Cours d'algebre pour la licence et le Capes

Jean-Étienne ROMBALDI

6 juillet 2007

Table des matières

Avant-propos

Ce livre est en construction.

Cet ouvrage destiné aux étudiants préparant le Capes externe de Mathématiques et aux enseignants préparant l'agrégation interne fait suite au livre « Éléments d'analyse réelle pour le Capes et l'Agrégation Interne de Mathématiques ».

Notations

$\mathbb{N}$	ensemble des entiers naturels.
$\mathbb{Z}$	l'anneau des entiers relatifs.
$\mathbb{Q}$	corps des nombres rationnels.
$\mathbb{R}$	corps des nombres réels.
$\mathbb{C}$	corps des nombres complexes.
$\Re(z)$	partie réelle du nombre complexe z.
$\Im(z)$	partie imaginaire du nombre complexe z.
$\mathbb{K}[X]$	algèbre des polynômes à une indéterminée à coefficients dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.
C_n^p	coefficient binomial.

1

Éléments de logique et de théorie des ensembles

Pour les exemples et exercices traités dans ce chapitre les ensembles usuels de nombres entiers, rationnels réels et complexes sont supposés connus, au moins de manière intuitive comme cela se passe au Lycée. Nous reviendrons plus loin sur les constructions de ces ensembles.

1.1 Quelques notions de logique

Nous allons préciser à un premier niveau quelques notions mathématiques qui sont relativement intuitives mais nécessitent quand même des définitions rigoureuses.

L'idée étant de préciser schématiquement comment se présente une théorie mathématique ainsi que la notion essentielle de démonstration.

La première notion est celle d'assertion. De manière intuitive, une assertion est un énoncé mathématique aussi rigoureux que possible qui ne peut prendre que deux valeurs de vérité à savoir « vrai » ou « faux » mais jamais entre les deux comme dans le langage courant.

Une assertion qui est toujours vraie est une tautologie.

Par exemple les énoncés suivantes sont des assertions : $2 < 15$ (elle est vraie), $\sqrt{2}$ est un nombre rationnel (elle est fausse), $\cos(n\pi) = (-1)^n$ (vraie), ...

Deux assertions sont dites logiquement équivalentes, ou plus simplement équivalentes, si elles sont toutes deux vraies ou toutes deux fausses.

Il y a ensuite les énoncés qui se démontrent. Pour ce faire, on se donne des règles précises (que nous verrons par la pratique) qui permettent de construire de nouvelles assertions à partir d'assertions données.

Remarque 1.1 *Il ne faut pas croire que dans une théorie donnée toute assertion P soit obligatoirement démontrable. En 1931 Kurt Gödel à démontré qu'il y a des assertions non démontrables (on dit aussi qu'elles sont indécidables) : il n'est pas possible de démontrer que P est vraie ni que P est fausse.*

À la base de toute théorie mathématique, on dispose d'un petit nombre d'assertions qui sont supposés vraies à priori (c'est-à-dire avant toute expérience) et que l'on nomme axiomes ou postulats. Ces axiomes sont élaborés par abstraction à partir de l'intuition et ne sont pas déduits d'autres relations.

Par exemple, la géométrie euclidienne est basée sur une quinzaine d'axiomes. L'un de ces axiomes est le postulat numéro 15 qui affirme que par un point donné passe une et une seule droite parallèle à une droite donnée.

Une autre exemple important est donné par la construction de l'ensemble noté $\mathbb{N}$ des entiers naturels. Cette construction peut se faire en utilisant les axiomes de Peano suivants :

- 0 est un entier naturel ;
- tout entier naturel n a un unique successeur noté $n+1$;
- deux entiers naturels ayant même successeur sont égaux ;
- une partie P de $\mathbb{N}$ qui contient 0 et telle que si n est dans P alors le successeur de n y est aussi, est égale à $\mathbb{N}$ (axiome de récurrence).

Nous reviendrons au paragraphe 1.6 sur l'ensemble $\mathbb{N}$ en partant sur une autre base.

La théorie des ensemble est basée sur le système d'axiomes de Zermelo-Fränkel.

La notion de définition nous permet de décrire un objet ou une situation précise à l'aide du langage courant.

Les énoncés qui se démontrent sont classés en fonction de leur importance dans une théorie comme suit :

- un théorème est une assertion vraie déduite d'autres assertions, il s'agit en général d'un résultat important à retenir ;
- un lemme est un résultat préliminaire utilisé pour démontrer un théorème ;
- un corollaire est une conséquence importante d'un théorème ;
- une proposition est de manière générale un résultat auquel on peut attribuer la valeur vraie ou fausse sans ambiguïté.

Pour rédiger un énoncé mathématique, on utilise le langage courant et les objets manipulés sont représentés en général par des lettres de l'alphabet latin ou grec. Usuellement, on utilise :

- les lettres minuscules $a, b, c, ...$ pour des objets fixés ;
- les lettres minuscules $x, y, z, t, ...$ pour des objets inconnus à déterminer ;
- les lettres majuscules $E, F, G, H, ...$ pour des ensembles ;
- des lettres de l'alphabet grecques minuscules ou majuscules $\alpha, \beta, \varepsilon, \delta, ...$ $\Lambda, \Gamma, \Omega, ...$

1.2 Les connecteurs logiques de base

L'élaboration de nouvelles assertions à partir d'autres se fait en utilisant les connecteurs logiques de négation, de conjonction, de disjonction, d'implication et d'équivalence définis comme suit, où P et Q désignent des assertions.

- La négation de P, notée $\neg P$, ou non P ou $\overline{P}$, est l'assertion qui est vraie si P est fausse et fausse si P est vraie.
  Par exemple la négation de l'assertion : « x est strictement positif » est « x est négatif ou nul ».
  En théorie des ensembles on admet qu'il n'existe pas d'assertion P telle que P et $\overline{P}$ soient toutes deux vraies. On dit que cette théorie est non contradictoire.
- La conjonction de P et Q, notée $P \wedge Q$ (lire P et Q), est l'assertion qui est vraie uniquement si P et Q sont toutes deux vraies (et donc fausse dans les trois autres cas).
  Par exemple $P \wedge \overline{P}$ est toujours faux (on se place dans des théories non contradictoires).
- La disjonction de P et Q, notée $P \vee Q$ (lire P ou Q), est l'assertion qui est vraie uniquement si l'une des deux assertions P ou Q est vraie (donc fausse si P et Q sont toutes deux fausses).
  Par exemple $P \vee \overline{P}$ est toujours vraie (c'est une tautologie).
  Il faut remarquer que le « ou » pour « ou bien » est inclusif, c'est-à-dire que P et Q peuvent être toutes deux vrais dans le cas où $P \vee Q$ est vraie.
  On peut aussi introduire le « ou exclusif », noté W, qui est vrai uniquement lorsque l'une des deux assertions, mais pas les deux simultanément, est vraie.

– L'implication, notée $P \to Q$, est l'assertion qui est fausse uniquement si P est vraie et Q fausse (donc vraie dans les trois autres cas).
On peut remarquer que si P est fausse, alors $P \to Q$ est vraie indépendamment de la valeur de vérité de Q.
L'implication est à la base du raisonnement mathématique. En partant d'une assertion P (ou de plusieurs), une démonstration aboutit à un résultat Q. Si cette démonstration est faite sans erreur, alors $P \to Q$ est vraie et on notera $P \Rightarrow Q$ (ce qui signifie que si P est vraie, alors Q est vraie). Dans ce cas, on dit que P est une condition suffisante et Q une condition nécessaire.
On peut remarquer que l'implication est transitive, c'est-à-dire que si P implique Q et Q implique R, alors P implique R.

– L'équivalence de P et Q, notée $P \leftrightarrow Q$, est l'assertion qui est vraie uniquement si $P \to Q$ et $Q \to P$ sont toutes deux vraies. Dans le cas où $P \leftrightarrow Q$ est vraie on dit que P est Q sont équivalentes et on note $P \Leftrightarrow Q$ (ce qui signifie que P et Q sont, soit toutes deux vraies, soit toutes deux fausses). Dans ce cas, on dit que Q est une condition nécessaire et suffisante de P.

On peut résumer ce qui précède, en utilisant la table de vérité suivante :

P	Q	$\overline{P}$	$P \wedge Q$	$P \vee Q$	$P \to Q$	$P \leftrightarrow Q$
V	V	F	V	V	V	V
V	F	F	F	V	F	F
F	V	V	F	V	V	F
F	F	V	F	F	V	V

Les tables de vérité peuvent être utilisées pour faire certaines démonstrations. On rappelle que deux assertions qui ont même table de vérité sont équivalentes.

Avec le théorème qui suit, on résume quelques règles de calcul.

Théorème 1.1 *Soient P, Q, R des propositions. On a les équivalences :*

1. *commutativité :*
$$(P \wedge Q) \Leftrightarrow (Q \wedge P)$$
$$(P \vee Q) \Leftrightarrow (Q \vee P)$$

2. *associativité*
$$(P \wedge (Q \wedge R)) \Leftrightarrow ((P \wedge Q) \wedge R)$$
$$(P \vee (Q \vee R)) \Leftrightarrow ((P \vee Q) \vee R)$$

3. *distributivité :*
$$(P \wedge (Q \vee R)) \Leftrightarrow ((P \wedge Q) \vee (P \wedge R))$$
$$(P \vee (Q \wedge R)) \Leftrightarrow ((P \vee Q) \wedge (P \vee R))$$

4. *négations :*
$$\left(\overline{\overline{P}}\right) \Leftrightarrow (P)$$
$$\left(\overline{P \wedge Q}\right) \Leftrightarrow \left(\overline{P} \vee \overline{Q}\right)$$
$$\left(\overline{P \vee Q}\right) \Leftrightarrow \left(\overline{P} \wedge \overline{Q}\right)$$
$$(P \to Q) \Leftrightarrow \left(\overline{Q} \to \overline{P}\right)$$
$$(P \to Q) \Leftrightarrow \left(\overline{P} \vee Q\right)$$
$$\left(\overline{P \to Q}\right) \Leftrightarrow \left(P \wedge \overline{Q}\right)$$

Démonstration. On utilise les tables de vérité (exercices). ■
Les équivalences $\left(\overline{P \wedge Q}\right) \Leftrightarrow \left(\overline{P} \vee \overline{Q}\right)$ et $\left(\overline{P \vee Q}\right) \Leftrightarrow \left(\overline{P} \wedge \overline{Q}\right)$ sont appelées lois de Morgan.

Exercice 1.1 *Montrer que les assertions $P \to Q$ et $\overline{P} \vee Q$ sont équivalentes.*

Solution 1.1 *On montre qu'elles ont même table de vérité.*

P	Q	$\overline{P}$	$\overline{P} \vee Q$	$P \to Q$
V	V	F	V	V
V	F	F	F	F
F	V	V	V	V
F	F	V	V	V

Exercice 1.2 *Montrer que les assertions $\overline{P \to Q}$ et $P \wedge \overline{Q}$ sont équivalentes.*

Solution 1.2 *On montre qu'elles ont même table de vérité.*

P	Q	$P \wedge \overline{Q}$	$\overline{P \to Q}$
V	V	F	F
V	F	V	V
F	V	F	F
F	F	F	F

Exercice 1.3 *Montrer que les assertions $P \leftrightarrow P$, $(P \wedge Q) \to P$, $P \to (P \vee Q)$, $P \vee (P \to Q)$, $P \to (Q \to P)$ et $((P \to Q) \to P) \to P$ sont des tautologies (i. e. toujours vraies).*

Solution 1.3 *Pour $P \leftrightarrow P$, $(P \wedge Q) \to P$, $P \to (P \vee Q)$, c'est évident et pour les autres, on utilise la table de vérité :*

P	Q	$P \to Q$	$Q \to P$	$P \vee (P \to Q)$	$P \to (Q \to P)$	$((P \to Q) \to P) \to P$
V	V	V	V	V	V	V
V	F	F	V	V	V	V
F	V	V	F	V	V	V
F	F	V	V	V	V	V

Exercice 1.4 *Simplifier l'expression :*

$$R = \left(\overline{P} \wedge Q\right) \vee \left(\overline{P} \wedge \overline{Q}\right) \vee (P \wedge Q).$$

Solution 1.4 *En utilisant les tables de vérité, on a :*

P	Q	$\overline{P} \wedge Q$	$\overline{P} \wedge \overline{Q}$	$\left(\overline{P} \wedge Q\right) \vee \left(\overline{P} \wedge \overline{Q}\right)$	$P \wedge Q$	R
V	V	F	F	F	V	V
V	F	F	F	F	F	F
F	V	V	F	V	F	V
F	F	F	V	V	F	V

Donc R a la même table de vérité que $P \to Q$, ce qui signifie que R est équivalent à $P \to Q$.

Exercice 1.5 *Soient P, Q, R trois assertions.*

1. *Écrire la négation de chacune de ces assertions :* $\overline{P}\wedge Q$, $\overline{P}\vee Q$, $P\vee(Q\wedge R)$, $P\wedge(Q\vee R)$, $P\to\overline{Q}$, $P\leftrightarrow Q$, $\overline{P\vee Q}\to R$, $P\vee Q\to\overline{R}$, $\overline{P}\wedge Q\Rightarrow R$ *et* $\overline{P}\vee Q\to\overline{R}$.
2. *Traduire chacune de ces assertions, ainsi sa négation, en langage courant où* P *correspond à « j'écris », Q à « je pense » et R à « je chante ».*

Solution 1.5 *On a :*

$$\overline{\overline{P}\wedge Q}=P\vee\overline{Q}$$

ce qui peut se traduire par la négation de « je n'écris pas et je pense » est « j'écris ou je ne pense pas » ;

$$\overline{\overline{P}\vee Q}=P\wedge\overline{Q}$$

$$\overline{P\vee(Q\wedge R)}=\overline{P}\wedge\overline{Q\wedge R}=\overline{P}\wedge\left(\overline{Q}\vee\overline{R}\right)=\left(\overline{P}\wedge\overline{Q}\right)\vee\left(\overline{P}\wedge\overline{R}\right)$$

et ainsi de suite.

Exercice 1.6 *Montrer les équivalences qui suivent.*

1. $(P\to(Q\to R))\Leftrightarrow((P\wedge Q)\to R)$
2. $((P\vee Q)\to R)\Leftrightarrow((P\to R)\wedge(Q\to R))$
3. $((P\wedge Q)\to R)\Leftrightarrow((P\to R)\vee(Q\to R))$
4. $(P\to(Q\wedge R))\Leftrightarrow((P\to Q)\wedge(P\to R))$
5. $(P\to(Q\vee R))\Leftrightarrow((P\to Q)\vee(P\to R))$

Solution 1.6 *On peut utiliser les tables de vérité ou utiliser l'équivalence* $(P\to Q)\Leftrightarrow\left(\overline{P}\vee Q\right)$. *Par exemple, on a :*

$$(P\to(Q\to R))\Leftrightarrow\overline{P}\vee\left(\overline{Q}\vee R\right)\Leftrightarrow\overline{P\wedge Q}\vee R\Leftrightarrow((P\wedge Q)\to R)$$

Exercice 1.7 *Montrer que les assertions PWQ (ou exclusif) et* $\left(P\wedge\overline{Q}\right)\vee\left(\overline{P}\wedge Q\right)$ *sont équivalentes.*

Solution 1.7 *On montre qu'elles ont même table de vérité.*

P	Q	PWQ	$\left(P\wedge\overline{Q}\right)\vee\left(\overline{P}\wedge Q\right)$
V	V	F	F
V	F	V	V
F	V	V	V
F	F	F	F

Exercice 1.8 *Soient* a,b *deux entiers naturels.*

1. *Donner un équivalent de* $(a<b)\to(a=b)$
2. *Donner la négation de* $(a\leq b)\to(a>b)$

Solution 1.8

1. $(a<b)\to(a=b)$ *est équivalent à* $(a\geq b)\vee(a=b)$ *encore équivalent à* $a\geq b$.
2. *La négation de* $(a\leq b)\to(a>b)$ *est* $(a\leq b)\wedge(a\leq b)$, *soit* $(a\leq b)$.

Exercice 1.9 *On dispose de* 6 *pièces de* 1 *euro dont une seule est fausse et plus lourde que les autres. Montrer qu'on peut la détecter en utilisant une balance de type Roberval en effectuant au plus deux pesées. Même question avec* 8 *pièces.*

Solution 1.9 *On numérote de* 1 *à* 6 *les pièces. On place les pièces* $1, 2, 3$ *sur le plateau* P_1 *de la balance et les pièces* $4, 5, 6$ *sur le plateau* P_2*. L'un des deux plateaux, disons* P_1 *est plus chargé, il contient donc la fausse pièce. On isole la pièce* 3 *et on place la pièce* 1 *sur le plateau* P_1 *et la pièce* 2 *sur* P_2*. Si les plateaux sont équilibrés c'est* 3 *qui est fausse, sinon le plateau le plus chargé contient la fausse pièce.*
Pour 8 *pièces, on isole les pièces* 7 *et* 8 *et on place les pièces* $1, 2, 3$ *sur le plateau* P_1 *et les pièces* $4, 5, 6$ *sur le plateau* P_2*. Si les plateaux sont équilibrés, on compare* 7 *et* 8 *avec la balance et on détermine la fausse pièce, sinon l'un des deux plateaux, disons* P_1 *est plus chargé, il contient donc la fausse pièce et le procédé utilisé pour les* 6 *pièces nous permet de trouver la fausse pièce.*

Exercice 1.10 *Des cannibales proposent à un touriste de décider lui même de son sort en faisant une déclaration : si celle-ci est vraie, il sera rôti, sinon il sera bouilli. Quelle déclaration peut faire ce touriste (malin) pour imposer une troisième solution ?*

Solution 1.10 ♠♠♠

Exercice 1.11 *Les habitants d'un village sont partagés en deux clans : ceux du clan* A *disent toujours la vérité et ceux du clan* B *mentent toujours. Un touriste passant par ce village rencontre trois habitants et souhaite savoir à quel clan appartient chacun d'eux. Il n'entend pas la réponse du premier, le deuxième répète ce qu'il a entendu, selon lui, du premier et le troisième lui indique le clan du premier et du second. Le touriste a la réponse à sa question. Pouvez-vous faire de même.*

Solution 1.11 ♠♠♠

On dit qu'une théorie est non contradictoire si $P \wedge \overline{P}$ est faux pour toute proposition P.

Exercice 1.12 *Montrer que si dans une théorie une propriété* P *est contradictoire, c'est-à-dire si* $P \wedge \overline{P}$ *est vraie, alors* $Q \wedge \overline{Q}$ *est vraie pour toute propriété* Q*.*

Solution 1.12 *Nous allons montrer que s'il existe un énoncé contradictoire* P*, alors tout énoncé* Q *est vrai, donc* $\overline{Q}$ *aussi et* $Q \wedge \overline{Q}$ *est vraie.*
On vérifie tout d'abord que $R = \overline{P} \to (P \to Q)$ *est une tautologie avec la table de vérité :*

P	Q	$\overline{P}$	$P \to Q$	$\overline{P} \to (P \to Q)$
V	V	F	V	V
V	F	F	F	V
F	V	V	V	V
F	F	V	V	V

Comme R *et* $\overline{P}$ *sont vraies,* $P \to Q$ *est vraie et* Q *est vraie puisque* P *est vraie.*

1.3 Quelques méthodes de raisonnement

En général l'énoncé d'une proposition à démontrer est formé d'une ou plusieurs hypothèses qui constituent l'assertion H et d'une ou plusieurs conclusions qui constituent l'assertion C. Il s'agit donc de montrer l'implication $H \Longrightarrow C$.

Si de plus, on peut montrer que $C \Longrightarrow H$, on dira alors que la réciproque de la proposition est vraie.

Les idées de base que l'on peut utiliser sont les suivantes.

- Une assertion peut toujours être remplacée par n'importe quelle assertion qui lui est équivalente.
- On peut effectuer une démonstration directe, c'est à dire de déduire logiquement C de H.
- L'implication étant transitive, on peut essayer de montrer que $C \Longrightarrow C'$ sachant par ailleurs que $C' \Longrightarrow H$.
- Dans le cas où une démonstration directe semble difficile, on peut essayer une démonstration par l'absurde qui consiste à étudier l'assertion $H \wedge \overline{C}$ équivalente à $\overline{H \longrightarrow C}$ et on montre qu'on aboutit à une impossibilité si cette dernière assertion est vraie (pratiquement, on suppose que la conclusion est fausse avec les hypothèses et on aboutit à une absurdité). Il en résulte alors que $\overline{H \longrightarrow C}$ est fausse, c'est à dire que $H \longrightarrow C$ est vraie, soit $H \Longrightarrow C$.
- On peut aussi essayer de montrer la contraposée $\overline{C} \Rightarrow \overline{H}$ puisque les implications $H \rightarrow C$ et $\overline{C} \rightarrow \overline{H}$ sont équivalentes.
- La démonstration par contre-exemple permet de montrer qu'une implication $H \rightarrow C$, où H est C sont des propriétés portant sur des variables x, est fausse. Pour ce faire on cherche une ou des valeurs de x pour lesquels $H(x)$ est vraie et $C(x)$ est fausse.
- La démonstration par récurrence permet de montrer qu'une propriété portant sur des entiers naturels est toujours vraie. Cette méthode de démonstration est décrite au paragraphe 1.6, où elle apparaît comme un théorème basé sur le fait que l'ensemble des entiers naturels est bien ordonné. Si on accepte l'axiome de Péano, le principe de récurrence en est une conséquence immédiate.

Exercice 1.13 *En raisonnant par l'absurde, montrer que $\sqrt{2}$ est irrationnel.*

Solution 1.13 *Supposons que $\sqrt{2} = \dfrac{p}{q}$ avec p, q entiers naturels non nuls premiers entre eux. On a alors $p^2 = 2q^2$ qui entraîne que p est pair, soit $p = 2p'$ et $q^2 = 2p'^2$ entraîne q pair, ce qui contredit p et q premiers entre eux.*

Exercice 1.14 *En raisonnant par l'absurde, montrer que $\dfrac{\ln(2)}{\ln(3)}$ est irrationnel.*

Solution 1.14 *Supposons que $\dfrac{\ln(2)}{\ln(3)} = \dfrac{p}{q}$ avec p, q entiers naturels non nuls premiers entre eux. On a alors $\ln(2^q) = \ln(3^p)$ et $2^p = 3^q$, ce qui est impossible puisque 2^p est un entier pair et 3^q est un entier impair.*

Exercice 1.15 *Soit n un entier naturel non carré, c'est-à-dire ne s'écrivant pas sous la forme $n = p^2$ avec p entier. En raisonnant par l'absurde et en utilisant le théorème de Bézout, montrer que $\sqrt{n}$ est irrationnel.*

Solution 1.15 *Si n est non carré, on a alors $n \geq 2$.*
Supposons que $\sqrt{n} = \dfrac{p}{q}$ avec p, q premiers entre eux dans $\mathbb{N}^*$. Le théorème de Bézout nous dit qu'il existe un couple (u, v) d'entiers relatifs tels que $up + vq = 1$. On a alors :

$$1 = (up + vq)^2 = u^2p^2 + 2uvpq + v^2q^2$$

avec $u^2p^2 = u^2nq^2$. L'égalité précédente s'écrit alors $qr = 1$ avec $r = u^2nq + 2uvp + v^2q$ dans $\mathbb{Z}$, ce qui implique que $q = 1$ et $\sqrt{n} = p$, en contradiction avec n non carré.

Exercice 1.16 *Sachant que tout entier supérieur ou égal à 2 admet un diviseur premier, montrer que l'ensemble $\mathcal{P}$ des nombres premiers est infini.*

Solution 1.16 *On sait déjà que $\mathcal{P}$ est non vide (il contient 2). Supposons que $\mathcal{P}$ soit fini avec :*

$$\mathcal{P} = \{p_1, \cdots, p_r\}.$$

L'entier $n = p_1 \cdots p_r + 1$ est supérieur ou égal à 2, il admet donc un diviseur premier $p_k \in \mathcal{P}$. L'entier p_k divise alors $n = p_1 \cdots p_r + 1$ et $p_1 \cdots p_r$, il divise donc la différence qui est égale à 1, ce qui est impossible. En conclusion $\mathcal{P}$ est infini.

Exercice 1.17 *Montrer que $x = \sqrt[3]{45 + 29\sqrt{2}} + \sqrt[3]{45 - 29\sqrt{2}}$ est un entier.*

Solution 1.17 *En posant $a = \sqrt[3]{45 + 29\sqrt{2}}$ et $b = \sqrt[3]{45 - 29\sqrt{2}}$, on a :*

$$\begin{cases} a^3 + b^3 = 90 \\ ab = \sqrt[3]{45^2 - 2 \cdot 29^2} = \sqrt[3]{343} = 7 \end{cases}$$

ce qui donne :

$$\begin{aligned} 90 &= (a+b)\left(a^2 - ab + b^2\right) \\ &= (a+b)\left((a+b)^2 - 3ab\right) = x\left(x^2 - 21\right) \end{aligned}$$

donc x est racine du polynôme :

$$P(X) = X\left(X^2 - 21\right) - 90$$

On regarde si ce polynôme a des racines entières. Comme $n^2 - 21$ est négatif pour $n \leq 4$, on cherche ces racines à partir de $n = 5$. On a $P(5) = -70$ et $P(6) = 0$. On a alors $P(X) = (X-6)(X^2 + 6X + 15)$ et $x = 6$, puis c'est la seule racine réelle de P.

1.4 Notions de base sur les ensembles. Quantificateurs

Nous nous contenterons d'une définition intuitive de la notion d'ensemble.

Un ensemble est une collection d'objets possédant des propriétés communes, ces objets sont les éléments de l'ensemble.

On utilisera les notations suivantes, pour les ensembles de nombres usuels :
- $\mathbb{N}$ est ensemble des entiers naturels ;
- $\mathbb{Z}$ est l'ensemble des entiers relatifs ;
- $\mathbb{Q}$ est l'ensemble des nombres rationnels
- $\mathbb{R}$ est l'ensemble des nombres réels ;
- $\mathbb{C}$ est l'ensemble des nombres complexes.

On admet l'existence d'un ensemble qui ne contient aucun élément. Cet ensemble est noté $\emptyset$ et on dit que c'est l'ensemble vide.

Nous serons souvent amenés à décrire un ensemble en précisant les propriétés que doivent vérifier tous ses éléments, ce que nous noterons de la façon suivante :

$$E = \{\text{description des propriétés des éléments de } E\}$$

(on dit que l'ensemble E est défini en compréhension).

Cette notion d'ensemble défini en compréhension peut conduire à des paradoxes liés au problème de « l'ensemble de tous les ensembles », mais à un premier niveau, on se contente de ce point de vue intuitif. Une étude approffondie de la théorie des ensembles peut mener assez loin. Le lecteur intéressé peut consulter le volume de Bourbaki sur les ensembles, ou tout autre ouvrage spécialisé.

On peut aussi décrire un ensemble en donnant la liste finie ou infinie de tous ces éléments, quand cela est possible, ce qui se note :

$$E = \{x_1, x_2, \cdots, x_n\}$$

s'il s'agit d'un ensemble fini ou :

$$E = \{x_1, x_2, \cdots, x_n, \cdots\}$$

s'il s'agit d'un ensemble infini pour lequel on peut numéroter les éléments (un tel ensemble est dit dénombrable). On dit alors que l'ensemble E est défini en extension.

Un singleton est un ensemble qui ne contient qu'un élément, soit $E = \{a\}$.

Si n, m sont deux entiers relatifs, l'ensemble des entiers relatifs compris entre n et m sera noté $\{n, \cdots, m\}$. Dans le cas où $m < n$, il ne peut y avoir d'entiers entre n et m et cet ensemble est tout simplement l'ensemble vide. Dans le cas où $n = m$, cet ensemble est le singleton $\{n\}$. Pour $n < m$, on notera aussi $\{n, n+1, \cdots, m\}$ cet ensemble.

Nous nous contentons dans un premier temps de définitions intuitives de ces notions d'ensemble fini ou dénombrable (voir les paragraphes 1.9 et 1.10 pour des définitions plus rigoureuses).

Si E est un ensemble, on notera $a \in E$ pour signifier que a est un élément de E, ce qui se lit « a appartient à E ». La négation de cette assertion est « a n'appartient pas à E » et se notera $a \notin E$.

Pour signifier qu'un ensemble F est contenu dans un ensemble E, ce qui signifie que tout élément de F est dans E, nous noterons $F \subset E$ qui se lit « F est contenu dans E ». On peut écrire de manière équivalent que $E \supset F$ pour dire que E contient F. La négation de cette assertion est notée $F \not\subset E$.

Deux ensembles E et F sont égaux si, et seulement si, ils ont les mêmes éléments, ce qui se traduit par $E \subset F$ et $F \subset E$.

On admet que si E est un ensemble, il existe un ensemble dont tous les éléments sont formés de tous les sous-ensembles (ou parties) de E. On note $\mathcal{P}(E)$ cet ensemble et on dit que c'est l'ensemble des parties de E. Ainsi $F \subset E$ est équivalent à $F \in \mathcal{P}(E)$. L'ensemble vide et E sont des éléments de $\mathcal{P}(E)$.

Par exemple pour $E = \{1, 2, 3\}$, on a :

$$\mathcal{P}(E) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$$

Pour décrire des ensembles, ou faire des raisonnements, nous utiliseront les deux quantificateurs suivants.

- Le quantificateur universel « quel que soit » ou « pour tout » noté $\forall$ utilisé pour signifier que tout élément x d'un ensemble E vérifie une propriété $P(x)$, la syntaxe étant :

$$(\forall x \in E)\ \ (P(x)). \tag{1.1}$$

- Le quantificateur existentiel « il existe » noté $\exists$ pour signifier qu'il existe au moins un élément x de E vérifiant la propriété $P(x)$, la syntaxe étant :

$$(\exists x \in E) \mid (P(x)). \tag{1.2}$$

Pour signifier qu'il existe un et un seul x dans E vérifiant la propriété $P(x)$, on utilisera la syntaxe :

$$(\exists! x \in E) \mid (P(x)).$$

La négation de l'assertion 1.1 est :

$$(\exists x \in E) \mid \left(\overline{P(x)}\right)$$

en utilisant le symbole $\mid$ qui se lit « tel que » utilisé pour traduire le fait que x est tel que la propriété $\overline{P(x)}$ est vérifiée et la négation de 1.2 est :

$$(\forall x \in E)\ \left(\overline{P(x)}\right).$$

Nous verrons qu'il n'est pas toujours facile de traduire la négation d'une assertion en utilisant les quantificateurs.

Par exemple pour traduire le fait qu'une suite $(u_n)_{n\in\mathbb{N}}$ de nombres réels est convergente vers un réel ℓ nous écrirons :

$$(\exists \ell \in \mathbb{R}) \mid (\forall \varepsilon > 0,\ \exists n_0 \in \mathbb{N} \mid \forall n \geq n_0,\ |u_n - \ell| < \varepsilon)$$

ce qui signifie qu'il existe un réel ℓ tel que quel que soit la précision $\varepsilon > 0$ que l'on choisisse l'écart entre u_n et ℓ (soit $|u_n - \ell|$) est inférieur à ε à partir d'un certain rang n_0.

La négation de cette assertion s'écrit :

$$(\forall \ell \in \mathbb{R}),\ (\exists \varepsilon > 0,\ \forall n_0 \in \mathbb{N},\ \exists n \geq n_0 \mid |u_n - \ell| \geq \varepsilon)$$

Nous étudierons plus loin les suites réelles ou complexes.

En utilisant les quantificateurs, il faudra faire attention à l'ordre d'apparition de ces derniers. Par exemple les assertions suivantes, où f est une fonction à valeurs réelles définie sur un ensemble E :

$$\forall x \in E,\ \exists M > 0 \mid\ f(x) < M$$

et

$$\exists M > 0 \mid \forall x \in E,\ f(x) < M.$$

ne sont pas équivalentes. La première assertion signifie que pour tout élément x de E il existe un réel $M > 0$ qui dépend à priori de x (il faudrait donc le noter $M(x)$) tel que $f(x) < M$ (par exemple $M(x) = f(x) + 1$ convient), alors que la seconde signifie qu'il existe un réel $M > 0$, indépendant de x dans E, tel que $f(x) < M$, ce qui n'est pas la même chose.

1.5 Les symboles $\sum$ et $\prod$

Si n est un entier naturel non nul et $x_1, x_2, \cdots, x_n$ des entiers, rationnels, réels ou complexes, on notera :

$$\sum_{k=1}^{n} x_k = x_1 + x_2 + \cdots + x_n \text{ et } \prod_{k=1}^{n} x_k = x_1 \cdot x_2 \cdot \cdots \cdot x_n$$

la somme et le produit des x_k.

Dans une telle somme ou produit l'indice est muet, c'est-à-dire que $\sum_{k=1}^{n} x_k = \sum_{i=1}^{n} x_i$ et $\prod_{k=1}^{n} x_k = \prod_{i=1}^{n} x_i$.

La manipulation d'un produit de réels strictement positifs se ramène à une somme en utilisant la fonction logarithme :

$$\ln\left(\prod_{k=1}^{n} x_k\right) = \sum_{k=1}^{n} \ln\left(x_k\right)$$

On peut également effectuer des changements d'indice. Par exemple, en posant $i = k+1$, on aura :

$$\sum_{k=1}^{n} x_k = \sum_{i=2}^{n+1} x_{i-1} = \sum_{k=2}^{n+1} x_{k-1}$$

On peut ajouter ou multiplier de telles sommes (ou produits). Par exemple, on a :

$$\sum_{k=1}^{n} x_k + \sum_{k=1}^{n} y_k = \sum_{k=1}^{n} \left(x_k + y_k\right)$$

$$\lambda \sum_{k=1}^{n} x_k = \sum_{k=1}^{n} \lambda x_k$$

$$\left(\sum_{k=1}^{n} x_k\right)\left(\sum_{k=1}^{m} y_k\right) = \left(\sum_{j=1}^{n} x_j\right)\left(\sum_{k=1}^{m} y_k\right) = \sum_{\substack{1\leq j\leq n\\ 1\leq k\leq m}} x_j y_k.$$

Pour vérifier ce résultat, on écrit que :

$$\begin{aligned} S &= \left(\sum_{k=1}^{n} x_k\right)\left(\sum_{k=1}^{m} y_k\right) \\ &= \left(x_1 + x_2 + \cdots + x_n\right)\left(\sum_{k=1}^{m} y_k\right) \\ &= x_1 \sum_{k=1}^{m} y_k + \cdots + x_n \sum_{k=1}^{m} y_k \\ &= \sum_{j=1}^{n} x_j \left(\sum_{k=1}^{m} y_k\right) = \sum_{j=1}^{n}\left(\sum_{k=1}^{m} x_j y_k\right) = \sum_{\substack{1\leq j\leq n\\ 1\leq k\leq m}} x_j y_k. \end{aligned}$$

Exercice 1.18 *Montrer que pour tout entier $n \geq 1$, on a :*

$$P_n = \prod_{k=1}^{n}\left(1 + \frac{1}{k}\right)^k = \frac{(n+1)^n}{n!}.$$

Solution 1.18 *Il revient au même de calculer $S_n = \ln\left(P_n\right)$. On a :*

$$\begin{aligned} S_n &= \ln\left(\prod_{k=1}^{n}\left(\frac{k+1}{k}\right)^k\right) = \sum_{k=1}^{n}\left(k\ln\left(k+1\right) - k\ln\left(k\right)\right) \\ &= \sum_{k=1}^{n} k\ln\left(k+1\right) - \sum_{k=1}^{n} k\ln\left(k\right) \end{aligned}$$

et le changement d'indice $j = k + 1$ dans la première somme donne :

$$\begin{aligned} S_n &= \sum_{j=2}^{n+1} (j-1)\ln(j) - \sum_{k=1}^{n} k\ln(k) \\ &= \sum_{j=2}^{n+1} j\ln(j) - \sum_{j=2}^{n+1} \ln(j) - \sum_{k=1}^{n} k\ln(k) \\ &= \sum_{k=2}^{n+1} k\ln(k) - \sum_{k=2}^{n+1} \ln(k) - \sum_{k=1}^{n} k\ln(k) \\ &= (n+1)\ln(n+1) - \sum_{k=2}^{n+1} \ln(k) \end{aligned}$$

(on a utilisé le fait que l'indice est muet dans une somme).
On a donc en définitive :

$$\begin{aligned} S_n = \ln(P_n) &= \ln\left((n+1)^{n+1}\right) - \sum_{k=2}^{n+1} \ln(k) \\ &= \ln\left((n+1)^{n+1}\right) - \ln\left(\prod_{k=2}^{n} k\right) = \ln\left((n+1)^{n+1}\right) - \ln(n!) \\ &= \ln\left(\frac{(n+1)^{n+1}}{n!}\right) \end{aligned}$$

et $P_n = \dfrac{(n+1)^n}{n!}$.
Une autre solution consiste à effectuer directement un changement d'indice dans le produit. Soit :

$$\begin{aligned} P &= \prod_{k=1}^{n} \left(\frac{k+1}{k}\right)^k = \frac{\prod_{k=1}^{n} (k+1)^k}{\prod_{k=1}^{n} k^k} \\ &= \frac{\prod_{j=2}^{n+1} j^{j-1}}{\prod_{k=1}^{n} k^k} = \frac{\prod_{k=2}^{n+1} k^{k-1}}{\prod_{k=1}^{n} k^k} = \frac{2 \cdot 3^2 \cdot 4^3 \cdot \cdots \cdot n^{n-1} \cdot (n+1)^n}{2^2 \cdot 3^3 \cdot 4^4 \cdot \cdots \cdot (n-1)^{n-1} \cdot n^n} \\ &= \frac{(n+1)^n}{2 \cdot 3 \cdot 4 \cdot \cdots \cdot (n-1) \cdot n} = \frac{(n+1)^n}{n!}. \end{aligned}$$

1.6 Les théorèmes de récurrence

On désigne par $\mathbb{N}$ l'ensemble des entiers naturels, soit :

$$\mathbb{N} = \{0, 1, 2, \cdots, n, \cdots\}.$$

La construction de cet ensemble avec les opérations usuelles d'addition et de multiplication est admise.

On note $\mathbb{N}^*$ l'ensemble $\mathbb{N}$ privé de 0.

Notre point de départ est l'axiome du bon ordre suivant : toute partie non vide de $\mathbb{N}$ admet un plus petit élément, ce qui signifie que si A est une partie non vide de $\mathbb{N}$, il existe alors un entier m tel que :

$$\begin{cases} m \in \mathbb{N}, \\ \forall n \in A,\ m \leq n. \end{cases}$$

Exercice 1.19 *On peut montrer que $\sqrt{3}$ est irrationnel en utilisant seulement le fait que $\mathbb{N}$ est bien ordonné. Pour ce faire on raisonne par l'absurde en supposant qu'il existe deux entiers strictement positifs a et b tels que $\sqrt{3} = \dfrac{a}{b}$.*
On introduit l'ensemble :

$$A = \left\{ q \in \mathbb{N} - \{0\} \mid \exists p \in \mathbb{N} \mid \sqrt{3} = \frac{p}{q} \right\}.$$

1. *Montrer que A a un plus petit élément q_1. On a donc $\sqrt{3} = \dfrac{p_1}{q_1}$ avec $p_1 \in \mathbb{N}$.*
2. *Montrer que $\sqrt{3} = \dfrac{3q_1 - p_1}{p_1 - q_1}$ et conclure.*

Solution 1.19

1. *Si on suppose $\sqrt{3}$ rationnel alors l'ensemble A est non vide dans $\mathbb{N}$ et en conséquence il admet un plus petit élément q_1. Comme $q_1 \in A$, il existe un entier $p_1 \geq 1$ tel que $\sqrt{3} = \dfrac{p_1}{q_1}$.*
2. *On a :*
$$\sqrt{3} + 1 = \frac{2}{\sqrt{3} - 1} = \frac{2q_1}{p_1 - q_1}$$
et :
$$\sqrt{3} = \frac{2q_1}{p_1 - q_1} - 1 = \frac{3q_1 - p_1}{p_1 - q_1} = \frac{p_2}{q_2}$$
où on a posé :
$$\begin{cases} p_2 = 3q_1 - p, \\ q_2 = p_1 - q_1. \end{cases}$$
Comme $1 < \sqrt{3} = \dfrac{p_1}{q_1} < 2$ (puisque $1 < 3 = \sqrt{3}^2 < 4$) on a $p_1 < 2q_1$, donc $p_2 > 0$ et $q_2 < q_1$. On a donc $\sqrt{3} = \dfrac{p_2}{q_2}$ avec $q_2 \in A$ et $q_2 < q_1$, ce qui contredit le fait que q_1 est le plus petit élément de A. On peut donc conclure à l'irrationalité de $\sqrt{3}$.

En fait l'exercice précédent peut se généraliser comme suit.

Exercice 1.20 *Soit n un entier naturel non carré (i. e. il n'existe pas d'entier p tel que $n = p^2$). On se propose, comme dans l'exercice précédent, de montrer que $\sqrt{n}$ est irrationnel en utilisant seulement le fait que $\mathbb{N}$ est bien ordonné.*
Pour ce faire on raisonne par l'absurde en supposant qu'il existe deux entiers strictement positifs a et b tels que $\sqrt{n} = \dfrac{a}{b}$.
On introduit l'ensemble :

$$A = \left\{ q \in \mathbb{N} - \{0\} \mid \exists p \in \mathbb{N} \mid \sqrt{n} = \frac{p}{q} \right\}.$$

1. *Montrer que A a un plus petit élément q_1. On a donc $\sqrt{n} = \dfrac{p_1}{q_1}$ avec $p_1 \in \mathbb{N}$.*
2. *Montrer qu'il existe un entier $m_1 \in [1, \sqrt{n}[$ tel que $\sqrt{n} = \dfrac{nq_1 - m_1 p_1}{p_1 - m_1 q_1}$ et conclure.*

Solution 1.20

1. *Si on suppose $\sqrt{n}$ rationnel alors l'ensemble A est non vide dans $\mathbb{N}$ et en conséquence il admet un plus petit élément q_1. Comme $q_1 \in A$, il existe un entier $p_1 \geq 1$ tel que $\sqrt{n} = \dfrac{p_1}{q_1}$.*
2. *L'ensemble :*
$$B = \left\{ m \in \mathbb{N}^* \mid m^2 < n \right\}$$
étant non vide dans $\mathbb{N}^*$ (1 est dans B car n non carré dans $\mathbb{N}$ entraîne $n \geq 2$) et majoré par n admet un plus grand élément $m_1 \in \mathbb{N} \cap [1, \sqrt{n}[$ et on a :
$$m_1^2 < n < (m_1 + 1)^2$$
(m_1 est en fait la partie entière de $\sqrt{n}$). On a alors :
$$\sqrt{n} + m_1 = \frac{n - m_1^2}{\sqrt{n} - m_1} = \frac{(n - m_1^2)\, q_1}{p_1 - m_1 q_1}$$
et :
$$\sqrt{n} = \frac{(n - m_1^2)\, q_1}{p_1 - m_1 q_1} - m_1 = \frac{nq_1 - m_1 p_1}{p_1 - m_1 q_1} = \frac{p_2}{q_2}$$
où on a posé :
$$\begin{cases} p_2 = nq_1 - m_1 p_1, \\ q_2 = p_1 - m_1 q_1. \end{cases}$$
En tenant compte de $\sqrt{n} = \dfrac{p_1}{q_1}$, on a :
$$p_2 = p_1 \left(n \frac{q_1}{p_1} - m_1 \right) = p_1 \left(\sqrt{n} - m_1 \right) > 0,$$
soit $p_2 \geq 1$ et $q_2 \geq 1$ puisque $\sqrt{n} = \dfrac{p_2}{q_2} > 0$. Ensuite de :
$$\sqrt{n} = \frac{p_1}{q_1} < m_1 + 1,$$
on déduit que :
$$q_2 = p_1 - m_1 q_1 < q_1.$$
On a donc $q_2 \in A$ et $q_2 < q_1$, ce qui contredit le fait que q_1 est le plus petit élément de A. On peut donc conclure à l'irrationalité de $\sqrt{n}$.

De l'axiome du bon ordre, on déduit les deux théorèmes fondamentaux qui suivent. Le premier résultat est souvent appelé théorème de récurrence faible et le second théorème de récurrence forte.

Théorème 1.2 *Soient $n_0 \in \mathbb{N}$ et $\mathcal{P}(n)$ une propriété portant sur les entiers $n \geq n_0$. La propriété $\mathcal{P}(n)$ est vraie pour tout entier $n \geq n_0$ si et seulement si :*

(i) $\mathcal{P}(n_0)$ est vraie ;

(*ii*) *pour tout* $n \geq n_0$ *si* $\mathcal{P}(n)$ *est vrai alors* $\mathcal{P}(n+1)$ *est vraie.*

Démonstration. La condition nécessaire est évidente.

En supposant les conditions (*i*) et (*ii*) vérifiées, on note A l'ensemble des entiers $n \geq n_0$ pour lesquels $\mathcal{P}(n)$ est faux. Si A est non vide il admet alors un plus petit élément $n > n_0$ (puisque $\mathcal{P}(n_0)$ est vraie). Mais alors $\mathcal{P}(n-1)$ est vraie ce qui implique, d'après (*ii*), que $\mathcal{P}(n)$ est vraie, soit une contradiction. En définitive A est vide et la propriété est vraie pour tout entier $n \geq n_0$. ■

Théorème 1.3 *Soient* $n_0 \in \mathbb{N}$ *et* $\mathcal{P}(n)$ *une propriété portant sur les entiers* $n \geq n_0$. *La propriété* $\mathcal{P}(n)$ *est vraie pour tout entier* $n \geq n_0$ *si et seulement si :*

(*i*) $\mathcal{P}(n_0)$ *est vraie ;*

(*ii*) *pour tout* $n \geq n_0$ *si* $\mathcal{P}(k)$ *est vrai pour tout entier* k *compris entre* n_0 *et* n, *alors* $\mathcal{P}(n+1)$ *est vraie.*

Démonstration. La condition nécessaire est évidente.

En supposant les conditions (*i*) et (*ii*) vérifiées, on note A l'ensemble des entiers $n \geq n_0$ pour lesquels $\mathcal{P}(n)$ est faux. Si A est non vide il admet alors un plus petit élément $n > n_0$ et $\mathcal{P}(k)$ est vraie pour tout k compris entre n_0 et $n-1$, ce qui implique que $\mathcal{P}(n)$ est vraie, soit une contradiction. En définitive A est vide et la propriété est vraie pour tout entier $n \geq n_0$. ■

Exercice 1.21 *Montrer que* $2^n > n^2$ *pour tout entier* $n \geq 5$.

Solution 1.21 *Pour* $n = 5$, *on a* $2^5 = 32 > 5^2 = 25$.
Supposant le résultat acquis au rang $n \geq 5$, *on a :*

$$2^{n+1} = 22^n > 2n^2 > (n+1)^2$$

puisque :

$$2n^2 - (n+1)^2 = n^2 - 2n - 1 = (n-1)^2 - 2 > 0$$

pour $n \geq 5$. *Le résultat est donc vrai au rang* $n+1$ *et il vrai pour tout* $n \geq 5$.

Exercice 1.22 *Montrer que si* φ *est une fonction strictement croissante de* $\mathbb{N}$ *dans* $\mathbb{N}$, *on a alors* $\varphi(n) \geq n$ *pour tout* n.

Solution 1.22 *Comme* φ *est une fonction de* $\mathbb{N}$ *dans* $\mathbb{N}$, $\varphi(0)$ *est un entier naturel et donc* $\varphi(0) \geq 0$. *Supposant le résultat acquis pour* $n \geq 0$, *sachant que* φ *est strictement croissante, on a* $\varphi(n+1) > \varphi(n) \geq n$, *donc* $\varphi(n+1) > n$, *ce qui équivaut à* $\varphi(n+1) \geq n+1$ *puisque* $\varphi(n+1)$ *est un entier.*

Le théorème de récurrence faible peut être utilisé pour montrer quelques identités classiques comme celles qui apparaissent avec les exercices qui suivent.

Exercice 1.23 *Montrer par récurrence que pour tout entier naturel non nul* n, *on a :*

$$U_n = \sum_{k=1}^{n} k = \frac{n(n+1)}{2},$$

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

$$W_n = \sum_{k=1}^{n} k^3 = \left(\frac{n(n+1)}{2}\right)^2 = U_n^2.$$

Solution 1.23 *Pour $n = 1$ c'est clair.*
En supposant les résultats acquis pour $n \geq 1$, on a :

$$U_{n+1} = \frac{n\,(n+1)}{2} + n + 1 = \frac{n^2 + 3n + 2}{2} = \frac{(n+1)\,(n+2)}{2}$$

$$\begin{aligned} V_{n+1} &= \frac{n\,(n+1)\,(2n+1)}{6} + (n+1)^2 = \frac{(n+1)\,(2n^2+7n+6)}{6} \\ &= \frac{(n+1)\,(n+2)\,(2n+3)}{6} \end{aligned}$$

$$\begin{aligned} W_{n+1} &= \left(\frac{n\,(n+1)}{2}\right)^2 + (n+1)^3 = \frac{(n+1)^2\,(n^2+4n+4)}{4} \\ &= \left(\frac{(n+1)\,(n+2)}{2}\right)^2 \end{aligned}$$

On a aussi :

$$\begin{aligned} U_n &= 1 + 2 + \cdots + (n-1) + n \\ &= n + (n-1) + \cdots + 2 + 1 \end{aligned}$$

et en additionnant terme à terme on obtient :

$$2U_n = n\,(n+1)\,.$$

Le calcul de U_n peut aussi se faire en passant par V_{n+1} et en utilisant l'identité :

$$(k+1)^2 = k^2 + 2k + 1$$

Précisément, en effectuant le changement d'indice $k = j + 1$, on a :

$$V_{n+1} = \sum_{k=1}^{n+1} k^2 = \sum_{j=0}^{n} (j+1)^2 = \sum_{j=0}^{n} j^2 + 2\sum_{j=0}^{n} j + \sum_{j=0}^{n} 1$$

soit :

$$V_{n+1} = V_n + 2U_n + n + 1$$

et :

$$2U_n = V_{n+1} - V_n - (n+1) = (n+1)^2 - (n+1) = n\,(n+1)$$

ce qui donne bien $U_n = \dfrac{n\,(n+1)}{2}$.
De même, le calcul de V_n peut aussi se faire en passant par W_{n+1} et en utilisant l'identité :

$$(k+1)^3 = k^3 + 3k^2 + 3k + 1$$

Précisément, en effectuant le changement d'indice $k = j + 1$, on a :

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

soit :

$$W_{n+1} = W_n + 3V_n + 3U_n + n + 1$$

et :

$$\begin{aligned} 3V_n &= W_{n+1} - W_n - 3U_n - (n+1) = (n+1)^3 - 3\frac{n\,(n+1)}{2} - (n+1) \\ &= \frac{n\,(n+1)\,(2n+1)}{2} \end{aligned}$$

ce qui donne bien $V_n = \dfrac{n\,(n+1)\,(2n+1)}{6}$.
Ce procédé peut en fait se généraliser.

Exercice 1.24 *Calculer, pour tout entier naturel* n*, la somme :*

$$I_n = 1 + 3 + 5 + \cdots (2n-1) + (2n+1)\,.$$

Solution 1.24 *On a :*

$$\begin{aligned} I_n = \sum_{k=0}^{n} (2k+1) = 2\sum_{k=0}^{n} k + \sum_{k=0}^{n} 1 &= n\,(n+1) + (n+1) \\ &= (n+1)^2\,. \end{aligned}$$

Exercice 1.25 *On appelle nombres triangulaires les sommes* $U_n = \sum_{k=1}^{n} k$ *et nombres pyramidaux les sommes* $P_n = \sum_{k=1}^{n} U_k$*. Montrer que :*

$$P_n = \frac{n\,(n+1)\,(n+2)}{6}.$$

Solution 1.25 *Pour* $n = 1$ *on a* $P_1 = U_1 = 1$ *et le résultat est acquis est vrai pour* $n = 1$*. En le supposant acquis pour* $n \geq 1$*, on a :*

$$\begin{aligned} P_{n+1} &= \frac{n\,(n+1)\,(n+2)}{6} + \frac{(n+1)\,(n+2)}{2} \\ = \frac{(n+1)\,(n+2)}{2}\left(\frac{n}{3}+1\right) &= \frac{(n+1)\,(n+2)\,(n+3)}{6}. \end{aligned}$$

Exercice 1.26 *Montrer par récurrence, que pour tout entier naturel* n *et tout nombre complexe* λ *différent de* 1*, on a :*

$$\sum_{k=0}^{n} \lambda^k = \frac{\lambda^{n+1} - 1}{\lambda - 1}.$$

Solution 1.26 *Pour* $n = 0$*, c'est clair. Si c'est vrai pour* $n \geq 0$*, alors :*

$$\sum_{k=0}^{n+1} \lambda^k = \frac{\lambda^{n+1} - 1}{\lambda - 1} + \lambda^{n+1} = \frac{\lambda^{n+2} - 1}{\lambda - 1}.$$

Plus généralement, on a l'identité (dite remarquable) suivante.

Exercice 1.27 *Montrer que pour tout entier naturel n et tous nombres complexes a et b on a :*

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

Solution 1.27 *Pour $n = 0$, c'est évident. En supposant le résultat acquis au rang $n \geq 0$, on a :*

$$\begin{aligned} b^{n+2} - a^{n+2} &= \left(b^{n+1} - a^{n+1}\right) b + ba^{n+1} - a^{n+2} \\ &= (b-a)\sum_{k=0}^{n} a^k b^{n+1-k} + (b-a)\, a^{n+1} \\ &= (b-a)\left(b^{n+1} + ab^n + \cdots + a^{n-1}b^2 + a^n b\right) + (b-a)\, a^{n+1} \\ &= (b-a)\sum_{k=0}^{n+1} a^k b^{n+1-k}. \end{aligned}$$

Le résultat est donc vrai pour tout $n \geq 0$.

Le théorème de récurrence nous permet de définir la fonction factorielle sur l'ensemble des entiers naturels de la façon suivante :

$$\begin{cases} 0! = 1 \\ \forall n \in \mathbb{N},\ (n+1)! = (n+1)\, n! \end{cases}$$

De manière plus générale, c'est le théorème de récurrence qui nous assure de l'existence et de l'unicité d'une suite (réelle ou complexe) définie par :

$$\begin{cases} u_0 \text{ est un scalaire donné,} \\ \forall n \in \mathbb{N},\ u_{n+1} = f\left(u_n\right) \end{cases}$$

où f est une fonction définie sur un ensemble I et à valeurs dans le même ensemble I. Une telle suite est dite définie par une relation de récurrence (d'ordre 1).

Une telle suite peut aussi se définir en donnant les premières valeurs $u_0, u_1, \cdots, u_p$ et une relation $u_{n+1} = f\left(u_n, \cdots, u_{n-(p-1)}\right)$ pour $n \geq p-1$. Une telle suite est dite définie par une relation de récurrence d'ordre p.

Exercice 1.28 *Montrer que pour tout entier naturel n et tous nombres complexes a et b on a :*

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

où $C_n^k = \dfrac{n!}{k!\,(n-k)!}$ pour k compris entre 0 et n avec la convention $0! = 1$ (formule du binôme de Newton).

Solution 1.28 *Pour $n = 0$ et $n = 1$, c'est évident. En supposant le résultat acquis au rang $n \geq 1$, on a :*

$$\begin{aligned}(a+b)^{n+1} &= (a+b)^n (a+b) = \left(\sum_{k=0}^{n} C_n^k a^{n-k} b^k\right)(a+b) \\ &= \sum_{k=0}^{n} C_n^k a^{n-(k-1)} b^k + \sum_{k=0}^{n} C_n^k a^{n-k} b^{k+1} \\ &= \sum_{k=0}^{n} C_n^k a^{n-(k-1)} b^k + \sum_{k=1}^{n+1} C_n^{k-1} a^{n-(k-1)} b^k \\ &= a^{n+1} + \sum_{k=1}^{n} \left(C_n^k + C_n^{k-1}\right) a^{n+1-k} b^k + b^{n+1}\end{aligned}$$

et tenant compte de $C_n^k + C_n^{k-1} = C_{n+1}^k$ (triangle de Pascal), cela s'écrit :

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

Le résultat est donc vrai pour tout $n \geq 0$.

Les coefficients C_n^k se notent aussi $\begin{pmatrix} n \\ k \end{pmatrix}$.

On peut remarquer que, pour k fixé :

$$\begin{pmatrix} n \\ k \end{pmatrix} = \frac{n!}{k!\,(n-k)!} = \frac{n\,(n-1)\cdots(n-k+1)}{k!}$$

est un un polynôme en n de degré k, ce qui permet d'étendre cette définition à $\mathbb{R}$ ou même $\mathbb{C}$.

Comme $(a+b)^n$, on a aussi :

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

Exercice 1.29 *Montrer par récurrence, que pour tout entier naturel non nul n et tout nombre complexe λ différent de 1, on a :*

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

Solution 1.29 *Pour $n = 1$, c'est clair. Si c'est vrai pour $n \geq 1$, alors :*

$$\begin{aligned}\sum_{k=1}^{n+1} k\lambda^k &= n\frac{\lambda^{n+1}}{\lambda-1} + \lambda\frac{1-\lambda^n}{(\lambda-1)^2} + (n+1)\lambda^{n+1} \\ &= n\frac{\lambda^{n+1}}{\lambda-1} + \lambda\frac{1-\lambda^n}{(\lambda-1)^2} + n\lambda^{n+1}\frac{\lambda-1}{\lambda-1} + \lambda^{n+1}\frac{(\lambda-1)^2}{(\lambda-1)^2} \\ &= \frac{n\lambda^{n+2}}{\lambda-1} + \frac{\lambda}{(\lambda-1)^2}\left(1 + \lambda^{n+1}(\lambda-2)\right) \\ &= \frac{(n+1)\lambda^{n+2}}{\lambda-1} + \frac{\lambda}{(\lambda-1)^2}\left(1-\lambda^{n+1}\right).\end{aligned}$$

Exercice 1.30 *Montrer que pour tout entier $n \geq 1$, on a :*

$$\sqrt{2+\sqrt{2+\sqrt{2+\sqrt{\cdots+\sqrt{2}}}}} = 2\cos\left(\frac{\pi}{2^{n+1}}\right)$$

(le nombre 2 apparaissant n fois sous la racine).

Solution 1.30 *Notons* $x_n = \sqrt{2+\sqrt{2+\sqrt{2+\sqrt{\cdots+\sqrt{2}}}}}$. *Pour $n = 1$, on a :*

$$x_1 = \sqrt{2} = 2\cos\left(\frac{\pi}{4}\right).$$

Supposant le résultat acquis au rang $n \geq 1$, on a :

$$x_{n+1}^2 = 2 + x_n = 2 + 2\cos\left(\frac{\pi}{2^{n+1}}\right)$$

et utilisant la formule $\cos(2\theta) = 2\cos^2(\theta) - 1$, *il vient :*

$$\cos\left(\frac{\pi}{2^{n+1}}\right) = \cos\left(2\frac{\pi}{2^{n+2}}\right) = 2\cos^2\left(\frac{\pi}{2^{n+2}}\right) - 1$$

on a :

$$x_{n+1}^2 = 4\cos^2\left(\frac{\pi}{2^{n+2}}\right).$$

Comme x_{n+1} est positif, on en déduit que $x_{n+1} = 2\cos\left(\frac{\pi}{2^{n+2}}\right)$.

Exercice 1.31 *Soit $x_1, x_2, \cdots, x_n$ des réels dans $[0,1]$. Montrer par récurrence que* $\prod_{k=1}^{n}(1-x_k) \geq 1 - \sum_{k=1}^{n} x_k$.

Solution 1.31 *Notons :*

$$u_n = \prod_{k=1}^{n}(1-x_k) \;\; \textit{et} \;\; v_n = 1 - \sum_{k=1}^{n} x_k.$$

Pour $n = 1$, on a $u_1 = v_1$.
Supposant le résultat acquis au rang $n \geq 1$ et tenant compte de $1 - x_{n+1} \geq 0$, on a :

$$u_{n+1} = u_n(1-x_{n+1}) \geq \left(1 - \sum_{k=1}^{n} x_k\right)(1-x_{n+1})$$

$$\geq 1 - \sum_{k=1}^{n} x_k - x_{n+1} + x_{n+1}\sum_{k=1}^{n} x_k \geq 1 - \sum_{k=1}^{n+1} x_k = v_{n+1}.$$

puisque tous les x_k sont positifs.

Les théorèmes de récurrence peuvent aussi être utilisés pour montrer les résultats fondamentaux d'arithmétique suivants.

Exercice 1.32 *Soit a, b deux entiers naturels avec b non nul. Montrer qu'il existe un unique couple d'entiers (q, r) tel que :*

$$\begin{cases} a = bq + r, \\ 0 \leq r \leq b - 1. \end{cases}$$

Solution 1.32 *On montre tout d'abord l'existence du couple (q, r) par récurrence sur l'entier $a \geq 0$.*
Pour $a = 0$, le couple $(q, r) = (0, 0)$ convient.
Supposant le résultat acquis pour tous les entiers a' compris entre 0 et $a - 1$, où a est un entier naturel non nul, on distingue deux cas. Si a est compris entre 1 et $b - 1$, le couple $(q, r) = (0, a)$ convient, sinon on a $a \geq b$, donc $0 \leq a - b \leq a - 1$ et l'hypothèse de récurrence nous assure de l'existence d'un couple d'entiers (q, r) tels que $a - b = bq + r$ et $0 \leq r \leq b - 1$, ce qui nous fournit le couple d'entiers $(q', r) = (q + 1, r)$.
L'unicité se montre facilement par l'absurde.

Exercice 1.33 *Soit n un entier naturel supérieur ou égal à 2. Montrer, par récurrence, que soit n est premier, soit n admet un diviseur premier $p \leq \sqrt{n}$.*

Solution 1.33 *Pour $n = 2$ et $n = 3$, le résultat est évident (n est premier).*
Supposons le acquis pour tous les entiers strictement inférieurs à $n \geq 3$. Si n est premier, c'est terminé, sinon il existe deux entiers a et b compris entre 2 et $n - 1$ tels que $n = ab$ et comme ces deux entiers jouent des rôles symétriques, on peut supposer que $a \leq b$. L'hypothèse de récurrence nous dit que soit a est premier et c'est alors un diviseur premier de n tel que $a^2 \leq ab \leq n$, soit a admet un diviseur premier $p \leq \sqrt{a}$ et p divise aussi n avec $p \leq \sqrt{n}$.

Exercice 1.34 *Montrer que tout entier naturel n supérieur ou égal à 2 se décompose de manière unique sous la forme :*

$$n = p_1^{\alpha_1} \cdots p_r^{\alpha r},$$

où les p_k sont des nombres premiers vérifiant :

$$2 \leq p_1 < p_2 < \cdots < p_r$$

et les α_k sont des entiers naturels non nuls (décomposition en nombres premiers).

Solution 1.34 *On démontre tout d'abord l'existence d'une telle décomposition par récurrence sur $n \geq 2$.*
Pour $n = 2$, on a déjà la décomposition.
Supposons que, pour $n \geq 2$, tout entier k compris entre 2 et n admet une telle décomposition. Si $n + 1$ est premier, on a déjà la décomposition, sinon on écrit $n + 1 = ab$ avec a et b compris entre 2 et n et il suffit d'utiliser l'hypothèse de récurrence pour a et b.
L'unicité d'une telle décomposition se montre également par récurrence sur $n \geq 2$. Le résultat est évident pour $n = 2$. Supposons le acquis pour tout entier k compris entre 2 et $n \geq 2$. Si $n + 1$ a deux décompositions :

$$n + 1 = p_1^{\alpha_1} \cdots p_r^{\alpha r} = q_1^{\beta_1} \cdots q_s^{\beta r},$$

où les p_j [resp. q_i] sont premiers deux à deux distincts et les α_j [resp. β_i] entiers naturels non nuls. L'entier p_1 est premier et divise le produit $q_1^{\beta_1} \cdots q_s^{\beta r}$, il divise donc nécessairement l'un des q_k. L'entier q_k étant également premier la seule possibilité est $p_1 = q_k$. En simplifiant par p_1 on se ramène à la décomposition d'un entier inférieur ou égal à n et il suffit d'utiliser l'hypothèse de récurrence pour conclure.

Exercice 1.35 *Pour tout entier naturel n supérieur ou égal à 2, on note* $H_n = \sum_{k=1}^{n} \frac{1}{k}$.

1. *Soit p un entier naturel non nul. Montrer que* $H_{2p} = \frac{1}{2}H_p + \frac{a}{2b+1}$ *où a, b sont des entiers naturels avec a non nul.*
2. *Montrer par récurrence que pour tout entier naturel non nul H_n est le quotient d'un entier impair par un entier pair et qu'en conséquence ce n'est pas un entier.*

Solution 1.35

1. *On a :*
$$H_{2p} = \sum_{k=1}^{p} \frac{1}{2k} + \sum_{k=0}^{p-1} \frac{1}{2k+1} = \frac{1}{2}H_p + \frac{N}{D}$$
avec $D = \text{ppcm}\,(1, 3, \cdots, 2p-1)$ *qui est impair et N entier naturel non nul.*
2. *On a* $H_2 = \frac{3}{2} \notin \mathbb{N}$. *Supposons le résultat acquis au rang $n \geq 2$. Si $n = 2p$, on a alors :*
$$\begin{aligned} H_{n+1} = H_n + \frac{1}{2p+1} &= \frac{2a+1}{2b} + \frac{1}{2p+1} \\ &= \frac{(2a+1)(2p+1)+2b}{2b(2p+1)} = \frac{2a'+1}{2b'} \end{aligned}$$
avec $a' = a + b + p + 2ap$ *et* $b' = b(2p+1)$. *Si $n = 2p+1$, on a alors :*
$$\begin{aligned} H_{n+1} = H_{2(p+1)} &= \frac{c}{2d+1} + \frac{1}{2}H_{p+1} \\ &= \frac{c}{2d+1} + \frac{1}{2}\frac{2a+1}{2b} = \frac{4bc+(2d+1)(2a+1)}{4b(2d+1)} = \frac{2a'+1}{2b'} \end{aligned}$$
avec $a' = a + d + 2ad + 2bc$ *et* $b' = 2b(2d+1)$.
Dans tous les cas, H_n est le quotient d'un entier impair par un entier pair et en conséquence, ce n'est pas un entier.

1.7 L'algèbre des parties d'un ensemble

Nous allons définir sur l'ensemble $\mathcal{P}(E)$ des parties d'un ensemble E des opérations qui vont traduire les idées intuitives de partie complémentaire, d'intersection et de réunion.

L'ensemble E étant donné et $A, B, C, \cdots$ désignant des parties de E (donc des éléments de $\mathcal{P}(E)$), on définit les ensembles suivant.

– le complémentaire de A dans E est l'ensemble noté $C_E A$, ou $E \setminus A$ (lire E moins A) ou $\overline{A}$ des éléments de E qui ne sont pas dans A, ce qui peut se traduire par :
$$\left(x \in \overline{A}\right) \Leftrightarrow ((x \in E) \wedge (x \notin A))$$
ou encore par :
$$\overline{A} = \{x \in E \mid x \notin A\}$$

– L'intersection de A et B, notée $A \cap B$, est l'ensemble des éléments de E qui sont dans A et dans B, soit :
$$(x \in A \cap B) \Leftrightarrow ((x \in A) \wedge (x \in B))$$

ou encore :

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

Si $A \cap B = \emptyset$, on dit alors que A et B sont disjointes.
Par exemple A et $\overline{A}$ sont disjointes.

– La réunion de A et B, notée $A \cup B$, est l'ensemble des éléments de E qui sont soit dans A, soit dans B (éventuellement dans A et B) soit :

$$(x \in A \cup B) \Leftrightarrow ((x \in A) \vee (x \in B))$$

ou encore :

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

– La différence de A et B, notée $A \setminus B$, est l'ensemble des éléments de E qui sont dans A et qui ne sont pas dans B, soit :

$$(x \in A \setminus B) \Leftrightarrow ((x \in A) \wedge (x \notin B))$$

ou encore :

$$A \setminus B = \{x \in A \mid x \notin B\}$$

Ainsi $\overline{A} = E \setminus A$.

– La différence symétrique de A et B, notée $A\Delta B$, est l'ensemble des éléments de E qui sont soit dans A et pas dans B soit dans B et pas dans A (c'est-à-dire dans A ou exclusif dans B), soit :

$$(x \in A\Delta B) \Leftrightarrow ((x \in A) \wedge (x \notin B)) \vee ((x \in B) \wedge (x \notin A))$$

Par exemple, on a $A\Delta\emptyset = A$, $A\Delta E = \overline{A}$.

Ces opérateurs de complémentarité, intersection, réunion et différence symétrique sont décrits à l'aide des connecteurs logiques non de négation, $\wedge$ de conjonction, $\vee$ de disjonction et W de disjonction exclusive.

Avec le théorème qui suit, on résume les résultats essentiels relatifs à ces opérateurs ensemblistes.

Théorème 1.4 *Soient E un ensemble et $A, B, C, \cdots$ des sous-ensembles de E. On a :*

1. *commutativité :*

$$A \cap B = B \cap A$$
$$A \cup B = B \cup A$$
$$A\Delta B = B\Delta A$$

2. *associativité :*

$$A \cap (B \cap C) = (A \cap B) \cap C$$
$$A \cup (B \cup C) = (A \cup B) \cup C$$
$$A\Delta (B\Delta C) = (A\Delta B) \Delta C$$

3. *distributivité :*

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$
$$A \cap (B\Delta C) = (A \cap B) \Delta (A \cap C)$$
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

4. différence symétrique :

$$A\Delta A = \emptyset$$
$$A\Delta B = (A \setminus B) \cup (B \setminus A)$$
$$A\Delta B = \left(A \cap \overline{B}\right) \cup \left(B \cap \overline{A}\right)$$
$$A\Delta B = (A \cup B) \setminus (A \cap B)$$

5. négations :

$$\overline{\overline{A}} = A$$
$$(A \subset B) \Leftrightarrow \left(\overline{B} \subset \overline{A}\right)$$
$$\overline{A \cap B} = \overline{A} \cup \overline{B}$$
$$\overline{A \cup B} = \overline{A} \cap \overline{B}$$

Démonstration. Laissée au lecteur. ■

On notera l'analogie entre ce théorème et le théorème 1.1 sur les règles de calculs avec les connecteurs logiques.

Toutes ces égalités entre ensembles se visualisent bien en utilisant les diagrammes d'Euler-Venn.

La propriété d'associativité de l'intersection et de la réunion nous permet d'écrire $A \cap B \cap C$ et $A \cup B \cup C$ l'intersection et la réunion de trois ensembles sans se soucier de parenthèses. De manière plus générale, grâce à cette associativité, on peut définir l'intersection ou la réunion de n sous-ensembles $A_1, A_2, \cdots, A_n$ de E par :

$$(x \in A_1 \cap A_2 \cap \cdots \cap A_n) \Leftrightarrow ((x \in A_1) \wedge (x \in A_2) \wedge \cdots \wedge (x \in A_n))$$

et :

$$(x \in A_1 \cup A_2 \cup \cdots \cup A_n) \Leftrightarrow ((x \in A_1) \vee (x \in A_2) \vee \cdots \vee (x \in A_n))$$

De façon condensée, on écrira $(A_k)_{1 \leq k \leq n}$ une telle famille de sous ensembles de E et :

$$\bigcap_{k=1}^{n} A_k = A_1 \cap A_2 \cap \cdots \cap A_n$$

l'intersection et :

$$\bigcup_{k=1}^{n} A_k = A_1 \cup A_2 \cup \cdots \cup A_n$$

la réunion.

On vérifie facilement que pour tout entier j compris entre 1 et n, on a :

$$\bigcap_{k=1}^{n} A_k \subset A_j \subset \bigcup_{k=1}^{n} A_k.$$

Définition 1.1 *On dit qu'une famille $(A_k)_{1 \leq k \leq n}$ de parties d'un ensemble E forme une partition de E si les A_k sont deux à deux disjoints, c'est-à-dire que $A_k \cap A_j = \emptyset$ pour $1 \leq k \neq j \leq n$ de réunion égale à E, soit $\bigcup_{k=1}^{n} A_k = E$.*

Dans le cas où (A_1, A_2) forme une partition de E, on a nécessairement $A_2 = \overline{A_1}$.

Exercice 1.36 *Simplifier les expressions suivantes, où A et B sont des sous-ensembles d'un ensemble E :*

1. $C = \left(\overline{A} \cap B\right) \cup \left(\overline{A} \cap \overline{B}\right) \cup (A \cap B)$
2. $\overline{C}$
3. $D = \overline{A \cap \overline{\overline{B}} \cap \left(\overline{A} \cap B\right)} \cup (A \cap B) \cap (A \cap B)$

Solution 1.36

1. *Avec la distributivité de $\cap$ sur $\cup$, on a :*

$$\overline{A} = \overline{A} \cap E = \overline{A} \cap \left(B \cup \overline{B}\right) = \left(\overline{A} \cap B\right) \cup \left(\overline{A} \cap \overline{B}\right)$$

(on a mis $\overline{A}$ en facteur) et avec la distributivité de $\cup$ sur $\cap$, on a :

$$C = \overline{A} \cup (A \cap B) = \left(\overline{A} \cup A\right) \cap \left(\overline{A} \cup B\right) = E \cap \left(\overline{A} \cup B\right) = \overline{A} \cup B.$$

2. $\overline{C} = A \cap \overline{B}$.
3. *En posant :*

$$X = A \cap \overline{B},\ Y = \overline{X} \cap \left(\overline{A} \cap B\right),\ Z = \overline{Y} \cup (A \cap B),\ T = \overline{Z} \cap (A \cap B)$$

on a :

$$\begin{aligned} D = \overline{T} &= Z \cup \left(\overline{A} \cup \overline{B}\right) = \overline{Y} \cup (A \cap B) \cup \left(\overline{A} \cup \overline{B}\right) \\ &= X \cup \left(A \cup \overline{B}\right) \cup (A \cap B) \cup \left(\overline{A} \cup \overline{B}\right) \end{aligned}$$

avec $\left(A \cup \overline{B}\right) \cup \left(\overline{A} \cup \overline{B}\right) = E$, donc $D = E$.

Exercice 1.37 *Soient $A_1, A_2, \cdots, A_p$ des ensembles deux à deux distincts. Montrer que l'un de ces ensembles ne contient aucun des autres.*

Solution 1.37 *On raisonne par l'absurde, c'est-à-dire qu'on que chacun des ensembles A_k contient un ensemble A_j différent de A_k. Donc A_1 contient un ensemble $A_{j_1} \neq A_1$, soit $A_{j_1} \subsetneqq A_1$, A_{j_1} contient un ensemble $A_{j_2} \neq A_{j_1}$, soit $A_{j_2} \subsetneqq A_{j_1}$, et on peut continuer indéfiniment, ce qui est impossible puisque la famille d'ensembles est finie.*

Exercice 1.38 *Que dire de deux ensembles A et B tels que $A \cap B = A \cup B$?*

Solution 1.38 *On a toujours $A \cap B \subset A \cup B$. Si de plus $A \cup B \subset A \cap B$, on a alors :*

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

ce qui donne $A = B$.

Exercice 1.39 *Soient A, B, C trois ensembles. Montrer que $A \cap C = A \cup B$ si, et seulement si, $B \subset A \subset C$.*

Solution 1.39 *Si $A \cap C = A \cup B$, alors :*

$$B \subset A \cup B = A \cap C \subset A \text{ et } A \subset A \cup B = A \cap C \subset C.$$

Réciproquement si $B \subset A \subset C$, alors :

$$A \cap C = A = A \cup B$$

Exercice 1.40 *Soient A, B, C trois ensembles. Montrer que si $A \cup B \subset A \cup C$ et $A \cap B \subset A \cap C$, alors $B \subset C$.*

Solution 1.40 *Soit $x \in B$. Comme $A \cup B \subset A \cup C$, x est dans $A \cup C$. S'il est dans C c'est fini, sinon il est dans A, donc dans $A \cap B \subset A \cap C$, donc dans C.*

Exercice 1.41 *Soient A, B, C trois ensembles. Montrer que :*

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

Solution 1.41 *On a :*

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

et, en notant $D = (A \cup B) \cap (B \cup C) \cap (C \cup A)$, on a :

$$D = ((B \cap C) \cup (A \cap B)) \cup (C \cap A) = (A \cap B) \cup (B \cap C) \cup (C \cap A)$$

Ou alors on part de $x \in D$ et on montre que $x \in E = (A \cap B) \cup (B \cap C) \cup (C \cap A)$, puis partant de $x \in E$, on montre que $x \in D$.

La notion de produit cartésien de deux ensembles sera très souvent utilisée. Elle correspond à l'idée de couples et se généralise pour aboutir à la notion de liste.

Définition 1.2 *Étant donné deux ensembles E et F, on appelle produit cartésien de E par F l'ensemble $E \times F$ des couples (x, y) formés d'un élément x de E et d'un élément y de F.*

Il est à noter que les couples sont ordonnés, c'est-à-dire que $(x, y) = (y, x)$ $E \times F$ si, et seulement si $x = y$. De manière plus générale, on a $(x, y) = (x', y')$ dans $E \times F$ si, et seulement si $x = x'$ et $y = y'$.

Dans le cas où $F = E$, on note E^2 pour $E \times E$.

On peut itérer le procédé et définir le produit cartésien $E_1 \times E_2 \times \cdots \times E_n$ de n ensembles comme l'ensemble des listes (ordonnées) $(x_1, x_2, \cdots, x_n)$ formées d'un élément x_1 de E_1 suivi d'un élément x_2 de E_2, $\cdots$, suivi d'un élément x_n de E_n. On notera de façon condensé :

$$\prod_{k=1}^{n} E_k = E_1 \times E_2 \times \cdots \times E_n.$$

Là encore, on a $(x_1, x_2, \cdots, x_n) = (x'_1, x'_2, \cdots, x'_n)$ dans $E \times F$ si, et seulement si $x_k = x'_k$ pour tout k compris entre 1 et n.

Dans le cas où tous les E_k sont égaux à un même ensemble E, on notera E^n pour $E \times E \times \cdots \times E$ (n fois).

Exercice 1.42 *Montrer que l'ensemble :*

$$C = \left\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\right\}$$

ne peut pas s'écrire comme produit cartésien de deux parties de $\mathbb{R}$.

Solution 1.42 *Si $C = E \times F$, où E et F sont deux parties de $\mathbb{R}$, on a alors $(1, 0) \in C = E \times F$ et $(1, 0) \in C = E \times F$, donc $1 \in E \cap F$ et $(1, 1) \in E \times F = C$, ce qui est faux.*

1.8 Applications. Notions d'injectivité, surjectivité et bijectivité

Les notations E, F, G désignent des ensembles.

Définition 1.3 *On appelle application, ou fonction, de E dans F (ou de E vers F) toute partie Γ du produit cartésien $E \times F$ telle que :*

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

En notant f une application de E dans F (c'est en réalité le triplet (E, F, Γ) avec la propriété énoncée ci-dessus), on notera pour tout $x \in E$, $f(x)$ l'unique élément de F tel que $(x, f(x)) \in \Gamma$ et on dira que $f(x)$ est l'image de x par f et x est un antécédent de y par f. Un antécédent de y par f n'est pas unique a priori.

On dira aussi que E est l'ensemble de départ (ou l'ensemble de définition), F l'ensemble d'arrivée et Γ le graphe de l'application f.

Deux applications f et g sont égales si, et seulement si, elles ont même ensemble de départ E, même ensemble d'arrivée F et même graphe Γ, c'est-à-dire que :

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

On a tout simplement précisé l'idée d'un procédé qui associe à tout élément de E un unique élément de F.

On notera :

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

une telle application (ou fonction). On utilisera aussi les notation $f : E \to F$ ou $f : x \mapsto f(x)$.

Remarque 1.2 *Nous ne faisons pas la distinction ici entre fonction et application. Usuellement, on distingue ces notions en disant qu'une fonction de E dans F toute partie Γ du produit cartésien $E \times F$ telle que pour tout élément x de E, il existe au plus un élément y de F tel que $(x, y) \in \Gamma$. Le sous-ensemble D de E pour lequel il existe un unique élément y de F tel que $(x, y) \in \Gamma$ est appelé l'ensemble de définition de la fonction. Une application est donc une fonction pour laquelle tout élément de l'ensemble de départ E a une image dans F.*

On notera $\mathcal{F}(E, F)$ ou F^E l'ensemble de toutes les applications de E dans F (la deuxième notation sera justifiée plus loin).

L'application qui associe à tout x d'un ensemble E le même x est l'application identique notée Id_E, où Id si l'ensemble E est fixé.

Si f est une fonction de E dans F et D un sous-ensemble non vide de E, on définit une application g de D dans F en posant :

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

et on dit que g est la restriction de f à D, ce qui se note $g = f_{|D}$.

Définition 1.4 *Soit f une application de E dans F. Pour toute partie A de E, l'image de A par f est le sous ensemble de F noté $f(A)$ et défini par :*

$$f(A) = \{f(x) \mid x \in A\}.$$

Pour toute partie B de F, l'image réciproque de B par f est le sous ensemble de E noté $f^{-1}(B)$ et défini par :

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

On a donc, pour tout $y \in F$:

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

et pour tout $x \in E$:

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

L'ensemble $f(E)$ est appelé l'image de f.

À propos de la notation $f^{-1}(B)$, on pourra lire la remarque 1.3 (qui n'engage que moi) plus loin.

Exemple 1.1 *On a $f(\emptyset) = \emptyset$, $f(\{x\}) = \{f(x)\}$ pour tout $x \in E$, $f^{-1}(\emptyset) = \emptyset$ et $f^{-1}(F) = E$.*

Pour tout $y \in F$, $f^{-1}\{y\}$ est l'ensemble des $x \in E$ tels que $f(x) = y$ et cet ensemble peut être vide ou formé de un ou plusieurs éléments. En fait $f^{-1}\{y\}$ est l'ensemble des solutions dans E de l'équation $f(x) = y$, où y est donné dans F et x l'inconnue dans E. Cette équation peut avoir 0 ou plusieurs solutions.

Exemple 1.2 *Pour $f : x \mapsto x^2$ avec $E = F = \mathbb{R}$, on a $f^{-1}\{0\} = \{0\}$, $f^{-1}\{-1\} = \emptyset$ et $f^{-1}\{1\} = \{-1, 1\}$.*

On vérifie facilement le résultat suivant.

Théorème 1.5 *Soit f une application de E dans F. Pour toutes parties A, B de E et C, D de F, on a :*

1. $A \subset B \Rightarrow f(A) \subset f(B)$
2. $f(A \cup B) = f(A) \cup f(B)$
3. $f(A \cap B) \subset f(A) \cap f(B)$
4. $C \subset D \Rightarrow f^{-1}(C) \subset f^{-1}(D)$
5. $f^{-1}(C \cup D) = f^{-1}(C) \cup f^{-1}(D)$
6. $f^{-1}(C \cap D) = f^{-1}(C) \cap f^{-1}(D)$
7. $f^{-1}\left(\overline{C}\right) = \overline{f^{-1}(C)}$

Démonstration. Vérification immédiate.

Par exemple, pour le point 2, on peut écrire que y est dans $f(A \cup B)$ si, et seulement si, il existe x dans $A \cup B$ tel que $y = f(x)$, ce qui implique que $y \in f(A)$ dans le cas où $x \in A$ ou $y \in f(B)$ dans le cas où $x \in B$, soit $y \in f(A) \cup f(B)$ dans tous les cas. Réciproquement si $y \in f(A) \cup f(B)$, il est dans $f(A)$ ou $f(B)$ et s'écrit donc $y = f(x)$ avec x dans A ou B, ce qui signifie que $y \in f(A \cup B)$. On a donc les inclusions $f(A \cup B) \subset f(A) \cup f(B)$ et $f(A) \cup f(B) \subset f(A \cup B)$, c'est-à-dire l'égalité souhaitée.

Pour le point 3, on a seulement une inclusion. Dire que $y \in f(A \cap B)$ équivaut à dire qu'il existe $x \in A \cap B$ tel que $y = f(x)$ et $y \in f(A) \cap f(B)$. Réciproquement, si $y \in f(A) \cap f(B)$, il existe $x_1 \in A$ et $x_2 \in B$ tels que $y = f(x_1) = f(x_2)$ et, a priori, il n'y a aucune raison pour que $x_1 = x_2$. ■

Exercice 1.43 *Vérifier sur un exemple que l'égalité $f(A \cap B) = f(A) \cap f(B)$ n'est pas toujours vérifiée.*

Solution 1.43 *Considérer* $f : x \mapsto \sin(x)$ *avec* $A = [-\pi, \pi]$ *et* $B = [0, 2\pi]$. *On a :*

$$f(A \cap B) = f([0, \pi]) = [0, 1] \subsetneq f(A) \cap f(B) = [-1, 1].$$

Exercice 1.44 *Soit* f *une application de* E *dans* F. *Vérifier que :*

1. *pour toute partie* A *de* E, $A \subset f^{-1}(f(A))$
2. *pour toute partie* B *de* F, $f(f^{-1}(B)) = B \cap f(E)$.

Solution 1.44 *Vérification immédiate.*

Exercice 1.45 *Soient* E *un ensemble et* f *une application de* $\mathcal{P}(E)$ *dans* $\mathbb{R}$ *telle que pour toutes parties disjointes de* E *on ait* $f(A \cup B) = f(A) + f(B)$.

1. *Montrer que* $f(\emptyset) = 0$.
2. *Montrer que pour toutes parties* A, B *de* E, *on a :*

$$f(A \cup B) + f(A \cap B) = f(A) + f(B).$$

Solution 1.45

1. *On a* $f(\emptyset) = f(\emptyset \cup \emptyset) = f(\emptyset) + f(\emptyset)$ *dans* $\mathbb{R}$, *donc* $f(\emptyset) = 0$.
2. *Avec les partitions* $A \cup B = A \cup (B \setminus A)$ *et* $B = (A \cap B) \cup (B \setminus A)$, *on a :*

$$\begin{cases} f(A \cup B) = f(A) + f(B \setminus A) \\ f(B) = f(A \cap B) + f(B \setminus A) \end{cases}$$

et par soustraction :

$$f(A \cup B) - f(B) = f(A) - f(A \cap B)$$

qui donne le résultat.

Après avoir défini le cardinal d'un ensemble et la notion d'ensemble fini (qui est quand même intuitive), nous verrons que si E est un ensemble fini alors la fonction f qui associe à une partie A de E son cardinal (c'est-à-dire le nombre de ses éléments) vérifie l'équation fonctionnelle de l'exercice précédent.

On dispose d'une opération importante sur les fonctions, c'est la composition des fonctions qui permet de construire de nouvelles fonctions à partir de fonctions données.

Définition 1.5 *Soient* f *une application de* E *dans* F *et* g *une application de* F *dans* G. *La composée de* f *par* g *est la fonction de* E *dans* G *notée* $g \circ f$ *et définie par :*

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

Ce qui peut se schématiser par :

$$\begin{array}{ccccc} E & \xrightarrow{f} & F & \xrightarrow{g} & G \\ x & \mapsto & f(x) & \mapsto & g(f(x)) \end{array}$$

On remarquera que $f \circ g$ n'est pas définie a priori (dans la situation de la définition).

Dans le cas ou f est définie de E dans F et g de F dans E, on peut définir les applications $f \circ g$ (de F dans F) et $g \circ f$ (de E dans E) et il n'y a aucune raison pour que ces applications soient égales, même si $F = E$.

Dans le cas où $E = F$, on dit que les applications f et g (définies de E dans E) commutent si $f \circ g = g \circ f$.

On vérifie facilement que la loi de composition est associative, c'est-à-dire que $f \circ (g \circ h) = (f \circ g) \circ h$, quand toutes ces composées ont un sens.

Cette propriété d'associativité permet de définir la composée de n applications $f_1 \circ f_2 \circ \cdots \circ f_n$ sans se soucier de parenthèses.

Si f est une application de E dans E, on peut définir la suite de ses itérées par la relation de récurrence suivante :

$$\begin{cases} f^1 = f \\ \forall n \in \mathbb{N}^*,\ f^{n+1} = f^n \circ f \end{cases}$$

On convient que $f^0 = Id_E$.

On vérifie facilement que $f^p \circ f^q = f^q \circ f^p = f^{p+q}$ pour tous entiers naturels p, q.

Exercice 1.46 *Soient E et F deux ensembles. Déterminer toutes les applications f de E dans E telles que $f \circ g = g \circ f$ pour toute application g de E dans E.*

Solution 1.46 *Soit $x \in E$ et g la fonction définie sur E par $g(y) = x$ pour tout $y \in E$ (la fonction constante égale à x). On a alors $x = g(f(x)) = f(g(x)) = f(x)$. Comme x est quelconque dans E, on déduit que $f = Id_E$.*

Les notions suivantes d'injectivité et de surjectivité sont aussi très importantes.

Définition 1.6 *Soient E, F deux ensembles et f une application de E dans F. On dit que f est :*

1. *injective (ou que c'est une injection) si deux éléments distincts de E on deux images distinctes dans F, soit :*

$$x_1 \neq x_2 \text{ dans } E \Rightarrow f(x_1) \neq f(x_2) \text{ dans } E \tag{1.3}$$

2. *surjective (ou que c'est une surjection) si tout élément de F a au moins un antécédent dans E, soit :*

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

3. *bijective (ou que c'est une bijection) si elle est à la fois injective ou surjective.*

Une injection peut aussi se caractériser en disant que tout élément de y a au plus un antécédent par f, encore équivalent à dire que pour tout $y \in F$ l'équation $y = f(x)$ a au plus une solution x dans E, ce qui revient à dire que si x_1 et x_2 sont deux éléments de E tels que $f(x_1) = f(x_2)$, alors $x_1 = x_2$ (contraposée de (1.3)).

Une surjection peut se caractériser en disant que pour tout $y \in F$ l'équation $y = f(x)$ a au moins une solution x dans E, encore équivalent à dire que $f(E) = F$.

Si f est une surjection de E dans F, on dit parfois que f est une surjection de E sur (pour surjection) F.

Une bijection peut se caractériser en disant que tout élément de y a un unique antécédent par f, encore équivalent à dire que pour que pour tout $y \in F$ l'équation $y = f(x)$ a une et une seule solution x dans E, ce qui permet de définir l'application réciproque de f, notée f^{-1}, de F dans E par :

$$\left(y \in F \text{ et } x = f^{-1}(y)\right) \Leftrightarrow (x \in E \text{ et } y = f(x)).$$

Cette application f^{-1} est une bijection de F dans E.

L'application $f \circ f^{-1}$ est alors l'application identité $y \mapsto y$ de F dans F et l'application $f^{-1} \circ f$ est alors l'application identité $x \mapsto x$ de E dans E, ce qui se note $f \circ f^{-1} = Id_F$ et $f^{-1} \circ f = Id_E$.

Définition 1.7 *On appelle permutation d'un ensemble E toute bijection de E dans lui même.*

On note en général $\mathfrak{S}(E)$ l'ensemble des permutations de E.

Exemple 1.3 *L'application $x \mapsto x^2$ est surjective de $\mathbb{R}$ dans $\mathbb{R}^+$, mais non injective. Elle bijective de $\mathbb{R}^+$ dans $\mathbb{R}^+$.*

Remarque 1.3 *Dans le cas où f est une application de E dans F, on a noté pour toute partie B de F, $f^{-1}(B)$ l'image réciproque de B par f, sans aucune hypothèse de bijectivité pour f. Dans le cas où f est bijective, $f^{-1}(B)$ est aussi l'image directe de B par f^{-1}, mais dans le cas général, il faut bien prendre garde, malgré la notation, que f n'a aucune raison d'être bijective. Il faudrait en réalité utiliser un autre symbole que f^{-1} (par exemple $f^*(B)$, $f^{(-1)}(B)$, ou $f^{\varsigma\xi\emptyset\boxtimes}(f)$), mais je préfère utiliser la notation $f^{-1}(B)$ rencontrée le plus souvent. Si l'on sait de quoi l'on parle il n'y a pas de véritable problème, il s'agit seulement d'une notation.*
On peut lire dans An introduction to the theory of numbers de Hardy et Wright, p. 7 : «We shall very often use A as in (vi), viz. an unspecified positive constant. Different A's have usually different values, even when they occur in the same formula ; and even when definite values can be assigned to them, these values are irrelevant to the argument.» C'est peut être excessif, mais l'essentiel est toujours de savoir de quoi l'on parle, on pourra ensuite écrire les choses en toute rigueur.

Exercice 1.47 *Montrer qu'une application f strictement monotone de $\mathbb{R}$ dans $\mathbb{R}$ est injective.*

Solution 1.47 *Supposons que f soit strictement croissante (au besoin on remplace f par $-f$). Si $x \neq y$, on a nécessairement $x > y$ ou $y > x$ et donc $f(x) > f(y)$ ou $f(x) < f(y)$, soit $f(x) \neq f(y)$ dans tous les cas.*

Exercice 1.48 *Soit m un entier naturel. Montrer que s'il existe un entier naturel n et une injection φ de $E_n = \{1, \cdots, n\}$ dans $E_m = \{1, \cdots, m\}$, on a alors nécessairement $n \leq m$.*

Solution 1.48 *On procède par récurrence sur $m \geq 0$.*
Si $m = 0$, on a alors $E_m = \emptyset$ et $E_n = \emptyset$ (en effet, si $E_n \neq \emptyset$, l'ensemble $f(E_n)$ est alors non vide et contenu dans l'ensemble vide, ce qui est impossible), donc $n = 0$.
Supposons le résultat acquis pour $m \geq 0$. Soit φ une injection de E_n dans E_{m+1}. Si $n = 0$, on a bien $n \leq m + 1$. Si $n \geq 1$, on distingue alors deux cas de figure :

- *soit $\varphi(n) = m + 1$ et dans ce cas φ induit une bijection de E_{n-1} dans E_m (la restriction de φ à E_{n-1}) et $n - 1 \leq m$, soit $n \leq m + 1$;*
- *soit $\varphi(n) \neq m + 1$ et dans ce cas, en désignant par ψ l'application de E_{m+1} dans lui même définie par $\psi(\varphi(n)) = m + 1$, $\psi(m + 1) = \varphi(n)$ et $\psi(k) = k$ pour $k \in E_{m+1} \setminus \{\varphi(n), m + 1\}$, l'application $\psi \circ \varphi$ est injective de E_n dans E_{m+1} (composée de deux injections puisque φ est injective et ψ bijective) avec $\psi \circ \varphi(n) = m + 1$, ce qui nous ramène au cas précédent.*

On déduit de l'exercice précédent que pour $n > m$ dans $\mathbb{N}$, il n'existe pas d'injection de $\{1, \cdots, n\}$ dans $\{1, \cdots, m\}$.

Exercice 1.49 *Soient n, m deux entiers naturels. Montrer que s'il existe une bijection φ de $E_n = \{1, \cdots, n\}$ sur $E_m = \{1, \cdots, m\}$, on a alors nécessairement $n = m$.*

Solution 1.49 *On a $n \leq m$ puisque φ est une injection de E_n dans E_m et $m \leq n$ puisque φ^{-1} est une injection de E_m dans E_n, ce qui donne $n = m$.*

Le résultat des deux exercices précédents nous seront utiles pour définir le cardinal (c'est-à-dire le nombre d'éléments) d'un ensemble fini.

Exercice 1.50 *Soient E, F deux ensembles et f une bijection de E sur F. Montrer que si g [resp. h] est une application de F sur E telle que $g \circ f = Id_E$ [resp. $f \circ h = Id_F$], alors g [resp. h] est bijective et $g = f^{-1}$ [resp. $h = f^{-1}$].*

Solution 1.50 *Résulte de $g = (g \circ f) \circ f^{-1} = Id_E \circ f^{-1} = f^{-1}$ et $h = f^{-1} \circ (f \circ h) = f^{-1} \circ Id_F = f^{-1}$.*

On vérifie facilement le résultat suivant.

Théorème 1.6 *Soient E, F, G des ensembles, f une application de E dans F et g une application de F dans G.*

1. *Si f et g sont injectives, alors $g \circ f$ est injective (la composée de deux injections est une injection).*
2. *Si f et g sont surjectives, alors $g \circ f$ est surjective (la composée de deux surjections est une surjection).*
3. *Si f et g sont bijectives, alors $g \circ f$ est bijective (la composée de deux injections est une bijection) et $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.*

Démonstration.

1. Supposons f et g injectives. Si $g \circ f(x_1) = g \circ f(x_2)$, alors $g(f(x_1)) = g(f(x_2))$, donc $f(x_1) = f(x_2)$ puisque g est injective et $x_1 = x_2$ puisque f est injective.
2. Supposons f et g surjectives. Pour tout $z \in G$, il existe $y \in F$ tel que $z = g(y)$ puisque g est surjective et $y \in F$ s'écrit $y = f(x)$ avec $x \in E$ puisque f est surjective. On a donc $z = g \circ f(x_1)$ avec $x \in E$. L'application $g \circ f$ est donc surjective.
   De manière plus compacte, on peut écrire que :
$$(g \circ f)(E) = g(f(E)) = g(F) = G.$$
3. Les deux premiers points nous disent que $g \circ f$ est bijective si f et g le sont. Puis avec $(f^{-1} \circ g^{-1}) \circ g \circ f = f^{-1} \circ Id_F \circ f = f^{-1} \circ f = Id_E$, on déduit que $f^{-1} \circ g^{-1}$ est l'inverse de $g \circ f$.

■

Exercice 1.51 *Soient E, F, G des ensembles, f une application de E dans F et g une application de F dans G. Montrer que :*

1. *si $g \circ f$ est injective, alors f est injective ;*
2. *si $g \circ f$ est surjective, alors g est surjective ;*
3. *si $g \circ f$ est surjective et g injective, alors f est surjective ;*
4. *Si $g \circ f$ est injective et f surjective, alors g est injective.*

Solution 1.51

1. *Si x, x' dans E sont tels que $f(x) = f(x')$, alors $g \circ f(x) = g \circ f(x')$ et $x = x'$ puisque $g \circ f$ est injective. L'application f est donc injective.*

2. *Pour tout z dans G, il existe x dans E tel que $z = g \circ f(x)$ puisque $g \circ f$ est surjective et en notant $y = f(x)$, on a $y \in F$ et $z = g(y)$, ce qui prouve que g est surjective.*
3. *Soit $y \in F$. Comme $g \circ f$ est surjective, il existe $x \in E$ tel que $z = g(y) = (g \circ f)(x) = g(f(x))$ et $y = f(x)$ si on suppose de plus que g est injective. En conséquence, f est surjective.*
4. *Soient y, y' dans F tels que $g(y) = g(y')$. Comme f est surjective, il existe x, x' dans E tels que $y = f(x)$ et $y' = f(x')$, ce qui donne $g \circ f(x) = g \circ f(x')$ et $x = x'$ puisque $g \circ f$ est injective, donc $y = y'$.*

Le résultat qui suit peut parfois être utile pour montrer l'injectivité, la surjectivité ou la bijectivité d'une application.

Théorème 1.7 *Soient E, F deux ensembles et f une application de E dans F.*

1. *S'il existe une application g de F dans E telle que $g \circ f = Id_E$, alors f est injective.*
2. *S'il existe une application h de F dans E telle que $f \circ h = Id_F$, alors f est surjective.*
3. *S'il existe deux applications g et h de F dans E telles que $g \circ f = Id_E$ et $f \circ h = Id_F$, alors f est bijective et $g = h = f^{-1}$.*

Démonstration.

1. Si x, x' dans E sont tels que $f(x) = f(x')$, alors $x = g \circ f(x) = g \circ f(x') = x'$ et f est injective.
2. Pour tout $y \in F$, on a $y = (f \circ h)(y) = f(h(y))$ avec $x = h(y) \in E$, donc f est surjective.
3. Les deux premiers points nous disent que f est bijective et de $g \circ f = Id_E$, on déduit que $f^{-1} = (g \circ f) \circ f^{-1} = g$. De même $h = g^{-1}$.

■

Exercice 1.52 *Soient m un entier naturel non nul et E un ensemble non vide. Montrer que s'il existe une surjection φ de $E_m = \{1, \cdots, m\}$ sur E, on peut alors construire une injection de E dans E_m.*

Solution 1.52 *Comme φ est surjective de E_m sur E, on a $\varphi^{-1}\{x\} \neq \emptyset$ pour tout $x \in E$ et chacun de ces sous-ensembles de E_m a un plus petit élément $j_x = \min \varphi^{-1}\{x\} \in E_m$, ce qui permet de définir l'application ψ de E dans E_m par :*

$$\forall x \in E,\ \psi(x) = j_x$$

On a alors :

$$\forall x \in E,\ \varphi \circ \psi(x) = \varphi(j_x) = x$$

c'est-à-dire que $\varphi \circ \psi = Id_E$ et l'application ψ est injective (théorème précédent).

Exercice 1.53 *Soient n, m deux entiers naturels non nuls. Montrer que s'il existe une surjection φ de $E_n = \{1, \cdots, n\}$ sur $E_m = \{1, \cdots, m\}$, on a alors nécessairement $n \geq m$.*

Solution 1.53 *En utilisant le résultat de l'exercice précédent, on peut construire une injection de E_m dans E_n et nécessairement $m \leq n$ (exercice 1.48).*

Exercice 1.54 *Soient E un ensemble et f une application de E dans E. Montrer que f est injective si, et seulement si, $f(A \cap B) = f(A) \cap f(B)$ pour toutes partie A et B de E.*

Solution 1.54 *On a toujours $f(A\cap B)\subset f(A)\cap f(B)$ pour toutes partie A et B de E, que f soit injective ou pas. En effet un élément y de $f(A\cap B)$ s'écrit $y=f(x)$ avec $x\in A\cap B$ et donc $y\in f(A)\cap f(B)$. Réciproquement si $y\in f(A)\cap f(B)$, il existe $x\in A$ et $x'\in B$ tels que $y=f(x)=f(x')$ et dans le cas où f est injective, on a nécessairement $x=x'\in A\cap B$, donc $y\in f(A\cap B)$.*
On a donc $f(A\cap B)=f(A)\cap f(B)$ pour toutes partie A et B de E, si f est injective.
Réciproquement supposons que $f(A\cap B)=f(A)\cap f(B)$ pour toutes partie A et B de E. Si f n'est pas injective, il existe $x\neq x'$ dans E tels que $f(x)=f(x')$ et :

$$\emptyset=f(\emptyset)=f(\{x\}\cap\{x'\})=f(\{x\})\cap f(\{x\})=f(\{x\})=\{f(x)\}$$

ce qui est impossible. Donc f est injective.

Exercice 1.55 *Soient E un ensemble et f une application de E dans E. Montrer que f est bijective si, et seulement si, $f\left(\overline{A}\right)=\overline{f(A)}$ pour toute partie A de E.*

Solution 1.55 *Supposons f bijective. Un élément y de E est dans $f\left(\overline{A}\right)$ si, et seulement si, il s'écrit $y=f(x)$ où x est uniquement déterminé dans $\overline{A}$, ce qui implique $y\notin f(A)$ (sinon $y=f(x')=f(x)$ avec $x'\in A$ et $x=x'\in A$, ce qui contredit $x\in\overline{A}$). On a donc $f\left(\overline{A}\right)\subset\overline{f(A)}$. Si $y\notin f(A)$, il s'écrit $y=f(x)$ (f est bijective) et $x\notin A$, donc $y\in f\left(\overline{A}\right)$. On a donc $\overline{f(A)}\subset f\left(\overline{A}\right)$ et $f\left(\overline{A}\right)=\overline{f(A)}$.*
Supposons que $f\left(\overline{A}\right)=\overline{f(A)}$ pour toute partie A de E. En particulier, on a $f(E)=f\left(\overline{\emptyset}\right)=\overline{f(\emptyset)}=\overline{\emptyset}=E$ et f est surjective. Si $x\neq x'$ dans E, en remarquant que $x'\in\overline{\{x\}}$, on a $f(x')\in f\left(\overline{\{x\}}\right)=\overline{\{f(x)\}}$ et $f(x)\neq f(x')$. Donc f est injective.

Exercice 1.56 *Soient E,F,G,H des ensembles, f une application de E dans F, g une application de F dans G et h une application de G dans H. Montrer que si $g\circ f$ et $h\circ g$ sont bijectives, alors f,g et h sont bijectives.*

Solution 1.56 *Si $g\circ f$ est bijective, elle est alors surjective et il en est de même de g (exercice 1.51). Si $h\circ g$ est bijective, elle est alors injective et il en est de même de g (exercice 1.51). Donc g est bijective. Il en résulte que $f=g^{-1}\circ(g\circ f)$ et $h=(h\circ g)\circ g^{-1}$ sont bijectives comme composées.*

Exercice 1.57 *On désigne par f l'application définie sur $\mathbb{N}^2=\mathbb{N}\times\mathbb{N}$ par :*

$$\forall(n,m)\in\mathbb{N}^2,\ f(n,m)=2^n3^m$$

Montrer que f est injective. Il résulte que $\mathbb{N}^2$ est en bijection avec le sous ensemble $f(\mathbb{N}^2)$ de $\mathbb{N}$. Ce résultat se traduit en disant que $\mathbb{N}^2$ est dénombrable.

Solution 1.57 *L'égalité $f(n,m)=f(n',m')$ avec (n,m) et (n',m') dans $\mathbb{N}^2$ équivaut à $2^n3^m=2^{n'}3^{m'}$ et l'unicité de la décomposition en facteurs premiers d'un entier naturel non nul nous dit que $(n,m)=(n',m')$. L'application f est donc injective de $\mathbb{N}^2$ dans $\mathbb{N}$ et bijective de $\mathbb{N}^2$ dans $f(\mathbb{N}^2)\subset\mathbb{N}$.*

Exercice 1.58 *Montrer que l'application $f:(n,m)\mapsto 2^{n+m+1}+2^m$ est injective de $\mathbb{N}^2$ dans $\mathbb{N}$.*

Solution 1.58 *L'égalité $f(n,m)=f(n',m')$ avec (n,m) et (n',m') dans $\mathbb{N}^2$ équivaut à $2^m\left(2^{n+1}+1\right)=2^{m'}\left(2^{n'+1}+1\right)$. Si $m>m'$, on a alors $2^{m-m'}\left(2^{n+1}+1\right)=2^{n'+1}+1$ qui est à la fois pair et impair, ce qui est impossible. De manière analogue, on voit que $m'>m$ est impossible. On a donc $m=m'$ et $2^{n+1}+1=2^{n'+1}+1$, ce qui équivaut à $n=n'$. L'application f est donc injective.*

1.9 Cardinal d'un ensemble fini

La notion d'ensemble fini est relativement intuitive, c'est un ensemble dont on peut numéroter les éléments de 1 (ou de 0) à n où n est un entier naturel non nul (si on numérote à partir de 1). Par exemple $\{1,2,3,4,5\}$ est fini et on a envie de dire qu'il a 5 éléments.

On rappelle que si n est un entier naturel, l'ensemble $\{1,\cdots,n\}$ est l'ensemble vide pour $n=0$ et l'ensemble des entiers compris entre 1 et n pour n non nul.

De manière précise, on peut donner la définition suivante.

Définition 1.8 *On dit qu'un ensemble E est fini s'il existe un entier naturel n et une bijection φ de $\{1,\cdots,n\}$ sur E.*
Un ensemble qui n'est pas fini est dit infini.

Remarque 1.4 *Si $n=0$ dans la définition ci-dessus, on dispose alors d'une application f de E dans l'ensemble vide (la bijection réciproque φ^{-1}) et l'ensemble E est nécessairement vide. En effet si $E\neq\emptyset$, on a alors $f(E)\neq\emptyset$ qui est contenu dans l'ensemble vide, ce qui est impossible.*

Si φ est une bijection φ de $\{1,\cdots,n\}$ sur E avec $n\geq 1$, on a alors $E=\varphi(\{1,\cdots,n\})=\{\varphi(1),\cdots,\varphi(n)\}$ et il semble naturel de dire que n est le nombre d'éléments de E. Pour valider cette définition, on a besoin du résultat suivant qui nous assure l'unicité d'un tel entier n.

Théorème 1.8 *Si un ensemble E est en bijection avec un ensemble $\{1,\cdots,n\}$ où n est un entier naturel n, alors cet entier n est unique.*

Démonstration. Si φ est une bijection de $E_n=\{1,\cdots,n\}$ sur E et ψ une bijection de $E_m=\{1,\cdots,m\}$ sur E, alors $\psi^{-1}\circ\varphi$ est une bijection de E_n sur E_m et $n=m$ (exercice 1.49). ■

On peut donc donner la définition suivante.

Définition 1.9 *Soit E un ensemble fini. Si φ est une bijection de $E_n=\{1,\cdots,n\}$ sur E, où n est un entier naturel, on dit alors que n est le cardinal (ou le nombre d'éléments) de E et on note $n=\operatorname{card}(E)$ (ou encore $\#(E)$).*

Une bijection φ de E_n sur un ensemble fini non vide E nous permet de numéroter les éléments de E et on peut noter :

$$E=\{x_1,x_2,\cdots,x_n\}$$

où $x_k=\varphi(k)$ pour k compris entre 1 et n (ces x_k sont deux à deux distincts).

Exemple 1.4 *L'ensemble vide $\emptyset=\{1,\cdots,0\}$ est de cardinal nul.*
Un singleton $\{a\}$ en bijection avec $\{1\}$ est de cardinal 1.
Bien entendu l'ensemble $\{1,\cdots,n\}$ est de cardinal n, pour tout entier naturel n.
De manière plus générale, l'ensemble $\{p+1,\cdots,p+n\}$ des entiers compris entre $p+1$ et $p+n$, où p est un entier relatif et n un entier naturel est de cardinal n (l'application $k\mapsto k+p$ réalise une bijection de $\{1,\cdots,n\}$ sur $\{p+1,\cdots,p+n\}$).
Si p,q sont deux entier relatifs avec $p\leq q$, l'ensemble $\{p,\cdots,q\}$ des entiers compris entre p et q, est de cardinal $q-p+1$.

Avec les deux théorèmes qui suivent, on donne les propriétés essentielles du cardinal d'un ensemble fini.

Théorème 1.9

1. *Si E, F sont deux ensembles finis disjoints, alors $E \cup F$ est fini et :*

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

2. *Si F est une partie d'un ensemble fini E, alors :*

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

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

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

5. *Si $(E_k)_{1 \leq k \leq p}$ est une famille finie d'ensembles finis et deux à deux disjoints, alors :*

$$\operatorname{card}\left(\bigcup_{k=1}^{p} E_k\right) = \sum_{k=1}^{p} \operatorname{card}(E_k)$$

6. *Si E, F sont deux ensembles finis, alors le produit cartésien $E \times F$ est fini et :*

$$\operatorname{card}(E \times F) = \operatorname{card}(E) \operatorname{card}(F)$$

7. *Si $(E_k)_{1 \leq k \leq p}$ est une famille finie d'ensembles finis, alors le produit cartésien $\prod_{k=1}^{p} E_k$ est fini et :*

$$\operatorname{card}\left(\prod_{k=1}^{p} E_k\right) = \prod_{k=1}^{p} \operatorname{card}(E_k)$$

Démonstration.

1. On désigne par n le cardinal de E et par m celui de F. On dispose donc d'une bijection f de E sur $E_n = \{1, \cdots, n\}$ et d'une bijection g de F sur $E_m = \{1, \cdots, m\}$. L'application h définie sur $E \cup F$ par :

$$h(x) = \begin{cases} f(x) \text{ si } x \in E \\ n + g(x) \text{ si } x \in F \end{cases}$$

   réalise alors une bijection de $E \cup F$ sur $E_{n+m} = \{1, \cdots, n+m\}$. En effet, elle est bien définie puisque E et F sont disjoints et pour tout $k \in E_{n+m}$ il existe un unique $x \in E \cup F$ tel que $k = h(x)$, cet élément étant $x = f^{-1}(k)$ si $1 \leq k \leq n$ ou $x = g^{-1}(k-n)$ si $n+1 \leq k \leq m$. L'ensemble $E \cup F$ est donc fini de cardinal $n + m = \operatorname{card}(E) + \operatorname{card}(F)$.
2. Avec la partition $E = (E \setminus F) \cup F$, on déduit que $\operatorname{card}(E) = \operatorname{card}(F) + \operatorname{card}(E \setminus F)$.
3. De l'égalité précédente, on déduit que $\operatorname{card}(F) \leq \operatorname{card}(E)$.
   Supposons que $\operatorname{card}(E) = \operatorname{card}(F)$. Si $F \neq E$, il existe $x \in E \setminus F$ et de l'inclusion $F \cup \{x\} \subset E$ avec $F \cap \{x\} = \emptyset$, on déduit $\operatorname{card}(F) + 1 \leq \operatorname{card}(E)$, ce qui contredit l'égalité $\operatorname{card}(E) = \operatorname{card}(F)$. On a donc $F = E$. La réciproque est évidente.

4. Des partitions :
$$E \cup F = (E \setminus F) \cup F \text{ et } E = (E \setminus F) \cup (E \cap F)$$
on déduit que :
$$\operatorname{card}(E \cup F) = \operatorname{card}(E \setminus F) + \operatorname{card}(F)$$
et :
$$\operatorname{card}(E) = \operatorname{card}(E \setminus F) + \operatorname{card}(E \cap F)$$
ce qui donne $\operatorname{card}(E \cup F) = \operatorname{card}(E) + \operatorname{card}(F) - \operatorname{card}(E \cap F)$.
5. Laissée au lecteur.
6. Laissée au lecteur.
7. Laissée au lecteur.

■

Exercice 1.59 *Montrer que si E, F, G sont trois ensembles finis, alors $E \cup F \cup G$ est fini et :*

$$\begin{aligned}\operatorname{card}(E \cup F \cup G) = \operatorname{card}(E) + \operatorname{card}(F) + \operatorname{card}(G) \\ - \operatorname{card}(E \cap F) - \operatorname{card}(E \cap G) - \operatorname{card}(F \cap G) \\ + \operatorname{card}(E \cap F \cap G)\end{aligned}$$

Solution 1.59 *Laissée au lecteur.*

Théorème 1.10 *Soient E, F deux ensembles finis non vides et φ une application de E dans F.*

1. *Si φ est injective, alors* $\operatorname{card}(E) \leq \operatorname{card}(F)$.
2. *Si φ est surjective, alors* $\operatorname{card}(E) \geq \operatorname{card}(F)$.
3. *Si φ est bijective, alors* $\operatorname{card}(E) = \operatorname{card}(F)$.
4. *On a* $\operatorname{card}(\varphi(E)) \leq \min(\operatorname{card}(E), \operatorname{card}(F))$ *et φ est injective si, et seulement si* $\operatorname{card}(\varphi(E)) = \operatorname{card}(E)$, *$\varphi$ est surjective si, et seulement si* $\operatorname{card}(\varphi(E)) = \operatorname{card}(F)$.
5. *Si E et F sont de même cardinal, alors :*
$$\varphi \text{ injective} \Leftrightarrow \varphi \text{ surjective} \Leftrightarrow \varphi \text{ bijective}$$
6. *S'il existe un entier naturel non nul p tel que pour tout $y \in F$, $\varphi^{-1}\{y\}$ est de cardinal p, alors φ est surjective et* $\operatorname{card}(E) = p \operatorname{card}(F)$ *(principe des bergers).*

Démonstration. On désigne par n le cardinal de E et par m celui de F. On dispose donc d'une bijection f de $E_n = \{1, \cdots, n\}$ sur E et d'une bijection g de $E_m = \{1, \cdots, m\}$ sur F.

1. Si $\varphi : E \to F$ est injective, alors $g^{-1} \circ \varphi \circ f$ est injective de E_n dans E_m et $n \leq m$ (exercice 1.48).
2. Si $\varphi : E \to F$ est surjective, alors $g^{-1} \circ \varphi \circ f$ est surjective de E_n dans E_m et $n \geq m$ (exercice 1.53).
3. Résulte des deux points précédents.

(a) Comme $\varphi(E) \subset F$, on a $\operatorname{card}(\varphi(E)) \leq \operatorname{card}(F)$. En notant $\varphi(E) = \{y_1, \cdots, y_p\}$ où les y_k, pour k compris entre 1 et p, sont deux à deux distincts, on a la partition $E = \bigcup_{k=1}^{p} \varphi^{-1}\{y_k\}$ et :

$$\operatorname{card}(E) = \sum_{k=1}^{p} \operatorname{card}\left(\varphi^{-1}\{y_k\}\right) \geq p = \operatorname{card}(\varphi(E))$$

puisque les sont tous non vides.

(b) Si φ est injective, elle induit alors une bijection de E sur $\varphi(E)$ et $\operatorname{card}(\varphi(E)) = \operatorname{card}(E)$.
Réciproquement si $p = \operatorname{card}(\varphi(E)) = \operatorname{card}(E)$ (notations du **4.a.**), les $\varphi^{-1}\{y_k\}$ sont tous de cardinal égal à 1, ce qui signifie que tout élément de $\varphi(E)$ a un unique antécédent dans E, donc φ est bijective de E sur $\varphi(E)$ et injective de E dans F.

(c) Si φ est surjective, on a alors $\varphi(E) = F$, donc $\operatorname{card}(\varphi(E)) = \operatorname{card}(F)$.
Réciproquement si $\operatorname{card}(\varphi(E)) = \operatorname{card}(F)$, on a $\varphi(E) = F$ et φ est surjective.

4. Si φ est injective, on alors $\operatorname{card}(\varphi(E)) = \operatorname{card}(E) = \operatorname{card}(F)$, donc $\varphi(E) = F$ et φ est surjective.
Si φ est surjective, on a alors $\operatorname{card}(\varphi(E)) = \operatorname{card}(F) = \operatorname{card}(E)$ et φ est injective, donc bijective.
Enfin si φ est bijective, elle injective.
Les trois propositions sont donc bien équivalentes.

5. Si $\varphi^{-1}\{y\}$ est de cardinal $p \geq 1$ pour tout $y \in F$, tous ces ensembles sont non vides et φ est surjective.
En notant $F = \{y_1, \cdots, y_m\}$ où les y_k, pour k compris entre 1 et m, sont deux à deux distincts, on a la partition :

$$E = \varphi^{-1}(F) = \varphi^{-1}\left(\bigcup_{k=1}^{m} \{y_k\}\right) = \bigcup_{k=1}^{m} \varphi^{-1}\{y_k\}$$

et :

$$\operatorname{card}(E) = \sum_{k=1}^{m} \operatorname{card}\left(\varphi^{-1}\{y_k\}\right) = mp = p \operatorname{card}(F).$$

■

Exercice 1.60 *Montrer qu'une partie de $\mathbb{N}$ est finie si, et seulement si, elle est majorée.*

Solution 1.60 *Si E est une partie majorée de $\mathbb{N}$, il existe alors un entier n tel que E soit contenue dans $\{1, \cdots, n\}$ et E est finie de cardinal au plus égal à n.*
Pour la réciproque, on procède par récurrence sur le cardinal.
Un ensemble de cardinal nul est vide et majoré par n'importe quel entier (l'assertion : $\forall x \in \emptyset$, $x \leq 27$ est vraie).
Supposons le résultat acquis pour les parties de $\mathbb{N}$ de cardinal $n \geq 0$ et soit E une partie de $\mathbb{N}$ de cardinal $n+1$. Pour $p \in E$ (E est non vide puisque de cardinal $n+1 \neq 0$) l'ensemble $E \setminus \{p\}$ est de cardinal n, donc majoré par un entier M et $M' = \max(M, p)$ est un majorant de E.

De cet exercice, on déduit que $\mathbb{N}$ est infini (est-ce une évidence ?), donc aussi $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ et $\mathbb{C}$.

1.10 Ensembles infinis dénombrables

Définition 1.10 *On dit qu'un ensemble E est infini dénombrable s'il existe une bijection de E sur $\mathbb{N}$.*

On dira simplement dénombrable pour infini dénombrable.

Si E est un ensemble dénombrable, une bijection $n \mapsto \varphi(n)$ de $\mathbb{N}$ sur E permet de numéroter les éléments de E :

$$E = \{\varphi(0), \varphi(1), \cdots, \varphi(n), \cdots\}$$

On notera plus simplement, $e_k = \varphi(k)$ où k est un entier naturel, les éléments de E.

Exercice 1.61 *Montrer que l'ensemble $\mathbb{Z}$ des entiers relatifs est dénombrable.*

Solution 1.61 *On peut vérifier que $\mathbb{Z}$ est dénombrable en ordonnant les entiers relatifs comme suit :*

$$\mathbb{Z} = \{0, -1, 1, -2, 2, \cdots, -k, k, \cdots\}$$

ce qui revient à vérifier que l'application :

$$\begin{array}{rccl} \varphi : & \mathbb{Z} & \to & \mathbb{N} \\ & n & \mapsto & \left\{ \begin{array}{l} 2n \text{ si } n \text{ est positif ou nul} \\ -2n-1 \text{ si } n \text{ est strictement négatif} \end{array} \right. \end{array}$$

est bijective.
Supposons que n, m soient deux entiers relatifs tels que $\varphi(n) = \varphi(m)$. Si $n \geq 0$ et $m < 0$, on a alors $2n = -2m-1$, soit $2(n+m) = 1$ dans $\mathbb{Z}$, ce qui est impossible. De même $n < 0$ et $m \geq 0$ est impossible. On a donc soit $n \geq 0$, $m \geq 0$, donc $2n = 2m$ et $n = m$, soit $n < 0, m < 0$, donc $-2n-1 = -2m-1$ et $n = m$. L'application φ est donc injective.
Si k est un entier naturel, il est soit pair, donc $k = 2n = \varphi(n)$ avec $n \in \mathbb{N}$, soit impair, donc $k = 2(-n) - 1 = \varphi(n)$ avec $n \in \mathbb{Z}^{*,-}$. L'application φ est donc surjective.

On peut montrer de plusieurs façons que l'ensemble $\mathbb{N}^2$ est dénombrable.

Exercice 1.62 *Montrer que l'application $\varphi : (n, m) \mapsto 2^n(2m+1) - 1$ est bijective de $\mathbb{N}^2$ sur $\mathbb{N}$.*

Solution 1.62 *L'égalité $\varphi(n, m) = \varphi(n', m')$ avec (n, m) et (n', m') dans $\mathbb{N}^2$ équivaut à $2^n(2m+1) = 2^{n'}(2m'+1)$. Si $n > n'$, on a alors $2^{n-n'}(2m+1) = 2m'+1$ qui est à la fois pair et impair, ce qui est impossible. De manière analogue, on voit que $n' > n$ est impossible. On a donc $n = n'$ et $2m+1 = 2m'+1$, ce qui équivaut à $m = m'$. L'application φ est donc injective.*
Soit $r \in \mathbb{N}$. Si $r = 0$, on a alors $r = \varphi(0, 0)$. Si $n \in \mathbb{N}^*$, alors $n+1 \geq 2$ et cet entier se décompose en facteurs premiers, ce qui s'écrit $r+1 = 2^n(2m+1)$ avec (n, m) dans $\mathbb{N}^2$. L'application φ est donc surjective et en définitive bijective.

Exercice 1.63 *Montrer que l'application $\varphi : (n, m) \mapsto \dfrac{1}{2}(n+m)(n+m+1) + m$ est bijective de $\mathbb{N}^2$ sur $\mathbb{N}$.*

Solution 1.63 *On remarque d'abord que pour tout $(n, m) \in \mathbb{N}^2$, les entiers $n+m$ et $n+m+1$ sont de parités différentes, donc $\dfrac{1}{2}(n+m)(n+m+1)$ est entier et φ est bien à valeurs dans $\mathbb{N}$.*

Supposons que $\varphi(n,m)=\varphi(n',m')$ *avec* (n,m) *et* (n',m') *dans* $\mathbb{N}^2$. *En notant* $N=n+m$ *et* $M=n'+m'$, *on a* $M\geq m'$ *et :*

$$\frac{N(N+1)}{2}\leq\varphi(n,m)=\varphi(n',m')\leq\frac{M(M+1)}{2}+M$$

ce qui entraîne

$$M^2+3M=\left(M+\frac{3}{2}\right)^2-\frac{9}{4}\geq N^2+N=\left(N+\frac{1}{2}\right)^2-\frac{1}{4}$$

soit :

$$(2M+3)^2-9\geq(2N+1)^2-1$$

ou encore :

$$(2M+3)^2-(2N+1)^2=(2(M+N)+4)(2(M-N)+2)\geq 8$$

c'est-à-dire :

$$(M+N+2)(M-N+1)\geq 2$$

et nécessairement $M\geq N$. *Comme* M *et* N *jouent des rôles symétriques, on a aussi* $M\leq N$ *et* $M=N$. *De* $\varphi(n,m)=\varphi(n',m')$, *on déduit alors que* $m=m'$ *puis* $n=n'$. *L'application* φ *est donc injective.*

2

Le corps $\mathbb{C}$ des nombres complexes

Les ensembles $\mathbb{Z}$ d'entiers relatifs et $\mathbb{Q}$ de nombres rationnels peuvent être construits à partir de problèmes analogues. Pour l'ensemble $\mathbb{Z}$ il s'agit des équations $x + a = 0$ qui n'ont pas de solution dans $\mathbb{N}$ pour a entier naturel non nul et pour l'ensemble $\mathbb{Q}$ il s'agit des équations $ax = 1$ qui n'ont pas de solution dans $\mathbb{Z}$ pour a entier relatif différent de $-1, 0$ et 1. Le passage de l'ensemble $\mathbb{Q}$ de nombres rationnels à l'ensemble $\mathbb{R}$ de nombres réels est plus délicat. Les problèmes sont de nature algébrique (par exemple l'équation $x^2 = 2$ n'a pas de solution dans $\mathbb{Q}$) mais aussi de nature topologique : l'existence de borne supérieure pour les ensembles non vides et majorés n'est pas assurée dans $\mathbb{Q}$ alors qu'elle l'est dans $\mathbb{R}$ (par exemple l'ensemble $A = \{r \in \mathbb{Q} \mid r^2 \leq 2\}$ n'a pas de borne supérieure dans $\mathbb{Q}$). On consultera le cours d'analyse pour de plus amples détails sur la construction de l'ensemble $\mathbb{R}$ des nombres réels.

La construction de l'ensemble des nombres complexes est motivée par le fait que certaines équations polynomiales telles que l'équation $x^2 + 1 = 0$ n'ont pas de solutions réelles.

Le but de ce chapitre est de construire un ensemble que nous noterons $\mathbb{C}$ qui contient $\mathbb{R}$ et qui est muni d'opérations d'addition et de multiplication ayant les mêmes propriétés que leurs analogues sur $\mathbb{R}$, ce qui se traduira en disant que $\mathbb{C}$ est un corps commutatif. De plus dans cet ensemble $\mathbb{C}$ toute équation algébrique $P(x) = 0$, où P est un polynôme non constant, a des solutions, ce qui se traduira en disant que $\mathbb{C}$ est algébriquement clos. Dans un premier temps, on se contentera de décrire les solutions des équations de degré 2, $x^2 + bx + c = 0$. Pour les équations de degré supérieur, on dispose du théorème de d'Alembert-Gauss dit théorème fondamental de l'algèbre dont la démonstration classique nécessite des outils d'analyse réelle tels que le fait qu'une fonction continue sur un compact de $\mathbb{C}$ est bornée et atteint ses bornes (voir le cours d'analyse et le problème du paragraphe **??**). Nous verrons aussi que contrairement à $\mathbb{R}$, l'ensemble $\mathbb{C}$ que nous aurons construit ne peut pas être muni d'une relation d'ordre compatible avec la multiplication, c'est-à-dire telle que si $x \leq y$ et $0 \leq z$, alors $x \cdot z \leq y \cdot z$.

2.1 Conditions nécessaires à la construction de $\mathbb{C}$

Supposons que nous ayons construit un ensemble $\mathbb{C}$ contenant $\mathbb{R}$ muni d'opérations d'addition et multiplication qui prolongent celles que nous connaissons sur les réels avec les même propriétés (mises à part celle relatives à la relation d'ordre $\leq$ sur $\mathbb{R}$) et tel que l'équation $x^2 + 1 = 0$ admette au moins une solution i dans $\mathbb{C}$.

Pour tous réels x, y, le nombre $z = x + iy$ sera alors dans $\mathbb{C}$ et l'égalité $z = 0$ est réalisée si, et seulement si, $x = y = 0$. En effet, si $y = 0$, alors $x = 0$ et si $y \neq 0$, alors $i = -\dfrac{x}{y}$ est réel, ce qui n'est pas possible puisque l'équation $x^2 + 1$ n'a pas de solution réelle. Il en résulte que pour x, x', y, y' réels l'égalité $x + iy = x' + iy'$ est réalisée si, et seulement si, $x = x'$ et $y = y'$.

De plus pour l'addition et la multiplication de deux éléments $z = x + iy$ et $z' = x' + iy'$ de $\mathbb{C}$, on doit avoir :

$$\begin{cases} z + z' = (x + x') + i\,(y + y') \\ zz' = (xx' - yy') + i\,(xy' + yx') \end{cases}$$

2.2 Construction de $\mathbb{C}$

Les considérations précédentes nous conduisent à définir sur l'ensemble $\mathbb{R}^2$ des couples de réels les opérations d'addition et de multiplication suivantes, où $z = (x, y)$ et $z' = (x', y')$ sont deux éléments quelconques de $\mathbb{R}^2$:

$$\begin{cases} z + z' = (x + x', y + y') \\ z \cdot z' = (xx' - yy', xy' + yx') \end{cases}$$

On notera $\mathbb{C}$ l'ensemble $\mathbb{R}^2$ muni de ces deux opérations (ou lois de composition interne) et on l'appelle ensemble des nombres complexes.

La multiplication de deux nombres complexes z et z' sera notée $z \cdot z'$ ou plus simplement zz'.

L'égalité de deux nombres complexes $z = (x, y)$ et $z' = (x', y')$ est réalisée si, et seulement si, on a les égalités $x = x'$ et $y = y'$ (c'est ce qui se passe dans tout produit cartésien $E \times F$).

En particulier, pour $y = y' = 0$, on a $(x, 0) = (x', 0)$ si, et seulement si $x = x'$, ce qui signifie que l'application :

$$\begin{array}{ccc} \mathbb{R} & \to & \mathbb{C} \\ x & \mapsto & (x, 0) \end{array}$$

est injective, ce qui permet de réaliser une bijection de $\mathbb{R}$ sur le sous ensemble $\mathbb{R}'$ de $\mathbb{C}$ formé des couples $(x, 0)$. Cette bijection permet d'identifier $\mathbb{R}$ à $\mathbb{R}'$, ce qui signifie qu'un nombre réel x est identifié à son image $(x, 0)$ dans $\mathbb{C}$. Cette identification $x = (x, 0)$ est bien compatible avec les opérations d'addition et de multiplication des réels dans le sens où :

$$\begin{cases} x + x' = (x, 0) + (x', 0) = (x + x', 0) = x + x' \\ xx' = (x, 0)\,(x', 0) = (xx', 0) = xx' \end{cases}$$

L'opération d'addition vérifie les propriétés suivantes, déduites des propriétés analogues sur $\mathbb{R}$:

- elle est commutative, c'est-à-dire que pour tous nombres complexes $z = (x, y)$ et $z' = (x', y')$, on a :
$$z + z' = (x + x', y + y') = (x' + x, y' + y) = z' + z$$
- elle est associative, c'est-à-dire que pour tous nombres complexes $z = (x, y)$, $z' = (x', y')$ et $z'' = (x'', y'')$, on a :
$$\begin{aligned} z + (z' + z'') &= (x, y) + (x' + x'', y' + y'') = (x + x' + x'', y + y' + y'') \\ &= ((x + x') + x'', (y + y') + y'') = (x + x', y + y') + (x'', y'') \\ &= (z + z') + z'' \end{aligned}$$
- le réel $0 = (0, 0)$ est un élément neutre, c'est-à-dire que pour tout nombre complexe $z = (x, y)$, on a :
$$z + 0 = 0 + z = z$$

– tout nombre complexe $z = (x, y)$ admet un opposé donné par $z' = (-x, -y)$, ce qui signifie que :

$$z + z' = z' + z = 0$$

On note $-z$ cet opposé.

Tout cela se traduit en disant que $(\mathbb{C}, +)$ est un groupe commutatif comme $(\mathbb{R}, +)$, $(\mathbb{Q}, +)$ et $(\mathbb{Z}, +)$.

La notion de groupe est étudiée plus en détails au chapitre suivant.

Comme pour n'importe quel groupe, on peut vérifier que :

– l'élément neutre est unique ;

– pour tout $z \in \mathbb{C}$, l'opposé est unique ;

– tout élément de $\mathbb{C}$ est simplifiable (ou régulier) pour l'addition, c'est-à-dire que si $z + z' = z + z''$, alors $z' = z''$.

Pour ce qui est de l'autre opération de multiplication, on a les propriétés suivantes, encore déduites des propriétés analogues sur $\mathbb{R}$:

– elle est commutative, c'est-à-dire que pour tous nombres complexes $z = (x, y)$ et $z' = (x', y')$, on a :

$$zz' = (xx' - yy', xy' + yx') = (x'x - y'y, y'x + x'y) = z'z$$

– elle est associative, c'est-à-dire que pour tous nombres complexes $z = (x, y)$, $z' = (x', y')$ et $z'' = (x'', y'')$, on a $z(z'z'') = (zz')z''$. En effet, on a :

$$\begin{aligned} z(z'z'') &= (x, y)(x'x'' - y'y'', x'y'' + y'x'') \\ &= (x(x'x'' - y'y'') - y(x'y'' + y'x''), x(x'y'' + y'x'') + y(x'x'' - y'y'')) \\ &= (xx'x'' - xy'y'' - yx'y'' - yy'x'', xx'y'' + xy'x'' + yx'x'' - yy'y'') \end{aligned}$$

et :

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

On peut remarquer que seule la commutativité de l'addition des réels a été utilisée ici.

– elle est distributive par rapport à l'addition, c'est-à-dire que pour tous nombres complexes z, z' et z'', on a $z(z' + z'') = zz' + zz''$, ce qui se vérifie encore sans problème.

– le réel $1 = (1, 0)$ est un élément neutre, c'est-à-dire que pour tout nombre complexe $z = (x, y)$, on a :

$$z \cdot 1 = 1 \cdot z = z$$

– tout nombre complexe $z = (x, y)$ différent de 0 admet un inverse donné par :

$$z' = \left(\frac{x}{x^2 + y^2}, -\frac{y}{x^2 + y^2} \right),$$

ce qui signifie que $zz' = z'z = 1$. En effet l'égalité $zz' = 1$ équivaut à :

$$\begin{cases} xx' - yy' = 1 \\ xy' + yx' = 0 \end{cases}$$

ce qui entraîne :

$$\begin{cases} x^2x' - xyy' = x \\ yxy' + y^2x' = 0 \end{cases} \text{ et } \begin{cases} yxx' - y^2y' = y \\ x^2y' + xyx' = 0 \end{cases}$$

et en additionnant les deux premières égalités [resp. en soustrayant les deux dernières], on obtient $(x^2+y^2)\,x' = x$, $(x^2+y^2)\,y' = -y$, ce qui donne compte tenu de $x^2+y^2 \neq 0$ pour $z \neq 0$, $x' = \dfrac{x}{x^2+y^2}$ et $y' = -\dfrac{y}{x^2+y^2}$. Réciproquement, on vérifie facilement que cette solution convient. On note z^{-1} ou $\dfrac{1}{z}$ cet inverse.

On peut remarquer qu'on a utilisé ici la commutativité du produit des réels.

Tout cela se traduit en disant que $(\mathbb{C}, +, \cdot)$ est un corps commutatif comme $(\mathbb{R}, +, \cdot)$, $(\mathbb{Q}, +, \cdot)$.

La notion de corps est étudiée plus en détails au chapitre suivant.

Là encore le neutre et l'inverse sont uniques et tout nombre complexe non nul est simplifiable pour le produit.

On note $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$ et ce qui précède nous dit que $(\mathbb{C}^*, \cdot)$ est un groupe commutatif.

On peut remarquer que pour tout réel non nul x, on a bien :

$$\frac{1}{x} = \left(\frac{1}{x}, 0\right) = \frac{1}{(x,0)}$$

Ces opérations d'addition et multiplication prolongent bien celles de $\mathbb{R}$.

L'associativité de la multiplication permet de définir les puissances n-ièmes d'un nombre complexe z par :

$$\begin{cases} z^0 = 1 \\ \forall n \in \mathbb{N}^*,\ z^{n+1} = z^n \cdot z \end{cases}$$

Pour $z \neq 0$ et $n \in \mathbb{N}^*$, on a $z^n \neq 0$ et :

$$(z^n)^{-1} = \frac{1}{z^n} = \left(z^{-1}\right)^n$$

On note alors $z^{-n} = \dfrac{1}{z^n}$.

Comme sur $\mathbb{R}$, on a $z^{p+q} = z^p z^q$ et $(z^p)^q = z^{pq}$ pour tous entiers relatifs p et q.

Comme sur $\mathbb{R}$, une égalité $zz' = 0$ équivaut à $z = 0$ ou $z' = 0$. En effet si $z \neq 0$, il admet un inverse et $0 = z^{-1} \cdot 0 = z^{-1}zz' = z'$.

En posant $i = (0,1)$, on a :

$$i^2 = (0,1)\,(0,1) = (-1,0) = -1$$

De cette égalité, on déduit que $\dfrac{1}{i} = -i$.

Le nombre complexe $-i$ est aussi solution de l'équation $x^2+1 = 0$ et i, $-i$ sont les seules solutions de cette équation. En effet si $\alpha^2+1 = 0$, on a $\alpha^2 = -1 = i^2$ et $\alpha^2 - i^2 = (\alpha - i)\,(\alpha + i) = 0$ de sorte que $\alpha = i$ ou $\alpha = -i$.

Théorème 2.1 *Tout nombre complexe s'écrit de manière unique $z = x + iy$, où x et y sont deux réels.*

Démonstration. Un nombre complexe s'écrit de manière unique :

$$\begin{aligned} z = (x,y) &= (x,0) + (0,y) \\ &= (x,0)\,(1,0) + (y,0)\,(0,1) \\ &= x \cdot 1 + y \cdot i = x + iy \end{aligned}$$

■

Définition 2.1 *Avec les notations du théorème qui précède, on dit que x est la partie réelle de z et y sa partie imaginaire, ce qui se note $x = \Re(z)$ et $y = \Im(z)$.*

Définition 2.2 *On dit qu'un nombre complexe est un imaginaire pur si sa partie réelle est nulle.*

En résumé un nombre complexe s'écrit $z = x+iy$ où $i^2 = -1$ et pour $z = x+iy$, $z' = x'+iy'$ dans $\mathbb{C}$, on a :

$$\begin{cases} z = z' \Leftrightarrow x = x' \text{ et } y = y' \\ z \in \mathbb{R} \Leftrightarrow y = \Im(z) = 0 \\ z + z' = (x + x') + i(y + y') \\ zz' = (xx' - yy') + i(xy' + x'y) \\ \dfrac{1}{z} = \dfrac{x}{x^2+y^2} - \dfrac{y}{x^2+y^2} i \text{ si } z \neq 0 \end{cases}$$

Exercice 2.1 *Écrire sous la forme $x + iy$ les nombres complexes suivants :*

$$u = \left(\frac{\sqrt{3}-i}{1+i\sqrt{3}}\right)^{27}, \ \left(\frac{\sqrt{3}+i}{\sqrt{3}-i} + \frac{\sqrt{3}-i}{\sqrt{3}+i} - 1\right)^{111}, \ w = \frac{\left(1+i\sqrt{3}\right)^4}{(1+i)^3}.$$

Solution 2.1 *On a* $\dfrac{\sqrt{3}-i}{1+i\sqrt{3}} = \dfrac{\sqrt{3}-i}{i\left(\sqrt{3}-i\right)} = \dfrac{1}{i} = -i$ *et* $u = -i^{27} = -\left(i^4\right)^7 \dfrac{1}{i} = i.$

On a

$$\begin{aligned} \frac{\sqrt{3}+i}{\sqrt{3}-i} + \frac{\sqrt{3}-i}{\sqrt{3}+i} &= \frac{\left(\sqrt{3}+i\right)^2 + \left(\sqrt{3}-i\right)^2}{3-i^2} \\ &= \frac{2\left(3+i^2\right)}{3-i^2} = 1 \end{aligned}$$

et $z = 0$.
On a :

$$\left(1+i\sqrt{3}\right)^4 = -8\left(1+i\sqrt{3}\right) \text{ et } (1+i)^3 = -2(1-i)$$

et :

$$w = 4\frac{1+i\sqrt{3}}{1-i} = 4\frac{\left(1+i\sqrt{3}\right)(1+i)}{(1-i)(1+i)} = 2\left(1-\sqrt{3}\right) + 2i\left(1+\sqrt{3}\right)$$

Exercice 2.2 *Calculer i^n pour tout entier relatif n.*

Solution 2.2 *Pour $n = 0$, on a $i^0 = 1$.*
Tout entier relatif s'écrit $n = 4q + r$ avec $r = 0, 1, 2$ ou 3 et :

$$i^n = \left(i^4\right)^q i^r = i^r = \begin{cases} 1 \text{ si } r = 0 \\ i \text{ si } r = 1 \\ -1 \text{ si } r = 2 \\ -i \text{ si } r = 3 \end{cases}$$

Exercice 2.3 *Montrer qu'il n'existe pas de relation d'ordre sur $\mathbb{C}$ qui prolonge la relation $\leq$ de $\mathbb{R}$ et qui soit compatible avec la somme et le produit, c'est à dire telle que $a \leq b$ et $c \leq d$ entraîne $a + c \leq b + d$ et $a \leq b$ et $0 \leq c$ entraîne $ac \leq bc$.*

Solution 2.3 *Supposant qu'une telle relation existe. Si* $0 \leq i$ *[resp.* $0 \leq -i$*], alors* $0 \leq i^2 \leq -1$ *[resp.* $0 \leq (-i)^2 = (-1)^2 i^2 = -1$*] dans* $\mathbb{R}$*, ce qui est incompatible avec la relation d'ordre sur* $\mathbb{R}$*.*

Exercice 2.4 *Soit* $j = -\dfrac{1}{2} + i\dfrac{\sqrt{3}}{2}$*. Calculer* j^2*,* j^3*,* $1 + j + j^2$ *et* j^n *pour tout entier relatif* n*.*

Solution 2.4 *On a* $j^0 = 1$*,* $j^1 = j$*,* $j^2 = -\dfrac{1}{2} - i\dfrac{\sqrt{3}}{2}$*,* $j^3 = 1$ *et* $1 + j + j^2 = 0$*. En écrivant* n *sous la forme* $n = 3q + r$ *avec* $r = 0, 1$ *ou* 2*, on a :*

$$j^n = j^r = \begin{cases} 1 \ si\ r = 0 \\ j \ si\ r = 1 \\ j^2 \ si\ r = 2 \end{cases}$$

Exercice 2.5 *Calculer* $(1+i)^2$*,* $(1+i)^6$ *et* $(1+i)^7$*. En utilisant la formule du binôme de Newton, en déduire les valeurs de* $1 - C_7^2 + C_7^4 - C_7^6$ *et* $C_7^1 - C_7^3 + C_7^5 - 1$*.*

Solution 2.5 *On a :*

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

donc :

$$(1+i)^6 = (2i)^3 = -8i$$

et :

$$(1+i)^7 = -8i\,(1+i) = 8 - 8i.$$

En utilisant la formule du binôme de Newton, on a aussi :

$$(1+i)^7 = \sum_{k=0}^{7} C_7^k i^k = \left(1 - C_7^2 + C_7^4 - C_7^6\right) + \left(C_7^1 - C_7^3 + C_7^5 - 1\right) i$$

ce qui nous donne $1 - C_7^2 + C_7^4 - C_7^6 = 8$ *et* $C_7^1 - C_7^3 + C_7^5 - 1 = -8$*.*

Exercice 2.6 *Calculer* $(1+i)^n$ *pour tout entier naturel* n*. En déduire les valeurs des sommes* $\sum_{j=0}^{p} C_{2p}^{2j} (-1)^j$ *et* $\sum_{j=0}^{p-1} C_{2p}^{2j+1} (-1)^j$ *pour tout entier naturel non nul* p*.*

Solution 2.6 *On a* $(1+i)^2 = 2i$*, donc* $(1+i)^{2p} = 2^p i^p$ *pour tout* $p \geq 0$ *et on connaît les* i^p*. Pour les entiers impairs, on a :*

$$(1+i)^{2p+1} = 2^p\,(1+i)\,i^p = 2^p\left(i^p + i^{p+1}\right).$$

On a donc :

$$(1+i)^{2p} = 2^p i^p = \begin{cases} 2^p \ si\ p = 4q \\ 2^p i \ si\ p = 4q+1 \\ -2^p \ si\ p = 4q+2 \\ -2^p i \ si\ p = 4q+3 \end{cases}$$

et :

$$(1+i)^{2p+1} = 2^p\left(i^p + i^{p+1}\right) = \begin{cases} 2^p\,(1+i) \ \ si\ p = 4q \\ 2^p\,(-1+i) \ \ si\ p = 4q+1 \\ -2^p\,(1+i) \ \ si\ p = 4q+2 \\ 2^p\,(1-i) \ \ si\ p = 4q+3 \end{cases}$$

En utilisant la formule du binôme de Newton, on a :

$$(1+i)^{2p} = \sum_{k=0}^{2p} C_{2p}^{k} i^{k} = \sum_{j=0}^{p} C_{2p}^{2j} i^{2j} + \sum_{j=0}^{p-1} C_{2p}^{2j+1} i^{2j+1}$$
$$= \sum_{j=0}^{p} C_{2p}^{2j} (-1)^{j} + i \sum_{j=0}^{p-1} C_{2p}^{2j+1} (-1)^{j}$$

En identifiant les parties réelles et imaginaires, on en déduit que :

$$\sum_{j=0}^{4q} C_{8q}^{2j} (-1)^{j} = 2^{4q} \text{ et } \sum_{j=0}^{4q-1} C_{8q}^{2j+1} (-1)^{j} = 0$$

$$\sum_{j=0}^{4q+1} C_{8q+2}^{2j} (-1)^{j} = 0 \text{ et } \sum_{j=0}^{4q} C_{8q+2}^{2j+1} (-1)^{j} = 2^{4q+1}$$

$$\sum_{j=0}^{4q+2} C_{8q+4}^{2j} (-1)^{j} = -2^{4q+2} \text{ et } \sum_{j=0}^{4q+1} C_{8q+4}^{2j+1} (-1)^{j} = 0$$

$$\sum_{j=0}^{4q+3} C_{8q+6}^{2j} (-1)^{j} = 0 \text{ et } \sum_{j=0}^{4q+2} C_{8q+6}^{2j+1} (-1)^{j} = -2^{4q+3}$$

Par exemple :

$$\sum_{j=0}^{4} C_{8}^{2j} (-1)^{j} = C_{8}^{4} - C_{8}^{2} - C_{8}^{6} + C_{8}^{8} + 1 = 2^{4} = 16$$

et :

$$\sum_{j=0}^{3} C_{8}^{2j+1} (-1)^{j} = C_{8}^{1} - C_{8}^{3} + C_{8}^{5} - C_{8}^{7} = 0$$

2.3 Conjugué et module d'un nombre complexe

Définition 2.3 *Le conjugué du nombre complexe $z = x+iy$ est le nombre complexe $\overline{z} = x-iy$.*

On déduit immédiatement de cette définition les propriétés suivantes du conjugué.

Théorème 2.2 *Pour tous nombres complexes z et z', on a :*

1. $\overline{\overline{z}} = z$.
2. $\overline{z+z'} = \overline{z} + \overline{z'}$.
3. $\overline{zz'} = \overline{z}\,\overline{z'}$.
4. $z + \overline{z} = 2\Re(z)$.
5. $z - \overline{z} = 2i\Im(z)$.
6. $z \in \mathbb{R} \Leftrightarrow z = \overline{z}$.
7. z *est imaginaire pur si, et seulement si,* $\overline{z} = -z$.
8. $z\overline{z} = (\Re(z))^2 + (\Im(z))^2 \in \mathbb{R}^+$.

Le dernier point du théorème précédent nous permet de donner la définition suivante.

Définition 2.4 *Le module du nombre complexe $z = x + iy$ est le réel :*

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

Dans le cas où z est réel, $|z|$ est la valeur absolue de z.
On vérifie facilement les propriétés suivantes liées au module.

Théorème 2.3 *Pour tous nombres complexes z et z', on a :*

1. $|z| = |\overline{z}|$
2. $z = 0$ *si, et seulement si,* $|z| = 0$
3. *si* $z \neq 0$, *alors* $\dfrac{1}{z} = \dfrac{\overline{z}}{|z|^2}$
4. $|z| = 1$ *si, et seulement si,* $z^{-1} = \overline{z}$.
5. $|zz'| = |z|\,|z'|$
6. *si* $z' \neq 0$, *alors* $\left|\dfrac{z}{z'}\right| = \dfrac{|z|}{|z'|}$
7. $|\Re(z)| \leq |z|\,,\ |\Im(z)| \leq |z|$
8. $|z + z'| \leq |z| + |z'|$
9. $||z| - |z'|| \leq |z - z'|$

Pour ce qui est du module d'un produit ou d'une somme de nombres complexes, on a de manière plus générale les résultats suivants qui se montre facilement par récurrence à partir des précédents.

Théorème 2.4 *Pour toute suite finie $z_1, \cdots, z_n$ de nombres complexes, on a :*

$$\begin{cases} \left|\displaystyle\prod_{k=1}^{n} z_k\right| = \displaystyle\prod_{k=1}^{n} |z_k| \\ \left|\displaystyle\sum_{k=1}^{n} z_k\right| \leq \displaystyle\sum_{k=1}^{n} |z_k| \end{cases}$$

En particulier, pour tout nombre complexe z et tout entier naturel n, on a $|z^n| = |z|^n$. Cette égalité étant encore valable pour n entier relatif et z non nul.

Exercice 2.7 *Montrer que si z est un nombre complexe de module égal à 1, alors $u = i\dfrac{1+z}{1-z}$ est réel.*

Solution 2.7 *Comme z est de module 1, on a $\overline{z} = \dfrac{1}{z}$ et :*

$$\overline{u} = -i\frac{1+\overline{z}}{1-\overline{z}} = -i\frac{1+\frac{1}{z}}{1-\frac{1}{z}} = -i\frac{z+1}{z-1} = u,$$

ce qui prouve que u est réel.

Exercice 2.8 *Montrer que si z et z' sont deux nombres complexes de module égal à 1 tels que $zz' \neq -1$, alors $u = \dfrac{z+z'}{1+zz'}$ est réel.*

Solution 2.8 *Comme z et z' sont de module 1, on a :*

$$\overline{u} = \frac{\overline{z} + \overline{z'}}{1 + \overline{z}\overline{z'}} = \frac{\frac{1}{z} + \frac{1}{z'}}{1 + \frac{1}{zz'}} = \frac{z + z'}{1 + zz'} = u,$$

ce qui prouve que u est réel.

Exercice 2.9 *Soient z, z' deux nombres complexes avec $z' \neq -1$. À quelle condition le nombre complexe $u = \dfrac{z + z'\overline{z}}{1 + z'}$ est-il réel ?*

Solution 2.9 *Dire que u est réel équivaut à dire que :*

$$\overline{u} = \frac{\overline{z} + \overline{z'}z}{1 + \overline{z'}} = \frac{z + z'\overline{z}}{1 + z'}$$

ce qui est encore équivalent à :

$$\left(\overline{z} + \overline{z'}z\right)(1 + z') = (z + z'\overline{z})\left(1 + \overline{z'}\right)$$

ou encore à :

$$\overline{z} + z'\overline{z'}z = z + z'\overline{z'}\overline{z}$$

soit à :

$$\overline{z}\left(1 - |z'|^2\right) = z\left(1 - |z'|^2\right).$$

En définitive, u est réel si, et seulement si, $|z'| = 1$ avec $z' \neq -1$ ou $z = \overline{z}$, ce qui signifie que z est réel.

Exercice 2.10 *Montrer que le produit de deux entiers naturels qui sont somme de deux carrés d'entiers est encore somme de deux carrés d'entiers.*

Solution 2.10 *Soient $n = a^2 + b^2$ et $m = c^2 + d^2$ où a, b, c, d sont des entiers relatifs. En écrivant que $n = |u|^2$ et $m = |v|^2$ où, $u = a + ib$ et $v = c + id$, on a :*

$$\begin{aligned} nm = |uv|^2 &= |(ac - bd) + (ad + bc)\,i|^2 \\ &= (ac - bd)^2 + (ad + bc)^2 \end{aligned}$$

(identité de Lagrange), c'est-à-dire que nm est somme de deux carrés d'entiers.

Exercice 2.11 *Soient $z_1, \cdots, z_n$ des nombres complexes non nuls deux à deux distincts et tels que $\sum_{k=1}^{n} z_k = 0$. Montrer qu'il existe deux indices $j \neq k$ compris entre 1 et n tels que $1 \leq \dfrac{|z_j|}{|z_k|} \leq 2$.*

Solution 2.11 *Quitte à réordonner, on peut supposer que :*

$$|z_1| \leq \cdots \leq |z_n|$$

En supposant que le résultat annoncé est faux, on a $\dfrac{|z_k|}{|z_{k-1}|} \notin [1, 2]$ pour tout k compris entre 2 et n et comme $\dfrac{|z_k|}{|z_{k-1}|} \geq 1$, on a nécessairement $\dfrac{|z_k|}{|z_{k-1}|} > 2$ pour tout k compris entre 2 et n et :

$$|z_n| > 2\,|z_{n-1}| > 2^2\,|z_{n-2}| > \cdots > 2^{n-1}\,|z_1|.$$

Mais l'hypothèse $\sum_{k=1}^{n} z_k = 0$ *nous donne :*

$$|z_n| = \left|\sum_{k=1}^{n-1} z_k\right| \leq \sum_{k=1}^{n-1} |z_k| < |z_n| \sum_{k=1}^{n-1} \frac{1}{2^k}$$

soit :

$$\sum_{k=1}^{n-1} \frac{1}{2^k} = \frac{1}{2} \frac{1 - \frac{1}{2^{n-1}}}{1 - \frac{1}{2}} = 1 - \frac{1}{2^{n-1}} > 1$$

ce qui est impossible.

Exercice 2.12 *Déterminer tous les nombres complexes* a *et* b*, tels que la fonction* $f : z \mapsto az + b\overline{z}$ *soit involutive (i.e. telle que* $f \circ f = Id_{\mathbb{C}}$*).*

Solution 2.12 *Dire que* f *est involutive équivaut à dire que pour tout nombre complexe* z*, on a :*

$$a\left(az + b\overline{z}\right) + b\left(\overline{az} + \overline{b}z\right) = z$$

ce qui est encore équivalent à :

$$\left(a^2 + |b|^2 - 1\right) z + b\left(a + \overline{a}\right)\overline{z} = 0 \tag{2.1}$$

Prenant respectivement $z = 1$ *et* $z = i$*, on aboutit à :*

$$\begin{cases} \left(a^2 + |b|^2 - 1\right) + b\left(a + \overline{a}\right) = 0 \\ \left(a^2 + |b|^2 - 1\right) - b\left(a + \overline{a}\right) = 0 \end{cases}$$

ce qui donne $a^2 + |b|^2 - 1 = 0$ *par addition et* $b\left(a + \overline{a}\right) = 0$ *par soustraction.*
Pour $b = 0$*, on a* $a^2 = 1$ *et* $a = \pm 1$*.*
Pour $b \neq 0$*, on a* $\overline{a} = -a$*, ce qui signifie que* a *est imaginaire pur, soit* $a = i\alpha$ *avec* α *réel et* $|b|^2 = 1 + \alpha^2 \geq 1$*, ce qui impose* $|b| \geq 1$ *et* $\alpha = \pm\sqrt{|b|^2 - 1}$*.*
Réciproquement si b *est un nombre complexe de module supérieur ou égal à* 1 *et* $a = \pm i\sqrt{|b|^2 - 1}$*, on a alors* $a^2 + |b|^2 = 1$ *et* $a + \overline{a} = 0$*, ce qui entraîne* (2.1) *pour tout nombre complexe* z*.*

Exercice 2.13 *Déterminer l'ensemble des nombres complexes* z *tels que* $u = (z - 1)(\overline{z} - i)$ *soit réel [resp. imaginaire pur].*

Solution 2.13 *En écrivant* $z = x + iy$*, on a*

$$\begin{aligned} u &= (x - 1 + iy)(x - i(y + 1)) \\ &= x^2 + y^2 + y - x + i(1 + y - x) \end{aligned}$$

et u *est réel [resp. imaginaire pur] si, et seulement si,* (x, y) *appartient à la droite d'équation* $y = x - 1$ *[resp. au cercle d'équation* $x^2 + y^2 + y - x = 0$*].*

Exercice 2.14 *Déterminer l'ensemble des nombres complexes* z *tels que* $|z - i| = |z - iz| = |z - 1|$*.*

Solution 2.14 *On note $z = x + iy$ un nombre complexe.*
L'égalité $|z - i| = |z - iz|$ équivaut à $x^2 + (y-1)^2 = (x+y)^2 + (y-x)^2$, ou encore à :

$$x^2 - 2y + y^2 + 1 = 2x^2 + 2y^2$$

soit à :

$$x^2 + y^2 + 2y - 1 = 0. \tag{2.2}$$

L'égalité $|z - i| = |z - 1|$ équivaut à $x^2 + (y-1)^2 = (x-1)^2 + y^2$, encore équivalent à $x = y$.
L'équation (2.2) nous dit alors que x est nécessairement solution de :

$$2x^2 + 2x - 1 = 0$$

ce qui donne deux solutions possibles :

$$z_1 = -\frac{1+\sqrt{3}}{2}(1+i) \text{ et } z_2 = \frac{\sqrt{3}-1}{2}(1+i)$$

Réciproquement, on vérifie que ces solutions conviennent bien.
Géométriquement, l'ensemble des points cherché est l'intersection du cercle $\mathcal{C}$ d'équation (2.2) et de la droite $\mathcal{D}$ d'équation $y = x$ (figure 2.1).

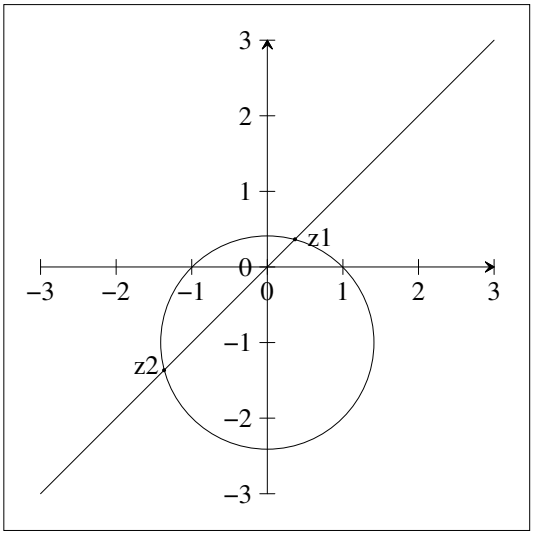

FIG. 2.1 – $\mathcal{C} \cap \mathcal{D}$

2.4 Les équations de degré 2

On s'intéresse ici aux équations algébriques de degré 2 à coefficients complexes, c'est-à-dire aux équations de la forme $ax^2 + bx + c = 0$ où a, b, c sont des nombres complexes avec $a \neq 0$.

Si on se place dans le cadre réel (i. e. les coefficients a, b, c sont réels et on cherche des solutions réelles), on sait qu'une telle équation n'a pas nécessairement de solution (c'est l'exemple de $x^2 + 1 = 0$ qui nous a conduit aux nombres complexes).

En divisant par a, on se ramène au cas où $a = 1$.

On remarque tout d'abord qu'une telle équation a au plus deux solutions complexes. En effet, si $z_1 \in \mathbb{C}$ est solution de $x^2 + bx + c = 0$, en désignant par z une autre solution, on a le système d'équations :

$$\begin{cases} z^2 + bz + c = 0 \\ z_1^2 + zx_1 + c = 0 \end{cases}$$

et par soustraction on aboutit à :

$$(z - z_1)(z + z_1 + b) = 0$$

qui donne $z = x_1$ ou $z = -z_1 - b$.

Il suffit donc de trouver une solution (s'il en existe) de cette équation pour avoir les deux.

Nous allons voir que sur $\mathbb{C}$ une équation de degré 2 a toujours deux solutions, distinctes ou confondues.

Connaissant $i \in \mathbb{C}$ solution de $x^2 + 1 = 0$, on déduit que pour tout réel a non nul l'équation $x^2 + a = 0$ a exactement deux solutions distinctes. En effet si a est négatif, cette équation équivaut à $x^2 = -a$ avec $-a > 0$, et dans l'ensemble des réels, on sait que cela équivaut à dire que $x = \pm\sqrt{-a}$, ce qui fourni deux solutions réelles distinctes. Si a est positif, cette équation équivaut à $x^2 = -a = i^2\left(\sqrt{a}\right)^2$, soit à $x^2 - \left(i\sqrt{a}\right)^2 = 0$ ce qui équivaut à $x = \pm i\sqrt{a}$.

On a donc le résultat suivant.

Théorème 2.5 *Pour tout nombre réel non nul a l'équation $x^2 + a = 0$ a exactement deux solutions données par :*

$$\begin{cases} x_1 = -\sqrt{-a} \text{ et } x_2 = \sqrt{-a} \text{ si } a < 0 \\ x_1 = -i\sqrt{a} \text{ et } x_2 = i\sqrt{a} \text{ si } a > 0 \end{cases}$$

Pour $a = 0$, $x = 0$ est la seule solution de cette équation.

Définition 2.5 *Si α est un nombre complexe, on dit que le nombre complexe u est une racine carrée de α si $u^2 = \alpha$.*

Le théorème précédent nous dit que tout nombre réel non nul a a exactement deux racines carrées complexes, ce sont les réels $\pm\sqrt{a}$ pour $a > 0$ et les complexes $\pm i\sqrt{-a}$ pour $a < 0$.

Ce résultat est en fait valable pour tout nombre complexe α.

Théorème 2.6 *Tout nombre complexe non nul $\alpha = a + ib$ a exactement deux racines carrées.*

Démonstration. Il s'agit de résoudre l'équation $z^2 = \alpha$ et pour ce faire il nous suffit de trouver une solution.

Si $b = 0$, alors $\alpha = a$ est réel et le problème a été résolu.

On suppose donc que $b \neq 0$.

En notant $z = x + iy$, l'équation $z^2 = \alpha = a + ib$ équivaut au système de deux équations aux inconnues x, y :

$$\begin{cases} x^2 - y^2 = a \\ 2xy = b \end{cases}$$

En utilisant le module de z, le système :

$$\begin{cases} x^2 - y^2 = a \\ |z^2| = |z|^2 = x^2 + y^2 = |\alpha| \end{cases}$$

nous donne immédiatement :

$$x^2 = \frac{a + |\alpha|}{2}$$

avec

$$a + |\alpha| = a + \sqrt{a^2 + b^2} > a + \sqrt{a^2} = a + |a| \geq 0$$

et en conséquence, $x = \pm\dfrac{\sqrt{a+|\alpha|}}{\sqrt{2}}$, la partie imaginaire y étant déterminée par l'équation $2xy = b$ avec $b \neq 0$ (donc $x \neq 0$ et $y \neq 0$). En définitive l'équation $z^2 = \alpha$ a deux solutions complexes données par :

$$z_1 = \sqrt{\frac{a+|\alpha|}{2}} + i\frac{b}{2\sqrt{\frac{a+|\alpha|}{2}}} = \frac{1}{\sqrt{2}}\frac{a+ib+|\alpha|}{\sqrt{a+|\alpha|}}$$
$$= \frac{1}{\sqrt{2}}\frac{\alpha+|\alpha|}{\sqrt{\Re(\alpha)+|\alpha|}} = \frac{1}{\sqrt{2}}\frac{a+ib+\sqrt{a^2+b^2}}{\sqrt{a+\sqrt{a^2+b^2}}}$$

et :

$$z_2 = -z_1$$

■

On en déduit alors le résultat suivant.

Théorème 2.7 *Toute équation de degré* 2, $az^2+bz+c=0$ *(*$a \neq 0$*) a deux solutions complexes distinctes ou confondues.*

Démonstration. En utilisant la forme réduite d'un polynôme de degré 2 :

$$az^2+bz+c = a\left(\left(z+\frac{b}{2a}\right)^2 - \frac{b^2-4ac}{4a^2}\right)$$

on est ramené à l'équation :

$$\left(z+\frac{b}{2a}\right)^2 = \frac{b^2-4ac}{4a^2}$$

qui a deux solutions distinctes (si $b^2-4ac \neq 0$) ou confondues (si $b^2-4ac=0$). ■

Avec les notations du théorème la quantité $\delta = b^2-4ac$ est appelé discriminant de l'équation $ax^2+bx+c=0$.

Pour $\delta = 0$, $z_1 = z_2 = -\dfrac{b}{2a}$ est la seule solution (on dit que c'est une racine double) et pour $\delta \neq 0$, les deux solutions sont $z_1 = \dfrac{-b-\gamma}{2a}$ et $z_1 = \dfrac{-b+\gamma}{2a}$ où γ est une racine carrée de δ (i.e. $\gamma^2 = \delta$).

Dans les deux cas, on $z_1+z_2 = -\dfrac{b}{a}$ et $z_1z_2 = \dfrac{b^2-\gamma^2}{4a^2} = \dfrac{b^2-\delta}{4a^2} = \dfrac{c}{a}$.

Réciproquement si z_1, z_2 sont deux nombres complexes tels que $z_1+z_2 = -\dfrac{b}{a}$ et $z_1z_2 = \dfrac{c}{a}$, on a $z_2 = -\dfrac{b}{a}-z_1$ et $z_1\left(-\dfrac{b}{a}-z_1\right) = \dfrac{c}{a}$, c'est-à-dire que z_1 est solution de l'équation $z^2+\dfrac{b}{a}z+\dfrac{c}{a}=0$, soit de l'équation $az^2+bz+c=0$. Comme z_1 et z_2 jouent des rôles symétriques, z_2 est également solution de cette équation. En résumé, on a le résultat suivant.

Théorème 2.8 *Étant donnés des nombres complexes* a, b, c *avec* $a \neq 0$*, les nombres complexes* z_1 *et* z_2 *sont les racines de l'équation* $az^2+bz+c=0$ *si, et seulement si,* $z_1+z_2 = -\dfrac{b}{a}$ *et* $z_1z_2 = \dfrac{c}{a}$.

Dans le cas où les coefficients a, b, c sont réels, il en est de même de δ et on distingue trois cas de figure :

- soit $\delta = 0$ et $x_1 = -\dfrac{b}{2a}$ est la seule solution réelle de cette équation ;
- soit $\delta > 0$ et $x_1 = \dfrac{-b-\sqrt{\delta}}{2a}$, $x_1 = \dfrac{-b+\sqrt{\delta}}{2a}$ sont les deux solutions réelles de cette équation ;
- soit $\delta < 0$ et $x_1 = \dfrac{-b-i\sqrt{-\delta}}{2a}$, $x_1 = \dfrac{-b+i\sqrt{-\delta}}{2a}$ sont les deux solutions complexes non réelles de cette équation.

Parfois le coefficient b s'écrit naturellement sous la forme $b = 2b'$ et on a :

$$\delta = 4\left((b')^2 - ac\right)$$

La quantité $\delta' = (b')^2 - ac$ est alors appelée discriminant réduit de l'équation $ax^2 + 2b'x + c = 0$. Dans le cas où les coefficients a, b, c sont réels, on a :

- soit $\delta' = 0$ et $x_1 = -\dfrac{b'}{a}$ est la seule solution réelle de cette équation ;
- soit $\delta' > 0$ et $x_1 = \dfrac{-b'-\sqrt{\delta'}}{a}$, $x_1 = \dfrac{-b'+\sqrt{\delta'}}{a}$ sont les deux solutions réelles de cette équation ;
- soit $\delta' < 0$ et $x_1 = \dfrac{-b'-i\sqrt{-\delta'}}{a}$, $x_1 = \dfrac{-b'+i\sqrt{-\delta'}}{a}$ sont les deux solutions complexes non réelles de cette équation.

De manière plus générale, on peut montrer le théorème suivant que nous admettrons.

Théorème 2.9 (d'Alembert-Gauss) *Toute équation polynomiale à coefficients complexes de degré non nul n admet n racines complexes distinctes ou confondues.*

On rappelle que si P est un polynôme non constant (i. e. de degré $n \geq 1$) à coefficients complexes [resp. réels], on dit que $\alpha \in \mathbb{C}$ [resp. $\alpha \in \mathbb{R}$] est racine de P si $P(\alpha) = 0$.

On peut donner plusieurs démonstrations du théorème de d'Alembert-Gauss mais toutes utilisent des outils d'algèbre ou d'analyse plus sophistiqués que le contenu de ce chapitre. Le problème du paragraphe **??** propose une démonstration classique qui utilise des outils d'analyse.

Il est par contre facile de montrer qu'un polynôme à coefficients complexes [resp. réels] de degré $n \geq 1$ a au plus n racines complexes[resp. réelles] distinctes ou confondues. En effet pour $n = 1$, l'unique racine du polynôme $az + b$ avec $a \neq 0$ est $z = -\dfrac{b}{a}$. En supposant le résultat acquis pour les polynômes de degré $n - 1 \geq 1$, on se donne un polynôme $P(z) = \sum_{k=0}^{n} a_k z^k$ de degré n (ce qui signifie que $a_n \neq 0$). S'il admet une racine α, on peut écrire, pour tout nombre complexe z :

$$P(z) = P(z) - P(\alpha) = \sum_{k=1}^{n} a_k \left(z^k - \alpha^k\right)$$

avec $z^k - \alpha^k = (z-\alpha)\sum_{j=1}^{k} z^{k-j}\alpha^{j-1}$ pour tout $k \geq 1$, ce qui donne $P(z) = (z-\alpha)\,Q(z)$, où Q est un polynôme de degré $n-1$, il a donc au plus $n-1$ racines et P a au plus n racines.

Une conséquence importante est que deux polynômes P et Q de degré $n \geq 1$ qui coïncident en $n+1$ points distincts sont nécessairement égaux. En effet $P - Q$ est nul ou de degré au plus n. S'il est non constant, il est degré $p \leq n$ avec $n + 1 > p$ racines, ce qui est impossible. Il est donc constant égal à $P(\alpha) - Q(\alpha) = 0$ où α est l'une des racines communes de P et Q.

Exercice 2.15 *Factoriser dans $\mathbb{R}$ puis dans $\mathbb{C}$, $x^4 + 1$.*

Solution 2.15 *Dans* $\mathbb{R}$ *on a :*

$$\begin{aligned} x^4+1 &= \left(x^2+1\right)^2-2x^2 \\ &= \left(x^2-\sqrt{2}x+1\right)\left(x^2+\sqrt{2}x+1\right) \end{aligned}$$

les deux polynômes $x^2 \pm \sqrt{2}x+1$ *de discriminant* $\delta=-2<0$ *étant sans racines réelles. Sur* $\mathbb{C}$ *ces polynômes ont pour racines* $\dfrac{\pm 1 \pm i1}{\sqrt{2}}$*, ce qui donne :*

$$x^4+1=\left(x-\frac{1-i}{\sqrt{2}}\right)\left(x-\frac{1+i}{\sqrt{2}}\right)\left(x+\frac{1+i}{\sqrt{2}}\right)\left(x+\frac{1-i}{\sqrt{2}}\right)$$

Exercice 2.16 *Déterminer les racines carrés complexes de* $\alpha=-7+24i$*.*

Solution 2.16 *En écrivant* $z=x+iy$*, l'équation* $z^2=\alpha$ *équivaut à :*

$$\begin{cases} x^2-y^2=-7 \\ xy=12 \end{cases}$$

Considérant que $|z|^2=x^2+y^2=|\alpha|=25$*, on déduit que* $x^2=9$*, donc* $x=3$ *et* $y=4$ *ou* $x=-3$ *et* $y=-4$*. Les deux racines carrés de* α *sont donc* $\pm(3+4i)$*.*

Exercice 2.17 *Résoudre dans* $\mathbb{C}$ *l'équation* $z^2-(1-2i)z+1-7i=0$*.*

Solution 2.17 *Cette équation est équivalente à :*

$$\left(z-\frac{1-2i}{2}\right)^2=\frac{(1-2i)^2}{4}-1+7i=\frac{-7+24i}{4}$$

soit à $Z^2=\alpha=-7+24i$*, où on a posé* $Z=2z-1+2i$*, ce qui donne* $Z=\pm(3+4i)$ *et* $z=\dfrac{1-2i}{2}\pm\dfrac{3+4i}{2}$*. Les deux solutions complexes de cette équation sont donc :*

$$z_1=2+i,\ z_2=-2-3i.$$

2.5 Les équations de degré 3 et 4

On s'intéresse tout d'abord aux équations polynomiales de degré 3 :

$$P(z)=z^3+az^2+bz+c=0$$

où a, b, c sont des nombres complexes.

Dans un premier temps, on effectue une translation en vue de supprimer le terme en z^2 de cette équation, c'est-à-dire qu'in cherche $\lambda\in\mathbb{C}$ qui permette de supprimer z^2 dans :

$$P(z-\lambda)=(z-\lambda)^3+a(z-\lambda)^2+b(z-\lambda)+c.$$

En développant, on a :

$$P(z-\lambda)=z^3+(a-3\lambda)z^2+\left(\lambda^2-2a\lambda+b\right)z+\left(c-b\lambda+a\lambda^2-\lambda^3\right).$$

Le choix de $\lambda = \dfrac{a}{3}$ donne :

$$P(z-\lambda) = z^3 + \left(b - \frac{a^2}{3}\right) z + \left(c - \frac{ab}{3} + \frac{2a^3}{27}\right).$$

On est donc ramené à l'équation :

$$Q(z) = z^3 + pz + q = 0$$

où on a noté $p = b - \dfrac{a^2}{3}$ et $q = c - \dfrac{ab}{3} + \dfrac{2a^3}{27}$. Si z est solution de $Q(z) = 0$, alors $z - \lambda$ est solution de $P(t) = 0$.

Si $p = 0$, alors les solutions de $Q(z) = 0$ sont les racines cubiques de $-q$.

Si $p \neq 0$, on cherche alors les solutions sous la forme $z = u + v$ en imposant une condition supplémentaire à u et v. En développant :

$$P(u+v) = u^3 + v^3 + (3uv + p)(u + v) + q$$

on est amené à imposer $3uv + p = 0$, ce qui donne le système de deux équations à deux inconnues :

$$\begin{cases} u^3 + v^3 = -q \\ 3uv = -p \end{cases}$$

Les nombres complexes u^3 et v^3 sont alors solutions de : :

$$\begin{cases} u^3 + v^3 = -q \\ u^3 v^3 = -\dfrac{p^3}{27} \end{cases}$$

ce qui revient à dire que ce sont les solutions de l'équation de degré 2 :

$$x^2 + qx - \frac{p^3}{27} = 0$$

Le discriminant de cette équation est :

$$\delta = \frac{4p^3 + 27q^2}{27}.$$

Notant ω une racine carrée de δ ($\omega^2 = \delta$), on a :

$$u^3 = \frac{-q - \omega}{2} \text{ et } v^3 = u^3 = \frac{-q + \omega}{2}$$

En désignant par w une racine cubique $\dfrac{-q-\omega}{2}$, les deux autres sont jw et $\overline{j}w$. Enfin la relation $3uv = -p$ avec $p \neq 0$, donne $u \neq 0$, $v \neq 0$ et $v = -\dfrac{p}{3u}$. On a donc ainsi trouvé trois solutions (u, v), à savoir :

$$\left(w, -\frac{p}{3w}\right),\ \left(jw, -\frac{p}{3jw}\right) = \left(jw, -\frac{p\overline{j}}{3w}\right) \text{ et } \left(\overline{j}w, -\frac{p}{3\overline{j}w}\right) = \left(\overline{j}w, -\frac{pj}{3w}\right)$$

ce qui donne trois solutions pour l'équation $Q(z) = 0$:

$$z_1 = w - \frac{p}{3w},\ z_2 = jw - \frac{p\overline{j}}{3w},\ z_3 = \overline{j}w - \frac{pj}{3w}$$

et on les a toutes.

Exercice 2.18 *Résoudre dans* $\mathbb{C}$ *l'équation :*

$$P(z) = z^3 - 3z^2 + 4z - 4 = 0$$

Solution 2.18 *On élimine tout d'abord le terme en* z^2*. On a :*

$$P(z-\lambda) = (z-\lambda)^3 - 3(z-\lambda)^2 + 4(z-\lambda) - 4$$
$$= z^3 - 3(\lambda+1)z^2 + \left(4 + 6\lambda + 3\lambda^2\right)z - \left(\lambda^3 + 3\lambda^2 + 4\lambda + 4\right)$$

et $\lambda = -1$ *donne :*

$$Q(z) = P(z+1) = z^3 + z - 2.$$

Cherchant les solutions sous la forme $z = u + v$*, on aboutit à :*

$$\begin{cases} u^3 + v^3 = 2 \\ u^3 v^3 = -\dfrac{1}{27} \end{cases}$$

qui nous conduit à résoudre :

$$x^2 - 2x - \frac{1}{27} = 0$$

de solutions :

$$u^3 = 1 + \frac{2\sqrt{7}}{3\sqrt{3}} \text{ et } v^3 = 1 - \frac{2\sqrt{7}}{3\sqrt{3}}.$$

Ce qui donne :

$$u \in \left\{ \sqrt[3]{1 + \frac{2\sqrt{7}}{3\sqrt{3}}}, j\sqrt[3]{1 + \frac{2\sqrt{7}}{3\sqrt{3}}}, \overline{j}\sqrt[3]{1 + \frac{2\sqrt{7}}{3\sqrt{3}}} \right\}$$

et $v = -\dfrac{1}{3u}$ *avec :*

$$\frac{1}{u} = \frac{1}{\sqrt[3]{1 + \dfrac{2\sqrt{7}}{3\sqrt{3}}}} = \frac{\sqrt[3]{\dfrac{2\sqrt{7}}{3\sqrt{3}} - 1}}{\sqrt[3]{\dfrac{28}{27} - 1}} = 3\sqrt[3]{\frac{2\sqrt{7}}{3\sqrt{3}} - 1}.$$

D'où les solutions de $Q(z) = 0$ *:*

$$z_1 = \sqrt[3]{\frac{2\sqrt{7}}{3\sqrt{3}} + 1} - \sqrt[3]{\frac{2\sqrt{7}}{3\sqrt{3}} - 1}$$
$$z_2 = j\sqrt[3]{\frac{2\sqrt{7}}{3\sqrt{3}} + 1} - \overline{j}\sqrt[3]{\frac{2\sqrt{7}}{3\sqrt{3}} - 1}$$
$$z_2 = \overline{j}\sqrt[3]{\frac{2\sqrt{7}}{3\sqrt{3}} + 1} - j\sqrt[3]{\frac{2\sqrt{7}}{3\sqrt{3}} - 1}$$

Comme 1 *est racine évidente de* $Q(z) = 1$ *et que* z_1 *est la seule solution réelle, on a nécessairement :*

$$\sqrt[3]{\frac{2\sqrt{7}}{3\sqrt{3}} + 1} - \sqrt[3]{\frac{2\sqrt{7}}{3\sqrt{3}} - 1} = 1$$

ce qui peut se vérifier en élevant au carré.
Les solutions de $P(z) = 0$ *sont alors :*

$$z_1 + 1 = 2,\ z_2 + 1,\ z_3 + 1.$$

2.6 Arguments d'un nombre complexe

On suppose connues du cours d'analyse les fonctions trigonométriques cos, sin et tan avec leurs principales propriétés. En particulier, les fonctions cos et sin sont définies sur $\mathbb{R}$, périodiques de période 2π, la fonction cos est paire, la fonction sin est impaire et on a les formules de trigonométrie suivantes valables pour tous réels a, b :

$$\begin{cases} \cos(a+b) = \cos(a)\cos(b) - \sin(a)\sin(b) \\ \sin(a+b) = \sin(a)\cos(b) + \cos(a)\sin(b) \end{cases}$$

desquelles on déduit les suivantes bien utiles :

$$\begin{cases} \cos(a-b) = \cos(a)\cos(b) + \sin(a)\sin(b) \\ \sin(a-b) = \sin(a)\cos(b) - \cos(a)\sin(b) \\ \cos(2a) = \cos^2(a) - \sin^2(a) \\ \sin(2a) = 2\sin(a)\cos(b) \\ \cos(a)\cos(b) = \dfrac{\cos(a+b) + \cos(a-b)}{2} \\ \sin(a)\sin(b) = \dfrac{\cos(a-b) - \cos(a+b)}{2} \\ \cos(a)\sin(b) = \dfrac{\sin(a+b) - \sin(a-b)}{2} \\ \sin(a)\cos(b) = \dfrac{\sin(a+b) + \sin(a-b)}{2} \end{cases}$$

enfin avec $\cos(0) = 1$, on déduit que :

$$\cos^2(a) + \sin^2(a) = 1.$$

La fonction tan est définie sur $\mathbb{R} \setminus \left\{ \dfrac{\pi}{2} + k\pi \mid k \in \mathbb{Z} \right\}$ par $\tan(x) = \dfrac{\sin(x)}{\cos(x)}$ et elle est impaire et π-périodique.

La fonction cos réalise une bijection de $[0, \pi]$ sur $[-1, 1]$, la fonction sin une bijection de $\left[-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right]$ sur $[-1, 1]$ et la fonction tan une bijection de $\left]-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right[$ sur $\mathbb{R}$. Les fonctions réciproques de ces fonctions trigonométriques sont notées respectivement arccos, arcsin et arctan. On a donc :

$$(x \in [0, \pi] \text{ et } y = \cos(x)) \Leftrightarrow (y \in [-1, 1] \text{ et } x = \arccos(y))$$

$$\left(x \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \text{ et } y = \sin(x)\right) \Leftrightarrow (y \in [-1, 1] \text{ et } x = \arcsin(y))$$

$$\left(x \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[\text{ et } y = \tan(x)\right) \Leftrightarrow (y \in \mathbb{R} \text{ et } x = \arctan(y))$$

Exercice 2.19 *Montrer que pour tout réel x on a :*

$$\sin\left(\frac{x}{2}\right)\left(\frac{1}{2} + \sum_{k=1}^{n} \cos(kx)\right) = \frac{1}{2}\sin\left(\frac{2n+1}{2}x\right).$$

Solution 2.19 *Pour tout entier naturel k et pour tout réel u on a :*

$$\sin\left(\frac{u}{2}\right)\cos(ku) = \frac{1}{2}\left(\sin\left(\left(k + \frac{1}{2}\right)u\right) - \sin\left(\left(k - \frac{1}{2}\right)u\right)\right),$$

on en déduit alors que pour tout réel u on a :

$$\sin\left(\tfrac{u}{2}\right)\left(\tfrac{1}{2}+\sum_{k=1}^{n}\cos\left(ku\right)\right)=\tfrac{1}{2}\left(\sin\left(\tfrac{u}{2}\right)+\sum_{k=1}^{n}\left(\sin\left(\tfrac{2k+1}{2}u\right)-\sin\left(\tfrac{2k-1}{2}u\right)\right)\right)$$
$$=\tfrac{1}{2}\sin\left(\tfrac{2n+1}{2}u\right)$$

Exercice 2.20 *Montrer que pour tout réel x on a :*

$$\sin\left(\frac{x}{2}\right)\left(\sum_{k=0}^{n-1}\sin\left(\left(2k+1\right)\frac{x}{2}\right)\right)=\sin^2\left(\frac{n}{2}x\right).$$

Solution 2.20 *Pour tout entier naturel k et pour tout réel u on a :*

$$\sin\left(\frac{u}{2}\right)\sin\left(\left(2k+1\right)\frac{u}{2}\right)=\frac{1}{2}\left(\cos\left(ku\right)-\cos\left(\left(k+1\right)u\right)\right),$$

on en déduit alors que pour tout réel u on a :

$$\sin\left(\tfrac{u}{2}\right)\left(\sum_{k=0}^{n-1}\sin\left(\left(2k+1\right)\tfrac{u}{2}\right)\right)=\tfrac{1}{2}\left(\sum_{k=0}^{n-1}\left(\cos\left(ku\right)-\cos\left(\left(k+1\right)u\right)\right)\right)$$
$$=\tfrac{1}{2}\left(1-\cos\left(nu\right)\right)=\sin^2\left(\tfrac{n}{2}u\right).$$

Théorème 2.10 *Si z est un nombre complexe de module 1, il existe un unique réel $\theta\in[-\pi,\pi[$ tel que $z=\cos\left(\theta\right)+i\sin\left(\theta\right)$.*

Démonstration. Le nombre complexe $z=x+iy$ est de module 1 si, et seulement si $x^2+y^2=1$. En particulier x est dans $[-1,1]$ et il existe un unique réel $\alpha\in[0,\pi]$ tel que $x=\cos\left(\alpha\right)$. Avec $y^2=1-x^2=\sin^2\left(\alpha\right)$, on déduit que $y=\pm\sin\left(\alpha\right)$, soit $y=\sin\left(\pm\alpha\right)$. Avec la parité de la fonction cos, on peut écrire que $x=\cos\left(\pm\alpha\right)$ et on aboutit à $\left(x,y\right)=\left(\cos\left(\theta\right),\sin\left(\theta\right)\right)$ avec $\theta\in[-\pi,\pi[$ (pour $\left(x,y\right)=\left(\cos\left(\pi\right),\sin\left(\pi\right)\right)=\left(-1,0\right)$, on écrit $\left(x,y\right)=\left(\cos\left(-\pi\right),\sin\left(-\pi\right)\right)$).

Si $\theta'\in[-\pi,\pi[$ est une autre solution, de $\cos\left(\theta\right)=\cos\left(\theta'\right)$, on déduit que $\theta'=\pm\theta$. Si $\theta'=\theta$, c'est terminé, sinon $\theta'=-\theta$ et de $\sin\left(\theta\right)=\sin\left(\theta'\right)=-\sin\left(\theta\right)$, on déduit que θ vaut 0 ou $-\pi$, 0 étant la seule solution puisque $\theta'=\pi\notin[-\pi,\pi[$. D'où l'unicité. ■

Corollaire 2.1 *Pour tout nombre complexe non nul z, il existe un unique réel $\theta\in[-\pi,\pi[$ tel que $\dfrac{z}{|z|}=\cos\left(\theta\right)+i\sin\left(\theta\right)$.*

Démonstration. On applique le théorème précédent à $\dfrac{z}{|z|}$ qui est de module égal à 1. ■

Définition 2.6 *Avec les notations du corollaire qui précède, on dit que le réel $\theta\in[-\pi,\pi[$ est l'argument principal du nombre complexe non nul z.*

Si $\theta\in[-\pi,\pi[$ est l'argument principal d'un nombre complexe $z\in\mathbb{C}^*$, les seuls réels θ' tels que $\dfrac{z}{|z|}=\cos\left(\theta'\right)+i\sin\left(\theta'\right)$ sont les réels $\theta'=\theta+2k\pi$, où k est un entier relatif. En effet ces réels conviennent et les égalités $\cos\left(\theta\right)=\cos\left(\theta'\right)$ et $\sin\left(\theta\right)=\sin\left(\theta'\right)$ sont réalisées si, et seulement si il existe un entier relatif k tel que $\theta'=\theta+2k\pi$ (on peut trouver un entier k tel que $\theta'-2k\pi$ soit dans $[-\pi,\pi[$, c'est-à-dire que k est tel que $-\pi\leq\theta'-2k\pi<\pi$, soit $k\leq\dfrac{\theta'+\pi}{2\pi}<k+1$, encore équivalent à $k=\left[\dfrac{\theta'+\pi}{2\pi}\right]$ et $\theta'-2k\pi$ est l'argument principal de z).

Définition 2.7 *On dit qu'un réel θ est un argument du nombre complexe non nul z si* $\dfrac{z}{|z|} = \cos(\theta) + i\sin(\theta)$.

Ce qui précède nous dit qu'un nombre complexe non nul admet une infinité d'arguments et que deux tels arguments différent d'un multiple entier de 2π, on dit alors qu'ils sont égaux modulo 2π.

On notera $\theta' \equiv \theta \ (2\pi)$ pour signifier que les réels θ' et θ sont égaux modulo 2π.

Si θ est un argument d'un nombre complexe non nul z, on notera $\arg(z) \equiv \theta \ (2\pi)$. La notation $\arg(z)$ signifie qu'on a choisi un argument de z, c'est donc un réel défini modulo 2π.

Par abus de langage, on écrira $\theta = \arg(z)$ quand il n'y a pas d'ambiguïté.

Avec le théorème qui suit on donne quelques propriétés des arguments d'un nombre complexe.

Théorème 2.11 *En désignant par z et z' des nombres complexes non nuls, λ un réel non nul et n un entier relatif, on a :*

1. $\arg(\overline{z}) \equiv -\arg(z) \ (2\pi)$
2. $\arg(zz') \equiv \arg(z) + \arg(z') \ (2\pi)$
3. $\arg\left(\dfrac{z}{z'}\right) \equiv \arg(z) - \arg(z') = \arg\left(z\overline{z'}\right) \ (2\pi)$
4. $\arg(z^n) \equiv n\arg(z) \ (2\pi)$
5. *si* $\lambda > 0$, *alors* $\arg(\lambda z) \equiv \arg(z) \ (2\pi)$, *si* $\lambda < 0$, *alors* $\arg(\lambda z) \equiv \arg(z) + \pi \ (2\pi)$
6. *z est réel si, et seulement si* $\arg(z) \equiv 0 \ (\pi)$ *(c'est-à-dire que les arguments de z sont de la forme $k\pi$ avec $k \in \mathbb{Z}$, l'argument principal étant 0 ou $-\pi$).*
7. *z est imaginaire pur si, et seulement si* $\arg(z) \equiv \dfrac{\pi}{2} \ (\pi)$ *(c'est-à-dire que les arguments de z sont de la forme $\dfrac{\pi}{2} + k\pi$ avec $k \in \mathbb{Z}$, l'argument principal étant $-\dfrac{\pi}{2}$ ou $\dfrac{\pi}{2}$).*

Démonstration. Il suffit de considérer le cas des nombres complexes de module 1 par définition des arguments.

1. Pour $z = \cos(\theta) + i\sin(\theta)$, on a :
$$\overline{z} = \cos(\theta) - i\sin(\theta) = \cos(-\theta) + i\sin(-\theta)$$
et $-\theta$ est un argument de $\overline{z}$, donc $\arg(\overline{z}) \equiv -\theta \ (2\pi)$.
2. Pour $z = \cos(\theta) + i\sin(\theta)$ et $z' = \cos(\theta') + i\sin(\theta')$, on a :
$$\begin{aligned} zz' &= (\cos(\theta)\cos(\theta') - \sin(\theta)\sin(\theta')) + i(\sin(\theta)\cos(\theta') + \cos(\theta)\sin(\theta')) \\ &= \cos(\theta+\theta') + i\sin(\theta+\theta') \end{aligned}$$
et donc $\arg(zz') \equiv \theta + \theta' \ (2\pi)$.
3. On a
$$\begin{aligned} \arg\left(\frac{z}{z'}\right) \equiv \arg\left(z\overline{z'}\right) &\equiv \arg(z) + \arg\left(\overline{z'}\right) \\ &\equiv \arg(z) - \arg(z') \ (2\pi) \end{aligned}$$
4. Se déduit de ce qui précède.

■

En désignant par φ l'application qui associe à tout réel θ le nombre complexe $\varphi(\theta) = \cos(\theta) + i\sin(\theta)$ on réalise une application surjective de $\mathbb{R}$ sur l'ensemble Γ des nombres complexes de module 1. Cette application n'est pas injective puisque l'égalité $\varphi(\theta) = \varphi(\theta')$ équivaut à $\theta' \equiv \theta\ (2\pi)$. En restriction à $[-\pi, \pi[$ cette application φ est bijective.

Théorème 2.12 *Avec les notations qui précèdent, on a $\varphi(0) = 1$ et pour tous réels θ, θ' :*

$$\varphi(\theta + \theta') = \varphi(\theta)\varphi(\theta').$$

Démonstration. On a $\varphi(0) = \cos(0) + i\sin(0) = 1$ et :

$$\begin{aligned}\varphi(\theta+\theta') &= \cos(\theta+\theta') + i\sin(\theta+\theta')\\ &= (\cos(\theta)\cos(\theta') - \sin(\theta)\sin(\theta')) + i(\sin(\theta)\cos(\theta') + \cos(\theta)\sin(\theta'))\\ &= (\cos(\theta) + i\sin(\theta))(\cos(\theta') + i\sin(\theta'))\\ &= \varphi(\theta)\varphi(\theta')\end{aligned}$$

■

La fonction φ vérifie donc la même équation fonctionnelle que la fonction exponentielle réelle. Cette remarque justifie la notation $\varphi(\theta) = e^{i\theta}$.

Avec $1 = \varphi(0) = \varphi(\theta - \theta) = \varphi(\theta)\varphi(-\theta)$, on déduit que $\dfrac{1}{\varphi(\theta)} = \varphi(-\theta) = \overline{\varphi(\theta)}$ (ce que l'on savait déjà : l'inverse d'un nombre complexe de module 1 est égal à son conjugué).

On a donc en résumé la notation :

$$\forall \theta \in \mathbb{R},\ e^{i\theta} = \cos(\theta) + i\sin(\theta)$$

ce qui définit une fonction 2π-périodique surjective de $\mathbb{R}$ sur l'ensemble Γ des nombres complexes de module 1 avec les propriétés suivantes :

$$\begin{cases} e^{i\cdot 0} = e^0 = 1\\ \forall (\theta, \theta') \in \mathbb{R}^2,\ e^{i(\theta+\theta')} = e^{i\theta}e^{i\theta'}\\ \forall \theta \in \mathbb{R},\ \dfrac{1}{e^{i\theta}} = e^{-i\theta} = \overline{e^{i\theta}}\\ \forall (\theta, \theta') \in \mathbb{R}^2,\ \left(e^{i\theta} = e^{i\theta'}\right) \Leftrightarrow (\exists k \in \mathbb{Z} \mid \theta' = \theta + 2k\pi)\\ \forall \theta \in \mathbb{R},\ \cos(\theta) = \Re\left(e^{i\theta}\right) = \dfrac{e^{i\theta} + e^{-i\theta}}{2} \text{ et } \sin(\theta) = \Im\left(e^{i\theta}\right) = \dfrac{e^{i\theta} - e^{-i\theta}}{2i} \end{cases}$$

Par récurrence sur $n \geq 0$, on déduit facilement que $\left(e^{i\theta}\right)^n = e^{in\theta}$. Puis pour $n < 0$ on a $e^{in\theta} = \dfrac{1}{e^{-in\theta}} = \left(\dfrac{1}{e^{i\theta}}\right)^{-n} = \left(e^{-i\theta}\right)^{-n} = e^{in\theta}$, c'est-à-dire que cette formule est valable pour tous les entiers relatifs. Nous verrons un peu plus loin l'intérêt de cette égalité.

On a en particulier les valeurs suivantes :

$$e^{i\pi} = -1,\ e^{i\frac{\pi}{2}} = i$$

les égalités $e^{i\theta} = 1$, $e^{i\theta} = -1$ et $e^{i\theta} = i$ étant réalisées respectivement si, et seulement si $\theta = 2k\pi$, $\theta = (2k+1)\pi$ et $\theta = \dfrac{\pi}{2} + 2k\pi$, où k est un entier relatif.

Un nombre complexe non nul peut donc s'écrire sous la forme $z = \rho e^{i\theta}$ où ρ est un réel strictement positif uniquement déterminé, c'est le module de z, et θ est un argument de z. Cette écriture est l'écriture polaire (ou trigonométrique) de z.

Exercice 2.21 *Soient z, z' deux nombres complexes non nuls. Montrer que $|z+z'| = |z| + |z'|$ si, et seulement si, $\arg(z) \equiv \arg(z')\ (2\pi)$.*

Solution 2.21 *L'égalité $|z+z'| = |z| + |z'|$ est équivalente à $|z+z'|^2 = (|z| + |z'|)^2$ avec :*

$$|z+z'|^2 = (z+z')\left(\overline{z} + \overline{z'}\right) = |z|^2 + |z'|^2 + z\overline{z'} + \overline{z}z'.$$

On a donc $|z+z'| = |z| + |z'|$ si, et seulement si, $z\overline{z'} + \overline{z}z' = 2|z||z'|$, ce qui équivaut encore à $\Re\left(z\overline{z'}\right) = |z||z'|$. En utilisant l'écriture polaire, $z = \rho e^{i\theta}$ et $z' = \rho' e^{i\theta'}$ avec $\rho > 0$, $\rho' > 0$ et θ, θ' réels, on déduit que $|z+z'| = |z| + |z'|$ si, et seulement si, $\cos(\theta - \theta') = 1$, ce qui équivaut à $\theta - \theta' \equiv 0\ (2\pi)$ et signifie que $\arg(z) \equiv \arg(z')\ (2\pi)$.

Plus généralement, on a le résultat suivant.

Exercice 2.22 *Démontrer que, pour tout entier naturel r non nul et toute famille $(z_1, \cdots, z_r)$ de complexes non nuls, l'égalité :*

$$\left|\sum_{k=1}^{r} z_k\right| = \sum_{k=1}^{r} |z_k|$$

est réalisée si, et seulement si :

$$\forall k \in \mathbb{N},\ (2 \leq k \leq r) \Rightarrow (\exists \lambda_k \in]0, +\infty[,\ z_k = \lambda_k z_1)$$

ce qui revient à dire que tous les z_k ont le même argument modulo 2π.

Solution 2.22 *Chaque nombre complexe non nul z_k $(1 \leq k \leq r)$ peut s'écrire $z_k = \rho_k e^{i\theta_k}$ avec $\rho_k = |z_k| > 0$ et $\theta_k \in [-\pi, \pi[$. On a alors :*

$$\begin{cases} \left|\sum_{k=1}^{r} z_k\right|^2 = \sum_{k=1}^{r} |z_k|^2 + 2 \sum_{1 \leq j < k \leq r} \rho_j \rho_k \cos(\theta_j - \theta_k), \\ \left(\sum_{k=1}^{r} |z_k|\right)^2 = \sum_{k=1}^{r} |z_k|^2 + 2 \sum_{1 \leq j < k \leq n} \rho_j \rho_k \end{cases}$$

et l'égalité $\left|\sum_{k=1}^{r} z_k\right| = \sum_{k=1}^{r} |z_k|$ est équivalente à :

$$\sum_{1 \leq j < k \leq r} \rho_j \rho_k (1 - \cos(\theta_j - \theta_k)) = 0.$$

Tous les termes de cette somme étant positifs ou nuls avec $\rho_j \rho_k > 0$, on en déduit que $\cos(\theta_j - \theta_k) = 1$ avec $\theta_j - \theta_k \in]-2\pi, 2\pi[$ pour $1 \leq j < k \leq r$ (on a $-\pi \leq \theta_j < \pi$ et $-\pi \leq \theta_k < \pi$ donc $-\pi < -\theta_k \leq \pi$ et $-2\pi < \theta_j - \theta_k < 2\pi$), ce qui donne $\theta_j = \theta_k$ et en notant θ cette valeur commune on a $z_k = \rho_k e^{i\theta} = |z_k| e^{i\theta}$ pour tout entier k compris entre 1 et r ou encore :

$$z_k = \frac{|z_k|}{|z_1|} |z_1| e^{i\theta} = \lambda_k z_1\ (1 \leq k \leq r)$$

où on a posé $\lambda_k = \dfrac{|z_k|}{|z_1|}$ pour tout k compris entre 1 et r.
Réciproquement si $z_k = \lambda_k z_1$ avec $\lambda_k > 0$ pour tout k compris entre 2 et r et $\lambda_1 = 1$, on a :

$$\left|\sum_{k=1}^{r} z_k\right| = |z_1| \sum_{k=1}^{r} \lambda_k = \sum_{k=1}^{r} \lambda_k |z_1| = \sum_{k=1}^{r} |z_k|.$$

On peut aussi démontrer ce résultat par récurrence sur $r \geq 1$, le cas $r = 2$ correspondant au cas d'égalité dans l'inégalité triangulaire sur $\mathbb{C}$.

Exercice 2.23 *Déterminer, pour tout couple de réel* (θ, θ') *tel que* $\cos\left(\dfrac{\theta-\theta'}{2}\right) \neq 0$, *le module et un argument de* $e^{i\theta} + e^{i\theta'}$.

Solution 2.23 *On a :*

$$e^{i\theta} + e^{i\theta'} = e^{i\frac{\theta+\theta'}{2}}\left(e^{i\frac{\theta-\theta'}{2}} + e^{-i\frac{\theta-\theta'}{2}}\right) = 2\cos\left(\frac{\theta-\theta'}{2}\right)e^{i\frac{\theta+\theta'}{2}}$$

donc $e^{i\theta} + e^{i\theta'} \neq 0$ *si* $\cos\left(\dfrac{\theta-\theta'}{2}\right) \neq 0$,

$$\left|e^{i\theta} + e^{i\theta'}\right| = 2\left|\cos\left(\frac{\theta-\theta'}{2}\right)\right|$$

et :

$$\arg\left(e^{i\theta} + e^{i\theta'}\right) \equiv \begin{cases} \dfrac{\theta+\theta'}{2}\ (2\pi) \ \ si\ \ \cos\left(\dfrac{\theta-\theta'}{2}\right) > 0 \\ \dfrac{\theta+\theta'}{2} + \pi\ (2\pi) \ \ si\ \ \cos\left(\dfrac{\theta-\theta'}{2}\right) < 0 \end{cases}$$

Exercice 2.24 *Pour tout réel* θ, *on désigne par* z_θ *le nombre complexe* $z_\theta = 1 - e^{i\theta}$.

1. *Exprimer* z_θ *en fonction de* $\sin\left(\dfrac{\theta}{2}\right)$ *et de* $e^{i\frac{\theta}{2}}$.
2. *Montrer que pour tout entier naturel* n *et tout réel* $\theta \in \mathbb{R} \setminus 2\pi\mathbb{Z}$, *on a :*

$$\sum_{k=0}^{n} e^{ik\theta} = e^{in\frac{\theta}{2}}\frac{\sin\left(\frac{n+1}{2}\theta\right)}{\sin\left(\frac{\theta}{2}\right)}$$

3. *En déduire des expressions de* $\sum_{k=0}^{n}\cos(k\theta)$ *et de* $\sum_{k=1}^{n}\sin(k\theta)$ *pour tout entier naturel non nul* n *et tout réel* θ.

Solution 2.24

1. *On a :*

$$z_\theta = e^{i\frac{\theta}{2}}\left(e^{-i\frac{\theta}{2}} - e^{i\frac{\theta}{2}}\right) = -2i\sin\left(\frac{\theta}{2}\right)e^{i\frac{\theta}{2}}.$$

On peut aussi écrire que :

$$\begin{aligned} z_\theta &= 1 - \cos(\theta) - i\sin(\theta) = 2\sin^2\left(\frac{\theta}{2}\right) - 2i\cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right) \\ &= -2i\sin\left(\frac{\theta}{2}\right)\left(\cos\left(\frac{\theta}{2}\right) + i\sin\left(\frac{\theta}{2}\right)\right) = -2i\sin\left(\frac{\theta}{2}\right)e^{i\frac{\theta}{2}}. \end{aligned}$$

2. *Comme* $e^{i\theta} \neq 1$ *pour* $\theta \in \mathbb{R} \setminus 2\pi\mathbb{Z}$, *on a :*

$$\begin{aligned} \sum_{k=0}^{n} e^{ik\theta} = \sum_{k=0}^{n}\left(e^{i\theta}\right)^k = \frac{1-e^{i(n+1)\theta}}{1-e^{i\theta}} &= \frac{-2i\sin\left(\frac{n+1}{2}\theta\right)e^{i\frac{n+1}{2}\theta}}{-2i\sin\left(\frac{\theta}{2}\right)e^{i\frac{\theta}{2}}} \\ &= e^{in\frac{\theta}{2}}\frac{\sin\left(\frac{n+1}{2}\theta\right)}{\sin\left(\frac{\theta}{2}\right)}. \end{aligned}$$

3. *Ce qui donne en identifiant les parties réelles et imaginaires dans l'identité précédente :*

$$\sum_{k=0}^{n}\cos(k\theta)=\cos\left(n\frac{\theta}{2}\right)\frac{\sin\left(\frac{n+1}{2}\theta\right)}{\sin\left(\frac{\theta}{2}\right)}$$

et :

$$\sum_{k=1}^{n}\sin(k\theta)=\sin\left(n\frac{\theta}{2}\right)\frac{\sin\left(\frac{n+1}{2}\theta\right)}{\sin\left(\frac{\theta}{2}\right)}$$

pour $n\geq 1$ et $\theta\in\mathbb{R}\setminus 2\pi\mathbb{Z}$. Pour $\theta\in 2\pi\mathbb{Z}$, on a $\cos(k\theta)=1$ et $\sin(k\theta)=0$ pour tout k, de sorte que $\sum_{k=0}^{n}\cos(k\theta)=n+1$ et $\sum_{k=1}^{n}\sin(k\theta)=0$.

Exercice 2.25 *Pour tout réel $\theta\in\left]-\pi,\pi\right[$, on désigne par z_θ le nombre complexe $z_\theta=1+e^{i\theta}$.*

1. *Exprimer z_θ en fonction de $\cos\left(\frac{\theta}{2}\right)$ et de $e^{i\frac{\theta}{2}}$.*
2. *En calculant, pour n entier naturel non nul, z_θ^n de deux manières différentes, montrer que :*

$$\sum_{k=0}^{n}C_n^k\cos(k\theta)=2^n\cos^n\left(\frac{\theta}{2}\right)\cos\left(n\frac{\theta}{2}\right)$$

et :

$$\sum_{k=1}^{n}C_n^k\sin(k\theta)=2^n\cos^n\left(\frac{\theta}{2}\right)\sin\left(n\frac{\theta}{2}\right)$$

Solution 2.25

1. *On a :*

$$z_\theta=e^{i\frac{\theta}{2}}\left(e^{-i\frac{\theta}{2}}+e^{i\frac{\theta}{2}}\right)=2\cos\left(\frac{\theta}{2}\right)e^{i\frac{\theta}{2}}.$$

On peut aussi écrire que :

$$\begin{aligned}z_\theta&=1+\cos(\theta)+i\sin(\theta)=2\cos^2\left(\frac{\theta}{2}\right)+2i\cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right)\\&=2\cos\left(\frac{\theta}{2}\right)\left(\cos\left(\frac{\theta}{2}\right)+i\sin\left(\frac{\theta}{2}\right)\right)=2\cos\left(\frac{\theta}{2}\right)e^{i\frac{\theta}{2}}.\end{aligned}$$

2. *On a d'une part :*

$$z_\theta^n=2^n\cos^n\left(\frac{\theta}{2}\right)e^{i\frac{n\theta}{2}}$$

et d'autre part, en utilisant la formule du binôme de Newton :

$$z_\theta^n=\left(1+e^{i\theta}\right)^n=\sum_{k=0}^{n}C_n^ke^{ik\theta}$$

ce qui donne en identifiant les parties réelles et imaginaires :

$$\sum_{k=0}^{n}C_n^k\cos(k\theta)=2^n\cos^n\left(\frac{\theta}{2}\right)\cos\left(n\frac{\theta}{2}\right)$$

et :

$$\sum_{k=1}^{n}C_n^k\sin(k\theta)=2^n\cos^n\left(\frac{\theta}{2}\right)\sin\left(n\frac{\theta}{2}\right)$$

Exercice 2.26 *Pour tout réel $\theta \in]-\pi,\pi[$, on désigne par z_θ le nombre complexe $z_\theta = 1 + e^{i\theta}$.*

1. *Déterminer le module et un argument de z_θ.*
2. *Déterminer, pour $\theta \in]-\pi,\pi[\setminus \{0\}$, le module et un argument de $u_\theta = \dfrac{1+e^{i\theta}}{1-e^{i\theta}}$.*
3. *Déterminer, pour tout entier naturel $n \geq 2$, le module et un argument de z_θ^n.*
4. *On prend $\theta = 2\dfrac{\pi}{3}$ et $n = 2006$. Déterminer le module et l'argument de z_θ^n qui est dans $]-\pi,\pi[$.*
5. *Calculer $e^{i\frac{\pi}{12}}$.*
6. *On prend $\theta = \dfrac{\pi}{6}$ et $n = 2006$. Déterminer le module et l'argument de z_θ^n qui est dans $]-\pi,\pi[$.*

Solution 2.26

1. *L'exercice 2.25 nous dit que $z_\theta = 2\cos\left(\dfrac{\theta}{2}\right)e^{i\frac{\theta}{2}}$ et tenant compte de $\cos\left(\dfrac{\theta}{2}\right) > 0$ pour $\theta \in]-\pi,\pi[$, on déduit que :*

$$|z_\theta| = 2\cos\left(\frac{\theta}{2}\right) \text{ et } \arg(z_\theta) \equiv \frac{\theta}{2}\ (2\pi)$$

2. *Pour $\theta \in]-\pi,\pi[\setminus \{0\}$, on a $\sin\left(\dfrac{\theta}{2}\right) \neq 0$ et :*

$$u_\theta = \frac{2\cos\left(\frac{\theta}{2}\right)e^{i\frac{\theta}{2}}}{-2i\sin\left(\frac{\theta}{2}\right)e^{i\frac{\theta}{2}}} = i\operatorname{cotan}\left(\frac{\theta}{2}\right)$$

(exercice 2.24), ce qui donne :

$$|u_\theta| = \left|\operatorname{cotan}\left(\frac{\theta}{2}\right)\right|$$

et :

$$\arg(z_\theta) \equiv \begin{cases} \dfrac{\pi}{2}\ (2\pi) & si\ \theta \in]0,\pi[\\ -\dfrac{\pi}{2}\ (2\pi) & si\ \theta \in]-\pi,0[\end{cases}$$

3. *On a $z_\theta^n = 2^n\cos^n\left(\dfrac{\theta}{2}\right)e^{i\frac{n\theta}{2}}$ avec $\cos\left(\dfrac{\theta}{2}\right) > 0$ pour $\theta \in]-\pi,\pi[$, donc :*

$$|z_\theta^n| = 2^n\cos^n\left(\frac{\theta}{2}\right) \text{ et } \arg(z_\theta^n) \equiv \frac{n\theta}{2}\ (2\pi)$$

4. *On a :*

$$z_\theta = 2\cos\left(\frac{\pi}{3}\right)e^{i\frac{\pi}{3}} = e^{i\frac{\pi}{3}} = \frac{1}{2} + i\frac{\sqrt{3}}{2}$$

$2006 = 6 * 334 + 2$, *de sorte que :*

$$|z_\theta^n| = 1 \text{ et } \arg(z_\theta^n) \equiv \frac{n\pi}{3} \equiv \frac{2\pi}{3}\ (2\pi)$$

5. *On a :*
$$\frac{\sqrt{3}}{2} = \cos\left(\frac{\pi}{6}\right) = 2\cos^2\left(\frac{\pi}{12}\right) - 1 = 1 - 2\sin^2\left(\frac{\pi}{12}\right)$$
avec $\cos\left(\frac{\pi}{12}\right) > 0$ *et* $\sin\left(\frac{\pi}{12}\right) > 0$, *donc :*
$$\cos\left(\frac{\pi}{12}\right) = \sqrt{\frac{1}{2} + \frac{\sqrt{3}}{4}} = \frac{\sqrt{2+\sqrt{3}}}{2}$$
$$\sin\left(\frac{\pi}{12}\right) = \sqrt{\frac{1}{2} - \frac{\sqrt{3}}{4}} = \frac{\sqrt{2-\sqrt{3}}}{2}$$
6. *On a*
$$|z_\theta^n| = 2^n \left|\cos\left(\frac{\pi}{12}\right)\right|^n = \left(\sqrt{2+\sqrt{3}}\right)^n = \left(2+\sqrt{3}\right)^{1003}$$
et $2006 = 24 * 83 + 14$ *de sorte que :*
$$\arg\left(z_\theta^n\right) \equiv n\frac{\pi}{12} \equiv \frac{7\pi}{6} \ (2\pi)$$

Exercice 2.27 *Pour cet exercice,* θ *est un réel fixé appartenant à* $]0, \pi[$.

1. *Déterminer le module et un argument de* $u_\theta = 1 + e^{i\theta}$.
2. *Déterminer le module et un argument de* $v_\theta = 1 - e^{i\theta}$.
3. *Résoudre dans* $\mathbb{C}$ *l'équation* $z^2 - 2ze^{i\theta} + 2ie^{i\theta}\sin(\theta) = 0$.

Solution 2.27

1. *On a :*
$$u_\theta = 1 + \cos(\theta) + i\sin(\theta) = 2\cos^2\left(\frac{\theta}{2}\right) + 2i\cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right)$$
$$= 2\cos\left(\frac{\theta}{2}\right)\left(\cos\left(\frac{\theta}{2}\right) + i\sin\left(\frac{\theta}{2}\right)\right) = 2\cos\left(\frac{\theta}{2}\right)e^{i\frac{\theta}{2}}$$
avec $\cos\left(\frac{\theta}{2}\right) > 0$ *pour* $\theta \in]0, \pi[$, *donc :*
$$|u_\theta| = 2\cos\left(\frac{\theta}{2}\right) \ \text{ et } \ \arg(u_\theta) \equiv \frac{\theta}{2} \ (2\pi)$$
2. *On a :*
$$v_\theta = 1 + e^{i(\theta+\pi)} = 2\cos\left(\frac{\theta}{2} + \frac{\pi}{2}\right)e^{i\left(\frac{\theta}{2}+\frac{\pi}{2}\right)}$$
$$= -2\sin\left(\frac{\theta}{2}\right)e^{i\left(\frac{\theta}{2}+\frac{\pi}{2}\right)} = 2\sin\left(\frac{\theta}{2}\right)e^{i\left(\frac{\theta}{2}+3\frac{\pi}{2}\right)}$$
avec $\sin\left(\frac{\theta}{2}\right) > 0$ *pour* $\theta \in]0, \pi[$, *donc :*
$$|v_\theta| = 2\sin\left(\frac{\theta}{2}\right) \ \text{ et } \ \arg(v_\theta) \equiv \frac{\theta}{2} + 3\frac{\pi}{2} \equiv \frac{\theta}{2} - \frac{\pi}{2} \ (2\pi)$$

3. On a :

$$\begin{aligned} P(z) = z^2 - 2ze^{i\theta} + 2ie^{i\theta}\sin(\theta) &= \left(z - e^{i\theta}\right)^2 + e^{i\theta}\left(2i\sin(\theta) - e^{i\theta}\right) \\ &= \left(z - e^{i\theta}\right)^2 + e^{i\theta}\left(i\sin(\theta) - \cos(\theta)\right) \\ &= \left(z - e^{i\theta}\right)^2 - e^{i\theta}\left(\cos(\theta) - i\sin(\theta)\right) \\ &= \left(z - e^{i\theta}\right)^2 - e^{i\theta}e^{-i\theta} = \left(z - e^{i\theta}\right)^2 - 1 \end{aligned}$$

et $P(z) = 0$ équivaut à $z = e^{i\theta} \pm 1$. Les solutions sont donc $u_\theta = 1 + e^{i\theta}$ et $-v_\theta = e^{i\theta} - 1$.

Exercice 2.28 *On désigne par $D(0,1)$ le disque unité fermé du plan complexe, soit :*

$$D(0,1) = \{z \in \mathbb{C} \mid |z| \leq 1\}$$

Montrer que l'application :

$$\begin{array}{rccc} \varphi : & \mathbb{C} \setminus D(0,1) & \to & \mathbb{C} \setminus [-1,1] \\ & z & \mapsto & \dfrac{1}{2}\left(z + \dfrac{1}{z}\right) \end{array}$$

est bijective.

Solution 2.28 *Tout $z \in \mathbb{C} \setminus D(0,1)$ s'écrit $z = \rho e^{i\theta}$ avec $\rho > 1$ et :*

$$\begin{aligned} \varphi(z) &= \frac{1}{2}\left(\rho e^{i\theta} + \frac{1}{\rho}e^{-i\theta}\right) \\ &= \frac{1}{2}\left(\left(\rho + \frac{1}{\rho}\right)\cos(\theta) + i\left(\rho - \frac{1}{\rho}\right)\sin(\theta)\right) \end{aligned}$$

Si $\varphi(z)$ est réel, nécessairement $\sin(\theta) = 0$, donc $\cos(\theta) = \pm 1$ et :

$$|\varphi(z)| = \frac{\rho^2 + 1}{2\rho} = \frac{(\rho - 1)^2 + 2\rho}{2\rho} > \frac{2\rho}{2\rho} = 1.$$

On a donc bien $\varphi(z) \in \mathbb{C} \setminus [-1,1]$ pour tout $z \in \mathbb{C} \setminus D(0,1)$.
Il s'agit maintenant de montrer que pour tout $Z \in \mathbb{C} \setminus [-1,1]$ l'équation $\frac{1}{2}\left(z + \frac{1}{z}\right) = Z$ a une unique solution $z \in \mathbb{C} \setminus D(0,1)$. Cette équation est équivalente à l'équation de degré 2 :

$$z^2 - 2Zz + 1 = 0$$

Le discriminant réduit de cette équation est $\delta' = Z^2 - 1 \neq 0$, ce qui donne deux solution distinctes $z_1 = \rho_1 e^{i\theta_1}$ et $z_2 = \rho_2 e^{i\theta_2}$ dans $\mathbb{C}$. De $z_1 z_2 = 1$, on déduit que $\rho_1\rho_2 e^{i(\theta_1+\theta_2)} = 1$, donc que $e^{i(\theta_1+\theta_2)}$ est réel positif et $\theta_1 + \theta_2 \equiv 0$ modulo 2π. On a donc $z_1 = \rho_1 e^{i\theta_1}$ et $z_2 = \frac{1}{\rho_1}e^{-i\theta_1}$. Si $\rho_1 = 1$, on a alors :

$$2Z = z_1 + z_2 = e^{i\theta_1} + e^{-i\theta_1} = 2\cos(\theta_1)$$

et $Z = \cos(\theta_1) \in [-1,1]$, ce qui n'est pas. On a donc $\rho_1 \neq 1$, de sorte que $z_1 \in \mathbb{C} \setminus D(0,1)$ et $z_2 \notin \mathbb{C} \setminus D(0,1)$ pour $\rho_1 > 1$, ou $z_2 \in \mathbb{C} \setminus D(0,1)$ et $z_1 \notin \mathbb{C} \setminus D(0,1)$ pour $\rho_1 < 1$. Il y a donc une unique racine dans $\mathbb{C} \setminus D(0,1)$.

L'écriture polaire des nombres complexes non nuls peut aussi être utilisée pour résoudre des équations de la forme $a\cos(x) + b\sin(x) = 0$. Pour ce faire, on utilise le résultat suivant.

Théorème 2.13 *Soient a, b des réels non tous deux nuls (c'est-à-dire que $(a,b) \neq (0,0)$). Il existe un unique couple (ρ, θ) de réels dans $\mathbb{R}^{+,*} \times [-\pi, \pi[$ tel que :*

$$\forall x \in \mathbb{R},\ a\cos(x) + b\sin(x) = \rho\cos(x-\theta).$$

ρ est le module de $u = a + ib$ et θ son argument principal.

Démonstration. Comme $(a,b) \neq (0,0)$, on a $u = a + ib \neq 0$ et ce nombre complexe s'écrit de manière unique $u = \rho e^{i\theta}$ avec $\rho = |u| = \sqrt{a^2+b^2} > 0$ et $\theta \in [-\pi, \pi[$. Pour tout réel x, on a alors d'un part :

$$\begin{aligned} ue^{-ix} &= (a+ib)(\cos(x) - i\sin(x)) \\ &= a\cos(x) + b\sin(x) + i(b\cos(x) - a\sin(x)) \end{aligned}$$

et d'autre part :

$$ue^{-ix} = \rho e^{i(\theta - x)} = \rho\cos(\theta - x) + i\rho\sin(\theta - x)$$

ce qui donne :

$$a\cos(x) + b\sin(x) = \rho\cos(\theta - x) = \rho\cos(x - \theta)$$

et aussi :

$$a\sin(x) - b\cos(x) = \rho\sin(x - \theta).$$

Pour ce qui est de l'unicité, supposons que (ρ', θ') soit une autre solution à notre problème. On a alors, pour tout réel x :

$$\rho\cos(x-\theta) = \rho'\cos(x-\theta').$$

Prenant $x = \theta$ [resp. $x = \theta'$], on en déduit que $\rho = \rho'\cos(\theta - \theta') \leq \rho'$ [resp. $\rho' = \rho\cos(\theta' - \theta) \leq \rho$] et $\rho = \rho'$.

Prenant $x = 0$ [resp. $x = \dfrac{\pi}{2}$], on a $\cos(\theta) = \cos(\theta')$ [resp. $\cos\left(\dfrac{\pi}{2} - \theta\right) = \sin(\theta) = \cos\left(\dfrac{\pi}{2} - \theta'\right) = \sin(\theta')$] avec θ, θ' dans $[-\pi, \pi[$, ce qui équivaut à $\theta = \theta'$. ■

Corollaire 2.2 *Soient a, b des réels non tous deux nuls. Les solutions de l'équation :*

$$a\cos(x) + b\sin(x) = 0 \tag{2.3}$$

sont les réels :

$$x = \theta + \frac{\pi}{2} + k\pi$$

où $\theta \in [-\pi, \pi[$ est l'argument principal de $a + ib$ et k un entier relatif.

Démonstration. Le théorème précédent nous que l'équation (2.3) est équivalente à $\cos(x-\theta) = 0$, soit à $(x-\theta) \equiv \dfrac{\pi}{2}\ (\pi)$, encore équivalent à dire que $x = \theta + \dfrac{\pi}{2} + k\pi$ avec $k \in \mathbb{Z}$. ■

Les formules suivantes, valables pour $\theta \in \mathbb{R}$ et $n \in \mathbb{Z}$:

$$\begin{cases} \cos(\theta) = \Re\left(e^{i\theta}\right) = \dfrac{e^{i\theta} + e^{-i\theta}}{2} \\ \sin(\theta) = \Im\left(e^{i\theta}\right) = \dfrac{e^{i\theta} - e^{-i\theta}}{2i} \end{cases}$$

(formules d'Euler) et :

$$(\cos(\theta) + i\sin(\theta))^n = \left(e^{i\theta}\right)^n = e^{in\theta} = \cos(n\theta) + i\sin(n\theta)$$

(formule de Moivre) permettent d'obtenir relativement facilement des formules de trigonométrie.

En utilisant la formule du binôme de Newton, la formule de Moivre s'écrit :

$$\cos(n\theta) + i\sin(n\theta) = \sum_{k=0}^{n} C_n^k \cos^k(\theta) \sin^{n-k}(\theta)\, i^{n-k}$$

et l'identification des parties réelles et imaginaires nous permet d'exprimer $\cos(n\theta)$ et $\sin(n\theta)$ comme combinaisons linéaires de puissances de $\cos(\theta)$ et $\sin(\theta)$.

Exercice 2.29 *Exprimer* $\cos(4\theta)$ *et* $\sin(4\theta)$ *comme combinaisons linéaires de puissances de* $\cos(\theta)$ *et* $\sin(\theta)$.

Solution 2.29 *On a :*

$$\begin{aligned}\cos(4\theta) + i\sin(4\theta) = e^{i4\theta} = \left(e^{i\theta}\right)^4 &= (\cos(\theta) + i\sin(\theta))^4 \\ &= \cos^4(\theta) + 4\cos^3(\theta)\sin(\theta)\, i - 6\cos^2(\theta)\sin^2(\theta) \\ &\quad -4\cos(\theta)\sin^3(\theta)\, i + \sin^4(\theta)\end{aligned}$$

et en conséquence :

$$\begin{aligned}\cos(4\theta) &= \cos^4(\theta) - 6\cos^2(\theta)\sin^2(\theta) + \sin^4(\theta) \\ &= \left(\cos^2(\theta) + \sin^2(\theta)\right)^2 - 8\cos^2(\theta)\sin^2(\theta) \\ &= 1 - 8\cos^2(\theta)\sin^2(\theta)\end{aligned}$$

et :

$$\begin{aligned}\sin(4\theta) &= 4\left(\cos^3(\theta)\sin(\theta) - \cos(\theta)\sin^3(\theta)\right) \\ &= 4\cos(\theta)\sin(\theta)\left(\cos^2(\theta) - \sin^2(\theta)\right)\end{aligned}$$

L'utilisation des formules d'Euler et de la formule du binôme de Newton nous permet d'exprimer des puissances $\cos(\theta)$ et $\sin(\theta)$ comme combinaisons linéaires de $\cos(p\theta)$ et $\sin(q\theta)$. On dit qu'on linéarise $\cos^n(\theta)$ ou $\sin^m(\theta)$, n et m étant des entiers naturels.

Pour ce faire, on écrit que :

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

et nécessairement :

$$\cos^n(\theta) = \frac{1}{2^n}\sum_{k=0}^{n} C_n^k \cos((2k-n)\theta). \tag{2.4}$$

On peut remarquer que l'on a aussi :

$$\sum_{k=0}^{n} C_n^k \sin((2k-n)\theta) = 0.$$

En réalité cette formule est évidente. Par exemple, pour $n = 2p$, le changement d'indice $k = 2p - j$ nous permet d'écrire la deuxième moitié de cette somme sous la forme :

$$\begin{aligned}\sum_{k=p+1}^{2p} C_{2p}^{k} \sin\left((2k-2p)\,\theta\right) &= \sum_{j=0}^{p-1} C_{2p}^{2p-j} \sin\left((2p-2j)\,\theta\right)\\ &= -\sum_{j=0}^{p} C_{2p}^{j} \sin\left((2j-2p)\,\theta\right)\\ &= -\sum_{k=0}^{p} C_{n}^{k} \sin\left((2k-n)\,\theta\right)\end{aligned}$$

La vérification étant analogue pour n impair.

La parité de $\cos^p(\theta)$ peut aussi justifier l'absence de termes en $\sin(q\theta)$ dans la formule (2.4).

Exercice 2.30 *Linéariser* $\cos^4(\theta)$ *et* $\sin^4(\theta)$.

Solution 2.30 *On a :*

$$\begin{aligned}\cos^4(\theta) &= \frac{1}{16}\left(e^{i\theta}+e^{-i\theta}\right)^4\\ &= \frac{1}{16}\left(e^{4i\theta}+4e^{3i\theta}e^{-i\theta}+6e^{2i\theta}e^{-2i\theta}+4e^{i\theta}e^{-3i\theta}+e^{-4i\theta}\right)\\ &= \frac{1}{16}\left(e^{4i\theta}+e^{-4i\theta}+4\left(e^{2i\theta}+e^{-2i\theta}\right)+6\right)\\ &= \frac{1}{8}\cos(4\theta)+\frac{1}{2}\cos(2\theta)+\frac{3}{8}\end{aligned}$$

et :

$$\begin{aligned}\sin^4(\theta) &= \frac{1}{16}\left(e^{i\theta}-e^{-i\theta}\right)^4\\ &= \frac{1}{16}\left(e^{4i\theta}-4e^{3i\theta}e^{-i\theta}+6e^{2i\theta}e^{-2i\theta}-4e^{i\theta}e^{-3i\theta}+e^{-4i\theta}\right)\\ &= \frac{1}{16}\left(e^{4i\theta}+e^{-4i\theta}-4\left(e^{2i\theta}+e^{-2i\theta}\right)+6\right)\\ &= \frac{1}{8}\cos(4\theta)-\frac{1}{2}\cos(2\theta)+\frac{3}{8}\end{aligned}$$

On peut aussi exprimer $\tan(n\theta)$ en fonction de $\tan(\theta)$ pour tout entier naturel n.

Exercice 2.31 *Exprimer* $\tan(4\theta)$ *en fonction de* $\tan(\theta)$.

Solution 2.31 *On a :*

$$\tan(4\theta) = \frac{\sin(4\theta)}{\cos(4\theta)} = 4\frac{\cos^3(\theta)\sin(\theta)-\cos(\theta)\sin^3(\theta)}{\cos^4(\theta)-6\cos^2(\theta)\sin^2(\theta)+\sin^4(\theta)}$$

et divisant numérateur et dénominateur par $\cos^4(\theta)$, *on obtient :*

$$\tan(4\theta) = 4\tan(\theta)\,\frac{1-\tan^2(\theta)}{1-6\tan^2(\theta)+\tan^4(\theta)}.$$

2.7 Racines n-ièmes d'un nombre complexe

La représentation des nombres complexes nous sera très utile pour résoudre des équations de la forme $x^n = \alpha$.

On rappelle que pour tout entier naturel non nul n, la fonction $f : x \mapsto x^n$ réalise une bijection de $\mathbb{R}^+$ sur lui même. L'application réciproque de f est notée $x \mapsto \sqrt[n]{x}$ ou $x \mapsto x^{\frac{1}{n}}$ et on l'appelle fonction racine n-ième. Donc pour tout réel positif a, l'unique solution de l'équation $x^n = a$ est le réel positif $\sqrt[n]{a}$.

On rappelle que si $z = re^{it}$ avec $r > 0$ et $t \in \mathbb{R}$, on a pour tout entier relatif n, $z^n = r^n e^{int}$ et pour $z' = r'e^{it'}$ avec $r' > 0$ et $t' \in \mathbb{R}$, l'égalité $z = z'$ est réalisée si, et seulement si $r = r'$ et $t \equiv t'$ modulo 2π.

Dans le cas où $n = 2$ et $\alpha \neq 0$, on écrit $\alpha = \rho e^{i\theta}$ avec $\rho > 0$ et $\theta \in [-\pi, \pi[$ et on cherche $z = re^{it}$ avec $r > 0$ et $t \in [-\pi, \pi[$ tel que :

$$z^2 = r^2 e^{2it} = \rho e^{i\theta}$$

ce qui équivaut à $r^2 = \rho$ et $2t = \theta + 2k\pi$ avec $k \in \mathbb{Z}$. On a donc $r = \sqrt{\rho}$ et $t = \dfrac{\theta}{2} + k\pi$ avec $k \in \mathbb{Z}$, ce qui donne $z = \sqrt{\rho} e^{i\frac{\theta}{2}} e^{ik\pi}$ et α a deux racines carrées qui sont :

$$z_1 = \sqrt{\rho} e^{i\frac{\theta}{2}} \text{ et } z_2 = \sqrt{\rho} e^{i\frac{\theta}{2}} e^{i\pi} = -z_1.$$

Définition 2.8 *Étant donné un nombre complexe α et un entier naturel non nul n, on appelle racine n-ième de α tout nombre complexe z tel que $z^n = \alpha$.*

Remarque 2.1 *Si $\alpha = 0$, l'équation $z^n = \alpha$ équivaut à $z = 0$, c'est-à-dire que 0 est l'unique racine n-ième de 0.*
Si $\alpha \neq 0$, une racine n-ième de α est nécessairement non nulle.

Remarque 2.2 *Si α est un nombre complexe non nul, il s'écrit $\alpha = \rho e^{i\theta}$ avec $\rho > 0$ et $\theta \in [-\pi, \pi[$ (ou $\theta \in \mathbb{R}$ si on se contente d'un quelconque argument de α) et le nombre complexe $z_0 = \sqrt[n]{\rho} e^{i\frac{\theta}{n}}$ nous fournit une solution de l'équation $z^n = \alpha$. Pour tout autre solution z de cette équation on aura $z^n = z_0^n$, soit $\left(\dfrac{z}{z_0}\right)^n = 1$ et la connaissance de toutes les racines n-ièmes de 1 nous fournira toutes les racines n-ièmes de α.*

Définition 2.9 *Étant donné un entier naturel non nul n, on appelle racine n-ième de l'unité toute racine n-ième de 1.*

Théorème 2.14 *Soit n un entier naturel non nul. Il y a exactement n racines n-ièmes de l'unité qui sont données par :*

$$\omega_k = e^{\frac{2ik\pi}{n}} = \cos\left(\frac{2k\pi}{n}\right) + i\sin\left(\frac{2k\pi}{n}\right) \ (0 \leq k \leq n-1)$$

Démonstration. Si $z^n = 1$, on a alors $|z|^n = |z^n| = 1$, donc $|z| = 1$ (c'est l'unique racine n-ième réelle positive de 1) et $z = e^{i\theta}$ avec $\theta \in \mathbb{R}$. L'équation $z^n = 1$ équivaut alors à $e^{in\theta} = 1$, encore équivalent à $n\theta \equiv 0$ modulo 2π. Les racines n-ièmes de l'unité sont donc les nombres complexes $e^{\frac{2ik\pi}{n}}$ où k décrit l'ensemble $\mathbb{Z}$ des entiers relatifs. En effectuant la division euclidienne par n, tout entier k s'écrit $k = qn + r$ avec $0 \leq r \leq n-1$ et $e^{\frac{2ik\pi}{n}} = e^{\frac{2ir\pi}{n}} = \omega_r$. De plus pour j, k entiers compris entre 0 et $n-1$, l'égalité $\omega_j = \omega_k$ est équivalente à $e^{\frac{2i(j-k)\pi}{n}} = 1$, encore

équivalent à $\dfrac{2(j-k)\pi}{n} \equiv 0$ modulo 2π, ce qui revient à dire que $j-k$ est divisible par n, soit $j-k=qn$ et avec $|j-k| \leq n-1$ (puisque j et k sont dans l'intervalle $[0, n-1]$), on déduit que $q=0$ est la seule possibilité, ce qui signifie que $j=k$. On a donc bien le résultat annoncé. ■

Le théorème précédent peut aussi s'énoncer comme suit.

Théorème 2.15 *Pour tout entier naturel non nul n et tout nombre complexe z, on a :*

$$z^n - 1 = \prod_{k=0}^{n-1} (z - \omega_k)$$

où les $\omega_k = e^{\frac{2ik\pi}{n}}$, pour k compris entre 0 et $n-1$, sont les racines n-ièmes de l'unité.

Démonstration. Le polynôme $P(z) = z^n - 1 - \prod_{k=0}^{n-1} (z - \omega_k)$ est nul ou de degré au plus égal à $n-1$ (le coefficient dominant de $\prod_{k=0}^{n-1} (z-\omega_k)$ est z^n) et s'annule en n points distincts (les ω_k), c'est donc le polynôme nul. ■

Exemple 2.1 *Les racines cubiques de l'unité sont :*

$$\begin{cases} \omega_0 = 1, \\ \omega_1 = e^{\frac{2i\pi}{3}} = \cos\left(\dfrac{2\pi}{3}\right) + i\sin\left(\dfrac{2\pi}{3}\right) = -\dfrac{1}{2} + i\dfrac{\sqrt{3}}{2} \\ \omega_2 = e^{\frac{4i\pi}{3}} = e^{-\frac{2i\pi}{3}} = \cos\left(\dfrac{2\pi}{3}\right) - i\sin\left(\dfrac{2\pi}{3}\right) = -\dfrac{1}{2} - i\dfrac{\sqrt{3}}{2} \end{cases}$$

La racine ω_1 est usuellement notée j et $\omega_2 = \overline{j}$.

Corollaire 2.3 *Soit n un entier naturel non nul. Tout nombre complexe non nul $\alpha = \rho e^{i\theta}$ a exactement n racines n-ièmes données par :*

$$u_k = u_0\omega_k = \sqrt[n]{\rho} e^{i\frac{\theta}{n}} e^{\frac{2ik\pi}{n}} \quad (0 \leq k \leq n-1)$$

Exercice 2.32 *Résoudre dans $\mathbb{C}$ l'équation $z^6 = (\overline{z})^2$. Combien l'équation a-t-elle de solutions ?*

Solution 2.32 *On voit que $z=0$ est solution.*
Si $z^6 = (\overline{z})^2$ avec $z \neq 0$, alors $|z|=1$, donc $z = e^{i\theta}$ et $e^{i8\theta} = 1$, soit $\theta = \dfrac{2k\pi}{8}$ avec $k \in \mathbb{Z}$, ce qui donne les 8 solutions $e^{i\frac{k\pi}{4}}$ où $k \in \{0, 1, \cdots, 7\}$. Donc 9 solutions au total.

Exercice 2.33 *Résoudre dans $\mathbb{C}$ l'équation $z^4 = (\overline{z})^4$.*

Solution 2.33 *On voit que $z=0$ est solution.*
Pour $z \neq 0$, on écrit que $z = \rho e^{i\theta}$ avec $\rho > 0$ et $\theta \in [0, 2\pi[$ et de $z^4 = (\overline{z})^4$, on déduit que $e^{i8\theta} = 1$, soit $\theta = \dfrac{2k\pi}{8}$ avec $k \in \{0, 1, \cdots, 7\}$. Les solutions non nulles de cette équation sont donc les nombres complexes de la forme $\rho e^{i\frac{k\pi}{4}}$ où ρ est un réel strictement positif et k est un entier compris entre 0 et 7. L'ensemble S des solutions est donc infini, c'est la réunion des quatre droites D_k d'équation polaire $\theta = k\dfrac{\pi}{4}$ où k est entier compris entre 0 et 3. D_0 est l'axe des x, D_1 la diagonale d'équation $y=x$, D_2 l'axe des y et D_3 la diagonale d'équation $y=-x$.

Exercice 2.34 *Déterminer, pour n entier naturel non nul, toutes les racines n-ièmes de −1.*

Solution 2.34 *Il s'agit de résoudre l'équation* $z^n = -1 = e^{i\pi}$. *Les solutions de cette équation sont les :*

$$u_k = e^{(2k+1)\frac{i\pi}{n}} \ (0 \leq k \leq n-1)$$

Exercice 2.35 *En notant, pour n entier naturel non nul,* $(\omega_k)_{0\leq k\leq n-1}$ *la suite de toutes les racines n-ièmes de l'unité, montrer que pour* $n \geq 2$, *on a* $\sum_{k=0}^{n-1} \omega_k = 0$ *et* $\prod_{k=0}^{n-1} \omega_k = (-1)^{n-1}$.

Solution 2.35 *On a :*

$$\sum_{k=0}^{n-1} \omega_k = \sum_{k=0}^{n-1} \omega_1^k = \frac{1-\omega_1^n}{1-\omega_1} = 0$$

(pour $n \geq 2$, *on a bien* $\omega_1 = e^{\frac{2i\pi}{n}} \neq 1$*) et :*

$$\begin{aligned}\prod_{k=0}^{n-1} \omega_k = \prod_{k=0}^{n-1} \omega_1^k = \omega_1^{\sum_{k=0}^{n-1} k} = \omega_1^{\frac{n(n-1)}{2}} \\ = \left(e^{\frac{2i\pi}{n}}\right)^{\frac{n(n-1)}{2}} = e^{\frac{n(n-1)}{2}\frac{2i\pi}{n}} = e^{i(n-1)\pi} = \left(e^{i\pi}\right)^{n-1} = (-1)^{n-1}\end{aligned}$$

($m = \dfrac{n(n-1)}{2}$ *étant entier, on a bien* $\left(e^{i\theta}\right)^m = e^{im\theta}$*). De la première identité, on déduit que :*

$$\sum_{k=0}^{n-1} \cos\left(\frac{2k\pi}{n}\right) = \sum_{k=0}^{n-1} \sin\left(\frac{2k\pi}{n}\right) = 0.$$

Remarque 2.3 *La parenthèse (*$m = \dfrac{n(n-1)}{2}$ *étant entier ...) est due à la remarque du pointilleux lecteur qui voudrait montrer que* $-1 = 1$ *comme suit :*

$$\left(1 = e^{2i\pi}\right) \Rightarrow \left(1 = 1^{\frac{1}{2}} = \left(e^{2i\pi}\right)^{\frac{1}{2}} = e^{\frac{1}{2}2i\pi} = e^{i\pi} = -1\right)$$

Exercice 2.36 *Montrer que, pour tout entier* $n \geq 1$, *l'ensemble des racines* $2n$*-ièmes de l'unité est aussi donnée par :*

$$\Gamma_{2n} = \{-1, 1\} \cup \left\{e^{\frac{ik\pi}{n}} \mid 1 \leq k \leq n-1\right\} \cup \left\{e^{\frac{-ik\pi}{n}} \mid 1 \leq k \leq n-1\right\}$$

En déduire que, pour tout nombre complexe z, *on a :*

$$z^{2n} - 1 = \left(z^2 - 1\right) \prod_{k=1}^{n-1} \left(z^2 - 2\cos\left(\frac{k\pi}{n}\right) z + 1\right).$$

Solution 2.36 *Ces racines* $2n$*-ièmes sont les :*

$$\omega_k = e^{\frac{2ik\pi}{2n}} = e^{\frac{ik\pi}{n}} \ (0 \leq k \leq 2n-1).$$

Pour $k = 0$, *on a* $\omega_0 = 1$, *pour* $k = n$, *on a* $\omega_n = e^{i\pi} = -1$ *et pour* $k = 2n - j$ *compris entre* $n+1$ *et* $2n-1$, *on a :*

$$\omega_k = e^{\frac{i\{2n-j\}\pi}{n}} = e^{\frac{-ij\pi}{n}}$$

ce qui donne le résultat attendu.
On a donc, pour tout nombre complexe z :

$$\begin{aligned} z^{2n}-1 &= (z-1)(z+1)\prod_{k=1}^{n-1}\left(z-e^{i\frac{k\pi}{n}}\right)\left(z-e^{-i\frac{k\pi}{n}}\right) \\ &= \left(z^2-1\right)\prod_{k=1}^{n-1}\left(z^2-\left(e^{i\frac{k\pi}{n}}+e^{-i\frac{k\pi}{n}}\right)z+1\right) \\ &= \left(z^2-1\right)\prod_{k=1}^{n-1}\left(z^2-2\cos\left(\frac{k\pi}{n}\right)z+1\right) \end{aligned}$$

Exercice 2.37 *Résoudre dans $\mathbb{C}$ l'équation $z^8+z^4+1=0$.*

Solution 2.37 *Si z est solution de cette équation, alors $t=z^4$ est solution de $t^2+t+1=0$, ce qui donne $t\neq 1$ et $\dfrac{t^3-1}{t-1}=0$, donc $t=j=e^{\frac{2i\pi}{3}}$ ou $t=\overline{j}=j^2$.*
Il s'agit alors de calculer les racines quatrièmes de j et de $\overline{j}$, ces racines sont les :

$$z_k=e^{i\left(\frac{\pi}{6}+k\frac{\pi}{2}\right)} \text{ et } \overline{z_k}=e^{-i\left(\frac{\pi}{6}+k\frac{\pi}{2}\right)} \ (0\leq k\leq 3)$$

soit :

$$z_0=e^{i\frac{\pi}{6}}=\frac{\sqrt{3}}{2}+\frac{i}{2},\ z_1=e^{2i\frac{\pi}{3}}=j=-\frac{1}{2}+i\frac{\sqrt{3}}{2}$$
$$z_2=e^{7i\frac{\pi}{6}}=-z_1,\ z_3=e^{5i\frac{\pi}{3}}=-z_1$$

et leurs conjugués.
On peut aussi procéder comme suit. Si z est solution de cette équation, il est alors non nul et $z^4+1+\dfrac{1}{z^4}=0$, soit $\left(z^2+\dfrac{1}{z^2}\right)^2=1$, donc $z^2+\dfrac{1}{z^2}=\pm 1$, soit $\left(z+\dfrac{1}{z}\right)^2-2=\pm 1$, c'est-à-dire $\left(z+\dfrac{1}{z}\right)^2=1$ ou $\left(z+\dfrac{1}{z}\right)^2=3$.
Si $\left(z+\dfrac{1}{z}\right)^2=1$, on a alors $z+\dfrac{1}{z}=\pm 1$, soit $z^2\pm z+1=0$ avec $z\neq\pm 1$ ou encore $\dfrac{z^3\pm 1}{z\pm 1}=0$, ce qui donne les 4 solutions, $j,\overline{j},-j,-\overline{j}$.
Si $\left(z+\dfrac{1}{z}\right)^2=3$, on a alors $z+\dfrac{1}{z}=\pm\sqrt{3}$, soit $z^2\pm\sqrt{3}z+1=0$ ou encore $\left(z\pm\dfrac{\sqrt{3}}{2}\right)^2=-\dfrac{1}{4}$, ce qui donne les 4 autres solutions, $z_0=\dfrac{\sqrt{3}}{2}+\dfrac{i}{2},\overline{z_0},-z_0,-\overline{z_0}$.

2.8 Représentation géométrique des nombres complexes

On suppose connu le plan affine euclidien que l'on note $\mathcal{P}$ et que l'on muni d'un repère orthonormé direct (pour utiliser les angles orientés de vecteurs) $\mathcal{R}=(O,\overrightarrow{u},\overrightarrow{v})$.

Un point M de ce plan est repéré par ses coordonnées $(x,y)\in\mathbb{R}^2$, ce qui signifie qu'on a l'égalité vectorielle $\overrightarrow{OM}=x\overrightarrow{u}+y\overrightarrow{v}$.

On notera AB la distance de A à B et $(\widehat{\overrightarrow{v_1},\overrightarrow{v_2}})$ l'angle orienté des vecteurs $\overrightarrow{v_1},\overrightarrow{v_2}$.

L'application φ qui associe à tout nombre complexe $z=x+iy$ le point $M(z)$ de coordonnées (x,y) dans le repère $\mathcal{R}$ réalise une bijection de $\mathbb{C}$ sur $\mathcal{P}$.

Tout point M du plan $\mathcal{P}$ s'écrit donc de manière unique $M = M(z)$ et peut ainsi être identifié au nombre complexe z. Le plan $\mathcal{P}$ muni de cette identification est appelé plan complexe ou plan d'Argand-Cauchy.

Si $M \in \mathcal{P}$ s'écrit $M = M(z)$, on dit que z est l'affixe de M et M le point image de z. Le vecteur $\overrightarrow{OM}$ est aussi appelé vecteur image de z et on dit que z est l'affixe de $\overrightarrow{OM}$.

En utilisant cette identification entre $\mathcal{P}$ et $\mathbb{C}$, on peut donner les interprétations géométriques suivantes où $a, b, z, z', \cdots$ désignent des nombres complexes et $A, B, M, M', \cdots$ leurs images respectives dans $\mathcal{P}$:

- l'axe O_x est identifié à l'ensemble des nombres réels ;
- l'axe O_y est identifié à l'ensemble des imaginaires purs ;
- $a + b$ est l'affixe du sommet C du parallélogramme $OACB$ défini par $\overrightarrow{OC} = \overrightarrow{OA} + \overrightarrow{OB}$;
- $|z| = OM$ est la distance de O à M ;
- $|b - a| = AB$ est la distance de A à B ;
- l'ensemble des nombres complexes z tels que $|z - \omega| = \rho$ est identifié au cercle de centre $\Omega(\omega)$ et de rayon $\rho \geq 0$;
- l'ensemble des nombres complexes z tels que $|z - \omega| < \rho$ [resp. $|z - \omega| \leq \rho$] est identifié au disque ouvert [resp. fermé] de centre $\Omega(\omega)$ et de rayon $\rho \geq 0$;
- pour $A \neq B$, le point M est sur la médiatrice du segment $[AB]$ si, et seulement si, $|z - a| = |z - b|$;
- $\Re(\overline{z}z')$ est égal au produit scalaire $\overrightarrow{OM} \cdot \overrightarrow{OM'}$;
- si θ est un argument de z, c'est alors une mesure de l'angle orienté $\left(\widehat{\overrightarrow{u}, \overrightarrow{OM}}\right)$;
- plus généralement si les points A, B, C sont deux à deux distincts alors un argument de $\dfrac{c-a}{b-a}$ est une mesure de l'angle orienté $\left(\widehat{\overrightarrow{AB}, \overrightarrow{AC}}\right)$;
- si les points A, B, C sont deux à deux distincts, alors les points A, B, C sont alignés si, et seulement si, $b - a$ et $c - a$ ont même argument modulo π, ce qui équivaut encore à dire que $\dfrac{b-a}{c-a}$ est réel ou encore que $(b-a)\left(\overline{c-a}\right)$ est réel ;
- si les points A, B, C sont deux à deux distincts, alors les vecteurs $\overrightarrow{AB}$ et $\overrightarrow{AC}$ sont orthogonaux si, et seulement si, $\arg(b-a) \equiv \arg(c-a) + \dfrac{\pi}{2}$ modulo π, ce qui équivaut encore à dire que $\dfrac{b-a}{c-a}$ est imaginaire pur ou encore que $(b-a)\left(\overline{c-a}\right)$ est imaginaire pur ;
- si les points A, B, C, D sont deux à deux distincts, alors ces points sont cocycliques si, et seulement si, les angles $\left(\widehat{\overrightarrow{AC}, \overrightarrow{AD}}\right)$ et $\left(\widehat{\overrightarrow{BC}, \overrightarrow{BD}}\right)$ sont égaux modulo π, ce qui équivaut encore à dire que $\dfrac{d-a}{c-a}$ et $\dfrac{d-b}{c-b}$ ont même argument modulo π, ou encore que l'argument de $\dfrac{(d-a)(c-b)}{(c-a)(d-b)}$ est nul modulo π, ce qui revient à dire que $(d-a)(c-b)\left(\overline{c-a}\right)\left(\overline{d-b}\right)$ est réel ;
- pour tout entier naturel non nul n, $\dfrac{1}{n}\sum_{k=1}^{n} z_k$ est l'affixe de l'isobarycentre de la famille de points $(M_k)_{1 \leq k \leq n}$.

Les nombres complexes peuvent être utilisé pour décrire quelques transformations géométriques de $\mathcal{P}$. Ainsi :

- $z \mapsto z + a$ est la translation de vecteur $\overrightarrow{OA}$;
- $z \mapsto -z$ est la symétrie par rapport à O ;
- $z \mapsto \overline{z}$ est la symétrie orthogonale par rapport à l'axe O_x ;

- $z \mapsto -\overline{z}$ est la symétrie orthogonale par rapport à l'axe O_y ;
- pour tout réel $\rho > 0$, $z \mapsto \rho z$ est l'homothétie de rapport ρ et de centre O ;
- pour tout réel θ, $z \mapsto e^{i\theta} z$ est la rotation de centre O et d'angle θ ;
- pour tous nombres complexes a, b avec $a \notin \{0, 1\}$ d'argument θ, l'application $z \mapsto az + b$ est la composée commutative de la rotation d'angle θ et de centre $\Omega\left(\dfrac{b}{1-a}\right)$ et de l'homothétie de centre O et de rapport $|a|$, on dit que cette application est la similitude directe de centre Ω, de rapport $|a|$ et d'angle $\arg(a)$.

Les résultats de ce paragraphe sont supposés acquis à la sortie du lycée (peut-être suis-je optimiste ?).

Exercice 2.38 *Soit $\alpha = \rho e^{i\theta}$ un nombre complexe non nul et n un entier naturel non nul. Montrer que les racines n-ièmes de α se déduisent des racines n-ième de l'unité par une similitude directe de centre 0, de rapport $\sqrt[n]{\rho}$ et d'angle $\dfrac{\theta}{n}$.*

Solution 2.38 *Les racines n-ièmes de α sont les :*

$$z_k = \sqrt[n]{\rho} e^{i\frac{\theta+2k\pi}{n}} = \sqrt[n]{\rho} e^{i\frac{\theta}{n}} e^{i\frac{2k\pi}{n}} \quad (0 \leq k \leq n-1)$$

où les $e^{i\frac{2k\pi}{n}}$, pour k compris entre 0 et $n-1$, sont les racines n-ième de l'unité.

Exercice 2.39 *Soient $0 < a < b$ et $\lambda > 0$ des réels. Déterminer l'ensemble des nombres complexes z tels que :*

$$\frac{|z-a|}{|z-b|} = \lambda$$

(lignes de niveau de l'application $z \mapsto \dfrac{|z-a|}{|z-b|}$).

Solution 2.39 *On a nécessairement $z \neq b$. En écrivant $z = x + iy$, l'égalité $\dfrac{|z-a|}{|z-b|} = \lambda$ qui est équivalente à $|z-a|^2 = \lambda^2 |z-b|^2$ s'écrit aussi :*

$$(x-a)^2 + y^2 = \lambda^2 \left((x-b)^2 + y^2\right)$$

ou encore :

$$\left(\lambda^2 - 1\right)\left(x^2 + y^2\right) + 2\left(a - b\lambda^2\right)x + \left(\lambda^2 b^2 - a^2\right) = 0.$$

Si $\lambda = 1$, on obtient $2(a-b)x + (b^2 - a^2) = 0$, soit

$$x = \frac{a^2 - b^2}{2(a-b)} = \frac{a+b}{2}$$

et y est un réel quelconque. L'ensemble des points cherché est donc la droite d'équation $x = \dfrac{a+b}{2}$. C'est la médiatrice du segment $[AB]$, où $A = (a, 0)$ et $B = (b, 0)$.
Si $\lambda \neq 1$, on a :

$$\left(x^2 + y^2\right) + 2\frac{a - b\lambda^2}{\lambda^2 - 1}x + \frac{\lambda^2 b^2 - a^2}{\lambda^2 - 1} = 0$$

soit :

$$\begin{aligned}\left(x-\frac{a-b\lambda^2}{1-\lambda^2}\right)^2+y^2&=\left(\frac{a-b\lambda^2}{\lambda^2-1}\right)^2-\frac{\lambda^2b^2-a^2}{\lambda^2-1}\\&=\frac{1}{\left(\lambda^2-1\right)^2}\left(\left(a-b\lambda^2\right)^2-\left(\lambda^2-1\right)\left(\lambda^2b^2-a^2\right)\right)\\&=\frac{1}{\left(\lambda^2-1\right)^2}\left(a^2\lambda^2-2ab\lambda^2+b^2\lambda^2\right)\\&=\frac{\lambda^2\left(b-a\right)^2}{\left(\lambda^2-1\right)^2}\end{aligned}$$

L'ensemble des points cherché est donc le cercle de centre $\Omega=\left(\frac{a-b\lambda^2}{1-\lambda^2},0\right)$ *et de rayon* $\frac{\lambda\left(b-a\right)}{\left|\lambda^2-1\right|}$.
Par exemple l'ensemble des nombres complexes z *tels :*

$$\frac{|z-1|}{|z-2|}=\frac{1}{\sqrt{2}}$$

est le cercle de centre $\Omega=(0,0)$ *et de rayon* $\sqrt{2}$.

3

Espaces vectoriels réels

3.1 L'espace vectoriel $\mathbb{R}^n$

On peut utiliser les nombres réels $x \in \mathbb{R}$ pour représenter tous les points d'une droite, les couples de réels $(x, y) \in \mathbb{R}^2$ pour représenter tous les points d'un plan et les triplets de réels $(x, y, z) \in \mathbb{R}^3$ pour représenter tous les points d'un espace.

De manière plus générale, étant donné un entier naturel non nul n, on appelle vecteur tout élément du produit cartésien $\mathbb{R}^n = \mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}$ (répété n fois). Ces vecteurs sont des listes ordonnées de n réels $x_1, x_2, \cdots, x_n$ et seront notés $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ ou plus simplement $x = (x_i)_{1 \leq i \leq n}$ et on dit que les x_i sont les composantes du vecteur x.

La représentation sous forme de vecteurs colonnes sera justifiée plus loin par l'utilisation du calcul matriciel.

On peut naturellement munir cet ensemble d'une opération interne d'addition notée $+$ et définie par :

$$\forall x = (x_i)_{1 \leq i \leq n} \in \mathbb{R}^n,\ \forall y = (y_i)_{1 \leq i \leq n} \in \mathbb{R}^n,\ x + y = (x_i + y_i)_{1 \leq i \leq n}$$

On dit que cette opération est interne car elle associe à deux éléments x et y de $\mathbb{R}^n$ un élément $x + y$ de $\mathbb{R}^n$.

Des propriétés de l'addition des réels, on déduit facilement que cette opération d'addition vérifie les propriétés suivantes :

(*i*) elle est commutative, ce qui signifie que pour tous vecteurs x et y, on a $x + y = y + x$;

(*ii*) elle est associative, ce qui signifie que pour tous vecteurs x, y et z, on a $x + (y + z) = (x + y) + z$;

(*iii*) le vecteur nul $0 = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ est un élément neutre pour cette addition, ce qui signifie que pour tout vecteur x on a $x + 0 = 0 + x = x$;

(*iv*) pour tout vecteur $x = (x_i)_{1 \leq i \leq n} \in \mathbb{R}^n$, le vecteur $x' = (-x_i)_{1 \leq i \leq n}$ est tel que $x + x' = x' + x = 0$, on dit que x' est un opposé de x et on le note $-x$.

Tout cela se résume en disant que l'ensemble $\mathbb{R}^n$ muni de l'addition, que l'on note $(\mathbb{R}^n, +)$, est un groupe commutatif.

De même, on peut naturellement munir $\mathbb{R}^n$ d'une multiplication externe définie par :

$$\forall \lambda \in \mathbb{R},\ \forall x = (x_i)_{1\leq i\leq n} \in \mathbb{R}^n,\ \lambda \cdot y = (\lambda \cdot x_i)_{1\leq i\leq n}$$

On dit que cette opération est externe car elle associe à un réel λ (en dehors de $\mathbb{R}^n$) et à un élément x de $\mathbb{R}^n$ un élément $\lambda \cdot x$ de $\mathbb{R}^n$.

On écrira plus simplement λx pour $\lambda \cdot x$.

Là encore des propriétés de l'addition et de la multiplication des réels, on déduit que cette opération externe vérifie les propriétés suivantes :

(v) pour tout réel λ et tous vecteurs x et y, on a $\lambda(x+y) = \lambda x + \lambda y$;

(vi) pour tous réels λ, μ et tout vecteur x, on a $(\lambda+\mu)x = \lambda x + \mu x$;

(vii) pour tous réels λ, μ et tout vecteur x, on a $\lambda(\mu x) = (\lambda\mu)x$;

$(viii)$ pour tout vecteur x, on a $1x = x$.

Tout cela se résume en disant que l'ensemble $\mathbb{R}^n$ muni de cette addition interne et de cette multiplication externe, que l'on note $(\mathbb{R}^n, +, \cdot)$, est un espace vectoriel.

3.2 Définition d'un espace vectoriel réel

De manière plus générale, on donne la définition suivante.

Définition 3.1 *On appelle espace vectoriel réel tout ensemble non vide E muni d'une addition interne $(x,y) \in E \times E \mapsto x+y \in E$ et d'une multiplication externe $(\lambda, x) \in \mathbb{R} \times E \mapsto \lambda x \in E$ vérifiant les propriétés (i) à $(viii)$ précédentes.*

Pour simplifier, on dira espace vectoriel pour espace vectoriel réel.

Les éléments d'un espace vectoriel sont appelés vecteurs.

Dans un espace vectoriel l'élément neutre pour l'addition est noté 0 et on dit que c'est le vecteur nul et le symétrique d'un vecteur x est noté $-x$ et on dit que c'est l'opposé de x.

On vérifie facilement que le neutre 0 est unique, c'est-à-dire que c'est l'unique élément e de E tel que $x+e = e+x = x$ pour tout $x \in E$ (on a $e = e+0$ puisque 0 est neutre et $e+0 = 0$ puisque e est neutre, donc $e = 0$), que pour tout $x \in E$ l'opposé $-x$ est unique (si x' est un autre opposé, de $x+x' = 0$, on déduit que $(-x)+(x+x') = x' = (-x)+0 = -x$) et que tout élément de E est simplifiable, c'est-à-dire que pour tous x, y, z dans E l'égalité $x+y = x+z$ équivaut à $y = z$ (il suffit d'ajouter $-x$ aux deux membres de cette égalité).

Pour x, y dans E, la somme $x+(-y)$ est simplement notée $x-y$.

Exemple 3.1 *L'ensemble $\mathbb{C}$ des nombres complexes est un espace vectoriel réel.*

Exemple 3.2 *On rappelle qu'une suite réelle est une application u définie sur $\mathbb{N}$ et à valeurs réelles. L'ensemble de toutes les suites réelles est un espace vectoriel réel.*

Exemple 3.3 *Plus généralement, étant donné une partie I non vide de $\mathbb{R}$, l'ensemble E de toutes les applications de I dans $\mathbb{R}$ est un espace vectoriel réel.*

Exemple 3.4 *L'ensemble noté $\mathbb{R}[x]$ des fonctions polynomiales réelles (on dira plus simplement polynômes réel), c'est-à-dire l'ensemble des fonctions P définies par $P(x) = \sum_{k=0}^{n} a_k x^k$ pour tout réel x, où les coefficients a_k sont réels, est un espace vectoriel réel.*

Exemple 3.5 *L'ensemble $\mathcal{F}$ des polynômes trigonométriques, c'est-à-dire l'ensemble des fonctions P définies par $P(x) = \sum_{k=0}^{n} (a_k \cos(kx) + b_k \sin(kx))$ pour tout réel x, où les coefficients a_k et b_k sont réels, est un espace vectoriel réel.*

Exercice 3.1 *Montrer que dans un espace vectoriel E, l'égalité $\lambda x = 0$ où λ est un réel et x un vecteur est équivalente à $\lambda = 0$ ou $x = 0$.*

Solution 3.1 *Pour tout vecteur x, on a :*

$$0 \cdot x + x = 0 \cdot x + 1 \cdot x = (0+1)\, x = 1 \cdot x = x$$

et simplifiant par x (ce qui revient à ajouter $-x$ aux deux membres de cette égalité), on aboutit à $0 \cdot x = 0$.
De même, pour tout réel λ, on a :

$$\lambda \cdot 0 = \lambda \cdot (0+0) = \lambda \cdot 0 + \lambda \cdot 0$$

et simplifiant par $\lambda \cdot 0$, on aboutit à $\lambda \cdot 0 = 0$.
Supposons que $\lambda x = 0$. Si $\lambda = 0$ c'est terminé, sinon λ est inversible dans $\mathbb{R}$ et :

$$x = 1 \cdot x = \left(\frac{1}{\lambda}\lambda\right) x = \frac{1}{\lambda}(\lambda x) = \frac{1}{\lambda} 0 = 0.$$

Exercice 3.2 *Montrer que dans un espace vectoriel E, on a $(-1)\, x = -x$ pour tout vecteur x.*

Solution 3.2 *Pour tout réel x, on a :*

$$x + (-1)\, x = 1 \cdot x + (-1)\, x = (1-1)\, x = 0 \cdot x = 0$$

et $(-1)\, x = -x$ puisque l'opposé de x est unique.

Définition 3.2 *Soient E un espace vectoriel réel, n un entier naturel non nul et $x, y, x_1, \cdots, x_n$ des éléments de E.*
On dit que y est colinéaire à x s'il existe un réel λ tel que $y = \lambda x$.
Plus généralement, on dit que y est combinaison linéaire de $x_1, \cdots, x_n$ s'il existe des réels $\lambda_1, \cdots, \lambda_n$ tels que $y = \sum_{k=1}^{n} \lambda_k x_k$.

On peut remarquer que, par définition, un espace vectoriel réel est stable par combinaison linéaire, c'est-à-dire que si $x_1, \cdots, x_n$ sont dans E et $\lambda_1, \cdots, \lambda_n$ dans $\mathbb{R}$, alors la combinaison linéaire $\sum_{k=1}^{n} \lambda_k x_k$ est encore dans E.

En s'inspirant de la construction l'espace vectoriel $\mathbb{R}^n$ comme produit cartésien de p exemplaires de l'espace vectoriel $\mathbb{R}$, on vérifie facilement que le produit cartésien $F = F_1 \times \cdots \times F_p$ de p espaces vectoriels $F_1, \cdots, F_p$ est naturellement muni d'une structure d'espace vectoriel avec les lois définies par :

$$\begin{cases} x + y = (x_1 + y_1, \cdots, x_p + y_p) \\ \lambda x = (\lambda x_1, \cdots, \lambda x_p) \end{cases}$$

où $x = (x_1, \cdots, x_p)$, $y = (y_1, \cdots, y_p)$ sont deux éléments de F et λ un réel.

3.3 Sous-espaces vectoriels

Définition 3.3 *Soit E un espace vectoriel. On dit qu'une partie F de E est un sous-espace vectoriel de E si :*

1. *le vecteur 0 est dans F ;*
2. *pour tous vecteurs x, y dans F et tout réel λ, les vecteurs $x+y$ et λx sont dans F.*

L'appellation sous-espace vectoriel est justifiée par le résultat suivant.

Théorème 3.1 *Tout sous-espace vectoriel d'un espace vectoriel est un espace vectoriel.*

Démonstration. Soit F un sous-espace vectoriel d'un espace vectoriel E.
Comme F contient 0, il est non vide.
Le deuxième point de la définition nous dit que l'addition des vecteurs et la multiplication d'un vecteur par un réel restreintes à F y définissent bien respectivement une opération interne et externe.
L'addition des vecteurs qui est commutative sur E l'est en particulier sur F.
L'élément neutre 0 pour l'addition est bien dans F.
Tout vecteur $x \in F$ admet un opposé $-x \in E$ et en écrivant que $-x = (-1)\,x$, on voit que $-x$ est bien dans F.
Les propriétés (v) à $(viii)$ vérifiées dans E le sont en particulier dans F. ■
De manière équivalente, on peut dire qu'une partie F d'un espace vectoriel E est un sous-espace vectoriel si, et seulement si, F est non vide et pour tous vecteurs x, y dans F, tous réel λ, μ, le vecteur $\lambda x + \mu y$ est dans F.

Exercice 3.3 *Justifier l'affirmation précédente.*

Solution 3.3 *Laissée au lecteur.*

De manière plus générale, un sous-espace vectoriel d'un espace vectoriel est une partie non vide stable par combinaison linéaire.

Exercice 3.4 *Justifier l'affirmation précédente.*

Solution 3.4 *Laissée au lecteur.*

Exemple 3.6 *Si E un espace vectoriel, alors $\{0\}$ et E sont des sous-espaces vectoriels de E.*

Exemple 3.7 *$\mathbb{R}$ et l'ensemble des imaginaires purs sont des sous-espaces vectoriels réels de $\mathbb{C}$.*

Exercice 3.5 *Montrer que l'intersection de deux sous-espaces vectoriels d'un espace vectoriel E est un sous-espace vectoriel. Qu'en est-il de la réunion ?*

Solution 3.5 *Soient F, G deux sous-espaces vectoriels de E. L'intersection $H = F \cap G$ contient 0 puisqu'ils sont dans F et G et pour tous x, y dans H, λ, μ dans $\mathbb{R}$, le vecteur $\lambda x + \mu y$ est dans F et G, donc dans H. En définitive, H est un sous-espace vectoriel de E.*
En général la réunion de deux sous-espaces vectoriels de E n'est pas un sous-espace vectoriel. Par exemple dans $\mathbb{R}^2$, les ensembles F et G définis par $F = \left\{ \begin{pmatrix} x \\ 0 \end{pmatrix} \mid x \in \mathbb{R} \right\}$ (l'axe des x) et $G = \left\{ \begin{pmatrix} 0 \\ y \end{pmatrix} \mid x \in \mathbb{R} \right\}$ (l'axe des y) sont des sous-espaces vectoriels, mais pas $F \cup G$ puisque $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ et $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ sont dans cette réunion, mais pas leur somme $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Exercice 3.6 *Soient F, G deux sous-espaces vectoriels de E. Montrer que $F \cup G$ est un sous-espace vectoriel de E si et seulement si $F \subset G$ ou $G \subset F$.*

Solution 3.6 *Si $F \subset G$ [resp. $G \subset F$], on a alors $F \cup G = G$ [resp. $F \cup G = F$] et c'est un sous-espace vectoriel de E.*
Réciproquement supposons que $F \cup G$ soit un sous-espace vectoriel de E. Si $F \nsubseteq G$ et $G \nsubseteq F$, il existe alors $x \in F \setminus G$ et $y \in G \setminus F$, donc comme x et y sont dans $F \cup G$, il en est de même de $x + y$, mais $x + y \in F$ entraîne $y = (x + y) - x \in F$, ce qui n'est pas et $x + y \in G$ entraîne $x = (x + y) - y \in G$, ce qui n'est pas, il y a donc une impossibilité et on a nécessairement $F \subset G$ ou $G \subset F$.

Exercice 3.7 *On désigne par $\mathcal{F}$ l'espace vectoriel des fonctions de $\mathbb{R}$ dans $\mathbb{R}$ et par $\mathcal{P}$ [resp. $\mathcal{I}$] le sous-ensemble de $\mathcal{F}$ formés de toutes les fonctions paires [resp impaires] de $\mathbb{R}$ dans $\mathbb{R}$.*

1. *Montrer que $\mathcal{P}$ et $\mathcal{I}$ sont des sous-espaces vectoriels de $\mathcal{F}$.*
2. *Calculer $\mathcal{P} \cap \mathcal{I}$.*
3. *Montrer que toute fonction $f \in \mathcal{F}$, s'écrit de manière unique comme somme d'une fonction paire et d'une fonction impaire. Ce résultat se traduit en disant que $\mathcal{F}$ est somme directe des sous-espaces $\mathcal{P}$ et $\mathcal{I}$ et on note $\mathcal{F} = \mathcal{P} \oplus \mathcal{I}$.*

Solution 3.7

1. *La fonction nulle est à la fois paire et impaire donc dans $\mathcal{P}$ et dans $\mathcal{I}$. Si f, g sont deux fonctions paires [resp impaires], il en est alors de même de $f + g$ et de λf pour tout réel λ. Les ensembles $\mathcal{P}$ et $\mathcal{I}$ sont donc bien des sous-espaces vectoriels de $\mathcal{F}$.*
2. *Dire qu'une fonction f est dans $\mathcal{P} \cap \mathcal{I}$ signifie qu'elle est à la fois paire et impaire et donc que pour tout réel x, on a :*
$$f(x) = f(-x) = -f(x)$$
ce qui revient à dire que $f(x) = 0$. On a donc $\mathcal{P} \cap \mathcal{I} = \{0\}$.
3. *Pour toute fonction $f \in \mathcal{F}$, la fonction g [resp. h] définie sur $\mathbb{R}$ par :*
$$g(x) = \frac{f(x) + f(-x)}{2} \text{ [resp. } h(x) = \frac{f(x) - f(-x)}{2}\text{]}$$
est paire [resp. impaire] et $f = g + h$. Si (g', h') est un autre couple dans $\mathcal{P} \times \mathcal{I}$ tel que $f = g' + h'$, la fonction $g - g' = h' - h$ est dans $\mathcal{P} \cap \mathcal{I}$ donc nulle et $g = g'$, $h = h'$. Une telle écriture est donc unique.
Par exemple si f est la fonction exp, *les fonctions g et h sont les fonctions hyperboliques* ch *et* sh *définies par* $\operatorname{ch}(x) = \dfrac{e^x + e^{-x}}{2}$ *et* $\operatorname{sh}(x) = \dfrac{e^x - e^{-x}}{2}$.

Définition 3.4 *Dans l'espace $\mathbb{R}^n$, où n est un entier naturel non nul, on appelle droite vectorielle tout sous-ensemble de la forme :*
$$D = \{\lambda a \mid \lambda \in \mathbb{R}\}$$
où a est un vecteur non nul donné.

On notera $D = \mathbb{R}a$ une telle droite.
On a donc, en notant $a = (a_k)_{1 \leq k \leq n}$, pour tout $x = (x_k)_{1 \leq k \leq n} \in \mathbb{R}^n$:
$$(x \in D) \Leftrightarrow (\exists \lambda \in \mathbb{R} \mid \forall k \in \{1, 2, \cdots, n\},\ x_k = \lambda a_k)$$

Une telle représentation est appelée représentation paramétrique de la droite D.

Cette définition correspond bien à la notion de droite vectorielle du plan $\mathbb{R}^2$ ou de l'espace $\mathbb{R}^3$ étudiée en Lycée.

On vérifie facilement qu'une droite de $\mathbb{R}^n$ est un sous-espace vectoriel.

Pour $n = 1$, $D = \mathbb{R}$ est la seule droite vectorielle puisque, pour tout $a \neq 0$, tout réel x peut s'écrire $x = \dfrac{x}{a}a = \lambda a$, donc $\mathbb{R} \subset \mathbb{R}a \subset \mathbb{R}$ et $\mathbb{R} = \mathbb{R}a$.

Définition 3.5 *Dans l'espace $\mathbb{R}^n$, où $n \geq 2$, on appelle plan vectoriel tout sous-ensemble de la forme :*

$$P = \left\{\lambda a + \mu b \mid (\lambda, \mu) \in \mathbb{R}^2\right\}$$

où a et b sont deux vecteurs non colinéaires donnés.

On notera $P = \mathbb{R}a \oplus \mathbb{R}b$ un tel plan.

On a donc, en notant $a = (a_k)_{1\leq k\leq n}$ et $b = (b_k)_{1\leq k\leq n}$, pour tout $x = (x_k)_{1\leq k\leq n} \in \mathbb{R}^n$:

$$(x \in P) \Leftrightarrow (\exists\,(\lambda, \mu) \in \mathbb{R} \mid \forall k \in \{1, 2, \cdots, n\},\ x_k = \lambda a_k + \mu b_k)$$

Une telle représentation est appelée représentation paramétrique du plan P.

Là encore cette définition correspond bien à la notion de plan vectoriel de l'espace $\mathbb{R}^3$ étudiée en Lycée.

On vérifie facilement qu'un plan de $\mathbb{R}^n$ est un sous-espace vectoriel.

Une partie finie d'un espace vectoriel E distincte de $\{0\}$ n'est pas un sous-espace vectoriel, mais à partir d'une partie finie de vecteurs de E, on peut engendrer un sous-espace vectoriel en s'inspirant des définitions de droites et plans.

Théorème 3.2 *Soient E un espace vectoriel et $x_1, \cdots, x_n$ des éléments de E. L'ensemble F de toutes les combinaisons linéaires de $x_1, \cdots, x_n$ est un sous-espace vectoriel de E.*

Démonstration. L'ensemble F contient $0 = \sum_{k=1}^{n} 0 \cdot x_k$ et pour $x = \sum_{k=1}^{n} \lambda_k x_k$, $y = \sum_{k=1}^{n} \mu_k x_k$ dans F et λ, μ dans $\mathbb{R}$, on a :

$$\lambda x + \mu y = \sum_{k=1}^{n} (\lambda\lambda_k + \mu\mu_k)\, x_k \in F$$

donc F est bien un sous-espace vectoriel de E. ∎

Définition 3.6 *Avec les notations du théorème précédent, on dit que F est le sous-espace vectoriel de E engendré par $x_1, \cdots, x_n$ et on le note $F = \langle x_1, \cdots, x_n\rangle$, ou $F = \mathrm{Vect}\,\{x_1, \cdots, x_n\}$ ou encore $F = \sum_{k=1}^{n} \mathbb{R}x_k$.*

Remarque 3.1 *Si tous les x_k sont nuls, alors $F = \{0\}$.*

Exemple 3.8 *Dans l'espace $\mathbb{R}[x]$ des fonctions polynomiales, pour tout entier naturel non nul n, le sous-espace vectoriel engendré par $1, x, \cdots, x^n$ est formé de l'ensemble des polynômes de degré au plus égal à n, on le note $\mathbb{R}_n[x]$ (le cas $n = 0$ correspond aux polynômes constants).*

De manière un peu plus générale, on peut définir le sous-espace vectoriel d'un espace vectoriel E engendré par une famille X non vide de E (non nécessairement finie) comme l'ensemble $F = \text{Vect}(X)$ (ou $F = \langle X \rangle$) de toutes les combinaisons linéaires d'éléments de X. Un vecteur x de E est donc dans $\text{Vect}(X)$ si, et seulement si, il existe un entier $p \geq 1$, des vecteurs $x_1, \cdots, x_p$ dans X et des réels $\lambda_1, \cdots, \lambda_p$ tels que $x = \sum_{k=1}^{p} \lambda_k x_k$.

Le théorème qui suit nous donne deux définitions équivalentes de $\text{Vect}(X)$.

Théorème 3.3 *Si X est une partie non vide d'un espace vectoriel E, $\text{Vect}(X)$ est l'intersection de tous les sous-espaces vectoriels de E qui contiennent X. C'est aussi le plus petit (pour l'ordre défini par l'inclusion) sous-espace vectoriel de E qui contient X, c'est-à-dire que $\text{Vect}(X)$ contient X et est contenu dans tout sous-espace vectoriel de E qui contient X.*

On peut aussi définir des sous-espaces vectoriels de $\mathbb{R}^n$ en utilisant des équations linéaires (on verra plus loin, avec la notion de base, que cela est encore possible pour n'importe quel espace vectoriel).

Théorème 3.4 *Étant donnés un entier naturel non nul n et n réels non tous nuls $a_1, \cdots, a_n$, l'ensemble :*

$$F = \left\{ x = (x_i)_{1 \leq i \leq n} \in \mathbb{R}^n \mid \sum_{k=1}^{n} a_k x_k = 0 \right\}$$

est un sous-espace vectoriel de $\mathbb{R}^n$.

Démonstration. Il suffit de vérifier. ∎

Remarque 3.2 *En fait si tous les a_k sont nuls, F est encore défini et c'est $\mathbb{R}^n$ tout entier.*

Pour $n = 1$, on a $a_1 \neq 0$ et cet espace F est réduit à $\{0\}$.

Pour $n = 2$, on a $(a_1, a_2) \neq (0, 0)$ et supposant par exemple que $a_2 \neq 0$, l'équation $a_1 x_1 + a_2 x_2 = 0$ équivaut à $x_2 = -\dfrac{a_1}{a_2} x_1$, ce qui signifie que F est l'ensemble des vecteurs de la forme :

$$x = x_1 \begin{pmatrix} 1 \\ -\frac{a_1}{a_2} \end{pmatrix}$$

où x_1 décrit $\mathbb{R}$, ce qui équivaut encore à dire que F est l'ensemble des vecteurs de la forme :

$$x = \frac{x_1}{a_2} \begin{pmatrix} a_2 \\ -a_1 \end{pmatrix} = \lambda \begin{pmatrix} a_2 \\ -a_1 \end{pmatrix}$$

où λ décrit $\mathbb{R}$. En définitive F est la droite engendrée par $\begin{pmatrix} a_2 \\ -a_1 \end{pmatrix}$.

Pour $n = 3$, on a $(a_1, a_2, a_3) \neq (0, 0, 0)$ et supposant par exemple que $a_3 \neq 0$, l'équation $a_1 x_1 + a_2 x_2 + a_3 x_3 = 0$ équivaut à $x_2 = -\dfrac{a_1}{a_3} x_1 - \dfrac{a_2}{a_3} x_2$, ce qui signifie que F est l'ensemble des vecteurs de la forme :

$$x = x_1 \begin{pmatrix} 1 \\ 0 \\ -\frac{a_1}{a_3} \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ -\frac{a_2}{a_3} \end{pmatrix}$$

où (x_1, x_2) décrit $\mathbb{R}^2$, ce qui équivaut encore à dire que F est l'ensemble des vecteurs de la forme :

$$x = \frac{x_1}{a_3}\begin{pmatrix} a_3 \\ 0 \\ -a_1 \end{pmatrix} + \frac{x_2}{a_3}\begin{pmatrix} 0 \\ a_3 \\ -a_2 \end{pmatrix} = \lambda\begin{pmatrix} a_3 \\ 0 \\ -a_1 \end{pmatrix} + \mu\begin{pmatrix} 0 \\ a_3 \\ -a_3 \end{pmatrix}$$

où (λ, μ) décrit $\mathbb{R}^2$, les vecteurs $\begin{pmatrix} a_3 \\ 0 \\ -a_1 \end{pmatrix}$ et $\begin{pmatrix} 0 \\ a_3 \\ -a_3 \end{pmatrix}$ étant non colinéaires puisque $a_3 \neq 0$.

En définitive F est le plan engendré par $\begin{pmatrix} a_3 \\ 0 \\ -a_1 \end{pmatrix}$ et $\begin{pmatrix} 0 \\ a_3 \\ -a_3 \end{pmatrix}$.

De manière générale, on donne la définition suivante.

Définition 3.7 *Étant donné un entier naturel non nul n, on appelle hyperplan de $\mathbb{R}^n$ tout sous-espace vectoriel de la forme :*

$$H = \left\{ x = (x_i)_{1\leq i\leq n} \in \mathbb{R}^n \mid \sum_{k=1}^{n} a_k x_k = 0 \right\}$$

où les réels $a_1, \cdots, a_n$ ne sont pas tous nuls.

On dit aussi que H est l'hyperplan d'équation $\sum_{k=1}^{n} a_k x_k = 0$.

Dans l'espace $\mathbb{R}^3$ l'intersection de deux plans vectoriels distincts est une droite.

Plus généralement, on a le résultat suivant.

Théorème 3.5 *Étant donnés deux entiers naturels non nuls n et p, des réels $a_{1,1}, \cdots, a_{1,n}$, $\cdots$, $a_{p,1}, \cdots, a_{p,n}$, l'ensemble :*

$$F = \left\{ x = (x_i)_{1\leq i\leq n} \in \mathbb{R}^n \mid \sum_{k=1}^{n} a_{i,k} x_k = 0 \ \ (1 \leq i \leq p) \right\}$$

est un sous-espace vectoriel de $\mathbb{R}^n$.

On dit que F est le sous-espace vectoriels de $\mathbb{R}^n$ d'équations linéaires :

$$\begin{cases} \sum_{k=1}^{n} a_{1,k} x_k = 0 \\ \qquad \vdots \\ \sum_{k=1}^{n} a_{p,k} x_k = 0 \end{cases}$$

3.4 Applications linéaires

Pour ce paragraphe, on désigne par E et F deux espaces vectoriels. On note 0 le vecteur nul de E et aussi celui de F. En toute rigueur il faudrait noter 0_E et 0_F ces vecteurs nuls, mais en fonction du contexte on sait en général de quel vecteur nul il s'agit.

Définition 3.8 *On dit qu'une application u de E dans F est linéaire (ou que c'est un morphisme d'espaces vectoriels) si pour tous vecteurs x, y de E et tout réel λ, on a :*

$$\begin{cases} u(x+y) = u(x) + u(y) \\ u(\lambda x) = \lambda u(x) \end{cases}$$

Remarque 3.3 *Une application linéaire u de E dans F transforme 0_E en 0_F et l'opposé de x dans E en l'opposé de $u(x)$ dans F. En effet, on a :*

$$u(0) = u(0+0) = u(0) + u(0)$$

donc $u(0) = 0$ et :

$$0 = u(0) = u(x + (-x)) = u(x) + u(-x)$$

et donc $u(-x) = -u(x)$.

Exemple 3.9 *Pour tout réel λ, l'application $u : x \mapsto \lambda x$ est linéaire de E dans E. On dit que c'est l'homothétie de rapport λ. Pour $\lambda = 1$, cette application est l'application identité et on la note Id_E ou Id.*

Exemple 3.10 *Pour tout entier j compris entre 1 et n, l'application u définie sur $\mathbb{R}^n$ par $u(x) = x_j$, si $x = (x_i)_{1 \leq i \leq n}$ est linéaire de E dans $\mathbb{R}$. On dit que c'est la j-ième projection canonique.*

Exemple 3.11 *Une translation de vecteur non nul $x \mapsto x + a$, définie de E dans E, n'est pas linéaire.*

Exemple 3.12 *Étant donné un intervalle réel I, la dérivation $f \mapsto f'$ est linéaire de l'espace vectoriel des fonctions dérivables de I dans $\mathbb{R}$ dans l'espace vectoriel des fonctions de I dans $\mathbb{R}$.*

On notera $\mathcal{L}(E, F)$ l'ensemble de toutes les applications linéaires de E dans F.

Si u et v sont deux applications linéaires de E dans F, $u + v$ est l'application définie sur E par :

$$\forall x \in E,\ (u+v)(x) = u(x) + v(x)$$

et pour tout réel λ, λu est l'application définie sur E par :

$$\forall x \in E,\ (\lambda u)(x) = \lambda u(x).$$

Il est facile de vérifier que $u + v$ et λu sont aussi des application linéaires de E dans F. On a donc ainsi défini une addition interne sur $\mathcal{L}(E, F)$ et une multiplication externe. Le résultat qui suit se démontre alors facilement.

Théorème 3.6 *L'ensemble $\mathcal{L}(E, F)$ de toutes les applications linéaires de E dans F muni de ces deux opérations est un espace vectoriel réel.*

Dans le cas où $F = E$, on note plus simplement $\mathcal{L}(E)$ pour $\mathcal{L}(E, F)$.

Les éléments de $\mathcal{L}(E)$ sont aussi appelés endomorphismes de E.

La composition des applications permet aussi de construire des applications linéaires à partir d'applications linéaires données.

Théorème 3.7 *Soient E, F, G des espaces vectoriels, u une application linéaire de E dans F et v une application linéaire de F dans G. La composée $v \circ u$ est alors une application linéaire de E dans G.*

Démonstration. Il suffit de vérifier. ■

Théorème 3.8 *Si u est une application linéaire de E dans F, on a alors pour tout entier naturel non nul n, tous vecteurs $x_1, \cdots, x_n$ de E et tous réels $\lambda_1, \cdots, \lambda_n$:*

$$u\left(\sum_{k=1}^{n} \lambda_k x_k\right) = \sum_{k=1}^{n} \lambda_k u\left(x_k\right).$$

Démonstration. On procède par récurrence pour $n \geq 1$. Pour $n = 1$, on a bien $u(\lambda x) = \lambda u(x)$ pour tout réel λ et tout vecteur x par définition d'une application linéaire.
Pour $n = 2$, toujours par définition d'une application linéaire, on a pour tous vecteurs x, y et tous réels λ, μ :

$$u(\lambda x + \mu y) = u(\lambda x) + u(\mu y) = \lambda u(x) + \mu u(y).$$

Supposant le résultat acquis au rang $n \geq 2$, on se donne $n+1$ vecteurs $x_1, \cdots, x_n, x_{n+1}$ et $n+1$ réels $\lambda_1, \cdots, \lambda_n, \lambda_{n+1}$ et on a :

$$\begin{aligned} u\left(\sum_{k=1}^{n+1} \lambda_k x_k\right) &= u\left(\sum_{k=1}^{n} \lambda_k x_k\right) + u\left(\lambda_{n+1} x_{n+1}\right) \\ &= \sum_{k=1}^{n} \lambda_k u\left(x_k\right) + \lambda_{n+1} u\left(x_{n+1}\right) \\ &= \sum_{k=1}^{n+1} \lambda_k u\left(x_k\right) \end{aligned}$$

■

Définition 3.9 *Soit u une application linéaire de E dans F.*
Le noyau de u est l'ensemble :

$$\ker(u) = \{x \in E \mid u(x) = 0\}$$

et l'image de u l'ensemble :

$$\operatorname{Im}(u) = \{u(x) \mid x \in E\}.$$

Théorème 3.9 *Le noyau d'une application linéaire u de E dans F est un sous-espace vectoriel de E et son image un sous-espace vectoriel de F.*

Démonstration. On a vu que $\ker(u)$ contient 0 et pour x, y dans $\ker(u)$, λ, μ dans $\mathbb{R}$, on a :

$$u(\lambda x + \mu y) = \lambda u(x) + \mu u(y) = 0$$

ce qui signifie que $\lambda x + \mu y \in \ker(u)$. Donc $\ker(u)$ est bien un sous-espace vectoriel de E.
De manière analogue, en utilisant la linéarité de u, on montre que $\operatorname{Im}(u)$ est un sous-espace vectoriel de F. ■

Théorème 3.10 *Soit u une application linéaire de E dans F.*

1. *L'application u est injective si, et seulement si,* $\ker(u)$ *est réduit à* $\{0\}$.
2. *L'application u est surjective si, et seulement si,* $\mathrm{Im}(u) = F$.

Démonstration.

1. Supposons u injective. Pour tout $x \in \ker(u)$, on a $u(x) = u(0)$ et nécessairement $x = 0$ puisque u est injective. Donc $\ker(u) = \{0\}$.
   Réciproquement, supposons que $\ker(u) = \{0\}$. Si x, y dans E sont tels que $u(x) = u(y)$, on a alors $u(x-y) = u(x) - u(y) = 0$, c'est-à-dire que $x - y \in \ker(u)$ et donc $x = y$.
2. Ce résultat est en fait valable pour toute application de E dans F (la linéarité de u n'intervient pas ici).

■

Les applications linéaires de E dans $\mathbb{R}$ ont un statut particulier.

Définition 3.10 *On appelle forme linéaire sur E toute application linéaire de E dans* $\mathbb{R}$.

Exemple 3.13 *Étant donné un segment $I = [a, b]$ non réduit à un point, l'application* $f \mapsto \displaystyle\int_a^b f(x)\,dx$ *est une forme linéaire sur l'espace vectoriel des fonctions continues de I dans* $\mathbb{R}$.

Exercice 3.8 *Montrer qu'une forme linéaire φ sur E non identiquement nulle est surjective.*

Solution 3.8 *Dire que $\varphi \neq 0$ signifie qu'il existe un vecteur $x_0 \in E$ tel que $\lambda = \varphi(x_0) \neq 0$. Pour tout réel y, on peut alors écrire :*

$$y = \frac{y}{\lambda}\lambda = \frac{y}{\lambda}\varphi(x_0) = \varphi\left(\frac{y}{\lambda}x_0\right)$$

soit $y = \varphi(x)$ avec $x = \dfrac{y}{\lambda}x_0 \in E$, ce qui signifie que φ est surjective.

Exercice 3.9 *On se donne un intervalle réel I non réduit à un point et on désigne par E l'ensemble de toutes les fonctions dérivables de I dans $\mathbb{R}$ et par F l'ensemble de toutes les fonctions de I dans* $\mathbb{R}$.

1. *Montrer que E est un espace vectoriel.*
2. *Déterminer le noyau de l'application linéaire $u : f \mapsto f'$ où f' est la fonction dérivée de f.*

Solution 3.9 *Laissée au lecteur.*

Exercice 3.10 *Soit u une application linéaire de E dans E (i. e. un endomorphisme de E). Montrer que :*

$$\mathrm{Im}(u) \subset \ker(u) \Leftrightarrow u \circ u = 0.$$

Solution 3.10 *Si* $\mathrm{Im}(u) \subset \ker(u)$, *on a alors $u(x) \in \ker(u)$ pour tout $x \in E$ et $u(u(x)) = 0$, ce qui signifie que $u \circ u = 0$.*
Réciproquement si $u \circ u = 0$, pour tout $y = u(x) \in \mathrm{Im}(u)$, on a $u(y) = u(u(x)) = u \circ u(x) = 0$, ce qui signifie que $y \in \ker(u)$ et $\mathrm{Im}(u) \subset \ker(u)$.

Exercice 3.11 *On appelle projecteur de E tout endomorphisme p de E tel que $p \circ p = p$.*

1. *Montrer que* $\mathrm{Im}(p)$ *est l'ensemble des vecteurs invariants de p.*

2. *Montrer que* $\mathrm{Im}\,(p) \cap \ker(p) = \{0\}$.
3. *Montrer que tout vecteur* $x \in E$ *s'écrit de manière unique comme somme d'un vecteur de* $\ker(p)$ *et d'un vecteur de* $\mathrm{Im}\,(p)$ *(on dit que* E *est somme directe de* $\ker(p)$ *et* $\mathrm{Im}\,(p)$, *ce qui se note* $E = \ker(p) \oplus \mathrm{Im}\,(p)$*).*
4. *Montrer que si* p *est un projecteur, il en est alors de même de* $q = Id_E - p$ *et on a* $\ker(q) = \mathrm{Im}\,(p)$ *et* $\mathrm{Im}\,(q) = \ker(p)$.

Solution 3.11 *Laissée au lecteur.*

On rappelle que si u est une bijection de E sur F, elle admet alors une application réciproque notée u^{-1} et définie par :

$$\left(y \in F \text{ et } x = u^{-1}(y)\right) \Leftrightarrow (x \in E \text{ et } y = u(x)).$$

Cette application u^{-1} est aussi l'unique application de F dans E telle que $u \circ u^{-1} = Id_F$ et $u^{-1} \circ u = Id_E$.

Dans le cas où u est linéaire, il en est de même de u^{-1}. En effet si y, y' sont deux éléments de F, ils s'écrivent de manière unique $y = u(x)$, $y' = u(x')$ et on a :

$$\begin{aligned} u^{-1}(y + y') &= u^{-1}(u(x) + u(x')) \\ &= u^{-1}(u(x + x')) = x + x' = u^{-1}(y) + u^{-1}(y') \end{aligned}$$

et pour tout réel λ :

$$u^{-1}(\lambda y) = u^{-1}(\lambda u(x)) = u^{-1}(u(\lambda x)) = \lambda x = \lambda u^{-1}(y).$$

Définition 3.11 *On appelle isomorphisme de* E *sur* F *toute application linéaire bijective de* E *sur* F.

Dans le cas où $E = F$, un isomorphisme de E sur E est appelé automorphisme de E.

On note $GL(E)$ l'ensemble de tous les automorphismes de E et on dit que $GL(E)$ est le groupe linéaire de E (l'appellation groupe sera justifiée plus loin).

Exercice 3.12 *L'ensemble* $GL(E)$ *est-il un espace vectoriel ?*

Solution 3.12 *Non, sauf dans le cas où* $E = \{0\}$.

3.5 La base canonique de $\mathbb{R}^n$ et expression matricielle des applications linéaires de $\mathbb{R}^n$ dans $\mathbb{R}^m$

Tout vecteur $x = (x_i)_{1 \leq i \leq n}$ de $\mathbb{R}^n$ s'écrit $\sum_{k=1}^{n} x_k e_k$, où on a noté, pour tout entier k compris entre 1 et n, e_k le vecteur dont toutes les composantes sont nulles sauf la k-ième qui vaut 1. L'espace vectoriel $\mathbb{R}^n$ est donc engendré par les vecteurs $e_1, e_2, \cdots, e_n$. On dit que le système $(e_1, e_2, \cdots, e_n)$, que l'on note plus simplement $(e_k)_{1 \leq k \leq n}$, est un système générateur de $\mathbb{R}^n$. De plus par définition du produit cartésien $\mathbb{R}^n$ une telle écriture est unique, ce qui se traduit en disant que le système $(e_k)_{1 \leq k \leq n}$ est un système libre.

On dit que $(e_k)_{1 \leq k \leq n}$ est la base canonique de $\mathbb{R}^n$. Nous définirons plus loin la notion de base d'un espace vectoriel.

Si m est un autre entier naturel non nul, on note $(f_k)_{1 \leq k \leq m}$ la base canonique de $\mathbb{R}^m$.

Étant donnée une application linéaire u de E dans F, on a pour tout vecteur $x = \sum_{j=1}^{n} x_j e_j$ de $\mathbb{R}^n$:

$$u(x) = u\left(\sum_{j=1}^{n} x_j e_j\right) = \sum_{j=1}^{n} x_j u(e_j)$$

et chacun des vecteurs $u(e_j)$, pour j compris entre 1 et n, étant dans $\mathbb{R}^m$, il s'écrit :

$$u(e_j) = \sum_{i=1}^{m} a_{ij} f_i$$

On a donc en définitive :

$$u(x) = \sum_{j=1}^{n} x_j \sum_{i=1}^{m} a_{ij} f_i = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} x_j\right) f_i$$

ce qui signifie que les composantes du vecteurs $u(x)$ sont données par :

$$y_i = \sum_{j=1}^{n} a_{ij} x_j \ \ (1 \leq i \leq m) \tag{3.1}$$

Par exemple pour $n = m = 2$, on a :

$$\begin{cases} u(e_1) = a_{11} f_1 + a_{21} f_2 \\ u(e_2) = a_{12} f_1 + a_{22} f_2 \end{cases}$$

et :

$$u(x) = y_1 f_1 + y_2 f_2$$

avec :

$$\begin{cases} y_1 = a_{11} x_1 + a_{12} x_2 \\ y_2 = a_{21} x_1 + a_{22} x_2 \end{cases} \tag{3.2}$$

L'application linéaire u est donc uniquement déterminée par les 4 réels a_{11}, a_{12}, a_{21} et a_{22}. On stocke ces réels dans un tableau à 2 lignes et 2 colonnes :

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

où la première colonne $\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}$ est le vecteur $u(e_1)$ et la deuxième colonne $\begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}$ le vecteur $u(e_2)$. Un tel tableau est appelé matrice à 2 lignes et 2 colonnes ou plus simplement matrice 2×2.

On traduit les deux égalités de (3.2) en utilisant le produit matriciel :

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

De manière générale, une application linéaire u de $\mathbb{R}^n$ dans $\mathbb{R}^m$ est donc uniquement déterminée par la matrice A à m lignes et n colonnes :

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

où la colonne numéro j, pour j compris entre 1 et n, est le vecteur :

$$u(e_j) = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

Les égalités (3.1) se traduisent alors par le produit matriciel :

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}$$

qui s'effectue comme suit :

$$y_i = \begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} = \sum_{j=1}^{n} a_{ij} x_j$$

pour tout i compris entre 1 et m.

Tout cela est compacté en :

$$y = u(x) \Leftrightarrow y = Ax$$

et on dit que A est la matrice de u dans les bases canoniques de $\mathbb{R}^n$ et $\mathbb{R}^m$. Si $n = m$, on dit simplement que A est la matrice de $u \in \mathcal{L}(\mathbb{R}^n)$ dans la base canonique de $\mathbb{R}^n$.

Exercice 3.13 *Soit* $u \in \mathcal{L}(\mathbb{R}^3)$ *de matrice* $A = \begin{pmatrix} 4 & 2 & -4 \\ -6 & -4 & 6 \\ -1 & -1 & 1 \end{pmatrix}$ *dans la base canonique* (e_1, e_2, e_3). *Déterminer le noyau* u.

Solution 3.13 *L'image du vecteur* $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ *est le vecteur :*

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 4 & 2 & -4 \\ -6 & -4 & 6 \\ -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4x_1 + 2x_2 - 4x_3 \\ -6x_1 - 4x_2 + 6x_3 \\ -x_1 - x_2 + x_3 \end{pmatrix}$$

Dire que $x \in \mathbb{R}^3$ *est dans le noyau de* u *équivaut à dire que ses composantes sont solutions du système linéaire de* 3 *équations à* 3 *inconnues :*

$$\begin{cases} 4x_1 + 2x_2 - 4x_3 = 0 \\ -6x_1 - 4x_2 + 6x_3 = 0 \\ -x_1 - x_2 + x_3 = 0 \end{cases}$$

L'équation (3) *donne* $x_3 = x_1 + x_2$ *qui reporté dans la première donne* $x_2 = 0$ *et* $x_1 = x_3$. *Réciproquement tout vecteur vérifiant ces conditions est solution du système linéaire. Le noyau de* u *est donc :*

$$\ker(u) = \left\{ \begin{pmatrix} x_1 \\ 0 \\ x_1 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = x_1 (e_1 + e_3) \mid x \in \mathbb{R} \right\}$$

c'est donc la droite vectorielle engendré par le vecteur $e_1 + e_3$.

3.6 Matrices réelles

On note $\mathcal{M}_{m,n}(\mathbb{R})$ l'ensemble de toutes les matrices à m lignes et n colonnes et à coefficients réels.

Une matrice $\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$ dans $\mathcal{M}_{m,n}(\mathbb{R})$ sera notée $A = ((a_{i,j}))_{\substack{1\leq i\leq m \\ 1\leq j\leq n}}$, le premier indice i étant le numéro de ligne et le deuxième j, le numéro de colonne.

Pour $n = m$, on note $\mathcal{M}_n(\mathbb{R})$ l'ensemble $\mathcal{M}_{n,n}(\mathbb{R})$ et dit que c'est l'ensemble des matrices carrées réelles d'ordre n.

3.6.1 Opérations sur les matrices

Théorème 3.11 *Si u et v sont deux applications linéaires de $\mathbb{R}^n$ dans $\mathbb{R}^m$ ayant pour matrices respectives $A = ((a_{i,j}))_{\substack{1\leq i\leq m \\ 1\leq j\leq n}}$ et $B = ((b_{i,j}))_{\substack{1\leq i\leq m \\ 1\leq j\leq n}}$ dans les bases canoniques, alors l'application linéaire $u+v$ a pour matrice dans ces bases canoniques, la matrice $((a_{i,j}+b_{ij}))_{\substack{1\leq i\leq m \\ 1\leq j\leq n}}$. Et pour tout réel λ, la matrice de λu dans les bases canoniques est la matrice $((\lambda a_{i,j}))_{\substack{1\leq i\leq m \\ 1\leq j\leq n}}$.*

Démonstration. Résulte de :

$$\begin{aligned}(u+v)(e_j) &= u(e_j) + v(e_j) \\ &= \sum_{i=1}^{m} a_{ij} f_i + \sum_{i=1}^{m} b_{ij} f_i \\ &= \sum_{i=1}^{m} (a_{ij} + b_{ij}) f_i \ \ (1 \leq j \leq n)\end{aligned}$$

et de :

$$(\lambda u)(e_j) = \lambda u(e_j) = \lambda \sum_{i=1}^{m} a_{ij} f_i = \sum_{i=1}^{m} \lambda a_{ij} f_i \ \ (1 \leq j \leq n)$$

■

On définit donc naturellement la somme de deux matrices $n \times m$ et le produit d'une telle matrice par un réel par :

$$A + B = ((a_{i,j} + b_{ij}))_{\substack{1\leq i\leq m \\ 1\leq j\leq n}} \text{ et } \lambda A = ((\lambda a_{i,j}))_{\substack{1\leq i\leq m \\ 1\leq j\leq n}}$$

en utilisant les notations précédentes.

On vérifie alors facilement le résultat suivant.

Théorème 3.12 *L'ensemble $\mathcal{M}_{m,n}(\mathbb{R})$ des matrices réelles à m lignes et n colonnes est un espace vectoriel.*

On note 0 la matrice nulle, c'est-à-dire l'élément de $\mathcal{M}_{m,n}(\mathbb{R})$ dont toutes les composantes sont nulles et pour toute matrice $A = ((a_{i,j}))_{\substack{1\leq i\leq m \\ 1\leq j\leq n}}$, on note $-A = ((-a_{i,j}))_{\substack{1\leq i\leq m \\ 1\leq j\leq n}}$ l'opposé de A.

En explicitant la matrice d'une composée de deux applications linéaires on définira le produit de deux matrices.

On se donne donc une application linéaire v de $\mathbb{R}^n$ dans $\mathbb{R}^m$ et une application linéaire u de $\mathbb{R}^m$ dans $\mathbb{R}^r$ de matrices respectives $B = ((b_{ij}))_{\substack{1\leq i\leq m\\1\leq j\leq n}}$ dans $\mathcal{M}_{m,n}(\mathbb{R})$ et $A = ((a_{i,j}))_{\substack{1\leq i\leq r\\1\leq j\leq m}}$ dans $\mathcal{M}_{r,m}(\mathbb{R})$.

On note toujours $(e_k)_{1\leq k\leq n}$ la base canonique de $\mathbb{R}^n$, $(f_k)_{1\leq k\leq m}$ celle de $\mathbb{R}^m$ et $(g_k)_{1\leq k\leq r}$ est celle de $\mathbb{R}^r$.

La matrice $C = ((c_{i,j}))_{\substack{1\leq i\leq r\\1\leq j\leq n}} \in \mathcal{M}_{r,n}(\mathbb{R})$ de $u\circ v \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^r)$ est obtenu en calculant les composantes des vecteurs $u\circ v(e_j)$, pour j compris entre 1 et n, dans la base $(g_k)_{1\leq k\leq r}$.

On a :

$$u\circ v(e_j) = u(v(e_j)) = u\left(\sum_{k=1}^{m} b_{kj} f_k\right) = \sum_{k=1}^{m} b_{kj} u(f_k)$$

avec, pour k compris entre 1 et m :

$$u(f_k) = \sum_{i=1}^{r} a_{ik} g_i$$

ce qui donne :

$$u\circ v(e_j) = \sum_{k=1}^{m} b_{kj}\left(\sum_{i=1}^{r} a_{ik} g_i\right) = \sum_{i=1}^{r}\left(\sum_{k=1}^{m} a_{ik} b_{kj}\right) g_i$$

et signifie que les coefficients de la matrice C sont donnés par :

$$c_{ij} = \sum_{k=1}^{m} a_{ik} b_{kj} \ (1\leq i\leq r,\ 1\leq j\leq n).$$

Au vu de ce résultat, on donne la définition suivante.

Définition 3.12 *Étant données une matrice $A = ((a_{i,j}))_{\substack{1\leq i\leq r\\1\leq j\leq m}} \in \mathcal{M}_{r,m}(\mathbb{R})$ et une matrice $B = ((b_{ij}))_{\substack{1\leq i\leq m\\1\leq j\leq n}} \in \mathcal{M}_{m,n}(\mathbb{R})$, le produit AB de A par B est la matrice $C = ((c_{i,j}))_{\substack{1\leq i\leq r\\1\leq j\leq n}}$ de $\mathcal{M}_{r,n}(\mathbb{R})$ définie par :*

$$c_{ij} = \sum_{k=1}^{m} a_{ik} b_{kj} \ (1\leq i\leq r,\ 1\leq j\leq n).$$

Et nous venons de montrer le résultat suivant.

Théorème 3.13 *Si v est une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^m$ de matrice $B = ((b_{ij}))_{\substack{1\leq i\leq m\\1\leq j\leq n}}$ dans $\mathcal{M}_{m,n}(\mathbb{R})$ et u une application linéaire de $\mathbb{R}^m$ dans $\mathbb{R}^r$ de matrice $A = ((a_{i,j}))_{\substack{1\leq i\leq r\\1\leq j\leq m}}$ dans $\mathcal{M}_{r,m}(\mathbb{R})$ dans les bases canoniques, alors la matrice dans les bases canoniques de l'application linéaire $u\circ v$ de $\mathbb{R}^n$ dans $\mathbb{R}^r$ est la matrice produit $C = AB$.*

Il est important de remarquer que l'on ne peut définir le produit AB que si le nombre de colonnes de la matrice A est égal au nombre de lignes de la matrice B, ce produit est donc défini de $\mathcal{M}_{r,m}(\mathbb{R})\times\mathcal{M}_{m,n}(\mathbb{R})$ dans $\mathcal{M}_{r,n}(\mathbb{R})$.

Exercice 3.14 *Soient $A = \begin{pmatrix} 1 & 2 & 3\\ -1 & 2 & 1\end{pmatrix}$ et $B = \begin{pmatrix} 1 & -1\\ 0 & 2\\ 2 & 3\end{pmatrix}$. Calculer AB et BA.*

Solution 3.14 *On a :* $AB = \begin{pmatrix} 7 & 12 \\ 1 & 8 \end{pmatrix}$ *et* $BA = \begin{pmatrix} 2 & 0 & 2 \\ -2 & 4 & 2 \\ -1 & 10 & 9 \end{pmatrix}$.

Exercice 3.15 *Pour tout réel* θ*, on désigne par* M_θ *la matrice réelle :*

$$M_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

Montrer que pour tous réels θ *et* θ'*, on a* $M_\theta M_{\theta'} = M_{\theta'} M_\theta = M_{\theta+\theta'}$.

Solution 3.15 *Laissée au lecteur.*

Exercice 3.16 *On se place dans l'espace* $\mathcal{M}_2(\mathbb{R})$ *des matrices carrées d'ordre* 2 *où on note* $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ *la matrice identité d'ordre* 2.

1. *Donner des exemples de matrices* A *et* B *telles que* $AB = 0$ *et* $BA \neq 0$.
2. *Montrer que si* $AB = I_2$*, alors* $BA = I_2$.

Solution 3.16 *Laissée au lecteur.*

Exercice 3.17 *Déterminer dans la base canonique* (e_1, e_2) *de* $\mathbb{R}^2$ *la matrice de l'endomorphisme* u *(s'il existe) tel que* $u(e_1) = ae_1 - e_2$ *et* $u \circ u = u$ *où* a *est un réel donné.*

Solution 3.17 *Laissée au lecteur.*

Exercice 3.18 *Donner une condition nécessaire et suffisante portant sur les réels* a *et* b *pour que l'endomorphisme* u *de* $\mathbb{R}^2$ *de matrice* $A = \begin{pmatrix} a+1 & a \\ b & b+1 \end{pmatrix}$ *soit un automorphisme.*

Solution 3.18 *Laissée au lecteur.*

Exercice 3.19 *Déterminer, par leur matrice dans la base canonique, tous les endomorphismes non nuls de* $\mathbb{R}^2$ *tels que* $\operatorname{Im}(u) \subset \ker(u)$. *Si* u *est un tel endomorphisme donner la matrice de* $v = Id_E + u$ *et montrer que c'est un automorphisme de* E.

Solution 3.19 *Laissée au lecteur.*

L'opération de multiplication des matrices est une opération interne sur l'espace $\mathcal{M}_n(\mathbb{R})$ des matrices carrées réelles d'ordre n vérifiant les propriétés suivante :

- elle est associative, c'est-à-dire que pour toutes matrices A, B, C dans $\mathcal{M}_n(\mathbb{R})$, on a $A(BC) = (AB)C$;
- elle est distributive par rapport à l'addition, c'est-à-dire que pour toutes matrices A, B, C dans $\mathcal{M}_n(\mathbb{R})$, on a $A(B+C) = AB + AC$;
- la matrice

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

est l'élément neutre pour ce produit, c'est-à-dire que $A \cdot I_n = I_n \cdot A = A$ pour toute matrice A dans $\mathcal{M}_n(\mathbb{R})$.

Ces propriétés ajoutées à celle de l'addition des matrices se traduisent en disant que $(\mathcal{M}_n(\mathbb{R}), +, .)$ est un anneau unitaire.

L'associativité du produit matriciel dans $\mathcal{M}_n(\mathbb{R})$ permet de définir les puissances successives d'une matrice A par la relation de récurrence :

$$\begin{cases} A^0 = I_n \\ \forall p \in \mathbb{N},\ A^{p+1} = A^p A = A A^p. \end{cases}$$

Exercice 3.20 *Calculer A^n pour tout entier naturel n, où* $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$.

Solution 3.20 *On a $A^0 = I_3$, $A^1 = A$ et :*

$$A^2 = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix},\ A^3 = \begin{pmatrix} 1 & 3 & 6 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}.$$

En supposant que, pour $n \geq 1$, on a $A^n = \begin{pmatrix} 1 & n & a_n \\ 0 & 1 & n \\ 0 & 0 & 1 \end{pmatrix}$, *où a_n est un entier à déterminer, on a :*

$$A^{n+1} = \begin{pmatrix} 1 & n & a_n \\ 0 & 1 & n \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & n+1 & n+a_n+1 \\ 0 & 1 & n+1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & n & a_{n+1} \\ 0 & 1 & n \\ 0 & 0 & 1 \end{pmatrix}$$

avec :

$$a_{n+1} = a_n + n + 1.$$

La suite $(a_n)_{n\geq 1}$ est donc définie par la relation de récurrence :

$$\begin{cases} a_1 = 1 \\ \forall n \geq 1,\ a_{n+1} = a_n + n + 1 \end{cases}$$

ce qui donne :

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

(vérification par récurrence sur $n \geq 1$). On a donc, pour tout $n \geq 0$:

$$A^n = \begin{pmatrix} 1 & n & \frac{n(n+1)}{2} \\ 0 & 1 & n \\ 0 & 0 & 1 \end{pmatrix}$$

Remarque 3.4 *Le produit des matrices dans $\mathcal{M}_n(\mathbb{R})$ n'est pas commutatif. Par exemple pour* $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ *et* $B = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$, *on a :*

$$AB = \begin{pmatrix} 5 & 11 \\ 11 & 25 \end{pmatrix} \neq BA = \begin{pmatrix} 10 & 14 \\ 14 & 20 \end{pmatrix}.$$

On dira que deux matrices A et B dans $\mathcal{M}_n(\mathbb{R})$ commutent si $AB = BA$.

3.6.2 Matrices inversibles

Définition 3.13 *On dit qu'une matrice A dans $\mathcal{M}_n(\mathbb{R})$ est inversible s'il existe une matrice A' dans $\mathcal{M}_n(\mathbb{R})$ telle que $AA' = A'A = I_n$. On dit alors que A' est un inverse de A.*

Si une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est inversible son inverse est alors unique. En effet si A'' est un autre inverse de A, on a :

$$A'' = A''I_n = A''(AA') = (A''A)A' = I_nA' = A'.$$

On note A^{-1} l'inverse de A quand il existe.

On peut aussi remarquer qu'une matrice inversible A n'est jamais nulle puisque $AA^{-1} = I_n \neq 0$.

Exercice 3.21 *Montrer que si la matrice $A \in \mathcal{M}_n(\mathbb{R})$ a une colonne [resp. une ligne] nulle, alors elle n'est pas inversible.*

Solution 3.21 *En notant $C_1, \cdots, C_n$ [resp. $L_1, \cdots, L_n$] les colonne [resp. lignes] de A, on a pour toute matrice $A' \in \mathcal{M}_n(\mathbb{R})$:*

$$A'A = (A'C_1, \cdots, A'C_n)$$

et pour $C_j = 0$, la colonne j de $A'A$ est nulle, ce qui interdit l'égalité $A'A = I_n$.
Pour ce qui est des lignes, on écrit que :

$$AA' = \begin{pmatrix} L_1A' \\ \vdots \\ L_nA' \end{pmatrix}.$$

Exercice 3.22 *Montrer que pour tout réel θ, la matrice $M_\theta = \begin{pmatrix} \cos(\theta) & -\cos(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$ est inversible d'inverse $M_{-\theta}$.*

Solution 3.22 *Laissée au lecteur.*

Théorème 3.14 *Si $A \in \mathcal{M}_n(\mathbb{R})$ est inversible, alors A^{-1} est aussi inversible et $(A^{-1})^{-1} = A$. Le produit de deux matrices inversibles A et B est inversible et $(AB)^{-1} = B^{-1}A^{-1}$.*

Démonstration. Le premier point résulte de la définition et le deuxième de :

$$(AB)\left(B^{-1}A^{-1}\right) = A\left(BB^{-1}\right)A^{-1} = AI_nA^{-1} = AA^{-1} = I_n$$

et :

$$\left(B^{-1}A^{-1}\right)(AB) = B^{-1}\left(A^{-1}A\right)B = B^{-1}I_nB = B^{-1}B = I_n.$$

■

Par récurrence, on déduit que le produit de p matrices inversibles $A_1, \cdots, A_p$ est inversible avec $(A_1 \cdots A_p)^{-1} = A_p^{-1} \cdots A_1^{-1}$.

Exercice 3.23 *Soit $P \in \mathcal{M}_n(\mathbb{R})$ inversible. Montrer que $A \in \mathcal{M}_n(\mathbb{R})$ est inversible si, et seulement si, AP [resp. PA] est inversible.*

Solution 3.23 *Le théorème précédent nous dit que la condition est nécessaire.*
Réciproquement si AP [resp. PA] est inversible, alors $A = (AP)\,P^{-1}$ [resp. $A = P^{-1}\,(PA)$] est inversible.

Théorème 3.15 *Un endomorphisme u de $\mathbb{R}^n$ est bijectif si, et seulement si, sa matrice A dans la base canonique est inversible et dans ce cas A^{-1} est la matrice de u^{-1} dans la base canonique.*

Démonstration. Supposons u bijectif et notons A' la matrice de u^{-1} dans la base canonique de $\mathbb{R}^n$. De $u \circ u^{-1} = u^{-1} \circ u = Id$, on déduit que $AA' = A'A = I_n$, ce qui signifie que A est inversible d'inverse A'.

Réciproquement supposons A inversible et désignons par u' l'endomorphisme de $\mathbb{R}^n$ de matrice A^{-1} dans la base canonique. La matrice de $u \circ u'$ [resp. de $u' \circ u$] est $AA^{-1} = I_n$ [resp. $A^{-1}A = I_n$], donc $u \circ u' = I_d$ [resp. de $u' \circ u = I_d$] et u est surjective [resp. injective]. L'endomorphisme u est donc bijectif. ■

Exercice 3.24 *On désigne par $(e_i)_{1\leq i\leq n}$ la base canonique de $\mathbb{R}^n$ avec $n \geq 2$ et pour tout entier j compris entre 1 et n, par M_i la matrice :*

$$M_j = (0, \cdots, 0, e_1, 0, \cdots, 0)$$

la colonne e_1 étant en position j.
Montrer que pour tout réel λ et tout entier j compris entre 2 et n, la matrice :

$$P_j(\lambda) = I_n + \lambda M_j$$

est inversible et déterminer son inverse.

Solution 3.24 *Soit, pour j et λ fixés, u l'endomorphisme de $\mathbb{R}^n$ canoniquement associé à $P_j(\lambda)$. Il est défini par :*

$$u(e_k) = \begin{cases} \lambda e_1 + e_j \text{ si } k = j \\ e_k \text{ si } k \neq j \end{cases}$$

Cet endomorphisme est inversible d'inverse u' défini par :

$$u'(e_k) = \begin{cases} -\lambda e_1 + e_j \text{ si } k = j \\ e_k \text{ si } k \neq j \end{cases}$$

donc $P_j(\lambda)$ est inversible d'inverse $P_j(-\lambda)$.

Les matrices $P_j(\lambda)$ sont des matrices de transvection (paragraphe 5.1). On peut vérifier que la multiplication à gauche par une matrice de transvection $P_j(\lambda)$ a pour effet de remplacer la ligne L_1 de A par $L_1 + \lambda L_j$, les autres lignes étant inchangées et la multiplication à droite par une matrice de transvection $P_j(\lambda)$ a pour effet de remplacer la colonne C_j par $C_j + \lambda C_1$, les autres colonnes étant inchangées (théorème 5.1).

Théorème 3.16 *Une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est inversible si, et seulement si, l'unique solution du système linéaire $Ax = 0$ est $x = 0$.*

Démonstration. Si A est inversible et x est solution de $Ax = 0$, on a alors $A^{-1}Ax = x = 0$, donc 0 est l'unique solution de $Ax = 0$.

Pour la réciproque, on procède par récurrence sur $n \geq 1$.

On désigne par $(e_i)_{1\leq i\leq n}$ la base canonique de $\mathbb{R}^n$

Pour $n=1$, $A=(a)$ est inversible si, et seulement si, $a\neq 0$ et le résultat est évident.

Supposons le résultat acquis pour $n-1\geq 1$ et soit $A\in\mathcal{M}_n(\mathbb{R})$ telle que $x=0$ soit l'unique solution du système linéaire $Ax=0$.

La première colonne de A n'est pas nulle puisque c'est $C_1=Ae_1$ avec $e_1\neq 0$, il existe donc un indice j tel que $a_{j1}\neq 0$. Montrer que A est inversible équivaut à montrer que $\dfrac{1}{a_{j1}}A$ est inversible, ce qui nous ramène à $a_{j1}=1$. Si $a_{11}=0$, alors $j\geq 2$ et il est équivalent de montrer que la matrice $P_j(1)A$ (notations de l'exercice 3.24) est inversible, ce qui nous ramène à $a_{11}=1$ ($P_j(1)A$ se déduit de A en ajoutant la ligne j à la ligne 1). Il est encore équivalent de montrer que $AP_2(-a_{12})$ est inversible, ce qui ramène à $a_{12}=0$ et multipliant à droite par $P_j(-a_{1j})$ pour $2\leq j\leq n$, on se ramène à $a_{1j}=0$. En résumé, il suffit de considérer le cas où :

$$A=\begin{pmatrix}1 & 0\\ c & B\end{pmatrix}$$

avec $0\in\mathcal{M}_{1,n-1}(\mathbb{R})$, $c\in\mathcal{M}_{n-1,1}(\mathbb{R})$ et $B\in\mathcal{M}_{n-1}(\mathbb{R})$. Si $x'\in\mathbb{R}^{n-1}\setminus\{0\}$ est solution de $Bx'=0$, alors $x=\begin{pmatrix}0\\ x'\end{pmatrix}\in\mathbb{R}^n\setminus\{0\}$ est solution de $Ax=0$, ce qui contredit l'hypothèse de départ. Donc $x'=0$ est l'unique solution de $Bx'=0$ et B' est inversible. En posant $A'=\begin{pmatrix}1 & 0\\ d & B^{-1}\end{pmatrix}$ où $d\in\mathcal{M}_{n-1,1}(\mathbb{R})$ est à préciser, on a :

$$AA'=\begin{pmatrix}1 & 0\\ c & B\end{pmatrix}\begin{pmatrix}1 & 0\\ d & B^{-1}\end{pmatrix}=\begin{pmatrix}1 & 0\\ c+Bd & I_{n-1}\end{pmatrix}$$

et :

$$A'A=\begin{pmatrix}1 & 0\\ d & B^{-1}\end{pmatrix}\begin{pmatrix}1 & 0\\ c & B\end{pmatrix}=\begin{pmatrix}1 & 0\\ d+B^{-1}c & I_{n-1}\end{pmatrix}$$

soit $AA'=A'A=I_n$ en prenant $d=-B^{-1}c$. La matrice A est donc inversible. ■

Corollaire 3.1 *Un endomorphisme u de $\mathbb{R}^n$ est bijectif si, et seulement si, son noyau est réduit à $\{0\}$.*

Démonstration. Si A est la matrice canoniquement associée à u, on a $u(x)=Ax$ et $\ker(u)=\{0\}$ équivaut à dire que 0 est l'unique solution de $Ax=0$, ce qui équivaut à A inversible encore équivalent à dire que u est un isomorphisme. ■

Corollaire 3.2 *Une matrice $A\in\mathcal{M}_n(\mathbb{R})$ est inversible d'inverse $A'\in\mathcal{M}_n(\mathbb{R})$ si, et seulement si, $A'A=I_n$ [resp. $AA'=I_n$].*

Démonstration. Supposons que $A'A=I_n$. Si x est solution de $Ax=0$, on a alors $x=I_nx=A'(Ax)=A'0=0$ et A est inversible avec $A^{-1}=(A'A)A^{-1}=A'$.

Si $AA'=I_n$, la matrice A' est alors inversible d'inverse A, donc $A=(A')^{-1}$ est inversible d'inverse A'. ■

Exercice 3.25 *Montrer que si $A\in\mathcal{M}_n(\mathbb{R})$ a une colonne [resp. une ligne] nulle, alors elle n'est pas inversible.*

Solution 3.25 *Si la colonne j [resp. la ligne i] de A est nulle, alors pour toute matrice $A'\in\mathcal{M}_n(\mathbb{R})$, la matrice $A'A$ [resp. AA'] a sa colonne j [resp. sa ligne i] nulle. En conséquence il ne peut exister de matrice $A'\in\mathcal{M}_n(\mathbb{R})$ telle que $A'A=I_n$ [resp. $AA'=I_n$], ce qui signifie que A n'est pas inversible.*

Montrer qu'une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est inversible et calculer son inverse revient à montrer que pour tout vecteur $y \in \mathbb{R}^n$ le système linéaire de n équation à n inconnues $Ax = y$ a une unique solution et à exprimer cette solution x en fonction de y. Nous verrons plus loin comment l'algorithme de Gauss nous permet d'effectuer une telle résolution.

Exercice 3.26 *Montrer que la matrice* $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ *est inversible et déterminer son inverse.*

Solution 3.26 *Il s'agit de résoudre, pour* $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ *donné dans* $\mathbb{R}^2$, *le système :*

$$\begin{cases} x_1 + 2x_2 = y_1 \\ 3x_1 + 4x_2 = y_2 \end{cases}$$

Multipliant la première équation par 3 *et retranchant la deuxième équation au résultat obtenu, on a* $2x_2 = 3y_1 - y_2$, *soit* $x_2 = \frac{3}{2}y_1 - \frac{1}{2}y_2$ *qui reporté dans la première équation donne* $x_1 = -y_1 + y_2$. *On a donc :*

$$\begin{cases} x_1 = -2y_1 + y_2 \\ x_2 = \frac{3}{2}y_1 - \frac{1}{2}y_2 \end{cases}$$

ce qui signifie que A *est inversible et que :*

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

De manière plus générale, on a le résultat suivant.

Théorème 3.17 *Une matrice* $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbb{R})$ *est inversible si, et seulement si,* $ad - bc \neq 0$ *et dans ce cas son inverse est donné par :*

$$A^{-1} = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Démonstration. Pour tous réels a, b, c, d, on a :

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = (ad - bc)\, I_2$$

Si $ad - bc \neq 0$, cela s'écrit $AA' = I_2$ avec $A' = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$, ce qui signifie que A est inversible d'inverse A'. Réciproquement si A est inversible, on a :

$$\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = A^{-1}\left(A\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}\right) = (ad - bc)\, A^{-1}$$

donc $ad - bc \neq 0$ (puisque $A \neq 0$) et $A^{-1} = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$. ∎

Exercice 3.27 *Soit* $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$. *Pour tout entier* $n \geq 1$, *on désigne par* A^n *la matrice* $A \cdot A \cdot \cdots \cdot A$, *ce produit ayant* n *termes. On note* $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ *et par convention* $A^0 = I_2$.

1. *Calculer* A^2 *et* A^3.
2. *Montrer que pour tout entier* $n \geq 1$, *la matrice* A^n *est de la forme :*
$$A^n = \begin{pmatrix} a_n + (-1)^n & a_n \\ a_n & a_n + (-1)^n \end{pmatrix}$$
où a_n *est un entier.*
3. *Calculer* $A^2 - 2A - 3I_2$.
4. *Montrer que* A *est inversible et calculer* A^{-1}.
5. *Montrer que pour tout entier* $n \geq 2$, *il existe un polynôme* Q_n *et deux entiers* α_n *et* β_n *tels que :*
$$X^n = Q_n(X)\left(X^2 - 2X - 3\right) + \alpha_n X + \beta_n. \tag{3.3}$$
6. *En évaluant* (3.3) *en* -1 *et* 3 *déterminer* α_n *et* β_n.
7. *En déduire* A^n *pour tout* $n \geq 2$.

Solution 3.27

1. *On a :*
$$A^2 = \begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix},\ A^3 = \begin{pmatrix} 13 & 14 \\ 14 & 13 \end{pmatrix}$$
2. *C'est vrai pour* $n = 1$ *avec* $a_1 = 2$ *et en supposant le résultat acquis pour* n, *on a :*
$$\begin{aligned} A^{n+1} &= \begin{pmatrix} a_n + (-1)^n & a_n \\ a_n & a_n + (-1)^n \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 3a_n + (-1)^n & 3a_n + 2(-1)^n \\ 3a_n + 2(-1)^n & 3a_n + (-1)^n \end{pmatrix} \\ &= \begin{pmatrix} 3a_n + 2(-1)^n + (-1)^{n+1} & 3a_n + 2(-1)^n \\ 3a_n + 2(-1)^n & 3a_n + 2(-1)^n + (-1)^{n+1} \end{pmatrix} \end{aligned}$$
puisque :
$$2(-1)^n + (-1)^{n+1} = (-1)^n(2-1) = (-1)^n$$
3. *On a* $A^2 - 2A - 3I_2 = 0$.
4. *On a :*
$$A^{-1} = \frac{1}{3}\begin{pmatrix} -1 & 2 \\ 2 & -1 \end{pmatrix}$$
par calcul direct ou avec la question précédente :
$$A^{-1} = \frac{1}{3}(A - 2I_2).$$
5. *Pour* $n = 2$, *on a :*
$$X^2 = \left(X^2 - 2X - 3\right) + 2X + 3$$
donc $Q_2 = 1$ *et* $(\alpha_2, \beta_2) = (2, 3)$. *Supposant le résultat acquis pour* $n \geq 2$, *on a :*
$$\begin{aligned} X^{n+1} &= XQ_n(X)\left(X^2 - 2X - 3\right) + \alpha_n X^2 + \beta_n X \\ &= XQ_n(X)\left(X^2 - 2X - 3\right) + \alpha_n\left(\left(X^2 - 2X - 3\right) + 2X + 3\right) + \beta_n X \\ &= (\alpha_n + XQ_n)(X)\left(X^2 - 2X - 3\right) + (2\alpha_n + \beta_n)X + 3\alpha_n \end{aligned}$$
soit le résultat au rang $n + 1$.

6. -1 *et* 3 *sont les racines de* X^2-2X-3, *donc :*

$$\begin{cases} -\alpha_n+\beta_n=(-1)^n \\ 3\alpha_n+\beta_n=3^n \end{cases}$$

et résolvant le système :

$$\begin{cases} \alpha_n=\dfrac{3^n-(-1)^n}{4} \\ \beta_n=\dfrac{3^n+3\,(-1)^n}{4} \end{cases}$$

7. On a :

$$A^n=\alpha_nA+\beta_nI_2=\frac{1}{4}\left(\left(3^n-(-1)^n\right)A+\left(3^n+3\,(-1)^n\right)I_2\right)$$
$$=\frac{1}{4}\begin{pmatrix} 3^n-(-1)^n+3^n+3\,(-1)^n & 2\left(3^n-(-1)^n\right) \\ 2\left(3^n-(-1)^n\right) & 3^n-(-1)^n+3^n+3\,(-1)^n \end{pmatrix}$$
$$=\frac{1}{2}\begin{pmatrix} 3^n+(-1)^n & 3^n-(-1)^n \\ 3^n-(-1)^n & 3^n+(-1)^n \end{pmatrix}=\begin{pmatrix} a_n+(-1)^n & a_n \\ a_n & a_n+(-1)^n \end{pmatrix}$$

avec :

$$a_n=\frac{3^n-(-1)^n}{2}\ \textit{et}\ a_n+(-1)^n=\frac{3^n+(-1)^n}{2}.$$

3.6.3 Déterminant d'une matrice d'ordre 2

Définition 3.14 *Le déterminant d'une matrice* $A=\begin{pmatrix} a & b \\ c & d \end{pmatrix}\in\mathcal{M}_2(\mathbb{R})$ *est le réel :*

$$\det(A)=ad-bc.$$

Ce déterminant est aussi noté :

$$\det(A)=\begin{vmatrix} a & b \\ c & d \end{vmatrix}.$$

Le théorème 3.17 nous dit qu'une matrice $A\in\mathcal{M}_2(\mathbb{R})$ est inversible si, et seulement si, son déterminant est non nul. Nous généraliserons plus loin cette définition du déterminant.

Avec le théorème qui suit on résume les propriétés fondamentales du déterminant des matrices d'ordre 2.

Théorème 3.18 *On désigne par* A,B *des matrices réelles d'ordre* 2.

1. $\det(I_2)=1$.
2. *Pour tout réel* λ, *on a,* $\det(\lambda A)=\lambda^2\det(A)$.
3. $\det(AB)=\det(A)\det(B)$.
4. *Si* A *est inversible, alors* $\det(A^{-1})=\dfrac{1}{\det(A)}$.
5. *Si l'une des lignes [resp. des colonnes] de* A *est nulle, alors* $\det(A)=0$.
6. *Si* A' *est la matrice déduite de* A *en permutant les deux lignes [resp. les deux colonnes], alors* $\det(A')=-\det(A)$.

Démonstration. On note $A=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ et $B=\begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}$.

1. Il suffit de vérifier.
2. On a $\lambda A = \begin{pmatrix} \lambda a & \lambda b \\ \lambda c & \lambda d \end{pmatrix}$ et :
$$\det(\lambda A) = \lambda^2 ad - bc = \lambda^2 \det(A).$$
3. On a :
$$AB = \begin{pmatrix} aa' + bc' & ab' + bd' \\ ca' + dc' & cb' + dd' \end{pmatrix}$$
et :
$$\begin{aligned} \det(AB) &= (aa' + bc')(cb' + dd') - (ab' + bd')(ca' + dc') \\ &= aa'cb' + aa'dd' + bc'cb' + bc'dd' - ab'ca' - ab'dc' - bd'ca' - bd'dc' \\ &= aa'dd' + bc'cb' - ab'dc' - bd'ca' \\ &= ad(a'd' - b'c') - bc(a'd' - b'c') \\ &= (ad - bc)(a'd' - b'c') = \det(A)\det(B). \end{aligned}$$
4. Dans le cas où A est inversible, on $AA^{-1} = I_2$ et :
$$\det(A)\det\left(A^{-1}\right) = \det\left(AA^{-1}\right) = \det(I_2) = 1,$$
ce qui donne $\det(A^{-1}) = \dfrac{1}{\det(A)}$.
5. Résulte de la définition.
6. On a $A' = \begin{pmatrix} c & d \\ a & b \end{pmatrix}$ [resp. $A' = \begin{pmatrix} c & d \\ a & b \end{pmatrix}$] et :
$$\det(A') = cb - ad = -\det(A).$$

■

3.6.4 Transposée d'une matrice

Définition 3.15 *Si $A = ((a_{ij}))_{\substack{1\le i\le n \\ 1\le j\le m}}$ est une matrice à n lignes et m colonnes, la transposée de A est la matrice à m lignes et n colonnes ${}^tA = \left(\left(a'_{ij}\right)\right)_{\substack{1\le i\le m \\ 1\le j\le n}}$ où $a'_{ij} = a_{ji}$.*

La transposée d'un vecteur colonne $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ est le vecteur ligne ${}^tX = (x_1, x_2, \cdots, x_n)$.

En représentant $A = ((a_{ij}))_{\substack{1\le i\le n \\ 1\le j\le m}}$ sous forme de lignes $A = \begin{pmatrix} L_1 \\ L_2 \\ \vdots \\ L_n \end{pmatrix}$ [resp. de colonnes $M = (C_1, C_2, \cdots, C_m)$] où :
$$L_i = (a_{i1}, a_{i2}, \cdots, a_{im}) \text{ [resp. } C_j = \begin{pmatrix} a_{1,j} \\ a_{2,j} \\ \vdots \\ a_{n,j} \end{pmatrix}]$$

est la ligne numéro i [resp. la colonne numéro j] de M, on a :

$$ {}^tA = \left({}^tL_1, {}^tL_2, \cdots, {}^tL_n\right) \text{ [resp. } {}^tA = \begin{pmatrix} {}^tC_1 \\ {}^tC_2 \\ \vdots \\ {}^tC_m \end{pmatrix}\text{]} $$

Exemple 3.14 *La transposée d'une matrice carrée triangulaire supérieure [resp. inférieure] est triangulaire inférieure [resp. supérieure].*

Définition 3.16 *On dit qu'une matrice carrée $A = ((a_{ij}))_{1\leq i,j\leq n}$ est symétrique si elle est égale à sa transposée, ce qui revient à dire que $a_{ij} = a_{ji}$ pour tous i, j compris entre 1 et n.*

Définition 3.17 *On dit qu'une matrice carrée $A = ((a_{ij}))_{1\leq i,j\leq n}$ est anti-symétrique si ${}^tA = -A$, ce qui revient à dire que $a_{ij} = -a_{ji}$ pour tous i, j compris entre 1 et n.*

Remarque 3.5 *Une matrice anti-symétrique a tous ses termes diagonaux nulles. En effet, pour tout i compris entre 1 et n, on $a_{ii} = -a_{ii}$ et en conséquence, $a_{ii} = 0$.*

Théorème 3.19 *Pour toutes matrices $A = ((a_{i,j}))_{\substack{1\leq i\leq n\\1\leq j\leq m}}$ et $B = ((b_{ij}))_{\substack{1\leq i\leq n\\1\leq j\leq m}}$ dans $\mathcal{M}_{n,m}(\mathbb{R})$ et tout réel λ, on a :*

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

Pour toutes matrices $A = ((a_{i,j}))_{\substack{1\leq i\leq p\\1\leq j\leq n}}$ dans $\mathcal{M}_{p,n}(\mathbb{R})$ et $B = ((b_{ij}))_{\substack{1\leq i\leq n\\1\leq j\leq m}}$ dans $\mathcal{M}_{n,m}(\mathbb{R})$:

$$ {}^t(AB) = {}^tB\, {}^tA $$

Si $A \in \mathcal{M}_n(\mathbb{R})$ est inversible, alors ${}^tA$ est aussi inversible et :

$$ \left({}^tA\right)^{-1} = {}^t\left(A^{-1}\right) $$

Démonstration. On a ${}^tA = \left(\left(a'_{ij}\right)\right)_{\substack{1\leq i\leq m\\1\leq j\leq n}}$ avec $a'_{ij} = a_{ji}$ et ${}^t({}^tA) = {}^tA' = \left(\left(a''_{ij}\right)\right)_{\substack{1\leq i\leq n\\1\leq j\leq m}}$ avec $a''_{ij} = a'_{ji} = a_{ij}$. Donc ${}^t({}^tA) = A$.

Les résultats sur les combinaisons linéaires de matrices sont évidents.

Les coefficients de $C = AB$ sont les $c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$ et ceux de ${}^tC$ les :

$$ c'_{ij} = c_{ji} = \sum_{k=1}^{n} a_{jk}b_{ki} = \sum_{k=1}^{n} b'_{ik}a'_{kj} $$

et on reconnaît là les coefficients de ${}^tB\, {}^tA$.

Si $A \in \mathcal{M}_n(\mathbb{R})$ est inversible, on a :

$$ I_n = {}^tI_n = {}^t\left(AA^{-1}\right) = {}^t\left(A^{-1}\right)\ {}^tA $$

ce qui signifie que ${}^tA$ est inversible avec $\left({}^tA\right)^{-1} = {}^t\left(A^{-1}\right)$. ∎

Cette notion de matrice transposée nous sera utile lors de l'étude des formes bilinéaires et quadratiques.

3.6.5 Trace d'une matrice carrée

Définition 3.18 *La trace d'une matrice carrée* $A = ((a_{i,j}))_{1\leq i,j\leq n} \in \mathcal{M}_n(\mathbb{R})$ *est le réel :*

$$\mathrm{Tr}(A) = \sum_{i=1}^{n} a_{ii}$$

(somme des termes diagonaux).

Exemple 3.15 *La trace de la matrice identité* I_n *est* $\mathrm{Tr}(I_n) = n$.

Théorème 3.20 *L'application trace est linéaire de* $\mathcal{M}_n(\mathbb{R})$ *dans* $\mathbb{R}$ *(on dit que c'est une forme linéaire) et pour toutes matrices* A, B *dans* $\mathcal{M}_n(\mathbb{R})$, *on a* $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$.

Exercice 3.28 *Montrer qu'une matrice et sa transposée ont même trace.*

Exercice 3.29 *Calculer* $\mathrm{Tr}({}^tAA)$ *pour* $A = ((a_{i,j}))_{1\leq i,j\leq n} \in \mathcal{M}_n(\mathbb{R})$.

3.7 Systèmes d'équations linéaires

On se donne une matrice $A = ((a_{ij}))_{\substack{1\leq i\leq m\\1\leq j\leq n}}$ dans $\mathcal{M}_{m,n}(\mathbb{R})$, un vecteur $b = (b_i)_{1\leq i\leq m}$ dans $\mathbb{R}^m$ et on s'intéresse au système linéaire $Ax = b$ d'inconnue $x = (x_i)_{1\leq i\leq n}$ dans $\mathbb{R}^n$. Un tel système s'écrit :

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{cases}$$

Pour $b = 0$, un tel système a au moins $x = 0$ comme solution. Le système $Ax = 0$ est le système homogène associé au système $Ax = b$.

En notant, pour j compris entre 1 et n, $C_j = (a_{ij})_{1\leq i\leq m}$ la colonne numéro j de la matrice A, résoudre le système $Ax = b$ revient à trouver tous les réels $x_1, \cdots, x_n$ tels que :

$$x_1C_1 + x_2C_2 + \cdots + x_nC_n = b$$

Dans le cas où le nombre d'inconnues n est strictement supérieur au nombre d'équations m, le système homogène $Ax = 0$ a une infinité de solutions.

Théorème 3.21 *Pour toute matrice* $A = ((a_{ij}))_{\substack{1\leq i\leq m\\1\leq j\leq n}} \in \mathcal{M}_{m,n}(\mathbb{R})$ *avec* $n > m$, *le système homogène* $Ax = 0$ *a une infinité de solutions.*

Démonstration. On sait déjà que $x = 0$ est solution.

On désigne par $B = \begin{pmatrix} A \\ 0_{n-m,n} \end{pmatrix} \in \mathcal{M}_n(\mathbb{R})$ la matrice carrée d'ordre n ayant pour m premières lignes celles de A, les suivantes étant nulles. Pour toute matrice $B' \in \mathcal{M}_n(\mathbb{R})$ la matrice BB' a sa dernière ligne nulle et ne peut donc égaler I_n, ce qui signifie que la matrice B n'est pas inversible. Il existe donc $x \in \mathbb{R}^n \setminus \{0\}$ tel que $Bx = \begin{pmatrix} Ax \\ 0_{n-m,n} \end{pmatrix} = 0$ (théorème 3.16). Le vecteur x est donc solution non nulle de $Ax = 0$ et la droite dirigée par x nous donne une infinité de solutions de ce système linéaire. ∎

Exercice 3.30 *Résoudre le système linéaire :*

$$\begin{cases} 2x+y-z=0 \ (1) \\ x+y+z=0 \ (2) \end{cases}$$

Solution 3.28 *On élimine l'inconnue z en additionnant les deux équations, ce qui donne $3x+2y=0$, soit $y=-\dfrac{3}{2}x$ qui reporté dans (1) nous donne $z=\dfrac{1}{2}x$. Une solution de ce système est donc de la forme $X=\dfrac{x}{2}u$ où u est le vecteur $u=\begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix}$ et x un réel. Réciproquement l'égalité $Au=0$ nous dit que tout vecteur X colinéaire à u est solution du système.*
En définitive l'ensemble des solutions de ce système est la droite $D=\mathbb{R}u$ dirigée par u.

Théorème 3.22 *Soit A dans $\mathcal{M}_n(\mathbb{R})$. Le système $Ax=b$ a une unique solution dans $\mathbb{R}^n$ pour tout vecteur b, si et seulement si, la matrice A est inversible.*

Démonstration. Supposons A inversible. Le vecteur $A^{-1}b$ est solution de $Ax=b$ et si x est solution de ce système, on a alors $A^{-1}(Ax)=A^{-1}b$, soit $x=A^{-1}b$. Notre système a bien une unique solution.

Réciproquement, supposons que, pour tout $b\in\mathbb{R}^n$ le système $Ax=b$ a unique solution. En désignant par $(e_j)_{1\leq j\leq n}$ la base canonique de $\mathbb{R}^n$ et, pour tout j compris entre 1 et n, par C_j la solution de $Ax=e_j$, la matrice $A'=(C_1,\cdots,C_n)$ est telle que :

$$AA'=(AC_1,\cdots,AC_n)=(e_1,\cdots,e_n)=I_n$$

ce qui signifie que A est inversible d'inverse A'. ■

La méthode des pivots de Gauss peut être utilisée pour résoudre un tel système. Nous décrivons dans un premier temps cette méthode sur un exemple avec l'exercice qui suit.

Exercice 3.31 *Résoudre le système linéaire :*

$$\begin{cases} x+y+z=3 \ (1) \\ 2x+y+z=4 \ (2) \\ x-y+z=1 \ (3) \end{cases}$$

Solution 3.29 *La première étape consiste à éliminer x dans les équations (2) et (3). Pour ce faire on remplace l'équation (2) par $(2)-2(1)$ et l'équation (3) par $(3)-(1)$, ce qui donne le système :*

$$\begin{cases} x+y+z=3 \ (1) \\ -y-z=-2 \ (2) \\ -2y=-2 \ (3) \end{cases}$$

La deuxième étape consiste à éliminer y dans l'équation (3) en remplaçant cette équation par $(3)-2(2)$, ce qui donne :

$$\begin{cases} x+y+z=3 \ (1) \\ -y-z=-2 \ (2) \\ 2z=2 \ (3) \end{cases}$$

Le système obtenue est alors un système triangulaire et il se résout en remontant les équations, ce qui donne :

$$\begin{cases} z=1 \\ y=2-z=1 \\ x=3-y-z=1 \end{cases}$$

3.8 Sommes et sommes directes de sous-espaces vectoriels

On se donne pour ce paragraphe un espace vectoriel réel E.

Définition 3.19 *Soient F et G deux sous-espaces vectoriels de E. On dit que E est somme des espaces F et G si l'application :*

$$\begin{array}{rccc} \varphi & F\times G & \to & E \\ & (x,y) & \mapsto & x+y \end{array}$$

est surjective et on note alors $E = F + G$.
Si cette application est bijective, on dit que E est somme directe des espaces F et G et on note $E = F \oplus G$.

Dire que $E = F + G$ signifie donc que l'on peut écrire tout vecteur $x \in E$ sous la forme $x = y + z$, où $y \in F$ et $z \in G$ et dire que $E = F \oplus G$ signifie donc que l'on peut écrire de manière unique tout vecteur $x \in E$ sous la forme $x = y + z$, où $y \in F$ et $z \in G$.

Théorème 3.23 *Soient F et G deux sous-espaces vectoriels de E. On a $E = F \oplus G$ si, et seulement si, $E = F + G$ et $F \cap G = \{0\}$.*

Démonstration. Supposons que $E = F \oplus G$, on a alors $E = F + G$ et tout $x \in F \cap G$ s'écrit $x = x + 0 = 0 + x$ avec $(x, 0)$ et $(0, x)$ dans $F \times G$, ce qui impose $x = 0$ puisque φ est bijective.

Réciproquement supposons que $E = F+G$ et $F\cap G = \{0\}$. Si $x \in E$ s'écrit $x = y+z = y'+z'$, avec y, y' dans F et z, z' dans G, on a alors $y - y' = z' - z \in F \cap G$, donc $y - y' = z - z' = 0$ et $(y, z) = (y', z')$. La somme est donc directe. ∎

Définition 3.20 *On dit que deux sous-espaces vectoriels F et G de E sont supplémentaires, si $E = F \oplus G$. On dit aussi que F est un supplémentaire de G ou que G est un supplémentaire de F dans E.*

Remarque 3.6 *E est l'unique supplémentaire de $\{0\}$, mais un sous-espace vectoriel F de E distinct de $\{0\}$ et de E admet une infinité de supplémentaires. Il suffit de considérer deux droites de $\mathbb{R}^2$ dirigées par deux vecteurs non colinéaires pour s'en convaincre.*

On peut définir la somme ou la somme directe de p sous-espaces de E comme suit.

Définition 3.21 *Soient $p \geq 2$ un entier et $F_1, \cdots, F_p$ des sous-espaces vectoriels de E. On dit que E est somme des espaces $F_1, \cdots, F_p$ si l'application :*

$$\begin{array}{rccc} \varphi & F_1\times\cdots\times F_p & \to & E \\ & (x_1,\cdots,x_p) & \mapsto & x_1+\cdots+x_p \end{array}$$

est surjective et on note alors $E = F_1 + \cdots + F_p$ ou de manière plus compacte $E = \sum_{k=1}^{p} F_k$.
Si cette application est bijective, on dit que E est somme directe des espaces $F_1, \cdots, F_p$ et on note $E = F_1 \oplus \cdots \oplus F_p$ ou $E = \bigoplus_{k=1}^{p} F_k$.

En fait la somme de deux sous-espaces vectoriels F et G peut se définir par :

$$F+G=\{y+z \mid y\in F \text{ et } z\in G\}.$$

Il est facile de vérifier que $F+G$ est un sous-espace vectoriel de E. Ce sous-espace n'est en général pas égal à E.

On peut aussi définir $F+G$ comme le sous-espace vectoriel de E engendré par la réunion $F\cup G$ (qui en général n'est pas un espace vectoriel).

Théorème 3.24 *Si F,G sont deux sous-espaces vectoriels de E, alors la somme $F+G$ est le sous-espace vectoriel de E engendré par $F\cup G$.*

Démonstration. Dire que $x\in \text{Vect}\,(F\cup G)$ équivaut à dire qu'il s'écrit $x=\sum_{k=1}^{p}\lambda_k x_k$ où les x_k sont des éléments de $F\cup G$ et les λ_k des réels. En séparant les x_k qui sont dans F deux ceux qui sont dans G, cette somme peut s'écrire $x=y+z$ avec $y\in F$ et $z\in G$, ce qui signifie que $x\in F+G$.

Réciproquement $x\in F+G$ s'écrit $x=y+z$ avec $(y,z)\in F\times G$ et il est bien dans $\text{Vect}\,(F\cup G)$. ■

De manière un peu plus générale, on a le résultat suivant.

Théorème 3.25 *Si $F_1,\cdots,F_p$ sont des sous-espaces vectoriels de E, alors la somme $\sum_{k=1}^{p}F_k$ est le sous-espace vectoriel de E engendré par $\bigcup_{k=1}^{p}F_k$.*

Exercice 3.32 *Montrer que si φ est une forme linéaire non nulle sur E, il existe alors un vecteur non nul a dans E tel que :*

$$E=\ker(\varphi)\oplus\mathbb{R}a.$$

Solution 3.30 *La forme linéaire φ étant non nulle, on peut trouver un vecteur a dans E tel que $\varphi(a)\neq 0$. Ce vecteur a est nécessairement non nul. Pour tout vecteur x dans E, le vecteur $h=x-\dfrac{\varphi(x)}{\varphi(a)}a$ est dans le noyau de φ et en écrivant que $x=h+\dfrac{\varphi(x)}{\varphi(a)}a$ on déduit que $E=\ker(\varphi)+\mathbb{R}a$. Si x est dans $\ker(\varphi)\cap\mathbb{R}a$ on a alors $x=\lambda a$ et $\lambda\varphi(a)=\varphi(x)=0$ avec $\varphi(a)\neq 0$ ce qui entraîne $\lambda=0$ et $x=0$. On a donc $\ker(\varphi)\cap\mathbb{R}a=\{0\}$ et $E=\ker(\varphi)\oplus\mathbb{R}a$.*

4

Espaces vectoriels réels de dimension finie

4.1 Systèmes libres, systèmes générateurs et bases

Nous avons déjà rencontré et utilisé la base canonique de $\mathbb{R}^n$. Nous allons donner une définition précise de cette notion dans le cadre des espaces vectoriels réels, ce qui nous mènera à la notion de dimension.

On se donne un espace vectoriel réel E.

Définition 4.1 *Soit $\mathcal{B} = (e_i)_{1\leq i\leq n}$ une famille (ou un système) de n vecteurs de E, où n est un entier naturel non nul. On dit que $\mathcal{B}$ est :*

- *une famille libre, ou que les vecteurs $e_1, \cdots, e_n$ sont linéairement indépendants, si pour toute famille $(\lambda_i)_{1\leq i\leq n}$ l'égalité $\sum_{i=1}^{n} \lambda_i e_i = 0$ est réalisée si, et seulement si, tous les λ_i sont nuls ;*
- *une famille liée, si ce n'est pas une famille libre (i. e. il existe des réels $\lambda_1, \cdots, \lambda_n$ non tous nuls tels que $\sum_{i=1}^{n} \lambda_i e_i = 0$) ;*
- *une famille génératrice si pour tout vecteur $x \in E$, il existe des réels $\lambda_1, \cdots, \lambda_n$ tels que $x = \sum_{i=1}^{n} \lambda_i e_i$;*
- *une base de E si elle est libre et génératrice.*

Remarque 4.1 *Dire que $\mathcal{B} = (e_i)_{1\leq i\leq n}$ est une famille génératrice de E équivaut à dire que $E = \text{vect}\,(\mathcal{B})$ (l'espace vectoriel engendré par $\mathcal{B}$).*

Avec le théorème qui suit, on résume quelques propriétés des familles libres ou liées.

Théorème 4.1 *Soit $\mathcal{B} = (e_i)_{1\leq i\leq n}$ une famille de n vecteurs de E, où n est un entier naturel non nul.*

1. *Si $n = 1$, dire que $\mathcal{B}$ est libre [resp. liée] signifie que $e_1 \neq 0$ [resp. $e_1 = 0$].*
2. *Si $\mathcal{B}$ est libre, alors tous les vecteurs e_i sont non nuls.*
3. *Si l'un des e_i est nul, alors $\mathcal{B}$ est liée.*
4. *Si $\mathcal{B}$ contient une famille liée, elle est elle même liée.*
5. *Si $\mathcal{B}$ est contenue dans une partie libre, elle est elle même libre.*
6. *Si $\mathcal{B}$ est liée, l'un des vecteurs e_j est combinaison linéaire des autres.*
7. *Si $\mathcal{B}$ est une base de E, alors tout vecteur x de E s'écrit de manière unique comme combinaison linéaire des vecteurs $e_1, \cdots, e_n$.*

Démonstration. Résultent des définitions. ∎

L'utilisation des déterminants, définis pour l'instant dans le seul cas des matrices d'ordre 2, nous donne un moyen élémentaire de vérifier que deux vecteurs de $\mathbb{R}^2$ sont linéairement indépendants.

Théorème 4.2 *Les vecteurs* $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ *et* $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ *sont linéairement indépendants si, et seulement si :*

$$\begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} \neq 0.$$

Démonstration. Il revient au même de montrer que x et y sont liés si, et seulement si, $\begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} = 0$.

Supposons le système (x, y) lié. On a alors $y = \lambda x$ ou $x = \lambda y$ pour un réel λ et, par exemple dans le premier cas :

$$\begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} = \begin{vmatrix} x_1 & \lambda x_1 \\ x_2 & \lambda x_2 \end{vmatrix} = \lambda (x_1 x_2 - x_1 x_2) = 0.$$

Réciproquement, on suppose que $\begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} = 0$.

Si $x = 0$ ou $y = 0$, le système (x, y) est alors lié.

Si x et y sont non nuls, en supposant que y_2 est non nul, l'égalité $\begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} = x_1 y_2 - y_1 x_2 = 0$ entraîne (c'est même équivalent) :

$$y_2 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - x_2 \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

avec $(y_2, -x_2) \neq (0, 0)$, ce qui signifie que x et y sont liés. Si $y_2 = 0$, on a alors $y_1 \neq 0$ et on écrit que l'égalité $x_1 y_2 - x_2 y_1 = 0$ entraîne :

$$x_1 \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} - y_1 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

avec $(x_1, -y_1) \neq (0, 0)$. ∎

Si $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ est une base de E, alors tout vecteur $x \in E$ s'écrit $x = \sum_{i=1}^{n} \lambda_i e_i$ et les réels λ_i qui sont uniquement déterminés sont appelés les composantes ou coordonnées de x dans la base $\mathcal{B}$. Réciproquement une telle famille $\mathcal{B}$ vérifiant cette propriété est une base de E.

La base canonique de $\mathbb{R}^n$ est bien une base au sens de la définition qu'on a donné.

Exemple 4.1 *Dans l'espace* $\mathbb{R}_n[x]$ *des fonctions polynomiales de degré au plus égal à* n*, la famille de polynômes* $(1, x, \cdots, x^n)$ *est une base puisque tout polynôme dans* $\mathbb{R}_n[x]$ *s'écrit sous la forme* $P = \sum_{k=0}^{n} a_k x^k$*, les réels* a_k *étant uniquement déterminés. On dit que cette base est la base canonique de* $\mathbb{R}_n[x]$.

Le résultat de l'exercice qui suit est à retenir.

Exercice 4.1 *Soient n un entier naturel et $\mathcal{B} = (P_0, P_1, \cdots, P_n)$ une famille de polynômes dans $\mathbb{R}_n[x]$ telle que P_k soit de degré k, pour tout k compris entre 0 et n (P_0 est constant non nul). On dit qu'une telle famille de polynômes est échelonnée en degrés. Montrer que $\mathcal{B}$ est une base de $\mathbb{R}_n[x]$.*

Solution 4.1 *Notons, pour k compris entre 0 et n :*

$$P_k(x) = a_{k,0} + a_{k,1}x + \cdots + a_{k,k}x^k = \sum_{j=0}^{k} a_{k,j}x^j$$

où le coefficient $a_{k,k}$ est non nul.
Nous allons montrer le résultat par récurrence sur $n \geq 0$.
Pour $n = 0$, $\mathbb{R}_0[x]$ est l'espace des fonctions (ou polynômes) constantes sur $\mathbb{R}$ et P_0 est non nul dans cet espace, donc libre, et en écrivant tout polynôme constant sous la forme :

$$P(x) = a = \frac{a}{P_0}P_0 = \lambda P_0,$$

on voit que P_0 engendre $\mathbb{R}_0[x]$. Donc (P_0) est une base de $\mathbb{R}_0[x]$.
Supposons le résultat acquis au rang $n \geq 0$ et soit $\mathcal{B} = (P_0, P_1, \cdots, P_{n+1})$ une famille de polynômes échelonnée en degrés dans $\mathbb{R}_{n+1}[x]$. La famille $(P_0, P_1, \cdots, P_n)$ est alors échelonnée en degrés dans $\mathbb{R}_n[x]$ et l'hypothèse de récurrence nous dit qu'elle forme une base de $\mathbb{R}_n[x]$. On se donne un polynôme P dans $\mathbb{R}_{n+1}[x]$. Si P est de degré inférieur ou égal à n, il est dans $\mathbb{R}_n[x]$ et s'écrit comme combinaison linéaire de $P_0, P_1, \cdots, P_n$, sinon il est de la forme $P(x) = Q(x) + a_{n+1}x^{n+1}$ avec Q dans $\mathbb{R}_n[x]$ et $a_{n+1} \neq 0$. En écrivant que :

$$x^{n+1} = \frac{1}{a_{n+1,n+1}}P_{n+1}(x) - \sum_{j=0}^{n} \frac{a_{n+1,j}}{a_{n+1,a_{n+1}}}x^j,$$

on déduit que $P(x) = R(x) + \dfrac{a_{n+1}}{a_{n+1,n+1}}P_{n+1}(x)$ avec R dans $\mathbb{R}_n[x]$, donc combinaison linéaire de $P_0, P_1, \cdots, P_n$ et P est combinaison linéaire de $P_0, P_1, \cdots, P_{n+1}$. Le système $\mathcal{B}$ est donc générateur de $\mathbb{R}_{n+1}[x]$.
Si on a l'égalité $\sum_{j=0}^{n+1} \lambda_j P_j = 0$, alors $\lambda_{n+1}P_{n+1} = -\sum_{j=0}^{n} \lambda_j P_j$ est dans $\mathbb{R}_n[x]$ et λ_{n+1} est nécessairement nul puisque P_{n+1} qui est de degré $n+1$ n'est pas dans $\mathbb{R}_n[x]$. On a alors $\sum_{j=0}^{n} \lambda_j P_j$ et tous les λ_j sont nuls puisque $(P_0, P_1, \cdots, P_n)$ une base de $\mathbb{R}_n[x]$. Le système $\mathcal{B}$ est donc libre et c'est une base de $\mathbb{R}_{n+1}[x]$.

Exercice 4.2 *Montrer que la famille $\mathcal{B} = (L_0, L_1, L_2)$ de polynômes de $\mathbb{R}_2[x]$ définie par :*

$$\begin{cases} L_0(x) = (x-1)(x-2) \\ L_1(x) = x(x-2) \\ L_2(x) = x(x-1) \end{cases}$$

forme une base de $\mathbb{R}_2[x]$. En déduire que pour tout polynôme P dans $\mathbb{R}_2[x]$ on a :

$$\int_0^2 P(t)\,dt = \frac{P(0) + 4P(1) + P(2)}{3}$$

(formule des trois niveaux).

Solution 4.2 *Supposons que* $\lambda_0 P_0+\lambda_1 P_1+\lambda_2 P_2 = 0$. *Prenant les valeurs successives* $x = 0, 1, 2$, *on déduit que* $\lambda_0 = \lambda_1 = \lambda_2 = 0$. *Le système* $\mathcal{B}$ *est donc libre. Étant donné* $P = ax^2 + bx + c$, *on cherche des réels* $\lambda_0, \lambda_1, \lambda_2$ *tels que :*

$$ax^2 + bx + c = \lambda_0 P_0 + \lambda_1 P_1 + \lambda_2 P_2.$$

Là encore les valeurs $x = 0, 1, 2$ *nous donne :*

$$\begin{cases} 2\lambda_0 = c \\ -\lambda_1 = a + b + c \\ 2\lambda_2 = 4a + 2b + c \end{cases}$$

ce qui détermine de manière unique les réels $\lambda_0, \lambda_1, \lambda_2$. *Le système* $\mathcal{B}$ *est donc générateur et c'est une base de* $\mathbb{R}_2[x]$.
Tout polynôme P *dans* $\mathbb{R}_2[x]$ *s'écrit donc de manière unique* $P = \lambda_0 P_0 + \lambda_1 P_1 + \lambda_2 P_2$ *avec* $\lambda_0 = \dfrac{P(0)}{2}$, $\lambda_1 = -P(1)$ *et* $\lambda_2 = \dfrac{P(2)}{2}$. *On a alors :*

$$\begin{aligned} \int_0^2 P(t)\,dt &= \frac{P(0)}{2}\int_0^2 L_0(t)\,dt - P(1)\int_0^2 L_1(t)\,dt + \frac{P(2)}{2}\int_0^2 L_2(t)\,dt \\ &= \frac{P(0) + 4P(1) + P(2)}{3} \end{aligned}$$

On se limitera dans ce chapitre aux familles libres ou génératrices qui sont finies. Mais en réalité, on est rapidement amené à considérer des familles qui peuvent être infinies. On donne donc les définitions suivantes (qui ne seront pas utilisées au niveau élémentaire où se situe ce cours).

Définition 4.2 *Soit* $\mathcal{B} = (e_i)_{i\in I}$ *une famille de vecteurs de* E, *où* I *est un ensemble non vide quelconque (fini ou infini) d'indices. On dit que* $\mathcal{B}$ *est :*

- *une famille libre, ou que les vecteurs* e_i, *pour* $i \in I$ *sont linéairement indépendants, si toute sous-famille finie de* $\mathcal{B}$ *est libre, ce qui signifie que pour tout sous-ensemble non vide et fini* J *de* I, *une combinaison linéaire* $\sum_{j\in J} \lambda_j e_j$, *où les* λ_j *pour* $j \in J$ *sont des réels, est nulle si, et seulement si, tous ces* λ_j *sont nuls ;*
- *une famille liée, si ce n'est pas une famille libre (i. e. il existe une partie fini* J *de* I *et des réels* λ_j *où* j *décrit* J *qui sont non tous nuls tels que* $\sum_{j\in J} \lambda_j e_j = 0$*) ;*
- *une famille génératrice si l'espace vectoriel engendré par* $\mathcal{B}$ *est l'espace* E *tout entier, ce qui signifie que pour tout vecteur* $x \in E$, *il existe une partie finie* J *de* I *et des réels* λ_j *où* j *décrit* J *tels que* $x = \sum_{j\in J} \lambda_j e_j$ *;*
- *une base de* E *si elle est libre et génératrice.*

Exercice 4.3 *On désigne par* E *l'espace vectoriel des applications de* $\mathbb{R}$ *dans* $\mathbb{R}$ *et pour tout entier* $k \geq 1$ *par* f_k *la fonction définie sur* $\mathbb{R}$ *par :*

$$\forall x \in \mathbb{R},\ f_k(x) = \sin(kx).$$

Montrer que la famille (f_1, f_2, f_3) *est libre dans* E.

Solution 4.3 *Soient $\lambda_1, \lambda_2, \lambda_3$ des réels tels que :*

$$\forall x \in \mathbb{R},\ \lambda_1 \sin(x) + \lambda_2 \sin(2x) + \lambda_3 \sin(3x) = 0.$$

En dérivant deux fois, on a :

$$\forall x \in \mathbb{R},\ \lambda_1 \sin(x) + 4\lambda_2 \sin(2x) + 9\lambda_3 \sin(3x) = 0.$$

et retranchant ces deux équations, on a :

$$\forall x \in \mathbb{R},\ 3\lambda_2 \sin(2x) + 8\lambda_3 \sin(3x) = 0.$$

Prenant $x = \dfrac{\pi}{2}$ dans cette troisième équation on obtient $\lambda_3 = 0$ et la première donne $\lambda_1 = \lambda_3 = 0$. Il reste donc $\lambda_2 \sin(2x) = 0$ et $\lambda_2 = 0$. La famille (f_1, f_2, f_3) est donc libre dans E.

Un peu plus généralement, on a le résultat suivant.

Exercice 4.4 *On désigne par E l'espace vectoriel des applications de $\mathbb{R}$ dans $\mathbb{R}$. Montrer, pour tous réels $0 < a_1 < a_2 < \cdots < a_n$, la famille :*

$$\mathcal{L} = \{f_{a_k} : x \longmapsto \sin(a_k x) \mid 1 \leq k \leq n\}$$

est libre dans E (ce qui peut se traduire en disant que la famille de fonctions $\{f_a : x \longmapsto \sin(ax) \mid a \in \mathbb{R}^{+,*}\}$ est libre dans E).

Solution 4.4 *On procède par récurrence sur $n \geq 1$.*
Pour $n = 1$, la fonction $f_a : x \longmapsto \sin(ax)$ n'est pas la fonction nulle, donc (f_a) est libre dans E.
Supposons le résultat acquis au rang $n - 1 \geq 1$ et soient $0 < a_1 < a_2 < \cdots < a_n$, $\lambda_1, \lambda_2, \cdots, \lambda_n$ des réels tels que :

$$\forall x \in \mathbb{R},\ \sum_{k=1}^{n} \lambda_k \sin(a_k x) = 0.$$

En dérivant deux fois, on a :

$$\forall x \in \mathbb{R},\ \sum_{k=1}^{n} \lambda_k a_k^2 \sin(a_k x) = 0.$$

Il en résulte que :

$$\forall x \in \mathbb{R},\ \sum_{k=1}^{n-1} \lambda_k \left(a_k^2 - a_n^2\right) \sin(a_k x) = 0.$$

et l'hypothèse de récurrence nous dit que $\lambda_k (a_k^2 - a_n^2) = 0$ pour tout k compris entre 1 et $n-1$, ce qui équivaut à dire que $\lambda_k = 0$ pour tout k compris entre 1 et $n-1$ puisque $a_k^2 \neq a_n^2$ pour $k \neq 0$. Il reste alors $\lambda_n f_{a_n} = 0$ dans E et $\lambda_n = 0$. On a donc ainsi montré que la famille $(f_{a_k})_{1 \leq k \leq n}$ est libre dans E.

4.2 Espaces vectoriels de dimension finie

Le théorème qui suit est essentiel pour définir la notion de dimension finie.

Théorème 4.3 *Si un espace vectoriel E admet une famille génératrice $\mathcal{B}$ formée de $n \geq 1$ éléments, alors toute famille libre dans E a au plus n éléments (ce qui équivaut à dire qu'un système de plus de $n+1$ vecteurs est lié).*

Démonstration. On procède par récurrence sur n.

Soit $\mathcal{B} = (e_i)_{1\leq i\leq n}$ une famille génératrice de E.

Si $n = 1$, alors pour tout couple (x, y) de vecteurs non nuls de E on peut trouver deux réels non nuls λ et μ tels que $x = \lambda e_1$ et $y = \mu e_1$ et on a la combinaison linéaire nulle $\mu x - \lambda y = 0$ avec μ et $-\lambda$ non nuls, ce qui signifie que le système (x, y) est lié. Il ne peut donc exister de famille libre à 2 éléments dans E et a fortiori il ne peut en exister à plus de 2 éléments.

Supposons le résultat acquis au rang $n - 1 \geq 1$, c'est-à-dire que dans tout espace vectoriel F admettant un système générateur de $n - 1$ vecteurs une famille de plus de n vecteurs est liée. Supposons que E soit un espace vectoriel admettant une famille génératrice à n éléments. Supposons qu'il existe une famille libre ayant $m \geq n + 1$ éléments. On peut extraire de cette famille une famille libre à $n+1$ éléments puisque toute sous-famille d'une famille libre est libre. Soit $\mathcal{L} = (f_i)_{1\leq i\leq n+1}$ une telle famille. Comme $\mathcal{B} = (e_i)_{1\leq i\leq n}$ est génératrice, il existe des réels a_{ij} tels que :

$$\begin{cases} f_1 = a_{11}e_1 + \cdots + a_{1n}e_n \\ \vdots \\ f_{n+1} = a_{n+1,1}e_1 + \cdots + a_{n+1,n}e_n \end{cases}$$

Si tous les $a_{i,n}$ sont nuls alors les f_i sont dans l'espace vectoriel F engendré par les $n-1$ vecteurs $e_1, \cdots, e_{n-1}$ et en conséquence liés (hypothèse de récurrence), en contradiction avec $\mathcal{L}$ libre. Il existe donc un indice i compris entre 1 et $n + 1$ tel que $a_{i,n} \neq 0$ et en changeant au besoin la numérotation des éléments de $\mathcal{L}$ on peut supposer que $i = n + 1$. Les n vecteurs :

$$\begin{cases} g_1 = a_{n+1,n}f_1 - a_{1n}f_{n+1} \\ \vdots \\ g_n = a_{n+1,n}f_1 - a_{nn}f_{n+1} \end{cases}$$

sont dans l'espace vectoriel F engendré par les $n - 1$ vecteurs $e_1, \cdots, e_{n-1}$ (on a annulé les composantes en e_{n+1}) et en conséquence liés (hypothèse de récurrence), c'est-à-dire qu'il existe des réels $\lambda_1, \cdots, \lambda_n$ non tous nuls tels que :

$$\lambda_1 g_1 + \cdots + \lambda_n g_n = 0$$

ce qui entraîne :

$$a_{n+1,n}\left(\lambda_1 f_1 + \cdots + \lambda_n f_n\right) - \left(\lambda_1 a_{1n} + \cdots + \lambda_n a_{nn}\right) f_{n+1} = 0$$

les réels $a_{n+1,n}\lambda_1, \cdots, a_{n+1,n}\lambda_n$ n'étant pas tous nuls. Ce qui nous dit encore que les f_i sont liés et est en contradiction avec $\mathcal{L}$ libre. Il est donc impossible de trouver un tel système $\mathcal{L}$ libre. ■

Définition 4.3 *On dit qu'un espace vectoriel est de dimension finie s'il est réduit à $\{0\}$ ou s'il est différent de $\{0\}$ et admet une base formée d'un nombre fini de vecteurs. Dans le cas contraire, on dit qu'il est de dimension infinie.*

On déduit alors du théorème précédent le suivant.

Théorème 4.4 *Si E un espace vectoriel non réduit à $\{0\}$ et de dimension finie, alors toutes les bases ont le même nombre d'éléments.*

Démonstration. Si $\mathcal{B} = (e_i)_{1\leq i\leq n}$ et $\mathcal{B}' = (e'_i)_{1\leq i\leq n'}$ sont deux bases de l'espace vectoriel E, ce sont alors deux familles génératrices et libres et le théorème précédent nous dit que $n' \leq n$ et $n \leq n'$, soit $n = n'$. ■

On peut alors donner la définition suivante.

Définition 4.4 *Si E est un espace vectoriel non réduit à $\{0\}$ et de dimension finie, alors sa dimension est le nombre de l'une quelconque de ses bases. On note* $\dim(E)$ *cette dimension.*

Par convention, on dira que l'espace vectoriel $\{0\}$ est de dimension 0.

Un espace vectoriel E est donc de dimension 0 si, et seulement si, il est réduit à $\{0\}$.

Dans le cas général on peut montrer, mais cela dépasse le niveau de ce cours d'introduction, que tout espace vectoriel admet une base (finie ou infinie).

On appelle droite tout espace vectoriel de dimension 1 et plan tout espace vectoriel de dimension 2.

Exemple 4.2 *Bien entendu l'espace $\mathbb{R}^n$ est de dimension n.*

Exemple 4.3 *L'ensemble $\mathbb{C}$ des nombres complexes est un espace vectoriel réel de dimension 2.*

Exemple 4.4 *Pour tout entier naturel n, l'espace $\mathbb{R}_n[x]$ des fonctions polynomiales de degré au plus égal à n est de dimension $n+1$.*

Exemple 4.5 *Pour tous entiers naturels non nuls n et m, l'espace $\mathcal{M}_{m,n}(\mathbb{R})$ des matrices réelles à m lignes et n colonnes est un espace vectoriel de dimension $m \cdot n$. En particulier l'espace $\mathcal{M}_n(\mathbb{R})$ des matrices carrées d'ordre n est de dimension n^2.*
La base canonique de $\mathcal{M}_{m,n}(\mathbb{R})$ est $(E_{i,j})_{\substack{1\leq i\leq m\\1\leq j\leq n}}$ où $E_{i,j}$ est la matrice dont tous les coefficients sont nuls sauf celui en position (i,j).

Exemple 4.6 *L'espace des fonctions définies sur un intervalle réel I et à valeurs réelles est de dimension infinie (l'exercice 4.4 nous montre qu'on peut trouver des familles libres ayant une infinité d'éléments). Il admet des bases, mais il n'est pas possible d'en expliciter une.*

Exercice 4.5 *Montrer que la dimension de l'espace des matrices carrées A d'ordre n qui sont symétriques (i. e. telles que ${}^tA = A$) est égale à $\dfrac{n(n+1)}{2}$ et que la dimension de l'espace des matrices carrées A d'ordre n qui sont anti-symétriques (i. e. telles que ${}^tA = -A$) est égale à $\dfrac{n(n-1)}{2}$.*

Solution 4.5 *Laissée au lecteur.*

Remarque 4.2 *Si $\mathcal{B} = (e_i)_{1\leq i\leq n}$ est une base E, on a alors $E = \text{vect}(\mathcal{B})$.*

Le théorème qui suit explique l'importance de l'espace vectoriel $\mathbb{R}^n$ étudié en début de chapitre.

Théorème 4.5 *Un espace vectoriel réel de dimension n est isomorphe à $\mathbb{R}^n$.*

Démonstration. C'est le choix d'une base $\mathcal{B} = (e_i)_{1\leq i\leq n}$ de E qui nous permet de définir un isomorphisme de E sur $\mathbb{R}^n$. En effet l'unicité de l'écriture de tout vecteur x de E sous la forme $x = \sum_{k=1}^{n} \lambda_k e_k$ se traduit en disant que l'application :

$$\begin{array}{ccc} E & \to & \mathbb{R}^n \\ x = \sum_{k=1}^{n} \lambda_k e_k & \mapsto & (\lambda_1, \lambda_2, \cdots, \lambda_n) \end{array}$$

est bijective et il est facile de vérifier que cette application est linéaire. ■

Théorème 4.6 *Soit E un espace vectoriel de dimension $n \geq 1$.*

1. *Une famille libre dans E a au plus n élément et c'est une base si, et seulement si, elle a exactement n éléments.*
2. *Une famille génératrice dans E a au moins n élément et c'est une base si, et seulement si, elle a exactement n éléments.*

Démonstration. Le cas $n = 1$ est laissé au lecteur et on suppose que $n \geq 2$.

1. Le théorème 4.3 nous dit qu'une famille libre dans E a au plus n élément et si c'est une base, elle a obligatoirement n éléments. Il reste à montrer qu'une famille libre de n éléments est une base. Pour ce faire il suffit de montrer qu'elle est génératrice. Notons $\mathcal{B} = (e_i)_{1\leq i\leq n}$ une telle famille libre. Pour tout vecteur $x \in E$ la famille $\mathcal{B} \cup \{x\}$ est liée puisque formée de $n+1$ éléments, il existe donc des réels $\lambda, \lambda_1, \cdots, \lambda_n$ non tous nuls tels que $\lambda x + \sum_{i=1}^{n} \lambda_i e_i = 0$. Si $\lambda = 0$, on a alors $\sum_{i=1}^{n} \lambda_i e_i = 0$ et tous les λ_i sont nuls puisque $\mathcal{B}$ est libre, ce qui n'est pas possible. On a donc $\lambda \neq 0$ et $x = -\sum_{i=1}^{n} \frac{\lambda_i}{\lambda} e_i$. On a donc ainsi montré que $\mathcal{B}$ est génératrice et que c'est une base.
2. Le théorème 4.3 nous dit qu'une famille génératrice dans E a au moins n élément et si c'est une base, elle a obligatoirement n éléments. Il reste à montrer qu'une famille génératrice de n éléments est une base. Pour ce faire il suffit de montrer qu'elle est libre. Notons $\mathcal{B} = (e_i)_{1\leq i\leq n}$ une telle famille génératrice. Si cette famille est liée, l'un des e_i, disons e_n, est combinaison linéaire des autres et la famille $(e_i)_{1\leq i\leq n-1}$ est génératrice, ce qui est en contradiction avec le théorème 4.3. La famille $\mathcal{B}$ est donc génératrice et c'est une base de E.

■

On retient de ces résultats que pour montrer qu'une famille finie $\mathcal{B}$ de vecteurs est une base de E, on peut procéder comme suit :

- si on ne connaît pas la dimension de E, on montre que $\mathcal{B}$ est génératrice et libre ;
- si on sait que E est de dimension n, on vérifie que $\mathcal{B}$ a exactement n éléments et on montre que $\mathcal{B}$ est libre ou qu'elle est génératrice (il est inutile de montrer les deux points).

On a défini un espace vectoriel de dimension finie comme un espace vectoriel admettant une base finie. Le théorème qui suit nous dit qu'on peut aussi définir un espace vectoriel de dimension finie comme un espace vectoriel admettant une famille génératrice finie.

Théorème 4.7 *Soit E un espace vectoriel admettant une famille génératrice finie. De cette famille on peut extraire une base et E est de dimension finie.*

Démonstration. Soit $\mathcal{G} = (u_i)_{1\leq i\leq p}$ une famille génératrice de E. Si cette famille est libre, elle constitue alors une base de E et E est de dimension p. Sinon, cette famille est liée et l'un de ses éléments, disons u_p est combinaison linéaire des autres (en changeant la numérotation des éléments de G on peut toujours se ramener à ce cas de figure), ce qui implique que la famille $\mathcal{G}' = (u_i)_{1\leq i\leq p-1}$ est encore génératrice. Si cette famille est libre, c'est alors une base et E est de dimension finie, sinon on recommence. En un nombre fini de telles opérations on construit ainsi une base de E formée de $n \leq p$ éléments. ■

Théorème 4.8 (base incomplète) *Soit E un espace vectoriel de dimension $n \geq 1$. Toute famille libre à p éléments dans E (nécessairement $1 \leq p \leq n$) peut se compléter en une base.*

Démonstration. Soient $\mathcal{B} = (e_i)_{1\leq i\leq n}$ une base de E et $\mathcal{L} = (u_i)_{1\leq i\leq p}$ une famille libre dans E. On sait déjà que $p \leq n$. Si $p = n$, $\mathcal{L}$ est alors une base.

Supposons que $p < n$. Il existe alors un vecteur e_k dans $\mathcal{B}$ tel que $\mathcal{L}' = \mathcal{L} \cup \{e_k\}$ soit libre. En effet si un tel système est lié pour tout entier k compris entre 1 et n, il existe alors, pour chaque entier k, des réels $\lambda_1, \cdots, \lambda_p, \lambda_{p+1}$ non tous nuls tels que $\sum_{i=1}^{p} \lambda_i u_i + \lambda_{p+1} e_k = 0$. Si $\lambda_{p+1} = 0$, on a alors $\sum_{i=1}^{p} \lambda_i u_i = 0$ et tous les λ_i sont nuls puisque $\mathcal{L}$ est libre. On a donc $\lambda_{p+1} \neq 0$ et $e_k = -\sum_{i=1}^{p} \frac{\lambda_i}{\lambda_{p+1}} u_i$. En conséquence, tous les vecteurs de la base $\mathcal{B}$ sont combinaisons linéaires des éléments de $\mathcal{L}$ et $\mathcal{L}$ est alors générateur de E, ce qui est impossible pour $p < n$. Le système $\mathcal{L}'$ est donc libre. Si $p + 1 = n$, c'est une base et sinon on recommence. On arrive ainsi à compléter $\mathcal{L}$ en une base de E au bout d'un nombre fini d'opérations. ■

Les deux corollaires qui suivent sont des résultats importants à retenir.

Corollaire 4.1 *Soit E un espace vectoriel de dimension n. Tout sous-espace vectoriel F de E est de dimension finie $m \leq n$ et $m = n$ si, et seulement si, $F = E$.*

Démonstration. Si $F = \{0\}$, il est alors de dimension $0 \leq n$ et $n = 0$ équivaut à $F = E$.

On suppose que $F \neq \{0\}$ (donc $n \geq 1$).

Montrons tout d'abord que F est de dimension finie. Comme $n + 1$ vecteurs de F sont nécessairement liés, on peut définir l'entier m comme le plus grand entier pour lequel on peut trouver m vecteurs de F linéairement indépendants. On a $m \geq 1$ puisque $F \neq \{0\}$ et $m \leq n$ d'après le théorème 4.3. Soient donc $(f_i)_{1\leq i\leq m}$ une famille libre dans F. Pour tout vecteur $x \in F$, la famille $(f_1, \cdots, f_m, x)$ est liée et x est combinaison linéaire des f_i puisque $(f_i)_{1\leq i\leq m}$ est libre. La famille $(f_i)_{1\leq i\leq m}$ est donc une base de F et cet espace est de dimension finie $m \leq n$.

Si $m = n$, une base de F est aussi une base de E et $F = E$. La réciproque est évidente. ■

Corollaire 4.2 *Tout sous-espace-vectoriel F d'un espace vectoriel E de dimension finie admet des supplémentaires et pour tout supplémentaire G de F dans E, on :*

$$\dim(E) = \dim(F) + \dim(G). \tag{4.1}$$

Démonstration. On suppose que F est un sous-espace vectoriel strict de E, c'est-à-dire que $F \neq \{0\}$ et $F \neq E$. On a alors $1 \leq p = \dim(F) \leq n - 1$.

Une base $\mathcal{L} = (u_i)_{1\leq i\leq p}$ de F se complète en une base $\mathcal{B} = (u_i)_{1\leq i\leq n}$ et on vérifie facilement que le sous-espace vectoriel G de E engendré par $\mathcal{L}' = (u_i)_{p+1\leq i\leq n}$ est un supplémentaire de F. Pour cet espace G, on a $\dim(G) = n - p = \dim(E) + \dim(F)$.

Réciproquement si $E = F \oplus G$, on vérifie facilement que la réunion d'une base de F et d'une base de G nous fourni une base de E, ce qui implique que $\dim(E) = \dim(F) + \dim(G)$. ■

De manière plus générale, on peut montrer que si E est un espace vectoriel de dimension finie ou non, alors tout sous-espace vectoriel de E admet une supplémentaire dans E.

L'égalité (4.1) pour $E = F \oplus G$ peut se généraliser.

Théorème 4.9 *Soit E un espace vectoriel de dimension n. Si E est somme directe de $p \geq 2$ sous espaces stricts $F_1, \cdots, F_p$, soit $E = \bigoplus_{k=1}^{p} F_k$, alors en désignant, pour tout k compris entre 1 et p, par $\mathcal{B}_k$ une base de F_k, la famille $\mathcal{B} = \bigcup_{k=1}^{p} \mathcal{B}_k$ est une base de E et :*

$$\dim(E) = \sum_{k=1}^{p} \dim(F_k).$$

Démonstration. On vient de voir que le résultat est vrai pour $p = 2$ et une récurrence nous montre qu'il est vrai pour tout $p \geq 2$. ■

Dans le cas de la somme, non nécessairement directe, de deux sous-espaces d'un espace de dimension finie, on a le résultat suivant.

Théorème 4.10 *Soient E un espace vectoriel de dimension finie et F, G deux sous espaces vectoriels de E. On a :*

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

Démonstration. Comme $F \cap G$ est un sous-espace de G, il admet un supplémentaire H dans G :

$$G = (F \cap G) \oplus H$$

Ce sous-espace H de G est aussi un sous-espace de $F + G$.

En fait H est un supplémentaire de F dans $F + G$.

En effet, on a $F + H \subset F + G$ puisque $H \subset G$ et tout $x \in F + G$ s'écrit $x = y + z$ avec $y \in F$ et $z \in G = (F \cap G) \oplus H$, donc $z = z_1 + z_2$ avec $z_1 \in F \cap G \subset F$ et $z_2 \in H$, ce qui donne $x = (y + z_1) + z_2 \in F + H$. On a donc $F + G \subset F + H$ et l'égalité $F + G = F + H$.

Si maintenant x est dans $F \cap H$, il est dans $F \cap G$ puisque $H \subset G$, donc dans $(F \cap G) \cap H = \{0\}$. On a donc $F \cap H = \{0\}$ et $F + G = F \oplus H$.

Il en résulte que :

$$\begin{aligned} \dim(F+G) &= \dim(F) + \dim(H) \\ &= \dim(F) + \dim(G) - \dim(F \cap G). \end{aligned}$$

■

Ce résultat a pour conséquence le résultat suivant très utile pour montrer qu'un espace est somme directe de deux sous-espaces.

Théorème 4.11 *Soient E un espace vectoriel de dimension finie et F, G deux sous espaces vectoriels de E. On a :*

$$\begin{aligned} (E = F \oplus G) &\Leftrightarrow E = F + G \text{ et } \dim(E) = \dim(F) + \dim(G) \\ &\Leftrightarrow F \cap G = \{0\} \text{ et } \dim(E) = \dim(F) + \dim(G) \end{aligned}$$

Démonstration. On sait déjà que si $E = F \oplus G$, alors $E = F + G$, $F \cap G = \{0\}$ et $\dim(E) = \dim(F) + \dim(G)$.

Supposons que $E = F + G$ et $\dim(E) = \dim(F) + \dim(G)$. Le théorème précédent nous dit alors que $\dim(F \cap G) = 0$, soit que $F \cap G = \{0\}$ et on a $E = F \oplus G$.

De même si $F \cap G = \{0\}$ et $\dim(E) = \dim(F) + \dim(G)$, on a alors $F + G = F \oplus G$ et cet espace a même dimension que E, donc $E = F \oplus G$. ■

Les propriétés des applications linéaires injectives, surjectives ou bijectives décrites par le théorème qui suit sont importantes.

Théorème 4.12 *Soient E, F deux espaces vectoriels de dimension finie et u une application linéaire de E dans F.*

1. *Si u est injective, elle transforme alors tout système libre de E en un système libre de F et* $\dim(E) \leq \dim(F)$.
2. *Si u est surjective, elle transforme alors tout système générateur de E en un système générateur de F et* $\dim(F) \leq \dim(E)$.
3. *Si u est bijective, elle transforme alors toute base de E en une base de F et* $\dim(E) = \dim(F)$ *(deux espaces vectoriels de dimension finie isomorphes ont la même dimension).*
4. *Si* $\dim(E) = \dim(F)$, *alors :*

$$u \text{ bijective} \Leftrightarrow u \text{ injective} \Leftrightarrow u \text{ surjective.}$$

Démonstration.

1. Soit $\mathcal{L} = (x_i)_{1 \leq i \leq p}$ un système libre dans E. Si $\sum_{i=1}^{p} \lambda_i u(x_i) = 0$, on a alors du fait de la linéarité de u, $u\left(\sum_{i=1}^{p} \lambda_i x_i\right) = 0$, ce qui signifie que $\sum_{i=1}^{p} \lambda_i x_i$ est dans le noyau de u, donc nul puisque u est injective, ce qui équivaut à la nullité de tous les coefficients λ_i puisque $\mathcal{L}$ est libre. La famille $u(\mathcal{L})$ est donc libre.
   Prenant pour $\mathcal{L}$ une base de E, elle est formée de $n = \dim(E)$ éléments et $u(\mathcal{L})$ est libre à n éléments dans F, donc $n \leq m = \dim(F)$.
2. Soit $\mathcal{L} = (x_i)_{1 \leq i \leq p}$ un système générateur de E. Comme u est surjective tout vecteur y de F s'écrit $y = u(x)$ avec x dans E qui s'écrit $x = \sum_{i=1}^{p} \lambda_i x_i$, ce qui donne $y = \sum_{i=1}^{p} \lambda_i u(x_i)$.
   Le système $u(\mathcal{L})$ est donc générateur de F.
   Prenant pour $\mathcal{L}$ une base de E, elle est formée de n éléments et $u(\mathcal{L})$ est générateur de F à n éléments, donc $n \geq m$.
   3. et 4.Résultent des deux points précédents.

■

Le théorème 4.5 et le point 3. du théorème précédent nous disent que deux espaces vectoriels de dimension finie ont même dimension si, et seulement si, ils sont isomorphes.

En vue de généraliser le théorème 4.11, on utilisera le résultat suivant.

Lemme 4.1 *Si $F_1, \cdots, F_p$ sont des espaces vectoriels de dimension finie, il en est de même de l'espace produit $F = F_1 \times \cdots \times F_p$ et on a :*

$$\dim(F) = \sum_{k=1}^{p} \dim(F_k).$$

Démonstration. Pour $p = 1$, il n'y a rien à montrer.

En procédant par récurrence sur $p \geq 2$, il suffit de montrer le résultat pour $p = 2$.

Pour ce faire, on vérifie que si $\mathcal{B}_1 = (e_i)_{1\leq i\leq n}$ est une base de F_1 et $\mathcal{B}_2 = (f_i)_{1\leq i\leq m}$ une base de F_2, alors la famille :

$$\mathcal{B} = \{(e_i, 0) \mid 1 \leq i \leq n\} \cup \{(0, f_j) \mid 1 \leq j \leq m\}$$

est une base de $F = F_1 \times F_2$. En effet tout vecteur $z = (x, y)$ dans F s'écrit :

$$z = \left(\sum_{i=1}^{n} \lambda_i e_i, \sum_{j=1}^{m} \mu_j f_j\right) = \sum_{i=1}^{n} \lambda_i (e_i, 0) + \sum_{j=1}^{m} \mu_j (0, f_j)$$

donc $\mathcal{B}$ engendre F et l'égalité :

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

est équivalente à :

$$\left(\sum_{i=1}^{n} \lambda_i e_i, \sum_{j=1}^{m} \mu_j f_j\right) = (0, 0)$$

soit à $\sum_{i=1}^{n} \lambda_i e_i = 0$ et $\sum_{j=1}^{m} \mu_j f_j = 0$ qui impose $\lambda_i = 0$ pour tout i compris entre 1 et n et $\mu_j = 0$ pour tout j compris entre 1 et m. La famille $\mathcal{B}$ est donc libre et c'est une base de F. L'espace F est donc de dimension finie égale au nombre d'éléments de $\mathcal{B}$, soit à $n + m$. ■

Théorème 4.13 *Si $F_1, \cdots, F_p$ sont des sous-espaces vectoriels d'un espace vectoriel E de dimension finie, on a alors $E = \bigoplus_{k=1}^{n} F_k$ si, et seulement si, $E = \sum_{k=1}^{p} F_k$ et $\dim(E) = \sum_{k=1}^{p} \dim(F_k)$.*

Démonstration. On sait déjà que la condition est nécessaire.

Dire que $E = \sum_{k=1}^{p} F_k$, équivaut à dire que l'application linéaire :

$$\begin{array}{rccc} \varphi & F = F_1 \times \cdots \times F_p & \to & E \\ & (x_1, \cdots, x_p) & \mapsto & x_1 + \cdots + x_p \end{array}$$

est surjective et si de plus $\dim(F) = \dim(E)$, cette application est en fait une bijection ce qui signifie que $E = \bigoplus_{k=1}^{n} F_k$. ■

Exercice 4.6 *Montrer que l'ensemble des matrices carrées d'ordre n de trace nulle est un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{R})$ et calculer sa dimension.*

4.3 Rang d'un système de vecteurs ou d'une application linéaire

On désigne par E un espace vectoriel réel de dimension finie.

On rappelle que le sous-espace vectoriel $F = \text{Vect}\{x_1, \cdots, x_p\}$ de E engendré par des vecteurs $x_1, \cdots, x_p$ de E est l'ensemble de toutes les combinaisons linéaires de ces vecteurs, soit :

$$F = \left\{ x = \sum_{k=1}^{p} \lambda_k x_k \mid (\lambda_1, \lambda_2, \cdots, \lambda_p) \in \mathbb{R}^p \right\}.$$

Définition 4.5 *Le rang de la famille* $\{x_1, \cdots, x_p\}$ *de vecteurs de* E *est la dimension de l'espace vectoriel engendré par ces vecteurs. On le note* $\operatorname{rg}(x_1, \cdots, x_p)$.

Théorème 4.14 *Le rang d'une famille* $\{x_1, \cdots, x_p\}$ *de* p *vecteurs de* E *est au maximum égal à* p *et ce rang vaut* p *si, et seulement si, cette famille est libre.*

Démonstration. Si le système $\mathcal{L} = \{x_1, \cdots, x_p\}$ est libre, il constitue une base de $F = \operatorname{Vect}(\mathcal{L})$ et $\operatorname{rg}(\mathcal{L}) = \dim(F) = p$. Réciproquement si le rang vaut p, la famille $\mathcal{L}$ est génératrice de F avec p éléments, c'est donc une base de F et en conséquence une famille libre. ■

Remarque 4.3 *Si* E *est de dimension* n*, le rang d'une famille de vecteurs de* E *est au plus égal à* n*. Si ce rang vaut* n*, on a alors* $\operatorname{Vect}\{x_1, \cdots, x_p\} = E$ *et* $\{x_1, \cdots, x_p\}$ *est un système générateur de* E*. Dans le cas où* $p = n$*, c'est une base.*

Définition 4.6 *Soient* E, F *deux espaces vectoriels de dimension finie et* u *une application linéaire de* E *dans* F*. Le rang de* u *est la dimension de* $\operatorname{Im}(u)$*. On le note* $\operatorname{rg}(u)$.

En désignant par $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ une base de E, l'image de u est le sous-espace vectoriel de F engendré par $\{u(e_1), \cdots, u(e_n)\}$ et :

$$\operatorname{rg}(u) = \operatorname{rg}(u(e_1), \cdots, u(e_n)).$$

Remarque 4.4 *Comme* $\operatorname{Im}(u)$ *est un sous-espace vectoriel de* F*, on a* $\operatorname{rg}(u) \leq \dim(F)$ *et comme* $\operatorname{rg}(u) = \operatorname{rg}(u(e_1), \cdots, u(e_n))$*, on a aussi* $\operatorname{rg}(u) \leq \dim(E)$*. Donc :*

$$\operatorname{rg}(u) \leq \min(\dim(E), \dim(F))$$

Plus précisément, on a les résultats suivants.

Théorème 4.15 *Soient* E, F *deux espaces vectoriels de dimension finie et* u *une application linéaire de* E *dans* F*. On a* $\operatorname{rg}(u) = \dim(F)$ *si, et seulement si,* u *est surjective.*

Démonstration. Dire que u est surjective équivaut à dire que $\operatorname{Im}(u) = F$, ce qui est encore équivalent à $\operatorname{rg}(u) = \dim(\operatorname{Im}(u)) = \dim(F)$ puisque $\operatorname{Im}(u)$ est un sous-espace vectoriel de F. ■

Théorème 4.16 (du rang) *Soient* E, F *deux espaces vectoriels de dimension finie et* u *une application linéaire de* E *dans* F*. On a :*

$$\dim(E) = \dim(\ker(u)) + \operatorname{rg}(u).$$

Démonstration. Soit H un supplémentaire de $\ker(u)$ dans E et v la restriction de u à H, c'est-à-dire l'application v définie sur H par :

$$\forall x \in H,\ v(x) = u(x).$$

Le noyau de cette application est :

$$\ker(v) = H \cap \ker(u) = \{0\}$$

ce qui signifie que v est injective de H dans F et réalise une bijection de H dans $\operatorname{Im}(v)$.

En écrivant tout vecteur y de $\operatorname{Im}(u)$ sous la forme $y = u(x)$ avec $x \in E$ qui s'écrit $x = x_1 + x_2$ où $x_1 \in \ker(u)$ et $x_2 \in H$, on a $y = u(x_1) + u(x_2) = v(x_2)$, c'est-à-dire que y est dans $\operatorname{Im}(v)$. On a donc $\operatorname{Im}(u) \subset \operatorname{Im}(v)$ et comme $\operatorname{Im}(v) \subset \operatorname{Im}(u)$, on a en fait $\operatorname{Im}(v) = \operatorname{Im}(u)$ et v réalise un isomorphisme de H sur $\operatorname{Im}(u)$. Il en résulte que :

$$\operatorname{rg}(u) = \dim(\operatorname{Im}(u)) = \dim(H) = \dim(E) - \dim(\ker(u)).$$

■

Remarque 4.5 *En montrant le théorème du rang, on a en fait montré que* $\mathrm{Im}\,(u)$ *est isomorphe à un supplémentaire de* $\ker(u)$ *dans* E.

Corollaire 4.3 *Soient* E, F *deux espaces vectoriels de dimension finie et* u *une application linéaire de* E *dans* F. *On a :*

1. $\mathrm{rg}\,(u) \leq \min(\dim(E), \dim(F))$;
2. $\mathrm{rg}\,(u) = \dim(E)$ *si, et seulement si,* u *est injective.*

Démonstration.

1. De la formule du rang, on déduit que $\mathrm{rg}\,(u) \leq \dim(E)$ et $\mathrm{rg}\,(u) \leq \dim(F)$ par définition.
2. Si $\mathrm{rg}\,(u) = \dim(E)$, la formule du rang nous dit que $\dim(\ker(u)) = 0$, soit que $\ker(u) = \{0\}$ et u est injective. Réciproquement si u est injective, on a $\ker(u) = \{0\}$ et $\mathrm{rg}\,(u) = \dim(E)$.

■

Exercice 4.7 *Soit* E *un espace vectoriel réel de dimension finie et* $u \in \mathcal{L}(E)$.

1. *Montrer que :*
$$\mathrm{Im}\,(u) = \mathrm{Im}\left(u^2\right) \Leftrightarrow \ker(u) = \ker\left(u^2\right)$$
où $u^2 = u \circ u$.
2. *Montrer que :*
$$\mathrm{Im}\,(u) = \mathrm{Im}\left(u^2\right) \Leftrightarrow E = \ker(u) \oplus \mathrm{Im}\,(u) \Leftrightarrow \ker(u) = \ker\left(u^2\right)$$

Solution 4.6

1. *On a toujours :*
$$\mathrm{Im}\left(u^2\right) \subset \mathrm{Im}\,(u),\ \ker(u) \subset \ker\left(u^2\right)$$
donc :
$$\mathrm{Im}\,(u) = \mathrm{Im}\left(u^2\right) \Leftrightarrow \mathrm{rg}\,(u) = \mathrm{rg}\left(u^2\right)$$
et :
$$\ker(u) = \ker\left(u^2\right) \Leftrightarrow \dim(\ker(u)) = \dim\left(\ker\left(u^2\right)\right)$$
D'autre part, le théorème du rang nous dit que :
$$\dim(E) = \dim(\ker(u)) + \mathrm{rg}\,(u) = \dim\left(\ker\left(u^2\right)\right) + \mathrm{rg}\left(u^2\right)$$
ce qui permet de déduire que :
$$\mathrm{Im}\,(u) = \mathrm{Im}\left(u^2\right) \Leftrightarrow \ker(u) = \ker\left(u^2\right)$$
2. *Il suffit de montrer que :*
$$\mathrm{Im}\,(u) = \mathrm{Im}\left(u^2\right) \Leftrightarrow E = \ker(u) \oplus \mathrm{Im}\,(u)$$
Si $\mathrm{Im}\,(u) = \mathrm{Im}\,(u^2)$, *alors pour tout* x *dans* E, *il existe* y *dans* E *tel que* $u(x) = u^2(y)$, *donc* $x - u(y) \in \ker(u)$ *et* $x = (x - u(y)) + u(y) \in \ker(u) + \mathrm{Im}\,(u)$. *On a donc* $E = \ker(u) + \mathrm{Im}\,(u)$ *et avec le théorème du rang, on déduit que* $E = \ker(u) \oplus \mathrm{Im}\,(u)$.
Si $E = \ker(u) \oplus \mathrm{Im}\,(u)$ *alors tout* $x \in \ker\left(u^2\right)$ *s'écrit* $x = x_1 + u(x_2)$ *avec* $u(x_1) = 0$ *et* $0 = u^2(x) = u^3(x_2)$ *entraîne que* $u^2(x_2) \in \ker(u) \cap \mathrm{Im}\,(u) = \{0\}$, *donc* $u(x_2) \in \ker(u) \cap \mathrm{Im}\,(u) = \{0\}$ *et* $x = x_1 \in \ker(u)$. *On a donc* $\ker\left(u^2\right) \subset \ker(u)$ *et* $\ker(u) = \ker\left(u^2\right)$, *ce qui équivaut à* $\mathrm{Im}\,(u) = \mathrm{Im}\,(u^2)$.

4.4 Expression matricielle des applications linéaires

Nous avons déjà introduit les matrices en utilisant les bases canoniques de $\mathbb{R}^n$ et $\mathbb{R}^m$. En fait tout peut être repris dans le cadre des espaces vectoriels de dimension fini en utilisant des bases de ces espaces, les démonstrations étant identiques à celles du paragraphe 3.5.

On se donne un espace vectoriel E de dimension n, une base $\mathcal{B} = (e_k)_{1\leq k\leq n}$ de E, un espace vectoriel F de dimension m, une base $\mathcal{B}' = (f_k)_{1\leq k\leq m}$ de F et une application linéaire u de E dans F.

Pour tout entier j compris entre 1 et n, il existe des réels a_{ij} tels que :

$$u(e_j) = \sum_{i=1}^{m} a_{ij} f_i$$

et pour tout vecteur $x = \sum_{j=1}^{n} x_j e_j$ dans E, on a :

$$u(x) = \sum_{j=1}^{n} x_j \sum_{i=1}^{m} a_{ij} f_i = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} x_j \right) f_i$$

c'est-à-dire que les composantes dans la base $\mathcal{B}'$ de $u(x)$ sont les :

$$y_i = \sum_{j=1}^{n} a_{ij} x_j \ \ (1 \leq i \leq m)$$

Définition 4.7 *Avec les notations qui précèdent, on dit que la matrice :*

$$A = ((a_{ij}))_{\substack{1\leq i\leq m \\ 1\leq j\leq n}}$$

est la matrice de u dans les base $\mathcal{B}$ et $\mathcal{B}'$.
Dans le cas ou $E = F$ et $\mathcal{B} = \mathcal{B}'$, on dira simplement que A est la matrice de u dans la base $\mathcal{B}$.

Pour tout j compris entre 1 et n, la colonne numéro j de la matrice A est le vecteur :

$$C_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

formé des composantes de $u(e_j)$ dans la base $\mathcal{B}'$.

L'égalité $y = u(x)$ se traduit alors par le produit matriciel :

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} = AX$$

où les x_j, pour j compris entre 1 et n sont les composantes de x dans la base $\mathcal{B}$ et les y_i, pour i compris entre 1 et m celles de $u(x)$ dans la base $\mathcal{B}'$.

Exercice 4.8 *Pour tout entier naturel non nul n, on désigne par $\mathbb{R}_n[x]$ l'espace vectoriel des fonctions polynomiales à coefficients réels et de degré au plus égal à n. Pour $n \geq 1$, on considère l'application :*

$$\begin{array}{cccc} u & \mathbb{R}_n[x] & \to & \mathbb{R}_n[x] \\ & P & \mapsto & xP' \end{array}$$

où on a noté, pour toute fonction polynomiale $P \in \mathbb{R}_n[x]$, P' le polynôme dérivé de P.

1. *Montrer que u est une application linéaire de E dans E.*
2. *Donner la matrice de u dans la base canonique $\mathcal{B} = (1, x, x^2, \cdots, x^n)$ de E.*
3. *L'application u est-elle injective ?*
4. *L'application u est-elle surjective ?*
5. *Calculer le noyau, l'image et le rang de u.*
6. *Soit F le sous-espace-vectoriel de E engendré par $(x, x^2, \cdots, x^n)$. Montrer que l'application u est bijective de F sur F.*

Solution 4.7

1. *Pour $P \in E$, on a $P' \in \mathbb{R}_{n-1}[x]$ et $xP' \in \mathbb{R}_n[x] = E$, donc u va bien de E dans E.*
   Pour P, Q dans E et λ, μ dans $\mathbb{R}$, on a :

$$\begin{aligned} u(\lambda P + \mu Q) = x(\lambda P + \mu Q)' &= \lambda(xP') + \mu(xQ') \\ &= \lambda u(P) + \mu u(Q) \end{aligned}$$

   donc u est linéaire.
2. *On a $u(1) = 0$ et pour k entier compris entre 1 et n :*

$$u\left(x^k\right) = kx^k.$$

   La matrice de u dans $\mathcal{B}$ est donc :

$$A = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & n-1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & n \end{pmatrix}$$

3. *On a $u(1) = 0$ avec $1 \neq 0$, donc u non injective.*
4. *Pour tout polynôme $P \in E$, on a $u(P)(0) = 0$ car $u(P) = xP'$, donc 1 n'est pas dans l'image de u et u est non surjective.*
5. *On a $P \in \ker(u)$ si, et seulement si, $xP' = 0$, ce qui équivaut encore à $P' = 0$, soit P constant. Donc $\ker(u)$ est la droite dirigée par le polynôme constant 1.*
   En utilisant le théorème du rang, on déduit que

$$\operatorname{rg}(u) = \dim(E) - 1 = n.$$

   L'image de u est contenu dans l'espace $x\mathbb{R}_{n-1}[x]$ des polynômes de $\mathbb{R}_n[x]$ multiples de x. Réciproquement tout polynôme dans $x\mathbb{R}_{n-1}[x]$ s'écrit xQ avec $Q \in \mathbb{R}_{n-1}[x]$ et désignant par $P \in \mathbb{R}_n[x]$ une primitive de Q, on a $xQ = xP' = u(P)$. Donc $\operatorname{Im}(u) = x\mathbb{R}_{n-1}[x]$.
   On retrouve le rang de u :

$$\operatorname{rg}(u) = \dim\left(x\mathbb{R}_{n-1}[x]\right) = \dim\left(\mathbb{R}_{n-1}[x]\right) = n.$$

6. *On remarque que* $F = x\mathbb{R}_{n-1}[x] = \mathrm{Im}(u)$. *Donc la restriction* v *de* u *à* F *est un endomorphisme de* F. *Son noyau est :*

$$\ker(v) = F \cap \ker(u) = \{0\}$$

il en résulte que u *est un isomorphisme.*

Comme dans le cas de $\mathbb{R}^n$ et $\mathbb{R}^m$, on a le résultat suivant.

Théorème 4.17 *Si* u *et* v *sont deux applications linéaires de* E *dans* F *de matrices respectives* $A = ((a_{i,j}))_{\substack{1\leq i\leq m\\1\leq j\leq n}}$ *et* $B = ((b_{i,j}))_{\substack{1\leq i\leq m\\1\leq j\leq n}}$ *dans les bases* $\mathcal{B}$ *et* $\mathcal{B}'$ *alors, pour tous réels* λ, μ, *l'application linéaire* $\lambda u + \mu v$ *a pour matrice dans ces bases la matrice* $\lambda A + \mu B = ((\lambda a_{i,j} + \mu b_{ij}))_{\substack{1\leq i\leq m\\1\leq j\leq n}}$.

Ce résultat peut se traduire en disant que l'application qui associe à toute application linéaire $u \in \mathcal{L}(E,F)$ sa matrice $A \in \mathcal{M}_{m,n}(\mathbb{R})$ dans les bases $\mathcal{B}$ et $\mathcal{B}'$ est linéaire. Plus précisément, on a le résultat suivant.

Théorème 4.18 *Étant données une base* $\mathcal{B} = (e_k)_{1\leq k\leq n}$ *de* E *et une base* $\mathcal{B}' = (f_k)_{1\leq k\leq m}$ *de* F, *l'application* φ *qui associe à toute application linéaire* $u \in \mathcal{L}(E,F)$ *sa matrice* $A \in \mathcal{M}_{m,n}(\mathbb{R})$ *dans les bases* $\mathcal{B}$ *et* $\mathcal{B}'$ *est un isomorphisme de* $\mathcal{L}(E,F)$ *sur* $\mathcal{M}_{m,n}(\mathbb{R})$.

Démonstration. On vient de voir que φ est linéaire et cette application est bijective du fait qu'une application linéaire $u \in \mathcal{L}(E,F)$ est uniquement déterminée par sa matrice dans les bases $\mathcal{B}$ et $\mathcal{B}'$. ■

En particulier, pour toute matrice $A \in \mathcal{M}_{m,n}(\mathbb{R})$ il existe une unique application linéaire $u \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ ayant pour matrice A dans les bases canoniques de $\mathbb{R}^n$ et $\mathbb{R}^m$. On dira que u est l'application linéaire canoniquement associée à la matrice A.

Pour ce qui est de la matrice d'une composée d'applications linéaires nous avons le résultat suivant où G est un espace vectoriel de dimension r et $\mathcal{B}'' = (g_k)_{1\leq k\leq r}$ une base de G.

Théorème 4.19 *Si* v *est une application linéaire de* E *dans* F *de matrice* $B \in \mathcal{M}_{m,n}(\mathbb{R})$ *dans les bases* $\mathcal{B}$ *et* $\mathcal{B}'$ *et* u *une application linéaire de* F *dans* G *de matrice* $A \in \mathcal{M}_{r,m}(\mathbb{R})$ *dans les bases* $\mathcal{B}'$ *et* $\mathcal{B}''$, *alors la matrice dans les bases* $\mathcal{B}$ *et* $\mathcal{B}''$ *de l'application linéaire* $u \circ v$ *de* E *dans* G *est la matrice produit* $C = AB \in \mathcal{M}_{r,n}(\mathbb{R})$.

On en déduit le résultat suivant, où $\mathcal{B}$ et $\mathcal{B}'$ sont deux bases de E.

Théorème 4.20 *Un endomorphisme* u *de* E *est bijectif si, et seulement si, sa matrice* A *dans les bases* $\mathcal{B}$ *et* $\mathcal{B}'$ *est inversible et dans ce cas* A^{-1} *est la matrice de* u^{-1} *dans les bases* $\mathcal{B}'$ *et* $\mathcal{B}$.

En regardant une matrice comme un ensemble de vecteurs colonnes, on peut donner la définition suivante.

Définition 4.8 *Soit* $A = ((a_{ij}))_{\substack{1\leq i\leq m\\1\leq j\leq n}}$ *une matrice dans* $\mathcal{M}_{m,n}(\mathbb{R})$. *En désignant, pour tout* j *compris entre* 1 *et* n, *par* $C_j = (a_{ij})_{1\leq i\leq n}$ *le vecteur de* $\mathbb{R}^n$ *représentant la colonne numéro* j *de* A, *le rang de* A *est le rang de la famille* $(C_1, C_2, \cdots, C_n)$ *de vecteurs de* $\mathbb{R}^n$.

Théorème 4.21 *Le rang d'une matrice* $A \in \mathcal{M}_{r,m}(\mathbb{R})$ *est égal au rang de l'application linéaire* $u \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ *canoniquement associé à la matrice* A.

Démonstration. En désignant par $\mathcal{B} = (e_k)_{1\leq k\leq n}$ une base de $\mathbb{R}^n$ et par $\mathcal{B}' = (f_k)_{1\leq k\leq m}$ une base de $\mathbb{R}^m$, pour tout j compris entre 1 et n, la colonne numéro j de A est $C_j = u(e_j)$ et :

$$\operatorname{rg}(A) = \operatorname{rg}(C_1, C_2, \cdots, C_n) = \operatorname{rg}(u(e_1), u(e_2), \cdots, u(e_n)) = \operatorname{rg}(u).$$

■

Théorème 4.22 *Si $u \in \mathcal{L}(E,F)$ a pour matrice $A \in \mathcal{M}_{r,m}(\mathbb{R})$ dans les bases $\mathcal{B}$ et $\mathcal{B}'$ (toujours avec les notations du début de ce paragraphe), alors le rang de u est égal au rang de A.*

Démonstration. On utilise la formule du rang.

Dire que $x = \sum_{j=1}^{n} x_j e_j$ est dans le noyau de u équivaut à dire que $u(x) = 0$, ce qui se traduit par $AX = 0$, où $X = (x_j)_{1\leq j\leq n}$ est le vecteur de $\mathbb{R}^n$ formé des composantes de x dans la base $\mathcal{B}$. En désignant par v l'application linéaire canoniquement associée à A, on a $v(X) = AX = 0$, c'est-à-dire que X est dans le noyau de v. Réciproquement si $X \in \ker(v)$, on a $AX = 0$, ce qui équivaut à $u(x) = 0$, soit $x \in \ker(u)$. L'application $x \mapsto X$ réalise donc un isomorphisme de $\ker(u)$ sur $\ker(v)$ et $\dim(\ker(u)) = \dim(\ker(v))$, ce qui équivaut à $\operatorname{rg}(u) = \operatorname{rg}(v)$ en utilisant le théorème du rang. Et donc $\operatorname{rg}(u) = \operatorname{rg}(v) = rg(A)$. ■

4.5 Formules de changement de base

On se donne un espace vectoriel E de dimension n et deux base de E, $\mathcal{B}_1 = (e_k)_{1\leq k\leq n}$ et $\mathcal{B}_2 = (e'_k)_{1\leq k\leq n}$. Une question naturelle est de savoir comment passer des composantes d'un vecteur de E d'une base à l'autre.

Pour tout vecteur $x = \sum_{j=1}^{n} x_j e_j = \sum_{j=1}^{n} x'_j e'_j$ dans E, on note $X = (x_j)_{1\leq j\leq n}$ et $X' = \left(x'_j\right)_{1\leq j\leq n}$ les vecteurs de $\mathbb{R}^n$ respectivement formés des composantes de x dans les bases $\mathcal{B}_1$ et $\mathcal{B}_2$.

Comme, pour j compris entre 1 et n le vecteur e'_j est dans E, il s'écrit :

$$e'_j = \sum_{i=1}^{n} p_{ij} e_i$$

où les p_{ij} sont des réels uniquement déterminés. On a alors :

$$\begin{aligned} x = \sum_{i=1}^{n} x_i e_i = \sum_{j=1}^{n} x'_j e'_j &= \sum_{j=1}^{n} x'_j \left(\sum_{i=1}^{n} p_{ij} e_i\right) \\ &= \sum_{i=1}^{n} \left(\sum_{j=1}^{n} p_{ij} x'_j\right) e_i \end{aligned}$$

et avec l'unicité de l'écriture de x dans la base $\mathcal{B}_1$, on déduit que :

$$x_i = \sum_{j=1}^{n} p_{ij} x'_j \ (1 \leq i \leq n)$$

ce qui peut se traduire par l'égalité matricielle :

$$X = PX'$$

où P est la matrice :

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix}$$

ayant pour colonne j le vecteur de $\mathbb{R}^n$ formé des composantes de e'_j dans la bases $\mathcal{B}_1$.

On peut retenir cette formule sous la forme :

$$X_{\mathcal{B}_1} = P_{\mathcal{B}_1 \to \mathcal{B}_2} X_{\mathcal{B}_2}$$

avec des notations évidentes.

Définition 4.9 *Avec les notations qui précèdent, on dit que P est la matrice de passage de la base $\mathcal{B}_1$ à la base $\mathcal{B}_2$.*

Théorème 4.23 *Avec les notations qui précèdent, la matrice de passage P de $\mathcal{B}_1$ à $\mathcal{B}_2$ est inversible et son inverse est la matrice de passage de $\mathcal{B}_2$ à $\mathcal{B}_1$.*

Démonstration. Il suffit de remarquer que la matrice P de $\mathcal{B}_1$ à $\mathcal{B}_2$ est la matrice dans les bases $\mathcal{B}_2$ et $\mathcal{B}_1$ de l'identité de E. En effet, pour tout j compris entre 1 et n on a :

$$Id_E\left(e'_j\right) = e'_j = \sum_{i=1}^{n} p_{ij} e_i.$$

Cette matrice est donc inversible et son inverse est la matrice dans les bases $\mathcal{B}_1$ et $\mathcal{B}_2$ de $(Id_E)^{-1} = Id_E$ c'est-à-dire la matrice de passage de $\mathcal{B}_2$ à $\mathcal{B}_1$. ■

On a donc :

$$\left(P_{\mathcal{B}_1 \to \mathcal{B}_2}\right)^{-1} = P_{\mathcal{B}_2 \to \mathcal{B}_1}.$$

Le calcul de l'inverse de la matrice de passage P de $\mathcal{B}_1$ à $\mathcal{B}_2$ peut se calculer en résolvant le système linéaire aux inconnues e_i :

$$e'_j = \sum_{i=1}^{n} p_{ij} e_i \ \ (1 \leq i \leq n)$$

Exercice 4.9 *On désigne par $\mathcal{B}_1 = (e_1, e_2, e_3)$ la base canonique de $\mathbb{R}^3$.*

1. *Montrer que $\mathcal{B}_2 = (e'_1 = e_1 - e_2, e'_2 = 2e_2 + e_3, e'_3 = e_1 + e_3)$ est une base de $\mathbb{R}^3$.*
2. *Calculer l'inverse de la matrice de passage P de $\mathcal{B}_1$ à $\mathcal{B}_2$.*

Solution 4.8

1. *L'égalité $\lambda_1 e'_1 + \lambda_2 e'_2 + \lambda_3 e'_3 = 0$ équivaut à :*

$$\begin{cases} \lambda_1 + \lambda_3 = 0 \\ -\lambda_1 + 2\lambda_2 = 0 \\ \lambda_2 + \lambda_3 = 0 \end{cases}$$

En additionnant les deux premières équations, on a $\lambda_3 + 2\lambda_2 = 0$ qui reporté dans la troisième donne $\lambda_2 = 0$. Les équations 2 et 3 donnent alors $\lambda_1 = \lambda_3 = 0$. La famille $\mathcal{B}_2$ est donc libre dans $\mathbb{R}^3$ et c'est une base puisqu'elle a trois éléments.

2. La matrice de passage de $\mathcal{B}_1$ à $\mathcal{B}_2$ est la matrice :

$$P = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 2 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

Avec :

$$\begin{cases} e'_1 = e_1 - e_2 \\ e'_2 = 2e_2 + e_3 \\ e'_3 = e_1 + e_3 \end{cases}$$

on déduit que :

$$\begin{cases} e_1 = 2e'_1 + e'_2 - e'_3 \\ e_2 = e'_1 + e'_2 - e'_3 \\ e_3 = -2e'_1 - e'_2 + 2e'_3 \end{cases}$$

ce qui signifie que :

$$P^{-1} = \begin{pmatrix} 2 & 1 & -2 \\ 1 & 1 & -1 \\ -1 & -1 & 2 \end{pmatrix}$$

On est maintenant en mesure d'exprimer la matrice d'un endomorphisme u de E dans la base $\mathcal{B}_2$ en fonction de sa matrice dans la base $\mathcal{B}_1$.

Théorème 4.24 *Si u est un endomorphisme de E de matrice $A = ((a_{ij}))_{1\leq i,j\leq n}$ dans la base $\mathcal{B}_1$ et de matrice $A' = \left(\left(a'_{ij}\right)\right)_{1\leq i,j\leq n}$ dans la base $\mathcal{B}_2$, on a alors $A' = P^{-1}AP$, où P est la matrice de passage de $\mathcal{B}_1$ à $\mathcal{B}_2$.*

Démonstration. PA' est la matrice de $Id_E \circ u = u$ dans les bases $\mathcal{B}_2$ et $\mathcal{B}_1$ et AP est aussi la matrice de $u \circ Id_E = u$ dans les bases $\mathcal{B}_2$ et $\mathcal{B}_1$. On a donc $PA' = AP$, ce qui équivaut à $A' = P^{-1}AP$. ∎

Cette formule de changement de base peut se retenir sous la forme :

$$A_{\mathcal{B}_2} = P_{\mathcal{B}_2 \to \mathcal{B}_1} A_{\mathcal{B}_1} P_{\mathcal{B}_1 \to \mathcal{B}_2}.$$

Dans le cas où la matrice A' de u dans la base $\mathcal{B}_2$ a une forme plus simple que celle de A, on peut l'utiliser pour calculer les puissances de u ou de A. L'idée étant que l'égalité $A' = P^{-1}AP$ entraîne pour tout entier naturel p, on a $(A')^p = P^{-1}A^pP$ ou encore $A^p = P\,(A')^p\,P^{-1}$. La vérification étant immédiate par récurrence sur $p \geq 0$ (toujours avec la convention $A^0 = I_n$).

Si par exemple, la matrice A' est diagonale, c'est-à-dire de la forme :

$$A' = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

on a alors, pour tout $p \geq 0$:

$$(A')^p = \begin{pmatrix} \lambda_1^p & 0 & \cdots & 0 \\ 0 & \lambda_2^p & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n^p \end{pmatrix}$$

Pour $p \geq 1$, l'endomorphisme u^p est défini par la formule de récurrence $u^p = u^{p-1} \circ u$ avec $u^0 = Id_E$. Si A est la matrice de u dans la base $\mathcal{B}_1$, alors celle de A^p dans cette même base est A^p.

Exercice 4.10 *On désigne par $\mathcal{B}_1$ et $\mathcal{B}_2$ les bases de $\mathbb{R}^3$ de l'exercice 4.9 et par u l'endomorphisme de $(\mathbb{R}^3)$ de matrice $A = \begin{pmatrix} 4 & 2 & -4 \\ -6 & -4 & 6 \\ -1 & -1 & 1 \end{pmatrix}$ dans la base canonique $\mathcal{B}_1$.*

1. *Déterminer la matrice de u dans la base $\mathcal{B}_2$.*
2. *Calculer, pour tout $p \in \mathbb{N}$, la matrice de u^p dans la base canonique de $\mathbb{R}^3$.*

Solution 4.9

1. *La matrice de u dans la base $\mathcal{B}_2$ est :*

$$A' = P^{-1}AP = \begin{pmatrix} 2 & 1 & -2 \\ 1 & 1 & -1 \\ -1 & -1 & 2 \end{pmatrix} \begin{pmatrix} 4 & 2 & -4 \\ -6 & -4 & 6 \\ -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ -1 & 2 & 0 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

2. *La matrice de u^p dans la base $\mathcal{B}_2$ est :*

$$(A')^p = \begin{pmatrix} 2^p & 0 & 0 \\ 0 & (-1)^p & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

et dans la base $\mathcal{B}_1$ c'est :

$$A^p = P\,(A')^p\,P^{-1} = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 2 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 2^p & 0 & 0 \\ 0 & (-1)^p & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 & 1 & -2 \\ 1 & 1 & -1 \\ -1 & -1 & 2 \end{pmatrix} = \begin{pmatrix} 2^{p+1} & 2^p & -2^{p+1} \\ 2\,(-1)^p - 2^{p+1} & 2\,(-1)^p - 2^p & 2^{p+1} - 2\,(-1)^p \\ (-1)^p & (-1)^p & (-1)^{p+1} \end{pmatrix}$$

Exercice 4.11 *Soit $u \in \mathcal{L}\,(\mathbb{R}^3)$ de matrice $A = \begin{pmatrix} -5 & 3 & 3 \\ -8 & 6 & 4 \\ -1 & 1 & 3 \end{pmatrix}$ dans la base canonique $\mathcal{B}_1 = (e_1, e_2, e_3)$.*

1. *Déterminer les images par u des vecteurs $e'_1 = e_1 + e_2$, $e'_2 = e_2 - e_3$, $e'_3 = e_1 + 2e_2 + e_3$.*
2. *Montrer que $\mathcal{B}_2 = (e'_1, e'_2, e'_3)$ est une base de $\mathbb{R}^3$ et donner la matrice de u dans cette base.*
3. *Calculer, pour tout $p \in \mathbb{N}$, la matrice de u^p dans la base canonique de $\mathbb{R}^3$.*

Solution 4.10 *Laissée au lecteur.*

5

Opérations élémentaires et déterminants

On se donne un entier $n \geq 1$ et $\mathcal{M}_n(\mathbb{R})$ désigne l'espace vectoriel des matrices réelles carrées d'ordre n.

Pour i, j entiers compris entre 1 et n, on note E_{ij} la matrice dont tous les coefficients sont nuls sauf celui d'indice (i, j) (ligne i et colonne j) qui vaut 1.

On rappelle que la famille $(E_{ij})_{1 \leq i,j \leq n}$ est une base de $\mathcal{M}_n(\mathbb{R})$ qui est donc de dimension n^2.

Pour toute matrice $A = ((a_{ij}))_{1 \leq i,j \leq n} \in \mathcal{M}_n(\mathbb{R})$, on note pour tout entier i compris entre 1 et n :

$$L_i = (a_{i1}, a_{i2}, \cdots, a_{in})$$

sa ligne numéro i (c'est une matrice à une ligne et n colonnes) et pour tout entier j compris entre 1 et n :

$$C_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix}$$

sa colonne numéro j (c'est une matrice à n lignes et une colonne).

On écrira :

$$A = \begin{pmatrix} L_1 \\ \vdots \\ L_n \end{pmatrix} \text{ ou } A = \begin{pmatrix} C_1 & \cdots & C_n \end{pmatrix}.$$

Définition 5.1 *On dit qu'une matrice $A = ((a_{ij}))_{1 \leq i,j \leq n} \in \mathcal{M}_n(\mathbb{R})$ est triangulaire inférieure [resp. supérieure] si $a_{ij} = 0$ pour $1 \leq i < j \leq n$ [resp. pour $1 \leq j < i \leq n$].*

Une matrice triangulaire inférieure est donc de la forme :

$$A = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

et une matrice triangulaire supérieure de la forme :

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}$$

Définition 5.2 *Une matrice diagonale est une matrice triangulaire inférieure et supérieure.*

Une matrice diagonale est donc de la forme :

$$A = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}$$

5.1 Opérations élémentaires. Matrices de dilatation et de transvection

On suppose que $n \geq 2$.

On appelle matrice déduite de A par opération élémentaire sur les lignes de A toute matrice de la forme :

$$A_i(\lambda) = \begin{pmatrix} L_1 \\ \vdots \\ L_{i-1} \\ \lambda L_i \\ L_{i+1} \\ \vdots \\ L_n \end{pmatrix},$$

avec $1 \leq i \leq n$ et $\lambda \in \mathbb{R}^*$, c'est-à-dire que la matrice $A_i(\lambda)$ est déduite de la matrice A en multipliant sa ligne numéro i par λ ou de la forme :

$$A_{ij}(\lambda) = \begin{pmatrix} L_1 \\ \vdots \\ L_{i-1} \\ L_i + \lambda L_j \\ L_{i+1} \\ \vdots \\ L_n \end{pmatrix}$$

$1 \leq i \neq j \leq n$ et $\lambda \in \mathbb{R}$, c'est-à-dire que la matrice $A_{ij}(\lambda)$ est déduite de la matrice A en ajoutant à la ligne numéro i la ligne numéro j multipliée par λ.

On appelle matrice déduite de A par opération élémentaire sur les colonnes de A toute matrice de la forme :

$$A'_j(\lambda) = \begin{pmatrix} C_1 & \cdots & C_{j-1} & \lambda C_j & C_{j+1} & \cdots & C_n \end{pmatrix},$$

avec $1 \leq j \leq n$ et $\lambda \in \mathbb{R}^*$, c'est-à-dire que la matrice $A'_i(\lambda)$ est déduite de la matrice A en multipliant sa colonne numéro j par λ ou de la forme :

$$A'_{ij}(\lambda) = \begin{pmatrix} C_1 & \cdots & C_{j-1} & C_j + \lambda C_i & C_{j+1} & \cdots & C_n \end{pmatrix}$$

$1 \leq i \neq j \leq n$ et $\lambda \in \mathbb{R}$, c'est-à-dire que la matrice $A'_{ij}(\lambda)$ est déduite de la matrice A en ajoutant à la colonne numéro j la colonne numéro i multipliée par λ.

Définition 5.3 *On appelle matrice de transvection toute matrice de la forme :*

$$T_{ij}(\lambda) = I_n + \lambda E_{ij},$$

avec $1 \leq i \neq j \leq n$ *et* $\lambda \in \mathbb{R}$.

Une matrice de transvection $T_{ij}(\lambda)$ est donc une matrice triangulaire dont tous les termes diagonaux valent 1 et de termes hors de la diagonale tous nuls sauf celui d'indice (i, j) (i. e. en ligne i et colonne j) qui vaut λ.

Définition 5.4 *On appelle matrice de dilatation toute matrice de la forme :*

$$D_i(\lambda) = I_n + (\lambda - 1) E_{ii},$$

avec $1 \leq i \leq n$ *et* $\lambda \in \mathbb{R}^*$.

Une matrice de dilatation $D_i(\lambda)$ est donc diagonale de termes diagonaux tous égaux à 1 sauf le numéro i qui vaut λ.

Théorème 5.1 *Avec les notations qui précèdent on a :*

$$A_i(\lambda) = D_i(\lambda) A,\ A_{ij}(\lambda) = T_{ij}(\lambda) A,$$
$$A'_j(\lambda) = AD_j(\lambda),\ A'_{ij}(\lambda) = AT_{ij}(\lambda).$$

C'est-à-dire que :

1. *la multiplication à gauche par une matrice de dilatation $D_i(\lambda)$ a pour effet de multiplier la ligne i par λ ;*
2. *la multiplication à droite par une matrice de dilatation $D_j(\lambda)$ a pour effet de multiplier la colonne j par λ ;*
3. *la multiplication à gauche par une matrice de transvection $T_{ij}(\lambda)$ a pour effet de remplacer la ligne L_i par $L_i + \lambda L_j$;*
4. *la multiplication à droite par une matrice de transvection $T_{ij}(\lambda)$ a pour effet de remplacer la colonne C_j par $C_j + \lambda C_i$.*

Démonstration. Le coefficient d'indice (p, q) du produit de matrices $D_i(\lambda) A$ est obtenu en faisant le produit de la ligne p de $D_i(\lambda)$ par la colonne q de A, ce qui donne en notant $\alpha_{p,q}$ ce coefficient :

$$\alpha_{p,q} = \begin{cases} a_{p,q} \text{ si } 1 \leq p \neq i \leq n,\ 1 \leq q \leq n, \\ \lambda a_{iq} \text{ si } p = i,\ 1 \leq q \leq n. \end{cases}$$

On a donc bien $A_i(\lambda) = D_i(\lambda) A$.

Les autres égalités se montrent de façon analogue. ■

Ce résultat justifie la définition suivante.

Définition 5.5 *On appelle matrice élémentaire une matrice de dilatation ou de transvection.*

Lemme 5.1 *Une matrice élémentaire est inversible avec :*

$$T_{ij}(\lambda)^{-1} = T_{ij}(-\lambda),$$

pour une matrice de transvection et :

$$D_i(\lambda)^{-1} = D_i\left(\frac{1}{\lambda}\right),$$

pour une matrice de dilatation.

Démonstration. Pour λ, μ réels et $i \neq j$ compris entre 1 et n, la matrice $T_{ij}(\lambda) T_{ij}(\mu)$ se déduit de $T_{ij}(\mu)$ en ajoutant à sa ligne i sa ligne j multipliée par λ, ce qui donne la matrice $T_{ij}(\lambda + \mu)$.

Prenant $\mu = -\lambda$, on a $T_{ij}(\lambda) T_{ij}(-\lambda) = T_0 = I_n$, ce qui signifie que $T_{ij}(\lambda)$ est inversible d'inverse $T_{ij}(-\lambda)$.

Le deuxième résultat est évident. ■

Avec l'exercice qui suit, on vérifie que toute matrice inversible d'ordre 2 est produit de matrices élémentaires. Ce résultat est en fait vrai pour les matrices inversibles d'ordre $n \geq 2$.

Exercice 5.1 *Soit* $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ *une matrice inversible.*

1. *On suppose que* $c \neq 0$.

   (a) *Déterminer un réel* λ_1 *tel que :*

   $$A_1 = T_{12}(\lambda_1) A = \begin{pmatrix} 1 & b_1 \\ c_1 & d_1 \end{pmatrix}$$

   (b) *Déterminer un réel* λ_2 *tel que :*

   $$A_2 = T_{21}(\lambda_2) A_1 = \begin{pmatrix} 1 & b_2 \\ 0 & d_2 \end{pmatrix}$$

   (c) *Déterminer un réel* λ_3 *tel que :*

   $$A_3 = A_2 T_{12}(\lambda_3) = \begin{pmatrix} 1 & 0 \\ 0 & d_1 \end{pmatrix}$$

   (d) *En déduire qu'il existe des matrices de transvection* P_1, P_2, Q_1 *et une matrice de dilatation* D *telles que :*

   $$A = P_1 P_2 D Q_1$$

2. *Donner un résultat analogue dans le cas où* $c = 0$.

Solution 5.1

1.

   (a) *Pour tout réel* λ_1, *on a :*

   $$T_{12}(\lambda_1) A = \begin{pmatrix} 1 & \lambda_1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a + c\lambda_1 & b + d\lambda_1 \\ c & d \end{pmatrix}$$

   et prenant λ_1 *tel que* $a + c\lambda_1 = 1$, *soit* $\lambda_1 = \dfrac{1 - a}{c}$, *on a :*

   $$T_{12}(\lambda_1) A = \begin{pmatrix} 1 & \dfrac{d - \det(A)}{c} \\ c & d \end{pmatrix}$$

(b) Pour tout réel λ_2, on a :

$$T_{21}(\lambda_2) A_1 = \begin{pmatrix} 1 & 0 \\ \lambda_2 & 1 \end{pmatrix} \begin{pmatrix} 1 & b_1 \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & b_1 \\ c+\lambda_2 & d+b_1\lambda_2 \end{pmatrix}$$

et prenant $\lambda_2 = -c$, on a :

$$T_{21}(\lambda_2) A_1 = \begin{pmatrix} 1 & b_1 \\ 0 & d-cb_1 \end{pmatrix} = \begin{pmatrix} 1 & \dfrac{d-\det(A)}{c} \\ 0 & \det(A) \end{pmatrix}$$

(c) Pour tout réel λ_3, on a :

$$A_2 T_{12}(\lambda_3) = \begin{pmatrix} 1 & b_2 \\ 0 & d_2 \end{pmatrix} \begin{pmatrix} 1 & \lambda_3 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & b_2+\lambda_3 \\ 0 & d_2 \end{pmatrix}$$

et prenant $\lambda_3 = -b_2$, on a :

$$A_2 T_{12}(\lambda_3) = \begin{pmatrix} 1 & 0 \\ 0 & d_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \det(A) \end{pmatrix}$$

(d) On a donc en définitive :

$$T_{21}(\lambda_2) T_{12}(\lambda_1) A T_{12}(\lambda_3) = D(\det(A))$$

et utilisant le fait que les matrices de transvections sont inversibles, on déduit que :

$$A = T_{12}(-\lambda_1) T_{21}(-\lambda_2) D(\det(A)) T_{12}(-\lambda_3)$$

où $\lambda_1 = \dfrac{1-a}{c}$, $\lambda_2 = -c$ et $\lambda_3 = \dfrac{\det(A)-d}{c}$.

2. Si $c = 0$, on a nécessairement $a \neq 0$ puisque A est inversible et :

$$T_{21}(1) A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} = \begin{pmatrix} a & b \\ a & b+d \end{pmatrix}$$

ce qui nous ramène au cas précédent et donne :

$$T_{21}(1) A = P_1 P_2 D Q_1$$

soit :

$$A = T_{21}(-1) P_1 P_2 D Q_1.$$

De manière plus générale, on a le résultat suivant.

Théorème 5.2 *Une matrice $A \in \mathcal{M}_n(\mathbb{R})$ (où $n \geq 2$) est inversible si, et seulement si, elle est produit de matrices élémentaires. Précisément si $A \in \mathcal{M}_n(\mathbb{R})$ est inversible, il existe alors des matrices de transvection $P_1, \cdots, P_r$ et $Q_1, \cdots, Q_s$ et une matrice de dilatation $D_n(\lambda)$ telles que :*

$$A = P_1 \cdots P_r D_n(\lambda) Q_1 \cdots Q_s.$$

Démonstration. On procède par récurrence sur $n \geq 2$.

Le cas $n = 2$ a été traité avec l'exercice précédent.

On le suppose vrai pour toutes les matrices inversibles d'ordre $n-1 \geq 2$ et on se donne une matrice inversible A d'ordre n.

On se ramène tout d'abord par opération élémentaire au cas où $a_{21} \neq 0$. Si $a_{21} = 0$, comme A est inversible, sa colonne 1 n'est pas nulle (cette colonne est Ae_1 où e_1 est le premier vecteur de base canonique et $x = 0$ est l'unique solution de $Ax = 0$), il existe donc un indice $i \in \{1, 3, \cdots, n\}$ tel que $a_{i1} \neq 0$ et la matrice $T_{2i}(1) A$ (déduite de A en ajoutant la ligne i à la ligne 2) est telle que son coefficient d'indice $(2, 1)$ est non nul.

Une fois ramené à $a_{21} \neq 0$, on se ramène à $a_{11} = 1$ en remplaçant la première ligne L_1 par $L_1 + \lambda L_2$ (multiplication à gauche par $T_{12}(\lambda)$) où le scalaire λ est choisi tel que $a_{11} + \lambda a_{21} = 1$.

Ensuite, pour tout $i \in \{2, 3, \cdots, n\}$, en remplaçant la ligne L_i par $L_i - a_{i1} L_1$ (multiplication à gauche par $T_{i1}(-a_{i1})$), on annule le coefficient d'indice $(i, 1)$.

On peut donc trouver des matrices de transvection $P_1, \cdots, P_k$ telles que :

$$P_k \cdots P_1 A = \begin{pmatrix} 1 & \alpha_{12} & \cdots & \alpha_{1n} \\ 0 & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \alpha_{n2} & \cdots & \alpha_{nn} \end{pmatrix}.$$

De manière analogue, en multipliant à droite par des matrices de transvection, $Q_1, \cdots, Q_m$, on obtient :

$$P_k \cdots P_1 A Q_1 \cdots Q_m = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \beta_{22} & \cdots & \beta_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \beta_{n2} & \cdots & \beta_{nn} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{pmatrix}$$

On peut alors conclure en appliquant l'hypothèse de récurrence à la matrice B qui est d'ordre $n-1$ et inversible. En effet, si B n'est pas inversible, il existe $x' \neq 0$ dans $\mathbb{R}^{n-1}$ tel que $Bx' = 0$, donc $x = \begin{pmatrix} 0 \\ x' \end{pmatrix} \in \mathbb{R}^n$ est non nul solution de $P_k \cdots P_1 A Q_1 \cdots Q_m x = 0$ qui équivaut à $Ay = 0$ avec $y = Q_1 \cdots Q_m x \neq 0$ puisque les matrices P_i et Q_j sont inversibles, en contradiction avec A inversible. ∎

Nous verrons plus loin (paragraphe 5) que, comme dans le cas $n = 2$, le réel λ qui intervient dans le théorème précédent est uniquement déterminé par la matrice A, c'est son déterminant.

Pour $n = 1$, le résultat est encore vrai avec $A = (a) = D(a)$.

5.2 Déterminants des matrices carrées

Nous avons déjà défini le déterminant d'une matrice réelle d'ordre deux, $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ par :

$$\det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

(définition 3.14).

Une matrice d'ordre 1 étant tout simplement un réel, son déterminant est lui même.

Le déterminant d'une matrice carrée $A = ((a_{ij}))_{1 \leq i,j \leq n}$ d'ordre $n \geq 3$ peut se définir par récurrence comme suit :

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+1} a_{i,1} \det(A_{i,1})$$

où $A_{i,1}$ est, pour i compris entre 1 et n, la matrice d'ordre $n-1$ déduite de A en supprimant la première colonne et la ligne numéro i.

Dans cette expression, on dit qu'on développe le déterminant suivant la première colonne.

On note :

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

Les lignes d'une matrice $A \in \mathcal{M}_n(\mathbb{R})$ étant notées $L_1, L_2, \cdots, L_n$, on écrira aussi :

$$\det(A) = \det \begin{pmatrix} L_1 \\ L_2 \\ \vdots \\ L_n \end{pmatrix}.$$

Exemple 5.1 *Pour $n = 3$ et $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$, on a :*

$$\det(A) = \begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} - 4 \begin{vmatrix} 2 & 3 \\ 8 & 9 \end{vmatrix} + 7 \begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix} = 0$$

Exercice 5.2 *Soient $\alpha_1, \alpha_2, \alpha_3$ des réels. Calculer le déterminant de la matrice :*

$$V(\alpha_1, \alpha_2, \alpha_3) = \begin{pmatrix} 1 & 1 & 1 \\ \alpha_1 & \alpha_2 & \alpha_3 \\ \alpha_1^2 & \alpha_2^2 & \alpha_3^2 \end{pmatrix}$$

Une telle matrice est dite de Vandermonde.

Solution 5.2 *On a :*

$$\begin{aligned} \det(V(\alpha_1, \alpha_2, \alpha_3)) &= \begin{vmatrix} \alpha_2 & \alpha_3 \\ \alpha_2^2 & \alpha_3^2 \end{vmatrix} - \alpha_1 \begin{vmatrix} 1 & 1 \\ \alpha_2^2 & \alpha_3^2 \end{vmatrix} + \alpha_1^2 \begin{vmatrix} 1 & 1 \\ \alpha_2 & \alpha_3 \end{vmatrix} \\ &= \alpha_2\alpha_3^2 - \alpha_2^2\alpha_3 - \alpha_1\left(\alpha_3^2 - \alpha_2^2\right) + \alpha_1^2(\alpha_3 - \alpha_2) \\ &= \alpha_2\alpha_3(\alpha_3 - \alpha_2) - \alpha_1(\alpha_3 - \alpha_2)(\alpha_3 + \alpha_2) + \alpha_1^2(\alpha_3 - \alpha_2) \\ &= (\alpha_3 - \alpha_2)\left(\alpha_2\alpha_3 - \alpha_1(\alpha_3 + \alpha_2) + \alpha_1^2\right) \\ &= (\alpha_3 - \alpha_2)(\alpha_2(\alpha_3 - \alpha_1) - \alpha_1(\alpha_3 - \alpha_1)) \\ &= (\alpha_3 - \alpha_2)(\alpha_3 - \alpha_1)(\alpha_2 - \alpha_1) \end{aligned}$$

Exercice 5.3 *Calculer le déterminant de la matrice :*

$$A = \begin{pmatrix} 5 & 4 & 2 & 1 \\ 2 & 3 & 1 & -2 \\ -5 & -7 & -3 & 9 \\ 1 & -2 & -1 & 4 \end{pmatrix}.$$

Solution 5.3 *On a* $\det(A) = 38$.

Théorème 5.3 *Si $A_i(\lambda)$ est la matrice déduite de $A \in \mathcal{M}_n(\mathbb{R})$ en multipliant sa ligne i par un réel λ, on a alors* $\det(A_i(\lambda)) = \lambda \det(A)$, *soit :*

$$\det \begin{pmatrix} L_1 \\ \vdots \\ L_{i-1} \\ \lambda L_i \\ L_{i+1} \\ \vdots \\ L_n \end{pmatrix} = \lambda \det \begin{pmatrix} L_1 \\ \vdots \\ L_{i-1} \\ L_i \\ L_{i+1} \\ \vdots \\ L_n \end{pmatrix}$$

Démonstration. On procède par récurrence sur $n \geq 1$.
Pour $n = 1$ c'est clair et pour $n = 2$, on a :

$$\begin{vmatrix} \lambda a & \lambda b \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ \lambda c & \lambda d \end{vmatrix} = \lambda (ad - bc) = \lambda \det(A).$$

Supposons le résultat acquis pour les matrices d'ordre $n - 1 \geq 2$. Soient A d'ordre n et $A' = A_i(\lambda)$ déduite de A en multipliant sa ligne i par λ. On a alors :

$$\det(A') = (-1)^{i+1} \lambda a_{i,1} \det(A_{i,1}) + \sum_{\substack{k=1 \\ k \neq i}}^{n} (-1)^{k+1} a_{k,1} \det(A'_{k1})$$

la matrice $A'_{k,1}$, pour $k \neq i$, étant déduite de $A_{k,1}$ en multipliant sa ligne i par λ. On a donc $\det\left(A'_{k,1}\right) = \lambda \det(A_{k,1})$ pour $k \neq i$ et $\det(A') = \lambda \det(A)$. ■

Corollaire 5.1 *Si $A \in \mathcal{M}_n(\mathbb{R})$ a une ligne nulle, alors* $\det(A) = 0$.

Démonstration. Supposons que la ligne i de A soit nulle. En désignant par $A' = A_i(\lambda)$ la matrice déduite de A en multipliant sa ligne i par $\lambda = 0$, on a $A' = A$ et $\det(A) = \det(A') = 0 \det(A) = 0$. ■

Corollaire 5.2 *Pour tout $A \in \mathcal{M}_n(\mathbb{R})$ et tout réel λ, on a* $\det(\lambda A) = \lambda^n \det(A)$.

Démonstration. En utilisant n fois le théorème précédent, on a :

$$\det(\lambda A) = \det \begin{pmatrix} \lambda L_1 \\ \lambda L_2 \\ \vdots \\ \lambda L_n \end{pmatrix} = \lambda \det \begin{pmatrix} L_1 \\ \lambda L_2 \\ \vdots \\ \lambda L_n \end{pmatrix}$$

$$= \cdots = \lambda^n \det \begin{pmatrix} L_1 \\ L_2 \\ \vdots \\ L_n \end{pmatrix}.$$

■

Théorème 5.4 *Le déterminant d'une matrice triangulaire est égale au produit de ses termes diagonaux, soit :*

$$\det(A) = \prod_{i=1}^{n} a_{ii}$$

Démonstration. Considérons tout d'abord le cas des matrices triangulaires inférieures.
On procède par récurrence sur $n \geq 1$.
Pour $n = 1$ c'est clair et pour $n = 2$, on a :

$$\begin{vmatrix} a & 0 \\ c & d \end{vmatrix} = ad - 0 \cdot c = ad.$$

Supposons le résultat acquis pour les matrices triangulaires inférieures d'ordre $n - 1 \geq 2$ et soit A triangulaire inférieure d'ordre n. La matrice A_{11} est triangulaire inférieure de diagonale $a_{22}, \cdots a_{nn}$ et pour i compris entre 2 et n, la matrice A_{i1} est telle que sa première ligne est nulle, elle est donc de déterminant nul et :

$$\det(A) = a_{1,1} \det(A_{1,1}) = \prod_{i=1}^{n} a_{ii}.$$

Pour le cas des matrices triangulaires supérieures, le cas $n = 1$ est encore évident et le cas $n = 2$ se vérifie par le calcul. Supposant le résultat acquis au rang $n-1 \geq 2$, pour A triangulaire supérieure d'ordre n, La matrice A_{11} est triangulaire supérieure de diagonale $a_{22}, \cdots a_{nn}$ et pour i compris entre 2 et n, les coefficients a_{i1} sont nuls de sorte que :

$$\det(A) = a_{1,1} \det(A_{1,1}) = \prod_{i=1}^{n} a_{ii}.$$

■

Exemple 5.2 *Si $A = I_n$ est la matrice identité, on a alors* $\det(I_n) = 1$.

Exemple 5.3 *Si $A = D_i(\lambda)$ est une matrice de dilatation, on a alors* $\det(D_i(\lambda)) = \lambda$.

Exemple 5.4 *Si $A = T_{ij}(\lambda)$ est une matrice de transvection, on a alors* $\det(T_{ij}(\lambda)) = 1$.

Théorème 5.5 *Soient A, A', A'' des matrices de lignes respectives L_i, L'_i, L''_i (pour i compris entre 1 et n) telles que $L_i = L'_i = L''_i$ pour $i \neq k$ et $L''_k = L_k + L'_k$ où k est un indice compris entre 1 et n. On a :*

$$\det(A'') = \det(A) + \det(A')$$

soit :

$$\det \begin{pmatrix} L_1 \\ \vdots \\ L_{k-1} \\ L_k + L'_k \\ L_{k+1} \\ \vdots \\ L_n \end{pmatrix} = \det \begin{pmatrix} L_1 \\ \vdots \\ L_{k-1} \\ L_k \\ L_{k+1} \\ \vdots \\ L_n \end{pmatrix} + \det \begin{pmatrix} L_1 \\ \vdots \\ L_{k-1} \\ L'_k \\ L_{k+1} \\ \vdots \\ L_n \end{pmatrix}$$

Démonstration. On procède par récurrence sur $n \geq 1$.
Pour $n = 1$ c'est clair et pour $n = 2$, il suffit de calculer.
Supposons le résultat acquis pour les matrices d'ordre $n - 1 \geq 2$. Soient A, A', A'' d'ordre n vérifiant les conditions du théorème. On a alors :

$$\det(A'') = (-1)^{k+1} (a_{k,1} + a'_{k1}) \det\left(A''_{k,1}\right) + \sum_{\substack{i=1 \\ i \neq k}}^{n} (-1)^{i+1} a_{i,1} \det\left(A''_{i,1}\right)$$

avec $A''_{k,1} = A_{k,1} = A'_{k,1}$ et pour $i \neq k$, les matrices $A_{i,1}, A'_{i,1}, A''_{i1}$ vérifiant les hypothèses du théorème au rang $n-1$ (avec des notations évidentes), donc :

$$\begin{aligned}\det(A'') &= (-1)^{k+1}\left(a_{k,1}\det(A_{k,1}) + a'_{k1}\det\left(A'_{k,1}\right)\right)\\ &+\sum_{\substack{i=1\\ i\neq k}}^{n}(-1)^{i+1}a_{i,1}\det(A_{i,1}) + \sum_{\substack{i=1\\ i\neq k}}^{n}(-1)^{i+1}a'_{i,1}\det\left(A'_{i,1}\right)\\ &= \det(A) + \det(A').\end{aligned}$$

■

Les théorèmes 5.3 et 5.5 se traduisent en disant que le déterminant est linéaire par rapport à chaque ligne.

Théorème 5.6 *Si A' est la matrice déduite de $A \in \mathcal{M}_n(\mathbb{R})$ en permutant deux lignes, on a alors* $\det(A') = -\det(A)$, *soit :*

$$\det\begin{pmatrix}\vdots\\ L_i\\ \vdots\\ L_j\\ \vdots\end{pmatrix} = -\det\begin{pmatrix}\vdots\\ L_j\\ \vdots\\ L_i\\ \vdots\end{pmatrix}$$

où les pointillés indiquent les lignes inchangées.

Démonstration. On procède par récurrence sur $n \geq 2$.

Pour $n = 2$, il suffit de calculer.

Supposons le résultat acquis pour les matrices d'ordre $n-1 \geq 2$.

La permutation de deux lignes se faisant avec un nombre impair de permutations de deux lignes successives (par exemple la permutation $(2,4)$ se fait par les trois permutations $(2,3,4) \to (3,2,4) \to (3,4,2) \to (4,3,2)$), il suffit de considérer le cas où $j = i+1$ (montrer ce point rigoureusement). On se donne donc A d'ordre n et A' est déduite de $A \in \mathcal{M}_n(\mathbb{R})$ en permutant les lignes i et $i+1$. Pour $k \neq i$ et $k \neq i+1$, on a $a'_{k1} = a_{k,1}$ et $\det\left(A'_{k,1}\right) = -\det(A_{k,1})$ par hypothèse de récurrence, et avec $a'_{i,1} = a_{(i+1),1}$, $a'_{(i+1),1} = a_{i,1}$, $A'_{i,1} = A_{(i+1),1}$, $A'_{(i+1),1} = A_{i,1}$, on déduit que :

$$\begin{aligned}\det(A') &= (-1)^{i+1}a_{(i+1),1}\det\left(A_{(i+1),1}\right) + (-1)^{i}a_{i,1}\det(A_{i,1})\\ &- \sum_{\substack{k=1\\ k\neq i, k\neq i+1}}^{n}(-1)^{k+1}a_{k,1}\det(A_{k,1}) = -\det(A).\end{aligned}$$

■

Le résultat précédent se traduit en disant que le déterminant est une forme alternée sur les lignes.

Corollaire 5.3 *Si la matrice $A \in \mathcal{M}_n(\mathbb{R})$ a deux lignes identiques, alors* $\det(A) = 0$.

Démonstration. Si $L_i = L_j$ avec $i \neq j$, alors matrice A' déduite de A en permutant ces deux lignes est égale à A et $\det(A) = -\det(A)$, donc $\det(A) = 0$. ■

Corollaire 5.4 *On ne change pas la valeur d'un déterminant si on ajoute à une ligne une combinaison linéaire des autres lignes.*

Démonstration. Il suffit de montrer le résultat quand on ajoute à la ligne L_i la ligne L_j multipliée par un réel λ où $i \neq j$. Dans ce cas, on a :

$$\det\begin{pmatrix} \vdots \\ L_i + \lambda L_j \\ \vdots \\ L_j \\ \vdots \end{pmatrix} = \det\begin{pmatrix} \vdots \\ L_j \\ \vdots \\ L_j \\ \vdots \end{pmatrix} + \lambda \det\begin{pmatrix} \vdots \\ L_j \\ \vdots \\ L_j \\ \vdots \end{pmatrix} = \det\begin{pmatrix} \vdots \\ L_j \\ \vdots \\ L_j \\ \vdots \end{pmatrix}$$

où les pointillés indiquent les lignes inchangées. ■

En effectuant des opérations élémentaires sur les lignes d'une matrice A, on peut se ramener à une matrice triangulaire supérieure de même déterminant que celui de A.

Exercice 5.4 *Calculer le déterminant de la matrice :*

$$A = \begin{pmatrix} 5 & 4 & 2 & 1 \\ 2 & 3 & 1 & -2 \\ -5 & -7 & -3 & 9 \\ 1 & -2 & -1 & 4 \end{pmatrix}$$

de l'exercice 5.3 en effectuant des opérations élémentaires.

Solution 5.4 *Les opérations* $L_2 \to L_2 - \frac{2}{5}L_1$, $L_3 \to L_3 + L_1$, $L_4 \to L_4 - \frac{1}{5}L_1$ *donnent :*

$$\begin{aligned} \det(A) &= \begin{vmatrix} 5 & 4 & 2 & 1 \\ 0 & \frac{7}{5} & \frac{1}{5} & -\frac{12}{5} \\ 0 & -3 & -1 & 10 \\ 0 & -\frac{14}{5} & -\frac{7}{5} & \frac{19}{5} \end{vmatrix} = 5 \begin{vmatrix} \frac{7}{5} & \frac{1}{5} & -\frac{12}{5} \\ -3 & -1 & 10 \\ -\frac{14}{5} & -\frac{7}{5} & \frac{19}{5} \end{vmatrix} \\ &= 5\frac{1}{5}\frac{1}{5} \begin{vmatrix} 7 & 1 & -12 \\ -3 & -1 & 10 \\ -14 & -7 & 19 \end{vmatrix} = \frac{1}{5} \begin{vmatrix} 7 & 1 & -12 \\ -3 & -1 & 10 \\ -14 & -7 & 19 \end{vmatrix} \end{aligned}$$

Puis les opérations $L_2 \to L_2 + \frac{3}{7}L_1$, $L_3 \to L_3 + \frac{14}{7}L_1 = L_3 + 2L_1$ *donnent :*

$$\begin{aligned} \det(A) &= \frac{1}{5} \begin{vmatrix} 7 & 1 & -12 \\ 0 & -\frac{4}{7} & \frac{34}{7} \\ 0 & -5 & -5 \end{vmatrix} = \frac{1}{5} \cdot 7 \cdot \frac{2}{7} \cdot 5 \begin{vmatrix} -2 & 17 \\ -1 & -1 \end{vmatrix} \\ &= 2 \cdot 19 = 38 \end{aligned}$$

Exercice 5.5 *Développer le déterminant de la matrice suivante sous la forme d'un produit de facteurs linéaires en* x *:*

$$A(x) = \begin{pmatrix} x+2 & 2x+3 & 3x+4 \\ 2x+3 & 3x+4 & 4x+5 \\ 3x+5 & 5x+8 & 10x+17 \end{pmatrix}.$$

Solution 5.5 *Les opérations $L_3 \to L_3 - L_2$, $L_2 \to L_2 - L_1$ (dans l'ordre indiqué) donnent :*

$$\det(A(x)) = \begin{vmatrix} x+2 & 2x+3 & 3x+4 \\ x+1 & x+1 & x+1 \\ x+2 & 2x+4 & 6x+12 \end{vmatrix}$$
$$= (x+1)(x+2) \begin{vmatrix} x+2 & 2x+3 & 3x+4 \\ 1 & 1 & 1 \\ 1 & 2 & 6 \end{vmatrix}$$

puis $L_3 \to L_3 - L_2$ donne :

$$\det(A(x)) = (x+1)(x+2) \begin{vmatrix} x+2 & 2x+3 & 3x+4 \\ 1 & 1 & 1 \\ 0 & 1 & 5 \end{vmatrix}$$
$$= (x+1)(x+2)\left((x+2)\begin{vmatrix} 1 & 1 \\ 1 & 5 \end{vmatrix} - \begin{vmatrix} 2x+3 & 3x+4 \\ 1 & 5 \end{vmatrix}\right)$$
$$= (x+1)(x+2)(4(x+2) - (7x+11))$$
$$= -(x+1)(x+2)(3x+3) = -3(x+1)^2(x+2)$$

Exercice 5.6 *Soient α, β deux réels et $A(\alpha,\beta) = ((a_{ij}))_{1\le i,j\le n}$ la matrice d'ordre $n \ge 3$ définie par :*

$$\forall i \in \{1,2,\cdots,n\}, \quad \begin{cases} a_{ii} = \beta, \\ a_{ij} = \alpha \text{ si } j \in \{1,2,\cdots,n\} - \{i\}. \end{cases}$$

Calculer $\Delta(\alpha,\beta) = \det(A(\alpha,\beta))$.

Solution 5.6 *La matrice $A(\alpha,\beta)$ est de la forme :*

$$A(\alpha,\beta) = \begin{pmatrix} \beta & \alpha & \alpha & \cdots & \alpha \\ \alpha & \beta & \alpha & \cdots & \alpha \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \alpha & \cdots & \alpha & \beta & \alpha \\ \alpha & \cdots & \alpha & \alpha & \beta \end{pmatrix}.$$

Si $\alpha = 0$, la matrice est diagonale et :

$$\Delta(0,\beta) = \beta^n.$$

On suppose que $\alpha \neq 0$.
En ajoutant les lignes 2 à n à la première ligne on a :

$$\Delta(\alpha,\beta) = (\beta + (n-1)\alpha) \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ \alpha & \beta & \alpha & \cdots & \alpha \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \alpha & \cdots & \alpha & \beta & \alpha \\ \alpha & \cdots & \alpha & \alpha & \beta \end{vmatrix}.$$

Puis en retranchant la première ligne multipliée par α aux lignes 2 à n on obtient :

$$\Delta(\alpha,\beta) = (\beta + (n-1)\alpha) \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & \beta-\alpha & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \beta-\alpha & 0 \\ 0 & \cdots & 0 & 0 & \beta-\alpha \end{vmatrix}$$
$$= (\beta + (n-1)\alpha)(\beta-\alpha)^{n-1}.$$

Exercice 5.7 *En admettant que* 1700, 1020, 1122 *et* 1309 *sont tous divisibles par* 17, *montrer sans le calculer que le déterminant :*

$$D = \begin{vmatrix} 1 & 7 & 0 & 0 \\ 1 & 0 & 2 & 0 \\ 1 & 1 & 2 & 2 \\ 1 & 3 & 0 & 9 \end{vmatrix}$$

est divisible par 17.

Solution 5.7 *On ne change pas la valeur de ce déterminant si on remplace la colonne* 4 *par* $C_4 + 10C_3 + 10^2C_2 + 10^3C_1$, *ce qui donne :*

$$D = \begin{vmatrix} 1 & 7 & 0 & 1700 \\ 1 & 0 & 2 & 1020 \\ 1 & 1 & 2 & 1122 \\ 1 & 3 & 0 & 1309 \end{vmatrix} = 17 \begin{vmatrix} 1 & 7 & 0 & 100 \\ 1 & 0 & 2 & 60 \\ 1 & 1 & 2 & 66 \\ 1 & 3 & 0 & 77 \end{vmatrix} = 17q$$

avec q *entier puisque tous les coefficients du déterminant sont entiers.*

Les théorèmes 5.3, 5.5 et le corollaire 5.1 se traduisent aussi par le résultat suivant.

Corollaire 5.5 *Pour toute matrice* $A \in \mathcal{M}_n(\mathbb{R})$, *toute matrice de dilatation* $D_i(\lambda)$ *et toute matrice de transvection* $T_{ij}(\lambda)$, *on a :*

$$\begin{cases} \det(D_i(\lambda) A) = \det(D_i(\lambda)) \det(A) = \lambda \det(A) \\ \det(T_{ij}(\lambda) A) = \det(T_{ij}(\lambda)) \det(A) = \det(A) \end{cases}$$

En utilisant le théorème 5.2, on obtient le résultat suivant.

Théorème 5.7 *Pour toute matrice inversible* A *et toute matrice* B *dans* $\mathcal{M}_n(\mathbb{R})$, *on a* $\det(AB) = \det(A)\det(B)$.

Démonstration. La matrice A étant inversible s'écrit $A = P_1 \cdots P_r D_n(\lambda) Q_1 \cdots Q_s$, où les matrices P_i et Q_j sont des matrices de transvection et la matrice $D_n(\lambda)$ une matrice de dilatation (théorème 5.2). Une utilisation répétée du corollaire précédent nous donne :

$$\det(A) = \det(D_n(\lambda)) = \lambda$$

et pour toute matrice B :

$$\det(AB) = \det(D_n(\lambda)) \det(B) = \det(A)\det(B).$$

■

Le résultat précédent est en fait valable pour toutes matrices A et B. Le cas où la matrice A n'est pas inversible se traite en utilisant le résultat suivant.

Théorème 5.8 *Une matrice* $A \in \mathcal{M}_n(\mathbb{R})$ *est inversible si, et seulement si, son déterminant est non nul et dans ce cas, on a* $\det(A^{-1}) = \dfrac{1}{\det(A)}$.

Démonstration. Si A est inversible d'inverse A^{-1}, on $AA^{-1} = I_n$ et le théorème précédent nous dit que $\det(A)\det(A^{-1}) = \det(I_n) = 1$, donc $\det(A) \neq 0$ et $\det(A^{-1}) = \dfrac{1}{\det(A)}$.

La réciproque se démontre par récurrence sur $n \geq 1$.

Pour $n = 1$, le résultat est évident car $\det(a) = a$ pour tout réel a.

Supposons le résultat acquis pour les matrices d'ordre $n-1 \geq 1$ et soit A d'ordre n telle que $\det(A) \neq 0$. La première colonne de A est nécessairement non nulle (définition du déterminant) et on peut reprendre la démonstration du théorème 5.2 pour trouver des matrices de transvection $P_1, \cdots, P_k$ telles que :

$$P_k \cdots P_1 A = \begin{pmatrix} 1 & \alpha_{12} & \cdots & \alpha_{1n} \\ 0 & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \alpha_{n2} & \cdots & \alpha_{nn} \end{pmatrix} = \begin{pmatrix} 1 & \alpha \\ 0 & B \end{pmatrix}$$

où α est un vecteur ligne à $n-1$ composantes et B une matrice carrée d'ordre $n-1$.

Comme les matrices P_k sont inversibles, on a :

$$\det(A) = \det(P_k \cdots P_1 A) = \det(B)$$

et $\det(B) \neq 0$. La matrice B est donc inversible, ce qui implique que A est aussi inversible. En effet si $Ax = 0$, en notant $x = \begin{pmatrix} x_1 \\ x' \end{pmatrix}$ avec $x_1 \in \mathbb{R}$ et $x' \in \mathbb{R}^{n-1}$, on a :

$$\begin{cases} x_1 + \alpha x' = 0 \\ Bx' = 0 \end{cases}$$

ce qui entraîne $x' = 0$ et $x_1 = 0$, soit $x = 0$. La matrice A est donc inversible. ∎

Théorème 5.9 *Pour toutes matrices A et B dans $\mathcal{M}_n(\mathbb{R})$, on a :*

$$\det(AB) = \det(BA) = \det(A)\det(B).$$

Démonstration. Il reste à traiter le cas où la matrice A n'est pas inversible. Dans ce cas la matrice AB ne peut être inversible (sinon, en notant C l'inverse de AB, on a $(AB)C = I_n$, soit $A(BC) = I_n$ et A est inversible) et on a :

$$0 = \det(AB) = \det(A)\det(B) = 0 \cdot \det(B).$$

L'égalité $\det(BA) = \det(B)\det(A)$ donne $\det(AB) = \det(BA)$. ∎

On peut remarquer que $\det(AB) = \det(BA)$ alors qu'en général $AB \neq BA$.

La multiplication à droite par une matrice élémentaire se traduisant par une action particulière sur les colonnes, on déduit de ce théorème et du théorème 5.1 les propriétés suivantes du déterminant.

Corollaire 5.6 *Si $A'_j(\lambda)$ est la matrice déduite de $A \in \mathcal{M}_n(\mathbb{R})$ en multipliant sa colonne j par un réel λ, on a alors* $\det\left(A'_j(\lambda)\right) = \lambda \det(A)$.
Si $A \in \mathcal{M}_n(\mathbb{R})$ a une colonne nulle, alors $\det(A) = 0$.

Pour l'instant, le déterminant d'une matrice se calcule en utilisant la première colonne de cette dernière. En réalité, on peut aussi utiliser la première ligne et nous en déduirons que cette première ligne ou colonne peut être remplacée par n'importe quelle autre. Précisément, on a les résultats suivants.

Théorème 5.10 *Pour toute matrice A dans $\mathcal{M}_n(\mathbb{R})$, on a* $\det({}^tA) = \det(A)$.

Démonstration. Comme d'habitude c'est trivial pour $n = 1$. On suppose donc que $n \geq 2$.

Si A n'est pas inversible, il en est de même de sa transposée (en effet si ${}^tA$ est inversible, il en est de même de $A = {}^t({}^tA)$ – théorème 3.19 –) et on a alors $\det({}^tA) = \det(A) = 0$.

Si A est inversible, elle s'écrit :

$$A = P_1 \cdots P_r D_n(\lambda) Q_1 \cdots Q_s$$

où les P_i, Q_j sont des matrices de transvection et $\lambda = \det(A)$, ce qui donne :

$${}^tA = {}^tQ_s \cdots {}^tQ_1 \, {}^tD_n(\lambda) \, {}^tP_r \cdots {}^tP_1$$

les transposées de matrices élémentaires étant des matrices élémentaires de même type avec ${}^tD_n(\lambda) = D_n(\lambda)$, ce qui donne :

$$\det\left({}^tA\right) = \det(D_n(\lambda)) = \lambda = \det(A).$$

■

De ce résultat, on déduit le développement du déterminant suivant la première ligne (pour $n \geq 2$) :

$$\det(A) = \det\left({}^tA\right) = \sum_{j=1}^{n} (-1)^{j+1} a_{1,j} \det(A_{1,j})$$

où $A_{1,j}$ est la matrice carrée d'ordre $n-1$ déduite de A en supprimant la ligne 1 et la colonne j.

On en déduit alors les propriétés suivantes relatives aux colonnes, ces propriétés étant analogues à celles obtenues pour les lignes.

Corollaire 5.7 *Si A' est la matrice déduite de $A \in \mathcal{M}_n(\mathbb{R})$ en permutant deux colonnes, on a alors* $\det(A') = -\det(A)$.
Soient A, A', A'' des matrices de lignes respectives C_j, C'_j, C''_j (pour j compris entre 1 et n) telles que $C_j = C'_j = C''_j$ pour $j \neq k$ et $C''_k = C_k + C'_k$ où k est un indice compris entre 1 et n. On a :

$$\det(A'') = \det(A) + \det(A').$$

On ne change pas la valeur d'un déterminant si on ajoute à une ligne une combinaison linéaire des autres lignes.

Ce corollaire se traduit en disant que le déterminant est linéaire par rapport à chaque colonne et que c'est une forme alternée sur les colonnes.

En général, pour calculer un déterminant, on essayera d'effectuer des opérations élémentaires sur les lignes ou les colonnes dans le but de faire apparaître un maximum de coefficients nuls, ce qui facilitera le calcul du déterminant de la matrice obtenue.

De tout ce qui précède, on déduit les différentes formes de développement d'un déterminant suivant une ligne ou une colonne (pour $n \geq 2$).

Théorème 5.11 *Pour toute matrice $A \in \mathcal{M}_n(\mathbb{R})$, on a :*

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{i,j} \det(A_{i,j}) \ \ (1 \leq j \leq n)$$

(développement suivant la colonne j) et

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{i,j} \det(A_{i,j}) \ (1 \leq i \leq n)$$

(développement suivant la ligne i) où A_{ij} est la matrice carrée d'ordre $n-1$ déduite de A en supprimant la ligne i et la colonne j.

Démonstration. Pour $j = 1$, c'est la définition première du déterminant et pour $i = 1$ c'est une conséquence immédiate de $\det({}^tA) = \det(A)$.

Fixons la colonne $j \geq 2$ et notons $(e_i)_{1 \leq i \leq n}$ la base canonique de $\mathbb{R}^n$.

La colonne C_j s'écrit $C_j = \sum_{i=1}^{n} a_{ij} e_i$ et en utilisant la linéarité du déterminant par rapport à la j-ième colonne, on a :

$$\det(A) = \sum_{i=1}^{n} a_{ij} \det(B_{i,j})$$

où B_{ij} est la matrice déduite de A en remplaçant C_j par e_i. En permutant la colonne j avec la colonne $j-1$, puis $j-1$ avec $j-2$, $\cdots$, 2 avec 1 et ensuite la ligne i avec la ligne $i-1$, $i-1$ avec $i-2$, $\cdots$, 2 avec 1 (on fait rien pour $i = 1$) on aboutit à :

$$\begin{aligned}\det(B_{i,j}) &= (-1)^{i+j} \begin{vmatrix} 1 & a_{11} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1n} \\ 0 & a_{21} & \cdots & a_{2,j-1} & a_{2,j+1} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ 0 & a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{vmatrix} \\ &= (-1)^{i+j} \det(A_{i,j})\end{aligned}$$

et on a le résultat annoncé.

On procède de manière analogue pour la deuxième formule. ∎

Avec les notations du théorème, on dit que $\det(A_{i,j})$ est le mineur d'indice (i,j) de la matrice A et que $(-1)^{i+j} \det(A_{i,j})$ est le cofacteur d'indice (i,j) de A.

Exercice 5.8 *Soient $n \geq 2$ un entier et $\alpha_1, \alpha_2, \cdots, \alpha_n$ des réels.*

1. *Calculer le déterminant $\Delta(\alpha_1, \cdots, \alpha_n)$ de la matrice :*

$$V(\alpha_1, \cdots, \alpha_n) = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{n-1} & \alpha_2^{n-1} & \cdots & \alpha_n^{n-1} \end{pmatrix}$$

   Une telle matrice est dite de Vandermonde.

2. *À quelle condition une telle matrice est-elle inversible ?*

Solution 5.8 *Pour $n = 2$, on a $\Delta(\alpha_1, \alpha_2) = \alpha_2 - \alpha_1$ et pour $n = 3$, on a fait le calcul avec l'exercice 5.2.*

1. *Le calcul de $\Delta(\alpha_1,\cdots,\alpha_n)$ se fait par récurrence sur $n\geq 2$.*
En retranchant, pour $i=n,n-1,\cdots,2$ à la ligne i la ligne $i-1$ multipliée par α_1, on obtient :

$$\Delta(\alpha_1,\cdots,\alpha_n)=\begin{vmatrix}1 & 1 & \cdots & 1\\ 0 & \alpha_2-\alpha_1 & \cdots & \alpha_n-\alpha_1\\ \vdots & \vdots & \ddots & \vdots\\ 0 & \alpha_2^{n-2}(\alpha_2-\alpha_1) & \cdots & \alpha_n^{n-2}(\alpha_n-\alpha_1)\end{vmatrix}$$
$$=\begin{vmatrix}\alpha_2-\alpha_1 & \alpha_3-\alpha_1 & \cdots & \alpha_n-\alpha_1\\ \alpha_2(\alpha_2-\alpha_1) & \alpha_3(\alpha_3-\alpha_1) & \cdots & \alpha_n(\alpha_n-\alpha_1)\\ \vdots & \vdots & \ddots & \vdots\\ \alpha_2^{n-2}(\alpha_2-\alpha_1) & \alpha_3^{n-2}(\alpha_3-\alpha_1) & \cdots & \alpha_n^{n-2}(\alpha_n-\alpha_1)\end{vmatrix}$$

soit :

$$\Delta(\alpha_1,\cdots,\alpha_n)=\left(\prod_{k=2}^{n}(\alpha_k-\alpha_1)\right)\begin{vmatrix}1 & \cdots & 1\\ \alpha_2 & \cdots & \alpha_n\\ \vdots & \ddots & \vdots\\ \alpha_2^{n-2} & \cdots & \alpha_n^{n-2}\end{vmatrix}$$
$$=\left(\prod_{k=2}^{n}(\alpha_k-\alpha_1)\right)\Delta(\alpha_2,\cdots,\alpha_n)$$

et par récurrence :

$$\det(A_n)=\prod_{k=2}^{n}(\alpha_k-\alpha_1)\prod_{2\leq i<j\leq n}^{n}(\alpha_j-\alpha_i)$$
$$=\prod_{1\leq i<j\leq n}^{n}(\alpha_j-\alpha_i).$$

2. *Cette matrice est inversible si, et seulement si, les α_i sont deux à deux distincts.*

Exercice 5.9 *Calculer le déterminant de la matrice :*

$$A_n=\begin{pmatrix}1 & 1 & \cdots & 1\\ 2 & 2^2 & \cdots & 2^n\\ \vdots & \vdots & \ddots & \vdots\\ n & n^2 & \cdots & n^n\end{pmatrix}.$$

Solution 5.9 *On a :*

$$\det(A_n)=n!\begin{vmatrix}1 & 1 & \cdots & 1\\ 1 & 2 & \cdots & 2^{n-1}\\ \vdots & \vdots & \ddots & \vdots\\ 1 & n & \cdots & n^{n-1}\end{vmatrix}=n!\Delta(1,2,\cdots,n)$$
$$=n!\prod_{1\leq j<i\leq n}(i-j)=n!\prod_{i=2}^{n}(i-1)!=\prod_{i=2}^{n}i!.$$

Exercice 5.10 *Soit*

$$A_n = \begin{pmatrix} a_1 & c_1 & 0 & \cdots & 0 \\ b_2 & a_2 & c_2 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & b_{n-1} & a_{n-1} & c_{n-1} \\ 0 & \cdots & 0 & b_n & a_n \end{pmatrix},$$

une matrice tridiagonale d'ordre $n \geq 3$ à coefficients réels.
Pour tout entier k compris entre 1 et n, on désigne par D_k le déterminant de la matrice d'ordre k formée des k premières lignes et k premières colonnes de A_n (les D_k sont les déterminants extraits principaux de A_n).

1. *Exprimer, pour tout k compris entre 3 et n, D_k en fonction de D_{k-1}et D_{k-2}.*
2. *Calculer le déterminant de :*

$$A_n = \begin{pmatrix} 2 & 1 & 0 & \cdots & 0 \\ 2^2 & 5 & 1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & (n-1)^2 & 2n-1 & 1 \\ 0 & \cdots & 0 & n^2 & 2n+1 \end{pmatrix}$$

Solution 5.10

1. *En développant D_k suivant la dernière ligne on a :*

$$D_k = a_k \begin{vmatrix} a_1 & c_1 & 0 & \cdots & 0 \\ b_2 & a_2 & c_2 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & b_{k-2} & a_{k-2} & c_{k-2} \\ 0 & \cdots & 0 & b_{k-1} & a_{k-1} \end{vmatrix} - b_k \begin{vmatrix} a_1 & c_1 & 0 & \cdots & 0 \\ b_2 & a_2 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & c_{k-3} & 0 \\ \vdots & \ddots & b_{k-2} & a_{k-2} & 0 \\ 0 & \cdots & 0 & b_{k-1} & c_{k-1} \end{vmatrix}$$
$$= a_k D_{k-1} - b_k c_{k-1} D_{k-2}.$$

   Ce qui donne, avec les valeurs initiales $D_1 = a_1$ et $D_2 = a_1 a_2 - b_2 c_1$, un algorithme de calcul de D_n.

2. *On a :*

$$D_n = (2n+1)\, D_{n-1} - n^2 D_{n-2}$$

   avec les valeurs initiales $D_1 = 2$, $D_2 = 6$. En calculant D_3 et D_4 on conjecture que $D_n = (n+1)!$ Ce qui se montre par récurrence sur $n \geq 2$. C'est vrai pour $n = 2$ et le supposant acquis jusqu'au rang $n - 1 \geq 2$, on a :

$$D_n = (2n+1)\, n! - n^2\, (n-1)! = (n+1)!$$

5.3 Déterminant d'une famille de vecteurs

Étant donnée une famille $(x_j)_{1 \leq j \leq n}$ de n vecteurs de $\mathbb{R}^n$, on définit le déterminant de cette famille comme le déterminant de la matrice A dont les colonnes sont formées de ces vecteurs. En notant, pour j compris entre 1 et n, $x_j = (x_{i,j})_{1 \leq i \leq n}$ (vecteur colonne), on a donc :

$$\det(x_1, \cdots, x_n) = \det\left((x_{ij})\right)_{1 \leq i,j \leq n}.$$

Du théorème 5.8 on déduit le résultat suivant bien utile pour vérifier qu'un système de n vecteurs dans $\mathbb{R}^n$ est libre et donc forme une base.

Théorème 5.12 *Une famille $(x_j)_{1\leq j\leq n}$ de n vecteurs de $\mathbb{R}^n$ est libre si, et seulement si, son déterminant est non nul.*

Démonstration. En utilisant les notations qui précèdent, on note $P = ((x_{ij}))_{1\leq i,j\leq n}$.

Dire que le système $(x_j)_{1\leq j\leq n}$ est libre équivaut à dire que l'unique solution $\lambda \in$dans $\mathbb{R}^n$ du système linéaire $\sum_{j=1}^{n} \lambda_j x_j$ est $\lambda = 0$, ce système s'écrivant aussi $P\lambda = 0$, cela revient à dire que la matrice P est inversible, ce qui est encore équivalent à $\det(P) \neq 0$. ■

5.4 Déterminant d'un endomorphisme

On désigne par E un espace vectoriel réel de dimension $n \geq 1$.

Si $\mathcal{B}_1$ et $\mathcal{B}_2$ sont deux bases de E, u un endomorphisme de E, A_1 la matrice de u dans la base $\mathcal{B}_1$ et A_2 sa matrice dans la base $\mathcal{B}_2$, on sait alors que $A_2 = P^{-1}AP$ où P est la matrice de passage de $\mathcal{B}_1$ à $\mathcal{B}_2$. Il en résulte alors que :

$$\begin{aligned}\det(A_2) &= \det\left(P^{-1}AP\right) = \det\left(P^{-1}\right)\det(A_1)\det(P)\\ &= \frac{1}{\det(P)}\det(A_1)\det(P) = \det(A_1).\end{aligned}$$

C'est-à-dire que ces déterminants ne dépendent pas du choix d'une base.

On peut alors donner la définition suivante.

Définition 5.6 *Le déterminant d'un endomorphisme u de E est le déterminant de la matrice de u dans une base de E.*

Du théorème 4.19, on déduit le résultat suivant.

Théorème 5.13 *Si u, v sont deux endomorphismes de E, alors :*

$$\det(u\circ v) = \det(v\circ u) = \det(u)\det(v).$$

Et du théorème 4.20, on déduit le résultat suivant.

Théorème 5.14 *Un endomorphisme u de E est inversible si, et seulement si, son déterminant est non nul.*