COURS D'ALGÈBRE AU MAGISTÈRE DE CACHAN

Marc HINDRY, Université Paris 7.
hindry@math.jussieu.fr

En un semestre A, B, C et D ont été traitées et il a été fait allusion aux parties E et F.

Quelques références choisies.

J'utiliserai beaucoup et je recommande comme référence (en particulier pour l'algèbre à l'agrégation) le

Cours d'algèbre, D. Perrin (collection Ellipses)

Sauf les parties D et F, ce livre traite tous les thèmes abordés dans ce cours. Parmi les traités généraux d'algèbre traitant également l'algèbre linéaire je signale

Algebra, S. Lang (collection Addison-Wesley), très dense et riche.

Cours d'algèbre, R. Godement (collection Hermann), la partie cours est du niveau des deux premières années d'université mais les exercices permettent d'aller au niveau licence-maîtrise.

Algebra, M. Artin (collection Prentice-Hall), très pédagogique et attrayant.

Algebra, Birkhoff & MacLane (collection Chelsea), un classique.

Pour approfondir la notion d'action de groupes, les applications à la géométrie citons

Eléments de géométrie, R. Mneimné (collection Cassini) foisonnant et instructif.

Pour une introduction aux représentations de groupes finis, il est difficile de surpasser

Représentations linéaires des groupes finis, J-P. Serre (collection Hermann).

Pour finir, je recommande de jeter un coup d'oeil au volume de l'encyclopédie russe

Basic notions of algebra, I. Shafarevic (collection Springer).

A. GROUPES ET ACTIONS ET GROUPES

Une présentation des groupes, de leurs quotients avec des exemples. La notion centrale présentée est celle d'action de groupe.

A.1. Généralités sur les groupes.

Définition. Un *groupe* est la donnée d'un ensemble G et d'une loi interne $G \times G \to G$ vérifiant
(i) (élément neutre) Il existe $e \in G$ tel que, pour tout $g \in G$, on ait $e * g = g * e = g$.
(ii) (associativité) Pour tout $g, g', g'' \in G$, on a $(g * g') * g'' = g * (g' * g'')$.
(iii) (inverse d'un élément) Pour tout $g \in G$, il existe $g' \in G$ tel que, $g' * g = g * g' = e$.

Remarques. L'ensemble G s'appelle l'*ensemble sous-jacent*; par abus de langage, on parlera du groupe G, sous-entendant ainsi la loi que l'on notera le plus souvent comme un produit; l'inverse de g sera alors noté g^{-1}. Lorsque la loi vérifie de plus $g * g' = g' * g$, on dira que le groupe est *commutatif* ou *abélien* et l'on notera alors parfois la loi comme une addition et l'inverse de g s'écrira $-g$.

Exemples. Vous connaissez déjà bien sûr des groupes comme $\mathbf{Z}$, $\mathbf{Z}/n\mathbf{Z}$ (munis de l'addition), ou $\mathcal{S}_n$ (le groupe des permutations sur n éléments) ou $\mathrm{GL}(n, \mathbf{R})$, le groupe des matrices de taille $n \times n$ inversibles à coefficients réels. Comme exemple initial, ajoutons l'ensemble des transformations linéaires préservant une figure dans le plan, l'espace ou plus généralement $\mathbf{R}^n$; ces transformations sont d'ailleurs des isométries. Concrètement l'ensemble des transformations linéaires du plan préservant un polygone régulier à n côtés est un groupe noté D_n (dont on montre ci-dessous qu'il est de cardinal $2n$); l'ensemble des transformations linéaires du plan préservant un cube est un groupe (dont on peut montrer qu'il est de cardinal 48);

Premiers calculs. Dans un groupe, "on peut toujours simplifier", c'est-à-dire que $xy = xz$ entraîne $y = z$. En effet il suffit de multiplier par x^{-1} :

$$y = ey = (x^{-1}x)y = x^{-1}(xy) = x^{-1}(xz) = (x^{-1}x)z = ez = z.$$

L'inverse de x^{-1} est x et l'inverse de xy est $y^{-1}x^{-1}$, en effet :

$$(xy)(y^{-1}x^{-1}) = x(yy^{-1})x^{-1} = xex^{-1} = xx^{-1} = e.$$

Définition.

$$x^n := \begin{cases} e & \text{si } n = 0 \\ \underbrace{x \dots x}_{(n \text{ fois})} & \text{si } n > 0 \\ \underbrace{x^{-1} \dots x^{-1}}_{(|n| \text{ fois})} & \text{si } n < 0 \end{cases}$$

On a $x^m.x^n = x^{m+n}$ et $(x^m)^n = x^{mn}$. Si $y = gxg^{-1}$ alors $y^n = gx^ng^{-1}$.

Un sous-ensemble H d'un groupe G est un *sous-groupe* si la loi de groupe sur G induit une loi de groupe sur H. C'est-à-dire si H est stable par multiplication, passage à l'inverse et contient l'élément neutre (l'associativité est alors automatique). On voit facilement que cette condition équivaut à dire que $e \in H$ et que $x, y \in H$ entraîne $xy^{-1} \in H$. De même il est immédiat de montrer que l'intersection de sous-groupes est un sous-groupe.

Si S est un sous-ensemble d'un groupe G on définit le *sous-groupe engendré par S* comme le plus petit sous-groupe de G contenant S, i.e. l'intersection de tous les sous-groupes contenant H. C'est un exercice facile de voir que c'est aussi l'ensemble des produits $x_1^{\epsilon_1} \cdots x_r^{\epsilon_r}$ avec $r \geq 0$, $x_i \in S$ et $\epsilon_i = \pm 1$.

Soit G_1 et G_2 deux groupes. On définit le *produit de groupes* qui a comme ensemble sous-jacent $G_1 \times G_2$ par la loi de composition :

$$(g_1, g_2) * (g_1', g_2') := (g_1 g_1', g_2 g_2').$$

Une application $f : G_1 \to G_2$ entre deux groupes est un *homomorphisme de groupes* si elle vérifie

$$\forall x, y \in G_1,\ f(xy) = f(x)f(y);$$

c'est un *isomorphisme* si elle est bijective, un *automorphisme* si de plus $G_1 = G_2$. On appelle *noyau* le sous-groupe $\mathrm{Ker}(f) = \{x \in G_1 \mid f(x) = e\}$ et *image* le sous-groupe $f(G_1) = \{y \in G_2 \mid \exists x \in G_1,\ f(x) = y\}$. Il est immédiat que le composé d'homomorphismes (resp. d'isomorphismes, resp. d'automorphismes) est encore un homomorphisme (resp. un isomorphisme, resp. un automorphisme). En particulier l'ensemble des automorphismes d'un groupe G est un groupe que l'on notera $\mathrm{Aut}(G)$. Remarquons aussi que la bijection réciproque d'un isomorphisme est automatiquement un homomorphisme.

Exemples. L'application $x \mapsto x^2$ est un homomorphisme de groupes si et seulement si le groupe G est *abélien* (i.e. commutatif). Soit $x \in G$, l'application $\phi_x : G \to G$ définie par $\phi_x(y) := xyx^{-1}$ est un automorphisme appelé *automorphisme intérieur* de G; de plus l'application $x \mapsto \phi_x$ de G dans $\mathrm{Aut}(G)$ est un homomorphisme de groupes. L'ensemble des images par automorphisme intérieur d'un élément $y \in G$ s'appelle la *classe de conjugaison* de y.

Décrivons maintenant l'exemple cité plus haut de groupe d'origine géométrique: le groupe dièdral D_n.

Théorème. *Le groupe des isométries planes d'un polygone régulier à n côtés ($n \geq 3$), de centre O a pour cardinal $2n$; il contient n rotations, les rotations d'angle $2k\pi/n$ et de centre O et n symétries, les symétries orthogonales fixant les droites passant par O et un sommet ou le milieu d'une arête.*

Preuve. On voit facilement que les isométries décrites dans l'énoncé laissent invariant le polygone, il s'agit de démontrer que ce sont les seules. Pour cela on va utiliser le lemme suivant (dont on laisse la preuve en exercice) :

Lemme. *Soit s une isométrie plane laissant invariant un polygone régulier à n côtés, de centre O et sommets $A_1, \ldots, A_n$ alors*

- *Si s fixe deux sommets adjacents, alors s est l'identité;*
- *Si s fixe un sommet A_i, alors s est soit l'identité soit la symétrie par rapport à la droite OA_i.*

Soit maintenant σ une isométrie du polygone, il existe une rotation r d'angle $2k\pi/n$ telle que $r \circ \sigma(A_1) = A_1$ (en effet ces rotations permutent circulairement les sommets). Donc, d'après le lemme, ou bien $r \circ \sigma = id$ et alors σ est une rotation d'angle $-2k\pi/n$ ou bien $r \circ \sigma$ est la symétrie s_1 par rapport à OA_1 et $\sigma = r^{-1} \circ s_1$. Cela suffit pour voir que $\mathrm{card}(D_n) = 2n$ et permet de vérifier (indirectement) que $r^{-1} \circ s_1$ est une des symétries décrites. □

Remarques. Les rotations forment un sous-groupe de D_n isomorphe à $\mathbf{Z}/n\mathbf{Z}$. Si r est une rotation et s une symétrie, alors $srs^{-1} = srs = r^{-1}$ (vérifiez-le). On peut utiliser cela pour montrer que le centre de D_n est trivial si n est impair et d'ordre 2 (engendré par la rotation d'angle π) si n est pair. On peut aussi interpréter D_2 comme le groupe des isométries planes laissant invariant un segment (il est isomorphe à $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$).

A.2. Quotient d'un groupe par un sous-groupe.

On intoduit les notations suivantes, si A et B sont des parties d'un groupe G. On pose $A.B := \{a.b \mid a \in A, b \in B\}$ et de même $A^{-1} := \{a^{-1} \mid a \in A\}$. On écrira $g.A$ pour $\{g\}.A$

Soit H un sous-groupe de G, on définit deux relations d'équivalence par

$$x\mathcal{R}y \Leftrightarrow xH = yH \Leftrightarrow y^{-1}x \in H$$
$$x\mathcal{R}'y \Leftrightarrow Hx = Hy \Leftrightarrow xy^{-1} \in H$$

On notera G/H l'ensemble quotient $G/\mathcal{R}$ (resp. $H\backslash G$ l'ensemble quotient $G/\mathcal{R}'$). Vérifions, par exemple, que $\mathcal{R}$ est une relation d'équivalence. On a $x^{-1}x = e \in H$ donc $x\mathcal{R}x$. Si $x\mathcal{R}y$ alors $y^{-1}x \in H$ donc $x^{-1}y = (y^{-1}x)^{-1} \in H$ et $y\mathcal{R}x$. Si $x\mathcal{R}y$ et $y\mathcal{R}z$ alors $y^{-1}x \in H$ et $z^{-1}y \in H$ donc $z^{-1}x = (z^{-1}y)(y^{-1}x) \in H$ et $x\mathcal{R}z$.

Remarque. Hormis ces relations d'équivalence "jumelles", la seule autre relation d'équivalence "intéressante" est la relation de conjugaison : $x\mathcal{R}y$ si il existe $g \in G$ avec $y = gxg^{-1}$. Les classes d'équivalence pour cette relation s'appelle naturellement *classes de conjugaison*.

Il faut faire attention qu'en général $gH \neq Hg$ (on verra plus loin que l'égalité n'est vraie pour tout g que si le sous-groupe H est distingué). Par contre la transformation $A \mapsto A^{-1}$ envoie gH sur Hg^{-1} donc il y a une bijection naturelle entre G/H et $H\backslash G$. Remarquons ensuite que les classes d'équivalence ont toutes le même cardinal que H. En effet l'application de H vers gH (resp. $H.g$) qui, à x associe gx (resp. xg) est visiblement une bijection. On a ainsi démontré le théorème suivant

Théorème. (Lagrange) *Soit G un groupe et H un sous-groupe, alors* $\mathrm{card}(G/H) = \mathrm{card}(H\backslash G)$ *et*

$$\mathrm{card}(G) = \mathrm{card}(H)\,\mathrm{card}(G/H).$$

Exemples. On tire facilement que si $x \in G$ et G fini, alors l'ordre de g divise $\mathrm{card}(G)$. Ainsi, comme $(\mathbf{Z}/p\mathbf{Z})^*$ a pour cardinal $p-1$ on en tire que, pour a entier premier avec p, on a $a^{p-1} \equiv 1 \bmod p$, ou encore que pour tout entier $a^p \equiv a \bmod p$ ("petit théorème" de Fermat). Plus généralement, si on note $\phi(n) = \mathrm{card}(\mathbf{Z}/n\mathbf{Z})^*$ on obtient que, pour a entier premier avec n, on a $a^{\phi(n)} \equiv 1 \bmod p$ (théorème d'Euler).

Définition. Un sous-groupe H de G est *distingué* si, pour tout $g \in G$, on a $H = gHg^{-1}$.

Remarquons qu'il est équivalent de demander que, pour tout $g \in G$, on ait $gH = Hg$ ou encore que, pour tout $g \in G$, on ait $H \subset gHg^{-1}$. Par ailleurs, le noyau d'un homomorphisme $f : G \to G'$ est toujours distingué; en effet si $y \in \mathrm{Ker}(f)$ alors $f(xyx^{-1}) = f(x)f(y)f(x)^{-1} = f(x)e'f(x)^{-1} = e'$ donc $xyx^{-1} \in \mathrm{Ker}(f)$.

Proposition. *L'intersection de sous-groupes distingués est un sous-groupe distingué. Si $f : G \to G'$ est un homomorphisme de groupes et si $H' \triangleleft G'$ alors $f^{-1}(H') \triangleleft G$; si $H \triangleleft G$ alors $f(H) \triangleleft f(G)$.*

Preuve. Immédiat. □

Remarquons que dans la dernière partie de la proposition, on ne peut pas conclure que $f(H)$ est distingué dans G', sauf si f est surjective.

Le principal intérêt des sous-groupes distingués est le suivant.

Proposition. *Soit H un sous-groupe de G. Il existe une structure de groupe sur l'ensemble G/H telle que la surjection canonique $s : G \to G/H$ soit un homomorphisme si et seulement si le sous-groupe H est distingué.*

Preuve. Supposons qu'une telle structure existe sur G/H alors H est le noyau de l'homomorphisme $s : G \to G/H$ donc est distingué dans G. Supposons inversement H distingué dans G, on est amené à définir une loi sur G/H par la formule $(xH) * (yH) = xyH$ (pour que s soit un homomorphisme) et le point est de vérifier que cette formule est bien définie, i.e. que si $x' \in xH$ et $y' \in yH$ alors $x'y'H = xyH$. Or on a bien, puisque H est distingué et $x' = xh$, $y' = yh'$, l'égalité $x'y'H = xhyh'H = xhyH = xhHy = xHy = xyH$. L'application $s : G \to G/H$ est surjective et vérifie donc $s(x) * s(y) = s(xy)$; on en tire immédiatement que G/H muni de la loi $*$ est un groupe. □

Théorème. (Propriété universelle du quotient) *Soit $f : G \to G'$ un homomorphisme de groupes. Soit H un sous-groupe et $s : G \to G/H$ la surjection canonique. Il existe une application $\hat{f} : G/H \to G'$ telle que*

$f = \hat{f} \circ c$ si et seulement si $H \subset \mathrm{Ker}(f)$. Dans ce cas, si de plus H est un sous-groupe distingué (et donc G/H un groupe), alors $\hat{f}$ est un homomorphisme de groupes, $\hat{f}(G/H) = f(G)$ et $\mathrm{Ker}(\hat{f}) = \mathrm{Ker}(f)/H$.

Preuve. La condition ensembliste garantissant l'existence de $\hat{f}$ est que $s(x) = s(y)$ entraîne $f(x) = f(y)$. Or $s(x) = s(y)$ équivaut à $xH = yH$ ou encore $x^{-1}y \in H$ alors que $f(x) = f(y)$ équivaut à $f(x^{-1}y) = e'$ ou encore $x^{-1}y \in \mathrm{Ker}(f)$. La deuxième partie est immédiate sauf peut-être la détermination du noyau de $\hat{f}$. Soit xH un élément de G/H qui soit dans le noyau de $\hat{f}$ alors $f(x) = \hat{f}(xH) = e'$ donc $x \in \mathrm{Ker}(f)$ d'où l'égalité $\mathrm{Ker}(\hat{f}) = \mathrm{Ker}(f)/H$. □

Corollaire. *Soit $f : G \to G'$ un homomorphisme de groupe, alors $f(G) \cong G/\mathrm{Ker}(f)$.*

Preuve. On applique la propriété universelle avec $H = \mathrm{Ker}(f)$ alors $\mathrm{Ker}(\hat{f}) = \mathrm{Ker}(f)/\mathrm{Ker}(f)$ est trivial donc $\hat{f}$ injective. □

Applications. a) Le sous-groupe $\langle x \rangle$ engendré par un élément $x \in G$ est isomorphe soit à $\mathbf{Z}$ (on dira que x est d'*ordre infini*) soit à $\mathbf{Z}/n\mathbf{Z}$ avec $n \geq 1$ (on dira que x est d'*ordre* n). En effet d'après le corollaire appliqué à l'homomorphisme défini par $f(m) := x^m$ de $\mathbf{Z}$ vers $\langle x \rangle \subset G$, on a $\langle x \rangle \cong \mathbf{Z}/\mathrm{Ker}(f)$.

b) Le noyau de l'homorphisme $G \to \mathrm{Aut}(G)$ qui a un élément associe l'automorphisme intérieur associé est le centre de G, noté $Z(G)$; si l'on note $\mathrm{Int}(G)$ le groupe des automorphismes intérieurs, on a donc $\mathrm{Int}(G) \cong G/Z(G)$.

A.3. Action de groupe.

La notion suivante est fondamentale; d'une part les groupes apparaissent naturellement dans la plupart des problèmes à travers leurs actions (ou représentations) et d'autre part, pour étudier les groupes eux-mêmes, on verra qu'il est souvent avantageux de les faire agir.

Définition. Une *action* d'un groupe G sur un ensemble X est une application $\Phi : G \times X \to X$ telle que
(i) $\Phi(e, x) = x$.
(ii) $\Phi(g, \Phi(g', x)) = \Phi(gg', x)$.

Remarque. Il est équivalent de se donner un homomorphisme $\rho : G \to \mathrm{Bij}(X)$. La correspondance est donnée par

$$\rho(g)(x) = \Phi(g, x).$$

On abrègera en général $\Phi(g, x)$ en $g.x$.

Exemple. Si ϕ est une bijection de X sur X, alors $\mathbf{Z}$ agit sur X par l'action $n \cdot x = \phi^n(x)$. Le groupe $\mathrm{GL}(2, \mathbf{R})$ agit naturellement sur $\mathbf{R}^2$; voici une action moins évidente. Choisissons $G = \mathrm{SL}(2, \mathbf{R})$ et $\mathcal{H} := \{z \in \mathbf{C} \mid \mathrm{Im}(z) > 0\}$ le demi-plan de Poincaré, l'application suivante est une action de groupe:

$$\begin{array}{ccc} G \times \mathcal{H} & \to & \mathcal{H} \\ \begin{pmatrix} a & b \\ c & d \end{pmatrix}, z & \mapsto & \frac{az+b}{cz+d} \end{array}$$

Une action définit une relation d'équivalence

$$x \mathcal{R} y \Leftrightarrow \exists g \in G,\ y = g.x$$

dont les classes d'équivalence $G.x = \{g.x \mid g \in G\}$ s'appellent les *orbites* de l'action. L'ensemble quotient $X/\mathcal{R}$ sera noté X/G, l'orbite de x sera notée $\mathcal{O}(x)$.

Définitions. Le *stabilisateur* d'un élément $x \in X$ est le sous-groupe de G des éléments qui fixe x, i. e. $G_x = \{g \in G \mid g \cdot x = x\}$. Le *noyau* d'une action est l'intersection des stabilisateurs de tous les points (c'est aussi le noyau de l'homomorphisme associé). Une action est dite *fidèle* si son noyau est trivial, *transitive* s'il n'y a qu'un orbite.

Exemples. Le noyau de l'action de $\mathrm{SL}(2,\mathbf{R})$ sur $\mathcal{H}$ donnée ci-dessus est $\pm I$, l'action de $\mathrm{SL}(2,\mathbf{R})$ est transitive, le stabilisateur de $i \in \mathcal{H}$ est $\mathrm{SO}(2,\mathbf{R}) = \left\{ \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \mid a,b \in \mathbf{R},\ a^2+b^2=1 \right\}$.

Remarques. Si on dispose d'une action $G \times X \to X$, on peut lui associer les actions suivantes:
(a) Pour tout sous-groupe H de G, une action de H sur X.
(b) Si $K = \cap_{x \in X} G_x$ est le noyau de l'action, alors on hérite d'une action de G/K sur X qui est fidèle.
(c) Si $\mathcal{P}(X)$ (resp. $\mathcal{P}_n(X)$) désigne l'ensemble des parties de X (resp. l'ensemble des parties de cardinal n) alors onpeut définir une action de G sur $\mathcal{P}(X)$ (resp. $\mathcal{P}_n(X)$) par $g \cdot A = \{g \cdot a \mid a \in A\}$.

Formule des classes (1ère forme).

$$\mathrm{card}(X) = \sum_{C \in X/G} \mathrm{card}(C)$$

Formule des classes (2ème forme).

$$\mathrm{card}(\mathcal{O}(x)) = \mathrm{card}(G/G_x).$$

En effet, considérons l'application $f : G \to \mathcal{O}(x)$ définie par $f(g) = g.x$. On a alors $f(g) = f(g')$ si et seulement si $g.x = g'.x$ ou encore $x = (g^{-1}g').x$ ou encore $g^{-1}g' \in G_x$ ou encore $gG_x = g'G_x$. Ainsi, d'après la propriété universelle du quotient, f passe au quotient pour donner une bijection $\hat{f} : G/G_x \to \mathcal{O}(x)$. On en tire

Théorème. (Formule des classes) *Soit G fini agissant sur X fini et soit R un système d'éléments de X représentant les classes de X/G, alors*

$$\mathrm{card}(X) = \sum_{x \in R} \mathrm{card}(G/G_x) = \sum_{x \in R} \frac{\mathrm{card}(G)}{\mathrm{card}(G_x)}.$$

On note X^G l'ensemble des points fixes, c'est-à-dire

$$X^G = \{x \in X \mid \forall g \in G,\ g \cdot x = x\} = \{x \in X \mid G_x = G\}$$

Un groupe de cardinal une puissance d'un nombre premier p sera appelé un p-groupe.

Corollaire. *Soit G un p-groupe agissant sur un ensemble fini X alors*

$$|X^G| \equiv |X| \bmod p.$$

En particulier, si $|X|$ n'est pas divisible par p, il existe un point fixe.

Preuve. On écrit la formule des classes en observant que l'orbite d'un point fixe est, bien sûr, réduite à un point et que les autres orbites ont pour cardinal $(G : G_x)$ avec $G_x \neq G$ donc ce cardinal est divisible par p et $|X| \equiv \sum_{x \in X^G} 1 \bmod p$. □

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

Preuve. Soit G un p-groupe. Considérons l'action de G sur lui-même par conjugaison (i.e. $X = G$ et $g \cdot x = gxg^{-1}$). On voit aisément que $X^G = Z(G)$ et donc $|Z(G)|$ est divisible par p d'après le corollaire précédent. □

Exercice. Montrer que si $(G : H) = p$ est le plus petit nombre premier divisant $\mathrm{card}(G)$ alors H est distingué dans G. (Indication : considérer l'action de G sur G/H par translation, introduire l'homomorphisme associé $\rho : G \to \mathcal{S}_p = \mathrm{Bij}(G/H)$ et montrer que $H = \mathrm{Ker}(\rho)$).

A.4. Théorèmes de Sylow.

Le théorème suivant recense essentiellement ce que l'on peut dire d'un groupe fini en ne connaissant que son cardinal.

Théorème. (Sylow) *Soit p un nombre premier et G un groupe de cardinal $p^r m$ avec m non divisible par p.*
(i) Il existe un sous-groupe P de cardinal p^r (un tel sous-groupe s'appelle un p-sous-groupe de Sylow de G).
(ii) Soit H un p-sous-groupe et P un p-sous-groupe de Sylow de G, alors il existe $g \in G$ tel que $H \subset gPg^{-1}$. En particulier deux p-sous-groupes de Sylow de G sont conjugués.
(iii) Soit n_p le nombre de p-sous-groupes de Sylow de G. Alors $n_p \equiv 1 \bmod p$ et n_p divise m.

Preuve. Il s'agit de variations sur le thème des actions de groupes et de la formule des classes.
(i) Considérons l'action de G sur lui-même par translation et l'action induite sur $X = \mathcal{P}_{p^r}(G)$. Si R désigne un ensemble des représentants des classes d'équivalence, on a par la formule des classes

$$C_{mp^r}^{p^r} = |X| = \sum_{A \in R} (G : G_A).$$

Admettons provisoirement (voir lemme ci-dessous) que p ne divise pas $|X|$. Alors il existe une orbite, disons celle de A_0 de cardinal premier avec p. On a donc $(G : G_{A_0})$ non divisible par p donc $|G_{A_0}|$ est divisible par p^r. Mais par ailleurs, si l'on choisit $a_0 \in A_0$, on peut considérer l'application $G_{A_0} \to A$ définie par $g \mapsto ga_0$ qui est clairement injective donc $|G_{A_0}|$ est majoré par p^r et divisible par p^r donc égal à p^r. Ainsi G_{A_0} est un p-sous-groupe de Sylow. La preuve sera complète grâce au lemme

Lemme. *Soit m non divisible par un nombre premier p, alors*

$$C_{mp^r}^{p^r} \equiv m \bmod p.$$

On peut démontrer cela directement, en effet

$$C_{mp^r}^{p^r} = \frac{(mp^r)!}{(p^r)!(mp^r - p^r)!} = m \prod_{k=1}^{p^r-1} \left(\frac{mp^r - k}{p^r - k} \right).$$

Or si $k = p^s \ell$ alors $(mp^r - k)(p^r - k)^{-1} = (mp^{r-s} - \ell)(p^{r-s} - \ell)^{-1} \equiv 1 \bmod p$ d'où le lemme.

Une deuxième preuve du lemme consiste à appliquer la formule des classes précédente avec $G = \mathbf{Z}/p^r\mathbf{Z} \times \mathbf{Z}/m\mathbf{Z}$, vérifier que $\mathbf{Z}/p^r\mathbf{Z} \times \{0\}$ est le seul sous-groupe à p^r éléments et que les seules parties à p^r éléments qu'il laisse stable sont les $\mathbf{Z}/p^r\mathbf{Z} \times \{x\}$; toutes ces parties forment une orbite unique de cardinal m, les autres parties vérifient $(G : G_A) \equiv 0 \bmod p$ et donc on a bien $C_{mp^r}^{p^r} = |\mathcal{P}_{p^r}(G)| \equiv m \bmod p$.
(ii) Soit P un p-sous-groupe de Sylow (dont l'existence est maintenant garantie) et H un p-sous-groupe de G. Nous faisons agir H sur G/P par la formule $(h, gP) \mapsto hgP$. Comme le cardinal de G/P n'est pas divisible par p et que H est un p-groupe, on en déduit l'existence d'un point fixe. Donc il existe $g_0 \in G$ tel que pour tout $h \in H$ on ait $hg_0P = g_0P$ ou encore $hg_0 \in g_0P$ ou encore $h \in g_0Pg_0^{-1}$. Ainsi $H \subset g_0Pg_0^{-1}$; si de plus H est un p-sous-groupe de Sylow, on a donc égalité.
(iii) Notons $X = Sy\ell_p$ l'ensemble des p-sous-groupes de Sylow de G et n_p son cardinal. Si $P \in X$ alors gPg^{-1} est de nouveau un p-sous-groupe de Sylow de G. On dispose ainsi d'une action par conjugaison de G sur X qui est transitive d'après le résultat précédent. Si P est un p-sous-groupe de Sylow de G, on a clairement $P \subset G_P$ puisque P est un sous-groupe, par conséquent

$$n_p = (G : G_P) = \frac{|G|}{(G_P : P)|P|} = \frac{m}{(G_P : P)}.$$

Ainsi n_p divise m. Considérons maintenant l'action de P sur $Sy\ell_p$, toujours par conjugaison. L'élément P est visiblement fixe; nous allons montrer qu'il est l'unique point fixe et nous pourrons alors conclure que

$$n_p \equiv |Sy\ell^P| \equiv 1 \bmod p$$

Soit donc $Q \in Sy\ell^P$ et introduisons $G_0 = \langle P, Q \rangle$ le sous-groupe engendré par P et Q (argument dit "de Frattini"). On constate que P et Q sont encore deux p-sous-groupes de Sylow de G_0 et par conséquent

sont conjugués dans G_0 : il existe $y \in G_0$ tel que $P = yQy^{-1}$ mais Q est fixé par P (par hypothèse) et, bien sûr, est fixé par Q donc par G_0 et on peut conclure que $P = Q$. □

Corollaire. *Soit G un groupe fini. Il existe un élément d'ordre p dans G si et seulement si p divise* $\mathrm{card}(G)$.

Preuve. La nécessité provient du théorème de Lagrange. Supposons que p divise $\mathrm{card}(G)$, alors il existe un p-sous-groupe non trivial H (par exemple un p-sous-groupe de Sylow) et $y \in H \setminus \{e\}$. L'élément y est d'ordre une puissance de p, disons p^r avec $r \geq 1$. On voit immédiatement que l'élément $x = y^{p^{r-1}}$ est d'ordre p. □

A.5. Produit semi-direct.

Nous voulons expliciter un peu et illustrer par des exemples la notion de produit semi-direct. Il s'agit d'étudier la situation suivante où l'on a un groupe G et deux sous-groupes K et H tels que :
(i) $H \cap K = \{e\}$
(ii) Tout élément g de G s'écrit $g = kh$ avec $k \in K$ et $h \in H$ (ou encore $K.H = G$).
Par exemple, si (i) est réalisée et si, de plus, G est fini et $\mathrm{card}(G) = \mathrm{card}(K)\,\mathrm{card}(H)$ alors la condition (ii) est aussi vérifiée.

Ces hypothèses entraînent que l'application

$$\begin{array}{ccc} K \times H & \stackrel{f}{\longrightarrow} & G \\ (k,h) & \longmapsto & kh \end{array}$$

est une bijection. Si on suppose de plus que les éléments de H et K commutent, alors f est un isomorphisme de groupe : on a $G \cong K \times H$. C'est évidemment faux en général et l'objet de ce paragraphe est d'étudier le groupe G dans l'hypothèse où l'un des sous-groupes, disons K est distingué dans G. On verra qu'alors G est isomorphe à un groupe que l'on peut fabriquer à partir de K et H, l'ensemble sous-jacent étant $K \times H$ mais la loi de groupe étant différente de la loi de groupe produit. On dira que G est un produit *semi-direct*.

Exercice. Vérifier les assertions suivantes concernant l'application $f : K \times H \to G$ donnée par $(k,h) \mapsto kh$. a) l'application est injective si et seulement si $H \cap K = \{e\}$. b) le sous-ensemble $K.H$ n'est pas toujours un sous-groupe (donner un contre-exemple). c) Si K est distingué (ou H) alors $K.H$ est un sous-groupe. d) f est un isomorphisme entre $K \times H$ et $K \cdot H$ si et seulement si les éléments de K commutent avec ceux de H et $H \cap K = \{e\}$. e) si $H \cap K = \{e\}$ et les deux sous-groupes sont distingués, alors f est un isomorphisme.

Exemples.

1. Considérons dans le groupe $\mathcal{S}_3$ les sous-groupes $K = \mathcal{A}_3$ et $H = \{id, (1,2)\}$ alors on a bien $K \lhd G$ et $H \cap K = \{e\}$ ainsi que $\mathcal{S}_3 = K.H$ mais $\mathcal{S}_3$ n'est pas isomorphe à $K \times H$ (qui est commutatif).

2. Soit D_n le groupe (de cardinal $2n$) des isométries d'un polygone régulier à n côtés. La rotation ρ de centre O le centre du polygone et d'angle $2\pi/n$ engendre un sous-groupe K distingué dans D_n et d'ordre n. Une symétrie s par rapport à une droite passant par O et un sommet engendre un sous-groupe H d'ordre 2 et on a $D_n = K.H$ et $H \cap K = \{e\}$. Cependant D_n n'est pas isomorphe à $K \times H$ (qui est commutatif).

3. Soit $Aff = Aff(\mathbf{R}^n)$ le groupe des transformations affines, c'est-à-dire:

$$Aff := \{f : \mathbf{R}^n \to \mathbf{R}^n \mid f(X) = AX + b\ ,\ A \in \mathrm{GL}(n, \mathbf{R})\ , b \in \mathbf{R}^n\}$$

Rappelons que $\mathrm{GL}(n, \mathbf{R})$ désigne le groupe des matrices $n \times n$ inversibles à coefficient dans le corps $\mathbf{R}$. Le sous-groupe des translations $K = \{f \in Aff \mid f(X) = X + b\}$ est distingué dans Aff et le sous-groupe des applications linéaires $H = \{f \in Aff \mid f(0) = 0\}$ est tel que $Aff = K.H$ et $H \cap K = \{id\}$; cependant Aff n'est pas isomorphe à $K \times H$.

Nous allons maintenant construire et définir le produit semi-direct et voir que ces trois exemples sont des produits semi-directs.

Premier point de vue (description). On suppose $K \lhd G$ et les conditions (i) et (ii) vérifiées. Pour décrire le groupe G, on utilise la bijection $f : K \times H \to G$ pour définir une nouvelle loi de groupe sur l'ensemble $K \times H$; on pose $(k,h) * (k',h') = f^{-1}\left(f(k,h).f(k',h')\right)$. On vérifie alors immédiatement que $f((k,h) * (k',h')) = f(k,h).f(k',h')$. On peut calculer $*$ en observant que $(kh).(k'h') = k(hk'h^{-1})hh'$ et que $hk'h^{-1} \in K$ puisque K est distingué dans G. Si l'on note $\phi_h(x) = hxh^{-1}$ on obtient :

$$(k,h) * (k',h') = (k\phi_h(k'), hh') \tag{1}$$

Ceci suggère que, inversement, on puisse reconstruire le groupe G comme l'ensemble $K \times H$ muni de la loi définie par (1) *; nous allons voir qu'il en est bien ainsi.*

Deuxième point de vue (construction). On considère deux groupes K et H avec un homomorphisme $\phi : H \to \mathrm{Aut}(K)$ (ainsi H agit sur K) ; on définit sur l'ensemble $K \times H$ la loi de composition :

$$(k,h) * (k',h') = (k\phi(h)(k'), hh') \tag{2}$$

Remarquons que cette loi est la loi de groupe produit "ordinaire" si et seulement si ϕ est l'homomorphisme "trivial" : $\forall h \in H,\ \phi(h) = id$.

Théorème : Soient K, H deux groupes et $\phi : H \to \mathrm{Aut}(K)$ un homomorphisme de groupes.
1) L'ensemble $K \times H$ muni de la loi $* = *_\phi$ définie par

$$(k,h) * (k',h') = (k\phi_h(k'), hh')$$

est un groupe, appelé produit semi-direct de K et H relativement à ϕ ; il se note

$$K \times_\phi H \quad \text{ou} \quad K \rtimes_\phi H.$$

2) Un groupe G est isomorphe à $K \rtimes_\phi H$ si et seulement si il contient deux sous-groupes K' et H' avec $K \cong K' \triangleleft G$ et $H \cong H'$ de sorte que l'action de H' sur K' par automorphismes intérieurs corresponde à l'homomorphisme $\phi : H \to \mathrm{Aut}(K)$.
(Remarque : la notation est faite pour rappeler que K est distingué dans le grand groupe).

Preuve. 1) L'élément neutre est (e, e') (où e est l'élément neutre de K et e' celui de H); l'inverse de (h,k) est $(\phi(h^{-1})(k^{-1}), h^{-1})$. On vérifie enfin l'associativité :

$$((k,h) * (k',h')) * (k'',h'') = (k\phi(h)(k'), hh') * (k'',h'') = (k\phi(h)(k')\phi(hh')(k''), hh'h'')$$

alors que

$$(k,h) * ((k',h') * (k'',h'')) = (k\phi(h)(k'\phi(h')(k'')), hh'h'') = (k\phi(h)(k')\phi(hh')(k''), hh'h'')$$

2) La discussion précédant le théorème montre que si G contient K', H' comme indiqués alors $G \cong K' \rtimes_\phi H'$. Inversement le sous-groupe $K' := K \times \{e'\}$ est distingué dans $K \rtimes_\phi H$ et l'action de $H' = \{e\} \times H$ sur K' par automorphismes intérieurs est donnée par ϕ puisque:

$$(e,h) * (k,e') * (e,h)^{-1} = (\phi(h)(k), h) * (e, h^{-1}) = (\phi(h)(k), e').$$

□

Exercices.

a) Montrer que $K \rtimes_\phi H$ est commutatif si et seulement si K et H sont commutatifs et ϕ trivial (produit "direct").

b) Plus généralement, décrire le centre de $K \rtimes_\phi H$ en terme de ϕ et des centres de K et H.

c) Soient H, K deux sous-groupes distingués de G avec $K \cap H = \{e\}$, montrer que les éléments de K et et H commutent et en déduire que le groupe engendré par H et K est isomorphe à $K \times H$.

Illustrations. Reprenons les trois exemples du début et explicitons ϕ sur chacun de ces exemples.

1. Notons $\tau = (1,2)$ et $\rho = (1,2,3) \in \mathcal{S}_3$ alors $\tau\rho\tau^{-1} = (2,1,3) = (1,3,2) = \rho^{-1}$ donc la conjugaison par τ agit sur $\mathcal{A}_3 = \{id, (1,2,3), (1,3,2)\} = \{id, \rho, \rho^{-1}\}$ comme $j : x \mapsto x^{-1}$. Si l'on pose $\phi(\tau) = j$, $\phi(id) = id$ on obtient un homomorphisme $\phi : H \to \mathrm{Aut}(\mathcal{A}_3)$ tel que

$$\mathcal{S}_3 \cong \mathcal{A}_3 \rtimes_\phi \mathbf{Z}/2\mathbf{Z}$$

Exercice. Décrire un homomorphisme $\phi : \mathbf{Z}/2\mathbf{Z} \to \mathrm{Aut}(\mathcal{A}_n)$ tel que $\mathcal{S}_n \cong \mathcal{A}_n \rtimes_\phi \mathbf{Z}/2\mathbf{Z}$.

2. Si ρ est la rotation plane d'angle $2\pi/n$, de centre l'origine, et s une symétrie (par rapport à la bissectrice d'un des angles formés par les côtés d'un polygone régulier à n côtés), un calcul laissé en exercice montre que $s\rho s^{-1} = s\rho s = \rho^{-1}$. Si l'on désigne par ϕ l'homomorphisme (de $\mathbf{Z}/2\mathbf{Z}$ dans $\mathrm{Aut}(\mathbf{Z}/n\mathbf{Z})$) qui associe à 1 l'automorphisme $\phi(1) : x \mapsto -x$ on obtient :

$$D_n \cong \mathbf{Z}/n\mathbf{Z} \rtimes_\phi \mathbf{Z}/2\mathbf{Z}.$$

3. Considérons l'action de $\mathrm{GL}(2, \mathbf{R})$ sur $\mathbf{R}^2$ donnée par $(A, X) \mapsto AX$ (où A est une matrice inversible et X un vecteur de $\mathbf{R}^2$) ; cette action induit un homomorphisme $\phi : \mathrm{GL}(2, \mathbf{R}) \to \mathrm{Aut}(\mathbf{R}^2)$ et on obtient ainsi :

$$Aff \cong \mathbf{R}^2 \rtimes_\phi \mathrm{GL}(2, \mathbf{R})$$

On voit qu'il est important en général de déterminer le groupe d'automorphisme d'un groupe K pour étudier ensuite les homomorphismes $H \to \mathrm{Aut}(K)$ et les produits semi-directs associés ; c'est en général assez difficile et nous le ferons ici seulement dans le cas des groupes finis abéliens de la forme $K = (\mathbf{Z}/n\mathbf{Z})^r$.

Proposition *L'application $f \mapsto f(1)$ induit un isomorphisme de groupes* $\mathrm{Aut}(\mathbf{Z}/n\mathbf{Z}) \cong (\mathbf{Z}/n\mathbf{Z})^*$.

En effet soit $f \in \mathrm{Aut}(\mathbf{Z}/n\mathbf{Z})$ alors $x = f(1)$ doit être un générateur de $\mathbf{Z}/n\mathbf{Z}$ et f est entièrement déterminé par $f(1)$ (puisque $f(n) = nx$). Inversement si $x \in (\mathbf{Z}/n\mathbf{Z})^*$ alors $f(n) = nx$ définit bien un automorphisme de $\mathbf{Z}/n\mathbf{Z}$. Enfin on vérifie sans difficulté que si f et g sont des automorphismes de $\mathbf{Z}/n\mathbf{Z}$ alors $(g \circ f)(1) = g(1)f(1)$.

Proposition. *Le groupe* $\mathrm{Aut}\left((\mathbf{Z}/p\mathbf{Z})^r\right)$ *est isomorphe au groupe* $\mathrm{GL}(r, \mathbf{Z}/p\mathbf{Z})$ *des matrices inversibles de taille $r \times r$ à coefficients dans* $\mathbf{Z}/p\mathbf{Z}$.

Preuve. En effet un homomorphisme de groupe f de $(\mathbf{Z}/p\mathbf{Z})^r$ vers $(\mathbf{Z}/p\mathbf{Z})^r$ est forcément $\mathbf{Z}/p\mathbf{Z}$-linéaire puisque $f(nx) = nf(x)$. Dire que f est bijectif équivaut à dire que la matrice associée est inversible. □

Exercice. Montrer que $\mathrm{Aut}\left((\mathbf{Z}/n\mathbf{Z})^r\right)$ est isomorphe au groupe $\mathrm{GL}(r, \mathbf{Z}/n\mathbf{Z})$ des matrices inversibles de taille $r \times r$ à coefficients dans $\mathbf{Z}/n\mathbf{Z}$. Plus généralement, pouvez-vous décrire le groupe $\mathrm{Aut}\left((\mathbf{Z}/p\mathbf{Z})^r \times (\mathbf{Z}/p^2\mathbf{Z})^s\right)$ ou encore $\mathrm{Aut}\left((\mathbf{Z}/p\mathbf{Z})^{r_1} \times \ldots \times (\mathbf{Z}/p^m\mathbf{Z})^{r_m}\right)$?

Application. Nous allons déterminer les classes d'isomorphismes de groupes de cardinal pq où p et q sont des nombres premier distincts.

Supposons $p < q$, alors les théorèmes de Sylow nous indique qu'il existe un unique q-Sylow (on ne peut avoir $p \equiv 1 \bmod q$) que l'on désignera par K ; appelons H un p-Sylow et $\phi : H \to \mathrm{Aut}(K)$ l'action par conjugaison de H sur K. Comme K est isomorphe à $\mathbf{Z}/q\mathbf{Z}$ et H est isomorphe à $\mathbf{Z}/p\mathbf{Z}$, on a $H \cap K = \{e\}$ et ensuite $G = K.H$ et $G \cong K \rtimes_\phi H$. On doit donc étudier les homomorphismes de $\mathbf{Z}/p\mathbf{Z}$ dans $\mathrm{Aut}(\mathbf{Z}/q\mathbf{Z}) = (\mathbf{Z}/q\mathbf{Z})^*$. On doit alors distinguer deux cas.

1er cas : $q \not\equiv 1 \bmod p$. Dans ce cas le seul homomorphisme $\phi : \mathbf{Z}/p\mathbf{Z} \to \mathrm{Aut}(\mathbf{Z}/q\mathbf{Z}) = (\mathbf{Z}/q\mathbf{Z})^*$ est trivial donc $G \cong \mathbf{Z}/q\mathbf{Z} \times \mathbf{Z}/p\mathbf{Z} \cong \mathbf{Z}/pq\mathbf{Z}$

2ème cas : $q \equiv 1 \bmod p$. Dans ce cas, le groupe $\mathrm{Aut}(\mathbf{Z}/q\mathbf{Z}) = (\mathbf{Z}/q\mathbf{Z})^*$ contient des éléments d'ordre p et il y a donc un homomorphisme non trivial $\phi : \mathbf{Z}/p\mathbf{Z} \to \mathrm{Aut}(\mathbf{Z}/q\mathbf{Z}) = (\mathbf{Z}/q\mathbf{Z})^*$ et l'on peut donc fabriquer le produit semi-direct $G \cong \mathbf{Z}/q\mathbf{Z} \rtimes_\phi \mathbf{Z}/p\mathbf{Z}$. On obtient ainsi deux groupes de cardinal pq non isomorphes (on laisse en exercice, voir lemme à la fin, la vérification du fait que deux homomorphismes non triviaux $\phi : \mathbf{Z}/p\mathbf{Z} \to \mathrm{Aut}(\mathbf{Z}/q\mathbf{Z}) = (\mathbf{Z}/q\mathbf{Z})^*$ induisent des groupes isomorphes).

Remarque : Lorsque $p = 2$, on a toujours $q \equiv 1 \bmod 2$ et on retrouve les deux groupes $\mathbf{Z}/2q\mathbf{Z}$ et D_q. Lorsque $p = 3$ on s'aperçoit par exemple qu'il n'y a qu'un groupe de cardinal 15 (c'est $\mathbf{Z}/15\mathbf{Z}$) alors qu'il y en a deux de cardinal 21 (ce sont $\mathbf{Z}/21\mathbf{Z}$ et le produit semi-direct $\mathbf{Z}/7\mathbf{Z} \rtimes_\phi \mathbf{Z}/3\mathbf{Z}$).

Exercices. Soit p un nombre premier impair, on se propose de décrire les groupes de cardinal p^2 et p^3.

A1) Soit G un groupe de cardinal p^2, montrer que, ou bien G est cyclique (et isomorphe à $\mathbf{Z}/p^2\mathbf{Z}$), ou bien tous les éléments différents de l'élément neutre sont d'ordre p.

A2) Soit G un groupe non cyclique d'ordre p^2, soit K un sous-groupe d'ordre p, montrer que $K \triangleleft G$ et qu'il existe H sous-groupe d'ordre p tel que $K \cap H = \{e\}$. En déduire que G est un produit semi-direct de $\mathbf{Z}/p\mathbf{Z}$ par $\mathbf{Z}/p\mathbf{Z}$.

A3) Montrer que tout groupe de cardinal p^2 est commutatif et isomorphe à $\mathbf{Z}/p^2\mathbf{Z}$ ou $\mathbf{Z}/p\mathbf{Z} \times \mathbf{Z}/p\mathbf{Z}$.

B1) Montrer que la matrice $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ est un élément d'ordre p de $\mathrm{GL}(2, \mathbf{Z}/p\mathbf{Z})$. En déduire que l'application ϕ de $\mathbf{Z}/p\mathbf{Z}$ dans $\mathrm{GL}(2, \mathbf{Z}/p\mathbf{Z})$ définie par $m \mapsto T^m$ est un homomorphisme de groupes.

B2) Vérifier que le groupe $G_1 = (\mathbf{Z}/p\mathbf{Z})^2 \rtimes_\phi \mathbf{Z}/p\mathbf{Z}$ (où ϕ est défini comme en B1) est de cardinal p^3, n'est pas commutatif bien que tous ses éléments distincts de e soient d'ordre p.

C1) Montrer que la classe de $p+1$ modulo p^2 est d'ordre p dans le groupe $\left(\mathbf{Z}/p^2\mathbf{Z}\right)^*$ et en déduire que l'application ϕ de $\mathbf{Z}/p\mathbf{Z}$ dans $(\mathbf{Z}/p^2\mathbf{Z})^*$ définie par $m \mapsto (1+p)^m$ est un homomorphisme de groupes.

C2) Montrer que le groupe $G_2 = \mathbf{Z}/p^2\mathbf{Z} \rtimes_\phi \mathbf{Z}/p\mathbf{Z}$ (où ϕ est défini comme en C1) n'est ni commutatif ni isomorphe à G_1.

D) On donne des indications pour montrer qu'un groupe de cardinal p^3 est isomorphe à l'un des cinq groupes suivants (N.B. p est supposé impair) :

$$G_1 = (\mathbf{Z}/p\mathbf{Z})^2 \rtimes_\phi \mathbf{Z}/p\mathbf{Z},\ G_2 = \mathbf{Z}/p^2\mathbf{Z} \rtimes_\phi \mathbf{Z}/p\mathbf{Z},\ G_3 = (\mathbf{Z}/p\mathbf{Z})^3,\ G_4 = \mathbf{Z}/p^2\mathbf{Z} \times \mathbf{Z}/p\mathbf{Z}$$

$$\text{ou } G_5 = \mathbf{Z}/p^3\mathbf{Z}$$

Si G non cyclique, montrer que ou bien il existe K sous-groupe cyclique de cardinal p^2 ou bien tous les éléments ($\neq e$) sont d'ordre p et alors il existe K sous-groupe isomorphe à $(\mathbf{Z}/p\mathbf{Z})^2$; dans les deux cas $K \triangleleft G$ avec $G/K \cong \mathbf{Z}/p\mathbf{Z}$. Montrer (c'est la partie difficile) qu'il existe un sous-groupe H de cardinal p tel que $K \cap H = \{e\}$ et en déduire que $G \cong K \rtimes_\phi \mathbf{Z}/p\mathbf{Z}$ et conclure en étudiant les homomorphismes de $\mathbf{Z}/p\mathbf{Z}$ vers $\mathrm{Aut}(K) = (\mathbf{Z}/p^2\mathbf{Z})^*$ ou $\mathrm{GL}(2, \mathbf{Z}/p\mathbf{Z})$. Si tous les éléments sont d'ordre p il n'y a pas de difficulté (et le fait que p est impair n'intervient pas) sinon choisir x générateur de $K \cong \mathbf{Z}/p^2\mathbf{Z}$ et $y \notin K$. On montre qu'il existe c tel que $y' = x^c y$ soit d'ordre p et alors le sous-groupe engendré par y' fournit le sous-groupe H cherché. Pour cela on observe qu'il existe a, b tels que

$$yxy^{-1} = x^a \qquad \text{et} \qquad y^p = x^b$$

parce que $< x >= K$ est distingué et parce que G/K est d'ordre p. On observe que $b = pb'$ car $e = y^{p^2} = x^{pb}$ et $a \equiv 1 \bmod p$ car $a^p \equiv 1 \bmod p^2$. On en tire d'abord que $x^c y^m = y^m x^{a^m c}$ puis enfin que

$$(x^c y)^p = x^{b + c(a^{p-1} + \ldots + a + 1)}.$$

On doit alors résoudre l'équation $b + c(a^{p-1} + \ldots + a + 1) = 0$ dans $\mathbf{Z}/p^2\mathbf{Z}$ sachant que p ne divise pas a mais divise b. Il suffit que $a^{p-1} + \ldots + a + 1 \not\equiv 0 \bmod p^2$. On a $a = 1 + pr$ et donc $a^{p-1} + \ldots + a + 1 \equiv p + p^2 r(p-1)/2 \equiv p \not\equiv 0 \bmod p^2$ (ceci est vrai car p est *impair* !).

Pour la vérification que tous les produits semi-directs non triviaux sont isomorphes à G_1 ou G_2 voir l'application du lemme plus loin.

E) Le groupe des matrices 3×3 triangulaires supérieures avec des 1 sur la diagonale à coefficients dans $\mathbf{Z}/p\mathbf{Z}$ est non commutatif de cardinal p^3. Si p est impair, est-il isomorphe à "G_1" ou "G_2" (Cf. exercice précédent); si $p = 2$, est-il isomorphe à "D_4" ou "$\mathbf{H}$" (Cf. exercice suivant).

Attention : Si K est un sous-groupe distingué de G, il n'est pas toujours vrai que G soit isomorphe à $K \rtimes_\phi G/K$; pour cela il faut qu'il existe un sous-groupe H tel que la surjection canonique $s : G \to G/K$ donne un isomorphisme $H \to G/K$.

Exemple. Soit K l'unique sous-groupe de cardinal p dans $G = \mathbf{Z}/p^2\mathbf{Z}$. On a G/H isomorphe à $\mathbf{Z}/p\mathbf{Z}$ mais G n'est pas isomorphe à un produit demi-direct $H \rtimes_\phi \mathbf{Z}/p\mathbf{Z}$. Comme autre exemple moins trivial on peut considérer le groupe quaternionique

$$\mathbf{H} := \{+1, -1, +i, -i, +j, -j, +k, -k\}.$$

On vérifie que tous les sous-groupes sont distingués mais si on prend un des sous-groupes d'ordre 4 (engendré par $\pm i$, $\pm j$ ou $\pm k$) le quotient de $\mathbf{H}$ par ce sous-groupe est $\mathbf{Z}/2\mathbf{Z}$ sans que l'on puisse écrire $\mathbf{H}$ comme produit semi-direct.

Exercice. Montrer qu'il y a 5 groupes d'ordre 8 (à isomorphisme près): $\mathbf{Z}/8\mathbf{Z}$, $\mathbf{Z}/4\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$, $(\mathbf{Z}/2\mathbf{Z})^3$, D_4 (le groupe des isométries du carré) et $\mathbf{H}$ (le groupe quaternionique précédent).

Supplément : isomorphismes entre produits semi-directs.

Lemme. *Soit K, H deux groupes, $\alpha \in \mathrm{Aut}(K)$ et $\beta \in \mathrm{Aut}(H)$; soient $\phi, \psi : H \to \mathrm{Aut}(K)$ tels que $\phi(h) = \alpha^{-1} \circ \psi(\beta(h)) \circ \alpha$ alors l'application (α, β) induit un isomorphisme de groupe de $K \rtimes_\phi H$ vers $K \rtimes_\psi H$.*

Preuve. On calcule, en notant $F = (\alpha, \beta)$ pour alléger

$$\begin{aligned} F\left((k,h) *_\phi (k',h')\right) &= (\alpha(k\phi(h)(k')), \beta(hh')) \\ &= (\alpha(k)\alpha \circ (\alpha^{-1} \circ \psi(\beta(h)) \circ \alpha)(k'), \beta(h)\beta(h')) \\ &= (\alpha(k)\psi(\beta(h))(\alpha(k')), \beta(h)\beta(h')) \\ &= (\alpha(k), \beta(h)) *_\psi (\alpha(k'), \beta(h')) \\ &= F(k,h) *_\psi F(k',h') \end{aligned}$$

□

Applications. 1) Soit M une matrice d'ordre p dans $\mathrm{GL}(2, \mathbf{Z}/p\mathbf{Z})$ alors il existe une matrice inversible P telle que $M = P \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} P^{-1}$ donc le produit semi-direct non trivial $(\mathbf{Z}/p\mathbf{Z})^2 \rtimes_\phi \mathbf{Z}/p\mathbf{Z}$ associé à $\phi(m) = M^m$ est isomorphe à celui obtenu en prenant $M_0 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. 2) Soit y un élément d'ordre p dans $(\mathbf{Z}/p^2\mathbf{Z})^*$ alors $y = (p+1)^a$ avec a premier avec p donc le produit semi-direct non trivial $\mathbf{Z}/p^2\mathbf{Z} \rtimes_\phi \mathbf{Z}/p\mathbf{Z}$ associé à $\phi(m) = y^m \in (\mathbf{Z}/p^2\mathbf{Z})^*$ est isomorphe à celui obtenu en choisissant $y = p+1$; en effet $\beta(m) = am$ définit un automorphisme de $\mathbf{Z}/p\mathbf{Z}$ puisque a est premier avec p.

Attention. L'énoncé du lemme ne dit pas que ce sont les seuls isomorphismes entre produits semi-directs. Par exemple considérons $\phi : \mathbf{Z}/2\mathbf{Z} \to \mathrm{Int}(\mathcal{S}_3)$ l'homorphisme qui associe à 1 la conjugaison par une transposition, alors ϕ est non trivial mais pourtant $\mathcal{S}_n \rtimes_\phi \mathbf{Z}/2\mathbf{Z}$ est isomorphe à $\mathcal{S}_n \times \mathbf{Z}/2\mathbf{Z}$ (vérification "à la main" ou voir l'exercice suivant).

Exercice. Soit x un élément d'ordre r dans un groupe G. On note $\phi : \mathbf{Z}/r\mathbf{Z} \to \mathrm{Aut}(G)$ l'homomorphisme qui à 1 associe l'automorphisme intérieur associé à x (i.e. $\phi(1)(g) = xgx^{-1}$). Le produit semi-direct $G \rtimes_\phi \mathbf{Z}/r\mathbf{Z}$ n'est pas trivial, montrer néanmoins que l'application $f : G \rtimes_\phi \mathbf{Z}/r\mathbf{Z} \to G \times \mathbf{Z}/r\mathbf{Z}$ définie par $f(g,m) = (gx^m, m)$ est un isomorphisme de groupes

Exercice. Soit $N = p_1^{m_1} \dots p_s^{m_s}$ un entier impair avec sa décomposition en facteurs premiers. Montrer que le sous-groupe $U := \{a \in (\mathbf{Z}/N\mathbf{Z})^* \mid a^2 = 1\}$ est isomorphe à $(\mathbf{Z}/2\mathbf{Z})^s$. En déduire une description des différents produits semi-directs $\mathbf{Z}/n\mathbf{Z} \rtimes_\phi \mathbf{Z}/2\mathbf{Z}$. (Indication : il y a, à isomorphisme près, 2^s tels groupes, dont $\mathbf{Z}/N\mathbf{Z}$ et D_N). Comment doit-on modifier l'énoncé si $N = 2^{m_0} p_1^{m_1} \dots p_s^{m_s}$?

A.6. Groupes abéliens.

Remarquons tout de suite qu'un groupe abélien est la même chose qu'un $\mathbf{Z}$-module (i.e. un "espace vectoriel" sur l'anneau $\mathbf{Z}$). Comme exemples de groupes abéliens nous citerons au départ $\mathbf{Z}$, $\mathbf{Z}/n\mathbf{Z}$, $(\mathbf{Z}/n\mathbf{Z})^*$, $\mathbf{Q}$, $\mathbf{Q}/\mathbf{Z}$. Un groupe abélien est *de type fini* s'il possède un nombre fini de générateurs; il est dit *libre* s'il possède une base sur $\mathbf{Z}$, *libre de rang fini* s'il possède une base finie (et est donc isomorphe à $\mathbf{Z}^r$). Les groupes abéliens en général ne sont pas libres, en effet $\mathbf{Z}/n\mathbf{Z}$, par exemple, ne peut pas être libre. Dans le groupe $\mathbf{Q}$ deux éléments sont toujours liés mais le groupe n'est pas isomorphe à $\mathbf{Z}$. Un élément $x \in G$ est dit de *torsion* s'il existe $m \geq 1$ tel que $x^m = e$. Tous les éléments de $\mathbf{Q}/\mathbf{Z}$ sont de torsion sans que le groupe soit fini, donc il ne peut pas être de type fini. L'ensemble des éléments de torsion dans G abélien forme un sous-groupe $G_{\text{torsion}} := \{g \in G \mid \exists m \geq 1,\ g^m = e\}$; en effet si x est d'ordre m et y d'ordre n alors $(xy)^{mn} = (x^m)^n(y^n)^m = e$. Observons d'ailleurs que, si de plus m et n sont premiers entre eux, alors l'ordre de xy est exactement mn; en effet si $(xy)^k = e$, alors $x^{kn} = e$ (resp. $y^{km} = e$) donc m divise kn (resp. n divise km) donc m divise k (resp. n divise k) et enfin mn divise k.

Notation. Dans ce chapitre nous noterons (sauf mention contraire) additivement les groupes abéliens; l'élément neutre de $(G, +)$ sera noté 0.

Les groupes Z et $\mathbf{Z}/n\mathbf{Z}$ (rappels).

Le groupe $\mathbf{Z}$ est l'unique groupe (à isomorphisme près) qui est cyclique (engendré par un élément) et infini. Tous ses sous-groupes sont du type $m\mathbf{Z}$ pour $m \geq 0$. L'ensemble $\mathbf{Z}$ est également muni d'une multiplication qui en fait un anneau commutatif. Dans cet anneau on a la notion de divisibilité et l'on suppose connue la notion de PGCD et PPCM (que l'on révisera dans le cadre plus général des anneaux). Dans le cas de $\mathbf{Z}$ on voit que la notion d'*idéal* (voir le chapitre sur les anneaux) coïncide avec celle de sous-groupe. On peut en déduire facilement le théorème suivant

Théorème. (Bézout) *Soit $m, n \in \mathbf{Z}$ et soit d leur PGCD, alors il existe $a, b \in \mathbf{Z}$ tels que*

$$d = am + bn.$$

Preuve. L'ensemble $H := m\mathbf{Z} + n\mathbf{Z} = \{am + bn \mid a, b \in \mathbf{Z}\}$ est clairement un sous-groupe; il est donc de la forme $d'\mathbf{Z}$ et il existe a, b tels que $d' = am + bn$. Comme d divise a et b, on voit que d divise $am + bn = d'$ mais a, b appartiennent à H donc d' divise a et b donc d' divise également d et on conclut que $d = d'$ (si l'on a pris soin de les prendre tous les deux positifs). □

Le groupe $\mathbf{Z}/n\mathbf{Z}$ est l'unique groupe cyclique à n éléments (à isomorphisme près) i.e. engendré par un élément d'ordre n. On peut déjà étudier ses générateurs

Proposition. *Soit $m \in \mathbf{Z}$ et $\bar{m}$ sa classe dans $\mathbf{Z}/n\mathbf{Z}$, les trois propriétés suivantes sont équivalentes*
(i) L'élément $\bar{m}$ est un générateur de $\mathbf{Z}/n\mathbf{Z}$.
(ii) Les éléments m et n sont premiers entre eux.
(iii) L'élément $\bar{m}$ est inversible modulo n, c'est-à-dire qu'il existe $m' \in \mathbf{Z}$ tel que $mm' \equiv 1 \bmod n$ ou encore $\bar{m}\bar{m}' = 1 \in \mathbf{Z}/n\mathbf{Z}$.

Preuve. Supposons que $\bar{m}$ engendre $\mathbf{Z}/n\mathbf{Z}$ alors il existe $m' \in \mathbf{Z}$ tel que $m'\bar{m} = 1 \in \mathbf{Z}/n\mathbf{Z}$; ainsi $mm' \equiv 1 \bmod n$ ce qui signifie que m est inversible modulo n. Si $mm' \equiv 1 \bmod n$ alors $mm' = 1 + an$ et donc m est premier avec n. Si m est premier avec n alors, d'après le théorème de Bézout, il existe a, b tels que $am + bn = 1$ donc $a\bar{m} = 1 \in \mathbf{Z}/n\mathbf{Z}$ et donc $\bar{m}$ engendre $\mathbf{Z}/n\mathbf{Z}$. □

En particulier on voit que l'ensemble des éléments inversibles de $\mathbf{Z}/n\mathbf{Z}$, qui forment automatiquement un groupe, est égal à

$$(\mathbf{Z}/n\mathbf{Z})^* = \{\bar{m} \in \mathbf{Z}/n\mathbf{Z} \mid m \text{ est premier avec } n\}.$$

On note $\phi(n) := \text{card}\,((\mathbf{Z}/n\mathbf{Z})^*)$ l'*indicatrice d'Euler*. On en déduit facilement que, si p est premier, $\phi(p^r) = p^r - p^{r-1} = (p-1)p^{r-1}$. Le calcul en général de $\phi(n)$ se fait grâce au lemme classique suivant.

Proposition. (Lemme chinois) *Soit $m, n \in \mathbf{Z}$, supposons m et n premiers entre eux, alors les groupes $\mathbf{Z}/mn\mathbf{Z}$ et $\mathbf{Z}/m\mathbf{Z} \times \mathbf{Z}/n\mathbf{Z}$ sont naturellement isomorphes. De plus cet isomorphisme est aussi un isomorphisme d'anneaux et, par conséquent induit un isomorphisme entre $(\mathbf{Z}/mn\mathbf{Z})^*$ et $(\mathbf{Z}/m\mathbf{Z})^* \times (\mathbf{Z}/n\mathbf{Z})^*$.*

Preuve. Considérons l'application $f : \mathbf{Z} \to \mathbf{Z}/m\mathbf{Z} \times \mathbf{Z}/n\mathbf{Z}$ donnée par $x \mapsto (x \bmod m, x \bmod n)$. C'est un homomorphisme de groupe de noyau $\mathrm{PPCM}(m,n)\mathbf{Z}$, d'où une injection

$$\hat{f} : \mathbf{Z}/\,\mathrm{PPCM}(m,n)\mathbf{Z} \hookrightarrow \mathbf{Z}/m\mathbf{Z} \times \mathbf{Z}/n\mathbf{Z}.$$

Comme m et n sont supposés premiers entre eux, on a $\mathrm{PPCM}(m,n) = mn$ et, pour des raisons de cardinalité, l'homomorphisme $\hat{f}$ doit être un isomorphisme. De manière générale, si A et B sont des anneaux, on a $(A \times B)^* = A^* \times B^*$ d'où la deuxième assertion. □

La description des sous-groupes de $\mathbf{Z}/n\mathbf{Z}$ est assez simple.

Proposition. *Pour chaque entier $d \geq 1$ divisant n, il existe un unique sous-groupe de $\mathbf{Z}/n\mathbf{Z}$ d'ordre d, c'est le sous-groupe cyclique engendré par la classe de n/d dans $\mathbf{Z}/n\mathbf{Z}$.*

Preuve. Supposons $n = dd'$ alors l'élément $x = \bar{d}' \in \mathbf{Z}/n\mathbf{Z}$ est d'ordre d car clairement $dx = 0$ et, si $cx = 0$ alors n divise cd' donc d divise c. Soit maintenant H un sous-groupe de $\mathbf{Z}/n\mathbf{Z}$ d'ordre d. Notons $s : \mathbf{Z} \to \mathbf{Z}/n\mathbf{Z}$ la surjection canonique. On sait que $s^{-1}(H) = m\mathbf{Z}$ est engendré par m donc H est engendré par $\bar{m} \in \mathbf{Z}/n\mathbf{Z}$. On a $d\bar{m} = 0$ donc n divise dm donc d' divise m donc le sous-groupe H est contenu dans le sous-groupe engendré par $\bar{d}'$ et donc égal à ce sous-groupe. □

Comme application, on peut en tirer la formule (que nous utiliserons plus bas)

$$n = \sum_{d \,|\, n} \phi(d).$$

En effet on écrit $\mathbf{Z}/n\mathbf{Z}$ comme union (disjointe) des ensembles d'éléments d'ordre d pour d divisant n. Le nombre de ces éléments est le nombre de générateurs de l'unique sous-groupe de cardinal d, et comme ce dernier est isomorphe à $\mathbf{Z}/d\mathbf{Z}$, le nombre de générateurs est $\phi(d)$.

Les groupes $(\mathbf{Z}/n\mathbf{Z})^*$.

On notera (à titre d'exception dans ce chapitre) multiplicativement la loi du groupe $(\mathbf{Z}/n\mathbf{Z})^*$. D'après ce que nous avons vu, si $n = p_1^{\alpha_1} \dots p_s^{\alpha_s}$ alors

$$(\mathbf{Z}/n\mathbf{Z})^* \cong (\mathbf{Z}/p_1^{\alpha_1}\mathbf{Z})^* \times \ldots \times (\mathbf{Z}/p_s^{\alpha_s}\mathbf{Z})^*$$

et en particulier

$$\phi(n) = \phi(p_1^{\alpha_1}) \dots \phi(p_s^{\alpha_s}) = \prod_{i=1}^{s} \left(p_i^{\alpha_i} - p_i^{\alpha_i - 1}\right) = n \prod_{i=1}^{s} \left(1 - \frac{1}{p_i}\right)$$

Il reste à décrire la structure des groupes $(\mathbf{Z}/p^\alpha\mathbf{Z})^*$.

Proposition. *Soit p premier et $\alpha \geq 1$ alors*
(i) Si p est impair $(\mathbf{Z}/p^\alpha\mathbf{Z})^*$ est cyclique.
(ii) Si $p = 2$ et $\alpha \geq 3$ alors $(\mathbf{Z}/2^{\alpha-2}\mathbf{Z})^* \cong \mathbf{Z}/2^\alpha\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$ n'est pas cylique. Par contre $(\mathbf{Z}/2\mathbf{Z})^* = \{1\}$ et $(\mathbf{Z}/4\mathbf{Z})^* \cong \mathbf{Z}/2\mathbf{Z}$ sont cycliques.

Preuve. Commençons par montrer que $(\mathbf{Z}/p\mathbf{Z})^*$ est cyclique, en fait plus généralement on a le résultat suivant.

Lemme. *Soit k un corps commutatif et G un sous-groupe fini de k^*, alors G est cyclique. En particulier $(\mathbf{Z}/p\mathbf{Z})^*$ est cyclique.*

Preuve du lemme. Notons $n := \mathrm{card}(G)$ et $\psi(d)$ le nombre d'éléments d'ordre d dans G. On a clairement $n = \sum_{d \,|\, n} \psi(d)$. Soit d divisant n, ou bien il n'y a pas d'élément d'ordre d dans G auquel cas $\psi(d) = 0$, ou

bien il en existe un qui engendre alors un sous-groupe cyclique H d'ordre d. Tous les éléments de H sont solutions de l'équation $X^d = 1$, mais, comme k est un corps commutatif, une telle équation possède au plus d racines dans k; tous les éléments d'ordre d sont donc dans H et il en a $\phi(d)$ puisque $H \cong \mathbf{Z}/d\mathbf{Z}$. Ainsi $\psi(d)$ vaut zéro ou $\phi(d)$, mais comme $n = \sum_{d\,|\,n} \psi(d) = \sum_{d\,|\,n} \phi(d)$, on voit que $\psi(d) = \phi(d)$ pour tout d divisant n. En particulier $\psi(n) = \phi(n) \geq 1$, ce qui implique bien que G est cyclique. □

Lemme. *Soit p premier impair, la classe de $p+1$ dans $(\mathbf{Z}/p^\alpha\mathbf{Z})^*$ est d'ordre $p^{\alpha-1}$.*

Preuve du lemme. Montrons d'abord par récurrence la congruence

$$(p+1)^{p^k} \equiv 1 + p^{k+1} \bmod p^{k+2}.$$

Pour $k = 0$, la congruence est triviale. Supposons donc $(p+1)^{p^{k-1}} = 1 + p^k + ap^{k+1}$ alors $(p+1)^{p^k} = \left(1 + p^k + ap^{k+1}\right)^p \equiv 1 + p(p^k + ap^{k+1}) \equiv 1 + p^{k+1} \bmod p^{k+2}$. Pour l'avant-dernière congruence, on a besoin de $p \neq 2$; en effet la formule du binôme de Newton fait apparaître des termes multiples de p^{kr} donc nuls modulo p^{k+2} sauf peut-être si $r = 2$ et $k = 1$ mais le terme s'écrit alors $C_p^2 p^2$ qui est bien nul modulo p^3 si p est impair. En particulier, on voit que $(p+1)^{p^{\alpha-1}} \equiv 1 \bmod p^\alpha$ mais $(p+1)^{p^{\alpha-2}} \equiv 1 + p^{\alpha-1} \not\equiv 1 \bmod p^\alpha$, ce qui implique bien que $p+1$ est d'ordre $p^{\alpha-1}$ dans $(\mathbf{Z}/p^\alpha\mathbf{Z})^*$.

On peut maintenant terminer la preuve de la proposition pour p impair. Soit $x \in \mathbf{Z}$ tel que x modulo p engendre $(\mathbf{Z}/p\mathbf{Z})^*$ i.e. est d'ordre $p-1$ dans $(\mathbf{Z}/p\mathbf{Z})^*$; alors $\bar{x}$ est d'ordre $m(p-1)$ dans $(\mathbf{Z}/p^\alpha\mathbf{Z})^*$ et donc $y = \bar{x}^m$ est d'ordre exactement $p-1$ dans $(\mathbf{Z}/p^\alpha\mathbf{Z})^*$. L'élément $y(p+1)$ est donc d'ordre $p^{\alpha-1}(p-1)$ donc est un générateur de $(\mathbf{Z}/p^\alpha\mathbf{Z})^*$ (car $p^{\alpha-1}$ et $p-1$ sont premiers entre eux).

Lemme. *La classe de 5 dans $(\mathbf{Z}/2^\alpha\mathbf{Z})^*$ est d'ordre $2^{\alpha-2}$. De plus la classe de -1 n'appartient pas au sous-groupe engendré par la classe de 5.*

Preuve du lemme. On montre d'abord par récurrence que

$$5^{2^k} \equiv 1 + 2^{k+2} \bmod 2^{k+3}.$$

La congruence est triviale pour $k = 0$, supposons donc que $5^{2^{k-1}} = 1 + 2^{k+1} + a2^{k+2}$ alors $5^{2^k} = (1 + 2^{k+1} + a2^{k+2})^2 = 1 + 2(2^{k+1} + a2^{k+2}) + 2^{2(k+1)}(1 + 2a)^2 \equiv 1 + 2^{k+2} \bmod 2^{k+3}$. En particulier $5^{2^{\alpha-2}} \equiv 1 \bmod 2^\alpha$ mais $5^{2^{\alpha-3}} \equiv 1 + 2^{\alpha-1} \not\equiv 1 \bmod 2^\alpha$ donc 5 est bien d'ordre $2^{\alpha-2}$. Supposons que $5^\beta \equiv -1 \bmod 2^\alpha$ alors $5^{2\beta} \equiv 1 \bmod 2^\alpha$ donc $2^{\alpha-2}$ divise 2β donc $2^{\alpha-3}$ divise β ou encore $\beta = \gamma 2^{\alpha-3}$. Comme 5 est d'ordre $2^{\alpha-2}$, on peut considérer β comme un entier modulo $2^{\alpha-2}$ et donc γ modulo 2. L'entier γ doit être impair donc on peut le supposer égal à 1, c'est-à-dire $5^{2^{\alpha-3}} \equiv 1 \bmod 2^\alpha$, mais $5^{2^{\alpha-3}} \equiv 1 + 2^{\alpha-1} \bmod 2^\alpha$ donc $-1 \equiv 1 + 2^{\alpha-1} \bmod 2^\alpha$ ou encore $2 + 2^{\alpha-1} \equiv \bmod 2^\alpha$ soit $1 + 2^{\alpha-2} \equiv \bmod 2^{\alpha-1}$, ce qui n'est pas possible. □

Pour la démonstration de la deuxième partie de la proposition, on peut supposer $\alpha \geq 3$ (en effet le calcul de $(\mathbf{Z}/2\mathbf{Z})^*$ et $(\mathbf{Z}/4\mathbf{Z})^*$ est immédiat). La classe de 5 engendre donc un sous-groupe isomorphe à $\mathbf{Z}/2^{\alpha-2}\mathbf{Z}$ et -1 engendre un sous-groupe d'ordre 2 non contenu dans le précédent donc $(\mathbf{Z}/2^\alpha\mathbf{Z})^* = \langle 5 \rangle \oplus \langle -1 \rangle \cong \mathbf{Z}/2^{\alpha-2}\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$. □

Exercice. Montrer que si la classe de $x \in \mathbf{Z}$ engendre $(\mathbf{Z}/p^2\mathbf{Z})^*$ alors elle engendre aussi $(\mathbf{Z}/p^\alpha\mathbf{Z})^*$ (pour p impair).

Remarque. Le sous-groupe quaternionique $H_8 = \{\pm 1, \pm i, \pm j, \pm k\}$ est un sous-groupe fini du groupe multiplicatif du corps $\mathbf{H}$ mais n'est pas cyclique (cela ne contredit pas le lemme vu car $\mathbf{H}$ n'est pas commutatif).

Théorèmes de structure.

Les produits finis de groupes cycliques sont évidemment abéliens de type fini. Nous allons voir réciproquement que tout groupe abélien de type fini est en fait isomorphe à un groupe de la forme $\mathbf{Z}^r \times \mathbf{Z}/m_1\mathbf{Z} \times \ldots \times \mathbf{Z}/m_s\mathbf{Z}$. Toutefois le lemme chinois indique qu'une telle décomposition n'est pas *a priori* unique. On peut néanmoins en extraire des éléments invariants ou canoniques.

Théorème. *Tout groupe abélien G de type fini est produit de groupes cycliques. Plus précisément il existe $r \geq 0$ et $a_1, \dots a_s$ avec $a_i \geq 2$ et a_i divise a_{i+1} tels que*

$$G \cong \mathbf{Z}^r \times \mathbf{Z}/a_1\mathbf{Z} \times \ldots \times \mathbf{Z}/a_s\mathbf{Z}.$$

De plus les entiers $r, s, a_1, a_2, \dots a_s$ sont uniques.

Nous allons utiliser dans la preuve un autre théorème de structure, décrivant les sous-groupes de $\mathbf{Z}^r$, qui est démontré au chapitre sur les modules sur les anneaux principaux.

Théorème. *Soit H un sous-groupe de $\mathbf{Z}^r$ alors*
(i) Le groupe H est libre de rang $s \leq r$.
(ii) Il existe $e_1, \dots, e_r$ base de $\mathbf{Z}^r$ et $a_1, \dots a_s \geq 1$ tels que a_i divise a_{i+1} et $a_1e_1, \dots a_se_s$ forment une base de H.

Preuve (du théorème antérieur). Supposons que G possède n générateurs, alors on en déduit un homomorphisme surjectif $f : \mathbf{Z}^n \to G$ et un isomorphisme $\mathbf{Z}^n/\operatorname{Ker}(f) \cong G$. On applique le théorème précédent à $\operatorname{Ker}(f)$ et on obtient des e_i et a_i tels que $\mathbf{Z}^n = \mathbf{Z}e_1 \oplus \dots \mathbf{Z}e_n$ et tels que $\operatorname{Ker}(f) = \mathbf{Z}a_1e_1 \oplus \dots \mathbf{Z}a_me_m$. D'où l'on tire aisément

$$G \cong \mathbf{Z}^n/\operatorname{Ker}(f) \cong \mathbf{Z}/a_1\mathbf{Z} \times \ldots \times \mathbf{Z}/a_m\mathbf{Z} \times (\mathbf{Z})^{n-m}$$

avec a_i divisant a_{i+1} et $a_i \geq 1$. En éliminant les facteurs correspondant à $a_i = 1$, on obtient l'existence de la décomposition annoncée. Montrons maintenant l'unicité. Nous allons utiliser le

Lemme. *Soit $M \geq 1$ alors le sous-groupe $M\mathbf{Z}/n\mathbf{Z}$ est cyclique de cardinal $n/\operatorname{PGCD}(n, M)$; le quotient $(\mathbf{Z}/n\mathbf{Z})/M(\mathbf{Z}/n\mathbf{Z})$ est cyclique de cardinal $\operatorname{PGCD}(n, M)$.* □

Preuve. Notons $d = PGCD(n, M)$ et $n = n'd$, $M = M'd$. Considérons la composée $\mathbf{Z} \xrightarrow{\times M} \mathbf{Z} \to \mathbf{Z}/n\mathbf{Z}$. Son noyau est le sous-groupe des $x \in \mathbf{Z}$ tels que n divise Mx ou encore tels que n' divise x d'où un isomorphisme entre $\mathbf{Z}/n'\mathbf{Z}$ et l'image, c'est-à-dire $M\mathbf{Z}/n\mathbf{Z}$. Enfin $(\mathbf{Z}/n\mathbf{Z})/M(\mathbf{Z}/n\mathbf{Z})$ est cyclique de cardinal d donc isomorphe à $\mathbf{Z}/d\mathbf{Z}$. □

Supposons maintenant

$$G \cong \mathbf{Z}^r \times \mathbf{Z}/a_1\mathbf{Z} \times \ldots \times \mathbf{Z}/a_s\mathbf{Z} \cong \mathbf{Z}^{r'} \times \mathbf{Z}/b_1\mathbf{Z} \times \ldots \times \mathbf{Z}/b_t\mathbf{Z}$$

avec $a_i, b_i \geq 2$ et a_i divise a_{i+1}, resp. b_i divise b_{i+1}. On commence par choisir un entier M multiple de a_s et b_t alors $MG \cong \mathbf{Z}^r \cong \mathbf{Z}^{r'}$ donc $r = r'$. En remplaçant G par G_{torsion} on peut maintenant supposer G fini (i.e. $r = r' = 0$). Choisissons p divisant a_1 (noter que $a_1 \geq 2$) alors $PGCD(p, a_i) = p$ et $PGCD(p, b_i) = p$ ou 1 suivant que p divise b_i ou non. Donc d'après le lemme $G/pG \cong (\mathbf{Z}/p\mathbf{Z})^s \cong (\mathbf{Z}/p\mathbf{Z})^{t-\operatorname{card}\{i \mid p \text{ ne divise pas } b_i\}}$. Ainsi $s \leq t$ et, par symétrie $t = s$ et donc p divise b_1. Ecrivons donc $a_i = pa'_i$ et $b_i = pb'_i$, alors $pG \cong \mathbf{Z}/a'_1\mathbf{Z} \times \ldots \times \mathbf{Z}/a'_s\mathbf{Z} \cong \mathbf{Z}/b'_1\mathbf{Z} \times \ldots \times \mathbf{Z}/b'_s\mathbf{Z}$. par récurrence sur $\operatorname{card}(G)$ on en tire que $a'_i = b'_i$ et donc $a_i = b_i$. □

Revenons aux groupes abéliens finis et montrons qu'on peut écrire une autre décomposition canonique.

Théorème. *Un groupe abélien fini G est somme directe de ses p-sous-groupes de Sylow. Un p-groupe abélien est isomorphe à un produit $(\mathbf{Z}/p\mathbf{Z})^{m_1} \times (\mathbf{Z}/p^2\mathbf{Z})^{m_2} \times \ldots \times (\mathbf{Z}/p^r\mathbf{Z})^{m_r}$ avec $m_i \geq 0$. De plus les m_i sont uniques.*

Le groupe G est abélien donc possède un unique p-sous-groupe de Sylow. On voit aisément que celui-ci est égal à $G_p := \{x \in G \mid \exists m \geq 0,\ p^m x = 0\}$. La première partie du théorème est alors une conséquence du lemme ci-dessous; la deuxième partie découle directement du théorème de structure précédent.

Lemme. *Soit G un groupe de cardinal MN avec M et N premiers entre eux. Soit $G_1 = \{x \in G \mid Mx = 0\}$ et $G_2 = \{x \in G \mid Nx = 0\}$, alors $G = G_1 \oplus G_2$.*

Preuve. D'après le théorème de Bézout, il existe $a, b \in \mathbf{Z}$ tels que $aM + bN = 1$. Si $x \in G_1 \cap G_2$ alors $x = (aM + bN)x = 0$. Si maintenant $x \in G$ alors $x = bNx + aMx$ et, comme MN est un exposant pour G, on a $bNx \in G_1$ et $aMx \in G_2$. □

Exercice. Soit une décomposition $G \cong \mathbf{Z}/a_1\mathbf{Z} \times \ldots \times \mathbf{Z}/a_s\mathbf{Z}$ avec $a_i \geq 2$ et a_i divise a_{i+1}. Montrer que l'exposant de G est égal à a_s et que le nombre minimal de générateurs de G est s.

A.7. Le groupe $\mathcal{S}_n$.

Le groupe $\mathcal{S}_n$ est le groupe des bijections de l'ensemble $[1,n] = \{1, 2, \ldots, n\}$, il est isomorphe au groupe des bijections d'un ensemble fini de cardinal n. Il intervient donc chaque fois qu'un groupe agit sur un ensemble fini, en particulier dans les questions de combinatoire. D'un autre côté, le groupe $\mathcal{S}_n$ est "trop" riche pour pouvoir être entièrement décrit; par exemple tout groupe fini est sous-groupe d'un $\mathcal{S}_n$: en effet, l'action de G par translation sur lui-même est fidèle et induit donc une injection de G dans les bijections de G.

Le *support* d'une permutation $\sigma \in \mathcal{S}_n$ est le sous-ensemble $\{i \in [1,n] \mid \sigma(i) \neq i\}$. Le groupe $\mathcal{S}_n$ agit transitivement sur $[1,n]$ et le stabilisateur de n est naturellement isomorphe à $\mathcal{S}_{n-1}$ donc la formule des classes nous dit que $\mathrm{card}(\mathcal{S}_n/\mathcal{S}_{n-1}) = n$ d'où l'on tire aisément par récurrence

$$\mathrm{card}(\mathcal{S}_n) = n!$$

Une première façon de noter les éléments de $\mathcal{S}_n$ est simplement d'écrire la liste des images, par exemple la permutation σ définie par $\sigma(1) = 2$, $\sigma(2) = 6$, $\sigma(3) = 3$, $\sigma(4) = 5$, $\sigma(5) = 8$, $\sigma(6) = 4$, $\sigma(7) = 10$, $\sigma(8) = 9$, $\sigma(9) = 1$, $\sigma(10) = 7$, peut être notée $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 2 & 6 & 3 & 5 & 8 & 4 & 10 & 9 & 1 & 7 \end{pmatrix}$. Cette notation est toutefois lourde et ne reflète pas vraiment les propriétés de σ (par exemple : quel est son ordre?). La situation est un peu similaire à l'écriture d'un nombre entier : l'écriture de la décomposition en facteurs premiers contient beaucoup plus d'information arithmétique que la donnée du nombre en base 10. Il est donc utile d'introduire une telle notion pour les permutations.

Définition. Un *cycle de longueur* m (ou m-cycle) est associé à un sous-ensemble ordonné $I = \{i_1 \ldots, i_m\}$ et est donné par $\sigma(i_1) = i_2, \ldots, \sigma(i_{m-1}) = i_m$, $\sigma(i_m) = i_1$ et, pour tout $j \notin I$, $\sigma(j) = j$. L'ensemble I s'appelle le *support* du cycle. On note une telle permutation $\sigma = (i_1 \ldots, i_m)$. Un cycle de longueur 2 est une *transposition*.

Remarquons que, avec la notation introduite $(i_1 \ldots, i_m) = (i_2 \ldots, i_m, i_1)$, etc. Un cycle de longueur m a clairement pour ordre m. L'intérêt de cette notion provient en bonne partie du résultat suivant.

Théorème. (Décomposition en cycles) *Soit $\sigma \in \mathcal{S}_n \setminus \{id\}$ il existe $\sigma_1, \ldots, \sigma_r$, cycles de longueurs $m_1, \ldots, m_r$ ayant des supports disjoints, tels que*

$$\sigma = \sigma_1 \cdots \sigma_r.$$

De plus, l'union des supports des σ_i est le support de σ, les σ_i commutent entre eux et sont uniques (à l'ordre près).

Preuve. On décompose l'ensemble $X = [1,n]$ sous l'action du groupe engendré par σ en orbites. Sur chaque orbite X_i de cardinal $m \geq 2$, la permutation σ agit comme un cycle σ_i de support X_i. Il est alors immédiat que σ est égale au produit des σ_i et celles-ci sont uniquement déterminées par σ. Deux permutations dont les supports sont disjoints commutent; le reste est clair. □

Si σ s'écrit $\sigma_1 \cdots \sigma_r$ comme dans l'énoncé du théorème, i.e. est produit de cycles à supports disjoints de longueur $m_1, \ldots, m_r$, on dira que σ est de *type* $(m_1, \ldots, m_r)$.

Corollaire. *Soit σ une permutation de type $(m_1, \ldots, m_r)$, alors son ordre est égal au PPCM de $m_1, \ldots, m_r$.*

Preuve. Notons $M := \mathrm{PPCM}(m_1, \ldots, m_r)$. Comme $\sigma = \sigma_1 \ldots \sigma_r$ on a $\sigma^M = \sigma_1^M \ldots \sigma_r^M = id$ et d'autre part si $\sigma^N = \sigma_1^N \ldots \sigma_r^N = id$, alors σ^N agit sur le support de σ_i comme σ_i^N et comme l'identité donc $\sigma_i^N = id$ et m_i divise N donc M divise N. □

Exemple. La décomposition de la permutation donnée ci-dessus s'écrit $\sigma = (1, 2, 6, 4, 5, 8, 9)(7, 10)$. Elle a donc pour ordre 14.

Corollaire. *La classe de conjugaison d'une permutation de type $(m_1, \ldots, m_r)$ est l'ensemble des permutations de même type.*

Preuve. Commençons par vérifier la "formule-clef" suivante où ρ désigne une permutation quelconque :

$$\rho(i_1,\dots,i_m)\rho^{-1} = (\rho(i_1),\dots,\rho(i_m)).$$

Notons $\sigma = (i_1,\dots,i_m)$. Si $j \notin \{\rho(i_1),\dots,\rho(i_m)\}$ alors $\rho^{-1}(j) \notin \{i_1,\dots,i_m\}$ donc $\rho\sigma\rho^{-1}(j) = j$. Si $j = \rho(i_k)$ alors $\rho^{-1}(j) = i_k$ donc $\sigma\rho^{-1}(j) = i_{k+1}$ (avec la convention que $m+1 = 1$) et $\rho\sigma\rho^{-1}(j) = \rho(i_{k+1})$ comme annoncé. Ainsi le conjugué d'un m-cycle est un m-cycle; de plus si $\sigma' = (j_1,\dots,j_m)$ est un autre m-cycle on peut choisir $\rho \in \mathcal{S}_n$ telle que $\rho(i_k) = j_k$ et donc $\sigma' = \rho\sigma\rho^{-1}$. Ainsi la classe de conjugaison d'un m-cycle est l'ensemble des m-cycles. Dans le cas général, si $\sigma = \sigma_1 \dots \sigma_r$, alors $\rho\sigma\rho^{-1} = (\rho\sigma_1\rho^{-1})\dots(\rho\sigma_r\rho^{-1})$ donc le conjugué d'une permutation de type $m_1,\dots,m_r$ est encore du même type et réciproquement. □

La *signature* d'une permutation $\sigma \in \mathcal{S}_n$ peut être définie par la formule

$$\epsilon(\sigma) := \prod_{1\le i<j\le n} .\frac{(\sigma(i)-\sigma(j))}{(i-j)}.$$

Proposition. *L'application $\epsilon : \mathcal{S}_n \to \{\pm 1\}$ est un homomorphisme de groupes. La signature d'une transposition est égale à -1. Son noyau est noté $\mathcal{A}_n$ et s'appelle le groupe alterné.*

Preuve. Observons que $\eta_\sigma(i,j) = (\sigma(i)-\sigma(j))/(i-j)$ ne dépend que de la paire $\{i,j\}$. On peut écrire

$$\epsilon(\sigma\tau) = \prod_{\{i,j\}} \frac{\sigma\tau(i))-\sigma\tau(j)}{\tau(i)-\tau(j)}.\frac{\tau(i)-\tau(j)}{i-j} = \left(\prod_{\{i,j\}} \eta_\sigma(\tau(i),\tau(j))\right)\epsilon(\tau) = \epsilon(\sigma)\epsilon(\tau).$$

Pour la deuxième affirmation, il suffit de vérifier que $\epsilon((1,2)) = -1$ ce qui est élémentaire. □

Remarques. On sait (Cf plus loin) que toute permutation peut s'écrire comme le produit d'un certain nombre de transpositions, disons $\sigma = \tau_1 \dots \tau_s$; on en déduit que $\epsilon(\sigma) = (-1)^s$. Un m-cycle est le produit de $m-1$ transpositions donc la signature d'un m-cycle est $(-1)^{m-1}$, la signature d'une permutation de type $(m_1,\dots,m_r)$ est $(-1)^{m_1+\dots+m_r-r}$.

Corollaire. *Le sous-groupe $\mathcal{A}_n$ est distingué dans $\mathcal{S}_n$ et* $\mathrm{card}(\mathcal{A}_n) = n!/2$.

Preuve. Immédiat. □

Générateurs de $\mathcal{S}_n$ et $\mathcal{A}_n$.

Tout d'abord l'ensemble des cycles est un ensemble de générateurs de $\mathcal{S}_n$ d'après le théorème de décomposition en cycles. Ensuite tout cycle peut s'écrire comme produit de transpositions car

$$(i_1,\dots,i_m) = (i_1,i_2)(i_2,i_3)\cdots(i_{m-1},i_m)$$

donc l'ensemble des transpositions est un ensemble de générateurs de $\mathcal{S}_n$. On peut même se restreindre au sous-ensemble des transpositions de la forme $(i,i+1)$ pour $1 \le i \le m-1$. En effet si $i < j$ et $\rho = (i+1,i+2)\dots(j-1,j)$ alors $\rho(i) = i$ et $\rho(j) = i+1$ donc $\rho(i,j)\rho^{-1} = (i,i+1)$. A titre d'exercice on pourra montrer qu'une transposition et un cycle de longueur n forme un système minimal de générateurs. Montrons que les cycles de longueur 3 engendrent $\mathcal{A}_n$. Un élément $\sigma \in \mathcal{A}_n$ s'écrit comme un produit d'un nombre pair de transpositions (puisque $\epsilon(\sigma) = +1$) donc $\mathcal{A}_n$ est engendré par les éléments de la forme $(i,j)(k,\ell)$, où l'on peut supposer $(i,j) \neq (k,\ell)$. Si $\mathrm{card}\,(\{i,j\}\{k,\ell\}) = 1$ alors $(i,j)(k,\ell)$ est un 3-cyle, sinon on peut écrire $(i,j)(k,\ell) = (i,j)(j,k)(j,k)(k,\ell)$ et chacune des permutations $(i,j)(j,k)$ et $(j,k)(k,\ell)$ est un 3-cycle.

Exemple de sous-groupes de $\mathcal{S}_n$ (resp. de $\mathcal{A}_n$).

(a) Si $n \le 2$, le groupe $\mathcal{S}_n$ est commutatif, cependant si $n \ge 3$, le centre de $\mathcal{S}_n$ est trivial. En effet si $\rho \in Z(\mathcal{S}_n)$ alors $(i,j) = (\rho(i),\rho(j))$ donc $\{\rho(i),\rho(j)\} = \{i,j\}$; supposons qu'il existe i avec $\rho(i) \neq i$, alors pour tout $j \neq i$ on a $\rho(i) = j$, ce qui est aburde dès que $n \ge 3$.

(b) Soit $m \leq n$, un cycle de longueur m dans $\mathcal{S}_n$ est déterminé par son support (il y a $C_n^m = \binom{n}{m}$ possibilités) et l'ordre donné à ce support (à permutation cyclique près, soit $(m-1)!$ possibilités). Ainsi $\mathcal{S}_n$ contient $(m-1)!C_n^m$ cycles de longueur m et le nombre de sous-groupes cycliques que ceux-ci engendrent est $(m-1)!C_n^m/\phi(m)$. Attention : ce n'est pas, en général, le nombre de sous-groupes cycliques de cardinal m, néanmoins, si p est premier et $p \leq n < 2p$, un sous-groupe de cardinal p est engendré par un p-cycle et il y a donc $(p-2)!C_n^p$ tels sous-groupes. (Exercice : vérifier dans ce cas un des théorèmes de Sylow qui affirme que $(p-2)!C_n^p \equiv 1 \bmod p$ et en déduire le théorème de Wilson $(p-2)! \equiv 1 \bmod p$).

(c) Soit $n = n_1 + n_2 + \ldots + n_r$ une partition de n, alors on dispose d'une injection $\mathcal{S}_{n_1} \times \ldots \times \mathcal{S}_{n_r} \hookrightarrow \mathcal{S}_n$ en associant à $(\sigma_1, \ldots, \sigma_r \in \mathcal{S}_{n_1} \times \ldots \times \mathcal{S}_{n_r}$ la permutation définie, pour $1 \leq i \leq r$ et $1 \leq j \leq n_i$, par $\sigma(n_1 + \ldots + n_{i-1} + j) = n_1 + \ldots + n_{i-1} + \sigma_i(j)$.

(d) Il existe néanmoins d'autres façons de plonger $\mathcal{S}_m$ dans $\mathcal{S}_n$. Ainsi le groupe $\mathcal{S}_5$ possède six 5-sous-groupes de Sylow d'où une injection $\mathcal{S}_5 \hookrightarrow \mathcal{S}_6$. Notons que l'image de $\mathcal{S}_5$ ne stabilise aucun élément puisqu'il agit transitivement.

(e) (centralisateur d'un élément) Soit $\sigma \in \mathcal{S}_n$, on veut déterminer le sous-groupe

$$C(\sigma) := \{\rho \in \mathcal{S}_n \mid \rho\sigma = \sigma\rho\}.$$

Si $\sigma = (i_1, \ldots, i_m)$ un m-cycle, un élément ρ commute avec σ si l'on a l'égalité de cycles $(\rho(i_1), \ldots, \rho(i_m))$ donc si et seulement si le sous-ensemble $\{i_1, \ldots, i_m\}$ est une orbite (sous l'action du sous-groupe engendré par σ) sur lequel σ agit par permutation circulaire. Si l'on identifie le sous-groupe des permutations de support $\{i_1, \ldots, i_m\}$ (resp. fixant le sous-ensemble $\{i_1, \ldots, i_m\}$)avec $\mathcal{S}_m$ (resp. $\mathcal{S}_{n-m}$) alors $\mathcal{S}_{n-m} \hookrightarrow C(\sigma)$; de plus le sous-groupe $\mathcal{S}_{n-m}$ est distingué dans $C(\sigma)$ et le quotient est isomorphe au sous-groupe engendré par σ (i.e. à $\mathbf{Z}/m\mathbf{Z}$); en particulier $\mathrm{card}(C(\sigma)) = (n-m)!m$. Montrer plus généralement que si σ est le produits de r_2 transpositions, r_3 cycles de longueur 3 etc (avec disons $n = r_1 + 2r_2 + 3r_3 + \ldots + sr_s$) alors

$$\mathrm{card}(C(\sigma)) = r_1!r_2! \ldots r_s!2^{r_1} \ldots s^{r_s}.$$

Le groupe $\mathcal{S}_1$ est trivial, le groupe $\mathcal{S}_2$ est commutatif. Le groupe $\mathcal{S}_3$ possède trois sous-groupes de cardinal 2 (autant que de transpositions), un unique sous-groupe de cardinal 3 : le sous-groupe $\mathcal{A}_3$ (puisque $\mathcal{A}_3 \triangleleft \mathcal{S}_3$) qui est cyclique. En particulier $\mathcal{S}_3$ est résoluble. Le groupe $\mathcal{S}_4$ contient quatre sous-groupes isomorphes à $\mathcal{S}_3$ qui sont tous conjugués (les stabilisateurs de $1, 2, 3, 4$) et donc quatre sous-groupes de cardinal 3 (qui sont tous conjugués). Les 2-sous-groupes de Sylow de $\mathcal{S}_4$ sont au nombre de 3 et sont isomorphes au groupe dièdral D_4. En effet l'action de D_4 sur les sommets d'un carré induit un isomorphisme de D_4 sur un sous-groupe de $\mathcal{S}_4$; ce sous-groupe ne peut être distingué car sinon il contiendrait tous les éléments d'ordre 2 ou 4 de $\mathcal{S}_4$ donc il y a 3 tels sous-groupes (qui sont tous conjugués). On peut en déduire un sous-groupe particulier

Le *sous-groupe de Klein* de $\mathcal{S}_4$ est l'intersection de ses 2-sous-groupes de Sylow, ou encore le sous-groupe constitué de l'élément neutre et des doubles transpositions

$$K := \{id, (1,2)(34), (1,3)(2,4), (1,4)(2,3)\}.$$

Ce sous-groupe est donc distingué dans $\mathcal{S}_4$ et isomorphe à $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$. En particulier la suite $\mathcal{S}_4 \supset \mathcal{A}_4 \supset K \supset \{id, (1,2)(34)\} \supset \{id\}$ est une suite de composition avec quotients successifs $\mathbf{Z}/2\mathbf{Z}$, $\mathbf{Z}/3\mathbf{Z}$, $\mathbf{Z}/2\mathbf{Z}$ et $\mathbf{Z}/2\mathbf{Z}$ donc $\mathcal{S}_4$ est résoluble. Le groupe quotient $\mathcal{S}_4/K$ est isomorphe $\mathcal{S}_3$; en effet, si l'on fait agir $\mathcal{S}_4$ sur ses 2-sous-groupes de Sylow, le stabilisateur (normalisateur) de chacun de ces sous-groupes de Sylow est égal à lui-même, donc leur intersection est K; l'homomorphime $\rho : \mathcal{S}_4 \to \mathcal{S}_3$ associé à cette action a donc pour noyau K et est donc surjectif.

Théorème. *Soit H un sous-groupe distingué non trivial de $\mathcal{S}_n$, alors ou bien $H = \mathcal{A}_n$ ou bien $n = 4$ et H est le sous-groupe de Klein. Le groupe $\mathcal{S}_n$ est résoluble si et seulement si $n \leq 4$, le groupe $\mathcal{A}_n$ est simple si et seulement si $n \geq 5$.*

Preuve. Montrons d'abord que, si $n \geq 5$ tous les 3-cycles sont conjugés dans $\mathcal{A}_n$ et donc un sous-groupe distingué qui contient un 3-cycle les contient tous et est donc égal à $\mathcal{A}_n$. Soit $\sigma = (i, j, k)$, dès que $\rho(1) = i, \rho(2) = j$ et $\rho(3) = k$, on a $\rho(1, 2, 3)\rho^{-1} = (i, j, k)$. A priori $\rho \in \mathcal{S}_n$ mais, si $n \geq 5$ on peut s'arranger pour que $\rho \in \mathcal{A}_n$, quitte à remplacer éventuellement ρ par $\rho(4, 5)$.

Montrons que $\mathcal{A}_5$ est simple. Soit $H \neq \{id\}$ sous-groupe distingué de $\mathcal{A}_5$. Si H contient un 3-cycle alors $H = \mathcal{A}_5$. Si H contient une double transposition $\sigma = (i,j)(k,\ell)$, alors, en posant $\rho = (k,\ell,m)$ avec m le cinquième élément, on a $\sigma\rho\sigma\rho^{-1} = (k,\ell,m) \in H$ donc $H = \mathcal{A}_5$. Si H contient un 5-cycle alors il contient un 5-Sylow de $\mathcal{A}_5$ et donc tous et donc les 24 cycles de longueur 5; mais 25 ne divise pas $\mathrm{card}(\mathcal{A}_5) = 60$ donc H contient d'autres éléments donc un 3-cycle ou une double transposition.

Montrons que $\mathcal{A}_{n-1}$ simple entraîne $\mathcal{A}_n$ simple (pour $n \geq 6$). Soit $H \triangleleft \mathcal{A}_n$ un sous-groupe non trivial. Considérons $G_i = \{\sigma \in \mathcal{A}_n \mid \sigma(i) = i\} \cong \mathcal{A}_{n-1}$, on a $H \cap G_i \triangleleft G_i$ donc $H \cap G_i = G_i$ ou $\{id\}$. Si $G_i \subset H$ alors H contient un 3-cycle et $H = \mathcal{A}_n$. Il nous reste à montrer qu'on ne peut avoir $H \cap G_i = \{id\}$. Soit donc $\sigma \in H \setminus \{id\}$. On a $\sigma(1) = i \neq 1$, choisissons $j \neq 1, i$ alors $\sigma(j) = k$ et on peut choisir $\ell, m \notin \{1,i,j,k\}$. Soit $\rho = (j,\ell,m) \in \mathcal{A}_n$ alors $\tau := \rho^{-1}\sigma^{-1}\rho\sigma$ est dans H et $\tau(1) = 1$ alors que $\tau(j) = \ell$ onc $\tau \in H \cap G_1 \setminus \{id\}$, ce qui est une contradiction.

Enfin montrons que $H \triangleleft \mathcal{S}_n$ et $H \neq \{id\}, \mathcal{S}_n$ entraîne $H = \mathcal{A}_n$ pour $n \geq 5$ (les cas $n \leq 4$ sont laissés en exercice). On a ou bien $H \cap \mathcal{A}_n = \mathcal{A}_n$ mais alors $H = \mathcal{A}_n$ ou bien $H \cap \mathcal{A}_n = \{id\}$ mais alors $\mathrm{card}(H) = 2$ ce qui est impossible car les conjugués d'un produit de transposition ne lui sont pas tous égaux. L'analyse des cas où $n \leq 4$ est laissée au lecteur. □

Remarques. Le groupe $\mathcal{A}_5$ a pour cardinal 60, c'est le plus petit groupe simple (non commutatif); Le groupe $\mathcal{A}_5$ contient 5 "copies" de $\mathcal{A}_4$ (les stabilisateurs de $1,2,3,4,5$) qui contiennent chacun une copie du groupe de Klein, ce qui fournit les cinq 2-sous-groupes de Sylow. En effet si on écrit $K \subset \mathcal{A}_4 \subset \mathcal{A}_5$ on sait que $\mathcal{A}_4$ normalise K et en fait doit être égal au normalisateur de K dans $\mathcal{A}_5$ car K ne peut pas être distingué; il y a donc bien $5 = (\mathcal{A}_5 : \mathcal{A}_4)$ sous-groupes de Sylow.

Montrons qu'un groupe simple G de cardinal 60 est isomorphe à $\mathcal{A}_5$. Un tel groupe n'admet pas d'homomorphisme non trivial vers $\mathcal{S}_4$ (sinon le noyau contredirait la simplicité de G) donc pas d'action non triviale sur des ensembles de cardinal ≤ 4. D'après les théorèmes de Sylow, le nombre de 2-sous-groupes de Sylow est donc *a priori* 5 ou 15, Le nombre de 5-sous-groupes de Sylow est 6 (donc il y a 24 éléments d'ordre 5) et le nombre de 3-sous-groupes de Sylow est 10 (donc il y a 20 éléments d'ordre 3). Supposons $n_2 = 5$, alors l'action de G sur les 2-sous-groupes de Sylow donne une injection $G \hookrightarrow \mathcal{S}_5$. L'image est d'indice deux donc distinguée donc c'est $\mathcal{A}_5$. Supposons $n_2 = 15$, alors un décompte des éléments montre qu'il existe deux 2-sous-groupes de Sylow tels que $\mathrm{card}(P_1 \cap P_2) > 1$ (sinon l'union des 2-sous-groupes de Sylow aurait pour cardinal $(15 \times 3) + 1 = 46$). Soit $x \in P_1 \cap P_2 \setminus \{e\}$, alors P_1 et P_2, étant commutatifs, sont dans le commutateur $C(\sigma)$ qui est donc de cardinal $4m$ avec $m > 1$. Le groupe G agit transitivement sur $G/C(\sigma)$ qui est de cardinal $15/m$. mais on a vu que $m > 1$ et que $15/m \geq 5$ donc $G/C(\sigma)$ a pour cardinal 5 et on en tire un homorphisme $\rho : G \to \mathcal{S}_5$ qui, comme précédemment doit être un isomorphisme avec $\mathcal{A}_5$. (Bien entendu la possibilité $n_2 = 15$ est impossible *a posteriori*).

Exercices (illustrations géométriques). 1) Soit K un corps commutatif, montrer que l'action naturelle de $\mathrm{SL}(2,K)$ sur K^2 induit une action transitive sur $\mathbf{P}^1(K)$ (l'ensemble des droites de K^2 passant par l'origine) et que son noyau est $\{\pm Id\}$. On note $\mathrm{PSL}(2,K)$ le quotient de $\mathrm{SL}(2,K)$ par $\{\pm Id\}$. En déduire les isomorphismes suivants :
(i) $\mathrm{PSL}(2,\mathbf{Z}/2\mathbf{Z}) \cong \mathcal{S}_3$
(ii) $\mathrm{PSL}(2,\mathbf{Z}/3\mathbf{Z}) \cong \mathcal{A}_4 \subset \mathcal{S}_4$
(iii) $\mathrm{PSL}(2,\mathbf{Z}/5\mathbf{Z}) \cong \mathcal{A}_5 \subset \mathcal{A}_6$

2) Considérons G le groupe du cube (qu'on peut supposer centré en l'origine) et faisons-le agir sur les quatre "grandes" diagonales. Montrer que cette action induit un homomorphisme $\rho : G \to \mathcal{S}_4$ dont le noyau est $\{\pm Id\}$ et en déduire que

$$G \cong \mathcal{S}_4 \times \{\pm Id\}.$$

Décrire les isométries correspondant aux transpositions, cycles, etc.

A.8. Le b-a-ba de la classification des groupes finis.

On donne quelques compléments "culturels" sur les groupes, leurs descriptions, pour la plupart sans preuves.

A.8.1. Théorème de Jordan-Holder.

Si un groupe G possède un sous-groupe distingué H non trivial (distinct de G et $\{e\}$), on peut écrire une suite exacte $1 \to H \to G \to G/H \to 1$ et considérer qu'on a ramené l'étude de G à l'étude de deux groupes plus petits : H et G/H. Toutefois il est inexact de penser que l'on sait tout sur G si l'on ne connait que H et G/H : par exemple si $\mathbf{Z}/3\mathbf{Z} \cong H \triangleleft G$ et $G/H \cong \mathbf{Z}/2\mathbf{Z}$ alors $G \cong \mathbf{Z}/6\mathbf{Z}$ ou $\mathcal{S}_3$. Ces considérations nous amènent naturellement aux deux définitions suivantes.

Définition. Un groupe est *simple* s'il n'admet aucun sous-groupe distingué non trivial.

L'exemple de groupe simple le plus facile à décrire est $\mathbf{Z}/p\mathbf{Z}$, ce sont d'ailleurs les seuls groupes simples abéliens; on les exclut parfois par convention (parce qu'ils sont trop simples!). On a vu que les groupes $\mathcal{A}_n$ étaient simples lorsque $n \geq 5$.

Définition. Une *suite de composition* d'un groupe G est la donnée d'une suite de sous-groupes emboités i.e. $G = G_0 \supset G_1 \supset \ldots \supset G_n = \{e\}$ telle que $G_{i+1} \triangleleft G_i$ et G_i/G_{i+1} est simple. Une autre suite de composition $G = G'_0 \supset G'_1 \supset \ldots \supset G'_m = \{e\}$ est dite *équivalente* à la première si $m = n$ et il existe une permutation $\sigma : [1, n] \to [1, n]$ telle que $G_{\sigma(i)}/G_{\sigma(i)+1} \cong G'_i/G'_{i+1}$.

Remarquons que demander que G_i/G_{i+1} soit simple équivaut à demander que la suite G_i soit maximale au sens que si $G_i \supset H \supset G_{i+1}$ avec $H \triangleleft G_i$ alors $H = G_i$ ou G_{i+1}.

Théorème. (Jordan-Holder) *Soit G un groupe fini, alors G admet une suite de composition qui est unique à équivalence près.*

Preuve. La première partie est claire, démontrons donc la deuxième. Supposons données deux suites de composition $G = H_0 \supset H_1 \supset \ldots \supset H_m$ et $G = K_0 \supset K_1 \supset \ldots \supset K_n$ et supposons (raisonnement par induction) que le théorème est déjà démontré pour les groupes admettant une suite de composition de longueur $\leq m - 1$. Si $H_1 = K_1$ alors on peut appliquer l'hypothèse de récurrence à H_1 et conclure. Dans le cas contraire on introduit une suite de composition de $H_1 \cap K_1$ notée (attention à la numérotation) $H_1 \cap K_1 = L_2 \supset L_3 \supset \ldots \supset L_r$ de sorte que l'on a le diagramme suivant où les flèches indiquent que le groupe en bas de la flèche est un sous-groupe distingué du groupe au-dessus.

$$
\begin{array}{ccccc}
 & & G & & \\
 & \swarrow & & \searrow & \\
H_1 & & & & K_1 \\
\downarrow & \searrow & & \swarrow & \downarrow \\
H_2 & & H_1 \cap K_1 & & K_2 \\
\downarrow & & \downarrow & & \downarrow \\
H_3 & & L_3 & & K_3 \\
\vdots & & \vdots & & \vdots \\
\downarrow & & \downarrow & & \downarrow \\
\{e\} = H_m & & L_r & & K_n = \{e\}
\end{array}
$$

De plus tous les quotients sont simples; c'est clair par construction, sauf pour les inclusions de $H_1 \cap K_1$ dans K_1 et H_1 où cela résulte du lemme suivant

Lemme. *Dans la situation ci-dessus, si $H_1 \neq K_1$ alors $G/H_1 \cong K_1/H_1 \cap K_1$ et $G/K_1 \cong H_1/H_1 \cap K_1$. En particulier $K_1/H_1 \cap K_1$ et $H_1/H_1 \cap K_1$ sont simples.*

Preuve. L'application $K_1 \hookrightarrow K_1H_1 \to K_1H_1/H_1$ a pour noyau $H_1 \cap K_1$ d'où l'isomorphisme classique $K_1/H_1 \cap K_1 \cong K_1H_1/H_1$. Par ailleurs on a $K_1 \triangleleft K_1H_1 \triangleleft G$, mais, vues les hypothèses, $K_1 \neq K_1H_1$ donc $H_1K_1 = G$. □

Suite de la preuve. On dispose donc de deux suites de composition de H_1 de longueur $m-1$ et $r-1$; on peut donc appliquer l'hypothèse de récurrence et conclure que $m = r$ et les quotients $\{H_1/H_2, \ldots, H_{m-1}/H_m\}$ et $\{H_1/H_1 \cap K_1, H_1 \cap K_1/L_3, \ldots, L_{r-1}/L_r\}$ sont isomorphes deux à deux. Le même raisonnement appliqué aux deux suites de composition de K_1 montre que $n = r$ et que les quotients $\{K_1/K_2, \ldots, K_{n-1}/K_n\}$ et $\{K_1/H_1 \cap K_1, H_1 \cap K_1/L_3, \ldots, L_{r-1}/L_r\}$ sont isomorphes deux à deux. On en tire, en se souvenant du lemme précédent, que les quotients $\{G/H_1, H_1/H_2, \ldots, H_{m-1}/H_m\}$ sont isomorphes (à permutation près) aux quotients $\{K_1/H_1 \cap K_1, H_1/H_1 \cap K_1, H_1 \cap K_1/L_3, \ldots, L_{r-1}/L_r\}$ donc également aux quotients $\{K_1/K_2, H_1/H_1 \cap K_1, K_2/K_3, \ldots, K_{n-1}/K_n\}$ et enfin aux quotients $\{K_1/K_2, G/K_1, K_2/K_3, \ldots, K_{n-1}/K_n\}$ comme annoncé. □

Il est naturel d'introduire la définition suivante qui a par ailleurs une grande importance historique : d'après Galois, les équations polynomiales $P(x) = 0$ dont on peut exprimer les racines à l'aides des opérations de corps et de radicaux $\sqrt[n]{}$ sont celles qui ont un groupe résoluble.

Définition. Un groupe G est *résoluble* s'il existe une suite $G = G_0 \supset G_1 \supset \ldots \supset G_n = \{e\}$ telle que $G_{i+1} \triangleleft G_i$ et G_i/G_{i+1} est abélien.

Si le groupe G est fini, il revient au même de demander que ses facteurs de Jordan-Holder soient isomorphes à $\mathbf{Z}/p\mathbf{Z}$. Un des théorèmes les plus difficiles de la théorie des groupes finis (Feit-Thomson) dit qu'un groupe de cardinal impair est toujours résoluble.

Exercices. Montrer que, si $H \triangleleft G$ alors G est résoluble si et seulement si H *et* G/H sont résolubles. Montrer qu'un groupe de cardinal ≤ 100 et $\neq 60$ est résoluble. Montrer qu'un groupe G de cardinal $2n$ avec n impair contient un sous-groupe distingué d'indice 2 et en particulier n'est pas simple (Indication : l'action par translation induit $\rho : G \to \mathcal{S}_{2n}$, montrer que $\mathrm{Ker}(\epsilon \circ \rho)$ est d'indice 2 dans G). En admettant le théorème de Feit-Thomson, montrer que G est résoluble.

A.8.2. Classification des petits groupes (début).

On peut chercher à classer les "petits" groupes à isomorphisme près. Si l'on note $\gamma(n)$ le nombre de classes d'isomorphisme de groupes de cardinal n, on a déjà vu que $\gamma(p) = 1$, $\gamma(p^2) = 2$ $\gamma(p^3) = 5$ et $\gamma(pq) = 2$ ou 1 suivant que $q \equiv 1 \bmod p$ ou non. Si on poursuit les calculs, on peut obtenir par exemple la table suivante pour $n \leq 34$:

n	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
$\gamma(n)$	1	1	2	1	2	1	5	2	2	1	5	1	2	1	14	1

18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
5	1	5	2	2	1	15	2	2	5	4	1	4	1	51	1	2

Pour $n = 24$ les calculs sont plus longs (voir exercice ci-dessous), pour $n = 16$ ou 32, ils deviennent plus délicats, pour les autres valeurs, donnons sans preuve une description des classes d'isomorphismes.

(a) Pour $n = 8$ les cinq groupes sont les trois groupes abéliens $(\mathbf{Z}/2\mathbf{Z})^3$, $\mathbf{Z}/4\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$ et $\mathbf{Z}/8\mathbf{Z}$ et les deux non commutatifs D_4 et $H_8 = \{\pm 1, \pm i, \pm j, \pm k\}$.

(b) Pour $n = 12$ les cinq groupes sont $\mathbf{Z}/12\mathbf{Z}$, $\mathbf{Z}/6\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$, $D_3 \times \mathbf{Z}/2\mathbf{Z}$, $\mathcal{A}_4$ et le produit semidirect $\mathbf{Z}/3\mathbf{Z} \rtimes_\phi \mathbf{Z}/4\mathbf{Z}$ où $\phi : \mathbf{Z}/4\mathbf{Z} \to \mathbf{Z}/2\mathbf{Z} \cong (\mathbf{Z}/3\mathbf{Z})^*$ est l'homomorphisme surjectif qui associe à $x \bmod 4$ l'élément $x \bmod 2$ (remarquer que D_6 est isomorphe à $D_3 \times \mathbf{Z}/2\mathbf{Z}$).

(c) Pour $n = 18$ les cinq groupes sont $\mathbf{Z}/18\mathbf{Z}$, $\mathbf{Z}/6\mathbf{Z} \times \mathbf{Z}/3\mathbf{Z}$, D_9, $D_3 \times \mathbf{Z}/3\mathbf{Z}$ et un produit semi-direct $(\mathbf{Z}/3\mathbf{Z})^2 \rtimes_\phi \mathbf{Z}/2\mathbf{Z}$ où $\phi : \mathbf{Z}/2\mathbf{Z} \to \mathrm{GL}(2, \mathbf{Z}/3\mathbf{Z})$ est donné par $\phi(1) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

(d) Pour $n = 20$ les cinq groupes sont $\mathbf{Z}/20\mathbf{Z}$, $\mathbf{Z}/10\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$, D_{10}, $D_5 \times \mathbf{Z}/2\mathbf{Z}$ et un produit semi-direct $\mathbf{Z}/5\mathbf{Z} \rtimes_\phi \mathbf{Z}/4\mathbf{Z}$ où $\phi : \mathbf{Z}/4\mathbf{Z} \to (\mathbf{Z}/5\mathbf{Z})^*$ est un isomorphisme.

(e) Pour $n = 28$ les quatre groupes sont $\mathbf{Z}/28\mathbf{Z}$, $\mathbf{Z}/14\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$, D_{14}, $D_7 \times \mathbf{Z}/2\mathbf{Z}$.
(f) Pour $n = 30$ les quatre groupes sont $\mathbf{Z}/30\mathbf{Z}$, $D_3 \times \mathbf{Z}/5\mathbf{Z}$, D_{15}, $D_5 \times \mathbf{Z}/3\mathbf{Z}$.

Exercices. Démontrer les affirmations $(a) \dots (f)$. Classifier les groupes de cardinal 24 ainsi:

(i) On appelle P (resp. Q) un 2-sous-groupe de Sylow (resp. un 3-sous-groupe de Sylow) et n_2 (resp. n_3) le nombre de 2-sous-groupes de Sylow (resp. de 3-sous-groupes de Sylow). Montrer que $n_2 = 1$ ou 3 (resp. $n_3 = 1$ ou 4). En déduire que soit $G = Q \rtimes_\phi P$ soit $G = P \rtimes_\phi Q$ soit $n_2 = 3$ et $n_3 = 4$ et alors $G \cong \mathcal{S}_4$.

(ii) Si $n_2 = 1$, montrer qu'il y a 8 groupes possibles. Si $P \cong (\mathbf{Z}/2\mathbf{Z})^3$ les groupes possibles sont $\mathbf{Z}/6\mathbf{Z} \times (\mathbf{Z}/2\mathbf{Z})^2$, $\mathbf{Z}/2\mathbf{Z} \times \mathcal{A}_4$ ou un produit semi-direct $(\mathbf{Z}/2\mathbf{Z})^3 \rtimes_\phi \mathbf{Z}/3\mathbf{Z}$ où $\phi : \mathbf{Z}/3\mathbf{Z} \to \mathrm{GL}(3, \mathbf{Z}/2\mathbf{Z})$ est donné par la permutation circulaire des coordonnées; si $P \cong \mathbf{Z}/4\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$ alors $G \cong \mathbf{Z}/12\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$; si $P \cong \mathbf{Z}/8\mathbf{Z}$ alors $G \cong \mathbf{Z}/24\mathbf{Z}$; si $P \cong D_4$ alors $G \cong D_4 \times \mathbf{Z}/3\mathbf{Z}$; si $P \cong H_8$ alors ou bien $G \cong H_8 \times \mathbf{Z}/3\mathbf{Z}$ ou bien
$$G = H_8 \rtimes_\phi \mathbf{Z}/3\mathbf{Z} \cong \left\{ \pm 1, \pm i, \pm j, \pm k, \frac{\pm 1 + \pm i + \pm j + \pm k}{2} \right\}.$$

(iii) Si $n_3 = 1$ et $n_2 = 3$ il ya 6 groupes possibles : $D_3 \times (\mathbf{Z}/2\mathbf{Z})^2$, $D_6 \times \mathbf{Z}/2\mathbf{Z}$, $D_3\mathbf{Z}/4\mathbf{Z}$, D_{12} et des produits semidirects $\mathbf{Z}/3\mathbf{Z} \rtimes_\phi P$ (avec $\phi : P \to \mathbf{Z}/2\mathbf{Z} = \mathrm{Aut}(\mathbf{Z}/3\mathbf{Z})$ surjective et $P = \mathbf{Z}/8\mathbf{Z}$, D_4 ou H_8).

A.8.3. Classification des groupes simples finis (aperçu).

La classification exhaustive des groupes simples finis (non abéliens) a été achevé au début des années 80. Le plus petit groupe simple est $\mathcal{A}_5$, il a cardinal 60; le suivant est $\mathrm{PSL}(2, \mathbf{Z}/7\mathbf{Z})$, il a cardinal 168. On peut répartir les groupes simples non abéliens en 17 familles (infinies) auxquelles il faut ajouter 26 groupes exceptionnels appelés *groupes sporadiques*. Je ne dirai presque rien sur les groupes sporadiques sinon écrire dans un tableau leurs cardinaux et donner les noms de leurs découvreurs : Mathieu, Held, Janko, Conway, Lyons, O'Nan, Fischer, Fischer-Griess, Higman-Sims, Suzuki, McLaughlin, Rudvalis. Je ne vais pas décrire toutes les familles mais les principales.

Le groupe alterné. Lorsque $n \geq 5$ on a vu que le groupe $\mathcal{A}_n$ est simple.

Les autres familles sont des groupes *de type de Lie*, c'est-à-dire qu'ils correspondent à des groupes de Lie comme $\mathrm{SL}(n, \mathbf{R})$ sauf qu'au lieu de considérer des coefficients dans $\mathbf{R}$ ou $\mathbf{C}$, on choisit les coefficients dans un corps fini $\mathbf{F}_q$ où $q = p^r$ désigne le cardinal du corps.

Le groupe spécial linéaire. Le groupe $\mathrm{PSL}(n, \mathbf{F}_q) := \mathrm{SL}(n, \mathbf{F}_q)/Z$ est simple pour $n \geq 2$ où Z désigne le centre, c'est-à-dire le sous-groupe $\{aI \mid a \in \mathbf{F}_q,\ a^n = 1\}$. [Exceptions : $n = 2$ et $q = 2$ ou 3]

Le groupe symplectique. Le groupe $\mathrm{PSp}(2n, \mathbf{F}_q) := \mathrm{Sp}(2n, \mathbf{F}_q)/Z$ est simple pour $n \geq 2$ où Z désigne le centre, c'est-à-dire le sous-groupe $\{\pm I\}$. Rappelons que le groupe symplectique est le groupe des matrices préservant une forme bilinéaire alternée non dégénérée; explicitement, si $J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$ on peut écrire $\mathrm{Sp}(2n, K) = \{A \in \mathrm{GL}(2n, K) \mid {}^tAJA = J\}$. [Exceptions : $n = 1$ et $q = 2$ ou 3; $n = 2$ et $q = 2$]

Le groupe orthogonal (on suppose ici que la caractéristique est $\neq 2$). Le commutateur du groupe orthogonal (groupe des isométries) d'une forme quadratique se note Ω. Sur un espace de dimension n sur $\mathbf{F}_q$ on distingue trois cas. Si n impair on choisit $Q(x) = x_1x_2 + \ldots + x_{n-2}x_{n-1} + x_n^2$, on note $\mathrm{SO}(n, \mathbf{F}_q)$ le groupe des rotations, $\Omega_n(\mathbf{F}_q)$ le sous-groupe des commutateurs et $\mathrm{P}\Omega_n(\mathbf{F}_q)$ le quotient par le centre. Alors $\mathrm{P}\Omega_n(\mathbf{F}_q)$ est simple pour $n \geq 5$ (impair). Si $n = 2m$ est pair on distingue deux formes quadratiques $Q^+(x) = x_1x_2 + \ldots + x_{n-3}x_{n-2} + x_{n-1}x_n$ et $Q^-(x) = x_1x_2 + \ldots + x_{n-3}x_{n-2} + x_{n-1}^2 - ax_n^2$ (avec $a \in \mathbf{F}_q^* \setminus \mathbf{F}_q^{*2}$), on note $\mathrm{SO}^+(n, \mathbf{F}_q)$ (resp. $\mathrm{SO}^-(n, \mathbf{F}_q)$) le groupe des rotations, $\Omega_n^+(\mathbf{F}_q)$ (resp. $\Omega_n^-(\mathbf{F}_q)$) le sous-groupe des commutateurs et $\mathrm{P}\Omega_n^+(\mathbf{F}_q)$ (resp. $\mathrm{P}\Omega_n^+(\mathbf{F}_q)$) le quotient par le centre. Alors $\mathrm{P}\Omega_n^+(\mathbf{F}_q)$ (resp. $\mathrm{P}\Omega_n^-(\mathbf{F}_q)$) est simple pour $n \geq 6$ (pair). Pour les petites dimensions, $\mathrm{SO}(2)$ est commutatif, $\mathrm{P}\Omega(3, \mathbf{F}_q) \cong \mathrm{PSL}(2, \mathbf{F}_q)$ alors que $\mathrm{P}\Omega^+(4, \mathbf{F}_q) \cong \mathrm{PSL}(2, \mathbf{F}_q) \times \mathrm{PSL}(2, \mathbf{F}_q)$ et $\mathrm{P}\Omega^-(4, \mathbf{F}_q) \cong \mathrm{PSL}(2, \mathbf{F}_{q^2})$. On a également $\mathrm{P}\Omega^+(6, \mathbf{F}_q) \cong \mathrm{PSL}(4, \mathbf{F}_q)$ et $\mathrm{P}\Omega^-(6, \mathbf{F}_q) \cong \mathrm{PSU}(4, \mathbf{F}_{q^2})$. Plus curieusement $\mathrm{card}(\mathrm{P}\Omega(2m+1, \mathbf{F}_q)) = \mathrm{card}(\mathrm{PSp}(2m, \mathbf{F}_q))$ et on a $\mathrm{P}\Omega(5, \mathbf{F}_q) \cong \mathrm{PSp}(4, \mathbf{F}_q))$ mais l'égalité de cardinaux ne correspond pas à un isomorphisme pour $m \geq 3$.

Le groupe unitaire. On note $x^\sigma := x^q$ l'automorphisme involutif de $\mathbf{F}_{q^2}$. On note $\mathrm{U}_n(\mathbf{F}_{q^2})$ le groupe des isométries de la forme hermitienne $H(x) = x_1x_1^\sigma + \ldots + x_nx_n^\sigma$, puis $\mathrm{SU}_n(\mathbf{F}_{q^2}) = \mathrm{U}_n(\mathbf{F}_{q^2}) \cap \mathrm{SL}_n(\mathbf{F}_{q^2})$ et enfin

$\mathrm{PSU}_n(\mathbf{F}_{q^2})$ le quotient par le centre $Z = \{aI \mid a^{q+1} = 1\}$. Le groupe $\mathrm{PSU}_n(\mathbf{F}_{q^2})$ est simple pour $n \geq 2$. [Exceptions : $n = 2$ et $q^2 = 4$ ou 9; $n = 3$ et $q^2 = 4$]

Il existe en plus des groupes de type de Lie exceptionnels G_2, F_4, E_6, E_7 et E_8 de dimension 14, 52, 78, 133 et 248 qui conduisent aussi à des groupes simples finis. Enfin il existe des formes "tordues" de certains de ces groupes que je ne décrirai pas. (Voir les deux tableaux).

A.8.4. Groupes définis par générateurs et relations.

La manière la plus commode de "décrire un groupe" à un ordinateur est de lui donner des générateurs avec les relations vérifiées par ceux-ci. Pour décrire cela plus précisément, commençons par construire le groupe "engendré par m éléments sans relations". Soit S un ensemble (qui sera fini dans nos applications), on définit l'ensemble des *mots* sur S comme l'ensemble des suites $x_1 x_2 \dots x_n$ avec $n \geq 1$ et $x_i \in S \times \{\pm 1\}$ [pour simplifier la notation on pourra écrire x pour $(x, +1)$ et x^{-1} pour $(x, -1)$] auquel on ajoute le "mot vide" que l'on note e. On notera $M(S)$ l'ensemble des mots. On peut définir une multiplication des mots par la formule $(x_1 x_2 \dots x_n) * (y_1 y_2 \dots y_s) = x_1 x_2 \dots x_n y_1 y_2 \dots y_s$; cette multiplication est clairement associative. On définit ensuite une relation d'équivalence $\mathcal{R}$ sur $M(S)$ comme celle engendrée par les relations $xx^{-1}\mathcal{R}e$ (c'est-à-dire que pour tout mot m, n on impose $mxx^{-1}n\mathcal{R}mn$ et on étend la relation par transitivité). Remarquons qu'on peut introduire des représentants canoniques des classes en choisissant le mot le plus court ou *mot réduit* (pourquoi est-il unique?). L'ensemble S est naturellement inclus dans $M(S)$ (resp. dans $M(S)/\mathcal{R}$) si l'on identifie un élément $x \in S$ et le mot à une lettre. On notera $i : S \hookrightarrow M(S)/\mathcal{R}$ cette inclusion. La multiplication sur $M(S)$ induit une multiplication sur $M(S)/\mathcal{R}$ (elle "passe au quotient"). On peut aussi définir le produit de deux mots réduits comme le mot réduit obtenu à partir du produit des deux mots.

Théorème. *L'ensemble $M(S)/\mathcal{R}$ muni de la loi $*$ est un groupe, appelé groupe libre sur S et noté $L(S)$. Il vérifie la propriété universelle suivante : pour tout groupe G et toute application $f : S \to G$, il existe une unique homomorphisme $\phi : L(S) \to G$ tel que $\phi \circ i = f$.*

Preuve. La loi est automatiquement associative et a pour élément neutre le mot vide (sa classe). L'inverse de la classe d'un mot $u_1 u_2 \dots u_m$ (avec u_i ou $u_i^{-1} \in S$) est la classe de $u_m^{-1} \dots u_2^{-1} u_1^{-1}$. Pour la deuxième partie, posons $\phi(x_1^{\epsilon_1} \dots x_m^{\epsilon_m}) = f(x_1)^{\epsilon_1} \dots f(x_m)^{\epsilon_m}$. On vérifie aisément que $\phi : L(S) \to G$ est bien défini et a les propriétés voulues. □

Corollaire. *Tout groupe peut s'écrire comme quotient d'un groupe libre. Plus précisément si G admet admet pour générateur un sous-ensemble S, alors G est un quotient de $L(S)$.*

Preuve. Il suffit de considérer l'application $\phi : L(S) \to G$ associée par la propriété universelle à l'injection $S \hookrightarrow G$ et de remarquer que $\phi(L(S))$ est un sous-groupe contenant S donc égal à G tout entier. On a donc bien $G \cong L(S)/\operatorname{Ker}(\phi)$. □

Si on écrit $G = L(S)/N$ et si R est un ensemble de générateurs de N on dit qu'on a une *présentation de G par générateurs et relations.* Précisons qu'il y deux notions de "générateurs" : le sous-groupe engendré par R n'est pas forcément distingué, ici on doit prendre pour N le plus petit sous-groupe distingué contenant R.

Exemples. Une présentation de $\mathbf{Z}/n_1\mathbf{Z} \times \dots \mathbf{Z}/n_s\mathbf{Z}$ est donnée par s générateurs $x_1, \dots, x_s$ et les relations $x_i^{n_i} = e$ et $x_i y_i x_i^{-1} y_i^{-1} = e$. Une présentation du groupe dièdral D_n est donnée par deux générateurs ρ et σ avec les relations $\rho^n = e$ et $\sigma^2 = (\sigma\rho)^2 = e$. Une présentation possible de $\mathcal{S}_n$ (penser à x_i comme la transposition $(i, i+1)$) est

$$\begin{cases} \text{générateurs : } x_i \text{ pour } 1 \leq i \leq n-1 \\ \text{relations : } x_i^2 \text{ pour } 1 \leq i \leq n-1, \qquad (x_i x_{i+1})^3 \text{ pour } 1 \leq i \leq n-1 \text{ et } x_i x_j x_i^{-1} x_j^{-1} \text{ pour } |i-j| \geq 2 \end{cases}$$

Les groupes libres présentent des analogies avec les espaces vectoriels (ou les groupes abéliens libres) mais réservent aussi quelques surprises. Par exemple deux groupes libres $L(X)$ et $L(Y)$ sont isomorphes si et seulement si $\operatorname{card}(X) = \operatorname{card}(Y)$ et un sous-groupe d'un groupe libre est encore libre (théorème de Nielsen-Schreier). Cependant un sous-groupe même de L_n (le groupe libre à n générateurs) peut être de n'importe quel rang : ainsi si $(L_n : H) = m$ alors H est un groupe libre à $(n-1)m + 1$ générateurs; pire, si $n > 1$

alors L_n contient des sous-groupes qui ne sont pas de type fini! Le sous-groupe des commutateurs est un tel exemple.

Le groupe libre à n générateurs est important en topologie puisque c'est le groupe fondamental du plan privé de n points. Le groupe fondamental d'une variété s'écrit souvent naturellement commme quotient d'un groupe libre. Par exemple le groupe fondamental d'une surface compacte à g trous est le quotient du groupe libre à $2g$ générateurs $x_1, y_1, x_2, y_2, \ldots, x_g, y_g$ par le sous-groupe distingué engendré par la relation $x_1 y_1 x_1^{-1} y_1^{-1} \ldots x_g y_g x_g^{-1} y_g^{-1}$.

A.8.5. Problèmes de Burnside.

Au début du siècle, Burnside a posé la question de savoir si un groupe de type fini et d'exposant fini est nécessairement fini; à défaut il a également demandé si le nombre de classes d'isomorphisme de groupes finis d'exposant n, possédant m générateurs est fini. On peut formaliser cela à l'aide de la notion de groupe libre à m générateurs.

Considérons L_m le groupe libre à m générateurs et $N = N_n$ le sous-groupe engendré par les éléments $\{g^n \mid g \in L_m\}$ (i.e. le plus petit sous-groupe de L_m contenant tous les g^n); remarquons que $N_n \triangleleft L_m$ puisque $x(g_1^n \ldots g_r^n)x^{-1} = \left(xg_1x^{-1}\right)^n \ldots \left(xg_rx^{-1}\right)^n$.

(1) Première question : pour quelles valeurs de m et n le groupe $B(m,n) := L_m/N_n$ est-il fini?
(2) Deuxième question : les quotients finis de $B(m,n)$ sont-ils en nombre fini (à isomorphisme près)?

La première question s'appelle traditionnellement *problème de Burnside* et la deuxième *problème restreint de Burnside*. Il est clair que si $B(m,n)$ est fini, alors ses quotients sont tous finis et qu'il n'y en a qu'un nombre fini. Cependant justement la réponse à la première question est négative en général. Plus précisément il est facile de voir que $B(1,n) = \mathbf{Z}/n\mathbf{Z}$ et $B(m,2) = (\mathbf{Z}/2\mathbf{Z})^m$. On sait (Burnside, Sarov et Hall) que $B(m,3)$, $B(m,4)$ et $B(m,6)$ sont finis; mais dans l'autre sens on sait que $B(m,n)$ est infini lorsque $m > 1$ et n est impair > 665 (Novikov et Adjan) ou n pair > 8000. Par contre le problème restreint de Burnside admet une réponse positive. Hall et Higman on montré en 1956 que, modulo des résultats de classification des groupes simples finis (qui ont été démontrés dans les années 80) le cas $n = p_1^{r_1} \ldots p_s^{r_s}$ découlait du cas $p_i^{r_i}$. A la fin des années 80, Zelmanov a ensuite prouvé que les quotients finis de $B(m,p^r)$ étaient, à isomorphisme près, en nombre fini.

Exercice. Montrer que les groupes de cardinal de la forme $N = p^a$, pq, pqr, $4pq$ (sauf 60) ou $2(2m+1)$ ne sont pas simples. Indications : les deux premiers cas a été traités en cours, pour les deux suivants utiliser les théorèmes de Sylow, dans le dernier cas considérer l'action de G sur lui-même par translation et le morphisme $\rho : G \to \mathcal{S}_{2(2m+1)}$ correspondant, montrer que l'image d'un élément d'ordre 2 a pour signature -1 et conclure. On montrera plus tard (chapitre F) qu'un groupe de cardinal $p^a q^b$ est résoluble. Montrer qu'un groupe de cardinal $N \leq 200$ n'est pas simple sauf si $N \in \{60, 168\}$.

APPENDICE : LA LISTE DES GROUPES SIMPLES FINIS

I. Les 17 familles infinies de groupes simples finis non abéliens et leurs cardinaux

Groupe	Autre nom	cardinal
$\mathcal{A}_n$		$\frac{n!}{2}$
$A_n(q)$ (1)	$PSL_{n+1}(q)$	$\frac{1}{(n+1,q-1)} q^{\frac{n(n+1)}{2}} \prod_{i=2}^{n+1}(q^i-1)$
${}^2A_n(q)$ (1)	$PSU_{n+1}(q)$	$\frac{1}{(n+1,q+1)} q^{\frac{n(n+1)}{2}} \prod_{i=2}^{n+1}(q^i-(-1)^i)$
$B_n(q)$ (2)	$P\Omega_{2n+1}(q)$	$\frac{1}{(2,q-1)} q^{n^2} \prod_{i=1}^{n}(q^{2i}-1)$
${}^2B_n(q)$ (3)	$Sz(q)$	$q^2(q-1)(q^2+1)$
$C_n(q)$	$PSp_{2n}(q)$	$\frac{1}{(2,q-1)} q^{n^2} \prod_{i=1}^{n}(q^{2i}-1)$
$D_n(q)$	$P\Omega^+_{2n}(q)$	$\frac{1}{(4,q^n-1)} q^{n(n-1)}(q^n-1)\prod_{i=1}^{n-1}(q^{2i}-1)$
${}^2D_n(q)$	$P\Omega^-_{2n}(q)$	$\frac{1}{(4,q^n+1)} q^{n(n-1)}(q^n+1)\prod_{i=1}^{n-1}(q^{2i}-1)$
${}^3D_4(q)$		$q^{12}(q^2-1)(q^8+q^4+1)(q^6-1)$
$G_2(q)$		$q^6(q^2-1)(q^6-1)$
${}^2G_2(q)$ (4)		$q^3(q-1)(q^3+1)$
$F_4(q)$		$q^{24}(q^2-1)(q^6-1)(q^8-1)(q^{12}-1)$
${}^2F_4(q)$ (5)		$q^{12}(q-1)(q^3+1)(q^4-1)(q^6+1)$
$E_6(q)$		$\frac{1}{(3,q-1)} q^{36}(q^2-1)(q^5-1)(q^6-1)(q^8-1)(q^9-1)(q^{12}-1)$
${}^2E_6(q)$		$\frac{1}{(3,q+1)} q^{36}(q^2-1)(q^5+1)(q^6-1)(q^8-1)(q^9+1)(q^{12}-1)$
$E_7(q)$		$\frac{1}{(2,q-1)} q^{63}(q^2-1)(q^6-1)(q^8-1)(q^{10}-1)(q^{12}-1)(q^{14}-1)(q^{18}-1)$
$E_8(q)$		$q^{120}(q^2-1)(q^8-1)(q^{12}-1)(q^{14}-1)(q^{18}-1)(q^{20}-1)(q^{24}-1)(q^{30}-1)$

(1) $A_1(2)$, $A_1(3)$ et ${}^2A_2(2)$ sont résolubles.
(2) $B_2(2) = C_2(2)$ et $G_2(2)$ ont un sous-groupe des commutateurs d'indice 2 qui est simple.
(3) définis seulement pour $q = 2^{2n+1}$; ${}^2B_2(2)$ est résoluble
(4) définis seulement pour $q = 3^{2n+1}$; ${}^2G_2(3)$ a un sous-groupe des commutateurs d'indice 3 qui est simple.
(5) définis seulement pour $q = 2^{2n+1}$; ${}^2F_4(2)$ a un sous-groupe des commutateurs d'indice 2 qui est simple.

Les notations de la colonne de gauche proviennent de la classification des algèbres de Lie simples. Les notations de la page suivante correspondent le plus souvent aux initiales des découvreurs des groupes sporadiques.

II. Les 26 groupes simples finis sporadiques et leurs cardinaux

groupe	cardinal du groupe et sa factorisation
M_{11}	$2^4.3^2.5.11 = 7920$
M_{12}	$2^6.3^3.5.11 = 95\ 040$
M_{22}	$2^7.3^2.5.7.11 = 443\ 520$
M_{23}	$2^7.3^2.5.7.11.23 = 10\ 200\ 960$
M_{24}	$2^{10}.3^3.5.7.11.23 = 244\ 823\ 040$
J_1	$2^3.3.5.7.11.19 = 175\ 560$
J_2	$2^7.3^3.5^2.7 = 604\ 800$
J_3	$2^7.3^5.5.17.19 = 50\ 232\ 960$
J_4	$2^{21}.3^3.5.7.11^3.23.29.31.37.43 = 86\ 775\ 571\ 046\ 077\ 562\ 880$
HS	$2^9.3^2.5^3.7.11 = 44\ 352\ 000$
He	$2^{10}.3^3.5^2.7^3.17 = 4\ 030\ 387\ 200$
Mc	$2^7.3^6.5^3.7.11 = 898\ 128\ 000$
Suz	$2^{13}.3^7.5^2.7.11.13 = 448\ 345\ 497\ 600$
Ly	$2^8.3^7.5^6.7.11.31.37.67 = 51\ 765\ 179\ 004\ 000\ 000$
Ru	$2^{14}.3^3.5^3.7.13.29 = 36\ 481\ 536\ 000$
$O'N$	$2^9.3^4.5.7^3.11.19.31 = 460\ 815\ 505\ 920$
Co_1	$2^{21}.3^9.5^4.7^2.11.13.23 = 4\ 157\ 776\ 806\ 543\ 360\ 000$
Co_2	$2^{18}.3^6.5^3.7.11.23 = 42\ 305\ 421\ 312\ 000$
Co_3	$2^{10}.3^7.5^3.7.11.23 = 495\ 766\ 656\ 000$
Fi_{22}	$2^{17}.3^9.5^2.7.11.13 = 64\ 561\ 751\ 654\ 400$
Fi_{23}	$2^{18}.3^{13}.5^2.7.11.13.17.23 = 4\ 089\ 460\ 473\ 293\ 004\ 800$
Fi'_{24}	$2^{21}.3^{16}.5^2.7^3.11.13.17.23.29 = 1\ 255\ 205\ 709\ 190\ 661\ 721\ 292\ 800$
F_5	$2^{14}.3^6.5^6.7.11.19 = 273\ 030\ 912\ 000\ 000$
F_3	$2^{15}.3^{10}.5^3.7^2.13.19.31 = 998\ 205\ 382\ 766\ 592\ 000$
F_2	$2^{41}.3^{13}.5^6.7^2.11.13.17.19.23.31.47 = 4\ 154\ 781\ 581\ 226\ 426\ 191\ 177\ 580\ 544\ 000\ 000$
F_1	$2^{46}.3^{20}.5^9.7^6.11^2.13^3.17.19.23.29.31.41.47.59.71$
	$= 808\ 017\ 424\ 794\ 512\ 875\ 886\ 459\ 904\ 961\ 710\ 757\ 005\ 754\ 368\ 000\ 000\ 000$

B. ANNEAUX.

B.1. Généralités et exemples.

Définition. Un *anneau* est la donnée d'un ensemble A et de deux lois internes $+$ et $\times$ telles que
(i) L'ensemble A muni de la loi $+$ est un groupe abélien d'élément neutre $0 = 0_A$.
(ii) La loi $\times$ est associative et possède un élément neutre $1 = 1_A$.
(ii) La loi $\times$ est *distributive* par rapport à l'addition, c'est-à-dire

$$\forall a, b, c \in A,\ a(b+c) = ab + ac \qquad \text{et} \qquad (a+b)c = ac + bc.$$

Si, de plus, la loi $\times$ est commutative, on dit que l'anneau est commutatif. L'anneau est un *corps* si tout élément distinct de 0_A est inversible (pour la loi $\times$).

L'exemple type d'anneau est $\mathbf{Z}$ muni de l'addition et de la multiplication usuelles, de même $\mathbf{Z}/n\mathbf{Z}$ est un anneau. Les ensembles $\mathbf{Q}$, $\mathbf{R}$ et $\mathbf{C}$ sont des corps. Nous traiterons surtout des anneaux commutatifs mais donnons néanmoins deux exemples d'anneaux non commutatifs. Si A est un anneau, $\mathrm{Mat}(n \times n, A)$ muni de l'addition et de la multiplication de matrices, est un anneau qui n'est pas commutatif dès que $n \geq 2$. Il est plus difficile de construire un corps non commutatif, voici l'exemple classique des quaternions, dû à Hamilton. L'ensemble sous-jacent est un $\mathbf{R}$-espace vectoriel de dimension 4 possédant une base qu'il est classique de noter $\{1, i, j, k\}$, l'addition est simplement l'addition de deux vecteurs, la multiplication est $\mathbf{R}$-bilinéaire et définie sur les éléments de la base par le fait que 1 est élément neutre et

$$ij = -ji = k,\ \ jk = -kj = i \ \text{ et } \ ki = -ik = j.$$

L'arithmétique (les nombres) fournit un grand nombre d'exemples d'anneaux, mais ces derniers sont présents aussi en théorie des ensembles, en analyse, etc. Si on note $\mathcal{P}(X)$ l'ensemble des parties d'un ensemble X et Δ la différence symétrique $A \,\Delta\, B := (A \cup B) \setminus (A \cap B)$ alors $(\mathcal{P}(X), \Delta, \cap)$ est un anneau commutatif qui a la particularité que pour tout x, on a $x + x = 0$ et $x \cdot x = x$. Si $K = \mathbf{R}$ ou $\mathbf{C}$ (ou plus généralement un anneau commutatif), l'ensemble $\mathcal{F}(X, K)$ des fonctions de X vers K est un anneau; si X est un espace topologique, l'ensemble $\mathcal{C}^0(X, \mathbf{R})$ des fonctions continues est également un anneau, idem avec les fonctions de classe $\mathcal{C}^k$ (si X est un ouvert de $\mathbf{R}^n$ ou plus généralement une variété différentielle); l'ensemble des fonctions de $\mathbf{R}$ dans $\mathbf{R}$ développables en série entière forme aussi un anneau. Si l'on considère l'espace vectoriel $L^1(\mathbf{R}^n)$ des fonctions intégrables (modulo les fonctions nulles presque partout) on peut le munir du produit de convolution $(f * g)(x) = \int_{\mathbf{R}^n} f(x-y)g(y)dm(y)$ et ce produit vérifie tous les axiomes de structure d'anneau commutatif sauf l'existence d'un élément neutre. Un analogue purement algébrique du produit de convolution est fourni par l'*algèbre de groupe* $A[G]$ (où A est un anneau commutatif et G un groupe) dont l'ensemble sous-jacent est l'ensemble des fonctions presque nulles de G vers A, la somme est la somme de fonctions et le produit est défini par la formule :

$$f * g(x) = \sum_{yz=x} f(y)g(z) = \sum_{y \in G} f(y)g(y^{-1}x).$$

Un élément a est *inversible* dans a s'il existe $a' \in A$ tel que $aa' = a'a = 1$.

Remarques. L'ensemble des éléments inversibles forme un groupe, pour la multiplication, noté A^*. Il faut distinguer A^* et $A \setminus \{0\}$; en effet ces deux ensembles ne sont égaux que lorsque A est un corps. Par exemple $\mathrm{Mat}(n \times n, A)^* = \mathrm{GL}(n, A) = \{U \in \mathrm{Mat}(n \times n, A) \mid \det(U) \in A^*\}$ et $(\mathbf{Z}/6\mathbf{Z})^* = \{\bar{1}, \bar{5}\}$ est un groupe à deux éléments.

Un sous-anneau B de A est un sous-ensemble tel que addition et multiplication de A induisent une structure d'anneau sur B avec même élément neutre 1_A. Par exemple $\mathbf{Z}[i] := \{a + bi \mid a, b \in \mathbf{Z}\}$ est un sous-anneau de $\mathbf{C}$ et $\mathbf{Z}[\sqrt[3]{2}] := \{a + b\sqrt[3]{2} + c\sqrt[3]{4} \mid a, b, c \in \mathbf{Z}\}$ est un sous-anneau de $\mathbf{R}$.

Soient A, B deux anneaux, on peut définir leur produit comme l'ensemble $A \times B$ muni de l'addition $(a, b) + (a', b') = (a + a', b + b')$ et de la multiplication $(a, b) \cdot (a', b') = (aa', bb')$. Il est immédiat de vérifier

qu'on obtient bien une structure d'anneau qui est commutatif si et seulement si A et B sont commutatifs. Remarquons que $A \times \{0_B\}$ est un sous-ensemble de $A \times B$ stable par addition et multiplication et possédant un élément neutre $(1_A, 0_B)$; c'est donc un anneau mais ce n'est pas un sous-anneau de $A \times B$ puisque son élément neutre n'est pas celui de $A \times B$. Observons qu'on a facilement l'égalité $(A \times B)^* = A^* \times B^*$.

Soient A, B deux anneaux, une application $f : A \to B$ est un *homomorphisme d'anneaux* si $f(x+y) = f(x) + f(y)$, $f(xy) = f(x)f(y)$ et $f(1_A) = 1_B$. C'est un *isomorphisme d'anneaux* si c'est de plus une bijection (en effet la bijection réciproque est automatiquement un homomorphisme).

On suppose désormais que les anneaux considérés sont commutatifs.

On a vu la construction du quotient d'un groupe par un sous-groupe; on peut se demander quelle propriété doit avoir un sous-groupe (additif) d'un anneau A pour que le quotient garde une structure d'anneau, la réponse est précisément la notion d'idéal.

Définition. Un *idéal* d'un anneau A est un sous-ensemble qui est un sous-groupe pour l'addition et vérifie

$$\forall a \in A, \forall j \in I,\ aj \in I.$$

Remarques et exemples. Si on ne supposait pas l'anneau A commutatif, il faudrait distinguer les idéaux à gauche (tels que $AI \subset I$) ou à droite (tels que $IA \subset I$) ou bilatère. Soit $a \in A$, l'ensemble $aA = \{ab \mid b \in A\}$ est un idéal de A appelé *idéal principal*. Tous les idéaux de l'anneau $\mathbf{Z}$ sont de la forme $a\mathbf{Z}$ puisque cela est déjà vrai pour les sous-groupes. Il est souvent intéressant de traduire les propriétés des éléments en des propriétés d'idéaux, par exemple:

Définition. Un idéal I distinct de A est *premier* si $ab \in I$ entraîne a ou b est dans I. Un élément $a \in A$ est *premier* si l'idéal aA est premier.

On voit facilement que, dans le cas $A = \mathbf{Z}$ les élément premiers sont les nombres $\pm p$ avec p nombre naturel premier (au sens usuel).

Il est immédiat de voir que le noyau d'un homomorphisme est un idéal, que l'intersection d'idéaux est un idéal, que l'image réciproque d'un idéal par un homomorphisme $f : A \to B$ est encore un idéal; par contre l'image d'un idéal n'est *a priori* un idéal que dans $f(A)$ et pas dans B. L'image réciproque d'un idéal premier par un homomorphisme d'anneaux est un idéal premier. Enfin les idéaux permettent de construire les anneaux quotient.

Théorème. *Soit A un anneau et I un idéal, il existe une unique structure d'anneau sur A/I telle que la surjection canonique $s : A \to A/I$ soit un homomorphisme d'anneaux. Ce quotient vérifie la propriété universelle suivante:*
Soit $f : A \to B$ un homomorphisme d'anneaux, l'application f se factorise en $f = \hat{f} \circ s$ si et seulement si $I \subset \mathrm{Ker}(f)$. Si cette condition est vérifiée, l'application $\hat{f} : A/I \to B$ est un homomorphisme d'anneaux et l'on a
(i) L'image $\hat{f}(A/I)$ est égale à l'image $f(A)$.
(ii) Le noyau $\mathrm{Ker}(\hat{f})$ est égal à $\mathrm{Ker}(f)/I$.

Preuve. Le théorème a déjà été prouvé en termes de groupes, il suffit donc de vérifier que $\hat{f}$, quand elle existe, est un homomorphisme d'anneaux, étant entendu que la multiplication est définie sur A/I par $(a+I)\cdot(b+I) = ab + I$ et que cette définition est indépendante du choix des représentants des classes précisément parce que I est un idéal. Soit donc $\bar{a} = a + I$ et $\bar{b} = b + I$ deux éléments de A/I; on a $\hat{f}(\bar{a} \cdot \bar{b}) = \hat{f}(\overline{ab}) = \hat{f} \circ s(ab) = f(ab) = f(a)f(b) = \hat{f}(\bar{a})\hat{f}(\bar{b})$. □

Remarquons que, si l'on avait pas supposé l'anneau commutatif, il aurait fallu considérer un idéal bilatère pour avoir $(a+i)(b+j) = ab + ib + aj + ij \in ab + I$.

On pourra vérifier à titre d'exercice que l'application $I \mapsto s^{-1}(I)$ fournit une correspondance bijective entre les idéaux de A/I et les idéaux de A contenant I. Comme pour les groupes, on peut en déduire de nombreux isomorphismes dont le plus fondamental est $f(A) \cong A/\mathrm{Ker}(f)$.

Pour n'importe quel anneau, on dispose d'un homomorphisme canonique d'anneaux $i_A : \mathbf{Z} \to A$ défini par $m \mapsto m1_A$. Le noyau est de la forme $\mathrm{Ker}(i_A) = m_A\mathbf{Z}$ avec m_A entier ≥ 0. D'après ce qui précède, $i_A(\mathbf{Z})$ est un sous-anneau de A isomorphe à l'anneau $\mathbf{Z}/m_A\mathbf{Z}$. L'entier m_A s'appelle la *caractéristique* de l'anneau A.

On a déjà observé que l'intersection d'idéaux est encore un idéal; on peut définir d'autres opérations sur les idéaux, notamment la *somme* de deux idéaux I, J est définie comme $I + J := \{i + j \mid i \in I \text{ et } j \in J\}$ alors que le *produit* de deux idéaux I, J est défini comme $IJ := \{i_1j_1 + \ldots + i_mj_m \mid i_h \in I \text{ et } j_h \in J\}$. On peut définir d'ailleurs la somme d'idéaux indexés par un ensemble quelconque et le produit d'idéaux indexés par un ensemble fini. On remarque que l'on a toujours $IJ \subset I \cap J$ mais en général on n'a pas égalité; en effet si $I = J = 2\mathbf{Z}$ dans $A = \mathbf{Z}$ alors $IJ = 4\mathbf{Z} \neq 2\mathbf{Z} = I \cap J$. Voici un énoncé classique d'isomorphisme qui est souvent utile.

Proposition. (Lemme chinois généralisé) *Soient I, J deux idéaux de A tels que $I + J = A$ alors $IJ = I \cap J$ et, de plus,*

$$A/IJ \cong A/I \times A/J.$$

Preuve. Considérons l'homomorphisme $f : A \to A/I \times A/J$ produit des deux surjections canoniques. Son noyau est clairement $I \cap J$. Montrons que f est surjective. Pour cela observons que, par hypothèse, il existe $i \in I$ et $j \in J$ tels que $i + j = 1$. Si $a, b \in A$, considérons $c := aj + bi$ on a $c = a(j + i) + i(b - a) \in a + I$ et de même $c = b(i + j) + j(a - b) \in b + J$ donc $f(c) = (s_I(a), s_J(b))$, ce qui prouve bien que f est surjective. On a donc $A/I \cap J \cong A/I \times A/J$ et il reste à voir que $IJ = I \cap J$. On a toujours $IJ \subset I \cap J$; soit donc $a \in I \cap J$, on peut écrire $a = ai + aj$ mais $a \in J$ donc $ai \in IJ$ et $a \in I$ donc $aj \in IJ$ donc $a \in IJ$. □

Remarque. Si $a, b \in \mathbf{Z}$ sont premiers entre eux, on remarque que $a\mathbf{Z} + b\mathbf{Z} = \mathbf{Z}$ et la proposition précédente est bien une généralisation du "lemme chinois" classique.

Terminons ce paragraphe en donnant deux constructions importantes d'anneaux.

(i) Si A est un anneau, on définit l'anneau des polynômes (à une variable) comme l'anneau des suites $P = (a_n)_{n\in\mathbf{N}}$ presque nulles (i.e. telles que $a_n = 0$ pour n assez grand) muni de l'addition et multiplication définies par

$$(a_n)_{n\in\mathbf{N}} + (b_n)_{n\in\mathbf{N}} = (a_n + b_n)_{n\in\mathbf{N}} \quad \text{et} \quad (a_n)_{n\in\mathbf{N}}(b_n)_{n\in\mathbf{N}} = (c_n)_{n\in\mathbf{N}} \text{ avec } c_n = \sum_{k=0}^{n} a_k b_{n-k}$$

On vérifie directement qu'on a bien défini un anneau. Posons $X := (0, 1_A, 0, \ldots, 0, \ldots)$, on montre aisément par récurrence que X^i est la suite dont le seul terme non nul est situé au cran i et vaut 1_A et on récupère ainsi la notation usuelle $P = (a_n)_{n\in\mathbf{N}} = a_0 + a_1X + \ldots + a_dX^d$. On note classiquement $A[X]$ l'anneau des polynômes à coefficients dans A. On peut définir le *degré* d'un polynôme par la formule

$$\deg(P) = \max\{d \in \mathbf{N} \mid a_d \neq 0\}$$

avec la convention que $\deg(0) = -\infty$ (ou n'est pas défini). On a immédiatement les deux formules: $\deg(P + Q) \leq \max\{\deg(P), \deg(Q)\}$ et $\deg(PQ) \leq \deg(P) + \deg(Q)$ avec égalité dans la deuxième formule dès que l'anneau A est intègre (ou plus généralement si le coefficient dominant de P n'est pas un diviseur de zéro). Il n'y a pas de difficulté (sauf peut-être l'empilement de notations) à généraliser cette construction aux polynômes en n variables ou même à une infinité de variables. L'ensemble $A[X_i \mid i \in I]$ est défini comme les "suites" presque nulles d'éléments de A indexées par $\mathbf{N}^{(I)} = \{(n_i)_{i\in I} \in \mathbf{N}^I \mid n_i = 0 \text{ pour presque tout } i\}$ et l'addition et la multiplication de manière analogue. On montre aisément qu'on a des isomorphismes canoniques comme $(A[X])[Y] = (A[Y])[X] = A[X, Y]$, c'est-à-dire que l'on peut voir un polynôme en X, Y à coefficients dans A comme un polynôme en X (resp. en Y) à coefficients dans $A[Y]$ (resp. dans $A[X]$).

(ii) Si S est une partie multiplicative de A (i.e. $1 \in S$ et S est stable par multiplication) on définit l'*anneau des fractions* $S^{-1}A$ ainsi : on définit une relation d'équivalence sur $A \times S$ par

$$(a, s)\mathcal{R}(a', s') \Leftrightarrow \exists s'' \in S,\ s''(as' - a's) = 0.$$

On note $[(a,s)] \in A \times S/\mathcal{R}$ la classe d'un couple $(a,s) \in A \times S$. On définit deux lois sur l'ensemble $S^{-1}A := A \times S/\mathcal{R}$ par

$$[(a,s)] + [(a',s')] = [(as' + a's, ss')] \quad \text{et} \quad [(a,s)] \cdot [(a',s')] = [(aa', ss')]$$

Remarquons que l'introduction de s'' dans la définition de $\mathcal{R}$ est inutile si l'anneau A est intègre (et $0 \notin S$) mais est nécessaire en général pour que $\mathcal{R}$ soit transitive. Par ailleurs, on dispose d'une application naturelle $i : A \to S^{-1}A$ donnée par $a \mapsto [(a, 1_A)]$, c'est un homomorphisme d'anneaux qui permet d'établir une bijection entre d'une part les idéaux propres de S^{-1} et d'autre part les idéaux de A ne rencontrant pas S (un sens de la bijection est donnée par $J \mapsto i^{-1}(J)$).

Une application classique de cette construction est la construction du *corps des fractions* d'un anneau intègre (i.e sans diviseur de zéro). En effet, si A est intègre, on peut choisir $S = A \setminus \{0\}$ comme partie multiplicative et on constate alors que l'anneau $S^{-1}A$ est un corps et que $i : A \to S^{-1}A$ est injective. En effet $[(a, 1_A)] = [(0_A, s)]$ équivaut à $sa = 0_A$ et donc $a = 0_A$; par ailleurs si $a \in A$ et $b \in A \setminus \{0\}$, alors l'élément $[(b,a)]$ est inverse de $[(a,b)]$. Cette construction est l'analogue de la construction de $\mathbf{Q}$ à partir de $\mathbf{Z}$. On note souvent $\mathrm{Frac}(A)$ le corps ainsi construit. Comme autre exemple citons $K(X_1, \ldots, X_n) = \mathrm{Frac}\,(K[X_1, \ldots, X_n])$ le corps des fractions rationnelles en n indéterminées.

L'exemple suivant est important en algèbre commutative. On prend pour S le complémentaire d'un idéal premier P, i.e. $S := A \setminus P$ (la définition d'un idéal premier dit précisément que S est multiplicative). L'anneau $S^{-1}A$ se note alors A_P et jouit une propriété remarquable : il possède un unique idéal maximal, l'idéal formé des éléments $[(a,s)]$ avec $a \in P$ (et $s \in S$); on dit que A_P est un *anneau local*.

Exercice. (Fonction polynôme) Soit $P \in A[X_1, \ldots, X_n]$, on lui associe une fonction $f_P : A^n \to A$ définie par $f_P(x) = P(x)$. L'application $P \mapsto f_P$ est un homomorphisme d'anneaux. Montrer sur un exemple que cette application n'est pas nécessairement injective. Si A est intègre et infini, montrer que l'application $P \mapsto f_P$ est injective (indication : si $n = 1$ montrer qu'un polynôme non nul n'a qu'un nombre fini de racine et procéder par induction sur le nombre de variables). Si $A = \mathbf{Z}/p\mathbf{Z}$ et $n = 1$, montrer que le noyau est engendré par $X^p - X$.

B.2. Divisibilité et idéaux.

On dit que a *divise* b dans l'anneau A s'il existe $c \in A$ tel que $b = ac$; il revient au même de dire que b est un *multiple* de a ou encore que $b \in aA$ ou encore que $bA \subset aA$. Le fait qu'une relation de divisibilité corresponde à une inclusion d'idéaux (principaux) est fondamental dans la suite. Si l'on se place dans un anneau intègre on voit facilement que

$$a \text{ divise } b \quad \text{et} \quad b \text{ divise } a \quad \Leftrightarrow \quad aA = bA \quad \Leftrightarrow \quad \exists u \in A^*,\ b = au.$$

En effet ($\Leftarrow$) est trivial et, si $b = ca$ et $a = c'b$ alors $b = (cc')b$ ou encore $b(1 - cc') = 0$ mais on peut supposer $b \neq 0$ (sinon on a $b = a = 0$ et l'énoncé est banal) et donc, comme A est intègre $cc' = 1$, ce qui signifie bien que $c, c' \in A^*$. On dira que a et b sont *associés* si $b = ua$ avec $u \in A^*$; cette relation est visiblement une relation d'équivalence.

Un élément $a \in A$ est *irréductible* s'il est non nul, non inversible et vérifie la propriété suivante : si $a = bc$ alors b ou c est inversible. On a vu qu'un élément a est premier si l'idéal aA est premier, ou encore si on a l'implication a divise bc entraîne a divise b ou c. Il est clair qu'un élément premier est irréductible (prouvez-le!) mais la réciproque est fausse en général.

On a vu au paragraphe précédent la définition d'un idéal premier; un idéal $I \subset A$ est *maximal* si $I \neq A$ et $I \subset J \subset A$ entraîne $J = I$ ou $J = A$.

Proposition. *Un idéal I est premier si et seulement si A/I est intègre. Un idéal est maximal si et seulement si A/I est un corps.*

Preuve. L'anneau A/I est intègre si et seulement si le produit de deux classes $\bar{a}$ et $\bar{b}$ est nul (i.e. $ab \in I$) dans le seul cas où $\bar{a} = 0$ (i.e. $a \in I$) ou $\bar{b} = 0$ (i.e. $b \in I$), ce qui signifie bien que I est premier. Si A/I

est un corps, ses seuls idéaux sont $\{0\}$ et A/I donc les seuls idéaux de A contenant I sont I et A, ce qui montre bien que I est maximal. Si I est maximal, soit $\bar{a} \in A/I \setminus \{0\}$, alors $a \notin I$ donc $I \neq I + aA \subset A$ donc $A = I + aA$ et il existe $b \in A$ et $i \in I$ tels que $1 = i + ab$ donc $\bar{a}\bar{b} = 1 \in A/I$. Ainsi A/I est bien un corps. □

Par analogie avec les propriétés déjà connues de l'anneau $\mathbf{Z}$ on définit les propriétés suivantes pour un anneau commutatif intègre A.

Définition. Un anneau A est *euclidien* s'il existe une application $\phi : A \setminus \{0\} \to \mathbf{N}$ telle que pour tout $a \in A$, $b \in A \setminus \{0\}$, il existe $q, r \in A$ tels que $a = bq + r$ et ou bien $r = 0$ ou bien $\phi(r) < \phi(b)$.

Définition. Un anneau A est *principal* si tout idéal de A est principal i.e. de la forme aA.

Définition. Un anneau A est *noethérien* si tout idéal est engendré par un nombre fini d'éléments.

Remarque. Cette propriété est équivalente à la propriété souvent donnée comme définition qui dit qu'une suite croissante d'idéaux de A doit être stationnaire. En effet si $I_1 \subset \ldots \subset I_n \subset I_{n+1} \subset \ldots$ est une suite croissante d'idéaux de A noethérien, considérons $I = \cup_{n \geq 1} I_n$, c'est un idéal engendré par disons $a_1, \ldots, a_r$ mais il existe n_0 tel que $a_1, \ldots, a_r \in I_{n_0}$ donc pour tout $n \geq n_0$, on a $I_n = I_{n_0}$. Inversement si toute suite croissante d'idéaux de A est stationnaire et si I est un idéal de A, effectuons la construction suivante. Soit $a_1 \in I$ et $I_1 = a_1 A$, si $I = I_1$ alors I est de type fini, sinon soit $a_2 \in I \setminus I_1$. Posons alors $I_2 = a_1 A + a_2 A$, ou bien $I = I_2$ ou bien il existe $a_3 \in I \setminus I_2$ etc. La construction doit d'arrêter au bout d'un nombre fini d'étapes car $I_1 \subset I_2 \ldots$ est stationnaire.

Définition. Un anneau A est *factoriel* si tout élément non nul et non inversible peut s'écrire comme produit (fini) d'éléments irréductibles et d'une unité et que cette décomposition est unique au sens suivant : si $a = u p_1^{m_1} \ldots p_r^{m_r} = u' q_1^{n_1} \ldots q_s^{n_s}$ avec $u, u' \in A^*$ et les p_i (resp. les q_j) sont irréductibles non associés deux à deux et $m_i, n_j \geq 1$ alors $r = s$ et il existe une permutation $\sigma \in \mathcal{S}_r$ telle que p_i est associé avec $q_{\sigma(i)}$ et $m_i = n_{\sigma(i)}$.

On peut écrire cette définition de manière un peu plus concrète en introduisant $\mathcal{P}$ un ensemble de représentants des éléments irréductibles modulo la relation d'équivalence "être associé". L'anneau A est alors factoriel si pour tout élément non nul, il existe une unité $u \in A^*$ et une unique suite presque nulle d'entier positifs $(m_p(a))_{p \in \mathcal{P}}$ telles que

$$a = u \prod_{p \in \mathcal{P}} p^{m_p(a)}.$$

En général on n'a pas de manière simple de choisir les éléments de $\mathcal{P}$ toutefois dans le cas de $\mathbf{Z}$ on choisit bien sûr l'élément irréductible positif et dans le cas de $K[X]$ on choisit le polynôme irréductible unitaire.

Définition. Un anneau A est *intégralement clos* si pour tout élément $x \in \mathrm{Frac}(A)$ le fait d'être racine d'une équation du type $x^d + a_{d-1} x^{d-1} + \ldots + a_0 = 0$ avec $a_i \in A$ entraîne $x \in A$.

Nous allons étudier les propriétés des anneaux de ce type et en particulier prouver les implications suivantes.

$$\text{Euclidien} \Longrightarrow \quad \begin{array}{c} \text{Principal} \\ \Downarrow \\ \text{Noethérien} \end{array} \quad \Longrightarrow \text{Factoriel} \Longrightarrow \text{Intégralement clos}$$

La notion de divisibilité introduit une notion d'ordre (partiel) sur les idéaux d'un anneau A; il est naturel d'examiner l'existence de majorant, borne supérieure, etc. au sens de cette relation d'ordre. Cette notion est traditionnellement formulée en termes des éléments, bien que, pour être précis il faudrait considérer les classes d'équivalence d'éléments associés.

Définition. Un élément $d \in A$ est *un PGCD* de a et b s'il vérifie les deux propriétés suivantes:
(i) L'élément d divise a et b,
(ii) Si un élément d' divise a et b, alors d' divise d.
Un élément $m \in A$ est *un PPCM* de a et b s'il vérifie les deux propriétés suivantes:
(i) L'élément m est un multiple de a et b,
(ii) Si un élément m' est un multiple de a et b, alors m' est un multiple de m.

Il est clair qu'un PGCD (resp. un PPCM), s'il existe est unique à un élément inversible près, i.e. l'idéal engendré est unique. Les premières propriétés du PGCD et PPCM, quand ils existent sont les suivantes.

Lemme. *Soit A un anneau dans lequel $PGCD$ et $PPCM$ existent, soient $a_1, \dots, a_m, a, b, c \in A$, alors*
(i) $\mathrm{PGCD}(aa_1, \dots, aa_m) = a\,\mathrm{PGCD}(a_1, \dots, a_m)$,
(ii) $\mathrm{PGCD}(a + bc, b) = \mathrm{PGCD}(a, b)$,
(iii) $\mathrm{PPCM}(aa_1, \dots, aa_m) = a\,\mathrm{PPCM}(a_1, \dots, a_m)$.

Preuve. Laissée en exercice. □

On peut traduire ces définitions en terme d'idéaux. En effet a divise b si et seulement si $bA \subset aA$; ainsi d est un diviseur de a et b si et seulement si $a, b \in dA$ ou encore si et seulement si $aA + bA \subset dA$, et m est un multiple de a et b si et seulement si $m \in aA \cap bA$ ou encore si et seulement si $mA \subset aA \cap bA$. On a ainsi prouvé :

Proposition. *Un PGCD de a et b existe dans A si et seulement si il existe un plus petit idéal principal contenant $aA + bA$ et dans ce cas* $\mathrm{PGCD}(a, b)A$ *est cet idéal. Un PPCM de a et b existe dans A si et seulement si il existe un plus grand idéal principal contenu dans $aA \cap bA$ et dans ce cas* $\mathrm{PPCM}(a, b)A$ *est cet idéal.*

Cette proposition rend transparent le prochain théorème.

Théorème. *Soit A un anneau principal, alors le* PGCD *et* PPCM *existent toujours et vérifient*
(i) $\mathrm{PGCD}(a, b)A = aA + bA$.
(ii) $PPCM(a, b) = aA \cap bA$.
De plus $\mathrm{PGCD}(a, b)\,\mathrm{PPCM}(a, b)A = abA$ *et la propriété de Bézout est vérifiée : si d est un* PGCD *de a et b alors*

$$\exists u, v \in A,\ au + bv = d.$$

Preuve. Le seul point qui reste à prouver est que ab et dm sont associés (où d est un PGCD et m un PPCM). Ecrivons $a = da'$ et $b = db'$ alors a' et b' sont premiers entre eux et il existe u, v tels que $a'u + b'v = 1$. L'élément $da'b'$ est un multiple de a et b; inversement si $m' = ac = bc'$ est un multiple de a et b alors $m' = m'(a'u + b'v) = bc'a'u + acb'v = da'b'(c'u + cv)$ est un multiple de $da'b'$ ce qui prouve que ce dernier est un PPCM de a et b. □

Lemme. *Soit A un anneau principal et a irréductible. Si a divise bc alors a divise b ou c. Si a et b sont premiers entre eux et a divise bc alors a divise c.*

Remarque. Le premier énoncé s'appelle le lemme d'Euclide, le second le lemme de Gauss. Le lemme d'Euclide dit qu'un élément irréductible est premier (dans un anneau principal).

Preuve. Si a irréductible divise bc et ne divise pas b, considérons $d = \mathrm{PGCD}(a, b)$. Comme a est irréductible, on a $d = 1$ ou $d = a$ (à un élément inversible près) donc $d = 1$ donc il existe u, v tels que $1 = au + bv$ donc $c = auc + bcv$ est bien divisible par a. Si a et b sont premiers entre eux, alors il existe u, v tels que $1 = au + bv$ et de même $c = auc + bcv$ est bien divisible par a. □

Théorème. *Un anneau principal est noethérien et factoriel.*

Preuve. Tout idéal est de type fini (même engendré par un élément) donc A est noethérien. Montrons maintenant l'existence d'une décomposition en éléments irréductibles dans un anneau noethérien. Supposons qu'il existe $a \in A$ sans décomposition en produit d'éléments irréductibles; comme toute suite croissante d'idéaux est finie, on peut supposer que l'idéal aA est maximal parmi les bA avec b sans décomposition en produit d'éléments irréductibles. L'élément a n'est pas irréductible donc il s'écrit $a = bc$ avec $b, c \notin A^*$. Ainsi $aA \subset bA$ et $aA \subset cA$ (inclusion sans égalité) donc b et c admettent une décomposition en produit d'éléments irréductibles et par conséquent $a = bc$ également. La démonstration du théorème est achevée par la preuve du lemme suivant.

Lemme. *Soit A un anneau tel que tout élément non inversible, non nul, puisse s'écrire comme produit d'irréductibles. Supposons que tout élément irréductible soit premier dans A alors A est factoriel.*

Preuve (du lemme). L'existence d'une décomposition en irréductibles étant acquise, il s'agit d'en prouver l'unicité. Supposons donc que $a = u\prod_{i=1}^{r} p_i^{m_i} = v\prod_{j=1}^{s} q_j^{n_j}$ avec p_i irréductibles non associés deux à deux, $m_i \geq 1$ (idem pour q_j et n_j) et $u, v \in A^*$. Raisonnons par récurrence sur la longueur d'une décomposition de a. Si a admet une décomposition de longueur 1, i.e. a irréductible, alors p_1 divise a donc a et p_1 sont associés et $a = u'p_1$ avec $u' \in A^*$ donc $p_1^{(m_1-1)}\prod_{i=2}^{r} p_i^{m_i} = u'$, ce qui n'est possible que si $r = 1$ et $m_1 = 1$. En général, p_1 divise a donc l'un des q_j; quitte à les renuméroter, on peut supposer que p_1 divise q_1 donc $q_1 = wp_1$ avec $w \in A^*$. En divisant par p_1 on obtient $up_1^{m_1-1}\prod_{i=2}^{r} p_i^{m_i} = vwq_1^{n_1-1}\prod_{j=2}^{s} q_j^{n_j}$ et on peut appliquer l'hypothèse de récurrence pour conclure. □

Remarques. On peut observer que si a est irréductible dans A principal alors A/aA est un corps; en effet l'idéal aA est maximal car si $aA \subset I \subset A$ alors $I = bA$ et donc b divise a donc est soit inversible (auquel cas $I = bA = A$) soit associé à a (auquel cas $I = bA = aA$). Cette propriété n'est plus vraie en général dans les anneaux factoriels. Par exemple dans $K[X, Y]$ qui est factoriel (voir plus loin) l'élément X est irréductible (et premier) mais $K[X, Y]/XK[X, Y] \cong K[Y]$ n'est pas un corps; de même le théorème de Bézout n'est plus vérifié dans cet anneau. Cependant un grand nombre des propriétés des anneaux principaux sont préservées dans le cadre des anneaux factoriels, en particulier:

Proposition. *Dans un anneau factoriel A, le PGCD et PPCM existe toujours et le produit ab est associé au produit* $\mathrm{PGCD}(a, b)\,\mathrm{PPCM}(a, b)A$. *Un élément irréductible est premier et les lemmes d'Euclide et de Gauss restent vrais.*

Preuve. Ecrivons chaque élément $a \in A$ sous la forme $a = u\prod_{p\in\mathcal{P}} p^{m_p(a)}$, où $u \in A^*$ et $\mathcal{P}$ désigne un ensemble de représentants des éléments irréductibles modulo les éléments inversibles. Il est alors clair que

$$\mathrm{PGCD}(a, b) = \prod_{p\in\mathcal{P}} p^{\min\{m_p(a),m_p(b)\}} \quad \text{et} \quad \mathrm{PPCM}(a, b) = \prod_{p\in\mathcal{P}} p^{\max\{m_p(a),m_p(b)\}}.$$

De plus ces formules montrent que le produit du PGCD par le PPCM est associé à ab. Si p irréductible divise ab, on a $ab = pc$ et en écrivant la décomposition en éléments irréductibles de a, b et c et en utilisant l'unicité, on voit que (un élément associé à) p apparaît dans la décomposition de ab donc dans celle de a ou b. Le même raisonnement permet de vérifier le lemme de Gauss. □

Proposition. *Un anneau factoriel est intégralement clos.*

Preuve. Soit $x \in \mathrm{Frac}(A)$ racine d'un polynôme unitaire $P(X) = X^d + a_{d-1}X^{d-1} + \ldots + a_0 \in A[X]$. On peut écrire $x = a/b$ avec $a, b \in A$, de plus, quitte à diviser a et b par leur PGCD, on peut supposer que a et b sont premiers entre eux. Mais alors l'équation $P(x) = 0$ s'écrit aussi $a^d + a_{d-1}ba^{d-1} + \ldots + a_0b^d = 0$ ce qui entraîne b divise a^d. Comme a^d et b sont premiers entre eux, on conclut que b est inversible, ce qui signifie bien que $x \in A$. □

Venons-en a des exemples concrets d'anneaux non factoriels que nous choisirons d'abord dans le cadre de l'arithmétique. Tout d'abord notons que les anneaux du type $A_1 = \mathbf{Z}\left[\sqrt{5}\right]$ ou $A_2 = \mathbf{Z}\left[i\sqrt{3}\right]$ ne sont pas intégralement clos car $(1 + \sqrt{5})/2 \notin A_1$ (bien qu'étant racine de $X^2 - X - 1 = 0$) et $(1 + i\sqrt{3})/2 \notin A_2$ (bien qu'étant racine de $X^2 - X + 1 = 0$) donc A_1 et A_2 ne sont pas factoriels. Cependant l'anneau $\mathbf{Z}\left[i\sqrt{5}\right]$ est intégralement clos (voir ci-dessous) mais n'est pas factoriel, en effet

$$6 = 2 \cdot 3 = (1 + i\sqrt{5})(1 - i\sqrt{5})$$

or chacun des éléments $2, 3, 1 + i\sqrt{5}, 1 - i\sqrt{5}$ est irréductible. Pour voir cela regardons leurs modules au carré que nous appellerons *norme*; si $2 = (a+bi\sqrt{5})(a' + b'i\sqrt{5})$ alors $4 = (a^2+5b^2)(a'^2+5b'^2)$ mais $a^2+5b^2 = 2$ est impossible donc on doit avoir $a^2+5b^2 = 1$ ou $a'^2+5b'^2 = 1$ e qui entraîne $a+bi\sqrt{5} = \pm 1$ ou $a' + b'i\sqrt{5} = \pm 1$. On remarquera que ces éléments sont irréductibles mais ne sont pas premiers.

Il est donc intéressant de regarder les anneaux les plus "complets" possible. Nous ne donnerons pas de définition générale mais observerons simplement que si $d \in \mathbf{Z} \setminus \{0, 1\}$ est sans facteurs carrés, et si nous introduisons

$$A_d := \{\alpha = x + y\sqrt{d} \in \mathbf{Q}(\sqrt{d}) \mid \alpha \text{ est racine d'une équation } a\alpha^2 + b\alpha + c = 0 \text{ avec } a, b, a \in \mathbf{Z}\}$$

alors A_d est forcément intégralement clos si c'est bien un anneau. On peut montrer que c'est bien le cas et que $A_d = \mathbf{Z}[\omega] = \{a + b\omega \mid a, b \in \mathbf{Z}\}$ où $\omega = \left(1 + \sqrt{d}\right)/2$ si $d \equiv 1 \bmod 4$ et $\omega = \sqrt{d}$ sinon. Pour étudier ces anneaux les propriétés suivantes (laissées en exercice) seront utiles:

Lemme. *La norme est multiplicative : $N(\alpha\beta) = N(\alpha)N(\beta)$. Les unités de $\mathbf{Z}[i\sqrt{d}]$ (ou $\mathbf{Z}[\omega]$) sont les éléments de norme 1. Enfin pour $\alpha \neq 0$ on a $N(\alpha) = \mathrm{card}(\mathbf{Z}[\omega]/\alpha\mathbf{Z}[\omega]$.*

Exercice. Soit K un *corps de nombres*, c'est-à-dire un sous-corps de $\mathbf{C}$ qui est de dimension finie sur $\mathbf{Q}$ (comme espace vectoriel). Si $\alpha \in K$ alors la multiplication par α définit un endomorphisme $\mathbf{Q}$-linéaire de K, on définit $N(\alpha)$ comme le déterminant de cet endomorphisme (resp. $Tr(\alpha)$ comme la trace de l'endomorphisme). Montrer que $N(\alpha\beta) = N(\alpha)N(\beta)$ et que N coïncide avec la norme précédemment définie. Montrer que $\alpha \in \mathbf{Q}(\sqrt{d})$ est racine d'une équation $X^2 + aX + b = 0$ avec $a, b \in \mathbf{Z}$ si et seulement si $N(\alpha)$ et $Tr(\alpha)$ sont entiers.

Théorème. *Soit $d \geq 1$, sans facteur carré et $\omega = \left(1 + i\sqrt{d}\right)/2$ si $d \equiv 3 \bmod 4$ et $\omega = i\sqrt{d}$ sinon. L'anneau $\mathbf{Z}[\omega]$ est eucliden si et seulement si $d \in \{1, 2, 3, 7, 11\}$.*

Preuve. Pour prouver que les anneaux cités sont euclidiens, on prouve qu'ils sont euclidiens pour la norme définie sur $\mathbf{Q}(i\sqrt{d})$ par $N(u + vi\sqrt{d}) = u^2 + dv^2$. Pour cela on établit le lemme suivant où l'on suppose que $d \in \{1, 2, 3, 7, 11\}$.

Lemme. *Soit $u + vi\sqrt{d} \in \mathbf{Q}(i\sqrt{d})$, il existe $\alpha \in \mathbf{Z}[\omega]$ tel que $N(u + vi\sqrt{d} - \alpha) < 1$.*

Preuve (du lemme). On utilise qu'un rationnel x (ou même un réel) possède un entier m (resp. un demi-entier $m/2$) tel que $|x - m| \leq 1/2$ (resp. $|x - m/2| \leq 1/4$). Donc, pour $m, n \in \mathbf{Z}$ bien choisis, $N(u + vi\sqrt{d} - (m + ni\sqrt{d})) = (u - m)^2 + d(v - n)^2 \leq (d + 1)/4 < 1$ si $d = 1$ ou 2 et par ailleurs $N(u + vi\sqrt{d} - (m + n\frac{1+i\sqrt{d}}{2})) = (u - m - n/2)^2 + d(v - n/2)^2 \leq 1/4 + d/16 < 1$ si $d = 3, 7$ ou 11. □

On en déduit aisément que ces anneaux sont euclidiens : si $z, z' \in \mathbf{Z}[\omega] \setminus \{0\}$ et si $\alpha \in \mathbf{Z}[\omega]$ est tel que $N(zz'^{-1} - \alpha) < 1$ alors $N(z - \alpha z') < N(z')$ donc en posant $r = z - \alpha z'$ on obtient bien une division euclidienne. Inversement, si $\mathbf{Z}[\omega]$ est euclidien, choisissons $\alpha \in \mathbf{Z}[\omega]$ non inversible et tel que $N(\alpha)$ soit minimale. La division par α donne toujours un reste nul ou inversible et on a donc

$$N(\alpha) = \mathrm{card}(\mathbf{Z}[\omega]/\alpha\mathbf{Z}[\omega]) \leq \mathrm{card}\,(\mathbf{Z}[\omega]^* \cup \{0\})$$

mais on voit aisément que (sauf pour $d = 1$ et 3) $\mathbf{Z}[\omega]^* = \{\pm 1\}$ donc $N(\alpha) \leq 2, 3$. Mais l'équation $a^2 + db^2 = 2$ ou 3 (avec $a, b \in \mathbf{Z}$) n'a pas de solution pour $d > 3$ et l'équation $(a + b/2)^2 + \frac{d}{4}b^2 = a^2 + ab + \frac{d+1}{4}b^2 = 2$ ou 3 (avec $a, b \in \mathbf{Z}$) n'a pas de solution pour $d > 12$ d'où le résultat. □

Citons sans démonstration le théorème suivant (dont la preuve dépasse le niveau de ce cours) :

Théorème. *Soit $d \geq 1$, sans facteur carré et $\omega = \left(1 + i\sqrt{d}\right)/2$ si $d \equiv 3 \bmod 4$ et $\omega = i\sqrt{d}$ sinon. L'anneau $\mathbf{Z}[\omega]$ est principal si et seulement si $d \in \{1, 2, 3, 7, 11, 19, 43, 67, 163\}$.*

Il est très difficile de démontrer que ce sont les seuls anneaux principaux; démontrer que ces anneaux sont principaux est plus élémentaire. Nous le ferons pour le premier non euclidien.

Proposition. *L'anneau $\mathbf{Z}\left[\frac{1+i\sqrt{19}}{2}\right]$ est principal et non euclidien.*

Preuve. Notons $A = \mathbf{Z}\left[\frac{1+i\sqrt{19}}{2}\right]$. Commençons par montrer que $2A$ est un idéal maximal dans A. Le polynôme minimal de $\omega = \frac{1+i\sqrt{19}}{2}$ est $P = X^2 - X + 5$ donc $A \cong \mathbf{Z}[X]/P\mathbf{Z}[X]$ (considérer l'évaluation $\mathbf{Z}[X] \to A$ donnée par $Q \mapsto Q(\omega)$) et $A/2A \cong \mathbf{Z}/2\mathbf{Z}[X]/\bar{P}\mathbf{Z}/2\mathbf{Z}[X]$. Le polynôme $\bar{P} \in \mathbf{Z}/2\mathbf{Z}[X]$ est irréductible (car de degré 2 et sans racine dans $\mathbf{Z}/2\mathbf{Z}$) donc $A/2A$ est un corps et $2A$ est maximal. Ensuite montrons que l'on peut toujours effectuer une division euclidienne (au sens de la norme) soit de a par b, soit de $2a$ par b. Soit $x + iy\sqrt{19} \in \mathbf{Q}(i\sqrt{19})$, il suffit de voir qu'il existe $m, n \in \mathbf{Z}$ tels que $N_1 = N(x + iy\sqrt{19} - m - n\frac{1+i\sqrt{19}}{2}) < 1$ ou $N_2 = N(2(x + iy\sqrt{19}) - m - n\frac{1+i\sqrt{19}}{2}) < 1$. Or ou bien il existe $n \in \mathbf{Z}$

tel que $|y - n/2| \leq 1/6$ et alors on peut choisir $m \in \mathbf{Z}$ tel que $N_1 \leq (1/4) + (19/36) < 1$ ou bien il existe $n \in \mathbf{Z}$ tel que $(n/2) + 1/6 < y < (n+1)/2 - 1/6$ et alors $|2y - (n+1/2)| \leq 1/6$ et alors on peut choisir $m \in \mathbf{Z}$ tel que $N_2 \leq (1/4) + (19/36) < 1$. Soit maintenant I un idéal non nul de A et b un élément non nul de I de norme minimale, nous allons montrer que $I = bA$. On a clairement $bA \subset I$. Inversement soit $a \in I$, si l'on peut effectuer la division euclidienne $a = bq + r$ alors $r \in I$ et $N(r) < N(b)$ entraîne $r = 0$ et $a = bq \in bA$; on peut donc supposer $2a = bq + r$ et donc, pour la même raison $2a = bq$. Comme 2 est premier, ou bien 2 divise q et alors $a \in bA$, ou bien 2 ne divise pas q et donc 2 divise b, i.e. $b = 2b'$. Mais alors comme $2A$ est maximal et $q \notin 2A$ on a $2A + qA = A$ ou encore il existe $u, v \in A$ tels que $2u + qv = 1$. On en tire $b' = 2ub' + qvb' = ub + va \in I$, mais $N(b') = N(b)/4 < N(b)$ ce qui contredit l'hypothèse que $N(b)$ est minimale et achève la démonstration. □

Exercice. Soit $d > 1$ Montrer que l'anneau $\mathbf{Z}[\omega] = \{a + b\omega \mid a, b \in \mathbf{Z}\}$ où $\omega = \left(1 + \sqrt{d}\right)/2$ si $d \equiv 1 \bmod 4$ et $\omega = \sqrt{d}$ sinon, est euclidien, donc principal pour les valeurs $d \in \{2, 3, 5, 6, 7, 11, 13, 14\}$. Montrer que l'anneau $\mathbf{Z}[\omega]$ n'est pas factoriel pour $d = 10$ ou 15.

On ignore si il existe une infinité de valeur $d > 1$ sans facteur carré tels que $\mathbf{Z}[\omega]$ soit principal (on sait que, pour ces anneaux, principal équivaut à factoriel).

Un autre exemple classique d'anneau principal est l'anneau des polynômes à une variable et à coefficients dans un corps. En fait cet anneau est euclidien, ce qui est un corollaire de l'énoncé classique suivant.

Proposition. (division euclidienne de polynômes) *Soit A un anneau et soient P, B deux polynômes de $A[X]$ tels que le coefficient dominant de B soit inversible dans A, alors il existe $Q, R \in A[X]$ tels que*
(i) $P = QB + R$
(ii) $\deg(R) < \deg(B)$.

Preuve. Ecrivons $B = b_d X^d + \ldots + b_0$ avec $b_d \in A^*$ donc $b_d^{-1} \in A$, et notons $P = a_n X^n + \ldots + a_0$. Raisonnons par récurrence sur le degré n de P. Si $n < d$ alors on peut choisir $Q = 0$ et $R = P$. Si $n \geq d$ considérons $P_1 := P - a_n b_d^{-1} X^{n-d} B$, on a clairement $\deg(P_1) < \deg(P)$ donc, par hypothèse de récurrence, il existe Q_1, R_1 tels que $P_1 = Q_1 B + R_1$ et $\deg(R_1) < \deg(B)$. On constate alors que $Q = Q_1 + a_n b_d^{-1} X^{n-d}$ et $R = R_1$ conviennent. □

Remarque. La démonstration fournit un algorithme qui est d'ailleurs l'algorithme usuel de calcul de la division de polynômes.

En particulier si A est un corps, la seule condition pour avoir une division euclidienne est que B soit non nul et l'anneau $A[X]$ est donc euclidien et principal. On peut aisément voir que $A[X]$ est principal seulement lorsque A est un corps. Si A n'est pas intègre, alors $A[X]$ non plus et n'est donc pas principal. Si maintenant A est intègre mais n'est pas un corps, il existe un élément a non nul et non inversible, montrons alors que l'idéal engendré par a et X dans $A[X]$ n'est pas principal. Si P était un générateur, on aurait $X = PQ$ et $a = PQ'$ donc P devrait être une constante inversible et donc $A[X] = aA[X] + XA[X]$. Mais une égalité $1 = aS + XR$ est impossible car, en regardant les coefficients constants, on en déduirait que a est inversible.

Application. L'idéal des polynômes de $K[X_1, \ldots, X_n]$ nuls en $x = (x_1, \ldots, x_n) \in K^n$ est engendré par $X_1 - x_1, \ldots, X_n - x_n$. En effet le résultat est connu pour $n = 1$ et on peut procéder par récurrence : si $P(x) = 0$, on effectue la division de P par $X_n - x_n$ dans l'anneau $A[X_n] = K[X_1, \ldots, X_{n-1}][X_n]$, soit $P = (X_n - x_n)Q + R$ avec $R \in K[X_1, \ldots, X_{n-1}]$. On constate que $R(x_1, \ldots, x_{n-1}) = 0$ donc, d'après l'hypothèse de récurrence, R est dans l'idéal engendré par $X_1 - x_1, \ldots, X_{n-1} - x_{n-1}$.

B.3. Anneaux de polynômes.

L'objet de ce paragraphe est d'étudier les propriétés de l'anneau des polynômes à plusieurs variables et de voir quelles propriétés des anneaux principaux sont conservées.

Théorème. *Soit A un anneau noethérien, alors $A[X_1, \ldots, X_n]$ est encore noethérien.*

Preuve. Il suffit bien sûr de prouver que $A[X]$ est noethérien. Soit I un idéal de $A[X]$, nous allons chercher un ensemble fini de générateurs. Introduisons les ensembles

$$I_n := \{a \in A \mid \exists P \in I,\ P = aX^n + \text{termes de degré} < n\},$$

on voit facilement que ce sont des idéaux de A et que $I_n \subset I_{n+1}$. Comme A est noethérien, il existe n_0 tel que, pour $n \geq n_0$ on ait $I_n = I_{n_0}$ et, par ailleurs, il existe un nombre fini d'éléments $a_1^{(n)}, \ldots, a_{m_n}^{(n)} \in I_n$ qui engendrent I_n. Pour simplifier les notations, on peut supposer que $m_n = m$ pour $n \leq n_0$. Considérons également, pour $n \leq n_0$ des polynômes $P_j^{(n)} = a_j^{(n)} X^n + \ldots \in I$ et montrons que l'ensemble fini $\{P_j^{(n)} \mid n \leq n_0,\ j \leq m\}$ engendre l'idéal I. Notons J l'idéal engendré par ces polynômes, on a $J \subset I$. Considérons donc $P \in I$ et raisonnons par récurrence sur $d = \deg(P)$. Si $d \leq n_0$ alors $P = aX^d + \ldots$ et $a \in I_d$ et donc il existe $b_j \in A$ tels que $a = b_1 a_1^{(d)} + \ldots + b_m a_m^{(d)}$. Le polynôme $P' := P - (b_1 P_1^{(d)} + \ldots + b_m P_m^{(d)})$ est donc dans I et a un degré $< d$, on peut donc supposer par récurrence qu'on sait déjà que P' est dans J et donc $P = P' + b_1 P_1^{(d)} + \ldots + b_m P_m^{(d)}$ est également dans J. Si maintenant $d > n_0$, on sait que $a \in I_{n_0}$, on écrit $a = b_1 a_1^{(n_0)} + \ldots + b_m a_m^{(n_0)}$ et on raisonne de même avec $P' := P - X^{d-n_0}(b_1 P_1^{(n_0)} + \ldots + b_m P_m^{(n_0)})$. □

Lorsque A est factoriel (plus généralement si un PGCD existe toujours dans A) on peut introduire le *contenu* d'un polynôme $P \in A[X]$ qui est, par définition, un PGCD de ses coefficients :

$$\text{si } P = a_0 + a_1 X + \ldots + a_d X^d, \text{ alors } c(P) := \mathrm{PGCD}(a_0, \ldots, a_d)$$

Un polynôme P est dit *primitif* si $c(P) = 1$. On peut toujours factoriser un polynôme comme $P = c(P)P'$ avec P' polynôme primitif du même degré que P.

Lemme. (lemme de Gauss) *Soit A un anneau factoriel, soient P, Q deux polynômes de $A[X]$, alors $c(PQ) = c(P)c(Q)$.*

Preuve. Ecrivons $P = c(P)P'$ et $Q = c(Q)Q'$ avec P', Q' primitifs, alors $PQ = c(P)c(Q)P'Q'$ et donc $c(PQ) = c(P)c(Q)c(P'Q')$ et on voit qu'il suffit de montrer que le produit de deux polynômes primitifs est primitif. Soit donc P, Q primitifs et supposons $c(PQ) \neq 1$ alors il existe p irréductible dans A qui divise $c(PQ)$. Comme A est factoriel, p est premier et $B = A/pA$ est intègre. Considérons l'application $A[X] \to B[X]$ qui, à un polynôme P, associe le polynôme $\bar{P}$ avec les coefficients réduits modulo pA; c'est un homomorphisme d'anneaux. On constate que $\bar{P} \neq 0$ et $\bar{Q} \neq 0$ alors que $\overline{PQ} = 0$ ce qui contredit le fait que $B[X]$ est intègre. □

Lemme. *Soit A un anneau factoriel, soit $K := \mathrm{Frac}(A)$, les éléments irréductibles de $A[X]$ sont, d'une part, les polynômes constants qui sont irréductibles dans A, d'autre part les polynômes de $A[X]$ qui sont primitifs et irréductibles dans $K[X]$.*

Preuve. Il est facile de vérifier que ces éléments sont irréductibles. En effet un polynôme constant ne peut se factoriser qu'en produit de deux polynômes constants, donc un élément a est irréductible dans A si et seulement si il est irréductible dans $A[X]$; si P est primitif et irréductible dans $K[X]$ et si $P = QR$ avec $Q, R \in A[X]$ alors Q ou R est inversible dans $K[X]$ donc constant donc inversible sinon il ne serait pas primitif. Inversement soit P un polynôme non constant. S'il n'est pas primitif, il n'est pas irréductible puisqu'on peut le factoriser $P = c(P)P'$ avec $c(P)$ et P' non inversibles. Si P est primitif et non irréductible dans $K[X]$, montrons qu'il n'est pas irréductible dans $A[X]$. En effet si $P = QR$ avec $Q, R \in K[X]$ et $\deg(Q), \deg(R) \geq 1$, on peut écrire $Q = (a/b)Q'$ et $R = (c/d)R'$ avec $a, b, c, d \in A$ et $Q', R' \in A[X]$ primitifs. On a donc $bdP = acQ'R'$ donc bd et ac diffèrent d'un élément inversible, disons u, donc $P = uQ'R'$ avec $u \in A^*$. □

Théorème. *Soit A un anneau factoriel, alors $A[X_1, \ldots, X_n]$ est encore factoriel.*

Remarque. L'énoncé reste vrai avec une infinité d'indéterminées indépendantes. En effet si I est un ensemble quelconque indexant des indéterminées X_i indépendantes, l'anneau $B := A[X_i \mid i \in I]$ est réunion des anneaux $B_J := A[X_i \mid i \in J]$ où J parcourt les sous-ensembles finis de I. Chacun des B_J est factoriel et

donc B aussi. On remarquera que, si I est infini, l'anneau B n'est pas noethérien, même si A est noethérien ou même est un corps.

Preuve. Il suffit de prouver que si A est factoriel, alors $A[X]$ est factoriel. Introduisons $K := \mathrm{Frac}(A)$ le corps des fractions de A. Montrons d'abord l'existence d'une décomposition en éléments irréductibles. Tout polynôme $P \in A[X]$, se factorise sous la forme $P = Q_1^{m_1} \dots Q_r^{m_r}$ avec Q_i irréductibles dans $K[X]$. Ecrivons comme précédemment $Q_i = (a_i/b_i)Q_i'$ avec $a_i, b_i \in A$ et $Q_i' \in A[X]$ primitifs. On obtient $b_1^{m_1} \dots b_r^{m_r} P = a_1^{m_1} \dots a_r^{m_r} Q_1'^{m_1} \dots Q_r'^{m_r}$ et, en observant qu'il existe $c \in A$ tel que $a_1^{m_1} \dots a_r^{m_r} = c b_1^{m_1} \dots b_r^{m_r}$ on conclut que $P = cQ_1'^{m_1} \dots Q_r'^{m_r}$. En écrivant c comme produit d'irréductibles de A on obtient la décomposition cherchée. Montrons maintenant l'unicité de la décomposition en éléments irréductibles. Soit donc P se décomposant en

$$P = a_1^{\ell_1} \dots a_t^{\ell_t} Q_1^{m_1} \dots Q_r^{m_r} = b_1^{h_1} \dots b_u^{h_u} R_1^{n_1} \dots R_s^{n_s}$$

avec a_i, b_j irréductibles dans A et Q_i, R_j non constants et irréductibles dans $A[X]$, donc primitifs et irréductibles dans $K[X]$. En utilisant l'unicité de la décomposition dans $K[X]$, on voit que $r = s$ et que, quitte à permuter les indices, $Q_i = \lambda_i R_i$ avec $\lambda_i \in K^*$ et $m_i = n_i$. Mais si l'on écrit $\lambda_i = \alpha_i/\beta_i$ avec $\alpha_i, \beta_i \in A$ on voit que $\beta_i Q_i = \alpha_i R_i \in A[X]$ donc, comme Q_i et R_i sont primitifs, β_i et α_i sont associés et λ est une unité de A. On en déduit que Q_i et R_i sont associés et que $\lambda a_1^{\ell_1} \dots a_t^{\ell_t} = b_1^{h_1} \dots b_u^{h_u}$ avec $\lambda \in A^*$. L'unicité de la décomposition dans A permet alors de conclure. □

B.4. Ensembles algébriques et idéaux de $K[X_1, \ldots, X_n]$

On désigne par K un corps quelconque. A un moment donné on supposera K *algébriquement clos* i.e. que tout polynôme non constant à coefficients dans K possède une racine dans K (par exemple $\mathbf{C}$ est algébriquement clos). Le point de départ de la géométrie algébrique est l'étude des ensembles de zéros communs d'une famille de polynômes.

Définition. Un *sous-ensemble algébrique* de K^n est un ensemble du type

$$Z = \{x \in K^n \mid \forall P \in S,\ P(x) = 0\}$$

où S est un sous-ensemble de $K[X_1, \ldots, X_n]$.

Il convient de remarquer tout de suite que si I est l'idéal engendré par S dans $K[X_1, \ldots, X_n]$ alors les zéros communs des polynômes de S sont les mêmes que les zéros communs des polynômes de I ou encore que les zéros communs de générateurs de I. On voit en particulier qu'un ensemble algébrique peut toujours être défini par un nombre fini de polynômes (théorème de Hilbert). On peut donc associer à un idéal I un ensemble algébrique que l'on notera $\mathcal{V}(I) = \{x \in K^n \mid \forall P \in I,\ P(x) = 0\}$.

Lemme. *On a les propriétés suivantes*
(i) *Si $I \subset J$, alors $\mathcal{V}(I) \supset \mathcal{V}(J)$*
(ii) $\mathcal{V}(I) \cup \mathcal{V}(J) = \mathcal{V}(IJ)$
(iii) $\cap_{t\in T} \mathcal{V}(I_t) = \mathcal{V}\left(\sum_{t\in T} I_t\right)$
(iv) $\mathcal{V}(\{0\}) = K^n$ *et* $\mathcal{V}(K[X_1, \ldots, X_n]) = \emptyset$.

La preuve est laissée en exercice.

Remarque. Les propriétés (ii), (iii) et (iv) peuvent être interprétées comme le fait que les sous-ensembles algébriques sont les fermés d'une topologie qu'on appelle *topologie de Zariski*.

Inversement, on peut associer à tout sous-ensemble Z de K^n un idéal $\mathcal{I}(Z)$ défini par :

$$\mathcal{I}(Z) := \{P \in K[X_1, \ldots, X_n] \mid \forall x \in Z,\ P(x)\}.$$

On a alors les propriétés suivantes.

Lemme. *Soient Z, Z' des sous-ensembles de K^n, soient I, J des idéaux de $K[X_1, \ldots, X_n]$.*
(i) *Si $Z \subset Z'$, alors $\mathcal{I}(Z) \supset \mathcal{I}(Z')$*
(ii) *$Z \subset \mathcal{V}(\mathcal{I}(Z))$ avec égalité si et seulement si Z est un sous-ensemble algébrique.*
(iii) $I \subset \mathcal{I}(\mathcal{V}(I))$.

Preuve. Le point (i) est facile. Si $P \in I$ alors pour tout $x \in \mathcal{V}(I)$ on a $P(x) = 0$ donc $P \in \mathcal{I}(\mathcal{V}(I))$ et (iii) est établi. Soit $z \in Z$ alors pour tout $P \in \mathcal{I}(Z)$ on a $P(x) = 0$ donc $z \in \mathcal{V}(\mathcal{I}(Z)$ et on a bien $Z \subset \mathcal{V}(\mathcal{I}(Z))$. Supposons de plus que $Z = \mathcal{V}(I)$ alors $\mathcal{I}(Z) = \mathcal{I}(\mathcal{V}(I)) \supset I$ donc $\mathcal{V}(\mathcal{I}(Z)) \subset \mathcal{V}(I) = Z$. □

Remarque. Il y a deux raisons simples qui font qu'en général l'inclusion $I \subset \mathcal{I}(\mathcal{V}(I))$ n'est pas une égalité.
(a) Si le corps K n'est pas algébriquement clos, il existe $P \in K[X]$, polynôme non constant et sans racine dans K; considérons donc $I = PK[X]$. Par hypothèse $\mathcal{V}(I) = \emptyset$ et donc $\mathcal{I}(\mathcal{V}(I)) = K[X] \neq I$.
(b) Même si K est algébriquement clos il y un obstacle dû au fait que "$P = 0$" et "$P^m = 0$" définissent les mêmes ensembles algébriques. Concrètement si P est un polynôme irréductible de $K[X_1, \ldots, X_n]$ tel que, si $I_1 = PK[X_1, \ldots, X_n]$, on a $\mathcal{I}(\mathcal{V}(I_1)) = I_1$, posons $I_m = P^m K[X_1, \ldots, X_n]$. On constate alors que $\mathcal{I}(\mathcal{V}(I_m)) = I_1 \neq I_m$.

On peut remédier à l'obstacle (a) en remplaçant K par sa clôture algébrique. On peut remédier à l'obstacle (b) en remplaçant I par son radical :

Définition. Le *radical* d'un idéal I dans un anneau (commutatif) A est l'ensemble des éléments dont une puissance est dans I; on le note $\sqrt{I}$. En symbole :

$$\sqrt{I} := \{a \in A \mid \exists m \geq 1,\ a^m \in I\}.$$

Remarque. On a clairement, pour tout idéal I légalité $\mathcal{V}(I) = \mathcal{V}(\sqrt{I})$ et les inclusions $I \subset \sqrt{I} \subset \mathcal{I}(\mathcal{V}(I))$. On a visiblement $\sqrt{\sqrt{I}} = \sqrt{I}$; cela suggère de considérer les idéaux I *réduits*, i.e. tels que $\sqrt{I} = I$. On va montrer que, lorsque K est algébriquement clos, les correspondances $I \mapsto \mathcal{V}(I)$ et $Z \mapsto \mathcal{I}(Z)$ induisent deux bijections réciproques l'une de l'autre entre idéaux réduits de $K[X_1, \ldots, X_n]$ d'une part et sous-ensembles algébriques de K^n d'autre part. Le point clef est le célèbre résultat:

Théorème. (Nullstellensatz ou Théorème des zéros de Hilbert) *Soit K un corps algébriquement clos. Soient $P_1, \ldots, P_m, Q \in K[X_1, \ldots, X_n]$ tels que, pour tout $x \in K^n$, on ait $P_1(x) = \ldots = P_m(x) = 0$ implique $Q(x) = 0$, alors il existe $t \geq 1$ et $A_1, \ldots, A_m \in K[X_1, \ldots, X_n]$ tels que*

$$Q^t = A_1 P_1 + \ldots + A_m P_m.$$

Ce résultat peut se traduire en le fait que, pour tout idéal I, on a $\mathcal{I}(\mathcal{V}(I)) = \sqrt{I}$ et en particulier, si I est réduit, $\mathcal{I}(\mathcal{V}(I)) = I$.

Preuve. Nous allons donner la preuve pour $K = \mathbf{C}$ en indiquant ce qui resterait à démontrer dans le cas général. Tout d'abord nous allons déduire le théorème du résultat apparemment plus faible.

Proposition. (Nullstellensatz "faible") *Soit K un corps algébriquement clos. Soient $P_1, \ldots, P_m$ des polynômes de $K[X_1, \ldots, X_n]$ sans zéros communs dans K^n, alors il existe $A_1, \ldots, A_m \in K[X_1, \ldots, X_n]$ tels que*

$$1 = A_1 P_1 + \ldots + A_m P_m.$$

Preuve (que la version "faible" entraîne le théorème). On introduit pour la preuve une indéterminée supplémentaire T et on observe que les polynômes $P_1, \ldots, P_m, 1 - TQ$ n'ont aucun zéro commun dans K^{n+1} et donc il existe $A_1, \ldots, A_m, B \in K[X_1, \ldots, X_n, T]$ tels que

$$1 = A_1(X,T)P_1(X) + \ldots + A_m(X,T)P_m(X) + B(X,T)(1 - TQ(X).$$

On note $L := K(X_1, \ldots, X_n)$ le corps des fractions de $K[X_1, \ldots, X_n]$ et on regarde l'identité précédente comme une égalité de polynômes dans $L[T]$. On peut "bien entendu" remplacer T par $1/Q \in L$ et conserver l'égalité. Si $t = \max_i \deg_T A_i(X,T)$ on voit que $Q(X)^t A_i(X, 1/Q(X)) = A_i'(X) \in K[X_1, \ldots, X_n]$ et on obtient

$$Q^t = A_1' P_1 + \ldots + A_m' P_m.$$

Montrons maintenant que la version "faible" se déduit de l'énoncé suivant

Proposition. *Soit K un corps et L une K-algèbre de type fini (i.e. il existe $x_1, \ldots x_n \in L$ tels que $L = K[x_1, \ldots, x_n]$) qui est également un corps, alors L est une extension algébrique finie de K. En particulier, si K est algébriquement clos, $L = K$.*

Preuve (que la proposition entraîne le Nullstellensatz faible). Supposons donc qu'il existe I un idéal non trivial tel que $\mathcal{V}(I) = \emptyset$. Quitte à remplacer I par un idéal maximal le contenant, on peut supposer I maximal. Mais alors $L = K[X_1, \ldots, X_n]/I$ est un corps et est une K-algèbre de type fini donc d'après la proposition c'est K. Ainsi on peut considérer les éléments $x_i :=$ la classe de X_i modulo I comme des éléments de K. Soit $P \in I$, on calcule $P(x_1, \ldots, x_n) = P(\bar{X}_1, \ldots, \bar{X}_n) = \bar{P} = 0$. Le point $x = (x_1, \ldots, x_n)$ est donc un zéro commun, ce qui contredit $\mathcal{V}(I) = \emptyset$. Remarquons au passage que l'on a prouvé :

Corollaire. *Soit K algébriquement clos, les idéaux maximaux de $A = K[X_1, \ldots, X_n]$ sont les idéaux de la forme*

$$I_x = (X_1 - x_1)A + \ldots + (X_n - x_n)A = \{P \in K[X_1, \ldots, X_n] \mid P(x) = 0\}.$$

En effet on a clairement $K[X_1, \ldots, X_n]/I_x \cong K$ donc I_x est maximal et on a montré précédemment que tout idéal non trivial était contenu dans un I_x.

Terminons par une preuve de la proposition dans le cas où $K = \mathbf{C}$ (ou plus généralement le cas où K n'est pas dénombrable). Les monômes $x_1^{i_1} \ldots x_n^{i_n}$ forme une partie génératrice dénombrable de L comme K-espace

vectoriel. Pour chaque x_i, ou bien x_i est algébrique sur K ou bien il est transcendant sur K. Mais, dans le cas où l'un des x_i serait transcendant, on aurait $K[T] \cong K[x_i] \subset L$ donc $K(T) \cong K(x_i) \subset L$ (puisque L est un corps). Or la théorie de la décomposition en éléments simples des fractions rationnelles nous apprend que les fractions

$$\left\{ \frac{1}{T-a} \mid a \in K \right\}$$

sont K-linéairement indépendantes. Si $\mathrm{card}(K) > \mathrm{card}(\mathbf{N})$ ceci entraîne une contradiction et termine la preuve. □

Terminons par quelques remarques concernant la topologie de Zariski. Tout d'abord chaque sous-ensemble algébrique de K^n est muni d'une topologie induite par celle de Zariski. La correspondance entre idéaux et sous-ensembles algébriques, plus le fait que $K[X_1, \ldots, X_n]$ est noethérien entraîne l'assertion suivante.

Proposition. *Soit $\ldots \subset Z_{n+1} \subset Z_n \subset \ldots \subset Z_1 \subset K^n$ une suite décroissante de sous-ensemble algébriques, alors cette suite est stationnaire.*

Preuve. En effet $Z_i = \mathcal{V}(\mathcal{I}(Z_i))$ et $\mathcal{I}(Z_n) \subset \mathcal{I}(Z_{n+1})$ donc la suite des $\mathcal{I}(Z_n)$ est stationnaire. □

Remarquons que cette propriété peut s'interpréter comme une propriété de compacité (sans la propriété d'être séparé) : d'une intersection vide de fermés, on peut extraire une intersection finie qui est encore vide. Par ailleurs, si K est fini, la topologie de Zariski est la topologie discrète sur K^n. On supposera donc K infini pour la suite. On peut étudier les notions classiques (connexité, compacité, etc.); en fait la notion suivante est plus naturelle dans ce contexte:

Définition. Un espace topologique Z est *irréductible* s'il n'est pas réunion de deux fermés non triviaux, c'est-à-dire que $Z = F_1 \cup F_2$ avec F_1, F_2 fermés entraîne $F_1 = Z$ ou $F_2 = Z$.

Remarque. Il revient au même de demander que tous les ouverts non vides soient denses dans Z (considérer $O_i = Z \setminus F_i$). On voit donc qu'un espace irréductible n'est jamais séparé (sauf s'il est réduit à un point).

Exemple. Lorsque K est infini, l'espace K^n, muni de la topologie de Zariski est irréductible. Il suffit, pour vérifier cela, de montrer qu'un polynôme $P \in K[X_1, \ldots, X_n]$ s'annulant sur le complémentaire des zéros d'un polynôme Q non nul est en fait identiquement nul. Mais, dans ce cas, le polynôme PQ s'annule sur K^n tout entier et est donc nul (ici l'on utilise l'hypothèse K infini) donc, comme Q n'est pas nul, on en tire bien $P = 0$.

Proposition. *Un ensemble algébrique $Z \subset K^n$ est irréductible si et seulement si l'idéal $\mathcal{I}(Z)$ est premier.*

Preuve. Supposons $Z = Z_1 \cup Z_2$ avec $Z_1 \not\subset Z_2$ et $Z_2 \not\subset Z_1$ et posons $I_i = \mathcal{I}(Z_i)$ pour $i = 1, 2$. On a donc $I_2 \not\subset I_1$ et $I_1 \not\subset I_2$ et on peut choisir $P_1 \in I_1 \setminus I_2$ et $P_2 \in I_2 \setminus I_1$ de sorte que $P_1 P_2$ s'annule sur Z donc $P_1 P_2 \in \mathcal{I}(Z)$ mais P_1 (resp. P_2) ne s'annule pas sur tout Z_2 (resp. sur tout Z_1) et donc $P_1 \notin \mathcal{I}(Z)$ (resp. $P_2 \notin \mathcal{I}(Z)$), ce qui montre que $\mathcal{I}(Z)$ n'est pas premier. Supposons maintenant que $\mathcal{I}(Z)$ ne soit pas premier et soit $P_1, P_2 \notin \mathcal{I}(Z)$ tels que $P_1 P_2 \in \mathcal{I}(Z)$. Posons $Z_i := \{x \in Z \mid P_i(x) = 0\}$ pour $i = 1, 2$. On a clairement Z_i fermé et $Z = Z_1 \cup Z_2$. Si on avait disons $Z_1 \subset Z_2$ alors P_2 s'annulant sur Z_2 s'annulerait sur Z et on aurait $P_2 \in \mathcal{I}(Z)$, ce qui est une contradiction. □

Tous les ensembles algébriques ne sont pas irréductibles. Par exemple si $Z \subset K^2$ est défini par $xy = 0$ on voit immédiatement que Z est réunion de deux fermés – les droites définies par $x = 0$ et $y = 0$ – qui sont irréductibles. Ce phénomène est général.

Proposition. *Soit Z un sous-ensemble algébrique de K^n, alors Z est réunion finie de sous-ensembles algébriques irréductibles $Z = Z_1 \cup \ldots \cup Z_m$. Si de plus on impose que $Z_i \not\subset Z_j$ pour $i \neq j$, alors les Z_i sont uniques.*

Preuve. Le fait que Z soit réunion finie de sous-ensembles irréductibles est immédiat à partir du caractère noethérien : si ce n'était pas le cas on pourrait écrire une suite infinie strictement décroissante de sous-ensembles algébriques. En effet si Z n'est pas irréductible $Z = Z_1 \cup Z_2$ et si, disons, Z_2 n'est pas irréductible on continue la décomposition. Il est clair que l'on peut construire (en éliminant les composantes "inutiles")

une décomposition $Z = Z_1 \cup \ldots \cup Z_m$ avec $Z_i \not\subset Z_j$ pour $i \neq j$. Si $Z = Y_1 \cup \ldots \cup Y_n$ est une autre décomposition en irréductibles, Observons que $Y_1 = (Z_1 \cap Y_1) \cup \ldots \cup (Z_m \cap Y_1)$ donc, comme Y_1 est irréductible il existe i tel que $Z_i \cup Y_1 = Y_1$ c'est-à-dire $Y_1 \subset Z_i$. Par symétrie Z_i doit être contenu dans un des Y_j donc dans Y_1 (sinon on aurait $Y_1 \subset Y_j$). On conclut que Y_1 est égal à Z_i. $\square$

Les sous-ensembles irréductibles de la proposition s'appelle les *composantes irréductibles* de Z.

C. CORPS.

C.1. Généralités et exemples.

On supposera ici les corps commutatifs. Pour les corps finis, cette hypothèse n'est pas nécessaire (voir appendice à ce chapitre). Il existe des corps non commutatifs, le plus célèbre est le corps des quaternions, il est étudié dans un chapitre spécial. Nous connaissons déjà un certain nombre de corps commutatif : $\mathbf{Z}/p\mathbf{Z}$, $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{C}$, si K est un de ces corps $K(X_1, \ldots, X_n)$ est encore un corps. Nous allons en construire d'autres.

Commençons par déterminer la caractéristique d'un corps K. L'homomorphisme $i_A : \mathbf{Z} \to K$ a une image qui est un sous-anneau intègre de K donc $\mathrm{Ker}(i_A)$ est un idéal premier. Ainsi soit $\mathrm{Ker}(i_A) = \{0\}$ et i_A est injectif et $\mathrm{car}(K) = 0$, soit il existe un nombre premier p tel que $\mathrm{Ker}(i_A) = p\mathbf{Z}$ et alors $\mathrm{car}(K) = p$. Dans le premier cas K contient un sous-anneau isomorphe à $\mathbf{Z}$ donc contient un sous-corps isomorphe à $\mathbf{Q}$, dans le second cas K contient un sous-corps isomorphe à $\mathbf{Z}/p\mathbf{Z}$. En caractéristique p le phénomène le plus remarquable est le suivant:

Lemme. *Soit K un corps de caractéristique p, alors l'application $\phi : K \to K$ définie par $\phi(x) = x^p$ est un homomorphisme de corps.*

Preuve. On a toujours $(xy)^p = x^p y^p$ (puisque l'on ne considère que les corps commutatifs); il suffit donc de prouver que, lorsque $\mathrm{car}(K) = p$, on a $(x + y)^p = x^p + y^p$. Ceci est en fait immédiat si l'on utilise la formule du binôme de Newton et l'observation que les coefficients binomiaux C_p^r sont divisibles par p pour $1 \leq r \leq p - 1$. □

Remarque. Le lemme ne dit pas que ϕ est un isomorphisme, et d'ailleurs il n'est pas en général surjectif (prendre par exemple $K = (\mathbf{Z}/p\mathbf{Z})(X)$); par contre ϕ est toujours injectif, comme le montre un lemme ci-dessous, et définit donc un isomorphisme avec un sous-corps de K que l'on note souvent K^p. Dans le cas $K = (\mathbf{Z}/p\mathbf{Z})(X)$ on voit aisément que $K^p = (\mathbf{Z}/p\mathbf{Z})(X^p) \neq (\mathbf{Z}/p\mathbf{Z})(X)$.

Un autre phénomène spécifique à la caractéristique p est la possibilité pour un polynôme d'avoir une dérivée identiquement nulle sans être constant. En effet si $\mathrm{car}(K) = p$ et si $P \in K[X]$ est non constant, posons $Q(X) := P(X^p)$ alors $Q'(X) \equiv 0$. On définit ici bien sûr formellement la dérivée de $P = a_n X^n + \ldots + a_0$ par $P' = n a_n X^{n-1} + (n-1) a_{n-1} X^{n-2} + \ldots + a_1$. Montrons que la dérivée permet néanmoins de caractériser les racines simples d'un polynôme même en caractéristique p.

Lemme. *Soit K un corps, $\alpha \in K$ et $P \in K[X]$. Alors $(X - \alpha)$ divise P si et seulement si $P(\alpha) = 0$; de plus $(X - \alpha)^2$ divise P si et seulement si $P(\alpha) = P'(\alpha) = 0$.*

Preuve. On écrit d'abord la division euclidienne $P = (X - \alpha)Q + R$ avec $\deg(R) < 1$ donc R est constant et $P(\alpha) = R$ d'où le premier énoncé. On écrit ensuite la division euclidienne $P = (X - \alpha)^2 Q + R$ avec $\deg(R) \leq 1$ donc $R(X) = aX + b$. On a donc $P'(\alpha) = R'(\alpha) = a$ donc $P'(\alpha) = 0$ entraîne $a = 0$ et alors $P(\alpha) = b = 0$. □

Lemme. *Soit $f : K \to L$ un homomorphisme de corps, alors f est injectif.*

Preuve. Par définition $f(1_K) = 1_L$ et par conséquent, si $x \in K \setminus \{0\}$ on en tire $1_L = f(xx^{-1}) = f(x)f(x^{-1})$ donc $f(x) \neq 0$. □

Lorsque $f : K \to L$ est un homomorphisme de corps, on peut identifier K avec un sous-corps de L; on peut aussi considérer L comme un K-espace vectoriel en introduisant l'application:

$$\begin{array}{ccc} K \times L & \to & L \\ (x, y) & \mapsto & f(x)y \end{array}$$

Dans ce contexte on notera $[L : K] = \dim_K L$ la dimension de L vu comme K-espace vectoriel. La notation est en bonne partie motivée par la propriété importante suivante.

Proposition. *Soit $K \subset L \subset F$ une tour de corps, alors $[F : K] = [F : L][L : K]$.*

Preuve. Nous donnons la preuve lorsque ces dimensions sont finies, en fait l'énoncé et même la preuve restent valables avec des cardinaux quelconques. Considérons $e_1, \ldots, e_m$ une base de L sur K et $f_1, \ldots, f_n$ une base de F sur L, nous allons montrer que $\{e_i f_j \mid 0 \leq i \leq m,\ 0 \leq j \leq n\}$ fournit une base de L sur K. Montrons d'abord que c'est une partie génératrice. Soit $x \in F$, alors il existe $\lambda_i \in L$ tels que $x = \sum_{i=1}^{n} \lambda_i f_i$ (car les f_j forment une L-base de F). Par ailleurs il existe $\alpha_{ij} \in K$ tels que $\lambda_i = \sum_{j=1}^{m} \alpha_{ij} e_j$ (car les e_j forment une K-base de L) et donc $x = \sum_{i,j} \alpha_{ij} e_j f_i$. Montrons maintenant l'indépendance linéaire. Si $\alpha_{ij} \in K$ et $\sum_{i,j} \alpha_{ij} e_j f_i = 0$ alors $\sum_i \left(\sum_j \alpha_{ij} e_j \right) f_i = 0$ donc $\sum_j \alpha_{ij} e_j = 0$ (puisque les f_i sont L-linéairement indépendants) et donc les α_{ij} sont nuls (puisque les e_j sont K-linéairement indépendants). □

Un corollaire "évident" est que, si $K \subset L \subset F$ alors $[L : K] \leq [F : K]$; de plus, si ces dimensions sont finies, on a $[L : K] = [F : K]$ si et seulement si $F = L$.

Terminons ce paragraphe en citant sans détail d'autres exemples de corps.

(i) Soit p premier, considérons

$$\mathbf{Z}_p := \left\{ (a_n)_{n \geq 1} \in \prod_{n \geq 1} \mathbf{Z}/p^n\mathbf{Z} \mid a_{n+1} \equiv a_n \bmod p^n \right\}.$$

C'est un anneau intègre, appelé l'anneau des *entiers p-adiques*, son corps des fractions $\mathbf{Q}_p$ appelé le corps des *nombres p-adiques*. On peut montrer que $\mathbf{Q}_p$ est un analogue de $\mathbf{R}$ au sens qu'il est la complétion de $\mathbf{Q}$ pour la valeur absolue $|x|_p := p^{-\mathrm{ord}_p(x)}$.

(ii) Soit U un ouvert connexe du plan complexe, alors l'ensemble $\mathcal{M}(U)$ des fonctions méromorphes sur U est un corps.

(i) Soit K un corps, l'ensemble des *séries formelles* $\sum_{n=0}^{\infty} a_n X^n$ peut être muni d'une structure d'anneau noté $K[\![X]\!]$. En rendant inversible X, on obtient un corps appelé *corps des séries formelles* et noté $K(\!(X)\!)$. On peut aussi le voir comme l'ensemble des séries $\sum_{n \geq -n_0}^{\infty} a_n X^n$.

C.2. Eléments algébriques et transcendants.

Soit $K \subset L$ une extension de corps et $\alpha \in L$. Considérons l'homomorphisme d'anneaux "évaluation en α" définie de la manière suivante:

$$\begin{array}{rccc} ev_\alpha : & K[X] & \to & L \\ & P & \mapsto & P(\alpha) \end{array}$$

Lorsque $\mathrm{Ker}(ev_\alpha) = \{0\}$, on dit que α est *transcendant* sur K. Lorsque $\mathrm{Ker}(ev_\alpha) \neq \{0\}$, on dit que α est *algébrique* sur K. Si $\mathrm{Ker}(ev_\alpha) = PK[X]$, on appellera P le *polynôme minimal* de α sur K (il n'est tout-à-fait unique que si on lui impose d'être unitaire).

Notons $K[\alpha]$ le plus petit sous-anneau de L contenant K et α et $K(\alpha)$ le plus petit sous-corps de L contenant K et α. Par construction $K[\alpha]$ est l'image de ev_α donc est isomorphe à $K[X]/\mathrm{Ker}(ev_\alpha)$. Si α est trancendant, on voit que $K[\alpha] \cong K[X]$ et $K(\alpha) \cong K(X)$; en particulier $K(\alpha)$ est de dimension infinie sur K. Si α est algébrique et P son polynôme minimal sur K, alors P est irréductible dans $K[X]$ donc l'idéal engendré par P est maximal et $K[\alpha] = K(\alpha) \cong K[X]/PK[X]$. De plus dans ce cas on a $[K(\alpha) : K] = \deg(P)$. En effet une base de $K[\alpha] = K(\alpha)$ sur K est donnée par $1, \alpha, \alpha^2, \ldots, \alpha^{\deg(P)-1}$. On a en particulier prouvé:

Proposition. *Soit $\alpha \in L \supset K$ alors α est algébrique sur K si et seulement si $[K(\alpha) : K] < \infty$. Dans ce cas $[K(\alpha) : K]$ est le degré du polynôme minimal de α sur K.*

Remarque. On peut en déduire que si $K \subset F \subset L$ alors $[F(\alpha) : F] \leq [K(\alpha) : K]$. En effet le membre de gauche est le degré du polynôme minimal de α sur F qui divise le polynôme minimal de α sur K dont le degré est le membre de droite.

Corollaire. *Soit $\alpha, \beta \in L \supset K$ et supposons α, β algébriques sur K alors $\alpha + \beta$, $\alpha\beta$ et α/β sont algébriques sur K.*

Preuve. Il suffit de montrer que $[K(\alpha,\beta):K]<\infty$. En effet on aura alors, pour tout élément $x\in K(\alpha,\beta)$ l'inégalité $[K(x):K]\leq[K(\alpha,\beta):K]<\infty$ et donc x algébrique sur K. Mais par ailleurs on a

$$[K(\alpha,\beta):K]=[K(\alpha,\beta):K(\alpha)][K(\alpha):K]\leq[K(\beta):K][K(\alpha):K]<\infty$$

ce qui achève la démonstration. □

Exemple. Soit $\delta=\sqrt[5]{2}+\sqrt[7]{3}+\sqrt[2]{5}$ alors δ est algébrique sur $\mathbf{Q}$. Illustrons les méthodes précédentes en montrant que $[\mathbf{Q}(\delta):\mathbf{Q}]=70$ donc son polynôme minimal est de degré 70 et serait fastidieux à écrire. Notons pour abréger $\alpha=\sqrt[5]{2}$, $\beta=\sqrt[7]{3}$ et $\gamma=\sqrt[2]{5}$. Alors le polynôme minimal sur $\mathbf{Q}$ de α (resp. β, resp. γ) est X^5-2 (resp. X^7-3, resp. X^2-5) donc $[\mathbf{Q}(\alpha):\mathbf{Q}]=5$ (resp. $[\mathbf{Q}(\beta):\mathbf{Q}]=7$, resp. $[\mathbf{Q}(\gamma):\mathbf{Q}]=2$). On a

$$[\mathbf{Q}(\delta):\mathbf{Q}]\leq[\mathbf{Q}(\alpha,\beta,\gamma):\mathbf{Q}]\leq[\mathbf{Q}(\alpha):\mathbf{Q}][\mathbf{Q}(\beta):\mathbf{Q}][\mathbf{Q}(\gamma):\mathbf{Q}]=5\cdot7\cdot2=70.$$

Mais $[\mathbf{Q}(\alpha,\beta,\gamma):\mathbf{Q}]=[\mathbf{Q}(\alpha,\beta,\gamma):\mathbf{Q}(\alpha)][\mathbf{Q}(\alpha):\mathbf{Q}]$ donc 5 (resp. 7, resp. 2) divise $[\mathbf{Q}(\alpha,\beta,\gamma):\mathbf{Q}]$, donc 70 également d'où $[\mathbf{Q}(\alpha,\beta,\gamma):\mathbf{Q}]=70$. Enfin on laisse en exercice de vérifier que $\mathbf{Q}(\delta)=\mathbf{Q}(\alpha,\beta,\gamma)$ et donc le polynome minimal de δ est de degré 70. On pourra procéder ainsi: a) Vérifier que $\mathbf{Q}(\alpha+\beta)=\mathbf{Q}(\alpha,\beta)$ etc. b) Montrer que $\mathbf{Q}(\delta,\gamma)=\mathbf{Q}(\alpha,\beta,\gamma)$. c) Montrer que γ ne peut être de degré 2 sur $\mathbf{Q}(\delta)$ car sinon α serait aussi de degré 2 et conclure.

Corollaire. *Soit $K\subset L$ une extension de corps. Le sous-ensemble*

$$F:=\{\alpha\in L\mid \alpha \textit{ est algébrique sur } K\}$$

est un sous-corps de L.

Preuve. L'ensemble F est stable par toutes les opérations de corps donc est un sous-corps de L. □

Exemple. Considérons $\bar{\mathbf{Q}}:=\{x\in\mathbf{C}\mid \text{ est algébrique sur } \mathbf{Q}\}$, c'est un sous-corps de $\mathbf{C}$. De plus $\bar{\mathbf{Q}}$ est algébriquement clos. En effet soit $P=X^n+a_{n-1}X^{n-1}+\ldots+a_0\in\bar{\mathbf{Q}}[X]$, montrons qu'il possède une racine dans $\bar{\mathbf{Q}}$. Introduisons $K=\mathbf{Q}(a_0,\ldots,a_{n-1})$ alors $[K:\mathbf{Q}]<\infty$. En effet

$$[\mathbf{Q}(a_0,\ldots,a_{n-1}):\mathbf{Q}]=[\mathbf{Q}(a_0,\ldots,a_{n-1}):\mathbf{Q}(a_0,\ldots,a_{n-2})]\ldots[\mathbf{Q}(a_0):\mathbf{Q}]$$

et $[\mathbf{Q}(a_0,\ldots,a_i):\mathbf{Q}(a_0,\ldots,a_{i-1})]\leq[\mathbf{Q}(a_i):\mathbf{Q}]<\infty$. Soit maintenant $x\in\mathbf{C}$ une racine de P (il en existe puisque $\mathbf{C}$ est algébriquement clos) alors, comme $P\in K[X]$ on a $[K(x):K]<\infty$ donc $[\mathbf{Q}(x):\mathbf{Q}]\leq[K(x):\mathbf{Q}]=[K(x):K][K:\mathbf{Q}]<\infty$. Donc x est algébrique sur $\mathbf{Q}$ et appartient donc bien à $\bar{\mathbf{Q}}$.

Nous disposons maintenant de tous les outils nécessaires pour construire des extensions de corps. Nous savons déjà construire, à partir de K le corps $K(X)=\mathrm{Frac}(K[X])$. Soit $P=a_0+a_1X+\ldots+a_nX^n$ un polynôme irréductible de $K[X]$ alors $L:=K[X]/PK[X]$ est un corps qui contient de manière naturelle un sous-corps isomorphe à K. En effet considérons

$$i=s\circ j:K\overset{j}{\hookrightarrow}K[X]\overset{s}{\to}K[X]/PK[X]$$

on obtient $K':=i(K)\cong K$. Montrons que l'élément $\alpha\in L$ égal à la classe de X dans $K[X]/PK[X]$ est racine de $P'=i(a_0)+i(a_1)X+\ldots+i(a_n)X^n\in K'$. En effet

$$\begin{aligned}P'(\alpha)&=i(a_0)+i(a_1)\alpha+\ldots+i(a_n)\alpha^n\\&=s\circ j(a_0)+s\circ j(a_1)s(X)+\ldots+s\circ j(a_n)s(X)^n\\&=s\left(j(a_0)+j(a_1)X+\ldots+j(a_n)X^n\right)\\&=s(P)\\&=0\end{aligned}$$

On voit qu'ainsi on peut fabriquer des extensions L d'un corps K quelconque, telles que des polynômes donnés à coefficients dans K admettent des racines dans L. On peut se demander si de telles constructions sont uniques en un certain sens. Voici la réponse.

Théorème. *Soit K un corps et $P \in K[X]$ non constant.*

(i) Il existe $L \supset K$ telle que L contienne une racine de P. De plus, si P est irréductible dans $K[X]$ et si L est minimale (i.e. si $K \subset L' \subset L$ et P possède une racine dans L' alors $L = L'$) alors L est unique à isomorphisme près et s'appelle un corps de rupture de P (en fait $L \cong K[X]/PK[X]$).

(ii) Il existe une extension $L \supset K$ telle que P soit scindé sur L c'est-à-dire $P = a(X-\alpha_1)\dots(X-\alpha_n)$ avec $a, \alpha_1, \dots, \alpha_n \in L$ et minimale; une telle extension est unique à isomorphisme près et s'appelle le corps de décomposition de P sur K.

Preuve. (i) Soit L un corps contenant une racine α de P, alors $K(\alpha) \subset L$ donc L est minimal si et seulement si $L = K(\alpha)$; dans ce cas l'évaluation en α induit un isomorphisme $K[X]/PK[X] \cong K(\alpha) = L$. Prouvons maintenant, par récurrence sur $n = \deg(P)$, l'existence d'un corps de décomposition. Soit P_1 un facteur irréductible de P et K_1 un corps de rupture minimal de P_1 dans lequel il acquiert une racine α_1. Alors, dans $K_1[X]$ on peut factoriser $P = (X-\alpha_1)Q$. On dispose, par hypothèse de récurrence, d'une extension $L_1 \supset K_1$ sur laquelle Q, et par conséquent P est scindé, i.e. $P = a(X-\alpha_1)\dots(X-\alpha_n)$ avec $a, \alpha_1, \dots, \alpha_n \in L_1$. On pose $L := K(\alpha_1, \dots, \alpha_n)$ et alors P est encore scindé sur L et L est minimal puisque si $K \subset L' \subset L$ et P scindé sur L' alors L' contient K et les racines de P, c'est-à-dire $\alpha_1, \dots, \alpha_n$ donc contient L. Prouvons maintenant, par récurrence sur $n = \deg(P)$, l'unicité (à isomorphisme près) d'un corps de décomposition. Pour faciliter l'induction, on va démontrer un résultat un tout petit plus général (qui achèvera la preuve du théorème) :

Lemme. *Soit $i : K \to K'$ un isomorphisme de corps. Soit P un polynôme de $K[X]$ et L un corps de décomposition de P sur K et soit L' un corps de décomposition de $i(P)$ sur K' alors il existe un isomorphisme $\phi : L \to L'$ qui prolonge i.*

Preuve. Tout d'abord on étend i en un isomorphisme $K[X] \to K'[X]$ que l'on note encore i. Soit $\alpha_1 \in L$ une racine de P et P_1 son polynôme minimal alors $P = P_1 Q$ et $i(P) = i(P_1)i(Q)$. Soit $\alpha_1' \in L'$ une racine de $i(P_1)$. Alors $L_1 = K(\alpha_1)$ est un corps de rupture de P_1 et $L_1' = K(\alpha_1')$ est un corps de rupture de $i(P_1)$ donc on peut prolonger i en un isomorphisme $\phi_1 : L_1 :\to L_1'$ qui envoie α_1 sur α_1'. La factorisation $P = (X-\alpha_1)R$ dans $L_1[X]$ se traduit par la factorisation $i(P) = (X-\alpha_1')\phi_1(R)$ dans $L_1'[X]$. Mais L est un corps de décomposition de R sur L_1 et L' est un corps de décomposition de $\phi_1(R)$ sur L_1' donc, par hypothèse de récurrence, l'isomorphisme ϕ_1 se prolonge en un isomorphisme $\phi : L \to L'$. □

Exemple de corps de décomposition. Soit $K = \mathbf{Q}$ et $P = X^n - 2$, alors un corps de rupture est $\mathbf{Q}(\sqrt[n]{2})$ et un corps de décomposition $L = \mathbf{Q}\left(\exp(2i\pi/n)\sqrt[n]{2}, k = 0, 1, \dots, n-1\right) = \mathbf{Q}\left(\sqrt[n]{2}, \exp(2i\pi/n)\right)$.

Ces théorèmes généraux montrent l'importance des polynômes irréductibles dans $K[X]$. Il est clair que les polynômes de degré 1 sont toujours irréductibles. De même un polynôme de degré 2 ou 3 est irréductible si et seulement si il ne possède pas de racine dans K. Déterminer les autres polynômes irréductibles est nettement plus délicat en général. Nous rappelons seulement ici que les seuls polynômes irréductibles de $\mathbf{R}[X]$ sont les polynômes de degré 1 et les polynômes du second degré sans racines réelles; nous donnons aussi deux critères d'irréductibilité et l'exemple des polynômes cyclotomiques.

Proposition. *Soit A un anneau factoriel et $K := \mathrm{Frac}(A)$, soit $P = a_n X^n + \ldots + a_0 \in A[X]$ et soit $p \in A$ un élément irréductible.*

(i) (Critère d'Eisenstein) Supposons que p ne divise pas a_n, que p divise $a_{n-1}, \dots, a_0$, mais que p^2 ne divise pas a_0, alors P est irréductible dans $K[X]$.

(ii) (Critère de réduction) Supposons que p ne divise a_n, et que $\bar{P} \in (A/pA)[X]$ soit irréductible, alors P est irréductible dans $K[X]$.

Preuve. Pour les deux critères, on considère l'homomorphisme de réduction des coefficients d'un polynôme $P \mapsto \bar{P}$ de $A[X]$ dans $(A/pA)[X]$. Supposons donc que $P = QR$ avec $Q, R \in A[X]$, on en déduit $\bar{P} = \bar{Q}\bar{R}$. L'hypothèse de (i) indique que $\bar{P} = uX^n$ avec $u \neq 0$. Ainsi $uX^n = \bar{Q}\bar{R}$ entraîne $\bar{Q} = vX^d$ et $\bar{R} = wX^{n-d}$, si $d \neq 0, n$ on en tirerait que $Q = q_d X^d + \ldots + q_0$ avec p divisant q_0 et $R = r_{n-d}X^{n-d} + \ldots + r_0$ avec p divisant r_0; d'où p^2 divise $q_0 r_0$, ce qui contredirait les hypothèses. On conclut que $\bar{Q}$ ou $\bar{R}$ est constant et donc Q ou R est constant. L'hypothèse de (ii) indique que $\bar{Q}$ ou $\bar{R}$ est inversible donc constant dans $(A/pA)[X]$. Mais l'hypothèse $a_n \notin pA$ entraîne que les coefficients dominants de Q et R ne sont pas non plus divisibles par p et donc que $\deg(Q) = \deg(\bar{Q})$ et $\deg(R) = \deg(\bar{R})$ donc l'un des deux est constant. □

Remarques et exemples. Si l'on sait de plus que $c(P) = 1$ alors, sous les hypothèses de l'un des deux critères, on a P irréductible dans $A[X]$. En utilisant le critère d'Eisenstein pour $A = \mathbf{Z}$ et $p = 2$, on voit que $X^n - 2$ est irréductible dans $\mathbf{Q}[X]$ (ou $\mathbf{Z}[X]$). En utilisant le critère d'Eisenstein pour $A = \mathbf{Z}[Y]$ et $p = Y$, on voit que $P = (Y-1)X^n - Y^2X + Y$ est irréductible dans $A[Y] = \mathbf{Z}[X, Y]$. Le polynôme $\bar{P} = X^4 + X + 1 \in \mathbf{Z}/2\mathbf{Z}[X]$ est irréductible, en effet il n'a pas de racine dans $\mathbf{Z}/2\mathbf{Z}$ et le seul polynôme irréductible sur $\mathbf{Z}/2\mathbf{Z}$ de degré deux est $X^2 + X + 1$ qui ne divise pas $\bar{P}$. Par conséquent le polynôme $P = 11X^4 - 6X^3 + 4X^2 + 7X - 5$ est irréductible dans $\mathbf{Q}[X]$ (ou $\mathbf{Z}[X]$).

Les polynômes cyclotomiques sont les facteurs irréductibles de $X^n - 1$ dans $\mathbf{Q}[X]$ (ou $\mathbf{Z}[X]$); on peut les définir ainsi:

Définition. Soit $n \geq 1$, le n-ème *polynôme cyclotomique* est défini par

$$\Phi_n(X) = \prod_{\zeta \in \mu_n^*} (X - \zeta)$$

où μ_n^* est l'ensemble des racines n-èmes primitives de l'unité (dans $\mathbf{C}$).

Avec la définition donnée $\Phi_n \in \mathbf{C}[X]$ et il est clair que $\deg(\Phi_n) = \phi(n)$ et que

$$X^n - 1 = \prod_{d \,|\, n} \Phi_d(X) \qquad (*)$$

Cependant il est moins évident qu'en fait $\Phi_n \in \mathbf{Z}[X]$ et que Φ_n est irréductible dans $\mathbf{Q}[X]$ (ou $\mathbf{Z}[X]$). Commençons par voir que les coefficients de Φ_n sont entiers. Il est clair que $\Phi_1(X) = X - 1 \in \mathbf{Z}[X]$. On peut alors démontrer ce que l'on veut par induction sur n en utilisant la formule $(*)$. En effet le polynôme $B := \prod_{d \,|\, n, d \neq n} \Phi_d(X)$ est unitaire et, par hypothèse de récurrence, à coefficients entiers; on peut donc effectuer dans $\mathbf{Z}[X]$ la division euclidienne $X^n = BQ + R$. La formule $(*)$ garantit alors que $R = 0$ et $Q = \Phi_n$. Nous concluons avec le résultat suivant:

Théorème. *Le polynôme Φ_n est irréductible dans $\mathbf{Z}[X]$.*

Preuve. Soit ζ une racine primitive n-ème de l'unité et P son polynôme minimal sur $\mathbf{Q}$, on veut montrer que $P = \Phi_n$. Observons d'abord que $P \in \mathbf{Z}[X]$. Choisissons ensuite p un nombre premier ne divisant pas n alors ζ^p est encore une racine primitive n-ème de l'unité. Soit Q son polynôme minimal qui est également dans $\mathbf{Z}[X]$. Si P et Q étaient distincts, le produit PQ diviserait Φ_n. Mais comme $Q(\zeta^p) = 0$ on voit que ζ est racine de $Q(X^p)$ et donc $Q(X^p) = P(X)R(X)$ pour un certain $R \in \mathbf{Z}[X]$. En réduisant les coefficients modulo p on obtient:

$$\bar{Q}(X^p) = \bar{Q}(X)^p = \bar{P}(X)\bar{R}(X).$$

ou encore $\bar{P}(X)$ divise $\bar{Q}(X)^p$ dans $(\mathbf{Z}/p\mathbf{Z})[X]$ mais les facteurs de $X^n - 1$ et donc de $\bar{P}(X)$ sont simples dans $(\mathbf{Z}/p\mathbf{Z})[X]$ (la dérivée de $X^n - 1$ est nX^{n-1} et on a pris soin de choisir p ne divisant pas n) donc en fait $\bar{P}(X)$ divise $\bar{Q}(X)$. Mais alors $\bar{P}(X)^2$ divise $\bar{\Phi}_n(X)$ dans $(\mathbf{Z}/p\mathbf{Z})[X]$, ce qui contredit le fait que les facteurs de $\bar{\Phi}_n(X)$ sont simples. En résumé on a prouvé que, pour p premier ne divisant pas n, le polynôme minimal de ζ annulait ζ^p. On en tire aisément que, si m est premier avec n alors $P(\zeta^m) = 0$. Ainsi $\deg(P) \geq \phi(n)$ et comme P divise Φ_n, on a donc $P = \Phi_n$ et ce dernier est irréductible. □

Corollaire. *Soit ζ une racine primitive m-ème, alors $[\mathbf{Q}(\zeta) : \mathbf{Q}] = \phi(m)$.*

Preuve. Le polynôme minimal sur $\mathbf{Q}$ de ζ est Φ_m qui est de degré $\phi(m)$. □

Exercices. Montrer les formule suivantes
(a) Si p est premier, $\Phi_p(X) = X^{p-1} + X^{p-2} + \ldots + X + 1$.
(b) Si p premier divise n, alors $\Phi_{np}(X) = \Phi_n(X^p)$.
(c) Si p premier ne divise pas n, alors $\Phi_{np}(X)\Phi_n(X) = \Phi_n(X^p)$.

Montrer que, si $n \geq 3$, on a $[\mathbf{Q}(\cos(2\pi/n)) : \mathbf{Q}] = \phi(n)/2$. Pouvez-vous déterminer $[\mathbf{Q}(\sin(2\pi/n)) : \mathbf{Q}]$?

C.3. Corps finis.

Nous verrons en appendice qu'un corps fini est nécessairement commutatif. Si K est fini, sa caractéristique est un nombre premier p et K est un espace vectoriel de dimension finie (disons n) sur $\mathbf{Z}/p\mathbf{Z}$. On en tire en particulier que $\mathrm{card}(K) = \mathrm{card}\left((\mathbf{Z}/p\mathbf{Z})^n\right) = p^n$. Nous allons démontrer

Théorème. *Soit p un nombre premier et un entier $n \geq 1$, alors il existe un corps de cardinal p^n, unique à isomorphisme près. On le note $\mathbf{F}_{p^n}$.*

Remarque. Si $n = 1$ on connait déjà ce résultat et en fait $\mathbf{F}_p = \mathbf{Z}/p\mathbf{Z}$. Cependant, si $n \geq 2$, on a $\mathbf{F}_{p^n} \cong (\mathbf{Z}/p\mathbf{Z})^n$ en tant que $\mathbf{Z}/p\mathbf{Z}$-espaces vectoriels ou en tant que groupes additifs mais pas en tant qu'anneaux. On a ainsi trois anneaux à ne pas confondre : $\mathbf{Z}/p^n\mathbf{Z}$, $(\mathbf{Z}/p\mathbf{Z})^n$ et $\mathbf{F}_{p^n}$.

Exemple. Le polynôme $X^2 + X + 1 \in \mathbf{F}_2[X]$ est irréductible donc $\mathbf{F}_2[X]/(X^2 + X + 1)\mathbf{F}_2[X]$ est un corps de dimension 2 sur $\mathbf{F}_2$ donc de cardinal 4 donc isomorphe à $\mathbf{F}_4$.

Revenons à un corps fini K de cardinal $q = p^n$. On sait donc que $\mathrm{card}(K^*) = q - 1$ et donc que pour tout $x \in K^*$ on a $x^{q-1} = 1$ et donc pour tout $x \in K$ on a $x^q - x = 0$. Remarquons que si $X^q - X$ est considéré comme un polynôme à coefficients dans $\mathbf{F}_p$ on obtient la factorisation $X^q - X = \prod_{\alpha \in K}(X - \alpha) \in K[X]$. Ceci suggère l'énoncé suivant:

Théorème. *Soit $q = p^n$ et K le corps de décomposition de $X^q - X$ sur $\mathbf{F}_p = \mathbf{Z}/p\mathbf{Z}$. C'est un corps de cardinal $q = p^n$ et tout corps de cardinal q lui est isomorphe.*

Preuve. Il suffit de prouver que si K est le corps de décomposition de $X^q - X$ sur $\mathbf{F}_p = \mathbf{Z}/p\mathbf{Z}$, c'est un corps de cardinal q. Dans $K[X]$ on a $X^q - X = \prod_{i=1}^{q}(X - \alpha_i)$. Posons $S := \{\alpha \in K \mid \alpha^q - \alpha = 0\}$. L'ensemble S des racines de $X^q - X$ dans K a pour cardinal q car $X^q - X$ est scindé sur K et les racines sont simples car la dérivée est le polynôme constant -1. Montrons que S est un sous-corps de K et donc $K = S$. En effet si $\alpha^q - \alpha = 0$ et $\beta^q - \beta = 0$ alors $(\alpha + \beta)^q - (\alpha + \beta) = \alpha^q + \beta^q - \alpha - \beta = 0$ et donc $\alpha + \beta \in S$; par ailleurs si $p \neq 2$, on a $(-\alpha)^q - (-\alpha) = -\alpha^q + \alpha = 0$ donc $-\alpha \in S$; enfin $(\alpha\beta)^q = \alpha^q\beta^q = \alpha\beta$ donc $\alpha\beta \in S$ et (si α est non nul) $(\alpha^{-1})^q = \alpha^{-q} = \alpha^{-1}$ donc $\alpha^{-1} \in S$. $\square$

Remarques. Il est clair que l'homomorphisme $\phi : \mathbf{F}_{p^n} \to \mathbf{F}_{p^n}$ défini par $\phi(x) = x^p$ est un isomorphisme car une application injective entre deux ensembles finis de même cardinal est une bijection. On a clairement $\phi^n = id_{\mathbf{F}_{p^n}}$ puisque $x^{p^n} = x$ pour tout $x \in \mathbf{F}_{p^n}$. Par ailleurs, nous avons vu qu'un sous-groupe fini de K^* (avec K corps commutatif) est cyclique, donc $\mathbf{F}_{p^n}^*$ est isomorphe (comme groupe) à $\mathbf{Z}/(p^n - 1)\mathbf{Z}$. On voit donc que l'application $x \mapsto x^m$ définit une bijection de $\mathbf{F}_{p^n}^*$ (ou $\mathbf{F}_{p^n}$) si et seulement si $\mathrm{PGCD}(m, p^n - 1) = 1$; c'est un homomorphisme de groupe sur $\mathbf{F}_{p^n}^*$ mais bien sûr pas un homomrphisme d'anneaux sur $\mathbf{F}_{p^n}$. Lorsque $d := \mathrm{PGCD}(m, p^n - 1)$ est différent de 1, le noyau est cyclique de cardinal d et on a $(\mathbf{F}_{p^n}^* : \mathbf{F}_{p^n}^{*m}) = d$.

Exercices. Montrer que $\mathbf{F}_q$ est (isomorphe à) un sous-corps de $\mathbf{F}_{q'}$ si et seulement si $q = p^m$ et $q' = p^n$ avec m divisant n.

Appendice : le théorème de Wedderburn.

Il s'agit du résultat suivant:

Théorème. (théorème de Wedderburn) *Soit K un corps fini, alors K est commutatif.*

Preuve. Soit $Z = \{x \in K \mid \forall y \in K,\ xy = yx\}$ alors Z est clairement un sous-corps commutatif de K; notons $q = \mathrm{card}(Z)$ et $n = \dim_Z K$. On va montrer par l'absurde qu'on ne peut avoir $n \geq 2$. Considérons le groupe K^* et son action sur lui-même par conjugaison. Soit $y \in K^*$, si on pose $C(y) = \{x \in K \mid xy = yx\}$ alors $C(y)$ est un sous-corps de K qui contient Z; notons $n_y = \dim_Z C(y)$. On a $C(y) = K$ si et seulement si $y \in Z$ et le stabilisateur de y sous l'action de K^* est $C(y)^* = C(y) \setminus \{0\}$, ainsi la formule des classes s'écrit:

$$q^n - 1 = \mathrm{card}(K^*) = \mathrm{card}(Z^*) + \sum_{y \in R} \frac{\mathrm{card}(K^*)}{\mathrm{card}(C(y)^*)} = q - 1 + \sum_{y \in R} \frac{q^n - 1}{q^{n_y} - 1}$$

où R désigne un ensemble de représentants des classes de conjugaison non réduites à un élément, ou encore telles que $1 \leq n_y < n$. On fait maintenant l'observation que $X^n - 1 = \prod_{d \,|\, n} \Phi_d(X)$ où $\Phi_d \in \mathbf{Z}[X]$ désigne

le polynôme cyclotomique. On voit donc que $q^n - 1 = \prod_{d\,|\,n} \Phi_d(q)$ et donc que $\Phi_n(q)$ divise $q^n - 1$ et même $(q^n - 1)/(q^{n_y} - 1)$ lorsque $n_y < n$. En revenant à l'équation des classes, on voit donc que $\Phi_n(q)$ divise $q - 1$. En particulier $|\Phi_n(q)| \leq q - 1$. Mais $|\Phi_n(q)| = \prod_\zeta |q - \zeta|$ où ζ parcourt les racines n-èmes primitives et l'on a $|q - \zeta| \geq q - 1$, d'où une contradiction si $n \geq 2$. □

Exercice (Théorème de Chevalley-Waring). Soit $k = \mathbf{F}_q$ un corps fini de caractéristique p. On veut montrer que si $P \in k[x_1, \ldots, x_n]$ avec $\deg(P) < n$ alors

$$\mathrm{card}\{x \in k^n \mid P(x) = 0\} \equiv 0 \bmod p.$$

En particulier, si P est homogène de degré $d < n$ alors P possède un zéro non trivial (i;e. distinct de 0). On pourra procéder ainsi :

(a) Montrer que $\sum_{x \in k} x^m$ est nul si $m = 0$ ou si $q - 1$ ne divise pas m mais vaut -1 dans les autres cas. [Comme le polynôme "X^0" est le polynôme constant, il est naturel de prendre ici la convention $0^0 = 1$].
(b) Soit $P \in k[x_1, \ldots, x_n]$ avec $\deg(P) < (q - 1)n$, en déduire que $\sum_{x \in k^n} P(x) = 0$.
(c) Appliquer le résultat précédent à $P(x)^{q-1}$ et conclure.
(d) Démontrer par une méthode analogue la généralisation suivante. Soient $P_1, \ldots, P_s$ des polynômes de degrés $d_1, \ldots, d_s$ avec $d_1 + \ldots + d_s < n$, montrer que

$$\mathrm{card}\{x \in k^n \mid P_1(x) = \ldots = P_s(x) = 0\} \equiv 0 \bmod p.$$

En particulier, si les polynômes sont homogènes, ils ont un zéro commun non trivial.

Exercice. Montrer que $\mathbf{F}_{p^m} \subset \mathbf{F}_{p^n}$ si et seulement si m divise n.

D. MODULES.

On donne une brève présentation de la théorie des modules sur un anneau commutatif avec comme objectif et motivation la description de la décomposition d'un endomorphisme d'espace vectoriel et la détermination de sa classe de similitude.

D.1. Modules : généralités et exemples.

Soit A un anneau commutatif, un A-*module* est un ensemble M muni d'une addition $M \times M \to M$ et d'une multiplication par les scalaires $A \times M \to M$ vérifiant les mêmes axiomes qu'un espace vectoriel, c'est-à-dire :
(i) $(M, +)$ est un groupe abélien
(ii) $\forall a, b \in A, \forall x \in M$ on a $a \cdot (b \cdot x) = (ab) \cdot x$
(ii) $\forall a \in A, \forall x, y \in M$ on a $a \cdot (x + y) = a \cdot x + a \cdot y$
(iv) $\forall a, b \in A, \forall x \in M$ on a $(a + b) \cdot x = a \cdot x + b \cdot x$
(v) $\forall x \in M$ on a $1 \cdot x = x$

Remarque. Si A n'est pas commutatif on peut néanmoins définir des modules à droite ou à gauche.

Exemples. Si A est un corps, un A-module n'est rien d'autre qu'un A-espace vectoriel. Un groupe abélien est, de manière "évidente", un $\mathbf{Z}$-module si l'on pose $n \cdot x = x + \cdots + x$ (n fois) pour $n > 0$ et $n \cdot x = -x - \cdots - x$ ($|n|$ fois) pour $n < 0$. Si A est un anneau commutatif et si I est un idéal, alors A/I est naturellement un A-module en posant $a \cdot (x + I) = ax + I$. En particulier A peut être vu comme un A-module.

Opérations sur les modules.

Un *sous-module* N d'un module M est un sous-ensemble tel que les opérations sur M induisent une structure de A-module sur N. C'est-à-dire :
(i) N est un sous-groupe de M
(ii) N est stable par multiplication par un scalaire
ou encore
(i') $\forall x, y \in N, \forall a, b \in A, ax + by \in N$.

Exemples. Les sous-modules de A sont les idéaux de A. Si $a \in A$, l'ensemble $aM := \{ax \mid x \in M\}$ est un sous-module de M; plus généralement, si I est un idéal de l'anneau A, l'ensemble $I \cdot M := \{x = a_1x_1 + \ldots a_rx_r \mid r \geq 0,\ a_i \in I \text{ et } x_i \in M\}$ est un sous-module.

Si N_1 et N_2 sont des sous-modules de M, l'intersection $N_1 \cap N_2$ est un sous-module, la *somme* est le sous-module $N_1 + N_2 = \{x_1 + x_2 \,|\, x_1 \in N_1 \text{ et } x_2 \in N_2\}$. Si de plus $N_1 \cap N_2 = \{0\}$ on dit que la somme est *directe* et on la note $N_1 \oplus N_2$. La notion de somme (directe ou non) se généralise à une famille quelconque de sous-modules $\{N_i\}_{i \in I}$.

Une application $f : M \to N$ est un *homomorphisme de modules*, si elle vérifie $f(x + y) = f(x) + f(y)$ et $f(a \cdot x) = a \cdot f(x)$. Si de plus f est bijective, on dit que c'est un *isomorphisme de modules*.

Remarques. La dernière appellation est justifiée car on vérifie immédiatement que la bijection réciproque f^{-1} est encore un homomorphisme de modules. Le composé de deux homomorphismes est encore un homomorphisme. L'image directe ou réciproque par un homomorphisme d'un sous-module est encore un sous-module. En particulier le noyau $\mathrm{Ker}(f)$ est un sous-module de M et l'image $\mathrm{Im}(f)$ est un sous-module de N. L'ensemble des endomorphismes $f : M \to M$ forme un anneau (non commutatif en général) en posant $(f + g)(x) = f(x) + g(x)$ et $(fg)(x) = f(g(x))$. Si $M = A^r$, alors $\mathrm{End}(M)$ est isomorphe à l'anneau des matrices $r \times r$ à coefficients dans A.

Si N_1 et N_2 sont des sous-modules de M, le *produit de modules* est définit comme l'ensemble $N_1 \times N_2$ muni des lois $(x_1, x_2) + (x_1', x_2') = (x_1 + x_1', x_2 + x_2')$ et $a \cdot (x_1, x_2) = (a \cdot x_1, a \cdot x_2)$.

Remarque. La notion de produit se généralise à une famille quelconque de modules $\{N_i\}_{i \in I}$. Lorsque les N_i sont des sous-modules en somme directe, on a $\prod_{i \in I} N_i \cong \oplus_{i \in I} N_i$ seulement lorsque I est fini.

Soit N un sous-module de M, on peut construire le *module quotient* M/N comme le groupe abélien M/N (déjà construit) muni de la multiplication par un scalaire $a \cdot (x+N) = (a \cdot x) + N$. On a alors la propriété universelle du quotient

Théorème *Soit $f : M \to M'$ un homomorphisme de A-modules et soit N un sous-module et $s : M \to M/N$ la surjection canonique.*
(i) Il existe une application $\hat{f} : M/N \to M'$ telle que $f = \hat{f} \circ s$ si et seulement si $N \subset \mathrm{Ker}(f)$.
(ii) Dans ce cas l'application $\hat{f}$ est un homomorphisme de modules, son image est égale à celle de f (i. e. $\hat{f}(M/N) = f(M)$) et son noyau est $\mathrm{Ker}(f)/N$.

Preuve. En terme de groupe quotient "tout" a déjà été prouvé; il reste seulement à vérifier que l'application $\hat{f}$, quand elle existe, est bien un homomorphisme de modules, ce qui est immédiat. □

Par exemple on en déduit que $M/\mathrm{Ker}(f) \cong \mathrm{Im}(f)$. Si N_1 et N_2 sont deux sous modules de M, l'application $x \mapsto (x, -x)$ identifie $N_1 \cap N_2$ à un sous-module de $N_1 \times N_2$ et l'on voit que $N_1 + N_2 \cong (N_1 \times N_2)/(N_1 \cap N_2)$.

Les notions de *combinaison linéaire*, *partie libre*, de *partie génératrice* ou de *base* se définissent comme en algèbre linéaire sur un corps. Néanmoins une différence notable est la non-existence de base d'un module en général. En fait on peut introduire la notion suivante (qui n'a d'intérêt que si M n'est pas un espace vectoriel ou encore si A n'est pas un corps).

Définition. Soit x élément d'un A-module M, on appelle *annulateur* de x l'idéal

$$\mathrm{Ann}(x) = \{a \in A \,|\, a \cdot x = 0\}.$$

Si N est un sous-module, son annulateur est

$$\mathrm{Ann}(N) = \bigcap_{x \in N} \mathrm{Ann}(x) = \{a \in A \,|\, \forall x \in N, a \cdot x = 0\}.$$

Remarquons qu'un A-module M est automatiquement un $A/\mathrm{Ann}(M)$-module en posant $\bar{a} \cdot x = ax$ (ce qui est loisible puisque ax ne dépend que de la classe $\bar{a}$ de a modulo l'idéal $\mathrm{Ann}(M)$).

Exemple. Soit $M = A/I$ vu comme A-module (avec I idéal de A), on a clairement $\mathrm{Ann}(M) = I$. Considérons $M = \mathbf{Q}/\mathbf{Z}$ vu comme $\mathbf{Z}$-module, pour tout élément x égal à la classe de a/b avec a et b premiers entre eux on a $\mathrm{Ann}(x) = b\mathbf{Z}$, néanmoins $\mathrm{Ann}(M) = \{0\}$. Remarquons que l'ensemble

$$M_{\mathrm{torsion}} := \{x \in M \mid \exists a \in A \setminus \{0\},\ a \cdot x = 0\} = \{x \in M \mid \mathrm{Ann}(x) \neq 0\}$$

est un sous-module de M.

Supposons A intègre, lorsque l'annulateur d'un élémént non nul de M n'est pas réduit à $\{0\}$ on voit tout de suite qu'il ne peut pas exister de base sur A. On donne donc un statut spécial aux modules possédant une base. On définit de même l'analogue de la dimension finie dans les espaces vectoriels.

Définition. Un A-module M est *libre* s'il possède une base (i. e. une partie libre et génératrice sur A). Il est de *type fini* s'il possède une partie génératrice finie.

Ainsi un module libre de type fini est isomorphe à A^n. Il n'est pas évident que l'entier n soit unique, même si cela est vrai ; au paragraphe suivant on vérifie que si A est principal et $A^n \cong A^m$ alors $m = n$. Remarquons aussi que A, considéré comme A-module, est libre de rang 1 et que ses sous-modules non nuls (c'est-à-dire ses idéaux non nuls) sont libres de rang 1 si et seulement si A est principal.

D.2. Modules de type fini sur les anneaux principaux.

Rappelons qu'un anneau commutatif unitaire A est *principal* s'il est intègre et tout idéal est de la forme aA. Nous commençons par montrer qu'on a bien une notion de "dimension", qu'on appellera plutôt *rang*, et on donne ensuite la description des sous-modules d'un module libre de type fini sur un tel anneau.

Proposition. *Soit A un anneau principal, M un A-module admettant deux bases $\mathcal{B}$ et $\mathcal{B}'$ alors* $\mathrm{card}(\mathcal{B}) = \mathrm{card}(\mathcal{B}')$. *Si $M = N \oplus N'$ et si $\mathcal{B}$ et $\mathcal{B}'$ sont des bases de N et N' respectivement, alors $\mathcal{B} \cup \mathcal{B}'$ est une base de M.*

Preuve. Si A est un corps, le résultat est la base de l'algèbre linéaire. Sinon, soit a un élément irréductible de A, alors $k = A/aA$ est un corps et le module quotient M/aM est annulé par aA donc peut être vu comme un k-module c'est-à-dire un k-espace vectoriel. Mais si $e_1, \ldots, e_r$ forment une base de M sur A et si l'on désigne par $\bar{e}_i$ la classe de e_i modulo aM, il est immédiat que $\bar{e}_1, \ldots, \bar{e}_r$ forment une base de M/aM sur k. L'entier r est donc la dimension du k-espace vectoriel M/aM et ne dépend donc pas de la base choisie. La deuxième affirmation est immédiate. □

Définition. Si M est un A-module libre de type fini, on appelle *rang* de M le cardinal d'une base.

Théorème *Soit A un anneau principal, M un A-module libre de rang r, et N un sous-module alors*
(i) Le module N est libre de rang $s \leq r$.
(ii) Il existe $e_1, \ldots, e_r$ base de M sur A et $a_1, \ldots, a_s \in A$ tels que a_i divise a_{i+1} et

$$N = Aa_1e_1 \oplus \ldots \oplus Aa_se_s.$$

Preuve. La preuve se fait par récurrence sur l'entier r, le cas $r = 1$ étant vérifié précisément parce que l'anneau A est supposé principal. Commençons par la preuve de (i). Si l'on note $e_1, \ldots, e_r$ une base de M on peut écrire $M = Ae_1 \oplus \ldots \oplus Ae_r$ et considérer l'homomorphisme de A-modules $e_r^* : M \to A$ défini par $e_r^*(a_1e_1 + \ldots + a_re_r) = a_r$. L'ensemble $e_r^*(N)$ est un sous-module, c'est-à-dire un idéal de A. Choisissons $x_0 \in N$ tel que $e_r^*(x_0) = a$ avec $e_r^*(N) = aA$. On va appliquer le lemme suivant

Lemme. *Soit $f : M \to A$ un homomorphisme non nul de modules et x tel que $f(x)A = f(M)$ alors $M = \mathrm{Ker}(f) \oplus Ax$.*

Preuve du lemme. Soit $y \in \mathrm{Ker}(f) \cap Ax$ alors $y = ax$ et $f(y) = af(x) = 0$, mais $f(x) \neq 0$ car sinon l'homomorphisme f serait nul, donc $a = 0$ (l'anneau A est intègre) et $y = 0$. Soit maintenant $y \in M$, on sait qu'il existe $b \in A$ tel que $f(y) = bf(x) = f(bx)$, donc $f(y - bx) = 0$ et $y - bx \in \mathrm{Ker}(f)$. On peut donc écrire $y = (y - bx) + (bx) \in \mathrm{Ker}(f) + Ax$. □

Si $N \subset \mathrm{Ker}(e_r^*)$ alors, comme $\mathrm{Ker}(e_r^*) = Ae_1 \oplus \ldots \oplus Ae_{r-1}$, on peut appliquer l'hypothèse de récurrence et conclure que N est libre de rang $\leq r - 1$. Sinon, en appliquant le lemme à $e_r^* : N \to A$ on obtient que $N = (\mathrm{Ker}(e_r^*) \cap N) \oplus Ax_0$. En appliquant l'hypothèse de récurrence au sous-module $\mathrm{Ker}(e_r^*) \cap N \subset \mathrm{Ker}(e_r^*)$, on obtient que $\mathrm{Ker}(e_r^*) \cap N$ est libre de rang $\leq r - 1$. Donc N est libre de rang $\leq r$.

Montrons maintenant (ii), toujours par récurrence sur r. Pour chaque homomorphisme de modules $f : M \to A$ tel que $f(N) \neq 0$, on choisit $a_f \in A$ tel que $f(N) = a_fA$ et $u_f \in N$ tel que $f(u_f) = a_f$. On choisit ensuite f_1 tel que $a_{f_1}A$ soit maximal parmi les a_fA. Remarque : cela signifie que si $a_{f_1}A \subset a_fA$ alors $a_{f_1}A = a_fA$ mais on ne peut pas, à ce stade de la preuve, affirmer que a_{f_1} divise tous les a_f. Pour alléger les notations on écrira $a_1 = a_{f_1}$; on choisit aussi $u_1 \in N$ tel que $f_1(u_1) = a_1$. Montrons d'abord que pour tout f on a a_1 divise $f(u_1)$. Appelons $d = \mathrm{PGCD}(a_1, f(u_1))$, alors, d'après le théorème de Bézout, il existe $b, c \in A$ tels que $d = ba_1 + cf(u_1)$. Considérons alors l'homomorphisme $f' = bf_1 + cf$, on a $f'(u_1) = d$ donc $a_{f'}$ divise d qui divise a_1 ou encore $a_1A \subset a_{f'}A$ d'où $a_1A = dA = a_{f'}A$. Mais $a_1 = \mathrm{PGCD}(a_1, f(u_1))$ signifie exactement que a_1 divise $f(u_1)$. On en tire l'existence de $e_1 \in M$ tel que $u_1 = a_1e_1$ et donc $f(e_1) = 1$; en effet si $y_1, \ldots, y_r$ est une base de M alors $y_i^*(u_1) = a_1b_i$ et donc $u_1 = \sum_i y_i^*(u_1)y_i = a_1(\sum_i b_iy_i)$. On applique alors le lemme précédent à $f_1 : M \to A$ avec l'élément e_1 puis à $f_1 : N \to A$ avec l'élément u_1, ce qui donne

$$M = Ae_1 \oplus \mathrm{Ker}(f_1) \qquad \text{et} \qquad N = Aa_1e_1 \oplus (N \cap \mathrm{Ker}(f_1)).$$

Comme, d'après (i), $\mathrm{Ker}(f_1)$ est libre de rang $r - 1$, on peut lui appliquer l'hypothèse de récurrence et conclure qu'il existe une base $e_2, \ldots, e_r$ de $\mathrm{Ker}(f_1)$ et des éléments $a_r, \ldots, a_r \in A$ tels que a_i divise a_{i+1} et

$$N \cap \mathrm{Ker}(f_1) = Aa_2e_2 \oplus \ldots \oplus Aa_re_r.$$

Il reste donc seulement à vérifier que a_1 divise a_2. Pour cela considérons $f = e_1^* + e_2^*$; on a $f(a_2 e_2) = a_2$ donc a_f divise a_2 et par ailleurs $f(u_1) = f(a_1 e_1) = a_1$ donc a_f divise a_1 mais on a vu que cela entraînait $a_1 A = a_f A$ donc on a bien a_1 qui divise a_2. □

Théorème *Soit A un anneau principal, M un A-module de type fini, il existe $r, m \in \mathbf{N}$ et $a_1, \ldots, a_m \in A$ éléments non nuls et non inversibles tels que a_i divise a_{i+1} et*

$$M \cong A^r \times A/a_1 A \times \ldots \times A/a_m A,$$

De plus, les entiers r, s et la suite d'idéaux $a_m A \subset \ldots \subset a_1 A$ sont uniques.

Preuve. Soient $x_1, \ldots, x_n$ des générateurs de M (comme A-module), on a donc un homomorphisme surjectif $\Phi : A^n \to M$ défini par $\Phi(b_1, \ldots, b_n) = b_1 x_1 + \ldots + b_n x_n$. Soit $N = \mathrm{Ker}(\Phi)$, on a $M \cong A^n/N$ et, d'après le théorème précédent il existe une base $e_1, \ldots, e_n$ de M sur A et $a_1, \ldots, a_n \in A$ que a_i divise a_{i+1} et $N = A a_1 e_1 \oplus \ldots \oplus A a_n e_n$. On montre aisément que

$$M \cong A^n/N = (Ae_1 \oplus \ldots \oplus Ae_n) \,/\, (Aa_1 e_1 \oplus \ldots \oplus Aa_n e_n) \cong A/a_1 A \times \ldots \times A/a_n A.$$

On peut omettre dans cette décomposition les facteurs avec a_i inversible et si $a_i = 0$ on peut écrire $A/a_i A \cong A$ d'où le résultat annoncé. L'unicité se démontre aisément à partir de l'observation que, d'une part $M/bM \cong (A/bA)^r \times A/\operatorname{PGCD}(a_1, b)A \times \ldots \times A/\operatorname{PGCD}(a_n, b)A$ et d'autre part $bM \cong A^r \times A/(a_1/\operatorname{PGCD}(a_1, b))A \times \ldots \times A/(a_n/\operatorname{PGCD}(a_n, b))A$. □

Pour accentuer le parallèle avec les groupes abéliens, définissons un *A-module cyclique* comme un A-module isomorphe à A/aA. Le théorème précédent affirme qu'un module de torsion et de type fini est isomorphe à un produit ou somme fini de modules cycliques. Ceci est bien une généralisation du théorème décrivant les groupes finis abéliens comme produit de groupes cycliques.

Terminons ce paragraphe en donnant une version utile du théorème de structure des sous-modules de A^n.

Lemme. *Soit $M \in \mathrm{Mat}(n \times m, A)$ avec A principal, il existe $U \in \mathrm{GL}_n(A)$ et $V \in \mathrm{GL}_m(A)$ et $a_1, \ldots, a_s \in A \setminus \{0\}$ avec $s = \mathrm{rang}(M) \leq \min(m, n)$ et a_i divisant a_{i+1} tels que*

$$M = U \begin{pmatrix} a_1 & & & O \\ 0 & a_2 & & \\ & & \ldots & 0 \\ & & a_s & 0 \\ & & & 0 \end{pmatrix} V.$$

Variante. *Soit un homomorphisme $f : A^n \to A^m$, il existe $e_1, \ldots, e_n$ base de A^n et $f_1, \ldots, f_m$ base de A^m et $a_1, \ldots, a_s \in A \setminus \{0\}$ avec $s = \mathrm{rang}(f) \leq \min(m, n)$ et a_i divisant a_{i+1} tels que*

$$f(e_i) = \begin{cases} a_i f_i & \text{si } 1 \leq i \leq s \\ 0 & \text{sinon} \end{cases}$$

Preuve. Prouvons par exemple la variante. Il existe a_i et f_i tels que le sous-module $f(A^n) \subset A^m$ soit égal à $a_1 A f_1 \oplus \ldots \oplus a_s A f_s$. Choisissons $e_i \in A^n$ tel que $f(e_i) = a_i f_i$ (pour $1 \leq i \leq s$); on a alors $A^n = Ae_1 \oplus \ldots \oplus Ae_s \oplus \mathrm{Ker}(f)$. En choisissant $e_{s+1}, \ldots, e_n$ une base de $\mathrm{Ker}(f)$ on obtient l'énoncé.

D.3. Facteurs invariants de matrices.

D.3.1. Le $K[X]$-module associé à un endomorphisme sur un espace vectoriel.

Définition. Soit E un K-espace vectoriel de dimension finie n et $u \in \mathrm{End}_K(E)$. On définit une structure de $K[X]$-module sur l'ensemble E de la façon suivante : l'addition est l'addition dans l'espace vectoriel et la multiplication par un polynôme $P = a_0 + a_1 X + \ldots + a_d X^d$ est définie par

$$P \cdot x = P(u)(x) = (a_0 I + a_1 u + \ldots + a_d u^d)(x) = a_0 x + a_1 u(x) + \ldots + a_d u^d(x).$$

On notera E_u le $K[X]$-module ainsi obtenu. On remarque tout de suite qu'il s'agit d'un module de type fini. De plus, $\mathrm{Ann}(E_u)$ est non trivial puisqu'il contient le polynôme caractéristique (théorème de Cayley-Hamilton) donc le module E_u est de torsion (on peut aussi utiliser le fait que, pour $x \in E$, les vecteurs $x, u(x), u^2(x), \ldots, u^n(x)$ sont liés).

Proposition. *Soit $u, v \in \mathrm{End}_k(E)$, alors les $K[X]$-modules E_u et E_v sont isomorphes si et seulement si les endomorphismes u et v sont semblables, c'est-à-dire qu'il existe une application K-linéaire inversible h telle que $v = h \circ u \circ h^{-1}$.*

Preuve. Pour distinguer les structures de $K[X]$-modules E_u et E_v dans cette preuve nous noterons $P \cdot_u x = P(u)(x)$ et $P \cdot_v x = P(v)(x)$. Supposons qu'il existe h linéaire inversible telle que $v = h \circ u \circ h^{-1}$, alors $v^m = h \circ u^m \circ h^{-1}$ et plus généralement $P(v) = h \circ P(u) \circ h^{-1}$ donc

$$h(P \cdot_u x) = h(P(u)(x)) = (h \circ P(u))(x) = (P(v) \circ h)(x) = P \cdot_v h(x).$$

Ainsi h est en fait un isomorphisme de $K[X]$-modules $h : E_u \to E_v$. Supposons inversement qu'il existe un tel isomorphisme de $K[X]$-modules $h : E_u \to E_v$. L'application h est en particulier K-linéaire et bijective et de plus

$$h(u(x)) = h(X \cdot_u x) = X \cdot_v h(x) = v(h(x))$$

donc on a $h \circ u = v \circ h$ et u et v sont semblables. □

Par ailleurs, avant d'appliquer à notre situation les théorèmes de structure du paragraphe précédent, observons qu'un $K[X]$-sous-module de E_u n'est rien d'autre qu'un sous-espace vectoriel stable par u. Ainsi une décomposition en somme de sous-modules correspond à une décomposition en somme de sous-espaces vectoriels stables par u. De même un sous-module cyclique correspond à un sous-espace vectoriel engendré par un vecteur x et ses images successives $u(x), u^2(x), \ldots$ par l'endomorphisme u.

D.3.2. Facteurs invariants d'un endomorphisme.

Le module E_u est isomorphe à $K[X]/P_1K[X] \times \ldots \times K[X]/P_rK[X]$ avec P_i non constants et P_i divise P_{i+1}, de plus les P_i sont uniques (à un scalaire près), ce qui justifie la

Définition. Les polynômes P_i s'appellent les *facteurs invariants* de u.

Remarquons qu'il est assez facile de voir (démontrez-le!) que P_r est le polynôme minimal de u, tandis que le polynôme caractéristique est égal au produit $P_1 \ldots P_r$. Nous allons généraliser cette observation ci-dessous.

D'après ce qui précède, u et v sont semblables si et seulement si ils ont mêmes facteurs invariants. Donnons maintenant une interprétation de ces invariants et une méthode de calcul (théorique). Le module E_u se décompose en $E_1 \oplus \ldots \oplus E_r$ avec E_i module cyclique de la forme $K[X]/PK[X]$. Ces sous-modules correspondent à des sous-espaces vectoriels stables par u sur lequel u agit comme la multiplication par X sur $K[X]/PK[X]$. Soit $P = X^d + p_{d-1}X^{d-1} + \ldots + p_0$, prenons comme K-base de $K[X]/PK[X]$ les éléments $1, X, \ldots, X^{d-1}$ et soit $e_1, \ldots, e_d$ la K-base correspondante de E_i, la matrice de u dans cette base est une matrice dite *compagnon* :

$$\mathrm{Mat}(u; e_1, \ldots, e_d) = \begin{pmatrix} 0 & & & -p_0 \\ 1 & & & \vdots \\ & \ddots & 0 & -p_{d-2} \\ & & 1 & -p_{d-1} \end{pmatrix}$$

On obtient en particulier que toute matrice est semblable à une matrice dont les blocs diagonaux sont les matrices compagnon associées à ses facteurs invariants.

Soit A la matrice de u dans une base. Définissons $D_i = D_i(A)$ comme le PGCD des mineurs d'ordre i de la matrice $A - XId$. En particulier D_n est le polynôme caractéristique de u ou A.

Théorème *Les matrices A et B sont semblables si et seulement si $D_i(A) = D_i(B)$ pour $1 \leq i \leq n$.*

Preuve. Posons $A = \mathrm{Mat}(u; (e_1, \ldots, e_n)$. La matrice $A - Xid$ définit un endomorphisme $\Phi : K[X]^n \to K[X]^n$; définissons également $\mu : K[X]^n \to E_u$ par

$$\mu(P_1, \ldots, P_n) = P_1 \cdot e_1 + \ldots + P_n \cdot e_n = P_1(u)(e_1) + \ldots + P_n(u)(e_n).$$

L'homomorphisme μ est clairement surjectif et $\Phi(K[X]^n) \subset \mathrm{Ker}\,\mu$; en effet

$$\begin{aligned}\mu(\Phi(0, \ldots, P_i, \ldots, 0) &= \mu(a_{1i}P_i, \ldots, a_{ii}P_i - XP_i, \ldots, a_{ni}P_i) \\ &= a_{1i}P_i(e_1) + \ldots + a_{ni}P_i(e_n) - uP_i(u)(e_i) \\ &= P_i(u)\,(a_{1i}e_1 + \ldots + a_{ni}e_n - u(e_i)) = 0\end{aligned}$$

Par ailleurs on a vu que le théorème de structure des sous-modules de modules libres peut s'interpréter comme l'existence de deux matrices de changement de base U et V (à coefficient dans $K[X]$) et de polynômes $Q_1, \ldots, Q_n$ avec Q_i divise Q_{i+1} et $A - Xid = U\,\mathrm{diag}(Q_1, \ldots, Q_n)V$. On voit, d'une part, que le PGCD des mineurs d'ordre i est $D_i = Q_1 \ldots Q_i$ et d'autre part que $K[X]^n/\Phi(K[X]^n) \cong K[X]/Q_1K[X] \times \ldots \times K[X]/Q_nK[X]$ d'où l'on tire que $K[X]^n/\Phi(K[X]^n)$ est un K-espace vectoriel de dimension $\sum \deg(Q_i) = \deg\det(A - Xid) = n$. Comme $K[X]^n/\,\mathrm{Ker}(\mu)$ est de même dimension, on en tire $\Phi(K[X]^n) \subset \mathrm{Ker}\,\mu$ et $E_u \cong K[X]^n/\Phi(K[X]^n)$. L'unicité des facteurs invariants de u, disons, $P_1, \ldots, P_r$, implique donc que $(Q_1, \ldots, Q_n) = (1, \ldots, 1, P_1, \ldots, P_r)$. Ainsi la donnée des facteurs invariants P_i équivaut à celle des D_i, ce qui achève la preuve. □

Commentaire. La théorie des $K[X]$-modules nous donne que deux matrices (ou endomorphismes) sont semblables si elles ont les mêmes polynômes "P_i" et le raisonnement précédent montre que la donnée des "P_i" équivaut à celle des "D_i". En fait explicitement $D_{n-i} = P_1 \ldots P_{r-i}$ et $D_{n-r} = \ldots = D_1 = 1$.

Corollaire. *Les matrices A et ${}^tA$ sont semblables.*

Preuve du corollaire. En effet on a clairement $D_i({}^tA) = D_i(A)$. □

Exercice. Fabriquer deux matrices 4×4 non semblables ayant les mêmes polynômes caractéristiques et minimaux (indication : choisir le polynôme minimal $(X - \lambda)^2$ et le polynôme caractéristique $(X - \lambda)^4$). Peut-on fabriquer de tels exemples en dimension 2 ou 3 ?

Exercice. Démontrer de deux façons (en utilisant les résultats précédents et directement) l'énoncé suivant : deux matrices $A, B \in \mathrm{Mat}(n \times n, \mathbf{R})$ sont semblables sur $\mathbf{C}$ (i. e. il existe $U \in \mathrm{GL}(n, \mathbf{C})$ telle que $B = UAU^{-1}$) si et seulement si elles sont semblables sur $\mathbf{R}$ (i. e. il existe $U \in \mathrm{GL}(n, \mathbf{R})$ telle que $B = UAU^{-1}$).

D.3.3. Classes de conjugaison de matrices sur un corps algébriquement clos.

On suppose dans ce paragraphe que le corps K est algébriquement clos et donc tout polynôme est scindé sur K.

Définition. On appelle *bloc de Jordan* de taille d et valeur propre λ la matrice carrée

$$J(d; \lambda) := \begin{pmatrix} \lambda & 0 & & \ldots & & \\ 1 & \lambda & & & & \\ 0 & 1 & \ddots & & & \\ & & \ddots & \ddots & & \\ & & & 1 & \lambda & 0 \\ & & & 0 & 1 & \lambda \end{pmatrix}$$

Si $\lambda = 0$ on note simplement $J(d) = J(d, 0)$. Remarquons que certains auteurs appellent bloc de Jordan la transposée de $J(d, \lambda)$; le principal intérêt de ces matrices est de fournir des représentants explicites des classes de conjugaison de matrices et d'après le corollaire précédent J et ${}^tJ$ sont semblables donc choisir l'une ou l'autre a peu d'influence sur le résultat fondamental suivant

Théorème (Décomposition de Jordan) *Toute matrice carrée est semblable à une matrice composée de blocs de Jordan sur la diagonale et de zéros ailleurs, i.e. du type*

$$J = \begin{pmatrix} J(d_1, \lambda_1) & & \\ & \ddots & \\ & & J(d_r, \lambda_r) \end{pmatrix}$$

De plus les blocs sont uniques, à l'ordre près.

Nous donnons une preuve en terme de $K[X]$-modules, pour une preuve uniquement en terme de K-espace vectoriel, voir le paragraphe suivant.

Preuve. Le $K[X]$-module peut être décomposé en produit de modules cycliques $K[X]/PK[X]$ et comme $P = \prod_\lambda (X-\lambda)^{m_\lambda}$, on peut, en utilisant le lemme chinois généralisé le décomposer en produit de modules cycliques de la forme $K[X]/(X-\lambda)^m K[X]$. Pour analyser ce dernier, quitte à faire le changement de variable $Y = X - \lambda$ (ce qui revient aussi à remplacer u par $u - \lambda id$) on peut supposer $\lambda = 0$. Si l'on note x la classe de X dans $K[X]/X^m K[X]$, une K-base de $K[X]/X^m K[X]$ est fournie par $e_1 = 1$, $e_2 = x$, ..., $e_m = x^{m-1}$ et dans cette base l'action de u, qui correspond à la multiplication par x est donnée par $u(e_1) = e_2$, $u(e_2) = e_3$, ..., $u(e_{m-1}) = e_m$ et enfin $u(e_m) = x^m = 0$. La matrice de $u - \lambda id$ dans cette base est donc bien un bloc de Jordan $J(m)$. L'unicité des blocs (à l'ordre près) est clair si l'on observe que les dimensions des $\mathrm{Ker}(u-\lambda)^j$ sont déterminées par les (d_i, λ_i) et vice versa. □

D.3.4. Supplément : les tableaux de Young

Reprenons l'étude d'un endomorphisme u de E et démontrons directement (i. e. en utilisant uniquement l'algèbre K-linéaire) que u possède une décomposition de Jordan. Si le polynôme caractéristique s'écrit $\det(u - Xid) = \prod_\lambda (X-\lambda)^{m_\lambda}$ alors $E = \oplus \mathrm{Ker}(u-\lambda)^{m_\lambda}$ donc on se ramène au cas d'un endomorphisme nilpotent. On peut donc supposer $u^n = 0$.

Posons $K_i = \mathrm{Ker}(u^i)$ et soit r le plus petit entier tel que $K_r = E$ alors

$$\{0\} \subset K_1 \subset \ldots \subset K_r.$$

Choisissons H_i un supplémentaire de K_{r-i} dans K_{r-i+1} c'est-à-dire tel que $K_{r-i+1} = K_{r-i} \oplus H_i$. Observons que la restriction de u à H_i est injective (en effet $\mathrm{Ker}(u) = K_1 \subset K_{r-i}$ donc $\mathrm{Ker}(u) \cap H_i = \{0\}$) et montrons qu'on peut de plus imposer $u(H_{i-1}) \subset H_i$ (ce qui montrera également que $h_{i-1} \leq h_i$). En effet, une fois choisi H_i, on remarque que si $x \in H_i \subset K_{r-i+1}$ est non nul, alors $u(x) \in K_{r-i}$ mais $u(x) \notin K_{r-i-1}$ (sinon $x \in K_{r-i}$); on a donc $u(H_i) \cap K_{r-i-1} = \{0\}$ et $u(H_i) \subset K_{r-i}$ et l'on peut construire un supplémentaire H_{i+1} de K_{r-i-1} dans K_{r-i} qui contienne $u(H_i)$. Le choix de H_1 est arbitraire.

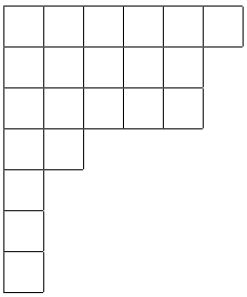
On peut maintenant choisir une base de $E = \oplus_i H_i$ ainsi (où l'on note $h_i = \dim(H_i)$

- On choisit $\mathcal{B}_1 = (e_{1,j})_{1 \leq j \leq h_1}$ base de H_1.
- On choisit $\mathcal{B}_2 = (e_{2,j})_{1 \leq j \leq h_2}$ base de H_2 en imposant $e_{2,j} = u(e_{1,j})$ pour $j \leq h_1$.
- Ayant construit $\mathcal{B}_i = (e_{i,j})_{1 \leq j \leq h_i}$ base de H_i, on choisit $\mathcal{B}_{i+1} = (e_{i+1,j})_{1 \leq j \leq h_{i+1}}$ base de H_{i+1} en imposant $e_{i+1,j} = u(e_{i,j})$ pour $j \leq h_i$.

On regroupe maintenant $E = \oplus_{j=1}^{h_r} E_j$ avec E_j le sous-espace vectoriel ayant pour base $\mathcal{B}'_j = (e_{i,j})_{i \in I_j}$ (où $I_j = \{i \,|\, 1 \leq i \leq r \text{ et } j \leq h_i\}$). On voit facilement que les E_j sont stables par u et que la matrice de $u_{|E_j}$ dans la base $\mathcal{B}'_j$ est une matrice de Jordan de taille $d_j = \mathrm{card}(I_j)$; en effet par construction $u(e_{i,j}) = e_{i+1,j}$ sauf le dernier qui est nul. On obtient donc la matrice de u dans la base $\mathcal{B}' = \mathcal{B}'_1 \cup \ldots \mathcal{B}'_{h_r}$

$$\mathrm{Mat}(u, \mathcal{B}') = \begin{pmatrix} J(d_1) & 0 & \ldots & \\ 0 & J(d_2) & & \\ & & \ddots & 0 \\ & & 0 & J(d_{h_r}) \end{pmatrix}$$

La combinatoire un peu embrouillée peut être clarifiée par l'introduction des *tableaux de Young*. On calcule $h_i = \dim(K_{r-i+1}) - \dim(K_{r-i})$. On dessine un premier tableau en rangeant h_r carrés sur la première ligne, puis h_{r-1} carrés sur la seconde et ainsi de suite (sur le dessin $h_7 = 6$, $h_6 = h_5 = 5$, $h_4 = 2$ et $h_3 = h_2 = h_1 = 1$), le tableau dual s'obtient en inversant ligne et colonnes. On obtient alors d_1 carrés sur la première ligne, d_2 sur la deuxième etc. (sur l'exemple $d_1 = 7$, $d_2 = 4$, $d_3 = d_4 = d_5 = 3$ et $d_6 = 1$), ce qui donne la taille des blocs de Jordan. Ce procédé permet, inversement, de calculer $\dim(K_i)$ à partir de la taille des blocs de Jordan.



Premier tableau

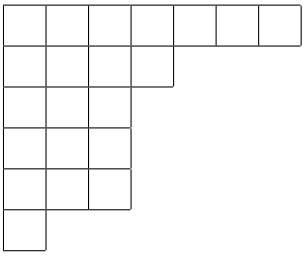


Tableau dual

E. GROUPES CLASSIQUES.

La géométrie "classique" ne considère souvent que la norme euclidienne sur $\mathbf{R}^n$ *donnée par* $||x||^2 = x_1^2 + \ldots + x_n^2$ *et le produit scalaire associé* $\langle x|y\rangle = x_1y_1 + \ldots + x_ny_n$ *ainsi que la norme sur* $\mathbf{C}^n$ *donnée par* $||z||^2 = z_1\bar{z}_1 + \ldots + z_n\bar{z}_n$ *et le produit hermitien associé* $\langle z|w\rangle = z_1\bar{w}_1 + \ldots + z_n\bar{w}_n$ *ainsi bien sûr que les isométries associées. Plusieurs théories amènent néanmoins à considérer des formes plus générales :*

• *La relativité (ou l'équation des ondes) conduit à considérer la forme quadratique de Minkowski qui, en notant* (x_1, x_2, x_3, t) *un vecteur de* $\mathbf{R}^4$*, s'écrit* $||x||^2 = x_1^2 + x_2^2 + x_3^2 - c^2t^2$*.*

• *Les équations de la mécanique de Lagrange (ou Hamilton) amène à introduire pour* $x, y \in \mathbf{R}^{2n}$ *le produit* $B(x, y) = (x_1y_2 - x_2y_1) + \ldots + (x_{2n-1}y_{2n} - x_{2n}y_{2n-1})$*. La parité de la dimension de l'espace s'explique par le fait qu'on considère ensemble la "position" et la "vitesse" d'une particule (espace des phases).*

• *L'arithmétique oblige à considérer d'autres corps que* $\mathbf{R}$ *ou* $\mathbf{C}$*, par exemple le problème de savoir si, pour* n *et* D *donnés, on peut trouver des solutions entières de* $x^2 + Dy^2 = n$ *ou* $x_1^2 + \ldots + x_r^2 = n$*.*

• *Les géométries finies (utiles en théorie des groupes, en combinatoire et bien sûr en informatique) requièrent de travailler sur des corps finis* $\mathbf{Z}/p\mathbf{Z}$ *ou plus généralement* $\mathbf{F}_q$*.*

Le paysage usuel se trouve ainsi démultiplié : si le corps K *possède une topologie (par exemple, si* $K = \mathbf{R}$ *ou* $\mathbf{C}$*) on peut se demander si le groupe est compact, connexe, etc ; si le corps* K *est fini, on peut se demander quel est le cardinal de ces groupes, etc. Néanmoins, une "surprise" est que cette diversification n'entraîne pas la multiplication des groupes associés. En fait à tous les étages, ce sont les groupes dit classiques, c'est-à-dire le groupe des matrices carrés inversibles, le groupe des matrices de déterminant 1, le groupe des matrices respectant une forme quadratique (groupe orthogonal), le groupe des matrices respectant une forme alternée (groupe symplectique), le groupe des matrices respectant une forme hermitienne (groupe unitaire) qui fournissent, à quelques exceptions près, les groupes "intéressants". Nous ne montrerons pas cela (voir l'article de Tits cité en bibliographie) mais incluons néanmoins cette affirmation pour motiver l'étude des groupes classiques.*

E.1. Formes sesqui-linéaires.

Définition. Soit E, F des K-espaces vectoriels, σ un automorphisme de K, une application $f : E \to F$ est *σ-linéaire* si $f(x + y) = f(x) + f(y)$ et $f(ax) = \sigma(a)f(x)$. Une forme $B : E \times F \to K$ est *σ-sesqui-linéaire* si, pour $y \in F$, l'application $B(., y) : E \to K$ est une forme linéaire et, pour $x \in E$, l'application $B(x, .) : F \to K$ est une forme σ-linéaire.

On appelle *noyau à gauche* de B (resp. à droite) l'ensemble

$$\mathrm{Ker}_g(B) = \{x \in E \,|\forall y \in F,\ B(x, y) = 0\} \qquad (\text{resp. } \mathrm{Ker}_d(B) = \{y \in F \,|\forall x \in E,\ B(x, y) = 0\})$$

Il est immédiat de voir que ce sont des sous-espaces vectoriels. On dit que B est *non dégénérée* si ses noyaux à gauche et à droite sont nuls. La forme B induit une forme $\bar{B} : E/\mathrm{Ker}_g(B) \times F/\mathrm{Ker}_d(B) \to K$ définie par $\bar{B}(x + \mathrm{Ker}_g(B), y + \mathrm{Ker}_d(B)) := B(x, y)$. La forme $\bar{B}$ est non dégénérée. Ces considérations permettent en général de se ramener au cas des formes non dégénérées. Dans le cas d'une forme non dégénérée $B : E \times F \to K$, on voit que l'application $y \to B(., y)$ induit une injection de F vers E^* donc $\dim(F) \leq \dim(E^*)$; l'application $x \to \sigma^{-1} \circ B(x, .)$ induit une injection de E vers F^* et donc $\dim(E) \leq \dim(F^*)$. Dans le cas où E (ou F) est de dimension finie, on en tire donc que $\dim(E) = \dim(F)$ et que B permet d'identifier E et F^* (ou F et E^*). Les espaces E et F étant donc isomorphes, on voit que le cas essentiel à considérer est celui d'une forme non dégénérée $B : E \times E \to K$, cas que nous considérons donc désormais.

Il est naturel de considérer la relation d'orthogonalité $x \perp y$ si $B(x, y) = 0$. Une condition naturelle à imposer est que cette relation soit symétrique (i. e. $x \perp y \Leftrightarrow y \perp x$) ; une telle relation est décrite par la proposition suivante

Proposition. *Soit* $B : E \times E \to K$ *une forme* σ*-sesqui-linéaire non dégénérée et vérifiant*

$$x \perp y \Leftrightarrow y \perp x \qquad (*)$$

alors on est dans un des trois cas suivant :
(i) (Forme symétrique) On a $\sigma = id$ *et* $\forall x, y \in E, B(x,y) = B(y,x)$.
(ii) (Forme anti-symétrique) On a $\sigma = id$ *et* $\forall x, y \in E, B(x,y) = -B(y,x)$.
(iii) (Forme σ*-hermitienne) On a* $\sigma \neq id$ *mais* $\sigma^2 = id$ *et il existe* $\alpha \in K^*$ *tel que, si* $B'(x,y) = \alpha B(x,y)$, *alors on a* $\forall x, y \in E, B'(x,y) = \sigma \circ B'(y,x)$.

Preuve. Si $\dim(E) = 1$, l'énoncé est trivial (et sans intérêt) ; on peut donc supposer $\dim(E) \geq 2$. Soit $x \in E \setminus \{0\}$, considérons les formes linéaires $f_x(y) = B(y,x)$ et $g_x(y) = \sigma^{-1} \circ B(x,y)$; elles ont, par hypothèse, même noyau donc sont proportionnelles, c'est-à-dire qu'il existe $\alpha(x) \in K^*$ tel que $f_x = \alpha(x) g_x$ ou encore $B(y,x) = \alpha(x)\sigma^{-1} \circ B(x,y)$. Montrons d'abord que α ne dépend pas de x. Considérons pour cela l'application $i : E \to E^*$ donnée par $x \mapsto f_x$ qui est σ-linéaire et bijective et l'application $j : E \to E^*$ donnée par $x \mapsto g_x$ qui est σ^{-1}-linéaire et bijective. Introduisons $h = j^{-1} \circ i : E \to E$, alors comme $i(x) = \alpha(x) j(x)$ on a $h(x) = j^{-1}(\alpha(x) j(x)) = \sigma(\alpha(x)) j^{-1} \circ j(x) = \sigma(\alpha(x))x$. Ainsi $h(x) = \lambda(x)x$ (en posant $\lambda(x) = \sigma(\alpha(x))$) ; de plus l'application h est σ^2-linéaire. Si x, y sont non colinéaires, on a d'une part $h(x+y) = \lambda(x+y)y + \lambda(x+y)x$ d'autre part $h(x+y) = h(x) + h(y) = \lambda(x)x + \lambda(y)y$ donc $\lambda(x) = \lambda(x+y) = \lambda(y)$. Si enfin x et y sont colinéaires, on peut choisir z non colinéaire avec x, y (car $\dim(E) \geq 2$) donc $\lambda(x) = \lambda(z) = \lambda(y)$ et ainsi $\alpha(x) = \alpha(y) = \alpha$. On voit ainsi que h est linéaire (donc $\sigma^2 = id$) et que $B(y,x) = \alpha\sigma^{-1} \circ B(x,y)$. Supposons d'abord $\sigma = id$ alors $B(y,x) = \alpha B(x,y) = \alpha^2 B(y,x)$ donc $\alpha^2 = 1$ ou encore $\alpha = \pm 1$, ce qui donne les deux premiers cas (i) et (ii). Supposons maintenant $\sigma \neq id$ et commençons par montrer qu'il existe $x_0 \in E$ tel que $B(x_0, x_0) = \beta \neq 0$. En effet sinon on aurait pour tout $x, y \in E$ l'égalité $0 = B(x+y, x+y) = B(x,x) + B(x,y) + B(y,x) + B(y,y) = B(x,y) + B(y,x)$ donc B antisymétrique et bilinéaire contredisant $\sigma \neq id$. Posons alors $B' = \beta^{-1} B$, remarquons que $\beta = B(x_0,x_0) = \alpha\sigma^{-1} \circ B(x_0,x_0) = \alpha\sigma^{-1}(\beta)$ et calculons

$$B'(y,x) = \beta^{-1}(B(y,x) = \beta^{-1}\alpha\sigma^{-1} \circ B(x,y) = \beta^{-1}\alpha\sigma^{-1}(\beta B'(x,y)) = \sigma^{-1}(B'(x,y)),$$

ce qui prouve bien que B' est σ-hermitienne. □

Si la caractéristique de K est deux, alors $+1 = -1$ et, par convention, on considèrera que si $B(x,y) = B(y,x)$ la forme est anti-symétrique (on exclut donc le cas (i)) en caractéristique 2). Dans le cas (i), la forme B est associée à une forme quadratique $Q(x) := B(x,x)$ et on parle de géométrie orthogonale ; le groupe $\{f \in \mathrm{GL}(E) \mid \forall x, y \in E, B(f(x), f(y)) = B(x,y)\}$ s'appelle le *groupe orthogonal* de la forme Q (ou B) et se note $\mathrm{O}(E,Q)$ ou $\mathrm{O}(Q)$. Le sous-groupe $\mathrm{O}(E,Q) \cap \mathrm{SL}(E)$ se note $\mathrm{SO}(Q)$; il est d'indice deux dans $\mathrm{O}(Q)$. Dans le cas (iii), la forme B est associée à une forme hermitienne $H(x) := B(x,x)$ et on parle de géométrie unitaire ; le groupe $\{f \in \mathrm{GL}(E) \mid \forall x, y \in E, B(f(x), f(y)) = B(x,y)\}$ s'appelle le *groupe unitaire* de la forme H (ou B) et se note $\mathrm{U}(E,H)$ ou $\mathrm{U}(H)$. Le sous-groupe $\mathrm{U}(E,H) \cap \mathrm{SL}(E)$ se note $\mathrm{SU}(H)$. Dans le cas (ii), on parle de géométrie symplectique ; le groupe $\{f \in \mathrm{GL}(E) \mid \forall x, y \in E, B(f(x), f(y)) = B(x,y)\}$ s'appelle le *groupe symplectique* de la forme B et se note $\mathrm{Sp}(E,B)$ ou $\mathrm{Sp}(B)$. On verra que $\mathrm{Sp}(B) \subset \mathrm{SL}(E)$.

Remarque. Si B est symétrique et on pose $Q(x) = B(x,x)$ on voit facilement que

$$B(x,y) = \frac{1}{2}\left(Q(x+y) - Q(x) - Q(y)\right)$$

et donc la donnée de la forme bilinéaire équivaut à la donnée de la forme quadratique Q. La même remarque vaut en fait quand on compare une forme σ-hermitienne B et $H(x) := B(x,x)$.

On peut classifier les involutions non triviales d'un corps K à l'aide de ses sous-extensions quadratiques. Si la caractéristique de K n'est pas 2, cette classification est donnée ci-dessous, voir les exercices pour le cas de caractéristique 2.

Proposition. *Soit* K *un corps de caractéristique* $\neq 2$ *et* σ *une involution non triviale de* K, *alors* $K_0 := \{x \in K \mid \sigma(x) = x\}$ *est un sous-corps avec* $[K : K_0] = 2$ *et il existe* $\alpha \in K \setminus K_0$ *tel que* $\alpha^2 = d \in K_0$ *et alors* σ *est donnée par* $\sigma(a + b\alpha) = a - b\alpha$ *(lorsque* $a, b \in K_0$*). Inversement, toute extension* K/K_0 *de degré 2 correspond à une telle involution.*

Preuve. Soit $\sigma \neq id$ une involution de K, il est immédiat de vérifier que $K_0 := \{x \in K \mid \sigma(x) = x\}$ est un sous-corps. Si $x \in K \setminus K_0$ alors $\sigma(x) + x$ et $\sigma(x)x$ sont dans K_0 donc le polynôme $X^2 - (\sigma(x) + x)X + \sigma(x)x$

est dans $K_0[X]$ et annule x et $[K_0(x) : K_0] = 2$. L'élément $\alpha := x - \sigma(x)$ vérifie $\sigma(\alpha) = -\alpha$ donc $\alpha \notin K_0$ et $K_0(\alpha) = K_0(x)$. Par ailleurs si $y \in K \setminus K_0$ et $\beta = y - \sigma(y)$ alors $\sigma(\beta) = -\beta$ et $K_0(\beta) = K_0(y)$, donc $\sigma(\alpha/\beta) = \alpha/\beta$ donc $\alpha/\beta \in K_0$ donc $\beta \in K_0(\alpha)$ et ainsi $K = K_0(\alpha)$. Inversement si $[K : K_0] = 2$, soit $x \in K \setminus K_0$, alors $K = K_0(x)$ et x est racine de $X^2 + aX + b \in K_0[X]$. Posons $\alpha = x + a/2$ (c'est ici que l'on doit supposer $\mathrm{car}(K) \neq 2$) alors α est racine de $X^2 - d = 0$ avec $d = (a^2 - 4b)/4$. On vérifie alors directement que la formule $\sigma(a + b\alpha) = a - b\alpha$ définit un automorphisme involutif de K tel que K_0 soit le sous-corps fixé. □

Exercice. Montrer qu'on peut reconstruire un produit hermitien à partir de la forme hermitienne en montrant que, si $\sigma(\alpha) = -\alpha$ alors

$$B(x,y) = \frac{1}{4}\left\{H(x+y) - H(x-y) - \frac{1}{\alpha}H(x+\alpha y) + \frac{1}{\alpha}H(x-\alpha y)\right\}.$$

Si $F \subset E$ on définit l'*orthogonal* de F comme $F^\perp = \{x \in E \mid \forall y \in F, B(x,y) = 0\}$ (noter que cette définition n'est vraiment raisonnable que si B est symétrique, anti-symétrique ou hermitienne).

Lemme. $\dim(F) + \dim(F^\perp) = \dim(E)$ *et par conséquent, si* $F \cap F^\perp = \{0\}$ *alors* $E = F \oplus F^\perp$.

Preuve. Soit $e_1, \ldots, e_r$ une base de F et $\Phi : E \to K^r$ définie par $\Phi(x) = (B(x,e_1), \ldots, B(x,e_r))$. On a $\mathrm{Ker}(\Phi) = F^\perp$ et Φ est surjective car sinon l'image serait contenue dans un hyperplan et l'on aurait une équation du type $\forall x \in E$, $\sum_{i=1}^r \lambda_i B(x,e_i) = B\left(x, \sum_{i=1}^r \lambda_i' e_i\right) = 0$ ce qui entraînerait $\sum_{i=1}^r \lambda_i' e_i = 0$, contredisant l'indépendance linéaire des e_i. On obtient bien $\dim(E) = \dim \mathrm{Im}(\Phi) + \dim \mathrm{Ker}(\Phi) = \dim(F) + \dim(F^\perp)$. □

Application. (Décomposition en somme orthogonale.)

1er exemple. Supposons B symétrique ou σ-hermitienne. On dit que x est *isotrope* si $B(x,x) = 0$. Si x non isotrope, alors $E = \langle x \rangle \oplus \langle x \rangle^\perp$. De plus il est clair que si B est non dégénérée, il existe un vecteur non isotrope; on voit donc, par récurrence sur la dimension, qu'il existe une base orthogonale ou encore une base $e_1, \ldots, e_n$ telle que $E = \langle e_1 \rangle \oplus \ldots \oplus \langle e_n \rangle$ soit une décomposition orthogonale (i.e. avec $B(e_i, e_j) = 0$ pour $i \neq j$). Dans une base convenable, la forme quadratique $Q(x) := B(x,x)$ (resp. la forme hermitienne $H(x) = B(x,x)$) s'écrit donc $Q(x) = a_1 x_1^2 + \ldots + a_n x_n^2$ avec $a_i \in K$ (resp. $H(x) = a_1 x_1 x_1^\sigma + \ldots + a_n x_n x_n^\sigma$ avec $a_i \in K_0$).

Si $K = \mathbf{C}$ (ou plus généralement K est algébriquement clos) on peut écrire $a_i = b_i^2$ et choisir une base où $Q(x) = x_1^2 + \ldots + x_n^2$. Si $K = \mathbf{R}$ on peut écrire $a_i = \pm b_i^2$ et choisir une base où $Q(x) = x_1^2 + \ldots + x_r^2 - x_{r+1}^2 - \ldots - x_n^2$. Si $K = \mathbf{C}$ et σ est la conjugaison complexe on peut trouver une base telle que $H(x) = x_1 x_1^\sigma + \ldots + x_r x_r^\sigma - x_{r+1} x_{r+1}^\sigma - \ldots - x_n x_n^\sigma$.

2ème exemple. Définissons un *plan hyperbolique* comme un espace P de dimension 2 possédant une base e_1, e_2 avec $B(e_1,e_2) = 1$ et $B(e_1,e_1) = B(e_2,e_2) = 0$. Remarquons que si $x \in E$ est isotrope, on peut trouver un deuxième vecteur y tel que $\langle x, y \rangle$ soit un plan hyperbolique. En effet, on commence par prendre x' tel que $B(x,x') \neq 0$, en remplaçant x' par un multiple, on se ramène à $B(x,x') = 1$; on calcule alors $B(x, ax + x') = B(x,x') = 1$ et on vérifie que $B(ax + x', ax + x') = 2a + B(x',x')$ donc $y = -\frac{1}{2}B(x',x')x + x'$ convient (si la caractéristique est 2, on a $B(x',x') = 0$ et $y = x'$). De plus on a clairement $P \cap P^\perp = \{0\}$ et donc en itérant le procédé, on voit que tout espace E muni d'une forme non dégénérée peut se décomposer sous la forme :

$$E = P_1 \perp \ldots \perp P_m \perp F$$

avec P_i plan hyperbolique et F sous-espace sur lequel la forme n'a aucun vecteur isotrope non nul. On remarquera que m est un invariant de la forme quadratique (i. e. une autre décomposition fera apparaître le même nombre de plans hyperboliques).

3ème exemple. Si $B(x,x) = B(y,y) = 0$ et $B(x,y) \neq 0$ alors x, y engendre un plan hyperbolique Π et $E = \Pi \oplus \Pi^\perp$. Si maintenant B est antisymétrique non dégénérée, alors E contient nécessairement un tel plan hyperbolique et en répétant le procédé, on voit que E sera somme orthogonale de plans hyperboliques. En

particulier la dimension de E est paire et il existe une base dite *symplectique* $e_1, \ldots, e_{2n}$ telle que $B(e_i, e_j) = 0$ sauf $B(e_i, e_{i+n}) = -B(e_{i+n}, e_i) = 1$ pour $1 \leq i \leq n$.

Si l'on identifie K^n et E, via une base $e_1, \ldots, e_n$ de E, puis $\mathrm{GL}(E)$ et $\mathrm{GL}(n, K)$ l'interprétation matricielle d'une forme σ-sesquilinéaire est donnée par une matrice carrée A dont les coefficients sont les $B(e_i, e_j)$:

$$\forall X, Y \in K^n, B(X, Y) = {}^tXA\sigma(Y).$$

On vérifie aisément que B est symétrique si et seulement si ${}^tA = A$, anti-symétrique si et seulement si ${}^tA = -A$, et σ-hermitienne si et seulement si ${}^tA = \sigma(A)$. Enfin si A est la matrice associée à la forme B le groupe orthogonal, symplectique ou unitaire s'écrit comme le groupe

$$G = \{M \in \mathrm{Mat}(n \times n; K) \mid {}^tMA\sigma(M) = A\}.$$

Dans le cas où la forme est symétrique ou hermitienne, on peut se ramener à une matrice A diagonale à coefficients dans K si la forme est bilinéaire, à coefficients dans $K_0 = \{x \in K \mid \sigma(x) = x\}$ si la forme est σ-hermitienne. Enfin si la forme est antisymétrique, on peut se ramener à $A = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$ et écrire le groupe symplectique

$$\mathrm{Sp}(2n, K) = \left\{ M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \mid {}^tAC \text{ et } {}^tBD \text{ sont symétriques et } {}^tAD - {}^tCB = I \right\}.$$

D'un point de vue géométrique, d'autres groupes sont naturels à considérer, notamment celui des similitudes et isométries affines, ainsi que les groupes résultant de l'action sur l'espace projectif; définissons-les brièvement.

Définition Soit $B : E \times E \to K$ une forme sesqui-linéaire, on appelle *similitude* un automorphisme f de E tel qu'il existe $\mu \in K^*$ tel que

$$\forall x, y \in E, \qquad B(f(x), f(y)) = \mu B(x, y),$$

l'élément μ s'appelle le *multiplicateur* de f.

Remarque. Le groupe des multiplicateurs contient les carrés (resp. les éléments "normes" de la forme $a\sigma(a)$) si B est symétrique ou antisymétrique (resp. hermitienne). Le groupe des similitudes contient évidemment les homothéties et les isométries. En fait, le groupe des multiplicateurs est exactement celui des carrés (resp. des normes) si et seulement si les similitudes sont les produits d'une isométrie par une homothétie. Si n est impair on vérifie facilement que le sous-groupe des multiplicateurs est égal au groupe des carrés (cas orthogonal) ou au groupe des normes $x\sigma(x)$ (cas hermitien) et donc une similitude est produit d'une isométrie par une homothétie. Quand la dimension est paire, ce n'est pas toujours le cas comme on peut le vérifier élémentairement dans le cas d'un plan hyperbolique.

Exercice. Les similitudes sont les seules applications linéaires $f : E \to E$ préservant la relation d'orthogonalité (i. e. telle que $x \perp y \Leftrightarrow f(x) \perp f(y)$).

Rappelons que si E est un K-espace vectoriel (de dimension $n + 1$), l'espace projectif $\mathbf{P}(E)$ correspondant est défini comme l'ensemble des droites vectorielles de E ou encore comme le quotient de $E \setminus \{0\}$ par la relation de colinéarité. Si $E = K^{n+1}$ on peut noter $\mathbf{P}(E) = \mathbf{P}^n(K)$. Le groupe $\mathrm{GL}(E) = \mathrm{GL}_{n+1}(K)$ possède une action naturelle sur l'espace projectif $\mathbf{P}(E) = \mathbf{P}^n(K)$ définie par $(f, D) \mapsto f(D)$. Le noyau de cette action est aussi le centre de $\mathrm{GL}_{n+1}(K)$, i.e. l'ensemble des homothéties. Ceci justifie les notations/définitions suivantes.

Définition Le quotient du groupe $\mathrm{GL}(E)$ par les homothéties de rapport $\alpha \in K^*$ s'appelle le *groupe projectif linéaire* ; il se note $\mathrm{PGL}(E)$. De manière générale, on note $\mathrm{PSL}(E)$ (resp. $\mathrm{PSO}(E, B)$, $\mathrm{PSp}(E, B)$, $\mathrm{PSU}(E, B)$, etc.) l'image dans $\mathrm{PGL}(E)$ de $\mathrm{SL}(E)$ (resp. $\mathrm{SO}(E, B)$, $\mathrm{Sp}(E, B)$, $\mathrm{SU}(E, B)$, etc.)

Remarque On peut aussi introduire l'espace des sous-espaces vectoriels de dimension r dans E (un K-espace vectoriel de dimension $n+1$). Celui-ci s'appelle *Grasmannienne* et est muni d'une action transitive de $\mathrm{GL}_{n+1}(K)$ ou encore de $\mathrm{PGL}_{n+1}(K)$, définie par $(\sigma, F) \mapsto \sigma(F)$.

Nous allons maintenant étudier brièvement les groupes orthogonaux, symplectiques et unitaires. Mais avant d'étudier ces groupes classiques donnons quelques raisonnements généraux concernant les groupes. En particulier, nous allons introduire un argument assez général dû essentiellement à Iwasawa qui permet de démontrer la simplicité de quelques groupes.

Rappelons tout d'abord deux notions.

(i) Une action $\rho : G \to \mathrm{Bij}(X)$ d'un groupe G sur X est *doublement transitive* si pour tout $x_1 \neq x_2$, $y_1 \neq y_2$ il existe $g \in G$ tel que $g \cdot x_1 = y_1$ et $g \cdot x_2 = y_2$ (on pourrait définir de même la notion d'action n-transitive).

(ii) Le sous-groupe des commutateurs $D(G)$ d'un groupe G est le sous-groupe engendré par les commutateurs $[x, y] = xyx^{-1}y^{-1}$. C'est un sous-groupe distingué et c'est le plus petit sous-groupe distingué de G tel que le quotient soit abélien.

Proposition *Soit $\rho : G \to \mathrm{Bij}(X)$ une action doublement transitive; supposons qu'il existe des sous-groupes abéliens $\{A_x\}_{x\in X}$ dont la réunion engendre G et tels que $gA_xg^{-1} = A_{g\cdot x}$. Si N est un sous-groupe distingué de G alors ou bien $N \subset \mathrm{Ker}(\rho)$, i.e. N agit trivialement, ou bien N agit transitivement et contient $D(G)$. En particulier, si de plus $G = D(G)$ alors le groupe $G/\mathrm{Ker}(\rho)$ est simple.*

Preuve. Soit $x \in X$ et $H := G_x$, la double transitivité se traduit par le fait que bien sûr G agit transitivement sur X mais aussi que H agit transitivement sur $X \setminus \{x\}$. On en tire en particulier que, dès que $g \notin H$, on a $G = H \cup HgH$ et en particulier que H est un sous-groupe (propre) maximal. Si maintenant $N \triangleleft G$ alors $N' = NH$ est un sous-groupe donc est égal soit à G (si $N \not\subset H$) soit à H (si $N \subset H$). Dans le premier cas, d'après ce qui précède, l'action de N est transitive, dans le second cas on a $N = gNg^{-1} \subset gHg^{-1} = G_{g\cdot x}$ donc N agit trivialement. Si N agit transitivement, il suffit de voir que $NA_x = G$ car alors $G/N = NA_x/N \cong A_x/A_x \cap N$ est abélien et donc $D(G) \subset N$; il suffit donc de montrer que tous les sous-groupes A_y sont contenus dans NA_x. Mais NA_x est aussi le sous-groupe engendré par N et A_x; soit $y \in X$ alors il existe $n \in N$ tel que $n \cdot x = y$ donc $A_y = nA_xn^{-1}$ est contenu dans NA_x. Si de plus $G = D(G)$, soit $s : G \to G/\mathrm{Ker}(\rho)$ la surjection canonique et $\{e\} \neq M \triangleleft G/\mathrm{Ker}(\rho)$; considérons $N := s^{-1}(M)$, il n'agit pas trivialement sur X et il est distingué dans G donc la démonstration précédente montre que $D(G)$ est contenu dans N et donc que $N = G$ et finalement que $M = G/\mathrm{Ker}(\rho)$. □

La proposition suivante peut être utile pour calculer le groupe $D(G)$.

Proposition. *Soit G un groupe, notons G^2 le sous-groupe engendré par les carrés d'éléments de G.*

(i) Le sous-groupe G^2 est distingué et contient $D(G)$.

(ii) Si G est engendré par des éléments d'ordre 2, alors $G^2 = D(G)$.

(iii) Si G est engendré par des éléments d'ordre 2 tous conjugués, alors $(G : D(G)) \leq 2$.

Preuve. Comme $yx^2y^{-1} = (xyx^{-1})^2$, on voit que G^2 est bien distingué dans G; de plus le quotient G/G^2 est d'exposant 2 donc abélien [en effet $(ab)^2 = abab = e$ entraîne $ab = ba$] et donc $D(G) \subset G^2$. Si $x_1, \ldots, x_m$ sont des éléments d'ordre 2 alors $(x_1 \ldots x_m)^2 = x_1 \ldots x_m x_1^{-1} \ldots x_m^{-1}$ est un produit de commutateurs donc si tout élément de G est de la forme $x_1 \ldots x_m$, on a bien $G^2 = D(G)$. Enfin si on note $\bar{x}$ l'image de x dans $G/D(G)$, on a $\overline{yxy^{-1}} = \bar{x}$ et donc, sous les hypothèses de (iii), l'image d'un des éléments d'ordre 2 engendre $G/D(G)$. □

E.2. Les groupes $\mathrm{GL}(n, K)$ et $\mathrm{SL}(n, K)$.

Les relations entre les groupes $\mathrm{GL}_n(k)$, $\mathrm{SL}_n(k)$, $\mathrm{PGL}_n(k)$ et $\mathrm{PSL}_n(k)$ peuvent être décrites par le diagramme suivant où les lignes et colonnes sont exactes et où la flèche $\mathrm{GL}_n(k) \to k^*$ dans la ligne centrale est le

déterminant et où on note $\mu_n(k) := \{x \in k^* \mid x^n = 1\}$ et $k^{*n} := \{x^n \mid x \in k^*\}$.

$$\begin{array}{ccccccccc}
 & & 0 & & 0 & & 0 & & \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \to & \mu_n(k) & \to & k^* & \to & k^{*n} & \to & 0 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \to & \mathrm{SL}_n(k) & \to & \mathrm{GL}_n(k) & \to & k^* & \to & 0 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \to & \mathrm{PSL}_n(k) & \to & \mathrm{PGL}_n(k) & \to & k^*/k^{*n} & \to & 0 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
 & & 0 & & 0 & & 0 & &
\end{array}$$

Si $k = \mathbf{F}_q$ on obtient aisément, en comptant les bases de $E = \mathbf{F}_q^n$:

$$\mathrm{card}\left(\mathrm{GL}_n(\mathbf{F}_q)\right) = (q^n - 1)\,(q^n - q)\ldots\left(q^n - q^{n-1}\right)$$

et on peut en déduire le cardinal des autres groupes $\mathrm{SL}_n(k)$, $\mathrm{PGL}_n(k)$ et $\mathrm{PSL}_n(k)$.

On va maintenant décrire des transformations donnant un ensemble très utile de générateurs de $\mathrm{SL}_n(k)$.

Définition. Une *transvection* est un élément de $\mathrm{SL}_n(k) \setminus \{Id\}$ laissant fixe un hyperplan.

On voit aisément qu'une telle application est nécessairement de la forme $u(x) = x + f(x)a$ avec a vecteur non nul et f forme linéaire non nulle contenant a dans son noyau. La droite engendré par a est caractérisée par u et on dira que u est une transvection de droite $\langle a\rangle$. L'ensemble des transvections de direction une droite forment un sous-groupe isomorphe à k^{n-1}. Si $\sigma = t(a, f)$ désigne la transvection telle que $\sigma(x) = x + f(x)a$ alors, pour $\rho \in \mathrm{GL}(E)$, on a : $\rho \circ t(a, f) \circ \rho^{-1} = t(\rho(a), f \circ \rho^{-1})$. Enfin, dans une base convenable la matrice d'une transvection est une matrice avec des 1 sur la diagonale et un unique coefficient non nul au dessus de la diagonale et que l'on peut prendre égal à 1. En particulier toutes les transvections sont conjuguées dans $\mathrm{GL}_n(k)$; on voit aisément que cela reste vrai dans $\mathrm{SL}_n(k)$ lorsque $n \geq 3$ (ce dernier point n'est plus vrai dans $\mathrm{SL}_2(k)$).

Thorme. *Le centre de* $\mathrm{GL}_n(k)$ *est le sous-groupe des homothéties que l'on peut identifier à* k^**; Le centre de* $\mathrm{SL}_n(k)$ *est le sous-groupe des homothéties de rapport une racine n-ième, que l'on peut identifier à* $\mu_n(k)$.

Preuve. Une matrice commutant avec la transvection $t(a, f)$ doit laisser stable la droite engendrée par a. On a vu qu'une telle application doit être une homothétie. Enfin l'homothétie de rapport λ est dans $\mathrm{SL}_n(k)$ si et seulement si $\lambda^n = 1$. □

Thorme. *Les transvections engendrent* $\mathrm{SL}_n(k)$.

Preuve. On prouve d'abord que si $x, y \in E$ non nuls, il existe un produit de transvections u tel que $u(x) = y$. Si x et y ne sont pas colinéaires, on choisi $a = y - x$ et f forme linéaire nulle sur a mais pas sur x (ni sur y donc); quitte à multiplier f par un scalaire on peut s'assurer que $f(x) = 1$. On a alors $t(a, f)(x) = x + f(x)(y - x) = y$. Si x et y sont colinéaires, on passe par un troisième vecteur z non colinéaire et deux transvections telles que $u_2(x) = z$ et $u_1(z) = y$.

Si maintenant $v \in \mathrm{SL}_n(k)$ et $x \in E$ non nul, il existe u produit de transvections tel que $v \circ u(x) = x$. Si la dimension est 2, on en déduit que $v \circ u$ est une transvection, sinon on procède par récurrence sur la dimension. L'application $v \circ u$ induit une application sur $\bar{E} = E/\langle x\rangle$ qui est encore de déterminant 1 et peut donc, par hypothèse de récurrence, s'écrire comme produit de transvections de $\bar{E}$. En relevant les transvections de $\bar{E}$ en des transvections de E on conclut (les détails sont laissés au lecteur). □

Thorme. *Le groupe des commutateurs de* $\mathrm{GL}_n(k)$ *est* $D(\mathrm{GL}_n(k)) = \mathrm{SL}_n(k)$ *sauf pour* $n = 2$ *et* $k = \mathbf{F}_2$*. Le groupe des commutateurs de* $\mathrm{SL}_n(k)$ *est* $D(\mathrm{SL}_n(k)) = \mathrm{SL}_n(k)$ *sauf pour* $n = 2$ *et* $k = \mathbf{F}_2$ *ou* $\mathbf{F}_3$.

Preuve. On a clairement $D(\mathrm{SL}_n(k)) \subset D(\mathrm{GL}_n(k)) \subset \mathrm{SL}_n(k)$. Soit σ une transvection. Si $\mathrm{car}(k) \neq 2$, alors σ^2 est encore une transvection (de même droite) donc s'écrit $\sigma^2 = \rho\sigma\rho^{-1}$ (si $n \geq 3$ on peut même choisir

$\rho \in \mathrm{SL}_n(k)$). Ainsi $\sigma = \sigma^2\sigma^{-1} = \rho\sigma\rho^{-1}\sigma^{-1}$ est un commutateur de $\mathrm{GL}_n(k)$ (et même un commutateur de $\mathrm{SL}_n(k)$ si $n \geq 3$). Pour examiner les cas $n = 2$ ou $\mathrm{car}(k) = 2$, il suffit essentiellement d'observer que si $k \neq \mathbf{F}_2$ ou $\mathbf{F}_3$ il y a des matrices diagonales dans $\mathrm{SL}_2(k)$ qui ne sont pas des homothéties, tout simplement les matrices $\rho = \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix}$ avec $a \in k \setminus \{0, 1, -1\}$. Si l'on choisit $\sigma = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ et on observe que $\rho\sigma\rho^{-1}\sigma^{-1} = \begin{pmatrix} 1 & a^2 - 1 \\ 0 & 1 \end{pmatrix}$ on voit que $D(\mathrm{SL}_2(k))$ contient une et donc toutes les transvections et donc tout $\mathrm{SL}_2(k)$). □

Thorme. *Soit k un corps (commutatif) et $n \geq 2$ alors le groupe $\mathrm{PSL}_n(k)$ est simple sauf pour $n = 2$ et $k = \mathbf{F}_2$ ou $\mathbf{F}_3$.*

Preuve. On va utiliser, pour varier, la méthode d'Iwasawa cité en introduction à ce chapitre. Soit E un k-espace vectoriel de dimension $n \geq 2$. Considérons l'action de $G = \mathrm{SL}(E)$ sur $X = \mathbf{P}(E)$, un élément $x \in X$ peut être vu comme une droite vectorielle de E; considérons A_x le sous-groupe des transvections de droite x. Il est immédiat de vérifier que $A_x \cong k^{n-1}$ est commutatif et on vérifie bien $gA_xg^{-1} = A_{g\cdot x}$: en effet si a est un vecteur non nul de direction x, tout élément $u \in A_x$ s'écrit $u(y) = y + f(y)a$ avec f forme linéaire nulle en a. Appelons donc $E_x^* = \{f \in E^* \mid f(a) = 0\}$ alors $f \mapsto u : y \to y + f(y)a$ définit un isomorphisme de groupes de E_x^* vers A_x et comme $vuv^{-1}(y) = y + f(v^{-1}(y))v(a)$, on a bien $vA_xv^{-1} = A_{v(x)}$. Par ailleurs, on sait que les transvections engendrent $\mathrm{SL}(E)$. On a vu que $D(\mathrm{SL}(E)) = \mathrm{SL}(E)$ sauf si $n = 2$ et $k = \mathbf{F}_2$ ou $\mathbf{F}_3$; on en déduit donc le théorème. □

Remarque. Si $k = \mathbf{F}_q$, $n = \dim(E)$, notons $d = \mathrm{PGCD}(n, q-1)$, alors

$$\mathrm{card}\,\mathrm{PSL}_n(\mathbf{F}_q) = \frac{\prod_{i=1}^{n}(q^n - q^{n-i})}{d(q-1)}$$

Remarque. Comme $\mathrm{PSL}_2(\mathbf{F}_2) \cong \mathcal{S}_3$ et $\mathrm{PSL}_2(\mathbf{F}_3) \cong \mathcal{A}_4$ ne sont pas égaux à leurs sous-groupes de commutateurs, ils ne sont pas simples. En considérant l'action de $\mathrm{PSL}_2(\mathbf{F}_4)$ sur $\mathbf{P}^1(\mathbf{F}_4)$ de cardinal 5, on voit que $\rho : \mathrm{PSL}_2(\mathbf{F}_4) \to \mathcal{S}_5$ induit un isomorphisme $\rho' : \mathrm{PSL}_2(\mathbf{F}_4) \to \mathcal{A}_5$.

E.3. Groupe orthogonal.

Commençons par décrire entièrement le cas de la dimension 2 qui est particulier.

Proposition. *Soit $Q(x_1, x_2) = x_1^2 + Dx_2^2$ avec $D \in K^*$, alors le groupe des isométries directes s'écrit*

$$\mathrm{SO}(Q) = \left\{ \begin{pmatrix} a & -cD \\ c & a \end{pmatrix} \mid a^2 + Dc^2 = 1 \right\}$$

Si $-D$ est un carré dans K alors $\mathrm{SO}(Q) \cong K^*$; si $-D$ n'est pas un carré dans K alors $\mathrm{SO}(Q) \cong \{x \in K(\sqrt{-D})^* \mid N_K^{K(\sqrt{-D})}(x) = 1\}$; en particulier le groupe $\mathrm{SO}(Q)$ est abélien. Les éléments de $\mathrm{O}^-(Q)$ sont des symétries par rapport à une droite et, si $s \in \mathrm{O}^-(Q)$ et $r \in \mathrm{SO}(Q)$, alors $srs^{-1} = r^{-1}$.

Preuve. Un calcul direct montre que $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ est dans $\mathrm{O}(Q)$ si et seulement si $a^2 + Dc^2 = 1$, $b^2 + Dd^2 = D$ et $ab + Dcd = 0$. Si $c = 0$ on voit que $b = 0$ et donc $a^2 = d^2 = 1$. Si $c \neq 0$, on en tire $d = -ab/cD$ puis $b^2 = c^2D^2$ ou encore $b = \epsilon cD$ (avec $\epsilon = \pm 1$), puis $d = -\epsilon a$ et $\det(M) = -\epsilon$. Si la forme est isotrope, on peut en fait se ramener à une matrice $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ et on voit alors aisément que $\mathrm{SO}(Q) = \left\{ \begin{pmatrix} a & 1 \\ 0 & a^{-1} \end{pmatrix} \mid a \in k^* \right\}$ est isomorphe à k^*. Si la forme est anisotrope, posons $\omega = \sqrt{-D}$, on a un homomorphisme $k(\omega)^* \to \mathrm{GL}_2(k)$ défini par $a + c\omega \mapsto \begin{pmatrix} a & -cD \\ c & a \end{pmatrix}$ qui induit l'isomorphisme annoncé entre $\{a + c\omega \mid a^2 + Dc^2 = 1\}$ et $\mathrm{SO}(Q)$. Les dernières affirmations se vérifient directement. □

Remarque. Le groupe $\mathrm{SO}(Q)$ est donc commutatif si E est un plan ($\dim(E) = 2$). On obtient ainsi une généralisation de la notion d'angle : si le plan contient un vecteur isotrope (i. e. si $-D$ est un carré) alors un "*angle*" est donné par un élément $\alpha \in K^*$; si le plan ne contient pas de vecteur isotrope (i. e. si $-D$ n'est pas un carré) alors un "*angle*" est donné par un élément $\alpha \in K_1 = \{x \in K(\sqrt{-D})^* \mid N_K^{K(\sqrt{-D})}(x) = 1\}$. Par exemple, si $K = \mathbf{R}$ on retrouve que dans le cas de la géométrie hyperbolique, un angle est donné par un réel non nul et dans le cas de la géométrie euclidienne un angle est donné par un complexe α de module 1. Dans le dernier cas on a un homomorphisme surjectif de $K = \mathbf{R}$ vers K_1 donné par $t \mapsto \exp(2\pi i t)$. On retrouve donc l'expression des rotations sous la forme

$$R(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

Cherchons maintenant les isométries orthogonales qui fixent un hyperplan.

Lemme. *Soit $H = \langle x \rangle^\perp$ un hyperplan de E et supposons que σ est une isométrie fixant H point par point, alors*
(i) Si $Q(x) = 0$ alors $\sigma = id_E$.
(ii) Si $Q(x) \neq 0$ alors ou bien $\sigma = id_E$ ou bien σ est la symétrie hyperplane définie par

$$\sigma(y) = y - 2\frac{B(x,y)}{B(x,x)}x.$$

Preuve. Dans le cas (ii) on a $E = \langle x \rangle \perp H$ et comme $\sigma(H) = H$ on en tire que $\sigma(x) \in \langle x \rangle$ donc il existe $\lambda \in K$ tel que $\sigma(x) = \lambda x$. Comme σ est une isométrie, on a $\lambda = \pm 1$. Si $\lambda = 1$ alors $\sigma = id$ et si $\lambda = -1$ alors σ est la symétrie par rapport à l'hyperplan H donné par la formule annoncée. En effet tout vecteur $y \in E$ se décompose en $y = (y - (B(y,x)/B(x,x))x) + (B(y,x)/B(x,x))x$ avec $y - (B(y,x)/B(x,x))x \in \langle x \rangle^\perp = H$ donc

$$\sigma(y) = (y - (B(y,x)/B(x,x))x) - (B(y,x)/B(x,x))x = y - 2(B(x,y)/B(x,x))x.$$

Dans le cas (i), on a $x \in H$ et il existe $z \in E$ tel que $B(x,z) = 1$. On a alors $\langle x, z \rangle^\perp = H_0 \subset H$ et $E = \langle x, z \rangle \perp H_0$. On sait que $\sigma(H_0) = H_0$ donc $\sigma(z) = \lambda x + \mu z$; on sait aussi que $\sigma(x) = x$. On a donc $1 = B(z,x) = B(\sigma(z), \sigma(x)) = \lambda B(x,x) + \mu B(z,x) = \mu$ d'où $\mu = 1$, ainsi que $B(z,z) = B(\sigma(z), \sigma(z)) = \lambda^2 B(x,x) + 2\lambda\mu B(z,x) + \mu^2 B(z,z) = 2\lambda + B(z,z)$ d'où $\lambda = 0$ et $\sigma(z) = z$ donc $\sigma = id$. □

Si $Q(x) \neq 0$, on notera s_x la symétrie hyperplane caractérisée par $\forall y \in \langle x \rangle^\perp, s_x(y) = y$ et $s_x(x) = -x$. Remarquons que $s_{ax} = s_x$ pour $a \in K^*$; en fait $s_x = s_y$ équivaut à x, y colinéaires. Par ailleurs, si ρ est une isométrie, $\rho s_x \rho^{-1} = s_{\rho(x)}$.

Thorme. *Les symétries hyperplanes s_x avec $Q(x) \neq 0$ engendrent $\mathrm{O}(Q)$.*

Preuve. On raisonne par récurrence, le résultat étant facile si $n = 1$ ou 2. Soit donc $\sigma \in \mathrm{O}(Q)$, si il existe $x \in E$ tel que $Q(x) \neq 0$ et $\sigma(x) = x$ alors on décompose $E = \langle x \rangle \oplus \langle x \rangle^\perp = \langle x \rangle \oplus H$ (disons). Alors $\sigma' := \sigma_{|H}$ s'écrit comme produit de symétries hyperplanes s'_{x_i} dans H, c'est-à-dire $\sigma' = s'_{x_1} \dots s'_{x_m}$. Notons donc s_{x_i} la symétrie dans E associée à $x_i \in H$ on a alors $\sigma = s_{x_1} \dots s_{x_m}$ puisque les deux applications sont des isométries et coïncident sur x et sur H. Soit maintenant x_1 non isotrope et $x_2 = \sigma(x_1)$, si $x_1 - x_2$ est non isotrope, alors $s_{x_1 - x_2}(x_1) = x_2$ donc $s_{x_1 - x_2} \circ \sigma$ fixe x_1 et s'écrit donc comme produit de symétries hyperplanes, donc σ également. Observons que $Q(x_1 + x_2) + Q(x_1 - x_2) = 2Q(x_1) + 2Q(x_2) = 4Q(x_1)$, donc, si $x_1 - x_2$ est isotrope, alors $x_1 + x_2$ n'est pas isotrope, et on a $s_{x_1+x_2}(x_1) = -x_2$ donc $s_{x_2} \circ s_{x_1+x_2}(x_1) = x_2$ et $s_{x_1+x_2} \circ s_{x_2} \circ \sigma$ fixe x_1 donc est produit de symétries hyperplanes, donc σ également. □

Remarque. On peut raffiner l'énoncé précédent en montrant que σ s'écrit comme le produit d'au plus $n = \dim(E)$ symétries hyperplanes (théorème de Cartan-Dieudonné).

Les symétries hyperplanes sont en quelque sorte les involutions les plus simples de $\mathrm{O}(Q)$ (sous-espace propre pour la valeur propre 1 de codimension 1) ; l'analogue pour $\mathrm{SO}(Q)$ est constitué par les *renversements*,

c'est-à-dire les isométries directes fixant un sous-espace F de codimension 2 non isotrope (i.e. $F \cap F^\perp = \{0\}$) et agissant par -1 sur son supplémentaire orthogonal ; ce sont bien sûr également des involutions.

Thorme. *Supposons* $\dim(E) \geq 3$, *alors les renversements engendrent* $\mathrm{SO}(Q)$.

Preuve. Remarquons que la conclusion de l'énoncé est fausse en général si $\dim(E) = 2$. Prouvons d'abord l'énoncé pour $n = 3$. D'après le théorème précédent, $\rho \in \mathrm{SO}(Q)$ peut s'écrire comme produit d'un nombre pair de symétries hyperplanes, il suffit donc de montrer que le produit $s_{x_1} s_{x_2}$ de deux symétries hyperplanes peut s'écrire comme le produit de deux renversements. Mais en dimension 3, $-s_x$ est un renversement et $s_{x_1} s_{x_2} = (-s_{x_1})(-s_{x_2})$, ce qui achève ce cas. Reprenons le cas général, on peut supposer x_1 et x_2 non colinéaires (sinon $s_{x_1} s_{x_2} = id$) et donc $L = \langle x_1, x_2 \rangle^\perp$ est de codimension 2 et, comme $Q(x_1)$ et $Q(x_2) \neq 0$, on a $\dim(L \cap L^\perp) \leq 1$ donc on peut choisir L_1 hyperplan de L tel que, si on note $L_2 = L_1^\perp$, on ait $E = L_1 \perp L_2$. Or $s_{x_1} s_{x_2}$ fixe L_1 point par point et agit sur L_2 comme $s'_{x_1} s'_{x_2}$ (où s'_{x_i} désigne la restriction de s_{x_i} à L_2). Mais L_2 est un espace de dimension 3 donc $s'_{x_1} s'_{x_2} = \rho'_1 \rho'_2$ avec ρ'_i renversement de L_2. Désignons par ρ_i l'isométrie agissant comme l'identité sur L_1 et comme ρ'_i sur L_2, alors ρ_i est un renversement et $s_{x_1} s_{x_2} = \rho_1 \rho_2$. □

Connaissant comme générateurs de $\mathrm{O}(Q)$ les symétries hyperplanes et comme générateurs de $\mathrm{SO}(Q)$ les renversements (lorsque $n \geq 3$), on en tire aisément que les éléments du centre sont des homothéties.

Proposition. *Le centre de* $\mathrm{O}(Q)$ *est le sous-groupe à deux éléments* $\{\pm I\}$*; le centre de* $\mathrm{SO}(Q)$ *est réduit au sous-groupe trivial si* $n \geq 3$ *est impair et égal à* $\{\pm I\}$ *si* $n \geq 4$ *est pair.*

Regardons maintenant deux cas particulièrement intéressants : celui des groupes orthogonaux réels usuels (i. e. pour la forme euclidienne) et celui des corps finis.

Thorme. *Le groupe* $\mathrm{SO}(3, \mathbf{R})$ *est simple. Plus généralement le groupe* $\mathrm{SO}(2n+1, \mathbf{R})$ *est simple alors que, pour* $n \geq 3$ *le groupe* $\mathrm{SO}(2n, \mathbf{R})$ *contient comme unique sous-groupe normal non trivial* $\{+1, -1\}$ *et donc* $\mathrm{PSO}(2n, \mathbf{R})$ *est simple.*

Preuve. On suppose d'abord $n = 3$ et on commence par un lemme géométriquement évident dont on laisse le lecteur formaliser la démonstration :

Lemme. *Soit* x_1, x_2, y_1, y_2 *des vecteurs de la sphère de* $\mathbf{R}^3$ *tels que* $||x_1 - x_2|| = ||y_1 - y_2||$*, alors il existe une rotation* $\rho \in \mathrm{SO}_3(\mathbf{R})$ *telle que* $\rho(x_1) = y_1$ *et* $\rho(x_2) = y_2$.

Soit maintenant H un sous-groupe distingué de $\mathrm{SO}_3(\mathbf{R})$ possédant un élément σ distinct de l'identité. Soit Δ l'axe de σ et e_1 un point de la sphère hors de l'axe. Posons $0 < \delta_0 = ||e_1 - \sigma(e_1)||$, lorsque x parcourt l'arc du méridien passant par e_1 et rejoignant l'axe Δ, la distance $||x - \sigma(x)||$ décroit continûment de δ_0 à 0. En particulier, si δ est assez petit ($\delta \leq \delta_0$ suffit), il existe x_1 sur la sphère tel que $||x_1 - \sigma(x_1)|| = \delta$. Soit maintenant y_1, y_2 deux points de la sphère tels que $||y_1 - y_2|| = \delta$; d'après le lemme, il existe $\rho \in \mathrm{SO}_3(\mathbf{R})$ telle que $\rho(x_1) = y_1$ et $\rho(\sigma(x_1)) = y_2$. Ainsi $\sigma' = \rho\sigma\rho^{-1} \in H$ et $\sigma'(y_1) = y_2$. En itérant ce procédé, on voit que H opère transitivement sur la sphère. En particulier, il existe $\sigma'' \in H$ telle que $\sigma''(e_1) = -e_1$, mais alors σ'' est un renversement et donc H contient tous les renversements et est donc égal à $\mathrm{SO}_3(\mathbf{R})$.

Dans le cas $n \geq 5$, on se ramène au cas de dimension 3 ainsi : chaque sous-espace F de dimension 3 induit une décomposition $\mathbf{R}^n = F \oplus F^\perp$ et une injection $\mathrm{SO}_3(\mathbf{R}) \hookrightarrow \mathrm{SO}_n(\mathbf{R})$; si H est un sous-groupe distingué contenant $\sigma \neq \pm I$, il suffit de voir que H rencontre l'un des $\mathrm{SO}_3(\mathbf{R})$ non trivialement car alors il le contiendra en entier et contiendra donc un renversement et donc tous et sera donc égal à $\mathrm{SO}_n(\mathbf{R})$ entier.

Comme $\sigma \neq \pm I$, σ doit bouger un plan, disons F. Notons ρ le renversement de plan F, alors $\sigma' = \rho\sigma\rho^{-1}\sigma^{-1}$ est dans $H \setminus \{\pm I\}$ et peut s'écrire comme le produit des deux renversements de plan F et $\sigma(F)$ donc possède un sous-espace de points fixes de dimension $\geq n - 4$. Il y a donc un point fixe disons x_1 non nul (puisque $n \geq 5$!). Soit maintenant x_2 tel que x_2 et $\sigma(x_2)$ ne soient pas colinéaires, posons $r = s_{x_2} s_{x_1}$ (produit de deux symétries hyperplanes). On a alors

$$\sigma'' := \sigma' r \sigma^{-1} r^{-1} = \sigma' s_{x_2} \sigma'^{-1} \sigma' s_{x_1} \sigma'^{-1} s_{x_1} s_{x_2} = (\sigma' s_{x_2} \sigma'^{-1}) s_{x_2}$$

est dans $H \setminus \{I\}$ et est un produit de deux symétries hyperplanes. Ainsi σ'' possède un sous-espace de points fixes de dimension $n - 2$ donc est contenu dans un $\mathrm{SO}_3(\mathbf{R})$, ce qui achève la preuve. □

Remarque. Le groupe $\mathrm{PSO}_4(\mathbf{R})$ n'est pas simple, voir le chapitre sur les quaternions (E.6) pour une preuve de ce fait et une description. De manière générale l'étude des quaternions et de leurs généralisations (algèbres de Clifford) permet d'approfondir l'étude des groupes orthogonaux (Cf ibidem); en particulier on peut ainsi élucider la structure de $\mathrm{SO}(Q)$ lorsqu'il existe au moins un vecteur isotrope non nul. Cette dernière condition est automatiquement vérifiée lorsque k est un corps fini et $n \geq 3$ comme l'indique le lemme suivant.

Lemme. *Une forme quadratique sur un espace de dimension $n \geq 3$ sur $\mathbf{F}_q$ possède un vecteur non nul isotrope.*

Preuve. On se ramène à trouver un zéro non trivial au polynôme $x^2 + ay^2 + bz^2$ avec $ab \neq 0$. Le nombre de carrés dans $\mathbf{F}_q$ est $(q+1)/2$ donc les fonctions $x^2 + a$ et $-bz^2$ ont une valeur commune au moins. □

Ainsi on voit que l'on peut toujours écrire une décomposition

$$E = P_1 \perp \ldots \perp P_m \perp F$$

avec P_i plan hyperbolique et F soit nul, soit de dimension 1, soit de dimension 2 sans vecteur isotrope non nul.

Si n est pair on note $\epsilon = +1$ si la forme quadratique est équivalente à $x_1x_2 + x_3x_4 + \ldots + x_{n-1}x_n$ et $\epsilon = -1$ si la forme quadratique est équivalente à $x_1x_2 + x_3x_4 + \ldots + x_{n-3}x_{n-2} + x_{n-1}^2 - ax_n^2$ avec $a \notin \mathbf{F}_q^{*2}$.

Thorme. *Soit G le groupe orthogonal d'une forme quadratique non dégénérée sur un espace de dimension n sur $\mathbf{F}_q$ alors*

$$\mathrm{card}(G) = \begin{cases} q^{\frac{(n-1)^2}{4}} \prod_{i=1}^{\frac{n-1}{2}} (q^{2i} - 1) & \text{si } n \text{ est impair} \\ q^{\frac{n(n-2)}{4}} (q^{\frac{n}{2}} - \epsilon) \prod_{i=1}^{\frac{n-2}{2}} (q^{2i} - 1) & \text{si } n \text{ est impair} \end{cases}$$

On donne la preuve sous forme d'exercice (ref. Artin, chapitre III, p 145-147).

Exercice. 1) Montrer que $\phi_n(Q) := \mathrm{card}\{x \in \mathbf{F}_q^n \mid Q(x) = 0\}$ vaut q^{n-1} si n impair et $(q^{n/2} - \epsilon)(q^{n/2-1} + \epsilon) + 1$ si n pair. 2) Montrer que le nombre de paires hyperboliques (i. e. de paires (e_1, e_2) telles que $Q(e_1) = Q(e_2) = 0$ et $B(e_1, e_2) = 1$) est $\lambda_n = q^{n-2}(\phi_n - 1)$. 3) Montrer que si $E = P \perp F$ avec P plan hyperbolique et si on note $Q' = Q_{|F}$ alors $\mathrm{card}\,\mathrm{SO}(Q) = \lambda_n \,\mathrm{card}\,\mathrm{SO}(Q')$.

E.4. Groupe symplectique.

Démontrons d'abord que si $B : E \times E \to K$ est anti-symétrique, non dégénérée, alors $\dim(E) = 2m$ et on peut choisir une base $(e_1, \ldots, e_m, f_1, \ldots, f_m)$ telle que $B(e_i, e_j) = B(f_i, f_j) = 0$ et $B(e_i, f_j) = \delta_i^j$, c'est-à-dire que $E = \langle e_1, f_1 \rangle \perp \ldots \perp \langle e_m, f_m \rangle$ et la matrice de B sur $\langle e_i, f_i \rangle$ est $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. En particulier, tous les formes antisymétriques non dégénérées sur E sont semblables.

En effet, soit $e_1 \neq 0$, alors $B(e_1, e_1) = 0$ et il existe $f_1 \in E$ tel que $B(e_1, f_1) = 1$ (il existe un vecteur f tel que $B(e_1, f) \neq 0$ et en le multipliant par un scalaire adéquat on obtient f_1). Le plan $\langle e_1, f_1 \rangle$ est donc hyperbolique et $\langle e_1, f_1 \rangle \cap \langle e_1, f_1 \rangle^\perp = \{0\}$ donc $E = \langle e_1, f_1 \rangle \perp \langle e_1, f_1 \rangle^\perp = \{0\}$. On peut appliquer une induction au sous-espace $F = \langle e_1, f_1 \rangle^\perp = \{0\}$ et l'écrire comme somme orthogonale de plans hyperboliques. □

Proposition. (Pfaffien) *Il existe un polynôme* Pf *à coefficients entiers (appelé Pfaffien) tel que si les coefficients x_{ij} d'une matrice A sont des variables telles que $x_{ij} = -x_{ji}$ alors*

$$\det(A) = \mathrm{Pf}(x_{ij})^2.$$

Si l'on impose $\mathrm{Pf}(J) = 1$, *le polynôme* Pf *est unique et vérifie de plus :*

$$\text{si } ((y_{ij})) = {}^tC((x_{ij}))C, \text{ alors} \qquad \mathrm{Pf}(y_{ij}) = \det(C)\,\mathrm{Pf}(x_{ij})$$

Preuve. En travaillant avec des coefficients dans le corps $\mathbf{Q}(\ldots, x_{ij}, \ldots)$, on voit que $A = P \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} {}^tP$ donc $\det(A) = (\det(P))^2$. A priori $\det(P) = Q/R$ avec $Q, R \in \mathbf{Z}[\ldots, x_{ij}, \ldots]$ mais comme ce dernier

anneau est factoriel et que $Q^2/R^2 \in \mathbf{Z}[\ldots, x_{ij}, \ldots]$, on en déduit que R divise Q dans $\mathbf{Z}[\ldots, x_{ij}, \ldots]$ et donc que $\det(P) \in \mathbf{Z}[\ldots, x_{ij}, \ldots]$. On pose alors $\mathrm{Pf}(\ldots, x_{ij}, \ldots) = \pm\det(P)$ en choisissant le signe de sorte que $\mathrm{Pf}(J) = +1$. Pour la deuxième formule on voit que $\det(y_{ij}) = \det(x_{ij})\det(C)^2$ donc $\mathrm{Pf}(y_{ij}) = \pm\det(C)\,\mathrm{Pf}(x_{ij})$. On détermine le signe en choisissant $C = I$. □

Corollaire. *On a* $\mathrm{Sp}(E) \subset \mathrm{SL}(E)$*, c'est-à-dire que pour toute matrice* $M \in \mathrm{Sp}(n, K)$ *on a* $\det(M) = +1$.

Preuve. Une matrice M est dans $\mathrm{Sp}(E)$ si $J = {}^tMJM$ donc $\mathrm{Pf}(J) = \det(M)\,\mathrm{Pf}(J)$ et $\det(M) = 1$. □

Etudions maintenant les transformation symplectiques qui fixent un hyperplan. On sait déjà que se sont des transvections de la forme $\sigma(x) = x + f(x)a$ avec f forme linéaire vérifiant $f(a) = 0$. Calculons donc $B(f(x), f(y)) - B(x, y) = B(f(x)a, y) + B(x, f(y)a) = B(f(y)x - f(x)y, a)$; en remarquant que l'ensemble des $f(y)x - f(x)y$ est $\mathrm{Ker}(f)$, on voit que la transvection σ est symplectique si et seulement si $B(\mathrm{Ker}(f), a) \equiv 0$, c'est-à-dire si $a \in \mathrm{Ker}(f) \cap \mathrm{Ker}(f)^\perp$. Ceci justifie la définition suivante

Définition. Une *transvection symplectique* est une application de la forme $\sigma(x) = x + f(x)a$ avec $0 \neq a \in \mathrm{Ker}(f) \cap \mathrm{Ker}(f)^\perp$. On peut aussi l'écrire $\sigma(x) = x + \lambda B(x, a)a$.

Thorme. *Les transvections symplectiques engendrent le groupe* $\mathrm{Sp}(E)$.

Remarque. Comme une transvection est clairement une matrice de $\mathrm{SL}_{2n}(k)$, on obtient ainsi une nouvelle preuve que $\mathrm{Sp}_{2n}(k) \subset \mathrm{SL}_{2n}(k)$.

Preuve. Le lemme suivant permet de démontrer le théorème par récurrence (passant d'un espace de dimension $2n$ à un espace de dimension $2n - 2$).

Lemme. *Soit* $P = \langle x_1, x_2\rangle$ *et* $P' = \langle y_1, y_2\rangle$ *deux plan hyperboliques (i.e.* $B(x_1, x_2) = B(y_1, y_2) = 1$*) alors il existe* ρ *un produit de transvections symplectiques tel que* $\rho(x_1) = y_1$ *et* $\rho(x_2) = y_2$.

Preuve du lemme. Commençons par envoyer x_1 sur y_1. Si $B(x_1, y_1) \neq 0$ alors une transvection suffit : on choisit $t(x) = x + \lambda B(x, a)a$ avec $\lambda = B(x_1, y_1)^{-1}$ et $a = y_1 - x_1$ de sorte que $t(x_1) = y_1$. Si jamais $B(x_1, y_1) = 0$, on prend un vecteur z tel que $B(x_1, z) = 1$ mais tel que $B(x_2, z)$ et $B(y_1, z)$ sont non nuls et on passera de x_1 à z puis à y_1 avec deux transvections. Envoyons maintenant x_2 sur y_2 en laissant fixe x_1. De nouveau, si $B(x_2, y_2) \neq 0$, une transvection suffit : on choisit $t(x) = x + \lambda B(x, a)a$ avec $\lambda = B(x_2, y_2)^{-1}$ et $a = y_2 - x_2$ de sorte que $B(x_1, a) = B(x_1, y_2) - B(x_1, x_2) = 1 - 1 = 0$ donc $t(x_1) = x_1$ et $t(x_2) = y_2$. Si jamais $B(x_2, y_2) = 0$ on va choisir z tel que $B(x_1, z) = 1$ mais $B(x_2, z)$ et $B(y_2, z)$ non nuls car alors on pourra passer de x_2 à z puis à y_2 en laissant fixe x_1 par deux transvections. On vérifie immédiatement que $z = x_1 + y_2$ convient. □

Thorme. *Le groupe* $\mathrm{PSp}_n(k)$ *est simple sauf pour* $n = 2$ *et* $k = \mathbf{F}_2, \mathbf{F}_3$ *ou* $n = 4$ *et* $k = \mathbf{F}_2$.

Remarque. On a clairement $\mathrm{Sp}_2(k) = \mathrm{SL}_2(k)$ donc le théorème est en fait déjà démontré dans le cas $n = 2$.

Soit H un sous-groupe normal de $\mathrm{Sp}_{2n}(k)$ contenant $\sigma \neq \pm Id$. Remarquons tout d'abord que si H contient toutes les transvections de direction a, il contiendra toutes les conjuguées et donc toutes les transvections et donc $H = \mathrm{Sp}_{2n}(k)$ (en particulier, si $k = \mathbf{F}_2$ ou $\mathbf{F}_3$, il suffit que H contienne une transvection). Choisissons a un vecteur tel que $b := \sigma(a)$ ne soit pas colinéaire avec a. Montrons d'abord qu'on peut supposer $B(a, b) \neq 0$. En effet si $B(a, b) = 0$, on peut choisir $c \in \langle b\rangle^\perp$ avec disons $B(c, a) = 1$; on choisit ensuite une transvection $t(x) = x + B(c - a, x)(c - a)$ qui vérifie $t(a) = c$ et $t(b) = b$ donc, si $\rho := t\sigma^{-1}t^{-1}\sigma$, on a $\rho \in H$ et $\rho(a) = c$.

On suppose donc $\sigma(a) = b$ et $B(a, b) \neq 0$. On choisit $t(x) = x + B(a, x)a$ et $\rho := t\sigma^{-1}t^{-1}\sigma$, on a $\rho \in H$ et $\rho(b) = b + B(b, a)a$ non colinéaire à b. Mais ρ est le produit de t qui laisse fixe $\langle a\rangle^\perp$ et de $\sigma^{-1}t^{-1}\sigma$ qui laisse fixe $\langle b\rangle^\perp$ donc ρ fixe l'orthogonal du plan hyperbolique P engendré par a et b. Si on décompose $E = P \oplus P^\perp$, on a donc $\rho = (\rho_1, I_{P^\perp}) \in \mathrm{Sp}_2(k) \times \{I_{P^\perp}\}$ et comme $\mathrm{PSp}_2(k) = \mathrm{PSL}_2(k)$ est simple (sauf si $k = \mathbf{F}_2$ ou $\mathbf{F}_3$), on conclut que H contient $\mathrm{Sp}_2(k) \times \{I_{P^\perp}\}$ et donc toutes les transvections de direction disons a et donc $H = \mathrm{Sp}_{2n}(k)$. Supposons que l'on ait montré que $\mathrm{PSp}_4(\mathbf{F}_3)$ et $\mathrm{PSp}_6(\mathbf{F}_2)$ sont simples, alors on peut appliquer le raisonnement précédent lorsque $k = \mathbf{F}_2$ et $n \geq 6$ (resp. $k = \mathbf{F}_3$ et $n \geq 4$) en incluant P dans une somme de trois plans hyperboliques (resp. de 2 plans hyperboliques). Nous renvoyons au livre de Artin pour la preuve concernant ces deux groupes particuliers (qui sont de cardinal 1 451 520 et 25 920). □

Exercice. Montrer, par une méthode similaire à celle suggérée pour les groupes orthogonaux, que :

$$\mathrm{card}(\mathrm{Sp}_{2n}(\mathbf{F}_q)) = q^{n^2}\prod_{i=1}^{n}(q^{2i}-1)$$

Exercice. Montrer que la réunion des matrices suivantes fournit un ensemble de générateurs de $\mathrm{Sp}_{2n}(k)$: les matrices $\begin{pmatrix} I & S \\ 0 & I \end{pmatrix}$ (où S est une matrice $n\times n$ symétrique), les matrices $\begin{pmatrix} A & 0 \\ 0 & {}^tA^{-1} \end{pmatrix}$ (où A une matrice $n\times n$ inversible) et $J := \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$. Montrer que si $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ est symplectique, alors $M^{-1} = \begin{pmatrix} {}^tD & -{}^tB \\ -{}^tC & {}^tA \end{pmatrix}$.

E.5. Groupe unitaire.

On examine maintenant le cas d'une forme σ-hermitienne $H : E\times E \to k$ avec E un k-espace vectoriel qu'on supposera souvent de dimension ≥ 2. On note k_0 le sous-corps fixé par l'involution σ. Les deux cas que nous traiterons comme exemple sont $k = \mathbf{C}$ (avec $\sigma(z) = \bar{z}$ et $k_0 = \mathbf{R}$) et $k = \mathbf{F}_{q^2}$ (avec $\sigma(x) = x^q$ et $k_0 = \mathbf{F}_q$). Dans le premier cas on a vu qu'on peut se ramener à une forme du type $B(x,x) = x_1\bar{x}_1 + \ldots + x_p\bar{x}_p - x_{p+1}\bar{x}_{p+1} - \ldots - x_n\bar{x}_n$ et on notera le groupe correspondant $\mathrm{U}_{p,n-p}(\mathbf{C})$ ou encore $\mathrm{U}_n(\mathbf{C})$ si $p = n$; il existe des vecteurs isotropes non nuls si et seulement si $1 \leq p \leq n-1$. Dans le second cas il faut observer que l'application de $\mathbf{F}_{q^2}^*$ vers $\mathbf{F}_q^*$ donnée par $x \mapsto x\sigma(x) = x^{q+1}$ est surjective (son noyau est de cardinal $q+1$) et donc on peut se ramener à la forme $B(x,x) = x_1\sigma(x_1) + \ldots + x_n\sigma(x_n)$. On notera $\mathrm{U}_n(\mathbf{F}_{q^2})$ le groupe correspondant. Observons enfin que, dans ce cas, il existe des vecteurs isotropes dès que $n \geq 2$.

Notons $S := \{x \in k^* \mid x\sigma(x) = 1\}$ le "*cercle unité* ". Le déterminant d'un élément unitaire est dans S. Inversement soit $\lambda \in S$, si on a mis la forme bilinéaire sous forme diagonale, la matrice $\mathrm{diag}(1,\ldots,1,\lambda)$ est unitaire de déterminant λ. On a donc la suite exacte :

$$0 \to \mathrm{SU}(B) \to \mathrm{U}(B) \to S \to 0.$$

Supposons $\mathrm{car}(k) \neq 2$ et $k = k_0(\omega)$ avec $\omega^2 = \delta \in k_0$ et $\sigma(\omega) = -\omega$. On peut décomposer $B(x,y) = R(x,y) + \omega I(x,y)$ en partie "réelle" et "imaginaire" à valeur dans k_0. On s'aperçoit facilement que R est k_0-bilinéaire symétrique et I est k_0-bilinéaire antisymétrique. De plus elles sont liés par la relation $I(x,\omega x) = -R(x,x)$. Ainsi la donnée de B équivaut à celle de R ou celle de I.

On peut définir l'analogue des symétries hyperplanes comme les transformations laissant fixes les points d'un hyperplan non isotrope. Ce sont les *quasi-symétries*; elles sont de la forme

$$u_{\lambda,e}(x) = x + (\lambda-1)\frac{B(x,e)}{B(e,e)}e \qquad \text{avec } \lambda \in S \text{ et } e \text{ non isotrope.}$$

On a bien sûr $u_{\lambda,e}(x) = x$ si $x \in \langle e\rangle^\perp$ et $u_{\lambda,e}(e) = \lambda e$; ainsi $\det(u_{\lambda,e}) = \lambda$. On peut calquer la démonstration faite pour le groupe orthogonal et montrer que les quasi-symétries engendrent le groupe unitaire (au moins en caractéristique différente de 2). On omet également la démonstration du théorème suivant, qui utilise des techniques assez similaires à celles utilisées dans les paragraphes précédents.

Thorme. *Supposons qu'il existe un vecteur isotrope non nul et que l'on est pas dans le cas $n = 2$ et $q = 2$ ou 3, ou $n = 3$ et $q = 2$. Alors le groupe $\mathrm{PSU}(B)$ est simple. Si $k = \mathbf{F}_{q^2}$ et $\dim E = n$ on a*

$$\mathrm{card}(\mathrm{SU}_n(\mathbf{F}_{q^2})) = q^{n(n-1)/2}(q^2-1)(q^3+1)\ldots(q^n-(-1)^n).$$

Il est également vrai que $\mathrm{PSU}_n(\mathbf{C})$ est simple pour $n \geq 2$.

Pour les deux dernières affirmations, on peut dénombrer dans le cas fini comme on l'a indiqué pour le groupe orthogonal et ramener la preuve de la simplicité de $\mathrm{PSU}_n(\mathbf{C})$ à celle de $\mathrm{PSU}_2(\mathbf{C})$. Or ce dernier groupe est isomorphe à $\mathrm{SO}_3(\mathbf{R})$ comme on le verra au paragraphe suivant en utilisant les quaternions.

Il existe bien sûr de nombreux liens entre les groupes que nous avons brièvement étudié. Par exemple, définissons, pour $n \geq 1$ l'espace de Siegel

$$\mathcal{H}_n = \{\tau \in \mathrm{Mat}(n \times n, \mathbf{C} \mid \tau \text{ est symétrique et } \mathrm{Im}(\tau) > 0\}.$$

La notation $\mathrm{Im}(\tau) > 0$ signifie ici que la matrice $\mathrm{Im}(\tau)$ est définie positive. On définit alors une action de $G = \mathrm{Sp}_{2n}(\mathbf{R})$ par la formule, où $\gamma = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$:

$$\gamma \cdot \tau = (A\tau + B)(C\tau + D)^{-1}.$$

Lorsque $n = 1$, on retrouve le demi-plan de Poincaré et l'action classique de $\mathrm{SL}_2(\mathbf{R})$. L'action de G sur $\mathcal{H}_n$ est transitive et le stabilisateur de iI est un sous-groupe compact isomorphe à $U_n(\mathbf{C})$ par l'application

$$\begin{array}{ccc} U_n(\mathbf{C}) & \to & \mathrm{Sp}_{2n}(\mathbf{R}) \\ A + iB & \mapsto & \begin{pmatrix} A & B \\ -B & A \end{pmatrix} \end{array}$$

Avec cette identification, on vérifie facilement que $\mathrm{U}_n(\mathbf{C}) = \mathrm{O}_{2n}(\mathbf{R}) \cap \mathrm{Sp}_{2n}(\mathbf{R})$.

Exercices de topologie. On termine en proposant un certain nombre de propriétés des groupes classiques sur $\mathbf{R}$ ou $\mathbf{C}$ dont on pourra trouver une preuve (et beaucoup plus!) dans le livre de Mneimné et Testard cité ci-dessous.

(a) Les groupes $\mathrm{O}_n(\mathbf{R})$, $\mathrm{SO}_n(\mathbf{R})$, $\mathrm{U}_n(\mathbf{C})$ et $\mathrm{SU}_n(\mathbf{C})$ sont compacts; les trois derniers sont connexes alors que le premier a deux composantes connexes.

(b) (décomposition polaire sur $\mathbf{R}$) Tout élément M de $\mathrm{GL}_n(\mathbf{R})$ se décompose de manière unique en produit $M = OS$ avec $O \in \mathrm{O}_n(\mathbf{R})$ et S symétrique définie positive. En déduire que $\mathrm{GL}_n(\mathbf{R})$ est *homéomorphe* à $\mathrm{O}_n(\mathbf{R}) \times \mathbf{R}^{n(n+1)/2}$ et possède donc deux composantes connexes.

(c) (décomposition polaire sur $\mathbf{C}$) Tout élément M de $\mathrm{GL}_n(\mathbf{C})$ se décompose de manière unique en produit $M = UH$ avec $U \in \mathrm{U}_n(\mathbf{C})$ et H hermitienne définie positive. En déduire que $\mathrm{GL}_n(\mathbf{C})$ est *homéomorphe* à $\mathrm{U}_n(\mathbf{C}) \times \mathbf{R}^{n^2}$ et connexe.

(d) (Décomposition d'Iwasawa). Soit $K = \mathrm{SO}_n(\mathbf{R})$, notons A le sous-groupe des matrices diagonales $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ avec $\lambda_i > 0$ et $\prod_i \lambda_i = 1$ et N le sous-groupe des matrices triangulaires possédant des 1 sur la diagonale. L'application de $K \times A \times N$ vers $\mathrm{SL}_n(\mathbf{R})$ définie par $(k, a, n) \to kan$ définit un homéomorphisme.

(e) Les groupes $\mathrm{SL}_n(\mathbf{R})$, $\mathrm{SL}_n(\mathbf{C})$, $\mathrm{Sp}_{2n}(\mathbf{R})$ et $\mathrm{Sp}_{2n}(\mathbf{C})$ sont connexes de même que $\mathrm{SO}_n(\mathbf{C})$ et $\mathrm{SO}_n(\mathbf{R})$.

(f) Par contre $\mathrm{SO}_{p,q}(\mathbf{R})$ possède deux composantes connexes si $p, q \geq 1$ (et le groupe n'est pas compact).

(g) Le groupe fondamental de $\mathrm{SO}_n(\mathbf{R})$ ou de $\mathrm{SL}_n(\mathbf{R})$ est $\mathbf{Z}/2\mathbf{Z}$ si $n \geq 3$ et $\mathbf{Z}$ si $n = 2$.

(h) Les groupes $\mathrm{SU}_n(\mathbf{C})$, $\mathrm{SL}_n(\mathbf{C})$ sont simplement connexes. Les groupes $\mathrm{U}_n(\mathbf{C})$, $\mathrm{GL}_n(\mathbf{C})$ et $\mathrm{Sp}_{2n}(\mathbf{R})$ ont pour groupe fondamental $\mathbf{Z}$. Le groupe $\mathrm{SU}_{p,q}(\mathbf{C})$ est connexe et son groupe fondamental est isomorphe à $\mathbf{Z}$.

On termine par quelques références spécifiques à ce chapitre (l'article de Tits contient notamment la classification des groupes simples avec les groupes exceptionnels E_6, E_7, E_8, F_4 et G_2).

Artin, E., Geometric Algebra, Interscience, 1957.

Dieudonné, J., La géométrie des groupes classiques, Ergebnisse d. Math. Springer, 1955.

Mneimné, R. et Testard, F., Introduction à la théorie des groupes de Lie classiques. Hermann, 1986.

Tits, J., Groupes simples et géométries associées, Actes du congrès international des mathématiciens de Stockholm (1962), pages 197-221.

E.6. Quaternions, arithmétique et groupe orthogonal.

Nous allons construire l'exemple classique de corps non commutatif : le corps des quaternions découvert par Hamilton, et développer deux applications, l'une arithmétique (le théorème des quatre carrés), l'autre géométrique (l'étude des groupes d'isométries $\mathrm{SO}(3,\mathbf{R})$*,* $\mathrm{SO}(4,\mathbf{R})$ *et* $\mathrm{SU}(2,\mathbf{C})$*). Nous montrerons aussi que le corps des quaternions est le "seul" corps non commutatif de dimension fini sur* $\mathbf{R}$*.*

E.6.1. Le corps des quaternions.

La façon la plus concrète de construire le corps des quaternions est comme un espace vectoriel réel de dimension 4 muni d'une base $\mathbf{1}, I, J, K$ et d'une multiplication $\mathbf{R}$-bilinéaire définie par le fait que $\mathbf{1}$ est élément neutre et les formules

$$I^2 = J^2 = K^2 = -\mathbf{1}, \quad IJ = -JI = K, \quad JK = -KJ = I \quad \text{et} \quad KI = -IK = J \qquad (*)$$

Il faut alors vérifier "à la main" l'associativité : par exemple $(IJ)K = K^2 = -\mathbf{1}$ et $I(JK) = I^2 = -\mathbf{1}$. Pour s'épargner cette vérification on peut aussi définir $\mathbf{H}$ comme sous-algèbre des matrices 2×2 complexes ou 4×4 réelles (l'associativité est alors immédiate mais il faut vérifier que les matrices introduites vérifient les formules $(*)$). On peut ainsi définir

$$\mathbf{H} = \left\{ \begin{pmatrix} \alpha & -\beta \\ \bar{\beta} & \bar{\alpha} \end{pmatrix} \middle| \ \alpha, \beta \in \mathbf{C} \right\}$$

avec $\mathbf{1} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, $I = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$, $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ et $K = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$ ou encore

$$\mathbf{H} = \left\{ \begin{pmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{pmatrix} \middle| \ a, b, c, d \in \mathbf{R} \right\}$$

avec

$$\mathbf{1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \ I = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \ J = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \ K = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

Remarque. Une fois construit $\mathbf{H}$, on peut remarquer que c'est une $\mathbf{R}$-algèbre engendrée par deux éléments i, j avec les relations $i^2 = j^2 = -\mathbf{1}$ et $ij = -ji$. En effet en posant $k := ij$ on en déduit la table de multiplication puisque $k^2 = ijij = -iijj = -\mathbf{1}$ et $ik = iij = -j = (ii)j = -iji = -ki$ etc. Le fait que $\mathbf{H}$ ne soit pas commutatif se lit déjà sur la table de multiplication, mais plus précisément nous avons le lemme suivant

Lemme. *Le centre de* $\mathbf{H}$ *est* $\mathbf{R1}$ *(que l'on identifiera, le cas échéant, à* $\mathbf{R}$*). Si* $z \in \mathbf{H} \setminus \mathbf{R1}$ *alors*

$$C(z) := \{z' \in \mathbf{H} \mid zz' = z'z\} = \mathbf{R1} + \mathbf{R}z$$

Preuve. Si $q = a\mathbf{1} + bI + cJ + dK$ et $q' = a'\mathbf{1} + b'I + c'J + d'K$ sont deux quaternions, leur multiplication s'écrit

$$qq' = (aa' - bb' - cc' - dd')\mathbf{1} + (ab' + ba' + cd' - dc')I + (ac' - bd' + ca' + db')J + (ad' + bc' - cb' + da')K$$

donc les deux éléments commutent si et seulement si

$$\begin{pmatrix} 0 & -d & c \\ d & 0 & -b \\ -c & b & 0 \end{pmatrix} \begin{pmatrix} b' \\ c' \\ d' \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

On voit donc que tous les éléments commutent avec $a\mathbf{1}+bI+cJ+dK$ si $b=c=d=0$ mais, si $z\in\mathbf{H}\setminus\mathbf{R}\mathbf{1}$, alors un calcul simple montre que le rang du système est égal à deux donc $\dim_{\mathbf{R}} C(z)=2$, or clairement $\mathbf{R}(z)\subset C(z)$ et $[\mathbf{R}(z):\mathbf{R}]=2$ donc $\mathbf{R}(z)=C(z)$. □

On introduit le *conjugué* d'un quaternion $z=a\mathbf{1}+bI+cJ+dK$ comme $\bar{z}=a\mathbf{1}-bI-cJ-dK$ ainsi que sa *trace* $\mathrm{Tr}(z)=z+\bar{z}$ et sa *norme* $\mathrm{N}(z)=z\bar{z}$. On vérifie alors

Lemme. *Soient $z,w\in\mathbf{H}$, $\overline{z+w}=\bar{z}+\bar{w}$, $\overline{zw}=\bar{w}\cdot\bar{z}$ et si $z=a\mathbf{1}+bI+cJ+dK$, alors $\mathrm{N}(z)=z\bar{z}=\bar{z}z=(a^2+b^2+c^2+d^2)\mathbf{1}$ et $\mathrm{Tr}(z)=2a\mathbf{1}$; de plus $\mathrm{Tr}(z+z')=\mathrm{Tr}(z)+\mathrm{Tr}(z')$, $\mathrm{N}(zz')=\mathrm{N}(z)\,\mathrm{N}(z')$ et z est racine du polynôme $X^2-\mathrm{Tr}(z)X+\mathrm{N}(z)\in\mathbf{R}[X]$.*

Preuve. Des calculs directs (laissés au lecteur) permettent de vérifier ces formules. Noter que la conjugaison est un *anti-isomorphisme* de corps, i.e. qu'elle renverse l'ordre de la multiplication. □

On voit immédiatement comme corollaire que $\mathbf{H}$ est un corps puisque, si $z=a\mathbf{1}+bI+cJ+dK$ est un quaternion non nul, alors $\mathrm{N}(z):=a^2+b^2+c^2+d^2\in\mathbf{R}^*$ et $z\bar{z}/\,\mathrm{N}(z)=\mathbf{1}$ donc $z^{-1}=\bar{z}/\,\mathrm{N}(z)$.

Remarque. On dispose donc d'une sorte de *décomposition polaire* des quaternions en notant G le groupe (multiplicatif) des quaternions de norme 1.

$$\mathbf{H}^*\cong\mathbf{R}_+^*\times G\qquad\text{(isomorphisme de groupes)}$$

donnée par $z\mapsto(\sqrt{\mathrm{N}(z)},z/\sqrt{\mathrm{N}(z)})$. Toutefois on prendra garde que le groupe G (analogue du cercle unité pour les complexes) n'est pas un groupe commutatif, en fait il est isomorphe au groupe $\mathrm{SU}(2,\mathbf{C})$ (alors que le cercle unité peut être vu comme $\mathrm{SU}(1,\mathbf{C})$). Ce dernier point peut se montrer facilement à partir de la représentation

$$\mathbf{H}=\mathbf{C}\oplus\mathbf{C}j$$

dans laquelle il faut faire attention que, si $z=a+bi\in\mathbf{C}$, alors

$$jz=\bar{z}j\qquad\text{(en général }\neq zj)$$

Exercice. Montrer que l'équation $X^2-1=0$ possède exactement deux solutions dans $\mathbf{H}$ mais que l'équation $X^2+1=0$ possède une infinité de solutions dans $\mathbf{H}$ (indication : on montrera que la sphère $a=b^2+c^2+d^2-1=0$ décrit l'ensemble des solutions).

Exercice. Montrer que $\mathbf{H}^*$ contient un sous-groupe fini non-cyclique (rappel : ceci est impossible dans le groupe multiplicatif d'un corps commutatif).

E.6.2 Sommes de carrés d'entiers.

L'énoncé du théorème suivant se situe entièrement dans $\mathbf{N}$, pourtant il sera commode, pour le démontrer, de travailler dans l'anneau des entiers de Gauss $\mathbf{Z}[i]$.

Thorme. *Un entier $n\in\mathbf{N}$ peut s'écrire comme somme de deux carrés d'entiers si et seulement si chaque nombre premier p congru à 3 modulo 4 apparaît avec un exposant pair dans la décomposition en facteurs premiers de n.*

L'énoncé du théorème suivant n'a rien à voir avec les quaternions mais nous allons le démontrer en étudiant l'arithmétique de sous-anneaux du corps $\mathbf{H}$.

Thorme. (Lagrange) *Soit $n\in\mathbf{N}$ alors il existe des entiers x,y,z,t tels que $n=x^2+y^2+z^2+t^2$.*

Remarque. On voit facilement qu'un carré est congru à $+1$, 0 ou 4 modulo 8 donc si n est congru à 7 modulo 8, il n'est pas somme de trois carrés. Le même raisonnement montre que si $n=4m=x^2+y^2+z^2$ alors chacun des entiers x,y,z est divisible par 2 et donc m est également somme de trois carrés. On peut donc conclure que les entiers de la forme $n=4^r(8m+7)$ ne sont pas somme de trois carrés. Il est vrai, mais

nous ne le démontrerons pas ici, que tous les autres entiers peuvent s'écrire comme somme de trois carrés d'entiers (par exemple $2.7 = 14 = 3^2 + 2^2 + 1^2$, $2^3.7 = 6^2 + 4^2 + 2^2$ et $30 = 5^2 + 2^2 + 1^2$).

Remarque. L'ensemble des sommes de deux carrés (resp. de quatre carrés) est stable par produit mais pas l'ensemble des sommes de trois carrés. En effet $18 = 2.3^2 = 4^2 + 1^2 + 1^2$ et $14 = 2.7 = 3^2 + 2^2 + 1^2$ mais $18.14 = 4.9.7$ n'est pas somme de trois carrés.

Si on pose

$$\mathcal{C}_2 := \{n \in \mathbf{N} \mid \exists x, y \in \mathbf{N},\ n = x^2 + y^2\} \text{ et } \mathcal{C}_4 := \{n \in \mathbf{N} \mid \exists x, y, z, t \in \mathbf{N},\ n = x^2 + y^2 + z^2 + t^2\}$$

on veut donc montrer que $n \in \mathcal{C}_2$ si et seulement si tout nombre premier congru à 3 modulo 4 apparaît avec un exposant pair et que $\mathcal{C}_4 = \mathbf{N}$. On va introduire l'anneau $\mathbf{Z}[i]$ et les deux anneaux

$$A_0 = \mathbf{Z}\mathbf{1} + \mathbf{Z}I + \mathbf{Z}J + \mathbf{Z}K \quad \text{ et } A = A_0 + \mathbf{Z}\left(\frac{1 + I + J + K}{2}\right).$$

Il est clair que $\mathcal{C}_2 = \{\mathrm{N}(z) \mid z \in B\}$ et $\mathcal{C}_4 = \{\mathrm{N}(z) \mid z \in A_0\}$, en fait on a aussi $\mathcal{C}_4 = \{\mathrm{N}(z) \mid z \in A\}$ car $N(A_0) = N(A)$. En effet d'une part, si $x, y, z, t \in \mathbf{Z} + 1/2$ alors $\mathrm{N}(x\mathbf{1} + yI + zJ + tK) \in \mathbf{Z}$, d'autre part si $\alpha \in A \setminus A_0$, on peut écrire $\alpha = 2\alpha_0 + \epsilon$ avec $\alpha_0 \in A_0$ et $\epsilon = (\pm 1 \pm I \pm J \pm K)/2$ et alors $\alpha\overline{\epsilon} = \alpha_0(2\overline{\epsilon}) + 1 \in A_0$. De plus, la norme étant multiplicative, il suffit de montrer que tout nombre premier p est une norme. Comme $2 = 1^2 + 1^2$ il suffit d'ailleurs de le faire pour p premier impair. Pour cela nous allons montrer d'abord que $\mathbf{Z}[i]$ est principal et A est *principal* à gauche (ou à droite).

Proposition. *L'anneau $\mathbf{Z}[i]$ est euclidien donc principal. L'anneau A est euclidien à gauche, donc principal à gauche (idem à droite).*

Preuve. Notons B l'anneau A ou $\mathbf{Z}[i]$, l'énoncé signifie que pour $\alpha \in B$ et $\beta \in B \setminus \{0\}$, il existe $q, r \in B$ tel que $\alpha = q\beta + r$ avec $\mathrm{N}(r) < \mathrm{N}(\beta)$ (lorsque l'anneau est A, il faut faire attention au sens des multiplications). Supposons ceci démontré, on en tire aussitôt que $\mathbf{Z}[i]$ est principal, en fait la "même" démonstration montre que A est principal (à gauche). Soit donc I un idéal à gauche non nul de A (i.e. $A.I \subset I$), il contient un élément $\beta \neq 0$ de norme minimale et on a clairement $A\beta \subset I$. Inversement, soit $\alpha \in I$, écrivons $\alpha = q\beta + r$ avec $\mathrm{N}(r) < \mathrm{N}(\beta)$, on a alors $r = \alpha - q\beta \in I$ donc r est nul et on a bien $I = A\beta$. Montrons maintenant que A et $\mathbf{Z}[i]$ sont euclidiens. La preuve est basée sur le lemme élémentaire suivant dont la preuve est laissée au lecteur.

Lemme. *Soit $x \in \mathbf{R}$, il existe $m \in \mathbf{Z}$ tel que $|x - m| \leq 1/2$ et il existe $n \in \mathbf{Z}$ tel que $|x - n/2| \leq 1/4$.*

Soit donc $\alpha \in \mathbf{Z}[i]$ et $\beta \in \mathbf{Z}[i] \setminus \{0\}$, alors $\alpha/\beta = x + iy \in \mathbf{Q}[i]$ et il existe $m, n \in \mathbf{Z}$ tels que $|x - m| \leq 1/2$ et $|y - n| \leq 1/2$ donc

$$\mathrm{N}\left((x + iy) - (m + in)\right) = (x - m)^2 + (y - n)^2 \leq \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

d'où, si l'on note $q := m + ni$ l'inégalité cherchée

$$\mathrm{N}(\alpha - q\beta) \leq \frac{\mathrm{N}(\beta)}{2} < \mathrm{N}(\beta).$$

Soit maintenant $\alpha \in A$ et $\beta \in A \setminus \{0\}$, alors $\alpha\beta^{-1} = x + yI + zJ + tK \in \mathbf{H}$ et il existe $m \in \mathbf{Z}$ tel que $|x - m/2| \leq 1/4$. On choisit alors $q = (m + nI + hJ + \ell K)/2$ avec m, n, h, ℓ entiers de même parité (de sorte que $q \in A$) et tel que $|y - n/2|$, $|z - h/2|$ et $|t - \ell/2|$ soient $\leq 1/2$. On obtient alors

$$\mathrm{N}(\alpha\beta^{-1} - q) = \left(x - \frac{m}{2}\right)^2 + \left(y - \frac{n}{2}\right)^2 + \left(z - \frac{h}{2}\right)^2 + \left(t - \frac{\ell}{2}\right)^2 \leq \frac{1}{16} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4} < 1$$

d'où l'inégalité cherchée

$$\mathrm{N}(\alpha - q\beta) < \mathrm{N}(\beta).$$

□

(Somme de deux carrés). L'anneau $\mathbf{Z}[i]$ est principal donc factoriel et on voit facilement que $\mathbf{Z}[i]^* = \{\pm 1, \pm i\}$ (voir exercice) ; déterminons maintenant les éléments irréductibles. Tout d'abord $2 = -i(1+i)^2$ et $1+i$ est irréductible car sa norme est 2. Un nombre premier p congru à 3 modulo 4 est irréductible dans $\mathbf{Z}[i]$ car si $p = \alpha\beta$ alors $\mathrm{N}(\alpha)\,\mathrm{N}(\beta) = p^2$ mais l'égalité $\mathrm{N}(\alpha) = p$ est impossible donc $\mathrm{N}(\alpha)$ ou $\mathrm{N}(\beta)$ vaut 1 et donc α ou β est inversible. Enfin, soit p un nombre premier congru à 1 modulo 4, on sait que le groupe $(\mathbf{Z}/p\mathbf{Z})^*$ est cyclique d'ordre $p-1$ donc contient un élément $\bar{a}$ d'ordre 4 donc tel que $a^2 \equiv -1 \bmod p$. En particulier on a donc $(a+i)(a-i) = a^2+1 \in p\mathbf{Z}[i]$ mais ni $a+i$ ni $a-i$ ne peuvent appartenir à $p\mathbf{Z}[i]$ qui n'est donc pas un idéal premier. Comme $\mathbf{Z}[i]$ est principal l'élément p n'est pas irréductible et on peut donc écrire $p = \alpha\beta$ avec $\mathrm{N}(\alpha) = \mathrm{N}(\beta) = p$ (on a en fait forcément $\beta = \bar{\alpha}$). On peut résumer cela en

Lemme. *Les éléments irréductibles de* $\mathbf{Z}[i]$ *(non associés deux à deux) sont :* $1+i$*, les premiers* p *congrus à* 3 *modulo* 4*, les deux facteurs* α_p, β_p *de* p *congrus à* 1 *modulo* 4 *décomposant* $p = \alpha_p\beta_p$*.*

La norme d'un entier de Gauss dont la factorisation s'écrit

$$q = i^r(1+i)^s \prod_{p\equiv 1 \bmod 4} \alpha_p^{m_p}\beta_p^{n_p} \prod_{p\equiv 3 \bmod 4} p^{\ell_p}$$

est donc égale à

$$\mathrm{N}(q) = 2^s \prod_{p\equiv 1 \bmod 4} p^{m_p+n_p} \prod_{p\equiv 3 \bmod 4} p^{2\ell_p}$$

ce qui démontre le théorème des deux carrés.

(Somme des quatres carrés). Il suffit de montrer que si p est un nombre premier impair, il est la norme d'un élément de A. Le nombre de carrés dans $\mathbf{Z}/p\mathbf{Z}$ est $(p+1)/2$ donc le polynôme $-1-X^2$ prend au moins une fois pour valeur un carré ; en d'autre termes, il existe $a, b \in \mathbf{Z}$ tels que $a^2+b^2+1 \in p\mathbf{Z}$. On en tire que $(\mathbf{1}+aI+bJ)(\mathbf{1}-aI-bJ) \in pA$. Considérons donc l'idéal (à gauche) I engendré par p et $\mathbf{1}+aI+bJ$, on a $I = A\beta$ puisque A est principal (à gauche) et d'autre part des inclusions strictes $pA \subset I \subset A$. Ainsi $p = \alpha\beta$ et $\mathrm{N}(p) = \mathrm{N}(\alpha)\,\mathrm{N}(\beta) = p^2$(avec β et α non inversibles) donc $\mathrm{N}(\alpha)$ et $\mathrm{N}(\beta)$ différents de 1 donc égaux à p. □

Exercice. Montrer qu'un élément de $\mathbf{Z}[i]$, A ou A_0 est inversible si et seulement si sa norme vaut 1. En déduire que

$$\mathbf{Z}[i]^* = \{\pm 1, \pm i\}, \qquad A_0^* = \{\pm 1, \pm I, \pm J, \pm K\} \qquad \text{et} \qquad A^* = A_0^* \cup \left\{\frac{\pm 1 \pm I \pm J \pm K}{2}\right\}$$

(A_0^* et A^* sont les groupes quaternioniques d'ordre 8 et 24 respectivement). Le groupe A^* est-il isomorphe à $\mathcal{S}_4$? Montrer que A_0 n'est pas principal (à gauche). En déduire également qu'un élément de norme égale à un nombre premier est irréductible.

E.6.3. Quaternions et isométries.

Nous allons voir que le groupe des quaternions de norme 1, noté G*, est isomorphe à* $\mathrm{SU}(2,\mathbf{C})$*, que* $\mathrm{SO}(3,\mathbf{R}) \cong G/\{\pm\mathbf{1}\}$ *et* $\mathrm{SO}(4,\mathbf{R}) \cong G \times G/\{\pm(\mathbf{1},\mathbf{1})\}$*, ce qui permettra de décrire les rotations de l'espace de dimension 3 ou 4.*

Dans toute la suite on identifiera librement $\mathbf{H}$ avec $\mathbf{R}^4$ (via la base $\mathbf{1}, I, J, K$) et $E := \mathbf{R}I + \mathbf{R}J + \mathbf{R}K$ avec $\mathbf{R}^3$ (via la base I, J, K). Le premier lien entre quaternions et isométries est l'observation simple que, si $x, y \in \mathbf{R}^4 = \mathbf{H}$, on a

$$||x||^2 = x\bar{x} = \mathrm{N}(x) \qquad \text{et} \qquad x \cdot y = \frac{1}{2}\,\mathrm{Tr}(x\bar{y})$$

Remarquons en particulier que $\mathbf{R1}$ est l'orthogonal de E dans $\mathbf{H}$. On étudie maintenant l'action par conjugaison $\Phi(q)(x) = qxq^{-1}$. Cette action fournit un homomorphisme $\Phi : \mathbf{H}^* \to \mathrm{GL}_{\mathbf{R}}(\mathbf{H})$. En fait $\mathrm{N}(\Phi(q)(x)) = \mathrm{N}(qxq^{-1}) = \mathrm{N}(x)$ donc $\Phi(q)$ est une isométrie et de plus $\Phi(q)(\mathbf{1}) = \mathbf{1}$ donc $\Phi(q)$ laisse stable $E \cong \mathbf{R}^3$ d'où en fait un homomorphisme de $\mathbf{H}^*$ vers $\mathrm{O}(3,\mathbf{R})$ qu'on note encore Φ. Comme $\mathbf{H}^*$ est connexe,

l'image de Φ est en fait dans $\mathrm{SO}(3,\mathbf{R})$. Par ailleurs on a $\mathbf{R}^* \subset \mathrm{Ker}(\Phi)$ (en fait on a égalité puisque $\mathbf{R}$ est le centre de $\mathbf{H}$) donc on peut restreindre Φ à G (le groupe des quaternions de norme 1) sans changer son image. On a donc défini un homomorphisme

$$\Phi : G \to \mathrm{SO}(3,\mathbf{R}) \qquad (*)$$

dont le noyau est $\mathbf{R}^* \cap G = \{+1,-1\}$; pour voir que Φ est surjectif on démontre le lemme suivant qui a son propre intérêt :

Lemme. *Soit x un quaternion de E de norme 1 (i.e. une solution de $x^2+1=0$), considérons $q = \cos(\theta)+\sin(\theta)x$, alors $\Phi(q)$ est la rotation d'angle 2θ et d'axe (orienté) x.*

Preuve. On a $qxq^{-1} = x$ puisque q commute avec x. Déterminons maintenant l'action de $\Phi(q)$ sur le plan P orthogonal à x (dans E). L'équation $2x\cdot y = \mathrm{Tr}(x\bar{y}) = x\bar{y}+y\bar{x} = 0$ équivaut, puisque dans E on a $\bar{y} = -y$, à l'équation $xy+yx=0$. Choisissons donc y de norme 1 tel que $xy=-yx$, alors $z=xy$ permet de compléter la famille libre $\{x,y\}$ en une base orthonormé $\{x,y,z\}$. On calcule alors (en remarquant que $xyx = y$ et $x^2 = y^2 = -1$)

$$\begin{aligned}\Phi(q)(y) &= (\cos(\theta)+\sin(\theta)x)y(\cos(\theta)-\sin(\theta)x)\\ &= \left(\cos^2(\theta)-\sin^2(\theta)\right)y + 2\cos(\theta)\sin(\theta)xy\\ &= \cos(2\theta)x+\sin(2\theta)xy\end{aligned}$$

et

$$\begin{aligned}\Phi(q)(xy) &= (\cos(\theta)+\sin(\theta)x)xy(\cos(\theta)-\sin(\theta)x)\\ &= \left(\cos^2(\theta)-\sin^2(\theta)\right)xy - 2\cos(\theta)\sin(\theta)y\\ &= \cos(2\theta)xy-\sin(2\theta)y\end{aligned}$$

ainsi $\Phi(q)$ est bien la rotation d'angle 2θ et d'axe x. □

Remarque. On voit en particulier que les *renversements*, i.e. les rotations d'angle π, correspondent à des quaternions imaginaires purs. Il n'est pas très difficile de voir que tout quaternion non nul peut s'écrire comme produit de quaternions imaginaires purs (exercice : le démontrer) et on en tire que les renversements sont des générateurs de $\mathrm{SO}(3,\mathbf{R})$ (Cf le paragraphe E.3). Remarquons enfin que, si on se restreint à $q \in G$ alors $\Phi(q)(x) = qx\bar{q}$ puisque $q^{-1} = \bar{q}$.

Pour étudier $\mathrm{SO}(4,\mathbf{R})$, on considère l'action de $\mathbf{H}^*\times\mathbf{H}^*$ sur $\mathbf{H}$ donnée par $\Psi(q,r)(x) = qxr^{-1}$. Il s'agit d'une similitude directe de rapport $\mathrm{N}(qr^{-1})$ puisque $\mathrm{N}(\Psi(q,r)(x)) = \mathrm{N}(qxr^{-1}) = \mathrm{N}(qr^{-1})\,\mathrm{N}(x)$; en particulier, si on se restreint à $G\times G$, on obtient un homomorphisme de groupes

$$\Psi : G\times G \to \mathrm{SO}(4,\mathbf{R})$$

dont on démontre de manière similaire qu'il est surjectif. Par ailleurs, il a pour noyau le sous-groupe $\mathrm{Ker}(\Psi) = \{(+1,+1),(-1,-1)\}$. En effet, si $u \in \mathrm{SO}(4,\mathbf{R})$ et $u(\mathbf{1}) = z$, posons $v(x) = z^{-1}u(x)$, alors $v(\mathbf{1}) = \mathbf{1}$ donc v est une rotation de E et on a vu qu'une telle rotation s'écrivait $v(x) = yxy^{-1}$ pour un certain $y \in G$. On a donc $u(x) = (zy)xy^{-1} = \Psi(zy,y)(x)$. Par ailleurs, si $\forall x \in \mathbf{H}, \Psi(q,r)(x) = x$ alors, en prenant $x = \mathbf{1}$ on obtient $q = r$ et donc q, de norme 1, est dans le centre de $\mathbf{H}$ donc vaut ± 1. On a donc bien démontré

Thorme. *Le groupe G est isomorphe à $\mathrm{SU}(2,\mathbf{C})$; l'homorphisme Φ induit un isomorphisme $G/\{+1,-1\} \cong \mathrm{SO}(3,\mathbf{R})$; l'homorphisme Ψ induit un isomorphisme $G\times G/\{(1,1),(-1,-1)\} \cong \mathrm{SO}(4,\mathbf{R})$. En particulier $\mathrm{SO}(4,\mathbf{R})/\{\pm id\}$ n'est pas simple.*

Exercices. a) Vérifier par un calcul direct que $\mathrm{SU}(2,\mathbf{C})$ est l'ensemble des matrices 2×2 à coefficients complexes de la forme $\begin{pmatrix}\alpha & -\beta\\ \bar{\beta} & \bar{\alpha}\end{pmatrix}$ telles que $|\alpha|^2+|\beta|^2 = 1$. Redémontrer ainsi que $G \cong \mathrm{SU}(2,\mathbf{C})$.

b) Donner, à l'aide des quaternions, une ("deuxième") preuve que le groupe $\mathrm{SO}(3,\mathbf{R})$ est simple. (indications : on pourra considérer H sous-groupe distingué de G, montrer que si H contient un quaternion pur imaginaire

de norme un, il les contient tous et que si H n'est pas contenu dans le centre $\{\pm 1\}$ il contient un tel quaternion).

E.6.4. Caractérisation et généralisation des quaternions.

On peut se demander si le corps des quaternions est un exemple unique, la réponse est oui si l'on souhaite conserver les propriétés d'associativité et d'existence d'un inverse. Cependant il existe d'autres algèbres de dimension finie sur $\mathbf{R}$ *intéressantes ; nous citons les octaves de Cayley (un "faux corps" au sens où tout élément non nul est inversible mais la multiplication n'est pas associative) et développons un peu les algèbres de Clifford associées à une forme quadratique (associatives mais en général non intègres).*

Thorme. (Frobenius) *Soit K un corps contenant dans son centre un sous-corps isomorphe à $\mathbf{R}$ et supposons K de dimension finie sur ce sous-corps, alors K est isomorphe à $\mathbf{R}$, $\mathbf{C}$ ou $\mathbf{H}$.*

Remarque. Si K est une $\mathbf{R}$-algèbre de dimension finie, c'est un corps si et seulement si elle est intègre (la multiplication par $a \in K$ est $\mathbf{R}$-linéaire donc injective si et seulement si elle est surjective). Si l'on enlève l'hypothèse de dimension finie on trouve d'autres corps comme $K = \mathbf{R}(X)$ le corps des fractions rationnelles.

Preuve. On identifie $\mathbf{R}$ et le sous-corps de $Z(K)$ isomorphe à $\mathbf{R}$. Commençons par quelques remarques préliminaires. Si a est un élément de K alors $\mathbf{R}(a)$ est un sous-corps commutatif de K et $[\mathbf{R}(a) : \mathbf{R}] = 1$ ou 2 avec ou bien $\mathbf{R}(a) = \mathbf{R}$ (si $a \in \mathbf{R}$) ou bien $\mathbf{R}(a) \cong \mathbf{C}$. De plus si $b \in C(a) := \{z \in K \mid za = az\}$ alors $\mathbf{R}(a, b)$ est un sous-corps commutatif et $[\mathbf{R}(a, b) : \mathbf{R}] = 1$ ou 2. Ainsi si $a \notin \mathbf{R}$, on a $C(a) = \mathbf{R}(a)$.

Si $K \neq \mathbf{R}$, alors K contient un sous-corps isomorphe à $\mathbf{C}$ et en particulier un élément i tel que $i^2 = -\mathbf{1}$. Si $K \neq \mathbf{R}(i) \cong \mathbf{C}$ nous allons utiliser le

Lemme. *Soit $b \in K \setminus \mathbf{R}(i)$ alors l'élément $c := bi - ib$ est non nul et vérifie $ic = -ci$ et $\exists r \in \mathbf{R}, c^2 = -r^2\mathbf{1}$; en particulier l'élément $\tilde{c} = cr^{-1}$ vérifie $i\tilde{c} = -\tilde{c}i$ et $\tilde{c}^2 = -\mathbf{1}$.*

Preuve du lemme. Comme $b \notin C(i)$, d'après les préliminaires, on a bien $c \neq 0$. Maintenant $ic = ibi + b$ alors que $ci = -b - ibi$ donc $ic = -ci$ et $c^2 i = ic^2$ donc $c^2 \in \mathbf{R}(i) \cap \mathbf{R}(c) = \mathbf{R}$. S'il existait $r \in \mathbf{R}$ tel que $c^2 = r^2$ (i.e. si $c^2 \in \mathbf{R}_+$) alors l'équation $X^2 - r^2 = 0$ possèderait au moins quatre racines $(\pm r, \pm c)$ dans $\mathbf{R}(c)$ corps commutatif, ce qui est impossible. Donc $c^2 < 0$ et le reste suit. □

Revenons à la démonstration du théorème. Si K n'est isomorphe ni à $\mathbf{R}$ ni à $\mathbf{C}$ alors il existe $b \in K \setminus \mathbf{R}(i)$ et le lemme construit un élément que nous notons j tel que $ij = -ji$ et $j^2 = -\mathbf{1}$. Notons donc $k := ij$ alors $K' := \mathbf{R}\mathbf{1} + \mathbf{R}i + \mathbf{R}j + \mathbf{R}k$ est un sous-corps de K isomorphe à $\mathbf{H}$. Montrons qu'en fait $K' = K$ et donc $K \cong \mathbf{H}$. Soit $b \in K$, si $b \in \mathbf{R}(i)$ alors $b \in K'$ et si $b \notin \mathbf{R}(i)$ le lemme permet de construire $c = bi - ib \neq 0$ tel que $ci = -ic$. Mais alors $(jc)i = j(-ic) = i(jc)$ donc $jc \in \mathbf{R}(i)$ et donc c est dans K'. Par ailleurs l'élément $d := bi + ib$ commute avec i donc est dans $\mathbf{R}(i)$ donc dans K' donc $2bi = c + d$ est dans K' et b également, ce qui achève la démonstration. □

Algèbre des octaves de Cayley.

Un présentation commode et rapide des *octaves de Cayley* est comme un espace vectoriel (disons à gauche) $\mathbf{Ca} = \mathbf{H} + \mathbf{H}\mathcal{L}$ de dimension 2 sur $\mathbf{H}$ (de dimension 8 sur $\mathbf{R}$) muni d'une base $1, \mathcal{L}$ et d'une multiplication (non associative)

$$(p_1 + p_2\mathcal{L})(q_1 + q_2\mathcal{L}) = p_1q_1 - \bar{q}_2p_2 + (q_2p_1 + p_2\bar{q}_1)\mathcal{L}$$

On définit la conjugaison comme $\overline{(p_1 + p_2\mathcal{L})} = \bar{p}_1 - p_2\mathcal{L}$ et la norme d'un octave de Cayley par la formule $\mathrm{N}(p_1 + p_2\mathcal{L}) = (p_1 + p_2\mathcal{L})\overline{(p_1 + p_2\mathcal{L})} = \mathrm{N}(p_1) + \mathrm{N}(p_2)$. On voit donc que l'inverse de $p_1 + p_2\mathcal{L}$ est $\mathrm{N}(p_1 + p_2\mathcal{L})^{-1}\overline{(p_1 + p_2\mathcal{L})}$. La multiplication n'est pas associative : par exemple $((j\mathcal{L})i)\,\mathcal{L} = -k \neq k = (j\mathcal{L})\,(i\mathcal{L})$, mais vérifie une propriété un peu plus faible (qu'on laisse vérifier en exercice) :

$$\forall q, r \in \mathbf{Ca}, \qquad q(qr) = (qq)r, \qquad (qr)q = q(rr) \qquad \text{et} \qquad q(rq) = (qr)q.$$

Exercice. Démontrer que la norme est multiplicative sur **Ca** (i. e. $\mathrm{N}(qq') = \mathrm{N}(q)\,\mathrm{N}(q')$) et en déduire une identité du type

$$\left(\sum_{i=1}^{8} x_i^2\right)\left(\sum_{i=1}^{8} y_i^2\right) = \sum_{i=1}^{8} B_i(x,y)^2$$

avec $B_i(x,y)$ formes bilinéaires en x, y.

Algèbres de Clifford.

Soit Q une forme quadratique sur un espace vectoriel E de dimension n sur un corps commutatif K de caractéristique $\neq 2$; on notera B la forme bilinéaire associée. On définit "à la main" l'algèbre de Clifford $C(Q) = C(E,Q)$ ainsi : on note $x \circ y$ le produit dans cette algèbre et on choisit $e_1, \ldots, e_n$ une base orthogonale (i.e. $Q(e_i + e_j) = Q(e_i) + Q(e_j)$ si $i \neq j$) et on prend comme générateurs de $C(Q)$ les éléments e_i avec les relations $e_i \circ e_j = -e_j \circ e_i$ (si $i \neq j$) et $e_i \circ e_i = Q(e_i)\mathbf{1}$. Pour chaque sous-ensemble I de l'intervalle $[1, \ldots, n]$ on ordonne les éléments $i_1 < i_2 < \ldots < i_r$ et on pose $e_I = e_{i_1} \circ e_{i_2} \ldots \circ e_{i_r}$ et $e_\emptyset = \mathbf{1}$. On peut alors décrire $C(Q)$ comme un espace vectoriel de dimension 2^n avec une base donnée par $\{e_I \mid I \subset [1, \ldots, n]\}$ et la règle de multiplication s'écrit

$$e_I \circ e_J = (-1)^{s(I,J)} \prod_{i \in I \cap J} Q(e_i) e_{I \Delta J}$$

avec $I \,\Delta\, J := (I \cup J) \setminus (I \cap J)$ et $s(I,J) = \mathrm{card}\{(i,j) \in I \times J \mid i > j\}$.

Remarque. En supposant connu le produit tensoriel, une définition plus intrinsèque (qui donne automatiquement l'associativité et l'indépendance par rapport au choix d'une base orthogonale) est la suivante. On définit $T^0(E) = K$, $T^n(E) = E \otimes_K E \otimes_K \ldots \otimes_K E$ (produit de n copies de E) et $T(E) = \oplus_{n \geq 0} T^n(E)$; ce dernier espace est une K-algèbre non commutative, le produit envoyant $T^m(E) \times T^n(E)$ dans $T^{m+n}(E)$ et en notant que, en général $e_1 \otimes e_2 \neq e_2 \otimes e_1$. On définit ensuite $I(Q)$ comme l'idéal bilatère engendré par les éléments $x \otimes x - Q(x)\mathbf{1}$ et $C(Q) = T(E)/I(Q)$.

Remarques. Dès qu'il existe $x \in E \setminus \{0\}$ tel que $Q(x) = 0$, l'algèbre $C(Q)$ n'est pas intègre ; en effet on a alors $x \circ x = Q(x)\mathbf{1} = 0$ alors que $x \neq 0$. Sur l'une ou l'autre des définitions, on voit qu'on a une injection $i : E \hookrightarrow C(Q)$. On peut montrer que $C(Q)$ est l'algèbre solution d'un problème universel : pour toute K-algèbre A munie d'un homomorphisme d'espace vectoriel $j : E \to A$ tel que $j(x)j(x) = Q(x)\mathbf{1}_A$, il existe un homomorphisme d'algèbre $\phi : C(Q) \to A$ tel que $j = \phi \circ i$. En particulier on peut identifier E à un K-sous-espace vectoriel de $C(Q)$. L'espace vectoriel engendré par les produits d'un nombre *pair* de vecteurs $e_1 \circ \ldots \circ e_{2m}$ est une sous-algèbre qu'on note $C^+(Q)$.

Soit alors $G_1 = \{\alpha \in C(E,Q)^* \mid \alpha E \alpha^{-1} \subset E\}$ et $G = G_1 \cap C^+(Q)$. Les ensembles G_1 et G forment des groupes et de plus on a un homomorphisme évident $\Phi : G_1 \to \mathrm{GL}(E)$. En fait on va voir que cette homomorphisme est à valeur dans $\mathrm{O}(E,Q)$. Pour cela on va utiliser deux résultats classiques vus au chapitre E.3 : soit $x \in E$ avec $Q(x) \neq 0$ alors il existe une isométrie notée s_x laissant fixe $(Kx)^\perp$ et changeant x en $-x$; elle est donnée par la formule $s_x(y) = y - 2\frac{B(x,y)}{Q(x)}x$ (vérification directe) et, de plus, ces isométries, appelées *symétries hyperplanes* engendrent $\mathrm{O}(Q)$ (théorème de Cartan-Dieudonné).

Soit $x, y \in E$ et $\alpha \in G_1$, on a $x \circ y + y \circ x = 2B(x,y)$ donc $\alpha \circ (x \circ y + y \circ x) \circ \alpha^{-1} = 2B(x,y)$ ou encore

$$2B(\alpha \circ x \circ \alpha^{-1}, \alpha \circ y \circ \alpha^{-1}) = (\alpha \circ x \circ \alpha^{-1}) \circ (\alpha \circ y \circ \alpha^{-1}) + (\alpha \circ y \circ \alpha^{-1}) \circ (\alpha \circ x \circ \alpha^{-1}) = 2B(x,y)$$

ce qui montre bien que $\Phi(\alpha)$ est une isométrie. L'analogue du lemme décrivant une rotation de $\mathbf{R}^3$ comme "$\Phi(q)$" est le suivant

Lemme. *Soit $x \in E$ non isotrope (i.e. $Q(x) \neq 0$) et s_x la symétrie hyperplane associée, alors $\Phi(x) = -s_x$. En particulier $\mathrm{SO}(Q)$ est contenu dans l'image $\Phi(G)$.*

Observons que $x \circ x = Q(x)\mathbf{1}$ donc $x^{-1} = Q(x)^{-1}x$ et comme $x \circ y + y \circ x = 2B(x,y)$, on en tire

$$x \circ y \circ x^{-1} = -y + 2B(x,y)x^{-1} = -y + 2\frac{B(x,y)}{Q(x)}x = -s_x(y).$$

Soit $r \in \mathrm{SO}(Q)$, alors $r = s_{x_1} \dots s_{x_{2m}}$ donc $r = \Phi(x_1 \circ \dots \circ x_{2m})$. $\square$

Pour calculer le noyau de $\Phi : G \to \mathrm{SO}(Q)$, il faut trouver les éléments inversibles commutant avec tout $x \in E$. Ce calcul est basé sur la formule suivante qui se vérifie directement à partir de la définition du produit de l'algèbre $C(Q)$:

$$e_I \circ e_J \circ e_I^{-1} = (-1)^{|I||J|-|I\cap J|} e_J$$

et on en tire

$$\{x \in C^+(Q) \mid \forall y \in E,\ x \circ y = y \circ x\} = K.$$

On obtient donc que $\Phi : G \to \mathrm{SO}(Q)$ est surjective de noyau K^*. Comme tout élément de $\mathrm{SO}(Q)$ est produit (d'un nombre pair de) symétries hyperplanes, on voit aussi que tout élément de G s'écrit $ax_1 \circ \dots \circ x_r$ avec $a \in K^*$ et $x_i \in E$.

L'analogue de la norme est le suivant : on définit une involution $x \mapsto \bar{x}$, de $C(Q)$ dans $C(Q)$ par la formule (pour $i_1 < \dots < i_r$)

$$\overline{e_{i_1} \circ \dots e_{i_r}} = e_{i_r} \circ \dots e_{i_1} = (-1)^{r(r-1)/2} e_{i_1} \circ \dots e_{i_r}$$

puis la *norme spinorielle*

$$\mathrm{N}_{sp}(\alpha) = \alpha \circ \bar{\alpha}$$

et on peut alors montrer

Lemme. *L'involution $x \mapsto \bar{x}$ est un anti-isomorphisme de $C(Q)$, pour les éléments de G (ou G_1) on a $\mathrm{N}_{sp}(\alpha \circ \beta) = \mathrm{N}_{sp}(\alpha)\,\mathrm{N}_{sp}(\beta)$.*

Preuve. La première affirmation est claire, la deuxième est un peu plus subtile (d'ailleurs l'énoncé n'est pas vrai pour tous les éléments de $C(Q)$). En fait un élément de G (ou G_1) s'écrit $\alpha = ax_1 \circ \dots \circ x_r$ (resp. $\beta = by_1 \circ \dots \circ y_r$) avec $a, b \in K$ et $x_i, y_i \in E$ donc $\mathrm{N}_{sp}(\alpha) = a^2 Q(x_1) \dots Q(x_r)$ (resp. $\mathrm{N}_{sp}(\beta) = b^2 Q(y_1) \dots Q(y_r)$), or $\mathrm{N}_{sp}(\alpha \circ \beta) = \alpha \circ \beta \circ \bar{\beta} \circ \bar{\alpha} = \mathrm{N}_{sp}(\beta)\alpha \circ \bar{\alpha} = \mathrm{N}_{sp}(\alpha)\,\mathrm{N}_{sp}(\beta)$. $\square$

Ceci permet de montrer que, si $x_1, \dots, x_r$ sont des vecteurs non isotropes de E et si $s_{x_1} \dots s_{x_r} = id_E$ alors $Q(x_1) \dots Q(x_r) \in K^{*2}$. En effet r doit être pair et donc $\Phi(x_1 \circ \dots \circ x_r) = id$ et donc $x_1 \circ \dots \circ x_r = a \in K^*$. Mais alors $a^2 = \mathrm{N}_{sp}(x_1 \circ \dots \circ x_r) = Q(x_1) \dots Q(x_r)$.

Introduisons maintenant $\Omega(Q) = \Omega(E, Q)$ le sous-groupe des *commutateurs* de $\mathrm{SO}(Q)$. On a démontré les parties *a*) et *b*) de l'énoncé suivant

Thorme. *a) L'application Φ induit un isomorphisme $G/K^* \to \mathrm{SO}(Q)$.*

b) Soit $\rho \in \mathrm{SO}(Q)$ décomposée en produit de symétries hyperplanes $\rho = s_{x_1} \dots s_{x_m}$, alors l'élément $M(\rho) := Q(x_1) \dots Q(x_m) \in K^*/K^{*2}$ ne dépend pas de la décomposition et l'application $M : \mathrm{SO}(Q) \to K^*/K^{*2}$ est un homomorphisme de groupes dont le noyau contient $\Omega(E, Q)$.

c) Supposons de plus qu'il existe $x \in E \setminus \{0\}$ tel que $Q(x) = 0$ (on dit que la forme est isotrope) alors la norme spinorielle induit un isomorphisme $\mathrm{SO}(E, Q)/\Omega(E, Q) \cong K^*/K^{*2}$. Si, de plus $n \geq 5$, alors $\mathrm{P}\Omega(E, Q) := \Omega(E, Q)/Z(\Omega(E, Q))$ est simple.

Plus exactement, on a démontré la première affirmation et la deuxième découle des remarques précédentes. Il est clair, puisque K^*/K^{*2} est commutatif que $\Omega(Q) \subset \mathrm{Ker}(M)$. Pour le point *c*) on renvoie aux livres d'Artin (*Algèbre géométrique*) et Dieudonné (*Géométrie des groupes classiques*). On notera qu'en général, l'hypothèse de l'existence d'un vecteur isotrope (i.e. d'un $x \neq 0$ tel que $Q(x) = 0$) est indispensable aux conclusions de *c*) (Cf ibidem), néanmoins dans le cas où $K = \mathbf{R}$ et Q est la forme quadratique définie positive, on a vu que le groupe $\mathrm{PO}(E, Q) = \mathrm{PO}(n, \mathbf{R})$ est simple dès que $n = \dim(E) = 3$ ou ≥ 5.

F. REPRÉSENTATIONS DES GROUPES FINIS.

Une *représentation* d'un groupe G est un homomorphisme ρ de G vers $\mathrm{GL}(E)$ où E est un K-espace vectoriel; par abus on parlera de la représentation E si le contexte est sans ambigüité. On peut voir ρ comme une action linéaire de G sur E. On s'intéressera exclusivement au cas où E est de dimension finie et principalement au cas où G est fini et K est de caractéristique zéro (voir ci-dessous pourquoi).

Une variante consiste à introduire l'*algèbre de groupe* $K[G]$ qui est l'algèbre ayant pour ensemble sous-jacent l'ensemble des fonctions de G dans K muni de la somme $(f+g)(x) = f(x)+g(x)$ et du *produit de convolution* :

$$(f * g)(x) = \sum_{y \in G} f(y)g(y^{-1}x).$$

En notant $e_g(x) = 1$ si $x = g$ et 0 sinon, on a $K[G] = \oplus_{g\in G} K e_g$ et le produit d'algèbre s'écrit $e_g \cdot e_{g'} = e_{gg'}$. L'algèbre $K[G]$ est commutative si et seulement si G est commutatif, en fait plus précisément on vérifie aisément que f est dans le centre de $K[G]$ si et seulement si $f(hgh^{-1}) = f(g)$ c'est-à-dire f est constante sur les classes de conjugaison (on dit que f est *centrale*). Une représentation est simplement un $K[G]$-module!

F.1. Généralités.

Donnons deux exemples pour commencer. Un homomorphisme $\phi : G \to \mathbf{C}^*$ est une représentation de dimension 1, son image est un groupe fini cyclique. Ensuite on peut définir $\rho : \mathcal{S}_n \to \mathrm{GL}(n, K)$ par $\rho(\sigma)(e_i) = e_{\sigma(i)}$.

Une représentation est dite *fidèle* si $\mathrm{Ker}(\rho) = \{e\}$; remarquons que ρ induit toujours une représentation fidèle de $G/\mathrm{Ker}(\rho)$.

Un *homomorphisme* $f : E_1 \to E_2$ entre deux représentations $\rho_i : G \to \mathrm{GL}(E_i)$ est un homomorphisme de $K[G]$-module, ou, en d'autres termes une application K-linéaire telle que pour tout $g \in G$ on ait $\rho_2(g) \circ f = f \circ \rho_1(g)$. L'ensemble des homomorphismes forment un groupe noté $\mathrm{Hom}_G(E_1, E_2)$; l'ensemble des endomorphismes d'une représentation E forme un anneau noté $\mathrm{End}_G(E)$. Un *isomorphisme de représentations* est un homomorphisme bijectif.

La *somme* de deux représentations $\rho_i : G \to \mathrm{GL}(E_i)$ est la représentation $\rho : G \to \mathrm{GL}(E_1 \oplus E_2)$ définie par $\rho(g)(x_1 + x_2) = \rho_1(g)(x_1) + \rho_2(g)(x_2)$. Si A_i est la matrice de $\rho_i(g)$ dans une base de E_i, la matrice de $\rho(g)$ dans la base de E obtenu en réunissant les vecteurs des bases de E_1, E_2 est $\begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$.

On note $\check{E}$ le dual de E. La *représentation duale* ou *contragrédiente* d'une représentation $\rho : G \to \mathrm{GL}(E)$ est la représentation $\check{\rho} : G \to \mathrm{GL}(\check{E})$ définie par

$$\forall x \in E,\ \forall \check{x} \in \check{E},\ \ (\check{\rho}(g)(\check{x}))\,(\rho(g)(x)) = \check{x}(x) \qquad \text{ou encore} \qquad \check{\rho}(g)(\check{x}) = \check{x} \circ \rho(g)^{-1}.$$

Si A est la matrice de $\rho(g)$ dans une base de E, la matrice de $\check{\rho}(g)$ dans la base duale est ${}^tA^{-1}$.

La *représentation régulière* de G est définie ainsi : on prend comme espace $E := K[G] = \oplus_{g\in G} K e_g$ et comme action $\rho(g)(e_h) = e_{gh}$. Dans la base des e_g les matrices des $\rho(g)$ sont des matrices de permutation.

Une *sous-représentation* est un sous-espace F de E stable sous l'action de G (c'est un $K[G]$-sous-module). La représentation quotient notée $\bar{\rho}$ ou $\bar{\rho}_{E/F}$ est la représentation qu'on obtient par action sur E/F (c'est le $K[G]$-module quotient). Si on choisit F' un supplémentaire (non nécessairement G-invariant) et une base de E respectant la décomposition $E = F \oplus F'$, si A est la matrice de $\rho_F(g)$ dans la base de F et B la matrice de $\bar{\rho}_{E/F}(g)$ dans la base de E/F déduite de celle de F' alors la matrice de $\rho(g)$ est de la forme $\begin{pmatrix} A & * \\ 0 & B \end{pmatrix}$.

Par exemple $E^G := \{x \in E \mid \forall g \in G,\ \rho(g)(x) = x\}$ est une sous-représentation.

Remarque. Il n'est pas vrai en général que E soit isomorphe à $F \oplus E/F$ comme le montre l'exemple du groupe T des matrices triangulaires de $\mathrm{GL}(2, K)$ agissant sur $E = K^2$. Il y a un sous-espace stable (une droite) mais aucun supplémentaire n'est stable.

Une représentation est *irréductible* si elle n'admet aucune sous-représentation autre qu'elle-même et la représentation nulle. Une représentation de dimension 1 est clairement irréductible, nous verrons plus loin que tout groupe non abélien admet au moins une représentation irréductible de dimension ≥ 2. Les deux énoncés suivants expliquent en grande partie l'importance de cette notion.

Lemme. (Lemme de Schur) *Soit $f : E_1 \to E_2$ un homomorphisme entre deux représentations irréductibles $\rho_i : G \to \mathrm{GL}(E_i)$, alors*
(i) Si les deux représentations ne sont pas isomorphes, alors $f = 0$.
(ii) Si $E_1 = E_2$ et $\rho_1 = \rho_2$, alors f est une homothétie.

Preuve. On observe que $\mathrm{Ker}(f)$ est une sous-représentation de E_1, et $\mathrm{Im}(f)$ une sous-représentation de E_2 donc ou bien $\mathrm{Ker}(f) = E_1$ et f est nulle, ou bien $\mathrm{Ker}(f) = \{0\}$ et f est injective. Dans le deuxième cas, $\mathrm{Im}(f)$ est non nul donc égal à E_2. Pour prouver le point (ii) on remplace f par $f - \lambda id$ avec λ valeur propre de f; comme $\mathrm{Ker}(f - \lambda id) \neq \{0\}$ on conclut que $\mathrm{Ker}(f - \lambda id) = E_1$ et donc $f = \lambda id$. □

Théorème. *Toute représentation d'un groupe fini G sur un corps de caractéristique première à $\mathrm{card}(G)$ se décompose en somme de représentations irréductibles.*

Preuve. On a recours au procédé de la moyenne dû à Weyl. Une première méthode, lorsque $K = \mathbf{R}$ ou $\mathbf{C}$ est d'introduire sur E un produit scalaire invariant par G. Précisément, si $(.,.)$ est un produit scalaire ou hermitien sur E, on pose

$$(x, y)_G := \sum_{g \in G} (\rho(g)(x), \rho(g)(y)).$$

On voit immédiatement que c'est de nouveau un produit scalaire ou hermitien et que $\rho(g)$ est une isométrie pour ce produit scalaire. Si F est une sous-représentation de E, on constate que $F' := F^{\perp}$ est invariant par $\rho(G)$ et $E = F \oplus F'$; bien entendu l'orthogonal est pris au sens du produit scalaire $(.,.)_G$. Une variante de ce procédé consiste à construire un projecteur G-invariant de E sur F. Si $p : E \to F$ est un projecteur quelconque de E sur F (i.e. pour $x \in F$ on a $p(x) = x$), posons $p_G = \frac{1}{\mathrm{card}(G)} \sum_{g \in G} \rho(g) p \rho(g)^{-1}$, on vérifie que, si $x \in F$ alors $p_G(x) = x$ puisque $\rho(g)(x) \in F$ pour tout $g \in G$. Le noyau du projecteur p_G fournit alors le supplémentaire cherché. On remarquera que cette variante nécessite uniquement comme hypothèse que $\mathrm{card}(G)$ soit inversible dans K. □

Remarques. a) L'exemple où G est le groupe des matrices 2×2 triangulaires supérieures à coefficients dans $\mathbf{F}_p$ agissant sur $V = \mathbf{F}_p^2$ montre que l'hypothèse du théorème n'est pas superflue. En effet, il y a bien dans ce cas une sous-représentation de dimension 1 mais pas de supplémentaire stable.

b) On peut traduire cet énoncé en disant que les $K[G]$-modules de type fini sont semi-simples. Soit $\rho : G \to \mathrm{GL}(E)$ une représentation, il existe des entiers $m_i \geq 1$ et des représentations irréductibles $E_1, \ldots, E_r$ deux à deux non isomorphes telles que $E \cong E_1^{m_1} \oplus \ldots \oplus E_r^{m_r}$. On dit que m_i est la *multiplicité* de E_i dans la représentation E (on verra plus loin que cette multiplicité est intrinsèque) et, dans ce cas, le lemme de Schur se traduit en le fait que

$$\mathrm{End}_G(E) \cong \mathrm{End}(E_1^{m_1}) \times \ldots \times \mathrm{End}(E_r^{m_r}) \cong \mathrm{Mat}(m_1 \times m_1, K) \times \ldots \times \mathrm{Mat}(m_r \times m_r, K).$$

On peut par exemple se demander quelle est la décomposition de la représentation régulière. Cette question est résolue plus loin à l'aide de la théorie des caractères mais notons tout de suite que la représentation régulière n'est jamais irréductible (sauf si $G = \{e\}$) puisque, si $e_G = \sum_{g \in G} e_g$, la droite $F = K e_G$ définit une sous-représentation triviale.

Le *produit tensoriel* de deux représentations $\rho_i : G \to \mathrm{GL}(E_i)$ est la représentation $\rho : G \to \mathrm{GL}(E_1 \otimes E_2)$ définie par

$$\rho(g)(x_1 \otimes x_2) := (\rho_1(g)(x_1)) \otimes (\rho_2(g)(x_2)).$$

Si $A = (a_{ij})$ (resp. $B = (b_{ij})$) est la matrice de $\rho_1(g)$ dans une base e_i (resp. de $\rho_2(g)$ dans une base f_j) alors la matrice de ρ dans la base $e_i \otimes f_j$ est le produit tensoriel des matrices A et B i.e. $c_{i_1,j_1;i_2,j_2} = a_{i_1 j_1} b_{i_2 j_2}$. En se rappelant que $\mathrm{Hom}(E, F) = \check{E} \otimes F$ on voit que si $\rho : G \to \mathrm{GL}(E)$ et $\rho' : G \to \mathrm{GL}(F)$ sont des représentations

de G on obtient une représentation de G dans $\mathrm{Hom}(E,F)$ en tensorisant la représentation contragrédiente de ρ par ρ'. On peut l'écrire explicitement : si $f \in \mathrm{Hom}(E,F)$, on a $(\check{\rho} \otimes \rho')(g)(f) = \rho'(g) \circ f \circ \rho(g^{-1})$. En particulier les éléments invariants de cette représentation sont les homomorphismes de représentation (i.e. les $f \in \mathrm{Hom}(E,F)$ tels que $\rho'(h) \circ f = f \circ \rho(h)$), en d'autres termes $\mathrm{Hom}(E,F)^G = \mathrm{Hom}_G(E,F)$.

Exercice. Soit $\rho : G \to \mathrm{GL}(E)$ une représentation de dimension n en caractéristique $\neq 2$ et $\rho^{(2)} : G \to \mathrm{GL}(E \otimes E)$ la représentation produit tensoriel de deux copies de ρ. Soit $\delta : E \otimes E \to E \otimes E$ linéaire telle que $\delta(x_1 \otimes x_2) = x_2 \otimes x_1$ et F^+ (resp; F^-) le sous-espace des éléments invariants (resp. anti-invariants) de l'involution δ. Montrer que F^+ et F^- sont des sous-représentations de E de dimensions respectivement $n(n+1)/2$ et $n(n-1)/2$ et que $E = F^+ \oplus F^-$. La représentation F^+ (resp. F^-) s'appelle le *carré symétrique* (resp. le *carré alterné*) et se note souvent $\mathrm{Sym}^2(E)$ (resp. $\Lambda^2(E)$).

On peut décrire les représentations irréductibles (et donc les autres) de $G_1 \times G_2$ à partir de celles de G_1 et G_2. Tout d'abord si $\rho_i : G_i \to \mathrm{GL}(E_i)$ sont des représentations de G, on définit $\rho_1 \otimes \rho_2 : G_1 \times G_2 \to \mathrm{GL}(E_1 \otimes E_2)$ par la formule $(\rho_1 \otimes \rho_2)(g_1, g_2) = \rho_1(g_1) \otimes \rho_2(g_2)$. On peut remarquer que l'application de $K[G_1] \otimes K[G_2]$ vers $K[G_1 \times G_2]$ définie par $e_{g_1} \otimes e_{g_2} \mapsto e_{(g_1,g_2)}$ est un isomorphisme. On a alors

Proposition. *La représentation $\rho_1 \otimes \rho_2$ est une représentation irréductible si et seulement si ρ_1 et ρ_2 sont irréductibles. De plus toute représentation irréductible de $G_1 \times G_2$ est de cette forme.*

La preuve est plus facile une fois qu'on a développé la théorie des caractères et est donc renvoyée au paragraphe suivant. □

Décrivons toutes les représentations d'un groupe abélien fini, en supposant pour simplifier que le corps de base est $\mathbf{C}$. D'après ce qui précède, il suffit de considérer les représentations de $\mathbf{Z}/m\mathbf{Z}$. Les matrices $\rho(g)$ sont toutes diagonalisables car leurs polynômes minimaux divisent $X^{\mathrm{card}(G)} - 1$ et possèdent donc des racines simples; elles sont mêmes simultanément diagonalisables. Ainsi, à changement de base près, il existe n homomorphismes $\phi_i : G \to \mathbf{C}^*$ tels que

$$\rho(g) = \begin{pmatrix} \phi_1(g) & 0 & & \\ 0 & \phi_2(g) & & \\ & & \ddots & \\ & & & \phi_n(g) \end{pmatrix}.$$

En particulier les représentations irréductibles sont celles de dimension 1 et la représentation régulière de G est la somme directe de toutes les représentations irréductibles de G (avec multiplicité 1).

F.2. Caractère d'une représentation.

On suppose dans tout ce paragraphe que $K = \mathbf{C}$.

Définition. Le *caractère* d'une représentation est l'application $\chi_\rho : G \to \mathbf{C}$ donnée par $\chi_\rho(g) = \mathrm{Tr}\,\rho(g)$.

On note tout de suite quelques propriétés évidentes : $\chi_\rho(e) = \dim(\rho)$, $\chi_\rho(g^{-1}) = \overline{\chi_\rho(g)}$ et χ_ρ est constant sur les classes de conjugaison. De plus on a les formules :

(i) $\chi_{\rho_1 \oplus \rho_2}(g) = \chi_{\rho_1}(g) + \chi_{\rho_2}(g)$.
(ii) $\chi_{\rho_1 \otimes \rho_2}(g) = \chi_{\rho_1}(g)\chi_{\rho_2}(g)$.
(iii) $\chi_{\check{\rho}}(g) = \chi_\rho(g^{-1})$.
(iv) Si ρ est isomorphe à ρ' alors $\chi_\rho = \chi_{\rho'}$.

Preuve. On a $\chi_\rho(e) = \mathrm{Tr}\, id_E = \dim(E)$. Ensuite les matrices $\rho(g)$ sont diagonalisables avec pour valeurs propres des racines de l'unité donc les valeurs propres de $\rho(g^{-1})$ sont les conjuguées d'icelles et l'on en déduit bien $\chi_\rho(g^{-1}) = \overline{\chi_\rho(g)}$. Par ailleurs $\chi_\rho(hgh^{-1}) = \mathrm{Tr}(\rho(h)\rho(g)\rho(h)^{-1}) = \mathrm{Tr}\,\rho(g) = \chi_\rho(g)$. Les formules (i), (ii) et (iii) découlent du fait que $\mathrm{Tr}\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} = \mathrm{Tr}\,A + \mathrm{Tr}\,B$, que $\mathrm{Tr}\,A \otimes B = (\mathrm{Tr}\,A)(\mathrm{Tr}\,B)$ et que $\mathrm{Tr}({}^tA) = \mathrm{Tr}\,A$. Enfin si $\rho'(g) = f \circ \rho(g) \circ f^{-1}$ on a $\chi_{\rho'}(g) = \mathrm{Tr}\,\rho'(g) = \mathrm{Tr}(f \circ \rho(g) \circ f^{-1}) = \mathrm{Tr}\,\rho(g) = \chi_\rho(g)$. □

On introduit maintenant un produit scalaire sur les fonctions de G dans $\mathbf{C}$:

$$(f, f') := \frac{1}{\mathrm{card}(G)} \sum_{g \in G} f(g)\overline{f'(g)}.$$

Une des propriétés essentielles de ce produit scalaire est la suivante :

Proposition. *Soit χ_1, χ_2 les caractères de deux représentations irréductibles de G. On a $(\chi_1, \chi_2) = 0$ si les deux représentations ne sont pas isomorphes et $(\chi_1, \chi_2) = 1$ si les deux représentations sont isomorphes.*

Preuve. Considérons $\pi := \frac{1}{\mathrm{card}(G)} \sum_{g \in G} \rho(g)$, on voit que π est un projecteur de E sur E^G (en effet pour tout $x \in E$ on a que $\pi(x)$ est invariant par G et si $x \in E^G$ alors $\pi(x) = x$) et donc que sa trace vaut $\dim(E^G)$. On a donc montré que

$$\frac{1}{\mathrm{card}(G)} \sum_{g \in G} \chi_\rho(g) = \dim(E^G).$$

On applique alors cela à $\rho = \check{\rho_1} \otimes \rho_2$ et on en tire

$$(\chi_2, \chi_1) = \frac{1}{\mathrm{card}(G)} \sum_{g \in G} \chi_1(g^{-1})\chi_2(g) = \frac{1}{\mathrm{card}(G)} \sum_{g \in G} \chi_\rho(g) = \dim(\mathrm{Hom}(E_1, E_2)^G) = \dim(\mathrm{Hom}_G(E_1, E_2)).$$

Mais le lemme de Schur indique que la dernière dimension est nulle si $\chi_1 \neq \chi_2$ et vaut un si $E_1 \cong E_2$. □

Corollaire. *Soit $\rho : G \to \mathrm{GL}(E)$ une représentation de caractère χ_ρ et soit χ_ψ le caractère d'une représentation irréductible $\psi : G \to \mathrm{GL}(F)$. La multiplicité de la représentation F dans E est égale à (χ_ρ, χ_ψ).*

Preuve. En effet écrivons une décomposition de E en somme d'irréductibles $E = E_1^{m_1} \oplus \ldots \oplus E_r^{m_r}$ et notons χ_i le caractère de E_i. On a donc $\chi_\rho = m_1\chi_1 + \ldots + m_r\chi_r$. D'après la proposition précédente, on a

$$(\chi_\rho, \chi_\psi) = m_1(\chi_1, \chi_\psi) + \ldots + m_r(\chi_r, \chi_\psi) = \begin{cases} 0 & \text{si } \chi_\psi \neq \chi_1, \ldots, \chi_r \\ m_i & \text{si } \chi_\psi = \chi_i \end{cases}$$

d'où le résultat. □

Remarque. En particulier cet énoncé permet de vérifier que la multiplicité ne dépend pas de la décomposition de la représentation E.

On en déduit facilement l'important résultat suivant :

Théorème. *Deux représentations sont isomorphes si et seulement si elles ont mêmes caractères.*

Preuve. En effet si $\chi_\rho = \chi_{\rho'}$ alors les deux représentations contiennent une représentation irréductible de caractère χ avec la même multiplicité $(\chi_\rho, \chi) = (\chi_{\rho'}, \chi)$ et sont donc isomorphes à la même somme de représentations irréductibles. □

On peut aussi observer que si $E = E_1^{m_1} \oplus \ldots \oplus E_r^{m_r}$ est une décomposition de la représentation ρ en irréductibles deux à deux non isomorphes alors $(\chi_\rho, \chi_\rho) = m_1^2 + \ldots m_r^2$ et en déduire l'énoncé suivant :

Proposition. *Soit $\rho : G \to \mathrm{GL}(E)$ une représentation, alors (χ_ρ, χ_ρ) est un entier strictement positif qui est égal à 1 si et seulement si ρ est irréductible.*

Nous sommes maintenant en mesure de calculer la décomposition de la représentation régulière.

Théorème. *Soit $Irr(G)$ l'ensemble des représentations irréductibles de G (à isomorphismes près), si χ est le caractère d'une de ces représentations, on note $\chi(e) = m_\chi$ sa dimension. On a lors*

$$\chi_{\mathrm{reg}} = \sum_{\chi \in Irr(G)} m_\chi \chi$$

ou encore

$$\mathrm{Reg}_G = \oplus_{E\in Irr(G)} E^{m_E}$$

et en particulier

$$|G| = \sum_{\chi\in Irr(G)} m_\chi^2.$$

Preuve. On voit directement que $\chi_{\mathrm{reg}}(g) = 0$ si $g \in G \setminus \{e\}$ (resp. $\chi_{\mathrm{reg}}(e) = \mathrm{card}(G)$). En effet pour $g \neq e$ la matrice de $\rho(g)$ est une matrice de permutation sans point fixe. On en tire

$$(\chi_{\mathrm{reg}}, \chi) = \frac{1}{\mathrm{card}(G)} \sum_{g\in G} \chi_{\mathrm{reg}}(g^{-1})\chi(g) = \frac{1}{\mathrm{card}(G)}(\mathrm{card}(G)\chi(e)) = \chi(e) = m_\chi$$

et on applique les résultats précédents. □

Exemple. Le nombre de représentation de G de dimension 1 est $(G : D(G))$ (où $D(G)$ désigne le sous-groupe des commutateurs). En effet $\mathrm{GL}(1)$ est commutatif et donc une telle représentation se factorise par $G/D(G)$.

Application. Nous sommes en mesure de démontrer les résultats annoncés au paragraphe précédent sur les représentations de $G_1 \times G_2$. Soient $\rho_i : G_i \to \mathrm{GL}(E_i)$ deux représentations des groupes G_i de cardinal N_i et χ_i leurs caractères respectifs, le caractère de $\rho = \rho_1 \otimes \rho_2$ est donné par $\chi(g_1, g_2) = \chi_1(g_1)\chi_2(g_2)$ donc

$$\begin{aligned}(\chi, \chi) &= \frac{1}{N_1 N_2} \sum_{(g_1,g_2)\in G_1\times G_2} \chi(g_1, g_2)\overline{\chi(g_1, g_2)} \\ &= \left(\frac{1}{N_1}\sum_{g_1\in G_1} \chi_1(g_1)\overline{\chi_1(g_1)}\right)\left(\frac{1}{N_2}\sum_{g_2\in G_2} \chi_2(g_2)\overline{\chi_2(g_2)}\right) \\ &= (\chi_1, \chi_1)(\chi_2, \chi_2).\end{aligned}$$

On en déduit que ρ est irréductible si et seulement si $(\chi, \chi) = 1$ donc si et seulement si $(\chi_1, \chi_1) = (\chi_2, \chi_2) = 1$ donc si et seulement si ρ_1 et ρ_2 sont irréductibles. Remarquons que l'application de $Irr(G_1) \times Irr(G_2)$ vers $Irr(G_1 \times G_2)$ définie par $(\rho_1, \rho_2) \mapsto \rho_1 \otimes \rho_2$ est injective car le caractère χ_1 (resp. χ_2) se récupère à partir de χ par la formule $\chi_1(g_1) = \chi(g_1, 1)$ (resp. $\chi_2(g_2) = \chi(1, g_2)$). Enfin on vérifie que l'application est également surjective car

$$\sum_{\rho_1\in Irr(G_1),\rho_2\in Irr(G_2)} m^2_{\rho_1\otimes\rho_2} = \sum_{\rho_1\in Irr(G_1)} m^2_{\rho_1} \sum_{\rho_2\in Irr(G_2)} m^2_{\rho_2} = |G_1| \cdot |G_2| = \sum_{\rho\in Irr(G_1\times G_2)} m^2_\rho.$$

Nous allons montrer deux autres résultats sur les représentations irréductibles.

Théorème. *Les représentations irréductibles sont en bijection avec les classes de conjugaison.*

Preuve. Considérons la sous-algèbre A de $\mathbf{C}[G]$ consituée des fonctions *centrales* i.e. constantes sur les classes de conjugaison, sa dimension sur $\mathbf{C}$ est égale au nombre de classes de conjugaison (c'est aussi le centre de l'algèbre $\mathbf{C}[G]$). Les caractères $\chi_1, \ldots, \chi_r$ des représentations irréductibles de G sont dans A et forment une famille orthonormale. Montrons qu'ils fournissent une base de A, ce qui achèvera la preuve. Il suffit de montrer que si $f : G \to \mathbf{C}$ est une fonction dans A orthogonale aux χ_i alors f est nulle. Introduisons, pour toute représentation $\rho : G \to \mathrm{GL}(E)$ l'application linéaire $\rho_f = \sum_{g\in G} f(g)\rho(g)$. On peut calculer ρ_f à l'aide du lemme suivant

Lemme. *Si ρ est de dimension n, irréductible de caractère χ alors ρ_f est une homothétie de rapport*

$$\frac{1}{n}\sum_{g\in G} f(g)\chi(g) = \frac{\mathrm{card}(G)}{n}(f, \bar{\chi}).$$

Preuve du lemme. On commence par montrer que ρ_f est un endomorphisme de la représentation, en effet :

$$\rho(h)^{-1}\rho_f\rho(h) = \sum_{g\in G} f(g)\rho(h)^{-1}\rho(g)\rho(h) = \sum_{g\in G} f(g)\rho(h^{-1}gh) = \sum_{k\in G} f(hkh^{-1})\rho(k) = \rho_f.$$

Le lemme de Schur garantit donc que ρ_f est une homothétie et sa trace vaut

$$\operatorname{Tr}\rho_f = \sum_{g\in G} f(g)\operatorname{Tr}\rho(g) = \sum_{g\in G} f(g)\chi(g)$$

d'où le résultat. □

La preuve montre que $\mathrm{card}(Irr(G))$ est égal au nombre de classes de conjugaison de G car chacun des deux ensembles est en bijection naturelle avec les éléments d'une base d'un même espace vectoriel, mais ne fournit pas de bijection naturelle entre ces classes et les représentations irréductibles; en fait on ne connait de telle bijection que pour certains groupes particuliers (par exemple les groupes $\mathcal{S}_n$). Revenons au cas où f est orthogonale aux caractères des représentations irréductibles donc à tous les caractères; on voit donc que $\rho_f = 0$ pour toutes les représentations et en particulier pour la représentation régulière. Ceci entraîne que, si ρ est la représentation régulière, on a

$$0 = \rho_f(e_h) = \sum_{g\in G} f(g)\rho(g)(e_h) = \sum_{g\in G} f(g)e_{gh}.$$

Puisque les e_g sont linéairement indépendants, on en déduit bien que $f(g) = 0$ pour tout $g \in G$. □

Corollaire. *Un groupe G est abélien si et seulement si toutes ses représentations irréductibles sont de dimension 1.*

Preuve. On a déjà vu que si G est abélien, alors ses représentations irréductibles sont de dimension 1. Inversement si ses représentations irréductibles sont de dimension 1, on a $\mathrm{card}(G) = 1^2 + \ldots + 1^2$ donc il y a $\mathrm{card}(G)$ représentations irréductibles, donc autant de classes de conjugaison, ce qui signifie que G est abélien. □

Théorème. *Soit m la dimension d'une représentation irréductible de G, alors*
(i) L'entier m divise $\mathrm{card}(G)$.
(ii) Soit A un sous-groupe abélien de G, alors $m \leq (G : A)$.

Preuve. Le point (i) est prouvé plus loin à l'aide de considérations d'intégralité. Pour le point (ii) considérons la restriction $\rho_A : A \to \mathrm{GL}(E)$; c'est une représentation de groupe abélien donc elle contient une sous-représentation F de dimension 1. Soit maintenant $g_1, \ldots, g_s$ des représentants de G/A (en particulier $s = (G : A)$) alors $E' = \rho(g_1)(F) + \ldots + \rho(g_s)(F)$ est stable par $\rho(G)$ car gg_i s'écrit aussi $g_j h$ avec $h \in A$ donc $\rho(g)\rho(g_i)(F) = \rho(g_j h)(F) = \rho(g_j)(F)$. On a donc $E = E'$ et comme la dimension de E' est $\leq s$ on a bien démontré l'énoncé. □

Exemples. Donnons maintenant pour quelques "petits" groupes finis G une description des représentations irréductibles et de leurs caractères.

(a) Si G est abélien de cardinal n, il existe n homomorphismes différents $G \to \mathbf{C}^* = \mathrm{GL}(1, \mathbf{C})$ et ainsi $n = 1^2 + \ldots + 1^2$. Si $G = \mathbf{Z}/n\mathbf{Z}$, ces homomorphismes s'écrivent $\phi_k(m) = \exp(2i\pi km/n)$. Plus généralement, si $G = \mathbf{Z}/a_1\mathbf{Z} \times \ldots \times \mathbf{Z}/a_r\mathbf{Z}$, ces homomorphismes s'écrivent

$$\phi_k(m_1, \ldots, m_r) = \exp\left(2i\pi\left(\frac{k_1 m_1}{a_1} + \ldots + \frac{k_r m_r}{a_r}\right)\right) \qquad \text{pour} \quad 0 \leq k_i \leq a_i - 1.$$

(b) Si $G = \mathcal{S}_3$, on a deux représentations de dimension 1, la représentation triviale et la signature. Il y a trois classes de conjugaison donc une autre représentation qui doit être de dimension 2 (puisque

$6 = 1^2+1^2+2^2$). On peut décrire cette représentation comme la sous-représentation de la représentation de permutation $\rho : \mathcal{S}_3 \to \mathrm{GL}(3, \mathbf{C})$ définie par le plan d'équation $x_1 + x_2 + x_3 = 0$.

(c) Si $G = D_n$, alors G contient un sous-groupe abélien (cyclique) d'indice 2 donc les représentations irréductibles sont de dimension 1 ou 2. On a $G/D(G) \cong \mathbf{Z}/2\mathbf{Z}$ si n est impair (resp. $\cong (\mathbf{Z}/2\mathbf{Z})^2$) si n est pair). Ainsi le nombre de représentations irréductibles de dimension 2 est $(2n-2)/4 = (n-1)/2$ si n est impair et $(2n-4)/4 = n/2-1$ si n est pair. Notons $x \in G$ une rotation d'ordre n et $y \in G$ une symétrie (donc $x^n = y^2 = e$ et $yxy = yxy^{-1} = x^{-1}$), on peut décrire les représentations de dimension 2 par les formules

$$\rho_k(x) = \begin{pmatrix} \exp(2k\pi i/n) & 0 \\ 0 & \exp(-2k\pi i/n) \end{pmatrix} \qquad \text{et} \qquad \rho_k(y) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Si n est impair on choisit $k = 1, \ldots, (n-1)/2$; si n est pair on choisit $k = 1, \ldots, (n-2)/2$. Remarquons que les valeurs des caractères s'écrivent

$$\chi_k(x^a) = 2\cos(2ka\pi/n) \qquad \text{et} \qquad \chi_k(yx^a) = 0.$$

(d) Plus généralement si p, q sont premiers tels que $q = mp + 1$, soit $G = \mathbf{Z}/q\mathbf{Z} \rtimes_\phi \mathbf{Z}/p\mathbf{Z}$ le groupe non commutatif de cardinal pq. Les représentations de G ont une dimension inférieure à p et qui divise pq donc égale à 1 ou p. Par ailleurs $G/D(G) \cong \mathbf{Z}/p\mathbf{Z}$ donc il y a p représentations de dimension 1 et m représentations de dimension p. On a bien $pq = p1^2 + mp^2$. On peut vérifier qu'il y a bien $m+p$ classes de conjugaison : la classe du neutre, les éléments d'ordre p se répartissent en $p-1$ classes de cardinal q, les éléments d'ordre q se répartissent en m classes de cardinal p. Pour construire les représentations de dimension p on peut procéder comme suit. On note $x \in G$ un générateur du sous-groupe d'ordre q et y un élément d'ordre p de sorte que $yxy^{-1} = x^u$ où u entier qui est d'ordre p dans $(\mathbf{Z}/q\mathbf{Z})^*$; on choisit α racine q-ième de l'unité et on pose :

$$\rho_\alpha(x) = \begin{pmatrix} \alpha & & & & \\ 0 & \alpha^u & & & \\ & & \ddots & & \\ & & & \alpha^{u^{p-2}} & \\ & & & & \alpha^{u^{p-1}} \end{pmatrix} \qquad \text{et} \qquad \rho_\alpha(y) = \begin{pmatrix} 0 & & & & 1 \\ 1 & 0 & & & \\ & & \ddots & & \\ & & & 0 & \\ 0 & & & 1 & 0 \end{pmatrix}.$$

Les caractères de ces représentations s'écrivent, en notant $\alpha = \exp(2k\pi i/q)$,

$$\chi_k(x^a) = \sum_{j=0}^{p-1} \exp(2kau^j\pi i/q) \qquad \text{et} \qquad \chi_k(y^b x^a) = 0.$$

On obtient un ensemble non redondant en choisissant $k_1, \ldots, k_m$ tels que l'ensemble des $k_i u^j$ (pour $i = 1, \ldots, m$ et $j = 0, \ldots, p-1$) décrive tout $(\mathbf{Z}/q\mathbf{Z})^*$.

(e) Si $G = H_8$. Il y a 5 classes de conjugaison. Le quotient de G par son centre $\{\pm 1\}$ est isomorphe à $(\mathbf{Z}/2\mathbf{Z})^2$ donc il y a quatre représentations de dimension 1 et une représentation de dimension 2 puisque $8 = 1^2 + 1^2 + 1^2 + 1^2 + 2^2$. La représentation de dimension 2 est la représentation "standard" donnée par :

$$\rho(\pm 1) = \pm Id, \quad \rho(i) = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad \rho(j) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \qquad \text{et} \qquad \rho(k) = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix},$$

le caractère s'écrivant $\chi(\pm 1) = \pm 2$ et $\chi(x) = 0$ si $x \neq \pm 1$.

(f) Si $G = \mathcal{A}_4$, il y a 4 classes de conjugaison et le quotient de G par le groupe de Klein est isomorphe à $\mathbf{Z}/3\mathbf{Z}$ et c'est $G/D(G)$, donc il y a exactement trois représentations de dimension 1 et donc une quatrième représentation irréductible qui doit être de dimension 3 puisque $12 = 1^2 + 1^2 + 1^2 + 3^2$. Considérons la représentation de permutation sur $\mathbf{C}^4$ et E le sous-espace $x_1 + x_2 + x_3 + x_4 = 0$, on vérifie facilement que c'est la représentation cherchée.

(g) Si $G = \mathcal{A}_5$, il y a 5 classes de conjugaison : la classe de l'élément neutre, la classe $\mathcal{C}_{2,2}$ des doubles transpositions (15 éléments), la classe $\mathcal{C}_3$ des 3-cycles (20 éléments) et deux classes $\mathcal{C}_5$ et $\mathcal{C}'_5$ décrivant les 5-cycles (12 éléments chacune). Comme G est simple, la seule représentation de dimension 1 est la représentation triviale. La représentation de permutation sur $\mathbf{C}^5$ contient une sous-représentation E : le sous-espace $x_1+x_2+x_3+x_4+x_5 = 0$, on vérifie facilement que c'est une représentation irréductible, on la note ρ_4. Le groupe $\mathcal{A}_5$ contient six sous-groupes de Sylow de cardinal 5 et on a donc une injection "exotique" $\mathcal{A}_5 \hookrightarrow \mathcal{S}_6$; on obtient une sous-représentation F de dimension 5 donnée par l'hyperplan somme des coordonnées égale zéro dans la représentation de permutation ; on vérifie également que c'est une représentation irréductible de $\mathcal{A}_5$, on la note ρ_5. Les deux autres représentations irréductibles sont de dimension 3 puisque $60 = 1^2+3^2+3^2+4^2+5^2$. On sait que $\mathcal{A}_5$ est le groupe d'isométrie de l'icosaèdre, d'où une première représentation ρ_2 de dimension 3, la deuxième s'obtient en modifiant celle-ci par l'automorphisme ϕ "conjugaison par une transposition" (qui n'est pas intérieur dans $\mathcal{A}_5$), c'est-à-dire $\rho_3 := \rho_2 \circ \phi$. Comme ϕ échange $\mathcal{C}_5$ et $\mathcal{C}'_5$ et comme $\chi_2(\mathcal{C}_5) \neq \chi_2(\mathcal{C}'_5)$ on voit que ρ_2 et ρ_3 ne sont pas isomorphes.

On peut vérifier que le tableau des valeurs des 5 caractères $\chi_i := \chi_{\rho_i}$ est le suivant :

	1	$\mathcal{C}_{2,2}$	$\mathcal{C}_3$	$\mathcal{C}_5$	$\mathcal{C}'_5$
χ_1	1	1	1	1	1
χ_2	3	-1	0	$\frac{1+\sqrt{5}}{2}$	$\frac{1-\sqrt{5}}{2}$
χ_3	3	-1	0	$\frac{1-\sqrt{5}}{2}$	$\frac{1+\sqrt{5}}{2}$
χ_4	4	0	1	-1	-1
χ_5	5	1	-1	0	0

Passons maintenant à des considérations d'*intégralité*. La première remarque est que les valeurs propres d'un $\rho(g)$ étant des racines de l'unité, les valeurs prises par un caractère sont des entiers algébriques. On peut être plus précis et montrer

Proposition. *Soit $f : G \to \mathbf{C}$ une fonction centrale dont les valeurs sont des entiers algébriques, et soit χ le caractère d'une représentation irréductible de dimension n alors $\frac{1}{n}\sum_{g\in G} f(g)\chi(g)$ est un entier algébrique.*

Preuve. Soit $C_1, \ldots, C_h$ les classes de conjugaison de G et $e_{C_i} = \sum_{g\in C_i} g \in \mathbf{Z}[G]$. Alors $B = \mathbf{Z}e_{C_1} + \ldots + \mathbf{Z}e_{C_h}$ est un sous-anneau commutatif de $\mathbf{Z}[G]$ (ou $\mathbf{C}[G]$) dont tous les éléments sont entiers sur $\mathbf{Z}$ (i.e. racine d'un polynôme *unitaire* à coefficients dans $\mathbf{Z}$). On en déduit que le polynôme caractéristique (ou minimal) de $\rho_f = \sum_{g\in G} f(g)\rho(g)$ est à coefficients entiers et donc que sa valeur propre $\frac{1}{n}\sum_{g\in G} f(g)\chi(g)$ est un entier algébrique. □

Corollaire. *La dimension n d'une représentation irréductible de G divise le cardinal de G.*

Preuve. On applique l'énoncé précédent à la fonction $f(g) = \chi(g^{-1})$ et on obtient que $\frac{1}{n}\sum_{g\in G}\chi(g^{-1})\chi(g) = \frac{|G|}{n}(\chi,\chi) = \frac{|G|}{n}$ est un entier algébrique donc un entier. □

Application. Montrons qu'un groupe G de cardinal $p^a q^b$ est résoluble (résultat dû à Burnside). Il suffit en fait de montrer qu'il n'est pas simple.

Lemme. *Soit ρ une représentation irréductible de G de dimension n, soit χ son caractère et soit C une classe de conjugaison de cardinal h avec $\mathrm{PGCD}(h,n) = 1$, alors ou bien $\chi(C) = 0$ ou bien $\chi(C) = n\zeta$ avec ζ racine de l'unité et $\rho(C)$ est dans le centre de la représentation.*

Preuve du lemme. Choisissons $x \in C$, ou bien toutes les valeurs propres de $\rho(x)$ sont égales et alors $\rho(x) = \zeta id$ et bien sûr $\chi(x) = n\zeta$, ou bien les valeurs propres $\zeta_1, \ldots, \zeta_n$ ne sont pas toutes égales. Dans le dernier cas on a donc $|\chi(x)| = |\zeta_1 + \ldots + \zeta_n| < n$. Par ailleurs nous avons vu que $h\chi(x)/n$ est un entier algébrique, or l'hypothèse entraîne par le théorème de Bézout l'existence de $u, v \in \mathbf{Z}$ tels que $uh + vn = 1$ donc

$$\frac{\chi(x)}{n} = u\left(\frac{h\chi(x)}{n}\right) + v\chi(x)$$

est encore un entier algébrique. Etant de module < 1 dans tout plongement, il est donc nul (car sa norme est un entier < 1). $\square$

Soit maintenant G de cardinal $p^a q^b$, choisissons $x \neq e$ dans le centre d'un q-sous-groupe de Sylow, alors ou bien x est dans le centre de G qui n'est donc pas simple, ou bien la classe de conjugaison de x a pour cardinal une puissance positive de p. Soit χ le caractère de ρ une des représentations irréductibles de G, d'après le lemme, ou bien p divise la dimension m_χ de la représentation, ou bien $\chi(x) = 0$, ou bien $\rho(x)$ est dans le centre de $\rho(G)$. Mais, en écrivant

$$0 = \chi_{\text{reg}}(x) = \sum_{\chi \in Irr(G)} m_\chi \chi(x) = 1 + \sum_{\chi \neq 1} m_\chi \chi(x)$$

on voit qu'on ne peut avoir $\chi(x) = 0$ pour tous les χ tels que p ne divise pas m_χ, sinon on obtiendrait une égalité du type $0 = 1 + p(\text{entier})$. Par conséquent pour une de ces représentations, $\rho(x)$ est dans le centre de $\rho(G)$ et donc G n'est pas simple.