M102 - Algèbre et géométrie

Franck Benoist

L1S1 - Université Paris-Saclay

Table des matières

1 Logique et raisonnements **5**
- 1.1 Énoncés . . . 5
- 1.2 Connecteurs logiques . . . 5
- 1.3 Tautologies . . . 8
- 1.4 Quelques types de raisonnement . . . 9

2 Ensembles et applications **13**
- 2.1 Ensembles . . . 13
- 2.2 Applications . . . 15
- 2.3 Cas des ensembles finis . . . 18

3 Relations sur un ensemble **23**
- 3.1 Généralités . . . 23
- 3.2 Relations d'équivalence . . . 23
- 3.3 Relations d'ordre . . . 25

4 Arithmétique **27**
- 4.1 Divisibilité . . . 27
- 4.2 Division euclidienne . . . 28
- 4.3 Algorithme d'Euclide . . . 28
- 4.4 Décomposition en facteurs premiers . . . 30
- 4.5 Congruences et $\mathbb{Z}/d\mathbb{Z}$ . . . 32

5 Nombres complexes **35**
- 5.1 Définitions et premières propriétés . . . 35
- 5.2 Interprétation géométrique . . . 37
- 5.3 Quelques équations . . . 38

6 Polynômes **41**
- 6.1 Définitions et premières propriétés . . . 41
- 6.2 Racines d'un polynôme . . . 43
- 6.3 Factorisation des polynômes . . . 45

7 Géométrie dans le plan et dans l'espace **49**
- 7.1 Géométrie dans le plan . . . 49
  - 7.1.1 Coordonnées . . . 49
  - 7.1.2 Droites . . . 49
  - 7.1.3 Distance . . . 50
- 7.2 Géométrie dans l'espace . . . 51
  - 7.2.1 Coordonnées . . . 51
  - 7.2.2 Plan dans l'espace . . . 51
  - 7.2.3 Droite dans l'espace . . . 52
  - 7.2.4 Distance . . . 52

Chapitre 1

Logique et raisonnements

Objectif : donner quelques règles logiques pour des raisonnements mathématiques rigoureux

1.1 Énoncés

On va considérer dans la suite des énoncés, qui peuvent être vrais ou faux. On utilisera aussi souvent les *valeurs de vérité* 0 (pour faux) et 1 (pour vrai).

Exemple 1.1

1. $P : 3 \leq 5$. *P est vrai.*
2. $Q : 5$ *est pair. Q est faux.*
3. R : *Le ciel est bleu.*

Parfois, les énoncés dépendent de variables (on notera $P(x)$ pour un énoncé qui dépend d'une variable x). Il faudra alors donner des valeurs (pas forcément numériques) à ces variables avant de pouvoir dire si ces énoncés sont vrais ou faux.

Exemple 1.2

1. $P(n)$: *n est un nombre premier.*
   $P(3)$ est vrai et $P(4)$ est faux.
2. $Q(X, Y)$: *X est un ensemble contenu dans l'ensemble Y.*
   $Q(\mathbb{N}, \mathbb{R})$ est vrai.

On appelle théorème un énoncé mathématique dont on a pu montrer qu'il est vrai. On utilise aussi parfois d'autres termes, selon l'importance qu'on accorde à ces théorèmes : proposition, propriété (pour des théorèmes de moindre importance), lemme (un résultat préliminaire pour démontrer un théorème), corollaire (une conséquence facile d'un théorème).

1.2 Connecteurs logiques

On peut utiliser différents connecteurs logiques pour construire des énoncés à partir d'autres énoncés.

Négation (non,$\neg$)
Si P est un énoncé, on peut considérer sa négation non P (notée aussi parfois $\neg P$). L'énoncé non P est vrai exactement quand P est faux. On utilise souvent des tables de vérité pour résumer ces propriétés :

P	non P
0	1
1	0

Exemple 1.3

1. *P : 5 est pair ; non P : 5 est impair.*
   P est faux ; non P est vrai.
2. *$Q(X,Y)$: X est contenu dans Y ; non $Q(X,Y)$: X n'est pas contenu dans Y (attention, ce n'est pas : Y est contenu X)*
3. *R : le chat est noir ; non R : le chat n'est pas noir.*

Conjonction (et, $\wedge$)
Si P et Q sont des énoncés, on peut construire l'énoncé P et Q (noté aussi parfois $P \wedge Q$). L'énoncé P et Q est vrai exactement quand P et Q sont tous les deux vrais :

P	Q	P et Q
0	0	0
1	0	0
0	1	0
1	1	1

Exemple 1.4
$P(n)$: n est un nombre impair et n est un nombre premier.
$P(2)$ est faux (car 2 est pair), $P(9)$ est faux (car 9 n'est pas premier), $P(3)$ est vrai.

On utilise aussi parfois une accolade, en particulier quand on écrit des systèmes d'équations :
$\begin{cases} x+y=3 \\ x-y=1 \end{cases}$ signifie $x+y=3$ et $x-y=1$.
Quand on utilise plusieurs connecteurs logiques et qu'il y a risque de confusion, on utilise des parenthèses.

Exemple 1.5

1. | *P* | *Q* | *non (P et Q)* | *(non P) et Q* |
   |---|---|---|---|
   | *0* | *0* | *1* | *0* |
   | *1* | *0* | *1* | *0* |
   | *0* | *1* | *1* | *1* |
   | *1* | *1* | *0* | *0* |

   Les énoncés non (P et Q) d'une part, (non P) et Q d'autre part, n'ont pas les mêmes valeurs de vérité, on évite donc d'écrire non P et Q, qui serait ambigu.
2. *Les énoncés (P et Q) et R d'une part, P et (Q et R) d'autre part, ont mêmes valeur de vérité (on le constate en écrivant les tables de vérité), on peut donc se permettre d'écrire P et Q et R.*

Disjonction (ou, $\vee$)
Si P et Q sont des énoncés, on peut construire l'énoncé P ou Q (noté aussi parfois $P \vee Q$). L'énoncé P ou Q est vrai exactement quand P est vrai ou Q est vrai (éventuellement les deux) :

P	Q	P ou Q
0	0	0
1	0	1
0	1	1
1	1	1

Si on veut exprimer le "ou exclusif" (P est vrai ou Q est vrai mais pas les deux), on peut utiliser les connecteurs précédents et écrire (P ou Q) et non (P et Q) :

P	Q	(P ou Q) et non (P et Q)
0	0	0
1	0	1
0	1	1
1	1	0

Implication ($\Rightarrow$, si ... alors ...)
Si P et Q sont des énoncés, on peut construire l'énoncé $P \Rightarrow Q$ (si P alors Q). L'énoncé $P \Rightarrow Q$ est vrai exactement quand P est faux ou Q est vrai :

P	Q	P $\Rightarrow$ Q
0	0	1
1	0	0
0	1	1
1	1	1

Dans l'implication $P \Rightarrow Q$, P s'appelle la prémisse et Q la conclusion.

Exemple 1.6 $P(n) : n \geq 4 \Rightarrow n \geq 2$. *(Si $n \geq 4$, alors $n \geq 2$.)*
L'énoncé $P(n)$ est vrai pour tout entier n (en utilisant les propriétés de l'ordre et le fait que $4 \geq 2$). On peut voir les différentes situations où une implication est vraie pour différentes valeurs de n :
$P(1)$ est vrai : la prémisse est fausse et la conclusion est fausse.
$P(3)$ est vrai : la prémisse est fausse et la conclusion est vraie.
$P(5)$ est vrai : la prémisse est vraie et la conclusion est vraie.

Exemple 1.7 *"Si je me couche tôt ce soir, alors le jour se lèvera demain". C'est vrai car la conclusion est vraie (il n'est même pas nécessaire de savoir si la prémisse est vraie ou fausse). L'implication au sens logique ne correspond donc pas à la recherche de cause.*

Une manière peut-être plus intuitive de justifier la table de vérité d'une implication est de considérer sa négation : $P \Rightarrow Q$ est faux exactement quand P est vrai mais Q est faux.

Exemple 1.8 *Si je joue au loto, alors je gagne au loto. C'est faux car il est possible que la prémisse soit vraie (je joue au loto) mais que la conclusion soit fausse (je ne gagne pas au loto).*

Attention : pour une implication $P \Rightarrow Q$, ne pas confondre sa négation non $(P \Rightarrow Q)$, l'implication réciproque $Q \Rightarrow P$, et l'implication contraposée (non Q) $\Rightarrow$ (non P) (que l'on verra plus tard).

P	Q	P $\Rightarrow$ Q	non (P $\Rightarrow$ Q)	Q$\Rightarrow$ P	(non Q) $\Rightarrow$ (non P)
0	0	1	0	1	1
1	0	0	1	1	0
0	1	1	0	0	1
1	1	1	0	1	1

Exemple 1.9 *L'implication "si j'ai gagné au loto, alors j'ai joué au loto" est toujours vraie.*
Sa négation "j'ai gagné au loto et je n'ai pas joué" est toujours fausse.
L'implication réciproque "si j'ai joué au loto, alors j'ai gagné au loto" peut être fausse (voir plus haut).
L'implication contraposée "si je n'ai pas joué au loto, alors je n'ai pas gagné au loto" est toujours vraie.

Équivalence ($\Leftrightarrow$, équivalent à, si et seulement si)
Si P et Q sont des énoncés, on peut construire l'énoncé $P \Leftrightarrow Q$ (P est équivalent à Q, P si et seulement si Q, qu'on écrit souvent P ssi Q). L'énoncé $P \Leftrightarrow Q$ est vrai exactement quand P et Q ont la même valeur de vérité (tous les deux vrais ou tous les deux faux), ou encore quand $P \Rightarrow Q$ et $Q \Rightarrow P$ sont vrais :

P	Q	P $\Leftrightarrow$ Q
0	0	1
1	0	0
0	1	0
1	1	1

Exemple 1.10

1. *x est supérieur ou égal à 4 ssi x n'est pas strictement inférieur à 4.*

2. L'eau gèle si et seulement si sa température est portée à $0°C$ ou moins.

Quantificateur universel ($\forall$, pour tout, quelque soit)
Si $P(x)$ est un énoncé qui dépend d'une variable x (et peut-être d'autres), on peut construire l'énoncé $\forall x\ P(x)$ (pour tout x, $P(x)$; quelque soit x, $P(x)$). L'énoncé $\forall x\ P(x)$ est vrai exactement quand $P(x)$ est vrai pour toutes les valeurs de x. L'énoncé $\forall x\ P(x)$ ne dépend plus de la variable x, on n'a pas besoin de donner une valeur à x pour savoir si $\forall x\ P(x)$ est vrai ou faux.

Exemple 1.11 *1.* $\forall x\ x^2 \geq 0$

2. $\forall a\ \forall b\ (a+b)^2 = a^2 + 2ab + b^2$

Souvent, on précise dans quel ensemble peut varier la variable x. On écrit $\forall x \in X\ P(x)$ pour dire $\forall x\ (x \in X \Rightarrow P(x))$. On peut aussi imposer des conditions sur x, par exemple $\forall x \geq 2\ x^2 \geq 4$ signifie $\forall x\ (x \geq 2 \Rightarrow x^2 \geq 4)$.

Quantificateur existentiel ($\exists$, il existe ... tel que)
Si $P(x)$ est un énoncé qui dépend d'une variable x (et peut-être d'autres), on peut construire l'énoncé $\exists x\ P(x)$ (il existe x tel que $P(x)$). L'énoncé $\exists x\ P(x)$ est vrai exactement quand il existe au moins une valeur de x telle que $P(x)$ soit vrai. L'énoncé $\exists x\ P(x)$ ne dépend plus de la variable x, on n'a pas besoin de donner une valeur à x pour savoir si $\exists x\ P(x)$ est vrai ou faux.
On utilisera aussi la variante suivante : $\exists! x\ P(x)$ (il existe un unique x tel que $P(x)$). Cet énoncé est vrai quand il existe exactement une valeur de x telle que $P(x)$ soit vrai.
Comme pour le quantificateur universel, on peut préciser un ensemble auquel appartient la variable x ou une condition qu'elle doit satisfaire : $\exists x \in X\ P(x)$ signifie $\exists x\ (x \in X$ et $P(x))$.

Exemple 1.12 $\exists x \in \mathbb{C}\ x^2 + 1 = 0$

Attention : quand on utilise plusieurs quantificateurs différents, l'ordre dans lequel on les écrit est important. Par exemple, $\forall x \in \mathbb{R}\ \exists y \in \mathbb{R}\ y < x$ est vrai : pour tout réel x, on peut trouver un réel y (qui dépend de x) tel que $y < x$, par exemple $y = x - 1$. Mais $\exists y \in \mathbb{R}\ \forall x \in \mathbb{R}\ y < x$ est faux : on ne peut pas trouver de réel y (choisi une fois pour toute) qui soit strictement plus petit que tous les réels x.

1.3 Tautologies

On dit qu'un énoncé est une tautologie s'il est toujours vrai, indépendamment des valeurs de vérité des énoncés qui le composent.

Exemple 1.13 *1. P ou (non P) est une tautologie*

2. P ou Q n'est pas une tautologie : cet énoncé est faux quand P et Q le sont

3. $((P \Rightarrow Q)$ et $(Q \Rightarrow R)) \Rightarrow (P \Rightarrow R)$ est une tautologie

On peut souvent vérifier qu'un énoncé est une tautologie en écrivant sa table de vérité :

P	Q	R	P ⇒ Q	Q ⇒ R	(P⇒ Q) et (Q⇒ R)	P ⇒ R	((P⇒ Q) et (Q⇒ R))⇒ (P⇒ R)
0	0	0	1	1	1	1	1
1	0	0	0	1	0	0	1
0	1	0	1	0	0	1	1
1	1	0	1	0	0	0	1
0	0	1	1	1	1	1	1
1	0	1	0	1	0	1	1
0	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1

Pour cet exemple, on peut aussi faire le raisonnement suivant. Pour montrer qu'une implication est vraie, on suppose que la prémisse est vraie, et on cherche à montrer que la conclusion est vraie (c'est une conséquence de la table de vérité de l'implication : si la prémisse est fausse, on sait déjà que l'implication est

vraie). Supposons que $(P \Rightarrow Q)$ et $(Q \Rightarrow R)$ est vrai, on veut montrer que $P \Rightarrow R$ est vrai. On suppose donc que P est vrai. Comme $P \Rightarrow Q$ est vrai, Q est vrai. Comme $Q \Rightarrow R$ est vrai, R est vrai. Ainsi, $P \Rightarrow R$ est vrai, et on conclut que $((P \Rightarrow Q)$ et $(Q \Rightarrow R)) \Rightarrow (P \Rightarrow R)$ est une tautologie.

Voici quelques autres tautologies :

1. $(P$ et $Q) \Rightarrow P$; $(P$ et $Q) \Rightarrow Q$
2. $P \Rightarrow (P$ ou $Q)$; $Q \Rightarrow (P$ ou $Q)$
3. $((\forall x\ P(x))$ ou $(\forall x\ Q(x))) \Rightarrow (\forall x\ (P(x)$ ou $Q(x)))$ (attention, l'implication réciproque n'est pas toujours vraie)
4. $(\exists x\ (P(x)$ et $Q(x)) \Rightarrow ((\exists x\ P(x))$ et $(\exists x\ Q(x)))$ (attention, l'implication réciproque n'est pas toujours vraie)
5. $((P \Leftrightarrow Q)$ et $(Q \Leftrightarrow R)) \Rightarrow (P \Leftrightarrow R)$
6. $P \Rightarrow (Q \Rightarrow P)$

Quand un énoncé de la forme $P \Leftrightarrow Q$ est une tautologie (pour des énoncés particuliers P et Q), on dit que P et Q sont équivalents entre eux. Il existe un grand nombre d'équivalences (en fait une infinité), voici une liste de celles qu'on utilise fréquemment dans les raisonnements mathématiques :

1. P est équivalent à non non P
2. Lois de Morgan : non $(P$ et $Q)$ est équivalent à (non P) ou (non Q) ; non $(P$ ou $Q)$ est équivalent à (non P) et (non Q)
3. Distributivité : $(P$ et $Q)$ ou R est équivalent à $(P$ ou $R)$ et $(Q$ ou $R)$; $(P$ ou $Q)$ et R est équivalent à $(P$ et $R)$ ou $(Q$ et $R)$
4. non $(\exists x\ P(x))$ est équivalent à $\forall x$ (non $P(x)$)
5. non $(\forall x\ P(x))$ est équivalent à $\exists x$ (non $P(x)$)
6. Contraposée : $P \Rightarrow Q$ est équivalent à (non Q) $\Rightarrow$ (non P)
7. $\forall x\ (P(x)$ et $Q(x))$ est équivalent à $(\forall x\ P(x))$ et $(\forall x\ Q(x))$ (attention, ça devient faux si on remplace et par ou)
8. $\exists x\ (P(x)$ ou $Q(x))$ est équivalent à $(\exists x\ P(x))$ ou $(\exists x\ Q(x))$ (attention, ça devient faux si on remplace ou par et)

1.4 Quelques types de raisonnement

On peut déduire des tautologies précédentes quelques types de raisonnement souvent utilisés.

Modus ponens
Si P est vrai et $P \Rightarrow Q$ est vrai, alors Q est vrai. Par exemple, pour montrer un théorème T, on commence par montrer un lemme L, puis on montre que le lemme implique le théorème $(L \Rightarrow T)$.

Raisonnement par contraposée
On a vu que $P \Rightarrow Q$ est équivalent à (non Q) $\Rightarrow$ (non P). Il est parfois plus facile, plutôt que de montrer directement $P \Rightarrow Q$ (on suppose que P est vrai et on cherche à prouver que Q est vrai), de montrer la contraposée (non Q) $\Rightarrow$ (non P) (on suppose que Q est faux et on cherche à prouver que P est faux).

Exemple 1.14 *Soit n un entier. Montrons que si n^2 est pair, alors n est pair. Cela revient à montrer la contraposée de cette implication : si n est impair, alors n^2 est impair. Supposons que n est impair, on peut l'écrire $n = 2k+1$ pour un certain entier k. On calcule alors $n^2 = (2k+1)^2 = 4k^2+4k+1 = 2(2k^2+2k)+1$: c'est un nombre impair.*

Raisonnement par l'absurde
C'est une variante du raisonnement par contraposée. Pour montrer qu'un énoncé P est vrai, on suppose que P est faux et on cherche à aboutir à un énoncé F dont on sait qu'il est faux. Si on y parvient, on a montré l'implication (non P) $\Rightarrow F$, et donc, par contraposée, l'implication (non F) $\Rightarrow$ (non non P) (et

donc (non F) $\Rightarrow P$ car non non P est équivalent à P). Comme F est faux, non F est vrai, et donc P aussi.

Exemple 1.15 *Soient a et b deux nombres réels positifs tels que $\frac{a}{1+b} = \frac{b}{1+a}$. Montrer que $a = b$.*
On suppose par l'absurde que $a \neq b$. Comme $\frac{a}{1+b} = \frac{b}{1+a}$, on obtient $a(1+a) = b(1+b)$, ou encore $a^2 - b^2 = b - a$, puis en factorisant $(a-b)(a+b) = -(a-b)$. Comme on a supposé que $a \neq b$, on peut diviser par $a - b$ (qui est non-nul), et on obtient $a + b = -1$: c'est faux car la somme de deux nombres positifs ne peut pas valoir -1. On a donc monté que $a = b$.

Raisonnement par contre-exemple
Comme non $(\forall x\ P(x))$ est équivalent à $\exists x$ (non $P(x)$), pour montrer que $\forall x\ P(x)$ est faux, il suffit de trouver un contre-exemple, c'est-à-dire trouver x tel que $P(x)$ soit faux.

Exemple 1.16 *Montrons que l'énoncé suivant est faux : pour tout entier n, si n est divisible par 6 et par 4, alors n est divisible par 24.*
On cherche une valeur particulière de n telle que cette implication soit fausse, c'est-à-dire telle que la prémisse soit vraie mais la conclusion fausse. On peut prendre $n = 12$: $12 = 6 \times 2 = 4 \times 3$ est divisible par 6 et par 4, mais il n'est pas divisible par 24.

Raisonnement par équivalence
D'après la tautologie $((P \Leftrightarrow Q) \text{ et } (Q \Leftrightarrow R)) \Rightarrow (P \Leftrightarrow R)$, si on montre une suite d'équivalences $P_1 \Leftrightarrow P_2 \Leftrightarrow \ldots \Leftrightarrow P_n$, alors on a obtenu $P_1 \Leftrightarrow P_n$. C'est ce qu'on utilise en particulier quand on cherche à résoudre des équations. Par exemple, $2x + 3 = 7 \Leftrightarrow 2x = 4 \Leftrightarrow x = 2$: la seule solution de l'équation $2x + 3 = 7$ est $x = 2$. Dans ce type de raisonnement, il faut prendre garde à bien écrire des équivalences. Par exemple, dans ce qui précède, l'équation 1 implique l'équation 2 (en retranchant 3) et l'équation 2 implique l'équation 1 (en ajoutant 3), l'équation 2 implique l'équation 3 (en divisant par 2) et l'équation 3 implique l'équation 2 (en multipliant par 2).

Exemple 1.17 *Le raisonnement suivant n'est pas correct :*

$$\begin{cases} x + y = 1 \\ x - y = 3 \\ x - 2y = 1 \end{cases} \Leftrightarrow \begin{cases} y = 1 - x \\ x - (1 - x) = 3 \end{cases}$$

$$\Leftrightarrow \begin{cases} 2x = 4 \\ y = 1 - x \end{cases}$$

$$\Leftrightarrow \begin{cases} x = 2 \\ y = -1 \end{cases}$$

Dans le passage du premier système au second, l'implication directe est vraie (on retranche x dans la première équation, puis on remplace y par sa valeur dans la deuxième équation), mais pas l'implication réciproque (on a oublié la troisième équation du premier système, et on ne peut pas la retrouver à partir du second système).
En fait, on obtient seulement

$$\begin{cases} x + y = 1 \\ x - y = 3 \\ x - 2y = 1 \end{cases} \Rightarrow \begin{cases} x = 2 \\ y = -1 \end{cases}$$

ce qui signifie que la seule solution possible au système est $x = 2$ et $y = -1$. Mais ce n'est pas une solution car $\begin{cases} 2 + (-1) = 1 \\ 2 - (-1) = 3 \\ 2 - 2 \times (-1) = 1 \end{cases}$ *est faux, donc le système n'a pas de solution.*

Raisonnement par récurrence
Ce type de raisonnement n'est pas aussi général que les précédents ; il ne s'applique que lorsque l'on veut démontrer une propriété $P(n)$ qui dépend d'un entier n. Nous le signalons ici car il est très souvent

utilisé.
Pour un énoncé $P(n)$ qui porte sur un entier n, et n_0 un certain entier fixé, si $P(n_0)$ est vrai et si pour tout $n \geq n_0$, $P(n) \Rightarrow P(n+1)$ est vrai, alors pour tout $n \geq n_0$, $P(n)$ est vrai. C'est une application en cascade du principe du modus ponens : $P(n_0)$ est vrai par hypothèse ; comme $P(n_0) \Rightarrow P(n_0+1)$ est vrai, $P(n_0+1)$ est vrai ; comme $P(n_0+1) \Rightarrow P(n_0+2)$ est vrai, $P(n_0+2)$ est vrai ; etc...

Exemple 1.18 *On cherche à donner une formule pour calculer la somme des n premiers nombres entiers.*
Notation : on note

$$\sum_{i=0}^{n} i = 0 + 1 + \ldots + n.$$

Plus généralement, pour une suite (u_i), on note la somme des u_i pour i allant de 0 à n

$$\sum_{i=0}^{n} u_i = u_0 + u_1 + \ldots + u_n.$$

On a une notation similaire pour le produit des u_i pour i allant de 0 à n :

$$\prod_{i=0}^{n} u_i = u_0 \times u_1 \times \ldots \times u_n.$$

Revenons à notre problème : on va montrer par récurrence sur $n \geq 0$ l'énoncé

$$P(n) : \sum_{i=0}^{n} i = \frac{n(n+1)}{2}$$

$P(0)$ est vrai : $0 = \frac{0\times 1}{2}$.
Pour tout $n \geq 0$, on montre que $P(n) \Rightarrow P(n+1)$: on suppose que $P(n)$ est vrai (pour ce n donné) et on cherche à montrer que $P(n+1)$ est vrai :

$$\begin{aligned}
\sum_{i=0}^{n+1} i &= \Big(\sum_{i=0}^{n} i\Big) + (n+1) \\
&= \frac{n(n+1)}{2} + (n+1) && \text{en utilisant } P(n) \\
&= (n+1)\big(\frac{n}{2} + 1\big) && \text{en factorisant par } n+1 \\
&= \frac{(n+1)(n+2)}{2}
\end{aligned}$$

C'est exactement l'énoncé $P(n+1)$. On a donc montré par récurrence que $P(n)$ est vrai pour tout entier $n \geq 0$.

Chapitre 2

Ensembles et applications

2.1 Ensembles

Les ensembles sont des collections d'objets, qu'on appelle éléments de l'ensemble (ceci n'est pas une définition, seulement une manière de décrire les ensembles). Si X est un ensemble, on note $x \in X$ pour dire que x appartient à X, ou encore x est un élément de X (et $x \notin X$ pour la négation de $x \in X$). Le point essentiel est qu'un ensemble est déterminé par ses éléments, c'est-à-dire que l'énoncé suivant est vrai :

$$X = Y \Leftrightarrow (\forall x\ (x \in X \Leftrightarrow x \in Y)),$$

ou encore : deux ensembles sont égaux ssi ils ont exactement les mêmes éléments.

On peut préciser : on dit que l'ensemble X est inclus (ou contenu) dans l'ensemble Y, ou encore que X est un sous-ensemble de Y, si $\forall x\ (x \in X \Rightarrow x \in Y)$. On note alors $X \subset Y$.

Proposition 2.1 *$X = Y$ ssi ($X \subset Y$ et $Y \subset X$).*

Preuve D'après les tautologies vues au chapitre précédent, $(\forall x\ (x \in X \Rightarrow x \in Y))$ et $(\forall x\ (x \in Y \Rightarrow x \in X))$ est équivalent à $\forall x \in X\ ((x \in X \Rightarrow x \in Y)$ et $(x \in Y \Rightarrow x \in X))$, c'est-à-dire $\forall x\ (x \in X \Leftrightarrow x \in Y)$, ce qui est équivalent à $X = Y$. □

On note souvent $X \subsetneq Y$ pour dire que X est inclus dans Y mais $X \neq Y$: tous les éléments de X appartiennent à Y mais il existe (au moins) un élément de Y qui n'est pas dans X.

On a plusieurs moyens pour décrire des ensembles. On peut déjà considérer certains ensembles qu'on considère bien connus : $\mathbb{N}$, $\mathbb{R}$, ... Pour des ensembles finis (c'est-à-dire des ensembles avec un nombre fini d'éléments), il suffit d'en faire la liste ; on note $\{a_1, \ldots, a_n\}$ l'ensemble dont les éléments sont exactement $a_1, \ldots, a_n$. L'ordre et les éventuelles répétitions n'ont pas d'importance : par exemple, $\{1, 2, 1\}$, $\{2, 1\}$ et $\{1, 2\}$ représentent le même ensemble.

On peut aussi utiliser des ensembles connus pour décrire de nouveaux ensembles. Si X est un ensemble et $P(x)$ un énoncé qui dépend d'une variable dans X, on peut regarder le sous-ensemble de X formé des éléments tels que $P(x)$ est vrai, on le note

$$\{x \in X ; P(x)\}.$$

Exemple 2.2 *$\{n \in \mathbb{N} ; \exists k \in \mathbb{N}\ n = 2k\}$ est l'ensemble des entiers naturels pairs.*
$\mathbb{R}_+ = \{x \in \mathbb{R} ; x \geq 0\}$

En particulier, on peut choisir pour $P(x)$ un énoncé qui est toujours faux :

$$\{x \in X ; x \neq x\}.$$

Cet ensemble n'a aucun élément, on l'appelle l'ensemble vide et on le note $\emptyset$. Il n'y a qu'un ensemble vide (il est caractérisé par le fait qu'il n'a pas d'élément).

Proposition 2.3 *Pour tout ensemble X, $\emptyset \subset X$.*

Preuve $\forall x\ (x \in \emptyset \Rightarrow x \in X)$ est vrai car $x \in \emptyset$ est toujours faux. □

Soient X et Y deux ensembles, on définit :

1. la réunion de X et Y, dont les éléments sont les objets qui appartiennent à X ou à Y (éventuellement aux deux) :
$$x \in X \cup Y \Leftrightarrow x \in X \text{ ou } x \in Y$$
2. l'intersection de X et Y, dont les éléments sont les objets qui appartiennent à la fois à X et à Y :
$$x \in X \cap Y \Leftrightarrow x \in X \text{ et } x \in Y$$
On peut aussi écrire $X \cap Y = \{x \in X; x \in Y\} = \{x \in Y; x \in X\}$.
On dit que X et Y sont disjoints si $X \cap Y = \emptyset$ (il n'y a pas d'éléments à la fois dans X et Y).
3. la différence X privé de Y, dont les éléments sont les éléments de X qui n'appartiennent pas à Y :
$$x \in X \setminus Y \Leftrightarrow x \in X \text{ et non } x \in Y$$
On peut aussi écrire $X \setminus Y = \{x \in X; x \notin Y\}$.

   Exemple 2.4 $\mathbb{R}^* = \mathbb{R} \setminus \{0\} = \{x \in \mathbb{R}; x \neq 0\}$

   Si Y est un sous-ensemble de X, on dit aussi que $X \setminus Y$ est le complémentaire de Y dans X. Parfois, l'ensemble X est sous-entendu, et on dit seulement que $X \setminus Y$ est le complémentaire de Y, et on note Y^c.

   Exemple 2.5 *Dans l'ensemble* $\mathbb{N}$*, soit* $\mathbb{P}$ *le sous-ensemble des entiers naturels pairs. Alors* $\mathbb{P}^c$*, le complémentaire de* $\mathbb{P}$ *(sous-entendu : dans* $\mathbb{N}$*) est le sous-ensemble des entiers naturels impairs. Dans* $\mathbb{R}$*,* $[1,3]^c =]-\infty, 1[\cup]3, +\infty[$.

4. la différence symétrique de X et de Y, dont les éléments sont les objets qui appartiennent à X ou à Y mais pas aux deux :
$$X \Delta Y = (X \cup Y) \setminus (X \cap Y) = (X \setminus Y) \cup (Y \setminus X)$$
5. l'ensemble des parties de X, dont les éléments sont les sous-ensembles de X :
$$Y \in \mathcal{P}(X) \Leftrightarrow Y \subset X$$

   Exemple 2.6 *Si* $X = \{1, 2\}$*, alors* $\mathcal{P}(X) = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}$.

Définition 2.7 *Pour deux objets* x *et* y*, on définit le couple* (x, y)*. Il est caractérisé par la propriété suivante :*
$$(x, y) = (x', y') \Leftrightarrow (x = x' \text{ et } y = y').$$

Remarque 2.8 *Ne pas confondre le couple* (x, y) *avec la paire* $\{x, y\}$*. Par exemple,* $(1, 2) \neq (2, 1)$ *mais* $\{1, 2\} = \{2, 1\}$.

Définition 2.9 *Pour deux ensembles* X *et* Y*, on définit le produit cartésien* $X \times Y$ *: c'est l'ensemble dont les éléments sont tous les couples* (x, y) *avec* $x \in X$ *et* $y \in Y$*.*

Définition 2.10 *Plus généralement, pour tout entier* $n \geq 1$*, on définit le* n*-uplet* $(x_1, \ldots, x_n)$*, caractérisé par le fait que*
$$(x_1, \ldots, x_n) = (x'_1, \ldots, x'_n) \Leftrightarrow (x_1 = x'_1 \text{ et } \ldots \text{ et } x_n = x'_n).$$
Pour n *ensembles* $X_1, \ldots, X_n$*, on définit le produit cartésien* $X_1 \times \ldots \times X_n$ *: c'est l'ensemble dont les éléments sont tous les* n*-uplets* $(x_1, \ldots, x_n)$ *avec* $x_1 \in X_1, \ldots, x_n \in X_n$*.*
On note $X^n = X \times \ldots \times X$ *(*n *fois).*

2.2 Applications

Soient X et Y deux ensembles. Une application f de X dans Y, notée $f : X \to Y$, est un objet f qui à tout élément $x \in X$ associe un élément de Y, noté $f(x)$. On écrit souvent $x \mapsto f(x)$. On dit que X est l'ensemble de départ de f et Y l'ensemble d'arrivée.

Exemple 2.11 *Pour tout ensemble X, on peut définir l'application identité, notée id ou id_X :*

$$\begin{array}{rccc} id & : & X & \to & X \\ & & x & \mapsto & x \end{array}$$

En choisissant un élément y de Y, on peut définir l'application constante égale à y :

$$\begin{array}{ccc} X & \to & Y \\ x & \mapsto & y \end{array}$$

On peut aussi utiliser le terme fonction au lieu d'application. De manière générale, une fonction f de X dans Y est une application de X' dans Y pour un certain sous-ensemble X' de X, appelé le domaine de définition de f.

Exemple 2.12 *La fonction $x \mapsto \frac{1}{x}$ de $\mathbb{R}$ dans $\mathbb{R}$ désigne l'application de $\mathbb{R}^*$ dans $\mathbb{R}$ qui à tout $x \in \mathbb{R}^*$ associe $\frac{1}{x}$.*

Définition 2.13 *Soit une application $f : X \to Y$. Le graphe de f est l'ensemble de tous les couples $(x, f(x))$ avec $x \in X$. C'est un sous-ensemble de $X \times Y$.*

Remarque 2.14 *On a l'habitude des applications de $\mathbb{R}$ dans $\mathbb{R}$ (souvent définies par des formules). Comme le graphe d'une application $f : \mathbb{R} \to \mathbb{R}$ est un sous-ensemble du plan $\mathbb{R}^2$, on peut le représenter graphiquement. Par exemple, si $f : \mathbb{R} \to \mathbb{R}$ est définie par $f(x) = x + 1$, le graphe de f est la droite d'équation $y = x + 1$.*
Mais la notion de graphe existe pour toutes les applications, même quand il n'existe pas d'interprétation graphique. Par exemple, si X est l'ensemble des étudiants de la promo et Y l'ensemble des lycées, on peut considérer l'application $f : X \to Y$ qui à chaque étudiant associe son lycée d'origine. Le graphe de f est alors l'ensemble des couples formés d'un étudiant et de son lycée d'origine.

Proposition 2.15 *Soit G un sous-ensemble de $X \times Y$. Alors G est le graphe d'une certaine application $f : X \to Y$ ssi $\forall x \in X \; \exists! y \in Y \; (x, y) \in G$.*

Preuve Si G est le graphe d'une application $f : X \to Y$, G est l'ensemble des couples $(x, f(x))$ avec $x \in X$. Donc, pour tout $x \in X$, le seul $y \in Y$ tel que $(x, y) \in G$ est $y = f(x)$.
Réciproquement, si G vérifie la propriété $\forall x \in X \; \exists! y \in Y \; (x, y) \in G$, on peut définir une application $f : X \to Y$, qui à tout $x \in X$ associe l'unique $y \in Y$ tel que $(x, y) \in G$. Alors le graphe de f est G : si $(x, y) \in G$, alors $y = f(x)$ donc (x, y) est dans le graphe de f ; réciproquement si (x, y) est dans le graphe de f, cela signifie que $y = f(x)$ et donc $(x, y) \in G$. □

Remarque 2.16 *On peut aussi voir par l'argument précédent qu'une application est complètement déterminée par son graphe.*

Remarque 2.17 *En prenant $X = \emptyset$, $\emptyset \times Y = \emptyset$ et $G = \emptyset$, vu comme sous ensemble de $\emptyset \times Y$ vérifie la propriété qui caractérise les graphes :*

$$\forall x (x \in \emptyset \Rightarrow (\exists! y \in Y \; (x, y) \in G))$$

puisque $x \in \emptyset$ est toujours faux. On appelle application vide l'application de $\emptyset$ dans Y associée à ce graphe, c'est la seule application de $\emptyset$ dans Y.

Définition 2.18 *Soit une application $f : X \to Y$, et X' un sous-ensemble de X. La restriction de f à X', notée $f_{|X'}$, est l'application de X' dans Y qui à tout x dans X associe $f(x)$.*

Exemple 2.19 *Soit l'application $f : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = x^2$. Alors $f_{|\mathbb{N}}$ est l'application*

$$\begin{array}{rccc} f_{|\mathbb{N}} & : & \mathbb{N} & \to & \mathbb{R} \\ & & x & \mapsto & x^2 \end{array}$$

Définition 2.20 *Soient deux applications $f : X \to Y$ et $g : Y \to Z$. On définit l'application composée $g \circ f$: c'est l'application de X dans Z qui à tout x dans X associe $g(f(x))$.*

Exemple 2.21 *Soient f et g les applications de $\mathbb{R}$ dans $\mathbb{R}$ définies par $f(x) = x+1$ et $g(x) = x^2$. Alors $f \circ g$ et $g \circ f$ sont des applications de $\mathbb{R}$ dans $\mathbb{R}$, définies par $f \circ g(x) = x^2 + 1$ et $g \circ f(x) = (x+1)^2$.*

Remarque 2.22 *Attention : on ne peut définir $g \circ f$ que si l'ensemble d'arrivée de f est le même que l'ensemble de départ de g.*

Proposition 2.23 *Soit une application $f : X \to Y$. Alors $f \circ id_X = f$ et $id_Y \circ f = f$.*

Preuve Il suffit de calculer, pour tout $x \in X$: $f \circ \mathrm{id}_X(x) = f(\mathrm{id}_X(x)) = f(x)$ et $\mathrm{id}_Y \circ f(x) = \mathrm{id}_Y(f(x)) = f(x)$. □

Définition 2.24 *Soit une application $f : X \to Y$.*

1. *Soit X' un sous-ensemble de X. On appelle image directe de X' par f, et on note $f(X')$, le sous-ensemble de Y dont les éléments sont tous les $f(x)$ avec $x \in X'$, ou encore*

$$f(X') = \{y \in Y; \exists x \in X' \ y = f(x)\}.$$

2. *Soit Y' un sous-ensemble de Y. On appelle image réciproque de Y' par f, et on note $f^{-1}(Y')$, le sous-ensemble suivant de X*

$$f^{-1}(Y') = \{x \in X; f(x) \in Y'\}$$

Exemple 2.25 *Soit l'application $f : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = x^2$.*
Montrons que $f([0,2]) = [0,4]$. Si $0 \leq x \leq 2$, alors $0 \leq x^2 \leq 4$, ce qui montre que $f([0,2]) \subset [0,4]$. Puis, pour l'inclusion réciproque, si $y \in [0,4]$, $\sqrt{y} \in [0,2]$, ce qui montre que $y = f(\sqrt{y}) \in f([0,2])$. On conclut que $f([0,2]) = [0,4]$.
Montrons que $f^{-1}([0,4]) = [-2,2]$. Si $x \in [-2,2]$, alors $f(x) \in [0,4]$, ce qui signifie que $x \in f^{-1}([0,4])$. Puis, plutôt que de montrer l'implication réciproque, on montre la contraposée de l'implication réciproque : si $x \notin [-2,2]$, alors $x > 2$ ou $x < -2$, et dans les deux cas, $f(x) > 4$ donc $x \notin f^{-1}([0,4])$. On conclut que $f^{-1}([0,4]) = [-2,2]$.

Définition 2.26 *Soit une application $f : X \to Y$. On dit que f est surjective, ou que f est une surjection, ssi $f(X) = Y$.*
Cela revient à dire, puisque l'inclusion $f(X) \subset Y$ est toujours vraie, que

$$\forall y \in Y \ \exists x \in X \ f(x) = y.$$

Remarque 2.27 *Attention, cette notion dépend de l'ensemble d'arrivée Y.*
Par exemple, si $f : \mathbb{R} \to \mathbb{R}$ est l'application définie par $f(x) = x^2$, on voit facilement que $f(\mathbb{R}) = \mathbb{R}_+ \subsetneq \mathbb{R}$, donc f n'est pas surjective. Mais pour $g : \mathbb{R} \to \mathbb{R}_+$ définie par $g(x) = x^2$ (définie par la même formule que f mais avec un ensemble d'arrivée différent), alors on a de la même manière $g(\mathbb{R}) = \mathbb{R}_+$ ce qui donne cette fois que g est surjective.

Définition 2.28 *Soit une application $f : X \to Y$. On dit que f est injective, ou que f est une injection, ssi :*

$$\forall x \in X \ \forall x' \in X \ (x \neq x' \Longrightarrow f(x) \neq f(x')),$$

ou, de manière équivalente, en considérant la contraposée :

$$\forall x \in X \ \forall x' \in X \ (f(x) = f(x') \Longrightarrow x = x').$$

Exemple 2.29 *1. L'application $f : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = x + 1$ est injective : pour tous réels x et x', si $x + 1 = x' + 1$, alors $x = x'$.*

2. L'application $f : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = x^2$ n'est pas injective : il suffit de trouver deux réels distincts qui ont la même image par f, par exemple $f(2) = f(-2) = 4$.

3. Soit X l'ensemble des étudiants de la promo et $f : X \to \mathbb{N}$ l'application qui à tout étudiant associe son numéro d'étudiant. Alors f est injective : deux étudiants différents ont leurs numéros d'étudiant différents. Et f n'est pas surjective : il suffit pour le voir de trouver un entier qui n'est le numéro d'étudiant d'aucun étudiant.

Définition 2.30 *Soit une application $f : X \to Y$. On dit que f est bijective, ou que f est une bijection, ssi f est à la fois injective et surjective.*

Exemple 2.31 *Pour tout ensemble X, l'application identité $id_X : X \to X$ est bijective. En effet, comme $id_X(x) = x$, si $id_X(x) = id_X(x')$, alors $x = x'$, donc id_X est injective. Et pour tout $x \in X$, $x = id_X(x) \in id_X(X)$: id_X est surjective.*

Remarque 2.32 *Soit $f : X \to Y$ une application. Dire que f est surjective, c'est dire que pour tout $y \in Y$, il existe au moins un $x \in X$ tel que $f(x) = y$. Dire que f est injective, c'est dire que si $f(x) = f(x')$, alors $x = x'$. Et donc dire que f est bijective est équivalent à :*

$$\forall y \in Y \ \exists! x \in X \ f(x) = y.$$

Théorème 2.33 *Soit une application $f : X \to Y$, où X et Y sont des ensembles non vides.*

1. *f est injective ssi il existe une application $g : Y \to X$ telle que $g \circ f = id_X$.*
2. *f est surjective ssi il existe une application $g : Y \to X$ telle que $f \circ g = id_Y$.*
3. *f est bijective ssi il existe une application $g : Y \to X$ telle que $g \circ f = id_X$ et $f \circ g = id_Y$.*

Preuve

1. On suppose que f est injective. Fixons un élément $a \in X$. On définit $g : Y \to X$ par

$$g(y) = \begin{cases} \text{l'unique } x \in X \text{ tel que } f(x) = y \text{ si un tel } x \text{ existe} \\ a \text{ sinon} \end{cases}$$

   Dans le premier cas, x est unique car f est injective, donc $f(x) = f(x') = y$ implique $x = x'$. Vérifions que $g \circ f = \mathrm{id}_X$: pour tout $x \in X$, si $y = f(x)$, $g(y) = x$ car x est l'unique élément tel que $f(x) = y$, et ainsi $g(f(x)) = x$.
   Réciproquement, on suppose qu'il existe $g : Y \to X$ telle que $g \circ f = \mathrm{id}_X$. Supposons que $f(x) = f(x')$. En appliquant g, on trouve que $g(f(x)) = g(f(x'))$, et donc que $x = x'$ puisque $g \circ f = \mathrm{id}_X$.

2. On suppose que f est surjective. On définit $g : Y \to X$ de la manière suivante : pour tout $y \in Y$, il existe au moins un $x \in X$ tel que $f(x) = y$ (car f est surjective), on en choisit un et on pose $g(y) = x$. On peut alors vérifier que $f \circ g = \mathrm{id}_Y$: pour tout $y \in Y$, posons $x = g(y)$, alors $f(x) = y$ d'après la construction de g, donc $f(g(y)) = y$. Remarque : on a utilisé sans le dire pour la définition de g un axiome appelé l'axiome du choix.
   Réciproquement, on suppose qu'il existe $g : Y \to X$ telle que $f \circ g = \mathrm{id}_Y$. Alors pour tout $y \in Y$, en posant $x = g(y)$, on voit que $y = f(g(y)) = f(x)$, donc f est surjective.

3. On suppose que f est bijective. On définit $g : Y \to X$ de la manière suivante : pour tout $y \in Y$, d'après la remarque précédente, il existe un unique $x \in X$ tel que $f(x) = y$, et on pose $g(y) = x$. Alors $g \circ f = \mathrm{id}_X$: si $x \in X$ et $y = f(x)$, $g(y) = x$ par définition de g, donc $g(f(x)) = x$. Et $f \circ g = \mathrm{id}_Y$: si $y \in Y$ et $x = g(y)$, alors $f(x) = y$ par définition de g, donc $f(g(y)) = y$.
   Réciproquement, s'il existe une application $g : Y \to X$ telle que $g \circ f = \mathrm{id}_X$ et $f \circ g = \mathrm{id}_Y$, on peut en particulier en déduire que f est injective d'après le cas 1, et que f est surjective d'après le cas 2, donc f est bijective.

□

Exemple 2.34 *Soit $f : \mathbb{R} \to \mathbb{R}$ l'application définie par $f(x) = x^2$. On a déjà vu que $f(\mathbb{R}) = \mathbb{R}_+$; on sait aussi que $f(x) = f(x')$ ssi $x = x'$ ou $x = -x'$. On en déduit les résultats suivants :*

1. *Soit $f : \mathbb{R}_+ \to \mathbb{R}$ définie par $f(x) = x^2$. Alors f est injective. Construisons $g : \mathbb{R} \to \mathbb{R}_+$ comme dans la preuve du théorème, avec 0 comme valeur "par défaut" : si $y < 0$, il n'existe pas de $x \in \mathbb{R}_+$ tel que $f(x) = y$, et on pose donc $g(y) = 0$; si $y \geq 0$, l'unique $x \in \mathbb{R}_+$ tel que $f(x) = y$ est $x = \sqrt{y}$, on pose donc $g(y) = \sqrt{y}$. Et on vérifie bien que $g \circ f = id_{\mathbb{R}_+}$: pour tout $x \in \mathbb{R}_+$, $g(f(x)) = \sqrt{x^2} = |x| = x$.*
2. *Soit $f : \mathbb{R} \to \mathbb{R}_+$ définie par $f(x) = x^2$. Alors f est surjective. Construisons $g : \mathbb{R}_+ \to \mathbb{R}$ comme dans la preuve du théorème : pour tout $y \in \mathbb{R}_+$, il existe deux (ou un seul si $y = 0$) $x \in \mathbb{R}$ tels que $f(x) = y$, à savoir $\sqrt{y}$ et $-\sqrt{y}$, et il faut en choisir un des deux. On peut par exemple définir :*

$$g : y \mapsto \sqrt{y} \quad ou \quad g : y \mapsto -\sqrt{y} \quad ou \quad g : y \mapsto \begin{cases} \sqrt{y} \text{ si } y \in \mathbb{Q} \\ -\sqrt{y} \text{ si } y \in \mathbb{R}_+ \setminus \mathbb{Q} \end{cases}$$

   Dans tous les cas, $f \circ g = id_{\mathbb{R}_+}$: pour tout $y \in \mathbb{R}_+$, $f(g(y)) = (\pm\sqrt{y})^2 = y$.
3. *Soit $f : \mathbb{R}_+ \to \mathbb{R}_+$ définie par $f(x) = x^2$. Alors f est bijective. Pour tout $y \in \mathbb{R}_+$, l'unique élément $x \in \mathbb{R}_+$ tel que $f(x) = y$ est $x = \sqrt{y}$. On définit donc $g : \mathbb{R}_+ \to \mathbb{R}_+$ par $g(y) = \sqrt{y}$, et on vérifie comme dans les deux cas précédents que $g \circ f = f \circ g = id_{\mathbb{R}_+}$.*

Proposition et définition 2.35 *Soit $f : X \to Y$ une application bijective. Alors il existe une unique application $g : Y \to X$ telle que $f \circ g = id_Y$ et $g \circ f = id_X$. On l'appelle application réciproque de f et on la note f^{-1}. Pour tout $x \in X$ et tout $y \in Y$, on a l'équivalence :*

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

Preuve On sait déjà d'après le théorème précédent qu'il existe une application $g : Y \to X$ telle que $f \circ g = \mathrm{id}_Y$ et $g \circ f = \mathrm{id}_X$. Reste à montrer que g est unique : si $h : Y \to X$ est une autre application telle que $f \circ h = \mathrm{id}_Y$ et $h \circ f = \mathrm{id}_X$, on calcule de deux manières $g \circ f \circ h = g \circ \mathrm{id}_Y = g$ et $g \circ f \circ h = \mathrm{id}_X \circ h = h$, et on conclut que $g = h$.
Puis, si $y = f(x)$, on obtient en appliquant f^{-1} que $f^{-1}(y) = f^{-1}(f(x)) = x$. Et réciproquement, si $x = f^{-1}(y)$, on obtient en appliquant $f(x) = f(f^{-1}(y)) = y$. □

Proposition 2.36 *Soit $f : X \to Y$ une bijection. Alors $f^{-1} : Y \to X$ est aussi une bijection, et $(f^{-1})^{-1} = f$.*

Preuve L'application réciproque est caractérisée par $f \circ f^{-1} = \mathrm{id}_Y$ et $f^{-1} \circ f = \mathrm{id}_X$. D'après le théorème 2.33, l'existence de f implique que f^{-1} est bijective, et $f = (f^{-1})^{-1}$. □

Exemple 2.37
1. *Dans l'exemple précédent, avec $f : \mathbb{R}_+ \to \mathbb{R}_+$ définie par $f(x) = x^2$, on a vu que $f^{-1} : \mathbb{R}_+ \to \mathbb{R}_+$ est définie par $f^{-1}(y) = \sqrt{y}$.*
2. *Soit $f = id_X$, alors $f^{-1} = id_X$ car $id_X \circ id_X = id_X$.*

2.3 Cas des ensembles finis

Soit X un ensemble fini, c'est-à-dire un ensemble avec un nombre fini d'éléments. On appelle cardinal de X le nombre d'éléments de X, et on note $\mathrm{card}(X)$, $|X|$ ou $\#X$. Par exemple, $\mathrm{card}(\emptyset) = 0$.

Proposition 2.38 *Soit une application $f : X \to Y$ où X et Y sont des ensembles finis.*

1. *Si f est injective, alors $card(X) \leq card(Y)$.*
2. *Si f est surjective, alors $card(X) \geq card(Y)$.*
3. *Si f est bijective, alors $card(X) = card(Y)$.*

Preuve

1. On raisonne par récurrence sur $\text{card}(X)$.
   Si $\text{card}(X) = 0$, le résultat est évident puisqu'on a nécessairement $\text{card}(Y) \geq 0$.
   On suppose que le résultat est vrai quand $\text{card}(X) = n$, et on charche à montrer le résultat pour un ensemble X de cardinal $n+1$. Fixons $x \in X$ et posons $X' = X \setminus \{x\}$, c'est un ensemble de cardinal n. Notons que pour tout $x' \in X$, $f(x') \neq f(x)$ car f est injective ; on peut donc définir

$$\begin{array}{rccc} g & : & X' & \to & Y \setminus \{f(x)\} \\ & & x' & \mapsto & f(x') \end{array}$$

   C'est encore une application injective : pour deux éléments distincts x' et x'' dans X', $f(x') \neq f(x'')$ car f est injective et donc $g(x') \neq g(x'')$. Donc d'après l'hypothèse de récurrence $\text{card}(X') \leq \text{card}(Y \setminus \{f(x)\})$. En ajoutant un élément de chaque côté, on trouve bien que $\text{card}(X) \leq \text{card}(Y)$, ce qui achève la démonstration par récurrence.
2. On va utiliser deux fois le théorème 2.33. Comme $f : X \to Y$ est surjective, il existe une application $g : Y \to X$ telle que $f \circ g = \text{id}_Y$. Puis, l'existence de f telle que $f \circ g = \text{id}_Y$ montre que g est injective. On peut donc appliquer le point précédent : $\text{card}(Y) \leq \text{card}(X)$.
3. Puisque f est bijective, elle est injective et surjective, et on peut donc appliquer les deux points précédents pour conclure que $\text{card}(X) = \text{card}(Y)$.

□

Remarque 2.39 *En prenant les contraposées des implications précédentes, on obtient aussi des résultats importants :*

1. *Si card$(X) >$ card(Y), il n'existe pas d'injection $X \to Y$.*
2. *Si card$(X) <$ card(Y), il n'existe pas de surjection $X \to Y$.*
3. *Si card$(X) \neq$ card(Y), il n'existe pas de bijection $X \to Y$.*

On appelle parfois ces résultats le principe des tiroirs : si on veut ranger m chaussettes dans n tiroirs, avec $m > n$, alors un des tiroirs contiendra plus d'une chaussette (car la fonction qui associe à une chaussette le tiroir dans laquelle on la range ne peut pas être injective). Et si on veut ranger m chaussettes dans n tiroirs, avec $m < n$, il y aura au moins un tiroir qui ne contiendra aucune chaussette (car la fonction qui associe à une chaussette le tiroir dans laquelle on la range ne peut pas être surjective).

Corollaire 2.40 *Soient X et Y deux ensembles finis tels que card$(X) =$ card(Y), et soit une application $f : X \to Y$.*
Alors f est bijective ssi (f est injective ou f est surjective).
En d'autres termes, si f est injective, elle est automatiquement surjective, et si f est surjective, elle est automatiquement injective.

Preuve Supposons que $f : X \to Y$ est injective. Supposons par l'absurde qu'elle n'est pas surjective. Alors il existe $y \in Y$ tel que pour tout $x \in X$, $f(x) \neq y$. L'application f permet donc de définir une application de $\tilde{f} : X \to Y \setminus \{y\}$ avec $\tilde{f}(x) = f(x)$ pour tout $x \in X$. Il est clair que $\tilde{f}$ est encore injective. Or $Y \setminus \{y\}$ a un élément de moins que Y, donc aussi que X : c'est impossible d'après le principe des tiroirs.
Si on suppose maintenant que f est surjective, on trouve d'près le théorème 2.33 une application $g : Y \to X$ telle que $f \circ g = \text{id}_Y$. En utilisant encore le théorème 2.33, on en déduit que g est injective, et donc qu'elle est bijective d'après le point précédent. En composant l'égalité $f \circ g = \text{id}_Y$ par g^{-1}, on trouve $f \circ g \circ g^{-1} = \text{id}_Y \circ g^{-1}$, et donc $f = g^{-1}$, qui est une application bijective d'après la proposition 2.36. □

Remarque 2.41 *Attention, le résultat est bien entendu faux si on ne suppose pas card$(X) =$ card(Y), puisqu'on sait qu'il n'existe pas de bijection entre deux ensembles finis de cardinal différent.*

Corollaire 2.42 *Soient deux ensembles finis X et Y de même cardinal tel que $X \subset Y$. Alors $X = Y$.*

Preuve Comme $X \subset Y$, on peut définir "l'application d'inclusion" :

$$\begin{array}{ccc} X & \to & Y \\ x & \mapsto & x \end{array}$$

Cette application est clairement injective, elle est donc surjective, ce qui signifie que pour tout $y \in Y$, il existe $x \in X$ tel que $y = x$. Ainsi $Y \subset X$ et donc $X = Y$. □

Remarque 2.43 *Les corollaires précédents peuvent être généralisés à d'autres situations qui seront vues par la suite, par exemple en algèbre linéaire.*

Le fait suivant est intuitivement clair.

Fait 2.44 *Soient X et Y deux ensembles finis disjoints. Alors card$(X \cup Y)$ = card(X) + card(Y).*

Corollaire 2.45 *Soient $X_1, \ldots, X_n$ des ensembles finis deux à deux disjoints. Alors*

$$\text{card}(X_1 \cup \ldots \cup X_n) = \text{card}(X_1) + \ldots + \text{card}(X_n).$$

Preuve On montre le résultat par récurrence sur $n \geq 1$.
Si $n = 1$, le résultat est évident : $\text{card}(X_1) = \text{card}(X_1)$.
Puis on suppose le résultat vrai pour un $n \geq 1$, et on veut le montrer pour $n + 1$. Si $X_1, \ldots, X_{n+1}$ sont deux à deux disjoints, X_{n+1} n'a pas délément commun avec $X_1, \ldots, X_n$, donc il est disjoint de $X_1 \cup \ldots \cup X_n$. On peut donc appliquer le fait 2.44 à $X_1 \cup \ldots \cup X_n$ et X_{n+1} : $\text{card}(X_1 \cup \ldots \cup X_n \cup X_{n+1}) = \text{card}(X_1 \cup \ldots \cup X_n) + \text{card}(X_{n+1})$. Or les ensembles $X_1, \ldots, X_n$ sont deux à deux disjoints, donc $\text{card}(X_1 \cup \ldots \cup X_n) = \text{card}(X_1) + \ldots + \text{card}(X_n)$ d'après l'hypothèse de récurrence, ce qui permet de conclure que $\text{card}(X_1 \cup \ldots \cup X_{n+1}) = \text{card}(X_1) + \ldots + \text{card}(X_{n+1})$. □

Corollaire 2.46 *Soient X et Y deux ensembles finis. Alors card$(X \cup Y)$ = card(X) + card(Y) − card$(X \cap Y)$.*

Preuve Il est facile de vérifier que $X \cup Y = (X \setminus Y) \cup (Y \setminus X) \cup (X \cap Y)$, et que ces trois ensembles sont deux à deux disjoints. Le corollaire précédent donne donc que $\text{card}(X \cup Y) = \text{card}(X \setminus Y) + \text{card}(Y \setminus X) + \text{card}(X \cap Y)$. De plus, $X = (X \setminus Y) \cup (X \cap Y)$, et ces deux ensembles sont disjoints, donc $\text{card}(X) = \text{card}(X \setminus Y) + \text{card}(X \cap Y)$, ou encore $\text{card}(X \setminus Y) = \text{card}(X) - \text{card}(X \cap Y)$. De la même manière, $\text{card}(Y \setminus X) = \text{card}(Y) - \text{card}(X \cap Y)$. En reportant dans la première égalité, on obtient :

$$\text{card}(X \cup Y) = \text{card}(X) - \text{card}(X \cap Y) + \text{card}(Y) - \text{card}(X \cap Y) + \text{card}(X \cap Y) = \text{card}(X) + \text{card}(Y) - \text{card}(X \cap Y).$$

□

Corollaire 2.47 *Soit X et Y deux ensembles finis et $f : X \to Y$ une surjection. On suppose qu'il existe un entier p tel que pour tout $y \in Y$, card$(f^{-1}(\{y\})) = p$. Alors card$(X) = p \times$ card(Y).*

Preuve Posons $\text{card}(Y) = n$ et $y_1, \ldots, y_n$ les éléments de Y. Pour i entre 1 et n, on pose $X_i = f^{-1}(\{y_i\})$. Par hypothèse, X_i est de cardinal p. On voit que $X = X_1 \cup \ldots \cup X_n$: pour tout $x \in X$, $f(x) = y_i$ pour un certain i entre 1 et n, ce qui nous dit que $x \in X_i$, et donc que $X \subset X_1 \cup \ldots \cup X_n$ (l'inclusion réciproque étant évidente puisque les X_i sont des sous-ensembles de X). De plus, les ensembles $X_1, \ldots, X_n$ sont deux à deux disjoints : si $i \neq j$, si un élément x est dans $X_i \cap X_j$, on devrait avoir $f(x) = y_i = y_j$, ce qui est impossible. On peut donc appliquer le corollaire précédent :

$$\text{card}(X) = \text{card}(X_1) + \ldots + \text{card}(X_n) = p \times n = p \times \text{card}(Y).$$

□

Corollaire 2.48 *Soient X et Y deux ensembles finis. Alors card$(X \times Y)$ = card$(X) \times$ card(Y).*

Preuve Soit l'application (appelée projection sur Y)

$$\begin{array}{rccc} f & : & X \times Y & \to & Y \\ & & (x,y) & \mapsto & y \end{array}$$

Cette application est surjective : pour tout $y \in Y$, on choisit $x \in X$, et alors $f(x,y) = y$ (le cas où $X = \emptyset$ se traite différemment mais est facile). De plus, pour tout $y \in Y$, $f^{-1}(\{y\}) = X \times \{y\}$, qui a le même cardinal que X (on peut le justifier par le fait que l'application $X \to X \times \{y\}, x \mapsto (x,y)$ est une bijection). En utilisant le corollaire précédent, on peut donc conclure que $\text{card}(X \times Y) = \text{card}(X) \times \text{card}(Y)$. □

Corollaire 2.49 *Soient X et Y deux ensembles finis, et Z l'ensemble de toutes les applications de X dans Y. Alors*

$$card(Z) = card(Y)^{card(X)}$$

Preuve On prouve le résultat par récurrence sur $\text{card}(X)$.
Si $\text{card}(X) = 0$, c'est-à-dire $X = \emptyset$, la seule application de $\emptyset$ dans Y est l'application vide, on a donc bien $\text{card}(Z) = 1 = \text{card}(Y)^0$.
On suppose le résultat connu pour $\text{card}(X) = n$, et on veut le montrer pour $\text{card}(X) = n+1$. Si $\text{card}(X) = n+1$, on choisit $x_0 \in X$ et on pose $X' = X \setminus \{x_0\}$, qui est de cardinal n. Posons Z' l'ensemble des applications de X' dans Y. Considérons l'application suivante :

$$\begin{array}{rccc} g & : & Z & \to & Z' \times Y \\ & & f & \mapsto & (f_{|X'}, f(x_0)) \end{array}$$

On montre que g est injective : si $g(f) = g(f')$, alors $f_{|X'} = f'_{|X'}$ et $f(x_0) = f'(x_0)$, ce qui permet de dire que $f(x) = f'(x)$ pour tout $x \in X$, et donc $f = f'$. Et on montre que g est surjective : soit $(f_0, y) \in Z' \times Y$, posons $f : X \to Y$ définie par $f(x) = f_0(x)$ si $x \in X'$ et $f(x_0) = y$. On a bien $g(f) = (f_0, y)$. Comme g est bijective, $\text{card}(Z) = \text{card}(Z' \times Y)$, et donc $\text{card}(Z) = \text{card}(Z') \times \text{card}(Y)$ d'après le corollaire précédent. Or, d'après l'hypothèse de récurrence, $\text{card}(Z') = \text{card}(Y)^{\text{card}(X')} = \text{card}(Y)^n$. Donc

$$\text{card}(Z) = \text{card}(Y)^n \times \text{card}(Y) = \text{card}(Y)^{n+1} = \text{card}(Y)^{\text{card}(X)}.$$

□

D'autres résultats de cardinalité sont importants à retenir. Se reporter aux feuilles d'exercices pour les démonstrations. On rappelle la notation pour la factorielle d'un entier n : $n! = 1 \times 2 \times \ldots \times n$, et par convention $0! = 1$.

Fait 2.50
1. *Soit X un ensemble fini de cardinal n, et Y l'ensemble des bijections de X dans X (on dit aussi : les permutations de X). Alors $card(Y) = n!$.*
2. *Soit X un ensemble fini de cardinal n, et p un entier tel que $0 \leq p \leq n$. Soit Y l'ensemble des sous-ensembles de X qui sont de cardinal p. Alors $card(Y) = \frac{n!}{p!(n-p)!}$, qu'on appelle coefficient binomial $\binom{n}{p}$. L'appellation "coefficient binomial" provient de la formule du binôme de Newton, que l'on énonce ci-dessous.*

Lemme 2.51 *Formule de Pascal.*
Soient des entiers $0 < p < n$. Alors $\binom{n}{p} = \binom{n-1}{p} + \binom{n-1}{p-1}$.

Preuve On peut donner deux preuves.
La première est un calcul de fractions :

$$\begin{aligned} \binom{n-1}{p} + \binom{n-1}{p-1} &= \frac{(n-1)!}{p!(n-1-p)!} + \frac{(n-1)!}{(p-1)!(n-p)!} \\ &= \frac{(n-1)!}{(p-1)!(n-1-p)!} \times (\frac{1}{p} + \frac{1}{n-p}) \\ &= \frac{(n-1)!}{(p-1)!(n-1-p)!} \times \frac{n}{p(n-p)} \\ &= \frac{n!}{p!(n-p)!} = \binom{n}{p} \end{aligned}$$

La deuxième preuve utilise la caractérisation de $\binom{n}{p}$ comme nombre de parties à p éléments d'un ensemble à n éléments. Soit X un ensemble à n éléments. Par hypothèse, $n \geq 1$ donc on peut fixer a un élément de X. L'ensemble des parties de X à p éléments est l'union de P_1, l'ensemble des parties de X à p éléments qui contiennent a, et de P_2, l'ensemble des parties de X à p éléments qui ne contiennent pas a. De plus, ces deux ensembles sont disjoints, donc par les résultats précédents, $\binom{n}{p}$, qui est le cardinal de l'ensembles des parties de X à p éléments, est égal à $\text{card}(P_1) + \text{card}(P_2)$. Puis, comme une partie de X (à p éléments) qui ne contient pas a est exactement une partie (à p éléments) de $X \setminus \{a\}$, qui est un ensemble à $n-1$ éléments, on obtient $\text{card}(P_2) = \binom{n-1}{p}$. Quant à P_1, choisir une partie de X à p éléments qui contient a revient exactement à choisir une partie de $X \setminus \{a\}$ à $p-1$ éléments et à lui joindre a. Si on veut être plus précis, on peut montrer que l'application $Y \mapsto Y \cup \{a\}$ est une bijection de l'ensemble des parties de $X \setminus \{a\}$ à $p-1$ éléments vers l'ensemble des parties de X à p éléments qui contiennent a. Cela montre que $\text{card}(P_1) = \binom{n-1}{p-1}$. On conclut donc que $\binom{n}{p} = \binom{n-1}{p} + \binom{n-1}{p-1}$. □

Théorème 2.52 *Formule du binôme de Newton.*
Soit un entier $n \geq 1$ et deux réels a et b (le théorème est en fait encore valable dans des contextes plus généraux, par exemple pour a et b deux nombres complexes). Alors

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

Preuve On montre la formule par récurrence sur n.
Pour $n = 1$, $\binom{1}{0} = \binom{1}{1} = 1$, on a donc bien $(a+b)^1 = a + b = \binom{1}{0} a^0 b^1 + \binom{1}{1} a^1 b^0$.
Passons à l'hérédité. On suppose que $(a+b)^n = \sum_{p=0}^{n} \binom{n}{p} a^p b^{n-p}$ et on veut montrer la formule au rang $n+1$. On développe et on regroupe les termes :

$$\begin{aligned}
(a+b)^{n+1} &= (a+b)(a+b)^n \\
&= a \times \sum_{p=0}^{n} \binom{n}{p} a^p b^{n-p} + b \times \sum_{p=0}^{n} \binom{n}{p} a^p b^{n-p} \\
&= \underbrace{\sum_{p=0}^{n} \binom{n}{p} a^{p+1} b^{n-p}}_{q=p+1} + \underbrace{\sum_{p=0}^{n} \binom{n}{p} a^p b^{n+1-p}}_{q=p} \\
&= \sum_{q=1}^{n+1} \binom{n}{q-1} a^q b^{n+1-q} + \sum_{q=0}^{n} \binom{n}{q} a^q b^{n+1-q} \\
&= \underbrace{\binom{n}{0}}_{=1=\binom{n+1}{0}} a^0 b^{n+1} + \sum_{q=1}^{n} \underbrace{\left(\binom{n}{q-1} + \binom{n}{q}\right)}_{=\binom{n+1}{q}} a^q b^{n+1-q} + \underbrace{\binom{n}{n}}_{=1=\binom{n+1}{n+1}} a^{n+1} b^0 \\
&= \sum_{q=0}^{n+1} \binom{n+1}{q} a^q b^{n+1-q}
\end{aligned}$$

□

Chapitre 3

Relations sur un ensemble

3.1 Généralités

Soit X un ensemble. Une relation binaire sur l'ensemble X est un énoncé $P(x, y)$ dépendant de deux éléments x et y de X. Si on appelle R cette relation, on note souvent xRy pour dire que $P(x, y)$ est vrai.

Exemple 3.1
1. *Soit X n'importe quel ensemble. L'égalité est une relation binaire sur X : $x = y$.*
2. *Sur $\mathbb{R}$, on a la relation d'ordre $x \leq y$.*
3. *Soit X l'ensemble des étudiants de la promo. On définit une relation binaire sur X par : x est plus jeune que y.*

Remarque 3.2 *On s'intéressera seulement ici aux relations binaires (on dit aussi : d'arité 2), c'est-à-dire aux énoncés qui dépendent de deux variables. Mais on peut aussi parler de relations de n'importe quelle autre arité. Par exemple, sur l'ensemble $\mathbb{R}$, on peut définir une relation ternaire (ou encore : d'arité 3) par $x < y < z$.*

Définition 3.3 *Soit R une relation binaire sur un ensemble X. On dit que :*
1. *R est réflexive si pour tout x dans X, xRx.*
   Par exemple, l'gálité est une relation réflexive (sur n'importe quel ensemble) car on a toujours $x = x$.
2. *R est une relation symétrique si pour tous x et y dans X, xRy équivaut à yRx.*
   Par exemple, l'gálité est une relation symétrique (sur n'importe quel ensemble) car $x = y$ est toujours équivalent à $y = x$. La relation $\leq$ sur $\mathbb{R}$ n'est pas symétrique car $x \leq y$ n'est pas équivalent à $y \leq x$.
3. *R est une relation antisymétrique si pour tous x et y dans X, (xRy et yRx) implique $x = y$.*
   Par exemple, la relation $\leq$ sur $\mathbb{R}$ est antisymétrique car si $x \leq y$ et $y \leq x$, alors $x = y$.
4. *R est une relation transitive si pour tous x, y, z dans X, (xRy et yRz) implique xRz.*
   Par exemple, la relation xRy définie par "x est plus jeune que y" est transitive : si x est plus jeune que y et y est plus jeune que z, alors x est plus jeune que z.

Remarque 3.4 *Dans la suite, nous étudierons des relations qui vérifient certaines de ces propriétés. Mais il existe aussi des relations binaires qui ne satisfont pas ces propriétés. Par exemple, sur l'ensemble $X = \{1, 2, 3\}$, définissons une relation R en posant que $1R2$, $2R3$ et $3R2$ sont vrais, et tous les autres énoncés xRy sont faux. Alors R n'est pas réflexive puisque $1R1$ est faux ; R n'est pas symétrique car $1R2$ est vrai mais $2R1$ est faux ; R n'est pas antisymétrique car $2R3$ et $3R2$ sont vrais mais $2 \neq 3$; R n'est pas transitive car $1R2$ et $2R3$ sont vrais mais $1R3$ est faux.*

3.2 Relations d'équivalence

Définition 3.5 *Soit R une relation binaire sur un ensemble X. On dit que R est une relation d'équivalence si elle est réflexive, symétrique et transitive.*

Exemple 3.6 *1. La relation d'égalité sur n'importe quel ensemble X est une relation déquivalence. En effet, pour tous x, y, z dans X, on a bien : $x = x$; $x = y$ ssi $y = x$; si $x = y$ et $y = z$ alors $x = z$.*

2. Fixons un entier $d \in \mathbb{Z}$. Sur l'ensemble $\mathbb{Z}$, on définit la relation R par : nRm ssi il existe $k \in \mathbb{Z}$ tel que $n = m + kd$. On note souvent cette relation $n \equiv m\ [d]$ (on lit : n est congru à m modulo d) plutôt que nRm. C'est une relation d'équivalence : $n \equiv n\ [d]$ car $n = n + 0 \times d$; si $n \equiv m\ [d]$, alors il existe $k \in \mathbb{Z}$ tel que $n = m + kd$, et donc $m = n + (-k)d$, c'est-à-dire $m \equiv n\ [d]$; enfin si $n \equiv m\ [d]$ et $m \equiv p\ [d]$, il existe k et l dans $\mathbb{Z}$ tels que $n = m + kd$ et $m = p + ld$, ce qui donne $n = p + (k + l)d$, et donc $n \equiv p\ [d]$.

3. On définit de la même façon la relation $x \equiv y\ [2\pi]$ sur $\mathbb{R}$ par $x \equiv y\ [2\pi]$ ssi il existe $k \in \mathbb{Z}$ tel que $x = y + k \times 2\pi$. C'est une relation d'équivalence par le même argument que pour l'exemple précédent.

Proposition 3.7 *Soit une application $f : X \to Y$. Soit R la relation sur X définie par : xRy ssi $f(x) = f(y)$. Alors R est une relation d'équivalence.*

Preuve La relation R est réflexive car $f(x) = f(x)$. Elle est symétrique car $f(x) = f(y)$ équivaut à $f(y) = f(x)$. Elle est transitive car si $f(x) = f(y)$ et $f(y) = f(z)$, alors $f(x) = f(z)$. □

Exemple 3.8 *Soit X l'ensemble des étudiants de la promo et $f : X \to \mathbb{R}$ l'application qui à un étudiant associe sa note à l'examen. La relation d'équivalence décrite ci-dessus est donnée par : xRy ssi x et y ont eu la même note à l'examen.*

Définition 3.9 *Soit R une relation d'équivalence sur un ensemble X, et $x \in X$. La classe d'équivalence de x, souvent notée $\overline{x}$ ou C_x, est le sous-ensemble de X formé des éléments y équivalents à x :*

$$\overline{x} = C_x = \{y \in X; yRx\}.$$

Proposition 3.10 *Soient x et y deux éléments de X, qui est muni d'une relation d'équivalence R.*
Si xRy, alors $\overline{x} = \overline{y}$.
Si non xRy, alors $\overline{x} \cap \overline{y} = \emptyset$.

Preuve Supposons que xRy. Pour tout $z \in X$, on obtient $zRx \Leftrightarrow zRy$ en utilisant la propriété de transitivité, c'est-à-dire $z \in \overline{x}$ ssi $z \in \overline{y}$. Donc $\overline{x} = \overline{y}$.
Pour le deuxième point, on va montrer la contraposée : si $\overline{x} \cap \overline{y} \neq \emptyset$, choisissons $z \in \overline{x} \cap \overline{y}$, alors zRx et zRy et donc xRy par transitivité. □

Définition 3.11 *Soit R une relation d'équivalence sur l'ensemble X. On définit l'ensemble quotient X/R comme l'ensemble formé des classes d'équivalences $\overline{x}$ pour $x \in X$. On dit que l'application*

$$\begin{array}{ccc} X & \to & X/R \\ x & \mapsto & \overline{x} \end{array}$$

est l'application canonique de X dans X/R.

Remarque 3.12 *L'application canonique $X \to X/R$ est surjective par construction.*
La définition précédente donne une construction réciproque à la proposition 3.7 : si $f : X \to X/R$ est l'application canonique, alors pour tous $x, y \in X$, $f(x) = f(y)$ ssi xRy (c'est une conséquence de la proposition 3.10).

Proposition et définition 3.13 *Soit X/R l'ensemble quotient d'un ensemble X par une relation R. Alors X/R forme une partition de X, au sens où :*

- *les éléments de X/R sont des sous-ensembles non vides de X (car $\overline{x}$ contient x)*
- *les éléments de X/R sont deux à deux disjoints (d'après la proposition 3.10, si $\overline{x} \neq \overline{y}$, alors $\overline{x} \cap \overline{y} = \emptyset$)*

— *les éléments de X/R recouvrent X : tout $x \in X$ est dans un élément de X/R, à savoir $\overline{x}$*

Exemple 3.14 *1. Sur l'ensemble $\mathbb{Z}$, considérons la relation d'équivalence R donnée par nRm ssi $n \equiv m$ [2] (voir l'exemple 3.6). L'ensemble quotient $\mathbb{Z}/R$ est habituellement noté $\mathbb{Z}/2\mathbb{Z}$. Remarquons que 0 et 1 ne sont pas équivalents, car il n'existe pas $k \in \mathbb{Z}$ tel que $1 = 0 + 2k$. Donc $\overline{0} \neq \overline{1}$. Par définition, $\overline{0} = \{2k; k \in \mathbb{Z}\}$: c'est l'ensemble des entiers relatifs pairs. Et $\overline{1} = \{2k+1; k \in \mathbb{Z}\}$: c'est l'ensemble des entiers relatifs impairs. S'il existait une autre classe d'équivalence, elle devrait être disjointe de $\overline{0}$ et de $\overline{1}$, ce qui est impossible puisque tout entier relatif est soit pair, soit impair. On a donc monté que $\mathbb{Z}/2\mathbb{Z} = \{\overline{0}, \overline{1}\}$. Notons que ce choix des représentants 0 et 1 pour ces classes d'équivalence est arbitraire : on voit par exemple que $0 \equiv 6$ [2] et $1 \equiv 3$ [2], on peut donc aussi écrire $\mathbb{Z}/2\mathbb{Z} = \{\overline{6}, \overline{5}\}$*

2. Soit l'application $f : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = x^2$, et R la relation d'équivalence sur $\mathbb{R}$ définie par xRy ssi $f(x) = f(y)$. On sait que xRy ssi $y = \pm x$. Alors les classes d'équivalence sont $\overline{0} = \{0\}$ et, pour tout $x > 0$, $\overline{x} = \overline{-x} = \{x, -x\}$.

On utilisera la proposition suivante dans la suite (on l'admet sans démonstration).

Proposition 3.15 *Soit X un ensemble muni d'une relation d'équivalence R, $f : X \to X/R$ l'application canonique et $g : X \to Y$ une application telle que pour tous $x, y \in X$, si xRy, alors $g(x) = g(y)$. Alors il existe une unique application $h : X/R \to Y$ telle que $g = h \circ f$.*

Exemple 3.16 *Soit X l'ensemble des étudiants de la promo, et R la relation d'équivalence définie par xRy ssi x et y ont la même note à l'examen, avec $f : X \to X/R$ l'application canonique. Soit Y l'ensemble des mentions, et $g : X \to Y$ telle que $g(x)$ est la mention obtenue par l'étudiant x. Chaque élément de X/R correspond à une note ; l'application h est l'application qui associe aux notes entre 10 et 12 la mention passable, aux notes entre 12 et 14 la mention assez bien, etc. Alors on a bien $g = h \circ f$.*

3.3 Relations d'ordre

Définition 3.17 *Soit R une relation binaire sur un ensemble X. On dit que R est une relation d'ordre si elle est réflexive, antisymétrique et transitive.*

Exemple 3.18 *1. Soit X l'ensemble $\mathbb{N}$ ou $\mathbb{Z}$ ou $\mathbb{R}$, et R la relation $\leq$. C'est une relation d'ordre : $x \leq x$; si $x \leq y$ et $y \leq x$ alors $x = y$; si $x \leq y$ et $y \leq z$ alors $x \leq z$.*

2. Soit Y un ensemble et $X = \mathcal{P}(Y)$ l'ensemble des sous-ensembles de Y. Soit la relation $\subset$ sur X. C'est une relation d'ordre : $A \subset A$; si $A \subset B$ et $B \subset A$ alors $A = B$; si $A \subset B$ et $B \subset C$ alors $A \subset C$.

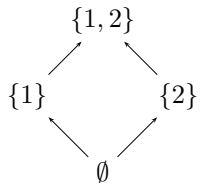
Remarque 3.19 *On a défini ici les relations d'ordre R au sens large, au sens où elles vérifient xRx. On peut toujours leur associer une relation d'ordre strict, définie par xRy et $x \neq y$. Par exemple, la relation d'ordre strict associée à $\leq$ est $<$, la relation d'ordre strict associée à $\subset$ est $\subsetneq$.*

Définition 3.20 *Soit R une relation d'ordre sur un ensemble X.*
On dit que R est une relation d'ordre total si pour tous x, y dans X, xRy ou yRx.
Sinon, on dit que R est une relation d'ordre partiel : il existe deux éléments x, y dans X tels qu'on n'a ni xRy, ni yRx (on dit alors que x et y sont incomparables).

Exemple 3.21 *1. Sur $\mathbb{N}$ (ou $\mathbb{Z}$, ou $\mathbb{R}$), l'ordre $\leq$ est total : on a toujours soit $x \leq y$, soit $y \leq x$.*

2. Soit $X = \mathcal{P}(\{1, 2\})$, l'ordre $\subset$ est partiel, car les ensembles $\{1\}$ et $\{2\}$ sont incomparables.

On peut représenter une relation d'ordre avec un graphe, tel que xRy ssi il existe un chemin allant de x à y en suivant les flèches.
Par exemple, pour l'ordre $\subset$ sur $X = \mathcal{P}(\{1, 2\})$:



On n'a pas besoin de tracer de flèche de $\emptyset$ vers $\{1,2\}$ car l'énoncé $\emptyset \subset \{1,2\}$ peut être déduit des autres flèches.
Dans le cas d'une relation d'ordre total (par exemple $\leq$ sur $\mathbb{N}$), on obtient un graphe linéaire :

$$0 \to 1 \to 2 \to \dots$$

Définition 3.22 *Soit X un ensemble avec une relation d'ordre $\leq$, et $x \in X$. On dit que :*
- *x est le plus grand élément de X si pour tout $y \in X$, $y \leq x$*
- *x est le plus petit élément de X si pour tout $y \in X$, $x \leq y$*
- *x est un élément maximal de X s'il n'existe pas de $y \in X$ tel que $x < y$*
- *x est un élément minimal de X s'il n'existe pas de $y \in X$ tel que $y < x$*

Exemple 3.23
1. *Dans $\mathbb{N}$ avec l'ordre $\leq$, 0 est le plus petit élément, et il n'y a pas de plus grand élément.*
2. *Dans l'ensemble $X = \{\emptyset, \{1\}, \{2\}\}$, avec l'ordre $\subset$: $\emptyset$ est un élément minimal, et aussi le plus petit élément de X ; $\{1\}$ et $\{2\}$ sont des éléments maximaux de X, mais ce ne sont pas des plus grands éléments (car $\{1\} \not\subset \{2\}$ et $\{2\} \not\subset \{1\}$).*

Proposition 3.24 *Si $\leq$ est une relation d'ordre total sur X, et x un élément minimal de X, alors c'est le plus petit élément de X.*
De même, si x est un élément maximal de X, alors c'est le plus grand élément de X.

Preuve On suppose que x un élément minimal de X. Soit $y \in X$. Comme x est un élément minimal de X, on ne peut pas avoir $y < x$. Or, comme l'ordre $\leq$ est total, $y \leq x$ ou $x \leq y$, il faut donc que $x \leq y$: x est bien le plus petit élément de X.
L'argument est le même pour le deuxième point. □

Chapitre 4

Arithmétique

On entend par arithmétique l'étude des entiers avec les opérations élémentaires (addition, multiplication...). Nous donnerons les définitions et les résultats pour l'ensemble des entiers naturels $\mathbb{N}$, avec éventuellement les adaptations pour l'ensemble des entiers relatifs $\mathbb{Z}$.

4.1 Divisibilité

Définition 4.1 *Soient m et n deux entiers. On définit la relation $m|n$: $m|n$ ssi il existe $k \in \mathbb{N}$ tel que $n = mk$. On dit alors que m divise n, ou encore que m est un diviseur de n, que n est divisible par m ou que n est un multiple de m.*

Exemple 4.2 *2 divise 6, 3 ne divise pas 5.*
Pour tout entier n, $1|n$ car $n = 1 \times n$.
Pour tout entier n, $n|0$ car $0 = n \times 0$.

Proposition 4.3 *Si $m|n$, alors $m \leq n$ ou $n = 0$.*
La relation $|$ est une relation d'ordre, avec 1 comme plus petit élément et 0 comme plus grand élément (d'après les exemples précédents).
Si $m|n$ et $m|n'$, alors $m|n + n'$.
Par définition, $m|mk$.

Preuve Si $m|n$, on trouve $k \in \mathbb{N}$ tel que $n = mk$. Il y a deux cas : soit $k = 0$ et alors $n = m \times 0 = 0$; soit $k \geq 1$ et alors $n = mk \geq m$.
La relation $|$ est réflexive : pour tout $n \in \mathbb{N}$, $n|n$ car $n = n \times 1$.
La relation $|$ est antisymétrique : si $n|m$ et $m|n$, on veut montrer que $n = m$. D'après la première propriété, comme $n|m$, alors $n \leq m$ ou $m = 0$. Si $m = 0$, comme $m|n$, on peut écrire $n = mk$, alors $n = 0 \times k = 0 = m$. Si $n \leq m$, comme $m|n$, alors $m \leq n$ ou $n = 0$: dans le premier cas, $n = m$ car la relation $\leq$ est antisymétrique, dans le deuxième cas, on conclut de $n = 0$ et de $n|m$ que $m = 0 = n$ comme précédemment.
La relation $|$ est transitive : si $m|n$ et $n|p$, on veut montrer que $m|p$. Comme $m|n$, on trouve $k \in \mathbb{N}$ tel que $n = mk$. Comme $n|p$, on trouve $l \in \mathbb{N}$ tel que $p = nl$. On en déduit que $p = nl = mkl$ donc $m|p$.
Si $m|n$ et $m|n'$, on trouve k et k' dans $\mathbb{N}$ tels que $n = mk$ et $n' = mk'$. Alors $n+n' = mk+mk' = m(k+k')$ donc $m|n + n'$. □

Remarque 4.4 *On définit la relation $|$ sur $\mathbb{Z}$: pour m et n dans $\mathbb{Z}$, $m|n$ s'il existe $k \in \mathbb{Z}$ tel que $n = mk$. La relation $|$ n'est pas une relation d'ordre sur $\mathbb{Z}$ car elle n'est pas antisymétrique : par exemple $3| - 3$ car $-3 = (-1) \times 3$ et $-3|3$ car $3 = (-1) \times (-3)$, mais $3 \neq -3$.*

Définition 4.5 *Soit $p \in \mathbb{N}$. On dit que p est un nombre premier si $p \geq 2$ et si les seuls diviseurs de p sont 1 et p.*

Exemple 4.6 2, 3, 5, 7,... *sont premiers. Dans la pratique, comme les diviseurs de* p *(pour* $p \neq 0$*) sont* $\leq p$*, il suffit de vérifier s'il existe entre* 1 *et* p *d'autres diviseurs de* p *que* 1 *et* p*.*
6 *n'est pas premier car* $6 = 2 \times 3$ *:* 2 *et* 3 *sont des diviseurs de* 6 *distincts de* 1 *et* 6*.*

4.2 Division euclidienne

Le résultat de division euclidienne sera essentiel pour les théorèmes suivants.

Proposition et définition 4.7 *Soit* m *et* n *deux entiers naturels avec* $m \neq 0$*. Alors il existe des entiers naturels* q *et* r*, avec* $r < m$*, tels que* $n = mq + r$*.*
Ces entiers q *et* r *sont uniques ;* q *s'appelle le quotient de la division euclidienne de* n *par* m*, et* r *le reste.*

Exemple 4.8 *Division euclidienne de* 311 *par* 12 *:* $311 = 12 \times 25 + 11$*, avec* $11 < 12$ *;* 25 *est le quotient, et* 11 *le reste.*

Preuve On fixe $m \neq 0$ et on montre par récurrence sur n la propriété $P(n)$: il existe q et r dans $\mathbb{N}$ avec $r < m$ tel que $n = mq + r$.
$P(0)$ est vraie : pour $q = 0$ et $r = 0 < m$, on a bien $0 = m \times 0 + 0$.
$P(n) \Rightarrow P(n+1)$: d'après $P(n)$, on trouve q et r dans $\mathbb{N}$ avec $r < m$ tels que $n = mq + r$. On cherche maintenant à faire la division euclidienne de de $n+1$ par m : on peut écrire $n+1 = mq+r+1$. Puisque $r < m$, il y a deux cas : soit $r < m-1$, et alors $r' = r+1 < m$ et on a bien $n+1 = mq + r'$; soit $r = m-1$, et alors $n+1 = mq + m = m(q+1) + 0$ avec $0 < m$.
Cela montre la propriété $P(n)$ pour tout $n \in \mathbb{N}$. Il faut encore montrer l'unicité de q et r. Supposons qu'on a deux écritures $n = mq + r = mq' + r'$ avec $r < m$ et $r' < m$. Si $r = r'$, on obtient en retranchant r que $mq = mq'$, et donc que $q = q'$ (en divisant par m, qui est non nul). Si $r \neq r'$, quitte à échanger r avec r' et q avec q', on peut supposer que $r' > r$. En faisant la différence entre les deux écritures de n, on obtient $m(q-q') + (r-r') = 0$, ou encore $m(q-q') = r' - r$, avec $0 < r' - r < m$. C'est impossible, car cela signifierait que m divise $r - r'$, ce qui implique $r - r' = 0$ ou $m \leq r - r'$. □

Remarque 4.9 $m|n$ *ssi le reste de la division euclidienne de* n *par* m *est* 0*.*

On utilisera aussi dans la suite une version de la division euclidienne pour les entiers relatifs.

Proposition 4.10 *Soient* m *et* n *dans* $\mathbb{Z}$ *avec* $m \neq 0$*. Alors il existe* q *et* r *dans* $\mathbb{Z}$*, uniques, tels que* $n = mq + r$ *et* $0 \leq r < |m|$*.*

4.3 Algorithme d'Euclide

On se pose la question suivante : soient a et b deux éléments de $\mathbb{N}$. Peut-on trouver parmi les diviseurs communs de a et de b un plus grand élément, pour la relation $|$?

Exemple 4.11 *Soient* a *et* b *deux nombres premiers distincts. Les seuls diviseurs de* a *sont* 1 *et* a*, les seuls diviseurs de* b *sont* 1 *et* b*, et comme* $a \neq b$*, le seul diviseur commun de* a *et de* b *est* 1*, et c'est donc évidemment le plus grand diviseur commun.*
Si $b = 0$*, tous les diviseurs de* a *sont aussi des diviseurs de* 0*. Or le plus grand diviseur de* a *est* a *pour l'ordre* $|$ *(car* a *divise* a *et si* n *divise* a *alors* $n|a$*), donc* a *est le plus grand diviseur commun de* a *et de* 0*.*

Lemme 4.12 *Soit* a *et* b *dans* $\mathbb{N}$ *avec* $b \neq 0$*. Soit* $a = bq + r$ *la division euclidienne de* a *par* b*. Alors pour tout* $n \in \mathbb{N}$*,* n *divise à la fois* a *et* b *ssi* n *divise à la fois* b *et* r*.*

Preuve Si n divise a et b, on peut trouver k et l tels que $a = nk$ et $b = nl$, ce qui donne $r = a - bq = n(k - lq)$, donc n divise r, et il divise b par hypothèse.
Si n divise b et r, on peut trouver k et l tels que $b = nk$ et $r = nl$, ce qui donne $a = bq + r = n(kq + l)$,

donc n divise a, et il divise b par hypothèse. □

On peut maintenant décrire l'algorithme d'Euclide. Posons $r_0 = a$ et $r_1 = b$. Si $b = 0$, on sait déjà que le plus grand diviseur commun de a et de b est a. Si $b \neq 0$, on peut faire la division euclidienne par b : $r_0 = q_0 r_1 + r_2$ avec $r_2 < r_1$. On continue à faire une suite de divisions euclidiennes tant que c'est possible, c'est-à-dire tant que le dernier reste obtenu est $\neq 0$:

$$\begin{aligned} r_0 &= q_0 r_1 + r_2 \text{ avec } r_2 < r_1 && (1)\\ r_1 &= q_1 r_2 + r_3 \text{ avec } r_3 < r_2 && (2)\\ &\vdots \\ r_{n-1} &= q_{n-1} r_n + r_{n+1} \text{ avec } r_{n+1} < r_n && (n) \end{aligned}$$

On s'arrête quand le dernier reste r_{n+1} obtenu est 0 ; cette situation se produit obligatoirement car les suites strictement décroissantes d'entiers naturels $r_1 > r_2 > \ldots > r_n > r_{n+1}$ atteignent nécessairement 0.
En appliquant n fois le lemme 4.12, on obtient que les diviseurs communs de r_0 et de r_1 sont les mêmes que ceux de r_1 et de r_2, qui sont les mêmes que ceux de r_2 et de r_3, ..., qui sont les mêmes que ceux de r_n et de r_{n+1}. Or, comme $r_{n+1} = 0$, on a vu que le plus grand diviseur commun de r_n et de r_{n+1} est r_n ; c'est donc aussi le plus grand diviseur commun de a et de b. On a donc démontré le résultat suivant :

Proposition et définition 4.13 *Soient a et b dans $\mathbb{N}$. Il existe un plus grand diviseur commun de a et de b (pour l'ordre $|$), que l'on note $pgcd(a, b)$.*
Il est obtenu comme le dernier reste non-nul dans l'algorithme d'Euclide.

Exemple 4.14 *Calculer $pgcd(51, 18)$.*
On utilise l'algorihme d'Euclide :

$$\begin{aligned} \underbrace{51}_{r_0} &= 2 \times \underbrace{18}_{r_1} + \underbrace{15}_{r_2} && (1)\\ \underbrace{18}_{r_1} &= 1 \times \underbrace{15}_{r_2} + \underbrace{3}_{r_3} && (2)\\ \underbrace{15}_{r_2} &= 5 \times \underbrace{3}_{r_3} + \underbrace{0}_{r_4} && (3) \end{aligned}$$

Le dernier reste non-nul est 3, donc $pgcd(51, 18) = 3$: 3 divise à la fois 51 et 18 ($51 = 3 \times 17$ et $18 = 3 \times 6$) et tout entier n qui divise à la fois 51 et 18 divise 3 (donc $n = 1$ ou $n = 3$).

Définition 4.15 *On dit que deux entiers a et b sont premiers entre eux si $pgcd(a, b) = 1$.*

Exemple 4.16 *On a vu dans l'exemple 4.11 que si a et b sont deux nombres premiers distincts, alors ils sont premiers entre eux.*
On peut avoir deux nombres premiers entre eux même s'ils ne sont pas premiers. Par exemple, on constate par l'algorithme d'Euclide que $pgcd(4, 9) = 1$ mais ni 4, ni 9 ne sont premiers.

Proposition 4.17 *Identité de Bézout.*
Soient a et b dans $\mathbb{N}$, et $d = pgcd(a, b)$. Alors il existe u et v dans $\mathbb{Z}$ tels que $d = au + bv$.

Preuve La preuve qu'on donne ici permet de calculer u et v. On pose $r_0 = a$, $r_1 = b$, et on applique

l'algorithme d'Euclide :

$$\begin{aligned} r_0 &= q_0 r_1 + r_2 \text{ avec } r_2 < r_1 && (1)\\ r_1 &= q_1 r_2 + r_3 \text{ avec } r_3 < r_2 && (2)\\ &\vdots \\ r_{n-2} &= q_{n-2} r_{n-1} + r_n \text{ avec } r_n < r_{n-1} && (n-1)\\ r_{n-1} &= q_{n-1} r_n + r_{n+1} \text{ avec } r_{n+1} < r_n && (n) \end{aligned}$$

où r_n est le dernier reste non-nul, c'est-à-dire $r_{n+1} = 0$ et $r_n = pgcd(a,b) = d$. On va montrer successivement, pour $i = n, n-1, \ldots, 0$, qu'il existe u_i et v_i dans $\mathbb{Z}$ tels que $d = r_i u_i + r_{i+1} v_i$ (c'est une "récurrence inverse").
L'énoncé est évident pour $i = n$: on a par construction $d = r_n = r_n \times 1 + r_{n+1} \times 0$.
On suppose que l'énoncé est vrai au rang i pour $i \geq 1$: $d = r_i u_i + r_{i+1} v_i$ et on veut le montrer au rang $i-1$. On utilise l'équation (i) : $r_{i-1} = q_{i-1} r_i + r_{i+1}$, donc $r_{i+1} = r_{i-1} - q_{i-1} r_i$, et donc

$$d = r_i u_i + r_{i+1} v_i = r_i u_i + (r_{i-1} - q_{i-1} r_i) v_i = r_{i-1} v_i + r_i (u_i - q_{i-1} v_i).$$

C'est l'identité qu'on voulait obtenir au rang $i-1$.
En particulier, au rang $i = 0$, avec $a = r_0$ et $b = r_1$, on trouve $d = a u_0 + b v_0$. □

Remarque 4.18 *On ne s'est intéressé qu'à des entiers naturels a et b, mais les coefficients u et v trouvés sont en général des entiers relatifs.*

Exemple 4.19 *On a vu dans l'exemple 4.14 que $pgcd(51, 18) = 3$. En utilisant les équations qu'on a données, on va trouver u et v dans $\mathbb{Z}$ tels que $3 = 51u + 18v$. L'équation (2) donne $3 = 18 - 15$, et l'équation (1) donne $15 = 51 - 18 \times 2$, donc $3 = 18 - (51 - 18 \times 2) = -51 + 18 \times 3 = 51u + 18v$ avec $u = -1$ et $v = 3$.*

Corollaire 4.20 *Soient a et b dans $\mathbb{N}$. Alors a et b sont premiers entre eux ssi il existe u et v dans $\mathbb{Z}$ tels que $1 = au + bv$.*

Preuve L'implication directe est la proposition précédente avec $d = 1$.
Pour l'implication réciproque, supposons qu'il existe u et v dans $\mathbb{Z}$ tels que $1 = au + bv$. Si n est un diviseur commun de a et de b, alors on obtient en utilisant la proposition 4.3 que $n|a|au$ et $n|b|bv$ et donc $n|au + bv$, c'est-à-dire $n|1$. On a vu que 1 est le plus petit élément pour l'ordre $|$, ce qui donne que le seul diviseur de 1 est 1. Ainsi, $n = 1$, 1 est le seul diviseur commun de a et de b et donc $pgcd(a,b) = 1$. □

4.4 Décomposition en facteurs premiers

Le résultat suivant est appelé lemme de Gauss, mais c'est un résultat essentiel.

Lemme 4.21 *Lemme de Gauss.*
Soient a, b, c dans $\mathbb{N}$. On suppose que $a|bc$ et que a et b sont premiers entre eux. Alors $a|c$.

Preuve On utilise le corollaire 4.20 : comme a et b sont premiers entre eux, il existe u et v dans $\mathbb{Z}$ tels que $au + bv = 1$. En multipliant par c, on trouve $c = acu + bcv$. Or $a|bc$ donc il existe $k \in \mathbb{N}$ tel que $bc = ak$. On obtient donc $c = acu + akv = a(cu + kv)$, ce qui dit que $a|c$. □

Corollaire 4.22 *Soit p un nombre premier et $a_1 \ldots a_n$ un produit de n entiers naturels ($n \geq 1$). Si $p|a_1 \ldots a_n$, alors p divise l'un des a_i.*

Preuve On montre le résultat par récurrence sur le nombre n de facteurs dans le produit.
Si $n = 1$, le résultat est trivial : si $p|a_1$, alors $p|a_1$.
On suppose le résultat vrai pour un certain n et on le montre pour $n + 1$. On suppose que p divise $a_1 \dots a_{n+1}$, qu'on voit comme $a_1(a_2 \dots a_{n+1})$. Si $p|a_1$, on a terminé. Sinon, p et a_1 sont premiers entre eux : en effet, comme p est premier, les seuls diviseurs de p sont 1 et p ; or p ne divise pas a_1 donc le seul diviseur commun de p et a_1 est 1. On peut donc appliquer le lemme de Gauss : p divise $a_2 \dots a_{n+1}$. Comme c'est un produit de n nombres entiers, on peut appliquer l'hypothèse de récurrence : p divise a_2 ou ... ou a_{n+1}. □

Théorème 4.23 *Décomposition en facteurs premiers.*
Soit un entier $n \geq 2$. Alors il existe des nombres premiers $p_1, \dots, p_k$ tels que $n = p_1 \dots p_k$.
De plus, cette décomposition est unique : si $n = p_1 \dots p_k$ et $n = q_1 \dots q_l$ sont deux décompositions en facteurs premiers, alors $k = l$ et, quitte à faire une permutation de $q_1, \dots, q_l$, $p_i = q_i$ pour tout i entre 1 et k.

Preuve On montre par récurrence sur $n \geq 2$ que n admet une décomposition en facteurs premiers.
Si $n = 2$, on a déjà la décomposition puisque 2 est premier.
On montre la propriété d'hérédité sous la forme suivante : on suppose que le résultat est vrai pour tout entier $< n$ et on le démontre pour n. Si n est premier, on a terminé. Sinon, on trouve deux entiers u et v tels que $n = uv$ avec u et v strictement compris entre 1 et n. On peut donc utiliser l'ypothèse de récurrence pour u et v : $u = p_1 \dots p_k$ et $v = q_1 \dots q_l$ pour des nombres premiers $p_1, \dots, p_k, q_1, \dots, q_l$. Et donc $n = uv = p_1 \dots p_k q_1 \dots q_l$.
On démontre maintenant l'unicité : on suppose que $n = p_1 \dots p_k = q_1 \dots q_l$ pour des nombres premiers $p_1, \dots, p_k, q_1, \dots, q_l$. On montre le résultat par récurrence sur le nombre de facteurs dans ces décompositions, c'est-à-dire sur $k + l$. Par construction, $k \geq 1$ et $l \geq 1$. Le cas d'initialisation est donc $k = l = 1$, et la situation est alors $n = p_1 = q_1$, et le résultat est évident.
On écrit maintenant $p_1 \dots p_k = q_1 \dots q_l$ et on suppose le résultat vrai quand le nombre de facteurs est $< k+l$. En particulier, p_1 divise $q_1 \dots q_l$, et p_1 est premier. Donc, par le corollaire 4.22, p_1 divise l'un des q_i. Quitte à faire une permutation de $q_1, \dots, q_l$, on peut supposer que $p_1|q_1$. Comme q_1 est premier et $p_1 \neq 1$, cela implique $p_1 = q_1$. On peut maintenant diviser l'égalité $p_1 \dots p_k = q_1 \dots q_l$ par p_1 : on obtient $p_2 \dots p_k = q_2 \dots q_l$ (dans le cas $k = 1$, le membre de gauche est en fait 1, donc celui de droite aussi, ce qui donne $l = 1$, et on a terminé ; même chose si $l = 1$). On obtient donc deux décompositions avec moins de facteurs premiers (à savoir $k + l - 2$), et l'hypothèse de récurrence donnne donc $k - 1 = l - 1$, et $p_2 = q_2, \dots, p_k = q_k$ quitte à réordonner $q_2, \dots, q_k$. □

Remarque 4.24 *Soit $n = p_1 \dots p_k$ la décomposition de n en facteurs premiers. Il est possible que certains des p_i soient égaux entre eux ; dans ce cas, on les regroupe. Par exemple, pour $n = 12 = 2 \times 2 \times 3$, on écrit $12 = 2^2 \times 3$.*
On peut faire apparaitre de manière artificielle un nombre premier dans une telle décomposition en lui assignant l'exposant 0. Par exemple, on peut écrire $12 = 2^2 \times 3^1 \times 5^0$. Cette idée est principalement utilisée pour comparer les décompositions en facteurs premiers de plusieurs nombres entiers. On pourra écrire 1 sous cette forme en fixant tous les exposants à 0.

Lemme 4.25 *Soit deux entiers naturels n et $m \geq 1$. D'après la remarque précédente, on peut écrire les décompositions en facteurs premiers $n = p_1^{n_1} \dots p_k^{n_k}$ et $m = p_1^{m_1} \dots p_k^{m_k}$ avec $p_1, \dots, p_k$ des nombres premiers deux à deux distincts et des exposants $n_1, \dots, n_k, m_1, \dots, m_k \geq 0$. Alors $n|m$ ssi pour tout i entre 1 et k, $n_i \leq m_i$.*

Preuve On suppose que $n|m$. Alors il existe $u \in \mathbb{N}$ tel que $m = nu$. On peut supposer que la décomposition en facteurs premiers de u s'écrit $u = p_1^{u_1} \dots p_k^{u_k}$ avec $u_i \geq 0$ pour tout i. Alors $m = nu$ s'écrit $p_1^{m_1} \dots p_k^{m_k} = p_1^{n_1+u_1} \dots p_k^{n_k+u_k}$. L'unicité de la décomposition en facteurs premiers implique que $m_i = n_i + u_i$ pour tout i entre 1 et n, et donc aussi $m_i \geq n_i$.
Réciproquement, si $m_i \geq n_i$ pour tout i, on peut considérer l'entier $u = p_1^{m_1-n_1} \dots p_k^{m_k-n_k}$. On calcule alors $nu = p_1^{n_1} \dots p_k^{n_k} p_1^{m_1-n_1} \dots p_k^{m_k-n_k} = p_1^{m_1} \dots p_k^{m_k} = m$, ce qui donne que $n|m$. □

Exemple 4.26 $12 = 2^2 \times 3^1$ *et* $72 = 2^3 \times 3^2$*, donc* $12|72$.
$12 = 2^2 \times 3^1 \times 5^0$ *et* $30 = 2^1 \times 3^1 \times 5^1$ *donc* 12 *ne divise pas* 30*, à cause de l'exposant de* 2.

Proposition 4.27 *Soient deux entiers* n *et* $m \geq 1$*, et les décompositions en facteurs premiers* $n = p_1^{n_1} \dots p_k^{n_k}$ *et* $m = p_1^{m_1} \dots p_k^{m_k}$ *avec* $p_1, \dots, p_k$ *des nombres premiers deux à deux distincts et des exposants* $n_1, \dots, n_k, m_1, \dots, m_k \geq 0$*. Pour tout* i*, posons* $u_i = \min(n_i, m_i)$*. Alors* $pgcd(n, m) = p_1^{u_1} \dots p_k^{u_k}$.

Preuve Posons $u = p_1^{u_1} \dots p_k^{u_k}$. On utilise le lemme précédent : par définition, $u_i \leq n_i$ pour tout i donc $u|n$. De la même manière, $u|m$; donc u est un diviseur commun de m et de n. Soit d un diviseur commun de m et n, écrivons sa décomposition $d = p_1^{d_1} \dots p_k^{d_k}$. Comme $d|n$, on doit avoir $d_i \leq n_i$ pour tout i ; et comme $d|m$, on doit avoir $d_i \leq m_i$ pour tout i. Cela donne $d_i \leq \min(n_i, m_i) = u_i$ pour tout i. Donc $d|u$ d'après le lemme précédent, ce qui signifie que u est bien le plus grand (au sens de la relation $|$) diviseur commun de m et n. □

Une démonstration semblable donne le résultat suivant.

Proposition et définition 4.28 *Soient deux entiers* n *et* $m \geq 1$*, et les décompositions en facteurs premiers* $n = p_1^{n_1} \dots p_k^{n_k}$ *et* $m = p_1^{m_1} \dots p_k^{m_k}$ *avec* $p_1, \dots, p_k$ *des nombres premiers deux à deux distincts et des exposants* $n_1, \dots, n_k, m_1, \dots, m_k \geq 0$*. Pour tout* i*, posons* $v_i = \max(n_i, m_i)$*, et* $v = p_1^{v_1} \dots p_k^{v_k}$*. Alors* v *est le plus petit multiple commun de* m *et* n *(au sens de* $|$*), et on le note* $v = ppcm(m, n)$.

Exemple 4.29 *Les décompositions en facteurs premiers de* 60 *et de* 14 *sont* $60 = 2^2 \times 3^1 \times 5^1 \times 7^0$ *et* $14 = 2^1 \times 3^0 \times 5^0 \times 7^1$ *donc* $pgcd(60, 14) = 2^1 \times 3^0 \times 5^0 \times 7^0 = 2$ *et* $ppcm(60, 14) = 2^2 \times 3^1 \times 5^1 \times 7^1 = 420$.

Proposition 4.30 *Soient deux entiers* m *et* $n \geq 1$*. Alors* $mn = pgcd(m, n) \times ppcm(m, n)$.

Preuve En écrivant les décompositions en facteurs premiers $n = p_1^{n_1} \dots p_k^{n_k}$ et $m = p_1^{m_1} \dots p_k^{m_k}$, on a $mn = p_1^{m_1+n_1} \dots p_k^{m_k+n_k}$ et, d'après les résultats précédents,

$$\text{pgcd}(m, n) \times \text{ppcm}(m, n) = p_1^{\min(m_1,n_1)+\max(m_1,n_1)} \dots p_k^{\min(m_k,n_k)+\max(m_k,n_k)}.$$

Or on a le résultat général : $a + b = \min(a, b) + \max(a, b)$; car si $a \leq b$, $\min(a, b) = a$ et $\max(a, b) = b$, donc $\min(a, b) + \max(a, b) = a + b$, et même raisonnement si $a > b$. En appliquant ce résultat à chacun des exposants, on trouve bien $mn = \text{pgcd}(m, n) \times \text{ppcm}(m, n)$. □

Corollaire 4.31 *Soient* a, b, c *des entiers* ≥ 1*. On suppose que* a *et* b *sont premiers entre eux, que* a *divise* c *et que* b *divise* c*. Alors* ab *divise* c.

Preuve Le fait que a et b sont premiers entre eux siginifie que $\text{pgcd}(a, b) = 1$, et donc $ppcm(a, b) = ab$ d'après la proposition précédente. Or, par hypothèse, c est un multiple commun de a et de b, et comme $ppcm(a, b)$ est le plus petit multiple commun de a et de b pour la relation $|$, on obtient $ab|c$. □

Remarque 4.32 *On a deux méthodes pour calculer le pgcd de deux entiers* m *et* n *: utiliser l'algorithme d'Euclide ou la proposition 4.27 avec les décompositions en facteurs premiers. Pour le calcul du ppcm, on dispose de la proposition et définition 4.28 avec les décompositions en facteurs premiers.*
Parfois, surtout pour de grands entiers m *et* n*, il est difficile de trouver la décomposition en facteurs premiers de* m *et* n*, et plus simple d'utiliser l'algorithme d'Euclide. Ainsi, pour calculer le ppcm de* m *et* n*, on peut calculer* $pgcd(m, n)$ *par l'algorithme d'Euclide et utiliser la formule* $mn = pgcd(m, n) \times ppcm(m, n)$ *pour en déduire* $ppcm(m, n)$.

4.5 Congruences et $\mathbb{Z}/d\mathbb{Z}$

On fixe pour ce paragraphe un entier $d \geq 1$.
On rappelle (voir chapitre 3) qu'on a la relation d'équivalence de congruence modulo d sur $\mathbb{Z}$: pour $n, m \in \mathbb{Z}$,

$$n \equiv m \ [d] \text{ ssi } \exists k \in \mathbb{Z} \ n = m + kd \text{ ssi } d|n - m$$

L'ensemble quotient pour cette relation d'équivalence est noté $\mathbb{Z}/d\mathbb{Z}$; pour $n \in \mathbb{Z}$, on note $\overline{n}$ sa classe d'équivalence.

Proposition 4.33 *L'ensemble $\mathbb{Z}/d\mathbb{Z}$ contient exactement d éléments : $\mathbb{Z}/d\mathbb{Z} = \{\overline{0}, \dots, \overline{d-1}\}$.*

Preuve Les éléments $\overline{0}, \dots, \overline{d-1}$ sont dans $\mathbb{Z}/d\mathbb{Z}$ par définition ; on montre que tous les éléments de $\mathbb{Z}/d\mathbb{Z}$ peuvent être écrits sous cette forme. Tout élément de $\mathbb{Z}/d\mathbb{Z}$ est la classe d'équivalence $\overline{n}$ d'un certain élément n de $\mathbb{Z}$. Écrivons la division euclidienne de n par d : il existe q et r dans $\mathbb{Z}$, avec $0 \leq r < d$, tels que $n = dq + r$. Cela siginifie que $n \equiv r\ [d]$, c'est-à-dire $\overline{n} = \overline{r}$, avec $0 \leq r \leq d-1$.
Puis on montre que les éléments $\overline{0}, \dots, \overline{d-1}$ sont deux à deux distincts. Soient i et j compris entre 0 et $d-1$, avec $i \neq j$. On peut, quitte à échanger i et j, supposer que $i > j$. Alors $0 < i - j < d$, donc d ne peut pas diviser $i - j$, donc $\overline{i} \neq \overline{j}$. □

Lemme 4.34 *On suppose que $n \equiv n'\ [d]$ et $m \equiv m'\ [d]$. Alors $n + m \equiv n' + m'\ [d]$ et $nm \equiv n'm'\ [d]$.*

Preuve On peut écrire $n' = n + kd$ et $m' = m + ld$ pour des éléments l et m de $\mathbb{Z}$. Alors $n' + m' = n + m + (k + l)d$ donc $n + m \equiv n' + m'\ [d]$ et $n'm' = (n + kd)(m + ld) = nm + (km + nl + kld)d$ donc $nm \equiv n'm'\ [d]$. □
On en déduit, en utilisant la proposition 3.15 :

Proposition et définition 4.35 *On peut définir une addition et une multiplication sur $\mathbb{Z}/d\mathbb{Z}$. Ces opérations sont caractérisées par :*
pour tout $n, m \in \mathbb{Z}$, $\overline{n} + \overline{m} = \overline{n+m}$ et $\overline{n} \times \overline{m} = \overline{n \times m}$.

Exemple 4.36 *On peut donner les tables d'addition et de multiplication de $\mathbb{Z}/4\mathbb{Z}$. On utilise la définition précédente, et la division euclidienne par 4 pour représenter tous les éléments de $\mathbb{Z}/4\mathbb{Z}$ comme la classe d'un entier entre 0 et 3 :*

$+$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{0}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{1}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{0}$
$\overline{2}$	$\overline{2}$	$\overline{3}$	$\overline{0}$	$\overline{1}$
$\overline{3}$	$\overline{3}$	$\overline{0}$	$\overline{1}$	$\overline{2}$

$\times$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$
$\overline{1}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{2}$	$\overline{0}$	$\overline{2}$	$\overline{0}$	$\overline{2}$
$\overline{3}$	$\overline{0}$	$\overline{3}$	$\overline{2}$	$\overline{1}$

Les règles de calculs élémentaires de $\mathbb{Z}$ se répercutent pour $\mathbb{Z}/d\mathbb{Z}$.

Fait 4.37 *Pour tous $\overline{m}, \overline{n}, \overline{p}$ dans $\mathbb{Z}/d\mathbb{Z}$:*
$\overline{m} + \overline{n} = \overline{n} + \overline{m}$ *et* $\overline{m} \times \overline{n} = \overline{n} \times \overline{m}$ *(commutativité)*
$\overline{m} + (\overline{n} + \overline{p}) = (\overline{m} + \overline{n}) + \overline{p}$ *et* $\overline{m} \times (\overline{n} \times \overline{p}) = (\overline{m} \times \overline{n}) \times \overline{p}$ *(associativité)*
$\overline{m} \times (\overline{n} + \overline{p}) = \overline{m} \times \overline{n} + \overline{m} \times \overline{p}$ *(distributivité)*

Remarque 4.38 *En utilisant une division euclidienne, on peut représenter tout élément de $\mathbb{Z}/d\mathbb{Z}$ comme la classe d'un entier compris entre 0 et $d-1$. Mais parfois, il peut être plus utile d'écrire cet élément comme la classe d'un entier hors de cet intervalle.*
Par exemple, on cherche à déterminer le chiffre des unités de 2019^{167}. Pour un entier n, le chiffre des unités de n correspond au reste de la division euclidienne de n par 10. Ainsi, $2019 = 201 \times 10 + 9$, ce qui dit aussi que $\overline{2019} = \overline{9}$ dans $\mathbb{Z}/10\mathbb{Z}$. Pour calculer le chiffre des unités de 2019^{167}, on va calculer $\overline{2019^{167}}$ dans $\mathbb{Z}/10\mathbb{Z}$. On a $\overline{2019^{167}} = (\overline{2019})^{167} = (\overline{9})^{167}$. Or il est difficile de calculer $(\overline{9})^{167}$; sachant que $\overline{9} = \overline{-1}$ dans $\mathbb{Z}/10\mathbb{Z}$, on calcule plutôt $(\overline{-1})^{167} = \overline{(-1)^{167}} = \overline{-1} = \overline{9}$. Ainsi, $\overline{2019^{167}} = \overline{9}$, ce qui dit que le chiffre des unités de 2019^{167} est 9.

Chapitre 5

Nombres complexes

5.1 Définitions et premières propriétés

Un nombre complexe z est donné par deux nombres réels : sa partie réelle, notée $\mathrm{Re}(z)$, et sa partie imaginaire, notée $\mathrm{Im}(z)$. Si $\mathrm{Re}(z) = a$ et $\mathrm{Im}(z) = b$, on note $z = a + ib$ (on le voit pour l'instant simplement comme une notation). On note $\mathbb{C}$ l'ensemble des nombres complexes.
On dit que z est réel si $\mathrm{Im}(z) = 0$. On peut ainsi identifier $\mathbb{R}$ à un sous-ensemble de $\mathbb{C}$; plus formellement, l'application suivante est une injection :

$$\begin{array}{ccc} \mathbb{R} & \to & \mathbb{C} \\ a & \mapsto & a + i0 \end{array}$$

On dit que z est imaginaire pur si $\mathrm{Re}(z) = 0$.

Opérations sur $\mathbb{C}$

Soient des nombres complexes $z = a + ib$ et $z' = a' + ib'$.
On définit l'addition par : $z + z' = (a + a') + i(b + b')$.
On définit la mutiplication par $zz' = (aa' - bb') + i(ab' + a'b)$.

Fait 5.1 *Pour tous z, z', z'' dans $\mathbb{C}$:*

1. *$z + z' = z' + z$ et $zz' = z'z$ (commutativité)*
2. *$z + (z' + z'') = (z + z') + z''$ et $z(z'z'') = (zz')z''$ (associativité)*
3. *$z(z' + z'') = zz' + zz''$ (distributivité)*

Remarque 5.2 *On note bien sûr i le nombre complexe $0+i1$. D'après la formule pour la multiplication, il vérifie $i^2 = -1$. Dans la pratique, pour multiplier deux nombres complexes, il suffit d'utiliser les propriétés énoncées dans le fait précédent et le fait que $i^2 = -1$.*

Conjugué d'un nombre complexe

Pour un nombre complexe $z = a + ib$ (on sous-entend toujours que a et b sont réels), le conjugué de z, noté $\overline{z}$, est défini par $\overline{z} = a - ib$.

Proposition 5.3

1. $\overline{\overline{z}} = z$
2. $\overline{z + z'} = \overline{z} + \overline{z}'$
3. $\overline{zz'} = \overline{z}\overline{z}'$

Preuve On note $z = a + ib$ et $z' = a' + ib'$.
$\overline{\overline{z}} = \overline{a - ib} = a - i(-b) = a + ib = z$.
$\overline{z + z'} = \overline{(a + a') + i(b + b')} = a - ib + a' - ib' = \overline{z} + \overline{z}'$.
$\overline{zz'} = \overline{(aa' - bb') + i(ab' + a'b)} = (aa' - bb') - i(ab' + a'b)$ et $\overline{z}\overline{z}' = (a - ib)(a' - ib') = (aa' - bb') - i(ab' + a'b)$.
□

Module et argument d'un nombre complexe

Soit $z = a + ib \in \mathbb{C}$, avec $a, b \in \mathbb{R}$. On calcule $z\overline{z} = (a+ib)(a-ib) = a^2 + b^2$: c'est un nombre réel positif ou nul, dont on peut prendre la racine carrée. On définit le module de z : $|z| = \sqrt{z\overline{z}} = \sqrt{a^2+b^2}$.

Proposition 5.4 *1.* $|z| = 0$ *ssi* $z = 0$
2. $|zz'| = |z||z'|$

Preuve Si $z = 0$, il est clair que $|z| = \sqrt{0 \times \overline{0}} = 0$.
Réciproquement, pour $z = a + ib$, supposons que $|z| = \sqrt{a^2+b^2} = 0$. Alors $a^2 + b^2$ est une somme de deux nombres réels positifs ou nuls, qui est nulle, ce qui oblige à ce que les deux termes soient nuls. Donc $a = b = 0$ et ainsi $z = 0$.
Pour calculer $|zz'|$, on utilise la proposition 5.3 : $|zz'| = \sqrt{zz'\overline{zz'}} = \sqrt{z\overline{z}z'\overline{z'}} = |z||z'|$. □

On rappelle le fait suivant. Noter que l'usage en mathématiques est d'exprimer les angles en radians (la mesure de l'angle plat π, celle de l'angle droit $\pi/2$,...).

Fait 5.5 *Soit* $x, y \in \mathbb{R}$ *tels que* $x^2 + y^2 = 1$. *Alors il existe un unique réel* $\theta \in]-\pi, \pi]$ *tel que* $x = \cos(\theta)$ *et* $y = \sin(\theta)$.

Soit $z = a + ib \in \mathbb{C}$, on suppose que $z \neq 0$, et donc aussi $|z| \neq 0$. On peut toujours diviser un nombre complexe par un nombre réel non nul (il suffit de diviser la partie réelle et la partie imaginaire par ce réel), regardons donc $\frac{z}{|z|} = \frac{a}{\sqrt{a^2+b^2}} + i\frac{b}{\sqrt{a^2+b^2}}$. Utilisons le fait précédent : puisque $(\frac{a}{\sqrt{a^2+b^2}})^2 + (\frac{b}{\sqrt{a^2+b^2}})^2 = \frac{a^2+b^2}{a^2+b^2} = 1$, il existe un unique $\theta \in]-\pi, \pi]$ tel que $\frac{a}{\sqrt{a^2+b^2}} = \cos(\theta)$ et $\frac{b}{\sqrt{a^2+b^2}} = \sin(\theta)$. On appelle θ l'argument de z, on le note $\mathrm{Arg}(z)$.

Remarque 5.6 *La réciproque du fait précédent est également vraie : pour tout* $\theta \in \mathbb{R}$, $\cos^2(\theta) + \sin^2(\theta) = 1$. *Cela signifie aussi que* $|\cos(\theta) + i\sin(\theta)| = \sqrt{\cos^2(\theta) + \sin^2(\theta)} = 1$.

Si $z \in \mathbb{C}$ est non nul, on pose $r = |z| \in \mathbb{R}_+^*$ et $\theta = \mathrm{Arg}(z) \in]-\pi, \pi]$, et on a alors $z = r(\cos(\theta) + i\sin(\theta))$. Cette écriture est unique : si $z = r(\cos(\theta) + i\sin(\theta)) = r'(\cos(\theta') + i\sin(\theta'))$, avec r, r' des réels > 0 et θ, θ' dans $]-\pi, \pi]$, alors $r = r' = |z|$ et $\theta = \theta' = \mathrm{Arg}(z)$.

Remarque 5.7 *Soit* $z \in \mathbb{C}$ *non nul,* $r = |z|$ *et* $\theta = Arg(z)$. *Si* $\theta' \equiv \theta\ [2\pi]$, *c'est-à-dire s'il existe* $k \in \mathbb{Z}$ *tel que* $\theta' = \theta + k \times 2\pi$, *alors* $\cos(\theta') = \cos(\theta)$ *et* $\sin(\theta') = \sin(\theta)$ *car les fonctions cosinus et sinus sont* 2π*-périodiques. On a donc également* $z = r(\cos(\theta') + i\sin(\theta'))$, *et on dit que* θ' *est un argument de* z.

Notation exponentielle

Soit $\theta \in \mathbb{R}$. On note $e^{i\theta} = \cos(\theta) + i\sin(\theta)$. Ainsi, si $z \in \mathbb{C}$ est non nul, avec $r = |z|$ et θ un argument de z, alors $z = re^{i\theta}$: c'est ce qu'on appelle l'écriture trigonométrique de z.
On voit ici cette écriture seulement comme une notation, mais elle a en fait un sens plus profond. On va voir maintenant que cette notation est conforme à l'usage habituel du calcul des puissances, on rappelle pour cela les formules trigonométriques.

Fait 5.8 $\cos(\theta + \theta') = \cos(\theta)\cos(\theta') - \sin(\theta)\sin(\theta')$
$\sin(\theta + \theta') = \cos(\theta)\sin(\theta') + \sin(\theta)\cos(\theta')$

On en déduit, en utilisant aussi les formules de multiplication, que si $z = re^{i\theta}$ et $z' = r'e^{i\theta'}$, alors

$$\begin{aligned} zz' = re^{i\theta}r'e^{i\theta'} &= rr'(\cos(\theta) + i\sin(\theta))(\cos(\theta') + i\sin(\theta')) \\ &= rr'(\cos(\theta)\cos(\theta') - \sin(\theta)\sin(\theta') + i(\cos(\theta)\sin(\theta') + \sin(\theta)\cos(\theta'))) \\ &= rr'(\cos(\theta + \theta') + i\sin(\theta + \theta')) \\ &= rr'e^{i(\theta+\theta')}. \end{aligned}$$

C'est bien la formule attendue pour un produit d'exponentielles. On a en particulier montré le résultat suivant.

Proposition 5.9 *Soient z, z' deux nombres complexes non nuls, θ un argument de z et θ' un argument de z', alors $\theta + \theta'$ est un argument de zz'.*

Exemple 5.10 *En prenant les valeurs des fonctions cosinus et sinus en différentes valeurs simples, on voit que $1 = e^{i\times 0}$, $i = e^{i\pi/2}$, $-1 = e^{i\pi}$ et $-i = e^{-i\pi/2}$.*
Attention : $Arg(-i) = -\pi/2 \in]-\pi, \pi]$, mais $-\pi/2 + (-\pi/2) = -\pi$ est seulement un argument de $(-i)\times(-i) = -1$; pour avoir l'argument de -1 il faut ajouter 2π pour obtenir $Arg(-1) = \pi \in]-\pi, \pi]$.

Remarque 5.11 *Attention : il n'existe pas de formule simple pour $|z + z'|$ et $Arg(z + z')$.*

Division

Soit $z \in \mathbb{C}$, on a vu que $z\overline{z} = |z|^2$. En particulier, si $z \neq 0$, on peut diviser cette égalité par le réel non nul $|z|^2$, et on obtient : $z \times \frac{\overline{z}}{|z|^2} = 1$. On dit alors que $\frac{\overline{z}}{|z|^2}$ est l'inverse de z, qu'on peut noter $\frac{1}{z}$. On peut donc aussi calculer, pour tous $z, z' \in \mathbb{C}$, avec $z' \neq 0$, $\frac{z}{z'} = z \times \frac{1}{z'} = \frac{z\overline{z'}}{|z'|^2}$.

Exemple 5.12

$$\frac{1+i}{1+2i} = \frac{(1+i)(1-2i)}{(1+2i)(1-2i)} = \frac{3}{5} - \frac{1}{5}i.$$

Si $|z| = 1$, alors $\frac{1}{z} = \frac{\overline{z}}{|z|^2} = \overline{z}$.

La méthode précédente permet de calculer le quotient de deux nombres complexes sous la forme algébrique. Si $z = re^{i\theta}$ et $z' = r'e^{i\theta'}$ sont données sous forme trigonométrique, avec $r' \neq 0$, les calculs précédents montrent que

$$r'e^{i\theta'} \times \frac{1}{r'}e^{-i\theta'} = \frac{r'}{r'}e^{i(\theta'-\theta')} = e^{i\times 0} = 1$$

Ainsi, $\frac{1}{z'} = \frac{1}{r'}e^{-i\theta'}$ et $\frac{z}{z'} = \frac{r}{r'}e^{i(\theta-\theta')}$.

5.2 Interprétation géométrique

On se place dans le plan avec un repère orthonormé $(O; \overrightarrow{OI}, \overrightarrow{OJ})$. Soit $z = a + ib \in \mathbb{C}$ avec $a, b \in \mathbb{R}$. On peut lui associer dans le plan le point M de coordonnées (a, b). On dit alors que M est le point d'affixe z, ou encore que z est l'affixe de M.

Fait 5.13 *L'application de $\mathbb{C}$ vers le plan, qui à tout $z \in \mathbb{C}$ associe le point M d'affixe $\mathbb{C}$, est une bijection.*

Exemple 5.14 *L'affixe du point O est 0, l'affixe de I est 1 et l'affixe de J est i.*

Soit M le point d'affixe z. En utilisant le théorème de Pythagore, on voit que la distance OM est égale à $\sqrt{a^2 + b^2}$, c'est-à-dire $OM = |z|$.
Si $z \neq 0$, $M \neq O$ et on peut considérer θ, la mesure de l'angle orienté entre les vecteurs $\overrightarrow{OI}$ et $\overrightarrow{OM}$. Les formules de trigonométrie dans le triangle donnent que $a = OM \times \cos(\theta)$ et $b = OM \times \sin(\theta)$, c'est-à-dire $\frac{z}{|z|} = \cos(\theta) + i\sin(\theta)$, et donc θ est un argument de z.
Le point M' d'affixe $\overline{z}$ est obtenu en prenant l'image de M (d'affixe z) par la symétrie orthogonale par rapport à la droite (OI).
On appelle cercle trigonométrique le cercle de centre O et de rayon 1. C'est l'ensemble des points d'affixe z avec $|z| = 1$; on peut écrire tous ces points sous la forme $z = e^{i\theta}$ avec $\theta \in \mathbb{R}$ (ou seulement $\theta \in]-\pi, \pi]$).

5.3 Quelques équations

Racines carrées

Soit $w \in \mathbb{C}$. On cherche les racines carrées de w, c'est-à-dire les éléments $z \in \mathbb{C}$ tels que $z^2 = w$.

Forme algébrique
On suppose que $w = c + id$ est donné sous forme algébrique et on cherche également $z = a + ib$ sous forme algébrique.
On calcule $z^2 = (a^2 - b^2) + i2ab$; on en déduit que

$$z^2 = w \Leftrightarrow \begin{cases} \mathrm{Re}(z^2) = \mathrm{Re}(w) \\ \mathrm{Im}(z^2) = \mathrm{Im}(w) \end{cases} \Leftrightarrow \begin{cases} a^2 - b^2 = c \\ 2ab = d \end{cases}$$

Pour rendre la résolution de ce système plus facile, on introduit une nouvelle équation : $z^2 = w$ implique que $|z^2| = |z|^2 = |w|$, c'est-à-dire $a^2 + b^2 = |w|$, et donc

$$z^2 = w \Leftrightarrow \begin{cases} a^2 - b^2 = c \ (1) \\ 2ab = d \\ a^2 + b^2 = |w| \ (3) \end{cases} \Leftrightarrow \begin{cases} a^2 = \frac{c+|w|}{2} \\ b^2 = \frac{|w|-c}{2} \\ 2ab = d \end{cases} \Leftrightarrow \begin{cases} a = \pm\sqrt{\frac{c+|w|}{2}} \\ b = \pm\sqrt{\frac{|w|-c}{2}} \\ 2ab = d \end{cases}$$

On a obtenu la valeur de a^2 en faisant la somme des équations (1) et (3), et celle de b^2 en faisant leur différence. L'équation $2ab = d$ est utilisée pour déterminer les signes : si $d > 0$, a et b sont de même signe ; si $d < 0$, a et b sont de signes différents ; et si $d = 0$, $a = 0$ ou $b = 0$.

Exemple 5.15 *Racines carrées de $w = 3 - 4i$. En reprenant les calculs précédents, avec $c = 3$, $d = -4$ et $|w| = \sqrt{3^2 + (-4)^2} = 5$, on trouve*

$$(a + ib)^2 = 3 - 4i \Leftrightarrow \begin{cases} a = \pm 2 \\ b = \pm 1 \\ 2ab = -4 \end{cases}$$

Comme on doit avoir $2ab = -4$, a et b sont de signes différents, ce qui donne deux racines carrées : $z_1 = 2 - i$ et $z_2 = -2 + i = -z_1$.

Forme trigonométrique
On suppose que $w = Re^{i\phi}$ est donné sous forme trigonométrique et on cherche également $z = re^{i\theta}$ sous forme trigonométrique.

Lemme 5.16 *Soit d et T des réels, avec $d \neq 0$. Alors pour tous x, y dans $\mathbb{R}$, $dx \equiv dy\ [T]$ ssi $x \equiv y\ [T/d]$.*

Preuve Si $dx \equiv dy\ [T]$, il existe $k \in \mathbb{Z}$ tel que $dx = dy + kT$, et donc $x = y + \frac{kT}{d}$ en divisant par d (qui est non nul), ce qui dit que $x \equiv y\ [T/d]$. L'implication réciproque se montre de la même façon, cette fois en multipliant par d. □
Par choix, r et R sont sictement positifs, et alors :

$$z^2 = w \Leftrightarrow r^2 e^{2i\theta} = Re^{i\phi} \Leftrightarrow \begin{cases} r^2 = R \\ 2\theta \equiv \phi\ [2\pi] \end{cases} \Leftrightarrow \begin{cases} r = \sqrt{R} \\ \theta \equiv \frac{\phi}{2}\ [\pi] \end{cases}$$

Cela donne bien deux racines carrées pour w : en écrivant $\theta = \frac{\phi}{2} + k\pi$, si $k = 2k'$ est pair, alors $\theta \equiv \frac{\phi}{2}\ [2\pi]$, ce qui donne $z = \sqrt{R}e^{i\phi/2}$, et si $k = 2k' + 1$ est impair, alors $\theta = \frac{\phi}{2} + \pi + 2k'\pi \equiv \frac{\phi}{2} + \pi\ [2\pi]$, ce qui donne $z = \sqrt{R}e^{i(\phi/2+\pi)}$. Puisque $e^{i\pi} = -1$, on peut aussi écrire ces deux solutions $z = \pm\sqrt{R}e^{i\phi/2}$.

Application : calcul de certaines valeurs de cosinus et sinus
Pour certains nombres complexes dont on connait à la fois la forme algébrique et la forme trigonométrique, on peut utiliser les deux méthodes et comparer les résultats.

Exemple 5.17 *Cherchons les racines carrées de i.*

On cherche $z = a + ib$ tel que $z^2 = i$, ce qui donne les équations $\begin{cases} a^2 - b^2 = 0 \\ 2ab = 1 \\ a^2 + b^2 = 1 \end{cases}$ *, ce qui équivaut à* $\begin{cases} a = \pm\frac{\sqrt{2}}{2} \\ b = \pm\frac{\sqrt{2}}{2} \\ 2ab = 1 \end{cases}$ *. D'après la dernière équation, a et b sont de même signe, ce qui donne les racines carrées $z = \pm(\frac{\sqrt{2}}{2} + i\frac{\sqrt{2}}{2})$.*

D'autre part, i s'écrit sous forme géométrique $i = e^{i\pi/2}$. On cherche $z = re^{i\theta}$ (avec $r > 0$) tel que $z^2 = e^{i\pi/2}$, ce qui donne les équations $\begin{cases} r^2 = 1 \\ 2\theta \equiv \pi/2\ [2\pi] \end{cases}$ *et donc $z = \pm e^{i\pi/4}$.*

On identifie les deux formes des racines carrées de i. Comme $\cos(\pi/4) > 0$ et $\sin(\pi/4) > 0$, on doit avoir $e^{i\pi/4} = \frac{\sqrt{2}}{2} + i\frac{\sqrt{2}}{2}$. On en déduit que $\cos(\pi/4) = \sin(\pi/4) = \frac{\sqrt{2}}{2}$.

Équations du second degré

On cherche les solutions de l'équation $\alpha z^2 + \beta z + \gamma = 0$, où α, β, γ sont des coefficients dans $\mathbb{C}$, avec $\alpha \neq 0$. Comme pour les équations du second degré à coefficients réels, on peut la mettre sous forme canonique :

$$\alpha z^2 + \beta z + \gamma = \alpha((z + \frac{\beta}{2\alpha})^2 - \frac{\beta^2 - 4\alpha\gamma}{4\alpha^2}).$$

On note encore le discriminant $\Delta = \beta^2 - 4\alpha\gamma$. C'est un élément de $\mathbb{C}$, on a vu précédemment qu'on peut trouver $\delta \in \mathbb{C}$ tel que $\delta^2 = \Delta$. On en déduit la factorisation

$$\alpha z^2 + \beta z + \gamma = \alpha((z + \frac{\beta}{2\alpha})^2 - (\frac{\delta}{2\alpha})^2) = \alpha(z - \frac{-\beta + \delta}{2\alpha})(z - \frac{-\beta - \delta}{2\alpha})$$

et donc les solutions de $\alpha z^2 + \beta z + \gamma = 0$ sont $z = \frac{-\beta \pm \delta}{2\alpha}$.

Racines n-ièmes de l'unité

Soit un entier $n \geq 2$. On cherche dans $\mathbb{C}$ les solutions de l'équation $z^n = 1$; on les appelle les racines n-ièmes de l'unité. On cherche $z = re^{i\theta}$ sous forme trigonométrique ; $z^n = 1$ équivaut à $\begin{cases} r^n = 1 \\ n\theta \equiv 0\ [2\pi] \end{cases}$.

Comme $r > 0$, $r^n = 1$ équivaut à $r = 1$; d'après le lemme 5.16, $n\theta \equiv 0\ [2\pi]$ équivaut à $\theta \equiv 0\ [2\pi/n]$, et il existe donc $k \in \mathbb{Z}$ tel que $\theta = \frac{2k\pi}{n}$. En écrivant la division euclidienne de k par n, $k = qn + s$ avec $0 \leq s \leq n - 1$, on trouve que $\theta = \frac{2s\pi}{n} + q2\pi$, avec $q \in \mathbb{Z}$, c'est-à-dire $\theta \equiv \frac{2s\pi}{n}\ [2\pi]$. On trouve donc les n racines n-ièmes de l'unité : ce sont les $e^{i2s\pi/n}$, pour s entier variant entre 0 et $n - 1$. Géométriquement, ces racines n-ièmes de l'unité sont situées sur le cercle trigonométrique ; pour $s = 0$, on obtient le point d'affixe 1, et les racines suivantes sont obtenues en appliquant succesivement une rotation d'angle $2\pi/n$. Ces n points forment un polygône régulier à n côtés.

Chapitre 6

Polynômes

On va s'intéresser dans ce chapitre aux polynômes, à coefficients dans $\mathbb{R}$ ou dans $\mathbb{C}$. Comme les définitions et la plupart des résultats sont les mêmes pour $\mathbb{R}$ et $\mathbb{C}$, on écrira $\mathbb{K}$ pour désigner soit $\mathbb{R}$, soit $\mathbb{C}$.

6.1 Définitions et premières propriétés

On se donne un nom de variable X. Un polynôme en la variable X, à coeffcients dans $\mathbb{K}$, est donné par une suite de coefficients $(a_n)_{n\in\mathbb{N}}$ dans $\mathbb{K}$, nuls à partir d'un certain rang, c'est-à-dire $\exists N \in \mathbb{N}\ \forall n \geq N\ a_n = 0$. On note un tel polynôme $P = a_N X^N + a_{N-1}X^{N-1} + \ldots + a_1 X + a_0$, ou encore $P = \sum_{n=0}^{N} a_n X^n$. On écrira parfois aussi $P(X)$ au lieu de P. On appelle polynôme nul le polynôme dont tous les coefficients sont nuls.
On note $\mathbb{K}[X]$ l'ensemble des polynômes en X à coefficients dans $\mathbb{K}$. Comme $\mathbb{R} \subset \mathbb{C}$, $\mathbb{R}[X] \subset \mathbb{C}[X]$.

Définition 6.1 *Soit $P \in \mathbb{K}[X]$ un polynôme non nul, et $(a_n)_{n\in\mathbb{N}}$ ses coefficients. On définit le degré de P, noté $\deg(P)$, comme le plus grand entier N tel que $a_N \neq 0$.*
Par convention, si $P = 0$, le polynôme nul, alors $\deg(P) = -\infty$.

Exemple 6.2 *Soit $P(X) = aX^2 + 3X - 1$. Si $a \neq 0$, alors $\deg(P) = 2$. Si $a = 0$, alors $\deg(P) = 1$.*

On définit maintenant la somme et le produit de deux polynômes. Soient $P(X) = a_N X^N + \ldots + a_1 X + a_0$ et $Q(X) = b_N X^N + \ldots + b_1 X + b_0$ (on peut toujours écrire des expressions de ce type avec le même N pour P et Q, quitte à remplir par des coefficients nuls). Alors on définit :

1. $$P(X) + Q(X) = \sum_{n=0}^{N} (a_n + b_n) X^n$$
2. $$P(X)Q(X) = \sum_{n=0}^{2N} c_n X^n \text{ avec } c_n = \sum_{i=0}^{n} a_i b_{n-i}$$

Fait 6.3 *La somme et le produit de polynômes vérifient les propriétés suivantes :*

1. *$P(X) + Q(X) = Q(X) + P(X)$ et $P(X)Q(X) = Q(X)P(X)$ (commutativité)*
2. *$(P(X) + Q(X)) + R(X) = P(X) + (Q(X) + R(X))$ et $(P(X)Q(X))R(X) = P(X)(Q(X)R(X))$ (associativité)*
3. *$P(X)(Q(X) + R(X)) = P(X)Q(X) + P(X)R(X)$ (distributivité)*

Remarque 6.4 *En pratique, plutôt que de calculer un produit à l'aide de la règle de la définition, on utilise les propriétés ci-dessus qui permettent de développer un produit selon les règles habituelles (avec bien sûr $a_n X^n b_m X^m = a_n b_m X^{n+m}$). La règle donnée dans la définition peut être utile si on n'a besoin de connaitre qu'un des coefficients du produit. Par exemple, si on fait le produit $(X^5 + 3X^4 - X^3 + 2X^2 + 5X - 2)(7X^2 + 2X - 4)$, le coefficient de X^6 sera $1 \times 2 + 3 \times 7 = 23$.*

Pour le résultat suivant, on utilise les conventions habituelles pour $-\infty$: si $x \in \mathbb{N}$ ou $x = -\infty$, alors $x + (-\infty) = -\infty$ et $\max(x, -\infty) = x$.

Proposition 6.5 *Soient P et Q dans $\mathbb{K}[X]$. Alors :*

1. $\deg(P+Q) \leq \max(\deg(P), \deg(Q))$ *; si* $\deg(P) \neq \deg(Q)$*, alors* $\deg(P+Q) = \max(\deg(P), \deg(Q))$
2. $\deg(PQ) = \deg(P) + \deg(Q)$

Preuve Les cas où P ou Q sont nuls sont faciles à vérifier, car $P + 0 = P$ et $P \times 0 = 0$. On suppose donc maintenant que $P(X)$, de coefficients $(a_n)_{n\in\mathbb{N}}$, et $Q(X)$, de coefficients $(b_n)_{n\in\mathbb{N}}$, sont non nuls.
Soit $N = \max(\deg(P), \deg(Q))$. Si $n > N$, alors $n > \deg(P)$, donc $a_n = 0$, et $n > \deg(Q)$, donc $b_n = 0$, et on obtient donc $a_n + b_n = 0$. Cela dit que le degré de $P + Q$ est $\leq N$. Si de plus $\deg(P) \neq \deg(Q)$, on peut supposer que $N = \deg(P) > \deg(Q)$ quitte à échanger P et Q. Comme $N > \deg(Q)$, $b_N = 0$, et donc $a_N + b_N = a_N \neq 0$, donc $N = \deg(P + Q)$.
Pour le produit, écrivons $P(X)Q(X) = \sum_{n=0}^{2N} c_n X^n$ avec $c_n = \sum_{i=0}^{n} a_i b_{n-i}$ (ici N est n'importe quel entier $\geq \max(\deg(P), \deg(Q))$). Regardons c_n pour $n > \deg(P)+\deg(Q)$. Dans la somme $c_n = \sum_{i=0}^{n} a_i b_{n-i}$, si $i > \deg(P)$ alors $a_i = 0$ et donc $a_i b_{n-i} = 0$, et si $i \leq \deg(P)$ alors $n-i > \deg(P)+\deg(Q)-i > \deg(Q)$ donc $b_{n-i} = 0$ et donc aussi $a_i b_{n-i} = 0$. Ainsi c_n est une somme de termes nuls, donc $c_n = 0$. Cela dit que $\deg(PQ) \leq \deg(P) + \deg(Q)$. De plus, pour $n = \deg(P) + \deg(Q)$, on voit par un raisonnement semblable au précédent que dans la somme $c_n = \sum_{i=0}^{n} a_i b_{n-i}$, le seul terme non nul est obtenu pour $i = \deg(P)$ et donc $n - i = \deg(P) + \deg(Q) - \deg(P) = \deg(Q)$, puisque dans ce cas $a_i \neq 0$ et $b_{n-i} \neq 0$ par définition du degré (les autres termes sont nuls car soit $a_i = 0$, soit $b_{n-i} = 0$). On a donc trouvé $c_{\deg(P)+\deg(Q)} = a_{\deg(P)} b_{\deg(Q)} \neq 0$, ce qui justifie que le degré de PQ est $\deg(P) + \deg(Q)$. □

Remarque 6.6 *Si P est un polynôme non nul de degré n, le coefficient devant X^n s'appelle le coefficient dominant de P. La preuve précédente nous dit que le coefficient dominant de PQ est le produit du coefficient dominant de P et de celui de Q.*

Exemple 6.7

1. $P(X) = 4X^3 - 2X + 1$ *et* $Q(X) = 5X^2 - X + 2$ *;* $\deg(P) = 3$ *et* $\deg(Q) = 2$.
   $P(X) + Q(X) = 4X^3 + 5X^2 - 3X + 3$ *;* $\deg(P + Q) = 3 = \max(\deg(P), \deg(Q))$.
   $P(X)Q(X) = 20X^5 - 4X^4 - 2X^3 + 7X^2 - 5X + 2$ *;* $\deg(PQ) = 5 = \deg(P) + \deg(Q)$*, le coefficient dominant de PQ est* $20 = 4 \times 5$.
2. $P(X) = 2X^2 - 3X + 1$ *et* $Q(X) = -2X^2 + X - 4$*, tous les deux de degré* 2.
   $P(X) + Q(X) = -2X - 3$ *;* $\deg(P + Q) = 1 \leq 2 = \max(\deg(P), \deg(Q))$.

Définition 6.8 *Soit un polynôme $P(X) = \sum_{n=0}^{N} a_n X^n$. On définit sa dérivée, notée P', par*

$$P'(X) = \sum_{n=1}^{N} n a_n X^{n-1}.$$

Remarque 6.9

1. *Cette définition est pour l'instant uniquement formelle, c'est une manipulation sur les coefficients du polynôme qui n'utilise pas la notion de dérivée au sens de l'analyse.*
2. *Si* $\deg(P) \leq 0$*, on dit que P est un polynôme constant. Dans ce cas, on peut écrire $P = a_0$ et alors $P' = 0 \times a_0 = 0$.*

Proposition 6.10 *Soient deux polynômes P et Q.*

1. *Si P est de degré $N \geq 1$, de coefficient dominant a_N, alors P' est de degré $N - 1$, de coefficient dominant Na_N.*
2. $(P + Q)' = P' + Q'$
3. $(PQ)' = PQ' + P'Q$

Preuve Si $N = \deg(P) \geq 1$, on peut écrire $P(X) = a_N X^N + \ldots + a_1 X + a_0$ avec $a_N \neq 0$, et alors $P'(X) = Na_N X^{N-1} + \ldots + a_1$ avec $Na_N \neq 0$, donc $\deg(P) = N - 1$ et le coefficient dominant de P' est Na_N.
La propriété $(P + Q)' = P' + Q'$ se vérifie immédiatement.

Pour le produit, écrivons $P(X) = \sum_{n=0}^{N} a_n X^n$ et $Q(X) = \sum_{n=0}^{N} b_n X^n$. Alors $P(X)Q(X) = \sum_{n=0}^{2N} c_n X^n$ avec $c_n = \sum_{j=0}^{n} a_j b_{n-j}$. Et donc $(PQ)'(X) = \sum_{n=1}^{2N} n c_n X^{n-1}$: en faisant le changement de variable $m = n-1$, on voit que le coefficient de X^m dans $(PQ)'$ est $(m+1)c_{m+1} = (m+1)\sum_{j=0}^{m+1} a_j b_{m+1-j}$.
D'autre part, faisons le produit $P'Q$: le coefficent de X^i dans P' est $(i+1)a_{i+1}$ donc le coefficient de X^m dans $P'Q$ est $\sum_{i=0}^{m}(i+1)a_{i+1}b_{m-i}$. Par le même argument, le coefficient de X^m dans PQ' est $\sum_{i=0}^{m} a_i(m-i+1)b_{m-i+1}$. Donc le coefficient de X^m dans $P'Q + PQ'$ est

$$\begin{aligned}\sum_{i=0}^{m}(i+1)a_{i+1}b_{m-i} + \sum_{i=0}^{m} a_i(m-i+1)b_{m-i+1} &= \underbrace{\sum_{j=1}^{m+1} j a_j b_{m-j+1}}_{j=i+1} + \underbrace{\sum_{j=0}^{m} a_j(m-j+1)b_{m-j+1}}_{j=i} \\ &= (m+1)a_{m+1}b_0 + \Big(\sum_{j=1}^{m}(j+m-j+1)a_j b_{m-j+1}\Big) + (m+1)a_0 b_{m+1} \\ &= (m+1)\sum_{j=0}^{m+1} a_j b_{m+1-j}\end{aligned}$$

On retrouve bien le coefficient de X^m dans $(PQ)'$, donc $(PQ)' = P'Q + PQ'$. □

Définition 6.11 *On définit les derivées successives de P par récurrence sur n :*
$P^{(0)} = P$ et pour tout $n \in \mathbb{N}$, $P^{(n+1)} = (P^{(n)})'$.
On appelle $P^{(n)}$ la dérivée n-ième de P.
Pour les petites valeurs de n, on utilise plutôt les notations $P' = P^{(1)}$, $P'' = P^{(2)}$, ...

Remarque 6.12 *Soit P un polynôme de degré $N \geq 1$, de coefficient dominant a_N. En appliquant la propriété 1 de la proposition 6.10 N fois successivement, on obtient que $P^{(N)}$ est le polynôme constant $N!a_N$. Et donc $P^{(N+1)} = 0$.*

Exemple 6.13 $P(X) = 3X^2 - 4X + 3$; $P'(X) = 6X - 4$; $P''(X) = P^{(2)}(X) = 6 = 2! \times 3$; $P^{(3)}(X) = 0$.

6.2 Racines d'un polynôme

Soit $P(X) = \sum_{n=0}^{N} a_n X^n$ un élément de $\mathbb{K}[X]$. Alors P permet de définir une application de $\mathbb{K}$ dans $\mathbb{K}$:

$$\begin{array}{ccc}\mathbb{K} & \to & \mathbb{K} \\ x & \mapsto & \sum_{n=0}^{N} a_n x^n\end{array}$$

On l'appelle la fonction polynomiale associée à P ; par simplicité, on la note encore P.
Comme $\mathbb{R}[X] \subset \mathbb{C}[X]$, si $P \in \mathbb{R}[X]$, on a aussi $P \in \mathbb{C}[X]$: P permet donc à la fois de définir une fonction polynomiale de $\mathbb{R}$ dans $\mathbb{R}$ et aussi de $\mathbb{C}$ dans $\mathbb{C}$.

Remarque 6.14 *Les opérations de somme, produit et dérivée sur les polynômes correspondent aux mêmes opérations sur les fonctions polynomiales associées.*

Définition 6.15 *Soit $P \in \mathbb{K}[X]$ et $a \in \mathbb{K}$. On dit que a est une racine de P (on précise parfois : dans $\mathbb{K}$) si $P(a) = 0$.*

Exemple 6.16
1. *Soit $P = a_0$ un polynôme constant ; la fonction polynomiale associée est la fonction constante égale à a_0. Si $a_0 \neq 0$, P n'admet pas de racine ; si $a_0 = 0$, tout élément $a \in \mathbb{K}$ est racine de P.*
2. *Soit $P(X) = aX + b \in \mathbb{K}[X]$ un polynôme de degré 1 (avec donc $a \neq 0$). Alors la seule racine de P dans $\mathbb{K}$ est $-\frac{b}{a}$.*

On a vu dans le chapitre précédent que toute équation du second degré à coefficients dans $\mathbb{C}$ admet au moins une solution dans $\mathbb{C}$, ce qui signifie exactement que tout polynôme $P \in \mathbb{C}[X]$ de degré 2 admet au moins une racine. On a en fait un résultat beaucoup plus général, que l'on admet sans démonstration.

Théorème 6.17 - Théorème de d'Alembert-Gauss, Théorème fondamental de l'algèbre
Tout polynôme $P \in \mathbb{C}[X]$ non constant admet au moins une racine dans $\mathbb{C}$.

La situation est différente pour les polynômes à coefficients réels. Le fait suivant est bien connu.

Fait 6.18 *Soit $P(X) = aX^2 + bX + c \in \mathbb{R}[X]$ un polynôme de degré 2, et $\Delta = b^2 - 4ac$ son discriminant. Alors P admet une racine dans $\mathbb{R}$ ssi $\Delta \geq 0$.*

Exemple 6.19 *$P(X) = X^2 + 1$ n'admet pas de racine réelle car son discriminant est -4.*

Proposition 6.20 *Soit $P(X) = a_N X^N + \ldots + a_0 \in \mathbb{R}[X]$ un poynôme de degré N impair. Alors P admet une racine réelle.*

Preuve On va utiliser des notions d'analyse. On peut supposer que $a_N > 0$ (si ce n'est pas le cas, il suffit de considérer $-P$, sachant que P et $-P$ ont exactement les mêmes racines). Alors la limite de $P(x)$ quand x tend vers $+\infty$ est $+\infty$, et celle en $-\infty$ est $-\infty$. En particulier, on peut trouver deux réels x_0 et x_1 tels que $P(x_0) < 0$ et $P(x_1) > 0$. D'autre part, les fonctions polynomiales sont continues, donc en appliquant le théorème des valeurs intermédiaires, on peut trouver x compris entre x_0 et x_1 tel que $P(x) = 0$: x est une racine de P. □

Définition 6.21 *Soit $P \in \mathbb{K}[X]$ et $a \in \mathbb{K}$. On dit que a est une racine de P de multiplicité m si $P(a) = P'(a) = \ldots = P^{(m-1)}(a) = 0$ et $P^{(m)}(a) \neq 0$.*
On dira aussi que a est une racine simple si a est une racine de multiplicité 1 (c'est-à-dire $P(a) = 0$ et $P'(a) \neq 0$), a est une racine double si elle est de multiplicité 2,...

Exemple 6.22
1. *Soit $P(X) = X^2 - 3X + 2$; 1 est une racine de P car $P(1) = 0$, $P'(X) = 2X - 3$ et $P'(1) \neq 0$ donc 1 est une racine simple de P.*
2. *Soit $P(X) = X^2 + 2X + 1$, $P'(X) = 2X + 2$, $P''(X) = 2$; $P(-1) = P'(-1) = 0$ et $P''(-1) \neq 0$ donc -1 est une racine double de P.*

Remarque 6.23 *Soit P un polynôme non nul, de degré N, de coefficient dominant a_N. On a vu que $P^{(N)}$ est le polynôme constant $N!a_N \neq 0$; en particulier, pour toute racine a de P, on pourra trouver un entier n tel que $P^{(n)}(a) \neq 0$.*

Proposition 6.24 *Soit $P(X) = a_N X^N + \ldots + a_1 X + a_0 \in \mathbb{R}[X]$ un polynôme à coefficients réels et $z \in \mathbb{C}$ une racine de P. Alors le conjugué $\overline{z}$ est aussi une racine de P, de même multiplicité que z.*

Preuve Prenons le conjugué de l'égalité $P(z) = a_N z^N + \ldots + a_1 z + a_0 = 0$. En utilisant les règles de calcul pour le conjugué, on obtient $\overline{a_N}\,\overline{z}^N + \ldots + \overline{a_1}\,\overline{z} + \overline{a_0} = \overline{0} = 0$. Comme tous les coefficients a_n sont réels, ils vérifient $\overline{a_n} = a_n$, et l'égalité précédente devient donc $a_N\overline{z}^N + \ldots + a_1\overline{z} + a_0 = 0$, c'est-à-dire $P(\overline{z}) = 0$.
Supposons maintenant que z est une racine de multiplicité m de P. Alors $P(z) = P'(z) = \ldots = P^{(m-1)}(z) = 0$ et $P^{(m)}(z) \neq 0$. En appliquant ce qu'on vient de montrer aux polynômes $P, P', \ldots, P^{(m-1)}$, qui sont des polynômes à coefficients réels, on trouve que $P(\overline{z}) = P'(\overline{z}) = \ldots = P^{(m-1)}(\overline{z}) = 0$. De plus, $P^{(m)}(\overline{z}) \neq 0$: en effet, si on suppose par l'absurde que $P^{(m)}(\overline{z}) = 0$, en appliquant encore une fois le premier résultat on obtient que $P^{(m)}(\overline{\overline{z}}) = 0$, c'est-à-dire $P^{(m)}(z) = 0$, ce qui est une contradiction. On a donc montré que $\overline{z}$ est une racine de P de multiplicité m. □

6.3 Factorisation des polynômes

Un certain nombre de notions et résultats d'arithmétique ont des analogues pour les polynômes. Nous ne donnerons pas toutes les démonstrations, certaines ne sont que des adaptations des démonstrations de l'arithmétique.

Définition 6.25 *Soient P et Q dans $\mathbb{K}[X]$. On dit que P divise Q, et on note $P|Q$, s'il existe $U \in \mathbb{K}[X]$ tel que $Q = PU$.*

Fait 6.26 *La relation $P|Q$ est réflexive et transitive.*
Si $P|Q$ et $P|R$, alors $P|Q+R$.

Proposition et définition 6.27 *Soient P et Q deux polynômes non nuls ; ($P|Q$ et $Q|P$) ssi il existe une constante non nulle $\lambda \in \mathbb{K}$ (c'est-à-dire un polynôme de degré 0) telle que $Q = \lambda P$.*
On dit dans ce cas que P et Q sont associés.

Preuve On suppose que $P|Q$ et $Q|P$. Comme $P|Q$, il existe $U \in \mathbb{K}[X]$ tel que $Q = PU$. Notons que $U \neq 0$ car $Q \neq 0$. D'après les propriétés du degré, $\deg(Q) = \deg(P) + \deg(U)$, avec $\deg(U) \geq 0$ car $U \neq 0$, donc $\deg(Q) \geq \deg(P)$. On montre de la même manière, en utilisant cette fois le fait que $Q|P$, que $\deg(P) \geq \deg(Q)$. On peut donc conclure que $\deg(Q) = \deg(P)$; cela implique donc que $\deg(U) = 0$, donc U est une constante non nulle λ, et $Q = \lambda P$.
Montrons maintenant la réciproque. S'il existe une constante non nulle $\lambda \in \mathbb{K}$ telle que $Q = \lambda P$, alors $P|Q$ par définition, et on obtient aussi que $P = \frac{1}{\lambda}Q$, donc $Q|P$. □

Définition 6.28 *Soit P un polynôme dans $\mathbb{K}[X]$ de degré ≥ 1. On dit que P est irréductible dans $\mathbb{K}[X]$ si pour tous P_1 et P_2 dans $\mathbb{K}[X]$ tels que $P = P_1P_2$, P_1 ou P_2 est associé à P.*

Exemple 6.29 *Si P est de degré 1, alors il est irréductible. En effet, si on a une factorisation $P = P_1P_2$, alors $\deg(P_1)+\deg(P_2) = \deg(P) = 1$, ce qui oblige à ce que $\deg(P_1) = 0$ et $\deg(P_2) = 1$ (ou vice-versa). Dans ce cas, P_1 est une constante non nulle et donc P est associé à P_2.*

Proposition et définition 6.30 - Division euclidienne
Soient P_1 et P_2 deux éléments de $\mathbb{K}[X]$, avec $P_2 \neq 0$. Alors il existe Q et R dans $\mathbb{K}[X]$, uniquement déterminés, tels que $P_1 = P_2Q + R$ avec $\deg(R) < \deg(P_2)$.
On dit que Q est le quotient de la division euclidienne de P_1 par P_2, et R son reste.

Preuve On se donne le polynôme P_2 non nul, de degré d et de coefficient dominant a_d. On va montrer l'existence de la division euclidienne par récurrence sur le degré de P_1 : on suppose qu'on sait faire la division euclidienne de P_3 par P_2 pour tout polynôme P_3 de degré $< n$ et on va montrer qu'on peut faire la division euclidienne de P_1 par P_2 pour tout P_1 de degré n.
Si $n < \deg(P_2)$, il n'y a rien à faire : $P_1 = P_2 \times 0 + P_1$ avec $\deg(P_1) = n < \deg(P_2)$: 0 est le quotient et P_1 le reste. Cela joue le rôle d'étape d'initialisation.
Supposons maintenant que $\deg(P_1) = n \geq d = \deg(P_2)$. Notons b_n le coefficient dominant de P_1. Alors $(\frac{b_n}{a_d}X^{n-d})P_2$ est un polynôme de degré n, et de coefficient dominant $\frac{b_n}{a_d}a_d = b_n$. En faisant la différence $P_3 = P_1 - \frac{b_n}{a_d}X^{n-d}P_2$, les coefficients de X^n se simplifient et on obtient un polynôme P_3 de degré $< n$. On peut alors appliquer l'hypothèse de récurrence : il existe Q et R dans $\mathbb{K}[X]$, avec $\deg(R) < d$ tels que $P_3 = QP_2 + R$. En revenant à la définition de P_3, on trouve que

$$P_1 = P_3 + \frac{b_n}{a_d}X^{n-d}P_2 = (Q + \frac{b_n}{a_d}X^{n-d})P_2 + R,$$

c'est ce qu'on voulait, avec $Q + \frac{b_n}{a_d}X^{n-d}$ comme quotient et R, de degré $< d$, comme reste.
Reste à voir l'unicité de Q et de R. Si $P_1 = Q_1P_2 + R_1 = Q_2P_2 + R_2$, avec R_1 et R_2 de degré $< \deg(P_2)$, on peut écrire que $R_2 - R_1 = (Q_1 - Q_2)P_2$. Donc P_2 divise $R_2 - R_1$, avec $\deg(R_2 - R_1) \leq \max(\deg(R_1), \deg(R_2)) < \deg(P_2)$: c'est impossible, sauf si $R_2 - R_1 = 0$. On a donc montré que $R_1 = R_2$, ce qui donne $(Q_1 - Q_2)P_2 = 0$, et donc $Q_2 - Q_1 = 0$ (on peut le justifier encore une fois avec les degrés).

□
La preuve donne aussi un algorithme pour faire la division euclidienne de P_1 par P_2. Si $\deg(P_1) < \deg(P_2)$, il n'y a rien à faire. Sinon, on élimine le terme de plus haut degré de P_1 en lui retranchant un polynôme de la forme $\frac{b_n}{a_d} X^{n-d} P_2$; et on recommence jusqu'à obtenir un reste dont le degré est $< \deg(P_2)$. On peut poser cette division euclidienne comme on le fait pour les entiers.

Exemple 6.31 *On veut faire la division euclidienne de* $P_1(X) = X^5 - 3X^4 + X^3 + 2X^2 + 5X - 1$ *par* $P_2(X) = X^3 + 2X + 1$ *:*

$$\begin{array}{rrrrrrrrrrrr|l}
 & X^5 & - & 3X^4 & + & X^3 & + & 2X^2 & + & 5X & - & 1 & X^3+2X+1 \\
- & (X^5 & & & + & 2X^3 & + & X^2) & & & & & \hline X^2-3X-1 \\
\hline
 & & - & 3X^4 & - & X^3 & + & X^2 & + & 5X & - & 1 & \\
 & - & (- & 3X^4 & & & - & 6X^2 & - & 3X) & & & \\
\hline
 & & & & - & X^3 & + & 7X^2 & + & 8X & - & 1 & \\
 & & & - & (- & X^3 & & & - & 2X & - & 1) & \\
\hline
 & & & & & & & 7X^2 & + & 10X & & & \\
\end{array}$$

Ce qui nous donne : $X^5 - 3X^4 + X^3 + 2X^2 + 5X - 1 = (X^3 + 2X + 1)(X^2 - 3X - 1) + 7X^2 + 10X$*,* $X^2 - 3X - 1$ *est le quotient et* $7X^2 + 10X$ *est le reste.*

Remarque 6.32 *Soient* P_1 *et* P_2 *dans* $\mathbb{K}[X]$*, avec* $P_2 \neq 0$*. Alors* $P_2 | P_1$ *ssi le reste de la division euclidienne de* P_1 *par* P_2 *est nul.*

Lemme 6.33 *Soient* $P \in \mathbb{K}[X]$*,* $a \in \mathbb{K}$*, et* $P = (X - a)Q + R$ *la division euclidienne de* P *par le polynôme* $X - a$*. Alors* R *est un polynôme constant, de valeur* $P(a)$*.*

Preuve Par définition de la division euclidienne, on doit avoir $\deg(R) < \deg(X - a) = 1$, donc R est un polynôme constant. Pour avoir sa valeur, on peut l'évaluer en n'importe quel point. Or, en prenant la valeur en a dans l'égalité $P = (X - a)Q + R$, on trouve $P(a) = 0 \times Q(a) + R(a) = R(a)$. □

Théorème 6.34 *Soient* $P \in \mathbb{K}[X]$ *et* $a \in \mathbb{K}$*. Alors* $(X - a)$ *divise* P *ssi* $P(a) = 0$*.*

Preuve Si $(X - a)|P$, on peut écrire $P(X) = (X - a)U(X)$ pour un certain polynôme $U(X)$, et donc $P(a) = 0 \times U(a) = 0$.
Réciproquement, supposons que $P(a) = 0$. D'après le lemme 6.33, le reste de la division euclidienne de P par $X - a$ est le polynôme constant nul, et donc, par la remarque 6.32, $X - a$ divise P. □

Corollaire 6.35 *Soient* $P \in \mathbb{K}[X]$*,* $a \in \mathbb{K}$ *et un entier* $m \geq 1$*. Alors* $(X - a)^m$ *divise* P *ssi* a *est une racine de* P *de multiplicité* $\geq m$*.*
On en déduit aussi que a *est une racine de* P *de multiplicité exactement* m *ssi* $(X - a)^m$ *divise* P *mais* $(X - a)^{m+1}$ *ne divise pas* P*.*

Preuve On ne donne pas tous les détails de la démonstration.
Pour $m = 1$, c'est exactement le théorème 6.34.
Considérons le cas $m = 2$. Si $(X - a)^2$ divise P, on écrit $P(X) = U(X)(X - a)^2$. On a déjà vu que $P(a) = 0$, et on calcule $P'(X) = U'(X)(X - a)^2 + U(X) \times 2(X - a)$, donc $P'(a) = 0$. Cela signifie que a est une racine au moins double de P. Réciproquement, supposons que a est une racine au moins double de P. Comme a est une racine de P, on sait déjà qu'on peut écrire $P(X) = (X - a)U(X)$ d'après le théorème 6.34. En dérivant, on obtient $P'(X) = U(X) + (X - a)U'(X)$, et en prenant la valeur en a, $P'(a) = U(a) + 0$. Or $P'(a) = 0$ puisque a est une racine au moins double de P, donc $U(a) = 0$, et en appliquant le théorème 6.34, on sait qu'on peut écrire $U(X) = (X - a)V(X)$. Donc $P(X) = (X - a)^2 V(X)$.
Le cas général s'obtient en répétant cet argument. □

Exemple 6.36 *Soit $P(X) = aX^2 + bX + c \in \mathbb{R}[X]$ un polynôme de degré 2, et soit $\Delta = b^2 - 4ac$ son discriminant.*
Si $\Delta \geq 0$, on sait que P admet une racine α, donc on sait par le thórème 6.34 que P peut se factoriser en $P(X) = (X-\alpha)U(X)$; de plus $\deg(X-\alpha) = \deg(U) = 1$ (car la somme des degrés doit être 2), donc ni $(X-\alpha)$ ni $U(X)$ ne sont associés à P, donc P n'est pas irréductible.
Supposons maintenant que $\Delta < 0$, on sait que P n'admet pas de racine réelle. Si on se donne une factorisation $P(X) = P_1(X)P_2(X)$ dans $\mathbb{R}[X]$, où ni P_1 ni P_2 ne sont des constantes, on doit avoir que $\deg(P_1) = \deg(P_2) = 1$. On écrit $P_1(X) = uX + v$ (avec $u \neq 0$), alors $P(X) = (uX+v)P_2(X) = (X + \frac{v}{u}) \times (uP_2(X))$. Donc $X + \frac{v}{u}$ divise P, ce qui donne que $-\frac{v}{u}$ est une racine réelle de P, une contradiction. Donc une telle factorisation de P n'existe pas, ce qui signifie que P est irréductible.

Soit un polynôme $P \in \mathbb{K}[X]$ de degré ≥ 1. On cherche à le factoriser dans $\mathbb{K}[X]$ en un produit de facteurs irréductibles. Si P est déjà irréductible, il n'y a rien à faire. Sinon, on peut écrire $P = P_1P_2$, avec P_1 et P_2 qui ne sont pas associés à P, ce qui implique qu'ils sont de degré compris entre 1 et $\deg(P)$ strictement. On peut donc recommencer la factorisation avec les polynômes "plus simples" P_1 et P_2. On obtient de cette manière un théorème analogue à celui de l'arithmétique.

Théorème 6.37 *Tout polynôme $P \in \mathbb{K}[X]$ de degré ≥ 1 peut être écrit comme un produit de facteurs irréductibles.*

Définition 6.38 *Un polynôme $P \in \mathbb{K}[X]$ de degré ≥ 1 est dit scindé (dans $\mathbb{K}[X]$) s'il peut être écrit*

$$P(X) = \lambda(X - a_1)\dots(X - a_n)$$

où $\lambda, a_1, \dots, a_n$ sont des éléments de $\mathbb{K}$.

Remarque 6.39

1. *Si on peut écrire $P(X) = \lambda(X - a_1)\dots(X - a_n)$, on vérifie facilement que n est le degré de P, $a_1, \dots, a_n$ sont les racines de P (éventuellement avec répétition), et λ est le coefficient dominant de P.*
2. *On regroupe souvent les racines quand elles sont égales entre elles : si P est scindé, on peut l'écrire sous la forme*
$$P(X) = \lambda(X - b_1)^{m_1}\dots(X - b_r)^{m_r},$$
où les $b_1, \dots, b_r$ sont deux à deux distincts. Ce sont les racines de P, et pour tout i, m_i est la multiplicité de la racine b_i. On constate en particulier que le degré de P vaut $m_1 + \dots + m_r$: on dit que le degré de P est égal au nombre de racines de P, comptées avec leur multiplicité (b_1 est compté m_1 fois, ...).
3. *Plus généralement, pour un polynôme non nul, on a toujours que le nombre de racines de P est inférieur ou égal au degré de P.*
4. *Certains polynômes ne sont pas scindés. Par exemple $X^2 + 1$ n'est pas scindé dans $\mathbb{R}[X]$ car il n'a pas de racine dans $\mathbb{R}$ (et il devrait avoir deux racines dans $\mathbb{R}$ s'il était scindé). Attention : ne pas confondre les cas $P(X) = X^2 + 1$, qui n'est pas scindé dans $\mathbb{R}[X]$, et $P(X) = (X+1)^2$, qui est scindé dans $\mathbb{R}[X]$, avec -1 comme racine double.*

Proposition 6.40 *Les polynômes irréductibles de $\mathbb{C}[X]$ sont exactement les polynômes de degré 1.*

Preuve On a déjà vu que les polynômes de degré 1 sont irréductibles (exemple 6.29).
Réciproquement, supposons que P est un polynôme irréductible (en particulier de degré ≥ 1 par définition). D'après le théorème 6.17, P admet une racine $a \in \mathbb{C}$, et donc, d'après le théorème 6.34, on peut factoriser P sous la forme $P(X) = (X - a)U(X)$. Comme P est irréductible, U doit être une constante non nulle, et donc P est de degré 1. □

En utilisant le théorème 6.37, on en déduit donc :

Corollaire 6.41 *Tout polynôme $P \in \mathbb{C}[X]$ est scindé dans $\mathbb{C}[X]$.*

Lemme 6.42 *Soient $P \in \mathbb{K}[X]$, et a et b deux racines distinctes de P. Alors $(X - a)(X - b)$ divise P.*

Preuve On utilise le théorème 6.34. Comme a est une racine de P, on peut écrire $P(X) = (X-a)U(X)$. De plus, b est une racine de P, donc $P(b) = (b-a)U(b) = 0$. Comme $b \neq a$, cela implique que $U(b) = 0$. On peut donc maintenant appliquer le théorème 6.34 pour U : il existe $V \in \mathbb{K}[X]$ tel que $U(X) = (X-b)V(X)$, et donc $P(X) = (X-a)(X-b)V(X)$. □

Remarque 6.43 *On pourrait aussi donner une démonstration analogue à celle d'arithmétique en disant que $(X-a)$ et $(X-b)$ sont "premiers entre eux" (on n'a pas défini ici cette notion pour les polynômes).*

Lemme 6.44 *Soit $z \in \mathbb{C}$. Alors $(X-z)(X-\overline{z}) \in \mathbb{R}[X]$.*

Preuve En développant, on trouve $P(X) = (X-z)(X-\overline{z}) = X^2 - (z+\overline{z})X + z\overline{z}$. Or $z+\overline{z} = 2\mathrm{Re}(z) \in \mathbb{R}$ et $z\overline{z} = |z|^2 \in \mathbb{R}$: tous les coefficients de P sont dans $\mathbb{R}$, ce qui signifie que $P \in \mathbb{R}[X]$. □

Corollaire 6.45 *Les polynômes irréductibles de $\mathbb{R}[X]$ sont exactement les polynômes de degré 1 et les polynômes de degré 2 de discriminant < 0.*

Preuve On a déjà vu que les polynômes de degré 1 sont irréductibles (exemple 6.29) et qu'un polynôme de degé 2 est irréductible dans $\mathbb{R}[X]$ ssi son discriminant est < 0 (exemple 6.36). Il reste donc à montrer que si $P \in \mathbb{R}[X]$ est un polynôme de degré ≥ 3, il n'est pas irréductible.
Si P admet une racine réelle a, il est divisible par $X-a$ et donc il n'est pas irréductible. Sinon, on sait que P admet une racine complexe z (par le théorème 6.17) qui n'est donc pas réelle. On a alors $z \neq \overline{z}$. Or, par la proposition 6.24, $\overline{z}$ est aussi une racine de P, et donc, en utilisant les deux lemmes précédents, P est divisible par $Q(X) = (X-z)(X-\overline{z})$, qui est un polynôme réel : on peut écrire $P(X) = Q(X)U(X)$, avec $\deg(P) = 3 > 2 = \deg(Q)$, donc U n'est pas une constante, et P n'est pas irréductible. □

Méthodes de factorisation des polynômes
On cherche à factoriser un polynôme $P \in \mathbb{K}[X]$, de degré ≥ 1, en facteurs irréductibles.
Le cas de la factorisation dans $\mathbb{C}[X]$ est le plus simple (au moins en théorie) : tous les facteurs irréductibles de P sont de degré 1, et pour les trouver, il suffit de chercher une racine a de P : on peut alors écrire $P(X) = (X-a)Q(X)$ (on trouve Q en faisant la division de P par $X-a$), et on cherche ensuite à factoriser Q. Mieux, on peut chercher la multiplicité m de la racine a, ce qui permet de factoriser $P(X) = (X-a)^m U(X)$, et le degré de U est plus petit que celui du polynôme Q précédent si $m \geq 2$.
En pratique, on sait bien trouver les racines de P s'il est de degré 1 ou 2, mais on n'a pas de méthode si $\deg(P) \geq 3$. Dans la plupart des exemples qui seront vus, il existera toujours une racine évidente, et on trouvera une racine de P en cherchant des valeurs simples : $0, 1, -1, 2, \ldots$.
Le cas de la factorisation dans $\mathbb{R}[X]$ peut être plus compliqué, car il est possible que P n'admette pas de racine réelle, sans pour autant être irréductible. On peut alors commencer à factoriser P dans $\mathbb{C}[X]$ comme vu précédemment. Et si on trouve une racine complexe non réelle z, P est divisible par $(X-z)(X-\overline{z})$, qui est un polynôme irréductible dans $\mathbb{R}[X]$.

Exemple 6.46

1. *On veut factoriser le polynôme $P(X) = X^4 - 12X^3 + 42X^2 - 52X + 21$. On voit que $P(1) = 0$: 1 est une racine évidente de P. Puis on calcule $P'(1) = 0$ et $P''(1) \neq 0$: 1 est une racinde double de P. On en déduit que P est divisible par $(X-1)^2$. En faisant la division de P par $(X-1)^2$ (ou encore $X^2 - 2X + 1$), on trouve que $P(X) = (X-1)^2(X^2 - 10X + 21)$. Puis on trouve facilement les racines de $Q(X) = X^2 - 10X + 21$, à savoir 3 et 7 (avec $\Delta = 16$), ce qui donne $Q(X) = (X-3)(X-7)$ (car le coefficient dominant de Q est 1). Donc la décomposition en facteurs irréductibles de P est $P(X) = (X-1)^2(X-3)(X-7)$; en particulier P est scindé, avec 4 racines comptées avec leur multiplicité.*
2. *On veut factoriser le polynôme $P(X) = X^4 + X^3 + 2X^2 + X + 1$ dans $\mathbb{R}[X]$. On sait que P n'est pas irréductible car il est de degré 4, mais on ne parvient pas à trouver de racine réelle. Cherchons des racines complexes : on voit que $P(i) = 1 - i - 2 + i + 1 = 0$. Ainsi, i est une racine de P, et donc $\overline{i} = -i$ aussi, et on sait que P est divisible par $Q(X) = (X-i)(X+i) = X^2 + 1$ (c'est un polynôme irréductible dans $\mathbb{R}[X]$). On fait la division de P par Q et on trouve $P(X) = (X^2+1)(X^2+X+1)$. Or $X^2 + X + 1$ est un polynôme irréductible dans $\mathbb{R}[X]$ puisque son discriminant est -3 ; on a donc bien trouvé la décomposition de P en facteurs irréductibles dans $\mathbb{R}[X]$.*

Chapitre 7

Géométrie dans le plan et dans l'espace

7.1 Géométrie dans le plan

7.1.1 Coordonnées

On considère le plan $\mathcal{P}$ muni d'un repère $(O, \overrightarrow{OI}, \overrightarrow{OJ})$: O est appelé l'origine du repère, et les vecteurs $\overrightarrow{OI}$ et $\overrightarrow{OJ}$ sont par définition non-colinéaires. Tout point M de $\mathcal{P}$ peut alors être décrit par ses coordonnées $(x_M, y_M) \in \mathbb{R}^2$ dans ce repère. Cela signifie que

$$\overrightarrow{OM} = x_M \overrightarrow{OI} + y_M \overrightarrow{OJ}$$

Soient deux points A et B du plan $\mathcal{P}$, et leurs coordonnées respectives (x_A, y_A) et (x_B, y_B). On peut considérer le vecteur d'origine A et d'extrémité B, noté $\overrightarrow{AB}$.
Pour un tel vecteur $\overrightarrow{AB}$, on définit ses coordonnées $(x_B - x_A, y_B - y_A)$ dans $\mathbb{R}^2$. Un vecteur est complètement déterminé par ses coordonnées : cela siginifie que $\overrightarrow{AB} = \overrightarrow{CD}$ si et seulement si $\begin{cases} x_B - x_A = x_D - x_C \\ y_B - y_A = y_D - y_C \end{cases}$; deux vecteurs égaux peuvent avoir des origines et des extrémités différentes.
On appelle le vecteur nul le vecteur dont les coordonnées sont nulles : $\overrightarrow{0} = (0, 0)$.
Dire que deux vecteurs $\overrightarrow{u}$ et $\overrightarrow{v}$ sont colinéaires revient à dire qu'il existe un réel λ tel que $\overrightarrow{v} = \lambda \overrightarrow{u}$ ou $\overrightarrow{u} = \lambda \overrightarrow{v}$. Le vecteur nul est colinéaire à tous les vecteurs puisqu'on a toujours $\overrightarrow{0} = 0\overrightarrow{u}$.

On suppose désormais que le repère $(O, \overrightarrow{OI}, \overrightarrow{OJ})$ est orthogonal, c'est-à-dire que les vecteurs $\overrightarrow{OI}$ et $\overrightarrow{OJ}$ sont orthogonaux (on peut aussi dire perpendiculaires). Pour deux vecteurs $\overrightarrow{u} = (x, y)$ et $\overrightarrow{u'} = (x', y')$, on définit leur produit scalaire :

$$\langle \overrightarrow{u}, \overrightarrow{u'} \rangle = xx' + yy'.$$

Fait 7.1 *Les vecteurs $\overrightarrow{u}$ et $\overrightarrow{u'}$ sont orthogonaux ssi leur produit scalaire $\langle \overrightarrow{u}, \overrightarrow{u'} \rangle$ est nul.*

7.1.2 Droites

Pour définir une droite, il suffit de deux points distincts : pour tous points distincts A et B, il existe une unique droite passant par A et B, que l'on note habituellement (AB).
Pour tout point M du plan, le point M appartient à la droite (AB) ssi le vecteur $\overrightarrow{AM}$ est colinéaire au vecteur $\overrightarrow{AB}$. On dit dans ce cas que A est une origine de la droite et $\overrightarrow{AB}$ en est un vecteur directeur.
Plus généralement, si on se donne un point A de coordonnées (x_A, y_A) et un vecteur non nul $\overrightarrow{u} = (a, b)$, on peut définir la droite D d'origine A et de vecteur directeur $\overrightarrow{u}$: c'est l'ensemble des points M du plan tels que $\overrightarrow{AM}$ est colinéaire à $\overrightarrow{u}$. Ainsi, en notant (x, y) les coordonnées de M, le point M appartient à D ssi il existe $\lambda \in \mathbb{R}$ tel que $\overrightarrow{AM} = \lambda u$, c'est-à-dire $\begin{cases} x - x_A = \lambda a \\ y - y_A = \lambda b \end{cases}$

. On obtient ainsi la description paramétrique de la droite D : c'est l'ensemble des points de coordonnées $\begin{cases} x = x_A + \lambda a \\ y = y_A + \lambda b \end{cases}$, où le paramètre λ parcourt $\mathbb{R}$.

De manière alternative, on peut décrire la droite D à partir d'une origine et d'un vecteur normal. Un vecteur normal à une droite est un vecteur non nul qui est orthogonal au vecteur directeur de la droite. Si A de coordonnée (x_A, y_A) est une origine de la droite D, et si $\overrightarrow{v} = (c, d)$ est un vecteur normal à D, alors le point M appartient à D ssi $\overrightarrow{AM}$ est orthogonal à $\overrightarrow{v}$. On a donc, en notant (x, y) les coordonnées de M :

$$M \in D \Leftrightarrow \langle \overrightarrow{AM}, \overrightarrow{v} \rangle = 0 \Leftrightarrow c(x - x_A) + d(y - y_A) = 0.$$

On a ainsi obtenu une équation cartésienne de la droite D :

$$D = \{M \in \mathcal{P}; cx + dy = cx_A + dy_A\}.$$

Notons que les coefficients devant les coordonnées x et y de M sont précisément les coordonnées (c, d) du vecteur normal $\overrightarrow{v}$.

Exemple 7.2 *Considérons les points A de coordonnées $(1, 2)$ et B de coordonnées $(3, 5)$, et D la droite (AB).*
La droite D admet A comme origine et $\overrightarrow{AB} = (2, 3)$ comme vecteur directeur. On en déduit sa description paramétrique : D est l'ensemble des points M de coordonnées $\begin{cases} x = 1 + 2\lambda \\ y = 2 + 3\lambda \end{cases}$ quand λ parcourt $\mathbb{R}$.
On constate facilement que $\overrightarrow{v} = (-3, 2)$ est un vecteur normal à la droite D, puisque $\langle \overrightarrow{AB}, \overrightarrow{v} \rangle = 2 \times (-3) + 3 \times 2 = 0$. En utilisant la méthode précédente, on peut alors obtenir une équation cartésienne de la droite D : $D = \{M \in \mathcal{P}; -3x + 2y = 1\}$.

On peut également faire le travail précédent dans le sens inverse. Soit D une droite donnée par son équation cartésienne $cx + dy = e$. Pour trouver une origine et un vecteur directeur de D, il suffit de trouver deux point distincts de A et B sur D : on peut alors considérer A comme l'origine de D et $\overrightarrow{AB}$ comme vecteur directeur. D'après ce qu'on a vu précédemment, il est encore plus facile de trouver un vecteur normal pour D, simplement en utilisant les coefficients de l'équation cartésienne : le vecteur $\overrightarrow{v} = (c, d)$ est un vecteur normal pour D.

Fait 7.3 *Les vecteurs directeurs de D sont tous colinéaires entre eux. Les vecteurs normaux à D sont tous colinéaires entre eux.*

Exemple 7.4 *Soit D la droite d'équation $2x + y = 3$. On voit que les points A de coordonnées $(0, 3)$ et B de coordonnées $(3, -3)$ sont deux points distincts de D ; ainsi D est la droite d'origine A et de vecteur directeur $\overrightarrow{AB} = (3, -6)$. Le vecteur $\overrightarrow{v} = (6, 3)$ est orthogonal au vecteur directeur $\overrightarrow{AB}$, c'est donc un vecteur normal à D. On peut aussi trouver un vecteur normal en regardant l'équation cartésienne de D : $\overrightarrow{v'} = (2, 1)$ est aussi un vecteur normal à D, et on constate bien que $\overrightarrow{v}$ et $\overrightarrow{v'}$ sont colinéaires.*

7.1.3 Distance

On suppose désormais que le repère $(O, \overrightarrow{OI}, \overrightarrow{OJ})$ est orthonormal (on dit aussi orthonormé), c'est-à-dire que c'est un repère orthogonal tel que de plus les vecteurs $\overrightarrow{OI}$ et $\overrightarrow{OJ}$ sont de longueur 1. En utilisant le théorème de Pythagore, on peut alors calculer la longueur de tout vecteur $\overrightarrow{OM}$.

Proposition 7.5 *Soit M un point de coordonnées (x, y). Alors la longueur du vecteur $\overrightarrow{OM}$ est $\sqrt{x^2 + y^2}$.*

Preuve Considérons le point N de coordonnée $(x, 0)$, alors le triangle OMN est rectangle en N (on peut faire le dessin). Le théorème de Pythagore donne donc $OM^2 = ON^2 + MN^2$, or $ON = |x|$ et $MN = |y|$ donc $OM = \sqrt{x^2 + y^2}$. □
Soient A et B deux points du plan, et M le point de coordonnées $(x_B - x_A, y_B - y_A)$. Alors $\overrightarrow{OM} = \overrightarrow{AB}$, et on en déduit :

Corollaire 7.6 *La distance entre deux points A et B est $\sqrt{(x_B - x_A)^2 + (y_B - y_A)^2}$.*

Cela permet d'écrire l'équation des cercles. Soit A un point du plan et R un réel positif. Le cercle de centre A et de rayon R est l'ensemble des points M tel que $AM = R$. En utilisant le corollaire ci-dessus, et en notant (x, y) les coordonnées de M, on trouve l'équation du cercle : $\sqrt{(x - x_A)^2 + (y - y_A)^2} = R$, qu'on écrit le plus souvent sous sa forme équivalente : $(x - x_A)^2 + (y - y_A)^2 = R^2$.
On peut aussi faire le travail dans l'autre sens, et reconnaître l'équation de cercles comme dans l'exemple suivant.

Exemple 7.7 *On cherche à caractériser l'ensemble des points M dont les coordonnées (x, y) satisfont l'équation $x^2 + y^2 + 4x + 6y + 2 = 0$. La méthode est de reconnaître dans les termes en x et les termes en y le début de développement d'un carré : $x^2 + y^2 + 4x + 6y + 2 = (x + 2)^2 - 4 + (y + 3)^2 - 9 + 2$. Donc $x^2 + y^2 + 4x + 6y + 2 = 0$ ssi $(x + 2)^2 + (y + 3)^2 = 11$: c'est l'équation du cercle de centre A, de coordonnées $(-2, -3)$, et de rayon $\sqrt{11}$.*

Exercice Déterminer l'ensemble des points M de coordonnées (x, y) qui vérifient $x^2 + y^2 - 2x + 4y \leq 30$.

7.2 Géométrie dans l'espace

7.2.1 Coordonnées

Les premières notions de géométrie dans l'espace sont très semblables à celles dnas le plan. On considère l'espace $\mathcal{E}$ muni d'un repère $(O, \overrightarrow{OI}, \overrightarrow{OJ}, \overrightarrow{OK})$: O est appelé l'origine du repère, et les vecteurs $\overrightarrow{OI}$, $\overrightarrow{OJ}$ et $\overrightarrow{OK}$ sont par définition non-coplanaires (ils ne sont pas contenus dans un même plan). Tout point M de $\mathcal{E}$ peut alors être décrit par ses coordonnées $(x_M, y_M, z_M) \in \mathbb{R}^3$ dans ce repère. Cela signifie que

$$\overrightarrow{OM} = x_M \overrightarrow{OI} + y_M \overrightarrow{OJ} + z_M \overrightarrow{OK}$$

Soient deux points A et B de l'espace $\mathcal{E}$, et leurs coordonnées respectives (x_A, y_A, z_A) et (x_B, y_B, z_B). On peut considérer le vecteur d'origine A et d'extrémité B, noté $\overrightarrow{AB}$.
Pour un tel vecteur $\overrightarrow{AB}$, on définit ses coordonnées $(x_B - x_A, y_B - y_A, z_B - z_A)$ dans $\mathbb{R}^3$. Un vecteur est complètement déterminé par ses coordonnées : cela siginifie que $\overrightarrow{AB} = \overrightarrow{CD}$ si et seulement si $\begin{cases} x_B - x_A = x_D - x_C \\ y_B - y_A = y_D - y_C \\ z_B - z_A = z_D - z_C \end{cases}$; deux vecteurs égaux peuvent avoir des origines et des extrémités différentes.
On appelle le vecteur nul le vecteur dont les coordonnées sont nulles : $\overrightarrow{0} = (0, 0, 0)$.

On suppose désormais que le repère $(O, \overrightarrow{OI}, \overrightarrow{OJ}, \overrightarrow{OK})$ est orthogonal, c'est-à-dire que les vecteurs $\overrightarrow{OI}$, $\overrightarrow{OJ}$ et $\overrightarrow{OK}$ sont orthogonaux deux à deux. Pour deux vecteurs $\overrightarrow{u} = (x, y, z)$ et $\overrightarrow{u'} = (x', y', z')$, on définit leur produit scalaire :

$$\langle \overrightarrow{u}, \overrightarrow{u'} \rangle = xx' + yy' + zz'.$$

Fait 7.8 *Les vecteurs $\overrightarrow{u}$ et $\overrightarrow{u'}$ sont orthogonaux ssi leur produit scalaire $\langle \overrightarrow{u}, \overrightarrow{u'} \rangle$ est nul.*

7.2.2 Plan dans l'espace

Contrairement à ce qui se passe dans le plan, un vecteur normal et une origine dans l'espace ne permettent pas de définir une droite, mais un plan. Plus précisément, considérons un point A de coordonnées (x_A, y_A, z_A) et un vecteur non nul $\overrightarrow{v} = (a, b, c)$. Alors l'ensemble des points M tels que $\overrightarrow{AM}$ est orthogonal à $\overrightarrow{v}$ forme un plan P ; on l'appelle le plan d'origine A et de vecteur normal $\overrightarrow{v}$. En écrivant le produit scalaire $\langle \overrightarrow{AM}, \overrightarrow{v} \rangle$, et en notant (x, y, z) les coordonnées de M, on obtient une équation cartésienne de P :

$$M \in P \Leftrightarrow ax + by + cz = ax_A + by_A + cz_A$$

Tous les vecteurs non nuls colinéaires à $\overrightarrow{v}$ sont aussi des vecteurs normaux à P.

On peut également définir un plan dans l'espace à partir de trois points non alignés : si A, B, C sont trois points non alignés, il existe un unique plan P qui contient A, B, C. Plus précisément, P est défini en choisissant A comme origine et les vecteurs $\overrightarrow{AB}$ et $\overrightarrow{AC}$ comme vecteurs générateurs :

$$P = P_{A,\overrightarrow{AB},\overrightarrow{AC}} := \{M \in \mathcal{E}; \exists \lambda \in \mathbb{R}\ \exists \mu \in \mathbb{R}\ \overrightarrow{AM} = \lambda\overrightarrow{AB} + \mu\overrightarrow{AC}\}$$

Si on permute les rôles joués par A, B et C, on obtient le même plan. Par exemple, en conservant la notation ci-dessus, vérifions que $P_{A,\overrightarrow{AB},\overrightarrow{AC}} = P_{B,\overrightarrow{BA},\overrightarrow{BC}}$. On va utiliser les relations de Chasles qui nous donnent $\overrightarrow{AB} = -\overrightarrow{BA}$ et $\overrightarrow{AC} = -\overrightarrow{BA} + \overrightarrow{BC}$. Alors si $M \in P_{A,\overrightarrow{AB},\overrightarrow{AC}}$, il existe par définition des réels λ et μ tels que $\overrightarrow{AM} = \lambda\overrightarrow{AB} + \mu\overrightarrow{AC}$, et donc $\overrightarrow{BM} = \overrightarrow{BA} + \overrightarrow{AM} = (1 - \lambda - \mu)\overrightarrow{BA} + \mu\overrightarrow{BC}$, ce qui prouve que $M \in P_{B,\overrightarrow{BA},\overrightarrow{BC}}$. On a donc montré que $P_{A,\overrightarrow{AB},\overrightarrow{AC}} \subset P_{B,\overrightarrow{BA},\overrightarrow{BC}}$; on montre l'inclusion réciproque de la même manière et ainsi que $P_{A,\overrightarrow{AB},\overrightarrow{AC}} = P_{B,\overrightarrow{BA},\overrightarrow{BC}}$. On emploie la même méthode si on choisit C comme origine.

7.2.3 Droite dans l'espace

Pour définir une droite dans l'espace, il suffit de deux points distincts : pour tous points distincts A et B, il existe une unique droite passant par A et B, notée (AB).
Pour tout point M de l'espace, le point M appartient à la droite (AB) ssi le vecteur $\overrightarrow{AM}$ est colinéaire au vecteur $\overrightarrow{AB}$. On dit dans ce cas que A est une origine de la droite et $\overrightarrow{AB}$ en est un vecteur directeur. Plus généralement, si on se donne un point A de coordonnées (x_A, y_A, z_A) et un vecteur non nul $\overrightarrow{u} = (a, b, c)$, on peut définir la droite D d'origine A et de vecteur directeur $\overrightarrow{u}$: c'est l'ensemble des points M du plan tels que $\overrightarrow{AM}$ est colinéaire à $\overrightarrow{u}$. Ainsi, en notant (x, y, z) les coordonnées de M, le point M appartient à D ssi il existe $\lambda \in \mathbb{R}$ tel que $\overrightarrow{AM} = \lambda u$, c'est-à-dire $\begin{cases} x - x_A = \lambda a \\ y - y_A = \lambda b \\ z - z_A = \lambda c \end{cases}$. On obtient ainsi la description paramétrique de la droite D : c'est l'ensemble des points de coordonnées $\begin{cases} x = x_A + \lambda a \\ y = y_A + \lambda b \\ z = z_A + \lambda c \end{cases}$, où le paramètre λ parcourt $\mathbb{R}$.
Cherchons maintenant des équations cartésiennes pour une telle droite D. Comme le vecteur $\overrightarrow{u} = (a, b, c)$ est non-nul, au moins une des coordonnées a, b, c est non-nulle. Supposons par exemple que $a \neq 0$ (si on avait $a = 0$, il faudrait adapter le raisonnement ci-dessous en faisant jouer à b ou c le rôle de a). On sait que le point M appartient à D ssi il existe $\lambda \in \mathbb{R}$ tel que $\begin{cases} x - x_A = \lambda a \\ y - y_A = \lambda b \\ z - z_A = \lambda c \end{cases}$. Comme $a \neq 0$, la seule valeur de λ possible pour que la première condition soit vérifiée est $\lambda = \frac{x - x_A}{a}$. Pour que cette valeur de λ soit bien une solution de tout le système, il faut aussi que les deux autres équations soient vérifiées, c'est-à-dire $\begin{cases} y = y_A + \frac{x - x_A}{a} b \\ z = z_A + \frac{x - x_A}{a} c \end{cases}$. En réécrivant ces équations, on trouve une condition nécessaire et suffisante pour que $M \in P$:

$$M \in P \Leftrightarrow \begin{cases} bx - ay = bx_A - ay_A \\ cx - az = cx_A - az_A \end{cases}$$

Ce sont des équations cartésiennes de D.
Notons que, d'après ce qu'on a vu précédemment, $bx - ay = bx_A - ay_A$ est l'équation cartésienne d'un plan dans l'espace, et de même pour $cx - az = cx_A - az_A$. La droite D est alors l'intersection de ces deux plans.

7.2.4 Distance

On suppose désormais que le repère $(O, \overrightarrow{OI}, \overrightarrow{OJ}, \overrightarrow{OK})$ est orthonormal (on dit aussi orthonormé), c'est-à-dire que c'est un repère orthogonal tel que de plus les vecteurs $\overrightarrow{OI}$, $\overrightarrow{OJ}$ et $\overrightarrow{OK}$ sont de longueur

1. En utilisant le théorème de Pythagore, on peut alors calculer la longueur de tout vecteur $\overrightarrow{OM}$.

Proposition 7.9 *Soit M un point de coordonnées (x, y, z). Alors la longueur du vecteur $\overrightarrow{OM}$ est $\sqrt{x^2 + y^2 + z^2}$.*

Preuve Considérons le point N de coordonnée $(x, y, 0)$, alors le triangle OMN est rectangle en N (on le vérifie en calculant le produit scalaire $\langle \overrightarrow{ON}, \overrightarrow{NM} \rangle = x \times 0 + y \times 0 + 0 \times z = 0$). Le théorème de Pythagore donne donc $OM^2 = ON^2 + MN^2$, or $MN = |z|$ et $ON^2 = x^2 + y^2$ (c'est le même calcul que dans le plan) donc $OM = \sqrt{x^2 + y^2 + z^2}$. □

Corollaire 7.10 *La distance entre deux points A et B est $\sqrt{(x_B - x_A)^2 + (y_B - y_A)^2 + (z_B - z_A)^2}$.*

Exemple 7.11 *On cherche à caractériser l'ensemble des points M dont les coordonnées (x, y, z) satisfont l'équation $(x-2)^2 + (y+1)^2 + (z-1)^2 = 9$. En posant A le point de coordonnées $(2, -1, 1)$, cette équation peu s'écrire $AM^2 = 9$. On a donc l'équation de la sphère de centre A et de rayon 3.*