MATHS AGREG

collection dirigée par Jacques MOISAN

COURS D'ALGÈBRE

Daniel PERRIN

ellipses

CAPES/AGREG

MATHÉMATIQUES

collection dirigée par Jacques MOISAN

COURS D'ALGÈBRE

Daniel PERRIN

Professeur à l'IUFM de Versailles
et à l'Université de Paris-Sud (Orsay)

ellipses

ISBN 2-7298-5552-1

INTRODUCTION

Ce livre est directement issu du Cours d'Algèbre paru voici maintenant plus de quatorze ans aux presses de feu l'École Normale Supérieure de Jeunes Filles.

La version initiale était la rédaction du cours oral de l'auteur donné dans le cadre de la préparation à l'agrégation de mathématiques à l'ENSJF, durant les années 1976-1981.

Par rapport au polycopié, ce livre ne diffère que sur des points mineurs : quelques corrections de détail, une rédaction un peu moins elliptique et surtout une typographie plus conforme aux normes actuelles.

Il bénéficiera désormais d'une diffusion moins clandestine qu'auparavant, ce qui était, je crois, le souhait de beaucoup et notamment des candidats à l'agrégation.

Les connaissances nécessaires pour aborder ce livre sont principalement celles des cours d'algèbre linéaire du premier cycle des universités. Il faut citer, en plus, quelques connaissances très élémentaires sur groupes, anneaux et corps, ainsi que le théorème de Zorn.

Le cours porte essentiellement sur les groupes "classiques" (groupes linéaires et orthogonaux), avec trois chapitres d'algèbre générale sur groupes, anneaux et corps. Voici le détail des différents chapitres :

Le chapitre I comprend les définitions générales sur les groupes, constamment utilisées dans la suite (sous-groupes distingués, centre, commutateurs, opérations, produits) assorties d'exemples et d'applications prises dans la théorie des groupes finis (Sylow, classification des groupes de petits ordres,...) Il se termine par l'étude plus approfondie des groupes symétriques et alternés (simplicité, automorphismes).

Le chapitre II porte sur les anneaux et particulièrement sur leurs propriétés arithmétiques, c'est-à-dire celles qui concernent la divisibilité (anneaux principaux, factoriels ...) avec comme application le théorème des deux carrés.

Le chapitre III contient la théorie élémentaire des corps (celle qui relève essentiellement de l'algèbre linéaire), à l'exclusion de la théorie de Galois. On y aborde les constructions à la règle et au compas, les corps finis et, en liaison avec ceux-ci, l'irréductibilité des polynômes à coefficients entiers ou rationnels, notamment les polynômes cyclotomiques.

Les chapitres suivants, consacrés aux groupes classiques sont construits sur un même plan : en vue d'obtenir un "dévissage" de ces groupes, on étudie leurs sous-groupes distingués (centre, commutateurs) à l'aide de générateurs appropriés. On termine en général par des théorèmes de simplicité.

Le chapitre IV concerne les groupes linéaires et comporte, outre ce qui précède, l'étude de ces groupes sur un corps de base fini.

Le chapitre V contient le vocabulaire des formes sesquilinéaires (quadratiques, hermitiennes et alternées), quelques résultats sur leur classification et des rudiments sur les groupes associés.

Dans le chapitre VI on étudie le groupe orthogonal euclidien réel en essayant de mettre en évidence le caractère particulier de ce groupe.

Le chapitre VII, voué à l'étude des quaternions, permet de terminer l'étude entreprise au chapitre précédent, en élucidant la structure du groupe orthogonal euclidien en dimension 4.

Enfin, au chapitre VIII, on étudie le groupe orthogonal général (pour un corps de caractéristique différente de 2 et une forme quelconque), avec entre autres l'étude des plans et des espaces hyperboliques, le théorème de Witt et celui de Cartan-Dieudonné. On termine par l'énoncé (sans démonstration) des résultats essentiels sur ces groupes, dans le cas où il y a des vecteurs isotropes.

De nombreux exercices sont rassemblés à la fin de chaque chapitre. Ils sont de difficultés variables et apportent souvent des compléments au cours (contre-exemples, applications, généralisations, démonstrations de résultats admis dans le cours). Le lecteur qui désire trouver une solution de certains de ces exercices pourra se reporter à l'excellent [FG].

Il y a presque dix ans maintenant que l'ENSJF a disparu, mais je voudrais profiter de l'occasion qui m'est offerte pour remercier les sévriennes à qui était destiné ce cours. J'ai conservé des relations amicales avec quelques unes, mais perdu de vue la plupart des autres, pourtant je voudrais qu'elles sachent, toutes, que les dix ans passés à l'ENSJF ont sans doute été les plus belles années de ma vie professionnelle.

Je voudrais remercier aussi les collègues avec qui je travaillais à l'époque et qui ont influencé ce texte et notamment Michel Broué, Pierre Mazet et Marie-France Vignéras.

Une mention spéciale va aux deux corédacteurs de la première version, Marc Cabanes et Martine Duchêne qui voient enfin leur travail publié au grand jour, ainsi qu'à tous ceux qui avaient permis, dans des conditions artisanales, la parution du premier Cours d'Algèbre et je pense notamment à Annie Iapteff.

Je remercie les éditions Ellipses et particulièrement Jacques Moisan pour avoir accepté ce texte sans réserves ni discussion. Je remercie enfin Jacques et Denise Moisan de l'aide qu'ils m'ont apporté pour la réalisation matérielle du manuscrit.

TABLE DES MATIÈRES

NOTATIONS

On désigne par $\mathbf{N}$ (resp. $\mathbf{Z}$, $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{C}$) l'ensemble des entiers ≥ 0 (resp. des entiers relatifs, des nombres rationnels, des nombres réels, des nombres complexes).

On note $|E|$ le cardinal d'un ensemble E. On note $[x]$ la partie entière du nombre réel x. On note C_n^p les coefficients binômiaux.

Si p et q sont des entiers (ou, plus généralement des éléments d'un anneau, cf. Chapitre II), la notation $p \mid q$ (resp. $p \nmid q$) signifie que p divise q (resp. p ne divise pas q).

Si p et q sont deux **entiers**, la notation $(p, q) = 1$ signifie que p et q sont premiers entre eux.

Si k est un corps, on note k^* l'ensemble des éléments non nuls de k et k^{*2} l'ensemble des carrés des éléments non nuls de k.

Nous utiliserons librement les notations usuelles de l'algèbre linéaire : $\dim E$ (dimension de l'espace vectoriel E), E^* (espace dual de E), $\operatorname{Ker} f$ (noyau de l'application linéaire f), $\operatorname{Im} f$ (image de l'application linéaire f), $\operatorname{rg} A$ ou $\operatorname{rg} u$ (rang d'une matrice ou d'une application linéaire), $\det A$ ou $\det u$ (déterminant d'une matrice ou d'un endomorphisme), $\operatorname{Tr} A$ ou $\operatorname{Tr} u$ (trace d'une matrice ou d'un endomorphisme).
Le sous-k-espace vectoriel engendré par un vecteur x sera noté $<x>$ ou kx.
Si (e_i), $(i \in I)$ est une base de E, (e_i^*), $(i \in I)$ désigne la base duale qui est une base de E^*.

I. GÉNÉRALITÉS SUR LES GROUPES,

groupes finis, groupe symétrique

0. Rappels.

Nous supposons connues les notions de groupe, groupe abélien (ou commutatif), de sous-groupe, d'homomorphisme de groupes.
Nous noterons généralement multiplicativement : $(g,h) \mapsto gh$ les lois de groupe, l'élément neutre sera alors noté 1, l'inverse de g sera noté g^{-1}. Cette règle aura dans ce livre une exception majeure, celle du groupe $\mathbf{Z}$, ses sous-groupes et ses quotients, ainsi que des groupes additifs des corps et des espaces vectoriels qui seront bien entendu notés additivement.

Le cardinal d'un groupe fini est aussi appelé son ***ordre***. Si p est un nombre premier, on appelle p-groupe un groupe dont le cardinal est une puissance de p. Si g est un élément de G, l'ordre de g est le plus petit entier $n > 0$ (s'il en existe) qui vérifie $g^n = 1$. C'est aussi l'ordre du sous-groupe engendré par g, cf. § 1 ci-dessous.

Le ***noyau*** d'un homomorphisme $f : G \to H$ est le sous-groupe de G défini par :

$$\operatorname{Ker} f = \{g \in G \mid f(g) = 1\}.$$

L'image de f est aussi un sous-groupe de H, noté $\operatorname{Im} f$. Un isomorphisme est un homomorphisme de groupes bijectif. Un ***automorphisme*** d'un groupe G est un isomorphisme de G sur G. Un exemple d'automorphisme est fourni par les ***automorphismes intérieurs***. Un tel automorphisme i_g est donné, pour $g \in G$, par la formule $i_g(x) = gxg^{-1}$.

Si H est un sous-groupe d'un groupe G on appelle ***classe à gauche*** de l'élément $a \in G$ relativement à H le sous-ensemble

$$aH = \{g \in G \mid g = ah, \quad h \in H\}$$

et on définit de même les classes à droite Ha. Les classes à gauche forment une partition de G. Leur ensemble est noté G/H. Ce n'est pas un groupe en général. Le cardinal de G/H est appelé l'***indice*** de H dans G et noté $(G : H)$.
Lorsque le groupe est fini, la considération des classes à gauche conduit au théorème suivant :

Théorème 0.1 (Lagrange).
Si H est un sous-groupe du groupe fini G, l'ordre de H et l'indice de H dans G divisent l'ordre de G. Précisément, on a

$$|G| = |H|\,|G/H| = |H|(G : H).$$

En particulier l'ordre d'un élément $g \in G$ divise l'ordre de G.

Le groupe des bijections (ou permutations) d'un ensemble E s'appelle le ***groupe symétrique*** de E et est noté $\mathfrak{S}(E)$. Si E et E' ont même cardinal les groupes symétriques associés sont isomorphes. Lorsque l'on a $E = \{1, 2, \ldots, n\}$ avec $n \in \mathbf{N}$ on pose $\mathfrak{S}(E) = \mathfrak{S}_n$ et on parle du groupe symétrique standard. Le cardinal de ce groupe est $n!$.

Le groupe symétrique contient des permutations remarquables : les ***cycles*** d'ordre k. Un tel cycle est noté $\sigma = (a_1, a_2, \ldots, a_k)$ avec les $a_i \in E$, distincts et la notation signifie que l'on a $\sigma(a) = a$ si a n'est pas l'un des a_i et $\sigma(a_i) = a_{i+1}$ (où l'indice est pris modulo k). Un tel cycle est un élément d'ordre k (i.e. vérifie $\sigma^k = 1$). Pour $k = 2$ on parle de ***transpositions***.

Le groupe $\mathfrak{S}_n$ est muni d'un homomorphisme surjectif, appelé signature, et noté $\varepsilon : \mathfrak{S}_n \to \{1, -1\}$, que l'on peut définir de multiples façons (cf. par exemple [L]) mais dont nous retiendrons les propriétés suivantes :
a) si τ est une transposition on a $\varepsilon(\tau) = -1$, plus généralement,
b) si σ est un cycle d'ordre k on a $\varepsilon(\sigma) = (-1)^{k+1}$.
Le noyau de ε est formé des permutations paires (i.e. celles qui vérifient $\varepsilon(s) = 1$). C'est un groupe de cardinal $n!/2$, appelé groupe ***alterné*** et noté $\mathfrak{A}_n$.

Enfin, le lecteur est supposé avoir une certaine familiarité avec quelques d'objets élémentaires comme les groupes additifs $\mathbf{Z}/n\mathbf{Z}$ des congruences modulo n ou le groupe de Klein $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}, \ldots$

1. Générateurs d'un groupe.

Proposition-définition 1.1.
*Soient G un groupe et $A \subset G$ une partie de G. Il existe un plus petit sous-groupe H de G contenant A. On dit que H est le sous-groupe engendré par A, ou que les éléments de A sont des **générateurs** de H. On note $H = <A>$.*

Démonstration. L'existence de H peut se voir de deux manières :
a) par « l'extérieur » : on considère tous les sous-groupes de G contenant A (il y a au moins G tout entier) et leur intersection convient ;
b) par « l'intérieur » : on suppose A non vide (sinon on a $H = \{1\}$), on pose $A^{-1} = \{x \in G \,|\, x^{-1} \in A\}$, puis $H = \{a_1 \ldots a_n \,|\, n \in \mathbf{N},\ a_i \in A \cup A^{-1}\}$. Alors H est un groupe, contient A et est évidemment le plus petit possible.

Exemples 1.2.

1) Groupes monogènes et cycliques.

Un groupe G engendré par un élément a, est dit ***monogène***. Il est isomorphe à $\mathbf{Z}$ ou $\mathbf{Z}/n\mathbf{Z}$, pour un $n \in \mathbf{N}$ (considérer l'homomorphisme surjectif $\varphi : n \mapsto a^n$, de $\mathbf{Z}$ dans G). Dans le second cas, G est dit ***cyclique***. En particulier, si $|G| = p$ est un nombre premier, G n'a pas de sous-groupe non trivial (en vertu du théorème de Lagrange), donc si $a \in G$ et $a \neq 1$, G est égal à $<a>$, donc cyclique et on a $G \simeq \mathbf{Z}/p\mathbf{Z}$.

2) Groupes symétrique $\mathfrak{S}_n$ et alterné $\mathfrak{A}_n$.

a) Les transpositions engendrent $\mathfrak{S}_n$, on peut même se limiter aux transpositions $(1,2), (1,3), \ldots (1,n)$ ou encore $(1,2), (2,3), \ldots (n-1,n)$ comme on le voit aisément par récurrence sur n.
Le lecteur montrera, à titre d'exercice, que la transposition (1,2) et le n-cycle (1,2,...,n) engendrent $\mathfrak{S}_n$ (cf. 4.10 ci-dessous).

b) Les cycles d'ordre 3 engendrent $\mathfrak{A}_n$, pour $n \geq 3$. En effet, $\mathfrak{A}_n$ est engendré par les produits pairs de transpositions et on a les formules :

$$(a,b)(b,c) = (a,b,c),$$

$$(a,b)(a,c) = (a,c,b)$$

(ce qui prouve au passage que tous les cycles d'ordre 3 sont dans $\mathfrak{A}_n$),

$$(a,b)(c,d) = (a,b)(a,c)\ (a,c)(c,d) = (a,c,b)(a,c,d).$$

Nous verrons, à propos des groupes classiques, de nombreux autres exemples de générateurs.

2. Sous-groupes distingués.

Définition 2.1.
Soit G un groupe et H un sous-groupe de G. On dit que H est ***distingué*** *dans G s'il est invariant par automorphisme intérieur i.e. si on a :*

$$\forall a \in G,\ \forall h \in H,\ aha^{-1} \in H.$$

On note alors : $H \lhd\!\lhd G$.

Remarques 2.2.
0) La condition ci-dessus équivaut à dire que pour tout $a \in G$ on a $aH = Ha$, i.e. l'égalité des classes à droite et à gauche modulo H.
1) Si $f : G \longrightarrow G'$ est un homomorphisme, son noyau $\operatorname{Ker} f$ est un sous-groupe distingué de G.
2) Réciproquement, si on a $H \lhd\!\lhd G$, le quotient G/H, ensemble des classes à gauche (ou à droite) est muni d'une structure de groupe et on a un homomorphisme surjectif $p : G \longrightarrow G/H$, de noyau H.
Dans la situation de 1) on a, de plus, un isomorphisme :

$$\operatorname{Im} f \simeq G/\operatorname{Ker} f.$$

3) Enfin on définit une ***suite exacte*** :

$$1 \longrightarrow N \xrightarrow{i} G \xrightarrow{p} H \longrightarrow 1.$$

Dans cette écriture, N, G, H sont des groupes, i, p des homomorphismes et la suite est dite exacte si : 1) i est injectif, 2) p est surjectif, 3) on a $\operatorname{Im} i = \operatorname{Ker} p$.
Lorsque les groupes sont abéliens et notés additivement on écrit les suites exactes avec des 0 :

$$0 \to N \to G \to H \to 0.$$

Exemples 2.3.

1) $\{1\}$ et G sont toujours des sous-groupes distingués (dits triviaux).

2) Si G est abélien, tout sous-groupe de G est distingué. Pour la réciproque, cf. Exercice B.3.

3) Étudions le groupe $\mathfrak{S}_3$ qui a 6 éléments :
$1 = \mathrm{Id}$, $\tau_c = (a,b)$, $\tau_b = (a,c)$, $\tau_a = (b,c)$, $\sigma = (a,b,c)$, $\sigma^2 = \sigma^{-1} = (a,c,b)$.
Le groupe $\mathfrak{S}_3$ contient un sous-groupe distingué d'ordre 3, $<\sigma>= \{1, \sigma, \sigma^2\} = \mathfrak{A}_3$, isomorphe à $\mathbf{Z}/3\mathbf{Z}$ et on a une suite exacte :

$$1 \longrightarrow \mathbf{Z}/3\mathbf{Z} \longrightarrow \mathfrak{S}_3 \longrightarrow \mathbf{Z}/2\mathbf{Z} \longrightarrow 1.$$

En revanche les sous-groupes $<\tau_a>= \{1, \tau_a\}$ ... ne sont pas distingués, on a :

$$\sigma\tau_a\sigma^{-1} = \tau_{\sigma(a)} = \tau_b.$$

Définition 2.4.
Un groupe $G \neq \{1\}$ est dit ***simple*** *si ses seuls sous-groupes distingués sont $\{1\}$ et G.*

Exemples 2.5.

1) $\mathbf{Z}/p\mathbf{Z}$ est simple si et seulement si p est premier.

2) $\mathfrak{A}_n$ est simple pour $n \geq 5$ (cf. §8).

2.6 Commentaire.

L'intérêt des sous-groupes distingués est de permettre le « dévissage » des groupes : si G est un groupe et si on a un sous-groupe distingué $H \lhd G$, on peut essayer de ramener l'étude de G à celle de H et de G/H (si G est fini, ces groupes sont de cardinal plus petit). Nous verrons des exemples de dévissages aux §6 et §7. Les groupes simples, eux, sont indévissables d'où l'intérêt particulier qu'on leur porte. La classification des groupes simples finis a été achevée en 1981, cf. [Pu].
Là encore, les groupes classiques nous fourniront beaucoup d'exemples de groupes simples.

3. Centre et commutateurs.

Nous exhibons maintenant deux sous-groupes distingués d'un groupe G, qui existent toujours, mais peuvent être triviaux : le centre et le groupe dérivé.

a) Le centre.

Définition 3.1.
Le ***centre*** *du groupe G est le sous-groupe de G formé des éléments qui commutent avec tous les autres :*

$$Z(G) = \{a \in G \mid \forall g \in G,\ ag = ga\}.$$

On a $Z(G) \lhd G$, ($Z(G)$ est même un sous-groupe ***caractéristique*** de G, i.e. invariant par tout automorphisme).

Exemples 3.2.

1) Si G est commutatif, on a $Z(G) = G$.

2) Si $G = \mathfrak{S}_n$, avec $n \geq 3$, on a $Z(G) = \{1\}$. En effet, soit $\sigma \in G$, $\sigma \neq 1$. Pour un certain i, on a $\sigma(i) = j \neq i$. Soit $k \neq i, j$ et $\tau = (j, k)$. Alors on a $\sigma\tau(i) = \sigma(i) = j$, $\tau\sigma(i) = \tau(j) = k$, donc $\sigma\tau \neq \tau\sigma$ et $\sigma \notin Z(G)$.

3) Soit $\mathbf{H_8} = \{\mp 1, \mp i, \mp j, \mp k\}$ le groupe des quaternions (cf. Chapitre VII). (La multiplication est définie par la règle des signes et les formules

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

Alors $Z(\mathbf{H_8}) = \{1, -1\}$ est non trivial, (cf. 4.15)

b) Les commutateurs.

Définition 3.3.

Le **groupe dérivé** $D(G)$ *est le sous-groupe* **engendré** *par les commutateurs de* G, *i.e. les éléments de la forme* $xyx^{-1}y^{-1}$ *avec* $x, y \in G$.(1)

Le groupe $D(G)$ est parfois appelé improprement groupe des commutateurs.
On a $D(G) \lhd\!\lhd G$, et même $D(G)$ caractéristique. En effet, si $\varphi \in \mathrm{Aut}\, G$, on a $\varphi(xyx^{-1}y^{-1}) = \varphi(x)\varphi(y)\varphi(x)^{-1}\varphi(y)^{-1}$, donc les commutateurs sont conservés. Notons que $G/D(G)$ est abélien. C'est même le plus grand quotient abélien de G, et ceci caractérise $D(G)$.

Exemples 3.4.

1) Si G est commutatif on a $D(G) = \{1\}$.
2) Si $G = \mathfrak{S}_3$ on a $D(G) = \{1, \sigma, \sigma^2\}$.
3) Si $G = \mathbf{H_8}$ on a $D(G) = \{1, -1\}$.
4) Si $G = \mathfrak{A}_5$ on a $D(G) = \mathfrak{A}_5$ (cf. §8).

4. Opération d'un groupe sur un ensemble.

Définition 4.1.

Soit G *un groupe,* X *un ensemble, on dit que* G **opère** *sur* X *si on s'est donné une application :*

$$\begin{aligned} G \times X &\longrightarrow X \\ (g, x) &\longmapsto g.x \end{aligned}$$

vérifiant les axiomes suivants :

1) $\forall g, g' \in G,\ \forall x \in X,\ g.(g'.x) = (gg').x$

2) $\forall x \in X,\ 1.x = x.$

Il revient au même de se donner un homomorphisme $\varphi : G \longrightarrow \mathfrak{S}(X)$, où $\mathfrak{S}(X)$ désigne le groupe des bijections de X (on pose alors : $g.x = \varphi(g)(x)$).

4.2 Commentaire.

C'est une notion essentielle ! D'abord, c'est la situation que l'on rencontre dans toute géométrie (affine, avec le groupe affine, projective, avec $PGL(n, k)$ cf. Chapitre IV, euclidienne avec le groupe des isométries, hyperbolique avec le groupe de Lorentz etc), ensuite, parce qu'au delà de l'intérêt de l'opération pour l'étude

(1) *Le commutateur de* x *et* y, $xyx^{-1}y^{-1}$, *est appelé ainsi car il vaut* 1 *si et seulement si* x *et* y *commutent. On le note parfois* $[x, y]$.

de l'ensemble X, elle permet souvent en retour d'obtenir des renseignements sur le groupe G comme nous le verrons au paragraphe suivant.
Nous appellerons « géométriques » les propriétés d'un élément de G relatives à une opération (points fixes ...) par opposition aux propriétés « algébriques » (ordre d'un élément, commutation ...).

Définition 4.3.
a) On dit que G opère ***transitivement*** *sur X si on a :*

$$\forall x \in X,\ \forall y \in X,\quad \exists g \in G,\quad g.x = y.$$

b) On dit que G opère ***fidèlement*** *si $\varphi : G \longrightarrow \mathfrak{S}(X)$ est injectif i.e. si $g.x = x$ pour tout $x \in X$ implique $g = 1$.*

Notons que $G/\mathrm{Ker}\,\varphi$ opère fidèlement sur X. Ainsi, si E est un espace vectoriel, le groupe $GL(E)$ opère non fidèlement sur l'ensemble $\mathbf{P}(E)$ des droites vectorielles de E mais son quotient $PGL(E)$ opère fidèlement (cf. IV, 2.8).

Si G n'opère pas transitivement, on introduit la relation d'équivalence suivante :

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

qui mesure le défaut de transitivité. Les classes pour cette relation sont les ***orbites*** de X sous G. L'orbite de $x \in X$ est notée $\omega(x)$. On notera que G opère transitivement sur $\omega(x)$.
Par exemple, les orbites du groupe orthogonal $O(n, \mathbf{R})$ dans son opération naturelle sur $\mathbf{R}^n$ sont les sphères de centre l'origine.

Exemple 4.4. Décomposition d'une permutation en produit de cycles disjoints.
Le groupe $\mathfrak{S}_n$ opère sur $X = \{1, 2, ..., n\}$. Soit $\sigma \in \mathfrak{S}_n$ et $<\sigma>$ le groupe cyclique engendré par σ qui opère aussi sur X. Soient $F_1, F_2, \ldots, F_r$ les orbites de X sous $<\sigma>$. Alors, les permutations σ_i définies par :

$$\sigma_i(x) = \begin{cases} x & \text{si } x \notin F_i \\ \sigma(x) & \text{si } x \in F_i \end{cases}$$

sont des cycles, d'ordre $|F_i|$, deux à deux permutables, et on a $\sigma = \sigma_1 \ldots \sigma_r$.
Par exemple si $X = \{1, \ldots, 8\}$ et

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 6 & 4 & 5 & 1 & 8 & 7 & 2 \end{pmatrix},$$

on a $\sigma = (1\,3\,4\,5)\,(2\,6\,8)\,(7) = (1\,3\,4\,5)\,(2\,6\,8)$, (en général, les cycles d'ordre 1 sont omis dans l'écriture de σ).

Définition 4.5.
Si G opère sur X et si $x \in X$, on définit $H_x = \{g \in G \mid g.x = x\}$.
C'est un sous-groupe de G (non distingué en général) appelé le ***stabilisateur*** *(ou encore fixateur) de x.*

Exemple 4.6. Dans l'opération de $\mathfrak{S}_n$ sur $X = \{1, ..., n\}$, le stabilisateur d'un point est isomorphe à $\mathfrak{S}_{n-1}$.

Orbites et stabilisateurs sont liés par la remarque évidente suivante (cf. aussi ci-dessous Exemple C) :

Proposition 4.7.
L'application $\overline{g} \mapsto g.x$ de G/H_x (ensemble des classes à gauche) dans $\omega(x)$ est bien définie et est une bijection.

Lorsque G est fini, on a donc $|\omega(x)| = |G|/|H_x|$, en particulier $|\omega(x)|$ divise $|G|$. Lorsque G et X sont munis de structures supplémentaires (par exemple topologiques) on peut souvent préciser la proposition. Ainsi, par exemple, on peut montrer que l'espace quotient $O(n, \mathbf{R})/O(n-1, \mathbf{R})$ est homéomorphe à la sphère unité $\mathbf{S}^{n-1}$ de $\mathbf{R}^n$.

4.8 Quelques exemples d'opérations.

Nous donnons ici seulement des exemples en théorie des groupes, les situations géométriques seront étudiées dans les chapitres suivants.

Exemple A.

On peut faire opérer G sur G par ***translation à gauche.***
On pose pour $g, a \in G$, $g.a = ga$. (Attention l'application $a \mapsto ga$ n'est pas un automorphisme de groupes).
On remarque que G opère alors simplement transitivement i.e.

$$\forall a, b \in G,\ \exists ! g \in G,\ g.a = b \qquad (\text{on prend } g = ba^{-1}).$$

A fortiori, G opère donc fidèlement et on a un homomorphisme injectif

$$\varphi :\ G \longrightarrow \mathfrak{S}(G).$$

En particulier on en déduit le théorème de Cayley :

Théorème 4.9.
Si G est fini de cardinal n, G est isomophe à un sous-groupe de $\mathfrak{S}_n$.

Exemple B.

On peut aussi faire opérer G sur lui-même par ***automorphisme intérieur*** en posant $g.a = gag^{-1}$. Les orbites s'appellent alors ***classes de conjugaison*** et si $a' = gag^{-1}$, a' est un conjugué de a.
Le stabilisateur de a s'appelle ***centralisateur*** :

$$H_a = \{g \in G \,|\, gag^{-1} = a\} = \{g \in G \,|\, ga = ag\}.$$

On définit de même le centralisateur d'une partie de A de G :

$$C_G(A) = \{g \in G \,|\, \forall a \in A,\ ga = ag\}.$$

En particulier $C_G(G)$ est le centre de G. Dans le cas du groupe symétrique, on a la proposition suivante :

Proposition 4.10.
1) Si $\sigma \in \mathfrak{S}_n$ est un cycle d'ordre p, $\sigma = (a_1, \ldots, a_p)$ et si $\tau \in \mathfrak{S}_n$, on a

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

2) Dans $\mathfrak{S}_n$ tous les cycles d'ordre p sont conjugués.
3) Si $n \geq 5$ les cycles d'ordres 3 sont conjugués dans $\mathfrak{A}_n$.

Démonstration.
1) Si $x \notin \{\tau(a_1), \ldots, \tau(a_p)\}$, $\tau^{-1}(x) \notin \{a_1, \ldots, a_p\}$ et donc :

$$\tau\sigma\tau^{-1}(x) = \tau\tau^{-1}(x) = x.$$

Si $x = \tau(a_i)$, $\tau\sigma\tau^{-1}(x) = \tau\sigma(a_i) = \tau(a_{i+1})$ (les indices sont pris modulo p).

Il faut retenir de cet exemple ce qu'on peut appeler le **principe de conjugaison**. Il se résume en deux idées essentielles :

a) Si $g \in G$ est un élément « d'un certain type » $g' = \tau g \tau^{-1}$ est un élément « de même type ». Par exemple si g est d'ordre k, g' aussi ; si g a k points fixes, g' aussi ; en géométrie si g est une symétrie hyperplane, g' aussi ...

b) Si g est caractérisé « géométriquement » par un certain ensemble Y (ici, c'est l'ensemble des points fixes de g), l'ensemble Y' correspondant pour $g' = \tau g \tau^{-1}$ s'obtient en transportant Y par τ : $Y' = \tau(Y)$.
Nous retrouverons très souvent ce principe dans la suite.

Nous reprenons la démonstration de 4.10 :

2) Soient $\sigma = (a_1, \ldots, a_p)$, $\tau = (b_1, \ldots, b_p)$ et soit $g \in \mathfrak{S}_n$ tel que $g(a_i) = b_i$ pour tout i (il existe évidemment un tel g, même si $p = n$, on dit que $\mathfrak{S}_n$ est n-fois transitif). Alors on a $\tau = g\sigma g^{-1}$ en vertu du principe énoncé ci-dessus.

3) On a besoin du

Lemme 4.11.
Le groupe $\mathfrak{A}_n$ est $n-2$ fois transitif sur $\{1, \ldots, n\}$ i.e., si on a $a_1, \ldots a_{n-2}$ distincts et $b_1, \ldots b_{n-2}$ distincts, il existe $\sigma \in \mathfrak{A}_n$ tel que $\sigma(a_i) = b_i$.

En effet, on écrit :

$$\{1, \ldots, n\} = \{a_1, \ldots, a_{n-2}, a_{n-1}, a_n\} = \{b_1, \ldots, b_{n-2}, b_{n-1}, b_n\}$$

et on considère $\sigma \in \mathfrak{S}_n$ telle que $\sigma(a_i) = b_i$ pour tout $i = 1, \ldots, n$. Si σ est paire c'est terminé, sinon on compose σ avec la transposition (a_{n-1}, a_n).

Le même raisonnement qu'en 2) montre alors que pour $n \geq 5$, les 3-cycles sont conjugués dans $\mathfrak{A}_n$.

Remarque 4.12. Si $n = 3$ ou 4 l'assertion précédente est fausse. En effet, pour $n = 3$ le groupe $\mathfrak{A}_3$ est abélien donc la conjugaison y est triviale. Pour $n = 4$ il y a 8 cycles d'ordre 3 et ils ne peuvent être tous conjugués dans $\mathfrak{A}_4$ sinon ils formeraient une orbite dont le cardinal devrait diviser 12 en vertu de 4.7.

Exemple 4.13. Soit k un corps commutatif, $GL(n, k)$ le groupe des matrices carrées d'ordre n inversibles à coefficients dans k. Alors, pour $A, B \in GL(n, k)$:
A, B conjuguées $\iff \exists P \in GL(n, k)$, $B = P^{-1}AP \iff A, B$ semblables.

4.14 Une application aux p-groupes.

Rappelons que si p est un nombre premier, on appelle p-groupe un groupe dont le cardinal est une puissance de p. On a alors la proposition suivante :

Proposition 4.15.
Le centre d'un p-groupe distinct de $\{1\}$ n'est pas réduit à $\{1\}$.

Démonstration. On prouve d'abord le lemme suivant :

Lemme 4.16.
Soit G un p-groupe opérant sur un ensemble X et soit X^G l'ensemble des points fixes de X sous G, i.e.

$$X^G = \{x \in X \mid \forall g \in G \;\; g.x = x\}$$

alors on a $|X| \equiv |X^G| \pmod{p}$

Démonstration. On écrit X comme réunion disjointe de ses orbites sous G en remarquant que l'on a :

$$x \in X^G \Longleftrightarrow \omega(x) = \{x\}.$$

Si $x \notin X^G$, on a donc $|\omega(x)| > 1$ et comme $|\omega(x)|$ divise $|G| = p^n$, p divise $|\omega(x)|$. Le résultat provient alors de l'égalité :

$$|X| = |X^G| + \sum_{x \notin X^G} |\omega(x)|.$$

La proposition 2 résulte du lemme en remarquant que si G opère sur G par automorphisme intérieur, les points fixes sont les éléments du centre de G. On a ainsi : $|G| \equiv |Z(G)| \pmod{p}$ et comme 1 est dans $Z(G)$, on a $|Z(G)| \geq p$ d'où le résultat.

Exemple C.

Soit H un sous-groupe de G, pas nécessairement distingué, G/H l'ensemble des classes à gauches modulo H (i.e. l'ensemble des parties aH pour $a \in G$) alors, G ***opère sur*** G/H ***par translation*** en posant :

$$g.(aH) = (ga)H.$$

Cette opération est transitive : on a, pour $a, b \in G$, $(ba^{-1})aH = bH$ et le stabilisateur de aH est le sous-groupe aHa^{-1} conjugué de H (cf. Exemple D).
Elle n'est pas fidèle en général, si $\varphi : G \longrightarrow \mathfrak{S}(G/H)$ est le morphisme associé, on a en effet :

$$\operatorname{Ker}\varphi = \bigcap_{a \in G} aHa^{-1}.$$

Cet exemple est, en fait, le cas « générique » d'une opération transitive.
En particulier modulo la bijection $G/H_x \xrightarrow{\sim} \omega(x)$ vue en 4.7, l'opération de G sur $\omega(x)$ n'est autre que son opération naturelle sur G/H_x. Voici un exemple d'utilisation de cette opération :

Proposition 4.17.
Soit G un groupe infini et H un sous-groupe de G, distinct de G et d'indice fini. Alors G n'est pas simple.

Démonstration.
En effet, G opère sur G/H, d'où un homomorphisme $\varphi : G \longrightarrow \mathfrak{S}(G/H) \simeq \mathfrak{S}_n$ dont le noyau $\operatorname{Ker}\varphi$ est un sous-groupe distingué, distinct de $\{1\}$ car G est infini et de G car $H \neq G$. Voir aussi, dans le même style, l'exercice B.4.

Exemple D.

Cette fois, X est l'ensemble des sous-groupes de G et G opère sur X par automorphisme intérieur : $g.H = gHg^{-1}$.
On dit encore que H et gHg^{-1} sont conjugués. Le stabilisateur de H est

$$N_G(H) = \{g \in G \,|\, gHg^{-1} = H\}.$$

On l'appelle le ***normalisateur*** de H dans G, on a $H \lhd N_G(H)$ et $N_G(H)$ est le plus grand sous-groupe de G qui ait cette propriété.
Par exemple, dans $\mathfrak{S}_3$, les trois sous-groupes à 2 éléments $\{1, \tau_a\}, \ldots$ sont conjugués et sont leurs propres normalisateurs.
Nous allons utiliser tous ces exemples dans le paragraphe suivant.

5. Les théorèmes de Sylow.

Le théorème de Lagrange affirme que si H est un sous-groupe du goupe fini G, son cardinal divise le cardinal de G. On peut se demander, à l'inverse, si dans un groupe de cardinal n il existe, pour tout diviseur d de n, un (ou plusieurs) sous-groupe d'ordre d.
Il n'en est rien en général comme le montre l'exemple de $\mathfrak{A}_4$: on a $|\mathfrak{A}_4| = 12$ et on voit aisément que $\mathfrak{A}_4$ n'a pas de sous-groupe d'ordre 6. Il est cependant un cas très important où la propriété est vraie, celui des sous-groupes de Sylow. Dans tout ce paragraphe, la lettre p désigne un nombre premier.

Définition 5.1.
Soit G un groupe fini de cardinal n et p un diviseur premier de n. Si $n = p^\alpha m$ avec $p \nmid m$, on appelle ***p-sous-groupe de Sylow*** *de G un sous-groupe de cardinal p^α.*

Remarque 5.2. Dire que P est un p-sous-groupe de Sylow de G signifie :
1) que P est un p-groupe,
2) que l'indice $(G : P)$ est premier à p.

Exemple 5.3. Soit $\mathbf{F}_p = \mathbf{Z}/p\mathbf{Z}$ le corps fini à p éléments (p premier) et soit $G = GL(n, \mathbf{F}_p)$, $n \in \mathbf{N}^*$. Alors G est un groupe fini de cardinal

$$|G| = (p^n - 1)(p^n - p) \ldots (p^n - p^{n-1})$$

(cf. Chapitre IV, il suffit de compter les bases de $\mathbf{F}_p^n$) d'où $|G| = m\, p^{n(n-1)/2}$ avec $p \nmid m$.
On exhibe alors aisément un p-sous-groupe de Sylow P de G, c'est l'ensemble des matrices triangulaires supérieures strictes :

$$P = \{A = (a_{ij}) \,|\, a_{ij} = 0 \text{ si } i > j \text{ et } a_{ii} = 1\}.$$

En effet, comme les a_{ij}, pour $i < j$, sont quelconques, on a bien

$$|P| = p \times p^2 \times \cdots \times p^{n-1} = p^{n(n-1)/2}.$$

Le théorème suivant atteste l'existence des sous-groupes de Sylow :

Théorème 5.4 (Sylow).
Soit G un groupe fini et p un diviseur (premier) de $|G|$, alors G contient au moins un p-sous-groupe de Sylow.

Démonstration. La démonstration qui suit est tirée de l'excellent (mais introuvable) [S3]. Elle repose sur un lemme qui permet, connaissant un Sylow d'un groupe G d'en trouver un pour un sous-groupe H :

Lemme 5.5.
Soit G un groupe avec $|G| = n = p^\alpha m$, avec $p \nmid m$ et soit H un sous-groupe de G. Soit S un p-Sylow de G. Alors il existe $a \in G$ tel que $aSa^{-1} \cap H$ soit un p-Sylow de H.

Démonstration. Le groupe G opère sur G/S par translation à gauche (cf. §4 Exemple C) et le stabilisateur de aS est aSa^{-1}. Mais H opère lui aussi sur G/S par restriction, avec comme stabilisateur de aS, $aSa^{-1} \cap H$.
Il reste à voir que l'un de ces groupes est un Sylow de H. Ce sont déjà des p-groupes et il suffit donc que, pour un $a \in G$, $|H/(aSa^{-1} \cap H)|$ soit premier à p.
Mais on a, d'après 4.7, $|H/aSa^{-1} \cap H| = |\omega(aS)|$, cardinal de l'orbite de aS dans G/S sous l'action de H. Si tous ces nombres étaient divisibles par p, il en serait de même de $|G/S|$ car G/S est réunion des orbites $\omega(aS)$. Mais ceci contredit le fait que S est un p-Sylow de G.

Le théorème 1 résulte du lemme via l'exemple 5.3.

Soit en effet G un groupe et p un diviseur de $|G| = n$. On plonge d'abord G dans $\mathfrak{S}_n$ (par Cayley, cf. 4.9), puis on plonge $\mathfrak{S}_n$ dans $GL(n, \mathbf{F}_p)$ de la manière classique à savoir que $\sigma \in \mathfrak{S}_n$ s'envoie sur l'endomorphisme u_σ défini dans la base canonique par $u_\sigma(e_i) = e_{\sigma(i)}$.
Finalement on a donc réalisé G comme un sous-groupe de $GL(n, \mathbf{F}_p)$ qui possède un p-Sylow (cf. 5.3 ci-dessus), donc G aussi par le lemme 5.5.

Il y a beaucoup d'autres démonstrations de ce théorème, nous en signalons quelques-unes en exercices.

Corollaire 5.6.
Si $|G| = p^\alpha m$, $p \nmid m$, G contient des sous-groupes d'ordre p^i pour tout $i \leq \alpha$.

Démonstration. On est ramené au cas des p-groupes qui se traite par récurrence, vu l'existence d'un centre non trivial.

Le deuxième théorème de Sylow étudie la conjugaison des p-sous-groupes de Sylow.

Théorème 5.7 (Sylow).
Soit G un groupe, de cardinal $|G| = p^\alpha m$, avec $p \nmid m$.
1) Si H est un sous-groupe de G qui est un p-groupe, il existe un p-Sylow S, avec $H \subset S$.
2) Les p-Sylow sont tous conjugués (et donc leur nombre k divise n).
3) On a $k \equiv 1 \pmod{p}$ (donc k divise m).

On notera que l'assertion $k \mid n$ résulte du fait que les p-Sylow forment une orbite sous G (cf. 4.7).

Corollaire 5.8.
Si S est un p-Sylow de G,on a :

$$S \lhd G \iff S \text{ est l'unique } p\text{-Sylow de } G \iff k = 1.$$

Démonstration (de 5.7) On prouve 1) et 2) ensemble. Si H est un p-sous-groupe et S un p-Sylow de G, il existe, en vertu du lemme 3, $a \in G$ tel que $aSa^{-1} \cap H$ soit un p-Sylow de H. Mais comme H est un p-groupe, on a $aSa^{-1} \cap H = H$,

donc H est inclus dans aSa^{-1} qui est un Sylow. Si de plus H est un Sylow, on a exactement $H = aSa^{-1}$.
Pour le point 3), on fait, bien entendu, opérer G par conjugaison sur l'ensemble X de ses p-Sylow (cf. §4 Exemple D). Soit S un p-Sylow, S opère lui aussi sur X et on a encore, en vertu de 4.16 la congruence :

$$|X| \equiv |X^S| \pmod{p}.$$

Il reste à voir que l'on a $|X^S| = 1$. Bien sûr si $s \in S$, on a $sSs^{-1} = S$, autrement dit $S \in X^S$, on doit donc montrer que c'est le seul.
Pour cela soit T un p-Sylow, et supposons que T soit normalisé par S :

$$\forall s \in S,\ sTs^{-1} = T.$$

On utilise alors un argument dû à Frattini : on considère le sous-groupe N de G engendré par S et T. On a $S \subset N$, $T \subset N$ et ce sont, *a fortiori*, des p-Sylow de N. Mais, comme S normalise T, on a $T \lhd N$ et donc (corollaire 5.8 !) T est l'unique Sylow de N, donc $S = T$, cqfd.

Voici une application du corollaire 5.8. On en trouvera de nombreuses autres au paragraphe 7 et dans les exercices.

Proposition 5.9.
Un groupe d'ordre 63 n'est pas simple.

Démonstration. On regarde les sous-groupes de Sylow d'ordre 7 de G, (car $63 = 3^2 \times 7$), leur nombre k est congru à 1 modulo 7 et divise 9, donc $k = 1$, il y a un seul 7-Sylow qui est donc distingué.
Le même type d'argument vaut aussi pour $255 = 3 \times 5 \times 17$ et pour bien d'autres.

6. Produits directs et semi-directs.

Soit G un groupe, $N \lhd G$ un sous-groupe distingué, G/N le groupe quotient. On cherche, connaissant N et G/N, à reconstituer G. Plus généralement, étant donnés deux groupes N et H, on cherche tous les groupes G tels qu'on ait une suite exacte :

$$1 \longrightarrow N \longrightarrow G \longrightarrow H \longrightarrow 1.$$

Un tel groupe G s'appelle une ***extension*** de N par H (certains auteurs disent de H par N). Le problème général est délicat (cf. [H] Ch. 15) et nous n'en étudierons que deux cas particuliers très élémentaires : les produits directs et semi-directs.

a) Produits directs.

Soient N et H deux groupes. Le produit direct $G = N \times H$ est le produit cartésien de N et H, muni de la loi produit :

$$(n, h)(n', h') = (nn', hh').$$

On a alors une projection $p : G \to H$ définie par $p(n, h) = h$. C'est un homomorphisme de groupes, surjectif, de noyau le sous-groupe (distingué) $\overline{N}$ avec

$$\overline{N} = \{(n, 1) \mid n \in N\}$$

et on a donc une suite exacte :

$$1 \to N \xrightarrow{i} N \times H \xrightarrow{p} H \to 1$$

avec $i(n) = (n, 1)$, de sorte que $N \times H$ est bien une extension de N par H.
Bien entendu, ici, N et H jouent des rôles symétriques et on a aussi un sous-groupe

$$\overline{H} = \{(1, h) \mid h \in H\},$$

noyau de la projection sur N. On notera que le sous-groupe $\overline{H}$ est tel que :

1) la restriction de la projection $p|_{\overline{H}} : \overline{H} \to H$ est un isomorphisme,

2) $\overline{H}$ est un sous-groupe distingué de $N \times H$.

Voici un exemple bien classique de produit direct :

Proposition 6.1 (lemme chinois).
Si p et q sont des entiers premiers entre eux on a un isomorphisme

$$\mathbf{Z}/pq\mathbf{Z} \simeq \mathbf{Z}/p\mathbf{Z} \times \mathbf{Z}/q\mathbf{Z}.$$

Démonstration. Si on désigne par $\overline{n}$ (resp. $\widehat{n}$, resp. $\dot{n}$) les classes de n modulo pq (resp. p, resp. q) on a un homomorphisme $\overline{n} \mapsto (\widehat{n}, \dot{n})$, injectif car $(p, q) = 1$ et on conclut grâce à l'égalité des cardinaux.
Bien sûr l'assertion est fausse si p et q ne sont pas premiers entre eux (par exemple $\mathbf{Z}/4\mathbf{Z}$ et $V_4 = \mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$ ne sont pas isomorphes).

b) Produits semi-directs.

Le produit semi-direct est une variante affaiblie du produit direct. Soit G un groupe et N un sous-groupe distingué de G. On a donc une suite exacte

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

(où i est l'inclusion) et supposons que, comme dans le cas du produit direct, il existe un sous-groupe H de G tel que $p|_H$ induise un isomorphisme de H sur G/N (on dit qu'on a un relèvement de G/N). Attention, contrairement au cas du produit direct, H n'est pas distingué, a priori. Les conséquences de cette hypothèse sont les suivantes :
1) on a $N \cap H = \{1\}$,
2) on a $G = NH = \{nh \mid n \in N \text{ et } h \in H\}$.
En effet, si $g \in G$ et si $\overline{g} = p(g)$, il existe $h \in H$ tel que $p(h) = \overline{g}$, donc $gh^{-1} \in N$. L'écriture de g sous la forme nh est unique (si on a $nh = n'h'$, on a $n'^{-1}n = h'h^{-1} = 1$ car $N \cap H = \{1\}$), de sorte que G est en bijection avec $N \times H$.
Attention, cependant, si on calcule le produit de deux éléments de G on a la formule :

$$(nh)(n'h') = nhn'h' = nhn'h^{-1}hh'$$

avec $hn'h^{-1} \in N$ car N est distingué dans G.
On a donc montré, sous l'hypothèse de l'existence d'un relèvement, les deux faits suivants :

1) comme dans le cas du produit direct, G est encore en bijection avec le produit ensembliste $N \times H$,

2) la multiplication n'est pas celle du produit direct, elle est "tordue" au moyen de l'opération de H sur N par conjugaison : $h.n = hnh^{-1}$. On a, en effet :

$$(n, h)(n', h') = (n(h.n'), hh').$$

On notera que cette opération de H sur N n'est pas seulement ensembliste, H opère sur N par automorphismes de groupes.

L'analyse précédente nous conduit alors à la définition d'un produit semi-direct :

Proposition-définition 6.2.
1) Soient N et H deux groupes et Aut N *le groupe des automorphismes* **de groupe** *de N. Soit $\varphi : H \longrightarrow \mathrm{Aut}\, N$ un homomorphisme qui définit une opération de H sur N par la formule : $h.n = \varphi(h)(n)$.*
On définit sur l'ensemble produit $N \times H$ une loi par :

$$(n,h)(n',h') = (n(h.n'), hh').$$

Alors, $N \times H$, muni de cette loi est un groupe appelé **produit semi-direct** *de N par H relativement à φ, et noté*

$$N \rtimes_\varphi H \quad \textit{ou simplement} \quad N \rtimes H.$$

2) On a une suite exacte :

$$\begin{array}{ccccccc} 1 & \longrightarrow & N & \xrightarrow{i} & N \rtimes H & \xrightarrow{p} & H \longrightarrow 1 \\ & & n & \overset{i}{\longmapsto} & (n,1) & & \\ & & & & (n,h) & \overset{p}{\longmapsto} & h \end{array}$$

de sorte que $N \rtimes H$ est une extension de N par H.

Démonstration. C'est une vérification facile. Le calcul de l'associativité de la loi met en évidence la nécessité que φ soit une opération par automorphismes de groupes (et pas seulement ensembliste). On notera aussi que l' élément neutre de $N \rtimes H$ est (1,1), et que l'inverse de (n,h) est $(h^{-1}.n^{-1}, h^{-1})$.

Remarques 6.3.
1) Le groupe $N \rtimes H$ contient deux sous-groupes isomorphes respectivement à N et H :

$$\overline{N} = \{(n,1) \,|\, n \in N\}, \qquad \overline{H} = \{(1,h) \,|\, h \in H\}.$$

On a $\overline{N} \lhd\!\lhd N \rtimes H$, mais $\overline{H}$ n'est pas distingué en général.
2) On a $\overline{N} \cap \overline{H} = \{1\}$ et $N \rtimes H = \overline{N}\,\overline{H}$ car $(n,1)(1,h) = (n,h)$.
3) Si l'opération φ est non triviale (i.e. si $\varphi(h)$ n'est pas toujours égal à Id_N), le groupe obtenu n'est pas commutatif car $(1,h)(n,1) = (h.n,h)$ est en général distinct de (n,h).
4) Si on identifie N et H à $\overline{N}$ et $\overline{H}$ l'opération φ est décrite par $h.n = hnh^{-1}$.

On cherche maintenant des conditions permettant d'assurer qu'un groupe G est un produit.

Proposition 6.4 (Critères de décomposition en produit).
a) *Si on a une suite exacte :*

$$1 \longrightarrow N \xrightarrow{i} G \xrightarrow{p} H \longrightarrow 1,$$

et s'il existe un relèvement $\overline{H}$ de H, i.e. un sous-groupe $\overline{H}$ de G tel que la restriction de la projection p à $\overline{H}$ soit un isomorphisme de $\overline{H}$ sur H, le groupe G est isomorphe à un produit semi-direct $N \rtimes H$. Il revient au même de dire que p possède une section i.e. qu'il existe un homomorphisme $s : H \longrightarrow G$ tel que $p \circ s = \mathrm{Id}_H$. On dit alors que l'extension est **scindée**.
b) *Si on a deux sous-groupes N et H de G avec*
1) $N \lhd\!\lhd G$,
2) $N \cap H = \{1\}$,

3) $G = NH$,
alors on a $G \simeq N \rtimes H$.

Démonstration. C'est essentiellement l'analyse menée au début de b).

On peut caractériser les produits directs parmi les semi-directs :

Proposition 6.5.
Soient N, H, φ *comme dans la définition 6.2, soit* $G = N \rtimes_\varphi H$ *et soit* $\overline{H}$ *le sous-groupe des éléments* $(1, h)$. *Les propriétés suivantes sont équivalentes :*
1) φ *est trivial (i.e. on a* $\varphi(h) = \mathrm{Id}_N$ *pour tout* $h \in H$*),*
2) le sous-groupe $\overline{H}$ *est distingué dans* G,
3) la loi de groupe sur G *est celle du produit direct.*
(C'est le cas, en particulier, si l'extension est centrale, i.e. si on a $N \subset Z(G)$*).*

Remarques 6.6.
1) Si G est extension non scindée de N par H il se peut cependant que G soit produit semi-direct, $G \simeq N' \rtimes H'$. Par exemple, le lecteur vérifiera sans peine que $\mathbf{Z}/4\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$ (resp. D_4, cf. ci-dessous) sont tous deux produit direct (resp. semi-direct) de $\mathbf{Z}/4\mathbf{Z}$ par $\mathbf{Z}/2\mathbf{Z}$ et cependant ils contiennent tous deux un sous-groupe distingué d'ordre 2 tel que l'extension correspondante soit non scindée.
2) Un produit direct peut être isomorphe à un produit semi-direct non trivial. Le lecteur vérifiera par exemple que l'on a un isomorphisme

$$\mathfrak{S}_3 \times \mathbf{Z}/2\mathbf{Z} \simeq \mathfrak{S}_3 \rtimes \mathbf{Z}/2\mathbf{Z}.$$

Exemples 6.7.
1) Le groupe symétrique $\mathfrak{S}_n$
On a la suite exacte définie par la signature :

$$1 \longrightarrow \mathfrak{A}_n \longrightarrow \mathfrak{S}_n \xrightarrow{\epsilon} \{-1, 1\} \longrightarrow 1.$$

Si τ est une transposition, on a une section s de ϵ en posant $s(1) = \mathrm{Id}$, $s(-1) = \tau$ et donc, d'après 6.4 :

$$\mathfrak{S}_n \simeq \mathfrak{A}_n \rtimes \{-1, 1\} \simeq \mathfrak{A}_n \rtimes \mathbf{Z}/2\mathbf{Z}$$

et le produit n'est pas direct.

2) Le groupe diédral D_n.

Il s'agit du groupe des isométries du plan euclidien conservant un polygone régulier à n côtés. Il contient les n rotations $\rho(0, 2k\pi/n)$ pour $k = 0, \ldots, n-1$ (0 désigne le centre du polygone) et les n réflexions (i.e. symétries) par rapport aux droites passant par 0 et les sommets ou les milieux des côtés du polygone (attention aux cas n pair ou impair). Le sous-groupe des rotations est distingué et isomorphe à $\mathbf{Z}/n\mathbf{Z}$. Comme $|D_n| = 2n$, on a donc une suite exacte :

$$1 \longrightarrow \mathbf{Z}/n\mathbf{Z} \longrightarrow D_n \xrightarrow{p} \mathbf{Z}/2\mathbf{Z} \longrightarrow 1$$

et un isomorphisme

$$D_n \simeq \mathbf{Z}/n\mathbf{Z} \rtimes \mathbf{Z}/2\mathbf{Z}$$

car n'importe quelle réflexion fournit une section de p.
On notera au passage l'isomorphisme $D_3 \simeq \mathfrak{S}_3$.

3) Deux contre-exemples.

1) Le groupe cyclique $\mathbf{Z}/8\mathbf{Z}$ n'est pas de la forme $N \rtimes H$. En effet, comme $\mathbf{Z}/8\mathbf{Z}$ est abélien, le produit serait direct, or $\mathbf{Z}/8\mathbf{Z}$ n'est isomorphe ni à $\mathbf{Z}/4\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$ ni à $(\mathbf{Z}/2\mathbf{Z})^3$ qui sont les seuls possibles.

2) Le groupe des quaternions $\mathbf{H}_8$ n'est pas non plus décomposable en produit $N \rtimes H$. En effet, si $|N| = 4$, on a $H \simeq \mathbf{Z}/2\mathbf{Z}$, et comme $-1 \in N$, il n'y a pas de section. Si $|N| = 2$, $N = \{1, -1\}$ et $\mathbf{H}_8/N \simeq V_4$, mais $\mathbf{H}_8$ n'a pas de sous-groupe isomorphe à V_4 (ou encore l'extension serait centrale et scindée, donc triviale).

7. Automorphismes de $\mathbf{Z}/n\mathbf{Z}$.

Lorsqu'on cherche à calculer les produits semi-directs $N \rtimes H$ on est tout d'abord amené à déterminer les opérations $\varphi : H \longrightarrow \mathrm{Aut}\, N$ et donc, avant tout, à préciser le groupe des automorphismes de groupe de N.
Le problème général n'est pas simple, là encore, et nous nous limiterons à quelques cas particuliers ($\mathbf{Z}/n\mathbf{Z}$, $\mathfrak{S}_n$,...), cf. aussi Exercices D 1,2.

Soit $n \in \mathbf{N}^*$, $n \geq 2$. Si $s \in \mathbf{Z}$, nous notons $\overline{s}$ son image dans $\mathbf{Z}/n\mathbf{Z}$.

Proposition 7.1.
Soit $s \in \mathbf{Z}$, les propriétés suivantes sont équivalentes :
1) s est premier avec n,
2) $\overline{s}$ est générateur du groupe $(\mathbf{Z}/n\mathbf{Z}, +)$,
3) $\overline{s} \in (\mathbf{Z}/n\mathbf{Z})^*$ groupe des éléments inversibles pour la multiplication de l'anneau $\mathbf{Z}/n\mathbf{Z}$.

Démonstration. Par Bézout, on a les équivalences :
s premier à $n \iff \exists \lambda, \mu \in \mathbf{Z},\ \lambda s + \mu n = 1 \iff \exists \lambda \in \mathbf{Z}\ \ \overline{\lambda}\overline{s} = \overline{1}$ dans $\mathbf{Z}/n\mathbf{Z} \iff \overline{s} \in (\mathbf{Z}/n\mathbf{Z})^*$.
D'autre part, si $\lambda \in \mathbf{Z}$ on a [2] $\overline{\lambda}\overline{s} = \lambda\overline{s} = \overline{1}$, et ceci signifie $\underbrace{\overline{s} + \cdots + \overline{s}}_{\lambda \text{ fois}} = \overline{1}$, donc $\overline{1} \in < \overline{s} >$ et donc $\mathbf{Z}/n\mathbf{Z} = < \overline{s} >$ autrement dit $\overline{s}$ est un générateur.

Définition 7.2 (La fonction d'Euler).
On appelle fonction d'Euler et on note $\varphi(n)$ le nombre d'entiers x tels que $1 \leq x \leq n$ et x premier avec n.

D'après 7.1, on a $\varphi(n) = |(\mathbf{Z}/n\mathbf{Z})^*|$.
Si p est premier, il est clair que l'on a $\varphi(p) = p - 1$ et $\varphi(p^\alpha) = p^{\alpha-1}(p-1), \alpha \in \mathbf{N}^*$.

Proposition 7.3.
On a un isomorphisme $\mathrm{Aut}\, \mathbf{Z}/n\mathbf{Z} \simeq (\mathbf{Z}/n\mathbf{Z})^*$. *En particulier,* $\mathrm{Aut}\, \mathbf{Z}/n\mathbf{Z}$ *est un groupe abélien, de cardinal $\varphi(n)$.*

Démonstration.
Soit $u \in \mathrm{Aut}\, \mathbf{Z}/n\mathbf{Z}$, alors $u(1)$ est un générateur de $(\mathbf{Z}/n\mathbf{Z}, +)$, donc $u(1)$ est dans $(\mathbf{Z}/n\mathbf{Z})^*$ et on vérifie que l'application $\tau : u \mapsto u(1)$ est un homomorphisme.

---

[2] Dans le cas de $\mathbf{Z}/n\mathbf{Z}$, comme dans celui de $\mathbf{Z}$, la loi de multiplication est définie à partir de l'addition répétée. En particulier un automorphisme de groupe préserve automatiquement la structure d'anneau.

En sens inverse, soit σ défini sur $(\mathbf{Z}/n\mathbf{Z})^*$ par $\sigma(s)x = sx$. On voit que $\sigma(s)$ est un endomorphisme de $(\mathbf{Z}/n\mathbf{Z}, +)$ car $s(x+y) = sx + sy$. C'est un automorphisme car $sx = 0 \implies x = 0$ puisque s est inversible. Enfin, il est clair que σ et τ sont réciproques l'un de l'autre.

Nous allons maintenant préciser la structure de $(\mathbf{Z}/n\mathbf{Z})^*$ suivant la décomposition en facteurs premiers de n.

Proposition 7.4.
Soit n un entier, $n = p_1^{\alpha_1} \dots p_r^{\alpha_r}$ avec les p_i premiers distincts et les α_i dans $\mathbf{N}^*$.
1) On a un isomorphisme d'anneaux

$$\mathbf{Z}/n\mathbf{Z} \simeq \prod_{i=1}^{r} \mathbf{Z}/p_i^{\alpha_i}\mathbf{Z},$$

2) on a un isomorphisme de groupes

$$(\mathbf{Z}/n\mathbf{Z})^* \simeq \prod_{i=1}^{r} (\mathbf{Z}/p_i^{\alpha_i}\mathbf{Z})^*,$$

3) on a

$$\varphi(n) = \prod_{i=1}^{r} \varphi(p_i^{\alpha_i}) = n \prod_{i=1}^{r} (1 - 1/p_i).$$

Démonstration. Le point 1) résulte du lemme chinois (6.1). On obtient alors le point 2) en passant aux éléments inversibles et le calcul de $\varphi(n)$ s'ensuit aussitôt.

Pour être complet il nous reste à examiner la structure des $(\mathbf{Z}/p^\alpha\mathbf{Z})^*$ pour p premier. On a d'abord le résultat suivant :

Lemme 7.5.
Si p est un nombre premier on a un isomorphisme $(\mathbf{Z}/p\mathbf{Z})^* \simeq \mathbf{Z}/(p-1)\mathbf{Z}$.

Démonstration. Voir Chapitre III 2.7 ou [S1], nous montrerons plus généralement que le groupe multiplicatif d'un corps fini est cyclique.

Il faut ensuite distinguer les cas $p = 2$ ou p impair.

Proposition 7.6.
Si p est un nombre premier ≥ 3 et α un entier ≥ 2 on a :

$$(\mathbf{Z}/p^\alpha\mathbf{Z})^* \simeq \mathbf{Z}/\varphi(p^\alpha)\mathbf{Z} \simeq \mathbf{Z}/p^{\alpha-1}(p-1)\mathbf{Z}.$$

Démonstration. On commence par établir le

Lemme 7.7.
Si $k \in \mathbf{N}^*$, on a $(1+p)^{p^k} = 1 + \lambda\, p^{k+1}$ avec $\lambda \in \mathbf{N}^*$, premier à p.

Démonstration (du lemme 7.7). On raisonne par récurrence sur k. Pour $k = 1$ on a la formule :

$$(1+p)^p = 1 + \mathrm{C}_p^1\, p + \dots + \mathrm{C}_p^i\, p^i + \dots + p^p$$

et, pour $1 \leq i < p$, p divise C_p^i donc, pour $i \geq 2$ et $i < p$, $p^3 \mid \mathrm{C}_p^i p^i$ et comme $p \geq 3$, p^3 divise aussi p^p de sorte qu'on a, en définitive :

$$(1+p)^p = 1 + p^2 + up^3 = 1 + p^2(1+up)$$

et $\lambda = 1 + up$ est bien premier à p.
Supposons le lemme prouvé au rang k. On a donc :

$$(1+p)^{p^k} = 1 + \lambda\, p^{k+1}$$

avec λ premier à p.
Il en résulte :

$$(1+p)^{p^{k+1}} = (1+\lambda\, p^{k+1})^p = 1 + \sum_{i=1}^{p-1} C_p^i \lambda^i p^{(k+1)i} + \lambda^p p^{(k+1)p}.$$

Pour $i = 1$ on a le terme λp^{k+2}, pour $i \geq 2$, p^{k+3} est en facteur et donc on a bien

$$(1+p)^{p^{k+1}} = 1 + p^{k+2}(\lambda + up).$$

De ce lemme résulte que $1+p$ est un élément d'ordre $p^{\alpha-1}$ de $(\mathbf{Z}/p^\alpha\mathbf{Z})^*$. On a en effet : $(1+p)^{p^{\alpha-1}} = 1 + \lambda\, p^\alpha \equiv 1 \pmod{p^\alpha}$ et $(1+p)^{p^{\alpha-2}} = 1 + \lambda\, p^{\alpha-1}$ avec $p \nmid \lambda$, donc $(1+p)^{p^{\alpha-2}} \neq 1$ dans $\mathbf{Z}/p^\alpha\mathbf{Z}$.
La proposition résulte alors de ce lemme : on considère l'homomorphisme surjectif naturel induit par l'identité de $\mathbf{Z}$:

$$\psi : (\mathbf{Z}/p^\alpha\mathbf{Z})^* \longrightarrow (\mathbf{Z}/p\mathbf{Z})^*.$$

Soit $x \in (\mathbf{Z}/p^\alpha\mathbf{Z})^*$ un élément tel que $\psi(x)$ engendre $\mathbf{Z}/(p-1)\mathbf{Z}$ (cf. 7.5). L'ordre de x est un multiple de $p-1$ et donc, dans le groupe $< x >$ il y a un élément y d'ordre $p-1$. Mais alors comme $(\mathbf{Z}/p^\alpha\mathbf{Z})^*$ est abélien, $y(1+p)$ est d'ordre $p^{\alpha-1}(p-1)$ en vertu du lemme ci-dessous et le groupe est cyclique.

Lemme 7.8.
Soient a et b des éléments d'ordre p et q d'un groupe G. On suppose que a et b commutent et que p et q sont premiers entre eux, alors ab est d'ordre pq.

Démonstration. Comme a et b commutent on a $(ab)^{pq} = a^{pq}b^{pq} = 1$ donc l'ordre de ab divise pq. Inversement, si $(ab)^n = 1 = a^n b^n$ on a, en élevant à la puissance q, $a^{nq}b^{nq} = 1$, donc $a^{nq} = 1$. Il en résulte que l'ordre p de a divise nq donc aussi n puisque p et q sont premiers entre eux. Le même raisonnement appliqué à b montre que q divise n et donc pq divise n.

Remarque 7.9. Attention le lemme 7.8 ne subsiste pas si a et b ne commutent pas (sinon, par exemple, le groupe $\mathfrak{S}_3$ serait cyclique), ni si p et q ont un facteur commun, même en remplaçant le produit pq par le ppcm de p et q (regarder le cas $b = a^{-1}$).

Il reste à traiter le cas $p = 2$:

Proposition 7.10.
On a $(\mathbf{Z}/2\mathbf{Z})^* = \{1\}$, $(\mathbf{Z}/4\mathbf{Z})^* = \{1, -1\} \simeq \mathbf{Z}/2\mathbf{Z}$.
Pour $\alpha \geq 3$ on a $(\mathbf{Z}/2^\alpha\mathbf{Z})^* \simeq \mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/2^{\alpha-2}\mathbf{Z}$.

On notera que le groupe n'est donc pas cyclique si $\alpha \geq 3$.
Démonstration. Les cas de 2 et 4 sont triviaux. Pour les autres on commence par prouver le lemme suivant :

Lemme 7.11.
Soit $k \in \mathbf{N}^*$, on a $(5)^{2^k} = 1 + \lambda 2^{k+2}$ avec λ impair.

Démonstration (de 7.11) On raisonne encore par récurrence ; pour $k = 1$ on a :

$$5^2 = (1+4)^2 = 1 + 3.8.$$

On suppose ensuite :

$$(5)^{2^k} = 1 + \lambda 2^{k+2}$$

et on a alors

$$(5)^{2^{k+1}} = (1 + \lambda 2^{k+2})^2 = 1 + \lambda 2^{k+3} + \lambda^2 2^{2k+4}$$

d'où le résultat.

Revenant à 7.10, on considère alors, si $\alpha \geq 3$, l'homomorphisme surjectif

$$\psi : (\mathbf{Z}/2^\alpha\mathbf{Z})^* \longrightarrow (\mathbf{Z}/4\mathbf{Z})^* = \{1, -1\} \simeq \mathbf{Z}/2\mathbf{Z}.$$

Si $N = \operatorname{Ker}\psi$, on a $|N| = 2^{\alpha-2}$ et l'élément 5 de N est d'ordre $2^{\alpha-2}$ par 7.11 donc N est cyclique et on a la suite exacte :

$$1 \longrightarrow \mathbf{Z}/2^{\alpha-2}\mathbf{Z} \longrightarrow (\mathbf{Z}/2^\alpha\mathbf{Z})^* \xrightarrow{\psi} \mathbf{Z}/2\mathbf{Z} \longrightarrow 1.$$

D'autre part, comme 1 et -1 ne sont pas égaux modulo 4, le sous-groupe $\{1, -1\}$ de $(\mathbf{Z}/2^\alpha\mathbf{Z})^*$ fournit un relèvement de $\mathbf{Z}/2\mathbf{Z}$, de sorte que l'extension est scindée, mais comme $(\mathbf{Z}/2^\alpha\mathbf{Z})^*$ est abélien, on a un produit **direct** :

$$(\mathbf{Z}/2^\alpha\mathbf{Z})^* \simeq \mathbf{Z}/2^{\alpha-2} \times \mathbf{Z}/2\mathbf{Z},$$

et ceci achève l'étude de $\operatorname{Aut}\mathbf{Z}/n\mathbf{Z}$.

Application : détermination des groupes d'ordre pq pour p, q premiers ; $p < q$.

Soit G un tel groupe, Q un q-Sylow de G (cf. 5.4). D'après 5.7, le nombre k des q-groupes de Sylow est un diviseur de p, congru à 1 modulo q, donc on a $k = 1$ et Q est distingué dans G.
Comme q est premier, on a $Q \simeq \mathbf{Z}/q\mathbf{Z}$, et de même $G/Q \simeq \mathbf{Z}/p\mathbf{Z}$. De plus si P est un p-Sylow quelconque, il fournit un relèvement de G/Q et donc G est produit semi-direct : $G \simeq \mathbf{Z}/q\mathbf{Z} \rtimes \mathbf{Z}/p\mathbf{Z}$. Il reste à calculer ces produits. On a vu ci-dessus l'isomorphisme $\operatorname{Aut}\mathbf{Z}/q\mathbf{Z} \simeq \mathbf{Z}/(q-1)\mathbf{Z}$.
L'opération de $\mathbf{Z}/p\mathbf{Z}$ sur $\mathbf{Z}/q\mathbf{Z}$ correspond donc à un homomorphisme

$$\varphi : \mathbf{Z}/p\mathbf{Z} \longrightarrow \mathbf{Z}/(q-1)\mathbf{Z}.$$

Deux cas sont possibles :

1) $p \nmid q-1$, le seul φ possible est trivial, le produit est direct et $G \simeq \mathbf{Z}/pq\mathbf{Z}$ est cyclique (cf. 6.1),

2) $p \mid q-1$, $\mathbf{Z}/(q-1)\mathbf{Z}$ possède un unique sous-groupe d'ordre p et il y a donc une opération non triviale, donc un vrai produit semi-direct. De plus deux telles opérations φ, ψ "diffèrent" d'un automorphisme de $\mathbf{Z}/p\mathbf{Z}$ et le lemme facile suivant montre qu'alors les produits correspondants sont isomorphes :

Lemme 7.12.
Soient φ, ψ deux opérations de H sur N et $\alpha \in \operatorname{Aut} H$ tel que le diagramme ci-dessous soit commutatif :

$$\begin{array}{ccc} H & \xrightarrow{\varphi} & \operatorname{Aut}(N) \\ {\scriptstyle\alpha}\downarrow & \nearrow_{\psi} & \\ H & & \end{array} \qquad \text{i.e.} \quad \varphi = \psi \circ \alpha.$$

Alors, l'application $(n,h) \mapsto (n,\alpha(h))$ *est un isomorphisme de* $N \rtimes_\psi H$ *sur* $N \rtimes_\varphi H$.

En conclusion on a montré le théorème :

Théorème 7.13.
Soient p,q *des nombres premiers avec* $p < q$.
1) Si $p \nmid q-1$ *tout groupe d'ordre* pq *est cyclique. C'est le cas par exemple si* $pq = 15,\ 35,\ 51, \ldots$
2) Si $p \mid q-1$ *il y a deux groupes d'ordre* pq *non isomorphes, le groupe cyclique et un produit semi-direct non commutatif. Par exemple si* $p=2$ *on a* $\mathbf{Z}/2q\mathbf{Z}$ *et le groupe diédral* D_q. *Ce cas se produit encore pour* $pq = 21,\ 39, \ldots$ ([3])

Pour d'autres exemples de ces techniques, voir les exercices.

8. Structures des groupes symétrique $\mathfrak{S}_n$ et alterné $\mathfrak{A}_n$.

Le théorème essentiel de ce paragraphe est le suivant :

Théorème 8.1.
Le groupe $\mathfrak{A}_n$ *est simple pour* $n \geq 5$.

Corollaire 8.2.
On a $D(\mathfrak{A}_n) = \mathfrak{A}_n$ *pour* $n \geq 5$ *et* $D(\mathfrak{S}_n) = \mathfrak{A}_n$ *pour* $n \geq 2$.

Ces résultats sont historiquement très importants, dûs à E. Galois et liés à l'impossibilité de résoudre par radicaux l'équation générale de degré ≥ 5 (en fait le corollaire suffit). On trouve ici un premier exemple non banal (i.e. autre que le groupe $\mathbf{Z}/p\mathbf{Z}$ avec p premier) de groupe simple.

Démonstration. Notons que le corollaire, qui est conséquence évidente du théorème, peut aussi se démontrer directement : il est clair que l'on a $D(\mathfrak{A}_n) \subset D(\mathfrak{S}_n) \subset \mathfrak{A}_n$. Comme $\mathfrak{A}_n$ est engendré par les 3-cycles, il suffit de prouver que tout 3-cycle est, dans $\mathfrak{A}_n$, un commutateur. Soit $\sigma = (a,b,c)$ un 3-cycle, $\sigma^2 = (a,c,b)$ en est un autre, donc (cf. 4.10) σ^2 et σ sont conjugués dans $\mathfrak{A}_n$: il existe $\tau \in \mathfrak{A}_n$ avec $\sigma^2 = \tau^{-1}\sigma\tau$ d'où $\sigma = \sigma^{-1}\tau^{-1}\sigma\tau$ cqfd ! Le lecteur montrera de même l'assertion concernant $D(\mathfrak{S}_n)$.

Passons maintenant au théorème :
Nous allons en donner une démonstration en deux temps : pour $n=5$ d'abord, par une méthode très élémentaire qui mettra en évidence l'intérêt de la connaissance des classes de conjugaison dans les questions de simplicité ; pour $n > 5$ ensuite, par réduction au cas $n=5$, suivant une technique que nous retrouverons pour les groupes orthogonaux.
Le principe des démonstrations de simplicité que nous donnerons dans cet ouvrage est le suivant. Soit H un sous-groupe distingué de G,

1) si $h \in H$, la classe de conjugaison de h est contenue dans H i.e. on a $\forall g \in G,\ ghg^{-1} \in H$,

([3]) *Il résulte du théorème de la progression arithmétique de Dirichlet, cf. [S1] Ch. VI, que, pour* p *fixé,* $p \geq 3$, *il existe une infinité de nombres premiers* q *satisfaisant l'une ou l'autre des deux conditions ci-dessus.*

2) si $h \in H$ et $g \in G$ le commutateur $c = ghg^{-1}h^{-1} = (ghg^{-1})h^{-1}$ est dans H et n'est pas, en général, conjugué de h, de sorte qu'on obtient ainsi une nouvelle classe de conjugaison, le but ultime étant de montrer qu'un système générateur de G est tout entier dans H.

1) Le théorème pour $n = 5$.

Le groupe $\mathfrak{A}_5$ a 60 éléments : le neutre, 15 éléments d'ordre 2 (produit de deux transpositions disjointes), 20 d'ordre 3 (3-cycles), 24 d'ordre 5 (5-cycles).
On a vu que les cycles d'ordre 3 sont conjugués dans $\mathfrak{A}_5$. Les éléments d'ordre 2 le sont aussi : si $\tau = (ab)(cd)(e)$ et $\tau' = (a'b')(c'd')(e')$ il existe $\sigma \in \mathfrak{A}_n$ tel que $\sigma(a) = a'$, $\sigma(b) = b'$, $\sigma(e) = e'$ et on a alors $\tau' = \sigma\tau\sigma^{-1}$.
Soit alors $H \lhd \mathfrak{A}_5$, $H \neq \{1\}$. Si H contient un élément d'ordre 3 (resp. 2) il les contient tous d'après ce qui précède. S'il contient un élément d'ordre 5, il contient le 5-Sylow engendré par cet élément, donc tous les 5-sous groupes de Sylow puisqu'ils sont conjugués, donc tous les éléments d'ordre 5.
Mais H ne peut contenir un seul des trois types d'éléments précédents (en plus du neutre) car ni $25 = 24+1$, ni $21 = 20+1$, ni $16 = 15+1$ ne divisent 60 (n'oublions pas que le cardinal de H divise $|\mathfrak{A}_5| = 60$). Donc H contient au moins deux des trois types, d'où $|H| \geq 15 + 20 + 1 = 36$ et donc $|H| = 60$, $H = \mathfrak{A}_5$.

Remarque 8.3. Les 24 éléments d'ordre 5 ne sont pas conjugués dans $\mathfrak{A}_5$ (sinon ils formeraient une orbite et 24 diviserait 60). On peut cependant éviter le recours à Sylow dans la démonstration précédente en remarquant que si a et b sont d'ordre 5, b est conjugué dans $\mathfrak{A}_5$ de a ou de a^2.

2) Le cas $n > 5$.

Posons $E = \{1, \ldots, n\}$. Soit $H \lhd \mathfrak{A}_n$, $H \neq \{1\}$ et soit $\sigma \in H$, $\sigma \neq 1$. On va se ramener au cas $n = 5$ et, pour ceci, fabriquer à partir de σ un élément non trivial de H qui n'agisse, en fait, que sur un ensemble à 5 éléments, donc qui ait $n - 5$ points fixes.
Vu les remarques ci-dessus, la méthode naturelle à notre disposition est de prendre un commutateur $\rho = \tau\sigma\tau^{-1}\sigma^{-1}$, avec $\tau \in \mathfrak{A}_n$. Ecrivant $\rho = (\tau\sigma\tau^{-1})\sigma^{-1}$, on a vu que ρ est dans H. Mais si on regarde ρ par l'autre bout : $\rho = \tau(\sigma\tau^{-1}\sigma^{-1})$ on constate que ρ est produit de deux éléments du type de τ de sorte que si τ a beaucoup de points fixes, il en sera de même de ρ.

Il ne reste plus qu'à mettre ces remarques en forme :

Comme $\sigma \neq 1$, il existe $a \in E$ tel que $b = \sigma(a) \neq a$. Soit $c \in E$ tel que $c \neq a, b, \sigma(b)$, soit τ le 3-cycle $\tau = (acb)$, de sorte que $\tau^{-1} = (abc)$ et soit $\rho = \tau\sigma\tau^{-1}\sigma^{-1} = (acb)(\sigma a, \sigma b, \sigma c)$. Comme $b = \sigma(a)$, l'ensemble $F = \{a, b, c, \sigma a, \sigma b, \sigma c\}$ a au plus 5 éléments et on a $\rho(F) = F$, $\rho|_{E-F} = \mathrm{Id}_{E-F}$.
Quitte à rajouter, au besoin, des éléments à F, on peut supposer $|F| = 5$. On note enfin que ρ est distinct de 1, car $\rho(b) = \tau\sigma(b) \neq b$ puisque $\sigma(b) \neq \tau^{-1}(b) = c$.
Soit alors $\mathfrak{A}(F)$ l'ensemble des permutations paires de F, $\mathfrak{A}(F)$ est isomorphe à $\mathfrak{A}_5$ et $\mathfrak{A}(F)$ se plonge dans $\mathfrak{A}_n$ par $u \mapsto \overline{u}$ avec :

$$\overline{u}|_F = u \ ; \ \overline{u}|_{E-F} = \mathrm{Id}_{E-F}.$$

Posons $H_0 = \{u \in \mathfrak{A}(F) \,|\, \overline{u} \in H\} = H \cap \mathfrak{A}(F)$. Il est clair que H_0 est distingué dans $\mathfrak{A}(F)$ et qu'on a $\rho|_F \in H_0$ et $\rho|_F \neq \mathrm{Id}_F$. Comme $\mathfrak{A}(F) \simeq \mathfrak{A}_5$ est simple, on

a $H_0 = \mathfrak{A}(F)$. Soit alors u un cycle d'ordre 3 de $\mathfrak{A}(F)$, il est dans H_0, donc $\overline{u}$ qui est encore un cycle d'ordre 3 est dans H.
Mais, comme les 3-cycles sont conjugués dans $\mathfrak{A}_n$, ils sont tous dans H, et comme ils engendrent $\mathfrak{A}_n$, on a montré $H = \mathfrak{A}_n$, ce qui achève la démonstration.

Remarque 8.4. Le groupe $\mathfrak{A}_4$ n'est pas simple car $D(\mathfrak{A}_4)$ est un sous-groupe distingué d'ordre 4 isomorphe à V_4 :

$$D(\mathfrak{A}_4) = \{\mathrm{id};\ (12)(34);\ (13)(24);\ (14)(23)\}.$$

Corollaire 8.5.
Pour $n \geq 5$ les sous-groupes distingués de $\mathfrak{S}_n$ sont $\{1\}, \mathfrak{A}_n, \mathfrak{S}_n$.

Démonstration. Soit $H \lhd \mathfrak{S}_n$. On a alors $H \cap \mathfrak{A}_n \lhd \mathfrak{A}_n$, donc $H \cap \mathfrak{A}_n = \{1\}$ ou $\mathfrak{A}_n$.
Si $H \cap \mathfrak{A}_n = \mathfrak{A}_n$, on a $H = \mathfrak{A}_n$ ou $\mathfrak{S}_n$.
Si $H \cap \mathfrak{A}_n = \{1\}$, la signature ε induit un isomorphisme de H sur $\varepsilon(H) \subset \{-1, 1\}$ de sorte que $|H| \leq 2$. Si $|H| = 2$, on a $H = \{1, \sigma\}$, mais si $\tau \in \mathfrak{S}_n$, comme $\tau\sigma\tau^{-1} \in H$ et $\tau\sigma\tau^{-1} \neq 1$, c'est que l'on a $\tau\sigma\tau^{-1} = \sigma$ donc σ est central, or le centre de $\mathfrak{S}_n$ est trivial (cf. 3.2.2) d'où une contradiction, donc $H = \{1\}$.

Corollaire 8.6.
Soit H un sous-groupe d'indice n de $\mathfrak{S}_n$, alors H est isomorphe à $\mathfrak{S}_{n-1}$.

Démonstration. Pour $n \leq 3$ l'assertion est évidente. Pour $n = 4$, si $H \not\simeq \mathfrak{S}_3$, H est cyclique (cf. 7.13) mais c'est absurde car $\mathfrak{S}_4$ ne contient pas d'élément d'ordre 6.
Supposons $n \geq 5$ et posons $G = \mathfrak{S}_n$, alors G, donc aussi H, opère par translation à gauche sur $E = G/H$ (cf. § 4 Exemple C) et on a donc un homomorphisme $\varphi : G \longrightarrow \mathfrak{S}(E) \simeq \mathfrak{S}_n$.
Montrons que φ est injectif : on a vu (*loc. cit.*) que l'on a $\operatorname{Ker}\varphi = \bigcap_{a \in G} aHa^{-1}$ et comme $\operatorname{Ker}\varphi$ est distingué dans $\mathfrak{S}_n$ et $\operatorname{Ker}\varphi \subset H$, le corollaire 8.5 montre qu'on a $\operatorname{Ker}\varphi = \{1\}$ (puisque $(n-1)! < n!/2$).
Pour une raison de cardinal φ est donc un isomorphisme. D'autre part, comme H est le stabilisateur de la classe $\overline{1} = H$ (cf. § 4 Exemple C), $\varphi(H)$ est le stabilisateur d'un point dans $\mathfrak{S}_n$, et c'est donc un sous-groupe isomorphe à $\mathfrak{S}_{n-1}$ (voir aussi ci-après, Proposition 8.10).

Calcul des automorphismes de $\mathfrak{S}_n$.

Les motivations sont encore celles du paragraphe 6 : le calcul des extensions.

Rappelons que parmi les automorphismes de G il y a déja les automorphismes intérieurs qui forment un groupe $\operatorname{Int} G \subset \operatorname{Aut} G$. De plus, on a la suite exacte :

$$\begin{array}{ccccccc} 1 & \longrightarrow & Z(G) & \longrightarrow & G & \longrightarrow & \operatorname{Int} G \longrightarrow 1, \\ & & & & g & \longmapsto & i_g \end{array}$$

où $Z(G)$ est le centre de G et i_g l'automorphisme donné par $i_g(x) = gxg^{-1}$.
Dans le cas présent, il n'y en a pas d'autres (ou presque) :

Théorème 8.7.
Pour $n \neq 6$, tout automorphisme de $\mathfrak{S}_n$ est intérieur : $\operatorname{Aut}\mathfrak{S}_n = \operatorname{Int}\mathfrak{S}_n$.

On notera que si on a, de plus, $n \geq 3$, on a $\operatorname{Aut}\mathfrak{S}_n \simeq \mathfrak{S}_n$ puisque le centre de $\mathfrak{S}_n$ est alors trivial (cf. 3.2.2).

Démonstration. La clé de la démonstration est une traduction des propriétés géométriques des éléments de $\mathfrak{S}_n$ en propriétés algébriques. On entend (cf. § 4) par propriétés géométriques celles relatives à l'opération de $\mathfrak{S}_n$ sur $\{1, \ldots, n\}$ (par exemple le nombre de points fixes, les cycles ...) et par algébriques celles qui s'expriment uniquement en termes de groupes (ordre des éléments, propriétés de commutation ...).
Les propriétés algébriques se conservent par automorphisme, mais pas, *a priori*, les propriétés géométriques. Ce sont pourtant celles qui nous intéressent ici, témoin la proposition ci-dessous :

Proposition 8.8.
Soit $\varphi \in \mathrm{Aut}\ \mathfrak{S}_n$; si φ transforme transposition en transposition, φ est intérieur.

Démonstration. Le groupe $\mathfrak{S}_n$ est engendré par les transpositions $\tau_i = (1, i)$ pour $i \geq 2$. Pour chaque i, $\varphi(\tau_i)$ est donc une transposition. De plus, si $i \neq j$, τ_i et τ_j ne commutent pas, donc $\varphi(\tau_i)$ et $\varphi(\tau_j)$ non plus, donc ces transpositions ne sont pas disjointes.
Si on pose $\varphi(\tau_2) = (\alpha_1, \alpha_2)$, on peut donc supposer $\varphi(\tau_3) = (\alpha_1, \alpha_3)$ et on en déduit $\varphi(\tau_i) = (\alpha_1, \alpha_i)$ pour $i > 3$ avec $\{\alpha_1, \ldots, \alpha_n\} = \{1, \ldots, n\}$ (si on avait $\varphi(\tau_1) = (\alpha_2, \alpha_3)$, comme $(\alpha_1\alpha_2)(\alpha_1\alpha_3)(\alpha_2\alpha_3) = (\alpha_1\alpha_3)$ on aurait par φ^{-1} $(12)(13)(1i) = (13)$ ce qui est faux). Les α_i sont tous distincts, sinon φ ne serait pas injective.
On a construit ainsi une permutation $\alpha \in \mathfrak{S}_n$ et on a $\alpha\tau_i\alpha^{-1} = (\alpha_1, \alpha_i) = \varphi(\tau_i)$, de sorte que φ et l'automorphisme intérieur i_α coïncident sur les α_i qui engendrent $\mathfrak{S}_n$, donc on a $\varphi = i_\alpha$.

Pour utiliser 8.8 il nous reste à caractériser ***algébriquement*** les transpositions. D'abord, ce sont des éléments d'ordre 2. Mais, dès que $n \geq 4$, ce ne sont pas les seuls, il y a aussi les produits $(ab)(cd)$ ...
L'idée est d'étudier le ***centralisateur*** d'un élément d'ordre 2. ([4]) Nous indiquons précisément deux méthodes :

1) On calcule le cardinal d'un centralisateur ; le résultat est le suivant :

Lemme 8.9.
Soit $s \in \mathfrak{S}_n$, on suppose que $n = k_1 + 2k_2 + \cdots + nk_n$ avec $k_i \in \mathbf{N}$ et que s est produit de $k_1 + \cdots + k_n$ cycles disjoints, k_1 cycles d'ordre 1, k_2 d'ordre 2, ..., k_n d'ordre n. Alors, si $c(s)$ est le centralisateur de s, on a :

$$|c(s)| = \prod_{i=1}^{n} k_i!\, i^{k_i}.$$

Nous laissons la démonstration de ce lemme en exercice. Le théorème 8.7 en résulte aisément. En effet, si τ est une transposition, $\varphi(\tau)$ est d'ordre 2, donc produit de k transpositions disjointes. Par ailleurs, on a $\varphi(c(\tau)) = c(\varphi(\tau))$ et donc $|c(\tau)| = |c(\varphi(\tau))|$ d'où $2(n-2)! = 2^k\, k!\, (n-2k)!$ On voit aisément que ceci impose $k = 1$, sauf si $n = 6$ et $k = 3$ auquel cas on a bien $2.4! = 48 = 2^3 3!$. Par conséquent, sauf (peut-être) pour $n = 6$, φ est intérieur.

([4]) Les spécialistes de théorie des groupes savent bien que c'est une bonne idée !

2) L'autre méthode, où le cas $n = 6$ apparaît de façon plus naturelle, consiste à étudier la structure de groupe du centralisateur.
Notons déjà que l'on peut supposer $n \geq 6$. En effet, pour $n \geq 2$ on a $D(\mathfrak{S}_n) = \mathfrak{A}_n$, et comme φ transforme commutateur en commutateur, il conserve $\mathfrak{A}_n$ de sorte que l'image d'une transposition est un produit d'un nombre impair k de transpositions disjointes. Si $k \neq 1$ on a donc $k \geq 3$ et $n \geq 6$.
Soit d'abord $\tau = (a, b)$ une transposition. Si $s \in c(\tau)$, on a $s\tau s^{-1} = (s(a)s(b)) = (a, b)$. Si on pose $E = \{1, \ldots, n\}$ et $F = E - \{a, b\}$, on a donc :

$$s \in c(\tau) \Longleftrightarrow s(\{a, b\}) = \{a, b\} \Longleftrightarrow s(F) = F.$$

On a ainsi un homomorphisme surjectif :

$$r : c(\tau) \longrightarrow \mathfrak{S}(F) = \mathfrak{S}_{n-2}$$
$$s \longmapsto s_{|F}$$

et, comme $\operatorname{Ker} r = \{1, \tau\}$, on a la suite exacte :

$$1 \longrightarrow \mathbf{Z}/2\mathbf{Z} \longrightarrow c(\tau) \longrightarrow \mathfrak{S}_{n-2} \longrightarrow 1.$$

Soit maintenant $\tau = (a_1, a_2) \ldots (a_{2k-1}, a_{2k})$ un produit de k transpositions disjointes et posons $\tau_i = (a_{2i-1}, a_{2i})$. Comme les τ_i commutent entre elles, on a $\tau_i \in c(\tau)$. Le sous-groupe N de $c(\tau)$ engendré par les τ_i est de cardinal 2^k, formé d'éléments d'ordre 2, il est donc isomorphe à $(\mathbf{Z}/2\mathbf{Z})^k$. De plus, N est distingué dans $c(\tau)$. En effet, si s est dans $c(\tau)$, on a :

$$s\tau s^{-1} = (s(a_1), s(a_2)) \ldots (s(a_{2k-1}), s(a_{2k})) = \tau$$

donc, vu l'unicité de la décomposition en cycles disjoints, la conjugaison par s permute les τ_i : $s\tau_i s^{-1} = \tau_j$.
(On peut vérifier aisément que $c(\tau)/N$ est isomorphe à $\mathfrak{S}_k \times \mathfrak{S}_{n-2k}$, mais nous n'utiliserons pas ce fait).

Revenons au théorème, et supposons que pour une transposition τ, $\varphi(\tau) = \tau'$ soit un produit de k transpositions avec $k \geq 3$. Les centralisateurs $c(\tau)$ et $c(\tau')$ sont alors isomorphes, donc $c(\tau)$ contient un sous-groupe distingué isomorphe à $(\mathbf{Z}/2\mathbf{Z})^k$. Par r on en déduit que $\mathfrak{S}_{n-2}$ possède un sous-groupe distingué isomorphe à $(\mathbf{Z}/2\mathbf{Z})^l$, avec $l = k$ ou $k - 1$, donc $l > 0$.
Mais on connaît les sous-groupes distingués de $\mathfrak{S}_{n-2}$ (cf. 8.5) et on voit que ceci n'est possible que si $n - 2 = 4$ (donc $n = 6$) ou $n - 2 = 2$ (donc $n = 4$). Comme on a supposé $n \geq 6$ on a donc prouvé que, sauf peut-être pour $n = 6$, l'image d'une transposition par φ est une transposition, de sorte que φ est intérieur.

Il reste à examiner le cas $n = 6$.
Nous allons prouver qu'il est exceptionnel : on a $\operatorname{Aut} \mathfrak{S}_6 \neq \operatorname{Int} \mathfrak{S}_6$.
On commence par remarquer que si on a une bijection $f : E \longrightarrow F$ entre deux ensembles, on en déduit un isomorphisme de groupes :

$$\varphi : \mathfrak{S}(E) \longrightarrow \mathfrak{S}(F) \quad \text{par} \quad \varphi(\sigma) = f\sigma f^{-1}.$$

En particulier, bien sûr, si $|E| = n$, on a $\mathfrak{S}(E) \simeq \mathfrak{S}_n$. On note, de plus, que φ transforme le stabilisateur de $e \in E$ en celui de $f(e) \in F$.
On pose ensuite $X = \{1, \ldots, n\}$, $\mathfrak{S}(X) = \mathfrak{S}_n$ et on note $S(i)$ le stabilisateur de i :

$$S(i) = \{\sigma \in \mathfrak{S}_n \mid \sigma(i) = i\}.$$

C'est un sous-groupe d'indice n de $\mathfrak{S}_n$, isomorphe à $\mathfrak{S}_{n-1}$ et les $S(i)$ sont conjugués car, si $\tau(i) = j$, on a $\tau S(i)\tau^{-1} = S(j)$.
On a alors la proposition suivante :

Proposition 8.10.
Pour $n \neq 4$ les propriétés suivantes sont équivalentes :
1) Aut $\mathfrak{S}_n = \text{Int } \mathfrak{S}_n$.
2) Les sous-groupes d'indice n de $\mathfrak{S}_n$ sont tous conjugués (donc ce sont les $S(i)$).

Démonstration. Prouvons seulement (non 2) $\Longrightarrow$ (non 1), le reste ne nous est pas utile et est laissé à titre d'exercice.
Soit donc H un sous-groupe d'indice n, non conjugué des $S(i)$. On a vu dans la démonstration de 8.6 qu'on a un isomorphisme :

$$\varphi : \mathfrak{S}_n \longrightarrow \mathfrak{S}(\mathfrak{S}_n/H)$$

obtenu en faisant opérer $\mathfrak{S}_n$ par translation sur $\mathfrak{S}_n/H$. Dans cet isomorphisme, $\varphi(H)$ est le stabilisateur de la classe $H = \overline{1}$.
Choisissant une bijection f de $\mathfrak{S}_n/H$ sur $X = \{1, \ldots, n\}$ telle que $f(\overline{1}) = 1$, on en déduit un isomorphisme :

$$\psi : \mathfrak{S}(\mathfrak{S}_n/H) \longrightarrow \mathfrak{S}_n$$

et $\varphi(H)$ s'envoie par ψ sur $S(1)$. On en déduit que $\psi \circ \varphi$ est un automorphisme de $\mathfrak{S}_n$ qui vérifie $\psi \circ \varphi(H) = S(1)$, mais, comme H et $S(1)$ ne sont pas conjugués, $\psi \circ \varphi$ ne saurait être intérieur.

On peut maintenant conclure :

Proposition 8.11.
On a Aut $\mathfrak{S}_6 \neq$ Int $\mathfrak{S}_6$.

Démonstration.
Il suffit de construire un sous-groupe H d'indice 6 de $\mathfrak{S}_6$, non conjugué des $S(i)$. Pour ceci, il suffit de trouver un tel H qui opère transitivement sur $\{1, \ldots, 6\}$.
Nous verrons une très belle méthode pour y parvenir à partir des groupes linéaires (cf. IV, 5.4.3). En voici une autre, fondée sur les théorèmes de Sylow :

Lemme 8.12.
Le groupe $\mathfrak{S}_5$ contient 6 sous-groupes de Sylow d'ordre 5.

Démonstration. Soit k le nombre de 5-Sylow de $\mathfrak{S}_5$. On a $k \equiv 1 \pmod 5$ et $k \mid 24$, donc $k = 1$ ou 6. Le cas $k = 1$ est à rejeter sinon P serait distingué, d'où la conclusion.

Revenons à 8.11. Soit X l'ensemble des 5-Sylow de $\mathfrak{S}_5$. Le groupe $\mathfrak{S}_5$ opère sur X par conjugaison, transitivement (cf. 5.7) et fidèlement (cf. 8.5). On a donc un homomorphisme injectif :

$$\varphi : \mathfrak{S}_5 \longrightarrow \mathfrak{S}(X) \simeq \mathfrak{S}_6$$

et comme $\mathfrak{S}_5$ est transitif sur X, le sous-groupe $\varphi(\mathfrak{S}_5)$ convient.

Remarque 8.13. On voit facilement que Int $\mathfrak{S}_6$ est un sous-groupe d'indice 2 de Aut $\mathfrak{S}_6$. En effet, si φ et ψ sont deux automorphismes "extérieurs" ils transforment tous deux la classe de conjugaison des transpositions (qui a 15 éléments) en celle des produits de 3 transpositions disjointes, mais alors, $\psi \circ \varphi$ conserve les transpositions donc est intérieur.

EXERCICES SUR LE CHAPITRE I

A. Groupes cycliques et $(\mathbf{Z}/n\mathbf{Z})^*$.

1) Montrer que tout sous-groupe (resp. tout quotient) d'un groupe cyclique est cyclique.

---

2) Montrer que, si d est un diviseur de n, $\mathbf{Z}/n\mathbf{Z}$ a exactement un sous-groupe et un quotient d'ordre d.

---

3) Soit G un groupe dont tout sous-groupe propre est cyclique. G est-il nécessairement cyclique, abélien ? Si G est de plus supposé abélien est-il cyclique ? Exemples.

---

4) Soit G un groupe, Z son centre. On suppose G/Z cyclique. Montrer que G est abélien.
Application : montrer que tout groupe d'ordre p^2, avec p premier, est isomorphe à $\mathbf{Z}/p^2\mathbf{Z}$ ou à $\mathbf{Z}/p\mathbf{Z} \times \mathbf{Z}/p\mathbf{Z}$.

---

5) Soient $a, n \in \mathbf{Z}$, avec $(a, n) = 1$, montrer qu'on a $a^{\varphi(n)} \equiv 1 \pmod{n}$.

---

6) Déterminer les entiers n pour lesquels $(\mathbf{Z}/n\mathbf{Z})^*$ est cyclique.

B. Sous-groupes distingués.

1) On suppose que H et K sont des sous-groupes distingués de G, montrer que le sous-groupe $H \bigcap K$ est distingué dans G.

---

2) Si on a $K \lhd H \lhd G$, a-t-on nécessairement $K \lhd G$? Et si K est un sous-groupe caractéristique de H ?

---

3) Un groupe dont tout sous-groupe est distingué est-il nécessairement abélien ?

---

4) Soit G un groupe fini, p le plus petit facteur premier de $|G|$, H un sous-groupe d'***indice*** p. Montrer que H est distingué dans G (utiliser § 4 Exemple C).
Étudier le cas $p = 2$. Application aux groupes d'ordre pq , $p < q$: retrouver le fait que le q-Sylow est distingué.

---

5) Soit $f : G \longrightarrow H$ un homomorphisme et soit $N \lhd\!\lhd H$, montrer qu'on a $f^{-1}(N) \lhd\!\lhd G$ (remarque très utile pour construire de gros sous-groupes distingués).
En déduire qu'un groupe d'ordre p^α a des sous-groupes distingués de tous les ordres p^i, avec $0 \leq i \leq \alpha$.

---

6) Soit G un groupe fini, p le plus petit facteur premier de $|G|$, H un sous-groupe d'***ordre*** p, distingué dans G. Montrer que H est central (faire opérer G sur H par conjugaison).

C. Opérations ; conjugaison ; Sylow.

1) Deux éléments de même ordre d'un groupe G sont-ils nécessairement conjugués ? Déterminer les groupes abéliens dans lesquels cette propriété est vraie.
Donner un exemple de groupe non abélien qui la vérifie (mais ne pas trop essayer de les trouver tous ...)

---

2) *Démonstration de Sylow* (Wielandt).
Soit G un groupe, on pose $|G| = n = p^\alpha m$, avec $p \nmid m$.
On considère l'ensemble X des parties de G de cardinal p^α et l'ensemble Y des p-sous-groupes de Sylow de G.

a) On fait opérer G sur X par translation à gauche. Soit $E \in X$, G_E le stabilisateur de E. Montrer qu'on a $|G_E| \leq p^\alpha$.

b) Montrer que $|G_E| = p^\alpha \iff E = Sx$ avec $x \in G$ et $S \in Y$. Montrer qu'alors on a $S = G_E$.

c) En déduire, en considérant les orbites de X sous G, la congruence

$$|X| \equiv m\,|Y| \pmod{p}.$$

d) Montrer qu'on a $|X| \equiv m \pmod{p}$ (soit par un calcul direct, soit en appliquant c) à $\mathbf{Z}/n\mathbf{Z}$).

e) Démontrer la congruence : $|Y| \equiv 1 \pmod{p}$ (qui prouve le premier théorème de Sylow et une partie du second).

---

3) Avec les notations et la méthode de 2), étudier les sous-groupes d'ordre p^i, avec $i \leq \alpha$.

---

4) *Lemme de Cauchy.*
Soit G un groupe abélien fini, p un diviseur premier de $|G|$. Montrer, sans utiliser le théorème de Sylow, que G contient un élément d'ordre p.

5) *Sylow par récurrence.*

a) Montrer le premier théorème de Sylow lorsque G est abélien (utiliser 4) et une récurrence).

b) Dans le cas général, on raisonne par récurrence sur n. On fait opérer G sur G par automorphismes intérieurs.
Montrer que, ou bien le centre Z de G est non trivial, ou bien il existe une orbite, non réduite à un point, de cardinal premier à p. Finir par récurrence, dans les deux cas, soit avec le centre, soit avec le stabilisateur.

---

6) Montrer qu'un groupe abélien est produit direct de ses sous-groupes de Sylow. (Pour des précisions sur la structure des groupes abéliens, cf. par exemple [Bbki] Alg. Ch. VII).

---

7) *Un exercice de botanique.*
On considère les groupes suivants : $\mathbf{Z}/n\mathbf{Z}$, D_n , $\mathfrak{S}_3$, $\mathfrak{S}_4$, $\mathfrak{S}_5$, $\mathfrak{A}_4$, $\mathfrak{A}_5$, $\mathbf{H}_8$, $GL(2, \mathbf{F}_p)$, avec p premier, $p \leq 5 \ldots$
Pour chacun d'entre eux on déterminera l'ordre du groupe, de ses éléments, les classes de conjugaison, les centralisateurs des éléments, les sous-groupes, éventuellement distingués (et, dans ce cas, les quotients), leurs classes de conjugaison et leurs normalisateurs, les sous-groupes de Sylow, leur nombre, les groupes $Z(G)$, $D(G)$, en reconnaissant pour chaque sous-groupe d'éventuels isomorphismes avec des groupes connus.

D. Automorphismes ; produits.

1) Montrer qu'on a $\mathrm{Aut}\left((\mathbf{Z}/p\mathbf{Z})^n\right) \simeq GL(n, \mathbf{F}_p)$.

---

2) Calculer $\mathrm{Aut}\, G$ pour $G = V_4$, $\mathfrak{S}_3$, D_4, $\mathbf{Z}/4\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$, $\mathbf{H}_8$ (les trois derniers sont plus difficiles).

---

3) Prouver le lemme 7.12.

---

4) Soient $\varphi, \psi : H \longrightarrow \mathrm{Aut}\, N$ deux homomorphismes. Soit $u \in \mathrm{Aut}\ N$, on suppose qu'on a pour tout $h \in H$, $\psi(h) = u\varphi(h)u^{-1}$, autrement dit φ et ψ "diffèrent" d'un automorphisme intérieur de $\mathrm{Aut}\, N$.
Montrer qu'alors les produits semi-directs $N \rtimes_\varphi H$ et $N \rtimes_\psi H$ sont isomorphes.

---

5) Montrer que si pour tous $a, b \in G$ on a $(ab)^2 = a^2b^2$, G est abélien, ceci vaut en particulier si pour tout a, on a $a^2 = 1$.
Construire, comme produit semi-direct, un groupe d'ordre 27 où tout élément est d'ordre 3 mais qui n'est pas commutatif, bien qu'il vérifie $a^3b^3 = (ab)^3$ pour tous a, b.

E. Classification des groupes de petits ordres.

Le thème des exercices suivants est de trouver pour n petit (soit en valeur absolue, soit par son nombre de facteurs premiers) tous les groupes d'ordre n, à isomorphisme près.
Pour cela, on utilisera deux types d'ingrédients :

a) Des théorèmes permettant d'affirmer qu'un groupe n'est pas simple, au seul vu de son cardinal. Les théorèmes de Sylow permettent parfois d'y parvenir. Nous utiliserons aussi des théorèmes de Burnside.
Il faut connaître le résultat le plus spectaculaire (mais très difficile !) dans cette direction : le théorème de Feit-Thompson qui affirme que tout groupe simple non banal (i.e. $\neq \mathbf{Z}/p\mathbf{Z}$) est d'ordre **pair**.

b) Les seules techniques de calcul des extensions à notre disposition, à savoir les produits directs ou semi-directs.

1) Déterminer les groupes d'ordre 8 (utiliser, par exemple, les exercices D5, B4 et A4).

---

2) *Utilisation de Sylow pour montrer qu'un groupe n'est pas simple.*

a) La première méthode est de prouver que le nombre k de p-sous-groupes de Sylow vaut 1 et ceci en utilisant les deux remarques suivantes :

1) on a $k \equiv 1 \pmod{p}$

2) on a $k \mid |G|$ et précisément, si $|G| = p^\alpha m$, $k \mid m$.

Exemples
$n = p^\alpha q$, p, q premiers $p > q$ (cf. aussi B4) (par exemple : $n = 18, 54, 50\ldots$) ;
$n = p^\alpha q^\beta$ avec $p^\alpha < q + 1$ (par exemple $n = 20, 28, 44, 88, 99, 100\ldots$) ;
$n = p^\alpha q^\beta$ lorsque, plus généralement, aucun p^i, $i \leq \alpha$ n'est congru à 1 modulo q (par exemple $n = 40, 45\ldots$) ; ou même n produit de plus de 3 facteurs premiers, dans certains cas favorables (par exemple : $n = 42, 255, 84\ldots$)

b) Lorsque cette méthode échoue, on peut parfois conclure en dénombrant pour un p fixé le nombre d'éléments qui sont dans les p-Sylow et en constatant qu'il reste peu de chose pour les autres.
Exemples
$n = 12, 30, 56\ldots$ ou encore $n = pqr$ avec p, q, r premiers distincts (cette méthode fonctionne bien pour $\alpha = 1$ et k proche de m.)

c) Si G a k sous-groupes de Sylow, comme ceux-ci forment une orbite, on a un homomorphisme $\varphi : G \longrightarrow \mathfrak{S}_k$ dont le noyau peut fournir un sous-groupe distingué non trivial. C'est le cas, par exemple, si $|G| > k!$ ou même si $|G| \nmid k!$
Exemples
$n = 12, 24, 36, 48$; $n = p^\alpha q$ avec : $p^\alpha \nmid (q-1)!$

d) *Applications :*
Montrer qu'un sous-groupe d'ordre $p^2 q$ a un sous-groupe distingué non trivial.
Montrer qu'aucun groupe d'ordre $n < 60$ n'est simple (non banal, i.e. $\neq \mathbf{Z}/p\mathbf{Z}$).

---

3) Soit G un groupe, S un 2-sous-groupe de Sylow. On suppose S cyclique et $|G| > 2$. Montrer que G n'est pas simple. En déduire que, si G est simple et $|G|$

pair, $|G|$ est multiple de 4, (ainsi, par exemple un groupe de cardinal 90 n'est pas simple).
(Faire opérer G sur G par translation et considérer la signature de la permutation induite sur G par le générateur s de S, on voit que $\varepsilon(s) = -1$ d'où un homomorphisme non trivial dans $\{-1, 1\}$).
Reprendre 2d) pour $n \leq 100$, $n \neq 60$.

---

4) Montrer que les groupes d'ordre 12 sont produits (directs ou semi-directs) de groupes à 3 ou 4 éléments ([1]). Déterminer tous les groupes d'ordre 12 à isomorphisme près (il y en a 5). Reconnaître parmi eux : $\mathbf{Z}/12\mathbf{Z}$, $\mathbf{Z}/6\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$, D_6, $\mathfrak{A}_4$, $\mathbf{Z}/6\mathbf{Z} \rtimes \mathbf{Z}/2\mathbf{Z}$, $\mathfrak{S}_3 \times \mathbf{Z}/2\mathbf{Z}$, $\mathfrak{S}_3 \rtimes \mathbf{Z}/2\mathbf{Z}$, pour d'éventuels nouveaux, traiter l'exercice C7.

---

4 bis) Même question pour les groupes d'ordre 18 (attention, il y a deux actions non conjuguées de $\mathbf{Z}/2\mathbf{Z}$ sur $(\mathbf{Z}/3\mathbf{Z})^2$, cf. D4).

---

5) Soit G un groupe de cardinal 24.
Montrer que, si aucun sous-groupe de Sylow (d'ordre 3 ou 8) de G n'est distingué, G est isomorphe à $\mathfrak{S}_4$ (faire opérer G sur ses 3-Sylow).
Déterminer les groupes d'ordre 24 (il y en a 15).

---

6) Montrer que tout groupe d'ordre 255 est cyclique (cf. aussi F4).

---

7) Déterminer les groupes d'ordre p^2q , $p \neq q$ (il y a divers cas suivant que $p \mid q-1$, $q \mid p-1$, $q \mid p+1$; il faut connaître un peu la structure de $GL(2, \mathbf{F}_p)$).

---

8) Soit p premier, $p \geq 3$ et G un groupe de cardinal p^3. On suppose que G contient un élément x d'ordre p^2. Soit $H = <x>$ le sous-groupe engendré.

a) Montrer que H est distingué dans G.

b) Soit $y \notin H$. Montrer qu'on a $y^p \in H$. On pose $y^p = x^m$. Montrer que $p \mid m$.

c) On pose $y^{-1}xy = x^n$. Calculer $y^{-l}x^r y^l$ en fonction de l, r, n.

d) On cherche un entier $k \in \mathbf{N}$ tel que yx^k soit d'ordre p.
Montrer que le problème se ramène à résoudre, en k , dans $\mathbf{Z}/p^2\mathbf{Z}$, l'équation :

$$m + k(1 + n + \cdots + n^{p-1}) = 0.$$

Montrer que cette équation a toujours des solutions.

e) Déduire de ce qui précède la liste des groupes d'ordres p^3 (il y a 3 groupes commutatifs et 2 produits semi-directs).

---

9) Compléter pour $n < 60$ la liste des groupes, à isomorphisme près (on évitera les cas $n = 16, 32, 48, 54$).

([1]) Plus généralement, toute extension G de deux groupes finis N et H de cardinaux premiers entre eux est scindée : théorème de Schur-Zassenhaus, cf. [H].

F. Le transfert. Théorème de Burnside.

Pour progresser un peu dans la classification nous introduisons dans les exercices suivants du matériel supplémentaire.

1) *L'homomorphisme de transfert.*
Soit G groupe fini, H un sous-groupe, G/H l'ensemble des classes à gauche. Soient $x_1, \ldots, x_n$ des représentants des classes de sorte qu'on a :

$$G/H = \{x_1H, \ldots, x_nH\}, \text{ avec } x_1 \in H.$$

Soit $X = \{x_1, \ldots, x_n\}$; G opère sur G/H par translation, donc sur X par : $g.x_i = x_j \iff (gx_i)H = x_jH$. Mieux, G opère sur $\{1, \ldots, n\}$ par : $g.x_i = x_{g.i}$.
On pose alors, pour $i \in \{1, \ldots, n\}$: $gx_i = x_{g.i}h_{i,g.i}$ avec $h_{i,g.i} \in H$.
Soit $D(H)$ le groupe dérivé de H, le quotient $H/D(H)$ est abélien et on note $\overline{h}$ la classe de h modulo $D(H)$. On pose alors :

$$V_{G\to H}(g) = \overline{\prod_{i=1}^{n} h_{i,g.i}}$$

Montrer que :

1) $V_{G\to H}$ est un homomorphisme de G dans $H/D(H)$, appelé **transfert** *(Verlagerung)* de G dans H.

2) $V_{G\to H}$ ne dépend pas du choix des x_i, (cf. [H] § 14).

---

2) *Théorème de Burnside.*
Soit G un groupe fini, P un p-Sylow de G. On suppose que P est contenu dans le centre de son normalisateur.
On va montrer que P possède un p-complément i.e. qu'il existe un sous-groupe $N \lhd G$ tel que $G/N \simeq P$.
Pour ceci, on établira les points suivants :

a) Soient $A_1, A_2 \subset P$ deux parties « distinguées » dans P (i.e. telles que $\forall g \in P$, $gA_ig^{-1} = A_i$). On suppose A_1, A_2 conjuguées dans G, montrer qu'elles le sont aussi dans $N_G(P)$, normalisateur de P dans G (considérer le normalisateur de A_2 et utiliser l'argument de Frattini, voir la démonstration du théorème 5.7).

b) On considère le transfert $V_{G\to P} : G \longrightarrow P$ (cf. 1) ci-dessus). Il suffit de prouver que V est surjectif. Pour ceci, si $u \in P$, on calculera $V(u)$ en utilisant pour représentants de G/P les éléments u^jx_i où $\omega(x_1), \ldots, \omega(x_\alpha)$ sont les orbites de G sous $< u >$. On trouvera $V(u) = \prod_{i=1}^{\alpha} x_i^{-1}u^{r_i}x_i$, avec $r_i = |\omega(x_i)|$, puis on utilisera a) et l'hypothèse $P \subset Z(N_G(P))$.

---

3) *Application.*
Soit G un groupe fini simple $\neq \mathbf{Z}/p\mathbf{Z}$, n son ordre, montrer qu'on a l'alternative suivante :

1) ou bien 12 divise n,

2) ou bien le plus petit facteur premier de n est au moins au cube.

---

4) Soit n un entier. On dit que n est **cyclique** si tout groupe d'ordre n est cyclique.
Montrer que pour qu'un entier n soit cyclique, il faut et il suffit que n soit premier à $\varphi(n)$, i.e. qu'on ait $n = p_1 \dots p_r$, avec les p_i premiers distincts et $p_i \nmid p_j - 1$, (utiliser 3).
Cas particulier : retrouver le cas des groupes d'ordre pq.

---

5) Même exercice qu'en 4) mais en remplaçant cyclique par abélien. Montrer que $n = p_1^{\alpha_1} \dots p_r^{\alpha_r}$ doit vérifier les conditions suivantes :
1) $1 \leq \alpha_i \leq 2$ pour tout i,
2) $\forall i, j,\ i \neq j,\ p_i \nmid p_j - 1$,
3) Pour tout i tel que $\alpha_i = 2$, et pour tout j, $p_j \nmid p_i + 1$.

---

6) Montrer que si G est simple et $|G| = 60$ on a $G \simeq \mathfrak{A}_5$ (montrer que G a cinq sous-groupes de Sylow d'ordre 4, cf. 2), et le faire opérer sur cet ensemble à cinq éléments.)

---

7) Si $|G| = 60$ et si G n'est pas simple, montrer qu'il est produit semi-direct (fabriquer des sous-groupes distingués d'ordre 12, 15 ou 20).
Déterminer tous les groupes d'ordre 60 (il y en a 13, difficile).

---

8) *Vers Feit-Thomson.*
Montrer qu'un groupe d'ordre n impair, avec $n < 2000$ est soit de la forme $\mathbf{Z}/p\mathbf{Z}$, soit non simple (utiliser les exercices 2) et 3) ci-dessus et un autre théorème de Burnside, que l'on admettra, cf. [S2] : tout groupe d'ordre $p^\alpha q^\beta$ est non simple pour $\alpha + \beta > 1$).

G. Groupes symétrique et alterné.

1) Montrer que les 2-Sylow de $\mathfrak{S}_4$ sont isomorphes à D_4. Combien y-en-a-t-il ?

---

2) Montrer que les 2-Sylow de $\mathfrak{A}_5$ sont isomorphes à V_4. Combien y-en-a-t-il ?

---

3) Montrer que $\mathfrak{A}_n$ est engendré par les carrés des éléments de $\mathfrak{S}_n$.

---

4) Montrer que $\mathfrak{A}_6$ est engendré par $\sigma = (1\ 2\ 3\ 4\ 5)$ et $\tau = (1\ 3\ 5)(2\ 4\ 6)$.
(Soit $H = <\sigma, \tau>$, calculer $\sigma\tau$ et en déduire une majoration de l'indice de H, puis utiliser 3a) et E 2c)).
Montrer qu'en revanche $\sigma = (1\ 2\ 3\ 4\ 5)$ et $\tau = (1\ 3\ 6)(2\ 5\ 4)$ n'engendrent par $\mathfrak{A}_6$.
Que peut-on dire du groupe $<\sigma, \tau>$? (difficile).

II. ANNEAUX,

propriétés arithmétiques

0. Rappels.

Nous supposons connue la notion d'anneau, et notamment d'anneau de polynômes. Les anneaux que nous étudierons sont supposés ***commutatifs*** (la théorie des anneaux et algèbres non commutatifs est intéressante, mais très différente, cf.[Bl], nous rencontrerons seulement, dans les prochains chapitres, l'algèbre de matrices $\mathbf{M}(n,k)$ et le corps des quaternions).
Les anneaux seront aussi supposés ***unitaires*** (les anneaux non unitaires, eux, n'ont guère d'intérêt et interviennent, essentiellement, comme idéaux des anneaux unitaires).
Nous supposons, de plus, que les homomorphismes conservent les éléments unités : si on a $f : A \longrightarrow B$, $f(1_A) = 1_B$ et que les sous-anneaux contiennent l'élément unité de l'anneau.
La notion d'idéal est supposée connue ainsi que celle d'anneau quotient. Nous utiliserons librement les résultats et les notions qui suivent :

a) Théorème d'isomorphisme.

Soit $f : A \to B$ un homomorphisme d'anneaux et $I = \mathrm{Ker}\, f$. Soit J un idéal de A contenu dans I et $p : A \to A/J$ la projection canonique. Alors :
1) il existe un unique homomorphisme $\overline{f} : A/J \to B$ tel que $f = \overline{f}p$ (on dit que f se factorise par A/J),
2) $\overline{f}$ est injectif si et seulement si on a $J = I$,
3) $\overline{f}$ est surjectif si et seulement si f l'est.
En particulier on a $\mathrm{Im}\, f \simeq A/\mathrm{Ker}\, f$.

b) Propriété universelle des anneaux de polynômes.

Soient A et B deux anneaux. La donnée d'un homomorphisme

$$f : A[X_1, \ldots, X_n] \to B$$

équivaut à la donnée de sa restriction à A (donc d'un homomorphisme de A dans B) et des images des X_i (donc de n éléments de B), cf. [L] Ch. V.

c) Opérations sur les idéaux.

On note (x) ou xA l'idéal engendré par x dans A i.e. l'ensemble des éléments de la forme xa pour $a \in A$. Un tel idéal est dit ***principal***

Une intersection quelconque d'idéaux est un idéal.

La ***somme*** d'une famille I_k d'idéaux de A est l'ensemble des sommes finies $\sum x_k$ avec $x_k \in I_k$. C'est un idéal qui est la borne supérieure des I_k pour l'inclusion. On le note $\sum I_k$. En particulier si $I_k = (f_k)$ on trouve ***l'idéal engendré*** par les f_k. Dans $\mathbf{Z}$ l'idéal somme de (x) et (y) est l'idéal engendré par le *pgcd* de x et y.
Un idéal I d'un anneau A est dit ***de type fini*** s'il est engendré par un nombre fini d'éléments i.e. s'il existe $x_1, x_2, \ldots, x_n \in I$ tels que tout $x \in I$ s'écrive $x = \sum a_i x_i$ avec des $a_i \in A$.

Le ***produit*** de deux idéaux I et J est l'idéal noté IJ engendré par les produits xy pour $x \in I$ et $y \in J$. On a $IJ \subset I \cap J$, mais la réciproque est fausse : dans $\mathbf{Z}$ l'idéal produit (resp. intersection) de (x) et (y) est l'idéal engendré par xy (resp. par le *ppcm* de x et y).

d) Algèbres.

Soit A un anneau. Une ***A-algèbre*** est un anneau B muni d'un homomorphisme (souvent injectif, mais ce n'est pas indispensable) $f : A \to B$. Elle est dite ***de type fini*** si elle est engendrée comme algèbre par un nombre fini d'éléments $x_1, \ldots, x_n$ de B i.e., si tout élément de B est une fonction polynômiale des x_i à coefficients dans A. Il revient encore au même (cf. b) de demander que B soit isomorphe à un quotient d'un anneau de polynômes $A[X_1, \ldots, X_n]$.

1. Quelques remarques sur les idéaux.

a) Idéaux d'un anneau quotient.

Soit A un anneau, I un idéal, $p : A \longrightarrow A/I$ la surjection canonique.
Alors les idéaux de A/I sont en bijection avec les idéaux de A contenant I via les deux applications, croissantes relativement à l'inclusion :

$$\begin{array}{rcl} J \supset I & \mapsto & p(J) \\ p^{-1}(J') & \hookleftarrow & J' \end{array}$$

Attention, l'image d'un idéal par un homomorphisme n'est pas en général un idéal. C'est vrai toutefois dans le cas où l'homomorphisme est surjectif. En revanche avec l'image réciproque tout se passe bien.

b) Anneaux intègres.

Définition 1.1.
Un anneau A est dit ***intègre*** *si on a :*
1) $A \neq \{0\}$,
2) $\forall a, b \in A,\ ab = 0 \implies a = 0$ ou $b = 0$.

Exemples 1.2.
1) Si A est un corps, A est intègre.
2) Un sous-anneau d'un anneau intègre l'est aussi.
3) Réciproquement, tout anneau intègre est un sous-anneau d'un corps, son corps des fractions $K = \mathrm{Fr}(A)$ (construit à partir de A comme $\mathbf{Q}$ à partir de $\mathbf{Z}$).

4) Si A est intègre, l'anneau des polynômes $A[X]$ est intègre.

5) En revanche, si I est un idéal de A, A/I n'est pas nécessairement intègre, même si A l'est. A cet égard on a la définition suivante :

Définition 1.3.
Soit A un anneau, I un idéal de A, I est dit **premier** *si et seulement si l'anneau A/I est intègre. Il revient au même d'imposer :*
1) $A \neq I$,
2) $\forall a, b \in A$, $ab \in I \implies a \in I$ ou $b \in I$.

Exemples 1.4.

1) Si $A = \mathbf{Z}$, $I = n\mathbf{Z}$ est premier si et seulement si $n = 0$ ou n premier (cf. § 3).

2) Si on a un homomorphisme $f : A \longrightarrow B$ et si I est un idéal premier de B, $f^{-1}(I)$ est un idéal premier de A.

c) Idéaux maximaux.

Définition 1.5.
Un idéal I de A est dit **maximal** *si $I \neq A$ et si I est maximal (relativement à l'inclusion) pour cette condition : si J est un idéal de A tel que $J \supset I$ et $J \neq A$, on a $J = I$.*

Exemple 1.6. Les idéaux maximaux de $\mathbf{Z}$ sont les $p\mathbf{Z}$ pour p premier.

Proposition 1.7.
Soit I un idéal de A. On a l'équivalence : I maximal $\iff$ A/I est un corps.

Corollaire 1.8.
Tout idéal maximal est premier.

Démonstration (de 1.7) On montre d'abord le lemme suivant :

Lemme 1.9.
Soit A un anneau, A est un corps si et seulement si on a :
1) $A \neq \{0\}$,
2) les seuls idéaux de A sont $\{0\}$ et A.

Démonstration. En effet si on a 1) et 2), et si $a \in A$ est non nul, l'idéal principal aA est différent de $\{0\}$, c'est donc A, donc on a $1 \in aA$ et a est inversible. Réciproquement, si A est un corps et I un idéal $\neq \{0\}$, soit $a \in I$, $a \neq 0$, alors on a $1 = a^{-1}a \in I$, donc $I = A$.

La proposition 1 résulte du lemme et de la description des idéaux de A/I, cf. a).

Notons que la réciproque du corollaire est fausse, par exemple $\{0\}$ est un idéal premier non maximal de $\mathbf{Z}$, ou encore, si k est un corps, l'idéal principal (X) de $k[X, Y]$ est premier mais non maximal.

Signalons sans démonstration le théorème suivant qui résulte aisément du théorème de Zorn :

Théorème 1.10 (Krull).
Soit I un idéal de A, $I \neq A$, il existe un idéal maximal m de A, contenant I.

2. Anneaux noethériens.

Il s'agit d'une propriété de finitude des anneaux, remarquablement stable par les opérations usuelles. Rappelons qu'un idéal I d'un anneau A est dit de type fini s'il est engendré par un nombre fini d'éléments :

$$I = (a_1, \ldots, a_n) = \{\lambda_1 a_1 + \cdots + \lambda_n a_n \mid \lambda_i \in A\}.$$

Proposition-définition 2.1.
Les propriétés suivantes sont équivalentes :
1) Tout idéal de A est de type fini.
2) Toute suite croissante $I_1 \subset I_2 \subset \ldots \subset I_n \subset \ldots$ d'idéaux de A est stationnaire, i.e. $(\exists N \in \mathbf{N},\ n \geq N \implies I_n = I_N)$.
3) Tout ensemble non vide d'idéaux de A a un élément maximal pour l'inclusion.
Un anneau qui vérifie 1) 2) ou 3) est dit ***noethérien.***

Démonstration.
1) $\implies$ 2). Comme la suite I_n est croissante, la réunion $I = \bigcup_{n \in \mathbf{N}} I_n$ est un idéal, donc est de la forme $I = (a_1, \ldots, a_k)$. Il existe $N \in \mathbf{N}$ tel que $a_1, \ldots, a_k \in I_N$, mais alors on a $I = I_N$ d'où le résultat.
2) $\implies$ 3). Soit E un ensemble non vide d'idéaux. Supposons que E n'ait pas d'élément maximal. On construit alors par récurrence une suite $(I_n)_{n \in \mathbf{N}}$ qui contredit 2) : on prend $I_1 \in E$ quelconque, puis, comme I_1 n'est pas maximal, on trouve $I_2 \in E$ avec $I_1 \subsetneq I_2$, etc. (on notera qu'on n'utilise pas le théorème de Zorn).
3) $\implies$ 1). Soit I un idéal et $E = \{\text{idéaux } J \text{ de } A \mid J \subset I \text{ et } J \text{ de type fini}\}$. L'ensemble E est non vide car (0) est dans E. Soit J un élément maximal de E et, si $J \neq I$, soit $a \in I - J$. Alors $J + (a)$ est encore de type fini, contenu dans I et contient strictement J ce qui contredit la maximalité de J. On a donc $J = I$ et I est bien de type fini.

Exemples 2.2.
1) Un anneau principal (i.e. intègre et dont tout idéal est principal ; cf. § 3) est noethérien. C'est le cas, par exemple, de $\mathbf{Z}$, des corps ...
2) Si A est noethérien, A/I l'est aussi (cf. § 1 a)).
3) L'anneau des polynômes à une infinité de variables $A = k[X_1, \ldots, X_n, \ldots]$ n'est pas noethérien car on a une suite croissante non stationnaire d'idéaux : $(X_1) \subset (X_1, X_2) \subset \ldots$ Pour d'autres exemples, voir les exercices.
4) Un sous-anneau d'un anneau noethérien ne l'est pas nécessairement ; par exemple si dans 3) k est un corps, l'anneau A est intègre, donc (cf. § 1 b)) c'est un sous-anneau d'un corps.

Le théorème de transfert de Hilbert fournit de nombreux exemples d'anneaux noethériens :

Théorème 2.3 (Hilbert).
Si A est noethérien, $A[X]$ est noethérien.

Démonstration. Soit I un idéal de $A[X]$. On définit, pour $n \in \mathbf{N}$:

$$d_n(I) = \{\text{coefficients dominants des éléments de } I \text{ qui sont de degré } n\} \bigcup \{0\}.$$

On voit aisément que $d_n(I)$ est un idéal de A. De plus, on a les propriétés suivantes :
1) si $I \subset J$, $d_n(I) \subset d_n(J)$ pour tout n,
2) si $n \in \mathbf{N}$, $d_n(I) \subset d_{n+1}(I)$,
3) si $I \subset J$, $I = J \iff \forall n \in \mathbf{N},\ d_n(I) = d_n(J)$.
Les deux premières propriétés sont évidentes (pour 2) il suffit de multiplier par X). Pour la troisième, supposons que l'on ait $d_n(I) = d_n(J)$ pour tout n et $I \neq J$. Soit $P \in J - I$, $P \neq 0$, de degré minimal k. Le coefficient dominant de P est dans $d_k(J) = d_k(I)$. Soit $Q \in I$ de même degré et de même coefficient dominant que P. Alors, on a $P - Q \in J - I$ et $d^\circ(P - Q) < k$ ce qui contredit la minimalité du degré de P.

Revenons au théorème : soit $I_0 \subset I_1 \subset \ldots \subset I_n \subset \ldots$ une suite croissante d'idéaux de $A[X]$. Comme A est noethérien la famille $d_k(I_n)$ pour $(k, n) \in \mathbf{N} \times \mathbf{N}$ admet un élément maximal $d_l(I_m)$. Pour $k \leq l$, la suite $d_k(I_n)$ est croissante par rapport à n, donc il existe n_k tel que pour $n \geq n_k$, $d_k(I_n) = d_k(I_{n_k})$. On pose alors :

$$N = \sup\,(m;\ n_0,\ n_1, \ldots, n_l).$$

Montrons alors qu'on a, pour $n \geq N$, $I_n = I_N$ ce qui établira le théorème. En effet comme $I_N \subset I_n$, il suffit, d'après le point 3) ci-dessus, de prouver les relations :

$$\forall k \in \mathbf{N},\ d_k(I_n) = d_k(I_N).$$

1) Si $k \geq l$ on a : $d_k(I_N) \supset d_k(I_m) \supset d_l(I_m)$ et $d_k(I_n) \supset d_k(I_m) \supset d_l(I_m)$ et comme $d_l(I_m)$ est maximal, on a bien $d_k(I_N) = d_k(I_n)$.
2) Si $k < l$ on a : $d_k(I_N) = d_k(I_{n_k}) = d_k(I_n)$ et le théorème est démontré.

Corollaire 2.4.
Si A est noethérien, $A[X_1, \ldots, X_n]$ l'est aussi.

Corollaire 2.5.
Si A est noethérien et si B est une A-algèbre de type fini (cf. 0.d), B est un anneau noethérien.

3. Propriétés arithmétiques.

On entend par propriétés arithmétiques des anneaux celles relatives à relation de la ***divisibilité***. On suppose $A \neq \{0\}$ dans toute la suite.

a) Eléments inversibles.

On pose :

$$A^* = \{a \in A \mid \exists b \in A,\ ab = 1\}.$$

Les éléments de A^* sont les éléments inversibles de A appelés parfois « unités de A ». On vérifie que A^* est un groupe pour la multiplication.

Exemples 3.1.
1) Si A est un corps, on a $A^* = A - \{0\}$.
2) $(\mathbf{Z}/n\mathbf{Z})^*$ a été déterminé au Chapitre I § 7.
3) Si k est un corps on a $k[X]^* = k^*$.

Remarque 3.2. Pour $a \in A$, on a : $a \in A^* \iff (a) = A$.

b) Divisibilité.

Définition 3.3.
Soient $a, b \in A$. On dit que a divise b et on écrit $a \mid b$ si et seulement si il existe $c \in A$ avec $b = ac$.

Proposition 3.4.
On a : $b \mid a \iff (a) \subset (b)$.

En particulier, pour tout $a \in A$ et tout $\alpha \in A^*$ on a : $\{0\} \subset (a) \subset (\alpha) = A$ de sorte que tout élément $a \in A$ divise 0 et qu'un inversible divise tous les éléments de A.

Remarque 3.5.
La relation $b \mid a$ est un préordre (i.e. elle est réflexive et transitive, mais pas antisymétrique en général). On lui associe la relation d'équivalence :

$$a\mathcal{R}b \iff a \mid b \text{ et } b \mid a \iff (a) = (b).$$

Proposition 3.6.
On suppose A intègre. Alors on a $a\mathcal{R}b \iff \exists u \in A^*,\ a = bu$.

Démonstration. En effet, on a $b = ac$, $a = bd$, d'où $b = bcd$, donc $b(1 - cd) = 0$ et comme A est intègre, si $b \neq 0$, on a $1 - cd = 0$ i.e. $c, d \in A^*$. Le cas $b = 0$ est trivial.

Remarque 3.7. Ce résultat est en défaut si A n'est pas intègre : on prend $A = k[X, Y, Z]/X(1 - YZ)$; si x, y, z sont les images de X, Y, Z dans A, on a $x = xyz$ donc $x \mid xy$ et $xy \mid x$ mais le lecteur vérifiera qu'il n'existe pas de $u \in A^*$ tel que $xy = ux$.

Nous supposerons désormais A **intègre**.

Définition 3.8.
*Soient $a, b \in A$, on dit que a et b sont **associés** si et seulement si on a $a\mathcal{R}b$.*

Les éléments associés jouent des rôles identiques pour la divisibilité ce qu'on peut encore exprimer en disant que les éléments inversibles sont négligeables dans les questions d'arithmétique. La proposition suivante établit un dictionnaire entre la divisibilité et l'inclusion des idéaux principaux :

Proposition 3.9.
L'application $a \mapsto (a)$ induit un isomorphisme d'ensembles ordonnés de $A/\mathcal{R}$ muni de la relation de divisibilité, sur $\mathcal{J}(A)$, ensemble des idéaux principaux de A, muni de l'inclusion inverse.

Définition 3.10.
*Soit $p \in A$. On dit que p est **irréductible** si et seulement si on a :*
1) $p \notin A^*$,
2) $p = ab \implies a \in A^*$ ou $b \in A^*$.

La propriété 2) peut encore s'exprimer en disant que les seuls diviseurs de p sont les éléments inversibles et les associés de p.

Remarques 3.11.
1) 0 n'est pas irréductible.
2) Si A n'est pas un corps : p irréductible $\iff (p)$ est maximal dans $\mathcal{J}(A) - \{A\}$.

Exemple 3.12. Dans **Z** les irréductibles sont les nombres premiers.

Définition 3.13.
Soient $a, b \in A$; on dit que a et b sont **premiers entre eux** *(ou encore étrangers) si on a :*

$$\forall d \in A, \quad d \mid a \text{ et } d \mid b \implies d \in A^*$$

Autrement dit, a et b n'ont pas de diviseurs communs non triviaux.

c) Anneaux factoriels.

La notion de factorialité généralise la propriété de décomposition unique en facteurs premiers de **Z**. Attention, les anneaux factoriels ne vérifient pas toutes les propriétés de **Z** (par exemple le théorème de Bézout y est, en général, faux).

Définition 3.14.
Soit A un anneau. On dit que A est **factoriel** *s'il vérifie les trois propriétés suivantes :*
(0) A est intègre,
(E) tout élément a non nul de A s'écrit $a = up_1 \dots p_r$ avec $u \in A^*$ et $p_1, \dots, p_r$ irréductibles,
(U) cette décomposition est unique, à permutation près et à des inversibles près : si $a = up_1 \dots p_r = vq_1 \dots q_s$, on a $r = s$ et il existe $\sigma \in \mathfrak{S}_r$ tel que p_i et $q_{\sigma(i)}$ soient associés.

On peut reformuler cette définition en introduisant un système de représentants P des irréductibles de A, i.e. un ensemble P d'irréductibles qui est tel que pour tout p irréductible il existe un unique $q \in P$ vérifiant $p \mathcal{R} q$:

Définition 3.14 bis.
L'anneau A est factoriel si et seulement si :
(0) A est intègre,
(E) $\forall a \in A$, $a \neq 0$, a s'écrit sous la forme $a = u \prod_{p \in P} p^{v_p(a)}$, avec $u \in A^*$, $v_p(a) \in \mathbf{N}$ et les $v_p(a)$ nuls, sauf un nombre fini,
(U) cette écriture est unique.
L'entier $v_p(a)$ s'appelle la valuation p-adique de a.

Exemples 3.15. On peut prendre comme système de représentants :
dans **Z** les nombres premiers > 0,
dans $k[X]$ (k désignant un corps) les polynômes unitaires irréductibles.

Remarque 3.16. Avec ces notations, pour $a, b \neq 0$ on a :

$$a \mid b \implies \forall p \in P, \quad v_p(a) \leq v_p(b).$$

Nous allons maintenant analyser un peu les deux conditions (E) et (U).

Si nous partons d'un élément $a \in A$, non irréductible, il s'écrit $a = bc$ où b et c ne sont pas associés à a. Si b ou c n'est pas irréductible, on peut réitérer l'opération. La condition (E) affirme essentiellement que cette dichotomie s'arrête. Par exemple, dans **Z**, c'est l'ordre de grandeur des nombres a, b, c avec $(|b|, |c| < |a|)$ qui force l'arrêt du processus. On peut donc s'attendre à ce que cette propriété de finitude soit valable pour un anneau noethérien. De fait :

Proposition 3.17.
Si A est noethérien et intègre, A vérifie (E).

Démonstration. On considère l'ensemble :

$$F = \{(a) \in \mathcal{J}(A) \mid a \neq 0 \text{ et } a \text{ n'est pas de la forme } a = up_1 \dots p_r\}.$$

En particulier, si $(a) \in F$, a n'est ni inversible, ni irréductible.
Supposons F non vide, et soit (a) un élément maximal de F (il en existe car A est noethérien). Comme a n'est pas irréductible on a $a = bc$, avec $b, c \notin A^*$. On en déduit $(a) \underset{\neq}{\subset} (b)$, $(a) \underset{\neq}{\subset} (c)$, sinon, si on a par exemple $(a) = (b)$, on a $a = bu$ avec $u \in A^*$ (cf. 3.5 et 3.6), d'où $a = bu = bc$ et, comme A est intègre, $c = u \in A^*$, contradiction. Mais alors, comme (a) est maximal dans F, (b) et (c) ne sont pas dans F et donc b et c s'écrivent sous la forme de (E) : $b = up_1 \dots p_r$; $c = vq_1 \dots q_s$, et on a aussi $a = bc = uvp_1 \dots p_r q_1 \dots q_s$ ce qui contredit le fait que a est dans F.

Remarques 3.18.
1) Attention, noethérien et intègre n'impliquent pas factoriel. Considérons en effet l'anneau $A = \mathbf{Z}[i\sqrt{5}] = \{a + bi\sqrt{5} \in \mathbf{C} \mid a, b \in \mathbf{Z}\}$. Cet anneau est noethérien (il est isomorphe à $\mathbf{Z}[T]/(T^2 + 5)$) et intègre, mais on a, dans A :

$$9 = 3 \times 3 = (2 + i\sqrt{5})(2 - i\sqrt{5})$$

et on vérifie aisément que 3, $2 + i\sqrt{5}$ et $2 - i\sqrt{5}$ sont irréductibles ce qui contredit la propriété (U). (On considérera pour cela la norme d'un élément $z = a + ib\sqrt{5}$, i.e. $N(z) = z\overline{z} = a^2 + 5b^2$.)
2) Remarquons aussi que factoriel n'implique pas noethérien. Ainsi, par exemple, l'anneau de polynômes à une infinité de variables $k[X_1, \dots, X_n, \dots]$ est factoriel et non noethérien, cf. § 2 et § 4.

On s'intéresse maintenant à la condition (U).

Proposition 3.19.
Soit A un anneau intègre vérifiant (E). Les conditions suivantes sont équivalentes :
1) A vérifie (U),
*2) **le lemme d'Euclide** : si p est irréductible et p divise ab, alors p divise a ou b,*
3) p irréductible $\Longleftrightarrow$ (p) premier, (on dit alors parfois que p est premier),
*4) **le théorème de Gauss** : si a divise bc et si a est premier avec b, a divise c.*

Démonstration.
a) Notons déjà que, sans aucune hypothèse, (p) premier $\Longrightarrow$ p irréductible. En effet, si $p = ab$, on a $ab \in (p)$, donc a ou $b \in (p)$ donc p divise a ou b.
b) L'équivalence de 2) et 3) est claire, sans l'hypothèse que A vérifie (E).
c) On a aussi, toujours sans (E), 4) $\Longrightarrow$ 2) (c'est clair) et 2) $\Longrightarrow$ 1) : en effet, si $a = u \prod p^{v_p(a)} = v \prod p^{w_p(a)}$, on choisit p tel que $v_p(a) > 0$, comme p divise le deuxième membre, il divise un de ses facteurs, et on a donc $w_p(a) > 0$. On divise alors les deux membres par p et on termine par récurrence.
d) En revanche, pour prouver 1) $\Longrightarrow$ 4) (ou aussi 1) $\Longrightarrow$ 2)) on a besoin de (E). On décompose a, b, c en facteurs premiers et il s'agit de montrer : $v_p(a) \leq v_p(c)$ pour tout p. Sinon, on a, pour un p, $v_p(a) > v_p(c)$, or, comme a divise bc, on a $v_p(b) \geq v_p(a) - v_p(c)$ donc $v_p(b)$ et $v_p(a)$ sont > 0 et p divise a et b, c'est une contradiction.

Cette proposition nous fournit un exemple important d'anneau factoriel :

Définition 3.20.
Un anneau A est dit ***principal*** *s'il est intègre et si tout idéal de A est principal.*

Corollaire 3.21.
Un anneau principal est factoriel.

Démonstration. On note d'abord qu'on a les implications :
A principal $\Longrightarrow$ A noethérien $\Longrightarrow$ A vérifie (E) (3.17).
Ensuite, si p est irréductible dans A, l'idéal (p) est maximal parmi les idéaux principaux de A distincts de A (3.11.2), donc maximal tout court, donc premier (cf. 1.8) et on conclut grâce à la Proposition 3.19.

Remarque 3.22.
On a déterminé tous les idéaux premiers d'un anneau A principal. Si A n'est pas un corps, les idéaux premiers de A sont l'idéal (0) et les idéaux maximaux (p) engendrés par les irréductibles.

d) ppcm et pgcd.

Là encore, la situation connue sur $\mathbf{Z}$ se généralise.

Proposition-définition 3.23.
Si A est un anneau factoriel, l'ensemble ordonné $A/\mathcal{R}$ (ou encore $\mathcal{J}(A)$, cf. Prop. 3.9) est réticulé, c'est-a-dire que deux éléments ont un sup et un inf.
On pose alors, si $\inf((a);(b)) = (c)$, $c = \operatorname{ppcm}(a,b)$ *(plus petit commun multiple), et si* $\sup((a);(b)) = (d)$, $d = \operatorname{pgcd}(a,b)$, *(plus grand commun diviseur).*

Démonstration.
Si $a = u\prod p^{v_p(a)}$ et $b = v\prod p^{v_p(b)}$, le ppcm est donné par $c = \prod p^{\sup(v_p(a),v_p(b))}$ et le pgcd par $d = \prod p^{\inf(v_p(a),v_p(b))}$.

Remarques 3.24.
0) Attention, le ppcm et le pgcd ne sont définis qu'aux éléments inversibles près.
1) On a $(a)\cap(b) = (c)$ où c est le ppcm de a et b, mais pas en général $(a)+(b) = (d)$ avec pour d le pgcd (c'est le théorème de Bézout, cf. plus loin).
2) On définit aisément par récurrence le pgcd et le ppcm de n éléments.
3) Pour des réciproques et contre-exemples, voir les exercices.

e) Le théorème de Bézout.

Proposition 3.25.
Soit A un anneau principal, soient $a, b \in A - \{0\}$ et soit $d = \operatorname{pgcd}(a,b)$, alors on a $(d) = (a) + (b)$ ce qui signifie encore qu'il existe $\lambda, \mu \in A$ tels que $d = \lambda a + \mu b$.

Démonstration. Par définition du pgcd, l'idéal (d) est le sup des idéaux (a) et (b) dans l'ensemble des idéaux principaux de A, donc dans l'ensemble de tous les idéaux, et c'est donc $(a)+(b)$.

Corollaire 3.26.
Soit A un anneau principal, soient $a, b \in A - \{0\}$, premiers entre eux. Alors, on a $(a)+(b) = (1)$ i.e. il existe $\lambda, \mu \in A$ tels que $1 = \lambda a + \mu b$.

Remarque 3.27. Le théorème de Bézout est en défaut dans un anneau factoriel non principal. Ainsi, l'anneau $k[X,Y]$ est factoriel (cf. § 4) et X et Y sont premiers entre eux, mais on a $(X)+(Y) = (X,Y) \neq (1)$.

f) Anneaux euclidiens.

Nous allons étudier maintenant une classe importante d'anneaux principaux.

Définition 3.28.
Un anneau A est dit ***euclidien*** *si :*
1) A est intègre,
2) A est muni d'une division euclidienne, i.e. il existe une fonction (appelée parfois stathme) $v : A - \{0\} \longrightarrow \mathbf{N}$ telle que si $a, b \in A - \{0\}$, il existe $q, r \in A$ avec $a = bq + r$ et ($r = 0$ ou $v(r) < v(b)$).

Théorème 3.29.
Un anneau euclidien est principal.

Démonstration. Soit I un idéal de A, $I \neq \{0\}$. Soit $b \in I$, $b \neq 0$, tel que $v(b)$ soit minimal. Soit $a \in I$, on effectue la division euclidienne de a par b :

$$a = bq + r \text{ avec } r = 0 \text{ ou } v(r) < v(b),$$

on a alors $r = a - bq \in I$ et comme $v(b)$ est minimal pour les éléments non nuls de I, on a $r = 0$, donc $a \in (b)$ et on a bien montré $I = (b)$.

Exemple 3.30. L'anneau $\mathbf{Z}$ muni de l'application $v(n) = |n|$ est euclidien.

Un autre exemple important est celui de $k[X]$, où k est un corps. On commence par prouver un lemme un peu plus général :

Lemme 3.31 : division euclidienne dans $A[X]$.
Soit A un anneau et soit $P \in A[X]$, $P \neq 0$, de coefficient dominant ***inversible.*** *Soit $F \in A[X]$, il existe $Q, R \in A[X]$, tels que l'on ait :*
1) $F = PQ + R$,
2) $d^\circ R < d^\circ P$ ou $R = 0$.

Démonstration.
On peut supposer P unitaire : $P(X) = X^n + a_{n-1}X^{n-1} + \cdots + a_0$. On considère l'anneau quotient $B = A[X]/(P)$. Soit x l'image de X dans B. Il suffit de prouver que tout élément de B est combinaison linéaire à coefficients dans A de $1, x, \ldots, x^{n-1}$. Par linéarité il suffit même de le faire pour les monômes x^i et c'est immédiat par récurrence en tenant compte de la relation :

$$x^n = -a_{n-1}x^{n-1} - \cdots - a_0.$$

Corollaire 3.32.
Si k est un corps, l'anneau $k[X]$ est euclidien (avec comme stathme $v(P) = d^\circ P$).

Exemple 3.33. Autres exemples d'anneaux euclidiens.
Nous verrons au § 6 l'anneau $\mathbf{Z}[i]$. Signalons également l'anneau $k[[X]]$ des séries formelles à coefficients dans un corps et l'anneau D des nombres décimaux (sous-anneau de $\mathbf{Q}$ engendré par $\mathbf{Z}$ et $1/10$), cf. § 3 Exercice 7. Nous donnons au § 5 un exemple d'anneau principal non euclidien.

4. Stabilité des notions étudiées ; théorème de Gauss.

a) Passage à l'anneau des polynômes.

Si A est principal, $A[X]$ ne l'est plus en général :

Proposition 4.1.
Soit A un anneau. $A[X]$ est principal si et seulement si A est un corps.

Démonstration. Si A est un corps, $A[X]$ est euclidien, donc principal. Réciproquement, si $A[X]$ est principal, il est intègre, donc A aussi. Il en résulte que X est irréductible dans A pour une raison de degré. Donc (X) est premier (car A est factoriel), donc maximal (car A est principal) donc $A \simeq A[X]/(X)$ est un corps. (On peut aussi considérer les idéaux (X, a) pour $a \in A$).

En revanche, la factorialité se conserve :

Théorème 4.2 (Gauss).
Si A est factoriel, $A[X]$ est factoriel.

Démonstration. Notons d'abord que $A[X]$ est intègre et qu'on a $A[X]^* = A^*$. Le ressort de la démonstration est l'utilisation du corps des fractions K de A et de l'anneau $K[X]$ qui, lui, est principal donc factoriel.
On définit d'abord, pour $P \in A[X]$, $P \neq 0$, le ***contenu*** de P, noté $c(P)$:
si $P(X) = a_n X^n + \cdots + a_0$, on pose $c(P) = \operatorname{pgcd}(a_0, \ldots, a_n)$, l'élément $c(P)$ est bien entendu défini modulo A^*.
Un polynôme P est dit ***primitif*** si on a $c(P) = 1$.

Lemme 4.3 (Gauss).
On a $c(PQ) = c(P)c(Q)$ modulo A^*.

Démonstration (de 4.3)
1) Supposons d'abord P, Q primitifs : $c(P) = c(Q) = 1$. Si $c(PQ) \neq 1$, il existe $p \in A$, irréductible qui divise $c(PQ)$ (car A est factoriel). Mais comme on a $c(P) = c(Q) = 1$, il existe $i_0, j_0 \in N$ tels que :

$$\forall i < i_0, \quad p \mid a_i \quad \text{mais} \quad p \nmid a_{i_0}$$
$$\forall j < j_0, \quad p \mid b_j \quad \text{mais} \quad p \nmid b_{j_0}.$$

Par hypothèse on a

$$p \mid c_{i_0+j_0} = \sum_{i+j=i_0+j_0} a_i b_i = a_{i_0} b_{j_0} + \sum_{\substack{i+j=i_0+j_0 \\ i<i_0 \text{ ou } j<j_0}} a_i b_j$$

mais alors $p \mid a_{i_0} b_{j_0}$ ce qui contredit le lemme d'Euclide (3.19).

2) Passons au cas général :
Posons $d = c(P)$, $e = c(Q)$, $P' = P/d$, $Q' = Q/e$, on a $c(P') = c(Q') = 1$ et $PQ = deP'Q'$ avec $c(P'Q') = 1$ d'après le premier cas. Il en résulte aussitôt que $c(PQ) = de$.

On peut alors déterminer les irréductibles de $A[X]$:

Proposition 4.4.
Les polynômes $P(X) \in A[X]$ irréductibles dans $A[X]$ sont :
1) les constantes $p \in A$, irréductibles dans A,
2) les polynômes $P(X)$, de degré ≥ 1, primitifs et irréductibles dans $K[X]$.

Démonstration (de 4.4)
a) Ces éléments sont des irréductibles :
1) Si $p \in A$ et si $p = P(X)Q(X)$, on a $d^\circ P = d^\circ Q = 0$, donc $P, Q \in A$ et l'un d'eux est dans A^*, donc est inversible dans $A[X]$ et p est irréductible dans $A[X]$.

2) Si $P(X) = Q(X)R(X)$ dans $A[X]$, donc aussi dans $K[X]$, comme P est irréductible dans $K[X]$, on a, par exemple, $Q \in K[X]^*$, donc $d^\circ Q = 0$ et $Q \neq 0 : Q = a \in A$. On en déduit $P(X) = aR(X)$, donc $a \mid c(P)$ et comme $c(P) = 1$, $a \in A^*$ et P est irréductible.

b) Ce sont les seuls irréductibles. En effet, soit P un irréductible de $A[X]$.

1) Si $d^\circ P = 0$, $P = p \in A$ est nécessairement irréductible dans A.

2) Si $d^\circ P > 0$, il est clair que l'on a $c(P) = 1$ et il reste à voir que P est irréductible dans $K[X]$. Sinon on aurait $P(X) = Q(X)R(X)$ avec $Q, R \in K[X]$. On peut alors écrire $Q(X)$ sous la forme $Q(X) = \frac{a}{b}Q'(X)$ avec $Q' \in A[X]$, $c(Q') = 1$ et $a, b \in A$, premiers entre eux : si $Q(X) = \sum_{i=1}^{n} \frac{a_i}{b_i} x^i$ on prend pour b un ppcm des b_i, d'où $Q(X) = \sum_{i=1}^{n} \frac{a'_i}{b} x^i$ puis pour a un pgcd des a'_i et enfin on simplifie éventuellement la fraction a/b.

De même on écrit $R(X) = \frac{c}{d}R'(X)$. On a $bd\ P(X) = ac\ Q'(X)R'(X)$, d'où en calculant les contenus par le lemme de Gauss : $c(bd\ P) = bd = c(ac\ Q'R') = ac$, modulo A^*.

On a donc $P = \lambda Q'R'$, $\lambda \in A^*$, mais comme P est irréductible dans $A[X]$, Q' ou R' est dans $A[X]^* = A^*$, donc de degré 0, et P est donc irréductible dans $K[X]$.

Revenons à 4.2 en prouvant les propriétés (E) et (U) :

Pour (E), on considère d'abord $P \in A[X]$, primitif, que l'on écrit : $P(X) = P_1(X)^{\alpha_1} \dots P_r(X)^{\alpha_r}$, avec $P_i \in K[X]$, irréductible dans $K[X]$. On écrit ensuite, comme ci-dessus :

$$P_i(X) = \frac{a_i}{b_i} P'_i(X) \text{ avec } c(P'_i) = 1.$$

Comme $\frac{a_i}{b_i} \in K^*$, P'_i est un polynôme irréductible dans $A[X]$ et on a :

$$(\prod b_i)P = \prod a_i \prod {P'_i}^{\alpha_i}$$

mais en passant aux contenus, on voit qu'on a :

$$P = u \prod {P'_i}^{\alpha_i} \text{ avec } u \in A^*, \text{ d'où (E).}$$

Si P n'est pas primitif, on a $P = dP'$ avec P' primitif et on conclut en décomposant d dans A et P' dans $A[X]$.

Pour (U), on prouve que si P est irréductible, l'idéal $(P) = PA[X]$ est premier.

1) Si $P = p \in A$ alors, on a :

$$A[X]/(p) \simeq A/(p)[X]$$

qui est intègre puisque $A/(p)$ l'est.

2) Si $d^\circ P \geq 1$, on considère le diagramme ci-dessous :

$$\begin{array}{ccc} A[X] & \hookrightarrow & K[X] \\ \downarrow & & | \\ A[X]/(P) & \xhookrightarrow{i} & K[X]/(P) \end{array}$$

comme $K[X]/(P)$ est intègre, puisque P est irréductible et $K[X]$ principal, il suffit de prouver que i est injectif ou encore que l'on a :

$$(PK[X]) \cap A[X] = PA[X].$$

Seule l'inclusion $\subset$ est non triviale. Soit $Q \in A[X]$, $Q = PR$, $R \in K[X]$, on écrit, comme d'habitude, $R = \frac{a}{b}R'$ et $Q = c\,Q'$, avec $R', Q' \in A[X]$ et primitifs. D'où : $cb\,Q' = a\,PR'$. Passant aux contenus, on voit que b divise a, donc que R est dans $A[X]$ et ceci achève la démonstration du théorème de Gauss.

Corollaire 4.5.
Si A est factoriel, $A[X_1, \ldots, X_n]$ est factoriel.

Démonstration. C'est clair par récurrence sur n.
On notera que les anneaux $\mathbf{Z}[X]$, $k[X, Y]$ pour k un corps, sont factoriels et non principaux.

Corollaire 4.6.
Si A est factoriel, $A[X_1, \ldots, X_n, \ldots]$ l'est aussi.

Démonstration. Cela résulte de 4.5 car tout se passe avec un nombre fini d'indéterminées.

b) Passage au quotient.

Bien entendu, on se restreint au cas des idéaux premiers afin que le quotient soit intègre.

Remarque 4.7. Pour un anneau A principal la situation est très simple : les quotients intègres A/I de A sont A, si $I = \{0\}$, ou des corps, si $I = (p)$, avec p irréductible.
Si $I = (a)$ avec a quelconque, que l'on écrit comme produit d'irréductibles distincts : $a = p_1^{\alpha_1} \ldots p_r^{\alpha_r}$, on peut calculer A/I :

$$A/(a) \simeq \prod_{i=1}^{r} A/(p_i^{\alpha_i}), \text{ c'est le lemme chinois.}$$

Pour un anneau factoriel, la situation est plus compliquée, mais en général les quotients intègres ne sont pas factoriels, par exemple
1) $\mathbf{Z}[i\sqrt{5}] \simeq \mathbf{Z}[X]/(X^2 + 5)$ n'est pas factoriel (cf. 3.18),
2) $\mathbf{C}[X, Y]/(Y^2 - X^3)$ n'est pas factoriel, (montrer que l'image $\overline{Y}$ de Y est un élément irréductible, mais que l'idéal $(\overline{Y})$ n'est pas premier).

c) Sous-anneaux.

Comme dans le cas noethérien, l'exemple d'un anneau intègre plongé dans son corps de fractions montre qu'il ne faut espérer aucune conservation par passage aux sous-anneaux.

5. Un exemple d'anneau principal non euclidien.

Nous allons prouver que l'anneau $\mathbf{Z}\left[\frac{1+i\sqrt{19}}{2}\right]$ est principal, mais n'est pas euclidien. Un autre exemple, $\mathbf{R}[X, Y]/(X^2+Y^2+1)$, est étudié dans les exercices.

a) Comment reconnaître qu'un anneau n'est pas euclidien.

Proposition 5.1.
Soit A un anneau euclidien. Il existe $x \in A$, $x \notin A^*$, tel que la restriction à $A^* \cup \{0\}$ de la projection canonique de A sur $A/(x)$ soit surjective.

Démonstration.
Si A est un corps, $x = 0$ convient. Sinon, parmi les éléments de A non nuls et non inversibles on choisit x tel que $v(x)$ soit minimal.
Alors, si $a \in A$, on a $a = xq + r$ avec $r = 0$ ou $v(r) < v(x)$, donc $a \equiv r \pmod{x}$.
Mais, si $r \neq 0$, comme $v(r) < v(x)$, r est inversible et a est bien égal, modulo (x), à 0 ou à un élément de A^*.

Remarques 5.2.
1) Comme l'image d'un inversible est inversible, $A/(x)$ est un corps, donc (x) est maximal.
2) Voici quelques exemples :
pour $A = \mathbf{Z}$, $A^* = \{1, -1\}$, on peut prendre $x = 2$ ou 3,
pour $A = k[X]$, $A^* = k^*$, on prend $x = X - a$, pour $a \in k$,
pour $A = \mathbf{Z}[i]$, $A^* = \{\mp 1, \mp i\}$ (cf. §6), on prend $x = 1 - i$ car $A/(1-i) \simeq \mathbf{Z}/2\mathbf{Z}$.

b) Application, l'anneau $A = \mathbf{Z}\left[\dfrac{1+i\sqrt{19}}{2}\right] = \mathbf{Z}[\alpha]$ *n'est pas euclidien.*

On pose $\alpha = \dfrac{1+i\sqrt{19}}{2}$, d'où $\overline{\alpha} = \dfrac{1-i\sqrt{19}}{2}$ et α vérifie l'équation $\alpha^2 - \alpha + 5 = 0$.
En effet, on a $\alpha + \overline{\alpha} = 1$, $\alpha\overline{\alpha} = 5$. On pose alors

$$A = \mathbf{Z}[\alpha] = \{z \in \mathbf{C} \mid z = a + b\alpha, \quad a, b \in \mathbf{Z}\}.$$

L'anneau $\mathbf{Z}[\alpha]$ est un sous-anneau de $\mathbf{C}$, donc est intègre. De plus A est stable par conjugaison, puisque $\overline{\alpha} = 1 - \alpha$. On définit alors, pour $z \in A$, $N(z) = z\overline{z} = a^2 + ab + 5b^2$ (la "norme" de z au sens des arithméticiens) et on a $N(z) \in \mathbf{N}$, $N(zz') = N(z)N(z')$ et $N(z) > 0$ pour $z \neq 0$.

La norme nous permet de calculer le groupe A^*. En effet, si $z \in A^*$, on a $N(zz^{-1}) = N(z)N(z^{-1}) = 1$, avec $N(z), N(z^{-1}) \in \mathbf{N}$, donc $N(z) = 1$. On a donc, si $z = a + b\alpha$, la relation $a^2 + ab + 5b^2 = 1$, avec $a, b \in \mathbf{Z}$. Mais, on a $b^2 + a^2 + ab \geq b^2 + a^2 - |ab| \geq (|b| - |a|)^2 \geq 0$, et donc, $1 = a^2 + ab + 5b^2 \geq 4b^2$. On en déduit $b = 0$ et $a = \pm 1$, donc $A^* = \{1, -1\}$.

Il en résulte que A n'est pas euclidien, sinon, en vertu de 5.1, il existerait $x \in A$ tel que $A/(x)$ soit un corps à 2 ou 3 éléments. On aurait donc un homomorphisme surjectif $\varphi : \mathbf{Z}[\alpha] \longrightarrow K$, avec $K = \mathbf{F}_2$ ou $\mathbf{F}_3$ (cf. Chapitre III §2). La restriction de φ à $\mathbf{Z}$ étant la projection canonique de $\mathbf{Z}$ sur $\mathbf{Z}/2\mathbf{Z}$ ou $\mathbf{Z}/3\mathbf{Z}$, on aurait donc un élément $\beta = \varphi(\alpha) \in K$ vérifiant $\beta^2 - \beta + 5 = 0$ dans K. Pour $K = \mathbf{F}_2$, on aurait $\beta^2 + \beta + 1 = 0$, qui n'a pas de solutions dans $\mathbf{F}_2$, pour $K = \mathbf{F}_3$, on aurait $\beta^2 - \beta - 1 = 0$ qui n'a pas de solutions dans $\mathbf{F}_3$, d'où une contradiction.

c) L'anneau $A = \mathbf{Z}[\alpha]$ est principal.

Proposition 5.3 (pseudo division euclidienne).
Soient $a, b \in A - \{0\}$, il existe $q, r \in A$ avec :
1) $r = 0$ ou $N(r) < N(b)$,
2) $a = bq + r$ ou $2a = bq + r$.

Démonstration.
Soit $x = \dfrac{a}{b} = \dfrac{a\bar{b}}{b\bar{b}} \in \mathbf{C}$, que l'on écrit $x = u + v\alpha$, avec $u, v \in \mathbf{Q}$. Soit $n = [v]$ la partie entière de v. On a $v \in [n,\ n+1[$.

1) Supposons $v \notin]n + \frac{1}{3},\ n + \frac{2}{3}[$ et soient alors s et t les entiers les plus proches de u et v respectivement. On a $|s-u| \leq \frac{1}{2}$, $|t-v| \leq \frac{1}{3}$. On pose alors $q = s + t\alpha$, de sorte que q est dans A et on a :

$$N(x-q) = (s-u)^2 + (s-u)(t-v) + 5(t-v)^2 \leq \frac{1}{4} + \frac{1}{6} + \frac{5}{9} = \frac{35}{36} < 1.$$

Si on pose $r = a - bq = b(x-q)$, on a bien $N(r) < N(b)$ et le résultat voulu.

2) Supposons $v \in]n + \frac{1}{3},\ n + \frac{2}{3}[$. On considère alors $2x = 2u + 2v\alpha$, et on a : $2v \in]2n + \frac{2}{3},\ 2n + 1 + \frac{1}{3}[$ et donc, si $m = [2v]$, on a $2v \notin]m + \frac{1}{3},\ m + \frac{2}{3}[$, on est ramené au cas précédent, et on a $2a = bq + r$ avec $N(r) < N(b)$, ce qui achève de prouver 5.3.

Montrons maintenant que A est principal.

1) L'idéal (2) est maximal dans A.
Il suffit de remarquer que l'on a $A \simeq \mathbf{Z}[T]/(T^2 - T + 5)$ (comme on le voit par division euclidienne, cf. 3.31) et donc, en vertu du théorème d'isomorphisme (cf. 0.a) on a

$$A/(2) \simeq \mathbf{Z}[T]/(2, T^2 - T + 5) \simeq (\mathbf{Z}/2\mathbf{Z})[T]/(T^2 + T + 1).$$

Or, le polynôme $T^2 + T + 1$ est irréductible sur $\mathbf{F}_2 = \mathbf{Z}/2\mathbf{Z}$, d'où le résultat par 3.32 et 3.22.

2) Soit I un idéal $\neq \{0\}$ de A et soit $a \in I$, $a \neq 0$, tel que $N(a)$ soit minimal. Si $I = (a)$, on a terminé, sinon, soit $x \in I - (a)$, et effectuons la pseudo-division euclidienne de x par a :

i) Si $x = aq + r$ avec $N(r) < N(a)$ ou $r = 0$, comme $r \in I$, on a $r = 0$, donc $x \in (a)$ et c'est une contradiction.

ii) On a donc $2x = aq + r$ avec $N(r) < N(a)$ ou $r = 0$ et, pour la même raison que ci-dessus, on a $r = 0$, donc $2x = aq$.
Comme (2) est maximal, donc premier, on a a ou $q \in (2)$. Si $q \in (2)$, $q = 2q'$ et $x \in (a)$, contradiction. Donc on a $q \notin (2)$ et $a \in (2)$, $a = 2a'$, d'où $x = a'q \in (a')$. Il suffit alors de montrer que a' est dans I car cela contredira la minimalité de $N(a)$ (puisque $a = 2a'$). Comme l'idéal (2) est maximal et ne contient pas q, l'idéal $(2, q)$ est égal à A tout entier. On a donc une relation de Bézout : $\lambda 2 + \mu q = 1$, avec $\lambda, \mu \in A$. On en déduit $a' = \lambda 2a' + \mu q a' = \lambda a + \mu x$, donc $a' \in I$ puisque a et x sont dans I.

Remarque 5.4. On montre que l'anneau des entiers du corps quadratique $\mathbf{Q}(i\sqrt{d})$, pour $d \in \mathbf{N}^*$ et d sans facteur carré, (cf. III, § 2, 3, Exercice 12) est
- euclidien pour $d = 1, 2, 3, 7, 11$,
- principal (non euclidien) pour $d = 19, 43, 67, 163$.

6. L'anneau $\mathbf{Z}[i]$ et le théorème des deux carrés.

a) Introduction.

Le problème est de déterminer quels entiers $n \in \mathbf{N}$ sont sommes de deux carrés : $n = a^2 + b^2$ avec $a, b \in \mathbf{N}$. On pose :

$$\Sigma = \{n \in \mathbf{N} \mid n = a^2 + b^2;\ \ a, b \in \mathbf{N}\}.$$

Exemples 6.1. On a 0, 1, 2, 4, 5, 8, 9, 10, ... $\in \Sigma$, mais 3, 6, 7, 11, 12, ... $\notin \Sigma$.

Remarque 6.2. Si $n \equiv 3$ (mod. 4) on a $n \notin \Sigma$. En effet, si a est pair, on a $a^2 \equiv 0$ (mod. 4), si a est impair, $a^2 \equiv 1$ (mod. 4), donc $a^2 + b^2 \equiv 0, 1, 2$ (mod. 4).

L'idée que nous allons utiliser pour étudier Σ et qui est sans doute due à Gauss est de noter que, si $n \in \Sigma$, $n = a^2 + b^2$, n s'écrit dans $\mathbf{C}$, $n = (a + ib)(a - ib)$, et que cette relation a lieu, en fait, dans l'anneau $\mathbf{Z}[i]$ des entiers de Gauss :

$$\mathbf{Z}[i] = \{a + ib \in \mathbf{C} \mid a, b \in \mathbf{Z}\}.$$

En particulier si p est un nombre premier de $\mathbf{N}$ qui est somme de deux carrés, il n'est plus irréductible dans $\mathbf{Z}[i]$, par exemple on a : $5 = (2 + i)(2 - i)$.

b) Etude de l'anneau $\mathbf{Z}[i]$.

Tout d'abord $\mathbf{Z}[i]$ est un anneau intègre, puisqu'inclus dans $\mathbf{C}$. On remarque ensuite qu'on dispose d'un automorphisme de $\mathbf{Z}[i]$ donné par la conjugaison :

$$\begin{aligned} \sigma : \mathbf{Z}[i] &\longrightarrow \mathbf{Z}[i] \\ z = a + ib &\longmapsto \overline{z} = a - ib. \end{aligned}$$

Comme au paragraphe précédent on a aussi la "norme" :

$$\begin{aligned} N : \mathbf{Z}[i] &\longrightarrow \mathbf{N} \\ z = a + ib &\longmapsto z\overline{z} = a^2 + b^2 \end{aligned}$$

et N est multiplicative : $N(zz') = N(z)N(z')$.
L'introduction de la norme permet de calculer les inversibles de $\mathbf{Z}[i]$:

Proposition 6.3.
On a $\mathbf{Z}[i]^* = \{\pm 1, \pm i\}$.

Démonstration. Si $z \in \mathbf{Z}[i]^*$, on a $zz' = 1$, d'où $N(z)N(z') = 1$ et comme $N(z), N(z') \in \mathbf{N}$, ce n'est possible qu'avec $N(z) = N(z') = 1$. Si $z = a + ib$, on a donc $a^2 + b^2 = 1$, de sorte que l'un des nombres a ou b est nul et que l'autre vaut ± 1. Notons au passage qu'on a $z \in \mathbf{Z}[i]^* \iff N(z) = 1$.

Proposition 6.4.
L'ensemble Σ *des sommes de deux carrés est stable par multiplication.*

Démonstration. On traduit la propriété $n \in \Sigma$ en termes d'entiers de Gauss :

$$n \in \Sigma \iff \exists z \in \mathbf{Z}[i],\ n = N(z).$$

Alors si $n, n' \in \Sigma$, on a $n = N(z)$, $n' = N(z')$ donc $nn' = N(zz') \in \Sigma$.
Si on écrit $z = a + ib$, $z' = c + id$, on retrouve la célèbre identité dite de Lagrange (mais sans doute connue d'Euclide) :

$$(a^2 + b^2)(c^2 + d^2) = (ac - bd)^2 + (ad + bc)^2.$$

Cette seule proposition est déjà révélatrice de l'efficacité de l'introduction de $\mathbf{Z}[i]$. De plus elle ramène essentiellement l'étude de Σ à la détermination des nombres p premiers de $\mathbf{N}$ qui sont dans Σ. Pour cela on va étudier la structure arithmétique de $\mathbf{Z}[i]$.

Proposition 6.5.
L'anneau $\mathbf{Z}[i]$ est ***euclidien*** *(relativement à la fonction N), donc principal.*

Démonstration. Soient $z, t \in \mathbf{Z}[i] - \{0\}$. Pour faire la division euclidienne de z par t, on commence par considérer $z/t \in \mathbf{C}$. On approxime ensuite z/t par un entier de Gauss q : si $z/t = x + iy$, on prend $q = a + ib$ où a et b sont les entiers les plus proches de x, y. On a ainsi $|z/t - q| \leq \dfrac{\sqrt{2}}{2} < 1$ (car $|x - a|$ et $|y - b|$ sont $\leq \dfrac{1}{2}$).
On pose alors $r = z - qt$, de sorte que r est dans $\mathbf{Z}[i]$ et on a $r = t(z/t - q)$ d'où $|r| = |t|\ |z/t - q| < |t|$ et en élevant au carré, $N(r) < N(t)$. On a donc bien écrit $z = qt + r$ avec $N(r) < N(t)$ comme annoncé.

On peut alors prouver le théorème principal de ce paragraphe :

Théorème 6.6.
Soit $p \in \mathbf{N}$ un nombre premier. On a l'équivalence :

$$p \in \Sigma \Longleftrightarrow p = 2 \text{ ou } p \equiv 1 \text{ (mod. 4)}.$$

Démonstration. La condition est nécessaire puisqu'on a vu que si $p \equiv 3$ (mod. 4) on a $p \notin \Sigma$.
Pour la réciproque on établit d'abord le lemme suivant :

Lemme 6.7.
On a : $p \in \Sigma \Longleftrightarrow p$ n'est pas irréductible dans $\mathbf{Z}[i]$.

Démonstration (du lemme 6.7).
Pour le sens $\Rightarrow$: si $p = a^2 + b^2$, on a $p = (a + ib)(a - ib)$ et a, b sont $\neq 0$, donc $a + ib, a - ib$ ne sont pas dans $\mathbf{Z}[i]^*$, de sorte que p n'est pas irréductible.
Pour le sens $\Leftarrow$: si $p = zz'$ avec $z, z' \neq \pm 1, \pm i$, on a $N(p) = N(z)N(z') = p^2$ et, comme $N(z), N(z') \neq 1$, c'est qu'on a nécessairement $p = N(z)$, donc $p \in \Sigma$ et le lemme est démontré.

Comme $\mathbf{Z}[i]$ est principal, donc factoriel, dire que p est non irréductible c'est dire, exactement, que l'idéal principal $(p) = p\mathbf{Z}[i]$ est non premier, donc que le quotient $\mathbf{Z}[i]/(p)$ est non intègre.
Pour étudier ce quotient on utilise l'isomorphisme :

$$\mathbf{Z}[i] \simeq \mathbf{Z}[X]/(X^2 + 1)$$

(que l'on prouve par division euclidienne, cf. 3.31), puis les isomorphismes suivants, qui résultent tous du théorème d'isomorphisme 0.a :

$$\mathbf{Z}[i]/(p) \simeq \mathbf{Z}[X]/(X^2 + 1, p) \simeq [\mathbf{Z}[X]/(p)]/(X^2 + 1) \simeq \mathbf{Z}/p\mathbf{Z}[X]/(X^2 + 1),$$

et ce dernier anneau n'est autre que $\mathbf{F}_p[X]/(X^2+1)$, où $\mathbf{F}_p = \mathbf{Z}/p\mathbf{Z}$ désigne le corps à p éléments.
On a donc montré les équivalences :
(p) non premier $\iff X^2+1$ non irréductible dans $\mathbf{F}_p[X] \iff X^2+1$ a une racine dans $\mathbf{F}_p$, et, en résumé :

$$p \in \Sigma \iff -1 \in \mathbf{F}_p^{*2}.$$

Il reste à montrer que -1 est un carré de $\mathbf{F}_p$ si et seulement si on a $p = 2$ ou $p \equiv 1 \pmod 4$, ce que nous ferons au Chapitre III, 2.13 et qui achèvera la démonstration.

Exemples 6.8.
On vérifie que les nombres premiers $41, 53, 61, \cdots$ qui sont $\equiv 1 \pmod 4$ sont bien dans Σ : $41 = 5^2 + 4^2$; $53 = 7^2 + 2^2$; $61 = 6^2 + 5^2 \ldots$

On peut maintenant achever l'étude de Σ :

Théorème 6.9.
Soit $n \in \mathbf{N}^*$, $n \neq 1$, on décompose n en facteurs premiers :

$$n = \prod_{p \in P} p^{v_p(n)}.$$

Alors on a $n \in \Sigma \iff v_p(n)$ pair pour $p \equiv 3 \pmod 4$.

Démonstration.
L'implication $\Rightarrow$ est claire avec 6.4 et 6.6, en notant qu'un carré est toujours dans Σ.
Pour $\Leftarrow$, soit $p \equiv 3 \pmod 4$. On montre par récurrence sur $v_p(n)$ que $v_p(n)$ est pair. Si $v_p(n) = 0$, c'est clair. Sinon, p divise $n = a^2 + b^2 = (a+ib)(a-ib)$, mais comme p est irréductible dans $\mathbf{Z}[i]$ (cf. 6.7), p divise, par exemple, $a + ib$. Mais alors, comme p est entier, on a $p \mid a$ et $p \mid b$ donc $p^2 \mid n$ et si on écrit $a = pa'$, $b = pb'$, on a $\dfrac{n}{p^2} = a'^2 + b'^2 \in \Sigma$. Mais $v_p\left(\dfrac{n}{p^2}\right) = v_p(n) - 2$ est pair d'après l'hypothèse de récurrence donc aussi $v_p(n)$.

On peut aussi achèver la détermination des irréductibles de $\mathbf{Z}[i]$:

Théorème 6.10.
Les irréductibles de $\mathbf{Z}[i]$ sont, aux éléments inversibles près :
1) les entiers premiers $p \in \mathbf{N}$ avec $p \equiv 3 \pmod 4$,
2) les entiers de Gauss $a + ib$ dont la norme $a^2 + b^2$ est un nombre premier.

Démonstration. D'abord les éléments ci-dessus sont irréductibles (pour 1) on l'a vu, pour 2) il suffit de considérer la norme). Réciproquement, si z est irréductible, on a $N(z) = z\overline{z} \in \mathbf{N}$. Soit p un nombre premier de $\mathbf{N}$ divisant $N(z)$. Si $p \equiv 3 \pmod 4$, p est premier dans $\mathbf{Z}[i]$ donc divise z ou $\overline{z}$ et on a $z = p$ à ± 1, $\pm i$ près. Si $p \in \Sigma$, on a $p = a^2 + b^2$ et l'entier de Gauss $t = a + ib$ est irréductible dans $\mathbf{Z}[i]$ donc divise z ou $\overline{z}$ et on a $z = t$ ou $\overline{t}$ à $\pm 1, \pm i$ près.

Remarque 6.11. Les éléments $1+i$ et $1-i$ sont associés. En revanche si $z = a+ib$ avec $N(z) = p$ premier et $p \equiv 1 \pmod 4$, on vérifie que z et $\overline{z}$ ne sont pas associés.

EXERCICES SUR LE CHAPITRE II

1. Idéaux.

1) Que peut-on dire de l'image réciproque d'un idéal maximal par un homomorphisme f ? Et si f est surjectif ?

---

2) Soit $\mathfrak{P}$ un idéal premier de A, $I_1, \ldots, I_n$ des idéaux. On suppose que $\mathfrak{P}$ contient le produit $I_1 \ldots I_n$, montrer que $\mathfrak{P}$ contient l'un des I_k.

---

3) Montrer que si I est un idéal non premier, il existe des idéaux I_1, I_2 distincts de I tels que $I \subset I_1$, $I \subset I_2$ et $I_1 I_2 \subset I$.

---

4) Soit A un anneau. Un élément $a \in A$ est dit ***nilpotent*** s'il existe $n \in \mathbf{N}$ tel que $a^n = 0$.

a) Montrer que l' ensemble $\mathrm{Nil}(A)$ des éléments nilpotents de A est un idéal.

b) Montrer que si $\mathfrak{P}$ est un idéal premier de A, on a $\mathrm{Nil}(A) \subset \mathfrak{P}$.

c) On va montrer, réciproquement, que $\mathrm{Nil}(A)$ est l'intersection des idéaux premiers de A. Pour ceci, soit $s \notin \mathrm{Nil}(A)$, et soit $S = \{1, s, s^2, \ldots, s^n, \ldots\}$. Montrer que l'ensemble des idéaux de A disjoints de S contient un élément maximal $\mathfrak{P}$ (utiliser le théorème de Zorn). Montrer que $\mathfrak{P}$ est premier (cf. Exercice 3). Conclure.

---

5) Soit A un anneau distinct de $\mathbf{Z}/4\mathbf{Z}$ et de $\mathbf{F}_2[X]/(X^2)$.
Montrer l'équivalence des propriétés suivantes :

a) A est intègre,

b) tout polynôme unitaire de degré n a au plus n racines dans A,

c) tout polynôme unitaire de degré 2 a au plus 2 racines dans A.

(Seul $c \implies a$ est délicat. Supposer A non intègre, soient $a, b \in A$ avec $ab = 0$, $a, b \neq 0$ et considérer les polynômes $X(X - a + b)$ et X^2. Puis, remarquer que si a est un diviseur de 0, et si x est dans A, xa est un diviseur de 0).
Étudier les cas $A = \mathbf{Z}/4\mathbf{Z}$ et $A = \mathbf{F}_2[X]/(X^2)$.

---

2. Anneaux noethériens.

1) Soit k un corps et $A = k[X, Y]$. Soit :

$$B = \{P \in k[X,Y] \mid P(X,Y) = \sum_{i>j} a_{ij} X^i Y^j\}.$$

a) Montrer que B est engendré, comme k-algèbre, par $X, X^2Y, \ldots, X^{i+1}Y^i, \ldots$

b) Montrer que B n'est pas noethérien.

---

2) Soit $A = \mathcal{H}(\mathbf{C})$ l'anneau des fonctions holomorphes dans tout le plan complexe.

a) Montrer que A est intègre. Quel est son corps des fractions ?

b) Montrer que A n'est pas noethérien (considérer, pour $k \in \mathbf{N}$, les idéaux :

$$I_k = \{f \in \mathcal{H}(\mathbf{C}) \mid f(z) = 0,\ \forall z \in \mathbf{N} - \{0, 1, \ldots, k\}\}$$

et penser à utiliser la fonction sin z).

D'autres propriétés de A sont étudiées dans les exercices du § 3.

---

3) On se propose de montrer que si A est noethérien, A n'a qu'un nombre fini d'idéaux premiers minimaux (au sens de l'inclusion).

a) Soit I un idéal de A. Montrer qu'il existe des idéaux premiers $\mathfrak{P}_1, \ldots, \mathfrak{P}_r$ tels que $I \supset \mathfrak{P}_1 \ldots \mathfrak{P}_r$ (raisonner par l'absurde et utiliser le point 3) de la définition de noethérien ; cf. aussi § 1 Exercice 3).

b) Conclure en appliquant a) à $I = \{0\}$ et en utilisant § 1 Exercice 2.

---

4) Soit A un anneau noethérien et M un A-module ([1]) de type fini.

a) Montrer que M est un module noethérien (i.e. toute suite croissante de sous-modules stationne).

b) Soit $f : M \longrightarrow M$, A-linéaire surjective. Montrer que f est bijective (considérer la suite $N_0 = f^{-1}\{0\}$, $N_1 = f^{-1}(N_0), \ldots$).

c) La même propriété est-elle encore vraie avec f injective ?

d) On ne suppose plus A noethérien. Montrer que b) reste vrai (si $x_1, \ldots, x_n$ engendrent M on écrira :

$$f(x_j) = \sum_{i=1}^{n} a_{ij}\, x_j, \quad \text{avec } a_{ij} \in A.$$

et on se ramènera au cas noethérien en considérant le sous-anneau de A engendré par les a_{ij} et d'autres éléments convenablement choisis).

---

3. Propriétés arithmétiques.

1) Calculer $A[X]^*$ lorsque A est un anneau quelconque (attention aux éléments nilpotents de A, cf. § 1 Exercice 4).

([1]) Un module est à un anneau ce qu'un espace vectoriel est à un corps, cf. [Bbki] Algèbre Ch. II ou [L] Ch. III.

2) Soit B un anneau et A un sous-anneau de B. Soit $b \in B$. On dit que b est ***entier*** sur A s'il vérifie une équation unitaire :

$$b^n + a_{n-1}b^{n-1} + \cdots + a_0 = 0 \text{ avec } a_0, \ldots, a_{n-1} \in A.$$

Un anneau intègre A est dit ***intégralement clos*** si pour tout élément $x \in K$ (où $K = \mathrm{Fr}(A)$ est le corps des fractions de A), x entier sur A implique $x \in A$.

a) Montrer qu'un anneau factoriel est intégralement clos.

b) Soit $d \in \mathbf{Z} - \{0\}$ un entier sans facteur carré. On pose :

$$\mathbf{Z}[\sqrt{d}] = \{a + b\sqrt{d} \in \mathbf{C} \mid a, b \in \mathbf{Z}\}$$

Montrer que si on a $d \equiv 1 \pmod 4$, $\mathbf{Z}[\sqrt{d}]$ n'est pas intégralement clos (considérer l'élément $\dfrac{1+\sqrt{d}}{2}$).

On peut montrer, en revanche, que si $d \equiv 2, 3 \bmod 4$, $\mathbf{Z}[\sqrt{d}]$ est intégralement clos, cf. [Sa] 2.5 ou III § 2,3 Exercice 12, mais pas nécessairement factoriel, comme le montre l'exemple de $d = -5$, cf. II, 3.18.1.

---

3) Retour sur l'anneau $A = \mathcal{H}(\mathbf{C})$, cf. § 2 Exercice 2.

a) Déterminer A^*. Montrer que $f \in A^* \iff \exists\, g \in A,\ f = \exp(g)$.

b) Montrer que f est irréductible dans A si et seulement si f a un seul zéro z qui, de plus, est simple (i.e. z vérifie $f(z) = 0$ et $f'(z) \neq 0$).

c) En déduire que A n'est pas factoriel. Retrouver aussi le b) de l'exercice 2 § 2.

d) Montrer que A est intégralement clos (cf. Exercice 2). On peut montrer que A vérifie le théorème de Bézout, i.e. que tout idéal de type fini de A est principal, cf. [R] Ch.15.

---

4) Dans $\mathbf{Z}[i\sqrt{5}]$, montrer que 3 et $2 + i\sqrt{5}$ n'ont pas de ppcm et que 9 et $3(2 + i\sqrt{5})$ n'ont pas de pgcd.

---

5) Soit A un anneau intègre et vérifiant la propriété (E) (cf. 3.14). On suppose que pour tous $a, b \in A$, a et b ont un ppcm.

a) Montrer que a et b ont aussi un pgcd et qu'on a :

$$ab = \mathrm{ppcm}\,(a, b) \times \mathrm{pgcd}\,(a, b).$$

b) Si $m = \mathrm{ppcm}\,(a, b)$, montrer que l'on a $(m) = (a) \cap (b)$.

c) Montrer que A vérifie le « lemme d'Euclide », cf. 3.19 donc que A est factoriel.

---

6) Soit A un anneau factoriel vérifiant le théorème de Bézout (i.e. tel que pour tous $a, b \in A$, l'idéal (a, b) est principal). Montrer que A est principal (attention, on ne suppose pas A noethérien).

---

7) Soit A un anneau euclidien, relativement à v. Soit $K = \mathrm{Fr}(A)$ et, pour $s \in A$, $s \neq 0$, soit $A_s = \left\{x \in K \mid x = \dfrac{a}{s^n},\ n \in \mathbf{N}\right\}$.

Montrer que A_s est un anneau euclidien (modifier la fonction v en posant $v(s) = 1$).

---

4. Stabilité des notions étudiées ; théorème de Gauss.

1) Soit k un corps. Etudier l'irréductibilité des polynômes ci-dessous dans $k[X, Y]$:
$Y - X^2$, $X^2 + Y^2 + 1$, $X^2 + Y^2 - 1$, $X^2 - Y^2 - 1$, $Y^2 - X^3$, $X^3 - Y^2 - X$, $XY^3 - X^2Y - Y^2 + X$.

---

2) Retour sur l'anneau $\mathbf{C}[X, Y]/(Y^2 - X^3) = A$ (cf. 4.7).
On désigne par x et y les images de X, Y dans A.

a) Montrer qu'on définit un homomorphisme φ de A dans $\mathbf{C}[T]$ en posant $\varphi(x) = T^2$, $\varphi(y) = T^3$. Montrer que φ est injectif. Quelle est son image ?

b) Montrer que $\mathrm{Fr}(A)$ est isomorphe à $\mathbf{C}(T)$, et en déduire que A n'est pas intégralement clos (cf. § 3 Exercice 2) (considérer l'élément T), donc pas factoriel. L'anneau A est l'anneau de la courbe plane C d'équation $Y^2 - X^3 = 0$; φ correspond au paramétrage $x = t^2$, $y = t^3$ de C cf. [Fu] Ch. II. Le fait que A ne soit pas intégralement clos est dû au point singulier de C à l'origine.

---

3) Montrer que $\mathbf{C}[X, Y]/(X^3 - Y^2 - X)$ n'est pas factoriel (en désignant par x, y les images de X, Y, on montrera que y est irréductible mais que l'idéal (y) n'est pas premier, on a :

$$y^2 = x^3 - x = x(x - 1)(x + 1)).$$

Ici la courbe correspondante est non singulière et l'anneau est intégralement clos.

---

4) Montrer que les anneaux $\mathbf{C}[X, Y]/(Y - X^2)$ et $\mathbf{C}[X, Y]/(XY - 1)$ sont factoriels et même principaux (trouver des anneaux plus simples à qui ils sont isomorphes, cf. § 3 Exercice 7).

---

5) Soient $P, Q \in \mathbf{C}[X, Y]$, sans facteurs communs.

a) Montrer qu'il existe des polynômes $A, B \in \mathbf{C}[X, Y]$ et $D \in \mathbf{C}[X]$, avec $D \neq 0$, tels que : $D = AP + BQ$ (travailler dans l'anneau $\mathbf{C}(X)[Y]$).

b) En déduire que l'ensemble :

$$V = \{(x, y) \in \mathbf{C}^2 \mid P(x, y) = Q(x, y) = 0\}$$

est fini. (L'intersection de deux courbes planes sans composante commune est finie, cf. [Fu] Ch.I § 6).

---

5. Anneaux principaux, factoriels, euclidiens.

1) Soit $A = \mathbf{R}[X, Y]/(X^2 + Y^2 + 1)$.

a) Soient ξ, η les images de X, Y dans A. Montrer que tout élément de A s'écrit de manière unique :

$$u = a(\xi)\eta + b(\xi) \text{ avec } a, b \in \mathbf{R}[X]$$

(utiliser une division euclidienne par X^2+Y^2+1, cf. 3.31).

b) Montrer que la donnée d'un homomorphisme de $\mathbf{R}$-algèbres de A dans $\mathbf{C}$ équivaut à la donnée d'un couple $(x,y) \in \mathbf{C}^2$ avec $x^2+y^2+1=0$.

c) Déduire de a) et b) que l'on a $A^* = \mathbf{R}^*$ (résoudre dans $\mathbf{C}^2$, $a(x)y+b(x)=0$ et $x^2+y^2+1=0$).

d) Montrer que A n'est pas euclidien (cf. 5.1).

---

2) Soit A un anneau intègre et noethérien. On suppose que tout idéal maximal de A est principal.

a) Montrer que A est factoriel.

b) Montrer que A est principal (commencer par étudier les idéaux $I=(a,b)$).

---

3) Soit k un corps infini. On se propose d'étudier les idéaux premiers de $k[X,Y]$.

a) Soit (P) un idéal principal. Montrer que (P) n'est pas maximal (considérer un idéal du type $(P, X-x)$ avec x convenablement choisi).

b) Montrer, en travaillant dans $k(X)[Y]$, que si on a F et $P \in k[X,Y]$ avec $P \neq 0$, il existe Q, $R \in k[X,Y]$ et $a \in k[X]$ tels que :

1) $a(X)F(X,Y) = P(X,Y)Q(X,Y) + R(X,Y)$,

2) $d^\circ_Y R < d^\circ_Y P$,

Soit m un idéal premier non principal de $k[X,Y]$.

c) Montrer que m contient deux polynômes $P(X)$ et $Q(Y)$ irréductibles.

d) En déduire qu'alors m est maximal et que $k[X,Y]/m$ est un k-espace vectoriel de dimension finie.

e) Si k est algébriquement clos, montrer que $m=(X-a,\ Y-b)$ avec $a,b \in k$. Quel est alors le corps résiduel ?

f) Si $k=\mathbf{R}$, montrer que le corps résiduel $K=\mathbf{R}[X,Y]/m$ est $\mathbf{R}$ ou $\mathbf{C}$.
Si $K=\mathbf{R}$, montrer que l'on a $m=(X-a,\ Y-b)$, avec $a,b \in \mathbf{R}$.
Si $K=\mathbf{C}$, montrer que l'on a $m=(P(X),Q(X,Y))$ (ou l'analogue en permutant X et Y), avec P irréductible de degré 2 et $Q(X,Y)=aX+bY+c$, pour $a,b,c \in \mathbf{R}$.

g) Déterminer tous les idéaux premiers de $k[X,Y]$.

---

4) On revient sur l'exemple de l'exercice 1) : $A=\mathbf{R}[X,Y]/(X^2+Y^2+1)$.

a) Soit $\overline{m}$ un idéal maximal de A, m son image réciproque dans $\mathbf{R}[X,Y]$. Montrer que $m=(X^2+Y^2+1,Q)$ avec $Q(X,Y)=aX+bY+c$ (utiliser Exercice 3). En déduire que $\overline{m}$ est principal.

b) Montrer que A est principal (cf. Exercice 2) et non euclidien (cf. Exercice 1).

---

5) Montrer que $\mathbf{C}[X,Y]/(X^2+Y^2+\varepsilon)$ est principal, pour $\varepsilon = \pm 1$ (faire un changement de variables pour se ramener à $XY-1$, cf. § 4 Exercice 5).

---

6) Montrer que $\mathbf{R}[X,Y]/(X^2+Y^2-1)$ n'est pas factoriel.

---

6. L'anneau $\mathbf{Z}[i]$ et le théorème des deux carrés.

1) Soit $x \in \mathbf{R}$. On suppose que $\operatorname{tg} 2\pi x$ est un rationnel distinct de $0, 1, -1$. Montrer que x est irrationnel (montrer que $z = e^{i4\pi x}$ est dans $\mathbf{Q}[i]$, et que, si $x \in \mathbf{Q}$, z est entier sur $\mathbf{Z}[i]$. Conclure à l'aide de § 3 Exercice 2).

---

2) Soit $n \in \mathbf{N}$, $n \geq 2$. Montrer que les trois conditions suivantes sont équivalentes :

a) -1 est un carré dans $\mathbf{Z}/n\mathbf{Z}$.

b) $n = x^2 + y^2$ avec $x, y \in \mathbf{N}$, x et y premiers entre eux.

c) $n = 2^\alpha \prod_k p_k^{\alpha_k}$ avec $\alpha \in \{0, 1\}$ et $p_k \equiv 1 \pmod 4$.

(On montrera les implications $a \implies b \implies c \implies a$. Le seul point délicat est $c \implies a$, on pourra considérer $z = x + iy = (1+i)^\alpha \prod_k z_k^{\alpha_k}$ avec $N(z_k) = p_k$ et montrer que x et y sont premiers entre eux en utilisant 6.9).

---

3) a) Montrer que si p est premier et $p \equiv 1 \pmod 4$, l'ensemble des irréductibles de $\mathbf{Z}[i]$ de norme p est, à la relation d'association près, consitué de deux éléments $z_p, \overline{z}_p$, non associés.

Dans la suite, on suppose choisi pour tout $p \equiv 1 \pmod 4$ un tel z_p avec $N(z_p) = p$. Tout élément z de $\mathbf{Z}[i]$ admet alors une décomposition unique :

$$z = u(1+i)^\alpha \prod q^{\alpha_q} \prod z_p^{\alpha_p} \prod \overline{z_p}^{\beta_p}$$

avec $u \in \mathbf{Z}[i]^*$ et $q \equiv 3 \pmod 4$.

b) Soit n un entier écrit sous la forme

$$n = 2^\alpha \prod p^{\alpha_p} \prod q^{\alpha_q}$$

avec $p \equiv 1 \pmod 4$ et $q \equiv 3 \pmod 4$, montrer que l'on a :

$$S(0, n) = \{z \in \mathbf{Z}[i] \mid N(z) = n\} = A_n \cup -A_n \cup iA_n \cup -iA_n$$

$$\text{où } A_n = \Big\{(1+i)^\alpha \prod q^{\alpha_q} \prod z_p^{\alpha_p - t_p}\, \overline{z_p}^{t_p} \mid t_p \in \{0, 1, \ldots, \alpha_p\}\Big\}.$$

c) En déduire que les éléments $n = x_0^2 + y_0^2$ de Σ tels que l'équation $n = x^2 + y^2$ n'ait pas d'autres solutions dans $\mathbf{N} \times \mathbf{N}$ que (x_0, y_0) et (y_0, x_0) sont ceux qui ont au plus un facteur premier $\equiv 1 \pmod 4$. ([2])

d) Soit $m \in \mathbf{N}^*$ et soit $z \in \mathbf{Z}[i]$ tels que $z^m \in \mathbf{Z}$. Montrer qu'on a l'une des possibilités suivantes : $z \in \mathbf{Z} \cup i\mathbf{Z}$, ou $\dfrac{z}{1+i} \in \mathbf{Z} \cup i\mathbf{Z}$ et retrouver ainsi le résultat de l'exercice 1).

e) Décomposer en produit de facteurs irréductibles dans $\mathbf{Z}[i]$, les entiers de Gauss $1 + 3i$, $70 + i$, $4 + 3i$, $201 + 43i$, $99 + i$.

([2]) Ce résultat était connu de Fermat.

III. CORPS : THÉORIE ÉLÉMENTAIRE

Sauf dans l'énoncé du théorème de Wedderburn (4.9), les corps seront toujours supposés ***commutatifs***.

1. Les techniques vectorielles.

a) Degré d'une extension. Eléments algébriques.

Définition 1.1.
Soient K, L des corps avec $K \subset L$. On dit que L est une **extension** *(sous-entendu : de corps) de K.*

Exemples 1.2.
On a ainsi $\mathbf{R} \subset \mathbf{C}$, $\mathbf{R} \subset \mathbf{R}(T)$, $\mathbf{Q} \subset \mathbf{Q}(i)\ldots$

Remarques 1.3.
1) Si K est un sous-corps de L, L est un K-espace vectoriel.
2) Si $\dim_K L$ est finie, on pose $[L : K] = \dim_K L$ et l'entier $[L : K]$ s'appelle le **degré** de L sur K.
3) Si K et L sont des corps finis, on a $|L| = |K|^n$ avec $n = [L : K]$.

Théorème 1.4 (de la base télescopique).
Soient $K \subset L \subset M$ des corps, $(e_i)_{i\in I}$ une base de L sur K, $(f_j)_{j\in J}$ une base de M sur L. Alors la famille $(e_i f_j)_{(i,j)\in I\times J}$ est une base de M sur K.

Corollaire 1.5 (multiplicativité du degré).
Dans la situation de 1.4, si les degrés sont finis, on a $[M : K] = [M : L][L : K]$.

Démonstration (de 1.4)
1) La famille $e_i f_j$ est libre sur K :
En effet, si on a $\sum_{i,j} \lambda_{ij} e_i f_j = 0$, avec $\lambda_{ij} \in K$ on a $\sum_j f_j(\sum_i \lambda_{ij} e_i) = 0$, donc, puisque f_j est une base de M sur L, on a pour tout j, $\sum_i \lambda_{ij} e_i = 0$, et puisque e_i est une base de L sur K, on a bien $\lambda_{ij} = 0$ pour tous i, j.
2) Elle engendre M : soit $x \in M$, on l'écrit $x = \sum_j \mu_j f_j$, $\mu_j \in L$, puis on décompose $\mu_j = \sum_i \lambda_{ij} e_i$ d'où finalement $x = \sum_{i,j} \lambda_{ij} e_i f_j$ avec $\lambda_{ij} \in K$.

Ce théorème très simple est, en fait, un outil très efficace en théorie des corps comme nous le verrons dans la suite.

Définition 1.6.
1) Soit $K \subset L$ une extension et A une partie de L. On dit que A engendre L sur K et on écrit alors $L = K(A)$ si L est le plus petit sous-corps de L contenant K et A. Si A est fini, $A = \{\alpha_1, \ldots, \alpha_n\}$, on note $L = K(\alpha_1, \ldots, \alpha_n)$.
2) L'extension $K \subset L$ est dite **monogène** *s'il existe $\alpha \in L$ tel que $L = K(\alpha)$.*

Remarques 1.7.

1) Si on a une extension $K \subset L$ et si $\alpha \in L$ on note $K[\alpha]$ le sous anneau de L engendré par K et α. On a $K[\alpha] \subset K(\alpha)$. ***Attention***, $K[\alpha]$ n'est pas en général isomorphe à l'anneau des polynômes $K[T]$, ni $K(\alpha)$ au corps des fractions rationnelles $K(T)$.

2) On peut décrire ainsi $K[\alpha]$ et $K(\alpha)$:
si $x \in K[\alpha]$, x s'écrit $x = P(\alpha)$ avec $P \in K[T]$ i.e. $x = a_n\alpha^n + \cdots + a_0$, $a_i \in K$,
si $x \in K(\alpha)$, on a $x = \dfrac{P(\alpha)}{Q(\alpha)}$ avec $P, Q \in K[T]$, $Q(\alpha) \neq 0$.
La différence entre $K[\alpha]$ et $K[T]$ (resp. $K(\alpha)$ et $K(T)$) vient du fait qu'on peut avoir $Q(\alpha) = 0$ avec $Q \in K[T]$, $Q \neq 0$. De façon précise, il y a deux types d'éléments de L relativement à K :

Définition 1.8.
Soit $K \subset L$ une extension et soit $\alpha \in L$. Soit $\varphi : K[T] \longrightarrow L$ l'homomorphisme défini par $\varphi|_K = \mathrm{id}_K$ et $\varphi(T) = \alpha$.
1) Si φ est injectif, on dit que α est **transcendant sur** *K,*
2) sinon, on dit que α est **algébrique sur** *K. Ceci signifie qu'il existe un polynôme $P(T)$, non nul, tel que $P(\alpha) = 0$. Plus précisément, si $I = \mathrm{Ker}\,\varphi$, I est un idéal principal non nul (cf. II, 3.29 et 3.32), et on a donc $I = (P)$, avec $P \neq 0$ et on peut supposer P unitaire. Le polynôme P est, par définition, le* **polynôme minimal** *de α sur K.*

Exemples 1.9.

1) On peut montrer que e et π sont transcendants sur $\mathbf{Q}$ (mais pas sur $\mathbf{R}$, bien entendu). Dans $K(T)$, l'élément T est transcendant sur K.

2) Les nombres $\sqrt{2}$, i, $\sqrt[3]{2}, \cdots$ sont algébriques sur $\mathbf{Q}$, de polynômes minimaux respectifs $X^2 - 2$, $X^2 + 1$, $X^3 - 2, \cdots$.

On peut alors préciser la structure de $K(\alpha)$ selon les deux cas de 1.8 :

Proposition 1.10.
Si α est transcendant, on a $K[\alpha] \simeq K[T]$ et $K(\alpha) \simeq K(T)$ (et donc $K(\alpha)$ est distinct de $K[\alpha]$).

Démonstration. C'est clair, car l'homomorphisme $\varphi : K[T] \to L$ est injectif et d'image $K[\alpha]$.

Théorème 1.11.
Soit $K \subset L$ une extension et soit $\alpha \in L$. Les propriétés suivantes sont équivalentes :
1) α est algébrique sur K,
2) on a $K[\alpha] = K(\alpha)$,
3) on a $\dim_K K[\alpha] < +\infty$.

Précisément, si P est le polynôme minimal de α, P est irréductible et on a $\dim_K K[\alpha] = [K[\alpha] : K] = d^\circ P$. Cet entier s'appelle le ***degré*** *de α.*

Démonstration. Prouvons 1) $\Longrightarrow$ 2). Supposons α algébrique de polynôme minimal P. D'après II §0, on a un isomorphisme :

$$\overline{\varphi} = K[T]/(P) \longrightarrow K[\alpha].$$

Comme l'anneau $K[\alpha]$ est inclus dans L, il est intègre, de sorte que l'idéal (P) est premier, donc P est irréductible dans l'anneau principal $K[T]$, donc (P) est maximal (cf. II, 3.22). Il en résulte que $K[\alpha]$ est un corps d'où $K[\alpha] = K(\alpha)$.
On a aussi 2) $\Longrightarrow$ 1) (par 1.10) et 3) $\Longrightarrow$ 1) est clair également car si α est transcendant on a $K[\alpha] \simeq K[T]$ par 1.10, et cet espace vectoriel est de dimension infinie sur K. Enfin 1) $\Longrightarrow$ 3) et la remarque sur la dimension résultent de l'isomorphisme $K[T]/(P) \longrightarrow K[\alpha]$, car si P est de degré n, on montre par division euclidienne, cf. II, 3.31, que $1, \alpha, \ldots, \alpha^{n-1}$ est une base de $K[\alpha]$ sur K .

On se reportera aux exercices pour d'autres démonstrations.

Définition 1.12.
1) Une extension $K \subset L$ est dite ***finie*** *si on a $\dim_K L (= [L : K]) < +\infty$.*
2) Une extension $K \subset L$ est dite ***algébrique*** *si pour tout $\alpha \in L$, α est algébrique sur K.*

Remarque 1.13. Le théorème 1.11 montre que toute extension finie est algébrique. Nous verrons plus loin que la réciproque est fausse.

Théorème 1.14.
Soit $K \subset L$ une extension et posons

$$M = \{x \in L \mid x \text{ est algébrique sur } K\}.$$

Alors M est un sous-corps de L.

Démonstration. Soient $\alpha, \alpha' \in M$. Considérons le sous-anneau $K[\alpha, \alpha']$ engendré par α et α'. On a $K[\alpha, \alpha'] = K[\alpha][\alpha']$ et donc comme α' est algébrique sur K, donc *a fortiori* sur $K[\alpha]$, le théorème 1.11 montre que $K[\alpha]$ et $K[\alpha, \alpha']$ sont des corps. De plus le théorème 1.11 et la multiplicativité des degrés donnent $[K[\alpha, \alpha'] : K] < +\infty$. Mais alors, comme $K[\alpha + \alpha']$ et $K[\alpha\alpha']$ sont inclus dans $K[\alpha, \alpha']$, ils sont eux aussi de dimension finie sur K et donc, cf. 1.11, $\alpha + \alpha'$ et $\alpha\alpha'$ sont algébriques donc sont dans M.

Remarque 1.15. Sans les techniques vectorielles, ce théorème n'est pas évident. On s'en convaincra aisément en cherchant un polynôme de $\mathbf{Q}[T]$ qui s'annule en $\sqrt[7]{5} + \sqrt[3]{7}\ \sqrt[5]{3}$.

Exemple 1.16. Soit $A = \{\alpha \in \mathbf{C} \mid \alpha \text{ algébrique sur } \mathbf{Q}\}$, A est un corps, algébrique sur $\mathbf{Q}$, mais l'extension $\mathbf{Q} \subset A$ n'est pas finie car il existe des éléments de A de degré arbitrairement grand, par exemple $\sqrt[n]{2}$, qui est de degré n car le polynôme $X^n - 2$ est irréductible sur $\mathbf{Q}$ (en vertu du critère d'Eisenstein, cf. ci-dessous 3.2).

Définition 1.17.
Un corps K est dit ***algébriquement clos*** *s'il vérifie l'une quelconque des propriétés équivalentes suivantes :*
1) tout polynôme $P \in K[X]$ de degré ≥ 1 admet une racine dans K,

2) tout polynôme $P \in K[X]$ est produit de polynômes de degré 1,
3) les éléments irréductibles de $K[X]$ sont les $X - a$, $a \in K$,
4) si une extension $K \subset L$ est algébrique, on a $L = K$.

Exemples 1.18.

1) Le corps $\mathbf{C}$ est algébriquement clos, (théorème de D'Alembert-Gauss).

2) Le corps A défini en 1.16 ci-dessus est lui aussi algébriquement clos. C'est même la clôture algébrique de $\mathbf{Q}$ (cf. ci-dessous 1.33). On montre aisément que A est dénombrable, ce qui, puisque $\mathbf{R}$ ne l'est pas, prouve l'existence dans $\mathbf{R}$ de nombres transcendants sur $\mathbf{Q}$, cf. Exercice 3.

b) Application : constructions à la règle et au compas.

Nous allons résoudre par la négative deux problèmes de construction posés par les grecs : la duplication du cube et la trisection de l'angle. On se rendra compte, là encore, de l'efficacité des méthodes vectorielles. Pour des détails, notamment historiques, on consultera l'excellent livre de J.-C. Carréga, cf. [Ca].

On considère le plan euclidien $\mathbf{R}^2$ muni de deux points $O = (0,0)$ et $I = (1,0)$. On cherche à construire de nouveaux points « à la règle et au compas ». Précisément, soit A une partie de $\mathbf{R}^2$, on considère trois types de figures construites à partir de A :

a) les ***droites affines*** $< P, Q >$ pour $P, Q \in A$, $P \neq Q$,

b) les ***cercles*** centrés en un point $P \in A$, passant par un point $Q \in A$, avec $P \neq Q$,

b') les ***cercles*** centrés en $P \in A$, de rayon $||QR||$, avec $Q, R \in A$, $Q \neq R$.

Définition 1.19.

1) Soit A une partie de $\mathbf{R}^2$ et soit $M \in \mathbf{R}^2$. On dit que M est constructible (sous-entendu, à la règle et au compas), en un pas, à partir de A s'il existe deux éléments distincts, droites ou cercles, de type a) ou b) [1] *ci-dessus, dont M soit un point d'intersection.*

Un point $M \in \mathbf{R}^2$ est dit ***constructible*** *s'il existe une suite $A_0 \subset \ldots \subset A_n$ de parties de $\mathbf{R}^2$ avec :*

a) $A_0 = \{O, I\}$, b) $M \in A_n$, c) $A_i = A_{i-1} \cup \{M_i\}$ où M_i est constructible en un pas à partir de A_{i-1}.

Un nombre réel x est dit constructible si et seulement si $(x, 0)$ l'est.

Proposition 1.20.

Les points suivants de $\mathbf{R}^2$ sont constructibles :
1) les $(n, 0)$ pour $n \in \mathbf{N}$,
2) les $(0, n)$ pour $n \in \mathbf{N}$,
3) les $(x, 0)$ pour $x \in \mathbf{Q}$.
Si le réel $x > 0$ est constructible, il en est de même de $\sqrt{x}$.

Démonstration.

L'assertion 1) est claire et 2) aussi car on construit l'axe des ordonnées comme médiatrice des points $1, -1$ de l'axe des abscisses. Pour 3) on montre d'abord que si on a trois points $P, Q, R \in \mathbf{R}^2$, distincts, on sait construire à la règle et au compas la parallèle à $< P, Q >$ passant par R.

---

[1] *Le lecteur vérifiera que l'on peut remplacer les cercles de type b) par ceux de type b') sans changer, en définitive, les points constructibles.*

On construit alors $\frac{p}{q} \in \mathbf{Q}$ en menant la parallèle au segment $< (p,0);(0;q) >$ passant par $(0,1)$ (c'est le théorème de Thalès!).
Enfin, pour le dernier point, on pose $a = \frac{x-1}{2}$, $b = a+1 = \frac{x+1}{2}$. On a $(b-a)(b+a) = b+a = b^2 - a^2 = x$, donc $b^2 = a^2 + c^2$ avec $c^2 = x$. On construit alors à partir de x les points $(0,a)$ et $(b,0)$ et on construit $(c,0)$ comme troisième sommet d'un triangle rectangle de sommets O et $(0,a)$ et dont l'hypothénuse a pour longueur b (c'est le théorème de Pythagore!).

Théorème 1.21.
Soit x un réel constructible. Alors x est algébrique sur $\mathbf{Q}$ et son degré $[\mathbf{Q}[x] : \mathbf{Q}]$ est une puissance de 2.

Remarque 1.22. Attention, la réciproque est fausse, par exemple il existe des x de degré 4 non constructibles. La condition suffisante est que la clôture normale de $\mathbf{Q}(x)$ soit de degré 2^n (cf. par exemple [St]).
Démonstration (de 1.21) Par hypothèse, on a une suite $A_0 \subset A_1 \subset \ldots \subset A_n$ comme en 1.19, avec $(x,0) \in A_n$. Soit K_i le sous-corps de $\mathbf{R}$ engendré sur $\mathbf{Q}$ par les coordonnées des points de A_i. On a donc $K_0 = \mathbf{Q}$ et $x \in K_n$.

Lemme 1.23.
On a $[K_i : K_{i-1}] = 1$, 2 ou 4.

Admettons un instant ce lemme, alors une récurrence immédiate montre que $[K_n : \mathbf{Q}]$ est une puissance de 2 en vertu de la multiplicativité des degrés (cf. 1.5) et comme $\mathbf{Q}[x]$ est un sous-corps de K_n, $[\mathbf{Q}[x] : Q]$ divise $[K_n : \mathbf{Q}]$ (toujours par 1.5) d'où le résultat. ([2])
Démonstration (de 1.23) On a $A_i = A_{i-1} \cup \{M_i\}$ avec $M_i = (x_i, y_i)$, donc $K_i = K_{i-1}(x_i, y_i)$.
Par définition M_i est intersection de droites ou de cercles, dont les équations sont dans $K_{i-1}[X,Y]$ de sorte que x_i et y_i vérifient des équations de degré ≤ 2 sur K_{i-1}. On a donc $[K_{i-1}(x_i) : K_{i-1}] \leq 2$ et $[K_{i-1}(x_i, y_i) : K_{i-1}(x_i)] \leq 2$ d'où le résultat.

En fait, une étude plus attentive montre que $[K_i : K_{i-1}] = 1$ ou 2 car l'intersection de deux cercles se ramène toujours à une intersection d'un cercle (l'un des deux) et d'une droite (l'axe radical).

Applications : où l'on surpasse les grecs.

i) Impossibilité de la duplication du cube.

Le problème (dit de Délos) est de construire à la règle et au compas un nombre a tel que le cube d'arête a ait un volume double du cube unité. Autrement dit, on a $a^3 = 2$ et il s'agit donc de construire $a = \sqrt[3]{2}$.

Proposition 1.24.
Le nombre $\sqrt[3]{2}$ n'est pas constructible.

Démonstration. En effet le polynôme $X^3 - 2$ est irréductible sur $\mathbf{Q}$ (car il n'a pas de racines dans $\mathbf{Q}$, voir aussi 3.2) donc c'est le polynôme minimal de $\sqrt[3]{2}$ et donc on a (cf. 1.11) $[\mathbf{Q}(\sqrt[3]{2}) : \mathbf{Q}] = 3$ qui n'est pas une puissance de 2.

([2]) Simple, mais efficace, n'est-ce-pas ?

ii) Impossibilité de la trisection de l'angle.

On cherche à « trisecter » l'angle $\pi/3$, donc, à construire $x = \cos \pi/9$. La formule $\cos 3\theta = 4\cos^3 \theta - 3\cos \theta$ montre que x vérifie $8x^3 - 6x - 1 = 0$ et, ce polynôme étant irréductible sur $\mathbf{Q}$, on a encore $[\mathbf{Q}(x) : \mathbf{Q}] = 3$.

iii) Impossibilité de la quadrature du cercle.

On cherche cette fois un carré de côté a dont l'aire soit celle du cercle unité i.e. on cherche a vérifiant $a^2 = \pi$, donc $a = \sqrt{\pi}$. Mais on a le théorème suivant :

Théorème 1.25 (Lindemann).
Le nombre π (donc aussi $\sqrt{\pi}$) est transcendant sur $\mathbf{Q}$, donc non constructible (cf. 1.21).

Le théorème de Lindemann est beaucoup moins élémentaire que ce qui précède et utilise des techniques d'analyse. On se reportera à [St] Chapitre VI pour une démonstration.

c) Corps de rupture, corps de décomposition.

Nous allons résoudre maintenant les deux problèmes suivants :

1) Étant donné $P \in K[X]$, irréductible de degré $d > 1$, construire une extension dans laquelle P admet une racine a (donc est divisible par $X - a$ et, en particulier, n'est plus irréductible).

2) Étant donné $P \in K[X]$, construire une extension dans laquelle P soit décomposé en produit de facteurs de degré 1.

Définition 1.26.
Soit K un corps, $P \in K[X]$ un polynôme ***irréductible****. Une extension $L \supset K$ est appelée un* ***corps de rupture*** *de P sur K si L est une extension monogène $L = K(\alpha)$ avec $P(\alpha) = 0$.*

Théorème 1.27.
Soit $P \in K[X]$, irréductible. Il existe un corps de rupture de P sur K, unique à isomorphisme près.

Démonstration.

a) *Existence.*

On prend $L = K[X]/(P)$, c'est un corps (cf. II, 3.22), dans lequel K s'injecte et si x est l'image de X dans L, on a bien $P(x) = 0$ et $L = K(x)$.

Ainsi, par exemple, $\mathbf{C} = \mathbf{R}[X]/(X^2 + 1)$ est un corps de rupture de $X^2 + 1$, $\mathbf{Q}(\sqrt[3]{2}) = \mathbf{Q}[X]/(X^3 - 2)$ est un corps de rupture de $X^3 - 2$, etc.

b) *Unicité.*

Plus précisément, on a le lemme suivant :

Lemme 1.28.
Soient K, K' deux corps, $i : K \longrightarrow K'$ un isomorphisme que l'on étend de manière évidente en un isomorphisme, noté encore i, de $K[X]$ sur $K'[X]$ en envoyant X sur X. Soit $P \in K[X]$ un polynôme irréductible et soit $P' = i(P)$. Soit L (resp. L') un corps de rupture de P sur K (resp. de P' sur K') engendré par une racine x de P (resp. x' de P'). Alors il existe un unique isomorphisme φ de L sur L', prolongeant i, et vérifiant $\varphi(x) = x'$.

Démonstration. On a un morphisme $u : K[X]/(P) \longrightarrow L$ défini par $u(\overline{X}) = x$, (où $\overline{X}$ désigne l'image de X dans le quotient). C'est un isomorphisme (il est

clairement surjectif par définition d'un corps de rupture et il est injectif car P étant irréductible est le polynôme minimal de x). On a, de même, $u' : K'[X]/(P') \longrightarrow L'$. Par ailleurs, l'isomorphisme $i : K[X] \to K'[X]$ induit encore un isomorphisme par passage au quotient :

$$\bar{i} : K[X]/(P) \longrightarrow K'[X]/(P').$$

Il suffit alors de prendre $\varphi = u' \bar{i} u^{-1}$ pour avoir l'isomorphisme cherché.

Si L est un corps de rupture de P, le polynôme P n'est pas, en général, entièrement factorisé sur L. Par exemple, si $K = \mathbf{Q}$ et $P(X) = X^3 - 2$, on a $L = \mathbf{Q}(\sqrt[3]{2})$ et on peut prendre la racine cubique réelle de 2 de sorte qu'on a $L \subset \mathbf{R}$, mais alors, les autres racines de P (i.e. $j\sqrt[3]{2}$ et $j^2\sqrt[3]{2}$) ne sont pas dans L. Ceci nous amène à poser la définition suivante :

Définition 1.29.
Soit $P \in K[X]$ un polynôme, irréductible ou non, de degré n. On appelle **corps de décomposition** *de P sur K une extension L de K qui est telle que :*
1) dans $L[X]$, P est produit de facteurs de degré 1 (ou encore, P a "toutes" ses racines dans L),
2) le corps L est minimal pour cette propriété (ou encore, les racines de P engendrent L).

Théorème 1.30.
Pour tout $P \in K[X]$, il existe un corps de décomposition de P sur K, unique à isomorphisme près. On le note $D_K(P)$.

Démonstration.
a) *Existence.*
On raisonne par récurrence sur $d^\circ P$. Si P est de degré 1, $L = K$ convient. Si P est produit de facteurs de degré 1, là encore $L = K$ convient. Sinon, P admet un facteur irréductible Q, de degré ≥ 2. Soit K' un corps de rupture de Q sur K et x une racine de Q, donc de P, dans K'. On a donc $P(X) = (X - x)P'(X)$ dans $K'[X]$. Comme on a $d^\circ P' < d^\circ P$, il existe un corps de décomposition L de P' sur K'. Mais alors, L est aussi un corps de décomposition de L sur K. En effet, dans L, P' admet $n-1$ racines $x_2, \ldots, x_n$ (éventuellement multiples) et $L = K'(x_2, \ldots, x_n)$ donc, comme $K' = K(x)$, les racines $x, x_2, \ldots, x_n$ de P sont toutes dans L et engendrent L sur K.

b) *Unicité.*
Comme précédemment, on a un lemme un peu plus précis :

Lemme 1.31.
Soient K, K' deux corps, $i : K \longrightarrow K'$ un isomorphisme, que l'on étend comme en 1.28 en un isomorphisme, noté encore i, de $K[X]$ sur $K'[X]$. Soient $P \in K[X]$ un polynôme, $P' = i(P)$ et L (resp. L') un corps de décomposition de P sur K (resp. de P' sur K'). Alors il existe un isomorphisme φ de L sur L', prolongeant i.

Démonstration.
On raisonne par récurrence sur $[L : K]$. Si $L = K$, on a $L' = K'$. Sinon, soit $\alpha \in L - K$ une racine de P et soit Q le polynôme minimal de α qui est un facteur irréductible de P. Soit $Q' = i(Q)$ qui est donc un facteur irréductible de P', et soit α' une racine de Q' dans L'. On pose $M = K(\alpha)$, $M' = K'(\alpha')$. Le corps M (resp. M') est un corps de rupture de Q sur K (resp. de $Q' = i(Q)$ sur K').

En vertu du lemme 1.28 (unicité du corps de rupture), il existe un isomorphisme $\psi : M \longrightarrow M'$ prolongeant i et tel que $\psi(\alpha) = \alpha'$. Dans $M[X]$ (resp. $M'[X]$) on a $P(X) = (X - \alpha)S(X)$ (resp. $P'(X) = (X - \alpha')S'(X)$) et, comme on a $\psi(\alpha) = \alpha'$, on a aussi $\psi(S) = S'$. Mais alors L est un corps de décomposition de S sur M (resp. L' de S' sur M') et par l'hypothèse de récurrence, on a un isomorphisme $\varphi : L \longrightarrow L'$ prolongeant ψ et cet isomorphisme convient.

Exemples 1.32.
1) Pour $K = \mathbf{Q}$, $P(X) = X^3 - 2$, on a $D_K(P) = Q(\sqrt[3]{2}, j)$.
2) Pour $K = \mathbf{Q}$, $P(X) = X^4 - 2$, $D_K(P) = Q(\sqrt[4]{2}, i)$.

Un prolongement naturel de la notion de corps de décomposition est celle de clôture algébrique :

Définition 1.33.
Une extension $\overline{K}$ de K est appelée une ***clôture algébrique*** *de K si elle vérifie :*
1) $\overline{K}$ est algébriquement clos,
2) $\overline{K}$ est algébrique sur K.

Par exemple, $\mathbf{C}$ est une clôture algébrique de $\mathbf{R}$, A (cf. 1.16) est une clôture algébrique de $\mathbf{Q}$.
Pour un énoncé d'existence et d'unicité de la clôture algébrique, cf. [J] 2 Ch. 8.

2. Les corps finis.

a) Caractéristique et cardinal.

Rappelons d'abord (cf. 1.3.3) que si K et L sont des corps finis avec $K \subset L$, et $\dim_K L = n$, on a $|L| = |K|^n$.

Définition 2.1.
Soit K un corps (quelconque). On appelle ***sous-corps premier*** *de K le plus petit sous-corps de K (contenant 1).*

Description du sous-corps premier.

Soit $\varphi : \mathbf{Z} \longrightarrow K$ l'homomorphisme d'anneaux défini par $\varphi(n) = n.1 = 1 + \cdots + 1$ (répété n fois). Le noyau de φ est alors un idéal de $\mathbf{Z}$, donc $\operatorname{Ker} \varphi$ est de la forme $p\mathbf{Z}$ et comme $\mathbf{Z}/p\mathbf{Z} \simeq \operatorname{Im} \varphi$ est inclus dans K donc est intègre, $p\mathbf{Z}$ est un idéal premier. Il y a donc deux cas : p est nul ou est un nombre premier.

Définition 2.2.
Le nombre p, générateur de $\operatorname{Ker} \varphi$ *est appelé la* ***caractéristique*** *du corps K. La caractéristique d'un corps est donc 0 ou un nombre premier. On la note* $\operatorname{car}(K)$.

Remarques 2.3.

0) Si le corps K est de caractéristique $p > 0$, on a par définition, dans K, $p.1 = 0$, mais aussi, pour tout x de K : $px = p(1x) = (p.1)x = 0$.

1) Si $\operatorname{car}(K) = 0$, on a $\varphi(\mathbf{Z}) \simeq \mathbf{Z} \subset K$, de sorte que K est infini. De plus K contient le corps des fractions de $\mathbf{Z}$, i.e. $\mathbf{Q}$ qui est le sous-corps premier de K.

2) Si K est fini, on a $\operatorname{car}(K) = p > 0$. Le sous-corps premier de K est $\mathbf{Z}/p\mathbf{Z}$. On le note aussi $\mathbf{F}_p$.

3) Si K est fini, on a $\operatorname{car}(K) = p > 0$ et la remarque initiale montre que l'on a $q = |K| = p^n$: ***le cardinal d'un corps fini est une puissance d'un nombre premier*** : sa caractéristique (par exemple, il n'y a pas de corps de cardinal 6).

Proposition 2.4.
Soit K un corps de caractéristique $p > 0$. L'application $F : K \longrightarrow K$ définie par $F(x) = x^p$ est un homomorphisme de corps appelé ***homomorphisme de Frobenius.***
Si K est fini, c'est un automorphisme.
Si K est le corps premier $\mathbf{F}_p$, c'est l'identité.

Démonstration. Il est clair qu'on a $F(xy) = F(x)F(y)$. Par ailleurs on calcule $(x+y)^p$ par la formule du binôme :

$$(x+y)^p = x^p + C_p^1 x^{p-1}y + \cdots + C_p^i x^{p-i}y^i + \cdots + y^p.$$

Mais, si $1 \leq i \leq n-1$, on sait que p divise C_p^i et donc $C_p^i x^{p-i}y^i = 0$ dans K puisque $\mathrm{car}(K) = p$, cf. 2.3.0. On a donc bien $(x+y)^p = x^p + y^p$.
Un homomorphisme de corps est toujours injectif (son noyau est un idéal, donc (0) ou K et ce n'est pas K car 1 s'envoie sur 1), donc bijectif si K est fini.
Enfin, si $K = \mathbf{F}_p$ et si $x \in \mathbf{F}_p^*$, comme $|\mathbf{F}_p^*|$ est de cardinal $p-1$, on a $x^{p-1} = 1$ par le théorème de Lagrange donc $x^p = x$ et ceci vaut aussi pour $x = 0$ ce qui signifie que F est l'identité de $\mathbf{F}_p$.

b) Existence et unicité des corps finis.

Théorème 2.5.
Soit p un nombre premier et soit $n \in \mathbf{N}^*$. On pose $q = p^n$.
1) Il existe un corps K à q éléments, c'est le corps de décomposition du polynôme $X^q - X$ sur $\mathbf{F}_p$.
2) En particulier, K est unique, à isomorphisme près, (cf. 1.30). On le note $\mathbf{F}_q$.

Démonstration.
1) Si K est un corps à q éléments et si $x \in K^*$, on a $x^{q-1} = 1$, de sorte que tout élément de K est racine du polynôme $X^q - X$. D'autre part le sous-corps premier de K est nécessairement $\mathbf{F}_p$, puisque le cardinal de K est une puissance de sa caractéristique (cf. 2.3.3). Comme $X^q - X$ est à coefficients dans $\mathbf{F}_p[X]$, on a bien $K = D_{\mathbf{F}_p}(X^q - X)$.
2) Réciproquement, soit $K = D_{\mathbf{F}_p}(X^q - X)$ et soit $k \subset K$ l'ensemble des racines de $X^q - X$. Alors, k est un corps, car si x, y sont dans k, on a $x^q = x$, $y^q = y$, donc aussi $(xy)^q = xy$ et $(x+y)^q = x + y$ car l'application $x \mapsto x^q$ de K dans K n'est autre que l'automorphisme de Frobenius itéré n fois.
De plus, si on pose $P(X) = X^q - X$, on a $P'(X) = qX^{q-1} - 1 = -1$ car q est une puissance de la caractéristique. Il en résulte que les racines de P sont simples et donc on a $|k| = q$: on a obtenu un corps à q éléments. Enfin, il est clair que l'on a $k = D_{\mathbf{F}_p}(X^q - X) = K$.

Remarque 2.6. En fait, l'unicité du corps à q éléments est vraie aussi en un sens plus fort : si on note $\overline{\mathbf{F}_p}$ une clôture algébrique de $\mathbf{F}_p$, il y a dans $\overline{\mathbf{F}_p}$ un unique corps à $q = p^n$ éléments : c'est, comme le montre la preuve de 2.5, l'ensemble des racines du polynôme $X^q - X$.

c) Etude du groupe multiplicatif $\mathbf{F}_q^*$.

Le théorème suivant a été annoncé au Chapitre I (lemme 7.5) :

Théorème 2.7.
Le groupe multiplicatif $\mathbf{F}_q^*$ est un groupe cyclique (donc isomorphe à $\mathbf{Z}/(q-1)\mathbf{Z}$).

Démonstration. (Voir aussi [S1] Chapitre I § 1). On montre d'abord un lemme :

Lemme 2.8.
Soit $n \in N^*$, on a la formule $n = \sum_{d|n} \varphi(d)$.

Démonstration (de 2.8) En effet, dans $\mathbf{Z}/n\mathbf{Z}$, tout élément a pour ordre un diviseur d de n et il y a exactement $\varphi(d)$ éléments d'ordre d puisqu'ils engendrent l'unique sous-groupe (cyclique) d'ordre d de $\mathbf{Z}/n\mathbf{Z}$ (cf. I, 7.1 et Exercice I A.2).

Revenant à $\mathbf{F}_q^*$, posons $n = q-1$ et soit d un diviseur de n. S'il existe un $x \in \mathbf{F}_q^*$ d'ordre d, on considère le sous-groupe $H = < x > \simeq \mathbf{Z}/d\mathbf{Z}$ de $\mathbf{F}_q^*$. On a $|H| = d$ et pour tout $y \in H$, $y^d = 1$. Comme le polynôme $Y^d - 1$ a au plus d racines dans $\mathbf{F}_q$, tout élément d'ordre d de $\mathbf{F}_q^*$ est donc dans H. Il en résulte (cf. I, 7.1) que le nombre $N(d)$ d'éléments d'ordre d de $\mathbf{F}_q^*$ vaut 0 ou $\varphi(d)$, donc, en tous cas, on a $N(d) \leq \varphi(d)$. Mais tout élément de $\mathbf{F}_q^*$ a pour ordre un diviseur de n ce qui implique $n = |\mathbf{F}_q^*| = \sum_{d|n} N(d)$ et en comparant avec l'égalité du lemme 2.8, on voit qu'on a nécessairement $N(d) = \varphi(d)$ pour tout d.
En particulier, on a $N(n) = \varphi(n) > 0$, donc $\mathbf{F}_q^*$ contient un élément d'ordre n, donc est cyclique.

Remarques 2.9.

1) On ne sait pas, en général, trouver explicitement un générateur de $\mathbf{F}_q^*$ (cf. § 2, 3 exercice 3).

2) La même démonstration prouve que tout sous-groupe fini du groupe multiplicatif d'un corps ([3]) est cyclique.

d) Les carrés de $\mathbf{F}_q$.

La connaissance des carrés de $\mathbf{F}_q$ va nous permettre de terminer la démonstration du théorème des deux carrés (cf. Chapitre II § 6). Plus généralement elle permet d'aborder l'étude des irréductibles des anneaux du type $\mathbf{Z}[i\sqrt{d}]$.

Le nombre $q = p^n$ est toujours une puissance du nombre premier p. On pose $\mathbf{F}_q^2 = \{x \in \mathbf{F}_q \mid \exists y \in \mathbf{F}_q,\ x = y^2\}$ et $\mathbf{F}_q^{*2} = \mathbf{F}_q^2 \cap \mathbf{F}_q^*$.

Proposition 2.10.
1) Pour $p = 2$, on a $\mathbf{F}_q^2 = \mathbf{F}_q$.
2) Pour $p > 2$, on a $|\mathbf{F}_q^2| = \dfrac{q+1}{2}$ et $|\mathbf{F}_q^{*2}| = \dfrac{q-1}{2}$.

Démonstration.
Le point 1) résulte de 2.4 (Frobenius).
Si on a $p > 2$ et donc $1 \neq -1$ dans $\mathbf{F}_p$, on a la suite exacte :

$$1 \longrightarrow \{1,-1\} \longrightarrow \mathbf{F}_q^* \longrightarrow \mathbf{F}_q^{*2} \longrightarrow 1$$
$$x \longmapsto x^2$$

qui donne $|\mathbf{F}_q^{*2}| = \dfrac{q-1}{2}$. L'autre égalité résulte de la formule $\mathbf{F}_q^2 = \mathbf{F}_q^{*2} \cup \{0\}$.

([3]) Il s'agit d'un corps commutatif, bien entendu, l'assertion serait fausse sinon, penser au groupe des quaternions contenu dans le corps du même nom.

Proposition 2.11 : (caractérisation des carrés).
On suppose $p > 2$. Alors on a :

$$x \in \mathbf{F}_q^{*2} \Longleftrightarrow x^{\frac{q-1}{2}} = 1.$$

Démonstration. Posons $X = \{x \in \mathbf{F}_q \,|\, x^{\frac{q-1}{2}} = 1\}$. On a $|X| \leq \dfrac{q-1}{2}$ (car un polynôme de degré $\dfrac{q-1}{2}$ a au plus $\dfrac{q-1}{2}$ racines). D'autre part, si x est dans $\mathbf{F}_q^{*2}$ on a $x = y^2$, donc $x^{\frac{q-1}{2}} = y^{q-1} = 1$ et donc on a $\mathbf{F}_q^{*2} \subset X$. Pour une raison de cardinal, il en résulte $X = \mathbf{F}_q^{*2}$.

Remarques 2.12.

1) On a la suite exacte

$$\begin{array}{ccccccc} 1 & \longrightarrow & \mathbf{F}_q^{*2} & \longrightarrow & \mathbf{F}_q^{*} & \longrightarrow & \{1,-1\} \longrightarrow 1 \\ & & & & x & \longmapsto & x^{\frac{q-1}{2}} \end{array}$$

En effet, si x est dans $\mathbf{F}_q^*$, $y = x^{\frac{q-1}{2}}$ vérifie $y^2 = 1$ donc y vaut ∓ 1, le reste est clair avec 2.11.

2) La proposition 2.11 fournit un critère pour savoir si un élément de $\mathbf{F}_q$ est un carré. Par exemple, si $q = 7$, donc $\mathbf{F}_q = \mathbf{Z}/7\mathbf{Z}$, on a $\dfrac{q-1}{2} = 3$ et $2^3 = 8 = 1 \pmod 7$, de sorte que 2 est un carré, $3^3 = 27 = -1 \pmod 7$, et 3 n'est pas un carré.
En particulier, on peut prouver le résultat annoncé au Chapitre II §6 au cours de la démonstration du Théorème 6.6 :

Corollaire 2.13.
Soit p un nombre premier, $p > 2$ et posons $q = p^n$, $n \in \mathbf{N}^*$. Alors, -1 est un carré dans $\mathbf{F}_q$ si et seulement si q est congru à 1 modulo 4.

Démonstration. C'est une conséquence immédiate de 2.11 :

$$-1 \in \mathbf{F}_q^{*2} \Longleftrightarrow (-1)^{\frac{q-1}{2}} = 1 \Longleftrightarrow \frac{q-1}{2} \text{ pair } \Longleftrightarrow q \equiv 1 \pmod 4.$$

Voici une autre démonstration de ce corollaire :
Dire que $q \equiv 1 \pmod 4$ revient à dire que le cardinal de $\mathbf{F}_q^{*2}$, i.e. $\dfrac{q-1}{2}$ est pair. Mais, d'après le Théorème de Sylow (cf. I, 5.4), ceci revient à dire qu'il y a un élément d'ordre 2 dans $\mathbf{F}_q^{*2}$. Un tel élément est un x tel que $x^2 = 1$, $x \neq 1$, c'est donc nécessairement -1 et on a donc bien $q \equiv 1 \pmod 4 \Longleftrightarrow -1 \in \mathbf{F}_q^{*2}$.

Remarque 2.14. En fait on utilise beaucoup moins que Sylow, voici l'ingrédient exact :

Lemme 2.15.
Soit G un groupe fini, $X = \{g \in G \,|\, g^2 = 1\}$, alors on a $|G| \equiv |X| \pmod 2$. En particulier si $|G|$ est pair, G contient des éléments d'ordre exactement 2, (car 1 est dans X).

Le lemme est évident, il suffit de remarquer que $g^2 = 1 \Longleftrightarrow g = g^{-1}$ et de grouper les éléments de $G - X$ par paires $\{g, g^{-1}\}$ (cf. aussi I, 4.16).

Application 2.16.
Il existe une infinité de nombres premiers de la forme $4m + 1$.

Démonstration. Soit n un entier et p un facteur premier de $(n!)^2 + 1$. On a $p > n$ (sinon p diviserait $n!$). D'autre part dans $\mathbf{F}_p = \mathbf{Z}/p\mathbf{Z}$, on a $(n!)^2 + 1 = 0$, donc $-1 = (n!)^2$ est un carré, et donc (cf. 2.13) on a $p \equiv 1 \pmod 4$. On a donc de tels nombres premiers arbitrairement grands (cf. aussi Exercice 4.14).

3. Irréductibilité des polynômes de $k[X]$.

Rappelons que l'anneau $k[X]$, lorsque k est un corps, est principal, donc factoriel. On s'intéresse aux éléments irréductibles de cet anneau.
Remarquons au passage que si A est factoriel et si $K = \mathrm{Fr}\,(A)$ la connaissance des éléments irréductibles de $A[X]$ passe par celle des irréductibles de $K[X]$ (cf. II, 4.4), $P(X)$ est irréductible sur A si et seulement si il l'est sur K et si son contenu est 1 (cf. aussi § 2, 3 Exercice 11.e).

Remarques 3.1. (remarques préliminaires)
1) Les polynômes $X - a$ sont irréductibles, pour tout $a \in k$.
2) Si $P \in k[X]$ est irréductible et $d^\circ P > 1$, P n'a pas de racines dans k.
En particulier si k est algébriquement clos, les $X - a$, $a \in k$ sont les seuls irréductibles.
3) La réciproque de 2) est inexacte : le polynôme $P(X) = (X^2 + 1)^2$ n'a pas de racines dans $\mathbf{R}$, mais est réductible. Elle est vraie toutefois si on a $d^\circ P \leq 3$.
4) Pour $k = \mathbf{R}$ les polynômes irréductibles sont
a) les polynômes $X - a$ avec $a \in \mathbf{R}$,
b) les polynômes de degré 2 sans racine réelle.
5) Bien entendu, l'irréductibilité d'un polynôme de $k[X]$ ne subsiste pas en général dans une extension de k (cf. § 1.c, le corps de rupture est fait pour cela).

Nous donnons d'abord quelques critères classiques d'irréductibilité :

Théorème 3.2 : (critère d'Eisenstein).
Soit A un anneau factoriel et soit $K = \mathrm{Fr}\,(A)$. Soit $P(X) = a_n X^n + \cdots + a_0$ avec $a_i \in A$. Soit $p \in A$ un élément irréductible. On suppose :
1) $p \nmid a_n$,
2) $\forall i = 0, \ldots, n-1$, $p \mid a_i$,
3) $p^2 \nmid a_0$.
Alors P est irréductible dans $K[X]$ (donc aussi dans $A[X]$, cf. II 4.4, pourvu que l'on ait $\mathrm{pgcd}\,(a_i) = 1$*).*

Démonstration. Si P n'est pas irréductible dans $K[X]$, on a $P(X) = Q(X)R(X)$ avec $Q, R \in A[X]$ et $d^\circ Q$, $d^\circ R < d^\circ P$ (cf. Chapitre II § 4 a)). Posons alors $Q(X) = b_q X^q + \cdots + b_0$, $R(X) = c_r X^r + \cdots + c_0$, avec $b_i, c_i \in A$ et $0 < q, r < n$. Comme A est factoriel et p irréductible, l'idéal (p) est premier (c'est le lemme d'Euclide, cf. II, 3.19), donc $B = A/(p)$ est intègre. Projetons l'égalité $P = QR$ dans $B[X]$. En désignant par $\overline{u}$ l'image de $u \in A$ dans B, on a :

$$\overline{P}(X) = \overline{a}_n X^n = (\overline{b}_q X^q + \cdots + \overline{b}_0)(\overline{c}_r X^r + \cdots + \overline{c}_0)$$

en vertu des hypothèses sur les a_i, avec $\overline{a}_n \neq 0$, donc aussi $\overline{b}_q, \overline{c}_r \neq 0$. Mais cette égalité est encore vraie dans $L[X]$, où L désigne le corps des fractions de

B. Comme $L[X]$ est principal et X irréductible, l'unicité de la décomposition en facteurs irréductibles dans $L[X]$ montre que X divise $\overline{Q}$ et $\overline{R}$, donc que l'on a $\overline{b}_0 = \overline{c}_0 = 0$ dans B, mais alors p^2 divise $a_0 = b_0 c_0$ contrairement à l'hypothèse.

Remarque 3.3. Attention, a priori $B[X]$ n'est pas factoriel car B ne l'est pas (cf. Chapitre II § 4 b)).

Applications 3.4.

1) Si p est un nombre premier de $\mathbf{N}$, le polynôme $X^{p-1} + \cdots + X + 1$ est irréductible sur $\mathbf{Z}$ (poser $X = Y + 1$ et appliquer Eisenstein avec p).

2) Soit $a \in \mathbf{Z}$, $a = p_1^{\alpha_1} \ldots p_r^{\alpha_r}$. On suppose que l'un des α_i est égal à 1. Alors $X^n - a$ est irréductible sur $\mathbf{Z}$.

3) Soit $\lambda \in k$, $\lambda \neq 0, 1$. Le polynôme $Y^2 - X(X-1)(X-\lambda)$ est irréductible (il définit une courbe "elliptique" du plan, cf. [P]).

Théorème 3.5 (réduction).
Soient A un anneau factoriel et $K = \mathrm{Fr}\, A$. Soit I un idéal premier de A et $B = A/I$ qui est un anneau intègre de corps de fractions L. Soit $P(X) = a_n X^n + \cdots + a_0$ un polynôme de $A[X]$ et $\overline{P}$ sa réduction modulo I. On suppose $\overline{a}_n \neq 0$ dans B. Alors, si $\overline{P}$ est irréductible sur B ou L, le polynôme P est irréductible sur K.

Démonstration. Supposons qu'on ait $P(X) = Q(X)R(X)$ dans $A[X]$, avec $Q(X) = b_q X^q + \cdots + b_0$ et $R(X) = c_r X_r + \cdots + c_0$. On a $\overline{P} = \overline{QR}$, d'où $\overline{a}_n = \overline{b}_q \overline{c}_r$ et donc $\overline{b}_q, \overline{c}_r \neq 0$.
Comme P est irréductible dans $L[X]$ (ou $B[X]$), l'un des polynômes $\overline{Q}$ ou $\overline{R}$ est de degré 0, disons $\overline{Q}$, donc Q est aussi de degré 0 et P est irréductible dans $K[X]$.

Remarques 3.6.

1) Attention $P(X)$ n'est pas nécessairement irréductible dans $A[X]$, exemple le polynôme $2X \in \mathbf{Z}[X]$ avec $I = (3)$.

2) Notons que P irréductible sur B implique $\overline{P}$ irréductible sur L lorsque B est factoriel, (cf. aussi § 2, 3 Exercice 11).

Applications 3.7.

1) $X^2 + Y^2 + 1$ est irréductible dans $\mathbf{R}[X, Y]$ (faire $Y = 0$, i.e. passer au quotient par l'idéal (Y)).

2) Le cas le plus fréquent d'utilisation de 3.5 est le cas $A = \mathbf{Z}$, $I = (p)$ avec p premier, $B = \mathbf{F}_p$ est alors un corps. Ainsi, le polynôme $X^3 + 462X^2 + 2433X - 67691$ est irréductible sur $\mathbf{Z}$. En effet, on le réduit modulo 2, il reste $X^3 + X + 1$ qui est irréductible sur $\mathbf{F}_2$ (sinon il aurait une racine dans $\mathbf{F}_2$).

3) Soit p un nombre premier, alors $X^p - X - 1$ est irréductible sur $\mathbf{Z}$.
Cela résulte en effet du lemme suivant :

Lemme 3.8.
Le polynôme $X^p - X - 1$ est irréductible sur $\mathbf{F}_p$.

Démonstration. Soit k un corps de décomposition de $X^p - X - 1 = P(X)$ sur $\mathbf{F}_p$ et soit $\alpha \in k$ une racine de P. Alors, pour $i = 0, 1, \ldots, p-1$, $\alpha + i$ est aussi une racine de P car on a, par Frobenius :

$$(\alpha + i)^p - (\alpha + i) - 1 = (\alpha^p - \alpha - 1) + i^p - i$$

et, comme i est dans $\mathbf{F}_p$, on a $i^p - i = 0$. Les racines de P sont donc exactement $\alpha, \alpha + 1, \ldots, \alpha + p - 1$.

Supposons P réductible sur $\mathbf{F}_p$, $P = QR$ avec $Q, R \in \mathbf{F}_p[X]$; $d^\circ Q$, $d^\circ R < p$. On a, dans $k[X]$:

$$Q(X) = \prod_{k=1}^{d} (X - \alpha - i_k), \text{ où } d = d^\circ Q.$$

Le coefficient du terme en X^{d-1} de Q est alors $-(d\alpha + i_1 + \cdots + i_d)$ qui est donc dans $\mathbf{F}_p$. On en déduit : $d\alpha \in \mathbf{F}_p$. Mais, comme on a $d < p$, d est dans $\mathbf{F}_p^*$, donc α est dans $\mathbf{F}_p$. Mais alors, on a $\alpha^p = \alpha$ et ceci contredit $\alpha^p - \alpha - 1 = 0$.

Pour progresser dans cette voie nous donnons un ***troisième critère***, souvent commode dans le cas des corps finis.

Théorème 3.9.
Soit $P \in k[X]$, de degré $n > 0$. Alors, P est irréductible sur k si et seulement si P n'a pas de racines dans les extensions K de k qui vérifient $[K : k] \leq n/2$.

Démonstration. Si P est irréductible et si $x \in K$ est racine de P, le corps $k[x]$ est un corps de rupture de P, donc il est de degré n et on a $[K : k] \geq n$.
Réciproquement, si P n'est pas irréductible, on a $P = QR$ et $d^\circ Q$ ou $d^\circ R \leq n/2$. Si, par exemple, on a $d^\circ Q \leq n/2$ et si Q' est un facteur irréductible de Q, P aura une racine dans un corps de rupture de Q', de degré $\leq n/2$.

Exemples 3.10.

1) Le polynôme $X^4 + X + 1$ est irréductible sur $\mathbf{F}_2$. Il suffit de vérifier qu'il n'a pas de racines dans $\mathbf{F}_2$ ni $\mathbf{F}_4$. Pour $\mathbf{F}_2$ c'est clair, pour $\mathbf{F}_4$ on note que l'on a $\mathbf{F}_4 = \mathbf{F}_2[j]$ avec $j^2 + j + 1 = 0$. Si $x \in \mathbf{F}_4 - \mathbf{F}_2$, on a $x = j$ ou $x = j + 1 = -j^2$, donc $x^3 = 1$, et $x^4 + x + 1 = 2x + 1 = 1 \neq 0$.

2) On déduit par exemple de 1), par réduction modulo 2, que $X^4 + 8X^2 + 17X - 1$ est irréductible sur $\mathbf{Z}$.

3) La méthode de réduction, si elle est très efficace, a pourtant des limites, par exemple :

Proposition 3.11.
Le polynôme $X^4 + 1$ est irréductible sur $\mathbf{Z}$ (donc sur $\mathbf{Q}$) mais est réductible sur $\mathbf{F}_p$ pour tout nombre premier p.

Démonstration.
Pour l'irréductibilité sur $\mathbf{Z}$, il y a de nombreuses méthodes. On peut noter, cf. § 4, que $X^4 + 1 = \Phi_8$ est un polynôme cyclotomique, donc irréductible, cf. 4.10. On peut aussi poser $X = 1 + Y$ et utiliser Eisenstein. On peut enfin procéder par identification.

Regardons maintenant modulo p. Pour $p = 2$, on a $X^4 + 1 = (X + 1)^4$. On suppose donc $p > 2$. On a $X^8 - 1 = (X^4 + 1)(X^4 - 1)$, de sorte que, si $X^4 + 1$ a une racine x dans un corps K, on a $x^8 = 1$, avec $x^4 \neq 1$, i.e. x est un élément d'ordre 8 de K^*. Pour prouver que $X^4 + 1$ est réductible sur $\mathbf{F}_p$, il suffit, d'après 3.9 de prouver qu'il a une racine dans $\mathbf{F}_{p^2}$ donc de montrer que $\mathbf{F}_{p^2}^*$ contient un élément d'ordre 8. Pour ceci, comme $\mathbf{F}_{p^2}^*$ est cyclique d'ordre $p^2 - 1$ (cf. 2.7) il suffit d'établir le lemme suivant :

Lemme 3.12.
Soit p un nombre premier > 2. Alors, $p^2 - 1$ est multiple de 8.

En effet on a $p^2 - 1 = (p-1)(p+1)$ et $p-1$ et $p+1$ sont deux nombres pairs consécutifs, donc l'un des deux est multiple de 4, cqfd.

Nous étudierons systématiquement au §4 le comportement des polynômes cyclotomiques sur $\mathbf{F}_p$.

L'exemple précédent montre qu'on ne peut espérer que la méthode de réduction aboutisse toujours à prouver l'irréductibilité d'un polynôme sur $\mathbf{Z}$. Cependant, même si cette méthode échoue en apparence (c'est-à-dire si $P \in \mathbf{Z}[X]$ est réductible sur tout $\mathbf{F}_p$), elle peut parfois, par une analyse plus fine des décompositions de $\overline{P}$, mener au résultat :

Exemples 3.13.
1) Le polynôme $X^5 + X^2 + X + 2 = P(X)$ est irréductible sur $\mathbf{Z}$.
On regarde sur $\mathbf{F}_2$, on a $\overline{P} = X(X^4 + X + 1)$ et on a vu (3.10) que $X^4 + X + 1$ est irréductible sur $\mathbf{F}_2$. Si P est réductible sur $\mathbf{Z}$, on a nécessairement $P = QR$ avec $d^\circ Q = 1$, $d^\circ R = 4$, i.e. P a une racine dans $\mathbf{Z}$ (sinon on aurait $P = QR$ avec $d^\circ Q = 2$, $d^\circ R = 3$ et par projection sur $\mathbf{F}_2$ ceci contredirait l'irréductibilité de $X^4 + X + 1$). Mais si $P(a) = 0$ avec $a \in \mathbf{Z}$, $\overline{a}$ est racine de $\overline{P}$ dans tous les $\mathbf{F}_p$ et on constate aisément que $\overline{P}$ n'a pas de racines dans $\mathbf{F}_3$. (On peut aussi remarquer que la racine a divise nécessairement le coefficient de degré 0 de P, ici 2, ce qui ramène le problème à un nombre fini de calculs, cf. [VdW] §32).
2) $P(X) = X^5 - 7$.
Modulo 2, on a $P(X) = X^5 + 1 = (X+1)(X^4 + X^3 + X^2 + X + 1)$ et le deuxième facteur est irréductible ; mais modulo 11, P n'a pas de racines ($\mathbf{F}_{11}^*$ est d'ordre 10, donc si $x \in \mathbf{F}_{11}^*$, on a $(x^5)^2 = 1$ donc $x^5 = \mp 1 \neq 7$), donc P est irréductible sur $\mathbf{Z}$.

Donnons pour finir un critère de ***conservation de l'irréductibilité par extension de corps***.

Théorème 3.14.
Soit $P \in k[X]$ un polynôme irréductible de degré n et soit K une extension de degré m avec $(m, n) = 1$. Alors P est encore irréductible sur K.

Démonstration. Supposons $P = QR$ avec $Q, R \in K[X]$, Q irréductible de degré q avec $0 < q < n$. Soit $L = K[X]/(Q)$ un corps de rupture de Q qui s'écrit $L = K(x)$ avec $Q(x) = 0$. On utilise alors la multiplicativité des degrés (1.5) : on a $[L : k] = [L : K][K : k] = qm = [L : k(x)]\,[k(x) : k] = rn$ (on a $[k(x) : k] = n$ puisque $k(x)$ est un corps de rupture de P sur k). Mais alors, comme n et m sont premiers entre eux, n diviserait q par le théorème de Gauss, et c'est absurde.

Remarque 3.15. Sans l'hypothèse $(m, n) = 1$, le théorème est faux : $X^4 + 1$ qui est irréductible sur $\mathbf{Q}$, ne l'est plus sur $\mathbf{Q}(i)$: on a $X^4 + 1 = (X^2 + i)(X^2 - i)$.

Exemples 3.16.
1) $X^3 + X + 1$ est irréductible sur $\mathbf{Q}(i)$ comme sur $\mathbf{Q}$.
2) Si $X^3 + X + 1$ n'a pas de racines dans $\mathbf{F}_p$, par exemple, si $p = 2$ ou $5 \ldots$, il est irréductible sur $\mathbf{F}_p$, donc aussi sur $\mathbf{F}_{p^n}$ si 3 ne divise pas n. Mais il est réductible dans $\mathbf{F}_{p^{3k}}$ pour tout $k \in \mathbf{N}^*$.

4. Cyclotomie.

a) Racines de l'unité, racines primitives.

Soient k un corps et $n \in \mathbf{N}^*$, on considère le polynôme $P_n(X) = X^n - 1$.

Remarque 4.1. La dérivée de P_n est $P_n'(X) = nX^{n-1}$. Si la caractéristique p de k ne divise pas n, la seule racine de P_n' est 0, qui n'annule pas P_n, donc P_n n'a que des racines simples.
Si p divise n, on a $n = mp$ et $X^n - 1 = (X^m - 1)^p$ (par Frobenius) donc P_n a des racines multiples dans tout corps de décomposition.

Dans toute la suite, on supposera toujours n premier à la caractéristique de k.

L'ensemble des racines n-èmes de l'unité dans k sera noté $\mu_n(k)$:

$$\mu_n(k) = \{\zeta \in k \mid \zeta^n = 1\}.$$

C'est un sous-groupe de k^*, de cardinal $\leq n$, donc cyclique (cf. 2.9.2).

Désignons par $K_n = D_k(P_n)$, corps de décomposition de P_n sur k. Alors, on a $|\mu_n(K_n)| = n$ et $\mu_n(K_n) \simeq \mathbf{Z}/n\mathbf{Z}$. De plus, comme $\mu_n(k)$ est inclus dans $\mu_n(K_n)$, on a $\mu_n(k) \simeq \mathbf{Z}/d\mathbf{Z}$ pour un d, diviseur de n.

Nous étudions maintenant le groupe $\mu_n(K_n)$.

Définition 4.2.
Une racine n-ème ***primitive*** *de 1 est un élément ζ de K_n tel que $\zeta^n = 1$ et $\zeta^d \neq 1$ pour $d < n$. Autrement dit, ζ est un générateur du groupe $\mu_n(K_n)$, de sorte qu'il y a $\varphi(n)$ racines primitives de 1 (cf. Chapitre I § 7). Leur ensemble sera noté $\mu_n^*(K_n)$.*

Définition 4.3.
Le n-ème ***polynôme cyclotomique*** *$\Phi_{n,k} \in K_n[X]$ est donné par la formule :*

$$\Phi_{n,k}(X) = \prod_{\zeta \in \mu_n^*(K_n)} (X - \zeta).$$

Remarques 4.4.

0) Si ζ est une racine n-ème primitive de l'unité, les autres sont les ζ^m avec $(m, n) = 1$.

1) Le polynôme $\Phi_{n,k}$ est unitaire, de degré $\varphi(n)$.

2) Soit k_0 le corps premier de k ($k_0 = \mathbf{Q}$ ou $\mathbf{F}_p$) et soit $K_0 = D_{k_0}(X^n - 1)$. Alors $\Phi_{n,k}$ est dans $K_0[X]$, de sorte qu'on peut toujours supposer $k = k_0$ (en fait, nous montrerons que $\Phi_{n,k}$ est même dans $k_0[X]$).

b) Etude de $\Phi_{n,k}$.

Proposition 4.5.
On a la formule (en omettant l'indice k) :

$$X^n - 1 = \prod_{d \mid n} \Phi_d(X).$$

Démonstration. Cela résulte de l'égalité : $\mu_n(K_n) = \bigcup_{d \mid n} \mu_d^*(K_n)$ (l'union est ici disjointe) qui exprime que si ζ est une racine n-ème de 1, l'ordre de ζ est un diviseur de n.

Remarque 4.6. En comparant les degrés, on retrouve la formule vue en 2.8 :

$$n = \sum_{d \mid n} \varphi(d).$$

Exemples 4.7.
On a $\Phi_1(X) = X - 1$, puis, comme $X^2 - 1 = \Phi_1(X)\Phi_2(X) = (X-1)(X+1)$, on a $\Phi_2(X) = X + 1$.
La formule 4.5 permet le calcul des Φ_n par récurrence en écrivant :

$$\Phi_n(X) = \frac{X^n - 1}{\prod\limits_{\substack{d \mid n \\ d \neq n}} \Phi_d(X)}.$$

On trouve ainsi successivement :
$\Phi_3 = X^2 + X + 1$, $\Phi_4 = X^2 + 1$, $\Phi_5 = X^4 + X^3 + X^2 + X + 1$, $\Phi_6 = X^2 - X + 1$, $\Phi_8 = X^4 + 1$, $\Phi_9 = X^6 + X^3 + 1 \ldots$

Proposition 4.8.
1) On a $\Phi_{n,\mathbf{Q}}(X) \in \mathbf{Z}[X]$.
2) Soit k *un corps quelconque et* $\sigma : \mathbf{Z} \longrightarrow k$ *l'homomorphisme canonique (cf. § 2.a). On a alors :*

$$\Phi_{n,k}(X) = \sigma(\Phi_{n,\mathbf{Q}}(X)).$$

En particulier $\Phi_{n,\mathbf{F}_p}$ *s'obtient à partir de* $\Phi_{n,\mathbf{Q}}$ *par réduction modulo* p.

Démonstration.

1) On raisonne par récurrence sur n. On a $\Phi_1(X) = X - 1 \in \mathbf{Z}[X]$. Supposons la propriété établie pour $d < n$. Soit $F(X) = \prod\limits_{\substack{d \mid n \\ d \neq n}} \Phi_d(X)$, on a $F(X) \in \mathbf{Z}[X]$ et F unitaire. On effectue alors la division euclidienne de $X^n - 1$ par $F(X)$ dans $\mathbf{Z}[X]$ (cf. II, 3.31) :

$$X^n - 1 = F(X)P(X) + R(X), \quad \text{avec} \quad P, R \in \mathbf{Z}[X] \quad \text{et} \quad d^\circ R < d^\circ F.$$

Mais on a $X^n - 1 = \Phi_n(X)F(X)$ dans $\mathbf{Q}[X]$, donc $F(X)(\Phi_n(X) - P(X)) = R(X)$ et, pour une raison de degré, on a nécessairement $\Phi_n = P \in \mathbf{Z}[X]$.

2) On raisonne par récurrence, le cas $n = 1$ étant trivial. Dans $\mathbf{Z}[X]$, on a :

$$X^n - 1 = \prod_{d \mid n} \Phi_{d,\mathbf{Q}}(X) = \Phi_{n,\mathbf{Q}}(X)F(X).$$

Comme σ est un homomorphisme, on a, dans $K_n[X]$, (avec $K_n = D_k(X^n - 1)$) : $X^n - 1 = \sigma(X^n - 1) = \sigma(\Phi_{n,\mathbf{Q}})\sigma(F)$. Mais, par l'hypothèse de récurrence, on a

$$\sigma(F) = \prod_{\substack{d \mid n \\ d \neq n}} \sigma(\Phi_{d,\mathbf{Q}}) = \prod_{\substack{d \mid n \\ d \neq n}} \Phi_{d,k},$$

et comme on a, par définition, $X^n - 1 = \prod\limits_{d \mid n} \Phi_{d,k}$, il en résulte, puisque $k[X]$ est intègre, qu'on a bien $\Phi_{n,k} = \sigma(\Phi_{n,\mathbf{Q}})$.

c) Application, le théorème de Wedderburn.

Théorème 4.9 (Wedderburn).
Tout corps fini est commutatif.

Démonstration.

1) Soit k un corps fini, pas nécessairement commutatif, Z le centre de k :

$$Z = \{a \in k \mid \forall x \in k,\ ax = xa\},$$

Z est un sous-corps commutatif de k de cardinal $q \geq 2$ et comme k est un Z-espace vectoriel, on a $|k| = q^n$.

2) Supposons k non commutatif, donc $n > 1$. Alors k^* opère sur lui-même par automorphismes intérieurs. Pour $x \in k^*$ on note $\omega(x)$ l'orbite de x. On pose par ailleurs : $k_x = \{y \in k \mid yx = xy\}$, k_x est un sous-corps de k (pas nécessairement commutatif) et le stabilisateur de x dans l'opération est k_x^*.
On a $|k_x| = q^d$ pour la même raison que ci-dessus. De plus, d est un diviseur de n (remarquer, soit que k est un k_x-espace vectoriel à gauche, donc $|k|$ une puissance de $|k_x|$, soit, si l'on veut éviter le recours aux espaces vectoriels sur les corps non commutatifs, que l'inclusion $k_x^* \subset k^*$ entraîne $q^d - 1 \mid q^n - 1$ et que, pour $q \in \mathbf{N}$, $q \geq 2$, ceci impose $d \mid n$).
Le cardinal de l'orbite de x est alors :

$$|\omega(x)| = \frac{|k^*|}{|k_x^*|} = \frac{q^n - 1}{q^d - 1}.$$

3) On a, dans $\mathbf{Z}$, par définition des polynômes cyclotomiques :

$q^n - 1 = \prod_{m \mid n} \Phi_m(q)$ et de même, $q^d - 1 = \prod_{m \mid d} \Phi_m(q)$, donc $\dfrac{q^n - 1}{q^d - 1} = \prod_{\substack{m \mid n \\ m \nmid d}} \Phi_m(q)$.

Pour $d \neq n$, on voit donc en particulier que $\Phi_n(q)$ divise $\dfrac{q^n - 1}{q^d - 1}$.

4) On écrit alors l'équation des classes :

$$|k^*| = |Z^*| + \sum_{x \notin Z} |\omega(x)|,$$

et dire que x n'est pas dans Z signifie que l'on a $d \neq n$, de sorte qu'on a :

$$q^n - 1 = q - 1 + \sum \frac{q^n - 1}{q^d - 1},$$

la somme portant sur un certain nombre de diviseurs stricts de n. Il en résulte que $\Phi_n(q)$ divise $q - 1$, en particulier on a $|\Phi_n(q)| \leq q - 1$.

5) On a $\Phi_n(q) = (q - \zeta_1) \ldots (q - \zeta_l)$ où $\zeta_1, \ldots, \zeta_l \in \mathbf{C}$ sont les racines primitives n-èmes de 1 et vérifient donc $|\zeta_i| = 1$ et $\zeta_i \neq 1$, puisque $n \neq 1$. Mais alors, on a pour tout i, $|q - \zeta_i| > q - 1$ (un dessin permet de mieux s'en convaincre) et donc $|\Phi_n(q)| > (q - 1)^l \geq q - 1$ et c'est une contradiction.

d) L'irréductibilité de Φ_n sur $\mathbf{Z}$.

Théorème 4.10.
Le polynôme cyclotomique $\Phi_n(X)$ (qui est dans $\mathbf{Z}[X]$, cf. 4.8.1) est irréductible sur $\mathbf{Z}$, donc sur $\mathbf{Q}$.

Corollaire 4.11.
Si ζ est une racine primitive n-ème de 1 dans un corps de caractéristique nulle, son polynôme minimal sur $\mathbf{Q}$ est Φ_n, et donc on a $[\mathbf{Q}(\zeta) : \mathbf{Q}] = \varphi(n)$.

Démonstration. Soit K un corps de décomposition de Φ_n sur $\mathbf{Q}$, $\zeta \in K$ une racine n-ème primitive de 1. Soit p un nombre premier, ne divisant pas n.

1) ζ^p est une autre racine primitive de 1. En effet, cf. 4.4.0, les générateurs de $\mu_n(K)$, c'est-à-dire les racines primitives, sont les ζ^m, avec $(m,n) = 1$.

2) Soit $f \in \mathbf{Q}[X]$ (resp. g) le polynôme minimal de ζ (resp. ζ^p) sur $\mathbf{Q}$. Alors on a $f, g \in \mathbf{Z}[X]$. En effet, comme $\mathbf{Z}[X]$ est factoriel (cf. Chapitre II, 4.2) on a $\Phi_n(X) = f_1(X)^{\alpha_1} \dots f_r(X)^{\alpha_r}$ avec $f_i \in \mathbf{Z}[X]$ irréductible. Comme Φ_n est unitaire, il en est de même des f_i (quitte à multiplier f_i par -1). Mais alors, ζ est racine de l'un des f_i et f_i qui est unitaire et irréductible sur $\mathbf{Z}$, donc sur $\mathbf{Q}$, n'est autre que f ; de même pour g.
Notons au passage que f et g divisent Φ_n, dans $\mathbf{Z}[X]$.

3) Nous allons montrer que l'on a $f = g$.
Sinon, comme f, g sont irréductibles et distincts, le produit $f.g$ divise Φ_n dans $\mathbf{Z}[X]$. Par ailleurs, comme $g(\zeta^p) = 0$, ζ est racine du polynôme $g(X^p)$, donc $f(X)$ divise $g(X^p)$, a priori dans $\mathbf{Q}[X]$, mais aussi dans $\mathbf{Z}[X]$:

$$g(X^p) = f(X)h(X) \text{ avec } h \in \mathbf{Z}[X].$$

(Si h est dans $\mathbf{Q}[X]$, on l'écrit $h = \frac{a}{b}h'$ de la manière standard et on utilise le Lemme de Gauss II 4.3).
Projetons alors cette égalité dans $\mathbf{F}_p$, on écrit :

$$g(X) = a_r X^r + \dots + a_0 \quad \text{avec} \quad a_i \in \mathbf{Z}, \quad \text{d'où} \quad g(X^p) = a_r X^{pr} + \dots + a_1 X^p + a_0$$

mais, modulo p, on a $\overline{a}_i = \overline{a}_i^p$, donc (par Frobenius)

$$\overline{g}(X^p) = (\overline{a}_r X^r + \dots + \overline{a}_0)^p = \overline{g}(X)^p.$$

Soit alors $\varphi(X)$ un facteur irréductible de $\overline{f}(X)$ sur $\mathbf{F}_p$. On a $\overline{g}(X)^p = \overline{f}(X)\overline{h}(X)$, donc, par le lemme d'Euclide, φ divise $\overline{g}$.
Comme fg divise Φ_n sur $\mathbf{Z}$, $\overline{f}\,\overline{g}$ divise $\overline{\Phi}_n$ sur $\mathbf{F}_p$, donc φ^2 divise $\overline{\Phi}_n = \Phi_{n,\mathbf{F}_p}$ (cf. 4.8.2). Mais alors dans un corps de décomposition de Φ_n sur $\mathbf{F}_p$, $\overline{\Phi}_n$ aura une racine double, ce qui contredit 4.1 : comme la caractéristique p du corps $\mathbf{F}_p$ ne divise pas n par hypothèse, $\Phi_{n,\mathbf{F}_p}$ n'a pas de racine multiple.

4) Si ζ' est une racine primitive m-ème on a $\zeta' = \zeta^m$ avec $m = p_1^{\alpha_1} \dots p_r^{\alpha_r}$ et $p_i \nmid n$. Il résulte alors de 3) par une récurrence immédiate que ζ' et ζ ont même polynôme minimal sur $\mathbf{Q}$, donc on a $f(\zeta') = 0$ de sorte que f admet toutes les racines primitives de l'unité comme zéros. On a donc $d^\circ f \geq \varphi(n)$ et comme $f \mid \Phi_n$ cela impose $f = \Phi_n$.
Il en résulte que Φ_n est irréductible sur $\mathbf{Q}$, donc sur $\mathbf{Z}$ (cf. II 4.4) puisque le contenu de Φ_n est 1 (Φ_n est unitaire).

Corollaire 4.12.
Soit α, (resp. α') une racine n-ème (resp. m-ème) primitive de 1 dans $\mathbf{C}$. On suppose $(m,n) = 1$. Alors on a $\mathbf{Q}(\alpha) \cap \mathbf{Q}(\beta) = \mathbf{Q}$.

Démonstration. On prouve d'abord un lemme :

Lemme 4.13.
Soit $k \subset M$ une extension, K et L deux corps intermédiaires, KL le sous-corps de M engendré par K et L. Alors on a $[KL : L] \leq [K : k]$.

Démonstration (de 4.13) En effet, tout élément de KL peut s'écrire $x = \Sigma u_i v_i$ avec $u_i \in K, v_i \in L$ et donc, si $x_1, \ldots, x_n$ est une base de K sur k, il est clair que $x_1, \ldots, x_n$ engendrent KL sur L.

Revenons à 4.12 et posons $k = \mathbf{Q}(\alpha) \cap \mathbf{Q}(\alpha')$. Comme $(m, n) = 1$, $\alpha\alpha'$ est une racine primitive mn-ème de 1 (cf. I, 7.8), on a donc $\mathbf{Q}(\alpha\alpha') = \mathbf{Q}(\alpha)\mathbf{Q}(\alpha')$ (et aussi $k(\alpha\alpha') = k(\alpha)k(\alpha')$), et $[\mathbf{Q}(\alpha\alpha') : \mathbf{Q}] = \varphi(nm)$ (cf. Corollaire 4.11).
Mais, grâce à la condition $(m, n) = 1$, on a $\varphi(mn) = \varphi(m)\varphi(n)$ (cf. I, 7.4). Comme on a $[\mathbf{Q}(\alpha') : \mathbf{Q}] = \varphi(m)$, il en résulte que $[\mathbf{Q}(\alpha\alpha') : \mathbf{Q}(\alpha')] = \varphi(n)$ par multiplicativité des degrés (cf. 1.5). Mais le lemme 4.13 implique alors $[\mathbf{Q}(\alpha) : k] \geq \varphi(n)$, et avec $[\mathbf{Q}(\alpha) : \mathbf{Q}] = \varphi(n)$ on en déduit $k = \mathbf{Q}$.

e) Comportement de Φ_n sur $\mathbf{F}_p$.

L'exemple de $\Phi_8 = X^4 + 1$ montre qu'il est possible que Φ_n soit réductible sur tous les corps $\mathbf{F}_p$. Nous allons étudier précisément ce phénomène. Nous utiliserons, sans démonstration, le théorème suivant, dont nous avons vu un cas particulier en 2.16, (cf. aussi § 4 Exercice 14) :

Théorème de la progression arithmétique (Dirichlet).
Soient $a, n \in \mathbf{N}^*$ tels que $(a, n) = 1$. Il existe une infinité de nombres premiers p vérifiant $p \equiv a \pmod{n}$ (cf. [S1] Chapitre VI).

On a alors le résultat suivant :

Théorème 4.14.
Soit $n \in \mathbf{N}^*$. Les propriétés suivantes sont équivalentes :
1) il existe p premier, avec $(p, n) = 1$, tel que $\Phi_{n,\mathbf{F}_p}$ soit irréductible sur $\mathbf{F}_p$,
2) $(\mathbf{Z}/n\mathbf{Z})^*$ est cyclique,
3) on a $n = 1, 2, 4, q^\alpha$ ou $2q^\alpha$ avec q premier impair.

Démonstration.
L'équivalence de 2) et 3) résulte de la description de $(\mathbf{Z}/n\mathbf{Z})^*$, cf. Chapitre I § 7.
En vertu de 3.9, dire que Φ_n est réductible sur $\mathbf{F}_p$, signifie qu'il existe $m \in \mathbf{N}$ avec $m \leq \dfrac{\varphi(n)}{2}$ tel que Φ_n ait une racine dans $\mathbf{F}_{p^m}$. Cela signifie encore que $\mathbf{F}_{p^m}^*$ contient un élément d'ordre n i.e., puisque $\mathbf{F}_{p^m}^*$ est cyclique d'ordre $p^m - 1$, qu'on a $n \mid p^m - 1$. Autrement dit, Φ_n est réductible sur $\mathbf{F}_p$ si et seulement si p est d'ordre strictement inférieur à $\varphi(n)$ dans $(\mathbf{Z}/n\mathbf{Z})^*$.
Si $(\mathbf{Z}/n\mathbf{Z})^*$ n'est pas cyclique, cette condition est toujours réalisée, ce qui prouve l'implication 1) $\Longrightarrow$ 2). Réciproquement, si $(\mathbf{Z}/n\mathbf{Z})^*$ est cyclique et si $a \in \mathbf{N}$ est tel que $\overline{a}$ engendre $(\mathbf{Z}/n\mathbf{Z})^*$, on peut, en vertu du théorème de la progression arithmétique, supposer $a = p$ premier. Mais alors, comme p est d'ordre $\varphi(n)$ dans $(\mathbf{Z}/n\mathbf{Z})^*$, Φ_n est irréductible sur $\mathbf{F}_p$.

Remarque 4.15. Si p divise n, le polynôme $\Phi_{n,\mathbf{F}_p}$ n'a pas été défini, mais on peut cependant étudier la réduction modulo p de $\Phi_{n,\mathbf{Z}}$, notée $\overline{\Phi}_n$. On a alors le résultat suivant :

Proposition 4.16.
Si p est un diviseur de n, la réduction $\overline{\Phi}_n$ est réductible sur $\mathbf{F}_p$, sauf, éventuellement, si on a $p = 2$ et $n = 2q^\alpha$, avec q premier impair. En particulier le théorème 4.14 reste vrai sans la restriction $(p, n) = 1$.

Démonstration. On pose $n = p^\alpha m$, avec $p \nmid m$. On a donc, sur $\mathbf{F}_p$, $X^n - 1 = (X^m - 1)^{p^\alpha}$ par Frobenius. Supposons $\overline{\Phi}_n$ irréductible sur $\mathbf{F}_p$. Comme $\overline{\Phi}_n$ divise $X^n - 1$, il divise $X^m - 1$ et même, puisqu'on a :

$$X^m - 1 = \prod_{d \mid m} \Phi_{d,\mathbf{F}_p},$$

$\overline{\Phi}_n$ divise l'un des $\Phi_{d,\mathbf{F}_p}$, pour d diviseur de m.
En particulier, on a alors $\varphi(n) \leq \varphi(d)$. Mais on a $\varphi(n) = \varphi(m)\varphi(p^\alpha)$ et comme d divise m, on a $\varphi(d) \leq \varphi(m)$. Cela n'est possible qu'avec $\varphi(p^\alpha) = 1$ ce qui impose $p^\alpha = 2$. On a donc $n = 2m$, avec m impair. De plus, si d divise m, et $d \neq m$, avec m impair, on a $\varphi(d) < \varphi(m) = \varphi(n)$.
La seule possibilité est donc que $\overline{\Phi}_n$ divise $\Phi_{m,\mathbf{F}_p}$, et comme $\varphi(m) = \varphi(n)$ on a alors $\overline{\Phi}_n = \Phi_{m,\mathbf{F}_p}$. Comme $\overline{\Phi}_n$ est irréductible sur $\mathbf{F}_p$ et que, cette fois, p ne divise pas m, on peut appliquer le théorème 4.14, m étant impair, on a $m = q^\alpha$, avec q premier impair.

Remarque 4.17. Il reste à étudier le comportement de $\Phi_{2q^\alpha} = \Phi_{q^\alpha}$ sur $\mathbf{F}_2$. Notons que, par exemple, Φ_6 est irréductible sur $\mathbf{F}_2$, mais que Φ_{14} en revanche est réductible sur $\mathbf{F}_2$, cf. § 4 Exercice 12.

EXERCICES SUR LE CHAPITRE III

1. Les techniques vectorielles.

Quelques-uns des exercices font appel à la théorie de Galois pour laquelle on se reportera, par exemple, à [St].

0) Pour quels nombres premiers p, q a-t-on $\mathbf{Q}(\sqrt{p}) \subset \mathbf{Q}(\sqrt[3]{q})$?

---

1) Soit $k \subset K$ une extension de degré p, p premier. Déterminer les corps intermédiaires L (i.e. tels que $k \subset L \subset K$).

---

2) *Une autre démonstration du théorème 1.11* (dont on reprend les notations). Pour $1 \implies 3$, montrer que $1, \alpha, \ldots, \alpha^{n-1}$ engendrent $K[\alpha]$ si α est de degré n. Pour $3 \implies 2$, considérer l'application linéaire $x \mapsto ax$ de $K[\alpha]$ dans lui-même, avec $a \in K[\alpha]$, $a \neq 0$. Enfin, pour $2 \implies 1$, écrire que α^{-1} est dans $K[\alpha]$.

---

3) Soit $A = \{\alpha \in \mathbf{C} \mid \alpha \text{ est algébrique sur } \mathbf{Q}\}$.

a) Montrer que A est un corps.

b) Montrer que A est algébriquement clos (pour $P \in A[X]$, considérer le sous-corps de A engendré par les coefficients de P.)

c) En déduire que A est une clôture algébrique de $\mathbf{Q}$.

d) Montrer que l'ensemble des polynômes de degré $\leq n$ à coefficients dans $\mathbf{Q}$ est dénombrable. En déduire que $\mathbf{Q}[X]$ et A sont dénombrables.

e) Montrer qu'il existe des nombres réels transcendants sur $\mathbf{Q}$.

f) Trouver une extension $\mathbf{Q} \underset{\neq}{\subset} K \underset{\neq}{\subset} \mathbf{R}$, avec $\mathbf{R}$ algébrique sur K, donc K non dénombrable (prendre, à l'aide du Théorème de Zorn, un sous-corps K maximal tel que $\sqrt{2} \notin K$).

---

4) Soient $K \subset L \subset M$ des extensions. On suppose $K \subset L$ et $L \subset M$ finies (resp. algébriques). Montrer que $K \subset M$ est finie (resp. algébrique).

---

5) Donner les polynômes minimaux sur $\mathbf{Q}$ des éléments suivants de $\mathbf{C}$: $\sqrt[3]{7}+\sqrt{2}$, $i+j$, $j+\sqrt{3}$, $j\sqrt{2}$, $i+\sqrt{2}$.

Même question pour ζ, racine primitive cinquième de 1, mais sur $\mathbf{Q}$, $\mathbf{Q}(i)$, $\mathbf{Q}(\sqrt{5})$ (considérer $\zeta + \zeta^{-1}$).

---

6) *Théorème de l'élément primitif.*
Soit K un corps de caractéristique nulle.

a) Soit $P \in K[X]$, irréductible, L un corps de décomposition de P. Montrer que P n'a que des racines simples dans L.

b) Soit L une extension finie de K. On suppose que l'on a $L = K[x, y]$. Soient P, Q les polynômes minimaux de x, y sur K et soit $M = D_K(PQ)$.
On écrit dans M :

$$P(X) = (X - x)\prod_{i=2}^{n}(X - x_i)$$

$$Q(X) = (X - y)\prod_{j=2}^{m}(X - y_j).$$

Montrer qu'il existe $t \in K$ tel que pour tous $i, j \geq 2$, $x + ty \neq x_i + ty_j$.

c) Avec les notations ci-dessus, soient $z = x + ty$ et $K' = K[z]$. On pose $F(X) = P(z - tX)$. Montrer que F est dans $K'[X]$ et que l'on a $X - y = \text{pgcd}\,(F, Q)$.

d) En déduire qu'on a $y \in K'$ et $L = K'$.
Montrer par récurrence que toute extension finie de K est monogène (théorème de l'élément primitif).

e) Trouver $x \in \mathbf{C}$ tel que $\mathbf{Q}(x) = \mathbf{Q}(i, j, \sqrt{2})$.

f) Soit k un corps de caractéristique $p > 0$, $K = k(T, U)$ (corps des fractions rationnelles en deux indéterminées) et $K_0 = k(T^p, U^p)$.
Montrer qu'on a $[K : K_0] = p^2$ et que, pour tout $x \in K$, x^p est dans K_0. En déduire que K n'est pas une extension monogène de K_0.

---

7) Soit K un corps de caractéristique nulle.

a) Montrer qu'on a, pour L extension de K, l'équivalence :

$$[L : K] \leq n \Longleftrightarrow \forall x \in L,\ [K(x) : K] \leq n.$$

(cf. Exercice 6).

b) Montrer qu'il est possible d'avoir à la fois :

$$[L : K] = +\infty \quad \text{et} \quad \forall x \in L, \quad [K(x) : K] < +\infty.$$

c) Si $\text{car}\,(K) = p$ est > 0, montrer que a) peut être en défaut (s'inspirer de Exercice 6.f), mais avec une infinité d'indéterminées).

---

8) Montrer que l'ensemble des nombres constructibles est un corps. En utilisant la théorie de Galois, montrer que si x est algébrique sur $\mathbf{Q}$, de polynôme minimal P, x est constructible si et seulement si $[D_{\mathbf{Q}}(P) : \mathbf{Q}]$ est une puissance de 2.

---

9) Soit K l'ensemble des réels définis comme suit : $x \in K$ si et seulement si il existe une suite de corps : $K_0 = \mathbf{Q} \subset K_1 \subset \ldots \subset K_n \subset \mathbf{R}$ tels que : $x \in K_n$ et $[K_i : K_{i-1}] \leq 3$ pour $1 \leq i \leq n$.

a) Montrer que K est un sous-corps de $\mathbf{R}$.

b) Préciser les degrés sur $\mathbf{Q}$ des éléments de K.

c) Montrer que si $x \in K$, $x > 0$, x admet une racine carrée dans K.

d) Montrer que toute équation de degré 3 à coefficients dans K a une racine dans K.

e) Montrer que K ne contient pas de racine 11-ème de 1 distincte de 1.

f) Donner, à l'aide de la théorie de Galois, une caractérisation des éléments de K (cf. Exercice 8). En déduire quelles sont les racines de l'unité qui sont dans K.

---

10) Soit $P \in K[X]$, $n = d^\circ P$. Montrer que $[D_K(P) : K]$ divise $n!$.

---

2 et 3. Corps finis et irréductibilité des polynômes de $k[X]$.

Les exercices sur ces deux paragraphes ont été rassemblés car ils mêlent souvent les deux thèmes.

1) Décrire les corps finis de cardinal 4, 8, 9, 16 ; en particulier déterminer l'ordre des éléments inversibles et leur polynôme minimal sur le corps premier.

---

2) Montrer qu'on a $2 \in \mathbf{F}_p^{*2} \Longleftrightarrow p \equiv \mp 1 \pmod 8$ (prendre ζ, racine huitième primitive de 1 dans $\mathbf{F}_p^{*2}$, cf. 3.12, et remarquer que $\zeta + \zeta^{-1}$ est une racine de 2.)

---

3) Déterminer les générateurs de $\mathbf{F}_p^*$ pour $p = 2, 3, 5, 7, 11, 31, 43, 71$. (Commencer par essayer les petits entiers : $\mp 2, \mp 3 \ldots$ et ne pas oublier que si x et y sont d'ordres premiers entre eux, on a (cf. I, 7.8) :

$$\text{ordre } (xy) = \text{ordre } x \times \text{ordre } y.)$$

---

4) *Théorème de l'élément primitif pour les corps finis.*
On considère l'extension $K = \mathbf{F}_q \subset L = \mathbf{F}_{q^n}$. Montrer qu'il existe $\alpha \in L$ tel que $L = K[\alpha]$ (cf. 2.7 et § 1 Exercice 6).

---

5) *Groupe de Galois de $\mathbf{F}_{q^n}$ sur $\mathbf{F}_q$.*
Soit $\sigma : \mathbf{F}_{q^n} \longrightarrow \mathbf{F}_{q^n}$ l'application définie par $\sigma(x) = x^q$.

a) Montrer que σ est un automorphisme de $\mathbf{F}_{q^n}$.

b) Montrer que l'on a $\sigma|_{\mathbf{F}_q} = \mathrm{Id}$.

c) On définit le groupe de Galois de $\mathbf{F}_{q^n}$ sur $\mathbf{F}_q$ comme le groupe des automorphismes de $\mathbf{F}_{q^n}$ dont la restriction à $\mathbf{F}_q$ est l'identité. En utilisant 4), montrer qu'on a $|\mathrm{Gal}\,(\mathbf{F}_{q^n}/\mathbf{F}_q)| \leq n$. En déduire que $\mathrm{Gal}\,(\mathbf{F}_{q^n}/\mathbf{F}_q)$ est le groupe cyclique engendré par σ.

---

6) *Clôture algébrique de $\mathbf{F}_p$.*
Lorsque q' est une puissance de q, on identifie $\mathbf{F}_q$ à l'unique sous-corps à q éléments de $\mathbf{F}_{q'}$ (cf. III, 2.6).

a) Soit $\overline{\mathbf{F}}_p$ une clôture algébrique de $\mathbf{F}_p$. Montrer que tout $x \in \overline{\mathbf{F}}_p$, $x \neq 0$ est une racine de l'unité.

b) Pour $m \leq n$, montrer que l'on a $\mathbf{F}_{p^m} \subset \mathbf{F}_{p^n}$. [1]

c) Soit q une puissance de p, montrer qu'il existe n tel que $\mathbf{F}_q \subset \mathbf{F}_{p^n}$

d) Montrer que $K = \bigcup_{n \in \mathbf{N}^*} \mathbf{F}_{p^n}$ est muni naturellement d'une structure de corps et que K est une clôture algébrique de $\mathbf{F}_p$ et même de tout corps fini de caractéristique p.

---

7) *Fonction de Möbius.*

On définit $\mu : \mathbf{N}^* \longrightarrow \{0, 1, -1\}$ comme suit : $\mu(1) = 1$; $\mu(n) = 0$ si n contient un facteur carré ; $\mu(p_1 \dots p_r) = (-1)^r$ si $p_1, \dots, p_r$ sont des nombres premiers distincts.

a) Montrer que μ est multiplicative au sens de l'arithmétique (i.e., si $n = n_1 n_2$ avec $(n_1, n_2) = 1$, on a $\mu(n) = \mu(n_1)\mu(n_2)$).

b) Montrer que pour tout $n \in \mathbf{N}^*$, $n \neq 1$, on a $\sum_{d \mid n} \mu(d) = 0$.

c) Soit $f : \mathbf{N}^* \longrightarrow A$ une application, où A désigne un groupe abélien noté additivement. On pose : $g(n) = \sum_{d \mid n} f(d)$.

Démontrer la "formule d'inversion de Möbius" : $f(n) = \sum_{d \mid n} \mu(\frac{n}{d}) g(d)$. Énoncer une variante multiplicative.

d) En déduire la formule : $\varphi(n) = \sum_{d \mid n} \mu(\frac{n}{d}) d$.

---

8) *Polynômes irréductibles sur* $\mathbf{F}_q$.

Soit $I(n, q)$ le nombre de polynômes de $\mathbf{F}_q[X]$, irréductibles, unitaires, de degré n.

a) Montrer la relation $\sum_{d \mid n} d\, I(d, q) = q^n$ (considérer $X^{q^n} - X$ et ses facteurs irréductibles).

b) En déduire la formule :

$$I(n, q) = \frac{1}{n} \sum_{d \mid n} \mu(\frac{n}{d})\, q^d \quad \text{(cf. Exercice 7).}$$

c) Montrer que $n\, I(n, q)$ est le nombre d'éléments de $\mathbf{F}_{q^n}$ qui ne sont dans aucun sous-corps propre de $\mathbf{F}_{q^n}$ contenant $\mathbf{F}_q$.

d) Calculer $I(6, 3)$ et $I(7, 5)$.

e) Trouver un équivalent de $I(n, q)$ quand n tend vers $+\infty$

f) Donner la liste des polynômes irréductibles :

- de degré ≤ 5 sur $\mathbf{F}_2$,
- de degré ≤ 3 sur $\mathbf{F}_3$,
- de degré ≤ 2 sur $\mathbf{F}_4$ et $\mathbf{F}_5$.

---

[1] Attention, en revanche, $\mathbf{F}_{p^m}$ n'est pas en général un sous-corps de $\mathbf{F}_{p^n}$, il faut pour cela que m divise n. Par exemple $\mathbf{F}_4$ n'est pas inclus dans $\mathbf{F}_8$.

9) *Irréductibilité de $X^p - a$.*
Soient k un corps, $a \in k$ et p un nombre premier. Montrer que $X^p - a$ est irréductible sur k si et seulement si il n'a pas de racines dans k. (Si $X^p - a = PQ$ avec $P, Q \in k[X]$ et $0 < n = d^\circ P < p$, montrer, en décomposant $X^p - a$ en facteurs de degré 1, que l'on a $a^n = b^p$ où b est, au signe près, le terme constant de P. Conclure avec Bézout).

---

10) Etudier l'irréductibilité de $X^{p^r} - a$ pour p premier et r entier, (cf. [L] Ch. VIII § 9 pour l'étude générale de $X^n - a$; attention le cas étudié en 9) manque dans certaines éditions de [L]).

---

11) *Irréductibilité d'un polynôme à coefficients dans un anneau.*
Soit A un anneau intègre, K son corps des fractions.

a) Soit $P \in A[X]$. Si P est irréductible dans $K[X]$, et si les coefficients de P sont premiers entre eux, montrer que P est irréductible dans $A[X]$ (cf. II § 4).

b) Soit L une extension de K et $x \in L$ un élément entier sur A (cf. Chapitre II, § 3 Exercice 2). Montrer qu'il existe $P \in A[X]$, unitaire, irréductible tel que $P(x) = 0$.

c) En déduire que si A est factoriel, le polynôme minimal de x sur K est dans $A[X]$ (voir la démonstration du théorème 4.10).

d) Montrer que si A n'est pas intégralement clos, il existe $P \in A[X]$ irréductible dans $A[X]$ et pas dans $K[X]$ (cf. Chapitre II *loc. cit.*).

e) Montrer que si A est intégralement clos et si $P \in A[X]$ est unitaire et irréductible sur A, il l'est aussi sur K (considérer les racines de P dans une extension L de K et montrer que ce sont des entiers algébriques sur A. En admettant que l'ensemble des éléments de L entiers sur A est un anneau (cf. [L] Ch. IX § 1), en déduire que les coefficients de tout diviseur de P sont entiers sur A).
Déduire de ce qui précède que c) est encore vrai si A est intégralement clos.

---

12) *Les entiers de $\mathbf{Q}(\sqrt{d})$.*
Soit $d \in \mathbf{Z}$ sans facteur carré, i.e. $d = \mp p_1 \dots p_r$ avec $p_1, \dots, p_r$ premiers distincts. Soit $\sqrt{d} \in \mathbf{C}$ une racine carrée de d et

$$K = \mathbf{Q}(\sqrt{d}) = \{a + b\sqrt{d} \mid a, b \in \mathbf{Q}\}.$$

Soit $A = \{z \in K \mid z \text{ est entier sur } \mathbf{Z}\}$ (cf. Chapitre II § 3 Exercice 2).

a) Soit $x = a + b\sqrt{d} \in \mathbf{Q}(\sqrt{d})$. Montrer l'équivalence :

$$x \in A \iff 2a \in \mathbf{Z} \quad \text{et} \quad a^2 - db^2 \in \mathbf{Z}.$$

b) On suppose $d \equiv 2, 3 \pmod 4$, montrer qu'on a :

$$A = \mathbf{Z}[\sqrt{d}] = \{a + b\sqrt{d} \mid a, b \in \mathbf{Z}\}.$$

c) On suppose $d \equiv 1 \pmod 4$, montrer qu'on a $A = \mathbf{Z}\left[\dfrac{1 + \sqrt{d}}{2}\right]$.

---

13) Déterminer tous les nombres $x \in \mathbf{Q}$ tels que $\cos 2\pi x \in \mathbf{Q}$ (utiliser 12).

14) Sur quels corps finis les polynômes cyclotomiques Φ_3, Φ_5, Φ_6, Φ_7, sont-ils irréductibles ?

---

15) *Ou comment construire des exemples :*
Soient $p_1, \ldots, p_r$ des nombres premiers et soient $P_1, \ldots, P_r \in \mathbf{Z}[X]$. Montrer qu'il existe $P \in \mathbf{Z}[X]$ irréductible tel que, pour $i = 1, \ldots, r$, on ait

$$P \equiv P_i \pmod{p_i}$$ (utiliser le lemme chinois).

Étudier sur $\mathbf{Z}$, $\mathbf{F}_2$, $\mathbf{F}_3$, $\mathbf{F}_5$, le polynôme : $X^4 - 10X^3 + 21X^2 - 10X + 11$.

---

16) Étudier l'irréductibilité des polynômes suivants sur $\mathbf{Z}$:
$X^3 + 4X^2 - 5X + 7$, $5X^3 + 3X^2 - 4X - 27$, $X^3 - 6X^2 - 4X - 13$,
$X^3 + 4X^2 - 4X + 25$, $X^3 - X^2 - X - 1$, $X^3 + 30X^2 + 6X + 1$,
$5X^4 + 17X^3 - 8X^2 - 6X + 23$, $X^4 + 5X^3 - 3X^2 - X + 7$, $X^4 + 7X^2 + 4X + 1$,
$X^4 + X^3 + 2X^2 + X + 1$, $X^4 + 4X^3 + 3X^2 + 7X - 4$,
$7X^5 + 4X^4 - 2X^3 + 5X^2 - 6X + 11$,
$X^5 + X^4 + X^3 + X^2 + X + 1 - X^i$ pour $i = 1, 2, 3, 4$.
$X^5 + 3X^4 - 2X^3 - 4X^2 + 5X + 4$, $X^5 + 4X^4 - 4X^3 + 11X^2 - 5X + 7$,
$X^6 + X^3 + 1$, $X^6 + X^2 + 1$, $X^6 + X + 1$,
$X^7 + X + 1$, $5X^7 - 2X^6 + 4X^5 - 12X^4 + 19X^3 - 6X^2 - 20X + 17$
(réduire systématiquement modulo 2, 3 ...).

---

17) Étudier l'irréductibilité des polynômes :

X^2+1, X^3+1, X^4+1, X^4-X^2+1, X^3-X-1, X^5-7, X^4-3, X^2+X+1,

sur $\mathbf{Z}$, $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{C}$, $\mathbf{Q}(\sqrt{d})$, $\mathbf{Q}(\zeta)$, où ζ désigne une racine de l'unité, $\mathbf{F}_q$ (difficile).

---

18) Soient p, l deux nombres premiers impairs. On suppose :
1) $l \equiv -1 \pmod{3}$,
2) p est un générateur de $(\mathbf{Z}/l\mathbf{Z})^*$.
Montrer que $X^{l+1} - X + p$ est irréductible sur $\mathbf{Z}$ (réduire modulo p et modulo 2).
Exemple : $p = 47$, $l = 71$.

---

19) Soit p premier impair et l un diviseur premier de $p - 1$. Soit $a \in \mathbf{Z}$ un générateur de $(\mathbf{Z}/p\mathbf{Z})^*$. Montrer que $X^l + pX^k - a$ est irréductible sur $\mathbf{Z}$ pour tout entier $k \geq 1$ (utiliser l'exercice 9).

---

4. Cyclotomie.

1) Pour p premier, calculer $\Phi_p(X)$, puis $\Phi_{p^\alpha}(X)$, pour $\alpha \in \mathbf{N}^*$.

---

2) Montrer que si n est impair, $n > 1$, on a $\Phi_{2n}(X) = \Phi_n(-X)$ et si n est pair, $\Phi_{2n}(X) = \Phi_n(X^2)$. Plus généralement, montrer que si $n = p_1^{\alpha_1} \dots p_r^{\alpha_r}$, on a $\Phi_n(X) = \Phi_{p_1 \dots p_r}(X^m)$ avec $m = p_1^{\alpha_1 - 1} \cdots p_r^{\alpha_r - 1}$.
Montrer que si p est premier et ne divise pas pas n on a :

$$\Phi_{p^n}(X) = \Phi_n(X^p)/\Phi_n(X).$$

---

3) On pose $\Phi_n(X) = a_n X^n + a_{n-1} X^{n-1} + \cdots + a_0$. Montrer que Φ_n est un polynôme réciproque, i.e. que l'on a $a_n = a_0$, $a_{n-1} = a_1$, $\dots, a_{n-i} = a_i$ pour $i = 0, \dots, n$ (considérer $\Phi_n(1/X)$).

---

4) Montrer la formule :

$$\Phi_n(X) = \prod_{d \mid n} (X^{n/d} - 1)^{\mu(d)}$$

où μ est la fonction de Möbius (cf. § 2,3 Exercice 7).

---

5) Pour quels entiers n a-t-on $\varphi(n) \leq 10$? Calculer Φ_n pour ces entiers (utiliser les exercices 1, 2, 3).

---

6) Trouver n tel que Φ_n ait d'autres coefficients que 0, 1, -1 (réponse : $n = 105$).

---

7) Soit $\zeta = x + iy$ l'unique racine primitive 5-ème de 1 dans $\mathbf{C}$ vérifiant $x > 0$ et $y > 0$. Calculer x (utiliser $\zeta + \zeta^{-1} = \zeta + \overline{\zeta} = 2x$ et l'exercice 3).
En déduire que x et y sont constructibles (cf. § 1) et donner une construction explicite du pentagone régulier.

---

8) a) Soit $\zeta \in \mathbf{C}$ une racine primitive d-ème de 1, avec $d \geq 1$ et soit $k = \mathbf{Q}(\zeta)$. Montrer que les seules racines de l'unité contenues dans k sont :

i) les racines d-èmes si d est pair,

ii) les racines $2d$-èmes si d est impair.

(Considérer le groupe G des racines de l'unité de k. Montrer que G est fini. Si $|G| = n$, prendre une racine primitive n-ème α et utiliser l'inclusion $\mathbf{Q}(\alpha) \subset \mathbf{Q}(\zeta)$).

b) Déduire de a) que si on se donne des entiers d et n avec $d \mid n$ et si on suppose (d pair) ou (d et n impairs) il existe un corps k avec $\mu_n(k) \simeq \mathbf{Z}/d\mathbf{Z}$.

c) Montrer que si $d = 2^r - 1$, et si d divise n il existe k avec $\mu_n(k) \simeq \mathbf{Z}/d\mathbf{Z}$ (exemple : $d = 3$, $n = 6$).

d) Montrer que, quel que soit le corps k, $\mu_{30}(k)$ n'est jamais isomorphe à $\mathbf{Z}/5\mathbf{Z}$. Étudier le cas général.

---

9) Soit K une extension finie de $\mathbf{Q}$. Montrer qu'il n'y a qu'un nombre fini de racines de l'unité dans K.

---

10) Dans quelles extensions $\mathbf{Q}(\sqrt{d})$, $d \in \mathbf{Q}$, y-a-t-il des racines de l'unité autres que 1 et -1 ?

11) Calculer $\Phi_n(1)$ et $\Phi_n(-1)$ pour $n \in \mathbf{N}^*$.

---

12) On considère le nombre $p = 1093$.

a) Montrer que p est premier.

b) Montrer qu'on a $2^{p-1} \equiv 1 \pmod{p^2}$.

(On pourra montrer successivement, dans $\mathbf{Z}/p^2\mathbf{Z}$, les égalités :

$$3^{14} = 4p + 1,\ 3^2 2^{26} = -469p - 1,\ 3^{14} 2^{182} = -4p - 1,\ 2^{182} = -1).$$

Les nombres premiers vérifiant b) sont rarissimes, 1093 et 3511 sont les deux seuls tels nombres plus petits que 10^5.

c) Montrer que Φ_{p^2} est réductible sur $\mathbf{F}_2$.

---

13) *Symbole de Legendre.*

Soit p un nombre premier > 2. On définit le symbole de Legendre $(\frac{x}{p})$ pour $x \in \mathbf{F}_p^*$ comme suit :

$$(\frac{x}{p}) = 1 \Longleftrightarrow x \in \mathbf{F}_p^{*2}, \qquad (\frac{x}{p}) = -1 \Longleftrightarrow x \notin \mathbf{F}_p^{*2}.$$

Il revient au même de poser (cf. § 2) :

$$(\frac{x}{p}) = x^{\frac{p-1}{2}}.$$

a) Montrer que l'application $x \mapsto (\frac{x}{p})$ est un caractère (i.e un homomorphisme) de $\mathbf{F}_p^*$ dans $\{1, -1\}$.

b) Montrer qu'on a $\sum_{x \in \mathbf{F}_p^*} (\frac{x}{p}) = 0$.

c) Soit ζ une racine primitive p-ème de 1 dans $\mathbf{C}$.

On pose $s = \sum_{x \in \mathbf{F}_p^*} (\frac{x}{p}) \zeta^x$ où ζ^x a un sens évident. Montrer qu'on a $s^2 = (\frac{-1}{p})p$.

(On pourra montrer la formule :

$$s^2 = \sum_{x,y} (\frac{y}{p}) \zeta^{x(y+1)}$$

et séparer les termes $y = -1$ et $y \neq -1$).

d) Déduire de c) que toute extension quadratique de $\mathbf{Q}$ (i.e. de la forme $\mathbf{Q}(\sqrt{d})$, avec $d \in \mathbf{Q}$) est contenue dans une extension cyclotomique (i.e. de la forme $\mathbf{Q}(\zeta)$, avec ζ racine de l'unité). Il s'agit du cas le plus élémentaire d'un célèbre théorème de Kronecker et Weber.

Pour des compléments sur le symbole de Legendre, en particulier la superbe loi de réciprocité quadratique, cf. [S1] Chapitre I § 3.

---

14) *Un "petit" théorème de Dirichlet.*

Soit n un entier ≥ 2.

a) Montrer qu'un nombre premier p est congru à 1 modulo n si et seulement si $\mathbf{F}_p$ contient une racine primitive n-ème de l'unité.

b) Soit k un entier et p un facteur premier de $\Phi_n(k!)$.

Montrer qu'on a $p \equiv 1 \pmod{n}$ et $p \geq k$. Conclure (cf. 4.14).

IV. LE GROUPE LINÉAIRE

Soit k un corps (commutatif, mais de caractéristique quelconque), et E un k-espace vectoriel de dimension n. ***Le groupe linéaire*** $GL(E)$ est le groupe des ***k-automorphismes*** de E, c'est-à-dire des applications k-linéaires bijectives de E dans E.
La donnée d'une base de E définit un isomorphisme de $GL(E)$ sur $GL(n,k)$, ***groupe des matrices $n \times n$, inversibles***, à coefficients dans k. Mais cet isomorphisme n'est pas canonique (i.e. dépend du choix de la base), rappelons que si $u \in GL(E)$ a pour matrice A dans une base $\mathcal{B}$, dans la base $\mathcal{B}'$ déduite de $\mathcal{B}$ par la matrice de passage P, il admet pour matrice $P^{-1}AP$. Remarquons que A et $P^{-1}AP$ sont conjuguées dans $GL(n,k)$.
L'intérêt de cet isomorphisme est de fournir un outil pour l'étude de $GL(E)$, à savoir, le calcul matriciel (cf. § 4 pour un exemple d'utilisation).
On suppose le lecteur familier avec les bases de l'algèbre linéaire, notamment le calcul matriciel et la notion de déterminant.

1. Déterminant et groupe $SL(E)$.

L'application déterminant est un homomorphisme (multiplicatif) de $GL(E)$ dans k^*. Son noyau est appelé groupe spécial linéaire et noté $SL(E)$, il est isomorphe au groupe $SL(n,k)$ des matrices de déterminant 1.

Proposition 1.1.
On a une suite exacte :

$$1 \longrightarrow SL(E) \longrightarrow GL(E) \xrightarrow{\text{dét}} k^* \longrightarrow 1.$$

De plus, $GL(E)$ est produit semi-direct de $SL(E)$ par k^*.

Démonstration. On peut travailler dans $GL(n,k)$. Soit H le sous-groupe de $GL(n,k)$ formé des matrices de la forme

$$A(\lambda) = \begin{pmatrix} \lambda & & & \\ & 1 & & 0 \\ & & \ddots & \\ & & & \ddots \\ & 0 & & \\ & & & 1 \end{pmatrix}, \quad \lambda \in k^*$$

Alors la restriction $\det|_H$ induit un isomorphisme de H sur k^*, ce qui prouve à la fois la surjectivité du déterminant et le produit semi-direct.

Nous avons ainsi un premier dévissage de $GL(E)$. Nous allons maintenant étudier, suivant le plan annoncé au chapitre I, des générateurs de $GL(E)$ et $SL(E)$, les centres, les groupes dérivés, et enfin, la simplicité de ces groupes.

2. Générateurs et centres de $GL(E)$ et $SL(E)$.

On cherche des générateurs les plus simples possibles, donc ayant, comme les transpositions dans le cas de $\mathfrak{S}_n$, le plus de points fixes possibles, c'est-à-dire ici, un hyperplan.

a) Les dilatations.

Proposition-définition 2.1.
Soit H un hyperplan de E et soit $u \in GL(E)$ tel que $u|_H = \mathrm{Id}_H$. Les conditions suivantes sont équivalentes :
1) on a $\det u = \lambda \neq 1$ *(i.e. $u \notin SL(E)$),*
2) u admet une valeur propre $\lambda \neq 1$ (donc une droite propre D pour λ) et u est diagonalisable,
3) on a $\mathrm{Im}(u - \mathrm{Id}) \not\subset H$,
4) dans une base convenable, u a pour matrice :

$$\begin{pmatrix} 1 & & & 0 \\ & \ddots & & \\ & & \ddots & \\ 0 & & & \lambda \end{pmatrix}, \text{ avec } \lambda \in k^*,\ \lambda \neq 1.$$

On dit que u est une ***dilatation d'hyperplan H, de droite D, de rapport*** *λ. On a alors $D = \mathrm{Im}\,(u - \mathrm{Id})$; $H = \mathrm{Ker}\,(u - \mathrm{Id})$. Lorsque l'on a $\lambda = -1$ et $\mathrm{car}\,(k) \neq 2$, u est appelée une* ***réflexion.***

Démonstration. Il est clair qu'on a 4) $\Longrightarrow$ 1) et 1) $\Longrightarrow$ 2) (car u a déjà la valeur propre 1 d'ordre $n-1$ avec le sous-espace propre H de dimension $n-1$).
2) $\Longrightarrow$ 3) : Soit x un vecteur propre non nul pour la valeur propre $\lambda \neq 1$, on a alors $u(x) = \lambda x$, donc $u(x) - x = (\lambda - 1)x \neq 0$ et $x \notin H$, de sorte que l'on a $\mathrm{Im}\,(u - \mathrm{Id}) \not\subset H$. Remarquons que comme on a $\dim(\mathrm{Im}\,(u - \mathrm{Id})) = 1$, on a bien $\mathrm{Im}\,(u - \mathrm{Id}) = (x) = D$.
3) $\Longrightarrow$ 4) : Notons déjà qu'on a $\dim(\mathrm{Im}(u - \mathrm{Id})) = 1$ (car le noyau est H, de dimension $n-1$). On prend une base $e_1, \ldots, e_{n-1}$ de H et un vecteur $e_n \in \mathrm{Im}\,(u - \mathrm{Id})$, $e_n \notin H$. Ce vecteur engendre donc $\mathrm{Im}\,(u - \mathrm{Id})$. Comme $e_n \notin H$, on a $u(e_n) - e_n \neq 0$, donc $u(e_n) - e_n = \mu e_n$ avec $\mu \neq 0$ et $u(e_n) = (1 + \mu)e_n = \lambda e_n$ avec $\lambda \neq 1$, d'où la matrice voulue dans la base $e_1, \cdots, e_n$.

b) Les transvections.

C'est le cas diamétralement opposé.

Proposition-définition 2.2.
Soit H un hyperplan de E, d'équation $f \in E^*$, (cela signifie qu'on a $H = \operatorname{Ker} f$, avec $f \neq 0$). Soit $u \in GL(E)$, $u \neq \mathrm{Id}$, tel que $u|_H = \mathrm{Id}_H$. Les conditions suivantes sont équivalentes :
1) on a $\det u = 1$ (i.e. $u \in SL(E)$),
2) u n'est pas diagonalisable,
3) on a $D = \operatorname{Im}(u - \mathrm{Id}) \subset H$,
4) l'homomorphisme induit, $\overline{u} : E/H \longrightarrow E/H$, est l'identité de E/H,
5) il existe $a \in H$, $a \neq 0$, tel que l'on ait :

$$\forall x \in E,\ u(x) = x + f(x)a,$$

6) dans une base convenable, u a pour matrice :

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

On dit alors que u est une ***transvection d'hyperplan H et de droite D.*** *On a, avec les notations ci-dessus, $D = (a)$ et $D \subset H$.*

Démonstration. Les implications suivantes sont claires : 6) $\Longrightarrow$ 1) $\Longrightarrow$ 2) $\Longrightarrow$ 3) (cf. a)).
Pour 5) $\Longrightarrow$ 6), on construit une base $e_1, \ldots, e_n$ de E en partant de $e_{n-1} = a$, que l'on complète en une base $e_1, \ldots, e_{n-1}$ de H et on prend enfin $e_n \notin H$, tel que $f(e_n) = 1$.
3) $\Longrightarrow$ 4) : Soit $x \in E$, on a $u(x) - x \in H$, donc, dans E/H, on a $\overline{u}(\overline{x}) = \overline{x}$. Notons que ceci montre en fait 3) $\Longleftrightarrow$ 4).
Reste enfin 3) $\Longrightarrow$ 5). Soit $x_0 \in E$ tel que $f(x_0) = 1$. L'élément $a = u(x_0) - x_0$ est dans l'image de $u - \mathrm{Id}$ donc est dans H. Comme x_0 n'est pas dans H, a est $\neq 0$. Alors, on a, pour tout x, $u(x) = x + f(x)\,a$. En effet, ces deux applications linéaires coïncident sur H et en $x_0 \notin H$ donc sont égales.

Remarques 2.3.

1) La caractérisation 5) est souvent la plus commode dans les calculs.

2) Dans le cas des dilatations, la donnée de H, D et λ est équivalente à celle de u. La situation est un peu compliquée pour les transvections. En effet, u détermine D et H, mais la réciproque est fausse (penser aux transvections $\begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}$). D'autre part la donnée d'un point $a \in D \subset H$ et d'une équation f de H détermine u (cf. 5)) ; mais u ne détermine f et a qu'à un scalaire près (si f et a conviennent, λf et a/λ aussi).

Pour $f \in E^*$, $f \neq 0$, et $a \in \operatorname{Ker} f$, $a \neq 0$, nous noterons $\tau(f, a)$ la transvection donnée par la formule :

$$\forall x \in E, \quad \tau(f,a)(x) = x + f(x)a.$$

Remarquons que si $\tau = \tau(f, a)$, on a $\tau^{-1} = \tau(f, -a)$ et aussi :

$$\tau(f,a)\tau(f,b) = \tau(f, a+b) \quad \text{(cf. § 2 Exercice 1).}$$

On a aussi une caractérisation duale des transvections :

Proposition 2.4.
Soit $u \in GL(E)$, $u \neq \mathrm{Id}$. Les propriétés suivantes sont équivalentes :
1) u est une transvection de droite D,
2) on a $u|_D = \mathrm{Id}$ et l'homomorphisme induit $\overline{u} : E/D \longrightarrow E/D$ est l'identité.

Démonstration. L'implication 1) $\Longrightarrow$ 2) est claire avec la caractérisation 5) des transvections.
2) $\Longrightarrow$ 1) : La condition sur le quotient, $\overline{u}(\overline{x}) = \overline{x}$, s'écrit encore :

$$\forall x \in E,\ u(x) - x \in D.$$

On a donc $\mathrm{Im}(u - \mathrm{Id}) \subset D$ et comme $u \neq \mathrm{Id}$, cela impose $\mathrm{Im}\,(u - \mathrm{Id}) = D$. Il en résulte que $\mathrm{Ker}\,(u - \mathrm{Id})$ est un hyperplan, contenant D puisque $u|_D = \mathrm{Id}$ et u est bien une transvection de droite D.

Proposition 2.5 (Comportement par conjugaison).
Soit τ une transvection de droite D et d'hyperplan H et soit $u \in GL(E)$. Alors, $u\tau u^{-1}$ est une transvection de droite $u(D)$ et d'hyperplan $u(H)$. Précisément, si on a $\tau = \tau(f, a)$, on a $u\tau u^{-1} = \tau(f \circ u^{-1}, u(a))$.

Démonstration. On a, pour $x \in E$, $\tau u^{-1}(x) = u^{-1}(x) + f(u^{-1}(x))a$ d'où $u\tau u^{-1}(x) = x + f(u^{-1}(x))\,u(a)$, d'où le résultat (on notera que si $H = \mathrm{Ker}\,f$, on a $u(H) = \mathrm{Ker}\,(f \circ u^{-1})$).

c) Application, calcul des centres.

Théorème 2.6.
Le centre Z de $GL(E)$ est formé des homothéties $x \mapsto \lambda x$, avec $\lambda \in k^*$. Il est donc isomorphe à k^*.
Le centre de $SL(E)$ est $Z \cap SL(E)$, il est isomorphe à $\mu_n(k) = \{\lambda \in k \mid \lambda^n = 1\}$ (cf. Chapitre III § 4).

Remarque 2.7. Pour $n = 1$, $GL(E) = k^*$ est commutatif et $SL(E) = \{1\}$.
Démonstration (de 2.6) On a d'abord un lemme qui caractérise géométriquement les homothéties :

Lemme 2.8.
Soit $u \in GL(E)$. Supposons que u laisse invariantes toutes les droites vectorielles de E, alors u est une homothétie.

Démonstration. En formules, ceci s'écrit comme une interversion de quantificateurs :

$$(\forall x \in E),\ (\exists \lambda \in k^*),\ (u(x) = \lambda x) \Longrightarrow (\exists \lambda \in k^*),\ (\forall x \in E),\ (u(x) = \lambda x).$$

Pour $n = 1$, c'est clair. Sinon, si x, y sont non colinéaires on a : $u(x) = \lambda x$, $u(y) = \mu y$, mais aussi $u(x + y) = \nu(x + y) = \lambda x + \mu y$, d'où $\lambda = \mu = \nu$.
Si x et y sont colinéaires le résultat est évident.

Le théorème 2.6 en résulte aussitôt, à l'aide de la proposition 2.5 : soit $u \in GL(E)$ qui centralise $SL(E)$. Alors si τ est une transvection de droite D, on a $u\tau u^{-1} = \tau$. Or $u\tau u^{-1}$ est une transvection de droite $u(D)$, de sorte qu'on a $u(D) = D$. Comme ceci vaut pour toute droite D, le lemme 2.8 montre que u est une homothétie.

Définition 2.9.
Le quotient de $GL(E)$ par son centre est appelé le ***groupe projectif linéaire*** *et est noté $PGL(E)$. De même le quotient de $SL(E)$ par son centre est noté $PSL(E)$. On note $PGL(n,k)$ et $PSL(n,k)$ les quotients des groupes matriciels correspondants.*

Remarque 2.10. Soit h_λ l'homothétie $x \mapsto \lambda x$, on a dét $h_\lambda = \lambda^n$, de sorte qu'on a une suite exacte :

$$1 \longrightarrow PSL(E) \longrightarrow PGL(E) \xrightarrow{\overline{dét}} k^*/k^{*n} \longrightarrow 1.$$

où on a posé $k^{*n} = \{\lambda \in k^* \mid \exists \mu \in k^*,\ \lambda = \mu^n\}$. En particulier, si k est algébriquement clos, on a un isomorphisme : $PSL(E) \simeq PGL(E)$.

d) Générateurs de $SL(E)$ et $GL(E)$.

Théorème 2.11.
Les transvections engendrent $SL(E)$.

Corollaire 2.12.
Les transvections et les dilatations engendrent $GL(E)$.

Démonstration. Le corollaire est immédiat via le théorème : soit $u \in GL(E)$ avec $\lambda = \text{dét}\, u$ et soit v une dilatation de rapport λ^{-1}. Alors on a $vu \in SL(E)$ et u est produit de v^{-1} et de transvections.
Le théorème 2.11 se prouve par récurrence sur n. Pour $n = 1$, c'est clair. On a ensuite un lemme qui décrit la transitivité des transvections :

Lemme 2.13.
Soient $x, y \in E - \{0\}$. Il existe une transvection u ou un produit de deux transvections uv, tels que $u(x) = y$ ou $uv(x) = y$.

Démonstration (de 2.13). Supposons x, y non colinéaires. On cherche u sous la forme $u(x) = x + f(x)a$. On prend $a = y - x$ et pour H un hyperplan contenant a, mais pas x. On choisit alors l'équation f de H de sorte que l'on ait $f(x) = 1$ et $u = \tau(f, a)$ convient.
Si x et y sont colinéaires, on prend z non colinéaire et on trouve, d'après ce qui précède, des transvections u, v telles que $u(x) = z$, $v(z) = y$.

Revenons au théorème 2.11 :
Soit $u \in SL(E)$ et soit $x \in E$, $x \neq 0$. Quitte à remplacer u par vu où v est un produit de transvections, on peut supposer que l'on a $u(x) = x$ (lemme 2.13).
Soit D la droite engendrée par x et soient $\pi : E \longrightarrow E/D$ la projection canonique et $\overline{u} : E/D \longrightarrow E/D$ l'automorphisme induit par u.
Montrons tout d'abord qu'on a $\overline{u} \in SL(E/D)$. Pour cela, on prend une base $e_1 = x, e_2, \ldots, e_n$ de E, de sorte que $\pi(e_2), \ldots, \pi(e_n)$ est une base de E/D. Si on écrit les matrices de u et $\overline{u}$ dans ces bases, en tenant compte de $u(e_1) = e_1$, le développement de dét u par rapport à la première colonne montre qu'on a aussi dét $\overline{u} = 1$.
On applique alors à $\overline{u}$ l'hypothèse de récurrence, on a $\overline{u} = \overline{\tau}_1 \ldots \overline{\tau}_r$, où $\overline{\tau}_i = \tau(\overline{f_i}, \overline{a_i})$ est une transvection de E/D. Soit alors $a_i \in E$ tel que $\pi(a_i) = \overline{a_i}$ et $f_i \in E^*$ définie par $f_i = \overline{f_i} \circ \pi$. Posons $\tau_i = \tau(f_i, a_i)$. Il est clair que τ_i induit $\overline{\tau_i}$ sur E/D. De plus, comme $f_i(x) = \overline{f_i} \circ \pi(x) = 0$, on a $\tau_i(x) = x$. Posons alors $v = \tau_1 \ldots \tau_r$,

on a $v(x) = u(x)$ et $\overline{v} = \overline{u}$ donc, en vertu de la proposition 2.4, $v^{-1}u$ est une transvection, de sorte que u est produit de transvections.

Variante.
On peut éviter le recours au quotient en démontrant le lemme suivant :

Lemme 2.14.
Soit $x \in E - \{0\}$ et soient H_1, H_2 deux hyperplans distincts, tels que $x \notin H_1 \cup H_2$. Alors, il existe une transvection u telle que $u(x) = x$ et $u(H_1) = H_2$.

Démonstration. On prend $(H_1 \cap H_2) + ka$ pour l'hyperplan de u.

On termine alors la démonstration du théorème 2.11 en se ramenant, par usage des lemmes 2.13 et 2.14, au cas $u(H) = H$ et $u(x) = x$, avec $x \notin H$, et en appliquant l'hypothèse de récurrence à $u|_H$.

Remarque 2.15. On peut prouver que tout élément u de $SL(E)$ est produit d'au plus $n = \dim E$ transvections sauf si u est une homothétie, auquel cas il en faut $n + 1$, (cf. §2 Exercices 5 à 8 pour des compléments sur ce théorème).

e) Conjugaison.

Nous cherchons maintenant des réciproques à la proposition 2.5.

Proposition 2.16.
Deux dilatations sont conjuguées dans $GL(E)$ si et seulement si elles ont même rapport.

Démonstration. C'est clair, car elles ont alors même matrice dans des bases convenables.

Proposition 2.17.
Deux transvections quelconques sont conjuguées dans $GL(E)$. Pour $n \geq 3$, elles le sont aussi dans $SL(E)$.

Démonstration. Dans $GL(E)$, c'est clair, elles ont même réduite de Jordan (cf. 2.2, caractérisation 6).
Supposons $n \geq 3$ et soient u, v deux transvections et $w \in GL(E)$, tel que $v = wuw^{-1}$. Si $\lambda = \det w$, il suffit de trouver $s \in GL(E)$, avec $\det s = \lambda^{-1}$ et $svs^{-1} = v$. En effet, alors, on aura $(sw)u(sw)^{-1} = v$ et $sw \in SL(E)$.
Pour ceci, on se place dans une base sur laquelle v a pour matrice :

$$\begin{pmatrix} 1 & & & & 0 \\ & \ddots & & & \\ & & \ddots & & \\ & & & 1 & 1 \\ 0 & & & & 1 \end{pmatrix} \quad \text{et on prend} \quad s = \begin{pmatrix} 1 & & & & & 0 \\ & \ddots & & & & \\ & & 1 & & & \\ & & & \lambda & & \\ & & & & 1/\lambda & \\ 0 & & & & & 1/\lambda \end{pmatrix}$$

ce qui est possible puisqu'on a $n \geq 3$. Il est clair alors que s convient.

Pour $n = 2$, la proposition analogue est fausse. On a le résultat matriciel suivant :

Proposition 2.18.
1) Dans $SL(2, k)$ toute transvection est conjuguée d'une matrice $\begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}$, avec

$\lambda \in k^*$.

2) Soient $\lambda, \mu \in k^*$, *alors* $s = \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}$ *et* $t = \begin{pmatrix} 1 & \mu \\ 0 & 1 \end{pmatrix}$ *sont conjuguées dans* $SL(2,k)$ *si et seulement si* λ/μ *est un carré dans* k.

Démonstration.

1) Soient u une transvection, e_1, e_2 une base de E, $k\varepsilon_1$ l'hyperplan de u, et soit $\varepsilon_2 \notin k\varepsilon_1$. Dans la base $(\alpha\varepsilon_1, \varepsilon_2)$, u a la matrice voulue et, pour un α convenable, on a dét $(\alpha\varepsilon_1, \varepsilon_2)/(e_1, e_2) = 1$, donc le changement de base est dans $SL(2,k)$.

2) Supposons qu'il existe $g = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ avec $\alpha\delta - \beta\gamma = 1$ vérifiant $gsg^{-1} = t$.

On a alors $gs = tg$ c'est-à-dire :

$$gs = \begin{pmatrix} \alpha & \alpha\lambda + \beta \\ \gamma & \gamma\lambda + \delta \end{pmatrix} = tg = \begin{pmatrix} \alpha + \mu\gamma & \beta + \mu\delta \\ \gamma & \delta \end{pmatrix}.$$

On voit donc que la relation $gs = tg$ implique $\gamma = 0$ et $\alpha\lambda = \mu\delta$, avec $\delta = 1/\alpha$ car g est de déterminant 1 et donc $\lambda/\mu = \delta^2$ est un carré de k.
Réciproquement, si $\lambda/\mu = \delta^2$, avec $\delta \in k^*$, on prend $\alpha = 1/\delta$, $\gamma = 0$ et β quelconque, et g convient pour passer de u à v.

Remarque 2.19. Les classes de conjugaison des transvections dans $SL(2,k)$ dépendent donc de manière essentielle de la structure de k. Par exemple, il y a une seule classe si k est algébriquement clos, deux si $k = \mathbf{R}$ ou $\mathbf{F}_q$, une infinité si $k = \mathbf{Q}$.

3. Commutateurs.

Les résultats de ce numéro sont conséquences de ceux du paragraphe suivant, mais nous en donnons des démonstrations directes.

Théorème 3.1.
1) On a $D(GL(n,k)) = SL(n,k)$, *sauf dans le cas :* $(n = 2, k = \mathbf{F}_2)$.
2) On a $D(SL(n,k)) = SL(n,k)$, *sauf dans les deux cas :* $(n = 2, k = \mathbf{F}_2)$ et $(n = 2, k = \mathbf{F}_3)$.

Démonstration. On note encore E un k-espace vectoriel de dimension n.

a) Si g, h sont dans $GL(E)$, on a dét $(ghg^{-1}h^{-1}) = 1$ et donc on a toujours $D(GL(E)) \subset SL(E)$ et $D(SL(E)) \subset SL(E)$.

b) Il suffit de prouver qu'**une** transvection u est un commutateur.
En effet, si on a $u = aba^{-1}b^{-1}$ avec $a, b \in SL(E)$, (resp. $a, b \in GL(E)$), et si v est une autre transvection, u et v sont conjuguées dans $GL(E)$ (cf. 2.17), donc on a $v = gug^{-1} = i_g(u)$, où i_g désigne l'automorphisme intérieur défini par $g \in GL(E)$. On a alors :

$$v = i_g(u) = i_g(a)i_g(b)i_g(a)^{-1}i_g(b)^{-1}$$

et comme $SL(E)$ est distingué dans $GL(E)$, on $i_g(a), i_g(b) \in SL(E)$ (resp. $GL(E)$). On voit ainsi que toutes les transvections sont des commutateurs, mais comme elles engendrent le groupe $SL(E)$ (cf. 2.11), on a $SL(E) \subset D(SL(E))$ (resp. $SL(E) \subset D(GL(E))$).

c) Un cas particulier bien agréable est le suivant : $n \geq 3$ et car $(k) \neq 2$.
En effet, si u est une transvection, u^2 en est aussi une (car $u^2 \neq \mathrm{Id}$ à cause de car $(k) \neq 2$, comme on le voit par exemple sur la matrice de u^2) de sorte que u et

u^2 sont conjuguées dans $SL(E)$ (cf. 2.17) et on a donc $u^2 = sus^{-1}$ avec $s \in SL(E)$, d'où $u = sus^{-1}u^{-1}$ et $u \in D(SL(E))$. ([1])
On a montré ainsi $D(SL(E)) = SL(E)$ et, *a fortiori*, $D(GL(E)) = SL(E)$.

d) Supposons $|k| > 3$ (i.e. $k \neq \mathbf{F}_2, \mathbf{F}_3$) et $n = 2$.
Si on pose $t = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ et $s = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix}$, $(\lambda \neq 0)$, on a

$$\tau = sts^{-1}t^{-1} = \begin{pmatrix} 1 & \lambda^2 - 1 \\ 0 & 1 \end{pmatrix},$$

et si on choisit $\lambda \neq \pm 1$ (ce qui est possible car on a supposé $|k| > 3$), τ est une transvection et on a $D(SL(2,k)) = D(GL(2,k)) = SL(2,k)$.

Pour $n > 2$, la même méthode fonctionne en choisissant un plan P, un supplémentaire S de P et en prolongeant les matrices ci-dessus par Id_S.

e) Si $k = \mathbf{F}_2, \mathbf{F}_3$, mais $n \geq 3$, on considère les matrices :

$$u = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}; \; t = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ et } s = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

de sorte que u est une transvection et qu'on a $u = tst^{-1}s^{-1}$, avec $s, t \in SL(E)$ et on conclut comme en d).

f) Il reste à montrer $D(GL(2,\mathbf{F}_3)) = SL(2,\mathbf{F}_3)$.
On regarde les matrices $t = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, $s = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. On a $sts^{-1}t^{-1} = t$ et donc t est dans $D(GL(E))$ (on prendra garde que s n'est pas dans $SL(E)$ et qu'on ne peut donc pas conclure pour $D(SL)$).

g) *Les cas exceptionnels :*
On a $GL(2,\mathbf{F}_2) = SL(2,\mathbf{F}_2) \simeq \mathfrak{S}_3$ et donc $D(SL(2,\mathbf{F}_2)) \simeq \mathfrak{A}_3$ (cf. § 5).
On a $|SL(2,\mathbf{F}_3)| = 24$, et $D(SL(2,\mathbf{F}_3)) \simeq H_8$ (cf. § 5 et § 3 Exercice 2).

4. La simplicité de $PSL(n,k)$.

Théorème 4.1.
Le groupe $PSL(n,k)$ est simple, sauf dans les deux cas suivants :
1) $n = 2$, $k = \mathbf{F}_2$,
2) $n = 2$, $k = \mathbf{F}_3$.

Démonstration. On utilise les techniques inaugurées au Chapitre I § 8.
Soit E un k-espace vectoriel de dimension n et soit $\overline{N}$ un sous-groupe distingué de $PSL(E)$, non réduit à l'élément neutre. Par image réciproque il lui correspond un sous-groupe distingué N de $SL(E)$, contenant le centre Z de $SL(E)$, et distinct de Z, et il faut montrer que l'on a $N = SL(E)$. Nous devons distinguer deux cas.
1) Premier cas : $n \geq 3$.
Comme les transvections engendrent $SL(E)$ (cf. 2.11) et sont toutes conjuguées (cf. 2.17), il suffit de montrer que l'une d'elles est dans N.

([1]) Le lecteur attentif aura noté que cet argument a déjà servi en I, 8.2 pour le groupe symétrique.

L'idée est la suivante : on dispose au départ d'un élément $\sigma \in N$, non trivial. On fabrique de nouveaux éléments de N comme commutateurs :

$$\text{si } \tau \in SL(E), \text{ alors } \rho = \sigma(\tau\sigma^{-1}\tau^{-1}) \in N.$$

Si τ est une transvection d'hyperplan H, $\sigma\tau\sigma^{-1}$ est une transvection d'hyperplan $\sigma(H)$, donc $\rho = (\sigma\tau\sigma^{-1})\tau^{-1}$ est produit de deux transvections et sera même une transvection si on a $\sigma(H) = H$ et $\rho \neq \mathrm{Id}$. On va donc chercher à construire un élément de N qui laisse globalement invariant un hyperplan.

Précisons maintenant tout cela : soit $\sigma \in N$, $\sigma \notin Z$. Comme σ n'est pas une homothétie, il existe $a \in E$ tel que $b = \sigma(a)$ ne soit pas colinéaire à a. Soit τ une transvection de droite $<a>$ et posons $\rho = \sigma\tau\sigma^{-1}\tau^{-1}$. Soit H un hyperplan de E contenant le plan $<a,b>$ (il en existe, puisqu'on a $n \geq 3$).

On a alors les trois propriétés suivantes :

1) $\rho \in N$ et $\rho \neq \mathrm{Id}$,

2) $\forall x \in E$, $\rho(x) - x \in H$,

3) $\rho(H) = H$.

En effet, il est clair que ρ est dans N. Si on avait $\rho = \mathrm{Id}$, on aurait $\tau = \sigma\tau\sigma^{-1}$, mais ces transvections ont respectivement pour droites $<a>$ et $<b>$ et on a $<a> \neq <b>$. Pour le point 2), on remarque (cf. 2.2.5) qu'on a $\rho(x) - x \in <a,b> \subset H$ et 3) en résulte aussitôt.

Deux éventualités sont alors possibles :

a) Il existe une transvection u, d'hyperplan H qui ne commute pas à ρ.

Alors, si on pose $v = \rho u \rho^{-1} u^{-1}$, on a $v \in N$, $v \neq \mathrm{Id}$ et v est produit des transvections u^{-1}, d'hyperplan H et $\rho u \rho^{-1}$, d'hyperplan $\rho(H) = H$, donc v est une transvection non triviale de N.

b) Sinon, ρ commute à toutes les transvections d'hyperplan H. Soit $f \in E^*$ une équation de H et u une transvection de vecteur $c \in H$ qui s'écrit :

$$u(x) = x + f(x)c.$$

On a $\rho u = u\rho$, donc, pour tout x de E :

$$\rho(x) + f(x)\rho(c) = \rho(x) + f(\rho(x))\, c.$$

Soit $x \notin H$, comme $\rho(x) - x \in H$, on a $f(\rho(x)) = f(x) \neq 0$, d'où $\rho(c) = c$. Mais ceci vaut pour tout $c \in H$, donc on a $\rho|_H = \mathrm{Id}$ et, comme ρ est de déterminant 1, ρ est déjà une transvection.

Dans les deux cas, on voit que N contient une transvection, donc $N = SL(E)$, ce qui achève la démonstration du cas $n \geq 3$.

2) *Deuxième cas : $n = 2$.*

Dans la démonstration précédente, deux points essentiels ne subsistent plus :

1) les transvections ne forment plus une seule classe de conjugaison (sauf si on a $k^* = k^{*2}$, cf. 2.18).

2) on a utilisé l'hypothèse $n \geq 3$ dans la construction d'un élément de N qui laisse invariant un hyperplan.

Notons que pour $n = 2$, l'existence d'un hyperplan stable par un $g \in SL(E)$ revient à celle d'un vecteur propre non nul, donc d'une valeur propre de g dans k. En particulier, si k est algébriquement clos, la méthode précédente s'applique sans modification.

Dans le cas général, on va construire d'abord un élément $g \in N$ qui possède une valeur propre, puis suffisamment de transvections, le tout, bien sûr, au moyen de commutateurs.

Dans tout ce qui suit on suppose $|k| \geq 7$.
Nous verrons en effet au § 5 que les cas $k = \mathbf{F}_2, \mathbf{F}_3$ sont exceptionnels et que $PSL(2, \mathbf{F}_4)$ et $PSL(2, \mathbf{F}_5)$ sont tous deux isomorphes à $\mathfrak{A}_5$, donc simples (cf. Chapitre I, 8.1).

Lemme 4.2.
On suppose $|k| \geq 7$. *Soit* $s = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, *avec* $ad - bc = 1$ *i.e.* $s \in SL(2,k)$, *avec* $c \neq 0$. *Alors il existe* $g \in SL(2,k)$ *tel que* $g^{-1}s^{-1}gs$ *admette une valeur propre* $\lambda \in k$, $\lambda \neq 0, 1, -1$.

Démonstration. On cherche sous la forme $g = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$. Soit $e_1 = (1,0)$ le premier vecteur de base. Il suffit de résoudre l'équation $g^{-1}s^{-1}gs(e_1) = \lambda e_1$, i.e. $gs(e_1) = \lambda sg(e_1)$, c'est-à dire encore :

$$a\alpha + c\beta = \lambda(a\alpha + b\gamma) \tag{1}$$

$$a\gamma + c\delta = \lambda(c\alpha + d\gamma). \tag{2}$$

Soit $\lambda \in k^{*2}$, $\lambda \neq \mp 1$ (un tel λ existe car on a $|k| \geq 7$, donc $|k^{*2}| \geq 3$). On prend alors $\gamma = 0$, $\delta = \sqrt{\lambda}$, $\alpha = 1/\delta$ et (2) est satisfaite, puis comme c est non nul, on prend $\beta = \dfrac{(\lambda - 1)a}{c\sqrt{\lambda}}$ et (1) est vérifiée.

Lemme 4.3.
Si $s \in SL(2,k)$ *a une valeur propre* $\lambda \in k^*$, *avec* $\lambda \neq \mp 1$, s *est conjugué dans* $SL(2,k)$ *de* $t = \begin{pmatrix} \lambda & 0 \\ 0 & 1/\lambda \end{pmatrix}$.

Démonstration. On note d'abord que s et t sont conjugués dans $GL(2,k)$. En effet, comme on a $\det s = 1$, les valeurs propres de s sont λ et $1/\lambda$, donc elles sont distinctes et s est diagonalisable, $s = utu^{-1}$, avec $u \in GL(E)$.
Ensuite, si on pose $d = \det u$, $d \in k^*$ et $v = \begin{pmatrix} 1/d & 0 \\ 0 & 1 \end{pmatrix}$, on voit que v commute à t et donc on a $t = v^{-1}tv = v^{-1}u^{-1}suv$ avec $\det(uv) = 1$.

Lemme 4.4.
Soit $\lambda \in k^*$, $\lambda \neq \pm 1$ *et posons* $s = \begin{pmatrix} \lambda & 0 \\ 0 & 1/\lambda \end{pmatrix}$; *soit* $\mu \in k$, $\mu \neq 0$ *et posons* $t = \begin{pmatrix} 1 & \mu \\ 0 & 1 \end{pmatrix}$.
Alors, il existe $g \in SL(2,k)$ *tel que l'on ait* $g^{-1}s^{-1}gs = t$.

Démonstration. On cherche g sous la forme $g = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ avec $\alpha\delta - \beta\gamma = 1$, vérifiant $gs = sgt$. Or, on a :

$$gs = \begin{pmatrix} \alpha\lambda & \beta/\lambda \\ \gamma\lambda & \delta/\lambda \end{pmatrix} \quad \text{et} \quad sgt = \begin{pmatrix} \alpha\lambda & \alpha\lambda\mu + \beta\lambda \\ \gamma/\lambda & \gamma\mu/\lambda + \delta/\lambda \end{pmatrix}.$$

La relation $\gamma\lambda = \gamma/\lambda$ implique $\gamma = 0$ car λ^2 est $\neq 1$. Il reste $\beta/\lambda = \alpha\lambda\mu + \beta\lambda$ et $\alpha\delta = 1$. On prend alors $\alpha = 1/\lambda - \lambda$ (de sorte que α est non nul) et $\beta = \lambda\mu$.

On peut maintenant prouver le théorème pour $n = 2$.
Soit $s \in N, s \neq \mp\mathrm{Id}$.

1) Si s a une valeur propre $\lambda \in k^*$, $\lambda \neq \mp 1$, s est conjugué dans $SL(E)$ de $s' = \begin{pmatrix} \lambda & 0 \\ 0 & 1/\lambda \end{pmatrix}$ (lemme 4.3). On en déduit que s' est dans N. Alors, pour tout $\mu \in k^*$, il existe $g \in SL(E)$ tel que $t = \begin{pmatrix} 1 & \mu \\ 0 & 1 \end{pmatrix} = g^{-1}s'^{-1}gs'$ (lemme 4.4). On a donc $t \in N$ et donc (cf. 2.11 et 2.18) $N = SL(2,k)$.

2) Si $s = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ avec $c \neq 0$, comme on a supposé $|k| \geq 7$, il existe $g \in SL(E)$ tel que $g^{-1}s^{-1}gs$ ait une valeur propre $\lambda \neq \mp 1$ (lemme 4.2). Comme $g^{-1}s^{-1}gs$ est dans N, on est ramené au cas précédent.

3) Avec les notations de 2), si on a $c = 0$ et si on n'est pas dans le cas 1), on a $s = \begin{pmatrix} \varepsilon & \mu \\ 0 & \varepsilon \end{pmatrix}$ avec $\varepsilon = \mp 1$ et $\mu \neq 0$ (car s n'est pas diagonalisable, donc a une valeur propre ε double). Soit alors $t = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, on a $t \in SL(E)$ et $tst^{-1} = \begin{pmatrix} \varepsilon & 0 \\ -\mu & \varepsilon \end{pmatrix}$, mais alors, comme tst^{-1} est dans N, on est ramené au cas 2), ce qui achève de prouver le Théorème 4.1.

Remarque 4.5. Pour une autre démonstration du Théorème 4.1, cf. Exercices 3 et 4. La démonstration proposée ci-dessus, qui repose sur des idées très simples, met en lumière l'efficacité du calcul matriciel, au moins pour ce qui concerne la dimension 2.

5. Le cas des corps finis.

Rappelons (cf. Chapitre III) que $\mathbf{F}_q$ désigne le corps à q éléments, où $q = p^\alpha$ avec p premier et $\alpha \in \mathbf{N}^*$.

Proposition 5.1.
Les cardinaux des groupes linéaires sur $\mathbf{F}_q$ sont les suivants :
1) $|GL(n,\mathbf{F}_q)| = (q^n - 1)(q^n - q)\dots(q^n - q^{n-1})$.
2) $|SL(n,\mathbf{F}_q)| = (q^n - 1)(q^n - q)\dots(q^n - q^{n-2})q^{n-1} = N$.
3) $|PGL(n,\mathbf{F}_q)| = |SL(n,\mathbf{F}_q)| = N$.
4) $|PSL(n,\mathbf{F}_q)| = N/d$ où $d = \mathrm{pgcd}\,(n, q-1)$.

Démonstration. Soit $e_1, \dots e_n$ la base canonique de $\mathbf{F}_q^n$. Si A est dans $GL(n,\mathbf{F}_q)$, $Ae_1, \dots, Ae_n$ est une base de $\mathbf{F}_q^n$ et on a ainsi une bijection de $GL(n,\mathbf{F}_q)$ sur l'ensemble des bases de $\mathbf{F}_q^n$. Pour choisir une telle base $a_1, \dots, a_n$, on peut prendre a_1 quelconque, non nul, on a donc $q^n - 1$ choix pour a_1. On doit ensuite prendre a_2 en dehors de la droite (a_1), d'où $q^n - q$ choix pour a_2. Plus généralement si $a_1, \dots, a_i$ sont choisis, a_{i+1} doit être pris en dehors du sous-espace $(a_1, \dots, a_i)$, d'où $q^n - q^i$ choix, ce qui prouve 1).
Les points 2) et 3) en résultent aussitôt puisque $\mathbf{F}_q^*$ a $q - 1$ éléments. Enfin, 4) résulte de 2.6 et du lemme suivant :

Lemme 5.2.
On a (avec les notations de III, § 4) : $|\mu_n(\mathbf{F}_q)| = d = \mathrm{pgcd}\,(n, q-1)$.

Démonstration. Par Bézout, on a $r, s \in \mathbf{Z}$ tels que $d = r(q-1) + sn$. Soit $x \in \mathbf{F}_q^*$, on a $x^{q-1} = 1$, donc, si $x \in \mu_n(\mathbf{F}_q)$, on a $x^d = x^{(q-1)r} x^{ns} = 1$. Réciproquement, si $x^d = 1$, on a *a fortiori* $x^n = 1$, et, en définitive, on a $\mu_n(\mathbf{F}_q) = \mu_d(\mathbf{F}_q)$. Mais, le polynôme $X^{q-1} - 1$ admet $q-1$ racines dans $\mathbf{F}_q$, donc $X^d - 1$ qui en est un diviseur en a d, et donc on a $|\mu_d(\mathbf{F}_q)| = d$.

Proposition 5.3.
On a les isomorphismes suivants :
1) $GL(2, \mathbf{F}_2) = SL(2, \mathbf{F}_2) = PSL(2, \mathbf{F}_2) \simeq \mathfrak{S}_3$.
2) $PGL(2, \mathbf{F}_3) \simeq \mathfrak{S}_4$; $PSL(2, \mathbf{F}_3) \simeq \mathfrak{A}_4$.
3) $PGL(2, \mathbf{F}_4) = PSL(2, \mathbf{F}_4) \simeq \mathfrak{A}_5$.
4) $PGL(2, \mathbf{F}_5) \simeq \mathfrak{S}_5$; $PSL(2, \mathbf{F}_5) \simeq \mathfrak{A}_5$.

Démonstration. On introduit l'espace projectif $\mathbf{P}(E)$ associé à E, ensemble des droites vectorielles de E. Le groupe $GL(E)$ opère sur $\mathbf{P}(E)$ de manière évidente, et les homothéties opérant trivialement, $PGL(E)$ opère aussi sur $\mathbf{P}(E)$. De plus, $PGL(E)$ opère fidèlement sur $\mathbf{P}(E)$ (cf. 2.8). Enfin, si $k = \mathbf{F}_q$ et $n = 2$, $\mathbf{P}(\mathbf{F}_q^2)$, que l'on note aussi $\mathbf{P}^1(\mathbf{F}_q)$ (la droite projective sur $\mathbf{F}_q$), a $q+1$ éléments ([2]), de sorte qu'on a un homomorphisme injectif :

$$\varphi : PGL(2, \mathbf{F}_q) \longrightarrow \mathfrak{S}_{q+1}.$$

1) Si $k = \mathbf{F}_2$, on a $\mathbf{F}_2^* = 1$, donc les groupes $GL(E)$, $SL(E)$, $PGL(E)$, $PSL(E)$ sont tous égaux, et de cardinal 6 (Proposition 5.1). Comme $PGL(2, \mathbf{F}_2)$ s'injecte dans $\mathfrak{S}_3$, il lui est isomorphe.

2) Comme on a $|PGL(2, \mathbf{F}_3)| = 24$, le même raisonnement que ci-dessus donne $PGL(2, \mathbf{F}_3) \simeq \mathfrak{S}_4$ et donc aussi $PSL(2, \mathbf{F}_3) \simeq \mathfrak{A}_4$, puisque $\mathfrak{A}_4$ est le seul sous-groupe d'indice 2 de $\mathfrak{S}_4$.

3) Si $k = \mathbf{F}_4$, corps de caractéristique 2, on a Id = −Id donc

$$SL(2, \mathbf{F}_4) = PSL(2, \mathbf{F}_4) = PGL(2, \mathbf{F}_4).$$

On a, là encore, un homomorphisme

$$\varphi : PSL(2, \mathbf{F}_4) \longrightarrow \mathfrak{S}_5,$$

injectif, et comme $PSL(2, \mathbf{F}_4)$ est de cardinal 60, donc distingué dans $\mathfrak{S}_5$, $PSL(2, \mathbf{F}_4)$ est isomorphe à $\mathfrak{A}_5$ (cf. Chapitre I, 8.5).

4) Pour $k = \mathbf{F}_5$, on a un homomorphisme injectif :

$$\varphi : PGL(2, \mathbf{F}_5) \longrightarrow \mathfrak{S}_6.$$

Comme on a $|PGL(2, \mathbf{F}_5)| = 120$, l'image de φ est un sous-groupe d'indice 6 de $\mathfrak{S}_6$, donc isomorphe à $\mathfrak{S}_5$ (cf. Chapitre I, 8.6) et il en résulte que $PSL(2, \mathbf{F}_5)$ est isomorphe à $\mathfrak{A}_5$.

Remarques 5.4.

1) Comme $\mathfrak{S}_3$ et $\mathfrak{A}_4$ ne sont pas simples, le théorème 4.1 est bien en défaut pour $n = 2$ et $k = \mathbf{F}_2$ ou $\mathbf{F}_3$. En revanche, il est vrai pour $n = 2$ et $k = \mathbf{F}_4$ ou $\mathbf{F}_5$, puisque $\mathfrak{A}_5$ est simple.

([2]) Ces éléments sont les q éléments de la droite affine, plus un point à l'infini, cf. § 5 Exercice 2.

2) Comme on a $D(\mathfrak{S}_3) = \mathfrak{A}_3$, le théorème 3.1 est en défaut pour $n = 2$, $k = \mathbf{F}_2$. Comme on a $PSL(2, \mathbf{F}_3) = \mathfrak{A}_4$, ce groupe admet un quotient de cardinal 3 (par le sous-groupe de Klein contenu dans $\mathfrak{A}_4$), donc abélien, qui est aussi un quotient de $SL(2, \mathbf{F}_3)$, et donc on a $D(SL(2, \mathbf{F}_3)) \neq SL(2, \mathbf{F}_3)$ (cf. aussi § 3 exercice 2).

3) Le plongement de $PGL(2, \mathbf{F}_5)$ dans $\mathfrak{S}_6$ nous fournit un sous-groupe isomorphe à $\mathfrak{S}_5$, mais non trivial, i.e. qui n'est pas le stabilisateur d'un point. En effet $GL(E)$, donc aussi $PGL(E)$, opère transitivement (et même triplement transitivement) sur $\mathbf{P}(E)$. Ceci nous fournit une autre démonstration pour I, 8.11.

4) Notons enfin, sans démonstration, les isomorphismes suivants, pour lesquels on renvoie à [D3] :

1) On a $PSL(2, \mathbf{F}_7) \simeq PSL(3, \mathbf{F}_2)$ (le groupe simple d'ordre 168 cf. § 5 exercice 3).

2) On a $PSL(2, \mathbf{F}_9) \simeq \mathfrak{A}_6$ et $PSL(4, \mathbf{F}_2) \simeq \mathfrak{A}_8$.

3) En revanche, $PSL(3, \mathbf{F}_4)$ qui est lui aussi de cardinal $20160 = 8!/2$, n'est pas isomorphe à $\mathfrak{A}_8$ (cf. § 5 exercice 1).

EXERCICES SUR LE CHAPITRE IV

1. Déterminant et groupe $SL(E)$.

1) Soient k un corps et $n \in \mathbf{N}^*$. On considère l'homomorphisme $e_n : k^* \longrightarrow k^*$, défini par $e_n(x) = x^n$.

a) Montrer que e_n est bijectif si et seulement si on a $\mu_n(k) = \{1\}$ et $k^* = k^{*n}$ (où k^{*n} est l'ensemble des éléments de k^* qui sont des puissances n-èmes).

b) Soit A un sous-groupe de k^* tel que $e_n(A) = k^*$. Montrer qu'on a $A = k^*$.

c) On suppose $k = \mathbf{R}$, pour quels n l'homomorphisme e_n est-il bijectif ?

d) Si $k = \mathbf{F}_q$, donner une condition nécessaire et suffisante pour que e_n soit bijectif.

---

2) On suppose qu'il existe un sous-groupe H de $GL(n,k)$ tel que l'application $(g,h) \mapsto gh$ de $SL(n,k) \times H$ dans $GL(n,k)$ soit un isomorphisme de groupes.

a) Montrer que $\det|_H : H \longrightarrow k^*$ est un isomorphisme.

b) Montrer que H est contenu dans le centre Z de $GL(n,k)$, puis que $H = Z$ (cf. Exercice 1 b)).

c) Donner une condition nécessaire et suffisante pour qu'il existe un tel H.

---

3) On suppose qu'il existe un sous-groupe H de $GL(n,k)$ tel que la projection p de $GL(n,k)$ sur $PGL(n,k)$ induise un isomorphisme de H sur $PGL(n,k)$.

a) Montrer qu'on a des isomorphismes :

$$GL(n,k) \simeq PGL(n,k) \times Z \simeq PGL(n,k) \times k^*$$

et que H est distingué dans $GL(n,k)$.

b) Montrer, en utilisant § 4 Exercices 1,2, que H contient $SL(n,k)$, sauf pour quelques cas exceptionnels que l'on examinera.

c) Montrer qu'on a alors $H = SL(n,k)$ et donner une condition nécessaire et suffisante pour qu'on ait un tel isomorphisme.

---

2. Générateurs et centres de $GL(E)$ et $SL(E)$.

1) Soit E un k-espace vectoriel de dimension n et H un hyperplan de E.

On pose :

$$D(H) = \{u \in GL(E) \mid u|_H = \mathrm{Id}_H\}, \quad T(H) = D(H) \cap SL(E).$$

a) Montrer que $D(H)$ et $T(H)$ sont des sous-groupes de $GL(E)$ et préciser leurs éléments.
Montrer que $T(H)$ est distingué dans $D(H)$, préciser le quotient (réponse : k^*) et étudier l'extension (est-elle produit semi-direct ?...)
Montrer que $T(H)$ est isomorphe au groupe additif de H, donc est commutatif. Comment k^* opère-t-il sur $H \simeq k^{n-1}$?

b) Calculer les conjugués de $D(H)$ (resp. de $T(H)$) dans $GL(E)$ (resp. $GL(E)$ et $SL(E)$).
Déterminer le normalisateur $ND(H)$ de $D(H)$ dans $GL(E)$; étudier le quotient $ND(H)/D(H)$.
Mêmes questions pour $T(H)$.

---

2) On reprend les notations de 1).

a) Soient $u \in T(H)$, $v \in D(H) - T(H)$, on suppose $u \neq \mathrm{Id}$. Montrer qu'on a $uv \neq vu$.

b) Si u, v sont deux dilatations d'hyperplan H, à quelle condition commutent-elles ?

c) Calculer les centralisateurs de $T(H)$ et $D(H)$ dans $GL(E)$.

d) A quelle condition deux transvections quelconques commutent-elles ?

---

3) a) Montrer que les groupes $(k^n, +)$ et $(k^*, \times)$ ne sont pas isomorphes, sauf si $n = 0$ et $k = \mathbf{F}_2$ (distinguer le cas de la caractéristique 2).

b) Montrer que $T(H)$ est un sous-groupe caractéristique de $D(H)$ (i.e. invariant par tout automorphisme) (utiliser les exercices 2 a) b) et 3 a)).

c) Montrer l'inclusion Int $(ND(H)) \subset$ Aut $D(H)$ et donner un exemple où cette inclusion est stricte.

---

4) Soit D une droite de E. On pose :

$$U(D) = \{\text{transvections de droite } D\} \cup \{\mathrm{Id}\}.$$

Montrer que $U(D)$ est un groupe commutatif isomorphe à $(k^{n-1}, +)$. Calculer ses conjugués et son normalisateur dans $GL(E)$.

---

5) Soit D une droite de E. On pose :

$$GL_D(E) = \{u \in GL(E) \mid u|_D = \mathrm{Id}_D\} \quad \text{et} \quad SL_D(E) = SL(E) \cap GL_D(E).$$

Soit $p : GL_D(E) \longrightarrow GL(E/D)$ l'homomorphisme canonique $u \mapsto \overline{u}$.

a) Montrer qu'on a $\mathrm{Ker}\, p = U(D)$.

b) Soit F un supplémentaire de D. On identifie $GL(F)$ à un sous-groupe de $GL_D(E)$ en prolongeant $v \in GL(F)$ par l'identité sur D.
Montrer que p induit des isomorphismes de $GL(F)$ sur $GL(E/D)$ et de $SL(F)$ sur $SL(E/D)$.

c) En déduire que p est surjectif et qu'on a des isomorphismes

$$GL_D(E) \simeq U(D) \times GL(F) \quad \text{et} \quad SL_D(E) \simeq U(D) \times SL(F).$$

d) Déduire de c) une autre démonstration du théorème 2.11.

---

6) *La méthode du pivot, ou une autre démonstration du théorème 2.11*
Pour $\lambda \in k^*$ et $i, j \in \{1, \ldots, n\}$, avec $i \neq j$, on considère :

1) la matrice de Kronecker E_{ij}, qui a pour seul terme non nul $a_{ij} = 1$,

2) la matrice $B_{ij}(\lambda) = I + \lambda E_{ij}$.

a) Soit $A \in GL(n,k)$, calculer en fonction de A, en termes de lignes ou de colonnes, les produits $B_{ij}(\lambda)A$ et $AB_{ij}(\lambda)$.

b) On pose $P_{ij} = B_{ij}(1)B_{ji}(-1)B_{ij}(1)$. Décrire $P_{ij}A$.

c) Montrer que, si A est dans $SL(n,k)$, A s'écrit comme un produit de matrices $B_{ij}(\lambda)$. (Procéder par récurrence en modifiant A par multiplication à gauche par des $B_{ij}(\lambda)$. Faire apparaître d'abord un 1 à la place a_{11}, puis des 0 aux places a_{i1}, $i \neq 1$ etc..., cf. [D] Ch. II § 1 ou [A] Th. 4.1.)

d) Utiliser c) pour donner une méthode pratique d'inversion des matrices (faire subir les mêmes manipulations à A et I pour aboutir à I et A^{-1}).

e) Utiliser les matrices $B_{ij}(\lambda)$ pour calculer le centre de $GL(n,k)$.

---

7) L'objet de cet exercice est d'étudier le nombre minimum de transvections nécessaires pour écrire un élément de $SL(E)$.

On désigne par E un k-espace vectoriel de dimension $n \geq 1$. Soit $u \in GL(E)$.

On pose $F_u = \{x \in E \mid u(x) = x\}$, espace des points fixes de u et $p_u = n - \dim F_u$. Soit $\overline{E} = E/F_u$ et soit $\overline{u}$ l'automorphisme de $\overline{E}$ induit par u.

a) Soit $u \in SL(E)$. On complète une base de F_u par des vecteurs $e_1 \ldots e_p$ avec $p = p_u$ de façon à obtenir une base $\mathcal{B}$ de E. Soit $\overline{e}_1, \ldots, \overline{e}_p$ les images de $e_1, \ldots, e_p$ dans $\overline{E}$. Montrer que ces vecteurs forment une base de $\overline{E}$. Montrer que la matrice de u dans la base $\mathcal{B}$ est de la forme :

$$\left(\begin{array}{c|c} I_{n-p} & B \\ \hline 0 & A \end{array}\right) \quad \text{avec } A \in SL(p,k).$$

Montrer que A est la matrice de $\overline{u}$ dans $\overline{e}_1, \ldots, \overline{e}_p$ et qu'on a $\overline{u} \in SL(\overline{E})$.

b) On suppose $u \in SL(E)$, $u \neq \mathrm{Id}$. Montrer que les conditions suivantes sont équivalentes :

i) $\overline{u}$ est une homothétie de rapport $\lambda \neq 1$.

ii) Dans une base convenable, u admet une matrice diagonale :

$$\begin{pmatrix} 1 & & & & \\ & \ddots & & 0 & \\ & & 1 & & \\ & & & \lambda & \\ & 0 & & & \ddots \\ & & & & & \lambda \end{pmatrix} \quad \text{avec } \lambda \neq 1.$$

Un élément $u \in SL(E)$, $u \neq \mathrm{Id}$, vérifiant i) ou ii) sera dit ***exceptionnel.***

Déterminer les éléments exceptionnels de $SL(E)$ lorsque $n = 2$. Pour p donné, existe-t-il toujours des $u \in SL(E)$, exceptionnels, tels que $p_u = p$? Étudier le cas $p = 1$.

c) Soit $u \in SL(E)$ une homothétie et τ une transvection. Calculer $p_{\tau u}$. Montrer que τu n'est pas exceptionnel.

d) On pose pour $u \in SL(E)$:

$$m_u = \inf\{m \in \mathbf{N} \mid u = \tau_1 \dots \tau_m \text{ où les } \tau_i \text{ sont des transvections de } E\}.$$

On convient que si $u = \mathrm{Id}$, on a $m_u = 0$. Montrer l'inégalité $m_u \geq p_u$.

e) Montrer que si u est exceptionnel on a $m_u \geq p_u + 1$.

f) On suppose dim $E \geq 3$. Soit $u \in SL(E)$; on suppose que u n'est pas une homothétie.

Montrer qu'il existe $a, b, c \in E$, linéairement indépendants, tels que l'on ait :

1) $u(b) = c$,

2) $u(a)$ n'est pas dans le plan $<a, c>$.

g) Soit $u \in SL(E)$ tel que $u \neq \mathrm{Id}$ et u non exceptionnel. On suppose que $\overline{u}$ n'est pas une homothétie. Montrer qu'il existe une transvection τ telle que $F_{\tau u} \supsetneq F_u$ et précisément, telle que $p_{\tau u} = p_u - 1$ (on pourra prendre $b \in E$ tel que $\overline{b}$ ne soit pas vecteur propre pour $\overline{u}$, et, si $c = u(b)$, prendre une transvection de droite $<b - c>$).

h) Avec les hypothèses de g), montrer qu'on peut, de plus, trouver τ telle que τu ne soit pas exceptionnel. (Distinguer les cas $p_u = 2$ et $p_u > 2$. Si $p_u > 2$, imposer à l'hyperplan de τ de contenir un élément a tel que $\overline{u}(\overline{a}) \notin <\overline{a}, \overline{c}>$, cf. f).)

i) Soit $u \in SL(E)$, tel que $u \neq \mathrm{Id}$ et u non exceptionnel. On suppose que $\overline{u}$ est une homothétie, donc nécessairement $\overline{u} = \mathrm{Id}$. Montrer qu'il existe une transvection τ telle que $F_{\tau u} \supsetneq F_u$, avec τu non exceptionnel.

j) Montrer par récurrence sur p_u que si u n'est pas exceptionnel on a $m_u = p_u$.

k) Montrer que si u est une homothétie, $u \neq \mathrm{Id}$, on a $m_u = n + 1$. En déduire que si u est exceptionnel on a $m_u = p_u + 1$.

---

8) Cet exercice est l'analogue de 7) mais avec les dilatations. On reprend les notations de 7).

a) Soit $u \in GL(E)$. On dit que u est une ***transgression*** si on a $u \neq \mathrm{Id}$ et $\overline{u} = \mathrm{Id}$. Montrer que ceci équivaut à $u \neq \mathrm{Id}$ et $(u - \mathrm{Id})^2 = 0$ et qu'alors u est dans $SL(E)$.

b) On pose pour $u \in GL(E)$:

$$s_u = \inf\{s \in \mathbf{N} \mid u = \tau_1 \dots \tau_s \text{ où les } \tau_i \text{ sont des dilatations}\}.$$

Montrer qu'on a $s_u \geq p_u$ et $s_u \geq p_u + 1$ si u est une trangression.

c) Soit $F \subset E$ un sous-espace et soient $x, y \in E$ dont les images sont indépendantes dans E/F. Montrer que pour tout $\lambda \in k$, distinct de 0 et 1, il existe une dilatation τ d'hyperplan H avec $F \subset H$, de rapport λ et telle que $\tau(y) = x$.

d) On suppose $k \neq \mathbf{F}_2$, $\mathbf{F}_3$. Montrer par récurrence sur p_u qu'on a $s_u = p_u$ si u n'est pas une transgression, $s_u = p_u + 1$ sinon. Étudier les cas $k = \mathbf{F}_2$, $\mathbf{F}_3$.

e) Soit $u \in GL(E)$, on pose :

$$r_u = \inf\{r \in \mathbf{N} \mid u = \tau_1 \dots \tau_r \text{ où } \tau_i \text{ est une dilatation ou une transvection}\}.$$

Montrer qu'on a $r_u = p_u$.

---

9) Soit E un k-espace vectoriel de dimension $n \geq 2$. On pose :

$T = \{u \in GL(E) \mid \mathrm{Tr}\,(u) = 0\}$ (l'ensemble des matrices de trace nulle).

a) Soient $\sigma \in \mathfrak{S}_n$, $e_1, \ldots, e_n$ une base de E et u_σ défini par $u_\sigma(e_i) = e_{\sigma(i)}$. A quelle condition a-t-on $u_\sigma \in T$?

b) Soit $\lambda \in k^*$. Montrer qu'il existe $u \in T$ tel que $\det u = \lambda$.

c) On suppose $n = 2$. Montrer que toute transvection de E est produit d'éléments de T (distinguer le cas de la caractéristique 2).

d) On suppose $n > 2$. Montrer que le résultat précédent subsiste (utiliser a)).

e) Montrer que tout élément de $GL(E)$ est produit d'éléments de T.

---

10) Soit $u \in GL(E)$, avec $\dim E = n$, et soit r un entier vérifiant $1 \leq r \leq n-1$. Montrer que si u laisse invariants tous les sous-espaces de dimension r de E, u est une homothétie.

---

11) Soit k un corps et σ un automorphisme de k. Soit E un k-espace vectoriel. Une application u de E dans E est dite σ-semi linéaire si on a :

1) $\forall x, y \in E,\ u(x+y) = u(x) + u(y)$.

2) $\forall x \in E,\ \forall \lambda \in k,\ u(\lambda x) = \sigma(\lambda) u(x)$.

a) Montrer que l'ensemble des applications semi-linéaires bijectives (pour σ variable dans $\mathrm{Aut}\,(k)$) forme un groupe qui sera noté $\Gamma L(E)$.

b) Montrer que l'application qui à u, σ-semi-linéaire, associe σ est un homomorphisme dont on calculera l'image et le noyau.

c) Montrer que $SL(E)$ est un sous-groupe distingué de $\Gamma L(E)$.

d) Montrer que le centre de $\Gamma L(E)$ est inclus dans le groupe k^* des homothéties, mais qu'il en est, en général, distinct. Déterminer ce centre dans les cas suivants : $k = \mathbf{C}$, $\mathbf{R}$, $\mathbf{Q}(\sqrt[3]{2})$, $\mathbf{F}_q$.

---

12) Montrer que $\mathbf{Q}^*/\mathbf{Q}^{*2}$ est infini (utiliser les nombres premiers).

---

13) Soit $u \in GL(E)$ diagonalisable. Calculer le centralisateur $C(u)$ de u dans $GL(E)$. Montrer que deux dilatations de même déterminant sont conjuguées par un élément de $SL(E)$. Généraliser.

---

3. Commutateurs.

1) Calculer $D(PGL(n,k))$ et $D(PSL(n,k))$ (pour les cas d'exceptions du théorème 3.1, cf. § 5).

---

2) *Étude de* $SL(2, \mathbf{F}_3)$.

a) Quels sont les cardinaux des groupes $GL(2,\mathbf{F}_3)$, $SL(2,\mathbf{F}_3)$, $PSL(2,\mathbf{F}_3)$, $PGL(2,\mathbf{F}_3)$ (cf. Proposition 5.1).

b) Montrer qu'on a $PGL(2,\mathbf{F}_3) \simeq \mathfrak{S}_4$ et $PSL(2,\mathbf{F}_3) \simeq \mathfrak{A}_4$ (cf. 5.3).

c) Déterminer le centre de $GL(2,\mathbf{F}_3)$ et celui, noté Z, de $SL(2,\mathbf{F}_3)$.

d) Montrer que $SL(2,\mathbf{F}_3)$ n'est pas isomorphe à $\mathfrak{S}_4$. On pose $G = SL(2,\mathbf{F}_3)$.

e) Déterminer les éléments d'ordre 2 de G.

f) Soit $P = PSL(2,\mathbf{F}_3)$. Calculer $D(P)$ et $P/D(P)$. En déduire qu'on a $|D(G)| \leq 8$, puis, en comparant $D(G)$, $D(P)$ et Z, montrer qu'on a $D(G) \simeq \mathbf{H}_8$.

g) Montrer qu'on a une décomposition : $G \simeq \mathbf{H}_8 \rtimes (\mathbf{Z}/3\mathbf{Z})$. Préciser l'opération.

h) Donner la liste des éléments d'ordre 2, 4, 8,... de G. Déterminer les 2-Sylow de G et retrouver ainsi le résultat de f).

i) Déterminer tous les sous-groupes distingués de G.

---

3) Avec les notations de § 2 Exercice 1, calculer le groupe des commutateurs de $D(H)$ (montrer que deux dilatations de même rapport sont conjuguées dans $D(H)$, attention au cas $k = \mathbf{F}_2$).

---

4) Avec les notations de § 2 Exercices 4 et 5, calculer le groupe des commutateurs de $GL_D(E)$ (utiliser la décomposition en produit semi-direct de $GL_D(E)$ et étudier la conjugaison dans $GL_D(E)$ des éléments de $U(D)$, attention aux cas $n = 2,3$, $k = \mathbf{F}_2$).

---

4. La simplicité de $PSL(n,k)$.

1) Déterminer tous les sous-groupes distingués de $SL(E)$ (utiliser le théorème 4.1 ; en dimension 2, on pourra chercher $u \in SL(E)$ tel que $u^2 = -\mathrm{Id}$, en dimension supérieure, on utilisera 2.17).

---

2) Déterminer tous les sous-groupes distingués de $GL(E)$ en fonction des sous-groupes de k^* (utiliser 1) et des commutateurs).

---

3) *La méthode d'Iwasawa pour démontrer la simplicité d'un groupe.*

Soit G un groupe opérant sur un ensemble X. Dans toute la suite, on suppose que G opère doublement transitivement sur X, c'est-à-dire : pour tous les couples (x_1,x_2), (y_1,y_2) de points de X avec $x_1 \neq x_2$ et $y_1 \neq y_2$, il existe $g \in G$ tel que $gx_1 = y_1$, $gx_2 = y_2$.

Soit $x \in X$ et $H \subset G$ le stabilisateur de x.

a) Montrer que G est doublement transitif si et seulement si il est transitif et si H est transitif sur $X - \{x\}$.

b) On a une bijection de G/H (classes à gauche) sur X, déduite de l'application $g \mapsto gx$, et l'opération de G sur X correspond, via cette bijection, à l'opération de G sur G/H par translation (cf. Chapitre I § 4 Exemple C).

Montrer, en utilisant cette description, qu'on a $G = H \cup HgH$ pour tout $g \in G - H$.

c) Montrer que H est un sous-groupe maximal de G (i.e. maximal parmi les sous-groupes propres de G).

d) Soient N un sous-groupe distingué de G et posons :

$$NH = \{nh \in G \mid n \in N,\ h \in H\}.$$

Montrer que NH est un sous-groupe de G. En déduire que N opère transitivement ou trivialement sur X.

e) On suppose qu'on a, pour tout $x \in X$, un sous-groupe T_x de G tel que :

i) T_x est abélien,

ii) $T_{gx} = gT_xg^{-1}$,

iii) les sous-groupes T_x engendrent G.

Soit N un sous-groupe distingué de G. On suppose que N n'opère pas trivialement sur X. Montrer qu'on a $G = NT_x$. En déduire que N contient le groupe des commutateurs $D(G)$.

---

4) *Application à $PSL(E)$*

Démontrer le Théorème 4.1 en appliquant l'exercice 3 avec $G = PSL(E)$, $X = \mathbf{P}(E)$, $T_x = U(D_x)$ (notation de §2 Exercice 4) où D_x est la droite de E d'image x dans $\mathbf{P}(E)$.

---

5. Le cas des corps finis.

1) On se propose de comparer les groupes $PSL(3, \mathbf{F}_4) = G$ et $PSL(4, \mathbf{F}_2) = H$.

a) Montrer qu'on a $|G| = |H| = 20160 = 8!/2$.

b) Montrer que $H = PSL(4, \mathbf{F}_2) = SL(4, \mathbf{F}_2)$ contient deux classes de conjugaison distinctes formées d'éléments d'ordre 2 : les transvections et une autre classe dont on donnera un représentant sous forme de Jordan.

c) Montrer que tout élément d'ordre 2 de G est image d'un élément d'ordre 2 de $SL(3, \mathbf{F}_4)$.

d) Montrer que tout élément d'ordre 2 de $SL(3, \mathbf{F}_4)$ est une transvection.

e) En déduire que G et H ne sont pas isomorphes.

---

2) *$PGL(2, k)$ ou les homographies.*

a) Montrer que l'espace projectif $\mathbf{P}(E)$ (cf. §5) est le quotient de $E - \{0\}$ par la relation d'équivalence $\mathcal{R}$ définie par :

$$x\mathcal{R}y \iff \exists \lambda \in k^*,\ x = \lambda y.$$

On notera $p : E - \{0\} \longrightarrow \mathbf{P}(E)$ la projection canonique.

b) Si $E = k^2$, on pose $\mathbf{P}(E) = \mathbf{P}^1(k)$, et on l'appelle droite projective sur k. Soit $\widehat{k} = k \cup \{\infty\}$ l'ensemble obtenu en adjoignant à k un point à l'infini. Montrer que l'application $\psi : k^2 - \{0\} \longrightarrow \widehat{k}$ définie par $\psi(x, y) = x/y$ si $y \neq 0$ et $\psi(x, 0) = \infty$ induit une bijection de $\mathbf{P}^1(k)$ sur $\widehat{k}$ de bijection réciproque φ, donnée par $\varphi(\lambda)) = p(\lambda, 1)$ pour $\lambda \in k$ et $\varphi(\infty) = p(1, 0)$.

On identifie dans toute la suite $\mathbf{P}^1(k)$ et $\widehat{k}$ par ces bijections.

c) Montrer que $PGL(2, k)$ opère fidèlement sur $\mathbf{P}^1(k)$. Les bijections de $\mathbf{P}^1(k)$ ainsi obtenues s'appellent les homographies de la droite projective $\mathbf{P}^1(k)$. Calculer, en termes de $\widehat{k}$, la bijection $\overline{u}$ induite par un élément $u = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ de $GL(2, k)$.

d) Soit $u \in GL(2,k)$. Donner une condition nécessaire et suffisante portant sur $\det u$ pour que son image $\overline{u}$ dans $PGL(2,k)$ soit dans $PSL(2,k)$. Pour des précisions sur les homographies, cf. [B] Chapitre 6.

---

3)*Les groupes simples d'ordre* 168, $PSL(2,\mathbf{F}_7)$ *et* $PSL(3,\mathbf{F}_2)$.
Le but de cet exercice est de prouver que tout groupe simple d'ordre 168 est isomorphe à $PSL(2,\mathbf{F}_7)$. En particulier, $PSL(3,\mathbf{F}_2)$ sera donc isomorphe à $PSL(2,\mathbf{F}_7)$.
Dans tout ce qui suit G désigne un groupe simple d'ordre 168.

a) Soit S l'ensemble des 7-sous-groupes de Sylow de G. Montrer qu'on a $|S| = 8$. Soient $P, Q \in S$, avec $P \neq Q$, et soit $N = N_G(P)$ le normalisateur de P dans G.
Calculer $|N|$. Montrer que P opère transitivement par conjugaison sur $S - \{P\}$.
Soit $M = N \cap N_G(Q)$. Montrer qu'on a $|M| = 3$.

b) Le groupe G opère par conjugaison sur S, on a donc un homomorphisme :

$$\varphi : G \longrightarrow \mathfrak{S}(S).$$

Montrer que φ est injectif. On pose $\overline{g} = \varphi(g)$ pour $g \in G$. Montrer que l'ordre d'un élément de G est inférieur ou égal à 12. En déduire que N n'est pas cyclique.

c) Montrer que G contient 28 sous-groupes de Sylow d'ordre 3 (considérer N pour éliminer la possibilité de 7 tels sous groupes).
Soit $H = N_G(M)$. Calculer $|H|$. Montrer que H n'est pas cyclique (compter les éléments de G d'ordres 7, 3, 6 et regarder les 2-Sylow).

d) Soit π un générateur de P. Montrer que l'application :

$$\theta : \{0, 1, \ldots, 6\} \longrightarrow S - \{P\}$$

donnée par : $i \longmapsto (\overline{\pi})^i(Q)$ est bijective.
On pose, de plus, $\theta(\infty) = P$, de sorte que θ est une bijection de $\widehat{\mathbf{F}}_7 = \mathbf{P}^1(\mathbf{F}_7)$ sur S (cf. Exercice 2 b)).

Dans toute la suite, on identifie S et $\mathbf{P}^1(\mathbf{F}_7)$ au moyen de cette bijection.

e) Montrer qu'alors, $\overline{\pi}$ est une homographie de $\mathbf{P}^1(\mathbf{F}_7)$ dont on donnera une matrice (cf. Exercice 2 c)).

f) Montrer que pour un choix convenable d'un générateur μ de M, on a : $\overline{\mu}(x) = 2x$, pour $x \in \mathbf{F}_7$, $\overline{\mu}(\infty) = \infty$ (on regardera $\mu\ \pi\ \mu^{-1}$).
Montrer que $\overline{\mu}$ est une homographie dont on précisera la matrice.

g) Soit $\tau \in H - M$. Montrer qu'on a $\tau\mu\tau^{-1} = \mu^{-1}$. En déduire que $\overline{\tau}$ est une homographie de $\mathbf{P}^1(\mathbf{F}_7)$ du type $\overline{\tau}(x) = a/x$. Préciser a et une matrice de τ.

h) Montrer que π, μ, τ engendrent G. En déduire que $\varphi(G)$ est inclus dans $PGL(2,\mathbf{F}_7)$, puis que l'on a $G \simeq PSL(2,\mathbf{F}_7)$, (cf. Exercice 2, d)).

V. FORMES SESQUILINÉAIRES,

généralités.

1. Définitions.

Soit k un corps commutatif et σ un **automorphisme** de k. On utilisera la notation exponentielle : $\sigma(\lambda) = \lambda^\sigma$, en remarquant que l'on a alors la formule :

$$(\lambda^\sigma)^\tau = \lambda^{\tau\sigma}.$$

Définition 1.1.
*Soit E un k-espace vectoriel. On appelle forme σ-**sesquilinéaire** sur E une application $f : E \times E \longrightarrow k$, qui vérifie :*
1) pour y fixé, l'application $x \mapsto f(x,y)$ est linéaire,
*2) pour x fixé, l'application $y \mapsto f(x,y)$ est σ-**semi-linéaire**, i.e. additive et telle que l'on ait :*

$$\forall \lambda \in k,\ f(x, \lambda y) = \lambda^\sigma f(x,y).$$

*Lorsque $\sigma = \mathrm{Id}_k$, on parle de forme **bilinéaire**.*

Si E est de dimension n, muni d'une base $e_1, \ldots, e_n$, la forme f est déterminée par les n^2 nombres $a_{ij} = f(e_i, e_j)$. En effet, si $A \in \mathbf{M}(n,k)$ désigne la matrice des a_{ij}, et si on pose $x = \sum_{i=1}^{n} x_i e_i,\ y = \sum_{j=1}^{n} y_j e_j$ on a :

$$f(x,y) = \sum_{i,j} a_{ij}\, x_i\, y_j^\sigma.$$

Si on convient d'appeler encore x et y les matrices colonnes des coefficients x_i, y_j, on a donc l'écriture matricielle de f [1] :

$$f(x,y) = {}^t x A y^\sigma.$$

La matrice A s'appelle la matrice de f relativement à la base (e_i).

---

[1] En identifiant le nombre λ et la matrice dont l'unique terme est λ.

Posons, pour $y \in E$, $f_y(x) = f(x, y)$; f_y est une forme linéaire sur E, donc un élément de E^*. La forme f définit donc une application semi-linéaire :

$$\begin{aligned} \overline{f} : E &\longrightarrow E^*, \\ y &\longmapsto f_y, \end{aligned}$$

dont la donnée équivaut à celle de f et dont la matrice, dans les bases (e_i) et (e_i^*) (où (e_i^*) désigne la base duale de (e_i)), n'est autre que A.

Définition 1.2.
On dit que f est ***non dégénérée*** *si l'application $\overline{f}$ est injective. Lorsque E est de dimension finie, il revient au même de dire que $\overline{f}$ est bijective. La forme f est non dégénérée si et seulement si* $\operatorname{Ker} \overline{f}$ *est nul, avec*

$$\operatorname{Ker} \overline{f} = \{y \in E \mid \forall x \in E,\ f(x, y) = 0\}.$$

Par abus de langage, le sous-espace $\operatorname{Ker} \overline{f}$ *s'appelle aussi le* ***noyau*** *de f et le* ***rang*** *de f est, par définition, celui de $\overline{f}$.*

Attention, a priori, cette définition n'est pas symétrique en x et y (cf. § 2).
En dimension finie, le critère d'injectivité par le déterminant est encore valable, même dans le cas semi-linéaire, et donc on a :

$$f \text{ non dégénérée} \iff \operatorname{Ker} \overline{f} = 0 \iff \det A \neq 0.$$

Définition 1.3.
Avec les notations précédentes le déterminant de A est appelé un ***discriminant*** *de f.*

On notera que si on fait un changement de base de matrice P, la nouvelle matrice de f est $A' = {}^tPAP^\sigma$ et qu'on a donc $\det A' = \delta\delta^\sigma \det A$ avec $\delta = \det P$, $\delta \neq 0$. Le discriminant de f n'est donc défini qu'à un élément du type $\delta\delta^\sigma$ près. Un tel élément s'appelle une ***norme***. Lorsque $\sigma = \mathrm{Id}_k$, c'est simplement un ***carré***.

2. Formes réflexives.

L'un des intérêts essentiels des formes sesquilinéaires est de permettre la définition d'une relation ***d'orthogonalité*** sur E, notée $x \perp y$, par la formule :

$$x \perp y \iff f(x, y) = 0.$$

Mais, bien sûr, on attend de cette relation qu'elle soit symétrique, ce qui nous conduit à la définition suivante :

Définition 2.1.
Soit f une forme sesquilinéaire sur E, f est dite ***réflexive*** [2] *si on a, pour tous $x, y \in E$:*

$$f(x, y) = 0 \iff f(y, x) = 0.$$

Les notions suivantes fournissent des exemples de formes réflexives.

[2] *La terminologie n'est peut-être pas très heureuse, mais elle est traditionnelle et, de toute manière, provisoire (cf. Théorème 2.9).*

Définition 2.2.
Soit f une forme **bilinéaire** *(i.e. relative à $\sigma = \mathrm{Id}_k$), f est dite* **symétrique** *(resp. antisymétrique) si pour tous $x, y \in E$, on a :*

$$f(x,y) = f(y,x), \quad (\text{resp. } f(x,y) = -f(y,x)).$$

Une forme symétrique ou antisymétrique est évidemment réflexive.

Définition 2.3.
Si f est une forme bilinéaire symétrique, l'application $q : E \longrightarrow k$ définie par $q(x) = f(x,x)$, est appelée forme **quadratique** *associée à f. Réciproquement, f est la forme* **polaire** *de q, cf. 4.3.*

Définition 2.4.
Soit f une forme σ-sesquilinéaire, f est dite **alternée** *si et seulement si on a, pour tout x de E, $f(x,x) = 0$.*

Lemme 2.5.
Si f est alternée et $f \neq 0$, on a les propriétés suivantes :
1) $\sigma = \mathrm{Id}_k$,
2) f est antisymétrique.

Démonstration.
Pour $x, y \in E$, on a : $f(x+y, x+y) = 0 = f(x,x) + f(x,y) + f(y,x) + f(y,y)$ d'où $f(x,y) = -f(y,x)$.
Soient maintenant $x, y \in E$ tels que $f(y,x) \neq 0$ et soit $\lambda \in k$. On a $f(\lambda x, y) = \lambda f(x,y) = -\lambda f(y,x)$, mais aussi : $f(\lambda x, y) = -f(y, \lambda x) = -\lambda^\sigma f(y,x)$ donc $\lambda = \lambda^\sigma$ pour tout $\lambda \in k$, d'où $\sigma = \mathrm{Id}$ et f est antisymétrique.

Remarque 2.6. Réciproquement, si k n'est pas de caractéristique 2, f antisymétrique implique f alternée car on a $f(x,x) = -f(x,x) = 0$.
En caractéristique 2, on notera que symétrique équivaut à antisymétrique.

Notons enfin un exemple avec $\sigma \neq \mathrm{Id}$:

Définition 2.7.
Une forme σ-sesquilinéaire, avec $\sigma \neq \mathrm{Id}$, est dite **hermitienne** *si on a, pour tous $x, y \in E$:*

$$f(y,x) = f(x,y)^\sigma.$$

Remarques 2.8.
1) Pour $f \neq 0$, cette définition entraîne que σ est une involution i.e. $\sigma^2 = \mathrm{Id}_k$. En effet, on a $f(x,y) = f(x,y)^{\sigma^2}$ et tout élément de k peut s'écrire sous la forme $f(x,y)$ (si $f(u,v) = a \neq 0$, $f(\lambda u, v) = a\lambda$ décrit k tout entier).
2) On définirait comme en 2.3, une forme quadratique hermitienne associée à f.

En fait, ces exemples décrivent à peu près toutes les formes réflexives, précisément :

Théorème 2.9.
On suppose E de dimension finie n, $n \geq 2$. Soit f une forme σ-sesquilinéaire, non dégénérée, réflexive. Alors :
1) σ est une involution, $\sigma^2 = \mathrm{Id}_k$,

2) a) Si $\sigma = \mathrm{Id}_k$, f est une forme bilinéaire symétrique ou antisymétrique,
b) Si $\sigma \neq \mathrm{Id}_k$, il existe $\alpha \in k^*$ tel que αf soit hermitienne.

Remarques 2.10.

1) En dimension 1, toute forme est réflexive.

2) Lorsque $\sigma \neq \mathrm{Id}_k$, f n'est pas hermitienne en général : si $E = \mathbf{C}^2$ avec pour σ la conjugaison et pour f la forme $x\overline{x'} + y\overline{y'}$, λf n'est hermitienne que pour λ réel.

3) Si f est dégénérée, le théorème reste vrai, sauf si $\dim E = \dim \mathrm{Ker}\, f + 1$, auquel cas f peut être quelconque, en particulier σ n'est pas nécessairement involutif dans ce cas.

Démonstration (de 2.9).
Soit $x \in E$, $x \neq 0$. Posons :

$$H_x = \{y \in E \mid f_x(y) = f(y,x) = 0\}.$$

Comme f_x est une forme linéaire non nulle, H_x est un hyperplan de E.
Comme f est réflexive on a les équivalences :

$$f(y,x) = 0 \Longleftrightarrow f(x,y) = 0 \Longleftrightarrow f(x,y)^{\sigma^{-1}} = 0.$$

Posons $g_x(y) = f(x,y)^{\sigma^{-1}}$, alors g_x est linéaire. En effet, si $\lambda \in k$, on a :

$$g_x(\lambda y) = f(x,\lambda y)^{\sigma^{-1}} = (\lambda^\sigma)^{\sigma^{-1}} f(x,y)^{\sigma^{-1}} = \lambda g_x(y).$$

Comme on a $H_x = \mathrm{Ker}\, f_x = \mathrm{Ker}\, g_x$, les formes linéaires f_x et g_x sont proportionnelles : $g_x = \alpha(x) f_x$, avec $\alpha(x) \in k^*$.
Ceci vaut pour tout x de E (pour $x = 0$ c'est trivial). On a donc deux applications, $\overline{f}, \overline{g} : E \longrightarrow E^*$ définies par $\overline{f}(x) = f_x$, $\overline{g}(x) = g_x$, avec $\overline{f}$ σ-semi-linéaire et $\overline{g}$, σ^{-1}-semi-linéaire. On pose alors, $u = (\overline{f})^{-1} \circ \overline{g}$, u est σ^{-2}-semi-linéaire de E dans E, et on a, pour tout $x \in E$:

$$u(x) = (\overline{f})^{-1}[\alpha(x) f(x)] = \alpha(x)^{\sigma^{-1}} x.$$

Mais on a alors le lemme suivant que l'on comparera à IV, 2.8 :

Lemme 2.11.
Soit $u : E \longrightarrow E$, une application τ-semi-linéaire, on suppose :
1) $\dim E \geq 2$,
2) $\forall x \in E$, $\exists \lambda \in k$ tel que $u(x) = \lambda x$,
alors u est une homothétie. En particulier, u est linéaire et on a $\tau = \mathrm{Id}$.

Démonstration. Soient $x, y \in E$, non colinéaires.
On a $u(x) = \lambda x$, $u(y) = \mu y$, $u(x+y) = \nu(x+y)$ d'où $\nu x + \nu y = \lambda x + \mu y$, et puisque x et y sont indépendants, on en déduit $\lambda = \mu = \nu$.
Si x et y sont colinéaires, on passe par l'intermédiaire d'un z non colinéaire (c'est possible car on a $\dim E \geq 2$).

Remarque 2.12. Attention, le résultat ne subsiste pas en dimension 1, si $E = ke$ l'application semi-linéaire définie par $u(\lambda e) = \lambda^\tau e$ convient, quel que soit τ.

Revenons au théorème, comme u est non nul, on a donc prouvé :

a) $\sigma^{-2} = \mathrm{Id}$, donc $\sigma^2 = \mathrm{Id}$, σ est bien une involution,

b) $\overline{g} = \lambda \overline{f}$, $\lambda \in k^*$, c'est-à-dire : $\forall x, y \in E$, $f(x,y)^{\sigma^{-1}} = \lambda f(y,x)$, ou encore $f(x,y) = \mu f(y,x)^\sigma$ avec $\mu = \lambda^\sigma$. Deux cas sont alors possibles :

1) $\sigma = \mathrm{Id}_k$; f est bilinéaire.
Soient x, y tels que $f(x,y) \neq 0$, on a $f(x,y) = \mu f(y,x) = \mu^2 f(x,y)$, d'où l'on déduit $\mu^2 = 1$. Si $\mu = 1$, f est symétrique, si $\mu = -1$, f est antisymétrique.

2) $\sigma^2 = \mathrm{Id}_k$, $\sigma \neq \mathrm{Id}_k$.
Alors, f n'est pas alternée (Lemme 2.5), donc, il existe $x_0 \in E$ tel que $\alpha = f(x_0, x_0)$ soit non nul. On a $\alpha = \mu\alpha^\sigma$ d'après ce qui précède.
Si on pose $g = \alpha^{-1} f$, on a : $\alpha g(x,y) = \mu\alpha^\sigma g(y,x)^\sigma$ et donc $g(x,y) = g(y,x)^\sigma$ pour tous x, y, de sorte que g est hermitienne.

Remarques 2.13.

1) Matriciellement, si A est une matrice de f, f est symétrique (resp. antisymétrique) si on a $\sigma = \mathrm{Id}$ et ${}^tA = A$ (resp. ${}^tA = -A$), f est hermitienne si $\sigma \neq \mathrm{Id}$ et ${}^tA = A^\sigma$.

2) Nous avons déjà rencontré une involution, à savoir la conjugaison sur $\mathbf{C}$. La proposition suivante montre qu'en caractéristique différente de 2, les involutions sont toutes de ce type. Il s'agit d'un cas très particulier du calcul des extensions cycliques (le célèbre théorème 90 de Hilbert, cf. [L] Ch. VIII §6).

Proposition 2.14.
Soit k un corps, $\sigma \in \mathrm{Aut}\, k$ tel que $\sigma^2 = \mathrm{Id}$, $\sigma \neq \mathrm{Id}$. Alors, il existe un sous-corps k_0 de k et un élément $a \in k$ tels que :
1) $[k : k_0] = 2$, $k = k_0[a] = \{\lambda + \mu a \mid \lambda, \mu \in k\}$,
2) si on suppose $\mathrm{car}\,(k) \neq 2$, a vérifie une équation $a^2 = \alpha$, avec $\alpha \in k_0$ et σ est donné par $\sigma|_{k_0} = \mathrm{Id}$ et $\sigma(a) = -a$,
3) si on suppose $\mathrm{car}\,(k) = 2$, a vérifie une équation $a^2 + a + \alpha = 0$, avec $\alpha \in k_0$ et σ est donné par $\sigma|_{k_0} = \mathrm{Id}$, $\sigma(a) = 1 + a$.

Remarques 2.15.

1) On notera que σ est alors l'unique automorphisme de k, distinct de Id_k, qui vérifie $\sigma|_{k_0} = \mathrm{Id}_{k_0}$.

2) Lorsque k est le corps des nombres complexes et σ la conjugaison, on a $k_0 = \mathbf{R}$, et par exemple, $a = i$, $\alpha = -1$.

Démonstration (de 2.14). Soit $k_0 = \{x \in k \mid x^\sigma = x\}$, k_0 est un sous-corps de k, distinct de k puisque $\sigma \neq \mathrm{Id}_k$. Soit $a \in k - k_0$. Il faut distinguer deux cas selon que k est de caractéristique 2 ou non :

1) On suppose $\mathrm{car}\,(k) \neq 2$. L'élément $a - a^\sigma$ est anti-invariant (i.e. vérifie $(a - a^\sigma)^\sigma = -(a - a^\sigma)$) et, quitte à remplacer a par $a - a^\sigma$, on peut donc supposer $a^\sigma = -a$. L'élément $\alpha = -aa^\sigma$ est alors un invariant, donc on a $\alpha \in k_0$ et $a^2 = \alpha$. D'autre part si x est anti-invariant, x/a est invariant, donc $x = \lambda a$, $\lambda \in k_0$.
Soit alors $b \in k$, on écrit $b = \dfrac{b + b^\sigma}{2} + \dfrac{b - b^\sigma}{2}$ et donc, puisque $b + b^\sigma$ est invariant et $b - b^\sigma$ anti-invariant, on a $b = \lambda + \mu a$ avec $\lambda, \mu \in k_0$ et on a donc $k = k_0[a] = \{\lambda + \mu a ;\ \lambda, \mu \in k_0\}$ avec toutes les propriétés annoncées.

2) On suppose $\mathrm{car}\,(k) = 2$. Quitte à remplacer a par $\dfrac{a}{a + a^\sigma}$, on peut supposer $a^\sigma = 1 + a$. L'élément $\alpha = aa^\sigma$ est dans k_0 et vérifie $a^2 + a + \alpha = 0$. Soit $b \in k$, $b \notin k_0$, et $c = \dfrac{b}{b + b^\sigma}$, on a $c^\sigma = 1 + c$, d'où $(a + c)^\sigma = 1 + a + 1 + c = a + c$, de sorte que $a + c$ est dans k_0. Il en résulte que c, donc aussi b, appartiennent à $k_0[a] = \{\lambda + \mu\alpha \mid \lambda, \mu \in k_0\}$.

3. Sous-espaces orthogonaux et isotropes.

Dans tout ce paragraphe, E désigne un k-espace vectoriel et f une forme sesquilinéaire sur E, ***non dégénérée*** et que nous supposerons soit ***symétrique***, soit ***hermitienne***, soit ***alternée***, donc (cf. § 2) réflexive. On peut donc adopter la définition suivante :

Définition 3.1.
Soient $x, y \in E$. On dit que x et y sont **orthogonaux** *(relativement à f) si on a $f(x,y) = 0$. On écrit alors $x \perp y$. On dit que deux parties A, B de E sont orthogonales, et on note $A \perp B$, si on a :*

$$\forall x \in A,\ \forall y \in B,\ x \perp y.$$

Définition 3.2.
Soit A une partie de E. ***L'orthogonal de*** *A est la partie $A^\perp$:*

$$A^\perp = \{x \in E \mid \forall a \in A,\ a \perp x\}.$$

C'est un sous-espace vectoriel de E.

Remarque 3.3. L'application $A \mapsto A^\perp$ est décroissante.

Proposition 3.4.
Si E est un espace vectoriel de dimension n et si V est un sous-espace de dimension p de E, on a $\dim V^\perp = n - p$.

Démonstration.
Soit $\overline{f} : E \longrightarrow E^*$, définie comme au § 1. Soit $e_1, \ldots, e_p$ une base de V, complétée par $e_{p+1}, \ldots, e_n$ en une base de E et soit $F = \overline{f}(V^\perp) = \{g \in E^* \mid g|_V = 0\}$.
Comme $\overline{f}$ est semi-linéaire bijective, il suffit de prouver qu'on a $\dim F = n - p$.
Soit (e_i^*), pour $i = 1, \ldots, n$, la base duale de (e_i) et soit $u \in E^*$, on a $u = \sum_{i=1}^{n} \lambda_i e_i^*$ et u est dans F si et seulement si $u(e_i) = 0$ pour $i = 1, \ldots, p$, ce qui signifie que les λ_i, pour $i = 1, \ldots, p$, sont nuls, donc que $e_{p+1}^*, \ldots, e_n^*$ est une base de F, d'où la conclusion.

Corollaire 3.5.
On suppose E de dimension finie. Soient V, W des sous-espaces de E, alors on a les formules suivantes :
1) $V^{\perp\perp} = V$,
2) $(V + W)^\perp = V^\perp \cap W^\perp$,
3) $(V \cap W)^\perp = V^\perp + W^\perp$.

Démonstration.
Il est clair que l'on a $V \subset V^{\perp\perp}$, d'où l'égalité, pour une raison de dimension.
L'égalité 2) est évidente, sans l'hypothèse de finitude sur E. L'égalité 3) résulte de 2) appliquée à $V^\perp$ et $W^\perp$ et de 1).

Remarque 3.6. Les égalités 1) et 3) ne subsistent plus si E est de dimension infinie, cf. Exercices 2 et 3.

Définition 3.7.
Soit $x \in E$, $x \neq 0$. On dit que x est ***isotrope*** *si et seulement si on a $f(x,x) = 0$.*

Définition 3.8.
Un sous-espace $V \subset E$ est dit ***isotrope*** *(ou singulier) si on a $V \cap V^\perp \neq \{0\}$, c'est-à-dire s'il existe $x \in V$, tel que pour tout $y \in V$ on ait $f(x,y) = 0$.*
Il revient au même de dire que $f|_V$ est dégénérée.

Remarque 3.9. Si $V = kx$, avec $x \neq 0$, x est isotrope si et seulement si V l'est.

Définition 3.10.
Soit V un sous-espace de E. On dit que V est ***totalement isotrope*** *si on a $V \subset V^\perp$, ou encore, si f, en restriction à V, est nulle.*

Remarque 3.11.
Si $\dim E = n$, on a vu en 3.4 qu'on a $\dim V^\perp = n - \dim V$ et donc, si V est totalement isotrope, on a $\dim V \leq n/2$.

Définition 3.12.
On appelle ***indice*** *de f l'entier ν, maximum des dimensions des sous-espaces totalement isotropes. Si $\dim E = n$, on a donc $\nu \leq n/2$.*

Remarques 3.13.
1) S'il existe un vecteur isotrope $x \in E$, $x \neq 0$, le sous-espace kx est totalement isotrope et donc l'indice ν de la forme est ≥ 1.
2) Si $\nu = 0$, la forme f est dite anisotrope (ou définie). La relation $f(x,x) = 0$ implique alors $x = 0$.
3) Si V est isotrope, $V \cap V^\perp$ est totalement isotrope.
4) Si V est non isotrope, on a $V \cap V^\perp = \{0\}$. En particulier, si E est de dimension finie, il est somme directe : $E = V \oplus V^\perp$ et on écrit alors $E = V \perp V^\perp$.
5) La définition de l'indice donnée ci-dessus est assez peu commode. Nous verrons (cf. VIII, 4.5) qu'en fait, les sous-espaces totalement isotropes maximaux ont tous même la dimension (qui vaut donc ν) ce qui fournit une meilleure définition de l'indice.

4. Groupes unitaire, orthogonal, symplectique.

Définition 4.1.
Soit f une forme sesquilinéaire sur E, non dégénérée que nous supposons soit hermitienne, soit symétrique, soit alternée.
On appelle ***isométries de*** *E (relativement à f) les automorphismes $u \in GL(E)$ qui vérifient :*

$$\forall x, y \in E,\ f(u(x), u(y)) = f(x,y).$$

Les isométries forment un sous-groupe de $GL(E)$ appelé :
1) $U(f)$ si f est hermitienne : ***groupe unitaire,***
2) $O(f)$ si f est symétrique : ***groupe orthogonal,***
3) $Sp(f)$ si f est alternée : ***groupe symplectique.***
Lorsque E est de dimension finie n sur k on note ces groupes $U_n(k,f)$, $O_n(k,f)$, $Sp_n(k,f)$.

Remarque 4.2. La définition des isométries s'étend de manière évidente dans les deux cas suivants :

1) si f est dégénérée,

2) si $u : (E, f) \longrightarrow (E', f')$ est une application linéaire entre deux espaces différents et/ou munis de deux formes différentes.

Proposition 4.3.
Si f est symétrique (resp. hermitienne) et si on a $\operatorname{car}(k) \neq 2$, un élément u de $GL(E)$ est une isométrie si et seulement si il conserve la forme quadratique (resp. quadratique hermitienne) q attachée à f, i.e. si on a :

$$\forall x \in E,\ q\big(u(x)\big) = f\big(u(x), u(x)\big) = q(x) = f(x, x).$$

Démonstration. L'un des sens est clair, l'autre résulte des formules :

1) pour f symétrique :

$$f(x, y) = \frac{1}{2}\,[q(x + y) - q(x) - q(y)].$$

2) pour f hermitienne, avec les notations de 2.14 :

$$f(x, y) = \frac{1}{4}\,[q(x + y) - q(x - y)] - \frac{1}{4a}\,[q(x + ay) - q(x - ay)].$$

Dans toute la suite, nous supposerons toujours $\operatorname{car}(k) \neq 2$ ***dans l'étude des formes symétriques ou hermitiennes.***

On notera éventuellement $U(q)$ ou $O(q)$ le groupe des isométries, q étant la forme quadratique associée à f.

Remarques 4.4.

1) Si $\alpha \in k^*$, f et αf ont même groupe d'isométries. Le théorème 2.9 montre donc que les groupes définis ci-dessus sont ceux de toutes les formes réflexives, au moins en dimension finie.

2) *Forme matricielle :*
En dimension finie, si f a pour matrice A et u pour matrice U, u est une isométrie si et seulement si on a ${}^tUAU^\sigma = A$.

3) Soit $u \in GL(E)$ et soit $e_1, \ldots, e_n$ une base de E. Alors u est une isométrie si et seulement si on a pour tous i, j :

$$f(u(e_i), u(e_j)) = f(e_i, e_j).$$

4) Si u est une homothétie de rapport λ, u est une isométrie si et seulement si on a $\lambda\lambda^\sigma = 1$. En particulier, les homothéties Id et $-\mathrm{Id}$ sont des isométries, et ce sont les seules dans le cas $\sigma = \mathrm{Id}$.

5) Si $\dim E = n$, la relation matricielle donne :

$$(\det u)(\det(u^\sigma)) = (\det u)(\det u)^\sigma = 1.$$

En particulier si $\sigma = \mathrm{Id}$, on a $\det u = \mp 1$. Cela nous conduit à la définition suivante :

Définition 4.5.
Le sous-groupe de $U(f)$ (resp. $O(f)$) formé des isométries de déterminant 1 est distingué et s'appelle groupe ***spécial*** *unitaire (resp. orthogonal), ou encore groupe unitaire (resp. orthogonal)* ***positif.***
Il est noté $SU(f)$ (resp. $SO(f)$) ou encore $U^+(f)$ (resp. $O^+(f)$).

Les éléments de $O^+(f)$ s'appellent des ***isométries positives***, *ou des* ***rotations***. *Ceux de $O^-(f)$ (i.e. de déterminant -1) des* ***isométries négatives***, *ou des retournements (mais cette dernière terminologie n'est pas universellement admise).*

Dans le cas symplectique, on démontre que tous les éléments $u \in Sp(f)$ sont de déterminant 1 (cf. [D]).

Un exemple d'isométries : les symétries.

Dans la suite du § 4 nous supposons f symétrique, soit q la forme quadratique associée et posons $\dim E = n$.
Rappelons que si un élément $u \in GL(E)$ vérifie $u^2 = \mathrm{Id}$, il existe deux sous-espaces $E^+(u)$ et $E^-(u)$ (ou simplement E^+ et E^-) qui vérifient :

1) $E = E^+ \oplus E^-$,

2) $u|_{E^+} = \mathrm{Id}_{E^+}$, $u|_{E^-} = -\mathrm{Id}_{E^-}$.

Dans une base $e_1, \ldots, e_n$ telle que $e_1, \ldots, e_p \in E^+$, $e_{p+1}, \ldots, e_n \in E^-$, u a donc pour matrice :

$$U = \begin{pmatrix} 1 & & & & \\ & \ddots & & 0 & \\ & & 1 & & \\ & & & -1 & \\ & 0 & & & \ddots \\ & & & & & -1 \end{pmatrix}$$

Si on a $u^2 = \mathrm{Id}$ et $u \neq \mathrm{Id}$, on dit que u est une involution (ou encore une ***symétrie***).
Si $\dim E^- = 1$ (resp. 2) on dit que u est une ***réflexion*** (resp. un ***renversement***).

Nous allons déterminer les symétries orthogonales, c'est-à-dire celles qui sont dans $O(q)$.

Proposition 4.6.
Soit $u \in GL(E)$ avec $u^2 = \mathrm{Id}$, et soient E^+ et E^- les sous-espaces associés à u. Alors, u est une isométrie si et seulement si E^+ et E^- sont orthogonaux.
On a alors $E^+ = (E^-)^\perp$, $E^- = (E^+)^\perp$, de sorte que E^+ et E^- sont non isotropes.
Réciproquement, si $F \subset E$ est un sous-espace non isotrope, il existe une et une seule symétrie orthogonale u telle que $F = E^+(u)$.

Démonstration.
Supposons que u soit une isométrie. Soient $x \in E^+$, $y \in E^-$. On a $f(u(x), u(y)) = f(x, -y) = -f(x, y) = f(x, y) = 0$, donc E^+ et E^- sont orthogonaux.
En sens inverse, si E^+ et E^- sont orthogonaux, soient $x, y \in E$ et soient $x = x' + x''$, $y = y' + y''$ leurs décompositions sur E^+ et E^-. On a :

$$u(x) = x' - x'', \; u(y) = y' - y''$$

et donc

$$f(x, y) = f(x', y') + f(x'', y'') = f(u(x), u(y)),$$

de sorte que u est dans $O(f)$.
Enfin, si F est non isotrope, on définit u par $u|_F = \mathrm{Id}_F$, $u|_{F^\perp} = -\mathrm{Id}_{F^\perp}$, u est une isométrie involutive qui vérifie $E^+(u) = F$.

Remarques 4.7.

1) Si u est une réflexion orthogonale, u est déterminée soit par son hyperplan $H = E^+(u)$, soit par sa droite $D = E^-(u)$. On la notera éventuellement $u = \tau_H$ ou τ_D ou encore τ_x si $x \in D$, $x \neq 0$. Il y a donc bijection entre les réflexions orthogonales, les hyperplans non isotropes et les droites non isotropes. Notons enfin que si u est une réflexion u est dans $O^-(q)$.

2) Si u est un renversement orthogonal, u est déterminé par son plan $P = E^-(u)$. Il y a bijection entre les renversements orthogonaux et les plans non isotropes. Si u est un renversement, u est dans $O^+(q)$.

3) Si x est non isotrope, on a, pour $y \in E$:

$$\tau_x(y) = y - 2\,\frac{f(x,y)}{f(x,x)}\,x.$$

(Il suffit de vérifier la formule pour $y = x$ et $y \perp x$).

Proposition 4.8.
Soit $F \subset E$ un sous-espace non isotrope et τ_F la symétrie orthogonale par rapport à F (c'est-à-dire vérifiant $E^+(\tau_F) = F$). Soit $u \in O(q)$. Alors $u\tau_F u^{-1}$ est la symétrie orthogonale par rapport à $u(F)$, i.e. $\tau_{u(F)}$. De plus, on a $E^-(\tau_{u(F)}) = u(E^-(\tau_F))$.

Démonstration.
Comme τ_F est une involution, il en est de même de $u\tau_F u^{-1}$. On vérifie aussitôt la formule $E^+(u\tau_F u^{-1}) = u(E^+(\tau_F))$ et de même pour E^-, d'où le résultat. On notera qu'on a là un nouvel exemple du principe de conjugaison, cf. I 4.10

5. Les similitudes.

Définition 5.1.
Soit f une forme sesquilinéaire non dégénérée sur E, hermitienne, symétrique ou alternée. Soit $u \in GL(E)$ et $\mu \in k^*$. On dit que u est une ***similitude*** *de* ***multiplicateur*** *μ (relativement à f) si on a :*

$$\forall x, y \in E,\ f(u(x),\ u(y)) = \mu f(x,y).$$

Le groupe des similitudes est noté $GU(f)$, $GO(f)$ ou $GS_p(f)$ dans les trois cas habituels. Le multiplicateur μ est bien déterminé par u et l'application qui à u associe son multiplicateur est un homomorphisme de groupes à valeurs dans k^*, dont le noyau est le groupe des isométries de f, de sorte qu'on a des suites exactes du type :

$$1 \longrightarrow U(f) \longrightarrow GU(f) \xrightarrow{\mu} k^*,$$

mais, attention, μ n'est pas surjectif en général (cf. Exercice 2).
On peut aussi définir des semi-similitudes qui sont seulement semi-linéaires (cf. [D] Ch. I § 9).

Parmi les similitudes, il y a, outre les isométries, les homothéties. En effet, l'homothétie de rapport λ est une similitude de multiplicateur $\lambda\lambda^\sigma$. Attention, cependant, en général, les isométries et les homothéties n'engendrent pas le groupe des similitudes (cf. Exercice 3).

Forme matricielle.

Si A est la matrice de f, U celle de u, l'application u est une similitude de multiplicateur μ si et seulement si on a ${}^tUAU^\sigma = \mu A$. On se reportera au § 7 et § 7 Exercice 2 pour des caractérisations géométriques des similitudes.

6. Bases orthogonales ; classification des formes sesquilinéaires.

Définition 6.1.
Soit f une forme sesquilinéaire réflexive (éventuellement dégénérée). Une base $e_1, \ldots, e_n$ de E est dite **orthogonale** *pour f si on a :*

$$\forall i, j,\ i \neq j \implies f(e_i, e_j) = 0.$$

Dans une telle base, la matrice de f est diagonale :

$$A = \begin{pmatrix} \gamma_1 & & & \\ & \ddots & 0 & \\ & 0 & \ddots & \\ & & & \gamma_n \end{pmatrix}$$

avec $\gamma_i = f(e_i, e_i)$.

Remarque 6.2. La forme f est non dégénérée si et seulement si tous les γ_i sont non nuls.

Théorème d'existence 6.3.
On suppose car $(k) \neq 2$ et f symétrique ou hermitienne sur E de dimension finie. Alors, il existe une base orthogonale pour f. Avec les notations de 6.1, on a alors $\gamma_i = \gamma_i^\sigma$ pour tout i.

Démonstration.

1) On se ramène au cas où f est non dégénérée en prenant une base de $N = \operatorname{Ker} f$, puis un supplémentaire V de N (on a alors $V \perp N$).

2) On raisonne par récurrence sur n. Pour $n = 1$, le théorème est clair. Comme f n'est pas alternée (cf. 2.5), il existe e_1 non isotrope. Soit $H = (ke_1)^\perp$, de sorte que l'on a $E = ke_1 \perp H$. Grâce à l'hypothèse de récurrence, H possède une base orthogonale pour f, soit $e_2, \ldots, e_n$. Alors, il est clair que $e_1, \ldots, e_n$ convient.

Les bases orthogonales vont nous permettre, dans certains cas, de classer les formes sesquilinéaires.

Définition 6.4.
Soient f, f' deux formes sesquilinéaires sur E. On dit que f et f' sont **équivalentes** *s'il existe $u \in GL(E)$ tel que l'on ait :*

$$\forall x, y \in E,\ f'(x, y) = f(u(x), u(y)).$$

Si les matrices de f et f' dans une base de E sont A et A', il revient au même de dire qu'il existe une matrice P inversible telle que $A' = {}^tPAP^\sigma$.
On peut encore dire que dans une base $e_1, \ldots, e_n$ (resp. $e'_1, \ldots, e'_n$) f (resp. f') a pour matrice A (resp. la même matrice A). Nous noterons $f \sim f'$ lorsque f et f' sont équivalentes.

Remarques 6.5.

0) La définition 6.4 s'étend immédiatement au cas de deux formes sur des espaces E, E' différents.

1) Si f et f' sont équivalentes, les groupes d'isométries correspondants sont isomorphes. En effet, si $f'(x,y) = f(u(x), u(y))$, on a :

$$v \in O(f') \Longleftrightarrow uvu^{-1} \in O(f),$$

donc $O(f)$ et $O(f')$ sont conjugués dans $GL(E)$.

2) En vertu de la remarque 4.4.1, on a donc le même résultat si $f \sim \lambda f'$ pour $\lambda \in k^*$.

Le problème de la classification des formes consiste à trouver les classes d'équivalence pour la relation $\sim$ et si possible à les caractériser par des invariants numériques. Sur un corps quelconque c'est en général un problème très difficile (sauf dans le cas alterné, cf. [D] Ch. I § 8).

On a des ***conditions nécessaires*** d'équivalence sur f, f' ; si $f \sim f'$ on a :

1) $\operatorname{rg}(f) = \operatorname{rg}(f')$ (égalité des rangs),

2) $\nu(f) = \nu(f')$ (égalité des indices),

3) si $\sigma = \mathrm{Id}$ et si on note $\delta(f)$ un discriminant de f, on a $\delta(f) = \delta(f')$ dans k^*/k^{*2}.

Mais ces conditions ne sont pas suffisantes en général (penser aux formes $x^2 + y^2$ et $-x^2 - y^2$ sur $\mathbf{R}$).

Nous allons résoudre entièrement le problème de la classification dans les cas suivants :

1) Pour f symétrique : k algébriquement clos, $k = \mathbf{R}$, $k = \mathbf{F}_q$.

2) Pour f hermitienne : $k = \mathbf{C}$, k fini.

On pourra consulter [S1] Ch. IV pour le cas $k = \mathbf{Q}$.

Théorème 6.6.

1) On suppose k algébriquement clos et $\dim E = n$. Alors, toutes les formes quadratiques q non dégénérées sur E sont équivalentes. Dans une base convenable, elles ont pour matrice l'identité et, si $x \in E$ s'écrit $x = (x_1, \ldots, x_n)$ dans cette base, on a $q(x) = x_1^2 + \ldots + x_n^2$. L'indice ν de q vaut $[n/2]$ (partie entière).

2) On suppose $k = \mathbf{R}$ et $\dim E = n$. Il y a $n+1$ classes d'équivalence de formes quadratiques non dégénérées sur E. Dans une base convenable, q a pour matrice :

$$\begin{pmatrix} 1 & & & & & \\ & \ddots & & & 0 & \\ & & 1 & & & \\ & & & -1 & & \\ & 0 & & & \ddots & \\ & & & & & -1 \end{pmatrix}$$

avec p signes $+$ et $(n-p)$ signes $-$, $0 \leq p \leq n$. Autrement dit, si $x = (x_1, \ldots, x_n)$ dans cette base, on a :

$$q(x) = \sum_{i=1}^{p} x_i^2 - \sum_{i=p+1}^{n} x_i^2.$$

L'indice de q est alors $\nu = \inf(p, n-p)$.

3) On suppose $k = \mathbf{C}$ et $\dim E = n$. Il y a $n+1$ classes de formes hermitiennes

relatives à la conjugaison, non dégénérées. Leurs matrices sont les mêmes qu'en 2) de sorte que si $z = (z_1, \ldots, z_n) \in E$, *on a :*

$$q(z) = \sum_{i=1}^{p} z_i \overline{z_i} - \sum_{i=p+1}^{n} z_i \overline{z_i}.$$

Définition 6.7.
Avec les notations de 6.6.2 (dans le cas symétrique réel), le couple $(p, n-p)$ *s'appelle la* ***signature de*** q. *Si* $p = n$, *la forme* q *est anisotrope et on a :*

$$\forall x \in E,\ x \neq 0 \implies q(x) > 0,$$

q *est dite* ***définie positive, ou euclidienne****, cf. Chapitre VI. Dans une base convenable,* q *a alors pour matrice la matrice identité* I_n. *Une telle base* $e_1, \ldots, e_n$, *telle que* $q(e_i) = 1$ *et* $f(e_i, e_j) = 0$ *pour* $i \neq j$ *est dite* ***orthonormée****.*

Démonstration (de 6.6).

1) Dans une base orthogonale $e_1, \ldots, e_n$, la forme q s'écrit, si $x = \sum_{i=1}^{n} x_i e_i$:

$$q(x) = \sum_{i=1}^{n} a_i x_i^2 \quad \text{avec } a_i \neq 0.$$

Mais comme k est algébriquement clos, on peut écrire $a_i = \alpha_i^2$ et si on pose $x_i' = \alpha_i x_i$, on a $q(x) = \sum_{i=1}^{n} x_i'^2$, donc q a pour matrice I dans la base $(\alpha_i^{-1} e_i)_{1 \leq i \leq n}$.
Cette base est dite orthonormée. L'indice ν est au plus $[n/2]$ (cf. 3.11). D'autre part, si $e_1 \ldots e_n$ est orthonormée, et si i est une racine de -1, $e_1 + ie_2$, $e_3 + ie_4, \ldots$ forment un système de $[n/2]$ vecteurs isotropes indépendants, d'où $\nu = [n/2]$.

2) On a encore une base orthogonale $e_1, \ldots, e_n$ avec, disons, $a_i = q(e_i) > 0$ pour $i = 1, \ldots, p$ et $a_i = q(e_i) < 0$ pour $i = p+1, \ldots, n$. On pose $a_i = \alpha_i^2$ pour $i \leq p$, $a_i = -\alpha_i^2$ pour $i > p$.
Dans la base $(\alpha_1{}^{-1} e_1, \ldots, \alpha_n{}^{-1} e_n)$, q a la matrice annoncée.
Les $n+1$ formes ainsi obtenues ne sont pas équivalentes. En effet, si on a une forme q et deux bases $e_1, \ldots, e_n$; $e_1', \ldots, e_n'$ avec :

$$q(e_1) = \ldots = q(e_p) = q(e_1') = \ldots = q(e_{p'}') = 1,$$

$$q(e_{p+1}) = \ldots = q(e_n) = q(e_{p'+1}') = \ldots = q(e_n') = -1,$$

nous allons montrer que et p et p' sont égaux.
Soit F (resp. G') le sous-espace engendré par $e_1, \ldots, e_p$ (resp. $e_{p'+1}', \ldots, e_n'$). Sur $F - \{0\}$, on a $q(x) > 0$, sur $G' - \{0\}$, $q(x) < 0$. Ceci montre que $F \cap G'$ est réduit à $\{0\}$, donc $F \oplus G'$ est un sous-espace de E, d'où $p + n - p' \leq n$ et donc $p \leq p'$. On obtient de même $p' \leq p$ en utilisant F' et G, donc $p = p'$.

Il reste l'assertion concernant l'indice. On voit aisément en utilisant les vecteurs $e_1 + e_{p+1}$, $e_2 + e_{p+2}, \ldots$, que l'on a $\nu \geq \inf(p, n-p)$. L'égalité sera prouvée au chapitre VIII, 3.7.

Théorème 6.8.
Soit $k = \mathbf{F}_q$ *un corps fini de caractéristique différente de 2 et* E *un* k*-espace de*

dimension n. Soit $\alpha \in \mathbf{F}_q^*$, $\alpha \notin \mathbf{F}_q^{*2}$, cf. III 2.10. Il y a deux classes d'équivalences de formes quadratiques non dégénérées sur E, de matrices :

$$Q_1 = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ & & \ddots \\ 0 & & 1 \end{pmatrix} \quad \text{ou} \quad Q_2 = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ & & 1 \\ 0 & & \alpha \end{pmatrix}$$

Une forme Q est de l'un ou l'autre type suivant que son discriminant $\delta(Q)$ est, ou non, un carré de $\mathbf{F}_q^*$.

Remarque 6.9. En dimension impaire on a $Q_2 \sim \alpha Q_1$ (car les discriminants de ces formes valent tous deux α à un carré près), de sorte qu'il y a un seul groupe orthogonal à isomorphisme près. En revanche, en dimension paire, les deux groupes orthogonaux ne sont pas isomorphes, cf. ci-dessous Exercice 6, ainsi que VIII, 6.5.2 et VIII, 9.2.

Démonstration (de 6.8). On prouve d'abord le lemme suivant :

Lemme 6.10.
L'équation en x, y, $ax^2 + by^2 = 1$, avec $a, b \in \mathbf{F}_q^*$, a des solutions dans $\mathbf{F}_q$.

Démonstration (de 6.10). On sait qu'il y a $\dfrac{q+1}{2}$ carrés dans $\mathbf{F}_q$ (cf. III, 2.10), donc, quand y décrit $\mathbf{F}_q$, la quantité $\dfrac{1-by^2}{a}$ prend $\dfrac{q+1}{2}$ valeurs, donc cette quantité prend nécessairement une valeur qui est un carré puisqu'on a $\dfrac{q+1}{2} + \dfrac{q+1}{2} > q$.

Revenant à 6.8, on va déduire du lemme que, pour $n \geq 2$, toute forme quadratique Q sur E a une matrice du type annoncé. On procède par récurrence.

Dans le cas $n = 2$, on choisit une base orthogonale pour Q dans laquelle on a $Q(x, y) = ax^2 + by^2$. Par le lemme, on a un vecteur $e_1 = (x, y)$ tel que $Q(e_1) = 1$. Soit e_2 un vecteur orthogonal à e_1. Si on a $Q(e_2) = \lambda^2 \in \mathbf{F}_q^{*2}$, on remplace e_2 par $\dfrac{1}{\lambda} e_2$. Si $Q(e_2)$ n'est pas un carré, comme $\mathbf{F}_q^{*2}$ est d'indice 2 dans $\mathbf{F}_q^*$ (cf. III, 2.10), on a $Q(e_2) = \lambda^2 \alpha$ et on remplace e_2 par $\dfrac{1}{\lambda} e_2$.

Pour $n > 2$, on prend une base orthogonale $e_1, \ldots, e_n$. Le lemme montre qu'il existe un vecteur ε_1 du plan $< e_1, e_2 >$ avec $Q(\varepsilon_1) = 1$. On applique alors l'hypothèse de récurrence à l'hyperplan $H = <\varepsilon_1>^\perp$ et on obtient la matrice voulue.

Comme α n'est pas un carré, les deux formes précédentes ne sont pas équivalentes car leurs discriminants ne sont pas égaux modulo $\mathbf{F}_q^{*2}$ et donc il y a bien deux classes d'équivalences de formes sur $\mathbf{F}_q$.
Pour le calcul de l'indice, voir Exercice **6**.

Théorème 6.11.
Soit k un corps fini, avec $\operatorname{car}(k) \neq 2$.
1) S'il existe une involution σ de k, non triviale, le cardinal de k est un carré, soit q^2, et on a $k = \mathbf{F}_{q^2}$.
2) Réciproquement il y a une unique involution σ non triviale de $\mathbf{F}_{q^2}$, donnée par $\sigma(x) = x^q$, et le corps des invariants de σ est égal à $\mathbf{F}_q$.
3) Dans ce cas, si E est un k-espace de dimension n, il y a une seule classe de

formes hermitiennes relatives à σ sur E. Une telle forme admet dans une base convenable la matrice identité.

Démonstration.

1) Notons d'abord que si k est fini, muni d'une involution σ, on a un sous-corps k_0 de k avec $[k : k_0] = 2$ (cf. 2.14). Si on pose $k_0 = \mathbf{F}_q$, on a donc $k = \mathbf{F}_{q^2}$. De plus σ est alors l'unique automorphisme de k laissant invariant k_0 et distinct de l'identité (cf. 2.14) et comme $\mathbf{F}_{q^2}$ n'a qu'un sous-corps de cardinal q, il a une unique involution non triviale. Cette involution n'est autre que l'automorphisme de Frobenius $\sigma : \mathbf{F}_{q^2} \longrightarrow \mathbf{F}_{q^2}$ défini par $\sigma(x) = x^q$ (cf. III, 2.4 ; q est une puissance de la caractéristique). On vérifie bien, en effet :

1) $\sigma(x) = x$ si $x \in \mathbf{F}_q$,

2) $\sigma^2(x) = x^{q^2} = x$ pour tout $x \in \mathbf{F}_{q^2}$,

3) $\sigma \neq \mathrm{Id}$ (car $x^q - x$ a au plus q racines dans $\mathbf{F}_{q^2}$).

Considérons alors l'application $N : \mathbf{F}_{q^2}^* \longrightarrow \mathbf{F}_q^*$ donnée par $N(x) = x\sigma(x) = x^{q+1}$. (On a bien $N(x) \in \mathbf{F}_q^*$ car $x\sigma(x)$ est invariant par σ). Alors, N est surjective. En effet, N est un homomorphisme de groupes de noyau :

$$\mathrm{Ker}\, N = \{x \in \mathbf{F}_{q^2}^* \mid x^{q+1} = 1\}.$$

Or, l'équation $x^{q+1} = 1$ a au plus $q+1$ racines, donc on a $|\mathrm{Ker}\, N| \leq q+1$ et donc $|\mathrm{Im}\, N| \geq \dfrac{q^2-1}{q+1} = q-1$ ce qui prouve $\mathrm{Im}\, N = \mathbf{F}_q^*$.

Soit alors f une forme hermitienne sur E, $e_1, \ldots, e_n$ une base orthogonale, $\gamma_i = f(e_i, e_i)$, avec $\gamma_i \in \mathbf{F}_q^*$, donc $\gamma_i = N(\lambda_i)$, $\lambda_i \in \mathbf{F}_{q^2}^*$. Alors, si on pose $\varepsilon_i = \dfrac{1}{\lambda_i} e_i$ on a $f(\varepsilon_i, \varepsilon_i) = \dfrac{\gamma_i}{\lambda_i \lambda_i^\sigma} = 1$ et donc (ε_i) est une base orthonormée pour f, ce qui achève la démonstration.

7. Caractérisation des similitudes.

Proposition 7.1.
Soit E un k-espace vectoriel de dimension n, muni d'une forme quadratique q non dégénérée. Soit $u \in GL(E)$. Alors, u est une similitude si et seulement si u conserve l'orthogonalité :

$$u \in GO(q) \Longleftrightarrow \big(\forall x, y \in E,\ \ x \perp y \implies u(x) \perp u(y)\big).$$

Démonstration.

Le sens $\Longrightarrow$ est trivial. Réciproquement, soit $e_1, \ldots, e_n$ une base orthogonale pour q. On pose $\varepsilon_i = u(e_i)$, de sorte que (ε_i) est aussi une base orthogonale. Comme on a $q(e_i) \neq 0$ et $q(\varepsilon_i) \neq 0$ puisque q est non dégénérée, on a $q(\varepsilon_i) = \mu_i q(e_i)$ avec $\mu_i \in k$, et il suffit de prouver que les μ_i sont tous égaux. Considérons le vecteur $e_i + e_j$, qui est orthogonal à $e_i + \lambda e_j$, lorsque $\lambda = -\dfrac{q(e_i)}{q(e_j)}$. Alors, par hypothèse, $\varepsilon_i + \varepsilon_j$ est orthogonal à $\varepsilon_i + \lambda\varepsilon_j$ et on en déduit $\lambda = -\dfrac{q(\varepsilon_i)}{q(\varepsilon_j)} = \lambda \dfrac{\mu_i}{\mu_j}$, donc $\mu_i = \mu_j$.

Remarque 7.2. Lorsque l'indice ν de q est ≥ 1, on a une autre caractérisation des similitudes. Si $u \in GL(E)$, u est une similitude si et seulement si u conserve l'isotropie :

$$u \in GO(q) \iff \Big(\forall x \in E,\ q(x) = 0 \implies q\big(u(x)\big)\Big) = 0,$$

(cf. Exercice 2).

EXERCICES SUR LE CHAPITRE V

1. Définitions.

1) Montrer que le seul automorphisme du corps $\mathbf{R}$ est l'identité.

---

2) Déterminer les automorphismes du corps $\mathbf{F}_q$ (cf. Chapitre III § 2 Exercice 5).

---

3) Déterminer les automorphismes du corps $\mathbf{Q}(\sqrt{d})$, avec $d \in \mathbf{Q}^*$.

---

4) Montrer qu'il existe des automorphismes du corps $\mathbf{C}$ autres que l'identité et la conjugaison (difficile : cf. Chapitre III § 1 Exercice 3.f ou [L] Ch. X § 1 et Ch. VII § 2 Th.2). En déduire qu'il existe d'autres involutions que la conjugaison, donc des sous-corps $K \subset \mathbf{C}$ avec $K \neq \mathbf{R}$ et $[\mathbf{C} : K] = 2$.

---

5) Généraliser aux applications semi-linéaires les techniques de calcul du rang au moyen des déterminants.

---

6) Soit k un corps et $\sigma \in \operatorname{Aut} k$. Montrer que l'application $N : k^* \longrightarrow k^*$ définie par $N(x) = xx^\sigma$ est un homomorphisme de groupes (appelé norme).
Déterminer l'image de N (le groupe des normes) lorsque $k = \mathbf{C}$ et lorsque σ est la conjugaison. Déterminer $\mathbf{C}^*/N(\mathbf{C}^*)$.

---

7) Soit $d \in \mathbf{Z}$ un entier sans facteur carré, i.e. tel que $d = \varepsilon p_1 \dots p_r$ avec $\varepsilon = \mp 1$ et les p_i premiers distincts. Soit σ l'automorphisme de $\mathbf{Q}(\sqrt{d})$ défini par $\sigma(\sqrt{d}) = -\sqrt{d}$.
Déterminer le groupe des normes relatif à σ (cf. Exercice 6). Montrer que ce groupe est strictement inclus dans $\mathbf{Q}^*$ (on admettra que si un nombre $d \in \mathbf{Z}$ n'est pas un carré de $\mathbf{Z}$, il existe un nombre premier p tel que $\overline{d} \notin \mathbf{F}_p^{*2}$).
Préciser ce groupe dans le cas $d = -1$ (cf. Chapitre II § 4).

---

2. Formes réflexives.

1) Déterminer toutes les formes réflexives dégénérées sur E, lorsque l'on a

$$\dim E = \dim \operatorname{Ker} f + 1.$$

---

3. Sous-espaces orthogonaux et isotropes.

1) Soit $E = l^2$ l'ensemble des suites $u = (u_n)_{n\in\mathbf{N}}$, de nombres réels, de carré sommable, i.e. telles que la série $|u_n|^2$ soit convergente : $\sum_{n\in\mathbf{N}} |u_n|^2 < +\infty$.

a) Montrer que E est un $\mathbf{R}$-espace vectoriel.

b) Soient $u = (u_n)_{n\in\mathbf{N}}$, $v = (v_n)_{n\in\mathbf{N}}$ des éléments de E. Montrer que la série de terme général $|u_n v_n|$ est convergente.

c) Pour $u, v \in E$, on pose $(u \mid v) = \sum_{n\in\mathbf{N}} u_n v_n$.
Montrer que $(u \mid v)$ est une forme bilinéaire symétrique sur E, définie positive.

d) Pour $u \in E$, on pose $||u|| = \sqrt{(u \mid u)}$. Montrer que $u \mapsto ||u||$ est une norme sur E.

e) Montrer que E est complet pour cette norme (donc que E est un espace de Hilbert).

f) Soit e_i l'élément de E défini par : $(e_i)_n = 0$ si $i \neq n$ et $(e_i)_i = 1$.
Montrer qu'on a, dans l'espace de Hilbert E, l'égalité :

$$u = (u_n)_{n\in\mathbf{N}} = \sum_{n\in\mathbf{N}} u_n e_n.$$

Montrer que les e_i sont deux à deux orthogonaux.

---

2) On reprend les notations de 1).
Soit F le sous-espace vectoriel de E engendré par les e_i, de sorte que F est l'ensemble des suites qui n'ont qu'un nombre fini de termes non nuls.

a) Montrer que F est strictement inclus dans E.

b) En déduire que F est strictement inclus dans $F^{\perp\perp}$.

---

3) On reprend les notations de 1) et 2).
Soit u la suite définie par $u_n = \dfrac{1}{n+1}$.

a) Montrer que u est élément de E.

b) Soit G le sous-espace $\mathbf{R}u$. Montrer que $F^\perp + G^\perp$ est strictement inclus dans $(F \cap G)^\perp$.

---

4) On reprend les notations de 1).
Soit $x \in E$ et $f_x \in E^*$ la forme linéaire définie par $f_x(y) = (x \mid y)$.

a) Montrer que f_x est continue sur l'espace normé E.

b) Soit $\overline{f} : E \longrightarrow E^*$ l'application définie par $\overline{f}(x) = f_x$. Montrer que $\overline{f}$ est injective, mais non bijective (construire une forme $g \in E^*$ non continue).

5) On reprend les notations de 1) et 2).
Montrer que l'on a $F \cap F^{\perp} = \{0\}$, mais que E n'est pas somme directe de F et de son orthogonal $F^{\perp}$.

---

6) Déterminer les vecteurs isotropes et les sous-espaces totalement isotropes pour les formes suivantes :
$k = \mathbf{R}$; $q = x^2+y^2$, x^2-y^2, $2xy$, $x^2+y^2-z^2$, $x^2+y^2+z^2-t^2$, $x^2+y^2-z^2-t^2$.
$k = \mathbf{C}$; $q = x^2+y^2$, $x^2+y^2+z^2$.
$k = \mathbf{F}_q$; $Q = x^2+y^2$, x^2-y^2.

---

4. Groupes unitaire, orthogonal, symplectique.

1) Déterminer les groupes unitaires, orthogonaux, symplectiques lorsque l'espace E est de dimension 1.

---

2) Déterminer les dilatations qui sont unitaires (resp. orthogonales, resp. symplectiques).

---

3) Soit E un k-espace vectoriel de dimension $n \geq 2$ et φ une forme non dégénérée sur E, hermitienne, symétrique ou alternée.
Soit τ une transvection de E, que l'on écrit sous la forme : $\tau(x) = x + f(x)a$, avec $a \in E$, $a \neq 0$, $f \in E^*$, $f \neq 0$ et $f(a) = 0$. On suppose que τ est une isométrie relativement à φ.

a) Montrer que a est isotrope.

b) Montrer qu'il existe $\lambda \in k^*$ tel que l'on ait : $\forall x \in E$, $f(x) = \lambda\varphi(x,a)$. Préciser l'hyperplan de τ.

c) Si φ est σ-hermitienne, avec $\sigma \neq \mathrm{id}$, montrer qu'on a $\lambda + \lambda^{\sigma} = 0$.

d) Montrer qu'il existe toujours des transvections symplectiques, jamais de transvections orthogonales (si $\mathrm{car}\,(k) \neq 2$), et des transvections unitaires si et seulement si l'indice ν est ≥ 1.

---

4) On suppose $\mathrm{car}\,(k) \neq 2$. Déterminer les involutions des groupes unitaires et symplectiques. Préciser le cas des réflexions.
Montrer que si k est de caractéristique 2 le groupe symplectique possède des involutions (cf. Exercice 3).

---

5) *Une caractérisation des isométries.*
Soit q une forme quadratique non dégénérée sur E et soit $u : E \longrightarrow E$ une application (que l'on ne suppose pas linéaire) qui vérifie :

i) $u(0) = 0$,

ii) $\forall x, y \in E$, $q\big(u(x) - u(y)\big) = q(x-y)$

(si par exemple q est euclidienne on a donc, en termes de normes, $||u(x)-u(y)|| = ||x-y||$).
Montrer que l'on a $u \in O(q)$, donc que u est linéaire (utiliser une base orthogonale).

6) *Une autre caractérisation des isométries.*
Soit q une forme quadratique sur E, et pour $\alpha \in k^*$, soit $S_\alpha = \{x \in E \mid q(x) = \alpha\}$ (la "sphère" dont le "carré du rayon" est α). Soit $\alpha \in k^*$, on suppose $S_\alpha \neq \emptyset$ et on se propose d'étudier l'équivalence, pour $u \in GL(E)$:

$$(*) \qquad u \in O(q) \Longleftrightarrow u(S_\alpha) = S_\alpha.$$

a) Montrer que, quitte à remplacer q par $\alpha^{-1}q$, on peut se ramener au cas $\alpha = 1$. On posera $S = S_1$. Montrer que $(*)$ est vraie si E est de dimension 1.

b) On suppose dim $E = 2$ et q dégénérée. Soit w un vecteur non nul du noyau de q et soit $v \in S$. Déterminer les vecteurs $\lambda v + \mu w$, avec $(\lambda, \mu \in k)$, qui sont dans S. En déduire que $(*)$ est vraie dans ce cas.

c) On suppose dim $E = 2$, q non dégénérée et $|k| > 5$.
Soit (v, w) une base orthogonale de E avec $v \in S$ et $q(w) = a \in k^*$. Déterminer les vecteurs $\lambda v + \mu w$ de S (on pourra, par exemple, couper la conique $\lambda^2 + a\mu^2 = 1$ par une droite variable issue de v).
En déduire que $(*)$ est vraie dans ce cas.

d) Montrer que $(*)$ est vraie aussi pour $k = \mathbf{F}_5$ et dim $E = 2$.

e) Lorsque l'on a $k = \mathbf{F}_3$ et dim $E = 2$, montrer que $(*)$ est vraie si $q \sim x^2 + y^2$, fausse si $q \sim x^2 - y^2$.

f) Montrer que $(*)$ est vraie si $k = \mathbf{F}_3$ et $n = 3$.

g) Déduire ce qui précède que $(*)$ est vraie si q est non dégénérée, sauf dans le cas $n = 2$, $k = \mathbf{F}_3$, $q \sim x^2 - y^2$.

---

7) Soit q une forme quadratique sur un corps de caractéristique différente de 2. Montrer qu'on a un isomorphisme :

$$O(q) \simeq O^+(q) \rtimes \{1, -1\}$$

(utiliser les symétries).

---

5. Les similitudes.

1) Soit u une similitude de multiplicateur μ, relative à f forme σ-sesquilinéaire non dégénérée, sur E de dimension n.
Montrer que l'on a $\mu^n = (\det u)(\det u)^\sigma$.

---

2) Soit q une forme quadratique non dégénérée sur E.
Montrer que l'on a $\mu(GO(q)) = \{\lambda \in k^* \mid q \sim \lambda q\}$ et que ce groupe contient k^{*2} (mais ne lui est pas nécessairement égal, cf. Exercice 3).

---

3) Soit q une forme quadratique non dégénérée sur un espace vectoriel réel E.

a) Montrer que l'on a $\mu(GO(q)) = \mathbf{R}^{+*}$ sauf si $n = 2p$ et si q est de signature (p, p) (cf. Exercice 2). Montrer que dans ce dernier cas, on a $\mu(GO(q)) = \mathbf{R}^*$.

b) Montrer que l'on a $GO(q) = \mathbf{R}^* O(q)$ (i.e. que les isométries et les homothéties engendrent le groupe des similitudes) si et seulement si on a l'égalité $\mu(GO(q)) = \mathbf{R}^{+*}$. Conclure.

6. Bases orthogonales ; classification des formes sesquilinéaires.

1) Soit E un espace de Hilbert réel de dimension infinie (par exemple $E = l^2$, cf. § 3 Exercice 1). On rappelle qu'un tel espace n'est jamais de dimension (algébrique) dénombrable (cf. par exemple [FAM] Ch. V § 9 Exercice 2).
On fait sur E l'hypothèse suivante :
(H) il existe une base orthogonale algébrique $(e_i)_{i\in I}$ de E, c'est-à-dire une famille de vecteurs qui vérifie :

1) $\forall i, j \in I, \quad i \neq j \implies e_i \perp e_j$,

2) $\forall x \in E, \quad x = \sum_{i\in I} x_i e_i$ où les x_i sont des réels nuls sauf un nombre fini.

Soit $J \subset I$ une partie dénombrable de I et V le sous-espace engendré par les e_i, pour $i \in J$.

a) Montrer que, si $\overline{V}$ est l'adhérence de V, on a $\overline{V}^{\perp} = V^{\perp}$. En déduire que les e_i, pour $i \in I - J$, sont dans $\overline{V}^{\perp}$.

b) Montrer que l'on a $\overline{V} \cap \overline{V}^{\perp} = \{0\}$ et en déduire que $(e_i)_{i\in J}$ est une base de $\overline{V}$.

c) Montrer que $\overline{V}$ est un espace de Hilbert de dimension dénombrable, et en déduire que l'hypothèse (H) est absurde.

---

2) Soit f une forme alternée sur le k-espace vectoriel E de dimension n.

a) On suppose $n = 2$ et f non dégénérée. Montrer qu'il existe une base e_1, e_2 de E telle que $f(e_1, e_1) = f(e_2, e_2) = 0$, $f(e_1, e_2) = 1$. Une telle base est dite symplectique.

b) Montrer que si n est impair, f est dégénérée (utiliser a) et raisonner par récurrence).

c) Si on a $\dim E = 2m$ et f non dégénérée, montrer qu'il existe une base symplectique pour f , i.e. une base $e_1, \varepsilon_1, \ldots, e_m, \varepsilon_m$ de E, telle que, si on pose $P_i = <e_i, \varepsilon_i>$, on ait $P_i \perp P_j$ pour $i \neq j$ et que (e_i, ε_i) soit une base symplectique de P_i.

d) En déduire que toutes les formes alternées de rang $2m$ sur k sont équivalentes.

e) Montrer que l'on a $Sp(2, k) = SL(2, k)$.

---

3) On considère des nombres premiers $p_1, \ldots, p_n$; $q_1, \ldots, q_n$ et les formes quadratiques $p_1 x_1^2 + \cdots + p_n x_n^2$ et $q_1 x_1^2 + \cdots + q_n x_n^2$.
A quelle condition ces formes sont-elles équivalentes sur $\mathbf{Q}$?

---

4) On considère la forme quadratique sur $\mathbf{R}^n$ définie par :

$$q(x_1, \ldots, x_n) = \sum_{i\neq j} x_i x_j.$$

Déterminer la signature de q (on pourra diagonaliser la matrice A de q dans une base orthonormée et noter que cette opération constitue à la fois une réduction de l'endomorphisme associé à A et de la forme q).

---

5) Donner la liste des groupes orthogonaux réels, à isomorphisme près, pour $n \leq 4$.

---

6) Soit p un nombre premier et soit $q = p^\alpha$, pour $\alpha \in \mathbf{N}^*$. Soit $m \in \mathbf{N}^*$. On considère la forme quadratique Q_0 sur $\mathbf{F}_q^{2m}$ définie par :

$$Q_0(x_1, y_1, \ldots, x_m, y_m) = \sum_{i=1}^{m} x_i^2 - y_i^2.$$

a) Montrer que l'indice $\nu(Q_0)$ est égal à m.

b) Montrer que toute forme de rang $2m+1$ sur $\mathbf{F}_q$ est équivalente soit à $Q_0 + x_{m+1}^2$, soit à $Q_0 + a x_{m+1}^2$, avec $a \notin \mathbf{F}_q^{*2}$.
En déduire que toute forme de rang $2m+1$ est d'indice m.

c) Montrer que toute forme Q de rang $2m$ sur $\mathbf{F}_q$ est équivalente à Q_0, ou à la forme Q_1 définie par :

$$Q_1(x_1, \ldots, y_m) = \sum_{i=1}^{m-1} x_i^2 - y_i^2 + x_m^2 - a y_m^2, \quad \text{avec } a \notin \mathbf{F}_q^{*2}.$$

En déduire que l'indice de Q est au moins $m-1$ et qu'il vaut m si Q est équivalente à Q_0 ou, ce qui revient au même, si $(-1)^m \delta(Q)$ est dans $\mathbf{F}_q^{*2}$ (où $\delta(Q)$ désigne le discriminant de Q).

d) En utilisant le lemme de plongement d'un sous-espace totalement isotrope dans un hyperbolique (cf. VIII, 3.5), montrer que si Q est équivalente à Q_1, on a $\nu(Q) = m - 1$.

---

7) a) Soit E un k-espace vectoriel de dimension n et soient q, q' deux formes quadratiques non dégénérées définies sur E. Soit $e_1, \cdots, e_n$ une base de E, orthogonale à la fois pour q et q'. On pose $a_i = q(e_i)$ et $a_i' = q'(e_i)$. On suppose que, pour tout i, $\dfrac{a_i}{a_i'}$ est un carré de k. Montrer que q et q' sont équivalentes.

b) Déduire de a) que si deux formes quadratiques ont respectivement pour matrices A et A^{-1} dans une même base elles sont équivalentes.

c) Déduire de a) que les propriétés suivantes sont équivalentes :

i) le groupe k^*/k^{*2} est fini,

ii) il n'y a qu'un nombre fini de classes d'équivalences de formes quadratiques de rang fixé sur k.

---

7. Caractérisation des similitudes.

1) Soit E un k-espace vectoriel de dimension n et soient q, q' deux formes quadratiques sur E, non dégénérées. On désigne par $C(q)$ (resp. $C(q')$) le cône isotrope de q (resp. q') :

$$C(q) = \{x \in E \mid q(x) = 0\}.$$

On suppose $C(q) = C(q')$ et $C(q) \neq \{0\}$ et on se propose de montrer que l'on a $q' = \lambda q$ pour un certain $\lambda \in k^*$.

a) Soit $x_0 \in C(q)$, $x_0 \neq 0$. En considérant, parmi les vecteurs $\alpha x_0 + x$, pour ($\alpha \in k$ et $x \in E$), ceux qui sont dans $C(q)$, montrer que l'hyperplan H orthogonal à x_0 est le même pour q et q'.

b) Montrer qu'il existe $\lambda \in k^*$ tel que, si $x \notin H$, on a $q'(x) = \lambda q(x)$.

c) Montrer que, pour $x \notin H$ et $y \notin H$, on a $\varphi'(x,y) = \lambda\varphi(x,y)$. Conclure.

d) Soit $V(q)$ l'image de $C(q) - \{0\}$ dans l'espace projectif $\mathbf{P}(E)$ ($V(q)$ est la quadrique projective associée à q). Montrer que si $V(q)$ est non vide $V(q)$ détermine q à un scalaire près.

---

2) On reprend les notations de 1) en supposant toujours $C(q) \neq \{0\}$. Soit $u \in GL(E)$. Montrer que l'on a :

$$u \in GO(q) \iff u\big(C(q)\big) = C(q).$$

---

3) Montrer que le normalisateur de $O(q)$ dans $GL(E)$ est $GO(q)$ (on pourra considérer les réflexions τ_x, pour $x \in E$, x non isotrope et utiliser 2) et/ou 7.1).

VI. LE GROUPE ORTHOGONAL

EUCLIDIEN

1. Notations et rappels.

Soient E un espace vectoriel ***réel***, de dimension finie n, q une forme quadratique sur E et f sa forme polaire.
Dans tout ce chapitre nous supposerons q ***définie positive (ou euclidienne)***, c'est-à-dire telle que :

$$\forall x \in E, \quad x \neq 0 \implies q(x) > 0.$$

En particulier, la forme q est non dégénérée et anisotrope.
Le couple (E, q) s'appelle un ***espace euclidien***. On pose $f(x, y) = (x \mid y)$, c'est le ***produit scalaire*** de x et y et $||x|| = \sqrt{(x \mid x)}$. L'application $x \mapsto ||x||$ est une ***norme*** sur E.
On a vu (cf. V, 6.7) qu'il existe pour q des ***bases orthonormées*** $e_1, \ldots, e_n$, c'est-à-dire telles que l'on ait :

$$\forall i = 1, \ldots, n, \quad q(e_i) = 1, \tag{1}$$

$$\forall i, j \in \{1, \ldots, n\}, \quad i \neq j \implies (e_i \mid e_j) = 0. \tag{2}$$

Le groupe $O(q)$ opère simplement transitivement sur l'ensemble des bases orthonormées de E : si $e_1, \ldots, e_n$ et $e'_1, \ldots, e'_n$ sont deux telles bases, il existe $u \in O(q)$, unique, tel que $u(e_i) = e'_i$ pour tout i (cf. V, 4.4.3).
Dans une base orthonormée, q a pour matrice I (matrice identité) et on a,

$$\text{si } x = \sum_{i=1}^{n} x_i e_i, \quad q(x) = \sum_{i=1}^{n} x_i^2.$$

Le groupe $O(q)$ est isomorphe au groupe des ***matrices orthogonales*** :

$$O(n, \mathbf{R}) = \{A \in \mathbf{M}(n, \mathbf{R}) \mid {}^tAA = I\},$$

le sous-groupe $O^+(q)$ étant isomorphe à $O^+(n, \mathbf{R}) = \{A \in O(n, \mathbf{R}) \mid \det A = 1\}$.

De même le groupe $GO(q)$ est isomorphe au groupe $GO(n, \mathbf{R})$:

$$GO(n, \mathbf{R}) = \{A \in M(n, \mathbf{R}) \mid \exists \mu \in \mathbf{R}^*, \ {}^tAA = \mu I\}.$$

Notons que dans le cas euclidien, toute similitude est produit d'une isométrie et d'une homothétie. En effet, si μ est le multiplicateur de $u \in GO(q)$, on a :

$$\big(u(x) \mid u(x)\big) = \mu(x \mid x), \text{ pour tout } x \in E,$$

donc $\mu > 0$ et donc, si on pose $\lambda = \sqrt{\mu}$, on a $u = h_\lambda v$ où $v \in O(q)$ et $h_\lambda(x) = \lambda x$ (cf. Chapitre V § 5 Exercice 3).

Signalons enfin deux propriétés remarquables des espaces euclidiens :

1) il n'y a ni vecteurs, ni sous-espaces isotropes dans E,

2) si F est un sous-espace de E, $q|_F$ est euclidienne.

2. Générateurs et centres de $O(q)$ et $O^+(q)$.

Au Chapitre V § 4, nous avons mis en évidence des éléments particuliers de $O(q)$ et $O^+(q)$, les symétries orthogonales, en particulier les réflexions et les renversements. Leur manipulation est ici grandement facilitée par l'absence d'éléments isotropes dans E. On les utilise d'abord pour le calcul des centres de $O(q)$ et $O^+(q)$.

Théorème 2.1.
1) Le centre de $O(q)$ est $Z = \{\mathrm{Id}, -\mathrm{Id}\}$. En particulier, pour $n \geq 2$, $O(q)$ n'est pas commutatif.
2) Pour $n \geq 3$, le centre de $O^+(q)$ est $Z \cap O^+(q)$, c'est-à-dire $\{\mathrm{Id}\}$ si n est impair, $\{\mathrm{Id}, -\mathrm{Id}\}$ si n est pair.

Démonstration.

1) Il est clair que l'on a $\{\mathrm{Id}, -\mathrm{Id}\} \subset Z$. Réciproquement, soit $u \in Z$ et τ_D une réflexion de droite D. On a $u\tau_D u^{-1} = \tau_D$ puisque u est central, mais aussi $u\tau_D u^{-1} = \tau_{u(D)}$ (cf. V, 4.8), de sorte qu'on a $u(D) = D$. Autrement dit, u laisse invariantes toutes les droites de E, donc est une homothétie (cf. IV, 2.8) donc $u = \mp\mathrm{Id}$ (pour voir que $\mp\mathrm{Id}$ est une isométrie, cf. V, 4.4.4).
Pour $n \geq 2$, on a $O(q) \neq \{\mathrm{Id}, -\mathrm{Id}\}$ (il y a par exemple les réflexions orthogonales) donc $O(q)$ n'est pas commutatif.
Pour $n = 1$, $O(q)$ est réduit à $\{\mathrm{Id}, -\mathrm{Id}\}$.

2) Remarquons que $-\mathrm{Id}$ est dans $O^+(q)$ si et seulement si n est pair. Si u est dans le centre de $O^+(q)$ et si τ_P est un renversement de plan P, on a,

$$u\tau_P u^{-1} = \tau_P = \tau_{u(P)} \ (\textit{loc. cit.}) \text{ donc } u(P) = P,$$

pour tout plan P. Mais comme on a supposé $n \geq 3$, toute droite est intersection de deux plans, donc u laisse aussi invariantes toutes les droites de E et u est une homothétie.

Remarques 2.2.

1) Nous verrons que, pour $n = 2$, $O^+(q)$ est commutatif.

2) Les résultats de ce théorème sont encore valables pour une forme quadratique quelconque, mais il faut prendre garde à ne considérer les réflexions τ_D que lorsque D est non isotrope (cf. Chapitre VIII).

3) On a un isomorphisme $O(q) \simeq O^+(q) \rtimes \{1, -1\}$ (grâce aux réflexions) et même, si n est impair, $O(q) \simeq O^+(q) \times \{1, -1\}$ (grâce à $-\mathrm{Id}$).

Définition 2.3.
Le groupe ***projectif orthogonal*** *est le groupe* $PO(q) = O(q)/\{\mathrm{Id}, -\mathrm{Id}\}$. *De même, on pose* $PO^+(q) = O^+(q)/(Z \cap O^+(q))$ *et on a les groupes matriciels correspondants* $PO(n, \mathbf{R})$ *et* $PO^+(n, \mathbf{R})$.

Théorème 2.4.
Le groupe $O(q)$ *est engendré par les réflexions orthogonales. Plus précisément, si* u *est dans* $O(q)$, u *est produit d'au plus* n *réflexions.*

Démonstration. Soit $u \in O(q)$ et $F_u = \{x \in E \mid u(x) = x\}$ l'espace des points fixes de u. Posons $p_u = n - \dim F_u$. Nous allons prouver que u est produit d'au plus p_u réflexions (on convient que Id est produit de zéro réflexion).
On raisonne par récurrence sur p_u, le cas $p_u = 0$ correspond à $u = \mathrm{Id}$.
Supposons $p_u > 0$, soit $x \in F_u^\perp$, $x \neq 0$, et soit $y = u(x)$. On a $y \neq x$ (car x n'est pas dans F_u) et $y \in F_u^\perp$ car F_u, donc aussi $F_u^\perp$, est stable par u. De plus, comme on a $||x|| = ||y||$, on en déduit $(x - y \mid x + y) = 0$ de sorte que $x - y$ est orthogonal à $x + y$.
Soit alors τ la réflexion définie par le vecteur $x - y$.
On a $\tau(x - y) = y - x$ et $\tau(x + y) = x + y$, donc $\tau(y) = x$. De plus, comme $x - y$ est dans $F_u^\perp$, on a $\tau|_{F_u} = \mathrm{Id}$. On a donc l'inclusion $F_{\tau u} \supset F_u$ et, comme x est dans $F_{\tau u}$ et pas dans F_u, on a $p_{\tau u} < p_u$. Par hypothèse de récurrence, on peut écrire $\tau u = \tau_1 \dots \tau_r$ où les τ_i sont des réflexions et $r \leq p_{\tau u}$, mais alors on a aussi $u = \tau\tau_1 \dots \tau_r$ et $r + 1 \leq p_u$, cqfd.

Remarques 2.5.

1) En fait p_u est aussi le nombre minimal de réflexions nécessaires pour écrire u, cf. Exercice 2.

2) Le théorème est encore vrai pour une forme quelconque, mais nettement plus délicat lorsqu'il y a des vecteurs isotropes (cf. Chapitre VIII).

3) Si τ est une réflexion, on a $\det \tau = -1$ et donc, si u est dans $O^+(q)$, u est produit d'un nombre pair de réflexions.

Théorème 2.6.
Pour $n \geq 3$, $O^+(q)$ *est engendré par les renversements ; précisément, tout élément* $u \in O^+(q)$ *est produit d'au plus* n *renversements.*

Démonstration.

1) Supposons $n = 3$. On a, si $u \neq \mathrm{Id}$, $u = \tau_1\tau_2$ où τ_1, τ_2 sont des réflexions (cf. 2.4 et 2.5.3). Mais, comme on a $n = 3$, $-\tau_i = \sigma_i$ est un renversement (comme on le voit aussitôt sur les matrices) et on a $u = \sigma_1\sigma_2$.

2) Pour n quelconque ≥ 3, soit $u \in O^+(q)$, on a $u = \tau_1 \dots \tau_{2p}$ avec $2p \leq n$, les τ_i étant des réflexions. Il suffit donc de prouver le lemme suivant :

Lemme 2.7.
Soit $n \geq 3$ *et soient* τ_1, τ_2 *des réflexions, il existe des renversements* σ_1, σ_2 *tels que* $\tau_1\tau_2 = \sigma_1\sigma_2$.

Démonstration (de 2.7). Posons $u = \tau_1\tau_2$. Soient H_1, H_2 les hyperplans de τ_1, τ_2 et soit V un sous-espace de dimension $n - 3$ de $H_1 \cap H_2$ (n'oublions pas qu'on a $n \geq 3$). On a $u|_V = \mathrm{Id}$ et donc $u(V^\perp) \subset V^\perp$. D'après le cas 1) on a $u|_{V^\perp} = \sigma_1\sigma_2$ où σ_i est un renversement de $V^\perp$, et on obtient le résultat en prolongeant les σ_i par l'identité sur V.

3. Conjugaison et commutateurs.

On étudie la conjugaison des générateurs précédents dans $O(q)$ et $O^+(q)$.
On commence par étudier la transitivité de l'opération de $O(q)$ sur les sous-espaces :

Lemme 3.1.
Soient V_1, V_2 deux sous-espaces de E de même dimension. Alors il existe $u \in O^+(q)$ tel que $u(V_1) = V_2$.

Démonstration. Soit $e_1, \ldots, e_p$ (resp. $\varepsilon_1, \ldots, \varepsilon_p$) une base orthonormée de V_1 (resp. de V_2) que l'on complète par une base orthonormée $e_{p+1}, \ldots, e_n$ (resp. $\varepsilon_{p+1}, \ldots, \varepsilon_n$) de $V_1^\perp$ (resp. $V_2^\perp$). Alors si $u \in GL(E)$ est défini par $u(e_i) = \varepsilon_i$ pour $i = 1, \ldots, n$, on a $u(V_1) = V_2$ et u, qui transforme une base orthonormée en une base orthonormée, est une isométrie (cf. V, 4.4.3).
De plus, quitte à remplacer ε_1 par $-\varepsilon_1$, on peut supposer $u \in O^+(q)$.

Remarque 3.2. Ce résultat ne subsiste absolument plus pour une forme quadratique quelconque (cf. Chapitre VIII, Th. de Witt). Par exemple si $E = \mathbf{R}^2$ et si q est la forme hyperbolique définie par $q(x, y) = x^2 - y^2$, le groupe $O(q)$ n'est plus transitif sur les droites (il faut distinguer selon que l'on a $q > 0$, $q < 0$, ou $q = 0$ sur D).
De plus, il ne faut pas croire que la situation est plus simple lorsque q est anisotrope sur un corps autre que $\mathbf{R}$. On s'en convaincra en étudiant le cas de la forme $x^2 + y^2$ sur $\mathbf{Q}$ (regarder, par exemple, les droites $y = 0$ et $y = x$) (cf. Exercice 1).

Proposition 3.3.
Soient s_1, s_2 deux symétries orthogonales de même nature (i.e. telles que l'on ait $\dim V_1 = \dim V_2$, avec $V_1 = E^+(s_1)$ et $V_2 = E^+(s_2)$, cf. V, 4.6).
Alors, s_1 et s_2 sont conjuguées « par $O^+(q)$ » (i.e. il existe $u \in O^+(q)$ tel que $us_1u^{-1} = s_2$).

Démonstration. On prend $u \in O^+(q)$ tel que $u(V_1) = V_2$ grâce à 3.1 et on conclut par V, 4.8.

Proposition 3.4.
1) Pour $n \geq 2$, on a $D(O(q)) = O^+(q)$.
2) Pour $n \geq 3$, on a $D(O^+(q)) = O^+(q)$.

Démonstration.
1) Il est clair que l'on a $D(O(q)) \subset O^+(q)$ car, pour un commutateur, on a $\det(uvu^{-1}v^{-1}) = 1$. Réciproquement si τ_1, τ_2 sont deux réflexions, il existe $u \in O^+(q)$ tel que $\tau_2 = u\tau_1u^{-1}$ (c'est 3.3) et donc $\tau_2\tau_1 = u\tau_1u^{-1}\tau_1 = u\tau_1u^{-1}\tau_1^{-1}$ ([1]) est un commutateur. Comme tout élément de $O^+(q)$ est un produit pair de réflexions, on a $O^+(q) \subset D(O(q))$.
2) On a évidemment $D(O^+(q)) \subset O^+(q)$. Réciproquement, il suffit de prouver que tout renversement est un commutateur, et même, puisque les renversements sont conjugués dans $O^+(q)$, qu'un renversement est un commutateur. Pour ceci, soit $V \subset E$ de dimension 3 (on a $n \geq 3$), muni d'une base orthonormée

([1]) Bien sûr, pour une symétrie τ, qui est un élément d'ordre 2, on a $\tau = \tau^{-1}$.

(e_1, e_2, e_3) et considérons les trois renversements $\sigma_1, \sigma_2, \sigma_3$ définis par $\sigma_i|_{V^\perp} = \mathrm{Id}$ et $\sigma_i(e_i) = e_i$ (et donc $\sigma_i(e_j) = -e_j$ pour $i \neq j$).
On a alors $\sigma_3 = \sigma_1\sigma_2$. Mais, toujours par 3.3, il existe $u \in O^+(q)$ tel que $\sigma_2 = u\sigma_1 u^{-1}$ et donc $\sigma_3 = \sigma_1 u\sigma_1 u^{-1} = \sigma_1 u\sigma_1^{-1}u^{-1}$ est un commutateur.

Remarques 3.5.

1) Pour $n = 2$, nous verrons que $O^+(q)$ est commutatif, de sorte que l'on a $D(O^+(q)) = \{\mathrm{Id}\}$.

2) La proposition 3.4 n'est pas vraie pour une forme quelconque. Par exemple, si q est une forme d'indice $\nu \geq 1$ sur $\mathbf{R}$ et si n est ≥ 3, le groupe dérivé $\Omega(q)$ de $O^+(q)$ est un sous-groupe d'indice 2 de $O^+(q)$ (cf. [D] Chapitre II § 8).

4. La dimension 2 et les angles.

Dans tout ce paragraphe, on a $n = \dim E = 2$. Rappelons, cf. V, 4.5 que les isométries positives sont aussi appelées rotations. Pour des précisions concernant la dimension 2, on se reportera à [PR].

Proposition 4.1.
1) Si u est une isométrie négative, u est une réflexion (i.e., ici, une symétrie orthogonale par rapport à une droite).
2) Si u est une rotation elle s'écrit sous la forme $u = \tau_1\tau_2$ où les τ_i sont des réflexions, l'une d'elle pouvant être choisie arbitrairement.
3) Soient $\rho \in O^+(E)$ et $\tau \in O^-(E)$. On a $\tau\rho\tau^{-1} = \rho^{-1}$.
4) Le groupe $O^+(E)$ des rotations est ***commutatif****, donc aussi le groupe $O^+(2, \mathbf{R})$ des matrices orthogonales directes.*

Démonstration. On notera que les points 1) et 2) sont des cas particuliers du théorème 2.4.
Montrons cependant le point 1) sans utiliser 2.4. Soit $u \in O^-(E)$. On a donc $\det u = -1$. Le polynôme caractéristique de u est alors $X^2 - (\mathrm{Tr}\, u)X - 1$. Son discriminant vaut donc $(\mathrm{Tr}\, u)^2 + 4 > 0$ de sorte que u a deux valeurs propres réelles. Comme u est une isométrie ces valeurs propres ne peuvent être que 1 ou -1 (conservation de la norme). Comme leur produit est -1 ce sont donc 1 et -1, et u qui a deux valeurs propres réelles distinctes est diagonalisable. En écrivant la matrice de u dans une base où elle est diagonale on voit alors que $u^2 = \mathrm{Id}_E$ donc u est une involution, donc une symétrie orthogonale, donc une réflexion (les seules symétries possibles en dimension 2 sont les réflexions et $-\mathrm{Id}$, mais cette dernière est positive).
2) On prend pour τ_1 une réflexion quelconque et on pose $\tau_2 = \tau_1^{-1}u$. Il est clair que τ_2 est une isométrie négative, donc une réflexion par 1) et on a alors $u = \tau_1\tau_2$ comme annoncé.
3) On écrit $\rho = \tau\tau'$ comme produit de deux réflexions en imposant la première. On a alors :

$$\tau\rho\tau^{-1} = \tau\tau\tau'\tau^{-1} = \tau'\tau = \rho^{-1}.$$

4) Soient $\rho, \rho' \in O^+(E)$. Posons $\rho' = \tau\tau'$. Il s'agit ici de montrer la formule $\rho'\rho\rho'^{-1} = \rho$. On obtient cette conjugaison par ρ' en conjuguant successivement par τ (ce qui change ρ en ρ^{-1}) puis par τ' (ce qui ramène en ρ).

Forme matricielle des éléments de $GO(q)$.
On détermine aisément par le calcul les éléments de $GO(2,\mathbf{R})$. On trouve deux types de matrices :

$$U_1 = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \quad et \quad U_2 = \begin{pmatrix} a & b \\ b & -a \end{pmatrix}$$

avec $a, b \in \mathbf{R}$ et $a^2 + b^2 > 0$.
Les matrices du type U_1 forment un groupe noté $GO^+(2,\mathbf{R})$ (similitudes directes), isomorphe au groupe multiplicatif $\mathbf{C}^*$ par l'application :

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} \mapsto a + ib.$$

Comme le multiplicateur des similitudes de matrices U_1 ou U_2 est $\mu = a^2 + b^2$, on trouve aussitôt les matrices de $O(2,\mathbf{R})$:

$$O^+(2,\mathbf{R}) = \left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \quad a, b \in \mathbf{R},\ a^2 + b^2 = 1. \right\}$$

$$O^-(2,\mathbf{R}) = \left\{ \begin{pmatrix} a & b \\ b & -a \end{pmatrix} \quad a, b \in \mathbf{R},\ a^2 + b^2 = 1. \right\}$$

L'isomorphisme ci-dessus se restreint en un isomorphisme de $O^+(2,\mathbf{R})$ sur le groupe $\mathbf{U}$ des nombres complexes de module 1.
On retrouve en particulier le fait que $O^+(2,\mathbf{R})$ est commutatif.

Rappelons le résultat suivant : (cf. par exemple [C] Chapitre I § 3) :

Proposition 4.2.
L'application φ définie par $\varphi(t) = e^{it}$ est un homomorphisme de groupes du groupe additif de $\mathbf{R}$ dans $\mathbf{U}$, surjectif, de noyau $2\pi\mathbf{Z}$ qui induit un isomorphisme de groupes topologiques de $\mathbf{R}/2\pi\mathbf{Z}$ sur $\mathbf{U}$.

On pose $e^{it} = \cos t + i \sin t$, ce qui permet d'écrire tout élément de $O^+(2,\mathbf{R})$ sous la forme :

$$R(t) = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}, \quad \text{avec } t \in \mathbf{R}.$$

Définition 4.3.
L'élément $\bar{t}$, image de t dans $\mathbf{R}/2\pi\mathbf{Z}$, s'appelle **l'angle** *de $R(t)$. Le groupe $\mathbf{R}/2\pi\mathbf{Z}$ s'appelle le groupe des angles.*

Remarques 4.4.
1) Si $\bar{t} \in \mathbf{R}/2\pi\mathbf{Z}$ est l'image de t, on dit que t est une mesure de l'angle $\bar{t}$, toute autre mesure de $\bar{t}$ étant alors de la forme $t + 2k\pi$, $k \in \mathbf{Z}$. Il faut utiliser ce terme avec prudence : par exemple, comme $\mathbf{R}/2\pi\mathbf{Z}$ contient des éléments d'ordre fini (images de $\pi, \pi/2, \ldots$), on ne peut parler d'angle plus petit qu'un autre, et les inégalités portant sur les mesures sont donc de peu d'intérêt (cf. [D1] Annexe 1 ou [B] 8.7).
2) On peut construire un corps k ordonné tel que, sur k, la plus grande partie de l'étude du groupe euclidien soit identique à celle que nous avons faite sur $\mathbf{R}$, mais où il n'existe pas d'homomorphisme du type de φ (cf. [D1] Annexe 1 ou ci-dessous Exercice 4).

3) Dans le cas du corps évoqué en 2), on peut cependant parler d'angles : le groupe des angles est alors simplement le groupe $\mathbf{U}_k$, ou $O^+(2,k)$, mais noté additivement (cf. [D1] Ch. V § 4).

Angle d'une rotation
On choisit une **orientation** de E, c'est-à-dire une base orthonormée particulière, (e_1, e_2). Si $(\varepsilon_1, \varepsilon_2)$ est une autre base orthonormée, elle sera dite ***directe***, ou positive (resp. ***indirecte*** ou négative) si l'unique élément $u \in O(q)$ défini par $u(e_i) = \varepsilon_i$ est dans $O^+(q)$ (resp. dans $O^-(q)$). Alors si $\rho \in O^+(q)$, on appelle angle de ρ, l'angle de la matrice de ρ dans n'importe quelle base orthomormée directe (comme $O^+(q)$ est commutatif, cette matrice R ne dépend pas du choix de la base directe). Attention, dans une base indirecte, ρ a pour matrice R^{-1}, d'angle opposé à celui de R. Pour des compléments géométriques sur les angles, angles de vecteurs, de droites, ... on se reportera à [D1] Ch. V § 4, à [B] 8.6 et 8.7, à [Fr] Ch.V § II, III, IV ou à [PR].

5. Structure des éléments de $O(q)$.

Dans ce paragraphe on revient au cas général d'un espace euclidien (E, q) de dimension n.

Théorème 5.1.
Soit $u \in O(q)$. L'espace E est somme directe orthogonale :

$$E = V \perp W \perp P_1 \perp \ldots \perp P_r$$

où r est un entier ≥ 0, où V, W, P_i sont des sous-espaces stables par u, avec $u|_V = \mathrm{Id}_V$, $u|_W = -\mathrm{Id}_W$, et où $\dim P_i = 2$ et $u|_{P_i}$ est une rotation plane, distincte de $\mp\mathrm{Id}_{P_i}$.

Remarque 5.2. Matriciellement, dans une base orthonormée convenable, u a donc pour matrice :

$$\begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & \begin{pmatrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{pmatrix} & & \\ & & & & & & & \ddots & \\ & & & & & & & & \begin{pmatrix} \cos\theta_r & -\sin\theta_r \\ \sin\theta_r & \cos\theta_r \end{pmatrix} \end{pmatrix}$$

avec $\theta_i \in [0, 2\pi[$, et $\theta_i \neq 0, \pi$.
Démonstration (de 5.1) On procède par récurrence sur n. Pour $n = 1$, c'est clair, et pour $n = 2$, cela résulte de 4.1 et 4.2. Supposons donc $n \geq 3$.
Il suffit de trouver un sous-espace F de E, non trivial, et stable par u. En effet, $F^\perp$ sera lui aussi stable par u et on appliquera la récurrence à F et $F^\perp$.

Si u admet une valeur propre réelle (nécessairement égale à ∓ 1), c'est terminé. Remarquons que ceci se produit en particulier si n est impair ou si on a $\det u = -1$. Sinon, soit λ une valeur propre de u sur $\mathbf{C}$, $\lambda \notin \mathbf{R}$, de sorte que $\overline{\lambda}$ est aussi valeur propre de u. Soit $x \in E^{\mathbf{C}}$ (complexifié de E) un vecteur propre de u relatif à λ et soit $\overline{x}$, son conjugué, qui est propre pour u relativement à $\overline{\lambda}$. Le plan (sur $\mathbf{C}$), $P = \mathbf{C}x + \mathbf{C}\overline{x}$ est invariant par $u^{\mathbf{C}}$. Mais les vecteurs $\dfrac{x+\overline{x}}{2}$ et $\dfrac{x-\overline{x}}{2i}$ sont réels et forment une base de P. Comme u est réel, le plan (sur $\mathbf{R}$) engendré par $\dfrac{x+\overline{x}}{2}$ et $\dfrac{x-\overline{x}}{2i}$ est invariant par u.

Remarques 5.3.

0) Les sous-espaces V et W sont bien déterminés par u (ce sont des sous-espaces propres). En revanche, les P_i peuvent ne pas l'être si les valeurs propres complexes de u ont des multiplicités ≥ 2.

1) On peut éviter le recours à la complexification dans la démonstration du théorème (cf. Exercice 1).

2) Si n est impair, V ou W est non nul.

3) Si on a $\det u = -1$, W est non nul.

4) Si n est impair et $\det u = 1$, V est non nul.

5) Si $u \in O^+(q)$ on a $\dim V \equiv n \pmod 2$.

Exemple 5.4. Le cas $n = 3$.

1) Soit $u \in O^+(q)$. D'après la remarque 4), u admet la valeur propre $+1$; on a les cas suivants :

$$\mathrm{Id} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \text{ (renversement)},$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \text{ avec } \theta \in \mathbf{R},\ \theta \notin \pi\mathbf{Z}, \text{ rotation d'axe } \mathbf{R}e_1, \text{ d'angle } \overline{\theta}.$$

2) Soit $u \in O^-(q)$, u a la valeur propre -1 (remarque 3). On a les cas suivants :

$$-\mathrm{Id} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ (réflexion)},$$

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \text{ (produit de trois réflexions)}.$$

6. La simplicité du groupe $O^+(3, \mathbf{R})$.

Théorème 6.1.
Le groupe $O^+(3, \mathbf{R})$ est simple.

Démonstration. Soit N un sous-groupe distingué de $O^+(3, \mathbf{R})$ avec $N \neq \{1\}$, il s'agit de prouver que N est égal à $O^+(3, \mathbf{R})$.

1) Comme les renversements engendrent $O^+(3)$ (cf. 2.6) et sont conjugués dans $O^+(3)$ (cf. 3.3), il suffit de prouver que N contient un renversement.

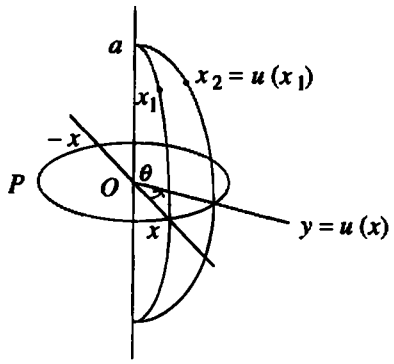
2) Soit $u \in N$, $u \neq \mathrm{Id}$. D'après 5.4.1, u est une rotation d'axe $\mathbf{R}a$, avec $||a|| = 1$, et d'angle θ, avec pour matrice, dans une base orthonormée convenable :

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

Si $\theta = \pi$, u est un renversement, sinon, comme on a aussi $u^{-1} \in N$, on peut supposer $0 < \theta < \pi$. Soit P le plan orthogonal à $\mathbf{R}a$ et notons $\mathbf{S}_2$ la sphère unité de E :

$$\mathbf{S}_2 = \{x \in E \mid ||x|| = 1\}.$$

Soit x un point de l'équateur de $\mathbf{S}_2$ (i.e. un $x \in P \cap \mathbf{S}_2$) et soit $y = u(x)$. On a la figure ci-dessous :



Posons $d = ||x - y||$. Un calcul immédiat fournit $d^2 = 2(1 - \cos\theta)$.

3) On a la propriété suivante :

$$\forall m,\ 0 \leq m \leq d,\ \exists\, x_1, x_2 \in \mathbf{S}_2,\ u(x_1) = x_2 \text{ et } ||x_1 - x_2|| = m.$$

Géométriquement, c'est clair. En effet, u transforme $\mathbf{S}_2$ en elle-même et conserve a. Le méridien de x se transforme donc en celui de y et lorsque x_1 varie sur ce méridien entre x et a, $||x_1 - u(x_1)||$ varie entre d et 0. De façon précise, on considère les vecteurs $x + \lambda a$, $\lambda \in \mathbf{R}$. On a :

$$||x + \lambda a||^2 = 1 + \lambda^2 \text{ donc } x_1 = \frac{x + \lambda a}{\sqrt{1 + \lambda^2}} \in \mathbf{S}_2 \text{ et}$$

$$||u(x_1) - x_1|| = \frac{d}{\sqrt{1 + \lambda^2}}.$$

Il suffit alors de prendre, si $m \neq 0$, $\lambda = \dfrac{\sqrt{d^2 - m^2}}{m}$.

4) Soit alors m tel que $0 \leq m \leq d$. Si y_1, y_2 sont dans $\mathbf{S}_2$ avec $||y_1 - y_2|| = m$, il existe $u' \in N$, tel que $u'(y_1) = y_2$.

En effet, il existe $s \in O^+(q)$ tel que $s(y_1) = x_1$, $s(y_2) = x_2$ (c'est le troisième « cas d'égalité des triangles » cf. Exercice 1). On pose alors $u' = s^{-1}us$, on a $u' \in N$, puisque N est distingué et $u'(y_1) = y_2$.

5) On va maintenant prouver que N contient un renversement en obtenant celui-ci comme composé de « petites » rotations, au sens de 4). Soit $n \in \mathbf{N}$ et ρ_n la rotation d'axe a et d'angle π/n. On a $||x - \rho_n(x)|| = 2(1 - \cos\pi/n)$. Comme $\mathbf{R}$ est archimédien, le rapport π/n tend vers 0 quand n tend vers $+\infty$ et donc pour n assez grand on a $||x - \rho_n(x)|| \leq d$.

On pose alors $x_0 = x$, $x_1 = \rho_n(x), \ldots, x_{i+1} = \rho_n(x_i) \ldots$, et on a $x_n = -x$.
Comme on a $||x_{i+1} - x_i|| = ||x_1 - x_0|| = ||x - \rho_n(x)|| \leq d$, il existe $u_i \in N$ tel que $u_i(x_i) = x_{i+1}$ (cf. 4)). Mais alors, si on pose $v = u_n \ldots u_1$, on a $v \in N$ et $v(x) = -x$.
Il en résulte que v est un renversement (cf. 5.4), et le théorème est démontré.

Remarques 6.2.

1) Nous avons déjà remarqué au § 3 que dans le cas d'une forme quadratique quelconque, les renversements ne sont pas conjugués en général. La démonstration précédente ne subsiste donc pas et, de fait, si on considère une forme non positive sur $\mathbf{R}$, ou une forme définie positive sur $\mathbf{Q}$, le groupe $O^+(3)$ correspondant n'est pas simple (cf. Exercice 4 pour l'étude du groupe euclidien rationnel, où le Théorème 6.1 est spectaculairement en défaut).

2) Si k est un corps ordonné et q une forme définie positive sur k, la Proposition 3.3 est encore valable si et seulement si k est euclidien, c'est-à-dire :

$$\forall a \in k,\ a > 0 \implies a \text{ est un carré dans } k$$

(en effet, il existe alors des bases orthonormées pour q et toutes ses restrictions).

3) Cependant, l'hypothèse k euclidien n'est pas suffisante pour assurer la validité de 6.1. En effet, nous avons utilisé de façon essentielle une autre propriété de $\mathbf{R}$: le fait que $\mathbf{R}$ est archimédien.
Si k n'est pas archimédien, le Théorème 6.1 est inexact : il existe dans k des éléments $\varepsilon > 0$ « infiniment petits », c'est-à-dire tels que : $\forall n \in \mathbf{N}$, $n\varepsilon < 1$. Pour un tel élément on écrit $\varepsilon \ll 1$.
Soit alors $N = \{u \in O^+(q) \mid \forall x \in \mathbf{S}_2 \text{ (sphère unité)},\ q(u(x) - x) \ll 1\}$.
Alors N est un sous-groupe distingué de $O^+(3, k)$, non trivial, cf. Exercice 2, ou aussi [A] Ch. V § 2,3.

4) Il n'est pas immédiat de construire un corps k euclidien et non archimédien. On peut partir de l'anneau $\mathbf{R}[T]$, ordonné par :

$$a_n T^n + \cdots + a_0 > 0 \iff a_n > 0,$$

de sorte qu'on a alors $T \gg 1$. On ordonne $\mathbf{R}(T)$ de l'unique manière compatible, $1/T$ est alors un infiniment petit. Mais $\mathbf{R}(T)$ n'est pas euclidien, et pour trouver k, il faut prendre une clôture réelle de $\mathbf{R}(T)$ (cf. [L] Ch. XI).

7. La simplicité de $PO^+(n, \mathbf{R})$ pour $n \geq 5$.

Théorème 7.1.
Le groupe $PO^+(n, \mathbf{R})$ est simple pour $n = 3$ et pour $n \geq 5$.

Remarque 7.2. Le cas $n = 4$ est exceptionnel et sera étudié au chapitre VII.
Démonstration (de 7.1)

1) Le cas $n = 3$ a été vu au § 6.

2) Soit $\overline{N}$ un sous-groupe distingué de $PO^+(n, \mathbf{R})$, avec $\overline{N} \neq \{1\}$. Il lui correspond un sous-groupe N de $O^+(n, \mathbf{R})$, distingué et contenant le centre $Z(O^+)$, avec $N \neq Z(O^+)$. Il s'agit de prouver que N est égal à $O^+(n, \mathbf{R})$. En vertu de 2.6 et 3.3, il suffit de montrer que N contient un renversement.

3) L'idée est de se ramener à la simplicité de $O^+(3, \mathbf{R})$ (cf. 6.1), par une méthode analogue à celle de I, 8.1.

Soit U un sous-espace de E, de dimension 3, on a un plongement de $O^+(U)$ dans $O^+(E)$ donné par $s \mapsto s \perp \mathrm{Id}_{U^\perp}$ et $O^+(U)$ s'identifie ainsi à un sous-groupe de $O^+(E)$.
Soit $N_1 = N \cap O^+(U)$, on a $N_1 \lhd O^+(U)$, donc, (cf. 6.1) si N_1 n'est pas réduit à $\{\mathrm{Id}\}$, on a $N_1 = O^+(U)$, en particulier si s est un renversement de U, s est dans N_1, donc $s \perp \mathrm{Id}_{U^\perp}$, qui est encore un renversement, est dans N, et on a gagné.

4) Tout revient donc à trouver U, de dimension 3, tel que $N \cap O^+(U) \neq \{\mathrm{Id}\}$, autrement dit, à trouver $s \in N$, $s \neq \mathrm{Id}$, dont l'espace des points fixes V_s soit de dimension $\geq n-3$.
Comme s est une rotation, on aura, en fait, $\dim V_s = n-2$, (cf. 5.3.5) et donc, s sera de la forme $s = \tau_b\tau_c$, où τ_b, τ_c sont des réflexions (cf. 2.4).

5) L'idée, comme au Chapitre I, 8.1 est de prendre pour s un commutateur $s = \rho\sigma\rho^{-1}\sigma^{-1}$ avec $\rho \in N$ et $\sigma \in O^+(n,\mathbf{R})$, ce qui assure que s est dans N, et de choisir σ avec beaucoup de points fixes.
(Attention, comme σ doit être une rotation, on ne peut prendre pour σ une réflexion τ_b qui donnerait pourtant un commutateur de la bonne forme :

$$\text{si } c = \rho(b),\ s = \rho\tau_b\rho^{-1}\tau_b = \tau_c\tau_b.$$

Mais pour affirmer alors $s \in N$, il faudrait que N soit distingué dans $O(n,\mathbf{R})$, et pas seulement dans $O^+(n,\mathbf{R})$).

6) On prend donc $\sigma = \tau_b\tau_a$, on a $\sigma \in O^+$ donc $s = \rho\sigma\rho^{-1}\sigma^{-1} \in N$ et s s'écrit :

$$s = \rho\tau_b\rho^{-1}\rho\tau_a\rho^{-1}\tau_a\tau_b = \tau_{\rho(b)}\tau_{\rho(a)}\tau_a\tau_b$$

et s sera de la forme voulue si on a $\tau_{\rho(a)} = \tau_a$, c'est-à-dire $\rho(a) = \mp a$.
Si de plus, ρ est distinct de $\mp\mathrm{Id}$, on choisira b tel que b et $c = \rho(b)$ ne soient pas colinéaires. Alors $s = \tau_c\tau_b$ sera distinct de Id et on aura réalisé le programme énoncé en 4).

7) On est donc ramené à trouver un élément $\rho \in N$, $\rho \neq \mp\mathrm{Id}$, et possédant **un** point fixe $a \neq 0$ (on mesurera les progrès par rapport à 4)).
Pour ceci, on part de $v \in N$, $v \neq \mp\mathrm{Id}$, et d'un renversement u de plan P. Alors $\rho = vuv^{-1}u^{-1}$ est dans N et ρ est le produit de deux renversements, $u^{-1} = u$ de plan P et vuv^{-1} de plan $v(P)$. Il en résulte que ρ laisse fixe $P^\perp \cap v(P)^\perp$ qui est de dimension $\geq n-4$, donc, (et c'est ici qu'apparaît la nécessité de l'hypothèse $n \geq 5$) strictement positive.
De plus, on aura $\rho \neq \mathrm{Id}$ pourvu que $v(P)$ ne soit pas égal à P, ce qui, puisque v n'est pas une homothétie, est vrai pour au moins un plan P (cf. Chapitre IV §2 Exercice 10). Comme ρ a un point fixe, il est clair qu'on a $\rho \neq -\mathrm{Id}$, et la démonstration est achevée.

Remarques 7.3.

1) Lorsque la dimension n est impaire, l'existence d'un point fixe pour ρ dans $O^+(n,\mathbf{R})$ est claire (cf. 5.3.4) ce qui simplifie la fin de la démonstration.

2) La démonstration ci-dessus montre aussi, en tenant compte de la remarque faite en 5), que si N est un sous-groupe distingué de $O(n,\mathbf{R})$, contenant le centre et distinct du centre, N contient $O^+(n,\mathbf{R})$ et ce, dès que $n \geq 3$ (cf. Chapitre VII pour comprendre le cas $n = 4$).

8. Les automorphismes de $O^+(3,\mathbf{R})$.

Dans ce paragraphe, nous utiliserons le théorème fondamental de la géométrie projective, pour lequel nous renvoyons à [B] Ch. 5 ou à [D] Ch. III § 1. Le lecteur est supposé avoir une certaine familiarité avec quelques notions élémentaires de géométrie projective (alignement, etc).

Théorème 8.1.
Tout automorphisme de $O^+(3,\mathbf{R})$ est intérieur.

Démonstration. On pose $G = O^+(3,\mathbf{R})$. Soit φ un automorphisme de G.
L'idée de la démonstration, analogue à celle de I, 8.7, est de considérer les involutions de G i.e. les $s \in G$ tels que $s^2 = \mathrm{Id}$ et $s \neq \mathrm{Id}$. Dans le cas présent, ce sont exactement les renversements. Comme $\varphi(s^2) = \varphi(s)^2 = \varphi(\mathrm{Id}) = \mathrm{Id}$, on voit que φ transforme renversement en renversement et de même φ^{-1}.
Mais, on peut associer, de manière bijective, à tout renversement, son axe, c'est-à-dire une droite vectorielle de $\mathbf{R}^3$, ou encore, un élément de $\mathbf{P}^2(\mathbf{R})$, espace projectif associé, (cf. Chapitre IV § 5).
On constate donc que φ induit une bijection, notée ψ, de $\mathbf{P}^2(\mathbf{R})$ dans $\mathbf{P}^2(\mathbf{R})$ donnée par la formule $\varphi(s_x) = s_{\psi(x)}$ et nous allons prouver que ψ conserve l'alignement.

Si les points x, y, z de $\mathbf{P}^2(\mathbf{R})$ correspondent aux droites D_x, D_y, D_z, par définition de l'alignement projectif, ces points sont alignés si et seulement si les droites D_x, D_y, D_z sont coplanaires, ce qu'on peut traduire en disant qu'il existe une droite Δ perpendiculaire commune à D_x, D_y, D_z.
On cherche donc à traduire algébriquement sur les renversements le fait que leurs axes soient orthogonaux. On a le lemme suivant, qui est laissé au lecteur :

Lemme 8.2.
Soient s_1, s_2 des renversements d'axes D_1, D_2 distincts. Alors on a $s_1 s_2 = s_2 s_1$ si et seulement si D_1 et D_2 sont orthogonaux.

Alors, si s_a est le renversement d'axe D_a, la condition d'alignement se traduit ainsi : x, y, z sont alignés si et seulement s'il existe un renversement s qui commute à s_x, s_y et s_z.
Mais alors, si on a $ss_x = s_x s$, on a, en appliquant φ, $\varphi(ss_x) = \varphi(s_x s)$ donc $\varphi(s)\varphi(s_x) = \varphi(s_x)\varphi(s)$, et donc, le renversement $\varphi(s)$ commute à $\varphi(s_x)$, $\varphi(s_y)$, $\varphi(s_z)$, i.e. à $s_{\psi(x)}$, $s_{\psi(y)}$, $s_{\psi(z)}$, de sorte que $\psi(x)$, $\psi(y)$, $\psi(z)$ sont alignés.

En vertu du théorème fondamental de la géométrie projective *(loc. cit.)* il existe alors $u \in GL(3,\mathbf{R})$ tel que, si $\overline{u}$ est l'homographie déduite de u, on a $\overline{u} = \psi$.

Si x est un point de $\mathbf{P}^3(\mathbf{R})$ et si D_x est la droite associée on a alors : $u(D_x) = D_{\overline{u}(x)} = D_{\psi(x)}$.
D'autre part, en vertu du lemme 8.2, u conserve l'orthogonalité, donc (cf. V, 7.1), u est une similitude : $u \in GO(3,\mathbf{R})$.
On a donc $u = h_\lambda . v$ où v est une isométrie et h_λ une homothétie (cf. § 1). Mais alors on a $\overline{u} = \overline{v}$ et on peut supposer $u \in O(3,\mathbf{R})$. De plus, si on a $u \in O^-(3,\mathbf{R})$, on a $-u \in O^+(3,\mathbf{R})$ et $\overline{-u} = \overline{u}$, donc on peut supposer $u \in O^+(3,\mathbf{R})$.

Mais alors, on a pour tout $g \in G$, $\varphi(g) = ugu^{-1}$ ce qui prouve que φ est intérieur.

En effet, il suffit de le vérifier sur les renversements, cf. 2.6, or, si s_x est le renversement d'axe D_x on a $\varphi(s_x) = s_{\psi(x)}$, renversement d'axe $D_{\psi(x)}$, par définition de ψ. Par ailleurs, us_xu^{-1} est le renversement d'axe $u(D_x) = D_{\psi(x)}$ (cf. V, 4.8), donc on a bien $\varphi(s_x) = us_xu^{-1}$, cqfd.

Remarque 8.3. Ce type de démonstration s'applique à de nombreux autres groupes (cf. [D] Ch. IV). Des difficultés supplémentaires proviennent du fait qu'il y a, en général plusieurs types d'involutions (l'exemple de $O^+(5,\mathbf{R})$ est instructif à cet égard, cf. Exercice 1).
Signalons enfin le cas $n = 8$ où se produit un phénomène très intéressant, appelé la « trialité ». On a, dans ce cas : $\mathrm{Aut}\, O^+(8)/\mathrm{Int}\, O^+(8) \simeq \mathfrak{S}_3$.

EXERCICES SUR LE CHAPITRE VI

1. Notations et rappels.

1) Soit k un corps totalement ordonné, E un k-espace vectoriel, q une forme positive sur E (i.e. vérifiant $\forall x \in E,\ q(x) \geq 0$) et non dégénérée.
Montrer que q est définie positive.

---

2) Démontrer l'inégalité de Minkowski :

$$||x+y|| \leq ||x|| + ||y||, \quad \text{avec } ||x|| = \sqrt{(x\,|\,x)}.$$

---

3) Avec les notations du § 1, on pose :

$$GO^+(q) = \{u \in GO(q)\,|\,\text{dét } u > 0\}.$$

a) Montrer que $GO^+(q)$ est un sous-groupe distingué de $GO(q)$.
b) Montrer qu'on a des isomorphismes :

$$GO(q) \simeq O(q) \times \mathbf{R}^{+*},$$

$$GO^+(q) \simeq O^+(q) \times \mathbf{R}^{+*},$$

(cf. Chapitre V § 5 Exercice 3).

---

4) Soit k un corps totalement ordonné et E un k-espace vectoriel de dimension n. A quelle condition sur k existe-t-il, pour toute forme q définie positive sur E, une base de E orthonormée pour q ?

---

5) Montrer que les groupes suivants ne sont pas isomorphes :

$$O^+(2,\mathbf{Q}) = \{A \in GL(2,\mathbf{Q})\,|\,{}^tAA = I \text{ et dét } A = 1\} \quad \text{et}$$

$$O_2^+(\mathbf{Q},J) = \{A \in GL(2,\mathbf{Q})\,|\,{}^tAJA = J \text{ et dét } A = 1\} \quad \text{avec} \quad J = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

(on cherchera des éléments d'ordre 4 dans ces groupes).

---

6) Montrer que si on a $k = \mathbf{Q}$ et s'il existe une base orthonormée relative à q sur l'espace E, de dimension ≥ 2, il existe $F \subset E$ tel que $q|_F$ ne possède pas de base orthonormée.

---

2. Générateurs et centres de $O(q)$ et $O^+(q)$.

1) Démontrer les théorèmes 2.1, 2.4, 2.6 pour une forme anisotrope sur un corps quelconque, (de caractéristique différente de 2, bien sûr).

---

2) On reprend les notations du théorème 2.4.
Soit $u \in O(q)$ et $m_u = \inf\{r \in \mathbf{N} \,|\, u = \tau_1 \dots \tau_r\}$ où les τ_i sont des réflexions orthogonales.

a) Montrer qu'on a $m_u = p_u$.

b) Montrer que, si $u \in O^+(q)$, on a $\dim F_u \equiv n \pmod 2$.

---

3. Conjugaison et commutateurs.

1) Soit q la forme quadratique sur $\mathbf{Q}^2$ définie par $q(x,y) = x^2 + y^2$. Soit $e_1 = (x_1, y_1)$, $e_2 = (x_2, y_2)$ deux vecteurs non nuls de $\mathbf{Q}^2$, $D_1 = <e_1>, D_2 = <e_2>$ les droites engendrées.
A quelle condition existe-t-il $u \in O(q)$ tel que $u(D_1) = D_2$?
Montrer que les orbites de $O(q)$ dans l'ensemble $\mathbf{P}^1(\mathbf{Q})$ des droites de $\mathbf{Q}^2$ sont en bijection avec une partie de $\mathbf{Q}^*/\mathbf{Q}^{*2}$ que l'on précisera, à l'aide de II, 6.9.

---

2) Soit k un corps ordonné euclidien (i.e. tel que tout $a \in k$, $a \geq 0$, est un carré dans k), et soit q une forme définie positive sur k.
Montrer que le lemme 3.1 est encore vrai. Réciproquement, si un corps ordonné k est tel que le lemme 3.1 soit vrai pour toute forme définie positive, montrer que k est euclidien.

---

3) Montrer que sur un corps ordonné euclidien les Propositions 3.3 et 3.4 sont encore valables.

---

4) *Un lemme général sur les groupes dérivés.*
Soit G un groupe, G^2 le sous-groupe de G engendré par les carrés des éléments de G.

a) Montrer qu'on a $D(G) \subset G^2$.

b) On suppose G engendré par des involutions. Montrer qu'on a $G^2 = D(G)$.

c) On suppose G engendré par des éléments $s \in S$, tous conjugués. Montrer que $G/D(G)$ est monogène ; cas où les éléments de S sont des involutions.

d) Appliquer ces résultats aux calculs des groupes dérivés de $\mathfrak{S}_n$, $\mathfrak{A}_n$, $O(q)$, $O^+(q)$... (remarquer que les cycles d'ordre 3, les renversements, ... sont des carrés).

5) Montrer que $D(O(2,\mathbf{Q}))$ n'est pas égal à $O^+(2,\mathbf{Q})$ (remarquer que $O^+(2,\mathbf{Q})$ est commutatif, isomorphe à $\mathbf{U}\cap\mathbf{Q}(i)$, cf. §4, et que i n'est pas un carré dans ce groupe, et conclure par Exercice 4,b)).

---

4. La dimension 2 et les angles.

1) Démontrer la Proposition 4.1 pour une forme anisotrope quelconque.

---

2) Vérifier matriciellement l'assertion 4.1.1.

---

3) Montrer que si k est euclidien et si q est définie positive sur E, k-espace de dimension 2, le groupe $O^+(q)$ est encore isomorphe au groupe des matrices $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$, avec $a,b\in k$, $a^2+b^2=1$, donc aussi au groupe $\mathbf{U}_k$ des éléments $a+bi$ de l'extension $K=k(i)$ avec $i^2=-1$, tels que $a^2+b^2=1$.
Pour le cas général, cf. Chapitre VIII §6.

---

4) On reprend les notations de l'exercice 9 du chapitre III §1.

a) Montrer que K est euclidien.

b) Montrer que K est archimédien.

c) Montrer que tous les théorèmes énoncés dans ce chapitre restent vrais pour une forme positive sur K, sauf :

- la proposition 4.2 (cf. ci-après),
- le théorème 5.1 qui n'est vrai qu'en dimension ≤ 3 (et sans cosinus dans la forme matricielle).

Vérifier soigneusement, en particulier, le théorème 6.1.

d) Montrer qu'il n'existe pas d'isomorphisme analogue à l'exponentielle, $\varphi : (K/a\mathbf{Z},+) \longrightarrow \mathbf{U}_K$ avec $a\in K^*$ et $\mathbf{U}_K=\{a+ib\in\mathbf{U}\,|\,a,b\in K\}$. (Montrer qu'il existe dans $K/a\mathbf{Z}$ des éléments de tout ordre $n\in\mathbf{N}^*$, mais que $\mathbf{U}_K$ ne contient pas d'élément d'ordre 11, cf. Chapitre III §1 Exercice 9, e)). On consultera [D1] Annexe 1 pour des compléments sur ce sujet.

---

5. Structure des éléments de $O(q)$.

1) *Une autre démonstration du théorème 5.1.*
Les notations sont celles du §1, on suppose $n\geq 3$. On désigne par S la sphère unité de E, on a donc $S=\{x\in E\,|\,||x||=\sqrt{q(x)}=1\}$. Soit $u\in O(q)$.

a) Montrer qu'il existe $x\in S$ tel que $||u(x)-x||$ soit minimum. Montrer que ceci équivaut à dire que $a=(x\,|\,u(x))$ est maximum.

b) On suppose $y=u(x)\neq\mp x$. Montrer que a est distinct de ∓ 1. Soit P le plan $<x,y>$, on pose $b=(x\,|\,u^2(x))$. Montrer qu'on a $1+b=2a^2$ (raisonner par l'absurde en montrant que si $1+b\neq 2a^2$, il existe $z\in S\cap P$ tel que $(z\,|\,u(z))>a$). Interpréter la condition $1+b=2a^2$ en termes d'angles.

c) Montrer que le plan P est stable par u (si $z = u(y)$ on écrira x, y, z dans un système orthonormé x, j, k, avec $j \in P$).

d) Montrer qu'il existe un sous-espace F de E, distinct de $\{0\}$ et E et stable par u et en déduire une nouvelle démonstration de 5.1.

---

2) Soit k un corps totalement ordonné, muni de la topologie de l'ordre. Soit C le cercle unité de k^2 :

$$C = \{(x,y) \in k^2 \mid x^2 + y^2 = 1\}, \quad \text{et soit} \quad \Gamma = \{(x,y) \in C \mid x \geq 0\}.$$

a) Montrer que l'application $\varphi : [-1,1] \longrightarrow \Gamma$ donnée par :

$$\varphi(t) = \left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}\right)$$

est un homéomorphisme et calculer φ^{-1}.

b) Déduire de a) que la sphère unité

$$S = \left\{(x_1, \ldots, x_n) \in k^n \;\middle|\; \sum_{i=1}^{n} x_i^2 = 1\right\}$$

est compacte si et seulement si les segments $[a,b]$ pour $a, b \in k$ le sont.

c) Montrer que si S est compacte, k est euclidien (utiliser b) et pour trouver $\sqrt{a}$, pour $a > 0$, une suite récurrente du type $x_{n+1} = x_n + \lambda(x_n^2 - a)$ avec λ convenable ou, si l'on préfère, la fonction $f(x) = |x^2 - a|$).

d) Montrer que si S est compacte, k est archimédien (utiliser b) et remarquer que si k n'est pas archimédien la suite des entiers est bornée mais n'a pas de valeur d'adhérence).

e) Déduire de ce qui précède que si S est compacte, le corps k est isomorphe à $\mathbf{R}$ (cf. par exemple, [LFA] Ch. 1 Th. 1.8.5).

---

3) Montrer que la démonstration du théorème 5.1 donnée dans le cours reste valable pour un corps k réellement clos (i.e. euclidien et tel que tout polynôme de degré impair admette une racine dans k, tout ceci impliquant que $k(\sqrt{-1})$ est algébriquement clos, cf. [L] Ch. XI § 2 Th. 1).
Comparer avec la démonstration de l'exercice 1 (cf. Exercice 2).

---

6. La simplicité de $O^+(3,\mathbf{R})$.

1) *Troisième cas d'égalité des triangles.*
On reprend les notations du § 1, avec $n \geq 2$.
Soient $x, y, x', y' \in E$. On suppose $||x|| = ||x'||$, $||y|| = ||y'||$, $||x-y|| = ||x'-y'||$.
Montrer qu'il existe $u \in O(q)$ tel que $u(x) = x'$, $u(y) = y'$ (si x et y ne sont pas colinéaires, considérer le plan $<x,y>$).
On suppose $n = 2$. Montrer que si x, y ne sont pas colinéaires u est unique.
On suppose $n \geq 3$. Montrer qu'on peut imposer $u \in O^+(q)$.

---

2) Soit k un corps ordonné non archimédien et q la forme définie sur k^3 par $q(x,y,z) = x^2 + y^2 + z^2$. On dit qu'un élément $\varepsilon > 0$ de k est infiniment petit si et seulement si pour tout $n \in \mathbf{N}$ on a $n\varepsilon < 1$ et on écrit alors $\varepsilon \ll 1$.
Soit $N = \{u \in O^+(q) \mid \forall x \in S,\ q(u(x) - x) \ll 1\}$ (où S désigne la sphère unité de l'espace vectoriel E).

a) Montrer que N est un sous-groupe distingué de $O^+(q)$.

b) Montrer qu'on a $N \neq O^+(q)$ (considérer un renversement).

c) Montrer qu'on a $N \neq \{\mathrm{Id}\}$ (utiliser une matrice de rotation « d'angle » θ avec $t =$ « $\tan\dfrac{\theta}{2}$ » infiniment petit, cf. § 5 Exercice 2).

d) En déduire que $O^+(q)$ n'est pas simple.

---

3) *Étude de l'anneau* $\mathbf{Z}_2$.
On rappelle que si x est un rationnel non nul, x s'écrit de manière unique $x = \dfrac{p}{q}$, avec $p, q \in \mathbf{Z}$, $(p,q) = 1$ et $q > 0$. Dans toute la suite, le symbole $\dfrac{p}{q}$ aura toujours cette signification.
On pose $\mathbf{Z}_2 = \{x = \dfrac{p}{q} \in \mathbf{Q}^* \mid q \text{ impair}\} \cup \{0\}$.

a) Montrer que $\mathbf{Z}_2$ est un sous-anneau de $\mathbf{Q}$ contenant $\mathbf{Z}$.

b) Calculer $\mathbf{Z}_2^*$ et montrer que l'idéal $2\mathbf{Z}_2$ est l'unique idéal maximal de $\mathbf{Z}_2$ (on dit que $\mathbf{Z}_2$ est un anneau local).

c) Montrer qu'on a $\mathbf{Z}_2/2\mathbf{Z}_2 \simeq \mathbf{F}_2$.

d) Soient $x, y, z \in \mathbf{Q}$ tels que $x^2 + y^2 + z^2 = 1$. Montrer qu'on a $x, y, z \in \mathbf{Z}_2$.

---

4) *Étude du groupe* $O_3(\mathbf{Q}, E)$.
Dans cet exercice, E désigne la forme euclidienne :

$$E(x,y,z) = x^2 + y^2 + z^2.$$

On pose $O(3,\mathbf{Q}) = O_3(\mathbf{Q}, E) = \{A \in \mathbf{M}(3,\mathbf{Q}) \mid {}^tAA = I\}$.

a) Montrer que $O(3,\mathbf{Q})$ est inclus dans $\mathbf{M}(3,\mathbf{Z}_2)$ (cf. Exercice 3).

b) Pour $r \in \mathbf{N}^*$, on pose :

$$G_r = \{A \in O(3,\mathbf{Q}) \mid A = I + 2^r B \text{ avec } B \in \mathbf{M}(3,\mathbf{Z}_2)\}.$$

Montrer que G_r est un sous-groupe distingué de $O(3,\mathbf{Q})$.

c) Montrer qu'on a $\bigcap_{r \in \mathbf{N}^*} G_r = \{I\}$.

d) Montrer que G_1 est distinct de $O(3,\mathbf{Q})$ et que G_1 n'est pas inclus dans $O^+(3,\mathbf{Q})$.

e) Montrer que pour $r \geq 1$, on a $G_{r+1} \subset G_r$ et $G_{r+1} \neq G_r$ (on pourra considérer $u = \tau_{e_1}\tau_{e_2}$ où τ_{e_i} est la réflexion relative à e_i, avec $e_1 = (1,0,0)$, $e_2 = (1, 2^{r-1}, 0)$).

f) Montrer que, pour $r \geq 2$, on a $G_r \subset O^+(3,\mathbf{Q})$ (considérer, sinon, un vecteur propre relatif à -1 de $A \in G_r$).

g) Montrer que, pour $r \geq 2$, on a $G_r/G_{r+1} \simeq (\mathbf{Z}/2\mathbf{Z})^3$ (si $A = I + 2^r B$ est dans G_r on lui associera la matrice $\overline{B}$ obtenue par réduction modulo 2 à partir de B et on montrera que $\overline{B}$ est une matrice alternée, i.e. antisymétrique avec en plus, caractéristique 2 oblige, $\overline{b}_{ii} = 0$).

h) Montrer qu'on a un isomorphisme $O(3,\mathbf{Q})/G_1 \simeq O(3,\mathbf{F}_2)$, où l'on a posé

$$O(3,\mathbf{F}_2) = \{A \in \mathbf{M}(3,\mathbf{F}_2) \mid {}^tAA = I\} \simeq \mathfrak{S}_3$$

(si A est dans $O(3,\mathbf{Q})$, on lui associera aussi sa réduction $\overline{A}$ modulo 2).

i) Montrer qu'on a $G_1/G_2 \simeq (\mathbf{Z}/2\mathbf{Z})^4$ (utiliser le même homomorphisme qu'en g), mais montrer que cette fois l'image est le sous-espace de $\mathbf{M}(3,\mathbf{F}_2)$ engendré par les matrices diagonales et par la matrice dont tous les termes valent 1).

j) Résumer les propriétés de $O(3,\mathbf{Q})$ sur un diagramme où l'on figurera les sous-groupes G_r et les divers quotients. Comparer au Théorème 6.1 du cas réel!

---

8. Les automorphismes de $O^+(3,\mathbf{R})$.

1) On désigne par E un espace vectoriel réel de dimension 5, muni d'une forme euclidienne q.

a) Soit $u \in O^+(q)$ tel que $u^2 = \mathrm{Id}$ et $u \neq \mathrm{Id}$. Montrer que u est un renversement ou que $-u$ est une réflexion (dans ce cas, u sera appelée une antiréflexion).

b) Montrer que les antiréflexions engendrent $O^+(q)$ et qu'elles sont toutes conjuguées.

c) Soit σ un renversement et τ une antiréflexion. Calculer les centralisateurs $c(\sigma)$ et $c(\tau)$ de σ et τ dans $O^+(q)$.

d) Montrer qu'on a $D(c(\sigma)) \simeq O^+(2,\mathbf{R}) \times O^+(3,\mathbf{R})$ et $D(c(\tau)) \simeq O^+(4,\mathbf{R})$. En déduire que $c(\sigma)$ et $c(\tau)$ ne sont pas isomorphes.

e) Soit φ un automorphisme de $O^+(q)$. Montrer que l'image par φ d'une antiréflexion est une antiréflexion.

f) Si τ est une antiréflexion, on lui associe la droite $D = \mathrm{Ker}\ (\tau - \mathrm{Id})$ et on écrit alors $\tau = \tau_D$. Montrer que φ induit une bijection ψ de $\mathbf{P}(E)$ par la formule $\varphi(\tau_D) = \tau_{\psi(D)}$.

g) Soient D_1, D_2, D_3 trois droites distinctes de E, $\tau_i = \tau_{D_i}$ l'antiréflexion associée à D_i et posons $C = c(\tau_1) \cap c(\tau_2) \cap c(\tau_3)$.
Préciser le groupe C et montrer en particulier qu'on a $D(C) \simeq O^+(3,\mathbf{R})$ si D_1, D_2, D_3 sont coplanaires et $D(C) \simeq O^+(2,\mathbf{R})$ sinon. En déduire que ψ conserve l'alignement.

h) En considérant le centre du groupe $c(\tau_1) \cap c(\tau_2)$, montrer que ψ conserve l'orthogonalité.

i) Montrer, comme en 8.1, que φ est intérieur.

j) Généraliser cet exercice au groupe $O^+(2n+1,\mathbf{R})$, avec $n \in \mathbf{N}^*$.

VII. QUATERNIONS

Les quaternions, inventés en 1843 par le mathématicien irlandais William Hamilton, jouent, en dimension 3 et 4, vis à vis des groupes orthogonaux, un rôle analogue à celui des nombres complexes en dimension 2, à la différence notoire qu'il n'y a plus commutativité.

1. Définition du corps $\mathbf{H}$.

a) Définition.

Théorème et définition 1.1.
Il existe une algèbre $\mathbf{H}$ *de dimension 4 sur* $\mathbf{R}$, *appelée algèbre des* **quaternions**, *munie d'une base* $1, i, j, k$ *telle que :*
a) *l'élément* 1 *est élément neutre pour la multiplication,*
b) *on a les formules :*

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

Le corps des nombres réels est isomorphe à la sous-algèbre des quaternions de la forme $a.1$, $a \in \mathbf{R}$, *avec laquelle on l'identifie. Un quaternion s'écrit alors :*

$$q = a + bi + cj + dk, \quad \text{avec } a, b, c, d \in \mathbf{R}.$$

Démonstration. Il y a plusieurs méthodes pour prouver l'existence de $\mathbf{H}$, mais toutes requièrent un minimum de calculs. C'est assez normal, dans la mesure où cette existence est non banale, voire exceptionnelle (cf. § 4 b,c) et où on en déduit presque directement des propriétés non triviales des groupes orthogonaux (cf. § 2 et § 3).

a) La première méthode consiste à prendre pour $\mathbf{H}$ l'espace $\mathbf{R}^4$ muni d'une base $1, i, j, k$ et à définir la multiplication par bilinéarité à partir des formules ci-dessus. Il faut alors vérifier l'associativité sur la base. C'est facile, mais fastidieux, encore qu'on puisse réduire le nombre de calculs en utilisant les symétries des formules (cf. [Bl] Ch. I § 1 ou ci-dessous Exercice 2).

b) Une deuxième méthode consiste à trouver une représentation de H comme sous-algèbre de $\mathbf{M}(4, \mathbf{R})$. Pour ceci, on suppose la construction déjà effectuée et on fait opérer $\mathbf{H}$ sur lui-même par multiplication à gauche : on pose, pour $q, q' \in \mathbf{H}$,

$T_q(q') = qq'$. L'application T_q est linéaire et si $q = a + bi + cj + dk$, la matrice de T_q dans la base $1, i, j, k$ est :

$$M(q) = \begin{pmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{pmatrix}$$

Cette analyse étant faite, on *définit* alors $\mathbf{H}$ comme l'ensemble des matrices de cette forme et on vérifie les conditions requises. En particulier, on a les formules :

$$i = \left(\begin{array}{cc|cc} 0 & -1 & \multicolumn{2}{c}{(0)} \\ 1 & 0 & & \\ \hline \multicolumn{2}{c|}{(0)} & 0 & -1 \\ & & 1 & 0 \end{array}\right) \quad ; \quad j = \left(\begin{array}{cc|cc} \multicolumn{2}{c|}{(0)} & -1 & 0 \\ & & 0 & 1 \\ \hline 1 & 0 & \multicolumn{2}{c}{(0)} \\ 0 & -1 & & \end{array}\right)$$

$$\text{et} \quad k = \left(\begin{array}{cc|cc} \multicolumn{2}{c|}{(0)} & 0 & -1 \\ & & -1 & 0 \\ \hline 0 & 1 & \multicolumn{2}{c}{(0)} \\ 1 & 0 & & \end{array}\right)$$

c) On peut aussi représenter $\mathbf{H}$ comme sous-algèbre de $\mathbf{M}(2, \mathbf{C})$, cf. Exercice 4 ou § 4 a).

b) Conjugué, norme, inverse.

Définition 1.2.
Soit $q \in \mathbf{H}$, $q = a + bi + cj + dk$ avec $a, b, c, d \in \mathbf{R}$. On définit le conjugué $\overline{q}$ de q par la formule $\overline{q} = a - bi - cj - dk$.

Proposition 1.3.
L'application $q \mapsto \overline{q}$ de $\mathbf{H}$ dans $\mathbf{H}$ est un antiautomorphisme de $\mathbf{H}$ (i.e. elle est $\mathbf{R}$-linéaire et vérifie, pour tous $q_1, q_2 \in \mathbf{H}$, $\overline{q_1 q_2} = \overline{q_2}\,\overline{q_1}$.)

Démonstration. On peut, soit le vérifier sur la base, soit utiliser la représentation matricielle car on a alors :

$$M(\overline{q}) = {}^t M(q) \text{ et on a bien } {}^t(AB) = {}^t B\, {}^t A.$$

Remarques 1.4.
1) On a $\overline{\overline{q}} = q$: comme sur les complexes la conjugaison est une involution.
2) On a $q \in \mathbf{R} \iff \overline{q} = q$.
3) Posons $P = \{q = bi + ci + dk \mid b, c, d \in \mathbf{R}\}$, on dit que P est l'espace des ***quaternions purs*** et on a : $q \in P \iff \overline{q} = -q$.

Définition-Proposition 1.5.
Soit $q = a + bi + cj + dk \in \mathbf{H}$, on définit la **norme** *$N(q)$ de q en posant :*

$$N(q) = q\overline{q} = \overline{q}q = a^2 + b^2 + c^2 + d^2.$$

Alors, on a $N(q) \in \mathbf{R}^+$.

En effet, comme on a $\overline{q\overline{q}} = \overline{\overline{q}}\,\overline{q} = q\overline{q}$, le quaternion $q\overline{q}$ est réel et un calcul immédiat montre qu'on a $q\overline{q} = a^2 + b^2 + c^2 + d^2$. De même, $\overline{q}q$ est réel et vaut encore $a^2 + b^2 + c^2 + d^2$.

Notons que $q \mapsto N(q)$ est une forme quadratique euclidienne sur $\mathbf{H}$, de forme polaire $\varphi(q_1, q_2) = \frac{1}{2}(q_1\overline{q_2} + q_2\overline{q_1})$. La base $1, i, j, k$ est orthonormée relativement à N et la conjugaison est une symétrie orthogonale, d'espaces propres $\mathbf{R}$ et P.

Théorème 1.6.
1) L'algèbre $\mathbf{H}$ *est un corps (non commutatif).*
2) Le centre de $\mathbf{H}$ *est égal à* $\mathbf{R}$.
3) On a $N(q_1q_2) = N(q_1)N(q_2)$, *donc* $N : \mathbf{H}^* \longrightarrow \mathbf{R}^{+*}$ *est un homomorphisme de groupes, surjectif, dont le noyau, groupe des quaternions de norme* 1, *sera noté* G.

Démonstration.
1) On a $N(q) = 0 \Longleftrightarrow q = 0$, donc, si $q \neq 0$, q a un inverse $q^{-1} = \frac{1}{N(q)}\overline{q}$. On notera que, comme $N(q)$ est réel, il commute à $\overline{q}$ et qu'on peut écrire :

$$\overline{q}N(q)^{-1} = N(q)^{-1}\overline{q} = \frac{1}{N(q)}\overline{q}.$$

2) Il est clair que $\mathbf{R}$ est central puisque $\mathbf{H}$ est une $\mathbf{R}$-algèbre. Réciproquement, si q est dans $Z(\mathbf{H})$ on écrit $qi = iq$, $qj = jq$ et un calcul immédiat montre qu'on a $q \in \mathbf{R}$.
3) On a $N(q_1q_2) = q_1q_2\overline{q_1}\overline{q_2} = q_1q_2\overline{q_2}\,\overline{q_1} = q_1N(q_2)\overline{q_1} = q_1\overline{q_1}N(q_2) = N(q_1)N(q_2)$ car $N(q_2)$ est central. Si a est dans $\mathbf{R}^+$, on a $N(\sqrt{a}) = a$, donc N est surjectif.

Remarques 1.7.
1) Si q est dans G, i.e. si on a $N(q) = 1$, on a $q^{-1} = \overline{q}$.
2) Identifions $\mathbf{H}$ à $\mathbf{R}^4$ muni de sa topologie naturelle. Alors, le groupe G des quaternions de norme 1 s'écrit $G = \{(a, b, c, d) \in \mathbf{R}^4 \mid a^2 + b^2 + c^2 + d^2 = 1\}$. Il est donc homéomorphe à la sphère $\mathbf{S}^3$ et en particulier G est connexe.
3) Pour un quaternion q on a les équivalences suivantes :

$$q \in \mathbf{R} \Longleftrightarrow q^2 \in \mathbf{R}^+, \qquad q \in P \Longleftrightarrow q^2 \in \mathbf{R}^-.$$

En effet si on a $q = a \in \mathbf{R}$, on a $N(q) = q^2 = a^2 \in \mathbf{R}^+$, tandis que si p est dans P, on a $N(p) = p\overline{p} = -p^2$, de sorte que $p^2 = -N(p)$ est dans $\mathbf{R}^-$.
Réciproquement, on écrit $q = a+p$, avec $a \in \mathbf{R}$ et $p \in P$ et on a $q^2 = a^2+p^2+2ap$, donc, si q^2 est réel, on a $2ap = 0$, et a ou p est nul, d'où le résultat.

2. Opération de $\mathbf{H}$ sur $\mathbf{R}^3$.

a) Quaternions et groupe orthogonal.

Le corps $\mathbf{H}$ n'est pas commutatif, donc l'opération de $\mathbf{H}^*$ sur $\mathbf{H}$ par automorphismes intérieurs est non banale. En fait, il suffit de faire opérer G, car si q est dans $\mathbf{H}^*$, il s'écrit $q = \lambda r$ avec $\lambda = \sqrt{N(q)} \in \mathbf{R}$ et $r \in G$ et comme λ est central, il ne donne rien dans les automorphismes intérieurs.

On pose donc, pour $q \in G$ et $q' \in \mathbf{H}$:

$$S_q(q') = qq'q^{-1} = qq'\overline{q}.$$

Nous allons étudier cette opération et montrer qu'elle donne une paramétrisation du groupe $O^+(3, \mathbf{R})$ par le groupe G, un peu comme on avait paramétré $O^+(2, \mathbf{R})$

par le groupe des nombres complexes de module 1, mais avec toutefois quelques différences.

1) L'application $S_q : \mathbf{H} \longrightarrow \mathbf{H}$ est $\mathbf{R}$-linéaire et bijective car on a $S_{\overline{q}} = (S_q)^{-1}$. On obtient donc ainsi une application :

$$S : G \longrightarrow GL(4, \mathbf{R}).$$

2) L'application S est un homomorphisme car on a $S_{q_1 q_2}(q') = q_1 q_2 q' \overline{q}_2 \overline{q}_1 = S_{q_1} S_{q_2}(q')$ et son noyau est $Z(\mathbf{H}) \cap G = \mathbf{R} \cap G = \{1, -1\}$.

3) On a, pour $a \in \mathbf{R}$, $S_q(a) = a$, et donc $S_q|_{\mathbf{R}} = \mathrm{Id}_{\mathbf{R}}$.

4) Par ailleurs, S_q conserve la norme, i.e. vérifie $N(S_q(q')) = N(q')$. En effet, on a $N(qq'\overline{q}) = N(q)N(q')N(\overline{q}) = N(q')$, car q est de norme 1. Autrement dit, S_q est donc un élément du groupe orthogonal euclidien défini par N : $S_q \in O(N) \simeq O(4, \mathbf{R})$.

5) Mais, pour N, l'espace des quaternions purs P est l'orthogonal de $\mathbf{R}$, et comme on a $S_q|_{\mathbf{R}} = \mathrm{Id}_{\mathbf{R}}$, P est stable par S_q.
On pose alors $s_q = S_q|_P$, on a encore $s_q \in O(N|_P) \simeq O(3, \mathbf{R})$ et $s : G \longrightarrow O(3, \mathbf{R})$ est un homomorphisme de noyau $\{1, -1\}$.

6) Munissons $O(3, \mathbf{R})$ de sa topologie naturelle, obtenue en le considérant comme sous-espace de $\mathbf{M}(3, \mathbf{R})$, lui-même identifié à $\mathbf{R}^9$. L'application s est alors continue, comme on le voit en calculant la matrice de s_q dans la base i, j, k. En effet, si $q = a + bi + cj + dk$, les coefficients de la matrice sont des polynômes homogènes de degré 2 en a, b, c, d. Par exemple, on a $s_{11} = a^2 + b^2 - c^2 - d^2$, $s_{12} = 2(bc - ad)\ldots$ Mais, le déterminant, dét : $O(3, \mathbf{R}) \longrightarrow \{1, -1\}$ est lui aussi une application continue, donc l'application composée dét $\circ\, s : G \longrightarrow \{1, -1\}$ est continue. Comme G est connexe (cf. 1.7.2), l'image de G par dét $\circ\, s$ est connexe, donc un singleton, et comme on a $s(1) = \mathrm{Id}$, c'est nécessairement $\{+1\}$. Autrement dit, on a $s(G) \subset O^+(3, \mathbf{R})$.

7) Montrons enfin l'égalité $s(G) = O^+(3, \mathbf{R})$. Soit $p \in P \cap G$. On calcule $s_p(p) = p\overline{p}p = p$, ce qui prouve que s_p fixe p, donc est une rotation d'axe (p). D'autre part, comme p est dans $P \cap G$ on a $\overline{p} = -p$, donc $p^2 = -p\overline{p} = -1$ et $(s_p)^2 = s_{p^2} = s_{-1} = \mathrm{Id}$, donc s_p est une involution et c'est donc le renversement d'axe $< p >$. On obtient ainsi tous les renversements de $O^+(3, \mathbf{R})$, et comme ils engendrent le groupe, on a bien $s(G) = O^+(3, \mathbf{R})$. En définitive, on a prouvé le théorème suivant :

Théorème 2.1.
Soit G le groupe des quaternions de norme 1. On a un isomorphisme :

$$\overline{s} : G/\{1, -1\} \xrightarrow{\sim} O^+(3, \mathbf{R}).$$

Remarques 2.2.
1) La suite exacte :

$$1 \longrightarrow \{1, -1\} \longrightarrow G \xrightarrow{s} O^+(3, \mathbf{R}) \longrightarrow 1$$

n'est pas scindée (cf. Chapitre I §6). Sinon, on aurait un sous-groupe H de G tel que $s|_H$ soit un isomorphisme de H sur $O^+(3, \mathbf{R})$. Mais alors, si g est dans G on

aurait g ou $-g \in H$. Si on prend $p \in P \cap G$, on a $p^2 = (-p)^2 = -1$, donc $-1 \in H$, contradiction.

2) On peut se demander comment décrire les isométries négatives de $\mathbf{H}$ à partir de G, en particulier les réflexions.
Soit $q \in G$, et τ_q la réflexion de droite $<q>$ (i.e. d'hyperplan $<q>^\perp$). On sait (cf. V, 4.7.3) que $\tau_q(x)$ est donné, pour tout $x \in \mathbf{H}$, par la formule :

$$\tau_q(x) = x - 2\frac{\varphi(q,x)}{\varphi(q,q)}\, q.$$

donc $\tau_q(x) = x - (q\overline{x} + x\overline{q})q = -q\overline{x}q$.
Si on a $x \in P$ et $q \in P \cap G$, on a donc : $\tau_q(x) = -qx\overline{q} = -s_q(x)$.
Ceci permet d'éviter le recours à l'argument topologique dans 2.1, 6). En effet, si pour un $q \in G$ on avait $s_q \in O^-(P)$, s_q s'écrirait $s_q = \tau_{p_1} \dots \tau_{p_r}$ avec $p_i \in P \cap G$ et r impair et on aurait donc $s_q(x) = qx\overline{q} = -p_1 \dots p_r\, x\, \overline{p_1 \dots p_r} = -q_1 x \overline{q_1}$ pour tout $x \in P$ (en posant $q_1 = p_1 \dots p_r$). Posons alors $q_2 = \overline{q_1} q$. On a donc, $q_2 x \overline{q_2} = -x = \overline{x}$ pour tout $x \in P$ donc aussi pour tout $x \in \mathbf{H}$. Mais, comme la conjugaison est un anti-automorphisme, ceci est impossible.

b) Application : Calcul des automorphismes de $\mathbf{H}$.

Théorème 2.3.
Tout automorphisme de corps de $\mathbf{H}$ *est intérieur, i.e. de la forme* S_q, *cf. ci-dessus.*

Démonstration. (cf. aussi [Bl] Prop. 1.5). Soit u un automorphisme de $\mathbf{H}$, u conserve le centre de $\mathbf{H}$, c'est-à-dire $\mathbf{R}$ et $u|_{\mathbf{R}}$ est un automorphisme de $\mathbf{R}$, donc $u|_{\mathbf{R}} = \mathrm{Id}_{\mathbf{R}}$ (cf. Chapitre V § 1 Exercice 1).
Par ailleurs, on a $q \in P \iff q^2 \in \mathbf{R}^-$ (cf. 1.7.3), donc, si on a $q \in P$, on en déduit $u(q)^2 = u(q^2) = q^2 \in \mathbf{R}^-$, donc $u(q) \in P$: u laisse P stable. De plus, pour $q \in P$, on a $N(q) = -q^2$ et donc, comme $u(q)^2 = q^2$, on a $N(q) = N(u(q))$: u conserve la norme. On a donc prouvé :

$$u|_P \in O(N|_P) = O(3,\mathbf{R}).$$

Or, i, j, k est une base orthonormée de P, relativement à N et donc $i' = u(i)$, $j' = u(j)$, $k' = u(k)$ est orthonormée. Il existe donc $\varepsilon = \mp 1$ tel que i, j, k et $i', j', \varepsilon k'$ soient de même orientation et, en vertu du théorème 2.1, il existe $q \in G$ tel que l'on ait $i' = s_q(i)$, $j' = s_q(j)$ et $\varepsilon k' = s_q(k)$ (avec les notations de 2.1). Mais, u et S_q sont des automorphismes de $\mathbf{H}$ et donc, on a :

$$u(ij) = u(k) = k' = u(i)u(j) = i'j',$$

et de même $s_q(ij) = \varepsilon k' = s_q(i)s_q(j) = i'j'$ donc $\varepsilon = +1$.
On a donc $u|_P = s_q = S_q|_P$, mais aussi $u|_{\mathbf{R}} = S_q|_{\mathbf{R}} = \mathrm{Id}$, donc $u = S_q$ et u est un automorphisme intérieur.

3. La structure de $O^+(4,\mathbf{R})$.

Il y a une autre opération de $\mathbf{H}$ sur $\mathbf{H}$, celle vue au § 1, à savoir la translation à gauche T_q, définie par $T_q(q') = qq'$. Si q est dans G, il est clair que T_q conserve N, donc T_q, comme S_q, est dans $O(N) = O(4,\mathbf{R})$.
Cependant, ni T_q, ni S_q ne permettent d'atteindre $O^+(4,\mathbf{R})$ tout entier, (cf. Exercice 1), mais la conjonction des deux va y parvenir.

Précisément, on fait opérer le produit direct $G \times G$ sur $\mathbf{H}$. Soient $q_1, q_2 \in G$. On pose, pour $q \in \mathbf{H}$, $S_{q_1,q_2}(q) = q_1 q \overline{q_2}$. Notons que l'on a $S_{q_1,q_2} = T_{q_1\overline{q_2}} \circ S_{q_2}$.
On a alors les propriétés suivantes :

1) S_{q_1,q_2} est linéaire bijective, donc est dans $GL_{\mathbf{R}}(\mathbf{H}) \simeq GL(4,\mathbf{R})$ et l'application $S : G \times G \longrightarrow GL(4,\mathbf{R})$, définie par $S(q_1,q_2) = S_{q_1,q_2}$ est un homomorphisme.

2) L'application S_{q_1,q_2} conserve N, donc est dans $O(N) \simeq O(4,\mathbf{R})$.

3) Comme S est continu et $G \times G$ connexe, le même raisonnement qu'au §2 montre qu'on a $S(G \times G) \subset O^+(N) \simeq O^+(4,\mathbf{R})$.

4) Calculons Ker S : si $S_{q_1,q_2} = \mathrm{Id}$, on a pour tout $q \in \mathbf{H}$, $q_1 q \overline{q_2} = q$. Pour $q = 1$, on trouve $q_1 = q_2$, on voit ensuite que q_1 est central, donc $q_1 = \mp 1 = q_2$. En définitive, on a montré :

$$\mathrm{Ker}\ S = \{(1,1),(-1,-1)\}.$$

5) On a $\mathrm{Im}\, S = O^+(N)$. En effet, soit $g \in O^+(N)$, il y a deux cas :
1) Si on a $g(1) = 1$, comme $P = 1^\perp$, on a $g(P) = P$ et $g|_P \in O^+(3,\mathbf{R})$, donc, en vertu de 2.1, il existe $q \in G$ tel que $g|_P = s_q$, donc $g = S_{q,q}$.
2) Si on a $g(1) = r$, on a $N(r) = N(1) = 1$, donc $r \in G$. On a alors :

$$S_{\overline{r},1} \circ g(1) = S_{\overline{r},1}(r) = \overline{r} r 1 = 1$$

(remarquer que $S_{\overline{r},1} = T_{\overline{r}}$). Donc, d'après le premier cas, il existe $q \in G$ tel que $S_{\overline{r},1} \circ g = S_{q,q}$, i.e. $g = S_{rq,q}$. On a donc démontré le :

Théorème 3.1.
On a un isomorphisme :

$$\overline{S} : (G \times G)/\{(1,1);\ (-1,-1)\} \xrightarrow{\sim} O^+(N) \xrightarrow{\sim} O^+(4,\mathbf{R}).$$

On s'intéresse maintenant au ***groupe projectif*** $PO^+(4,\mathbf{R}) = O^+(4,\mathbf{R})/\{\mathrm{Id}, -\mathrm{Id}\}$. Pour ceci, on cherche les couples $(q_1,q_2) \in G \times G$ tels que $S_{q_1,q_2} = -\mathrm{Id}$, c'est-à-dire, tels que l'on ait, pour tout $q \in H$, $q_1 q \overline{q_2} = -q$.
Faisant $q = 1$, on trouve $q_1 = -q_2$, puis on voit que q_1 est central d'où $q_1 = \mp 1$. Les seuls couples convenables sont donc $(1,-1)$ et $(-1,1)$. Soit alors V le groupe formé par les 4 éléments $(1,1)$, $(-1,-1)$, $(1,-1)$ et $(-1,1)$, le théorème d'isomorphisme montre que $\overline{S}$ induit un isomorphisme :

$$\widehat{S} : (G \times G)/V \xrightarrow{\sim} PO^+(4,\mathbf{R}).$$

Mais, on a aussi un isomorphisme :

$$\varphi : (G \times G)/V \xrightarrow{\sim} G/\{1,-1\} \times G/\{1,-1\}$$

obtenu à partir de l'homomorphisme $(q_1,q_2) \mapsto (\pi(q_1),\pi(q_2))$ où π désigne la projection canonique de G sur $G/\{1,-1\}$.
En comparant maintenant avec 2.1, on obtient le théorème suivant :

Théorème 3.2.
On a un isomorphisme :

$$PO^+(4,\mathbf{R}) \simeq O^+(3,\mathbf{R}) \times O^+(3,\mathbf{R}).$$

On notera, en particulier, que le groupe $PO^+(4,\mathbf{R})$ n'est pas simple, contrairement à tous les autres groupes $PO^+(n,\mathbf{R})$ pour $n \geq 3$ (cf. Chapitre VI, 7.1, voir aussi les exercices 2, 3, 4 ci-dessous).

4. Quelques compléments sur **H**.

a) Relation avec $SU(2,\mathbf{C})$.

Le corps $\mathbf{C}$ se plonge comme sous-corps de $\mathbf{H}$, par exemple comme l'ensemble des $a+bi$, ([1]) avec $a,b \in \mathbf{R}$ et $\mathbf{H}$ est alors un $\mathbf{C}$-espace vectoriel pour la loi extérieure $(\lambda,q) \mapsto q\lambda$ pour $q \in \mathbf{H}$, et $\lambda \in \mathbf{C}$ (attention au sens). Une base de $\mathbf{H}$ sur $\mathbf{C}$ est alors $1, j$ et si q est un quaternion, $q = a+bi+cj+dk$, on a $q = 1(a+bi)+j(c-di) = 1\lambda + j\mu$, avec $\lambda,\mu \in \mathbf{C}$.

On fait alors opérer G sur $\mathbf{H}$ par multiplication à gauche : $T_q(q') = qq'$. On a vu que T_q est $\mathbf{R}$-linéaire, mais elle est aussi $\mathbf{C}$-linéaire (à cause du sens d'écriture), et inversible, donc T_q est dans $GL(2,\mathbf{C})$ et on a un homomorphisme $T : G \longrightarrow GL(2,\mathbf{C})$, injectif.
La matrice de T_q sur la base $1, j$, pour $q = 1\lambda + j\mu$ se calcule aisément :

$$\text{on a} \quad T_q = \begin{pmatrix} \lambda & -\overline{\mu} \\ \mu & \overline{\lambda} \end{pmatrix} \quad \text{avec } \lambda\overline{\lambda} + \mu\overline{\mu} = 1, \text{ (puisque } q \in G).$$

On a les formules ${}^t\overline{T}_q = T_q^{-1}$ et $\det T_q = 1$, de sorte que T_q appartient à $SU(2,\mathbf{C})$, groupe spécial unitaire relatif à la forme hermitienne standard $\lambda\overline{\lambda} + \mu\overline{\mu}$ sur $\mathbf{C}^2$. De plus, on obtient ainsi toutes les matrices de $SU(2,\mathbf{C})$ et on a donc le théorème suivant :

Théorème 4.1.
On a un isomorphisme :

$$T : G \xrightarrow{\sim} SU(2,\mathbf{C}).$$

Avec le théorème 2.1, on a aussitôt :

Corollaire 4.2.
On a un isomorphisme $SU(2,\mathbf{C})/\{\mathrm{Id}, -\mathrm{Id}\} \simeq O^+(3,\mathbf{R})$.
En particulier, le groupe $SU(2,\mathbf{C})/\{\mathrm{Id}, -\mathrm{Id}\}$ est simple (cf. Chapitre VI, 6.1).

b) Le théorème de Frobenius.

On peut se demander s'il existe d'autres corps que les quaternions qui soient des $\mathbf{R}$-algèbres de dimension finie, afin d'obtenir, par exemple, les paramétrisations analogues à celles des § 2, 3 en dimension supérieure. Sous cette forme, la réponse est négative :

([1]) Mais aussi comme l'ensemble des $a+bj$ ou des $a+bk$,... On notera que dans le corps (non commutatif) $\mathbf{H}$, l'équation $x^2 = -1$ a beaucoup de solutions !

Théorème 4.3 : (Frobenius).
Tout corps K contenant $\mathbf{R}$ dans son centre et de dimension finie sur $\mathbf{R}$ est isomorphe à $\mathbf{R}$, $\mathbf{C}$, ou $\mathbf{H}$.

Démonstration.

1) Supposons tout d'abord K commutatif et $K \neq \mathbf{R}$. Soit $a \in K - \mathbf{R}$. Comme on a $[K : \mathbf{R}] < +\infty$, a est algébrique sur $\mathbf{R}$ (cf. III, 1.13), et comme son polynôme minimal est irréductible sur $\mathbf{R}$, il est de degré 2 (cf. III, 3.1.4). Le discriminant Δ de ce polynôme, qui est < 0, est donc un carré dans K, mais alors -1 aussi, de sorte que K contient un sous-corps isomorphe à $\mathbf{C}$. Mais K est *a fortiori* algébrique sur ce sous-corps, donc, puisque $\mathbf{C}$ est algébriquement clos, on a $K = \mathbf{C}$ (cf. III, 1.17).

2) Supposons donc K non commutatif. Soit $a \in K - \mathbf{R}$, alors $\mathbf{R}[a]$ est un corps commutatif de dimension finie sur $\mathbf{R}$, donc isomorphe à $\mathbf{C}$. On trouve ainsi dans K un sous-corps isomorphe à $\mathbf{C}$, que l'on désigne par $\mathbf{C}$ et on note i une racine de -1 dans $\mathbf{C}$.
Notons que $\mathbf{C}$ est alors un sous-corps commutatif maximal de K (cf. 1)), en particulier, si x est dans K et commute à i, on a $x \in \mathbf{C}$.
Soit alors $y \in K - \mathbf{C}$, y ne commute pas à i et on construit un élément z qui anticommute à i. ([2]) On prend pour cela $z = yi - iy$, z est non nul et on a bien $iz = -zi$.
Mais on a alors $iz^2 = z^2 i$, donc z^2, qui commute à i, est dans $\mathbf{C}$. Comme on a $[\mathbf{R}(z) : \mathbf{R}] = 2$ (car $\mathbf{R}(z)$ est commutatif) et $\mathbf{R}(z) \neq \mathbf{C}$, $\mathbf{R}(z) \cap \mathbf{C}$ ne peut être que $\mathbf{R}$, donc z^2 est dans $\mathbf{R}$.
Comme $\mathbf{R}(z)$ est un corps commutatif, z^2 est même dans $\mathbf{R}^-$, sinon, si $z^2 = a > 0$, cette équation aurait au moins 4 racines dans $\mathbf{R}(z)$: $z, -z, \sqrt{a}, -\sqrt{a}$, or c'est interdit car $\mathbf{R}(z)$ est commutatif. On a donc $z^2 = -\alpha$, avec $\alpha \in \mathbf{R}^+$. Posons $j = z/\sqrt{\alpha}$, on a toujours $ij = -ji$, mais aussi $j^2 = -1$. Si on pose $k = ij$, le sous-espace de K engendré par $1, i, j, k$ est un sous-corps de K, isomorphe à $\mathbf{H}$. On l'appelle donc $\mathbf{H}$, on a $\mathbf{H} \subset K$ et il reste à voir qu'on a égalité.

3) Supposons $K \neq \mathbf{H}$, soit $u \in K - \mathbf{H}$, on réitère le procédé du 2). On pose $v = ui - iu$, v anticommute à i, et si $l = v/\sqrt{-v^2}$, on a $li = -il$ et $l^2 = -1$. Considérons alors jl, on a $jli = ijl$, donc $jl \in \mathbf{C}$, et on en déduit que l et v sont dans $\mathbf{H}$. Enfin, soit $w = ui + iu$, w commute à i, donc w est dans $\mathbf{C} \subset \mathbf{H}$, mais on a $ui = \dfrac{v + w}{2}$ donc $ui \in \mathbf{H}$, et $u \in \mathbf{H}$, contradiction.

c) Les octaves de Cayley.

Il existe en dimension 8, une « algèbre », non associative, qui est presque un corps, c'est **Ca**, algèbre des octaves de Cayley ou octonions. En voici une description sommaire :
On prend $\mathbf{Ca} = \mathbf{H} \times \mathbf{H}$, avec sa structure naturelle de $\mathbf{R}$-espace vectoriel, et la multiplication définie par

$$(q_1, q_2)\ (q_1', q_2') = (q_1 q_1' - \overline{q_2'} q_2\ ,\ q_2 \overline{q_1'} + q_2' q_1).$$

([2]) Pour expliquer un peu les choses, on cherche à reconstituer le corps $\mathbf{H}$ dans K, en particulier, on cherche un j.

L'algèbre **Ca** n'est ni commutative, ni associative, on a toutefois, la propriété plus faible suivante, pour $r_1, r_2 \in \mathbf{Ca}$, on a :

$$r_1^2 r_2 = r_1(r_1 r_2) \quad \text{et} \quad r_1 r_2^2 = (r_1 r_2) r_2.$$

On définit un conjugué $\overline{r}$ de r par : $\overline{(q_1, q_2)} = (\overline{q_1}, -\overline{q_2})$ et une norme par $N(r) = r\overline{r} = \overline{r}r$. On vérifie que, pour $r = (q_1, q_2)$, on a $N(r) = N(q_1) + N(q_2) \geq 0$. On a encore $N(r_1 r_2) = N(r_1)N(r_2)$ et, si on a $r \in \mathbf{Ca}$, $r \neq 0$, r a un inverse $r^{-1} = \overline{r}/N(r)$ (cf. [Bbki] Algèbre 1,2,3 p. 176 ou [AH] p. 107).

d) Les algèbres de Clifford.

On peut tout de même, pour étudier une forme quadratique q, construire une algèbre $C(q)$, algèbre de Clifford de q, qui rend à peu près les mêmes services que **H** dans l'étude de $O(q)$, mais cette algèbre n'est pas intègre. (cf. [D] Ch.II. §7,8,9 ou [A] Ch.V. §4).

e) Un peu de topologie.

Le groupe G, dont on a vu qu'il est isomorphe à $SU(2, \mathbf{C})$, est un groupe topologique (et même un groupe de Lie) homéomorphe à la sphère $\mathbf{S}^3$. C'est un phénomène exceptionnel, les seules sphères munies de telles structures sont $\mathbf{S}^0 = \{1, -1\}$, $\mathbf{S}^1 \simeq \mathbf{U}$ et $\mathbf{S}^3 \simeq G$, ces structures étant liées à la structure des corps **R**, **C**, **H** (cependant la sphère $\mathbf{S}^7$ a aussi une structure, plus faible, liée à **Ca**). Ce résultat est dû à Adams, cf. [K] Ch. V §1 Th.1.6).

Par ailleurs, l'isomorphisme $\overline{s}$ du théorème 2.1 qui est un isomorphismes de groupes topologiques renseigne aussi sur la structure topologique de $O^+(3, \mathbf{R})$. En effet, $O^+(3, \mathbf{R})$ apparaît comme le quotient de $\mathbf{S}^3$ par l'antipodie, relation qui identifie les points x et $-x$. Il est donc homéomorphe à l'espace projectif $\mathbf{P}^3(\mathbf{R})$. Le groupe $O^+(3, \mathbf{R})$ est compact, connexe, car $\mathbf{S}^3$ l'est, mais non simplement connexe, puisque G en est un revêtement non trivial de degré 2. Comme $G \simeq \mathbf{S}^3$ est simplement connexe, c'est même le revêtement universel de $O^+(3, \mathbf{R})$, et le groupe fondamental de $O^+(3, \mathbf{R})$ est $\mathbf{Z}/2\mathbf{Z}$, voir exercices 1,2.
(Pour des précisions sur ces questions, cf. [B] Ch. VIII §9,10 ou [G] Ch. VI §6 et Ch. X. §6).

5. Les quaternions généralisés.

Dans tout ce qui suit, K désigne un corps commutatif de caractéristique différente de 2.

Théorème et définition 5.1.
Soient $\alpha, \beta \in K^*$. Il existe une K-algèbre $H_{\alpha,\beta}$, de dimension 4, appelée **algèbre des quaternions généralisés** *relatifs à α, β (et K), munie d'une base $1, i, j, k$ telle que :*
1) 1 est élément neutre pour la multiplication,
2) on a $i^2 = \alpha$, $j^2 = \beta$, $ij = -ji = k$.
On identifie K avec le sous-corps des éléments $a.1$, pour $a \in K$.

Remarque 5.2. On déduit aussitôt par associativité les formules :

$$k^2 = -\alpha\beta, \quad jk = -kj = -\beta i, \quad ki = -ik = -\alpha j.$$

Démonstration (de 5.1) On définit $H_{\alpha,\beta}$ comme ci-dessus, et on vérifie l'associativité (cf. Exercice 1 pour une démonstration matricielle).

On vérifie aisément que le centre de $H_{\alpha,\beta}$ est K.

On définit ensuite, comme dans $\mathbf{H}$, le conjugué $\overline{q}$ du quaternion $q = a+bi+cj+dk$ par $\overline{q} = a-bi-cj-dk$, et l'application $q \mapsto \overline{q}$ est un antiautomorphisme de $H_{\alpha,\beta}$.

On pose enfin :

$$N(q) = q\overline{q} = \overline{q}q = a^2 - \alpha b^2 - \beta c^2 + \alpha\beta d^2.$$

On a $N(q) \in K$ et N est encore une forme quadratique, mais, à la différence de ce qui se passe sur $\mathbf{H}$, elle peut être d'indice $\nu \geq 1$ i.e. avoir des isotropes. Précisément, on a le théorème suivant :

Théorème 5.3.
L'algèbre $H_{\alpha,\beta}$ est un corps si et seulement si la forme N est anisotrope.

Démonstration.

1) Supposons N anisotrope et soit $q \in H_{\alpha,\beta}$, $q \neq 0$. On a alors $N(q) \neq 0$, donc q admet pour inverse $q^{-1} = \dfrac{1}{N(q)}\overline{q}$.

2) Si q est un vecteur isotrope non nul, on a $q\overline{q} = 0$ avec $q \neq 0$, $\overline{q} \neq 0$, de sorte que $H_{\alpha,\beta}$ n'est pas intègre.

Corollaire 5.4.
Pour $K = \mathbf{F}_q$, l'algèbre $H_{\alpha,\beta}$ n'est jamais un corps.

Ce résultat est un corollaire, soit du théorème de Wedderburn (Chapitre III, 4.9) soit, plus simplement, du fait que si N est une forme quadratique sur $\mathbf{F}_q$ en au moins 3 variables, N admet des vecteurs isotropes non triviaux (V, 6.10, ou V, § 6, Exercice 6).

Lorsque N est d'indice $\nu \geq 1$, on a le théorème suivant, pour lequel nous renvoyons à [Bl] Th. 1.5.

Théorème 5.5.
Si la forme N est d'indice ≥ 1, l'algèbre $H_{\alpha,\beta}$ est isomorphe à l'algèbre des matrices $\mathbf{M}(2, K)$.

Exemple 5.6. On prend $\alpha = \beta = 1$, on a donc $-\alpha\beta = -1$. On obtient un isomorphisme de $H_{\alpha,\beta}$ sur $M(2,K)$ en prenant :

$$i = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad j = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad k = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Si $q = a + bi + cj + dk$, la matrice correspondante est :

$$\begin{pmatrix} a+b & c+d \\ c-d & a-b \end{pmatrix}.$$

Revenons maintenant au cas général, on pose :

$$H^*_{\alpha,\beta} = H^* = \{q \in H_{\alpha,\beta} \mid N(q) \neq 0\}.$$

On a donc $H^* = H_{\alpha,\beta} - \{0\}$ si N est anisotrope, et sinon $H^* \simeq GL(2,K)$, (cf. 5.5). Si q est dans H^*, q est inversible et $q^{-1} = \dfrac{1}{N(q)}\overline{q}$.

On fait alors, comme au §2, opérer H^* sur $H = H_{\alpha,\beta}$ par automorphismes intérieurs. On pose pour $q \in H^*$, $q' \in H$:

$$S_q(q') = qq'q^{-1}.$$

Un certain nombre de propriétés subsistent : S_q conserve N, $S_q|_K = \mathrm{Id}_K$ et $S_q(P) = P$ où $P = \{bi + cj + dk \mid b, c, d \in K\}$ est l'espace des quaternions purs.
Si on pose $s_q = S_q|_P$ on a donc un homomorphisme $s : H^* \longrightarrow O(N|_P)$, et le noyau de s est $K \cap H^* = K^*$.
On va montrer ensuite que s est à valeurs dans $O^+(N|_P)$ en utilisant la technique évoquée en 2.2.2.
Soit $p \in P$ avec $N(p) \neq 0$, la réflexion τ_p associée à p est donnée par la formule de V, 4.7.3 :

$$\forall x \in P, \quad \tau_p(x) = p\overline{x}p^{-1} = -pxp^{-1}.$$

Notons déjà que si σ_p est le renversement d'axe p, on a $\sigma_p(x) = -\tau_p(x) = s_p(x)$.
Soit $u \in O_3(K, N|_P)$, nous verrons au Chapitre VIII, théorème 5.7, que u est produit de réflexions, $u = \tau_{p_1} \dots \tau_{p_r}$ pour $p_1, \dots, p_r \in P$, avec $N(p_i) \neq 0$, l'entier r étant pair si u est dans O^+, impair sinon. On a donc, en admettant VIII, 5.7 :

$$u(x) = (-1)^r p_1 \dots p_r \, x \, (p_1 \dots p_r)^{-1} \text{ pour tout } x \in P.$$

On peut maintenant revenir à la question posée : soit $q \in H^*$ et supposons $s_q \in O^-(N|_P)$, on aurait alors pour tout $x \in P$:

$$s_q(x) = qxq^{-1} = -q'xq'^{-1}, \text{ en posant } q' = p_1 \dots p_r.$$

Mais alors, si on pose $u = q^{-1}q'$, on a, pour tout $x \in P$, $-x = uxu^{-1}$, ou encore, $\overline{x} = uxu^{-1}$.
Comme cette dernière formule vaut aussi pout tout $x \in K$, on aurait donc pour tout x de H : $\overline{x} = uxu^{-1}$, avec $u \in H^*$, mais $x \mapsto \overline{x}$ est un antiautomorphisme et $x \mapsto uxu^{-1}$ un automorphisme, c'est absurde.

En définitive, on a un homomorphisme $s : H^* \mapsto O_3^+(K, N|_P)$ et s est surjectif car les renversements sont dans l'image comme on l'a noté ci-dessus et ils engendrent O^+ (cf. VIII, 5.9) : on obtient le théorème suivant :

Théorème 5.7.
On a un isomorphisme $\overline{s} : H^*/K^* \xrightarrow{\sim} O_3^+(K, N|_P)$.

Une application.
On reprend les notations de l'exemple 5.6 : $\alpha = \beta = 1$, et on note encore N la restriction de N à P : on a donc $N(bi + cj + dk) = -b^2 - c^2 + d^2$ et N est une forme d'indice 1 (de type Lorentz, au signe près, cf. VIII, 4.10). On désigne par $\Omega_3(K, N)$ le groupe des commutateurs de $O_3^+(K, N)$. On a alors le résultat suivant, qui découle aussitôt du théorème 5.7 et du calcul des commutateurs du groupe linéaire, cf. IV, 3.1 :

Corollaire 5.8.
1) Avec les notations précédentes, on a un isomorphisme :

$$GL(2, K)/K^* = PGL(2, K) \simeq O_3^+(K, N).$$

2) On a un isomorphisme :

$$PSL(2, K) \simeq \Omega_3(K, N).$$

En particulier, on a : $O_3^+/\Omega_3 \simeq K^*/K^{*2}$ et ce groupe est non trivial en général (comparer au cas euclidien, cf. VI, 3.4) et $\Omega_3(K,N)$ est simple, sauf si $K = \mathbf{F}_3$ (cf. IV, 4.1).

Le théorème 5.5 ci-dessus, montre que cet isomorphisme vaut pour tous les α, β tels que $N|_P$ soit d'indice $\nu \geq 1$, ce que nous retrouverons aisément au Chapitre VIII en montrant que pour un corps k donné, en dimension 3, les formes N d'indice ≥ 1 sont toutes équivalentes à un scalaire près, donc ont des groupes orthogonaux isomorphes, cf. VIII, 4.10.

Sur ce paragraphe, on pourra consulter [D] Ch. II § 9, [V] Ch. I § 3 et [Bl] Ch. I § 3, ainsi que l' exercice 3 ci-dessous pour l'étude de la dimension 4.

EXERCICES SUR LE CHAPITRE VII

1. Définition du corps H.

1) Soit K un corps, A un K-espace vectoriel muni d'une base $(e_i)_{i\in I}$. Soit $\nu : A \times A \longrightarrow A$ une loi de composition sur A, supposée bilinéaire.

a) Montrer que, pour prouver l'associativité ou la commutativité de ν, il suffit de le faire sur la base i.e. de montrer les formules du type :

$$\nu(e_i, \nu(e_j, e_k)) = \nu(\nu(e_i, e_j), e_k), \quad \text{pour tous } i, j, k \in I.$$

b) Soit $\varphi : A \longrightarrow A$ une application K-linéaire. Montrer que φ est un homomorphisme relativement à ν si et seulement si on a, pour tous $i, j \in I$:

$$\varphi(\nu(e_i, e_j)) = \nu(\varphi(e_i), \varphi(e_j)).$$

---

2) Soit $\mathbf{H}$ défini comme au § 1. On se propose de prouver l'associativité de la multiplication sur $\mathbf{H}$.

a) Soit $\sigma : \mathbf{H} \longrightarrow \mathbf{H}$ l'application $\mathbf{R}$-linéaire définie par : $\sigma(1) = 1$, $\sigma(i) = j$, $\sigma(j) = k$, $\sigma(k) = i$. Montrer que σ est un automorphisme de $\mathbf{H}$ (utiliser 1)).

b) Même question pour τ, défini par : $\tau(1) = 1$, $\tau(i) = -i$, $\tau(j) = k$, $\tau(k) = j$.

c) Montrer que la vérification de l'associativité de $\mathbf{H}$ se ramène à 5 cas (au lieu de 27), conclure.

---

3) On prend la définition de $\mathbf{H}$ comme sous-algèbre de $\mathbf{M}(4, \mathbf{R})$ (cf. § 1).

a) vérifier les formules $i^2 = -1$, $ij = k \ldots$

b) Montrer que, si q est non nul, $M(q)/\sqrt{N(q)}$ est dans $O(4, \mathbf{R})$.

c) En déduire l'expression de $M(q)^{-1}$, pour $q \neq 0$, puis la valeur de dét $M(q)$.

---

4) Soit L le sous-ensemble de $\mathbf{M}(2, \mathbf{C})$ défini par :

$$L = \left\{ M(a,b) = \begin{pmatrix} a & -\overline{b} \\ b & \overline{a} \end{pmatrix}, \text{ avec } a, b \in \mathbf{C} \right\}.$$

a) Montrer que L est un sous-$\mathbf{R}$-espace vectoriel de $\mathbf{M}(2, \mathbf{C})$, de dimension 4, dont une base est formée des matrices $M(1,0)$, $M(i,0)$, $M(0,1)$, $M(0,i)$.

b) Montrer que L est une sous-$\mathbf{R}$-algèbre de $\mathrm{M}(2,\mathbf{C})$ et donner la table de L dans la base précédente. En déduire que L est isomorphe à $\mathbf{H}$.

c) Si $q = \alpha+\beta i+\gamma j+\delta k$ est un quaternion, on pose $R_q = M(\alpha+\beta i,\ \gamma-\delta i)$. Montrer qu'on a $R_{\overline{q}} = {}^t\overline{R_q}$ et $N(q) = \det R_q$. Montrer que q est un quaternion pur si et seulement si la trace de R_q est nulle.

---

5) On identifie $\mathbf{H}$ au produit $\mathbf{R}\times P$, (P espace des quaternions purs) et on munit P du produit scalaire $(v\,|\,w)$ et du produit vectoriel $v\wedge w$ associés à la base orthonormée i,j,k.
Montrer que, pour $x,y\in\mathbf{R}$ et $v,w\in P$, la multiplication sur $\mathbf{H}$ s'écrit alors :

$$(x,v)(y,w) = (xy-(v\,|\,w),\ xw+yv+v\wedge w).$$

---

6) Soit $\Sigma = \{n\in\mathbf{N}\,|\,\exists\ x,\ y,\ z,\ t\in\mathbf{N},\ n = x^2+y^2+z^2+t^2\}$ (les sommes de quatre carrés).
Montrer, en utilisant les quaternions, que Σ est stable par multiplication. En fait, on a $\Sigma=\mathbf{N}$, cf. par exemple [Bl] Th. 1.4.

---

2. Opération de $\mathbf{H}$ sur $\mathbf{R}^3$.

1) On reprend les notations du théorème 2.1. Soit $q\in G$, on écrit $q = a+p$, avec $a\in\mathbf{R}$, $p\in P$. On a alors $N(q) = a^2+N(p) = a^2+\|p\|^2 = 1$. On pose $a=\cos\theta$, $\|p\| = \sin\theta$, pour $\theta\in[0,\pi]$.

a) Montrer que s_q est une rotation de P, d'axe $<p>$.

b) On suppose qu'on a $p = bi$, avec $b\in\mathbf{R}$, $b\geq 0$. Montrer que s_q est la rotation d'axe $<p>$ et d'angle 2θ.

c) Dans le cas général, montrer qu'il existe $g\in G$ tel que $s_g(q) = a+bi$, avec $b\in\mathbf{R}$, $b\geq 0$. En déduire que la conclusion de b) reste valable.

d) Déduire de ce qui précède une autre démonstration de la surjectivité de s.

---

3. La structure de $O^+(4,\mathbf{R})$.

1) Montrer que les images respectives des homomorphismes $q\mapsto T_q$ et $q\mapsto S_q$ de G dans $O^+(N)$ correspondent dans $G\times G$ au premier facteur et à la diagonale.

---

2) Soit $G = G_1\times G_2$ un groupe produit de deux groupes simples non abéliens .

a) Montrer que si N est un sous-groupe distingué de G et si $(x,y)\in N$ avec $x\neq 1$, il existe $z\in G_1$, $z\neq 1$ tel que $(z,1)\in N$.

b) Montrer que les seuls sous-groupes distingués de G sont $\{1\}$, $G_1\times\{1\}$, $\{1\}\times G_2$ et G.

c) Déterminer les sous-groupes distingués de $PO^+(4,\mathbf{R})$.

---

3) Soit H un sous-groupe distingué de $O^+(4,\mathbf{R})$, non contenu dans le centre.

a) Déterminer l'image $\overline{H}$ de H dans $PO^+(4,\mathbf{R})$.

b) Montrer qu'on a $-\mathrm{Id} \in H$ (utiliser un élément $(i,1)$ ou $(1,i)$ de $G \times G$).

c) Déterminer tous les sous-groupes distingués de $O^+(4,\mathbf{R})$.

---

4) Soient $q, r \in G$, T_r la multiplication à gauche par r dans H : $T_r(x) = rx$ et soit τ_q la réflexion de droite $<q>$. On rappelle qu'on a $\tau_q(x) = -q\overline{x}q$ (cf. 2.2.2).

a) Montrer qu'on a $\tau_q T_r \tau_q(x) = q\overline{q}x\overline{q}\overline{r}q = x\overline{q}\overline{r}q$.

b) En déduire que les sous-groupes facteurs du produit

$$PO^+(4,\mathbf{R}) \simeq O^+(3,\mathbf{R}) \times O^+(3,\mathbf{R})$$

sont échangés par conjugaison par les isométries négatives.
Montrer que $PO^+(4,\mathbf{R})$ est le seul sous-groupe distingué non trivial de $PO(4,\mathbf{R})$ (cf. VI, 7.3.2).

---

5) Soit φ l'automorphisme de $G \times G$ défini par $\varphi(x,y) = (y,x)$, $\overline{\varphi}$ l'automorphisme de $O^+(4,\mathbf{R})$ induit par φ et $\widehat{\varphi}$ l'automorphisme de $PO^+(4,\mathbf{R})$ induit par $\overline{\varphi}$.

a) Montrer que $\widehat{\varphi}$ échange les deux facteurs isomorphes à $O^+(3,\mathbf{R})$ et en déduire que $\widehat{\varphi}$ n'est pas intérieur.

b) Montrer que $\overline{\varphi}$ n'est pas intérieur (cf. VI, 8.1 ou VI, § 8, Exercice 1).

---

4. Compléments sur H.

Pour les exercices de topologie qui suivent, on pourra par exemple se référer à [G] Ch. V, Ch. VI § 4 et Ch. VII.

1) Soit G le groupe des quaternions de norme 1. On pose

$$U_1 = G - \{k\}, \quad U_2 = G - \{-k\}.$$

a) Montrer que l'on a $G = U_1 \cup U_2$, que U_1 et U_2 sont homéomorphes à $\mathbf{R}^3$ et que $U_1 \cap U_2$ est connexe.

b) Soit $\gamma : [0,1] \longrightarrow G$ un lacet continu. Montrer qu'il existe des points $t_0, \ldots, t_n \in [0,1]$ tels que $t_0 = 0 < t_1 < \ldots < t_n = 1$ et tels que, si $\gamma = \gamma|_{[t_i, t_{i+1}]}$, on ait pour tout i :

1) $\mathrm{Im}\ \gamma_i \subset U_1$ ou $\mathrm{Im}\ \gamma_i \subset U_2$,

2) $\gamma_i(t_i) \in U_1 \cap U_2$.

c) Montrer qu'il existe, pour tout i, un chemin δ_i contenu dans $U_1 \cap U_2$ et joignant $\gamma_i(0)$ et $\gamma_i(t_i)$.

d) Si γ et δ sont des chemins, on désigne par $\overline{\gamma}$ le chemin γ parcouru à l'envers et, si l'origine de δ est égale à l'extrémité de γ, par $\gamma \vee \delta$ le chemin obtenu en parcourant γ puis δ.
Montrer que les lacets $\gamma_1 \vee \delta_1$, $\delta_1 \vee \gamma_2 \vee \delta_2 \ldots$ sont homotopes à zéro. En déduire que γ est homotope à zéro, donc que G est simplement connexe.

---

2) Soit $p : G \longrightarrow G/\{1,-1\} = \overline{G}$ la projection canonique.

a) Soit $\gamma : [0,1] \longrightarrow \overline{G}$ un lacet continu tel que $\gamma(0) = \gamma(1) = \overline{1}$. Montrer qu'il existe $\delta : [0,1] \longrightarrow G$, unique, tel que δ soit continu, qu'on ait $\delta(0) = 1$ et pour

tout $t \in [0,1]$, $p\delta(t) = \gamma(t)$ (on recouvrira $\gamma([0,1])$ par des ouverts suffisamment petits).
(Il s'agit du lemme de *relèvement des chemins* dans un revêtement, attention, δ n'est plus un lacet, *a priori*).

b) *Relèvement des homotopies.*
Soit $H : [0,1]^2 \longrightarrow \overline{G}$ une application continue telle que, pour tout $u \in [0,1]$, on ait $H(0,u) = H(1,u) = \overline{1}$. Montrer qu'il existe une application continue $K : [0,1]^2 \longrightarrow G$ unique, telle que $K(0,u) = 1$ pour tout u et $p(K(t,u)) = H(t,u)$ pour tous $t, u \in [0,1]$.

c) Soit γ un lacet de $\overline{G}$, d'origine $\overline{1}$, δ son relèvement dans G (cf. a)). Montrer que si on a $\delta(1) = \delta(0) = 1$, γ est homotope à 0.
Montrer que si on a $\delta(1) = -1$, γ n'est pas homotope à 0 (cf. b)).

d) Montrer que si γ, γ' ne sont pas homotopes à 0, ils sont homotopes entre eux et que $\gamma \vee \gamma'$ est homotope à 0.

e) En déduire que $O^+(3, \mathbf{R})$ n'est pas simplement connexe et que son groupe fondamental est isomorphe à $\mathbf{Z}/2\mathbf{Z}$.

f) Montrer que le lacet $t \mapsto \rho(k, 2\pi t)$, pour $t \in [0,1]$ (rotations d'axe $< k >$, d'angle $2\pi t$) n'est pas homotope à 0 dans $O^+(3, \mathbf{R})$.

---

5. Les quaternions généralisés.

1) Soit K un corps de caractéristique différente de 2, soient $\alpha, \beta \in K$ et L une extension de K dans laquelle α est un carré, $\alpha = a^2$, avec $a \in L$. On considère la sous-K-algèbre A de $\mathbf{M}(2, L)$ engendrée par les matrices :

$$i = \begin{pmatrix} a & 0 \\ 0 & -a \end{pmatrix} \quad \text{et} \quad j = \begin{pmatrix} 0 & \beta \\ 1 & 0 \end{pmatrix}.$$

a) Montrer que A est de dimension 4 sur K, de base $1, i, j, k$ avec $k = ij$.

b) Montrer que A est isomorphe à $H_{\alpha,\beta}$. Étudier le cas où a est dans K.

---

2) Soit K un corps, de caractéristique différente de 2 et soient $a, b \in K^*$.

a) Montrer que si $-ab$ est un carré la forme quadratique $ax^2 + by^2$ possède des vecteurs isotropes.

b) Montrer que si q est une forme quadratique sur un espace E de dimension 4 et si on a $\nu(q) = 1$, le discriminant $\Delta(q)$ n'est pas un carré (utiliser a) et VIII, 2.4).

c) Déduire de b) que, pour $\alpha, \beta \in K$, la forme $x^2 - \alpha y^2 - \beta z^2 + \alpha\beta t^2$ est d'indice 0 ou 2 (i.e. anisotrope ou hyperbolique, cf. Chapitre VIII § 3 et 4).

---

3) Soit K un corps, de caractéristique différente de 2, soient $\alpha, \beta \in K^*$ et $H = H_{\alpha,\beta}$.

a) Soit $q \in H^*$ et soit τ_q la réflexion orthogonale (relativement à N), de droite $< q >$. Montrer que, pour tout $x \in H$ on a :

$$\tau_q(X) = -q\overline{x}q/N(q) = -\overline{q}^{-1}\overline{x}q.$$

b) Soit $u \in O_4^-(N)$, montrer qu'il existe $a, b \in H^*$ tels que, pour tout $x \in H$ on ait $u(x) = a\overline{x}b$ (cf. VIII, 5.7).

c) Soit $u \in O_4^+(N)$, montrer qu'il existe $a, b \in H^*$ tels que, pour tout $x \in H$ on ait $u(x) = axb$. Montrer qu'on a alors $N(a)N(b) = 1$.

d) Soit $U = \{(a,b) \in H^* \times H^* \mid N(a) = N(b)\}$. On pose pour $(a,b) \in U$ et $q \in H$: $S_{a,b}(q) = aqb^{-1}$.
Montrer que S est un homomorphisme de U dans $O_4(N)$ et calculer $\operatorname{Ker} S$.

e) Montrer qu'on a $S(U) \subset O_4^+(N)$ (cf. b)) puis $S(U) = O_4^+(N)$ (cf. c)). En déduire qu'on a un isomorphisme :

$$\overline{S} : U/\operatorname{Ker} S \simeq O_4^+(N).$$

f) On suppose $\alpha = \beta = 1$. On a alors (cf. VII, 5.6) $H \simeq \mathbf{M}(2,K)$, $H^* \simeq GL(2,K)$, et N, qui correspond au déterminant, est une forme hyperbolique (cf. Exercice 2).
Montrer alors les points suivants :

1) L'ensemble U correspond à $\{(u,v) \in GL(2,K) \times GL(2,K) \mid \text{dét } u = \text{dét } v\}$.

2) On a $\operatorname{Ker} S \simeq \{(\lambda,\lambda) \mid \lambda \in K^*\}$.

3) Le groupe dérivé $D(U)$, est isomorphe à $SL(2,K) \times SL(2,K)$.

4) Le groupe $\Omega_4(N) = D(O_4^+(N))$ est isomorphe au quotient du précèdent par $\{(\mathrm{Id},\mathrm{Id}), (-\mathrm{Id},-\mathrm{Id})\}$.

5) On a un isomorphisme $P\Omega_4(N) \simeq PSL(2,K) \times PSL(2,K)$, cf. VIII, §8 et VIII, 9.3.

VIII. LE GROUPE ORTHOGONAL,

CAS GÉNÉRAL

1. Introduction.

Dans tout ce chapitre k désigne un corps commutatif de caractéristique différente de 2, E un k-espace vectoriel de dimension finie n, q une forme quadratique ***non dégénérée*** sur E, f sa forme polaire. Sauf mention expresse du contraire, ces notations valent pour l'ensemble du chapitre.
Hormis la restriction sur la caractéristique, nous étudions donc ici le cas de la forme quadratique la plus générale. On notera en particulier, comparant au Chapitre VI, que :

1) le corps k est quelconque (et plus nécessairement $\mathbf{R}$),
2) la forme q est quelconque (donc plus nécessairement anisotrope).

La plupart des résultats de ce chapitre sont valables dans ce cadre général. On aura cependant parfois besoin de l'hypothèse $\nu(q) \geq 1$ (i.e. l'existence de vecteurs isotropes). Les vecteurs et sous-espaces isotropes sont d'ailleurs au centre des techniques utilisées dans ce chapitre.
Nous renvoyons au chapitre V pour toutes définitions et propriétés élémentaires concernant l'orthogonalité, l'isotropie, les groupes $O(q)$ et $O^+(q)\ldots$

2. Plans hyperboliques.

Définition 2.1.
Soit P un plan (i.e. un k-espace vectoriel de dimension 2), q une forme quadratique sur P de forme polaire f. On dit que (P, q) est un ***plan hyperbolique*** *s'il existe une base e_1, e_2 de P telle que l'on ait :*

$$q(e_1) = q(e_2) = 0 \quad \text{et} \quad f(e_1, e_2) = 1.$$

La base e_1, e_2 est dite ***hyperbolique*** *et on dit aussi que q est une forme hyperbolique.*

Remarques 2.2.

1) La matrice de q dans la base e_1, e_2 est $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. En particulier le discriminant $\Delta(q)$ vaut -1 et q n'est pas dégénérée. Pour $v \in E$, écrit dans cette base $v = xe_1 + ye_2 \in E$, on a $q(v) = 2xy$.

2) Si on a deux vecteurs e_1, e_2, avec $q(e_1) = q(e_2) = 0$ et $f(e_1, e_2) \neq 0$, ces vecteurs sont indépendants et le plan engendré est hyperbolique (remplacer e_2 par $e_2/f(e_1, e_2)$).

3) Si e_1, e_2 est une base hyperbolique de P, les droites $< e_1 >$ et $< e_2 >$ sont isotropes et non orthogonales. De plus, ce sont les seules droites isotropes de P, comme le montre la formule : $q(xe_1 + ye_2) = 2xy$. Cette remarque, jointe à la précédente, montre que dans un plan hyperbolique tout vecteur isotrope non nul peut toujours être pris comme premier vecteur d'une base hyperbolique.

4) Si e_1, e_2 est une base hyperbolique, et si on pose $\varepsilon_1 = e_1 + \frac{e_2}{2}$, $\varepsilon_2 = e_1 - \frac{e_2}{2}$, $\varepsilon_1, \varepsilon_2$ est une base de P dans laquelle q a pour matrice $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Si $v \in E$ est écrit dans cette base $v = x\varepsilon_1 + y\varepsilon_2$, on a $q(v) = x^2 - y^2$.

5) Lorsque k est le corps des nombres réels, l'ensemble $H = \{v \in P \mid q(v) = 1\}$ est une hyperbole, ce qui explique la terminologie.

La proposition suivante caractérise les plans hyperboliques :

Proposition 2.3.
Soit P un plan muni d'une forme quadratique q. Les conditions suivantes sont équivalentes :
1) (P, q) est un plan hyberbolique,
2) q est non dégénérée et d'indice $\nu(q) \geq 1$,
3) le discriminant de q est égal à -1 à un carré près.

Démonstration. Il est clair que 1) implique 2) et 3). Si on a 2), soit x un vecteur isotrope non nul, comme on a $\operatorname{Ker} f = 0$ il existe $z \in E$ tel que $f(x, z) \neq 0$. Posons $y = z + \lambda x$, avec $\lambda \in k$, on a encore $f(x, y) = f(x, z) \neq 0$. Mais pour un λ convenable, on a aussi $q(y) = 0$ (prendre $\lambda = -q(z)/2f(x, z)$) et on conclut alors à l'aide de la remarque 2.2.2.
Si on a 3), la forme q est non dégénérée car son discriminant est non nul. Par ailleurs, dans une base orthogonale on a $q(xe_1 + ye_2) = ax^2 + by^2$ et l'hypothèse affirme que $-ab$ est un carré non nul, donc aussi $-\frac{b}{a} = -ab(\frac{1}{a^2})$ et on a alors le vecteur isotrope $\sqrt{-\frac{b}{a}}\, e_1 + e_2$.

L'une des propriétés essentielles des plans hyperboliques est qu'on peut toujours englober une droite isotrope dans un tel plan :

Proposition 2.4.
Soit $x \in E$ un vecteur isotrope (non nul). Alors, il existe un plan P contenant x tel que $q|_P$ soit hyperbolique.

Démonstration. Comme q est non dégénérée, il existe $y \in E$ tel que $f(x, y) \neq 0$. Comme x est isotrope, y n'est pas colinéaire à x, donc $P = < x, y >$ est un plan.

Posons $a = f(x,y)$, dans la base x, y, $q|_P$ a pour matrice $\begin{pmatrix} 0 & a \\ a & b \end{pmatrix}$ avec $b = q(y)$, mais alors $q|_P$ est non dégénérée et d'indice ≥ 1, donc est hyperbolique d'après la proposition 2.3.

Proposition 2.5 (Classification des plans suivant leurs isotropes).
Soit P un plan muni d'une forme quadratique q (éventuellement dégénérée). Les cas suivants sont alors les seuls possibles :
a) q est anisotrope (donc non dégénérée), P ne contient aucune droite isotrope,
b) q est hyperbolique, P contient exactement deux droites isotropes,
c) q est de rang 1, P contient une seule droite isotrope (i.e. $\operatorname{Ker} q$), donc au moins trois droites non isotropes,
d) $q = 0$, toute droite de P est isotrope (et il y a au moins 4 droites).

Corollaire 2.6.
Si P contient 3 droites isotropes, on a $q|_P = 0$: P est totalement isotrope.

Démonstration (de 2.5). Si q est non dégénérée, on est dans les cas a) ou b) (cf. 2.3). Si q est de rang 1, soit e_1 une base de $\operatorname{Ker} q$, complétée par e_2. On a $q(e_2) = a \neq 0$ (car q est non nulle) et donc $q(xe_1 + ye_2) = ay^2$ ce qui montre que $< e_1 >$ est l'unique droite isotrope de P.
Comme on a $\operatorname{car}(k) \neq 2$, il y a au moins 4 droites dans un plan, portées par $e_1, e_2, e_1 + e_2, e_1 - e_2$ (cf. Chapitre IV § 5 Exercice 2), donc, ici, au moins trois droites non isotropes. Enfin, si $q = 0$, toutes les droites sont isotropes et il y en a au moins 4.

On a, avec les plans de rang 1, un résultat analogue à 2.4 :

Proposition 2.7.
On suppose $n \geq 3$. Soit $x \in E$ un vecteur isotrope non nul. Il existe un plan P contenant x, tel que $q|_P$ soit de rang 1.

Démonstration. Soit $H = (x)^\perp$ l'hyperplan orthogonal à x. L'hyperplan H n'est pas totalement isotrope, sinon on aurait $n - 1 \leq \nu(q) \leq \frac{n}{2}$, (cf. V, 3.12), donc $n \leq 2$. Soit donc $y \in H$ non isotrope donc non colinéaire à x. Alors le plan $P = < x, y >$ convient.

3. Espaces hyperboliques.

Définition 3.1.
Avec les notations du § 1, (E, q) est appelé un **espace hyperbolique** *si on a une décomposition en somme directe orthogonale : $E = P_1 \perp \ldots \perp P_r$ où chaque P_i, muni de $q|_{P_i}$, est un plan hyperbolique. La forme q est dite hyperbolique.*

Remarques 3.2.
1) Si E est hyperbolique, la dimension de E est égale à $2r$ donc est paire.
2) Si on prend une base hyperbolique (e_i, ε_i) de chaque P_i la matrice de q dans la base $e_1, \ldots, e_r, \varepsilon_1, \ldots, \varepsilon_r$ est de la forme : $\left(\begin{array}{c|c} 0 & I_r \\ \hline I_r & 0 \end{array}\right)$, où I_r est la matrice unité de dimension r.

3) Dans une base orthogonale convenable, q a pour matrice $\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & -I_r \end{array}\right)$, (cf. 2.2.4).

4) Dans un espace hyperbolique de dimension $n = 2r$ il y a un sous-espace totalement isotrope de dimension r (prendre un vecteur isotrope dans chaque P_i). Un tel sous-espace est parfois appelé un lagrangien. Notons que l'indice $\nu(q)$ vaut $\frac{n}{2}$ c'est-à-dire le maximum possible.

5) Lorsque l'on a $k = \mathbf{R}$ et $n = 2p$, la forme q est hyperbolique si et seulement si elle est de signature (p, p).

La proposition suivante est une généralisation de 2.4.

Proposition 3.3.
Soit F un sous-espace de E, $F_0 = \operatorname{Ker} q|_F$, U un supplémentaire de F_0 dans F, de sorte que l'on a $F = F_0 \perp U$. Soit $u_1, \ldots, u_r$ une base de F_0. Il existe alors des vecteurs $v_1, \ldots, v_r \in E$ tels que :
1) pour $i = 1, \ldots, r$, $P_i = < u_i, v_i >$ est un plan hyperbolique (pour $q|_{P_i}$) et u_i, v_i en est une base hyperbolique,
2) les sous-espaces $U, P_1, \ldots, P_r$ sont en somme directe dans E et deux à deux orthogonaux.

Commentaire 3.4.

Posons $G = P_1 \perp \ldots \perp P_r \perp U$. Comme U et les P_i sont des sous-espaces non isotropes de E, G est non isotrope (il suffit de regarder le discriminant pour s'en convaincre). Comme G contient F, on a plongé le sous-espace F (qui est isotrope si F_0 est non nul) dans un sous-espace non isotrope. De plus, la dimension de G (i.e. $\dim F + \dim F_0$) est minimale pour cette propriété. En effet, si on a $F \subset G'$, avec G' non isotrope, soit H l'orthogonal de F_0 dans G', on a $F \subset H$ donc : $\dim G' = \dim F_0 + \dim H \geq \dim F_0 + \dim F$.
Cette proposition permet dans de nombreux cas de ramener un problème concernant les sous-espaces au cas des sous-espaces non isotropes (cf. § 4). Elle admet de nombreux corollaires que nous énonçons avant de la démontrer.
Dans le cas particulier où F est totalement isotrope, la proposition 3.3 fournit le résultat suivant :

Corollaire 3.5.
Si $F \subset E$ est totalement isotrope, il existe $H \subset E$ hyperbolique tel que $F \subset H$ et $\dim H = 2 \dim F$, on dit que H est une extension hyperbolique de F.

Démonstration. On applique 3.3 à F, avec $F_0 = F$, $U = \{0\}$ et $H = P_1 \perp \ldots \perp P_r$.

On en déduit une caractérisation des espaces hyperboliques :

Corollaire 3.6.
Un espace (E, q) est hyperbolique si et seulement si il possède un lagrangien (i.e. un sous-espace F totalement isotrope avec $\dim E = 2 \dim F$).

Démonstration. Dans un sens c'est la remarque 3.2.4 précédente, dans l'autre le Corollaire 3.5 où on a $H = E$ pour une raison de dimension.

On peut aussi achever la démonstration de V, 6.6 (calcul de l'indice) :

Corollaire 3.7.
On suppose $k = \mathbf{R}$ et q de signature $(p, n-p)$. Alors on a $\nu(q) = \inf(p, n-p)$.

Démonstration. On a vu au Chapitre V l'inégalité $\nu(q) \geq \inf(p, n-p)$. Si F est totalement isotrope de dimension ν, soit H une extension hyperbolique de F. On a $E = H \perp H^\perp$ et, si (r, s) est la signature de $q|_{H^\perp}$, la signature de $q|_H$ étant (ν, ν), celle de q est donc $\nu + r, \nu + s$, d'où $\nu \leq \inf(p, n-p)$, (cf. aussi Chapitre V § 6 Exercice 6).

Il reste tout de même à prouver la proposition 3.3 :

On raisonne par récurrence sur $r = \dim F_0$.

1) Si on a $r = 1$, on a $U \subset F$ et $U \neq F$, donc $F^\perp \subset U^\perp$ et $F^\perp \neq U^\perp$.
Soit $a \in U^\perp - F^\perp$ de sorte que $f(a, u_1)$ est non nul. Alors, $P = < a, u_1 >$ est un plan, hyperbolique (cf. 2.3) dont on trouve le vecteur v_1 par 2.2.3. Enfin, P est orthogonal à U.

2) Supposons $r > 1$ et posons $U_1 = < u_2, \ldots, u_r > \perp U$. On a $U_1 \subset F$ et $U_1 \neq F$, donc aussi $F^\perp \subset U_1^\perp$ et $F^\perp \neq U_1^\perp$. Soit $a \in U_1^\perp - F^\perp$, on a $f(a, u_1) \neq 0$ et $P_1 = < a, u_1 >$ est un plan hyperbolique orthogonal à U_1. Posons $F_1 = U_1 \perp P_1 = < u_2, \ldots, u_r > \perp U \perp P_1$.
On a $\mathrm{Ker}\, q|_{F_1} = < u_2, \ldots, u_r >$ (car $U \perp P_1$ est non isotrope). On peut donc appliquer à F_1 l'hypothèse de récurrence. On trouve alors des vecteurs $v_2, \ldots, v_r$ de E tels que $P_i = < u_i, v_i >$ soit hyperbolique, tels qu'on ait $P_i \perp P_j$ pour $i, j \geq 2$, $i \neq j$, et $P_i \perp (U \perp P_1)$ pour $i \geq 2$ et on trouve enfin v_1 par 2.2.3.

4. Le théorème de Witt.

a) Le théorème de Witt, énoncé et démonstration.

Nous avons vu, au Chapitre VI, Lemme 3.1, que, dans le cas euclidien, le groupe $O(q)$ opère transitivement sur les sous-espaces de dimension donnée de E. Mais nous avons remarqué que le résultat ne subsistait pas dans le cas général. Le théorème de Witt décrit cette situation en ramenant la détermination des orbites de $O(q)$ dans les sous-espaces de E à un problème d'équivalence des formes :

Théorème 4.1 : (Witt).
Soient F, F' deux sous-espaces de E. Les conditions suivantes sont équivalentes :
1) il existe $u \in O(q)$ tel que $u(F) = F'$,
2) les formes $q|_F$ et $q|_{F'}$, sont équivalentes,
3) il existe une isométrie $\sigma : F \longrightarrow F'$, relative à $q|_F$ et $q|_{F'}$.

Remarques 4.2.

1) On peut aussi chercher les orbites sous $O^+(q)$, cf. § 4 Exercice 1.

2) Le théorème de Witt est valable dans le cas d'une forme alternée ou hermitienne, cf. [D] Ch. I § 11.

3) La variante 2) est particulièrement commode quand on connaît bien la classification des formes sur k (par exemple, si k est algébriquement clos ou $k = \mathbf{R}$, ou $k = \mathbf{F}_q$). Plus généralement, on peut appliquer le théorème à des sous-espaces F tels que $q|_F$ soit bien connue ; on en déduit par exemple que $O(q)$ est transitif sur les droites isotropes, ou sur les plans hyperboliques ...

Démonstration (de 4.1).

1) Il est clair que 2) et 3) sont équivalents et que 1) implique 3). Pour prouver que 3) implique 1), nous prouverons en fait le résultat plus précis suivant :

Théorème 4.3 (Witt).
Si on a $\sigma : F \longrightarrow F'$ comme en 4.1.3, il existe $u \in O(q)$ tel que $u|_F = \sigma$.

Démonstration (de 4.3).

a) *Réduction au cas où F est non isotrope.*

Supposons le théorème prouvé dans ce cas, et soit F isotrope. On utilise la Proposition 3.3 en écrivant : $F = F_0 \perp U$, avec $F_0 = \operatorname{Ker} q|_F$. Comme σ est une isométrie, on a $\sigma(F) = F' = \sigma(F_0) \perp \sigma(U)$ et $\sigma(F_0) = \operatorname{Ker} q|_{F'}$. Soit $u_1, \ldots, u_r$ une base de F_0 et $v_1, \ldots, v_r$ des vecteurs, comme en 3.3. Alors si on pose $u'_i = \sigma(u_i)$, $u'_1 \ldots u'_r$ est une base de $\sigma(F_0)$ et la même proposition 3.3 fournit des vecteurs $v'_1, \ldots, v'_r$ avec les propriétés analogues. Si on pose $P_i = <u_i, v_i>$, $P'_i = <u'_i, v'_i>$, les plans P_i, P'_i sont hyperboliques, de bases hyperboliques (u_i, v_i) et (u'_i, v'_i) et on a $P_i \perp P_j$, pour $i \neq j$, et $P_i \perp U$ et de même, $P'_i \perp P'_j$, pour $i \neq j$, et $P'_i \perp \sigma(U)$.
On pose alors $G = P_1 \perp \ldots \perp P_r \perp U$ et $G' = P'_i \perp \ldots \perp P'_r \perp \sigma(U)$. Le sous-espace G est non isotrope et on prolonge σ en σ_1 définie sur G en posant $\sigma_1(v_i) = v'_i$. Il est clair que σ_1 est une isométrie de G sur G' et comme G est non isotrope, σ_1 s'étend en une isométrie de E, donc aussi σ.

b) *Démonstration dans le cas où F est non isotrope.*
On raisonne par récurrence sur $\dim F$.

1) On suppose $\dim F = 1$, de sorte qu'on a $F = < x >$, avec $q(x) \neq 0$. Soit $y = \sigma(x)$, on a donc $q(x) = q(y)$. On a alors le lemme suivant :

Lemme 4.4.
Si on a $q(x) = q(y) \neq 0$, l'un des vecteurs $x + y$ ou $x - y$ est non isotrope.

(Sinon on a $q(x+y) = 0 = 2q(x) + 2f(x,y)$, $q(x-y) = 0 = 2q(x) - 2f(x,y)$, d'où, en ajoutant, $4q(x) = 0$ et c'est absurde).

Soit alors $\varepsilon = \mp 1$ tel que $x + \varepsilon y$ soit non isotrope (cf. 4.4), et $H = <x + \varepsilon y>^{\perp}$ qui est un hyperplan non isotrope. Soit τ_H la réflexion orthogonale par rapport à H (cf. V, 4.6). Comme $x - \varepsilon y$ est dans H, on a $\tau_H(x) = -\varepsilon y$, donc $-\varepsilon \tau_H(x) = y$. Mais alors, $-\varepsilon \tau_H$ est dans $O(q)$ et prolonge σ, cqfd.

2) Le cas $\dim F > 1$.
On suppose avoir prouvé que pour tout espace (E, q), tout sous-espace F non isotrope de E de dimension $\leq r$ et toute isométrie σ de F dans E, σ s'étend en une isométrie de E. Soit alors $F \subset E$ non isotrope avec $\dim F = r + 1$ et $\sigma : F \longrightarrow F'$. On peut écrire $F = F_1 \perp F_2$ avec $F_i \subset F$, $F_i \neq F$ et F_i non isotrope (on prend une base orthogonale de F et on la partage en deux).
Alors, $\sigma_1 = \sigma|_{F_1}$ se prolonge à E en $\tau_1 \in O(q)$ par l'hypothèse de récurrence. Quitte à remplacer σ par $\tau_1^{-1}\sigma$, on est ramené au cas où $\sigma|_{F_1} = \mathrm{Id}$ (si $u \in O(q)$ prolonge $\tau_1^{-1}\sigma$, $\tau_1 u$ prolongera σ).
Mais, si $\sigma|_{F_1} = \mathrm{Id}$, comme on a $F_2 \subset F_1^{\perp}$, on en déduit $\sigma(F_2) \subset F_1^{\perp}$. On applique alors l'hypothèse de récurrence à F_2 vu comme sous-espace de $F_1^{\perp}$. Alors, $\sigma_2 = \sigma|_{F_2}$ s'étend en une isométrie τ_2 de $F_1^{\perp}$. Mais alors, comme on a $E = F_1 \perp F_1^{\perp}$, l'isométrie $\mathrm{Id}_{F_1} \perp \tau_2$ prolonge σ ce qui achève de prouver 4.3.

b) Le théorème de Witt : les corollaires classiques.

Notations : Nous utiliserons, outre les notations du § 1, les abrévations suivantes :
seti pour sous-espace totalement isotrope,

setim pour sous-espace totalement isotrope maximal.

Le premier corollaire permet de mieux comprendre ce qu'est l'indice $\nu(q)$:

Corollaire 4.5.
a) Tout seti est contenu dans un setim,
b) Si F est un seti et F' un setim, on a $\dim F \leq \dim F'$,
c) Tous les setim ont même dimension, à savoir $\nu(q)$.

Démonstration.
Le point a) est clair car E est de dimension finie.
Pour b), supposons $\dim F > \dim F'$. Soit $F_1 \subset F$ tel que $\dim F_1 = \dim F'$. Comme $q|_{F_1} = q|_{F'} = 0$, il existe une isométrie u de E telle que $u(F_1) = F'$ (th. de Witt). Mais alors on a $u(F) \supset F'$, $u(F) \neq F'$ (dimension) et $u(F)$ est totalement isotrope, ce qui contredit le fait que F' est un *setim*. Enfin, c) résulte aussitôt de b).

Corollaire 4.6.
Si $F, F' \subset E$ sont isométriques (i.e. si on a $q|_F \sim q|_{F'}$), $F^\perp$ et $F'^\perp$ sont isométriques.

Démonstration. Par Witt, on a $u \in O(q)$ tel que $u(F) = F'$ donc on a aussi $u(F^\perp) = F'^\perp$, ce qui prouve que ces sous-espaces sont isométriques.

Corollaire 4.7.
1) L'espace E admet une décomposition orthogonale $E = H \perp G$ avec H hyperbolique et G anisotrope.
2) Si on a une telle décomposition, on a $\dim H = 2\nu(q)$.
3) Si on a deux décompositions du type de 1), $E = H \perp G = H' \perp G'$, il existe $u \in O(q)$ tel que $u(H) = H'$ et $u(G) = G'$.
Si on pose $q_a = q|_G$, la forme anisotrope q_a est bien définie par q, à équivalence près.

Démonstration.
1) Soient F un *setim* de E, H une extension hyperbolique de F (cf. 3.5) et $G = H^\perp$. On a $E = H \perp G$ car H est non isotrope et G est anisotrope (sinon F ne serait pas un *setim*).
2) Soit F un *setim* de H, on a $\dim H = 2 \dim F$ et $F \perp G \subset F^\perp$, donc il y a égalité pour une raison de dimension. Soit $x \in F^\perp$, on a $x = y + z$, avec $y \in F$ et $z \in G$, donc $q(x) = q(z)$. Comme $q|_G$ est anisotrope, si $x \notin F$, x n'est donc pas isotrope, ce qui prouve que F est aussi un *setim* de E, donc qu'on a $\dim F = \nu$.
3) Comme on a $\dim H = \dim H'$, les espaces hyperboliques H et H' sont isométriques, donc d'après Witt, il existe $u \in O(q)$ tel que $u(H) = H'$, donc aussi $u(G) = u(H^\perp) = H'^\perp = G'$.

Corollaire 4.8 (équivalence des formes).
Soient q, q' deux formes non dégénérées sur E. Avec les notations de 4.7, on a :

$$q \sim q' \iff \begin{cases} 1)\ \nu(q) = \nu(q'), \\ 2)\ q_a \sim q'_a. \end{cases}$$

(Cela résulte aussitôt du Corollaire 4.7).

Remarque 4.9. Le corollaire 4.8 ramène la classification des formes quadratiques à celle des formes anisotropes, mais il faut être conscient du fait que dans les cas non triviaux (par exemple pour $k = \mathbf{Q}$), le plus gros du travail reste à faire ! (cf. [S1] Chapitre IV).

Corollaire 4.10.
On suppose $\dim E = 3$ *et* $\nu(q) = 1$. *Alors* q *est équivalente, à un scalaire près, à la forme de Lorentz* L *(i.e. la forme* $x^2 + y^2 - z^2$*). En particulier on a* $O(q) \simeq O(L)$.

Démonstration. On écrit $E = P \perp D$ où P est un plan hyperbolique et D une droite non isotrope (cf. Corollaire 4.7). Si on a $D = <e>$ et $q(e) = \lambda$, la forme q/λ est équivalente à la forme de Lorentz. En effet, le plan P est encore hyperbolique pour cette forme et, dans une base convenable, q/λ a donc pour matrice :

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Ce corollaire montre qu'en dimension 3, pour les formes d'indice 1, il y a un seul groupe orthogonal à étudier (qui l'a d'ailleurs été à la fin du Chapitre VII). Pour la dimension 4, cf. Exercice 2 § 4.

5. Générateurs et centres de $O(q)$ et $O^+(q)$.

Les techniques sont les mêmes qu'au Chapitre VI, mais il faut faire attention aux sous-espaces isotropes. Dans toute la suite de ce chapitre les réflexions et renversements considérés seront toujours orthogonaux (cf. V, Proposition 4.6 et Remarques 4.7).

a) Les centres de $O(q)$ *et de* $O^+(q)$.

Théorème 5.1.
Pour $n \geq 3$, *le centre de* $O(q)$ *est* $Z = \{\mathrm{Id}, -\mathrm{Id}\}$.

Démonstration. Soit D une droite non isotrope et τ_D la réflexion définie par D. Si u centralise $O(q)$, on a $u\tau_D u^{-1} = \tau_{u(D)} = \tau_D$ (cf. V, 4.8) donc on a $u(D) = D$ et u laisse invariantes toutes les droites non isotropes.
Si P est un plan non isotrope, il est déterminé par deux de ses droites non isotropes (par exemple une base orthogonale). On a donc aussi $u(P) = P$. La conclusion résulte alors de IV, 2.8 et du lemme suivant :

Lemme 5.2.
On suppose $n \geq 3$. *Alors, toute droite isotrope est intersection de deux plans hyperboliques, donc non isotropes.*

Démonstration (de 5.2) Soit $x \neq 0$, un vecteur isotrope et $y \in E$ tel que $f(x, y) \neq 0$, de sorte que le plan $P = <x, y>$ est hyperbolique (cf. 2.3). Alors, P et donc aussi $P^\perp$, sont non isotropes. Soit $z \in P^\perp$, non isotrope. On a $f(x, y + z) = f(x, y) \neq 0$, donc $<x, y + z>$ est un autre plan hyperbolique et on a :

$$<x> = <x, y> \cap <x, y + z>.$$

Théorème 5.3.
Pour $n = 2$, le centre de $O(q)$ est $\{\mathrm{Id}, -\mathrm{Id}\}$, sauf si E est un plan hyperbolique sur le corps $\mathbf{F}_3$.

Démonstration. Si u centralise $O(q)$, u laisse invariantes toutes les droites non isotropes de E. Si q est anisotrope, il est clair que u est une homothétie. Sinon, E est un plan hyperbolique, dont deux droites seulement sont isotropes (cf. 2.2.3) et qui contient donc, si $|k| > 3$, au moins 3 droites non isotropes (si e_1, e_2 sont isotropes, $e_1 + e_2$, $e_1 - e_2$, $e_1 + \lambda e_2$ avec $\lambda \neq 0$, 1, -1, sont non isotropes).
Mais alors u est une homothétie (car u admet trois droites propres distinctes). Le cas du plan hyperbolique sur $\mathbf{F}_3$ est effectivement exceptionnel, cf. § 6.

Théorème 5.4.
Pour $n \geq 3$ le centre de $O^+(q)$ est $Z \cap O^+(q)$, c'est-à-dire $\{\mathrm{Id}\}$ si n est impair, $\{\mathrm{Id}, -\mathrm{Id}\}$ si n est pair.

Démonstration. Soit P un plan non isotrope, σ_P le renversement de plan P. Si u centralise $O^+(q)$, on a : $u\sigma_P u = \sigma_P = \sigma_{u(P)}$, donc $u(P) = P$. La conclusion résulte alors du lemme habituel (cf. IV, 2.8), du Lemme 5.2 ci-dessus, et du lemme suivant :

Lemme 5.5.
On suppose $n \geq 3$. Toute droite non isotrope est intersection de deux plans non isotropes.

Démonstration. Soit $D = \langle x \rangle$, avec x non isotrope et soient $y, z \in D^\perp$, non isotropes et linéairement indépendants (par exemple des vecteurs d'une base orthogonale de $D^\perp$, qui est non isotrope et de dimension ≥ 2 par hypothèse). Alors, on a $\langle x \rangle = \langle x, y \rangle \cap \langle x, z \rangle$ et les plans sont non isotropes.

Remarques 5.6.
1) Pour $n = 2$, nous verrons au § 6 que $O^+(q)$ est commutatif.
2) Remarquons que dans les théorèmes précédents, nous démontrons en fait que le centralisateur de $O(q)$ ou $O^+(q)$ dans $GL(E)$ est le groupe k^* des homothéties.

b) Générateurs des groupes $O(q)$ et $O^+(q)$.

Théorème 5.7.
Le groupe $O(q)$ est engendré par les réflexions.

Démonstration. On procède par récurrence sur n, la propriété est claire pour $n = 1$. Soit $n > 1$ et supposons le résultat établi jusqu'à $n - 1$. Soit $u \in O(q)$.

1) Supposons qu'il existe $x \in E$, $x \neq 0$, non isotrope, tel que $u(x) = x$. Soit $H = \langle x \rangle^\perp$ l'hyperplan orthogonal, non isotrope lui aussi. On a $u(H) = H$ et on peut appliquer l'hypothèse de récurrence à $u|_H$. On a : $u|_H = \tau_1 \ldots \tau_r$ où τ_i est une réflexion de H. Mais, si on pose $\sigma_i = \tau_i \perp \mathrm{Id}_{H^\perp}$, σ_i est une réflexion de E, et comme $u(x) = x$, on a $u = \sigma_1 \ldots \sigma_r$.

2) Soit $x \in E$, $x \neq 0$, non isotrope et soit $y = u(x)$. Supposons $x - y$ non isotrope et soit $H = (x - y)^\perp$. Comme $x + y$ est dans H, on a, si τ_H est la réflexion par rapport à H, $\tau_H(y) = x$, donc $\tau_H \circ u(x) = x$. On est donc ramené au cas 1), on a $\tau_H u = \tau_1 \ldots \tau_r$, donc $u = \tau_H \tau_1 \ldots \tau_r$.

3) Avec les notations de 2), si le vecteur $x - y$ est isotrope, alors $x + y$ est non isotrope (cf. 4.4). Alors, si $H = \langle x + y \rangle^\perp$, on a $\tau_H(y) = -x$. Soit alors $L = \langle x \rangle^\perp$, on a $\tau_L(x) = -x$, d'où $\tau_L \tau_H u(x) = x$ et on est ramené au cas 1).

Remarque 5.8. Nous montrerons au §7 que tout élément de $O(q)$ est produit d'au plus n réflexions.

Théorème 5.9.
Pour $n \geq 3$, le groupe $O^+(q)$ est engendré par les renversements.

Démonstration. Si τ est une réflexion, on a $\det \tau = -1$, donc si u est dans $O^+(q)$, u est produit d'un nombre pair de réflexions. Le théorème 5.9 résulte alors du lemme suivant :

Lemme 5.10.
Si on a $n \geq 3$ et si τ_1, τ_2 sont des réflexions, il existe des renversements σ_1, σ_2 tels que $\sigma_1\sigma_2 = \tau_1\tau_2$.

Démonstration (de 5.10).
1) Si $n = 3$, $-\tau_i$ est un renversement et on a $\tau_1\tau_2 = (-\tau_1)(-\tau_2)$.
2) Dans le cas général, soient H_1, H_2 les hyperplans de τ_1, τ_2. On a $H_i^\perp = <x_i>$, avec x_i non isotrope, donc $(H_1 \cap H_2)^\perp = <x_1, x_2>$. On peut supposer $H_1 \neq H_2$, donc $\dim(H_1 \cap H_2) = n - 2$. Le sous-espace $H_1 \cap H_2$ est peut-être isotrope, mais comme $\operatorname{Ker} q|_{H_1 \cap H_2} = \operatorname{Ker} q|_{<x_1, x_2>}$, et comme x_1, x_2 sont non isotropes, on a $\dim \operatorname{Ker} q|_{H_1 \cap H_2} \leq 1$. Il existe donc un sous-espace V de $H_1 \cap H_2$ non isotrope de dimension $n - 3$, et on a alors $E = V \perp V^\perp$. Sur V, on a $\tau_1 = \tau_2 = \mathrm{Id}$, et sur $V^\perp$, qui est de dimension 3, on a $\tau_1\tau_2 = \sigma_1\sigma_2$, σ_i renversement de $V^\perp$. Il suffit alors de prolonger σ_i par l'identité sur V en σ_i' qui est encore un renversement, pour avoir $\tau_1\tau_2 = \sigma_1'\sigma_2'$.

6. La dimension 2.

Dans tout ce paragraphe, on suppose $\dim E = 2$.

a) Les éléments de $O(q)$.

Le premier résultat est le même que dans le cas euclidien, cf. VI, 4.1 :

Théorème 6.1.
1) Si u est dans $O^-(q)$, u est une réflexion.
2) Si u est dans $O^+(q)$, on a $u = \tau_1\tau_2$, où $\tau_1\tau_2$ sont des réflexions, l'une d'entre elles pouvant être choisie arbitrairement.
3) Soient $u \in O^+(q)$ et $\tau \in O^-(q)$, on a $\tau u \tau^{-1} = \tau u \tau = u^{-1}$.
4) Le groupe $O^+(q)$ est commutatif.
5) Soit $u \in O^+(q)$. On suppose qu'il existe $x \in E$, non nul et non isotrope, tel que $u(x) = x$. Alors on a $u = \mathrm{Id}$.

Démonstration.
1) Soit $u \in O^-(q)$ et $x \in E$, non nul et non isotrope. Il existe une réflexion τ telle que $\tau u(x) = \mp x$ (réflexion associée au vecteur $x - u(x)$ ou $x + u(x)$, cf. 5.7). Alors, τu est dans $O^+(q)$ et admet la valeur propre $\varepsilon = \mp 1$. Comme on a $\det \tau u = 1$ l'autre valeur propre de τu est aussi ε. De plus comme x est non isotrope, la droite orthogonale à x est distincte de $< x >$ et stable par τu donc propre pour τu, relativement à la valeur propre ε, mais alors τu est l'homothétie de rapport ε i.e. $\tau u = \mp \mathrm{Id}$, donc $u = \mp \tau$. Mais, comme on a $n = 2$, $-\tau$ est aussi une réflexion, donc u est une réflexion. On notera que cette démonstration prouve aussi 5).

2) Si u est dans $O^+(q)$ et si τ est une réflexion quelconque, τu est dans $O^-(q)$, donc est une réflexion et on a $u = \tau(\tau u)$.

3) Soient $u \in O^+(q)$ et τ une réflexion, on peut écrire $u = \tau\tau'$, où τ' est une autre réflexion. Alors, on a $\tau u \tau^{-1} = \tau u \tau = \tau^2\tau'\tau = \tau'\tau = u^{-1}$.

4) Soient $u, v \in O^+(q)$, avec $u = \tau_1\tau_2$. On a d'après 3), $uvu^{-1} = \tau_1\tau_2 v \tau_2\tau_1 = \tau_1 v^{-1}\tau_1 = v$ donc $O^+(q)$ est commutatif.

b) Détermination du groupe $O^+(q)$: le cas hyperbolique.

Supposons q hyperbolique et soit e_1, e_2 une base hyperbolique de E. Dans cette base q a pour matrice $J = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ et si $u \in O(q)$ a pour matrice A, on a ${}^tAJA = J$. Un calcul immédiat fournit les matrices :

$$A = \begin{pmatrix} 0 & \alpha \\ 1/\alpha & 0 \end{pmatrix} \in O^-(q) \quad \text{et} \quad A = \begin{pmatrix} \alpha & 0 \\ 0 & 1/\alpha \end{pmatrix} \in O^+(q), \quad \text{avec } \alpha \in k^*.$$

On en déduit la proposition suivante :

Proposition 6.2.
Le groupe $O^+(q)$, pour q hyperbolique, est isomorphe au groupe des matrices $\begin{pmatrix} \alpha & 0 \\ 0 & 1/\alpha \end{pmatrix}$ *avec $\alpha \in k^*$, lui-même isomorphe à k^*.*

Remarque 6.3. Dans le cas $k = \mathbf{F}_3$, on a $|O^+(q)| = |\mathbf{F}_3^*| = 2$, donc $|O(q)| = 4$ et le groupe $O(q)$ est commutatif composé de Id, $-$Id et de deux réflexions. Comme tous ses éléments $\neq$ Id sont d'ordre 2, $O(q)$ est isomorphe au groupe de Klein, $V_4 = \mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$.
Rappelons (cf. 5.3) que dans tous les autres cas, $O(q)$ est non commutatif.

c) Détermination du groupe $O^+(q)$: le cas anisotrope.

La forme q est équivalente, à un scalaire près, à la forme $x^2 + \alpha y^2$, de matrice $J = \begin{pmatrix} 1 & 0 \\ 0 & \alpha \end{pmatrix}$, avec $-\alpha \notin k^{*2}$ (puisque q est anisotrope). Notons que $\alpha = \Delta(q)$ est le discriminant de q. Soit K l'extension de k obtenue en lui adjoignant une racine carrée w de $-\alpha$, on a $K = k[w] = k[\sqrt{-\alpha}]$ et K est donc de degré 2 sur k. Si z est dans K, il s'écrit $z = x + yw$ avec $x, y \in k$.

Soit σ l'automorphisme de K, appelé conjugaison, et défini par $\sigma(x+yw) = x - yw$. On note $\sigma(z) = \overline{z}$ et on définit, pour $z = x + yw \in K$, la "norme" de z par $N(z) = z\overline{z} = x^2 + \alpha y^2$. On a $N(z) \in k$ et N induit un homomorphisme de K^* dans k^*, de noyau $\mathbf{U}_K = \{z \in k \mid N(z) = 1\}$.

Comme dans le cas des quaternions (cf. Chapitre VII § 2), on identifie K à k^2 et N à q et on fait opérer $\mathbf{U}_K$ sur K par

$$(\lambda, z) \mapsto \lambda z = u_\lambda(z), \text{ pour } \lambda \in \mathbf{U}_K,\ z \in K.$$

Il est clair que u_λ est dans $GL(2, k)$ et, comme $N(\lambda z) = N(\lambda)N(z) = N(z)$ puisque $\lambda \in \mathbf{U}_K$, on a même $u_\lambda \in O(N) = O(q)$. De plus l'application $\lambda \mapsto u_\lambda$ est un homomorphisme de groupes de $\mathbf{U}_K$ dans $O(q)$.
Si, pour un $z \neq 0$, on a $u_\lambda(z) = z = \lambda z$, on en déduit $\lambda = 1$. Il en résulte :

1) que l'homomorphisme $u : \lambda \mapsto u_\lambda$ est injectif,

2) que u_λ n'est pas une réflexion, donc (cf. 6.1) que u_λ est dans $O^+(q)$.

On a donc un homomorphisme injectif :

$$u : \mathbf{U}_K \longrightarrow O^+(q).$$

Montrons enfin que u est surjectif. Soit en effet $\rho \in O^+(q)$, si on a $\rho(1) = \lambda$ (avec l'identification de k^2 et K), on a $N(\lambda) = N(1) = 1$, donc $\lambda \in \mathbf{U}_K$ et $\rho(1) = u_\lambda(1)$. On a alors $u_\lambda^{-1}\rho(1) = 1$, mais comme $u_\lambda^{-1}\rho$ est dans $O^+(q)$ et fixe 1, on a $u_\lambda^{-1}\rho = \mathrm{Id}$ (cf. 6.1.5) donc $\rho = u_\lambda$.

En définitive, on a donc, comme dans le cas euclidien, cf. Chapitre VI § 4, un isomorphisme :

Proposition 6.4.
Soit q une forme anisotrope, $\alpha = \Delta(q)$ un discriminant de q, $K = k[\sqrt{-\alpha}]$ et $\mathbf{U}_K$ le groupe des éléments de norme 1 de K. On a un isomorphisme :

$$\mathbf{U}_K \simeq O^+(q).$$

Remarques 6.5.

1) Dans la base $1, w$ de K, la matrice de u_λ, pour $\lambda = a + bw \in \mathbf{U}_K$ est :

$$\begin{pmatrix} a & -\alpha b \\ b & a \end{pmatrix} \quad \text{avec} \quad a^2 + \alpha b^2 = 1.$$

2) Pour $k = \mathbf{F}_q$ on connaît parfaitement le groupe $O^+(Q)$.

a) Si Q est hyperbolique, on a vu que $O^+(Q)$ est isomorphe à $\mathbf{F}_q^*$, donc cyclique de cardinal $q - 1$ (cf. III, 2.7).

b) Si Q est anisotrope, on a un isomorphisme $O^+(Q) \simeq \mathbf{U}_K$ avec ici $K = \mathbf{F}_{q^2}$ et on a la suite exacte (cf. V, 6.11) :

$$1 \longrightarrow \mathbf{U}_K \longrightarrow K^* \xrightarrow{N} k^* \longrightarrow 1.$$

On a donc $|\mathbf{U}_K| = q + 1$ et comme K^* est cyclique, $\mathbf{U}_K$, donc aussi $O^+(Q)$, est un groupe cyclique de cardinal $q + 1$.

7. Le théorème de Cartan-Dieudonné.

Nous précisons ici le théorème 5.7. Les notations sont toujours celles du § 1, E est de nouveau de dimension n quelconque.

Théorème 7.1 (E. Cartan, J. Dieudonné).
Soit $u \in O(q)$, alors u est produit d'au plus n réflexions.

Démonstration. On raisonne par récurrence sur n. Pour $n = 1$, le théorème est clair, avec la convention usuelle que Id est produit de zéro réflexion. Notons que, pour $n = 2$, le théorème a été prouvé au § 6 (6.1.1 et 6.1.2).
Supposons donc $n \geq 3$. On reprend la démonstration de 5.7, mais avec plus de précision.

1) *Premier cas :*
S'il existe $x \in E$, $x \neq 0$ et non isotrope tel que $u(x) = x$, et si $H = <x>^\perp$, on a $u(H) = H$. Appliquant l'hypothèse de récurrence à $u|_H$, on voit que u est produit d'au plus $n - 1$ réflexions.
Notons que si l'on n'est pas dans ce premier cas, c'est que, si $v = u - \mathrm{Id}$, $\mathrm{Ker}\, v$ est totalement isotrope.

2) *Deuxième cas :*
Supposons qu'il existe $x \in E$, non isotrope, tel que $v(x) = x - u(x)$ soit non isotrope. Alors, si $H = < v(x) >^{\perp}$ et si τ_H est la réflexion par rapport à H, on a $\tau_H.u(x) = x$ et $\tau_H.u$ est dans le cas 1), donc u est produit de n réflexions au plus.

3) Remarquons que si l'on n'est pas dans le cas 1) ou 2), la technique de 5.7 n'aboutit pas, puisqu'elle permet seulement de prouver que u est produit de $n+1$ réflexions, donc n'assure pas le fonctionnement de la récurrence.

4) Analysons donc la situation lorsque 2) n'est pas vérifié. On a $v(x)$ isotrope, pour tout x *non isotrope* de E. En fait, $v(E)$ est totalement isotrope : il suffit de montrer que si x est isotrope, $v(x)$ est isotrope. En vertu de 2.7, on peut plonger x dans un plan P, de rang 1. Alors, P contient au moins 3 droites non isotropes (cf. 2.5.c) donc $v(P)$ contient au moins 3 droites isotropes, donc est totalement isotrope (cf. 2.6) donc $v(x)$ est isotrope. ([1])

5) Lorsque ni 1) ni 2) ne sont vérifiés, on a donc $\operatorname{Ker} v$ et $\operatorname{Im} v$ totalement isotropes. Si ν est l'indice de q, il en résulte qu'on a

$$r = \dim \operatorname{Ker} v \leq \nu, \qquad s = \dim \operatorname{Im} v \leq \nu,$$

donc $r + s \leq 2\nu \leq n$. Mais, la formule usuelle des dimensions :

$$\dim \operatorname{Ker} v + \dim \operatorname{Im} v = n$$

implique qu'on a $n = 2\nu$ (donc que q est hyperbolique) et que $\operatorname{Ker} v$ et $\operatorname{Im} v$ en sont deux *setim* (cf. §4).

Soit alors $e_1, \ldots, e_\nu$ une base de $\operatorname{Ker} v$. Il existe des éléments $\varepsilon_1, \ldots, \varepsilon_\nu \in E$ tels que $< e_i, \varepsilon_i >= P_i$ soit un plan hyperbolique, de base hyperbolique (e_i, ε_i) et que les P_i soient orthogonaux avec $E = P_1 \perp \ldots \perp P_\nu$ (cf. 3.3 et 3.5).
Vu le choix des e_i, on a, pour tout i, $u(e_i) = e_i$. Posons $u(\varepsilon_i) = \sum_j a_{ij} e_j + \sum_j b_{ij} \varepsilon_j$.
On a $f(u(\varepsilon_i), u(e_j)) = b_{ij} = f(\varepsilon_i, e_j) = \delta_{ij}$, (où δ_{ij} est l'indice de Kronecker).
Autrement dit, dans la base $(e_1, \ldots, e_\nu, \varepsilon_1, \ldots, \varepsilon_\nu)$, u a pour matrice $\left(\begin{array}{c|c} I & A \\ \hline 0 & I \end{array}\right)$, en particulier on a $\det u = 1$.

6) Lorsque 1) et 2) ne sont pas vérifiés, on a donc $n = 2\nu$ et $\det u = 1$.
Soit alors τ une réflexion quelconque. Comme $\det \tau u = -1$, τu n'est pas dans le cas exceptionnel, donc τu est produit d'au plus n réflexions, mais comme n est pair et $\tau u \in O^-(q)$, τu est produit d'au plus $n-1$ réflexions, donc u d'au plus n, et la démonstration est achevée.

On en déduit aussitôt avec 5.10 :

Corollaire 7.2.
On suppose $n \geq 3$. Si u est dans $O^+(q)$, u est produit d'au plus n renversements.

Remarque 7.3. Si le résultat du Théorème 7.1 est le même que dans le cas euclidien, en revanche on ne peut plus affirmer que, pour $u \in O(q)$ donné, le nombre minimal de réflexions nécessaires pour écrire u est égal à p_u (codimension de l'espace des points fixes de u, ou encore rang de $v = u - \mathrm{Id}$), cf. §5,7, Exercice 2.

([1]) Le lecteur sourcilleux qui aurait remarqué que $v(P)$ peut être de dimension < 2 montrera que même dans ce cas $v(P)$ est totalement isotrope.

8. Le groupe des commutateurs.

Le groupe des commutateurs de $O(q)$ sera noté $\Omega(q)$ (ou $\Omega_n(k,q)$ si on veut préciser le corps de base et la dimension).

Remarque 8.1. Si g et h sont dans $O(q)$, leur commutateur $ghg^{-1}h^{-1}$ est de déterminant 1, de sorte qu'on a $\Omega(q) \subset O^+(q)$.

Alors que, jusqu'à présent, la structure des groupes orthogonaux était assez voisine entre le cas euclidien et le cas général, elle va devenir sensiblement différente avec l'étude du groupe $\Omega(q)$. En particulier, on aura, en général, $\Omega(q) \neq O^+(q)$.

Théorème 8.2.
1) Le groupe $\Omega(q)$ est engendré par les produits $s(wsw^{-1})$ de deux réflexions conjuguées,
2) le groupe $\Omega(q)$ est engendré par les commutateurs $sts^{-1}t^{-1} = (st)^2$, où s et t sont des réflexions,
3) si u est dans $O(q)$, u^2 est dans $\Omega(q)$ et les carrés des éléments de $O(q)$ engendrent $\Omega(q)$,
4) le groupe $O(q)/\Omega(q)$ est commutatif et formé d'éléments d'ordre ≤ 2,
5) pour $n \geq 3$, on a aussi $\Omega(q) = D(O^+(q))$, groupe des commutateurs de $O^+(q)$.

Remarque 8.3. Le point 3) est conséquence d'un lemme plus général, cf. Chapitre VI § 3 Exercice 4.
Démonstration (de 8.2).
1) Soit $c = uvu^{-1}v^{-1}$ un commutateur, avec $u, v \in O(q)$. On écrit $u = \tau_1 \dots \tau_p$, où les τ_i sont des réflexions et on montre, par récurrence sur p, que u est produit de commutateurs du type 1).
C'est clair si $p = 1$; pour $p > 1$, on pose $\alpha = \tau_1 \dots \tau_{p-1}$, et on a $c = \alpha\tau_p v \tau_p \alpha^{-1} v^{-1}$, ou encore, en posant $\beta = \tau_p v \tau_p$, $c = \alpha\beta\alpha^{-1}\beta^{-1}\beta v^{-1}$. Comme α est produit de $p-1$ réflexions, on peut appliquer l'hypothèse de récurrence à $\alpha\beta\alpha^{-1}\beta^{-1}$, qui est donc produit de commutateurs de type 1). Quant à $\beta v^{-1} = \tau_p v \tau_p v^{-1}$, il est lui-même du type annoncé.

2) Soit $c = s(wsw^{-1})$ un commutateur de type 1) (c'est-à-dire que s est une réflexion) et écrivons $w = \tau_1 \dots \tau_p$ comme produit de réflexions. En vertu de 1), il suffit de prouver que c est produit de commutateurs de type 2), ce que nous montrons par récurrence sur p, le cas $p = 1$ étant clair. Soit $\alpha = \tau_1 \dots \tau_{p-1}$, le commutateur $c = s\alpha\tau_p s\tau_p\alpha^{-1}$ peut encore s'écrire : $c = (s\alpha s\alpha^{-1})(\alpha s\tau_p s\tau_p\alpha^{-1})$. Comme α est produit de $p-1$ réflexions, on peut appliquer l'hypothèse de récurrence à $s\alpha s\alpha^{-1}$. D'autre part, le deuxième terme est conjugué par α de $s\tau_p s\tau_p$ donc est de la même forme : $s'\tau'_p s'\tau'_p$ avec $s' = \alpha s\alpha^{-1}$, $\tau'_p = \alpha\tau_p\alpha^{-1}$ qui sont des réflexions.

3) Soit $u = \tau_1 \dots \tau_p \in O(q)$, on montre par récurrence sur p qu'on a $u^2 \in \Omega(q)$. C'est clair pour $p = 1$. Sinon, on pose, là encore, $\alpha = \tau_1 \dots \tau_{p-1}$ et on a $u^2 = \alpha\tau_p\alpha\tau_p = \alpha^2\alpha^{-1}\tau_p\alpha\tau_p$, d'où le résultat en appliquant à α l'hypothèse de récurrence.
Le fait que les carrés engendrent $\Omega(q)$ résulte de 2).

4) Le groupe $O(q)/\Omega(q)$ est évidemment commutatif, et, d'après 3), ses éléments sont d'ordre ≤ 2.

5) Soit $C(q) = D(O^+(q))$, il est clair que $C(q)$ est inclus dans $\Omega(q)$. Réciproquement, si s et t des réflexions, comme on a $n \geq 3$, il existe des renversements σ, τ tels que l'on ait $st = \sigma\tau$ (cf. 5.10). Alors $(st)^2 = (\sigma\tau)^2 = \sigma\tau\sigma^{-1}\tau^{-1}$ est dans $C(q)$ et la conclusion résulte de 2).

Remarques 8.4.

1) Pour $n = 2$, $O^+(q)$ est commutatif, donc $D(O^+(q)) = \{\mathrm{Id}\}$.

2) Pour $n = 2$, $\Omega(q)$ est engendré par les carrés de $O(q)$. Comme les éléments de $O^-(q)$ sont de carré l'identité (cf. 6.1.1), $\Omega(q)$ est donc le groupe des carrés de $O^+(q)$. Si q est hyperbolique, on a donc $\Omega(q) = k^{*2}$ (cf. 6.2) et si q est anisotrope, $\Omega(q) = \mathbf{U}_K^2$ (cf. 6.4).

Théorème 8.5.

Pour $n \geq 3$, le centre de $\Omega(q)$ est $Z \cap \Omega(q)$ (où $Z = \{\mathrm{Id}, -\mathrm{Id}\}$).

Remarques 8.6.

1) Si n est impair, comme $\Omega(q)$ est inclus dans $O^+(q)$, le centre est réduit à $\{\mathrm{Id}\}$.

2) Si n est pair le centre n'est pas nécessairement égal à $\{\mathrm{Id}, -\mathrm{Id}\}$, cf. 9.2 ci-dessous.

Démonstration (de 8.5). Nous aurons besoin de deux lemmes :

Lemme 8.7.

Supposons $\dim E = 2$ et soit $u \in O^+(q)$, $u \neq \mathrm{Id}$, alors u n'admet pas la valeur propre 1.

Démonstration. Sinon, cette valeur propre serait double (car on a $\det u = 1$) et on aurait deux cas :

a) Si on avait $u(x) = x$ avec x non isotrope, l'orthogonal $<x>^\perp$ serait une autre droite propre pour la valeur propre 1 et on aurait $u = \mathrm{Id}$ (ce cas a déjà été vu en 6.1.5).

b) Si on avait $u(x) = x$ avec x isotrope, le plan serait hyperbolique, et l'autre droite isotrope serait stable par u (car il n'y a que deux droites isotropes) donc on aurait aussi $u = \mathrm{Id}$.

Lemme 8.8.

Si k est distinct de $\mathbf{F}_3$ et si $P \subset E$ est un plan non isotrope, il existe $u \in O(q)$ tel que $u|_{P^\perp} = \mathrm{Id}$ et $u^2 \neq \mathrm{Id}$. L'espace des points fixes de u^2 est alors $P^\perp$.

Démonstration. Comme on a $k \neq \mathbf{F}_3$, $O(q|_P)$ n'est pas commutatif (cf. 5.3) donc $\Omega(q|_P)$ n'est pas réduit à $\{\mathrm{Id}\}$, donc il existe $v \in O(q|_P)$ tel que $v^2 \neq \mathrm{Id}$ (cf. ci-dessus 8.2.3). Alors $u = v \perp \mathrm{Id}_{P^\perp}$ convient.

Revenons au Théorème 8.5 et supposons d'abord $k \neq \mathbf{F}_3$.

Soit $v \in O(q)$ centralisant $\Omega(q)$ et, si P est un plan non isotrope, soit $u \in O(q)$ comme au Lemme 8.8. Comme u^2 est dans $\Omega(q)$ (cf. 8.2.3), on a $vu^2 = u^2v$. Soit $x \in P^\perp$, $x \neq 0$. On a $u^2v(x) = v(x)$. Ecrivant $v(x) = y + z$ avec $y \in P$, $z \in P^\perp$, on a $u^2(y) = y$. Mais, comme $u^2|_P \neq \mathrm{Id}$, le Lemme 8.7 implique alors $y = 0$.

Autrement dit, on vient de montrer qu'on a $v(P^\perp) = P^\perp$, donc aussi $v(P) = P$, de sorte que v laisse invariant tous les plans non isotropes, et c'est donc une homothétie (cf. 5.2 et 5.5).

Lorsque $k = \mathbf{F}_3$, le Lemme 8.8 est en défaut si P est un plan hyperbolique (cf. 6.2). Cependant la démonstration précédente permet de montrer que si v est dans le centre de $\Omega(q)$, v laisse invariants tous les plans anisotropes.

Lorsque n est ≥ 4 la conclusion résulte du lemme suivant et du Lemme 5.2.

Lemme 8.9.
On suppose $k = \mathbf{F}_3$ et $n \geq 4$. Toute droite non isotrope est intersection de deux plans anisotropes.

Démonstration (de 8.9). Soit $D = < x >$ une droite non isotrope, on a donc $q(x) = \mp 1$. Soit $D^\perp$ l'orthogonal de D, on a $\dim D^\perp \geq 3$, donc (cf. V, 6.8), $D^\perp$ admet une base orthogonale $e_2, e_3, \ldots, e_n$ avec $q(e_2) = q(e_3) = 1$.
Si $q(x)$ vaut $+1$, on a $D = < x, e_2 > \cap < x, e_3 >$ et ces plans sont anisotropes (car leur discriminant est 1).
Si $q(x)$ vaut -1, on a $D = < x, \varepsilon_2 > \cap < x, \varepsilon_3 >$, avec $\varepsilon_2 = \dfrac{e_2 + e_3}{2}$, $\varepsilon_3 = \dfrac{e_2 - e_3}{2}$, qui vérifient $q(\varepsilon_2) = q(\varepsilon_3) = 2 = -1$, donc les deux plans sont anisotropes et ceci achève de prouver 8.9.

Enfin, il reste le cas $k = \mathbf{F}_3$, $n = 3$. La forme q est équivalente à $x^2 + y^2 + z^2$ où à $x^2 + y^2 - z^2$ (cf. Chapitre V *loc. cit.*). Dans le deuxième cas, $-q$ est de discriminant 1, donc équivalente à $x^2 + y^2 + z^2$. Comme $O(q) = O(-q)$, on peut donc supposer $q \sim x^2 + y^2 + z^2$.
Soit $u \in Z(\Omega(q))$. Il s'agit de montrer que u est l'identité (dimension impaire). Les trois plans de coordonnées étant anisotropes sont invariants par u. Il en résulte que les droites engendrées par les vecteurs de base sont elles aussi invariantes par u. Comme on a $\det u = 1$, à une permutation près des vecteurs de base, on a

$$u = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad \text{ou } u = \mathrm{Id}.$$

Soit alors $v = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$, on a $v \in O(q)$ et $v^2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$, mais alors, si $u \neq \mathrm{Id}$ on a $uv^2 \neq v^2 u$, donc, comme v^2 est dans $\Omega(q)$ cela contredit $u \in Z(\Omega(q))$.

9. Compléments.

Nous avons donc exhibé une suite de sous-groupes de $O(q)$, chacun d'eux distingué dans le suivant :

$$\{1\} \lhd Z(\Omega(q)) \lhd \Omega(q) \lhd O^+(q) \lhd O(q).$$

Le problème est maintenant de calculer les quotients successifs.
On sait déja que $O(q)/O^+(q)$ est isomorphe à $\{1, -1\}$, que $O^+(q)/\Omega(q)$ est commutatif et formé d'éléments d'ordre 2 et que $Z(\Omega(q))$ est inclus dans $\{\mathrm{Id}, -\mathrm{Id}\}$.
Il reste essentiellement à élucider les trois questions suivantes :

1) décrire le groupe $O^+(q)/\Omega(q)$,
2) étudier le quotient $P\Omega(q) = \Omega(q)/Z(\Omega(q))$, en particulier sa simplicité,
3) pour n pair, préciser à quelle condition $-\mathrm{Id} \in \Omega(q)$.

Nous allons donner, ***sans démonstrations***, les résultats les plus significatifs sur ces questions. Ils nécessitent l'introduction d'objets nouveaux (algèbre de Clifford, norme spinorielle) ou de techniques différentes (par exemple sur $\mathbf{R}$ et $\mathbf{C}$, des méthodes topologiques, voire de groupes de Lie). On renvoie à [D] Ch. II § 7, 8, 9 pour des démonstrations, ainsi qu'aux exercices des § 8 et 9 ci-dessous.

Le trait commun de tous les résultats suivants est qu'ils nécessitent l'hypothèse $\nu(q) \geq 1$, i.e. ***l'existence de vecteurs isotropes.*** Le cas anisotrope est beaucoup plus compliqué (cf. [D] Ch. II § 12). A cet égard, l'exemple des formes euclidiennes sur $\mathbf{R}$ est trompeur, c'est le seul cas à notre connaissance où l'indice 0 est plus simple que l'indice ≥ 1.

Théorème 9.1.
On suppose $\nu(q) \geq 1$. *On a un isomorphisme* $\overline{\theta} : O^+(q)/\Omega(q) \xrightarrow{\sim} k^*/k^{*2}$.

L'isomorphisme $\overline{\theta}$ provient d'un homomorphisme de $O^+(q)$ dans k^*/k^{*2} appelé norme spinorielle, cf. Exercices 2, 3.

Théorème 9.2.
On suppose $\nu(q) \geq 1$, *soit* $\Delta(q)$ *un discriminant de* q. *On a :*

$$-\mathrm{Id} \in \Omega(q) \Longleftrightarrow n \text{ pair et } \Delta(q) \in k^{*2}.$$

Voir Exercice 3.

Théorème 9.3.
On suppose $n = 3$ *et* $\nu = 1$, *on a* $O^+(q) \simeq PGL(2,k)$ *et* $\Omega(q) \simeq PSL(2,k)$. *En particulier,* $\Omega(q)$ *est simple pour* $k \neq \mathbf{F}_3$.

Ce théorème a été démontré : cf. VII 5.8 et VIII, 4.10, et, pour la simplicité, cf. IV, 4.1.

Théorème 9.4.
On suppose $n = 4$ *et* $\nu = 1$, *soit* $\Delta = \Delta(q)$ *un discriminant de* q.
1) Δ *n'est pas un carré,*
2) On a $\Omega(q) \simeq PSL(2, k[\sqrt{\Delta}])$. *En particulier,* $\Omega(q)$ *est simple.*

Voir VII, § 5 Exercice 2 pour le point 1) et ci-dessous Exercice 4 pour le point 2) dans le cas réel.

Théorème 9.5.
On suppose $n = 4$ *et* $\nu = 2$. *On a* $-\mathrm{Id} \in \Omega(q)$ *et un isomorphisme :*

$$P\Omega(q) \simeq PSL(2,k) \times PSL(2,k).$$

En particulier, $P\Omega(q)$ *n'est pas simple.*

Voir Chapitre VII § 5 Exercice 3.

Enfin, le dernier théorème :

Théorème 9.6.
On suppose $n \geq 5$ *et* $\nu \geq 1$, *alors* $P\Omega(q)$ *est simple.*

Voir Exercice 5 dans le cas du corps des nombres réels.

EXERCICES SUR LE CHAPITRE VIII

Dans tous les exercices, les notations sont celles du § 1, en particulier, sauf mention expresse du contraire, les formes sont supposées non dégénérées.

1, 2, 3. Plans hyperboliques. Espaces hyperboliques.

1) Soit P un plan muni d'une forme quadratique q (non dégénérée), $\Delta(q)$ le discriminant de q.

a) Montrer que P est hyperbolique si et seulement si $-\Delta(q)$ est un carré non nul (cf. 2.3).

b) Si q' est une autre forme sur P, montrer que l'on a :

$$\exists \lambda \in k^*,\ q \sim \lambda q' \iff \Delta(q) = \Delta(q') \text{ dans } k^*/k^{*2}.$$

---

2) On suppose $\nu(q) \geq 1$. Montrer que l'application $q : E \to k$ est surjective (regarder sur un plan hyperbolique). Montrer qu'il existe une base de E formée de vecteurs isotropes (prendre un plan P hyperbolique, une base $e_3, \ldots, e_n$ de $P^\perp$ et modifier les e_i par des vecteurs de P).

---

3) On suppose $n \geq 3$ et $\nu(q) \geq 1$. Soit $u \in O(q)$. Montrer que si u laisse invariantes toutes les droites isotropes, on a $u = \mp\mathrm{Id}$ (prendre une base $e_1, \ldots, e_n$ d'isotropes, cf. Exercice 2, et, si a est anisotrope, distinguer selon qu'il existe ou non deux des e_i non orthogonaux à a).

---

4) On suppose $n \geq 3$ et $\nu(q) \geq 1$. Soit $u \in O(q)$. Montrer que si u laisse invariants tous les plans hyperboliques, on a $u = \mp\mathrm{Id}$. Même question avec les plans de rang 1.

---

5) On suppose $k = \mathbf{R}$ et $\nu(q) \geq 1$. Soit $u \in O(q)$. Montrer que si l'une des hypothèses suivantes est vérifiée, on a $u = \mp\mathrm{Id}$:

a) u laisse invariantes toutes les droites définies positives (i.e. les droites $D = \mathbf{R}a$, avec $q(a) > 0$),

b) *idem* avec les droites définies négatives,

c) Si $\nu \geq 2$, u laisse invariants tous les plans définis positifs (resp. négatifs).

---

6) Soit τ une réflexion orthogonale d'hyperplan H, de droite D et F un sous-espace de E. Montrer que F est stable par τ si et seulement si on a $F \subset H$ ou $F \supset D$.

---

7) Soit F un sous-espace de E stable par tous les éléments de $O(q)$.
Montrer que l'on a $F = \{0\}$ ou $F = E$, sauf si E est un plan hyperbolique sur $\mathbf{F}_3$; étudier ce dernier cas. (On montrera d'abord que si F est non nul, F est non isotrope en utilisant l'exercice 6 et une base orthogonale de E. On considérera ensuite $\mathrm{Id}_F \perp -\mathrm{Id}_{F^\perp}$, cf. 5.1).

---

8) *Une variante de 7)*
On suppose $k \neq \mathbf{F}_3$ et $n \geq 2$. Soit $F \subset E$ un sous-espace non trivial.

a) Montrer qu'il existe $x \in E$, non isotrope, tel que $x \notin F \cup F^\perp$ (après avoir traité les cas où F ou $F^\perp$ sont totalement isotropes on pourra chercher x sous la forme $\lambda y + \mu z$ avec $\lambda, \mu \in k$, $y \in F - F^\perp$ et $z \in F^\perp - F$).

b) On suppose $n \geq 3$ et $\dim F \geq 2$. Montrer qu'il existe $x \in E$, non isotrope, tel que $\tau_x(F) \neq F$ et $\tau_x(F) \cap F \neq \{0\}$.

---

9) *Un lemme de dénombrement.*
Soient $a, b \in k^*$. Montrer que si on a $|k| > 5$, le nombre de couples $(x, y) \in k^2$ solutions de $ax^2 + by^2 = a$ est au moins égal à 6.
Etudier les cas $k = \mathbf{F}_3$, $\mathbf{F}_5$, suivant que $-a/b$ est ou non dans k^{*2}.

---

10) *Généralisation de plusieurs lemmes,* (cf. Exercices 3, 4, 5 et Théorèmes 5.1 et 5.4).
On suppose $n \geq 2$ et $|k| > 5$. Deux sous-espaces F et G de E sont dits de même type si et seulement s'il existe $v \in O(q)$ tel que $G = v(F)$. En vertu du théorème de Witt, ceci revient à dire que l'on a $q|_F \sim q|_G$.
On se propose de prouver l'assertion (A) suivante :

Soit $u \in O(q)$ et F un sous-espace non trivial de E. Si u laisse invariants tous les sous-espaces de même type que F, on a $u = \mp\mathrm{Id}$, sauf si $n = 2$ et si F est une droite isotrope.

a) Prouver (A) pour $n = 2$ et F anisotrope (utiliser l'exercice 9).

b) Prouver (A) pour $n \geq 3$ et $\dim F = 1$ (utiliser les exercices 3 et 9).

c) Montrer (A) dans le cas général par récurrence sur $\dim F$ (prendre $v \in O(q)$ tel que $G = F \cap v(F)$ soit non nul et distinct de F, cf. Exercice 8 b) et lui appliquer la récurrence).

d) On pose pour $F \subset E$:

$$N_F = \{u \in O(q) \mid u(F') = F' \text{ pour tout } F' \text{ du type de } F\}.$$

Montrer que N_F est un sous-groupe distingué de $O(q)$.

e) Calculer N_F lorsque l'on a $n = 2$ et F isotrope.

f) Etudier les cas $k = \mathbf{F}_3, \mathbf{F}_5$. En particulier, calculer N_F lorsque $k = \mathbf{F}_3$ et $n = 2$; $k = \mathbf{F}_5$, $n = 2$ et $q \sim x^2 + y^2$; $k = \mathbf{F}_3$, $n = 3$, $q = x^2 + y^2 + z^2$, $F = <a>$ avec $q(a) = 1$.

---

11) Soient $u \in O(q)$ et H un hyperplan de E, on suppose qu'on a $u|_H = \mathrm{Id}_H$. Montrer que, si H est isotrope, u est l'identité et que, si H n'est pas isotrope, u est l'identité ou la réflexion par rapport à H.

---

4. Le théorème de Witt.

1) Soient F, F' deux sous-espaces de E, $\sigma : F \longrightarrow F'$ une isométrie. Soit $K = \mathrm{Ker}\ q|_F$.

a) Montrer qu'on a $\dim K + \dim F \leq \dim E$ et que l'égalité a lieu si et seulement si $F^\perp$ est totalement isotrope.

b) On suppose $F^\perp$ non totalement isotrope. Montrer qu'il existe $u \in O^+(q)$ (resp. $u \in O^-(q)$) tel que $u|_F = \sigma$ (Théorème de Witt raffiné).

c) On suppose E hyperbolique et F totalement isotrope maximal. Soit $u \in O(q)$ tel que $u(F) = F$. Montrer que l'on a $u \in O^+(q)$ (utiliser 3.3 et écrire une matrice par blocs, voir aussi la démonstration de 7.1).

d) On suppose $\dim F + \dim K = \dim E$, cf. a). Soit $u \in O(q)$, tel que $u|_F = \mathrm{Id}_F$. Montrer que u est dans $O^+(q)$. En déduire que le raffinement du théorème de Witt vu en b) est en défaut dans ce cas.

e) Soient $F \subset E$, $u \in O(q)$ et $F' = u(F)$, montrer qu'il existe $v \in O^+(q)$ tel que $v(F) = F'$, sauf si E est hyperbolique et si F en est un lagrangien.
Si Ω est l'ensemble des lagrangiens de E, montrer que Ω est une orbite sous $O(q)$, mais se décompose en deux orbites sous $O^+(q)$, cf. [D] Ch. II §6 ou [B] tome IV, Ch. 13 §7.

---

2) On suppose $n = 4$, $\nu > 0$. Soit $\Delta(q)$ un discriminant de q.

a) Montrer que l'on a :

q hyperbolique $\Longleftrightarrow \nu = 2 \Longleftrightarrow \Delta(q) \in k^{*2}$, et $\nu = 1 \Longleftrightarrow \Delta(q) \notin k^{*2}$

(cf. §1, 2, 3 Exercice 1 et VII, §5 Exercice 2).

b) Montrer que si q' est une autre forme d'indice $\nu' > 0$, on a,

$$\exists \lambda \in k^*,\ q' \sim \lambda q \Longleftrightarrow \Delta(q) = \Delta(q') \text{ dans } k^*/k^{*2}.$$

5, 7. Générateurs et centres de $O(q)$ et $O^+(q)$; théorème de Cartan-Dieudonné.

1) *Valeurs propres des éléments de $O(q)$.*
Soit $u \in O(q)$.

a) Montrer que si u a un vecteur propre x non isotrope la valeur propre associée est ∓ 1. Que se passe-t-il si x est isotrope ?

b) On pose :

$$F_u = \{x \in E \mid u(x) = x\} = \mathrm{Ker}\ (u - \mathrm{Id}), \quad G_u = \mathrm{Ker}\ (u + \mathrm{Id}).$$

Montrer que si n est impair et dét $u = 1$, on a $F_u \neq 0$, (utiliser 7.1).
Montrer que si n est impair et dét $u = -1$, on a $G_u \neq 0$ (considérer $-u$).
Montrer que si n est pair et dét $u = -1$, on a $F_u \neq 0$ (et donc, si dét $u = -1$, u admet toujours une valeur propre égale à ∓ 1).

c) On suppose k algébriquement clos. Montrer que u n'est pas nécessairement diagonalisable si $n \geq 3$ (chercher, par exemple, avec $q = 2xz + y^2$).

---

2) *Complément sur le théorème de Cartan-Dieudonné.*
Soit $u \in O(q)$, $F_u = \{x \in E \mid u(x) = x\} = \text{Ker}\,(u - \text{Id})$, on pose $p_u = \text{codim}\, F = \text{rg}\,(u - \text{Id})$ et $K_u = \text{Ker}\, q|_{F_u}$. On définit :

$$m_u = \inf\{r \in \mathbf{N} \mid \exists \tau_1, \ldots, \tau_r \text{ réflexions orthogonales avec } u = \tau_1 \ldots \tau_r\}.$$

On convient que m_u vaut zéro si $u = \text{Id}$. Le théorème de Cartan-Dieudonné (7.1) assure que l'on a $m_u \leq n$ et on se propose ici de comparer m_u et p_u (cf. VI, § 2).

a) Montrer que l'on a $p_u \leq m_u$.

b) On dit que u est **exceptionnel** si $F_u^\perp$ est totalement isotrope (cf. aussi § 4 Exercice 1 a)). Montrer que, si u est exceptionnel, on a $p_u < m_u$.

c) On suppose F_u totalement isotrope. Montrer que les conditions suivantes sont équivalentes (s'inspirer de la démonstration de 7.1) :

1) u est exceptionnel,
2) F_u est un lagrangien de E (et donc E est hyperbolique),
3) $\text{Im}\,(u - \text{Id})$ est totalement isotrope,
4) $\forall x \in E$, non isotrope, $u(x) - x$ est isotrope.

Montrer de plus qu'on a alors :

i) dét $u = 1$,
ii) $p_u = \dim F_u = \nu$ est pair (donc n est multiple de 4),

(on écrira la matrice de u dans une base convenable, sous la forme $\begin{pmatrix} I & A \\ 0 & I \end{pmatrix}$, avec A antisymétrique).

d) Déduire de c) que si u est exceptionnel, on a dét $u = 1$ et p_u pair (prendre G_u supplémentaire de K_u dans F_u et appliquer c) à $G_u^\perp$)).

e) Montrer que l'on a $m_u = p_u$ si $n = 2$ ou si $p_u = 1$.

f) On suppose u non exceptionnel, montrer que l'on a $m_u = p_u$ (on raisonnera par récurrence sur le couple (n, p_u) ; on s'inspirera de la démonstration de 7.1 en distinguant 3 cas :

1) F_u non totalement isotrope,
2) F_u totalement isotrope et $\dim F_u^\perp \geq 2 + \dim F_u$,
3) F_u totalement isotrope et $\dim F_u^\perp = 1 + \dim F_u$,

dans ce dernier cas on pourra prouver, en travaillant dans une base appropriée, qu'il existe $x \in E$ non isotrope tel que $x \notin F_u$, $u(x) - x$ non isotrope, et $F_u^\perp \cap <x>^\perp$ non totalement isotrope).

g) On suppose u exceptionnel et F_u totalement isotrope. Soit τ une réflexion quelconque. Montrer que τu n'est pas exceptionnel et vérifie $p_{\tau u} = p_u + 1$. En déduire que l'on a $m_u = p_u + 2$.

h) On suppose u exceptionnel. Montrer que l'on a $m_u = p_u + 2$ (on se ramènera à g) en travaillant dans ${G_u}^\perp$, cf. d)).

---

6. La dimension 2.

1) a) Montrer que l'on a $O(q) \simeq O^+(q) \rtimes \{1,-1\}$. Dans quel cas le produit est-il direct ?

b) Si k est fini et $n = 2$, montrer que $O(q)$ est un groupe diédral que l'on précisera.

---

2) Soit (P,q) un plan de rang 1, $O(q)$ son groupe orthogonal, e_1, e_2 une base de P avec $e_1 \in \operatorname{Ker} q$. Soit $u \in O(q)$. Montrer que la matrice de u dans la base (e_1,e_2) est de la forme :

$$u = \begin{pmatrix} a & b \\ 0 & \varepsilon \end{pmatrix} \quad \text{avec } a \in k^*,\ b \in k,\ \varepsilon = \mp 1.$$

Montrer que l'application $u \mapsto (a,\varepsilon)$ de $O(q)$ dans $k^* \times \{1,-1\}$ est un homomorphisme surjectif dont le noyau est isomorphe à $(k,+)$. En déduire que $O(q)$ est produit semi-direct de $(k,+)$ par $k^* \times \{1,-1\}$. Le produit est-il direct ?

---

3) On suppose $n = 3$. Soit D une droite de E, et $P = D^\perp$. On pose :

$$G = \{u \in O(q) \mid u(P) = P\} = \{u \in O(q) \mid u(D) = D\}.$$

a) On suppose D non isotrope. Montrer que l'on a

$$G \simeq O(P) \times O(D) = O(P) \times \{1,-1\}.$$

b) On suppose D isotrope. Soit $\varphi : G \longrightarrow O(P)$ l'homomorphisme de restriction. Montrer que φ est injectif (cf. § 1, 2, 3 Exercice 11) et surjectif (prendre une base e_1, e_2, e_3 avec $e_1 \in D, e_2 \in P$ et (e_1,e_2) hyperbolique et utiliser l'exercice 2) ; terminer le calcul de G à l'aide de l'exercice 2).

8, 9. Le groupe des commutateurs.

1) Soit (P,q) un plan hyperbolique. Montrer que l'on a $O^+(q)/\Omega(q) \simeq k^*/k^{*2}$.

---

2) On suppose $\nu \geq 1$. Soit $u \in O(q)$. On dit que u est une ***transformation hyperbolique*** s'il existe un plan hyperbolique $P \subset E$ tel que $u|_{P^\perp} = \mathrm{Id}$.

a) Montrer que les réflexions orthogonales sont des transformations hyperboliques (plonger un vecteur non isotrope dans un plan hyperbolique).

b) Soit P un plan hyperbolique et u une transformation hyperbolique de plan P'. Montrer que u s'écrit $u = sv$ avec s hyperbolique de plan P et $v \in \Omega(q)$ (utiliser le théorème de Witt).

c) Soit P un plan hyperbolique et $u \in O(q)$. Montrer que l'on a $u = sv$, avec s et v comme en b).

d) On garde les notations de c). On identifie le groupe $O(P) = O(q|_P)$ à un sous-groupe de $O(q)$ au moyen de l'application $v \mapsto v \perp \mathrm{Id}_{P^\perp}$. Soit $u \in O^+(q)$. On écrit $u = sv$ comme en c). On a $s \in O^+(P)$ et on désigne par $\overline{s}$ la classe de s dans $O^+(P)/O^+(P) \cap \Omega(q)$.

Montrer que l'application ψ qui à u associe $\overline{s}$ est bien définie (i.e. que $\overline{s}$ ne dépend que de u et pas de la décomposition $u = sv$). Montrer que ψ est un homomorphisme surjectif et qu'il induit un isomorphisme :

$$\overline{\psi} : O^+(q)/\Omega(q) \xrightarrow{\sim} O^+(P)/O^+(P) \cap \Omega(q).$$

Montrer que $\Omega(P)$ est inclus dans $O^+(P) \cap \Omega(q)$ et donc que $O^+(q)/\Omega(q)$ est isomorphe à un quotient de $O^+(P)/\Omega(P) \simeq k^*/k^{*2}$ (cf. Exercice 1).

Il semble qu'on ne sache pas prouver directement (i.e. sans la norme spinorielle, cf. Exercice 3) l'égalité $\Omega(P) = O^+(P) \cap \Omega(q)$, qui donnerait l'isomorphisme annoncé en 9.1 (cf. par exemple [D2] § 13, Prop. 12 et Rem. 28).

---

3) *La norme spinorielle.*

Soit $u \in O^+(q)$, on écrit $u = \tau_1 \dots \tau_r$, où $\tau_i = \tau_{x_i}$ est la réflexion définie par le vecteur non isotrope x_i.

a) Montrer que la classe de $q(x_1) \dots q(x_r)$ dans k^*/k^{*2} ne dépend pas du choix de x_i dans la droite de la réflexion τ_i. *On admettra* qu'elle ne dépend, en fait, que de u et non de la décomposition $u = \tau_1 \dots \tau_r$. (La démonstration de ce résultat requiert la construction de l'algèbre de Clifford $C(q)$).

On pose alors $\theta(u) = \overline{q(x_1) \dots q(x_r)} \in k^*/k^{*2}$.

b) Montrer que θ est un homomorphisme de $O^+(q)$ dans k^*/k^{*2}, appelé **norme spinorielle**, et que l'on a $\Omega(q) \subset \operatorname{Ker} \theta$.

c) On suppose $\nu \geq 1$. Montrer que θ est surjective (cf. § 1 Exercice 2).

d) Soit (P, q) un plan hyperbolique, muni d'une base hyperbolique (e_1, e_2). Soit $u \in O^+(q)$, de matrice $\begin{pmatrix} \alpha & 0 \\ 0 & 1/\alpha \end{pmatrix}$ dans cette base. Montrer que l'on a $\theta(u) = \overline{\alpha}$ dans k^*/k^{*2}. En déduire l'égalité $\operatorname{Ker} \theta = \Omega(q)$ dans ce cas.

e) Montrer que, si $\nu \geq 1$, on a $\operatorname{Ker} \theta = \Omega(q)$. (Utiliser d) et Exercice 2.c), cf. 9.1).

f) On suppose $\nu \geq 1$. Soit $\Delta(q)$ un discriminant de q. Montrer que l'on a $-\mathrm{Id} \in \Omega(q) \iff n$ pair et $\Delta(q) \in k^{*2}$ (décomposer $-\mathrm{Id}$ en produit de réflexions à l'aide d'une base orthogonale et utiliser e), cf. 9.2).

---

4) *PSL(2,C) et le groupe de Lorentz.*

Soit E l'espace des matrices hermitiennes 2×2 complexes :

$$E = \{m \in \mathrm{M}(2, \mathbf{C}) \mid {}^t\overline{m} = m\}.$$

On identifie $\mathbf{R}^4$ et E au moyen de l'application :

$$(t,\ x,\ y,\ z) \mapsto \begin{pmatrix} t+x & y+iz \\ y-iz & t-x \end{pmatrix}.$$

On pose, pour $m \in E$, $q(m) = \det m$. Si m est identifié à (t, x, y, z) comme ci-dessus on a $q(m) = t^2 - x^2 - y^2 - z^2$ (le déterminant est donc une forme de Lorentz).

Soient $V = \{m \in E \mid q(m) = 1\}$, $V^+ = \{m \in V \mid t > 0\}$, $V^- = \{m \in V \mid t < 0\}$.

a) Montrer que l'on a $V = V^+ \cup V^-$ et que V^+ et V^- sont connexes. En déduire que si u est dans $O(q)$, on a $u(V^+) = V^+$ ou V^-, puis, que si u est dans $\Omega(q)$ on a $u(V^+) = V^+$.

b) Soit $s \in SL(2,\mathbf{C})$. On définit, pour $m \in E$, $\sigma(s)(m) = sm^t\overline{s}$. Montrer que $\sigma(s)(m)$ est dans E et que σ est un homomorphisme de groupes de $SL(2,\mathbf{C})$ dans $O(q)$. Calculer Ker σ.

c) On rappelle qu'on a $D(SL(2,\mathbf{C})) = SL(2,\mathbf{C})$ (cf. IV, 3.1). En déduire l'inclusion $\mathrm{Im}\,\sigma \subset \Omega(q)$.

d) Soit $e \in E$ la matrice identité, qui est dans V^+, et soit $m \in V^+$. Montrer qu'il existe $s \in SL(2,\mathbf{C})$ tel que $\sigma(s)(m) = e$ (chercher s sous forme triangulaire).

e) Soit $G = \{u \in \Omega(q) \,|\, u(e) = e\}$. Montrer que l'on a $G = O^+(q|_H) \simeq O^+(3,\mathbf{R})$ où l'on a posé $H = <e>^\perp$.
Montrer que si s appartient à $SU(2,\mathbf{C})$, on a $\sigma(s) \in G$. Déterminer les s de $SU(2,\mathbf{C})$ qui vérifient $s^2 = \mp\mathrm{Id}$. Pour un tel s, montrer que l'on a $-is \in H$ et que $\sigma(s)|_H$ est le renversement d'axe $-is$.

f) Déduire de d) et e) que σ est surjectif, donc qu'on a un isomorphisme $\Omega(q) \simeq PSL(2,\mathbf{C})$ (cf. 9.4).

---

5) *La simplicité de $P\Omega(q)$ pour $k = \mathbf{R}$, $n \geq 5$, $\nu \geq 1$.*
On suppose $k = \mathbf{R}$, $n \geq 5$, $\nu \geq 1$ et, précisément, q de signature (r,s) avec avec $r \geq s \geq 1$. Soit N un sous-groupe distingué de $\Omega(q)$ contenant strictement le centre de $\Omega(q)$.

a) Soient P, P' deux plans hyperboliques de E. Montrer qu'il existe $u \in \Omega(q)$ tel que $u(P) = P'$ (utiliser Witt et Exercice 2.c)).

b) Soit $P \subset E$ un plan non totalement isotrope. Montrer qu'il existe un sous-espace $U \subset E$, non isotrope, de dimension 3 et d'indice 1 tel que $P \subset U$.

Un tel sous-espace U sera dit un L–sous-espace (la restriction de q à U est équivalente au signe près à la forme de Lorentz $x^2 + y^2 - z^2$).

c) Soit U un L–sous-espace. On identifie $\Omega(U) = \Omega(q|_U)$ à un sous-groupe de $\Omega(q)$ par l'application $u \mapsto u \perp \mathrm{Id}_{U^\perp}$. On suppose alors $N \cap \Omega(U) \neq \{\mathrm{Id}\}$.
Montrer que l'on a $\Omega(U) \subset N$ (cf. 9.3).

d) Soit U' un autre L–sous-espace. Montrer que sous l'hypothèse de c), on a aussi $\Omega(U') \cap N \neq \{\mathrm{Id}\}$, donc $\Omega(U') \subset N$ (utiliser a), attention $q|_U$ et $q|_{U'}$ ne sont équivalentes qu'au signe près).

e) Déduire de c) et d) que si, pour un L–sous-espace U, on a $N \cap \Omega(U) \neq \{\mathrm{Id}\}$, alors on a $N = \Omega(q)$ (considérer des générateurs $(st)^2$ de $\Omega(q)$, où s,t sont des réflexions et utiliser b)).

f) Montrer que si on a $\nu = 1$ (resp. $\nu \geq 2$), et si P est un plan hyperbolique (resp. un plan défini négatif), tout sous-espace de dimension $n-4$ de $P^\perp$ contient un vecteur a avec $q(a) > 0$.

g) Montrer qu'il existe $u \in N$, avec $u \neq \mp\mathrm{Id}$ et $a \in E$ tels que $q(a) > 0$ et $u(a) = a$ (partir d'un plan P comme en f) et de $s = (\tau_b\tau_c)^2$, pour $b,c \in P$ et fabriquer le commutateur de s et d'un élément non trivial de N, cf. aussi § 1, 2 ,3 Exercices 4 et 5).

h) Déduire des questions g) et b) qu'il existe un L–sous-espace U tel que $\Omega(U) \cap N \neq \{\mathrm{Id}\}$ (prendre $b \in E$ tel que $q(a) = q(b)$ et $u(b)$ non colinéaire à b; utiliser ensuite un commutateur de u et de $(\tau_d\tau_a)^2$, avec $d = a - b$ ou $a + b$).

i) Conclure (cf. 9.6).

BIBLIOGRAPHIE

[A] E. ARTIN, Algèbre géométrique, Gauthier-Villars, Paris (1978).

[AH] J.DIEUDONNÉ *et al.*, Abrégé d'histoire des mathématiques, Hermann, Paris (1978).

[B] M. BERGER, Géométrie, Cedic - Fernand Nathan, Paris (1977).

Tome 1, Action de groupes, espaces affines et projectifs.

Tome 2, Espaces euclidiens, triangles, cercles et sphères.

Tome 4, Formes quadratiques, quadriques et coniques.

[Bl] A. BLANCHARD, Les corps non commutatifs, PUF, Paris (1972).

[Bbki] N. BOURBAKI, Algèbre Chapitres 1,2,3, Hermann, Paris nouvelle édition (1970).

Algèbre Ch. VI et VII, Hermann, Paris (1964).

[C] H. CARTAN, Théorie élémentaire des fonctions analytiques d'une ou plusieurs variables complexes, Hermann, Paris (1961).

[Ca] J.-C. CARRÉGA, Théorie des corps, la règle et le compas, Hermann, Paris (1981)

[D] J. DIEUDONNÉ, La géométrie des groupes classiques, Springer, Berlin (1971).

[D1] J. DIEUDONNÉ, Algèbre lineaire et géométrie élémentaire, Hermann, Paris (1964).

[D2] J. DIEUDONNÉ, Sur les groupes classiques, Hermann, Paris (1948).

[D3] J. DIEUDONNÉ, Les isomorphismes exceptionnels entre les groupes classiques finis, Canad. J. of math. 6, 305 - 315 (1954).

[FAM] J. DIEUDONNÉ, Fondements de l'analyse moderne, Gauthier-Villars, Paris (1963).

[Fr] J. FRENKEL, Géométrie pour l'élève professeur, Hermann, Paris (1973).

[Fu] W. FULTON, Algebraic curves, Benjamin, New-York (1969).

[FG] S. FRANCINOU, H. GIANELLA, Exercices de mathématiques pour l'agrégation, Algèbre 1, Masson, Paris (1994).

[G] C. GODBILLON, Éléments de topologie algébrique, Hermann, Paris (1971).

[H] M. HALL Jr, The theory of groups, Mac Millan, New-York (1959).

[J] N. JACOBSON, Basic algebra, 2 vol., Freeman, San Francisco (1974, 1980).

[K] M. KAROUBI, K-theory, an introduction, Springer, Berlin (1978).

[L] S. LANG, Algebra, Addison-Wesley, New-York (1965).

[LFA] J. LELONG-FERRAND, J.-M. ARNAUDIÈS, Cours de mathématiques, tome 2, analyse, Dunod, Paris (1972).

[P] D. PERRIN, Géométrie algébrique, une introduction, Interéditions, Paris, 1995.

[PR] D. PERRIN, A. ROBERT, Géométrie pour l'écrit du CAPES, Ellipses, Paris (en préparation).

[Pu] L. PUIG, La classification des groupes finis simples : bref apercu et quelques conséquences internes, Séminaire Bourbaki, Exp. 584. Nov. 1981.

[R] W. RUDIN, Real and complex analysis, Mc Graw and Hill, New York (1966).

[Sa] P. SAMUEL, Théorie algébrique des nombres, Hermann, Paris (1967).

[S1] J.-P. SERRE, Cours d'arithmétique, PUF, Paris (1970).

[S2] J.-P. SERRE, Représentations linéaires des groupes finis, Hermann, Paris (1967, 3ème éd.).

[S3] J.-P. SERRE, Groupes finis, cours à l'ENSJF, notes polycopiées, Montrouge (1979).

[St] I. STEWART, Galois theory, Chapmann et Hall, London (1973).

[VdW] B.-L. VAN DER WAERDEN, Modern algebra, Vol. 1, 7ème édition, Springer, Berlin (1966).

[V] M.-F. VIGNÉRAS, Arithmétique des algèbres de quaternions, Springer, Berlin (1980).

INDEX TERMINOLOGIQUE

Aubin Imprimeur
LIGUGÉ, POITIERS

IMPRESSION - FINITION
Achevé d'imprimer en mars 2000
N° d'impression L 59844
Dépôt légal mars 2000
Imprimé en France

Cette collection regroupe des ouvrages variés dont le but est de compléter la formation scientifique des candidats aux concours d'Agrégation et de CAPES de Mathématiques, et éventuellement de leur donner une préparation spécifique à une épreuve ou un type d'épreuve.

Ce volume est directement issu du Cours d'Algèbre paru sous forme photocopiée aux presses de l'École Normale Supérieure de Jeunes Filles et connu des candidats à l'agrégation de mathématiques comme « le Perrin ». Il a permis à de très nombreux agrégatifs de compléter leur formation en algèbre, et d'arriver au concours avec des idées claires. Il s'adresse donc avant tout aux candidats à l'agrégation, mais peut être abordé avec profit dès le début du deuxième cycle de l'enseignement supérieur. Il devrait faire partie de la bibliothèque de base de tout enseignant de mathématiques.

Professeur à l'IUFM de Versailles et à l'Université de Paris-Sud (Orsay), Daniel Perrin s'est occupé pendant quinze ans de la préparation des normaliennes et normaliens à l'agrégation de mathématiques, d'abord à l'École Normale Supérieure de Jeunes Filles, puis à l'École Normale Supérieure.

ISBN 2-7298-5552-1