X+1$ divise $A(X)=X^{3n+2}+X^{3m+1}+X^{3p}$.

Exercice 3.9. — ★☆☆

On fixe $\theta\in\mathbb{R}$ et $n\in\mathbb{N}^\star$. Montrer que le polynôme $X^2-2\cos(\theta)X+1$ divise $X^{2n}-2\cos(n\theta)X^n+1$.

Exercice 3.10. — ★★☆

On fixe $n\in\mathbb{N}^\star$. Démontrer que $B(X)=(X-1)^2$ divise $A(X)=nX^{n+1}-(n+1)X^n+1$.

Exercice 3.11. — ★★☆

Dans cet exercice, on considère le polynôme $P(X)=(X^2-X+1)^2+1$.

1. Montrer que i est racine de P.
2. En déduire que P est divisible par X^2+1.
3. Déterminer la décomposition de P en produit de facteurs irréductibles dans $\mathbb{R}[X]$ et dans $\mathbb{C}[X]$.

Exercice 3.12. — ★☆☆

Dans cet exercice, on considère les polynômes $P(X)=X^4+1$ et $Q(X)=X^3+1$.

1. Déterminer les racines de P et Q, écrites sous forme algébrique et sous forme trigonométrique.
2. Décomposer P et Q en produits de facteurs irréductibles dans $\mathbb{R}[X]$.
3. Calculer le PGCD de P et Q.
4. En déduire un couple (U,V) de polynômes vérifiant : $UP+VQ=1$, $\deg(P)\leq 3$ et $\deg(Q)\leq 3$.

Notion de racine : utilisation de résultats d'analyse. —

Exercice 3.13. — ★☆☆

1. À l'aide d'un tableau de variations, montrer que le polynôme $P(X)=X^3+3X-5$ admet une et une seule racine réelle.
2. À l'aide d'un tableau de variations, montrer que le polynôme $P(X)=X^3+4X^2-2X$ admet trois racines réelles distinctes.
3. Montrer que tout polynôme de degré impair admet au moins une racine réelle.

Exercice 3.14. — ★★☆

1. Grâce à une étude de variations, montrer que le polynôme $P(X)=X^5-X^2+1$ admet une seule racine réelle.
2. Montrer que cette racine est irrationnelle.

Exercice 3.15. — ★★☆
Soient p et q deux réels. On considère le polynôme $P(X) = X^3 + pX + q$.
Le but de cet exercice est de déterminer, selon les valeurs de p et q, le nombre de racines réelles de P.

1. En utilisant le théorème des valeurs intermédiaires, montrer que P admet au moins une racine réelle.
2. Grâce à l'étude des variations de $x \mapsto P(x)$, montrer que si $p > 0$, alors P admet une racine réelle et deux racines complexes conjuguées.
3. On suppose désormais $p \leq 0$. On note $a = \sqrt{-p/3}$.
   (a) Par une étude de variations, relier le nombre de racines réelles de P au signe de $P(a)P(-a)$.
   (b) En déduire que lorsque $p \leq 0$, le polynôme P admet 1, 2 ou 3 racines réelles selon que $4p^3+27q^2$ est > 0, $= 0$ ou < 0.

Pour s'exercer à la décomposition en produits de facteurs irréductibles. —

Exercice 3.16. — ★☆☆

1. Déterminer la factorisation, dans $\mathbb{C}[X]$ et dans $\mathbb{R}[X]$, du polynôme $P(X) = X^4 + 3X^3 + 3X^2$.
2. Déterminer la factorisation, dans $\mathbb{C}[X]$ et dans $\mathbb{R}[X]$, du polynôme $P(X) = X^3 + 2X^2 + 2X + 1$.

Exercice 3.17. — ★☆☆

1. Déterminer la factorisation, dans $\mathbb{C}[X]$ et dans $\mathbb{R}[X]$, du polynôme $P(X) = X^4 + 2X^3 - X - 2$.
2. Déterminer la factorisation, dans $\mathbb{C}[X]$ et dans $\mathbb{R}[X]$, de $P(X) = X^5 + X^4 + 2X^3 + 2X^2 + X + 1$.

Exercice 3.18. — ★★☆
On considère le polynôme suivant :

$$P(X) = X^4 + 6X^3 + 16X^2 + 22X + 15.$$

1. Montrer qu'il est possible de trouver deux réels λ et μ tels que
$$P(X) = (X^2 + 3X + \lambda)^2 - (X + \mu)^2.$$
2. En déduire la décomposition de P en produit de facteurs irrédutibles, dans $\mathbb{C}[X]$ et dans $\mathbb{R}[X]$.

Exercice 3.19. — ★★☆

1. Déterminer la factorisation, dans $\mathbb{C}[X]$ et dans $\mathbb{R}[X]$, du polynôme $P(X) = X^2 + X + 1$.
2. En déduire la factorisation, dans $\mathbb{C}[X]$ et dans $\mathbb{R}[X]$, du polynôme $Q(X) = X^4 + X^2 + 1$.
3. Déterminer la factorisation, dans $\mathbb{C}[X]$ et dans $\mathbb{R}[X]$, du polynôme $R(X) = X^8 + X^4 + 1$.

Exercice 3.20. — ★☆☆
Dans cet exercice, on fixe $\theta \in \mathbb{R}$.

1. Déterminer la factorisation, dans $\mathbb{C}[X]$ et dans $\mathbb{R}[X]$, du polynôme $P(X) = X^2 - 2\cos(\theta)X + 1$.
2. En déduire la factorisation, dans $\mathbb{C}[X]$ et dans $\mathbb{R}[X]$, du polynôme $P(X) = X^4 - 2\cos(\theta)X^2 + 1$.

Exercice 3.21. — ★☆☆

1. Déterminer la factorisation, dans $\mathbb{C}[X]$ et dans $\mathbb{R}[X]$, du polynôme $P_1(X) = X^3 - 1$.
2. Déterminer la factorisation, dans $\mathbb{C}[X]$ et dans $\mathbb{R}[X]$, du polynôme $P_2(X) = X^3 - 1$.
3. Déterminer la factorisation, dans $\mathbb{C}[X]$ et dans $\mathbb{R}[X]$, du polynôme $P_3(X) = X^3 + 1$.
4. Déterminer la factorisation, dans $\mathbb{C}[X]$ et dans $\mathbb{R}[X]$, du polynôme $P_4(X) = X^4 + 1$.
5. Déterminer la factorisation, dans $\mathbb{C}[X]$ et dans $\mathbb{R}[X]$, du polynôme $P_5(X) = 1+X+X^2+X^3+X^4+X^5$.

Exercice 3.22. — ★★☆

On fixe un entier $n \in \mathbb{N}^\star$. En distinguant selon la parité de n, déterminer la factorisation dans $\mathbb{R}[X]$:

1. du polynôme $X^n - 1$,
2. du polynôme $1 + X + \cdots + X^n$,
3. du polynôme $X^n - 2$,
4. du polynôme $X^n + 1$.

Exercice 3.23. — ★★☆

Le but de cet exercice est de déterminer tous les polynômes $P \in \mathbb{R}[X]$ qui sont scindés sur $\mathbb{R}$ et qui vérifient :

$$P(X)P(X+2) + P(X^2) = 0.$$

1. Montrer que si α est une racine de P, alors α^2 est aussi une racine de P.
2. En déduire que la seule racine possible de P est 1.
3. Conclure.

Polynôme dérivé, lien avec la multiplicité des racines. —

Exercice 3.24. — ★☆☆

1. Existe-t-il un polynôme non nul $P \in \mathbb{R}[X]$ vérifiant $P' = P$?
2. Existe-t-il un polynôme non nul $P \in \mathbb{R}[X]$ vérifiant $P(X) = XP'(X)$?

Exercice 3.25. — ★☆☆

Dans cet exercice, on fixe deux réels a et b et on considère le polynôme $P(X) = X^4 + aX^3 + b$.

1. Montrer que si P admet une racine multiple au point $z \in \mathbb{C}$, alors on a soit $z = 0$, soit $z = -\frac{3a}{4}$.
2. En déduire une condition nécessaire et suffisante sur le couple $(a, b) \in \mathbb{R}^2$ pour que P admette une racine multiple.

Exercice 3.26. — ★☆☆

On fixe un entier $n \geq 2$ et on considère le polynôme

$$P(X) = X^{2n} - nX^{n+1} + nX^{n-1} - 1.$$

Déterminer l'ordre de multiplicité de 1 comme racine de P, en distinguant les cas $n = 2$, $n = 3$ et $n \geq 4$.

Exercice 3.27. — ★☆☆

Dans cet exercice, on note $P(X) = X^3 - 3X + 1$.

1. Montrer que P n'admet pas de racine multiple dans $\mathbb{C}$.
2. On note a, b et c les trois racines complexes de P. Sans chercher à calculer a, b et c, déterminer la valeur des nombres suivants :

$$s = a + b + c, \quad m = ab + ac + bc \quad \text{et} \quad p = abc.$$

Exercice 3.28. — ★★☆

Soit P un polynôme à coefficients réels. Montrer que si P est de degré 3 et si P admet une racine double dans $\mathbb{C}$, alors P admet en fait trois racines réelles (éventuellement confondues).

Exercice 3.29. — ★☆☆

Dans cet exercice, on considère le polynôme $P(X) = (X+1)^7 - X^7 - 1$.

1. Quel est le degré de P ?

2. Montrer que P est divisible par $(X-j)^2$, où j est le nombre complexe $e^{\frac{2i\pi}{3}}$.
3. Donner deux racines réelles "simples" de P, en précisant leurs multiplicité.
4. En remarquant que P est à coefficients réels, en déduire la factorisation de P dans $\mathbb{C}[X]$.

Exercice 3.30. — ★☆☆
On fixe trois réels a, b, c et on considère $P(X) = X^6 + aX^4 + bX^4 + c$, vu comme élément de $\mathbb{C}[X]$.

1. Déterminer a, b et c tels que 1 soit racine double de P et $j = e^{i\frac{2\pi}{3}}$ soit racine de P.
2. Montrer que si ces conditions sont vérifiées, alors P est en fait à coefficients réels et j est en fait racine double de P.
3. Trouver la factorisation de P en produit de facteurs irréductibles, dans $\mathbb{C}[X]$ et dans $\mathbb{R}[X]$.

Exercice 3.31. — ★☆☆
Dans cet exercice, on considère le polynôme

$$P(X) = X^8 + 2X^6 + 3X^4 + 2X^2 + 1.$$

1. Montrer que le nombre complexe $j = e^{\frac{2i\pi}{3}}$ est racine de P, en précisant son ordre de multiplicité.
2. Quelles conséquences sur les racines de P peut-on tirer du fait que P soit pair et à coefficients réels ?
3. En déduire la décomposition de P en facteurs irréductibles, dans $\mathbb{C}[X]$ et dans $\mathbb{R}[X]$.

Formule de Taylor pour les polynômes. —

Exercice 3.32. — ★☆☆

1. (a) Montrer qu'il existe un polynôme $P \in \mathbb{R}[X]$ de degré 3 vérifiant : $P(1) = 3$, $P'(1) = 5$ et $P''(1) = 1$.
   (b) Existe-t-il plusieurs polynômes vérifiant la propriété ci-dessus ?
2. Même question si l'on fixe un entier $n > 3$ et si l'on cherche P de degré n vérifiant : $P(1) = 3$, $P'(1) = 5$ et $P''(1) = 1$.
3. Même question si l'on fixe un entier $n \geq 3$ et si l'on cherche P de degré n vérifiant : $P(1) = 3$, $P'(1) = 5$ et $P''(1) = 1$ et $P^{(k)}(1) = 1$ pour tout $k \in \{3, \dots, n\}$.

Exercice 3.33. — ★★☆
Dans cet exercice, on fixe $n \in \mathbb{N}^\star$ et on considère le polynôme $P(X) = \frac{X^n(4-2X)^n}{n!}$

1. Montrer que les $(n-1)$ premières dérivées de P s'annulent toutes en $x = 0$ ainsi qu'en $x = 2$.
2. Écrire la formule de Taylor pour P au point 0. Écrire la formule de Taylor pour P au point 2.
3. En déduire que toutes les dérivées de P ont des valeurs entières en $x = 0$ et en $x = 2$.

CHAPITRE 4

MATRICES

1. Vocabulaire de base

1.1. Définition et notations. —

Définition 4.1 – Matrice à n lignes et p colonnes à coefficients réels ou complexes

Soient n et p deux entiers naturels non nuls. Une *matrice à n lignes et p colonnes à coefficients réels* (resp. complexes) est un tableau de la forme

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & \dots & a_{2p} \\ \vdots & & & & \vdots \\ a_{n1} & a_{n2} & \dots & \dots & a_{np} \end{pmatrix}$$

où $a_{11}, a_{12}, \dots, a_{np}$ sont des nombres réels (resp. complexes), appelés les *coefficients* de A.

Par exemple, la matrice $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 0 & -1 & 8 & 9 \end{pmatrix}$ est une matrice de format 3×4 (trois lignes, quatre colonnes) à coefficients réels.

Notation **$\mathcal{M}_{n,p}(\mathbb{R})$.** — On notera $\mathcal{M}_{n,p}(\mathbb{R})$ l'ensemble des matrices de format $n \times p$ à coefficients réels, et $\mathcal{M}_{n,p}(\mathbb{C})$ l'ensemble des matrices de format $n \times p$ à coefficients complexes.

Par exemple, on a $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \in \mathcal{M}_{2,3}(\mathbb{R})$, alors que $\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \in \mathcal{M}_{3,2}(\mathbb{R})$. On a bien sûr $\begin{pmatrix} 5i & -3 \\ -2 & 4+8i \end{pmatrix} \in \mathcal{M}_{2,2}(\mathbb{C})$. Quant à la matrice $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, on peut la voir comme un élément de $\mathcal{M}_{2,2}(\mathbb{R})$, mais aussi si nécessaire comme un élément de $\mathcal{M}_{n,p}(\mathbb{C})$.

Notation **(a_{ij}).** — Lorsque $A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{11} & a_{12} & \dots & a_{1p} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{np} \end{pmatrix}$, on écrira $A = (a_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$. Ainsi, écrire $A = (i+j)_{\substack{1 \leq i \leq 3 \\ 1 \leq j \leq 2}}$, c'est dire que A est la matrice à 3 lignes et 2 colonnes suivantes : $A = \begin{pmatrix} 2 & 3 & 4 \\ 3 & 4 & 5 \\ 4 & 5 & 6 \end{pmatrix}$.

Attention. — Veiller à ne pas confondre l'indice de ligne et l'indice de colonne, ou le nombre de lignes et le nombres de colonnes : dans ce cours, on écrira toujours *l'indice de ligne en premier.*

Définition 4.2 – Matrice-ligne, matrice-colonne

Soit $n \in \mathbb{N}^\star$.
Lorsque A est une matrice $n \times 1$ (à n lignes et une colonne), on dit que A est une *matrice-colonne à n coefficients.*
Lorsque A est une matrice $1 \times n$ (à une ligne et n colonnes), on dit que A est une *matrice-ligne à n coefficients.*

Par exemple, la matrice $\begin{pmatrix} 1 & 2 & 3 & 4 \end{pmatrix}$ est une matrice-ligne à quatre coefficients, tandis que $\begin{pmatrix} 1 \\ 8 \\ -4 \end{pmatrix}$ est une matrice-colonne à trois coefficients.

Définition 4.3 – Matrice carrée, coefficients diagonaux

Soit $n \in \mathbb{N}^\star$. Lorsque A est une matrice $n \times n$, on dit que A est une *matrice carrée de taille n.*
Si $A = (a_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$, les coefficients a_{11}, a_{22}, …, a_{nn} sont appelés *coefficients diagonaux de A.* On dit qu'ils sont situés sur la *diagonale* de A.

Notation. — On notera $\mathcal{M}_n(\mathbb{R})$ plutôt que $\mathcal{M}_{n,n}(\mathbb{R})$. Par exemple, la matrice $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ appartient à $\mathcal{M}_3(\mathbb{R})$. Ses coefficients diagonaux sont 1, 5 et 9.

Par ailleurs, Si A est une matrice carrée de taille n, on notera souvent $A = (a_{ij})_{1 \leq i,j \leq n}$ plutôt que $A = (a_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$. Par exemple, si $A = (i-j)_{1 \leq i,j \leq n}$, alors $A = \begin{pmatrix} 0 & -1 & -2 \\ 1 & 0 & -1 \\ 2 & 1 & 0 \end{pmatrix}$.

Définition 4.4 – Matrice diagonale

Soit A une matrice carrée de taille $n \times n$, où $n \in \mathbb{N}^\star$. On dit que A est *diagonale* lorsque A est de la forme

$$A = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & \dots & a_{nn} \end{pmatrix}$$

où $a_{11}, \dots, a_{nn}$ sont des nombres réels ou complexes.
Si on note $A = (a_{ij})_{1 \leq i,j \leq n}$, dire que A est diagonale revient à dire qu'on a $a_{ij} = 0$ dès que $i \neq j$.

On remarquera que dans la définition ci-dessus, on impose que les coefficients non diagonaux soient nuls, mais on n'impose rien sur les coefficients a_{11}, a_{22}, …, a_{nn} : ils peuvent très bien être nuls aussi. Ainsi, la matrice $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ est diagonale.

***Remarque 4.5* (Notation diag$(\cdots)$).** — Si α_1, α_2, …, α_n sont des nombres réels ou complexes, on note $\mathrm{diag}(\alpha_1, \ldots, \alpha_n)$ pour la matrice $\begin{pmatrix} \alpha_1 & 0 & \ldots & 0 \\ 0 & \alpha_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \ldots & \ldots & \alpha_n \end{pmatrix}$.

***Exemple 4.6* (Matrice identité).** — Si $n \in \mathbb{N}^\star$, on notera

$$I_n = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \ldots & \ldots & 1 \end{pmatrix} = \mathrm{diag}(1, \ldots, 1).$$

Cette matrice est l'une de celles qui apparaîtront le plus souvent dans toute discussion sur les matrices, pour des raisons expliquées plus bas.

Définition 4.7 – Matrice triangulaire supérieure

Soit A une matrice carrée de taille n. On dit que A est *triangulaire supérieure* si A est de la forme

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \ldots & a_{1n} \\ 0 & a_{22} & a_{23} & & a_{2n} \\ 0 & 0 & a_{33} & & a_{3n} \\ \vdots & & & \ddots & \vdots \\ 0 & \ldots & \ldots & 0 & a_{nn} \end{pmatrix},$$

autrement dit, lorsque pour tout i, j de $\{1, \ldots, n\}$ vérifiant $i > j$, on a $a_{ij} = 0$.

On dit de même que A est *triangulaire inférieure* si A est de la forme

$$\begin{pmatrix} a_{11} & 0 & & \ldots & 0 \\ a_{21} & a_{22} & 0 & & 0 \\ a_{31} & a_{32} & a_{33} & & \vdots \\ \vdots & & & \ddots & 0 \\ a_{n1} & \ldots & & \ldots & a_{nn} \end{pmatrix},$$

autrement dit, lorsque pour tout i, j de $\{1, \ldots, n\}$ vérifiant $i < j$, on a $a_{ij} = 0$

Par exemple, les matrices

$$B = \begin{pmatrix} 4 & 5 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 2 \end{pmatrix} \text{ et } C = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 0 & 5 \\ 0 & 0 & 2 \end{pmatrix} \text{ sont des matrices triangulaires supérieures,}$$

tandis que

$$D = \begin{pmatrix} 4 & 0 & 0 \\ 5 & 1 & 0 \\ 0 & 2 & 2 \end{pmatrix} \text{ est une matrice triangulaire inférieure.}$$

1.2. Somme et produit par un scalaire. —

Définition 4.8 – Somme de deux matrices

Soient A et B deux matrices à coefficients réels ou complexes.
On suppose que A et B ont le même format $n \times p$ où $(n,p) \in (\mathbb{N}^\star)^2$.
On définit la matrice $A+B$ comme suit :
si $A = (a_{ij})_{\substack{1\leq i\leq n\\1\leq j\leq p}}$ et $B = (b_{ij})_{\substack{1\leq i\leq n\\1\leq j\leq p}}$, alors $A+B = (a_{ij}+b_{ij})_{\substack{1\leq i\leq n\\1\leq j\leq p}}$

Par exemple, la matrice $\begin{pmatrix}1&2&3\\4&5&6\\7&8&9\end{pmatrix} + \begin{pmatrix}2&2&2\\-1&-1&-1\\10&10&10\end{pmatrix}$ n'est autre que $\begin{pmatrix}2&3&4\\3&4&5\\17&18&19\end{pmatrix}$.

Définition 4.9 – Multiplication d'une matrice par un coefficient réel ou complexe

Soit A une matrice à coefficients réels ou complexes et λ un nombre réel ou complexe.
On définit une matrice λA comme suit : si $A = (a_{ij})_{\substack{1\leq i\leq n\\1\leq j\leq p}}$, alors $\lambda A = (\lambda a_{ij})_{\substack{1\leq i\leq n\\1\leq j\leq p}}$

Par exemple, si $A = \begin{pmatrix}1&2&3\\3&2&1\\5&5&5\end{pmatrix}$, alors $5A = \begin{pmatrix}5&10&15\\15&10&5\\25&25&25\end{pmatrix}$.

Proposition 4.10 – Propriétés simples de ces deux opérations

Soient A, B, C trois matrices à coefficients réels ou complexes de taille $n \times p$, et λ, μ deux nombre réels ou complexes. On a les égalités suivantes :

- $A+B = B+A$;
- $(A+B)+C = A+(B+C)$;
- $A + \mathbf{0}_{n,p} = A$ *(où $\mathbf{0}_{n,p}$ est la matrice nulle de format $n \times p$)* ;
- $\lambda(A+B) = (\lambda A)+(\lambda B)$, $(\lambda+\mu)A = (\lambda A)+(\mu A)$ *et* $(\lambda\mu)A = \lambda(\mu A)$;
- $1A = A$ *et* $0A = \mathbf{0}_{n,p}$

2. Produit de matrices

2.1. Définition et premières remarques. —

Dans ce paragraphe, nous définissons une matrice AB lorsque A et B sont des matrices vérifiant une certaine condition de compatibilité.

Nous commençons par définir le produit dans le cas très particulier où A est une matrice-ligne, où B est une matrice-colonne, et où il y a autant de coefficients dans A que dans B.

Définition 4.11 – Produit ligne-colonne

Soit $n \in \mathbb{N}^\star$. Soit A une matrice-ligne à n coefficients

$$A = \begin{pmatrix} a_1 & a_2 & \dots & a_n \end{pmatrix}$$

et B une matrice-colonne à n coefficients

$$B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

On définit le produit AB comme la matrice 1×1 dans laquelle l'unique coefficient est donné par la somme $a_1b_1 + \cdots + a_nb_n = \sum_{i=1}^{n} a_ib_i$.

Par exemple, si x, y, z sont trois réels et si $A = \begin{pmatrix} x & y & z \end{pmatrix}$ et $B = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, alors $AB = \begin{pmatrix} x + 2y + 3z \end{pmatrix}$.

On commettra souvent l'abus de notation consistant à identifier une matrice 1×1 à son unique coefficient : dans l'exemple précédent, on pourra voir AB soit comme une matrice 1×1, soit comme le nombre $x+2y+3z$.

Attention. — Pour qu'un tel produit soit bien défini, il faut que les tailles se correspondent. Si $A = \begin{pmatrix} 1 & 2 & 3 & 4 \end{pmatrix}$ et $B = \begin{pmatrix} 4 \\ 2 \\ 1 \end{pmatrix}$, le produit AB n'est pas défini.

Définition 4.12 – Produit de deux matrices

Soient n, p, q, m quatre entiers naturels non nuls. Soit A une matrice $n \times p$ à coefficients réels ou complexes et B une matrice $m \times q$ à coefficients réels ou complexes.
On suppose vérifiée la condition suivante : $p = m$, c'est-à-dire

le nombre de colonnes de A est égal au nombre de lignes de B.

On note alors AB la matrice $(m_{ij})_{\substack{1\leq i\leq n \\ 1\leq j\leq q}}$ définie par : pour tout $i \in \{1, \dots, n\}$ et $j \in \{1, \dots, q\}$,

$$m_{ij} = L_i^A C_j^B$$

où L_i^A est la i-ème ligne de A et C_j^B est la j-ème colonne de B. On a en fait

$$m_{ij} = \sum_{k=1}^{p} a_{ik}b_{kj}.$$

Attention. — Le produit de deux matrices n'est pas toujours bien défini : il y a une condition de compatibilité sur les tailles des matrices.

Le produit d'une matrice $n \times p$ et d'une matrice $m \times q$ n'est défini que si $m = p$.
Dans ce cas, c'est une matrice $n \times q$.

Par exemple,

- si A est une matrice 3×4 et si B est une matrice 5×2, le produit AB n'a pas de sens.
- si A est une matrice 3×4 et B est une matrice 4×7, alors le produit AB est bien défini, et c'est une matrice 3×7.

Remarque 4.13 **(Disposition pratique des calculs).** — Pour éviter les erreurs , il est conseillé d'adopter la présentation suivante des calculs, proposée ici sur un exemple : si $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \end{pmatrix}$ et $B = \begin{pmatrix} 2 & 1 & 0 & 1 \\ 0 & 1 & 0 & 3 \\ 1 & 0 & 1 & 1 \end{pmatrix}$ pour calculer le produit AB, on peut disposer A et B comme suit :

$$\begin{pmatrix} 2 & 1 & 0 & 1 \\ 0 & 1 & 0 & 3 \\ 1 & 0 & 1 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \end{pmatrix} \quad \begin{pmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix}$$

ce qui permet de visualiser la taille de la matrice AB, puis on remplit les coefficients de AB « place par place » par exemple, pour trouver le coefficient sur la première ligne et la deuxième colonne, on utilise la ligne L_1^A et la colonne C_2^B :

$$\begin{pmatrix} 2 & 1 & 0 & 1 \\ 0 & 1 & 0 & 3 \\ 1 & 0 & 1 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \end{pmatrix} \quad \begin{pmatrix} \cdot & 5 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix}$$

On trouve alors, après 8 calculs : $AB = \begin{pmatrix} 5 & 3 & 3 & 10 \\ 2 & 1 & 2 & 5 \end{pmatrix}$.

Le produit matriciel vérifie plusieurs propriétés qui en facilitent la manipulation :

Proposition 4.14 – Propriétés du produit

Soient A, B, C trois matrices à coefficients réels ou complexes, alors

- *on a $(AB)C = A(BC)$ dès que les produits qui interviennent dans cette égalité sont bien définis ;*
- *on a $A(B + C) = AB + AC$ et $(A + B)C = AC + BC$ dès que les produits qui interviennent dans ces égalités sont bien définis ;*
- *pour tout réel λ, on a $A(\lambda B) = (\lambda A)B = \lambda(AB)$ dès que le produit AB est bien défini.*

Il y a par contre des différence cruciales avec les opérations de type « produit » que vous avez manipulées jusqu'ici, qui concernent le rapport entre les produits AB et BA.

1. Le produit AB peut avoir un sens alors que BA n'en a pas : c'est le cas lorsque A est une matrice (n, p) et B une matrice (p, q) avec $n \neq q$.

2. Même si les produits AB et BA ont un sens, les matrices AB et BA ne sont en général pas du même format, donc certainement pas égales ; par exemple $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 \\ -1 & -1 \\ 0 & 1 \end{pmatrix}$ alors :
$$AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad BA = \begin{pmatrix} 1 & 1 & 1 \\ -2 & -2 & -2 \\ 1 & 1 & 1 \end{pmatrix}$$
3. Enfin même dans le cas a priori le plus favorable, c'est-à-dire si A et B sont des matrices carrées de même ordre n, les deux matrices AB et BA sont aussi des matrices carrées de même ordre n, mais en général elles ne sont pas égales. Par exemple $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$ alors :
$$AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad BA = \begin{pmatrix} 2 & 2 \\ -2 & -2 \end{pmatrix}$$

Proposition 4.15 – En général, on a $AB \neq BA$

Si A et B sont deux matrices, il est possible que le produit AB existe et pas le produit BA.
Même lorsque les produits AB et BA sont bien définis, on n'a généralement pas $AB = BA$.

Remarque 4.16. — Lorsque A et B sont deux matrices vérifiant $AB = BA$, on dit que A et B *commutent*.

Par ailleurs, le troisième exemple ci-dessus indique une autre différence cruciale entre la multiplication des matrices et celle des nombres réels ou complexes ou celle des polynômes :

Attention. — On peut très bien avoir $AB = 0$ sans qu'aucune des deux matrices A, B soit nulle.

Une conséquence de cette propriété est qu'on ne peut pas « simplifier par A lorsque $A \neq 0$ » :

Si A, B et C sont trois matrices et si $AB = AC$, alors on ne peut pas en déduire $B = C$.

En effet, si $AB = AC$, alors $A(B - C) = 0$, mais comme on vient de le voir, cela n'implique pas forcément $B - C = 0$. Voici un exemple concret : si $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 \\ -1 & -1 \end{pmatrix}$, $C = \begin{pmatrix} 2 & 2 \\ -2 & -3 \end{pmatrix}$ alors on a : $AB = AC = \begin{pmatrix} 0 & -1 \\ 0 & -1 \end{pmatrix}$ et pourtant $B \neq C$.

Relevons deux d'usage fréquent où la forme générale du produit AB est connue d'avance à partir des propriétés de A et B :

Proposition 4.17 – Produit de deux matrices diagonales, de deux matrices triangulaires

Soient A et B deux matrices carrées de format $n \times n$.

- *Si A et B sont diagonales, alors AB aussi. De plus, si $A = \operatorname{diag}(\alpha_1, \ldots, \alpha_n)$ et $B = \operatorname{diag}(\beta_1, \ldots, \beta_n)$, alors $AB = \operatorname{diag}(\alpha_1\beta_1, \ldots, \alpha_n\beta_n)$.*
- *Si A et B sont triangulaires supérieures, alors AB aussi. De plus, dans ce cas les coefficients diagonaux de AB s'expriment simplement en fonction de ceux de A et de B.*

Précisons un peu le second point : si A et B sont triangulaires supérieures et si nous écrivons $A = \begin{pmatrix} \alpha_1 & \star & \dots & \star \\ 0 & \alpha_2 & & \star \\ \vdots & & \ddots & \vdots \\ 0 & \dots & \dots & \alpha_n \end{pmatrix}$ et $B = \begin{pmatrix} \beta_1 & \star & \dots & \star \\ 0 & \beta_2 & & \star \\ \vdots & & \ddots & \vdots \\ 0 & \dots & \dots & \beta_n \end{pmatrix}$, où les étoiles sont des coefficients non précisés (pas forcément égaux les uns aux autres), alors $AB = \begin{pmatrix} \alpha_1\beta_1 & \star & \dots & \star \\ 0 & \alpha_2\beta_2 & & \star \\ \vdots & & \ddots & \vdots \\ 0 & \dots & \dots & \alpha_n\beta_n \end{pmatrix}$, où les étoiles désignent des coefficients « compliqués » qui ne s'expriment pas généralement de façon simple à partir de ceux de A et B.

2.1.1. *Cas du produit matrice-vecteur.* —

On sera souvent amenés à considérer des produits matriciels de la forme AX où $A = \begin{pmatrix} a_{11} & \dots & a_{1p} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{np} \end{pmatrix}$ est une matrice $n \times p$ quelconque et où $X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$ est une matrice-colonne telle que le produit AX soit bien défini. Comme A est une matrice $n \times p$ et X une matrice $p \times 1$, le produit AX est alors une matrice $n \times 1$, donc une matrice-colonne comportant n coefficients. Pour des raisons qui apparaîtront clairement au début second semestre, on dira parfois « le vecteur X » plutôt que « la matrice-colonne X ».

Le résultat qui vient n'est qu'une reformulation de la définition du produit matriciel, mais il est utile.

Proposition 4.18 – Produit matrice-vecteur = combinaison des colonnes de la matrice

Soient $A = \begin{pmatrix} a_{11} & \dots & a_{1p} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{np} \end{pmatrix}$ *une matrice* $n \times p$ *quelconque et* $X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$ *une matrice-colonne comportant* p *coefficients. Notons* $C_1^A = \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix}$, …, $C_p^A = \begin{pmatrix} a_{1p} \\ \vdots \\ a_{np} \end{pmatrix}$ *les colonnes de* A*. Le produit* AX *peut être décrit par l'égalité suivante :*

$$A \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} = x_1 C_1^A + \cdots + x_p C_p^A.$$

Exemple 4.19. — Si $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$, alors

- Le produit $A \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ est une matrice-colonne donnée par la somme des colonnes de A : c'est $\begin{pmatrix} 6 \\ 15 \\ 24 \end{pmatrix}$.

- Le produit $A\begin{pmatrix}1\\0\\-1\end{pmatrix}$ est la différence entre la première et la troisième colonne de A : c'est $\begin{pmatrix}-2\\-2\\-2\end{pmatrix}$.

2.1.2. *Image et noyau d'une matrice.* —

Définition 4.20 – Image d'une matrice à coefficients réels

Soient n et p deux entiers naturels non nuls. Soit A une matrice de $\mathcal{M}_{n,p}(\mathbb{R})$.
On appelle *ensemble image* de A, et on note $\mathbf{Im}(A)$, le sous-ensemble de $\mathcal{M}_{n,1}(\mathbb{R})$ formé des matrices-colonnes Y telles qu'il existe une matrice-colonne $X \in \mathcal{M}_{p,1}(\mathbb{R})$ vérifiant $AX = Y$:

$$\mathbf{Im}(A) = \left\{ Y = \begin{pmatrix}y_1\\\vdots\\y_n\end{pmatrix} \in \mathcal{M}_{n,1}(\mathbb{R}) \quad / \quad \exists \begin{pmatrix}x_1\\\vdots\\x_p\end{pmatrix} \in \mathcal{M}_{p,1}(\mathbb{R}) \ : \ \begin{pmatrix}y_1\\\vdots\\y_n\end{pmatrix} = A\begin{pmatrix}x_1\\\vdots\\x_p\end{pmatrix} \right\}.$$

Remarque 4.21. — Compte tenu du §2.1.1, dire que Y appartient à $\mathbf{Im}(A)$ revient à dire qu'il est possible d'exprimer Y comme une combinaison linéaire des colonnes de A. En effet, si l'on note $C_1^A, \dots, C_p^A$ les colonnes de A, dire qu'on peut trouver des réels $x_1, \dots, x_p$ vérifiant $Y = A\begin{pmatrix}x_1\\\vdots\\x_p\end{pmatrix}$ est équivalent dire qu'on peut écrire $Y = x_1C_1^A + \cdots + x_pC_p^A$ avec $x_1, \dots, x_p$ réels.

Exemple 4.22. — Considérons la matrice $A = \begin{pmatrix}1&2&3\\1&2&3\\6&7&8\end{pmatrix}$. Fixons une colonne $Y = \begin{pmatrix}a\\b\\c\end{pmatrix}$. Dire que Y appartient à $\mathbf{Im}(A)$, c'est dire qu'il existe une colonne $X = \begin{pmatrix}x\\y\\z\end{pmatrix}$ vérifiant $\begin{pmatrix}a\\b\\c\end{pmatrix} = A\begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}x+2y+3z\\x+2y+3z\\6x+7y+8z\end{pmatrix}$.

En observant les deux premiers coefficients, on constate que l'existence d'un tel $\begin{pmatrix}x\\y\\z\end{pmatrix}$ n'est imaginable que dans le cas où $a = b$. On en déduit, par exemple, que $\begin{pmatrix}1\\-3\\0\end{pmatrix}$ ne peut pas appartenir à $\mathbf{Im}(A)$.

En revanche, toutes les colonnes $\begin{pmatrix}a\\b\\c\end{pmatrix}$ qui peuvent s'exprimer comme une combinaison des colonnes de A appartiennent à $\mathbf{Im}(A)$. Par exemple, $\begin{pmatrix}1\\1\\6\end{pmatrix} = A\begin{pmatrix}1\\0\\0\end{pmatrix}$ appartient à $\mathbf{Im}(A)$; de même $\begin{pmatrix}-1\\-1\\-3\end{pmatrix} = A\begin{pmatrix}1\\1\\0\end{pmatrix}$ appartient à $\mathbf{Im}(A)$, etc.

Définition 4.23 – Noyau d'une matrice à coefficients réels

Soient n et p deux entiers naturels non nuls. Soit A une matrice de $\mathcal{M}_{n,p}(\mathbb{R})$.
On appelle *noyau* de A, et on note $\mathbf{Ker}(A)$, le sous-ensemble de $\mathcal{M}_{p,1}(\mathbb{R})$ formé des matrices-colonnes X vérifiant $AX = \mathbf{0}_{n,1}$:

$$\mathbf{Ker}(A) = \left\{ X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} \in \mathcal{M}_{p,1}(\mathbb{R}) \quad / \quad A \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} = \mathbf{0}_{n,1} \right\}.$$

Exemple 4.24. — Si $A = \begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix}$, alors $A \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ (voir le §2.1.1). Ainsi, on a $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \in \mathbf{Ker}(A)$.

Remarque 4.25. — Si l'on note C_1^A, …, C_p^A les colonnes de A, et si l'on fixe $\begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} \in \mathcal{M}_{p,1}(\mathbb{R})$, alors dire que $A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \mathbf{0}_{n,1}$ revient à dire que la combinaison $x_1 C_A^1 + \cdots + x_p C_p^A$ est nulle. Cette remarque est très pratique pour trouver des éléments de $\mathbf{Ker}(A)$ en repérant des relations entre les colonnes de A.

Par exemple, si $A = \begin{pmatrix} 1 & -3 & 0 \\ 2 & -6 & 5 \\ 4 & -12 & 1 \end{pmatrix}$, alors on constate que la colonne C_2^A est égale à (-3) fois la colonne C_1^A ; ainsi $3C_1^A + C_2^A + 0C_3^A = \mathbf{0}$, d'où $\begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} \in \mathbf{Ker}(A)$.

2.1.3. *Calculs colonne par colonne.* —

Mentionnons une conséquence, qui porte sur les produits de matrices quelconques, de la règle pour le produit matrice-vecteur vue au §2.1.1 :

Proposition 4.26 – Calcul des colonnes du produit

Soient A une matrice de taille $n \times p$ et B une matrice de taille $p \times q$.
Si nous notons $C_1^B, \ldots, C_q^B$ les colonnes de B, alors les colonnes de la matrice AB sont les produits AC_1^B, AC_2^B, …, AC_p^B.

Considérons $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{pmatrix}$ et $B = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$. Pour calculer AB, procédons colonne par colonne :

- La première colonne de AB est $AC_1^B = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{pmatrix} \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 5 \\ 8 \end{pmatrix}$;

- La deuxième colonne de AB est $AC_2^B = \begin{pmatrix}1&2&3\\2&3&1\\3&1&2\end{pmatrix}\begin{pmatrix}0\\0\\0\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}$;
- La troisième colonne de AB est $AC_3^B = \begin{pmatrix}1&2&3\\2&3&1\\3&1&2\end{pmatrix}\begin{pmatrix}0\\1\\0\end{pmatrix} = \begin{pmatrix}2\\3\\1\end{pmatrix}$.

On constate donc que $AB = \begin{pmatrix}5&0&2\\5&0&3\\8&0&1\end{pmatrix}$.

Avec un peu d'habitude, cette stratégie de calcul peut permettre de repérer des simplifications : par exemple si l'une des colonnes de B est nulle, on constate qu'il y a pour A une colonne nulle. De plus, l'habitude du produit matrice-vecteur peut conduire à effectuer vite les produits AC_1^B, AC_2^B, …, AC_p^B.

3. Trace et transposée

3.1. Transposée d'une matrice quelconque. —

Définition 4.27 – Transposée

Soit A une matrice $n \times p$ à coefficients réels ou complexes.
On appelle matrice transposée de A, et on note A^T, la matrice de format $p \times n$ définit comme suit : si $A = (a_{ij})_{\substack{1\leq i\leq n\\1\leq j\leq p}}$, alors $A^T = (a_{ji})_{\substack{1\leq j\leq p\\1\leq i\leq n}}$.

Exemple 4.28. —

- Si $A = \begin{pmatrix}1&2&3&4\end{pmatrix}$, alors $A^T = \begin{pmatrix}1\\2\\3\\4\end{pmatrix}$.
- Si $A = \begin{pmatrix}1&2&3&4\\5&6&7&8\end{pmatrix}$, alors $A^T = \begin{pmatrix}1&5\\2&6\\3&7\\4&8\end{pmatrix}$.
- Si $A = \begin{pmatrix}1&2&3\\4&5&6\\7&8&9\end{pmatrix}$, alors $A^T = \begin{pmatrix}1&4&7\\2&5&8\\3&6&9\end{pmatrix}$.

Proposition 4.29 – Propriétés simples de la transposition

Soient A et B deux matrices de même taille ; soit λ un nombre réel. On a les égalités suivantes :

- $(A^T)^T = A$,
- $(A+B)^T = A^T + B^T$,
- $(\lambda A)^T = \lambda(A^T)$.

Proposition 4.30 – Transposée d'un produit

Si A et B sont deux matrices telles que le produit AB soit bien défini, alors on a l'égalité

$$(AB)^T = B^T A^T.$$

Définition 4.31 – Matrices symétriques et antisymétriques

Soit A une matrice carrée de $\mathcal{M}_n(\mathbb{R})$.

- On dit que A est symétrique lorsqu'on a $A^T = A$.
- On dit que A est antisymétrique lorsqu'on a $A^T = (-A)$.

Exemple 4.32. —

- Les matrices $\begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix}$ et $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 6 \\ 3 & 6 & 9 \end{pmatrix}$ sont symétriques.
- La matrice $\begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & 3 \\ -2 & -3 & 0 \end{pmatrix}$ est antisymétrique.
- Par contre, la matrice $\begin{pmatrix} 4 & 1 & 2 \\ -1 & 4 & 3 \\ -2 & -3 & 4 \end{pmatrix}$ n'est pas antisymétrique : si l'on observe les matrices $A^T = \begin{pmatrix} 4 & -1 & -2 \\ 1 & 4 & -3 \\ 2 & 3 & 4 \end{pmatrix}$ et $-A = \begin{pmatrix} -4 & -1 & -2 \\ 1 & -4 & -3 \\ 2 & 3 & -4 \end{pmatrix}$, on constate que les coefficients diagonaux sont différents.

Remarque 4.33. — Pour qu'une matrice carrée A puisse être antisymétrique, il est nécessaire que tous les coefficients diagonaux de A sont nuls, comme dans l'exemple ci-dessus. En effet, si $A = (a_{ij})_{1\leq i,j\leq n}$ et si l'on fixe $i \in \{1, \dots, n\}$, alors le i-ème coefficient diagonal de A^T est a_{ii} et le coefficient diagonal de $-A$ est $-a_{ii}$: si $A^T = -A$ on doit nécessairement avoir $a_{ii} = 0$.

Démonstration des proposition 4.29 et 4.30. —

- Si A est une matrice $n \times p$ et si nous notons $A = (a_{ij})_{\substack{1\leq i\leq n \\ 1\leq j\leq p}}$, la matrice A^T a pour terme général $(a_{ji})_{\substack{1\leq j\leq p \\ 1\leq i\leq n}}$ et pour format $p \times n$. Donc la matrice $(A^T)^T$ a pour terme général $(a_{ij})_{\substack{1\leq i\leq n \\ 1\leq j\leq p}}$ et pour format $n \times p$. On en déduit l'égalité $(A^T)^T = A$.
- Si A et B sont de format $n \times p$ et si nous notons $A = (a_{ij})_{\substack{1\leq i\leq n \\ 1\leq j\leq p}}$ et $B = (b_{ij})_{\substack{1\leq i\leq n \\ 1\leq j\leq p}}$, alors la matrice $A + B$ a pour terme général $(a_{ij} + b_{ij})_{\substack{1\leq i\leq n \\ 1\leq j\leq p}}$ et pour format $n \times p$, et donc la matrice $(A + B)^T$ a pour terme général $(a_{ji} + b_{ji})_{\substack{1\leq j\leq p \\ 1\leq i\leq n}}$ et pour format $p \times n$. D'autre part, les matrices A^T et B^T ont pour terme général respectivement $(a_{ji})_{\substack{1\leq j\leq p \\ 1\leq i\leq n}}$ et $(b_{ji})_{1\leq i\leq n}^{1\leq j\leq p}$, donc la matrice $A^T + B^T$ a pour terme général $(a_{ji} + b_{ji})_{\substack{1\leq j\leq p \\ 1\leq i\leq n}}$ et pour format $p \times n$. Les matrices $(A + B)^T$ et $A^T + B^T$ sont donc égales.
- On montre de même l'égalité $(\lambda A)^T = \lambda A^T$.
- Si A est une matrice $n \times p$ et si B est une matrice $p \times q$ et si nous notons $A = (a_{ij})_{\substack{1\leq i\leq n \\ 1\leq j\leq p}}$ et $B = (b_{ij})_{\substack{1\leq i\leq p \\ 1\leq j\leq q}}$, alors AB est une matrice $n \times q$: il s'agit de la matrice $(\sum_{k=1}^{p} a_{ik}b_{kj})_{1\leq j\leq q}^{1\leq i\leq n}$. Par

conséquent, la matrice $C = (AB)^T$ a pour format $q \times n$ et pour terme général

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

D'autre part, les matrices A^T et B^T ont pour format respectif $p \times n$ et $q \times p$ et pour terme général $(a_{ji})_{\substack{1 \leq j \leq p \\ 1 \leq i \leq n}}$ et $(b_{ji})_{\substack{1 \leq j \leq q \\ 1 \leq i \leq p}}$. Le produit $D = B^T A^T$ existe donc, a pour format $q \times n$ et pour terme général

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

Les matrices C et D ont même format $q \times n$, et on a $c_{ij} = d_{ij}$ pour tous i, j. On a donc $C = D$, comme espéré.

□

3.2. Trace d'une matrice carrée. —

Définition 4.34 – Trace

Soit A une matrice carrée de format $n \times n$.
On appelle *trace* de A, et on note $\mathrm{Tr}(A)$, le nombre donné par la somme des coefficients diagonaux de A : si $A = (a_{ij})_{1 \leq i,j \leq n}$, alors

$$\mathrm{Tr}(A) = a_{11} + a_{22} + \cdots + a_{nn}.$$

Exemple 4.35. — Si $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$, alors $\mathrm{Tr}(A) = 1 + 5 + 9 = 15$.

Proposition 4.36 – Propriétés simples de la notion de trace

Soient A et B deux matrices de même taille ; soit λ un nombre réel. On a les égalités suivantes :

- $\mathrm{Tr}(A + B) = \mathrm{Tr}(A) + \mathrm{Tr}(B)$;
- $\mathrm{Tr}(\lambda A) = \lambda \mathrm{Tr}(A)$;
- $\mathrm{Tr}(A^T) = \mathrm{Tr}(A)$.

Démonstration. — Si A et B sont deux matrices $n \times p$ et si nous notons $A = (a_{ij})_{1 \leq i,j \leq n}$ et $B = (b_{ij})_{1 \leq i,j \leq n}$, alors on a les égalités suivantes :

- $A + B = (a_{ij} + b_{ij})_{1 \leq i,j \leq n}$, donc $\mathrm{Tr}(A + B) = \sum_{i=1}^{n} (a_{ii} + b_{ii}) = \mathrm{Tr}(A) + \mathrm{Tr}(B)$;
- $\lambda A = (\lambda a_{ij})_{1 \leq i,j \leq n}$, donc $\mathrm{Tr}(\lambda A) = \sum_{i=1}^{n} \lambda a_{ii} = \lambda \mathrm{Tr}(A)$;
- $A^T = (a_{ji})_{1 \leq i,j \leq n}$, donc $\mathrm{Tr}(A^T) = \sum_{j=1}^{n} a_{jj} = \mathrm{Tr}(A)$.

□

L'utilité de la notion de trace repose essentiellement sur la remarque suivante :

Théorème 4.37 – Trace et produit

Si A et B sont deux matrices carrées de même taille, alors on a l'égalité

$$\mathrm{Tr}(AB) = \mathrm{Tr}(BA).$$

Ce résultat peut paraître surprenant : tout ce que nous avons vu jusqu'ici montre que si A et B sont deux matrices carrées de même taille, les matrices AB et BA sont généralement très différentes, et n'ont apparemment rien à voir l'une avec l'autre en général. Ce que montre la proposition ci-dessus, c'est qu'elles partagent tout de même certaines caractéristiques, dont la trace. Cette propriété de la trace la rend extrêmement utile dans beaucoup de discussions faisant intervenir le produit de matrices.

Démonstration. — Si A et B sont deux matrices $n\times n$ et si nous notons $A = (a_{ij})_{1\leq i,j\leq n}$ et $B = (b_{ij})_{1\leq i,j\leq n}$, alors AB est la matrice $(\sum_{k=1}^{n} a_{ik}b_{kj})_{1\leq i,j\leq n}$. Sa trace est la somme de ses coefficients diagonaux : on a donc $\mathrm{Tr}(AB) = \sum_{i=1}^{n}\sum_{k=1}^{n} a_{ik}b_{ki}$.

Par ailleurs, la matrice BA n'est autre que $(\sum_{k=1}^{n} b_{ik}a_{kj})_{1\leq i,j\leq n}$. Sa trace est la somme de ses coefficients diagonaux : on a donc $\mathrm{Tr}(BA) = \sum_{i=1}^{n}\sum_{k=1}^{n} b_{ik}a_{ki}$.

En comparant les expressions obtenues pour les sommes donnant $\mathrm{Tr}(AB)$ et $\mathrm{Tr}(BA)$, et en tenant compte du fait que les indices i, k y sont muets, on constate qu'elles sont égales. □

4. Puissances d'une matrice carrée

4.0.1. *Définition et exemples.* —

Définition 4.38 – Puissances d'une matrice

Soit A une matrice carrée de taille n.
On note $A^0 = I_n$, $A^1 = A$, $A^2 = AA$ et pour tout $k \in \mathbb{N}$, on note $A^{k+1} = AA^k = A^kA$.

Exemple 4.39. — Si $A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$, alors le calcul donne $A^2 = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$, $A^3 = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}$, etc.

Exemple 4.40. — Si A est une matrice carrée diagonale, et si l'on écrit

$$A = \mathrm{diag}(\alpha_1, \ldots, \alpha_n) = \begin{pmatrix} \alpha_1 & 0 & \ldots & 0 \\ 0 & \alpha_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \ldots & \ldots & \alpha_n \end{pmatrix}.$$

alors la proposition 4.17 montre que le calcul des puissances de A est très simple : pour tout $k \in \mathbb{N}$, on a

$$A^k = \mathrm{diag}(\alpha_1^k, \ldots, \alpha_n^k) = \begin{pmatrix} \alpha_1^k & 0 & \ldots & 0 \\ 0 & \alpha_2^k & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \ldots & \ldots & \alpha_n^k \end{pmatrix}.$$

Exemple 4.41. — Si A est une matrice triangulaire supérieure, alors la proposition 4.17 montre que pour tout $k \in \mathbb{N}$, la matrice A^k est aussi triangulaire supérieure.

De plus, en appliquant la remarque **??**, on constate que si $A = \begin{pmatrix} \alpha_1 & \star & \dots & \dots & \star \\ 0 & \alpha_2 & \star & & \star \\ 0 & 0 & \alpha_3 & & \star \\ \vdots & & & \ddots & \vdots \\ 0 & \dots & \dots & 0 & \alpha_n \end{pmatrix}$, alors pour tout $k \in \mathbb{N}$, on a $A^k = \begin{pmatrix} \alpha_1^k & \star & \dots & \dots & \star \\ 0 & \alpha_2^k & \star & & \star \\ \vdots & & & \ddots & \vdots \\ 0 & \dots & \dots & 0 & \alpha_n^k \end{pmatrix}$, où les étoiles désignent des coefficients ayant une expression « compliquée ».

Nous constatons donc que si A est triangulaire supérieure, il est possible d'exprimer simplement les coefficients diagonaux de A^k en fonction de ceux de A. En revanche, les coefficients de A^k situés strictement au-dessus de la diagonale n'ont pas d'expression « simple » qui soit systématique.

4.0.2. *Matrices nilpotentes.* —

Définition 4.42 – Matrice nilpotente, indice de nilpotence

Soit A une matrice carrée de taille n.
On dit que A est nilpotente s'il existe un entier $k \in \mathbb{N}^\star$ vérifiant $A^k = \mathbf{0}_{n,n}$.

Lorsque A est nilpotente, on appelle indice de nilpotence de A le plus petit entier p vérifiant : $A^{p-1} \neq 0$ et $A^p = \mathbf{0}_{n,n}$.

Exemple 4.43 **(Exemple fondamental : bloc de Jordan).** — Considérons la matrice $J = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$.

Le calcul donne $J^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ et $J^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. La matrice J est donc nilpotente d'indice 3.

Remarque 4.44. — Dire que A est nilpotente d'indice p revient à dire qu'on a $A^k = \mathbf{0}_{n,n}$ pour tout $k \geq p$ mais que $A^k \neq \mathbf{0}_{n,n}$ pour $k < p$. Dans l'exemple ci-dessus, les puissances J^0, J^1 et J^2 sont non-nulles, mais on a $J^k = \mathbf{0}$ dès que $k \geq 3$.

4.0.3. *Puissances de $I + A$.* —

Commençons par remarquer que si A et B sont deux matrices carrées de même taille et si l'on souhaite développer $(A + B)^2$, alors on doit tenir compte de l'ordre des produits : $(A + B)^2 = (A + B)(A + B) = A^2 + AB + BA + B^2$. En général, on a $AB \neq BA$, donc on ne peut pas « regrouper » AB et BA en $2AB$ comme dans le cas des nombres.

De même, le calcul de $(A+B)^3$ peut se faire en « développant tout » et on trouve $(A+B)^3 = A^3+AB^2+BAB+B^2A+A^2B+ABA+BA^2+B^3$... mais lorsque $AB \neq BA$, on ne peut pas regrouper davantage

les termes et on a $(A+B)^3 \neq A^3+3A^2B+3B^2A+B^3$. On retiendra que *la formule du binôme est fausse lorsque* $AB \neq BA$.

Cependant, lorsque A et B commutent, les regroupements qui sont à la base de la formule du binôme sont possibles :

Proposition 4.45 – Formule du « binôme-quand-ça-commute »

Soient A et B deux matrices carrées de même taille.

- *Si $AB = BA$. alors pour tout $n \in \mathbb{N}$, on a l'égalité $(A+B)^n = \sum_{k=0}^{n} \binom{n}{k} A^k B^{n-k}$.*
- *Si $AB \neq BA$, alors il n'existe pas de formule condensée pour $(A+B)^n$.*

***Exemple 4.46* (Puissances de $A+I$).** — Si A est une matrice carrée et si I est la matrice identité de même taille que A, alors les matrices A et I commutent. On peut donc appliquer la formule ci-dessus, et pour tout $n \in \mathbb{N}$, on a donc $(A+I)^n = \sum_{k=0}^{n} \binom{n}{k} A^k B^{n-k} = \sum_{k=1}^{n} \binom{n}{k} A^k$.

5. Matrices inversibles

5.1. Définition et premiers exemples. —

Définition 4.47

Soit A une matrice carrée de taille $n \times n$. On dit que A est inversible dans $\mathcal{M}_n(\mathbb{R})$ lorsqu'il existe une matrice $B \in \mathcal{M}_n(\mathbb{R})$ vérifiant les conditions suivantes :

$$AB = BA = I_n.$$

De plus, s'il existe une matrice B vérifiant ces conditions, alors il en existe une seule. Dans ce cas, on dit que B est l'inverse de A et on écrit $B = A^{-1}$.

Exemple 4.48. — Si $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 7 \end{pmatrix}$, alors la matrice $B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & \frac{1}{7} \end{pmatrix}$ vérifie $AB = BA = I_3$, donc A est inversible.

Exemple 4.49. — La matrice $A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ n'est pas inversible : si B est une matrice quelconque, on a toujours $AB = BA = \mathbf{0}_{3,3}$, donc il est impossible que AB soit égal à I_3.

***Exemple 4.50* (Un exemple de matrice non inversible).** — Vérifions que $J = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ n'est pas inversible. S'il existait une matrice K vérifiant $JK = KJ = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, la troisième colonne de KJ serait égale à $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$. Or, nous avons vu que la première colonne de KJ est égale à KC_1^J, où C_1^J est la

première colonne de J. Comme on a toujours $K\begin{pmatrix}0\\0\\0\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}$, il est impossible de choisir K de façon à avoir $K\begin{pmatrix}0\\0\\0\end{pmatrix} = \begin{pmatrix}1\\0\\0\end{pmatrix}$, donc impossible de trouver K vérifiant $KJ = I_3$. .

Justification de l'unicité de l'inverse. — Supposons qu'il existe deux matrices B_1 et B_2 vérifiant à la fois $AB_1 = B_1A = I_n$ et $AB_2 = B_2A = I_n$; vérifions qu'alors on a nécessairement $B_1 = B_2$. Pour cela, calculons B_2AB_1 de deux façons différentes :

- D'une part, on a $(B_1A) = I_n$, donc $B_1AB_1 = I_nB_1 = B_2$;
- D'autre part, on a $(AB_1) = I_n$, donc $B_1AB_2 = B_1I_n = B_1$.

On a donc nécessairement $B_1 = B_2$. □

5.2. Remarques théoriques. —

Attention. — On ne peut employer la notation A^{-1} que si on *sait déjà* que A est inversible.

Remarque 4.51. — A, B et C sont trois matrices carrées de même taille et si $AB = AC$, alors on ne peut pas en général en déduire $B = C$. Cette « simplification » est cependant possible dans le cas où A est inversible :

> Si A est inversible et si $AB = AC$, alors $B = C$.

En effet, si $AB = AC$ et si A est inversible, on peut multiplier les deux membres de l'égalité par A^{-1} : on obtient alors $A^{-1}(AB) = A^{-1}(AC)$, c'est-à-dire $(A^{-1}A)B = (A^{-1}A)C$; comme $A^{-1}A = I_n$, on obtient alors $B = C$.

Remarque 4.52. — Si on a $AB = BA = I_n$, alors A est inversible et $A^{-1} = B$. Mais l'égalité $AB = BA = I_n$ est « symétrique en A et B » : elle indique également que B est inversible et que $B^{-1} = A$. Nous constatons donc le fait suivant :

> Si A est inversible, alors A^{-1} est aussi inversible et $(A^{-1})^{-1} = A$.

Remarque 4.53. — Si A est inversible et si B est l'inverse de A, alors on a $AB = I_n$ et $BA = I_n$. Observons l'effet de la transposition sur cette égalité : en appliquant l'opération de transposition à chacun des termes de l'égalité, on obtient $(AB)^T = I_n^T$. Or, $I_n^T = I_n$ et $(AB)^T = B^TA^T$. On obtient donc $B^TA^T = I_n$. Le même raisonnement montre, partant de $BA = I_n$, que $A^TB^T = I_n$. Nous constatons donc que A^T est d'inversible, d'inverse B^T. Comme B est l'inverse de A, nous obtenons le résultat suivant :

> Si A est inversible, alors A^T est aussi inversible ; de plus $(A^T)^{-1} = (A^{-1})^T$.

Nous démontrerons au second semestre le résultat suivant, qui est très loin d'être évident :

Théorème 4.54 – Si $AB = I_n$, alors l'égalité $BA = I_n$ est automatique

Soient A et B deux matrices carrées de même taille. Si $AB = I_n$, alors on a automatiquement $BA = I_n$, si bien que A et B sont inversibles.

5.3. Inverse du produit. —

Théorème 4.55 – Inverse du produit

Soient A et B deux matrices carrées de même taille. Si A et B sont inversibles, alors AB est inversible et on a l'égalité suivante :

$$(AB)^{-1} = B^{-1}A^{-1}.$$

Démonstration. — Supposons que A et B sont inversibles et si notons $C = B^{-1}A^{-1}$. Pour vérifier que AB est inversible d'inverse C, calculons les produits $C(AB)$ et $(AB)C$:

$$C(AB) = B^{-1}A^{-1}AB = B^{-1}I_nB = B^{-1}B = I_n$$

tandis que

$$(AB)C = ABB^{-1}A^{-1} = AI_nA^{-1} = AA^{-1} = I_n.$$

On obtient bien le renseignement espéré. □

***Remarque 4.56* (Notation $\mathrm{GL}(n,\mathbb{R})$).** — On note $\mathrm{GL}(n,\mathbb{R})$ l'ensemble des matrices $n \times n$ qui sont inversibles dans $\mathcal{M}_n(\mathbb{R})$. D'après la proposition ci-dessus, si A et B sont deux matrices appartenant à $\mathrm{GL}(n,\mathbb{R})$, alors AB appartient à $\mathrm{GL}(n,\mathbb{R})$.

Exercices du chapitre 4

Dans toute la feuille, si n et p ne sont pas précisés dans l'énoncé de l'exercice, c'est qu'il s'agit d'éléments quelconques de $\mathbb{N}^\star$.

***Exercice 4.1* (Pour s'entraîner au produit...)** — ★☆☆
On donne les matrices :

$$A = \begin{pmatrix}1 & 2 & 1\end{pmatrix}, \quad B = \begin{pmatrix}1\\-1\\0\end{pmatrix}, \quad C = \begin{pmatrix}1 & -1\\0 & 1\end{pmatrix}, \quad D = \begin{pmatrix}0 & 1 & -1\\1 & 1 & 0\\0 & 1 & 0\end{pmatrix}, \quad E = \begin{pmatrix}1 & 0\\0 & 1\\-1 & -1\end{pmatrix}, \quad F = \begin{pmatrix}1 & 1 & 0\\0 & 1 & 1\end{pmatrix}.$$

1. Dans le tableau suivant, entourer les produits qui sont bien définis et rayer ceux qui n'ont pas de sens.

A^2	AB	AC	AD	AE	AF
BA	B^2	BC	BD	BE	BF
CA	CB	C^2	CD	CE	CF
DA	DB	DC	D^2	DE	DF
EA	EB	EC	ED	E^2	EF
FA	FB	FC	FD	FE	F^2

2. Calculer, dans le tableau ci-dessus, les produits qui sont bien définis.

Exercice 4.2. — ★☆☆
Soient n, p, m, q trois entiers. On considère deux matrices $A \in \mathcal{M}_{np}(\mathbb{R})$ et $B \in \mathcal{M}_{mq}(\mathbb{R})$.
À quelle condition les matrices AB et BA existent-elles toutes les deux ? À quelle condition ont-elles la même taille ?

Exercice 4.3. — ★☆☆ Soient A et B deux matrices carrées de même taille. Simplifier au maximum les produits suivants :

$$M_1 = (A+B)^2, \quad M_2 = (A-B)(A+B), \quad M_3 = (A-B)^2, \quad M_4 = (AB)^2, \quad M_5 = (I+A+\cdots+A^k)(I-A).$$

Exercice 4.4. — ★☆☆

On fixe des réels quelconques $x_1, \ldots, x_n$ et on considère les matrices $U = \begin{pmatrix}1\\\vdots\\1\end{pmatrix}$ (matrice-colonne à n coefficients) et $X = \begin{pmatrix}x_1, \ldots, x_n\end{pmatrix}$ (matrice-ligne). Calculer les produits UX et XU.

***Exercice 4.5* (Exemples de transvections et de dilatations).** — ★☆☆
Dans cet exercice, on fixe des réels a, b, …, h, i et on considère les matrices

$$A = \begin{pmatrix}a & b & c\\d & e & f\\g & h & i\end{pmatrix}, \quad T_1 = \begin{pmatrix}1 & 0 & 2\\0 & 1 & 0\\0 & 0 & 1\end{pmatrix}, \quad T_2 = \begin{pmatrix}1 & 0 & 0\\0 & 1 & 5\\0 & 0 & 1\end{pmatrix} \quad \text{et} \quad D = \begin{pmatrix}1 & 0 & 0\\0 & 4 & 0\\0 & 0 & 1\end{pmatrix}.$$

1. Calculer les produits T_1A et T_2A.
2. Calculer les produits AT_1 et AT_2.
3. Calculer les produits AD et DA.

***Exercice 4.6* (Produit de matrices triangulaires). — ★★☆**
Soient A et B deux matrices de $\mathcal{M}_n(\mathbb{R})$. On suppose que A et B sont triangulaires supérieures.
Montrer que $A+B$ et AB sont aussi triangulaires supérieures.
On écrira la démonstration formelle (avec symboles $\sum$).

***Exercice 4.7* (Matrices de rotation dans $\mathbb{R}^2$). — ★☆☆**
Dans cet exercice, on fixe un réel θ et on considère la matrice

$$A = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

Montrer que pour tout entier $n \geq 1$, on a

$$A^n = \begin{pmatrix} \cos(n\theta) & -\sin(n\theta) \\ \sin(n\theta) & \cos(n\theta) \end{pmatrix}.$$

***Exercice 4.8.* — ★☆☆**
On considère la matrice $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$.

1. Soit $J = A - I_3$. Écrire explicitement J et calculer J^n pour $n \in \mathbb{N}$.
2. En déduire la matrice A^n pour $n \in \mathbb{N}$.

Trace et transposée. —

***Exercice 4.9.* — ★☆☆**
Soit A une matrice $n \times n$ à coefficients réels.
Montrer qu'il existe un unique couple (λ, B) vérifiant les trois conditions suivantes :

- $\mathrm{Tr}(B) = 0$,
- λ est un nombre réel,
- et $A = B + \lambda I_n$.

***Exercice 4.10.* — ★★☆**
Dans cet exercice, on fixe deux matrices A et B de format $n \times n$ à coefficients réels, et on suppose que $\mathrm{Tr}(A) \neq (-1)$.
Déterminer toutes les matrices $X \in \mathcal{M}_n(\mathbb{R})$ vérifiant la condition suivante :

$$X + (\mathrm{Tr}(X))A = B.$$

***Exercice 4.11.* — ★☆☆**
Soient A et B deux matrices triangulaires inférieures de même taille. Montrer, sans calcul mais en se ramenant au résultat de l'exercice 4.6, que la matrice AB est triangulaire inférieure.

***Exercice 4.12.* — ★☆☆**
1) Soit X une matrice colonne à coefficient réels.

1. Calculer $X^T X$. Montrer que $X^T X = 0$ si et seulement si $X = 0$.
2. Soit A une matrice à coefficient réels et X une matrice colonne à coefficient réels telle que le produit AX existe. Montrer que $AX = 0$ si et seulement si $X^T A^T AX = 0$.
3. Montrer qu'ainsi énoncé ces résultats sont faux pour des matrices à coefficients complexes. Comment peut-on les généraliser au cas des matrices à coefficients complexes ?

Exercice 4.13. — ★☆☆
Soit $A \in \mathcal{M}_{n,p}(\mathbb{R})$.

1. Montrer que les produits $A(A^T)$ et $(A^T)A$ sont des matrices carrées symétriques.
2. Montrer que si l'une de ces deux matrices est nulle, alors A est nulle.

Exercice 4.14. — ★☆☆

1. Soit M une matrice $n \times n$ à coefficients réels. Montrer qu'il existe un unique couple (A, B) de matrices $n \times n$ vérifiant les deux conditions suivantes :
   - A est symétrique et B est antisymétrique,
   - $M = A + B$.
2. Déterminer A et B si $M = \begin{pmatrix} -2 & 3 & -1 \\ 5 & 4 & -1 \\ 1 & -3 & 2 \end{pmatrix}$.

Exercice 4.15. — ★★☆
Soit A une matrice de $\mathcal{M}_n(\mathbb{R})$; on suppose que A est symétrique et que A n'est pas la matrice nulle.

1. Montrer que A^2 est symétrique.
2. Exprimer les coefficients de A^2 en fonction de ceux de A.
3. Montrer que la trace de A^2 est strictement positive, puis en déduire que A^2 est non nulle.
4. La matrice A peut-elle être nilpotente ?

Matrices inversibles. —

Exercice 4.16. — ★☆☆

1. La somme de deux matrices inversibles est-elle toujours inversible ?
2. Soit A une matrice inversible. Montrer que pour tout $n \in \mathbb{N}$, la matrice A^n est inversible.

Exercice 4.17. — ★☆☆
Soit $A = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}$.

1. Calculer A^2.
2. En déduire que A n'est pas inversible.
3. Calculer A^n pour tout entier naturel n.

Exercice 4.18. — ★☆☆
Soit $B = \begin{pmatrix} 5 & -4 \\ 4 & -3 \end{pmatrix}$. On note N la matrice telle que $B = I + N$.

1. Calculer $(I - N)(I + N)$.
2. En déduire que B est inversible et calculer son inverse, puis B^{100}.

Exercice 4.19. — Soit $A = \begin{pmatrix} 2 & 0 & 3 \\ 0 & 2 & 0 \\ 0 & 3 & 2 \end{pmatrix}$.

a) Calculer $A^3 - 6A^2 + 12A$.
b) En déduire que A est inversible et calculer A^{-1}.

Exercice 4.20. — ★☆☆
On note I la matrice identité de $\mathcal{M}_n(\mathbb{R})$, et 0 la matrice nulle de $M_n(\mathbb{R})$.

Soit $A \in \mathcal{M}_n(\mathbb{R})$ vérifiant :

$$A^2 + A + I = 0.$$

1. Montrer que A est inversible et que $A^{-1} = -A - I$.
2. Montrer que $A^3 = I$.
3. Calculer, pour tout p de $\mathbb{N}^*$, A^p en fonction de A et I.

Exercice 4.21. — ★☆☆

Soit $A = \begin{pmatrix} 2 & 1 \\ 5 & -2 \end{pmatrix}$.

1. Calculer A^2. En déduire que A est inversible et déterminer son inverse.
2. Calculer A^n pour $n \in \mathbb{N}$.
3. On considère les les suites réelles $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ définies par u_0, v_0 et la relation de récurrence :

$$\begin{cases} u_{n+1} = 2u_n + v_n \\ v_{n+1} = 5u_n - 2v_n \end{cases}$$

Calculer u_n et v_n pour tout $n \in \mathbb{N}$.

Exercice 4.22. — On considère la matrice $A = \begin{pmatrix} 0 & -1 & 1 \\ 2 & 1 & 1 \\ 0 & 1 & -1 \end{pmatrix}$.

1. Calculer A^2 puis A^3.
2. La matrice A est-elle inversible ?
3. Calculer $(A + I_3)^{10}$.
4. On considère les suites les suites réelles $(u_n)_{n\in\mathbb{N}}$, $(v_n)_{n\in\mathbb{N}}$ et $(w_n)_{n\in\mathbb{N}}$ définies par u_0, v_0 et w_0 et par la relation de récurrence :

$$\begin{cases} u_{n+1} = u_n - v_n + w_n \\ v_{n+1} = 2u_n + 2v_n + w_n \\ w_{n+1} = v_n \end{cases}$$

Calculer v_{10} quand $u_0 = 1$, $v_0 = 0$ et $w_0 = -1$.

Exercices divers. —

***Exercice 4.23* (Noyau et image : exemples concrets).** — Déterminer le noyau et l'image de chacune des matrices suivantes.

$$1)\quad A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \qquad 2)\quad A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad 3)\quad A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \qquad 4)\quad A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

$$5)\quad A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \qquad 6)\quad A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 0 & 0 \end{pmatrix} \qquad 7)\quad A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$$

***Exercice 4.24* (Racines carrés de matrices).** — Soit A une matrice 2×2 à coefficients réels ou complexes. Soit M une matrice 2×2. On dit que M est une racine carrée de A dans $\mathcal{M}_2(\mathbb{C})$ si M est à coefficients complexes et si $M^2 = A$, et on dit que M est une racine carrée de A dans $M_2(\mathbb{R})$ si M est à coefficients réels et si $M^2 = A$.

1. Montrer que si $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, alors A n'admet aucune racine carrée dans $\mathcal{M}_2(\mathbb{C})$.
2. Montrer que la matrice $B = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}$ n'a aucune racine carrée dans $\mathcal{M}_2(\mathbb{R})$ mais a exactement deux racines carrées dans $\mathcal{M}_2(\mathbb{C})$.

3. Dans $\mathcal{M}_2(\mathbb{R})$, montrer que la matrice I_2 admet une infinité de racines carrées dont les coefficients diagonaux sont nuls.
4. Déterminer toutes les racines carrées de la matrice I_2 dans $\mathcal{M}_2(\mathbb{C})$. Donner un exemple de racine carrées de I_2 dans $M_2(\mathbb{C})$ qui n'appartient pas à $M_2(\mathbb{R})$.

***Exercice 4.25* (Suite des noyaux itérés). — ★★☆**

On fixe une matrice carrée $A \in \mathcal{M}_n(\mathbb{R})$.

1. Montrer que pour tout k dans $\mathbb{N}$, on a l'inclusion $\mathrm{Ker}(A^k) \subset \mathrm{Ker}(A^{k+1})$.
2. Supposons qu'il existe un entier naturel k vérifiant $\ker(A^{k_0+1}) = \ker(A^{k_0})$.
   Montrer que pour tout entier $q \geq k_0$, on a $\ker(A^q) = \ker(A^{k_0})$.
3. En déduire que si pour tout X dans $\mathcal{M}_{n,1}(\mathbb{R})$, on a l'implication $\big(A^2X = 0 \Leftrightarrow AX = 0\big)$, alors pour tout entier $k \geq 1$ et pour tout X dans $\mathcal{M}_{n,1}(\mathbb{R})$, on a l'équivalence $\big(A^kX \Leftrightarrow AX = 0\big)$.

***Exercice 4.26* (Groupe symplectique $Sp(4,\mathbb{R})$). — ★★☆ à ★★★**

Dans cet exercice, on considère la matrice

$$J = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}$$

et le sous-ensemble de $\mathcal{M}_4(\mathbb{R})$ suivant :

$$\mathcal{G} = \left\{ M \in \mathcal{M}_4(\mathbb{R}) \ / \ M^T J M = J \right\}.$$

1. Montrer que si A et B sont deux matrices appartenant à $\mathcal{G}$, alors AB appartient à G.
2. Calculer J^2 et en déduire que J est inversible.
3. Montrer que si M est une matrice appartenant à $\mathcal{G}$, alors JM est inversible.
4. En déduire que si M est une matrice appartenant à $\mathcal{G}$, alors M est inversible.
5. Si X et X' sont deux éléments de $\mathcal{M}_{4,1}(\mathbb{R})$, on définit
$$\omega(X, X') = (X^T)J(X').$$
   (a) Si $X = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}$ et $X' = \begin{pmatrix} a' \\ b' \\ c' \\ d' \end{pmatrix}$, exprimer $\omega(X, X')$ à l'aide des coefficients $a, b, \ldots, c', d'$.
   (b) Montrer que si M est une matrice appartenant à $\mathcal{G}$, alors
$$\forall X \in \mathcal{M}_{4,1}(\mathbb{R}), \forall X' \in \mathcal{M}_{4,1}(\mathbb{R}), \qquad \omega(MX, MX') = \omega(X, X'). \tag{$\star$}$$
   (c) (*Plus difficile*). Montrer réciproquement que si M est une matrice 4×4 pour laquelle la propriété $(\star)$ est vérifiée, alors on a $M \in \mathcal{G}$.

CHAPITRE 5

SYSTÈMES LINÉAIRES

1. Généralités

1.1. Système linéaire ; matrice des coefficients. —
Vous savez résoudre des systèmes de *deux équations linéaires à deux inconnues*, du type

$$\begin{cases} x + 3y = 5 \\ 3x - 5y = 8 \end{cases}$$

d'inconnue $(x, y) \in \mathbb{R}^2$. Dans ce chapitre, nous apprenons à résoudre des systèmes d'équations pouvant comporter trois équations et trois inconnues, mais ayant une forme analogue, par exemple

$$\begin{cases} 2x + y + z = 2 \\ y + 2z = -1 \\ 5z = 5 \end{cases}$$

ou plus généralement des systèmes de n équations à p inconnues ayant une forme très particulière : on fixe des nombres réels a_{11}, …, a_{1p}, …, a_{n1}, …, a_{np}, ainsi que des nombres $b_1, \dots, b_n$, et on cherche tous les p-uplets de nombres réels $x_1, \dots, x_p$ vérifiant les conditions suivantes :

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1p}x_p = b_1 \\ \vdots \qquad\qquad\qquad\qquad\quad \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{np}x_p = b_p \end{cases} \tag{S}$$

On peut reformuler le système (S) au moyen d'une égalité de matrices : il est équivalent à l'égalité de matrices

$$\begin{pmatrix} a_{11} & \dots & a_{1p} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{np} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \tag{S_{mat}}$$

Pour cette raison, on associe généralement au système (S) deux matrices :

- la *matrice des coefficients du membre de gauche*, qui est la matrice $A = \begin{pmatrix} a_{11} & \dots & a_{1p} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{np} \end{pmatrix}$,
- la *colonne des seconds membres*, qui est la matrice $\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$.

Par exemple, pour le système

$$\begin{cases} 2x + y + z = 2 \\ y + 2z = -1 \\ 5z = 5 \end{cases}$$

la matrice des coefficients du membre de gauche (ou, en abrégé, la « matrice du système ») est $A = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 5 \end{pmatrix}$, et la colonne des seconds membres est $\begin{pmatrix} 2 \\ -1 \\ 5 \end{pmatrix}$.

1.2. Cas d'un système triangulaire. — Certains systèmes sont faciles à résoudre : par exemple, pour trouver les réels x, y, z vérifiant

$$\begin{cases} 2x + y + z = 2 \\ y + 2z = -1 \\ 5z = 5 \end{cases}$$

on constate que la troisième équation impose $z = 1$; en combinant cette information avec la deuxième équation, on constate qu'elle impose $y = -3$. La première équation permet alors de trouver $x = 2$.

Plus généralement, il est facile de résoudre les systèmes du type suivant.

Définition 5.1 – Système triangulaire

Soit (S) un système de n équations linéaires à n inconnues. On dit que (S) est un système triangulaire lorsque la matrice A des coefficients du membre de gauche vérifie la propriété suivante :

- A est triangulaire supérieure,
- il n'y a aucun zéro sur la diagonale de A.

Exemple 5.2. — Le système $\begin{cases} 5x + 8z = 0 \\ 2y - 9z = -1 \\ z = 0 \end{cases}$, de matrice $\begin{pmatrix} 5 & 0 & 8 \\ 0 & 2 & -9 \\ 0 & 0 & 1 \end{pmatrix}$, est triangulaire. Par contre, le système $\begin{cases} 5x + 2y + 8z = 0 \\ -9z = -1 \\ z = 0 \end{cases}$ ne l'est pas : sa matrice $A = \begin{pmatrix} 5 & 2 & 8 \\ 0 & 0 & -9 \\ 0 & 0 & 1 \end{pmatrix}$ est triangulaire supérieure, mais il y a un zéro sur la diagonale de A.

Remarque 5.3. — On notera qu'un système ne peut être triangulaire que si sa matrice A est une matrice carrée : cette appellation est donc réservée aux systèmes comportant *autant d'équations que d'inconnues.*

Proposition 5.4 – Pour les systèmes triangulaires, il y a une unique solution

Si (S) est un système triangulaire, alors (S) admet une unique solution.

1.3. Notion de système échelonné. — Considérons le système suivant

$$\begin{cases} x + y + 2z + t = 0 \\ 2z - 4t = 0 \end{cases} \qquad (S_{\text{ech}})$$

d'inconnue $(x, y, z, t) \in \mathbb{R}^4$. Le système (S_{ech}) n'est pas triangulaire : sa matrice est

$$A = \begin{pmatrix} 1 & 1 & 2 & 1 \\ 0 & 0 & 2 & -4. \end{pmatrix}$$

On peut néanmoins le résoudre, pourvu qu'on mène la discussion avec soin.

- On remarque que la dernière équation ne permet pas de trouver la valeur de z, ni la valeur de t. Elle exprime simplement un *lien* entre z et t.
- De même, la première équation ne permet pas de trouver x, y, z ou t : elle exprime simplement un *lien* entre ces quatre inconnues. Si on utilise le renseignement fourni par la deuxième équation, on constate cependant que (S_{ech}) est équivalent à

$$\begin{cases} x &= -y - 3t \\ z &= 2t. \end{cases} \tag{S_{fin}}$$

  On remarque que *si l'on fixe des valeurs pour y et pour t*, on pourra
  - trouver, avec la deuxième équation, une valeur pour x de façon à ce que la deuxième équation soit satisfaite,
  - et trouver, avec la première équation une valeur pour x de façon à ce que la première équation soit satisfaite.

  Autrement dit, si l'on *choisit* deux réels α et β *arbitraires* et si l'on considère la colonne

$$\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} -\alpha - 3\beta \\ \alpha \\ 2\beta \\ \beta \end{pmatrix} \tag{1.1}$$

  on obtient une solution de (S_{ech}).
- Nous constatons donc que le système (S_{ech}) admet *une infinité de solutions* : $\begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$ en est une, $\begin{pmatrix} -3 \\ 0 \\ 2 \\ 2 \end{pmatrix}$ en est une autre, $\begin{pmatrix} -11 \\ 5 \\ 4 \\ 2 \end{pmatrix}$ une autre...
- De plus, *toute* solution du système (S_{ech}) est de la forme (1.1), puisque si (x, y, z, t) est un quadruplet donnant une solution du système et si l'on note α pour y et β pour t, on aura $\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} -\alpha - 3\beta \\ \alpha \\ 2\beta \\ \beta \end{pmatrix}$.
- Le point crucial de la discussion précédente est le fait que la forme du système (S_{ech}) permet de *choisir librement des valeurs pour y et t* et, ensuite, d'en *déduire les valeurs que doivent prendre x et z, en fonction du choix fait pour y et t, pour que (x, y, z, t) soit une solution du système.*

  On dit, pour résumer cela, que les variables y et t sont des *variables libres dans* (S_{ech}), tandis que x et t sont des *variables liées* dans (S_{ech}).

  Attention cependant, cette terminologie (variables « libres » et variables « liées ») n'est pas universelle.

Les systèmes ayant une forme analogue à (S_{ech}) joureront un rôle crucial dans ce chapitre. On dit qu'ils sont *échelonnés*, ou que leur matrice est *échelonnée en lignes* :

Définition 5.5 – Matrice échelonnée

Soit A une matrice $n \times p$ à coefficients réels. On dit que A est *échelonnée en lignes* (ou simplement *échelonnée*) lorsque chacune des lignes non nulles de A commence par strictement plus de zéros que la précédente.

Par exemple, les matrices $A = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 5 & 2 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ et $C = \begin{pmatrix} -7 & 2 & 0 & 4 \\ 0 & 8 & 0 & 3 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ sont échelonnées, mais les matrices $D = \begin{pmatrix} 1 & 2 & 0 & 4 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ $E = \begin{pmatrix} 0 & 2 & 0 & 4 \\ 5 & 0 & 1 & 2 \end{pmatrix}$, $F = \begin{pmatrix} 0 & 2 & 3 & 1 \\ 0 & 4 & 0 & 5 \\ 6 & 0 & 0 & 1 \end{pmatrix}$ ne sont pas échelonnées.

Définition 5.6 – Matrice échelonnée réduite

Soit A une matrice $n \times p$ à coefficients réels. On dit que A est *échelonnée réduite* lorsque A est échelonnée et lorsque sur chaque ligne non nulle, le premier coefficient non nul est égal à 1.

Par exemple, $B = \begin{pmatrix} 5 & 2 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ n'est pas échelonnée réduite, mais $B' = \begin{pmatrix} 1 & 2/5 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ l'est. De même, $C = \begin{pmatrix} -7 & 2 & 0 & 4 \\ 0 & 8 & 0 & 3 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ n'est pas échelonnée réduite, mais $C' = \begin{pmatrix} 1 & -2/7 & 0 & -4/7 \\ 0 & 1 & 0 & 3/8 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ est échelonnée réduite.

On dira qu'un système linéaire (S) est *échelonné réduit* lorsque sa matrice est échelonnée réduite : par exemple, le système

$$\begin{cases} x + y + 2z + t = 0 \\ \qquad\quad 2z - 4t = 0 \end{cases} \tag{1.2}$$

est échelonné mais n'est pas réduit ; en revanche, le système

$$\begin{cases} x + y + 2z + t = 0 \\ \qquad\quad z - 2t = 0 \end{cases} \tag{$S_{\text{ech}-\text{r}}$}$$

est échelonné réduit.

Nous verrons au §3.2 et au §3.3 qu'on peut résoudre tous les systèmes échelonnés réduits en imitant la discussion menée pour S_{ech}.

2. Opérations élémentaires et méthode du pivot

2.1. Exemples et principe général. —

***Exemple 5.7* (Un système qui se ramène à un système triangulaire)**

Nous commençons par considérer le système linéaire

$$\begin{cases} x + 3y + 5z = 1 \\ 3x + 6y + 9z = -2 \\ -2x - 3y - 7z = 2 \end{cases} \tag{S_1}$$

d'inconnue $(x, y, z) \in \mathbb{R}^3$.

Nous saurions le résoudre s'il était triangulaire, c'est-à-dire si x n'apparaissait pas dans la deuxième équation et x, y n'apparaissaient pas dans la troisième. Mais (S_1) n'est pas triangulaire.

Cependant, on peut faire la remarque suivante. Notons L_1 la première ligne du système, L_2 la deuxième et L_3 la troisième. Si les équations L_1 et L_2 sont vérifiées, alors assurément l'équation $-3y-6z=-5$ est vérifiée, puisqu'elle est obtenue comme la différence L_2-3L_1.

Si (x,y,z) est une solution de (S_1), on constate donc que c'est aussi une solution de

$$\left\{\begin{array}{rcrcrcr} x & + & 3y & + & 5z & = & 1 \\ & - & 3y & - & 6z & = & -5 \\ -2x & - & 3y & - & 7z & = & 9 \end{array}\right. \qquad (S_1')$$

En fait, et ceci est crucial, (S_1) et (S_1') sont *équivalents*. En effet, si (x,y,z) vérifie les équations de (S_1'), en ajoutant trois fois la première ligne de (S_1') à la deuxième, on retrouve la deuxième ligne de (S_1) qui n'apparaissait plus dans (S_1').

Nous avons transformé (S_1) en un système (S_1') qui est *équivalent*, mais dans lequel y n'apparaît plus dans la deuxième équation.

De même, on peut faire disparaître x de la troisième équation en remplaçant la ligne L_3 de (S_1) par L_3+2L_1, et on obtient un système *équivalent à* (S_1) :

$$(S_1) \iff \left\{\begin{array}{rcrcrcrl} x & + & 3y & + & 5z & = & 1 & \\ & - & 3y & - & 6z & = & -5 & \\ & & 3y & + & 3z & = & 11 & L_3 \leftarrow L_3+2L_1 \end{array}\right.$$

Nous pouvons alors, en ajoutant la deuxième ligne à la troisième, obtenir un système équivalent à (S_1), mais triangulaire :

$$(S_1) \iff \left\{\begin{array}{rcrcrcrl} x & + & 3y & + & 5z & = & 1 & \\ & & -3y & - & 6z & = & -5 & \\ & & & - & 3z & = & 6 & L_3 \leftarrow L_3+L_2 \end{array}\right.$$

Nous voilà en mesure de résoudre (S_1) : pour que le dernier système soit vérifié, il est nécessaire qu'on ait $z=2$, $y=\frac{-3}{7}$ et $x=-\frac{12}{7}$. C'est également suffisant, donc (S_1) admet une unique solution, donnée par $(x,y,z)=(-\frac{12}{7},\frac{-3}{7},-2)$.

***Exemple 5.8* (Un système qui se ramène à un système échelonné)**

Considérons à présent le système suivant :

$$\left\{\begin{array}{rcrcrcrcr} x & + & y & + & 2z & + & t & = & 0 \\ 3x & + & 3y & + & 6z & + & 3t & = & 0 \\ -2x & - & 2y & - & 7z & + & 4t & = & 0 \\ x & + & y & + & 4z & - & 3t & = & 0 \end{array}\right. \qquad (S_2)$$

d'inconnue $(x,y,z,t)\in\mathbb{R}^4$.

En imitant les manipulations de l'exemple ci-dessus, on peut s'appuyer sur la première ligne pour "chasser x de toutes les autres équations", et on obtient un système équivalent à (S_2) :

$$(S_2) \iff \left\{\begin{array}{rcrcrcrcrl} x & + & y & + & 2z & + & t & = & 0 & \\ & & & & & & 0 & = & 0 & L_2 \leftarrow L_2-3L_1 \\ & & & - & 3z & + & 6t & = & 0 & L_3 \leftarrow L_3+2L_1 \\ & & & & 2z & - & 4t & = & 0 & L_4 \leftarrow L_4-L_1 \end{array}\right. \qquad (2.1)$$

Nous constatons cependant que y a disparu de toutes les équations, sauf la première ! Il est donc impossible d'obtenir un système triangulaire équivalent à (S_2).

On peut cependant remarquer que si l'on échange la deuxième et la troisième ligne,

$$(S_2) \iff \left\{ \begin{array}{rcrcrcrcll} x & + & y & + & 2z & + & t & = & 0 & \\ & & & - & 3z & + & 6t & = 0 & & L_2 \leftrightarrow L_3 \\ & & & & & & 0 & = & 0 & \\ & & & & 2z & - & 4t & = 0 & & \end{array} \right. \tag{2.2}$$

puis qu'on utilise la deuxième équation pour "chasser z de la dernière équation", on obtient

$$(S_2) \iff \left\{ \begin{array}{rcrcrcrcll} x & + & y & + & 2z & + & t & = & 0 & \\ & & & - & 3z & + & 6t & = 0 & & \\ & & & & & & 0 & = & 0 & \\ & & & & & & 0 & = 0 & & L_4 \leftarrow L_4 + \frac{2}{3}L_2 \end{array} \right. \tag{2.3}$$

On constate que la variable t, à son tour, a disparu... Cependant, en divisant de part et d'autre de la deuxième équation par -3, on constate l'équivalence

$$(S_2) \iff \left\{ \begin{array}{rcrcrcrcll} x & + & y & + & 2z & + & t & = & 0 & \\ & & & & z & & -2t & = 0 & & \\ & & & & & & 0 & = & 0 & \\ & & & & & & 0 & = 0 & & L_4 \leftarrow L_4 + L_2 \end{array} \right. \tag{2.4}$$

entre (S_2) et un système qui, à défaut d'être triangulaire, est *échelonné réduit* : c'est le même que le système échelonné réduit ($S_{\text{ech}-\text{r}}$) que nous avons résolu dans le paragraphe précédent. Nous savons donc résoudre (S_2) !

Les deux exemples précédents indiquent une stratégie possible pour résoudre tous les systèmes linéaires : à l'aide des opérations que nous avons effectuées (ajouter une ligne aux autres, échanger des lignes, multiplier de part et d'autre d'une équation), on peut espérer ramener tout système à un système "plus simple" et résoudre ce système "plus simple".

Nous verrons dans la suite qu'on peut en fait *ramener tout système linéaire à un système échelonné réduit*, et qu'il est possible en imitant les remarques du §1.3 de *résoudre tous les systèmes échelonnés réduits*.

2.2. Transvection, dilatation, permutation. —

Nous commençons par donner un nom aux opérations sur les systèmes linéaires effectuées dans les exemples ci-dessus.

Définition 5.9 – Opérations élémentaires sur les lignes

Fixons $n \in \mathbb{N}^\star$ et $p \in \mathbb{N}^\star$. Soit (S) un système linéaire de n équations à p inconnues.

- On dit qu'on effectue sur (S) une *transvection de lignes* lorsqu'on choisit i et j dans $\{1, \dots, n\}$ avec $i \neq j$, ainsi qu'un nombre $\alpha \in \mathbb{R}$, et lorsqu'on remplace la ligne L_i de (S) par $L_i + \alpha L_j$.

  ⇝ L'écriture « $L_i \leftarrow L_i + \alpha L_j$ » résume cette opération.
- On dit qu'on effectue sur (S) une *dilatation de ligne* lorsqu'on choisit $i \in \{1, \dots, n\}$ ainsi qu'un nombre $\beta \in \mathbb{R}$ avec $\beta \neq 0$, et lorsqu'on remplace la ligne L_i de (S) par βL_i.

  ⇝ L'écriture « $L_i \leftarrow \beta L_i$ » résume cette opération.
- On dit qu'on effectue sur (S) une *permutation de lignes* lorsqu'on choisit $i \in \{1, \dots, n\}$ et $j \in \{1, \dots, n\}$ avec $i \neq j$, et qu'on échange les lignes L_i et L_j de (S)

  ⇝ L'écriture « $L_i \leftrightarrow L_j$ » résume cette opération.

Nous avons déjà remarqué le fait suivant :

Proposition 5.10 – Opérations élémentaires sur les lignes → système équivalent

Si (S) est un système linéaire, et si (S') est le système obtenu à partir de (S') en effectuant sur les lignes de (S) l'une des opérations élémentaires décrites ci-dessus, alors (S) et (S') sont équivalents.

Dans la suite de ce paragraphe, nous montrons que partant d'un système linéaire quelconque, il est possible d'obtenir, par une succession d'opérations élémentaires, un système échelonné réduit : d'après la proposition ci-dessus, le système initial sera donc équivalent à un système échelonné réduit.

2.3. Interprétation matricielle des opérations élémentaires. —

Plutôt que d'écrire les opérations élémentaires sur les systèmes linéaires, il sera très avantageux de travailler sur la matrice des coefficients du membre de gauche. Dans ce but, nous introduisons les matrices suivantes.

Définition 5.11 – Matrice de transvection, de dilatation, de permutation

Fixons $n \in \mathbb{N}^\star$. Si i et j sont deux éléments distincts de $\{1, \dots, n\}$, si α est un nombre réel et si β est un nombre réel vérifiant $\beta \neq 0$,

- On note $T_{i,j,\alpha}$ la matrice $n \times n$ obtenue à partir de l'identité en effectuant l'opération élémentaire $L_i \leftarrow L_i + \alpha L_j$
- On note $D_{i,\beta}$ la matrice $n \times n$ obtenue à partir de l'identité en effectuant l'opération élémentaire $L_i \leftarrow \beta L_i$
- On note $P_{i,j}$ la matrice $n \times n$ obtenue à partir de l'identité en effectuant l'opération élémentaire $L_i \leftrightarrow L_j$

Exemple 5.12. — Si l'on fixe $n = 3$, alors la matrice $T_{2,1,-4}$ est égale à $\begin{pmatrix} 1 & 0 & 0 \\ -4 & 0 & 0 \end{pmatrix}$, tandis que la matrice $T_{3,2,5}$ est égale à $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 5 & 1 \end{pmatrix}$.

L'importance de ces matrices est due à la proposition suivante :

Proposition 5.13 – Multiplication à gauche par une matrice d'opération élémentaire

Soit A une matrice $n \times p$. Fixons i et j dans $\{1, \dots, n\}$, ainsi que $a \in \mathbb{R}$ et $\beta \in \mathbb{R}^\star$.

- *La matrice $T_{i,j,\alpha}A$ est la matrice obtenue à partir de A en effectuant l'opération élémentaire $L_i \leftarrow L_i + \alpha L_j$*
- *La matrice $D_{i,\beta}A$ est la matrice obtenue à partir de A en effectuant l'opération élémentaire $L_i \leftarrow \beta L_i$*
- *La matrice $P_{i,j}A$ est la matrice obtenue à partir de A en effectuant l'opération élémentaire $L_i \leftrightarrow L_j$*

Exemple 5.14. — Si $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$, alors $T_{2,1,-4}A$ est la matrice obtenue en effectuant sur A l'opération $L_2 \leftarrow L_2 - 4L_1$: elle est égale à $\begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 7 & 8 & 9 \end{pmatrix}$.

Grâce à la proposition 5.13, on peut utiliser la multiplication à gauche par des matrices élémentaires pour "encoder" les opérations sur un systèmes linéaires que nous rencontrions dans les exemples du §**??**.

Nous avons vu au paragraphe précédent, dans la discussion de l'exemple 5.7, que

- l'opération élémentaire $L_i \leftarrow L_i + \alpha L_j$ peut être "annulée" en effectuant ensuite l'opération $L_i \leftarrow L_i - \alpha L_j$

Les deux autres opérations élémentaires sont elles aussi réversibles :

- l'opération élémentaire $L_i \leftarrow \beta L_i$, $\beta \neq 0$, peut être "annulée" en effectuant ensuite l'opération $L_i \leftarrow \frac{1}{\beta} L_i$
- l'opération élémentaire de permutation $L_i \leftrightarrow L_j$ peut être annulée en effectuant une deuxième fois la permutation $L_i \leftrightarrow L_j$.

Sur les matrices correspondantes, cela se traduit par le résultat suivant :

Proposition 5.15 – Multiplication à gauche par une matrice d'opération élémentaire

Soit A une matrice $n \times p$. Fixons i et j dans $\{1, \dots, n\}$, ainsi que $a \in \mathbb{R}$ et $\beta \in \mathbb{R}^\star$.

- *La matrice $T_{i,j,\alpha}A$ est inversible, d'inverse $T_{i,j,-\alpha}$*
- *La matrice $D_{i,\beta}$ est inversible, d'inverse $D_{i,\frac{1}{\beta}}$.*
- *La matrice $P_{i,j}$ est inversible, et son inverse est $P_{i,j}$ elle-même.*

Nous remarquons de plus que parmi les transvections, nous n'avons eu besoin ci-dessus que de transvections du type suivant.

Définition 5.16 – Transvection descendante

On dit qu'une transvection $L_i \leftarrow L_i + \alpha L_j$ est *descendante* lorsque $i > j$ (c'est-à-dire lorsqu'elle consiste à ajouter à une ligne L_i une ligne située *au-dessus* de L_i.

Si T est une matrice de transvection $T = T_{i,j,\alpha}$, nous avons vu que T est inversible et que son inverse est $T_{i,j,-\alpha}$. On peut donc remarquer que si T est une matrice de transvection descendante, alors T^{-1} est aussi une matrice de transvection descendante.

2.4. Notion de réduite échelonnée et théorème principal. — Nous pouvons maintenant annoncer que tout système peut être « rendu échelonné réduit » par une succession d'opérations élémentaires. Le plus simple, pour les besoins de ce cours, est probablement d'utiliser les matrices et de montrer que que toute matrice peut être « rendue échelonnée réduite » à l'aide d'une succession d'opérations élémentaires. C'est le résultat principal de ce chapitre.

Théorème 5.17 – Le théorème de réduction

Soient n et p deux éléments de $\mathbb{N}^\star$. Soit A une matrice de $\mathcal{M}_{n,p}(\mathbb{R})$.
Il existe une matrice inversible $U \in \mathcal{M}_{n,n}(\mathbb{R})$ vérifiant les deux propriétés suivantes :
(a) la matrice $S = UA$ est échelonnée réduite,
(b) la matrice U peut s'obtenir comme produit d'un nombre fini de matrices de transvection descendante, de dilatation et de permutation.

Corollaire 5.18 – Conséquence pour les systèmes linéaires

Tout système linéaire est équivalent à un système échelonné réduit. De plus, partant d'un système linéaire (S) quelconque, on peut obtenir un système échelonné réduit équivalent à (S) par une succession de transvections, de dilatations et de permutations de lignes.

Proposition 5.19 – Réduite échelonnée d'une matrice

Si A est une matrice de $\mathcal{M}_{n,p}(\mathbb{R})$, alors dans l'égalité $S = UA$ (avec U inversible) du théorème ci-dessus, la matrice S est uniquement déterminée par A (mais la matrice U, en général, ne l'est pas).
On dit que S est la réduite échelonnée de A.

La vérification du théorème de réduction n'est pas très agréable à écrire. Nous incluons cependant, pour que ce cours soit complet, des démonstrations des deux résultats ci-dessus (il n'est pas nécessaire de retenir ces preuves).

Démonstration de l'unicité de la réduite échelonnée. — S'il existe des matrices inversibles U, U', S et S' telles que $S = UA$ et $S' = U'A$ soient échelonnées réduites, alors on peut écrire $S' = UA = U(U^{-1}S) = (UU^{-1})S$. Or, la matrice U^{-1} est l'inverse d'un produit de matrices . De plus, si T est une matrice de transvection descendante, alors T^{-1} est aussi une matrice de transvection descendante ; de même, si D est une matrice de dilatation, D^{-1} est une matrice de dilatation, et si P est une matrice de permutation, P^{-1} en est une aussi.

D'après la formule pour l'inverse d'un produit de matrices, nous constatons que U^{-1} est produit d'un nombre fini de matrices de transvection descendante, de dilatation et de permutation. C'est donc aussi le cas de $U'U^{-1}$.

Ainsi, on peut obtenir S' à partir de S par une succession de transvections descendantes, de dilatations et de permutations. Mais il est clair que chacune de ces opérations « brise » le caractère échelonné réduit. Il est donc nécessaire qu'on ait $U'U^{-1} = I_n$, d'où $S = S'$, comme annoncé. □

Démonstration du théorème de réduction. — Nous allons raisonner par récurrence sur le nombre p de colonnes de la matrice A. Dans tout le raisonnement, on fixe le nombre de lignes $n \in \mathbb{N}^\star$.

- Si $A = (a_{i1})_{1\leq i\leq n}$ est de taille $n \times 1$ et si A n'est pas la colonne nulle, notons i_0 le plus petit $i \in \{1, \dots, n\}$ tel que a_{i_0} soit non nul. En effectuant la dilatation $L_{i_0} \leftarrow \frac{1}{a_{i_0}} L_{i_0}$, on obtient une matrice-colonne avec un 1 en position j_0 : ainsi, la matrice $D_{1,\frac{1}{a_{i_0}}} A$ commence par $(i_0 - 1)$ zéros, suivis d'un 1, suivis de coefficients arbitraires. Pour $i > i_0$, effectuons alors l'opération élémentaire $L_i \leftarrow L_i - \frac{a_i}{L}{}_{i_0}$: on obtient une matrice ayant des zéros en position i pour tout $i > i_0$. Cette matrice

est échelonnée réduite ! Ainsi, si on note

$$U = T_{i_0+1,i_0,-a_{i_0+1}} \dots T_{n,i_0,-a_n} D_{i_0,\frac{1}{a_{i_0}}},$$

on constate que UA est échelonnée réduite et que U est du type indiqué dans (b) du théorème.

- Soit $n \in \mathbb{N}^\star$. Supposons que pour toute matrice $B \in \mathcal{M}_{n-1,p}$, il existe une matrice inversible $V \in \mathcal{M}_{n-1,n-1}(\mathbb{R})$ telle que VB soit échelonnée réduite et V soit un produit de matrices de transvection descendante, de dilatation et de permutation.

  Considérons à présent une matrice $A \in \mathcal{M}_{n,p}(\mathbb{R})$, et notons $A = (a_{ij})_{\substack{1\leq i\leq n\\ 1\leq j\leq p}}$.

  Distinguons deux cas :

  - Si la première colonne de A est formée de zéros, appliquons l'hypothèse de récurrence à $B = (a_{ij})_{\substack{2\leq i\leq n\\ 1\leq j\leq p}}$, qui est une matrice à $(p-1)$ colonnes obtenue en "retirant" la première colonne de A. On constate qu'il existe une matrice $U \in \mathcal{M}_{n,n}(\mathbb{R})$, du type indiqué dans (b) du théorème, telle que UB soit échelonnée réduite. Mais UA est la matrice dont la première colonne est nulle et dont les suivantes sont celles de UB. On constate qu'elle est échelonnée réduite.
  - Si la première colonne de A n'est pas formée de zéros, répétons les étapes effectuées dans le cas d'une matrice à une seule colonne : en notant i_0 le plus petit $j \in \{1, \dots, p\}$ tel que $a_{i_0 1}$ soit non nul, on constate que la matrice $A' = T_{i_0+1,i_0,-a_{i_0+1}} \dots T_{n,i_0,-a_n} D_{i_0,\frac{1}{a_{i_0}}} A$ a, sur sa première colonne, un coefficient 1 en position i_0 et des zéros partout ailleurs.

    En permutant la première ligne et la i_0-ème, on constate que

    $$A'' = P_{i_0 1} T_{i_0+1,i_0,-a_{i_0+1}} \dots T_{n,i_0,-a_n} D_{i_0,\frac{1}{a_{i_0}}} A$$

    est une matrice dont la première colonne commence par un 1, suivi uniquement de zéros. Appliquons alors l'hypothèse de récurrence à la matrice B de taille $n \times (p-1)$ obtenue en retirant la première colonne de A''.

    On obtient l'existence d'une matrice $U \in \mathcal{M}_{n,n}(\mathbb{R})$, du type indiqué dans (b) du théorème, telle que UB soit échelonnée réduite. On observe alors que

    - $VA'' = VP_{i_0 1} T_{i_0+1,i_0,-a_{i_0+1}} \dots T_{n,i_0,-a_n} D_{i_0,\frac{1}{a_{i_0}}} A$ est échelonnée réduite, puisque sa première colonne est $\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ et les suivantes sont celles de VB
    - et $U = VP_{i_0 1} T_{i_0+1,i_0,-a_{i_0+1}} \dots T_{n,i_0,-a_n} D_{i_0,\frac{1}{a_{i_0}}}$ est du type indiqué dans (b) du théorème, puisque V l'est.

□

2.5. Réduction : en pratique. — Si A est une matrice $n \times p$, comment trouver la réduite échelonnée de A ?

Nous mènerons la discussion à partir de l'exemple suivant :

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

Notre but est de trouver la réduite échelonnée S de A, ainsi qu'une matrice inversible U, produit de transvections descendantes, de dilatations et de permutations, telle que $S = UA$.

Nous construirons la matrice U, et trouverons la matrice S, par plusieurs étapes successives inspirées de l'exemple §5.7. Nous commençons avec la remarque suivante :

$$\text{si } U_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{alors} \quad U_0 A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

Pour pouvoir obtenir une matrice comportant un zéro au début de la deuxième ligne, on peut effectuer la transvection $L_2 \leftarrow L_2 - 4L_1$ sur les lignes de la matrice A. Les résultats du paragraphe précédent indiquent qu'il faut, pour cela, multiplier la matrice A sur la gauche par la matrice $T_{2,1,-4}$. Nous constatons donc que

$$\text{si } U_1 = \begin{pmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{alors} \quad U_1 A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 7 & 8 & 9 \end{pmatrix} \qquad L_2 \leftarrow L_2 - 4L_1$$

Pour poursuivre, on peut effectuer la transvection $L_3 \leftarrow L_3 - 7L_1$ sur la matrice $U_1 A$ que nous venons d'obtenir. La matrice obtenue sera égale à $T_{3,1,-7} U_0 A$: on peut la réécrire comme $U_2 A$ où $U_2 = T_{3,1,-7} U_0$. Cette dernière matrice est celle qu'on obtient à partir de U_1 en effectuant la transvection $L_3 \leftarrow L_3 - 7L_1$, c'est-à-dire en effectuant sur U_1 l'opération que nous souhaitons faire ! Nous obtenons la conclusion suivante :

$$\text{si } U_2 = \begin{pmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ -7 & 0 & 1 \end{pmatrix}, \quad \text{alors} \quad U_3 A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 0 \end{pmatrix} \qquad L_3 \leftarrow L_3 - 7L_1$$

On peut ensuite poursuivre les opérations, en gardant trace des modifications sur la matrice par laquelle on multiplie A à chaque étape :

$$\text{si } U_3 = \begin{pmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 1 & -2 & 1 \end{pmatrix}, \quad \text{alors} \quad U_3 A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 0 \end{pmatrix} \qquad L_3 \leftarrow L_3 - 2L_2$$

La dernière matrice est échelonnée. Pour obtenir la réduite échelonnée de A, il suffit alors d'effectuer une dilatation :

$$\text{si } U_4 = \begin{pmatrix} 1 & 0 & 0 \\ 4/3 & -1/3 & 0 \\ 1 & -2 & 1 \end{pmatrix}, \quad \text{alors} \quad U_4 A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 0 \end{pmatrix} \qquad L_2 \leftarrow \tfrac{1}{-3} L_2$$

Les remarques ci-dessus permettent de ramener de façon relativement efficace tout système linéaire à un système échelonné. Considérons par exemple le système suivant :

$$\left\{ \begin{array}{rcrcrcrcrcl} x_1 & + & 2x_2 & + & 3x_3 & + & 4x_4 & + & 5x_5 & = & 1 \\ 2x_1 & + & 4x_2 & + & 6x_3 & + & 8x_4 & + & 10x_5 & = & 1 \\ 4x_1 & + & 12x_2 & + & 12x_3 & + & 15x_4 & + & 16x_5 & = & 1 \\ -3x_1 & - & 2x_2 & - & 9x_3 & - & 14x_4 & - & 23x_5 & = & 1 \end{array} \right. \tag{2.5}$$

d'inconnue $(x_1, x_2, x_3, x_4, x_5) \in \mathbb{R}^5$.

La matrice des coefficients du membre de gauche est

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 6 & 8 & 10 \\ 4 & 12 & 12 & 15 & 16 \\ -3 & -2 & -9 & -14 & -23 \end{pmatrix}$$

Suivons la stratégie ci-dessus pour trouver la réduite échelonnée de A ainsi qu'une matrice inversible U telle que $S = UA$:

$$\text{si } U_0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \text{alors} \quad U_0 A = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 6 & 6 & 7 & 6 \\ 4 & 12 & 12 & 15 & 16 \\ -3 & -2 & -9 & -14 & -23 \end{pmatrix}$$

$$\text{si } U_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ -4 & 0 & 1 & 0 \\ 3 & 0 & 0 & 1 \end{pmatrix}, \quad \text{alors} \quad U_1 A = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & -1 & -4 \\ 0 & 4 & 0 & -2 & -8 \end{pmatrix} \qquad \begin{array}{l} L_2 \leftarrow L_2 - 2L_1 \\ L_3 \leftarrow L_3 - 4L_1 \\ L_4 \leftarrow L_4 + 3L_1 \end{array}$$

$$\text{si } U_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 3 & 0 & 0 & 1 \\ -4 & 0 & 1 & 0 \\ -2 & 1 & 0 & 0 \end{pmatrix}, \quad \text{alors} \quad U_2 A = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 4 & 0 & -2 & -8 \\ 0 & 4 & 0 & -1 & -4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \qquad L_2 \leftrightarrow L_4$$

$$\text{si } U_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 3 & 0 & 0 & 1 \\ -1 & 0 & 1 & -1 \\ -2 & 1 & 0 & 0 \end{pmatrix}, \quad \text{alors} \quad U_3 A = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 4 & 0 & -2 & -8 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \qquad L_3 \leftarrow L_3 - L_2$$

$$\text{si } U_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 3/4 & 0 & 0 & 1/4 \\ -1 & 0 & 1 & -1 \\ -2 & 1 & 0 & 0 \end{pmatrix}, \quad \text{alors} \quad U_3 A = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 0 & -1/2 & -2 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \qquad L_2 \leftarrow \tfrac{1}{4} L_2$$

On constate ainsi que (2.5) est équivalent à un système échelonné réduit de la forme

$$\begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 + 5x_5 = ? \\ \qquad x_2 \qquad - \tfrac{1}{2}x_4 - 2x_5 = ? \\ \qquad\qquad\quad x_4 + 4x_5 = ? \\ \qquad\qquad\qquad\quad 0 = ? \end{cases}$$

Pour résoudre ce dernier système, il nous manque l'information sur les seconds membres.

Comment trouver les seconds membres ? Deux possibilités :

- *Première stratégie : utilisation de la matrice U.*

  On peut remarquer que le système (2.5) équivaut à l'égalité de matrices

$$A \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \tag{2.6}$$

et que la réduite échelonnée S est reliée à A par $S = UA$, avec $U = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 3/4 & 0 & 0 & 1/4 \\ -1 & 0 & 1 & -1 \\ -2 & 1 & 0 & 0 \end{pmatrix}$. Si nous multiplions sur la gauche par U dans (2.6), on obtient l'égalité

$$S \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = U \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

En calculant $U \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$, on trouve $\begin{pmatrix} 1 \\ 1 \\ 1 \\ -1 \end{pmatrix}$. Nous constatons donc que (2.5) est équivalent à

$$\left\{ \begin{array}{rcrcrcrcrcl} x_1 & + & 2x_2 & + & 3x_3 & + & 4x_4 & + & 5x_5 & = & 1 \\ & & x_2 & & & - & \tfrac{1}{2}x_4 & - & 2x_5 & = & 1 \\ & & & & & & x_4 & + & 4x_5 & = & 1 \\ & & & & & & & & 0 & = & -1 \end{array} \right. \tag{2.7}$$

ce qui nous permet de conclure qu'il n'a pas de solution !

- *Deuxième stratégie : utilisation d'une matrice "augmentée".*

On peut aussi remarquer que la colonne $U \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ est celle qu'on obtient en appliquant à la colonne $\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ les mêmes opérations élémentaires que celles effectuées sur la matrice A. D'où l'idée d'effectuer les opérations "au fur et à mesure", en ajoutant sur la droite de A une colonne gardant trace de ces opérations. Au lieu d'effectuer les opérations élémentaires sur $A = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 6 & 8 & 10 \\ 4 & 12 & 12 & 15 & 16 \\ -3 & -2 & -9 & -14 & -23 \end{pmatrix}$, on peut les effectuer sur la "matrice augmentée"

$$\tilde{A} = \left(\begin{array}{ccccc|c} 1 & 2 & 3 & 4 & 5 & 1 \\ 2 & 4 & 6 & 8 & 10 & 1 \\ 4 & 12 & 12 & 15 & 16 & 1 \\ -3 & -2 & -9 & -14 & -23 & 1 \end{array} \right).$$

Si nous avions effectué sur $\tilde{A}$ les mêmes opérations que celles que nous avons faites pour mettre A sous forme échelonnée réduite, on aurait obtenu successivement les matrices "augmentées"

$$U_1\tilde{A} = \left(\begin{array}{ccccc|c} 1 & 2 & 3 & 4 & 5 & 1 \\ 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 4 & 0 & -1 & -4 & -3 \\ 0 & 4 & 0 & -2 & -8 & 4 \end{array} \right),$$

$$U_2\tilde{A} = \left(\begin{array}{ccccc|c} 1 & 2 & 3 & 4 & 5 & 1 \\ 0 & 4 & 0 & -2 & -8 & 4 \\ 0 & 4 & 0 & -1 & -4 & -3 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{array}\right),$$

$$U_3\tilde{A} = \left(\begin{array}{ccccc|c} 1 & 2 & 3 & 4 & 5 & 1 \\ 0 & 4 & 0 & -2 & -8 & 4 \\ 0 & 0 & 0 & 1 & 4 & 1 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{array}\right),$$

$$U_4\tilde{A} = \left(\begin{array}{ccccc|c} 1 & 2 & 3 & 4 & 5 & 1 \\ 0 & 1 & 0 & -1/2 & -4 & 1 \\ 0 & 0 & 0 & 1 & 4 & 1 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{array}\right).$$

Cette dernière matrice augmentée, de partie gauche échelonnée réduite, contient toute l'information nécessaire pour retrouver le système final (2.7).

3. Résolution des systèmes linéaires

3.1. Rang d'un système linéaire ou d'une matrice. —

Définition 5.20 – Rang d'une matrice

Soient n et p deux éléments de $\mathbb{N}^\star$. Soit A une matrice de $\mathcal{M}_{n,p}(\mathbb{R})$.
On appelle *rang* de A, et on note $\mathrm{rg}(A)$, le nombre de lignes non nulles de la réduite échelonnée de A.

Si (S) est un système linéaire de n équations à p inconnues, on dit que (S) est de rang r lorsque la matrice donnant les coefficients des membres de gauche est de rang r.

Exemple 5.21. — La matrice $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 3 & 4 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ est de rang 3, la matrice $\begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ est de rang 2. Pour ces deux matrices, qui sont échelonnées, le rang est particulièrement facile à lire.

Pour une matrice qui n'est pas échelonnée, il faut passer par le théorème de réduction du paragraphe précédent pour trouver le rang. Par exemple, la matrice $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ a pour réduite échelonnée la matrice $\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$, qui est de rang 2 : c'est donc que A est de rang 2.

Proposition 5.22 – Bornes sur le rang

Si A est une matrice $n \times p$, alors on a $\mathrm{rg}(A) \leq n$ et $\mathrm{rg}(A) \leq p$.

Démonstration. — Soit A une matrice $n \times p$ et S la réduite échelonnée de A. Le nombre de lignes de S est n et le nombre de lignes non nulles de S est $\mathrm{rg}(A)$: on a donc bien sûr $\mathrm{rg}(A) \leq n$.

De plus, pour chaque $k \in \{1, \dots, r\}$, on sait qu'il y a un coefficient non nul sur la k-ème ligne de S ; si note j_k la position du premier coefficient non nul de la k-ème ligne de S, alors on a $1 \leq j_1 < j_2 < \cdots < j_r$ par définition de la notion de matrice échelonnée. Comme j_r est un indice de colonne, on a $j_r \leq p$. Or, comme les j_k sont des entiers, l'inégalité $1 \leq j_1 < j_2 < \dots j_r$ implique qu'on a $j_r \geq r$. C'est donc que $r \leq p$, comme annoncé. □

3.2. Résolution des systèmes linéaires homogènes. — Nous avons vu comment ramener tout système linéaire à un système échelonné. Il nous reste à apprendre à résoudre les systèmes linéaires échelonnés. Pour cela, il faut procéder en deux étapes :

1. Apprendre à résoudre les systèmes linéaires échelonnés qui sont *homogènes*, c'est-à-dire dont les seconds membres ne sont que des zéros,
2. Apprendre à résoudre les systèmes *non homogènes*, en s'appuyant sur un lien avec les systèmes homogènes.

Nous abordons à présent le cas des systèmes échelonnés et homogènes. Nous partons donc d'un système linéaire homogène (S), de n équations à p inconnues $x_1, \dots, x_p$, et nous supposons que (S) est échelonné : il est ainsi de la forme

$$\left\{\begin{array}{rcccccccccc} x_{j_1} & + & a_{1(j_1+1)}x_{j_1+1} & + & \dots & + & \dots & + & a_{1p}x_p & = & 0 \\ & & x_{j_2} & + & a_{1(j_2+1)}x_{j_2+1} & + & \dots & + & a_{2p}x_p & = & 0 \\ & & \ddots & & & & & & \vdots & = & 0 \\ & & & & x_{j_r} & + & \dots & + & a_{rp}x_p & = & 0 \\ & & & & & & & & 0 & = & 0 \\ & & & & & & & & \vdots & = & 0 \\ & & & & & & & & 0 & = & 0 \end{array}\right. \tag{3.1}$$

où r est un entier inférieur ou égal à n et à p donnant le rang de A, où $j_1, \dots, j_k$ sont des entiers vérifiant $1 \leq j_1 < j_2 < \cdots < j_r \leq p$. On note A la matrice de ce système ; la matrice A est bien sûr échelonnée réduite.

Tout au long de la discussion qui va suivre, et qui est assez technique, nous garderons en tête l'exemple du système

$$\left\{\begin{array}{rcccccccc} x_1 & + & 2x_2 & + & 3x_3 & + & 4x_4 & + & 5x_5 & = & 0 \\ & & & & & & x_4 & + & 3x_5 & = & 0 \\ & & & & & & & & 0 & = & 0 \\ & & & & & & & & 0 & = & 0 \end{array}\right. \tag{S_{ex}}$$

d'inconnues réelles x_1, x_2, x_3, x_4 et x_5, et de matrice $A_{\text{ex}} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$.

Pour chaque $k \in \{1, \dots, r\}$, nous noterons j_k la position du premier coefficient non nul sur la ligne k de A.

Par exemple, pour A_{ex}, on a $r = 2$, $j_1 = 1$ et $j_2 = 4$.

Nous notons alors $\mathcal{L}$ l'ensemble des indices $j \in \{1, \dots, p\}$ qui ne sont égaux à aucun des j_k. Dans l'exemple précédent, on a $\mathcal{L} = \{2, 3\}$.

Nous adopterons le vocabulaire suivant (ceci est une terminologie qui n'est pas universelle) : on dira que

- Les variables $x_{j_1}, \dots, x_{j_r}$ sont des « variables liées » dans (S),
- Les variables x_j, $j \in \mathcal{L}$, sont des « variables libres » dans (S).

Dans notre exemple (S_{ex}), les variables « libres » sont x_2, x_3 et x_5, les variables « liées » sont x_1 et x_4.

Dans le système échelonné (S), on peut alors exprimer toutes les « variables liées » $x_{j_1}, \dots, x_{j_r}$ en fonction des« variables libres » x_j, $j \in \mathcal{L}$. Les coefficients exacts s'obtiennent en « remontant les calculs. Par exemple, dans (S_{ex}), le système initial est équivalent à

$$\begin{cases} x_1 = -2x_2 - 3x_3 + 7x_5 \\ x_4 = -3x_5 \\ 0 = 0 \\ 0 = 0. \end{cases} \tag{3.2}$$

Dans le cas général, on constate que (S) est équivalent à un système de la forme

$$\begin{cases} x_{j_1} = \displaystyle\sum_{j \in \mathcal{L}} \alpha_{1,j} x_j \\ x_{j_2} = \displaystyle\sum_{j \in \mathcal{L}} \alpha_{2,j} x_j \\ \vdots \\ x_{j_r} = \displaystyle\sum_{j \in \mathcal{L}} \alpha_{r,j} x_j \\ 0 = 0 \\ \vdots \\ 0 = 0. \end{cases} \tag{3.3}$$

où les coefficients $\alpha_{i,j}$, $i \in \{1, \dots, r\}$, $j \in \mathcal{L}$, s'obtiennent en « remontant les calculs » pour exprimer toutes les variables liées en fonction des variables libres.

On constate alors qu'en choisissant l'un des x_j, $j \in \mathcal{L}$ égal à 1 et tous les autres x_j, $j \in \mathcal{L}$, égaux à 0, on obtient une solution de (S).

Par exemple, pour (S_{ex}), en choisissant $x_2 = 1$, $x_3 = 0$ et $x_5 = 1$, on obtient une solution de (S_{ex}), à savoir

$$X_1 = \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

en choisissant $x_2 = 0$, $x_3 = 1$ et $x_5 = 0$, on obtient une solution de (S), à savoir

$$X_2 = \begin{pmatrix} -3 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

et en choisissant $x_2 = x_3 = 0$ et $x_5 = 1$, on obtient une troisième solution, donnée par

$$X_3 = \begin{pmatrix} 7 \\ 0 \\ 0 \\ -3 \\ 1 \end{pmatrix}$$

Dans le cas général, on obtient ainsi $(p-r)$ solutions X_j, $j \in \mathcal{L}$: pour $j \in \mathcal{L}$ la solution numéro j est obtenue en choisissant la variable libre x_j égale à 1 et les autres variables libres égales à 0.

Si l'on se donne, pour chaque $j \in \mathcal{L}$, un nombre réel β_j, la colonne $X = \sum_{j\in\mathcal{L}} \beta_j X_j$ est alors la colonne obtenue en choisissant β_j à toutes les positions $j \in \mathcal{L}$ correspondant à des variables libres, et en insérant, à chacune des positions $j_1, \dots, j_r$ correspondant à des variables liées, la valeur $x_{j_k} = \sum_{j\in\mathcal{L}} \alpha_{j,k}\beta_j$.

Cette colonne X fournit, vu la forme de (S_{fin}), une solution du système initial (S). De plus, en reprenant les remarques ci-dessus et les exemples vus au §S_{ech}, on constate que *toute* solution de (S) est de cette forme.

Nous pouvons résumer la discussion par l'énoncé suivant.

Théorème 5.23 – Résolution des systèmes échelonnés

Soit (S) un système linéaire homogène de n équations à p inconnues. Soit r le rang du système (S). Rappelons qu'on a $r \leq p$.

- *si $r = p$, alors l'unique solution de (S) est $\mathbf{0}_{1,p}$.*
- *si $r < p$, alors le système (S) admet une infinité de solutions.*

De plus, si X_j, $j \in \mathcal{L}$, sont les les $(p-r)$ solutions construites en choisissant « l'une des variables libres égales à 1 et les autres égales à 0 », alors l'ensemble des solutions de (S) est

$$\left\{ \sum_{j\in\mathcal{L}} \beta_j X_j, \text{ où } \beta_j \text{ est un réel pour chaque } j \in \mathcal{L} \right\}$$

3.3. Cas des systèmes linéaires non homogènes. — Cherchons à résoudre le système linéaire

$$\left\{ \begin{array}{rcrcrcl} x & + & 2y & + & 3z & = & 1 \\ 4x & + & 5y & + & 6z & = & 1 \\ 7x & +8y & + & 9z & = & 2 & \end{array} \right. \tag{3.4}$$

d'inconnue $(x, y, z) \in \mathbb{R}^3$.

Nous connaissons la réduite échelonnée de A : $U = \begin{pmatrix} 1 & 0 & 0 \\ 4/3 & -1/3 & 0 \\ 1 & -2 & 1 \end{pmatrix}$, alors $UA = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$ (§2.5).

Or, le système (3.4) est équivalent à l'égalité de matrices

$$A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \tag{3.5}$$

Cette égalité implique

$$UA \begin{pmatrix} x \\ y \\ z \end{pmatrix} = U \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}. \tag{3.6}$$

En fait, les égalités (3.5) et (3.6) sont *équivalentes*, parce que la matrice U est inversible : on peut donc retrouver la première à partir de la seconde en multipliant sur la gauche par U^{-1}.

En insérant le fait que nous connaissons la matrice UA, ainsi que le calcul de $U\begin{pmatrix}1\\1\\2\end{pmatrix} = \begin{pmatrix}1\\1\\1\end{pmatrix}$, nous constatons que le système initial (3.4) est équivalent à

$$\begin{cases} x & + & 2y & + & 3z & = & 1 \\ & & y & + & 2z & = & 1 \\ 0 & = 1 & & & & & \end{cases}$$

Ce dernier système n'a visiblement pas de solution ! Donc (3.4) n'en a pas non plus.

Si nous avions démarré, en revanche, avec le système

$$\begin{cases} x & + & 2y & + & 3z & = & 1 \\ 4x & + & 5y & + & 6z & = & 1 \\ 7x & +8y & + & 9z & = & 1 & \end{cases} \tag{3.7}$$

la discussion ci-dessus, notamment le fait que $U\begin{pmatrix}1\\1\\1\end{pmatrix} = \begin{pmatrix}1\\1\\0\end{pmatrix}$, montre que (3.7) est équivalent à

$$\begin{cases} \boxed{\text{x}} & + & 2y & + & 3z & = & 1 \\ & & \boxed{\text{y}} & + & 2\boxed{\text{z}} & = & 1 \\ & & & & 0 & = 1 & \end{cases} \tag{3.8}$$

et ce système a visiblement des solutions : par exemple, si l'on choisit $\boxed{z} = \boxed{0}$, on obtient une solution, à savoir

$$\begin{pmatrix}\boxed{x_0}\\ \boxed{y_0}\\ \boxed{z_0}\end{pmatrix} = \begin{pmatrix}\boxed{-1}\\ \boxed{2}\\ \boxed{0}\end{pmatrix} \tag{3.9}$$

Plus généralement, pour savoir si un système admet une solution, il est nécessaire de répondre à la question suivante : si, une fois trouvée une forme réduite échelonnée pour les membres de gauche, il y a des lignes entières de 0 sur la gauche, la transformation du membre de droite donne-t-elle également des zéros ? Nous pouvons résumer la discussion ci-dessus par l'énoncé formel suivant.

Proposition 5.24 – Conditions de compatibilité pour les systèmes inhomogènes

Soit (S) un système linéaire de n équations à p inconnues. Notons $\mathbf{b} = \begin{pmatrix}b_1\\ \vdots \\ b_n\end{pmatrix}$ la colonne des seconds membres de (S) et A la matrice des coefficients du membre de gauche de (S).
Soient r le rang de A et U une matrice inversible telle que UA soit une matrice échelonnée réduite. La matrice UA se termine par $(n-r)$ lignes de zéros.

- *Pour que le système (S) admette au moins une solution, il faut et il suffit que dans la colonne $U\mathbf{b}$, les $(n-r)$ dernières coordonnées soient toutes égales à zéro.*
- *Dans ce cas, si l'on note $\boxed{x_{j_1}}, \dots, \boxed{x_{j_r}}$ les "inconnues liées" et $\boxed{x_j}, j \in \mathcal{L}$ les "inconnues libres" discutées dans le §??, alors on obtient une solution particulière de (S) en choisissant toutes les inconnues "libres" $\boxed{x_j}$, $j \in \mathcal{L}$ égales à zéro, et en trouvant des valeurs pour les "inconnues liées" $\boxed{x_{j_1}}, \dots, \boxed{x_{j_r}}$ grâce à la réécriture de (S) sous forme échelonnée réduite.*

Maintenant que nous savons à quelle condition un système linéaire admet *au moins* une solution, voyons comment on peut *les trouver toutes.*

Nous reprenons l'exemple de

$$\left\{\begin{array}{rcrcrcl} x & + & 2y & + & 3z & = & 1 \\ 4x & + & 5y & + & 6z & = & 1 \\ 7x & + & 8y & + & 9z & = & 1 \end{array}\right. \tag{3.10}$$

et nous remarquons que nous avons trouvé, grâce à la réduite échelonnée, une solution $\begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}$:

elle vérifie bien sûr le système, on a donc

$$\left\{\begin{array}{rcrcrcl} x_0 & + & 2y_0 & + & 3z_0 & = & 1 \\ 4x_0 & + & 5y_0 & + & 6z_0 & = & 1 \\ 7x_0 & + & 8y_0 & + & 9z_0 & = & 1 \end{array}\right. \tag{3.11}$$

Comment trouver toutes les solutions (x, y, z) du système (3.10) ? Pour le voir, effectuons la différence entre (3.10) et (3.12) : on constate que si un triplet (x, y, z) est solution de (3.10), alors on doit avoir

$$\left\{\begin{array}{rcrcrcl} (x-x_0) & + & 2(y-y_0) & + & 3(z-z_0) & = & 1-1 \\ 4(x-x_0) & + & 5(y-y_0) & + & 6(z-z_0) & = & 1-1 \\ 7(x-x_0) & + & 8(y-y_0) & + & 9(z-z_0) & = & 1-1 \end{array}\right. \tag{3.12}$$

Nous constatons donc que le vecteur $\begin{pmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{z} \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} - \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}$ vérifie le système *homogène*

$$\left\{\begin{array}{rcrcrcl} \tilde{x} & + & 2\tilde{y} & + & 3\tilde{z} & = & 0 \\ 4\tilde{x} & + & 5\tilde{y} & + & 6\tilde{z} & = & 0 \\ 7\tilde{x} & + & 8\tilde{y} & + & 9\tilde{z} & = & 0 \end{array}\right. \tag{3.13}$$

qui est le même que le système initial (3.10), mais en "mettant à zéro tous les seconds membres". Nous avons vu comment résoudre ce système homogène au §3.2 (comme il s'agit d'un système homogène 3×3 dont le rang est 2, nous avons montré, notamment, qu'il admet une infinité de solutions).

Toute solution $\begin{pmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{z} \end{pmatrix}$ du système homogène (3.13) fournit une solution de (3.10), à savoir

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{z} \end{pmatrix} + \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}.$$

Nous constatons donc que (3.10) admet une infinité de solutions, et que nous les avons toutes trouvées à partir de la résolution du système homogène (3.13) et de la solution particulière (3.9).

Les arguments utilisés dans l'exemple précédent n'ont rien de spécifique et sont valables pour tout système linéaire. Nous pouvons résumer la discussion par l'énoncé général suivant.

Proposition 5.25 – Solution générale d'un système inhomogène : solution particulière + solution générale du système homogène

Soit (S) un système linéaire et (S_{hom}) le système homogène associé, obtenu en remplaçant les seconds membres de (S) par des zéros.

- *Si (S) admet une solution $\begin{pmatrix} u_1 \\ \vdots \\ u_p \end{pmatrix}$, alors un vecteur $\begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$ est solution de (S) si et seulement si la différence $\begin{pmatrix} x_1 - u_1 \\ \vdots \\ x_p - u_p \end{pmatrix}$ est solution du système (S_{hom}).*
- *Ainsi, si l'on parvient à trouver une solution particulière $\begin{pmatrix} u_1 \\ \vdots \\ u_p \end{pmatrix}$ de (S), alors l'ensemble des solutions de (S) est l'ensemble des vecteurs $\begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$ de la forme $\begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} = \begin{pmatrix} x_1^{\text{hom}} \\ \vdots \\ x_p^{\text{hom}} \end{pmatrix} + \begin{pmatrix} u_1 \\ \vdots \\ u_p \end{pmatrix}$, où $\begin{pmatrix} x_1^{\text{hom}} \\ \vdots \\ x_p^{\text{hom}} \end{pmatrix}$ est une solution du système homogène (S_{hom}).*

4. Conséquences pour l'inversibilité et le rang des matrices

4.1. Retour sur le noyau et l'image d'une matrice. — Considérons la matrice

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

et rappelons que nous avons introduit au chapitre 4 deux ensembles, le noyau $\text{Ker}(A)$ et l'image $\text{Im}(A)$. Ces deux ensembles ont des liens avec les systèmes linéaires :

- Le noyau $\text{Ker}(A)$ est l'ensemble des $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ de $\mathbb{R}^3$ pour lesquels on a l'égalité $A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$. C'est donc l'ensemble des solutions du système linéaire homogène

$$\begin{cases} x + 2y + 3z = 0 \\ 4x + 5y + 6z = 0 \\ 7x + 8y + 9z = 0 \end{cases}$$

Nous avons vu des outils pour résoudre ce système. Nous avons en fait montré qu'il est équivalent à

$$\begin{cases} x + 2y + 3z = 0 \\ y + 2z = 0 \\ 0 = 0 \end{cases}$$

et ce système admet une infinité de solutions : tous les $\begin{pmatrix} \alpha \\ -2\alpha \\ \alpha \end{pmatrix}$, $\alpha \in \mathbb{R}$.

- Quant à l'image $\mathrm{Ker}(A)$, il s'agit de l'ensemble est l'ensemble des $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$ de $\mathbb{R}^3$ pour lesquels il existe $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3$ vérifiant l'égalité $A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$. C'est donc l'ensemble des $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$ pour lesquels le système linéaire

$$\left\{ \begin{array}{rcrcrcl} x & + & 2y & + & 3z & = & a \\ 4x & + & 5y & + & 6z & = & b \\ 7x & + & 8y & + & 9z & = & c \end{array} \right. \tag{4.1}$$

de second membre (a, b, c) et d'inconnue $(x, y, z) \in \mathbb{R}^3$, admet au moins une solution.

Or, nous avons vu comment trouver la réponse à cette question, en utilisant le fait que la réduite échelonnée de $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ est $S = UA = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$, où $U = \begin{pmatrix} 1 & 0 & 0 \\ 4/3 & -1/3 & 0 \\ 1 & -2 & 1 \end{pmatrix}$. Le système (4.1) est donc équivalent à $S \begin{pmatrix} x \\ y \\ z \end{pmatrix} = U \begin{pmatrix} a \\ b \\ c \end{pmatrix}$, ou encore à

$$\left\{ \begin{array}{rcrcrcl} x & + & 2y & + & 3z & = & a \\ 4x & + & y & + & 2z & = & (4/3)a - (1/3)b \\ & & & & 0 & = & a - 2b + c \end{array} \right. .$$

La dernière ligne permet de savoir si (4.1) admet ou non une solution : si $a - 2b + c = 0$, il en admet une, et si $a - 2b + c \neq 0$, il n'en admet pas.

Nous avons trouvé une condition nécessaire et suffisante sur (a, b, c) pour qu'il appartienne à $\mathrm{Im}(A)$! Ainsi, $\mathrm{Im}(A)$ est l'ensemble des $\begin{pmatrix} a \\ b \\ c \end{pmatrix} \in \mathbb{R}^3$ vérifiant la condition $a-2b+c = 0$. Par exemple, on constate que $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ appartient à $\mathrm{Im}(A)$, mais $\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$ n'appartient pas à $\mathrm{Im}(A)$, ce qui n'est pas surprenant après la discussion du §3.3.

En général, il est possible de déterminer le noyau et l'image d'une matrice grâce aux méthodes de résolution des systèmes linéaires vues ci-dessus. Nous n'écrirons pas d'énoncé formel ici.

Nous terminons par une remarque importante (et inattendue) sur le cas des *matrices carrées*. Les ensembles $\mathrm{Ker}(A)$ et $\mathrm{Im}(A)$ sont de nature très différente en apparence. Mais leur détermination se ramène à l'étude des systèmes linéaires dont la matrice est A.

Considérons une matrice A de format $n \times n$. Lorsque la réduite échelonnée S de A ne comporte *pas de ligne de zéros*, le rang de A est n. On peut alors faire les remarques suivantes :

- un système échelonné de matrice S admet *toujours* une solution, quels que soient le membre de droite (puisqu'il n'y a pas de "condition de compatibilité" venant des lignes de zéros).
- de plus, il admet toujours une *unique* solution : si A est une matrice $n \times n$ et si le rang de A est n, on s'aperçoit en analysant la notion de matrice échelonnée que la réduite échelonnée de A est nécessairement *triangulaire supérieure avec des* 1 *sur la diagonale.*

Une conséquence de la première remarque est que toute colonne $\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ appartient à $\mathrm{Im}(A)$. On a donc $\mathrm{Im}(A) = \mathbb{R}^n$ dans ce cas.

Une conséquence de la seconde remarque est que la colonne $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0\dot{:}0$ est la seule solution de $A \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ (puisque c'en est une, et qu'il y a au plus une solution). On a donc $\mathrm{Ker}(A) = \mathbf{0}_{n,1}$ dans ce cas.

Ces remarques admettent des réciproques, résumées dans la proposition suivante.

Proposition 5.26 – Lien entre noyau et image pour les matrices carrées

Si A est une matrice carrée de taille n, alors les trois assertions suivantes sont équivalentes :

(i) $\mathrm{rg}(A) = n$.
(ii) $\mathrm{Im}(A) = \mathbb{R}^n$
(iii) $\mathrm{Ker}(A) = \{\mathbf{0}_{\mathbb{R}^n}\}$

La proposition ci-dessus est extrêmement importante dans l'étude des matrices. Afin d'expliquer pourquoi, introduisons l'application

$$\begin{array}{rccc} f_A & : & \mathbb{R}^n & \to & \mathbb{R}^n \\ & & \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} & \mapsto & A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}. \end{array} \tag{4.2}$$

Il y a un lien entre cette application et les propriétés du noyau et de l'image apparues dans la proposition précédente :

- Dire que $\mathrm{Im}(A) = \mathbb{R}^n$ revient à dire que toute colonne $\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ peut s'écrire sous la forme $A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ pour un certain $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$: cela revient à dire que f_A est surjective.
- Si f_A est injective, alors on doit avoir $\mathrm{Ker}(A) = \{\mathbf{0}_{n,1}\}$: si ce n'était pas le cas, l'élément $\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ de l'espace d'arrivée de f_A admettrait plusieurs antécédents par f_A.
- Réciproquement, si $\mathrm{Ker}(A) = \{\mathbf{0}_{n,1}\}$, alors f_A est injective. En effet, si X, X' sont deux colonnes vérifiant $f_A(X) = f_A(X')$, alors $AX = AX'$, donc $A(X - X') = \mathbf{0}_{n,1}$. Dans ce cas $(X - X')$ appartient à $\mathrm{Ker}(A)$, et sous notre hypothèse $\mathrm{Ker}(A) = \{\mathbf{0}_{n,1}\}$, on peut en déduire $X - X' = \mathbf{0}_{n,1}$, d'où $X = X'$.

Nous constatons donc que dans la proposition 5.26, la condition (ii) signifie que f_A est injective, alors que la condition (iii) signifie que f_A est surjective. Nous obtenons ainsi le résultat suivant :

Corollaire 5.27 – Un lien inattendu entre injectivité et surjectivité

Soit A une matrice carrée de taille n. Considérons l'application $f_A : \mathbb{R}^n \to \mathbb{R}^n$ définie en (4.2) Les deux assertions suivantes sont équivalentes :

1. *f_A est injective*
2. *f_A est surjective.*

Ce résultat est surprenant, car pour une application $f : E \to F$ entre deux ensembles quelconques, il n'existe habituellement aucun lien entre l'injectivité de f et la surjectivité de f : nous avons vu au chapitre 2 des exemples où f est surjective sans être injective, surjective sans être injective... Le corollaire ci-dessus montre que les application f_A associées à des matrices comme en (4.2), ont des propriétés extrêmement particulières. Vous étudierez en détail beaucoup de ces propriétés dans le cours d'algèbre du second semestre.

4.2. Inversibilité et méthode du "pivot". — Il y a un lien simple entre la notion d'inversibilité pour une matrice carrée A et les systèmes linéaires. En effet, considérons une matrice A de format $n \times n$ et supposons A inversible. Il existe alors une matrice B vérifiant $AB = BA = I_n$. Mais alors, pour tout $\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in \mathbb{R}^n$, le système linéaire

$$A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

admet *une et une seule solution* : en multipliant sur la gauche par B, on trouve $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = B \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$.

De plus, on peut trouver la matrice $B = A^{-1}$ grâce à la résolution de tels systèmes linéaires : en effet, nous avons vu au chapitre précédent (proposition 4.26) que $B \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ donne la première colonne de B, que $B \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}$ donne la deuxième, etc. Ainsi, si l'on résout le système linéaire

$$A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

la solution $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ donnera la première colonne de A^{-1}, et en variant la position du 1 au second membre on pourra trouver (après beaucoup de calculs) toutes les colonnes de A^{-1}.

Nous montrons à présent qu'on peut en fait, sans passer par les systèmes linéaires à proprement parler, déterminer efficacement à l'aide de la "méthode du pivot" si A est inversible, et calculer A^{-1} lorsque cette matrice existe. Nous montrons cela par deux exemples simples.

Exemple 5.28. — Considérons la matrice

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}.$$

Nous avons vu au §**??** que la réduite échelonnée de A est

$$S = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

et qu'on a

$$S = UA, \quad \text{où } U \text{ est une matrice inversible.}$$

Si A était inversible, comme U l'est aussi, la matrice $S = UA$ serait inversible. Or ce n'est pas le cas : S n'est pas inversible. En effet, S est une matrice dont la dernière ligne ne comporte que des zéros, et pour toute matrice S' de taille $n \times n$, le calcul de SS' montre que la dernière ligne de SS' est nulle, on ne peut donc jamais avoir $SS' = I_3$.

Il est donc impossible que A soit inversible.

exem

Exemple 5.29 **(Un calcul d'inverse).** — Considérons à présent la matrice

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 3 & 5 \\ -2 & -2 & 1 \end{pmatrix}.$$

Cherchons d'abord la réduite échelonnée de A à l'aide d'opérations élémentaires.

$$\text{si } U_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{alors} \quad U_0 A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 3 & 5 \\ -2 & -2 & 1 \end{pmatrix}$$

$$\text{si } U_1 = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}, \quad \text{alors} \quad U_1 A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 4 \\ 0 & 0 & 3 \end{pmatrix} \qquad \begin{matrix} L_2 \leftarrow L_2 - L_1 \\ L_3 \leftarrow L_3 + 2L_1 \end{matrix}$$

$$\text{si } U_2 = \begin{pmatrix} 1 & 0 & 0 \\ -1/2 & 1/2 & 0 \\ 2/3 & 0 & 1/3 \end{pmatrix}, \quad \text{alors} \quad U_2 A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \qquad \begin{matrix} L_2 \leftarrow \frac{1}{2} L_2 \\ L_3 \leftarrow \frac{1}{3} L_3 \end{matrix}$$

La réduite échelonnée de A est ainsi une matrice triangulaire supérieure avec une diagonale ne comportant que des 1. D'après les critères vus au paragraphe précédent, il est clair que pour $\begin{pmatrix} a \\ b \\ c \end{pmatrix} \in \mathbb{R}^3$, le système $A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ admet une unique solution, donc A est inversible.

On peut aller un cran plus loin pour trouver l'inverse de A. En effet, si nous trouvons une matrice B telle que $BA = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, alors nous aurons trouvé A^{-1}. Or, nous avons trouvé une matrice U telle que UA soit égale à $A' = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$. Cette dernière matrice n'est pas égale à I_3... mais il n'est pas difficile de voir qu'il est possible d'obtenir I_3 à partir de cette dernière matrice, en effectuant des *transvections remontantes*. En effet, on peut faire apparaître des zéros dans la dernière colonne en s'appuyant sur la dernière ligne de A' :

$$\text{si } U_3 = \begin{pmatrix} 1/3 & 0 & -1/3 \\ -11/6 & 1/2 & -2/3 \\ 2/3 & 0 & 1/3 \end{pmatrix}, \quad \text{alors} \quad U_3A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \begin{matrix} L_1 \leftarrow L_1 - L_3 \\ L_2 \leftarrow L_2 - 2L_3 \end{matrix}$$

On peut ensuite faire apparaître des zéros dans la deuxième colonne en effectuant une nouvelle « transvection remontante » :

$$\text{si } U_4 = \begin{pmatrix} 13/6 & -1/2 & 1/3 \\ -11/6 & 1/2 & -2/3 \\ 2/3 & 0 & 1/3 \end{pmatrix}, \quad \text{alors} \quad U_4A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad L_1 \leftarrow L_1 - L_2$$

Nous avons réussi à trouver une matrice U_4 telle que $U_4A = I_3$! C'est donc que A est inversible et que A^{-1} est donnée par $U_4 = \begin{pmatrix} 3 & -1/2 & 1/3 \\ -11/6 & 1/2 & -2/3 \\ 2/3 & 0 & 1/3 \end{pmatrix}$.

En adoptant la même stratégie, on constate le fait suivant.

Proposition 5.30

Soit A une matrice carrée de format $n \times n$. Il y a équivalence entre :

- *la matrice A est inversible*
- *la réduite échelonnée de A est une matrice triangulaire supérieure dont la diagonale ne comporte que des 1.*

Lorsque A est inversible, on peut trouver l'inverse de A en cherchant la réduite échelonnée de A, puis en effectuant des « transvections remontantes » pour passer de la réduite échelonnée de A à la matrice I_n.

Exercices du chapitre 5

***Exercice 5.1* (Systèmes 2×2).** — ★☆☆

Trouver, pour chacun des systèmes suivants, tous les couples $(x_1, x_2) \in \mathbb{C}^2$ vérifiant les conditions données.

1. $\begin{cases} 2x_1 + x_2 = 0 \\ x_1 + 2x_2 = 0 \end{cases}$
2. $\begin{cases} 2x_1 + x_2 = 3 \\ x_1 + 2x_2 = 3 \end{cases}$
3. $\begin{cases} 2x_1 + 4x_2 = 10 + 2i \\ 2x_1 + (4+2i)x_2 = 6 + 2i \end{cases}$

***Exercice 5.2* (Systèmes 2×2 avec paramètres).** — ★☆☆

Dans cet exercice, on fixe deux réels a et b et on considère les systèmes linéaires suivants d'inconnue $(x_1, x_2) \in \mathbb{R}^2$:

1. $\begin{cases} 2x_1 + x_2 = 1 \\ 4x_1 + 2x_2 = b \end{cases}$
2. $\begin{cases} 2ax_1 + ax_2 = 0 \\ x_1 + 2x_2 = b \end{cases}$

Déterminer pour chacun d'eux l'ensemble des solutions, en discutant selon les valeurs de a et b.

***Exercice 5.3* (Résolution de systèmes explicites).** — ★☆☆

Résoudre les systèmes suivants, d'inconnue $(x, y, z) \in \mathbb{R}^3$ pour les deux premiers, d'inconnue $(x, y, z, t) \in \mathbb{R}^4$ pour le dernier :

$$(S_1) \begin{cases} 2x + y + 2z = 7 \\ x + y + z = 4 \\ -2x + y - 2z = -4 \end{cases} \qquad (S_2) \begin{cases} 2x + y + 2z = 7 \\ x + y + 2z = 4 \\ -2x + y - z = -3 \end{cases}$$

$$(S_3) \begin{cases} x - 2y + 3z - 4t = 4 \\ y - z + t = -3 \\ x + 3y - 3t = 1 \\ x + 2y + z - 4t = 4 \end{cases}$$

***Exercice 5.4* (Systèmes avec paramètres).** — ★★☆

1. On fixe deux réels a et b. Résoudre le système suivant, d'inconnue $(x, y, z) \in \mathbb{R}^3$, en discutant selon la valeur de a et b :
$$\begin{cases} x + y + 2z = 1 \\ x + 2y + z = 2 \\ 3x + 4y + 5z = a \\ y + 3z = b \end{cases}$$
2. On fixe un réel m. Résoudre le système suivant, d'inconnue $(x, y, z) \in \mathbb{R}^3$, en discutant selon la valeur de m :
$$\begin{cases} x + y + (2m-1)z = 1 \\ mx + y + z = -1 \\ x + my + z = 3(m+1) \end{cases}$$

Exercice 5.5. — ★☆☆

1. Un système linéaire peut-il avoir exactement trois solutions ? Pourquoi ?
2. Fixons $n \in \mathbb{N}^\star$. Un système linéaire de n équations à n inconnues a-t-il toujours exactement une solution ? au moins une solution ? au plus une solution ?

***Exercice 5.6* (Notion de rang d'un système). — ★☆☆**

1. Considérons un système linéaire (S) de 7 équations à 5 inconnues. Si le rang de (S) est 4, le système a-t-il nécessairement au moins une solution ? au plus une solution ? une unique solution ?
2. Mêmes questions si (S) est un système linéaire de 7 équations à 5 inconnues et si le rang de (S) est 5.
3. Mêmes questions si (S) est un système linéaire de 5 équations à 7 inconnues et si le rang de (S) est 4, puis si (S) est un système de 5 équations à 7 inconnues dont le rang est 5.

***Exercice 5.7.* — ★☆☆**

Soit M une des matrices suivantes. Calculer la réduite échelonnée S de M et une matrice inversible U telle que $UM = S$.

$$\begin{pmatrix} 5 & 2 & 3 \\ 2 & -2 & 5 \\ 3 & 4 & 2 \end{pmatrix}, \quad \begin{pmatrix} 2 & -1 & 3 \\ 3 & -5 & 1 \\ 4 & -7 & 1 \end{pmatrix}, \quad \begin{pmatrix} 2 & 7 & 3 & 1 \\ 1 & 3 & 5 & -2 \\ 1 & 5 & -9 & 8 \\ 5 & 18 & 4 & 5 \end{pmatrix}$$

***Exercice 5.8.* — ★☆☆**

Résoudre les systèmes suivants, d'inconnue $(x, y, z) \in \mathbb{R}^3$ pour les deux premiers, d'inconnue $(x, y, z, t) \in \mathbb{R}^4$ pour le dernier :

$$(S_1) \begin{pmatrix} 2 & 3 & 5 \\ 3 & 7 & 4 \\ 1 & 2 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 6 \\ 1 \end{pmatrix} \qquad (S_2) \begin{pmatrix} 2 & -1 & 3 \\ 3 & -5 & -7 \\ 4 & -7 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$(S_3) \begin{pmatrix} 2 & 7 & 3 & 1 \\ 1 & 3 & 5 & -2 \\ 1 & 5 & -9 & 8 \\ 5 & 18 & 4 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 13 \\ 7 \\ 5 \\ 32 \end{pmatrix}.$$

***Exercice 5.9.* — ★☆☆**

Soit M une des matrices de l'exercice 5.7.

Trouver le rang r de M et des matrices inversibles P et Q vérifiant : $M = \begin{pmatrix} \mathbf{I}_r & 0 \\ 0 & 0 \end{pmatrix}$.

***Exercice 5.10* (Inversibilité et méthode du pivot). — ★☆☆**

Les matrices suivantes sont-elles inversibles ? Si oui, quel est leur inverse ?

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 2 \\ 2 & 0 & 1 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 1 & 1 & 2 \\ 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & 1 \\ 2 & -1 & -1 & 1 \end{pmatrix} \quad C = \begin{pmatrix} 1 & 1 & 1 & 2 \\ 1 & 0 & 1 & 2 \\ 0 & 1 & -2 & 1 \\ 2 & -1 & -1 & 1 \end{pmatrix}$$

***Exercice 5.11* (Inversibilité et méthode du pivot). — ★☆☆**

Soient s et t deux nombres réels. Pour quelles valeurs de s et t les matrices suivantes sont-elles inversibles ? Déterminer leur inverse dans ce cas.

$$A = \begin{pmatrix} 1 & t \\ t & 1 \end{pmatrix}, \quad B = \begin{pmatrix} s & 1-s \\ s & t \end{pmatrix}, \quad C = \begin{pmatrix} t & 1 & -1 \\ 1 & t & 1 \\ 1 & 1 & t \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 1 & 0 \\ 4s^2 & 1-2s^2 & -2st \\ 2s & -s & t \end{pmatrix}.$$