UVSQ 2025/2026
Licence de sciences et technologie, santé
LSMA610 (groupes et géométrie)

Notes de cours

Table des matières

1 Premiers éléments sur la structure de groupe **2**
1.1 La structure . . . 2
1.2 Sous-groupes engendrés . . . 4
1.3 Classes à droite et à gauche, théorème de Lagrange . . . 7
1.4 Groupe-quotient, théorèmes d'isomorphismes . . . 8

2 Le groupe symétrique **13**
2.1 Permutations d'un ensemble fini . . . 13
2.2 Signature d'une permutation . . . 17
2.3 Le groupe alterné . . . 18

3 Groupes abéliens de type fini **20**
3.1 Prélude à l'unicité des facteurs invariants . . . 20
3.2 GALTF, rang . . . 21
3.3 GAF et GATF, rang et facteurs invariants . . . 24

4 Groupes linéaires **26**
4.1 Petit mémento sur le déterminant . . . 26
4.2 Transvections et dilatations . . . 32
4.3 Le groupe linéaire sur les corps finis . . . 36
4.4 Le groupe orthogonal euclidien . . . 38
4.5 Un tout petit peu sur le groupe modulaire . . . 42

5 Action d'un groupe sur un ensemble **44**
5.1 Généralités, premiers exemples . . . 44
5.2 L'équation aux classes . . . 47
5.3 Produits semi-directs de groupes . . . 50
5.4 Théorèmes de Sylow . . . 53

6 Polynômes symétriques **56**
6.1 Théorème des polynômes symétriques . . . 56
6.2 Polynômes antisymétriques, polynômes invariants par le groupe alterné . . . 59
6.3 Séries formelles et formules de Newton . . . 61

1 Premiers éléments sur la structure de groupe

1.1 La structure

Lecture du polycopié *Structures abstraites* : groupes, sous-groupes, homomorphismes de groupes, exemples fondamentaux.

Définition (ordre d'un groupe, ordre d'un élément dans un groupe)
Soit G un groupe. L'*ordre de* G est son cardinal — fini ou non. On note $|G|$ l'ordre de G. Si $x \in G$, l'*ordre de* x est le cardinal de $\{x^n,\ n \in \mathbb{Z}\}$.

Exemple
Si n est un entier naturel non nul, l'ordre du groupe $(\mathbb{Z}/n\mathbb{Z}, +)$ est n. Si $k \in \mathbb{Z}$, l'ordre de la classe de k modulo n est $\frac{n}{PGCD(n,k)}$.

Exercice 1
Montrer que le groupe $\mathrm{GL}(2, \mathbb{Z}/2\mathbb{Z})$ des matrices 2×2 inversibles à coefficients dans le corps $\mathbb{Z}/2\mathbb{Z}$ est d'ordre 6 en faisant la liste exhaustive de ses éléments. Montrer que ce groupe n'est pas abélien et calculer l'ordre de chacun de ses éléments — on trouve trois éléments d'ordre 2 et deux éléments d'ordre 3.

Proposition (images directe ou inverse d'un sous-groupe)
Soit $f : G \to G'$ un homomorphisme de groupes.
(i) Si H est un sous-groupe de G, alors $f(H)$ est un sous-groupe de G'.
(ii) Si H' est un sous-groupe de G', alors $f^{-1}(H')$ est un sous-groupe de G.
(iii) En particulier, le noyau et l'image d'un homomorphisme de groupes $G \to G'$ sont des sous-groupes respectifs de G et de G'.

Preuve. On note les lois multiplicativement. (i) $f(1_G) \in f(H)$, $f(x)f(y) = f(xy) \in f(H)$ et $f(x)^{-1} = f\left(x^{-1}\right) \in f(H)$, pour tous $x, y \in H$. Donc $f(H)$ est un sous-groupe de G'. (ii) $f(1_G) = 1_{G'}$ et $1_{G'} \in H'$; donc $1_G \in f^{-1}(H')$. Par ailleurs, si $x, y \in f^{-1}(H')$, alors $f(xy) = f(x)f(y) \in H'$ et $f\left(x^{-1}\right) = f(x)^{-1} \in H'$; cela montre que $f^{-1}(H')$ est un sous-groupe de G. (iii) Si Γ est un groupe, $\{1_\Gamma\}$ et Γ en sont des sous-groupes. ■

Définition (conjugaison dans un groupe)
Soit G un groupe. Si $g, g' \in G$, on dit que g et g' sont *conjugués (dans G)* lorsqu'il existe $h \in G$ tel que $g' = hgh^{-1}$. Deux sous-groupes H et H' de G sont dits *conjugués (dans G)* lorsqu'il existe $g \in G$ tel que $H' = gHg^{-1}$.

Exercice 2
(i) Si H est un sous-groupe d'un groupe G et si $g \in G$, alors le conjugué gHg^{-1} est encore un sous-groupe de G.
(ii) La conjugaison est une relation d'équivalence sur les éléments d'un groupe. C'est aussi une relation d'équivalence sur l'ensemble des sous-groupes de G. Lorsque le groupe est abélien, les classes d'équivalences pour ces deux relations sont des singletons – autrement dit, dans un groupe abélien, la conjugaison est triviale.
(iii) Calculer les classes de conjugaison des éléments de $\mathrm{GL}(2, \mathbb{Z}/2\mathbb{Z})$ — on trouve une classe à un élement, une classe à deux élements et une classe à trois élements.
(iv) Dans un groupe, deux éléments conjugués ont le même ordre et deux sous-groupes conjugués sont isomorphes — et ont donc le même ordre.

Définition (sous-groupe distingué)
Soit G un groupe. Un sous-groupe H de G est *distingué* (ou *normal*) lorsqu'il est stable par conjugaison. On note alors $H \triangleleft G$. Autrement dit, $H \triangleleft G$ si, et seulement si $\forall g \in G$, $gHg^{-1} = H$.

Exemple
Dans le groupe $\mathrm{GL}(2, \mathbb{Z}/Z\mathbb{Z})$, le sous-groupe $\left\{I_2, \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}\right\}$ est distingué alors que le sous-groupe $\left\{I_2, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right\}$ ne l'est pas.

A noter Evidemment, tous les sous-groupes d'un groupe abélien sont distingués.

Exercice 3
Soient G un groupe et H un sous-groupe de G. Les assertions suivantes sont équivalentes.
(i) $H \triangleleft G$
(ii) $\forall g \in G$, $gHg^{-1} \subseteq H$
(iii) $\forall x \in G$, $\forall h \in H$, $\exists h' \in H$, $xh = h'x$.

Proposition (image inverse d'un sous-groupe distingué)
Soient $f : G \to G'$ un homomorphisme de groupes.
(i) Si H' un sous-groupe distingué de G', alors, $f^{-1}(H')$ est un sous-groupe distingué de G.
(ii) Si H est un sous-groupe distingué de G, alors $f(H)$ est un sous-groupe distingué de $f(G)$ — mais pas de G' en général.
(iii) En particulier, le noyau d'un homomorphisme de groupes est un sous-groupe distingué du groupe de départ.

Preuve. (i) Si $g \in G$ et si $h \in f^{-1}(H')$, alors $f\left(ghg^{-1}\right) = f(g)f(h)f(g)^{-1}$ est dans H' puisque $f(h) \in H'$ et puisque $H' \triangleleft G'$. Donc $ghg^{-1} \in f^{-1}(H')$. (ii) Si $g \in G$ et $h \in H$, alors $f(g)f(h)f(g)^{-1} = f\left(ghg^{-1}\right) \in f(H)$ puisque $H \triangleleft G$. (iii) est une conséquence immédiate de (i).

A noter
Pour montrer qu'un sous-groupe est distingué, il suffit donc de le faire apparaître comme le noyau d'un homomorphisme de groupes. Cette remarque est à mettre au rang de méthode. On verra que, réciproquement, tout sous-groupe distingué est le noyau d'un homomorphisme de groupes.

Exemple
Soit $\mathbb{F}$ un corps et $\mathbb{F}^\times$ son groupe multiplicatif. Si V est un $\mathbb{F}$-espace vectoriel de dimension finie, l'ensemble des applications linéaires bijectives $V \to V$ est un sous-groupe de l'ensemble des applications bijectives $V \to V$ pour la composition des applications (exercice). On le note $\mathrm{GL}(V)$: c'est le *groupe linéaire* de V.
L'application déterminant

$$\begin{array}{rccc} \det : & \mathrm{GL}(V) & \longrightarrow & \mathbb{F}^\times \\ & f & \longmapsto & \det(f) \end{array}$$

est un homomorphisme de groupes — paraphrase du fait que le déterminant d'une composée est le produit des déterminants. Son noyau est le *groupe spécial linéaire de V* ; on le note $\mathrm{SL}(V)$. La proposition précédente assure que $\mathrm{SL}(V) \triangleleft \mathrm{GL}(V)$.

Exercice 4
Soit $\mathbb{F}$ un corps et $d \in \mathbb{N} \setminus \{0\}$. On note $\mathcal{M}_d(\mathbb{F})$ le $\mathbb{F}$-espace vectoriel des matrices carrées à d lignes, d colonnes et à coefficients dans $\mathbb{F}$ et $\mathrm{GL}(d, \mathbb{F})$ le groupe des matrices inversibles de $\mathcal{M}_d(\mathbb{F})$ pour la multiplication des matrices – exercice dans l'exercice : c'est bien un groupe. On note également $T_d(\mathbb{F})$ le sous-ensemble de $\mathcal{M}_d(\mathbb{F})$ formé des matrices triangulaires supérieures inversibles. Montrer que $T_d(\mathbb{F})$ est un sous-groupe non distingué de $\mathrm{GL}(d, \mathbb{F})$.

Exercice 5 (Groupe des automorphismes d'un groupe ; automorphisme intérieur)
Soit G un groupe.
(i) Un automorphisme de G est un homomorphisme bijectif $G \to G$. Montrer que, muni de la composition des applications, l'ensemble des automorphismes de G est un groupe que l'on note usuellement $\mathrm{Aut}(G)$.
(ii) Si $g \in G$, on définit l'application $i_g : G \to G$, $h \mapsto ghg^{-1}$; montrer que $i_g \in \mathrm{Aut}(G)$. Les automorphismes i_g sont appelés *automorphismes intérieurs* de G.
(iii) Montrer que si G est un groupe, l'application $G \to \mathrm{Aut}(G)$, $g \mapsto i_g$ est un homomorphisme de groupes.
(iv) Montrer que l'ensemble des automorphismes intérieurs de G est un sous-groupe distingué du groupe des automorphismes de G.

Définition (centre d'un groupe)
Soit G un groupe. Le centre de G est l'ensemble de ses éléments qui commutent avec tous les éléments de G. On le note généralement $Z(G)$. Ainsi,

$$Z(G) = \{g \in G,\ \forall h \in G,\ hg = gh\}.$$

A noter
Le centre d'un groupe en est toujours un sous-groupe distingué (exercice). Bien sûr, un groupe est abélien si, et seulement s'il égale son centre.

Proposition (centre du groupe linéaire)
(i) Si V est un espace vectoriel de dimension finie sur un corps $\mathbb{F}$, alors le centre de $\mathrm{GL}(V)$ est le groupe $\mathbb{F}^\times \mathrm{id}_V$ de ses homothéties.

(ii) Si $\mathbb{F}$ est un corps et si d est un entier naturel non nul, le centre de $\mathrm{GL}(d,\mathbb{F})$ est $\mathbb{F}^\times I_d$.

PREUVE. Avec les notations de (i) et (ii), si V est de dimension d, le choix d'une base $\mathcal{B}$ de V rend les groupes $\mathrm{GL}(V)$ et $\mathrm{GL}(d,\mathbb{F})$ isomorphes *via* l'isomorphisme de groupes $u \in \mathrm{GL}(V) \longmapsto \mathrm{Mat}_{\mathcal{B}}(u)$ qui envoie toute homothétie $x\,\mathrm{id}_v$ sur xI_d — les notations sont évidentes. Ainsi, il suffit de montrer (i).
Si V est de dimension 1, alors $\mathrm{GL}(V) = \mathbb{F}^\times \mathrm{id}_V$ est évidemment isomorphe au groupe abélien $\mathbb{F}^\times$.
On suppose que $\dim V \geq 2$ et on prend $c \in Z(\mathrm{GL}(V))$. Soit $v \in V \setminus \{0\}$. On suppose que v et $w = c(v)$ sont linéairement indépendants. En complétant (v,w) en une base de V, soient p et q dans $\mathrm{GL}(V)$ tels que $p(v) = w$ et $p(w) = v$ d'une part, $q(v) = w$ et $q(w) = v + w$ d'autre part — on notera que w et $v + w$ sont aussi linéairement indépendants. Alors, les égalités $pc = cp$ et $qc = cq$ appliquées à v assurent que $v = c(w)$ et $v + w = c(w)$, ce qui contredit l'indépendance de v et w. On en déduit que pour tout $v \in V$, les vecteurs $c(v)$ et v sont colinéaires.
Autrement dit, tout vecteur de v est un vecteur propre de c. Cela oblige c à être une homothétie. En effet, pour tout $v \in V$, soit $\lambda(v) \in \mathbb{F}$ tel que $c(v) = \lambda(v)v$. Alors, si $v, w \in V$ sont indépendants, $\lambda(v+w)(v+w) = \lambda(v)v + \lambda(w)w$, ce qui entraîne que $\lambda(v) = \lambda(w)$. Ainsi, l'application λ est constante sur toute base de V, ce qui prouve que c est une homothétie. ■

Proposition (produit direct de groupes)
Soient G_1 et G_2 deux groupes notés multiplicativement. Alors, la loi de composition

$$\begin{array}{ccc} (G_1 \times G_2) \times (G_1 \times G_2) & \longrightarrow & G_1 \times G_2 \\ \big((g_1,g_2),(h_1,h_2)\big) & \longmapsto & (g_1h_1, g_2h_2) \end{array}$$

confère au produit cartésien $G_1 \times G_2$ une structure de groupe dont l'élément neutre est $(1_{G_1}, 1_{G_2})$. Dans ce groupe, l'inverse d'un élément (g_1,g_2) est $\big(g_1^{-1}, g_2^{-1}\big)$.

PREUVE. Exercice. ■

Définition (produit direct de groupes)
Le groupe décrit dans la proposition précédente est le *produit direct des groupes G_1 et G_2*. Le produit direct $G \times G$ est noté G^2. De façon analogue, on définit le produit direct $G_1 \times \cdots \times G_n$ d'une famille finie de groupes, la composition se faisant terme à terme. Lorsque tous les G_k sont égaux à un même groupe G, ce produit est noté G^n.

Exercice 6 Soient G, H et K des groupes.
(i) Les groupes produits $G \times H$ et $H \times G$ sont (canoniquement) isomorphes.
(ii) Les groupes produits $(G \times H) \times K$, $G \times (H \times K)$ et $G \times H \times K$ sont (canoniquement) isomorphes. Généraliser au cas d'une famille finie quelconque de groupes.
(iii) Si n et m sont des entiers naturels premiers entre eux, les groupes $\mathbb{Z}/mn\mathbb{Z}$ et $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ sont isomorphes — c'est le théorème chinois, ils sont même isomorphes en tant qu'anneaux.

1.2 Sous-groupes engendrés

Proposition (intersection d'une famille de sous-groupes)
Soient G un groupe et $(H_i)_{i\in I}$ une famille de sous-groupes de G. Alors, $\bigcap_{i\in I} H_i$ est un sous-groupe de G.

PREUVE. Exercice. ■

Définition (sous-groupe engendré par une partie)
Soient G un groupe et X une partie de G. Le *sous-groupe de G engendré par X* est l'intersection des sous-groupes de G qui contiennent X. On le notera $\langle X \rangle$.

Proposition (minimalité des sous-groupes engendrés)
Soit G un groupe et X une partie de G. Alors, le sous-groupe engendré $\langle X\rangle$ est le plus petit sous-groupe de G contenant X au sens de l'inclusion, i.e. *si H est un sous-groupe de G, alors $X \subseteq H \Longrightarrow \langle X\rangle \subseteq H$.*

Preuve. Exercice. ■

Proposition (caractérisation des sous-groupes engendrés)
Soient G un groupe et X une partie de G. Alors, le sous-groupe engendré par X est l'ensemble des produits finis d'éléments de X ou de leurs inverses Autrement dit,

$$\langle X\rangle = \left\{x_1x_2\dots x_n,\ n\in\mathbb{N}\setminus\{0\},\ \forall k\in\{1,\dots,n\},\ x_k\in X \text{ ou } x_k^{-1}\in X\right\}.$$

Preuve. On note $E = \left\{x_1x_2\dots x_n,\ n\in\mathbb{N}\setminus\{0\},\ \forall k\in\{1,\dots,n\},\ x_k\in X \text{ ou } x_k^{-1}\in X\right\}$. D'abord, E est un sous-groupe de G. En effet, il n'est pas vide et est évidemment stable par produit et par passage à l'inverse (attention au renversement dans la formule $(x_1x_2\dots x_n)^{-1} = x_n^{-1}\dots x_2^{-1}x_1^{-1}$). Par ailleurs, il contient X. Ainsi, par minimalité des sous-groupes engendrés, $\langle X\rangle \subseteq E$. Enfin, puisque $\langle X\rangle$ est un sous-groupe, il est stable par produit et par passage à l'inverse ; puisqu'il contient X, il contient donc E — raisonner par récurrence sur le n de la définition de E. Ainsi, $\langle X\rangle \supseteq E$. ■

Le slogan : *le sous-groupe engendré par X est l'ensemble des mots formés d'éléments de X ou de leurs inverses (sous-entendu : entre deux lettres du mot, on met le symbole de la loi de groupe).*

A noter
Dans les conditions de la proposition, on note parfois X^{-1} l'ensemble des inverses dans G des éléments de X. Alors, la proposition s'énonce ainsi :

$$\langle X\rangle = \left\{x_1x_2\dots x_n,\ n\in\mathbb{N}\setminus\{0\},\ \forall k\in\{1,\dots,n\},\ x_k\in X\cup X^{-1}\right\}.$$

Définition (groupes cycliques et monogènes)
Un groupe est *monogène* lorsqu'il est engendré par un singleton. Un groupe est *cyclique* lorsqu'il est monogène et fini.

A noter
(i) Ainsi, si un groupe G, dont la loi est notée multiplicativement, est monogène et si $g\in G$ est un générateur de G, alors $G = \langle g\rangle = \{g^n,\ n\in\mathbb{Z}\}$. Si le groupe est noté additivement, alors $G = \{ng,\ n\in\mathbb{Z}\}$.

(ii) Si G est un groupe et si $g\in G$, alors l'ordre de g est l'ordre du groupe monogène $\langle g\rangle$.

Exemples
Le groupe additif $\mathbb{Z}$ est monogène, engendré par le singleton $\{1\}$. Si $n\in\mathbb{N}\setminus\{0\}$, le groupe additif $\mathbb{Z}/n\mathbb{Z}$ est cyclique, engendré par le singleton $\left\{\overline{1}\right\}$ – on a noté $\overline{1}$ la classe du nombre 1 modulo n.

Exercice 7 (en forme de révision du cours d'algèbre de L2)
(i) Si $x\in\mathbb{Z}$, alors $\mathbb{Z} = \langle\{x\}\rangle$ si, et seulement si $x = \pm 1$.

(ii) Soit $n\in\mathbb{N}\setminus\{0\}$. Si $x\in\mathbb{Z}$, alors $\mathbb{Z}/n\mathbb{Z} = \langle\{\overline{x}\}\rangle$ si, et seulement si x et n sont premiers entre eux — on a noté $\overline{x}$ la classe de x modulo n.

(iii) Si $n\in\mathbb{N}\setminus\{0\}$, alors $\mathbb{Z}/n\mathbb{Z}$ est simple si, et seulement si n est un nombre premier.

(iv) Tout groupe monogène infini est isomorphe à $\mathbb{Z}$. Si $n\in\mathbb{N}\setminus\{0\}$, tout groupe cyclique d'ordre n est isomorphe au groupe additif de $\mathbb{Z}/n\mathbb{Z}$, ou encore au groupe multiplicatif $\mathbb{U}_n$ des racines complexes n^{e} de l'unité.

Noter que considérer un groupe cyclique d'ordre n comme étant isomorphe à $(\mathbb{Z}/n\mathbb{Z},+)$ ou à $(\mathbb{U}_n,\times)$ constitue un choix de point de vue, l'un ou l'autre pouvant s'avérer le plus pertinent selon l'argumentation que l'on cherche à apporter.

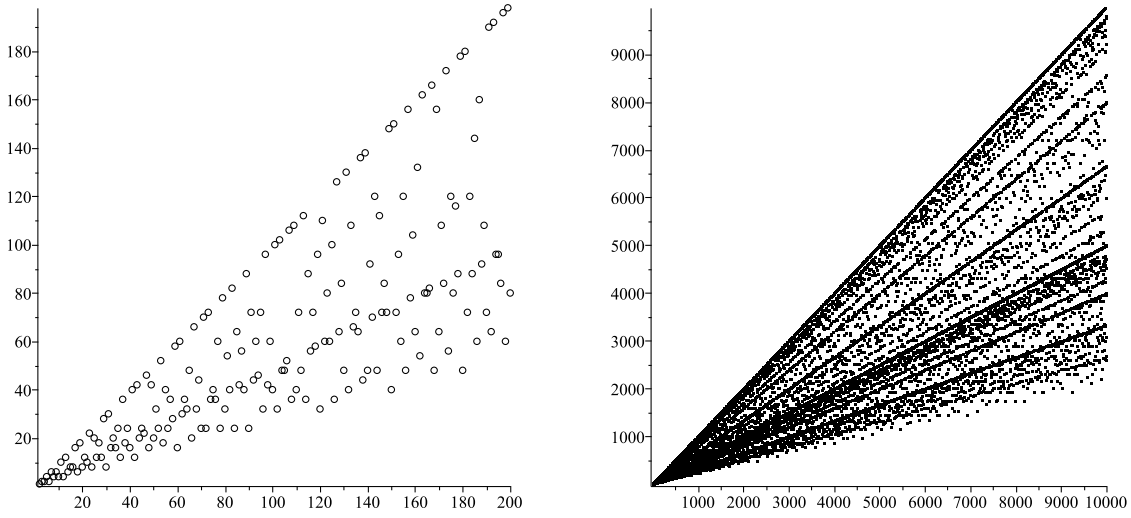
(v) Soient (a,b) et (c,d) dans $\mathbb{Z}^2$. Alors, la paire $\{(a,b),(c,d)\}$ engendre le groupe additif $\mathbb{Z}^2$ si, et seulement si $\det\begin{pmatrix} a & c\\ b & d\end{pmatrix} = ad-bc = \pm 1$.

Définition (fonction d'Euler)
Si n est un entier naturel non nul, on note $\varphi(n)$ le nombre de nombre entiers naturels de $\{1,\dots,n\}$ qui sont premiers avec n. La fonction $\mathbb{N}^* \to \mathbb{N}^*$, $n \mapsto \varphi(n)$ est appelé *fonction (indicatrice) d'Euler*.

A noter
(i) L'exemple (ii) ci-dessus montre que $\varphi(n)$ est le nombre d'éléments de $\mathbb{Z}/n\mathbb{Z}$ qui engendrent le groupe additif $\mathbb{Z}/n\mathbb{Z}$.
Autrement dit, $\varphi(n)$ est le nombre de racines primitives n^{e} de l'unité dans $\mathbb{C}$.
[Une racine n^{e} complexe de l'unité est dite *primitive* lorsqu'elle engendre le groupe $\mathbb{U}_n$ de toutes les racines n^{e} de l'unité. Les racines primitives n^{e} de l'unité sont ainsi les $\exp\left(\frac{2ik\pi}{n}\right)$ où k est premier avec n.]
C'est aussi l'ordre du groupe $(\mathbb{Z}/n\mathbb{Z})^\times$ des inversibles de l'anneau $\mathbb{Z}/n\mathbb{Z}$.

(ii) On dessine ci-dessous le tableau des premières valeurs de φ et son graphe sur $\{1,\dots,200\}$:

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	1	2	2	4	2	6	4	6	4	10	4	12	6	8	8



Graphe de la fonction d'Euler

(iii) Soit p un nombre premier. Alors $\varphi(p) = p-1$. Plus généralement, si n est un entier naturel non nul, $\varphi(p^n) = p^{n-1}(p-1)$.
(iv) Le théorème chinois assure que lorsque n et m sont premiers entre eux, alors l'anneau $\mathbb{Z}/mn\mathbb{Z}$ est isomorphe à l'anneau produit $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$. L'isomorphisme entre les groupes des inversibles de ces deux anneaux assure alors, en considérant les seuls cardinaux, que *si m et n sont premiers entre eux, alors $\varphi(mn) = \varphi(m)\varphi(n)$.*

Proposition (somme des φ)
Pour tout $n \geq 1$, $n = \sum_{d|n} \varphi(d)$.

Preuve. On se place dans le groupe $\mathbb{U}_n$ des racines n^{e} complexes de l'unité. Toute racine n^{e} est une racine primitive d^{e}, où d est un diviseur de n, puisque $\exp\left(\frac{2ik\pi}{n}\right) = \exp\left(\frac{2i(k/\delta)\pi}{n/\delta}\right)$ où $\delta = \text{pgcd}(n,k)$ est une racine primitive n/δ^{e} de l'unité. Comme une racine primitive d^{e} n'est pas une racine primitive d'^{e} si $d \neq d'$, cela montre que la famille des racines primitives d^{e}, lorsque d parcourt l'ensemble des diviseurs de n, constitue une partition de $\mathbb{U}_n$. ■

A noter
On peut aussi partitionner les éléments du groupe additif $\mathbb{Z}/n\mathbb{Z}$ en éléments d'ordre d lorsque d parcourt l'ensemble des diviseurs de n. Cela fournit version légèrement différente de la preuve qui précède.

Exercice 8
(i) La *fonction μ de Möbius* est définie sur $\mathbb{N}^*$ par $\mu(1) = 1$, $\mu(n) = 0$ lorsque n est divisible par le carré d'un entier supérieur ou égal à 2, et $\mu(p_1 p_2 \dots p_r) = (-1)^r$ si les p_k sont des nombres premiers distincts.
Montrer que $\sum_{k|n} \mu(k) = 0$, pour tout entier $n \geq 2$ — et que si $n = 1$, cette somme vaut 1.

(ii) Soient $(a_n)_{n\in\mathbb{N}^*}$ et $(b_n)_{n\in\mathbb{N}^*}$ des suites à valeurs dans un groupe additif (abélien). Montrer qu'il y a équivalence entre :
① pour tout $n \geq 1$, $b_n = \sum_{k|n} a_k$;
② pour tout $n \geq 1$, $a_n = \sum_{k|n} \mu\left(\frac{n}{k}\right) b_k$.

(iii) En déduire que pour tout $n \geq 1$, $\varphi(n) = n \sum_{d|n} \frac{\mu(d)}{d}$.

Définition (commutateur, groupe dérivé)
Soit G un groupe. Un *commutateur* de G en est un élément de la forme $ghg^{-1}h^{-1}$, où $g, h \in G$. Le *groupe dérivé* de G est le sous-groupe de G engendré par ses commutateurs. On le note généralement $D(G)$. Ainsi,

$$D(G) = \left\langle \left\{ ghg^{-1}h^{-1},\ (g,h) \in G^2 \right\} \right\rangle .$$

A noter
(i) En général, le produit de deux commutateurs n'est pas un commutateur.
[Mais trouver des exemples n'est pas si simple !]
(ii) Bien sûr, G est abélien si, et seulement si $D(G) = \{1_G\}$.

Exemple
$D\left(\mathrm{GL}\left(2, \mathbb{Z}/2\mathbb{Z}\right)\right) = \left\langle \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \right\rangle$. A ce stade, ce calcul est un peu fastidieux. Il aura une interprétation limpide une fois le groupe symétrique mis en place.

Proposition
Soit G un groupe. Alors, $D(G) \triangleleft G$.

Preuve. $z\left(xyx^{-1}y^{-1}\right)z^{-1} = \left(zxz^{-1}\right)\left(zyz^{-1}\right)\left(zx^{-1}z^{-1}\right)\left(zy^{-1}z^{-1}\right)$: le conjugué d'un commutateur est encore un commutateur. Cela suffit à prouver le résultat. ■

1.3 Classes à droite et à gauche, théorème de Lagrange

Définitions
Soient G un groupe, H un sous-groupe de G et $x \in G$. La *classe à gauche de x modulo H* est la partie xH de G : la *classe à droite de x modulo H* est la partie Hx de G. Pour préciser les notations évidentes,

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

Exercice 9
Dans les conditions des définitions précédentes, la relation binaire sur G définie par $x \sim y \Leftrightarrow x^{-1}y \in H$ est une relation d'équivalence dont les classes sont les classes à gauche modulo H. De même, la relation binaire sur G définie par $x \sim y \Leftrightarrow xy^{-1} \in H$ est une relation d'équivalence dont les classes sont les classes à droite modulo H.

On notera respectivement $(G/H)_g$ l'ensemble des classes à gauche modulo H et $(G/H)_d$ l'ensemble des classes à droite modulo H. Ces deux ensembles de parties de G forment deux partitions de G, en général distinctes.

Exemples
(i) On note T le sous-groupe de $\mathrm{GL}(2, \mathbb{R})$ formé des matrices trigonales inférieures — le coefficient en haut à droite est nul, les coefficients diagonaux sont non nuls. On note également $j = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \mathrm{GL}(2, \mathbb{R})$. Alors,

les classes à droite et à gauche de j modulo T sont différentes. En effet, jT est le sous-ensemble de $\mathrm{GL}(2,\mathbb{R})$ formé des matrices dont la seconde colonne est proportionnelle à $\begin{pmatrix}1\\1\end{pmatrix}$ alors que Tj est est le sous-ensemble de $\mathrm{GL}(2,\mathbb{R})$ formé des matrices dont la première ligne est proportionnelle à $\begin{pmatrix}1 & 1\end{pmatrix}$.

(ii) Soient $M \in \mathrm{GL}(2,\mathbb{R})$ et $S \in \mathrm{SL}(2,\mathbb{R})$. Alors, $SM = M\left(M^{-1}SM\right)$ avec $\det\left(M^{-1}SM\right) = 1$. De même, $MS = \left(MSM^{-1}\right)M$ avec $\det\left(MSM^{-1}\right) = 1$. Cela montre que les classes à droite et à gauche de M modulo $\mathrm{SL}(2,\mathbb{R})$ sont égales — quel que soit $M \in \mathrm{GL}(2,\mathbb{R})$.

Exercice 10
Soient G un groupe et H un sous-groupe de G. Alors, les assertions suivantes sont équivalentes.
(i) $H \lhd G$
(ii) $xH = Hx$ pour tout $x \in G$
(iii) $xH \subseteq Hx$ pour tout $x \in G$
[L'équivalence entre (ii) et (iii), immédiate lorsque H est fini, est toujours vraie. En effet, si on suppose (iii) vraie, soient $x \in G$ et $h \in H$. On écrit $hx = x\left(x^{-1}hx\right)$. Comme $x^{-1}h \in x^{-1}H \subseteq Hx^{-1}$, soit $h' \in H$ tel que $x^{-1}h = h'x^{-1}$. Alors, $hx = xh' \in xH$, cqfd.]

Proposition (théorème de Lagrange)
Soient G un groupe et H un sous-groupe de G.
(i) Toutes les classes à droites et toutes les classes à gauche modulo H ont un cardinal commun : celui de H.
(ii) Le cardinal de l'ensemble des classe à gauche et le cardinal de l'ensemble des classe à droite sont égaux. Ce cardinal commun est noté $[G : H]$.
(iii) $|G| = |H| \times [G : H]$ — égalité entre cardinaux.
(iv) Si G est un groupe fini, alors $|H|$ divise $|G|$ — et le quotient égale $[G : H]$.

Preuve. (i) Soit $x \in G$. Alors, l'application $H \to xH$, $h \mapsto xh$ est une bijection — sa surjectivité résulte de la définition de xH, son injectivité de l'existence de x^{-1}. Il en va de même pour $H \to Hx$, $h \mapsto hx$. (ii) et (iii) Les classes à gauche forment une partition de G dont toutes les parts ont le même cardinal — celui de H. On conclut avec le théorème des bergers. *Idem* pour les classes à droite. (iv) est une conséquence immédiate de (iii) puisque ces cardinaux sont des nombres. ■

Définition (indice d'un sous-groupe)
Dans les conditions de la proposition précédente, $[G : H]$ est l'*indice de H dans G*.

Exemples
(i) Il n'y a pas de sous-groupe d'ordre 35 dans $\mathbb{Z}/27\mathbb{Z} \times \mathbb{Z}/135\mathbb{Z}$.
(ii) Soit p un nombre premier. Tout homomorphisme de groupes $\mathbb{Z}/p\mathbb{Z} \to G$ est constant ou injectif.

Exercice 11 Tout sous-groupe d'indice 2 est distingué.

Définition (groupe simple)
Un groupe G est dit *simple* lorsque ses seuls sous-groupes distingués sont G et 1_G.

Exemple
Si n est un entier naturel non nul, le groupe $(\mathbb{Z}/n\mathbb{Z}, +)$ est simple si, et seulement si n est un nombre premier.

1.4 Groupe-quotient, théorèmes d'isomorphismes

Lemme (compatibilité de la multiplication modulo un sous-groupe)
Soient G un groupe et H un sous-groupe de G. Les assertions suivantes sont équivalentes.
(i) Pour tous $x, y \in G$, la classe à gauche de xy modulo H ne dépend que de xH et de yH.
(ii) Pour tous $x, y \in G$, la classe à droite de xy modulo H ne dépend que de Hx et de Hy.
(iii) $H \lhd G$

Preuve. On précise ce que signifie l'assertion (i) :

$$\forall x, y, x', y' \in G,\ \left(xH = x'H \text{ et } yH = y'H\right) \Longrightarrow \left(xyH = x'y'H\right). \tag{1}$$

Autrement dit, (i) signifie que la congruence à droite modulo H est compatible avec la loi du groupe. On montre l'équivalence de (i) et de (iii) ; celle entre (ii) et (iii) est du même acabit.
(iii)⇒(i) On suppose que $H \triangleleft G$. Soient $x, x', y, y' \in G$ tels que $xH = x'H$ et $yH = y'H$. Soient alors $h, k \in H$ tels que $x' = xh$ et $y' = yk$. Puisque H est distingué, soit $\ell \in H$ tel que $hy = y\ell$. Dans ces conditions, $x'y' = xhyk = xy(\ell k) \in xyH$. Cela montre que $x'y'H = xyH$.
(i)⇒(iii) On suppose que (1) est vraie. Soient $h \in H$ et $x \in G$. Puisque $h \in 1H$ et $x \in xH$, alors $hx \in 1xH$, ce qui signifie que $hx \in xH$. L'arbitraire sur x et h montre que $H \triangleleft G$. ■

Lorsque $H \triangleleft G$, les classes à gauche et les classes à droite sont les mêmes, au sens où $xH = Hx$ pour tout $x \in G$. On parle alors de *classe modulo* H, sans préciser s'il s'agit d'une classe à droite ou à gauche. Dans ces conditions, le lemme montre qu'on peut définir une loi de composition interne sur les classes modulo H. C'est cette loi qui confère à l'ensemble des classes modulo H une structure de groupe.

Proposition (définition du groupe-quotient modulo un sous-groupe distingué)
Soient G un groupe et H un sous-groupe distingué de G. On note G/H l'ensemble des classes modulo H des élément de G. Alors, l'application $G/H \times G/H \to G/H$, $(xH, yH) \mapsto xyH$ est bien définie et confère à l'ensemble G/H une structure de groupe. Son élément neutre est $H = 1.H$ et pour tout $x \in G$, l'inverse de xH est $x^{-1}H$.

Preuve. Que la loi soit bien définie est conséquence du lemme. Si $x \in G$, on note $\overline{x} = xH = Hx$ la classe de x modulo H. En particulier, $\overline{1} = H$. La définition de la loi sur G/H s'écrit alors $\overline{x} \cdot \overline{y} = \overline{xy}$. Si $x, y, z \in G$, alors $(\overline{x} \cdot \overline{y}) \cdot \overline{z} = \overline{xy} \cdot \overline{z} = \overline{(xy)z} = \overline{x(yz)} = \overline{x} \cdot \overline{yz} = \overline{x} \cdot (\overline{y} \cdot \overline{z})$, ce qui montre l'associativité de la loi. Par ailleurs, $\overline{1} \cdot \overline{x} = \overline{1 \cdot x} = \overline{x}$ et $\overline{x} \cdot \overline{1} = \overline{x \cdot 1} = \overline{x}$: la classe $\overline{1} = H$ est élément neutre. Enfin, si $x \in G$, alors $\overline{x} \cdot \overline{x^{-1}} = \overline{x \cdot x^{-1}} = \overline{1}$ ce qui montre que $\overline{x^{-1}}$ est inverse à droite de x. Un calcul analogue montre que $\overline{x^{-1}}$ est également inverse à gauche de $\overline{x}$. ■

A noter
Le lemme précédent montre aussi que si H n'est pas distingué, il est vain de chercher à définir une loi de groupe sur les classes à droite ou à gauche modulo H par une formule du type $xH.yH = xyH$ ou $Hx.Hy = Hxy$.

Définition (projection canonique)
Soient G un groupe et H un sous-groupe distingué de G. L'homomorphisme de groupes

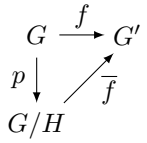
$$\begin{array}{rcl} p : G & \longrightarrow & G/H \\ x & \mapsto & \overline{x} = xH = Hx \end{array}$$

est appelé *surjection canonique* ou *projection canonique* — que ce soit un homomorphisme de groupes résulte immédiatement de la définition de la loi de groupe sur G/H dont c'est une paraphrase.

La propriété universelle du quotient, qui suit, est l'énoncé opératoire des groupes-quotient. C'est elle qui permet notamment de définir des homomorphismes de groupes dont l'ensemble de départ est un quotient avec le confort argumentaire procuré par son automatisme technique.

Proposition (propriété universelle des groupes-quotient)
Soient G et G' des groupes, $f : G \to G'$ un homomorphisme de groupes et H un sous-groupe distingué de G. On note $p : G \to G/H$ la projection canonique. On suppose que $H \subseteq \ker f$. Alors, il existe un unique homomorphisme de groupes $\overline{f} : G/H \to G'$ tel que $f = \overline{f} \circ p$. En outre,
(i) $\ker \overline{f} = p(\ker f) = \{xH,\ x \in \ker f\}$ *; en particulier, $\overline{f}$ est injectif si, et seulement si* $\ker f = H$ *;*
(ii) $\operatorname{im} \overline{f} = \operatorname{im} f$ *; en particulier, $\overline{f}$ est surjectif si, et seulement si f l'est.*

Le diagramme commutatif (et même cartésien) standard est le suivant :



Preuve. On prouve d'abord l'unicité. On suppose que $\overline{f} : G/H \to G'$ existe et vérifie $f = \overline{f} \circ p$. Alors, pour tout $x \in G$, $\overline{f}(\overline{x}) = f(x)$. Cela montre que l'application $\overline{f}$ cherchée est déterminée par f, ce qui prouve

l'unicité. Maintenant l'existence. On n'a pas le choix, le raisonnement précédent oblige à définir $\overline{f}$ par la formule $\overline{f}(\overline{x}) = f(x)$. Il s'agit de montrer que cette formule a du sens, ce qui est garanti par l'hypothèse $H \subseteq \ker f$. En effet, si $xH = yH$, alors $x^{-1}y \in H \subseteq \ker f$ et donc $f(x) = f(y)$: l'application f est constante sur les classes de congruence modulo H. Il reste à vérifier que $\overline{f}$ est un homomorphisme de groupes, ce qui est une pure routine. (i) et (ii) sont immédiates. ■

Exemple
Soient m et n deux entiers naturels non nuls. On suppose que m divise n, c'est-à dire que $n\mathbb{Z} \subseteq m\mathbb{Z}$. Le groupe $\mathbb{Z}$ est abélien, inutile de se préoccuper de distinction pour passer au quotient : la projection canonique $\mathbb{Z} \to \mathbb{Z}/mZ$ se factorise, *via* la projection canonique $\mathbb{Z} \to \mathbb{Z}/nZ$, en un homomorphisme surjectif d'anneaux $\mathbb{Z}/nZ \to \mathbb{Z}/m\mathbb{Z}$. Ce dernier envoie la classe modulo n d'un entier sur sa classe modulo m. Son noyau est l'ensemble des classes modulo n des multiples de m. C'est le sous-groupe cyclique $\langle\overline{m}\rangle$ de $\mathbb{Z}/n\mathbb{Z}$, où $\overline{m}$ désigne la classe de m modulo n. Il est d'ordre $\frac{n}{m}$.

Théorème (premier théorème d'isomorphisme pour les groupes)
Soit $f : G \to G'$ un homomorphisme de groupes. Alors, f induit un isomorphisme de groupes

$$\boxed{G/\ker f \simeq \operatorname{im} f}$$

PREUVE. C'est une application directe de la PUQ. ■

Exemples
(i) Si V est un espace vectoriel de dimension finie sur un corps $\mathbb{F}$, le déterminant induit un isomorphisme de groupes

$$\mathrm{GL}(V)/\mathrm{SL}(V) \simeq \mathbb{F}^*.$$

En effet, l'homomorphisme de groupes $\det : \mathrm{GL}(V) \to \mathbb{F}^\times$ a pour noyau $\mathrm{SL}(V)$. En outre, il est surjectif. Pour montrer ce dernier point, on peut prendre une base de V qui établit une application linéaire bijective $\mathbb{F}^{\dim V} \xrightarrow{\sim} V$ et aussi un isomorphisme de groupes $\mathrm{GL}(V) \simeq \mathrm{GL}(\dim V, \mathbb{F})$. Trouver un automorphisme de V dont le déterminant $x \in \mathbb{F}^\times$ est prescrit devient alors un jeu d'enfants : il suffit de considérer la matrice diagonale $\mathrm{diag}(x, 1, \dots 1)$.

(ii) L'exponentielle imaginaire $(\mathbb{R}, +) \mapsto (\mathbb{C}^\times, \times)$, $t \mapsto \exp(it)$ est un homomorphisme de groupes (l'exponentielle d'une somme est le produit des exponentielles). Son image est le groupe multiplicatif $S^1 = \{z \in \mathbb{C},\ |z| = 1\}$ des nombres complexes de module 1. Son noyau est le sous-groupe $2\pi\mathbb{Z}$ de $\mathbb{R}$. Le premier théorème d'isomorphisme montre que l'exponentielle imaginaire induit un isomorphisme de groupes

$$\mathbb{R}/2\pi\mathbb{Z} \simeq S^1.$$

A vrai dire, l'exponentielle jouit d'autre propriétés intéressantes, notamment topologiques. Ces propriétés se transmettent ou se traduisent presque toujours sur l'isomorphisme $\mathbb{R}/2\pi\mathbb{Z} \simeq S^1$. Par exemple, à condition de définir proprement une topologie sur le quotient (par exemple, la *topologie quotient* !), on obtient un homéomorphisme.

(iii) Soit $F = \left\{ \begin{pmatrix} 1 & a \\ 0 & b \end{pmatrix} \in \mathcal{M}_2(\mathbb{R}),\ a, b \in \mathbb{R},\ b \neq 0 \right\}$ l'ensemble des matrices de $\mathrm{GL}(2, \mathbb{R})$ qui fixent le vecteur-colonne ${}^t(1, 0)$. Alors, F est un sous-groupe de $\mathrm{GL}(2, \mathbb{R})$ et l'application $F \to \mathbb{R}^\times$, $\begin{pmatrix} 1 & a \\ 0 & b \end{pmatrix} \mapsto b$ est un homomorphisme de groupes surjectif dont le noyau est le sous-groupe $U = \left\{ \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix},\ a \in \mathbb{R} \right\}$ de F, isomorphe à $(\mathbb{R}, +)$ (exercice). Le premier théorème d'isomorphisme montre que les groupes F/U et $\mathbb{R}^\times$ sont isomorphes.

Exercice 12
Sur le mode de l'exemple (ii) ci-dessus, trouver des isomorphismes entre les groupes suivants.
(i) $\mathbb{Z}/n\mathbb{Z} \simeq \mathbb{U}_n$ si n est un entier naturel non nul et si $\mathbb{U}_n$ désigne le groupe des racines n^{e} complexes de l'unité, qui est un sous-groupe fini de S^1.
(ii) $(\mathbb{C}/2i\pi\mathbb{Z}, +) \simeq (\mathbb{C}^\times, \times)$.

(iii) $\mathbb{Q}/\mathbb{Z} \simeq \mathbb{U}_\infty$ où $\mathbb{U}_\infty$ désigne le groupe multiplicatif de *toutes* les racines complexes de l'unité (les racines carrées, cubiques, 4^{e}, 5^{e}, etc), qui est un sous-groupe dense de S^1.

A noter
Soient G un groupe, H un sous-groupe distingué de G et $p : G \to G/H$ la surjection canonique. Alors, les sous-groupes de G/H sont les classes modulo H des sous-groupes de G contenant H, en le sens suivant.

① Si $\mathcal{K}$ est un sous-groupe de G/H, alors $p^{-1}(\mathcal{K})$ est un sous-groupe de G contenant H.

② Inversement, si K est un sous-groupe de G contenant H alors $H \lhd K$ et le groupe K/H, qui est l'ensemble des classes modulo H des éléments de K, est un sous-groupe de G/H.

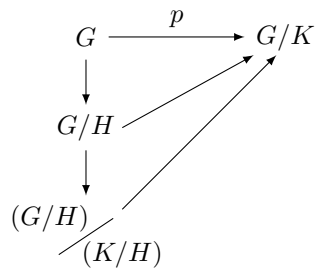
③ Les deux opérations $\mathcal{K} \mapsto p^{-1}(\mathcal{K})$ et $K \mapsto K/H$ définissent deux bijections réciproques l'une de l'autre entre l'ensemble des sous-groupes de G/H et l'ensemble des sous-groupes de G contenant H.

Théorème (deuxième théorème d'isomorphisme pour les groupes)
Soient G un groupe, H et K deux sous-groupes distingués de G, tels que $H \subseteq K$. Alors, $K/H \lhd G/H$ et la projection canonique $G \mapsto G/K$ induit un isomorphisme de groupes

$$\boxed{(G/H)\,/\,(K/H) \simeq G/K}$$

Preuve. Que H soit un sous-groupe distingué de K est immédiat et donne du sens au groupe-quotient K/H. Soit $p : G \mapsto G/K$ la projection canonique. Puisque $H \lhd G$ et $H \subseteq \ker p = K$, la propriété universelle du quotient induit un homomorphisme surjectif de groupes $G/H \to G/K$ dont le noyau est $p(K) = \{kH,\ k \in K\} = K/H$. On conclut avec le premier théorème d'isomorphisme. ■

A noter
On peut résumer la situation et la preuve du deuxième théorème d'isomorphisme par le diagramme commutatif suivant. S'assurer de bien comprendre comment les flèches sont définies (quels homomorphismes d'anneaux elles représentent).



Théorème (troisième théorème d'isomorphisme pour les groupes)
Soient G un groupe, H et K deux sous-groupes de G. On suppose que H normalise K, ce qui signifie que $hKh^{-1} = K$, pour tout $h \in H$. Alors,
(i) $H \cap K$ est un sous-groupe distingué de H ;
(ii) si on note $HK = \{hk,\ h \in H,\ k \in K\}$, alors $HK = KH$ et HK est un sous-groupe de G ;
(iii) K est un sous-groupe distingué de HK ;
(iv) on a un isomorphisme de groupes

$$\boxed{HK/K \simeq H/H \cap K}$$

Preuve. D'abord, l'hypothèse que H normalise K assure que $HK = KH$. En effet, si $h \in H$ et $k \in K$, alors $k' = hkh^{-1} \in K$ et donc $hk = k'h \in KH$. De même, $k'' = h^{-1}kh \in K$ et donc $kh = hk'' \in HK$: on a montré que $KH = HK$.
Si $h, h' \in H$ et $k, k' \in K$, soit $k'' \in K$ tel que $kh' = h'k''$. Alors, $hk \cdot h'k' = hh' \cdot k''k \in HK$ et $(hk)^{-1} = k^{-1}h^{-1} \in KH = HK$: on a montré que HK est un sous-groupe de G.

Si $k' \in K$ et si $hk \in HK$, alors $(hk)k'(hk)^{-1} = h\left(kk'k^{-1}\right)h^{-1} \in K$, ce qui montre que $K \triangleleft HK$. On peut donc considérer le groupe-quotient HK/K.
Enfin, on compose l'inclusion $H \hookrightarrow HK$ et la projection canonique $HK \to HK/K$, ce qui fournit un homomorphisme de groupes $f : H \to HK/K$ dont le noyau est $H \cap K$. En outre, si $hk \in HK$, alors $f(h) = f(hk)$ est la classe de hk modulo K, ce qui montre que f est surjectif. On conclut en appliquant le premier théorème d'isomorphisme à f. ■

Exercice 13 Soient G un groupe.

(i) Montrer que le groupe-quotient $G/D(G)$ est abélien. On appelle ce groupe l'*abélianisé* de G.

(ii) Soit H un sous-groupe distingué de G. On suppose que le groupe-quotient G/H est abélien. Montrer que $D(G) \subseteq H$ et que G/H est isomorphe à un quotient du groupe abélien $G/D(G)$.

[Le slogan, littéralement impropre mais parlant et souvent entendu, est le suivant : *$G/D(G)$ est le plus grand quotient abélien de G.* Son sens précis est : tout quotient abélien d'un groupe est (isomorphe à) un quotient de son abélianisé.]

2 Le groupe symétrique

2.1 Permutations d'un ensemble fini

Définition (groupe symétrique d'un ensemble)
Si E est un ensemble non vide, toute application bijective $E \to E$ est appelée *permutation de E*. Le *groupe symétrique* de E est l'ensemble des permutations de E muni de la composition des applications. On le notera $\mathfrak{S}_E$.

A noter
(i) L'unique bijection $\emptyset \to \emptyset$ est celle dont le graphe est vide. On convient de dire que $\mathfrak{S}_\emptyset = \{\emptyset\}$ est le groupe trivial.

(ii) L'élément neutre de $\mathfrak{S}_E$ est l'application identique de E, que l'on notera id_E, ou même parfois 1. Le symétrique d'un élément $\sigma \in \mathfrak{S}_E$ est sa réciproque. Bien souvent, on omettra de noter la composition : si $\sigma, \tau \in \mathfrak{S}_E$, on notera $\sigma \circ \tau = \sigma\tau$. De même, si $n \in \mathbb{N} \setminus \{0\}$, on définit par récurrence $\sigma^n = \sigma \circ \cdots \circ \sigma$ (n fois) et $\sigma^{-n} = \left(\sigma^{-1}\right)^n = (\sigma^n)^{-1}$. Par commodité, on note aussi $\sigma^0 = \mathrm{id}_E$.

Proposition (le groupe symétrique ne dépend que du cardinal)
Soient E et F deux ensembles non vides équipotents. Alors, les groupes symétriques $\mathfrak{S}_E$ et $\mathfrak{S}_F$ sont isomorphes.

Preuve. Soit $\varphi : E \to F$ une application bijective. Alors, l'application $\mathfrak{S}_E \to \mathfrak{S}_F$, $\sigma \mapsto \varphi \circ \sigma \circ \varphi^{-1}$ est un isomorphisme de groupes dont la réciproque est $\mathfrak{S}_F \to \mathfrak{S}_E$, $\sigma \mapsto \varphi^{-1} \circ \sigma \circ \varphi$. ■

Définition (groupe $\mathfrak{S}_n$)
Si n est un entier naturel non nul, on note $\mathfrak{S}_n = \mathfrak{S}_{\{1,\dots,n\}}$. Fort de la proposition précédente, on l'appelle parfois *groupe des permutations de n objets*.

Exercice 14
(i) Si n est un entier naturel non nul, l'ordre de $\mathfrak{S}_n$ est $n!$.

(ii) *Tout groupe est (isomorphe à) un sous-groupe d'un groupe de permutations.*
En effet, soit G un groupe. Si $x \in G$, on note $s_x : G \to G$ l'application définie par $s_x(y) = xy$, pour tout $y \in G$; montrer que s_x est une permutation de G. Montrer que l'application $G \to \mathfrak{S}_G$, $x \mapsto s_x$ est un homomorphisme injectif de groupes. En déduire que G est isomorphe à un sous-groupe de $\mathfrak{S}_G$.

(iii) Montrer que si $\emptyset \neq E \subseteq F$, alors $\mathfrak{S}_E$ est (canoniquement) isomorphe à un sous-groupe de $\mathfrak{S}_F$.
[Pour prolonger à F une permutation de E, fixer tous les éléments de $F \setminus E$.]

Définition (support d'une permutation)
Soient E un ensemble et σ une permuttation de E. Le *support de σ* est le complémentaire dans E de l'ensemble de ses points fixes. On le note $\mathrm{Supp}(\sigma)$. Ainsi,

$$\mathrm{Supp}(\sigma) = \{x \in E,\ \sigma(x) \neq x\}.$$

Exercice 15
Le support d'une permutation ainsi que son complémentaire sont *stables* par cette permutation. Cela signifie que si $s \in \mathfrak{S}_E$ alors $s(x) \in \mathrm{Supp}(s)$, pour tout $x \in \mathrm{Supp}(s)$ et $s(x) \in E \setminus \mathrm{Supp}(s)$, pour tout $x \in E \setminus \mathrm{Supp}(s)$.

On se concentre davantage sur le calcul des permutations d'un ensemble fini, c'est-à-dire sur la structure du groupe $\mathfrak{S}_n$. Dans cette étude, la notion de cycle est essentielle.

Définition (cycle, ou permutation cyclique)
Soient E un ensemble, p un entier naturel non nul et $a_1, \dots, a_p$ des éléments distincts de E. Le *p-cycle noté* $(a_1, \dots, a_p)$ est la permutation c de E définie par :

$$\begin{cases} \forall k \in \{1, \dots, p-1\},\ c(a_k) = a_{k+1} \\ c(a_p) = a_1 \\ \forall x \in E \setminus \{a_1, \dots, a_p\},\ c(x) = x. \end{cases}$$

Le nombre p est la *longueur* du cycle $(a_1, \dots, a_p)$. Lorsqu'il n'y a pas d'ambiguïté, on note parfois les cycles sans virgule : $(a_1, \dots, a_p) = (a_1 \dots a_p)$.

A noter

(i) Si $a \in E$, le 1-cycle (a) est l'application identique : son support est vide. Si $p \geq 2$, le support d'un p-cycle $(a_1, \ldots, a_p)$ est l'ensemble $\{a_1, \ldots, a_p\}$. Cela montre en particulier que la notion de longueur est bien définie — attention au cycle de longueur 1, un peu stupide.

(ii) Il n'y a pas d'unicité de l'écriture d'un cycle noté à l'aide des parenthèses. Par exemple, dans $\mathfrak{S}_5$, $(12) = (21)$ et $(12345) = (23451) = (34512) = (45123) = (51234)$. Dans $\mathfrak{S}_n$, lorsqu'on a besoin d'une forme univoque, on choisit souvent de placer le plus petit élément du support en première position : on préférera ainsi (12), (12345) et (26374) aux autres écritures.

(iii) Prendre garde à la notation lorsque l'on compose des cycles. Par exemple, dans $\mathfrak{S}_6$, $(1543)(246) = (154623)$ et $(246)(1543) = (156243)$.

(iv) $\mathfrak{S}_2$ est isomorphe à $\mathbb{Z}/2\mathbb{Z}$, unique groupe d'ordre 2 à isomorphisme près. En revanche, dès que $n \geq 3$, le groupe $\mathfrak{S}_n$ n'est pas commutatif puisque $(12)(23) = (123)$ et $(23)(12) = (132)$ sont distincts.

Exercice 16

Soient n et p des entiers naturels non nuls. Si $c = (a_0, \ldots, a_{p-1}) \in \mathfrak{S}_n$ et si $m \in \mathbb{Z}$, alors $c^m(a_k) = a_{\widetilde{k+m}}$ où $\widetilde{k+m}$ est le reste de la division euclidienne de $k+m$ par p.

Proposition (ordre d'un p-cycle)

Pour tout $p \in \mathbb{N} \setminus \{0\}$, l'ordre d'un p-cycle est p.

Preuve. C'est immédiat. On peut le voir comme une conséquence de l'exercice précédent. ■

Proposition (permutations à supports disjoints)

Deux permutations dont les supports sont disjoints commutent.

Preuve. Soient s et t deux permutations d'un ensemble E. On suppose que $\mathrm{Supp}(s) \cap \mathrm{Supp}(t) = \emptyset$. Soit $x \in E$. Si $x \in \mathrm{Supp}(s)$, alors $x \notin \mathrm{Supp}(t)$ et $s(x) \notin \mathrm{Supp}(t)$ puisque $s(x) \in \mathrm{Supp}(s)$; ainsi, $st(x) = s(t(x)) = s(x)$ et $ts(x) = t(s(x)) = s(x)$. Dans ce cas, $st(x) = ts(x)$. De la même façon, si $x \in \mathrm{Supp}(t)$, alors $st(x) = ts(x) = t(x)$. Enfin, si $x \notin \mathrm{Supp}(s) \cup \mathrm{Supp}(t)$, alors $st(x) = s(t(x)) = s(x) = x$ et $ts(x) = t(s(x)) = t(x) = x$, ce qui montre encore que $st(x) = ts(x)$. Dans tous les cas, $st(x) = ts(x)$ pour tout $x \in E$, ce qui montre que $st = ts$. ■

Exercice 17

Montrer que si s et t sont deux permutations à supports disjoints de $\mathfrak{S}_n$, alors l'ordre de st est le PPCM des ordres de s et de t. Etendre ce résultat au produit d'un nombre quelconque de cycles à supports disjoints.

Proposition (formule de conjugaison des cycles)

Soient E un ensemble, p un entier naturel non nul, $a_1, \ldots, a_p$ des éléments distincts de E et $\sigma \in \mathfrak{S}_E$. Alors,

$$\boxed{\sigma(a_1, \ldots, a_p)\sigma^{-1} = (\sigma(a_1), \ldots, \sigma(a_p))}$$

En particulier, le conjugué d'un p-cycle est un p-cycle.

Preuve. On note $c = (a_1, \ldots, a_p)$ et $c' = (\sigma(a_1), \ldots, \sigma(a_1))$. Soit $x \in E$. Si $x = \sigma(a_k)$ où $k \in \{1, \ldots, p-1\}$, alors $\sigma c \sigma^{-1}(x) = \sigma c(a_k) = \sigma(a_{k+1}) = c'(x)$. De même, si $x = \sigma(a_p)$, alors $\sigma c \sigma^{-1}(x) = \sigma(a_1) = c'(x)$. Enfin, si $x = \sigma(y)$ où $y \in E \setminus \mathrm{Supp}(c)$, alors $\sigma c \sigma^{-1}(x)\, \sigma c(y) = \sigma(y) = x$ et, puisque x n'est pas dans le support de c', on a également $c'(x) = x$. Dans tous les cas, $c(x) = c'(x)$: on a montré que $c = c'$. ■

Proposition (centre de $\mathfrak{S}_n$)

Soit n un entier naturel. Si $n \geq 3$, le centre de $\mathfrak{S}_n$ est trivial.

A noter

En abrégé, $Z(\mathfrak{S}_n) = \{\mathrm{id}\}_{\{1,\ldots,n\}}$ si $n \geq 3$. Le cas $n = 2$ est à part puisque $\mathfrak{S}_2 \simeq \mathbb{Z}/2\mathbb{Z}$ est abélien : $Z(\mathfrak{S}_2) = \mathfrak{S}_2$.

Preuve. Soit $\sigma \in \mathfrak{S}_n \setminus \left\{\mathrm{id}_{\{1,\ldots,n\}}\right\}$. Soit alors $a \in \{1, \ldots, n\}$ tels que $\sigma(a) \neq a$. On note $b = \sigma(a)$. Puisque $n \geq 3$, soit $c \in \{1, \ldots, n\} \setminus \{a, b\}$. Alors, $\sigma(ac)\sigma^{-1} = (b\sigma(c)) \neq (ac)$ puisque $b \notin \{a, c\}$. Cela montre que σ ne commute pas avec le cycle (ab) donc n'est pas dans $Z(\mathfrak{S}_n)$. ■

Proposition (dans le groupe symétrique, tous les p-cycles sont conjugués)

Soient E un ensemble, p un entier naturel non nul, c et c' deux p-cycles de $\mathfrak{S}_E$. Alors, il existe $\sigma \in \mathfrak{S}_E$ tel que $c' = \sigma c \sigma^{-1}$.

PREUVE. On note $c = (a_1, \dots, a_p)$ et $c' = (b_1, \dots, b_p)$. Soit σ une permutation de E qui vérifie $\sigma(a_k) = b_k$ pour tout $k \in \{1, \dots, p\}$. Il en existe puisque les ensembles $E \setminus \mathrm{Supp}(c)$ et $E \setminus \mathrm{Supp}(c')$ sont équipotents : il suffit de prolonger l'application σ ainsi définie sur $\mathrm{Supp}(c)$ à E tout entier en choisissant une bijection quelconque $E \setminus \mathrm{Supp}(c) \to E \setminus \mathrm{Supp}(c')$. Alors, la formule de conjugaison des cycles montre que $c' = \sigma c \sigma^{-1}$. ■

Deux exemples

(i) Une permutation de $\mathfrak{S}_n$ est parfois notée en dressant la liste des images successives de 1, 2, etc. Ainsi, on notera entre crochets $\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 2 & 3 & 5 & 1 \end{bmatrix}$ ou encore $[4, 2, 3, 5, 1]$ la permutation $s \in \mathfrak{S}_5$ qui vérifie $s(1) = 4$, $s(2) = 2$, $s(3) = 3$), $s(4) = 5$ et $s(5) = 1$.
Par exemple, $[6, 13, 12, 14, 4, 1, 7, 10, 2, 9, 11, 3, 8, 5] = (1, 6)(2, 13, 8, 10, 9)(3, 12)(4, 14, 5)$. Cette égalité se vériife en montrant que les deux permutations de $\mathfrak{S}_{14}$ écrites, l'une en donnant la suite des images, l'autre en produit de cycles, envoient tout élément de $\{1, \dots, 14\}$ sur la même image. Noter que le membre de droite est un produit de cycles à supports disjoints.

(ii) Autre exemple dans $\mathfrak{S}_5$: $(123)(325)(153)(1254) = (154)(23)$. Là encore, montrer image par image que les deux permutations fournies sont égales. A noter : dans le membre de droite, on écrit le produit de cycles du membre de gauche comme un produit de cycles *à supports disjoints*.

Proposition (décomposition en produit de cycles disjoints)
Soit n un entier naturel non nul. Toute permutation de $\mathfrak{S}_n$ se décompose en un produit de cycles à supports disjoints. A l'ordre près des facteurs, cette décomposition est unique.

Autrement dit, en termes plus explicites :
pour tout $\sigma \in \mathfrak{S}_n$, il existe $m \in \mathbb{N}$ et un ensemble $\{c_1, \dots, c_m\}$ de cycles de $\mathfrak{S}_n \setminus \{\mathrm{id}_E\}$ tels que
(i) $\forall j, k \in \{1 \dots, m\}$, $\big(j \neq k\big) \Longrightarrow \big(\mathrm{Supp}\,(c_j) \cap \mathrm{Supp}\,(c_j) = \emptyset\big)$;
(ii) $\sigma = c_1 \dots c_m$.
En outre, si $\sigma = c_1 \dots c_m = c'_1 \dots c'_\ell$ sont deux telles décompositions de σ en produits de cycles à supports disjoints, alors $m = \ell$ et $\exists s \in \mathfrak{S}_m$, $\forall k \in \{1, \dots, m\}$, $c'_k = c_{s(k)}$.
Noter que le cas $\sigma = \mathrm{id}_E$ correspond au cas où $m = 0$, c'est-à-dire au cas où l'ensemble des cycles qui décomposent σ est vide.

PREUVE. Opératoire, la preuve présentée ici reprend l'idée algorithmique des deux exemples qui précèdent l'énoncé du théorème de décomposition.
Pour toute permutation $\sigma \in \mathfrak{S}_n$ et pour tout $x \in \{1, \dots, n\}$, on appelle *orbite de x sous σ* le sous-ensemble $\big\{\sigma^k(x),\ k \in \mathbb{Z}\big\}$ de $\{1, \dots, n\}$. On la notera $\omega(x, \sigma)$, ou simplement $\omega(x)$; c'est une partie de $\{1, \dots, n\}$ stable par σ au sens où $\forall y \in \omega(x)$, $\sigma(y) \in \omega(x)$. En particulier, grâce à cette stabilité, la restriction de σ à $\omega(x)$ définit une permutation de $\omega(x)$, que l'on notera $\sigma_{\omega(x)}$. En outre, $\sigma_{\omega(x)}$ est un cycle de longueur maximale de $\mathfrak{S}_{\omega(x)}$. En effet, si p est le nombre d'éléments de $\omega(x)$, alors $p \geq 1$, $\sigma^p(x) = x$ et $\omega(x) = \big\{x, \sigma(x), \dots, \sigma^{p-1}(x)\big\}$, ce qui implique que $\sigma_{\omega(x)}$ soit le p-cycle $\big(x, \sigma(x), \dots, \sigma^{p-1}(x)\big)$ de $\mathfrak{S}_{\omega(x)}$. On prolonge $\sigma_{\omega(x)}$ en une permutation de $\mathfrak{S}_n$ en fixant tous les nombres qui ne sont pas dans $\omega(x)$: on obtient ainsi un cycle $c_{\omega(x)}$ de $\mathfrak{S}_n$ qui vérifie : $\forall k \in \omega(x)$, $c_{\omega(x)}(k) = \sigma(k)$ et $\forall x \in \{1, \dots, n\} \setminus \omega(x)$, $c_{\omega(x)}(k) = k$.
Ainsi, pour chaque orbite ω, on a construit un cycle c_ω de $\mathfrak{S}_n$ qui vérifie :

$$\begin{cases} \forall x \in \omega,\ c_\omega(x) = \sigma(x) \\ \forall x \in \{1, \dots, n\} \setminus \omega,\ c(x) = x. \end{cases}$$

Par ailleurs, les orbites sous σ forment une partition de $\{1, \dots, n\}$ — elles sont les classes de la relation d'équivalence $x \sim y \Leftrightarrow \exists k \in \mathbb{Z},\ y = \sigma^k(x)$. Noter en passant que les orbites constituées de singleton sont celles des points fixes de σ : pour une telle orbite $\omega = \omega(x) = \{x\}$, $\sigma(x) = x$ et $c_\omega = \mathrm{id}_{\{1,\dots,n\}}$.
Pour prouver l'existence de la décomposition cherchée, on montre que σ est le produit des cycles des orbites non triviales :

$$\sigma = \prod_{\#\omega \geq 2} c_\omega.$$

Noter que ce produit a du sens parce que les orbites forment une partition de $\{1, \dots, n\}$, ce qui entraîne que les cycles c_ω commutent entre eux : l'ordre dans lequel on effectue ce produit est indifférent. Une fois cette

construction faite, prouver l'égalité est élémentaire : on note $\tau = \prod_{\#\omega \geq 2} c_\omega$ et on montre que $\tau = \sigma$. Soit $x \in \{1, \dots, n\}$. Si $\sigma(x) = x$, alors $\omega(x) = \{x\}$: x est fixé par tous les c_ω et donc $\tau(x) = x = \sigma(x)$. Si $\sigma(x) \neq x$, alors $\#\omega(x) \geq 2$, $c_{\omega(x)}(x) = \sigma(x)$ et $c_\omega(x) = x$ pour tout orbite ω différente de $\omega(x)$; ainsi, là encore, $\tau(x) = \sigma(x)$. Ainsi, $\sigma(x) = \tau(x)$ pour tout $x \in \{1, \dots, n\}$, ce qui prouve que $\sigma = \tau$.

Il reste à prouver l'unicité de la décomposition. On suppose que $\sigma = c_1 \dots c_m$ où les c_k sont des cycles de longueurs supérieures ou égales à 2 et à supports disjoints. Alors, pour tout $k \in \{1, \dots, m\}$, le support de c_k est une orbite sous σ que l'on note ω_k. Par construction, c_k est le cycle c_{ω_k} défini plus haut, ce qui prouve l'unicité cherchée. ■

Le slogan *Les cycles qui décomposent une permutation sont (induits par) les restrictions à ses orbites.*

Proposition (CNS pour que deux permutations soient conjuguées)

Soient n un entier naturel non nul, $\sigma, \tau \in \mathfrak{S}_n$. Alors, σ et τ sont conjuguées si, et seulement si leurs décompositions en produits de cycles à supports disjoints ont le même forme *en le sens suivant : il existe $r \in \mathbb{N}$, des cycles $s_1, \dots s_r$ à supports disjoints et des cycles $t_1, \dots t_r$ à supports disjoints tels que :*

(i) $\forall k \in \{1, \dots, r\}$, les cycles s_k et t_k ont la même longueur ;

(ii) $\sigma = s_1 \dots s_r$ et $\tau = t_1, \dots t_r$.

Preuve. Que le conjugué d'un produit $s_1 \dots s_r$ ait la même forme vient de la formule de conjugaison des cycles : si σ et τ sont conjugués, ils ont la même forme. Inversement, soient $s_1, \dots s_r, t_1, \dots t_r$ comme dans l'énoncé. Puisque les supports des s_k sont disjoints et ceux des t_k aussi, il existe une permutation $\pi \in \mathfrak{S}_n$ telle que $t_k = \pi s_k \pi^{-1}$, pour tout $k \in \{1, \dots, r\}$ — en toute rigueur, faire une récurrence sur r. Alors, $t_1, \dots t_r = \pi s_1 \dots s_r \pi^{-1}$. ■

Exercice 18

Calculer les classes de conjugaison dans $\mathfrak{S}_2$, $\mathfrak{S}_3$, $\mathfrak{S}_4$, $\mathfrak{S}_5$ et $\mathfrak{S}_6$. Comment calculer le nombre de classes de conjugaisons dans $\mathfrak{S}_n$?

Définition (transposition) Une *transposition* est un 2-cycle.

Proposition (les transpositions engendrent $\mathfrak{S}_n$)

Soit n un entier naturel non nul. Le groupe $\mathfrak{S}_n$ est engendré par ses transpositions.

Preuve. Il s'agit de montrer que toute permutation est un produit de transpositions. Grâce au théorème de décomposition en produit de cycles disjoints, il suffit de montrer que tout cycle de $\mathfrak{S}_n$ est un produit de transpositions. Le calcul élémentaire $(a_1, \dots a_p) = (a_1 a_2)\,(a_2 a_3) \dots (a_{p-1} a_p)$ le montre. ■

Exercice 19

(i) Les transpositions de la forme $(k, k+1)$ engendrent $\mathfrak{S}_n$.

(ii) Si $n \geq 4$, les 4-cycles engendrent $\mathfrak{S}_n$.

[On pourra s'appuyer sur le calcul $(12) = (1324)(1234)^2$).]

Proposition (groupe de Klein)

$K = \{\mathrm{id}, (12)(34), (13)(24), (14)(23)\}$ *est un sous-groupe distingué de $\mathfrak{S}_4$, isomorphe à $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$.*

Ce groupe est appelé *groupe de Klein*[♫] On le verra, la situation $K \triangleleft \mathfrak{S}_4$ est exceptionnelle.

Preuve. Pour montrer que K est un sous-groupe de $\mathfrak{S}_4$ et qu'il est isomorphe à $(\mathbb{Z}/2\mathbb{Z})^2$, il suffit d'en faire la table. Pour montrer qu'il est distingué dans $\mathfrak{S}_4$, il suffit de montrer qu'il est invariant par la conjugaison des transpositions puisque ces dernières engendrent $\mathfrak{S}_4$. A renumérotation près, les deux calculs suivants suffisent pour conclure : $(12)(12)(34)(12) = (12)(34)$ et $(13)(12)(34)(13) = (14)(23)$. ■

A noter

Pour montrer que $K \triangleleft \mathfrak{S}_4$, on peut aussi utiliser directement la formule de conjugaison des cycles en remarquant que K contient *tous* les produits de deux transpositions à supports disjoints de $\mathfrak{S}_4$: si $\sigma \in \mathfrak{S}_4$, alors $\sigma(12)(34)\sigma^{-1} = (\sigma(1)\sigma(2))(\sigma(3)\sigma(4))$, qui est encore un produit de transpositions à supports disjoints, est encore dans K.

[♫] Felix Klein, 1849 – 1925. Sa contribution aux liens entre groupes et géométrie est déterminante. Lire son *Programme d'Erlangen*.

2.2 Signature d'une permutation

Proposition (existence et unicité de la signature)
Soit n un entier naturel supérieur ou égal à 2.
(i) Il existe un unique homomorphisme de groupes non trivial $\varepsilon : \mathfrak{S}_n \to \{-1, 1\}$.
(ii) Si c est un p-cycle, alors $\varepsilon(c) = (-1)^{p-1}$.
(iii) Soit $\sigma \in \mathfrak{S}_n$. Si m est le nombre d'orbites de $\{1, \dots, n\}$ sous σ, alors $\varepsilon(\sigma) = (-1)^{n-m}$.

Preuve. (i) D'abord, l'unicité. Soit $\epsilon : \mathfrak{S}_n \to \{\pm 1\}$ un homomorphisme de groupes. Puisque le groupe multiplicatif $\{\pm 1\}$ est abélien, deux permutations conjuguées ont la même image par ϵ. Comme les transpositions sont toutes conjuguées, elles ont donc toutes la même image par ϵ. Mais puisque elles engendrent $\mathfrak{S}_n$, si cette image commune est 1, alors ϵ est constant. Ainsi, si ϵ n'est pas trivial, $\epsilon(\tau) = -1$ pour toute transposition τ. A nouveau, comme les transpositions engendrent $\mathfrak{S}_n$, la valeur de ϵ sur les transpositions détermine la valeur de ϵ sur toutes les permutations. Cela démontre l'unicité.
Ensuite, l'existence. Si $\sigma \in \mathfrak{S}_n$, on note

$$\varepsilon(\sigma) = \prod_{(i,j)\in\{1,\dots n\}^2,\ i<j} \frac{\sigma(j) - \sigma(i)}{j - i}. \tag{2}$$

Comme chaque terme de ce produit est symétrique en i et j, le nombre rationnel $\varepsilon(\sigma)$ s'écrit aussi en sommant sur les *paires* d'éléments distincts

$$\varepsilon(\sigma) = \prod_{\{i,j\}\subseteq\{1,\dots n\},\ i\neq j} \frac{\sigma(j) - \sigma(i)}{j - i} = \frac{\displaystyle\prod_{\{i,j\},\ i\neq j} (\sigma(j) - \sigma(i))}{\displaystyle\prod_{\{i,j\},\ i\neq j} (j - i)}.$$

Puisque l'application $\{i, j\} \to \{\sigma(i), \sigma(j)\}$ est une bijection de l'ensemble des parties à deux éléments de $\{1, \dots, n\}$ sur lui-même — sa réciproque s'exhibe aisément —, le numérateur et le dénominateur de cette dernière fraction ont la même valeur absolue. On en déduit que $\varepsilon(\sigma) \in \{-1, 1\}$. Mieux encore, en revenant à la première formulation (2) en termes de couples, on obtient la célèbre formule

$$\varepsilon(\sigma) = (-1)^{I(\sigma)} \tag{3}$$

où $I(\sigma)$ est le *nombre d'inversions de σ* :

$$I(\sigma) = \text{Card}\left\{(i, j) \in \{1, \dots n\}^2,\ i < j \text{ et } \sigma(i) > \sigma(j)\right\}.$$

Il reste à montrer que ε ainsi défini est un homomorphisme de groupes non trivial. Soient $s, t \in \mathfrak{S}_n$. Alors,

$$\varepsilon(st) = \prod_{\{i,j\},\ i\neq j} \frac{s\circ t(j) - s\circ t(j)}{j - i} = \prod_{\{i,j\},\ i\neq j} \frac{s[t(j)] - s[t(i)]}{t(j) - t(i)} \times \prod_{\{i,j\},\ i\neq j} \frac{t(j) - t(i)}{j - i} = \varepsilon(s)\varepsilon(t)$$

la dernière égalité venant encore du fait que $\{i, j\} \mapsto \{t(i), t(j)\}$ est une bijection de l'ensemble des parties à deux éléments de $\{1, \dots, n\}$ sur lui-même. Enfin, il suffit pour conclure de montrer que $\varepsilon(12) = -1$ ce qui est immédiat à partir de (3) puisque $I((12)) = 1$.
(ii) Tous les p-cycles ont la même signature puisqu'ils sont tous conjugués dans $\mathfrak{S}_n$. Or, $\varepsilon(1, 2 \dots, p) = \varepsilon\left((1, 2)(2, 3) \dots (p - 1, p)\right) = \varepsilon\left((1, 2)\right)\varepsilon\left((2, 3)\right) \dots \varepsilon\left((p - 1, p)\right) = (-1)^{p-1}$.
(iii) Soit $\sigma \in \mathfrak{S}_n$. On décompose σ en produit de cycles à supports disjoints, y compris les cycles de longueur 1, triviaux, qui correspondent aux points fixes de σ : ainsi, $\sigma = c_1 \dots c_m$ où m est le nombre d'orbites sous σ et où chaque c_k est un cycle dont on note la longueur ℓ_k — pour tout $k \in \{1, \dots m\}$, $\ell_k \geq 1$. Alors,

$$\varepsilon(\sigma) = \prod_{k=1}^{m} (-1)^{\ell_k - 1} = (-1)^{\sum_{k=1}^m \ell_k} \times (-1)^{-m} = (-1)^{n-m}$$

puisque la somme des longueurs des cycles de σ, triviaux compris, égale la somme des cardinaux des orbites, c'est-à-dire n. ■

Définition (signature, permutations paires ou impaires)
On appelle ε la *signature*. Une permutation est dite *paire* si sa signature est 1, *impaire* sinon.

A noter
La preuve donnée ci-dessus est constructive. Elle exhibe la signature *via* son expression en termes de nombres d'inversions d'une permutation.

En passant Se faire raconter l'histoire du jeu de taquin de Lloyd.

2.3 Le groupe alterné

Définition (groupe alterné)
Si E est un ensemble fini, le *groupe alterné* de E est le sous-groupe de $\mathfrak{S}_E$ formé par ses permutations paires. On le note $\mathfrak{A}_E$ — cette lettre est un A gothique. Le sous-groupe des permutations paires de $\mathfrak{S}_n$ est noté $\mathfrak{A}_n$.

A noter
$\mathfrak{A}_n$ est le noyau de la signature. C'est donc un sous-groupe distingué d'indice 2 de $\mathfrak{S}_n$. En particulier, $|\mathfrak{A}_n| = \frac{n!}{2}$.

Proposition (les 3-cycles engendrent $\mathfrak{A}_n$)
Soit n un entier naturel non nul. Le groupe $\mathfrak{A}_n$ est engendré par ses 3-cycles.

Preuve. Toute permutation paire est un produit d'un nombre pair de transpositions. Or, $(12)(23) = (123)$ et $(12)(34) = (123)(234)$. Ces deux calculs, grâce à la formule de conjugaison des cycles, suffit pour conclure. ■

Corollaire (groupe dérivé de $\mathfrak{S}_n$)
Pour tout $n \geq 2$, $D(\mathfrak{S}_n) = \mathfrak{A}_n$.

Preuve. Si $n = 2$, c'est idiot. On suppose $n \geq 3$. Tout commutateur est une permutation paire ; donc $D(\mathfrak{S}_n) \subseteq \mathfrak{A}_n$. Puisque les 3-cycles engendrent $\mathfrak{A}_n$, il suffit de montrer que tout 3-cycle est un commutateur pour obtenir l'inclusion inverse. Or, $(12)(23)(12)(23) = (132)$. Cela suffit pour conclure. ■

Proposition (les 3-cycles sont conjugués dans $\mathfrak{A}_n$ lorsque $n \geq 5$)
On suppose que $n \geq 5$. Alors, si c et c' sont deux 3-cycles de $\mathfrak{A}_n$, il existe $\sigma \in \mathfrak{A}_n$ tel que $c' = \sigma c \sigma^{-1}$.

Preuve. Soit (abc) un 3-cycle de $\mathfrak{S}_n$. Puisque les 3-cycles sont conjugués dans $\mathfrak{S}_n$, soit $\tau \in \mathfrak{S}_n$ tel que $(abc) = \tau(123)\tau^{-1}$. Si τ est paire, c'est fini, $\sigma = \tau$ convient. Si τ est impaire, alors $\tau(45)$ est paire et $\sigma = \tau(45)$ convient puisque (123) et (45) commutent. ■

A noter
(i) Dans le groupe abélien (et même cyclique) $\mathfrak{A}_3 = \{1, (123), (132)\} \simeq \mathbb{Z}/3\mathbb{Z}$, les deux 3-cycles ne sont pas conjugués.
(ii) Dans $\mathfrak{A}_4$, les huit 3-cycles sont répartis en deux classes de conjugaison qui sont $\{(123), (142), (134), (243)\}$ et $\{(132), (124), (143), (234)\}$ — pour argumenter, par exemple, conjuguer (123) par les éléments du groupe de Klein.

Corollaire (groupe dérivé de $\mathfrak{A}_n$)
Pour tout $n \geq 5$, $D(\mathfrak{A}_n) = \mathfrak{A}_n$.

Preuve. Comme dans le calcul de $D(\mathfrak{S}_n)$, il suffit de montrer que tout 3-cycle est un commutateur dans $\mathfrak{A}_n$. Soit c un 3-cycle. Puisque $n \geq 5$, il est conjugué au 3-cycle c^2 dans $\mathfrak{A}_n$. Soit donc $s \in \mathfrak{A}_n$ tel que $scs^{-1} = c^2$. Alors, $c = scs^{-1}c^{-1}$. ■

Exercice 20 Montrer que $D(\mathfrak{A}_3) = (1)$ et que $D(\mathfrak{A}_4) = K$ (groupe de Klein).

On a la chaîne de sous-groupes distingués $(1) \triangleleft K \triangleleft \mathfrak{A}_4 \triangleleft \mathfrak{S}_4$. Cette situation est exceptionnelle, comme le montre le résultat suivant.

Théorème (simplicité de $\mathfrak{A}_n$ lorsque $n \neq 4$)
Si $n \neq 4$, le groupe alterné $\mathfrak{A}_n$ est simple.

PREUVE. Le groupe $\mathfrak{A}_2$ est trivial et $\mathfrak{A}_3$, cyclique d'ordre 3, est simple.

(i) On prouve d'abord que $\mathfrak{A}_5$ est simple. Soit $H \lhd \mathfrak{A}_5$. Si H contient un 3-cycle, il les contient tous puisque les 3-cycles sont conjugués dans $\mathfrak{A}_5$; donc $H = \mathfrak{A}_5$ puisque les 3-cycles engendrent $\mathfrak{A}_5$. Si H contient un élément d'ordre 2, quitte à renuméroter, H contient $(12)(34)$; en conjuguant par (13245), le groupe H contient $(34)(25)$, donc le produit $(12)(34)(34)(25) = (125)$. Alors $H = \mathfrak{A}_5$ puisqu'il contient un 3-cycle. Enfin, si H contient un élément d'ordre 5, *i.e.*, quitte à renuméroter, s'il contient (12345), on conjugue par (254) ce qui montre que H contient (15324) ; donc H contient le produit $(12345)(15324) = (254)$ et donc $H = \mathfrak{A}_5$. puisqu'il contient un 3-cycle. On a fait le tour des cas possibles, ce qui montre que $\mathfrak{A}_5$ est simple.

(ii) On suppose $n \geq 6$. Soit $H \lhd \mathfrak{A}_n$. On suppose que $H \neq (1)$. Soit alors $\sigma \in H$ et $a \in \{1, \dots, n\}$ tels que $b = \sigma(a) \neq a$. Soit $c \in \{1, \dots, n\} \setminus \{a, b, \sigma(b)\}$. Soit $\tau = (acb)$ et soit ρ le commutateur $\rho = \tau\sigma\tau^{-1}\sigma^{-1}$. D'une part $\rho = (\tau\sigma\tau^{-1})\sigma^{-1} \in H$. D'autre part, par la formule de conjugaison des cycles, $\rho = \tau(\sigma\tau^{-1}\sigma^{-1}) = (acb)(\sigma(a)\sigma(b)\sigma(c))$. Soit $E \subseteq \{1, \dots, n\}$ tel que $\operatorname{Card} E = 5$ et $E \supseteq \{a, b, c, \sigma(a), \sigma(b), \sigma(c)\}$ — un tel E existe car $b = \sigma(a)$. On termine la preuve par les deux points suivants.

① $\rho \neq 1$; en effet, $\rho = 1$ si, et seulement si $(abc) = (\sigma(a)\sigma(b)\sigma(c))$, *i.e.* si, et seulement si $(bca) = (b\sigma(b)\sigma(c))$ ce qui n'est pas puisque $c \neq \sigma(b)$.

② [On veut dire correctement que ρ "appartient à $H \cap \mathfrak{A}_E \lhd \mathfrak{A}_E$". Puisque $\mathfrak{A}_E$ est simple et $\rho \neq 1$, cela impose $H \cap \mathfrak{A}_E = \mathfrak{A}_E$. Donc H contient un 3-cycle. Donc $H = \mathfrak{A}_n$. On le dit correctement dans ce qui suit.] Soit $\pi : \mathfrak{A}_E \to \mathfrak{A}_n$ le prolongement par l'identité hors de E. Alors, $1 \neq (acb)(\sigma(a)\sigma(b)\sigma(c)) \in \pi^{-1}(H) \lhd \mathfrak{A}_E$. Par simplicité de $\mathfrak{A}_E$, cela impose $\pi^{-1}(H) = \mathfrak{A}_E$. Donc H contient le 3-cycle $\pi(abc)$. Donc $H = \mathfrak{A}_n$. ■

Exercice 21 Refaire une preuve du calcul des groupes dérivés de $\mathfrak{S}_n$ et $\mathfrak{A}_n$ en utilisant la simplicité de $\mathfrak{A}_n$.

Proposition (sous-groupes normaux de $\mathfrak{S}_n$)
Si $n \neq 4$, les seuls sous-groupes distingués de $\mathfrak{S}_n$ sont (1), $\mathfrak{A}_n$ et $\mathfrak{S}_n$.

PREUVE. Pour $n \in \{2, 3\}$, c'est immédiat. On suppose $n \geq 5$. Soit $H \lhd \mathfrak{S}_n$. Alors, $H \cap \mathfrak{A}_n \lhd \mathfrak{A}_n$ et donc, puisque $\mathfrak{A}_n$ est simple, $H \cap \mathfrak{A}_n \in \{(1), \mathfrak{A}_n\}$. Si $H \cap \mathfrak{A}_n = \mathfrak{A}_n$, alors $H \supseteq \mathfrak{A}_n$ et donc $H \in \{\mathfrak{A}_n, \mathfrak{S}_n\}$. On suppose que $H \cap \mathfrak{A}_n = (1)$ et que $H \neq (1)$. Alors, la restriction de la signature à H est un isomorphisme et H est d'ordre 2. Soit s l'unique permutation impaire telle que $H = \{1, s\}$. Alors, si $t \in \mathfrak{S}_n$, tst^{-1} est une permutation impaire de H. Donc $tst^{-1} = s$, ce qui montre que s et t commutent. Puisque cela est vrai pour tout $t \in \mathfrak{S}_n$, cela entraîne que s est central. Or, le centre de $\mathfrak{S}_n$ est trivial : nécessairement, $\sigma = 1$ et l'hypothèse $H \neq (1)$ ne tient pas. ■

Exercice 22 Trouver les sous-groupes distingués de $\mathfrak{S}_4$.

3 Groupes abéliens de type fini

Définition (GATF, GALTF et GAF)
Un groupe abélien est dit *de type fini* — en abrégé, *GATF* — lorsqu'il admet une partie finie qui l'engendre. Les groupes abéliens finis — en abrégé, *GAF* — en sont un cas particulier. Lorsqu'un groupe est isomorphe au groupe additif $\mathbb{Z}^r$ pour un $r \in \mathbb{N}$, on dit que c'est un groupe abélien *libre* de type fini — en abrégé, *GALTF*, ou encore un *réseau*.

Exemples
Si r est un entier naturel et G un GAF, alors $\mathbb{Z}^r \times G$ est un GATF. On verra que tous les GATF ont cette forme.

A noter
Tout GATF est isomorphe au quotient d'un GALTF.
En effet, si $G = \langle g_1, \ldots, g_n \rangle$, alors l'application $\mathbb{Z}^n \to G$, $(x_1, \ldots x_n) \mapsto \sum_{k=1}^n x_k g_k$ est un homomorphisme surjectif de groupes auquel il suffit d'appliquer le premier théorème d'isomorphisme. Son noyau est le *sous-groupe des relations* de G.

Exercice 23
(i) Si G est un groupe abélien, l'ensemble des éléments d'ordre fini de G est un sous-groupe de G.
(ii) L'ensemble des éléments d'ordre fini d'un groupe non abélien n'est en général pas un sous-groupe.

Définition (sous-groupe de torsion)
Si G est un groupe abélien, l'ensemble G_T des éléments d'ordre fini de G est le *sous-groupe de torsion* de G.

Exemples
(i) Si G est un GAF, le groupe de torsion de $G \times \mathbb{Z}^r$ est $G_T = G$.
(ii) Si G est un GALTF, il n'a pas de torsion : $G_T = \{0\}$.
(iii) Tous les éléments du groupe additif $\mathbb{Q}/\mathbb{Z}$ sont de torsion (d'ordre fini). Autrement dit, $(\mathbb{Q}/\mathbb{Z})_T = \mathbb{Q}/\mathbb{Z}$.
[On verra, une fois le théorème de structure des GATF installé, qu'un GATF est libre si, et seulement s'il est sans torsion. On voit avec cet exemple que ce résultat tombe en défaut si on ne suppose pas le groupe abélien finiment engendré.]

3.1 Prélude à l'unicité des facteurs invariants

On commence par un lemme d'apparence technique dont on fournit une preuve combinatoire. Il constitue un point crucial dans l'argumentaire choisi pour établir la structure des GAF et des GALTF

Lemme (régularité du produit pour les groupes finis)
Soient G, G' et H des groupes finis. On suppose que $G \times H \simeq G' \times H$. Alors, $G \simeq G'$.

Preuve. On compte. Si L et M sont deux groupes finis, on note $h(L, M)$ le nombre d'homomorphismes de groupes $L \to M$ et $i(L, M)$ le nombre d'homomorphismes injectifs de groupes $L \to M$.

① *Si L, G et H sont des groupes finis, alors $h(L, G \times H) = h(L, G) \times h(L, H)$.*

En effet, si note p_1 et p_2 les projections $p_1 : G \times H \to G$, $(g, h) \mapsto g$ et $p_2 : G \times H \to H$, $(g, h) \mapsto h$; et si on note aussi F et G les applications définies par les formules

$$\begin{array}{ccc} \mathrm{Hom}(L, G \times H) & \longrightarrow & \mathrm{Hom}(L, G) \times \mathrm{Hom}(L, H) \\ f & \overset{F}{\longmapsto} & (p_1 \circ f, p_2 \circ f) \\ \varphi \times \psi & \overset{G}{\longmapsto} & (\varphi, \psi) \end{array}$$

où $\varphi \times \psi$ est défini par $\varphi \times \psi(\ell) = (\varphi(\ell), \psi(\ell))$, alors, F et G sont des bijections réciproques l'une de l'autre.

② *Si L et G sont deux groupes finis, alors $h(L, G) = \sum_{N \lhd L} i\,(L/N, G)$.*

La somme ci-dessus porte sur tous les sous-groupes distingués N de L. En effet, soit

$$\mathcal{I} = \{(N, i),\ N \lhd L,\ i \in \mathrm{Hom}(L/N, G),\ i \text{ injectif}\}$$

La propriété universelle du quotient assure que tout homomorphisme de groupes $f : L \to G$ induit un homomorphisme injectif $\overline{f} : L/\ker(f) \to G$. Alors, l'application Φ définie par

$$\begin{array}{rccc} \Phi : & \mathrm{Hom}(L,G) & \longrightarrow & \mathcal{I} \\ & f & \longmapsto & \left(\ker f, \overline{f}\right) \end{array}$$

est une bijection, dont la réciproque est $(N, g) \mapsto p_N \circ g$ où $p_N : L \to L/N$ est la projection canonique. Donc $h(L,G) = \#\mathcal{I}$. En comptant le cardinal de $\mathcal{I}$ par sa première composante, on obtient la somme souhaitée.

③ *Si L, G et G' sont deux groupes finis tels que $h(L,G) = h(L,G')$, alors $i(L,G) = i(L,G')$.*

On procède par récurrence (forte) sur $|L|$. Si $|L| = 1$, il n'y a rien à démontrer. On suppose que $|L| \geq 2$. Alors, la formule ② fournit $h(L,G) = i(L,G) + \sum_{N \lhd L,\ N \neq (0)} i\left(L/N, G\right)$. Par récurrence, on peut remplacer G par G' dans cette dernière somme, ce qui montre le résultat.

④ Fin de la preuve : on se place dans les hypothèses du lemme. Alors, ① assure que $h(G,G) = h(G,G')$. En appliquant ③ pour $L = G$, on obtient alors que $i(G,G) = i(G,G')$. Puisque $i(G,G) \geq 1$ (l'identité est une injection $G \to G$), on en déduit que $i(G,G') \geq 1$. Mais l'hypothèse $G \times H \simeq G' \times H$ assure que $|G| = |G'|$: puisque ces cardinaux sont finis et égaux, tout homomorphisme injectif $G \to G'$ est un isomorphisme. ■

Exercice 24 Si on enlève l'hypothèse de finitude, la conclusion du lemme tombe en défaut.

Par exemple, si $\mathbb{F}$ est un corps, les $\mathbb{F}$-espaces vectoriels $\mathbb{F}[X]$ et $\mathbb{F} \times \mathbb{F}[X]$ sont isomorphes puisqu'ils ont la même dimension (infinie, dénombrable ; exercice : expliciter un tel isomorphisme). Un tel isomorphisme, en abandonnant la loi externe, est aussi un isomorphisme entre les groupes additifs $\{0\} \times \mathbb{F}[X]$ et $\mathbb{F} \times \mathbb{F}[X]$. Pourtant, le groupe additif $\mathbb{F}$ n'est pas trivial.

3.2 GALTF, rang

Proposition (les GALTF ont un rang)
Soient $n, m \geq 1$. Les groupes $\mathbb{Z}^m$ et $\mathbb{Z}^n$ sont isomorphes si, et seulement si $m = n$.

Preuve. Soit $f : \mathbb{Z}^m \to \mathbb{Z}^n$ un isomorphisme de groupes. Une récurrence immédiate montre que f est $\mathbb{Z}$-linéaire. On prolonge f à $\mathbb{Q}^m$ par la formule $f\left(\frac{1}{D}\left(x_1 \ldots, x_m\right)\right) = \frac{1}{D} f\left(x_1 \ldots, x_m\right)$ où $(x_1 \ldots, x_m) \in \mathbb{Z}^m$ et $D \in \mathbb{Z} \setminus \{0\}$. D'une part, cette formule a du sens puisque, avec ces notations, $\frac{1}{D}\left(x_1 \ldots, x_m\right) = \frac{1}{D'}\left(x'_1 \ldots, x'_m\right)$ implique $\frac{1}{D} f\left(x_1 \ldots, x_m\right) = \frac{1}{D'} f\left(x'_1 \ldots, x'_m\right)$ — en effet, si la prémice est vérifiée, alors la $\mathbb{Z}$-linéarité de f assure que $D' f\left(x_1 \ldots, x_m\right) = f\left(D' x_1 \ldots, D' x_m\right) = f\left(D x'_1 \ldots, D x'_m\right) = D f\left(x'_1 \ldots, x'_m\right)$. D'autre part, ladite formule définit l'image de n'importe quel élément de $\mathbb{Q}^m$ par réduction au même dénominateur (D) des coordonnées. Une fois ce prolongement $f : \mathbb{Q}^m \to \mathbb{Q}^n$ défini, sa $\mathbb{Q}$-linéarité et sa bijectivité sont immédiates — on exhibe sa réciproque qui a la même forme. Alors, les $\mathbb{Q}$-espaces vectoriels $\mathbb{Q}^m$ et $\mathbb{Q}^n$ étant isomorphes, ils ont la même dimension, ce qui montre que $m = n$. ■

Exercice 25
Faire une autre preuve de cette proposition en prenant un nombre premier p (par exemple 2) et en construisant un isomorphisme de $\mathbb{Z}/p\mathbb{Z}$-espaces vectoriels à partir d'un isomorphisme de groupes $\mathbb{Z}^m \to \mathbb{Z}^n$ — pour conclure, on argumentera à l'aide la dimension des espaces vectoriels construits.

Définition (rang d'un GALTF)
Si G est un GALTF isomorphe à $\mathbb{Z}^r$, le nombre r est appelé *rang* de G — cette définition est rendue possible par la proposition précédente.

Exercice 26
En passant par le corps des fractions, montrer que si $\mathcal{A}$ est un anneau intègre et si les anneaux $\mathcal{A}^m$ et $\mathcal{A}^n$ sont isomorphes, alors $m = n$.

Définition (base d'un GALTF)
Si G est un GALTF de rang r, une *base de G* est un r-uplet $(v_1, \ldots, v_r)$ d'éléments de G qui engendrent G et qui sont *$\mathbb{Z}$-linéairement indépendants*, ce qui signifie que pour tout $x_1, \ldots x_s \in \mathbb{Z}$, $x_1 v_1 + \cdots + x_s v_s = 0 \Longrightarrow x_1 = \cdots = x_s = 0$.

Exercice 27
Si G est un GALTF, un famille $(v_1, \ldots, v_r)$ d'éléments de G est une base de G si, et seulement si tout élément de G s'écrit de manière unique sous la forme $\sum_{k=1}^{r} x_k v_k$ où $x_1, \ldots, x_r \in \mathbb{Z}$.

Définition (formes coordonnées dans une base d'un GALTF)
Avec les notations de la définition d'une base d'un GALTF, pour tout $k \in \{1, \dots, r\}$, la k^{e} *forme coordonnée relative à la base* $(v_1, \dots, v_r)$ est l'homomorphisme de groupes $v_k^* : G \to \mathbb{Z}$ défini par $v_k^*(\gamma) = x_k$, pour tout $\gamma = \sum_{k=1}^r x_k v_k \in \Gamma$.

A noter
Comme dans les affaires de dualité dans les espaces vectoriels de dimension finie, avec les notations précédentes, les formes coordonnées sont définies par les relations $v_i^*(v_j) = \delta_{i,j}$ pour tous i, j — notation de Kronecker.

Définition (somme directe de sous-groupes abéliens)
Soient G un groupe abélien noté additivement, H et K des sous-groupes de G. On dit que G est *somme directe de H et K* lorsque tout élément de G s'écrit, de manière unique, comme la somme d'un élément de H et d'un élément de K. On note alors $G = H \oplus K$ (comme une somme directe de sous-espaces vectoriels).

Exercice 28
Dans les conditions de la définition ci-dessus, $G = H \oplus K$ si, et seulement si $H \cup K$ engendre G et $H \cap K = (0)$.

Théorème (théorème de la base adaptée)
Soient Γ un GALTF de rang r et G un sous-groupe de Γ. Alors :

(i) G est aussi un GALTF, de rang s inférieur ou égal à r

(ii) Il existe une base $(v_1, \dots, v_r)$ de Γ et des entiers naturels $a_1, \dots a_s$ non nuls tels que $a_1 | a_2 | \dots | a_s$ et tels que $(a_1 v_1, \dots, a_s v_s)$ soit une base de G

(iii) (unicité) si $a_1, \dots a_s$ et $b_1, \dots b_s$ sont des suites d'entiers naturels et si $(v_1, \dots, v_r)$ et $(w_1, \dots, w_r)$ sont des bases de Γ telles que $a_1 | a_2 | \dots | a_s$, $b_1 | b_2 | \dots | b_s$, $(a_1 v_1, \dots, a_s v_s)$ et $(b_1 w_1, \dots, b_s w_s)$ sont des bases de G, alors $a_k = b_k$ pour tout $k \in \{1, \dots, s\}$.

(iv) Γ / G est isomorphe à $\mathbb{Z}^{r-s} \times (\mathbb{Z}/a_1\mathbb{Z}) \times (\mathbb{Z}/a_2\mathbb{Z}) \cdots \times (\mathbb{Z}/a_s\mathbb{Z})$

(v) En particulier, $[\Gamma : G]$ est fini si, et seulement si $r = s$. Dans ces conditions, $[\Gamma : G] = a_1 a_2 \dots a_r$.

Preuve. Soit $(e_1, \dots, e_r)$ une base de Γ. On note $(e_1^*, \dots, e_r^*)$ les formes coordonnées relatives à cette base.
(i) On procède par récurrence sur $r = \operatorname{rg} \Gamma$. L'hypothèse de récurrence au rang r est la suivante : *si Γ est un GALTF de rang r et si G est un sous-groupe de Γ, alors G est un GALTF de rang inférieur ou égal à r.*
Si $r = 1$, on peut supposer que $\Gamma = \mathbb{Z}$ puisqu'il lui est isomorphe. Or, les sous-groupes de $\mathbb{Z}$ sont tous de la forme $a\mathbb{Z}$ où $a \in \mathbb{N}$. Si $a = 0$, $a\mathbb{Z} = \{0\}$ est un GALTF de rang 0 ; si $a \geq 1$, $a\mathbb{Z}$ est un GALTF de rang 1.
On suppose que $r \geq 2$ et on note $G_{r-1} = G \cap \bigoplus_{k=1}^{r-1} \mathbb{Z} e_k$. Par hypothèse de récurrence, puisque G_{r-1} est un sous-groupe du GALTF $\bigoplus_{k=1}^{r-1} \mathbb{Z} e_k$, c'est un GALTF de rang inférieur ou égal à $r-1$. Soit $e_r^* \in \operatorname{Hom}(\Gamma, \mathbb{Z})$ la r^{e} forme coordonnée relative à la base $(e_1, \dots, e_r)$. Puisque $e_r^*(G)$ est un sous-groupe de $\mathbb{Z}$, soit $a \in \mathbb{N}$ tel que $e_r^*(G) = a\mathbb{Z}$. Si $a = 0$, alors $G = G_{r-1}$ est libre, de rang inférieur ou égal à $r-1$. Si $a \neq 0$, soit $w \in G$ tel que $e_r^*(w) = a$. Pour tout $g \in G$, $e_r^*(g) \in a\mathbb{Z}$, ce qui entraîne qu'il existe $c \in \mathbb{Z}$ tel que $g - cw \in G \cap \ker(e_r^*) = G_{r-1}$. Cela montre que $G = G_{r-1} \oplus \mathbb{Z} w$, le fait que l'intersection $G_{r-1} \cap \mathbb{Z} w$ soit nulle étant immédiat. Or, G_{r-1} est un GALTF de rang inférieur ou égal à $r-1$; donc G est un GALTF de rang inférieur ou égal à r. On a montré (i).
(ii) On procède par récurrence sur $r = \operatorname{rg} \Gamma$. L'hypothèse de récurrence au rang r est la suivante : *si Γ est un GALTF de rang r et si G est un sous-GALTF de rang s, alors il existe une base $(v_1, \dots, v_r)$ de Γ et des entiers naturels non nuls $a_1, \dots a_s$ tels que $a_1 | a_2 | \dots | a_s$ et tels que $(a_1 v_1, \dots, a_s v_s)$ soit une base de G.*
Si $r = 1$, on peut supposer que $\Gamma = \mathbb{Z}$ puisqu'il lui est isomorphe. Comme G est alors un sous-groupe de $\mathbb{Z}$, soit $a \in \mathbb{N}$ tel que $G = a\mathbb{Z}$. Si $a = 0$, alors $G = \{0\}$ est libre de rang 0. Si $a \neq 0$, alors G est libre de rang 1 ; en outre, $v_1 = 1$ et $a_1 = a$ conviennent.
On suppose que $r \geq 2$. Si $G = \{0\}$, il n'y a rien à démontrer. On suppose que $G \neq \{0\}$. Soit donc $g \in G$, $g \neq 0$. Alors, une au moins des coordonnées de g dans la base $(e_1, \dots, e_r)$ est non nulle : il existe $k \in \{1, \dots, r\}$ tel que $e_k^*(g) \neq 0$. Ainsi, $\{f(G),\ f \in \operatorname{Hom}(\Gamma, \mathbb{Z})\} \neq \{\{0\}\}$. Soit alors a_1 l'entier naturel, non nul, défini par

$$a_1 = \min \{a \in \mathbb{N}^*,\ \exists f \in \operatorname{Hom}(\Gamma, \mathbb{Z}),\, f(G) = a\mathbb{Z}\}.$$

Soient alors $f_1 \in \operatorname{Hom}(\Gamma, \mathbb{Z})$ tel que $f_1(G) = a_1\mathbb{Z}$, et $w_1 \in G$ tel que $f_1(w_1) = a_1$. On montre alors l'assertion suivante :

$$\forall f \in \operatorname{Hom}(\Gamma, \mathbb{Z}),\ a_1 \text{ divise } f(w_1). \tag{4}$$

En effet, si $f \in \mathrm{Hom}(\Gamma, \mathbb{Z})$, on note $d = PGCD(a_1, f(w_1)) = PGCD(f_1(w_1), f(w_1))$. Soit alors $b, c \in \mathbb{Z}$ tels que $d = ba_1 + cf(w_1)$ — c'est une relation de Bézout. Dans ces conditions, la forme linéaire $bf_1 + cf \in \mathrm{Hom}(\Gamma, \mathbb{Z})$ vérifie $d = (bf_1 + cf)(w_1)$. Par minimalité de a_1, puisque $w_1 \in G$, cela montre que $a_1 \leq d$. Comme d est un diviseur de a_1, cela entraîne que $a_1 = d$ ce qui implique que a_1 est un diviseur de $f(w_1)$: on a montré (4).
On applique alors (4) aux formes coordonnées relatives à la base $(e_1, \dots, e_r)$ du GALTF Γ. il s'ensuit que toutes les coordonnées de w_1 sont divisibles par a_1. Soit alors $v_1 \in \Gamma$ tel que $w_1 = a_1 v_1$. En particulier, puisque $a_1 \neq 0$ et $f_1(w_1) = a_1 = a_1 f_1(v_1)$, cela entraîne que $f_1(v_1) = 1$.
On est alors dans la situation suivante :
(a) $\Gamma = \ker(f_1) \oplus \mathbb{Z}v_1$
(b) $G = (G \cap \ker(f_1)) \oplus a_1\mathbb{Z}v_1$
L'assertion (a) est garantie par la formule $\gamma = (\gamma - f_1(\gamma)v_1) + f_1(\gamma)v_1$ pour tout $\gamma \in \Gamma$, puisque $\gamma - f_1(\gamma)v_1 \in \ker(f_1)$, le fait que l'intersection $\ker(f_1) \cap \mathbb{Z}v_1$ soit nulle étant immédiat. L'assertion (b) est du même acabit en écrivant $g = (g - f_1(g)v_1) + f_1(g)v_1$ pour tout $g \in G$ et en remarquant que $f_1(g)$ est un multiple de a_1, puisque $f_1(G) = a_1\mathbb{Z}$ par définition de a_1.
Le groupe $\ker(f_1)$ est un sous-groupe du GALTF Γ. D'après (i), c'est donc lui-même un GALTF, de rang inférieur ou égal à r. Mais (a) impose que ce rang soit exactement $r-1$. En outre, $G \cap \ker(f_1)$ est un sous-groupe du GALTF $\ker(f_1)$. Toujours d'après (i), $G \cap \ker(f_1)$ est encore un GALTF ; soit $s \in \{1, \dots, r\}$ tel que $s - 1 = \mathrm{rg}\, G \cap \ker(f_1)$. Par hypothèse de récurrence, soient $a_2, \dots, a_s \in \mathbb{N}^*$ et $(v_2, \dots, v_r)$ une base de $\ker(f_1)$ tels que $a_2 | \dots | a_s$ et tels que $(a_2 v_2, \dots, a_s v_s)$ soit une base de $G \cap \ker(f_1)$. Alors, (a) et (b) assurent que $(v_1, \dots, v_r)$ est une base de Γ et que $(a_1 v_1, \dots, a_s v_s)$ est une base de G. Il reste à montrer que $a_1 | a_2$. Pour cela, soit $f = v_1^* + v_2^* \in \mathrm{Hom}(\Gamma, \mathbb{Z})$, somme des formes coordonnées v_1^* et v_2^* relatives à la base $(v_1, \dots, v_r)$ de Γ. D'une part, $f(a_1 v_1) = a_1$, ce qui entraîne par minimalité de a_1 que $f(G) = a_1\mathbb{Z}$. D'autre part, $f(a_2 v_2) = a_2$ ce qui implique que $a_2 \in a_1\mathbb{Z}$ puisque $a_2 v_2 \in G$. On a montré (ii).
(iv) et (v) En reprenant les notations du théorème, $\Gamma = \bigoplus_{k=1}^r \mathbb{Z}v_k$ et $G = \bigoplus_{k=1}^s \mathbb{Z}a_k v_k$. Alors, l'homomorphisme de groupes

$$\begin{array}{ccc} \Gamma & \longrightarrow & \mathbb{Z}/a_1\mathbb{Z} \times \cdots \times \mathbb{Z}/a_s\mathbb{Z} \times (\mathbb{Z}^{r-s}) \\ \sum_{k=1}^r x_k v_k & \longmapsto & (x_1 + a_1\mathbb{Z}, \dots, x_s + a_s\mathbb{Z}, (x_{s+1} \dots, x_r)) \end{array}$$

est surjectif et a pour noyau G, ce qui montre (iv) en appliquant le premier théorème d'isomorphisme. En particulier, Γ/G est fini si, et seulement si $r - s = 0$. Dans ce cas, Γ/G est isomorphe au groupe produit $\mathbb{Z}/a_1\mathbb{Z} \times \cdots \times \mathbb{Z}/a_r$ qui est d'ordre $a_1 \dots a_r$.
(iii) Dans la situation du (iv), s est le rang de G et le groupe de torsion de Γ/G est isomorphe au produit $\mathbb{Z}/a_1\mathbb{Z} \times \cdots \times \mathbb{Z}/a_s\mathbb{Z}$. Il suffit donc de montrer que si $a_1 | \dots | a_s$ et si $b_1 | \dots | b_s$, alors $\mathbb{Z}/a_1\mathbb{Z} \times \cdots \times \mathbb{Z}/a_s\mathbb{Z}$ et $\mathbb{Z}/b_1\mathbb{Z} \times \cdots \times \mathbb{Z}/b_s\mathbb{Z}$ sont isomorphes seulement si $a_k = b_k$ pour tout k, ce que l'on montre par récurrence sur s. Si $s = 1$, il n'y a rien à démontrer : deux groupes cycliques sont isomorphes si, et seulement s'ils ont le même ordre. On suppose donc que que $s \geq 2$ et que $\mathbb{Z}/a_1\mathbb{Z} \times \cdots \times \mathbb{Z}/a_s\mathbb{Z}$ et $\mathbb{Z}/b_1\mathbb{Z} \times \cdots \times \mathbb{Z}/b_s\mathbb{Z}$ sont isomorphes. Dans cette situation, a_s est l'ordre maximum d'un élément de $\mathbb{Z}/a_1\mathbb{Z} \times \cdots \times \mathbb{Z}/a_s\mathbb{Z}$. Donc $a_s = b_s$. Le lemme de régularité du produit pour les groupes abéliens assure alors que les groupes $\mathbb{Z}/a_1\mathbb{Z} \times \cdots \times \mathbb{Z}/a_{s-1}\mathbb{Z}$ et $\mathbb{Z}/b_1\mathbb{Z} \times \cdots \times \mathbb{Z}/b_{s-1}\mathbb{Z}$ sont isomorphes, les hypothèse de divisibilité sur les a_k et les b_k demeurant. On conclut par récurrence que $a_k = b_k$, pour tout k. ■

Pour aller plus loin

(i) Cette preuve s'appuie sur la principalité de l'anneau $\mathbb{Z}$. Le résultat du théorème de la base adaptée s'étend au cas des sous-modules d'un module libre sur un anneau principal — un *module* a les mêmes axiomes que ceux d'un espace vectoriel, hormis l'anneau des scalaires dont on ne suppose plus que c'est un corps. En particulier, si $\mathbb{F}$ est un corps, le théorème de la base adaptée dans le cadre de l'anneau principal $\mathbb{F}[X]$ et des polynômes d'endomorphismes est un outil parfait pour l'étude de la réduction des endomorphismes. Y trouve une réponse complète la question des classes de similitude des endomorphismes — ou des matrices carrées.

(ii) En utilisant la division euclidienne dans $\mathbb{Z}$, on peut aussi adopter un point de vue algorithmique — c'est une adaptation de l'algorithme du pivot de Gauss – qui fournit à la fois une autre preuve du puissant théorème de la base adaptée, mais aussi un mode de calcul effectif.

3.3 GAF et GATF, rang et facteurs invariants

Théorème (structure des GATF)
Soit G un GATF. Il existe un unique couple d'entiers naturels (r, s) et une unique suite $a_1, \dots, a_s$ d'entiers naturels tels que
(i) $a_1 \geq 2$ *;*
(ii) $a_1 | a_2 | \dots | a_s$ *;*
(iii) $G \simeq \mathbb{Z}^r \times (\mathbb{Z}/a_1\mathbb{Z}) \times (\mathbb{Z}/a_2\mathbb{Z}) \cdots \times (\mathbb{Z}/a_s\mathbb{Z})$.

PREUVE. Puisque G est un GATF, il admet un système générateur fini $\langle g_1 \dots, g_n \rangle$. Alors, l'application $(x_1, \dots x_n) \mapsto \sum_{k=1}^n x_k g_k$ est un homomorphisme surjectif de groupes $f : \mathbb{Z}^n \to G$. *Via* le premier théorème d'isomorphisme, il induit un isomorphisme entre G et le groupe $\mathbb{Z}^n / \ker f$, qui est le quotient d'un GALTF par un de ses sous-groupes. On conclut avec le théorème de la base adaptée qui fournit à la fois l'existence et l'unicité. ■

Définition (rang et facteurs invariants d'un GATF)
Dans la situation du théorème de structure des GATF, l'entier r est ***le*** *rang* de G et les nombres $a_1, \dots a_s$ sont ***les*** *facteurs invariants* de G.

A noter
(i) Un GATF est un GAF si, et seulement s'il est de rang 0, ou encore si tous ses éléments sont d'ordres finis.

(ii) Avec les notations du théorème de structure des GATF, le sous-groupe de torsion de G est isomorphe à $\mathbb{Z}/a_1\mathbb{Z} \times \cdots \times \mathbb{Z}/a_s\mathbb{Z}$.

(iii) Un GATF est un GALTF si, et seulement s'il n'a aucun facteur invariant ; autrement dit, lorsque $s = 0$, ou encore lorsque son sous-groupe de torsion est nul.

(iv) Ainsi deux GATF sont isomorphes si, et seulement s'ils ont le même rang et les mêmes facteurs invariants. Dans la même veine, deux GAF sont isomorphes si, et seulement s'ils ont les mêmes facteurs invariants.

Exemple
Les facteurs invariants du groupe abélien fini $\mathbb{Z}/60\mathbb{Z} \times \mathbb{Z}/90\mathbb{Z} \times \mathbb{Z}/150\mathbb{Z}$ sont $(30, 30, 900)$. En effet, cette suite est croissante pour l'ordre de divisibilité, et le théorème chinois montre successivement, sachant que $60 = 2^2.3.5$, $90 = 2.3^3.5$ et $150 = 2.3.5^2$, en détricotant les facteurs puis en les retricotant, que

$$\begin{aligned}
\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/45\mathbb{Z} \times \mathbb{Z}/60\mathbb{Z} \times \mathbb{Z}/150\mathbb{Z} & \\
&\simeq (\mathbb{Z}/2\mathbb{Z}) \times \left(\mathbb{Z}/3^2\mathbb{Z} \times \mathbb{Z}/5\mathbb{Z}\right) \times \left(\mathbb{Z}/2^2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/5\mathbb{Z}\right) \times \left(\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/5^2\mathbb{Z}\right) \\
&\simeq \left(\mathbb{Z}/2^2\mathbb{Z} \times \mathbb{Z}/3^2\mathbb{Z} \times \mathbb{Z}/5^2\mathbb{Z}\right) \times (\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/5\mathbb{Z}) \times (\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/5\mathbb{Z}) \\
&\simeq \mathbb{Z}/30\mathbb{Z} \times \mathbb{Z}/30\mathbb{Z} \times \mathbb{Z}/900\mathbb{Z}.
\end{aligned}$$

Exercice 29
Si a et b sont des entiers naturels non nuls, les facteurs invariants de $\mathbb{Z}/a\mathbb{Z} \times \mathbb{Z}/b\mathbb{Z}$ sont $\text{pgcd}(a, b)$ et $\text{ppcm}(a, b)$.

Définition (composante de p-torsion d'un groupe abélien)
Soient G un groupe abélien et p un nombre premier. La *composante de p-torsion* de G est son sous-groupe $G(p) := \left\{x \in G,\ \exists a \geq 0,\ x^{p^a} = 1\right\}$. Autrement dit, $G(p)$ est l'ensemble des éléments de G dont l'ordre est une puissance de p.

Exercice 30 Soient G un groupe abélien et p un nombre premier.
(i) $G(p)$ est un sous-groupe de G.

(ii) Si le groupe $G(p)$ est fini, son ordre est une puissance de p.

Théorème (décomposition des GAF en composantes primaires)
Soit G un GAF.
(i) G est somme directe de ses composantes de p-torsion :

$$G = \bigoplus_{\substack{p \text{ premier} \\ p \text{ divise } |G|}} G(p).$$

(ii) Si l'ordre de G est la puissance d'un nombre premier p, il existe une unique suite finie croissante d'entiers naturels non nuls $a_1 \leq \cdots \leq a_s$ telle que G soit isomorphe au produit

$$G \simeq \mathbb{Z}/p^{a_1}\mathbb{Z} \times \cdots \times \mathbb{Z}/p^{a_r}\mathbb{Z}.$$

(iii) La décomposition de G sous la forme

$$G \simeq \left(\mathbb{Z}/p_1^{a_{(p_1,1)}}\mathbb{Z} \times \cdots \times \mathbb{Z}/p_1^{a_{(p_1,s_{p_1})}}\mathbb{Z}\right) \times \cdots \times \left(\mathbb{Z}/p_m^{a_{(p_m,1)}}\mathbb{Z} \times \cdots \times \mathbb{Z}/p_m^{a_{(p_m,s_{p_m})}}\mathbb{Z}\right).$$

où $p_1 \leq \cdots \leq p_m$ est une suite finie croissante de nombres premiers et où $a_{(p_k,1)} \leq \cdots \leq a_{(p_k,s_{p_k})}$ est une suite finie croissante d'entiers naturels non nuls pour chaque $k \in \{1, \ldots, m\}$, dont l'existence est garantie par (i) et (ii), est unique.

PREUVE. Tout est conséquence directe du théorème de décomposition des GAF en facteurs invariants, et d'applications répétées du théorème chinois. ■

Définition (composantes primaires d'un GAF)
La décomposition de G selon le (iii) du théorème précédent est la *décomposition de G en composantes primaires.*

A noter

(i) La décomposition en composantes primaires est caractérisée par la donnée de la famille presque nulle d'entiers $\left(\left(a_{(p,k)}\right)_{1\leq k\leq s_p}\right)_{p \text{ premier}}$ avec les conditions de monotonie sur les $\left(a_{(p,k)}\right)_{1\leq k\leq s_p}$ énoncées dans le théorème.

(ii) Ainsi, deux GAF sont isomorphes si, et seulement s'ils ont la même décomposition en composantes primaires.

Exemple
On reprend l'exemple ci-dessus : $G = \mathbb{Z}/60\mathbb{Z} \times \mathbb{Z}/90\mathbb{Z} \times \mathbb{Z}/150\mathbb{Z}$. Comme dans celui-ci, on décompose chaque facteur en composantes primaires à l'aide du théorème chinois, puis on recompose nombre premier par nombre premier pour obtenir la décomposition en composantes primaires, qui s'écrit ici (troisième ligne)

$$\begin{aligned}\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/45\mathbb{Z} \times \mathbb{Z}/60\mathbb{Z} \times \mathbb{Z}/150\mathbb{Z} &\\ &\simeq (\mathbb{Z}/2\mathbb{Z}) \times \left(\mathbb{Z}/3^2\mathbb{Z} \times \mathbb{Z}/5\mathbb{Z}\right) \times \left(\mathbb{Z}/2^2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/5\mathbb{Z}\right) \times \left(\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/5^2\mathbb{Z}\right)\\ &\simeq \left(\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2^2\mathbb{Z}\right) \times \left(\mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/3^2\mathbb{Z}\right) \times \left(\mathbb{Z}/5\mathbb{Z} \times \mathbb{Z}/5\mathbb{Z} \times \mathbb{Z}/5^2\mathbb{Z}\right).\end{aligned}$$

Ainsi, le sous-groupe de p-torsion de G est isomorphe à $(\mathbb{Z}/p\mathbb{Z})^2 \times \mathbb{Z}/p^2\mathbb{Z}$ pour $p \in \{2, 3, 5\}$, nul pour tous les autres nombres premiers.

Exercice 31
Déduire des théorèmes de structure des GAF le résultat suivant : *soit G un groupe abélien fini dont l'ordre est un multiple d'un nombre premier p. Alors, G contient un élément d'ordre p.*

[Faire également une preuve directe de ce résultat, par récurrence sur l'ordre de G, en utilisant un groupe-quotient $G/\langle x\rangle$ pour un x non nul de G.]

Exercice 32
En utilisant le théorème de Bézout, faire une preuve directe du fait que *tout GAF est somme directe de ses sous-groupes de p-torsion*, par récurrence sur le nombre de facteurs premiers distincts de l'ordre du groupe.

Exercice 33

(i) Soient G un groupe abélien, L un GALTF et $f : G \to L$ un homomorphisme surjectif de groupes. Montrer qu'il existe un sous-groupe abélien libre de type fini H de G tel que $G = H \oplus \ker(f)$.

(ii) Soient G un GATF et G_T son sous-groupe de torsion. Montrer qu'il existe un sous-groupe L de G qui soit un GALTF et tel que $G = L \oplus G_T$.

4 Groupes linéaires

4.1 Petit mémento sur le déterminant

Définition (polynôme déterminant, déterminant d'une matrice carrée)
Soit n un entier naturel non nul. Le *déterminant à n^2 indéterminées* est le polynôme

$$\det = \det\left(X_{1,1}, X_{1,2}, \ldots, X_{n,n}\right) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{k=1}^{n} X_{k,\sigma(k)} \in \mathbb{Z}\left[X_{1,1}, X_{1,2}, \ldots, X_{n,n}\right].$$

On note souvent les n^2 indéterminées sous forme matricielle, si bien que cette formule de définition devient

$$\det M = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{k=1}^{n} X_{k,\sigma(k)}$$

où la *matrice $n \times n$ générique* $M = M\left(X_{1,1}, X_{1,2}, \ldots, X_{n,n}\right)$, dont les coefficients sont des indéterminées, est

$$M = M\left(X_{1,1}, X_{1,2}, \ldots, X_{n,n}\right) = \begin{pmatrix} X_{1,1} & X_{1,2} & \ldots & X_{1,n} \\ X_{2,1} & X_{2,2} & \ldots & X_{2,n} \\ \vdots & \vdots & & \vdots \\ X_{n,1} & X_{n,2} & \ldots & X_{n,n} \end{pmatrix}.$$

Si A est une matrice carrée à coefficients dans n'importe quel anneau commutatif $\mathcal{A}$, le *déterminant de A*, noté $\det(A)$, est l'élément de $\mathcal{A}$ obtenu en spécialisant le déterminant générique $\det M$ en les coefficients de A.

Exemples
Pour $n = 1$, $\det(X) = 1$.

Pour $n = 2$, $\det\left(X_{1,1}, X_{1,2}, X_{2,1}, X_{2,2}\right) = \det \begin{pmatrix} X_{1,1} & X_{1,2} \\ X_{2,1} & X_{2,2} \end{pmatrix} = X_{1,1}X_{2,2} - X_{1,2}X_{2,1}$.

Pour $n = 3$. D'abord id, puis les deux 3-cycles, puis les trois transpositions :

$$\begin{aligned} \det \begin{pmatrix} X_{1,1} & X_{1,2} & X_{1,3} \\ X_{2,1} & X_{2,2} & X_{2,3} \\ X_{3,1} & X_{3,2} & X_{3,3} \end{pmatrix} &= X_{1,1}X_{2,2}X_{3,3} + X_{1,2}X_{2,3}X_{3,1} + X_{1,3}X_{2,1}X_{3,2} \\ &\quad - X_{1,1}X_{2,3}X_{3,2} - X_{1,3}X_{2,2}X_{3,1} - X_{1,2}X_{2,1}X_{3,3} \end{aligned}$$

A noter
(i) En spécialisant, la formule de définition du déterminant, on obtient que $\det I_n = 1$.
(ii) Le polynôme déterminant à n^2 indéterminées est homogène de degré n et est composé de $n!$ monômes sans carrés, précédés de ± 1.
(iii) La sommation peut se faire en faisant agir les permutations sur le premier indice puisque $\sigma \mapsto \sigma^{-1}$ est une bijection de $\mathfrak{S}_n$ sur lui même qui préserve la signature. Ainsi, on a aussi

$$\det \begin{pmatrix} X_{1,1} & X_{1,2} & \ldots & X_{1,n} \\ X_{2,1} & X_{2,2} & \ldots & X_{2,n} \\ \vdots & \vdots & & \vdots \\ X_{n,1} & X_{n,2} & \ldots & X_{n,n} \end{pmatrix} = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{k=1}^{n} X_{\sigma(k),k}.$$

Cela montre en passant que la matrice générique et sa transposée ont le même déterminant.
(iv) En spécialisant le déterminant à n^2 indéterminées de façon *ad hoc*, on obtient la formule

$$\det \begin{pmatrix} X_{1,1} & \ldots & X_{1,n-1} & 0 \\ \vdots & & \vdots & \vdots \\ X_{n-1,1} & \ldots & X_{n-1,n-1} & 0 \\ 0 & \ldots & 0 & 1 \end{pmatrix} = \det \begin{pmatrix} X_{1,1} & \ldots & X_{1,n-1} \\ \vdots & & \vdots \\ X_{n-1,1} & \ldots & X_{n-1,n-1} \end{pmatrix} \tag{5}$$

En effet, le déterminant spécialisé du membre de gauche s'écrit $\sum_{\sigma\in\mathfrak{S}_n,\sigma(n)=n}\varepsilon(\sigma)\prod_{k=1}^{n-1}X_{k,\sigma(k)}$ alors que le prolongement à $\{1,\dots n\}$ par $\sigma(n)=n$ induit une bijection qui préserve la signature entre $\mathfrak{S}_{n-1}$ et l'ensemble des permutations de $\{1,\dots n\}$ qui fixent n. La préservation de la signature est immédiate grâce au théorème de décomposition des permutations en produits de cycles à supports disjoints.

Proposition (irréductibilité du déterminant)
Pour tout $n\geq 1$, le déterminant à n^2 indéterminées est irréductible dans $\mathbb{Z}[X_{1,1},X_{1,2},\dots,X_{n,n}]$.

PREUVE. Les inversibles de $\mathbb{Z}[X_{1,1},X_{1,2},\dots,X_{n,n}]$ sont ±1, polynômes de degré 0. On suppose que $\det = PQ$ où $P,Q\in\mathbb{Z}[X_{1,1},X_{1,2},\dots,X_{n,n}]$. Il s'agit de montrer que $P=\pm1$ ou $Q=\pm1$. On utilise plusieurs fois le fait que si un produit de deux polynômes est homogène, alors les deux facteurs sont également homogènes. On commence par isoler les indéterminées de la première ligne $X_{1,1},\dots X_{1,n}$. Comme det est homogène de degré 1 en ces n indéterminées, P et Q sont également homogènes en $X_{1,1},\dots X_{1,n}$; comme la somme des degrés de P et Q égale 1, cela force l'un des deux, disons Q, à être de degré 0 en $X_{1,1},\dots X_{1,n}$ — autrement dit, les indéterminées $X_{1,1},\dots X_{1,n}$ n'apparaissent pas dans l'écriture de Q. Pour chaque k, on isole alors les indéterminées de la colonne k, autrement dit $X_{1,k},X_{2,k},\dots X_{n,k}$. Là encore, det est homogène en $X_{1,k},X_{2,k},\dots X_{n,k}$, de degré 1, ce qui impose que P et Q soient également homogènes en ces indéterminées, le degré de P ou de Q étant nul. Comme le degré de P en $X_{1,k}$ égale 1, c'est Q qui est de degré 0. Cela étant vrai pour tout $k\in\{1,\dots,n\}$, on a montré que le degré de Q est nul en *toutes* les indéterminées $X_{1,1},\dots,X_{n,n}$. Cela signifie que Q est un polynôme constant. Le coefficient de det en le monôme $X_{1,1}X_{2,2}\dots X_{n,n}$ valant 1, cela impose que $Q=\pm1$. ■

On note $C_1,\dots C_n$ les colonnes de la matrice générique à n^2 indéterminées, et $L_1,\dots,L_n$ ses lignes. On note alors aussi

$$\det=\det(X_{1,1},X_{1,2},\dots,X_{n,n})=\det(C_1,C_2,\dots,C_n)=\det(L_1,L_2,\dots,L_n).$$

Proposition (le déterminant est n-linéaire alterné en ses lignes et ses colonnes)
Soit $n\geq 1$.
(i) Pour toute permutation $\tau\in\mathfrak{S}_n$, $\det\left(C_{\tau(1)},C_{\tau(2)},\dots,C_{\tau(n)}\right)=\varepsilon(\tau)\det(C_1,C_2,\dots,C_n)$.
(ii) Pour toute permutation $\tau\in\mathfrak{S}_n$, $\det\left(L_{\tau(1)},L_{\tau(2)},\dots,L_{\tau(n)}\right)=\varepsilon(\tau)\det(L_1,L_2,\dots,L_n)$.
(iii) Dans $\mathbb{Z}[X_{1,1},X_{1,2},\dots,X_{n,n},Y_1,\dots,Y_n,Z]$, si on note C le vecteur-colonne des indéterminées $Y_1,\dots,Y_n$, alors

$$\det(C_1+C,C_2,\dots C_n)=\det(C_1,C_2,\dots C_n)+\det(C,C_2,\dots C_n)$$
$$\text{et}\quad \det(ZC_1,C_2,\dots C_n)=Z\det(C_1,C_2,\dots C_n).$$

(iv) Dans $\mathbb{Z}[X_{1,1},X_{1,2},\dots,X_{n,n},Y_1,\dots,Y_n,Z]$, si on note L le vecteur-ligne des indéterminées $Y_1,\dots,Y_n$, alors

$$\det(L_1+L,L_2,\dots L_n)=\det(L_1,L_2,\dots L_n)+\det(L,L_2,\dots L_n)$$
$$\text{et}\quad \det(ZL_1,L_2,\dots L_n)=Z\det(L_1,L_2,\dots L_n).$$

(v) En particulier, dans la matrice générique, si on substitue une ligne ou une colonne à une autre, on obtient un déterminant nul.

PREUVE. On montre les assertions sur les colonnes. Celles sur les lignes s'en déduisent par transposition, ou par un raisonnement analogue.
(i) $\det\left(C_{\tau(1)},C_{\tau(2)},\dots,C_{\tau(n)}\right)=\sum_{\sigma\in\mathfrak{S}_n}\varepsilon(\sigma)\prod_{k=1}^n X_{k,\sigma\tau(k)}$. Puisque $\sigma\mapsto\sigma\tau$ est une bijection de $\mathfrak{S}_n$ sur lui-même, $\det\left(C_{\tau(1)},C_{\tau(2)},\dots,C_{\tau(n)}\right)=\varepsilon(\tau)\sum_{\sigma\in\mathfrak{S}_n}\varepsilon(\sigma\tau)\prod_{k=1}^n X_{k,\sigma\tau(k)}=\varepsilon(\tau)\det(C_1,C_1,\dots,C_N)$.
(iii) En isolant les indices de la première colonne, on obtient $\det=\sum_{\sigma\in\mathfrak{S}_n}\varepsilon(\sigma)X_{\sigma(1),1}\prod_{k=2}^n X_{\sigma(k),k}$. Alors, $\det(C_1+C,C_2,\dots C_n)=\sum_{\sigma\in\mathfrak{S}_n}\varepsilon(\sigma)\left(X_{\sigma(1),1}+Y_{\sigma(1),1}\right)\prod_{k=2}^n X_{\sigma(k),k}$ et le résultat s'en suit. De la même façon, $\det(ZC_1,C_2,\dots C_n)=\sum_{\sigma\in\mathfrak{S}_n}\varepsilon(\sigma)\left(ZX_{\sigma(1),1}\right)\prod_{k=2}^n X_{\sigma(k),k}=Z\det(C_1,C_2,\dots C_n)$.
(v) Il suffit de montrer que le déterminant d'une matrice dont les deux premières colonnes sont égales est nul. Or, en appliquant (i) à la transposition $\tau=(12)$ dont la signature est -1, on obtient, dans $\mathbb{Z}[X_{1,1},X_{1,2},\dots,X_{n,n}]$, que $\det(C_1,C_1,\dots,C_n)=-\det(C_1,C_1,\dots,C_n)$, ce qui entraîne que $\det(C_1,C_1,\dots,C_n)=0$. ■

A noter

(i) En combinant (i) et (iii), on montre le linéarité du déterminant en chacune de ses colonnes. *Idem* pour les lignes avec (ii) et (iv).

(ii) La n-linéarité ent le caractère alterné entraînent que si on substitue une colonne à une combinaison linéaire des autres, on ne modifie pas le déterminant. Autre formulation : si, après substitution, les colonnes d'une matrice sont liées, alors son déterminant est nul. *Idem* pour les lignes, bien sûr.

(iii) Pour montrer (v), on utilise le fait que $2 \neq 0$ dans $\mathbb{Z}$. Cela n'empêche pas sa spécialisation dans un anneau de caractéristique 2, par exemple $\mathbb{Z}/2\mathbb{Z}$. En effet, le fait que l'assertion (v) soit vraie sur $\mathbb{Z}$ permet de la transporter telle quelle sur n'importe quel anneau commutatif $\mathcal{A}$ *via* l'homomorphisme de caractéristique $\mathbb{Z} \to \mathcal{A}$, $1 \mapsto 1_{\mathcal{A}}$.

(iv) Après spécialisation, la proposition montre notamment que le déterminant d'une matrice à coefficients dans n'importe quel anneau commutatif est nul dès que l'une des conditions suivantes est vérifiée (la dernière englobe les deux autres) :
- une ligne ou une colonne est nulle ;
- deux lignes ou deux colonnes sont égales ;
- les lignes ou les colonnes sont linéairement dépendantes (exercice : cette dernière condition équivaut à la nullité du déterminant).

(v) Ni la formule de sa définition (sommer sur les permutations) ni l'utilisation récursive des formules de développement selon une ligne ou une colonne (voir plus bas) ne sont adaptées à un calcul algorithmique effectif d'un déterminant. Il suffit pour s'en convaincre de calculer en fonction de n le nombre d'opérations (multiplications et additions) que nécessite le calcul sous ces formes-là du déterminant d'une matrice générique de taille n. En revanche, grâce à l'invariance du déterminant par transformations élémentaires sur les lignes ou les colonnes d'une matrice, l'algorithme du pivot de Gauss est toujours une manière efficace de calculer un déterminant.

Théorème (déterminant d'un produit)
Soient $M = M(X_{1,1}, X_{1,2}, \ldots, X_{n,n})$ et $N = N(Y_{1,1}, Y_{1,2}, \ldots, Y_{n,n})$ deux matrices génériques dont les coefficients sont les indéterminées de $\mathbb{Z}[X_{1,1}, X_{1,2}, \ldots, X_{n,n}, Y_{1,1}, Y_{1,2}, \ldots, Y_{n,n}]$. Alors,

$$\det(MN) = \det M \times \det N.$$

PREUVE. Pour tout $j \in \{1, \ldots, n\}$, on note respectivement $C_j(X)$ et $C_j(Y)$ la j^{e} colonne de M et la j^{e} colonne de N. Par définition du produit matriciel, la j^{e} colonne de MN est $MC_j(Y)$, produit de la matrice carrée M par le vecteur-colonne $C_j(Y)$. Toujours selon la définition du produit matriciel, le produit $MC_j(Y)$ se développe en la somme de vecteurs-colonne $MC_j(Y) = \sum_{i=1}^{n} Y_{i,j} C_i(X)$. On a ainsi successivement

$$\begin{aligned}\det(MN) &= \det\left(MC_1(Y), \ldots, MC_n(Y)\right) = \det\left(\sum_{i=1}^{n} Y_{i,1} C_i(X), \ldots, \sum_{i=1}^{n} Y_{i,n} C_i(X)\right) \\ &= \sum_{i_1,\ldots,i_n \in \{1,\ldots n\}} Y_{i_1,1} \ldots Y_{i_n,n} \det\left(C_{i_1}(X) \ldots, C_{i_n}(X)\right),\end{aligned}$$

la dernière égalité venant de la n-linéarité du déterminant. En utilisant le fait que deux colonnes égales annulent le déterminant, il ne reste plus que les multi-indices de sommation contenant des i_k distincts. Cela permet de ré-écrire cette somme à l'aide de permutations :

$$\det(MN) = \sum_{\sigma \in \mathfrak{S}_n} Y_{\sigma(1),1} \ldots Y_{\sigma(n),n} \det\left(C_{\sigma(1)}(X) \ldots, C_{\sigma(n)}(X)\right).$$

Enfin, le caractère alterné du déterminant permet de conclure :

$$\det(MN) = \sum_{\sigma \in \mathfrak{S}_n} Y_{\sigma(1),1} \ldots Y_{\sigma(n),n} \varepsilon(\sigma) \det\left(C_1(X) \ldots, C_n(X)\right) = \det(M) \det(N).$$

en mettant $\det(M)$ en facteur dans la dernière égalité. ■

Exercice 34
Si $\mathcal{A}$ est un anneau commutatif et si $A \in \mathcal{M}_n(\mathcal{A})$ est inversible dans l'anneau $\mathcal{M}_n(\mathcal{A})$, alors $\det A$ est inversible dans l'anneau $\mathcal{A}$ et $\det\left(A^{-1}\right) = (\det A)^{-1}$.

Pour établir le développement du déterminant d'une matrice carrée selon ses lignes ou ses colonnes, on adopte la notation suivante.

Définition (cofacteurs)
Dans $\mathbb{Z}[X_{1,1}, X_{1,2}, \ldots, X_{n,n}]$ où $n \geq 2$, pour chaque couple d'indices $(i,j) \in \{A, \ldots, n\}^2$, on note $\mathrm{Cof}_{i,j}$ le *cofacteur* d'indice (i,j) qui est le polynôme

$$\mathrm{Cof}_{i,j} = (-1)^{i+j} \det \begin{pmatrix} X_{1,1} & \ldots & X_{1,j-1} & X_{1,j+1} & \ldots & X_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ X_{i-1,1} & \ldots & X_{i-1,j-1} & X_{i-1,j+1} & \ldots & X_{i-1,n} \\ X_{i+1,1} & \ldots & X_{i+1,j-1} & X_{i+1,j+1} & \ldots & X_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ X_{n,1} & \ldots & X_{n,j-1} & X_{n,j+1} & \ldots & X_{n,n} \end{pmatrix}.$$

La matrice dont on prend le déterminant, qui est de taille $(n-1) \times (n-1)$, est obtenue à partir de la matrice générique en supprimant sa i^{e} ligne et sa j^{e} colonne.

A noter
Le cofacteur d'indice (i,j) est aussi le déterminant de la matrice

$$\mathrm{Cof}_{i,j} = \det \begin{pmatrix} X_{1,1} & \ldots & X_{1,j-1} & 0 & X_{1,j+1} & \ldots & X_{1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ X_{i-1,1} & \ldots & X_{i-1,j-1} & 0 & X_{i-1,j+1} & \ldots & X_{i-1,n} \\ 0 & \ldots & 0 & 1 & 0 & \ldots & 0 \\ X_{i+1,1} & \ldots & X_{i+1,j-1} & 0 & X_{i+1,j+1} & \ldots & X_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ X_{n,1} & \ldots & X_{n,j-1} & 0 & X_{n,j+1} & \ldots & X_{n,n} \end{pmatrix}. \tag{6}$$

Pour montrer cela, il suffit de permuter les lignes selon le i-cycle $(n, n-1, \ldots, i)$ et les colonnes selon le j-cycle $(n, n-1, \ldots, j)$ et d'utiliser la formule (5). Le facteur $(-1)^{i+j}$ dans la définition des cofacteurs apparaît-il ainsi comme une signature.

Proposition (développement du déterminant selon une ligne ou une colonne)
Soit $n \geq 2$.
(i) Pour tout $k \in \{1, \ldots, n\}$, le déterminant se développe par rapport à la k^{e} ligne sous la forme du produit matriciel

$$\det(X_{1,1}, X_{1,2}, \ldots, X_{n,n}) = \begin{pmatrix} X_{k,1} & \ldots & X_{k,n} \end{pmatrix} \cdot {}^t\begin{pmatrix} \mathrm{Cof}_{k,1} & \ldots & \mathrm{Cof}_{k,n} \end{pmatrix}.$$

(ii) Pour tout $k \in \{1, \ldots, n\}$, le déterminant se développe par rapport à la k^{e} colonne sous la forme du produit matriciel

$$\det(X_{1,1}, X_{1,2}, \ldots, X_{n,n}) = \begin{pmatrix} \mathrm{Cof}_{1,k} & \ldots & \mathrm{Cof}_{n,k} \end{pmatrix} \cdot {}^t\begin{pmatrix} X_{1,k} & \ldots & X_{n,k} \end{pmatrix}.$$

Preuve. Les deux formules s'obtiennent en combinant la n-linéarité et la formule (6), puisque, si $(\delta_1, \ldots, \delta_n)$ désigne la base canonique de $\mathcal{M}_{n,1}$, on a $\begin{pmatrix} X_{k,1} & \ldots & X_{k,n} \end{pmatrix} = \sum_{i=1}^n X_{k,j}\, {}^t\delta_k$ et *idem* pour les colonnes. ■

Exemple (déterminant de Vandermonde)
Soit $n \geq 2$. Dans l'anneau de polynômes $\mathbb{Z}[X_1, \ldots, X_n]$,

$$V(X_1, \ldots X_n) = \det \begin{pmatrix} 1 & X_1 & X_1^2 & \ldots & X_1^{n-1} \\ 1 & X_2 & X_2^2 & \ldots & X_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & X_n & X_n^2 & \ldots & X_n^{n-1} \end{pmatrix} = \prod_{\substack{(i,j) \in \{1,\ldots,n\}^2 \\ i<j}} (X_j - X_i).$$

Pour prouver cela, on procède par récurrence sur n. Pour $n=2$, c'est la formule du déterminant 2×2. Si $n\geq 2$, pour tout $k\geq 2$, on remplace la colonne C_k de la matrice de Vandermonde par $C_k - X_n C_{k-1}$. On obtient

$$V(X_1,\dots X_n) = \det\begin{pmatrix} 1 & X_1-X_n & X_1(X_1-X_n) & \dots & X_1^{n-2}(X_1-X_n) \\ 1 & X_2-X_n & X_2(X_2-X_n) & \dots & X_2^{n-2}(X_2-X_n) \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & X_{n-1}-X_n & X_{n-1}(X_{n-1}-X_n) & \dots & X_{n-1}^{n-2}(X_{n-1}-X_n) \\ 1 & 0 & 0 & \dots & 0 \end{pmatrix}$$

En développant selon la dernière ligne et en utilisant la n-linéarité,

$$V(X_1,\dots X_n) = \left((-1)^{n-1}\prod_{k=1}^{n-1}(X_k-X_n)\right) V(X_1,\dots X_{n-1}),$$

ce qui permet de conclure par récurrence. ■

Exercice 35
Soit $\mathbb{F}$ un corps. Démontrer, avec le déterminant de Vandermonde, que tout polynôme de $\mathbb{F}[X]$ de degré d ayant au moins $d+1$ racines distinctes dans $\mathbb{F}$ est nécessairement nul — cela nécessite de montrer auparavant qu'un système linéaire homogène ayant autant d'inconnues que d'équations admet une solution non triviale (si, et) seulement si son déterminant est nul.

Définition (comatrice)
La *comatrice* de la matrice générique à n^2 indéterminées est la matrice de ses cofacteurs :

$$\mathrm{Com}(X_{1,1}, X_{1,2}, \dots, X_{n,n}) = (\mathrm{Cof}_{i,j})_{1\leq i,j\leq n}.$$

Proposition (matrice et comatrice)
Pour tout $n\geq 1$,

$$M\times {}^t\mathrm{Com} = {}^t\mathrm{Com}\times M = \det(M)\cdot I_n. \tag{7}$$

A noter
Ces produits matriciels doivent être lus comme deux fois n^2 identités poynomiales.

Preuve. Le coefficient de la i^{e} ligne et de la j^{e} colonne de $M\,{}^t\mathrm{Com}$ est $\begin{pmatrix} X_{i,1} & \dots & X_{i,n}\end{pmatrix}\cdot {}^t\begin{pmatrix}\mathrm{Cof}_{j,1} & \dots & \mathrm{Cof}_{j,n}\end{pmatrix}$ puisque la j^{e} colonne de ${}^t\mathrm{Com}$ est ${}^t\begin{pmatrix}\mathrm{Cof}_{j,1} & \dots & \mathrm{Cof}_{j,n}\end{pmatrix}$. Lorsque $i=j$, ce produit égale det, c'est une redite du développement de det $=\det M$ par rapport à la i^{e} ligne. Lorsque $i\neq j$, en substituant L_i à L_j dans det, on obtient, en développant par rapport à la j^{e} ligne, que

$$\det(L_1,\dots,L_i,\dots,L_i,\dots,L_n) = L_i\cdot {}^t\begin{pmatrix}\mathrm{Cof}_{j,1} & \dots & \mathrm{Cof}_{j,n}\end{pmatrix} = \begin{pmatrix} X_{i,1} & \dots & X_{i,n}\end{pmatrix}\cdot {}^t\begin{pmatrix}\mathrm{Cof}_{j,1} & \dots & \mathrm{Cof}_{j,n}\end{pmatrix},$$

le déterminant du terme de gauche étant nul puisque deux lignes sont égales — l'une des deux lignes L_i écrites est au rang j, l'autre au rang i. ■

Exercice 36
(i) Si $\mathcal{A}$ est un anneau commutatif et si $A\in\mathcal{M}_n(\mathcal{A})$, alors A est inversible à droite (resp. à gauche) dans l'anneau $\mathcal{M}_n(\mathcal{A})$ si, et seulement si $\det A$ est inversible dans l'anneau $\mathcal{A}$. En particulier, A est inversible à droite si, et seulement si A est inversible à gauche. Dans ces conditions, $A^{-1} = (\det A)^{-1}\times {}^t\mathrm{Com}(A)$.
(ii) Si $\mathbb{F}$ est un corps et si $A\in\mathcal{M}_n(\mathbb{F})$, alors $A\in \mathrm{GL}(n,\mathbb{F})$ si, et seulement si $\det A\neq 0$.

Théorème (Cayley-Hamilton)
Soient $n\geq 1$ et M la matrice $n\times n$ générique. On note $\chi_M(X)\in\mathbb{Z}[X, X_{1,1}, X_{1,2},\dots,X_{n,n}]$ le polynôme caractéristique de M, i.e. $\chi_M(X) = \det(XI_n - M)$. *Alors, en notant O_n la matrice $n\times n$ nulle,*

$$\chi_M(M) = O_n$$

A noter
Egalité entre matrices à coefficients polynomiaux, le théorème de Cayley-Hamilton doit être lu comme n^2 identités polynomiales.

PREUVE. Attention au raisonnement hâtif et erroné qui consiste à remplacer X par M dans χ_M, penser que XI_n devient alors M (c'est là l'erreur) et conclure. L'idée qui consiste à spécialiser [X=M] n'est pas stupide du tout, mais requiert davantage d'attention. On part de (7) que l'on applique à la matrice $XI_n - M$ dont les coefficients sont dans $\mathbb{Z}[X, X_{1,1}, X_{1,2}, \ldots, X_{n,n}]$. On obtient

$$\mathrm{Com}\,(XI_n - M) \times \left(XI_n - {}^tM\right) = \chi_M(X) I_n. \tag{8}$$

On note $\mathbb{Z}[M]$ le sous-anneau — commutatif — de $\mathcal{M}_n(\mathbb{Z}[X_{1,1}, X_{1,2}, \ldots, X_{n,n}])$ engendré par M, qui est l'anneau des polynômes en M à coefficients entiers. Soit alors $s : \mathbb{Z}[X, X_{1,1}, X_{1,2}, \ldots, X_{n,n}] \longrightarrow \mathbb{Z}[M]$ la substitution $X \mapsto M$. On prolonge s aux matrices $n \times n$ à coefficients dans $\mathbb{Z}[X, X_{1,1}, X_{1,2}, \ldots, X_{n,n}]$ et on le note encore s. Ainsi, l'homomorphisme d'anneaux

$$s : \mathcal{M}_n(\mathbb{Z}[X, X_{1,1}, X_{1,2}, \ldots, X_{n,n}]) \qu\longrightarrow \quad \mathcal{M}_n(\mathbb{Z}[M])$$

envoie une matrice à coefficients polynomiaux en $X, X_{1,1}, \ldots X_{n,n}$ sur une matrice dont les coefficients sont eux-même des matrices à coefficients polynomiaux en $X_{1,1}, \ldots X_{n,n}$, en remplaçant X par M. Pour tout $k \in \{1, \ldots, n\}$, on note $\delta_k \in \mathcal{M}_{n,1}(\mathbb{Z})$ le vecteur-colonne canonique $\delta_k = {}^t(0, \ldots, 0, 1, 0, \ldots, 0)$, le 1 étant placé au rang k. Alors, si $A \in \mathcal{M}_n(\mathbb{Z}[M])$, et si V est un vecteur-colonne dont les coefficients sont eux-même des vecteurs-colonne, le produit matriciel $A \cdot V$ a encore du sens par les règles habituelles de calcul matriciel ; c'est encore un vecteur-colonne à coefficients vecteurs-colonne. En particulier,

$$s(\chi_M(X) I_n) \cdot \begin{pmatrix} \delta_1 \\ \vdots \\ \delta_n \end{pmatrix} = \begin{pmatrix} \chi_M(M)\delta_1 \\ \vdots \\ \chi_M(M)\delta_n \end{pmatrix}. \tag{9}$$

Par ailleurs, en notant $M_{i,j} = X_{i,j}$ le coefficient de la i^{e} ligne et de la j^{e} colonne de M,

$$s\left(XI_n - {}^tM\right) \cdot \begin{pmatrix} \delta_1 \\ \vdots \\ \delta_n \end{pmatrix} = \begin{pmatrix} M\delta_1 - \sum_{k=1}^n M_{k,1}\delta_k \\ \vdots \\ M\delta_n - \sum_{k=1}^n M_{k,n}\delta_k \end{pmatrix} = \begin{pmatrix} 0_n \\ \vdots \\ 0_n \end{pmatrix} \tag{10}$$

où 0_n désigne le vecteur-colonne nul de $\mathcal{M}_{n,1}(\mathbb{Z})$. Combiner (8), (9) et (10) conduit au résultat puisqu'une matrice carrée est nulle si, et seulement son produit par tous les δ_k est nul. ■

Exercice 37
Faire une autre preuve du théorème de Cayley-Hamilton ainsi énoncé en le prouvant d'abord pour les matrices diagonalisables à coefficients complexes, puis en utilisant le théorème de prolongement analytique pour les polynômes à plusieurs indéterminées.

Exercice 38
Avec cette étude sur le déterminant générique, retrouver tous les résultats standard d'algèbre linéaire des premières années d'enseignement supérieur qui font intervenir un déterminant (sur un corps), notamment ceux qui concernent la résolution des systèmes linéaires. En voici quelques exemples.

(i) Deux matrices semblables ont le même déterminant. On définit ainsi le déterminant d'un endomorphisme d'un espace vectoriel de dimension finie, qui est le déterminant commun à toutes les matrices qui le représentent *via* le choix d'une base.
Mieux encore, deux matrices semblables ont le même polynôme caractéristique — mais la réciproque est fausse, chercher un contre exemple le plus parlant possible. On définit ainsi le polynôme caractéristique d'un endomorphisme comme le polynôme caractéristique commun à toutes les matrices qui le représentent *via* le choix d'une base.

(ii) Si $\mathbb{F}$ est un corps, une matrice carrée est inversible dans $\mathcal{M}_n(\mathbb{F})$ si, et seulement si son déterminant est non nul. Un endomorphisme d'une espace vectoriel de dimension finie est bijectif si, et seulement si son déterminant est non nul.

(iii) Une base $\mathcal{B}$ d'un espace vectoriel V de dimension finie n étant donnée, on définit le déterminant d'une n-uplet de vecteurs comme étant le déterminant de la matrice de leurs coordonnées dans la base $\mathcal{B}$. On le note

$$\det_{\mathcal{B}}(v_1, \ldots, v_n).$$

Si V est un espace vectoriel de dimension finie n, l'espace vectoriel des applications n-linéaires alternées[♫] sur V est une droite vectorielle. Si $\mathcal{B}$ est n'importe quelle base de V, l'application $(v_1, \ldots, v_n) \mapsto \det_{\mathcal{B}}(v_1, \ldots, v_n)$ est une base de cet espace ; en outre, $\det_{\mathcal{B}}(\mathcal{B}) = 1$.
Enfin, si f est un endomorphisme de V,

$$\det_{\mathcal{B}}(f(v_1), \ldots, f(v_n)) = \det(f) \times \det_{\mathcal{B}}(v_1, \ldots, v_n).$$

(iv) Sur un corps, un système linéaire homogène admet une solution non triviale si, et seulement si son déterminant est nul.

(v) Sur un corps $\mathbb{F}$, si $A \in \mathrm{GL}(n, \mathbb{F})$ est une matrice inversible et si $B \in \mathcal{M}_{n,1}(\mathbb{F})$, alors le système linéaire $AX = B$ admet une unique solution $X = {}^t(x_1, \ldots, x_n) \in \mathcal{M}_{n,1}(\mathbb{F})$, dont la k^{e} coordonnée s'écrit selon les *formules de Cramer*

$$x_k = \frac{\det A_k}{\det A}$$

où A_k est la matrice obtenue en remplaçant la k^{e} colonne de A par le vecteur-colonne B.

(vi) Si r est un entier naturel non nul, les *mineurs d'ordre r d'une matrice (rectangulaire)* sont les déterminants de ses sous-matrices carrées $r \times r$. Alors, une matrice est de rang r si, et seulement si elle admet un mineur d'ordre r non nul alors que tous ses mineurs d'ordre $r+1$ sont nuls.

(vii) Sur un corps $\mathbb{F}$, deux matrices carrées $n \times n$ (ou deux endomorphismes) sont semblables si, et seulement si elles ont les mêmes les facteurs invariants dits encore *invariants de similitude*, qui sont des suites de polynômes unitaires à une indéterminée qui se divisent les uns les autres, sous la forme $P_1|P_2|\ldots|P_m$.
Dire que la suite $P_1|P_2|\ldots|P_m$ est la suite des invariants de similitude d'un endomorphisme f d'un espace vectoriel de dimension finie V signifie que V se décompose en une somme directe $V = \bigoplus_{k=1}^{m} V_k$ de sous-espaces stables par f et que, pour chaque k, le polynôme minimal et le polynôme caractéristique de l'endomorphisme de V_k induit par f sont tous les deux égaux à P_k. En particulier, P_m est le polynôme minimal de f et $P_1 \ldots P_m$ son polynôme caractéristique, la somme des degrés des P_k valant n.
Calculer les invariants de similitude d'une matrice $A \in \mathcal{M}_n(\mathbb{F})$ revient essentiellement à effectuer l'algorithme du pivot de Gauss sur la matrice $XI_n - A$ dans l'anneau euclidien $\mathbb{F}[X]$, les invariants de similitudes apparaissant alors sur la diagonale de la matrice échelonnée réduite obtenue à la fin de l'algorithme.

4.2 Transvections et dilatations

Dans toute cette section, V est un espace vectoriel de dimension finie n sur un corps $\mathbb{F}$. On note $\mathrm{End}(V)$ l'espace vectoriel des applications linéaires $V \to V$ et V^* l'espace dual de V, qui est l'espace des formes linéaires $V \to \mathbb{F}$.
Un *hyperplan* de V en est un sous-espace de dimension $n-1$. Une forme linéaire $u \in V^* \setminus \{0\}$ est une *équation* d'un hyperplan H lorsque $H = \ker u$.

Les transvections et les dilatations de V sont les endomorphismes qui ont un hyperplan de vecteurs fixes. L'objectif principal est de montrer que les transvections engendrent le groupe spécial linéaire et que les transvections et les dilatations engendrent le groupe linéaire.

Définition (vecteur fixe)
Si f est un endomorphisme de V et si $v \in V$, on dit que v est un *vecteur fixe de f* lorsque $f(v) = v$.

A noter
L'ensemble des vecteurs fixes de V en est un sous-espace vectoriel, qui est $\ker(f - \mathrm{id}_V)$. Autrement dit, le sous-espace de points fixes de f est l'espace propre de f associé à la valeur propre 1.

[♫] Attention, sur un corps de caractéristique 2, les formes multilinéaires antisymétriques sont symétriques et ne coïncident pas avec les formes alternées.

Définition (transvection)
Un endomorphisme $t \in \mathrm{End}(V)$ est une *transvection* de V lorsqu'il existe une base $\mathcal{B}$ de V pour laquelle

$$\mathrm{Mat}_{\mathcal{B}}(t) = \begin{pmatrix} 1 & & & & & 0 \\ & 1 & & & & 0 \\ & & 1 & & & 0 \\ & & & \ddots & & \vdots \\ & & & & 1 & 1 \\ 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix}. \tag{11}$$

A noter
Soit t une transvection qui admet la matrice (11) dans la base $\mathcal{B} = (v_1, \dots, v_n)$ de V.
Alors le sous-espace des vecteurs fixes de t est l'hyperplan $H = \ker(t - \mathrm{id}_V) = \mathrm{Vect}(v_1, \dots, v_{n-1})$ et l'image de $t - \mathrm{id}_V$ est la droite vectorielle $D = \mathrm{im}(t - \mathrm{id}_V) = \mathrm{Vect}(v_{n-1})$. Cette dernière est incluse dans H. On appelle respectivement H et D *l'hyperplan de* t et *la droite de* t.
Si $(x_1, \dots, x_n)$ est le système générique des coordonnées dans $\mathcal{B}$, l'hyperplan de t a pour équation $x_n = 0$ et la transvection s'écrit $t(v) = v + x_n v_{n-1}$, pour tout $v \in V$ — il suffit de vérifier cela sur les vecteurs de la base $\mathcal{B}$.
Enfin, une transvection n'est pas diagonalisable.

Proposition (caractérisation des transvections)
Soient V un $\mathbb{F}$-espace vectoriel de dimension finie et $f \in \mathrm{End}(V)$. Les assertions suivantes sont équivalentes.

(i) f est une transvection

(ii) $\ker(f - \mathrm{id}_V)$ *est un hyperplan et* $\det f = 1$

(iii) $\ker(f - \mathrm{id}_V)$ *est un hyperplan et f n'est pas diagonalisable*

(iv) $\ker(f - \mathrm{id}_V)$ *est un hyperplan et la droite* $\mathrm{im}(f - \mathrm{id}_V)$ *est incluse dans* $\ker(f - \mathrm{id}_V)$

(v) il existe $u \in V^* \setminus \{0\}$ et $h \in \ker u \setminus \{0\}$ tels que $f(v) = v + u(v)h$, pour tout $v \in V$.

Preuve. (i)⇒(ii) est immédiat. (ii)⇒(iii) Si $\ker(f - \mathrm{id}_V)$ est un hyperplan, alors 1 est une valeur propre de f de multiplicité au moins $d-1$. Puisque $\det f = 1$, la valeur propre 1 est de multiplicité d alors que l'espace de vecteurs fixes est de dimension $d-1$. Donc f n'est pas diagonalisable. (iii)⇒(iv) Puisque $H = \ker(f - \mathrm{id}_V)$ est un hyperplan, grâce au théorème du rang, $D = \mathrm{im}(f - \mathrm{id}_V)$ est une droite vectorielle stable par f, c'est-à-dire une droite de vecteurs propres pour f. Si elle n'était pas dans H, alors f serait diagonalisable puisque la concaténation d'une base de H et d'une base de D formerait une base de vecteurs propres de f. (iv)⇒(v) On note encore $H = \ker(f - \mathrm{id}_V)$ et $D = \mathrm{im}(f - \mathrm{id}_V)$ et on suppose que H est un hyperplan (ce qui implique que D est une droite) et que $D \subseteq H$. Soit $u \in V^*$ une équation de H et soit $w \in V$ tel que $u(w) = 1$ — un tel w existe puisque $u \neq 0$. Soit alors $h = f(w) - w$. Comme $h \in D$, $h \in H$; en outre, $h \neq 0$ puisque $f(w) \neq w$. Alors, $V = H \oplus \mathbb{F}w$ et $f(v) = v + u(v)h$ pour tout $v \in V$ puisque cette formule est vraie sur H et en w. (v)⇒(i) Soient $H = \ker u = \ker(f - \mathrm{id}_V)$ et $w \in u^{-1}(1)$. On complète h en une base $(h_1, \dots, h_{n-2}, h)$ de H. Alors, $(h_1, \dots, h_{n-2}, h, w)$ est une base de V et la matrice de f dans cette base a la forme requise, puisque les h_k et h sont fixes et puisque $f(w) = h + w$. ∎

A noter
(i) *L'inverse d'une transvection est encore une transvection.*
En effet, du point de vue matriciel, l'inverse de la matrice (11) est de la même forme en remplaçant le 1 non diagonal de la dernière colonne par un -1, ce qui donne encore une matrice de transvection puisqu'elle est semblable à (11). Du point de vue géométrique, si f est une transvection de la forme (v) ci-dessus, son inverse est la transvection $v \mapsto v - u(v)h$.

(ii) *Toute conjuguée dans* $\mathrm{GL}(V)$ *d'une transvection est encore une transvection.*
C'est une conséquence directe de la définition d'une transvection, qui dit même que *deux transvections quelconques sont conjuguées dans* GL.
Mieux, si $g \in \mathrm{GL}(V)$ et si t est une transvection d'hyperplan H et de droite $D \subseteq H$, alors gtg^{-1} est une transvection d'hyperplan $g(H)$ et de droite $g(D)$. Plus précisément encore, si $u \in V^* \setminus \{0\}$ et $h \in \ker u \setminus \{0\}$, on note $t(u,h)$ la transvection de V définie par la caractérisation (v), c'est-à-dire par la formule $\forall v \in V$, $t(u,h)(v) = v + u(v)h$. Avec cette notation, $gt(u,h)g^{-1} = t\left(u \circ g^{-1}, g(h)\right)$ pour toute $g \in \mathrm{GL}(V)$ — bien noter que $u \circ g^{-1} \in V^* \setminus \{0\}$ et que $g(h) \in \ker\left(u \circ g^{-1}\right) \setminus \{0\}$.

Lemme (géométrique, pour la preuve du théorème qui suit)
On suppose que V est de dimension au moins 2.
(i) Si $x, y \in V \setminus \{0\}$, il existe $\tau \in \mathrm{End}(V)$ tel que τ soit un produit d'une ou deux transvections et $\tau(x) = y$.
(ii) Si H et K sont deux hyperplans distincts de V et si $x \in V \setminus H \cup K$, alors il existe une transvection τ telle que $\tau(x) = x$ et $\tau(H) = K$.

Preuve. (i) On suppose d'abord que x et y ne sont pas colinéaires. On pose $h = y - x$ et on prend un hyperplan H de V qui contienne h mais pas x, ce qui est possible puisque $n = \dim V \geq 2$ et $\{x, h\}$ est libre. Soit alors u une équation de H telle que $u(x) = 1$. Alors, $\tau(v) = v + u(v)h$ définit une transvection qui envoie x sur y. Ensuite, si x et y sont colinéaires, soit $z \in V \setminus \mathbb{F}x$. Un tel z existe puisque $\dim V \geq 2$. Selon ce qui précède, soient t_1 et t_2 deux transvections telles que $t_1(x) = z$ et $t_2(z) = y$. Alors, $\tau = t_2 \circ t_1$ envoie x sur y.
(ii) $H \cap K$ est un sous-espace de V de dimension $n - 2$. Puisque $x \notin H \cap K$, l'espace $\mathbb{F}x \oplus (H \cap K)$ est un hyperplan de V ; soit $u \in V^*$ une équation de cet hyperplan. Soient alors $h \in H$ et $k \in K$ tels que $H = H \cap K \oplus \mathbb{F}h$ et $K = H \cap K \oplus \mathbb{F}k$. Puisque $u(h) \neq 0 \neq u(k)$, quitte à remplacer h et k par $h/u(h)$ et $k/u(k)$, on peut supposer que $u(h) = u(k) = 1$. Alors, la transvection $\tau : v \mapsto v + u(v)(k - h)$ envoie h sur k et fixe $H \cap K$: elle envoie H sur K. En outre, elle fixe x puisque x est dans son hyperplan. ■

Théorème (les transvections engendrent $\mathrm{SL}(V)$)
Soit V un espace vectoriel de dimension finie. Alors, le groupe $\mathrm{SL}(V)$ est engendré par ses transvections.

Preuve. On procède par récurrence sur $n = \dim_{\mathbb{F}} V$. Si $n = 1$, alors $\mathrm{SL}(V) = \{\mathrm{id}_V\}$ et il n'y a rien à démontrer. On suppose que $n \geq 2$.
Soient $f \in \mathrm{SL}(V)$ et $x \in V \setminus \{0\}$. On cherche à montrer que f est un produit (une composée) de transvections. En appliquant le (i) du lemme géométrique, soit τ_1 un produit d'une ou deux transvections tel que $\tau_1(f(x)) = x$. Alors, $\tau_1 f \in \mathrm{SL}(V)$ et $\tau_1 f$ fixe x. On peut donc supposer que f fixe x. Soit H un hyperplan de V tel que $H \oplus \mathbb{F}x = V$. En particulier, $x \notin H \cap f(H)$ puisque $f^{-1}(x) = x \notin H$. Si $f(H) \neq H$, en appliquant le (i) du lemme géométrique, soit τ_2 une transvection telle que $\tau_2(f(H)) = H$ et $\tau_2(x) = x$. Alors, $\tau_2 f \in \mathrm{SL}(V)$, fixe x et vérifie $\tau_2 f(H) = H$ — on dit que $\tau_2 f$ stabilise H.
Ainsi, on peut supposer que f fixe x et stabilise un hyperplan H tel que $H \oplus \mathbb{F}x = V$.
Dans cette situation où $H \oplus \mathbb{F}x = V$, si $\varphi \in \mathrm{End}(H)$, on note $\varphi \oplus \mathrm{id}$ l'endomorphisme de V défini par $\varphi \oplus \mathrm{id}(h + \xi x) = \varphi(h) + \xi x$ (notations évidentes, $\xi \in \mathbb{F}$). L'ensemble des endomorphismes de V de déterminant 1, qui fixent x et qui stabilisent H forme un sous-groupe $\mathrm{SL}_{H,x}(V)$ de $\mathrm{SL}(V)$ et l'application $\mathrm{SL}(H) \to \mathrm{SL}_{H,x}(V)$, $\varphi \mapsto \varphi \oplus \mathrm{id}$ est un isomorphisme de groupes dont la réciproque envoie l'endomorphisme f sur l'endomorphisme $f_H \in \mathrm{End}(H)$ induit sur H par restriction. En outre, cet isomorphisme transforme toute transvection de H en une transvection de V. [On peut, si l'on veut, adopter un point de vue matriciel pour argumenter tous ces derniers points.]
On revient au $f \in \mathrm{SL}_{H,x}(V)$ dont on cherche à montrer qu'il est produit de transvections. Par récurrence, f_H est un produit de transvections de H. Par le mécanisme décrit au paragraphe précédent, on prolonge ces transvections en des transvections de V dont le produit égale f. ■

Corollaire (centre de SL)
Soit $\mathbb{F}$ un corps et $n \in \mathbb{N} \setminus \{0\}$.
Le centre de $\mathrm{SL}(n, \mathbb{F})$ est le groupe des homothéties de la forme $\{xI_n,\ x \in \mathbb{F},\ x^n = 1\}$, isomorphe au groupe des racines n^{e} de l'unité dans $\mathbb{F}$.

Preuve. On raisonne comme dans le calcul du centre de GL déjà fait au chapitre de généralités sur les groupes. La première partie de la preuve est simplifiée par la formule de conjugaison des transvections : soit $f \in Z(\mathrm{SL}(\mathbb{F}^n))$. Alors, f commute avec toutes les transvections de $\mathbb{F}^n$. Ainsi, d'après la formule de conjugaison (par f) des transvections, f stabilise toutes les droites de $\mathbb{F}^n$. Donc f est une homothétie, comme le montre la fin de la preuve déjà faite du calcul du centre de GL. ■

Définition (dilatation)
Un endomorphisme $d \in \mathrm{End}(V)$ est une *dilatation* de V lorsqu'il existe une base $\mathcal{B}$ de V et $x \in \mathbb{F} \setminus \{0,1\}$ tels que

$$\mathrm{Mat}_{\mathcal{B}}(d) = \mathrm{diag}(1,\dots,1,x) = \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & x \end{pmatrix}. \tag{12}$$

A noter
Soit d une dilatation qui admet la matrice $\mathrm{diag}(1,\dots,1,x)$ dans la base $\mathcal{B} = (v_1,\dots,v_n)$ de V.
Alors le sous-espace des vecteurs fixes de d est l'hyperplan $H = \ker(d - \mathrm{id}_V) = \mathrm{Vect}(v_1,\dots,v_{n-1})$ et l'image de $d - \mathrm{id}_V$ est la droite vectorielle $D = \mathrm{im}(d - \mathrm{id}_V) = \mathrm{Vect}(v_n)$. Cette dernière, qui est aussi les sous-espace propre de d associé à la valeur propre x, est un supplémentaire de H dans V : $V = H \oplus D$. On appelle respectivement x, H et D *le rapport*, *l'hyperplan* et *la droite de* d.
Ces trois données caractérisent une dilatation : lorsque $x \in \mathbb{F} \setminus \{0,1\}$, H un hyperplan et D est une droite qui n'est pas contenue dans H, on parle de *la* dilatation de rapport x, d'hyperplan H et de droite D. Avec les notations de la preuve du fait que les transvections engendrent SL, la dilatation de rapport x, d'hyperplan H et de droite D est $\mathrm{id}_H \oplus x\,\mathrm{id}_D : v_H + v_D \mapsto v_H + xv_D$ (notations évidentes).
Enfin, une dilatation est diagonalisable.

Proposition (caractérisation des dilatations)
Soient V un $\mathbb{F}$-espace vectoriel de dimension finie et $f \in \mathrm{GL}(V)$. Les assertions suivantes sont équivalentes.

(i) f est une dilatation

(ii) $\ker(f - \mathrm{id}_V)$ *est un hyperplan et* $\det f \neq 1$

(iii) $\ker(f - \mathrm{id}_V)$ *est un hyperplan et f est diagonalisable*

(iv) $\ker(f - \mathrm{id}_V)$ *est un hyperplan et la droite* $\mathrm{im}(f - \mathrm{id}_V)$ *n'est pas incluse dans* $\ker(f - \mathrm{id}_V)$

(v) Il existe $x \in \mathbb{F} \setminus \{0,1\}$, un hyperplan H et une droite supplémentaire D telles que $f = \mathrm{id}_H \oplus x\,\mathrm{id}_D$.

Preuve. (i)⇒(ii) Si le rapport de f est x, alors $\det f = x$. (ii)⇒(iii) Si le sous-espace des vecteurs fixes par f est un hyperplan, son polynôme caractérisitique est $(X-1)^{n-1}(X - \det f)$. Si $\det f \neq 1$, alors f est diagonalisable. (iii)⇒(iv) Puisque f est diagonalisable et puisque $H = \ker(f - \mathrm{id}_V)$ est un hyperplan, soit $D = \ker(f - x\,\mathrm{id}_V)$ la droite de vecteurs propres associée à la valeur propre différente de 1, que l'on nomme x. Alors, $D \neq H$. En outre, le théorème du rang assure que $\mathrm{im}(f - \mathrm{id}_V)$ est une droite. Enfin, puisque $x \neq 1$, tout vecteur v de D s'écrit $v = \frac{(f-\mathrm{id})(v)}{x-1}$. Donc la droite D est incluse dans la droite $\mathrm{im}(f - \mathrm{id}_V)$: ces deux droites sont donc égales ; elles ne sont pas dans H. (iv)⇒(v) La droite $D = \mathrm{im}(f - \mathrm{id}_V)$ est stable par f : c'est une droite propre, associée à une valeur propre x. Puisque D n'est pas incluse dans l'hyperplan $H = \ker(f - \mathrm{id}_V)$, alors $x \neq 1$, $V = H \oplus D$ et $f = \mathrm{id}_H \oplus x\,\mathrm{id}_D$. (v)⇒(i) C'est immédiat, voir le *à noter* qui suit la définition d'une dilatation. ■

A noter
(i) L'inverse d'une dilatation est encore une dilatation : même hyperplan, même droite, rapports inverses l'un de l'autre.

(ii) *Toute conjuguée dans* $\mathrm{GL}(V)$ *d'une dilatation est encore une dilatation.*
Plus précisément, soient $g \in \mathrm{GL}(V)$ et d la dilatation de rapport x, d'hyperplan H et de droite D. Alors, gdg^{-1} est la dilation de rapport x, d'hyperplan $g(H)$ et de droite $g(D)$.
Enfin, conséquence immédiate de la définition, *deux dilatations sont conjuguées dans* GL *si, et seulement si elles ont le même rapport.*

Proposition (les transvections et les dilatations engendrent $\mathrm{GL}(V)$)
Soit V un espace vectoriel de dimension finie. Alors, le groupe $\mathrm{GL}(V)$ *est engendré par ses transvections et ses dilatations.*

Preuve. Soit $f \in \mathrm{GL}(V)$. Si $f \in \mathrm{SL}(V)$, alors f est produit de transvections. Sinon, soit d n'importe quelle dilatation de rapport $\det f$. Alors, $d^{-1}f \in \mathrm{SL}(V)$ est un produit de transvections. ■

Définition (homographies, groupes PGL **et** PSL**)**
Soit V un espace vectoriel de dimension finie n sur un corps $\mathbb{F}$. On l'a vu, les groupes $\mathrm{GL}(V)$ et $\mathrm{SL}(V)$ admettent respectivement pour centres les groupes d'homothéties $Z(\mathrm{GL}(V)) = \mathbb{F}^\times \mathrm{id}_V$ et $Z(\mathrm{SL}(V)) = \mu_n(\mathbb{F})\,\mathrm{id}_V$, où $\mu_n(\mathbb{F})$ désigne le groupe des racines n^{e} de l'unité dans $\mathbb{F}$. On note $\mathrm{PGL}(V)$ et $\mathrm{PSL}(V)$ les groupes-quotient suivants :

$$\mathrm{PGL}(V) = \mathrm{GL}(V)/\mathbb{F}^\times \mathrm{id}_V \quad \text{et} \quad \mathrm{PSL}(V) = \mathrm{SL}(V)/\mu_n(\mathbb{F})\,\mathrm{id}_v\,.$$

Un élément de $\mathrm{PGL}(V)$ est une *homographie* sur V. Un élément de $\mathrm{PSL}(V)$ est une *homographie spéciale* sur V. [A vrai dire, ce vocabulaire trouve tout son sens lorsqu'on considère ces objets comme les transformations de *l'espace projectif* $\mathbb{P}(V)$ qui est l'ensemble des droites vectorielles de V.]
Du côté des matrices, de façon analogue, on note

$$\mathrm{PGL}(n,\mathbb{F}) = \mathrm{GL}(n,\mathbb{F})/\mathbb{F}^\times I_n \quad \text{et} \quad \mathrm{PSL}(n,\mathbb{F}) = \mathrm{SL}(n,\mathbb{F})/\mu_n(\mathbb{F})\,I_n.$$

Exercice 39
(i) Montrer que $\mathrm{GL}(n,\mathbb{Z}/2\mathbb{Z}) = \mathrm{SL}(n,\mathbb{Z}/2\mathbb{Z}) \simeq \mathrm{PGL}(n,\mathbb{Z}/2\mathbb{Z}) = \mathrm{PSL}(n,\mathbb{Z}/2\mathbb{Z})$.
(ii) Montrer que selon que n est pair ou impair, $\mathrm{PSL}(n,\mathbb{R}) = \mathrm{SL}(n,\mathbb{R})/\{\pm I_n\}$ ou $\mathrm{PSL}(n,\mathbb{R}) \simeq \mathrm{SL}(n,\mathbb{R})$.

Théorème (simplicité de PSL**, sauf cas sporadiques)**
Soient $\mathbb{F}$ un corps et $n \geq 1$. Alors, sauf lorsque $(n,\mathbb{F}) = (2,\mathbb{Z}/2\mathbb{Z})$ ou $(n,\mathbb{F}) = (2,\mathbb{Z}/3\mathbb{Z})$,

$$\textit{le groupe } \mathrm{PSL}(n,\mathbb{F}) \textit{ est simple.}$$

Les deux cas sporadiques
Un fois intallée la notion d'action d'un groupe sur un ensemble, on verra que $\mathrm{PSL}(2,\mathbb{Z}/2\mathbb{Z}) \simeq \mathfrak{S}_3$ et que $\mathrm{PSL}(2,\mathbb{Z}/3\mathbb{Z}) \simeq \mathfrak{A}_4$ dont on sait qu'ils ne sont pas simples.

Preuve. Voir liste d'exercices numéro 2 et 3. ■

4.3 Le groupe linéaire sur les corps finis

Théorème (corps finis)
(i) Le cardinal d'un corps fini est nécessairement la puissance d'un nombre premier.
(ii) Si p est un nombre premier et si $d \in \mathbb{N}^*$, il existe un corps de cardinal p^d, unique à isomorphisme (d'anneaux) près.

Preuve. (i) Soient $\mathbb{F}$ un corps fini et p sa caractéristique. Alors p est premier puisque l'homomorphisme d'anneaux $\mathbb{Z} \to \mathbb{F}$ $(1 \mapsto 1_{\mathbb{F}})$ a pour noyau $p\mathbb{Z}$ et se factorise donc en un homomorphisme injectif $i : \mathbb{Z}/p\mathbb{Z} \to \mathbb{F}$, ce qui oblige $\mathbb{Z}/p\mathbb{Z}$ à être intègre. En outre, $\mathbb{Z}/p\mathbb{Z}$ est un corps ainsi que son image par i qui est un sous-corps de $\mathbb{F}$ contenu dans tous les sous-corps de $\mathbb{F}$. On appelle $\mathbb{F}_p = i(\mathbb{Z}/p\mathbb{Z})$ le *sous-corps premier de* $\mathbb{F}$. Dans ces conditions, pour l'addition et la multiplication dans $\mathbb{F}$, le corps $\mathbb{F}$ est un espace vectoriel sur $\mathbb{F}_p$, de dimension finie d puisque l'ensemble $\mathbb{F}$ lui-même est fini. Ainsi, en tant qu'espace vectoriel, $\mathbb{F}$ est isomorphe à $\mathbb{F}_p^d$, dont le cardinal est p^d.
On admet le (ii), ☹. Pour l'essentiel, retenir que le groupe multiplicatif d'un corps $\mathbb{F}$ à $q = p^d$ éléments est d'ordre $q-1$. Ainsi, d'après le théorème de Lagrange, tout élément de $\mathbb{F}$ est racine du polynôme $X^q - X$ dont les coefficients (± 1) sont dans le sous-corps premier $\mathbb{F}_p$ de $\mathbb{F}$. En procédant par quotients successifs de l'anneau principal $\mathbb{F}_p[X]$ par les facteurs irréductibles de $X^q - X$, on obtient ce que l'on appelle *le corps de décomposition de* $X^q - X$ qui a les propriétés voulues par l'assertion (ii), y compris l'unicité. ■

Exemple Il n'y pas de corps à 3773 éléments.

Définition ("le" corps $\mathbb{F}_q$)
Si q est la puissance d'un nombre premier, on note $\mathbb{F}_q$ la classe d'isomorphisme des corps finis à q éléments, ou le plus souvent n'importe quel corps de cette classe.

Par exemple, on pourra noter $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ lorsque p est un nombre premier, ou $\mathbb{F}_4 = \mathbb{Z}/2\mathbb{Z}[X]/\left(X^2+X+1\right)$ ou encore $\mathbb{F}_{27} = \mathbb{Z}/3\mathbb{Z}[X]/\left(X^3-X+1\right)$, même si le statut de ces signes "=" peut être lu de plusieurs façons, dont l'ambiguïté n'est levée que par le contexte — souvent implicite — dans lequel on travaille.

Proposition (les groupes multiplicatifs des corps finis sont cycliques)
Soit q la puissance d'un nombre premier. Alors,
(i) le groupe $\left(\mathbb{F}_q^\times, \times\right)$ est cyclique
(ii) Lorsque $d|q-1$, l'unique sous-groupe d'ordre d de $\mathbb{F}_q^\times$ est l'ensemble des racines d^{e} de l'unité dans $\mathbb{F}$. Si ξ est un générateur de $\mathbb{F}_q^\times$, cet unique sous-groupe d'ordre d est engendré par $\xi^{\frac{q-1}{d}}$.

PREUVE. Pour tout diviseur d de $q-1$, on node $N(d)$ le nombre d'éléments d'ordre d du groupe $\mathbb{F}_q^\times$. Il s'agit de montrer que $N(q-1) \geq 1$: l'existence d'un élément d'ordre $q-1$ assurera que $\mathbb{F}_q^\times$ est cyclique.
Soit d un diviseur de $q-1$ tel que $N(d) \geq 1$. Soit alors $x \in \mathbb{F}_q^\times$, d'ordre d. Le groupe engendré par x est d'ordre d et est contenu dans l'ensemble des racines du polynôme $X^d - 1 \in \mathbb{F}_q[X]$. Or, ces dernières sont au plus au nombre de d. Donc le groupe engendré par x est exactement l'ensemble des racines d^{e} de l'unité de $\mathbb{F}$. En outre, il contient $\varphi(d)$ générateurs, comme tous les groupes cycliques d'ordre d. Donc $N(d) = \varphi(d)$.
On a montré que pour tout diviseur de $q-1$, $N(d) \in \{0, \varphi(d)\}$. Comme tout élément de $\mathbb{F}_q^\times$ a pour ordre un diviseur de $q-1$, on a l'égalité $q-1 = \sum_{d|q-1} N(d)$. Par ailleurs, $q-1 = \sum_{d|q-1} \varphi(d)$. Ces trois conditions entraînent que $N(d) = \varphi(d)$, pour tout diviseur d de $q-1$. En particulier, $N(q-1) = \varphi(q-1) \geq 1$. ■

Proposition (cardinaux des groupes linéaires sur des corps finis)
Soient q la puissance d'un nombre premier et $n \geq 1$. Alors,
(i) $|\mathrm{GL}(n, \mathbb{F}_q)| = (q^n - 1)(q^n - q)\left(q^n - q^2\right) \dots \left(q^n - q^{n-1}\right)$
(ii) $|\mathrm{SL}(n, \mathbb{F}_q)| = (q^n - 1)(q^n - q)\left(q^n - q^2\right) \dots \left(q^n - q^{n-2}\right) q^{n-1}$
(iii) $|\mathrm{PGL}(n, \mathbb{F}_q)| = |\mathrm{SL}(n, \mathbb{F}_q)|$
(iv) $|Z(\mathrm{SL}(n, \mathbb{F}_q))| = \mathrm{pgcd}(n, q-1)$
(v) $|\mathrm{PSL}(n, \mathbb{F}_q)| = \frac{|\mathrm{SL}(n, \mathbb{F}_q)|}{\mathrm{pgcd}(n, q-1)}$

PREUVE. (i) On considère $\mathrm{GL}(n, \mathbb{F}_q)$ comme le groupe des automorphismes linéaires de $\mathbb{F}_q^n$. Soit $\mathcal{C} = (e_1, \dots, e_n)$ la base canonique de $\mathbb{F}_q^n$. Choisir un élément de $\mathrm{GL}(n, \mathbb{F}_q)$ revient à choisir une base de $\mathbb{F}_q^n$. En effet, l'image de $\mathcal{C}$ par un élément de $\mathrm{GL}(n, \mathbb{F}_q)$ est une base de $\mathbb{F}_q^n$ et toute base de $\mathbb{F}_q^n$ est l'image de $\mathcal{C}$ par un unique élément de $\mathrm{GL}(n, \mathbb{F}_q)$. Il suffit donc de compter le nombre de bases de $\mathbb{F}_q^n$.
Or, en notant qu'un sous-espace vectoriel de diùension d de $\mathbb{F}_q^n$ a q^d éléments (il est isomorphe à $\mathbb{F}_q^d$), fabriquer une base $(v_1, \dots, v_n)$ de $\mathbb{F}_q^n$ consiste successivement à :
- choisir v_1 parmi les vecteurs non nuls : $q^n - 1$ choix possibles ;
- choisir v_2 dans $\mathbb{F}_q^n \setminus \mathrm{Vect}(v_1)$: $q^n - q$ choix possible ;
- choisir v_3 dans $\mathbb{F}_q^n \setminus \mathrm{Vect}(v_1, v_2)$: $q^n - q^2$ choix possible ;
- etc jusqu'au vecteur v_n qu'il faut choisir hors de l'hyperplan $\mathrm{Vect}(v_1, \dots, v_{n-1})$.

(ii) L'homomorphisme de groupes det : $\mathrm{GL}(n, \mathbb{F}_q) \to \mathbb{F}_q^\times$ a pour noyau $\mathrm{SL}(n, \mathbb{F}_q)$. En outre, il est surjectif puisque tout $x \in \mathbb{F}_q^\times$ est l'image par det de la matrice de dilatation $\mathrm{diag}(1, \dots, 1, x)$. Le premier théorème d'isomorphisme montre alors que det induit un isomorphisme de groupes $\mathrm{GL}(n, \mathbb{F}_q) / \mathrm{SL}(n, \mathbb{F}_q) \simeq \mathbb{F}_q^\times$, ce qui implique en particulier que $|\mathrm{GL}(n, \mathbb{F}_q)| = |\mathrm{SL}(n, \mathbb{F}_q)| \times \left|\mathbb{F}_q^\times\right|$. Cela prouve le résultat.
(iii) Le centre de $\mathrm{SL}(n, \mathbb{F}_q)$ est isomorphe au groupe des racines n^{e} de l'unité dans $\mathbb{F}_q$. On montre que ce groupe est d'ordre $d = \mathrm{pgcd}(n, q-1)$. Or, le groupe $\mathbb{F}_q^\times$ est cyclique d'ordre $q-1$. Il a donc un unique sous-groupe d'ordre d qui est exactement le groupe des racines d^{e} de l'unité dans $\mathbb{F}_q$. ■

A noter
Plus encore que le résultat lui-même, ce qu'il importe de retenir de l'énoncé précédent, c'est que le calcul de l'ordre de GL revient à calculer le nombre de bases de $\mathbb{F}_q^n$ et comment on mène ce calcul.

Exemple
L'ordre du groupe simple $\mathrm{PSL}(2, \mathbb{F}_5)$ est 60. On a déjà rencontré un autre groupe simple d'ordre 60, savoir $\mathfrak{A}_5$. On montrera que ces deux groupes sont isomorphes et, mieux encore, que tout groupe simple d'ordre 60 est isomorphe à $\mathfrak{A}_5$.

4.4 Le groupe orthogonal euclidien

Dans tout ce chapitre, le corps de base est $\mathbb{R}$.

En bref
Un espace vectoriel réel de dimension finie est dit *euclidien* lorsqu'on le munit d'un produit scalaire $\langle\cdot|\cdot\rangle$ — c'est-à-dire d'une forme bilinéaire symétrique définie positive. La *norme* associée à un produit scalaire est définie par $\|v\| = \sqrt{\langle v|v\rangle}$.
Le carré de la norme $q : v \mapsto \|v\|^2$ est une forme quadratique définie positive. Le produit scalaire, qui se retrouve à partir de la norme avec les formules standard $\langle v|w\rangle = \frac{1}{2}\left(q(v+w) - q(v) - q(w)\right) = \frac{1}{4}\left(q(v+w) - q(v-w)\right)$, est la *forme polaire* de la forme quadratique q. Ainsi, les données d'un produit scalaire ou d'une forme quadratique définie positive sont équivalentes.
Une *base orthonormée* d'un espace euclidien en est une base $(v_1, \ldots, v_n)$ formée de vecteurs unitaires et deux à deux orthogonaux : $\langle v_i|v_j\rangle = \delta_{i,j}$ (Kronecker). Il en existe toujours, comme l'assure l'algorithme d'orthonormalisation de Gram-Schmidt.
Si W est un sous-espace vectoriel d'un espace euclidien V, son orthogonal est l'ensemble des vecteurs orthogonaux à tous les vecteurs de W ; on le note $W^\perp = \{v \in V,\ \forall w \in W,\ \langle v|w\rangle = 0\}$. C'est un sous-espace supplémentaire de W : on a toujours $W \oplus W^\perp = V$.
Si W est un sous-espace vectoriel d'un espace euclidien V et si $v \in V$ se décompose en $v = w + w'$ où $w \in W$ et $w' \in W^\perp$, alors le vecteur w est le *projeté orthogonal de v sur W*. L'application $p_W : v \mapsto w$ est la *projection orthogonale sur W* ; elle est évidemment linéaire. Si $(v_1, \ldots, v_d)$ est une base orthonormée de W, le projeté orthogonal de $v \in V$ sur W est le vecteur

$$p_W(v) = \sum_{k=1}^{d} \langle v_k|v\rangle\, v_k.$$

En complétant la base $(v_1, \ldots, v_d)$ en une base orthonormée de V — c'est possible en combinant le théorème de la base incomplète et l'algorithme de Gram-Schmidt —, on obtient que $\|p_W(v)\| \leq \|v\|$.

Définition (isométrie, matrice orthogonale)
Soit V un espace euclidien. Un endomorphisme $f \in \mathrm{End}(V)$ est une *isométrie* de V lorsque f conserve la norme, c'est-à-dire lorsque $\|f(v)\| = \|v\|$, pour tout $v \in V$. On note $\mathrm{O}(V)$ l'ensemble des isométries de V.
Une matrice $M \in \mathcal{M}_n(\mathbb{R})$ est *orthogonale* lorsque $M\,{}^tM = I_n$. On note $\mathrm{O}(n)$ l'ensemble des matrices orthogonales $n \times n$.

Exercice 40
(i) f est une isométrie si, et seulement si elle conserve le produit scalaire, c'est-à-dire si, et seulement si $\langle f(v)|w\rangle = \langle v|w\rangle$, pour tous $v, w \in V$. Autre point de vue : si f est un endomorphisme, les assertions suivantes sont équivalentes :
(a) f est une isométrie ;
(b) f transforme toute base orthonormée en une base orthonormée ;
(c) il existe une base orthonormée que f transforme en une base orthonormée.

(ii) Soit $M \in \mathcal{M}_n(\mathbb{R})$. Les quatre assertions suivantes sont équivalentes :
(a) M est une matrice orthogonale
(b) ${}^tMM = I_n$
(c) les vecteurs-colonne de M forment une base orthonormée de $\mathcal{M}_{n,1}(\mathbb{R})$ pour le produit scalaire standard sur $\mathcal{M}_{n,1}(\mathbb{R})$, défini par $\langle X|Y\rangle = {}^tXY$
(d) les vecteurs-ligne de M forment une base orthonormée de $\mathcal{M}_{1,n}(\mathbb{R})$ pour le produit scalaire standard sur $\mathcal{M}_{1,n}(\mathbb{R})$, défini par $\langle X|Y\rangle = X{}^tY$.

(iii) Si f est une isométrie de V et si W est un sous-espace vectoriel de V stable par f, alors $W^\perp$ est un sous-espace de V stable par f qui vérifie $W \oplus W^\perp = V$. En outre, les endomorphismes de W et de $W^\perp$ induits par f sont aussi des isométries.
[C'est cette propriété de stabilité de l'orthogonal qui rend très facile la réduction des isométries ou des matrices orthogonales.]

(iv) Si f est une isométrie et si M est sa matrice (orthogonale) dans une base orthonormée, changer de base orthonormée revient à conjuguer la matrice de f par une matrice orthogonale. Autrement dit, les matrices de changement de bases orthonormées sont les matrices orthogonales.

(v) Muni de la composition des applications, $\mathrm{O}(V)$ est un sous-groupe de $\mathrm{GL}(V)$. Muni de la multiplication matricielle, $\mathrm{O}(n)$ est un sous-groupe de $\mathrm{GL}(n,\mathbb{R})$. Le choix d'une base orthonormée $\mathcal{B}$ de V induit un isomorphisme de groupes (non canonique : si on change de base, on change d'isomorphisme)

$$\begin{array}{ccc} \mathrm{O}(V) & \xrightarrow{\sim} & \mathrm{O}(n) \\ f & \longmapsto & \mathrm{Mat}_{\mathcal{B}}(f). \end{array}$$

(vi) Si ψ est une isométrie ou une matrice orthogonale, alors $\det\psi \in \{-1,1\}$. L'ensemble des isométries de déterminant 1 est un sous-groupe distingué de $\mathrm{O}(V)$. L'ensemble des matrices orthogonales de déterminant 1 est un sous-groupe distingué de $\mathrm{O}(n)$.

(vii) Si λ est une valeur propre réelle d'une isométrie, alors $\lambda \in \{-1,1\}$. Toute valeur propre complexe d'une isométrie est de module 1 (dans le cadre des matrices et des vecteurs-colonne, on pourra remarquer que si $MX = \lambda X$, alors $M\overline{X} = \overline{\lambda X}$ puisque M est une matrice réelle).

Vocabulaire
Le groupe $\mathrm{O}(V)$ est le *groupe orthogonal de V*. Le groupe $\mathrm{O}(n)$ est le *groupe orthogonal en dimension n* — ou *groupe orthogonal* tout court, la dimension n étant sous-entendue lorsque le contexte le permet. Une *rotation* de V est une isométrie dont le déterminant vaut 1 ; on dit aussi que c'est une *isométrie positive*. Une *matrice de rotation* est une matrice orthogonale de déterminant 1 ; on dit aussi que c'est une *matrice orthogonale positive*.

Notation
On note $\mathrm{SO}(V)$ le groupe des rotations de V et $SO(n)$ le groupe des matrices de rotations en dimension n. On note SO pour *groupe spécial orthogonal*.

A noter
Comme le déterminant est un homomorphisme surjectif de groupes $\mathrm{O}(V) \to \{-1,1\}$ dont le noyau est $SO(V)$,

$$[O(V) : \mathrm{SO}(V)] = 2 \text{ et } \quad [O(n) : \mathrm{SO}(n)] = 2.$$

Notation
Pour tout $\theta \in \mathbb{R}$, on note R_θ et S_θ les matrices orthogonales

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \in \mathrm{SO}(2) \text{ et } S_\theta = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} \in \mathrm{O}(2) \setminus \mathrm{SO}(2).$$

Exercice 41
(i) Pour tous $s,t \in \mathbb{R}$, $R_sR_t = R_{s+t}$, $S_sS_t = R_{s-t}$, $R_sS_t = S_{s+t}$ et $S_sR_t = R_{s-t}$.
(ii) Pour tout $\theta \in \mathbb{R}$, le polynôme caractéristique de R_θ est $X^2 - 2X\cos\theta + 1 = \left(X - e^{i\theta}\right)\left(X - e^{-i\theta}\right)$. En particulier, R_θ est diagonalisable (sur $\mathbb{R}$) si, et seulement si $R_\theta = \pm I_2$.
(iii) Pour tout $\theta \in \mathbb{R}$, S_θ est diagonalisable, semblable à $\mathrm{diag}(1,-1)$, et vérifie $S_\theta^2 = I_2$.

Proposition (classification des isométries en dimension 2)
(i) Pour tout $R \in \mathrm{SO}(2)$, il existe $\theta \in \mathbb{R}$ tel que $R = R_\theta$.
(ii) Pour tout $S \in \mathrm{O}(2) \setminus \mathrm{SO}(2)$, il existe $\theta \in \mathbb{R}$ tel que $S = S_\theta$.

Preuve. Soit $\begin{pmatrix} a & c \\ b & d \end{pmatrix} \in \mathrm{O}(2)$. Alors $a^2 + b^2 = 1$. Soit[♫] alors $\theta \in \mathbb{R}$ tel que $\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$. Comme le vecteur-colonne $\begin{pmatrix} c \\ d \end{pmatrix}$ est unitaire et orthogonal à $\begin{pmatrix} a \\ b \end{pmatrix}$, nécessairement, $\begin{pmatrix} c \\ d \end{pmatrix} = \pm\begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}$. On conclut selon que le déterminant de $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$ vaut 1 ou -1. ■

Proposition (le groupe $O(2)$)
(i) Le groupe $SO(2)$ est abélien, isomorphe au groupe multiplicatif S^1 des nombres complexes de module 1.
(ii) Pour tout $S \in \mathrm{O}(2) \setminus \mathrm{SO}(2)$, la paire $\{\mathrm{SO}(2), S \cdot \mathrm{SO}(2)\}$ est une partition de $\mathrm{O}(2)$.

[♫] L'existence d'un tel θ résulte des propriétés élémentaires de l'exponentielle. Voir par exemple les quatre premières pages du livre de W. Rudin : *Real and complex analysis*, ou sa traduction française si on préfère.

[On a noté $S \cdot \mathrm{SO}(2) = \{SR,\ R \in \mathrm{SO}(2)\}$.]

PREUVE. Que SO(2) soit abélien résulte de la formule d'addition des angles $R_s R_t = R_{s+t}$. Cette formule montre que l'application $\theta \mapsto R_\theta$ est un homomorphisme de groupes $(\mathbb{R}, +) \to (\mathrm{SO}(2), \times)$, surjectif d'après la classification des isométries en dimension 2. Son noyau est $2\pi\mathbb{Z}$, c'est encore une propriété des fonctions sinus et cosinus, elle-même héritée des propriétés de l'exponentielle. Le premier théorème d'isomorphisme induit alors un isomorphisme de groupes $\mathbb{R}/2\pi\mathbb{Z} \to \mathrm{SO}(2)$. Par ailleurs, l'exponentielle $\mathbb{R} \to S^1$, $\theta \mapsto e^{i\theta}$ induit aussi un isomorphisme de groupes $\mathbb{R}/2\pi\mathbb{Z} \to S^1$. La partition indiquée de O(2) est celle de ses classes modulo son sous-groupe distingué SO(2). ■

A noter

(i) L'isomorphisme de groupes de la preuve entre S^1 et SO(2) est l'application suivante, bien définie grâce au raisonnement tenu :

$$\begin{array}{ccc} S^1 & \xrightarrow{\sim} & \mathrm{SO}(2) \\ e^{i\theta} & \longmapsto & R_\theta. \end{array}$$

La structure de groupe commune sur S^1 ou sur SO(2) contient toutes le formules trigonométriques d'addition.

(ii) Lorsque $n \geq 3$, le groupe SO(n) n'est pas abélien : la commutativité de $SO(2)$ est une situation exceptionnelle. On peut la voir comme étant responsable de l'existence des angles orientés de vecteurs en dimension 2, notion qui disparaît en dimension supérieure.

Proposition (classification des isométries en dimension quelconque)

Soit V un espace euclidien de dimension $n \geq 2$ et f une isométrie de V. Il existe une base orthonormale de V dans laquelle la matrice de f s'écrit par blocs sous la forme

$$\begin{pmatrix} I_r & & & & & & \\ & -I_s & & & & & \\ & & R_{\theta_1} & & & & \\ & & & R_{\theta_2} & & & \\ & & & & \ddots & & \\ & & & & & R_{\theta_t} \end{pmatrix}$$

avec $r, s, t \in \mathbb{N}$, $\theta_1, \ldots, \theta_t \in \mathbb{R}$.

PREUVE. On procède par récurrence sur n. Pour $n \geq 2$, c'est fait. On suppose que $n \geq 3$. Si 1 ou -1 est valeur propre de f, puisque le supplémentaire orthogonal $W^\perp$ du sous-espace stable non nul $W = \ker(f \pm \mathrm{id}_V)$ est aussi stable par f, il suffit de mettre bout à bout des bases orthonormales de W et, par récurrence, de $W^\perp$ pour obtenir la base cherchée. Si ni 1 ni -1 n'est valeur propre de f, alors toutes les valeurs propres de f sont des nombres complexes non réels (de module 1). Là encore, il suffit de trouver un plan de V stable par f pour pouvoir appliquer l'hypothèse de récurrence à son supplémentaire orthogonal et conclure. On se place dans un cadre matriciel : soit $M \in \mathrm{O}(n)$ la matrice de f dans une base orthonormée quelconque ; ses valeurs propres sont des nombres complexes non réels. Puisque $\mathbb{C}$ est algébriquement clos, de telles valeurs propres existent. Soient $\zeta \in S^1 \setminus \{-1, 1\}$ et $X \in \mathcal{M}_{n,1}(\mathbb{C})$ tels que $MX = \zeta X$ et $X \neq 0$. Comme ζ n'est pas reélle, $X \notin \mathcal{M}_{n,1}(\mathbb{R})$. Puisque M est une matrice réelle, il en résulte que $M\overline{X} = \overline{\zeta} \cdot \overline{X}$, où $\overline{X}$ désigne le vecteur-colonne dont les coordonnées sont les conjuguées des coordonnées de X. Alors, les vecteurs $X + \overline{X}$ et $X - \overline{X}$ ont des coordonnées réelles et sont linéairement indépendants : ils engendrent un plan vectoriel de $\mathcal{M}_{n,1}(\mathbb{R})$, stable par M. C'est ce que l'on cherchait. ■

Définition (réflexions et renversements)

Une isométrie est une *réflexion* lorsque sous sous-espace des vecteurs fixes est un hyperplan. L'hyperplan $\ker(f - \mathrm{id}_V)$ est appelé *hyperplan de la réflexion*. Une isométrie f d'un espace euclidien de dimension n est un *renversement* lorsque $\dim\ker(f - \mathrm{id}_V) = n - 2$ et $\dim\ker(f + \mathrm{id}_V) = 2$.

Autrement dit, les réflexions et les renversements sont les isométries qui admettent respectivement pour matrices par blocs, dans une base orthonormée convenable,

$$\begin{pmatrix} I_{n-1} & \\ & -1 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} I_{n-2} & \\ & -I_2 \end{pmatrix}.$$

A noter

(i) Les renversements sont des isométries positives. Les réflexions sont des isométries négatives. En dimension 2 et en dimension 2 seulement, la réciproque est vraie : toute isométrie négative est une réflexion.

(ii) Si f est une réflexion ou un renversement de V, alors f est une involution : $f^2 = \mathrm{id}_V$.

Exercice 42

Dans le groupe orthogonal, toute conjuguée d'une réflexion est une réflexion et toute conjuguée d'un renversement est un renversement.

Plus précisément, si $s \in \mathrm{O}(V)$ est la réflexion d'hyperplan H et si $g \in \mathrm{O}(V)$, alors gsg^{-1} est la réflexion d'hyperplan $g(H)$. De même, si r est le renversement dont l'espace des points fixes est le sous-espace W de codimension 2, alors grg^{-1} est le renversement dont l'espace des points fixes est $g(W)$.

Proposition (centres de O et de SO)

(i) Pour tout $n \geq 2$, le centre de $\mathrm{O}(n)$ est $\{-I_n, I_n\}$.

(ii) Pour tout $n \geq 3$, le centre de $\mathrm{SO}(n)$ est $\{-I_n, I_n\}$ ou $\{I_n\}$, selon que n est respectivement pair ou impair.

Preuve. (i) Soit V un espace euclidien de dimension n. Soit f une isométrie du centre de $\mathrm{O}(V)$. Si s est n'importe quelle réflexion dont la droite des points fixes est D, alors fsf^{-1} est encore une réflexion dont la droite des points fixes est $f(D)$. Comme $fsf^{-1} = s$ puisque f est central, on en déduit que $f(D) = D$. On a montré que f stabilise toutes les droites vectorielles de V, ce qui entraîne que f est une homothétie, comme dans le calcul du centre de $\mathrm{GL}(V)$. Or, les seules homothéties de $\mathrm{O}(V)$ sont $\pm\,\mathrm{id}_V$.

(ii) Soit V un espace euclidien de dimension n. Soit f une isométrie du centre de $\mathrm{SO}(V)$. Si r est n'importe quel renversement dont le plan des points fixes est P, alors frf^{-1} est encore un renversement dont le plan des points fixes est $f(P)$. Comme $frf^{-1} = r$ puisque f est central, on en déduit que $f(P) = P$. On a montré que f stabilise tous les plans vectoriels de V. Or, puisque $n \geq 3$, toute droite est intersection de deux plans. Donc f stabilise toutes les droites : c'est une homothétie. ■

Exercice 43

Montrer que l'application $\{-I_3, I_3\} \times \mathrm{SO}(3) \xrightarrow{\sim} \mathrm{O}(3)$, $(\varepsilon I_3, M) \longmapsto \varepsilon M$ est un isomorphisme de groupes. Montrer que cet exemple se généralise à des isomorphismes de groupes $\mathrm{O}(2n+1) \simeq \{\pm 1\} \times \mathrm{SO}(2n+1)$, pour tout $n \in \mathbb{N}$, mais que les groupes $\mathrm{O}(2n)$ et $\{\pm 1\} \times \mathrm{SO}(2n)$ ne sont jamais isomorphes (on pourra considérer leurs centres).

Proposition (les réflexions engendrent le groupe orthogonal)

Soit V un espace euclidien de dimension n. Alors, toute isométrie de V est produit d'au plus n réflexions.

Preuve. Soit $f \in \mathrm{O}(V)$. On note $\mathrm{Fix}(f) = \ker(f - \mathrm{id}_V)$ le sous-espace des vecteurs fixes de f. On montre par récurrence (forte) sur $\operatorname{codim}\mathrm{Fix}(f) = n - \dim\mathrm{Fix}(f)$ que f est produit d'au plus $\operatorname{codim}\mathrm{Fix}(f)$ réflexions, ce qui est plus fort que le résultat annoncé. Si $\operatorname{codim}\mathrm{Fix}(f) = 0$, alors $f = \mathrm{id}_V$ est produit de 0 réflexions (l'énoncé de la proposition suggère cette convention de façon implicite, l'ajouter s'il faut pour lever l'ambiguïté). On suppose que $\operatorname{codim}\mathrm{Fix}(f) \geq 1$. Soient alors $x \in \mathrm{Fix}(f)^\perp \setminus \{0\}$ et $y = f(x)$. Puisque $x \notin \mathrm{Fix}(f)$, $x - y \neq 0$. En outre, puisque f est une isométrie, $\|x\|^2 = \|y\|^2$, ce qui entraîne que $x - y \perp x + y$. Soit r la réflexion d'hyperplan $H = \mathrm{Vect}(x-y)^\perp$. Alors, $x + y \in \mathrm{Fix}(r)$ et $x - y \in H^\perp$. Ainsi, $\mathrm{Fix}(f) \subseteq \mathrm{Vect}(x-y)^\perp = H = \mathrm{Fix}(r)$. Cela entraîne immédiatement que $\mathrm{Fix}(f) \subseteq \mathrm{Fix}(rf)$. Or, $r(y) = x$ comme le montrent les égalités $r(x+y) = x + y$ et $r(x-y) = -x + y$. Ainsi, $x \in \mathrm{Fix}(rf) \setminus \mathrm{Fix}(f)$. Il en résulte que $\operatorname{codim}\mathrm{Fix}(rf) < \operatorname{codim}\mathrm{Fix}(f)$. Par récurrence, rf est produit d'au plus $\operatorname{codim}\mathrm{Fix}(rf)$ réflexions, ce qui entraîne que f est produit d'au plus $\operatorname{codim}\mathrm{Fix}(rf) + 1 \leq \operatorname{codim}\mathrm{Fix}(f)$ réflexions, puisque les réflexions sont des involutions. ■

Exercice 44

Faire une autre preuve du même résultat en adoptant un point de vue matriciel et en utilisant le théorème de classification des isométries en dimension quelconque. On pourra aussi s'appuyer sur les formules de multiplication entre les R_θ et les S_θ établies dans l'exercice qui suit leurs définitions.

Proposition (les renversements engendrent SO en dimension ≥ 3)

Soit V un espace euclidien de dimension $n \geq 3$. Alors, toute rotation de V est produit d'au plus n renversements.

Preuve. Soit $f \in \mathrm{SO}(V)$.

On commence par le cas $n = 3$. Dans ce cas, f est produit d'au plus 3 réflexions. Comme f est positive, à moins d'être égale à l'identité, f est produit de 2 réflexions $f = r_1 r_2$. En remarquant que, en dimension 3, si r est une réflexion, alors $-r$ est un renversement, l'égalité $f = (-r_1)(-r_2)$ permet de conclure.

On suppose $n \geq 4$. Comme f est produit d'un nombre pair de réflexions, il suffit de montrer que tout produit de 2 réflexions est un produit de 2 renversements. Soient dont r_1 et r_2 deux réflexions, d'hyperplans respectifs H_1 et H_2. Alors, $(H_1 \cap H_2)^\perp$ est de dimension 1 ou 2. Puisque $n \geq 3$, soit W, sous-espace de dimension 3 de V contenant $(H_1 \cap H_2)^\perp$. Alors, $W^\perp$ est contenu dans $H_1 \cap H_2$, ce qui implique que $W^\perp$ est constitué de vecteurs fixes de $r_1 r_2$. Donc W est un espace de dimension 3, stable par $r_1 r_2$. Alors, en notant encore r_1 et r_2 les endomorphismes de W induits par r_1 et r_2, il existe deux renversements s_1 et s_2 de W tels que $r_1 r_2 = s_1 s_2$, comme l'assure l'étude préalable de la dimension 3. En prolongeant s_1 et s_2 à $V = W \oplus W^\perp$ par l'identité sur $W^\perp$, on obtient encore des renversements s_1 et s_2 qui vérifient l'égalité $r_1 r_2 = s_1 s_2$: on a écrit $r_1 r_2$ comme un produit de deux renversements. ■

Définition (groupes projectifs orthogonaux)
On note $\mathrm{PO}(V) = \mathrm{O}(V)/\{-1, 1\}$ et $PSO(V) = \mathrm{SO}(V)/Z(\mathrm{SO}(V))$ le *groupe projectif orthogonal* et le *groupe projectif spécial orthogonal* de V. De même, si $n \geq 2$, on note $\mathrm{PO}(n) = \mathrm{O}(n)/\{\pm I_n\}$ et $\mathrm{PSO}(n) = \mathrm{SO}(n)/Z(\mathrm{SO}(n))$.
Là encore, ces quotients sont pris à leur pleine mesure lorsqu'on les considère comme des transformations de la droite projective $\mathbb{P}(V)$. Bien sûr, lorsque n est impair, $\mathrm{PSO}(n) = \mathrm{SO}(n)$ puisque les centres sont alors triviaux.

Théorème (simplicité des groupes projectifs spéciaux orthogonaux)
(i) Le groupe $\mathrm{SO}(3)$ *est simple.*
(ii) $\mathrm{PSO}(4) \simeq \mathrm{SO}(3) \times \mathrm{SO}(3)$ *n'est pas simple.*
(iii) Pour $n \geq 5$*, le groupe* $\mathrm{PSO}(n)$ *est simple.*

Preuve. Voir Perrin page 150. Le cas exceptionnel de PSO(4) est à relier à la géométrie euclidienne de $\mathbb{R}^3$ et à l'étude du corps gauche des quaternions de Hamilton $\mathbb{H}(\mathbb{R})$. ■

4.5 Un tout petit peu sur le groupe modulaire

Proposition (matrices inversibles sur $\mathbb{Z}$)
Soit $n \geq 1$*. Une matrice de* $M \in \mathcal{M}_n(\mathbb{Z})$ *est inversible dans* $\mathcal{M}_n(\mathbb{Z})$ *si, et seulement si* $\det(M) \in \{-1, 1\}$.

Preuve. On utilise la formule $M \times {}^t\mathrm{Com}\, M = {}^t\mathrm{Com}\, M \times M = \det(M) I_n$. Puisque les coefficients de la comatrice générique sont des polynômes à coefficients entiers, $\mathrm{Com}(M) \in \mathcal{M}_n(\mathbb{Z})$. Si $\det(M) = \pm 1$, alors M admet $\pm {}^t\mathrm{Com}\, M$ pour inverse dans $\mathcal{M}_n(\mathbb{Z})$. Inversement, si M est inversible dans $\mathcal{M}_n(\mathbb{Z})$, il existe $N \in \mathcal{M}_n(\mathbb{Z})$ tel que $MN = I_n$. Alors, $\det(M)\det(N) = 1$ ce qui impose que $\det(M)$ soit inversible dans $\mathbb{Z}$. ■

Définition (groupes linéaires sur $\mathbb{Z}$)
On note $\mathrm{GL}(n, \mathbb{Z})$ le groupe des matrices inversibles de $\mathcal{M}_n(\mathbb{Z})$, c'est-à-dire le groupe des matrices à coefficients entiers dont le déterminant vaut ± 1. On note $\mathrm{SL}(n, \mathbb{Z})$ son sous-groupe des matrices de déterminant 1.

A noter
Puisque c'est le noyau du déterminant, $\mathrm{SL}(n, \mathbb{Z}) \triangleleft \mathrm{GL}(n, \mathbb{Z})$.

Notations
On notera S et T les éléments suivants du groupe $\mathrm{SL}(2, \mathbb{Z})$:

$$S = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{et} \quad T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

Exercice 45 Dans le groupe $\mathrm{SL}(2, \mathbb{Z})$, S est d'ordre 4 et T d'ordre infini.

Proposition (centre de $\mathrm{SL}(2, \mathbb{Z})$)
Le centre de $\mathrm{SL}(n, \mathbb{Z})$ *est* $\{\pm I_2\}$.

Preuve. Si M est dans le centre de $\mathrm{SL}(n, \mathbb{Z})$, alors M commute avec S et T. Donc $M = \pm I_2$, comme le montre un calcul élémentaire. ■

Définition (groupe modulaire)

On appelle *groupe modulaire* le groupe $\mathrm{PSL}(2,\mathbb{Z}) = \mathrm{SL}(2,\mathbb{Z}) / \{-I_2, I_2\}$. On note s et t les classes respectives de S et T dans le quotient $\mathrm{SL}(2,\mathbb{Z}) / \{-I_2, I_2\}$.

Exercice 46 Dans le groupe modulaire, s est d'ordre 2 et t est d'ordre infini.

Théorème (s et t engendrent le groupe modulaire)
(i) Le groupe $\mathrm{SL}(2,\mathbb{Z})$ *est engendré par* S *et* T.
(ii) Le groupe modulaire est engendré par s *et* t.

Preuve. Il suffit de montrer (i). On note G le sous-groupe de $\mathrm{SL}(2,\mathbb{Z})$ engendré par S et T. Soit $M = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in \mathrm{SL}(2,\mathbb{Z})$. Il s'agit de montrer que $M \in G$. On regarde d'abord l'effet de la multiplication à gauche de M par S, S^3 et T^q, où $q \in \mathbb{Z}$:

$$T^q M = \begin{pmatrix} a+bq & c+dq \\ b & d \end{pmatrix},\ SM = \begin{pmatrix} b & d \\ -a & -c \end{pmatrix} \text{ et } S^3 M = \begin{pmatrix} -b & -d \\ a & c \end{pmatrix}$$

Si $b = 0$, alors $M = T^c \in G$ ou $M = S^2 T^{-c} \in G$, selon le signe commun de $a = d = \pm 1$; donc $M \in G$. Si $a = 0$, on se ramène au cas $b = 0$ en remplaçant M par SM, ce qui montre que $M \in G$.
On suppose ainsi que $ab \neq 0$. La relation $\det M = 1$, qui est une relation de Bézout entre a et b, montre que ces derniers sont premiers entre eux. On adapte l'algorithme d'Euclide appliqué à a et b de la façon suivante : On procède à la suite de divisions euclidiennes

$$\begin{cases} a = bq_0 + r_0 \ \text{ où } \ q_0 \in \mathbb{Z} \ \text{ et } \ 0 \leq r_0 \leq |b| - 1 \\ -b = r_0 q_1 + r_1 \ \text{ où } \ q_1 \in \mathbb{Z} \ \text{ et } \ 0 \leq r_1 \leq r_0 - 1 \\ -r_0 = r_1 q_2 + r_2 \ \text{ où } \ q_2 \in \mathbb{Z} \ \text{ et } \ 0 \leq r_2 \leq r_1 - 1 \\ -r_1 = r_2 q_3 + r_3 \ \text{ où } \ q_3 \in \mathbb{Z} \ \text{ et } \ 0 \leq r_3 \leq r_2 - 1 \\ \qquad \vdots \\ -r_{m-2} = r_{m-1} q_m + r_m \ \text{ où } \ q_m \in \mathbb{Z} \ \text{ et } \ 0 \leq r_m \leq r_{m-1} - 1 \\ -r_{m-1} = r_m q_{m+1} + 1 \ \text{ où } \ q_{m+1} \in \mathbb{Z} \end{cases}$$

où r_m est le dernier reste strictement supérieur à 1. Transposé en termes matriciels, en notant $\star$ un nombre entier dont on n'a pas besoin d'expliciter la valeur en fonction des données, cela s'écrit successivement $T^{-q_0} M = \begin{pmatrix} r_0 & \star \\ b & \star \end{pmatrix}$, $S^3 T^{-q_0} M = \begin{pmatrix} -b & \star \\ r_0 & \star \end{pmatrix}$, $S^3 T^{-q_1} S^3 T^{-q_0} M = \begin{pmatrix} -r_0 & \star \\ r_1 & \star \end{pmatrix}$, $S^3 T^{-q_2} S^3 T^{-q_1} S^3 T^{-q_0} M = \begin{pmatrix} -r_1 & \star \\ r_2 & \star \end{pmatrix}$, jusqu'à la dernière ligne qui montre qu'il existe $h \in G$ tel que $hM = \begin{pmatrix} -r_m & \star \\ 1 & \star \end{pmatrix}$. La dernière opération $T^{r_m} hM = \begin{pmatrix} 0 & \star \\ 1 & \star \end{pmatrix}$ ramène l'affaire au cas où $a = 0$ et permet de conclure qu'il existe $g \in G$ tel que $gM \in G$. La multiplication à gauche par g^{-1} montre alors que $M \in G$. ∎

Exercice 47
(i) On note $T' = STS^{-1}$ et $t' = sts$. Calculer T', calculer l'ordre de t' et montrer que le groupe modulaire est engendré par t et t'.
(ii) On note $U = TS$ et $u = ts$. Calculer U, calculer l'ordre de u et montrer que le groupe modulaire est engendré par t et u.

5 Action d'un groupe sur un ensemble

5.1 Généralités, premiers exemples

C'est au travers de la notion d'action d'un groupe sur un ensemble que la structure de groupe prend tout son sens et montre son efficacité opératoire. Il faut mentionner à cet endroit l'œuvre de Felix Klein[♫], synthétisée dans son *Programme d'Erlangen*, qui donne une définition définitive à la *géométrie* en mathématiques : étudier la géométrie d'un objet, c'est le considérer comme subissant l'action d'un groupe, que l'on regarde ainsi comme un groupe de transformations de l'objet lui-même.
Ainsi, par exemple, l'objet $\mathbb{R}^3$, n'a pas la même géométrie selon le groupe de transformations (ici "naturelles") qu'on s'autorise. Faire de la topologie, c'est faire agir le groupe des homéomorphismes de $\mathbb{R}^3$ sur lui-même. Faire du calcul différentiel, c'est faire agir le groupe des difféomorphismes de $\mathbb{R}^3$ sur lui-même. Faire de la géométrie vectorielle (ou affine en ajoutant les translations), c'est faire agir le groupe $\mathrm{GL}\left(\mathbb{R}^3\right)$. Faire de la géométrie (vectorielle) euclidienne, c'est faire agir le groupe $\mathrm{O}\left(\mathbb{R}^3\right)$, etc.

Définition (action à gauche)
Soient G un groupe et X un ensemble non vide. Une *action à gauche de G sur X* est un homomorphisme de groupes $\varphi : G \to \mathfrak{S}_X$. On dit aussi *opération à gauche*.

A noter
Une action φ comme ci-dessus étant donnée, on note le plus souvent $g \cdot x = \varphi(g)(x)$, pour tous $g \in G$ et $x \in X$ — parfois, on enlève le point et on note simplement gx. Avec ces notations, le fait que φ soit un homomorphisme de groupes implique immédiatement que

$$\begin{cases} \forall x \in X,\ 1_G \cdot x = x \\ \forall g, g' \in G,\ \forall x \in X,\ g \cdot (g' \cdot x) = (gg') \cdot x. \end{cases} \tag{13}$$

Inversement, toute application $G \times X \to X$ notée $(g, x) \mapsto g \cdot x$ et qui vérifie les deux axiomes (13) définit une action de G sur X *via* l'application $G \to \mathfrak{S}_X$, $g \mapsto g\cdot$, où $g\cdot$ est l'application $X \to X$, $x \mapsto g \cdot x$.

Exercice 48
Ecrire tous les détails de ce qu'affirme le *à noter* ci-dessus. En particulier, s'assurer de bien comprendre la nécessité d'ajouter l'axiome $1 \cdot x = x$ pour obtenir une équivalence.

Définition (action à droite)
Une *action à droite* d'un groupe G sur un ensemble X est une application $X \times G \to G$ qui vérifie :

$$\begin{cases} \forall x \in X,\ x \cdot 1_G = x \\ \forall g, g' \in G,\ \forall x \in X,\ (x \cdot g)\, g' = x \cdot (gg')\,. \end{cases}$$

Exercice 49
(i) Montrer que la donnée d'une action à droite est équivalente à la donnée d'une application $\psi : G \to \mathfrak{S}_X$ qui vérifie : $\psi(gg') = \psi(g')\psi(g)$, pour tous $g, g' \in G$.
(ii) Si φ est une action à droite de G sur X, on obtient une action à gauche de G sur X en posant $\psi(g) = \varphi\left(g^{-1}\right)$.

A noter
Dans ce cours, lorsqu'on parle d'*action d'un groupe sur un ensemble* sans spécifier s'il s'agit d'une action à droite ou à gauche, c'est d'une action à gauche qu'il s'agit.

Exemples
(i) Si X est un ensemble et n un entier naturel, on note $\mathcal{P}_n(X)$ l'ensemble des parties de cardinal n de X. Le groupe $\mathfrak{S}_X$ agit sur $\mathcal{P}_n(X)$ par son *action naturelle*, définie par

$$\forall \sigma \in \mathfrak{S}_X,\ \forall Y \in \mathcal{P}_n(X),\ \sigma \cdot Y = \sigma(Y).$$

En effet, si $\sigma \in \mathfrak{S}_X$ et si $Y \in \mathcal{P}_n(X)$, alors $\sigma(Y)$ est encore dans $\mathcal{P}_n(X)$ puisque σ est une bijection $X \to X$. Les axiomes d'action à gauche sont immédiatement vérifiés.

[♫] Felix Klein, 1849 – 1925

(ii) Soit V un espace vectoriel. Pour tout entier naturel d, on note $\mathcal{G}_d(V)$ l'ensemble des sous-espaces vectoriels de dimension d de V. Le groupe $\mathrm{GL}(V)$ agit sur $\mathcal{G}_d(V)$ par son *action naturelle*, définie par

$$\forall g \in \mathrm{GL}(V),\ \forall W \in \mathcal{G}_d(V),\ g \cdot W = g(W).$$

En effet, si $g \in \mathrm{GL}(V)$ et si $W \in \mathcal{G}_d(V)$, alors $g(W)$ est encore dans $\mathcal{G}_d(V)$ puisque g est une application linéaire bijective. Les axiomes d'action à gauche sont immédiatement vérifiés.

(iii) Si G est un groupe, G agit sur lui-même *par translation à gauche* : il s'agit de l'opération

$$\begin{array}{ccc} G \times G & \longrightarrow & G \\ (g,h) & \longmapsto & g \cdot h = gh. \end{array}$$

Elle est aussi décrite par l'homorphisme de groupes $\varphi : G \to \mathfrak{S}_G$, défini par $\forall g, h \in G$, $\varphi(g)(h) = gh$. Noter que sa sœur, l'action à droite par translation définie par $(g,h) \mapsto hg$ est une action à droite.

(iv) Si G est un groupe, G agit sur lui-même *par conjugaison* : il s'agit de l'opération

$$\begin{array}{ccc} G \times G & \longrightarrow & G \\ (g,h) & \longmapsto & g \cdot h = ghg^{-1}. \end{array}$$

Elle est aussi décrite encore par l'homorphisme de groupes $\varphi : G \to \mathrm{Aut}(G) \subseteq \mathfrak{S}_G$, défini par $\forall g, h \in G$, $\varphi(g)(h) = ghg^{-1}$. Pour tout $g \in G$, la bijection $\varphi(g)$ est l'automorphisme intérieur $h \mapsto ghg^{-1}$ déjà rencontré. Là encore, sa sœur qu'est la conjugaison dans l'autre sens $(h,g) \mapsto g^{-1}hg$ définit une action à droite.

(v) Le groupe $\mathrm{SL}(2,\mathbb{R})$ agit sur le *demi-plan de Poincaré*

$$\mathfrak{H} = \{z \in \mathbb{C},\ \Im(z) > 0\}$$

par *homographies*. Il s'agit de l'action définie par

$$\forall \begin{pmatrix} a & b \\ c & d \end{pmatrix},\ \forall z \in \mathfrak{H},\ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot z = \frac{az+b}{cz+d}.$$

Il s'agit de vérifier que si $M \in \mathrm{SL}(2,\mathbb{R})$ et si $z \in \mathfrak{H}$, le nombre complexe $M \cdot z$ ainsi défini est bien dans $\mathfrak{H}$, ce qui est garanti par le calcul

$$\Im\left(\frac{az+b}{cz+d}\right) = \Im\left(\frac{(az+b)\,(c\overline{z}+d)}{|cz+d|^2}\right) = \Im\left(\frac{adz+bc\overline{z}}{|cz+d|^2}\right) = \frac{\Im(z)}{|cz+d|^2},$$

la dernière égalité venant du fait que le déterminant de la matrice égale 1. Enfin, les axiomes de l'action à gauche sont immédiatement vérifiées par un calcul élémentaire qui revient à simplifier la fraction

$$\frac{a'\frac{az+b}{cz+d} + b'}{c'\frac{az+b}{cz+d} + d'} = \frac{(a'a+b'c)\,z + (a'b+b'd)}{(c'a+d'c)\,z + (c'b+d'd)}.$$

(vi) D'autre exemples en vrac d'*actions naturelles* :
- Le groupe $\mathrm{O}(2)$ agit sur le cercle unité $\{v,\ \|v\| = 1\}$, le groupe $\mathrm{O}(3)$ agit sur la sphère unité
- Le groupe des isométries du plan qui préservent un polygone régulier agit sur l'ensemble des milieux des arêtes
- Le groupe des isométries de l'espace qui préservent un cube agit sur ses quatre diagonales

(vii) Soit X un espace topologique — par exemple, un espace métrique. On note $\mathrm{Aut}(X)$ le groupe des homéomorphismes $X \to X$ pour la composition des applications. Un *lacet tracé sur* X est une application continue $\ell : [0,1] \to X$ qui vérifie $\ell(0) = \ell(1)$. Le groupe $\mathrm{Aut}(X)$ agit sur l'ensemble $\mathcal{L}(X)$ des lacets tracés sur X par l'action $\mathrm{Aut}(X) \times \mathcal{L}(X) \to \mathcal{L}(X)$, $(f,\ell) \mapsto f \circ \ell$.

Définitions (orbite, stabilisateur)
Soit G un groupe agissant sur un ensemble X. Pour tout $x \in X$, l'*orbite de x sous l'action de G* , notée $G \cdot x$ ou Gx est la partie de X

$$Gx = \{gx,\ g \in G\}$$

et le *stabilisateur* ou encore le *groupe d'isotropie de x sous l'action de G* est le sous-groupe de G, noté G_x ou $\mathrm{Stab}(x)$, défini par

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

Exercice 50 Montrer que G_x est bien un sous-groupe de G.

Exemples

(i) Soit c un p-cycle de $\in \mathfrak{S}_n$. On fait agir le groupe cyclique $\langle c\rangle$ sur l'ensemble $\{1,\dots,n\}$ par son action naturelle. Si $x \in \{1,\dots,n\}$, l'orbite de x sous $\langle c\rangle$ est le support de c.
De même, si $\sigma \in \mathfrak{S}_n$ est une permutation quelconque et si $x \in \{1,\dots,n\}$, l'orbite de x sous l'action naturelle de $\langle\sigma\rangle$ est le support du cycle qui contient x (dans son support) dans la décomposition de σ en produit de cycles à supports disjoints.

(ii) Si $n \geq 2$, on fait encore agir le groupe $\mathfrak{S}_n$ sur $\{1,\dots,n\}$ par son action naturelle. Si $x \in \{1,\dots,n\}$, le stabilisateur de x est le sous-groupe des permutations de $\{1,\dots,n\}$ qui fixent x. Il est isomorphe à $\mathfrak{S}_{n-1}$. En effet, si on note $\overline{x} = \{1,\dots,n\} \setminus \{x\}$, alors l'application

$$\begin{array}{ccc} \mathfrak{S}_{\overline{x}} & \xrightarrow{\Phi} & \mathrm{Stab}(x) \\ s & \longmapsto & \begin{cases} y \mapsto s(y) \text{ si } y \neq x \\ x \mapsto x. \end{cases} \end{array}$$

est un isomorphisme de groupes ; en outre, les groupes $\mathfrak{S}_{\overline{x}}$ et $\mathfrak{S}_{n-1}$ sont isomorphes puisque l'ensemble $\overline{x}$ a pour cardinal $n-1$. Noter que pour prouver que Φ est une bijection, on peut exhiber sa réciproque de la façon suivante. Si $t \in \mathrm{Stab}(x)$, alors t stabilise aussi le complémentaire $\overline{x}$ de x ; autrement dit, $t(\overline{x}) \subseteq \overline{x}$ — cette inclusion est à vrai dire une égalité, par considération sur les cardinaux finis. En particulier, la restriction de t à $\overline{x}$ induit une permutation $\overline{x}$ que l'on note t_x. La réciproque de Φ est l'application $\mathrm{Stab}(x) \to \mathfrak{S}_{\overline{x}}$, $t \mapsto t_x$.

(iii) Soit V un espace vectoriel. On note $\mathcal{G}_1(V)$ l'ensemble des droites de V. On fait agir $\mathrm{GL}(V)$ sur $\mathcal{G}_1(V)$ par son action naturelle. Si $D \in \mathcal{G}_1(V)$, le stabilisateur de D est l'ensemble des applications linéaires bijectives $V \to V$ pour lesquelles D est une droite de vecteurs propres.

(iv) Soit $\mathcal{C} = [-1,1]^3$ le cube de l'espace euclidien standard $\mathbb{R}^3$. L'ensemble $\mathrm{O}(\mathcal{C}) = \left\{f \in \mathrm{O}\left(\mathbb{R}^3\right),\ f(\mathcal{C}) \subseteq \mathcal{C}\right\}$ est un sous-groupe de $\mathrm{O}\left(\mathbb{R}^3\right)$ — c'est lui-même un stabilisateur pour une certaine action de $\mathrm{O}\left(\mathbb{R}^3\right)$; exercice : laquelle ? On fait agir le groupe $\mathrm{O}(\mathcal{C})$ sur l'ensemble des point de $\mathcal{C}$ par son action naturelle $(f,x) \mapsto f(x)$. Alors, l'orbite d'un sommet de $\mathcal{C}$ par cette action est l'ensemble de tous les sommets de $\mathcal{C}$. Pour voir cela, il suffit de considérer l'action des rotations de $\mathrm{O}(\mathcal{C})$ dont les axes passent par les centres des faces.

Proposition (les orbites forment une partition)
Soit G un groupe agissant sur un ensemble X. La relation binaire sur X définie par : $x \sim y \iff \exists g \in G,\ y = gx$ est une relation d'équivalence. Ses classes sont les orbites sous l'action de G.

Preuve. C'est immédiat, conséquence directe des axiomes de l'action. ■

Proposition (les stabilisateurs de deux points d'une même orbite sont conjugués)
Soit G un groupe agissant sur un ensemble X. Pour tous $x \in X$ et $g \in G$, les groupes d'isotropie de x et de gx sont conjugués. Plus précisément,

$$G_{gx} = gG_xg^{-1}.$$

Preuve. Si $h \in G_x$, alors $ghg^{-1} \cdot gx = ghx = gx$. Ainsi, $gG_xg^{-1} \subseteq G_{gx}$. En conjuguant par g^{-1}, cela entraîne que $G_x \subseteq g^{-1}G_{gx}g$. On applique cette dernière formule en remplaçant x par gx et g par g^{-1} ; on obtient $G_{gx} \subseteq gG_xg^{-1}$, ce qui achève de montrer l'égalité cherchée. ■

Définitions (actions transitives, fidèles)
Soit G un groupe agissant sur un ensemble X. On dit que l'action est *transitive*, ou que G *agit transitivement* lorsqu'elle n'a qu'une seule orbite. Autrement dit, lorsque $\forall x, y \in X,\ \exists g \in G,\ y = gx$.
On dit que l'action est *fidèle*, ou que G *agit fidèlement* lorsque seul 1_G fixe tous les élements de X. Autrement dit, l'action $G \to \mathfrak{S}_X$ est fidèle lorsqu'elle est injective, ou encore lorsque $\forall g \in G,\ \left(\forall x \in X,\ gx = x\right) \Longrightarrow \left(g = 1_G\right)$.

Exemples

(i) L'action naturelle de $\mathrm{GL}(V)$ sur les droites de V est transitive, mais pas fidèle — sauf si le corps de base est $\mathbb{Z}/2\mathbb{Z}$.

(ii) Si $p \in \{1, \dots, n\}$, l'action naturelle de $\mathfrak{S}_n$ sur les parties à p éléments de $\{1, \dots, n\}$ est transitive et fidèle.

(iii) L'action de $\mathfrak{S}_3$ par conjugaison sur ses sous-groupes n'est pas transitive, mais elle est fidèle.

(iv) Si $\varphi : G \to \mathfrak{S}_X$ est une action de G sur X, elle induit, par propriété universelle du quotient, une action fidèle $\overline{\varphi} : G/\ker(\varphi) \to \mathfrak{S}_X$ du groupe-quotient $G/\ker(\varphi)$ sur X.

Par exemple, l'action naturelle de $\mathrm{GL}(V)$ sur les droites de V (exemple (i)) induit une action fidèle et transitive de $\mathrm{PGL}(V)$ sur les droites de V.

Définitions (partie stable ou fixe ; stabilisateur et fixateur d'une partie)
Soient G un ensemble agissant sur un ensemble X, et Y une partie de X.

(i) On dit que Y est *stable* sous l'action de G lorsque $\forall g \in G,\ \forall y \in Y,\ g \cdot y \in Y$. On dit que Y est *fixe* sous l'action de G lorsque $\forall g \in G,\ \forall y \in Y,\ g \cdot y = y$.

(ii) Le *stabilisateur* de Y est le sous-groupe des éléments de G qui stabilisent Y, au sens restreint où la partie et son image son *égales* ; on le note (encore) $\mathrm{Stab}(Y)$. Le *fixateur* de Y est le sous-groupe des éléments de G qui fixent tous les éléments de Y ; on le note $\mathrm{Fix}(Y)$. Ainsi, en notant $g \cdot Y = \{g \cdot y,\ y \in Y\}$ pour tout $g \in G$,

$$\mathrm{Stab}(Y) = \{g \in G,\ g \cdot Y = Y\} \quad \text{et} \quad \mathrm{Fix}(Y) = \{g \in G,\ \forall y \in Y,\ g \cdot y = y\}\,.$$

Exercice 51

(i) S'assurer que $\mathrm{Stab}(Y)$ et $\mathrm{Fix}(Y)$ sont bien des sous-groupes de G.

(ii) Montrer que dans le cas général, $\{g \in G,\ g \cdot Y \subseteq Y\}$ n'est pas un sous-groupe de G, mais que c'est le cas si Y est une partie finie de X.

Exemple
On considère l'action naturelle de $\mathrm{GL}(V)$ sur un espace vectoriel V. Soient D une droite de V et $g \in \mathrm{GL}(V)$. Dire que D est stable par g (ou par le groupe $\langle g \rangle$) signifie que D est une droite propre de g. Dire que D est fixe par g signifie que D est une droite propre de g associée à la valeur propre 1.

Définitions (normalisateur, centralisateur)
Soient G un groupe et H un sous-groupe de G. Le *normalisateur* et le *centralisateur* de H sont les sous-groupes de G, notés respectivement $\mathrm{Norm}_G(H)$ et $Z_G(H)$, définis par

$$\mathrm{Norm}_G(H) = \left\{g \in G,\ gHg^{-1} = H\right\} \quad \text{et} \quad Z_G(H) = \{g \in G,\ \forall h \in H,\ gh = hg\}\,.$$

Exercice 52

(i) Montrer que si H est un groupe fini, $\mathrm{Norm}_G(H) = \left\{g \in G,\ gHg^{-1} \subseteq H\right\}$.

(ii) Montrer, dans les conditions de la définition, que $H \lhd \mathrm{Norm}_G(H)$ et que $\mathrm{Norm}_G(H)$ est le plus grand sous-groupe de G dans lequel H est distingué.

(iii) $Z_G(G)$ est le centre de G.

5.2 L'équation aux classes

Proposition (le cardinal d'une orbite est l'indice du stabilisateur)

Soit G un groupe agissant sur un ensemble X. Alors, pour tout $x \in X$, $\boxed{\mathrm{Card}\,(G \cdot x) = [G : G_x]}$

Preuve. Soit $x \in X$. L'application $G \to X$, $g \mapsto gx$ a pour image l'orbite de x et est constante sur les classes à gauche modulo G_x. Mieux, ses fibres non vides sont exactement lesdites classes à gauche. Par propriété universelle du quotient pour les applications, elle induit une bijection entre l'ensemble quotient $(G/G_x)_g$ et l'orbite de x. ■

A noter
(i) C'est une égalité entre cardinaux que l'on peut aussi écrire sous la forme

$$|G| = |G_x| \times \mathrm{Card}\,(G \cdot x)\,.$$

Dans le cas où G est un groupe fini, cette formule dit notamment que les cardinaux des orbites sont finis et divisent l'ordre de G.

(ii) En particulier, lorsque l'action est transitive, il n'y a qu'une seule orbite ω, les groupes d'isotropie sont tous conjugués au même groupe G_ω, et $|G| = |G_\omega| \times \mathrm{Card}(\omega)$.

Proposition (partition en orbites)
Soit G un groupe agissant sur un ensemble fini X. On note $\mathcal{R}$ un système de représentants des orbites, c'est-à-dire une partie de X qui contient un élément de chaque orbite et un seul. Alors,

$$\boxed{\mathrm{Card}(X) = \sum_{x \in \mathcal{R}} [G : G_x]}$$

PREUVE. Les orbites forment une partition de X. La somme de leurs cardinaux est donc le cardinal de X. On conclut avec la proposition précédente. ■

Notation (ensemble des points fixes)
Soit G un groupe agissant sur un ensemble X. On note X^G l'ensemble des élément de X qui sont fixés par G. Si $g \in G$, on note aussi X^g l'ensemble des points fixes par g. Ainsi,

$$X^g = \{x \in X, gx = x\} \quad \text{et} \quad X^G = \{x \in X,\ \forall g \in G,\ gx = x\} = \bigcap_{g \in G} X^g;$$

Proposition (formule de Burnside, nombres d'orbites)
Soit G un groupe fini agissant sur un ensemble fini X. On note Ω l'ensemble des orbites de l'action. Pour tout $g \in G$ on note aussi X^g l'ensemble des éléments de X fixes par g. Alors,

$$\boxed{\mathrm{Card}\,(\Omega) = \frac{1}{|G|} \sum_{g \in G} \mathrm{Card}\,(X^g)}$$

PREUVE. Soit $\mathcal{I} = \{(g, x) \in G \times X,\ gx = x\}$. On compte le cardinal de $\mathcal{I}$ de deux façons, d'abord selon les g, puis selon les x. Pour chaque $g \in G$, le nombre de couples (g, x) qui sont dans $\mathcal{I}$ est le nombre de points fixes de g. Cela montre, d'une part, que $\mathrm{Card}\,(\mathcal{I})$ est la somme de la formule. D'autre part, pour chaque $x \in X$, le nombre de couples (g, x) qui sont dans $\mathcal{I}$ est l'ordre du groupe d'isotropie G_x, ce qui montre que $\mathrm{Card}\,\mathcal{I} = \sum_{x \in X} |G_x|$. Or, les groupes d'isotropie de deux points d'une même orbite ω sont conjugués ; ils ont donc le même ordre, qui est $|G|/\mathrm{Card}(\omega)$. En regroupant les termes de la somme précédente en orbites, on obtient que $\mathrm{Card}\,(\mathcal{I}) = \sum_{\omega \in \Omega} \mathrm{Card}(\omega) \times |G|/\mathrm{Card}(\omega) = \mathrm{Card}\,(\Omega) \times |G|$. ■

Définition (p-groupe)
Soit p un nombre premier. Un *p-groupe* est un groupe fini dont l'ordre est une puissance de p.

Proposition (action d'un p-groupe)
Soient X un ensemble fini, p un nombre premier et G un p-groupe. On note X^G l'ensemble des points fixes de l'action. Alors, $\mathrm{Card}(X) \equiv \mathrm{Card}\left(X^G\right)\ [p]$.

PREUVE. On écrit la formule de partition de X en orbites : $\mathrm{Card}(X) = \sum_{x \in \mathcal{R}} [G : G_x]$ où $\mathcal{R}$ est un système de représentants des orbites. Puisque G est un p-groupe, tous les indices $[G : G_x]$ sont divisible par p, à l'exception des orbites réduites à un singleton, qui sont exactement les éléments de X^G. ■

Proposition (le centre d'un p-groupe est non trivial)
Soient p un nombre premier et G un p-groupe. Alors, le centre de G n'est pas réduit au sous-groupe trivial $\{1\}$.

PREUVE. On fait agir G par conjugaison sur lui-même. Un élement de G est fixe pour l'action si, et seulement s'il est dans le centre de G ; autrement dit, avec les notations de la proposition précédente, $G^G = Z(G)$. Ainsi, ladite proposition affirme que $|Z(G)| = |G|\ [p]$. Or, G est un p-groupe ; donc p divise $|Z(G)|$. Comme $Z(G)$ contient 1_G, son ordre n'est pas nul ; donc $|Z(G)|$ est un multiple non nul de p : le centre est non trivial. ■

Exemple (groupe quaternionique $\mathbb{H}_8$)
Le groupe $\mathbb{H}_8$ est le groupe d'ordre 8 dont les éléments sont notés $\{\pm 1, \pm i, \pm j, \pm k\}$, soumis à la table de loi donnée par la règle habituelle des signes et par les relations

$$i^2 = j^2 = k^2 = -1,\ ij = -ji = k,\ jk = -kj = i,\ ki = -ik = j.$$

On donne aussi une représentation de $\mathbb{H}_8$ comme sous-groupe de $\mathrm{SL}(2,\mathbb{C})$ engendré par les matrices :

$$\mathbb{H}_8 = \left\langle \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \right\rangle.$$

Le groupe $\mathbb{H}_8$ est un 2-groupe dont le centre est $\{-1,1\}$, comme on le vérifie immédiatement à partir de la table de la loi de $\mathbb{H}_8$. On peut aussi le représenter par les matrices de $\mathrm{SL}(4,\mathbb{Z})$ suivantes :

$$\mathbb{H}_8 = \left\langle \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \right\rangle.$$

Noter que les signes "=" des lignes précédentes sont à prendre au sens des isomorphismes de groupes. Le mot *représentation* a une signification précise en mathématiques, qui ne sera pas développée ici. Pour une introduction lumineuse, voir par exemple le livre de Jean-Pierre Serre *Représentation linéaire des groupes finis*.

Proposition (un théorème de Cauchy)
Soient G un groupe fini et p un nombre premier divisant l'ordre de G. Alors, G contient un élément d'ordre p.

Preuve. On note n l'ordre de G — noter que $n \neq 1$ puisque p divise n. Soit

$$\mathcal{I} = \{(g_1, g_2 \ldots, g_p) \in G^p,\ g_1 g_2 \ldots g_p = 1\}.$$

Puisqu'un choix arbitraire de $g_1, \ldots, g_{p-1}$ conduit à un unique élément de $\mathcal{I}$ en posant $g_p = g_{p-1}^{-1} \ldots g_1^{-1}$, le cardinal de $\mathcal{I}$ est n^{p-1}. En particulier, ce nombre est un multiple de p. Soient $c \in \mathfrak{S}_p$ le p-cycle $(1,2,\ldots,p)$, et C le sous-groupe cyclique d'ordre p de $\mathfrak{S}_p$ que c engendre. On fait agir C sur les éléments de $\mathcal{I}$ par l'action naturelle $\sigma \cdot (g_1, g_2 \ldots, g_p) = \left(g_{\sigma(1)}, g_{\sigma(2)} \ldots, g_{\sigma(p)}\right)$ — que cela définisse une action est élémentaire. Les points fixes de cette action sont les $(g,g,\ldots,g)$ pour lesquels $g \in G$ vérifie $g^p = 1$, c'est-à-dire pour lesquels $g = 1$ ou g est d'ordre p, puisque p est un nombre premier. Or, C est un p-groupe. Donc le nombre de points fixes de l'action égale $\#\mathcal{I}$ modulo p. Comme le cardinal de $\mathcal{I}$ est un multiple de p, c'est vrai aussi du nombre de points fixes de l'action. Or, ce nombre est non nul puisque $(1,\ldots,1)$ est un point fixe. Il y donc au moins p points fixes de l'action, et en particulier au moins $p-1$ éléments d'ordre p dans G. ■

Exercice 53 Montrer que tous les sous-groupes de $\mathbb{H}_8$ sont distingués.

Exemple — paradigmatique au sens où l'action d'un groupe sur un ensemble renseigne sur le groupe lui-même
On notre $\mathrm{SO}(\mathcal{C})$ le *groupe positif du cube*, qui est le groupe des rotations de l'espace euclidien de dimension 3 qui stabilisent un cube — disons le cube $\mathcal{C} = [0,1]^3$ déjà rencontré, tous les cubes sont semblables et ont donc des groupes conjugués. Le groupe $G = \mathrm{SO}(\mathcal{C})$ agit naturellement sur les six centres des faces du cube — ces centres sont les points d'intersection du bord du cube et de sa sphère inscrite, qui sont toutes les deux stabilisées par le groupe. En considérant les rotations d'angles multiples de $\pi/2$ et dont les axes passent par les centres des faces, on montre que cette action est transitive. Enfin, si c est le centre d'une face et si G_c est le stabilisateur de c, alors tout élément de G_c est une rotation dont l'axe passe par c et qui agit sur les sommets de la face — pour des raisons de distance maximale, par exemple. Cela montre que G_c est le groupe cyclique d'ordre 4 engendré par n'importe laquelle des deux rotations d'ordre 4 et d'axe c. En appliquant la formule de partition en orbites, cela montre que l'ordre de $\mathrm{SO}(\mathcal{C})$ est $6 \times 4 = 24$.

On note $\mathrm{O}(\mathcal{C})$ le *groupe total du cube*, qui est le groupe de toutes les isométries qui le stabilisent. La symétrie centrale $-I_3$ est une isométrie négative qui stabilise le cube. Ainsi, le déterminant $\mathrm{O}(\mathcal{C}) \to \{-1,1\}$ est surjectif et son noyau est d'ordre 24. Cela montre que $\mathrm{O}(\mathcal{C})$ est d'ordre 48.

Exercice : une fois l'ordre de $\mathrm{SO}(\mathcal{C})$ et de $\mathrm{O}(\mathcal{C})$ connus, faire la liste de toutes ces isométries. Pour cela, considérer celles qui se conçoivent aisément, les compter, et se rendre compte qu'on les connaît toutes.

Pour aller plus loin : en faisant agir le groupe positif du cube sur les quatre diagonales, on montre que

$$\mathrm{SO}(\mathcal{C}) \simeq \mathfrak{S}_4 \quad \text{et} \quad \mathrm{O}(\mathcal{C}) \simeq \mathfrak{S}_4 \times \mathbb{Z}/2\mathbb{Z}.$$

5.3 Produits semi-directs de groupes

Définition (suite exacte d'homomorphismes de groupes)

Soient $i : N \to G$ et $p : G \to Q$ deux homomorphismes de groupes. On dit que *la suite*

$$1 \longrightarrow N \xrightarrow{i} G \xrightarrow{p} Q \longrightarrow 1 \tag{14}$$

est exacte lorsque ① i est injectif ② p est surjectif ③ $\mathrm{im}(i) = \ker(p)$. Lorsque les groupes sont abéliens, on remplace souvent les 1 par des 0.

Autrement dit, la suite (14) est exacte si, et seulement si $N \triangleleft G$ et $Q \simeq G/N$.

Exemples

(i) Soient G et H deux groupes. On note comme d'habitude $G \times H$ leur produit direct. Soient $i_1 : G \to G \times H$ l'injection $g \mapsto (g, 1)$ et $p_2 : G \times H \to H$ la second projection $(g, h) \mapsto h$. Alors, i et p sont des homomorphismes de groupes et la suite

$$1 \longrightarrow G \xrightarrow{i_1} G \times H \xrightarrow{p_2} H \longrightarrow 1$$

est exacte.

(ii) Le déterminant induit la suite exacte

$$1 \longrightarrow \mathrm{SO}(3) \xrightarrow{i} \mathrm{O}(3) \xrightarrow{\det} \{-1, 1\} \longrightarrow 1$$

où i est l'inclusion.

(iii) Si $n \geq 2$, la signature ε induit la suite exacte

$$1 \longrightarrow \mathfrak{A}_n \xrightarrow{i} \mathfrak{S}_n \xrightarrow{\varepsilon} \{-1, 1\} \longrightarrow 1$$

où i est l'inclusion. En remplaçant le groupe multiplicatif $\{-1, 1\}$ par sa version additive $\mathbb{Z}/2\mathbb{Z}$ qui lui est isomorphe, cette suite exactee s'écrit aussi $1 \to \mathfrak{A}_n \to \mathfrak{S}_n \to \mathbb{Z}/2\mathbb{Z} \to 1$.

(iv) Si V est un espace vectoriel de dimension finie sur un corps $\mathbb{F}$, le déterminant induit la suite exacte

$$1 \longrightarrow \mathrm{SL}(V) \xrightarrow{i} \mathrm{GL}(V) \xrightarrow{\det} \mathbb{F}^\times \longrightarrow 1$$

où i est (encore) l'inclusion. Pour s'assurer de la surjectivité du déterminant, prendre une version matricielle de cette suite exacte et considérer les matrices $\mathrm{diag}(\lambda, 1, \ldots, 1)$.

(v) La projection canonique $\mathbb{Z} \to \mathbb{Z}/2\mathbb{Z}$ annule $4\mathbb{Z}$, si bien qu'elle induit un homomorphisme surjectif de groupes $p : \mathbb{Z}/4\mathbb{Z} \to \mathbb{Z}/2\mathbb{Z}$, $x + 4\mathbb{Z} \mapsto x + 2\mathbb{Z}$. Par ailleurs, la multiplication par 2 induit un homomorphisme de groupes $\mathbb{Z} \to \mathbb{Z}/4\mathbb{Z}$, $x \mapsto 2x + 4\mathbb{Z}$ dont le noyau est $2\mathbb{Z}$; elle induit donc un homomorphisme injectif de groupes $m : \mathbb{Z}/2\mathbb{Z} \to \mathbb{Z}/4\mathbb{Z}$.

On considère l'homomorphisme de groupes $f : \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z} \to \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, $(x, y) \mapsto (x, p(y))$. Il est surjectif et son noyau est le sous-groupe de $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}$ engendré par $(0, 2)$. Autrement dit, f fournit une suite exacte

$$0 \longrightarrow \mathbb{Z}/2\mathbb{Z} \xrightarrow{i} \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z} \xrightarrow{f} (\mathbb{Z}/2\mathbb{Z})^2 \longrightarrow 0$$

où $i(x) = (0, m(x))$ — attention à la notation additive qui invite à remplacer les 1 extrémaux habituels des suites exactes par des 0. En résumé, en omettant de différencier les classes par des notations une fois que l'on s'est bien assuré du sens des objets, les deux flèches centrales de la suite sont $x \mapsto (0, 2x)$ et $(x, y) \mapsto (x, y)$. Cette suite exacte n'est pas scindée car $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}$ contient un unique sous-groupe isomorphe à $(\mathbb{Z}/2\mathbb{Z})^2$ et la restriction de f à ce dernier n'est pas un isomorphisme puisque son image est $\mathbb{Z}/2\mathbb{Z} \times \{0\}$.

Définition (produit semi-direct défini par une action par automorphismes)

Soient Q et N deux groupes, et $\varphi : Q \to \mathrm{Aut}(N)$ un homomorphisme de groupes — autrement dit, une action de Q sur N par automorphismes. Le *produit semi-direct de N et Q induit par φ* est la loi de groupe définie sur le produit cartésien $N \times Q$ par

$$(n, q) \cdot (n', q') = (n \cdot \varphi(q)\,(n')\,, q \cdot q')\,. \tag{15}$$

On note $N \rtimes_\varphi Q$ ou simplement $N \rtimes Q$ cette loi de groupe sur le produit cartésien $N \times Q$. L'inverse de (n, q) est $\left(\varphi\left(q^{-1}\right)\left(n^{-1}\right), q^{-1}\right)$.

Exercice 54

(i) Vérifier que cela définit bien une loi de groupe sur $N \times Q$ pour laquelle $N \lhd N \rtimes_\varphi Q$.

(ii) Avec les notations de la définition, l'injection $i_1 : N \to N \rtimes_\varphi Q$, $n \mapsto (n, 1)$ et la projection $p_2 : N \rtimes_\varphi Q \to Q$, $(n, q) \mapsto q$ induisent la suite exacte

$$1 \longrightarrow N \xrightarrow{i_1} N \rtimes_\varphi Q \xrightarrow{p_2} Q \longrightarrow 1. \tag{16}$$

A noter

Une lecture possible de la définition du produit semi-direct (15) : la loi de groupe sur $N \times Q$ n'est pas celle du produit direct qui se fait coordonnée par coordonnée, mais le produit sur la première coordonnée est "tordu" par l'action.

Proposition (une suite exacte scindée est un produit semi-direct)

Soit

$$1 \longrightarrow N \xrightarrow{i} G \xrightarrow{p} Q \longrightarrow 1$$

une suite exacte de groupes. Les assertions suivantes sont équivalentes.

(i) Il existe un homomorphisme de groupes $s : Q \to G$ tel que $p \circ s = \mathrm{id}_Q$.

$$1 \longrightarrow N \xrightarrow{i} G \underset{s}{\overset{p}{\rightleftarrows}} Q \longrightarrow 1$$

(ii) Il existe un sous-groupe Q' de G tel que la restriction de p à Q' soit un isomorphisme de groupes $Q' \xrightarrow{\sim} Q$.

(iii) Il existe un sous-groupe distingué N' de G, un isomorphisme de groupes $I : N' \xrightarrow{\sim} N$, un sous-groupe Q' de G et un isomorphisme de groupes $P : Q' \xrightarrow{\sim} Q$ tels que l'application

$$\begin{array}{rccc} \pi : & N' \rtimes_\psi Q' & \xrightarrow{\sim} & G \\ & (n, q) & \longmapsto & nq \end{array}$$

soit un isomorphisme de groupes pour l'action ψ de Q' sur N' définie par $\psi(q)(n) = qnq^{-1}$, et tels que le diagramme suivant soit commutatif

$$\begin{array}{ccccccccc} 1 & \longrightarrow & N' & \xrightarrow{i_1} & N' \rtimes_\psi Q' & \xrightarrow{p_2} & Q' & \longrightarrow & 1 \\ & & \downarrow{\scriptstyle I} & & \downarrow{\scriptstyle \pi} & & \downarrow{\scriptstyle P} & & \\ 1 & \longrightarrow & N & \xrightarrow{i} & G & \xrightarrow{p} & Q & \longrightarrow & 1 \end{array}$$

(iv) Il existe une action de Q sur N par automorphismes $\varphi : Q \to \mathrm{Aut}(N)$ et un isomorphisme de groupes $f : N \rtimes_\varphi Q \xrightarrow{\sim} G$ tels que le diagramme suivant soit commutatif

$$\begin{array}{ccccccccc} 1 & \longrightarrow & N & \xrightarrow{i_1} & N \rtimes_\varphi Q & \xrightarrow{p_2} & Q & \longrightarrow & 1 \\ & & \downarrow{\scriptstyle \mathrm{id}_N} & & \downarrow{\scriptstyle f} & & \downarrow{\scriptstyle \mathrm{id}_Q} & & \\ 1 & \longrightarrow & N & \xrightarrow{i} & G & \xrightarrow{p} & Q & \longrightarrow & 1 \end{array}$$

A noter

Dire que le diagramme est commutatif signifie que les applications composées symbolisées par des suites de flèches du diagramme ne dépendent pas du chemin choisi entre deux groupes du diagramme. Par exemple, dans ce dernier diagramme, $f \circ i_1 = i \circ \mathrm{id}_N = i$.

PREUVE. (i)⇒(ii) Soit $Q' = \mathrm{im}(s)$, sous-groupe de G. Alors, la restriction de p à Q' est un isomorphisme entre Q' et Q, dont la réciproque est s.

(ii)⇒(iii) On note $P : Q' \to Q$ la restriction de p à Q' et $s = P^{-1}$ sa réciproque. On note aussi $N' = i(N)$ et $I : N' \to N$ la réciproque de l'isomorphisme $i : N \to N'$. Puisque la suite est exacte, $N' = \ker p$ est un

sous-groupe distingué de G ce qui entraîne que ψ, définie comme dans l'énoncé, est bien une action sur N. Pour l'action ψ, la loi de groupe de $N' \rtimes_\psi Q'$ s'écrit $(n,q)\cdot(n',q') = \left(nqn'q^{-1}, qq'\right)$. Il en résulte immédiatement que π est un homomorphisme de groupes. Soit alors $\rho : G \to N' \rtimes_\psi Q'$ l'application définie par $\rho(g) = \left(g\cdot(s\circ p(g))^{-1}, s\circ p(g)\right)$, pour tout $g \in G$. D'une part, la définition de s assure que $s\circ p(g) \in Q'$. D'autre part, $g\cdot(s\circ p(g))^{-1} \in \ker p = N'$. Enfin, soit $g \in G$; on note $\rho(g) = (n,q) \in N'\times Q'$. En particulier, $g = nq$, si bien que $\pi\circ\rho(g) = g$: on a montré que $\pi\circ\rho = \mathrm{id}_G$. Inversement, si $(n,q) \in N'\times Q'$, alors $\rho\circ\pi(n,q) = \rho(nq)$. Comme $q \in Q'$, $s\circ p(nq) = s\circ p(q) = q$, et $(nq)\,(s\circ p(nq))^{-1} = nqq^{-1} = n$, ce qui montre que $\rho\circ\pi(n,q) = (n,q)$. On a montré que π et ρ sont des bijections réciproques l'une de l'autre. Que le diagramme commute résulte immédiatement des définitions de I, i_1, π, p_2 et P.
(iii)⇒(iv) Dans la situation de (iii), soient $I : N' \xrightarrow{\sim} N$ et $P : Q' \xrightarrow{\sim} Q$ deux isomorphismes de groupes. Puisque le diagramme commute, la réciproque de I est $i : N \xrightarrow{\sim} N'$ et P est la restriction de p à Q' ; on note $s : Q \xrightarrow{\sim} Q'$ la réciproque de P. On définit alors une action $\varphi : Q \to \mathrm{Aut}(N)$ de Q sur N par la formule $\varphi(q)(n) = I\left[\psi\left(s(q)\right)\left(i(n)\right)\right] = I\left[s(q)i(n)s(q)^{-1}\right]$. Alors, $i\times s : N\rtimes_\varphi Q \to N'\rtimes_\psi Q'$, $(n,q)\mapsto(i(n),s(q))$ est un homomorphisme de groupes, comme le montre un calcul élémentaire, évidemment bijectif. Puisque π est un isomorphisme, la composée $f = \pi\circ(i\times s) : N\rtimes_\varphi Q \to G$, $(n,q)\mapsto i(n)s(q)$ est aussi un isomorphisme de groupes. Que le diagramme commute résulte immédiatement des définitions de f et de s. ■

Définition (suite exacte scindée, section)
Dans la situation de la proposition, on dit que la suite exacte est *scindée* et que l'homomorphisme $s : Q \to G$ est une *section* (de la suite, ou de p). On appelle aussi parfois *section* le sous-groupe $Q' = s(Q)$ de G lui-même.

A noter
(i) Une lecture opératoire de la proposition est la suivante : *lorsqu'on a un suite exacte* $1 \to N \to G \to Q \to 1$, *le groupe* G *est isomorphe à un produit semi-direct* $N\rtimes_\varphi Q$ *si, et seulement si la suite est scindée,* i.e. *si, et seulement si elle admet une section.*

(ii) La version (iii) d'un suite exacte scindée est celle du *produit semi-direct interne* : N' et Q' sont des sous-groupes de G, N' est distingué dans G et tout élément de G s'écrit de manière unique sous la forme nq où $n \in N'$ et $q \in Q'$. En outre, le produit de deux éléments de G ainsi décomposés se lit au travers de l'action par conjugaison de Q sur N, *via* la formule $nq\cdot n'q' = n\left(qn'q'^{-1}\right)\cdot qq'$.

Autrement dit, si G possède deux sous-groupes N' et Q' tels que
① $N' \triangleleft G$
② $N'\cap Q' = \{1\}$
③ $G = N'Q'$,
alors G est un produit semi-direct (interne) $G \simeq N'\rtimes Q'$.

(iii) La version (iii) d'un suite exacte scindée est celle du *produit semi-direct externe* : les groupes N et Q ne sont pas des sous-groupes de G, mais Q agit sur N par automorphismes et le produit semi-direct qui résulte de cette action est isomorphe à G.

Exemples
On reprend les exemples ci-dessus.

(i) Le groupe O(3) contient des matrices d'isométries négatives, comme par exemple $D = \mathrm{diag}(-1,1,1)$, mais aussi $-I_3$. Une telle matrice engendre un sous-groupe d'ordre 2 de O(3), et la restriction du déterminant à ce sous-groupe est un isomorphisme.
On choisit d'abord la section $\langle D\rangle$. Elle scinde la suite exacte $1 \longrightarrow \mathrm{SO}(3) \longrightarrow \mathrm{O}(3) \xrightarrow{\det} \{-1,1\} \longrightarrow 1$, ce qui montre qu'on a un produit semi-direct

$$\mathrm{O}(3) \simeq \mathrm{SO}(3)\rtimes \mathbb{Z}/2\mathbb{Z}.$$

On choisit ensuite la section $\langle -I_3\rangle$ du déterminant. Cette fois, puisque $-I_3$ est dans le centre de O(3), le (iii) de la proposition montre que l'action ψ de $\langle -I_3\rangle$ sur SO(3), qui est la conjugaison, est triviale. Ainsi, le produit semi-direct pour cette action est un produit direct, et on a aussi un isomorphisme

$$\mathrm{O}(3) \simeq \mathrm{SO}(3)\times \mathbb{Z}/2\mathbb{Z}.$$

(ii) Lorsque $n \geq 2$, le groupe symétrique $\mathfrak{S}_n$ contient des permutations négatives d'ordre 2 — par exemple, les transpositions. Chacune d'elle fournit une section de la suite exacte $1 \to \mathfrak{A}_n \to \mathfrak{S}_n \to \mathbb{Z}/2\mathbb{Z} \to 1$ qui est donc

scindée. On a un produit semi-direct

$$\mathfrak{S}_n \simeq \mathfrak{A}_n \rtimes \mathbb{Z}/2\mathbb{Z}.$$

Noter que si $n \geq 5$, les seuls sous-groupes distingués de $\mathfrak{S}_n$ sont $\{1\}$, $\mathfrak{A}_n$ et $\mathfrak{S}_n$. En particulier, $\mathfrak{S}_n$ ne contient pas de sous-groupe distingué d'ordre 2. Cela montre qu'il n'y a pas d'isomorphisme entre $\mathfrak{S}_n$ et le produit direct $\mathfrak{A}_n \times \mathbb{Z}/2\mathbb{Z}$. Exercice : montrer que cela vaut aussi pour $n = 3$ ou 4.

(iii) Si $n \geq 1$ et si $\mathbb{F}$ est un corps, les matrices diagonales $\{\mathrm{diag}(x, 1, \ldots, 1),\ x \in \mathbb{F}^\times\}$ forment un sous-groupe de $\mathrm{GL}(n, \mathbb{F})$ qui fournit une section du déterminant. Ainsi, la suite $1 \to \mathrm{SL}(n, \mathbb{F}) \to \mathrm{GL}(n, \mathbb{F}) \to \mathbb{F}^\times \to 1$ est scindée. On a un produit semi-direct

$$\mathrm{GL}(n, \mathbb{F}) \simeq \mathrm{SL}(n, \mathbb{F}) \rtimes \mathbb{F}^\times.$$

5.4 Théorèmes de Sylow

Définition (p-Sylow d'un groupe fini)
Soient G un groupe fini et p un nombre premier. Un *p-sous-groupe de Sylow*[♫] *de G* est un p-sous-groupe S de G dont l'indice n'est pas divisible par p. On dit parfois simplement un *p-Sylow*.

Autrement dit, si $|G| = p^a q$ où $a \geq 0$ et où p ne divise pas q, un sous-groupe S de G est un p-Sylow de G si, et seulement si $|S| = p^a$.

Exemples

(i) Si G est un p-groupe, il a un unique p-Sylow qui est G lui-même.

(ii) Si G est un groupe abélien fini et si p est un nombre premier, G admet un unique p-Sylow qui est sa composante de p-torsion.

(iii) Soient p un nombre premier et n un entier naturel non nul. On note $S_{n,p}$ le sous-groupe de $\mathrm{GL}(n, \mathbb{F}_p)$ formé des matrices triangulaires supérieures avec des 1 sur la diagonale : si δ désigne le symbole de Kronecker,

$$S_{n,p} = \{M \in \mathrm{GL}(n, \mathbb{F}_p),\ \forall j, k \in \{1, \ldots, n\},\ j \geq k \Longrightarrow M_{j,k} = \delta_{j,k}\}.$$

Autrement dit, $S_{n,p}$ est formé des matrices de $\mathrm{GL}(n, \mathbb{F}_p)$ de la forme

$$\begin{pmatrix} 1 & \star & \cdots & \star \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \star \\ 0 & \cdots & 0 & 1 \end{pmatrix}$$

Que $S_{n,p}$ soit un sous-groupe de $\mathrm{GL}(n, \mathbb{F}_p)$ est immédiat. Puisque les $n(n-1)/2$ coefficients au dessus de la diagonale sont librement choisis dans $\mathbb{F}_p$, l'ordre de $S_{n,p}$ est $p^{n(n-1)/2}$. Par ailleurs, l'ordre de $\mathrm{GL}(n, \mathbb{F}_p)$ est

$$(p^n - 1)(p^n - p)\left(p^n - p^2\right) \ldots \left(p^n - p^{n-1}\right) = p^{\frac{n(n-1)}{2}} q$$

où q est un entier premier avec p — on a vu ce calcul, il suffit de compter les bases de $\mathbb{F}_p^n$. Cela montre que $S_{n,p}$ est un p-Sylow de $\mathrm{GL}(n, \mathbb{F}_p)$.

Exercice 55 Le conjugué d'un p-Sylow est toujours un p-Sylow.

Théorème (premier théorème de Sylow)
Si G est un groupe fini et si p est un nombre premier, alors G contient au moins un p-Sylow.

Preuve. On note $n = |G|$.
Le premier geste consiste à voir G comme un sous-groupe d'un groupe de permutations — ce résultat est parfois évoqué sous le nom de théorème de Cayley. Pour cela, on fait agir G sur lui-même par translation à gauche. Cela fournit l'homomorphisme de groupes $\varphi : G \to \mathfrak{S}_G$, $g \mapsto t_g$ où t_g est la permutation de G définie par $t_g(h) = gh$, pour tout $h \in G$. Le noyau de φ est trivial (l'action est fidèle), si bien que G est isomorphe à son image par φ qui est un sous-groupe de $\mathfrak{S}_G$.

[♫] Ludwig Sylow, 1832–1918.

On suppose ainsi que G est un sous-groupe de $\mathfrak{S}_n$. Le deuxième geste consiste à voir à son tour $\mathfrak{S}_n$ comme un sous-groupe de $\mathrm{GL}(n,\mathbb{F}_p)$, ou plutôt de $\mathrm{GL}\left(\mathbb{F}_p^n\right)$. Si $\sigma \in \mathfrak{S}_n$, on note $f_\sigma \in \mathrm{GL}\left(\mathbb{F}_p^n\right)$ l'endomorphisme de $\mathbb{F}_p^n$ qui permute sa base canonique $(v_1,\ldots,v_n)$ *via* σ : il est défini par

$$\forall k \in \{1,\ldots,n\},\ f(v_k) = v_{\sigma(k)}.$$

Un calcul élémentaire montre que l'application $f : \mathfrak{S}_n \to \mathrm{GL}\left(\mathbb{F}_p^n\right)$, $\sigma \mapsto f_\sigma$ est un homomorphisme injectif de groupes. Ainsi, G, qui est un sous-groupe de $\mathfrak{S}_n$, est isomorphe à son image par f qui est elle-même un sous-groupe de $\mathrm{GL}\left(\mathbb{F}_p^n\right)$.
On suppose ainsi que G est un sous-groupe de $\mathrm{GL}(n,\mathbb{F}_p)$. On conclut alors avec le lemme suivant, appliqué à $\mathcal{G} = \mathrm{GL}(n,\mathbb{F}_p)$, qui contient le p-Sylow $S_{n,p}$ de l'exemple ci-dessus. ■

Lemme
Soient $\mathcal{G}$ un groupe fini, p un nombre premier, $\mathcal{S}$ un p-Sylow de $\mathcal{G}$ et G un sous-groupe de $\mathcal{G}$. Alors, il existe $h \in \mathcal{G}$ tel que $h\mathcal{S}h^{-1} \cap G$ soit un p-Sylow de G.

Preuve. On fait agir G sur l'ensemble $(\mathcal{G}/\mathcal{S})_g$ des classes à gauche de $\mathcal{G}$ modulo $\mathcal{S}$, par translation à gauche. Si $g \in G$ et si $h\mathcal{S} \in (\mathcal{G}/\mathcal{S})_g$, l'action s'écrit $g \cdot h\mathcal{S} = (gh)\mathcal{S}$ — que ce soit une action à gauche a déjà été vu : c'est la restriction à G de l'action de $\mathcal{G}$ sur $(\mathcal{G}/\mathcal{S})_g$ par translation à gauche. Si $h \in \mathcal{G}$, le groupe d'isotropie de $h\mathcal{S}$ pour cette action est $G_{h\mathcal{S}} = \{g \in G,\ gh\mathcal{S} = h\mathcal{S}\} = h\mathcal{S}h^{-1} \cap G$. C'est un p-sous-groupe de G puisque c'est à la fois un sous-groupe de G et un sous-groupe du p-groupe $h\mathcal{S}h^{-1}$. Il suffit ainsi de trouver $h \in \mathcal{G}$ tel que p ne divise pas l'indice $[G : G_{h\mathcal{S}}]$, puisque dans ces conditions, $G_{h\mathcal{S}} = h\mathcal{S}h^{-1} \cap G$ sera un p-Sylow de G. On écrit l'équation aux classes :

$$[\mathcal{G} : \mathcal{S}] = \sum_{h \in \mathcal{R}} [G : G_{h\mathcal{S}}]$$

où $\mathcal{R} \subseteq (\mathcal{G}/\mathcal{S})_g$ désigne un système de représentant des orbites. Puisque $\mathcal{S}$ est un p-Sylow de $\mathcal{G}$, le nombre $[\mathcal{G} : \mathcal{S}]$ n'est pas un multiple de p. Par conséquent, l'un au moins des termes de la somme n'est pas un multiple de p : soit $h \in \mathcal{R}$ tel que p ne divise pas $[G : G_{h\mathcal{S}}]$. Alors, $G_{h\mathcal{S}} = h\mathcal{S}h^{-1} \cap G$ est un p-Sylow de G. ■

A noter
(i) Ce que montre le lemme, ce n'est pas que l'intersection d'un p-Sylow $\mathcal{S}$ avec un sous-groupe G est un p-Sylow de G, mais que l'intersection de G et d'un certain conjugué de $\mathcal{S}$ — qui est encore un p-Sylow — est un p-Sylow de G.

(ii) Le premier théorème de Sylow montre l'existence de p-Sylow dans n'importe quel groupe fini. Ces p-Sylow sont évidemment des p-sous-groupes maximaux (pour l'inclusion). Un autre point de vue est parfois pris pour introduire les p-Sylow en les définissant comme étant les p-sous-groupes maximaux dont l'existence est immédiate, le premier théorème de Sylow consistant alors à montrer que ce sont les p-sous-groupes dont p ne divise pas l'indice.

Théorème (deuxième théorème de Sylow)
Soient G est un groupe fini et p un nombre premier. Alors,
(i) tout p-sous-groupe de G est contenu dans un p-Sylow de G ;
(ii) tous les p-Sylow de G sont conjugués.

Preuve. Soit H un p-sous-groupe de G. Soit aussi S un p-Sylow de G — on sait qu'il en existe grâce au premier théorème de Sylow. On applique le lemme à cette situation : soit $g \in G$ tel que $gSg^{-1} \cap H$ soit un p-Sylow de H. Comme H est un p-groupe, il est son unique p-Sylow ; autrement dit, $gSg^{-1} \cap H = H$. En particulier, H est contenu dans le p-Sylow gSg^{-1}, ce qui démontre (i). Si en outre H est lui-même un p-Sylow, alors $H = gSg^{-1}$, ce qui montre que tout p-Sylow est conjugué à S : (ii) est démontré. ■

A noter
En particulier, le fait que les p-Sylow soient tous conjugués entraîne le résultat suivant. *Soient G un groupe, p un nombre premier et S un p-Sylow de G. Alors*

$$S \triangleleft G \iff S \text{ est l'unique } p\text{−Sylow de } G.$$

Théorème (troisième théorème de Sylow)

Soient G est un groupe fini et p un nombre premier. On note s_p le nombre de p-Sylow de G. Alors,

$$s_p \text{ divise } |G| \quad \text{et} \quad s_p \equiv 1\ [p].$$

PREUVE. Soient $\mathcal{S}$ l'ensemble des p-Sylow de G et $S \in \mathcal{S}$. Alors, $s_p = \operatorname{Card} \mathcal{S}$.
① On fait agir G sur $\mathcal{S}$ par conjugaison. Le deuxième théorème de Sylow assure que l'action est transitive. Autrement dit, l'action admet $\mathcal{S}$ tout entier pour unique orbite. L'équation aux classes montre alors que $s_p || G|$.
② On fait agir cette fois S sur $\mathcal{S}$ par conjugaison. Comme S est un p-groupe, l'ensemble $\mathcal{S}^S$ des points fixes de cette action vérifie $s_p \equiv \operatorname{Card}\left(\mathcal{S}^S\right)\ [p]$ — voir la proposition *Action d'un p-groupe*, page 48. Bien sûr, $S \in \mathcal{S}^S$. On montre que $\mathcal{S}^S = \{S\}$, ce qui permet de conclure. Pour tout $T \in \mathcal{S}$, on note $\operatorname{Norm}_G(T)$ le normalisateur de T dans G, assavoir $\operatorname{Norm}_G(T) = \left\{g \in G,\ gTg^{-1} = T\right\}$. Alors, pour tout $T \in \mathcal{S}$, le sous-groupe T de $\operatorname{Norm}_G(T)$ en est l'unique p-Sylow, puisque c'en est un p-Sylow distingué. Si en outre $T \in \mathcal{S}^S$, alors S est aussi un p-Sylow de $\operatorname{Norm}_G(T)$, ce qui impose que $S = T$. On a montré que S est l'unique point fixe de l'action. ■

Exemples d'application

(i) *Il n'y a pas de groupe simple d'ordre* 91.
En effet, soit G un groupe d'ordre 91. Alors, le nombre de 7-Sylow de G vérifie $s_7|13$ et $s_p \equiv 1\ [7]$, ce qui impose que $s_7 = 1$. Comme le conjugué d'un 7-Sylow est encore un 7-Sylow, on en déduit que l'unique 7-Sylow de G en est un sous-groupe distingué propre : G n'est pas simple.

(ii) L'unique 3-Sylow de $\mathfrak{S}_3$ est $\mathfrak{A}_3 = \langle(123)\rangle$. En revanche, $\mathfrak{S}_3$ contient trois 2-Sylow qui sont les sous-groupes engendrés par une transposition.

Exercice 56 Compter et décrire tous les Sylow de $\mathfrak{A}_4$, $\mathfrak{S}_4$, $\mathfrak{A}_5$, $\mathfrak{S}_5$.

Une application classique, rapidement

Proposition *Tout groupe simple d'ordre* 60 *est isomorphe à* $\mathfrak{A}_5$.

Une preuve : soit G un groupe simple d'ordre 60. Le nombre de ses 5-Sylow est 1 ou 6 ; c'est 6 puisque G est simple. L'action de G sur ses 5-Sylow par conjugaison fournit un homomorphisme $\varphi : G \to \mathfrak{S}_6$, injectif puisque l'action sur les 5-Sylow est transitive. En passant aux groupes dérivés, on obtient que $\varphi(D(G))$ est un sous-groupe de $D(\mathfrak{S}_6) = \mathfrak{A}_6$. Mais comme G est simple, $D(G) = G$. Donc l'image de φ est incluse dans $\mathfrak{A}_6$. Ainsi, G est isomorphe à un sous-groupe simple d'indice 6 de $\mathfrak{A}_6$.
Or, tout sous-groupe H simple d'indice 6 de $\mathfrak{A}_6$ est isomorphe à $\mathfrak{A}_5$.
En effet, on fait agir H sur les six classes à gauche de $\mathfrak{A}_6$ modulo H par translation à gauche : $h \cdot (\sigma H) = (h\sigma)H$. Cela fournit un homomorphisme de groupes $\psi : H \to \mathfrak{S}_6$, injectif puisque H est simple et puisque l'action n'est pas triviale — si elle était triviale, H serait un sous-groupe distingué propre de $\mathfrak{A}_6$ alors que ce dernier est simple. Le groupe d'isotropie H_H de la classe H contient H. Donc $\psi(H)$, qui est isomorphe à H, est contenu dans H_H. Or H_H est le fixateur dans $\mathfrak{S}_{(\mathfrak{A}_6/H)_g} \simeq \mathfrak{S}_6$ d'un point de $\{1, \ldots, 6\}$ — savoir H — : il est isomorphe à $\mathfrak{S}_5$. On a la situation suivante : H est un sous-groupe d'indice 2 de $H_H \simeq \mathfrak{S}_5$. Donc H est isomorphe à $\mathfrak{A}_5$.

Exercice 57 Démontrer le théorème de Cauchy page 49 à l'aide de la théorie de Sylow.

6 Polynômes symétriques

Dans tout le chapitre, les anneaux considérés sont commutatifs et unitaires.

6.1 Théorème des polynômes symétriques

Définition (action naturelle du groupe symétrique sur un anneau de polynôme)
Soient n un entier naturel non nul et $\mathcal{A}$ un anneau. Pour toute $\sigma \in \mathfrak{S}_n$ et pour tout $P \in \mathcal{A}[X_1, \ldots, X_n]$, on note $\sigma \cdot P$ ou simplement σP le polynôme

$$\sigma \cdot P(X_1, \ldots, X_n) = P\left(X_{\sigma(1)}, \ldots, X_{\sigma(n)}\right).$$

On voit immédiatement que cela définit une action (à gauche) de $\mathfrak{S}_n$ sur $\mathcal{A}[X_1, \ldots, X_n]$, qui vérifie en outre $\sigma \cdot (P+Q) = (\sigma \cdot P) + (\sigma \cdot Q)$ et $\sigma \cdot (P \times Q) = (\sigma \cdot P) \times (\sigma \cdot Q)$, pour tous $P, Q \in \mathcal{A}[X_1, \ldots, X_n]$.

Définition (polynôme symétrique)
Soient n un entier naturel non nul et $\mathcal{A}$ un anneau. Un polynôme P de $\mathcal{A}[X_1, \ldots, X_n]$ est dit *symétrique* lorsqu'il est invariant par toutes les permutations, c'est-à-dire lorsque $\sigma \cdot P = P$, pour tout $\sigma \in \mathfrak{S}_n$. On note

$$\mathcal{A}[X_1, \ldots, X_n]^{\mathfrak{S}_n}$$

le sous-anneau des polynômes symétriques de $\mathcal{A}[X_1, \ldots, X_n]$.

Exercice 58

(i) Vérifier que les polynômes symétriques forment bien un sous-anneau de l'anneau de tous les polynômes.

(ii) Un polynôme est symétrique si, et seulement s'il est invariant par toute transposition.

Exemples

(i) Pour tout $n \in \mathbb{N}^*$ et pour tout $p \in \mathbb{N}$, le *polynôme de Newton*

$$S_p = S_p(X_1, \ldots, X_n) = \sum_{k=1}^{n} X_k^p$$

est (évidemment) symétrique.

(ii) Dans $\mathbb{Z}[X, Y, Z]$, les polynômes $S_1 = X + Y + Z$, $\sigma_2 = XY + XZ + YZ$ et $S_2 = X^2 + Y^2 + Z^2$ sont symétriques et sont reliés par la relation $S_1^2 = S_2 + 2\sigma_2$.

Définition (polynômes symétriques élémentaires)
Soit $n \geq 1$. Pour tout $p \in \{0, \ldots, n\}$, on note $\sigma_p \in \mathbb{Z}[X_1, \ldots, X_n]$ le p^{e} *polynôme symétrique élémentaire à n indéterminées*, défini par le développement polynomial dans $\mathbb{Z}[X_1, \ldots, X_n, T]$ suivant :

$$\prod_{k=1}^{n} (1 + TX_k) = \sum_{p=0}^{n} \sigma_p(X_1, \ldots, X_n)\, T^p \tag{17}$$

A noter

(i) En particulier, $\sigma_0(X_1, \ldots, X_n) = 1$, σ_1 est la somme $\sigma_1(X_1, \ldots, X_n) = X_1 + \cdots + X_n$ et σ_n est le produit $\sigma_n(X_1, \ldots, X_n) = X_1 X_2 \ldots X_n$.

(ii) En faisant agir une permutation $\sigma \in \mathfrak{S}_n$ sur les membres de droite et de gauche de (17) dans l'anneau $\mathbb{Z}[T][X_1, \ldots, X_n]$, on vérifie que les σ_k sont bien des polynômes symétriques.

(iii) Une conséquence immédiate de (17) est la *relation entre racines et coefficients d'un polynôme*

$$\prod_{k=1}^{n} (T - X_k) = \sum_{p=0}^{n} (-1)^p \sigma_{n-p}(X_1, \ldots, X_n)\, T^p.$$

Si on donne ce nom à cette identité polynomiale, c'est parce que si $\mathcal{A}$ est un anneau intègre et si $P = \sum_{k=0}^{n} p_k T^k \in \mathcal{A}[T]$ est un polynôme unitaire de degré n admettant pour racines $x_1, \ldots, x_n \in \mathcal{A}$, alors les racines et les coefficients de P sont liés par les relations

$$\forall k \in \{1, \ldots n\},\ \sigma_k(x_1, \ldots, x_n)\, p_n = (-1)^{n-k} p_k.$$

Ainsi, les coefficients d'un polynômes sont des polynômes explicites en ses racines. Noter en passant que la démarche inverse qui consiste à exprimer les racines d'un polynôme en fonction de ses coefficients est autrement plus épineuse. C'est la question de la résolution des équations polynomiales. Par exemple, la question de la résolubilité des équations polynomiales "par radicaux" trouve une réponse dans le magnifique cadre de la théorie de Galois.

(iv) On peut aussi écrire les σ_k par une formule close, obtenue en développant (17). Cette formule est la suivante :

$$\begin{aligned}\sigma_p(X_1,\dots,X_n) &= \sum_{\substack{i_1,\dots,i_p\in\{1,\dots,n\}\\ i_1<i_2<\dots<i_p}} \prod_{k=1}^{p} X_{i_k}\\ &= X_1X_2\dots X_p + X_1X_2\dots X_{p+1} + \dots + X_{n-p+1}X_{n-p+2}\dots X_n\end{aligned}$$

C'est le polynôme homogène de degré p, somme de tous les monômes de degré p sans carré. Par exemple, $\sigma_3(X_1,X_2,X_3,X_4) = X_1X_2X_3 + X_1X_2X_4 + X_1X_3X_4 + X_2X_3X_4$.

Exercice 59 Montrer que le nombre de monômes de $\sigma_p(X_1,\dots,X_n)$ est le coefficient du binôme $\binom{n}{p}$.

Exemple
Le polynôme de $\mathbb{Z}[X_1,X_2,X_3,X_4]$

$$\begin{aligned}P = {} & X_1^2X_2X_3 + X_1^2X_2X_4 + X_1^2X_3X_4\\ & + X_2^2X_3X_4 + X_2^2X_3X_1 + X_2^2X_4X_1\\ & + X_3^2X_4X_1 + X_3^2X_4X_2 + X_3^2X_1X_2\\ & + X_4^2X_1X_2 + X_4^2X_1X_3 + X_4^2X_2X_3\end{aligned}$$

est symétrique, comme on le vérifie immédiatement en voyant qu'il est invariant par les six transpositions de $\mathfrak{S}_4$. Un calcul élémentaire montre que, pour ces quatre indéterminées, $\sigma_1\sigma_3 = P + 4\sigma_4$, ce qui permet d'exprimer P comme un polynôme en les σ_k, selon la formule $P = \sigma_1\sigma_3 - 4\sigma_4$.

Avant d'aborder la question des anneaux de polynômes invariants par le groupe symétrique ou par le groupe alterné, on prouve un lemme élémentaire dans les anneaux de polynômes généraux, qui sera bien utile pour prouver les théorèmes qui suivent.

Lemme
Soient n un entier naturel non nul, $\mathcal{A}$ un anneau commutatif unitaire et $P \in \mathcal{A}[X_1,\dots,X_n]$.
(i) Si $P(X_1,\dots,X_{n-1},0) = 0$, alors X_n divise P.
(ii) Si $X_1P = 0$, alors $P = 0$.
(iii) Si $X_1P \neq 0$, alors $P \neq 0$ et $\deg(X_1P) = 1 + \deg P$.
(iv) Si $X_1, X_2, \dots, X_k$ divisent P, alors le produit $X_1X_2,\dots X_k$ divise P, pour tout $k \in \{1,\dots,n\}$.
(v) Si $(X_1 - X_2)P = 0$, alors $P = 0$.
(vi) Si $X_1 - X_2$ et $X_1 - X_3$ divisent P, alors $(X_1 - X_2)(X_1 - X_3)$ divise P.
(vii) Si $X_1 - X_2$ et $X_3 - X_4$ divisent P, alors $(X_1 - X_2)(X_3 - X_4)$ divise P.

Preuve. Tous ces résultats sont immédiats si $\mathcal{A}$ est factoriel, puisqu'alors, $\mathcal{A}[X_1,\dots,X_n]$ l'est aussi, en vertu du théorème de transfert de Gauss. Ils restent cependant valides dans le cas général. (i) On fait la division euclidienne de P dans $\mathcal{A}[X_1,\dots,X_{n-1}][X_n]$ par le polynôme unitaire X_n, puis on spécialise $X_n = 0$. (ii) et (iii) On considère l'égalité $X_1P = 0$ dans l'anneau $\mathcal{A}[X_2,\dots,X_n][X_1]$. Vue ainsi, les résultats de (ii) et (iii) sont immédiats : la multiplication par X_1 n'est qu'un décalage des coefficients. (iv) En procédant par récurrence sur k (et n), il suffit de montrer le résultat pour $k = 2$. On suppose que X_1 et X_2 divisent P. Soit $Q \in \mathcal{A}[X_1,\dots,X_n]$ tel que $P = X_2Q$. On fait la division euclidienne de Q dans $\mathcal{A}[X_2,\dots,X_n][X_1]$ par le polynôme unitaire X_1. Soient $R \in \mathcal{A}[X_1,\dots,X_n]$ et $S \in \mathcal{A}[X_2,\dots,X_n]$ tels que $Q = X_1R + S$. Alors, $P = X_1X_2R + X_2S$. En spécialisant $X_1 = 0$, on obtient que $X_2S = 0$, et donc, en utilisant (ii), que $S = 0$. (v) Il suffit de le montrer pour $n = 2$, quitte à remplacer $\mathcal{A}[X_3,\dots,X_n]$ par $\mathcal{A}$. On explicite les coefficients de $P(X,Y) = \sum_{p,q\geq 0} a_{p,q}X^pY^q$, on écrit l'égalité $XP = YP$, on obtient des formules de récurrence sur les $a_{p,q}$ qui aboutissent directement au résultat. (vi) D'abord, $P = (X_1 - X_3)Q_1(X_1,X_2,\dots,X_n)$. Ensuite,

par division euclidienne de Q_1 dans $\mathcal{A}[X_2,\dots,X_n][X_1]$ par le polynôme unitaire X_1-X_2, on obtient $P=(X_1-X_3)(X_1-X_2)Q(X_1,X_2,\dots,X_n)+(X_1-X_3)R(X_2,\dots,X_n)$, ce qui implique en spécialisant $X_1=X_2$ que $(X_2-X_3)R(X_2,\dots,X_n)=0$. En appliquant (v), on conclut que $R=0$ et ainsi que $(X_1-X_2)(X_1-X_3)$ divise P. (vii) D'abord, $P=(X_3-X_4)Q_1(X_1,X_2,\dots,X_n)$. Ensuite, par division euclidienne de Q_1 dans $\mathcal{A}[X_2,\dots,X_n][X_1]$ par le polynôme unitaire X_1-X_2, on obtient $P=(X_3-X_4)(X_1-X_2)Q(X_1,X_2,\dots,X_n)+(X_3-X_4)R(X_2,\dots,X_n)$, ce qui implique en spécialisant $X_1=X_2$ que $(X_3-X_4)R(X_2,\dots,X_n)=0$. En appliquant (v), on conclut que $R=0$ et ainsi que $(X_1-X_2)(X_3-X_4)$ divise P. ■

Théorème (théorème des polynômes symétriques, version 1)
Soient n un entier naturel non nul et $\mathcal{A}$ un anneau commutatif unitaire. Alors, pour tout $P\in\mathcal{A}[X_1,\dots,X_n]^{\mathfrak{S}_n}$, il existe un unique $Q\in\mathcal{A}[X_1,\dots,X_n]$ tel que

$$P(X_1,\dots,X_n)=Q(\sigma_1,\dots,\sigma_n).$$

Preuve. Dans toute la preuve, si $k\in\{0,\dots,n\}$, on notera $(\sigma_k)_0$ le polynôme spécialisé

$$(\sigma_k)_0(X_1,\dots,X_{n-1})=\sigma_k(X_1,\dots,X_{n-1},0).$$

— on vérifie immédiatement que $(\sigma_1)_0$ est aussi le k^{e} polynôme symétrique élémentaire en les $n-1$ indéterminées $X_1,\dots,X_{n-1}$.
① On commence par l'existence. On procède par récurrence sur n. Si $n=1$, il n'y a rien à démontrer puisque $\sigma_1=X_1$. On suppose que $n\geq 2$ et que tout polynôme symétrique à $n-1$ indéterminées est un polynôme en les polynômes symétriques élémentaires en lesdites $n-1$ indéterminées. On montre par récurrence sur d que tout polynôme symétrique non nul de degré d de $\mathcal{A}[X_1,\dots,X_n]$ est un polynôme en $\sigma_1,\dots,\sigma_n$. Si $d=0$, il n'y a rien à faire. On suppose que $d\geq 1$ et que P est un polynôme symétrique de degré d de $\mathcal{A}[X_1,\dots,X_n]$. Alors, $P(X_1,\dots,X_{n-1},0)$ est un polynôme symétrique (de degré d) de $\mathcal{A}[X_1,\dots,X_{n-1}]$. Par hypothèse de récurrence, soit $Q_1\in\mathcal{A}[X_1,\dots,X_{n-1}]$ tel que $P(X_1,\dots,X_{n-1},0)=Q_1((\sigma_1)_0,\dots,(\sigma_{n-1})_0)$. On pose alors $P_1=P(X_1,\dots,X_n)-Q_1(\sigma_1,\dots,\sigma_{n-1})\in\mathcal{A}[X_1,\dots,X_n]^{\mathfrak{S}_n}$. Puisque $P_1(X_1,\dots,X_{n-1},0)=0$, le (i) du lemme montre que $X_n|P_1$. Puisque P_1 est symétrique, cela entraîne que tous les X_k divisent P. Ainsi, selon le (iv) du lemme, σ_n divise P. Soit donc $P_2\in\mathcal{A}[X_1,\dots,X_n]$ tel que $P_1=\sigma_nP_2$. L'application n fois du (ii) du lemme précédent montre alors que P_2 est également symétrique. Si $P_1=0$, on a montré que $P=Q_1(\sigma_1,\dots,\sigma_{n-1})$ et c'est fini. Si $P_1\neq 0$, alors $\deg(P_1)\leq d$ puisque $\deg(Q_1(\sigma_1,\dots,\sigma_{n-1}))=\deg(Q_1((\sigma_1)_0,\dots,(\sigma_{n-1})_0))$. Alors, l'application n fois du (iii) du lemme précédent montre que $\deg(P_2)\leq d-n$. On applique l'hypothèse de récurrence à P_2 : soit $Q_2\in\mathcal{A}[X_1,\dots,X_n]$ tel que $P_2=Q_2(\sigma_1,\dots,\sigma_n)$. Alors, on a montré que P s'écrit $P=Q_1(\sigma_1,\dots,\sigma_{n-1})+\sigma_nQ_2(\sigma_1,\dots,\sigma_n)$ qui est de la forme voulue.
② Unicité. Il suffit de montrer que $\forall P\in\mathcal{A}[X_1,\dots,X_n]$, $P(\sigma_1,\dots,\sigma_n)=0\Rightarrow P=0$. On procède par récurrence sur n. Si $n\geq 1$, il n'y a rien à démontrer puisque $P(\sigma_1)=P$. On suppose que $n\geq 2$. On suppose, par l'absurde, qu'il existe $P\neq 0$ tel que $P(\sigma_1,\dots,\sigma_n)=0$. Soit P un tel polynôme, de degré minimal. On ordonne $P=P_0+X_nP_1+\cdots+X_n^dP_d$ où $P_k\in\mathcal{A}[X_1,\dots,X_{n-1}]$ pour tout k. On substitue $X_1=\sigma_1,\dots,X_n=\sigma_n$ et on spécialise $X_n=0$. On obtient $0=P_0\left((\sigma_1)_0,\dots,(\sigma_1)_{n-1}\right)$. Par récurrence, cela entraîne que $P_0=0$ et donc que X_n divise P. Soit donc $Q\in\mathcal{A}[X_1,\dots,X_n]$ tel que $P=X_nQ$. D'après le lemme, $\deg(Q)=\deg(P)-1$. En outre, $\sigma_nQ(\sigma_1,\dots,\sigma_n)=0$. Le (ii) du lemme assure alors que $Q(\sigma_1,\dots,\sigma_n)=0$, ce qui contredit le caractère minimal du degré de P. ■

Définition (indépendance algébrique, transcendance)
Soient $\mathcal{B}$ un anneau, $\mathcal{A}$ un sous-anneau de $\mathcal{B}$, n un entier naturel non nul et $b_1,\dots,b_n\in\mathcal{B}$. On dit que $b_1,\dots,b_n$ sont *algébriquement indépendants sur $\mathcal{A}$*, ou que la famille $\{b_1,\dots,b_n\}$ est *algébriquement libre* lorsque

$$\forall P\in\mathcal{A}[X_1,\dots,X_n],\ P(b_1,\dots,b_n)=0\Longrightarrow P=0.$$

Sinon, les b_k sont *algébriquement dépendants sur $\mathcal{A}$*, ou encore la famille $\{b_1,\dots,b_n\}$ est *algébriquement liée*. Si $b\in\mathcal{B}$, on dit que b est *algébrique sur $\mathcal{A}$* lorsque $\{b\}$ est algébriquement liée. Sinon, on dit que b est *transcendant sur $\mathcal{A}$*.

Exemples
(i) Dans $\mathbb{C}$, tout élément est algébrique sur $\mathbb{R}$.

(ii) Si $z \in \mathbb{C}$, z est algébrique sur $\mathbb{Q}$ si, et seulement si le $\mathbb{Q}$-espace vectoriel $\mathbb{Q}[z]$ est de dimension finie.
(iii) Dans les conditions de la définition, $b_1, \ldots, b_n$ sont algébriquement indépendants si, et seulement si le sous-anneau $\mathcal{A}[b_1, \ldots, b_n]$ de $\mathcal{B}$ engendré par $\mathcal{A} \cup \{b_1, \ldots, b_n\}$ est isomorphe à l'anneau de polynômes $\mathcal{A} \cup [X_1, \ldots, X_n]$.

Notation
Dans les conditions de la définition des polynômes symétriques, on note $\mathcal{A}[\sigma_1, \ldots, \sigma_n]$ le sous-anneau de $\mathcal{A}[X_1, \ldots, X_n]$ engendré par $\mathcal{A} \cup \{\sigma_1, \ldots, \sigma_n\}$. Le théorème de caractérisation des sous-anneaux engendrés assure que

$$\mathcal{A}[\sigma_1, \ldots, \sigma_n] = \{P(\sigma_1, \ldots, \sigma_n),\ P \in A[X_1, \ldots, X_n]\}.$$

Théorème (théorème des polynômes symétriques, version 2)
Soient n un entier naturel non nul et $\mathcal{A}$ un anneau commutatif unitaire.
(i) $\mathcal{A}[X_1, \ldots, X_n]^{\mathfrak{S}_n} = \mathcal{A}[\sigma_1, \ldots, \sigma_n]$.
(ii) $\sigma_1, \ldots, \sigma_n$ sont algébriquement indépendants sur $\mathcal{A}[X_1, \ldots, X_n]$.

PREUVE. C'est une paraphrase de l'existence et de l'unicité de la version 1. ■

A noter
Dans les conditions du théorème des polynômes symétriques, on a la situation suivante :

$$\mathcal{A}[X_1, \ldots, X_n]^{\mathfrak{S}_n} = \mathcal{A}[\sigma_1, \ldots, \sigma_n] \simeq \mathcal{A}[X_1, \ldots, X_n].$$

6.2 Polynômes antisymétriques, polynômes invariants par le groupe alterné

Définition (polynôme invariant sous l'action du groupe alterné)
Soient n un entier naturel non nul et $\mathcal{A}$ un anneau. Un polynôme P de $\mathcal{A}[X_1, \ldots, X_n]$ est dit *invariant sous l'action du groupe alterné* lorsque $\sigma \cdot P = P$, pour tout $\sigma \in \mathfrak{A}_n$. On note

$$\mathcal{A}[X_1, \ldots, X_n]^{\mathfrak{A}_n}$$

le sous-anneau de $\mathcal{A}[X_1, \ldots, X_n]$ formé des polynômes invariants sous l'action du groupe alterné.

Exercice 60 $\mathcal{A}[X_1, \ldots, X_n]^{\mathfrak{A}_n}$ est bien un sous-anneau de $\mathcal{A}[X_1, \ldots, X_n]$. Il contient $\mathcal{A}[X_1, \ldots, X_n]^{\mathfrak{S}_n}$.

Exemple
Pour tout entier naturel non nul n, soit

$$V = V(X_1, \ldots X_n) = \prod_{\substack{(i,j) \in \{1,\ldots,n\}^2 \\ i<j}} (X_j - X_i)$$

le polynôme de Vandermonde. Son expression déterminantale montre immédiatement que pour tout $\sigma \in \mathfrak{S}_n$, $\sigma \cdot V = \varepsilon(\sigma) V$, où ε désigne la signature. En particulier, $V \in \mathbb{Z}[X_1, \ldots, X_n]^{\mathfrak{A}_n}$. L'action de $\mathfrak{S}_n$ sur V montre aussi que V^2 est un polynôme symétrique.

En deux indéterminées, $V(X,Y)^2 = (X-Y)^2 = (X+Y)^2 - 4(XY)^2 = \sigma_1^2 - 2\sigma_2$. Relier cette formule avec le discriminant du polynôme $(T-X)(T-Y)$, qui est précisément V^2. En d'autres termes, on dit que le discriminant du polynôme $T^2 - aT + b$ est $a^2 - 4b$.

Définition (discriminant)
Soit n un entier naturel non nul. Le *discriminant* à n indéterminées est le polynôme symétrique

$$\Delta(X_1, \ldots, X_n) = V(X_1, \ldots X_n)^2 = (-1)^{\frac{n(n-1)}{2}} \prod_{\substack{(i,j) \in \{1,\ldots,n\}^2 \\ i \neq j}} (X_i - X_j).$$

En appliquant le théorème des spolynômes symétriques, on note δ l'unique polynôme à n indéterminées tel que

$$\Delta(X_1, \ldots, X_n) = \delta(\sigma_1(X_1, \ldots, X_n), \ldots, \sigma_n(X_1, \ldots, X_n)).$$

Définition (polynôme antisymétrique)
Soient n un entier naturel non nul et $\mathcal{A}$ un anneau. Un polynôme P de $\mathcal{A}[X_1, \ldots, X_n]$ est dit *antisymétrique* lorsque $\sigma \cdot P = \varepsilon(\sigma) P$, pour toute $\sigma \in \mathfrak{S}_n$.

Proposition (structure des polynômes antisymétriques)
Soient n un entier naturel non nul, $\mathcal{A}$ un anneau commutatif unitaire dont la caractéristique est nulle ou impaire. Alors, les polynômes antisymétriques sont les polynômes de la forme VQ où $Q \in \mathcal{A}[X_1, \ldots, X_n]$ est un polynôme symétrique.

PREUVE. On suppose $n \geq 2$, sans quoi il n'y a rien à faire. Que ces polynômes soient antisymétriques est évident. Inversement, soit P un polynôme antisymétrique. On fait la division euclidienne de P dans l'anneau $\mathcal{A}[X_2, \ldots, X_n][X_1]$ par le polynôme unitaire $X_1 - X_2$: soient $Q \in \mathcal{A}[X_1, \ldots, X_n]$ et $R \in \mathcal{A}[X_2, \ldots, X_n]$ tels que $P = (X_1 - X_2)Q + R$. En spécialisant, $X_1 = X_2$, on obtient que $R = P(X_2, X_2, \ldots, X_n)$. Or, puisque P est antisymétrique, $(12) \cdot P = -P$ ce qui, en spécialisant $X_1 = X_2$ assure que $2P(X_2, X_2, \ldots, X_n) = 0$. Ainsi, $2R = 0$. Comme la caractéristique de $\mathcal{A}$ est nulle ou impaire, cela implique que $R = 0$. On a montré que $X_1 - X_2$ divise P. De la même façon, tous les $X_i - X_j$, $i \neq j$ divisent P. En faisant une récurrence qui utilise (vi) et (vii) du lemme précédent, on conclut que V divise P. Soit ainsi $Q \in \mathcal{A}[X_1, \ldots, X_n]$ tel que $P = VQ$. Il reste à montrer que Q est symétrique. Si $\tau \in \mathfrak{S}_n$ une transposition. On fait agir τ sur les polynômes de l'égalité $P = VQ$. On obtient $-P = -V \times \tau \cdot Q$, ce qui entraîne que $V(Q - \tau \cdot Q) = 0$. Par application répétée (récurrence) du (v) du lemme, on conclut que $Q - \tau \cdot Q = 0$ ce qui prouve que Q est symétrique. ■

Théorème (théorème des polynômes invariants par le groupe alterné)
Soient n un entier naturel non nul et $\mathcal{A}$ un anneau commutatif unitaire dans lequel le nombre 2 est inversible.
(i) $\mathcal{A}[X_1, \ldots, X_n]^{\mathfrak{A}_n} = \mathcal{A}[\sigma_1, \ldots, \sigma_n, V]$.
(ii) Plus précisément, tout élément de $\mathcal{A}[X_1, \ldots, X_n]^{\mathfrak{A}_n}$ s'écrit de manière unique sous la forme $P + QV$ où P et Q sont des polynômes symétriques.
(iii) L'anneau $\mathcal{A}[X_1, \ldots, X_n]^{\mathfrak{A}_n}$ est isomorphe à l'anneau- quotient

$$\mathcal{A}[X_1, \ldots, X_n]^{\mathfrak{A}_n} \simeq \mathcal{A}[X_1, \ldots, X_n, T] / \left(T^2 - \delta(X_1, \ldots X_n)\right).$$

PREUVE. Grâce au théorème des polynômes symétriques, (i) découle de (ii). On montre (ii).
① Unicité. Si $n = 1$, il n'y a rien à faire. On suppose $n \geq 2$ et que $P + QV = 0$ où P et Q sont symétriques. On fait agir la transposition (12) sur cette égalité. On obtient $P - QV = 0$, ce qui entraîne immédiatement que $2P = 0$ et $2QV = 0$. Puisque la caractéristique de $\mathcal{A}$ est nulle ou impaire, cela impose que $P = 0$ et $QV = 0$, ce qui implique $Q = 0$ par applications répétées (récurrence) des points (vi) et (vii) du lemme. On a montré l'unicité.
② Existence. Soit $A \in \mathcal{A}[X_1, \ldots, X_n]^{\mathfrak{A}_n}$. On note $P = A + (12) \cdot A$ et $Q' = A - (12) \cdot A$. On montre successivement que P' est symétrique, que Q' est antisymétrique. En particulier, grâce au théorème de structure des polynômes antisymétriques, il en résulte que $Q' = VQ$ où Q est symétrique. Comme 2 est inversible, puisque $A = \frac{1}{2}P + \frac{1}{2}QV$, le résultat s'en trouve démontré.
Le groupe $\mathfrak{S}_n$ est engendré par son sous-groupe $\mathfrak{A}_n$ et par la transposition (12) — plus précisément, on a la partition $\mathfrak{S}_n = \mathfrak{A}_n \cup (12)\mathfrak{A}_n$. Si $\sigma \in \mathfrak{A}_n$, puisque $\mathfrak{A}_n \lhd \mathfrak{S}_n$, il existe $\tau \in \mathfrak{A}_n$ tel que $\sigma(12) = (12)\tau$; ainsi, $\sigma \cdot P' = \sigma \cdot A + (12)\tau \cdot A = P'$ puisque A est invariant par σ et par τ. Par ailleurs, $(12) \cdot P' = (12) \cdot A + A = P'$: on a montré que P' est symétrique. De même, $\sigma \cdot Q' = Q'$ et $(12)Q' = -Q'$, ce qui suffit à prouver que Q' est antisymétrique.
(iii) Soit $s : \mathcal{A}[X_1, \ldots, X_n, T] \to \mathcal{A}[X_1, \ldots, X_n]^{\mathfrak{A}_n}$ la spécialisation $P \mapsto P(\sigma_1, \ldots, \sigma_n, V)$. Sa surjectivité est garantie par (i). Son noyau contient $T^2 - \delta$ par définition du discriminant. Inversement, soit $P \in \ker s$. On fait la division euclidienne de P par le polynôme unitaire $T^2 - \delta$ dans l'anneau $\mathcal{A}[X_1, \ldots, X_n][T]$: soient $Q \in \mathcal{A}[X_1, \ldots, X_n, T]$ et $R, S \in \mathcal{A}[X_1, \ldots, X_n]$ tels que $P = \left(T^2 - \delta\right)Q + RT + S$. En spécialisant *via* s, il vient $R(\sigma_1, \ldots, \sigma_n)V(X_1, \ldots, X_n) + R(\sigma_1, \ldots, \sigma_n) = 0$. L'unicité du (ii) et l'indépendance algébrique des σ_k entraînent alors que $R = S = 0$. On a montré que le noyau de s est l'idéal engendré par $T^2 - \delta$. On conclut avec le premier théorème d'isomorphisme pour les anneaux. ■

A noter
(i) Dans $\mathbb{Z}/4\mathbb{Z}[X, Y]$, on a l'égalité $2\sigma_1 + 2V = 2(X + Y) + 2(Y - X) = 0$ alors que $2(X + Y)$ est symétrique et non nul. L'unicité tombe en défaut sur cet anneau de caractéristique 4.
(ii) Si la caractéristique de $\mathcal{A}$ est nulle sans que 2 ne soit inversible, l'unicité subsiste.
En revanche, l'existence tombe en défaut. Par exemple, le polynôme $A = X^2Y + Y^2Z + Z^2X$ est invariant par le groupe $\mathfrak{A}_3$. Sa décomposition sur $\mathbb{Q}$ s'écrit $A = P + VQ$ où $Q = \frac{1}{2}$ et où $P = \frac{1}{2}(A + (12)A) =$

$\frac{1}{2}\left(X^2Y+Y^2Z+Z^2X+XY^2+YZ^2+ZX^2\right)$. Le polynôme $A\in\mathbb{Z}[X,Y,Z]^{\mathfrak{A}_3}$ n'a pas de décomposition sous la forme $P+VQ$ où $P,Q\in\mathbb{Z}[X,Y,Z]^{\mathfrak{S}_3}$.

(iii) Sur un corps (ou plus généralement sur un anneau factoriel) de caractéristique différente de 2, les arguments des preuves sont simplifiés par la factorialité de $\mathcal{A}[X_1,\ldots,X_n]$.

6.3 Séries formelles et formules de Newton

Définition (série formelle)
Soit $\mathcal{A}$ un anneau. Une *série formelle à coefficents dans* $\mathcal{A}$ est une suite (indexée par $\mathbb{N}$) d'éléments de $\mathcal{A}$. Si $(a_n)_{n\in\mathbb{N}}$ est une série formelle, on note

$$(a_n)_{n\in\mathbb{N}}=\sum_{n\geq 0}a_nX^n.$$

Les a_n sont les *coefficients* de la série formelle $\sum a_nX^n$. L'ensemble des séries formelles à coefficients dans $\mathcal{A}$ est noté

$$\mathcal{A}[[X]].$$

On définit sur l'ensemble des séries formelles deux lois de composition interne notées additivement et multiplicativement par les formules suivantes, qui miment l'addition et la multiplication des développements à l'origine des fonctions holomorphes (ou analytiques) : si $A=\sum a_nX^n$ et $B=\sum b_nX^n$, on définit

$$A+B=\sum_{n\geq 0}\left(a_n+b_n\right)X^n\quad\text{et}\quad AB=\sum_{n\geq 0}\left(\sum_{p+q=n}a_pb_q\right)X^n$$

où l'addition et la multiplication utilisées dans ces formules sont celles de l'anneau $\mathcal{A}$.

Exercice 61 Ces sommes sont bien définies : ces formules ont un sens.

Proposition (anneau des séries formelles)
Soit $\mathcal{A}$ un anneau (commutatif unitaire).
(i) L'addition et la multiplication ainsi définies confèrent à $\mathcal{A}[[X]]$ une structure d'anneau commutatif unitaire. Son zéro est la série formelle, notée 0, dont tous les coefficients sont nuls. Son unité est la série formelle, notée 0, dont tous les coefficients sont nuls à l'exception du premier qui est l'unité de $\mathcal{A}$.
(ii) Si $\mathcal{A}$ est intègre, alors $\mathcal{A}[[X]]$ est aussi intègre.

Preuve. Exercice. ■

A noter
(i) Les opérations dans l'anneau $\mathcal{A}[[X]]$ se comportent comme si la notation $\sum$ était une somme de série entière convergente. Autrement dit, les règles de calcul dans $\mathcal{A}[[X]]$ sont celles des séries entières convergentes.
(ii) On vérifie immédiatement que l'anneau des polynômes $\mathcal{A}[X]$ est un sous-anneau de $\mathcal{A}[[X]]$.

Exercice 62 (Substitution dans une série formelle)
Soient $\mathcal{A}$ un anneau et $Q\in\mathcal{A}[[X]]$ une série formelle *dont le terme constant est nul*. Alors, l'application

$$\begin{array}{ccc}\mathcal{A}[[X]] & \longmapsto & \mathcal{A}[[X]]\\ A=\sum_{n\geq 0}a_nX^n & \longmapsto & A\circ Q=\sum_{n\geq 0}a_nQ^n\end{array}$$

est bien définie et est un homomorphisme d'anneaux. C'est la *substitution de Q à l'indéterminée*.

Exemple Dans $\mathbb{Z}[[X]]$, le polynôme $1-X$ est inversible et son inverse est

$$(1-X)^{-1}=\sum_{n\geq 0}X^n.$$

Proposition (inverser les séries formelles à valuation non nulle)
Soit $\mathcal{A}$ un anneau et $Q \in \mathcal{A}[[X]]$. On suppose que le coefficient constant de Q est nul. Alors, $1-Q$ est inversible dans $\mathcal{A}[[X]]$ et

$$(1-Q)^{-1} = \frac{1}{1-Q} = \sum_{n\geq 0} Q^n.$$

PREUVE. Une fois le sens de la série formelle $\sum_{n\geq 0} Q^n$ assuré, il suffit de substituer Q à X dans la formule $(1-X)\left(\sum_{n\geq 0} X^n\right) = 1$. ■

Corollaire (développement des fractions rationnelles en séries formelles)
*Soient $\mathbb{F}$ un corps et $F \in \mathbb{F}(X)$. On suppose que 0 n'est pas un pôle de F (*i.e. *on suppose que X ne divise pas le dénominateur de F). Alors, F "est" une série formelle dans le sens suivant : il existe une unique $S_F \in \mathbb{F}[[X]]$ telle que, pour tous $N, D \in \mathbb{F}[X]$,*

$$F = \frac{N}{D} \Longrightarrow DS_F = N.$$

PREUVE. Puisque 0 n'est pas un pôle de F, soient $A, B \in \mathbb{F}[X]$ tels que $F = \frac{A}{B}$ et $B(0) \neq 0$. Alors, $B = B(0)(1-Q)$ où $Q \in \mathbb{F}[X]$ vérifie $Q(0) = 0$. Dans ces conditions, la série formelle B est inversible et son inverse est $B^{-1} = \frac{1}{B(0)} \sum_{n\geq 0} Q^n$. Alors, $S_F = AB^{-1}$ est l'unique série formelle qui convienne. ■

Notation Si $n \in \mathbb{N}^*$ et si $k \in \mathbb{N}$, dans $\mathbb{Z}[X_1, \ldots, X_n]$, on note $S_k = \sum_{i=1}^n X_i^k$ la k^{e} somme de Newton. Si $0 \leq k \leq n$, on note σ_k le k^{e} polynôme symétrique élémentaire ; si $k \geq n+1$, on note aussi $\sigma_k = 0$.

Proposition (formules de Newton)
Soient n un entier naturel non nul. Alors, pour tout $k \in \mathbb{N}$, dans l'anneau $\mathbb{Z}[X_1, \ldots, X_n]$,

$$\sum_{\substack{(p,q)\in\mathbb{N}^2 \\ p+q=k}} (-1)^q \sigma_q S_{p+1} = (-1)^k (k+1)\sigma_{k+1}.$$

PREUVE. Dans l'anneau de séries formelles $\mathbb{Z}[X_1, \ldots, X_n][[T]]$, soit

$$F = \prod_{k=1}^n (1 - TX_k) = \sum_{\ell\geq 0} (-1)^\ell \sigma_\ell(X_1, \ldots, X_n) T^\ell.$$

On calcule la dérivée logarithmique $S = F'/F \in \mathbb{Q}(X_1, \ldots, X_n, T)$ de F à partir de sa forme produit :

$$S = \frac{F'}{F} = \sum_{k=1}^n \frac{-X_k}{1-TX_k} = -\sum_{k=1}^n X_k \sum_{\ell\geq 0} T^\ell X_k^\ell = -\sum_{\ell\geq 0} S_{\ell+1} T^\ell.$$

Alors, les formules de l'énoncé sont les égalités terme à terme des coefficients de la série $FS = F'$: écrire

$$\left(\sum_{\ell\geq 0} (-1)^\ell \sigma_\ell(X_1, \ldots, X_n) T^\ell\right)\left(-\sum_{\ell\geq 0} S_{\ell+1} T^\ell\right) = \sum_{\ell\geq 0} (-1)^{\ell+1} \sigma_{\ell+1}(X_1, \ldots, X_n) T^\ell,$$

développer le produit, identifier les coefficients de T^k. ■

A noter
Avec des pointillés, ces formules s'écrivent

$$\begin{cases} S_1 = \sigma_1 \\ S_2 - \sigma_1 S_1 = -2\sigma_2 \\ S_3 - \sigma_1 S_2 + \sigma_2 S_1 = 3\sigma_3 \\ S_4 - \sigma_1 S_3 + \sigma_2 S_2 - \sigma_3 S_1 = -4\sigma_4 \\ \qquad\qquad\qquad \vdots \\ S_{n-1} - \sigma_1 S_{n-2} + \sigma_2 S_{n-3} - \cdots + (-1)^{n-1}\sigma_{n-1} S_1 = (-1)^{n-1} n\sigma_n \\ S_n - \sigma_1 S_{n-1} + \sigma_2 S_{n-2} - \cdots + (-1)^n \sigma_n S_1 = 0 \\ S_{n+1} - \sigma_1 S_n + \sigma_2 S_{n-1} - \cdots + (-1)^{n+1}\sigma_{n+1} S_1 = 0 \\ \qquad\qquad\qquad \vdots \end{cases}$$

Elles fournissent un système triangulaire (infini) qui permet de calculer les sommes de Newton en fonction des polynômes symétriques élémentaires. Ainsi, $S_2 = \sigma_1^2 - 2\sigma_2$, $S_3 = \sigma_1^3 - 3\sigma_1\sigma_2 + 3\sigma_3$, $S_4 = \sigma_1^4 - 4\sigma_1^2\sigma_2 + 4\sigma_1\sigma_3 + 2\sigma_2^2$, etc.

Exercice 63
(i) Calculer $S_3 = X^3 + Y^3$ et $S_4 = X^4 + Y^4$ en fonction de $X + Y$ et XY dans $\mathbb{Z}[X, Y]$.
(ii) Calculer $S_4 = X^4 + Y^4 + Z^4$ et $S_5 = X^5 + Y^5 + Z^5$ en fonction de $X + Y + Z$, $XY + XZ + YZ$ et XYZ dans $\mathbb{Z}[X, Y, Z]$.

Exercice 64
Montrer que le déterminant du système linéaire, obtenu à partir des formules de Newton, dont les inconnues sont les S_k, $1 \leq k \leq n$ et les paramètres les σ_k, $1 \leq k \leq n$, a pour déterminant 1. Montrer que le déterminant du système linéaire dont les inconnues sont les σ_k, $1 \leq k \leq n$ et les paramètres les S_k, $1 \leq k \leq n$, a pour déterminant $n!$.

Proposition (les sommes de Newton forment une base algébrique des polynômes symétriques)
Soient n un entier naturel non nul, $\mathbb{F}$ un corps de caractéristique nulle*, et $S_1, \ldots, S_n$ les n premières sommes de Newton dans $\mathbb{F}[X_1, \ldots, X_n]$. Alors,*

$$\mathbb{F}[X_1, \ldots, X_n]^{\mathfrak{S}_n} = \mathbb{F}[S_1, \ldots, S_n]$$

et $S_1, \ldots, S_n$ sont algébriquement indépendants.

Preuve. Grâce au théorème des polynômes symétriques, il suffit de montrer que $\mathbb{F}[S_1, \ldots, S_n] = \mathbb{F}[\sigma_1, \ldots, \sigma_n]$. Les formules de Newton et l'exercice précédent fournissent la clef de l'argumentation. ■

A noter
(i) S'assurer de bien comprendre à quel endroit l'hypothèse sur le corps intervient.
(ii) Plus précisément,

$$\mathbb{Z}[S_1, \ldots, S_n] \subsetneq \mathbb{Z}[X_1, \ldots, X_n]^{\mathfrak{S}_n} = \mathbb{Z}[\sigma_1, \ldots, \sigma_n] \subsetneq \mathbb{Q}[\sigma_1, \ldots, \sigma_n] = \mathbb{Q}[S_1, \ldots, S_n] = \mathbb{Q}[X_1, \ldots, X_n]^{\mathfrak{S}_n}.$$