Université Blaise Pascal
U.F.R. Sciences et Technologies
Département de Mathématiques et Informatique

Licence de Mathématiques

Troisième année

ALGÈBRE ET GÉOMÉTRIE

Polycopié du cours

2011-2012

François Dumas

Université Blaise Pascal - U.F.R. Sciences et Technologies - Département de mathématiques et informatique

Licence de Mathématiques
L3 S5 "Algèbre et Géométrie"

année 2011-2012

Plan du cours

François DUMAS

Première partie

- Leçon 1 - **Groupes et sous-groupes**
- Leçon 2 - **Morphismes de groupes, produit direct de groupes**
- Leçon 3 - **Groupes monogènes, groupes cycliques**
- Leçon 4 - **Groupe symétrique**
- Leçon 5 - **Conjugaison et sous-groupes normaux**
- Leçon 6 - **Groupes quotients**
- Leçon 7 - **Groupe opérant sur un ensemble**
- Leçon 8 - **Groupes d'isométries**
- Leçon 9 - **Sous-groupes d'isométries laissant invariante une partie du plan ou de l'espace**
- Leçon 10 - **Quelques compléments sur les groupes quotients**

Seconde partie

- Leçon 11 - **Anneaux, sous-anneaux, morphismes d'anneaux**
- Leçon 12 - **Eléments inversibles dans un anneau, corps, intégrité**
- Leçon 13 - **Idéal d'un anneau**
- Leçon 14 - **Anneaux quotients**
- Leçon 15 - **Anneaux euclidiens, anneaux principaux**
- Leçon 16 - **Divisibilité**
- Leçon 17 - **Arithmétique dans les anneaux principaux**
- Leçon 18 - **Arithmétique dans les anneaux de polynômes**
- Leçon 19 - **Polynômes d'endomorphismes**
- Leçon 20 - **Anneaux de polynômes en plusieurs indéterminées**

---

version révisée au 16 août 2011

Ces notes de cours contiennent inévitablement des coquilles, des erreurs, des points à améliorer ou modifier. Merci de me les signaler.

Francois.Dumas@univ-bpclermont.fr http://math.univ-bpclermont.fr/∼fdumas

Leçon 1

Groupes et sous-groupes

1.1 Notion de groupe

1.1.1 Définition. Soit G un ensemble non-vide. On appelle *loi de composition interne* dans G, ou *opération interne* dans G, toute application $* : G \times G \to G$.

Une telle loi de composition interne permet donc d'associer à tout couple (x, y) d'éléments de G un autre élément de G, noté $x * y$, et appelé le produit de x par y pour la loi $*$.

1.1.2 Définition. On appelle *groupe* tout ensemble non-vide G muni d'une loi de composition interne $*$, vérifiant les 3 propriétés suivantes (appelées axiomes de la structure de groupe):

(A1) la loi $*$ est associative dans G ;
rappelons que cela signifie que $x * (y * z) = (x * y) * z$ pour tous $x, y, z \in G$.

(A2) la loi $*$ admet un élément neutre dans G ;
rappelons que cela signifie qu'il existe $e \in G$ tel que $x * e = e * x = x$ pour tout $x \in G$.

(A3) tout élément de G admet un symétrique dans G pour la loi $*$;
rappelons que cela signifie que, pour tout $x \in G$, il existe $x' \in G$ tel que $x * x' = x' * x = e$.

1.1.3 Définition. On appelle *groupe commutatif*, ou *groupe abélien*, tout groupe G dont la loi $*$ vérifie de plus la condition supplémentaire de commutativité: $x * y = y * x$ pour tous $x, y \in G$.

1.1.4 Exemples.

(a) Pour tout ensemble X, l'ensemble $\mathcal{S}(X)$ des bijections de X sur X muni de la loi $\circ$ de composition des bijections est un groupe, appelé groupe symétrique sur X.

Le neutre en est l'identité de X, car $f \circ \mathrm{id}_X = \mathrm{id}_X \circ f = f$ pour toute $f \in \mathcal{S}(X)$. Pour toute $f \in \mathcal{S}(X)$, le symétrique de f pour la loi $\circ$ est la bijection réciproque f^{-1}, car $f \circ f^{-1} = f^{-1} \circ f = \mathrm{id}_X$. Dès lors que X contient au moins trois éléments, le groupe $\mathcal{S}(X)$ n'est pas abélien (montrez-le).

(b) Pour tout entier $n \geq 1$, l'ensemble $\mathrm{GL_n}(\mathbb{R})$ des matrices carrées d'ordre n inversibles à coefficients réels est un groupe pour la multiplication des matrices.

Le neutre en est la matrice identité I_n, car $M \times I_n = I_n \times M = M$ pour toute $M \in \mathrm{GL_n}(\mathbb{R})$. Pour toute $M \in \mathrm{GL_n}(\mathbb{R})$, le symétrique de M pour la loi $\times$ est la matrice inverse M^{-1}, car $M \times M^{-1} = M^{-1} \times M = I_n$. Dès lors que $n \geq 2$, le groupe $\mathrm{GL_n}(\mathbb{R})$ n'est pas abélien (montrez-le).

(c) L'ensemble $\mathbb{C}$ des nombres complexes muni de l'addition est un groupe abélien.

Le neutre en est le nombre complexe nul 0, car $z + 0 = 0 + z = z$ pour tout $z \in \mathbb{C}$. Pour tout $z \in \mathbb{C}$, le symétrique de z pour l'addition est son opposé $-z$, car $z + (-z) = (-z) + z = 0$.

(d) L'ensemble $\mathbb{C}^*$ des nombres complexes non-nuls muni de la multiplication est un groupe abélien.

Le neutre en est le nombre complexe 1, car $z.1 = 1.z = z$ pour tout $z \in \mathbb{C}^*$. Pour tout $z \in \mathbb{C}^*$, le symétrique de z pour la multiplication est son inverse z^{-1}, car $z.z^{-1} = z^{-1}.z = 1$.

1.1.5 Remarques et conventions de notation. Afin d'éviter la lourdeur de la notation $*$, on convient généralement de noter la loi de composition interne d'un groupe quelconque G, soit comme une multiplication (par un point.), soit comme une addition (par un $+$). Dans le premier cas, le symétrique d'un élément est appelé son inverse, dans le second cas, son opposé. Usuellement, on réserve la notation additive au cas des groupes abéliens. C'est pourquoi, dans toute la suite de ce polycopié, on adoptera pour les groupes quelconques, conformément à l'usage courant, la notation multiplicative.

(a) Un groupe G sera donc un ensemble non-vide G muni d'une loi de composition interne .
→ associative ($x.(y.z) = (x.y).z$ pour tous $x, y, z \in G$),
→ admettant un élément neutre e ($x.e = e.x = x$ pour tout $x \in G$),
→ et telle que tout élément $x \in G$ admette un symétrique x^{-1} pour la loi . ($x.x^{-1} = x^{-1}.x = e$).

(b) De plus, l'éventuelle commutativité de G se traduira par: $x.y = y.x$ pour tous $x, y \in G$.

(c) On utilisera la notation $x^n = x.x.x \cdots x$ (n facteurs) pour tous $x \in G$ et $n \in \mathbb{N}^*$, ainsi que les conventions $x^0 = e$, et $x^{-n} = (x^n)^{-1}$.

(d) Pour tous $x, y \in G$, on a $(x.y)^{-1} = y^{-1}.x^{-1}$ (montrez-le, attention à l'ordre !)

1.1.6 *Quelques remarques techniques, mais parfois utiles, sur les axiomes de la structure de groupe.*

(a) Dans un groupe G, l'élément neutre e est nécessairement unique, et le symétrique d'un élément quelconque est nécessairement unique.

(b) Si G est un ensemble non-vide muni d'une loi de composition interne . qui est supposée associative, il suffit que G admette un élément neutre e à droite (ce qui signifie que $x.e = x$ pour tout $x \in G$) et que tout élément $x \in G$ admette un symétrique $x' \in G$ à droite (ce qui signifie que $x.x' = e$) pour conclure que G est un groupe.

1.2 Sous-groupe

1.2.1 Exemple introductif. Considérons le groupe $\mathbb{C}^*$ pour la multiplication. Dans $\mathbb{C}^*$, considérons le sous-ensemble $\mathbb{R}^*$. En restreignant à $\mathbb{R}^*$ la multiplication dans $\mathbb{C}^*$, on obtient une loi de composition interne dans $\mathbb{R}^*$ (car le produit de deux réels non-nuls est encore un réel non-nul). La question de savoir si $\mathbb{R}^*$ est lui-même un groupe pour la loi . est donc fondée.

L'associativité de . dans $\mathbb{R}^*$ est évidemment vérifiée (la relation $x.(y.z) = (x.y).z$ étant vraie pour tous $x, y, z \in \mathbb{C}^*$, elle est a fortiori vraie pour tous $x, y, z \in \mathbb{R}^*$).

Le nombre complexe 1 est un élément de $\mathbb{R}^*$, et il est neutre pour la loi . dans $\mathbb{R}^*$ (la relation $x.1 = 1.x = x$ étant vraie pour tout $x \in \mathbb{C}^*$, elle est a fortiori vraie pour tout $x \in \mathbb{R}^*$).

Pour tout $x \in \mathbb{R}^*$, l'inverse x^{-1} de x dans $\mathbb{C}^*$ appartient à $\mathbb{R}^*$ et est donc l'inverse de x dans $\mathbb{R}^*$ (les égalités $x.x^{-1} = x^{-1}.x = 1$ étant alors vraies dans $\mathbb{R}^*$ comme dans $\mathbb{C}^*$).

On conclut que le sous-ensemble $\mathbb{R}^*$ est lui-même un groupe pour la multiplication déduite de celle de $\mathbb{C}^*$ par restriction. On dit alors que $\mathbb{R}^*$ est un sous-groupe de $\mathbb{C}^*$.

Le même raisonnement s'applique si on remplace $\mathbb{R}^*$ par $\mathbb{Q}^*$, mais pas si on le remplace par l'ensemble des nombres imaginaires purs (car le produit de deux imaginaires purs n'est pas un imaginaire pur), ou par l'ensemble $\mathbb{Z}^*$ (car l'inverse d'un entier non-nul peut ne pas être un entier).

1.2.2 Définition. Soit G un groupe muni d'une loi de composition interne . et soit H un sous-ensemble non-vide de G. On dit que H est un *sous-groupe* de G lorsque les deux conditions suivantes sont vérifiées:

(1) H est stable pour la loi . (ce qui signifie $x.y \in H$ pour tous $x, y \in H$),

(2) H est stable par passage à l'inverse (ce qui signifie $x^{-1} \in H$ pour tout $x \in H$).

Dans ce cas, la restriction à H de la loi . de G définit une loi de composition interne dans H, pour laquelle H est lui-même un groupe.

1.2.3 Exemples.

(a) $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ sont des sous-groupes du groupe $\mathbb{C}$ muni de l'addition, mais pas $\mathbb{N}$ (car l'opposé d'un élément de $\mathbb{N}$ n'est pas nécessairement un élément de $\mathbb{N}$).

(b) L'ensemble $\mathbb{U}$ des nombres complexes de module égal à 1 est un sous-groupe de $\mathbb{C}^*$ muni de la multiplication. Pour tout entier $n \geq 1$, l'ensemble $\mathbb{U}_n$ des racines n-ièmes de l'unité est un sous-groupe de $\mathbb{U}$.

(c) Pour tout $n \geq 2$, l'ensemble des matrices triangulaires supérieures d'ordre n à coefficients réels sans 0 sur la diagonale est un sous-groupe non-abélien de $\mathrm{GL}_n(\mathbb{R})$. L'ensemble des matrices diagonales d'ordre n à coefficients réels sans 0 sur la diagonale en est un sous-groupe abélien.

(d) *Important:* tout groupe G contient toujours au moins pour sous-groupes le sous-groupe trivial $\{e\}$ formé du seul élément neutre, et le groupe G lui-même.

1.2.4 Remarques.

(a) Les deux conditions de la définition 1.2.2 peuvent être synthétisées en une seule: soit H un sous-ensemble non-vide d'un groupe G, alors

(H est un sous-groupe de G) si et seulement si (pour tous $x, y \in H$, on a $x.y^{-1} \in H$).

(b) Si H est un sous-groupe de G, alors l'élément neutre e de G appartient nécessairement à H (car pour tout $x \in H$, on a $x^{-1} \in H$, et $x.x^{-1} = e \in H$). A contrario, un sous-ensemble de G qui ne contient pas le neutre de G ne peut en aucun cas être un sous-groupe (ce qui est dans la pratique une façon pratique très fréquente de vérifier qu'un sous-ensemble d'un groupe connu n'est pas un sous-groupe).

(c) Tout sous-groupe d'un groupe abélien est lui-même abélien, mais un groupe non abélien peut contenir des sous-groupes abéliens aussi bien que des sous-groupes non-abéliens (voir 1.2.3.c).

(d) Dans la pratique, dans la plupart des cas, pour montrer qu'un ensemble donné est un groupe, on ne revient pas à la définition par les trois axiomes, mais on cherche à montrer qu'il est un sous-groupe d'un groupe déjà connu.

(e) Attention, pour vérifier qu'un sous-ensemble donné d'un groupe est un sous-groupe, on n'oubliera pas de vérifier au préalable qu'il est non-vide; d'après la remarque (b) ci-dessus, le plus naturel pour cela est de s'assurer qu'il contient le neutre.

1.2.5 Exemples.

(a) Soit E un espace vectoriel de dimension finie. L'ensemble $\mathrm{GL}(E)$ des automorphismes d'espace vectoriel de E est un groupe appelé groupe linéaire de E; pour le montrer, il suffit de vérifier que c'est un sous-groupe de $\mathcal{S}(E)$. Les éléments de $\mathrm{GL}(E)$ qui ont un déterminant égal à 1 forment un sous-groupe de $\mathrm{GL}(E)$, dit groupe spécial linéaire, noté $\mathrm{SL}(E)$.

(b) Supposons de plus que E est euclidien. L'ensemble $\mathrm{O}(E)$ des isométries vectorielles de E est un groupe appelé groupe orthogonal de E; pour le montrer, il suffit de vérifier que c'est un sous-groupe de $\mathrm{GL}(E)$. L'ensemble $\mathrm{SO}(E)$ des isométries vectorielles positives de E est un sous-groupe de $\mathrm{O}(E)$, et l'on a $\mathrm{SO}(E) = \mathrm{O}(E) \cap \mathrm{SL}(E)$. L'ensemble des isométries vectorielles négatives de E n'est pas un sous-groupe de $\mathrm{O}(E)$ (il ne contient pas le neutre id_E).

(c) L'ensemble des bijections continues et strictement croissantes de $\mathbb{R}$ dans $\mathbb{R}$ est un groupe pour la loi $\circ$; pour le montrer, il suffit de vérifier que c'est un sous-groupe du groupe $\mathcal{S}(\mathbb{R})$ de toutes les bijections de $\mathbb{R}$ sur $\mathbb{R}$.

1.2.6 Proposition. *L'intersection de deux sous-groupes d'un groupe G est un sous-groupe de G. Plus généralement, l'intersection d'une famille quelconque de sous-groupes d'un groupe G est un sous-groupe de G.*

Preuve. Il suffit pour le montrer de prouver le second point. Soit donc $(H_i)_{i\in I}$ une famille de sous-groupes d'un groupe G. Posons $K = \bigcap_{i\in I} H_i$ l'intersection de tous les H_i. L'ensemble K est non-vide, car il contient le neutre e puisque celui-ci appartient à chacun des sous-groupes H_i. Soient x et y deux éléments de K. Pour tout $i \in I$, on a $x.y^{-1} \in H_i$ puisque H_i est un sous-groupe. Donc $x.y^{-1} \in K$. Ce qui prouve que K est un sous-groupe de G. □

1.2.7 Remarque. Attention, la réunion de deux sous-groupes n'est en général pas un sous-groupe.

Contre-exemple. Dans le groupe $\mathbb{C}^*$ muni de la multiplication, considérons le sous-groupe $\mathbb{U}_2 = \{1, -1\}$ des racines carrées de l'unité et le sous-groupe $\mathbb{U}_3 = \{1, j, j^2\}$ des racines cubiques de l'unité. Notons $K = \mathbb{U}_2 \cup \mathbb{U}_3 = \{1, -1, j, j^2\}$. On a $j \in K$ et $-1 \in K$, mais le produit $(-1).j = -j \notin K$. Donc K n'est pas stable par la multiplication, et ce n'est donc pas un sous-groupe de $\mathbb{C}^*$. □

1.2.8 Proposition et définition. *Soit G un groupe et X un sous-ensemble non-vide de G.*

(i) *L'intersection de tous les sous-groupes de G contenant X est un sous-groupe de G.*

(ii) *C'est le plus petit sous-groupe de G qui contient X (pour la relation d'inclusion).*

(iii) *Ses éléments sont les produits d'un nombre fini d'éléments de X et d'inverses d'éléments de X.*

Ce sous-groupe est appelé le sous-groupe de G engendré par X ; on le note $\langle X \rangle$.

Preuve. Les points (i) et (ii) découlent directement de 1.2.6. Pour (iii), notons G' le sous-ensemble de G formé de tous les produits d'un nombre fini d'éléments de X et d'inverses d'éléments de X. Il est clair que G' est un sous-groupe de G et qu'il contient X. Soit maintenant H un sous-groupe quelconque de G contenant X. Comme H est un sous-groupe, il contient tout produit d'un nombre fini d'éléments de H et d'inverses d'éléments de H ; puisque les éléments de X sont des éléments de H, il en résulte que $G' \subseteq H$. Ceci prouve que G' est le plus petit sous-groupe de G qui contient X, d'où le point (iii). □

1.3 Cas particulier des groupes finis

1.3.1 Définitions et notation. On appelle *groupe fini* un groupe G qui, en tant qu'ensemble, n'a qu'un nombre fini d'éléments. Ce nombre d'éléments (qui n'est autre que le cardinal de l'ensemble G) est appelé l'*ordre* du groupe G, noté $o(G)$ ou $|G|$.

1.3.2 Exemples.

(a) Soit n un entier strictement positif. Soit X un ensemble fini à n éléments. Le groupe $\mathcal{S}(X)$ des bijections de X sur X est alors un groupe fini d'ordre $n!$, que l'on appelle (indépendamment de l'ensemble X) le *groupe symétrique* sur n éléments, et que l'on note S_n.

La leçon 4 sera intégralement consacrée à l'étude de ce groupe. On trouvera ci-dessous en 1.3.5.d une description élémentaire du groupe S_3.

(b) Soit n un entier strictement positif. Le sous-groupe $\mathbb{U}_n$ des racines n-ièmes de l'unité dans $\mathbb{C}^*$ est fini, d'ordre n. On peut expliciter $\mathbb{U}_n = \{1, e^{2i\pi/n}, e^{4i\pi/n}, e^{6i\pi/n}, \ldots, e^{2(n-1)i\pi/n}\}$

1.3.3 Théorème (dit théorème de Lagrange) *Soit H un sous-groupe d'un groupe fini G. Alors H est fini, et l'ordre de H divise l'ordre de G.*

Preuve. Notons $|G| = n$. Il est clair que H est fini. Notons $|H| = m$. Pour tout $x \in G$, notons $xH = \{xh\,;\, h \in H\}$ (ce sous-ensemble est appelé la classe de x à gauche modulo H).

D'une part, pour tout $x \in G$, l'ensemble xH est formé de m éléments.

> En effet, si l'on note $H = \{h_1, h_2, \ldots, h_m\}$, alors xH est l'ensemble des éléments de la forme xh_i pour $1 \leq i \leq m$, et $xh_i \neq xh_j$ lorsque $i \neq j$ (car $xh_i = xh_j$ implique $x^{-1}xh_i = x^{-1}xh_j$ donc $h_i = h_j$ donc $i = j$).

D'autre part, l'ensemble des classes xH distinctes obtenues lorsque x décrit G est une partition de G.

En effet, tout $x \in G$ s'écrit $x = xe$ avec $e \in H$, donc $x \in xH$; ceci prouve que G est inclus dans la réunion des classes xH, et donc lui est égal puisque l'inclusion réciproque est triviale. Il reste à vérifier que deux classes xH et yH distinctes sont forcément disjointes. Pour cela, supposons qu'il existe $z \in xH \cap yH$, c'est-à-dire qu'il existe $h', h'' \in H$ tels que $z = xh' = yh''$. Tout élément xh de xH (avec $h \in H$) s'écrit alors $xh = (yh''h'^{-1})h = y(h''h'^{-1}h)$ avec $(h''h'^{-1}h) \in H$, et donc $xh \in yH$. On conclut que $xH \subseteq yH$. L'inclusion réciproque s'obtient de même et l'on déduit que $xH = yH$. On a ainsi prouvé que deux classes non disjointes sont égales, d'où le résultat voulu par contraposée.

On conclut que $n = mq$, où q désigne le nombre de classes xH distinctes obtenues lorsque x décrit G.□

1.3.4 Remarques. On peut représenter un groupe fini G d'ordre n par un tableau à n lignes et n colonnes portant dans la case d'intersection de la ligne indexé par un élément x de G et de la colonne indexé par un élément y de G la valeur du produit $x.y$. Il est facile de vérifier que tout élément de G apparaît une fois et une seule dans chaque ligne et chaque colonne de la table. Il est clair enfin qu'un groupe fini est abélien si et seulement si sa table est symétrique par rapport à la diagonale principale.

1.3.5 Exemples.

(a) Les tables des groupes $\mathbb{U}_2 = \{-1, 1\}$, $\mathbb{U}_3 = \{1, j, j^2\}$, $\mathbb{U}_4 = \{1, i, -1, -i\}$ sont:

	1	-1
1	1	-1
-1	-1	1

	1	j	j^2
1	1	j	j^2
j	j	j^2	1
j^2	j^2	1	j

	1	i	-1	$-i$
1	1	i	-1	$-i$
i	i	-1	$-i$	1
-1	-1	$-i$	1	i
$-i$	$-i$	1	i	-1

(b) Dans $\mathrm{GL}_2(\mathbb{R})$, notons:
$e = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, $a = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$, $b = \left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, $c = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$.
Alors $G_1 = \{e, a, b, c\}$ est un sous-groupe de $\mathrm{GL}_2(\mathbb{R})$ dont la table est:

G_1	e	a	b	c
e	e	a	b	c
a	a	b	c	e
b	b	c	e	a
c	c	e	a	b

(c) Dans $\mathrm{GL}_2(\mathbb{R})$, notons:
$e = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, $a = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, $b = \left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, $c = \left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$.
Alors $G_2 = \{e, a, b, c\}$ est un sous-groupe de $\mathrm{GL}_2(\mathbb{R})$ dont la table est:

G_2	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

(d) Le groupe symétrique S_3 est d'ordre $3! = 6$. On peut décrire explicitement ses six éléments. On convient pour cela de noter chaque élément $\sigma \in S_3$ comme une matrice à 2 lignes et 3 colonnes, où la seconde ligne indique les images respectives par σ de trois éléments arbitraires désignés par les entiers 1,2,3 sur la première ligne. On a alors $S_3 = \{e, \gamma, \gamma^2, \tau_1, \tau_2, \tau_3\}$ avec:

$$e = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{smallmatrix}\right), \quad \gamma = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{smallmatrix}\right), \quad \gamma^2 = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{smallmatrix}\right), \quad \tau_1 = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{smallmatrix}\right), \quad \tau_2 = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{smallmatrix}\right), \quad \tau_3 = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{smallmatrix}\right),$$

On peut alors dresser la table du groupe symétrique S_3.
On en déduit en particulier que le groupe S_3 n'est pas abélien.
On en tire aussi que le groupe S_3 admet trois sous-groupes d'ordre 2 qui sont $\{e, \tau_1\}$, $\{e, \tau_2\}$ et $\{e, \tau_3\}$, et un sous-groupe d'ordre 3 qui est $\{e, \gamma, \gamma^2\}$.
D'après le théorème de Lagrange, ce sont, avec le sous-groupe trivial $\{e\}$ et S_3 lui-même, ses seuls sous-groupes.

	e	γ	γ^2	τ_1	τ_2	τ_3
e	e	γ	γ^2	τ_1	τ_2	τ_3
γ	γ	γ^2	e	τ_3	τ_1	τ_2
γ^2	γ^2	e	γ	τ_2	τ_3	τ_1
τ_1	τ_1	τ_2	τ_3	e	γ	γ^2
τ_2	τ_2	τ_3	τ_1	γ^2	e	γ
τ_3	τ_3	τ_1	τ_2	γ	γ^2	e

Leçon 2

Morphismes de groupes, produit direct de groupes

2.1 Notion de morphisme de groupes

2.1.1 Définition. Soient G un groupe muni d'une loi de composition interne . et G' un groupe muni d'une loi de composition interne $*$. On appelle *morphisme de groupes*, ou *homomorphisme de groupes* de G dans G' toute application $f : G \to G'$ telle que:

$$f(x.y) = f(x) * f(y) \quad \text{pour tous } x, y \in G.$$

2.1.2 Exemples.
(a) L'application dét : $\mathrm{GL}_n(\mathbb{R}) \to \mathbb{R}^*$ qui à toute matrice carrée d'ordre n inversible associe son déterminant est un morphisme de groupes de $\mathrm{GL}_n(\mathbb{R})$ muni du produit matriciel dans $\mathbb{R}^*$ muni de la multiplication, car:

$$\text{dét}(A \times B) = \text{dét}\, A \,.\, \text{dét}\, B, \text{ pour toutes } A, B \in \mathrm{GL}_n(\mathbb{R}).$$

(b) L'application $\exp : \mathbb{R} \to \mathbb{R}_+^*$ qui à tout nombre réel associe son exponentielle est un morphisme de groupes de $\mathbb{R}$ muni de l'addition dans $\mathbb{R}_+^*$ muni de la multiplication, car:

$$\exp(x + y) = \exp x \,.\, \exp y, \text{ pour tous } x, y \in \mathbb{R}.$$

2.1.3 Convention et remarques. Comme on a convenu précédemment de noter les groupes multiplicativement, on continuera à utiliser le point de multiplication pour désigner aussi bien la loi de groupe de G que celle de G'. La condition caractérisant le fait qu'une application $f : G \to G'$ est un morphisme de groupes devient alors:

$$f(x.y) = f(x).f(y) \quad \text{pour tous } x, y \in G,$$

en prenant garde que le point désigne à gauche la loi de G et à droite celle de G'. Il est immédiat de vérifier alors que, pour un tel morphisme de groupes $f : G \to G'$, on a:

(i) $f(e) = e'$, où e désigne le neutre de G et e' celui de G',
(ii) $f(x^{-1}) = f(x)^{-1}$, pour tout $x \in G$,
(iii) $f(x^n) = f(x)^n$, pour tout $x \in G$ et tout $n \in \mathbb{Z}$.

2.1.4 Proposition. *L'image directe d'un sous-groupe et l'image réciproque d'un sous-groupe par un morphisme de groupes sont des sous-groupes. Plus précisément, si G et G' sont deux groupes et si $f : G \to G'$ est un morphisme de groupes, on a:*

(i) *pour tout sous-groupe H de G, l'image directe*

$$f(H) = \{x' \in G' \,;\, \exists\, x \in H,\ f(x) = x'\} = \{f(x) \,;\, x \in H\}$$

est un sous-groupe de G';

(ii) *pour tout sous-groupe H' de G', l'image réciproque*

$$f^{-1}(H') = \{x \in G \,;\, f(x) \in H'\}$$

est un sous-groupe de G.

Preuve. On montre le point (ii) en laissant au lecteur le soin de rédiger de même la preuve du (i). Considérons donc un sous-groupe H' de G', posons $H = f^{-1}(H')$, et montrons que H est un sous-groupe de G. Comme $f(e) = e'$ d'après 2.1.3.(i) et que $e' \in H'$ puisque H' est un sous-groupe de G', on a $e \in H$, et en particulier H n'est pas vide. Soient x et y deux éléments quelconques de H. On a donc $f(x) \in H'$ et $f(y) \in H'$, d'où $f(x).f(y)^{-1} \in H'$ car H' est un sous-groupe de G'. Or en utilisant 2.1.3.(ii), on a $f(x).f(y)^{-1} = f(x.y^{-1})$. On conclut que $f(x.y^{-1}) \in H'$, c'est-à-dire $x.y^{-1} \in H$, ce qui prouve le résultat voulu. □

2.1.5 PROPOSITION. *La composée de deux morphismes de groupes est encore un morphisme de groupes. Plus précisément, si G, G' et G'' sont trois groupes, et si $f : G \to G'$ et $g : G' \to G''$ sont des morphismes de groupes, alors $g \circ f : G \to G''$ est un morphisme de groupes.*

Preuve. Evidente, laissée au lecteur. □

2.2 Image et noyau

2.2.1 PROPOSITION ET DÉFINITION. *Soit $f : G \to G'$ un morphisme de groupes.*

(i) *l'ensemble $f(G) = \{x' \in G' \,;\, \exists\, x \in G,\ f(x) = x'\} = \{f(x) \,;\, x \in G\}$ est un sous-groupe de G' appelé l'image de f, et noté* $\operatorname{Im} f$*;*

(ii) *l'ensemble $f^{-1}(\{e'\}) = \{x \in G \,;\, f(x) = e'\}$ est un sous-groupe de G, appelé le noyau de f, et noté* $\operatorname{Ker} f$*.*

Preuve. Il suffit d'appliquer la proposition 2.1.4 avec $H = G$ et $H' = \{e'\}$. □

2.2.2 PROPOSITION. *Soit $f : G \to G'$ un morphisme de groupes.*

(i) *f est surjective si et seulement si* $\operatorname{Im} f = G'$*;*

(ii) *f est injective si et seulement si* $\operatorname{Ker} f = \{e\}$*.*

Preuve. Le point (i) est immédiat par définition même de la surjectivité. Pour montrer le (ii), supposons d'abord que f est injective. Soit $x \in \operatorname{Ker} f$. On a $f(x) = e'$, et puisque $f(e) = e'$ comme on l'a vu en 2.1.3.(i), on déduit $f(x) = f(e)$, qui implique $x = e$ par injectivité de f. On conclut que $\operatorname{Ker} f = \{e\}$. Réciproquement, supposons que $\operatorname{Ker} f = \{e\}$ et montrons que f est injective. Pour cela, considérons $x, y \in G$ tels que $f(x) = f(y)$. On a alors $f(x).f(y)^{-1} = e'$, donc $f(x.y^{-1}) = e'$, c'est-à-dire $x.y^{-1} \in \operatorname{Ker} f$. L'hypothèse $\operatorname{Ker} f = \{e\}$ implique alors $x.y^{-1} = e$, d'où $x = y$. L'injectivité de f est ainsi montrée, ce qui achève la preuve. □

2.2.3 EXEMPLES. Reprenons les exemples 2.1.2.

(a) Le noyau du morphisme dét : $\mathrm{GL}_n(\mathbb{R}) \to \mathbb{R}^*$ est $\operatorname{Ker} \text{dét} = \{A \in \mathrm{GL}_n(\mathbb{R}) \,;\, \text{dét}\, A = 1\} = \mathrm{SL}_n(\mathbb{R})$ qui, dès lors que $n \geq 2$, n'est pas réduit à $\{I_n\}$; donc le morphisme dét n'est pas injectif. En revanche, il est clair que, quel que soit un réel $x \in \mathbb{R}^*$, on peut trouver une matrice $A \in \mathrm{GL}_n(\mathbb{R})$ telle que $\text{dét}\, A = x$, ce qui prouve l'égalité $\operatorname{Im} \text{dét} = \mathbb{R}^*$ et la surjectivité de dét.

(b) Le morphisme $\exp : \mathbb{R} \to \mathbb{R}_+^*$ est surjectif; son noyau est $\operatorname{Ker} \exp = \{x \in \mathbb{R} \,;\, \exp(x) = 1\} = \{0\}$, ce qui prouve qu'il est aussi injectif.

2.3 Isomorphismes de groupes

2.3.1 DÉFINITION. Soient G et G' deux groupes. On appelle *isomorphisme de groupe* de G sur G' tout morphisme de groupes $f : G \to G'$ qui est de plus une bijection de G sur G'.
L'exemple 2.2.3.(b) ci-dessus est un exemple d'isomorphisme de groupes.

2.3.2 PROPOSITION. *Si f est un isomorphisme de groupes de G sur G', alors la bijection réciproque f^{-1} est un isomorphisme de groupes de G' sur G.*

Preuve. Soient x' et y' deux éléments quelconques de G'. Posons $x = f^{-1}(x')$ et $y = f^{-1}(y')$. Parce que f est un morphisme de groupes, on a $f(x.y) = f(x).f(y)$, donc $f(x.y) = x'.y'$, d'où $x.y = f^{-1}(x'.y')$, c'est-à-dire $f^{-1}(x').f^{-1}(y') = f^{-1}(x'.y')$. Ceci prouve que f^{-1} est un morphisme de groupes de G' sur G, ce qui achève la preuve. □

2.3.3 DÉFINITION. Soient G et G' deux groupes. On dit que G et G' sont *isomorphes* lorsqu'il existe un isomorphisme de groupes de G sur G'. On note $G \simeq G'$.

2.3.4 REMARQUES IMPORTANTES.

(a) Soient G et G' deux groupes isomorphes, et f un isomorphisme de G sur G'. Tout élément de G correspond par f à un et un seul élément de G' (et réciproquement), et ceci de telle façon que toute égalité vérifiée dans G par certains éléments sera vérifiée à l'identique dans G' par les images de ces derniers par f.

Si par exemple x est d'ordre fini n dans G, alors $x^n = e$ et $x^m \neq e$ pour tout $1 \leq m < n$; l'élément $f(x)$ de G' vérifie $f(x)^n = e'$ et $f(x)^m \neq e'$ pour tout $1 \leq m < n$, et donc $f(x)$ est aussi d'ordre n.

Si par exemple deux éléments x et y commutent dans G, c'est-à-dire vérifient $x.y = y.x$, alors on a dans G' l'égalité $f(x).f(y) = f(y).f(x)$, de sorte que les éléments $f(x)$ et $f(y)$ commutent dans G'.

(b) Il en résulte que deux groupes isomorphes ont exactement les mêmes propriétés algébriques. C'est pourquoi on exprime souvent l'isomorphisme de deux groupes G et G' en disant qu'il s'agit du même groupe (en fait c'est le même groupe "à isomorphisme près"), indépendamment de la réalisation concrète que l'on rencontre.

• Par exemple, le sous-groupe $G_1 = \{\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right), \left(\begin{smallmatrix}0&1\\-1&0\end{smallmatrix}\right), \left(\begin{smallmatrix}-1&0\\0&-1\end{smallmatrix}\right), \left(\begin{smallmatrix}0&-1\\1&0\end{smallmatrix}\right)\}$ de $\mathrm{GL}_2(\mathbb{R})$ étudié en 1.3.5.b et le sous-groupe $\mathbb{U}_4 = \{1, i, -1, -i\}$ de $\mathbb{C}^*$ sont isomorphes.

Ce sont en fait deux réalisations concrètes du même groupe abstrait, à savoir le groupe cyclique $C_4 = \{e, x, x^2, x^3\}$ d'ordre 4 tel qu'on l'étudiera dans la leçon 3 suivante (il suffit de poser $x = i$ dans le premier cas, et $x = \left(\begin{smallmatrix}0&1\\-1&0\end{smallmatrix}\right)$ dans le second).

• En revanche, le sous-groupe $G_2 = \{\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right), \left(\begin{smallmatrix}1&0\\0&-1\end{smallmatrix}\right), \left(\begin{smallmatrix}-1&0\\0&-1\end{smallmatrix}\right), \left(\begin{smallmatrix}-1&0\\0&1\end{smallmatrix}\right)\}$ étudié en 1.3.5.c ne leur est pas isomorphe (car il contient trois éléments d'ordre 2 alors que C_4 n'en contient qu'un seul). Il s'agit donc réellement d'un autre groupe.

2.4 Automorphismes de groupes

2.4.1 Définition. Soit G un groupe. On appelle *automorphisme* de G tout morphisme de groupes de G dans G qui est une bijection de G sur G.

• En d'autres termes, un automorphisme de groupe est un isomorphisme de groupes dont le groupe d'arrivée est le même que le groupe de départ.

• Il est clair, d'après la proposition 2.3.2, que la bijection réciproque d'un automorphisme de G est elle-même un automorphisme de G.

2.4.2 Exemples. L'application $\gamma : \mathbb{C} \to \mathbb{C}$ définie par $z \mapsto \gamma(z) = \overline{z}$ est un automorphisme du groupe $\mathbb{C}$ muni de l'addition; il vérifie $\gamma^{-1} = \gamma$. L'application $c : \mathbb{R}_+^* \to \mathbb{R}_+^*$ définie par $x \mapsto c(x) = x^2$ est un automorphisme du groupe $\mathbb{R}_+^*$ muni de la multiplication; sa bijection réciproque est l'automorphisme $c^{-1} : \mathbb{R}_+^* \to \mathbb{R}_+^*$ défini par $x \mapsto c^{-1}(x) = \sqrt{x}$.

2.4.3 Proposition et définition. *Soit G un groupe. L'ensemble des automorphismes du groupe G est un groupe pour la loi $\circ$, dont l'élément neutre est id_G. On le note $\mathrm{Aut}\, G$.*

Preuve. On montre que $\mathrm{Aut}\, G$ est un sous-groupe du groupe $\mathcal{S}(G)$ des bijections de l'ensemble G sur lui-même. Il est clair que $\mathrm{Aut}\, G \subset \mathcal{S}(G)$. L'ensemble $\mathrm{Aut}\, G$ n'est pas vide car $\mathrm{id}_G \in \mathrm{Aut}\, G$. Si $f, g \in \mathrm{Aut}\, G$, alors $f \circ g$ est bijectif (comme composé de deux bijections) et est un morphisme de groupes (d'après 2.1.5), donc $f \circ g \in \mathrm{Aut}\, G$. Ainsi $\mathrm{Aut}\, G$ est stable pour la loi $\circ$. Enfin, si $f \in \mathrm{Aut}\, G$, on a $f^{-1} \in \mathrm{Aut}\, G$ d'après la seconde remarque de 2.4.1, ce qui achève la preuve.

CONVENTION. – On a précédemment convenu de noter les groupes multiplicativement. Désormais, on s'autorisera aussi à ne pas écrire le point de multiplication s'il n'est pas absolument nécessaire à la compréhension; on notera donc xy pour $x.y$ le produit de deux éléments par la loi du groupe.

Comme pour toutes les structures algébriques, on a pour les groupes une notion de structure produit permettant, à partir de deux groupes (ou plus), de construire un nouveau groupe contenant chacun des groupes de départ comme sous-groupe (en un sens précisé ci-dessous en 2.6.4). Il existe plusieurs façons de faire le produit de deux groupes ; la plus simple est le produit direct, à laquelle sont consacrées les deux derniers paragraphes de cette leçon (voir leçon 9 pour le produit semi-direct)

2.5 Produit direct (externe) de deux groupes

2.5.1 Proposition et définition. *Soient G_1 et G_2 deux groupes, de neutres respectifs e_1 et e_2.*

(i) *Le produit cartésien $G_1 \times G_2 = \{(x_1, x_2),\ x_1 \in G_1,\ x_2 \in G_2\}$ est un groupe pour la loi:*

$$(x_1, x_2).(y_1, y_2) = (x_1 y_1, x_2 y_2) \quad \text{pour tous } x_1, y_1 \in G_1,\ x_2, y_2 \in G_2.$$

Ce groupe est appelé le produit direct de G_1 par G_2. On le note $G = G_1 \times G_2$. Son neutre est (e_1, e_2).

(ii) *L'application $p_1 : G_1 \times G_2 \to G_1$ qui, à tout élément $(x_1, x_2) \in G_1 \times G_2$, associe sa première composante x_1, est un morphisme de groupes (appelé première projection).*

(iii) *L'application $p_2 : G_1 \times G_2 \to G_2$ qui, à tout élément $(x_1, x_2) \in G_1 \times G_2$, associe sa seconde composante x_2, est un morphisme de groupes (appelé seconde projection).*

Preuve. Simple vérification, laissée au lecteur. □

2.5.2 Remarques. Il est clair que:

(a) $G_1 \times G_2$ est fini si et seulement si G_1 et G_2 le sont; on a alors $|G_1 \times G_2| = |G_1| \times |G_2|$;
(b) $G_1 \times G_2$ est abélien si et seulement si G_1 et G_2 le sont;
(c) le produit direct $G_1 \times G_2$ est isomorphe au produit direct $G_2 \times G_1$;
(d) on définit de même de façon évidente le produit direct d'un nombre fini quelconque de groupes.

2.5.3 Exemples. Le produit direct $\mathbb{Z} \times \mathbb{Z}$ n'est autre que le groupe additif $\mathbb{Z}^2 = \{(x, y)\,;\, x, y \in \mathbb{Z}\}$ des couples d'entiers muni de l'addition classique $(x, y) + (x', y') = (x + x', y + y')$. De même pour les groupes additifs de couples $\mathbb{R}^2$ ou $\mathbb{C}^2$, et plus généralement $\mathbb{Z}^n$, $\mathbb{R}^n$ ou $\mathbb{C}^n$ quel que soit $n \geq 2$.

2.6 Produit direct (interne) de deux sous-groupes

2.6.1 Notation et remarques. Soient G un groupe, H et K deux sous-groupes de G. On note HK le sous-ensemble de G formé des éléments qui s'écrivent comme le produit d'un élément de H par un élément de K.

$$HK = \{hk\,;\, h \in H, k \in K\}.$$

(a) Si $H \cap K = \{e\}$, tout élément de HK s'écrit *de façon unique* sous la forme hk avec $h \in H, k \in K$.

En effet, si $h_1 k_1 = h_2 k_2$ avec $h_1, h_2 \in H$ et $k_1, k_2 \in K$, on a $h_2^{-1} h_1 = k_2 k_1^{-1}$. Le premier produit est dans H puisque H est un sous-groupe, et le second est dans K puisque K est un sous-groupe. Donc $h_2^{-1} h_1 = k_2 k_1^{-1} \in H \cap K$, c'est-à-dire $h_2^{-1} h_1 = k_2 k_1^{-1} = e$, et donc $h_2 = h_1$ et $k_2 = k_1$.

(b) Si $H \cap K = \{e\}$, et si H et K sont finis, alors HK est fini et $|HK| = |H| \times |K|$.

En effet, il résulte du point précédent que $H \times K$ est alors équipotent à HK, via la bijection $(h, k) \mapsto hk$.

(c) On a $HK = KH$ si et seulement si, quels que soient $h \in H, k \in K$, il existe $h' \in H, k' \in K$ tels que $hk = k'h'$, et il existe $h'' \in H, k'' \in K$ tels que $kh = h''k''$.
Attention, cela n'implique pas que tout élément de H commute avec tout élément de K.

2.6.2 Définition. Soient G un groupe, H et K deux sous-groupes de G. On dit que G est le produit direct (interne) de H par K lorsque les trois conditions suivantes sont vérifiées:

$$(1)\ \ G = HK, \qquad (2)\ \ H \cap K = \{e\}, \qquad (3)\ \ \forall\ h \in H,\ \forall\ k \in K,\ hk = kh.$$

Il est clair qu'alors, on a aussi $G = KH$ et que G est donc le produit direct de K par H.

(a) *Exemple.* Soient: $G = \{ \left(\begin{smallmatrix} 1 & 0 & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{smallmatrix}\right) \, ; \, b, c \in \mathbb{R} \}$, $H = \{ \left(\begin{smallmatrix} 1 & 0 & 0 \\ 0 & 1 & b \\ 0 & 0 & 1 \end{smallmatrix}\right) \, ; \, b \in \mathbb{R} \}$, $K = \{ \left(\begin{smallmatrix} 1 & 0 & c \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right) \, ; \, c \in \mathbb{R} \}$,

Montrer que G est un sous-groupe de $\mathrm{GL}_3(\mathbb{R})$, que H et K sont des sous-groupes de G, et que G est le produit direct de H par K.

(b) *Exemple.* Soit: $G = \{ \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 & 0 \\ 0 & j \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 & 0 \\ 0 & j^2 \end{smallmatrix}\right), \left(\begin{smallmatrix} j & 0 \\ 0 & 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} j^2 & 0 \\ 0 & 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} j & 0 \\ 0 & j \end{smallmatrix}\right), \left(\begin{smallmatrix} j & 0 \\ 0 & j^2 \end{smallmatrix}\right), \left(\begin{smallmatrix} j^2 & 0 \\ 0 & j \end{smallmatrix}\right), \left(\begin{smallmatrix} j^2 & 0 \\ 0 & j^2 \end{smallmatrix}\right) \}$.

Montrer que G est un sous-groupe de $\mathrm{GL}_2(\mathbb{C})$, que les puissance de $x = \left(\begin{smallmatrix} 1 & 0 \\ 0 & j \end{smallmatrix}\right)$ forment un sous-groupe H d'ordre 3, que les puissances de $y = \left(\begin{smallmatrix} j & 0 \\ 0 & 1 \end{smallmatrix}\right)$ forment un autre sous-groupe K d'ordre 3, et que G est le produit direct de H par K.

2.6.3 Remarque. Soient G un groupe, H et K deux sous-groupes de G. Si G est le produit direct de H et K, alors tout élément de G s'écrit de façon unique comme le produit d'un élément de H par un élément de K. Cela découle des conditions (1) et (2) de la définition 2.6.2, et de la remarque (a) de 2.6.1. Attention, la réciproque est fausse (voir plus loin à la leçon 9 la notion de produit semi-direct).

Les notions de produit direct externe de deux groupes (vue en 2.5) et de produit direct interne de deux sous-groupes d'un groupe (vue ci-dessus) sont en fait deux formulations d'une même notion, comme le montre la proposition suivante.

2.6.4 Proposition.

(i) *Soient G_1 et G_2 deux groupes de neutres respectifs e_1 et e_2, et $G = G_1 \times G_2$ leur produit direct. Posons:*

$$H = G_1 \times \{e_2\} = \{(x_1, e_2) \, ; \, x_1 \in G_1\} \quad \text{et} \quad K = \{e_1\} \times G_2 = \{(e_1, x_2) \, ; \, x_2 \in G_2\}.$$

Alors H est un sous-groupe de G isomorphe à G_1, K est un sous-groupe de G isomorphe à G_2, et G est le produit direct interne de ses sous-groupes H et K.

(ii) *Réciproquement, si H et K sont deux sous-groupes d'un groupe G tels que G soit le produit direct interne de H par K, alors G est isomorphe au produit direct externe des groupes H et K.*

Preuve. Simple vérification, laissée au lecteur. □

Leçon 3

Groupes monogènes, groupes cycliques

3.1 Sous-groupe engendré par un élément

3.1.1 Remarque introductive. On a vu en 1.2.8 la notion de sous-groupe engendré par un sous-ensemble X dans un groupe G. L'objet de cette leçon est de développer le cas particulier où X est formé d'un seul élément.

3.1.2 Définition et proposition. *Soit G un groupe. Soit x un élément de G. On appelle sous-groupe monogène engendré par x dans G le sous-groupe engendré par le singleton $\{x\}$. On le note $\langle x\rangle$. C'est le plus petit sous-groupe de G contenant x, et l'on a:*

$$\langle x\rangle = \{x^m\,;\, m \in \mathbb{Z}\}.$$

Preuve. Il suffit a priori d'appliquer la proposition 1.2.8 avec $X = \{x\}$. Pour une bonne compréhension, reprenons la preuve dans ce cas particulier.

Le sous-groupe $\langle x\rangle$ contient x, donc (par stabilité pour la loi de G) il contient aussi $xx = x^2$, $x^2x = x^3$, et par récurrence x^m pour tout entier $m \geq 1$. Il contient aussi nécessairement le symétrique x^{-1} de x, donc aussi $x^{-1}x^{-1} = x^{-2}$, et par récurrence x^{-m} pour tout entier $m \geq 1$. Enfin il contient le neutre $e = xx^{-1}$ que l'on note par convention x^0. Ceci montre que $\langle x\rangle \supset \{x^m\,;\, m \in \mathbb{Z}\}$. Il est clair réciproquement que $\{x^m\,;\, m \in \mathbb{Z}\}$ est un sous-groupe de G contenant x. □

3.1.3 Remarque. Attention: l'énoncé précédent est formulé pour la notation multiplicative du groupe G. Dans le cas d'une loi notée comme une addition, il faut remplacer x^n par $nx = x + x + \cdots + x$ et x^{-1} par $-x$. Par exemple, dans le groupe $\mathbb{Z}$ muni de l'addition, $\langle x\rangle = \{mx\,;\, m \in \mathbb{Z}\}$.

3.1.4 Définition. Soit G un groupe. Soit x un élément de G. On dit que x est d'ordre fini dans G lorsqu'il existe des entiers $m \geq 1$ tel que $x^m = e$. Dans ce cas, on appelle ordre de x le plus petit d'entre eux. En d'autres termes:

$$(x \text{ est d'ordre } n \text{ dans } G) \Leftrightarrow (x^n = e \text{ et } x^m \neq e \text{ si } 1 \leq m < n).$$

Remarquons qu'alors le symétrique de x est $x^{-1} = x^{n-1}$.

3.1.5 Proposition. *Soit G un groupe. Soit x un élément de G. Si x est d'ordre fini $n \geq 1$ dans G, alors le sous-groupe $\langle x\rangle$ est fini d'ordre n, et l'on a:*

$$\langle x\rangle = \{e, x, x^2, x^3, \ldots, x^{n-1}\}.$$

Preuve. Soit x^m avec $m \in \mathbb{Z}$ un élément quelconque de $\langle x\rangle$. Par division euclidienne de m par n, il existe des entiers uniques q et r tels que $m = nq + r$ avec $0 \leq r \leq n-1$. On a $x^m = x^{nq+r} = (x^n)^q x^r = e^q x^r = x^r$, ce qui prouve que $\langle x\rangle$ est inclus dans l'ensemble $E := \{x^r\,;\, 0 \leq r \leq n-1\}$. La réciproque étant claire, on a $\langle x\rangle = E$. Il reste à vérifier que E est formé des n éléments distincts $e, x, x^2, x^3, \ldots, x^{n-1}$. Pour cela, supposons que $x^i = x^j$ avec $0 \leq i, j \leq n-1$; alors $x^{i-j} = e$ avec $-n < i - j < n$, ce qui, par minimalité de l'ordre n de x, implique $i - j = 0$ et donc $i = j$. On a donc bien $E = \{e, x, x^2, x^3, \ldots, x^{n-1}\}$, ce qui achève la preuve. □

3.1.6 Remarques.

(a) Il résulte de la proposition précédente et du théorème de Lagrange que, si le groupe G est fini, tout élément est d'ordre fini divisant $|G|$.

(b) Si x n'est pas d'ordre fini, le sous-groupe $\langle x\rangle$ n'est pas fini, ce qui ne peut se produire que si G est lui-même infini.

(c) Mais réciproquement, un groupe G infini peut contenir des sous-groupes du type $\langle x\rangle$ finis ou infinis. Par exemple, dans le groupe $\mathbb{C}^*$ pour la multiplication, le groupe $\langle i\rangle = \{1, i, -1, -i\}$ est fini et le groupe $\langle 5\rangle = \{5^m\,;\, m \in \mathbb{Z}\}$ est infini.

3.2 Groupes monogènes, groupes cycliques.

3.2.1 Définitions. Un groupe G est dit *monogène* lorsqu'il est engendré par un de ses éléments, c'est-à-dire lorsqu'il existe un élément $x \in G$ tel que $G = \langle x \rangle$.

Si de plus x est d'ordre fini $n \geq 1$, alors on dit que le groupe G est *cyclique* d'ordre n, et l'on a d'après ce qui précède:

$$G = \{e, x, x^2, x^3, \ldots, x^{n-1}\}.$$

Sinon, $x^i \neq x^j$ pour tous $i \neq j$ dans $\mathbb{Z}$, et $G = \{x^m\,;\, m \in \mathbb{Z}\}$ est monogène infini.
Il est clair qu'un groupe monogène (en particulier un groupe cyclique) est toujours abélien.

3.2.2 Proposition (Sous-groupe d'un groupe monogène infini). *Tout sous-groupe non-trivial d'un groupe monogène infini est monogène infini.*

Preuve. On a $G = \{x^m\,;\, m \in \mathbb{Z}\}$ avec $x \neq e$ qui n'est pas d'ordre fini. Soit H un sous-groupe de G distinct de $\{e\}$. Il existe donc dans H des éléments de la forme x^ℓ avec $\ell \in \mathbb{Z}^*$. Comme l'inverse d'un élément de H appartient à H, on peut préciser qu'il existe dans H des éléments de la forme x^ℓ avec $\ell \in \mathbb{N}^*$. Soit alors d le plus petit entier strictement positif tel que $x^d \in H$. Posons $K = \{x^{dm}\,;\, m \in \mathbb{Z}\}$. Il est clair que $K \subseteq H$ (car $x^d \in H$ et H est stable par produit et passage à l'inverse). Réciproquement, soit x^m un élément quelconque de H (avec $m \in \mathbb{Z}$). Par division euclidienne de m par d, il existe $a, r \in \mathbb{Z}$ uniques tels que $m = ad + r$ avec $0 \leq r < d$. On a $x^r = x^{m-ad} = x^m (x^d)^{-a}$ avec $x^m \in H$ et $(x^d)^{-a} \in K \subset H$, et donc $x^r \in H$. Par minimalité de d, on a donc forcément $r = 0$; d'où $x^m = x^{ad}$ et donc $x^m \in K$. Ceci prouve que $H \subseteq K$. On conclut que $H = K = \langle x^d \rangle$. □

3.2.3 Proposition (Sous-groupe d'un groupe cyclique). *Tout sous-groupe d'un groupe cyclique est cyclique. Plus précisément, si $G = \langle x \rangle$ est un groupe cyclique d'ordre $n \geq 1$, alors il existe pour tout diviseur q de n un et un seul sous-groupe d'ordre q, et c'est le sous-groupe cyclique engendré par x^p où $n = pq$.*

Preuve. On a $G = \{e, x, x^2, x^3, \ldots, x^{n-1}\}$. Il est clair que, si q est un diviseur de n, et si l'on pose $n = pq$ avec $p \in \mathbb{N}^*$, alors $\langle x^p \rangle = \{e, x^p, x^{2p}, x^{3p}, x^{(q-1)p}\}$ est un sous-groupe de G cyclique d'ordre q. Réciproquement, soit H un sous-groupe de G. D'après le théorème de Lagrange, l'ordre q de H doit diviser l'ordre de G. Notons $n = qp$. On peut supposer $H \neq \{e\}$, c'est-à-dire $q \neq 1$. Comme dans la preuve de la proposition précédente, on peut considérer d le plus petit entier $1 \leq d \leq n-1$ tel que $x^d \in H$, et montrer que $H = \langle x^d \rangle = \{e, x^d, x^{2d}, x^{3d}, \ldots, x^{(q-1)d}\}$. En notant par division euclidienne $n = d\ell + s$ avec $0 \leq s < d$, l'égalité $e = x^n = (x^d)^\ell x^s$ implique $x^s = (x^d)^{-\ell}$, qui est un élément de H, d'où $s = 0$ par minimalité de d. Finalement $n = d\ell$. Comme x est d'ordre n [ie. $x^n = e$ et $x^m \neq e$ si $1 \leq m < n$], alors x^d est d'ordre ℓ [ie. $(x^d)^\ell = e$ et $(x^d)^m \neq e$ si $1 \leq m < \ell$]. En d'autres termes $\ell = q$ et donc $d = p$, ce qui achève la preuve. □

3.3 Générateurs d'un groupe cyclique.

3.3.1 Exemple préliminaire.

Figure 1

Dans $\mathbb{C}^*$, considérons $x = \mathrm{e}^{i\pi/3}$, et $G = \{e, x, x^2, x^3, x^4, x^5\}$ le groupe cyclique d'ordre 6 engendré par x. Ce groupe G est le groupe $\mathbb{U}_6$ des racines sixièmes de l'unité dans $\mathbb{C}^*$. Ses éléments $e = 1$, $x = -j^2$, $x^2 = j$, $x^3 = -1$, $x^4 = j^2$, $x^5 = -j$ peuvent être représentés dans le plan complexe comme les sommets respectifs A, B, C, D, E, F d'un hexagone régulier centré en l'origine et inscrit dans le cercle unité.
Considérons dans G les sous-groupes cycliques qu'engendrent les différents éléments. On a bien sûr $\langle e \rangle = \{e\}$ et $\langle x \rangle = G$. De plus $\langle x^2 \rangle = \langle x^4 \rangle = \{e, x^2, x^4\}$ est le sous-groupe d'ordre 3 de G (cf. proposition 3.2.3), qui correspond au triangle ACE. De même $\langle x^3 \rangle = \{e, x^3\}$ est le sous-groupe d'ordre 2 de G, qui correspond au segment AD.

Considérons enfin le sous-groupe engendré par x^5. Il contient $(x^5)^2 = x^{10} = x^4$, $(x^5)^3 = x^{15} = x^3$, $(x^5)^4 = x^{20} = x^2$, $(x^5)^5 = x^{25} = x$, $(x^5)^6 = x^{30} = e$, et donc $\langle x^5 \rangle = G$. L'élément x^5 est, comme x, un générateur du groupe G.
Ce résultat est un cas particulier du théorème suivant.

3.3.2 Théorème. *Soit $G = \langle x \rangle$ un groupe cyclique d'ordre $n \geq 2$. Alors les générateurs de G sont les éléments x^k tels que les entiers k et n soient premiers entre eux.*

Preuve. On a $G = \{e, x, x^2, x^3, \ldots, x^{n-1}\}$. Soit $k \in \mathbb{Z}^*$ et $H = \langle x^k \rangle$. On a $H = G$ si et seulement si $x \in H$ (puisqu'alors H contient toutes les puissances de x et donc tous les éléments de G). Or:

$x \in H \Leftrightarrow$ il existe $u \in \mathbb{Z}$ tel que $x = x^{ku}$
$x \in H \Leftrightarrow$ il existe $u \in \mathbb{Z}$ tel que $x^{ku-1} = e$
$x \in H \Leftrightarrow$ il existe $u \in \mathbb{Z}$ tel que $ku - 1$ est multiple de l'ordre n de x
$x \in H \Leftrightarrow$ il existe $u, v \in \mathbb{Z}$ tel que $ku + nv = 1$.

Cette dernière condition équivaut, d'après le théorème de Bézout, au fait que k et n sont premiers entre eux, ce qui achève la preuve. □

3.3.3 Remarque (Indicatrice d'Euler). On appelle *fonction indicatrice d'Euler* l'application $\varphi : \mathbb{N}^* \to \mathbb{N}^*$ définie par $\varphi(1) = 1$ et, pour tout entier $n \geq 2$:

$\varphi(n)$ est le nombre d'entiers k tels que $1 \leq k \leq n-1$ et k est premier avec n.

D'après le théorème précédent, $\varphi(n)$ est le nombre de générateurs d'un groupe cyclique d'ordre n. Par définition de φ, on peut calculer:

$$\varphi(2) = 1, \ \varphi(3) = 2, \ \varphi(4) = 2, \ \varphi(5) = 4, \ \varphi(6) = 2, \ \varphi(7) = 6, \ \varphi(8) = 4, \ \ldots$$

Il est clair que, pour tout nombre premier p, on a $\varphi(p) = p - 1$.
Montrer en exercice que $\varphi(p^\alpha) = p^\alpha - p^{\alpha-1}$ pour tout nombre premier p et tout entier $\alpha \geq 1$.
On verra plus loin une formule générale permettant de calculer $\varphi(n)$ pour tout entier $n \geq 1$.

3.3.4 Exercice. Montrer que, si $G = \langle x \rangle$ est un groupe monogène infini, alors les seuls générateurs de G sont x et x^{-1}.

3.4 Groupes finis d'ordre premier

Proposition. *Soit G un groupe fini d'ordre premier p. Alors:*

1. *G est cyclique,*
2. *les seuls sous-groupes de G sont $\{e\}$ et G,*
3. *tous les éléments de G distincts de e sont des générateurs de G.*

Preuve. Comme $p > 1$, $G \neq \{e\}$. Soit $x \in G$ quelconque distinct de e. Posons $H = \langle x \rangle$ le sous-groupe de G engendré par x. D'après le théorème de Lagrange, l'ordre q de H doit diviser p. Comme p est premier, et comme $q \neq 1$ puisque $x \neq e$, on a forcément $q = p$. Donc $H = G$, c'est-à-dire $G = \langle x \rangle = \{e, x, x^2, x^3, \ldots, x^{p-1}\}$. Ceci prouve les points 1 et 2, et le point 3 résulte alors immédiatement du théorème 3.3.2. □

3.5 Quelques conséquences concrètes à retenir.

3.5.1. *Tout groupe monogène infini est isomorphe au groupe $\mathbb{Z}$ muni de l'addition.*

En effet, si G est un groupe monogène infini, il existe $x \in G$ tel que $G = \{x^m \, ; \, m \in \mathbb{Z}\}$. Cet élément x n'est pas d'ordre fini dans G (sinon G ne serait pas infini). L'application:

$$\begin{array}{rccc} f: & \mathbb{Z} & \longrightarrow & G \\ & m & \longmapsto & x^m \end{array}$$

est alors un morphisme du groupe $\mathbb{Z}$ muni de l'addition dans le groupe G muni de la loi . (car $f(m+n) = x^{m+n} = x^m.x^n = f(m).f(n)$). Ce morphisme est surjectif par construction. De plus, si $m \in \operatorname{Ker} f$, alors $x^m = e$, ce qui implique $m = 0$ puisque x n'est pas d'ordre fini; ceci prouve que $\operatorname{Ker} f = \{0\}$, donc que f est injectif. On conclut que f est un isomorphisme de groupes.

3.5.2. *Pour tout entier $n \geq 1$, il existe à isomorphisme près un et un seul groupe cyclique d'ordre n. On le note C_n.*

De façon abstraite, on le note multiplicativement $C_n = \{e, x, x^2, x^3, \ldots, x^{n-1}\}$. Une réalisation concrète en est le groupe $\mathbb{U}_n$ des racines n-ièmes de l'unité de $\mathbb{C}^*$, que l'on peut représenter géométriquement comme un polygone régulier à n côtés. On verra plus loin dans le cours une autre réalisation du groupe C_n, notée $\mathbb{Z}/n\mathbb{Z}$, avec une loi additive.

Comme deux groupes finis isomorphes ont évidemment le même ordre, il est clair que $C_n \not\simeq C_m$ dès lors que $n \neq m$.

3.5.3. *Pour tout nombre premier p, il existe à isomorphisme près un et un seul groupe fini d'ordre p ; c'est le groupe cyclique C_p.*

Cela résulte immédiatement de la proposition 3.4 et de ce qui précède.

3.5.4. *Il existe à isomorphisme près deux groupes d'ordre 4, et deux seulement:* l'un est le groupe cyclique C_4, l'autre est noté V et appelé le groupe de Klein; ils sont tous les deux abéliens, et leurs tables respectives sont:

C_4	e	a	b	c
e	e	a	b	c
a	a	b	c	e
b	b	c	e	a
c	c	e	a	b

V	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

En effet, soit $G = \{e, a, b, c\}$ un groupe d'ordre 4 de neutre e. Deux cas peuvent se présenter.

Premier cas: G contient un élément d'ordre 4. Supposons par exemple que ce soit a. Alors G doit contenir a^2 et $a^3 = a^{-1}$ qui sont distincts de a. On a donc $b = a^2$ et $c = a^3$ (ou le contraire), ce qui donne la première table.

Second cas: G ne contient pas d'élément d'ordre 4. Comme e est le seul élément d'ordre 1, et que G ne peut pas contenir d'éléments d'ordre 3 d'après le théorème de Lagrange, c'est que a, b, c sont tous les trois d'ordre 2. Donc $a^2 = b^2 = c^2 = e$, et chacun des trois est son propre inverse. Considérons le produit ab. Si l'on avait $ab = a$, on aurait $b = e$, ce qui est exclu. Si l'on avait $ab = b$, on aurait $a = e$, ce qui est exclu. Si l'on avait $ab = e$, on aurait $b = a^{-1}$, c'est-à-dire $b = a$, ce qui est exclu. On a donc forcément $ab = c$. On calcule de même les autres produits. On obtient la seconde table.

3.5.5. On peut de même démontrer (laissé en exercice) qu'*il existe à isomorphisme près deux groupes d'ordre 6, et deux seulement:* l'un est le groupe cyclique C_6 (et est donc abélien), l'autre est non abélien et est isomorphe par exemple au groupe symétrique S_3 (voir 1.3.5.d).

3.6 Produit direct de groupes cycliques, théorème chinois

3.6.1 Question. Si G_1 et G_2 sont deux groupes cycliques, le produit direct $G_1 \times G_2$ est-il cyclique? Le théorème suivant, dit théorème chinois, répond à cette question.

3.6.2 Premier exemple introductif.
Considérons le groupe cyclique $C_2 = \{e, x\}$ avec $x^2 = e$, et formons le produit direct $C_2 \times C_2$. On établit aisément sa table.

On reconnaît le groupe de Klein $V = \{e, a, b, c\}$ pour: $e = (e, e)$, $a = (e, x)$, $b = (x, e)$ et $c = (x, x)$.

Donc $C_2 \times C_2$ n'est pas cyclique.

	(e,e)	(e,x)	(x,e)	(x,x)
(e,e)	(e,e)	(e,x)	(x,e)	(x,x)
(e,x)	(e,x)	(e,e)	(x,x)	(x,e)
(x,e)	(x,e)	(x,x)	(e,e)	(e,x)
(x,x)	(x,x)	(x,e)	(e,x)	(e,e)

3.6.3 Second exemple introductif.

Considérons les groupes cycliques $C_2 = \{e, x\}$ et $C_3 = \{\varepsilon, y, y^2\}$ et formons le produit direct $C_2 \times C_3$. On établit aisément sa table.

	(e,ε)	(x,y)	(e,y^2)	(x,ε)	(e,y)	(x,y^2)
(e,ε)	(e,ε)	(x,y)	(e,y^2)	(x,ε)	(e,y)	(x,y^2)
(x,y)	(x,y)	(e,y^2)	(x,ε)	(e,y)	(x,y^2)	(e,ε)
(e,y^2)	(e,y^2)	(x,ε)	(e,y)	(x,y^2)	(e,ε)	(x,y)
(x,ε)	(x,ε)	(e,y)	(x,y^2)	(e,ε)	(x,y)	(e,y^2)
(e,y)	(e,y)	(x,y^2)	(e,ε)	(x,y)	(e,y^2)	(x,ε)
(x,y^2)	(x,y^2)	(e,ε)	(x,y)	(e,y^2)	(x,ε)	(e,y)

En posant $z = (x, y)$, on a: $(e, y^2) = z^2$, $(x, \varepsilon) = z^3$, $(e, y) = z^4$, $(x, y^2) = z^5$ et $(e, \varepsilon) = z^6$.
On conclut que $C_2 \times C_3 \simeq C_6$ est cyclique.

3.6.4 THÉORÈME (dit théorème chinois). *Soient G_1 et G_2 deux groupes cycliques d'ordres respectifs n et m. Alors, le produit direct $G_1 \times G_2$ est cyclique si et seulement si les entiers n et m sont premiers entre eux.*

Preuve. Supposons que n et m sont premiers entre eux. Notons $G_1 = \langle x \rangle \simeq C_n$ avec x d'ordre n, et $G_2 = \langle y \rangle \simeq C_m$ avec y d'ordre m. Soit $z = (x, y)$ dans $G_1 \times G_2$. Quel que soit $k \in \mathbb{Z}$, on a $z^k = (e_1, e_2)$ si et seulement si $x^k = e_1$ et $y^k = e_2$, ce qui équivaut à dire que k est multiple à la fois de n et de m. Or le ppcm de n et m est ici nm puisque n et m sont premiers entre eux. Donc $z^{nm} = (e_1, e_2)$ et $z^k \neq (e_1, e_2)$ pour tout $1 \leq k < nm$. On conclut que l'élément z est d'ordre nm dans $G_1 \times G_2$. Or on sait que $G_1 \times G_2$ est formé de nm éléments; on conclut que $G_1 \times G_2 = \langle z \rangle \simeq C_{nm}$. La réciproque est laissée au lecteur. □

Remarque. Avec les notations multiplicatives utilisées ici, le théorème chinois s'énonce donc sous la forme:

$$C_n \times C_m \simeq C_{nm} \iff m \text{ et } n \text{ premiers entre eux.}$$

3.6.5 EXEMPLE D'APPLICATION : structure des groupes abéliens finis

Un résultat classique (non démontré dans ce cours, mais qui pourra faire l'objet d'exercice en TD) établit que tout groupe abélien fini est isomorphe à un produit direct de groupes cycliques. Dès lors, en admettant ce résultat, on déduit du théorème chinois par exemple qu'il existe à isomorphisme près quatre groupes abéliens d'ordre 36, qui sont:

$$C_{36} \simeq C_9 \times C_4, \qquad C_3 \times C_{12} \simeq C_3 \times C_3 \times C_4,$$
$$C_2 \times C_{18} \simeq C_2 \times C_2 \times C_9, \qquad C_6 \times C_6 \simeq C_3 \times C_3 \times C_2 \times C_2 \simeq C_6 \times C_3 \times C_2,$$

ou encore que tous les groupes abéliens d'ordre ≤ 12 sont, à isomorphisme près, donnés par:

$n = 1$	C_1		
$n = 2$	C_2		
$n = 3$	C_3		
$n = 4$	C_4	$C_2 \times C_2$	
$n = 5$	C_5		
$n = 6$	$C_6 \simeq C_2 \times C_3$		
$n = 7$	C_7		
$n = 8$	C_8	$C_4 \times C_2$	$C_2 \times C_2 \times C_2$
$n = 9$	C_9	$C_3 \times C_3$	
$n = 10$	$C_{10} \simeq C_2 \times C_5$		
$n = 11$	C_{11}		
$n = 12$	$C_{12} \simeq C_3 \times C_4$	$C_2 \times C_6 \simeq C_2 \times C_2 \times C_3$	

Leçon 4

Groupe symétrique

4.1 Notion de groupe symétrique.

4.1.1 Remarque préliminaire. Soit n un entier strictement positif. Soit X un ensemble fini à n éléments. On sait que le groupe $\mathcal{S}(X)$ des bijections de X sur X est alors un groupe fini d'ordre $n!$. Si Y est un autre ensemble de même cardinal n, il existe une bijection f de X sur Y et l'on construit de façon évidente un isomorphisme de groupes φ de $\mathcal{S}(X)$ sur $\mathcal{S}(Y)$ en posant $\varphi(\sigma) = f \circ \sigma \circ f^{-1}$ pour tout $\sigma \in \mathcal{S}(X)$. Le groupe $\mathcal{S}(X)$ est donc, à isomorphisme près, indépendant du choix de l'ensemble X, et ne dépend donc que de son cardinal.

4.1.2 Définition et remarque. Pour tout entier $n \geq 1$, on appelle *groupe symétrique sur n éléments*, ou *n-ième groupe symétrique*, le groupe des bijections d'un ensemble fini à n éléments quelconque sur lui-même. On le note S_n.

(a) S_n est un groupe fini, d'ordre $n!$.

(b) Les éléments de S_n sont appelés les permutations sur n éléments. On note une telle permutation sous la forme $\sigma = \begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ \sigma(1) & \sigma(2) & \sigma(3) & \cdots & \sigma(n) \end{pmatrix}$.

(c) Sa loi de composition interne, qui est la composition $\circ$ des bijections est notée multiplicativement, c'est-à-dire que l'on note $\sigma\tau$ au lieu de $\sigma \circ \tau$ pour toutes $\sigma, \tau \in S_n$. On note e l'élément neutre de S_n, qui est l'identité de $\{1, 2, \ldots, n\}$.

(d) Pour $n = 1$, le groupe S_1 est le groupe trivial $\{e\}$ d'ordre 1. Pour $n = 2$, le groupe S_2 est d'ordre 2, donc $S_2 = C_2 = \{e, \tau\}$ où $e = \left(\begin{smallmatrix} 1 & 2 \\ 1 & 2 \end{smallmatrix}\right)$ et $\tau = \left(\begin{smallmatrix} 1 & 2 \\ 2 & 1 \end{smallmatrix}\right)$, qui vérifie bien $\tau^2 = e$.

(e) Dès lors que $n \geq 3$, le groupe S_n n'est pas abélien.

En effet, considérons trois entiers $1 \leq i, j, k \leq n$ distincts deux à deux (ce qui est possible car $n \geq 3$). Posons $\gamma = \begin{pmatrix} i & j & k \\ j & k & i \end{pmatrix}$ et $\tau = \begin{pmatrix} i & j & k \\ i & k & j \end{pmatrix}$, où la notation sous-entend que les éléments autres que i, j, k ont eux-mêmes pour image. On a $\gamma\tau = \begin{pmatrix} i & j & k \\ j & i & k \end{pmatrix}$ et $\tau\gamma = \begin{pmatrix} i & j & k \\ k & j & i \end{pmatrix}$. Donc $\gamma\tau \neq \tau\gamma$.

(f) Pour $n = 3$, le groupe S_3 est d'ordre 6. On a: $S_3 = \{e, \gamma, \gamma^2, \tau_1, \tau_2, \tau_3\}$ avec:

$e = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{smallmatrix}\right)$, $\gamma = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{smallmatrix}\right)$, $\gamma^2 = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{smallmatrix}\right)$, $\tau_1 = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{smallmatrix}\right)$, $\tau_2 = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{smallmatrix}\right)$, $\tau_3 = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{smallmatrix}\right)$.

	e	γ	γ^2	τ_1	τ_2	τ_3
e	e	γ	γ^2	τ_1	τ_2	τ_3
γ	γ	γ^2	e	τ_3	τ_1	τ_2
γ^2	γ^2	e	γ	τ_2	τ_3	τ_1
τ_1	τ_1	τ_2	τ_3	e	γ	γ^2
τ_2	τ_2	τ_3	τ_1	γ^2	e	γ
τ_3	τ_3	τ_1	τ_2	γ	γ^2	e

On a déjà vu en 1.3.5.(d) la table de S_3, rappelée ci-contre.

Le groupe S_3 est le plus petit groupe non abélien (car les groupes d'ordre 2, 3 ou 5 sont abéliens car cycliques d'après 3.5.3, et les deux seuls groupes d'ordre 4 sont abéliens comme on l'a vu en 3.5.4).

S_3 admet trois sous-groupes d'ordre 2 qui sont $\{e, \tau_1\}$, $\{e, \tau_2\}$ et $\{e, \tau_3\}$, et un sous-groupe d'ordre 3 qui est $\{e, \gamma, \gamma^2\}$.

4.2 Décomposition d'une permutation en produit de transpositions.

4.2.1 Définition. On appelle *transposition* de S_n toute permutation τ qui échange deux éléments i et j en laissant fixes les $n - 2$ autres. On note alors $\tau = [i, j]$. On a de façon évidente $\tau^2 = e$, c'est-à-dire $\tau^{-1} = \tau$.

4.2.2 Théorème. *Toute permutation de S_n est un produit d'un nombre fini de transpositions. En d'autres termes, le groupe S_n est engendré par ses transpositions.*

Preuve. On raisonne par récurrence sur n. C'est clair si $n = 2$. Supposons (H.R.) le résultat vrai pour S_{n-1} où $n \geq 3$. Prenons $\sigma \in S_n$ quelconque. Distinguons deux cas. Si $\sigma(n) = n$, notons σ' la restriction de σ à $\{1, 2, \ldots, n-1\}$. Il est clair que $\sigma' \in S_{n-1}$. Donc par H.R., $\sigma' = \tau'_1\tau'_2 \ldots \tau'_m$ où τ'_k est une transposition de $\{1, 2, \ldots, n-1\}$ pour tout $1 \leq k \leq m$. Chaque τ'_k se prolonge en une transposition τ_k de $\{1, 2, \ldots, n\}$ en posant $\tau_k(i) = \tau'_k(i)$ pour tout $1 \leq i \leq n-1$ et $\tau_k(n) = n$. Il est clair que l'on a alors $\sigma = \tau_1\tau_2 \ldots \tau_m$. Si maintenant $\sigma(n) = p \neq n$, posons $\tau = [n, p]$ et $\eta = \tau\sigma$. On a $\eta(n) = n$. En appliquant le premier cas, η se décompose en produit de transpositions. Donc $\sigma = \tau\eta$ aussi. □

4.2.3 Remarque. Il n'y a pas unicité de cette décomposition. Par exemple, dans S_4, on a $\left(\begin{smallmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 2 & 4 \end{smallmatrix}\right) = [2, 4][1, 4][4, 2][1, 3] = [2, 3][1, 2]$.

4.3 Signature

4.3.1 Définitions. Soit $n \geq 2$ un entier. Pour toute permutation $\sigma \in S_n$, on appelle *nombre d'inversions* de σ l'entier:

$$I(\sigma) = |\ \{\ (i, j) \in \{1, 2, \ldots, n\}^2\ ;\ i < j \ \text{ et } \ \sigma(i) > \sigma(j)\ \}\ |.$$

On appelle *signature* de σ l'entier valant $+1$ ou -1 défini par:

$$\varepsilon(\sigma) = (-1)^{I(\sigma)}.$$

4.3.2 Exemple. Si τ est une transposition de S_n, on a $\varepsilon(\tau) = -1$.

En effet, si $\tau = [i, j]$ avec $i < j$, les couples $(u, v) \in \{1, 2, \ldots, n\}^2$ vérifiant $u < v$ et $\sigma(u) > \sigma(v)$ sont exactement les $j-i$ couples $(i, i+1), (i, i+2), \ldots, (i, j)$ et les $j-i-1$ couples $(i+1, j), (i+2, j), \ldots, (j-1, j)$. Donc $I(\sigma) = j - i + j - i - 1 = 2(j - i) - 1$ est impair.

4.3.3 Proposition. *Quelles que soient deux permutations $\sigma, \gamma \in S_n$, on a: $\varepsilon(\gamma\sigma) = \varepsilon(\gamma)\varepsilon(\sigma)$. En d'autres termes, l'application:*

$$\begin{aligned} \varepsilon :\ S_n &\longrightarrow \{+1, -1\} \\ \sigma &\longmapsto \varepsilon(\sigma) \end{aligned}$$

est un morphisme de groupes.

Preuve. Pour toute permutation $\sigma \in S_n$ et toute application $f : \mathbb{Q}^n \to \mathbb{Q}$, on note $\sigma * f$ l'application $\mathbb{Q}^n \to \mathbb{Q}$ définie par: $\sigma * f(x_1, x_2, \ldots, x_n) = f(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(n)})$. Il est clair que, pour toutes $\gamma\sigma \in S_n$, on a $\gamma * (\sigma * f) = (\gamma\sigma) * f$. Considérons en particulier l'application $\Delta : \mathbb{Q}^n \to \mathbb{Q}$ définie par:

$$\Delta(x_1, x_2, \ldots, x_n) = \prod_{1 \leq i < j \leq n} (x_i - x_j).$$

Par définition du nombre d'inversions, on a $\sigma * \Delta(x_1, x_2, \ldots, x_n) = (-1)^{I(\sigma)}\Delta(x_1, x_2, \ldots, x_n)$ pour tous $x_1, x_2, \ldots, x_n \in \mathbb{Q}^n$ et toute $\sigma \in S_n$. Donc $\sigma * \Delta = \varepsilon(\sigma)\Delta$. On déduit que, pour toutes $\sigma, \gamma \in S_n$, on a: $\varepsilon(\gamma\sigma)\Delta = (\gamma\sigma) * \Delta = \gamma * (\sigma * \Delta) = \gamma * (\varepsilon(\sigma)\Delta) = \varepsilon(\sigma)\gamma * \Delta = \varepsilon(\sigma)\varepsilon(\gamma)\Delta$. Comme l'application Δ n'est évidemment pas identiquement nulle, on conclut que $\varepsilon(\gamma\sigma) = \varepsilon(\sigma)\varepsilon(\gamma) = \varepsilon(\gamma)\varepsilon(\sigma)$. □

4.3.4 Corollaire. *Soit $\sigma \in S_n$.*

(i) *Si σ se décompose d'une part en un produit de m transpositions, d'autre part en un produit de m' transpositions, alors les entiers naturels m et m' sont de même parité.*

(ii) *On a $\epsilon(\sigma) = (-1)^m$, où m désigne le nombre de transpositions d'une décomposition quelconque de σ en produit de transpositions.*

Preuve. Résulte immédiatement de 4.2.2, 4.3.2 et 4.3.3. □

4.4 Groupe alterné.

4.4.1. Définition. Pour tout entier $n \geq 2$, le noyau de ε est appelé n-ième groupe alterné. On le note A_n.

Le sous-groupe A_n de S_n est donc l'ensemble des permutations de S_n qui sont de signature 1 (c'est-à-dire qui se décomposent en un nombre pair de transpositions).

4.4.2 PROPOSITION. *Pour tout entier $n \geq 2$, le groupe A_n est fini d'ordre $\frac{n!}{2}$.*

Preuve. Notons X l'ensemble des permutations de S_n qui sont de signature -1. Le sous-ensemble X est non-vide (il contient par exemple les transpositions, voir 4.3.2). Si l'on fixe $\tau \in X$, il est facile de vérifier d'après 4.3.3 que l'application $\sigma \mapsto \tau\sigma$ réalise une bijection de A_n sur X. Comme $S_n = A_n \cup X$ et $A_n \cap X = \emptyset$, on conclut que $|X| = |A_n| = \frac{1}{2}\,|S_n| = \frac{1}{2}\,n!$. □

4.4.3 EXEMPLE. Pour $n = 2$, on a $A_2 = \{e\}$. Pour $n = 3$, on a $A_3 = \{e, \gamma, \gamma^2\}$ avec $\gamma = \left(\begin{smallmatrix}1&2&3\\2&3&1\end{smallmatrix}\right)$.

4.4.4 EXEMPLE. Pour $n = 4$, le groupe alterné A_4 est d'ordre 12. Donnons quelques précisions.

Le groupe A_4 contient les trois produits de deux transpositions disjointes:

$$a = [1,2][3,4] = \left(\begin{smallmatrix}1&2&3&4\\2&1&4&3\end{smallmatrix}\right), \quad b = [1,3][2,4] = \left(\begin{smallmatrix}1&2&3&4\\3&4&1&2\end{smallmatrix}\right), \quad c = [1,4][2,3] = \left(\begin{smallmatrix}1&2&3&4\\4&3&2&1\end{smallmatrix}\right).$$

Il contient aussi les huit permutations qui permutent circulairement trois éléments i, j, k en fixant le quatrième, et qui sont donc de la forme $[i,k][i,j]$. (De tels éléments sont appelés des 3-cycles).

$$x_1 = \left(\begin{smallmatrix}1&2&3&4\\1&3&4&2\end{smallmatrix}\right), \quad y_1 = \left(\begin{smallmatrix}1&2&3&4\\1&4&2&3\end{smallmatrix}\right) = x_1^2, \quad x_2 = \left(\begin{smallmatrix}1&2&3&4\\3&2&4&1\end{smallmatrix}\right), \quad y_2 = \left(\begin{smallmatrix}1&2&3&4\\4&2&1&3\end{smallmatrix}\right) = x_2^2,$$
$$x_3 = \left(\begin{smallmatrix}1&2&3&4\\2&4&3&1\end{smallmatrix}\right), \quad y_3 = \left(\begin{smallmatrix}1&2&3&4\\4&1&3&2\end{smallmatrix}\right) = x_3^2, \quad x_4 = \left(\begin{smallmatrix}1&2&3&4\\2&3&1&4\end{smallmatrix}\right), \quad y_4 = \left(\begin{smallmatrix}1&2&3&4\\3&1&2&4\end{smallmatrix}\right) = x_4^2.$$

	e	a	b	c	x_1	y_1	x_2	y_2	x_3	y_3	x_4	y_4
e	e	a	b	c	x_1	y_1	x_2	y_2	x_3	y_3	x_4	y_4
a	a	e	c	b	x_3	x_4	y_3	y_4	x_1	x_2	y_1	y_2
b	b	c	e	a	y_4	x_2	y_1	x_3	y_2	x_4	y_3	x_1
c	c	b	a	e	y_2	y_3	x_4	x_1	y_4	y_1	x_2	x_3
x_1	x_1	y_4	y_2	x_3	y_1	e	c	x_4	x_2	a	b	y_3
y_1	y_1	y_3	x_4	x_2	e	x_1	x_3	b	c	y_4	y_2	a
x_2	x_2	x_4	y_3	y_1	b	y_4	y_2	e	a	x_1	x_3	c
y_2	y_2	x_3	x_1	y_4	y_3	c	e	x_2	x_4	b	a	y_1
x_3	x_3	y_2	y_4	x_1	x_4	a	b	y_1	y_3	e	c	x_2
y_3	y_3	y_1	x_2	x_4	c	y_2	y_4	a	e	x_3	x_1	b
x_4	x_4	x_2	y_1	y_3	a	x_3	x_1	c	b	y_2	y_4	e
y_4	y_4	x_1	x_3	y_2	x_2	b	a	y_3	y_1	c	e	x_4

Les trois éléments a, b, c sont d'ordre 2, et le sous-groupe $V = \{e, a, b, c\}$ de A_4 est le groupe de Klein.

Les huit 3-cycles x_i, y_i pour $1 \leq i \leq 4$ sont d'ordre 3. On obtient donc quatre sous-groupes cycliques $G_i = \{e, x_i, y_i\}$, pour $1 \leq i \leq 4$.

On observe au passage que, bien que 4 et 6 soient des diviseurs de $|A_4| = 12$, le groupe A_4 ne contient pas d'élément d'ordre 4 ni 6.

4.5 Support et orbites.

4.5.1 DÉFINITION. Pour toute $\sigma \in S_n$, on appelle *support* de σ l'ensemble des éléments de $\{1, 2, \ldots, n\}$ qui ne sont pas fixés par σ:

$$\text{Supp}\ \sigma = \{i \in \{1, 2, \ldots, n\}\,;\, \sigma(i) \neq i\}.$$

En particulier, $\text{Supp}\ \sigma = \emptyset$ si et seulement si $\sigma = e$.

4.5.2 LEMME. *Pour toute $\sigma \in S_n$ non triviale, la restriction de σ à $\text{Supp}\ \sigma$ est une permutation de $\text{Supp}\ \sigma$.*

Preuve. Soit $i \in \text{Supp}\ \sigma$; notons $j = \sigma(i)$. Si on avait $j \notin \text{Supp}\ \sigma$, on aurait $\sigma(j) = j$, donc $\sigma(j) = \sigma(i)$, donc $i = j$, c'est-à-dire $i = \sigma(i)$, ce qui contredirait $i \in \text{Supp}\ \sigma$. C'est donc que $\text{Supp}\ \sigma$ est stable par σ. La restriction σ' de σ à $\text{Supp}\ \sigma$ est une application de $\text{Supp}\ \sigma$ dans lui-même, injective car σ l'est, et donc bijective. □

4.5.3 PROPOSITION. *Deux permutations de S_n dont les supports sont disjoints commutent.*

Preuve. On peut supposer $n \geq 2$. Soient $\sigma, \eta \in S_n$ tels que $\text{Supp}\ \sigma \cap \text{Supp}\ \eta = \emptyset$. Soit $i \in \mathbb{N}_n$ quelconque. Si $i \notin \text{Supp}\ \sigma \cup \text{Supp}\ \eta$; alors $\sigma(i) = i = \eta(i)$; donc $\sigma\eta(i) = \eta\sigma(i)$. Supposons maintenant $i \in \text{Supp}\ \sigma$. D'une part, $i \notin \text{Supp}\ \eta$, donc $\eta(i) = i$, donc $\sigma\eta(i) = \sigma(i)$. D'autre part, $i \in \text{Supp}\ \sigma$ implique $\sigma(i) \in \text{Supp}\ \sigma$ d'après le lemme précédent, donc $\sigma(i) \notin \text{Supp}\ \eta$, donc $\eta\sigma(i) = \sigma(i)$. On conclut que $\sigma\eta(i) = \eta\sigma(i)$. Le dernier cas est celui où $i \in \text{Supp}\ \eta$, que l'on traite de façon analogue en échangeant les rôles de σ et η. □

4.5.4 DÉFINITION. Pour toute $\sigma \in S_n$ et tout $i \in \{1, 2, \ldots, n\}$, on appelle *$\sigma$-orbite* de i l'ensemble des images de i par les différents éléments du groupe cyclique $\langle\sigma\rangle$; on note

$$\Omega_\sigma(i) = \{\sigma^k(i)\,;\, k \in \mathbb{Z}\}, \quad \text{pour tout } 1 \leq i \leq n.$$

Il est clair que les différentes σ-orbites forment une partition de $\{1, 2, \ldots, n\}$, et que $\Omega_\sigma(i) = \{i\}$ si et seulement $i \notin \text{Supp } \sigma$, (on dit alors que c'est une σ-orbite ponctuelle).

Exemple: soit $\sigma = \left(\begin{smallmatrix} 1&2&3&4&5&6&7&8 \\ 5&8&3&6&4&1&7&2 \end{smallmatrix}\right) \in S_8$; on a: $\Omega_\sigma(3) = \{3\}$, $\Omega_\sigma(7) = \{7\}$, $\Omega_\sigma(2) = \Omega_\sigma(8) = \{2, 8\}$, $\Omega_\sigma(1) = \{1, 5, 4, 6\} = \Omega_\sigma(5) = \Omega_\sigma(4) = \Omega_\sigma(6)$. Donc Supp $\sigma = \{1, 2, 4, 5, 6, 8\}$.

4.6 Décomposition d'une permutation en produit de cycles disjoints.

4.6.1 DÉFINITION. Une permutation $\sigma \in S_n$ est appelée un cycle lorsqu'il existe une σ-orbite et une seule qui n'est pas ponctuelle.

Exemple: soit $\sigma = \left(\begin{smallmatrix} 1&2&3&4&5&6 \\ 5&2&3&6&4&1 \end{smallmatrix}\right) \in S_6$; on a: $\Omega_2 = \{2\}$, $\Omega_3 = \{3\}$, $\Omega_1 = \{1, 5, 4, 6\} = \Omega_5 = \Omega_4 = \Omega_6$. Donc σ est un cycle.

4.6.2 PROPOSITION ET DÉFINITION. *Soit $\sigma \in S_n$ un cycle. On note p l'ordre de σ dans S_n.*

(i) *L'unique σ-orbite non ponctuelle est égale au support de σ.*

(ii) *Le cardinal du support de σ est égal à l'ordre p de σ.*

(iii) *Il existe $j_1, j_2, \ldots, j_p$ distincts dans $\{1, 2, \ldots, n\}$ tels que:*

$$\sigma(j_1) = j_2,\ \sigma(j_2) = j_3,\ \ldots,\ \sigma(j_p) = j_1 \quad \text{et} \quad \sigma(i) = i \ \text{ si } \ i \notin \{j_1, \ldots, j_p\}.$$

Preuve. Notons Ω l'unique σ-orbite non ponctuelle. Soit j_1 un représentant quelconque de Ω. Donc: $\Omega = \Omega_\sigma(j_1) = \{i \in \{1, 2, \ldots, n\}\,;\, \Omega_\sigma(i) \neq \{i\}\}$ c'est-à-dire $\Omega = \text{Supp } \sigma$, par définition même du support. Soit $q = |\Omega| = |\text{ Supp } \sigma|$. Donc:

$$\Omega = \text{Supp } \sigma = \{j_1, \sigma(j_1), \sigma^2(j_1), \ldots, \sigma^{q-1}(j_1)\},$$

les éléments étant deux à deux distincts. On a alors $\sigma^q(j_1) = j_1$, et ceci étant vrai pour tout représentant j_1 dans $\Omega = \text{Supp } \sigma$, on a $\sigma^q(i) = i$ pour tout $i \in \text{Supp } \sigma$. Mais l'égalité $\sigma^q(i) = i$ est claire si $i \notin \text{Supp } \sigma$ puisqu'alors $\sigma(i) = i$. Ainsi $\sigma^q = e$ dans S_n. Comme $\sigma^k \neq e$ pour $1 \leq k < q$, (puisque j_1 et $\sigma^k(j_1)$ sont alors deux éléments distincts de Ω), on conclut que q est exactement l'ordre de σ dans S_n. □

On dit que σ est un *p-cycle*, ou *cycle d'ordre p*. On note: $\sigma = [j_1, j_2, \ldots, j_p]$.

Remarque. On a aussi: $\sigma = [j_k, j_{k+1}, \ldots, j_p, j_1, \ldots, j_{k-1}]$ pour tout $1 < k \leq p$.

4.6.3 EXEMPLES ET PREMIÈRES PROPRIÉTÉS.

1. Les 2-cycles sont les transpositions $[i, j]$. Par convention, le seul 1-cycle est e.

2. Le n-cycle $[1, 2, \ldots, n] = \left(\begin{smallmatrix} 1&2&3&\ldots&n-1&n \\ 2&3&4&\ldots&n&1 \end{smallmatrix}\right)$ s'appelle la permutation circulaire de S_n. Il existe des n-cycles qui ne sont pas la permutation circulaire, par exemple $[1, 3, 4, 2] \in S_4$.

3. L'inverse d'un p-cycle est un p-cycle: $[j_1, j_2, \ldots, j_p]^{-1} = [j_p, j_{p-1}, \ldots, j_1]$.

4. Attention: si $\gamma \in S_n$ est un r-cycle, et si $2 \leq k \leq r-2$, alors γ^k n'est pas nécessairement un cycle. Par exemple, si γ est la permutation circulaire $[1, 2, 3, 4] \in S_4$, alors $\gamma^2 = [1, 3][2, 4]$ n'est pas un cycle.

5. Tout conjugué d'un p-cycle est un p-cycle. Plus précisément:

$$\text{si } \gamma = [j_1, j_2, \ldots, j_p] \ \text{ et } \ \sigma \in S_n, \ \text{ alors } \ \sigma\gamma\sigma^{-1} = [\sigma(j_1), \sigma(j_2), \ldots, \sigma(j_p)].$$

Preuve. Soit $i \in \{1, 2, \ldots, n\}$. Si $\sigma^{-1}(i) \in \text{Supp } \gamma$, il existe $1 \leq k \leq p$ tel que $i = \sigma(j_k)$. On a $\sigma\gamma\sigma^{-1}(i) = \sigma\gamma(j_k) = \sigma(j_{k+1})$ si $1 \leq k < p$, et $\sigma\gamma\sigma^{-1}(i) = \sigma\gamma(j_p) = \sigma(j_1)$ si $k = p$. Si maintenant $\sigma^{-1}(i) \notin \text{Supp } \gamma$, alors $\gamma\sigma^{-1}(i) = \sigma^{-1}(i)$ et donc $\sigma\gamma\sigma^{-1}(i) = i$. Ceci prouve par définition même que $\sigma\gamma\sigma^{-1}$ est le p-cycle $[\sigma(j_1), \sigma(j_2), \ldots, \sigma(j_p)]$. □

6. Pour $n \geq 3$, le groupe alterné A_n est engendré par les 3-cycles de S_n.

Preuve. Un produit de deux transpositions est nécessairement de l'un des deux types suivants (où i, j, k, l sont distincts deux à deux): ou bien $[i,j][i,k] = [i,k,j]$, ou bien $[i,j][k,l] = [i,l,k][i,j,k]$. Ce qui prouve le résultat voulu puisque A_n est l'ensemble des produits d'un nombre pair de transpositions. □

7. Si γ est un p-cycle, alors $\varepsilon(\gamma) = (-1)^{p-1}$.

Preuve. Si $\gamma = [j_1, j_2, \ldots, j_p]$, alors $\gamma = [j_1, j_p][j_1, j_{p-1}] \cdots [j_1, j_2]$. □

On a vu en 4.2.3 que la décomposition d'une permutation en produit de transpositions n'est pas unique. En revanche, comme on va le voir maintenant, toute permutation se décompose en produits de cycles disjoints (et donc commutant deux à deux), et ceci de façon unique.

4.6.4 Théorème (Décomposition en produit de cycles disjoints).

(i) *Toute $\sigma \in S_n$ non triviale se décompose en un produit de cycles non triviaux à supports disjoints.*

(ii) *Les cycles dans une telle décomposition commutent deux à deux.*

(iii) *Cette décomposition est unique à l'ordre près des facteurs.*

Preuve. Soit $\sigma \in S_n$ non triviale. Il existe donc au moins une σ-orbite non ponctuelle. Désignons par $\Omega_1, \ldots, \Omega_q$ les σ-orbites non ponctuelles deux à deux distinctes (et donc deux à deux disjointes). Pour tout $1 \leq k \leq q$, définissons $\gamma_k : \{1, 2, \ldots, n\} \longrightarrow \{1, 2, \ldots, n\}$ par: $\gamma_k(i) = \sigma(i)$ si $i \in \Omega_k$ et $\gamma_k(i) = i$ sinon. Alors γ_k est un cycle dans S_n, (car si l'on note $r_k = |\Omega_k|$, on a $\Omega_k = \{j, \sigma(j), \sigma^2(j), \ldots, \sigma^{r_k-1}(j)\}$ quel que soit $j \in \Omega_k$), de support égal à Ω_k. Il en résulte que les supports des γ_i sont deux à deux disjoints, donc (d'après la proposition 4.5.3), que les γ_i commutent deux à deux dans S_n. Posons $\sigma' = \gamma_1 \gamma_2 \ldots \gamma_q$; on va montrer que $\sigma' = \sigma$.

En effet, soit $j \in \{1, 2, \ldots, n\}$; distinguons deux cas.

• Si $j \in \Omega_1 \cup \ldots \cup \Omega_q$, alors j appartient à une seule de ces orbites: il existe $1 \leq k \leq q$ tel que $j \in \Omega_k$ et $j \notin \Omega_i$ si $i \neq k$. Puisque les γ_i commutent deux à deux, on peut écrire $\sigma' = \gamma_k \gamma_1 \ldots \gamma_{k-1} \gamma_{k+1} \ldots \gamma_q$. Pour tout indice $i \neq k$, on $\gamma_i(j) = j$ car $j \notin \Omega_i = \text{Supp } \gamma_i$; donc $\gamma_1 \ldots \gamma_{k-1} \gamma_{k+1} \ldots \gamma_q(j) = j$, d'où $\sigma'(j) = \gamma_k(j)$. Or $\gamma_k(j) = \sigma(j)$ puisque $j \in \Omega_k$. On conclut finalement que $\sigma'(j) = \sigma(j)$.

• Si $j \notin \Omega_1 \cup \ldots \cup \Omega_q$, alors, pour tout $1 \leq k \leq q$, on a $j \notin \text{Supp } \gamma_k$ donc $\gamma_k(j) = j$, de sorte que $\sigma'(j) = j$. Mais par ailleurs, $j \notin \Omega_1 \cup \ldots \cup \Omega_q$ signifie que la σ-orbite de j est ponctuelle, c'est-à-dire que $\sigma(j) = j$. Dans ce cas aussi, on a vérifié que $\sigma'(j) = \sigma(j)$.

On a ainsi prouvé les points (i) et (ii) du théorème. Pour prouver (iii), supposons que l'on a une décomposition $\sigma = \gamma'_1 \gamma'_2 \ldots \gamma'_p$ en produit de cycles non triviaux à supports deux à deux disjoints (donc commutant deux à deux). Pour tout $1 \leq i \leq p$, notons $\Omega'_i = \text{Supp } \gamma'_i$. Chaque Ω'_i est une σ-orbite non ponctuelle, plus précisément:

$$\text{pour tout } 1 \leq k \leq p \text{ et pour tout } j \in \Omega'_k, \text{ on a } \Omega'_k = \Omega_\sigma(j). \qquad (*)$$

En effet. Fixons $1 \leq k \leq p$ et $j \in \Omega'_k$. Il en résulte que $j \notin \Omega'_i$ si $1 \leq i \neq k \leq p$ (puisque les supports des γ'_i sont deux à deux disjoints). En d'autres termes, $\gamma'_i(j) = j$ pour tout $1 \leq i \neq k \leq p$. Donc en écrivant $\sigma = \gamma'_k \gamma'_1 \ldots \gamma'_{k-1} \gamma'_{k+1} \ldots \gamma'_p$, suivant la méthode déjà employée ci-dessus, on calcule $\sigma(j) = \gamma'_k(j)$. Comme $\gamma'_k(j)$ appartient à Ω'_k et n'appartient pas à Ω'_i pour $1 \leq i \neq k \leq p$, on réitère pour obtenir $\sigma^2(j) = (\gamma'_k)^2(j)$. Et finalement $\sigma^m(j) = (\gamma'_k)^m(j)$ pour tout entier $m \geq 1$. On conclut que: $\Omega'_k = \Omega_\sigma(j)$.

Réciproquement, on obtient ainsi toutes les σ-orbites non ponctuelles, plus précisément:

$$\text{pour tout } j \in \{1, 2, \ldots, n\} \text{ telle que } \Omega_\sigma(j) \neq \{j\}, \text{ il existe } 1 \leq k \leq p \text{ tel que } \Omega'_k = \Omega_\sigma(j). \qquad (**)$$

En effet. Par contraposée, si l'on suppose que quel que soit $1 \leq k \leq p$, on a $j \notin \Omega'_k$, alors $\gamma'_k(j) = j$ pour tout $1 \leq k \leq p$, de sorte que $\sigma(j) = j$, c'est-à-dire $\Omega_\sigma(j) = \{j\}$.

Il résulte de $(*)$ et $(**)$ que la décomposition $\sigma = \gamma'_1 \gamma'_2 \ldots \gamma'_p$ est, à l'ordre près, celle que l'on a construite dans la preuve du point (i), c'est-à-dire que $p = q$ et $\{\gamma_1, \ldots, \gamma_p\} = \{\gamma'_1, \ldots, \gamma'_p\}$. □

4.6.5 Exemples d'applications (en exercices)

(a) *Une définition équivalente de la signature.* Montrer que, pour toute permutation $\sigma \in S_n$, la signature de σ vérifie $\varepsilon(\sigma) = (-1)^{n-t}$, où t désigne le nombre de σ-orbites distinctes dans S_n.

Indication: utiliser 4.6.4 et le point 7 de 4.6.3

(b) *Ordre d'un élément quelconque de S_n.* Montrer que l'ordre dans S_n d'un élément quelconque σ non trivial est égal au p.p.c.m. des longueurs des cycles disjoints de la décomposition canonique de σ.

Indication: utiliser 4.6.4 et le point (ii) de 4.6.2

Leçon 5

Conjugaison et sous-groupes normaux

5.1 Conjugaison.

5.1.1 Définition. Soit G un groupe. Soient g et g' deux éléments de G. On dit que g' est *conjugué* avec g lorsqu'il existe un élément $x \in G$ tel que $g' = xgx^{-1}$.

(a) Si $g' = xgx^{-1}$, alors $g = x^{-1}g'x = yg'y^{-1}$ pour $y = x^{-1}$, de sorte que g est conjugué avec g'. On dira donc simplement que g et g' sont conjugués.

(b) Tout élément $g \in G$ est conjugué à lui-même (car $g = ege^{-1}$).

(c) La notion de conjugaison n'a bien sûr d'intérêt que pour un groupe G non abélien, car si G est abélien, le seul élément conjugué à un élément quelconque g de G est g lui-même.

(d) Avec la terminologie introduite un peu plus loin en 5.2.3, dire que g' est conjugué avec g se traduit par l'existence d'un automorphisme intérieur σ_x tel que $g' = \sigma_x(g)$, ou encore $g = \sigma_x^{-1}(g')$.

5.1.2 Proposition. *Soit G un groupe. La relation "être conjugué" dans G est une relation d'équivalence.*

Preuve. La réflexivité et la symétrie découlent des remarques (a) et (b) ci-dessus. Pour la transitivité, supposons que g' est conjugué avec g, et que g'' est conjugué avec g'. Il existe donc x et y dans G tels que $g' = xgx^{-1}$ et $g'' = yg'y^{-1}$. On a alors $g'' = yxgx^{-1}y^{-1} = (yx)g(yx)^{-1}$, ce qui prouve que g'' est conjugué avec g. □

5.1.3 Remarques et notations. Pour tout $g \in G$, la classe d'équivalence de g pour la relation de conjugaison est appelé la classe de conjugaison de g. On la note $\mathrm{cl}(g)$. Rappelons que:

$$\mathrm{cl}(g) = \{g' \in G\,;\, g' \text{ conjugué avec } g\} = \{xgx^{-1}\,;\, x \in G\}.$$

Comme on l'a dit en 5.1.1.(c), cette notion n'a d'intérêt que si G n'est pas abélien car, si G est abélien, on a $\mathrm{cl}(g) = \{g\}$ pour tout $g \in G$. Rappelons aussi que, comme pour toute relation d'équivalence, les classes de conjugaison forment une partition de G.

5.2 Automorphismes intérieurs et centre.

5.2.1 Proposition et définition. *Soit G un groupe.*

(i) *L'ensemble des éléments de G qui commutent avec tous les éléments de G est un sous-groupe de G, appelé le centre du groupe G, et noté $Z(G)$:*

$$Z(G) = \{x \in G\,;\, gx = xg \text{ pour tout } g \in G\}.$$

(ii) *Le sous-groupe $Z(G)$ est abélien.*

(iii) *G est abélien si et seulement si $Z(G) = G$.*

Preuve. Pour tout $g \in G$, on a $eg = ge = g$, donc $e \in Z(G)$, et $Z(G)$ n'est donc pas vide. Soient $x, y \in Z(G)$; pour tout $g \in G$, on a: $(xy)g = x(yg) = x(gy) = (xg)y = (gx)y = g(xy)$, et donc $xy \in Z(G)$. Soit $x \in Z(G)$; pour tout $g \in G$, on multiplie les deux membres de l'égalité $xg = gx$ par x^{-1} à gauche et à droite, et l'on obtient $gx^{-1} = x^{-1}g$, ce qui prouve que $x^{-1} \in Z(G)$. Ceci prouve (i). Les points (ii) et (iii) sont alors évidents. □

5.2.2 Exercice. Montrer que, pour tout $x \in G$ l'ensemble $C(x) = \{g \in G\,;\, gx = xg\}$ des éléments de G qui commutent avec x est un sous-groupe de G. On l'appelle le centralisateur de x. Montrer que $Z(G) = \bigcap_{x \in G} C(x)$.

5.2.3 Proposition et définition. *Soit G un groupe.*

(i) *Pour tout $x \in G$, l'application $\sigma_x : G \to G$ définie par:*
$$\sigma_x(g) = xgx^{-1} \quad \text{pour tout } g \in G$$
est un automorphisme du groupe G; on l'appelle l'automorphisme intérieur déterminé par x.

(ii) *L'ensemble $\operatorname{Int} G = \{\sigma_x\,;\, x \in G\}$ de tous les automorphismes intérieurs de G est un sous-groupe du groupe $\operatorname{Aut} G$ de tous les automorphismes de G.*

(iii) *L'application $\sigma : G \to \operatorname{Aut} G$ qui, à tout élément $x \in G$, associe l'automorphisme intérieur σ_x, est un morphisme de groupes, d'image $\operatorname{Int} G$ et de noyau le centre $Z(G)$.*

Preuve. (i) Fixons $x \in G$. Pour tout $g \in G$, on a:
$$\sigma_{x^{-1}}(\sigma_x(g)) = x^{-1}(xgx^{-1})x = g = x(x^{-1}gx)x^{-1} = \sigma_x(\sigma_{x^{-1}}(g)).$$
Ceci montre que $\sigma_x \circ \sigma_{x^{-1}} = \sigma_{x^{-1}} \circ \sigma_x = \mathrm{id}_G$, ce qui prouve que σ_x est une bijection de G sur G, dont la bijection réciproque est $\sigma_{x^{-1}}$. En d'autres termes, $\sigma_x^{-1} = \sigma_{x^{-1}}$. Par ailleurs, quels que soient $g, h \in G$, on a:
$$\sigma_x(gh) = x(gh)x^{-1} = xg(x^{-1}x)hx^{-1} = (xgx^{-1})(xhx^{-1}) = \sigma_x(g)\sigma_x(h),$$
ce qui montre que σ_x est un morphisme de groupes. On conclut que $\sigma_x \in \operatorname{Aut} G$.

(ii) L'ensemble $\operatorname{Int} G$ n'est pas vide: il contient en particulier $\mathrm{id}_G = \sigma_e$. Soient $x, y \in G$. Pour tout $g \in G$, on a: $\sigma_x(\sigma_y(g)) = x(ygy^{-1})x^{-1} = (xy)g(y^{-1}x^{-1}) = (xy)g(xy)^{-1} = \sigma_{xy}(g)$. Donc $\sigma_x \circ \sigma_y = \sigma_{xy}$. Ceci prouve que $\operatorname{Int} G$ est stable pour la loi $\circ$. Par ailleurs, on a déjà observé dans la preuve du point (i) que, pour tout $x \in G$, $\sigma_x^{-1} = \sigma_{x^{-1}} \in \operatorname{Int} G$, de sorte que $\operatorname{Int} G$ est aussi stable par passage à l'inverse. Ce qui achève la preuve.

(iii) On vient de voir que $\sigma_{xy} = \sigma_x \circ \sigma_y$ pour tous $x, y \in G$, ce qui montre que σ est un morphisme de groupes. Le fait que $\operatorname{Im}\sigma = \operatorname{Int} G$ découle de la définition même de $\operatorname{Int} G$. Soit maintenant $x \in \operatorname{Ker}\sigma$. Cela équivaut à $\sigma_x = \mathrm{id}_G$, c'est-à-dire à $xgx^{-1} = g$ pour tout $g \in G$, ou encore (en multipliant à droite par x) à $xg = gx$ pour tout $g \in G$. On conclut que $\operatorname{Ker}\sigma = Z(G)$. □

Le point (iii) de cette proposition sera utilisé plus loin de façon cruciale au corollaire 6.4.3. .

5.3 Notion de sous-groupe normal.

5.3.1 Notation. Soit G un groupe. Pour tout sous-groupe H de G, et pour tout $x \in G$, on note xHx^{-1} l'image de H par l'automorphisme intérieur σ_x :
$$xHx^{-1} = \{xhx^{-1}\,;\, h \in H\} = \sigma_x(H) \quad \text{pour tout } x \in G.$$
D'après la prop. 2.1.4, c'est un sous-groupe de G. Cette notion n'a d'intérêt que si G n'est pas abélien car, si G est abélien, on a $xhx^{-1} = h$ pour tous $x \in G, h \in H$.

5.3.2 Remarques. Soient G un groupe et H un sous-groupe de G. Considérons les quatre assertions suivantes:

(1) pour tout $h \in H$, pour tout $x \in G$, on a $xhx^{-1} = h$,
(2) pour tout $h \in H$, pour tout $x \in G$, on a $xhx^{-1} \in H$,
(3) pour tout $h \in H$, pour tout $x \in G$, il existe $h' \in H$ tel que $xhx^{-1} = h'$,
(4) pour tout $x \in G$, on a $xHx^{-1} = H$.

(a) Il est clair que (1) implique (2), mais la réciproque est fausse. On peut avoir (2) sans avoir (1).

En effet, considérons par exemple, dans le groupe symétrique $S_3 = \{e, \gamma, \gamma^2, \tau_1, \tau_2, \tau_3\}$, le sous-groupe $H = \{e, \gamma, \gamma^2\}$. On a $\tau_1\gamma\tau_1^{-1} = \tau_1\gamma\tau_1 = \gamma^2$. Donc $\tau_1\gamma\tau_1^{-1} \in H$ mais $\tau_1\gamma\tau_1^{-1} \neq \gamma$. Ce qui prouve que l'on n'a pas (1). Mais on a $\tau_1\gamma\tau_1^{-1} = \tau_2\gamma\tau_2^{-1} = \tau_3\gamma\tau_3^{-1} = \gamma^2$ et $\tau_1\gamma^2\tau_1^{-1} = \tau_2\gamma^2\tau_2^{-1} = \tau_3\gamma^2\tau_3^{-1} = \gamma$, qui suffit à prouver que l'on a (2).

(b) Il est clair que, si G est abélien, alors on a (1), mais la réciproque est fausse.

En effet, il suffit de prendre G non abélien et $H = \{e\}$, ou plus généralement $H = Z(G)$ le centre de G.

(c) Il est clair que (2) est équivalent à (3).

(d) Les assertions (2) et (4) sont équivalentes.

En effet, comme (2) équivaut à $xHx^{-1} \subset H$ pour tout $x \in G$, il est clair que (4) implique (2). Réciproquement, supposons que l'on a (2). On a donc l'inclusion $xHx^{-1} \subset H$; pour l'inclusion réciproque, tout $h \in H$ s'écrit $h = x(x^{-1}hx)x^{-1}$, et comme $x^{-1}hx \in H$ d'après l'hypothèse (2), on a $h \in xHx^{-1}$, ce qui prouve que $H \subset xHx^{-1}$.

5.3.3 Définition. Soient G un groupe et H un sous-groupe de G. On dit que H est *normal dans* G, ou encore *distingué dans* G, lorsque $xHx^{-1} = H$ pour tout $x \in G$. On note alors $H \triangleleft G$.

$$(H \triangleleft G) \Leftrightarrow (xHx^{-1} = H \text{ pour tout } x \in G) \Leftrightarrow (xhx^{-1} \in H \text{ pour tous } h \in H, x \in G)$$

• Par exemple, les calculs effectués à la remarque 5.3.2(a) ci-dessus montrent que le sous-groupe $H = \{e, \gamma, \gamma^2\}$ de S_3 est normal dans S_3.

• Ré-insistons sur le fait que la notion de sous-groupe normal n'a d'intérêt que pour les groupes non-abéliens puisqu'il résulte des remarques 5.3.2(a) et 5.3.2(b) que:

tout sous-groupe d'un groupe abélien G est normal dans G.

• Ré-insistons sur le fait vu en 5.3.2.(a) que, si $H \triangleleft G$, on a $xHx^{-1} = H$ pour tout $x \in G$, mais pas nécessairement $xhx^{-1} = h$ pour tous $x \in G, h \in H$.

5.4 Premiers exemples.

5.4.1 Proposition. *Pour tout groupe G, les sous-groupes $\{e\}$ et G sont normaux dans G.*

Preuve. Evident □

5.4.2 Proposition. *Pour tout groupe G, le centre $Z(G)$ est un sous-groupe normal dans G.*

Preuve. Quels que soient $h \in Z(G)$ et $x \in G$, on a $xhx^{-1} = h$ par définition du fait que h est dans le centre de G, donc $xhx^{-1} \in Z(G)$. □

5.4.3 Proposition. *Pour tout morphisme f d'un groupe G dans un groupe G', le noyau* $\operatorname{Ker} f$ *est un sous-groupe normal dans G.*

Preuve. Soient $h \in \operatorname{Ker} f$ et $x \in G$ quelconques. Il s'agit de vérifier que $xhx^{-1} \in \operatorname{Ker} f$. Pour cela, calculons $f(xhx^{-1}) = f(x)f(h)f(x)^{-1}$. Mais $f(h) = e'$, donc $f(xhx^{-1}) = f(x)f(x)^{-1} = e'$, ce qui prouve le résultat voulu. □

Dans la pratique, reconnaître un sous-groupe donné comme le noyau d'un morphisme est un moyen immédiat et fréquent de montrer qu'il est normal. Par exemple:

(i) $A_n \triangleleft S_n$,

En effet, le groupe alterné A_n n'est autre que le noyau du morphisme signature $\varepsilon : S_n \to \{1, -1\}$.

(ii) $\mathrm{SL}_n(\mathbb{R}) \triangleleft \mathrm{GL}_n(\mathbb{R})$.

En effet, le groupe spécial linéaire $\mathrm{SL}_n(\mathbb{R})$ n'est autre que le noyau du morphisme déterminant $\mathrm{GL}_n(\mathbb{R}) \to \mathbb{R}^*$.

5.4.4 Proposition. *Si G est un groupe fini d'ordre pair $2n$ et si H est un sous-groupe d'ordre n de G, alors H est normal dans G.*

Preuve. Notons $H = \{h_1, h_2, h_3, \ldots, h_n\}$, avec $h_1 = e$. Fixons un élément $y \in G$ tel que $y \notin H$ (il en existe car $|G| = 2n$ alors que $|H| = n$). Notons yH l'ensemble des produits par y à gauche des éléments de H. Comme $yh_i = yh_j$ si et seulement si $h_i = h_j$ (en multipliant à gauche par y^{-1}), on déduit que $yH = \{yh_1, yh_2, \ldots, yh_n\}$ est formé de n éléments distincts. S'il existait un élément commun à H et yH, il s'écrirait $yh_i = h_j$, ce qui est impossible car on aurait alors $y = h_j h_i^{-1} \in H$, ce qui est contraire au choix de y. On conclut donc que $G = H \cup yH$ et $H \cap yH = \emptyset$.

Dès lors, soient h_i un élément quelconque de H et x un élément quelconque de G. Si $x \in H$, alors xh_ix^{-1} est le produit de trois éléments de H, donc appartient à H. Si $x \notin H$, alors $x \in yH$, donc il existe $h_j \in H$ tel que $x = yh_j$. Donc $xh_ix^{-1} = yh_jh_ih_j^{-1}y^{-1} = yh_\ell y^{-1}$ où l'on a posé $h_\ell = h_jh_ih_j^{-1} \in H$. Si $yh_\ell y^{-1}$ n'appartenait pas à H, il appartiendrait à yH, donc on aurait $yh_\ell y^{-1} = yh_k$ pour un certain $h_k \in H$, d'où $h_\ell y^{-1} = h_k$, donc $y = h_k^{-1}h_\ell$ appartiendrait à H. Comme ce n'est pas le cas, c'est que $yh_\ell y^{-1} \in H$, c'est-à-dire $xh_ix^{-1} \in H$. □

- Ce résultat donne une nouvelle preuve du fait que $A_n \lhd S_n$, d'après la proposition 4.4.2.

- On verra plus loin en 6.1.5 que cette même preuve permet de montrer un résultat analogue dans un cadre un peu plus général pour G non nécessairement fini.

5.4.5 Proposition. *Si H et K sont deux sous-groupes d'un groupe G tels que $H \lhd G$ et $K \lhd G$, alors $H \cap K \lhd G$.*

Preuve. Immédiate; laissée au lecteur en exercice. □

5.4.6 Remarque. Attention: si H et K sont deux sous-groupes d'un groupe G tels que $K \subset H$,

$$(\ K \lhd H \ \text{ et } \ H \lhd G\ ) \ \text{ n'implique pas } \ (\ K \lhd G\ ).$$

Contre-exemple. Considérons le groupe alterné A_4, en reprenant pour ses éléments toutes les notations du paragraphe 4.4.4 du chapitre 4. Rappelons d'abord que:

pour toute permutation $\sigma \in S_4$ et toute transposition $[i,j]$, on a: $\sigma[i,j]\sigma^{-1} = [\sigma(i),\sigma(j)]$.

Considérons dans A_4 le sous-groupe $V = \{e, a, b, c\}$. Pour tout $\sigma \in A_4$, on a:

$$\sigma a\sigma^{-1} = \sigma[1,2][3,4]\sigma^{-1} = \sigma[1,2]\sigma^{-1}\sigma[3,4]\sigma^{-1} = [\sigma(1),\sigma(2)][\sigma(3),\sigma(4)] \in V.$$

On montre de même que $\sigma b\sigma^{-1} \in V$ et $\sigma c\sigma^{-1} \in V$ pour tout $\sigma \in A_4$. On conclut que $V \lhd A_4$.

Considérons dans A_4 le sous-groupe $K = \{e, a\}$ de V, donc de A_4. Il n'est pas normal dans A_4, car par exemple, $x_1ax_1^{-1} = x_1ay_1 = b \notin K$. Et pourtant K est normal dans V puisque V est abélien.

5.4.7 Proposition. *Pour tout groupe G, le groupe $\operatorname{Int} G$ des automorphismes intérieurs de G est un sous-groupe normal du groupe $\operatorname{Aut} G$ de tous les automorphismes de G.*

Preuve. Soit $\sigma_x \in \operatorname{Int} G$ quelconque, avec $x \in G$. Soit $\gamma \in \operatorname{Aut} G$. On calcule pour tout $g \in G$: $\gamma \circ \sigma_x \circ \gamma^{-1}(g) = \gamma\big(\sigma_x(\gamma^{-1}(g))\big) = \gamma\big(x\gamma^{-1}(g)x^{-1}\big) = \gamma(x)\,\gamma(\gamma^{-1}(g))\,\gamma(x^{-1}) = \gamma(x)g\gamma(x)^{-1} = ygy^{-1}$ en posant $y = \gamma(x)$. On déduit que $\gamma \circ \sigma_x \circ \gamma^{-1} = \sigma_y \in \operatorname{Int} G$. On a ainsi vérifié que, pour tous $\sigma_x \in \operatorname{Int} G$ et $\gamma \in \operatorname{Aut} G$, le conjugué $\gamma \circ \sigma_x \circ \gamma^{-1}$ appartient à $\operatorname{Int} G$. Ceci prouve que $\operatorname{Int} G \lhd \operatorname{Aut} G$. □

5.5 Normalisateur.

5.5.1 Proposition et définition. *Soient G un groupe et H un sous-groupe de G. L'ensemble $N_G(H) = \{x \in G\,;\, xHx^{-1} = H\}$ est un sous-groupe de G, appelé le normalisateur dans G de H.*

Preuve. Vérification immédiate, laissée au lecteur. □

5.5.2 Proposition. *Soient G un groupe et H un sous-groupe de G. On a:*

(i) *$H \lhd N_G(H)$,*

(ii) *$N_G(H)$ est le plus grand sous-groupe de G dans lequel H est normal;*

(iii) *en particulier $H \lhd G$ si et seulement si $N_G(H) = G$.*

Preuve. Vérification immédiate, laissée au lecteur. □

Leçon 6

Groupes quotients

6.1 Classe modulo un sous-groupe, indice.

6.1.1 Définition. Soit G un groupe. Soit H un sous-groupe. Pour tout $x \in G$, on note:

$$xH = \{xh\,;\, h \in H\} \quad \text{et} \quad Hx = \{hx\,;\, h \in H\}.$$

Le sous-ensemble xH s'appelle la classe à gauche de x modulo H. Le sous-ensemble Hx s'appelle la classe à droite de x modulo H.

(a) Pour tout $x \in G$, H est en bijection avec xH via $h \mapsto xh$, et en bijection avec Hx via $h \mapsto hx$.

(b) En particulier, pour $x = e$, on a: $eH = He = H$.

(c) Pour tout $x \in G$, on a $x \in xH$ et $x \in Hx$, (car $x = xe = ex$ et $e \in H$).

6.1.2 Lemme. *Soient G un groupe et H un sous-groupe de G.*

(i) Pour tous $x, y \in G$, on a:

$$(\ xH = yH\ ) \Leftrightarrow (\ x^{-1}y \in H\ ) \quad \text{et} \quad (\ Hx = Hy\ ) \Leftrightarrow (\ xy^{-1} \in H\ ).$$

(ii) *Les classes à gauche modulo H forment une partition de G, (de même que les classes à droite).*

(iii) *L'ensemble des classes à droite modulo H est en bijection avec l'ensemble des classes à gauche modulo H, via la bijection $Hx \mapsto x^{-1}H$.*

(iv) *Pour tout sous-groupe H de G, on a:*

$$(\ H \lhd G\ ) \Leftrightarrow (\ xH = Hx \text{ pour tout } x \in G\ ).$$

Preuve. (i) Supposons que $xH = yH$. En particulier, comme $y \in yH$, on a $y \in xH$, donc il existe $h \in H$ tel que $y = xh$, d'où $x^{-1}y = h \in H$. Réciproquement, supposons que $x^{-1}y \in H$. Posons $h_0 = x^{-1}y$. Soit $z \in yH$ quelconque. Il existe $h \in H$ tel que $z = yh = xh_0h$, qui appartient à xH puisque $h_0h \in H$. D'où $yH \subseteq xH$. Soit $z' \in xH$ quelconque. Il existe $h' \in H$ tel que $z' = xh' = yh_0^{-1}h'$, qui appartient à yH puisque $h_0^{-1}h \in H$. D'où $xH \subseteq yH$, et finalement $xH = yH$. On raisonne de même pour les classes à droite.

(ii) On a vu que tout $x \in G$ vérifie $x \in xH$, donc G est égal à la réunion des classes à gauche. Il reste à montrer que deux classes distinctes sont disjointes, ou encore que deux classes non disjointes sont égales. Considérons donc xH et yH (avec $x, y \in G$) tel que $xH \cap yH$ contienne au moins un élément z. Il existe donc $h, h' \in H$ tels que $z = xh = yh'$. Dans ce cas, $x^{-1}y = h(h')^{-1} \in H$, d'où $xH = yH$ d'après le point (i). Ainsi $xH = yH$ dès lors que $xH \cap yH \neq \emptyset$. La preuve à droite est identique.

(iii) Soit φ l'application $Hx \mapsto x^{-1}H$ de l'ensemble des classes à droite dans l'ensemble des classes à gauche. Il est clair qu'elle est surjective. De plus, si $x, y \in G$ vérifient $x^{-1}H = y^{-1}H$, alors $xy^{-1} \in H$ d'après la première équivalence du (i), d'où $Hx = Hy$ d'après la seconde équivalence du (i). Ce qui prouve que φ est injective, et donc finalement bijective.

(iv) Supposons que $xH = Hx$ pour tout $x \in G$. Alors quels que soient $x \in G$ et $h \in H$, il existe $h' \in H$ tel que $xh = h'x$, d'où $xhx^{-1} \in H$. Ceci prouve que $H \lhd G$. Réciproquement, supposons que $H \lhd G$. Fixons $x \in G$ quelconque. Tout élément z de xH s'écrit $z = xh$ avec $h \in H$, donc $z = xhx^{-1}x$. Mais $xhx^{-1} \in H$ par normalité de H, de sorte que $z = (xhx^{-1})x \in Hx$. Ceci prouve que $xH \subseteq Hx$. L'inclusion réciproque se montre de même, ce qui achève la preuve. □

6.1.3 Définition. Soient G un groupe et H un sous-groupe de G. On appelle *indice de H dans G*, noté $[G : H]$, le cardinal de l'ensemble des classes modulo H (à droite ou à gauche indifféremment d'après le (iii) du lemme précédent). On dit que H est d'indice fini lorsque ce cardinal est fini.

6.1.4 Théorème. *Si G est un groupe fini, alors tout sous-groupe H de G est d'indice fini dans G, et on a d'après le théorème de Lagrange l'égalité:*

$$|G| = |H| \times [G : H].$$

Preuve. Notons $|G| = m$, $|H| = n$, et $[G : H] = p$. Par définition, p est le nombre de classes (à gauche par exemple) modulo H. Or, d'après la remarque (a) de 6.1.1, chacune des classes est en bijection avec H, donc admet exactement n éléments. Il résulte alors du point (ii) du lemme 6.1.2 que $m = pn$. □

A noter qu'un sous-groupe infini H d'un groupe infini G peut très bien être d'indice fini (prendre par exemple $G = \mathrm{O}_n(\mathbb{R})$ et $H = \mathrm{SO}_n(\mathbb{R})$).

6.1.5 Proposition. *Soit G un groupe (fini ou non). Tout sous-groupe H d'indice 2 dans G est normal dans G.*

Preuve. Par hypothèse, il n'y a que deux classes à gauche modulo H; l'une est $eH = H$ qui, d'après le point (i) de 6.1.2, est aussi la classe de tout élément de H, et l'autre yH (avec $y \notin H$) est donc la classe commune à tous les éléments de G qui ne sont pas dans H. De même, il n'y a que deux classes à droite modulo H; l'une est $He = H$ qui, d'après le point (i) de 6.1.2, est aussi la classe de tout élément de H, et l'autre Hz (avec $z \notin H$) est donc la classe commune à tous les éléments de G qui ne sont pas dans H. Comme les classes forment une partition de G, il en résulte que $yH = Hz$. Dès lors, quel que soit $x \in G$, on a $xH = H = Hx$ si $x \in H$, et $xH = yH = Hz = Hx$ si $x \notin H$. Dans les deux cas, on a $xH = Hx$. Ce qui prouve que $H \lhd G$. □

Dans le cas particulier où G est fini d'ordre m, si H est d'indice 2 dans G, on a (d'après le théorème 6.1.4) m pair et H d'ordre $n = m/2$, de sorte que l'on retrouve la proposition 5.4.4.

6.2 Congruence modulo un sous-groupe normal.

6.2.1 Proposition et définition. *Soient G un groupe et H un sous-groupe de G. On suppose que $H \lhd G$. La relation binaire définie sur G par:*

$$\text{pour tous } x, y \in G,\ \ x \equiv y \ \ \text{lorsque}\ \ xy^{-1} \in H$$

est une relation d'équivalence dans G, appelée la congruence modulo H, ou encore l'équivalence modulo H, dont les classes d'équivalence vérifient:

$$\text{pour tout } x \in G,\ \ \overline{x} = xH = Hx.$$

Preuve. Pour tout $x \in G$, on a $x \equiv x$ puisque $xx^{-1} = e \in H$. Donc $\equiv$ est réflexive. Si $x, y \in G$ vérifient $x \equiv y$, alors $xy^{-1} \in H$, donc par passage à l'inverse $yx^{-1} \in H$, d'où $y \equiv x$. Donc $\equiv$ est symétrique. Si $x, y, z \in G$ vérifient $x \equiv y$ et $y \equiv z$, alors $xy^{-1} \in H$ et $yz^{-1} \in H$, donc par produit $xy^{-1}yz^{-1} \in H$, c'est-à-dire $xz^{-1} \in H$, d'où $x \equiv z$. Donc $\equiv$ est transitive, ce qui achève de montrer que $\equiv$ est une relation d'équivalence.

Pour tout $x \in G$, la classe d'équivalence de x est par définition $\overline{x} = \{y \in G\,;\, y \equiv x\}$. Or $y \equiv x$ si et seulement si $yx^{-1} \in H$, ce qui équivaut à l'existence d'un élément $h \in H$ tel que $y = hx$. Ceci prouve que $\overline{x} = Hx$, et comme $H \lhd G$, il résulte de 6.1.2.(iv) que l'on a aussi $\overline{x} = xH$. □

6.2.2 Remarques. Soient G un groupe et H un sous-groupe normal de G.

(a) Pour tout $x \in G$, $\overline{x}$ est par définition l'ensemble des éléments $y \in G$ tels que $y \equiv x$. Tout élément y de $\overline{x}$ s'appelle un représentant de $\overline{x}$.

$$(\ y \in \overline{x}\ ) \Leftrightarrow (\ y \equiv x\ ) \Leftrightarrow (\ yx^{-1} \in H\ ) \Leftrightarrow (\ \exists\ h \in H,\ y = hx) \Leftrightarrow (y \in Hx\ )$$

Comme $Hx = xH$ puisque $H \lhd G$, on a aussi:

$$(\ y \in \overline{x}\ ) \Leftrightarrow (\ y \in xH\ ) \Leftrightarrow (\ \exists\ h' \in H,\ y = xh') \Leftrightarrow (\ x^{-1}y \in H\ )$$

En particulier, x lui-même est un représentant de sa classe: $x \in \overline{x}$ pour tout $x \in G$.

(b) Deux éléments de G ont la même classe si et seulement s'ils sont congrus modulo H:

$$\text{pour tous } x, y \in G,\ \ \overline{x} = \overline{y}\ \ \text{si et seulement si}\ \ x \equiv y.$$

(c) On a en particulier: $\overline{e} = H$.

6.2.3 NOTATIONS. L'ensemble quotient de G par la relation d'équivalence $\equiv$ (qui est par définition l'ensemble des classes d'équivalence des éléments de G) est ordinairement noté $G/\equiv$. Comme ici la relation $\equiv$ est défini à partir du sous-groupe normal H, on convient de noter G/H l'ensemble quotient.

$$G/H = \{\, \overline{x} \;;\; x \in G \,\}.$$

Rappelons que l'on appelle surjection canonique l'application $G \to G/H$ qui, à tout élément de G, associe sa classe d'équivalence.

$$\begin{aligned} p :\ G &\longrightarrow G/H \\ x &\longmapsto \overline{x} \end{aligned}$$

L'application p est surjective par construction, mais en général non injective (car $p(x) = p(y)$ dès lors que $x \equiv y$, même si $x \neq y$).

Notons enfin que, d'après 6.1.3, le cardinal (fini ou infini) de G/H n'est autre que l'indice $[G : H]$ de H dans G.

6.3 Notion de groupe quotient.

6.3.1 COMMENTAIRE PRÉLIMINAIRE. Soient G un groupe et H un sous-groupe normal de G. Le but du théorème fondamental suivant est de munir l'ensemble quotient G/H d'une structure de groupe, déduite de celle de G. L'idée la plus naturelle pour cela est de définir la loi interne dans G/H par: $\overline{x}.\overline{y} = \overline{xy}$ pour tous $\overline{x}, \overline{y} \in G/H$. Mais il y a un point important auquel il faut faire attention ! Le produit de deux classes ainsi défini ne dépend-il pas des représentants x et y que l'on choisit pour poser $\overline{xy}$? En d'autres termes, si l'on prend d'autres représentants $x' \in \overline{x}$ et $y' \in \overline{y}$, (il n'y a aucune raison alors pour que $xy = x'y'$) est-il clair que $\overline{xy} = \overline{x'y'}$? C'est bien sûr indispensable pour que la définition de la loi dans G/H ait un sens. Et c'est effectivement le cas comme le montrent les calculs ci-dessous.

Supposons que $x' \in \overline{x}$ et $y' \in \overline{y}$. Alors $x'x^{-1} \in H$ et $y'y^{-1} \in H$. On a:

$$(x'y')(xy)^{-1} = x'y'y^{-1}x^{-1} = x'y'y^{-1}(x')^{-1}x'x^{-1} = \big[x'(y'y^{-1})(x')^{-1}\big]x'x^{-1}.$$

Or, $y'y^{-1} \in H$ par hypothèse et donc, parce que H est supposé normal dans G (c'est là qu'intervient cette hypothèse fondamentale), on a aussi $x'(y'y^{-1})(x')^{-1} \in H$. Finalement comme par ailleurs $x'x^{-1} \in H$, on conclut que $[x'(y'y^{-1})(x')^{-1}]x'x^{-1}$ appartient à H comme produit de deux éléments de H. On a ainsi vérifié que $(x'y')(xy)^{-1} \in H$, donc $\overline{xy} = \overline{x'y'}$.

6.3.2 THÉORÈME ET DÉFINITION. *Soient G un groupe et H un sous-groupe de G. On suppose que H est normal dans G.*

(i) *On définit une loi interne dans G/H en posant, indépendamment des représentants choisis:*

$$\overline{x}.\overline{y} = \overline{xy} \quad \textit{pour tous } x, y \in G.$$

(ii) *Cette loi munit l'ensemble G/H d'une structure de groupe, appelé le groupe quotient de G par H.*

(iii) *La surjection canonique $p : G \to G/H$ est alors un morphisme du groupe G dans le groupe quotient G/H.*

Preuve. Le point (i) a été montré ci-dessus en 6.3.1. Pour (ii), l'associativité de la loi définie dans G/H est évidente, car pour tous $x, y, z \in G$ on a $\overline{x}.(\overline{y}.\overline{z}) = \overline{x(yz)} = \overline{(xy)z} = (\overline{x}.\overline{y}).\overline{z}$. De même, pour tout $x \in G$, on a $\overline{x}.\overline{e} = \overline{xe} = \overline{x}$ et $\overline{x}.(\overline{x})^{-1} = \overline{xx^{-1}} = \overline{e}$, ce qui montre que G/H est un groupe. Enfin, le point (iii) est clair puisque, par définition, on a $p(xy) = \overline{xy} = \overline{x}.\overline{y} = p(x)p(y)$ pour tous $x, y \in G$. □

Retenons en particulier que l'élément neutre du groupe quotient G/H est $\overline{e} = H$ et, pour tout $x \in G$, le symétrique dans G/H de $\overline{x}$ est $\overline{x}^{-1} = \overline{x^{-1}}$.

6.3.3 EXEMPLES.

(a) Soit G un groupe. Si l'on prend $H = \{e\}$, alors $\overline{x} = \{x\}$ pour tout $x \in G$ (car $x \equiv y$ est alors équivalent à $xy^{-1} = e$, c'est-à-dire $y = x$). Il en résulte $G/\{e\} \simeq G$, via l'isomorphisme $x \mapsto \overline{x}$.

(b) Soit G un groupe. Si l'on prend $H = G$, alors $\overline{x} = \overline{e}$ pour tout $x \in G$ (car on a trivialement $xe^{-1} \in G$ c'est-à-dire $x \equiv e$ pour tout $x \in G$). Il n'y a donc qu'une seule classe, d'où $G/G = \{\overline{e}\}$ est le groupe trivial à un élément.

(c) Prenons $G = S_3 = \{e, \gamma, \gamma^2, \tau_1, \tau_2, \tau_3\}$ et $H = A_3 = \{e, \gamma, \gamma^2\} \lhd S_3$. On a $\overline{e} = H = \{e, \gamma, \gamma^2\}$ et $\overline{\tau_1} = \tau_1 H = \{\tau_1 e, \tau_1\gamma, \tau_1\gamma^2\} = \{\tau_1, \tau_2, \tau_3\} = \overline{\tau}_2 = \overline{\tau}_3$. Il n'y a que deux classes distinctes, donc $S_3/A_3 = \{\overline{e}, \overline{\tau}_1\} \simeq C_2$.

(d) Prenons $G = \mathrm{O}_n(\mathbb{R})$ et $H = \mathrm{SO}_n(\mathbb{R}) \lhd \mathrm{O}_n(\mathbb{R})$. On a $\overline{e} = \mathrm{SO}_n(\mathbb{R})$. Soit $s \in \mathrm{O}_n(\mathbb{R})$ tel que $s \notin \mathrm{SO}_n(\mathbb{R})$ quelconque. Pour tout élément $t \in \mathrm{O}_n(\mathbb{R})$ tel que $t \notin \mathrm{SO}_n(\mathbb{R})$, on a $st^{-1} \in \mathrm{SO}_n(\mathbb{R})$ (le produit de deux isométries négatives est une isométrie positive). Il n'y a que deux classes distinctes (la classe de toutes les isométries positives qui est égale à $\mathrm{SO}_n(\mathbb{R})$ et la classe de toutes les isométries négatives), donc $\mathrm{O}_n(\mathbb{R})/\mathrm{SO}_n(\mathbb{R}) \simeq C_2$.

Les exemples ci-dessus ne sont que des cas particuliers de résultats généraux que l'on verra un peu plus loin.

6.3.4 Remarques.

(a) Il est clair que, si G est abélien, alors G/H est abélien. Mais réciproquement on peut avoir G/H abélien sans que G le soit (voir les exemples (c) et (d) ci-dessus).

(b) Si H est normal et d'indice fini dans G, alors G/H est fini et l'on a d'après la définition 6.1.3:
$$|G/H| = [G : H].$$

(c) Si G est fini et H est normal dans G, alors G/H est fini et l'on a d'après le théorème 6.1.4:
$$|G/H| = |G|/|H|.$$
Attention, on peut avoir G/H fini sans que ni G ni H le soit (voir l'exemple (d) ci-dessus).

6.3.5 Exercice. Soient G un groupe et $Z(G)$ son centre. On sait que $Z(G) \lhd G$, donc on peut considérer le groupe quotient $G/Z(G)$. Montrer que, si $G/Z(G)$ est monogène, alors G est abélien.

6.4 Premier théorème d'isomorphisme.

6.4.1 Théorème. *Soit G un groupe. Pour tout groupe G' et tout morphisme de groupes $f : G \to G'$, le groupe quotient de G par le sous-groupe normal* $\operatorname{Ker} f$ *est isomorphe au sous-groupe* $\operatorname{Im} f$ *de G'. On note:*
$$\operatorname{Ker} f \lhd G \qquad \textit{et} \qquad G/\operatorname{Ker} f \simeq \operatorname{Im} f.$$

Preuve. On a déjà montré en 5.4.3 que $\operatorname{Ker} f \lhd G$. Pour tout $\overline{x} \in G/\operatorname{Ker} f$, posons $\varphi(\overline{x}) = f(x) \in \operatorname{Im} f$. Cette définition est indépendante du choix du représentant dans $\overline{x}$; en effet, si l'on choisit un autre représentant $y \in \overline{x}$, on a par définition $xy^{-1} \in \operatorname{Ker} f$, donc $f(xy^{-1}) = e$, d'où $f(x)f(y)^{-1} = e$, c'est-à-dire $f(x) = f(y)$, ou encore $\varphi(\overline{x}) = \varphi(\overline{y})$. On définit donc bien une application:
$$\begin{aligned} \varphi : \ G/\operatorname{Ker} f &\longrightarrow \operatorname{Im} f \\ \overline{x} &\longmapsto f(x) \end{aligned}$$
L'application φ est surjective par construction. Il est clair que c'est un morphisme de groupes puisque, pour tous $\overline{x}, \overline{y} \in G/\operatorname{Ker} f$, on a $\varphi(\overline{x}.\overline{y}) = \varphi(\overline{xy}) = f(xy) = f(x)f(y) = \varphi(\overline{x})\varphi(\overline{y})$. Vérifions qu'elle est injective. Pour cela, considérons $\overline{x} \in \operatorname{Ker}\varphi$. On a alors $\varphi(\overline{x}) = e'$, le neutre du groupe d'arrivée G'. D'où $f(x) = e'$, c'est-à-dire $x \in \operatorname{Ker} f$, ou encore $\overline{x} = \overline{e}$. Ceci montre que $\operatorname{Ker}\varphi = \{\overline{e}\}$, donc φ est injective. On conclut que φ est un isomorphisme de groupes de $G/\operatorname{Ker} f$ sur $\operatorname{Im} f$. □

6.4.2 Remarque et exemples. Le théorème ci-dessus peut se déduire d'une forme plus générale que l'on verra plus loin en 10.1.2. La forme particulière $G/\operatorname{Ker} f \simeq \operatorname{Im} f$ est cependant d'un usage tellement fréquent qu'il nous a semblé utile de la dégager immédiatement.

(a) Exemple: considérons le morphisme déterminant dét : $\mathrm{GL}_n(\mathbb{R}) \to \mathbb{R}^*$. Il est clairement surjectif (car pour tout réel non-nul λ, on peut trouver des matrices $A \in \mathrm{GL}_n(\mathbb{R})$ telles que $\det(A) = \lambda$), de sorte que $\operatorname{Im}\det = \mathbb{R}^*$. Par ailleurs, $\operatorname{Ker}\det = \mathrm{SL}_n(\mathbb{R})$ par définition. On conclut que:

$$\mathrm{GL}_n(\mathbb{R})/\mathrm{SL}_n(\mathbb{R}) \simeq \mathbb{R}^*.$$

(b) Exemple: considérons pour $n \geq 2$ le morphisme signature $\epsilon : S_n \to C_2 = \{+1, -1\}$. Il est clairement surjectif, donc $\operatorname{Im}\epsilon = C_2$. Par ailleurs, $\operatorname{Ker}\epsilon = A_n$ par définition. On conclut que:

$$S_n/A_n \simeq C_2.$$

6.4.3 Corollaire. *Pour tout groupe G, on a: $Z(G) \lhd G$ et $G/Z(G) \simeq \operatorname{Int} G$.*

Preuve. Résulte immédiatement de 5.4.2, et du point (iii) de la proposition 5.2.3. □

6.4.4 Corollaire. *Soit G un groupe. On suppose que G est le produit direct de deux sous-groupes H et K. Alors:*

$$(\ H \lhd G \ \text{ et } \ G/H \simeq K\ ) \quad \text{et} \quad (\ K \lhd G \ \text{ et } \ G/K \simeq H\ ).$$

Preuve. D'après la remarque 2.6.3, pour tout élément $x \in G$, il existe $h \in H$ et $k \in K$ uniques tels que $x = hk = kh$; posons $f_1(x) = h$ et $f_2(x) = k$. Il est facile de vérifier (écrivez les détails) que $f_1 : G \to H$ et $f_2 : G \to K$ sont des morphismes de groupes, qu'ils sont surjectifs, de noyaux respectifs $\operatorname{Ker} f_1 = K$ et $\operatorname{Ker} f_2 = H$. D'où le résultat en appliquant le théorème 6.4.1. □

6.5 Exemple: groupe dérivé et abélianisé.

6.5.1 Définitions et notations. Soit G un groupe. Pour tous $x, y \in G$, on appelle *commutateur* de x et y l'élément:

$$[x, y] := x^{-1}y^{-1}xy.$$

L'inverse d'un commutateur est un commutateur, mais le produit de deux commutateurs n'est a priori pas un commutateur. Les commutateurs ne constituent donc pas un groupe; on considère alors le sous-groupe engendré par les commutateurs (qui est ici l'ensemble de tous les produits d'un nombre fini de commutateurs).

On appelle *groupe dérivé* de G, noté $D(G)$, le sous-groupe de G engendré par les commutateurs.

Cette notion n'a d'intérêt que pour des groupes non abéliens, puisqu'il est clair que:

$$(\ G \text{ abélien }) \Leftrightarrow (\ [x, y] = e \text{ pour tous } x, y \in G\ ) \Leftrightarrow (\ D(G) = \{e\}\ )$$

6.5.2 Proposition et définition. *Soit G un groupe.*

(i) *$D(G) \lhd G$,*

(ii) *Pour tout sous-groupe $N \lhd G$, on a: (G/N abélien) $\Leftrightarrow$ ($D(G) \subseteq N$).*

(iii) *En particulier, $G/D(G)$ est un groupe abélien, appelé l'abélianisé de G.*

Preuve. Soient $x, y \in G$ quelconques. Considérons le commutateur $c = x^{-1}y^{-1}xy$. Pour tout $z \in G$, on calcule le conjugué de c par z:

$$zcz^{-1} = zx^{-1}y^{-1}xyz^{-1} = zx^{-1}z^{-1}zy^{-1}z^{-1}zxz^{-1}zyz^{-1} = (zxz^{-1})^{-1}(zyz^{-1})^{-1}(zxz^{-1})(zyz^{-1}).$$

On déduit que zcz^{-1} est un commutateur, et ceci pour tout commutateur c et tout $z \in G$. Soit alors $d \in D(G)$ quelconque. Comme on l'a remarqué en 6.5.1, $d = c_1c_2c_3 \dots c_p$, avec $c_1, c_2, c_3, \dots, c_p$ des commutateurs. Pour tout $z \in G$, il vient $zdz^{-1} = zc_1c_2c_3 \dots c_pz^{-1} = zc_1z^{-1}zc_2z^{-1}zc_3z^{-1} \dots zc_pz^{-1}$. Donc zdz^{-1} est, d'après l'étape précédente, un produit de commutateurs. On conclut que $zdz^{-1} \in D(G)$ pour tous $d \in D(G)$ et $z \in G$. Ce qui prouve (i).

Pour (ii), fixons un sous-groupe N normal dans G. Supposons G/N abélien. Pour tous $x, y \in G$, on a $\overline{xy} = \overline{yx}$, donc $\overline{x^{-1}y^{-1}xy} = \overline{e}$, ou encore $x^{-1}y^{-1}xy \in N$. Ainsi, le sous-groupe N contient tous les commutateurs d'éléments de G. Comme $D(G)$ est par définition le plus petit sous-groupe de G qui contient les commutateurs, on conclut que $D(G) \subseteq N$. La réciproque s'obtient en remontant les mêmes calculs. Ceci prouve (ii), et (iii) s'en déduit immédiatement pour $N = D(G)$. □

6.6 Exemple: quotients $\mathbb{Z}/n\mathbb{Z}$.

6.6.1 Attention ! Ce paragraphe étant consacré aux sous-groupes de $\mathbb{Z}$ et aux quotients correspondants, on abandonne provisoirement la notation multiplicative: le groupe $\mathbb{Z}$ est muni de l'addition usuelle des entiers, on note naturellement $x+y$ ce que l'on notait xy dans le cadre général, le neutre que l'on notait e est ici 0, le symétrique de x (que l'on notait x^{-1} dans le cadre général) est ici l'opposé $-x$, et on note $nx = x+x+\cdots+x$ ce que l'on notait x^n (pour $n \in \mathbb{Z}$). On pose enfin:

$$n\mathbb{Z} = \{nx\,;\, x \in \mathbb{Z}\}, \text{ pour tout } n \in \mathbb{Z}.$$

6.6.2 Proposition. *Le groupe additif $\mathbb{Z}$ vérifie les propriétés suivantes.*

(i) *Le groupe $\mathbb{Z}$ muni de l'addition est monogène infini, les générateurs de $\mathbb{Z}$ sont 1 et -1, et tout groupe monogène infini est isomorphe à $\mathbb{Z}$.*

(ii) *Pour tout $n \in \mathbb{Z}$, l'ensemble $n\mathbb{Z}$ est un sous-groupe de $\mathbb{Z}$, qui est le sous-groupe engendré par n, et l'on a: $n\mathbb{Z} = n'\mathbb{Z}$ si et seulement si $n' = n$ ou $n' = -n$.*

(iii) *Réciproquement, pour tout sous-groupe H de $\mathbb{Z}$, il existe un unique $n \in \mathbb{N}$ tel que $H = n\mathbb{Z}$.*

Preuve. L'isomorphisme du (i) a été vu au 3.5.1. Le reste se déduit alors immédiatement des résultats analogues vus en notation multiplicative à la leçon 3, en particulier 3.3.4 et 3.2.2. □

6.6.3 Remarques. Le groupe $\mathbb{Z}$ étant abélien, tous ses sous-groupes sont normaux; d'après la proposition précédente, ils sont de la forme $n\mathbb{Z}$ avec $n \in \mathbb{N}$. Par définition de la congruence modulo le sous-groupe $n\mathbb{Z}$, on a pour tous $x, y \in \mathbb{Z}$:

$$(\ x \text{ et } y \text{ congrus modulo } n\mathbb{Z}\ ) \Leftrightarrow (\ x - y \in n\mathbb{Z}\ ) \Leftrightarrow (\ \text{il existe } z \in \mathbb{Z} \text{ tel que } x - y = nz\ ).$$

On retrouve donc la notion de congruence modulo n de l'arithmétique élémentaire.
Le groupe quotient est naturellement noté $\mathbb{Z}/n\mathbb{Z}$. Ses éléments sont notés $\overline{x}$, avec $x \in \mathbb{Z}$. Sa loi est l'addition déduite de celle de $\mathbb{Z}$ par passage aux classes, c'est-à-dire:

$$\overline{x+y} = \overline{x} + \overline{y} \text{ pour tous } x, y \in \mathbb{Z}.$$

En particulier, son neutre est $\overline{0} = n\mathbb{Z}$.

Convention. Puisqu'il résulte des exemples 6.3.3.(a) et 6.3.3.(b) que $\mathbb{Z}/0\mathbb{Z} = \mathbb{Z}/\{0\} \simeq \mathbb{Z}$ et que $\mathbb{Z}/1\mathbb{Z} = \mathbb{Z}/\mathbb{Z} = \{\overline{e}\}$, on ne considérera plus dans la suite les cas triviaux $n = 0$ et $n = 1$.

6.6.4 Théorème. *Fixons un entier $n > 1$.*

(i) *Le groupe $\mathbb{Z}/n\mathbb{Z}$ est fini d'ordre n, et l'on a $\mathbb{Z}/n\mathbb{Z} = \{\overline{0}, \overline{1}, \overline{2}, \ldots, \overline{n-1}\}$.*

(ii) *Le groupe additif $\mathbb{Z}/n\mathbb{Z}$ est cyclique d'ordre n, c'est-à-dire isomorphe au groupe cyclique C_n.*

(iii) *Les générateurs du groupe $\mathbb{Z}/n\mathbb{Z}$ sont les classes $\overline{k}$ des entiers k qui sont premiers avec n.*

(iv) *Pour tout diviseur q de n, il existe un et un seul sous-groupe de $\mathbb{Z}/n\mathbb{Z}$ d'ordre q, qui est le sous-groupe cyclique engendré par $\overline{d}$, où $n = dq$.*

Preuve. D'après 3.5.2, il n'existe à isomorphisme près qu'un groupe cyclique d'ordre n, noté $C_n = \{e, x, x^2, x^3, \ldots, x^{n-1}\}$ suivant la proposition 3.1.5. Définissons alors l'application:

$$\begin{array}{rccc} f: & \mathbb{Z} & \longrightarrow & C_n \\ & k & \longmapsto & x^k \end{array}$$

On a $f(k+h) = x^{k+h} = x^k x^h = f(k)f(h)$ pour tous $h, k \in \mathbb{Z}$, ce qui prouve que f est un morphisme de groupes. Il est clair que f est surjective puisque tout élément de C_n est de la forme x^k pour un entier k. Enfin un entier k appartient à $\operatorname{Ker} f$ si et seulement si $x^k = e$, ce qui, puisque x est d'ordre n, équivaut au fait que k est multiple de n; en d'autres termes $\operatorname{Ker} f = n\mathbb{Z}$. On déduit alors du théorème 6.4.1 que $\mathbb{Z}/n\mathbb{Z} \simeq C_n$. Les points (i), (iii) et (iv) se déduisent alors immédiatement des résultats analogues démontrés en notation multiplicative (en particulier 3.2.3 et 3.3.2). □

EXEMPLES ($2 \leq n \leq 5$):

	$\overline{0}$	$\overline{1}$
$\overline{0}$	$\overline{0}$	$\overline{1}$
$\overline{1}$	$\overline{1}$	$\overline{0}$

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

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

	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$
$\overline{0}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$
$\overline{1}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{0}$
$\overline{2}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{0}$	$\overline{1}$
$\overline{3}$	$\overline{3}$	$\overline{4}$	$\overline{0}$	$\overline{1}$	$\overline{2}$
$\overline{4}$	$\overline{4}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$

6.6.5 COROLLAIRE (écriture additive du théorème chinois). *Pour tous $n > 1$ et $m > 1$, on a:*

$$(\mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z} \simeq \mathbb{Z}/nm\mathbb{Z}) \Leftrightarrow (n \text{ et } m \text{ premiers entre eux}).$$

EXEMPLES:

	$(\overline{0},\overline{0})$	$(\overline{0},\overline{1})$	$(\overline{1},\overline{0})$	$(\overline{1},\overline{1})$
$(\overline{0},\overline{0})$	$(\overline{0},\overline{0})$	$(\overline{0},\overline{1})$	$(\overline{1},\overline{0})$	$(\overline{1},\overline{1})$
$(\overline{0},\overline{1})$	$(\overline{0},\overline{1})$	$(\overline{0},\overline{0})$	$(\overline{1},\overline{1})$	$(\overline{1},\overline{0})$
$(\overline{1},\overline{0})$	$(\overline{1},\overline{0})$	$(\overline{1},\overline{1})$	$(\overline{0},\overline{0})$	$(\overline{0},\overline{1})$
$(\overline{1},\overline{1})$	$(\overline{1},\overline{1})$	$(\overline{1},\overline{0})$	$(\overline{0},\overline{1})$	$(\overline{0},\overline{0})$

$\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ n'est pas cyclique; c'est le groupe de Klein $V = \{e, a, b, c\}$ pour:
$e = (\overline{0},\overline{0})$, $a = (\overline{0},\overline{1})$, $b = (\overline{1},\overline{0})$ et $c = (\overline{1},\overline{1})$.

	$(\overline{0},\widetilde{0})$	$(\overline{1},\widetilde{1})$	$(\overline{0},\widetilde{2})$	$(\overline{1},\widetilde{0})$	$(\overline{0},\widetilde{1})$	$(\overline{1},\widetilde{2})$
$(\overline{0},\widetilde{0})$	$(\overline{0},\widetilde{0})$	$(\overline{1},\widetilde{1})$	$(\overline{0},\widetilde{2})$	$(\overline{1},\widetilde{0})$	$(\overline{0},\widetilde{1})$	$(\overline{1},\widetilde{2})$
$(\overline{1},\widetilde{1})$	$(\overline{1},\widetilde{1})$	$(\overline{0},\widetilde{2})$	$(\overline{1},\widetilde{0})$	$(\overline{0},\widetilde{1})$	$(\overline{1},\widetilde{2})$	$(\overline{0},\widetilde{0})$
$(\overline{0},\widetilde{2})$	$(\overline{0},\widetilde{2})$	$(\overline{1},\widetilde{0})$	$(\overline{0},\widetilde{1})$	$(\overline{1},\widetilde{2})$	$(\overline{0},\widetilde{0})$	$(\overline{1},\widetilde{1})$
$(\overline{1},\widetilde{0})$	$(\overline{1},\widetilde{0})$	$(\overline{0},\widetilde{1})$	$(\overline{1},\widetilde{2})$	$(\overline{0},\widetilde{0})$	$(\overline{1},\widetilde{1})$	$(\overline{0},\widetilde{2})$
$(\overline{0},\widetilde{1})$	$(\overline{0},\widetilde{1})$	$(\overline{1},\widetilde{2})$	$(\overline{0},\widetilde{0})$	$(\overline{1},\widetilde{1})$	$(\overline{0},\widetilde{2})$	$(\overline{1},\widetilde{0})$
$(\overline{1},\widetilde{2})$	$(\overline{1},\widetilde{2})$	$(\overline{0},\widetilde{0})$	$(\overline{1},\widetilde{1})$	$(\overline{0},\widetilde{2})$	$(\overline{1},\widetilde{0})$	$(\overline{0},\widetilde{1})$

$\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$ est cyclique, engendré par $x = (\overline{1},\widetilde{1})$.

Leçon 7

Groupe opérant sur un ensemble

7.1 Actions de groupes : les notions de base.

7.1.1. Définition. Soient G un groupe et E un ensemble non vide. On dit que G *opère* (à gauche) sur E s'il existe une loi externe:

$$\begin{array}{rcl} G \times E & \longrightarrow & E \\ (g,x) & \longmapsto & g.x \end{array}$$

qui satisfait les deux conditions:

$$(1) \quad \forall\, g,g' \in G,\ \forall\, x \in E,\ g.(g'.x) = (gg').x; \qquad (2) \quad \forall\, x \in E,\ e.x = x.$$

On dit aussi que l'on a une <u>action</u> de G sur E.

7.1.2. Théorème. *La donnée d'une action d'un groupe G sur un ensemble non-vide E équivaut à la donnée d'un morphisme de groupes de G dans le groupe symétrique $S(E)$.*

Preuve. Supposons donné un morphisme de groupes $\gamma : G \longrightarrow S(E)$; notons: $g \longmapsto \gamma_g$ (où γ_g désigne ici l'image de g par γ). On définit une loi externe $G \times E \longrightarrow E$ en posant $g.x = \gamma_g(x)$ pour tous $g \in G, x \in E$. En utilisant le fait que $\gamma_g \circ \gamma_{g'} = \gamma_{gg'}$ et $\gamma_e = \mathrm{id}_E$, on vérifie sans problème que les deux conditions (1) et (2) d'une action sont vérifiées.

Réciproquement, supposons que G opère sur E par $(g,x) \longmapsto g.x$. Définissons pour tout $g \in G$ une application $\gamma_g : E \longrightarrow E$ par $\gamma_g(x) = g.x$ quel que soit $x \in E$. On a alors:

$$\forall\ x \in E,\ \forall\ g,h \in G,\ \ \gamma_g\gamma_h(x) = g.(h.x) = (gh).x = \gamma_{gh}(x) \quad \text{et} \quad \gamma_e(x) = e.x = x,$$

ce qui prouve que $\gamma_g\gamma_h = \gamma_{gh}$ et $\gamma_e = \mathrm{id}_E$. On en déduit que γ_g est bijective pour tout $g \in G$ (en prenant $h = g^{-1}$), puis que l'application $\gamma : G \to S(E)$ est un morphisme de groupes. □

7.1.3. Définition et proposition. *Soit G un groupe opérant sur un ensemble non-vide E. Pour tout $x \in E$, on appelle <u>stabilisateur</u> de x l'ensemble:*

$$G_x = \{g \in G\,;\, g.x = x\} \ .$$

C'est un sous-groupe de G ; on le note aussi parfois $\mathrm{Stab}_G(x)$.

Preuve. On a $e.x = x$ donc $e \in G_x$. De plus, quels que soient $g,h \in G_x$, on calcule: $(gh^{-1}).x = (gh^{-1}).(h.x) = (gh^{-1}h).x = g.x = x$. D'où $gh^{-1} \in G_x$. □

7.1.4. Proposition et définition. *Soit G un groupe opérant sur un ensemble non-vide E. Pour tout $x \in E$, on appelle <u>orbite</u> de x l'ensemble:*

$$\Omega_x = \{y \in E\,;\, \exists\, g \in G,\ y = g.x\} = \{g.x\,;\, g \in G\} \ .$$

Les orbites des éléments de E sous l'action de G forment une partition de E.

Preuve. On vérifie aisément que la relation $\mathcal{R}$ définie sur E par: $(\forall\, x,y \in E,\ x\,\mathcal{R}\,y \ \Leftrightarrow\ \exists\, g \in G,\ y = g.x)$ est une relation d'équivalence. La classe d'équivalence pour $\mathcal{R}$ d'un élément $x \in E$ n'est autre par définition que l'orbite Ω_x. D'où le résultat puisque les classes d'équivalence forment une partition de E. □

7.1.5. Définitions. Soit G un groupe opérant sur un ensemble non-vide E.

(a) Le noyau du morphisme $\gamma : G \longrightarrow S(E)$ canoniquement associé à l'action de G sur E est appelé le noyau de l'action. Il est clair que c'est l'intersection des stabilisateurs:

$$\operatorname{Ker}\gamma = \{g \in G\,;\, \forall\, x \in E,\ g.x = x\} = \bigcap_{x\in E} G_x \quad .$$

On dit que l'action est <u>fidèle</u> (ou que G opère fidèlement sur E) lorsque γ est injectif, c'est-à-dire lorsque le noyau de l'action est $\{e\}$.

Si G opère fidèlement sur E, alors G est isomorphe à un sous-groupe de $S(E)$, à savoir $\operatorname{Im}\gamma$.

(b) Un élément $x \in E$ est appelé un <u>point fixe</u> de l'action de G lorsque $g.x = x$ pour tout $g \in G$,

On note E^G ou $\operatorname{Fix}_G(E)$ l'ensemble des points fixes de l'action de G sur E. Pour tout $x \in E$, on a:

$$(x \in E^G) \ \Leftrightarrow\ (G_x = G) \ \Leftrightarrow (\Omega_x = \{x\}), \qquad \text{(orbite ponctuelle)}.$$

On dit que l'action est sans point fixe lorsque $E^G = \emptyset$.

(c) On dit que G opère <u>transitivement</u> sur E, ou encore que l'action de G sur E est <u>transitive</u>, lorsqu'il n'y a qu'une seule orbite:

$$\text{(action transitive)} \ \Leftrightarrow\ (\forall\, x \in E,\ \Omega_x = E) \ \Leftrightarrow\ (\forall\, x, y \in E,\ \exists\, g \in G,\ y = g.x).$$

7.1.6 Exercice. Montrer que $G = \{\left(\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right)\ ;\ a \in \mathbb{R}_+^*, b \in \mathbb{R}\}$ est un sous-groupe de $\mathrm{GL}_2(\mathbb{R})$. Montrer que, en posant $g.z = az + b$ pour toute $g = \left(\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right)$ et tout $z \in \mathbb{C}$, on définit une action du groupe G sur l'ensemble $\mathbb{C}$. Montrer qu'il existe 3 orbites pour cette action: $\Omega_0 = \mathbb{R}$, $\Omega_i = \{z \in \mathbb{C}\,;\, \operatorname{Im} z > 0\}$ et $\Omega_{-i} = \{z \in \mathbb{C}\,;\, \operatorname{Im} z < 0\}$. Expliciter le stabilisateur G_z pour tout $z \in \mathbb{C}$, et montrer que $G_z = \{I_2\}$ lorrsque $z \notin \mathbb{R}$. Montrer que l'action est fidèle et sans point fixe.

7.1.7 Exercice. Soit E l'ensemble $\mathbb{C} \setminus \{-1, 0, 1\}$. Montrer que l'application $\sigma : E \to \mathbb{C}$ définie par $\sigma(z) = \frac{z-1}{z+1}$ est une bijection de E sur E et que σ est d'ordre 4 dans le groupe $S(E)$. On note G le sous-groupe cyclique d'ordre 4 engendré par σ. Montrer que G opère sur E par $s.z = s(z)$ pour tous $s \in G, z \in \mathbb{C}$. Montrer que le stabilisateur d'un élément $z \in \mathbb{C}$ est $G_z = G$ si $z = \pm i$, et $G_z = \{\mathrm{id}_E\}$ sinon. En déduire que $E^G = \{-i, +i\}$.

7.2 Exemples classiques.

7.2.1. Action d'un groupe sur lui-même par translation.

- *Tout groupe G opère sur lui-même par translation à gauche:* $\left|\begin{array}{ccc} G \times G & \longrightarrow & G \\ (g, x) & \longmapsto & g.x = gx \end{array}\right.$

(a) Pour tout $x \in G$, $G_x = \{g \in G\,;\, gx = x\} = \{e\}$. On en déduit que $\operatorname{Ker}\gamma = \bigcap_{x\in G} G_x = \{e\}$, donc l'action est fidèle.

(b) Pour tout $x \in G$, $\Omega_x = \{gx\,;\, g \in G\} = G$ (car tout $y \in G$ s'écrit $y = (yx^{-1})x$). On en déduit qu'il n'y a qu'une seule orbite, donc l'action est transitive. De plus, dès lors que $G \neq \{e\}$, on a $\Omega_x = G$ non ponctuelle pour tout $x \in G$, donc l'action est sans point fixe.

7.2.2. Action d'un groupe sur lui-même par conjugaison.

- *Tout groupe G opère sur lui même par conjugaison:* $\left|\begin{array}{ccc} G \times G & \longrightarrow & G \\ (g, x) & \longmapsto & g.x = gxg^{-1} \end{array}\right.$

(a) Pour tout $x \in G$, $G_x = \{g \in G\,;\, gx = xg\}$ est le *centralisateur* $C_G(x)$ de x (voir 5.2.2).

Donc $\operatorname{Ker}\gamma = \bigcap_{x\in G} G_x = \{g \in G\,;\, \forall\, x \in G,\ gx = xg\}$ est le *centre* $Z(G)$ (voir 5.2.1).

En particulier l'action est fidèle si et seulement si G est de centre trivial, ie. $Z(G) = \{e\}$.

(b) Pour tout $x \in G$, l'orbite $\Omega_x = \{gxg^{-1}\,;\, g \in G\}$ est la *classe de conjugaison* de x (voir 5.1.3).

On en déduit que Ω_x est ponctuelle si et seulement si x appartient au centre de G, donc l'ensemble des points fixes E^G est non-vide et égal au centre $Z(G)$.

De plus $\Omega_e = \{e\}$, donc $\Omega_e \neq G$ dès lors que $G \neq \{e\}$, et l'action n'est alors pas transitive.

7.2.3. Action d'un groupe sur l'ensemble de ses parties par conjugaison.

• *Tout groupe G opère sur $\mathcal{P}(G)$ par conjugaison:* $\left|\begin{array}{ccc} G \times \mathcal{P}(G) & \longrightarrow & \mathcal{P}(G) \\ (g, X) & \longmapsto & g.X = gXg^{-1} \end{array}\right.$

(a) Pour tout $X \in \mathcal{P}(G)$, $G_X = \{g \in G\,;\, gXg^{-1} = X\}$ est le normalisateur de X, noté $N_G(X)$.
Dans le cas où $X = H$ est un sous-groupe de G, on retrouve la notion de normalisateur vue en 5.5.

(b) On en déduit que $X \in \mathcal{P}(G)^G$ si et seulement si $N_G(X) = G$.
Dans le cas où $X = H$ est un sous-groupe de G, cela signifie que $H \lhd G$.

(c) On en déduit aussi que $\operatorname{Ker}\gamma = Z(G)$.
En effet, $\operatorname{Ker}\gamma = \bigcap_{X\subseteq G} G_X = \bigcap_{X\subseteq G} N_G(X)$. En considérant parmi les $X \subseteq G$ celles qui sont des singletons, il vient: $\operatorname{Ker}\gamma \subseteq \bigcap_{x\in G} N_G(\{x\}) = \bigcap_{x\in G} C_G(x) = Z(G)$. La réciproque est claire.

(d) Pour $X \in \mathcal{P}(G)$, $\Omega_X = \{gXg^{-1}\,;\, g \in G\}$ est la classe de conjugaison de X dans $\mathcal{P}(G)$.
En particulier l'action n'est pas transitive, car $\Omega_\emptyset = \{\emptyset\} \neq \mathcal{P}(G)$.

7.3 Indice des stabilisateurs et équation aux classes.

7.3.1. Proposition. *Soit G un groupe opérant sur un ensemble E. Si deux éléments x et y de E appartiennent à une même orbite, alors leurs stabilisateurs G_x et G_y sont conjugués dans G.*

Preuve. Soient $x \in E$ et $y \in \Omega_x$; il existe donc $g \in G$ tel que $y = g.x$. Montrons que: $G_y = gG_xg^{-1}$.

Soit $h \in G_y$. On a $y = h.y$, c'est-à-dire $g.x = h.(g.x) = hg.x$. On en tire: $x = e.x = g^{-1}g.x = g^{-1}.(g.x) = g^{-1}.(hg.x) = (g^{-1}hg).x$, donc $g^{-1}hg \in G_x$, ou encore $h \in gG_xg^{-1}$.

Réciproquement, soit $k \in gG_xg^{-1}$. On a $g^{-1}kg \in G_x$, donc $(g^{-1}kg).x = x$, d'où $kg.x = g.x$, c'est-à-dire $k.y = y$, ou encore $k \in G_y$. □

7.3.2. Théorème. *Soit G un groupe opérant sur un ensemble E.*

(i) *Pour tout $x \in E$, le cardinal de l'orbite Ω_x est égal à l'indice du stabilisateur G_x. On note:*

$$|\Omega_x| = [G : G_x] \quad .$$

(ii) *En particulier, si G est fini, $|\Omega_x|$ divise $|G|$.*

Preuve. On fixe $x \in E$. On note Q_{G_x} l'ensemble des classes à gauche modulo le sous-groupe G_x (voir 6.1). On montre que Ω_x et Q_{G_x} sont équipotents en construisant explicitement une bijection λ de Ω_x sur Q_{G_x}. Pour cela, à tout élément de Ω_x, donc de la forme $g.x$ pour un certain $g \in G$, on associe $\lambda(g.x) := gG_x$. L'application $\lambda : \Omega_x \longrightarrow Q_{G_x}$ ainsi construite est bien définie [en effet, pour tout autre $h \in G$ tels que $g.x = h.x$, on a $h^{-1}.(g.x) = h^{-1}.(h.x)$, donc $(h^{-1}g).x = (h^{-1}.h).x = e.x = x$, donc $h^{-1}g \in G_x$, d'où $gG_x = hG_x$, ie. $\lambda(g.x) = \lambda(h.x)$]. L'application λ est surjective par construction [tout élément de Q_{G_x} est de la forme gG_x pour un $g \in G$, et donc $\lambda(g.x) = gG_x$]. Montrons l'injectivité : si $g, h \in G$ vérifient $\lambda(g.x) = \lambda(h.x)$, alors $gG_x = hG_x$, donc $h^{-1}g \in G_x$, c'est-à-dire $(h^{-1}g).x = x$, d'où $h.x = h.((h^{-1}g).x) = (hh^{-1}g).x = g.x$. Ceci prouve (i). Le point (ii) découle alors de 6.1.4. □

7.3.3. Corollaire. *Soit G un groupe opérant sur un ensemble fini E. Soit $\{x_i\}_{1\leq i\leq r}$ une famille de représentants des orbites distinctes. On a:*

$$|E| = \sum_{i=1}^{r} [G : G_{x_i}] \quad .$$

Preuve. L'ensemble E étant fini, il y a un nombre fini r d'orbites distinctes. Choisissons un représentant x_i dans chacune de ces r orbites. Les Ω_{x_i} pour $1 \leq i \leq r$ forment une partition de E, donc $|E| = \sum_{i=1}^{r} |\Omega_{x_i}|$, d'où le résultat en appliquant le théorème précédent. □

7.3.4. Exemples d'application.

(i) *Si G est un groupe fini d'ordre 33 opérant sur un ensemble E fini de cardinal 19, alors l'action admet forcément des points fixes.*

En effet, toute orbite est d'ordre 1, 3, 11 ou 33. Comme $33 > |E|$, seuls 1, 3 et 11 restent possibles. Si E^G était vide, il n'y aurait pas d'orbite ponctuelle, et on aurait donc en tout et pour tout n orbites à 3 éléments et m orbites à 11 éléments, d'où $3n + 11m = 19$. Cette équation n'ayant pas de solutions dans $\mathbb{N}$, on conclut que $E^G \neq \emptyset$.

(ii) *Soit G un groupe fini d'ordre 15 opérant sans point fixe sur un ensemble E fini de cardinal 17; donner le nombre d'orbites et le cardinal de chacune d'elles.*

Toute orbite est de cardinal 3, 5 ou 15, puisqu'il n'y a pas de points fixes, donc pas d'orbites ponctuelles. S'il y avait une orbite à 15 éléments (il ne peut de toute façon pas y en avoir plus...), les deux éléments restants de E ne pourraient pas former une orbite. C'est donc qu'il n'y a pas d'orbites à 15 éléments. On a donc en tout et pour tout n orbites à 3 éléments et m orbites à 5 éléments, d'où $3n + 5m = 17$. La seule solution dans $\mathbb{N}$ est $n = 4$ et $m = 1$.

7.4. Equation aux classes pour un groupe fini, application aux p-groupes.

7.4.1. Théorème. *Soit G un groupe fini. Pour tout $x \in G$, on note $C_G(x)$ le centralisateur de x dans G. On note $Z(G)$ le centre de G.*

(i) *Le cardinal de la classe de conjugaison de tout élément de G divise $|G|$.*

(ii) *Soit $\{x_i\}_{1\leq i\leq r}$ une famille de représentants des classes de conjugaison distinctes dans G. Alors:*
$$|G| = \sum_{i=1}^{r} [G : C_G(x_i)] \quad .$$

(iii) *Soit $\{x_i\}_{1\leq i\leq k}$ une famille de représentants des classes de conjugaison distinctes* non ponctuelles *dans G. Alors:*
$$|G| = |Z(G)| + \sum_{i=1}^{k} [G : C_G(x_i)] \quad .$$

Preuve. Pour l'action de G sur lui-même par conjugaison (voir 7.2.2), l'orbite d'un élément $x \in G$ est sa classe de conjugaison $\Omega_x = \mathrm{cl}x$, et son stabilisateur est son centralisateur $C_G(x)$. Les points (i) et (ii) sont donc des conséquences immédiates du théorème 7.3.2 et du corollaire 7.3.3.

Pour montrer (iii), rappelons d'abord (voir 7.2.2.b) que l'orbite d'un élément $x \in G$ est ponctuelle si et seulement si $x \in Z(G)$. Notons alors $x_1, \ldots, x_k$ des représentants des orbites non ponctuelles (il peut ne pas y en avoir, auquel cas $k = 0$ et G est abélien), et $x_{k+1}, \ldots, x_r$ des représentants des orbites ponctuelles (il y en a toujours au moins une, celle de e, donc $k < r$). Pour $k + 1 \leq i \leq r$, on a $x_i \in Z(G)$ et $|\Omega_{x_i}| = 1$. Pour $1 \leq i \leq k$, on a $x_i \notin Z(G)$ et $|\Omega_{x_i}| = [G : C_G(x_i)] \neq 1$. Donc: $|G| = \sum_{i=1}^{k} |\Omega_{x_i}| + \sum_{i=k+1}^{r} |\Omega_{x_i}| = \sum_{i=1}^{k} [G : C_G(x_i)] + |Z(G)|$. □

On développe ci-dessous, comme exemples d'applications concrètes des résultats précédents à la théorie des groupes finis, quelques propriétés des groupes finis les plus simples, ceux dont l'ordre n'a qu'un seul diviseur premier. C'est le point de départ d'une théorie plus élaborée (celle des théorèmes de Sylow), qui pourra être vue dans des cursus ultérieurs.

7.4.2. Définition. Soit p un nombre premier. On appelle p-groupe tout groupe fini dont l'ordre est une puissance de p.

7.4.3. Lemme. *Si G est un p-groupe non trivial opérant sur un ensemble fini non-vide E, alors: $|E^G| \equiv |E| \bmod p$.*

Preuve. On sait (voir 7.1.5.b) que $|E^G|$ est le nombre d'orbites ponctuelles. Si toutes les orbites sont ponctuelles, alors $|E| = |E^G|$ et le résultat est clair. Sinon, on note $\Omega_{x_1}, \ldots, \Omega_{x_k}$ les orbites non ponctuelles. Donc: $|E| = |E^G| + \sum_{i=1}^{k} |\Omega_{x_i}|$. Pour tout $1 \leq i \leq k$, on a $|\Omega_{x_i}|$ divise $|G|$ d'après 7.3.2. Puisque $|G|$ est de la forme p^n avec $n \in \mathbb{N}^*$, on déduit que $|\Omega_{x_i}| = p^{m_i}$ avec $m_i \leq n$. Mais de plus $m_i \geq 1$ puisque $|\Omega_{x_i}| \neq 1$. On conclut que $|E| - |E^G| = \sum_{i=1}^{k} p^{m_i}$ est divisible par p. □

7.4.4. Proposition. *Le centre d'un p-groupe non trivial est non trivial.*

Preuve. On applique ce qui précède à l'action de G sur lui-même par conjugaison (voir 7.2.2): on a $E = G$ et $E^G = Z(G)$. Le lemme 7.4.3 implique donc que $|G| - |Z(G)|$ est divisible par p. Comme $|G|$ est divisible par p, on conclut que $|Z(G)|$ est divisible par p. Donc $|Z(G)| \neq 1$. □

7.4.5. Corollaire. *Si p est un nombre premier, tout groupe d'ordre p^2 est abélien.*

Preuve. Soit G un groupe d'ordre p^2. D'après le théorème de Lagrange, $|Z(G)|$ divise p^2. Comme $|Z(G)| \neq 1$ d'après la proposition précédente, on a donc $|Z(G)| = p$ ou $|Z(G)| = p^2$. Si $|Z(G)| = p^2$, alors $G = Z(G)$, donc G est abélien. Si $|Z(G)| = p$, alors $|G/Z(G)| = p$ (voir 6.3.4.c), donc le groupe $G/Z(G)$ est cyclique (voir 3.4). Or on a vu (voir 6.3.5) que $G/Z(G)$ monogène implique G abélien. □

7.4.6. Exercice. Montrer que: *Si G est un groupe non-abélien d'ordre p^3 avec p premier, alors, quels que soient $x \in G$ et $y \in G$ tels que $xy \neq yx$, le groupe G est engendré par x et y.*

Solution. Soit $H = \langle x, y \rangle$ le sous-groupe engendré par x et y dans G (voir 3.1.1). D'après le théorème de Lagrange son ordre divise p^3. En appliquant le corollaire précédent, $|H| \neq p^2$ car H non abélien puisque $xy \neq yx$. De plus $|H| \neq p$ car sinon H serait cyclique donc abélien. Enfin $|H| \neq 1$ car H contient au moins x et y. On conclut que $|H| = p^3$, donc $H = G$. □

Leçon 8

Groupes d'isométries

En vue de la leçon suivante, ce chapitre présente une synthèse des principaux résultats à connaître sur les isométries (voir le cours de géométrie de L2 pour les preuves, qui ne sont pas rappelées ici), et donne une interprétation de certains d'entre eux en termes de théorie des groupes, à la lumière des résultats nouveaux des leçons précédentes.

8.1 Groupes d'isométries vectorielles.

8.1.1. Rappels sur le groupe linéaire. Soit E un $\mathbb{R}$-espace vectoriel de dimension finie n.

(a) - L'ensemble des automorphismes d'espaces vectoriels de E (ie. des applications linéaires bijectives de E dans E) est un groupe (pour le loi de composition $\circ$) appelé le *groupe linéaire* de E, noté $\mathrm{GL}(E)$. Pour une base fixée de E, l'application $m : \mathrm{GL}(E) \to \mathrm{GL}_n(\mathbb{R})$ qui, à tout automorphisme de E associe sa matrice dans cette base, est un isomorphisme de groupes; donc:

$$\mathrm{GL}(E) \simeq \mathrm{GL}_n(\mathbb{R}).$$

(b) - Pour un même automorphisme $f \in \mathrm{GL}(E)$, les matrices de f relativement à deux bases de E sont conjuguées dans $\mathrm{GL}_n(\mathbb{R})$, donc admettent le même déterminant ; cela permet de définir le déterminant de f comme le déterminant de la matrice de f dans toute base de E. Il est clair (voir aussi 2.1.2.a) que $\det(g \circ f) = \det(f).\det(g)$ pour tous $f, g \in \mathrm{GL}(E)$, et donc:

$$\det : \mathrm{GL}(E) \to \mathbb{R}^* \text{ est un morphisme de groupes.}$$

(c) - Le noyau de $\det$, c'est-à-dire le sous-ensemble de $\mathrm{GL}(E)$ formé des automorphismes de déterminant 1, est un sous-groupe normal de $\mathrm{GL}(E)$ (voir 5.4.3) ; on l'appelle le *groupe spécial linéaire* de E et on le note $\mathrm{SL}(E)$. Comme le morphisme $\det : \mathrm{GL}(E) \to \mathbb{R}^*$ est évidemment surjectif, on déduit aussi de 6.4.1 que $\mathrm{GL}(E)/\mathrm{SL}(E) \simeq \mathbb{R}^*$. Il est clair enfin que, via le choix d'une base de E par rapport à laquelle on prend les matrices des automorphismes de E, il est isomorphe au groupe $\mathrm{SL}_n(\mathbb{R})$ des matrices carrées d'ordre n de déterminant 1 (voir 2.2.3.a).

$$\mathrm{SL}(E) = \{f \in \mathrm{GL}(E)\,;\, \det f = 1\} \simeq \mathrm{SL}_n(\mathbb{R}), \quad \mathrm{SL}(E) \lhd \mathrm{GL}(E) \text{ et } \mathrm{GL}(E)/\mathrm{SL}(E) \simeq \mathbb{R}^*.$$

(d) - C'est un exercice classique (faites-le !) que de vérifier que les seuls automorphismes de E qui commutent (pour la loi $\circ$) avec tous les automorphismes de E sont les multiples scalaires de l'identité, ie. les applications $\lambda\,\mathrm{id}_E$ où $\lambda \in \mathbb{R}^*$ (on les appelle parfois les homothéties vectorielles). On en déduit que:

$$Z(\mathrm{GL}(E)) \simeq \mathbb{R}^*.$$

Le quotient de $\mathrm{GL}(E)$ par son centre est appelé le *groupe projectif linéaire*, noté $\mathrm{PGL}(E)$.

8.1.2. Rappels sur le groupe orthogonal. On suppose maintenant de plus que l'espace vectoriel E est euclidien, c'est-à-dire muni d'un produit scalaire. On note $\cdot$ ce produit scalaire et $\|\ \|$ la norme associée. On abrège "base orthonormée" en b.o.n.

(a) - Soit f un endomorphisme de E. Les conditions suivantes sont équivalentes:

(i) f conserve la norme (ie. $\|f(u)\| = \|u\|$ pour tout u dans E),
(ii) f conserve le produit scalaire (ie. $f(u) \cdot f(v) = u \cdot v$ pour tous $u, v \in E$),
(iii) f est bijective et f^{-1} est égal à l'endomorphisme adjoint f^* de f (donc $f^* \circ f = \mathrm{id}_E$),
(iv) f transforme toute b.o.n. de E en une b.o.n. de E,
(v) la matrice M de f dans toute b.o.n. de E est orthogonale (ie. ${}^tMM = I_n$).

On appelle *isométrie vectorielle* de E tout endomorphisme de E satisfaisant l'une des conditions équivalentes précédentes. Il en résulte immédiatement que toute isométrie vectorielle est bijective

et que les isométries vectorielles forment un sous-groupe de $\mathrm{GL}(E)$, que l'on appelle le *groupe orthogonal* de E et que l'on note $\mathrm{O}(E)$.

$$\mathrm{O}(E) = \{f \in \mathrm{End}(E)\,;\, \|f(u)\| = \|u\| \text{ pour tout } u \in E\} \text{ sous-groupe de } \mathrm{GL}(E)$$

(b) - Il résulte en outre de (iii) ou (v) que le déterminant d'une isométrie vectorielle ne peut valoir que 1 ou -1. Le noyau du morphisme surjectif dét : $\mathrm{O}(E) \to \{1, -1\}$ s'appelle le *groupe spécial orthogonal* de E, noté $\mathrm{SO}(E)$ ou $\mathrm{O}^+(E)$; il est d'indice 2 dans $O(E)$.

$$\mathrm{SO}(E) = \mathrm{O}^+(E) = \mathrm{SL}(E) \cap \mathrm{O}(E) = \{f \in \mathrm{O}(E)\,;\, \text{dét } f = 1\} \text{ sous-groupe de } \mathrm{O}(E),$$
$$\mathrm{O}^+(E) \lhd \mathrm{O}(E), \quad \mathrm{O}(E)/\mathrm{O}^+(E) \simeq \{1, -1\}, \quad [\mathrm{O}(E) : \mathrm{O}^+(E)] = 2.$$

Les éléments du sous-groupe $\mathrm{O}^+(E)$ sont appelées les *isométries positives*, ou les *isométries directes*. Les isométries de déterminant -1 sont dites *isométries négatives*, ou les *isométries indirectes*. Attention: l'ensemble $\mathrm{O}^-(E)$ des isométries indirectes n'est pas un sous-groupe de $\mathrm{O}(E)$!

(c) - Il résulte aussi de (i) que les seules valeurs propres possibles pour une isométrie vectorielle sont 1 et -1.

8.1.3. Exemple : symétries orthogonales.

(a) - Soit F un ss-e.v. de E. On note $F^\perp$ le supplémentaire orthogonal de F dans E. Tout vecteur $u \in E$ se décompose de façon unique sous la forme $u = v + w$ avec $v \in F$ et $w \in F^\perp$. On pose alors $s_F(u) = v - w$. L'application $s_F : E \to E$ ainsi définie s'appelle la *symétrie orthogonale* par rapport à F.

C'est un endomorphisme de E qui vérifie $s_F(u) = u$ pour tout u de F et $s_F(u) = -u$ pour tout vecteur orthogonal à F.

Il en résulte que, si l'on note $p = \dim F$, et si l'on choisit une b.o.n. $\mathcal{B}$ de E formée d'une base de F complétée par une base de $F^\perp$, la matrice de s_F dans la base $\mathcal{B}$ est diagonale, avec p fois le coefficient 1 et $n - p$ fois le coefficient -1 sur la diagonale.

Figure 1

(b) - Les propriétés suivantes des symétries orthogonales ont été vues dans le cours de géométrie de L2 (on pourra en reprendre les preuves en exercice):

(i) Les symétries orthogonales sont des isométries vectorielles [en d'autres termes, pour tout ss-e.v. F de E, on a $s_F \in \mathrm{O}(E)$].

(ii) Pour tout ss-e.v. F de E, on a: $\text{dét}\, s_F = (-1)^{n-\dim F}$. Une symétrie orthogonale peut donc suivant les cas être une isométrie directe ou une isométrie indirecte.

(iii) Pour toute isométrie vectorielle $f \in \mathrm{O}(E)$, on a:
[f est une symétrie orthogonale] $\Leftrightarrow$ [$f \circ f = \mathrm{id}_E$] $\Leftrightarrow$ [f est diagonalisable].

(iv) Pour tout ss-e.v. F de E et pour toute $g \in \mathrm{O}(E)$, on a: $g \circ s_F \circ g^{-1} = s_{g(F)}$.

(c) - La propriété (iv) ci-dessus, particulièrement important dans la pratique pour l'étude géométrique des symétries orthogonales, s'interprète algébriquement en utilisant les notions vues à la leçon 7 :

Proposition. *Le groupe* $\mathrm{O}(E)$ *opère par conjugaison sur l'ensemble des symétries orthogonales.*

(d) - On appelle *symétrie hyperplane* de E toute symétrie orthogonale par rapport à un hyperplan de E (ie. un ss-e.v. de dimension $n - 1$). C'est une isométrie indirecte d'après (b) (ii) ci-dessus.

Une propriété fondamentale (voir cours de L2 pour la preuve) est le fait que toute isométrie vectorielle est la composée d'un nombre fini (au plus égal à n) de symétries hyperplanes. Algébriquement, conformément à 1.2.8, cela signifie que:

PROPOSITION. *Le groupe $O(E)$ est engendré par l'ensemble des symétries hyperplanes de E.*

Preuve. Soit X le sous-ensemble de $\mathrm{O}(E)$ formé de toutes les symétries hyperplanes. Soit G un sous-groupe de $\mathrm{O}(E)$ contenant X. Comme G est stable par la loi $\circ$, il contient aussi tous les produits d'éléments de X. Or, d'après la propriété rappelée ci-dessus, tout élément de $\mathrm{O}(E)$ est un produit d'éléments de X. Donc $\mathrm{O}(E) \subseteq G$, c'est-à-dire $\mathrm{O}(E) = G$. Ceci prouve que $\mathrm{O}(E)$ est le plus petit sous-groupe de $\mathrm{O}(E)$ contenant X, et donc $\mathrm{O}(E) = \langle X \rangle$. □

8.1.4. ACTION CANONIQUE DU GROUPE ORTHOGONAL SUR L'ESPACE E. Considérons l'application $\mathrm{O}(E) \times E \to E$ définie par $(f, u) \mapsto f.u = f(u)$. Il est clair que, quel que soit $u \in E$, on a $g.(f.u) = (g \circ f)(u)$ pour tous $f, g \in \mathrm{O}(E)$, et $\mathrm{id}_E .u = u$. On a donc une action du groupe $\mathrm{O}(E)$ sur l'ensemble E, que l'on appelle action canonique. La proposition suivante décrit les orbites et les stabilisateurs pour cette action. On rappelle d'abord dans le lemme suivant quelques propriétés utiles des isométries vectorielles (les preuves, faciles, sont laissées en exercice).

LEMME. *Soit F un ss-e.v. de E. On note $F^\perp$ son supplémentaire orthogonal.*

(i) *Pour toute isométrie $f \in \mathrm{O}(E)$, on a $f(F)^\perp = f(F^\perp)$.*
(ii) *Si F est stable par une isométrie $f \in \mathrm{O}(E)$, alors $f_{|F}$ est une isométrie de F, et $F^\perp$ est également stable par f.*
(iii) *Pour toute isométrie r de F, il existe une unique isométrie f de E telle que $f(u) = r(u)$ pour tout $u \in F$ et $f(u) = u$ pour tout $u \in F^\perp$.*

PROPOSITION. *On considère l'action canonique de $\mathrm{O}(E)$ sur E.*

(i) *Pour tout $u \in E$, l'orbite de u est la sphère de rayon $\|u\|$.*
(ii) *Pour tout $u \in E$ non-nul, le stabilisateur de u est isomorphe au groupe orthogonal $O(H_u)$ de l'hyperplan H_u orthogonal à la droite $\mathbb{R}u$.*

Preuve. Soit $u \in E$, notons $\Omega_u = \{f(u)\,;\, f \in \mathrm{O}(E)\}$ l'orbite de u, $d = \|u\|$ et $S_d = \{v \in E\,;\, \|v\| = d\}$. Comme tout élément de $\mathrm{O}(E)$ conserve la norme, il est clair que $\Omega_u \subseteq S_d$. Réciproquement, soit $v \in S_d$. Si v est colinéaire à u, on a $v = \pm u$, donc $f = \pm \mathrm{id}_E$ vérifie $v = f(u)$, et donc $v \in \Omega_u$. Sinon, notons F le plan de base $\{u, v\}$. Comme $\|u\| = \|v\|$, on sait (voir rappel 8.1.5 ci-dessous) qu'il existe dans le plan F une rotation r telle que $r(u) = v$. D'après le point (iii) du lemme ci-dessus, r se prolonge en une isométrie $f \in \mathrm{O}(E)$. Celle-ci vérifie donc $v = f(u)$, donc $v \in \Omega_u$. On conclut que $\Omega_u = S_d$.

Fixons $u \neq 0_E$ et notons $G_u = \{f \in \mathrm{O}(E)\,;\, f(u) = u\}$ le stabilisateur de u. Soient $\mathbb{R}u$ la droite de base $\{u\}$ et $H_u = (\mathbb{R}u)^\perp$ son hyperplan orthogonal. Quels que soient $v \in H_u$ et $f \in G_u$, on a $f(v) \cdot u = f(v) \cdot f(u) = v \cdot u = 0$, donc $f(v) \perp u$. Ceci prouve que $f(v) \in H_u$ pour tout $v \in H_u$, c'est-à-dire que H_u est stable sous l'action de toute $f \in G_u$; la restriction $f_{|H_u}$ de f à H_u est alors un élément de $\mathrm{O}(H_u)$. L'application $\varphi : G_u \to \mathrm{O}(H_u)$ définie par $f \mapsto f_{|H_u}$ est clairement un morphisme de groupes. Si $f \in \mathrm{Ker}\,\varphi$, on a $f(v) = v$ pour tout $v \in H_u$; comme par ailleurs $f(v) = v$ pour tout $v \in \mathbb{R}u$ (puisque $f \in G_u$) et que $E = \mathbb{R}u \oplus H_u$, on déduit que $f = \mathrm{id}_E$. Ceci prouve que $\mathrm{Ker}\,\varphi = \{\mathrm{id}_E\}$ et donc φ est injective. La surjectivité de φ découle du point (iii) du lemme précédent. On conclut que φ réalise un isomorphisme entre les groupes G_u et $\mathrm{O}(H_u)$. □

8.1.5. LE GROUPE ORTHOGONAL EN DIMENSION 2. On suppose ici que l'espace euclidien E est de dimension 2 et orienté. On résume sans démonstration (voir cours de L2) les principaux éléments de la description géométrique concrète des isométries de E.

(a) - Les symétries hyperplanes dans E sont les symétries orthogonales par rapport aux droites. Pour toute droite Δ de E, on note s_Δ la symétrie orthogonale par rapport à Δ. On a $s_\Delta \in \mathrm{O}^-(E)$. On peut choisir une b.o.n. de E dans laquelle la matrice de s_Δ est $\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$. La matrice de s_Δ dans une base quelconque de E est de la forme $\left(\begin{smallmatrix} a & b \\ b & -a \end{smallmatrix}\right)$ avec $a, b \in \mathbb{R}$ tels que $a^2 + b^2 = 1$.

(b) - Pour tout $\theta \in \mathbb{R}$, l'endomorphisme r_θ de E dont la matrice dans toute b.o.n. directe de E est $R_\theta = \left(\begin{smallmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{smallmatrix}\right)$ s'appelle la rotation d'angle θ. On a $r_\theta \in \mathrm{O}^+(E)$.
On vérifie par le calcul que, pour tous $\theta, \theta' \in \mathbb{R}$, on a $R_\theta R_{\theta'} = R_{\theta+\theta'}$, et donc $r_\theta \circ r_{\theta'} = r_{\theta+\theta'}$.

(c) - Théorème.

(i) *Le groupe* $\mathrm{O}^+(E)$ *des isométries directes du plan* E *est formé des rotations.*

(ii) *L'ensemble* $\mathrm{O}^-(E)$ *des isométries indirectes du plan* E *est formé des symétries orthogonales par rapport aux droites.*

Figure 2

Figure 3

Parmi les conséquences algébriques de ce théorème, citons:

1. L'application $\theta \mapsto r_\theta$ est un morphisme de groupes de $\mathbb{R}$ (muni de l'addition) dans $\mathrm{O}^+(E)$ (muni de la loi $\circ$), surjectif et de noyau $2\pi\mathbb{Z}$.
2. $\mathrm{O}^+(E) \simeq \mathbb{R}/2\pi\mathbb{Z} \simeq \mathbb{U}$ (où $\mathbb{U}$ est le groupe multiplicatif des nombres complexes de module 1).

Rappelons enfin deux propriétés géométriques utiles dans de nombreuses situations:

3. Si u et v sont deux vecteurs non-nuls de même norme dans le plan E, il existe une unique rotation $r \in \mathrm{O}^+(E)$ telle que $v = r(u)$ et il existe une unique symétrie $s \in \mathrm{O}^-(E)$ telle que $v = r(u)$ (voir figure 4).
4. Si Δ et Δ' sont deux droites du plan E, alors $s_{\Delta'} \circ s_\Delta = r_{2\theta}$ où θ est l'angle de la rotation envoyant un vecteur unitaire de Δ sur un vecteur unitaire de Δ' (voir figure 5).

Figure 4

Figure 5

8.1.6. Le groupe orthogonal en dimension 3. On suppose ici que l'espace euclidien E est de dimension 3 et orienté. On résume sans démonstration les principaux éléments de la description géométrique concrète des isométries de E.

(a) - Soit Δ une droite de E orientée (par le choix d'un vecteur directeur unitaire e_3). Soit P le plan vectoriel $\Delta^\perp$ muni de l'orientation déduite de celle de $\Delta^\perp$ (ie. qu'une b.o.n. $\{e_1, e_2\}$ de P est directe si la b.o.n. $\{e_1, e_2, e_3\}$ de E est directe dans E).

Pour tout $\theta \in \mathbb{R}$, on appelle *rotation* d'axe Δ et d'angle θ l'endomorphisme de E tel que $f_{|\Delta}$ soit l'identité et $f_{|P}$ soit la rotation de P d'angle θ.

La matrice dans la base $\{e_1, e_2, e_3\}$ de cette rotation est:

$$\begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Une telle rotation est une isométrie directe de E.

Figure 6

Remarques:

1. L'application id_E est considérée comme une rotation, d'angle nul et d'axe quelconque.
2. Pour toute droite Δ, la rotation d'axe Δ et d'angle π est la symétrie orthogonale par rapport à Δ. On l'appelle le demi-tour d'axe Δ (voir figure 7).

Figure 7

(b) - THÉORÈME.

(i) *Le groupe $\mathrm{O}^+(E)$ des isométries directes de l'espace E de dimension 3 est formé des rotations.*

(ii) *L'ensemble $\mathrm{O}^-(E)$ des isométries indirectes de l'espace E de dimension 3 est formé des symétries orthogonales par rapport à une plan, et des composées (dans un sens ou dans l'autre) d'une rotation par une symétrie orthogonale par rapport au plan orthogonal à l'axe de la rotation.*

Figure 8

$f = s \circ r$

Figure 9

$f = r \circ s$

8.2 Groupes d'isométries affines.

Les développements précédents se situaient dans le cadre de la géométrie vectorielle (ie. de l'algèbre linéaire) et les vecteurs (les éléments de E) étaient notés simplement par une lettre minuscule $(u, v, \ldots)$. On va maintenant travailler dans des espaces affines et l'on convient donc, pour distinguer les points des vecteurs, de noter les vecteurs avec une flèche $(\overrightarrow{u}, \overrightarrow{v}, \overrightarrow{AB}, \ldots)$.

8.2.1 ESPACE AFFINE ET GROUPE AFFINE. Soit $\mathcal{E}$ un espace affine sur $\mathbb{R}$, d'e.v. directeur E.

(a) - Rappelons que cela signifie que :

(i) E est un $\mathbb{R}$-espace vectoriel ; ses éléments sont des vecteurs, on les notera $\overrightarrow{u}, \overrightarrow{v}, \overrightarrow{w}, \ldots$.

(ii) $\mathcal{E}$ est un ensemble non vide ; ses éléments sont des points, on les notera $A, B, C, M, N, \ldots$.

(iii) Il existe une application $\mathcal{E} \times \mathcal{E} \to E; (A, B) \mapsto \overrightarrow{AB}$ vérifiant les deux axiomes suivants:

(A1) pour tout $A \in \mathcal{E}$, pour tout $\overrightarrow{u} \in E$, il existe un unique $M \in \mathcal{E}$ tel que $\overrightarrow{AM} = \overrightarrow{u}$;

(A2) pour tous $A, B, C \in \mathcal{E}$, on a la relation de Chasles $\overrightarrow{AC} = \overrightarrow{AB} + \overrightarrow{BC}$.

Pour tous $A \in \mathcal{E}$ et $\overrightarrow{u} \in E$, l'unique point $M \in \mathcal{E}$ tel que $\overrightarrow{AM} = \overrightarrow{u}$ est noté $M = A + \overrightarrow{u}$.

PROPOSITION. *Soit $\mathcal{E}$ un espace affine de $\mathbb{R}$-e.v. directeur E. L'application*

$$\begin{array}{ccc} E \times \mathcal{E} & \longrightarrow & \mathcal{E} \\ (\overrightarrow{u}, A) & \longmapsto & A + \overrightarrow{u} \end{array}$$

définit une action du groupe additif de E sur l'ensemble $\mathcal{E}$; cette action est transitive et fidèle.

Preuve. Soient $\overrightarrow{u}$ et $\overrightarrow{v}$ deux éléments quelconques de E et A un point de $\mathcal{E}$. Notons $B = A + \overrightarrow{u}$ et $C = (A + \overrightarrow{u}) + \overrightarrow{v} = B + \overrightarrow{v}$. On a $\overrightarrow{u} = \overrightarrow{AB}$ et $\overrightarrow{v} = \overrightarrow{BC}$ donc par la relation de Chasles $\overrightarrow{AC} = \overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{u} + \overrightarrow{v}$, d'où $C = A + (\overrightarrow{u} + \overrightarrow{v})$. On conclut que $(A + \overrightarrow{u}) + \overrightarrow{v} = A + (\overrightarrow{u} + \overrightarrow{v})$. Comme par ailleurs $A + \overrightarrow{0} = A$ (car $\overrightarrow{AA} = \overrightarrow{0}$), les deux conditions requises pour définir une action (voir 7.1.1) sur $\mathcal{E}$ du groupe additif de E sont vérifiées.

Il est clair par ailleurs que le seul vecteur $\overrightarrow{u}$ tel que $A + \overrightarrow{u} = A$ pour tout $A \in \mathcal{E}$ est le vecteur nul, donc l'action est fidèle (voir 7.1.5.a). Enfin, quels que soient $A, B \in \mathcal{E}$, il existe un vecteur $\overrightarrow{u}$ (à savoir le vecteur $\overrightarrow{AB}$) tel que $B = A + \overrightarrow{u}$, ce qui prouve que l'action est transitive (voir 7.1.5.c). □

(b) - Par définition, une application $\varphi : \mathcal{E} \to \mathcal{E}$ est dite affine lorsqu'il existe un endomorphisme d'espace vectoriel $f : E \to E$ tel que:

$$\text{pour tous } A, B \in \mathcal{E},\ \overrightarrow{\varphi(A)\varphi(B)} = f(\overrightarrow{AB})$$

On dit que f est l'endomorphisme vectoriel associé à l'endomorphisme affine φ.

(c) - Le résultat suivant sur la détermination d'un endomorphisme affine à partir de l'endomorphisme vectoriel associé est fondamental:

THÉORÈME. *Pour tout endomorphisme f de E et tout couple de points (A, B) de $\mathcal{E}$, il existe une unique application affine $\varphi : \mathcal{E} \to \mathcal{E}$ dont f est l'application linéaire associée, et qui envoie A sur B, ie. telle que $\varphi(A) = B$.*

(d) - On montre que: φ est bijective de $\mathcal{E}$ sur $\mathcal{E}$ si et seulement si f est bijective de E sur E. On en déduit que les endomorphismes bijectifs de $\mathcal{E}$ forment un groupe pour la loi $\circ$ (un sous-groupe du groupe des bijections). On l'appelle le *groupe affine* de $\mathcal{E}$, noté $\mathrm{GA}(\mathcal{E})$.

$$\mathrm{GA}(\mathcal{E}) = \{\varphi : \mathcal{E} \to \mathcal{E}\,;\, \varphi \text{ affine et bijective}\}.$$

1. On montre aisément (voir cours L2) que tout élément de $\mathrm{GA}(\mathcal{E})$ transforme un sous-espace affine de $\mathcal{E}$ en un sous-espace affine de même dimension, conserve le parallélisme, conserve les barycentres...

2. De plus, si $\varphi \in \mathrm{GA}(\mathcal{E})$, l'ensemble des points fixes de φ est soit vide, soit un sous-espace affine de $\mathcal{E}$, dirigé par le ss-e.v. $\mathrm{Ker}(f - \mathrm{id}_E)$ des vecteurs de E fixés par l'endomorphisme vectoriel f associé à φ.

3. L'application $\vartheta : \mathrm{GA}(\mathcal{E}) \to \mathrm{GL}(E)$ associant à tout φ son application linéaire associée est un morphisme de groupes. Ce morphisme est surjectif d'après l'observation (c) précédente.

8.2.2 GROUPE DES ISOMÉTRIES AFFINES. On suppose maintenant de plus que l'espace affine $\mathcal{E}$ est euclidien, ce qui signifie que l'e.v. directeur E est euclidien. On définit à partir de la norme euclidienne sur E la distance dans $\mathcal{E}$ par:

$$\text{pour tous } A, B \in \mathcal{E},\ d(A, B) = \|\overrightarrow{AB}\| \text{ ; on note encore } d(A, B) = AB.$$

Une application affine $\mathcal{E} \to \mathcal{E}$ est appelée une *isométrie affine* de $\mathcal{E}$ lorsqu'elle conserve la distance:

$$\varphi \text{ isométrie} \quad \Leftrightarrow \quad d(\varphi(A), \varphi(B)) = d(A, B) \text{ pour tous } A, B \in \mathcal{E}.$$

Il est facile de vérifier que φ est une isométrie affine de $\mathcal{E}$ si et seulement si son application linéaire associée f est une isométrie vectorielle de E. Toute isométrie affine est donc une bijection de $\mathcal{E}$ sur $\mathcal{E}$, et les isométries affines forment un sous-groupe de $\mathrm{GA}(\mathcal{E})$. On note $\mathrm{Is}(\mathcal{E})$ ce groupe.

$$\mathrm{Is}(\mathcal{E}) = \{\varphi \in \mathrm{GA}(\mathcal{E})\,;\, \varphi \text{ conserve la distance}\} = \{\varphi \in \mathrm{GA}(\mathcal{E})\,;\, f \in \mathrm{O}(E)\},$$ sous-groupe de $\mathrm{GA}(\mathcal{E})$.

On appelle *isométrie directe* de $\mathcal{E}$, ou *déplacement* de $\mathcal{E}$, toute isométrie $\varphi \in \mathrm{Is}(\mathcal{E})$ dont l'application linéaire associée f est une isométrie vectorielle directe de E, ie. telle que $f \in \mathrm{O}^+(E)$. On note $\mathrm{Is}^+(\mathcal{E})$ l'ensemble des déplacements de $\mathcal{E}$; il est clair que c'est un sous-groupe de $\mathrm{Is}(\mathcal{E})$. On a:

$$[\mathrm{Is}(\mathcal{E}) : \mathrm{Is}^+(\mathcal{E})] = 2 \quad ; \quad \mathrm{Is}^+(\mathcal{E}) \lhd \mathrm{Is}(\mathcal{E}).$$

De même, une isométrie $\varphi \in \mathrm{Is}(\mathcal{E})$ telle que $f \in \mathrm{O}^-(E)$ s'appelle une *isométrie indirecte* ou un *antidéplacement* de $\mathcal{E}$. Bien sûr, l'ensemble $\mathrm{Is}^-(\mathcal{E})$ des antidéplacements de $\mathcal{E}$ n'est pas un sous-groupe de $\mathrm{Is}(\mathcal{E})$.

8.2.3 Exemples (translations, symétries orthogonales)

(a) - Pour tout $\overrightarrow{u} \in E$, on appelle *translation* de vecteur $\overrightarrow{u}$ l'application $\tau_{\overrightarrow{u}} : \mathcal{E} \to \mathcal{E}$ définie par:

pour tout $M \in \mathcal{E}$, $\tau_{\overrightarrow{u}}(M) = M'$ est l'unique point de $\mathcal{E}$ tel que $\overrightarrow{MM'} = \overrightarrow{u}$.

Il est clair que : $\tau_{\overrightarrow{0}} = \mathrm{id}_{\mathcal{E}}$, $\tau_{\overrightarrow{u}} \circ \tau_{\overrightarrow{v}} = \tau_{\overrightarrow{u}+\overrightarrow{v}} = \tau_{\overrightarrow{v}} \circ \tau_{\overrightarrow{u}}$, et que $\tau_{\overrightarrow{u}}$ est bijective avec $\tau_{\overrightarrow{u}}^{-1} = \tau_{-\overrightarrow{u}}$.

Figure 10

Toute translation est affine, d'application linéaire associée id_E, et est donc une isométrie directe de $\mathcal{E}$.
On conclut que:
l'ensemble $T(\mathcal{E})$ des translations de $\mathcal{E}$ est un sous-groupe abélien de $\mathrm{Is}^+(\mathcal{E})$, isomorphe au groupe additif de E.

En fait, non seulement l'application linéaire associée à une translation de $\mathcal{E}$ est id_E, mais réciproquement toute application affine $\mathcal{E} \to \mathcal{E}$ d'application linéaire associée id_E est nécessairement une translation. Cela signifie que $T(\mathcal{E})$ est le noyau du morphisme surjectif $\vartheta : GA(\mathcal{E}) \to \mathrm{GL}(E)$ consistant à associer à tout $\varphi \in \mathrm{GA}(\mathcal{E})$ son application linéaire associée $f \in \mathrm{GL}(E)$. Il est clair que $\varphi \in \mathrm{Is}(\mathcal{E})$ si et seulement si $\vartheta(\varphi) \in \mathrm{O}(E)$. En résumant et en utilisant 6.4.1, on a:

Proposition.

(i) $T(\mathcal{E}) = \{\varphi \in GA(\mathcal{E})\,;\, f = \mathrm{id}_E\} = \mathrm{Ker}\,\vartheta$,

(ii) $T(\mathcal{E}) \lhd \mathrm{GA}(\mathcal{E})$ et $\mathrm{GA}(\mathcal{E})/T(\mathcal{E}) \simeq \mathrm{GL}(E)$; $T(\mathcal{E}) \lhd \mathrm{Is}(\mathcal{E})$ et $\mathrm{Is}(\mathcal{E})/T(\mathcal{E}) \simeq \mathrm{O}(E)$.

(b) - Soit $\mathcal{F}$ un sous-espace affine de $\mathcal{E}$, soit F le ss-e.v. de E directeur de $\mathcal{F}$, et soit $F^\perp$ le supplémentaire orthogonal de F dans E.
- Pour tout $M \in \mathcal{E}$, il existe un unique point $M' \in \mathcal{F}$ tel que $\overrightarrow{MM'} \in F^\perp$; ce point M', qui est le point d'intersection de $\mathcal{F}$ avec le sous-espace affine passant par M et dirigé par $F^\perp$, s'appelle le projeté orthogonal de M sur $\mathcal{F}$.
- Il existe alors un unique point $M'' \in \mathcal{E}$ tel que M' soit le milieu de $[M, M'']$. Ce point M' s'appelle le symétrique orthogonal de M par rapport à $\mathcal{F}$.
- L'application qui, à tout point de $\mathcal{E}$, associe son symétrique orthogonal par rapport à $\mathcal{F}$ s'appelle la *symétrie orthogonale affine par rapport à $\mathcal{F}$*. On la note $\sigma_{\mathcal{F}}$.
- Toute symétrie orthogonale affine $\sigma_{\mathcal{F}}$ est une application affine, dont l'application linéaire associée est la symétrie orthogonale vectorielle s_F par rapport au ss-e.v. F de E qui est l'espace directeur de $\mathcal{F}$ (voir 8.1.3). On en déduit:

les symétries orthogonales affines sont des isométries de $\mathcal{E}$.

Le fait que $\sigma_{\mathcal{F}}$ soit directe ou indirecte dépend des valeurs de n et $\dim \mathcal{F}$ [voir 8.1.3.b.(ii)].

Dans le cas où $\mathcal{F}$ est un hyperplan de $\mathcal{E}$, on dit que $\sigma_{\mathcal{F}}$ est une symétrie hyperplane affine ; les symétries hyperplanes affines sont des antidéplacements [voir 8.1.3.d].

Dans le cas où $\mathcal{F}$ est un singleton $\{A\}$, on dit que σ_A est la symétrie centrale de centre A ; à noter que σ_A est simplement l'application qui, à tout point $M \in \mathcal{E}$, associe le point $M'' \in \mathcal{E}$ tel que A est le milieu de $[MM'']$.

Figure 11, exemple en dimension 3	Figure 12, exemple en dimension 3	Figure 13, exemple en dimension 3
symétrie p/r à un plan	symétrie p/r à une droite	symétrie p/r à un point

8.2.4 Le groupe des isométries affines en dimension 2. On suppose ici que l'espace affine euclidien $\mathcal{E}$ est de dimension 2 et orienté. En particulier, les symétries orthogonales par rapport aux droites sont des antidéplacements. On résume sans démonstration (voir cours de L2) les principaux éléments de la description géométrique concrète des isométries de $\mathcal{E}$.

(a) - Pour tout $A \in \mathcal{E}$ et $\theta \in \mathbb{R}$, on *appelle rotation affine* de centre A et d'angle θ l'application affine $\rho_{A,\theta} : \mathcal{E} \to \mathcal{E}$ telle que son application linéaire associée soit la rotation vectorielle r_θ de E et telle que $\rho_{A,\theta}(A) = A$.

Figure 14

Elle existe et est unique d'après l'observation 8.2.1.c.

$$M' = \rho_{A,\theta}(M) \Leftrightarrow \overrightarrow{AM'} = r_\theta(\overrightarrow{AM}).$$

Il est clair que :

$$\rho_{A,\theta} \in \mathrm{Is}^+(\mathcal{E}) \text{ pour tous } A \in \mathcal{E},\ \theta \in \mathbb{R}.$$

Si $\theta \equiv 0$ modulo 2π, alors $\rho_{A,\theta} = \mathrm{id}_\mathcal{E}$,
sinon, A est l'unique point fixe de $\rho_{A,\theta}$

Théorème. *Le groupe* $\mathrm{Is}^+(\mathcal{E})$ *des déplacements du plan affine* $\mathcal{E}$ *euclidien est formé par les translations et les rotations.*

(b) - Pour toute droite affine $\mathcal{D}$ de $\mathcal{E}$ et tout vecteur $\overrightarrow{u}$ de E appartenant à la droite vectorielle Δ directrice de $\mathcal{D}$, on appelle symétrie glissée de vecteur $\overrightarrow{u}$ par rapport à $\mathcal{D}$ la composée φ de la translation $\tau_{\overrightarrow{u}}$ de vecteur $\overrightarrow{u}$ et de la symétrie orthogonale $\sigma_\mathcal{D}$ par rapport à $\mathcal{D}$. On a:

Figure 15

$$\varphi = \tau_{\overrightarrow{u}} \circ \sigma_\mathcal{D} = \sigma_\mathcal{D} \circ \tau_{\overrightarrow{u}}.$$

Il est clair que φ est affine d'application linéaire associée $f = \mathrm{id}_E \circ s_\Delta$, donc $f = s_\Delta \in \mathrm{O}^-(E)$, d'où $\varphi \in \mathrm{Is}^-(\mathcal{E})$.

Si $\overrightarrow{u} = \overrightarrow{0}$, alors φ est la symétrie $\sigma_\mathcal{D}$, donc l'ensemble des points fixes de φ est $\mathcal{D}$. Sinon, φ n'a pas de point fixe.

Théorème. *L'ensemble* $\mathrm{Is}^-(\mathcal{E})$ *des antidéplacements du plan affine euclidien* $\mathcal{E}$ *est formé par les symétries glissées.*

(c) - On termine en rappelant le résultat de composition suivant.

Proposition. *Soient* $\mathcal{D}$ *et* $\mathcal{D}'$ *deux droites du plan affine euclidien* $\mathcal{E}$.

(i) *Si* $\mathcal{D}$ *et* $\mathcal{D}'$ *sont parallèles, alors* $\sigma_{\mathcal{D}'} \circ \sigma_\mathcal{D} = \tau_{2\overrightarrow{u}}$, *où* $\overrightarrow{u}$ *est le vecteur orthogonal à la direction commune de* $\mathcal{D}$ *et* $\mathcal{D}'$ *tel que* $\tau_{\overrightarrow{u}}(\mathcal{D}) = \mathcal{D}'$.

(ii) *Si* $\mathcal{D}$ *et* $\mathcal{D}'$ *ne sont pas parallèles, alors* $\sigma_{\mathcal{D}'} \circ \sigma_\mathcal{D} = \rho_{A,2\theta}$, *où* A *est le point d'intersection de* $\mathcal{D}$ *et* $\mathcal{D}'$, *et* $\theta = \widehat{(\mathcal{D}, \mathcal{D}')}$ *modulo* π.

(iii) *En particulier, si* $\mathcal{D}$ *et* $\mathcal{D}'$ *sont perpendiculaires en* A, *alors* $\sigma_{\mathcal{D}'} \circ \sigma_\mathcal{D}$ *est la symétrie centrale de centre* A.

Figure 16

Figure 17

Figure 18

COROLLAIRE. *Tout déplacement du plan affine euclidien $\mathcal{E}$ est produit de deux symétries orthogonales par rapport à des droites. Tout antidéplacement du plan affine euclidien $\mathcal{E}$ est produit de une ou trois symétries orthogonales par rapport à des droites.*

Il en résulte que: *le groupe* $\mathrm{Is}\,\mathcal{E}$ *est engendré par les symétries orthogonales par rapport aux droites.*

8.2.5 LE GROUPE DES ISOMÉTRIES AFFINES EN DIMENSION 3. On suppose ici que l'espace affine euclidien $\mathcal{E}$ est de dimension 3 et orienté. En particulier, les symétries orthogonales par rapport aux plans sont des antidéplacements. On résume sans démonstration les principaux éléments de la description géométrique concrète des isométries de $\mathcal{E}$.

(a) - Pour toute droite affine orientée $\mathcal{D}$ de $\mathcal{E}$ et $\theta \in \mathbb{R}$, on appelle *rotation affine* d'axe $\mathcal{D}$ et d'angle θ l'application affine $\rho_{\mathcal{D},\theta} : \mathcal{E} \to \mathcal{E}$ telle que son application linéaire associée soit la rotation vectorielle $r_{\Delta,\theta}$ de E d'axe la droite vectorielle orientée Δ de E directrice de $\mathcal{D}$ et d'angle θ, et telle que $\rho_{\mathcal{D},\theta}(M) = M$ pour tout point M de la droite $\mathcal{D}$.

Figure 19

Elle existe et est unique d'après l'observation 8.2.1.c. Il est clair que :

$\rho_{\mathcal{D},\theta} \in \mathrm{Is}^+(\mathcal{E})$ pour tous $\theta \in \mathbb{R}$, $\mathcal{D}$ droite orientée de $\mathcal{E}$.

Si $\theta \equiv 0$ modulo 2π, alors $\rho_{\mathcal{D},\theta} = \mathrm{id}_{\mathcal{E}}$,
sinon, les points fixes de $\rho_{A,\theta}$ sont les points de $\mathcal{D}$.

(b) - Pour toute droite affine orientée $\mathcal{D}$ de $\mathcal{E}$ et $\theta \in \mathbb{R}$, on appelle *vissage* de $\mathcal{E}$ tout produit d'une rotation $\rho = \rho_{\mathcal{D},\theta}$ au sens précédent par une translation $\tau = \tau_{\overrightarrow{u}}$ telle que le vecteur $\overrightarrow{u}$ de translation appartienne à la direction Δ de $\mathcal{D}$. Parce que $\overrightarrow{u} \in \Delta$, on a $\tau \circ \rho = \rho \circ \tau$, et τ et ρ sont uniques. Il est clair qu'un vissage est un déplacement de $\mathcal{E}$.

Figure 20

Une rotation est un vissage dont le vecteur est nul. Une translation est un vissage dont l'angle est nul (modulo 2π). Un vissage qui n'est pas une rotation n'admet aucun point fixe.

THÉORÈME. *Le groupe* $\mathrm{Is}^+(\mathcal{E})$ *des déplacements de l'espace affine euclidien $\mathcal{E}$ de dimension 3 est formé par les vissages.*

THÉORÈME. *L'ensemble* $\mathrm{Is}^-(\mathcal{E})$ *des antidéplacements de l'espace affine euclidien $\mathcal{E}$ de dimension 3 est formé par :*

1. *les symétries orthogonales par rapport à un plan,*
2. *les composées d'une symétrie orthogonale par rapport à un plan avec une rotation dont l'axe est une droite perpendiculaire au plan de la symétrie,*
3. *les composées d'une symétrie orthogonale par rapport à un plan avec une translation dont le vecteur appartient au plan vectoriel directeur du plan de la symétrie (symétrie glissée).*

COROLLAIRE. *Toute isométrie de l'espace affine euclidien $\mathcal{E}$ de dimension 3 est un produit d'un nombre fini de symétries orthogonales par rapport à des plans.*

En d'autres termes: *le groupe* $\mathrm{Is}\,\mathcal{E}$ *est engendré par les symétries orthogonales par rapport aux plans.*

Leçon 9

Sous-groupes d'isométries laissant invariante une partie du plan ou de l'espace

9.1 Quelques principes généraux.

9.1.1. On fixe un $\mathbb{R}$-espace affine euclidien $\mathcal{E}$ de dimension finie n et on note E son e.v. directeur. Dans les applications qui suivent, on aura toujours $n = 2$ ou $n = 3$. Le problème central étudié dans cette leçon est de déterminer, pour un sous-ensemble $\mathcal{X}$ de points de $\mathcal{E}$, l'ensemble $G_{\mathcal{X}}$ des isométries affines de $\mathcal{E}$ qui laissent globalement invariant $\mathcal{X}$; il est clair que c'est un sous-groupe de $\mathrm{Is}(\mathcal{E})$, et l'intersection $\mathrm{Is}^+(\mathcal{E}) \cap G_{\mathcal{X}}$ est alors un sous-groupe de $\mathrm{Is}^+(\mathcal{E})$, que l'on notera ${G_{\mathcal{X}}}^+$.

$$G_{\mathcal{X}} = \{\, \varphi \in \mathrm{Is}(\mathcal{E}) \;;\; \varphi(\mathcal{X}) = \mathcal{X} \,\} \text{ sous-groupe de } \mathrm{Is}(\mathcal{E}) \;;$$
$${G_{\mathcal{X}}}^+ = \{\, \varphi \in \mathrm{Is}^+(\mathcal{E}) \;;\; \varphi(\mathcal{X}) = \mathcal{X} \,\} \text{ sous-groupe de } \mathrm{Is}^+(\mathcal{E}).$$

On note aussi ${G_{\mathcal{X}}}^- = G_{\mathcal{X}} \cap \mathrm{Is}^-(\mathcal{E})$. On a donc la réunion disjointe $G_{\mathcal{X}} = {G_{\mathcal{X}}}^+ \cup {G_{\mathcal{X}}}^-$.

> Exemple 1. Prenons $n = 2$ et $\mathcal{X} = \{A\}$ un singleton formé d'un unique point du plan. En utilisant les théorèmes de 8.2.4, et en rappelant qu'une translation de vecteur non-nul n'admet pas de point fixe, il est clair que ${G_{\mathcal{X}}}^+$ est le groupe de toutes les rotations de centre A, et que ${G_{\mathcal{X}}}^-$ est l'ensemble des symétries orthogonales par rapport aux droites passant par A.

9.1.2. Bien sûr, ${G_{\mathcal{X}}}^+$ est un sous-groupe de $G_{\mathcal{X}}$ alors que ${G_{\mathcal{X}}}^-$ n'est qu'un sous-ensemble (en particulier, ${G_{\mathcal{X}}}^-$ peut être vide). Le lemme suivant, très utile dans la pratique, précise ce point.

Lemme. *Supposons que ${G_{\mathcal{X}}}^-$ n'est pas vide. Alors:*

(i) *Pour tout $\sigma \in {G_{\mathcal{X}}}^-$, on a $ {G_{\mathcal{X}}}^- = \sigma\, {G_{\mathcal{X}}}^+ = \{\sigma \circ \varphi \,;\, \varphi \in {G_{\mathcal{X}}}^+\}$.*

(ii) *Le sous-groupe ${G_{\mathcal{X}}}^+$ est d'indice 2 dans $G_{\mathcal{X}}$.*

(iii) *Si ${G_{\mathcal{X}}}^-$ est fini, alors ${G_{\mathcal{X}}}^+$ est fini, et $|{G_{\mathcal{X}}}^+| = |{G_{\mathcal{X}}}^-| = \frac{1}{2}\,|G_{\mathcal{X}}|$.*

(iv) *Si ${G_{\mathcal{X}}}^+$ est réduit à $\{\mathrm{id}_{\mathcal{E}}\}$,* alors $G_{\mathcal{X}} = \{\mathrm{id}_{\mathcal{E}}, \sigma\}$ avec $\sigma \in {G_{\mathcal{X}}}^-$ d'ordre 2.

> *Preuve.* Fixons $\sigma \in {G_{\mathcal{X}}}^-$. Il est clair que $\sigma \circ \varphi \in {G_{\mathcal{X}}}^-$ pour tout $\varphi \in {G_{\mathcal{X}}}^+$. Réciproquement, toute $\psi \in {G_{\mathcal{X}}}^-$ s'écrit $\psi = \sigma \circ (\sigma^{-1} \circ \psi)$, avec $(\sigma^{-1} \circ \psi)$ qui appartient à ${G_{\mathcal{X}}}^+$ en tant que produit de deux éléments de ${G_{\mathcal{X}}}^-$. En résumé, l'application $\varphi \mapsto \sigma \circ \varphi$ définit une bijection de ${G_{\mathcal{X}}}^+$ sur ${G_{\mathcal{X}}}^-$, de bijection réciproque $\psi \mapsto \sigma^{-1} \circ \psi$. Les différents points du lemme en découlent, en observant que ${G_{\mathcal{X}}}^-$ n'est autre que la classe à gauche de σ (et de tous les éléments de ${G_{\mathcal{X}}}^-$) modulo le sous-groupe ${G_{\mathcal{X}}}^+$.□

Concrètement, la détermination de $G_{\mathcal{X}}$ repose donc sur la détermination de ${G_{\mathcal{X}}}^+$ et d'un élément de ${G_{\mathcal{X}}}^-$ s'il en existe.

> Exemple 2. Prenons $n = 2$ et $\mathcal{X} = \mathcal{D}$ une droite du plan $\mathcal{E}$. On sait que $\mathrm{Is}^+(\mathcal{E})$ est formé des translations et des rotations. Une translation laisse $\mathcal{D}$ globalement invariante lorsque son vecteur appartient à la direction Δ de $\mathcal{D}$. Une rotation laisse $\mathcal{D}$ globalement invariante lorsque son centre est sur $\mathcal{D}$ et que son angle est 0 modulo π (ie. lorsque c'est une symétrie centrale de centre sur $\mathcal{D}$). On conclut que :
>
> $G_{\mathcal{D}}^+$ est formé des symétries centrales dont le centre appartient à $\mathcal{D}$ et des translations dont le vecteur appartient à Δ.
>
> La symétrie $\sigma_{\mathcal{D}}$ par rapport à $\mathcal{D}$ appartient à $G_{\mathcal{D}}^-$. D'après le point (i) du lemme précédent, tout élément $\psi \in G_{\mathcal{D}}^-$ est de la forme $\psi = \sigma_{\mathcal{D}} \circ \varphi$ avec $\varphi \in G_{\mathcal{D}}^+$. Si φ est une symétrie centrale de centre A appartenant à $\mathcal{D}$, alors ψ est la symétrie orthogonale par rapport à la droite $\mathcal{D}'$ perpendiculaire en A à $\mathcal{D}$. Si φ est une translation de vecteur $\overrightarrow{u}$ appartenant à Δ, alors ψ est la symétrie glissée d'axe $\mathcal{D}$ et de vecteur $\overrightarrow{u}$. On conclut que :
>
> $G_{\mathcal{D}}^-$ est formé des symétries orthogonales par rapport aux droites perpendiculaires à $\mathcal{D}$ et des symétries glissées d'axe $\mathcal{D}$ et de vecteur appartenant à Δ.
>
> On pourra à titre d'exercice expliciter les composées deux à deux de chacun des 4 types d'isométries intervenant dans le groupe $G_{\mathcal{D}}$.

9.1.3. Un autre argument général d'usage très fréquent est le suivant, qui s'applique au cas où l'ensemble $\mathcal{X}$ est fini.

PROPOSITION. *Soit $\mathcal{X} = \{A_1, A_2, \ldots, A_m\}$ un ensemble fini non vide de m points de $\mathcal{E}$. Toute isométrie $\varphi \in G_{\mathcal{X}}$ fixe l'isobarycentre des points de $\mathcal{X}$.*

Preuve. Notons G l'isobarycentre des points de $\mathcal{X}$ et f l'application linéaire associée à φ. On a:

$$\sum_{i=1}^{m} \overrightarrow{\varphi(G)\varphi(A_i)} = \sum_{i=1}^{m} f(\overrightarrow{GA_i}) = f(\sum_{i=1}^{m} \overrightarrow{GA_i}) = f(\overrightarrow{0}) = \overrightarrow{0}.$$

Comme φ laisse $\mathcal{X}$ invariant, elle permute entre eux les m points A_i, de sorte que :

$$\sum_{i=1}^{m} \overrightarrow{\varphi(G)\varphi(A_i)} = \sum_{i=1}^{m} \overrightarrow{\varphi(G)A_i} = \overrightarrow{0}.$$

Par unicité du barycentre, cette dernière égalité prouve que $\varphi(G) = G$. □

Le cas le plus simple d'application de cette propriété est le suivant.

EXEMPLE 3. Prenons $n = 2$ et $\mathcal{X} = \{A, B\}$ une paire de points distincts de $\mathcal{E}$. Soit $\varphi \in {G_{\mathcal{X}}}^+$. D'après la proposition ci-dessus, φ fixe le milieu I de $[A, B]$. En utilisant le théorème 8.2.4.(a), φ est forcément soit l'identité, soit une rotation de centre I ; dans ce dernier cas, elle doit échanger A et B donc être d'angle π. On a donc ${G_{\mathcal{X}}}^+ = \{\mathrm{id}_{\mathcal{E}}, \rho\}$ où ρ est la symétrie centrale de centre I.
Par ailleurs, en notant $\mathcal{D}$ la droite (AB), il est clair que $\sigma_{\mathcal{D}} \in {G_{\mathcal{X}}}^-$. En utilisant le point (i) du lemme 9.1.2, et on observant que $\sigma_{\mathcal{D}} \circ \rho = \sigma_{\mathcal{D}'}$ où $\mathcal{D}'$ est la droite orthogonale à $\mathcal{D}$ passant par I (la médiatrice de $[A, B]$), on obtient ${G_{\mathcal{X}}}^- = \{\sigma_{\mathcal{D}}, \sigma_{\mathcal{D}'}\}$. D'où finalement:

$$G_{\mathcal{X}} = \{\mathrm{id}_{\mathcal{E}}, \rho, \sigma_{\mathcal{D}}, \sigma_{\mathcal{D}'}\}.$$

Comme $\rho^2 = \sigma_{\mathcal{D}}^2 = \sigma_{\mathcal{D}'}^2 = \mathrm{id}_E$, on conclut avec 3.5.4 que:

$G_{\mathcal{X}}$ est le groupe de Klein.

Figure 1

9.1.4. Par définition même de $G_{\mathcal{X}}$, il est clair que l'on peut considérer l'action canonique de $G_{\mathcal{X}}$ ou de ${G_{\mathcal{X}}}^+$ sur $\mathcal{X}$, définie par $\varphi.A = \varphi(A)$ pour toute $\varphi \in G_{\mathcal{X}}$ et tout $A \in \mathcal{X}$. Cela permet d'utiliser certains résultats sur les actions de groupes (dont l'important théorème 7.3.2) qui, combinés aux arguments 9.1.2 et 9.1.3, peuvent permettre la détermination explicite de $G_{\mathcal{X}}$. C'est le cas pour l'exemple fondamental ci-dessous.

9.2 Groupes diédraux.

9.2.1 DONNÉES. Dans ce qui suit, on se place dans le plan affine euclidien orienté $\mathcal{E}$. On fixe un entier $n \geq 3$ et on considère l'ensemble $\mathcal{X}$ des sommets d'un polygone régulier à n côtés. On note $\mathcal{X} = \{A_0, A_1, \ldots, A_{n-1}\}$ et O l'isobarycentre des points de $\mathcal{X}$ qui est aussi le centre du cercle circonscrit à $\mathcal{X}$. On a pour tout $1 \leq k \leq n-1$:

$$A_{k-1}A_k = A_0A_1 = A_{n-1}A_0 \quad \text{et} \quad \widehat{A_{k-1}OA_k} = \widehat{A_0OA_1} = \widehat{A_{n-1}OA_0} = \frac{2\pi}{n}.$$

Figure 2

cas $n = 5$

Figure 3

cas $n = 6$

Les sommets de $\mathcal{X}$ sont les affixes des nombres complexes $\mathrm{e}^{2ik\pi/n}$ pour $1 \leq k \leq n$.

9.2.2 Lemme. *$G_{\mathcal{X}}{}^+$ est le groupe cyclique $\{\mathrm{id}_{\mathcal{E}}, \rho, \rho^2, \ldots, \rho^{n-1}\}$ d'ordre n engendré par la rotation ρ de centre O et d'angle $2\pi/n$.*

Preuve. Il est clair $\rho \in G_{\mathcal{X}}{}^+$ et que ρ est d'ordre n. Donc le sous-groupe cyclique $\langle\rho\rangle$ d'ordre n engendré par ρ est un sous-groupe de $G_{\mathcal{X}}{}^+$.

Considérons l'action canonique (voir 9.1.4) de $G_{\mathcal{X}}{}^+$ sur $\mathcal{X}$, définie par $\varphi.A = \varphi(A)$ pour tout $\varphi \in G_{\mathcal{X}}{}^+$ et tout $A \in \mathcal{E}$. Prenons un point de $\mathcal{X}$, par exemple A_0, et notons Ω son orbite pour cette action. Comme $\Omega \subseteq \mathcal{X}$, on a $|\Omega| \leq n$ (en fait il est clair que l'on a l'égalité). Notons H le stabilisateur de A_0 pour cette action, ie. le sous-groupe des éléments de $G_{\mathcal{X}}{}^+$ fixant A_0. Soit $\varphi \in H$. D'après la proposition 9.1.3, φ fixe O. Comme φ ne peut être qu'une rotation ou une translation (voir 8.2.4), le fait qu'elle admette deux points fixes distincts A_0 et O implique que $\varphi = \mathrm{id}_{\mathcal{E}}$. Ainsi $H = \{\mathrm{id}_{\mathcal{E}}\}$.

En appliquant le théorème 7.3.2, on a $|\Omega| = |G_{\mathcal{X}}{}^+|/|H|$, d'où ici $|G_{\mathcal{X}}{}^+| = |\Omega| \leq n$. Comme $G_{\mathcal{X}}{}^+$ contient $\langle\rho\rangle$ d'ordre n, on conclut que $G_{\mathcal{X}}{}^+ = \langle\rho\rangle$. □

9.2.3 Lemme. *$G_{\mathcal{X}}{}^-$ est l'ensemble $\{\sigma, \sigma\circ\rho, \sigma\circ\rho^2, \sigma\circ\rho^3, \ldots, \sigma\circ\rho^{n-1}\}$, où σ est la symétrie orthogonale par rapport à la droite (OA_0). Il est formé de n symétries par rapport à des droites.*

(i) *Si n est impair, ces symétries sont les n symétries dont les axes sont les droites passant le centre O et chaque sommet A_i.*

(ii) *Si $n = 2p$ est pair, ces symétries sont:*
- d'une part les p symétries dont les axes sont les droites (A_iA_{i+p}) pour $0 \leq i \leq p-1$ joignant les sommets opposés,
- d'autre part les p symétries dont les axes sont les droites joignant les milieux des côtés opposés $[A_iA_{i+1}]$ et $[A_{i+p}A_{i+p+1}]$ pour $0 \leq i \leq p-1$.

Preuve. Soit $\sigma \in \mathrm{Is}^-(\mathcal{E})$ la symétrie orthogonale par rapport à la droite (OA_0). On désigne toujours par ρ la rotation de centre O et d'angle $2\pi/n$. Pour tout entier k, la composée $\sigma\circ\rho^k$ appartient à $\mathrm{Is}^-(\mathcal{E})$. Comme elle admet un point fixe O, il résulte du théorème 8.2.4.b que c'est une symétrie par rapport à une droite. Donc $(\sigma\circ\rho^k)\circ(\sigma\circ\rho^k) = \mathrm{id}_{\mathcal{E}}$, ou encore $\sigma\circ\rho^k = \rho^{n-k}\circ\sigma$.

Soit A_k un point quelconque de $\mathcal{X}$, avec $0 \leq k \leq n-1$. On a $\sigma(A_k) = \sigma(\rho^k(A_0)) = \rho^{n-k}(\sigma(A_0))$ d'après la relation ci-dessus. Mais $\sigma(A_0) = A_0$ par définition de σ, et $\rho^{n-k}(A_0) = A_{n-k} \in \mathcal{X}$ puisque $\rho \in G_{\mathcal{X}}{}^+$. On en déduit que $\sigma(A_k) = A_{n-k} \in \mathcal{X}$. Ceci prouve que $\mathcal{X}$ est stable par σ. On conclut que $\sigma \in G_{\mathcal{X}}{}^-$.

On applique le lemme 9.1.2 pour déduire que $G_{\mathcal{X}}{}^- = \{\sigma, \sigma\circ\rho, \sigma\circ\rho^2, \sigma\circ\rho^3, \ldots, \sigma\circ\rho^{n-1}\}$. Ces n isométries sont indirectes et fixent O, donc ce sont des symétries par rapport à des droites passant par O (voir 8.2.4.b). Le résultat s'en déduit aisément en distinguant suivant la parité de n, et ten utilisant le fait que l'axe d'une telle symétrie σ est la médiatrice du segment $[A_k\sigma(A_k)]$ dès lors que A_k n'est pas fixé par σ. □

Remarquons que le groupe $G_{\mathcal{X}}$ de l'exemple 3 de 9.1.3 correspond au cas $n = 2$ du groupe étudié ici aux lemmes 9.2.2 et 9.2.3. On conviendra donc dans ce qui suit d'englober le cas $n = 2$, en considérant une paire de points $\{A_1, A_2\}$ comme un polygône régulier à 2 sommets (et 2 côtés).

9.2.4 Théorème et définition. *Pour tout entier $n \geq 2$, on appelle groupe diédral d'ordre $2n$, noté D_n, le sous-groupe des isométries affines conservant un polygone régulier à n côtés.*

(i) *Le groupe D_n est engendré par deux éléments ρ et σ, et formé de $2n$ éléments distincts:*

$$D_n = \{e, \rho, \rho^2, \rho^3, \ldots, \rho^{n-1}, \sigma, \sigma\circ\rho, \sigma\circ\rho^2, \sigma\circ\rho^3, \ldots, \sigma\circ\rho^{n-1}\},$$

vérifiant les relations: $\rho^n = e$, $\sigma^2 = e$, $\sigma\circ\rho^k = \rho^{n-k}\circ\sigma$ pour tout $1 \leq k \leq n$.

Le sous-groupe $\langle\rho\rangle$ engendré par ρ est cyclique d'ordre n, d'indice 2 et normal dans D_n.

Le groupe D_n est non abélien pour $n \geq 3$; le groupe D_2 est isomorphe au groupe de Klein.

(ii) *Réciproquement, tout groupe engendré par deux éléments ρ et σ tels que ρ soit d'ordre n, σ soit d'ordre 2, et $\sigma\rho = \rho^{n-1}\sigma$, est isomorphe au groupe diédral D_n.*

Preuve. Le point (i) résulte immédiatement des résultats de 9.2.2 et 9.2.3. Le point (ii) est alors clair, en observant que $\rho^n = e$ et $\sigma\rho = \rho^{n-1}\sigma$ impliquent $\sigma\rho^k = \rho^{n-k}\sigma$ pour tout $1 \leq k \leq n$. □

9.2.5 Exemple : groupe du triangle. D_3 est le groupe des isométries du plan affine euclidien conservant un triangle équilatéral (ABC). Il est formé de l'identité $\mathrm{id}_{\mathcal{E}} = e$, de la rotation ρ de centre l'isobarycentre O de (ABC) et d'angle $2\pi/3$, de la rotation ρ^2 de centre O et d'angle $4\pi/3$, et des symétries orthogonales $\sigma_1, \sigma_2, \sigma_3$ par rapport aux trois médianes (ou hauteurs) du triangle. Le groupe D_3 est d'ordre 6, non abélien, engendré par les deux éléments ρ et σ_1 (on a $\sigma_3 = \rho \circ \sigma_1$ et $\sigma_2 = \rho^2 \circ \sigma_1$), et sa table est donnée ci-dessous.

Figure 4

D_3	e	ρ	ρ^2	σ_1	σ_2	σ_3
e	e	ρ	ρ^2	σ_1	σ_2	σ_3
ρ	ρ	ρ^2	e	σ_3	σ_1	σ_2
ρ^2	ρ^2	e	ρ	σ_2	σ_3	σ_1
σ_1	σ_1	σ_2	σ_3	e	ρ	ρ^2
σ_2	σ_2	σ_3	σ_1	ρ^2	e	ρ
σ_3	σ_3	σ_1	σ_2	ρ	ρ^2	e

Cette table est identique à celle du groupe symétrique S_3 donnée en 1.3.5.(d), donc: $D_3 \simeq S_3$.

9.2.6 Exemple: groupe du carré. D_4 est le groupe des isométries du plan affine euclidien conservant un carré $(ABCD)$. Il est formé de l'identité $e = \mathrm{id}_{\mathcal{E}}$, de la rotation ρ de centre le centre O du carré $(ABCD)$ et d'angle $\pi/2$, de la symétrie centrale ρ^2 de centre O, de la rotation ρ^3 de centre O et d'angle $3\pi/2$, des symétries orthogonales σ_1, σ_2 par rapport aux deux médianes du carré, et des symétries orthogonales τ_1, τ_2 par rapport aux deux diagonales du carré.
Le groupe D_4 est d'ordre 8, non abélien, engendré par les deux éléments ρ et σ_1 (on a $\tau_1 = \rho \circ \sigma_1$, $\sigma_2 = \rho^2 \circ \sigma_1$ et $\tau_2 = \rho^3 \circ \sigma_1$), et dont la table est donnée ci-dessous.

Figure 5

D_4	e	ρ	ρ^2	ρ^3	σ_1	σ_2	τ_1	τ_2
e	e	ρ	ρ^2	ρ^3	σ_1	σ_2	τ_1	τ_2
ρ	ρ	ρ^2	ρ^3	e	τ_1	τ_2	σ_2	σ_1
ρ^2	ρ^2	ρ^3	e	ρ	σ_2	σ_1	τ_2	τ_1
ρ^3	ρ^3	e	ρ	ρ^2	τ_2	τ_1	σ_1	σ_2
σ_1	σ_1	τ_2	σ_2	τ_1	e	ρ^2	ρ^3	ρ
σ_2	σ_2	τ_1	σ_1	τ_2	ρ^2	e	ρ	ρ^3
τ_1	τ_1	σ_1	τ_2	σ_2	ρ	ρ^3	e	ρ^2
τ_2	τ_2	σ_2	τ_1	σ_1	ρ^3	ρ	ρ^2	e

9.2.7 Remarques.

(a) On a vu en 9.1.3 que D_2, qui est d'ordre 4, est isomorphe au groupe de Klein V. On a vu en 9.2.5 que D_3, qui est d'ordre 6, est isomorphe au groupe symétrique S_3 (en fait, comme on l'a déjà dit en 3.5.5, il n'existe à isomorphisme près qu'un seul groupe non abélien d'ordre 6). Le groupe D_4 est d'ordre 8, non abélien, mais on pourra voir en exercice qu'il existe d'autres groupes non abéliens d'ordre 8 non isomorphes à D_4, comme le groupe de quaternions Q_8.

(b) Reprenons les notations de 9.2.5. Notons $H = \langle\rho\rangle = \{\mathrm{id}_{\mathcal{E}}, \rho, \rho^2, \ldots, \rho^{n-1}\}$ le sous-groupe cyclique de D_n engendré par ρ. Notons $K = \langle\sigma\rangle = \{\mathrm{id}_{\mathcal{E}}, \sigma\}$ le sous-groupe cyclique de D_n engendré par σ. On a bien $D_n = HK$ et $H \cap K = \{\mathrm{id}_{\mathcal{E}}\}$, mais les éléments de H ne commutent pas nécessairement avec les éléments de K. Par exemple $\sigma\rho \neq \rho\sigma$ lorsque $n \geq 3$. Ainsi, D_n n'est pas le produit direct de H par K (au sens de la définition 2.6.2), car la condition (3) n'est pas vérifiée. Elle est ici remplacée par la condition plus faible $HK = KH$, ce qui est équivaut au fait que, quels que soient $h \in H$ et $k \in K$, il existe $h' \in H$ et $k' \in K$ tels que $hk = k'h'$, mais sans avoir nécessairement $h = h'$ et $k = k'$. On verra plus loin en 10.4.4 que cette situation correspond à une notion plus faible que le produit direct (appelée produit semi-direct), et que D_n est le produit semi-direct de H par K.

9.2.8 Proposition (application aux sous-groupes finis du groupe des isométries du plan).

Soit $\mathcal{E}$ le plan affine euclidien orienté. Tout sous-groupe fini de $\mathrm{Is}(\mathcal{E})$ *est cyclique ou diédral.*

Plus explicitement, cela signifie que, pour tout groupe fini G de $\mathrm{Is}(\mathcal{E})$, il existe un entier n tel que $G \simeq C_n$ (auquel cas $n = |G|$) ou $G \simeq D_n$ (auquel cas $|G|$ est pair et $n = |G|/2$).

Preuve. On introduit d'abord les notations suivantes:

(1) pour tout point O de $\mathcal{E}$ et tout entier $n \geq 1$, on note $G(O, n)$ le groupe des rotations de centre O et d'angle $2k\pi/n, 0 \leq k \leq n-1$. C'est un groupe cyclique d'ordre n, engendré par la rotation ρ de centre O et d'angle $2\pi/n$. D'après le lemme 9.2.2, c'est le groupe des déplacements de $\mathcal{E}$ conservant un polygone régulier à n côtés centré en O. A noter que, par convention, $G(O, 1) = \{\mathrm{id}_\mathcal{E}\}$.

(2) pour tout point O de $\mathcal{E}$ et toute droite $\mathcal{D}$ de $\mathcal{E}$ passant par O, et pour tout entier $n \geq 1$, on note $G(O, \mathcal{D}, n)$ le groupe engendré par $G(O, n)$ et la symétrie orthogonale σ par rapport à $\mathcal{D}$. D'après 9.2.4, c'est le groupe des isométries de $\mathcal{E}$ conservant un polygone régulier à n côtés centré en O et dont un des sommets est sur $\mathcal{D}$. Il est donc isomorphe au groupe diédral D_n. A noter que, par convention, $G(O, \mathcal{D}, 1) = \{\mathrm{id}_\mathcal{E}, \sigma\}$.

Fixons un sous-groupe fini G de $\mathrm{Is}(\mathcal{E})$. Posons $G^+ = G \cap \mathrm{Is}^+(\mathcal{E})$, qui est un sous-groupe de G.

Supposons d'abord que $G \subseteq \mathrm{Is}^+(\mathcal{E})$, c'est-à-dire $G = G^+$. Choisissons $A \in \mathcal{E}$ quelconque, et considérons l'orbite $\mathcal{X}_A = \{\varphi(A)\,;\, \varphi \in G\}$ de A sous l'action canonique de G. C'est un ensemble fini de points $\{A, A_1, \ldots, A_m\}$ de $\mathcal{E}$, laissé globalement invariant par G. Tout élément φ de G est donc un déplacement de $\mathcal{E}$ qui fixe l'isobarycentre O de $\mathcal{X}_A$ (voir 9.1.3), et donc (voir 8.2.4) une rotation de centre O. Comme $\varphi^n = \mathrm{id}_\mathcal{E}$ en notant $|G| = n$, on déduit que φ est d'angle $2k\pi/n, 0 \leq k \leq n-1$. On conclut que $G \subseteq G(O, n)$, et donc $G = G(O, n)$ puisque les deux groupes sont de même ordre n.

Supposons maintenant qu'il existe $\sigma \in \mathrm{Is}^-(\mathcal{E}) \cap G$. Tout élément de G est alors soit un élément de G^+, soit le produit de σ par un élément de G^+. D'après la première étape de la preuve, il existe $O \in \mathcal{E}$ et $n = |G^+| \geq 1$ tel que $G^+ = G(O, n)$. Supposons que $n \geq 2$. En notant ρ la rotation de centre O et d'angle $2\pi/n$, on a $\rho \in G^+$ et $\sigma^{-1} \circ \rho \circ \sigma \in G^+$. Donc $\sigma^{-1}(\rho(\sigma(O))) = O$, donc $\rho(\sigma(O)) = \sigma(O)$, donc $\sigma(O) = O$ puisque O est le seul point fixe de ρ. Ainsi $\varphi \in \mathrm{Is}^-(\mathcal{E}))$ fixe O, donc σ est une symétrie orthogonale par rapport à une droite $\mathcal{D}$ de $\mathcal{E}$. On conclut que $G = G(O, \mathcal{D}, n)$. Enfin, dans le cas où $n = 1$, on a $G^+ = \{\mathrm{id}_\mathcal{E}\}$ et $G = \{\mathrm{id}_\mathcal{E}, \sigma\}$, donc $\sigma^2 \in G^+$ c'est-à dire $\sigma^2 = \mathrm{id}_\mathcal{E}$, donc σ est une symétrie orthogonale par rapport à une droite $\mathcal{D}'$, et l'on a $G = G(O', \mathcal{D}', 1)$ pour tout point O' de $\mathcal{D}'$. □

9.3 Groupe du tétraèdre.

9.3.1 Données et notations. Dans ce qui suit, on se place dans l'espace affine euclidien orienté $\mathcal{E}$ de dimension 3.
On considère l'ensemble $\mathcal{X} = \{A_1, A_2, A_3, A_4\}$ des sommets d'un tétraèdre régulier.
Les six arêtes de $\mathcal{X}$ sont les segments $[A_1A_2]$, $[A_1A_3]$, $[A_1A_4]$, $[A_2A_3]$, $[A_2A_4]$ et $[A_3A_4]$, de même longueur.
Les quatre faces $A_1A_2A_3$, $A_2A_3A_4$, $A_3A_4A_1$ et $A_4A_1A_2$ sont des triangles équilatéraux.
On note O l'isobarycentre des points de $\mathcal{X}$.

Figure 6

Pour tout $1 \leq i \leq 4$, la droite (OA_i) coupe la face opposée au sommet A_i en son centre de gravité (associativité des barycentres).

Pour tout $1 \leq i \leq 4$, on note ρ_i la rotation d'axe (OA_i) et d'angle $2\pi/3$.
Il est clair que ρ_i et ρ_i^2 sont des éléments de $G^+_\mathcal{X}$, ce qui donne déjà 8 éléments de $G^+_\mathcal{X}$.

Notons par ailleurs (en rappelant qu'un demi-tour est une rotation d'angle π) :
α le demi-tour d'axe la droite passant par les milieux de $[A_1A_2]$ et $[A_3A_4]$,
β le demi-tour d'axe la droite passant par les milieux de $[A_1A_3]$ et $[A_2A_4]$,
γ le demi-tour d'axe la droite passant par les milieux de $[A_1A_4]$ et $[A_2A_3]$.
Il est clair que α, β et γ sont trois nouveaux éléments de $G^+_\mathcal{X}$.

9.3.2 Théorème. *Le groupe $G_{\mathcal{X}}$ des isométries de l'espace conservant un tétraèdre régulier est d'ordre 24, isomorphe au groupe au groupe symétrique S_4. Le sous-groupe $G_{\mathcal{X}}^+$ des isométries positives est isomorphe au groupe alterné A_4 et formé des 12 rotations suivantes :*

$$G_{\mathcal{X}}^+ = \{\mathrm{id}_{\mathcal{E}}, \alpha, \beta, \gamma, \rho_1, \rho_1^2, \rho_2, \rho_2^2, \rho_3, \rho_3^2, \rho_4, \rho_4^2\}.$$

Preuve. Toute isométrie $\varphi \in G_{\mathcal{X}}$ permute entre eux les 4 sommets A_1, A_2, A_3, A_4 et induit donc par restriction à l'ensemble de ces 4 sommets une permutation dans S_4, que l'on notera $g(\varphi)$. On définit ainsi une application $g : G_{\mathcal{X}} \to S_4$, qui est de façon évidente un morphisme de groupes.

Par exemple, pour tout $1 \leq i \leq 4$, la permutation $g(\rho_i)$ est un des huit 3-cycles de S_4 (voir 4.4.4) et les rotations α, β, γ correspondent aux trois produits de deux transpositions dans S_4 (voir aussi 4.4.4). Remarquons que ce sont dans les deux cas des éléments du groupe alterné A_4.

Déterminons le noyau de g. Soit $\varphi \in \mathrm{Ker}\, g$. Cela signifie que $\varphi(A_i) = A_i$ pour tout $1 \leq i \leq 4$. L'application linéaire f associée à φ vérifie donc, pour tout $2 \leq j \leq 4$:

$$f(\overrightarrow{A_1A_j}) = \overrightarrow{\varphi(A_1)\varphi(A_j)} = \overrightarrow{A_1A_j}.$$

Or, les 4 sommets du tétraèdre n'étant pas coplanaires dans $\mathcal{E}$, les 3 vecteurs $\overrightarrow{A_1A_2}$, $\overrightarrow{A_1A_3}$, $\overrightarrow{A_1A_4}$ sont libres dans l'e.v. associé E, donc forment une base de E. Ainsi f fixe les vecteurs d'une base de E, d'où $f = \mathrm{id}_E$. Il en résulte (voir 8.2.3) que φ est une translation de $\mathcal{E}$. Mais on sait (voir 9.1.3) que φ fixe O. On conclut que $\varphi = \mathrm{id}_{\mathcal{E}}$.

On a ainsi prouvé que le morphisme $g : G_{\mathcal{X}} \to S_4$ est injectif. Il en résulte que $|G_{\mathcal{X}}| \leq 24$. On a déjà trouvé directement 12 éléments dans $G_{\mathcal{X}}^+$ (les 8 rotations ρ_i et ρ_i^2, les 3 rotations α, β, γ, et l'identité de $\mathcal{E}$), d'où $|G_{\mathcal{X}}^+| \geq 12$. Comme $G_{\mathcal{X}}^- \neq \emptyset$ (il contient par exemple la symétrie orthogonale par rapport au plan (OA_1A_2)), ceci implique (voir 9.1.2) que $|G_{\mathcal{X}}| \geq 24$. Finalement, on a $|G_{\mathcal{X}}| = 24$, et g étant un morphisme injectif entre deux groupes finis de même ordre, il est nécessairement bijectif, ie. un isomorphisme.

On a déjà observé au début de la preuve que les images par g des différents éléments de $G_{\mathcal{X}}^+$ sont les éléments de A_4, et donc la restriction de g à $G_{\mathcal{X}}^+$ réalise un isomorphisme entre $G_{\mathcal{X}}^+$ et A_4. □

9.3.3 Remarque. La table du groupe $G_{\mathcal{X}}^+$ s'obtient directement via l'isomorphisme g à partir de celle du groupe A_4 détaillée en 4.4.4.

9.3.4 Remarque. Les 12 éléments de $G_{\mathcal{X}}^-$ (qui sont d'après 9.1.2 les produits de chacun des 12 déplacements de $G_{\mathcal{X}}^+$ par un antidéplacement choisi dans $G_{\mathcal{X}}^-$) sont les images réciproques par g^{-1} des permutations impaires dans S_4. Il s'agit donc des images réciproques des six transpositions $[i, j]$ et des six 4-cycles $[i, j, k, \ell]$ de S_4. Donnons-en une description géométrique :

$\varphi \in G_{\mathcal{X}}^-$	$g(\varphi) \in S_4$
pour $1 \leq i, j, k, \ell \leq 4$ deux à deux distincts :	(6 éléments)
$\sigma_{(OA_kA_\ell)} = \tau_{ij}$	
$=$ symétrie orthogonale p/r	
au plan médiateur de $[A_iA_j]$	$[ij]$
ie. p/r au plan (OA_kA_ℓ)	
elle fixe A_k et A_ℓ en échangeant A_i et A_j	

$\varphi \in G_{\mathcal{X}}^-$	$g(\varphi) \in S_4$
$\sigma_{(OA_1A_2)} \circ \gamma$	[1423]
$\sigma_{(OA_1A_2)} \circ \beta$	[1324]
$\sigma_{(OA_1A_3)} \circ \alpha$	[1234]
$\sigma_{(OA_1A_3)} \circ \gamma$	[1432]
$\sigma_{(OA_1A_4)} \circ \alpha$	[1342]
$\sigma_{(OA_1A_4)} \circ \beta$	[1243]

Observons que l'engendrement de S_4 par les transpositions correspond à l'engendrement de $G_{\mathcal{X}}$ par des symétries orthogonales par rapport à un plan.

9.4 Groupe du cube.

9.4.1 On se place toujours dans l'espace affine euclidien orienté $\mathcal{E}$ de dimension 3. On considère l'ensemble $\mathcal{X} = \{A, B, C, D, A', B', C', D'\}$ des 8 sommets d'un cube.

6 faces:	$ABCD$, $A'B'C'D'$, $ABB'A'$, $DCC'D'$, $AA'D'D$, $BB'C'C$	
12 arêtes:	$[AA']$, $[BB']$, $[CC']$, $[DD']$, $[AD]$, $[BC]$, $[B'C']$, $[A'D']$, $[AB]$, $[A'B']$, $[D'C']$, $[DC]$	longueur a
12 petites diagonales:	$[AC]$, $[BD]$, $[A'C']$, $[B'D']$, $[A'B]$, $[AB']$, $[D'C]$, $[DC']$, $[AD']$, $[A'D]$, $[BC']$, $[B'C]$	longueur $\sqrt{2}a$
4 grandes diagonales:	$[AC']$, $[A'C]$, $[DB']$, $[D'B]$	longueur $\sqrt{3}a$

On note O l'isobarycentre des points de $\mathcal{X}$; O est le point de concours des 4 grandes diagonales. On considère les deux tétraèdres réguliers $\mathcal{T}_1 = AB'CD'$ et $\mathcal{T}_2 = A'BC'D$ inscrits dans le cube. Ils sont symétriques l'un de l'autre par rapport à O. La longueur de leur arête est $\sqrt{2}a$.

Figure 7

Figure 8

On considère :

- les quatre droites $\mathcal{D}_1 = (AC')$, $\mathcal{D}_2 = (BD')$, $\mathcal{D}_3 = (CA')$ et $\mathcal{D}_4 = (DB')$ portant les grandes diagonales du cube,
- les trois droites $\Delta_1, \Delta_2, \Delta_3$ passant par les isobarycentres des trois paires de faces opposées du cube ($ABCD$ et $A'B'C'D'$, $ABB'A'$ et $DCC'D'$, $AA'D'D$ et $BB'C'C$),
- les six droites δ_i (pour $1 \leq i \leq 6$) joignant les milieux des six paires d'arêtes opposées ($[AB]$ et $[D'C']$, $[CD]$ et $[A'B']$, $[AA']$ et $[CC']$, $[BB']$ et $[DD']$, $[AD]$ et $[B'C']$, $[A'D']$ et $[BC]$).

On en déduit les rotations suivantes de $G_{\mathcal{X}}^+$:

- pour $1 \leq i \leq 4$, ρ_i la rotation d'axe $\mathcal{D}_i$ et d'angle $2\pi/3$, ainsi que ρ_i^2 d'angle $4\pi/3$, qui conserve le tétraèdre $\mathcal{T}_1$,
- pour $1 \leq i \leq 3$, φ_i la rotation d'axe Δ_i et d'angle $\pi/2$, ainsi que φ_i^2 d'angle π et φ_i^3 d'angle $3\pi/2$,
- pour $1 \leq i \leq 6$, ψ_i d'axe δ_i et d'angle π (demi-tours).

On obtient ainsi $8 + 9 + 6 = 23$ éléments de $G_{\mathcal{X}}^+$, auxquels il faut ajouter bien sûr $\mathrm{id}_{\mathcal{E}}$. Donc $|G_{\mathcal{X}}^+| \geq 24$. Le théorème suivant montre que l'on a en fait ici tous les éléments de $G_{\mathcal{X}}^+$.

9.4.2 Théorème. *Le groupe $G_{\mathcal{X}}$ des isométries de l'espace conservant un cube est d'ordre 48. Le sous-groupe $G_{\mathcal{X}}^+$ des isométries positives est isomorphe au groupe symétrique S_4, et formé des 24 rotations décrites ci-dessus.*

Preuve. Considérons l'action canonique (voir 9.1.4) de ${G_{\mathcal{X}}}^+$ sur $\mathcal{X}$, définie par $\varphi.M = \varphi(M)$ pour tout $\varphi \in {G_{\mathcal{X}}}^+$ et tout $M \in \mathcal{E}$. Prenons un point de $\mathcal{X}$, par exemple A, et notons Ω son orbite pour cette action. Comme $\Omega \subseteq \mathcal{X}$, on a $|\Omega| \leq 8$. Notons H le stabilisateur de A pour cette action, ie. le sous-groupe des éléments de ${G_{\mathcal{X}}}^+$ fixant A. Soit $\varphi \in H$. Comme φ conserve la distance en fixant A, et que B', C et D' sont les seuls sommets du cube situés à la distance $\sqrt{2}a$ de A, l'isométrie φ laisse le

tétraèdre $\mathcal{T}_1$ globalement invariant. De plus φ fixe l'isobarycentre O de $\mathcal{X}$ d'après la proposition 9.1.3. En résumé φ est un élément de $G^+_{\mathcal{T}_1}$ fixant à la fois le sommet A du tétraèdre $\mathcal{T}_1$ et le point O, donc la droite $(AO) = \mathcal{D}_1$. D'après l'étude faite en 9.3, φ ne peut être que ρ_1, ρ_1^2 ou $\rho_1^3 = \mathrm{id}_\mathcal{E}$. Ceci prouve que $|H| \leq 3$. En appliquant le théorème 7.3.2, on a $|\Omega| = |{G_\mathcal{X}}^+|/|H|$ avec $|\Omega| \leq 8$ et $|H| \leq 3$. On déduit que $|{G_\mathcal{X}}^+| \leq 8 \times 3 = 24$. Comme on a ci-dessus explicité déjà 24 élément de ${G_\mathcal{X}}^+$, on conclut que $|{G_\mathcal{X}}^+| = 24$. Il en résulte que $|G_\mathcal{X}| = 48$ d'après 9.1.2, puisque $G_\mathcal{X}$ contient par exemple la symétrie centrale de centre O.

Il reste à montrer que ${G_\mathcal{X}}^+ \simeq S_4$. Soit $\varphi \in {G_\mathcal{X}}^+$. On a l'égalité des distances $\varphi(A)\varphi(C') = AC' = \sqrt{3}a$ puisque φ est une isométrie. Comme les seuls sommets de $\mathcal{X}$ deux à deux distants de $\sqrt{3}a$ sont les paires $\{A, C'\}$, $\{B, D'\}$, $\{C, A'\}$ et $\{D, B'\}$, la paire $\{\varphi(A), \varphi(C')\}$ coïncide avec l'une de ces quatre paires de points. Il en résulte que l'image par φ de la droite $\mathcal{D}_1$ est l'une des quatre droites $\mathcal{D}_1$, $\mathcal{D}_2$, $\mathcal{D}_3$ et $\mathcal{D}_4$. En d'autres termes, toute isométrie de ${G_\mathcal{X}}^+$ permute ces quatre droites, ce qui définit naturellement un morphisme de groupes $h : {G_\mathcal{X}}^+ \to S_4$. Si $\varphi \in \mathrm{Ker}\, h$, alors φ transforme chacune des quatre droites $\mathcal{D}_i$ en elle-même; donc l'application linéaire f associée à φ transforme un vecteur directeur $\overrightarrow{u_i}$ de chacune des droites $\mathcal{D}_i$ en lui-même ou son opposé : notons $\varphi(\overrightarrow{u_i}) = \varepsilon_i \overrightarrow{u_i}$ avec $\varepsilon_i = \pm 1$. Il est clair que $\{\overrightarrow{u_1}, \overrightarrow{u_2}, \overrightarrow{u_3}\}$ est une base de E. La matrice de f par rapport à cette base est diagonale avec $\varepsilon_1, \varepsilon_2, \varepsilon_3$ comme coefficients diagonaux. Le vecteur $\overrightarrow{u_4} = \alpha\overrightarrow{u_1} + \beta\overrightarrow{u_2} + \gamma\overrightarrow{u_3}$ vérifie alors:

$$\varphi(\overrightarrow{u_4}) = \alpha\varphi(\overrightarrow{u_1}) + \beta\varphi(\overrightarrow{u_2}) + \gamma\varphi(\overrightarrow{u_3}) = \alpha\varepsilon_1\overrightarrow{u_1} + \beta\varepsilon_2\overrightarrow{u_2} + \gamma\varepsilon_3\overrightarrow{u_3}.$$

En identifiant avec $\varphi(\overrightarrow{u_4}) = \varepsilon_4\overrightarrow{u_4} = \varepsilon_4\alpha\overrightarrow{u_1} + \varepsilon_4\beta\overrightarrow{u_2} + \varepsilon_4\gamma\overrightarrow{u_3}$, il vient $\varepsilon_4 = \varepsilon_3 = \varepsilon_2 = \varepsilon_1 = \pm 1$. On déduit que $f = \pm\,\mathrm{id}_E$. Mais $f \in \mathrm{O}^+(E)$ puisque $\varphi \in \mathrm{Is}^+(\mathcal{E})$, de sorte que seul le cas $f = \mathrm{id}_E$ est possible. Il en résulte (voir 8.2.3) que φ est une translation. Il ne reste qu'à rappeler que O est un point fixe de φ (voir 9.1.3) pour déduire que $\varphi = \mathrm{id}_\mathcal{E}$.

On a ainsi montré que $\mathrm{Ker}\, h = \{\mathrm{id}_\mathcal{E}\}$, et donc que le morphisme de groupes $h : {G_\mathcal{X}}^+ \to S_4$ est injectif. Comme on sait déjà que $|{G_\mathcal{X}}^+| = |S_4| = 24$, on conclut que h est un isomorphisme. □

9.4.3 Exercice. Montrer que ${G_\mathcal{X}}^+ \simeq S_4$ est aussi le groupe des déplacements de $\mathcal{E}$ laissant invariant l'octaèdre régulier dont les six sommets sont les centres des six faces du cube $\mathcal{X}$.

9.4.4 Remarque. Conformément à 9.1.2, tout élément de ${G_\mathcal{X}}^-$ est le produit d'une des 24 rotations de ${G_\mathcal{X}}^+$ par un élément de ${G_\mathcal{X}}^-$, par exemple la symétrie centrale σ de centre O. On peut en déduire, géométriquement une description des 24 isométries indirectes de ${G_\mathcal{X}}^-$, et algébriquement une description du groupe $G_\mathcal{X}$ comme un produit semi-direct (voir plus loin en 10.4) du groupe symétrique S_4 (isomorphe à ${G_\mathcal{X}}^+$) par le groupe cyclique C_2 (isomorphe à $\langle\sigma\rangle$).

9.5 Autres exemples.

Les modes de raisonnement développés sur les exemples précédents s'appliquent, avec des degrés de sophistication divers, à des configurations géométriques nombreuses et diverses. Parmi les exemples classiques qui pourront faire l'objet d'exercices de travaux dirigés, citons le groupe des isométries laissant invariant :
- un parallélogramme dans le plan de dimension 2 (distinguer suivant qu'il s'agit d'un carré, d'un losange, d'un rectangle, ou d'aucun de ces cas particuliers),
- la réunion de deux droites sécantes dans le plan de dimension 2 (distinguer suivant qu'elles sont ou non perpendiculaires),
- la réunion de deux droites non coplanaires dans l'espace de dimension 3 (distinguer suivant qu'elles sont ou non orthogonales),
- la réunion de deux plans perpendiculaires dans l'espace de dimension 3,
- un octaèdre (resp. un dodécaèdre, resp. un icosaèdre) régulier dans l'espace de dimension 3.
...

Leçon 10

Quelques compléments sur les groupes quotients

On expose dans cette leçon quelques sujets complémentaires sur les groupes, que l'on peut considérer comme "à la limite du programme". Les notes ci-dessous, que l'on pourra prendre comme des sujets de travaux dirigés ou d'exercices personnels, ont pour principal objectif de donner de ces résultats classiques une présentation qui soit dans le prolongement direct des développements précédents, avec des notations cohérentes et des références explicites aux résultats antérieurs utilisés.

10.1 Propriété universelle du groupe quotient.

Le théorème ci-dessous apporte une réponse à la question naturelle (et fréquente dans les situations que l'on peut rencontrer) de savoir si tout morphisme $f : G \to G'$ donne naissance à un morphisme $G/H \to G'$ pour tout quotient G/H du groupe de départ G.

10.1.1 Théorème. *Soient G un groupe, H un sous-groupe normal dans G, et p la surjection canonique $G \to G/H$.*

(i) *Pour tout groupe G' et tout morphisme de groupes $f : G \to G'$ tel que $H \subseteq \operatorname{Ker} f$, il existe un unique morphisme de groupes $\varphi : G/H \to G'$ tel que $f = \varphi \circ p$.*

$$\begin{array}{ccc} G & \xrightarrow{\ f\ } & G' \\ {\scriptstyle p}\downarrow & \nearrow_{\varphi} & \\ G/H & & \end{array}$$

(ii) *Avec les données ci-dessus, on a de plus:*

$$(\ f \text{ surjectif } \Rightarrow\ \varphi \text{ surjectif }) \quad \text{et} \quad (\ H = \operatorname{Ker} f\ \Rightarrow\ \varphi \text{ injectif }).$$

Preuve. Pour tout $\overline{x} \in G/H$, posons $\varphi(\overline{x}) = f(x) \in G'$. Montrons que cette définition est indépendante du choix du représentant dans $\overline{x}$. Pour cela, considérons $y \in G$ tel que $\overline{y} = \overline{x}$; on a par définition $xy^{-1} \in H$. Puisque $H \subseteq \operatorname{Ker} f$, on déduit que $xy^{-1} \in \operatorname{Ker} f$, donc $f(xy^{-1}) = e$ d'où $f(x)f(y)^{-1} = e'$, c'est-à-dire $f(x) = f(y)$, ou encore $\varphi(\overline{x}) = \varphi(\overline{y})$. On définit donc bien une application:

$$\begin{array}{rccc} \varphi : & G/H & \longrightarrow & G' \\ & \overline{x} & \longmapsto & f(x) \end{array}$$

qui, par définition, vérifie $\varphi \circ p = f$ puisque $\varphi(p(x)) = \varphi(\overline{x}) = f(x)$ pour tout $x \in G$. Il est clair que φ est un morphisme de groupes puisque, pour tous $\overline{x}, \overline{y} \in G/H$, on a $\varphi(\overline{x}.\overline{y}) = \varphi(\overline{xy}) = f(xy) = f(x)f(y) = \varphi(\overline{x})\varphi(\overline{y})$. Il reste à montrer l'unicité de φ. Soit donc ψ un morphisme $G/H \to G'$ tel que $\psi \circ p = f$. Alors, pour tout $\overline{x} \in G/H$, on a: $\psi(\overline{x}) = \psi(p(x)) = (\psi \circ p)(x) = f(x) = \varphi(\overline{x})$. D'où $\psi = \varphi$, ce qui achève de montrer le point (i).

Pour (ii), supposons d'abord que f est surjective. Soit $x' \in G'$ quelconque. Par surjectivité de f, il existe $x \in G$ tel que $x' = f(x)$. Comme $f(x) = \varphi(\overline{x})$, on déduit qu'il existe $\overline{x} \in G/H$ tel que $x' = \varphi(\overline{x})$. Ce qui prouve que φ est surjective.

Supposons enfin que $H = \operatorname{Ker} f$. Soit $\overline{x} \in G/H$ tel que $\overline{x} \in \operatorname{Ker} \varphi$. On a $e' = \varphi(\overline{x}) = f(x)$, d'où $x \in \operatorname{Ker} f$, c'est-à-dire $x \in H$; par suite $\overline{x} = \overline{e}$. Ceci prouve que $\operatorname{Ker} \varphi = \{\overline{e}\}$, et l'injectivité de φ. □

10.1.2 Remarque. Dans le cas où l'on prend dans le théorème ci-dessus $H = \operatorname{Ker} f$ et $G' = \operatorname{Im} f$, on a φ à la fois injectif et surjectif, qui réalise donc un isomorphisme de $G/\operatorname{Ker} f$ sur $\operatorname{Im} f$, et l'on retrouve le premier théorème d'isomorphisme vu en 6.4.1.

La question traitée au théorème 10.1.1 pour le groupe de départ du morphisme a son analogue pour le groupe d'arrivée. C'est l'objet du lemme suivant.

10.1.3 Lemme (fondamental de factorisation). *Soient G un groupe, H un sous-groupe normal dans G, et p la surjection canonique $G \to G/H$. Soient G' un groupe, H' un sous-groupe normal dans G', et p' la surjection canonique $G' \to G'/H'$. Alors, pour tout morphisme de groupes $f : G \to G'$ vérifiant la condition $f(H) \subseteq H'$, il existe un unique morphisme $\varphi : G/H \to G'/H'$ tel que $\varphi \circ p = p' \circ f$.*

$$\begin{array}{ccc} G & \xrightarrow{f} & G' \\ {\scriptstyle p}\downarrow & \overset{g}{\searrow} & \downarrow{\scriptstyle p'} \\ G/H & \xrightarrow[\varphi]{} & G'/H' \end{array}$$

Preuve. Posons $g = p' \circ f$, qui est un morphisme de groupes $G \to G'/H'$, comme composé de deux morphismes. Afin d'appliquer le théorème 10.1.1, montrons que $H \subseteq \operatorname{Ker} g$. Soit $x \in H$. On a $f(x) \in f(H)$. L'hypothèse $f(H) \subseteq H'$ implique donc $f(x) \in H'$. D'où $p'(f(x)) = \overline{e}'$. On déduit que $g(x) = \overline{e}'$, c'est-à-dire $x \in \operatorname{Ker} g$. Ainsi $g : G \to G'/H'$ est un morphisme vérifiant $H \subseteq \operatorname{Ker} g$; le théorème 10.1.1 assure l'existence d'un unique morphisme $\varphi : G/H \to G'/H'$ tel que $\varphi \circ p = g$, d'où le résultat. □

10.1.4 Remarque. Avec les données et notations ci-dessus, on a:

$$(\ p' \circ f \text{ surjectif } \Rightarrow\ \varphi \text{ surjectif }) \quad \text{et} \quad (\ H = \operatorname{Ker}(p' \circ f) \ \Leftrightarrow\ f^{-1}(H') = H \ \Rightarrow\ \varphi \text{ injectif }).$$

10.2 Quotient par une intersection et deuxième théorème d'isomorphisme.

10.2.1 Théorème. *Soient G un groupe et H un sous-groupe normal dans G. Pour tout sous-groupe K de G, le sous-ensemble HK est un sous groupe de G, et l'on a:*

$$H \cap K \lhd K, \qquad H \lhd HK, \qquad \text{et} \qquad K/(H \cap K) \simeq HK/H.$$

Preuve. Rappelons que $HK = \{hk\,;\, h \in H, k \in K\}$.

Vérifions que HK est un sous-groupe de G. On a clairement $e \in HK$. Soient $x, y \in HK$. Il existe $h, h' \in K$ et $k, k' \in K$ tels que $x = hk$ et $y = h'k'$. Donc $x^{-1}y = k^{-1}h^{-1}h'k' = k^{-1}(h^{-1}h')k(k^{-1}k')$. Puisque $h^{-1}h' \in H$ et $H \lhd G$, on a $k^{-1}(h^{-1}h')k \in H$. Comme par ailleurs $k^{-1}k' \in K$, on a bien $x^{-1}y \in HK$. On conclut que HK est un sous-groupe de G.

Vérifions que $H \cap K \lhd K$. Soit $h \in H \cap K$ et $x \in K$. On a $xhx^{-1} \in H$ puisque $H \lhd G$. On a aussi $xhx^{-1} \in K$ puisque x et h appartiennent au sous-groupe K. On conclut que $xhx^{-1} \in H \cap K$. Ce qui pouve que $H \cap K \lhd K$. On peut donc considérer le groupe quotient $K/H \cap K$. Notons $p : K \to K/H \cap K$ la surjection canonique.

Vérifions que $H \lhd HK$. Soit $\ell \in H$ et $x = hk \in HK$, avec $h \in H, k \in K$. On a $x\ell x^{-1} = hk\ell k^{-1}h^{-1}$. Puisque $H \lhd G$ et $\ell \in H$, on a $k\ell k^{-1} \in H$. Donc $x\ell x^{-1} = h(k\ell k^{-1})h^{-1} \in H$ comme produit de trois éléments de H. Ce qui prouve que $H \lhd HK$. On peut donc considérer le groupe quotient HK/H. Notons $p' : HK \to HK/H$ la surjection canonique.

Notons j l'injection canonique $K \to HK$. Rappelons que j est le morphisme défini par $j(k) = ke = k$ pour tout $k \in K$. On a bien sûr $j(H \cap K) \subseteq H$, de sorte que l'application directe du lemme 10.1.3 assure l'existence d'un morphisme de groupes $\varphi : K/H \cap K \to HK/H$ tel que $\varphi \circ p = p' \circ j$:

$$\begin{array}{ccc} K & \xrightarrow{j} & HK \\ {\scriptstyle p}\downarrow & & \downarrow{\scriptstyle p'} \\ K/H \cap K & \xrightarrow{\varphi} & HK/H \end{array}$$

Montrons que φ est surjective. Soit $\overline{hk}$ un élément quelconque de HK/H, avec $h \in H, k \in K$. On a $\overline{hk} = \overline{h}\,\overline{k} = \overline{k}$; on déduit que $HK/H = p'(K) = (p' \circ j)(K)$. On conclut avec 10.1.4 que φ est surjective.

Montrons que φ est injective. Soit $k \in K$ un élément quelconque de $\operatorname{Ker}(p' \circ j)$. On a $\overline{e} = p'(j(k)) = p'(k) = \overline{k}$, c'est-à-dire $k \in H$. Donc $k \in H \cap K$; on déduit que $\operatorname{Ker}(p' \circ j) \subseteq H \cap K$. L'inclusion réciproque étant claire, on déduit que $\operatorname{Ker}(p' \circ j) = H \cap K$. On conclut avec 10.1.4 que φ est injective. On conclut que φ réalise un isomorphisme de $K/H \cap K$ sur HK/H. □

10.2.2 Remarque. En notation additive, l'isomorphisme 10.2.1 devient $K/(H\cap K) \simeq (H+K)/H$.

10.3 Sous-groupes d'un groupe quotient et troisième théorème d'isomorphisme.

La question naturelle résolue par le théorème suivant est celle de la description des sous-groupes d'un groupe quotient G/H en fonction des sous-groupes de G.

10.3.1 Proposition. *Soient G un groupe et H un sous-groupe normal dans G. L'ensemble des sous-groupes de G/H est en bijection avec l'ensemble des sous-groupes de G contenant H.*
Plus précisément, si l'on note $p : G \to G/H$ la surjection canonique, il existe pour tout sous-groupe $\overline{K}$ de G/H un unique sous-groupe K de G contenant H tel que $\overline{K} = p(K) = K/H$.

Preuve. Soit $\overline{K}$ un sous-groupe de G/H. Posons $K = p^{-1}(\overline{K}) = \{x \in G\,;\, p(x) \in \overline{K}\}$. En tant qu'image réciproque d'un sous-groupe par un morphisme de groupes, K est un sous-groupe de G. Si $h \in H$, on a $p(h) = \overline{e}$, donc $p(h) \in \overline{K}$, de sorte que $h \in p^{-1}(\overline{K})$, c'est-à-dire $h \in K$. Ceci montre que $H \subseteq K$. Par définition de K, on a $p(K) \subseteq \overline{K}$. Réciproquement, soit $\overline{x} \in \overline{K}$, avec $x \in G$; comme $p(x) = \overline{x} \in \overline{K}$, on a clairement $x \in p^{-1}(\overline{K}) = K$, et donc $\overline{x} = p(x) \in p(K)$. En résumé, $\overline{K} = p(K)$. Enfin, $H \lhd G$ implique $H \lhd K$, et il est clair alors que $K/H = p(K)$.

Montrons maintenant l'unicité. Soit donc K' un sous-groupe de G tel que $H \subseteq K'$ et $\overline{K} = p(K')$. On a donc $p(K) = p(K')$. Quel que soit $k' \in K'$, il existe alors $k \in K$ tel que $p(k') = p(k)$, donc $k'k^{-1} \in H$; on a $k' = hk$ avec $h \in H$, et l'hypothèse $H \subseteq K$ implique $h \in K$, d'où $k' \in K$ comme produit de deux éléments de K. On conclut que $K' \subseteq K$. L'inclusion réciproque s'obtient de même. □

10.3.2 Théorème. *Soient G un groupe et H un sous-groupe normal dans G. Pour tout sous-groupe K normal dans G contenant H, on a:*

$$K/H \lhd G/H, \qquad \text{et} \qquad (G/H)\Big/(K/H) \simeq G/K.$$

Preuve. Notons q la surjection canonique $G \to G/K$ et p la surjection canonique $G \to G/H$. Il est clair que $K/H = p(K)$ est un sous-groupe de G/H (comme image du sous-groupe K par le morphisme de groupes p). Quels que soient $\overline{x} \in G/H$ et $\overline{k} \in K/H$, on a $\overline{x}\overline{k}\overline{x}^{-1} = p(xkx^{-1})$; or $xkx^{-1} \in K$ puisque $K \lhd G$, donc $\overline{x}\overline{k}\overline{x}^{-1} \in p(K)$. Ceci prouve que $K/H \lhd G/H$. On peut donc considérer le groupe quotient $(G/H)/(K/H)$; notons q' la surjection canonique $G/H \to (G/H)/(K/H)$. Puisque $p(K) = K/H$, on applique le lemme 10.1.3 pour conclure qu'il existe un morphisme $\varphi : G/K \to (G/H)/(G/K)$ tel que $\varphi \circ q = q' \circ p$.

$$\begin{array}{ccc} G & \xrightarrow{\;p\;} & G/H \\ {\scriptstyle q}\downarrow & & \downarrow{\scriptstyle q'} \\ G/K & \xrightarrow{\;\varphi\;} & \dfrac{(G/H)}{(K/H)} \end{array}$$

Le morphisme $q' \circ p$ est surjectif comme composé de deux surjections, et il résulte donc de 10.1.4 que φ est surjectif. On a $\mathrm{Ker}(q' \circ p) = K$, d'où l'on déduit avec 10.1.4 que φ est injectif. On conclut que φ est un isomorphisme de groupes de G/K sur $(G/H)/(K/H)$. □

10.4 Produit semi-direct.

10.4.1 Rappel. On a vu en 2.6.2 qu'un groupe G est produit direct (interne) de deux de ses sous-groupes H par K lorsque les trois conditions suivantes sont vérifiées:

$$(1)\ \ G = HK, \qquad (2)\ \ H \cap K = \{e\}, \qquad (3)\ \ \forall\ h \in H,\ \forall\ k \in K,\ hk = kh.$$

Tout élément de G s'écrit de façon unique comme le produit d'un élément de H par un élément de K. On en a déduit au corollaire 6.4.4 que $H \lhd G$ et $G/H \simeq K$.
De plus, la condition (3) implique que les deux sous-groupes H et K jouent dans un tel produit direct des rôles absolument symétriques. On a donc également $K \lhd G$ et $G/K \simeq H$.

10.4.2 Définition. Soient G un groupe, H et K deux sous-groupes de G. On dit que G est le produit semi-direct (interne) de H par K lorsque les trois conditions suivantes sont vérifiées:

$$(1)\ \ G = HK, \qquad (2)\ \ H \cap K = \{e\}, \qquad (3')\ \ H \lhd G.$$

10.4.3 Remarques. Avec les notations ci-dessus:

(a) Si G est produit semi-direct de H par K, on a aussi $G = KH$.

En effet, d'après la condition (1), tout élément x de G s'écrit $x = hk$ avec $h \in H$ et $k \in K$. Donc $g = kk^{-1}hk$, et comme $H \lhd G$, le produit $h' = k^{-1}hk$ est un élément de H. On a donc $g = kh'$ avec $k \in K$ et $h' \in H$. □

(b) Si G est produit semi-direct de H par K, tout élément x de G s'écrit de façon unique $x = hk$ avec $h \in H, k \in K$, et s'écrit aussi de façon unique $x = k'h'$ avec $h' \in H, k' \in K$, mais pas forcément avec $h = h'$.

Preuve. Le fait qu'un élément quelconque $x \in G$ s'écrive $x = hk = kh'$ avec $h' = k^{-1}hk$ a été vu à la remarque ci-dessus. L'unicité des décompositions découle de la seule condition (2) comme on l'a vu à la remarque 2.6.1.a. □

(c) Si G est produit semi-direct de H par K, alors $G/H \simeq K$.

Preuve. Analogue à celle de 6.4.4. □

(d) Si G est produit direct de H par K, alors a fortiori G est produit semi-direct de H par K.

Preuve. Il s'agit de vérifier que, si les conditions (1) et (2) sont vérifiées, alors la condition (3) implique la condition (3′). Pour cela, soient $\ell \in H$ et $x \in G$ quelconques. Il existe $h \in H, k \in K$ tels que $x = hk$. D'après la condition (3), on a: $x\ell x^{-1} = hk\ell k^{-1}h^{-1} = h\ell kk^{-1}h^{-1} = h\ell h^{-1}$, qui est un élément de H comme produit de trois éléments de H. Ceci prouve que $H \lhd G$. □

(e) La réciproque de (d) est fausse.

Preuve. Prenons par exemple comme en 6.3.3.c le groupe symétrique $S_3 = \{e, \gamma, \gamma^2, \tau_1, \tau_2, \tau_3\}$. Le sous-groupe alterné est $A_3 = H = \{e, \gamma, \gamma^2\}$. On a $\tau_1 = e\tau_1$, $\tau_3 = \gamma\tau_1$ et $\tau_2 = \gamma^2\tau_1$. En posant $K = \{e, \tau_1\}$, on a donc $S_3 = HK$ et $H \cap K = \{e\}$. On conclut que S_3 est le produit semi-direct de H par le sous-groupe $K = \{e, \tau_1\}$. Et pourtant la condition (3) d'un produit direct n'est pas vérifiée puisque par exemple $\tau_1\gamma = \tau_2 \neq \tau_3 = \gamma\tau_1$. □

Parce que $S_3 \simeq D_3$, l'exemple de S_3 n'est qu'un cas particulier du résultat suivant.

10.4.4 Exemple (groupes diédraux). En reprenant les notations de la leçon 9, considérons le groupe diédral:

$$\begin{aligned} D_n &= \{e, r, r^2, r^3, \dots, r^{n-1},\ s\ ,\ sr\ ,\ sr^2\ ,\ sr^3, \dots,\ sr^{n-1}\} \\ &= \{e, r, r^2, r^3, \dots, r^{n-1},\ r^{n-1}s\ ,\ r^{n-2}s\ ,\ \dots,\ r^2s\ ,\ rs\ ,\ s\}, \end{aligned}$$

et les sous-groupes $C_n = \{e, r, r^2, r^3, \dots, r^{n-1}\}$ et $K = \{e, s\}$. Il est clair que $D_n = C_nK$ et $C_n \cap K = \{e\}$. D'après les propositions 5.4.4 ou 6.1.5, on a de plus que $C_n \lhd D_n$. On conclut que:

Le groupe diédral D_n est produit semi-direct du groupe cyclique C_n par le sous-groupe d'ordre deux K.

En particulier, $D_n/C_n \simeq K \simeq C_2$.

Si $n > 2$, ce produit semi-direct n'est pas direct car $sr^k = r^{n-k}s \neq r^ks$, de sorte que la condition (3) n'est pas vérifiée. Dans le cas particulier où $n = 2$, D_2 (qui n'est autre que le groupe de Klein) est abélien, et produit direct de C_2 par K (qui sont tous les deux isomorphes au groupe d'ordre 2).

10.4.5 Exercice. Montrer que, dans $\mathrm{GL}_3(\mathbb{R})$, les matrices triangulaires supérieures dont les termes diagonaux valent 1 forment un sous-groupe, que l'on notera U. Montrer que:

$$H = \{\begin{pmatrix} 1 & 0 & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix},\ b, c \in \mathbb{R}\}, \quad K = \{\begin{pmatrix} 1 & a & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},\ a \in \mathbb{R}\}, \quad L = \{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix},\ b \in \mathbb{R}\}, \quad C = \{\begin{pmatrix} 1 & 0 & c \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},\ c \in \mathbb{R}\},$$

sont des sous-groupes de U, que U est le produit semi-direct de H par K, et que H est le produit direct de L par C.

On a vu à la proposition 2.6.4 que la notion de produit direct interne de deux sous-groupes internes était directement liée à la notion de produit direct externe construit à partir de deux groupes quelconques. C'est aussi le cas pour la notion plus générale de produit semi-direct, comme le montrent la proposition 10.4.7 suivante, et le lemme préliminaire 10.4.6.

10.4.6 Lemme. *Soient G un groupe, H un sous-groupe normal de G, et K un sous-groupe de G. On suppose que G est le produit semi-direct de H par K. Soient x, x' deux éléments quelconques de G. Si $x = hk$ et $x' = h'k'$ sont les décompositions (uniques) de x et x' (avec $h, h' \in H, k, k' \in K$), alors la décomposition du produit xx' est donnée par:*

$$xx' = h\gamma_k(h')kk', \qquad \text{avec } h\gamma_k(h') \in H \text{ et } kk' \in K$$

où γ_k désigne l'automorphisme intérieur de G défini par $y \mapsto kyk^{-1}$.

Preuve. Résulte simplement du calcul $xx' = hkh'k' = hkh'k^{-1}kk'$, et du fait que $kh'k^{-1}$ appartient à H puisque $H \lhd G$. □

10.4.7 Proposition et définition (produit semi-direct externe). *Soient G_1 et G_2 deux groupes. Soit $\gamma : G_2 \to \operatorname{Aut} G_1$ un morphisme de groupes. Pour tout $x_2 \in G_2$, on note γ_{x_2} l'automorphisme de G_1 image de x_2 par γ.*

(i) *Le produit cartésien $G_1 \times G_2 = \{(x_1, x_2),\ x_1 \in G_1,\ x_2 \in G_2\}$ est un groupe pour la loi définie par:*

$$(x_1, x_2).(y_1, y_2) = (x_1\gamma_{x_2}(y_1), x_2y_2) \quad \text{pour tous } x_1, y_1 \in G_1,\ x_2, y_2 \in G_2.$$

Ce groupe est appelé le produit semi-direct de G_1 par G_2. On le note $G = G_1 \times_\gamma G_2$, ou $G = G_1 \rtimes G_2$.

(ii) *Si l'on note $H = G_1 \times \{e_2\}$ et $K = \{e_1\} \times G_2$, alors H est un sous-groupe de G normal dans G et isomorphe à G_1, K est un sous-groupe de G isomorphe à G_2, et G est le produit semi-direct interne de H par K.*

Preuve. La vérification des axiomes de groupes pour (i) et des isomorphismes pour (ii) est technique et fastidieuse, mais élémentaire. C'est un excellent exercice, à faire absolument ! □

10.4.8 Remarques

1. Le produit direct de G_1 et G_2 est un cas particulier de produit semi-direct, correspondant au cas où γ_{x_2} est l'identité de G_1 pour tout $x_2 \in G_2$, c'est-à dire au cas où $\gamma : G_2 \to \operatorname{Aut} G_1$ est le morphisme constant $x_2 \mapsto \mathrm{id}_{G_1}$.
2. Dans le produit semi-direct $G_1 \rtimes G_2$, les groupes G_1 et G_2 ne jouent a priori pas des rôles symétriques. En particulier, même si G_1 et G_2 sont abéliens, $G_1 \rtimes G_2$ n'est en général pas abélien (et de fait il ne l'est que lorsque γ est trivial, c'est-à-dire lorsque le produit est direct). Par exemple, pour $n \geq 3$: les groupes cycliques C_n et C_2 sont abéliens, mais le groupe diédral $D_n \simeq C_n \rtimes C_2$ ne l'est pas.

10.4.9 Théorème (le groupe affine comme produit semi-direct). *Soit $\mathcal{E}$ un $\mathbb{R}$-e.v. de dimension finie, d'e.v. directeur E. Le groupe affine $\mathrm{GA}(\mathcal{E})$ est isomorphe à un produit semi-direct du groupe additif de E par le groupe linéaire $\mathrm{GL}(E)$.*

Preuve. On renvoie pour les rappels relatifs au groupe affine à 8.2.1. Considérons le morphisme de restriction γ du groupe $\mathrm{GL}(E)$ dans le groupe $\operatorname{Aut} E$ des automorphismes additifs de E, ie. $\gamma(f) = f$ pour tout $f \in \mathrm{GL}(E)$. On définit alors sur $E \times \mathrm{GL}(E)$ la structure de produit semi-direct associée à γ, c'est-à-dire:

$$(\vec{u}, f) \cdot (\vec{v}, g) = (\vec{u} + f(\vec{v}), f \circ g) \ \text{ pour tous } \vec{u}, \vec{v} \in E,\ f, g \in \mathrm{GL}(E).$$

Fixons un point $A \in \mathcal{E}$, et définissons l'application $\Gamma : \mathrm{GA}(\mathcal{E}) \to E \times \mathrm{GL}(E)$ qui, à tout $\varphi \in \mathrm{GA}(\mathcal{E})$, associe $\Gamma(\varphi) = (\overrightarrow{A\varphi(A)}, f)$, où f est l'application linéaire associée à φ. On montre que Γ est un isomorphisme de groupes.

Pour montrer que c'est un morphisme, considérons $\varphi, \psi \in \mathrm{GA}(\mathcal{E})$, d'applications linéaires respectives f et g. On calcule:

$\Gamma(\varphi)\cdot\Gamma(\psi) = (\overrightarrow{A\varphi(A)},\, f)\cdot(\overrightarrow{A\psi(A)},\, g) = (\overrightarrow{A\varphi(A)}+f(\overrightarrow{A\psi(A)}),\, f\circ g) = (\overrightarrow{A\varphi(A)}+\overrightarrow{\varphi(A)\varphi(\psi(A))},\, f\circ g) = (\overrightarrow{A\varphi(\psi(A))},\, f\circ g) = \Gamma(\varphi\circ\psi)$

Pour la surjectivité de Γ, fixons $\overrightarrow{u} \in E$ et $f \in \mathrm{GL}(E)$ quelconques. Notons $B \in \mathcal{E}$ le point tel que $\overrightarrow{AB} = \overrightarrow{u}$. On sait [voir 8.2.1.(b).3] qu'il existe $\varphi \in \mathrm{GA}(\mathcal{E})$ telle que $\varphi(A) = B$ et f est l'application linéaire associée à φ. Il est clair que l'on a alors $\Gamma(\varphi) = (\overrightarrow{u}, f)$.

Enfin, prenons $\varphi \in \operatorname{Ker}\Gamma$. On a $f = \mathrm{id}_E$ et $\overrightarrow{A\varphi(A)} = \overrightarrow{0}$. Le premier point implique que φ est une translation (voir 8.2.3.a) et le second qu'elle admet un point fixe. C'est donc $\mathrm{id}_{\mathcal{E}}$, ce qui prouve l'injectivité de Γ et achève la preuve. □

Leçon 11

Anneaux, sous-anneaux, morphismes d'anneaux

11.1 Notion d'anneau

11.1.1 Définition. Un *anneau* est un ensemble non vide muni de deux lois de composition internes, l'une notée comme une addition et l'autre comme une multiplication, vérifiant les propriétés:

(1) A est un groupe abélien pour l'addition, (on note 0 son élément neutre),

(2) la multiplication est associative, c'est-à-dire:
$$x(yz) = (xy)z \quad \text{pour tous } x, y, z \in A.$$

(3) la multiplication est distributive sur l'addition à gauche et à droite, c'est-à-dire:
$$x(y+z) = xy + xz \quad \text{et} \quad (x+y)z = xz + yz \quad \text{pour tous } x, y, z \in A.$$

On dit que l'anneau A est *commutatif* si de plus la multiplication est commutative, c'est-à-dire:
$$xy = yx \quad \text{pour tous } x, y \in A.$$
On dit que A est *unitaire* si de plus la multiplication admet un élément neutre 1.
$$x.1 = 1.x = x \quad \text{pour tout } x \in A.$$

11.1.2 Premiers exemples.

(a) L'ensemble $\mathbb{Z}$ des entiers est un anneau commutatif unitaire. Il en est de même de $\mathbb{Q}$, $\mathbb{R}$ et $\mathbb{C}$.

(b) L'ensemble des matrices carrées d'ordre $n \geq 2$ à coefficients réels est un anneau non-commutatif (pour le produit matriciel) unitaire (de neutre multiplicatif la matrice identité). Il en est de même de l'anneau des endomorphismes d'un espace vectoriel (pour la loi $\circ$).

(c) L'anneau nul est l'anneau $\{0\}$ formé d'un unique élément.

(d) Pour tout intervalle I de $\mathbb{R}$, l'ensemble $\mathcal{F}(I, \mathbb{R})$ des applications de I dans $\mathbb{R}$ est un anneau commutatif (la multiplication étant le produit des fonctions défini par $(fg)(x) = f(x)g(x)$ pour tout $x \in I$) unitaire (de neutre multiplicatif la fonction constante égale à 1). Il en est de même pour l'ensemble $\mathbb{R}^{\mathbb{N}}$ des suites de réels.

11.1.3 Exemple de $\mathbb{Z}/n\mathbb{Z}$. Fixons un entier $n \geq 2$.

Considérons le groupe additif $\mathbb{Z}/n\mathbb{Z} = \{\overline{0}, \overline{1}, \ldots, \overline{n-1}\}$. Rappelons que l'addition est définie par:
$$\overline{x} + \overline{y} = \overline{x+y} \quad \text{pour tous } \overline{x}, \overline{y} \in \mathbb{Z}/n\mathbb{Z}.$$
On a vu que cette définition est indépendante des représentants choisis, et que le groupe additif $\mathbb{Z}/n\mathbb{Z}$ est abélien. On définit une multiplication dans $\mathbb{Z}/n\mathbb{Z}$ à partir de celle de $\mathbb{Z}$ en posant:
$$\overline{x}\ \overline{y} = \overline{xy} \quad \text{pour tous } \overline{x}, \overline{y} \in \mathbb{Z}/n\mathbb{Z}.$$
Cette multiplication est bien définie, indépendamment des représentants choisis.

> *En effet,* si $\overline{x} = \overline{x'}$ et $\overline{y} = \overline{y'}$, alors $x' = x + nu$ et $y' = y + nv$ pour deux entiers $u, v \in \mathbb{Z}$, de sorte que $x'y' = xy + n(uy + vx + nuv)$, d'où $\overline{x'y'} = \overline{xy}$.

Il est immédiat de vérifier que $\mathbb{Z}/n\mathbb{Z}$ satisfait les conditions (2) et (3) de 11.1.1, que $\overline{1}$ est neutre pour la multiplication, et que la multiplication est commutative. On conclut que:

$\mathbb{Z}/n\mathbb{Z}$ est un anneau commutatif unitaire.

On verra plus loin que l'anneau $\mathbb{Z}/n\mathbb{Z}$, que nous avons souhaité introduire dès le début afin de disposer d'exemples significatifs pour les différentes notions que l'on va voir, est un exemple d'anneau quotient.

11.1.4 Exemple des anneaux de polynômes. On fixe un anneau commutatif unitaire A.

Notons (provisoirement) $B = A^{(\mathbb{N})}$ l'ensemble des suites d'éléments de A qui sont "à support fini" c'est-à-dire dont tous les termes sont nuls sauf un nombre fini d'entre eux.

On note $0_B = (0_A, 0_A, \ldots)$. Pour tout $f = (a_n)_{n\in\mathbb{N}}$ distinct de 0_B, on appelle degré de f le plus grand des entiers $n \in \mathbb{N}$ tels que $a_n \neq 0$. On définit une addition et une multiplication dans B en posant, pour tous $f = (a_n)_{n\in\mathbb{N}}$ et $g = (b_n)_{n\in\mathbb{N}}$ dans B,

$$f + g = (a_n + b_n)_{n\in\mathbb{N}} \quad \text{et} \quad fg = (c_n)_{n\in\mathbb{N}}, \text{ avec } c_n = \sum_{i=0}^{n} a_i b_{n-i}.$$

On peut montrer (vérification technique et fastidieuse, mais élémentaire) que, pour ces opérations, B est un anneau commutatif unitaire, avec $0_B = (0_A, 0_A, \ldots)$ et $1_B = (1_A, 0_A, 0_A, \ldots)$. On l'appelle l'anneau des polynômes en une indéterminée à coefficients dans A.

On définit aussi le produit externe d'un élément $\alpha \in A$ par un élément $f = (a_n)_{n\in\mathbb{N}}$ en posant $\alpha f = (\alpha a_n)_{n\in\mathbb{N}}$. A noter que le produit externe αf n'est autre que le produit interne de f par $(\alpha, 0_A, 0_A, \ldots)$. C'est pourquoi on convient de noter encore α l'élément $(\alpha, 0_A, 0_A, \ldots)$ de B. En particulier $0_B = 0_A$ et $1_B = 1_A$.

En posant $e_i = (0_A, 0_A, \ldots, 0_A, 1_A, 0_A, 0_A, \ldots)$, avec le 1_A en $i+1$-ième position, pour tout $i \in \mathbb{N}$, tout élément de B s'écrit de façon unique $f = \sum_{n\in\mathbb{N}} a_n e_n$ avec les $a_n \in A$ nuls sauf un nombre fini d'entre eux (de sorte que la somme est finie). Il est clair que $e_n e_m = e_{n+m}$ pour tous $n, m \in \mathbb{N}$, et donc $e_n = e_1^n$ pour tout $n \in \mathbb{N}$. On note traditionnellement $X = e_1$ et $B = A[X]$, et l'on retrouve les notations usuellement utilisées pour désigner les polynômes.

On retiendra que:

(a) Pour tout anneau commutatif unitaire A, les polynômes en une indéterminée à coefficients dans A forment un anneau commutatif unitaire, noté $A[X]$. Son neutre pour l'addition est 0_A. Son neutre pour la multiplication est 1_A.

(b) Pour tout élément non-nul P de $A[X]$, il existe un unique entier naturel n et un unique $(n+1)$-uplet $(a_0, a_1, \ldots, a_n)$ d'éléments de A, appelés les *coefficients* de P tels que:

$$P = a_n X^n + a_{n-1} X^{n-1} + \cdots + a_1 X + a_0 \quad \text{et} \quad a_n \neq 0.$$

L'entier n est appelé le *degré* de P, noté $\deg P$. L'élément non-nul a_n de A est appelé le *coefficient dominant* de P, noté $\mathrm{cd}(P)$. L'élément a_0 est appelé le *terme constant* de P. Par convention, un polynôme est nul si et seulement si tous ses coefficients sont nuls, et l'on pose $\deg 0 = -\infty$ et $\mathrm{cd}(0) = 0$.

(c) Deux polynômes non-nuls $P = \sum_{i=0}^{n} a_i X^i$ et $Q = \sum_{i=0}^{m} b_i X^i$ sont égaux si et seulement si $n = m$ et $a_i = b_i$ pour tout $0 \leq i \leq n$.

(d) Si $P = \sum_{i=0}^{n} a_i X^i$ et $Q = \sum_{i=0}^{m} b_i X^i$, on a: $\quad P+Q = \sum_{i=0}^{\max(n,m)} (a_i+b_i) X^i$ et $PQ = \sum_{i=0}^{n+m} (\sum_{j=0}^{i} a_j b_{i-j}) X^i$.

Sous forme développée explicite, la formule du produit est donc:
$PQ = (a_n X^n + a_{n-1} X^{n-1} + a_{n-2} X^{n-2} + \cdots + a_1 X + a_0)(b_m X^m + b_{m-1} X^{m-1} + b_{m-2} X^{m-2} + \cdots + b_1 X + b_0) =$
$a_n b_m X^{n+m} + (a_n b_{m-1} + a_{n-1} b_m) X^{n+m-1} + (a_n b_{m-2} + a_{n-1} b_{m-1} + a_{n-2} b_m) X^{n+m-2} + \cdots + (a_1 b_0 + a_0 b_1) X + a_0 b_0.$

(e) On en déduit que, pour tous P et Q dans $A[X]$, on a:

$$\deg(P + Q) \leq \max(\deg P, \deg Q) \quad \text{et} \quad \deg(PQ) \leq \deg P + \deg Q.$$

11.2 Sous-anneau.

11.2.1 Définition. Soit A un anneau. On appelle *sous-anneau* de A toute partie non-vide B de A qui vérifie les deux conditions suivantes:

(1) B est un sous-groupe du groupe additif A.

(2) B est stable par la multiplication de A, c'est-à-dire que l'on a:

$$xy \in B \text{ quels que soient } x \in B \text{ et } y \in B.$$

11.2.2 Définition. Soit A un anneau unitaire. On appelle *sous-anneau unitaire* de A tout sous-anneau de A qui contient 1_A.

11.2.3 Remarques.

(a) Si B est un sous-anneau de A, alors B est lui-même un anneau (pour les lois déduites de celles de A par restriction à B). De fait, dans la pratique, pour montrer qu'un ensemble donné est un anneau, on cherche souvent à montrer que c'est un sous-anneau d'un anneau déjà connu.

(b) Si B est un sous-anneau unitaire d'un anneau unitaire A, alors B est lui-même un anneau unitaire, et l'on a $1_B = 1_A$.

(c) Si l'anneau A est commutatif, alors tout sous-anneau de A est commutatif.

(d) Dans la pratique, pour montrer qu'un sous-ensemble non-vide B d'un anneau A est un sous-anneau de A, il suffit de vérifier que:

$$\text{pour tous } x \in B \text{ et } y \in B, \text{ on a } \ x - y \in B \ \text{ et } \ xy \in B.$$

Pour montrer qu'un sous-ensemble B d'un anneau unitaire A est un sous-anneau unitaire de A, il suffit de vérifier que:

$$(\ 1_A \in B\ ) \quad \text{et} \quad (\ \text{pour tous } x \in B \text{ et } y \in B, \text{ on a } \ x - y \in B \ \text{ et } \ xy \in B\ ).$$

11.2.4 Premiers exemples.

(a) Si A est un anneau, alors $\{0\}$ et A lui-même sont des sous-anneaux de A.

(b) Tout anneau unitaire A est un sous-anneau unitaire de $A[X]$ (le produit dans $A[X]$ de deux polynômes réduits à leur terme constant est égal à leur produit dans l'anneau A).

(c) $\mathbb{Z}$ est un sous-anneau unitaire de $\mathbb{Q}$ (et de $\mathbb{R}$, et de $\mathbb{C}$). Pour tout $n \geq 2$, l'ensemble $n\mathbb{Z} = \{nx\,;\, x \in \mathbb{Z}\}$ est un sous-anneau non unitaire de $\mathbb{Z}$.

(d) Dans $\mathcal{F}(I, \mathbb{R})$ les fonctions continues forment un sous-anneau unitaire.

11.2.5 Exemple des entiers de Gauss.

On appelle entier de Gauss tout nombre complexe dont la partie réelle et la partie imaginaire sont des entiers. On note $\mathbb{Z}[i]$ leur ensemble:

$$\mathbb{Z}[i] = \{a + ib\ ;\ a, b \in \mathbb{Z}\}.$$

$\mathbb{Z}[i]$ est un anneau commutatif unitaire, contenant $\mathbb{Z}$ comme sous-anneau.

En effet, quels que soient $x = a+ib$ et $x' = c+id$ avec $a, b, c, d \in \mathbb{Z}$, les complexes $x-x' = (a-c)+i(b-d)$ et $xx' = (ac-bd)+i(ad+bc)$ ont des parties réelles et imaginaires dans $\mathbb{Z}$, donc appartiennent à $\mathbb{Z}[i]$. Ceci prouve que $\mathbb{Z}[i]$ est un sous-anneau de $\mathbb{C}$ (donc en particulier un anneau commutatif). Il est clair que $\mathbb{Z}$ est un sous-anneau de $\mathbb{Z}[i]$, d'où il résulte en particulier que $1 \in \mathbb{Z}[i]$. □

L'application $N : \mathbb{Z}[i] \to \mathbb{N}$ définie par $N(a + ib) = (a + ib)(a - ib) = a^2 + b^2$ jouera dans l'étude de l'anneau $\mathbb{Z}[i]$ un rôle important. Bornons-nous pour l'instant à observer que, puisque $N(x) = x\overline{x} = |x|^2$ pour tout $x \in \mathbb{Z}[i]$, on a clairement $N(xx') = N(x)N(x')$ pour tous $x, x' \in \mathbb{Z}[i]$.

11.2.6 Généralisation.

Soit d un entier non-nul, que l'on suppose sans facteurs carrés (c'est-à-dire que d n'est divisible par aucun carré d'entier hormis 1). On désigne par ω une racine carrée dans $\mathbb{C}$ de d. On vérifie (la preuve est laissée en exercice):

$\mathbb{Z}[\omega] = \{a + \omega b\ ;\ a, b \in \mathbb{Z}\}$ *est un anneau commutatif unitaire, contenant $\mathbb{Z}$ comme sous-anneau,*

et que l'application $N : \mathbb{Z}[\omega] \to \mathbb{Z}$ définie par $N(a + \omega b) = (a + \omega b)(a - \omega b) = a^2 - db^2$ vérifie $N(xx') = N(x)N(x')$ pour tous $x, x' \in \mathbb{Z}[\omega]$.

11.3 Morphisme d'anneaux.

11.3.1 Définitions. Soient A et B deux anneaux commutatifs unitaires. On appelle *morphisme d'anneaux unitaires* de A dans B toute application $f : A \to B$ vérifiant les trois propriétés suivantes:

$$(\ f(x+y) = f(x) + f(y) \ \text{ et } \ f(xy) = f(x)f(y) \ \text{ pour tous } x, y \in A\ ) \text{ et } (\ f(1_A) = 1_B\ ).$$

Il résulte de la première condition qu'un morphisme d'anneaux unitaires est a fortiori un morphisme de groupes additifs. Les propriétés générales des morphismes d'anneaux unitaires sont de fait analogues à celles que nous avons démontrées pour les morphismes de groupes à la leçon 2. C'est pourquoi nous synthétisons ci-dessous les plus usuelles en laissant au lecteur le soin d'adapter les démonstrations.

11.3.2 Propriétés.

(a) Si $f : A \to B$ est un morphisme d'anneaux unitaires, alors l'image directe par f de tout sous-anneau unitaire de A est un sous-anneau unitaire de B, et l'image réciproque par f de tout sous-anneau unitaire de B est un sous-anneau unitaire de A.

(b) Si $f : A \to B$ et $g : B \to C$ sont des morphismes d'anneaux unitaires, alors $g \circ f : A \to C$ est un morphisme d'anneaux unitaires.

(c) Si $f : A \to B$ est un morphisme d'anneaux unitaires bijectif, alors sa bijection réciproque $f^{-1} : B \to A$ est un morphisme d'anneaux unitaires; on dit dans ce cas que f est un *isomorphisme*, et que les deux anneaux A et B sont *isomorphes*.

11.4 Anneaux produits.

11.4.1 Proposition et définition. *Soient A_1 et A_2 deux anneaux commutatifs unitaires.*

(i) *Le produit cartésien $A_1 \times A_2 = \{(x_1, x_2),\ x_1 \in A_1,\ x_2 \in A_2\}$ est un anneau commutatif unitaire pour les lois définies par:*

$$(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2) \ \text{ et } \ (x_1, x_2).(y_1, y_2) = (x_1y_1, x_2y_2),$$

pour tous $x_1, y_1 \in A_1,\ x_2, y_2 \in A_2$, et l'on a $1_{A_1 \times A_2} = (1_{A_1}, 1_{A_2})$. Cet anneau est appelé le produit direct de A_1 par A_2. On le note $A = A_1 \times A_2$.

(ii) *L'application $p_1 : A_1 \times A_2 \to A_1$ qui, à tout élément $(x_1, x_2) \in A_1 \times A_2$, associe sa première composante x_1, est un morphisme d'anneaux unitaires (appelé première projection).*

(iii) *L'application $p_2 : A_1 \times A_2 \to A_2$ qui, à tout élément $(x_1, x_2) \in A_1 \times A_2$, associe sa seconde composante x_2, est un morphisme d'anneaux unitaires (appelé seconde projection).*

Preuve. Simple vérification, laissée au lecteur. □

11.4.2 Remarques.

(a) Le produit direct $A_1 \times A_2$ est isomorphe au produit direct $A_2 \times A_1$.

(b) On définit de même de façon évidente le produit direct d'un nombre fini quelconque d'anneaux.

11.4.3 Proposition (théorème chinois). *Soient deux entiers $n \geq 2$ et $m \geq 2$. L'anneau produit $\mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$ est isomorphe à l'anneau $\mathbb{Z}/nm\mathbb{Z}$ si et seulement si n et m sont premiers entre eux.*

Preuve. D'après le corollaire 6.6.5 (voir aussi théorème 3.6.4), on sait que, pour n et m premiers entre eux, l'application $\overline{x} \mapsto (\widetilde{x}, \widehat{x})$ réalise un isomorphisme de groupes de $\mathbb{Z}/nm\mathbb{Z}$ sur $\mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$. Il est clair, par définition même des multiplications dans ces différents anneaux, que c'est aussi un isomorphisme d'anneaux unitaires. La réciproque est évidente. □

Leçon 12

Eléments inversibles d'un anneau, corps, intégrité

CONVENTION. – Bien que les anneaux non-commutatifs interviennent dans de nombreuses situations variées et intéressantes en mathématiques, on se limitera dans la suite de ce cours (en fonction des applications visées par les programmes) à l'étude des anneaux commutatifs et unitaires.
C'est pourquoi, dans les pages qui suivent, même lorsqu'on ne le précisera pas dans les énoncés, tous les anneaux seront supposés commutatifs, unitaires, et de plus non triviaux (c'est-à-dire distincts de $\{0\}$).

12.1 Groupe des unités.

12.1.1 Définition. Soit A un anneau commutatif unitaire. On appelle *unité* de A, ou *élément inversible dans* A, tout élément $x \in A$ tel qu'il existe un élément $y \in A$ vérifiant $xy = 1$.

Remarques.

(a) Si $x \in A$ est inversible dans A, il est facile de vérifier (faites-le...) qu'il n'existe qu'un seul élément $y \in A$ tel que $xy = 1$. On note $y = x^{-1}$; on l'appelle l'inverse de x dans A.

(b) Les éléments 1 et -1 sont toujours inversibles dans A, avec $1^{-1} = 1$ et $(-1)^{-1} = -1$. L'élément 0 n'est jamais inversible (dès lors que l'anneau A n'est pas trivial, c'est-à-dire $1 \neq 0$) car on a (vérifiez-le) $0x = 0 \neq 1$ pour tout $x \in A$.

12.1.2 Proposition et définition. *Soit A un anneau commutatif unitaire. L'ensemble des éléments de A inversibles dans A est un groupe pour la multiplication, appelé groupe des unités de A, et noté $U(A)$.*

Preuve. D'après la remarque (b) ci-dessus, $U(A)$ n'est pas vide, car il contient 1. Soient x et y deux éléments de $U(A)$. Il existe x' et y' dans A tels que $xx' = 1 = yy'$. Donc $(xy)(y'x') = x(yy')x' = x1x' = xx' = 1$, ce qui prouve que $xy \in U(A)$ (et que $(xy)^{-1} = y^{-1}x^{-1}$). On a ainsi vérifié que la multiplication de A se restreint en une loi de composition interne de $U(A)$. Elle est associative, et admet comme neutre 1 qui, comme on l'a observé, appartient à $U(A)$. Il reste à vérifier que tout élément $x \in U(A)$ admet un inverse dans $U(A)$, ce qui est évident puisque l'inverse $x' = x^{-1}$ d'un élément $x \in U(A)$ est lui-même dans $U(A)$, d'inverse $(x')^{-1} = x$. □

12.1.3 Exemples.

(a) $U(\mathbb{Z}) = \{-1, 1\}$.

(b) $U(\mathbb{Z}[i]) = \{1, -1, i, -i\}$.

Preuve. Reprenons les notations de 11.2.5. Soient $x = a + ib$ et $y = c + id$ avec $a, b, c, d \in \mathbb{Z}$ tels que $xy = 1$. On a alors $1 = N(xy) = N(x)N(y)$ avec $N(x), N(y) \in \mathbb{N}^*$, d'où $N(x) = N(y) = 1$ d'après l'exemple précédent. Or $N(x) = 1$ équivaut à $a^2 + b^2 = 1$ ce qui, dans $\mathbb{Z}$, se produit si et seulement si (a, b) est l'un des quatre couples $(1, 0)$, $(-1, 0)$, $(0, 1)$ ou $(0, -1)$. □

(c) Pour tout entier $n \geq 2$, $U(\mathbb{Z}/n\mathbb{Z}) = \{\ \overline{x}\ ;\ 0 \leq x \leq n-1,$ et x premier avec $n\}$.

Preuve. Soit $\overline{x}$ un élément quelconque de $\mathbb{Z}/n\mathbb{Z}$, avec $0 \leq x \leq n-1$. On a:

$$\begin{aligned}
(\ \overline{x} \text{ inversible dans } \mathbb{Z}/n\mathbb{Z}\ ) &\Leftrightarrow (\text{ il existe } u \in \mathbb{Z} \text{ tel que } \overline{x}\,\overline{u} = \overline{1}\ ) \\
&\Leftrightarrow (\text{ il existe } u \in \mathbb{Z} \text{ tel que } \overline{xu - 1} = \overline{0}\ ) \\
&\Leftrightarrow (\text{ il existe } u, v \in \mathbb{Z} \text{ tels que } xu - 1 = nv\ ) \\
&\Leftrightarrow (\text{ il existe } u, v \in \mathbb{Z} \text{ tels que } xu + n(-v) = 1\ )
\end{aligned}$$

d'où le résultat par le théorème de Bézout dans $\mathbb{Z}$. □

Remarquons que les éléments $\overline{x}$ qui sont inversibles dans l'anneau $\mathbb{Z}/n\mathbb{Z}$ sont aussi, d'après le point (iii) du théorème 6.6.4, ceux qui engendrent le groupe additif $\mathbb{Z}/n\mathbb{Z}$. En particulier le groupe $U(\mathbb{Z}/n\mathbb{Z})$ est fini d'ordre $\varphi(n)$, (où φ est l'indicatrice d'Euler).

12.2 Corps.

12.2.1 Définition. On appelle *corps commutatif* (ou plus simplement corps) tout anneau commutatif unitaire dans lequel tout élément non-nul est inversible.
En notant, pour tout anneau A commutatif unitaire $A^* = A \setminus \{0\}$, on a donc:

$$(A \text{ corps }) \Leftrightarrow (U(A) = A^*)$$

12.2.2 Définition. Soit K un corps. On appelle *sous-corps* de K tout sous-anneau unitaire F de K tel que l'inverse de tout élément non-nul de F appartient à F.

12.2.3 Exemples.

(a) $\mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$ sont des corps; ils contiennent comme sous-anneau $\mathbb{Z}$ qui, lui, n'est pas un corps.

(b) $\mathbb{Q}(i) = \{p + qi \ ; \ p, q \in \mathbb{Q}\}$ est un sous-corps de $\mathbb{C}$; il contient $\mathbb{Z}[i]$ comme sous-anneau qui, lui, n'est pas un corps.

> *Preuve.* On vérifie aisément que $\mathbb{Q}(i)$ est un sous-anneau de $\mathbb{C}$; pour tout $x = p+qi \in \mathbb{Q}(i)$ non-nul, son inverse x^{-1} dans $\mathbb{C}$ est égal à $\frac{p}{p^2+q^2} + \frac{-q}{p^2+q^2}i$ et appartient donc à $\mathbb{Q}(i)$. Ce qui prouve que $\mathbb{Q}(i)$ est un sous-corps de $\mathbb{C}$. Il est clair que $\mathbb{Z}[i]$ est un sous-anneau de $\mathbb{Q}(i)$, et le fait que ce n'est pas un corps découle immédiatement de 12.1.3.(b). □

(c) Pour tout entier $n \geq 2$, $(\mathbb{Z}/n\mathbb{Z}$ est un corps $) \Leftrightarrow (n$ est un nombre premier $)$.

> *Preuve.* Résulte immédiatement de 12.1.3.(c). □

12.3 Intégrité.

12.3.1 Définition. Soit A un anneau commutatif. On dit que A est *intègre*, ou encore que A est un *domaine d'intégrité*, lorsqu'il est non-nul et vérifie la propriété suivante:

$$\text{pour tous } x, y \in A, \ (xy = 0) \Leftrightarrow (x = 0 \ \text{ ou } \ y = 0).$$

> Un élément x de A est appelé un *diviseur de zéro* dans A lorsque $x \neq 0$ et lorsque qu'il existe $y \neq 0$ dans A tel que $xy = 0$. Donc A est intègre si et seulement s'il n'admet pas de diviseurs de zéro.

12.3.2 Premiers exemples et contre-exemples.

(a) Tout corps est un anneau intègre.

> *Preuve.* Soit K un corps. Soient $x, y \in K$ tels que $xy = 0$. Si $x \neq 0$, alors x est inversible dans K par définition d'un corps. Donc $x^{-1}xy = x^{-1}0$, c'est-à-dire $y = 0$. De même $y \neq 0$ implique $x = 0$. En résumé l'un au moins des deux facteurs x et y est nul. □

(b) Tout sous-anneau d'un anneau intègre est intègre. En particulier tout sous-anneau d'un corps est intègre. Par exemple, $\mathbb{Z}$ et $\mathbb{Z}[i]$ sont intègres (bien que ce ne soient pas des corps).

(c) Attention: un anneau produit $A_1 \times A_2$ n'est pas intègre (même si A_1 et A_2 le sont). *En effet,* les éléments $(1_{A_1}, 0_{A_2})$ et $(0_{A_1}, 1_{A_2})$ sont non-nuls, alors que leur produit l'est.

(d) Considérons les tables de multiplication des anneaux $\mathbb{Z}/5\mathbb{Z}$ et $\mathbb{Z}/6\mathbb{Z}$.

	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$
$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$
$\overline{1}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$
$\overline{2}$	$\overline{0}$	$\overline{2}$	$\overline{4}$	$\overline{1}$	$\overline{3}$
$\overline{3}$	$\overline{0}$	$\overline{3}$	$\overline{1}$	$\overline{4}$	$\overline{2}$
$\overline{4}$	$\overline{0}$	$\overline{4}$	$\overline{3}$	$\overline{2}$	$\overline{1}$

	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$
$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$
$\overline{1}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$
$\overline{2}$	$\overline{0}$	$\overline{2}$	$\overline{4}$	$\overline{0}$	$\overline{2}$	$\overline{4}$
$\overline{3}$	$\overline{0}$	$\overline{3}$	$\overline{0}$	$\overline{3}$	$\overline{0}$	$\overline{3}$
$\overline{4}$	$\overline{0}$	$\overline{4}$	$\overline{2}$	$\overline{0}$	$\overline{4}$	$\overline{2}$
$\overline{5}$	$\overline{0}$	$\overline{5}$	$\overline{4}$	$\overline{3}$	$\overline{2}$	$\overline{1}$

L'anneau $\mathbb{Z}/5\mathbb{Z}$ est un corps puisque 5 est un nombre premier ; il est donc a fortiori intègre. Au contraire, l'anneau $\mathbb{Z}/6\mathbb{Z}$ n'est pas intègre car, par exemple, $\overline{2}.\overline{3} = \overline{0}$ bien que $\overline{2} \neq \overline{0}$ et $\overline{3} \neq \overline{0}$; a fortiori, ce n'est pas un corps. Ces exemples sont des cas particuliers de la proposition suivante.

12.3.3 Proposition (cas des anneaux $\mathbb{Z}/n\mathbb{Z}$). *Pour tout entier $n \geq 2$, on a:*

(l'anneau $\mathbb{Z}/n\mathbb{Z}$ est intègre) $\Leftrightarrow$ *(n est un nombre premier)* $\Leftrightarrow$ *(l'anneau $\mathbb{Z}/n\mathbb{Z}$ est un corps)*

Preuve. D'après 12.2.3.(c) et 12.3.2.(a), le seul point à montrer est que $\mathbb{Z}/n\mathbb{Z}$ intègre implique n premier. Par contraposée, supposons que n n'est pas premier; il existe donc $k, m \in \mathbb{Z}$ tels que $n = km$ avec $1 < k < n$ et $1 < m < n$. On a alors $\overline{k}.\overline{m} = \overline{n} = \overline{0}$, bien que $\overline{k} \neq \overline{0}$ et $\overline{m} \neq \overline{0}$. □

12.3.4 Proposition (cas des anneaux de polynômes). *Soit A un anneau commutatif unitaire.*

(i) *Si A est intègre, alors pour tous polynômes $P, Q \in A[X]$, on a:*
$$\deg(PQ) = \deg P + \deg Q \quad \text{et} \quad \mathrm{cd}(PQ) = \mathrm{cd}(P)\,\mathrm{cd}(Q)$$

(ii) *$A[X]$ est intègre si et seulement si A est intègre.*

(iii) *En particulier, si K est un corps, alors l'anneau $K[X]$ est intègre.*

Preuve. Les égalités $\deg(PQ) = \deg P + \deg Q$ et $\mathrm{cd}(PQ) = \mathrm{cd}(P)\,\mathrm{cd}(Q)$ sont claires si P ou Q est nul. Supposons-les tous les deux non-nuls, et écrivons $P = a_nX^n + \cdots + a_1X + a_0$ et $Q = b_mX^m + \cdots + b_1X + b_0$, avec $\mathrm{cd}(P) = a_n \neq 0$ et $\mathrm{cd}(Q) = b_m \neq 0$. Alors:
$$PQ = a_nb_mX^{n+m} + (a_nb_{m-1} + a_{n-1}b_m)X^{n+m-1} + \cdots + (a_1b_0 + a_0b_1)X + a_0b_0.$$
L'intégrité de A implique $a_nb_m \neq 0$, donc $\mathrm{cd}(PQ) = a_nb_m$, d'où $\deg(PQ) = n + m$, ce qui prouve (i). Il résulte immédiatement de (i) que, si A est intègre, le produit de deux éléments non-nuls de $A[X]$ est non-nul, ce qui prouve que $A[X]$ est intègre. L'implication réciproque étant triviale d'après 12.3.2.(b), le point (ii) est établi. Le point (iii) en découle d'après 12.3.2.(a). □

12.3.5 Corollaire (groupe des unités des anneaux de polynômes). *Soit A un anneau commutatif unitaire. Si A est intègre, alors: $U(A[X]) = U(A)$.*

Preuve. L'inclusion $U(A) \subset U(A[X])$ est claire puisque A est un sous-anneau unitaire de $A[X]$. Pour la réciproque, considérons $P \in U(A[X])$. Il existe donc $Q \in A[X]$ tel que $PQ = 1$. Ces deux polynômes sont nécessairement non-nuls, donc il résulte du point (i) de la proposition précédente que $\deg P + \deg Q = 0$. On en tire $\deg P = \deg Q = 0$, c'est-à-dire $P \in A$ et $Q \in A$, et donc l'égalité $PQ = 1$ implique $P \in U(A)$ et $Q \in U(A)$. □

Remarque. $A[X]$ n'est jamais un corps.

En effet, que A soit ou non intègre, l'élément X de $A[X]$ vérifie $\deg PX = \deg P + 1$ pour tout $P \in A[X]$, de sorte que l'on ne peut pas avoir $PX = 1$, ce qui montre que X n'est jamais inversible. □

12.4 Corps des fractions d'un anneau intègre.

12.4.1 Construction. Il existe, on l'a vu, des anneaux intègres qui ne sont pas des corps. Le but de ce qui suit est de montrer que, néanmoins, on peut construire de façon canonique pour tout anneau intègre A un corps K qui le contient, et qui est (en un sens que l'on précisera) le plus petit corps qui le contient. Evidemment, la question ne se pose pas pour des anneaux non intègres [d'après les remarques 12.3.2.(a) et 12.3.2.(b)].

Fixons A un anneau commutatif unitaire intègre. Posons $A^* = A \setminus \{0\}$. On définit dans $A \times A^*$ la relation $\sim$ par:
$$(a, b) \sim (c, d) \;\Leftrightarrow\; ad = bc.$$

Etape 1: *la relation $\sim$ est une relation d'équivalence dans $A \times A^*$.*

Preuve. La réflexivité et la symétrie sont évidentes. Pour la transitivité, considérons trois couples (a,b), (c,d) et (e,f) dans $A \times A^*$. Supposons que $(a,b) \sim (c,d)$ et $(c,d) \sim (e,f)$. On a donc: $ad = bc$ et $cf = de$. Il vient $adf = bcf = bde$, et comme $d \neq 0$, l'intégrité de A implique $af = be$, d'où $(a,b) \sim (e,f)$. □

Pour tout couple $(a,b) \in A \times A^*$, on note $\frac{a}{b}$ la classe d'équivalence de (a,b) pour la relation $\sim$:

$$\tfrac{a}{b} = \{(c,d) \in A \times A^* \,;\, (c,d) \sim (a,b)\} = \{(c,d) \in A \times A^* \,;\, ad = bc\}.$$

Une telle classe s'appelle une fraction. On note $K = (A \times A^*)/\sim$ l'ensemble quotient de $A \times A^*$ par la relation $\sim$, c'est-à-dire l'ensemble des fractions. Tout couple (c,d) appartenant à $\frac{a}{b}$ s'appelle un représentant de la fraction $\frac{a}{b}$. On a:

$$(\ \tfrac{a}{b} = \tfrac{c}{d} \text{ dans } K\ ) \ \Leftrightarrow\ (\ (a,b) \sim (c,d) \text{ dans } A \times A^*\ ) \ \Leftrightarrow\ (\ ad = bc \text{ dans } A\ ).$$

Etape 2: *L'application $\phi : A \to K$ qui, à un élément $a \in A$ associe la fraction $\phi(a) = \frac{a}{1}$, est injective, et est appelée injection canonique de A dans K.*

Preuve. Soient $a, c \in A$ tels que $\phi(a) = \phi(c)$. Alors $\frac{a}{1} = \frac{c}{1}$, d'où $a.1 = 1.c$, donc $a = c$. □

On convient d'identifier A avec le sous-ensemble $\phi(A)$ de K, qui lui est équipotent. Via cette identification, A est un sous-ensemble de K, et on pose $a = \frac{a}{1}$, pour tout $a \in A$. En d'autres termes:

$$\text{quel que soit } a \in A, \text{ on a: } \ a = \tfrac{a}{1} = \{(c,d) \in A \times A^* \,;\, c = ad\} = \tfrac{ad}{d} \ \text{ pour tout } d \in A^*.$$

En particulier: $0 = \frac{0}{1} = \frac{0}{b}$ pour tout $b \in A^*$ et $1 = \frac{1}{1} = \frac{b}{b}$ pour tout $b \in A^*$.

Etape 3: *Les lois de composition internes dans K définies par:*

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd} \qquad \textit{et} \qquad \frac{a}{b} \times \frac{c}{d} = \frac{ac}{bd}$$

sont bien définies (indépendamment des représentants choisis), munissent K d'une structure d'anneau commutatif unitaire, et prolongent celles de A (ce qui signifie que l'injection canonique est un morphisme d'anneaux unitaires, ou encore que A peut être considéré, en l'identifiant avec son image par ϕ, comme un sous-anneau unitaire de K).

Preuve. Supposons que $\frac{a}{b} = \frac{a'}{b'}$ et $\frac{c}{d} = \frac{c'}{d'}$. Un calcul évident montre que $ab' = a'b$ et $cd' = c'd$ impliquent:

- d'une part: $(ad + bc)b'd' = (a'd' + b'c')bd$, et donc $\frac{ad+bc}{bd} = \frac{a'd'+b'c'}{b'd'}$,
- d'autre part: $(ac)(b'd') = (a'c')(bd)$, et donc $\frac{ac}{bd} = \frac{a'c'}{b'd'}$.

Ce qui prouve que les deux lois sont bien définies. Qu'elles satisfont alors tous les axiomes de la structure d'anneau commutatif unitaire (avec $0 = \frac{0}{1}$ pour neutre additif et $1 = \frac{1}{1}$ pour neutre multiplicatif) est une simple vérification, qu'on laisse au lecteur. Enfin quels que soient deux éléments $a, c \in A$, on a:

$$\phi(a + c) = \tfrac{a+c}{1} = \tfrac{a}{1} + \tfrac{c}{1} = \phi(a) + \phi(c) \quad \text{ et } \quad \phi(ac) = \tfrac{ac}{1} = \tfrac{a}{1}.\tfrac{c}{1} = \phi(a).\phi(c),$$

ce qui achève la preuve. □

Etape 4: *Tout élément non-nul de K est inversible dans K. Plus précisément, tout élément $\frac{a}{b} \in K$ avec $(a,b) \in A^* \times A^*$ admet $\frac{b}{a}$ pour inverse.*

Preuve. Evident puisque $\frac{a}{b}.\frac{b}{a} = \frac{ab}{ba} = 1$. □

En particulier, tout élément non-nul $a \in A$ admet dans K l'inverse $\frac{1}{a}$.

On déduit de cette construction et des vérifications faites aux différentes étapes le théorème suivant.

12.4.2 Théorème. *Soit A un anneau commutatif unitaire intègre.*

(i) *L'ensemble $K = (A \times A^*)/\sim$ des fractions sur A, muni des lois construites ci-dessus, est un corps commutatif, qui contient A comme sous-anneau unitaire.*

(ii) *Si K' est un sous-corps tel que $A \subseteq K' \subseteq K$, alors $K' = K$.*

Preuve. Le (i) a été montré en 12.4.1. Pour le (ii), supposons que $A \subseteq K' \subseteq K$ avec K' un corps. Soit $x \in K$. Par définition, il existe $a \in A$ et $b \in A^*$ tel que $x = \frac{a}{b} = a.\frac{1}{b}$. On a $b \in A$ donc $b \in K'$, avec $b \neq 0$; comme l'inverse de b dans K est $\frac{1}{b} \in K$, et que cet inverse doit appartenir à K' puisque K' est un sous-corps, on a $\frac{1}{b} \in K'$. Par ailleurs $a \in A$ donc $a \in K'$. Le sous-corps K' est stable par produit, donc $a.\frac{1}{b} \in K'$, c'est-à-dire $\frac{a}{b} = x \in K'$. Cela prouve que $K \subseteq K'$, donc $K = K'$. □

12.4.3 Exemples. Deux exemples classiques ont déjà été rencontrés lors des années précédentes:

(a) Le corps de fractions de l'anneau intègre $\mathbb{Z}$ est appelé corps des rationnels et est noté $\mathbb{Q}$.

(b) Le corps de fractions de l'anneau intègre de polynômes $\mathbb{R}[X]$ est appelé corps des fractions rationnelles à coefficients réels, et est noté $\mathbb{R}(X)$.

Plus généralement, pour tout anneau intègre A, le corps de fractions de l'anneau intègre $A[X]$ (voir 12.3.4) est appelé corps des fractions rationnelles à coefficients dans A. Ses éléments sont de la forme: $F(X) = \frac{P(X)}{Q(X)}$ avec $P, Q \in A[X]$, $Q \neq 0$.

A titre d'exercice, montrer que le corps de fractions de $\mathbb{Z}[i]$ est $\mathbb{Q}(i) = \{p + qi\,;\, p \in \mathbb{Q}, q \in \mathbb{Q}\}$.

Leçon 13

Idéal d'un anneau

13.1 Notion d'idéal.

13.1.1 Définition. Soit A un anneau commutatif unitaire. On appelle *idéal* de A toute partie non-vide I de A qui vérifie les deux conditions suivantes:

(1) I est un sous-groupe du groupe additif A,

(2) pour tous $x \in I$ et $a \in A$, on a $xa \in I$.

Exemples.

(a) $\{0\}$ et A sont des idéaux de A.

(b) Pour tout $n \in \mathbb{Z}$, l'ensemble $n\mathbb{Z}$ des multiples de n est un idéal de l'anneau $\mathbb{Z}$.

(c) Dans l'anneau $\mathcal{F}(\mathbb{R}, \mathbb{R})$, l'ensemble des fonctions qui s'annulent en 0 est un idéal.

13.1.2 Lemme (très utile dans la pratique). *Soit A un anneau commutatif unitaire.*

(i) *si I est un idéal de A qui contient 1, alors $I = A$.*

(ii) *si I est un idéal de A qui contient un élément de $U(A)$, alors $I = A$.*

Preuve. Supposons $1 \in I$. Tout $a \in A$ s'écrit $a = 1.a$ donc, comme $1 \in I$, il résulte de la propriété (2) que $a \in I$. On a alors $A \subseteq I$, donc $A = I$, ce qui prouve (i). Supposons maintenant que I contienne un élément x inversible dans A. On a $1 = xx^{-1}$ avec $x \in I$ et $x^{-1} \in A$, donc $1 \in I$, et on applique (i) pour conclure que $I = A$. □

13.1.3 Proposition. *Soient A et B des anneaux commutatifs unitaires. Soit $f : A \to B$ un morphisme d'anneaux unitaires. On a:*

(i) *Pour tout idéal J de B, l'image réciproque $f^{-1}(J)$ est un idéal de A.*

(ii) *En particulier,* $\operatorname{Ker} f = \{x \in A\,;\, f(x) = 0_B\}$ *est un idéal de A.*

(iii) *Pour tout idéal I de A, l'image directe $f(I)$ est un idéal de l'anneau $f(A) = \operatorname{Im} f$; (attention, ce n'est pas en général un idéal de B).*

Preuve. Sous les hypothèses de (i), on sait déjà que $f^{-1}(J)$ est un sous-groupe additif de A (voir 2.1.4). Soit $x \in f^{-1}(J)$ et $a \in A$. On a $f(xa) = f(x)f(a)$ avec $f(a) \in B$ et $f(x) \in J$, donc $f(xa) \in J$ puisque J est un idéal de B, c'est-à-dire $xa \in f^{-1}(J)$, ce qui prouve que $f^{-1}(J)$ est un idéal de A. On obtient (ii) en appliquant ce qui précède à $J = \{0_B\}$.

Pour (iii), considérons un idéal I de A. On sait que $f(I)$ est un sous-groupe additif de B (voir 2.1.4). Soit $y \in f(I)$, de sorte qu'il existe $x \in I$ tel que $y = f(x)$. Pour tout élément $b \in B$ qui appartient à $\operatorname{Im} f$, il existe $a \in A$ tel que $b = f(a)$; on a alors $yb = f(a)f(x) = f(ax)$ avec $ax \in I$ puisque $x \in I$ et que I est un idéal, et donc $yb \in f(I)$. Ceci prouve que $f(I)$ est un idéal de l'anneau $\operatorname{Im} f$. □

13.1.4 Proposition. *Soit A un anneau commutatif unitaire. L'intersection de deux idéaux de A est un idéal de A. Plus généralement, l'intersection d'une famille quelconque d'idéaux de A est un idéal de A.*

Preuve. Il suffit de montrer le second point. Soit donc $(I_j)_{j\in X}$ une famille d'idéaux de A. Posons $I = \bigcap_{j\in X} I_j$ l'intersection de tous les I_j. On sait déjà que I est un sous-groupe additif (proposition 1.2.6). Soient $x \in I$ et $a \in A$. On a $xa \in I_j$ pour tout $j \in X$ puisque I_j est un idéal, et donc $xa \in I$. Ce qui prouve que I est un idéal de A. □

L'intérêt majeur de la notion d'idéal est qu'elle permet de construire des anneaux quotients (voir leçon suivante). Donc, de même que la description des sous-groupes normaux d'un groupe donné est une question cruciale de la théorie des groupes, la question naturelle se pose de déterminer lorsque c'est possible les idéaux d'un anneau donné. En amont des procédés standards permettant de déduire de nouveaux idéaux à partir d'idéaux connus (voir 13.1.3, 13.1.4, 13.3.1, 13.4.1...), le type d'idéal le plus simple que l'on peut toujours considérer est décrit au paragraphe suivant.

13.2 Idéal principal, anneau principal.

13.2.1 Proposition et définition. *Soit A un anneau commutatif unitaire. Pour $x \in A$, on note:*

$$xA = \{xy\,;\, y \in A\} = \{z \in A\,;\, \text{il existe } y \in A \text{ tel que } z = xy\}.$$

Alors:

(i) *xA est un idéal de A, appelé l'idéal principal engendré par x;*

(ii) *xA est le plus petit idéal de A contenant x;*

(iii) *on a: ($xA = A$) $\Leftrightarrow$ ($x \in U(A)$).*

Preuve. Il est clair que xA est non-vide (il contient x puisque $x = x.1$). Soient $y \in xA$ et $z \in xA$ quelconques; il existe $a, b \in A$ tels que $y = xa$ et $z = xb$, donc $y - z = x(a - b) \in xA$, ce qui prouve que xA est un sous-groupe additif. Soient $y \in xA$ et $c \in A$ quelconques; il existe $a \in A$ tel que $y = xa$, donc $yc = xac = x(ac) \in xA$. On conclut que xA est un idéal de A.

Soit I un idéal de A contenant x. Comme $x \in I$, on a $xa \in I$ pour tout $a \in A$. Donc $xA \subseteq I$, d'où (ii).

Si $xA = A$, alors $1 \in xA$, de sorte qu'il existe $y \in A$ tel que $xy = 1$, ce qui prouve $x \in U(A)$. L'implication réciproque découle de 13.1.2.(ii). □

Bien que très simple, le corollaire suivant est important, et montre que la notion d'idéal n'a d'intérêt que pour des anneaux qui ne sont pas des corps.

13.2.2 Corollaire. *Soit A un anneau commutatif unitaire.*

$$(\ A \text{ est un corps}\ ) \Leftrightarrow (\ \text{les seuls idéaux de } A \text{ sont } \{0\} \text{ et } A\ ).$$

Preuve. Supposons que A est un corps. Soit I un idéal de A. Si $I \neq \{0\}$, il existe dans I un élément non-nul, donc inversible dans A puisque A est un corps. On conclut avec 13.1.2.(ii) que $I = A$. Supposons réciproquement que A n'admette que $\{0\}$ et A comme idéaux. Soit $x \in A$ quelconque non-nul. L'idéal xA étant alors distinct de $\{0\}$, on a nécessairement $xA = A$, d'où $x \in U(A)$ d'après 13.2.1.(iii). Ainsi tout élément non-nul de A est inversible dans A: on conclut que A est un corps. □

13.2.3 Définitions. On appelle *idéal principal* d'un anneau commutatif unitaire A tout idéal I de A pour lequel il existe un élément $x \in A$ tel que $I = xA$. On appelle *anneau principal* tout anneau commutatif unitaire A qui est intègre et dans lequel tout idéal est principal.
De nombreux exemples d'anneaux principaux seront donnés à la leçon 15. Bornons-nous ici à citer:

$$\mathbb{Z} \text{ est un anneau principal,}$$

ce qui découle immédiatement du lemme suivant:

13.2.4 Lemme (idéaux de $\mathbb{Z}$). *Pour tout idéal I de l'anneau $\mathbb{Z}$, il existe un unique $k \in \mathbb{N}$ tel que $I = k\mathbb{Z}$. Les seuls entiers m tels que $I = m\mathbb{Z}$ sont alors $m = k$ et $m = -k$.*

Preuve. Comme un idéal est en particulier un sous-groupe additif, cela résulte de la proposition 6.6.2.□

13.2.5 Contre-exemple. *L'anneau $\mathbb{Z}[X]$ n'est pas principal.*

Preuve. On le montre de façon élémentaire en vérifiant que, par exemple, dans $A = \mathbb{Z}[X]$, l'idéal $I = 2A + XA$ (qui n'est autre que l'idéal engendré par 2 et X) n'est pas un idéal principal.

Par l'absurde, supposons qu'il existe $P \in A$ tel que $I = PA$. Comme $2 \in I$, il existerait $Q \in A$ tel que $2 = PQ$, ce qui impliquerait par un raisonnement sur les degrés que $P \in \mathbb{Z}$. Comme de plus $X \in I$, il existerait $R \in A$ tel que $X = PR$, ce qui impliquerait $P = \pm 1$ (et $R = \pm X$). On aurait donc $1 = \pm P \in I$, de sorte qu'il existerait $S, T \in A$ tels que $1 = 2S + TX$, ce qui est clairement impossible dans $A = \mathbb{Z}[X]$, puisque le coefficient constant de $2S + TX$ est pair. □

13.3 Idéal engendré par une partie, somme d'idéaux.

13.3.1 Proposition et définition. *Soit A un anneau commutatif unitaire.*

(i) *Si I et J sont des idéaux de A, alors l'ensemble $I+J=\{x+y\,;\,x\in I, y\in J\}$ est un idéal de A, appelé l'idéal somme de I et J, et c'est le plus petit idéal contenant I et J;*

(ii) *En particulier, si x et y sont des éléments de A, l'ensemble $xA+yA=\{xa+yb\,;\,a,b\in A\}$ est le plus petit idéal de A contenant x et y.*

Preuve. Soient I et J deux idéaux de A. Il est clair que $I+J$ est un sous-groupe additif de A (c'est le sous-groupe engendré par $I\cup J$). Soit $z\in I+J$ et $a\in A$ quelconques; il existe $x\in I$ et $y\in J$ tels que $z=x+y$, d'où $za=xa+ya$. Or $xa\in I$ car $x\in I$ et I est un idéal; de même $ya\in J$. On conclut que $za\in I+J$, ce qui prouve que $I+J$ est un idéal de A. Il est clair que $I\subseteq I+J$, puisque tout $x\in I$ s'écrit $x=x+0$ avec $0\in J$; de même $J\subseteq I+J$. Pour montrer que c'est le plus petit, supposons que K est un idéal de A contenant I et J. En particulier, K est stable par addition, et donc, quels que soient $x\in I\subseteq K$ et $y\in J\subseteq K$, on a $x+y\in K$. Donc $I+J\subseteq K$. Ce qui achève de prouver (i). Le point (ii) s'en déduit avec $I=xA$ et $J=yA$. □

13.3.2 Remarques. Soit A un anneau commutatif unitaire.

(a) On définit généralement, pour toute partie non-vide X de A, l'idéal engendré par X comme l'intersection de tous les idéaux contenant X; c'est le plus petit idéal de A contenant X.

La proposition 13.2.1 correspond à $X=\{x\}$, le point (i) de 13.3.1 à $X=I\cup J$, et le point (ii) de 13.3.1 à $X=\{x,y\}$.

(b) L'intérêt de la notion d'idéal somme réside bien sûr dans le fait que la réunion de deux idéaux n'est en général pas un idéal (ce n'est pas en général un sous-groupe additif; prendre par exemple $A=\mathbb{Z}$, $I=2\mathbb{Z}$ et $J=3\mathbb{Z}$).

13.4 Produit d'idéaux, opérations sur les idéaux.

13.4.1 Définition et proposition. *Soit A un anneau commutatif unitaire. Si I et J sont des idéaux de A, on appelle produit des idéaux I et J, et on note IJ, l'ensemble des éléments de A qui sont somme d'un nombre fini de produits d'un élément de I par un élément de J.*

$$(\ x\in IJ\ )\Leftrightarrow(\ \text{il existe } n\in\mathbb{N}^*,\ y_1,\dots,y_n\in I,\ z_1,\dots,z_n\in J,\ \text{tels que } x=\sum_{i=1}^{n}y_iz_i\ ).$$

Alors IJ est un idéal de A ; c'est le plus petit idéal contenant l'ensemble $\{yz\,;\,y\in I, z\in J\}$, et il vérifie: $IJ\subset I\cap J$.

Preuve. Il est clair que IJ est un sous-groupe additif de A. Soit $x=\sum_{i=1}^{n}y_iz_i$ un élément quelconque de IJ, avec $y_1,\dots,y_n\in I$ et $z_1,\dots,z_n\in J$. Pour tout $a\in A$, on a $ay_i\in I$ quel que soit $1\le i\le n$, donc $ax=\sum_{i=1}^{n}(ay_i)z_i$ appartient encore à IJ. Ceci prouve que IJ est un idéal. Il est clair qu'il contient $X=\{yz\,;\,y\in I, z\in J\}$. Soit maintenant K un idéal qui contient X. Il contient aussi les sommes d'éléments de X, et donc $IJ\subseteq K$. Ceci s'applique en particulier à $K=I\cap J$, qui contient bien X. □

13.4.2 Proposition. *Soit A un anneau commutatif unitaire. Si I, J et K sont des idéaux de A, on a:*

$$I+(J+K)=(I+J)+K,\qquad I(JK)=(IJ)K,\qquad I(J+K)=IJ+IK.$$

13.5 Caractéristique d'un anneau.

13.5.1 Remarques préliminaires.

(a) Soit A un anneau commutatif unitaire. Pour tout $x\in A$, on note $2x=x+x$, $3x=x+x+x$ et de même $nx=x+x+\cdots+x$ (avec n termes) pour tout entier $n\ge 2$. On pose naturellement $1x=x$ et $0x=0$, ce qui définit la notation nx pour tout $n\in\mathbb{N}$. Si l'on considère maintenant un entier $m\le 0$, on convient que $mx=n(-x)=-(nx)$ où $n=-m\in\mathbb{N}$. On a ainsi défini la notation nx pour tout $x\in A$ et tout $n\in\mathbb{Z}$.

(b) Soit A un anneau commutatif unitaire. On vérifie aisément que, pour tout $n \in \mathbb{Z}$, on a:

$$(\ n1_A = 0_A \ ) \Leftrightarrow (\ nx = 0_A \ \text{ pour tout } x \in A \ ).$$

13.5.2 Lemme et définition. *Soit A un anneau commutatif unitaire. Il existe un unique morphisme d'anneaux unitaires $f : \mathbb{Z} \to A$. Il est défini par $f(n) = n1_A$ pour tout $n \in \mathbb{Z}$. On l'appelle le morphisme canonique de $\mathbb{Z}$ dans A.*

Preuve. Si f est un morphisme d'anneaux unitaires $\mathbb{Z} \to A$, on doit avoir $f(1) = 1_A$, d'où par additivité $f(2) = f(1) + f(1) = 1_A + 1_A = 2\,1_A$, et par récurrence $f(n) = n1_A$ pour tout entier $n \geq 1$. Comme f est un morphisme de groupes additifs, on a aussi $f(0) = 0_A$ et $f(m) = f(-n) = -f(n) = -(n1_A) = (-n)1_A = m1_A$ pour tout entier $m \leq 0$ et en posant $n = -m$. En résumé, on a $f(n) = n1_A$ pour tout $n \in \mathbb{Z}$. Réciproquement, il est facile de vérifier (faites-le) que f ainsi défini est bien un morphisme d'anneaux unitaires. □

13.5.3 Définition. Soit A un anneau commutatif unitaire. On appelle *caractéristique* de A, notée $\operatorname{car} A$, l'unique entier $k \in \mathbb{N}$ tel que $\operatorname{Ker} f = k\mathbb{Z}$, où f est le morphisme canonique de $\mathbb{Z}$ dans A.

Comme $f : \mathbb{Z} \to A$ est un morphisme d'anneaux unitaires d'après 13.5.2, $\operatorname{Ker} f$ est un idéal de $\mathbb{Z}$ d'après 13.1.3.(ii), et il est donc de la forme $k\mathbb{Z}$ pour un unique $k \in \mathbb{N}$ d'après 13.2.4.

Grâce à la remarque 13.5.1.(b), cette définition se traduit par:

$$\operatorname{car} A = 0 \ \Leftrightarrow \ \Big[\ (nx = 0_A \text{ pour tout } \ x \in A \ ) \Leftrightarrow (n = 0) \ \Big]$$

$$\operatorname{car} A = k > 0 \ \Leftrightarrow \ \Big[\ (nx = 0_A \text{ pour tout } \ x \in A \ ) \Leftrightarrow (n \in k\mathbb{Z}) \ \Big]$$

13.5.4 Exemples.

(a) L'anneau $\mathbb{Z}$ est de caractéristique nulle, ainsi que les corps $\mathbb{Q}, \mathbb{R}, \mathbb{C}$.

(b) Pour tout $n \geq 2$, l'anneau $\mathbb{Z}/n\mathbb{Z}$ est de caractéristique n. En particulier, pour tout nombre premier p, le corps $\mathbb{Z}/p\mathbb{Z}$ est de caractéristique p.

(c) Soit A un anneau commutatif unitaire. Pour tout sous-anneau unitaire B de A, on a:

$$\operatorname{car} A = \operatorname{car} B.$$

Leçon 14

Anneaux quotients

14.1 Quotient d'un anneau par un idéal.

14.1.1 Remarques préliminaires. Soit A un anneau commutatif unitaire. Soit I un idéal de A.

(a) L'idéal I est en particulier un sous-groupe du groupe additif A, et il est trivialement normal puisque A est abélien. On peut considérer le groupe additif quotient A/I. Rappelons que, si l'on note $\overline{a}$ la classe dans A/I d'un élément a de A, on a par définition:

$$\overline{a} = \{b \in A\,;\, a - b \in I\} := a + I,$$

et que l'addition dans A/I est définie par:

$$\overline{a} + \overline{b} = \overline{a+b} \quad \text{pour tous } a, b \in A,$$

d'où en particulier A/I abélien, de neutre additif $\overline{0} = I$. La surjection canonique $p : A \to A/I$, qui à tout élément a de A associe sa classe $\overline{a}$ est alors un morphisme de groupes pour l'addition.

(b) On définit dans A/I une multiplication en posant:

$$\overline{a}.\overline{b} = \overline{ab} \quad \text{pour tous } a, b \in A,$$

1. Elle est bien définie, indépendamment des représentants choisis.

   En effet. Soient $x' \in \overline{x}$ et $y' \in \overline{y}$. Alors $x' - x \in I$ et $y' - y \in I$.
   On a: $x'y' - xy = x'(y' - y) + (x' - x)y$.
   Comme $x' - x \in I$ et que I est un idéal, on a $(x' - x)y \in I$; de même $x'(y' - y) \in I$ puisque $y' - y \in I$. On conclut que $x'y' - xy \in I$ comme somme de deux éléments de I, et donc $\overline{x'y'} = \overline{xy}$.

2. Elle est associative, commutative, distributive sur l'addition dans A/I, et admet $\overline{1}$ comme élément neutre.

   En effet. Quels que soient $x, y, z \in A$, on a $(\overline{x}.\overline{y}).\overline{z} = \overline{(xy)z} = \overline{x(yz)} = \overline{x}.(\overline{y}.\overline{z})$, ce qui montre l'associativité. Le reste se montre de même. □

3. La surjection canonique p vérifie $p(1) = \overline{1}$ et $p(xy) = p(x).p(y)$ pour tous $x, y \in A$.

   En effet. Par définition de p d'une part, et de la multiplication dans A/I d'autre part, on a $p(xy) = \overline{xy} = \overline{x}.\overline{y} = p(x).p(y)$. □

On a ainsi démontré:

14.1.2 Théorème. *Soit A un anneau commutatif unitaire. Pour tout idéal I de A, le quotient A/I est un anneau commutatif unitaire, et la surjection canonique $p : A \to A/I$ est un morphisme d'anneaux unitaires.*

On a pour les anneaux quotients des résultats de même nature que ceux que l'on a montrés à la leçon 6 pour les groupes quotients, en particulier:

14.1.3 Théorème (dit premier théorème d'isomorphisme). *Soient A et A' deux anneaux commutatifs unitaires, et $f : A \to A'$ un morphisme d'anneaux unitaires. Alors l'anneau quotient de A par l'idéal* $\operatorname{Ker} f$ *est isomorphe au sous-anneau* $\operatorname{Im} f = f(A)$ *de A'. On note:*

$$A/\operatorname{Ker} f \simeq \operatorname{Im} f.$$

Preuve. En reprenant la preuve du théorème 6.4.1, on sait déjà que l'application:

$$\begin{array}{rccc} \varphi : & A/\operatorname{Ker} f & \longrightarrow & \operatorname{Im} f \\ & \overline{x} & \longmapsto & f(x) \end{array}$$

est bien définie et réalise un isomorphisme de groupes additifs de $A/\operatorname{Ker} f$ sur $\operatorname{Im} f$. Par ailleurs, en utilisant le fait que f est un morphisme d'anneaux unitaires, on a clairement $\varphi(\overline{1_A}) = f(1_A) = 1_{A'}$ et $\varphi(\overline{x}\,\overline{y}) = \varphi(\overline{xy}) = f(xy) = f(x)f(y) = \varphi(\overline{x})\varphi(\overline{y})$ pour tous $x, y \in A$, ce qui achève de prouver que φ est un isomorphisme d'anneaux unitaires. □

14.1.4 Proposition (idéaux d'un anneau quotient). *Soient A un anneau commutatif et I un idéal de A. Tout idéal de A/I est de la forme J/I pour J un unique idéal de A contenant I, avec la notation naturelle $J/I = p(J)$.*

Preuve. Soit K un idéal de A/I. Posons $J = p^{-1}(K) = \{x \in A\,;\, p(x) \in K\}$. En tant qu'image réciproque d'un idéal par un morphisme d'anneaux, J est un idéal de A. Si $x \in I$, on a $p(x) = \overline{0}$, donc $p(x) \in K$, de sorte que $x \in p^{-1}(K)$, c'est-à-dire $x \in J$. Ceci montre que $I \subseteq J$. Par définition de J, on a $p(J) \subseteq K$. Réciproquement, soit $\overline{x} \in K$, avec $x \in A$; comme $p(x) = \overline{x} \in K$, on a clairement $x \in p^{-1}(K) = J$, et donc $\overline{x} = p(x) \in p(J)$. En résumé, $K = p(J)$, ce que l'on note $K = J/I$.

On renvoie pour l'unicité à la preuve de la proposition 10.3.1. □

14.1.5 Application (idéaux de $\mathbb{Z}/n\mathbb{Z}$). Fixons un entier $n \geq 2$. Alors $n\mathbb{Z}$ est un idéal de $\mathbb{Z}$, et l'anneau quotient n'est autre que l'anneau commutatif unitaire $\mathbb{Z}/n\mathbb{Z}$ déjà considéré en 11.1.3. Pour tout diviseur q de n, il existe un et un seul idéal de $\mathbb{Z}/n\mathbb{Z}$ d'ordre q, qui est $d\mathbb{Z}/n\mathbb{Z}$ où $n = dq$ [voir aussi 6.6.4.(iv)]. Réciproquement tout idéal de $\mathbb{Z}/n\mathbb{Z}$ est de ce type.

Exemple: dans $\mathbb{Z}/12\mathbb{Z}$, les idéaux sont: $\{\overline{0}\} = 12\mathbb{Z}/12\mathbb{Z}$, $\{\overline{0}, \overline{6}\} = 6\mathbb{Z}/12\mathbb{Z}$, $\{\overline{0}, \overline{4}, \overline{8}\} = 4\mathbb{Z}/12\mathbb{Z}$, $\{\overline{0}, \overline{3}, \overline{6}, \overline{9}\} = 3\mathbb{Z}/12\mathbb{Z}$, $\{\overline{0}, \overline{2}, \overline{4}, \overline{6}, \overline{8}, \overline{10}\} = 2\mathbb{Z}/12\mathbb{Z}$ et $\mathbb{Z}/12\mathbb{Z}$. □

14.2 Idéaux premiers, idéaux maximaux.

14.2.1 Définitions. Soit A un anneau commutatif unitaire.

Un idéal P de A est dit *premier* lorsque $P \neq A$ et vérifie:

quels que soient deux éléments x et y de A, si $xy \in P$, alors $x \in P$ ou $y \in P$.

Un idéal M de A est dit *maximal* lorsque $M \neq A$ et vérifie:

quel que soit I un idéal de A, si M est strictement inclus dans I, alors $I = A$.

Remarque. Par définition, ($\{0\}$ premier) $\Leftrightarrow$ (A intègre). Si A est un corps, l'idéal $\{0\}$ est l'unique idéal maximal de A, et si A n'est pas un corps, $\{0\}$ n'est pas maximal (résulte de 13.2.2).

Exercice. La définition d'un idéal premier que l'on a donné ci-dessus en termes de produits d'éléments est équivalente (parce qu'on se limite à des anneaux commutatifs) à la caractérisation suivante en termes de produits d'idéaux, dont on laisse la démonstration au lecteur à titre d'exercice.

Soit A un anneau commutatif unitaire. Un idéal P de A distinct de A est premier si et seulement s'il vérifie: (I et J idéaux de A et $IJ \subseteq P$) $\Rightarrow$ ($I \subseteq P$ ou $J \subseteq P$).

14.2.2 Théorème (fondamental !) *Soit I un idéal d'un anneau commutatif unitaire A. On a:*

$$\begin{array}{ccc} I \text{ maximal} & \Longleftrightarrow & A/I \text{ corps} \\ \Downarrow & & \Downarrow \\ I \text{ premier} & \Longleftrightarrow & A/I \text{ intègre} \end{array}$$

Preuve. Supposons que M est un idéal maximal de A. Comme $M \neq A$, l'anneau A/M est non-nul. Considérons un idéal quelconque K de A/M. D'après la proposition 14.1.4, il existe un idéal J de A tel que $M \subseteq J$ et $K = J/M$. Mais, par maximalité de M, l'inclusion $M \subseteq J$ implique que $J = M$ ou $J = A$, c'est-à-dire $J/M = \{\overline{0}\}$ ou $J/M = A/M$. Ceci prouve que les seuls idéaux de A/M sont $\{\overline{0}\}$ et A/M. On conclut avec 13.2.2 que A/M est un corps. L'implication réciproque découle des mêmes calculs. L'équivalence de la première ligne est donc vérifiée.

Supposons que P est un idéal premier de A. Comme $P \neq A$, l'anneau A/P est non-nul. Considérons $\overline{x}, \overline{y} \in A/P$ tels que $\overline{x}\ \overline{y} = \overline{0}$. On a $\overline{xy} = \overline{0}$, c'est-à-dire $xy \in P$. Comme P est premier, on a $x \in P$ ou $y \in P$, c'est-à-dire $\overline{x} = \overline{0}$ ou $\overline{y} = \overline{0}$. Donc A/P est intègre. L'implication réciproque découle des mêmes calculs. L'équivalence de la seconde ligne est donc vérifiée.

Il suffit de rappeler que tout corps est un anneau intègre, voir 12.3.2.(a), pour achever la preuve. □

14.2.3 Remarques (lien entre primalité et maximalité)

(a) D'après le théorème ci-dessus, tout idéal maximal est premier.

(b) Il existe des anneaux commutatifs unitaires A possédant des idéaux premiers non-nuls qui ne sont pas maximaux.

Exemple. Prenons $A = \mathbb{Z}[X]$ et $I = XA$ l'idéal principal engendré par X. Soit $f : A \to \mathbb{Z}$ l'application qui à tout polynôme $P = a_m X^m + \cdots + a_1 X + a_0$, avec les $a_i \in \mathbb{Z}$, associe le terme constant a_0. Il est facile de vérifier que f est un morphisme d'anneaux unitaires, qu'il est surjectif, et que son noyau est $I = XA$. D'après 14.1.3, on a alors $A/I \simeq \mathbb{Z}$. Comme $\mathbb{Z}$ est intègre sans être un corps, l'idéal I est premier sans être maximal. □

(c) Dans l'anneau $\mathbb{Z}$, considérons un idéal quelconque I. D'après 13.2.4, il existe $k \in \mathbb{N}$ unique tel que $I = k\mathbb{Z}$. Si $k = 1$, alors $I = \mathbb{Z}$ n'est ni premier, ni maximal. Si $k = 0$, alors $I = \{0\}$ est premier mais non maximal. Si maintenant $k \geq 2$, il résulte de la proposition 12.3.3 et du théorème 14.2.2 que:

($k\mathbb{Z}$ est premier) $\Leftrightarrow$ (k est un nombre premier) $\Leftrightarrow$ ($k\mathbb{Z}$ est maximal)

Ainsi dans l'anneau $\mathbb{Z}$, les notions d'idéal maximal et d'idéal premier non-nul coïncident. C'est en fait le cas pour tous les anneaux principaux, comme le montre la proposition suivante.

14.2.4 Proposition. *Dans un anneau principal, tout idéal premier non-nul est maximal (et donc, pour les idéaux non-nuls, les notions de premier et de maximal coïncident).*

Preuve. Soit I un idéal premier non-nul de A. Il existe donc $a \in A$, $a \neq 0$, tel que $I = aA$. Soit J un idéal de A tel que $I \subset J$. Comme A est un anneau principal, il existe $b \in A$, $b \neq 0$, tel que $J = bA$. Comme $a \in I$, on a $a \in J$ donc il existe $x \in A$ tel que $a = bx$. Supposons que $I \neq J$, c'est-à-dire que $b \notin I$. On a $a = bx \in I$ avec $b \notin I$, donc le fait que I soit premier implique que $x \in I$. Donc il existe $y \in A$ tel que $x = ay$. On déduit que $a = bx = bay$, ou encore $a(1 - by) = 0$. L'intégrité de A implique, puisque $a \neq 0$, que $1 - by = 0$, d'où $by = 1$, ce qui prouve que $b \in U(A)$. D'après 13.2.1.(iii), on conclut que $J = A$. Ainsi, pour tout idéal J de A tel que $I \subset J$ et $J \neq I$, on a $J = A$. Donc I est maximal. □

On a vu en 14.2.3.(b) que $\mathbb{Z}[X]$ possède des idéaux premiers non-nuls non maximaux, ce qui donne une nouvelle preuve du fait (déjà montré en 13.2.5) que $\mathbb{Z}[X]$ n'est pas principal.

14.3 Idéaux premiers et morphismes d'anneaux

14.3.1 Lemme. *Soient A et B des anneaux commutatifs unitaires.*

(i) *Si $f : A \to B$ est un morphisme d'anneaux unitaires, alors, quel que soit Q un idéal premier de B, l'image réciproque $f^{-1}(Q)$ est un idéal premier de A, qui contient* $\operatorname{Ker} f$.

(ii) *Si $f : A \to B$ est un morphisme d'anneaux unitaires* surjectif, *alors, quel que soit P un idéal premier de A contenant* $\operatorname{Ker} f$, *l'image directe $f(P)$ est un idéal premier de B.*

(iii) *Soit A un anneau commutatif unitaire et I un idéal de A, distinct de A. Les idéaux premiers de A/I sont les idéaux de la forme P/I où P est un idéal premier de A contenant I.*

Preuve. On prouve (ii), en laissant au lecteur le soin de rédiger de même la preuve de (i). Comme f est supposée surjective, on sait d'après 13.1.3.(iii) que $f(P)$ est un idéal de $B = f(A)$. Montrons d'abord que $f(P) \neq B$. Par l'absurde, supposons $B = f(P)$. Quel que soit $a \in A$, il existerait alors $x \in P$ tel que $f(a) = f(x)$, d'où $a - x \in \operatorname{Ker} f$. Puisque $\operatorname{Ker} f \subseteq P$, on aurait $a - x \in P$, ce qui impliquerait $a \in P$; on obtiendrait $A = P$, ce qui contredit la primalité de P dans A. On conclut donc $f(P) \neq B$.

Soient maintenant $a, b \in B$ tels que $ab \in f(P)$. Par surjectivité de f, il existe $x, y \in A$ tels que $a = f(x)$ et $b = f(y)$. On a $f(xy) = f(x)f(y) = ab \in f(P)$, donc il existe $c \in P$ tel que $f(xy) = f(c)$, d'où $xy - c \in \operatorname{Ker} f$. Comme $\operatorname{Ker} f \subseteq P$, ceci implique $xy - c \in P$, et en rappelant que $c \in P$, il vient $xy \in P$. La primalité de P implique $x \in P$ ou $y \in P$, d'où $a \in f(P)$ ou $b \in f(P)$. Ceci prouve (ii).

D'après 14.1.4, le point (iii) se déduit immédiatement de (ii) en prenant $B = A/I$ et f la surjection canonique $A \to A/I$. □

14.3.2 Exemple d'application (cas des polynômes). Soit A un anneau commutatif unitaire. Pour tout idéal I de A, on note $I[X]$ le sous-ensemble de $A[X]$ formé des polynômes à coefficients dans I, c'est-à-dire de la forme: $\sum_{i=0}^{n} a_i X^i$, avec $n \geq 0$ et $a_0, a_1, \ldots, a_n \in I$. Alors on a:

(i) $I[X]$ est un idéal de $A[X]$;

(ii) les anneaux $(A/I)[X]$ et $A[X]/I[X]$ sont isomorphes;

(iii) $I[X]$ est un idéal premier de $A[X]$ si et seulement si I est un idéal premier de A.

Preuve. Le point (i) est une simple vérification. Pour le (ii), considérons la surjection canonique $p : A \to A/I$ et définissons son extension canonique:

$$\begin{array}{rccc} f : & A[X] & \longrightarrow & (A/I)[X] \\ & P = \sum_{i=0}^{n} a_i X^i & \longmapsto & f(P) = \sum_{i=0}^{n} p(a_i) X^i \end{array}$$

Il est clair que f est un morphisme d'anneaux unitaires, qu'il est surjectif, et que son noyau est $\operatorname{Ker} f = I[X]$. L'isomorphisme $A[X]/\operatorname{Ker} f \simeq \operatorname{Im} f$ devient donc $A[X]/I[X] \simeq (A/I)[X]$. Pour (iii), rappelons que I est premier si et seulement si A/I est intègre, ce qui équivaut d'après 12.3.4.(ii) à $(A/I)[X]$ intègre, c'est-à-dire $A[X]/I[X]$ intègre d'après le point (ii), ou encore $I[X]$ premier dans $A[X]$. □

14.4 Théorème de Krull.

Le théorème suivant est un résultat important et non trivial qui démontre l'existence d'idéaux maximaux dans tout anneau unitaire commutatif. Sa preuve utilise des arguments d'algèbre générale sur les structures ordonnées, dont le lemme de Zorn, et on ne donnera ci-dessous que le plan général de la preuve, sans entrer dans le détail des justifications de chaque étape.

14.4.1 Théorème (de Krull). *Tout anneau commutatif unitaire a au moins un idéal maximal.*

Grandes lignes de la preuve. Soit A un anneau commutatif unitaire. Soit E l'ensemble de tous les idéaux de A distincts de A. Il est non vide, car contient au moins $\{0\}$. L'inclusion définit une relation d'ordre dans E. Ce n'est pas un ordre total, mais seulement un ordre partiel (c'est-à-dire que, si $I, J \in E$ quelconques, on n'a pas forcément $I \subseteq J$ ou $J \subseteq I$).

Soit $F = (I_k)_{k \in X}$ une famille d'éléments de E totalement ordonnée par l'inclusion (quels que soient $k, \ell \in X$, on a $I_k \subseteq I_\ell$ ou $I_\ell \subseteq I_k$). On peut facilement vérifier qu'alors $I = \bigcup_{k \in X} I_k$ est un idéal de A. (Rappelons qu'en général une réunion d'idéaux n'est pas un idéal, mais le fait que I soit ici un idéal provient du fait que tous les I_k sont emboités puisque la famille est totalement ordonnée). L'idéal I est distinct de A (car sinon on aurait $1 \in I$, donc il existerait $k \in X$ tel que $1 \in I_k$, d'où $I_k = A$, ce qui contredirait $I_k \in E$). Donc $I \in E$, et il est clair que tout $I_k \in F$ vérifie $I_k \subseteq I$. En résumé, toute famille déléments de E totalement ordonnée admet un plus grand élément. On traduit cette propriété en disant que l'ensemble partiellement ordonné E est *inductif.*

Or un résultat d'algèbre très général (et non trivial) sur les structures ordonnées (le lemme de Zorn) affirme que tout ensemble (non vide) ordonné inductif admet (au moins) un élément maximal. Soit donc M un élément maximal de E. Cela signifie que, quel que soit un $J \in E$ tel que $M \subseteq J$, on a $J = M$. En d'autres termes, quel que soit un idéal J de A tel que $J \neq A$ et $M \subseteq J$, on a $J = M$. Ceci prouve que M est un idéal maximal de A, ce qui achève la preuve. □

14.4.2 Corollaire. *Soit A un anneau commutatif unitaire.*

(i) *Pour tout idéal I de A, distinct de A, les idéaux maximaux de A/I sont de la forme M/I où M est un idéal maximal de A contenant I.*

(ii) *Tout idéal distinct de A est contenu dans un idéal maximal de A.*

(iii) *Tout élément de A non inversible dans A est contenu dans un idéal maximal de A.*

Preuve. Soit I un idéal de A tel que $I \neq A$. D'après 14.4.1, l'anneau A/I admet un idéal maximal N. D'après 14.1.4, il existe un unique idéal M de A contenant I tel que $N = M/I$, où M/I désigne l'image $p(M)$ de M par la surjection canonique $p : A \to A/I$. On se propose de montrer que M est un idéal maximal de A. Pour cela, soit J un idéal de A tel que $M \subseteq J$. On a donc $I \subseteq M \subseteq J \subseteq A$, ce qui implique pour les images par p que $M/I \subseteq J/I \subseteq A/I$. La maximalité de l'idéal $N = M/I$ implique $J/I = M/I$ ou $J/I = A/I$, c'est-à-dire $J = M$ ou $J = A$. On conclut que M est un idéal maximal de A, ce qui prouve à la fois (i) et (ii). Le point (iii) résulte immédiatement du (ii) et de 13.2.1.(iii). □

14.5 Résultats complémentaires sur les anneaux quotients

Citons encore les deux résultats généraux suivants (on ne détaille pas les preuves, qui sont de simples adaptations de celles de 10.1.1 et 10.1.3 pour les groupes), et précisons que les théorèmes 10.2.1 et 10.3.2 ont aussi leurs analogues pour les anneaux (on laisse au lecteur le soin d'en préciser l'énoncé et la preuve).

14.5.1 Théorème (propriété universelle de l'anneau quotient) *Soient A un anneau commutatif unitaire, I un idéal de A, et p la surjection canonique $A \to A/I$.*

(i) *Pour tout anneau commutatif unitaire A' et tout morphisme d'anneaux unitaires $f : A \to A'$ tel que $I \subseteq \operatorname{Ker} f$, il existe un unique morphisme d'anneaux unitaires $\varphi : A/I \to A'$ tel que $f = \varphi \circ p$.*

$$\begin{array}{ccc} A & \xrightarrow{f} & A' \\ {\scriptstyle p}\downarrow & \nearrow_{\varphi} & \\ A/I & & \end{array}$$

(ii) *De plus: (f surjectif $\Rightarrow$ φ surjectif) et ($I = \operatorname{Ker} f$ $\Rightarrow$ φ injectif).*

14.5.2 Lemme (fondamental de factorisation). *Soient A un anneau commutatif unitaire, I un idéal de A, et p la surjection canonique $A \to A/I$. Soient A' un anneau commutatif unitaire, I' un idéal de A', et p' la surjection canonique $A' \to A'/I'$. Alors, pour tout morphisme d'anneaux unitaires $f : A \to A'$ vérifiant la condition $f(I) \subseteq I'$, il existe un unique morphisme $\varphi : A/I \to A'/I'$ tel que $\varphi \circ p = p' \circ f$.*

$$\begin{array}{ccc} A & \xrightarrow{f} & A' \\ {\scriptstyle p}\downarrow & \searrow & \downarrow{\scriptstyle p'} \\ A/I & \xrightarrow[\varphi]{} & A'/I' \end{array}$$

Leçon 15

Anneaux euclidiens, anneaux principaux

15.1 Multiples, diviseurs et idéaux principaux.

15.1.1 Définitions. Soit A un anneau commutatif unitaire. Soient x et y deux éléments de A. On dit que x est un *diviseur* de y dans A, ou encore que x *divise* y dans A, ou encore que y est un *multiple* de x dans A, lorsqu'il existe $a \in A$ tel que $y = xa$. On note alors: $x|y$.

15.1.2 Proposition. *Soit A un anneau commutatif unitaire. Pour tous $x, y \in A$, on a:*

$$(\ x|y \ ) \Leftrightarrow (\ y \in xA \ ) \Leftrightarrow (\ yA \subseteq xA \ ).$$

Preuve. Supposons que $x|y$. Il existe $a \in A$ tel que $y = xa$. Donc $y \in xA$. De plus, tout élément de yA est de la forme yb avec $b \in A$, donc de la forme xab, et donc appartient à xA, ce qui montre que $yA \subseteq xA$. La réciproque est claire. □

15.2 Notion d'anneau euclidien.

15.2.1 Proposition (exemple préliminaire de l'anneau $\mathbb{Z}$). *Quels que soient des entiers a et b, avec $b \neq 0$, il existe $q \in \mathbb{Z}$ et $r \in \mathbb{N}$ uniques tels que $a = bq + r$ et $0 \leq r < |b|$.*

Preuve. Pour montrer l'unicité, supposons l'existence de deux couples (q, r) et (q', r') dans $\mathbb{Z} \times \mathbb{N}$ satisfaisant aux conditions $a = bq + r$ avec $0 \leq r < |b|$, et $a = bq' + r'$ avec $0 \leq r' < |b|$. On a alors $b(q - q') = r' - r$ et $-|b| < r' - r < |b|$. Donc $-|b| < b(q - q') < |b|$. Comme $b \neq 0$, on en déduit que $-1 < q - q' < 1$, ce qui, puisque $q - q'$ est un entier, implique $q - q' = 0$. Ainsi $q = q'$, d'où $r = r'$.

Pour montrer l'existence, supposons d'abord $b > 0$. Posons $B = \{k \in \mathbb{Z}\,;\, kb \leq a\}$. C'est une partie de $\mathbb{Z}$ qui est non-vide (car $0 \in B$ si $a \geq 0$ et $a \in B$ si $a < 0$) et qui est majorée (par le maximum des entiers a et 0). Donc elle admet un plus grand élément. Notons-le q. On a par définition de q la double inégalité $qb \leq a < (q + 1)b$, de sorte que l'entier $r = a - qb$ vérifie $0 \leq r < b$.

Supposons maintenant $b < 0$. D'après ce qui précède, il existe $(q, r) \in \mathbb{Z} \times \mathbb{N}$ tel que $a = (-b)q + r$ et $0 \leq r < |b|$. Le couple $(-q, r) \in \mathbb{Z} \times \mathbb{N}$ vérifie alors $a = b(-q) + r$ et $0 \leq r < |b|$. □

15.2.2 Proposition (exemple préliminaire de l'anneau $K[X]$). *Soit K un corps commutatif. Quels que soient des polynômes F et G dans $K[X]$, avec $G \neq 0$, il existe $Q \in K[X]$ et $R \in K[X]$ uniques tels que $F = GQ + R$ et $\deg R < \deg G$.*

Preuve. Pour montrer l'unicité, supposons que deux couples (Q, R) et (Q', R') dans $K[X] \times K[X]$ satisfassent aux conditions $F = GQ + R = GQ' + R'$ avec $\deg R < \deg G$ et $\deg R' < \deg G$. On a alors: $G(Q - Q') = R' - R$. Comme K est un corps (en particulier intègre), on a d'après 12.3.4.(i) l'égalité $\deg G + \deg(Q - Q') = \deg(R' - R)$. Or $\deg R < \deg G$ et $\deg R' < \deg G$ implique d'après 11.1.4.(e) que $\deg(R' - R) < \deg G$. Donc $\deg G + \deg(Q - Q') < \deg G$, ce qui n'est possible que si $\deg(Q - Q') = -\infty$, c'est-à-dire $Q = Q'$. On a alors forcément aussi $R = R'$.

Pour montrer l'existence, notons $n = \deg F$ et $m = \deg G \in \mathbb{N}$. Si $n < m$, on a le résultat voulu en prenant $Q = 0$ et $R = F$. On suppose donc désormais que $n \geq m \geq 0$. Notons:

$$F = a_n X^n + \cdots + a_1 X + a_0 \quad \text{et} \quad G = b_m X^m + \cdots + b_1 X + b_0$$

avec les a_i et les b_j dans K, tels que $a_n \neq 0 \neq b_m$.

Si $n = m = 0$, alors $F = a_0 \neq 0$ et $G = b_0 \neq 0$, donc $F = (a_0 b_0^{-1})G$, ce qui prouve le résultat voulu avec $Q = a_0 b_0^{-1}$ et $R = 0$.

Par récurrence sur n, supposons la propriété voulue vraie pour G et tout polynôme F_1 de degré n_1 tel que $n > n_1 \geq m \geq 0$. Or on peut écrire $F = a_n b_m^{-1} X^{n-m} G + F_1$ avec $\deg F_1 \leq n - 1 < n$. Par hypothèse de récurrence, il existe $Q_1, R_1 \in K[X]$ tels que $F_1 = Q_1 G + R_1$ et $\deg R_1 < \deg G$. On déduit que $F = (a_n b_m^{-1} X^{n-m} + Q_1)G + R_1$, ce qui prouve le résultat voulu avec $Q = a_n b_m^{-1} X^{n-m} + Q_1$ et $R = R_1$. □

15.2.3 Définition. On appelle *anneau euclidien* un anneau commutatif unitaire qui est intègre, et pour lequel il existe une application $\delta : A^* \to \mathbb{N}$ vérifiant les deux conditions suivantes:

1. pour tous $a, b \in A^*$, $(\ a|b\ ) \Rightarrow (\ \delta(a) \leq \delta(b)\ )$;
2. pour tout $a \in A$ et $b \in A^*$, il existe $q, r \in A$ tels que:
$$(\ a = bq + r\ ) \quad \text{et} \quad (\ r = 0 \ \text{ ou } \ \delta(r) < \delta(b)\ ).$$

Une application δ vérifiant ces deux conditions s'appelle un *stathme* euclidien. Dans la condition 2, on dit que q est un *quotient* et r un *reste* dans la *division euclidienne* de a par b.

15.2.4 Exemples.

(a) L'anneau $\mathbb{Z}$ est euclidien, pour le stathme défini par $\delta(x) = |x|$ pour tout $x \in \mathbb{Z}^*$.

(b) Si K est un corps, l'anneau $K[X]$ est euclidien, pour le stathme défini par $\delta(F) = \deg F$ pour tout $F \in K[X]$ non-nul.

(c) L'anneau $\mathbb{Z}[i]$ est euclidien, pour le stathme défini par $\delta(z) = z\,\overline{z}$ pour tout $z \in \mathbb{Z}[i]$ non-nul.

Preuve. Les exemples (a) et (b) découlent directement des propositions 15.2.1 et 15.2.2 respectivement. L'exemple (c) est laissé en exercice. □

15.2.5 Remarque. La définition d'un stathme n'impose pas de conditions d'unicité de q et r dans la seconde condition.

Et effectivement, ils ne sont pas forcément uniques. Par exemple, pour $a = 19$ et $b = 3$, on a: $19 = 6 \times 3 + 1 = 7 \times 3 + (-2)$ avec $r = 1$ et $r' = -2$ qui vérifient tous les deux
$$\delta(r) = |1| = 1 < \delta(3) = 3 \ \text{ et } \ \delta(r') = |-2| = 2 < \delta(3) = 3.$$

L'unicité de q et r qui apparaît dans la proposition 15.2.1 tient au fait qu'on y a remplacé la condition ($r = 0$ ou $|r| < |b|$), qui correspond à la définition du stathme, par la condition plus forte $0 \leq r < |b|$.

Rappelons la notion d'anneau principal introduite en 13.2.3.

15.3 Liens entre anneaux euclidiens et anneaux principaux.

15.3.1 Définition. On appelle *anneau principal* un anneau commutatif unitaire qui est intègre, et dans lequel tout idéal est principal.

En d'autres termes, quel que soit I un idéal de A, il existe $x \in A$ (non unique a priori) tel que $I = xA$.

Le théorème suivant fournit une vaste classe d'anneaux principaux.

15.3.2 Théorème. *Tout anneau euclidien est principal.*

Preuve. Soit A un anneau euclidien, de stathme δ. Il est intègre, et il s'agit donc de montrer que tout idéal I de A est principal. C'est clair si $I = \{0\}$ (alors $I = 0A$) ou si $I = A$ (alors $I = 1A$). On suppose donc $I \neq \{0\}$ et $I \neq A$. On considère $E = \{\delta(x)\,;\, x \in I, x \neq 0\}$. C'est une partie non-vide de $\mathbb{N}$, elle admet donc un plus petit élément n. Il existe $x \in I$, $x \neq 0$ tel que $n = \delta(x)$. Soit alors $a \in I$ quelconque; par division euclidienne de a par x, il existe $q, r \in A$ tels que $a = xq + r$ avec $r = 0$ ou $\delta(r) < \delta(x) = n$. Or $r = a - xq$ avec $a \in I$ et $x \in I$, donc $r \in I$ par définition d'un idéal. Par minimalité de n, on ne peut donc pas avoir $\delta(r) < n$, et donc nécessairement $r = 0$, d'où $a = xq$. Ceci prouve que tout $a \in I$ appartient à xA, c'est-à-dire $I \subseteq xA$. Comme par ailleurs $xA \subseteq I$ puisque $x \in I$, on conclut que $I = xA$. □

15.3.3 Exemples, contre-exemples, remarques.

(a) $\mathbb{Z}$, $\mathbb{Z}[i]$, et $K[X]$ lorsque K est un corps, sont des anneaux principaux car euclidiens.

Preuve. Résulte immédiatement du théorème ci-dessus et des exemples 15.2.4. □

(b) La réciproque du théorème 15.3.2 est fausse. Il existe des exemples d'anneaux principaux qui ne sont pas euclidiens. On pourra par exemple montrer en TD que:

l'anneau $\mathbb{Z}[\omega] = \{a + \omega b\,;\, a, b \in \mathbb{Z}\}$ *pour* $\omega = \frac{1+i\sqrt{19}}{2}$ *est principal et non euclidien.*

(c) On a montré de deux façons différentes (en 13.2.5 et en 14.2.4) que l'anneau $\mathbb{Z}[X]$ n'est pas principal. Il en résulte en particulier que:

$$(\ A \text{ euclidien} \not\Rightarrow A[X] \text{ euclidien}\ ) \quad \text{et} \quad (\ A \text{ principal} \not\Rightarrow A[X] \text{ principal}\ ).$$

On a en fait le résultat général suivant:

15.3.5 Théorème. *Soit A un anneau commutatif unitaire. Les trois conditions suivantes sont équivalentes.*

(i) *A est un corps.* (ii) *$A[X]$ est euclidien;* (iii) *$A[X]$ est principal.*

Preuve. On a déja vu que (i) ⇒ (ii) ⇒ (iii). Supposons donc maintenant $A[X]$ principal. En particulier, $A[X]$ est intègre, et donc, d'après 12.3.4.(ii), A est intègre. Considérons l'application $f : A[X] \to A$ qui, à tout polynôme $P = \sum_{i=0}^{n} a_i X^i$, associe le coefficient a_0. Il est facile de voir que f est un morphisme d'anneaux, qui est clairement surjectif. Donc le premier théorème d'isomorphisme 14.1.3 conduit à $A[X]/\operatorname{Ker} f \simeq A$. L'intégrité de A implique que $A[X]/\operatorname{Ker} f$ est intègre, donc, d'après 14.2.2, $\operatorname{Ker} f$ est un idéal premier non-nul de $A[X]$. Mais comme $A[X]$ est supposé principal, $\operatorname{Ker} f$ est alors, d'après 14.2.4, un idéal maximal de $A[X]$, et donc, d'après 14.2.2, $A[X]/\operatorname{Ker} f$ est un corps. On conclut via l'isomorphisme $A[X]/\operatorname{Ker} f \simeq A$ que A est un corps. □

15.4 Une application de la division euclidienne dans $K[X]$.

15.4.1 Fonctions polynomiales. Soit K un corps. Pour tout polynôme $P \in K[X]$, on note $\Phi(P) = p$ la fonction poynomiale associée dans l'anneau $\mathcal{F}(K, K)$ des applications de K dans K. Rappelons que, par définition, si $P = a_n X^n + \cdots + a_1 X + a_0 \in K[X]$, alors p est la fonction $x \mapsto a_n x^n + \cdots + a_1 x + a_0$ de K dans K. Il est facile de voir que l'application $\Phi : K[X] \to \mathcal{F}(K, K)$ ainsi définie est un morphisme d'anneaux. Rappelons quelques notions et résultats établis en première année.

15.4.2 Définition. On appelle *zéro* (ou *racine*) dans le corps K d'un polynôme $P \in K[X]$ tout élément $a \in K$ tel que la fonction polynomiale $p : K \to K$ associée à P vérifie $p(a) = 0$.

15.4.3 Théorème. *Soit K un corps, et P un polynôme de $K[X]$.*

(i) *Un élément a de K est un zéro de P si et seulement si $X - a$ divise P dans $K[X]$.*

(ii) *Si P est non-nul de degré n, alors P admet au plus n zéros dans K.*

Preuve. On effectue la division euclidienne de P par $X - a$ dans $K[X]$. On obtient $P = (X - a)Q + R$ avec $\deg R < 1$, c'est-à-dire $R \in K$. En passant aux fonctions polynomiales associées, il vient $p(x) = (x-a)q(x)+r$, où $r = R \in K$, donc $R = p(a)$. Ainsi, on a dans $K[X]$ l'égalité $P = (X-a)Q+p(a)$. Dès lors, $X - a$ divise P si et seulement $p(a) = 0$, ce qui prouve (i). On en déduit aisément par récurrence sur n qu'un polynôme de degré n dans $K[X]$ possédant au moins $n+1$ zéros dans K est nécessairement nul, ce qui prouve (ii). □

15.4.4 Lien entre polynômes et fonctions polynomiales. Reprenons le morphisme d'anneaux $\Phi : K[X] \to \mathcal{F}(K, K)$ associant à un polynôme sa fonction polynomiale associée. Peut-il être non-injectif ? En d'autres termes, deux polynômes distincts dans $K[X]$ peuvent-ils avoir la même fonction polynomiale associée ?

En toute généralité, la réponse est oui. Par exemple, dans le corps $K = \mathbb{Z}/3\mathbb{Z}$, le polynôme $P = X(X-\overline{1})(X-\overline{2}) = X^3+\overline{2}X$ est non-nul bien que sa fonction polynomiale associée $p(x) = x(x-\overline{1})(x-\overline{2}) = x^3 + \overline{2}x$ soit la fonction nulle (puisqu'elle s'annule en tout point de K). Mais, dès lors que le corps K est infini, un polynôme P tel que $p = \Phi(P)$ est la fonction nulle admet une infinité de zéros dans K, donc est le polynôme nul d'après le point (ii) du théorème 15.4.3. En résumé:

si le corps K est infini, le morphisme Φ est injectif, et l'on peut sans inconvénient identifier un polynôme avec sa fonction polynomiale associée.

Leçon 16

Divisibilité

16.1 Multiples et diviseurs.

16.1.1 Rappel (voir aussi 15.1.1). Soit A un anneau commutatif unitaire. Soient x et y deux éléments de A. On dit que x est un *diviseur* de y dans A, ou encore que x *divise* y dans A, ou encore que y est un *multiple* de x dans A, lorsque il existe $a \in A$ tel que $y = xa$. On note alors: $x|y$. On a montré que: pour tous $x, y \in A$, on a:

$$(x|y) \Leftrightarrow (y \in xA) \Leftrightarrow (yA \subseteq xA).$$

16.1.2 Remarques. On déduit immédiatement que:

(i) *Pour tous* $x, y, z \in A$, $(x|y$ et $y|z) \Rightarrow (x|z)$.

(ii) *Pour tout* $u \in A$, $(u \in U(A)) \Leftrightarrow (uA = A) \Leftrightarrow (u|y$ quel que soit $y \in A)$.

(iii) *Pour tous* $x, u \in A$, $(u \in U(A)$ et $x|u) \Rightarrow (x \in U(A))$.

16.1.3 Commentaire. La notion de multiple et de diviseur, et donc toute la théorie de la divisibilité que l'on va détailler dans les leçons suivantes, n'a d'intérêt que dans un anneau qui n'est pas un corps. En effet, dans un corps K, deux éléments x et y non-nuls sont toujours à la fois multiple et diviseur l'un de l'autre puisque $x = zy$ avec $z = xy^{-1} \in K$ et $y = tx$ avec $t = yx^{-1} = z^{-1} \in K$.

16.2 Eléments associés.

16.2.1 Définition. Soit A un anneau commutatif unitaire *intègre*. Soient x et y deux éléments de A. On dit que x et y sont *associés* lorsqu'on a à la fois $x|y$ et $y|x$. On note alors $x \sim y$.

16.2.2 Proposition. *Soit A un anneau commutatif unitaire intègre. Soient x et y deux éléments de A. On a:*

$$(x \sim y) \Leftrightarrow (x|y \text{ et } y|x) \Leftrightarrow (xA = yA) \Leftrightarrow (\text{il existe } u \in U(A) \text{ tel que } x = uy).$$

Preuve. La première équivalence est vraie par définition, la seconde découle directement de 16.1.1. Pour la dernière, supposons que $x \sim y$. Il existe $u, v \in A$ tels que $x = uy$ et $y = vx$, donc $x = uvx$. Si $x = 0$, alors $y = 0$, et on a $x = uy$ pour tout $u \in U(A)$. Si $x \neq 0$, on écrit $x(1 - uv) = 0$ et on utilise l'intégrité de A pour déduire que $uv = 1$, d'où $u \in U(A)$, ce qui montre le résultat voulu. Réciproquement, supposons $x = uy$ avec $u \in U(A)$; on a $y|x$ et, puisque $y = u^{-1}x$ avec $u^{-1} \in A$, on a aussi $x|y$. On conclut que $x \sim y$. □

16.2.3 Exemples.

(a) Dans $\mathbb{Z}$, deux entiers m et n sont associés si et seulement si $m = \pm n$; (rappelons en effet que $U(\mathbb{Z}) = \{1, -1\}$).

(b) Pour tout anneau intègre A, deux polynômes P et Q de $A[X]$ sont associés si et seulement s'il existe $c \in U(A)$ tel que $P = cQ$, (rappelons que $U(A[X]) = U(A)$), et l'on a alors $Q = c^{-1}P$.

(c) En particulier, si K est un corps, deux polynômes P et Q de $K[X]$ sont associés si et seulement s'il existe $c \in K^*$ tel que $P = cQ$.

16.2.4 Remarque. Deux éléments associés ont les mêmes multiples et les mêmes diviseurs dans l'anneau A.

En effet. Supposons $x \sim y$. On a $x = uy$ avec $u \in U(A)$. On sait déjà (voir 16.2.2) que $xA = yA$. Soit z un diviseur de y; il existe $a \in A$ tel que $y = za$. Donc $x = uy = uza$, et donc z divise x. □

16.3 Eléments irréductibles, éléments premiers.

16.3.1 Définitions. Soit A un anneau commutatif unitaire *intègre*. Soit x un élément de A.

(a) x est dit irréductible dans A lorsqu'il n'est pas inversible dans A, et vérifie la condition:

si $x = ab$ avec $a, b \in A$, alors $a \in U(A)$ ou $b \in U(A)$.

(b) x est dit premier dans A lorsqu'il est non-nul et non inversible dans A, et vérifie la condition:

si x divise ab avec $a, b \in A$, alors x divise a ou x divise b.

16.3.2 Remarques.

(a) 0 n'est pas irréductible dans A.

(b) Dans la définition (a), le "ou" est exclusif. En d'autres termes, si x est irréductible dans A et s'écrit $x = ab$, alors un seul des deux éléments a, b appartient à $U(A)$ (car si les deux étaient dans $U(A)$, alors x appartiendrait aussi à $U(A)$, ce qui est contraire à la définition).

(c) Un élément de A peut être irréductible dans A mais ne plus l'être dans un anneau contenant A. Par exemple, 3 est irréductible dans $\mathbb{Z}$, mais ne l'est pas dans $\mathbb{Q}$ puisqu'il est inversible dans $\mathbb{Q}$.

16.3.3 Proposition (caractérisation en termes d'idéaux principaux). *Soit A un anneau commutatif unitaire intègre. Pour tout $x \in A$, on a:*

(i) (*x irréductible dans A*) $\Leftrightarrow$ (*xA maximal parmi les idéaux principaux distincts de A*).

(ii) (*x premier dans A*) $\Leftrightarrow$ (*xA idéal premier non-nul de A*).

Preuve. Supposons x irréducible. L'idéal principal $M = xA$ est distinct de A puisque $x \notin U(A)$. Soit $J = aA$ un idéal principal de A distinct de A, c'est-à-dire tel que $a \notin U(A)$, et supposons que $M \subseteq J$. Alors en particulier $x \in J$, donc il existe $b \in A$ tel que $x = ab$. Puisque $a \notin U(A)$, l'irréductibilité de x implique que $b \in U(A)$. Donc $x \sim a$, d'où $M = J$. Ceci prouve que M est maximal parmi les idéaux principaux distincts de A.

Réciproquement soit $x \in A$ tel que xA soit maximal parmi les idéaux principaux distincts de A. Soient $a, b \in A$ tels que $x = ab$. Alors $x \in aA$, et donc $xA \subseteq aA$. Si $a \in U(A)$, alors $aA = A$. Sinon, $aA \neq A$ et la maximalité de xA implique alors que $xA = aA$, donc $x \sim a$, d'où l'existence de $u \in U(A)$ tel que $x = ua$. Mais $x = ua = ba$ implique par intégrité de A que $b = u$, et donc $b \in U(A)$.

Ceci prouve (i). L'équivalence (ii) est quant à elle évidente par définition même d'un idéal premier et la traduction de la divisibilité en termes d'idéaux principaux rappelée en 16.1.1. □

16.3.4 Corollaire. *Soit A un anneau commutatif unitaire intègre.*

(i) *Tout élément de A associé à un élément irréductible dans A est encore irréductible dans A.*

(ii) *Tout élément de A associé à un élément premier dans A est encore premier dans A.*

Preuve. Découle de 16.3.3 puisque deux éléments associés engendrent le même idéal principal. □

16.3.5 Proposition. *Soit A un anneau commutatif unitaire intègre. Tout élément premier dans A est irréductible dans A.*

Preuve. Soit $x \in A$ premier dans A. On a $x \notin U(A)$. Supposons que $x = ab$ avec $a, b \in A$. En particulier $x|ab$, donc puisque x est premier, $x|a$ ou $x|b$. Supposons que $x|a$. Il existe $y \in A$ tel que $a = xy$, d'où $x = xyb$, ou encore $x(1 - yb) = 0$. Comme x est non-nul car premier, et que A est intègre, on conclut que $yb = 1$, et donc que $b \in U(A)$. On prouve de même que $a \in U(A)$ si $x|b$. □

16.3.6 Remarque. La réciproque est fausse en général. Par exemple, dans l'anneau $\mathbb{Z}[i\sqrt{5}] = \{a + ib\sqrt{5}\,; a, b \in \mathbb{Z}\}$, l'élément 3 est irréductible, mais non premier.

En effet. Posons $A = \mathbb{Z}[i\sqrt{5}]$. C'est un anneau commutatif unitaire intègre (vérifiez-le) qui contient $\mathbb{Z}$ comme sous-anneau.

Montrons d'abord que 3 n'est pas premier dans A. Observons d'abord que 3 ne divise pas $2+i\sqrt{5}$ dans A (en effet, on aurait sinon $(2+i\sqrt{5}) = 3(a+ib\sqrt{5})$ avec $a, b \in \mathbb{Z}$, d'où $3a = 2$ et $1 = 3b$, ce qui est impossible), et que de même 3 ne divise pas $2-i\sqrt{5}$. Et pourtant 3 divise $9 = (2+i\sqrt{5})(2-i\sqrt{5})$ dans A, puisque $9 = 3.3$. On conclut que 3 n'est pas premier dans A.

Montrons maintenant que 3 est irréductible dans A. Il est clair que 3 n'est pas inversible dans A. Supposons que $3 = xy$ avec $x = a+ib\sqrt{5}$ et $y = c+id\sqrt{5}$, où $a, b, c, d \in \mathbb{Z}$. Posons $N(x) = |x|^2 = a^2+5b^2$ et $N(y) = |y|^2 = c^2+5d^2$. On a: $9 = N(xy) = N(x)N(y)$ dans $\mathbb{N}^*$, et donc trois cas seulement sont possibles: $N(x) = N(y) = 3$, ou $N(x) = 1$ et $N(y) = 9$, ou $N(x) = 9$ et $N(y) = 1$. Or le premier cas est impossible (car $a^2+5b^2 = 3$ n'a pas de solutions entières), le second implique que $x \in U(A)$ (car $a^2+5b^2 = 1$ implique $a = \pm 1$ et $b = 0$, et donc $x = \pm 1$), et le troisième implique de même que $y \in U(A)$. On conclut que 3 est irréductible dans A. □

Néanmoins, on a le résultat suivant:

16.3.7 Proposition (cas particulier des anneaux principaux). *Si A est un anneau principal, tout élément irréductible est premier, et donc les notions d'élément premier et d'élément irréductible coïncident dans ce cas.*

Preuve. Soit x un élément irréductible de A. Il est non-inversible (par définition), non-nul (voir remarque (a) de 16.3.2), et l'idéal $M = xA$ est maximal parmi les idéaux principaux de A distincts de A. Mais ici, tout idéal de A est par hypothèse principal. Donc M est tout simplement un idéal maximal de A. Donc M est un idéal premier de A (voir 14.2.2), et comme il est non-nul, on déduit de 16.3.3.(ii) que x est un élément premier dans A. □

16.3.8 Exemples.

(a) Dans $\mathbb{Z}$, les éléments premiers (ou irréductibles) sont les nombres premiers et leurs opposés.

(b) Pour tout corps K, les polynômes de degré un sont toujours irréductibles dans $K[X]$.

(c) Si $K = \mathbb{C}$, les éléments irréductibles de $\mathbb{C}[X]$ sont les polynômes de degré un.

(d) Si $K = \mathbb{R}$, les éléments irréductibles de $\mathbb{R}[X]$ sont les polynômes de degré un, et les polynômes de degré deux de discriminant strictement négatif.

Les assertions (c) et (d) découlent du théorème 15.4.3, et ont été montrées en première année (repris plus loin leçon 18.2.2).

16.4 Eléments premiers entre eux, plus grand commun diviseur.

16.4.1 Définition. Soit A un anneau commutatif unitaire intègre. Soient x et y deux éléments de A. On dit que x et y sont *premiers entre eux*, ou *étrangers*, lorsque les seuls éléments de A qui divisent à la fois x et y sont les éléments de $U(A)$.

Exemple. Dans $\mathbb{Z}$ l'ensemble des diviseurs de 10 est $D_{10} = \{-10, -5, -2, -1, 1, 2, 5, 10\}$ et l'ensemble des diviseurs de 9 est $D_9 = \{-9, -3, -1, 1, 3, 9\}$. On a donc $D_{10} \cap D_9 = \{-1, 1\} = U(\mathbb{Z})$, donc 10 et 9 sont premiers entre eux. □

Exercice. Montrer que, dans $\mathbb{R}[X]$, les polynômes $P = X + 2$ et $Q = X - 1$ sont premiers entre eux. □

Remarque: si x est premier avec y, alors x est premier avec tout élément associé à y.

16.4.2 Proposition. *Tout élément irréductible est premier avec tout élément qu'il ne divise pas.*

Preuve. Soit x irréductible dans A. Soit $y \in A$ tel que x ne divise pas y. Par l'absurde, supposons que u soit un diviseur commun de x et y non inversible dans A. On aurait alors $x = ua$ et $y = ub$ avec $a, b \in A$. Comme $x = ua$ et $u \notin U(A)$, l'irréductibilité de x impliquerait que $a \in U(A)$. On obtiendrait $u = xa^{-1}$ avec $a^{-1} \in A$, de sorte que $y = xa^{-1}b$, ce qui contredit le fait que x ne divise pas y.

16.4.3 Définition. Soit A un anneau commutatif unitaire intègre. Soient x et y deux éléments de A. On dit que x et y admettent un plus grand commun diviseur dans A lorsqu'il existe un élément $d \in A$ tel que:

$$d \text{ divise } x, \ \ d \text{ divise } y, \ \text{ et tout élément qui divise à la fois } x \text{ et } y \text{ divise aussi } d.$$

On dit alors que d est un pgcd de a et b.

16.4.4 Proposition. *Soit A un anneau commutatif unitaire intègre. Soient $x, y \in A$.*

(i) *Si x et y admettent un pgcd d, alors un élément quelconque $d' \in A$ est un pgcd de x et y si et seulement si d' est associé à d.*

(ii) *x et y sont premiers entre eux si et seulement si 1 est un pgcd de x et y.*

(iii) *x et y sont premiers entre eux si et seulement si $U(A)$ est l'ensemble des pgcd de x et y.*

Preuve. Montrons (i). Si d' est un pgcd de x et y, il divise x et y, et donc puisque d est un pgcd de x et y, on a $d'|d$. De même, $d|d'$, et donc $d \sim d'$. Comme deux éléments associés ont les mêmes multiples et les mêmes diviseurs (voir 16.2.4), la réciproque est claire. Les points (ii) et (iii) se déduisent alors immédiatement de (i) et de 16.4.1. □

16.4.5 Remarques.

(a) Si $x = 0$, alors l'ensemble des diviseurs de x est A. Donc, pour tout $y \in A$, un pgcd de x et y est y. Les autres pgcd sont les éléments associés à y. En particulier, si $x = y = 0$, le seul pgcd de x et y est 0.

(b) On définit de même le pgcd d'un nombre fini quelconque d'éléments de A.

16.4.6 Proposition. *Soit A un anneau commutatif unitaire intègre. Soient $x, y \in A$ non-nuls. Si x et y admettent un pgcd d, alors les deux éléments x' et y' tels que $x = dx'$ et $y = dy'$ sont premiers entre eux dans A.*

Preuve. Soit z un diviseur commun à x' et y'. Il existe $a, b \in A$ tels que $x' = za$ et $y' = zb$. Donc $x = dza$ et $y = dzb$. Ceci prouve que dz est un diviseur commun à x et y, donc un diviseur de leur pgcd d. Il existe donc $u \in A$ tel que $d = dzu$, ou encore $d(1 - zu) = 0$. Comme A est intègre et $d \neq 0$ (car x et y sont non-nuls), on a $zu = 1$. On conclut que $z \in U(A)$. □

16.4.7 Exemples et remarque.

(a) Dans $\mathbb{Z}$, les pgcd de 12 et 30 sont 6 et -6.

(b) Dans $\mathbb{R}[X]$, les pgcd de $X^2 - 3X + 2$ et $X^2 - 1$ sont tous les polynômes $\alpha(X - 1)$ où $\alpha \in \mathbb{R}^*$.

De fait, dans les situations que l'on connaît bien de l'arithmétique dans $\mathbb{Z}$ ou $\mathbb{R}[X]$ ou $\mathbb{C}[X]$, deux éléments quelconques ont toujours des pgcd, et l'on a des résultats importants dans la pratique (théorèmes de Bézout, de Gauss,...) qui leur sont liés. On va les retrouver à la leçon suivante dans le cadre général des anneaux principaux, et plus particulièrement des anneaux euclidiens.

Leçon 17

Arithmétique dans les anneaux principaux

17.1 Théorème de Bézout et applications.

17.1.1 Théorème. *Soit A un anneau principal. Deux éléments quelconques de A admettent toujours des pgcd dans A. Plus précisément, quels que soient a et b dans A, tout générateur de l'idéal principal $aA + bA$ est un pgcd de a et b.*

$$(\ d \text{ est un pgcd de } a \text{ et } b\ ) \Leftrightarrow (\ aA + bA = dA\ )$$

Preuve. Soient $a, b \in A$ fixés. Comme A est principal, l'idéal $aA + bA$ est principal. Il existe $d \in A$ tel que $aA + bA = dA$. Montrons que d est un pgcd de a et b. On a d'abord $aA \subseteq aA + bA$, donc $aA \subseteq dA$, donc $d|a$. De même, $d|b$. Soit maintenant $c \in A$ tel que $c|a$ et $c|b$. Alors $aA \subseteq cA$ et $bA \subseteq cA$, donc, puisque cA est stable par addition, $aA + bA \subseteq cA$, c'est-à-dire $dA \subseteq cA$, et donc $c|d$. Ceci prouve que d est un pgcd de a et b. Réciproquement, soit d' un pgcd de a et b. D'après 16.4.4.(i), on a $d' \sim d$, donc $dA = d'A$, c'est-à-dire $d'A = aA + bA$. □

17.1.2 Théorème de Bézout. *Soit A un anneau principal. Pour tous a et b dans A, on a:*

$$(\ a \text{ et } b \text{ premiers entre eux dans } A\ ) \Leftrightarrow (\ \text{il existe } u, v \in A \text{ tels que } au + bv = 1\ )$$

Preuve. Soient $a, b \in A$ fixés. Supposons que a et b sont premiers entre eux. D'après 16.4.4.(ii), 1 est un pgcd de a et b. Il résulte alors du théorème précédent que $aA + bA = A$. En particulier $1 \in aA + bA$, et donc il existe $(u, v) \in A^2$ tel que $au + bv = 1$. Supposons réciproquement qu'il existe $u, v \in A$ tels que $au + bv = 1$; alors 1 appartient à $aA + bA$, donc $aA + bA = A$. Or, si d est un pgcd de a et b, on a $dA = aA + bA$. On déduit que $dA = A$, donc $d \in U(A)$, c'est-à-dire a et b premiers entre eux. □

17.1.3 Lemme de Gauss. *Soit A un anneau principal. Pour tous a, b, c dans A, on a:*

$$(\ a \text{ divise } bc, \text{ et } a \text{ premier avec } b\ ) \Rightarrow (\ a \text{ divise } c\ )$$

Preuve. Comme a et b sont premiers entre eux, il existe d'après le théorème de Bézout $u, v \in A$ tels que $au + bv = 1$. Donc $c = cau + cbv$. Comme a divise bc, on a $bc \in aA$, donc $bcv \in aA$. Par ailleurs il est clair que $acu \in aA$. Par stabilité de l'idéal aA pour l'addition, on conclut que $c = acu + bcv \in aA$. □

17.1.4 Remarques.

(a) Attention, si d est un pgcd de a et b, il existe d'après le théorème 17.1.1 des éléments $u, v \in A$ tels que $d = au + bv$. Le théorème de Bézout montre que la réciproque est vraie aussi si $d = 1$ (et donc plus généralement si $d \in U(A)$). Mais si $d \neq 1$, l'existence d'un couple (u, v) tel que $au + bv = d$ n'implique pas que d est le pgcd de a et b. Par exemple, dans $\mathbb{Z}$, on a $3 \times 10 + (-2) \times 14 = 2$, mais 2 n'est pas le pgcd de 3 et -2.

(b) Attention, dans le théorème de Bézout, il n'y a pas unicité du couple (u, v); le corollaire ci-dessous du lemme de Gauss détermine tous les couples (u, v) solutions.

17.1.5 Corollaire. (Une précision sur le théorème de Bézout). *Soit A un anneau principal. Soient a et b premiers entre eux dans A. Pour tout couple $(u, v) \in A^2$ tel que $au + bv = 1$, l'ensemble de tous les couples $(x, y) \in A^2$ tels que $ax + by = 1$ est égal à $\{(u, v) + c(-b, a)\,;\, c \in A\}$.*

Preuve. Soit $(u, v) \in A^2$ tel que $au + bv = 1$. Pour tout $c \in A$, le couple $(x, y) = (u, v) + c(-b, a) = (u - cb, v + ca)$ vérifie $ax + by = a(u - cb) + b(v + ca) = au - acb + bv + bca = au + bv = 1$. Réciproquement, quel que soit $(x, y) \in A^2$ tel que $ax + by = 1$, on a $ax + by = au + bv$, d'où $a(u - x) = b(y - v)$. Comme a et b sont premiers entre eux, il résulte du lemme de Gauss que a divise $y - v$. Il existe donc $c \in A$ tel que $y - v = ca$. On a alors: $cab = b(y - v) = a(u - x)$. Si $a \neq 0$, on déduit par intégrité de A que $u - x = cb$; on obtient donc bien $x = u - cb$ et $y = v + ca$. Si $a = 0$, alors $b \in U(A)$ et la propriété est claire. □

17.1.6 Remarques. Comme on le fait dans $\mathbb{Z}$, on définit naturellement la notion de ppcm (plus petit commun multiple) dans tout anneau principal A.

(a) On appelle ppcm de deux éléments $a, b \in A$ tout élément $m \in A$ tel que $aA \cap bA = mA$.

(b) (m est un ppcm de a et b) $\Leftrightarrow$ ($a|m$, $b|m$, et tout multiple de a et b est multiple de m).

(c) Le produit de tout pgcd de a et b par tout ppcm de a et b est associé à ab.

(d) En particulier (a et b sont premiers entre eux) $\Leftrightarrow$ (ab est un ppcm de a et b).

Comme la notion de pgcd, la notion de ppcm est clairement définie à l'association près, et s'étend naturellement à un nombre fini quelconque d'éléments de A.

17.2 Cas particulier des anneaux euclidiens.

17.2.1. Remarque préliminaire. Ce que l'on vient de voir sur l'arithmétique dans les anneaux principaux, basé sur les idéaux, s'applique en particulier aux anneaux euclidiens (voir 15.3.2). Néanmoins, dans ce cas particulier, on dispose de plus d'un processus algorithmique important basé sur la division euclidienne, appelé algorithme d'Euclide, qui permet entre autres de calculer les pgcd.

Quels que soient $a, b \in A$, on convient de désigner par $\mathrm{pgcd}(a, b)$ un pgcd quelconque de a et b. En d'autres termes, un élément $d \in A$ est un pgcd de a et b si et seulement si $d \sim \mathrm{pgcd}(a, b)$.

17.2.2. Lemme (fondamental de l'algorithme d'Euclide). *Soit A un anneau euclidien. Soient $a, b \in A$ tels que $b \neq 0$. Alors, pour tout reste r d'une division euclidienne de a par b, tout pgcd de a et b est associé à tout pgcd de b et r. En d'autres termes, en notant δ le stathme de A:*

$$(\ a = bq + r,\ \text{avec } r = 0 \ \text{ ou } \ \delta(r) < \delta(b)\ ) \ \Rightarrow\ (\ \mathrm{pgcd}(a, b) \sim \mathrm{pgcd}(b, r)\ ).$$

Preuve. Il résulte de l'égalité $a = bq + r$ que $a \in bA + rA$; comme $bA + rA$ est un idéal, on en déduit que $ax \in bA + rA$ pour tout $x \in A$, c'est-à-dire $aA \subset bA + rA$. Comme par ailleurs $bA \subset bA + rA$, la stabilité de $bA + rA$ pour l'addition implique alors $aA + bA \subset bA + rA$. En écrivant ensuite $r = a - bq$, on montre de même que $bA + rA \subset aA + bA$. Finalement $aA + bA = bA + rA$. Donc, en notant d un pgcd de a et b, et d' un pgcd de b et r, on a $dA = d'A$, c'est-à-dire $d \sim d'$. □

17.2.3 Théorème (algorithme d'Euclide). *Soient $a, b \in A$ non-nuls.*

(i) *il existe $k \in \mathbb{N}^*$ et des éléments $q_1, \ldots, q_k, r_0, r_1, \ldots, r_k \in A$, avec*

$$r_0 = b \neq 0,\ \ r_1 \neq 0,\ \ r_2 \neq 0,\ \ \ldots\ \ \ r_{k-2} \neq 0,\ \ r_{k-1} \neq 0,\ \ r_k = 0,$$

vérifiant la condition:

$$\delta(r_{k-1}) < \delta(r_{k-2}) < \cdots < \delta(r_2) < \delta(r_1) < \delta(r_0) = \delta(b),$$

et les égalités:

$$\begin{aligned}
a &= bq_1 + r_1 = r_0 q_1 + r_1,\\
r_0 &= r_1 q_2 + r_2,\\
r_1 &= r_2 q_3 + r_3,\\
&\cdots\cdots\cdots\\
r_{k-3} &= r_{k-2} q_{k-1} + r_{k-1},\\
r_{k-2} &= r_{k-1} q_k + r_k = r_{k-1} q_k.
\end{aligned}$$

(ii) *On a alors:* $\mathrm{pgcd}(a, b) \sim r_{k-1}$.

Preuve. On effectue la division euclidienne de a par b. Notons $a = bq_1 + r_1$ avec $r_1 = 0$ ou $\delta(r_1) < \delta(b)$.
Si $r_1 = 0$, on arrête.
Si $r_1 \neq 0$, on a $\delta(r_1) < \delta(b)$, et on effectue la division euclidienne de b par r_1.
Notons $b = r_1 q_2 + r_2$ avec $r_2 = 0$ ou $\delta(r_2) < \delta(r_1)$.
Si $r_2 = 0$, on arrête.

Si $r_2 \neq 0$, on a $\delta(r_2) < \delta(r_1) < \delta(b)$, et on effectue la division euclidienne de r_1 par r_2.

Notons $r_1 = r_2 q_3 + r_3$ avec $r_3 = 0$ ou $\delta(r_3) < \delta(r_2)$.

Si $r_3 = 0$, on arrête.

Si $r_3 \neq 0$, on a $\delta(r_3) < \delta(r_2) < \delta(r_1) < \delta(b)$, et on effectue la division euclidienne de r_2 par r_3.

On itère ainsi le processus. Comme il n'existe pas de suite strictement décroissante dans $\mathbb{N}$, il existe un rang $k \in \mathbb{N}^*$ tel que $r_k = 0$. En notant $r_0 = b$ pour la cohérence des notations, ceci prouve le point (i).

Pour (ii), remarquons que le lemme 17.2.2 appliqué dans la première égalité de (i) donne $\text{pgcd}(a,b) \sim \text{pgcd}(b,r_1) \sim \text{pgcd}(r_0,r_1)$. De même dans la deuxième égalité, on obtient $\text{pgcd}(r_0,r_1) \sim \text{pgcd}(r_1,r_2)$. Puis $\text{pgcd}(r_1,r_2) \sim \text{pgcd}(r_2,r_3)$, et par une récurrence évidente, $\text{pgcd}(a,b) \sim \text{pgcd}(r_{k-1},r_k)$. Or puisque r_k est nul, $\text{pgcd}(r_{k-1},r_k) \sim r_{k-1}$, ce qui achève la preuve. □

On traduit le point (ii) en disant que les pgcd de a et b sont les éléments associés au dernier reste non-nul dans la suite des divisions successives de a par b.

17.2.4 Exemple et remarque.

Dans l'anneau euclidien $\mathbb{Z}$, soient $a = 33810$ et $b = 4116$. La suite des divisions successives donne:

$$\underbrace{33810}_{a} = \underbrace{4116}_{b=r_0} \times 8 + \underbrace{882}_{r_1} \quad ; \quad \underbrace{4116}_{r_0} = \underbrace{882}_{r_1} \times 4 + \underbrace{588}_{r_2} \quad ; \quad \underbrace{882}_{r_1} = \underbrace{588}_{r_2} \times 1 + \underbrace{294}_{r_3} \quad ; \quad \underbrace{588}_{r_2} = \underbrace{294}_{r_3} \times 2 + 0$$

On conclut que $\text{pgcd}(a,b) \sim r_3$ donc $\text{pgcd}(33810, 4116) \sim 294$.

Remarque. L'algorithme d'Euclide permet non seulement de calculer $d = \text{pgcd}(a,b)$ mais aussi, en remontant les calculs dans la suite des divisions successives, de déterminer un couple (u,v) d'éléments de A tel que $d = au + bv$, faisant ainsi apparaître effectivement d comme un élément de $aA + bA$.

Ainsi, en reprenant l'exemple ci-dessus, on a:

$$d = 294 = 882 - 588 = 882 + (4 \times 882) - 4116 = 5 \times (33810 - 8 \times 4116) - 4116 = 5 \times 33810 - 41 \times 4116.$$

C'est un point crucial pour une bonne compréhension du théorème de Bézout.

17.3 Décomposition en produits d'éléments irréductibles.

17.3.1 Exemple préliminaire: décomposition d'un entier en produits de facteurs premiers.

On commence par rappeler un résultat classique sur les entiers, en en donnant une preuve directe, indépendante de la démonstration plus complexe de sa généralisation développée en 17.3.2.

Lemme. Dans l'anneau $\mathbb{Z}$:

(i) *Tout entier différent de 1 et de -1 admet au moins un diviseur premier.*

(ii) *Si un nombre premier divise un produit d'entiers, alors il divise l'un des facteurs de ce produit. En particulier, si un nombre premier divise un produit de nombres premiers, alors il est égal à l'un d'entre eux.*

Preuve. Il est clair qu'il suffit de prouver (i) pour un entier naturel n supérieur ou égal à 2. Notons $D(n)$ l'ensemble des diviseurs de n supérieurs ou égaux à 2. Il est non-vide car il contient n. Comme $\mathbb{N}$ est bien ordonné, $D(n)$ admet un plus petit élément p. C'est un diviseur de n, supérieur à 2; montrons qu'il est premier. Pour cela, considérons un entier naturel a divisant p. Par transitivité de la divisibilité, a divise n. Si $a \geq 2$, alors $a \in D(n)$, donc $p \leq a$ par minimalité de p; ainsi a divise p et $p \leq a$, donc $a = p$. Sinon, $a = 1$. Ceci prouve (i).

Soit p un nombre premier. Supposons que p divise le produit bc de deux entiers b et c. Supposons que p ne divise pas b. Alors, d'après 16.4.2, p est premier avec b. Ce qui, d'après lemme de Gauss, implique que p divise c. Ceci prouve la première assertion ; la seconde s'en déduit immédiatement. □

Théorème. *Soit n un entier supérieur ou égal à 2. Il existe, et ceci de façon unique, un entier $s \geq 1$, des nombres premiers $p_1, p_2, \ldots, p_s$ vérifiant $p_1 < p_2 < \cdots < p_s$, et des entiers naturels non-nuls $\alpha_1, \alpha_2, \ldots, \alpha_s$ tels que:*

$$n = p_1^{\alpha_1} p_2^{\alpha_2} \ldots p_s^{\alpha_s}.$$

Preuve. Montrons l'existence d'une telle décomposition par récurrence sur n. C'est clair si $n = 2$. Supposons-la vraie pour tout entier strictement inférieur à n. Soit p un diviseur premier de n (il en existe d'après le point (i) du lemme précédent). Si $n = p$, il n'y a rien à démontrer. Sinon, il existe $2 \leq n_0 \leq n-1$ tel que que $n = pn_0$. En appliquant l'hypothèse de récurrence à n_0, on a $n = pp_1^{\alpha_1}p_2^{\alpha_2}\dots p_s^{\alpha_s}$. S'il existe $1 \leq j \leq s$ tel que $p = p_j$, alors $n = p_1^{\alpha_1}p_2^{\alpha_2}\dots p_j^{\alpha_j+1}\dots p_s^{\alpha_s}$, d'où le résultat. Sinon, on obtient une décomposition du type voulu (en ordonnant p relativement aux p_i), avec $s+1$ facteurs, et un exposant 1 pour le facteur p.

Montrons l'unicité. Supposons pour cela que $n = p_1^{\alpha_1}p_2^{\alpha_2}\dots p_s^{\alpha_s} = q_1^{\beta_1}q_2^{\beta_2}\dots q_t^{\beta_t}$, avec $s \in \mathbb{N}^*$, $p_1 < p_2 < \dots < p_s$ premiers, $\alpha_i \in \mathbb{N}^*$ pour tout $1 \leq i \leq s$, et $t \in \mathbb{N}^*$, $q_1 < q_2 < \dots < q_t$ premiers, $\beta_j \in \mathbb{N}^*$ pour tout $1 \leq j \leq t$. D'après le point (ii) du lemme précédent, chaque p_i ($1 \leq i \leq s$) est égal à un des q_j ($1 \leq j \leq t$), et chaque q_j est égal à l'un des p_i. Comme les p_i sont à deux distincts, ainsi que les q_j, on a nécessairement $s = t$. De plus la condition de croissance sur les p_i et les q_j implique que l'on a précisément $p_1 = q_1, p_2 = q_2, \dots, p_s = q_s$. Donc finalement: $n = p_1^{\alpha_1}p_2^{\alpha_2}\dots p_s^{\alpha_s} = p_1^{\beta_1}p_2^{\beta_2}\dots p_s^{\beta_s}$. Si $\beta_1 > \alpha_1$, on en déduit l'égalité $p_2^{\alpha_2}\dots p_s^{\alpha_s} = p_1^{\beta_1-\alpha_1}p_2^{\beta_2}\dots p_s^{\beta_s}$; celle-ci implique que p_1 divise le produit $p_2^{\alpha_2}\dots p_s^{\alpha_s}$, ce qui est impossible d'après le point (ii) du lemme précédent. De même $\beta_1 < \alpha_1$ conduit à une contradiction. C'est donc que $\alpha_1 = \beta_1$. On prouve de façon analogue que $\alpha_i = \beta_i$ pour tout $1 \leq i \leq s$. □

Quitte à multiplier par ± 1, le théorème s'applique à tout entier différent de 0, 1 et -1. Ainsi, tout entier non-nul et non inversible s'écrit au signe près comme un produit de nombres premiers, de façon unique à l'ordre près des facteurs. En rappelant 16.3.8.(a), cette propriété est un cas particulier du théorème général suivant.

17.3.2 Théorème. *Soit A un anneau principal. Tout élément non-nul et non inversible se décompose en un produit d'un nombre fini d'éléments irréductibles dans A. Cette décomposition est unique à l'ordre près des facteurs et au produit par un élément inversible près.*

Explicitement, cela signifie que, dans tout anneau principal A, on a:

(1) tout élément $a \in A$, $a \neq 0$, $a \notin U(A)$, s'écrit $a = r_1r_2\dots r_n$, avec $r_1, r_2, \dots, r_n$ irréductibles dans A;

(2) si $r_1r_2\dots r_n = s_1s_2\dots s_m$, avec $r_1, \dots, r_n, s_1, \dots, s_m$ irréductibles dans A, alors $m = n$, et il existe une permutation $\sigma \in S_n$ telle que $s_i \sim r_{\sigma(i)}$ pour tout $1 \leq i \leq n$.

Preuve de l'unicité de la décomposition. Supposons que $r_1r_2\dots r_n = s_1s_2\dots s_m$ avec r_i et s_j irréductibles dans A pour tous $1 \leq i \leq n$ et $1 \leq j \leq m$. Comme A est principal, il résulte de 16.3.7 que l'élément r_1 est premier car irréductible, et comme il divise $s_1s_2\dots s_m$, il existe $1 \leq j \leq m$ tel que r_1 divise s_j. On a donc $s_j = ar_1$ pour un certain $a \in A$. Comme s_j est irréductible et que $r_1 \notin U(A)$, on a $a \in U(A)$, c'est-à-dire $r_1 \sim s_j$. Par intégrité, on simplifie par r_1 pour obtenir $r_2\dots r_n \sim s_1\dots s_{j-1}s_{j+1}\dots s_m$. On réitère, et le résultat voulu s'en déduit par récurrence. □

Preuve de l'existence de la décomposition. Raisonnons par l'absurde, en supposant que A ne vérifie pas (1). Cela signifie que l'ensemble:

$$R = \{a \in A,\ a \neq 0,\ a \notin U(A),\ a \text{ n'est pas produit d'éléments irréductibles dans } A\}$$

est non-vide. Il en résulte qu'est également non-vide l'ensemble $E = \{aA\ ;\ a \in R\}$ des idéaux principaux de A engendrés par les éléments de R.

On montre que E est inductif (voir 3.3.1 du chapitre 3) pour l'inclusion. Pour cela, considérons $F = (I_k)_{k\in X}$ une famille d'éléments de E totalement ordonnée par l'inclusion. Pour tout $k \in X$, considérons un élément $a_k \in R$ tel que $I_k = a_kA$. La réunion $I = \bigcup_{k\in X} I_k$ est un idéal non-nul de A. Comme A est principal, il existe $b \in A$, $b \neq 0$, tel que $I = bA$. Puisque $b \in I$, il existe $a_k \in R$ tel que $b \in a_kA$, et donc $bA \subseteq a_kA$. Comme par ailleurs $a_kA \subseteq I = bA$, on conclut que $I = a_kA$, et donc $I \in E$. En résumé, toute famille d'éléments de E totalement ordonnée admet un plus grand élément. On conclut que E est inductif.

D'après le lemme de Zorn, E admet (au moins) un élément maximal; notons-le cA, avec $c \in R$. Parce que $c \in R$, il est non-nul, non-inversible, et non-irréductible. Donc il existe $x, y \in A$ tel que $c = xy$ avec $x \notin U(A)$ et $y \notin U(A)$. Il en résulte que $cA \subset xA$ avec $cA \neq xA$, et $cA \subset yA$ avec $cA \neq yA$. De plus, il est clair que $x \in R$ ou $y \in R$ (en effet, sinon, x et y seraient produits d'irréductibles, et donc $c = xy$ aussi), d'où $xA \in E$ ou $yA \in E$. Dans l'un ou l'autre cas, il y a contradiction avec la maximalité de cA dans E. □

17.4 Exemples d'applications de la décomposition en éléments irréductibles.

17.4.1 Remarques préliminaires sur les notations. Soit A un anneau principal.

(a) Dans l'ensemble des éléments irréductibles de A, l'association définit d'après 16.3.4 une relation d'équivalence. En choisissant dans chaque classe d'équivalence un représentant particulier, on définit un *système de représentants* $\mathcal{R}$ des éléments irréductibles. En d'autres termes, tout élément irréductible de A est équivalent à un unique élément irréductible de la famille $\mathcal{R}$.

quel que soit r irréductible dans A, il existe $r' \in \mathcal{R}$ et $u \in U(A)$ uniques tels que $r = ur'$.

Exemples.

1. Dans l'anneau $\mathbb{Z}$, on choisit généralement comme système de représentants des éléments irréductibles l'ensemble $\mathcal{P}$ des nombres premiers positifs. Tout élément irréductible de $\mathbb{Z}$ est de la forme εp avec $p \in \mathcal{P}$ et $\varepsilon \in U(\mathbb{Z}) = \{-1, +1\}$.
2. Dans l'anneau $\mathbb{C}[X]$, on choisit généralement comme système de représentants des éléments irréductibles l'ensemble $\mathcal{R}$ des polynômes de degré 1 unitaires (ie. de coefficient dominant égal à 1). Tout élément irréductible est de la forme $\alpha(X-\beta)$ avec $X-\beta \in \mathcal{R}$ et $\alpha \in U(\mathbb{C}[X]) = \mathbb{C}^*$.

(b) Soit $\mathcal{R}$ un système de représentants des éléments irréductibles dans A. Soit $a \in A$ non-nul et non-inversible. Il résulte de la condition (1) de 17.3.2 que a s'écrit de façon unique, à l'ordre près des facteurs:

$$a = u\, r_1^{n_1} r_2^{n_2} \dots r_s^{n_s}, \quad \text{où } u \in U(A),\ r_i \in \mathcal{R} \text{ (avec } r_i \neq r_j \text{ si } i \neq j),\ n_i \in \mathbb{N}^*.$$

Exemples.

1. Dans $\mathbb{Z}$, tout élément a non-nul et distinct de ± 1 s'écrit de façon unique $a = \varepsilon\, p_1^{n_1} p_2^{n_2} \dots p_s^{n_s}$, où $\varepsilon = \pm 1$, $n_i \in \mathbb{N}^*$, $p_i \in \mathcal{P}$ (avec $p_i \neq p_j$ si $i \neq j$) pour tout $1 \leq i \leq s$..
2. Dans $\mathbb{C}[X]$, tout polynôme $P(X)$ de degré ≥ 1 s'écrit de façon unique:
$$P(X) = \alpha(X-\beta_1)^{n_1}(X-\beta_2)^{n_2} \dots (X-\beta_s)^{n_s},$$
où $\alpha \in \mathbb{C}^*$, $n_i \in \mathbb{N}^*$, $\beta_i \in \mathbb{C}$ (avec $\beta_i \neq \beta_j$ si $i \neq j$) pour tout $1 \leq i \leq s$.

(c) Soit $\mathcal{R}$ un système de représentants des éléments irréductibles dans A. Soient $a, b \in A$ non-nuls et non-inversibles. En réunissant les facteurs irréductibles intervenant dans l'écriture ci-dessus de a et dans celle de b, et en autorisant alors des exposants nuls dans l'une des décompositions, a et b s'écrivent de façon unique:

$$a = u\, r_1^{n_1} r_2^{n_2} \dots r_q^{n_q} \quad \text{et} \quad b = v\, r_1^{m_1} r_2^{m_2} \dots r_q^{m_q}, \quad \text{où } u, v \in U(A),$$
$$r_i \in \mathcal{R} \text{ (avec } r_i \neq r_j \text{ si } i \neq j),\ \ n_i \in \mathbb{N},\ \ m_i \in \mathbb{N},\ \ (n_i, m_i) \neq (0,0) \ \text{ pour tout } 1 \leq i \leq q.$$

17.4.2 Lemme (expression des diviseurs d'un élément). *Soit A un anneau principal. Soit $a \in A$, non-nul et non-inversible. Avec la notation du* 17.4.1.(b) *ci-dessus, les diviseurs de a dans A sont tous les éléments de la forme:*

$$w r_1^{p_1} r_2^{p_2} \dots r_s^{p_s}, \quad 0 \leq p_i \leq n_i \text{ pour tout } 1 \leq i \leq s,\ \ w \in U(A).$$

Preuve. Soit b un diviseur de a. Si $b \in U(A)$, le résultat est clair avec $b = w$ et $p_1 = p_2 = \dots = p_s = 0$. Supposons donc maintenant que $b \notin U(A)$. Soit r un des facteurs irréductibles intervenant dans la décomposition de b. Comme $b|a$, on a $r|a$, c'est-à-dire que r divise $r_1^{n_1} r_2^{n_2} \dots r_s^{n_s}$. Puisque r est premier (car irréductible, voir 16.3.7), on en tire que r est associé à l'un des r_i. Ceci prouve que b est de la forme $b = w\, r_1^{p_1} r_2^{p_2} \dots r_s^{p_s}$, avec $w \in U(A)$ et $p_i \geq 0$ pour tout $1 \leq i \leq s$.

Pour montrer que $p_i \leq n_i$ pour tout $1 \leq i \leq s$, raisonnons par l'absurde. Supposons par exemple (pour fixer les idées) que $p_1 > n_1$. En notant $a = xb$ avec $x \in A$, on aurait donc: $u\, r_2^{n_2} \dots r_s^{n_s} = xw\, r_1^{p_1 - n_1} r_2^{p_2} \dots r_s^{p_s}$, avec $p_1 - n_1 > 0$, ce que contredirait la condition (2) de 17.3.2. Ce qui achève la preuve. □

17.4.3 Proposition (expression du pgcd et du ppcm). *Soit A un anneau principal. Si a et b sont deux éléments de A non-nuls et non-inversibles donnés par les notations* 17.4.1.(c)*, on a:*

$$\mathrm{pgcd}(a,b) \sim r_1^{h_1} r_2^{h_2} \dots r_q^{h_q} \quad \textit{et} \quad \mathrm{ppcm}(a,b) \sim r_1^{\ell_1} r_2^{\ell_2} \dots r_q^{\ell_q},$$

avec $h_i = \min(n_i, m_i)$ et $\ell_i = \max(n_i, m_i)$ pour tout $1 \leq i \leq q$.

Preuve. Soient $a, b \in A$. Si $a = 0$, on a $\mathrm{pgcd}(a,b) \sim b$. Si $a \in U(A)$, on a $\mathrm{pgcd}(a,b) \sim a \sim 1$. De même si $b = 0$ ou $b \in U(A)$. Sinon, a et b sont non-nuls et non-inversibles: le résultat résulte alors de 17.4.2 et de la définition des pgcd et ppcm. □

Leçon 18

Arithmétique dans les anneaux de polynômes

18.1 Rappels.

• Soit A un anneau commutatif unitaire. On note $A[X]$ l'anneau des polynômes en une indéterminée X à coefficients dans A. Pour tout $P \in A[X]$ non-nul, il existe un unique entier naturel n et un unique $(n+1)$-uplet $(a_0, a_1, \ldots, a_n)$ d'éléments de A tels que:

$$P = a_n X^n + a_{n-1} X^{n-1} + \cdots + a_1 X + a_0 \quad \text{et} \quad a_n \neq 0.$$

L'entier n est le degré de P, noté $\deg P$. L'élément non-nul a_n de A est le coefficient dominant de P, noté $\mathrm{cd}(P)$. Par convention, on pose $\deg 0 = -\infty$. On a pour tous P et Q dans $A[X]$:

$$\deg(P+Q) \leq \max(\deg P, \deg Q) \quad \text{et} \quad \deg(PQ) \leq \deg P + \deg Q.$$

Un polynôme de $A[X]$ est dit unitaire lorsque son coefficient dominant est égal à 1.

• Si A est intègre, alors : $\deg(PQ) = \deg P + \deg Q$ pour tous $P, Q \in A[X]$,
d'où il résulte :

d'une part que $A[X]$ est intègre (mais n'est jamais un corps),

d'autre part que $U(A[X]) = U(A)$.

• On déduit immédiatement du second point que:

PROPOSITION. *Pour tout anneau intègre A, deux polynômes P et Q de $A[X]$ sont associés si et seulement s'il existe $c \in U(A)$ tel que $P = cQ$ (et alors $Q = c^{-1}P$).*

On en déduit aussi que la définition 16.3.1 se reformule de la façon suivante:

Si A est intègre, un polynôme P est irréductible dans $A[X]$ lorsque $P \notin U(A)$ et P vérifie la condition suivante:

$$\textit{si } P = QR \textit{ avec } Q, R \in A[X], \textit{ alors } Q \in U(A) \textit{ ou } R \in U(A).$$

EXERCICE. Montrer que, si A est intègre, tout polynôme de la forme $X - a$, avec $a \in A$, est irréductible dans $A[X]$.

• Concrètement, l'arithmétique dans $A[X]$ diverge fortement entre la situation la plus favorable où A est un corps (et où tout se passe plus ou moins comme dans $\mathbb{Z}$ en raison de la division euclidienne) et les cas beaucoup plus délicats où A n'est pas un corps (et où donc $A[X]$ n'est même pas principal). Ceci résulte évidemment de l'équivalence démontrée en 15.3.5 :

$$A[X] \text{ est euclidien } \Leftrightarrow \ A[X] \text{ est principal } \Leftrightarrow \ A \text{ est un corps}$$

18.2 Arithmétique dans $K[X]$ lorsque K est un corps.

18.2.1 SYNTHÈSE. On suppose dans tout ce paragraphe que K est un corps. Listons à titre d'exercices de révision quelques unes des principales propriétés arithmétiques des anneaux euclidiens vues dans les leçons précédentes lorsqu'on les traduit pour l'anneau $K[X]$.

(a) On a dans $K[X]$ une division euclidienne: pour tous $F, G \in K[X]$, avec $G \neq 0$, il existe $Q, R \in K[X]$ uniques tels que $F = GQ + R$ et $\deg R < \deg G$.

(b) Un polynôme G divise un polynôme F dans $K[X]$ (ou encore F est un multiple de G) lorsqu'il existe $Q \in K[X]$ tel que $F = QG$, ce qui, lorsque $G \neq 0$, équivaut à dire que le reste de la division euclidienne de F par G est nul. Si G divise F, l'idéal principal engendré par F est inclus dans l'idéal principal engendré par G.

(c) Un polynôme P est inversible dans $K[X]$ si et seulement s'il appartient à K^*, ce qui équivaut à $\deg P = 0$.

(d) Deux polynômes $F, G \in K[X]$ sont associés lorsque F divise G et G divise F, c'est-à-dire lorsqu'ils engendrent le même idéal. Comme $U(K[X]) = K^*$, il est clair que F et G sont associés si et seulement s'il existe $c \in K^*$ tel que $F = cG$. En particulier, tout polynôme non-nul de $K[X]$ est associé à un polynôme unitaire (il suffit de le multiplier par l'inverse de son coefficient dominant).

(e) Il y a équivalence dans $K[X]$ entre élément premier et élément irréductible. Un polynôme irréductible dans $K[X]$ est un élément $P \in K[X]$ avec $P \notin K$, tel que les seuls diviseurs de P dans $K[X]$ sont les éléments de K^* et les polynômes cP ($c \in K^*$) associés à P.

(f) Tout polynôme $P \in K[X]$ tel que $P \notin K$ se décompose en un produit d'un nombre fini d'éléments irréductibles dans $K[X]$; cette décomposition est unique à l'ordre près des facteurs et au produit par un élément de K^* près.

(g) Deux polynômes $P, Q \in K[X]$ sont premiers entre eux lorsque les seuls diviseurs communs à P et Q sont les éléments de K^*. On a dans $K[X]$ le théorème de Bézout:

(P et Q premiers entre eux dans $K[X]$) $\Leftrightarrow$ (il existe $F, G \in K[X]$ tels que $PF + QG = 1$),

et le lemme de Gauss: (P divise QR et P premier avec R dans $K[X]$) $\Rightarrow$ (P divise Q dans $K[X]$).

(h) Deux polynômes quelconques $P, Q \in K[X]$ admettent un pgcd (défini à la multiplication par un élément de K^* près), que l'on peut calculer en utilisant l'algorithme d'Euclide dans $K[X]$, ou la décomposition de P et de Q en produit de facteurs irréductibles.

18.2.2 Eléments irréductibles dans $K[X]$.

Proposition. *Soit K un corps.*

(i) *Tout polynôme de degré 1 est irréductible dans $K[X]$.*

(ii) *Si $K = \mathbb{C}$, les polynômes irréductibles dans $\mathbb{C}[X]$ sont les polynômes de degré 1.*

(iii) *Si $K = \mathbb{R}$, les polynômes irréductibles dans $\mathbb{R}[X]$ sont les polynômes de degré 1 et les polynômes de degré 2 n'admettant pas de zéros dans $\mathbb{R}$.*

Preuve. Soit P de degré 1 dans $K[X]$ tel que $P = QR$; on a $1 = \deg P = \deg QR = \deg Q + \deg R$, donc Q ou R est de degré 0 c'est-à-dire est inversible dans $K[X]$. Supposons maintenant que $K = \mathbb{C}$. Si $\deg P \geq 2$, il admet forcément un zéro dans $\mathbb{C}$ (c'est la propriété fondamentale de $\mathbb{C}$), donc d'après 15.4.3, on a $P = (X-a)Q$ avec $a \in \mathbb{C}$ et $Q \in \mathbb{C}[X]$, $\deg Q \geq 1$. Ceci prouve que P n'est pas irréductible. Par contraposée, on en déduit le point (ii). Supposons enfin que $K = \mathbb{R}$ et considérons $P \in \mathbb{R}[X]$. On a $P \in \mathbb{C}[X]$ et, pour tout zéro a de P dans $\mathbb{C}$ qui n'appartient pas à $\mathbb{R}$, on a $P(\overline{a}) = \overline{P(a)} = 0$, donc $\overline{a}$ est aussi un zéro de P dans $\mathbb{C}$, et le polynôme $(X-a)(X-\overline{a})$ est alors un élément de $\mathbb{R}[X]$ sans zéro réel. Le point (iii) s'en déduit. □

Donnons un exemple de situation où la recherche des éléments irréductibles dans $A[X]$ se ramène à celle des éléments irréductibles dans $K[X]$ pour K un corps déduit de A (par passage au quotient). On fixe un nombre premier p. A tout polynôme F dans $\mathbb{Z}[X]$, on associe sa réduction modulo p, qui est le polynôme $\overline{F}$ dans $(\mathbb{Z}/p\mathbb{Z})[X]$ défini par:

$$\text{si } F = \sum_{i=0}^{n} a_i X^i \text{ avec } a_i \in \mathbb{Z}, \text{ alors } \quad \overline{F} = \sum_{i=0}^{n} \overline{a}_i X^i, \text{ avec } \overline{a}_i \in \mathbb{Z}/p\mathbb{Z}.$$

Remarquons d'une part que $\mathbb{Z}/p\mathbb{Z}$ est un corps (car p premier), et d'autre part que si F est unitaire et s'écrit $F = PQ$ avec P, Q de degrés ≥ 1 dans $\mathbb{Z}[X]$, alors $\overline{F} = \overline{P}\,\overline{Q}$ avec $\deg \overline{P} = \deg P$ et $\deg \overline{Q} = \deg Q$. Par contraposée, on en déduit:

Proposition. *Soit F un polynôme unitaire dans $\mathbb{Z}[X]$. S'il existe un nombre premier p tel que la réduction de F modulo p soit irréductible dans $(\mathbb{Z}/p\mathbb{Z})[X]$, alors F est irréductible dans $\mathbb{Z}[X]$.*

Exemple. Soit $F = X^5 - 2X^4 - 4X^3 + 3X^2 + 6X + 5 \in \mathbb{Z}[X]$. Sa réduction modulo 2 est $\overline{F} = X^5 + X^2 + \overline{1}$ dans $(\mathbb{Z}/2\mathbb{Z})[X]$. Il est facile de vérifier par identifictaion que $\overline{F}$ ne peut s'exprimer dans $(\mathbb{Z}/2\mathbb{Z})[X]$ ni comme le produit d'un polynôme de degré 1 par un polynôme de degré 4, ni comme le produit d'un polynôme de degré 2 par un polynôme de degré 3 ; il est donc irréductible dans $(\mathbb{Z}/2\mathbb{Z})[X]$, et donc F est irréductible dans $\mathbb{Z}[X]$.

Attention: cette méthode de réduction modulo p est loin de tout régler ! Il existe des polynômes irréductibles dans $\mathbb{Z}[X]$ mais dont la réduction modulo p n'est pas irréductible dans $(\mathbb{Z}/p\mathbb{Z})[X]$, quel que soit le nombre premier p.

18.3 Eléments irréductibles dans $A[X]$ lorsque A est principal.

18.3.1 Définition. Soit A un anneau principal. Soit P un élément de $A[X]$ tel que $P \notin A$. On appelle *contenu* de P, noté $c(P)$, un pgcd dans A des coefficients de P.

Remarque. La notion de contenu n'est définie qu'à l'association dans A près. Lorsque l'on écrit $c(P) = a$, on a aussi $c(P) = ua$ pour tout $u \in U(A)$. On peut aussi écrire $c(P) \sim a$.

18.3.2 Définition. Soit A un anneau principal. Un polynôme P dans $A[X]$ est dit *primitif* lorsque $\deg P \geq 1$ et lorsque ses coefficients sont premiers entre eux.

$$(\ P \text{ primitif }) \ \Leftrightarrow \ (\ \deg P \geq 1 \ \text{ et } \ c(P) = 1\ )$$

Remarques.

(1) Tout polynôme unitaire est primitif.

(2) Tout polynôme $P \in A[X]$ tel que $P \notin A$ s'écrit $P = c(P)P_1$ avec P_1 primitif.

(3) Si $A = K$ tout polynôme de degré ≥ 1 est primitif dans $K[X]$.

18.3.3 Lemme. *Soit A un anneau principal. Soient P_1 et P_2 primitifs dans $A[X]$. Soient a_1 et a_2 non-nuls dans A. Si $a_1P_1 = a_2P_2$, alors a_1 et a_2 sont associés dans A, et P_1 et P_2 sont associés dans $A[X]$.*

Preuve. Comme P_1 est primitif, on a $c(a_1P_1) = a_1$. De même $c(a_2P_2) = a_2$. Donc a_1 et a_2 sont deux pgcd des coefficients du polynôme $a_1P_1 = a_2P_2$. Ils sont donc associés dans A: il existe $u \in U(A)$ tel que $a_2 = ua_1$. On a alors $a_1P_1 = ua_1P_2$, ce qui par intégrité de $A[X]$ (puisque A est intègre) implique que $P_1 = uP_2$. Comme u est un élément inversible de $A[X]$, on conclut que P_1 et P_2 sont associés. □

18.3.4 Lemme (Gauss). *Soit A un anneau principal. Soient P et Q deux éléments de $A[X]$. D'une part P et Q sont primitifs si et seulement si PQ est primitif. D'autre part $c(PQ) = c(P)c(Q)$.*

Preuve. Supposons que P et Q soient primitifs et que PQ ne le soit pas. Comme $c(PQ)$ n'est pas inversible dans l'anneau principal A, il est divisible par au moins un élément p irréductible et donc premier. Considérons l'anneau intègre $B = A/pA$. La surjection canonique $\pi : A \to B$ se prolonge canoniquement en un morphisme d'anneaux $\hat{\pi} : A[X] \to B[X]$ défini par $\hat{\pi}(\sum a_iX^i) = \sum \pi(a_i)X^i$. Comme $c(P) = 1$, l'élément p ne divise pas tous les coefficients de P, donc $\hat{\pi}(P) \neq 0$. De même, $\hat{\pi}(Q) \neq 0$. L'intégrité de B impliquant celle de $B[X]$, on en déduit que $\hat{\pi}(P)\hat{\pi}(Q) \neq 0$, c'est-à-dire $\hat{\pi}(PQ) \neq 0$. Or, p divise $c(PQ)$, donc tous les coefficients de PQ, donc $\hat{\pi}(PQ) = 0$. D'où une contradiction. On a ainsi montré que P et Q primitifs implique PQ primitif.

Réciproquement, supposons PQ primitif. On peut toujours écrire P et Q sous la forme $P = c(P)P_1$ et $Q = c(Q)Q_1$ avec P_1 et Q_1 primitifs. Alors P_1Q_1 est primitif d'après ce qui précède, et l'égalité $PQ = c(P)c(Q)P_1Q_1$ implique avec le lemme 18.3.3 que $c(P)c(Q)$ est associé à 1 dans A, c'est-à-dire inversible dans A. D'où $c(P) \in U(A)$ et $c(Q) \in U(A)$, de sorte que P et Q sont primitifs.

Enfin, plus généralement, en notant $P = c(P)P_1$, $Q = c(Q)Q_1$ et $PQ = c(PQ)S_1$ avec P_1, Q_1, S_1 primitifs, l'égalité $c(PQ)S_1 = c(P)c(Q)P_1Q_1$ implique, puisque P_1Q_1 est primitif d'après le début de la preuve, que $c(PQ)$ est associé à $c(P)c(Q)$ dans A, ce que l'on a convenu d'écrire aux éléments inversibles près $c(PQ) = c(P)c(Q)$. □

18.3.5 Lemme. *Soient A un anneau principal et K son corps de fractions. Tout polynôme $P \in K[X]$ tel que $P \notin K$ peut s'écrire $P = qP_1$, avec $q \in K^*$ et $P_1 \in A[X]$ primitif dans $A[X]$.*

Preuve. Notons $P = \sum_{i=0}^{n} \frac{a_i}{s_i}X^i$ avec $n \geq 1$, $a_i \in A$, s_i non-nuls dans A, et $a_n \neq 0$. Quitte à multiplier le numérateur et le dénominateur de chaque fraction $\frac{a_i}{s_i}$ par un même élément non-nul de A, on peut écrire toutes les fractions $\frac{a_i}{s_i}$ avec un même dénominateur s (par exemple un ppcm des s_i puisque cette notion existe dans l'anneau principal A, ou encore simplement le produit des s_i), sous la forme $\frac{a_i}{s_i} = \frac{a'_i}{s}$, avec $a'_i \in A$. Donc $P = \frac{1}{s}Q$ où $Q = \sum_{i=0}^{n} a'_iX^i \in A[X]$. Alors $Q = c(Q)Q_1$ avec Q_1 primitif, et donc $P = qQ_1$ en posant $q = \frac{c(Q)}{s} \in K^*$. □

18.3.6 Théorème. *Soient A un anneau principal et K son corps de fractions. Soit R un élément non-nul de $A[X]$.*

(i) *Si $R \in A$; alors R est irréductible dans $A[X]$ si et seulement s'il est irréductible dans A.*

(ii) *Si $R \notin A$; alors R est irréductible dans $A[X]$ si et seulement s'il est primitif dans $A[X]$ et irréductible dans $K[X]$.*

Preuve. Rappelons que $U(A[X]) = U(A)$ puisque A est intègre.

(i) Supposons $R \in A$. Notons alors $R = r$. Supposons d'abord r irréductible dans A. En particulier $r \notin U(A)$ donc $r \notin U(A[X])$. Si P et Q dans $A[X]$ sont tels que $r = PQ$, on a $0 = \deg r = \deg P + \deg Q$ donc $P \in A$ et $Q \in A$, de sorte que l'irréductibilité de r dans A implique $P \in U(A)$ ou $Q \in U(A)$, c'est-à-dire $P \in U(A[X])$ ou $Q \in U(A[X])$, ce qui prouve que r est irréductible en tant qu'élément de $A[X]$. Supposons maintenant que r est irréductible dans $A[X]$. En particulier $r \notin U(A[X])$ donc $r \notin U(A)$. Si $a, b \in A$ sont tels que $r = ab$, alors cette égalité dans $A[X]$ implique $a \in U(A[X])$ ou $b \in U(A[X])$, c'est-à-dire $a \in U(A)$ ou $b \in U(A)$, ce qui prouve que r est irréductible en tant qu'élément de A.

(ii) Supposons R de degré non-nul dans $A[X]$, primitif dans $A[X]$, et irréductible dans $K[X]$. Si P et Q dans $A[X]$ sont tels que $R = PQ$, comme R est irréductible dans $K[X]$, on a P ou Q dans $U(K[X]) = K^*$. Mais P et Q étant à coefficients dans A, cela signifie que P ou Q appartient à A^*. Considérons le cas où $P \in A$, $P \neq 0$. Dans $A[X]$, on peut toujours écrire $Q = c(Q)Q_1$ avec Q_1 primitif. On a l'égalité $R = Pc(Q)Q_1$ avec $Pc(Q) \in A$, Q_1 primitif dans $A[X]$ et R primitif dans $A[X]$. On en déduit avec le lemme 18.3.3 que $Pc(Q) \in U(A)$. D'où a fortiori $P \in U(A)$, ou encore $P \in U(A[X])$. De même $Q \in A$, $Q \neq 0$, implique $Q \in U(A[X])$. On a ainsi montré que R est irréductible dans $A[X]$.

Récipoquement, supposons R de degré non-nul irréductible dans $A[X]$. Ecrivons-le sous la forme $R = c(R)R_1$ avec R_1 primitif dans $A[X]$, de même degré que R; l'irréductibilité de R implique alors R_1 ou $c(R)$ inversible dans $A[X]$. Comme $\deg R_1 = \deg R \geq 1$, le premier cas est exclu, donc $c(R) \in U(A[X])$, c'est-à-dire $c(R) \in U(A)$, et donc R est primitif dans $A[X]$. Pour montrer maintenant que R est irréductible dans $K[X]$, considérons P et Q dans $K[X]$ tels que $R = PQ$. Raisonnons par l'absurde en supposant que P et Q ne sont pas dans K; ils sont d'après le lemme 18.3.5 de la forme $P = \frac{a}{b}P_1$ et $Q = \frac{c}{d}Q_1$ avec a, b, c, d non-nuls dans A, et P_1, Q_1 primitifs dans $A[X]$, de mêmes degrés strictement positifs que P et Q respectivement. L'égalité $R = PQ$ devient $bdR = acP_1Q_1$. Or R est primitif dans $A[X]$ comme on vient de le voir, et P_1Q_1 l'est aussi d'après le lemme 18.3.4. En appliquant le lemme 18.3.3, on déduit que R est associé à P_1Q_1 dans $A[X]$. Il existe donc $u \in U(A[X]) = U(A)$ tel que $R = uP_1Q_1$. Comme R est supposé irréductible dans $A[X]$, il en résulte que P_1 ou Q_1 appartient à $U(A[X]) = U(A)$, ce qui contredit l'hypothèse faite selon laquelle P et Q sont de degrés strictement positifs. C'est donc que P ou Q appartient à $U(K[X]) = K^*$, ce qui achève de prouver que R est irréductible dans $K[X]$. □

18.4 Une application : critère d'irréductibilité d'Eisenstein

18.4.1 Théorème. *Soit A un anneau principal. Soit $P = a_nX^n + a_{n-1}X^{n-1} + \cdots + a_1X + a_0$ un élément de $A[X]$ de degré $n \geq 1$. On suppose qu'il existe dans A un élément p, premier dans A, et satisfaisant les trois conditions suivantes:*

$$p \text{ divise } a_0, a_1, \ldots, a_{n-1}, \qquad p \text{ ne divise pas } a_n, \qquad p^2 \text{ ne divise pas } a_0.$$

(i) *Alors P est irréductible dans $K[X]$, où K désigne le corps de fractions de A.*

(ii) *Si de plus P est primitif dans $A[X]$ (en particulier s'il est unitaire dans $A[X]$), alors P est irréductible dans $A[X]$.*

Preuve. On montre d'abord le point (ii). Supposons donc P primitif. Par l'absurde, supposons P non irréductible dans $A[X]$: il existe donc $Q, R \in A[X]$ tels que $P = QR$, avec $0 < \deg Q < \deg P$ et $0 < \deg R < \deg P$. Comme P est primitif, le lemme 18.3.4 implique que Q et R le sont.

Posons $Q = \sum_{i=0}^{q} b_i X^i$ et $R = \sum_{i=0}^{r} c_i X^i$, avec $b_i, c_i \in A$, et $0 < q < n$ et $0 < r < n$. On a $a_n = b_q c_r \neq 0$, et l'hypothèse p ne divise pas a_n implique que p ne divise pas b_q et ne divise pas c_r. On a aussi $a_0 = b_0 c_0$, et donc par hypothèse p divise $b_0 c_0$ mais p^2 ne divise pas $b_0 c_0$, ce qui implique que p ne divise pas b_0 ou p ne divise pas c_0. Si l'on est dans le cas où p ne divise pas b_0, alors p divise c_0 en utilisant le fait que p est premier dans A. On a vu que p ne divise pas c_r, et on peut donc considérer le plus petit entier $k \in \{1, \ldots, r\}$ tel que p ne divise pas c_k. Par construction, p ne divise pas $b_0 c_k$, et p divise $b_i c_{k-i}$ pour tout $i \in \{1, \ldots, k\}$. Il en résulte que p ne divise pas $a_k = b_0 c_k + b_1 c_{k-1} + \cdots + b_k c_0$. Comme $1 \leq k \leq r < n$, ceci est contraire aux hypothèses faites au départ sur P. C'est donc que P est irréductible dans $A[X]$.

On ne suppose plus maintenant que P est primitif. Notons $P = c(P)P_1$ avec P_1 primitif. Comme $c(P)$ est un pgcd des a_i (pour $0 \leq i \leq n$), il existe $a'_0, a'_1, \ldots, a'_n$ premiers entre eux dans leur ensemble tels que $a_i = c(P)a'_i$ pour tout $0 \leq i \leq n$. Donc $P_1 = a'_n X^n + \cdots + a'_1 X + a'_0$. On a clairement p qui ne divise pas a'_n (sinon il diviserait $a_n = c(P)a'_n$) et p^2 qui ne divise pas a'_0 (par le même argument). Pour $0 \leq i \leq n-1$, p divise $a_i = c(P)a'_i$ avec p qui ne divise pas $c(P)$ (car sinon p diviserait en particulier a_n, ce qui est exclu), et donc p divise a'_i. Les coefficients a'_i du polynôme primitif P_1 vérifiant donc les conditions du critère, on peut appliquer à P_1 la première étape, et conclure que P_1 est irréductible dans $A[X]$. D'après le point (ii) du théorème 18.3.6, il s'ensuit que P_1 est irréductible dans $K[X]$. En multipliant par $c(P) \in K^* = U(K[X])$, il en est de même de $c(P)P_1 = P$. □

18.4.2 Exemples: $P = X^5 + 4X^3 + 12X + 2$ est unitaire donc primitif dans $\mathbb{Z}[X]$, et il est irréductible dans $\mathbb{Z}[X]$ par application du critère d'Eisenstein.

Pour tout entier $n \geq 1$ et tout nombre premier p, le polynôme unitaire $X^n - p$ est irréductible dans $\mathbb{Z}[X]$ et $\mathbb{Q}[X]$. Il existe donc en tout degré une infinité de polynômes irréductibles deux à deux non associés.

18.4.3 Exemple: polynômes cyclotomiques. Pour tout nombre premier p, on appelle p-ième polynôme cyclotomique le polynôme $\Phi_p = 1 + X + X^2 + \cdots + X^{p-1}$. C'est une notion qui intervient dans de nombreux problèmes liés aux solutions des équations algébriques et des extensions de corps.

Notons P le polynôme défini par $P(X) = \Phi_p(X+1)$. Il est clair par contraposition que P est irréductible dans $\mathbb{Q}[X]$ si et seulement si Φ_p l'est. L'irréductibilité de Φ_p ou P dans $\mathbb{Q}[X]$ équivaut à leur irréductibilité dans $\mathbb{Z}[X]$ puiqu'il s'agit de polynômes unitaires donc primitifs. On a :

$$P(X) = 1 + (1+X) + (1+X)^2 + \cdots + (1+X)^{p-1} = \tfrac{1-(1+X)^p}{1-(1+X)} = \sum_{i=1}^{p} \binom{p}{i} X^{i-1} = \sum_{j=0}^{p-1} \binom{p}{j+1} X^j.$$

Or il est bien connu (rappelez la preuve...) que le coefficient binomial $\binom{p}{i}$ est divisible par p pour tout $1 \leq i \leq p-1$. On peut donc appliquer directement le critère d'Eisenstein pour déduire que P est irréductible dans $\mathbb{Q}[X]$. On conclut:

pour tout nombre premier p, le polynôme cyclotomique Φ_p est irréductible dans $\mathbb{Q}[X]$.

Un dernier commentaire. Conformément aux programmes retenus pour la troisième année de licence, nous nous sommes limités dans les paragraphes 18.3 et 18.4 à des anneaux de polynômes $A[X]$ à coefficients dans un anneau principal A. Mais tout le contenu de ces paragraphes est en fait vrai sans aucun changement dans une classe d'anneaux plus vaste (contenant strictement les anneaux principaux), celle des anneaux factoriels.

Leçon 19

Polynômes d'endomorphismes

L'objet de cette leçon est de développer une application des propriétés de l'anneau des polynômes en une indéterminée à coefficients dans un corps K (dans la pratique $\mathbb{R}$ ou $\mathbb{C}$) pour des questions d'algèbre linéaire liées à la réduction des endomorphismes d'un K-espace vectoriel. On fixe donc dans toute cette leçon un K-espace vectoriel E que l'on suppose de dimension finie $n = \dim E \geq 1$.

19.1 Polynômes d'endomorphismes, polynômes de matrices.

19.1.1. Algèbre d'endomorphismes. Considérons l'ensemble $\operatorname{End} E$ des endomorphismes de l'espace vectoriel E. On sait (cours d'algèbre linéaire de deuxième année) que $\operatorname{End} E$ est un K-espace vectoriel (on peut additionner deux endomorphismes et considérer le produit externe d'un endomorphisme $u \in \operatorname{End} E$ par un scalaire $\lambda \in K$). Mais $\operatorname{End} E$ est également (voir 11.1.2.b) un anneau (non-commutatif) unitaire pour les lois $+$ et $\circ$. Cette double structure (de K-espace vectoriel et d'anneau), avec la propriété de cohérence (évidente à vérifier):

$$\lambda.(u \circ v) = (\lambda.u) \circ v = u \circ (\lambda.v) \text{ pour tous } u, v \in \operatorname{End} E, \lambda \in K$$

correspond à ce l'on appelle une structure de K-algèbre. On retiendra que:

$$(\operatorname{End} E, +, \circ, .) \textit{ est une } K\textit{-algèbre, non-commutative, unitaire.}$$

Rappelons en particulier que dans $\operatorname{End} E$, l'élément neutre pour le produit interne est id_E, et que l'on note $u^m = u \circ u \circ \ldots \circ u$, avec m facteurs.

19.1.2 Algèbre de matrices carrées. De même, en notant $\mathcal{M}_n(K)$ l'ensemble des matrices carrées d'ordre n à coefficients dans K, on a:

$$(\mathcal{M}_n(K), +, \times, .) \textit{ est une } K\textit{-algèbre, non-commutative, unitaire, isomorphe à } \operatorname{End} E.$$

En effet, tout choix d'une base $\mathcal{B}$ de E permet de considérer la bijection $m : \operatorname{End} E \to \mathcal{M}_n(K)$ associant à tout endormorphisme u sa matrice $m(u)$ par rapport à la base $\mathcal{B}$, qui vérifie $m(u+v) = m(u) + m(v)$, $m(u \circ v) = m(u) \times m(v)$, $m(\lambda.u) = \lambda.m(u)$ pour tous $u, v \in \operatorname{End} E, \lambda \in K$, et réalise donc un isomorphisme d'algèbres unitaires. En particulier, concernant l'élément neutre du produit interne, on a bien $m(\mathrm{id}_E) = I_n$.

19.1.3 Algèbre de polynômes. Considérons maintenant l'anneau commutatif unitaire $K[X]$ des polynômes en une indéterminée X à coefficients dans K. On peut aussi définir le produit externe d'un scalaire $\lambda \in K$ par un polynôme P, en posant $\lambda.P = \lambda \times P$, produit du polynôme constant λ par le polynôme P dans $K[X]$. On vérifie aisément qu'alors:

$$K[X] \textit{ est une } K\textit{-algèbre, commutative, unitaire.}$$

19.1.4 Théorème.

(i) *Pour tout $u \in \operatorname{End} E$ fixé, il existe un unique morphisme d'algèbres $\varphi_u : K[X] \to \operatorname{End} E$ tel que $\varphi_u(X) = u$. Il est défini par $P \mapsto P(u)$ où l'on pose:*

$$\textit{si } P = \alpha_m X^m + \alpha_{m-1} X^{m-1} + \cdots + \alpha_1 X + \alpha_0, \quad \textit{avec } \alpha_i \in K \textit{ pour tout } 0 \leq i \leq m, \textit{ alors:}$$
$$P(u) = \alpha_m.u^m + \alpha_{m-1}.u^{m-1} + \cdots + \alpha_1.u + \alpha_0.\,\mathrm{id}_E.$$

(ii) *Pour tout $A \in \mathcal{M}_n(K)$ fixé, il existe un unique morphisme d'algèbres $\psi_A : K[X] \to \mathcal{M}_n(K)$ tel que $\psi_A(X) = A$. Il est défini par $P \mapsto P(A)$ où l'on pose:*

$$\textit{si } P = \alpha_m X^m + \alpha_{m-1} X^{m-1} + \cdots + \alpha_1 X + \alpha_0, \quad \textit{avec } \alpha_i \in K \textit{ pour tout } 0 \leq i \leq m, \textit{ alors:}$$
$$P(A) = \alpha_m.A^m + \alpha_{m-1}.A^{m-1} + \cdots + \alpha_1.A + \alpha_0.I_n.$$

Preuve. Montrons d'abord l'unicité. Supposons que φ soit un morphisme d'algèbres $K[X] \to \operatorname{End} E$ tel que $\varphi(X) = u$. Si $P = \sum_{i=1}^{m} \alpha_i X^i$ est un élément de $K[X]$ quelconque, alors $\varphi(P) = \sum_{i=1}^{m} \varphi(\alpha_i X^i) = \sum_{i=1}^{m} \alpha_i \varphi(X)^i = \sum_{i=1}^{m} \alpha_i u^i$, ce qui prouve que $\varphi = \varphi_u$. Il est clair réciproquement que φ_u est bien un morphisme d'algèbre $K[X] \to \operatorname{End} E$ tel que $\varphi_u(X) = u$, ce qui prouve (i). Le (ii) est analogue. □

19.1.5 Remarques

(1) *Attention* au fait que $\varphi_u(\alpha_0) = \varphi_u(\alpha_0.1_K) = \alpha_0.\varphi_u(1_K) = \alpha_0.\operatorname{id}_E$, et de même $\psi_A(\alpha_0) = \alpha_0.I_n$.

(2) Avec les notations de 19.1.2, on a $\psi_A = m \circ \varphi_u$ pour tout choix d'une base de E avec $m(u) = A$.

19.2 Idéal d'annulation et polynôme minimal.

19.2.1 Proposition et définition.

– Pour tout endomorphisme $u \in \operatorname{End} E$, l'ensemble N_u des polynômes $P \in K[X]$ tels que $P(u) = O$ est un idéal non-nul de $K[X]$, appelé l'idéal d'annulation de l'endomorphisme u.

– Pour toute matrice $A \in \mathcal{M}_n(K)$, l'ensemble N_A des polynômes $P \in K[X]$ tels que $P(A) = O_n$ est un idéal non-nul de $K[X]$, appelé l'idéal d'annulation de la matrice A.

Preuve. Avec les notations de 19.1.4, et en rappelant que O désigne l'endomorphisme nul de E, on a $N_u = \operatorname{Ker} \varphi_u$. Comme φ_u est un morphisme d'anneaux, son noyau N_u est un idéal de $K[X]$. En tant que K-espaces vectoriels, $K[X]$ n'est pas de dimension finie alors que $\operatorname{End} E$ est de dimension n^2, donc φ_u n'est pas injectif, et donc l'idéal J_u n'est pas nul. La preuve est identique pour une matrice A, en rappelant que O_n désigne la matrice nulle. □

19.2.2 Proposition et définition.

– Pour tout endomorphisme $u \in \operatorname{End} E$, il existe un unique polynôme unitaire $Q_u \in K[X]$ tels que N_u soit l'idéal principal engendré par Q_u dans $K[X]$. Le polynôme unitaire Q_u est appelé le polynôme minimal de l'endormorphisme u.

– Pour toute matrice $A \in \mathcal{M}_n(K)$, il existe un unique polynôme unitaire $Q_A \in K[X]$ tels que N_A soit l'idéal principal engendré par Q_A dans $K[X]$. Le polynôme unitaire Q_A est appelé le polynôme minimal de la matrice A.

Preuve. Comme K est un corps, l'anneau $K[X]$ est principal (voir 15.3.5) donc l'idéal N_u est principal non-nul. Il résulte de 16.2.2 et 16.2.3.c qu'il existe un unique polynôme unitaire Q_u tel que $N_u = Q_u K[X]$. La preuve est identique pour une matrice A. □

19.2.3 Remarques

(1) Tout endomorphisme u de E annule son polynôme minimal Q_u, et un polynôme $P \in K[X]$ est annulé par u si et seulement s'il est multiple dans $K[X]$ du polynnôme minimal Q_u de u :

$$\begin{cases} Q_u(u) = O, \\ \text{et} \\ \text{pour tout } P \in K[X],\ (P(u) = O) \Leftrightarrow (\text{ il existe } R \in K[X] \text{ tel que } P = RQ_u), \end{cases}$$

avec une formulation analogue en termes de matrices.

(2) Pour tout choix d'une base de E, si l'on a $u \in \operatorname{End} E$ et $A \in \mathcal{M}_n(K)$ avec $m(u) = A$ (voir remarque (2) de 19.1.5), alors $N_u = N_A$ et $Q_u = Q_A$.

19.2.4 Exemples.

(1) Soient F et H deux sous-espaces vectoriels non-nuls de E tels que $E = F \oplus H$. Soit s la symétrie par rapport à F parallèlement à H. On sait que l'on a $s \circ s = \operatorname{id}_E$ dans $\operatorname{End} E$, donc s annule $X^2 - 1$, d'où $X^2 - 1 \in N_s$, et donc Q_s divise $X^2 - 1$. Mais ici $s \neq \operatorname{id}_E$ puisque $H \neq \{0_E\}$ et $s \neq -\operatorname{id}_E$ puisque $F \neq \{0_E\}$, d'où $X - 1 \notin N_s$ et $X + 1 \notin N_s$, et donc Q_s ne divise pas $X - 1$ ni $X + 1$. On conclut que *le polynôme minimal de la symétrie s est* $Q_s = X^2 - 1$.

(2) En supposant toujours que $E = F \oplus H$ avec F et H non-nuls, et en considérant cette fois la projection p de E sur F parallèlement à H, qui vérifie $p \circ p = p$, $p \neq \operatorname{id}_E$ et $p \neq O$, on vérifie de même que *le polynôme minimal de la projection p est* $Q_p = X^2 - X$.

(3) Si u est un endomorphisme nilpotent d'ordre p (ie. $u^p = O$ et $u^k \neq O$ pour $1 \leq k \leq p - 1$), alors le polynôme minimal de u est X^p. (On peut montrer que nécessairement $p \leq n = \dim E$).

19.3 Polynôme minimal et valeurs propres.

19.3.1 Quelques rappels (voir cours d'algèbre linéaire des années précédentes). Soit u un endomorphisme de E. Une *valeur propre* de E dans K est un scalaire $\lambda \in K$ tel qu'il existe un vecteur $x \in E$ non-nul vérifiant $u(x) = \lambda.x$. Un tel vecteur non-nul x est alors appelé un *vecteur propre* associé à la valeur propre λ. Si λ est une valeur propre de E dans K, alors le sous-espace vectoriel $E_\lambda = \mathrm{Ker}(u - \lambda\,\mathrm{id}_E) = \{x \in E\,;\, u(x) = \lambda.x\}$, qui est donc l'ensemble des vecteurs propres associés à λ auquel on adjoint le vecteur nul 0_E, est appelé le *sous-espace propre* associé à la valeur propre λ. On appelle *polynôme caractéristique* de u le polynôme $P_u = \det(u - X.\,\mathrm{id}_E) \in K[X]$. Son degré est égal à la dimension n de E. Les zéros de P_u dans K sont exactement les valeurs propres de u dans K. La multiplicité d'une valeur propre λ dans K est l'exposant avec lequel le facteur $(X - \lambda)$ apparaît dans la décomposition du polynôme P_u en produit de facteurs irréductibles dans $K[X]$.

Le résultat suivant a également été démontré en deuxième année:

19.3.2 Théorème (de Cayley-Hamilton). *Tout endomorphisme u de E annule son polynôme caractéristique.*

En d'autres termes, pour tout $u \in \mathrm{End}\,E$, on a $P_u(u) = O$, ou encore $P_u \in N_u$. D'où l'on déduit immédiatement, par définition de Q_u comme polynôme engendrant l'idéal principal N_u (voir 19.2.3.1):

19.3.3 Corollaire. *Pour tout endomorphisme u de E, le polynôme caractéristique P_u est un multiple du polynôme minimal Q_u dans $K[X]$.*

19.3.4 Proposition. *Pour tout endomorphisme u de E, les zéros dans K du polynôme minimal Q_u sont exactement les valeurs propres de u dans K.*

Preuve. D'après le corollaire 19.3.2, il existe un polynôme R tel que $P_u = RQ_u$ dans $K[X]$. Si $\lambda \in K$ un zéro de Q_u, on a $Q_u(\lambda) = 0$, donc $P_u(\lambda) = R(\lambda)Q_u(\lambda)$, et donc λ est une valeur propre de u.

Réciproquement, soit λ une valeur propre de u. Il existe donc un vecteur non-nul x de E tel que $u(x) = \lambda.x$. En composant par u, on en déduit $u^2(x) = u(\lambda.x) = \lambda.u(x) = \lambda^2.x$, puis $u^3(x) = \lambda^3.x$ et finalement $u^j(x) = \lambda^j.x$ pour tout $j \geq 0$.

Posons $Q_u = X^m + \beta_{m-1}X^{m-1} + \cdots + \beta_1 X + \beta_0$, où $m = \deg Q_u \leq \deg P_u = n$ et où les β_i appartiennent à K. Par définition du polynôme minimal, $Q_u(u)$ est l'endomorphisme nul O de E. Donc, pour tout vecteur $x \in E_\lambda$, on a:

$$\begin{aligned} 0_E = Q_u(u)(x) &= \left(u^m + \beta_{m-1}u^{m-1} + \cdots + \beta_1 u + \beta_0\,\mathrm{id}_E\right)(x) \\ &= u^m(x) + \beta_{m-1}u^{m-1}(x) + \cdots + \beta_1 u(x) + \beta_0 x = \lambda^m.x + \beta_{m-1}\lambda^{m-1}.x + \cdots + \beta_1\lambda.x + \beta_0.x \\ &= Q_u(\lambda).x. \end{aligned}$$

Il suffit de choisir le vecteur x non-nul dans E_λ pour conclure que $Q_u(\lambda)$ est nul dans K. □

En d'autres termes: *P_u et Q_u ont exactement les mêmes zéros dans K, qui sont les valeurs propres de u dans K.* En outre, P_u étant un multiple de Q_u, on a pour toute valeur propre λ de u:

$$1 \;\leq\; \begin{pmatrix}\text{la multiplicité de } \lambda \text{ en}\\ \text{tant que zéro de } Q_u\end{pmatrix} \;\leq\; \begin{pmatrix}\text{la multiplicité de } \lambda \text{ en}\\ \text{tant que zéro de } P_u\end{pmatrix}.$$

Notons enfin que tout ce que l'on vient de formuler en termes de polynôme minimal et de valeurs propres d'un endomorphisme peut être exprimé de façon analogue en termes de polynôme minimal et de valeurs propres d'une matrice carrée.

Exemple. Supposons que l'on ait une matrice $A \in \mathcal{M}_5(K)$ tel que $P_A = -(X-2)^3(X+1)^2$.

Alors, a priori, Q_A peut valoir:

$$\begin{array}{llllll} (X-2)(X+1), & \text{ou} & (X-2)^2(X+1), & \text{ou} & (X-2)^3(X+1), \\ (X-2)(X+1)^2, & \text{ou} & (X-2)^2(X+1)^2, & \text{ou} & (X-2)^3(X+1)^2. \end{array}$$

Pour déterminer ce que vaut effectivement Q_A:

on calcule $(A - 2I_5)(A + I_5)$.

Si ce produit est nul, c'est fini: $Q_A = (X-2)(X+1)$.

Sinon, on calcule $(A-2I_5)^2(A+I_5)$.

Si ce produit est nul, c'est fini: $Q_A = (X-2)^2(X+1)$.

Sinon, on calcule $(A-2I_5)^3(A+I_5)$.

Si ce produit est nul, c'est fini: $Q_A = (X-2)^3(X+1)$.

Sinon, on calcule $(A-2I_5)(A+I_5)^2$.

Si ce produit est nul, c'est fini: $Q_A = (X-2)(X+1)^2$.

Sinon, on calcule $(A-2I_5)^2(A+I_5)^2$.

Si ce produit est nul, c'est fini: $Q_A = (X-2)^2(X+1)^2$.

Sinon, on sait d'après le théorème de Cayley-Hamilton que $(A-2I_5)^3(A+I_5)^2 = O_5$.

Donc dans ce dernier cas, $Q_A = (X-2)^3(X+1)^2 = -P_A$.

On conçoit que de tels calculs directs sont vite fastidieux, voire inextricables à la main pour des matrices un peu grandes. D'où l'importance d'arguments théoriques plus généraux.

19.4 Lemme des noyaux et diagonalisabilité

19.4.1 Lemme fondamental (dit "lemme des noyaux"). *Soient $F_1, F_2, \dots, F_p$ des polynômes deux à deux premiers entre eux dans $K[X]$. Soit u un endomorphisme de E. On a:*

$$\operatorname{Ker}\Big(F_1(u)\circ F_2(u)\circ\cdots\circ F_p(u)\Big) = \operatorname{Ker} F_1(u) \oplus \operatorname{Ker} F_2(u) \oplus \cdots \oplus \operatorname{Ker} F_p(u)$$

Preuve. On considère dans $K[X]$ le polynôme $F = F_1 \times F_2 \times \cdots \times F_p$ et, pour tout $1 \leq i \leq p$, $G_i = F/F_i = \prod_{j\neq i} F_j$. Parce que les F_j sont deux à deux premiers entre eux, les G_i sont premiers entre eux dans leur ensemble, c'est-à-dire qu'il n'existe pas de polynôme non constant dans $K[X]$ qui divise tous les G_i. La propriété de Bézout (qui se généralise sans difficulté à plus de deux termes, voir 17.1) implique qu'il existe des polynômes $H_1, H_2, \dots, H_p$ tels que $H_1G_1 + H_2G_2 + \cdots + H_pG_p = 1$ dans $K[X]$. On a donc:

$$H_1(u)\circ G_1(u) + H_2(u)\circ G_2(u) + \cdots + H_p(u)\circ G_p(u) = \mathrm{id}_E \quad \text{dans } \operatorname{End} E. \qquad (*)$$

Ceci étant, montrons l'égalité voulue sur les noyaux. Posons d'abord $F(u) = F_1(u)\circ F_2(u)\circ\cdots\circ F_p(u)$. Comme les $F_i(u)$ commutent entre eux pour la loi $\circ$ dans $\operatorname{End} E$, on a $F(u) = G_i(u)\circ F_i(u)$ pour tout $1 \leq i \leq n$. Si un vecteur x appartient à $\operatorname{Ker} F_i(u)$, on a $F(u)(x) = G_i(u)(F_i(u)(x)) = G_i(u)(0_E) = 0_E$, donc $x \in \operatorname{Ker} F(u)$. Réciproquement, prenons $x \in \operatorname{Ker} F(u)$. D'après (*), on a $x = \sum_{i=1}^p H_i(u)(G_i(u)(x))$. Mais $F_i(u)\circ H_i(u)\circ G_i(u) = H_i(u)\circ G_i(u)\circ F_i(u) = H_i(u)\circ F(u)$ pour tout $1 \leq i \leq p$, et donc $F_i(u)(H_i(u)(G_i(u)(x))) = H_i(u)(F(u)(x)) = H_i(u)(0_E) = 0_E$ en utilisant le fait que $x \in \operatorname{Ker} F(u)$. Ceci prouve que $H_i(u)(G_i(u)(x)) \in \operatorname{Ker} F_i(u)$ pour tout $1 \leq i \leq p$. On en déduit que $x \in \operatorname{Ker} F_1(u) + \cdots + \operatorname{Ker} F_p(u)$, ce qui prouve l'inclusion $\operatorname{Ker} F(u) \subset \operatorname{Ker} F_1(u) + \cdots + \operatorname{Ker} F_p(u)$. L'inclusion rciproque provient de $\operatorname{Ker} F_i(u) \subset \operatorname{Ker} F(u)$ pour tout $1 \leq i \leq p$ établie ci-dessus. Ainsi, on obtient:

$$\operatorname{Ker} F(u) = \operatorname{Ker} F_1(u) + \operatorname{Ker} F_2(u) + \cdots + \operatorname{Ker} F_p(u).$$

Il reste à vérifier que cette somme est directe. Prenons $x_1 \in \operatorname{Ker} F_1(u), \cdots, x_p \in \operatorname{Ker} F_p(u)$ tels que $x_1 + \cdots + x_p = 0_E$. Pour tous $1 \leq i \neq j \leq n$, on a $G_i(u)(x_j) = 0_E$ puisque $G_i = \prod_{k\neq i} F_k$ fait nécessairement intervenir le facteur F_j. Dès lors, à partir de (*), on a:

$$x_j = \sum_{i=1}^{p} H_i(u)(G_i(u)(x_j)) = H_j(u)(G_j(u)(x_j)) = H_j(u)\Big(G_j(u)(-\sum_{k\neq j} x_k)\Big)$$

$$= -H_j(u)\Big(\sum_{k\neq j} G_j(u)(x_k)\Big) = -H_j(u)(0_E) = 0_E, \qquad \square$$

19.4.2 Théorème. *Un endomorphisme u de E est diagonalisable sur K si et seulement s'il existe des éléments $\lambda_1, \lambda_2, \dots, \lambda_p$ de K tous distincts tels que le polynôme $(X-\lambda_1)(X-\lambda_2)\cdots(X-\lambda_p)$ de $K[X]$ soit annulé par u dans* $\operatorname{End} E$.

Preuve. Supposons que u est diagonalisable. Rappelons (voir cours d'algèbre linéaire de deuxième année) que cela signifie que u admet des valeurs propres $\lambda_1, \lambda_2, \dots, \lambda_s$ (supposées par notation deux à deux distinctes) telles que $E = E_1 \oplus E_2 \oplus \cdots \oplus E_s$, où $E_j = \operatorname{Ker}(u - \lambda_j.\,\mathrm{id}_E)$ est le sous-espace propre associé à λ_j, pour tout $1 \leq j \leq s$. Introduisons dans $K[X]$ les polynômes $F = (X-\lambda_1)(X-\lambda_2)\cdots(X-\lambda_s)$

et $G_i = F/(X-\lambda_i) = \prod_{j\neq i}(X-\lambda_j)$ pour tout $1 \leq i \leq s$. On a $F(u) = G_i(u) \circ (u - \lambda_i.\,\mathrm{id}_E)$. Donc $F(u)(x) = 0_E$ pour tout $x \in E_i$. Donc $F(u)(x) = 0_E$ pour tout $x \in E$ puisque E est somme directe des E_i. On conclut que $F(u)$ est l'endomorphisme nul, ce qui prouve le résultat voulu (avec $p = s$).

Réciproquement, supposons satisfaite la condition de l'énoncé. Les polynômes $(X - \lambda_i)$ sont deux à deux premiers entre eux dans $K[X]$. On applique le lemme des noyaux pour déduire que $E = \bigoplus_{i=1}^{p} \mathrm{Ker}(u-\lambda_i.\,\mathrm{id}_E)$, donc u est diagonalisable, ses sous-espaces propres étant ceux des $\mathrm{Ker}(u-\lambda_i.\,\mathrm{id}_E)$ qui ne sont pas nuls (leur nombre s peut être a priori $\leq p$). □

19.4.3 Corollaire. *Un endomorphisme u de E est diagonalisable sur K si et seulement si son polynôme minimal est de la forme $Q_u = (X-\lambda_1)(X-\lambda_2)\cdots(X-\lambda_s)$, où $\lambda_1, \lambda_2, \ldots, \lambda_s$ sont les valeurs propres distinctes de u dans K.*

Preuve. Découle immédiatement de 19.4.2, 19.3.4, et de la définition du polynôme minimal. □

19.5 Sous-espaces caractéristiques

19.5.1 Définition. Soit u un endomorphisme de E. Soit λ une valeur propre de u dans K. On note q la multiplicité de λ. Rappelons que cela signifie que le polynôme caractéristique P_u est divisible dans $K[X]$ par $(X-\lambda)^q$, mais pas par $(X-\lambda)^{q+1}$. En outre, on sait (voir cours de deuxième année) que la multiplicité q est supérieure ou égal à la dimension du sous-espace propre E_λ, que u est diagonalisable sur K si et seulement si ces deux entiers coïncident pour chacune des valeurs propres de u.

On définit le *sous-espace caractéristique* associé à la valeur propre λ, noté F_λ, comme le noyau de l'endomorphisme $(u - \lambda.\,\mathrm{id}_E)^q$.

Il est clair que, pour toute valeur propre λ de u dans K, on a:

$$F_\lambda = \mathrm{Ker}(u-\lambda.\,\mathrm{id}_E)^q \quad \supseteq \quad E_\lambda = \mathrm{Ker}(u-\lambda.\,\mathrm{id}_E) \quad \neq \quad \{0_E\}$$

19.5.2 Proposition. *Soient u un endomorphisme de E. Pour toute valeur propre λ de u dans K, la dimension du sous-espace caractéristique F_λ est égale à la multiplicité q de la valeur propre λ.*

Preuve. Par définition de la multiplicité q, le polynôme caractéristique de u est de la forme $P_u = (X-\lambda)^q F$ avec $F \in K[X]$ tel que $F(\lambda) \neq 0$. Introduisons le sous-espace vectoriel $H = \mathrm{Ker}\, F(u)$ de E. Les polynômes $(X-\lambda)^q$ et F étant premiers entre eux dans $K[X]$, il résulte du lemme des noyaux que $E = F_\lambda \oplus H$. Le sous-espace vectoriel H est stable par u ; en effet, si $x \in H$, alors $F(u)(x) = 0_E$, or $F(u)(u(x)) = u(F(u)(x)) = u(0_E) = 0_E$ donc $u(x) \in H$. De même F_λ est stable par u. On peut donc considére la restriction v de u à F_λ et la restriction w de u à H, et l'on a (propriété classique du polynôme caractéristique, il suffit de choisir une base adaptée à la décomposition en somme directe et de faire la calcul des déterminants par blocs): $P_u = P_v P_w$ dans $K[X]$ (avec $P_w = 1$ dans le cas où $H = \{0_E\}$).

Notons $d = \dim F_\lambda$. Considérons $v' = v - \lambda.\,\mathrm{id}_{F_\lambda}$, qui est la restriction de $u - \lambda.\,\mathrm{id}_E$ à F_λ ; il est clair que c'est un endomorphisme nilpotent de F_λ, d'ordre $\leq q$, et donc $P_{v'} = (-1)^d X^d$, ce qui revient à dire que $P_v = (-1)^d (X-\lambda)^d$. Par ailleurs, par définition de H et de w, on a $F(w) = O$. Donc F est un multiple dans $K[X]$ du polynôme minimal Q_w de w. Comme $F(\lambda) \neq 0$, on a forcément $Q_w(\lambda) \neq 0$, ce qui prouve avec 19.3.4 que λ n'est pas une valeur propre de w, d'où $P_w(\lambda) \neq 0$.

En résumé, $P_u = (-1)^d (X-\lambda)^d P_w$ avec $P_w(\lambda) \neq 0$, ce qui signifie que d est la multiplicité de la valeur propre λ. □

19.5.3 Données. On suppose que P_u se décompose sur K en produit de facteurs de degré 1 (c'est-à-dire que A est trigonalisable sur K, voir cours d'algèbre linéaire de deuxième année) ce qui est toujours possible si $K = \mathbb{C}$, et en désignant par $\lambda_1, \lambda_2, \ldots, \lambda_s$ les valeurs propres *distinctes* de u dans K, on a:

$$P_u = (-1)^n (X-\lambda_1)^{q_1}(X-\lambda_2)^{q_2}\cdots(X-\lambda_s)^{q_s}, \quad n = \deg P_u = q_1 + q_2 + \cdots + q_s,$$
$$Q_u = \qquad (X-\lambda_1)^{p_1}(X-\lambda_2)^{p_2}\cdots(X-\lambda_s)^{p_s}, \quad m = \deg Q_u = p_1 + p_2 + \cdots + p_s,$$

avec $1 \leq p_i \leq q_i \leq n$ pour tout $1 \leq i \leq s$.

Pour tout $1 \leq i \leq s$, on note

$E_i = \mathrm{Ker}(u - \lambda_i.\,\mathrm{id}_E)$ le sous-espace propre associé à λ_i, qui vérifie $\dim E_i \leq q_i$,

$F_i = \mathrm{Ker}(u - \lambda_i.\,\mathrm{id}_E)^{q_i}$ le sous-espace caractéristique associé à λ_i, qui vérifie $\dim F_i = q_i$.

19.5.4 Proposition. *On reprend toutes les données et notations précédentes.*

(1) *E est somme directe des sous-espaces caractéristiques: $E = F_1 \oplus F_2 \oplus \cdots \oplus F_s$.*

(2) *Pour toute valeur propre λ_i de A, on a: $F_i = \mathrm{Ker}(u - \lambda_i.\,\mathrm{id}_E)^{p_i}$.*

Preuve. D'après le théorème de Cayley-Hamilton, on a $P_u(u) = O$, donc $E = \mathrm{Ker}\, P_u(u)$. On applique alors le lemme des noyaux pour conclure que $E = \bigoplus_{i=1}^{s} \mathrm{Ker}(u - \lambda_i.\,\mathrm{id}_E)^{q_i}$, ce qui prouve le point (1).

On a aussi $Q_u(u) = O$, donc de la même façon avec le lemme des noyaux, on a $E = \bigoplus_{i=1}^{s} \mathrm{Ker}(u - \lambda_i.\,\mathrm{id}_E)^{p_i}$. Ainsi, en rappelant que $F_i = \mathrm{Ker}(u - \lambda_i.\,\mathrm{id}_E)^{q_i}$ et en introduisant $H_i = \mathrm{Ker}(u - \lambda_i.\,\mathrm{id}_E)^{p_i}$ pour tout $1 \leq i \leq s$, on a $H_i \subseteq F_i$ puisque $p_i \leq q_i$, et $E = \bigoplus_{i=1}^{s} F_i = \bigoplus_{i=1}^{s} H_i$. On conclut que $F_i = H_i$ pour tout $1 \leq i \leq s$. □

19.5.5 Corollaire *Avec les hypothèses et données précédentes, les conditions suivantes sont équivalentes:*

(i) *u est diagonalisable,*

(ii) *pour toute valeur propre λ_i de u, on a $F_i = E_i$,*

(iii) *pour toute valeur propre λ_i de u, on a $p_i = 1$,*

(iv) *le polynôme minimal de u n'a que des termes de degré 1: $Q_u = (X - \lambda_1) \cdots (X - \lambda_s)$.*

Preuve. Il est clair que (iii) ⇔ (iv). Comme u diagonalisable signifie que $E = E_1 \oplus E_2 \oplus \cdots \oplus E_s$, et comme chaque E_i est inclus dans F_i, l'équivalence (i) ⇔ (ii) résulte du point (1) du théorème précédent. L'implication (iii) ⇒ (ii) résulte du point (2) du théorème précédent. La réciproque (i) ⇒ (iv) découle de 19.4.2. □

19.5.6 Suite des noyaux. Pour toute valeur propre λ_i de u, on appelle suite des noyaux associée à λ_i la suite croissante des sous-espaces vectoriels:

$$\underbrace{E_i}_{\dim = r_i} = \mathrm{Ker}(u - \lambda_i.\,\mathrm{id}_E) \subseteq \mathrm{Ker}(u - \lambda_i.\,\mathrm{id}_E)^2 \subseteq \cdots \subseteq \mathrm{Ker}(u - \lambda_i.\,\mathrm{id}_E)^{q_i} = \underbrace{F_i}_{\dim = q_i}.$$

Cette suite est donc formée de q_i sous-espaces, mais d'après le point (2) du théorème 19.5.4, elle est stationnaire à partir de $\mathrm{Ker}(u - \lambda_i.\,\mathrm{id}_E)^{p_i} = \cdots = \mathrm{Ker}(u - \lambda_i.\,\mathrm{id}_E)^{q_i} = F_i$.

19.5.7 Lemme. *On reprend les hypothèses et données précédentes. Si pour un certain $\ell \geq 1$ on a $\mathrm{Ker}(u - \lambda_i.\,\mathrm{id}_E)^{\ell+1} = \mathrm{Ker}(u - \lambda_i.\,\mathrm{id}_E)^{\ell}$ alors $\mathrm{Ker}(u - \lambda_i.\,\mathrm{id}_E)^{s} = \mathrm{Ker}(u - \lambda_i.\,\mathrm{id}_E)^{\ell}$ pour tout $s \geq \ell$.*

Preuve. On fait une récurrence sur $s \leq \ell + 1$ pour montrer $(P_s) : \mathrm{Ker}(u - \lambda_i.\,\mathrm{id}_E)^{t} = \mathrm{Ker}(u - \lambda_i.\,\mathrm{id}_E)^{\ell}$ pour tout $\ell \leq t \leq s$. L'initialisation pour $s = \ell+1$ est vérifiée par hypothèse. Supposons pour un certain $s \leq \ell + 1$ que (P_s) est vérifiée. Alors on a $\mathrm{Ker}(u - \lambda_i.\,\mathrm{id}_E)^{s} = \mathrm{Ker}(u - \lambda_i.\,\mathrm{id}_E)^{s-1} = \mathrm{Ker}(u - \lambda_i.\,\mathrm{id}_E)^{\ell}$. Soit $x \in \mathrm{Ker}(u - \lambda_i.\,\mathrm{id}_E)^{s+1}$, alors $(u - \lambda_i.\,\mathrm{id}_E)^{s+1}(x) = (u - \lambda_i.\,\mathrm{id}_E)^{s}(u - \lambda_i.\,\mathrm{id}_E)(x) = 0_E$, donc $(u - \lambda_i.\,\mathrm{id}_E)(x) \in \mathrm{Ker}(u - \lambda_i.\,\mathrm{id}_E)^{s}$, donc appartient par hypothèse de récurrence $\mathrm{Ker}(u - \lambda_i.\,\mathrm{id}_E)^{s-1}$, d'où l'on déduit que $(u - \lambda_i.\,\mathrm{id}_E)^{s-1}(u - \lambda_i.\,\mathrm{id}_E)(x) = (u - \lambda_i.\,\mathrm{id}_E)^{s}(x) = 0_E$, donc $x \in \mathrm{Ker}(u - \lambda_i.\,\mathrm{id}_E)^{s}$, ce qui prouve que $\mathrm{Ker}(u - \lambda_i.\,\mathrm{id}_E)^{s+1} \subset \mathrm{Ker}(u - \lambda_i.\,\mathrm{id}_E)^{s}$. L'inclusion inverse est évidente, donc $\mathrm{Ker}(u - \lambda_i.\,\mathrm{id}_E)^{s+1} = \mathrm{Ker}(u - \lambda_i.\,\mathrm{id}_E)^{s} = \mathrm{Ker}(u - \lambda_i.\,\mathrm{id}_E)^{\ell}$ ce qui montre (P_{s+1}) et achève la preuve lemme. □

19.5.8 Exemple. Soit $A = \begin{pmatrix} -4 & 1 & 0 & 1 \\ -2 & -1 & 0 & 1 \\ -12 & 6 & 3 & 1 \\ -2 & 1 & 0 & -1 \end{pmatrix}$. Soit u l'endomorphisme de $\mathbb{R}^4$ dont la matrice dans la base canonique est A. On calcule $P_u = P_A = (X - 3)(X + 2)^3$.

Par les méthodes habituelles, on détermine $E_1 = \mathrm{Ker}(u - 3.\,\mathrm{id}_E)$ (on sait qu'il est de dimension 1) et $E_2 = \mathrm{Ker}(u + 2.\,\mathrm{id}_E)$. On trouve que $\dim E_2 = 1$ ce qui, comme -2 est v.p. triple, prouve que A n'est pas diagonalisable. A priori, le polynôme minimal de A peut valoir: $(X - 3)(X + 2)^3$, ou $(X - 3)(X + 2)^2$, ou $(X - 3)(X + 2)$. Mais ce dernier cas est exclu puisque A n'est pas diagonalisable (voir 19.5.5).

Donc $(A-3I_4)(A+2I_4)$ est non-nulle. Comme par ailleurs on sait que $(A-3I_4)(A+2I_4)^3$ est nulle d'après le théorème de Cayley-Hamilton, c'est le calcul de $(A-3I_4)(A+2I_4)^2$ qui permet de trancher. On fait le calcul de ce produit matriciel:

$$\begin{pmatrix} -7 & 1 & 0 & 1 \\ -2 & -4 & 0 & 1 \\ -12 & 6 & 0 & 1 \\ -2 & 1 & 0 & -4 \end{pmatrix} \begin{pmatrix} -2 & 1 & 0 & 1 \\ -2 & 1 & 0 & 1 \\ -12 & 6 & 5 & 1 \\ -2 & 1 & 0 & 1 \end{pmatrix}^2 = \begin{pmatrix} -7 & 1 & 0 & 1 \\ -2 & -4 & 0 & 1 \\ -12 & 6 & 0 & 1 \\ -2 & 1 & 0 & -4 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -50 & 25 & 25 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = O_4.$$

On trouve $(A-3I_4)(A+2I_4)^2 = O_4$; on conclut que le polynôme minimal est $Q_A = (X-3)(X+2)^2$.

19.5.9 EXEMPLE. Soit $A = \begin{pmatrix} 1 & -1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & -1 & 2 & 0 & 1 \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 2 & -3 \end{pmatrix}$. Soit u l'endomorphisme de $\mathbb{R}^5$ dont la matrice dans la base canonique est A. On calcule $P_u = P_A = -(X-1)^3(X+1)^2$.

On détermine les sous-espaces propres; on obtient:

$\lambda_1 = 1$, $\quad E_1 = \operatorname{Ker}(u - \mathrm{id}_E)$, de dimension $r_1 = 2$ [une base est $(e_1, e_2 + e_3)$];

$\lambda_2 = -1$, $E_2 = \operatorname{Ker}(u + \mathrm{id}_E)$, de dimension $r_2 = 1$ [une base est $(e_1 + e_2 + e_3 - 2e_4 - 2e_5)$].

Donc A n'est pas diagonalisable. En particulier, $Q_A(x) \neq (x+1)(x-1)$.

On forme la suite des noyaux:

$E_1 = \operatorname{Ker}(u - \mathrm{id}_E) \subseteq \operatorname{Ker}(u - \mathrm{id}_E)^2 \subseteq \operatorname{Ker}(u - \mathrm{id}_E)^3 = F_1$, avec $\dim E_1 = r_1 = 2$ et $\dim F_1 = q_1 = 3$,

$E_2 = \operatorname{Ker}(u + \mathrm{id}_E) \subseteq \operatorname{Ker}(u + \mathrm{id}_E)^2 = F_2$, avec $\dim E_2 = r_2 = 1$ et $\dim F_2 = q_2 = 2$.

Le seul noyau à déterminer est $\operatorname{Ker}(u - \mathrm{id}_E)^2$ qui, au vu des dimensions, est égal à E_1 ou à F_1. Pour trancher, on peut faire la calcul direct de $(A - I_5)^2$. On peut aussi sans calcul utiliser le lemme 19.5.6 : si l'on avait $\operatorname{Ker}(u - \mathrm{id}_E) = \operatorname{Ker}(u - \mathrm{id}_E)^2$, on aurait aussi $\operatorname{Ker}(u - \mathrm{id}_E) = \operatorname{Ker}(u - \mathrm{id}_E)^3$, c'est-à-dire $E_1 = F_1$, ce qui est absurde. Donc $\operatorname{Ker}(u - \mathrm{id}_E)^2 = F_1$. En résumé:

$$\underbrace{E_1}_{r_1=2} = \operatorname{Ker}(u - \mathrm{id}_E) \subsetneq \operatorname{Ker}(u - \mathrm{id}_E)^2 = \operatorname{Ker}(u - \mathrm{id}_E)^3 = \underbrace{F_1}_{q_1=3},$$
$$\underbrace{E_2}_{r_2=1} = \operatorname{Ker}(u + \mathrm{id}_E) \subsetneq \operatorname{Ker}(u + \mathrm{id}_E)^2 = \underbrace{F_2}_{q_2=2}.$$

D'après le lemme des noyaux, $\operatorname{Ker}\big[(u - \mathrm{id}_E) \circ (u + \mathrm{id}_E)^2\big] = \operatorname{Ker}(u - \mathrm{id}_E) \oplus \operatorname{Ker}(u + \mathrm{id}_E)^2 = E_1 \oplus F_2$. Ce noyau est donc de dimension $r_1 + q_2 = 4 < 5$, de sorte que l'endomorphisme $(u - \mathrm{id}_E) \circ (u + \mathrm{id}_E)^2$ n'est pas nul, ou encore $(A - I_5)(A + I_5)^2 \neq O_5$.

De même, $\operatorname{Ker}\big[(u-\mathrm{id}_E)^2 \circ (u+\mathrm{id}_E)\big] = \operatorname{Ker}(u-\mathrm{id}_E)^2 \oplus \operatorname{Ker}(u+\mathrm{id}_E) = F_1 \oplus E_2$ est de dimension $q_1 + r_2 = 4 < 5$, de sorte que l'endomorphisme $(u-\mathrm{id}_E)^2 \circ (u+\mathrm{id}_E)$ n'est pas nul, ou encore $(A-I_5)^2(A+I_5) \neq O_5$.

En revanche, $\operatorname{Ker}\big[(u - \mathrm{id}_E)^2 \circ (u + \mathrm{id}_E)^2\big] = \operatorname{Ker}(u - \mathrm{id}_E)^2 \oplus \operatorname{Ker}(u + \mathrm{id}_E)^2 = F_1 \oplus F_2 = \mathbb{R}^5$, de sorte que l'endomorphisme $(u - \mathrm{id}_E)^2 \circ (u + \mathrm{id}_E)^2$ est nul, c'est-à-dire $(A - I_5)^2(A + I_5)^2 = O_5$.

On conclut que le polynôme minimal est $Q_A = (x-1)^2(x+1)^2$.

Leçon 20

Anneaux de polynômes en plusieurs indéterminées

Dans tout ce chapitre, le mot "anneau" désigne toujours un anneau commutatif unitaire.

20.1 Définitions et premières propriétés.

20.1.1 Anneaux de polynômes en deux indéterminées. On fixe un anneau A, et on considère l'anneau $B = A[X]$ des polynômes en une indéterminée X à coefficients dans A. On peut ensuite former l'anneau $B[Y]$ des polynômes en une indéterminée Y à coefficients dans B. L'anneau obtenu est appelé *anneau des polynômes en deux indéterminées à coefficients dans* A, et est noté $A[X, Y]$.

Par exemple: considérons dans $A[X, Y] = A[Y][X]$ le polynôme
$P = (X^3 + 4X^2 - 1)Y^3 + (2X - 5)Y^2 + (X^4 + X - 2)Y + (4X^3 - X + 3)$.
On a : $P = X^3Y^3 + 4X^2Y^3 - Y^3 + 2XY^2 - 5Y^2 + X^4Y + XY - 2Y + 4X^3 - X + 3$
$= YX^4 + (Y^3 + 4)X^3 + 4Y^3X^2 + (2Y^2 + Y - 1)X + (-Y^3 - 5Y^2 - 2Y + 3)$,
ce qui permet de réécrire P comme un élément de $A[Y][X]$.

Il est clair que ceci est général et permet d'écrire tout polynôme P dans $A[X, Y]$
- comme une somme finie de polynômes en Y à coefficients des polynômes en X,
- comme une somme finie de polynômes de la forme X^iY^j (dits *monômes*) à coefficients dans A,
- comme une somme finie de polynômes en X à coefficients des polynômes en Y.

On retiendra que $A[X, Y] = A[X][Y] \simeq A[Y][X]$, et tout élément de $A[X, Y]$ est une somme finie:

$$P = \sum_{(i,j)\in\mathbb{N}^2} a_{i,j}X^iY^j, \quad \text{avec } a_{i,j} \in A.$$

20.1.2 Anneaux de polynômes en trois indéterminées. De la même façon, on peut ensuite considérer l'anneau $B = A[X]$ des polynômes en une indéterminée X à coefficients dans l'anneau A, puis l'anneau $B[Y, Z]$ des polynômes en deux indéterminées Y et Z à coefficients dans B. L'anneau obtenu est appelé *anneau des polynômes en trois indéterminées à coefficients dans* A, et est noté $A[X, Y, Z]$. On a $A[X, Y, Z] = A[X][Y, Z] \simeq A[X][Y][Z] \simeq A[X, Y][Z]$, et tout élément de $A[X, Y, Z]$ est une somme finie:

$$P = \sum_{(i,j,k)\in\mathbb{N}^3} a_{i,j,k}X^iY^jZ^k, \quad \text{avec } a_{i,j,k} \in A.$$

20.1.3 Anneaux de polynômes en n indéterminées. On définit de même par récurrence l'anneau $A[X_1, X_2, \ldots, X_n]$ des *polynômes en n indéterminées à coefficients dans* A, comme l'anneau:

$$A[X_1, X_2, \ldots, X_n] = A[X_1, X_2, \ldots, X_{n-1}][X_n] \simeq A[X_1][X_2]\ldots[X_n] \simeq A[X_1][X_2, \ldots, X_n].$$

Tout élément de $A[X_1, X_2, \ldots, X_n]$ est une somme finie:

$$P = \sum_{(i_1,i_2,\ldots,i_n)\in\mathbb{N}^n} a_{i_1,i_2,\ldots,i_n}X_1^{i_1}X_2^{i_2}\ldots X_n^{i_n}, \quad \text{avec } \ a_{i_1,i_2,\ldots,i_n} \in A.$$

• *Définitions.* Un polynôme de la forme $aX_1^{i_1}X_2^{i_2}\ldots X_n^{i_n}$, avec $a \in A$, est appelé un *monôme*. Si $a \neq 0$, l'entier $i_1 + i_2 + \cdots + i_n$ est appelé le *degré total* de ce monôme. Tout polynôme est une somme finie de monômes, et on appelle *degré total d'un polynôme non-nul* le maximum des degrés totaux des monômes dont il est la somme. Par convention, le degré total du polynôme nul est strictement inférieur au degré total de tout polynôme non-nul; on le note $-\infty$.

• *Définitions.* Un polynôme non-nul de $A[X_1, X_2, \dots, X_n]$ est dit *homogène* de degré d (où d est un entier naturel) s'il est une somme finie de monômes qui sont tous de même degré total d. Pour tout polynôme P non-nul de $A[X_1, X_2, \dots, X_n]$ et tout entier naturel d, on appelle *composante homogène* de degré d de P la somme des monômes de P de degré total d.

(i) – Si $n = 1$, on note $X = X_1$ et on retrouve l'anneau $A[X]$ bien connu, étudié au chapitre précédent. Le degré total est le degré usuel.

(ii) – Si $n = 2$, on note souvent $X = X_1$ et $Y = X_2$ comme on l'a fait en 20.1.1.

EXEMPLE. Considérons par exemple dans $\mathbb{R}[X, Y]$ les polynômes:

$$P = 3X^3Y + 5X^3 - 2XY + 7 \quad \text{et} \quad Q = XY + 5X - 6Y + 1.$$

Le degré total de P est 4, celui de Q est 2. Dans Q, la composante homogène de degré 2 est XY, la composante homogène de degré 1 est $5X - 6Y$, la composante homogène de degré 0 est 1. On calcule:

$$PQ = \underbrace{3X^4Y^2}_{6} + \underbrace{20X^4Y - 18X^3Y^2}_{5} + \underbrace{25X^4 - 27X^3Y - 2X^2Y^2}_{4} + \underbrace{5X^3 - 10X^2Y + 12XY^2}_{3} + \underbrace{5XY}_{2} + \underbrace{35X - 42Y}_{1} + \underbrace{7}_{0}.$$

(iii) – Si $n = 3$, on note souvent $X = X_1$, $Y = X_2$ et $Z = X_3$ comme on l'a fait en 20.1.2.

EXEMPLE. Considérons par exemple dans $\mathbb{Z}[X, Y, Z]$ les polynômes:

$$P = X^3 + XYZ + X^2Z \quad \text{et} \quad Q = X + Y - Z.$$

P est homogène de degré 3 et Q est homogène de degré 1. Leur produit est homogène de degré 4 : $PQ = X^4 + X^3Y + 2X^2YZ - X^2Z^2 + XY^2Z - XYZ^2$.

On donnera en appendice à ce chapitre une définition plus formelle de l'anneau $A[X_1, \dots, X_n]$. Auparavant, on conclut ce paragraphe en donnant une première propriété (simple mais fondamentale) de ce type d'anneau.

20.1.4 PROPOSITION. *Si l'anneau est intègre, alors l'anneau $A[X_1, X_2, \dots, X_n]$ est intègre.*

Preuve. C'est évident par récurrence en utilisant le point (ii) de la proposition 12.3.4. □

Sur le plan arithmétique, il est important de noter que, d'après 15.3.5 :

dès lors que $n \geq 2$, l'anneau $A[X_1, \dots, X_n]$ n'est jamais principal (et donc jamais euclidien.)

En effet, même si K est un corps, $K[X, Y] \simeq K[X][Y]$ n'est pas principal puisque $K[X]$ n'est pas un corps.

En fait, après l'intégrité (proposition ci-dessus) la seule propriété arithmétique fondamentale qui se transmet d'un anneau A à un anneau $A[X_1, \dots, X_n]$ est la factorialité, évoquée dans le commentaire final de la leçon 18, et qui ne figure pas dans le programme de L3.

20.2 Polynômes symétriques.

On fixe un anneau commutatif unitaire intègre A, et un entier $n \geq 2$.

20.2.1 ACTION CANONIQUE DU GROUPE SYMÉTRIQUE SUR L'ANNEAU DES POLYNÔMES. Pour tout polynôme $P \in A[X_1, X_2, \dots, X_n]$ et toute permutation $\sigma \in S_n$, on note P_σ le polynôme de $A[X_1, X_2, \dots, X_n]$ obtenu en permutant les indéterminées $X_1, X_2, \dots, X_n$ suivant σ, c'est-à-dire:

$$P_\sigma(X_1, X_2, \dots, X_n) = P(X_{\sigma(1)}, X_{\sigma(2)}, \dots, X_{\sigma(n)}).$$

Exemple. Prenons $n = 3$ et $P(X, Y, Z) = X^2 + YZ - 3XY$. Si $\sigma = [1, 3, 2]$, alors $P_\sigma(X, Y, Z) = P(Z, X, Y) = Z^2 + XY - 3ZX$. Si $\tau = [1, 2]$, alors $P_\tau(X, Y, Z) = Y^2 + XZ - 3XY$.

Remarque. Si P est constant (c'est-à-dire de degré nul), alors $P = P_\sigma$ pour toute $\sigma \in S_n$.

Proposition.

(i) *Le groupe S_n opère sur $A[X_1, X_2, \ldots, X_n]$ par l'action: $(\sigma, P) \mapsto P_\sigma$.*

(ii) *Pour tout $\sigma \in S_n$, l'application $P \mapsto P_\sigma$ définit un automorphisme de l'anneau $A[X_1, X_2, \ldots, X_n]$.*

Preuve. Simple vérification, sans aucune difficulté. □

20.2.2 Notion de polynôme symétrique.

Définition. Un polynôme $P \in A[X_1, X_2, \ldots, X_n]$ est dit symétrique si $P_\sigma = P$ pour toute $\sigma \in S_n$.

Remarque. L'ensemble des polynômes symétriques n'est autre que l'ensemble des points fixes pour l'action du groupe S_n sur l'ensemble $A[X_1, X_2, \ldots, X_n]$.

Proposition. L'ensemble des polynômes symétriques est un sous-anneau de $A[X_1, X_2, \ldots, X_n]$.

Preuve. Simple vérification, sans aucune difficulté, utilisant le point (ii) de la proposition précédente.□

- Exemples avec $n = 2$. Les polynômes suivants sont des polynômes symétriques dans $A[X, Y]$:

(1) $S_1 = X + Y, \quad S_2 = X^2 + Y^2, \quad S_3 = X^3 + Y^3, \quad \ldots$

(2) $W_1 = X + Y, \quad W_2 = X^2 + XY + Y^2, \quad W_3 = X^3 + X^2Y + XY^2 + Y^3, \quad \ldots$

(3) $D = (X - Y)^2$.

(4) $\Sigma_1 = X + Y, \quad \Sigma_2 = XY$.

- Exemples avec $n = 3$. Les polynômes suivants sont des polynômes symétriques dans $A[X, Y, Z]$:

(1) $S_1 = X + Y + Z, \quad S_2 = X^2 + Y^2 + Z^2, \quad S_3 = X^3 + Y^3 + Z^3, \quad \ldots$

(2) $W_1 = X + Y + Z, \quad W_2 = X^2 + XY + XZ + Y^2 + YZ + Z^2,$
$W_3 = X^3 + Y^3 + Z^3 + X^2Y + XY^2 + X^2Z + XZ^2 + Y^2Z + YZ^2 + XYZ, \quad \ldots$

(3) $D = (X - Y)^2(X - Z)^2(Y - Z)^2$.

(4) $\Sigma_1 = X + Y + Z, \quad \Sigma_2 = XY + XZ + YZ, \quad \Sigma_3 = XYZ$.

Ces exemples sont des cas particuliers des exemples classiques suivants.

20.2.3 Exemples. Pour tout $n \geq 2$, les polynômes suivants sont symétriques dans $A[X_1, X_2, \ldots, X_n]$:

(1) les sommes de Newton: $\boxed{S_k = X_1^k + X_2^k + \cdots + X_n^k}$ pour tout $k \in \mathbb{N}$;

(2) les polynômes de Wronski: $\boxed{W_k = \sum_{i_1+i_2+\cdots+i_n=k} X_1^{i_1} X_2^{i_2} \ldots X_n^{i_n}}$ pour tout $k \in \mathbb{N}$;

(3) le discriminant des indéterminées: $\boxed{D = \prod_{1 \leq i < j \leq n} (X_i - X_j)^2}$;

(4) les polynômes symétriques élémentaires:

$\Sigma_1 = X_1 + X_2 + \cdots + X_n,$

$\Sigma_2 = X_1X_2 + X_1X_3 + \cdots + X_1X_n + X_2X_3 + \cdots + X_2X_n + \cdots + X_{n-1}X_n,$

$\ldots$

$\boxed{\Sigma_k = \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} X_{i_1} X_{i_2} \ldots X_{i_k}}$ pour tout $1 \leq k \leq n$, (somme de C_n^k termes),

$\ldots$

$\Sigma_n = X_1X_2 \ldots X_n.$

Remarque. Dans $A[X_1, X_2, \ldots, X_n][Z]$, le polynôme $P(Z) = (Z - X_1)(Z - X_2)\ldots(Z - X_n)$ vérifie:
$$P(Z) = Z^n - \Sigma_1 Z^{n-1} + \Sigma_2 Z^{n-2} - \cdots + (-1)^{n-1}\Sigma_{n-1}Z + (-1)^n \Sigma_n.$$

20.3 Engendrement par les polynômes symétriques élémentaires.

On reprend toutes les notations et hypothèses de 20.2. En particulier, on note $\Sigma_1, \Sigma_2, \ldots, \Sigma_n$ les polynômes symétriques élémentaires.

• *Premier exemple introductif.* Considérons dans $\mathbb{Z}[X, Y, Z]$ le polynôme symétrique:
$$P(X, Y, Z) = X^2Y + XY^2 + Y^2Z + YZ^2 + Z^2X + ZX^2.$$

On calcule: $\Sigma_1\Sigma_2 = (X + Y + Z)(XY + YZ + ZX)$
$= X^2Y + XYZ + X^2Z + XY^2 + Y^2Z + XYZ + XYZ + YZ^2 + Z^2X$
$= P(X, Y, Z) + 3XYZ = P(X, Y, Z) + 3\Sigma_3.$

On conclut que: $P(X, Y, Z) = \Sigma_1\Sigma_2 - 3\Sigma_3$, ou encore:
$$P(X, Y, Z) = F(\Sigma_1, \Sigma_2, \Sigma_3), \text{ avec } F = XY - 3Z \in \mathbb{Z}[X, Y, Z].$$

• *Second exemple introductif.* Considérons dans $\mathbb{Z}[X, Y, Z]$ le polynôme symétrique:
$$P(X, Y, Z) = (2X - Y - Z)(2Y - Z - X)(2Z - X - Y).$$

On a: $P(X, Y, Z) = (3X - \Sigma_1)(3Y - \Sigma_1)(3Z - \Sigma_1)$
$= (9XY - 3X\Sigma_1 - 3Y\Sigma_1 + \Sigma_1^2)(3Z - \Sigma_1).$
$= 27XYZ - 9XY\Sigma_1 - 9XZ\Sigma_1 + 3X\Sigma_1^2 - 9YZ\Sigma_1 + 3Y\Sigma_1^2 + 3Z\Sigma_1^2 - \Sigma_1^3$
$= 27XYZ - 9(XY + XZ + YZ)\Sigma_1 + 3(X + Y + Z)\Sigma_1^2 - \Sigma_1^3.$

On conclut que: $P(X, Y, Z) = 27\Sigma_3 - 9\Sigma_2\Sigma_1 + 2\Sigma_1^3$, ou encore:
$$P(X, Y, Z) = F(\Sigma_1, \Sigma_2, \Sigma_3), \text{ avec } F = 27Z - 9XY + 2X^3 \in \mathbb{Z}[X, Y, Z].$$

THÉORÈME FONDAMENTAL. *Soit $n \geq 2$ un entier. Soit A un anneau intègre. Pour tout polynôme symétrique $P \in A[X_1, X_2, \ldots, X_n]$, il existe un unique polynôme $F \in A[X_1, X_2, \ldots, X_n]$ tel que:*
$$P(X_1, X_2, \ldots, X_n) = F(\Sigma_1, \Sigma_2, \ldots, \Sigma_n),$$
où $\Sigma_1, \Sigma_2, \ldots, \Sigma_n$ sont les polynômes symétriques élémentaires en les X_i, $1 \leq i \leq n$.

La preuve de ce théorème est relativement longue et technique, et on ne la donnera pas ici, préférant développer quelques applications des polynômes symétriques à certaines questions concrètes sur les solutions des équations algébriques.

20.4 Formules de Newton.

On reprend toutes les notations et hypothèses de 20.2. En particulier, on note $\Sigma_1, \Sigma_2, \ldots, \Sigma_n$ les polynômes symétriques élémentaires, et $S_1, S_2, \ldots$ les sommes de Newton.

THÉORÈME. *Soit $n \geq 2$ un entier. Soit A un anneau intègre. On a dans l'anneau $A[X_1, X_2, \ldots, X_n]$ les relations suivantes:*

(i) $S_k - \Sigma_1 S_{k-1} + \Sigma_2 S_{k-2} - \cdots + (-1)^{k-1}\Sigma_{k-1}S_1 + (-1)^k k\Sigma_k = 0$, *pour tout* $1 \leq k \leq n$,

(ii) $S_\ell - \Sigma_1 S_{\ell-1} + \Sigma_2 S_{\ell-2} + \cdots + (-1)^n \Sigma_n S_{\ell-n} = 0$, *pour tout* $\ell > n$.

Preuve. Considérons dans $A[X_1, X_2, \ldots, X_n][Z]$ le polynôme $P(Z) = (Z - X_1)(Z - X_2)\ldots(Z - X_n)$. Comme on l'a vu à la fin de 20.2, on a:
$$P(Z) = Z^n - \Sigma_1 Z^{n-1} + \Sigma_2 Z^{n-2} - \cdots + (-1)^{n-1}\Sigma_{n-1}Z + (-1)^n\Sigma_n.$$
• Par définition de P, on a $P(X_i) = 0$ pour tout $1 \leq i \leq n$, et donc:
$$X_i^n - \Sigma_1 X_i^{n-1} + \Sigma_2 X_i^{n-2} - \cdots + (-1)^{n-1}\Sigma_{n-1}X_i + (-1)^n\Sigma_n = 0.$$
On fait la somme membre à membre de ces n égalités pour $1 \leq i \leq n$; il vient:

$$S_n - \Sigma_1 S_{n-1} + \Sigma_2 S_{n-2} - \cdots + (-1)^{n-1}\Sigma_{n-1}S_1 + (-1)^n n\Sigma_n = 0,$$

ce qui est l'assertion (i) pour $k = n$.

• Pour $\ell > n$, on considère dans $A[X_1, X_2, \ldots, X_n][Z]$ le polynôme $Z^{\ell-n}P(Z)$. Pour tout $1 \leq i \leq n$, il vérifie $X_i^{\ell-n}P(X_i) = 0$, donc:

$$X_i^{\ell-n}\big(X_i^n - \Sigma_1 X_i^{n-1} + \Sigma_2 X_i^{n-2} - \cdots + (-1)^{n-1}\Sigma_{n-1}X_i + (-1)^n\Sigma_n\big) = 0,$$

ou encore:

$$X_i^\ell - \Sigma_1 X_i^{\ell-1} + \Sigma_2 X_i^{\ell-2} - \cdots + (-1)^{n-1}\Sigma_{n-1}X_i^{\ell-n+1} + (-1)^n\Sigma_n X_i^{\ell-n} = 0.$$

On fait la somme membre à membre de ces n égalités pour $1 \leq i \leq n$; on obtient l'assertion (ii).

• Pour $k = 1$, la formule (i) est triviale, puisque $S_1 = \Sigma_1$.

• Il reste à prouver (i) pour $1 < k < n$. On raisonne pour cela par récurrence sur le nombre n d'indéterminées. C'est clair pour $n = 3$, car alors $k = 2$ et l'on a bien: $S_2 - \Sigma_1 S_1 + 2\Sigma_2 = 0$. On suppose maintenant la relation (i) vraie dans $A[X_1, X_2, \ldots, X_{n-1}]$, et on fixe $1 < k < n$.

On considère dans $A[X_1, X_2, \ldots, X_n]$ le polynôme $S_k - \Sigma_1 S_{k-1} + \Sigma_2 S_{k-2} - \cdots + (-1)^{k-1}\Sigma_{k-1}S_1 + (-1)^k k\Sigma_k$. Notons-le $Q(X_1, X_2, \ldots, X_{n-1}, X_n)$. Il est clairement homogène de degré k.

Introduisons enfin dans $A[X_1, X_2, \ldots, X_{n-1}]$ le polynôme $Q_0(X_1, \ldots, X_{n-1}) = Q(X_1, \ldots, X_{n-1}, 0)$.

Il est clair que, pour tout $1 \leq i \leq n-1$, on a: $\Sigma_i(X_1, \ldots, X_{n-1}, 0) = \Sigma_i(X_1, \ldots, X_{n-1})$, le i-ième polynôme symétrique élémentaire dans $A[X_1, X_2, \ldots, X_{n-1}]$. Et de même $S_i(X_1, \ldots, X_{n-1}, 0) = S_i(X_1, \ldots, X_{n-1})$. L'hypothèse de récurrence se traduit donc par: $Q_0(X_1, \ldots, X_{n-1}) = 0$ dans l'anneau $A[X_1, X_2, \ldots, X_{n-1}]$.

En d'autres termes, $Q(X_1, \ldots, X_{n-1}, 0) = 0$ dans $A[X_1, X_2, \ldots, X_{n-1}][X_n]$. On en déduit que Q est divisible par X_n dans $A[X_1, X_2, \ldots, X_{n-1}, X_n]$. Comme Q est symétrique, cela implique que Q est aussi divisible par X_i pour tout $1 \leq i \leq n-1$. Finalement Q est divisible par le produit $X_1X_2\ldots X_n$. Comme Q est homogène de degré $k < n$, ce n'est possible que si $Q = 0$, ce qui prouve le résultat voulu, et achève la preuve. □

Application 1 (*Sommes de Newton en fonction des polynômes symétriques élémentaires*). Soit $n \geq 2$ un entier. Soit A un anneau intègre. On a dans l'anneau $A[X_1, X_2, \ldots, X_n]$ les relations suivantes:

$$S_1 = \Sigma_1, \quad S_2 = S_1\Sigma_1 - 2\Sigma_2 = \Sigma_1^2 - 2\Sigma_2, \quad S_3 = S_2\Sigma_1 - S_1\Sigma_2 + 3\Sigma_3 = \Sigma_1^3 - 3\Sigma_1\Sigma_2 + 3\Sigma_3, \quad \cdots$$

et ainsi, de proche en proche, l'expression de tous les S_i comme des polynômes en les Σ_j.

Application 2 (*Polynômes symétriques élémentaires en fonction des sommes de Newton; cas d'un corps de caractéristique zéro*). Soit $n \geq 2$ un entier. Soit K un corps de caractéristique zéro. On a dans l'anneau $K[X_1, X_2, \ldots, X_n]$ les relations suivantes:

$$\Sigma_1 = S_1, \qquad \Sigma_2 = \tfrac{1}{2}S_1^2 - \tfrac{1}{2}S_2, \qquad \Sigma_3 = \tfrac{1}{6}S_1^3 - \tfrac{1}{2}S_1S_2 + \tfrac{1}{3}S_3, \quad \cdots$$

et, de proche en proche, l'expression de tous les Σ_j comme des polynômes en les S_i, à coefficients dans K.

Corollaire (une autre forme du théorème fondamental; cas d'un corps de caractéristique zéro). *Soit $n \geq 2$ un entier. Soit A un anneau intègre de caractéristique nulle. Soit K son corps de fractions. Pour tout polynôme symétrique $P \in A[X_1, X_2, \ldots, X_n]$, il existe un unique polynôme $G \in K[X_1, X_2, \ldots, X_n]$ tel que:*

$$P(X_1, X_2, \ldots, X_n) = G(S_1, S_2, \ldots, S_n),$$

où $S_1, S_2, \ldots, S_n$ sont les n premières sommes de Newton en les X_i, $1 \leq i \leq n$.

Preuve. Il suffit de combiner la seconde remarque ci-dessus avec le théorème fondamental de 20.3. □

20.5 Application aux relations entre coefficients et zéros d'un polynôme de $K[X]$.

On fixe un corps K *algébriquement clos.* Rappelons que cela signifie que tout polynôme dans $K[X]$ admet un zéro dans K, et donc que tout polynôme de degré n dans $K[X]$ se décompose comme un produit de n facteurs de degré 1 dans $K[X]$. L'exemple type de corps algébriquement clos à connaître est le corps $\mathbb{C}$ des nombres complexes.

Soit $P = \sum_{i=0}^{n} a_i X^i$ un polynôme de $K[X]$, de degré $n \geq 1$. Il a alors n zéros dans K, que l'on notera $\alpha_1, \alpha_2, \ldots, \alpha_n$, et se factorise en:

$$P(X) = \sum_{i=0}^{n} a_i X^i = a_n \prod_{j=1}^{n} (X - \alpha_j), \quad \text{avec } a_n \neq 0.$$

Pour tout $1 \leq k \leq n$, notons $\Sigma_k = \Sigma_k(\alpha_1, \alpha_2, \ldots, \alpha_n)$ le k-ième polynôme symétrique élémentaire en les α_i. On a alors (voir remarque finale de 20.2):

$$\prod_{j=1}^{n} (X - \alpha_i) = X^n - \Sigma_1 X^{n-1} + \Sigma_2 X^{n-2} - \cdots + (-1)^{n-1} \Sigma_{n-1} X + (-1)^n \Sigma_n.$$

On en déduit par identification que:

$$a_{n-1} = -a_n \Sigma_1, \quad a_{n-2} = a_n \Sigma_2, \quad \ldots, \quad a_1 = (-1)^{n-1} a_n \Sigma_{n-1}, \quad a_0 = (-1)^n a_n \Sigma_n.$$

On a ainsi établi:

PROPOSITION. *Si K est un corps algébriquement clos, alors pour tout polynôme $P(X) = \sum_{i=0}^{n} a_i X^i$ de degré $n \geq 1$, les n zéros $\alpha_1, \alpha_2, \ldots, \alpha_n$ de P vérifient:*

$$\Sigma_k(\alpha_1, \alpha_2, \ldots, \alpha_n) = (-1)^k \frac{a_{n-k}}{a_n}, \quad \text{pour tout } 1 \leq k \leq n.$$

COROLLAIRE. *Soit K un corps. Soient $\alpha_1, \alpha_2, \ldots, \alpha_n$ des éléments quelconques de K. Pour tout $1 \leq i \leq n$, posons $\lambda_i = \Sigma_i(\alpha_1, \alpha_2, \ldots, \alpha_n)$. Alors $\alpha_1, \alpha_2, \ldots, \alpha_n$ sont les zéros du polynôme:*

$$X^n - \lambda_1 X^{n-1} + \lambda_2 X^{n-2} - \cdots + (-1)^n \lambda_n.$$

Exemple 1. Pour $P(X) = aX^2 + bX + c \in \mathbb{C}[X]$, avec $a \neq 0$, on retrouve le résultat bien connu:

$$\Sigma_1 = \alpha_1 + \alpha_2 = -\frac{b}{a} \quad \text{et} \quad \Sigma_2 = \alpha_1 \alpha_2 = \frac{c}{a}.$$

Exemple 2. Pour $P(X) = X^3 + pX + q \in \mathbb{C}[X]$, on retrouve le résultat bien connu:

$$\Sigma_1 = \alpha_1 + \alpha_2 + \alpha_3 = 0, \quad \Sigma_2 = \alpha_1\alpha_2 + \alpha_1\alpha_3 + \alpha_2\alpha_3 = p \quad \text{et} \quad \Sigma_3 = \alpha_1\alpha_2\alpha_3 = -q.$$

20.6. Résultant de deux polynômes

20.6.1 POSITION DU PROBLÈME. Dans toute cette partie, K est un corps algébriquement clos. On cherche à résoudre la question suivante:

> étant donnés deux polynômes P et Q de degré ≥ 1 dans $K[X]$, distincts, trouver une condition nécessaire et suffisante pour qu'ils admettent au moins un zéro commun.

Dire que P et Q admettent un zéro commun $\alpha \in K$ équivaut à dire que le polynôme $X - \alpha$ divise à la fois P et Q dans $K[X]$, ce qui équivaut à dire que leur pgcd dans $K[X]$ est de degré ≥ 1.

20.6.2 PROPOSITION. *Soit K un corps algébriquement clos. Soient P et Q deux polynômes dans $K[X]$ de degrés respectifs $m \geq 1$ et $n \geq 1$. Ils admettent un zéro commun dans K si et seulement s'il existe des polynômes R et S dans $K[X]$ tels que:*

$$\deg R \leq m - 1, \qquad \deg S \leq n - 1, \qquad RQ = SP.$$

Preuve. Supposons les trois conditions de la proposition vérifiées. Notons $P = \gamma \prod_{i=1}^{k} (X - \alpha_i)^{m_i}$ et $Q = \delta \prod_{i=1}^{h} (X - \beta_i)^{n_i}$, où $\alpha_1, \alpha_2, \ldots, \alpha_k$ sont les zéros distincts de P dans K avec $m_1 + \cdots + m_k = m$, $\beta_1, \beta_2, \ldots, \beta_h$ sont les zéros distincts de Q dans K avec $n_1 + \cdots + n_h = n$, et $\gamma, \delta \in K^*$. On a donc: $\delta(X-\beta_1)^{n_1}(X-\beta_2)^{n_2}\cdots(X-\beta_k)^{n_h} R = \gamma(X-\alpha_1)^{m_1}(X-\alpha_2)^{m_2}\cdots(X-\alpha_k)^{m_k} S$. Puisque $\deg R < m$, il existe au moins un indice $1 \leq i_0 \leq m$ tel que $X - \alpha_{i_0}$ soit égal à l'un des $X - \beta_j$. Ainsi α_{i_0} est un zéro commun à P et Q dans K.

Réciproquement, supposons que P et Q aient un zéro commun $\alpha \in K$. Si D est un pgcd de P et Q dans $K[X]$, on a donc $\deg D \geq 1$, et il existe des polynômes non-nuls R et S dans $K[X]$ tels que $P = RD$ et $Q = SD$, avec $\deg R < \deg P$ et $\deg S < \deg Q$. On a alors l'égalité $RQ = RSD = SP$. □

20.6.3 Définition. Soit K un corps algébriquement clos. Soient P et Q deux polynômes non-nuls dans $K[X]$ de degrés respectifs $m \geq 1$ et $n \geq 1$. Notons:

$$P = \sum_{i=0}^{m} a_i X^i \quad \text{et} \quad Q = \sum_{i=0}^{n} b_i X^i, \qquad a_i, b_i \in K, \ a_m \neq 0, \ b_n \neq 0.$$

On appelle résultant de P et Q le déterminant d'ordre $m + n$ suivant:

$$R(P,Q) = \begin{vmatrix}
a_m & 0 & . & 0 & 0 & b_n & 0 & . & 0 & 0 \\
a_{m-1} & a_m & . & . & . & b_{n-1} & b_n & . & . & . \\
a_{m-2} & a_{m-1} & . & . & . & b_{n-2} & b_{n-1} & . & . & . \\
. & . & . & . & . & . & . & . & . & . \\
. & . & . & a_m & 0 & . & . & . & . & . \\
a_{m-n+1} & a_{m-n+2} & . & a_{m-1} & a_m & b_1 & . & . & . & . \\
a_{m-n} & a_{m-n+1} & . & a_{m-2} & a_{m-1} & b_0 & b_1 & . & . & . \\
a_{m-n-1} & a_{m-n} & . & . & a_{m-2} & 0 & b_0 & . & . & . \\
. & . & . & . & . & 0 & 0 & . & . & . \\
. & . & . & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . & 0 & . \\
a_2 & a_3 & . & a_n & a_{n+1} & 0 & . & . & b_n & 0 \\
a_1 & a_2 & . & a_{n-1} & a_n & 0 & . & . & b_{n-1} & b_n \\
a_0 & a_1 & . & a_{n-2} & a_{n-1} & 0 & . & . & b_{n-2} & b_{n-1} \\
0 & a_0 & . & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . & . & . \\
. & . & . & a_0 & a_1 & . & . & . & b_0 & b_1 \\
0 & 0 & . & 0 & a_0 & 0 & 0 & . & 0 & b_0
\end{vmatrix}$$

$\leftarrow 1$, $\leftarrow 2$, $\leftarrow 3$, $\leftarrow n$, $\leftarrow n+1$, $\leftarrow n+2$, $\leftarrow m-1$, $\leftarrow m$, $\leftarrow m+1$, $\leftarrow m+n$

(n colonnes de gauche, m colonnes de droite)

Remarque. On a ci-dessus, pour fixer clairement l'écriture du déterminant, supposé que $m > n$. Mutatis mutandis, l'analogue pour $m \leq n$ s'en déduit de façon évidente.

20.6.4 Théorème. *Soit K un corps algébriquement clos. Deux polynômes de $K[X]$ non-nuls et non constants ont au moins un zéro commun dans K si et seulement si leur résultant est nul.*

Preuve. Soient $P = \sum_{i=0}^{m} a_i X^i$ et $Q = \sum_{i=0}^{n} b_i X^i$ dans $K[X]$, de degrés respectifs $m \geq 1$ et $n \geq 1$. D'après la proposition 20.6.2, l'existence d'un zéro commun à P et Q équivaut à l'existence de deux polynômes non-nuls $R = \sum_{i=0}^{m-1} \lambda_i X^i$, de degré $\leq m-1$, et $S = \sum_{i=0}^{n-1} \mu_i X^i$, de degré $\leq n-1$, tels que $RQ = SP$. Par identification, cette égalité équivaut aux relations:

$$\begin{cases}
a_m\mu_{n-1} & = & b_n\lambda_{m-1} \\
a_{m-1}\mu_{n-1} + a_m\mu_{n-2} & = & b_{n-1}\lambda_{m-1} + b_n\lambda_{m-2} \\
a_{m-2}\mu_{n-1} + a_{m-1}\mu_{n-2} + a_m\mu_{n-3} & = & b_{n-2}\lambda_{m-1} + b_{n-1}\lambda_{m-2} + b_n\lambda_{m-3} \\
\cdots & \cdots & \cdots \\
a_0\mu_1 + a_1\mu_0 & = & b_0\lambda_1 + b_1\lambda_0 \\
a_0\mu_0 & = & b_0\lambda_0
\end{cases}$$

En faisant "tout passer" dans le premier membre, on obtient un système linéaire homogène, de $m+n$ équations à $m+n$ inconnues, ces inconnues étant $\mu_{n-1}, \mu_{n-2}, \ldots, \mu_0, -\lambda_{m-1}, -\lambda_{m-2}, \ldots, -\lambda_0$. Il admet une solution non-nulle si et seulement si son déterminant est nul. Or ce dernier n'est autre que le résultant $R(P,Q)$, d'où le résultat. □

20.6.5 Corollaire. *Soit K un corps algébriquement clos. Deux polynômes de $K[X]$ non-nuls et non constants sont premiers entre eux dans $K[X]$ si et seulement si leur résultant est non-nul.*

Preuve. On a déjà observé dans la preuve de 20.6.2 que l'existence d'un zéro commun à P et Q équivaut au fait que leur pgcd est de degré ≥ 1, c'est-à-dire que P et Q ne sont pas premiers entre eux. D'où le résultat d'après le théorème précédent. □

Exemples en petits degrés.

- Si $P = aX + b$ et $Q = cX + d$, alors $R(P,Q) = ad - bc$.
- Si $P = aX^2 + bX + c$ et $Q = pX + q$, alors $R(P,Q) = p^2c + q^2a - pqb$.
- Si $P = aX^2 + bX + c$ et $Q = pX^2 + qX + r$, alors $R(P,Q) = (ar - cp)^2 - (aq - bp)(br - cq)$.

Exercice. Soit $a \in K$ un paramètre quelconque. Montrer qu'il existe au plus 7 valeurs de a pour lesquelles les polynômes $P = X^4 + X^3 + X + a + 1$ et $Q = aX^3 + X + a$ ont un zéro commun. (Indication: vérifier que $R(P,Q) = a^7 + 2a^4 + 3a^3 + a^2 + a + 1$.)

20.6.6 Théorème (expression du résultant en fonction des zéros). *Soit K un corps algébriquement clos. Soient P et Q deux polynômes non-nuls dans $K[X]$ de degrés respectifs $m \geq 1$ et $n \geq 1$. Notons:*

$$P = \sum_{i=0}^{m} a_i X^i = a_m \prod_{j=1}^{m} (X - \alpha_j) \quad \text{et} \quad Q = \sum_{i=0}^{n} b_i X^i = b_n \prod_{j=1}^{n} (X - \beta_j),$$

où les a_i $(0 \leq i \leq m)$ et les b_i $(0 \leq i \leq n)$ sont les coefficients dans K de P et Q respectivement, avec $a_m \neq 0$ et $b_n \neq 0$, et où les α_j $(1 \leq j \leq m)$ et les β_j $(1 \leq j \leq n)$ sont les zéros dans K de P et Q respectivement.

Alors le résultant $R(P,Q)$ est donné par:

$$R(P,Q) = a_m^n b_n^m \prod_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} (\alpha_i - \beta_j).$$

Preuve. On raisonne en plusieurs étapes.

• *Première étape.* Soient P_1 et Q_1 les polynômes unitaires dans $K[X]$ définis par $P = a_m P_1$ et $Q = b_n Q_1$. On a:

$$P_1 = \prod_{j=1}^{m} (X - \alpha_j) \qquad \text{et} \qquad Q_1 = \prod_{j=1}^{n} (X - \beta_j).$$

Notons $\Sigma_1, \ldots, \Sigma_m$ les fonctions symétriques élémentaires en les zéros $\alpha_1, \ldots, \alpha_m$ de P, et $\Sigma'_1, \ldots, \Sigma'_n$ les fonctions symétriques élémentaires en les zéros $\beta_1, \ldots, \beta_n$ de Q. D'après les résultats de 20.5, on a:

$$\Sigma_1 = -\frac{a_{m-1}}{a_m},\ \Sigma_2 = \frac{a_{m-2}}{a_m},\ \ldots,\ \Sigma_m = (-1)^m \frac{a_0}{a_m},\ \Sigma'_1 = -\frac{b_{n-1}}{b_n},\ \Sigma'_2 = \frac{b_{n-2}}{b_n},\ \ldots,\ \Sigma'_n = (-1)^n \frac{b_0}{b_n},$$

et donc:

$$P_1 = X^m - \Sigma_1 X^{m-1} + \Sigma_2 X^{m-2} - \ldots + (-1)^{m-1} \Sigma_{m-1} X + (-1)^m \Sigma_m,$$
$$Q_1 = X^n - \Sigma'_1 X^{n-1} + \Sigma'_2 X^{n-2} - \ldots + (-1)^{n-1} \Sigma'_{n-1} X + (-1)^n \Sigma'_n.$$

• *Deuxième étape.* Reprenons maintenant les expressions développées $P = \sum_{i=0}^{m} a_i X^i$ et $Q = \sum_{i=0}^{n} b_i X^i$. L'expression du déterminant $R(P,Q)$ vu en 20.6.3 permet de voir $R(P,Q)$ comme un polynôme en les $n+m+2$ indéterminées $a_0, a_1, \ldots, a_m, b_0, b_1, \ldots, b_n$. Plus précisément, la forme du déterminant permet d'observer que ce polynôme est homogène de degré n en les indéterminées $a_0, a_1, \ldots, a_m$ et homogène de degré m en les indéterminées $b_0, b_1, \ldots, b_n$. Dans l'anneau $K[a_0, a_1, \ldots, a_m, b_0, b_1, \ldots, b_n]$, notons:

$$R(P,Q) = F(a_0, a_1, \ldots, a_m, b_0, b_1, \ldots, b_n).$$

D'après les observations précédentes:

$$R(P,Q) = a_m^n b_n^m F(\frac{a_0}{a_m}, \frac{a_1}{a_m}, \ldots, \frac{a_{m-1}}{a_m}, 1, \frac{b_0}{b_n}, \frac{b_1}{b_n}, \ldots, \frac{b_{n-1}}{b_n}, 1).$$

Cette relation appliquée aux polynômes P_1 et Q_1 développés comme à la fin de la première étape s'écrit:

$$R(P_1,Q_1) = F\Big((-1)^m\Sigma_m,\ (-1)^{m-1}\Sigma_{m-1},\ \ldots,\ -\Sigma_1,\ 1,\ (-1)^n\Sigma'_n,\ (-1)^{n-1}\Sigma'_{n-1},\ \ldots,\ -\Sigma'_1,\ 1\Big).$$

On en déduit d'abord que $R(P_1,Q_1)$ est un polynôme symétrique en $\alpha_1, \alpha_2, \ldots, \alpha_m$ d'une part, et en $\beta_1, \beta_2, \ldots, \beta_n$ d'autre part (c'est le sens évident du théorème 20.3).

On en déduit d'autre part que $R(P,Q) = a_m^n b_n^m R(P_1,Q_1)$, d'où $R(P_1,Q_1) = 0$ si et seulement si $R(P,Q) = 0$, c'est-à-dire d'après le théorème 20.6.4 si et seulement s'il existe un couple (i,j) avec $1 \leq i \leq m$ et $1 \leq j \leq n$ tel que $\alpha_i = \beta_j$.

Ces deux remarques impliquent que, dans $K[\alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_n]$, le polynôme $R(P_1,Q_1)$ est divisible par $(\alpha_i - \beta_j)$ pour tous $1 \leq i \leq m$ et $1 \leq j \leq n$, et donc par le produit $\Pi = \prod_{1\leq i\leq m, 1\leq j\leq n}(\alpha_i - \beta_j)$. En comparant pour chacun des polynômes $R(P_1,Q_1)$ et Π les degrés en $\alpha_1, \alpha_2, \ldots, \alpha_m$ et en $\beta_1, \beta_2, \ldots, \beta_n$, ainsi que le coefficient de $(\alpha_1\alpha_2\cdots\alpha_m)^n$, on conclut finalement que $R(P_1,Q_1) = \Pi$.

On en tire que $R(P,Q) = a_m^n b_n^m \Pi$, ce qui achève la preuve. □

Remarque. On a donc les expressions suivantes du résultant:

$$R(P,Q) = a_m^n b_n^m \prod_{\substack{1\leq i\leq m\\ 1\leq j\leq n}} (\alpha_i - \beta_j) = a_m^n \prod_{1\leq i\leq m} Q(\alpha_i) = (-1)^{mn}\, b_n^m \prod_{1\leq j\leq n} P(\beta_j).$$

qui rendent explicite la conclusion du théorème 20.6.4

20.7 Discriminant d'un polynôme.

20.7.1 Définition. Soit K un corps algébriquement clos. On appelle discriminant d'un polynôme P de degré au moins égal à 2 dans $K[X]$ le résultant de P et de son polynôme dérivé P'. On note:

$$\Delta(P) = R(P,P').$$

20.7.2 Théorème. *Soit K un corps algébriquement clos. Soient P un polynôme de degré au moins égal à 2 dans $K[X]$ et P' son polynôme dérivé. Les conditions suivantes sont équivalentes.*

(i) *Le polynôme P a au moins un zéro multiple dans K.*

(ii) *Les polynômes P et P' ne sont pas premiers entre eux.*

(iii) *Les polynômes P et P' ont au moins un zéro commun dans K.*

(iv) *Le discriminant $\Delta(P)$ du polynôme P est nul.*

Preuve. L'équivalence de (ii) et (iii) résulte de la remarque préliminaire de 20.6.1. Par définition même du discriminant, l'équivalence de (iii) et (iv) est une conséquence du théorème 20.6.2. Il suffit donc de montrer l'équivalence de (i) et (ii).

Supposons d'abord que P a un zéro multiple α. Alors il existe un polynôme Q de degré $n-2$, où n désigne le degré de P, tel que $P(X) = (X-\alpha)^2 Q(X)$. On a alors $P'(X) = 2(X-\alpha)Q(X) + (X-\alpha)^2 Q'(X)$, de sorte que $X - \alpha$ est un diviseur commun de P et P' de degré non-nul. Donc P et P' ne sont pas premiers entre eux.

Supposons réciproquement que P et P' ne sont pas premiers entre eux. Leur pgcd D est de degré strictement positif. Comme K est algébriquement clos, il admet au moins un zéro $\alpha \in K$. Le polynôme $X - \alpha$ divise D, donc divise P et P'. Il existe en particulier un polynôme Q de degré $n-1$, où n désigne le degré de P, tel que $P(X) = (X-\alpha)Q(X)$. On a alors $P'(X) = Q(X) + (X-\alpha)Q'(X)$. Mais comme $X - \alpha$ divise aussi P', on en déduit qu'il divise Q. Et donc P est divisible par $(X-\alpha)^2$, ce qui achève la preuve. □

20.7.3 Corollaire. *Soit K un corps algébriquement clos. Un polynôme P de degré au moins égal à 2 dans $K[X]$ n'admet que des zéros simples dans K si et seulement si son discriminant est non-nul.*

Exemple 1. Dans $\mathbb{C}[X]$, considérons $P(X) = aX^2 + bX + c$, avec $a \neq 0$. On a $P'(X) = 2aX + b$, et donc:

$$\Delta(P) = R(P, P') = \begin{vmatrix} a & 2a & 0 \\ b & b & 2a \\ c & 0 & b \end{vmatrix} = a(4ac - b^2).$$

L'application du corollaire ci-dessus montre que P n'a que des zéros simples si et seulement si $b^2 - 4ac \neq 0$, résultat bien connu !

Exemple 2. Dans $\mathbb{C}[X]$, considérons $P(X) = X^3 + pX + q$. On a $P'(X) = 3X^2 + p$, et donc:

$$\Delta(P) = R(P, P') = \begin{vmatrix} 1 & 0 & 3 & 0 & 0 \\ 0 & 1 & 0 & 3 & 0 \\ p & 0 & p & 0 & 3 \\ q & p & 0 & p & 0 \\ 0 & q & 0 & 0 & p \end{vmatrix} = 4p^3 + 27q^2.$$

Supposons que $\Delta(P) = 0$, et notons α le zéro multiple de P dans K (il est forcément unique puisque P est de degré 3). Comme α est alors aussi zéro de P', on a $\alpha^2 = -\frac{p}{3}$.

Si $p = 0$, alors la nullité de $\Delta(P) = 4p^3 + 27q^2$ implique que l'on a aussi $q = 0$, donc $P(X) = X^3$, qui admet 0 comme zéro triple.

Si $p \neq 0$, alors la nullité de $\Delta(P) = 4p^3 + 27q^2$ implique que l'on a $p = -\frac{27q^2}{4p^2}$, d'où $\alpha^2 = -\frac{p}{3} = \frac{9q^2}{4p^2}$. On vérifie que seul $\alpha = -\frac{3q}{2p}$ est zéro de P, et c'est un zéro double.

20.7.4 Proposition (expression du discriminant en fonction des zéros). *Soit K un corps algébriquement clos. Soit $P = a_n X^n + \cdots + a_1 X + a_0$ un polynôme de degré $n \geq 2$ dans $K[X]$. Soient $\alpha_1, \ldots, \alpha_n$ les zéros de P dans K. Alors le discriminant de P est donné par:*

$$\Delta(P) = (-1)^{\frac{n(n-1)}{2}} a_n^{2n-1} \prod_{1 \leq i < j \leq n} (\alpha_i - \alpha_j)^2.$$

Preuve. On a: $P(X) = a_n \prod_{1 \leq k \leq n} (X - \alpha_k)$, donc: $P'(X) = a_n \sum_{1 \leq j \leq n} \Big(\prod_{1 \leq k \leq n, k \neq j} (X - \alpha_k) \Big)$

d'où, pour tout $1 \leq i \leq n$, l'égalité: $P'(\alpha_i) = a_n(\alpha_i - \alpha_1)(\alpha_i - \alpha_2) \ldots (\alpha_i - \alpha_{i-1})(\alpha_i - \alpha_{i+1}) \ldots (\alpha_i - \alpha_n)$.

On calcule alors en utilisant les expressions de la remarque suivant le théorème 20.6.6 :

$$\begin{aligned}
\Delta(P) = R(P, P') &= a_n^{n-1} \prod_{1 \leq i \leq n} P'(\alpha_i) \\
&= a_n^{n-1} \prod_{1 \leq i \leq n} a_n(\alpha_i - \alpha_1)(\alpha_i - \alpha_2) \ldots (\alpha_i - \alpha_{i-1})(\alpha_i - \alpha_{i+1}) \ldots (\alpha_i - \alpha_n) \\
&= a_n^{2n-1} \prod_{1 \leq i \leq n} (\alpha_i - \alpha_1)(\alpha_i - \alpha_2) \ldots (\alpha_i - \alpha_{i-1})(\alpha_i - \alpha_{i+1}) \ldots (\alpha_i - \alpha_n) \\
&= a_n^{2n-1} \prod_{1 \leq i \leq n} (-1)^{i-1} (\alpha_1 - \alpha_i)(\alpha_2 - \alpha_i) \ldots (\alpha_{i-1} - \alpha_i)(\alpha_i - \alpha_{i+1}) \ldots (\alpha_i - \alpha_n) \\
&= a_n^{2n-1} (-1)^{n(n-1)/2} \prod_{1 \leq i < j \leq n} (\alpha_i - \alpha_j)^2.
\end{aligned}$$

□

Remarque. Cette relation justifie le nom de discriminant des indéterminées donné à l'exemple (3) du paragraphe 20.2.3

• On fixe un anneau A (commutatif, unitaire), et un entier naturel $n \geq 1$. Considérons une application f:

$$\begin{array}{rcl} \mathbb{N}^n & \to & A \\ (i_1, i_2, \ldots, i_n) & \mapsto & a_{i_1,i_2,\ldots,i_n} \end{array}$$

que l'on notera sous forme de suite (indexée sur $\mathbb{N}^n$), c'est-à-dire $(a_{i_1,i_2,\ldots,i_n})_{(i_1,i_2,\ldots,i_n)\in\mathbb{N}^n}$. On dit que f est à support fini lorsque $a_{i_1,i_2,\ldots,i_n} = 0$ sauf pour un nombre fini d'éléments $(i_1, i_2, \ldots, i_n)$ de $\mathbb{N}^n$. On note $\mathcal{R}_n(A)$ l'ensemble de toutes les suites f de ce type qui sont à support fini.

• Il est technique mais élémentaire de vérifier que $\mathcal{R}_n(A)$ est un anneau commutatif pour la somme et le produit définis par

$$(a_{i_1,i_2,\ldots,i_n})_{(i_1,i_2,\ldots,i_n)\in\mathbb{N}^n} + (b_{i_1,i_2,\ldots,i_n})_{(i_1,i_2,\ldots,i_n)\in\mathbb{N}^n} = (a_{i_1,i_2,\ldots,i_n} + b_{i_1,i_2,\ldots,i_n})_{(i_1,i_2,\ldots,i_n)\in\mathbb{N}^n},$$

$$(a_{i_1,i_2,\ldots,i_n})_{(i_1,i_2,\ldots,i_n)\in\mathbb{N}^n} \times (b_{i_1,i_2,\ldots,i_n})_{(i_1,i_2,\ldots,i_n)\in\mathbb{N}^n} = (c_{i_1,i_2,\ldots,i_n})_{(i_1,i_2,\ldots,i_n)\in\mathbb{N}^n},$$

avec

$$c_{i_1,i_2,\ldots,i_n} = \sum_{r=1}^{n} \sum_{j_r+k_r=i_r} a_{j_1,j_2,\ldots,j_n} b_{k_1,k_2,\ldots,k_n}.$$

Pour tout $a \in A$, on note encore a la suite $(a_{i_1,i_2,\ldots,i_n})_{(i_1,i_2,\ldots,i_n)\in\mathbb{N}^n}$ dont tous les termes $a_{i_1,i_2,\ldots,i_n}$ sont nuls, sauf $a_{0,0,\ldots,0} = a$. La définition du produit dans $\mathcal{R}_n(A)$ permet de vérifier que:

$$a \times (b_{i_1,i_2,\ldots,i_n})_{(i_1,i_2,\ldots,i_n)\in\mathbb{N}^n} = (ab_{i_1,i_2,\ldots,i_n})_{(i_1,i_2,\ldots,i_n)\in\mathbb{N}^n},$$

de sorte que A peut être identifié à un sous-anneau de $\mathcal{R}_n(A)$. En particulier, le neutre additif et le neutre multiplicatif de l'anneau $\mathcal{R}_n(A)$ sont (via l'identification ci-dessus) le zéro 0_A et le un 1_A de l'anneau A.

• Pour tout $(j_1, j_2, \ldots, j_n) \in \mathbb{N}^n$, on note $X_1^{j_1} X_2^{j_2} \ldots X_n^{j_n}$ la suite $(a_{i_1,i_2,\ldots,i_n})_{(i_1,i_2,\ldots,i_n)\in\mathbb{N}^n}$ dont tous les termes sont nuls, sauf $a_{j_1,j_2,\ldots,j_n} = 1$. La définition du produit dans $\mathcal{R}_n(A)$ permet de vérifier que:

$$(X_1^{j_1} X_2^{j_2} \ldots X_n^{j_n})(X_1^{k_1} X_2^{k_2} \ldots X_n^{k_n}) = X_1^{j_1+k_1} X_2^{j_2+k_2} \ldots X_n^{j_n+k_n}.$$

En particulier, $X_1^0 X_2^0 \ldots X_n^0 = 1$ et tout élément de l'anneau $\mathcal{R}_n(A)$ s'écrit comme une somme finie:

$$(a_{i_1,i_2,\ldots,i_n})_{(i_1,i_2,\ldots,i_n)\in\mathbb{N}^n} = \sum_{(i_1,i_2,\ldots,i_n)\in\mathbb{N}^n} a_{i_1,i_2,\ldots,i_n} X_1^{i_1} X_2^{i_2} \ldots X_n^{i_n}.$$

Définition. L'anneau $\mathcal{R}_n(A)$ est appelé l'anneau des polynômes en n indéterminées à coefficients dans A. On le note $A[X_1, X_2, \ldots, X_n]$.

Lemme (propriété universelle des anneaux de polynômes). *Soient A et B deux anneaux et φ un morphisme d'anneaux $A \to B$. Alors, quels que soient un entier $n \geq 1$ et des éléments $b_1, b_2, \ldots, b_n$ de B, il existe un unique morphisme d'anneaux $\Phi : A[X_1, X_2, \ldots, X_n] \to B$ qui prolonge φ (c'est-à-dire $\Phi(a) = \varphi(a)$ pour tout $a \in A$), et tel que $\Phi(X_i) = b_i$ pour tout $1 \leq i \leq n$.*

Preuve. Si Φ existe, il est nécessairement défini par:

$$\Phi\Big(\sum a_{i_1,i_2,\ldots,i_n} X_1^{i_1} X_2^{i_2} \ldots X_n^{i_n}\Big) = \sum \varphi(a_{i_1,i_2,\ldots,i_n}) b_1^{i_1} b_2^{i_2} \ldots b_n^{i_n},$$

ce qui réciproquement définit bien un morphisme d'anneaux $A[X_1, X_2, \ldots, X_n] \to B$. □

Proposition (fondamentale). *Soit A un anneau. Pour tout $n \geq 2$, il existe dans $A[X_1, X_2, \ldots, X_n]$ un sous-anneau isomorphe à $A[X_1, X_2, \ldots, X_{n-1}]$, et l'on a alors:*

$$A[X_1, X_2, \ldots, X_n] \simeq A[X_1, X_2, \ldots, X_{n-1}][X_n].$$

Preuve. Dans $B = A[X_1, X_2, \ldots, X_n]$, considérons l'ensemble C des polynômes $P = \sum a_{i_1,i_2,\ldots,i_n} X_1^{i_1} X_2^{i_2} \ldots X_n^{i_n}$ tels que $a_{i_1,i_2,\ldots,i_n} = 0$ pour tout multi-indice $(i_1, i_2, \ldots, i_n) \in \mathbb{N}^n$ tel que $i_n \neq 0$. Il est clair que C est un sous-anneau de A isomorphe à $A[X_1, X_2, \ldots, X_{n-1}]$. D'après le lemme précédent (appliqué avec $n = 1$), l'injection canonique $\varphi : C \to B$ se prolonge en un unique morphisme d'anneaux $\Phi : C[X] \to B$ tel que $\Phi(P) = \varphi(P) = P$ pour tout $P \in C$ et $\Phi(X) = X_n$, qui réalise de façon évidente un isomorphisme $C[X] \simeq B$, d'où le résultat en revenant aux notations de l'énoncé. □

Cette proposition permet de justifier la façon dont on a introduit en 20.1 les anneaux de polynômes en plusieurs indéterminées.