ALGÈBRE 1 (ENS, PREMIÈRE ANNÉE)

par

Antoine Ducros

Table des matières

1. Relations et quotients

1.1. — Nous ne ferons pour l'essentiel ici aucun rappel sur les notions de base de la théorie des ensembles (appartenance, réunion, intersection…) que nous supposerons connues au moins informellement. Nous allons toutefois passer un peu de temps sur le concept de *relation*, qui vous apparaît sans doute intuitif mais dont nous allons donner une définition rigoureuse.

Définition 1.1.1. — Une *relation* $\mathscr{R}$ est un triplet (X, Y, Γ) où X et Y sont deux ensembles et où Γ est un sous-ensemble de $X \times Y$, appelé le *graphe* de la relation $\mathscr{R}$.

On dit qu'un élément x de X et un élément y de Y sont en relation relativement à $\mathscr{R}$ si $(x, y) \in \Gamma$; on écrit alors $x\mathscr{R}y$.

Si X est un ensemble, une relation *sur* X est une relation de la forme (X, X, Γ).

Commentaires 1.1.2. — On voit que formellement, une relation est définie par son graphe, c'est-à-dire par la liste des couples d'éléments en relation. Mais en pratique, une relation est évidemment le plus souvent définie par une condition logique ou une formule.

Exemples 1.1.3. — Sur tout ensemble X on dispose de la relation $=$ (l'égalité) dont le graphe est la *diagonale* $\{(x, x)\}_{x\in X}$.

Sur l'ensemble $\mathbf{N}$ des entiers naturels, on dispose de la relation $\leqslant$ (l'ordre usuel) dont le graphe est $\{(x, y) \in \mathbf{N}^2, x \leqslant y\}$.

On peut aussi bien entendu se donner une relation directement par son graphe : par exemple, la relation sur $\{1, 2, 3\}$ de graphe $\{(1, 1), (2, 3), (3, 2), (1, 3)\}$ (choisi arbitrairement, et qui ne présente *a priori* strictement aucun intérêt).

1.2. Digression : la notion d'application. — Lorsqu'on écrit rigoureusement les fondements de la théorie des ensembles (et donc des mathématiques), la notion d'application est définie à partir de la notion de relation.

Définition 1.2.1. — Soient X et Y deux ensembles. Une *application* de X vers Y est une relation f de la forme (X, Y, Γ) possédant la propriété suivante : pour tout $x \in X$, il existe un unique $y \in Y$ tel que $(x, y) \in \Gamma$, c'est-à-dire encore tel que xfy. Cet unique élément y est le plus souvent noté $f(x)$. On dit que X est l'ensemble de départ (ou la source) de f, et que Y est son ensemble d'arrivée (ou son but).

Commentaires 1.2.2. — Une application est donc définie en théorie par son graphe, c'est-à-dire par la liste de ses valeurs ; en pratique, elle l'est évidemment le plus souvent par une formule.

Exemple 1.2.3. — Sur tout ensemble X, la relation d'égalité de graphe $\{(x, x)\}_{x\in X}$ est une application, appelée *identité* de X et notée Id_X ou Id si le contexte est clair ; on a $\mathrm{Id}_X(x) = x$ pour tout $x \in X$.

1.2.4. *Application vide.* — La définition d'une application par son graphe permet notamment de donner un sens rigoureux à la notion d'application de source *vide*, qui peut à première vue susciter un certain malaise. En effet, soit Y un ensemble quelconque. Une application de $\varnothing$ vers Y est d'après notre définition un sous-ensemble Γ de $\varnothing \times Y$ tel que

$$\forall x \in \varnothing, \ \ \exists! y \in Y \quad \text{t.q.} \ \ (x, y) \in \Gamma.$$

Regardons de plus près ce que cela signifie. L'ensemble $\varnothing \times Y$ est vide ; son seul sous-ensemble est donc l'ensemble vide. Or la phrase logique ci-dessus est satisfaite lorsque $\Gamma = \varnothing$, pour une raison très simple : *toute phrase logique commençant par* $\forall x \in \varnothing$ *est vraie.* En effet, il suffit pour s'en convaincre de remarquer que sa négation est de la forme «$\exists x \in \varnothing$ t.q. ...» et est donc évidemment fausse.

Il y a donc une et une seule application de $\varnothing$ vers Y, qu'on appelle *l'application vide.*

1.2.5. — Il y a par contre très peu d'applications de but vide. Plus précisément, si X est un ensemble non vide, alors il n'existe pas d'application f de X vers $\varnothing$: en effet si c'était le cas le choix d'un élément x de X (possible car $X \neq \varnothing$) fournirait un élément $f(x)$ de $\varnothing$, ce qui est absurde.

Notez par contre qu'il existe une (unique) application de $\varnothing$ dans lui-même : l'application vide vue au 1.2.4 ci-dessus (où Y était quelconque, et pouvait donc être vide). Elle coïncide nécessairement avec l'identité, par unicité ou encore parce que si on la note f on a alors $\forall x \in \varnothing, \ \ f(x) = x$ par le principe général rappelé en 1.2.4.

1.2.6. *Remarques générales à propos de l'ensemble vide.* — Dans ce texte, tout ensemble figurant dans les hypothèses d'un énoncé est autorisé à être vide *sauf exclusion expresse de ce cas.*

Lorsqu'elle est avérée, la validité d'un énoncé donné pour l'ensemble vide est sauf exception tautologique, et ne demande souvent aucun argument spécifique, grâce au fait qu'une proposition commençant par $\forall x \in \varnothing$ est toujours vraie.

Ce principe qui peut apparaître loufoque ou déstabilisant simplifie donc en fait la vie : il permet en effet d'inclure sans même y penser le cas de l'ensemble vide dans une bonne partie des raisonnements, évitant par là des distinctions aussi fastidieuses qu'inutiles qui alourdiraient la rédaction.

1.2.7. *Familles.* — Une *famille* $(x_i)_{i \in I}$ d'éléments d'un ensemble X paramétrée par un ensemble I est simplement une application $i \mapsto x_i$ de I dans X. Conformément à nos conventions générales rappelées ci-dessus, l'ensemble I peut être vide. Il y a plus précisément une et une seule famille d'éléments de X indexée par l'ensemble vide, à savoir l'application vide de $\varnothing$ dans X, qui est appelée aussi la *famille vide* d'éléments de X.

Si n est un entier, nous écrirons parfois $(e_1, \ldots, e_n)$ pour désigner une famille (e_i) paramétrée par $\{i \in \mathbf{N} \text{ t.q. } 1 \leqslant i \leqslant n\}$. Cette expression garde un sens si $n = 0$: l'ensemble $\{i \in \mathbf{N} \text{ t.q. } 1 \leqslant i \leqslant n\}$ étant alors vide, $(e_1, \ldots, e_n)$ désigne dans ce cas la famille vide.

1.3. Relations d'équivalence. — Nous allons maintenant introduire et étudier une classe particulière de relations qui joue un rôle absolument central en mathématiques.

Définition 1.3.1. — Soit X un ensemble. Une relation $\mathscr{R}$ sur X est appelée une *relation d'équivalence* si elle possède les propriétés suivantes :

- $\diamond$ $\mathscr{R}$ est *réflexive*, *i.e.* $x\mathscr{R}x$ pour tout $x \in X$;
- $\diamond$ $\mathscr{R}$ est *symétrique*, *i.e.* $x\mathscr{R}y \iff y\mathscr{R}x$ pour tout $(x, y) \in X^2$;
- $\diamond$ $\mathscr{R}$ est *transitive*, *i.e.* $(x\mathscr{R}y \text{ et } y\mathscr{R}z) \Rightarrow x\mathscr{R}z$ pour tout $(x, y, z) \in X^3$.

Exemples 1.3.2. — Sur tout ensemble X on a deux exemples extrêmes de relations d'équivalence, à savoir l'égalité d'une part, et la relation grossière $\mathscr{R}$ telle que $x\mathscr{R}y$ pour tout couple (x, y) (notez que ces deux relations coïncident si $X = \varnothing$ ou si X est un singleton).

Si n est un entier, la relation $\mathscr{C}_n$ de congruence modulo n sur $\mathbf{Z}$, définie par la condition $x\mathscr{C}_n y \iff (n \text{ divise } x - y)$ est une relation d'équivalence.

Si $f : X \to Y$ est une application quelconque entre deux ensembles X et Y, elle induit une relation d'équivalence $\mathscr{R}_f$, donnée par la condition

$$x\mathscr{R}_f x' \iff f(x) = f(x').$$

et appelée la relation d'équivalence définie par f. Nous verrons en fait un peu plus bas que toute relation d'équivalence est de la forme $\mathscr{R}_f$ pour une application f bien choisie.

1.3.3. *Classes d'équivalence.* — Soit X un ensemble et soit $\mathscr{R}$ une relation d'équivalence sur X. Si x est un elément de X, la *classe (d'équivalence) de x pour la relation $\mathscr{R}$* est l'ensemble des éléments y de X tels que $x\mathscr{R}y$.

1.3.3.1. — Soit $x \in X$ et soit C sa classe d'équivalence. Par réflexivité, C contient x (en particulier, C est non vide). Soit $y \in C$ et soit $z \in X$. Comme $x\mathscr{R}y$ on a par symétrie et transitivité

$$x\mathscr{R}z \iff y\mathscr{R}z.$$

Par conséquent, C est aussi la classe de y.

On a donc montré que *toute classe d'équivalence de la relation $\mathscr{R}$ est la classe de chacun de ses éléments.*

1.3.3.2. — Soient C et C' deux classes d'équivalence pour $\mathscr{R}$. Supposons que $C \cap C'$ soit non vide. Choisissons un élément x de cette intersection. On déduit de 1.3.3.1 que la classe de x est égale à C aussi bien qu'à C'. Par conséquent $C = C'$.

Deux classes d'équivalences distinctes de $\mathscr{R}$ sont donc disjointes.

1.3.4. *Relations d'équivalences et partitions.* — Soit X un ensemble ; nous noterons $\mathscr{P}(X)$ l'ensemble des parties de X. Une *partition* de X est un sous-ensemble de $\mathscr{P}(X)$ constitué de parties non vides, deux à deux disjointes, et dont la réunion est égale à X.

Soit $\mathscr{R}$ une relation d'équivalence sur X. L'ensemble des classes d'équivalence pour $\mathscr{R}$ est alors une partition de X. En effet toute classe d'équivalence est non vide

(1.3.3.1), les classes d'équivalence sont deux à deux disjointes (1.3.3.2) et leur réunion est égale à X car chaque élément de X appartient à sa propre classe.

Réciproquement, soit P une partition de X. Notons $\mathscr{R}$ la relation définie par la condition suivante : $x\mathscr{R}y$ si et seulement si x et y appartiennent au même élément de la partition P. Il est immédiat que $\mathscr{R}$ est une relation d'équivalence.

Nous laissons le lecteur vérifier que les deux constructions ci-dessus mettent en bijection l'ensemble des relations d'équivalences sur X et celui des partitions de X.

1.3.5. *Quotient par une relation d'équivalence.* — Soit X un ensemble et soit $\mathscr{R}$ une relation d'équivalence sur X. On appelle *quotient de X par $\mathscr{R}$*, et l'on note $X/\mathscr{R}$, l'ensemble des classes d'équivalences pour $\mathscr{R}$. L'application $p\colon X \to X/\mathscr{R}$ qui envoie un élément x sur sa classe d'équivalence est appelée *l'application quotient (relative à $\mathscr{R}$).* Par construction, p est surjective et l'on a $p(x) = p(x') \iff x\mathscr{R}x'$; on voit donc que $\mathscr{R}$ peut se décrire comme la relation d'équivalence naturellement associée à une application de source X (en l'occurrence, à l'application p).

***Théorème 1.3.6* (Propriété universelle du quotient)**
Soit X un ensemble, soit $\mathscr{R}$ une relation d'équivalence sur X et soit p l'application quotient de X vers $X/\mathscr{R}$. Soit f une application de X vers un ensemble Y telle que pour tout $(x, x') \in X^2$ on ait l'implication

$$x\mathscr{R}x' \Rightarrow f(x) = f(x').$$

Il existe alors une unique application g de $X/\mathscr{R}$ vers Y telle que $g \circ p = f$, ce qu'on décrit aussi de façon un peu plus imagée en disant que le diagramme

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[d, "p"'] & Y \\ X/\mathscr{R} \arrow[ur, "g"'] & \end{tikzcd}$$

est commutatif, ou commute. Nous dirons que g est l'application induite par f par passage au quotient (par $\mathscr{R}$).

Démonstration. — Commençons par l'unicité. Supposons qu'une telle g existe et soit $c \in X/\mathscr{R}$. Par surjectivité de p il existe $x \in X$ tel que $c = p(x)$; on a alors nécessairement $g(c) = g(p(x)) = f(x)$, si bien que g est uniquement déterminée.

Montrons maintenant l'existence. Soit $c \in X/\mathscr{R}$. Choisissons un antécédent x de c par p (il y en a au moins un par surjectivité de p). L'image de x par f ne dépend alors que de c, et pas de x ; en effet, si x' est un (autre) antécédent de c par p, l'égalité $p(x) = p(x')$ signifie que $x\mathscr{R}x'$, ce qui entraîne par hypothèse que $f(x) = f(x')$. Il est donc licite de poser $g(c) = f(x)$.

On a alors par construction $g(p(x)) = f(x)$ pour tout $x \in X$, et g fait donc bien commuter le diagramme. □

Commentaires 1.3.7. — Conservons les notations du théorème ci-dessus. Nous résumerons l'implication

$$x\mathscr{R}x' \Rightarrow f(x) = f(x')$$

en disant que f est *$\mathscr{R}$-invariante.*

Il est clair que p est $\mathscr{R}$-invariante, et donc que $g \circ p$ est $\mathscr{R}$-invariante pour toute application $g\colon X/\mathscr{R} \to Y$.

On peut donc reformuler comme suit le théorème ci-dessus : *l'application $g \mapsto g \circ p$ établit une bijection entre l'ensemble des applications de $X/\mathscr{R}$ vers Y et l'ensemble des applications $\mathscr{R}$-invariantes de X vers Y.*

On peut résumer la situation de manière un peu informelle par le slogan suivant : *se donner une application de $X/\mathscr{R}$ vers Y, c'est se donner une application $\mathscr{R}$-invariante de X vers Y.*

Ce dernier est à retenir absolument, et son usage doit devenir un réflexe : lorsque vous aurez besoin de construire une application *depuis* un ensemble quotient $X/\mathscr{R}$, vous devrez chercher à construire une application $\mathscr{R}$-invariante de source X (il y a cela dit des exceptions à ce principe, mais elles sont rares ; elles peuvent par exemple se produire si l'on dispose d'une description de $X/\mathscr{R}$ différente de sa présentation comme quotient).

1.3.8. *Comment travailler avec le quotient ?* — Il est fréquent en mathématiques qu'il y ait un certain hiatus entre la définition d'un objet et l'intuition qu'il convient de s'en faire ; cela ne signifie pas que la définition est mauvaise, mais que son rôle est avant tout *technique* (elle assure l'existence d'un objet ayant les propriétés requises, elle permet de raisonner rigoureusement avec celui-ci), et qu'elle ne permet pas de, ou disons ne suffit pas à, comprendre en profondeur ce qu'elle décrit.

C'est typiquement le cas en ce qui concerne les quotients : si l'on se contente de voir $X/\mathscr{R}$ comme un ensemble de classes d'équivalence (ce qu'il est *stricto sensu*), on risque très vite de ne plus rien comprendre à ce qui se passe, les classes d'équivalence étant elles-mêmes des sous-ensembles de X. Il vaut beaucoup mieux penser à $X/\mathscr{R}$ comme à un ensemble construit à partir de X en *décrétant* que deux éléments en relation au sens de $\mathscr{R}$ sont égaux, et en n'imposant *aucune autre contrainte* ; ou encore comme l'ensemble *le plus général* construit à partir de X en imposant que deux éléments en relation au sens de $\mathscr{R}$ soient égaux.

On observe ici un cas particulier d'un phénomène fréquent en mathématiques, qu'on rencontrera à de nombreuses reprises dans ce cours : on traduit une propriété intuitive du type « X est l'objet le plus général qui vérifie la condition C » par la satisfaction d'une propriété universelle qui consiste toujours à décrire une bijection entre un ensemble d'applications de source ou de but X (selon les cas) et un autre ensemble étroitement relié à la condition C. L'expérience a montré la fécondité de cette approche. L'un de ses nombreux mérites est qu'une propriété universelle détermine toujours l'objet qui la vérifie de manière essentiellement unique ; la proposition ci-dessous illustre ce phénomène dans le cas des quotients par les relations d'équivalence.

***Proposition 1.3.9* (Unicité du quotient à bijection unique près)**

Soit X un ensemble, soit $\mathscr{R}$ une relation d'équivalence sur X et soit p l'application quotient de X vers $X/\mathscr{R}$. Soit $q\colon X \to S$ une application $\mathscr{R}$-invariante et telle que pour tout ensemble Y, l'application $g \mapsto g \circ q$ établisse une bijection entre l'ensemble des applications de S vers Y et celui des applications $\mathscr{R}$-invariantes de X

vers Y. Il existe alors une unique application $\varphi\colon X/\mathscr{R} \to S$ telle que le diagramme

$$\begin{tikzcd} X \arrow[r, "q"] \arrow[d, "p"'] & S \\ X/\mathscr{R} \arrow[ur, "\varphi"'] & \end{tikzcd}$$

commute, et φ est une bijection.

Démonstration. — L'existence et l'unicité de φ résultent de la propriété universelle du quotient. Il reste à s'assurer que φ est bijective. En vertu de notre hypothèse sur q, il existe une (unique) application $\psi\colon S \to X/\mathscr{R}$ telle que le diagramme

$$\begin{tikzcd} X \arrow[r, "q"] \arrow[d, "p"'] & S \arrow[dl, "\psi"'] \\ X/\mathscr{R} & \end{tikzcd}$$

commute. On a $\psi\circ\varphi\circ p = \psi\circ q = p$; par injectivité de $g \mapsto g\circ p$ il vient $\psi\circ\varphi = \mathrm{Id}_{X/\mathscr{R}}$. On a également $\varphi\circ\psi\circ q = \varphi\circ p = q$; par injectivité de $g \mapsto g\circ q$ il vient $\varphi\circ\psi = \mathrm{Id}_S$. En conséquence φ est bijective (et son inverse est ψ). □

1.3.10. *Quelques propriétés des applications induites.* — Soit X un ensemble et soit $\mathscr{R}$ une relation d'équivalence sur X ; notons $x \mapsto \overline{x}$ l'application quotient de X vers $X/\mathscr{R}$. Soit $f\colon X \to Y$ une application $\mathscr{R}$-invariante, et soit $\overline{f}\colon X/\mathscr{R} \to Y$ l'application induite (les notations $x \mapsto \overline{x}$ et $\overline{f}$ sont relativement standard).

On a pour tout $x \in X$ l'égalité $f(x) = \overline{f}(\overline{x})$, ce qui entraîne que $\mathrm{Im}(f) \subset \mathrm{Im}(\overline{f})$. Par ailleurs comme $x \mapsto \overline{x}$ est une surjection de X sur $X/\mathscr{R}$, tout élément de $\mathrm{Im}(\overline{f})$ est de la forme $\overline{f}(\overline{x}) = f(x)$ pour un certain $x \in X$. Ainsi $\mathrm{Im}(f) \supset \mathrm{Im}(\overline{f})$, si bien qu'on a finalement $\mathrm{Im}(f) = \mathrm{Im}(\overline{f})$. En particulier, $\overline{f}$ est surjective si et seulement si f est surjective.

Par ailleurs, en vertu de la surjectivité de $x \mapsto \overline{x}$, l'application $\overline{f}$ est injective si et seulement si on a pour tout $(x, y) \in X^2$ l'équivalence

$$\overline{f}(\overline{x}) = \overline{f}(\overline{y}) \iff \overline{x} = \overline{y}$$

qu'on peut récrire

$$f(x) = f(y) \iff x\mathscr{R}y.$$

Autrement dit, $\overline{f}$ est injective si et seulement si $\mathscr{R}$ est la relation d'équivalence définie par f (exemple 1.3.2).

En particulier, si f est une surjection et si la relation d'équivalence associée à f coïncide avec $\mathscr{R}$, l'application $\overline{f}$ est une bijection de $X/\mathscr{R}$ sur Y.

1.3.11. — Soit $f\colon X \to Y$ une application. Si l'on désigne par $\mathscr{R}_f$ la relation d'équivalence associée à f (1.3.2) alors f est $\mathscr{R}_f$-invariante par définition, et il résulte de 1.3.10 que f induit par passage au quotient une bijection $X/\mathscr{R}_f \simeq \mathrm{Im}(f)$.

1.3.12. *Relation produit.* — Soit $(X_i)_{i\in I}$ une famille d'ensembles. On note X le produit $\prod_i X_i$, c'est-à-dire l'ensemble des familles $(x_i)_{i\in I}$ telles que $x_i \in X_i$ pour tout i.

Supposons donné pour tout i une relation d'équivalence $\mathscr{R}_i$ sur X_i, et notons $\xi \mapsto \overline{\xi}$ l'application quotient correspondante (l'indice i ne figure pas explicitement dans cette notation mais cela ne créera pas de confusion). Soit $\mathscr{R}$ le produit des $\mathscr{R}_i$, c'est-à-dire la relation sur X définie par la condition

$$(x_i)_i\mathscr{R}(y_i)_i \iff (\forall i,\ x_i\mathscr{R}_iy_i).$$

On vérifie immédiatement que $\mathscr{R}$ est une relation d'équivalence.

L'application

$$\pi\colon X \to \prod_i X_i/\mathscr{R}_i,\ (x_i)_i \mapsto (\overline{x_i})_i$$

est visiblement surjective, et $\mathscr{R}$ coïncide par construction avec la relation d'équivalence associée à π. On déduit alors de 1.3.10 que π induit par passage au quotient une bijection $X/\mathscr{R} \simeq \prod_i X_i/\mathscr{R}_i$.

1.4. Quotient par une relation arbitraire. — Soit X un ensemble. Si $\mathscr{R}$ est une relation d'équivalence sur X, nous avons vu comment «construire un ensemble à partir de X en décrétant que deux éléments en relation au sens de $\mathscr{R}$ ont égaux, et en n'imposant aucune autre contrainte» : on forme le quotient $X/\mathscr{R}$.

Une question naturelle se pose : si on part maintenant d'une relation $\mathscr{S}$ *quelconque* sur X, peut-on «construire un ensemble à partir de X en décrétant que deux éléments en relation au sens de $\mathscr{S}$ ont égaux, et en n'imposant aucune autre contrainte» ? Techniquement, cela revient (par analogie avec ce qui a été vu dans le cas d'une relation d'équivalence) à poser la question suivante : existe-t-il un ensemble Z et une application $\mathscr{S}$-invariante $p\colon X \to Z$ telle que pour tout ensemble Y, l'application $g \mapsto g \circ p$ établisse une bijection entre l'ensemble des applications de Z dans Y et l'ensemble des applications $\mathscr{S}$-invariantes de X dans Y ? Nous allons voir que la réponse est affirmative, mais cela va nécessiter un peu de travail.

1.4.1. *La relation d'équivalence engendrée par $\mathscr{S}$.* — Convenons de dire qu'une relation $\mathscr{T}$ sur X est *plus fine* qu'une relation $\mathscr{T}'$ (ou que $\mathscr{T}'$ est *plus grossière* que $\mathscr{T}$) si le graphe de $\mathscr{T}$ est contenu dans celui de $\mathscr{T}'$, c'est-à-dire encore si

$$x\mathscr{T}y \Rightarrow x\mathscr{T}'y$$

pour tout couple (x, y) d'éléments de X ; si c'est le cas, on écrira $\mathscr{T} \le \mathscr{T}'$.

Soit maintenant E l'ensemble des relations d'équivalences sur X qui sont plus grossières que $\mathscr{S}$, et soit $\mathscr{R}$ la relation sur X définie par la condition suivante :

$$x\mathscr{R}y \iff \forall \mathscr{T} \in \mathsf{E},\ \ x\mathscr{T}y.$$

On vérifie aussitôt que $\mathscr{R}$ est une relation d'équivalence ; on dit que c'est la *relation d'équivalence engendrée par $\mathscr{S}$*. Notez qu'on a par définition $\mathscr{R} \le \mathscr{T}$ pour toute $\mathscr{T} \in \mathsf{E}$. Notez aussi que si x et y sont deux éléments de X tels que $x\mathscr{S}y$ alors $x\mathscr{T}y$ pour toute $\mathscr{T} \in \mathsf{E}$, si bien que $x\mathscr{R}y$; par conséquent, $\mathscr{S} \le \mathscr{R}$. Remarquons enfin que si $\mathscr{S}$ est déjà une relation d'équivalence elle appartient à E et on vérifie aussitôt que cela entraîne l'égalité $\mathscr{R} = \mathscr{S}$.

1.4.2. *Description plus concrète de $\mathscr{R}$.* — La définition théorique de $\mathscr{R}$ va nous suffire, mais pour la curiosité du lecteur nous allons en donner une description plus tangible.

Soit $\mathscr{S}'$ la relation sur X définie par la condition suivante : $x\mathscr{S}'y$ si et seulement s'il existe un entier $n \geqslant 1$ et une suite $x = x_1, x_2, \ldots, x_n = y$ telle que pour tout i vérifiant $1 \leqslant i \leqslant n-1$ on ait $x_i\mathscr{S}x_{i+1}$ ou $x_{i+1}\mathscr{S}x_i$. On a alors $\mathscr{R} = \mathscr{S}'$; nous allons brièvement expliquer pourquoi.

On vérifie tout d'abord immédiatement que $\mathscr{S} \leq \mathscr{S}'$ et que $\mathscr{S}'$ est une relation d'équivalence (la transitivité et la symétrie découlent de la définition ; pour la réflexivité, notez que si $x \in X$ la suite singleton $x_1 = x$ montre que $x\mathscr{S}'x$, car la condition à vérifier commence par $\forall i \in \varnothing$). Autrement dit, $\mathscr{S}' \in \mathsf{E}$, ce qui entraîne que $\mathscr{R} \leq \mathscr{S}'$.

Vérifions réciproquement que $\mathscr{S}' \leq \mathscr{R}$, ce qui montrera que $\mathscr{R} = \mathscr{S}'$. Soient donc x et y deux éléments de X tels que $x\mathscr{S}'y$, et donnons-nous une suite $x = x_1, \ldots, x_n = y$ comme ci-dessus. Soit $\mathscr{T} \in \mathsf{E}$. Pour tout i entre 1 et $n-1$ on a $x_i\mathscr{S}x_{i+1}$ ou $x_{i+1}\mathscr{S}x_i$ et donc $x_i\mathscr{T}x_{i+1}$ puisque $\mathscr{T}$ est une relation d'équivalence (et en particulier symétrique) plus grossière que $\mathscr{S}$; par transitivité il vient $x\mathscr{T}y$. Ceci valant pour tout $\mathscr{T} \in \mathsf{E}$ on a $x\mathscr{R}y$, ce qu'il fallait démontrer.

1.4.3. — Nous allons maintenant montrer que l'application $p\colon X \to X/\mathscr{R}$ répond au problème posé au début de 1.4. On sait déjà que $\mathscr{S} \leq \mathscr{R}$, ce qui signifie exactement que l'application p est $\mathscr{S}$-invariante. Il s'agit maintenant de montrer que pour tout ensemble Y, l'application $g \mapsto g \circ p$ établit une bijection entre l'ensemble des applications de $X/\mathscr{R}$ dans Y et l'ensemble des applications $\mathscr{S}$-invariantes de X dans Y. Compte-tenu de la propriété universelle du quotient (appliquée à la relation $\mathscr{R}$), il suffit de prouver que pour tout ensemble Y et toute application $f\colon X \to Y$, l'application f est $\mathscr{S}$-invariante si et seulement si elle est $\mathscr{R}$-invariante.

Comme $\mathscr{S} \leq \mathscr{R}$, toute application $\mathscr{R}$-invariante de source X est $\mathscr{S}$-invariante. Réciproquement, soit f une application $\mathscr{S}$-invariante de source X et soit $\mathscr{R}_f$ la relation d'équivalence associée à f. Dire que f est $\mathscr{S}$-invariante signifie exactement que $\mathscr{S} \leq \mathscr{R}_f$, c'est-à-dire encore que $\mathscr{R}_f \in \mathsf{E}$. Mais on a alors $\mathscr{R} \leq \mathscr{R}_f$, et f est donc $\mathscr{R}$-invariante.

Commentaires 1.4.4. — L'ensemble $X/\mathscr{R}$ peut donc s'interpréter comme l'ensemble construit à partir de X en décrétant que deux éléments en relation au sens de $\mathscr{S}$ sont égaux (et en n'imposant aucune autre contrainte). Mais on sait par ailleurs que $X/\mathscr{R}$ peut s'interpréter comme l'ensemble construit à partir de X en décrétant que deux éléments en relation au sens de $\mathscr{R}$ sont égaux (et en n'imposant aucune autre contrainte).

Intuitivement, cette double interprétation a le sens suivant. Lorsqu'on décrète que deux éléments coïncident dès qu'ils sont en relation au sens de $\mathscr{S}$, cela entraîne des *dommages collatéraux* (en raison des propriétés formelles de la relation d'égalité) et force en fait deux éléments à coïncider dès qu'ils sont en relation au sens de $\mathscr{R}$. (Bien sûr si $\mathscr{S}$ est déjà une relation d'équivalence, $\mathscr{R} = \mathscr{S}$, il n'y a pas de vrais dommages collatéraux et les seules identifications qu'on obtient sont celles qu'on a imposées ; mais si $\mathscr{S}$ n'est pas une relation d'équivalence, la relation $\mathscr{R}$ est strictement plus grossière que $\mathscr{S}$ et des identifications apparaissent qui n'étaient pas explicitement spécifiées.)

2. Généralités sur les groupes

2.1. Définitions et propriétés de base. — Nous allons maintenant présenter la notion de groupe, dont l'étude occupera une grande partie de ce cours.

Définition 2.1.1. — Un *groupe* est un ensemble G *muni* d'une loi de composition interne $*: G \times G \to G$ qui satisfait les propriétés suivantes.

(i) La loi $*$ est *associative* : pour tout $(g, g', g'') \in G^3$ on a $g*(g'*g'') = (g*g')*g''$.

(ii) La loi $*$ admet un *élément neutre*, c'est-à-dire un élément e tel que $e*g = g*e = g$ pour tout $g \in G$ (un tel e est nécessairement unique, *cf.* ci-dessous).

(iii) Tout élément g de G admet un *symétrique* pour la loi $*$, c'est-à-dire un élément g' de G tel que $g * g' = g' * g = e$ (un tel g' est nécessairement unique, *cf.* ci-dessous).

Commentaires 2.1.2. — Insistons sur le sens de l'adjectif «muni» : il signifie que la loi $*$ *fait partie des données*. En toute rigueur, on devrait donc écrire «soit $(G, *)$» un groupe et non «soit G un groupe». Bien entendu, respecter ce principe conduirait à alourdir épouvantablement la rédaction, et l'on s'en affranchit donc le plus souvent ; mais il faut garder en tête que l'on commet un petit abus, pour les rares cas où il pourrait y avoir une ambiguïté sur la loi de groupe.

2.1.3. *Justification des énoncés d'unicité dans la définition 2.1.1.* — Si e et e' sont deux éléments neutres dans G on a alors $e * e' = e$ car e' est neutre, et $e * e' = e'$ car e est neutre ; ainsi, $e = e'$.

Si g' et g'' sont deux symétriques d'un même élément g de G on a

$$g' = g' * e = g' * (g * g'') = (g' * g) * g'' = e * g'' = g''$$

et partant $g' = g''$ (notez qu'on utilise ici l'associativité).

Remarque 2.1.4. — Un groupe est toujours *non vide*, puisqu'il possède un élément neutre.

Notations 2.1.5. — Lorsqu'on écrira «soit G un groupe» sans mention explicite de sa loi interne, celle-ci n'aura droit le plus souvent à aucune symbole spécifique et sera simplement notée $(g, h) \mapsto gh$; en général, on désignera par e l'élément neutre de G (s'il y a plusieurs groupes en jeu, il arrivera qu'on le note e_G pour éviter toute confusion). Si $(g, h) \in G^2$, on parlera de gh comme du *produit* de g et h et l'on notera g^{-1} le symétrique de g, qu'on appellera également son *inverse*. On a les formules $(g^{-1})^{-1} = g$ et $(gh)^{-1} = h^{-1}g^{-1}$ (attention au renversement de l'ordre des facteurs).

Comme la loi interne de G est associative, on peut définir le produit de toute famille finie *ordonnée* d'éléments de G (sans avoir à spécifier un parenthésage) ; lorsque la famille est *vide*, ce produit est égal à e.

On écrira souvent $g_1 \ldots, g_n$ pour désigner le produit de la famille ordonnée $(g_i)_{1 \leqslant i \leqslant n}$ d'éléments de G. Lorsque $n = 0$ cette famille est vide et l'on a donc $g_1 \cdots g_n = e$.

Si $g \in G$ et si $n \geqslant 0$ on posera

$$g^n = \underbrace{g \dots g}_{n \text{ facteurs}}$$

(si $n = 0$ on a donc $g^n = e$). Si $n < 0$ on posera $g^n = (g^{-n})^{-1}$. Ces définitions assurent la validité des formules usuelles $g^{n+m} = g^n g^m$ et $(g^n)^m = g^{nm}$.

Si g et h sont deux éléments de G et si E est un sous-ensemble de G on désignera par gE (resp. Eh, resp. gEh) l'ensemble des éléments de G de la forme gx (resp. xh, resp. gxh) avec $x \in E$. On a évidemment $eE = Ee = eEe = E$. Si g' et h' sont deux autres éléments de G, l'associativité de G entraîne les égalités

$$(g'g)E = g'(gE), \quad E(h)h' = E(hh') \text{ et } g'(gEh)h' = (g'g)E(hh').$$

2.1.6. *Simplification.* — Dans un groupe, on peut «simplifier à gauche à droite». Plus précisément, soit G un groupe et soient g, g' et h trois éléments de G. Si $hg = hg'$ alors $g = g'$ (multiplier à gauche par h^{-1}) ; si $gh = g'h$ alors $g = g'$ (multiplier à droite par h^{-1}).

Mentionnons quelques cas particuliers que nous utiliserons systématiquement :

(i) si $hg = g$ $(= eg)$ alors $h = e$;

(ii) si $gh = g$ $(= ge)$ alors $h = e$;

(iii) si $gh = e$ $(= gg^{-1})$ alors $h = g^{-1}$ (et donc $h^{-1} = g$).

2.1.7. *Groupes abéliens.* — Soit G un groupe. On dit qu'il est *commutatif*, ou *abélien*, si $gh = hg$ pour tout $(g, h) \in G^2$. Si c'est le cas, on adopte parfois pour G la notation *additive* : la loi interne est notée $+$, l'élément neutre 0, le symétrique d'un élément g est noté $-g$ et parfois appelé son *opposé*, et l'on écrit ng au lieu de g^n.

***Exemple 2.1.8* (Le groupe trivial).** — Le singleton $\{e\}$ muni de la seule loi interne possible (celle pour laquelle $ee = e$) est un groupe, qui est dit *trivial.*

Exemple 2.1.9. — L'ensemble $\mathbf{Z}$ muni de l'addition est un groupe abélien, sur lequel nous reviendrons en détail au chapitre suivant.

***Exemple 2.1.10* (Groupes de permutations).** — Soit X un ensemble. L'ensemble $\mathfrak{S}_X$ des bijections de X dans lui-même (qu'on appelle parfois aussi *permutations de* X), muni de la composition des applications, est un groupe.

Nous aurons l'occasion de revenir longuement sur le groupe $\mathfrak{S}_X$ lorsque X est fini. Indiquons simplement ici que si $|X| \geqslant 3$ alors $\mathfrak{S}_X$ est non abélien. En effet, choisissons trois éléments a, b et c de X, deux à deux distincts. Soit σ la permutation de X qui échange a et b et fixe tous les autres éléments de X, et soit τ celle qui échange a et c et fixe tous les autres éléments de X. Alors $\sigma \circ \tau \neq \tau \circ \sigma$: la première envoie a sur c, la seconde l'envoie sur b.

***Exemple 2.1.11* (Produits de groupes).** — Soient G et H deux groupes. L'ensemble $G \times H$ muni de la loi interne définie par la formule

$$(g, h)(g', h') = (gg', hh')$$

est un groupe. Son élément neutre est (e_G, e_H), et l'inverse d'un élément (g, h) est (g^{-1}, h^{-1}). On l'appelle le *produit direct* de G et H et on le note le plus souvent simplement $G \times H$ (on dit parfois simplement «produit» ; l'épithète «direct» est là pour lever toute ambiguïté, car il existe d'autres produits plus généraux, dits *semi-directs*, que nous rencontrerons et étudierons plus loin).

Plus généralement, si $(G_i)_{\in\in I}$ est une famille de groupes, l'ensemble $\prod_i G_i$ muni du produit composante par composante est un groupe (son élément neutre est la famille $(e_i)_{i\in I}$, où i désigne pour tout i l'élément neutre de G_i ; et l'inverse se calcule composante par composante). On l'appelle le produit des G_i, et on le note $\prod_i G_i$.

2.1.12. *Brefs rappels sur les anneaux.* — Un *anneau* est un groupe abélien $(A, +)$ muni d'une loi interne supplémentaire $\times$ qui est associative, possède un élément neutre 1 (nécessairement unique par le même raisonnement qu'en 2.1.3), et est *distributive par rapport à l'addition*, ce qui signifie que

$$a \times (b + c) = (a \times b) + (a \times c) \text{ et } (a + b) \times c = (a \times c) + (b \times c) \text{ pour tout } (a, b, c) \in A^3.$$

Remarque 2.1.13. — Dans les axiomes qui définissent un anneau, *on n'impose pas que* $1 \neq 0$. On démontre très facilement que dans un anneau donné A, on a $1 = 0$ si et seulement si $A = \{0\}$ – on dit alors que A est *l'anneau nul.*

Notation 2.1.14. — Il est fréquent, lorsqu'on travaille dans un anneau, que l'on omette le symbole mutiplicatif $\times$ et qu'on écrive simplement ab au lieu de $a \times b$.

2.1.15. *Anneaux commutatifs.* — On dit qu'un anneau A est *commutatif* si $ab = ba$ pour tout $(a, b) \in A^2$.

Nous supposerons connues les bases de la théorie des anneaux commutatifs, et notamment les notions d'anneau intègre, de corps, d'idéal, d'idéal premier, d'idéal maximal, d'idéal et d'anneau principal. Nous ferons parfois de brefs rappels sur l'un ou l'autre de ces points, mais la plupart du temps nous les utiliserons librement ; nous vous invitons à consulter si nécessaire vos cours antérieurs sur ces questions.

2.1.16. *Groupes des inversibles d'un anneau.* — Soit A un anneau. Un élément a de A est dit *inversible* s'il existe un élément b de A, nécessairement unique par le même raisonnement qu'en 2.1.3, tel que l'on ait $ab = ba = 1$.

L'ensemble des éléments inversibles de A est noté $A^\times$; il est stable sous la multiplication et cette dernière fait de $A^\times$ un groupe d'élément neutre 1, abélien dès que A est commutatif.

Exemple 2.1.17. — Si $A = \{0\}$ alors $A^\times = \{0\} = \{1\}$ (notez que l'anneau nul est le seul anneau dans lequel 0 soit inversible).

Exemple 2.1.18. — Si k est un corps, $k^\times$ est égal à $k\backslash\{0\}$.

Exemple 2.1.19. — Considérons $\mathbf{Z}$ comme un anneau (commutatif) *via* les lois usuelles + et $\times$. On a alors $\mathbf{Z}^\times = \{-1, 1\}$.

Exemple 2.1.20. — Soit k un corps et soit n un entier $\geqslant 0$. L'addition et la multiplication font de l'ensemble de matrices $\mathrm{M}_n(k)$ un anneau, non commutatif dès que $n \geqslant 2$; son groupe des éléments inversibles est traditionnellement noté $\mathrm{GL}_n(k)$, et est appelé le *groupe linéaire* de k.

Modulo l'identification d'une matrice 1×1 à son unique coefficient on a $\mathrm{M}_1(k) = k$ et $\mathrm{GM}_1(k) = k^\times$.

Vous vous demandez peut-être ce qu'il en est de $\mathrm{M}_0(k)$ et $\mathrm{GL}_0(k)$. Pour le comprendre, il faut revenir à la définition d'un élément de $\mathrm{M}_n(k)$: c'est une famille $(a_{ij})_{i,j}$ d'éléments de k indexée par l'ensemble des couples (i, j) d'entiers tels que $1 \leqslant i \leqslant n$ et $1 \leqslant j \leqslant n$. Lorsque $n = 0$, c'est donc une famille indexée par l'ensemble vide. Il s'ensuit que $\mathrm{M}_0(k)$ contient un seul élément (la famille vide). Par conséquent $\mathrm{M}_0(k)$ est l'anneau nul et $\mathrm{GL}_0(k)$ est le groupe trivial.

2.2. Sous-groupes d'un groupe donné. — Lorsqu'on souhaite étudier un ensemble muni d'une structure donnée, il est naturel et utile de s'intéresser à ses sous-ensembles qui «se comportent bien» vis-à-vis de cette structure (dans un sens à préciser selon la nature de celle-ci). Dans le cas d'un groupe, cela conduit à la notion de *sous-groupe.*

Définition 2.2.1. — Soit G un groupe et soit H une partie de G. On dit que H est un *sous-groupe* de G si $e \in H$ et si H est stable par produit et par inversion.

Remarque 2.2.2. — Pour que H soit un sous-groupe de G, il faut et il suffit que H soit non vide et que gh^{-1} appartienne à H pour tout $(g, h) \in H$. C'est en effet clairement nécessaire.

Supposons maintenant que ce soit vérifié, et montrons que H est un sous-groupe de G. Comme H est non vide, il existe $h_0 \in H$. Par hypothèse, $e = h_0 h_0^{-1}$ appartient alors à H. Soient g et h deux éléments de H. Puisque $e \in H$, le produit $eh^{-1} = h^{-1}$ appartient à H, qui est donc stable par inversion. Il s'ensuit que $gh = g(h^{-1})^{-1} \in H$, et H est stable par produit ; ainsi, H est un sous-groupe de G.

2.2.3. — Soit G un groupe et soit H un sous-groupe de G. Comme H est stable par produit, l'application $(h, h') \mapsto hh'$ définit une loi interne sur H. Elle fait de H un groupe : elle hérite en effet de l'associativité, elle admet un élément neutre e (qui appartient à H par hypothèse) et tout élément h de H a un inverse pour cette loi, à savoir h^{-1} (qui appartient à H par hypothèse). On dit que la structure de groupe de H est *héritée* de celle de G.

Il découle alors des définitions qu'une partie de H est un sous-groupe de H si et seulement si c'est un sous-groupe de G.

Exemples 2.2.4 **(Les cas triviaux).** — Si G est un groupe, les ensembles G et $\{e\}$ de G sont des sous-groupes de G.

Exemple 2.2.5. — Soit k un corps et soit E un k-espace vectoriel. Le sous-ensemble de $\mathfrak{S}_E$ formé des bijections k-linéaires de E dans lui-même est un sous-groupe de $\mathfrak{S}_E$.

Exemple 2.2.6. — Soit G un groupe. Si (H_i) est une famille de sous-groupes de G, l'intersection $\bigcap H_i$ est un sous-groupe de G (c'est immédiat).

2.2.7. *Sous-groupe engendré par une partie.* — Soit G un groupe et soit P une partie de G.

2.2.7.1. — D'après l'exemple 2.2.6, l'intersection de tous les sous-groupes de G contenant P est un sous-groupe de G. C'est manifestement le plus petit sous-groupe de G contenant P ; on l'appelle le sous-groupe de G *engendré* par P et on le note souvent $\langle P\rangle$.

2.2.7.2. — La définition de $\langle P\rangle$ peut sembler très théorique et peu tangible, mais nous allons en donner une description plus concrète. Posons $P^{-1} = \{g^{-1}, g \in P\}$. Le sous-groupe $\langle P\rangle$ coïncide alors avec l'ensemble Q des produits finis d'éléments de $P \cup P^{-1}$. Il est en effet clair que $Q \subset \langle P\rangle$ puisque $\langle P\rangle$ est un sous-groupe de G contenant P. Et par ailleurs il découle immédiatement de la définition de Q qu'il contient e (c'est le *produit vide* d'éléments de $P \cup P^{-1}$) et est stable par produit et inversion. Par conséquent, Q est un sous-groupe de G contenant évidemment P ; il vient $\langle P\rangle \subset Q$.

Exemple 2.2.8. — Soit G un groupe. Le sous-groupe de G engendré par la partie *vide* est égal à $\{e\}$: on peut le déduire aussi bien de sa définition directe que de sa description «concrète» donnée au 2.2.7.2

Exemple 2.2.9. — Soit g un élément d'un groupe G. On déduit de 2.2.7.2 que le sous-groupe $\langle g\rangle$ de G est égal à $\{g^n\}_{n\in\mathbf{Z}}$.

2.3. Morphismes de groupes. — Pour étudier les ensembles munis d'une structure d'un type donné, l'expérience a montré qu'il était fondamental de considérer les applications entre deux tels ensembles «compatibles» avec leurs structures, dans un sens à préciser (qui dépend du type de structure considéré). Ainsi, lorsqu'on travaille avec des espaces vectoriels, on s'intéresse aux applications linéaires ; lorsqu'on travaille avec des espaces topologiques, on s'intéresse aux applications continues, etc. Et dans le cas des groupes, on va s'intéresser aux *morphismes de groupes.*

Définition 2.3.1. — Soient G et H deux groupes. Un *morphisme de groupes* de H dans G est une application φ de H dans G telle que $\varphi(hh') = \varphi(h)\varphi(h')$ pour tout $(h, h') \in H^2$.

2.3.2. *Commentaires à propos de la terminologie.* — Le terme «morphisme» est générique : il désigne une application compatible avec un certain type de structure qui est précisé ensuite (sauf s'il est clairement indiqué par le contexte). Il y a ainsi des morphismes de groupes comme ici, mais aussi des morphisme d'anneaux, et plein d'autres types de morphismes que vous croiserez plus tard, dans ce cours ou ailleurs. Seules quelques classes vénérables d'objets échappent pour des raisons historiques à ce vocabulaire standardisé : on parle d'application linéaire et non de morphisme d'espaces vectoriels, ou d'application continue et non de morphisme d'espaces topologiques.

On rencontre encore parfois le terme «homomorphisme» au lieu de «morphisme», mais il a tendance à tomber en désuétude. Il perdure dans les notations : l'ensemble des morphismes d'un groupe G dans un groupe H est ainsi noté $\mathrm{Hom}(G, H)$.

Un *endomorphisme* désigne (quel que soit le type de structure en jeu) un morphisme d'un objet dans lui-même. Si G est un groupe, un endomorphisme (de groupes) de G est donc un morphisme (de groupes) de G dans G. L'ensemble des endomorphismes de G est noté $\mathrm{End}(G)$.

***Exemples 2.3.3* (Les cas triviaux).** — Soient G un groupe. L'application Id_G est un endomorphisme de groupes. Plus généralement, si H est un sous-groupe de G, l'inclusion de H dans G est un morphisme de groupes.

2.3.4. *Composition de morphismes.* — Soient $\varphi\colon G \to G'$ et $\psi\colon G' \to G''$ deux morphismes de groupes. On vérifie immédiatement que la composée $\psi \circ \varphi\colon G \to G''$ est un morphisme de groupes.

2.3.5. — Soit $\varphi\colon H \to G$ un morphisme de groupes. Par définition, φ est une application qui se comporte bien vis-à-vis des lois internes en jeu. Mais cette condition entraîne en fait automatiquement que φ se comporte également bien vis-à-vis des éléments neutres et des inversions, comme nous allons le voir ci-dessous ; ce petit miracle est dû à la propriété de simplification des égalités ans un groupe (2.1.6).

2.3.5.1. — On a $\varphi(e_H) = e_G$. En effet, l'égalité $e_H^2 = e_H$ implique que

$$\varphi(e_H)^2 = \varphi(e_H^2) = \varphi(e_H),$$

ce qui entraîne que $\varphi(e_H) = e_G$.

2.3.5.2. — Soit $h \in H$. On a $\varphi(h^{-1}) = \varphi(h)^{-1}$. En effet,

$$\varphi(h)\varphi(h^{-1}) = \varphi(hh^{-1}) = \varphi(e_H) = e_G,$$

ce qui entraîne que $\varphi(h^{-1}) = \varphi(h)^{-1}$.

2.4. Sous-groupes et morphismes. — Soit $\varphi\colon H \to G$ un morphisme de groupes ; nous allons décrire l'effet de φ sur les sous-groupes de H et de G.

2.4.1. — On vérifie immédiatement que si H' est un sous-groupe de H alors $\varphi(H')$ est un sous-groupe de G ; en particulier, $\mathrm{Im}\,\varphi = \varphi(H)$ est un sous-groupe de G.

Soit E un sous-ensemble de G. Le sous-groupe $\langle\varphi(E)\rangle$ de G est égal à $\varphi(\langle E\rangle)$: c'est par exemple une conséquence immédiate de la description concrète du sous-groupe engendré par une partie.

2.4.2. — On vérifie immédiatement que si G' est un sous-groupe de G alors $\varphi^{-1}(G')$ est un sous-groupe de H ; en particulier, $\varphi^{-1}(e_G)$ est un sous-groupe de H, appelé le *noyau* de φ et souvent noté $\mathrm{Ker}\varphi$.

2.4.3. — Le morphisme φ est injectif si et seulement si son noyau est trivial. C'est en effet clairement nécessaire. Réciproquement, supposons que $\mathrm{Ker}\varphi = \{e_H\}$ et soient h_1 et h_2 deux éléments de H tels que $\varphi(h_1) = \varphi(h_2)$. On a alors

$$\varphi(h_1h_2^{-1}) = \varphi(h_1)\varphi(h_2)^{-1} = e_G,$$

et donc $h_1h_2^{-1} = e_H$ en vertu de notre hypothèse sur $\mathrm{Ker}\varphi$; il vient $h_1 = h_2$.

2.5. Isomorphismes. — Nous allons maintenant introduire une classe fondamentale de morphismes de groupes.

2.5.1. — Soit $\varphi \colon H \to G$ un morphisme de groupes. Supposons que φ est bijectif. Sa réciproque *ensembliste* φ^{-1} est alors également un morphisme de groupes : en effet si g_1 et g_2 sont deux éléments de g, on a

$$g_1 g_2 = \varphi(\varphi^{-1}(g_1))\varphi(\varphi^{-1}(g_2)) = \varphi(\varphi^{-1}(g_1)\varphi^{-1}(g_2))$$

ce qui entraîne que $\varphi^{-1}(g_1 g_2) = \varphi^{-1}(g_1)\varphi^{-1}(g_2))$, par définition de φ^{-1}.

On dit alors que φ est un *isomorphisme (de groupes).*

***Exemple 2.5.2* (Un cas trivial).** — Si G est un groupe, Id_G est un isomorphisme.

2.5.3. — Il résulte immédiatement des définitions que la composée de deux isomorphismes de groupes est un isomorphisme de groupes ; la bijection réciproque d'un isomorphisme de groupes est un isomorphisme de groupes.

2.5.4. *Automorphismes.* — Si G est un groupe, un *automorphisme (de groupe)* de G est un isomorphisme de G dans G. L'ensemble des automorphismes de Gest un sous-groupe de $\mathfrak{S}_G$; on le note $\mathrm{Aut}(G)$.

Remarque 2.5.5. — Soient G et H deux groupes. On dit qu'ils sont *isomorphes* s'il existe un isomorphisme $\varphi \colon G \simeq H$. Supposons que ce soit le cas. Comme φ et sa réciproque φ^{-1} respectent le produit, l'élément neutre, et l'inversion des éléments, *toute propriété du groupe G qui ne met en jeu que la loi interne, l'élément neutre et l'inversion est également satisfaite par H.* Nous utiliserons très fréquemment, et implicitement, ce principe général.

Par exemple : si G est trivial, H est trivial ; si G est abélien, H est abélien ; si les seuls sous-groupes de G sont G et $\{e\}$, les seuls sous-groupes de H sont H et $\{e\}$, etc.

2.6. Structures sur les ensembles de morphismes. — Si G et H sont deux groupes, $\mathrm{Hom}(G, H)$ n'hérite en général d'aucune structure particulière ; c'est simplement un ensemble. Mais nous allons voir que la situation s'améliore lorsque H est abélien.

2.6.1. — Supposons donc donné deux groupes G et H, en supposant de plus H abélien et noté additivement. On vérifie immédiatement que pour tout couple (φ, ψ) d'éléments de $\mathrm{Hom}(G, H)$, l'application

$$\varphi + \psi := \quad (G \to H, g \mapsto \varphi(g) + \psi(g))$$

est un morphisme de groupes, et que $(\varphi, \psi) \mapsto \varphi + \psi$ fait de $\mathrm{Hom}(G, H)$ un groupe abélien. Son élément neutre est l'application nulle ; l'opposé d'un morphisme φ de G dans H est le morphisme $g \mapsto -\varphi(g)$.

2.6.2. — Soit G un groupe abélien noté additivement. Par ce qui précède, on dispose d'une loi de groupe abélien naturelle sur $\mathrm{End}(G)$, notée $+$. Nous laissons le lecteur vérifier que $(\mathrm{End}(G), +, \circ)$ est un anneau ; son élément neutre multiplicatif est l'identité ; son groupe des éléments inversibles coïncide avec $\mathrm{Aut}(G)$.

2.7. Les automorphismes intérieurs. — Soit G un groupe. Le but de ce qui suit est de construire un groupe particulier d'automorphismes de G, dits *intérieurs.*

2.7.1. — Soit $g \in G$. Notons ι_g l'application de G dans G qui envoie un élément h sur ghg^{-1}. Si h et h' sont deux éléments de G, on a $ghh'g^{-1} = ghg^{-1}gh'g^{-1}$, soit encore $\iota_g(hh') = \iota_g(h)\iota_g(h')$. Ainsi, ι_g est un morphisme de groupes, appelé la *conjugaison* par g ; on a clairement $\iota_e = \mathrm{Id}$.

On dira que deux sous-ensembles E et E' de G sont conjugués s'il existe $g \in G$ tel que $E' = \iota_g(E) = gEg^{-1}$; en pratique, on appliquera surtout cette notion lorsque E et E' sont des singletons (on parlera alors plus simplement d'éléments conjugués) ou lorsque ce sont des sous-groupes de G.

2.7.2. — Pour tout $(g, g', h) \in G$ on a

$$(gg')h(gg')^{-1} = gg'h(g')^{-1}g^{-1} = g[g'h(g')^{-1}]g^{-1},$$

soit encore $\iota_{gg'}(h) = \iota_g(\iota_{g'}(h))$; par conséquent, $\iota_{gg'} = \iota_g \circ \iota_{g'}$.

On a en particulier pour tout élément g de G les égalités $\iota_g \circ \iota_{g^{-1}} = \iota_e = \mathrm{Id}$ et $\iota_{g^{-1}} \circ \iota_g = \iota_e = \mathrm{Id}$; il s'ensuit que pour tout $g \in G$, le morphisme ι_g est un automorphisme de G, de réciproque $\iota_{g^{-1}}$.

Ainsi $g \mapsto \iota_g$ apparaît comme une application de G dans le groupe Aut G des automorphismes de G. La formule $\iota_{gg'} = \iota_g \circ \iota_{g'}$ signifie que cette application est un morphisme de groupes.

2.7.3. — Les automorphismes de la forme ι_g pour $g \in G$ sont appelés les automorphismes *intérieurs* de G. Par définition, l'ensemble des automorphismes intérieurs de G est l'image du morphisme $g \mapsto \iota_g$ de G dans $\mathrm{Aut}(G)$; c'est donc un sous-groupe de $\mathrm{Aut}(G)$.

2.7.4. *Le centre de G.* — Le noyau du morphisme $g \mapsto \iota_g$ est un sous-groupe de G, noté $\mathrm{Z}(G)$ et appelé le *centre* de G. Par définition, un élément g de G appartient à $\mathrm{Z}(G)$ si et seulement si $\iota_g = \mathrm{Id}$, ce qui veut dire que $ghg^{-1} = h$ pour tout $h \in G$, soit encore que $gh = hg$ pour tout $h \in G$.Un élément de G est donc dans le centre de G si et seulement s'il commute avec tout le monde. En particulier les éléments du centre commutent entre eux : par conséquent $\mathrm{Z}(G)$ est abélien.

2.7.5. *Le cas abélien.* — Si G est abélien, on a $\mathrm{Z}(G) = G$, et tout automorphisme intérieur de G est trivial.

Cette remarque permet d'exhiber facilement des automorphismes non intérieurs : par exemple, l'application $z \mapsto -z$ est un automorphisme non trivial du groupe abélien $\mathbf{Z}$, et il n'est donc pas intérieur.

Remarque 2.7.6. — Soient h et h' deux éléments conjugués de G ; soit $g \in G$ tel que $ghg^{-1} = h'$. Soit φ un morphisme de G vers un groupe G'. On a alors l'égalité $\varphi(h') = \varphi(ghg^{-1}) = \varphi(g)\varphi(h)\varphi(g)^{-1}$. Ainsi, $\varphi(h)$ et $\varphi(h')$ sont eux aussi conjugués. *Si de plus G' est abélien*, il vient $\varphi(h) = \varphi(h')$.

2.8. Relations de congruence. — Soit G un groupe et soit H un sous-groupe de G. Le but de ce qui suit est d'introduire et d'étudier deux relations d'équivalences sur G associées à H.

Définition 2.8.1. — La *congruence à gauche modulo H* est la relation $\mathscr{R}$ sur G définie par la condition

$$x\mathscr{R}y \iff x^{-1}y \in H.$$

La *congruence à droite modulo H* est la relation $\mathscr{S}$ sur G définie par la condition

$$x\mathscr{S}y \iff xy^{-1} \in H.$$

2.8.2. *Commentaires et premières propriétés.* — On vérifie aisément que $\mathscr{R}$ et $\mathscr{S}$ sont des relations d'équivalence. Nous allons tout d'abord décrire leurs classes.

Soit x un élément de G. La classe d'équivalence de x pour $\mathscr{R}$ est par définition l'ensemble des $y \in G$ tels que $x\mathscr{R}y$, c'est-à-dire tels que $x^{-1}y \in H$, c'est-à-dire encore tels que $y \in xH$. On dit que xH est la *classe à gauche de x modulo H*. La classe d'équivalence de x pour $\mathscr{S}$ est par définition l'ensemble des $y \in G$ tels que $y\mathscr{S}Sx$, c'est-à-dire tels que $yx^{-1} \in H$, c'est-à-dire encore tels que $y \in Hx$. On dit que Hx est la *classe à droite de x modulo H*.

Remarquons que les applications $y \mapsto xy$ et $y \mapsto yx$ de G dans lui-même sont injectives. Par conséquent, $y \mapsto xy$ établit une *bijection* entre H et xH ; de même, $y \mapsto yx$ établit une *bijection* entre H et Hx.

L'ensemble quotient de G par $\mathscr{R}$, c'est-à-dire l'ensemble des classes à gauche modulo H, est noté G/H. L'ensemble quotient de G par $\mathscr{S}$, c'est-à-dire l'ensemble des classes à droite de G modulo H, est noté $H\backslash G$.

2.8.3. *Une bijection naturelle $G/H \simeq H\backslash G$.* — Soient x et y deux éléments de G. On a

$$x\mathscr{R}y \iff x^{-1}y \in H \iff x^{-1}(y^{-1})^{-1} \in H \iff x^{-1}\mathscr{S}y^{-1}.$$

Ainsi, l'application $x \mapsto Hx^{-1}$ de G dans $H\backslash G$ est $\mathscr{R}$-invariante et elle induit donc par passage au quotient une application de G/H vers $H\backslash G$. De même, l'application $x \mapsto x^{-1}H$ de G dans G/H est $\mathscr{S}$-invariante et induit donc par passage au quotient une application de $H\backslash G$ vers G/H.

Il est immédiat que ces deux applications sont réciproques l'une de l'autre ; on a ainsi construit une *bijection* de G/H sur $H\backslash G$.

***Exemple 2.8.4* (Deux cas extrêmes).** — Les quotients G/G et $G\backslash G$ sont des singletons. Les applications quotient $G \to G/\{e\}$ et $G \to \{e\}\backslash G$ sont des bijections.

Remarque 2.8.5. — Soit (G_i) une famille de groupes. Pour tout i, soit H_i un sous-groupe de G_i. Il est immédiat que $\prod H_i$ est un sous-groupe de $\prod G_i$, et que la relation produit (1.3.12) des relations de congruence à gauche (resp. à droite) modulo les H_i est la relation de congruence à gauche (resp. à droite) modulo $\prod H_i$. On en déduit en vertu de 1.3.12 que les passages au quotient composante par composante induisent deux bijections

$$\prod G_i/\prod H_i \simeq \prod (G_i/H_i) \quad \text{et} \quad \prod H_i\backslash \prod G_i \simeq \prod (H_i\backslash G_i).$$

2.9. Le lemme de Lagrange. — Comme application immédiate de la théorie des congruences modulo un sous-groupe nous allons établir un lemme dont la preuve est très simple, qui joue un rôle fondamental en théorie des groupes finis.

Notation 2.9.1. — Si E est un ensemble fini, on notera $|E|$ son cardinal.

Lemme 2.9.2. — *Soit G un groupe fini et soit H un sous-groupe de G.*

(1) $|H|$ *divise* $|G|$ *(cet énoncé est souvent appelé le «lemme de Lagrange»).*

(2) $|G|/|H| = |G/H| = |H\backslash G|$.

Démonstration. — La congruence à gauche modulo H étant une relation d'équivalence, G est la réunion disjointe des classes à gauche modulo H. On a vu au 2.8.2 que toute classe à gauche modulo H est en bijection avec H, et en particulier de cardinal égal à $|H|$; il vient $|G| = |G/H| \cdot |H|$. On montre de même que $|G| = |H\backslash G| \cdot |H|$ (on pourrait également déduire l'égalité $|G/H| = |H\backslash G|$ de 2.8.3). □

Définition 2.9.3. — On appelle *indice* de H dans G, et l'on note $[G : H]$, le cardinal (éventuellement infini) de G/H, qui est aussi le cardinal de $H\backslash G$ d'après 2.8.3. Lorsque G est fini, on a $[G : H] = |G|/|H|$ d'après le lemme 2.9.2 ci-dessus.

2.10. Sous-groupes distingués. — Si G est un groupe et si H est un sous-groupe de G, les quotients G/H et $H\backslash G$ sont en général simplement des ensembles, sans structure naturelle de groupes. Le but de ce qui suit est d'introduire une classe particulière de sous-groupes de G pour lesquels G/H et $H\backslash G$ héritent tous deux d'une telle structure.

Lemme 2.10.1. — *Soit X un ensemble muni d'une loi de composition interne $(x, y) \mapsto xy$ et soit f une surjection de X sur un ensemble E. Il existe alors au plus une loi de composition interne $*$ sur E telle que $f(xy) = f(x) * f(y)$ pour tout $(x, y) \in X^2$.*

Supposons qu'une telle loi existe et que X soit un groupe. Sous cette hypothèse $(E, *)$ est un groupe, abélien si X est abélien, et l'application f est un morphisme de groupes.

Démonstration. — Supposons qu'il existe une loi $*$ comme dans l'énoncé et soient α et β deux éléments de E. Par surjectivité de f il existe x et y dans X tels que $f(x) = \alpha$ et $f(y) = \beta$. On a alors nécessairement $\alpha * \beta = f(x) * f(y) = f(xy)$, et $*$ est bien uniquement déterminée.

Faisons de plus l'hypothèse que X est un groupe. Soient α, β et γ trois éléments de E. Soient x, y et z des antécédents respectifs de α, β et γ dans X. On a alors

$$\begin{aligned}
\alpha * (\beta * \gamma) &= f(x) * (f(y) * f(z)) \\
&= f(x) * (f(yz)) \\
&= f(x(yz)) \\
&= f((xy)z) \\
&= f(xy) * f(z) \\
&= (f(x) * f(y)) * f(z) \\
&= (\alpha * \beta) * \gamma.
\end{aligned}$$

Ainsi, $*$ est associative ; on montre de même qu'elle est commutative si X est abélien. Soit ξ l'élément neutre de X et soit α un élément de E ; soit x un antécédent de α dans X. On a

$$f(\xi) * \alpha = f(\xi) * f(x) = f(\xi x) = f(x) = \alpha$$

et de même $\alpha * f(\xi) = \alpha$; ainsi, $f(\xi)$ est un élément neutre pour $*$. Enfin on a $\alpha * f(x^{-1}) = f(x) * f(x^{-1}) = f(xx^{-1}) = f(\xi)$ et de même $f(x^{-1}) * \alpha = f(\xi)$. Ainsi, α possède un symétrique pour $*$, à savoir $f(x^{-1})$. On a donc bien montré que $(E, *)$ est un groupe, abélien si X est abélien. La formule $f(xy) = f(x) * f(y)$ assure alors que f est un morphisme de groupes.

□

Théorème 2.10.2. — *Soit G un groupe et soit H un sous-groupe de G.*

(A) *Les assertions suivantes sont équivalentes :*

  (i) *il existe un groupe G' et un morphisme $f : G \to G'$ tel que $H = \operatorname{Ker} f$;*

  (ii) *pour tout $g \in G$ on a $gHg^{-1} \subset H$;*

  (iii) *pour tout $g \in G$ on a $gHg^{-1} = H$;*

  (iv) *pour tout $g \in G$ on a $gH = Hg$;*

  (v) *il existe une loi de groupe sur G/H telle que la flèche quotient $x \mapsto xH$ de G vers G/H soit un morphisme ;*

  (vi) *il existe une loi de groupe sur $H\backslash G$ telle que la flèche quotient $x \mapsto Hx$ de G vers $H\backslash G$ soit un morphisme.*

(B) *Si* (v) *est satisfaite, la loi de groupe en question sur G/H est unique, et le noyau du morphisme quotient $G \to G/H$ est égal à H. Si* (vi) *est satisfaite, la loi de groupe en question sur $H\backslash G$ est unique, et le noyau du morphisme quotient $G \to H\backslash G$ est égal à H.*

Lorsque les assertions équivalentes (i) *à* (vi) *sont satisfaites, on dit que H est* distingué *(ou parfois* normal*) dans G, et l'on écrit $H \lhd G$.*

Démonstration. — Commençons par montrer (B). Remarquons tout d'abord que si $*$ est une loi de groupe sur G/H, l'application quotient $G \to G/H$ est un morphisme si et seulement si on a $(xH)*(yH) = (xy)H$ pour tout $(x, y) \in G^2$; or il résulte du lemme

2.10.1 qu'il existe au plus une loi interne $*$ sur G/H vérifiant cette dernière formule, et que si une telle loi existe c'est automatiquement une loi de groupe. Plaçons-nous dans le cas où une telle loi existe. Le noyau du morphisme $G \to G/H$ est alors l'ensemble des éléments de G ayant même image que e dans G/H, c'est-à-dire la classe à gauche de e modulo H, à savoir $eH = H$. La partie de l'assertion (B) relative à (v) est donc démontrée ; sa partie relative à (vi) se prouve de façon analogue.

Venons-en à (A). Il est immédiat que (i)$\Rightarrow$(ii) et que (iii)$\Rightarrow$(ii). Montrons que (ii)$\Rightarrow$(iii), et partant que (ii) $\Longleftrightarrow$ (iii). Supposons que (ii) est vraie et soit $g \in G$. On a alors $gHg^{-1} \subset H$. Mais on a aussi $g^{-1}Hg = g^{-1}H(g^{-1})^{-1} \subset H$. En conjuguant par g les deux membres de l'égalité, on obtient l'inclusion $H \subset gHg^{-1}$, d'où finalement l'égalité $gHg^{-1} = H$.

Il est clair que (iii) $\Longleftrightarrow$ (iv). Montrons maintenant que (ii)$\Rightarrow$ (v). On suppose donc que (ii) est vraie et l'on veut montrer (v). D'après le rappel fait en début de preuve, il suffit de prouver qu'il existe une loi de composition interne $*$ sur G/H telle que $(xH) * (yH) = (xy)H$ pour tout (x, y) dans G^2. Soient x et y deux éléments de G et soient x' et y' deux éléments de G tels que $x'H = xH$ et $y'H = yH$, c'est-à-dire tels que $x^{-1}x' \in H$ et $y^{-1}y' \in H$. Comme $x^{-1}x' \in H$ l'hypothèse (ii) assure que $y^{-1}x^{-1}x'y \in H$. On a alors

$$(xy)^{-1}x'y' = y^{-1}x^{-1}x'y' = \underbrace{y^{-1}x^{-1}x'y}_{\in H}\,\underbrace{y^{-1}y'}_{\in H} \in H.$$

Par conséquent, $(x'y')H = (xy)H$. L'application de $G \times G$ vers G/H qui envoie (x, y) sur $(xy)H$ est donc invariante par le produit des congruences à gauche modulo H sur chacun des facteurs ; elle induit de ce fait par passage au quotient une application de $G/H \times G/H$ vers G/H qui envoie (xH, yH) sur $(xy)H$ pour tout couple (x, y) d'éléments de G, ce qu'on souhaitait établir.

L'implication (ii)$\Rightarrow$(vi) se montre de manière analogue. Et il est immédiat en vertu de (B) que (v)$\Rightarrow$(i) et que (vi)$\Rightarrow$(i). $\square$

Commentaires 2.10.3. — On prendra garde que l'assertion (v) (resp. (vi)) ne se contente pas d'affirmer l'existence d'une loi de groupe sur le quotient G/H (resp. $H\backslash G$), ce qui n'aurait rien de bien palpitant (on peut en effet démontrer que tout ensemble non vide possède une loi de groupe, essayez à l'occasion de faire l'exercice). Elle affirme plus précisément l'existence d'une telle loi *faisant de l'application quotient un morphisme*, et c'est cette dernière propriété qui est fondamentale.

Remarque 2.10.4. — En pratique, pour montrer qu'un sous-groupe est distingué, on utilise le plus souvent la caractérisation par les assertions (i) ou (ii) de l'énoncé du théorème 2.10.2 ; ainsi, les exemples 2.10.5, 2.10.6 et 2.10.7 ci-dessous, présentés sans justification, se traitent aisément à l'aide de (ii).

***Exemple 2.10.5* (Les cas triviaux).** — Soit G un groupe. Les sous-groupes G et $\{e\}$ de G sont distingués.

***Exemple 2.10.6* (Le cas abélien).** — Soit G un groupe abélien. Tout sous-groupe de G est alors distingué.

Exemple 2.10.7 **(Intersection de sous-groupes distingués)**
Soit G un groupe et soit (H_i) une famille de sous-groupes distingués de G. Le sous-groupe $\bigcap H_i$ de G est distingué.

Si E est un sous-ensemble de G, l'intersection des sous-groupes distingués de G contenant E est donc un sous-groupe distingué H de G, qui est le plus petit sous-groupe distingué de G contenant E. Nous vous laissons démontrer à titre d'exercice que H est le sous-groupe de G engendré par $\{geg^{-1}\}_{g\in G, e\in E}$.

Exemple 2.10.8 **(Le centre).** — Soit G un groupe. Le centre $\mathrm{Z}(G)$ de G (2.7.4) est distingué, puisque c'est le noyau du morphisme $g \mapsto (h \mapsto ghg^{-1})$ de G dans $\mathrm{Aut}(G)$.

Vous pouvez à titre l'exercice le vérifier en utilisant la définition de $\mathrm{Z}(G)$ comme l'ensemble des $g \in G$ tels que $gh = hg$ pour tout $h \in G$, et en utilisant la caractérisation des sous-groupes distingués par l'assertion (ii) du théorème 2.10.2.

Contre-exemple 2.10.9. — Le sous-groupe T de $\mathrm{GL}_2(\mathbf{Q})$ constitué des matrices diagonales n'est pas distingué. En effet,

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \in T,$$

mais

$$\begin{aligned} &\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix} \notin T. \end{aligned}$$

Exemple 2.10.10 **(Produit de sous-groupes distingués)**
Soit (G_i) une famille de groupes ; pour tout i, soit H_i un sous-groupe de G_i. Nous laissons le lecteur vérifier que $\prod H_i$ est distingué dans $\prod G_i$ si et seulement si H_i est distingué dans G_i pour tout i, et que si c'est le cas la bijection naturelle

$$\prod G_i / \prod H_i \simeq \prod G_i/H_i$$

est un isomorphisme de groupes.

2.11. Propriété universelle du quotient par un sous-groupe distingué. — Nous allons voir que le quotient d'un groupe par un sous-groupe distingué possède une propriété universelle qui est l'analogue dans le monde des groupes de la propriété universelle ensembliste du quotient par une relation d'équivalence (théorème 1.3.6).

Si G est un groupe et si H est un sous-groupe distingué de G, le quotient G/H sera toujours considéré comme muni de son unique loi de groupe faisant de $G \to G/H$ un morphisme, et celle-ci sera sauf mention expresse du contraire notée sans symbole (par concaténation).

Lemme 2.11.1. — *Soit $\varphi\colon G \to G'$ un morphisme de groupes, soit H un sous-groupe distingué de G et soit $\mathscr{R}$ la relation de congruence à gauche modulo H.*

(1) *L'application φ est $\mathscr{R}$-invariante si et seulement si $H \subset \operatorname{Ker} \varphi$.*

(2) *Supposons que φ est $\mathscr{R}$-invariante, c'est-à-dire par ce qui précède que H est contenu dans* Kerφ*. L'application de G/H vers G' déduite de φ par passage au quotient est alors un morphisme de groupes.*

Démonstration. — Supposons que φ est $\mathscr{R}$-invariante et soit $h \in H$. Comme on a $h\mathscr{R}e$ il vient $\varphi(h) = \varphi(e)$ et $h \in \operatorname{Ker}\varphi$; par conséquent, $H \subset \operatorname{Ker}\varphi$. Réciproquement, supposons que $H \subset \operatorname{Ker}\varphi$ et soient g et g' deux éléments de G tels que $g\mathscr{R}g'$. On a alors $g^{-1}g' \in H \subset \operatorname{Ker}\varphi$, si bien que $\varphi(g)^{-1}\varphi(g') = \varphi(g^{-1}g') = e$, ce qui entraîne que $\varphi(g) = \varphi(g')$; par conséquent, φ est $\mathscr{R}$-invariante, d'où (1).

Montrons maintenant (2). Soit $\pi\colon G \to G/H$ et soit $\psi\colon G/H \to G'$ l'application déduite de φ par passage au quotient ; par définition, on a $\psi(\pi(x)) = \varphi(x)$ pour tout $x \in G$. Soient α et β deux éléments de G/H. Choisissons x et y dans G tels que $\pi(x) = \alpha$ et $\pi(y) = \beta$ (ce qui est possible par surjectivité de π). On a alors

$$\begin{aligned} \psi(\alpha\beta) &= \psi(\pi(x)\pi(y)) \\ &= \psi(\pi(xy)) \\ &= \varphi(xy) \\ &= \varphi(x)\varphi(y) \\ &= \psi(\pi(x))\psi(\pi(y)) \\ &= \psi(\alpha)\psi(\beta), \end{aligned}$$

et ψ est donc bien un morphisme de groupes. □

On déduit du lemme 2.11.1 et de la propriété universelle du quotient dans le cas ensembliste (théorème 1.3.6) l'énoncé suivant.

***Théorème 2.11.2* (Propriété universelle du groupe quotient)**
Soit G un groupe, soit H un sous-groupe distingué de G et soit $\pi\colon G \to G/H$ le morphisme quotient. Pour tout groupe G' et tout morphisme φ de G vers G' tel que $H \subset$ Kerφ*, il existe un unique morphisme de groupes ψ de G/H vers G' tel que $\varphi = \psi \circ \pi$, c'est-à-dire tel que le diagramme*

$$\begin{tikzcd} G \arrow[r, "\varphi"] \arrow[d, "\pi"'] & G' \\ G/H \arrow[ur, "\psi"'] & \end{tikzcd}$$

commute. On dira que ψ est le morphisme de G/H vers G' induit par *φ.*

Commentaires 2.11.3. — Soient G, H et π comme ci-dessus. Supposons donné un morphisme ψ de G/H dans un groupe G' et posons $\varphi = \psi \circ \pi$. On déduit de l'égalité $H = \operatorname{Ker}\pi$ (et même déjà de l'inclusion $H \subset \operatorname{Ker}\pi$) que $H \subset \operatorname{Ker}\varphi$.

À la lueur de cette remarque, on peut reformuler comme suit la propriété universelle du quotient : *pour tout groupe G', la formule $\psi \mapsto \psi \circ \pi$ établit une bijection entre* $\operatorname{Hom}(G/H, G')$ *et l'ensemble des φ appartenant à* $\operatorname{Hom}(G, G')$ *tels que $H \subset$* Kerφ*.*

Autrement dit, *se donner un morphisme de G/H vers un groupe G', c'est se donner un morphisme de G vers G' qui est trivial sur H.*

Remarque 2.11.4. — La propriété universelle du morphisme quotient $\pi\colon G \to G/H$ caractérise ce dernier à isomorphisme unique près, dans le sens suivant. Supposons donné un morphisme $p\colon G \to \Gamma$ satisfaisant la même propriété universelle, c'est-à-dire tel que $H \subset \operatorname{Ker} p$ et tel que pour tout groupe G' l'application $\psi \mapsto \psi \circ p$ établisse une bijection entre $\operatorname{Hom}(\Gamma, G')$ et le sous-ensemble de $\operatorname{Hom}(G, G')$ formé des morphismes dont le noyau contient H. Il existe alors un unique morphisme $\iota\colon G/H \to \Gamma$ tel que le diagramme

$$\begin{tikzcd} & G \arrow[dl, "\pi"'] \arrow[dr, "p"] & \\ G/H \arrow[rr, "\iota"] & & \Gamma \end{tikzcd}$$

commute, c'est-à-dire tel que $\iota \circ \pi = p$, et ι est un isomorphisme. La preuve est analogue à celle de l'énoncé ensembliste correspondant (proposition 1.3.9).

2.11.5. *Comment penser au quotient ?* — Intuitivement, G/H est le groupe construit à partir de G en *décrétant* que les éléments de H sont triviaux, et en n'imposant *aucune autre contrainte*, sinon bien sûr celles qui en découlent par la théorie des groupes ; ou encore comme au groupe le plus général construit à partir de G en décrétant que les éléments de H sont triviaux.

2.12. Quotient d'un groupe par un sous-ensemble arbitraire. — Soit G un groupe et soit E un simple *sous-ensemble* E de G. Peut-on encore définir le «groupe le plus général construit à partir de G en décrétant que les éléments de E sont triviaux» ? Techniquement, cela revient à demander s'il existe un groupe Γ et un morphisme $\pi\colon G \to \Gamma$ tel que $E \subset \operatorname{Ker}\pi$ et tel que pour tout groupe G', l'application $\psi \mapsto \psi \circ \pi$ établisse une bijection entre $\operatorname{Hom}(\Gamma, G')$ et le sous-ensemble de $\operatorname{Hom}(G, G')$ formé des morphismes φ tels que $E \subset \operatorname{Ker}\varphi$. Nous allons-voir ci-dessous que la réponse est affirmative.

2.12.1. — Soit H le plus petit sous-groupe distingué de G contenant E (2.10.7) et soit π le morphisme quotient $G \to G/H$. Si φ est un morphisme de groupes de source G, son noyau $\operatorname{Ker}\pi$ est un sous-groupe distingué de G, si bien que $E \subset \operatorname{Ker}\pi$ ei et seulement si $H \subset \operatorname{Ker}\pi$. La propriété universelle du morphisme π peut alors de récrire en disant que $\psi \mapsto \psi\circ\pi$ établit une bijection entre $\operatorname{Hom}(G/H, G')$ et l'ensemble des φ appartenant à $\operatorname{Hom}(G, G')$ tels que $E \subset \operatorname{Ker}\varphi$; c'est précisément la propriété universelle cherchée. En termes un peu plus informels : *se donner un morphisme de G/H vers un groupe G', c'est se donner un morphisme de G vers G' qui est trivial sur E.*

Commentaires 2.12.2. — Par ce qui précède, on peut penser à G/H comme au groupe le plus général construit à partir de G en décrétant que les éléments de E sont triviaux ; or nous avons par ailleurs dit plus haut qu'on pouvait le voir comme le groupe le plus général construit à partir de G en décrétant que les éléments de H

sont triviaux. On peut informellement expliquer cette double description comme suit : lorsqu'on décrète que les éléments de E sont triviaux, la théorie des groupes impose des «dommages collatéraux» : cette opération trivialise non seulement les éléments de E, mais également ceux de H. Notons toutefois que comme le noyau du morphisme quotient $G \to G/H$ est exactement H, les dommages collatéraux en question ne se propagent pas au-delà de H.

2.13. Quotients, images et noyaux. — Soit $\varphi\colon G \to G'$ un morphisme de groupes.

2.13.1. — Soit H un sous-groupe distingué de G contenu dans $\mathrm{Ker}\varphi$. Notons π le morphisme quotient de G vers G/H et $\psi\colon G/H \to G'$ le morphisme déduit de φ par passage au quotient. Il résulte de 1.3.10 que $\mathrm{Im}\psi = \mathrm{Im}\varphi$. Par ailleurs comme π est surjective, ψ est injective si et seulement si on a

$$\psi(\pi(x)) = e \iff \pi(x) = e$$

pour tout $x \in G$, équivalence que l'on peut récrire $\varphi(x) = e \iff x \in H$. Autrement dit, ψ est injective si et seulement si $\mathrm{Ker}\varphi = H$ (on aurait pu aussi le déduire de la condition d'injectivité énoncée en 1.3.10, mais l'argument donné ici est un peu plus rapide).

2.13.2. *L'isomorphisme fondamental.* — Appliquons ce qui précède lorsque H est égal à $\mathrm{Ker}\varphi$ tout entier. Le morphisme ψ est alors injectif, et son image est $\mathrm{Im}\varphi$. Autrement dit, φ *induit par passage au quotient un isomorphisme* $(G/\mathrm{Ker}\varphi) \simeq \mathrm{Im}\varphi$. Il s'ensuit d'après le lemme 2.9.2 que si G est fini alors

$$|G| = |\mathrm{Im}\varphi| \cdot |\mathrm{Ker}\varphi|.$$

2.14. Sous-groupes d'un quotient. — Nous allons maintenant étudier les sous-groupes du quotient d'un groupe donné par un sous-groupe distingué.

Lemme 2.14.1. — *Soit G un groupe, soit H un sous-groupe distingué de G et soit $\pi\colon G \to G/H$ le morphisme quotient.*

1. *Soit Γ un sous-groupe de G ; on a $H \cap \Gamma \lhd \Gamma$ et $\pi(\Gamma) \simeq \Gamma/(H \cap \Gamma)$.*
2. *Les formules $\Gamma \mapsto \pi(\Gamma)$ et $\Delta \mapsto \pi^{-1}(\Delta)$ établissent une bijection croissante (pour l'inclusion) entre l'ensemble des sous-groupes de G contenant H et l'ensemble des sous-groupe de G/H.*

Démonstration. — Montrons d'abord (1). Comme $H = \mathrm{Ker}\pi$, le noyau de $\pi|_\Gamma$ est égal à $H \cap \Gamma$. Ce dernier est donc distingué dans Γ, et $\pi(\Gamma) \simeq \Gamma/(H \cap \Gamma)$ en vertu de 2.13.2.

Montrons maintenant (2). Soit Γ un sous-groupe de G contenant H ; montrons que $\pi^{-1}(\pi(\Gamma)) = \Gamma$. Il est clair que $\Gamma \subset \pi^{-1}(\pi(\Gamma))$. Réciproquement, soit $g \in G$ tel que $\pi(g) \in \pi(\Gamma)$. Il existe alors $\gamma \in \Gamma$ tel que $\pi(g) = \pi(\gamma)$, c'est-à-dire tel que $\pi(g\gamma^{-1}) = e$. Ainsi, $g\gamma^{-1} \in \mathrm{Ker}\pi = H \subset \Gamma$. Puisque $g = (g\gamma^{-1})\gamma$, on a $g \in \Gamma$.

La surjectivité de π implique par ailleurs que $\pi(\pi^{-1}(\Delta)) = \Delta$ pour toute partie Δ de G/H ; c'est en particulier le cas lorsque Δ est un sous-groupe de G/H.

Ainsi, les formules données en (2) établissent bien une bijection entre l'ensemble des sous-groupes de G contenant H et l'ensemble des sous-groupe de G/H. Il est par ailleurs immédiat qu'elles définissent des applications croissantes, ce qui achève la démonstration. □

2.14.2. *Le noyau d'un morphisme induit par passage au quotient.* — Soit G un groupe, soit H un sous-groupe distingué de G, et soit φ un morphisme de G vers G' tel que $H \subset \mathrm{Ker}\varphi$; soit $\psi \colon G/H \to G'$ le morphisme induit par φ. Nous avons vu au 1.3.10 que $\mathrm{Im}\psi = \mathrm{Im}\varphi$. Grâce au lemme précédent, nous pouvons maintenant décrire $\mathrm{Ker}\psi$: c'est précisément $\pi(\mathrm{Ker}\varphi)$. En effet, soit $g \in G$; on a

$$\pi(g) \in \mathrm{Ker}\psi \iff \psi(\pi(g)) = e \iff \varphi(g) = e \iff g \in \mathrm{Ker}\varphi\ ;$$

autrement dit, $\pi^{-1}(\mathrm{Ker}\psi) = \mathrm{Ker}\varphi$. On déduit alors du lemme 2.14.1 l'égalité $\mathrm{Ker}\psi = \pi(\mathrm{Ker}\varphi)$.

Lemme 2.14.3. — *Soit G un groupe, soit H un sous-groupe distingué de G et soit π le morphisme quotient de G dans G/H. Soit Γ un sous-groupe de G.*

(1) *Si $\Gamma \lhd G$ alors $\pi(\Gamma) \lhd G/H$.*

(2) *Si $\pi(\Gamma) \lhd G/H$ et si de plus $H \subset \Gamma$ alors $\Gamma \lhd G$, et le morphisme composé $G \to (G/H) \to (G/H)/\pi(\Gamma)$ induit un isomorphisme $G/\Gamma \simeq (G/H)/\pi(\Gamma)$.*

Démonstration. — Supposons que Γ soit distingué dans G. Soit $h \in \pi(\Gamma)$ et soit $x \in G/H$. Écrivons $h = \pi(\gamma)$ avec $\gamma \in \Gamma$ et $x = \pi(g)$ avec $g \in G$. On a

$$xhx^{-1} = \pi(g)\pi(\gamma)\pi(g^{-1}) = \pi(g\gamma g^{-1}).$$

Or $g\gamma g^{-1} \in \Gamma$ puisque $\Gamma \lhd G$; ainsi, $xhx^{-1} \in \pi(\Gamma)$ et $\pi(\Gamma) \lhd G/H$.

Supposons maintenant que $\pi(\Gamma) \lhd G/H$ et que Γ contient H. Le morphisme

$$G \to G/H \to (G/H)/\pi(\Gamma)$$

est surjectif comme composé de surjections ; et son noyau est égal à $\pi^{-1}(\pi(\Gamma))$, c'est-à-dire à Γ d'après le lemme 2.14.1. Ce dernier est donc distingué dans G, et il découle de 1.3.10 que le morphisme $G \to (G/H)/\pi(\Gamma)$ induit un isomorphisme $G/\Gamma \simeq (G/H)/\pi(\Gamma)$. □

Remarque 2.14.4. — Dans ce qui précède, nous avons, pour les besoins des énoncés et des démonstrations, donné un nom à la flèche quotient $G \to G/H$ (nous l'avons appelée π). Mais en pratique, on évite le plus souvent de le faire, pour ne pas introduire trop de notations ; dans ce cas, on désigne en général par $\overline{g}$ l'image d'un élément g de G dans G/H.

Et si Γ est un sous-groupe de G, son image dans G/H sera le plus souvent notée $\Gamma/(H \cap \Gamma)$, auquel elle s'identifie canoniquement (lemme 2.14.1). Avec cette convention, l'isomorphisme du lemme 2.14.3 (2) prend la forme plus suggestive

$$G/\Gamma \simeq (G/H)/(\Gamma/H).$$

2.15. Deux exemples matriciels. — Soit k un corps. Nous allons décliner 2.13.2 dans deux cas particuliers, où G sera à chaque fois un groupe de matrices.

Exemple 2.15.1. — Soit n un entier supérieur ou égal à 1. L'application déterminant induit un morphismes de groupes $\det\colon \mathrm{GL}_n(k) \to k^\times$ qui est surjectif car $n \geqslant 1$. Son noyau, c'est-à-dire l'ensemble des matrices de $\mathrm{GL}_n(k)$ dont le déterminant vaut 1, est noté $\mathrm{SL}_n(k)$ (c'est le groupe *spécial linéaire* en dimension n) ; il est évidemment distingué, et il découle de 2.13.2 que det induit un isomorphisme $\mathrm{GL}_n(k)/\mathrm{SL}_n(k) \simeq k^\times$.

Exemple 2.15.2. — Commençons par définir la *droite projective sur* k : c'est l'ensemble noté $\mathbf{P}^1(k)$ obtenu en adjoignant formellement à k un élément noté ∞ (nous verrons plus loin qu'on peut plus généralement définir pour tout entier n l'*espace projectif de dimension* n *sur* k, noté $\mathbf{P}^n(k)$; mais le cas $n = 1$ nous suffira pour le moment). Soit

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

une matrice appartenant à $\mathrm{GL}_2(k)$. On note h_M l'application de $\mathbf{P}^1(k)$ dans lui-même définie comme suit :

- ◇ $h_M(x) = \dfrac{ax+b}{cx+d}$ si $x \in k$ et si $cx + d \neq 0$;
- ◇ $h_M(x) = \infty$ si $x \in k$ et si $cx + d = 0$;
- ◇ $h_M(\infty) = a/c$ si $c \neq 0$;
- ◇ $h_M(\infty) = \infty$ si $c = 0$.

On vérifie aussitôt que $h_{\mathrm{I}_2} = \mathrm{Id}_{\mathbf{P}^1(k)}$, et que $h_{MN} = h_M \circ h_N$ pour tout couple (M, N) d'éléments de $\mathrm{GL}_2(k)$. En particulier, $h_M \circ h_{M^{-1}} = h_{M^{-1}} \circ h_M = \mathrm{Id}_{\mathbf{P}^1(k)}$ pour toute $M \in \mathrm{GL}_2(k)$. Par conséquent h_M est pour tout M une bijection de $\mathbf{P}^1(k)$ sur lui-même, et la formule $h_{MN} = h_M \circ h_N$ signifie dès lors que $M \mapsto h_M$ définit un morphisme de groupes de $\mathrm{GL}_2(k)$ sur $\mathfrak{S}_{\mathbf{P}^1(k)}$. L'image de ce morphisme est appelé le groupe des *homographies* de $\mathbf{P}^1(k)$; notons-le $\mathrm{H}(k)$ (ce n'est pas une notation standard).

Soit $a \in k^\times$. Il est immédiat que $h_{a\mathrm{I}_2} = \mathrm{Id}_{\mathbf{P}^1(k)}$. Réciproquement, soit

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

une matrice appartenant à $\mathrm{GL}_2(k)$ et telle que $h_M = \mathrm{Id}_{\mathbf{P}^1(k)}$. Comme $h_M(\infty) = \infty$ on a $c = 0$. Dans ce cas d est nécessairement non nul, et $h_M(x)$ est égal à $(ax + b)/d$ pour tout $x \in k$. On a donc $ax + b = dx$ pour tout $x \in k$, ou encore $(a - d)x + b = 0$. Ceci entraîne que $a = d$ et que $b = 0$ (faire par exemple $x = 0$ et $x = 1$). Autrement dit, $a \neq 0$ et la matrice M est égale à $a\mathrm{I}_2$.

Le noyau de $M \mapsto h_M$ est donc l'ensemble $\{a\mathrm{I}_2\}_{a\in k^\times}$ des matrices scalaires inversibles. On déduit alors de 2.13.2 que $M \mapsto h_M$ induit un isomorphisme

$$\mathrm{GL}_2(k)/\{a\mathrm{I}_2\}_{a\in k^\times} \simeq \mathrm{H}(k).$$

Remarque 2.15.3. — Pour tout entier n, l'ensemble $\{a\mathrm{I}_n\}_{a\in k^\times}$ est un sous-groupe distingué de $\mathrm{GL}_n(k)$ (c'est un exercice très facile). Le quotient $\mathrm{GL}_n(k)/\{a\mathrm{I}_n\}_{a\in k^\times}$ est en général noté $\mathrm{PGL}_n(k)$. Nous avons donc montré ci-dessus que $M \mapsto h_M$ induit un isomorphisme $\mathrm{PGL}_2(k) \simeq \mathrm{H}(k)$.

3. Propriétés du groupe Z et quelques conséquences

3.1. Brefs rappels sur les anneaux. — Nous allons commencer cette section par quelques rappels en théorie des anneaux commutatifs, sans démonstration ; vous pouvez essayer de les faire à titre d'exercice. Soit A un anneau commutatif (unitaire) et soit I un idéal de A ; c'est en particulier un sous-groupe additif de A.

3.1.1. *Structure d'anneau sur A/I.* — Si x et y sont deux éléments de A, la classe modulo I de xy ne dépend que des classes de x et y modulo I. La multiplication de A induit donc une loi interne supplémentaire sur le groupe abélien $(A/I, +)$, qui fait de celui-ci un anneau commutatif.

3.1.2. *Propriété universelle du quotient.* — Le morphisme quotient $\pi\colon A \to A/I$ est alors un morphisme d'anneaux de noyau I, et il satisfait la propriété universelle suivante : *pour tout anneau commutatif B, la formule $\psi \mapsto \psi \circ \pi$ établit une bijection entre l'ensemble des morphismes d'anneaux de A/I vers B et l'ensemble des morphismes de A vers B dont le noyau contient I.* Elle caractérise le morphisme $A \to A/I$ à unique isomorphisme près.

3.1.3. *Quotient par un sous-ensemble.* — Si E désigne un ensemble de générateurs de l'idéal I, la formule $\psi \mapsto \psi \circ \pi$ établit également une bijection entre l'ensemble des morphismes d'anneaux de A/I vers B et l'ensemble des morphismes de A vers B dont le noyau contient E.

Commentaires 3.1.4. — On se retrouve avec un phénomène analogue à ceux constatés en théorie des ensembles et en théorie des groupes : A/I apparaît à la fois comme l'anneau commutatif le plus général construit à partir de A en décrétant que les éléments de I sont nuls, et comme l'anneau commutatif le plus général construit à partir de A en se contentant de décréter que les éléments de E sont nuls. On observe donc ici aussi des «dommages collatéraux» : quand on force les éléments de E à être triviaux, on trivialise du même coup tous les éléments de I (mais les dégâts s'arrêtent là, le noyau de $A \to A/I$ étant précisément I).

3.1.5. *Idéaux de A/I.* — Les formules $J \mapsto \varphi(J)$ et $K \mapsto \varphi^{-1}(K)$ établissent une bijection entre l'ensemble des idéaux de A contenant I et l'ensemble des idéaux de A/I.

3.1.6. *L'isomorphisme fondamental.* — Soit $f\colon A \to B$ un morphisme d'anneaux commutatifs. Il induit un isomorphisme d'anneaux $(A/\mathrm{Ker} f) \simeq \mathrm{Im} f$.

3.2. Sommes et sommes directes internes dans les groupes abéliens. — Soit G un groupe abélien noté additivement et soit (G_i) une famille de sous-groupes de G.

3.2.1. — Il découle de 2.2.7.2 que le sous-groupe de G engendré par les G_i est l'ensemble des éléments de G de la forme $\sum g_i$ où (g_i) est une famille d'éléments de G presque tous nuls (en algèbre, *on ne sait faire que des sommes finies*) tels que $g_i \in G_i$ pour tout i. On dit que ce sous-groupe est la *somme* des G_i, et on le note

$\sum G_i$ On dit que les G_i sont en *somme directe* si tout élément de $\sum G_i$ a une unique écriture sous la forme $\sum g_i$ comme ci-dessus, et l'on écrit alors $\sum G_i = \bigoplus G_i$.

3.2.2. *Propriétés élémentaires.* — Nous laissons les preuves des faits suivants au lecteur ; elles sont analogues à celles qu'il a sans doute déjà rencontrées en algèbre linéaire.

3.2.2.1. — Pour que $\sum G_i = \bigoplus G_i$ il suffit de vérifier que $\sum g_i = 0 \Rightarrow (\forall i\ g_i = 0)$ pour toute famille (g_i) comme ci-dessus.

3.2.2.2. — Si G_1 et G_2 sont deux sous-groupes de G alors $G_1 + G_2 = G_1 \oplus G_2$ si et seulement si $G_1 \cap G_2 = \{0\}$.

3.2.3. *Somme d'idéaux.* — Soit A un anneau commutatif et soit $(J_i)_{i\in I}$ une famille d'idéaux de A. Il est immédiat que la somme $\sum J_i$ (au sens de 3.2.1) est encore un idéal de A.

3.3. Somme directe externe de groupes abéliens. — Nous allons maintenant définir une notion de somme directe qui diffère de la précédente. Cette dernière était *interne* : elle portait sur les sous-groupes d'un groupe donné. Celle que nous allons présenter maintenant est externe : elle porte sur une famille de groupes qui ne sont pas *a priori* plongés dans un même groupe.

3.3.1. *La somme directe externe : construction.* — Soit (G_i) une famille de groupes abéliens notés additivement. Soit H le sous-ensemble de $\prod_i G_i$ formé des éléments (g_i) tels que les g_i soient presque tous nuls ; c'est un sous-groupe de $\prod_i G_i$. Pour tout j, notons h_j l'application de G_j dans $\prod_i G_i$ qui envoie un élément γ sur la famille (g_i) telle que $g_j = \gamma$ et $i = 0$ si $i \neq j$. L'application h_j est un morphisme injectif de groupes.

Il résulte immédiatement de sa définition que le groupe H ci-dessus est la somme directe des $h_i(G_i)$. Comme h_i induit pour tout i un isomorphisme entre G_i et $h_i(G_i)$, on se permettra de dire que le groupe H est la somme directe *externe* des G_i, et d'écrire $H = \bigoplus_i G_i$. Cette construction *force* en quelque sorte les G_i à être contenus dans un même groupe H, et à être en somme directe dans ce dernier.

Remarque 3.3.2. — Attention : rien n'interdit à plusieurs des G_i d'être égaux à un même groupe G. Ils seront néanmoins considérés comme des sommandes distincts de la somme directe externe $\bigoplus G_i$; pour cette raison, on décrira parfois ces sommandes comme des *copies* de G.

3.3.3. — Supposons donné pour tout i un sous-groupe H_i de G_i. On vérifie immédiatement que la somme directe $\bigoplus H_i$ s'identifie à un sous-groupe de $\bigoplus G_i$, et que l'on a un isomorphisme naturel $\bigoplus G_i / \bigoplus H_i \simeq \bigoplus (G_i/H_i)$.

Remarque 3.3.4. — Lorsque la famille (G_i) est finie, la somme directe $\bigoplus G_i$ coïncide avec le produit $\prod G_i$. On pourra selon le contexte (ou selon ses goûts) préférer l'une ou l'autre des notations.

3.4. Étude du groupe Z : premières propriétés. — Nous allons maintenant entamer l'étude du groupe abélien $\mathbf{Z}$ et établir quelques unes de ses propriétés qui sont à la base de l'arithmétique.

Remarque 3.4.1. — Soit G un groupe abélien noté additivement, soit $g \in G$ et soit $n \in \mathbf{Z}$. L'élément ng de G est alors par définition la somme de n termes égaux à g si $n \geqslant 0$, et la somme de $(-n)$ termes égaux à $(-g)$ sinon. Mais lorsque G est lui-même égal à $\mathbf{Z}$, cet élément coïncide le produit de n et g au sens de la multiplication de $\mathbf{Z}$. En particulier, tout sous-groupe de $(\mathbf{Z}, +)$ est automatiquement stable par multiplication externe par les éléments de $\mathbf{Z}$, et est donc un idéal de $\mathbf{Z}$.

Soit $d \in \mathbf{Z}$. Le sous-groupe de $\mathbf{Z}$ engendré par d (qui est aussi en vertu de ce qui précède l'idéal principal de $\mathbf{Z}$ engendré par d) n'est autre que $d\mathbf{Z} := \{dn\}_{n\in\mathbf{Z}}$. Soit $d' \in \mathbf{Z}$. Il est immédiat qu'on a les équivalences

$$d\mathbf{Z} \subset d'\mathbf{Z} \iff d'|d$$

et

$$(d\mathbf{Z} = d'\mathbf{Z}) \iff (d|d' \text{ et } d'|d) \iff \exists \varepsilon \in \{-1, 1\}, d' = \varepsilon d.$$

Par conséquent, le générateur d'un idéal principal de $\mathbf{Z}$ est uniquement déterminé au signe près (ce fait s'étend à tout anneau commutatif intègre, à condition de remplacer «au signe près» par «à un inversible près»). Il peut donc toujours être choisi dans $\mathbf{N}$, et est alors *unique.*

Lemme 3.4.2. — *Tout sous-groupe de* $\mathbf{Z}$ *est de la forme* $d\mathbf{Z}$ *pour un unique* $d \in \mathbf{N}$.

Démonstration. — L'unicité de d a été mentionnée à la remarque 3.4.1 ci-dessus. Montrons maintenant son existence. Soit G un sous-groupe de $\mathbf{Z}$. Si $G = \{0\}$ alors $G = 0\mathbf{Z}$. Supposons maintenant G non nul. Soit G^+ l'ensemble des éléments strictement positifs de G. Puisque $G \neq \{0\}$ il existe $g \neq 0$ dans G. Si $g > 0$ il appartient à G^+, et si $g < 0$ alors $-g \in G^+$; ainsi, G^+ est non vide. Soit d son plus petit élément. Puisque $d \in G$, on a l'inclusion $d\mathbf{Z} \in G$. Nous allons établir l'inclusion réciproque.

Soit $g \in G$. Par division euclidienne, il existe $q \in \mathbf{Z}$ et $r \in \{0, \ldots, d-1\}$ tel que $g = qd + r$. Puisque $d\mathbf{Z} \subset G$, l'entier positif $r = g - qd$ appartient à G ; comme $r < d$, la définition même de d entraîne que $r = 0$. Ainsi $g = qd$ et $G \subset d\mathbf{Z}$. □

Le lemme précédent assure en particulier que tout idéal de l'anneau commutatif intègre $\mathbf{Z}$ est principal ; l'anneau $\mathbf{Z}$ est donc ce qu'on appelle un *anneau principal.* Les propriétés que nous allons maintenant énoncer et démontrer pour $\mathbf{Z}$ valent en fait pour tout anneau principal, avec essentiellement les mêmes preuves.

3.4.3. *Le PGCD.* — Soit $(a_i)_{i\in I}$ une famille d'éléments de $\mathbf{Z}$ et soit n un élément de $\mathbf{Z}$. L'élément n divise chacun des a_i si et seulement si $a_i \in n\mathbf{Z}$ pour tout i. Cela revient à demander que $n\mathbf{Z}$ contienne l'idéal $\sum_i a_i\mathbf{Z}$ engendré par les a_i. Ce dernier est de la forme $d\mathbf{Z}$ pour un entier $d \in \mathbf{Z}$ uniquement déterminé au signe près. Par conséquent, n divise chacun des a_i si et seulement si $d\mathbf{Z} \subset n\mathbf{Z}$, c'est-à-dire si et seulement si n divise d. L'entier d est appelé le plus grand commun diviseur (PGCD) des a_i.

Remarquons que pour que le PGCD d des a_i soit non nul, il faut et il suffit que $\sum a_i\mathbf{Z}$ soit non nul, c'est-à-dire qu'il existe i tel que $a_i \neq 0$. Si c'est le cas, l'égalité $\sum a_i\mathbf{Z} = d\mathbf{Z}$ implique que $\sum(a_i/d)\mathbf{Z} = \mathbf{Z}$; le PGCD des (a_i/d) vaut donc 1.

Si a et b sont deux éléments de $\mathbf{Z}$ et si d désigne leur PGCD, on a par définition l'égalité $a\mathbf{Z} + b\mathbf{Z} = d\mathbf{Z}$; il s'ensuit qu'il existe u et v dans $\mathbf{Z}$ tels que $au + bv = d$ (relation de Bezout). On dit que a et b sont premiers entre eux si $d = 1$. Cela revient à demander que $a\mathbf{Z} + b\mathbf{Z} = \mathbf{Z}$; il suffit pour cela que $a\mathbf{Z} + b\mathbf{Z}$ contienne 1, c'est-à-dire qu'il existe u et v dans $\mathbf{Z}$ tels que $au + bv = 1$.

3.4.4. *Le PPCM.* — Soit $(a_i)_{i\in I}$ une famille d'éléments de $\mathbf{Z}$ et soit n un élément de $\mathbf{Z}$. L'élément n est multiple de chacun des a_i si et seulement si $n \in a_i\mathbf{Z}$ pour tout i. Cela revient à demander que $n\mathbf{Z}$ soit contenu dans $\bigcap_i a_i\mathbf{Z}$. Ce dernier est de la forme $d\mathbf{Z}$ pour un entier $d \in \mathbf{Z}$ uniquement déterminé au signe près. Par conséquent, n est multiple de chacun des a_i si et seulement si $n\mathbf{Z} \subset d\mathbf{Z}$, c'est-à-dire si et seulement si n est multiple de d. L'entier d est appelé le plus petit commun multiple (PPCM) des a_i.

Si l'un des a_i est nul, le PPCM des a_i est nul. La réciproque est fausse : par exemple, le PPMC de tous les entiers > 0 est mutliple de tout entier > 0, donc est nul. Par contre si la famille (a_i) est finie et si le PPCM des a_i est nul, le produit des a_i est nul (car il est multiple de leur PPCM), si bien que l'un des a_i au moins est nul.

Lemme de Gauß 3.4.5. — *Si a, b et c sont trois éléments de $\mathbf{Z}$ tels que $a|bc$ et tels que a soit premier avec b, alors a divise c.*

Démonstration. — Soit $n \in \mathbf{Z}$ tel que $bc = an$; choisissons une relation de Bézout $au + bv = 1$. On a alors

$$c = c(au + bv) = auc + bcv = auc + anv = a(uc + nv).$$

□

Corollaire 3.4.6. — *Soient $a_1, \dots, a_r$ et b des éléments de $\mathbf{Z}$. Supposons que b est premier avec chacun des a_i ; il est alors premier avec $a_1 \dots a_r$.*

Démonstration. — On raisonne par récurrence sur r. Si $r = 0$ alors b est premier avec $a_1 \dots a_r$ car ce dernier est égal à 1, et le corollaire est vrai. Supposons $r > 0$ et le corollaire vrai pour les entiers $< r$. Soit d le PGCD de b et de $\prod a_i$. Par définition, d divise $a_1 \dots a_r$. Par ailleurs tout diviseur commun de d et de l'un des a_j est un diviseur commun de b et a_j, et vaut donc 1 ou (-1) ; par conséquent, d est premier avec chacun des a_j. Puisque d est premier avec a_1 et divise $\prod a_i$, le lemme de Gauß assure que d divise $a_2 \dots a_r$. Comme d est premier avec chacun des a_j, l'hypothèse de récurrence assure que d est premier avec $a_2 \dots a_r$; puisqu'il divise ce dernier, d vaut 1 ou (-1). □

Définition 3.4.7. — Rappelons qu'on appelle *nombre premier* tout élément p de $\mathbf{N}$ qui est > 1 et qui n'admet pour seuls diviseurs que 1 et lui-même.

Théorème 3.4.8 **(Écriture comme produit de nombres premiers)**
Soit n un élément non nul de $\mathbf{Z}$. Il possède une écriture sous la forme $\varepsilon \prod_{i=1}^{n} p_i^{n_i}$

où $\varepsilon \in \{1, -1\}$, où les p_i sont des nombres premiers deux à deux distincts et où les n_i sont des entiers > 1. Une telle écriture est unique à permutation près des p_i.

Démonstration. — Montrons d'abord l'existence d'une telle écriture par récurrence sur $|n|$. Si $|n| = 1$ alors $n = 1$ ou $n = -1$ et le théorème est vrai (dans les deux cas $\varepsilon = n$ et la famille des p_i est *vide*).“ Supposons $|n| > 1$ et le théorème vrai pour les entiers de valeur absolue $< |n|$. Posons $\varepsilon = 1$ si n est positif, et $\varepsilon = -1$ sinon. Si $|n|$ est premier l'écriture $n = \varepsilon|n|$ est du type souhaité. Sinon on peut écrire $|n|n = m\ell$ où m et ℓ sont deux entiers strictement compris entre 1 et $|n|$. En vertu de l'hypothèse de récurrence, m et ℓ sont tous deux produits d'un nombre fini de nombres premiers, et $n = \varepsilon|n| = \varepsilon m\ell$ possède donc une écriture de la forme requise.

Montrons maintenant l'unicité. Celle de ε est claire : c'est le signe de n. Il reste à s'assurer que si $p_1 \dots p_r = q_1 \dots q_s$, où les p_i et les q_j sont des nombres premiers (pas forcément deux à deux distincts) alors $r = s$ et il existe une permutation σ de $\{1, \dots, r\}$ telle que $q_i = p_{\sigma(i)}$ pour tout i. On procède par récurrence sur m. Si $m = 0$ la famille des p_i est vide, et $q_1 \dots q_s = 1$; comme un nombre premier est par définition strictement supérieur à 1, cette dernière égalité force s à être nul, et la famille des q_j à être vide, ce qu'il fallait établir. Supposons maintenant $m > 0$ et l'assertion vraie pour $m - 1$. L'entier p_1 divise $q_1 \dots q_s$. Il est alors égal à l'un des q_j : en effet, dans le cas contraire p_1 serait premier à chacun des q_j et partant premier à $q_1 \dots q_s$ (corollaire 3.4.6). On divise alors par p_1 les deux membres de l'égalité et on conclut en appliquant l'hypothèse de récurrence. □

Remarque 3.4.9. — La preuve de l'*existence* de la décomposition en produit de facteurs premiers est élémentaire et n'utilise pas la principalité de $\mathbf{Z}$. Cette existence n'a en fait rien de particulièrement remarquable : on peut démontrer plus généralement que dans n'importe quel anneau commutatif intègre noethérien, tout élément non nul est produit d'une famille finie d'éléments irréductibles.

C'est l'*unicité* de la décomposition qui fait sa force ; sa preuve repose sur le lemme de Gauß, c'est-à-dire *in fine* sur les relations de Bézout et donc la principalité de $\mathbf{Z}$.

Lemme 3.4.10. — *Soient a et b deux éléments de $\mathbf{Z}$. On a (au signe près) l'égalité $ab = \mathrm{PGCD}(a, b) \cdot \mathrm{PPCM}(a, b)$.*

Démonstration. — Soit d le PGCD de a et b et soit m leur PPCM. Choisissons une relation de Bézout $au + bv = d$. Si a et b sont nuls alors $d = m = 0$ et le lemme est évident. Supposons que a et b ne sont pas tous deux nuls. Dans ce cas $d \neq 0$; posons $\alpha = a/d$ et $\beta = b/d$. Nous allons démontrer que le PPCM de (a, b) est égal à $d\alpha\beta$ (au signe près), ce qui permettra de conclure car $ab = d^2\alpha\beta$. Comme $d\alpha = a$ et $d\beta = b$, le produit $d\alpha\beta$ est à la fois multiple de a et de b, et est donc multiple de m. Il suffit dès lors de prouver que m est multiple de $d\alpha\beta$. Par définition, m est multiple de a et de b (et *a fortiori* de d) ; écrivons $m = xa = yb$ avec x et y dans $\mathbf{Z}$. On a alors

$$m = d\frac{m}{d} = (au + bv)\frac{m}{d} = \underbrace{yu\frac{ab}{d}}_{\text{car } m=yb} + \underbrace{xv\frac{ab}{d}}_{\text{car } m=xa} = (yu + xv)d\alpha\beta.$$

□

3.4.11. — Soient $a_1, \dots, a_m$ des éléments de $\mathbf{Z}$. La famille des réductions modulo les différents a_i définit un morphisme d'anneaux $\mathbf{Z} \to \mathbf{Z}/a_1\mathbf{Z} \times \dots \times \mathbf{Z}/a_m\mathbf{Z}$. Comme son noyau contient visiblement $a_1 \dots a_m\mathbf{Z}$, il induit un morphisme d'anneaux

$$\mathbf{Z}/(a_1 \dots a_m\mathbf{Z}) \to \mathbf{Z}/a_1\mathbf{Z} \times \dots \times \mathbf{Z}/a_m.$$

Lemme chinois 3.4.12. — *Soit $(a_1, \dots, a_m)$ une famille d'éléments de $\mathbf{Z}$ deux à deux premiers entre eux. Le morphisme d'anneaux naturel*

$$\mathbf{Z}/(a_1 \dots a_m\mathbf{Z}) \to \mathbf{Z}/a_1\mathbf{Z} \times \dots \times \mathbf{Z}/a_m$$

est un isomorphisme.

Démonstration. — On procède par récurrence sur m, le cas $m = 0$ étant trivial (on a l'anneau nul des deux côtés). Supposons $m \geqslant 1$ et le résultat vrai pour les entiers strictement inférieurs à m. Posons $b = a_2 \dots a_m$. L'hypothèse de récurrence assure que le morphisme naturel $\mathbf{Z}/b\mathbf{Z} \to \mathbf{Z}/a_2\mathbf{Z} \times \dots \times \mathbf{Z}/a_m\mathbf{Z}$ est un isomorphisme. Il suffit donc de montrer que le morphisme naturel $\pi \colon \mathbf{Z}/(a_1b)\mathbf{Z} \to \mathbf{Z}/a_1\mathbf{Z} \times \mathbf{Z}/b\mathbf{Z}$ est un isomorphisme.

Comme a_1 est premier avec les a_j pour $j > 1$, il est premier avec b (corollaire 3.4.6). Choisissons une relation de Bézout $a_1u + bv = 1$.

Injectivité de π. Soit n un entier. L'image de n dans $\mathbf{Z}/a_1\mathbf{Z} \times \mathbf{Z}/b\mathbf{Z}$ est nulle si et seulement si n est à la fois multiple de a_1 et de b, donc si et seulement si n est multiple du PPCM de a_1 et b. Mais comme a_1 et b sont premiers entre eux ce PPCM vaut a_1b d'après le lemme 3.4.10. Le noyau de $\mathbf{Z} \to \mathbf{Z}/a_1\mathbf{Z} \times \mathbf{Z}/b\mathbf{Z}$ est donc égal à à $a_1b\mathbf{Z}$, ce qui entraîne l'injectivité de π.

Surjectivité de π. Soient x et y deux éléments de $\mathbf{Z}$. Posons $z = ya_1u + xbv$. En écrivant $bv = 1 - a_1u$ on voit que z est égal à x modulo a_1 ; en écrivant $a_1u = 1 - bv$ on voit que z est égal à y modulo b. Par conséquent, π est surjective. □

3.4.13. *Description de* $\mathbf{Z}/n\mathbf{Z}$. — Le morphisme quotient $\mathbf{Z} \to \mathbf{Z}/0\mathbf{Z}$ est un isomorphisme, (en particulier, $\mathbf{Z}/0\mathbf{Z}$ est infini). Soit $n \geqslant 1$. Par la théorie de la division euclidienne, tout élément de $\mathbf{Z}$ est congru à un et un seul élément de $\{0, \dots, n-1\}$. Par conséquent, les éléments de $\mathbf{Z}/n\mathbf{Z}$ sont les classes $\overline{0}, \dots, \overline{n-1}$ qui sont deux à deux distinctes ; le cardinal de $\mathbf{Z}/n\mathbf{Z}$ est donc égal à n.

3.5. Propriété universelle de $\mathbf{Z}/d\mathbf{Z}$. — Le but de ce qui suit est de décrire les morphismes de $\mathbf{Z}/d\mathbf{Z}$ vers un groupe donné G, en commençant par le cas où $d = 0$, c'est-à-dire par les morphismes de $\mathbf{Z}$ dans G.

Définition 3.5.1. — Soit G un groupe, soit $g \in G$ et soit $n \in \mathbf{Z}$. On dit que g est *de n-torsion* si $g^n = e$.

3.5.2. *Le cas abélien.* — Soit G un groupe abélien noté additivement et soit $n \in \mathbf{Z}$. L'application $g \mapsto ng$ de multiplication par n est alors un endomorphisme de G. Son image est notée nG, et son noyau est précisément l'ensemble des éléments de n-torsion de G. Ce dernier est donc un sous-groupe de G.

Remarque 3.5.3. — On prendra garde que si G n'est pas abélien, les éléments de n-torsion de G ne forment pas un sous-groupe en général. Par exemple, le lecteur vérifiera que dans $\mathrm{GL}_2(\mathbf{R})$, les matrices

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \text{ et } B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

sont de 2-torsion, mais que le produit AB n'est pas de 2-torsion.

3.5.4. *La propriété universelle de* $\mathbf{Z}$. — Soit G un groupe. Soit f un morphisme de $\mathbf{Z}$ dans G, et soit g l'image de 1. Pour tout $n \in \mathbf{Z}$ on a alors nécessairement $f(n) = f(n \cdot 1) = g^n$.

Réciproquement, soit $g \in G$. Comme $g^{n+m} = g^n g^m$ pour tout $(n, m) \in \mathbf{Z}^2$, l'application $n \mapsto g^n$ de $\mathbf{Z}$ dans G est un morphisme de groupes envoyant 1 sur g.

Récapitulons : $f \mapsto f(1)$ établit une bijection entre l'ensemble des morphismes de groupes de $\mathbf{Z}$ dans G et l'ensemble des éléments de G ; la bijection réciproque associe à un élément g de G le morphisme $n \mapsto g^n$.

Si G est abélien on vérifie immédiatement que cette bijection est un morphisme de groupes (2.6.1).

3.5.5. *La propriété universelle de* $\mathbf{Z}/d\mathbf{Z}$. — Soit $d \in \mathbf{Z}$ et soit G un groupe. Il découle de 2.12.1 (et du fait que comme $\mathbf{Z}$ est abélien, $d\mathbf{Z}$ est le plus petit sous-groupe distingué de $\mathbf{Z}$ contenant d) que $\psi \mapsto (n \mapsto \psi(\overline{n}))$ établit une bijection entre l'ensemble des morphismes de $\mathbf{Z}/d\mathbf{Z}$ dans G et l'ensemble des morphismes de $\mathbf{Z}$ dans G s'annulant sur d.

On en déduit à l'aide de 3.5.4 que $f \mapsto f(\overline{1})$ établit une bijection entre l'ensemble des morphismes de $\mathbf{Z}/d\mathbf{Z}$ dans G et l'ensemble des éléments g de d-torsion ; la bijection réciproque envoie un élément g tel que $g^d = e$ sur le morphisme $\overline{n} \mapsto g^n$ (comme $g^d = e$ l'élément g^n de G ne dépend bien que de la classe de n modulo d).

Si G est abélien on vérifie immédiatement que cette bijection est un morphisme de groupes (*cf.* 2.6.1 et 3.5.2).

3.5.6. *Énoncés informels.* — Les propriétés universelles énoncées en 3.5.4 et 3.5.5 peuvent se résumer peu ou prou par les slogans suivants : *se donner un morphisme de* $\mathbf{Z}$ *dans* G*, c'est choisir un élément de* G *– l'image de* 1 *; se donner un morphisme de* $\mathbf{Z}/d\mathbf{Z}$ *dans* G*, c'est choisir un élément de* d*-torsion de* G *– l'image de* $\overline{1}$.

3.5.7. *Endomorphismes de* $\mathbf{Z}$. — Il résulte de 3.5.4, appliqué avec $G = \mathbf{Z}$, que $a \mapsto h_a$ établit un isomorphisme de groupes entre $\mathbf{Z}$ et $\mathrm{End}\ \mathbf{Z}$, où h_a désigne l'endomorphisme $x \mapsto ax$ (l'homothétie de rapport a).

On vérifie aisément que cet isomorphisme de groupes est même un isomorphisme d'anneaux (2.6.2). Le groupe $\mathrm{Aut}\ \mathbf{Z}$ s'identifie donc *via* $a \mapsto h_a$ à $\mathbf{Z}^\times = \{-1, 1\}$.

3.5.8. *Endomorphismes de* $\mathbf{Z}/d\mathbf{Z}$. — Il résulte de 3.5.5, appliqué avec $G = \mathbf{Z}/p\mathbf{Z}$, que $a \mapsto h_a$ établit une bijection entre $\mathbf{Z}/d\mathbf{Z}$ et $\mathrm{End}\ \mathbf{Z}/d\mathbf{Z}$, où h_a désigne l'endomorphisme $x \mapsto ax$ (l'homothétie de rapport a).

On vérifie aisément que cet isomorphisme de groupes est même un isomorphisme d'anneaux (2.6.2). Le groupe $\mathrm{Aut}\ \mathbf{Z}/d\mathbf{Z}$ s'identifie donc *via* $a \mapsto h_a$ à $(\mathbf{Z}/d\mathbf{Z})^\times$.

3.6. Ordre d'un élément, groupes monogènes et groupes cycliques. — Nous allons appliquer les résultats que nous venons d'obtenir sur $\mathbf{Z}$ et ses quotients à l'étude de tous les groupes.

Définition 3.6.1. — Soit G un groupe et soit $g \in G$. On appelle *ordre* de g le cardinal du sous-groupe $\langle g \rangle = \{g^n\}_{n \in \mathbf{Z}}$, vu comme élément de $\mathbf{N} \cup \{+\infty\}$.

3.6.2. *Premières propriétés.* — Soit G un groupe et soit g un élément de G. Soit φ l'unique morphisme de $\mathbf{Z}$ dans G envoyant 1 sur g ; on a $\varphi(g) = g^n$ pour tout $n \in \mathbf{Z}$, si bien que $\mathrm{Im}\varphi = \langle g \rangle$.

Le noyau de φ est un sous-groupe de $\mathbf{Z}$; il s'écrit donc $d\mathbf{Z}$ pour un unique $d \in \mathbf{N}$. Il s'ensuit que φ induit un isomorphisme $\mathbf{Z}/d\mathbf{Z} \simeq \langle g \rangle$. Il résulte dès lors de 3.4.13 que l'ordre de g est infini si $d = 0$, et égal à d sinon. Plaçons-nous dans ce dernier cas. L'ordre d de g peut alors être caractérisé comme le plus petit entier $n > 0$ tel que $g^n = e$, et si $n \in \mathbf{Z}$ on a $g^n = e$ si et seulement si $d|n$.

3.6.3. — Soit G un groupe fini dont on note n le cardinal. Si $g \in G$, le groupe $\langle g \rangle$ est fini, et l'ordre de g est donc nécessairement fini, et divise n d'après le lemme 2.9.2. Ceci entraîne en vertu de 3.6.2 que $g^n = e$: *dans un groupe de cardinal n, tout élément est de n-torsion.*

Définition 3.6.4. — Un groupe G est dit *monogène* s'il existe $g \in G$ tel que $G = \langle g \rangle$.

Exemples 3.6.5. — Le groupe $\mathbf{Z}$ est monogène (il est engendré par 1). Soit $d \in \mathbf{Z}$. Puisque le morphisme $n \mapsto \overline{n}$ de $\mathbf{Z}$ dans $\mathbf{Z}/d\mathbf{Z}$ est surjectif, le groupe $\mathbf{Z}/d\mathbf{Z}$ est engendré par $\overline{1}$ et est donc monogène.

3.6.6. — Les exemples ci-dessus sont en fait les seuls exemples de groupes monogènes. Il résulte en effet de 3.6.2 qu'un groupe monogène G est isomorphe à $\mathbf{Z}/d\mathbf{Z}$ pour un certain $d \in \mathbf{N}$. Si c'est le cas, G est fini si et seulement si $d > 0$; l'entier d est alors égal au cardinal de G, et l'on dit que G est *cyclique.*

3.6.7. *À propos des groupes cycliques.* — Soit G un groupe fini et soit d son cardinal. Si G cyclique, il est engendré par un élément dont l'ordre est nécessairement égal à d. Réciproquement, si G possède un élément g d'ordre d alors le cardinal de $\langle g \rangle$ est égal à d, ce qui entraîne que $G = \langle g \rangle$; ainsi, G est cyclique.

Supposons maintenant que d est premier et soit g un élément de G différent de e. Comme l'ordre de g divise d et est différent de 1 (puisque $g \neq e$), il est exactement égal à d ; par ce qui précède, G est cyclique.

3.7. Sous-groupes de $\mathbf{Z}/n\mathbf{Z}$. — Soit n un entier $\geqslant 1$; nous allons faire une étude détaillée des sous-groupes de $\mathbf{Z}/n\mathbf{Z}$ et de l'ordre de ses éléments. Nous utiliserons très souvent implicitement le fait suivant : si a et b sont deux éléments de $\mathbf{Z}$ on a dans $\mathbf{Z}/n\mathbf{Z}$ les égalités

$$a\overline{b} = \overline{ab} = \overline{a}\overline{b}.$$

Précisons que la notation $a\overline{b}$ est ici une simple occurrence de la notation ag qui a un sens pour tout élément g d'un groupe abélien noté additivement, et que $\overline{a}\overline{b}$ désigne le produit de $\overline{a}$ et $\overline{b}$ dans l'anneau $\mathbf{Z}/n\mathbf{Z}$. L'égalité $a\overline{b} = \overline{ab}$ vient du fait que la réduction modulo n est un morphismes entre groupes abéliens notés additivement, et partant commute à la multiplication par a ; la seconde égalité provient du fait que la réduction modulo n est un morphisme d'anneaux ; et l'on a par ailleurs implicitement utilisé la double interprétation du produit ab (remarque 3.4.1).

3.7.1. — Soit G un sous-groupe de $\mathbf{Z}/n\mathbf{Z}$ et soit Γ son image réciproque dans $\mathbf{Z}$. On peut écrire $\Gamma = a\mathbf{Z}$ pour un (unique) $a \in \mathbf{N}$. Le groupe G étant égal à l'image de Γ (lemme 2.14.1), il vient $G = \langle \overline{a} \rangle$.

3.7.2. — Soit $a \in \mathbf{Z}$ et soit $r \in \mathbf{N}$ le PGCD de a et n. L'image réciproque de $\langle \overline{a} \rangle$ dans $\mathbf{Z}$ est égale à $a\mathbf{Z} + n\mathbf{Z} = r\mathbf{Z}$. En vertu du lemme 2.14.1, le groupe $\langle \overline{a} \rangle$ est alors égal à l'image de $r\mathbf{Z}$ dans $\mathbf{Z}/n\mathbf{Z}$, c'est-à-dire à $\langle \overline{r} \rangle$. Et d'après le lemme 2.14.3, le quotient $(\mathbf{Z}/n\mathbf{Z})/\langle \overline{a} \rangle)$ s'identifie canoniquement à $\mathbf{Z}/r\mathbf{Z}$.

L'intérêt de cette remarques est le suivant. Comme r divise n, l'ordre de $\overline{r}$ dans $\mathbf{Z}/n\mathbf{Z}$ est très facile à calculer ; en effet, si m est un entier on a

$$m\overline{r} = \overline{0} \iff n \text{ divise } mr \iff (n/r) \text{ divise } m.$$

L'ordre de $\overline{r}$ dans $\mathbf{Z}/n\mathbf{Z}$ est donc égal n/r.

3.7.3. *Récapitulation : description des sous-groupes de* $\mathbf{Z}/n\mathbf{Z}$. — Il résulte de ce qui précède que pour tout diviseur d de n il existe un et un seul sous-groupe de cardinal d de $\mathbf{Z}/n\mathbf{Z}$; il est cyclique, engendré par $\overline{n/d}$; le quotient correspondant de $\mathbf{Z}/n\mathbf{Z}$ s'identifie canoniquement à $\mathbf{Z}/\frac{n}{d}\mathbf{Z}$.

3.7.4. *Relations d'inclusion entre les sous-groupes de* $\mathbf{Z}/n\mathbf{Z}$. — Soient d et d' deux diviseurs de n ; soit G_d (resp. $G_{d'}$) l'unique sous-groupe de $\mathbf{Z}/n\mathbf{Z}$ de cardinal d (resp. d'). On a alors $G_d \subset G_{d'}$ si et seulement si $d|d'$.

En effet, G_d est engendré par $\overline{n/d}$; son image réciproque Γ_d dans $\mathbf{Z}$ est donc égale à $(n/d)\mathbf{Z} + n\mathbf{Z} = (n/d)\mathbf{Z}$ (car $n\mathbf{Z} \subset (n/d)\mathbf{Z}$). De même, l'image réciproque $\Gamma_{d'}$ de $G_{d'}$ dans $\mathbf{Z}$ est égale à $(n/d')\mathbf{Z}$.

D'après le lemme 2.14.1, $G_d \subset G_{d'}$ si et seulement si $\Gamma_d \subset \Gamma_{d'}$, c'est-à-dire si et seulement si $(n/d)\mathbf{Z} \subset (n/d')\mathbf{Z}$. Mais ceci revient à demander que n/d' divise n/d, c'est-à-dire encore que d divise d'.

3.7.5. *Sous-groupe de r-torsion de* $\mathbf{Z}/n\mathbf{Z}$. — Soit r un entier, et soit d le PGCD de n et r ; posons $\nu = n/d$ et $\rho = r/d$ (notons que d est non nul car $n \neq 0$) ; les entiers ν et ρ sont premiers entre eux.

Soit T le sous-groupe de $\mathbf{Z}/n\mathbf{Z}$ formé des éléments de r-torsion. Son image réciproque Θ dans $\mathbf{Z}$ est l'ensemble des entiers relatifs m tels que n divise rm, c'est-à-dire encore tels que ν divise ρm. Comme ν est premier avec ρ, on peut par le lemme de Gauß décrire également Θ comme l'ensemble des éléments m de $\mathbf{Z}$ tels que ν divise m. On a donc $\Theta = \nu\mathbf{Z} = (n/d)\mathbf{Z}$; on déduit alors du lemme 2.14.1 que T est le sous-groupe de $\mathbf{Z}/n\mathbf{Z}$ engendré par $\overline{n/d}$, c'est-à-dire l'unique sous-groupe de $\mathbf{Z}/n\mathbf{Z}$ de cardinal d.

3.7.6. *Générateurs de* $\mathbf{Z}/n\mathbf{Z}$. — Soit $a \in \mathbf{Z}$. On déduit de 3.7.2 que $\overline{a}$ engendre $\mathbf{Z}/n\mathbf{Z}$ si et seulement si a est premier avec n, c'est-à-dire encore si et seulement si il existe

u et v dans $\mathbf{Z}$ tels que $au + nv = 1$; autrement dit, c'est le cas si et seulement si $\overline{a}$ est inversible modulo n.

Le nombre de générateurs de $\mathbf{Z}/n\mathbf{Z}$ est donc égal au nombre d'entiers compris entre 0 et $n-1$ qui sont premiers à n, ou encore inversibles modulo n. Nous noterons ce nombre $\Phi(n)$; la fonction Φ est appelée *l'indicateur d'Euler*.

Si n est de la forme p^m avec p premier et $m \geqslant 1$ un calcul direct (fondé sur le fait qu'un entier est premier avec p^m si et seulement si il n'est pas multiple de p) montre que $\Phi(n) = p^{m-1}(p-1)$. En général, écrivons $n = \prod p_i^{m_i}$ avec les p_i premiers et deux à deux distincts, et les $m_i \geqslant 1$. On dispose d'un isomorphisme d'anneaux $\mathbf{Z}/n\mathbf{Z} \simeq \prod(\mathbf{Z}/p_i^{m_i}\mathbf{Z})$ donné par le lemme chinois. L'interprétation de $\Phi(n)$ en termes d'éléments inversibles assure alors que

$$\Phi(n) = \prod_i \Phi(p_i^{m_i}) = \prod_i p_i^{m_i-1}(p_i - 1).$$

3.7.7. — Soit d un diviseur de n. Un élément de $\mathbf{Z}/n\mathbf{Z}$ est d'ordre d si et seulement si il engendre un sous-groupe de cardinal d, donc si et seulement si il engendre l'unique sous-groupe de $\mathbf{Z}/n\mathbf{Z}$ de cardinal d, à savoir $\langle\overline{n/d}\rangle$. Les éléments d'ordre d de $\mathbf{Z}/n\mathbf{Z}$ sont donc exactement les générateurs du groupe cyclique $\langle\overline{n/d}\rangle$, qui est isomorphe à $\mathbf{Z}/d\mathbf{Z}$. Il y en a en conséquence exactement $\Phi(d)$.

Puisque tout élément de $\mathbf{Z}/n\mathbf{Z}$ est d'ordre divisant n, il vient

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

3.8. Exposant d'un groupe abélien fini. — Nous abordons maintenant l'étude générale des groupes abéliens fini. Nous allons commencer par l'étude d'un invariant important d'un tel groupe, son *exposant*.

Définition 3.8.1. — Soit G un groupe abélien fini noté additivement, et soit I l'ensemble des entiers d tels que $dg = 0$ pour tout $g \in G$. On vérifie immédiatement que I est un idéal de $\mathbf{Z}$. Soit e l'entier $\geqslant 0$ tel que $I = e\mathbf{Z}$. On dit que e est l'*exposant* de G.

3.8.2. *Exposant et cardinal.* — Soit G un groupe fini noté additivement et soit n son cardinal. Comme $ng = 0$ pour tout $g \in G$ (3.6.3) l'entier e divise n. Cette relation de divisibilité peut être stricte : par exemple si $G = \mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$, on a $e = 2$ et $n = 4$.

3.8.3. *Autre expression de l'exposant.* — Pour tout $g \in G$, soit I_g l'ensemble des entiers d tels que $dg = 0$. C'est un idéal de $\mathbf{Z}$ dont le générateur positif est l'ordre de g (3.6.2). Puisque I est l'intersection des I_g pour g parcourant G, l'exposant de G est le PPCM des ordres des éléments de G.

Lemme 3.8.4. — *Soit G un groupe abélien noté additivement et soient g et h deux éléments de G dont les ordres respectifs a et b sont finis et premiers entre eux. L'ordre de $g+h$ est alors égal à ab.*

Démonstration. — Soit d l'ordre de $g+h$. On a $ab(g+h) = bag + abh = 0$; par conséquent, d divise ab. Pour conclure, il suffit de montrer que ab divise d.

Par définition de d, on a $d(g+h)=0$, c'est-à-dire $dg=-dh$. L'élément $dg=-dh$ de G appartient donc au sous-groupe $H:=\langle g\rangle\cap\langle h\rangle$ de G. Comme H est contenu à la fois dans $\langle g\rangle$ et dans $\langle h\rangle$, son cardinal divise a et b. Puisque a et b sont premiers entre eux, $|H|=1$ et $H=\{0\}$. Ainsi, $dg=0$ et $-dh=0$. Ceci entraîne que $a|d$ et $b|d$. Par conséquent, d est multiple du PPCM de a et b, qui n'est autre que ab puisque a et b sont premiers entre eux. □

Lemme 3.8.5. — *Soit G un groupe abélien fini et soit e son exposant. Il existe un élément de G d'ordre exactement e.*

Démonstration. — Notons le groupe G additivement. Comme e est non nul (il divise $|G|$) on peut le décomposer en produit de facteurs premiers ; écrivons $e=\prod p_i^{n_i}$. Puisque e est le PPCM des ordres des éléments de G, il existe pour tout i un élément g_i de G dont l'ordre est de la forme $p_i^{n_i}m_i$ pour un certain $m_i>0$.

Fixons i. Pour tout $n\in\mathbf{Z}$ on a $nm_ig_i=0 \iff p_i^{n_i}m_i|nm_i \iff p_i^{n_i}|n$. Ainsi, g_i est d'ordre exactement $p_i^{n_i}$.

Une application répétée du lemme 3.8.4 assure alors que la somme $\sum_i g_i$ est d'ordre $\prod p_i^{n_i}=e$, ce qui achève la démonstration. □

Corollaire 3.8.6. — *Soit K un corps commutatif et soit G un sous-groupe fini de $K^\times$; le groupe G est cyclique.*

Démonstration. — Soit e l'exposant de G. On a $g^e=1$ pour tout $g\in G$; par conséquent, le polynôme $X^e-1\in k(X]$ a au moins $|G|$ racines distinctes dans K, ce qui implique que $|G|\leqslant e$. Par ailleurs, le lemme 3.8.5 assure qu'il existe un élément g de G d'ordre e. Le sous-groupe $\langle g\rangle$ de G a alors pour cardinal e ; puisque $|G|\leqslant e$ il vient $|G|=e$ et $G=\langle g\rangle$. □

Commentaires 3.8.7. — Il résulte du corollaire précédent que $K^\times$ est cyclique pour tout corps fini K. En particulier, si p est un nombre premier le groupe $\mathbf{F}_p^\times$ est cyclique. Il existe donc un entier n dont la classe $\overline{n}$ modulo p engendre $\mathbf{F}_p^\times$. Mais notez bien que ce résultat *n'est pas effectif* : le corollaire 3.8.6 repose en effet sur le lemme 3.8.5, et si celui-ci affirme que l'exposant d'un groupe abélien fini G est l'ordre d'un certain élément g de G, sa preuve ne fournit pas de méthode *pratique* pour exhiber un tel g.

3.9. Classification des groupes abéliens finis. — Nous nous proposons de terminer cette section par un théorème de classification de tous les groupes abéliens finis à isomorphisme près. Nous allons commencer par quelques lemmes techniques qui peuvent avoir leur intérêt propre, et qui sont essentiellement des énoncés de prolongements de morphismes. Le premier d'entre eux, ci-dessous, est essentiellement formel.

Lemme 3.9.1. — *Soient G et H deux groupes abéliens notés additivement, et soient G_1 et G_2 deux sous-groupes de G tel que $G=G_1+G_2$. Soit φ_1 un morphisme de G_1 dans H et soit φ_2 un morphisme de G_2 dans H. Supposons que φ_1 et φ_2 coïncident sur $G_1\cap G_2$. Il existe alors un unique morphisme $\varphi\colon G\to H$ tel que $\varphi|_{G_1}=\varphi_1$ et $\varphi|_{G_2}=\varphi_2$.*

Démonstration. — L'unicité est claire : si φ est un morphisme comme ci-dessus et si $g \in G$, on écrit $g = g_1 + g_2$ avec $g_1 \in G_1$ et $g_2 \in G_2$, et on a alors nécessairement $\varphi(g) = \varphi(g_1) + \varphi(g_2) = \varphi_1(g_1) + \varphi_2(g_2)$.

Montrons l'existence de φ. Soit $g \in G$. Écrivons $g = g_1 + g_2$ avec $g_1 \in G_1$ et $g_2 \in G_2$. Vérifions tout d'abord que l'élément $\varphi_1(g_1) + \varphi_2(g_2)$ de H ne dépend que de g, et pas de la décomposition choisie. Écrivons donc $g = g_1' + g_2'$ avec $g_1' \in G_1$ et $g_2' \in G_2$. L'égalité $g_1 + g_2 = g_1' + g_2'$ peut se récrire

$$\underbrace{g_1 - g_1'}_{\in G_1} = \underbrace{g_2' - g_2}_{\in G_2}.$$

Puisque φ_1 et φ_2 coïncident sur $G_1 \cap G_2$, il vient $\varphi_1(g_1 - g_1') = \varphi_2(g_2' - g_2)$, soit encore $\varphi_1(g_1) - \varphi_1(g_1') = \varphi_2(g_2') - \varphi_2(g_2)$, et finalement

$$\varphi_1(g_1) + \varphi_2(g_2) = \varphi_1(g_1') + \varphi_2(g_2'),$$

comme annoncé. Il est donc licite de *poser* $\varphi(g) = \varphi(g_1) + \varphi(g_2)$. On a en particulier $\varphi(g) = \varphi(g + 0) = \varphi_1(g)$ si $g \in G_1$, et $\varphi(g) = \varphi(0 + g) = \varphi_2(g)$ si $g \in G_2$.

Soient g et γ dans G. Écrivons $g = g_1 + g_2$ et $\gamma = \gamma_1 + \gamma_2$ où g_1 et γ_1 appartiennent à G_1, et g_2 et γ_2 à G_2. On a alors $g + \gamma = g_1 + \gamma_1 + g_2 + \gamma_2$. Comme $g_1 + \gamma_1 \in G_1$ et $g_2 + \gamma_2 \in G_2$, il résulte de la définition de φ que

$$\begin{aligned}
\varphi(g+\gamma) &= \varphi_1(g_1 + \gamma_1) + \varphi_2(g_2 + \gamma_2) \\
&= \varphi_1(g_1) + \varphi_2(g_2) + \varphi_1(\gamma_1) + \varphi_2(\gamma_2) \\
&= \varphi(g) + \varphi(\gamma)
\end{aligned}$$

et φ est un morphisme de groupes. □

Le lemme de prolongement suivant est plus subtil ; son énoncé et sa preuve font intervenir un peu d'arithmétique, et notamment les résultats vus plus haut sur les sous-groupes des groupes cycliques.

Lemme 3.9.2. — *Soit d un entier $\geqslant 1$, soit n un multiple de d, soit G un sous-groupe de $\mathbf{Z}/d\mathbf{Z}$ et soit ψ un morphisme de G dans $\mathbf{Z}/n\mathbf{Z}$. Le morphisme ψ s'étend en un morphisme de $\mathbf{Z}/d\mathbf{Z}$ vers $\mathbf{Z}/n\mathbf{Z}$.*

Démonstration. — Il existe un diviseur a de d tel que $G = \langle \overline{a} \rangle$; écrivons $d = ab$ et $n = dm$ avec a et m dans $\mathbf{N}$. Comme l'élément $\overline{a}$ de $\mathbf{Z}/d\mathbf{Z}$ est de b-torsion, l'élément $\psi(\overline{a})$ de $\mathbf{Z}/n\mathbf{Z}$ est de b-torsion ; comme $n = abm$, cela signifie que $\psi(\overline{a})$ est égal à $\overline{ram}$ pour un certain entier r (3.7.5). L'élément $\overline{rm}$ de $\mathbf{Z}/n\mathbf{Z}$ est de d-torsion (car $n = dm$) ; il existe donc un (unique) morphisme χ de $\mathbf{Z}/d\mathbf{Z}$ dans $\mathbf{Z}/n\mathbf{Z}$ envoyant $\overline{1}$ sur $\overline{rm}$. On a alors

$$\underbrace{\chi(\overline{a}) = \chi(a\overline{1})}_{\text{les classes sont prises modulo } d} \quad = \quad \underbrace{a\overline{rm} = \overline{arm}}_{\text{les classes sont prises modulo } n}.$$

Ainsi, $\chi(\overline{a}) = \psi(\overline{a})$. Comme $\overline{a}$ engendre G, la restriction de χ à G est égale à ψ. □

Lemme 3.9.3. — *Soit G un groupe abélien fini et soit $n > 0$ un entier tel que $ng = 0$ pour tout $g \in G$. Soit H un sous-groupe de G et soit φ un morphisme de H dans $\mathbf{Z}/n\mathbf{Z}$. Le morphisme φ s'étend alors en un morphisme de G dans $\mathbf{Z}/n\mathbf{Z}$.*

Démonstration. — On procède par récurrence sur l'indice $[G : H]$. S'il vaut 1 on a $H = G$ et il n'y a rien à démontrer. Supposons donc que $[G : H] > 1$, et que le lemme est vrai pour les sous-groupes de G d'indice strictement inférieur à $[G : H]$.

Comme $[G : H] > 1$, il existe un élément g de G qui n'appartient pas à H. Puisque $ng = 0$, l'ordre d de g est un diviseur de n. Le groupe $\langle g\rangle$ étant isomorphe à $\mathbf{Z}/d\mathbf{Z}$, il découle du lemme 3.9.2 ci-dessus que $\varphi|_{H\cap\langle g\rangle}$ se prolonge en un morphisme θ de $\langle g\rangle$ vers $\mathbf{Z}/n\mathbf{Z}$. Par construction, φ et θ coïncident sur $H \cap \langle g\rangle$. Par le lemme 3.9.1, il existe alors un (unique) morphisme Φ de $H + \langle g\rangle$ dans $\mathbf{Z}/n\mathbf{Z}$ dont la restriction à H est égale à φ et dont la restriction à $\langle g\rangle$ est égale à θ. Puisque $H + \langle g\rangle$ contient strictement H (car $g \notin H$), son indice dans G est strictement inférieur à $[G : H]$. L'hypothèse de récurrence assure alors l'existence d'un morphisme de G dans $\mathbf{Z}/n\mathbf{Z}$ qui prolonge Φ, et *a fortiori* φ. □

Nous pouvons maintenant énoncer le théorème de classification des groupes abéliens finis.

Théorème 3.9.4. — *Soit G un groupe abélien fini. Il existe une unique famille finie $(d_1, \ldots, d_n)$ d'entiers > 1 telle que $d_1|d_2|\ldots|d_n$ et telle que*

$$G \simeq \mathbf{Z}/d_1\mathbf{Z} \oplus \mathbf{Z}/d_2\mathbf{Z} \oplus \ldots \oplus \mathbf{Z}/d_n\mathbf{Z}.$$

Démonstration. — Nous allons tout d'abord prouver l'existence des d_i, puis leur unicité.

Existence de $(d_1, \ldots, d_n)$. — On procède par récurrence sur $|G|$. Si $|G| = 1$ le groupe G est trivial et la famille *vide* d'entiers convient. Supposons maintenant que $|G| > 1$ et que l'existence a été établie pour tout groupe abélien de cardinal strictement inférieur à celui de $|G|$; notons que comme il existe un élément non nul dans G, l'exposant e de G est > 1.

Le lemme 3.8.5 assure qu'il existe un élément $g \in G$ dont l'ordre est égal à e. Il existe alors un isomorphisme $\varphi \colon \langle g\rangle \simeq \mathbf{Z}/e\mathbf{Z}$. En vertu du lemme 3.9.3, l'isomorphisme φ se prolonge en un morphisme Φ de G dans $\mathbf{Z}/e\mathbf{Z}$. Soit H le noyau de Φ. Comme φ est surjectif, Φ l'est *a fortiori* et l'on a donc $e|H| = |G|$. Par ailleurs l'injectivité de φ assure que $H \cap \langle g\rangle = \{0\}$. Les sous-groupes H et $\langle g\rangle$ de G sont donc en somme directe, et le sous-groupe $H \oplus \langle g\rangle$ de G est de cardinal $e|H|$; puisque $e|H| = |G|$, on a $G = H \oplus \langle g\rangle \simeq H \oplus \mathbf{Z}/e\mathbf{Z}$.

Comme $e > 1$, le cardinal de H est strictement inférieur à celui de G. D'après l'hypothèse de récurrence il existe alors une famille finie $(d_1, \ldots, d_r)$ d'entiers > 1 tels que $d_1|d_2|\ldots|d_r$ et tels que H soit isomorphe à $\mathbf{Z}/d_1\mathbf{Z}\oplus\mathbf{Z}/d_2\mathbf{Z}\oplus\ldots\oplus\mathbf{Z}/d_r\mathbf{Z}$. Comme e est l'exposant de G, tous les éléments de H sont de e-torsion ; ceci entraîne notamment que d_r divise e si $r > 0$ (considérer l'élément $\overline{1}$ du dernier sommande $\mathbf{Z}/d_r\mathbf{Z}$). Posons $n = r + 1$ et $d_n = e$; la famille $(d_1, \ldots, d_n)$ satisfait alors les conditions de l'énoncé.

Unicité de $(d_1, \dots, d_n)$. — Pour montrer que $(d_1, \dots, d_n)$ est unique, nous allons montrer qu'elle peut être reconstituée à partir des propriétés intrinsèques du groupe G. Pour cela, commençons par une remarque : pour connaître $(d_1, \dots, d_n)$ il suffit de connaître, pour tout nombre premier p et tout entier $m > 0$, le cardinal $\ell(p, m)$ de l'ensemble des indices i tels que $p^m | d_i$. En effet, compte-tenu du fait que si p^m divise d_i il divise aussi d_j pour tout $j > i$, on vérifie aisément [1] que la famille $(d_1, \dots, d_n)$ s'obtient à partir des $\ell(p, m)$ par l'algorithme récursif suivant :

- ◇ si $\ell(p, m) = 0$ pour tout p et tout $m > 0$ la famille $(d_1, \dots, d_n)$ est vide ;
- ◇ sinon, d_n est égal au produit $\prod_{p \in P} p^{n_p}$, où P est l'ensemble des nombres premiers p tels que l'ensemble $E_p := \{m > 0, \ell(p, m) \neq 0\}$ soit non vide, et où n_p est le plus grand élément de E_p ; on remplace alors pour tout $p \in P$ et tout $m \in E_p$ l'entier $\ell(p, m)$ par $\ell(p, m) - 1$ (et on ne touche pas aux autres $\ell(p, m)$, qui sont de toutes façons nuls), puis l'on détermine $(d_1, \dots, d_{n-1})$ en appliquant l'algorithme à la nouvelle liste des $\ell(p, m)$.

Il suffit donc maintenant d'expliquer comment calculer les $\ell(p, m)$ à partir de G. Fixons un nombre premier p et un entier $m > 0$.

Soit i un entier compris entre 1 et n et soit e_i l'exposant de p dans la décompostion de d_i en produit de facteurs premiers. Le PGCD de d_i et p^m est alors égal à $p^{\min(m, e_i)}$. Le sous-groupe $p^m(\mathbf{Z}/d_i\mathbf{Z})$ de $\mathbf{Z}/d_i\mathbf{Z}$ est aussi son sous-groupe engendré par $\overline{p^m}$; c'est donc l'unique sous-groupe de $\mathbf{Z}/d_i\mathbf{Z}$ d'ordre $d_i/p^{\min(m, e_i)}$ (3.7.2). De même, $p^{m-1}(\mathbf{Z}/d_i\mathbf{Z})$ est l'unique sous-groupe de $\mathbf{Z}/d_i\mathbf{Z}$ d'ordre $d_i/p^{\min(m-1, e_i)}$, et il vient

$$|p^{m-1}(\mathbf{Z}/d_i\mathbf{Z})/p^m(\mathbf{Z}/d_i\mathbf{Z})| = \frac{p^{\min(m, e_i)}}{p^{\min(m-1, e_i)}}.$$

Le terme de droite vaut p si $m \leqslant e_i$, c'est-à-dire si p^m divise d_i ; et il vaut 1 si $m > e_i$.

En appliquant ce qui précède sommande par sommande, on en déduit que le quotient $p^{m-1}G/p^mG$ est de cardinal $p^{\ell(p,m)}$. L'entier $\ell(p, m)$ peut donc bien se décrire en termes des propriétés intrinsèques du groupe G. □

1. Si vous n'êtes pas convaincus et avez du mal à voir ce qui se passe, prenez un exemple concret, par exemple la famille $(2, 2, 2, 6, 12, 24, 120, 240, 240)$, calculez les $\ell(p, m)$ et appliquez l'algorithme.

4. Groupes opérant sur un ensemble et applications

4.1. Définitions et premières propriétés. — Historiquement, les groupes sont d'abord apparus comme «groupes de transformations» c'est-à-dire, en termes contemporains, comme sous-groupes de certains groupes de bijection. On a ensuite progressivement compris l'intérêt d'axiomatiser la notion, ce qui a débouché sur la notion de «groupe abstrait», celle que nous connaissons aujourd'hui.

Toutefois, l'expérience montre que pour comprendre un groupe abstrait, il peut être utile de le voir, éventuellement de plusieurs façons différentes, comme un groupe de transformations. Pour donner un sens rigoureux à ce slogan, on doit introduire la notion d'*opération* d'un groupe sur un ensemble.

Définition 4.1.1. — Soit G un groupe et soit X un ensemble. Une *opération à gauche* de G sur X est une application $(g,x) \mapsto g * x$ de $G \times X$ vers X telle que les propriétés suivantes soient satisfaites pour tout $(gh,x) \in G^2 \times X$:

(i) $e * x = x$;

(ii) $(gh) * x = g \cdot (h * x)$.

Commentaires 4.1.2. — Il y a aussi une notion d'opération *à droite* de G sur X : c'est une application $(x,g) \mapsto x * g$ de $X \times G$ dans X telle que $x * e = x$ et $x * (gh) = (x * g) * h$ pour tout $x \in X$ et tout $(g,h) \in G^2$.

Convention 4.1.3. — Sauf mention expresse du contraire, lorsque nous parlerons d'opération d'un groupe sur un ensemble, il s'agira toujours d'opération *à gauche.* Cette restriction simplifie la rédaction, et est bénigne. Les résultats que nous établirons s'étendent en effet *mutatis mutandis* aux opérations à droite, soit en en transcrivant les preuves, soit en remarquant qu'une opération à droite de G peut aussi s'interpréter comme une opération à gauche du groupe «opposé» G^{op}, qui est le groupe dont l'ensemble sous-jacent est G et dont la loi est $(g,h) \mapsto hg$.

Si G est un groupe et si X est un ensemble, on emploiera souvent l'expression «G opère sur X» pour signifier qu'on s'est donné une opération (à gauche, donc) de G sur X.

Lorsqu'on écrira «Soit G un groupe opérant sur un ensemble X» sans mentionner explicitement l'opération en jeu, celle-ci n'aura le plus souvent droit à aucun symbole spécifique et sera simplement notée $(g,x) \mapsto gx$.

4.1.4. *Un autre point de vue sur les opérations.* — Soit G un groupe et soit X un ensemble. Supposons donnée une opération de G sur X ; pour tout $g \in G$, notons $\iota(g)$ l'application $x \mapsto gx$ de X dans lui-même. La condition (i) de la définition 4.1.1 signifie que $\iota(e) = \mathrm{Id}_X$, et la condition (ii) que $\iota(gh) = \iota(g) \circ \iota(h)$ pour tout $(g,h) \in G^2$.

Soit $g \in G$. Par ce qui précède,

$$\iota(g) \circ \iota(g^{-1}) = \iota(g^{-1}) \circ \iota(g) = \iota(e) = \mathrm{Id}_X.$$

Autrement dit, $\iota(g)$ est une bijection de bijection réciproque $\iota(g^{-1})$. La formule $\iota(gh) = \iota(g) \circ \iota(h)$ peut alors se traduire en disant que ι est un morphisme de groupes de G vers le groupe $\mathfrak{S}_X$ des bijections de X dans lui-même.

Réciproquement, supposons donnée un morphisme de groupes $\upsilon\colon G \to \mathfrak{S}_X$. Nous laissons le lecteur vérifier que l'application $(g,x) \mapsto \upsilon(g)(x)$ est une opération de G sur X, et que nos deux constructions sont réciproques l'une de l'autre.

Il revient donc au même de se donner une opération de G sur X ou un morphisme de groupes de G dans $\mathfrak{S}_X$.

Remarque 4.1.5. — On peut faire le même jeu que ci-dessus avec les opérations à droite : à toute opération à droite d'un groupe G sur un ensemble X on peut associer l'application $\iota\colon G \to \mathfrak{S}_X, g \mapsto (x \mapsto xg)$. *Mais ce n'est plus un morphisme de groupes en général* : on a en effet $\iota(gh) = \iota(h) \circ \iota(g)$ pour tout (g,h). Si l'on veut obtenir un morphisme de groupes, on doit ou bien considérer ι comme une application de G^{op} dans $\mathfrak{S}_X$, ou bien remplacer ι par $x \mapsto (x \mapsto xg^{-1})$.

Les opérations à droite semblent donc en un sens un peu moins naturelles que les opérations à gauche. Vous vous demandez peut-être quelle est l'origine de cette dissymétrie *a priori* surprenante ; elle vient de la définition de la composition des applications (qui est la loi de groupe de $\mathfrak{S}_X$), qui repose elle-même sur un choix arbitraire, celui du sens dans lequel on compose ; c'est ce choix qui explique que la droite et la gauche ne jouent pas tout à fait le même rôle ici.

4.1.6. — On peut maintenant donner un sens rigoureux à ce qui a été annoncé en introduction. Lorsqu'un groupe G opère sur un ensemble X, chaque élément de g peut être vu comme à une bijection de X dans lui-même – la bijection $\iota(g)$ du 4.1.4 ci-dessus ; le produit de G correspond alors précisément à la composition des bijections. Mais il faut tout de même faire un peu attention : rien n'oblige le morphisme ι à être injectif, et deux éléments différents de G peuvent donc parfois s'interpréter comme la même bijection de X dans X.

4.1.7. — Soit X un ensemble et soit Γ un sous-groupe de $\mathfrak{S}_X$. Soit G un groupe opérant sur X. On dira que G opère *par automorphismes appartenant à* Γ si le morphisme de G vers $\mathfrak{S}_X$ qui définit l'opération considérée est à valeurs dans Γ ; cela revient à demander que $x \mapsto gx$ appartienne à Γ pour tout $x \in X$. Ainsi, si X est un groupe (resp. un espace vectoriel sur un corps k, resp. un anneau, resp. un espace topologique), on a une notion naturelle de groupes opérant sur X par automorphismes de groupes (resp. par automorphismes k-linéaires, resp. par automorphismes d'anneaux, resp. par homéomorphismes).

***Exemple 4.1.8* (opération tautologique).** — Si X est un ensemble, le groupe $\mathfrak{S}_X$ opère tautologiquement sur X par la formule $(\sigma,x) \mapsto \sigma(x)$; le morphisme correspondant de $\mathfrak{S}_X$ dans lui-même est l'identité.

***Exemple 4.1.9* (opération par translations).** — Soit G un groupe. La formule $(g,x) \mapsto gx$ définit une opération de G sur lui-même, dite *par translations (à gauche).* Attention : ici g est vu comme un élément du *groupe* G (celui qui opère) et x comme un élément de l'*ensemble* G.

Plus généralement, soit H un sous-groupe de G. Nous laissons le lecteur vérifier que la formule $(g,xH) \mapsto g(xH) = (gx)H$ définit une opération de $G \times G/H$ vers

G/H. On l'appelle encore l'opération *par translations (à gauche)* (lorsque $H = \{e\}$ on retrouve l'opération par translations de G sur lui-même).

***Exemple 4.1.10* (opération par automorphismes intérieurs)**
Soit G un groupe. L'application $g \mapsto (h \mapsto ghg^{-1})$ définit un morphisme de G dans $\mathrm{Aut}(G)$ (2.7 *et sq.*), qui est lui-même un sous-groupe de $\mathfrak{S}_G$. Elle définit donc une opération de G sur lui-même, dite *par automorphismes intérieurs* ou par *conjugaison.*

4.1.11. *Quelques exemples d'opérations induites.* — Soit X un ensemble sur lequel un groupe G opère et soit $\iota \colon G \to \mathfrak{S}_X$ le morphisme correspondant.

4.1.11.1. — Soit Y un sous-ensemble de X stable sous l'action de G . L'opération de G sur X induit par restriction à Y une opération de G sur Y.

4.1.11.2. — Soit $\varphi \colon H \to G$ un morphisme de groupes. La composition $\iota \circ \varphi$ est un morphisme de H vers $\mathfrak{S}_X$; elle définit donc une opération de H sur X, telle que $hx = \varphi(h)x$ pour tout $h \in H$; nous dirons que c'est l'opération de H sur X induite par celle de G (et par φ, s'il est nécessaire de le préciser). Notez le cas important où H est un sous-groupe de G et où ι est l'inclusion : on a alors simplement restreint l'opération initiale à H.

4.1.11.3. — Pour tout $P \in \mathscr{P}(X)$, on pose $gP = \{gx\}_{x \in P}$. On vérifie immédiatement que $(g, P) \mapsto gP$ définit une opération de G sur $\mathscr{P}(X)$; on dira qu'elle est *induite* par l'opération de G sur X.

Définition 4.1.12. — Soit G un groupe. Un G*-ensemble* est un ensemble X muni d'une opération de G. Une application $\varphi \colon Y \to X$ entre deux G-ensembles est dite *équivariante* (ou G*-équivariante* si la précision s'impose) si l'on a $\varphi(gy) = g\varphi(y)$ pour tout $(g, y) \in G \times Y$.

4.1.13. — On vérifie immédiatement les faits suivants : l'identité d'un G-ensemble est équivariante ; la composée de deux applications équivariantes est équivariante ; si une bijection entre deux G-ensembles est équivariante, sa réciproque est également équivariante.

4.2. Orbites d'une opération de groupes. — Soit X un ensemble. Nous allons expliquer dans ce qui suit comment associer à une opération donnée d'un groupe sur X une relation d'équivalence sur X qui joue un rôle fondamental dans toute la suite de la théorie.

4.2.1. — Soit G un groupe opérant sur un ensemble X et soit $\mathscr{R}$ la relation sur X définie par la condition

$$x\mathscr{R}y \iff \exists g \in G \text{ t.q. } gx = y.$$

Montrons que $\mathscr{R}$ est une relation d'équivalence, que l'on dira *induite par l'action de G sur X.*

Soit $x \in X$. On a $x = ex$; ainsi, $x\mathscr{R}x$ et $\mathscr{R}$ est réflexive.

Soient x et y deux éléments de X. Supposons que $x\mathscr{R}y$. Il existe alors $g \in G$ tel que $y = gx$. Il vient $g^{-1}y = (g^{-1}g)x = x$; ainsi, $y\mathscr{R}x$ et $\mathscr{R}$ est symétrique.

Enfin, soient x, y et z trois éléments de X tels que $x\mathscr{R}y$ et $y\mathscr{R}z$. Il existe alors deux éléments g et h de G tels que l'on ait $gx = y$ et $hy = z$. Il vient

$$(hg)x = h \cdot (gx) = hy = z.$$

Ainsi, $x\mathscr{R}z$ et $\mathscr{R}$ est transitive.

Définition 4.2.2. — Soit G un groupe opérant sur un ensemble X. Les classes d'équivalences de la relation d'équivalence sur X induite par l'action de G sont appelées *orbites (de l'action de G sur X, ou de X sous G).*

Remarque 4.2.3. — Une orbite est *par définition* non vide, et c'est l'orbite de chacun de ses éléments.

4.2.4. — Soit G un groupe opérant sur un ensemble X. Soit $x \in X$. L'orbite de x est égale par définition à $\{gx\}_{g \in G}$ et sera notée Gx. C'est visiblement le plus petit sous-ensemble de X stable sous l'action de G contenant x.

L'ensemble des orbites de X sous G sera noté $G\backslash X$.

Exemple 4.2.5. — Soit G un groupe et soit H un sous-groupe de G. L'action de G sur lui-même par translations à gauche définit par restriction une action de H sur G. Il résulte immédiatement des définitions que la relation d'équivalence induite par cette action est la relation de *congruence à droite modulo H*. L'ensemble des orbites s'identifie donc à $H\backslash G$.

Exemple 4.2.6. — Soit G un groupe opérant sur lui-même par automorphismes intérieurs. Deux éléments de G appartiennent alors à la même orbite si et seulement si ils sont conjugués ; les orbites sont appelées *classes de conjugaison*.

4.3. Stabilisateur d'un élément. — Soit X un ensemble sur lequel opère un groupe G. Nous allons expliquer dans ce qui suit comment associer à un élément x de X un sous-groupe de G puis expliquer les liens entre ce sous-groupe et l'orbite de x.

Définition 4.3.1. — Soit G un groupe opérant sur un ensemble X . Soit $x \in X$. Le *stabilisateur* de x (dans G) est l'ensemble des éléments g de G tels que $gx = x$. On le note $\mathrm{Stab}(x)$. On vérifie immédiatement qu'il s'agit d'un sous-groupe de G.

4.3.2. *Variation du stabilisateur dans une orbite fixée.* — Soit G un groupe opérant sur un ensemble X. Soit $x \in X$ et soit $y \in Gx$. Soit $g \in G$ tel que $gx = y$. Pour tout $h \in G$, on a les équivalences

$$\begin{aligned} h \in \mathrm{Stab}(x) &\iff hx = x \\ &\iff h(g^{-1}y) = g^{-1}y \\ &\iff (ghg^{-1})y = y \\ &\iff ghg^{-1} \in \mathrm{Stab}(y) \end{aligned}$$

Autrement dit $\mathrm{Stab}(y) = g\mathrm{Stab}(x)g^{-1}$. Les stabilisateurs de x et y sont donc conjugués ; mais notez que ce qu'on a obtenu est un peu plus précis : $\mathrm{Stab}(y)$ est égal au conjugué de $\mathrm{Stab}(x)$ par n'importe quel élément de G qui envoie x sur y.

Exemple 4.3.3. — Soit G un groupe ; faisons-le opérer sur lui-même par automorphismes intérieurs.

Stabilisateur d'un élément de G. Soit $h \in G$. Par définition, le stabilisateur de h est l'ensemble des éléments g de G tels que $ghg^{-1} = h$, c'est-à-dire tels que $gh = hg$; c'est donc l'ensemble des éléments de G qui commutent avec h.

Stabilisateur d'un sous-groupe de G. L'opération considérée en induit une sur $\mathscr{P}(G)$, puis par restriction sur l'ensemble des sous-groupes de G. Soit H un tel sous-groupe. Son stabilisateur sous cette action est par définition $\{g \in G, gHg^{-1} = H\}$. On l'appelle le *normalisateur* de H, et on le note $\mathrm{N}(H)$ (ou $\mathrm{N}_G(H)$ s'il y a besoin de préciser que le groupe ambiant est G). Il est immédiat que $H \subset \mathrm{N}(H)$ et que $\mathrm{N}(H)$ est le plus grand sous-groupe de G dans lequel H est distingué.

4.3.4. *Stabilisateur d'une partie.* — Soit G un groupe opérant sur un ensemble X ; on considère $\mathscr{P}(X)$ comme munie de l'opération induite. Soit P une partie de X. Soit M_P l'ensemble des éléments g de G tels que $gP \subset P$, c'est-à-dire tels que $gx \in P$ dès que $x \in P$.

Il est clair que $\mathrm{Stab}(P) \subset M_P$, mais l'inclusion est stricte en général. Par exemple, considérons l'opération de $\mathbf{Z}$ sur lui-même par translations. On a alors $1 + \mathbf{N} \subset \mathbf{N}$, mais $1 + \mathbf{N} \neq \mathbf{N}$ car $0 \notin 1 + \mathbf{N}$; ainsi, 1 n'appartient pas au stabilisateur de $\mathbf{N}$. Nous allons toutefois donner différentes conditions permettant de s'assurer que certains éléments de M_P (voire tous) appartiennent à G_P.

4.3.4.1. — Si Γ est un sous-groupe de G contenu dans M_P, il est contenu dans $\mathrm{Stab}(P)$. En effet, soit $g \in \Gamma$. On a $gP \subset P$, mais aussi $g^{-1}P \subset P$ car $g^{-1} \in \Gamma \subset M_P$ puisque Γ est un sous-groupe de G. En appliquant g aux deux termes de l'inclusion $g^{-1}P \subset P$ on voit que $P \subset gP$; par conséquent, $gP = P$ et $g \in G_P$.

4.3.4.2. — Donnons deux cas importants en pratique dans lesquels M_P est automatiquement égal à $\mathrm{Stab}(P)$.

Le cas où P *est fini.* En effet, plaçons-nous sous cette hypothèse et donnons-nous un élément g de M_P. La bijection $x \mapsto gx$ de X sur lui-même stabilise P et induit donc une injection de P dans lui-même. Comme P est fini cette injection est surjective et $gP = P$; par conséquent, $g \in \mathrm{Stab}(P)$.

Le cas où X *est un espace vectoriel sur un corps* k*, où* G *opère par automorphismes* k*-linéaires, et où* P *est un sous-espace vectoriel de dimension finie de* X. En effet, plaçons-nous sous cette hypothèse et donnons-nous un élément g de M_P. La bijection linéaire $x \mapsto gx$ de X sur lui-même stabilise P et induit donc une injection linéaire de P dans lui-même. Comme P est de dimension finie cette injection est surjective et $gP = P$; par conséquent, $g \in \mathrm{Stab}(P)$.

4.3.5. *Noyau d'une opération et stabilisateurs.* — Soit X un ensemble et soit G un groupe opérant sur X ; soit $\varphi \colon G \to \mathfrak{S}_X$ le morphisme correspondant. Il résulte des définitions que $\mathrm{Ker}\varphi = \bigcap_{x \in X} \mathrm{Stab}(x)$.

Définitions 4.3.6. — Soit G un groupe opérant sur un ensemble X.

On dit que G opère *fidèlement* sur X, ou bien que l'opération de G sur X est *fidèle*, si $\mathrm{Ker}\varphi$ est trivial, c'est-à-dire encore si φ est injectif. Si c'est le cas, on peut identifier G *via* φ à un groupe de bijections de X dans lui-même : les ambiguïtés éventuelles signalées en 4.1.6 sont alors inexistantes.

On dit que G opère *transitivement* sur X, ou bien que l'opération de G sur X est *transitive*, si X est non vide et si pour tout couple (x, y) d'éléments de X il existe $g \in G$ tel que $gx = y$. Cela revient à demander qu'il y ait exactement une orbite de X sous G, ou encore que X lui-même soit une telle orbite (notez que si X est vide il n'y a *aucune* orbite).

Exemple 4.3.7. — Soit G un groupe opérant sur un ensemble X et soit O une orbite de X sous G. Comme O est stable sous G, elle hérite par restriction d'une opération de G ; celle-ci est alors transitive par définition.

Soit $x \in O$. Il résulte de 4.3.2 et de 4.3.5 que le noyau de $G \to \mathfrak{S}_O$ est égal à $\bigcap_{g\in G} g\mathrm{Stab}(x)g^{-1}$. Nous vous laissons vérifier qu'il s'agit du plus grand sous-groupe distingué de G contenu dans $\mathrm{Stab}(x)$.

***Exemple 4.3.8* (le cas de l'action d'un groupe par translation sur lui-même)**

Soit G un groupe ; on le fait opérer sur lui-même par translations à gauche. Cette opération est transitive : en effet pour tout $g \in G$ on a $g = ge$, et G est donc égal à l'orbite de e. Si h est un élément de G son stabilisateur est alors $\{g \in G, gh = h\}$, qui est réduit à $\{e\}$. L'opération considérée est *a fortiori* fidèle (4.3.5). Elle induit en particulier un morphisme injectif $G \hookrightarrow \mathfrak{S}_G$; ainsi, G s'identifie à un sous-groupe de $\mathfrak{S}_G$.

***Exemple 4.3.9* (le cas de l'action d'un groupe par translation sur un quotient)**

L'exemple 4.3.8 ci-dessus se généralise comme suit. Soit G un groupe et soit H un sous-groupe de G. On fait opérer G sur G/H par translations à gauche (exemple 4.1.9). Cette opération est transitive : en effet pour tout g appartenant à G on a $gH = (ge)H = g(eH)$ et G/H est donc égal à l'orbite de eH.

Le stabilisateur de eH est égal à l'ensemble des $g \in G$ tels que $geH = eH$, c'est-à-dire à l'ensemble des $g \in G$ tels que $gH = eH$, qui n'est autre que $eH = H$. On déduit alors de 4.3.2 et 4.3.5 que l'ensemble des stabilisateurs des éléments de G/H est l'ensemble des conjugués de H, et que le noyau de $G \to \mathfrak{S}_{G/H}$ est égal à $\bigcap_{g\in G} gHg^{-1}$, qui est le plus grand sous-groupe distingué de G contenu dans H (c'est un cas particulier de 4.3.7).

***Exemple 4.3.10* (le cas de l'action d'un groupe sur lui-même par automorphismes intérieurs)**

Soit G un groupe. L'action de G sur lui-même par automorphismes intérieurs est induite par un morphisme $G \to \mathrm{Aut}\, G \subset \mathfrak{S}_G$ dont le noyau est $\mathrm{Z}(G)$ (2.10.8) ; elle est donc fidèle si et seulement si $\mathrm{Z}(G)$ est trivial.

Proposition 4.3.11. — *Soit G un groupe et soit X un ensemble sur lequel G opère. Soit $x \in X$. On fait opérer G sur le quotient $G/\mathrm{Stab}(x)$ par translations à gauche (4.1.9) et sur l'orbite Gx par restriction.*

(i) *L'application $\pi\colon G \to Gx$, $g \mapsto gx$ induit une bijection équivariante $\overline{\pi}\colon (G/\mathrm{Stab}(x)) \simeq Gx$*

(ii) *Si de plus G est fini, l'orbite Gx est finie et l'on a $|Gx| = |G|/|\mathrm{Stab}(x)|$; en particulier, $|Gx|$ divise $|G|$.*

Démonstration. — Soient g et h deux éléments de G. On a

$$\begin{aligned} \pi(g) = \pi(h) &\iff gx = hx \\ &\iff (h^{-1}g)x = x \\ &\iff h^{-1}g \in \mathrm{Stab}(x) \\ &\iff g\mathrm{Stab}(x) = h\mathrm{Stab}(x). \end{aligned}$$

Il en résulte que π est invariante par la relation de congruence à gauche modulo $\mathrm{Stab}(x)$, et que l'application induite $\overline{\pi}\colon G/\mathrm{Stab}(x) \to Gx$ est injective. Par ailleurs, π est surjective par définition de Gx ; par conséquent, $\overline{\pi}$ est surjective, et finalement bijective.

Soient g et h deux éléments de G. On a

$$\overline{\pi}(g(h\mathrm{Stab}(x))) = \overline{\pi}((gh)\mathrm{Stab}(x)) = (gh)x = g(hx) = g(\pi(h)) = g(\overline{\pi}(h\mathrm{Stab}(x))),$$

ce qui montre que $\overline{\pi}$ est équivariante, d'où (i).

Si G est fini alors $|\mathrm{Stab}(x)|$ divise $|G|$ et $G/\mathrm{Stab}(x)$ est fini de cardinal $|G|/|\mathrm{Stab}(x)|$, d'où (ii). □

Exemple 4.3.12. — Soit k un corps et soit n un entier. On définit *l'espace projectif de dimension n sur k*, noté $\mathbf{P}^n(k)$, comme l'ensemble des droites vectorielles de k^{n+1} (nous avons donné plus haut une définition de $\mathbf{P}^1(k)$ qui diffère de celle proposée ici, voir l'exemple 2.15.2 ; mais nous verrons au 4.3.13 ci-dessous que ces deux définitions sont en fait compatibles).

Soit $(e_1, \ldots, e_{n+1})$ la base canonique de k^{n+1} ; le choix de celle-ci permet d'identifier $\mathrm{GL}_{n+1}(k)$ au groupe des bijections linéaires de k^{n+1} sur lui-même, d'où une opération naturelle de $\mathrm{GL}_{n+1}(k)$ sur k^{n+1}. Cette opération en induit une sur l'ensemble des parties de k^{n+1} puis, par restriction, sur $\mathbf{P}^n(k)$.

L'opération de $\mathrm{GL}_{n+1}(k)$ *sur* $\mathbf{P}^n(k)$ *est transitive.* En effet, soit D une droite vectorielle de k^{n+1}. Choisissons un vecteur directeur v de D, et complétons-le en une base $(v = v_1, v_2, \ldots, v_{n+1})$. Soit u l'élément de $\mathrm{GL}_{n+1}(k)$ qui envoie e_i sur v_i pour tout i. On a alors $u(ke_1) = kv_1 = D$. Il s'ensuit que l'orbite de ke_1 est égale à $\mathbf{P}^n(k)$ tout entier, d'où notre assertion.

Le stabilisateur de la droite ke_1 est le sous-ensemble de $\mathrm{GL}_{n+1}(k)$ formé des éléments u tels que $u(e_1)$ soit de la forme λe_1 avec $\lambda \in k^\times$. C'est donc le sous-groupe H de $\mathrm{GL}_{n+1}(k)$ constitué des matrices dont tous les termes de la première colonne sauf celui en haut à gauche sont nuls, c'est-à-dire encore des matrices (a_{ij}) telles que $a_{i1} = 0$ pour tout $i > 1$.

Il résulte dès lors de la proposition 4.3.11 que l'application $u \mapsto u(ke_1)$ induit une bijection $\mathrm{GL}_{n+1}(k)$-équivariante

$$\mathrm{GL}_{n+1}(k)/H \simeq \mathbf{P}^n(k).$$

4.3.13. *Compatibilité entre les deux définitions de* $\mathbf{P}^1(k)$. — Nous avons *a priori* deux définitions de $\mathbf{P}^1(k)$: celle donnée à l'exemple 2.15.2, à savoir $k \cup \{\infty\}$; et celle donnée à l'exemple 4.3.12, à savoir l'ensemble des droites vectorielles de k^2.

Mais il n'y a pas à s'inquiéter, car ces deux définitions sont compatibles : on dispose en effet d'une bijection naturelle entre l'ensemble des droites vectorielles de k^2 et $k \cup \{\infty\}$, consistant à envoyer une droite $D = k(a, b)$ sur sa pente, définie comme étant égale à b/a si $a \neq 0$ et à ∞ sinon (il est clair que la pente ne dépend que de D et pas du vecteur directeur (a, b) choisi).

4.4. Opérations, orbites et arithmétique élémentaire. — Si G est un groupe fini opérant sur un ensemble X et si x est un élément de X, nous avons vu plus haut que le cardinal de l'orbite Gx est égal à $|G|/|\mathrm{Stab}(x)|$ (prop. 4.3.11). Nous nous proposons de conclure cette section en donnant quelques applications de cette formule.

Notations 4.4.1. — Soit G un groupe opérant sur un ensemble X. Pour tout sous-groupe H de G, on notera X^H l'ensemble des éléments x de X qui sont fixes sous H, c'est-à-dire tels que $hx = x$ pour tout $h \in H$, c'est-à-dire encore tels que $H \subset \mathrm{Stab}(x)$.

Si $g \in G$, on notera $\mathrm{Fix}(g)$ l'ensemble des points fixes de g, c'est-à-dire l'ensemble des $x \in X$ tels que $gx = x$, ou encore tels que $g \in \mathrm{Stab}(x)$.

Remarque 4.4.2. — Soit G un groupe opérant sur un ensemble X. Si $x \in X$, l'orbite Gx est réduite au singleton $\{x\}$ si et seulement si $x \in X^G$ (ce qui revient à demander que $\mathrm{Stab}(x) = G$).

4.4.3. — Soit G un groupe opérant sur un ensemble *fini* X.

4.4.3.1. — Comme X est réunion disjointe des orbites, on a $|X| = \sum_{O \in G\backslash X} |O|$. En séparant les orbites singleton des autres, on obtient au vu de la remarque 4.4.2 l'égalité

$$|X| = |X^G| + \sum_{O \in G\backslash X,\ |O|>1} |O|.$$

4.4.3.2. — On suppose de plus que G est fini de cardinal p^n pour un certain p premier et un certain $n \in \mathbf{N}$ (un tel G est appelé un *p-groupe*). Pour toute orbite O le cardinal $|O|$, qui divise $|G|$, est alors une puissance de p ; en particulier, ou bien $|O| = 1$, ou bien $|O|$ est multiple de p. On déduit alors de 4.4.3.1 que $|X| = |X^G|$ modulo p.

***Lemme 4.4.4* (formule de Burnside).** — *Soit G un groupe fini opérant sur un ensemble fini X. On a*

$$\frac{1}{|G|} \sum_{g \in G} |\mathrm{Fix}(g)| = |G\backslash X|.$$

Démonstration. — Soit E l'ensemble des couples $(g,x) \in G \times X$ tels que $gx = x$. On a trivialement $|E| = \sum_{g \in G} |\mathrm{Fix}(g)|$. Mais on a aussi

$$\begin{aligned}
|E| &= \sum_{x \in X} |\mathrm{Stab}(x)| \\
&= \sum_{O \in G \backslash X} \sum_{x \in O} |\mathrm{Stab}(x)| \\
&= \sum_{O \in G \backslash X} \sum_{x \in O} \frac{|G|}{|O|} \\
&= |G| \sum_{O \in G \backslash X} \sum_{x \in O} \frac{1}{|O|} \\
&= |G| \sum_{O \in G \backslash X} \frac{1}{|O|} \cdot |O| \\
&= |G| \sum_{O \in G \backslash X} 1 = |G| \cdot |G \backslash X|,
\end{aligned}$$

d'où la formule souhaitée. □

Remarque 4.4.5. — On peut retenir la formule de Burnside sous forme de l'énoncé suivant : *lorsqu'un groupe fini opère sur un ensemble fini, le nombre moyen de points fixes d'un élément du groupe est égal au nombre d'orbites.*

Commentaires 4.4.6. — Nombre de résultats en théorie des groupes finis se démontrent de manière particulièrement astucieuse et efficace en introduisant une opération de groupe judicieusement choisie puis en appliquant la formule fondamentale $|Gx| = |G|/|\mathrm{Stab}(x)|$ ou bien directement, ou bien à travers ses conséquences directes comme la congruence établie au 4.4.3.2 ou la formule de Burnside (lemme 4.4.4). Nous allons donner deux illustrations de cette méthode et nous en reverrons d'autres à la section 7.

Lemme 4.4.7. — *Soit p un nombre premier et soit G un p-groupe non trivial. Le centre de G est non trivial.*

Démonstration. — Faisons opérer G sur lui-même par automorphismes intérieurs. Le nombre de points fixes sous cette action est égal à $|G|$ modulo p d'après 4.4.3.2 et est donc nul modulo p puisque G est un p-groupe non trivial.

Or un élément g de G est fixe sous l'action considérée si et seulement si $hgh^{-1} = g$ pour tout $h \in G$, soit encore si et seulement si $h \in \mathrm{Z}(G)$. Ainsi le cardinal $\mathrm{Z}(G)$ est nul modulo p ; comme il est au moins égal à 1 puisque $\mathrm{Z}(G)$ est un groupe, il est au moins égal à p et est en particulier > 1. □

Corollaire 4.4.8. — *Soit p un nombre premier et soit G un groupe de cardinal p^2. Le groupe G est alors abélien, et est plus précisément isomorphe ou bien à $(\mathbf{Z}/p\mathbf{Z})^2$ ou bien à $\mathbf{Z}/p^2\mathbf{Z}$, ces deux cas étant exclusifs l'un de l'autre.*

Démonstration. — Montrons tout d'abord que G est abélien, c'est-à-dire que G est égal à son centre. On raisonne par l'absurde. Supposons que $\mathrm{Z}(G)$ soit un sous-groupe strict de G ; il est non trivial d'après le lemme 4.4.7 ci-dessus, et est donc de cardinal p. Soit g un élément de G qui n'appartient pas à $\mathrm{Z}(G)$. Le stabilisateur S de g sous l'action de G sur lui-même par automorphismes intérieurs est exactement l'ensemble des éléments de G qui commutent avec g ; en conséquence, S est un sous-groupe de G qui contient $\mathrm{Z}(G)$ et g. Il vient $|S| \geqslant p+1$, puis $|S| = p^2$ puisque $|S|$ divise p^2. Ainsi $S = G$; l'élément g commute donc avec tous les éléments de G, ce qui veut dire qu'il appartient à $\mathrm{Z}(G)$ et aboutit à une contradiction.

L'assertion sur le type d'isomorphie de G se déduit alors du théorème 3.9.4 : le cardinal de G étant égal à p^2, on voit que la liste d'entiers $d_1, \ldots, d_r$ dont ce théorème affirme l'existence et l'unicité ne peut être que p, p ou la liste singleton p^2, ce qui donne précisément le résultat souhaité.

Donnons également une preuve directe. Si G possède un élément g d'ordre p^2 il est isomorphe à $\mathbf{Z}/p^2\mathbf{Z}$. Sinon, tous les éléments de G sont de p-torsion ; il en résulte que pour tout entier n et tout élément g de G, l'élément ng de G ne dépend que de la classe de n modulo p. La formule $(\overline{n}, g) \mapsto ng$ définit donc sans ambiguïté une application de $\mathbf{Z}/p\mathbf{Z} \times G$ dans G, dont on vérifie aussitôt qu'elle fait de G un $\mathbf{Z}/p\mathbf{Z}$-espace vectoriel. Puisque G est fini, il est de dimension finie m, et est de ce fait isomorphe à $(\mathbf{Z}/p\mathbf{Z})^m$ comme $\mathbf{Z}/p\mathbf{Z}$-espace vectoriel, et *a fortiori* comme groupe. Puisque $|G| = p^2$, on a $m = 2$ et $G \simeq (\mathbf{Z}/p\mathbf{Z})^2$; et il n'est pas isomorphe à $\mathbf{Z}/p^2\mathbf{Z}$ puisqu'il n'a pas d'élément d'ordre p^2. □

Lemme 4.4.9. — *Soit k un corps de caractéristique $p > 0$, soit V un k-espace vectoriel non nul et soit G un p-groupe agissant linéairement sur V. L'espace vectoriel V^G est alors non nul.*

Démonstration. — Comme V est non nul, il existe un vecteur $v \neq 0$ dans V. Soit V_0 le sous $\mathbf{F}_p$-espace vectoriel de V engendré par $\{gv\}_{g \in G}$ (rappelons que $\mathbf{F}_p \subset k$ puisque k est de caractéristique p). Le $\mathbf{F}_p$-espace vectoriel V_0 est de type fini, et donc de dimension finie $m > 0$ (il contient v qui est non nul) ; son cardinal est alors égal à p^m. En vertu de 4.4.3.2 le cardinal de V_0^G est égal à p^m modulo p, et est donc multiple de p (c'est ici qu'on utilise l'hypothèse que G est un p-groupe). Or V_0^G contient au moins un élément, à savoir 0 ; il en contient dès lors au moins p, et possède en particulier un vecteur w non nul ; par conséquent, $V^G \neq \{0\}$. □

Corollaire 4.4.10. — *Soit k un corps de caractéristique $p > 0$, soit V un k-espace vectoriel de dimension finie et soit G un p-groupe agissant linéairement sur V. Il existe une base $\mathscr{B}$ de V telle que pour tout $g \in G$, la matrice de $v \mapsto gv$ dans $\mathscr{B}$ soit triangulaire supérieure avec des 1 sur la diagonale.*

Démonstration. — Posons $n = \dim V$. On procède par récurrence sur n. Si $n = 0$ la seule base de V est la base vide, qui possède la propriété requise. Supposons $n > 0$ et la propriété vraie en dimension $< \dim V$.

D'après le lemme 4.4.9, il existe un vecteur e non nul de V tel que $ge = e$ pour tout $g \in G$. Soit W le groupe quotient V/ke. Comme ke est un sous-espace vectoriel de V, il est immédiat que pour tout $(\lambda, v) \in k \times V$, la classe de λv modulo ke ne dépend que de celle de V. La loi externe de V induit donc par passage au quotient une loi externe $k \times W \to W$, qui fait de W un k-espace vectoriel ; l'application quotient $V \to W$ est linéaire, et la formule du rang assure que $\dim W = n-1$. Par ailleurs comme $e \in V^G$ le sous-espace ke de V est stable sous l'action de G ; il en résulte que celle-ci induit par passage au quotient une action linéaire de G sur W.

Par hypothèse de récurrence, il existe une base $(\varepsilon_2, \ldots, \varepsilon_n)$ de W telle que pour tout $g \in G$, la matrice de $w \mapsto gw$ dans $(\varepsilon_2, \ldots, \varepsilon_n)$ soit triangulaire supérieure avec des 1 sur la diagonale. Cela veut précisément dire que pour tout j compris entre 2 et n il existe une famille $(a_{ij})_{j<i\leqslant n}$ d'éléments de k telle que $g\varepsilon_j = \varepsilon_j + \sum_{i>j} a_{ij}\varepsilon_i$.

Posons $e_1 = e$ et, pour tout j compris entre 2 et n, choisissons un antécédent de e_j de ε_j dans V. La famille $(e_1, \ldots, e_n)$ est alors une base de V. En effet, vérifions qu'elle est libre et génératrice (il suffit en fait de vérifier l'une des deux propriétés, puisqu'elle est de cardinal n).

Cette famille est libre. Soit (λ_j) une famille de scalaires telle que $\sum \lambda_j e_j = 0$. En réduisant modulo ke il vient $\sum_{j\geqslant 2} \lambda_j \varepsilon_j = 0$. Comme la famille (ε_j) est libre, on en déduit que $\lambda_j = 0$ pour tout $j \geqslant 2$. On a alors $\lambda_1 e_1 = 0$, puis $\lambda_1 = 0$ car $e_1 = e \neq 0$.

Cette famille est génératrice. Soit $v \in V$. Comme (ε_j) est une famille génératrice de W, il existe une famille $(\lambda_j)_{j\geqslant 2}$ de scalaires telle que l'image $\overline{v}$ de v dans W s'écrive $\sum \lambda_j \varepsilon_j$. Cela signifie que $v - \sum_{j\geqslant 2} \lambda_j e_j$ appartient à ke, c'est-à-dire s'écrit $\lambda_1 e = \lambda_1 e_1$ pour un certain scalaire λ_1 ; on a alors $v = \sum_{j\geqslant 1} \lambda_j e_j$.

Soit $g \in G$. On a $ge_1 = e_1$. Soit $j \geqslant 2$. L'égalité $g\varepsilon_j = \varepsilon_j + \sum_{i>j} a_{ij}\varepsilon_i$ signifie que $ge_j - e_j - \sum_{i>j} a_{ij} e_i$ appartient à ke, c'est-à-dire s'écrit $a_{1j}e = a_{1j}e_1$ pour un certain $a_{1j} \in k$. On a alors $ge_j = e_j + \sum_{j<i\leqslant n} a_{ij} e_i$ pour tout j. La matrice de $v \mapsto gv$ dans la base (e_j) est donc triangulaire supérieure avec des 1 sur la diagonale, ce qui achève la démonstration. □

5. Groupes de permutations

5.1. — L'objet de cette section est l'étude du groupe de permutation $\mathfrak{S}_X$ d'un ensemble fini X ; nous allons commencer par quelques remarques d'ordre général.

Notation 5.1.1. — Soit n un entier. Nous noterons $\mathfrak{S}_n$ le groupe $\mathfrak{S}_{\{1,\dots,n\}}$.

Remarque 5.1.2. — Si $n = 0$ l'ensemble $\{1, \dots, n\}$ est vide (voir 1.2.7) et $\mathfrak{S}_0$ est donc réduit à $\{\mathrm{Id}\}_\varnothing$ d'après 1.2.4.

5.1.3. — Soit X un ensemble. La composition de deux permutations σ et τ de X sera souvent simplement notée $\sigma\tau$ au lieu de $\sigma \circ \tau$. Si σ est une permutation de X le *support* $\mathrm{Supp}(\sigma)$ de σ est l'ensemble des éléments x de X tels que $\sigma(x) \neq x$, c'est-à-dire encore le complémentaire de l'ensemble $\mathrm{Fix}(\sigma)$ des points fixes de σ. Les sous-ensembles $\mathrm{Supp}(\sigma)$ et $\mathrm{Fix}(\sigma)$ de X sont tous deux stables sous $\langle\sigma\rangle$; il suffit en effet de le vérifier pour l'un des deux, et x est un point fixe de σ on a bien entendu $\sigma^i(x) = x$ pour tout $i \in \mathbf{Z}$, si bien que $\mathrm{Fix}(\sigma)$ est stable sous $\langle\sigma\rangle$.

5.1.4. — Soient X et Y deux ensembles et soit $\varphi\colon X \to Y$ une bijection. On vérifie immédiatement que la formule $\sigma \mapsto \varphi\sigma\varphi^{-1}$ définit un isomorphisme de groupes de $\mathfrak{S}_X$ sur $\mathfrak{S}_Y$, de réciproque $\tau \mapsto \varphi^{-1}\tau\varphi$. En particulier, si X est un ensemble fini de cardinal n, le groupe $\mathfrak{S}_X$ est isomorphe à $\mathfrak{S}_n$, mais *non canoniquement* dès que $n \geqslant 2$: la construction d'un isomorphisme $\mathfrak{S}_X \simeq \mathfrak{S}_n$ repose en effet sur le choix d'une bijection $X \simeq \{1, \dots, n\}$).

5.1.5. — Soit G un groupe. L'opération de G sur lui-même par translations est fidèle et induit donc un morphisme injectif $G \hookrightarrow \mathfrak{S}_G$ (exemple 4.3.8). Si de plus G est fini alors $\mathfrak{S}_G \simeq \mathfrak{S}_{|G|}$ d'après le 5.1.4 ci-dessus. En conséquence, *tout groupe fini se plonge dans* $\mathfrak{S}_n$ *pour un certain* n. Ce résultat est parfois cité sous le nom de «théorème de Cayley». Mais il ne faut surestimer ni sa profondeur (il est essentiellement tautologique), ni son importance pratique. Il peut en effet donner l'impression que la théorie des groupes finis se ramène à celle des seuls groupes $\mathfrak{S}_n$ et n'est donc somme toute pas bien méchante. Mais c'est illusoire : en effet, l'étude de tous les sous-groupes de tous les groupes $\mathfrak{S}_n$ n'est pas un projet réaliste.

Rappel 5.1.6. — Si n est un entier, on note $n!$ le produit $\prod_{1\leqslant i\leqslant n}$. Notez que si $n = 0$ on a affaire au *produit vide*, qui est par convention égal à l'élément neutre de la multiplication, c'est-à-dire à 1.

On a pour tout entier n l'égalité $(n+1)\cdot n! = (n+1)!$ (remarquez que cette formule vaut aussi pour $n = 0$, puisque $0! = 1$).

Lemme 5.1.7. — *Soient X et Y deux ensembles finis de même cardinal n. L'ensemble des bijections de X vers Y a pour cardinal $n!$. En particulier, $|\mathfrak{S}_X| = n!$.*

Démonstration. — On procède par récurrence sur n. Le résultat est vrai si $n = 0$ puisque $\mathfrak{S}_\varnothing = \{\mathrm{Id}_\varnothing\}$ (remarque 5.1.2).

On suppose $n > 0$ et le résultat vrai pour $n-1$. Comme $n > 0$ il existe un élément $x \in X$. Pour tout $y \in Y$, notons B_y l'ensemble des bijections f de X vers Y qui envoient x sur y. L'ensemble B_y est lui-même en bijection avec l'ensemble des bijections de

$X\backslash\{x\}$ sur $Y\backslash\{y\}$ (par l'application qui associe à une bijection $f \in B_y$ sa restriction à $X\backslash\{x\}$; sa réciproque consiste à prolonger une bijection de $X\backslash\{x\}$ sur $Y\backslash\{y\}$ à X en envoyant x sur y). Il résulte alors de notre hypothèse de récurrence que le cardinal de B_y est égal à $(n-1)!$.

L'ensemble des bijections de X vers Y étant égal à $\coprod_y B_y$, son cardinal vaut

$$\sum_{y \in Y} |B_y| = \sum_{y \in Y} (n-1)! = |Y| \cdot (n-1)! = n \cdot (n-1)! = n!,$$

ce qu'il fallait démontrer. □

5.1.8. *Description explicite de $\mathfrak{S}_X$ en petit cardinal.* — On a vu à la remarque 5.1.2 que $\mathfrak{S}_\varnothing = \{\mathrm{Id}_\varnothing\}$.

Soit X un singleton. Il est immédiat qu'il y a une seule application de X dans lui-même, à savoir Id_X ; par conséquent, $\mathfrak{S}_X = \mathrm{Id}_X$.

Soit X un ensemble ayant exactement deux éléments x et y. En considérant les valeurs possibles que peut prendre une bijection sur x et sur y, on voit que $\mathfrak{S}_X$ a deux éléments : l'identité et la permutation qui échange x et y.

5.2. La notion de cycle. — Nous allons maintenant introduire une classe de permutations particulière, les *cycles*, qui jouent un rôle majeur dans l'étude des permutations quelconques d'un ensemble fini.

5.2.1. — Soit X un ensemble, soit k un entier > 1 et soient $x_1, \ldots, x_k$ des éléments deux à deux distincts de X. On note $(x_1 x_2 \ldots x_k)$ la permutation de X qui envoie x_i sur x_{i+1} pour tout $i < k$, qui envoie x_k sur x_1, et qui fixe tout point de X n'appartenant pas à $\{x_1, \ldots, x_k\}$.

Remarque 5.2.2. — L'écriture $(x_1 \ldots x_k)$ n'est pas unique. Il résulte en effet de la définition que pour tout i compris entre 1 et k on a

$$(x_1 \ldots x_k) = (x_i x_{i+1} \ldots x_{k-1} x_k x_1 \ldots x_{i-1}).$$

5.2.3. — On conserve les notations de 5.2.1. Nous allons énoncer quelques propriétés élémentaires de la permutation $(x_1 \ldots x_k)$.

5.2.3.1. — Le support de $(x_1 \ldots x_k)$ est égal à $\{x_1, \ldots, x_k\}$.

5.2.3.2. — On a $(x_1 \ldots x_k)^k = \mathrm{Id}_X$; par ailleurs on a pour tout $\ell < k$ l'égalité $(x_1 \ldots x_k)^\ell (x_1) = x_{1+\ell} \neq x_1$, et $(x_1 \ldots x_k)^\ell$ n'est donc pas égal à Id_X. Il en résulte que $(x_1 \ldots x_k)$ est d'ordre k.

5.2.3.3. — L'inverse de $(x_1 \ldots x_k)$ est égal à $(x_k x_{k-1} \ldots x_1)$.

5.2.3.4. — Soit φ une bijection de X sur un ensemble Y. On a

$$\varphi (x_1 \ldots x_k) \varphi^{-1} = (\varphi(x_1) \ldots \varphi(x_k)).$$

En effet, soit $y \in Y$. Si y n'appartient pas à $\{\varphi(x_1), \ldots, \varphi(x_n)\}$ alors $\varphi^{-1}(y)$ n'appartient pas à $\{x_1, \ldots, x_n\}$. On a dans ce cas $(x_1 \ldots x_k)(\varphi^{-1}(y)) = \varphi^{-1}(y)$, et $(\varphi (x_1 \ldots x_k) \varphi^{-1})(y) = y$.

Si $y = \varphi(x_i)$ pour un certain i on a $\varphi^{-1}(y) = x_i$, et $((x_1 \ldots x_k)\varphi^{-1})(y)$ est donc égal à x_{i+1} si $i < k$ et à x_1 si $i = k$. Il s'ensuit que $(\varphi(x_1 \ldots x_k)\varphi^{-1})(y)$ est égal à $\varphi(x_{i+1})$ si $i < k$, et à $\varphi(x_1)$ si $i = k$, ce qu'il fallait démontrer.

Définition 5.2.4. — Soit k un entier > 1. Une permutation σ d'un ensemble X est appelée un *cycle* si elle est de la forme $(x_1 \ldots x_k)$ où k est un entier $\geqslant 2$ et où les x_i sont des éléments deux à deux distincts de X.

Commentaires 5.2.5. — Soit σ un cycle sur X et soit $(x_1 \ldots x_k)$ une écriture de σ comme dans la définition 5.2.4. L'ensemble $\{x_1, \ldots, x_k\}$ est alors uniquement déterminé, car c'est le support de σ ; l'entier k l'est donc aussi puisque c'est le cardinal de cet ensemble (on peut également remarquer que c'est l'ordre de σ) ; on l'appelle la *longueur* de σ. La numérotation des x_i n'est par contre pas unique : on peut la décaler, *cf.* remarque 5.2.2. Toutefois, une fois x_1 choisi la valeur des x_i pour $i > 1$ est imposée : on a nécessairement $x_2 = \sigma(x_1), x_2 = \sigma(x_2) = \sigma^2(x_1)$, etc.

Un cycle de longueur k est souvent plus brièvement qualifié de *k-cycle.* Un 2-cycle est également appelé une *transposition.*

5.3. Décomposition d'une permutation en cycles. — Nous nous proposons dans cette section de montrer que toute permutation d'un ensemble fini possède une écriture d'un type particulier, sa *décomposition en cycles*, qui est extrêmement utile (aussi bien en pratique qu'en théorie) et que l'on peut obtenir par un algorithme très simple.

5.3.1. *Produit de permutations à supports deux à deux disjoints.* — Soit X un ensemble et soient $\sigma_1, \ldots, \sigma_r$ des permutations de X à supports deux à deux disjoints. On pose $\sigma = \sigma_1 \ldots \sigma_r$.

5.3.1.1. — Si $x \notin \bigcup \mathrm{Supp}(\sigma_i)$ on a alors $\sigma_i(x) = x$ pour tout i, et $\sigma(x)$ est donc égal à x.

5.3.1.2. — Supposons que x appartienne à $\mathrm{Supp}(\sigma_i)$ pour un certain i (nécessairement unique puisque les σ_i sont à supports deux à deux disjoints). Si $j > i$ alors $\sigma_j(x) = x$ car $x \notin \mathrm{Supp}(\sigma_j)$; et si $j < i$ alors $\sigma_i(x) \notin \mathrm{Supp}(\sigma_j)$ et l'on a donc $\sigma_j(\sigma_i(x)) = \sigma_i(x)$. Il vient $\sigma(x) = \sigma_i(x)$.

5.3.1.3. — Récapitulons : on a $(\sigma_1 \ldots \sigma)_r(x) = x$ si $x \notin \bigcup \mathrm{Supp}(\sigma_i)$, et $(\sigma_1 \ldots \sigma_r)(x) = \sigma_i(x)$ si x appartient à $\mathrm{Supp}(\sigma_i)$. Cette description montre que le support de $\sigma_1 \ldots \sigma_r$ est égal à $\bigcup \mathrm{Supp}(\sigma_i)$, que $\sigma_1 \ldots \sigma_r$ est égale à l'identité si et seulement si chacune des σ_i est égale à l'identité, et que $\sigma_1 \ldots \sigma_r$ *ne dépend pas de l'ordre des* σ_i ; le produit de permutations *à supports deux à deux disjoints* est donc commutatif.

Proposition 5.3.2. — *Soit X un ensemble fini et soit $\sigma \in \mathfrak{S}_X$. Il existe une famille finie $(C_1, \ldots, C_r)$ de cycles à supports deux à deux disjoints tels que $\sigma = C_1 \ldots C_r$. Cette famille est unique à permutation près des C_i*

Démonstration. — Commençons par l'existence. La formule $(k,x) \mapsto \sigma^k(x)$ définit une opération de $\mathbf{Z}$ sur X ; dans le reste de la preuve les termes «orbite» et «stabilisateur» feront implicitement référence à cette action.

Les orbites singleton sont les singletons de la forme $\{x\}$ avec $\sigma(x) = x$. Soient $S_1, \ldots, S_r$ les orbites *non singleton*; chacune d'elle est stable sous l'action de $\mathbf{Z}$, c'est-à-dire sous $\langle\sigma\rangle$. Fixons i, notons d le cardinal de S_i et choisissons un élément x dans S_i. Le stabilisateur de x dans $\mathbf{Z}$ ayant pour indice $|S_i| = d$, il est égal à $d\mathbf{Z}$. Si i et j sont deux éléments de $\mathbf{Z}$ on a

$$(\sigma^i(x) = \sigma^j(x)) \iff (\sigma^{i-j}(x) = x) \iff (i-j) \in d\mathbf{Z}.$$

On en déduit que $x, \sigma(x), \ldots, \sigma^{d-1}(x)$ sont des éléments deux à deux distincts de S_i. Comme $|S_i| = d$, on a $S_i = \{x, \sigma(x), \ldots, \sigma^{d-1}(x)\}$. Si i est compris entre 0 et $d-1$ alors $\sigma(\sigma^i(x)) = \sigma^{i+1}(x)$; et $\sigma(\sigma^{d-1}(x)) = \sigma^d(x) = x$ car d appartient au stabilisateur de x. Si l'on pose $C_i = (x\sigma(x)\ldots\sigma^{d-1}(x))$ on a donc $\sigma|_{S_i} = C_i|_{S_i}$.

Les C_i sont des cycles dont les supports S_i sont deux à deux disjoints ; si $a \in X$ on a $\sigma(a) = a$ si $a \notin \bigcup S_i$; sinon, a appartient à S_i pour un unique i et $\sigma(a) = C_i(a)$. Par conséquent, $\sigma = C_1 \ldots C_r$ (5.3.1.3).

Établissons maintenant l'unicité. Donnons-nous une écriture $\sigma = D_1 \ldots D_s$ où les D_j sont des cycles à supports deux à deux disjoints. Fixons j. Le support $\mathrm{Supp}(D_j)$ est stable sous $\langle\sigma\rangle$, et $\sigma^i(x) = D_j^i(x)$ pour tout $x \in \mathrm{Supp}(D_j)$ et tout $i \in \mathbf{Z}$. Il en résulte immédiatement que si x appartient à $\mathrm{Supp}(D_j)$ alors $\mathrm{Supp}(D_j) = \{\sigma^i(x)\}_{i\in\mathbf{Z}}$; autrement dit, $\mathrm{Supp}(D_j)$ est une orbite non singleton (son cardinal est égal à la longueur de D_j qui est $\geqslant 2$ par définition d'un cycle). Par ailleurs, le complémentaire de $\bigcup \mathrm{Supp}(D_j)$ est l'ensemble $\mathrm{Fix}(\sigma)$, c'est-à-dire l'ensemble des éléments x de $\{1, \ldots, n\}$ tels que $\{x\}$ soit une orbite. En conséquence, les supports des D_j sont *exactement* les orbites non singleton. Autrement dit on a $r = s$, et quitte à renuméroter les D_i on peut supposer que D_i a pour support S_i pour tout i. Mais on a alors pour tout i les égalités $D_i(x) = \sigma(x) = C_i(x)$ si $x \in S_i$, et $D_i(x) = C_i(x) = x$ si $x \notin S_i$. Ainsi $C_i = D_i$ pour tout i, ce qui achève la démonstration. □

5.3.3. *Commentaires et premiers exemples.* — Soit X un ensemble fini et soit σ une permutation appartenant à $\mathfrak{S}_X$.

5.3.3.1. — La proposition 5.3.2 assure que σ possède une écriture comme produit de cycles à supports deux à deux disjoints. Cette écriture n'est unique qu'à permutation des cycles près ; mais comme le produit de permutations à supports deux à deux disjoints est commutatif (5.3.1.3), on ne pouvait espérer mieux.

Pour alléger un peu l'expression, cette écriture sera simplement appelée la *décomposition en cycles* de la permutation.

5.3.3.2. *Les cas triviaux.* — Si $\sigma = \mathrm{Id}_X$, sa décomposition en cycles est le *produit vide* de cycles. Si σ est un cycle, sa décomposition en cycles est tout simplement σ.

5.3.3.3. — On ne fait plus d'hypothèse particulière sur σ. Soit $C_1 \ldots C_r$ sa décomposition en cycles. Pour tout i, notons ℓ_i la longueur de C_i.

Le support de σ est la réunion des supports des C_i (5.3.1.3).

Comme les C_i commutent deux à deux d'après *loc. cit.*, on a pour tout entier m l'égalité $\sigma^m = C_1^m C_2^m \dots C_r^m$ pour tout m. Par ailleurs si m est un entier, le support de C_i^m est contenu pour tout i dans le support de C_i ; les C_i^m sont donc à supports deux à deux disjoints, et il résulte alors de *loc. cit.* que $\sigma^m = \mathrm{Id}$ si et seulement si $C_i^m = \mathrm{Id}$ pour tout i. L'ordre de σ est par conséquent égal au PPCM des ordres des C_i, c'est-à-dire au PPCM des ℓ_i.

5.3.4. *L'algorithme de décomposition : description théorique.* — Soit σ une permutation d'un ensemble fini X. Pour obtenir sa décomposition en cycles, on suit peu ou prou la preuve de l'existence de cette décomposition. Donnons quelques précisions.

Introduisons tout d'abord une notation. Si x est un élément de X qui n'est pas un point fixe de σ, le plus petit entier $d > 0$ tel que $\sigma^d(x) = x$ (qu'on obtient en appliquant σ de façon répétée jusqu'à retomber sur x) est $\geqslant 2$, et $(x\sigma(x)\dots,\sigma^{d-1}(x))$ est un cycle, noté $C(x)$.

Pour obtenir l'écriture cherchée de σ on procède alors comme suit. si $\sigma = \mathrm{Id}$, il n'y a rien à faire. Sinon, il existe x tel que $\sigma(x) \neq x$, et on construit récursivement une suite finie $(x_1,\dots,x_r)$, constituée de points qui n'appartiennent pas à Fix σ, par le procédé suivant. On pose $x_1 = x$. Si $x_1,\dots,x_i$ ont été construits on distingue deux cas : ou bien $\sigma(x) = x$ pour tout x en dehors de $\bigcup_{1\leqslant j\leqslant i} \mathrm{Supp}(C(x_j))$, et on arrête ; ou bien il existe un élément y en dehors de $\bigcup_{1\leqslant j\leqslant i} \mathrm{Supp}(C(x_j))$ tel que $\sigma(y) \neq y$, et l'on pose $x_{i+1} = y$.

L'écriture cherchée est alors $\sigma = C(x_1)C(x_2)\dots C(x_r)$.

Exemple 5.3.5. — Soit $\sigma \in \mathfrak{S}_{15}$ la permutation donnée par le tableau des valeurs suivant :

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 \\ 5 & 8 & 7 & 11 & 9 & 6 & 12 & 4 & 2 & 13 & 3 & 1 & 14 & 10 & 15 \end{pmatrix}.$$

L'élément 1 n'est pas fixe. Le cycle $C(1)$ est égal à

$$(\ 1\ \ 5\ \ 9\ \ 2\ \ 8\ \ 4\ \ 11\ \ 3\ \ 7\ \ 12\ ).$$

L'élément 10 n'appartient pas au support de $C(1)$, et n'est pas fixe sous σ. Le cycle $C(10)$ est égal à

$$(\ 10\ \ 13\ \ 14\ ).$$

Le complémentaire de $\mathrm{Supp}(C(1)) \cup \mathrm{Supp}(C(2))$ est égal à $\{6, 15\}$, et 6 et 15 appartiennent tous deux à $\mathrm{Fix}(\sigma)$. L'algorithme s'arrête donc, et la décomposition en cycles de σ est par conséquent

$$(\ 1\ \ 5\ \ 9\ \ 2\ \ 8\ \ 4\ \ 11\ \ 3\ \ 7\ \ 12\ )\ (\ 10\ \ 13\ \ 14\ ).$$

L'ordre de σ est dès lors égal à PPCM(10,3)=30.

Remarque 5.3.6. — Observez à quel point cette méthode est efficace pour obtenir l'ordre, en regard de l'algorithme brutal qui consisterait à calculer les puissances successives de σ.

Notation 5.3.7. — Soit X un ensemble fini et soit $\sigma \in \mathfrak{S}_X$. Soit $C_1 \dots C_r$ la décomposition de σ en cycles. Pour tout entier ℓ on désigne par $e(\sigma, \ell)$ le cardinal de l'ensemble des indices i compris entre 1 et r tels que C_i soit de longueur ℓ (ce n'est pas une notation standard) ; on a $e(\sigma, \ell) = 0$ pour presque tout ℓ.

Proposition 5.3.8 **(Classes de conjugaison de $\mathfrak{S}_X$).** — *Soit X un ensemble fini et soient σ et τ deux éléments de $\mathfrak{S}_X$. Les assertions suivantes sont équivalentes :*

(i) *les éléments σ et τ de $\mathfrak{S}_X$ sont conjugués ;*

(ii) *on a $e(\sigma, \ell) = e(\tau, \ell)$ pour tout ℓ.*

Démonstration. — Écrivons σ comme un produit $C_1 \dots C_r$ de cycles à supports deux à deux disjoints ; pour tout i, on note ℓ_i la longueur de C_i, et l'on choisit une écriture $(a_{i1} \dots a_{i\ell_i})$ de C_i.

Supposons que σ et τ soient conjugués. Il existe alors $\varpi \in \mathfrak{S}_X$ telle que $\tau = \varpi\sigma\varpi^{-1}$. Pour tout i on pose $D_i = \varpi C_i \varpi^{-1}$. On a $\tau = D_1 \dots D_r$. Il résulte de 5.2.3.4 que $D_i = (\varpi(a_{i1}) \dots \varpi(a_{i\ell_i}))$ pour tout i ; c'est un cycle de longueur ℓ_i et de support $\varpi(\mathrm{Supp}(C_i))$. Les D_i sont donc des cycles à supports deux à deux disjoints (ϖ est injective) ; puisque chacun d'eux a pour longueur ℓ_i, on a $e(\tau, \ell) = e(\sigma, \ell)$ pour tout ℓ.

Réciproquement, supposons que $e(\tau, \ell) = e(\sigma, \ell)$ pour tout ℓ. Soit $D_1 \dots D_s$ l'écriture de τ comme produit de cycles à supports deux à deux disjoints. Notre hypothèse signifie que $s = r$ et que l'on peut renuméroter les D_i de sorte que D_i soit de longueur ℓ_i pour tout i. Choisissons pour tout i une écriture $(b_{i1} \dots b_{i\ell_i})$ de D_i. Les a_{ij} sont deux à deux distincts, et les b_{ij} aussi. La formule $a_{ij} \mapsto b_{ij}$ définit donc sans ambiguïté une bijection de $\{a_{ij}\}$ sur $\{b_{ij}\}$, que l'on prolonge arbitrairement en une permutation $\varpi \in \mathfrak{S}_X$. On a alors $\varpi_i \varpi^{-1} = D_i$ pour tout i (toujours d'après 5.2.3.4) ; il vient $\varpi\sigma\varpi^{-1} = \tau$. □

5.3.9. — Soit X un ensemble, soit ℓ un entier $\geqslant 2$ et soit $C = (a_1 \dots a_\ell)$ un cycle de longueur ℓ de $\mathfrak{S}_X$. On vérifie immédiatement l'égalité

$$C = (a_1 a_2)(a_2 a_3) \dots (a_{\ell-2} a_{\ell-1})(a_{\ell-1} a_\ell).$$

Le cycle C est donc le produit de $\ell - 1$ transpositions.

Supposons de plus que X est fini. Dans ce cas tout élément de $\mathfrak{S}_X$ est un produit de cycles (que l'on peut même choisir à supports deux à deyx disjoints), et est donc par ce qui précède un produit de transpositions. Cette écriture est loin d'être unique dès que $|X| \geqslant 2$, par exemple parce que l'identité est égal au produit vide de transpositions, mais aussi à $(xy)(xy)$ si x et y sont deux éléments distincts de X (si vous n'aimez pas ça, vous pouvez «décaler» cet exemple et remarquer que $(xy) = (xy)(xy)(xy)$). Nous verrons toutefois à la section suivante que les différentes écritures d'une même permutation comme produit de transpositions font toutes intervenir le même nombre de transpositions *modulo 2*.

5.4. Signature d'une permutation. — Nous allons maintenant définir pour tout ensemble fini X un morphisme de groupes de $\mathfrak{S}_X$ dans $\{-1, 1\}$ appelé la *signature*. Il y a différentes façons de procéder ; nous présentons ici celle de Bourbaki, concise et très astucieuse.

5.4.1. *Construction de la signature sur* $\mathfrak{S}_n$. — Soit n un entier. Pour tout $\sigma \in \mathfrak{S}_n$, il existe un unique morphisme d'anneaux de $\mathbf{Z}[X_1, \ldots, X_n]$ dans lui-même qui envoie X_i sur $X_{\sigma(i)}$ pour tout i. On le note $P \mapsto \sigma \cdot P$. On vérifie immédiatement (il suffit de s'en assurer sur chacun des X_i) que $\mathrm{Id} \cdot P = P$ pour tout P, et que $\sigma \cdot (\tau \cdot P) = (\sigma\tau) \cdot P$ pour tout (σ, τ, P) ; on a ainsi défini une opération de $\mathfrak{S}_n$ sur $\mathbf{Z}[X_1, \ldots, X_n]$ par automorphismes d'anneaux.

5.4.1.1. — Soit Δ l'élément $\prod_{i<j}(X_j - X_i)$ de $\mathbf{Z}[X_1, \ldots, X_n]$. On a

$$\Delta^2 = \prod_{i<j}(X_j - X_i)^2 = \prod_{i<j}(-1)(X_j - X_i)(X_i - X_j) = (-1)^{n(n-1)/2}\prod_{i \neq j}(X_i - X_j).$$

Cette dernière écriture montre que Δ^2 est invariant par permutation des indéterminées, c'est-à-dire que $\sigma \cdot (\Delta^2) = \Delta^2$ pour tout $\sigma \in \mathfrak{S}_n$.

5.4.1.2. — Soit $\sigma \in \mathfrak{S}_n$. On a $\Delta^2 = \sigma \cdot (\Delta^2) = (\sigma \cdot \Delta)^2$. Comme $\mathbf{Z}[X_1, \ldots, X_n]$ est intègre, il s'ensuit qu'il existe $\varepsilon(\sigma) \in \{-1, 1\}$ tel que $\sigma \cdot \Delta = \varepsilon(\sigma)\Delta$.

5.4.1.3. — Soient σ et $\tau \in \mathfrak{S}_n$. On a

$$\begin{aligned} \varepsilon(\sigma\tau)\Delta &= (\sigma\tau) \cdot \Delta \\ &= \sigma \cdot (\tau \cdot \Delta) \\ &= \sigma \cdot (\varepsilon(\tau)\Delta) \\ &= \varepsilon(\tau)\sigma \cdot \Delta \\ &= \varepsilon(\tau)\varepsilon(\sigma)\Delta. \end{aligned}$$

En utilisant encore l'intégrité de $\mathbf{Z}[X_1, \ldots, X_n]$ (et la commutativité de $\{-1, 1\}$) on voit que $\varepsilon(\sigma\tau) = \varepsilon(\sigma)\varepsilon(\tau)$. Par conséquent, ε est un morphisme de groupes de $\mathfrak{S}_n$ vers $\{-1, 1\}$, appelé la *signature*.

Proposition-définition 5.4.2. — *Soit X un ensemble fini de cardinal n et soit Φ une bijection de X sur $\{1, \ldots, n\}$. Le morphisme de groupes*

$$\begin{aligned} \mathfrak{S}_X &\to \{-1, 1\} \\ \sigma &\mapsto \varepsilon(\Phi\sigma\Phi^{-1}) \end{aligned}$$

ne dépend pas de Φ. On le note le plus souvent encore ε, et on l'appelle encore la signature.

Démonstration. — Soit Ψ une (autre) bijection de X sur $\{1, \ldots, n\}$. Soit $\sigma \in \mathfrak{S}_X$. On a

$$\Psi\sigma\Psi^{-1} = (\Psi\Phi^{-1})\Phi\sigma\Phi^{-1}(\Phi\Psi^{-1}) = \tau(\Phi\sigma\Phi^{-1})\tau^{-1},$$

où τ désigne la bijection $\Psi\Phi^{-1}$ de $\{1, \ldots, n\}$ sur lui-même. Les permutations $\Phi\sigma\Phi^{-1}$ et $\Psi\sigma\Psi^{-1}$ sont donc conjugées dans $\mathfrak{S}_n$; par conséquent, leurs images par le morphisme ε sont conjuguées dans $\{-1, 1\}$ (remarque 2.7.6), et finalement égales puisque ce dernier est abélien. □

Remarque 5.4.3. — Lorsque $X = \{1, \ldots, n\}$, on retrouve bien la signature ε définie au 5.4.1.3 (prendre $\Phi = \mathrm{Id}_{\{1,\ldots,n\}}$) ; nos choix en matière de terminologie et de notation sont donc cohérents.

***Exemple 5.4.4* (signature d'une transposition).** — Soit X un ensemble fini et soit $\tau = (x\ y)$ une transposition de X.

Choisissons une bijection Φ de X sur $\{1, \ldots, n\}$ qui envoie x sur 1 et y sur 2. On a

$$\varepsilon(\tau) = \varepsilon(\Phi(x\ y)\Phi^{-1}) = \varepsilon((1\ 2)).$$

Il reste à calculer ce dernier terme, en reprenant la définition de $\varepsilon\colon \mathfrak{S}_n \to \{-1, 1\}$ donnée au 5.4.1.3.

On a

$$\begin{aligned} \Delta &= \prod_{i<j}(X_j - X_i) \\ &= (X_2 - X_1)\prod_{j>2}(X_j - X_1)\prod_{j>2}(X_j - X_2)\prod_{j>i>2}(X_j - X_i). \end{aligned}$$

L'application (1 2) remplace $(X_2 - X_1)$ par $(X_1 - X_2)$, échange les deux facteurs $\prod_{j>2}(X_j - X_1)$ et $\prod_{j>2}(X_j - X_2)$, et laisse invariant le produit $\prod_{j>i>2}(X_j - X_i)$. Par conséquent $(1\ 2) \cdot \Delta = -\Delta$; ainsi, $\varepsilon((1\ 2)) = -1$. Il s'ensuit que $\varepsilon(\tau) = -1$.

5.4.5. *Signature et décomposition en produit de transpositions.* — Soit X un ensemble fini et soit $\sigma \in \mathfrak{S}_X$. On sait d'après 5.3.9 qu'elle s'écrit comme un produit $\tau_1\tau_2\ldots\tau_r$ de transpositions. Il résulte alors de l'exemple 5.4.4 que $\varepsilon(\sigma) = (-1)^r$. En particulier, *la classe de r modulo 2 ne dépend pas de l'écriture $\tau_1\ldots\tau_r$ choisie.*

On dit que σ est *paire* (resp. *impaire*) si sa signature est 1 (resp. -1). Par ce qui précède, σ est pâire (resp. impaire) si et seulement si elle s'écrit comme le produit d'un nombre pair (resp. impair) de transpositions.

5.4.6. *Calcul de la signature en général.* — Soit X un ensemble fini. Soit C un cycle de $\mathfrak{S}_X$ et soit ℓ sa longueur. On a vu au 5.3.9 que C s'écrit comme un produit de $\ell - 1$ transpositions. On déduit alors de 5.4.5 que $\varepsilon(C) = (-1)^{\ell-1}$.

Soit maintenant σ une permutation quelconque de $\mathfrak{S}_X$. Soit $C_1 \ldots C_s$ la décomposition de σ en cycles. Pour tout i, notons ℓ_i la longueur de C_i. Il résulte de ce qui précède que

$$\varepsilon(\sigma) = \prod_i (-1)^{\ell_i - 1} = (-1)^{\sum_i \ell_i - r}.$$

En pratique, on calcule le plus souvent la signature d'une permutation en effectuant sa décomposition en cycles et en appliquant la formule ci-dessus. Par exemple, supposons que $X = \{1, \ldots, 15\}$ et soit σ la permutation

$$\left|\begin{array}{ccccccccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 \\ 7 & 15 & 1 & 2 & 8 & 11 & 9 & 10 & 6 & 14 & 3 & 4 & 5 & 12 & 13 \end{array}\right. .$$

Sa décomposition en cycles est

$$\left(\begin{array}{cccccc} 1 & 7 & 9 & 6 & 11 & 3 \end{array}\right)\left(\begin{array}{ccccccccc} 2 & 15 & 13 & 5 & 8 & 10 & 14 & 12 & 4 \end{array}\right).$$

Il vient $\varepsilon(\sigma) = (-1)^{6+9-2} = -1$.

5.4.7. — Soit X un ensemble fini et soit n son cardinal. On note $\mathfrak{A}_X$ le sous-groupe distingué $\operatorname{Ker}\varepsilon$ de $\mathfrak{S}_X$ (c'est donc l'ensemble des permutations paires de X).

Si $n = 0$ ou $n = 1$ alors $\mathfrak{S}_X = \{\mathrm{Id}\}$, l'image de ε est $\{1\}$ et $\mathfrak{A}_X = \{\mathrm{Id}\}$.

Si $n \geqslant 2$ alors ε est surjective, puisque $(x\ y)$ est de signature (-1) pour tout couple (x, y) d'éléments distincts de X. Par conséquent, $\mathfrak{A}_X$ est d'indice 2 et $|\mathfrak{A}_X| = \dfrac{n!}{2}$.

5.4.8. — Nous allons maintenant décrire les groupes $\mathfrak{S}_3, \mathfrak{A}_3, \mathfrak{S}_4$ et $\mathfrak{A}_4$ en donnant la liste de leurs permutations, présentées *via* leurs décompositions en cycles.

- Liste des éléments de $\mathfrak{S}_3$ selon la nature de leur décomposition.
  - ◇ Produit vide de cycles : Id.
  - ◇ Transpositions : $(12), (13), (23)$.
  - ◇ 3-cycles : $(123), (132)$.
- Liste des éléments de $\mathfrak{A}_3$: $\mathrm{Id}, (123), (132)$.
- Liste des éléments de $\mathfrak{S}_4$ selon la nature de leur décomposition.
  - ◇ Produit vide de cycles : Id.
  - ◇ Transpositions : $(12), (13), (14), (23), (24), (34)$.
  - ◇ 3-cycles : $(234), (243), (134), (143), (124), (142), (123), (132)$.
  - ◇ 4-cycles : $(1234), (1243), (1324), (1342), (1423), (1432)$.
  - ◇ Produits de 2 transpositions : $(12)(34), (13)(24), (14)(23)$.
- Liste des éléments de $\mathfrak{A}_4$:

$$\mathrm{Id}, (234), (243), (134), (143), (124), (142), (123), (132), (12)(34), (13)(24), (14)(23).$$

5.5. Quelques isomorphismes exceptionnels. — Nous allons pour terminer ce chapitre exhiber quelques isomorphismes entre certains groupes de permutations et certains groupes linéaires (sur un corps fini). Ces isomorphismes sont dits *exceptionnels* car ils ne s'inscrivent dans aucune série systématique ; ce sont en quelques sorte d'heureuses coïncidences en petit cardinal et basse dimension.

5.5.1. *Un calcul de cardinal.* — Soit p un nombre premier et soit n un entier $\geqslant 1$. Se donner une matrice appartenant à $\mathrm{GL}_n(\mathbf{F}_p)$ revient à choisir une première colonne non nulle (il y a $p^n - 1$ choix), puis une seconde colonne qui n'est pas dans la droite engendrée par la première (ce qui fait $p^n - p$ choix), puis une troisième colonne qui n'est pas dans le plan engendré par les deux premières (ce qui fait $p^n - p^2$ choix), etc. Par conséquent,

$$|\mathrm{GL}_n(\mathbf{F}_p)| = (p^n - 1)(p^n - p)(p^n - p^2)\dots(p^n - p^{n-1}).$$

5.5.2. *Une remarque générale.* — On désigne toujours par p un nombre premier. On identifie $\mathrm{GL}_2(\mathbf{F}_p)$ à $\mathrm{GL}(\mathbf{F}_p^2)$. L'action tautologique de $\mathrm{GL}_2(\mathbf{F}_p)$ sur $\mathbf{F}_p^2$ en induit une sur l'ensemble $\mathbf{P}^1(\mathbf{F}_p)$ des droites vectorielles de $\mathbf{F}_p^2$, d'où un morphisme de $\mathrm{GL}_2(\mathbf{F}_p)$ dans $\mathfrak{S}_{\mathbf{P}^1(\mathbf{F}_p)}$. Son noyau est l'ensemble des bijections linéaires u de $\mathbf{F}_p^2$ dans lui-même telles que $u(D) = D$ pour toute droite vectorielle D de $\mathbf{F}_p^2$, c'est-à-dire encore telles que $u(v)$ soit colinéaire à v pour tout vecteur v de $\mathbf{F}_p^2$. On vérifie facilement (faites l'exercice !) que cette dernière condition est satisfaite si et seulement si u est une homothétie (de rapport nécessairement non nul, puisque u est bijective). Autrement dit, le noyau du morphisme $\mathrm{GL}_2(\mathbf{F}_p) \to \mathfrak{S}_{\mathbf{P}^1(\mathbf{F}_p)}$ considéré ici est $\{\lambda \mathrm{I}_2\}_{\lambda \in \mathbf{F}_p^\times}$. Ce morphisme induit donc un morphisme injectif $\mathrm{PGL}_2(\mathbf{F}_p) \hookrightarrow \mathfrak{S}_{\mathbf{P}^1(\mathbf{F}_p)}$ (rappelons que par définition, $\mathrm{PGL}_2(\mathbf{F}_p)$ est le quotient de $\mathrm{GL}_2(\mathbf{F}_p)$ par $\{\lambda \mathrm{I}_2\}_{\lambda \in \mathbf{F}_p^\times}$; *cf.* la remarque 2.15.3).

L'ensemble $\mathbf{P}^1(\mathbf{F}_p)$ s'identifie par ailleurs à $\mathbf{F}_p \cup \{\infty\}$ (4.3.13) ; il est donc de cardinal $p+1$. En en choisissant une numérotation arbitraire, on obtient ainsi par ce qui précède un morphisme injectif $\mathrm{PGL}_2(\mathbf{F}_p) \hookrightarrow \mathfrak{S}_{p+1}$.

5.5.3. *L'isomorphisme* $\mathrm{GL}_2(\mathbf{F}_2) \simeq \mathfrak{S}_3$. — En appliquant la construction précédente lorsque $p = 2$, on obtient un morphisme injectif $\varphi \colon \mathrm{PGL}_2(\mathbf{F}_2) \hookrightarrow \mathfrak{S}_3$. Mais puisque $\mathbf{F}_2^\times = \{1\}$, le quotient $\mathrm{PGL}_2(\mathbf{F}_2)$ est simplement $\mathrm{GL}_2(\mathbf{F}_2)$. De plus, en vertu de 5.5.1, on a $|\mathrm{GL}_2(\mathbf{F}_2)| = (2^2-1)(2^2-2) = 6$; les groupes $\mathrm{GL}_2(\mathbf{F}_2)$ et $\mathfrak{S}_3$ ont donc même cardinal. Par conséquent, le morphisme injectif φ est un isomorphisme.

5.5.4. *L'isomorphisme* $\mathrm{PGL}_2(\mathbf{F}_3) \simeq \mathfrak{S}_4$. — En appliquant la construction précédente lorsque $p = 3$, on obtient un morphisme injectif $\psi \colon \mathrm{PGL}_2(\mathbf{F}_3) \hookrightarrow \mathfrak{S}_4$. De plus, en vertu de 5.5.1 on a $|\mathrm{GL}_2(\mathbf{F}_3)| = (3^2-1)(3^2-3) = 48$; et comme $\mathbf{F}_3^\times = \{-1, 1\}$, on en déduit que le quotient $\mathrm{PGL}_2(\mathbf{F}_3)$ est de cardinal $48/2 = 24$. Les groupes $\mathrm{PGL}_2(\mathbf{F}_3)$ et $\mathfrak{S}_4$ ont donc même cardinal. Par conséquent, le morphisme injectif ψ est un isomorphisme.

5.5.5. *Le plongement de* PGL_5 *dans* $\mathfrak{S}_6$. — En appliquant la construction précédente lorsque $p = 5$, on obtient un morphisme injectif $\chi \colon \mathrm{PGL}_2(\mathbf{F}_5) \hookrightarrow \mathfrak{S}_6$.

Le morphisme χ *n'est pas surjectif.* En effet, en vertu de 5.5.1 on a l'égalité $|\mathrm{GL}_2(\mathbf{F}_5)| = (5^2-1)(5^2-5) = 480$; et comme $\mathbf{F}_5^\times = \{-2, -1, 1, 2\}$, on en déduit que le quotient $\mathrm{PGL}_2(\mathbf{F}_5)$ est de cardinal $480/4 = 120$. Le cardinal de $\mathfrak{S}_6$ étant égal à 720, χ ne peut être surjectif.

Remarque 5.5.6. — On peut démontrer que $\mathrm{PGL}_5 \simeq \mathfrak{S}_5$. Mais le morphisme χ ci-dessus n'est pas l'un des plongements standard de $\mathfrak{S}_5$ dans $\mathfrak{S}_6$, consistant à identifier $\mathfrak{S}_5$ au stabilisateur d'un entier $n \in \{1, \ldots, 6\}$ (un tel stabilisateur est canoniquement isomorphe à $\mathfrak{S}_{\{1,\ldots,6\} \setminus \{n\}}$, et donc non canoniquement isomorphe à $\mathfrak{S}_5$). Pour le voir, on remarque que l'action de $\mathfrak{S}_5$ sur $\{1, \ldots, 6\}$ fournie par un tel plongement possède par construction un point fixe. Or l'action de $\mathfrak{S}_5$ sur $\mathbf{P}^1(\mathbf{F}_5)$ qui définit χ est sans point fixe ; en effet, si D est une droite vectorielle de $\mathbf{F}_5^2$ il existe toujours une bijection linéaire u de $\mathbf{F}_5^2$ sur lui-même telle que $u(D) \neq D$.

6. Le produit semi-direct

6.1. Le produit semi-direct interne. — Dans cette section, nous allons introduire et étudier le *produit semi-direct*. On peut y penser comme à une variante non (nécessairement) commutative de la somme directe de deux groupes abéliens ; comme cette dernière, le produit semi-direct se décline en deux versions, l'une interne et l'autre externe. Nous allons commencer par la version interne, que nous définirons après quelques préliminaires.

6.1.1. *Sous-groupe engendré par deux sous-groupes.* — Soit G un groupe et soient H et F deux sous-groupes de G. Comme H et F sont stables par inversion, on a $(H \cup F)^{-1} = H \cup F$. Le sous-groupe $< H \cup F >$ de G est en conséquence constitué des produits finis d'éléments de $H \cup F$. Puisque le produit de deux éléments de H (resp. F) appartient à H (resp. F), tout produit fini d'éléments de $H \cup F$ peut s'écrire en faisant alterner un élément de H et un élément de F. De plus, quitte à rajouter l'élément neutre (qui appartient à H et à F) au début et/ou à la fin d'un tel produit, on peut toujours supposer qu'il porte sur une famille non vide, et qu'il commence (resp. se termine) par un élément de H (resp. de F). Autrement dit, $< H \cup F >$ est l'ensemble des éléments de G qui sont de la forme $h_1 f_1 \ldots h_r f_r$ où $r \geqslant 1$ et où h_i (resp. f_i) appartient à H (resp. F) pour tout i.

Nous noterons HF le sous-ensemble de G formé des éléments qui peuvent s'écrire comme un produit hf avec $h \in H$ et $f \in F$. Il est clair que HF est contenu dans $\langle H \cup F \rangle$, mais cette inclusion est stricte en général ; notez bien que HF n'a aucune raison d'être un sous-groupe de G, car il n'est *a priori* stable ni par l'inversion, ni par le produit.

Lemme 6.1.2. — *Soit G un groupe et soient H et F deux sous-groupes de G avec $H \lhd G$. On a l'égalité $< H \cup F > = HF$.*

Démonstration. — Il suffit de montrer que $< H \cup F >$ est contenu dans HF. Soit r un entier non nul et soient $h_1, \ldots, h_r$ (resp. $f_1, \ldots, f_r$) des éléments de H (resp. F). Nous allons prouver par récurrence sur r que $h_1 f_1 \ldots h_r f_r \in HF$, ce qui établira le lemme.

Pour $r = 1$ c'est évident. On suppose donc $r > 1$ et l'assertion établie pour $r - 1$. En vertu de l'hypothèse de récurrence, $h_1 f_1 \ldots h_{r-1} f_{r-1} \in HF$; il existe donc h_0 dans H et f_0 dans F tels que $h_1 f_1 \ldots h_{r-1} f_{r-1} = h_0 f_0$. Dès lors

$$h_1 f_1 \ldots h_r f_r = h_0 f_0 h_r f_r = h_0 f_0 h_r f_0^{-1} f_0 f_r = hf,$$

où l'on a posé $h = h_0 f_0 h_r f_0^{-1}$ et $f = f_0 f_r$. *Comme H est distingué dans G*, $f_0 h_r f_0^{-1}$ appartient à H, d'où il découle que $h \in H$. Il est par ailleurs clair que $f \in F$, et $h_1 f_1 \ldots h_r f_r = hf$ appartient de ce fait à HF. □

6.1.3. *Définition du produit semi-direct interne.* — Soit G un groupe et soient H et F deux sous-groupes de G. On suppose que le sous-groupe H est distingué dans G, que $H \cap F = \{e\}$, et que $< H \cup F > = G$, cette dernière hypothèse pouvant se réécrire $HF = G$ d'après le lemme 6.1.2 ci-dessus.

6.1.3.1. — Comme $HF = G$, tout élément de G a une écriture de la forme hf avec $h \in H$ et $f \in F$. *Montrons qu'une telle écriture est unique.* Supposons donc que $h_1 f_1 = h_2 f_2$, où h_1 et h_2 appartiennent à H, et f_1 et f_2 à F. On a alors $h_1^{-1}h_2 = f_2^{-1}f_1$. Le terme de gauche est un élément de H, celui de droite un élément de F. Comme $H \cap F = \{e\}$, ces termes sont tous deux égaux à e ; en conséquence, $h_1 = h_2$ et $f_1 = f_2$.

6.1.3.2. — Tout élément de G a donc une *unique* écriture de la forme hf avec $h \in H$ et $f \in F$. On se propose maintenant de comprendre l'effet de la loi de groupe de G sur ce type de décomposition. Pour cela, il est commode d'introduire la notation suivante : si $f \in F$, on désignera par $\varphi(f)$ l'automorphisme $h \mapsto fhf^{-1}$ de H (que cette formule définisse un automorphisme de H résulte du fait que ce dernier est distingué) ; l'application φ est un morphisme de groupes de F dans Aut H.

Soient $h \in H$ et $f \in F$. On peut écrire $fh = fhf^{-1}f = \varphi(f)(h)f$. Notons que $fh = hf$ si et seulement si $\varphi(f)(h) = h$. Le morphisme φ mesure donc en un sens le défaut de commutation des sous-groupes H et F de G : il est trivial (*i.e.* $\varphi(f) = \mathrm{Id}_H$ pour tout $f \in F$) si et seulement si $hf = fh$ pour tout couple $(h, f) \in H \times F$. La loi de groupe de G peut maintenant se décrire facilement. Prenons deux éléments h_1 et h_2 de H, et deux éléments f_1 et f_2 de F. De l'égalité $f_1 h_2 = \varphi(f_1)(h_2) f_1$ il vient :

$$(*) \quad h_1 f_1 h_2 f_2 = (\underbrace{h_1 \varphi(f_1)(h_2)}_{\in H})(\underbrace{f_1 f_2}_{\in F}).$$

On dit que G est le *produit semi-direct interne* de H et F, et que F *opère sur* H via φ. Il est immédiat, au vu des égalités ci-dessus, que φ est trivial si et seulement si l'on a $h_1 f_1 h_2 f_2 = h_1 h_2 f_1 f_2$ pour tout $(h_1, h_2, f_1, f_2) \in H^2 \times F^2$; si c'est le cas, on dit que G est le produit *direct* interne de H et F.

6.1.4. *Un critère utile.* — Soient n et m deux entiers premiers entre eux, soit G un groupe de cardinal nm, soit H un sous-groupe distingué de G de cardinal n, et soit F un sous-groupe de G de cardinal m. *On est alors dans la situation décrite ci-dessus.* En effet, $F \cap H$ est à la fois un sous-groupe de H et un sous-groupe de F, donc son cardinal est un diviseur commun à n et m ; il est en conséquence égale à 1, ce qui signifie que $H \cap F = \{e\}$. De plus, HF est un sous-groupe de G qui contient H et F ; son cardinal est de ce fait multiple de m et de n, et partant multiple de nm puisque $\mathrm{PGCD}(n, m) = 1$. Comme le cardinal de G est égal à mn, on a nécessairement $G = HF$.

6.2. Le produit semi-direct externe. — Soient H et F deux groupes, et soit φ un morphisme de groupes de F dans Aut H. On se propose de construire un groupe G contenant F (resp. H) comme sous-groupe (resp. comme sous-groupe distingué), tel que $H \cap F = \{e\}$, que $HF = G$, et que pour tout $(h, f) \in H \times F$ l'on ait $fh = \varphi(f)(h)f$, ou $\varphi(f)(h) = fhf^{-1}$ si l'on préfère ; on diposera alors de la formule $(*)$ vue au 6.1.3.2.

Remarque 6.2.1. — L'expression « on se propose de construire un groupe G *contenant* H et F... » constitue un abus de langage. Ce que l'on va en réalité chercher à fabriquer, c'est un groupe G muni de deux morphismes *injectifs* $i : H \hookrightarrow G$ et $j : F \hookrightarrow G$, tels que les propriétés énoncées ci-dessus soient satisfaites *modulo les identifications respectives de* H *à* $i(H)$ *et de* F *à* $j(F)$. Cela signifie précisément que les assertions suivantes devront être vérifiées (le lecteur conviendra aisément avec nous que le caractère rebutant de la seconde justifie l'abus commis ci-dessus) :

◇ $i(H)$ est distingué, $i(H) \cap j(F) = \{e\}$ et $G = i(H)j(F)$;

◇ pour tout $(h, f) \in H \times F$, on a $j(f)i(h) = i(\varphi(f)(h))j(f)$.

6.2.2. *La construction.* — Pour construire G (ainsi que les injections i et j), on se contente essentiellement de décalquer la formule $(*)$ de 6.1.3.2. On définit donc G comme étant l'*ensemble* produit $H \times F$, et on le munit d'une loi interne, notée multiplicativement, en posant

$$(h_1, f_1)(h_2, f_2) = (h_1\varphi(f_1)(h_2), f_1 f_2)$$

pour tout $(h_1, h_2, f_1, f_2) \in H^2 \times F^2$.

On vérifie (c'est un tout petit peu fastidieux, mais sans aucune difficulté) que l'on a bien ainsi construit un groupe ; son élément neutre est (e, e). Soit (h_1, h_2, f_1, f_2) appartenant à $H^2 \times F^2$. Il est immédiat que $(h_1, e)(h_2, e) = (h_1 h_2, e)$ et que $(e, f_1)(e, f_2) = (e, f_1 f_2)$. Ceci montre que l'application i (resp. j) qui envoie un élément h de H (resp. un élément f de F) sur (h, e) (resp. (e, f)) est un morphisme de groupes, trivialement injectif. Notons que $i(H)$ (resp. $j(F)$ est l'ensemble des éléments de G de la forme (h, e) (resp. (e, f)) avec $h \in H$ (resp. $f \in F$).

Il reste à s'assurer que les conditions énoncées à la remarque 6.2.1 sont remplies. Il résulte immédiatement de la définition de la loi de groupe définie sur $H \times F$ que $(h, f) \mapsto f$ est un morphisme de G dans F. Son noyau est visiblement égal à $i(H)$, lequel est par conséquent distingué ; l'égalité $i(H) \cap j(F) = \{e\}$ est triviale ; si h (resp. f) appartient à H (resp. F), alors $(h, f) = (h, e)(e, f)$ dans G, et l'on a donc bien $G = i(H)j(F)$; de plus,

$$j(f)i(h) = (e, f)(h, e) = (\varphi(f)(h), f) = (\varphi(f)(h), e)(e, f) = i(\varphi(f)(h))j(f).$$

Le groupe G (muni des injections i et j) jouit donc de toutes les propriétés requises. On l'appelle *produit semi-direct (externe)* de H et F relativement à φ, et on le note $H \rtimes_\varphi F$. Si φ est trivial, on retrouve le groupe produit habituel, que l'on note simplement $H \times F$.

Remarque 6.2.3. — Dans un certain nombre de cas, on a réellement intérêt, pour des raisons de confort psychologique, à penser au produit semi-direct en les termes un peu abusifs utilisés au début de ce paragraphe, c'est-à-dire à oublier la construction et les injections i et j, et à le voir comme un groupe contenant H et F, dans lequel chaque élément a une unique écriture sous la formes hf, et dont la loi est décrite par la formule $(*)$ de 6.1.3.2 (et se retrouve à partir de l'égalité $fh = \varphi(f)(h)f$, un peu plus simple à retenir) ; toutefois dans d'autres circonstances, il peut être recommandé de travailler avec les couples (h, f) ; c'est par exemple plus prudent si $F = H$, situation dans laquelle il faut être particulièrement soigneux pour éviter toute confusion.

Remarque 6.2.4. — Mentionnons une différence fondamentale entre le produit semi-direct interne et le produit semi-direct externe.

◇ Dans le cas interne, le morphisme $\varphi : F \to \mathrm{Aut}\, H$ est imposé par la situation, il est égal à $f \mapsto (h \mapsto fhf^{-1})$.

◇ Dans le cas externe, le morphisme φ est donné *a priori*, et l'on construit G de sorte que $\varphi(f) = h \mapsto fhf^{-1}$ pour tout $f \in F$; d'une certaine manière, on *force* $\varphi(f)$ à être la restriction à H de l'automorphisme intérieur de G associé à f.

6.2.5. *Liens avec le produit semi-direct interne.* — Donnons-nous comme au 6.2, un groupe G, et deux sous-groupes H et F de G tels que $H \triangleleft G$, que H et F engendrent G, et que $H \cap F = \{e\}$. Soit φ le morphisme de F dans Aut H qui envoie f sur $h \mapsto fhf^{-1}$. Les résultats de 6.1.3.2 peuvent alors, à la lumière de ce qui précède, se récrire comme suit : *l'application* $(h, f) \mapsto hf$ *établit un isomorphisme entre* $H \rtimes_\varphi F$ *et* G *.*

6.2.6. *Produits semi-direct et isomorphismes.* — Mentionnons un fait que nous utiliserons implicitement de façon répétée tout au long de ce texte. Soient G et H deux groupes, et soit φ un morphisme de H dans Aut G. Soient G' et H' deux groupes, et soient $i : G \simeq G'$ et $j : H \simeq H'$ deux isomorphismes. Soit φ' l'application $h \mapsto i \circ \varphi(j^{-1}(h)) \circ i^{-1}$ de H' dans Aut G'. C'est un morphisme de groupes, caractérisé par le fait que $\varphi'(j(h))(i(g)) = i(\varphi(h)(g))$ pour tout $(h, g) \in H \times G$; en termes un peu plus imagés, φ' est le morphisme correspondant à φ si l'on identifie G à G' *via* i et H à H' *via* j. Nous laissons le lecteur vérifier que l'application $(h, g) \mapsto (i(h), j(g))$ établit un isomorphisme entre $G \rtimes_\varphi H$ et $G' \rtimes_{\varphi'} H'$.

6.2.7. *Produits semi-directs de deux groupes cycliques.* — Soient n et m deux entiers strictement positifs. Nous nous proposons de décrire tous les produits semi-directs de la forme $\mathbf{Z}/n\mathbf{Z} \rtimes_\varphi \mathbf{Z}/m\mathbf{Z}$. Pour tout $a \in \mathbf{Z}/n\mathbf{Z}$ on notera h_a l'endomorphisme $x \mapsto ax$ de $\mathbf{Z}/n\mathbf{Z}$.

6.2.7.1. — Rappelons (3.5.8) que $a \mapsto h_a$ induit un isomorphisme de groupes de $(\mathbf{Z}/n\mathbf{Z})^\times$ sur Aut $\mathbf{Z}/n\mathbf{Z}$. Par ailleurs, il résulte de 3.5.5 que se donner un morphisme de groupes φ de $\mathbf{Z}/m\mathbf{Z}$ dans Aut $\mathbf{Z}/n\mathbf{Z}$ revient à se donner un élément u de m-torsion de Aut $\mathbf{Z}/n\mathbf{Z}$, le lien entre les deux étant le suivant : on a $u = \varphi(\overline{1})$; on a $\varphi(\overline{r}) = u^r$ pour tout $r \in \mathbf{Z}$.

En combinant les deux remarques qui précèdent, on voit que se donner un morphisme de groupes φ de $\mathbf{Z}/m\mathbf{Z}$ dans Aut $\mathbf{Z}/n\mathbf{Z}$ revient à se donner un élément a de m-torsion de $(\mathbf{Z}/n\mathbf{Z})^\times$, le lien entre les deux étant le suivant : on a $a = \varphi(\underbrace{\overline{1}}_{\mathrm{mod}\ m})(\underbrace{\overline{1}}_{\mathrm{mod}\ n})$; on a $\varphi(\underbrace{\overline{r}}_{\mathrm{mod}\ m})(x) = a^r x$ pour tout $r \in \mathbf{Z}$ et tout $x \in \mathbf{Z}/n\mathbf{Z}$.

Pour tout élément a d'ordre m de $(\mathbf{Z}/n\mathbf{Z})^\times$, on notera φ_a le morphisme de $\mathbf{Z}/m\mathbf{Z}$ dans Aut $\mathbf{Z}/n\mathbf{Z}$ qui correspond à a.

6.2.7.2. — Par ce qui précède, les produits semi-directs de la forme $\mathbf{Z}/n\mathbf{Z} \rtimes_\varphi \mathbf{Z}/m\mathbf{Z}$ sont exactement les $\mathbf{Z}/n\mathbf{Z} \rtimes_{\varphi_a} \mathbf{Z}/m\mathbf{Z}$ pour a parcourant l'ensemble des éléments de m-torsion de $(\mathbf{Z}/n\mathbf{Z})^\times$. Fixons un tel a. Le groupe $\mathbf{Z}/n\mathbf{Z} \rtimes_{\varphi_a} \mathbf{Z}/m\mathbf{Z}$ est alors l'ensemble $\mathbf{Z}/n\mathbf{Z} \times \mathbf{Z}/m\mathbf{Z}$, et sa loi interne est donnée par la formule

$$(\overline{u}, \overline{r}) \cdot (\overline{v}, \overline{s}) = (\overline{u} + a^r \overline{v}, \overline{r} + \overline{s}).$$

Si $a = \overline{1}$ on retrouve bien entendu le produit direct $\mathbf{Z}/n\mathbf{Z} \times \mathbf{Z}/m\mathbf{Z}$ qui est abélien. Si $a \neq \overline{1}$ alors $\mathbf{Z}/n\mathbf{Z} \rtimes_{\varphi_a} \mathbf{Z}/m\mathbf{Z}$ n'est pas abélien puisque φ_a est alors non trivial, *cf.*

la discussion au 6.1.3.2 ; on peut aussi directement constater que sous cette hypothèse

$$(\overline{1},\overline{0})\cdot(\overline{0},\overline{1})=(\overline{1},\overline{1})\neq(a,\overline{1})=(\overline{0},\overline{1})\cdot(\overline{1},\overline{0}).$$

En particulier, $\mathbf{Z}/n\mathbf{Z}\rtimes_{\varphi_a}\mathbf{Z}/m\mathbf{Z}$ n'est pas isomorphe à $\mathbf{Z}/n\mathbf{Z}\rtimes_{\varphi_{\overline{1}}}\mathbf{Z}/m\mathbf{Z}$ dès que $a\neq\overline{1}$. Mais précisons que $\mathbf{Z}/n\mathbf{Z}\rtimes_{\varphi_a}\mathbf{Z}/m\mathbf{Z}$ et $\mathbf{Z}/n\mathbf{Z}\rtimes_{\varphi_b}\mathbf{Z}/m\mathbf{Z}$ peuvent très bien être isomorphes sans que a soit égal à b, *cf.* l'exemple 6.2.8 ci-dessous.

6.2.7.3. — L'élément $-\overline{1}$ de $(\mathbf{Z}/n\mathbf{Z})^\times$ est de 2-torsion ; il donne donc lieu à un produit semi-direct $(\mathbf{Z}/n\mathbf{Z})\rtimes_{\varphi_{(-\overline{1})}}\mathbf{Z}/2\mathbf{Z}$, qui est souvent noté D_n et est appelé le *groupe diédral de rang n*. Son cardinal est $2n$. Son ensemble sous-jacent est $\mathbf{Z}/n\mathbf{Z}\times\mathbf{Z}/2\mathbf{Z}$ et sa loi interne est donnée par la formule

$$(\overline{u},\overline{r})\cdot(\overline{v},\overline{s})=(\overline{u}+(\overline{-1})^r\overline{v},\overline{r}+\overline{s}).$$

Si $n\geqslant 3$ alors $-\overline{1}\neq\overline{1}$ dans $\mathbf{Z}/n\mathbf{Z}$ et D_n n'est donc pas abélien, *cf.* 6.2.7.2 ci-dessus.

Par contre $\overline{1}=-\overline{1}$ dans $\mathbf{Z}/2\mathbf{Z}$; le groupe diédral D_2 est par conséquent le produit *direct* $\mathbf{Z}/2\mathbf{Z}\times\mathbf{Z}/2\mathbf{Z}$.

6.2.7.4. — Supposons que m est premier au cardinal $\Phi(n)$ de $(\mathbf{Z}/n\mathbf{Z})^\times$. Dans ce cas tout élément de m-torsion de $(\mathbf{Z}/n\mathbf{Z})^\times$ est trivial ; le seul produit semi-direct de la forme $\mathbf{Z}/n\mathbf{Z}\rtimes_\varphi\mathbf{Z}/m\mathbf{Z}$ est donc le produit *direct* $\mathbf{Z}/n\mathbf{Z}\times\mathbf{Z}/m\mathbf{Z}$.

C'est par exemple le cas si $n=3$ et $m=5$, car $\Phi(3)=2$.

Exemple 6.2.8. — Nous allons décrire tous les produits semi-directs de la forme $\mathbf{Z}/7\mathbf{Z}\rtimes_\varphi\mathbf{Z}/3\mathbf{Z}$. Le groupe $(\mathbf{Z}/7\mathbf{Z})^\times$ est égal à $\{-\overline{3},-\overline{2},-\overline{1},\overline{1},\overline{2},\overline{3}\}$; il est cyclique (corollaire 3.8.6). On a $\overline{2}\neq\overline{1}$ et $\overline{2}^3=\overline{8}=\overline{1}$. En conséquence, $\overline{2}$ est d'ordre 3, et $\langle\overline{2}\rangle=\{\overline{1},\overline{2},\overline{4}\}$ est donc l'unique sous-groupe de cardinal 3 de $(\mathbf{Z}/7\mathbf{Z})^\times$, qui est aussi le groupe des éléments de 3-torsion de $(\mathbf{Z}/7\mathbf{Z})^\times$. Les produits semi-directs cherchés sont en conséquence les suivants.

(1) Le produit $\mathbf{Z}/7\mathbf{Z}\rtimes_{\varphi_{\overline{1}}}\mathbf{Z}/3\mathbf{Z}=\mathbf{Z}/7\mathbf{Z}\times\mathbf{Z}/3\mathbf{Z}\simeq\mathbf{Z}/21\mathbf{Z}$.

(2) Le produit $\mathbf{Z}/7\mathbf{Z}\rtimes_{\varphi_{\overline{2}}}\mathbf{Z}/3\mathbf{Z}$. Son ensemble sous-jacent est $\mathbf{Z}/7\mathbf{Z}\times\mathbf{Z}/3\mathbf{Z}$ et sa loi interne est donnée par la formule

$$(\overline{u},\overline{r})\cdot(\overline{v},\overline{s})=(\overline{u}+\overline{2}^r\overline{v},\overline{r}+\overline{s}).$$

(3) Le produit $\mathbf{Z}/7\mathbf{Z}\rtimes_{\varphi_{\overline{4}}}\mathbf{Z}/3\mathbf{Z}$. Son ensemble sous-jacent est $\mathbf{Z}/7\mathbf{Z}\times\mathbf{Z}/3\mathbf{Z}$ et sa loi interne est donnée par la formule

$$(\overline{u},\overline{r})\cdot(\overline{v},\overline{s})=(\overline{u}+\overline{4}^r\overline{v},\overline{r}+\overline{s}).$$

On a vu en 6.2.7.2 que $\mathbf{Z}/7\mathbf{Z}\rtimes_{\varphi_{\overline{2}}}\mathbf{Z}/3\mathbf{Z}$ et $\mathbf{Z}/7\mathbf{Z}\rtimes_{\varphi_{\overline{4}}}\mathbf{Z}/3\mathbf{Z}$ sont tous deux non abéliens, et en particulier tous deux non isomorphes au produit direct $\mathbf{Z}/7\mathbf{Z}\times\mathbf{Z}/3\mathbf{Z}$. Mais le lecteur est invité à vérifier que $(\overline{u},\overline{r})\mapsto(\overline{u},2\overline{r})$ définit un isomorphisme de groupes de $\mathbf{Z}/7\mathbf{Z}\rtimes_{\varphi_{\overline{2}}}\mathbf{Z}/3\mathbf{Z}$ sur $\mathbf{Z}/7\mathbf{Z}\rtimes_{\varphi_{\overline{4}}}\mathbf{Z}/3\mathbf{Z}$, de réciproque donnée par la même formule.

6.2.9. *Compléments : interprétation géométrique du groupe diédral.* — Identifions $\mathbf{C}$ à $\mathbf{R}^2$, fixons un entier $n \geqslant 2$ et notons Γ le sous-groupe de $\mathbf{C}^\times$ formé des racines n-ièmes de l'unité dans $\mathbf{C}$ (notez que si $n \geqslant 3$ alors Γ est l'ensemble des sommets d'un polygone régulier à n côtés) ; le morphisme de groupes $k \mapsto e^{2ik\pi/n}$ de $\mathbf{Z}$ dans Γ passe au quotient modulo n et induit un isomorphisme $\mathbf{Z}/n\mathbf{Z} \simeq \Gamma$. Soit G le groupe des isométries affines g de $\mathbf{R}^2$ qui stabilisent Γ, c'est-à-dire telles que $g(\Gamma) = \Gamma$, ou encore telles que $g(\Gamma) \subset \Gamma$ (puisque Γ est fini, *cf.* 4.3.4.2).

6.2.9.1. — Comme l'origine O est l'isobarycentre de Γ, toutes les isométries de G fixent O ; autrement dit, elles sont linéaires. Une isométrie $\mathbf{R}$-linéaire de $\mathbf{C}$ est de la forme $z \mapsto uz$ ou $z \mapsto u\overline{z}$, avec $|u| = 1$; elle est directe dans le premier cas, indirecte dans le second, et u est uniquement déterminé dans les deux cas (c'est l'image de 1). Une isométrie de la forme $z \mapsto uz$ ou $z \mapsto u\overline{z}$ fixe Γ si et seulement si $u \in \Gamma$: c'est en effet clairement suffisant, et on voit que c'est nécessaire en appliquant l'isométrie considérée à $z = 1$.

6.2.9.2. — Soit G^+ le sous-groupe de G formé des isométries directes. C'est un sous-groupe distingué de G (en tant que noyau du déterminant). Par ce qui précède, l'application $u \mapsto [z \mapsto uz]$ établit un isomorphisme de Γ sur G^+. Par composition, on en déduit un isomorphisme $\mathbf{Z}/n\mathbf{Z} \simeq G^+$.

6.2.9.3. — Soit c la conjugaison complexe. Comme elle est d'ordre 2, le sous-groupe $\langle c\rangle$ de G est égal à $\{\mathrm{Id}, c\}$, et l'application $\overline{0} \mapsto \mathrm{Id}, \overline{1} \mapsto c$ est un isomorphisme de groupes de $\mathbf{Z}/2\mathbf{Z}$ sur $\langle c\rangle$. L'intersection $\langle c\rangle \cap G^+$ est égale à $\{\mathrm{Id}\}$, et la description explicite des éléments de G assure que $G = G^+ \cdot \langle c\rangle$. Par conséquent, G s'identifie à un produit semi-direct de G^+ par $\langle c\rangle$ (6.2.5) ; il reste à déterminer l'action de $\langle c\rangle$ sur G^+.

L'action de Id sur G^+ est triviale ; déterminons maintenant l'action de c. Soit $g \in G^+$; il existe un unique élément $u \in \Gamma$ tel que $g(z) = uz$ pour tout $z \in \mathbf{C}$. Soit $z \in \mathbf{C}$. On a

$$cgc^{-1}(z) = cg(\overline{z}) = c(u\overline{z}) = \overline{u\overline{z}} = \overline{u}z = u^{-1}z$$

car u est de module 1. Ainsi, $cg^{-1}c = g^{-1}$.

6.2.9.4. — En conséquence, G s'identifie à $G^+ \rtimes_\psi \langle c\rangle$, où $\psi\colon \langle c\rangle \to \mathrm{Aut}\, G^+$ est le morphisme

$$\mathrm{Id} \mapsto \mathrm{Id}, \quad c \mapsto (g \mapsto g^{-1}).$$

À l'aide des isomorphismes $\mathbf{Z}/n\mathbf{Z} \simeq \Gamma$ et $\mathbf{Z}/2\mathbf{Z} \simeq \langle c\rangle$ évoqués ci-dessus (6.2.9.2, 6.2.9.3), on en déduit que G est isomorphe à $\mathbf{Z}/n\mathbf{Z} \rtimes_\varphi \mathbf{Z}/2\mathbf{Z}$, où φ envoie $\overline{0}$ sur Id et $\overline{1}$ sur $\overline{a} \mapsto -\overline{a}$; autrement dit, G est isomorphe au groupe D_n décrit au 6.2.7.3.

6.3. Suites exactes. — Nous allons maintenant introduire une notion élémentaire mais extrêmement importante, celle de *suite exacte* ; après quelques considérations générales, nous verrons qu'elle est liée au produit semi-direct.

Définition 6.3.1. — Soit I un intervalle de $\mathbf{Z}$ (dont les bornes peuvent être finies ou infinies) et soit

$$\dots \xrightarrow{u_{i-2}} G_{i-1} \xrightarrow{u_{i-1}} G_i \xrightarrow{u_i} G_{i+1} \xrightarrow{u_{i+1}} \dots$$

une suite de morphismes de groupes, l'indice i parcourant I.

Soit i un élément de I tel que $i-1$ et $i+1$ appartiennent à I. On dit que la suite ci-dessus est *exacte en* G_i si $\mathrm{Im}(u_{i-1}) = \mathrm{Ker}(u_i)$. On dit qu'elle est *exacte* si elle est exacte en G_i pour tout $i \in I$ tel que $i-1$ et $i+1$ appartiennent à I.

Exemple 6.3.2. — Donnons-nous trois groupes F, G et H, ainsi que deux morphismes $u : H \to G$ et $p : G \to F$. La suite

$$1 \longrightarrow H \xrightarrow{u} G \xrightarrow{p} F \longrightarrow 1$$

est exacte en H si et seulement si u est injective ; elle est exacte en G si et seulement si $\mathrm{Im}(u) = \mathrm{Ker}(p)$, et telle est exacte en F si et seulement si p est surjective.

Cette suite est donc exacte si et seulement si u est injective, p est surjective, et $\mathrm{Im}(u) = \mathrm{Ker}(p)$.

La plupart des suites exactes que nous rencontrerons dans ce cours seront de ce type.

Remarque 6.3.3. — Si

$$1 \longrightarrow H \xrightarrow{u} G \xrightarrow{p} F \longrightarrow 1$$

est une suite exacte, alors $\mathrm{Im}(u)$, qui coïncide avec le noyau de p, est un sous-groupe *distingué* de G.

Exemple 6.3.4. — Soit G un groupe, soit H un sous-groupe de G, soit i l'inclusion de H dans G et soit $\pi : G \to G/H$ le morphisme quotient. La suite

$$1 \longrightarrow H \xrightarrow{i} G \xrightarrow{\pi} G/H \longrightarrow 1$$

est exacte.

Exemple 6.3.5. — Soient H et F deux groupes et soit $\varphi : F \to \mathrm{Aut}\, H$ un morphisme. Soit i le morphisme $h \mapsto (h, e)$ de H dans $H \rtimes_\varphi F$ et soit q le morphisme $(h, f) \mapsto f$ de $H \rtimes_\varphi F$ dans F ; la suite

$$1 \longrightarrow H \xrightarrow{i} H \rtimes_\varphi F \xrightarrow{q} F \longrightarrow 1$$

est exacte.

Définition 6.3.6. — On dit qu'une suite exacte

$$1 \longrightarrow H \xrightarrow{u} G \xrightarrow{p} F \longrightarrow 1$$

est *scindée* si le morphisme p possède une *section*, c'est-à-dire un morphisme s de F dans G tel que $p \circ s = \mathrm{Id}_F$.

Exemple 6.3.7. — La suite exacte de l'exemple 6.3.5 est scindée ; en effet, en reprenant les notations de *loc. cit.*, on voit facilement que $f \mapsto (e, f)$ est une section de $q : (h, f) \mapsto f$.

6.3.8. *Le cas où le groupe de droite est monogène.* — On se donne une suite exacte

$$1 \longrightarrow H \xrightarrow{u} G \xrightarrow{p} F \longrightarrow 1$$

et l'on suppose que F est monogène, c'est-à-dire engendré par un élément f. Dans ce cas, tout morphisme de groupes défini de source F est entièrement déterminé par sa valeur en f ; en particulier, si s est un morphisme de groupes de F dans G, alors $p \circ s = \mathrm{Id}_F \iff p(s(f)) = f$; autrement dit, s est une section de p si et seulement si $s(f)$ est un antécédent de f pour p.

6.3.8.1. — Supposons que f est d'ordre infini, auquel cas $n \mapsto f^n$ établit un isomorphisme $\mathbf{Z} \simeq F$. Dans cette situation, $s \mapsto s(f)$ définit une bijection entre $\mathrm{Hom}(F, G)$ et G ; compte-tenu de ce qui précède, on en déduit que $s \mapsto s(f)$ établit une bijection entre l'ensemble des sections de p et l'ensemble des antécédents de f pour p. *Comme f a au moins un antécédent pour p, l'ensemble des sections de p est non vide, et la suite exacte étudiée est donc scindée.*

Remarquons que si g est un antécédent de f pour p, la section s qui lui correspond est très simple à décrire : tout élément de F a une unique écriture sous la forme f^n, avec n dans $\mathbf{Z}$; son image par s est alors précisément g^n.

6.3.8.2. — Supposons que f est d'ordre fini m, auquel cas $n \mapsto f^n$ induit un isomorphisme $(\mathbf{Z}/m\mathbf{Z}) \simeq F$. Dans cette situation, $s \mapsto s(f)$ définit une bijection entre $\mathrm{Hom}(F, G)$ et l'ensemble des élements g de G tels que $g^m = e$; compte-tenu de ce qui précède, on en déduit que $s \mapsto s(f)$ définit une bijection entre l'ensemble des sections de p et l'ensemble des antécédents g de f pour p tels que $g^m = e$. *Contrairement à ce qui se produit lorsque f est d'ordre infini, cet ensemble peut très bien être vide, comme l'atteste l'exemple 6.3.9 ci-dessous.*

Remarquons que si g est un antécédent de f pour p tel que $g^m = e$, la section s qui lui correspond est très simple à décrire : tout élément de F a une écriture sous la forme f^n, où n appartient à $\mathbf{Z}$ *et est uniquement déterminé modulo m* ; son image par s est alors précisément g^n, qui ne dépend bien, en vertu de l'hypothèse faite sur g, que de la classe de n modulo m.

Exemple 6.3.9. — Nous allons exhiber une suite exacte dont le terme de droite est monogène et qui n'est pas scindée. Pour tout entier $d > 0$, on note μ_d le sous-groupe de $\mathbf{C}^*$ formé des racines d-ièmes de l'unité ; notons que μ_d est cyclique de cardinal d, engendré par $e^{2i\pi/d}$.

Fixons un entier $m > 0$. Soit u l'inclusion de μ_m dans μ_{m^2}. Soit p le morphisme de μ_{m^2} dans μ_m qui envoie un élément z sur z^m. Son noyau est par définition égal au sous-groupe μ_m de μ_{m^2}, et il est surjectif : tout élément de μ_m a une écriture sous la forme $e^{2ik\pi/m}$, où $k \in \mathbf{Z}$, et est ainsi l'image par p de l'élément $e^{2ik\pi/m^2}$ de μ_{m^2}. La suite

$$1 \longrightarrow \mu_m \xrightarrow{u} \mu_{m^2} \xrightarrow{p} \mu_m \longrightarrow 1$$

est donc exacte. *Si* $m > 1$ *elle n'est pas scindée.* Il suffit en effet, d'après ce qui a été vu plus haut, de vérifier que l'ensemble des $z \in \mu_{m^2}$ tels que $p(\mu) = e^{2i\pi/m}$ et tels que $z^m = 1$ est vide. Or si $z \in \mu_{m^2}$ vérifie $z^m = 1$, alors

$$p(\mu) = z^m = 1 \neq e^{2i\pi/m},$$

ce dernier fait résultant de l'hypothèse $m > 1$.

6.4. Sections, sous-groupes d'un certain type et structure de produit semi-direct. — On fixe une suite exacte

$$1 \longrightarrow H \xrightarrow{u} G \xrightarrow{p} F \longrightarrow 1$$

et l'on se propose de montrer que se donner une section de p revient à se donner une écriture de G comme produit semi-direct de H par F «compatible avec la suite exacte ci-dessus».

Lemme 6.4.1. — *Soit s une section de p. On a les égalités $u(H) \cap s(F) = \{e\}$ et $G = u(H)s(F)$. L'application s induit un isomorphisme $F \simeq s(F)$ dont la réciproque est $p_{|s(F)}$.*

Démonstration. — Soit g un élément de $s(F) \cap u(H)$. Comme $u(H) = \mathrm{Ker} p$, on a $p(g) = e$. Comme g appartient à $s(F)$, il s'écrit $s(f)$ pour un certain f dans F. On a alors $f = p(s(f)) = p(g) = e$, et donc $g = s(f) = e$.

Soit maintenant g un élément de G. On a

$$p(gs(p(g))^{-1}) = p(g)p(s(p(g)))^{-1} = p(g)p(g)^{-1} = e$$

(l'avant-dernière égalité provient du fait que $p \circ s = \mathrm{Id}_F$). On peut dès lors écrire $gs(p(g))^{-1} \in \mathrm{Ker} p = \mathrm{Im} u$. Il existe donc $h \in H$ tel que $g = u(h)s(p(g))$; en conclusion, $g \in u(H)s(F)$.

Soit $f \in F$. Si $s(f) = e$, alors $f = p(s(f)) = e$ et s est donc injective. Elle induit en conséquence un isomorphisme $F \simeq s(F)$. On a $p_{|s(F)} \circ s = p \circ s = \mathrm{Id}_F$, ce qui montre que $p_{|s(F)}$ est la réciproque de l'isomorphisme $F \simeq s(F)$ défini par s. □

Lemme 6.4.2. — *Soit Γ un sous-groupe de G tel que $\Gamma \cap u(H) = \{e\}$ et tel que $G = u(H)\Gamma$. La restriction de p à Γ induit un isomorphisme $\Gamma \simeq F$ dont la réciproque, vue comme morphisme de F dans G, est une section de p.*

Démonstration. — Soit $\gamma \in \Gamma$ tel que $p(\gamma) = e$. On a alors $\gamma \in \mathrm{Ker} p = \mathrm{Im} u$; comme $\Gamma \cap u(H) = \{e\}$, on a $\gamma = e$ et $p_{|\Gamma}$ est injectif. Soit $f \in F$. Comme p est surjectif, il existe g dans G tel que $f = p(g)$. Puisque $G = u(H)\Gamma$, on peut écrire $g = u(h)\gamma$ avec $h \in H$ et $\gamma \in \Gamma$; dès lors $f = p(g) = p(u(h)))p(\gamma) = p(\gamma)$ puisque $u(H) = \mathrm{Ker} p$. En conséquence, $p_{|\Gamma}$ est surjectif, et finalement bijectif.

Soit s la réciproque de $p_{|\Gamma}$, qui va de F dans Γ et que l'on voit comme étant à valeurs dans G. Si f appartient à F, alors $p(s(f)) = p_{|\Gamma}(s(f)) = f$. On a bien démontré que $p \circ s = \mathrm{Id}_F$. □

Soit $\mathscr{S}$ l'ensemble des sections de p, et soit $\mathscr{G}$ l'ensemble des sous-groupes Γ de G tels que $\Gamma \cap u(H) = \{e\}$ et tels que $G = u(H)\Gamma$. Il résulte des deux lemmes ci-dessus que si $s \in \mathscr{S}$, alors $s(F) \in \mathscr{G}$, et que si $\Gamma \in \mathscr{G}$, alors $(p_{|\Gamma})^{-1} \in \mathscr{S}$. On a ainsi construit une application Φ de $\mathscr{S}$ dans $\mathscr{G}$ et une seconde application Ψ de $\mathscr{G}$ dans $\mathscr{S}$.

Proposition 6.4.3. — *Les applications Φ et Ψ sont deux bijections réciproques l'une de l'autre.*

Démonstration. — Soit s une section de p. Le groupe $\Phi(s)$ n'est autre que $s(F)$; la section $\Psi(\Phi(s))$ est la réciproque de $p_{|\Phi(s)} = p_{|s(F)}$; en vertu du premier des deux lemmes ci-dessus, c'est précisément s.

Soit $\Gamma \in \mathscr{G}$. La section $\Psi(\Gamma)$ est égale à $p_{|\Gamma}^{-1}$; comme c'est un isomorphisme de F sur Γ, son image est précisément Γ. Or cette image est par définition le groupe $\Phi(\Psi(\Gamma))$, ce qui achève la démonstration. ☐

6.4.4. — Soit (φ, ι) un couple où φ est un morphisme de F dans Aut H et où ι est un isomorphisme entre $H \rtimes_\varphi F$ et G tel que le diagramme

$$\begin{tikzcd}
& & G \arrow[dr, "p"] & & \\
1 \arrow[r] & H \arrow[ur, "u"] \arrow[dr, "i"'] & & F \arrow[r] & 1 \\
& & H \rtimes_\varphi F \arrow[uu, "\iota"] \arrow[ur, "q"'] & &
\end{tikzcd}$$

commute, où i et q désignent respectivement les applications $h \mapsto (h, e)$ et $(h, f) \mapsto f$. Cela signifie que $\iota(h, e) = u(h)$ pour tout $h \in H$ et que $p(\iota(h, f)) = f$ pour tout $(h, f) \in H \rtimes_\varphi F$.

Soit σ la section $f \mapsto (e, f)$ du morphisme q. La composée $\iota \circ \sigma$ est une section s de p ; remarquons que le groupe $s(F)$ qui correspond à cette section n'est autre que $\iota(\{e\} \times F)$.

À tout couple (φ, ι) comme ci-dessus on sait ainsi associer une section s de p.

6.4.5. — Réciproquement, soit s une section de p. On va lui associer un couple (φ, ι) comme ci-dessus. D'après le paragraphe précédent, $s(F)$ est un sous-groupe de G tel que $s(F) \cap u(H) = \{e\}$ et tel que $G = s(F)u(H)$. En vertu de 6.2.5, il existe un morphisme $\psi : s(F) \to \operatorname{Aut} u(H)$ tel que $(a, b) \mapsto ab$ établisse un isomorphisme entre $u(H) \rtimes_\psi s(F)$ et G.

Compte-tenu du fait que u (resp. s) induit un isomorphisme entre H (resp. F) et $u(H)$ (resp. $s(F)$), il existe un morphisme φ de F vers Aut H tel que

$$(h, f) \mapsto u(h)s(f)$$

établisse un isomorphisme ι entre $H \rtimes_\varphi F$ et G.

On a en particulier $\iota(h, e) = u(h)$ et $p(\iota(h, f)) = p(s(f)) = f$ pour tout couple $(h, f) \in H \rtimes_\varphi F$, et la condition de commutativité du diagramme est ainsi satisfaite.

Commentaires 6.4.6. — L'isomorphisme ι est défini à partir de s par une formule simple, puisqu'on vient de voir qu'il envoie tout couple (h, f) sur $u(h)s(f)$. Le morphisme φ n'admet par contre pas de description aussi agréable : il est seulement *caractérisé* par l'égalité $s(f)u(h) = u(\varphi(f)(h))s(f)$ ou, si l'on préfère,

$$u(\varphi(f)(h)) = s(f)u(h)s(f)^{-1},$$

qui est valable pour tout (h, f).

Lemme 6.4.7. — *Les deux constructions $(\varphi, \iota) \mapsto s$ et $s \mapsto (\varphi, \iota)$ que nous venons de détailler sont réciproques l'une de l'autre.*

Démonstration. — Partons d'un couple (φ, ι). On lui associe la section s qui envoie un élément f de F sur $\iota(e, f)$; à cette section est associé à son tour un couple (ψ, η). Notre but est de montrer que $(\psi, \eta) = (\varphi, \iota)$. Par construction de (ψ, η), on a

$$\eta(h, e) = u(h) = \iota(h, e)$$

pour tout $h \in H$ et $\eta(e, f) = s(f) = \iota(e, f)$ pour tout $f \in F$; pour tout couple (h, f) on a donc $\eta(h, f) = \eta(h, e)\eta(e, f) = \iota(h, e)\iota(e, f) = \iota(h, f)$ (on a utilisé le fait que $(h, f) = (h, e)(e, f)$ dans $H \rtimes_\varphi F$ aussi bien que dans $H \rtimes_\psi F$). Les applications *ensemblistes* η et ι (toutes deux définies sur *l'ensemble* $H \times F$) coïncident donc. Par ailleurs, ι (resp. η) est un morphisme de $H \rtimes_\varphi F$ (resp. $H \rtimes_\psi F$) vers G ; on a donc $\iota(f, e)\iota(e, h) = \iota(\varphi(f)(h), f)$ et $\eta(f, e)\eta(e, h) = \eta(\psi(f)(h), f)$ pour tout $h \in H$ et tout $f \in F$. Compte-tenu du fait que $\eta(h, f) = \iota(h, f)$ pour tout (h, f), et que ι et η sont injectifs, on a $\psi(f)(h) = \varphi(f)(h)$ pour tout (h, f). En conséquence, $\psi = \varphi$; comme $\eta = \iota$, on a finalement $(\psi, \eta) = (\varphi, \iota)$.

Réciproquement, partons maintenant d'une section s. Il lui correspond un couple (φ, ι) ; à ce couple est associé à son tour une section t. Notre but est de montrer que $t = s$. Par construction de t, on a pour tout $f \in F$ l'égalité $t(f) = \iota(e, f)$; mais ce dernier terme est égal, par construction de ι, à $u(e)s(f)$, soit à $s(f)$. En conséquence, $t = s$. □

6.4.8. *Récapitulation.* — En vertu de la proposition 6.4.3 et du lemme 6.4.7 il revient au même de se donner :

- ◇ une section s de p ;
- ◇ un sous-groupe Γ de G tel que $\Gamma \cap u(H) = \{e\}$ et tel que $u(H)\Gamma = G$;
- ◇ un couple (φ, ι) formé d'un morphisme $\varphi : F \to \operatorname{Aut} H$ et d'un isomorphisme $\iota : H \rtimes_\psi F \simeq G$ tel que le diagramme

$$\begin{tikzcd}
 & & G \arrow[dr, "p"] & & \\
1 \arrow[r] & H \arrow[ur, "u"] \arrow[dr, "i"'] & & F \arrow[r] & 1 \\
 & & H \rtimes_\varphi F \arrow[uu, "\iota"] \arrow[ur, "q"'] & &
\end{tikzcd}$$

  commute.

Les liens entre ces objets sont les suivants.

(i) Si la section s est donnée, Γ est simplement le groupe $s(F)$; l'isomorphisme ι envoie (h,f) sur $u(h)s(f)$, et φ est caractérisé par le fait que pour tout (h,f), on a l'égalité

$$s(f)u(h) = u(\varphi(f)(h))s(f) \text{ ou encore } u(\varphi(f)(h)) = s(f)u(h)s(f)^{-1}.$$

(ii) Si Γ est donné, alors $p_{|\Gamma}$ induit un isomorphisme $\Gamma \simeq F$, et s est simplement la réciproque de cet isomorphisme, vue comme morphisme de F dans G.

(iii) Si (φ,ι) est donné, alors $s(f) = \iota(e,f)$ pour tout f dans F, et le groupe Γ est simplement $\iota(\{e\} \times F)$.

Remarque 6.4.9. — Cet énoncé qui porte sur l'équivalence entre trois types d'objets ne préjuge en rien de l'existence de tels objets (rappelez-vous qu'on a vu plus haut un exemple de suite exacte non scindée). Il implique par contre que s'il existe un objet (resp. s'il n'existe pas d'objet) de l'un des trois types fixés, alors il existe un objet de chacun des deux autres types (resp. il n'existe aucun objet de l'un ou l'autre des deux autres types).

Remarque 6.4.10. — L'équivalence entre la donnée de s et celle du couple (φ,ι) a pour corollaire le principe suivant, dont l'énoncé est volontairement un peu vague : *toute suite exacte scindée est, à isomorphisme près, de la forme*

$$1 \longrightarrow H \xrightarrow{\ i\ } H \rtimes_{\varphi} F \xrightarrow{\ q\ } F \longrightarrow 1 \ .$$

Exemple 6.4.11. — Soit k un corps, soit $\overrightarrow{E}$ un k-espace vectoriel et soit E un espace affine d'espace directeur $\overrightarrow{E}$. Notons τ l'application qui envoie un vecteur $\overrightarrow{u}$ sur la translation $t_{\overrightarrow{u}}$, et ℓ l'application qui envoie un élément du groupe affine $\mathrm{GA}(E)$ sur l'application linéaire associée, qui appartient à $\mathrm{GL}(\overrightarrow{E})$. La suite

$$1 \longrightarrow (\overrightarrow{E},+) \xrightarrow{\ \tau\ } \mathrm{GA}(E) \xrightarrow{\ \ell\ } \mathrm{GL}(\overrightarrow{E}) \longrightarrow 1$$

est exacte.

Soit $O \in E$. Pour toute application $\overrightarrow{f}$ appartenant à $\mathrm{GL}(\overrightarrow{E})$, on note $\overrightarrow{f}_O$ l'application affine qui fixe O et dont l'application linéaire associée est $\overrightarrow{f}$, à savoir $M \mapsto O + \overrightarrow{f}(\overrightarrow{OM})$. Il est immédiat que $\overrightarrow{f} \mapsto \overrightarrow{f}_O$ constitue une section de ℓ. Le sous-groupe qui lui est associé est l'image de $\mathrm{GL}(\overrightarrow{E})$ sous cette section ; c'est exactement le sous-groupe de $\mathrm{GA}(E)$ formé des bijections affines qui fixent O.

Le couple (φ,ι) correspondant à cette situation se décrit comme suit : ι envoie un couple $(\overrightarrow{u},\overrightarrow{f})$ sur $t_{\overrightarrow{u}} \circ \overrightarrow{f}_O$; le morphisme $\varphi : \mathrm{GL}(\overrightarrow{E}) \to \mathrm{Aut}\ (\overrightarrow{E},+)$ est tel que l'on ait pour tout couple $(\overrightarrow{u},\overrightarrow{f})$ l'égalité $\overrightarrow{f}_O \circ t_{\overrightarrow{u}} = t_{\varphi(\overrightarrow{f})(\overrightarrow{u})} \circ \overrightarrow{f}_O$. En l'appliquant à O, il vient $O + \varphi(\overrightarrow{f})(\overrightarrow{u}) = O + \overrightarrow{f}(\overrightarrow{u})$ et donc $\varphi(\overrightarrow{f})(\overrightarrow{u}) = \overrightarrow{f}(\overrightarrow{u})$. Ceci valant pour tout $(\overrightarrow{u},\overrightarrow{f})$, le morphisme φ est *l'inclusion naturelle* de $\mathrm{GL}(\overrightarrow{E})$ (groupe des bijections de $\overrightarrow{E}$ dans lui-même respectant l'addition et la multiplication par les scalaires) dans $\mathrm{Aut}\ (\overrightarrow{E},+)$ (groupe des bijections de $\overrightarrow{E}$ dans lui-même respectant simplement l'addition).

7. Théorèmes de Sylow

7.1. — Le but de ce chapitre est de démontrer un théorème de structure fondamental sur les groupes finis, et d'en donner quelques applications. Nous allons tout d'abord procéder à quelques préliminaires arithmétiques.

Lemme 7.1.1. — *Soit G un groupe opérant sur lui-même par translations ; on considèrera $\mathscr{P}(G)$ comme muni de l'action induite. Pour toute partie non vide X de G on a $|\mathrm{Stab}(X)| \leqslant |X|$.*

Remarque 7.1.2. — L'énoncé de ce lemme est à prendre au sens de la théorie des cardinaux, mais si vous n'êtes pas à l'aise avec elle vous pouvez supposer G fini – c'est le seul cas qui nous sera utile.

Démonstration du lemme 7.1.1. — Soit X une partie non vide de G et soit H son stabilisateur. Par hypothèse, il existe $x \in X$. Comme H stabilise X on a l'inclusion $Hx \subset X$, si bien que $|Hx| \leqslant |X|$. Et puisque l'application $h \mapsto hx$ est injective (par simplification dans le groupe G), on a $|Hx| = |H|$; par conséquent, $|H| \leqslant |X|$. □

Le lemme suivant est certainement bien connu. Nous le démontrons pour la commodité du lecteur, et aussi pour sa curiosité car nous en allons en donner la preuve classique reposant sur le lemme de Gauß mais également une seconde preuve probablement plus originale.

Lemme 7.1.3. — *Soit p un nombre premier et soit n un entier tel que $1 \leqslant n \leqslant p-1$. L'entier $\binom{p}{n}$ est nul modulo p.*

Démonstration. — Comme nous l'avons annoncé, nous allons donner deux preuves différentes.

Première preuve. On a l'égalité $n!\binom{p}{n} = p(p-1)\dots(p-n+1)$. Le produit de droite comprend $p-n$ termes ; comme $p-n > 0$, il en comprend au moins un et est donc multiple de p. Comme $n < p$, les facteurs premiers de $n!$ sont tous $< p$, et $n!$ est dès lors premier à p. Il en résulte que p divise $\binom{p}{n}$.

Seconde preuve. Soit E l'ensemble des parties de $\mathbf{Z}/p\mathbf{Z}$ de cardinal n. On considère l'action de $\mathbf{Z}/p\mathbf{Z}$ sur E déduite de l'action de $\mathbf{Z}/p\mathbf{Z}$ sur lui-même par translations.

Soit $X \in E$. Comme X est non vide, il résulte du lemme 7.1.1 que $|\mathrm{Stab}(X)| \leqslant |X|$. Ceci entraîne, puisque $n < p$, que $|\mathrm{Stab}(X)| < p$. Comme $|\mathrm{Stab}(X)|$ divise p, il vient $|\mathrm{Stab}(X)| = 1$; l'orbite de X a donc pour cardinal $p/1 = p$. Les orbites de l'action de $\mathbf{Z}/p\mathbf{Z}$ sur E sont ainsi toutes de cardinal p ; si N désigne le nombre d'orbites on a par conséquent $\binom{p}{n} = |E| = Np$. □

7.1.4. *Une conséquence importante.* — Si A est une $\mathbf{F}_p$-algèbre et si a et b sont deux éléments de A on a alors

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

Par une récurrence immédiate sur m, on en déduit que $(a+b)^{p^m} = a^{p^m} + b^{p^m}$ pour tout m.

Lemme 7.1.5. — *Soit p un nombre premier, soit n un entier et soit m un entier premier à p. L'entier $\binom{p^n m}{p^n}$ est premier à p.*

Démonstration. — Nous allons plus précisément montrer que $\binom{p^n m}{p^n}$ est égal à m modulo p. Pour cela, écrivons la formule du binôme

$$(X+Y)^{p^n m} = \sum_{k=0}^{p^n m} \binom{p^n m}{k} X^{p^n m-k} Y^k,$$

qui est valable dans $\mathbf{Z}[X,Y]$, où X et Y sont des indéterminées. L'entier $\binom{p^n m}{p^n}$ est le coefficient de $X^{p^n m - p^n} Y^{p^n} = X^{p^n(m-1)} Y^{p^n}$ dans le polynôme ci-dessus.

Le polynôme $(X+Y)^{p^n m}$ est égal à $((X+Y)^{p^n})^m$; son image dans $\mathbf{F}_p[X,Y]$ est donc égale à $(X^{p^n} + Y^{p^n})^m$ (*cf.* 7.1.4). Ce dernier terme se récrit

$$X^{p^n m} + mX^{p^n(m-1)} Y^{p^n} + \ldots + Y^{p^n m}.$$

Dans ce polynôme à coefficients dans $\mathbf{F}_p[X,Y]$, le coefficient de $X^{p^n(m-1)} Y^{p^n}$ est égal à m ; comme il coïncide avec la réduction de $\binom{p^n m}{p^n}$ modulo p, l'entier $\binom{p^n m}{p^n}$ est égal à m modulo p, comme annoncé. □

7.2. — Nous allons maintenant pouvoir démontrer le théorème de structure fondamental sur les groupes finis que nous avons évoqué plus haut. Il comporte plusieurs assertions, que l'on présente parfois comme les *théorèmes de Sylow.*

Théorème 7.2.1. — *Soit p un nombre premier et soit G un groupe fini. Écrivons $|G| = p^n m$ où $n \in \mathbf{N}$ et où m est premier à p.*

(1) *Il existe un sous-groupe S de G de cardinal p^n.*
(2) *Soit H un sous-groupe de G de cardinal p^ℓ avec $\ell \leqslant n$. Le sous-groupe H de G est conjugué à un sous-groupe de S, et à S lui-même si $\ell = n$.*
(3) *Le nombre de sous-groupes de G de cardinal p^n divise m, et est égal à* 1 *modulo p.*

Démonstration. — Nous allons montrer ces assertions séparément.

Preuve de (1). Soit E l'ensemble des parties à p^n éléments de G. L'action de G sur lui-même par translations induit une action de G sur E. On a

$$|E| = \sum_{O \in G\backslash E} |O|.$$

Le cardinal de E est égal à $\binom{p^n m}{p^n}$. En vertu du lemme 7.1.5, cet entier est premier à p. Il existe donc une orbite O de cardinal premier à p ; soit $P \in O$ et soit S le stabilisateur de P. Puisque le cardinal de O est égal à $p^n m/|S|$, l'entier $|S|$ est multiple de p^n. Et comme P est non vide, le lemme 7.1.1 assure que $|S| \leqslant |P| = p^n$. Il vient $|S| = p^n$.

Preuve de (2). Faisons opérer G sur G/S par translations à gauche, et restreignons cette opération à H. Comme H est un p-groupe, le nombre de points de G/S fixes sous H est égal modulo p au cardinal de G/S, c'est-à-dire à m (4.4.3.2). Puisque m est premier à p, il y a donc au moins un élément de G/S qui est fixe sous H. Autrement dit, H est contenu dans le stabilisateur d'un élément de G/S, stabilisateur qui est en vertu de 4.3.9 de la forme gSg^{-1} avec $g \in G$. L'inclusion $H \subset gSg^{-1}$ peut se récrire $gHg^{-1} \subset S$, et H est conjugué au sous-groupe gHg^{-1} de S. Si $\ell = n$ le cardinal de $g^{-1}Hg$ est égal à p^n, c'est-à-dire à $|S|$; en conséquence, $g^{-1}Hg = S$.

Preuve de (3). Soit F l'ensemble des sous-groupes de G de cardinal p^n. On fait opérer G sur F par conjugaison. Il résulte de l'assertion (2) déjà établie que cette action est transitive. Si T désigne le stabilisateur de S on a donc $|F| = |G|/|T|$. Or on a évidemment $gSg^{-1} = S$ pour tout $g \in S$; par conséquent, $S \subset T$. Il s'ensuit que $|S|$ divise $|T|$. On déduit alors de l'égalité

$$m = |G|/|S| = (|G|/|T|)(|T|/|S|)$$

que $|F| = |G|/|T|$ divise m. Restreignons maintenant l'action de G sur F au sous-groupe S de G. Comme S est un p-groupe, le cardinal de F est égal modulo p au nombre de points fixes de F sous S (4.4.3.2). Il suffit dès lors pour conclure de montrer qu'il y a exactement un tel point fixe, c'est-à-dire exactement un sous-groupe Γ de G de cardinal p^n tel que $g\Gamma g^{-1} = \Gamma$ pour tout $g \in S$. Il est clair que $gSg^{-1} = S$ pour tout $g \in S$. Soit maintenant Γ un sous-groupe de G de cardinal p^n tel que $g\Gamma g^{-1} = \Gamma$ pour tout $g \in S$; nous allons prouver que $\Gamma = S$, ce qui achèvera la démonstration. Soit Γ' le normalisateur de Γ dans G. C'est par définition (exemple 4.3.3) l'ensemble des éléments g de G tels que $g\Gamma g^{-1} = \Gamma$; il est clair que $\Gamma \subset \Gamma'$, et notre hypothèse sur Γ signifie que $S \subset \Gamma'$. Ainsi, S et Γ' apparaissent comme deux sous-groupes de Γ' de cardinal p^n. Puisque p^n est la plus grande puissance de p qui divise $|G|$, c'est *a fortiori* la plus grande puissance de p qui divise $|\Gamma'|$. Il résulte alors de l'assertion (2) déjà démontrée, appliquée au groupe Γ', qu'il existe $g \in \Gamma'$ tel que $g\Gamma g^{-1} = S$. Mais par définition de Γ' on a aussi $g\Gamma g^{-1} = \Gamma$; par conséquent, $\Gamma = S$. □

Nous allons maintenant faire quelques commentaires ; on conserve les notations du théorème 7.2.1 ci-dessus.

7.2.2. — L'assertion (1) du théorème affirme l'existence de sous-groupes de G de cardinal p^n. Un tel sous-groupe est appelé un *p-sous-groupe de Sylow* de G. Notons qu'on n'a pas supposé $n \geqslant 1$; si $n = 0$ (c'est-à-dire si p ne divise pas $|G|$) le groupe G admet un unique p-sous-groupe de Sylow, à savoir $\{e\}$.

7.2.3. — Soit φ un automorphisme de G. Si S est un p-sous-groupe de Sylow de G alors $|\varphi(S)| = |S| = p^n$; ainsi, $\varphi(S)$ est un p-sous-groupe de Sylow de G.

7.2.4. — L'assertion (2) du théorème assure que les p-sous-groupes de Sylow de G sont deux à deux conjugués. Notez deux conséquences importantes de ce point :

- ◇ si S et T sont deux p-sous-groupes de Sylow de G alors $S \simeq T$;
- ◇ s'il existe un p-sous-groupe de Sylow S de G qui est distingué, c'est l'unique p-sous-groupe de Sylow de G.

7.2.5. — Si G possède un unique p-sous-groupe de Sylow S, celui-ci est distingué ; c'est même en vertu de 7.2.3 un sous-groupe *caractéristique* de G, c'est-à-dire que $\varphi(S) = S$ pour tout automorphisme φ de G (notez que distingué signifie simplement que $\varphi(S) = S$ pour tout automorphisme *intérieur* φ de G).

7.2.6. — Faisons une remarque incidente sur la notion de sous-groupe caractéristique. Soit Γ un groupe quelconque et soit Δ un sous-groupe de Γ. Pour que Δ soit un sous-groupe caractéristique de Γ, il suffit que $\varphi(\Delta) \subset \Delta$ pour tout automorphisme φ de Γ. En effet si c'est le cas on a aussi pour un tel φ l'inclusion $\varphi^{-1}(\Delta) \subset \Delta$, et en appliquant φ aux deux membres de l'inclusion il vient $\Delta \subset \varphi(\Delta)$, et finalement $\varphi(\Delta) = \Delta$ (notez que cette égalité peut s'obtenir directement pour des raisons de cardinal si Δ est fini).

7.3. — Nous allons maintenant donner quelques exemples de groupes dont les sous-groupes de Sylow peuvent être décrits explicitement.

***Exemple 7.3.1* (Le cas d'un groupe abélien fini).** — Soit G un groupe abélien fini. Le théorème 3.9.4 assure l'existence d'une (unique) famille finie $(d_1, \ldots, d_r)$ d'entiers strictement supérieurs à 1 avec $d_1|d_2|\ldots|d_r$ telle que $G \simeq \bigoplus_i \mathbf{Z}/d_i\mathbf{Z}$. En décomposant chacun des d_i en produit de facteurs premiers et en appliquant le lemme chinois, on voit qu'on peut également écrire $G \simeq \bigoplus_{j\in J} \mathbf{Z}/p_j^{n_j}\mathbf{Z}$ où J est un ensemble fini, où chacun des p_j est un nombre premier et chacun des n_j un entier strictement positif ; on n'impose pas aux p_j d'être deux à deux distincts. (À titre d'exercice, vous pouvez établir l'unicité de cette écriture à permutation près des termes, par exemple en montrant comment reconstruire les d_i à partir des p_j et des n_j, puis comment exprimer ces derniers à partir des entiers $\ell(p, m)$ introduits dans la preuve du théorème 3.9.4).

Soit p un nombre premier et soit J_0 le sous-ensemble de J formé des indices j tels que $p_j = p$. La plus grande puissance de p divisant $|G|$ est alors $p^{\sum_{j\in J_0} n_j}$, et le sommande $\bigoplus_{j\in J_0} \mathbf{Z}/p^{n_j}\mathbf{Z}$ est un p-sous-groupe de Sylow de G. Il est évidemment distingué puisque G est abélien et c'est en conséquence le seul p-sous-groupe de Sylow de G. (On peut aussi le voir directement en remarquant que ce sous-groupe est exactement l'ensemble des éléments de G de $p^{\sum j\in J_0 n_j}$-torsion ; de ce fait il contient tout p-sous-groupe de Sylow de G, et on conclut avec un argument de cardinal).

***Exemple 7.3.2* (Le cas de $\mathfrak{S}_4$ et de $\mathfrak{A}_4$).** — Pour tout nombre premier p, on note ν_p et ν'_p les nombres respectifs de p-sous-groupes de Sylow de $\mathfrak{S}_4$ et $\mathfrak{A}_4$. Le groupe $\mathfrak{S}_4$ a pour cardinal $24 = 2^3 \cdot 3$, et le groupe $\mathfrak{A}_4$ a pour cardinal $12 = 2^2 \cdot 3$.

On en déduit les faits suivants, à l'aide de l'assertion (3) du théorème 7.2.1 :

- ◇ si p est un nombre premier différent de 2 et de 3 alors $\nu_p = \nu'_p = 1$ (aussi bien $\mathfrak{S}_4$ que $\mathfrak{A}_4$ n'ont qu'un p-sous-groupe de Sylow, à savoir $\{\mathrm{Id}\}$) ;
- ◇ ν_3 divise 8 et est congru à 1 modulo 3, donc vaut 1 ou 4 ; et ν'_3 divise 4 et est congru à 1 modulo 3, donc vaut 1 ou 4 ;
- ◇ ν_2 divise 3 et est congru à 1 modulo 2, donc vaut 1 ou 3 ; et ν'_2 divise 3 est est congru à 1 modulo 2, donc vaut 1 ou 3.

Un 3-sous-groupe de Sylow de $\mathfrak{S}_4$ est un sous-groupe de $\mathfrak{S}_4$ de cardinal 3, c'est-à-dire isomorphe à $\mathbf{Z}/3\mathbf{Z}$, c'est-à-dire encore engendré par un élément d'ordre 3 ; comme les seuls éléments d'ordre 3 de $\mathfrak{S}_4$ sont les 3-cycles, on en déduit que les 3-sous-groupes de Sylow de $\mathfrak{S}_4$ sont :

- ◇ $\{\mathrm{Id}, (1\ 2\ 3), (1\ 3\ 2)\}$;
- ◇ $\{\mathrm{Id}, (1\ 2\ 4), (1\ 4\ 2)\}$;
- ◇ $\{\mathrm{Id}, (1\ 3\ 4), (1\ 4\ 3)\}$;
- ◇ $\{\mathrm{Id}, (2\ 3\ 4), (2\ 4\ 3)\}$.

Il s'ensuit que $\nu_3 = 4$. On remarque de surcroît que tous ces groupes sont contenus dans $\mathfrak{A}_4$; ce sont donc également les 3-sous-groupes de Sylow de $\mathfrak{A}_4$, si bien que $\nu'_3 = 4$.

Nous allons maintenant construire les 2-sous-groupes de Sylow de $\mathfrak{S}_4$. Pour ce faire, on introduit l'ensemble E des partitions de $\{1, 2, 3, 4\}$ en deux sous-ensembles à deux éléments. L'ensemble E lui-même possède trois éléments, à savoir les partitions

$$P_1 := \{1, 2\} \coprod \{3, 4\}, \quad P_2 := \{1, 3\} \coprod \{2, 4\}, \quad \text{et} \quad P_3 := \{1, 4\} \coprod \{2, 3\}.$$

L'opération tautologique de $\mathfrak{S}_4$ sur E en induit une de $\mathfrak{S}_4$ sur E. Elle est transitive : en effet, $(2\ 3)$ envoie la partition P_1 sur P_2, et $(2\ 4)$ l'envoie sur P_3. On en déduit que $\mathrm{Stab}(P_1)$ est de cardinal $24/3 = 8$. C'est donc un 2-sous-groupe de Sylow de $\mathfrak{S}_4$; l'ensemble des 2-sous-groupes de Sylow de $\mathfrak{S}_4$ est alors l'ensemble des conjugués de $\mathrm{Stab}(P_1)$, c'est-à-dire $\{\mathrm{Stab}(P_1), \mathrm{Stab}(P_2), \mathrm{Stab}(P_3)\}$ (*cf.* 4.3.2). On calcule explicitement ces stabilisateurs sans difficulté.

- ◇ Le stabilisateur de P_1 est
$$\{\mathrm{Id}, (1\ 2)(3\ 4), (1\ 3)(2\ 4), (1\ 4)(2\ 3), (1\ 3\ 2\ 4), (1\ 4\ 2\ 3), (1\ 2), (3\ 4)\}.$$
- ◇ Le stabilisateur de P_2 est
$$\{\mathrm{Id}, (1\ 2)(3\ 4), (1\ 3)(2\ 4), (1\ 4)(2\ 3), (1\ 2\ 3\ 4), (1\ 4\ 3\ 2), (1\ 3), (2\ 4)\}.$$
- ◇ Le stabilisateur de P_3 est
$$\{\mathrm{Id}, (1\ 2)(3\ 4), (1\ 3)(2\ 4), (1\ 4)(2\ 3), (1\ 2\ 4\ 3), (1\ 3\ 4\ 2), (1\ 4), (2\ 3)\}.$$

Ces groupes sont donc les 2-sous-groupes de Sylow de $\mathfrak{S}_4$. Ils sont deux à deux distincts ; par conséquent $\nu_2 = 3$.

On vérifie immédiatement que l'intersection

$$\mathrm{Stab}(P_1) \cap \mathrm{Stab}(P_2) \cap \mathrm{Stab}(P_3)$$

est égale à $\{\mathrm{Id}, (1\ 2)(3\ 4), (1\ 3)(2\ 4), (1\ 4)(2\ 3)\}$. On sait que cette intersection coïncide avec le noyau du morphisme $\mathfrak{S}_4 \to \mathfrak{S}_E \simeq \mathfrak{S}_3$ induit par l'action de $\mathfrak{S}_4$ sur E. C'est en particulier un sous-groupe distingué de $\mathfrak{S}_4$, dont on observe qu'il est contenu dans $\mathfrak{A}_4$; il est *a fortiori* distingué dans $\mathfrak{A}_4$. Comme son cardinal est 4, c'est un 2-sous-groupe de Sylow de $\mathfrak{A}_4$; comme il est distingué dans $\mathfrak{A}_4$, c'est le seul (7.2.4) ; par conséquent, $\nu'_2 = 1$.

Mentionnons par ailleurs que comme le noyau du morphisme $\mathfrak{S}_4 \to \mathfrak{S}_E \simeq \mathfrak{S}_3$ est de cardinal 4, son image est de cardinal $24/4 = 6$; ce morphisme est donc surjectif.

***Exemple 7.3.3* (Le cas de $\mathrm{GL}_n(\mathbf{F}_p)$).** — Soit p un nombre premier et soit n un entier $\geqslant 1$. Se donner une matrice appartenant à $\mathrm{GL}_n(\mathbf{F}_p)$ revient à choisir une première colonne non nulle (il y a $p^n - 1$ choix), puis une seconde colonne qui n'est pas multiple de la première (ce qui fait $p^n - p$ choix), puis une troisième colonne qui n'est pas combinaison des deux premières (ce qui fait $p^n - p^2$ choix), etc. Par conséquent,

$$\begin{aligned}|\mathrm{GL}_n(\mathbf{F}_p)| &= (p^n-1)(p^n-p)(p^n-p^2)\dots(p^n-p^{n-1})\\ &= p\cdot p^2\cdot\ldots\cdot p^{n-1}(p^n-1)(p^{n-1}-1)(p^{n-2}-1)\dots(p-1)\\ &= p^{n(n-1)/2}\underbrace{(p^n-1)(p^{n-1}-1)(p^{n-2}-1)\dots(p-1)}_{\text{premier à } p}.\end{aligned}$$

Un sous-groupe de $\mathrm{GL}_n(\mathbf{F}_p)$ est donc un p-sous-groupe de Sylow si et seulement si son cardinal est égal à $p^{n(n-1)/2}$. Or le sous-groupe U de $\mathrm{GL}_n(\mathbf{F}_p)$ constitué des matrices triangulaires supérieures avec des 1 sur la diagonale est précisément de cardinal $p^{n(n-1)/2}$: se donner une telle matrice revient en effet à choisir ses coefficients surdiagonaux, qui sont au nombre de $(n-1)+(n-2)+\ldots+1 = n(n-1)/2$.

Les p-sous-groupes de Sylow de $\mathrm{GL}_n(\mathbf{F}_p)$ sont donc les sous-groupes de la forme $P\mathrm{U}P^{-1}$, où $P \in \mathrm{GL}_n(\mathbf{F}_p)$. Cela se déduit évidemment du fait que les p-sous-groupes de Sylow d'un groupe donné sont deux à deux conjugués. mais dans ce cas particulier, on peut aussi le déduire du corollaire 4.4.10. En effet, ce dernier assure que si G est un sous-groupe de $\mathrm{GL}_n(\mathbf{F}_p)$ qui est un p-groupe, il existe une base de $\mathbf{F}_p^n$ dans laquelle tous les éléments de ce sous-groupe ont une matrice triangulaire supérieure avec des 1 sur la diagonale. Mais cela signifie exactement qu'il existe $P \in \mathrm{GL}_n(\mathbf{F}_p)$ telle que $P^{-1}GP \subset \mathrm{U}$. Et si le cardinal de G est précisément $p^{n(n-1)/2}$ on a alors nécessairement $P^{-1}GP = \mathrm{U}$, soit encore $G = P\mathrm{U}P^{-1}$.

7.4. — Nous allons maintenant illustrer la puissance des théorèmes de Sylow en montrant comment ils permettent parfois de classer à isomorphisme près tous les groupes de cardinal donné.

7.4.1. *Classification des groupes de cardinal* 15. — Soit G un groupe de cardinal 15. On a $15 = 3\cdot 5$. Le nombre de 5-sous-groupes de Sylow de G divise 3, et est congru à 1 modulo 5 ; il y en a donc exactement un, que l'on note H et qui est distingué dans G ; étant de cardinal 5, le groupe H est isomorphe à $\mathbf{Z}/5\mathbf{Z}$. Soit K un 3-sous-groupe de Sylow de G ; il est isomorphe à $\mathbf{Z}/3\mathbf{Z}$.

Le sous-groupe H est distingué dans G, les cardinaux de H et K sont premiers entre eux, et $|H|\cdot|K| = |G|$. Il résulte alors de 6.1.4 que G s'identifie à $H \rtimes_\psi K$ pour un certain morphisme ψ de K vers Aut H. Il existe donc un morphisme
phi de $\mathbf{Z}/3\mathbf{Z}$ vers Aut $\mathbf{Z}/5\mathbf{Z}$ tel que $G \simeq \mathbf{Z}/5\mathbf{Z} \rtimes_\varphi \mathbf{Z}/3\mathbf{Z}$. Comme 3 est premier à $\Phi(5) = 4$, le morphisme φ est trivial et G est isomorphe au produit *direct* $\mathbf{Z}/5\mathbf{Z}\times\mathbf{Z}/3\mathbf{Z}$ (6.2.7.4), c'est-à-dire à $\mathbf{Z}/15\mathbf{Z}$.

7.4.2. *Classification des groupes de cardinal* 21. — Soit G un groupe de cardinal 21. On a $21 = 3\cdot 7$. Le nombre de 7-sous-groupes de Sylow de G divise 3, et est congru à 1 modulo 7 ; il y en a donc exactement un, que l'on note H et qui est distingué dans G ; étant de cardinal 7, le groupe H est isomorphe à $\mathbf{Z}/7\mathbf{Z}$. Soit K un 3-sous-groupe de Sylow de G ; il est isomorphe à $\mathbf{Z}/3\mathbf{Z}$.

Le sous-groupe H est distingué dans G, les cardinaux de H et K sont premiers entre eux, et $|H|\cdot|K| = |G|$. Il résulte alors de 6.1.4 que G s'identifie à $H \rtimes_\psi K$ pour un certain morphisme $\psi\colon K \to \operatorname{Aut} H$. Il existe donc un morphisme φ de $\mathbf{Z}/3\mathbf{Z}$ vers $\operatorname{Aut} \mathbf{Z}/7\mathbf{Z}$ tel que $G \simeq \mathbf{Z}/7\mathbf{Z} \rtimes_\varphi \mathbf{Z}/3\mathbf{Z}$. Il résulte alors de l'exemple 6.2.8 que l'on est dans l'un des deux cas suivants, exclusifs l'un de l'autre.

- *Premier cas.* Le groupe G est isomorphe à $\mathbf{Z}/7\mathbf{Z} \times \mathbf{Z}/3\mathbf{Z} \simeq \mathbf{Z}/21\mathbf{Z}$.
- *Second cas.* Le groupe G est isomorphe à $\mathbf{Z}/7\mathbf{Z} \rtimes_\varphi \mathbf{Z}/3\mathbf{Z}$, où φ est donné par la formule $\varphi(\underbrace{\quad\overline{r}\quad}_{\text{mod } 3})(x) = \underbrace{\quad\overline{2}^r\quad}_{\text{mod } 7} x$.

8. Suites de Jordan-Hölder, groupes résolubles et nilpotents

8.1. — Lorsqu'on étudie les objets d'une théorie mathématique on cherche souvent, lorsque la théorie en question permet de donner un sens à cette quête, à les «dévisser» en objets aussi élémentaires que possible. C'est ce que nous allons faire dans le cas des groupes. Nous allons commencer par expliquer ce que sont, dans ce contexte, les objets élémentaires et les dévissages.

8.1.1. *Les «dévissages» en théorie des groupes.* — En théorie des groupes, l'opération de base permettant de «dévisser» un groupe est le passage au quotient par un sous-groupe distingué : si G est un groupe et si $H \lhd G$, la suite exacte $1 \to H \to G \to G/H \to 1$ sera considérée comme un dévissage de G en deux constituants, le sous-groupe H et le groupe quotient G/H ; si G est fini et si H est un sous-groupe strict de G, chacun de ces deux constituants est de cardinal strictement inférieur à $|G|$. Ce principe nous conduit à la définition qui suit.

Définition 8.1.2. — Soit G un groupe et soit n un entier. On appellera *suite de composition de longueur n de G* toute suite

$$\{e\} = G_n \lhd G_{n-1} \lhd \ldots \lhd G_{n-1} \lhd G_0 = G$$

de sous-groupes de G, chacun étant distingué dans le suivant. Notons qu'on ne demande pas qu'ils soient deux à deux distincts, et que le cas limite $n = 0$ est autorisé (mais n'est possible que si G est trivial).

Les groupes $G_{n-1}/G_n, \ldots, G_0/G_1$ seront appelés les *quotients successifs* du dévissage $\{e\} = G_n \lhd G_{n-1} \lhd \ldots \lhd G_1 \lhd G_0 = G$.

8.1.3. *Rappels.* — Soit G un groupe et soit H un sous-groupe distingué de G. Rappelons quelques faits (lemme 2.14.1, lemme 2.14.3, remarque 2.14.4) que nous utiliserons librement dans la suite de cette section.

Si Γ est un sous-groupe de G, $\Gamma \cap H$ est distingué dans Γ et l'application naturelle $\Gamma \to G/H$ induit une injection $\Gamma/\Gamma \cap H \hookrightarrow G/H$, et permet donc de voir $\Gamma/\Gamma \cap H$ comme un sous-groupe de G/H. Si Γ est distingué dans G alors $\Gamma/\Gamma \cap H$ est distingué dans G/H.

Si Δ est un sous-groupe de G/H, il est de la forme Γ/H pour un unique sous-groupe Γ de G contenant H, égal à l'image réciproque de Δ dans G. De surcroît, $\Delta = \Gamma/H$ est distingué dans G/H si et seulement si Γ est distingué dans G, et si c'est le cas alors $G/\Gamma = (G/H)/(\Gamma/H)$.

8.1.4. — Soit G un groupe et soit $\{e\} = G_n \lhd G_{n-1} \lhd \ldots \lhd G_1 \lhd G_0 = G$ une suite de composition de G. Supposons donné pour tout i compris entre 0 et $n-1$ une suite de composition $\mathscr{S}_i$ de G_i/G_{i+1}. En vertu du 8.1.3 ci-dessus, chacune des $\mathscr{S}_i$ peut être vu comme une suite de groupes $G_{i+1} = \Gamma_{i,r_i} \lhd \Gamma_{i,r_i-1} \lhd \ldots \lhd \Gamma_{i,0} = G_i$. La suite

$$\{e\} = \Gamma_{n,r_n} \lhd \Gamma_{n,r_n-1} \lhd \ldots \lhd \Gamma_{n,0} = \Gamma_{n-1,r_{n-1}} \lhd \Gamma_{n-1,r_{n-1}-1} \lhd \ldots \lhd \Gamma_{1,0} = \Gamma_{0,r_0} \lhd \ldots \lhd \Gamma_{0,0} = G$$

est alors une suite de composition $\mathscr{S}$ de G, appelé la *concaténation* des $\mathscr{S}_i$; la liste des quotients successifs de $\mathscr{S}$ est la concaténation des listes des quotients successifs des $\mathscr{S}_i$; la longueur de $\mathscr{S}$ est la somme des longueurs des $\mathscr{S}_i$.

Définition 8.1.5. — Un groupe G est dit *simple* si $G \neq \{e\}$ et si les seuls sous-groupes distingués de G sont $\{e\}$ et G.

Exemples sans justification 8.1.6. — Les faits suivants seront démontrés en TD : le groupe $\mathfrak{A}_n$ est simple dès que $n \geqslant 5$; le groupe $\mathrm{SO}_3(\mathbf{R})$ est simple.

Remarque 8.1.7. — Soit $\varphi \colon G \to \Gamma$ un morphisme de groupes surjectif. Supposons que G soit simple. Le noyau de φ étant distingué dans G, il est égal à G ou $\{e\}$; par conséquent, ou bien $\Gamma = \{e\}$ ou bien φ est un isomorphisme.

Lemme 8.1.8. — *Soit G un groupe abélien fini. Le groupe G est simple si et seulement si G est isomorphe à $\mathbf{Z}/p\mathbf{Z}$ pour un certain p premier.*

Démonstration. — Soit p un nombre premier et soit H un sous-groupe de $\mathbf{Z}/p\mathbf{Z}$. Comme $|H|$ divise p le cardinal de H est ou bien égal à 1 (auquel cas H est trivial) ou bien égal à p (auquel cas $H = \mathbf{Z}/p\mathbf{Z}$). Ainsi, $\mathbf{Z}/p\mathbf{Z}$ est simple.

Soit G un groupe abélien fini simple. Cela signifie que $G \neq \{e\}$ et que les seuls sous-groupes de G sont G et $\{e\}$. Comme G est non trivial, il existe $g \neq 0$ dans G ; notons n l'ordre de g. Le sous-groupe $\langle g \rangle$ de G est non trivial et donc égal à G. Il est par ailleurs isomorphe à $\mathbf{Z}/n\mathbf{Z}$; l'ensemble de ses sous-groupes est donc en bijection avec l'ensemble des diviseurs de n. Comme G a exactement deux sous-groupes, n a exactement deux diviseurs et est en conséquence premier. □

Définition 8.1.9. — Soit G un groupe. Une *suite de Jordan-Hölder* de G est une suite de composition de G dont les quotients successifs sont simples.

Les suites de Jordan-Hölder sont les «dévissages en objets élémentaires» en théorie des groupes. Nous allons commencer par montrer qu'un groupe fini admet toujours une telle suite.

Lemme 8.1.10. — *Soit G un groupe fini. Il possède une suite de Jordan-Hölder.*

Démonstration. — On raisonne par récurrence sur $|G|$. Si $|G| = 1$ alors G possède une suite de Jordan-Höder de longueur nulle (constituée du groupe $G = \{e\}$ lui-même) ; la liste de ses quotients successifs est *vide*, et répond donc tautologiquement aux conditions de l'énoncé.

Supposons que $|G| > 1$ et que le résultat a été montré pour tout groupe de cardinal strictement inférieur à $|G|$. Si G est simple la suite de Jordan-Hölder $\{e\} \lhd G$ répond aux conditions de l'énoncé. Sinon, il existe un sous-groupe distingué H de G qui est non trivial et strict. Les cardinaux de H et G/H sont alors tous deux strictement inférieurs à $|G|$. Par hypothèse de récurrence, chacun d'eux admet une suite de Jordan-Hölder. En concaténant ces deux suites on obtient une suite de Jordan-Hölder de G. □

Corollaire 8.1.11. — *Soit G un groupe abélien fini. Il possède une suite de composition dont tous les quotients successifs sont de la forme $\mathbf{Z}/p\mathbf{Z}$ avec p premier.*

Démonstration. — Le lemme 8.1.10 ci-dessus assure l'existence d'une suite de Jordan-Hölder $\mathscr{S}$ de G. Par ailleurs les quotients successifs de $\mathscr{S}$ sont abéliens puisque G est abélien ; chacun d'eux est donc à la fois simple et abélien, c'est-à-dire cyclique de cardinal premier en vertu du lemme 8.1.8. □

Notons que tous les groupes n'admettent pas de suite de Jordan-Hölder, comme en atteste le contre-exemple très simple suivant.

Contre-exemple 8.1.12. — Le groupe $\mathbf{Z}$ n'admet pas de suite de Jordan-Hölder. En effet, supposons qu'il admette une telle suite $\mathscr{S}$. Chacun des quotients successifs de $\mathscr{S}$ serait alors à la fois simple et abélien (car $\mathbf{Z}$ est abélien), et donc cyclique d'ordre premier d'après le lemme 8.1.8 ; en particulier, les quotients successifs de $\mathscr{S}$ seraient tous finis, et $\mathbf{Z}$ serait donc fini, ce qui est absurde.

Remarque 8.1.13. — Le lemme 8.1.10 ramène en un sens l'étude des groupes finis à celle des groupes simples. Cela dit, la classification des groupes finis simples (à isomorphisme près) est achevée mais est extrêmement ardue. Et même en la prenant pour acquise, elle ne permet pas d'obtenir la classification de tous les groupes finis. En effet, trouver tous les groupes admettant une suite de Jordan-Hölder ayant une suite donnée de quotients successifs n'a rien d'évident. Par exemple si l'on se donne deux groupes H et K, trouver tous les groupes admettant une suite de Jordan-Hölder ayant comme quotients successifs H et K revient à trouver tous les groupes G s'insérant dans une suite exacte $1 \to H \to G \to K \to 1$. La réponse est connue si l'on suppose de plus que la suite est scindée : il s'agit des produits semi-directs $H \rtimes_{\varphi} K$ (mais il reste tout de même à déterminer tous les morphismes de K dans $\operatorname{Aut} H$, et à comprendre quand deux tels morphismes conduisent à des groupes isomorphes) ; mais sans cette hypothèse, c'est déjà un problème redoutable.

8.2. Un résultat d'unicité sur les suites de Jordan-Hölder. — Nous allons considérer la question suivante. Soit G un groupe ; supposons données deux suites de Jordan-Hölder $\mathscr{S}$ et $\mathscr{S}'$ de G ; que peut-on dire de $\mathscr{S}$ et $\mathscr{S}'$? Pour avoir une idée de ce qu'il est réaliste d'attendre, considérons le cas du groupe $\mathbf{Z}/6\mathbf{Z}$ (dont tous les sous-groupes sont distingués, puisqu'il est abélien !). Son sous-groupe $\langle\overline{2}\rangle$ est isomorphe à $\mathbf{Z}/3\mathbf{Z}$, et le quotient $(\mathbf{Z}/6\mathbf{Z})/\langle\overline{2}\rangle$ est isomorphe à $\mathbf{Z}/2\mathbf{Z}$; son sous-groupe $\langle\overline{3}\rangle$ est isomorphe à $\mathbf{Z}/2\mathbf{Z}$, et le quotient $(\mathbf{Z}/6\mathbf{Z})/\langle\overline{3}\rangle$ est isomorphe à $\mathbf{Z}/3\mathbf{Z}$. Comme $\mathbf{Z}/2\mathbf{Z}$ et $\mathbf{Z}/3\mathbf{Z}$ sont simples, on a deux suites de Jordan-Hölder pour $\mathbf{Z}/6\mathbf{Z}$, à savoir

$$\mathbf{Z}/6\mathbf{Z} \rhd \langle\overline{2}\rangle \rhd \{0\} \text{ et } \mathbf{Z}/6\mathbf{Z} \rhd \langle\overline{3}\rangle \rhd \{0\}.$$

Ces deux suites sont différentes, et les listes ordonnées de leurs quotients (à isomorphisme près) le sont aussi : pour la première c'est $\mathbf{Z}/3\mathbf{Z}, \mathbf{Z}/2\mathbf{Z}$ et pour la seconde $\mathbf{Z}/2\mathbf{Z}, \mathbf{Z}/3\mathbf{Z}$. Par contre, ces listes non ordonnées sont les mêmes.

Le mieux qu'on puisse espérer est donc que deux suites de Jordan-Hölder d'un groupe donné aient même liste de quotients successifs *à permutation près.* Nous allons voir un peu plus bas que c'est effectivement le cas.

Lemme 8.2.1. — *Soit G un groupe et soient H et K deux sous-groupes distingués de G. On suppose que G/K est simple. Le sous-groupe $H/H \cap K$ de G/K est alors* *ou bien trivial, ou bien égal à G/K. Il est trivial si et seulement si $H \subset K$. Si c'est le cas et si G/H est également simple alors $K = H$.*

Démonstration. — Comme H est distingué dans G, le quotient $H/H \cap K$ est un sous-groupe distingué de G/K ; ce dernier étant simple, $H/H \cap K$ est ou bien trivial, ou bien égal à G/K.

ll est clair qu'il est trivial si et seulement si $H \subset K$. Supposons que c'est le cas et que G/H est simple. Le quotient K/H est alors un sous-groupe distingué de G/H (puisque K est distingué dans G). Comme G/H est simple, K/H est ou bien trivial, ce qui équivaut à l'égalité $K = H$, ou bien égal à G/H. Mais si l'on avait $K/H = G/H$ on aurait $K = G$ et le quotient G/K serait trivial, ce qui est absurde car il est simple. □

Définition 8.2.2. — Soit G un groupe et soit

$$\mathscr{S} = (G = G_n \triangleright G_{n-1} \triangleright \ldots \triangleright G_1 \triangleright G_0 = \{e\})$$

une suite de Jordan-Hölder de G. Soit H un groupe simple. On appelle *multiplicité de H dans $\mathscr{S}$*, et l'on note $\mu(\mathscr{S}, H)$, le cardinal de $\{i \in \{1, \ldots, n\}$ t.q. $G_i/G_{i-1} \simeq H\}$.

Remarque 8.2.3. — Soient $\mathscr{S}$ une suite de Jordan-Hölder d'un groupe G. Il résulte de la définition que pour tout groupe simple H, la multiplicité $\mu(\mathscr{S}, H)$ ne dépend que de la classe d'isomorphie de H. Si l'on se donne un «système de représentants» $\mathscr{H}$ des classes d'isomorphie de groupes simples alors $\mu(\mathscr{S}, H) = 0$ pour presque tout $H \in \mathscr{H}$ et la longueur de $\mathscr{S}$ est égale à $\sum_{H \in \mathscr{H}} \mu(\mathscr{S}, H)$.

Nous pouvons maintenant énoncer le résultat annoncé sur l'unicité à permutation près de la liste des quotients d'une suite de Jordan-Hölder.

Théorème 8.2.4. — *Soit G un groupe. Pour tout couple $(\mathscr{S}, \mathscr{S}')$ de suites de Jordan-Hölder de G et tout groupe simple H les multiplicités $\mu(\mathscr{S}, H)$ et $\mu(\mathscr{S}', H)$ coïncident ; en particulier, $\mathscr{S}$ et $\mathscr{S}'$ ont même longueur.*

Démonstration. — On suppose que G admet au moins une suite de Jordan-Hölder (sans quoi le théorème est vrai, mais vide), et on raisonne par récurrence sur la longueur minimale n d'une telle suite. Si $n = 0$ le groupe G possède une suite de Jordan-Hölder de longueur nulle, et est donc trivial ; par conséquent, toute suite de Jordan-Hölder de G est de longueur nulle (G ne peut pas avoir de quotient simple puisqu'un groupe simple est non trivial). On suppose maintenant $n > 0$ et le théorème vrai pour les entiers $< n$. On fixe une suite de Jordan-Hölder

$$\mathscr{S} = (G = G_n \triangleright G_{n-1} \triangleright \ldots \triangleright G_1 \triangleright G_0 = \{e\})$$

de longueur n, et l'on se donne une seconde suite de Jordan-Hölder

$$\mathscr{S}' = (G = L_m \triangleright L_{m-1} \triangleright \ldots \triangleright L_1 \triangleright L_0 = \{e\}) \ ;$$

notons que par définition de n on a nécessairement $m \geqslant n$.

Nous allons montrer que $\mu(\mathscr{S}, H) = \mu(\mathscr{S}', H)$ pour tout groupe simple H, ce qui permettra de conclure. On pose

$$\Gamma = (G_{n-1} \triangleright \ldots \triangleright G_1 \triangleright G_0 = \{e\})$$

et

$$\Lambda = (L_{m-1} \rhd \ldots \rhd L_1 \rhd L_0 = \{e\}).$$

On remarque que Γ est une suite de Jordan-Hölder de G_{n-1}, et que Λ est une suite de Jordan-Hölder de L_{m-1}. Comme la longueur de Γ est strictement inférieure à n, l'hypothèse de récurrence assure que G_{n-1} satisfait les conclusions du théorème ; en particulier, toute suite de Jordan-Hölder de G_{n-1} est de longueur $n-1$.

Supposons que $L_{m-1} = G_{n-1}$, ce qui entraîne au vu de ce qui précède l'égalité $\mu(\Lambda, H) = \mu(\Gamma, H)$ pour tout groupe simple H ; donnons-nous un tel H.

- $\diamond$ Si $H \simeq G/G_{n-1} = G/L_{m-1}$ alors
$$\mu(\mathscr{S}, H) = \mu(\Gamma, H) + 1 \text{ et } \mu(\mathscr{S}', H) = \mu(\Lambda, H) + 1 = \mu(\Gamma, H) + 1.$$
- $\diamond$ Si $H \not\simeq G/G_{n-1} = G/L_{m-1}$ alors
$$\mu(\mathscr{S}, H) = \mu(\Gamma, H) \text{ et } \mu(\mathscr{S}', H) = \mu(\Lambda, H) = \mu(\Gamma, H).$$

On voit ainsi que $\mu(\mathscr{S}, H) = \mu(\mathscr{S}', H)$; le théorème est donc démontré lorsque $L_{m-1} = G_{n-1}$.

Supposons à partir de maintenant que $L_{m-1} \neq G_{n-1}$. En vertu du lemme 8.2.1, le groupe quotient $G_{n-1}/(L_{m-1} \cap G_{n-1})$ est ou bien trivial, ou bien égal à G/L_{m-1}, et il ne peut être trivial que si $L_{m-1} = G_{n-1}$; comme ce dernier cas est exclu par hypothèse, il vient $G_{n-1}/(L_{m-1} \cap G_{n-1}) = G/L_{m-1}$. Par symétrie des arguments, $L_{m-1}/(L_{m-1} \cap G_{n-1}) = G/G_{n-1}$.

Posons

$$\mathscr{T} = (G_{n-1} \rhd (L_{m-1} \cap G_{n-1}) \rhd (L_{m-2} \cap G_{n-1}) \rhd \ldots \rhd (L_1 \cap G_{n-1}) \rhd \{e\}).$$

C'est une suite de composition de G_{n-1}. Et pour tout i compris entre 1 et m le quotient $(L_i \cap G_{n-1})/(L_{i-1} \cap G_{n-1})$ est ou bien égal à L_i/L_{i-1}, ou bien trivial : c'est une conséquence du lemme 8.2.1 appliqué au groupe L_i et à ses sous-groupes distingués L_{i-1} et $L_i \cap G_{n-1}$. On voit donc qu'en éliminant de la suite $\mathscr{T}$ les termes redondants de façon à n'avoir aucun quotient trivial, on obtient une suite de composition Γ' de G_{n-1} dont tout quotient est de la forme L_i/L_{i-1} pour un certain i, et est en particulier simple ; c'est en conséquence une suite de Jordan-Hölder de G_{n-1}, et sa longueur est dès lors égale à $n-1$. Puisque le quotient $G_{n-1}/(L_{m-1} \cap G_{n-1})$ s'identifie à G/L_{m-1} il est non trivial, et la suite Γ' est donc de la forme

$$G_{n-1} \rhd (L_{m-1} \cap G_{n-1}) \rhd M_{n-3} \rhd \ldots \rhd M_1 \rhd M_0 = \{e\}.$$

On a vu plus haut que le quotient $L_{m-1}/(L_{m-1} \cap G_{n-1})$ est égal au groupe simple G/G_{n-1} ; par conséquent,

$$\Lambda' := (L_{m-1} \rhd (L_{m-1} \cap G_{n-1}) \rhd M_{n-3} \rhd \ldots \rhd M_1 \rhd M_0 = \{e\})$$

est une suite de Jordan-Hölder de L_{m-1}. Puisqu'elle est de longueur $n-1$, le groupe L_{m-1} satisfait les conclusions du théorème (en particulier Λ et Λ' ont même longueur, si bien que $m = n$).

Notons Ω la suite de Jordan-Hölder $M_{n-3} \rhd \ldots \rhd M_1 \rhd M_0 = \{e\}$. Nous allons terminer notre raisonnement à l'aide des quatre suites de Jordan-Hölder ci-dessous

(les accolades supérieures indiquent le nom des suites, les accolades inférieures la valeur des quotients)

$$\overbrace{G \rhd \overbrace{G_{n-1} \rhd \ldots \rhd G_1 \rhd \{e\}}^{\Gamma}}^{\mathscr{S}}$$

$$\overbrace{\underbrace{G_{n-1} \rhd (L_{n-1} \cap G_{n-1})}_{G/L_{n-1}} \rhd \overbrace{M_{n-3} \rhd \ldots \rhd \{e\}}^{\Omega}}^{\Gamma'}$$

$$\overbrace{G \rhd \overbrace{L_{n-1} \rhd \ldots \rhd L_1 \rhd \{e\}}^{\Lambda}}^{\mathscr{S}'}$$

$$\overbrace{\underbrace{L_{n-1} \rhd (L_{n-1} \cap G_{n-1})}_{G/G_{n-1}} \rhd \overbrace{M_{n-3} \rhd \ldots \rhd \{e\}}^{\Omega}}^{\Lambda'}$$

Soit H un groupe simple ; rappelons que $\mu(\Lambda, H) = \mu(\Lambda', H)$ et $\mu(\Gamma, H) = \mu(\Gamma', H)$ puisque G_{n-1} et L_{n-1} satisfont les conclusions du théorème.

◇ *Supposons que $H, G/G_{n-1}$ et G/L_{n-1} sont deux à deux non isomorphes.* On a alors

$$\begin{array}{rcl} \mu(\mathscr{S}, H) & = & \mu(\Gamma, H) \\ & = & \mu(\Gamma', H) \\ & = & \mu(\Omega, H) \\ & = & \mu(\Lambda', H) \\ & = & \mu(\Lambda, H) \\ & = & \mu(\mathscr{S}', H). \end{array}$$

◇ *Supposons que H et G/G_{n-1} sont isomorphes, mais que G/G_{n-1} et G/L_{n-1} ne le sont pas.* On a alors

$$\begin{array}{rcl} \mu(\mathscr{S}, H) & = & \mu(\Gamma, H) + 1 \\ & = & \mu(\Gamma', H) + 1 \\ & = & \mu(\Omega, H) + 1 \\ & = & \mu(\Lambda', H) \\ & = & \mu(\Lambda, H) \\ & = & \mu(\mathscr{S}', H). \end{array}$$

◇ *Supposons que H et G/L_{n-1} sont isomorphes, mais que G/G_{n-1} et G/L_{n-1} ne le sont pas.* On a alors

$$\begin{aligned}\mu(\mathscr{S},H) &= \mu(\Gamma,H)\\ &= \mu(\Gamma',H)\\ &= \mu(\Omega,H)+1\\ &= \mu(\Lambda',H)+1\\ &= \mu(\Lambda,H)+1\\ &= \mu(\mathscr{S}',H).\end{aligned}$$

◇ *Supposons que $H, G/L_{n-1}$ et G/G_{n-1} sont deux à deux isomorphes.* On a alors

$$\begin{aligned}\mu(\mathscr{S},H) &= \mu(\Gamma,H)+1\\ &= \mu(\Gamma',H)+1\\ &= \mu(\Omega,H)+2\\ &= \mu(\Lambda',H)+1\\ &= \mu(\Lambda,H)+1\\ &= \mu(\mathscr{S}',H).\end{aligned}$$

On a donc bien prouvé que $\mu(\mathscr{S},H)=\mu(\mathscr{S}',H)$ dans tous les cas, ce qui achève la démonstration. □

Commentaires 8.2.5. — Au cours de la preuve ci-dessus, c'est lorsqu'on établit les égalités $G_{n-1}/(L_{m-1}\cap G_{n-1})=G/L_{m-1}$ et $L_{m-1}/(L_{m-1}\cap G_{n-1})=G/G_{n-1}$ (à l'aide du lemme 8.2.1) qu'il se «passe vraiment quelque chose». C'est en effet à cette occasion qu'on exprime l'un des quotients successifs de la seconde suite de Jordan-Hölder (à savoir G/L_{m-1}) comme quotient d'un groupe apparaissant dans la première (à savoir G_{n-1}), et vice-versa. Le reste de la démonstration consiste essentiellement en des manipulations formelles.

Notation 8.2.6. — Soit G un groupe. Si G admet une suite de Jordan-Hölder $\mathscr{S}$ alors pour tout groupe simple H, l'entier $\mu(\mathscr{S},H)$ est indépendant du choix de $\mathscr{S}$; nous le noterons simplement $\mu(G,H)$.

Proposition 8.2.7. — *Soit G un groupe et soit K un sous-groupe distingué de G. Les assertions suivantes sont équivalentes :*

(i) *chacun des groupes K et G/K possède une suite de Jordan-Hölder ;*

(ii) *G possède une suite de Jordan-Hölder.*

De plus si elles sont satisfaites alors $\mu(G,H)=\mu(K,H)+\mu(G/K,H)$ pour tout groupe simple H.

Démonstration. — Supposons (i). En concaténant une suite de Jordan-Hölder $\mathscr{S}$ de K et une suite de Jordan-Hölder $\mathscr{S}'$ de G/K on obtient une suite de Jordan-Hölder $\mathscr{S}''$ de G, et on a pour tout groupe simple H l'égalité $\mu(\mathscr{S}'',H)=\mu(\mathscr{S},H)+\mu(\mathscr{S}',H)$, et par conséquent $\mu(G,H)=\mu(\mathscr{K},H)=\mu(G/K,H)$.

Supposons maintenant (ii). Soit

$$G = G_n \rhd G_{n-1} \rhd \ldots \rhd G_1 \rhd G_0 = \{e\}$$

une suite de Jordan-Hölder de G. Pour tout i, il résulte du lemme 8.2.1 que le quotient $G_i \cap K/(G_{i-1} \cap K)$ est ou bien trivial, ou bien isomorphe à G_i/G_{i-1}, qui est simple. Par conséquent, en partant de la suite de composition

$$K \rhd (K \cap G_{n-1}) \rhd \ldots \rhd (K \cap G_1) \rhd \{e\}$$

et en éliminant ses termes redondants on obtient une suite de Jordan-Hölder de K.

Désignons pour tout i par G'_i l'image de G_i dans le groupe quotient G/K. Fixons i entre 1 et n. La flèche composée $G_i \to G'_i \to G'_i/G'_{i-1}$ est surjective et son noyau contient G_{i-1} ; elle induit donc une surjection $(G_i/G_{i-1}) \to (G'_i/G'_{i-1})$. Comme G_i/G_{i-1} est simple, on en déduit que G'_i/G'_{i-1} est ou bien trivial ou bien isomorphe à G_i/G_{i-1} et en particulier simple (remarque 8.1.7). Par conséquent, en partant de la suite de composition

$$(G/K) = G'_n \rhd G'_{n-1} \rhd \ldots \rhd G'_1 \rhd \{e\}$$

et en éliminant ses termes redondants on obtient une suite de Jordan-Hölder de G/K. □

8.3. Commutateurs. — Le reste de la section 8 va être consacré à l'étude des groupes *résolubles*, c'est-à-dire des groupes qui possèdent une suite de composition à quotients abéliens, avec une focalisation sur une classe particulière de groupes résolubles, les groupes *nilpotents*. Mais nous allons tout d'abord introduire divers «groupes de commutateurs» dont nous nous servirons de manière cruciale.

Définition 8.3.1. — Soit G un groupe et soient a et b deux éléments de G. Le *commutateur* de a et b est l'élément $aba^{-1}b^{-1}$ de G, qui est aussi noté $[a, b]$. Il est trivial si et seulement si a et b commutent.

8.3.2. — Soit G un groupe et soient a et b deux éléments de G. On a alors

$$[a, b]^{-1} = bab^{-1}a^{-1} = [b, a].$$

Si φ est un morphisme de groupes de source G on a $\varphi([a, b]) = [\varphi(a), \varphi(b)]$.

8.3.3. — Soit G un groupe et soient H et K deux sous-groupes de G. On notera $[H, K]$ le sous-groupe engendré par les commutateurs de la forme $[h, k]$ avec $h \in H$ et $k \in K$.

8.3.3.1. — Si H' est un sous-groupe de H et si K' est un sous-groupe de K, il résulte de la définition que $[H', K'] \subset [H, K]$.

8.3.3.2. — On déduit de 8.3.2 que pour tout morphisme φ de G dans un groupe G', on a $\varphi([H, K]) = [\varphi(H), \varphi(K)]$.

Lemme 8.3.4. — *Supposons que G s'écrive comme un produit (direct) fini de groupes $G_1 \times \ldots \times G_n$, et que H et K soient respectivement de la forme $\prod H_i$ et $\prod K_i$ où H_i et K_i sont pour tout i des sous-groupes de G_i. On a alors*

$$[H, K] = \prod [H_i, K_i].$$

Démonstration. — Nous allons montrer la double inclusion. Soit $h = (h_1, \ldots, h_n) \in H$ et soit $k = (k_1, \ldots, k_n) \in K$; notre hypothèse sur H et K assure que $h_i \in H_i$ et $k_i \in K_i$ pour tout i. On a alors $[h, k] = ([h_i, k_i])_i$. Ainsi, $[h, k] \in \prod [H_i, K_i]$; ceci valant pour tout (h, k), il vient $[H, K] \subset \prod [H_i, K_i]$. *Notons que l'on n'a pas utilisé ici le fait que les produits en jeu sont finis ; c'est pour l'autre inclusion que nous en aurons besoin.*

Réciproquement, soit (a_i) un élément de $\prod [H_i, K_i]$; nous allons montrer qu'il appartient à $[H, K]$, ce qui permettra de conclure. Pour tout i, notons α_i l'élément de G dont la i-ème coordonnées est égale à a_i et dont les autres sont nulles. On a alors $(a_i) = \alpha_1 \alpha_2 \ldots \alpha_n$; il suffit donc de montrer que chacun des α_i appartient à $[H, K]$. Fixons i. Puisque $a_i \in [H_i, K_i]$, il existe une famille $(h_j, k_j)_{1 \leqslant j \leqslant m}$ d'éléments de $H_i \times K_i$, et une famille (ε_j) d'éléments de $\{-1, 1\}$, telles que
$$a_i = [h_1, k_1]^{\varepsilon_1} [h_2, k_2]^{\varepsilon_2} \ldots [h_m, k_m]^{\varepsilon_m}.$$
Pour tout j, notons λ_j (resp. μ_j) l'élément de G dont la i-ème coordonnée est égale à h_j (resp. k_j) et dont les autres coordonnées sont triviales. Par construction, on a $\lambda_j \in H$ et $\mu_j \in K$ pour tout j, et
$$\alpha_i = [\lambda_1, \mu_1]^{\varepsilon_1} [\lambda_2, \mu_2]^{\varepsilon_2} \ldots [\lambda_m, \mu_m]^{\varepsilon_m}.$$
Ainsi $\alpha_i \in [H, K]$, ce qu'il fallait démontrer. □

8.3.5. — Soit G un groupe. On définit récursivement deux suites $(\mathrm{D}^n(G))$ et $(\mathrm{C}^n(G))$ de sous-groupes de G comme suit :

◇ $\mathrm{D}^0(G) = G$ et $\mathrm{D}^n(G) = [\mathrm{D}^{n-1}(G), \mathrm{D}^{n-1}(G)]$ pour tout $n \geqslant 1$.

◇ $\mathrm{C}^0(G) = G$ et $\mathrm{C}^n(G) = [G, \mathrm{C}^{n-1}(G)]$ pour tout $n \geqslant 1$.

On écrira souvent $\mathrm{D}(G)$ au lieu de $\mathrm{D}^1(G)$, et $\mathrm{C}(G)$ au lieu de $\mathrm{C}^1(G)$. Notez que $\mathrm{C}(G) = \mathrm{D}(G) = [G, G]$; on l'appelle le *groupe dérivé* de G.

8.3.6. — Soit G un groupe. Nous allons mentionner quelques propriétés élémentaires des groupes $\mathrm{D}^n(G)$ et $\mathrm{D}^n(G)$.

8.3.6.1. — Le groupe $\mathrm{D}(G)$ étant engendré par les commutateurs, il est trivial si et seulement si chaque commutateur est trivial, c'est-à-dire si et seulement si G est abélien.

8.3.6.2. — On déduit de 8.3.3.1 que pour tout entier n et tout sous-groupe H de G on a les inclusions
$$\mathrm{D}^n(G) \subset C^n(G), \;\; \mathrm{D}^n(H) \subset \mathrm{D}^n(G) \;\text{ et }\; \mathrm{C}^n(H) \subset \mathrm{C}^n(G).$$

8.3.6.3. — Soit $\varphi \colon G \to G'$ un morphisme surjectif. On voit à l'aide de 8.3.3.2 et d'une récurrence immédiate sur n que
$$\varphi(\mathrm{D}^n(G)) = \mathrm{D}^n(G') \text{ et } \varphi(\mathrm{C}^n(G)) = \mathrm{C}^n(G')$$
pour tout n.

Ceci s'applique notamment lorsque φ est un automorphisme de G. Les sous-groupes $\mathrm{D}^n(G)$ et $\mathrm{C}^n(G)$ de G sont donc caractéristiques, et en particulier distingués.

8.3.6.4. — Si G est un produit direct de groupes $G_1 \times G_2 \times \ldots \times G_m$ il résulte du lemme 8.3.4 que $\mathrm{D}^n(G) = \prod \mathrm{D}^n(G_i)$ et $\mathrm{C}^n(G) = \prod \mathrm{C}^n(G_i)$ pour tout entier n.

8.3.7. *Abélianisé d'un groupe.* — Soit G un groupe et soit H un sous-groupe distingué de G. Le groupe $\mathrm{D}(G/H)$ est égal à l'image de $\mathrm{D}(G)$ par le morphisme quotient $G \to G/H$. Le groupe $\mathrm{D}(G/H)$ est donc trivial si et seulement si $\mathrm{D}(G)$ est contenu dans H ; autrement dit, G/H est abélien si et seulement si $\mathrm{D}(G) \subset H$. On voit ainsi que $G/\mathrm{D}(G)$ est *le plus grand quotient abélien de* G ; on l'appelle *l'abélianisé de* G. On peut y penser intuitivement comme le groupe le plus général construit à partir de G en imposant en plus la propriété d'abélianité : quotienter par $\mathrm{D}(G)$ signifie exactement qu'on force par décret les commutateurs à être triviaux. Cela se traduit rigoureusement par la propriété universelle suivante.

Lemme 8.3.8 **(propriété universelle de $G \to \mathrm{D}(G)$).** — *Soit φ un morphisme de G vers un groupe abélien G'. Il existe un unique morphisme $\psi\colon G/\mathrm{D}(G) \to G'$ tel que le diagramme*

$$\begin{tikzcd} G \arrow[rr, "\varphi"] \arrow[dr] & & G' \\ & G/D(G) \arrow[ur, "\psi"'] & \end{tikzcd}$$

commute.

Démonstration. — On a $\varphi(\mathrm{D}(G)) = \varphi([G,G]) = [\varphi(G), \varphi(G)] = \{e\}$ car G' est abélien. Cela signifie que $\mathrm{D}(G) \subset \mathrm{Ker}\varphi$, et le lemme découle donc de la propriété universelle du morphisme quotient. □

Remarque 8.3.9. — Si l'on note p le morphisme quotient $G \to D(G)$ la propriété universelle énoncée ci-dessus peut se reformuler comme suit : pour tout groupe abélien G', l'application $\psi \mapsto \psi \circ p$ établit une bijection entre l'ensemble des morphismes de G vers G' et l'ensemble des morphismes de $G/D(G)$ vers G'. En termes plus informels : se donner un morphisme de $G/D(G)$ vers G', c'est se donner un morphisme de G vers G'.

8.4. Groupes résolubles. — Nous avons maintenant suffisamment d'outils en main pour introduire la notion de groupe résoluble.

Proposition 8.4.1. — *Soit G un groupe et soit n un entier. Les assertions suivantes sont équivalentes :*

(i) *G possède une suite de composition de longueur n dont les quotients successifs sont abéliens et qui ne fait intervenir que des sous-groupes distingués dans G.*

(ii) *G possède une suite de composition de longueur n dont les quotients successifs sont abéliens.*

(iii) $\mathrm{D}^n(G) = \{e\}$.

Remarque 8.4.2. — Insistons bien sur la différence (*a posteriori* seulement apparente) entre (i) et (ii) : en général, dans une suite de composition, chaque groupe est simplement distingué *dans le précédent* ; or l'assertion (i) requiert qu'ils soient tous distingués *dans G*.

Démonstration de la proposition 8.4.1. — Il est évident que (i)⇒(ii). Supposons (ii). Le groupe G possède sous cette hypothèse une suite de composition

$$\{e\} = G_n \lhd G_{n-1} \lhd \ldots \lhd G_0 = G$$

tel que le quotient G_i/G_{i+1} soit abélien pour tout i compris entre 0 et $n-1$. On a alors $\mathrm{D}(G_i) \subset G_{i+1}$ pour tout i entre 0 et $n-1$. Par une récurrence immédiate sur n, ceci entraîne que $\mathrm{D}^i(G) \subset G_i$ pour tout i compris entre 0 et n. En particulier $\mathrm{D}^n(G) \subset G_n = \{e\}$, d'où (iii).

Supposons que (iii) est vraie. La suite

$$G = D^0(G) \rhd \mathrm{D}^1(G) \rhd \ldots \rhd \mathrm{D}^n(G) = \{e\}$$

est alors une suite de composition de G, ne faisant intervenir que des sous-groupes distingués de G. Pour tout i compris entre 0 et $n-1$ le groupe $\mathrm{D}^{i+1}(G)$ est égal à $\mathrm{D}(\mathrm{D}^i(G))$, et le quotient $\mathrm{D}^i(G)/\mathrm{D}^{i+1}(G)$ est en conséquence abélien, d'où (i). □

Définition 8.4.3. — Soit G un groupe. On dit que G est *résoluble* s'il existe un entier n tel que les propriétés équivalentes de la proposition 8.4.1 soient satisfaites ; le plus petit entier pour lequel c'est le cas est alors appelé la *classe de résolubilité de G*.

Exemples 8.4.4. — Un groupe est résoluble de classe nulle si et seulement si il est trivial. Un groupe est résoluble de classe $\leqslant 1$ si et seulement s'il est abélien. Un groupe est résoluble de classe $\leqslant 2$ si et seulement si son sous-groupe dérivé est abélien.

8.4.5. — Soit $(G_1, \ldots, G_m)$ une famille finie de groupes et soit n un entier. Le groupe $G_1 \times G_2 \ldots \times G_m$ est résoluble de classe $\leqslant n$ si et seulement si c'est le cas de chacun des G_i : c'est une conséquence immédiate de 8.3.6.4.

8.4.6. — Si un groupe G est résoluble, il possède en fait une suite de composition dont les quotients successifs sont de la forme $\mathbf{Z}/p\mathbf{Z}$ avec p premier. En effet, donnons-nous un dévissage $G = G_n \rhd G_{n-1} \rhd \ldots \rhd G_0 = \{e\}$ de G tels que les G_i/G_{i+1} soient abéliens. Pour tout $i \leqslant n-1$, le quotient G_i/G_{i+1} admet d'après le corollaire 8.1.11 une suite de composition $\mathscr{S}_i$ dont les quotients successifs sont de la forme $\mathbf{Z}/p\mathbf{Z}$ avec p premier. La concaténation des $\mathscr{S}_i$ fournit une suite de composition de G répondant à nos vœux.

Lemme 8.4.7. — *Soit G un groupe et soit H un sous-groupe de G.*

(1) *Si G est résoluble et si n désigne sa clase de résolubilité alors H est résoluble de classe $\leqslant n$.*

(2) *Supposons H distingué. Le groupe G est alors résoluble si et seulement si H et G/H le sont ; si c'est le cas, et si l'on désigne par n, r et s les classes de résolubilité de G, H et G/H on a alors*

$$r \leqslant n, \ s \leqslant n, \text{ et } n \leqslant r + s.$$

Démonstration. — Supposons G résoluble, et soit n sa classe de résolubilité. On a alors

$$\mathrm{D}^n(H) \subset \mathrm{D}^n(G) = \{e\}$$

et H est donc résoluble de classe $\leqslant n$.

On fait à partir de maintenant l'hypothèse que H est distingué dans G. Supposons que G est résoluble, et soit n sa classe de résolubilité. On a déjà vu que H est résoluble de classe $\leqslant n$. Le groupe $\mathrm{D}^n(G/H)$ est l'image du groupe $\mathrm{D}^n(G) = \{e\}$ par le morphisme quotient $G \to G/H$; il est donc trivial, ce qui montre que G/H est lui aussi résoluble de classe $\leqslant n$.

Supposons maintenant H et G/H résolubles, de classes respectives r et s. Il existe une suite de composition $\mathscr{S}$ de H de longueur r à quotients successifs abéliens, et une suite de composition $\mathscr{S}'$ de G/H de longueur s à quotients successifs abéliens. La concaténation de $\mathscr{S}$ et $\mathscr{S}'$ est une suite de composition de G de longueur $r+s$ à quotients successifs abéliens ; par conséquent, G est résoluble de classe $\leqslant r+s$. □

Exemple 8.4.8. — Nous allons maintenant démontrer que $\mathfrak{S}_4$ est résoluble de classe 3. On note K le sous-groupe $\{\mathrm{Id}, (12)(34), (13)(24), (14)(23)\}$ de $\mathfrak{A}_4$. Il est distingué dans $\mathfrak{A}_4$ (exemple 7.3.2). On vérifie immédiatement qu'il est isomorphe à $(\mathbf{Z}/2\mathbf{Z})^2$; en particulier, il est abélien. La suite

$$\mathfrak{S}_4 \rhd \mathfrak{A}_4 \rhd K \rhd \{\mathrm{Id}\}$$

est une suite de composition de $\mathfrak{S}_4$ de longueur 3. Le quotient $\mathfrak{S}_4/\mathfrak{A}_4$ est isomorphe à $\{-1,1\}$ *via* l'application signature, et est en particulier abélien. Le quotient $\mathfrak{A}_4/K$ est de cardinal 3 et donc isomorphe à $\mathbf{Z}/3\mathbf{Z}$; il est en particulier abélien. Enfin, on a vu ci-dessus que K est abélien. Les quotients successifs de la suite de composition considérée sont donc abéliens ; par conséquent, $\mathfrak{S}_4$ est résoluble de classe $\leqslant 3$.

Nous allons vérifier que sa classe de résolubilité est exactement 3. On a

$$[(12),(23)] = (12)(23)(12)(23) = \left| \begin{array}{cccc} 1 & 2 & 3 & 4 \\ 3 & 1 & 2 & 4 \end{array} \right. = (132).$$

Le sous-groupe distingué $\mathrm{D}(\mathfrak{S}_4)$ contient donc un 3-cycle. Il contient dès lors tous ses conjugués, c'est-à-dire tous les 3-cycles. Comme il y a huit 3-cycles dans $\mathfrak{S}_4$ (5.4.8), le cardinal de $D(\mathfrak{S}_4)$ est au moins 8. Par ailleurs $\mathrm{D}(\mathfrak{S}_4)$ est contenu dans $\mathfrak{A}_4$ puisque le quotient $\mathfrak{S}_4/\mathfrak{A}_4$ est abélien ; il s'ensuit que $|\mathrm{D}(\mathfrak{S}_4)|$ divise 12, et comme $|\mathrm{D}(\mathfrak{S}_4)| \geqslant 8$ on a $|\mathrm{D}(\mathfrak{S}_4)| = 12$ et $\mathrm{D}(\mathfrak{S}_4) = \mathfrak{A}_4$. Le groupe $\mathfrak{A}_4$ n'est pas abélien (vérifiez que (123) et (234) ne commutent pas, par exemple) ; par conséquent

$$\mathrm{D}^2(\mathfrak{S}_4) = \mathrm{D}(\mathfrak{A}_4) \neq \{\mathrm{Id}\}$$

et la classe de résolubilité de $\mathfrak{S}_4$ est donc strictement supérieure à 2, et partant exactement égale à 3.

8.5. Groupes nilpotents. — Nous nous proposons d'introduire et étudier les groupes *nilpotents* (qui sont des exemples de groupes résolubles). Nous utiliserons pour ce faire la suite de groupes $\mathrm{C}^n(G)$ associée à un groupe G donné (8.3.5), mais également deux autres suites que nous allons maintenant présenter.

8.5.1. — Soit G un groupe. On définit récursivement une suite $(\Lambda_n(G))$ comme suit :

- ◇ $\Lambda_0(G) = G$;
- ◇ pour tout entier $n \geqslant 1$ on a $\Lambda_n(G) = \Lambda_{n-1}(G)/\mathrm{Z}(\Lambda_{n-1}(G))$.

On a par construction pour tout entier n une suite de morphismes surjectifs

$$G \to \Lambda_1(G) \to \Lambda_2(G) \to \ldots \Lambda_n(G).$$

Leur composée est un morphisme surjectif $G \to \Lambda_n(G)$, dont le noyau est noté $\mathrm{Z}_n(G)$; on a donc $\Lambda_n(G) = G/\mathrm{Z}_n(G)$.

8.5.1.1. — La suite $(\mathrm{Z}_n(G))_n$ est par construction une suite croissante de sous-groupes distingués de G, et l'on a $\mathrm{Z}_0(G) = \{e\}$ et $\mathrm{Z}_1(G) = \mathrm{Z}(G)$. Pour tout $n \geqslant 1$, on peut écrire la surjection $G \to \Lambda_n(G)$ comme la composée $G \to \Lambda_{n-1}(G) \to \Lambda_n(G)$; son noyau est donc égal à l'image réciproque dans G du noyau de $\Lambda_{n-1}(G) \to \Lambda_n(G)$, c'est-à-dire de $\mathrm{Z}(\Lambda_{n-1}(G))$. Autrement dit, $\mathrm{Z}_n(G)$ est l'image réciproque dans G de $\mathrm{Z}(G/\mathrm{Z}_{n-1}(G))$, ce qui équivaut à l'égalité

$$\mathrm{Z}_n(G)/\mathrm{Z}_{n-1}(G) = \mathrm{Z}(G/\mathrm{Z}_{n-1}(G)).$$

Cette formule peut permettre de définir récursivement la suite $(\mathrm{Z}_n(G))_n$ (en imposant l'égalité $\mathrm{Z}_0(G) = \{e\}$), sans passer par les $\Lambda_n(G)$.

8.5.1.2. — Soit n un entier. Il résulte immédiatement des définitions que $\Lambda_n(G/\mathrm{Z}(G)) = \Lambda_{n+1}(G)$.

Proposition 8.5.2. — *Soit G un groupe et soit n un entier. Les assertions suivantes sont équivalentes :*

(i) $\Lambda_n(G) = \{e\}$.

(ii) $\mathrm{Z}_n(G) = G$.

(iii) *G possède une suite de composition $G = G_n \rhd G_{n-1} \rhd \ldots \rhd G_0 = \{e\}$ telle que $G_i \lhd G$ pour tout i et tel que G_i/G_{i-1} soit contenu dans le centre de G/G_{i-1} pour tout $i \geqslant 1$.*

(iv) $\mathrm{C}^n(G) = \{e\}$.

Démonstration. — Il est clair que (i)⇒(ii) (et même que (i) $\iff$ (ii)). Supposons que (ii) est vraie. Dans ce cas

$$G = \mathrm{Z}_n(G) \rhd \mathrm{Z}_{n-1}(G) \rhd \ldots \rhd \mathrm{Z}_0(G) = \{e\}$$

est une suite de composition de G répondant aux conditions spécifiées dans (iii).

Supposons que (iii) est vraie. Fixons i. Le groupe G_i/G_{i+1} est contenu dans le centre de G/G_{i+1} ; cela signifie que tout élément de G_i/G_{i+1} commute avec tout élément de G/G_{i+1}, soit encore que tout élément de G_i commute avec tout élément de G modulo G_{i+1}. Autrement dit, $[G, G_i] \subset G_{i+1}$. Il en résulte par une récurrence immédiate sur i que $\mathrm{C}^i(G) \subset G_i$ pour tout i. En particulier, $\mathrm{C}^n(G) \subset G_n = \{e\}$, d'où (iv).

Montrons par récurrence sur n que (iv)⇒(i). Cette implication est vraie si $n = 0$. En effet on a $\mathrm{C}^0(G) = G$ donc si $\mathrm{C}^0(G) = \{e\}$ alors $G = \{e\}$ et $\Lambda_0(G) = \{e\}$.

Supposons que $n > 0$, que l'implication est vraie pour les entiers $< n$, et que $\mathrm{C}^n(G) = \{e\}$. Comme $\mathrm{C}^n(G) = [G, \mathrm{C}^{n-1}(G)]$, tout élément de G commute avec tout élément de $\mathrm{C}^{n-1}(G)$. Autrement dit, $\mathrm{C}^{n-1}(G) \subset \mathrm{Z}(G)$. L'image de $\mathrm{C}^{n-1}(G)$ dans $G/\mathrm{Z}(G)$ est donc nulle ; par conséquent, $\mathrm{C}^{n-1}(G/\mathrm{Z}(G)) = \{e\}$. En vertu de l'hypothèse de récurrence, ceci entraîne que $\Lambda_{n-1}(G/\mathrm{Z}(G)) = \{e\}$. Puisque $\Lambda_{n-1}(G/\mathrm{Z}(G))$ est égal à $\Lambda_n(G)$, il vient $\Lambda_n(G) = \{e\}$. □

Définition 8.5.3. — Soit G un groupe. On dit que G est *nilpotent* s'il existe un entier n tel que les propriétés équivalentes de la proposition 8.5.2 soient satisfaites ; le plus petit entier pour lequel c'est le cas est alors appelé la *classe de nilpotence de G*.

Exemples 8.5.4. — Un groupe est nilpotent de classe nulle si et seulement si il est trivial. Un groupe est nilpotent de classe $\leqslant 1$ si et seulement si il est abélien. Un groupe G est nilpotent de classe $\leqslant 2$ si et seulement si $G/\mathrm{Z}(G)$ est abélien.

8.5.5. — Soit G un groupe. S'il est nilpotent de classe n, il est résoluble de classe $\leqslant n$: cela provient par exemple du fait que $\mathrm{D}^n(G) \subset \mathrm{C}^n(G)$, ou qu'une suite de décomposition comme dans la condition (iii) de la proposition 8.5.2 a tous ses quotients successifs abéliens (le centre d'un groupe est toujours abélien !).

8.5.6. — Soit G un groupe. Si $\mathrm{Z}(G) = \{e\}$ une récurrence immédiate sur n montre que $\Lambda^n(G) = G$ pour tout n ; par conséquent, G est nilpotent si et seulement s'il est trivial. Par contraposition, on voit qu'un groupe nilpotent non trivial a un centre non trivial.

8.5.7. — Soit $(G_1, \ldots, G_m)$ une famille finie de groupes et soit n un entier. Le groupe $G_1 \times G_2 \ldots \times G_m$ est nilpotent de classe $\leqslant n$ si et seulement si c'est le cas de chacun des G_i : c'est une conséquence immédiate de 8.3.6.4. On peut donner une preuve alternative de ce fait en démontrant que l'on a pour tout i l'égalité

$$\mathrm{Z}_i(G_1 \times \ldots \times G_m) = \mathrm{Z}_i(G_1) \times \ldots \times \mathrm{Z}_i(G_m).$$

L'exercice est laissé au lecteur.

Lemme 8.5.8. — *Soit G un groupe, soit H un sous-groupe de G et soit n un entier.*

(1) *Si G est nilpotent de classe n alors H est nilpotent de classe $\leqslant n$. Si de plus H est distingué, G/H est lui aussi nilpotent de classe $\leqslant n$.*

(2) *Le groupe G est nilpotent de classe n si et seulement si $G/\mathrm{Z}(G)$ est nilpotent de classe $n-1$.*

Démonstration. — Supposons G nilpotent de classe n. On a alors

$$\mathrm{C}^n(H) \subset \mathrm{C}^n(G) = \{e\}$$

et H est donc nilpotent de classe $\leqslant n$.

Supposons de plus H distingué. Le groupe $\mathrm{C}^n(G/H)$ est l'image du groupe $\mathrm{C}^n(G) = \{e\}$ par le morphisme quotient $G \to G/H$; il est donc trivial, ce qui entraîne que G/H est nilpotent de classe $\leqslant n$ et achève la démonstration de (1).

L'assertion (2) découle de l'égalité $\Lambda_n(G) = \Lambda_{n-1}(G/\mathrm{Z}(G))$. □

Proposition 8.5.9. — *Soit p un nombre premier et soit G un p-groupe. Le groupe G est nilpotent, et en particulier résoluble.*

Démonstration. — On raisonne par récurrence sur $|G|$. Le résultat est clair si $|G| = 1$. On suppose désormais $|G| > 1$, et la proposition vraie pour tout p-groupe de cardinal strictement inférieur à celui de G. Le lemme 4.4.7 assure que $\mathrm{Z}(G)$ est non trivial. Le quotient $G/\mathrm{Z}(G)$ est donc de cardinal strictement inférieur à $|G|$, et c'est évidemment un p-groupe ; il est en conséquence nilpotent d'après l'hypothèse de récurrence. L'assertion (2) du lemme 8.5.8 ci-dessus assure alors que G est nilpotent. □

Nous nous proposons maintenant de donner une caractérisation des groupes nilpotents finis ; nous avons pour cela besoin de deux lemmes. Rappelons que si H est un sous-groupe d'un groupe G, on désigne par $\mathrm{N}_G(H)$ le normalisateur de H dans G (exemple 4.3.3).

Lemme 8.5.10. — *Soit G un groupe nilpotent et soit H un sous-groupe strict de G. L'inclusion $H \subset \mathrm{N}_G(H)$ est stricte.*

Démonstration. — On procède par récurrence sur la classe de nilpotence n de G. Si $n = 0$ le groupe G est trivial et n'a pas de sous-groupe strict ; le lemme est donc vrai (et par ailleurs vide) dans ce cas. Supposons $n > 0$, et le lemme vrai pour les groupes nilpotents de classe $< n$. On distingue maintenant deux cas.

Supposons que $\mathrm{Z}(G)$ *est contenu dans* H. Le groupe $G/\mathrm{Z}(G)$ est alors nilpotent de classe $n-1$ (lemme 8.5.8) et $H/\mathrm{Z}(G)$ en est un sous-groupe strict. L'hypothèse de récurrence assure alors que $\mathrm{N}_{G/\mathrm{Z}(G)}(H/\mathrm{Z}(G))$ contient strictement $H/\mathrm{Z}(G)$. L'image réciproque K de $\mathrm{N}_{G/\mathrm{Z}(G)}(H/\mathrm{Z}(G))$ dans G contient donc strictement H. Par ailleurs, le fait que $H/\mathrm{Z}(G)$ soit distingué dans $\mathrm{N}_{G/\mathrm{Z}(G)}(H/\mathrm{Z}(G))$ assure que H est distingué dans K, et donc que $K \subset \mathrm{N}_G(H)$; en conséquence, $\mathrm{N}_G(H)$ contient *strictement* H.

Supposons que $\mathrm{Z}(G)$ *n'est pas contenu dans* H. Il existe alors $g \in \mathrm{Z}(G)$ tel que $g \notin H$. Comme $g \in \mathrm{Z}(G)$ on a $ghg^{-1} = h$ pour tout $h \in G$, et en particulier pour tout $h \in H$. Il vient $gHg^{-1} = H$, et g appartient donc à $\mathrm{N}_G(H)$; ainsi, ce dernier contient *strictement* H. □

Lemme 8.5.11. — *Soit G un groupe fini, soit p un nombre premier et soit S un p-sous-groupe de Sylow de G. On a l'égalité* $\mathrm{N}_G(\mathrm{N}_G(S)) = \mathrm{N}_G(S)$.

Démonstration. — Par définition, S est distingué dans $\mathrm{N}_G(S)$. Comme S est un p-sous-groupe de Sylow de G, c'est *a fortiori* un p-sous-groupe de Sylow de $\mathrm{N}_G(S)$. Comme il est distingué dans $\mathrm{N}_G(S)$, c'est son seul p-sous-groupe de Sylow et c'est donc un sous-groupe caractéristique de $\mathrm{N}_G(S)$.

Soit g un élément de $\mathrm{N}_G(\mathrm{N}_G(S))$. Puisque $\mathrm{N}_G(S)$ est distingué dans $\mathrm{N}_G(\mathrm{N}_G(S))$, la conjugaison par g induit un automorphisme φ de $\mathrm{N}_G(S)$. Comme S est un sous-groupe caractéristique de ce dernier, on a $\varphi(S) = S$; autrement dit, $gSg^{-1} = S$ et $g \in \mathrm{N}_G(S)$. □

Théorème 8.5.12. — *Soit G un groupe fini. Les assertions suivantes sont équivalentes.*

(i) *Le groupe G est nilpotent.*

(ii) *Pour tout nombre premier p, le groupe G admet un unique p-sous-groupe de Sylow.*

(iii) *Il existe une famille finie* $(p_1, \ldots, p_r)$ *de nombres premiers deux à deux distincts, et pour tout* i *compris entre* 1 *et* r *un* p_i*-groupe non trivial* G_i*, tels que* G *soit isomorphe à* $G_1 \times G_2 \times \ldots \times G_r$.

Démonstration. — On suppose que (i) est vraie. Soit p un nombre premier. Soit S un p-sous-groupe de Sylow de G. Le lemme 8.5.11 assure que $\mathrm{N}_G(\mathrm{N}_G(S)) = \mathrm{N}_G(S)$. Il en résulte en vertu du lemme 8.5.10 que $\mathrm{N}_G(S)$ ne peut être un sous-groupe strict du groupe nilpotent G ; par conséquent, $\mathrm{N}_G(S) = G$, ce qui veut dire que S est distingué dans G. C'est donc l'unique p-sous-groupe de Sylow de G, et (ii) est vraie.

On suppose que (ii) est vraie. Soient $p_1, \ldots, p_r$ les diviseurs premiers deux à deux distincts de $|G|$; pour tout i, on note G_i l'unique p_i-sous-groupe de Sylow de G, qui est un p_i-groupe non trivial distingué dans G. On a $|G| = \prod_i |G_i|$. Soient i et j deux entiers distincts compris entre 1 et r. Soit $g \in G_i$ et soit $h \in G_j$. Comme G_i est distingué dans G, le commutateur

$$[g, h] = ghg^{-1}h^{-1} = (ghg^{-1})h^{-1} = g(hg^{-1}h^{-1})$$

appartient à $G_j \cap G_j$; son ordre divise donc à la fois $|G_i|$ et $|G_j|$, qui sont premiers entre eux ; il s'ensuit que $[g, h] = e$, c'est-à-dire que g et h commutent. On note φ l'application du produit $G_1 \times G_2 \times \ldots G_r$ dans G qui envoie un r-uplet $(g_1, \ldots, g_r)$ sur le produit $g_1 g_2 \ldots g_r$. On a vu ci-dessus que si i et j sont deux entiers distincts compris entre 1 et r, tout élément de G_i commute avec tout élément de G_j. Il en résulte que φ est un morphisme de groupes. *Montrons que* φ *est surjectif.* Fixons i. Pour tout g appartenant à G_i, on a $g = \varphi(e, \ldots, e, g, e, \ldots, e)$ (où g est évidemment placé au rang i) ; par conséquent, $G_i \subset \mathrm{Im}\,\varphi$, et $|\mathrm{Im}\,\varphi|$ est donc multiple de $|G_i|$. Ceci valant pour tout i, le cardinal de $\mathrm{Im}\,\varphi$ est multiple du PPCM des cardinaux des G_i, qui est égal à $|G|$. Il en résulte que $\mathrm{Im}\,\varphi = G$, et φ est bien surjective. Comme G et $G_1 \times G_2 \times \ldots \times G_r$ ont même cardinal, le morphisme φ est également injectif, et G est isomorphe à $G_1 \times G_2 \times \ldots \times G_r$; ainsi, (iii) est vraie.

Supposons que (iii) est vraie. La proposition 8.5.9 assure que chacun des G_i est nilpotent ; le groupe G est donc nilpotent d'après 8.5.7 et (i) est vraie. □

Remarque 8.5.13. — On peut démontrer directement que (iii)⇒(ii). En effet, soient $p_1, \ldots, p_r$ des nombres premiers deux à deux distincts, et pour tout i compris entre 1 et r, soit G_i un p_i-groupe non trivial. Le cardinal du groupe produit $G := G_1 \times G_2 \times \ldots \times G_r$ est égal à $\prod |G_i|$. Par conséquent, si p est un nombre premier n'appartenant pas à $\{p_1, \ldots, p_r\}$ alors p ne divise pas $|G|$, et G a un unique p-sous-groupe de Sylow, à savoir $\{e\}$. Soit maintenant i compris entre 1 et r et soit Γ_i le sous-groupe de G constitué des éléments (g_j) tels que $g_j = e$ pour tout $j \neq i$. Il est immédiat que Γ_i est isomorphe à G_i ; c'est par conséquent un p_i-sous-groupe de Sylow de G. Il est visiblement distingué dans G, et est donc son unique p_i-sous-groupe de Sylow.

8.6. Une suite de composition matricielle. — On fixe un corps k et un entier $n \geqslant 1$. On note T le sous-groupe de $\mathrm{GL}_n(k)$ formé des matrices triangulaires supérieures, et U le sous-groupe de T constitué des matrices triangulaires supérieures n'ayant que des 1 sur la diagonale. On note $(e_1, \ldots, e_n)$ la base canonique de k^n ; nous identifierons un élément de $\mathrm{GL}_n(k)$ à l'endomorphisme de k^n dont il est la matrice dans $(e_1, \ldots, e_n)$.

8.6.1. — Pour tout entier $\ell \geqslant 1$ on note N_ℓ le sous-espace vectoriel de $\mathrm{M}_n(k)$ formé des matrices A telles que $Ae_i \in \mathrm{Vect}(e_j)_{1 \leqslant j \leqslant i-\ell}$.

8.6.1.1. — Les faits suivants sont des conséquences immédiates de la définition, que nous utiliserons librement dans la suite.

- Le sous-espace N_1 de $\mathrm{M}_n(k)$ est l'ensemble des matrices triangulaires supérieures n'ayant que des zéros sur la diagonale (on peut donc écrire $\mathrm{U} = \mathrm{I}_n + \mathrm{N}_1$) ; plus généralement, N_ℓ est l'ensemble des matrices dont tous les coefficients situés en-dessous de la «ℓ-ième surdiagonale» sont nuls.
- On a $\mathrm{N}_\ell \subset \mathrm{N}_{\ell'}$ dès que $\ell \geqslant \ell'$.
- On a $\mathrm{N}_\ell = 0$ dès que $\ell \geqslant n$.
- On a $\mathrm{N}_\ell \cdot \mathrm{N}_{\ell'} \subset \mathrm{N}_{\ell+\ell'}$ pour tout (ℓ, ℓ'). Il s'ensuit que $\mathrm{N}_\ell^m \subset \mathrm{N}_{\ell m}$ pour tout $m \geqslant 1$ puis que
$$\mathrm{N}_\ell^{\lceil \frac{n}{\ell} \rceil} \subset \mathrm{N}_n = \{0\}$$
où $\lceil \frac{n}{\ell} \rceil$ désigne la «partie entière supérieure» de n/ℓ, c'est-à-dire le plus petit entier supérieur ou égal à n/ℓ. En particulier, $\mathrm{N}_1^n = \{0\}$.

8.6.1.2. — Soit ℓ un entier $\geqslant 1$ et soient A et B deux éléments de N_ℓ. Comme A est nilpotente, la matrice $\mathrm{I}_n + A$ est inversible d'inverse $\sum_i (-1)^{i-1} A^i$ (la nilpotence de A garantit que cette somme est finie) ; notons que $(\mathrm{I}_n + A)^{-1} \in \mathrm{I}_n + \mathrm{N}_\ell$ puisque $\mathrm{N}_\ell^i \subset \mathrm{N}_{i\ell} \subset \mathrm{N}_\ell$ pour tout $i \geqslant 1$. Et on a par ailleurs
$$(\mathrm{I}_n + A)(\mathrm{I}_n + B) = \mathrm{I}_n + A + B + AB \in \mathrm{I}_n + \mathrm{N}_\ell,$$
la dernière inclusion provenant du fait que $AB \in \mathrm{N}_{2\ell} \subset \mathrm{N}_\ell$.

On déduit de ce qui précède (et du fait que $\mathrm{I}_n \in \mathrm{I}_n + \mathrm{N}_\ell$) que $\mathrm{I}_n + \mathrm{N}_\ell$ est un sous-groupe de $\mathrm{U} = \mathrm{I}_n + \mathrm{N}_1$.

8.6.1.3. — Soient ℓ et ℓ' deux entiers. Nous allons démontrer que
$$[(\mathrm{I}_n + \mathrm{N}_\ell), (\mathrm{I}_n + \mathrm{N}_{\ell'})] \subset \mathrm{I}_n + \mathrm{N}_{\ell+\ell'}.$$
Soit $A \in \mathrm{N}_\ell$ et soit $B \in \mathrm{N}_{\ell'}$. Vérifions que le commutateur
$$(\mathrm{I}_n + A)(\mathrm{I}_n + B)(\mathrm{I}_n + A)^{-1}(\mathrm{I}_n + B)^{-1}$$
appartient à $\mathrm{I}_n + \mathrm{N}_{\ell+\ell'}$. En écrivant
$$(\mathrm{I}_n + A)^{-1} = \sum_i (-1)^{i-1} A^i \text{ et } (\mathrm{I}_n + B)^{-1} = \sum_i (-1)^{i-1} B^i$$
et en développant le commutateur ci-dessus, on voit que celui-ci s'écrit comme un «polynôme non commutatif» en A et B, qui comprend différents types de monômes :

- un «terme constant» égal à I_n ;
- des monômes faisant intervenir au moins une fois A et au moins une fois B ; un tel monôme appartient toujours à $\mathrm{N}_{\ell+\ell'}$;

- ◇ des monômes de la forme A^j avec $j > 0$; ce sont plus précisément les monômes non constants qui apparaissent dans le développement du produit $(\mathrm{I}_n + A)(\sum_i (-1)^{i-1} A^i) = \mathrm{I}_n$, et *il n'y en a donc pas* ;
- ◇ des monômes de la forme B^j avec $j > 0$; ce sont plus précisément les monômes non constants qui apparaissent dans le développement du produit $(\mathrm{I}_n + B)(\sum_i (-1)^{i-1} B^i) = \mathrm{I}_n$, et *il n'y en a donc pas*.

Par conséquent, on peut écrire

$$(\mathrm{I}_n + A)(\mathrm{I}_n + B)(\mathrm{I}_n + A)^{-1}(\mathrm{I}_n + B)^{-1} = \mathrm{I}_n + C$$

où C est un élément de $\mathrm{N}_{\ell+\ell'}$, ce qu'on souhaitait établir.

8.6.1.4. — Notons une première conséquence importante de 8.6.1.3 : *pour tout ℓ, le groupe $\mathrm{I}_n + \mathrm{N}_\ell$ est distingué dans* $\mathrm{U} = \mathrm{I}_n + \mathrm{N}_1$. En effet, fixons ℓ et donnons-nous $U \in \mathrm{U}$ et $V \in \mathrm{I}_n + \mathrm{N}_\ell$. On déduit de 8.6.1.3 que $UVU^{-1}V^{-1} \in \mathrm{I}_n + \mathrm{N}_{\ell+1}$. Il vient

$$UVU^{-1} = (UVU^{-1}V^{-1})V \in (\mathrm{I}_n + \mathrm{N}_{\ell+1})(\mathrm{I}_n + \mathrm{N}_\ell) \subset (\mathrm{I}_n + \mathrm{N}_\ell) \cdot (\mathrm{I}_n + \mathrm{N}_\ell) \subset (\mathrm{I}_n + \mathrm{N}_\ell),$$

d'où notre assertion.

8.6.1.5. — Soit $\ell \geqslant 1$. En vertu de 8.6.1.3, on a

$$[\mathrm{U}, (\mathrm{I}_n + \mathrm{N}_\ell)] = [(\mathrm{I}_n + \mathrm{N}_1), (\mathrm{I}_n + \mathrm{N}_\ell)] \subset \mathrm{I}_n + \mathrm{N}_{\ell+1},$$

ce qui signifie que le groupe quotient $(\mathrm{I}_n + \mathrm{N}_\ell)/(\mathrm{I}_n + \mathrm{N}_{\ell+1})$ est contenu dans le centre de $\mathrm{U}/(\mathrm{I}_n + \mathrm{N}_{\ell+1})$. La suite de composition

$$\mathrm{U} = (\mathrm{I}_n + \mathrm{N}_1) \rhd (\mathrm{I}_n + \mathrm{N}_2) \rhd \ldots \rhd (\mathrm{I}_n + \mathrm{N}_{n-1}) \rhd (\mathrm{I}_n + \mathrm{N}_n) = \{\mathrm{I}_n\}$$

est donc du type décrit à la condition (iii) de l'énoncé de la proposition 8.5.2. Il s'ensuit que le groupe U est nilpotent.

8.6.1.6. — Nous allons maintenant décrire les quotients successifs de la suite de composition construite au 8.6.1.5. Fixons ℓ entre 1 et $n - 1$. Une matrice (a_{ij}) de $\mathrm{M}_n(k)$ appartient à $\mathrm{I}_n + \mathrm{N}_\ell$ si et seulement si les conditions suivantes sont satisfaites :

- ◇ $a_{ij} = 0$ dès que $j < i$ ou $i < j < i + \ell$;
- ◇ $a_{ii} = 1$ pour tout i.

Donnons-nous deux matrices $A = (a_{ij})$ et $B = (b_{ij})$ appartenant à $\mathrm{I}_n + \mathrm{N}_\ell$, et notons (c_{ij}) la matrice produit AB. Pour tout i compris entre 1 et $n - \ell$ on a

$$c_{i,i+\ell} = \sum_\lambda a_{i\lambda} b_{\lambda,i+\ell} = b_{i,i+\ell} + a_{i,i+\ell}$$

car $a_{i\lambda=0}$ si $\lambda < i$ ou $i < \lambda < i + \ell$ et $b_{\lambda,i+\ell} = 0$ si $\lambda > i + \ell$.

L'application de $\varphi_\ell \colon \mathrm{I}_n + \mathrm{N}_\ell \to k^{n-\ell}$ qui envoie une matrice (a_{ij}) sur $(a_{i,i+\ell})_{1 \leqslant i \leqslant n-\ell}$ est donc un morphisme de groupes. Le morphisme φ_ℓ est surjectif : si $(\alpha_i)_{1 \leqslant i \leqslant n-\ell}$ est un élément de $k^{n-\ell}$ la matrice de terme général (β_{ij}) avec $\beta_{ij} = 1$ si $i = j$, $\beta_{ij} = \alpha_i$ si $j = i + \ell$ et $\beta_{ij} = 0$ sinon, est un antécédent de (α_i) par φ. Le noyau de φ_ℓ est l'ensemble des matrices $(a_{ij}) \in \mathrm{I}_n + \mathrm{N}_\ell$ telles que $a_{i,i+\ell} = 0$ pour tout i compris entre 1 et $n - \ell$; c'est donc exactement $\mathrm{I}_n + \mathrm{N}_{\ell+1}$.

Par conséquent, φ_ℓ induit un isomorphisme

$$(\mathrm{I}_n + \mathrm{N}_\ell)/(\mathrm{I}_n + \mathrm{N}_{\ell+1}) \simeq k^{n-\ell}.$$

8.6.1.7. — Le groupe nilpotent U possède donc une suite de composition dont la liste des quotients successifs est $\{k^N\}_{n-1\geqslant N\geqslant 1}$. Fixons N. Le groupe k^N admet lui-même une suite de décomposition de longueur N dont les quotients successifs sont tous isomorphes à k, comme par exemple $G_0 = k^N \rhd G_1 \rhd \ldots \rhd G_N = \{0\}$, où G_i désigne pour tout i l'ensemble des éléments de k^N dont les i premières composantes sont nulles.

Par concaténation, il s'ensuit que U possède une suite de composition de longueur $1+2+\ldots+n-1 = \frac{n(n-1)}{2}$ dont tous les quotients successifs sont isomorphes à k.

8.6.2. — Soit ψ l'application de T dans $(k^\times)^n$ qui envoie une matrice (a_{ij}) sur $(a_{ii})_{1\leqslant i\leqslant n}$. C'est clairement un morphisme de groupes surjectif de noyau U. Ce dernier est donc distingué dans T, et le quotient T/U est isomorphe à $(k^\times)^n$. Le groupe $(k^\times)^n$ admet une suite de composition de longueur n dont tous les quotients successifs sont isomorphes à $k^\times$: il suffit par exemple de prendre $H_0 = (k^\times)^n \rhd H_1 \rhd \ldots \rhd H_n = \{1\}$, où H_i désigne pour tout i l'ensemble des éléments de $(k^\times)^n$ dont les i premières composantes sont égales à 1.

Par concaténation de cette suite de composition et de celle exhibée au 8.6.1.7 on obtient une suite de composition de T de longueur $n + \frac{n(n-1)}{2} = \frac{n(n+1)}{2}$ dont les n premiers quotients sont isomorphes à $k^\times$ et les $\frac{n(n-1)}{2}$ suivants à k. Tous ces groupes étant abéliens, T est résoluble.

9. Groupes libres, groupes définis par générateurs et relations

9.1. Groupe libre sur un ensemble. — Soit X un ensemble. Le but de ce qui suit est de construire le «groupe libre sur l'ensemble X». C'est informellement le groupe le plus général qu'on peut fabriquer à partir de X ; il est obtenu en décrétant qu'on sait multiplier et inverser les éléments de X, et en n'imposant à ces opérations aucune autre règle que celles données par la théorie générale des groupes.

Bien entendu, cette description est vague et non rigoureuse, et nous allons maintenant procéder à la construction détaillée de ce groupe. Nous montrerons qu'il est caractérisé par une propriété universelle, comme tout «objet le plus général tel que...» qui se respecte.

Cette contruction requiert pour des raisons techniques un certain nombre de contorsions. Nous allons notamment devoir introduire la notion de *monoïde* (qui ne nous servira pas ailleurs) puis celle de *monoïde libre*, avant d'en arriver enfin à celle de groupe libre.

Définition 9.1.1. — Un *monoïde* est un ensemble M muni d'une loi de composition interne associative et possédant un élément neutre e (nécessairement unique, par le même raisonnement qu'au 2.1.3).

Si M et N sont deux monoïdes de neutres respectifs e_M et e_N, un *morphisme de monoïdes* de M dans N est une application f de M vers N telle que l'on ait $f(e_M) = e_N$ et $f(ab) = f(a)f(b)$ pour tout $(a, b) \in M^2$.

Exemple 9.1.2. — Tout groupe est en particulier un monoïde. Plus précisément, un groupe est par définition un monoïde dans lequels tout élément a un inverse.

Exemple 9.1.3. — Si A est un anneau alors $(A, \times)$ est un monoïde ; ce n'est pas un groupe si $A \neq \{0\}$, car 0 n'a alors pas d'inverse.

Remarquons que l'application nulle de A dans A commute au produit mais n'est pas un morphisme de monoïdes si $A \neq \{0\}$ car elle n'envoie pas 1 sur 1. Contrairement à ce qui se passe pour les groupes, il est donc indispensable d'imposer dans la définition d'un morphisme de monoïdes que l'élément neutre soit envoyé sur l'élément neutre. (Rappelons que le caractère automatique de cette propriété dans le cas des groupes résulte de la *simplification des égalités* (2.1.6) qui n'existe pas dans un monoïde en général : si $A \neq \{0\}$ elle est justement en défaut dans $(A, \times)$ à cause de 0).

Exemple 9.1.4. — L'ensemble $\mathbf{N}$ muni de l'addition est un monoïde. Ce n'est pas un groupe : 1 n'a pas d'opposé.

Définition 9.1.5. — Soit X un ensemble. Un *mot sur l'alphabet* X est une suite finie $x_1 \ldots x_n$ d'éléments de X ; l'entier n est appelé la *longueur* du mot en question. Il existe un et un seul mot de longueur nulle sur l'alphabet X : c'est la suite vide, appelée également *mot vide* et notée $\varnothing$.

Soit $\Lambda(X)$ l'ensemble des mots sur l'alphabet X. La concaténation définit une loi de composition interne sur $\Lambda(X)$; elle est associative, et possède un élément neutre : le mot vide. Elle fait donc de $\Lambda(X)$ un monoïde, appelé le *monoïde libre* sur l'ensemble X ; nous identifierons dans ce qui suit X à un sous-ensemble de $\Lambda(X)$, en voyant un élément de X comme un mot de longueur 1.

Lemme 9.1.6 **(propriété universelle du monoïde libre)**
Soit X un ensemble, soit M un monoïde et soit $f\colon X \to M$ une application ensembliste. Il existe un unique morphisme de monoïdes de $\Lambda(X)$ dans M qui prolonge f.

Démonstration. — Un tel morphisme est nécessairement donné par la formule

$$x_1x_2\dots x_n \mapsto f(x_1)f(x_2)\dots f(x_n).$$

Réciproquement, il est immédiat que cette formule définit un morphisme de monoïdes de $\Lambda(X)$ dans M, qui prolonge f. □

9.1.7. *Construction du groupe libre.* — Soit X un ensemble. On introduit un ensemble X^{-1}, disjoint de X et muni d'une bijection $x \mapsto x^{-1}$ de X sur X^{-1} (attention : X^{-1} et x^{-1} sont de simples *notations*). Pour tout groupe G, on note $h(G)$ l'ensemble des morphismes (de monoïdes) f de $\Lambda(X\coprod X^{-1})$ dans G tels que $f(x^{-1}) = f(x)^{-1}$ pour tout $x \in X$. Soit $\mathscr{R}$ la relation sur $\Lambda(X\coprod X^{-1})$ telle que $m\mathscr{R}n$ si et seulement si pour tout groupe G et tout $f \in h(G)$ on a $f(m) = f(n)$.

9.1.7.1. — On vérifie aussitôt que $\mathscr{R}$ est une relation d'équivalence, et l'on note $F(X)$ le quotient $\Lambda(X\coprod X^{-1})/\mathscr{R}$. Soient m, n, m' et n' des éléments de M tels que $m\mathscr{R}n$ et $m'\mathscr{R}n'$. Soit G un groupe et soit $f \in h(G)$. Comme $m\mathscr{R}n$ et $m'\mathscr{R}n'$ on a $f(m) = f(n)$ et $f(m') = f(n')$. Il vient $f(mm') = f(m)f(m') = f(n)f(n') = f(nn')$. Ainsi, $mm'\mathscr{R}nn'$. Il s'ensuit que la loi interne de $\Lambda(X\coprod X^{-1})$ passe au quotient ;par $\mathscr{R}$ et induit une loi interne sur $F(X)$; nous vous laissons vérifier que celle-ci fait de $F(X)$ un monoïde, et que l'application quotient $\Lambda(X\coprod X^{-1}) \to F(X)$ est un morphisme de monoïdes (il suffit essentiellement de reprendre la preuve du lemme 2.10.1 en se limitant à ce qui concerne l'associativité et l'élément neutre).

9.1.7.2. *Le monoïde $F(X)$ est un groupe.* — Il s'agit de vérifier que chacun de ses éléments est inversible.

Tout élément de $F(X)$ est de la forme $\overline{x_1\dots x_k} = \overline{x_1}\dots\overline{x_k}$ où les x_i appartiennent à $X\coprod X^{-1}$. Il suffit donc de vérifier que $\overline{x}$ est inversible pour tout $x \in X\coprod X^{-1}$. Nous allons montrer que si $x \in X$ alors $\overline{x}$ est inversible d'inverse $\overline{x^{-1}}$, ce qui permettra de conclure.

Soit $x \in X$, soit G un groupe et soit $f \in h(G)$. On a $f(x^{-1}) = f(x)^{-1}$, et donc $f(xx^{-1}) = f(x^{-1}x) = e = f(\varnothing)$. Par conséquent, $\overline{xx^{-1}} = \overline{x^{-1}x} = \overline{\varnothing}$, ce qui montre que $\overline{x}$ est inversible d'inverse $\overline{x^{-1}}$, comme annoncé.

Lemme 9.1.8 **(propriété universelle du groupe $F(X)$)**
Soit X un ensemble, soit G un groupe et soit $f : X \to G$ une application. Il existe un unique morphisme de groupes φ de $F(X)$ vers G qui envoie $\overline{x}$ sur $f(x)$ pour tout $x \in X$.

Démonstration. — Commençons par l'unicité. Soit φ un morphisme satisfaisant les propriétés de l'énoncé. Comme $\overline{x}^{-1} = \overline{x^{-1}}$ pour tout $x \in X$ d'après 9.1.7.2, et comme tout élément de $F(X)$ est de la forme $\overline{x_1\dots x_k} = \overline{x_1}\dots\overline{x_k}$ où les x_i appartiennent à $X\coprod X^{-1}$, on voit que $F(X)$ est engendré en tant que groupe par l'ensemble des $\overline{x}$ pour $x \in X$. Par conséquent, φ est entièrement déterminé par sa restriction à cet ensemble, laquelle est imposée par hypothèse (puisque $\varphi(\overline{x}) = f(x)$ pour tout $x \in X$) ; ainsi, φ est unique.

Prouvons maintenant l'existence de φ. Soit g l'application de $X \coprod X^{-1}$ dans G qui envoie x sur $f(x)$ et x^{-1} sur $f(x)^{-1}$ pour tout $x \in X$. Par la propriété universelle du monoïde $\Lambda(X)$ (lemme 9.1.6), l'application g se prolonge en un morphisme de monoïdes $\psi : \Lambda(X) \to G$, qui appartient par construction à $h(G)$. Par conséquent, $\psi(m) = \psi(n)$ dès que $m\mathscr{R}n$, et ψ induit ainsi par passage au quotient une application $\varphi\colon F(X) \to G$, qui envoie par construction $\overline{x}$ sur $f(x)$ pour tout $x \in X$. Nous laissons le lecteur vérifier qu'il s'agit d'un morphisme de groupes. □

Le groupe $F(X)$ a été défini comme le quotient de $\Lambda(X \coprod X^{-1})$ par une relation d'équivalence *a priori* peu explicite : elle est donnée par des conditions portant sur *tous* les morphismes de monoïdes de source $\Lambda(X)$ dont le but est un groupe. Nous allons voir qu'il est néanmoins possible de décrire $F(X)$ de manière tangible. Pour ce faire, nous allons avoir besoin de la définition suivante.

Définition 9.1.9. — Soit X un ensemble. Un mot $m \in \Lambda(X \coprod X^{-1})$ est dit *réduit* s'il ne contient aucune suite de deux termes consécutifs de la forme xx^{-1} ou $x^{-1}x$ avec $x \in X$.

Théorème 9.1.10. — *Soit X un ensemble et soit $\mathscr{R}$ la relation d'équivalence sur $\Lambda(X \coprod X^{-1})$ définie au 9.1.7. Toute classe de $\mathscr{R}$ contient un unique mot réduit.*

Démonstration. — Nous allons traiter séparément l'existence (qui est facile) et l'unicité (qui est plus délicate).

Commençons par l'existence. Soit $m \in \Lambda(X)$. Nous allons montrer par récurrence sur la longueur de m l'existence d'un mot réduit équivalent à m. Si la longueur de m est nulle, m est le mot vide et est déjà réduit. Supposons que la longueur de m est strictement positive et que le résultat est vrai pour les mots de longueur strictement inférieure. Si m est réduit, il n'y a rien à faire. Sinon, m est de la forme $m'xx^{-1}m''$ ou $m'x^{-1}xm''$; par l'hypothèse de récurrence, $m'm''$ est équivalent à un mot réduit (sa longueur est strictement inférieure à celle de m). Il suffit maintenant de montrer que m est équivalent à $m'm''$. Supposons par exemple que $m = m'xx^{-1}m''$. On a

$$\overline{m} = \overline{m'} \cdot \overline{x} \cdot \overline{x^{-1}} \cdot \overline{m''} = \overline{m} \cdot \overline{m''} = \overline{mm''},$$

puisque $\overline{x}$ et $\overline{x^{-1}}$ sont inverses l'un de l'autre. Par conséquent, $m\mathscr{R}m'm''$, ce qu'il fallait démontrer ; la preuve dans le cas où $m = m'x^{-1}xm''$ est analogue.

Nous allons maintenant nous assurer que deux mots réduits équivalents coïncident. Soit E l'ensemble des mots réduits. Pour tout $x \in X$, soit σ_x l'application de E dans E qui envoie un mot réduit m sur xm si m n'est pas de la forme $x^{-1}m'$, et sur m' si m est de la forme $x^{-1}m'$ (nous laissons le lecteur s'assurer que les mots obtenus sont bien réduits). C'est une bijection : on vérifie aisément que sa réciproque envoie un mot réduit m sur $x^{-1}m$ si m n'est pas de la forme xm', et sur m' si m est de la forme xm'. Cette application ensembliste de X dans $\mathfrak{S}_E$ induit en vertu de la propriété universelle de $F(X)$ un morphisme de groupes de $F(X)$ vers $\mathfrak{S}_E$, c'est-à-dire une action $(\mu, m) \mapsto \mu \cdot m$ de $F(X)$ sur E.

Soit m un mot réduit. Montrons que $\overline{m} \cdot \varnothing = m$. On le vérifie par récurrence sur la longueur de m. Si m est de longueur nulle c'est le mot vide et $\overline{m}$ est donc l'élément neutre de $F(X)$, qui agit trivialement sur E ; l'assertion en découle. Supposons m de longueur strictement positive, et la propriété vraie pour les mots de longueur strictement inférieure à celle de m. On peut écrire $m = xm'$ avec $x \in X \coprod X^{-1}$. Comme m est réduit, m' est réduit. On a l'égalité

$$\overline{m} \cdot \varnothing = \overline{x} \cdot (\overline{m'} \cdot \varnothing).$$

Par hypothèse de récurrence, $\overline{m'} \cdot \varnothing$ est égal à m'. Si x appartient à X alors comme m est réduit, m' n'est pas de la forme $x^{-1}m''$, et l'on a donc

$$\overline{x} \cdot m' = \sigma_x(m') = xm' = m.$$

Si $x = y^{-1}$ pour un certain $y \in X$ alors comme m est réduit, m' n'est pas de la forme ym'', et l'on a donc

$$\overline{x} \cdot m' = \sigma_y^{-1}(m') = y^{-1}m' = m.$$

Conclusion. Si m et n sont deux mots réduits tels que $m\mathscr{R}n$, on a $\overline{m} = \overline{n}$ et donc $m = \overline{m} \cdot \varnothing = \overline{n} \cdot \varnothing = n$, ce qui termine la démonstration. □

Commentaires 9.1.11. — Le théorème 9.1.10 ci-dessus signifie que le passage au quotient par $\mathscr{R}$ permet d'identifier $F(X)$ à l'ensemble des mots *réduits* sur l'alphabet $X \coprod X^{-1}$ (en particulier, on peut voir $X \coprod X^{-1}$ comme un sous-ensemble de $F(X)$). Pour faire le produit de deux éléments de $F(X)$, on les concatène, puis on simplifie le mot obtenu en éliminant tous les termes de la forme xx^{-1} ou $x^{-1}x$, et l'on recommence jusqu'à obtention d'un mot réduit.

Signalons par ailleurs que si $n > 1$ on écrira souvent x^n (resp. x^{-n}) à la place d'une chaîne de n termes x (resp. x^{-1}) consécutifs.

Exemple 9.1.12. — Supposons que $X = \{a, b, c, d\}$. Les deux mots réduits

$$m = a^2b^{-1}c^3dada \text{ et } n = a^{-1}d^{-1}a^{-1}d^{-1}b^2ca^4$$

sont alors deux éléments de $F(X)$. La concaténation des deux mots est égale à

$$a^2b^{-1}c^3dadaa^{-1}d^{-1}a^{-1}d^{-1}b^2ca^4.$$

En quatre étapes (élimination de aa^{-1}, puis dd^{-1}, puis aa^{-1}, puis dd^{-1}), on obtient le mot réduit $a^2b^{-1}c^3b^2ca^4$, qui est donc le produit de m et n.

Remarque 9.1.13. — Vous vous demandez peut-être pourquoi on a construit $F(X)$ de façon passablement alambiquée (comme quotient du monoïde libre $\Lambda(X \coprod X^{-1})$ par une relation d'équivalence semblant de prime abord inexploitable) au lieu de le *définir* comme l'ensemble des mots réduits sur l'alphabet $X \coprod X^{-1}$, avec la concaténation-simplification comme loi interne. Pour le comprendre, essayez de démontrer directement l'associativité de cette loi ...

Remarque 9.1.14. — Nous nous autoriserons à considérer tout mot sur l'aphabet $X \coprod X^{-1}$ comme un élément de $F(X)$ même s'il n'est pas réduit, en l'identifiant à son image par l'application quotient ; autrement dit, nous omettrons désormais la barre de réduction modulo $\mathscr{R}$.

9.1.15. *À propos de la propriété universelle du groupe libre.* — Soit X un ensemble et soit ι l'inclusion de X dans $F(X)$.

9.1.15.1. — La propriété universelle de $F(X)$ (lemme 9.1.8) est en fait plus précisément une propriété universelle de l'application $\iota\colon X \to F(X)$, et peut se formuler ainsi : pour tout groupe G, l'application $\varphi \mapsto \varphi|_X = \varphi \circ \iota$ établit une bijection entre l'ensemble des morphismes de groupes de $F(X)$ dans G et l'ensemble des applications de X dans G.

Ce qu'on peut traduire d'une manière un peu plus informelle comme suit : se donner un morphisme de $F(X)$ dans un groupe G revient à se donner une application de X dans G, ou encore à choisir librement les images des éléments de X.

9.1.15.2. — L'application $\iota\colon X \to F(X)$ est caractérisée à unique isomorphisme près par sa propriété universelle, dans le sens suivant. Soit υ une application de X vers un groupe Γ telle que pour tout groupe G, l'application $\varphi \mapsto \varphi \circ \upsilon$ établisse une bijection entre l'ensemble des morphismes de groupes de Γ dans G et l'ensemble des applications de X dans G. Il existe alors un unique isomorphisme $f\colon F(X) \to \Gamma$ tel que le diagramme

$$\begin{tikzcd} F(X) \arrow[rr, "f"] & & \Gamma \\ & X \arrow[ul, "\iota"] \arrow[ur, "\upsilon"'] & \end{tikzcd}$$

commute.

En effet, la propriété universelle de ι assure qu'il existe un morphisme f de $F(X)$ dans Γ faisant commuter le diagramme ci-dessus. La propriété universelle de υ assure qu'il existe un morphisme g de Γ dans $F(X)$ tel que $g \circ \upsilon = \iota$. La composée $g \circ f$ est une application de $F(X)$ dans lui-même, et l'on a

$$(g \circ f) \circ \iota = g \circ \varepsilon = \iota = \mathrm{Id}_{F(X)} \circ \iota.$$

Par la partie «unicité» (ou «injectivité») de la propriété universelle de ι on a nécessairement $g \circ f = \mathrm{Id}_{F(X)}$. On montre de même que $f \circ g = \mathrm{Id}_\Gamma$.

9.1.16. *Évaluation des mots.* — Soit$X$ un ensemble, soit G un groupe et soit $(g_x)_{x\in X}$ une famille d'éléments de G. Soit $\varphi\colon F(X) \to G$ l'unique morphisme de groupes tel que $\varphi(x) = g_x$ pour tout $x \in X$. Soit m un mot sur l'alphabet $X \coprod X^{-1}$. On notera souvent $m(g_x)_x$ l'élément $\varphi(m)$ de G – ici, m est vu comme appartenant à $F(X)$, *cf.* 9.1.14 ; Si $m = x_1^{\varepsilon_1} \dots x_n^{\varepsilon_n}$ avec $\varepsilon_i \in \{-1, 1\}$ pour tout i, on a $m(g_x) = g_{x_1}^{\varepsilon_1} \dots g_{x_n}^{\varepsilon_n}$. On dira que φ est *le morphisme d'évaluation en la famille* (g_x).

Exemple 9.1.17. — Le seul mot sur un alphabet vide étant le mot vide, le groupe $F(\varnothing)$ est trivial.

Exemple 9.1.18. — Soit X un singleton $\{a\}$. Un mot réduit sur $X \coprod X^{-1}$ est de la forme a^n pour $n \in \mathbf{Z}$; on voit ainsi que $n \mapsto a^n$ établit un isomorphisme entre $\mathbf{Z}$ et $F(\{a\})$: le groupe libre sur un singleton s'identifie à $\mathbf{Z}$.

Exemple 9.1.19. — Soit X un ensemble à deux éléments $\{a, b\}$. Il n'y a rien de plus à dire sur $F(X)$ que les généralités mentionnées plus haut : ses éléments seront les mots réduits en les lettres a, b, a^{-1}, b^{-1}, et on les multiplie en concaténant et simplifiant.

Pour ceux qui connaissent un peu de topologie algébrique, indiquons que ce groupe s'identifie au groupe fondamental du plan privé de deux points. Plus précisément, soit P l'espace topologique $\mathbf{R}^2 - \{-1, 1\}$; soit $f : [0; 1] \to P$ le lacet $t \mapsto -1 + \exp(2i\pi t)$ (basé en l'origine) et soit $g : [0; 1] \to P$ le lacet $t \mapsto 1 - \exp(2i\pi t)$ (basé en l'origine). L'application ensembliste $\{a, b\} \to \pi_1(P, O)$ qui envoie a sur f et b sur g induit un morphisme de groupes $F(\{a, b\}) \to \pi_1(P, O)$; on démontre que c'est un isomorphisme.

9.2. Groupes définis par générateurs et relations. — Soit X un ensemble, et soit R un ensemble de mots sur l'alphabet $X \coprod X^{-1}$. On se propose informellement de construire le groupe le plus général fabriqué à partir de X (l'ensemble des générateurs) et dans lequel les mots appartenant à R (l'ensemble des relations) sont triviaux.

Définition 9.2.1. — On appelle *groupe défini par l'ensemble de générateurs X et l'ensemble de relations R*, et l'on note $\langle X|R\rangle$, le quotient de $F(X)$ par le plus petit sous-groupe distingué de $F(X)$ contenant R ; si $X = \{x_1, \ldots, x_n\}$ on omettra parfois les accolades ensemblistes : on écrira $\langle x_1, \ldots, x_n|R\rangle$ au lieu de $\langle\{x_1, \ldots, x_n\}|R\rangle$.

Proposition 9.2.2 **(propriété universelle d'un groupe défini par générateurs et relations)**

Soit X un ensemble, soit R un ensemble de mots sur l'alphabet $X \coprod X^{-1}$ et soit p l'application composée $X \to F(X) \to \langle X|R\rangle$. Soit G un groupe et soit $(g_x)_{x\in X}$ une famille d'éléments de G telle que $m(g_x)_x = e$ pour tout $m \in R$. Il existe un unique morphisme de groupes $\varphi\colon \langle X|R\rangle$ tel que $\varphi(p(x)) = g_x$ pour tout $x \in X$.

Démonstration. — C'est une simple combinaison du lemme 9.1.8 et de 2.12.1. □

Commentaires 9.2.3. — Soit X, R, p et G comme dans l'énoncé de la proposition 9.2.2 et soit φ un morphisme de groupes de $\langle X|R\rangle$ vers G ; posons $g_x = \varphi(p(x))$ pour tout $x \in X$. La composée $F(X) \to \langle X|R\rangle \to G$ envoie x sur g_x pour tout x, et coïncide donc avec $m \mapsto m(g_x)_x$; comme R est contenu dans le noyau du morphisme quotient $F(X) \to \langle X|R\rangle$, il vient $m(g_x)_{x\in X} = e$ pour tout $m \in R$.

Il s'ensuit que la propriété universelle décrite par la proposition 9.2.2 peut se reformuler comme suit : l'application $\varphi \mapsto (\varphi(p(x))_{x\in X}$ établit une bijection entre l'ensemble des morphismes de groupes de $\langle X|R\rangle$ vers G et l'ensemble des familles $(g_x)_{x\in X}$ d'éléments de G telles que $m(g_x)_x = e$ pour tout $m \in R$.

En termes un peu plus informels, se donner un morphisme de $\langle X|R\rangle$ vers un groupe G, c'est choisir une famille $(g_x)_{x\in X}$ d'éléments de G qui annulent chacune des relations appartenant à R.

Cette propriété universelle de l'application $p\colon X \to \langle X|R\rangle$ caractérise celle-ci à unique isomorphisme près. Nous laissons au lecteur le soin de formuler et de démontrer l'énoncé précis traduisant cette affirmation (il faut s'inspirer de 9.1.15.2).

9.2.4. *Le problème du mot.* — Soit X un ensemble et soit R un ensemble de mots sur l'alphabet $X \coprod X^{-1}$. Le morphisme quotient de $F(X)$ vers $\langle X|R\rangle$ envoie un mot m sur $m(\overline{x})_x$, et est surjectif. Le groupe $\langle X|R\rangle$ est donc constitué d'éléments de la forme $m(\overline{x})_x$ où m est un mot sur l'alphabet $X \coprod X^{-1}$. Mais cette description est bien entendu très insuffisante : elle ne précise pas à quelle condition sur les mots m et n on a $m(\overline{x})_x = n(\overline{x})_x$, c'est-à-dire encore à quelle condition sur un mot m on a $m(\overline{x})_x = e$. Certes, la réponse théorique est connue : on a $m(\overline{x})_x = e$ si et seulement si m appartient au plus petit sous-groupe distingué de $F(X)$ contenant R. Mais décider *en pratique* si c'est le cas, disons lorsque X et R sont finis, est souvent extrêmement difficile. C'est même *impossible* en toute généralité. On démontre en effet qu'il *n'existe pas* d'algorithme permettant de résoudre le *problème du mot*, c'est-à-dire de répondre en un temps fini pour n'importe quel ensemble fini X, n'importe quel ensemble fini R de mots sur l'alphabet $X \coprod X^{-1}$, et n'importe quel mot m sur ce même alphabet, à la question «m appartient-il au plus petit sous-groupe distingué de $F(X)$ contenant R ?».

Bien entendu, dans bon nombre de cas rencontrés, on sait tout de même résoudre ce problème ; ce qui est affirmé ici est simplement l'inexistence d'un algorithme marchant *dans tous les cas.*

Exemple 9.2.5. — Soit X un singleton $\{a\}$. Le morphisme $\mathbf{Z} \to F(X), n \mapsto a^n$ est un isomorphisme (9.1.18).

Soit n un entier. Comme $F(\{a\})$ est abélien, son plus petit sous-groupe distingué contenant a^n est le groupe engendré par a^n. Il s'ensuit que $\langle a|a^n\rangle$ est une présentation de $\mathbf{Z}/n\mathbf{Z}$ par générateurs et relations.

Exemple 9.2.6 **(Une présentation de $\mathbf{Z}^2$ par générateurs et relations)**
Nous allons démontrer que le groupe $\langle a, b|aba^{-1}b^{-1}\rangle$ est isomorphe à $\mathbf{Z}^2$. Pour cela, considérons l'application ensembliste de $\{a, b\}$ dans $\mathbf{Z}^2$ qui envoie a sur $(1, 0)$ et b sur $(0, 1)$. Comme

$$(1, 0) + (0, 1) - (1, 0) - (0, 1) = (0, 0),$$

cette application induit un morphisme φ de $\langle a, b|aba^{-1}b^{-1}\rangle$ vers $\mathbf{Z}^2$.

Par ailleurs, $\langle a, b|aba^{-1}b^{-1}\rangle$ est engendré par $\overline{a}$ et $\overline{b}$ qui commutent, puisque $\overline{a}\overline{b} = \overline{b}\overline{a}$ en vertu de la relation imposée. L'application $\psi : \mathbf{Z}^2 \to \langle a, b|aba^{-1}b^{-1}\rangle$ donnée par la formule $(n, m) \mapsto \overline{a}^n\overline{b}^m$ est par conséquent un morphisme de groupes. On vérifie immédiatement (sur les générateurs $\overline{a}$ et $\overline{b}$ d'une part, $(0, 1)$ et $(1, 0)$ de l'autre) que $\psi \circ \varphi = \mathrm{Id}$ et $\varphi \circ \psi = \mathrm{Id}$; ainsi, $\langle a, b|aba^{-1}b^{-1}\rangle$ est isomorphe à $\mathbf{Z}^2$.

Remarque 9.2.7. — Soit F le groupe libre sur l'alphabet $\{a, b\}$. On peut alors également décrire $\mathbf{Z}^2$ comme l'abélianisé de F. Pour le voir, considérons l'application ensembliste de $\{a, b\}$ dans $\mathbf{Z}^2$ qui envoie a sur $(1, 0)$ et b sur $(0, 1)$. Cette application induit un morphisme φ de F vers $\mathbf{Z}^2$. Comme $\mathbf{Z}^2$ est abélien, ce morphisme φ induit à son tour un morphisme ψ de $F/\mathrm{D}(F)$ vers $\mathbf{Z}^2$. Comme $F/\mathrm{D}(F)$ est abélien, les classes $\overline{a}$ et $\overline{b}$ de a et b modulo $\mathrm{D}(F)$ commutent. L'application $\chi : \mathbf{Z}^2 \to F/\mathrm{D}(F)$ donnée par la formule $(n, m) \mapsto \overline{a}^n\overline{b}^m$ est par conséquent un morphisme de groupes. On vérifie immédiatement (sur les générateurs $\overline{a}$ et $\overline{b}$ d'une part, $(0,1)$ et $(1,0)$ de l'autre) que $\chi \circ \psi = \mathrm{Id}$ et $\psi \circ \chi = \mathrm{Id}$; ainsi, $F/\mathrm{D}(F)$ est isomorphe à $\mathbf{Z}^2$.

Exercice 9.2.8. — Soit $n \geqslant 1$. Montrez que $\langle a, b | a^2, b^n, abab \rangle$ s'identifie au groupe diédral D_n (6.2.7.3).

10. Algèbre linéaire, dualité, produit tensoriel

Nous supposerons connues les notions de base de l'algèbre linéaire (théorie de la dimension, calcul matriciel, un peu de réduction des endomorphismes) que nous utiliserons librement ; nous vous renvoyons pour l'essentiel à vous cours antérieurs sur le sujet. Nous commencerons ce chapitre quelques compléments sur le sujet, suivis d'une étude détaillée de la *dualité* (dont vous avez probablement déjà entendu un peu parler) puis de la présentation d'une notion probablement complètement nouvelle : le *produit tensoriel*. On fixe pour toute la suite du chapitre un corps k.

10.1. Espaces d'applications linéaires. — Soient V et W deux k-espaces vectoriels. On vérifie que les formules

$$\varphi + \psi = v \mapsto \varphi(v) + \psi(v) \ \text{et} \lambda\varphi = v \mapsto \lambda\varphi(v)$$

font de l'ensemble $\mathrm{Hom}_k(V, W)$ des applications k-linéaires de V dans W un k-espace vectoriel.

10.2. Sommes directes. — Il y a, comme dans le cas des groupes abéliens deux notions distinctes de somme directe de k-espaces vectoriels.

10.2.1. *La somme directe interne.* — Soit V un k-espace vectoriel et soit (V_i) une famille de sous-espaces vectoriels de V. On a défini en 3.2.2.2 la somme des V_i *comme sous-groupes abéliens de* V ; elle est notée $\sum V_i$, et $\bigoplus V_i$ lorsqu'elle est directe ; on vérifie immédiatement que $\sum V_i$ (c'est-à-dire $\bigoplus V_i$ lorsque la somme est directe) est un sous-espace vectoriel de V.

10.2.2. *La somme directe externe.* — Soit (V_i) une famille de k-espaces vectoriels, qui ne sont pas *a priori* plongés dans un même k-espace vectoriel. La somme directe abstraite des *groupes abéliens* V_i, notée $\bigoplus V_i$, a été définie au 3.3.1. Nous laissons le lecteur vérifier que la formule $(\lambda, (v_i)_i) \mapsto (\lambda v_i)_i$ fait du groupe abélien $\prod V_i$ un k-espace vectoriel dont $\bigoplus V_i$ est un sous-espace vectoriel appelé la somme directe externe des V_i ; pour tout j, l'injection h_j de V_j dans $\bigoplus V_i$ qui envoie un élément v sur la famille (v_i) telle que $v_i = 0$ pour $i \neq j$ et telle que $v_j = v$ est k-linéaire, et $\bigoplus V_i$ s'identifie à la somme directe interne des $h_i(V_i)$.

Commentaires 10.2.3. — Soit (V_i) une famille *finie* de k-espaces vectoriels. La somme directe externe $\bigoplus V_i$ s'identifie alors au produit $\prod V_i$. L'emploi de l'une ou l'autre notation peut être une affaire de goût ou de contexte, mais on préfère souvent la première. Il est en effet conseillé de penser à l'opération $(V_i) \mapsto \bigoplus V_i$ comme à une *addition* d'espaces vectoriels (ainsi, si les V_i sont tous de dimension finie, la dimension de $\bigoplus V_i$ est la *somme* des dimensions des V_i) ; le rôle de la *multiplication* d'espaces vectoriels est quant à lui joué par une autre opération, le *produit tensoriel*, que nous introduirons plus loin.

10.3. Quotient par un sous-espace vectoriel. — Soit V un k-espace vectoriel et soit W un sous-espace vectoriel de V. Soit $\lambda \in k$. Le morphisme de groupes

$$\begin{tikzcd} V \arrow[r, "v\mapsto\lambda v"] \arrow[rr, bend left] & V \arrow[r] & V/W \end{tikzcd}$$

est trivial sur W donc induit un endomorphisme de V/W ; en faisant varier λ on obtient une loi externe $k \times V/W \to V/W$ qui envoie par construction $(\lambda, \overline{v})$ sur $\overline{\lambda v}$ pour tout (λ, v). On vérifie aussitôt que cette loi fait du groupe abélien V/W un k-espace vectoriel et que l'application quotient $V \to V/W$ est linéaire (nous avions déjà signalé et utilisé ce fait lors de la preuve du corollaire 4.4.10). Si φ est une application linéaire de V dans un k-espace vectoriel V' et si $W \subset \mathrm{Ker}(\varphi)$, on vérifie aussitôt que l'application induite $\overline{\varphi}\colon V/W \to V'$ est linéaire (elle hérite de cette propriété de φ grâce à la surjectivité de $V \to V/W$; la preuve est parallèle à celle du lemme 2.11.1 (2)). Par conséquent si π désigne l'application quotient $V \to V/W$ alors $\psi \mapsto \psi \circ \pi$ induit une bijection (visiblement k-linéaire) entre $\mathrm{Hom}_k(V/W, V')$ et le sous-espace vectoriel de $\mathrm{Hom}_k(V, V')$ formé des applications s'annulant sur W.

Si S est un supplémentaire de W dans V nous laissons le lecteur vérifier que $V \to V/W$ induit un isomorphisme $S \simeq V/W$.

10.4. Bases et calcul matriciel. — Vous n'avez probablement rencontré encore ces concepts qu'en dimension finie. Et même si c'est essentiellement dans ce cas là que nous les utiliserons, il est commode de savoir qu'ils s'étendent sans la moindre difficulté en dimension infinie (cela évite d'avoir à rajouter des hypothèses de finitude inutiles dans certaines énoncés). Nous allons brièvement expliquer comment, en laissant les vérifications au lecteur – les preuves sont la plupart du temps *mutatis mutandis* les mêmes qu'en dimension finie.

10.4.1. — Soit V un k-espace vectoriel. Si (e_i) est une famille de vecteurs de V, le sous-espace vectoriel qu'elle engendre (*i.e* le plus petit sous-espace vectoriel de V contenant les e_i) est l'ensemble des combinaisons linéaires $\sum_i a_i e_i$ où les a_i appartiennent à k et sont *presque tous nuls* (sinon, la somme considérée n'aurait *a priori* aucun sens). On dit que (e_i) est *génératrice* si cet ensemble de combinaisons linéaires est V tout entier ; on dit qu'elle est *libre* si

$$\sum a_i e_i = 0 \Rightarrow (\forall i, a_i = 0)$$

pour toute famille (a_i) d'éléments presque tous nuls de k. On dit que (e_i) est une *base* si elle est à la fois libre et génératrice. Il revient au même de demander que tout élément de V admette une *unique* écriture sous la forme $\sum a_i e_i$, où (a_i) est une famille d'éléments presque tous nuls de k.

Toute famille libre de V est contenue dans une famille libre maximale, qui est une base. En particulier, V possède des bases (appliquer l'énoncé précédent à la famille vide, qui est toujours libre). On démontre qu'elles ont toutes même cardinal (au sens de la théorie des cardinaux éventuellement infinis, dont nous n'aurons pas besoin ici ; nous nous servirons simplement du fait que si V possède une base finie, alors toutes ses bases sont finies de même cardinal, qui est appelé la *dimension* de V ; nous avons indiqué l'énoncé général, dont la signification et la preuve sont plus délicates, pour la curiosité du lecteur).

Soit $(e_i)_{i\in I}$ une base de V et soit W un k-espace vectoriel. Pour toute famille $(f_i)_{i\in I}$ d'éléments de W il existe une unique application liénaire $\ell\colon V \to W$ telle que $\ell(e_i) = f_i$ quel que soit i ; elle est évidemment donnée par la formule $\ell(\sum a_i e_i) = \sum a_i f_i$. Cette application est injective, resp. surjective, resp. bijective, si et seulement si (f_i) est libre, resp. génératrice, resp. une base.

10.4.2. — Si I et J sont deux ensembles arbitraires, on note $\mathrm{M}_{I,J}(k)$ l'ensemble des familles $(a_{ij})_{i\in I, j\in J}$ d'éléments de k telles que pour tout j, les éléments a_{ij} pour i variable soient presque tous nuls. Les éléments de $\mathrm{M}_{I,J}(k)$ sont appelées *matrices* de taille $I \times J$ à coefficients dans k. L'ensemble $\mathrm{M}_{I,J}(k)$ a une structure naturelle de k-espace vectoriel : l'addition et la multiplication par un scalaire se font coefficients par coefficients. Si L est un troisième ensemble on dispose d'une application bilinéaire de $\mathrm{M}_{IJ}(k) \times \mathrm{M}_{JL}(k)$ vers $\mathrm{M}_{IL}(k)$ donnée par la formule

$$(a_{ij})_{i,j}(b_{j\ell})_{j,\ell} = \left(\sum_j a_{ij} b_{j\ell}\right)_{i,\ell}.$$

Ce produit est associatif, c'est-à-dire que $(MN)P = M(NP)$ à chaque fois que cela a un sens.

10.4.3. — Pour tout ensemble I, on écrit $\mathrm{M}_I(k)$ au lieu de $\mathrm{M}_{I,I}(k)$. Le produit défini ci-dessus fait de $\mathrm{M}_I(k)$ une k-algèbre ; son groupe des éléments inversibles est noté $\mathrm{GL}_I(k)$.

10.4.4. — Soient V et W deux k-espaces vectoriels, soit $(f_j)_{j\in J}$ une base de V et soit $(e_i)_{i\in I}$ une base de W. Soit $\varphi\colon V \to W$ une application linéaire. La *matrice* de φ dans les bases (f_j) et (e_i) est la famille $(a_{ij})_{i,j} \in \mathrm{M}_{I,J}(k)$ telle que $\varphi(f_j) = \sum_i a_{ij} e_i$ pour tout j. L'application qui envoie φ sur sa matrice dans les bases (f_j) et (e_i) établit une bijection k-linéaire $\mathrm{Hom}_k(V, W) \simeq \mathrm{M}_{I,J}(k)$ (qui dépend des bases !). On vérifie par ailleurs que «la matrice de la composée est le produit des matrices» (nous laissons le lecteur énoncer soigneusement l'assertion rigoureuse résumée par ce slogan, en faisant attention aux espaces de départ et d'arrivée, au choix des bases, etc.).

En particulier en associant à un endomorphisme de V sa matrice «dans la base (e_i)» (c'est-à-dire qu'on prend (e_i) comme base *de départ et d'arrivée*), on établit un isomorphisme de k-algèbres $\mathrm{End}_k(V) \simeq \mathrm{M}_I(k)$, qui induit un isomorphisme de groupes $\mathrm{GL}(V) \simeq \mathrm{GL}_I(k)$.

La notion de matrice de passage et les formules de changement de base s'étendent à ce cadre ; nous laissons le lecteur formuler et prouver les énoncés correspondants.

Remarque 10.4.5. — Il y a une opération du calcul matriciel classique qui ne s'étend pas au cadre plus général que nous venons d'introduire : c'est la transposition. En effet dans notre définition de matrice, on demande que chaque colonne n'est qu'un nombre fini de termes non nuls et si $M = (a_{ij})_{(i,j)\in I\times J}$ est une matrice, la famille $(b_{j,i})_{(j,i)\in J\times I}$ définie par la formule $b_{ji} = a_{ij}$ n'a donc aucune raison d'être une matrice. C'en est toutefois automatiquement une dès que J est fini, et on l'appelle bien évidemment la *transposée* de M ; on la note ${}^tM$. Si I, J et K sont trois ensembles d'indices avec J et K finis, si M appartient à $\mathrm{M}_{IJ}(k)$ et N à $\mathrm{M}_{JK}(k)$ on a ${}^t(MN) = {}^tN{}^tM$.

10.4.6. — Soit V un k-espace vectoriel et soit W un sous-espace vectoriel de V. Soit $(e_i)_{i\in I}$ une base de W.

10.4.6.1. — Soit I' un ensemble d'indices disjoint de I et soit $(e_i)_{i\in I'}$ une famille d'éléments de V. Les assertions suivantes sont équivalentes :

(i) la famille $(\overline{e_i})_{i\in I'}$ est une base de V/W ;

(ii) la famille $(e_i)_{i\in I\coprod II'}$ est une base de V.

En effet, supposons que (i) soit vraie. Si $\sum_{i\in I\coprod J}\lambda_i e_i = 0$ on voit en réduisant cette égalité modulo W que $\sum_{i\in j}\lambda_i\overline{e_i} = 0$, ce qui force les λ_i à pour $i \in J$ être tous nuls car $(\overline{e_i})_{i\in I'}$ est libre. On a alors $\sum_{i\in I}\lambda_i e_i = 0$ et les λ_i sont également tous nuls puisque $(e_i)_{i\in I}$ est libre. Par ailleurs, soit $v \in V$. Comme $(\overline{e_i})_{i\in I'}$ engendre V/W on peut écrire $\overline{v} = \sum_{i\in J}\lambda_i\overline{e_ij}$ pour une certaine famille $(\lambda_i)_{i\in I}$ de scalaires ; cela signifie que $v - \sum_{i\in I} f_j$ appartient à W ; comme le e_i pour $i \in I$ engendrent W on peut écrire $v-w = \sum\lambda_i e_i$ pour une certaine famille $(\lambda_i)_{i\in I}$ de scalaires. Il vient $v = \sum_{i\in I\coprod I'}\lambda_i e_i$, d'où (ii).

Supposons maintenant (ii) vraie. Les (e_i) pour $i \in J$ engendrent alors un supplémentaire S de V (dont ils forment une base). Comme l'application quotient $V \to V/W$ induit un isomorphisme $S \simeq V/W$, les $\overline{e_i}$ pour $i \in J$ forment une base de V/W.

10.4.6.2. — Soit I' un ensemble d'indices disjoint de I et soit $(e_i)_{i\in I'}$ une famille d'éléments de V satisfaisant les conditions équivalentes ci-dessus. Soit φ un endomorphisme de V. Soit $A = (a_{ij})$ la matrice de φ dans la base $(e_i)_{i\in I\coprod I'}$.

L'espace W est stable sous φ si et seulement si $a_{ij} = 0$ pour tout couple (i, j) appartenant à $I' \times I$. Supposons que ce soit le cas. L'application linéaire composée

$$\begin{tikzcd} V \arrow[r, "\varphi"] \arrow[rr, bend left] & V \arrow[r] & V/W \end{tikzcd}$$

est alors triviale sur W (car $\varphi(W) \subset W$) et induit donc un endomorphisme $\overline{\varphi}$ de l'espace vectoriel quotient V/W ; on a par construction $\overline{\varphi}(\overline{v}) = \overline{\varphi(v)}$ pour tout $v \in V$.

On déduit des égalités $\varphi(e_j) = \sum_i a_{ij}e_i$ les deux faits suivants :

◇ la matrice de $\varphi|_W : W \to W$ dans la base $(e_i)_{i\in I}$ est le «bloc $I \times I$ de A», c'est-à-dire $(a_{ij})_{(i,j)\in I\times I}$;

◇ la matrice de $\overline{\varphi}$: $V/W \to V/W$ dans la base $(\overline{e_i})_{i\in I'}$ est le «bloc $I' \times I'$ de A», c'est-à-dire $(a_{ij})_{(i,j)\in I'\times I'}$.

Remarque 10.4.7. — Bien entendu, lorsque $I = \{1, \ldots, n\}$ et $J = \{1, \ldots, m\}$ on emploiera plutôt les notations $\mathrm{M}_{n,m}(k)$, $\mathrm{M}_n(k)$ et $\mathrm{GL}_n(k)$ au lieu de $\mathrm{M}_{I,J}(k)$, $\mathrm{M}_I(k)$ et $\mathrm{GL}_I(k)$.

10.4.8. *Linéarisé d'un ensemble X.* — Soit X un ensemble. On a vu à la section 9 comment définir le «groupe le plus général construit à partir de X». Il y a une construction analogue (et bien plus simple) dans le monde des k-espaces vectoriels, que nous allons décrire. Nous noterons kX, et appellerons *k-linéarisé* de l'ensemble X, le k-espace vectoriel des applications de f de X dans k qui sont nulles pour presque tout x (l'addition et la multiplication par un scalaire étant définies de façon évidente). Si l'on note $[x]$ l'application qui vaut 1 en x et 0 ailleurs alors les $[x]$ forment une base de kX (si $f \in kX$, son écriture dans la base $([x])$ est $f = \sum_x f(x)[x]$, ce qui a bien un sens car les $f(x)$ sont presque tous nuls). Un élément de kX n'est autre qu'une combinaison linéaire finie $\sum_x \lambda_x[x]$: en quelque sorte, on a simplement «décrété» qu'on savait faire des combinaisons k-linéaires d'éléments de X (on pourrait rendre ce principe encore plus manifeste en écrivant par abus x au lieu de $[x]$; c'est parfois faisable, mais dans plusieurs situtations que nous rencontrerons cela pourrait être source de confusion, et nous nous en tiendrons donc à la notation $[x]$).

10.4.8.1. — Mentionnons la propriété universelle de kX : pour tout k-espace vectoriel V et toute application φ de X dans V, il existe une unique application k-linéaire de kX dans V qui envoie $[x]$ sur $\varphi(x)$ pour tout x (c'est simplement dû au fait que la famille $([x])_x$ est une base de kX ; l'application linéaire en question est évidemment $\sum \lambda_x[x] \mapsto \sum \lambda_x\varphi(x)$).

10.4.8.2. — En particulier, si φ est une application de X vers un autre ensemble Y, il existe une unique application k-linéaire de X vers Y qui envoie $[x]$ sur $[\varphi(x)]$ pour tout x ; on l'appelle la *k-linéarisée de φ* et nous la noterons $\varphi_{k\text{-lin}}$. Si $(a_{yx})_{y,x}$ désigne sa matrice dans les bases $([x])_{x\in X}$ et $([y])_{y\in Y}$ alors $a_{yx} = \delta_{y,\varphi(x)}$ pour tout (y, x) (où $\delta_{a,b}$ est le symbole de Kronecker qui vaut 1 si $a = b$ et 0 sinon).

Il est immédiat que $(\mathrm{Id}_X)_{k\text{-lin}} = \mathrm{Id}_{kX}$ et que si ψ est une application de Y vers un ensemble Z alors $(\psi \circ \varphi)_{k\text{-lin}} = \psi_{k\text{-lin}} \circ \varphi_{k\text{-lin}}$). Il s'ensuit que si φ est bijective alors $\varphi_{k\text{-lin}}$ l'est aussi, et $(\varphi_{k\text{-lin}})^{-1} = (\varphi^{-1})_{k\text{-lin}}$.

On a même un résultat plus fort : comme l'application $\varphi_{k\text{-lin}}$ est injective (resp. surjective, resp. bijective) si et seulement si la famille $(\varphi([x]))_{x\in X}$ est une famile libre (resp. une famille génératrice, resp. une base) de kY, on voit que $\varphi_{k\text{-lin}}$ est injective (resp. surjective, resp. bijective) *si et seulement si* c'est le cas de φ.

10.4.8.3. — Notons par ailleurs qu'on peut retrouver φ à partir de $\varphi_{k\text{-lin}}$: pour tout $x \in X$, l'image de x par φ est l'unique élément y de Y tel que $\varphi([x]) = [y]$. Par conséquent, l'application $\chi \mapsto \chi_{k\text{-lin}}$ de $\hom(X, Y)$ vers $\hom_k(kX, kY)$ est injective.

10.5. Projections. — Soit V un k-espace vectoriel et soient W et W' deux sous-espaces vectoriels de V tels que $V = W \oplus W'$. La *projection sur W parallèlement à W'* est l'application linéaire p de V dans V qui envoie un élément $v = w + w'$ de V (avec $w \in W$ et $w' \in W'$) sur w. Il est immédiat que si q désigne la projection sur W' parallèlement à W alors $p + q = \mathrm{Id}$, c'est-à-dire encore $q = p - \mathrm{Id}$; on a les égalités

$$W = \mathrm{Im}(p) = \mathrm{Ker}(p - \mathrm{Id}) \text{ et } W' = \mathrm{Ker}(p) = \mathrm{Im}(p - \mathrm{Id}).$$

On dit emploie assez souvent le terme *projecteur* au lieu de *projection.*

Lemme 10.5.1. — *Soit V un k-espace vectoriel et soit p un endomorphisme de V. Les assertions suivantes sont équivalentes :*

(i) *p est une projection ;*

(ii) $p^2 = p$.

Démonstration. — Si p est une projection il est immédiat que $p^2 = p$. Réciproquement supposons que $p^2 = p$. Comme p annule polynôme $X^2 - X = X(X-1)$ et comme X et $X-1$ sont premiers entre eux le lemme des noyau assure que $V = \mathrm{Ker}(p) \oplus \mathrm{Ker}(p - \mathrm{Id})$. Nous allons toutefois le montrer directement ici, car la preuve est élémentaire (c'est lié au fait que la relation de Bezout entre X et $X-1$ est très simple).

Soit $v \in V$; montrons qu'il possède une unique écriture de la forme $w + w'$ avec $w \in \mathrm{Ker}(p)$ et $w' \in \mathrm{Ker}(p - \mathrm{Id})$. Commençons par l'unicité. Supposons donc que $v = w + w'$ avec $w \in \mathrm{Ker}(p)$ et $w' \in \mathrm{Ker}(p - \mathrm{Id})$. En appliquant p aux deux membres de cette égalité il vient $p(v) = p(w) + p(w') = w'$, si bien que nécessairement $w' = p(v)$ et $w = v - w' = v - p(v)$; l'unicité est donc établie. Pour l'existence, on s'inspire comme il se doit des formules exhibées par la preuve de l'unicité. On pose donc $w = v - p(v)$ et $w' = p(v)$. On a $v = w + w'$; il reste à s'assurer que $w \in \mathrm{Ker}(p)$ et $w' \in \mathrm{Ker}(p - \mathrm{Id})$. On a $p(w) = p(v) - p^2(v) = 0$ puisque $p^2 = p$; ainsi $w \in \mathrm{Ker}(p)$. On a par ailleurs $p(w') = p^2(v) = p(v) = w'$ (la seconde égalité utilise encore le fait que $p^2 = p$, si bien que $w' \in \mathrm{Ker}(p - \mathrm{Id})$.

Si $v \in V$ il possède donc par ce qui précède une unique écriture sous la forme $w + w'$ avec $w \in \mathrm{Ker}(p)$ et $w' \in \mathrm{Ker}(p - \mathrm{Id})$, et on a vu qu'on a $w' = p(v)$. Par conséquent, p est la projection sur $\mathrm{Ker}(p - \mathrm{Id})$ parallèlement à $\mathrm{Ker}(p)$. □

Lemme 10.5.2. — *Soit V un k-espace vectoriel, et soit W un sous-espace vectoriel de V. Soit p un endomorphisme k-linéaire de V satisfaisant les deux propriétés suivantes :*

(a) *$p(v) \in W$ pour tout $v \in V$;*

(b) *$p(v) = v$ pour tout $v \in W$.*

L'endomorphisme p est alors une projection d'image W.

Démonstration. — Soit $v \in V$. On a alors $p(v) \in W$ d'après (a), et $p(p(v)) = p(v)$ d'après (b). Ainsi $p^2 = p$ et p est donc une projection d'après le lemme 10.5.1.

Il résulte par ailleurs de (a) que $\mathrm{Im}(p) \subset W$, et de (b) que $W \subset \mathrm{Im}(p)$; par conséquent, $\mathrm{Im}(p) = W$. □

10.6. Dualité. — Nous allons maintenant introduire une notion importante en algèbre linéaire, la *dualité*. Comme nous le verrons, les définitions de base en la matière ont un sens pour un espace vectoriel quelconque, mais la théorie ne fonctionne vraiment bien qu'en dimension finie (et c'est uniquement dans ce contexte qu'elle mérite vraiment son nom de dualité).

La dualité n'est pas un sujet très ardu en soi, et les raisonnements qu'on y rencontre sont le plus souvent assez formels ; mais peut s'y heurter à des difficultés psychologiques car il est parfois délicat de bien comprendre où vivent les objets qu'elle met en jeu, et l'on peut pour cette raison être un peu perdu devant certaines formules même lorsqu'elles sont essentiellement tautologiques.

10.6.1. *Remarque sur les espaces d'applications linéaires.* — Soient V et W deux k-espaces vectoriels. Soit $(f_j)_{j\in J}$ une base de V et soit $(e_i)_{i\in I}$ une base de W. Pour tout (i,j), notons $\varphi_{i,j}$ l'application linéaire de V dans W qui envoie f_j sur e_i et f_ℓ sur 0 pour $\ell \neq j$. Sa matrice dans les bases (f_j) et (e_i) est la matrice E_{ij} définie comme suit : son coefficient d'indice (i,j) vaut 1 et tous les autres sont nuls.

Il est clair que la famille $(E_{ij})_{i,j}$ est une famille libre d'éléments de $\mathrm{M}_{I,J}(k)$. Si de plus J est fini elle est également génératrice : une matrice (a_{ij}) peut en effet dans ce cas s'écrire $\sum_{i,j} a_{ij}E_{ij}$, ce qui a un sens car la somme précédente est finie puisque presque tous les a_{ij} sont nuls (il n'y a qu'un nombre fini de coefficients non nuls dans chaque colonne, et un nombre fini de colonnes car J est fini).

Il s'ensuit que (φ_{ij}) est une famille libre de $\mathrm{Hom}_k(V,W)$, et que c'en est une base si V est de dimension finie.

Définition 10.6.2. — Soit V un k-espace vectoriel. On appelle *dual* de V, et l'on note V^*, le k-espace vectoriel des formes linéaires sur V, c'est-à-dire encore le k-espace vectoriel $\mathrm{Hom}_k(V,k)$.

10.6.3. *Formes linéaires coordonnées et base duale.* — Soit V un k-espace vectoriel et soit $(e_i)_{i\in I}$ une bas de V. On désigne traditionnellement par e_j^* la «j-ème forme linéaire » qui envoie un vecteur $\sum a_ie_i$ sur a_j. Attention : consacrée par l'usage, la notation e_j^* est trompeuse, car e_j^* *ne dépend pas que de* e_j*, mais bien de toute la base* (e_i)*, et le symbole* v^* *pour un vecteur «isolé»* v *de* V *n'a aucun sens.*

Il résulte de 10.6.1, appliqué en munissant V de la base (e_i) et k de la base 1, que (e_i^*) est une famille libre de V^*, et même une base de ce dernier si V est de dimension finie (c'est-à-dire si I est fini) ; on l'appelle alors la *base duale* de la base (e_i).

Ainsi si V est de dimension finie V^* a même dimension que V, et lui est en particulier isomorphe. Mais il n'y a pas en général d'isomorphisme canonique entre V et V^* ; le choix de (e_i) en fournit un (qui envoie e_i sur e_i^*) mais il dépend de (e_i).

Si V est de dimension infinie, (e_i^*) n'est jamais génératrice (on peut en fait montrer que la dimension de V^*, au sens des cardinaux, est strictement supérieure à celle de V). Pour le voir, considérons la forme linéaire φ qui envoie un vecteur $\sum a_ie_i$ sur $\sum a_i$. Nous allons montrer par l'absurde qu'elle n'appartient pas à $\mathrm{Vect}(e_i^*)_i$. Supposons que φ s'écrive $\sum a_ie_i^*$ pour une certaine famille (a_i) d'éléments presque tous nuls de k. Comme I est infini il existe $j\in I$ tel que $a_j = 0$. On a alors

$$1 = \varphi(e_j) = \sum_i a_ie_i^*(e_j) = \sum_i a_i\delta_{ij} = a_j = 0,$$

ce qui est absurde. Notons que si I est fini, la forme φ est évidemment égale à $\sum_{i\in i} e_i^*$, expression qui n'a *aucun sens* lorsque I est infini.

Lemme 10.6.4. — *Soit* V *un* k*-espace vectoriel, soit* W *un sous-espace vectoriel de* V *et soit* v *un vecteur de* V *n'appartenant pas à* W*. Il existe une forme linéaire* φ *sur* V *telle que* $\varphi|_W = 0$ *et* $\varphi(v)\neq 0$*.*

Démonstration. — Comme v n'appartient pas à W il est non nul et $kv \cap W = \{0\}$. Soit S un supplémentaire de $W \oplus kv$ dans V. L'unique forme linéaire sur V valant 0 sur $W \oplus S$ et 1 sur v convient alors. □

10.6.5. *Le bidual.* — Soit V un k-espace vectoriel. On appelle *bidual* de V le dual $(V^*)^*$ de V^*, qu'on note plus simplement V^{**}. Pour tout $v \in V$, l'application de V^* dans k qui envoie une forme linéaire φ sur $\varphi(v)$ est elle-même une forme linéaire, donc est un élément $\theta_V(v)$ de V^{**}. Il est immédiat que l'application $\theta_V \colon V \to V^{**}$ est linéaire. *L'application θ_V est injective.* En effet, soit v un vecteur non nul de V. Le lemme 10.6.4, appliqué au vecteur v et au sous-espace $\{0\}$ de V, assure l'existence d'une forme $\varphi \in V^*$ telle que $\varphi(v) \neq 0$; puisque $\varphi(v) = \theta_V(v)(\varphi)$, on a $\theta_V(v) \neq 0$ et θ_V est injective.

Supposons maintenant V de dimension finie. Il résulte de 10.6.3 que

$$\dim V = \dim V^* = \dim V^{**}.$$

Par conséquent, l'application linéaire θ_V est *bijective.*

On dispose ainsi *lorsque V est de dimension finie* d'un isomorphisme $\theta_V \colon V \simeq V^{**}$ qui a l'avantage d'être *canonique* : il ne repose sur aucun choix et est donné par une formule universelle. Par contre, sa réciproque θ_V^{-1} n'admet pas de description par une formule explicite ; c'est simplement un argument de dimension qui en assure l'existence.

Précisons que la notation θ_V n'est pas standard, mais que c'est celle que nous emploierons systématiquement.

Lemme 10.6.6. — *Soit V un k-espace vectoriel de dimension finie, soit $(\varphi_i)_{i\in I}$ une base de V^* et soit $(e_i)_{i\in I}$ une base de V. Les assertions suivantes sont équivalentes :*

(i) $(\varphi_i)_i$ *est la base duale de* (e_i) *;*

(ii) $(\theta_V(e_i))_i$ *est la base duale* (φ_i^*) *de* (φ_i) *dans* $V^{**} = (V^*)^*$.

Démonstration. — Dire que (φ_i) est la base duale de (e_i) signifie exactement que $\varphi_i(e_j) = \delta_{ij}$ pour tout (i,j). Mais $\varphi_i(e_j) = \theta_V(e_j)(\varphi_i)$ pour tout (i,j)par définition de θ_V ; par conséquent, (φ_i) est la base duale de (e_i) si et seulement si $\theta_V(e_j)(\varphi_i) = \delta_{ij}$ pour tout (i,j), c'est-à-dire encore si et seulement si $(\theta_V(e_i))_i$ est la base duale de (φ_i). □

Corollaire 10.6.7. — *Soit V un k-espace vectoriel de dimension finie et soit (φ_i) une base de V^*. Il existe une et une seule base de V dont (φ_i) est la base duale, à savoir $(\theta_V^{-1}(\varphi_i^*))_i$.*

Remarque 10.6.8. — La base dont le corollaire ci-dessus affirme l'existence et l'unicité est parfois appelée la base *antéduale* de (φ_i). Cela dit, il peut arriver qu'on utilise θ_V pour identifier V à V^{**} et modulo cette identification, la base antéduale de (φ_i) dans V est alors simplement sa base duale (φ_i^*) dans $V^{**} = V$, et l'on peut écrire pour toute base $(e_i)_i$ de V l'égalité $(e_i^{**})_i = (e_i)_i$.

Définition 10.6.9. — Soit V un k-espace vectoriel, soit W un sous-espace vectoriel de V, et soit E un sous-espace vectoriel de V^*.

On appelle *orthogonal* du sous-espace W, et l'on note $W^{\perp}$, l'ensemble des formes φ appartenant à V^* telles que $\varphi(v) = 0$ pour tout $v \in V$.

On appelle *orthogonal* du sous-espace E, et l'on note $E^{\perp}$, l'ensemble des vecteurs v appartenant à V tels que $\varphi(v) = 0$ pour tout $\varphi \in E$.

Commentaires 10.6.10. — On prendra garde que la notation $\perp$ s'applique aussi bien à des sous-espaces de V qu'à des sous-espaces de V^*, avec des significations différentes dans les deux cas, et une certaine ambiguïté : en effet si E est un sous-espace de V^* il possède un orthogonal $E^{\perp}$ dans V défini comme ci-dessus, mais également un orthogonal dans $V^{**} = (V^*)^*$, qui devrait aussi être noté $E^{\perp}$ d'après nos conventions. En pratique, cela ne posera aucun problème : nous réserverons la notation $E^{\perp}$ à l'orthogonal de E dans V, et introduirons une notation spécifique pour son orthogonal dans V^{**} si nous en avons besoin.

10.6.11. *Premières propriétés des orthogonaux.* — Soit V un k-espace vectoriel, soit W un sous-espace vectoriel de V et soit E un sous-espace vectoriel de V^*. On vérifie aussitôt que $W^{\perp}$ est un sous-espace vectoriel de V^*, et que $E^{\perp}$ est un sous-espace vectoriel de V. Il est tautologique que $V \subset (V^{\perp})^{\perp}$ et $E \subset (E^{\perp})^{\perp}$, que $\{0_V\}^{\perp} = V^*$, que $\{0_{V^*}\}^{\perp} = V$ et que $V^{\perp} = \{0_{V^*}\}$.

Proposition 10.6.12. — *Soit V un k-espace vectoriel et soit W un sous-espace vectoriel de V.*

(1) *On a $(W^{\perp})^{\perp} = W$. En particulier – prendre $W = \{0_V\}$ – on a $(V^*)^{\perp} = \{0_V\}$.*

(2) *Si V est de dimension infinie, il existe un sous-espace vectoriel E de V^* tel que l'inclusion $E \subset (E^{\perp})^{\perp}$ soit stricte.*

(3) *Supposons V de dimension finie et soit E un sous-espace vectoriel de V^*.*

  (3a) $\dim W^{\perp} = \dim V - \dim W$ *;*

  (3b) $\dim E^{\perp} = \dim V - \dim E$ *;*

  (3c) $(E^{\perp})^{\perp} = E$ *;*

  (3d) *le sous-espace vectoriel $\theta_V(E^{\perp})$ de V^{**} est égal à l'orthogonal de E dans V^{**}.*

Démonstration. — On sait que $W \subset (W^{\perp})^{\perp}$. Soit $v \in V \backslash W$. Le lemme 10.6.4 assure qu'il existe une forme linéaire φ sur V, nulle sur W, et telle que $\varphi(v) \neq 0$. La forme φ appartient alors à $W^{\perp}$, et comme $\varphi(v) \neq 0$ le vecteur v n'appartient pas à $(W^{\perp})^{\perp}$. Par conséquent $(W^{\perp})^{\perp} \subset W$, d'où (1).

Supposons que l'espace V est de dimension infinie, choisissons une base $(e_i)_{i \in I}$ de V, et posons $E = \mathrm{Vect}(e_i^*)_i$. On sait que E est un sous-espace strict de V^* (10.6.3). Par ailleurs, l'orthogonal de E est constitué des vecteurs v de V dont toutes les coordonnées dans la base (e_i) sont nulles ; c'est donc l'espace nul, et il vient

$$(E^{\perp})^{\perp} = \{0_V\}^{\perp} = V^* \supsetneq E,$$

d'où (2).

Supposons maintenant V de dimension finie, et fixons un sous-espace vectoriel E de V^*. Notons n la dimension de V et m celle de W, et choisissons une base $(e_1, \ldots, e_n)$

de V telle que $(e_1, \ldots, e_m)$ soit une base de W. Soit $\varphi \in V^*$; écrivons $\varphi = \sum a_i e_i^*$. La forme φ appartient à $W^\perp$ si et seulement si $\varphi(e_j) = a_j = 0$ pour tout j compris entre 1 et m ; par conséquent , $W^\perp = \mathrm{Vect}(e_i^*)_{m+1 \leqslant i \leqslant n}$ et $W^\perp$ est donc de dimension $n - m$, ce qui montre (3a).

Soit $v \in V$. Le vecteur v appartient à $E^\perp$ si et seulement si $\varphi(v) = 0$ pour toute $\varphi \in E$. Puisque $\varphi(v) = \theta_V(v)(\varphi)$, il s'ensuit que v appartient à $E^\perp$ si et seulement si $\theta_V(v)(\varphi) = 0$ pour toute $\varphi \in E$, ce qui revient à demander que $\theta_V(v)$ appartienne à l'orthogonal E' de E dans V^{**}. On a donc $E = \theta_V^{-1}(E')$, soit encore $\theta_V(E) = E'$, d'où (3d).

En appliquant (3a) au couple (V^*, E) au lieu de (V, W), il vient

$$\dim E' = \dim V^* - \dim E = \dim V - E.$$

Mais comme $E' = \theta_V(E^\perp)$ on a $\dim E' = \dim E^\perp$ et partant $\dim E^\perp = \dim V - \dim E$, et (3b) est démontré.

L'assertion (3b) assure que $\dim E^\perp = \dim V - \dim E$. L'assertion (3a) appliquée avec $W = E^\perp$ assure que $\dim(E^\perp)^\perp = \dim V - \dim E^\perp$. Il vient $\dim(E^\perp)^\perp = \dim E$. Combiné à l'inclusion $E \subset (E^\perp)^\perp$, ceci entraîne que $(E^\perp)^\perp = E$, ce qui prouve (3c) et achève la démonstration. □

Remarque 10.6.13. — Supposons V de dimension finie et soit E un sous-espace vectoriel de V^*. Si on utilise θ_V pour identifier V à V^{**}, l'ambiguïté de la notation $E^\perp$ disparaît : l'assertion (3d) ci-dessus peut en effet se reformuler en disant que modulo cette identification, les orthogonaux de E dans V et dans $V^{**} = V$ sont égaux.

Définition 10.6.14. — Soient V et W deux k-espaces vectoriels et soit $f \colon V \to W$ une application linéaire. On appelle *transposée de* f, et l'on note ${}^t\!f$, l'application linéaire $\varphi \mapsto \varphi \circ f$ de W^* vers V^*.

10.6.15. *Premières propriétés de la transposée.* — Soit V un k-espace vectoriel. Il est clair que ${}^t(\mathrm{Id}_V) = \mathrm{Id}_{V^*}$ et que pour tout k-espace vectoriel W, l'application $f \mapsto {}^t\!f$ de $\mathrm{Hom}_k(V, W)$ dans $\mathrm{Hom}_k(W^*, V^*)$ est linéaire. Si $f \colon V \to W$ et $g \colon U \to V$ sont deux applications linéaires, on vérifie que ${}^t(f \circ g) = {}^t\!g \circ {}^t\!f$ (ce sont des applications linéaires de W^* dans U^*). On en déduit que si f est bijective alors ${}^t\!f$ aussi, et que dans ce cas $({}^t\!f)^{-1} = {}^t(f^{-1})$.

Lemme 10.6.16. — *Soit $f \colon V \to W$ une application linéaire entre deux k-espaces vectoriels. On a* $(\mathrm{Im}(f))^\perp = \mathrm{Ker}({}^t\!f)$ *et* $(\mathrm{Ker}(f))^\perp = \mathrm{Im}({}^t\!f)$. *L'application f est injective (resp. surjective) si et seulement si ${}^t\!f$ est surjective (resp. injective).*

Démonstration. — Soit φ une forme linéaire sur W. La forme φ appartient à $(\mathrm{Im}(f))^\perp$ si et seulement si $\varphi(w) = 0$ pour tout $w \in \mathrm{Im}(f)$, c'est-à-dire encore si et seulement si $\varphi(f(v)) = 0$ pour tout $v \in V$. Mais comme $\varphi(f(v)) = {}^t\!f(\varphi)(v)$, on voit que φ appartient à $(\mathrm{Im}(f))^\perp$ si et seulement si ${}^t\!f(\varphi) = 0$, c'est-à-dire si et seulement si $\varphi \in \mathrm{Ker}({}^t\!f)$.

Posons $K = \mathrm{Ker}(f)$ et soit ψ une forme linéaire sur V. Si ψ appartient à $\mathrm{Im}({}^t\!f)$ il existe une forme linéaire φ sur W telle que $\psi = \varphi \circ f$; il est alors clair que $K \subset \mathrm{Ker}(\psi)$, c'est-à-dire encore que $\psi \in K^\perp$. Réciproquement, supposons que $\psi \in K^\perp$. Elle induit alors une forme linéaire $\overline{\psi}$ sur V/K, qu'on identifie au sous-espace $\mathrm{Im}(f)$ de W ; par construction, on a $\overline{\psi}(f(v)) = \psi(v)$ pour tout $v \in V$. Choisissons un supplémentaire S de $\mathrm{Im}(f)$ dans W, et désignons par φ l'unique forme linéaire sur W qui vaut 0 sur S et $\overline{\psi}$ sur $\mathrm{Im}(f)$. On a par construction $\psi = \varphi \circ f$, si bien que $\psi \in \mathrm{Im}({}^t f)$.

L'application f est injective si et seulement si $\mathrm{Ker}(f) = \{0_V\}$. En vertu de l'assertion (1) de la proposition 13.2.5, cela revient à demander que $(\mathrm{Ker}(f))^\perp = V^*$; puisque $(\mathrm{Ker}(f))^\perp = \mathrm{Im}({}^t f)$ par ce qui précède, f est injective si et seulement si ${}^t f$ est surjective.

L'application f est surjective si et seulement si $\mathrm{Im}(f) = W$. En vertu de l'assertion (1) de la proposition 13.2.5, cela revient à demander que $(\mathrm{Im}(f))^\perp = \{0_{W^*}\}$; puisque $(\mathrm{Im}(f))^\perp = \mathrm{Ker}({}^t f)$ par ce qui précède, f est surjective si et seulement si ${}^t f$ est injective. □

Proposition 10.6.17. — *Soit $f \colon V \to W$ une application linéaire entre k-espaces vectoriels.*

(1) *Le diagramme*

$$\begin{array}{ccc} V & \xrightarrow{\theta_V} & V^{**} \\ {\scriptstyle f}\big\downarrow & & \big\downarrow{\scriptstyle {}^t({}^t f)} \\ W & \xrightarrow{\theta_W} & W^{**} \end{array}$$

est commutatif. Si V et W sont de dimension finie on a donc ${}^t({}^t f) = \theta_W \circ f \circ \theta_V^{-1}$, ce qui signifie que ${}^t({}^t f) = f$ si on identifie V^{**} et W^{**} à V et W respectivement via *θ_V et θ_W.*

(2) *Supposons V et W de dimension finie. Le rang de ${}^t f$ est égal à celui de f.*

(3) *Supposons V et W de dimension finie. Soit (f_j) une base de V et soit (e_i) une base de W. Soit M la matrice de f dans les bases (f_j) et (e_i) et soit N la matrice de ${}^t f$ dans les bases (e_i^*) et (f_j^*). On a $N = {}^t M$.*

Démonstration. — Soit $v \in V$. Par définition, on a

$$\begin{aligned} {}^t({}^t f)(\theta_V(v)) &= \theta_V(v) \circ {}^t f \\ &= \varphi \mapsto \theta_V(v)({}^t f(\varphi)) \\ &= \varphi \mapsto \theta_V(v)(\varphi \circ f) \\ &= \varphi \mapsto (\varphi \circ f)(v)) \\ &= \varphi \mapsto \varphi(f(v)) \\ &= \varphi \mapsto \theta_W(f(v))(\varphi) \\ &= \theta_W(f(v)), \end{aligned}$$

ce qui achève de prouver (1).

L'assertion (2) provient de l'égalité $(\mathrm{Im}(f))^\perp = \mathrm{Ker}({}^t f)$ (lemme 10.6.16), du fait que $\dim((\mathrm{Im}(f))^\perp) = \dim W - \dim(\mathrm{Im}(f))$ (proposition 13.2.5 (3a)), et de la formule du rang.

Montrons maintenant (3). Soit a_{ij} le terme général de M et soit b_{ji} le terme général de N. Fixons i. On a ${}^t f(e_i^*) = \sum_\ell b_{\ell i} f_\ell^*$. Fixons j et appliquons les deux membres de l'égalité précédente (qui sont des formes linéaires sur V) au vecteur f_j. Il vient

$$(e_i^* \circ f)(f_j) = b_{ji}$$

et donc

$$\begin{aligned} b_{ji} &= e_i^*(f(f_j)) \\ &= e_i^*\left(\sum_\lambda a_{\lambda j} e_\lambda\right) \\ &= a_{ij}, \end{aligned}$$

ce qu'il fallait démontrer. □

Remarque 10.6.18. — Une fois connue (3), la partie de l'assertion (1) relative au cas de la dimension finie est équivalente à l'égalité ${}^t({}^tM) = M$, et l'assertion (2) est équivalente au fait que $\mathrm{rg}({}^tM) = \mathrm{rg}(M)$.

10.7. Le produit tensoriel. — Nous allons maintenant introduire une opération sur les k-espaces vectoriels qui est probablement nouvelle pour bon nombre d'entre vous. Étant donnés deux k-espaces vectoriels V et W, on cherche à définir une application bilinéaire $(vw) \mapsto v \otimes w$ sur $V \times W$, à valeur dans un k-espace vectoriel, et qui soit «la plus générale possible» : on souhaite, en quelque sorte, qu'elle ne vérifie rien d'autre que ce qui est imposé par la bilinéarité et la théorie générale des espaces vectoriels. Comme d'habitude, le fait d'être «le plus général possible» s'exprime rigoureusement en termes de satisfaction d'une propriété universelle.

Proposition 10.7.1. — *Soient V et W deux k-espaces vectoriels. Il existe un k-espace vectoriel $V \otimes_k W$ et une application bilinéaire $(v, w) \mapsto v \otimes w$ de $V \times W$ dans $V \otimes_k W$ telle que la propriété universelle suivante soit satisfaite : pour toute application bilinéaire b de $V \times_k W$ dans un k-espace vectoriel U, il existe une unique application* linéaire *f de $V \otimes_k W$ vers B telle que le diagramme*

$$\begin{tikzcd} V \times W \arrow[r, "b"] \arrow[d, "\otimes"'] & U \\ V \otimes_k W \arrow[ur, "f"'] & \end{tikzcd}$$

commute.

De plus, si (M, β) est un couple formé d'un k-espace vectoriel et d'une application bilinéaire $\beta : V \times W \to M$ qui satisfait la même propriété universelle, il existe une unique application linéaire i de $V \otimes_k W$ dans M telle que $i \circ \otimes = \beta$, et i est un isomorphisme.

Démonstration. — Commençons par montrer l'existence de $(V \otimes_k W, \otimes)$. La preuve consiste essentiellement à *imposer par décret* les relations souhaitées.

Soit E le k-linéarisé de *l'ensemble* $V \times_k W$ (10.4.8) ; rappelons que E possède une base $[v, w]_{(v,w) \in V \times W}$ indexée par $V \times W$. Soit $\mathscr{S}$ la réunion des sous-ensembles

$$\{[v, w + \lambda w'] - [v, w] - \lambda[v, w']\}_{v \in V, (w, w') \in W^2, \lambda \in K}$$

et
$$\{[v+\lambda v',w]-[v,w]-\lambda[v',w]\}_{(v,v')\in V^2,w\in W,\lambda\in K}$$
de E, et soit F le sous-espace vectoriel de E engendré par $\mathscr{S}$. On pose$V\otimes_k W = E/F$; pour tout $(v,w)\in V\times W$, on note $v\otimes w$ la classe de $[v,w]$ dans $V\otimes_k W$.

L'application $\otimes$ *est bilinéaire.* Il résulte en effet de la définition de F que l'on a, pour tout $(v,w)\in V\times W$ et tout $\lambda\in k$, les égalités
$$v\otimes(w+\lambda w')-v\otimes w+\lambda v\otimes w'=0 \text{ et } (v+\lambda v')\otimes w - v\otimes w+\lambda v'\otimes w=0$$
dans $V\otimes_k W=E/F$, d'où l'assertion.

Le couple $(V\otimes_k W,\otimes)$ *satisfait la propriété universelle requise.* Commençons par remarquer que comme $([v,w])_{v,w}$ est une base de E, et en particulier une partie génératrice de ce dernier, les $v\otimes w$ engendrent $V\otimes_k W$.

Soit U un k-espace vectoriel et soit b une application bilinéaire de $V\times W$ dans U. Nous allons montrer qu'il existe une unique application f comme dans l'énoncé. Commençons par l'unicité ; si f existe, la commutativité du diagramme assure que $f(v\otimes w)=b(v,w)$ pour tout $(v,w)\in V\times W$. La valeur de f est ainsi uniquement déterminée sur chacun des éléments de la forme $v\otimes w$ avec $v\in V$ et $w\in W$. Comme ceux-ci engendrent $V\otimes_k W$, on en déduit que f est uniquement déterminée.

Montrons maintenant l'existence de f. Soit ℓ l'unique application linéaire de E dans U qui envoie $[v,w]$ sur $b(v,w)$ pour tout $(v,w)\in V\times W$. La bilinéarité de b implique immédiatement que ℓ s'annule sur les éléments du système générateur $\mathscr{S}$ de F ; par conséquent, ℓ induit par passage au quotient une application linéaire $f: V\otimes_k W\to B$. Soit $(v,w)\in V\times W$. Comme $v\otimes w$ est la classe de $[v,w]$ modulo F, on a $f(v\otimes w)=\ell(e_{v,w})=b(v,w)$. Ainsi, $f\circ\otimes=b$, ce qu'il fallait démontrer.

Unicité de $V\times_k W$. Soit (M,β) comme dans l'énoncé. La propriété universelle de $(V\otimes_k W,\otimes)$ assure l'existence d'une unique application linéaire $i: V\otimes_k W\to M$ telle que $i\circ\otimes=\beta$. Il reste à s'assurer que i est un isomorphisme.

La propriété universelle de (M,β) assure l'existence d'une unique application linéaire $j: M\to V\otimes_k W$ telle que $j\circ\beta=\otimes$. La composée $j\circ i$ est une application linéaire de $V\otimes_K W$ dans lui même, et l'on a $(j\circ i)\circ\otimes=\otimes$. Comme on a par ailleurs $\mathrm{Id}_{V\otimes_k W}\circ\otimes=\otimes$, la partie «unicité» de la propriété universelle de $(V\otimes_k W,\otimes)$ garantit que $j\circ i=\mathrm{Id}_{V\otimes_k W}$. On montre de même que $i\circ j=\mathrm{Id}_M$. □

Commentaires 10.7.2. — Si U est un k-espace vectoriel et si $f: V\otimes_k W\to U$ est une application k-linéaire, il est clair que $f\circ\otimes: V\times W\to U$ est k-bilinéaire. On peut donc reformuler comme suit la propriété universelle de $(V\otimes_k W,\otimes)$: *pour tout k-espace vectoriel U, l'application $f\mapsto f\circ\otimes$ établit une bijection entre l'ensemble des applications k-linéaires de $V\otimes_k W$ dans U et l'ensemble des applications k-bilinéaires de $V\times W$ dans U.*

En termes un peu plus informels : se donner une application k-linéaire de $V\otimes_k W$ vers U, c'est se donner une application k-bilinéaire de $V\times W$ vers U.

10.7.3. — Soient V et W deux k-espaces vectoriels. Il résulte de la construction que $V\otimes_k W$ est engendré comme k-espace vectoriel par les éléments de la forme $v\otimes w$, qu'on appellent les *tenseurs purs* ; il est en fait même engendré *en tant que groupe abélien* par les tenseurs purs, car pour tout $(\lambda,v,w)\in k\times V\times W$, l'élément $\lambda(v\otimes w)$

de $V \otimes_k W$ est égal au tenseur pur $(\lambda v) \otimes w$. Notons une conséquence : $V \otimes_k W$ est nul dès que l'un des deux facteurs est nul.

10.7.4. *Produit tensoriel par une droite.* — Soit V un k-espace vectoriel et soit D une droite vectorielle sur k. Soit e une base de D (c'est-à-dire un vecteur non nul de D). L'application $f : V \to D \otimes_k V$ qui envoie v sur $e \otimes v$ est k-linéaire. L'application de $D \times V$ vers V qui envoie un couple (ae, v) sur av est bilinéaire ; elle induit donc une application k-linéaire g de $D \otimes_k V$ dans V, qui satisfait l'égalité $g(ae \otimes v) = av$ pour tout (a, v). On vérifie immédiatement que $g \circ f = \mathrm{Id}_V$, et que $f \circ g = \mathrm{Id}_{D \otimes_k V}$ (il suffit de vérifier cette dernière égalité sur les tenseurs purs). L'application k-linéaire f est donc bijective de réciproque g. Mentionnons deux conséquences importantes de ce fait.

10.7.4.1. — Appliquons ce qui précède lorsque $D = k$ et lorsque $e = 1$. On en déduit que l'application $v \mapsto 1 \otimes v$ établit un isomorphisme $V \simeq k \otimes_k V$, dont la réciproque envoie $\lambda \otimes v$ sur λv pour tout $\lambda \in k \times V$.

10.7.4.2. — Revenons au cas où (D, e) est quelconque, et supposons que V est lui-même une droite vectorielle, et donnons-nous une base f de V. L'application $v \mapsto e \otimes v$ établit un isomorphisme $V \simeq D \otimes_k V$. Comme $a \mapsto af$ établit un isomorphisme $k \simeq V$, il s'ensuit que $a \mapsto ae \otimes f$ établit un isomorphisme $k \simeq D \otimes_k V$. Autrement dit, $D \otimes_k V$ est une droite vectorielle dont $e \otimes f$ est une base.

10.7.5. *Commutativité du produit tensoriel.* — Soient V et W deux k-espaces vectoriels. L'application de $V \times_k W$ vers $W \otimes_k V$ qui envoie un couple (v, w) sur $w \otimes v$ est k-bilinéaire et induit donc une application k-linéaire $f \colon V \otimes_k W$ dans $W \otimes_k V$, qui satisfait l'égalité $f(v \otimes w) = w \otimes v$ pour tout (v, w). On construit de même une application k-linéaire g de $W \otimes_k V$ vers $V \otimes_k W$ qui satisfait l'égalité $g(w \otimes v) = v \otimes w$ pour tout (w, v). On vérifie sur les tenseurs purs que les applications f et g sont deux bijections réciproques l'une de l'autre. On a ainsi exhibé un isomorphisme naturel $V \otimes_k W \simeq W \otimes_k V$. En termes informels, *le produit tensoriel est commutatif.*

10.7.6. *Associativité du produit tensoriel.* — Soit U un k-espace vectoriel. Pour tout $u \in U$, soit f_u l'application de $V \times W$ dans $(U \otimes_k V) \otimes_k W$ qui envoie un couple (v, w) sur $(u \otimes v) \otimes w$. Elle est k-bilinéaire, et induit donc une application k-linéaire m_u de $V \otimes_k W$ dans $(U \otimes_k V) \otimes_k W$, qui envoie $v \otimes w$ sur $(u \otimes v \otimes w)$ pour tout (v, w).

L'application qui envoie u sur m_u est k-linéaire (pour tout triplet (λ, u, u') appartenant à $k \times U^2$, l'égalité $m_{u+\lambda u'} = m_u + \lambda m_{u'}$ se teste sur les tenseurs purs de $V \otimes_k W$). L'application de $U \times (V \otimes_k W)$ vers $(U \otimes_k V) \otimes_k W$ qui envoie (u, t) sur $m_u(t)$ est donc k-bilinéaire. Elle induit par conséquent une application k-linéaire μ de $U \otimes_k (V \otimes_k W)$ vers $(U \otimes_k V) \otimes_k W$ qui par construction envoie $u \otimes (v \otimes w)$ sur $(u \otimes v) \otimes w$ pour tout (u, v, w).

Pour tout $w \in W$, soit g_w l'application de $U \times V$ dans $U \otimes_k (V \otimes_k W)$ qui envoie un couple (u, v) sur $u \otimes (v \otimes w)$. Elle est k-bilinéaire, et induit donc une application k-linéaire n_w de $U \otimes_k V$ dans $U \otimes_k (V \otimes_k W)$, qui envoie $u \otimes v$ sur $u \otimes (v \otimes w)$ pour tout (v, w).

L'application qui envoie w sur n_w est k-linéaire (pour tout triplet (λ, w, w') appartenant à $k \times W^2$, l'égalité $n_{w+\lambda w'} = n_w + \lambda n_{w'}$ se teste sur les tenseurs purs de $U \otimes_k V$). L'application de $(U \otimes_k V) \otimes_k W$ vers $U \otimes_k (V \otimes_k W)$ qui envoie (t, w) sur $n_w(t)$ est donc k-bilinéaire. Elle induit par conséquent une application k-linéaire ν de $(U \otimes_k V) \otimes_k W$ vers $U \otimes_k (V \otimes_k W)$ qui par construction envoie $(u \otimes v) \otimes w$ sur $u \otimes (v \otimes w)$ pour tout (u, v, w).

Les applications k-linéaires μ et ν sont des bijections réciproques l'une de l'autre : on teste l'égalité $\nu \circ \mu = \mathrm{Id}_{U \otimes_k (V \otimes_k W)}$ sur les termes de la forme $u \otimes (v \otimes w)$ qui engendrent $U \otimes_k (V \otimes_k W)$; on teste l'égalité $\mu \circ \nu = \mathrm{Id}_{(U \otimes_k V) \otimes_k W}$ sur les termes de la forme $(u \otimes v) \otimes w$ qui engendrent $(U \otimes_k V) \otimes_k W$.

On a ainsi montré l'existence d'un isomorphisme naturel entre $(U \otimes_k V) \otimes_k W$ et $U \otimes_k (V \otimes W)$ qui envoie $(u \otimes v) \otimes w$ sur $u \otimes (v \otimes w)$ pour tout (u, v, w). En termes informels, *le produit tensoriel est associatif* ; et l'on pourra se contenter d'écrire $U \otimes_k V \otimes W$ et $u \otimes v \otimes w$, sans parenthèses.

10.7.7. — Il convient de penser à $\otimes$ comme à la loi bilinéaire la plus générale possible définie sur $V \times_k W$. Indiquons une propriété tangible qui apparaît comme une manifestation de ce principe un peu vague : si (v_i, w_i) est une famille d'éléments de $V \times_k W$ alors $\sum v_i \otimes w_i = 0$ si et seulement si $\sum b(v_i, w_i) = 0$ pour toute application bilinéaire b définie sur $V \times W$ et à valeurs dans un k-espace vectoriel U.

En effet, l'implication $\Leftarrow$ est claire, puisque $\otimes$ est une application bilinéaire de $V \times W$ vers $V \otimes_k W$. Réciproquement, supposons que $\sum v_i \otimes w_i = 0$, et soit b une application bilinéaire définie sur $V \times W$ et à valeurs dans un k-espace vectoriel U. La propriété universelle du produit tensoriel garantit l'existence (et l'unicité) d'une application linéaire $f : V \otimes_k W \to U$ telle que $f \circ \otimes = b$. On a alors

$$\sum b(v_i, w_i) = \underbrace{\sum f(v_i \otimes w_i) = f\left(\sum v_i \otimes w_i\right)}_{\text{car } f \text{ est linéaire}} = 0.$$

10.7.8. *Applications induites au niveau tensoriel.* — Soit k un corps et soient V, V', W et W' des k-espaces vectoriels. Soient $f \colon V \to V'$ et soit $g \colon W \to W'$ des applications k-linéaires. L'application de $V \times W$ vers $V' \otimes_k W'$ qui envoie (v, w) sur $f(v) \otimes g(w)$ est k-bilinéaire. Il existe donc une unique application k-linéaire de $V \otimes_k W$ vers $V' \otimes_k W'$ qui envoie $v \otimes w$ sur $f(v) \otimes g(w)$ pour tout (v, w) ; on la note $f \otimes g$.

10.7.8.1. — On a $\mathrm{Id}_V \otimes \mathrm{Id}_W = \mathrm{Id}_{V \otimes_k W}$: on le vérifie immédiatement sur les tenseurs purs.

10.7.8.2. — Si f' est une application k-linéaire de V' vers un k-espace vectoriel V'', et si g' est une application k-linéaire de W' vers un k-espace vectoriel W'' on a alors $(f' \circ f) \otimes (g' \circ g) = (f' \otimes g') \circ (f \otimes g)$; là encore, on le vérifie sur les tenseurs purs.

10.7.8.3. — Il résulte de 10.7.8.1 et 10.7.8.2 que si f et g sont bijectives alors $f \otimes g$ est bijective, de réciproque $f^{-1} \otimes g^{-1}$.

Proposition 10.7.9. — *Soit k un corps, soit V un k-espace vectoriel et soit (V_i) une famille de sous-espaces vectoriels de V tels que $V = \bigoplus_i V_i$; pour tout i, on note f_i l'inclusion $V_i \hookrightarrow V$. Soit W un k-espace vectoriel. Pour tout i, l'application $g_i := f_i \otimes \mathrm{Id}_W \colon V_i \otimes_k W \to V \otimes_k W$ est injective, et $V \otimes_k W$ est la somme directe des $g_i(V_i \otimes_k W)$.*

Démonstration. — Pour tout i, notons p_i la projection de V sur V_i, relative à la décomposition en somme directe que nous nous sommes donnée ; on désigne par q_i l'application k-linéaire $p_i \otimes \mathrm{Id}_W$ de $V \otimes_k W$ vers $V_i \otimes_k W$. Indiquons quelques propriétés des applications p_i et f_i.

(1) On a $p_i \circ f_i = \mathrm{Id}_{V_i}$ pour tout i.

(2) On a $p_i \circ f_j = 0$ pour tout (i, j) avec $i \neq j$.

(3) Pour tout $v \in V$ on a $p_i(v) = 0$ pour presque tout i, et $\sum_i f_i(p_i(v)) = v$.

Nous allons maintenant énoncer quelques propriétés des applications q_i et g_i ; nous justifierons nos assertions ensuite.

(a) On a $q_i \circ g_i = \mathrm{Id}_{V_i \otimes_k W}$ pour tout i.

(b) On a $q_i \circ g_j = 0$ pour tout (i, j) avec $i \neq j$.

(c) Pour tout $x \in V \otimes_k W$ on a $q_i(x) = 0$ pour presque tout i, et $\sum g_i(q_i(x)) = x$.

Les assertions (a) et (b) découlent respectivement des assertions (1) et (2) ci-dessus. En ce qui concerne l'assertion (c), on remarque que l'élément x est une somme finie de tenseurs purs, et qu'il suffit donc de montrer (c) lorsque x est un tenseur pur, auquel cas elle découle directement de (3).

Nous allons maintenant expliquer comment les assertions (a), (b) et (c) permettent de conclure. On déduit de l'assertion (a) que g_i est injective pour tout i (car $\mathrm{Id}_{V_i \otimes_k W}$ est injective). L'assertion (c) assure que $V \otimes_k W$ est égal à $\sum g_i(V_i \otimes_k W)$. Il reste à s'assurer que la somme est directe. Soit (x_i) une famille d'éléments presque tous nuls de $V \otimes_k W$, telle que $x_i \in g_i(V_i \otimes_k W)$ pour tout i et que $\sum x_i = 0$. Nous allons montrer que tous les x_i sont nuls. Par définition, x_i est pour tout i égal à $g_i(y_i)$ pour un certain $y_i \in V_i \otimes_k W$. Fixons j. On a les égalités

$$\begin{aligned} 0 &= q_j\left(\sum_i x_i\right) \\ &= q_j\left(\sum_i g_i(y_i)\right) \\ &= \sum_i q_j(g_i(y_i)) \\ &= y_j, \end{aligned}$$

la dernière égalité découlant de (a) et (b). Comme $y_j = 0$ on a $x_j = g_j(y_j) = 0$, ce qui achève la démonstration. □

Corollaire 10.7.10. — *Soit k un corps, soient V et W deux k-espaces vectoriels et soit V' un sous-espace vectoriel de k. Soit f l'inclusion de V' dans V. L'application $f \otimes \mathrm{Id}_W \colon V' \otimes_k W \to V \otimes_k W$ est alors injective.*

Démonstration. — On choisit un supplémentaire V'' de V' dans V et on applique la proposition 10.7.9 ci-dessus à la famille constituée des deux sous-espaces vectoriels V' et V'' de V. □

Remarque 10.7.11. — Reprenons les notations de la proposition 10.7.9. Chacune des applications g_i étant injective, elle permet de voir $V_i \otimes_k W$ comme un sous-espace vectoriel de $V \otimes_k W$. Modulo ces identifications la dernière assertion de la proposition se récrit sous la forme

$$\left(\bigoplus_i V_i\right) \otimes_k W = \bigoplus_i V_i \otimes_k W.$$

En termes un peu informels, le produit tensoriel commute aux sommes directes.

Proposition 10.7.12. — *Soit k un corps et soient V et W deux k-espaces vectoriels. Soit (e_i) une base de V et (f_j) une base de W. La famille $(e_i \otimes f_j)_{i,j}$ est alors une base de $V \otimes_k W$.*

Démonstration. — Pour tout i, on note D_i la droite vectorielle ke_i, et f_i son inclusion dans V ; pour tout j, on note Δ_j la droite vectorielle kf_j, et g_j son inclusion dans W. Pour tout (i, j), on note $e_i \otimes^\sharp f_j$ le produit tensoriel de e_i et e_j *calculé dans* $D_i \otimes_k \Delta_j$, afin de ne pas le confondre avec l'élément $e_i \otimes f_j$ de $V \otimes_k W$.

En appliquant deux fois la commutation du produit tensoriel aux sommes directes (dont l'énoncé précis est la proposition 10.7.9), on établit l'assertion suivante : *pour tout (i, j), l'application $h_{ij} := f_i \otimes g_j : D_i \otimes_k \Delta_j \to V \otimes_k W$ est injective, et l'on a $V \otimes_k W = \bigoplus_{i,j} h_{ij}(D_i \otimes_k \Delta_j)$.*

En vertu de 10.7.4.2, le produit tensoriel $D_i \otimes_k \Delta_j$ est pour tout (i, j) une droite vectorielle de base $e_i \otimes^\sharp f_j$; il s'ensuit que $h_{ij}(D_i \otimes_k \Delta_j)$ est une droite vectorielle de base $h_{ij}(e_i \otimes^\sharp f_j) = e_i \otimes f_j$. L'égalité $V \otimes_k W = \bigoplus_{i,j} h_{ij}(D_i \otimes_k \Delta_j)$ signifie alors précisément que $(e_i \otimes f_j)_{i,j}$ est une base de $V \otimes_k W$. □

10.7.13. *Quelques conséquences.* — Soit k un corps et soient V et W deux k-espaces vectoriels. Indiquons deux conséquences importantes de la proposition 10.7.12 ci-dessus.

10.7.13.1. — Le produit tensoriel $V \otimes_k W$ est nul si et seulement si l'un des facteurs est nul.

10.7.13.2. — Si V et W sont de dimension finie sur k il en va de même de $V \otimes_k W$ et $\dim(V \otimes_k W) = \dim V \cdot \dim W$.

10.7.14. *Aspect matriciel du produit tensoriel.* — Soient V, V', W et W' quatre k-espaces vectoriels. Donnons-nous une base (e_j) de V, une base (f_i) de V', une base (g_s) de W et une base (h_r) de W'. Soit φ une application k-linéaire de V dans V', et soit ψ une application k-linéaire de W dans W'. Soit (a_{ij}) la matrice de φ dans les bases (e_j) et (f_i), et soit (b_{rs}) la matrice de ψ dans les bases (g_s) et (h_r).

La proposition 10.7.12 assure que $(e_j \otimes g_s)_{j,s}$ est une base de $V \otimes_k W$, et que $(f_i \otimes h_r)_{i,r}$ est une base de $V' \otimes_k W'$. Soit (c_{irjs}) la matrice de $\varphi \otimes \psi$ dans ces deux bases. Le but de ce qui suit est de calculer explicitement la matrice (c_{irjs}). Pour cela, fixons j et s. On a les égalités

$$\begin{aligned}(\varphi \otimes \psi)(e_j \otimes g_s) &= \varphi(e_j) \otimes \psi(g_s) \\ &= (\sum_i a_{ij} f_i) \otimes (\sum_r b_{rs} h_r) \\ &= \sum_{i,r} a_{ij} b_{rs} f_i \otimes h_r.\end{aligned}$$

Il en résulte l'égalité $c_{irjs} = a_{ij} b_{rs}$ pour tout (i, r, j, s).

10.7.15. *Quelques objets classiques revisités.* — Soit k un corps et soient V et W deux k-espaces vectoriels.

10.7.15.1. — L'application de $V^* \times W$ dans $\mathrm{Hom}_k(V, W)$ qui envoie un couple (φ, w) sur l'endomorphisme $v \mapsto \varphi(v)w$ est k-bilinéaire ; elle induit donc une application k-linéaire $u \colon V^* \otimes_k W \to \mathrm{Hom}_k(V, W)$.

10.7.15.2. — L'application de $V^* \times V$ dans k qui envoie un couple (φ, v) sur $\varphi(v)$ est k-bilinéaire ; elle induit donc une forme k-linéaire $\tau \colon V^* \otimes_k V \to k$.

10.7.15.3. *Le cas de la dimension finie.* — Supposons que V et W sont de dimension finie. Choisissons une base $(f_1, \ldots, f_m)$ de V et une base $(e_1, \ldots, e_n)$ de W ; soit soit $(f_1^*, \ldots, f_m^*)$ la base duale de $(f_1, \ldots, f_m)$. La proposition 10.7.12 assure que $(f_j^* \otimes e_i)$ est une base de $V \otimes V^*$. Fixons deux indices i et j et posons $g_{ij} = u(f_j^* \otimes e_i)$. Pour tout ℓ on a alors $g_{ij}(f_\ell) = f_j^*(f_\ell)e_i = \delta_{\ell j} e_i$. Autrement dit, g_{ij} est l'application k-linéaire de V vers W dont la matrice dans les bases $(f_1, \ldots, f_m)$ et $(e_1, \ldots, e_n)$ est la matrice E_{ij} (dont le coefficient à la place (i, j) vaut 1, et dont tous les autres coefficients sont nuls).

Les (g_{ij}) forment donc une base de $\mathrm{Hom}_k(V, W)$. Ainsi, u transforme une base de $V^* \otimes_k W$ en une base de $\mathrm{Hom}_k(V, W)$; par conséquent, u est bijective. Notons que nous avons utilisées les bases choisies pour montrer que u est bijective, mais que la définition de u elle-même ne dépend pas de ces choix.

Plaçons-nous dans le cas particulier où $W = V$, et où $(f_j)_j = (e_i)_i$; l'application u induit alors un isomorphisme $V^* \otimes_k V \simeq \mathrm{End}_k(V)$. La composée $\tau \circ u^{-1}$ est une forme k-linéaire sur $\mathrm{End}_k(V)$, qui envoie pour tout (i, j) l'endomorphisme $g_{ij} = u(e_j^* \otimes e_i)$ sur $\tau(e_j^* \otimes e_i) = e_j^*(e_i) = \delta_{ij}$; il s'agit donc de la trace. On a ainsi donné une définition intrinsèque de celle-ci, *i.e.* indépendante de tout choix de base.

10.8. Quelques compléments. — Nous nous proposons de donner quelques autres applications du produit tensoriel, et notamment d'expliquer comment il peut être utilisé pour donner une description intrinsèque du déterminant (un peu dans l'esprit de celle de la trace vue en 10.7.15.3, mais en nettement plus délicate).

10.8.1. *Produit tensoriel de n espaces vectoriels.* — Soit n un entier $\geqslant 2$ et soient $V_1, \ldots, V_n$ des k-espaces vectoriels. L'associativité du produit tensoriel permet de définir $V_1 \otimes_k V_2 \ldots \otimes_k V_n$ (sans parenthèses). Cet espace est muni d'une application

n-linéaire (c'est-à-dire linéaire en chaque variable)

$$\begin{array}{rcl} V_1 \times \ldots \times V_n & \to & V_1 \otimes_k \ldots \otimes_k V_n \\ (v_1, \ldots, v_n) & \mapsto & v_1 \otimes \ldots \otimes v_n. \end{array}$$

L'espace vectoriel $V_1 \otimes_k V_2 \ldots \otimes_k V_n$ est engendré comme groupe abélien par les tenseurs purs $v_1 \otimes \ldots \otimes v_n$; si $(e_{ij})_j$ désigne pour tout i une base de V_i alors les éléments de la forme $e_{1j_1} \otimes \ldots \otimes e_{nj_n}$ forment une base de $V_1 \otimes_k V_2 \ldots \otimes_k V_n$.

Tout ceci s'étend au cas où n vaut 0 ou 1. Si $n = 1$ on n'a qu'un seul espace V_1, le produit tensoriel $V_1 \otimes_k V_2 \ldots \otimes_k V_n$ désigne alors simplement V_1, le produit cartésien $V_1 \otimes_k V_2 \ldots \otimes_k V_n$ désigne également V_1 et l'application $(v_1, \ldots, v_n) \mapsto v_1 \otimes \ldots \otimes v_n$ est simplement l'identité de V_1 (qui est bien 1-linéaire c'est-à-dire linéaire).

Si $n = 0$ on définit le produit tensoriel vide $V_1 \otimes \ldots \otimes V_n$ comme étant égal à k (qui est le «neutre» pour le produit tensoriel). Le produit cartésien vide $V_1 \times \ldots \times V_n$ est quant à lui l'espace nul $\{0\}$ et l'application $(v_1, \ldots, v_n) \mapsto v_1 \otimes \ldots \otimes v_n$ est alors $0 \mapsto 1$. Elle est 0-linéaire pour la bonne raison que cette condition est *vide*, puisqu'il n'y *aucune* composante à considérer.

Comme d'habitude, ces conventions ne sont pas introduites pour le plaisir d'être loufoque ni pour celui de se faire mal à la tête ; elles apportent au contraire justement assez de souplesse au formalisme pour ne pas avoir à distinguer sans cesse les cas particuliers $n = 0$ ou $n = 1$ dans les raisonnements. On a ainsi par exemple grâce à elles pour *tout* entier p compris entre 1 et n (et donc même pour $p = 1$ ou n) un isomorphisme canonique

$$(V_1 \otimes_k \ldots \otimes_k V_p) \otimes_k (V_{p+1} \otimes_k \ldots \otimes_k V_n) \simeq V_1 \otimes_k \ldots \otimes_k V_p \otimes V_{p+1} \otimes_k \ldots \otimes_k V_n$$

qui envoie $(v_1 \otimes \ldots \otimes v_p) \otimes (v_{p+1} \otimes \ldots \otimes v_n)$ sur $v_1 \otimes \ldots \otimes v_p \otimes v_{p+1} \otimes \ldots \otimes v_n$.

Proposition 10.8.2. — *Soit n un entier $\geqslant 0$ et soient $V_1, \ldots, V_n$ des k-espaces vectoriels. Soit P un k-espace vectoriel et soit $b \colon V_1 \times \ldots V_n \to P$ une application n-linéaire. Il existe une unique application k-linéaire $f \colon V_1 \otimes_k V_2 \ldots \otimes_k V_n \to P$ telle que le diagramme*

$$\begin{tikzcd} V_1 \times \ldots \times V_n \arrow[r, "b"] \arrow[d, "\otimes"'] & P \\ V_1 \otimes_k V_2 \ldots \otimes_k V_n \arrow[ur, "f"'] & \end{tikzcd}$$

commute.

Démonstration. — On procède par récurrence sur n. Si $n = 0$ on a $V_1 \times \ldots V_n = \{0\}$ et b est simplement une application quelconque de $\{0\}$ dans P, donnée par un élément p de P (à savoir $b(0)$). L'espace vectoriel $V_1 \otimes_k V_2 \ldots \otimes_k V_n$ est quant à lui égal à k et l'application $\otimes$ (la flèche verticale du diagramme) envoie 0 sur 1. Il suffit donc de s'assurer qu'il existe une unique application k-linéaire f de k dans P envoyant 1 sur p, ce qui est clair : c'est $\lambda \mapsto \lambda p$.

Supposons maintenant $n > 0$ et le résultat vrai pour $n-1$. L'unicité de l'application linéaire f est claire car les tenseurs purs engendrent $V_1 \otimes_k \ldots \otimes_k V_n$ et la commutativité du diagramme impose la formule $f(v_1 \otimes \ldots \otimes v_n) = b(v_1, \ldots, v_n)$. Montrons son existence. Pour tout $v \in V_n$ on dispose d'une application de $V_1 \times \ldots \times V_{n-1}$ vers P qui envoie $(v_1, \ldots, v_{n-1})$ sur $b(v_1, \ldots, v_{n-1}, v)$. Cette application est $(n-1)$-linéaire et induit donc par hypothèse de récurrence une application linéaire β_v du produit tensoriel $V_1 \otimes_k \ldots \otimes_k V_{n-1}$ vers P telle que $\beta_v(v_1 \otimes \ldots \otimes v_{n-1}) = b(v_1, \ldots, v_{n-1}, v)$ pour tout $(v_1, \ldots, v_{n-1})$. L'application β_v dépend linéairement de v (par l'égalité précédente, par la linéarité de b en la dernière variable, et par le fait que les tenseurs purs engendrent $V_1 \otimes_k \ldots \otimes_k V_{n-1}$). L'application de $(V_1 \otimes_k \ldots \otimes_k V_{n-1}) \times V_n$ dans P qui envoie (x, v) sur $\beta_v(x)$ est de ce fait bilinéaire, et induit en conséquence une application linéaire f de $(V_1 \otimes_k \ldots \otimes_k V_{n-1}) \otimes_k V_n$ dans P caractérisée par la formule $f(x \otimes v) = \beta_v(x)$ pour tout (x, v). En appliquant cette égalité lorsque x est de la forme $v_1 \otimes \ldots \otimes v_{n-1}$ et en identifiant $(V_1 \otimes_k \ldots \otimes_k V_{n-1}) \otimes_k V_n$ à $V_1 \otimes_k V_2 \ldots \otimes_k V_n$ il vient

$$f(v_1 \otimes v_2 \ldots \otimes v) = \beta_v(v_1 \otimes \ldots \otimes v_{n-1}) = b(v_1, \ldots, v_{n-1}, v),$$

ce qui achève la démonstration. □

10.8.3. *Quelques définitions.* — Nous allons maintenant présenter quelques notions générales d'algèbre que nous allons utiliser par la suite. Dans ce qui suit, *les anneaux ne sont pas supposés commutatifs.* Si A est un anneau, son centre $Z(A)$ est l'ensemble des éléments a de A tels que $ab = ba$ pour tout $b \in A$; c'est un sous-anneau de A qui est commutatif. Une *k-algèbre* est un anneau A muni d'un morphisme $k \to \mathrm{Z}(A)$.

10.8.3.1. *Idéaux bilatères.* — Soit A un anneau. Un *idéal bilatère* de A est un sous-groupe I de $(A, +)$ tel que $ai \in I$ et $ia \in I$ pour tout $i \in I$ et tout $a \in A$.

Si I est un idéal bilatère de A, la multiplication de A passe au quotient par I et fait de A/I un anneau (exercice !). Le morphisme quotient $\pi \colon A \to A/I$ est un morphisme d'anneaux et il satisfait la propriété universelle que vous imaginez probablement : pour tout anneau B, l'appplication $f \mapsto f \circ \pi$ établit une bijection entre l'ensemble des morphismes d'anneaux de A/I dans B et l'ensemble des morphismes d'anneaux de A vers B s'annulant sur I.

Si X est une partie de A, il existe un plus petit idéal bilatère contenant X , qui est dit engendré par X ; c'est l'ensemble des sommes (finies) $\sum a_i x_i b_i$ avec $x_i \in X$ et a_i et b_i dans A pour tout i.

Si A a une structure de k-algèbre, A/I en hérite (par composition des morphismes $k \to A \to A/I$).

10.8.3.2. *Anneaux gradués.* — Un anneau *gradué* est un anneau A muni d'une *graduation*, c'est-à-dire d'une écriture comme une somme directe $\bigoplus_{n \in \mathbf{N}} A_n$ telle que $1 \in A_0$ et $A_n A_m \subset A_{n+m}$ pour tout (n, m).

Soit $A = \bigoplus A_n$ un anneau gradué. On dit que A_n est l'ensemble des éléments *homogènes de degré n* de A ; si $a \in A$ s'écrit $\sum_n a_n$ avec $a_n \in A_n$ pour tout n, on dit que les a_n sont les *composantes homogènes* de a.

Un idéal bilatère I de A est dit *homogène* s'il est engendré par un ensemble $X \subset A$ composé d'éléments homogènes. Si c'est le cas, les composantes homogènes de tout élément de I appartiennent à I. En effet, il suffit de le vérifier pour un élément de I de la forme axb où x appartient à X (et est donc homogène d'un certain degré d)

et où $a = \sum a_n$ et $b = \sum b_n$ sont des éléments de A ; or la composante homogène de degré p d'un tel élément est $\sum_{n+m=p-d} a_n x b_m$, qui appartient à I (notons que si elle est non nulle on a forcément $p \geqslant d$).

On a donc $I = \bigoplus_n (I \cap A_n)$. L'anneau quotient A/I hérite alors d'une graduation naturelle $A/I = \bigoplus A_n/(I \cap A_n)$.

10.8.3.3. *Notion de k-algèbre graduée.* — Une *k-algèbre graduée* est un anneau gradué $A = \bigoplus A_n$ muni d'un morphisme d'anneaux de k dans $\mathrm{Z}(A)$ prenant ses valeurs dans A_0.

10.8.4. *L'algèbre tensorielle.* — Soit V un k-espace vectoriel. On note TV la somme directe $\bigoplus_{n \in \mathbf{N}} V^{\otimes n}$ où $V^{\otimes n}$ désigne le produit tensoriel de n copies de V (avec les conventions de la section précédente lorsque $n = 0$ ou $n = 1$).

Soient n et m deux entiers. La composée de l'isomorphisme canonique

$$V^{\otimes n} \otimes_k V^{\otimes m} \simeq V^{\otimes (n+m)}$$

et de l'application bilinéaire $\otimes \colon V^{\otimes n} \times V^{\otimes m} \to V^{\otimes n} \otimes_k V^{\otimes m}$ fournit une application bilinéaire $V^{\otimes n} \times V^{\otimes m} \to V^{\otimes (n+m)}$ qui envoie $((x_1 \otimes \ldots \otimes x_n), (y_1 \otimes \ldots \otimes y_m))$ sur $x_1 \otimes \ldots \otimes x_n \otimes y_1 \otimes \ldots \otimes y_m$ et qu'on notera encore $\otimes$. La collection (pour n et m variables) de ces applications définit une application bilinéaire $\otimes \colon TV \times TV \to TV$; on vérifie immédiatement sur les tenseurs purs qu'elle est associative (elle consiste essentiellement à les concaténer). Elle fait donc de TV une k-algèbre (son élément neutre multiplicatif est $1 \in k = T^0V$) qu'on appelle *l'algèbre tensorielle* de V.

Posons $T^nV = V^{\otimes n}$. La décomposition $TV = \bigoplus_n T^nV$ fait de TV une k-algèbre graduée. Notons que $T^1V = V$ - on identifiera désormais V par ce biais à un sous-espace vectoriel de TV.

***Proposition 10.8.5* (Propriété universelle de l'algèbre tensorielle)**
Soit A une k-algèbre et soit $f \colon V \to A$ une application k-linéaire. Il existe un unique morphisme de k-algèbres de TV dans A dont la restriction à V est égale à f.

Démonstration. — La description explicite de TV montre qu'elle est engendrée par V comme k-algèbre ; par conséquent, tout morphisme de k-algèbres de TV dans A est déterminé par sa restriction à V, d'où l'unicité.

Montrons maintenant l'existence. Pour tout entier n l'application de V^n dans A donnée par la formule

$$(v_1, \ldots, v_n) \mapsto \underbrace{f(v_1)f(v_2)\ldots f(v_n)}_{\text{produit calculé dans l'algèbre } A}$$

est n-linéaire. Elle induit donc en vertu de la proposition 10.8.2 une application k-linéaire f_n de $V^{\otimes n}$ dans A qui vérifie la formule

$$f_n(v_1 \otimes \ldots \otimes v_n) = f(v_1)f(v_2)\ldots f(v_n).$$

On vérifie alors que l'application de $TV = \bigoplus T^nV$ vers A qui envoie un élément $a = \sum_n a_n$ sur $\sum_n f_n(a_n)$ répond aux conditions de l'énoncé. □

Commentaires 10.8.6. — On peut reformuler la proposition universelle ci-dessus sous forme du slogan suivant : *se donner un morphisme de TV dans une k-algèbre A, c'est se donner une application k-linéaire de V dans A.*

Intuitivement, l'algèbre tensorielle de V est la k-algèbre la plus générale fabriquée à partir de V.

10.8.7. *L'algèbre symétrique.* — Soit V un k-espace vectoriel. Soit I l'idéal bilatère de TV engendré par les éléments $x \otimes y - y \otimes x$ pour x et y parcourant V. L'idéal bilatère I étant engendré par des éléments homogènes de degré 2, il est homogène et satisfait de surcroît les égalités $I \cap T^0V = I \cap T^1V = \{0\}$.

On appelle *algèbre symétrique de* V et l'on note SV, la k-algèbre graduée quotient $TV/I = \bigoplus_n T^nV/(I \cap T^nV)$; le sommande $T^nV/(I \cap T^nV)$ est noté S^nV. On a $S^0V = k$ et $S^1V = V$; on identifie désormais V par ce biais à un sous-espace vectoriel de SV (le morphisme quotient $TV \to SV$ induit un isomorphisme $T^1V \simeq S^1V$ et est donc compatible aux plongements de V dans les deux algèbres). La k-algèbre TV étant engendrée par V, l'algèbre SV l'est également. Or si x et y sont deux éléments de V, l'image de $x \otimes y - y \otimes x$ dans SV est nulle ; autrement dit, x et y commutent lorsqu'on les voit comme appartenant à $SV \supset V$; puisque V engendre SV, celle-ci est une k-algèbre commutative. Nous noterons son produit comme un produit usuel.

Proposition 10.8.8 **(Propriété universelle de l'algèbre symétrique)**
Soit A une k-algèbre commutative et soit $f: V \to A$ une application k-linéaire. Il existe un unique morphisme de k-algèbres de SV dans A dont la restriction à V est égale à f.

Démonstration. — Comme V engendre SV comme k-algèbre, un morphisme de k-algèbres de SV dans A est entièrement déterminé par sa restriction à V, d'où l'unicité.

Montrons maintenant l'existence. La propriété universelle de TV assure qu'il existe un morphisme de k-algèbres $\varphi: TV \to A$ tel que $\varphi|_V = f$. Comme A est commutative on a $\varphi(x \otimes y - y \otimes x) = f(x)f(y) - f(y)f(x) = 0$ pour tout $(x, y) \in V^2$. Ainsi φ passe au quotient par I et l'application induite de SV dans A répond aux conditions de l'énoncé. $\square$

Commentaires 10.8.9. — On peut reformuler la proposition universelle ci-dessus sous forme du slogan suivant : *se donner un morphisme de SV dans une k-algèbre commutative A, c'est se donner une application k-linéaire de V dans A.*

Intuitivement, l'algèbre symétrique de V est la k-algèbre commutative la plus générale fabriquée à partir de V.

10.8.10. — Soit V un k-espace vectoriel. Pour tout entier n l'application de V^n dans S^nV qui envoie un n-uplet $(v_1, \dots, v_n)$ sur le produit $v_1 \cdots v_n$ est n-linéaire symétrique (c'est-à-dire invariante par permutation des facteurs) par commutativité de SV. La proposition ci-dessous montre que celle opération est en un sens l'opération n-linéaire symétrique de source V^n la plus générale.

Proposition 10.8.11 **(Propriété universelle de S^nV)**
Soit V un k-espace vectoriel et soit n un entier $\geqslant 0$. Soit P un k-espace

vectoriel et soit $b\colon V^n \to P$ une application n-linéaire symétrique. Il existe une unique application linéaire f de S^nV dans P telle que $f(v_1 \cdots v_n) = b(v_1, \dots, v_n)$ pour tout $(v_1, \dots, v_n) \in V^n$.

Démonstration. — Puisque T^nV est engendré comme k-espace vectoriel par les tenseurs purs à n facteurs, S^nV est engendré comme k-espace vectoriel par les produits de n éléments de V, d'où l'unicité de f. Montrons maintenant son existence. La propriété universelle de $V^{\otimes n}$ (prop. 10.8.2) fournit une application k-linéaire $g\colon V^{\otimes n} \to P$ telle que $g(v_1 \otimes \ldots \otimes v_n) = b(v_1, \dots, v_n)$ tout $(v_1, \dots, v_n) \in V^n$.

Montrons que $I \cap T^nV \subset \mathrm{Ker}(g)$. Tout élément de $I \cap T^nV$ est somme de termes de la forme

$$v_1 \otimes \ldots \otimes v_m \otimes x \otimes y \otimes w_1 \otimes \ldots \otimes w_{m'} - v_1 \otimes \ldots \otimes v_m \otimes y \otimes x \otimes w_1 \otimes \ldots \otimes w_{m'}$$

où les v_i et les w_i ainsi que x et y appartiennent à V (et où $m + m' + 2 = n$). Or lorsqu'on applique g à un tel terme on trouve

$$b(v_1, \dots, v_m, x, y, w_1, \dots, w_{m'}) - b(v_1, \dots, v_m, y, x, w_1, \dots, w_{m'})$$

qui vaut 0 car b est symétrique. Par conséquent g passe au quotient par $I \cap T^nV$ et l'application induite $f\colon S^nV \to P$ répond aux conditions de l'énoncé. $\square$

10.8.12. *Algèbre symétrique et polynômes.* — Soit V un k-espace vectoriel et soit $(e_\ell)_{\ell\in\Lambda}$ une base de V. Soit A la k-algèbre des polynômes en une famille (T_ℓ) d'indéterminées indexée par Λ (notons que Λ peut être infini mais chaque polynôme appartenant à A ne fait intervenir qu'un nombre fini d'indéterminées). On dispose d'un morphisme d'évaluation $\varphi\colon P \mapsto P(e_\ell)_\ell$ de A dans SV. Par ailleurs, la propriété universelle de l'algèbre symétrique assure que l'application k-linéaire $\sum x_\ell e_\ell \mapsto \sum x_\ell T_\ell$ de V dans A se prolonge d'une unique manière en un morphisme de k-algèbres $\psi\colon SV \to A$. On vérifie que φ et ψ sont inverses l'un de l'autre (il suffit de s'en assurer sur les T_ℓ d'une part et les e_ℓ d'autre part). Par conséquent, $SV \simeq k[T_\ell]_\ell$.

Une algèbre symétrique est donc une algèbre de polynômes, mais non canoniquement : l'identification requiert le choix d'une base. (On aurait pu de même montrer que le choix d'une base permet d'identifier l'algèbre tensorielle à une algèbre de «polynômes non commutatifs»).

Supposons V de dimension finie d. Par ce qui précède, une fois qu'on a choisi une base $(e_1, \dots, e_d)$ de V on dispose pour tout n d'un isomorphisme entre l'espace vectoriel S^nV et l'espace vectoriel des polynômes en $T_1, \dots, T_d$ homogènes de degré n. Nous laissons le lecteur démontrer à titre d'exercice que sa dimension est $\binom{d+n-1}{n}$ (si $d + n > 0$; dans le cas $d = n = 0$ la formule donnée n'a pas de sens, mais la dimension est 1 puisque $S^0V = k$).

10.8.13. *L'algèbre extérieure.* — Soit V un k-espace vectoriel. Soit J l'idéal bilatère de TV engendré par les éléments $x \otimes x$ pour x parcourant V. L'idéal bilatère J étant engendré par des éléments homogènes de degré 2, il est homogène et satisfait de surcroît les égalités $J \cap T^0V = J \cap T^1V = \{0\}$.

On appelle *algèbre extérieure de* V et l'on note ΛV, la k-algèbre graduée quotient $TV/J = \bigoplus_n T^nV/(J \cap T^nV)$; le sommande $T^nV/(J \cap T^nV)$ est noté Λ^nV. On a $\Lambda^0V = k$ et $\Lambda^1V = V$; on identifie désormais V par ce biais à un sous-espace vectoriel de ΛV (le morphisme quotient $TV \to \Lambda V$ induit un isomorphisme $T^1 V \simeq \Lambda^1 V$ et est donc compatible aux plongements de V dans les deux algèbres). La k-algèbre TV étant engendrée par V, l'algèbre ΛV l'est également. Son produit est appelé le *produit extérieur* et est noté par le symbole $\wedge$ (le même que celui du produit vectoriel ; on le lit souvent «extérieur», ou «wedge» si on est d'humeur anglophone).

10.8.14. — On a $x \wedge x = 0$ pour tout $x \in V$, puisque $x \otimes x \in J$. Ceci entraîne que $x \wedge y = -y \wedge x$ pour tout $(x, y) \in V$. On a en effet $(x+y) \wedge (x+y) = 0$; en développant il vient

$$x \wedge x + x \wedge y + y \wedge x + y \wedge y = 0,$$

et l'égalité cherchée découle du fait que $x \wedge x = y \wedge y = 0$.

Soient $x_1, \ldots, x_n$ des éléments de V et soit $\sigma \in S_n$ On peut écrire σ comme composée de transpositions de la forme $(i \;\; i+1)$ (exercice !) ; on en déduit à l'aide de ce qui précède que $x_{\sigma(1)} \wedge \ldots \wedge x_{\sigma(n)} = \varepsilon(\sigma) x_1 \wedge \ldots \wedge x_n$.

Supposons qu'il existe $i \neq j$ avec $x_i = x_j$; en choisissant σ de sorte que $\sigma(i) = i$ et $\sigma(j) = i + 1$ on voit que $x_1 \wedge \ldots \wedge x_n = \varepsilon(\sigma) x_{\sigma(1)} \wedge \ldots \wedge x_{\sigma(n)} = 0$.

Par n-linéarité il s'ensuit plus généralement que si l'un des x_i est combinaison linéaire des autres alors $x_1 \wedge \ldots \wedge x_n = 0$.

Proposition 10.8.15 **(Propriété universelle de l'algèbre extérieure)**
Soit A une k-algèbre et soit $f \colon V \to A$ une application k-linéaire telle que $f(x)^2 = 0$ pour tout $x \in V$. Il existe un unique morphisme de k-algèbres de ΛV dans A dont la restriction à V est égale à f.

Démonstration. — Comme V engendre ΛV comme k-algèbre, un morphisme de k-algèbres de ΛV dans A est entièrement déterminé par sa restriction à V, d'où l'unicité.

Montrons maintenant l'existence. La propriété universelle de TV assure qu'il existe un morphisme de k-algèbres $\varphi \colon TV \to A$ tel que $\varphi|_V = f$. Comme $f(x)^2 = 0$ pour tout $x \in V$ on a $\varphi(x \otimes x) = 0$ pour tout $x \in V$. Ainsi φ passe au quotient par J et l'application induite de ΛV dans A répond aux conditions de l'énoncé. □

Commentaires 10.8.16. — On peut reformuler la proposition universelle ci-dessus sous forme du slogan suivant : *se donner un morphisme de ΛV dans une k-algèbre A, c'est se donner une application k-linéaire de V dans A telle que l'image de tout élément de V soit de carré nul.*

Intuitivement, l'algèbre extérieure V est la k-algèbre la plus générale fabriquée à partir de V en requérant de sucroît que chaque élément de V soit de carré nul.

10.8.17. — Soit V un k-espace vectoriel. Pour tout entier n l'application de V^n dans $\Lambda^n V$ qui envoie un n-uplet $(v_1, \ldots, v_n)$ sur le produit $v_1 \wedge \ldots \wedge v_n$ est n-linéaire *alternée*, c'est-à-dire qu'elle s'annule sur tout n-uplet ayant au moins deux coordonnées identiques (10.8.14). La proposition ci-dessous montre que celle opération est en un sens l'opération n-linéaire alternée de source V^n la plus générale.

Proposition 10.8.18 **(Propriété universelle de $\Lambda^n V$)**
Soit V un k-espace vectoriel et soit n un entier $\geqslant 0$. Soit P un k-espace vectoriel et soit $b \colon V^n \to P$ une application n-linéaire alternée. Il existe une unique

application linéaire f de $\Lambda^n V$ dans P telle que $f(v_1 \wedge \ldots \wedge v_n) = b(v_1, \ldots, v_n)$ pour tout $(v_1, \ldots, v_n) \in V^n$.

Démonstration. — Puisque $T^n V$ est engendré comme k-espace vectoriel par les tenseurs purs à n facteurs, $\Lambda^n V$ est engendré comme k-espace vectoriel par les produits extérieurs de n éléments de V, d'où l'unicité de f. Montrons maintenant son existence. La propriété universelle de $V^{\otimes n}$ (prop. 10.8.2) fournit une application k-linéaire $g\colon V^{\otimes n} \to P$ telle que $g(v_1 \otimes \ldots \otimes v_n) = b(v_1, \ldots, v_n)$ tout $(v_1, \ldots, v_n) \in V^n$.

Montrons que $J \cap T^n V \subset \operatorname{Ker}(g)$. Tout élément de $J \cap T^n V$ est somme de termes de la forme $v_1 \otimes \ldots \otimes v_m \otimes x \otimes x \otimes w_1 \otimes \ldots \otimes w_{m'}$ où les v_i et les w_i ainsi que x appartiennent à V (et où $m + m' + 2 = n$). Or lorsqu'on applique g à un tel terme on trouve $b(v_1, \ldots, v_m, x, x, w_1, \ldots, w_{m'})$ qui vaut 0 car b est alternée. Par conséquent g passe au quotient par $J \cap T^n V$ et l'application induite $f\colon \Lambda^n V \to P$ répond aux conditions de l'énoncé. □

10.8.19. *Algèbre extérieure en dimension finie.* — Soit V un k-espace vectoriel de dimension finie d. Choisissons une base $(e_1, \ldots, e_d)$ de V. Soit n un entier. Tout élément de l'espace vectoriel $\Lambda^n V$ est somme de termes de la forme $v_1 \wedge \ldots \wedge v_n$. En écrivant chacun des v_i dans la base $(e_1, \ldots, e_d)$, en utilisant la n-linéarité et les formules vues au 10.8.14, on voit qu'un tel élément est lui-même combinaison linéaires de vecteurs de la forme $e_{i_1} \wedge \ldots \wedge e_{i_n}$ où $i_1 < i_2 < \ldots < i_n$ est une suite strictement croissante de n éléments de $\{1, \ldots, d\}$. La famille $(e_{i_1} \wedge \ldots \wedge e_{i_n})_{i_1 < i_2 < \ldots < i_n}$ engendre donc $\Lambda^n V$. L'ensemble des suites strictement croissantes de n éléments de $\{1, \ldots, d\}$ est en bijection avec l'ensemble des parties à n éléments de $\{1, \ldots, d\}$ (on associe à une telle partie la suite de ses éléments donnés dans l'ordre croissant) ; il y en a donc $\binom{d}{n}$. Par conséquent, la dimension de $\Lambda^n V$ est majorée par $\binom{d}{n}$, avec égalité si et seulement si la famille $(e_{i_1} \wedge \ldots \wedge e_{i_n})_{i_1 < i_2 < \ldots < i_n}$ est libre (et en est donc une base). Notons deux conséquences importantes de ce fait :

- ◇ $\Lambda^n V$ est nul dès que $n > d$;
- ◇ $\dim \Lambda V = \sum_n \dim \Lambda^n V \leqslant \sum_n \binom{d}{n} = 2^d$, avec égalité si et seulement si $\dim \Lambda^n V = \binom{d}{n}$ pour tout n, c'est-à-dire encore si et seulement si la famille $(e_{i_1} \wedge \ldots \wedge e_{i_n})_{i_1 < i_2 < \ldots < i_n}$ est une base de $\Lambda^n V$ pour tout n.

Nous nous proposons maintenant de montrer qu'on a effectivement $\dim \Lambda V = 2^d$; puisqu'on sait que $\dim \Lambda V \leqslant 2^d$, il suffit d'exhiber une surjection linéaire de ΛV vers un k-espace vectoriel de dimension 2^d. C'est l'objet de ce qui suit.

10.8.20. *Construction d'une algèbre auxiliaire et application.* — Soit d un entier, soit r un entier et soient $E_1, \ldots, E_r$ des sous-ensembles deux à deux disjoints de $\{1, \ldots, d\}$; notons E leur réunion. Notons $\leqslant$ l'ordre usuel sur les entiers, et notons $\preceq$ l'ordre sur E défini comme suit : si $x \in E_i$ et $y \in E_j$ alors $x \preceq y$ si et seulement si $i < j$ ou $i = j$ et $x \leqslant y$.

Notons $\sigma_{E_1, \ldots, E_r}$ la permutation de E qui pour tout entier ℓ compris entre 1 et $|E|$ associe au ℓ-ième élément de E pour l'ordre $\leqslant$ son ℓ-ième élément pour l'ordre $\preceq$. Intuitivement, $\sigma_{E_1, \ldots, E_r}$ consiste à partir des éléments de E rangés par ordre croissant, et à les réordonner en plaçant d'abord les éléments de E_1 (rangés dans l'ordre croissant), puis ceux de E_2 (rangés dans l'ordre croissant), etc.

Donnons un exemple : on prend $d \geqslant 17$ et $E_1 = \{3, 7, 14\}, E_2 = \{4, 6, 8, 13\}$ et $E_3 = \{5, 9, 10, 12, 17\}$. La permutation σ_{E_1,E_2,E_3} est alors égale à

$$\begin{pmatrix} 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 12 & 13 & 14 & 17 \\ 3 & 7 & 14 & 4 & 6 & 8 & 13 & 5 & 9 & 10 & 12 & 17 \end{pmatrix}.$$

Si E, F et G sont trois sous-ensembles deux à deux disjoints de $\{1, \ldots, d\}$ il découle des définitions que

$$\sigma_{E,F,G} = (\sigma_{E,F} \coprod \mathrm{Id}_G) \circ \sigma_{E \coprod F, G} = (\mathrm{Id}_E \coprod \sigma_{F,G}) \circ \sigma_{E, F \coprod G},$$

où $\sigma_{E,F} \coprod \mathrm{Id}_G$ désigne la permutation de $E \coprod F \coprod G$ qui vaut $\sigma_{(E,F)}$ sur $E \coprod F$ et l'identité sur G, et de même pour $\mathrm{Id}_E \coprod \sigma_{F,G}$. En particulier, on a

$$(*) \quad \varepsilon(\sigma_{E,F})\varepsilon(\sigma_{E \coprod F,G}) = \varepsilon(\sigma_{(F,G)})\varepsilon(\sigma_{E,F \coprod G}) = \varepsilon(\sigma_{E,F,G}).$$

10.8.20.1. — Soit A le k-linéarisé de l'ensemble $\mathscr{P}(\{1, \ldots, d\})$. On définit sur A une loi de composition interne bilinéaire par les formules

$$[E] \cdot [F] = \begin{cases} 0 & \text{si } E \cap F \neq \varnothing \\ \varepsilon(\sigma_{E,F})[E \coprod F] & \text{sinon.} \end{cases}$$

Cette loi est associative. En effet, il suffit de le vérifier sur les éléments de la forme $[E]$ avec $E \in \mathscr{P}(\{1, \ldots, d\})$. Or si E, F et G sont trois parties de $\{1, \ldots, d\}$ on a par construction $[E] \cdot ([F] \cdot [G]) = ([E] \cdot [F]) \cdot [G] = 0$ si E, F et G ne sont pas deux à deux disjoints ; et s'ils le sont on a

$$[E] \cdot ([F] \cdot [G]) = \varepsilon(\sigma_{E,F \coprod G})\varepsilon(\sigma_{F,G})[E \coprod F \coprod G]$$

et

$$([E] \cdot [F]) \cdot [G] = \varepsilon(\sigma_{E \coprod F,G})\varepsilon(\sigma_{E,F})[E \coprod F \coprod G]$$

et on conclut à l'aide de $(*)$.

Par ailleurs, celle loi possède un élément neutre, à savoir $[\varnothing]$. Elle fait donc de A une k-algèbre.

Soit E une partie de $\{1, \ldots, d\}$. Soient $i_1 < \ldots < i_n$ les éléments de E rangés dans l'ordre croissant. On a alors par définition de la multiplication sur l'algèbre A l'égalité $[E] = [i_1] \cdot [i_2] \cdot \ldots \cdot [i_n]$ (en écrivant par abus $[i]$ au lieu de $[\{i\}]$ pour tout $i \in \{1, \ldots, d\}$). Il s'ensuit, la famille $([E])_{E \in \mathscr{P}(\{1,\ldots,d\})}$ étant une base de A, que A est engendrée comme k-algèbre par les éléments de la forme $[i]$ avec $i \in \{1, \ldots, d\}$.

Soit B le sous-espace vectoriel de A de base $([i])_{1\leqslant i\leqslant d}$ et soit $b = \sum \lambda_i[i]$ un élément de B. On a

$$\begin{aligned}
b^2 &= (\sum_i \lambda_i[i])^2 \\
&= \sum_{i,j} \lambda_i\lambda_j[i]\cdot[j] \\
&= \sum_i \lambda_i^2 \underbrace{[i]\cdot[i]}_{0} + \sum_{i<j} \lambda_i\lambda_j([i]\cdot[j]+[j]\cdot[i]) \\
&= \sum_{i<j} \lambda_i\lambda_j([\{i,j\}]-[\{i,j\}]) \\
&= 0.
\end{aligned}$$

10.8.20.2. — Soit V un k-espace vectoriel de dimension d et soit $(e_1,\ldots,e_d)$ une base de V. Soit f l'application k-linéaire de V vers A qui envoie e_i sur $[i]$ pour tout i. On a $f(x)\in B$ pour tout $x\in V$, si bien que $f(x)^2=0$ pour tout $x\in V$ par ce qui précède. La propriété universelle de l'algèbre extérieure assure alors que f s'étend de manière unique en un morphisme de k-algèbres $\varphi\colon \Lambda V\to A$. Puisque $f(e_i)=[i]$ pour tout i et puisque les éléments de la forme $[i]$ avec $i\in\{1,\ldots,d\}$ engendrent A, l'application φ est surjective. Or la dimension de A est égale au cardinal de $\mathscr{P}(\{1,\ldots,d\})$, c'est-à-dire à 2^d. On en déduit en vertu de 10.8.19 que $\dim \Lambda V = 2^d$ (et que $\varphi\colon \Lambda V\to A$ est un isomorphisme). Il s'ensuit (là encore d'après 10.8.19) que pour tout entier n l'espace vectoriel $\Lambda^n V$ est de dimension $\binom{d}{n}$ et que $(e_{i_1}\wedge\ldots\wedge e_{i_n})_{i_1<i_2<\ldots<i_n}$ en est une base.

10.8.21. *Applications induites par une application linéaire.* — Soit $f\colon V\to W$ une application k-linéaire entre k-espaces vectoriels.

10.8.21.1. — Pour tout n l'application

$$V^n\to W^{\otimes n},\ \ (v_1,\ldots,v_n)\mapsto f(v_1)\otimes\ldots f(v_n)$$

est n-linéaire ; elle induit donc une application k-linéaire $f^{\otimes n}$ de $V^{\otimes n}$ vers $W^{\otimes n}$, qui envoie $v_1\otimes\ldots v_n$ sur $f(v_1)\otimes\ldots f(v_n)$ pour tout $(v_1,\ldots,v_n)$. La collection des $f^{\otimes n}$ induit une application $Tf\colon TV\to TW$ qui par construction respecte les graduations (c'est-à-dire qu'elle envoie T^nV dans T^nW pour tout n – l'application $T^nV\to T^nW$ qu'elle induit est précisément $f^{\otimes n}$, qu'on notera aussi T^nf). On vérifie aussitôt que Tf est un morphisme de k-algèbres (car elle commute visiblement à la «concaténation des tenseurs»).

Il résulte de la construction que $T\mathrm{Id}_V=\mathrm{Id}_{TV}$, que $T(f\circ g)=Tf\circ Tg$ lorsque cela a un sens, et que si f est bijective alors Tf est bijective d'inverse $T(f^{-1})$; on a pour tout n des énoncés analogues avec T^n au lieu de T.

10.8.21.2. — Comme $Tf(x\otimes y-y\otimes x)=f(x)\otimes f(y)-f(y)\otimes f(x)$ pour tout $(x,y)\in V^2$, l'application Tf induit un morphisme de k-algèbres $Sf\colon SV\to SW$, qui préserve les graduations et induit par conséquent pour tout n une application k-linéaire S^nf de S^nV vers S^nW. On a $S\mathrm{Id}_V=\mathrm{Id}_{SV}$, $S(f\circ g)=Sf\circ Sg$ lorsque cela a un sens, et si f est bijective alors Sf est bijective d'inverse $S(f^{-1})$; on a pour tout n des énoncés analogues avec S^n au lieu de S.

10.8.21.3. — Comme $Tf(x \otimes x) = f(x) \otimes f(x)$ pour tout $x \in V$, l'application Tf induit un morphisme de k-algèbres $\Lambda f \colon \Lambda V \to \Lambda W$, qui préserve les graduations et induit par conséquent pour tout n une application k-linéaire $\Lambda^n f$ de $\Lambda^n V$ vers $\Lambda^n W$. On a $\Lambda \mathrm{Id}_V = \mathrm{Id}_{\Lambda V}$, $\Lambda(f \circ g) = \Lambda f \circ \Lambda g$ lorsque cela a un sens, et si f est bijective alors Λf est bijective d'inverse $\Lambda(f^{-1})$; on a pour tout n des énoncés analogues avec Λ^n au lieu de Λ.

10.8.22. *La définition intrinsèque du déterminant.* — Soit V un k-espace vectoriel de dimension finie d. D'après 10.8.20.2, $\Lambda^d V$ est de dimension $\binom{d}{d} = 1$. Il en résulte que $\mathrm{End}_k(\Lambda^d V) = k$ (tout endomorphisme k-linéaire de $\Lambda^d V$ est une homothétie). Pour tout endomorphisme f de V, l'endomorphisme $\Lambda^d f$ de $\Lambda^d V$ peut donc être vu comme un scalaire, que l'on appelle le *déterminant* de f et qu'on note $\det(f)$.

Comme $\Lambda^d \mathrm{Id}_V = \mathrm{Id}_{\Lambda^d V}$, on a $\det(\mathrm{Id}_{\mathrm{V}}) = 1$; si f et g sont deux endomorphismes de V, l'égalité $\Lambda^d(f \circ g) = \Lambda^d f \circ \Lambda^d g$ se récrit $\det(f \circ g) = \det(f) \cdot \det(g)$; on déduit de ce qui précède que si f est bijective alors $\det(f) \neq 0$ et $(\det(f))^{-1} = \det(f^{-1})$.

10.8.22.1. *Lien avec la définition classique.* — On vient donc de donner une définition intrinsèque du déterminant, qui ne fait pas appel au choix d'une base. Vérifions qu'elle est compatible avec la définition traditionnelle. Fixons donc une base $(e_1, \ldots, e_d)$ de V ; soit (a_{ij}) la matrice de f dans cette base.

La droite vectorielle $\Lambda^d V$ a pour base $e_1 \wedge \ldots \wedge e_d$ et l'on a

$$\det(f) e_1 \wedge \ldots \wedge e_d = \Lambda^d f(e_1 \wedge \ldots \wedge e_d) = f(e_1) \wedge \ldots \wedge f(e_d)$$

(la première égalité découle de la définition du déterminant et la seconde de celle de $\Lambda^d f$). En remplaçant pour tout j le terme $f(e_j)$ par $\sum_i a_{ij} e_i$, en développant par n-linéarité, et en utilisant le fait que $e_{\sigma(1)} \wedge \ldots \wedge e_{\sigma(n)} = \varepsilon(\sigma) e_1 \wedge \ldots \wedge e_n$ pour toute permutation σ de $\{1, \ldots, n\}$ et qu'un produit extérieur de plusieurs vecteurs est nul dès que deux d'entre eux coïncident (10.8.14), on obtient la formule usuelle pour $\det(f)$.

10.8.22.2. — On a mentionné en 10.8.22 que si f est bijective alors $\det(f) \neq 0$. Montrons la réciproque. On raisonne par contraposition ; on suppose donc que f n'est pas bijective. La famille $(f(e_1), \ldots, f(e_d))$ est alors liée ; il résulte dès lors de 10.8.14 que

$$\Lambda^d f(e_1 \wedge \ldots \wedge e_d) = f(e_1) \wedge \ldots \wedge f(e_d) = 0,$$

et partant que $\det(f) = 0$.

10.8.22.3. *À propos des $\Lambda^n f$.* — Soit n un entier. La famille $(e_{i_1} \wedge \ldots \wedge e_{i_n})_{i_1 < \ldots < i_n}$ est une base de $\Lambda^n V$. Nous laissons le lecteur vérifier (par le même type de calcul qu'au 10.8.22.1) que le coefficient d'indice $((i_1 < \ldots < i_n), (j_1 < \ldots < j_n))$ de la matrice de $\Lambda^n f$ dans cette base est le déterminant (au sens classique !) de la matrice (n, n) extraite de (a_{ij}) définie par les colonnes $j_1, \ldots, j_n$ et les lignes $i_1, \ldots, i_n$.

11. Représentations des groupes : généralités

11.1. Définition, premiers exemples, premières propriétés. — Nous allons introduire dans ce chapitre la notion de *représentation* d'un groupe, qui est essentiellement la variante linéaire de celle d'opération sur un ensemble. Nous établirons un certain nombre de propriétés valables dans un cadre très général (pour des groupes et des corps souvent quelconques) ; nous consacrerons ensuite un chapitre spécifique au cas des représentations complexes des groupes finis .

On fixe un corps k et un groupe G.

Définition 11.1.1. — Une *représentation k-linéaire de G* est un couple (V, ρ) où V est un k-espace vectoriel et ρ un morphisme de groupes de G dans le groupe $\mathrm{GL}(V)$ des automorphismes k-linéaires de V.

Une *représentation matricielle* de G (à coefficients dans k) est un morphisme de groupes de G dans $\mathrm{GL}_I(k)$, pour un certain ensemble I.

Commentaires 11.1.2. — Il résulte de la discussion du 4.1.4 que se donner une représentation k-linéaire de G revient à se donner un k-espace vectoriel V et une opération de G sur V *par automorphismes k-linéaires.*

11.1.3. *Expression matricielle d'une représentation.* — Soit (V, ρ) une représentation k-linéaire de G. Soit $\mathscr{B} = (e_i)_{i \in i}$ une base de V. La donnée de la base $\mathscr{B}$ fournit un isomorphisme $\iota_{\mathscr{B}} \colon \mathrm{GL}(V) \simeq \mathrm{GL}_I(k)$, et la composée $\iota_{\mathscr{B}} \circ \rho$ définit une représentation matricielle de G, que l'on appellera *l'expression de ρ dans la base $\mathscr{B}$.*

Soit $\pi \colon G \to \mathrm{GL}_I(k)$ une représentation matricielle de G. Il existe alors une unique représentation de G d'espace sous-jacent V dont π est l'expression dans la base $\mathscr{B}$; elle est donnée par le morphisme $\iota_{\mathscr{B}}^{-1} \circ \pi$.

11.1.4. — Une représentation linéaire (V, ρ) d'un groupe G sur un corps k est dite *fidèle* si elle est fidèle en tant qu'opération de G sur l'ensemble V, c'est-à-dire si ρ est injective.

Notons qu'on pourrait être tenté de définir la transitivité de façon analogue, mais ce concept n'a guère d'intérêt ici : en effet $\{0\}$ est toujours une orbite de l'action de G sur V, si bien que cette action n'est transitive que si $V = \{0\}$.

Convention 11.1.5. — Soit G un groupe et soit k un corps. Le plus souvent, une représentation k-linéaire (V, ρ) de G sera simplement désignée par son espace vectoriel sous-jacent V, sans symbole spécifique pour le morphisme $G \to \mathrm{GL}(V)$ ni pour l'opération $G \times V \to V$, que l'on notera $(g, v) \mapsto gv$. La *dimension* d'une représentation sera simplement celle de son espace vectoriel sous-jacent (on l'appelle parfois aussi son *degré*).

***Exemple 11.1.6* (La représentation triviale).** — Soit V un k-espace vectoriel. L'opération triviale de G sur V, décrite par la formule $(g, x) \mapsto x$, définit une représentation de G ; le morphisme correspondant $G \to \mathrm{GL}(V)$ est le morphisme trivial. Lorsque V est l'espace nul, on obtient ainsi la *représentation nulle* de G. Lorsque $V = k$, on obtient la *représentation unité* de G (sur le corps k), qu'on notera $\mathbf{1}_{G,k}$.

***Exemple 11.1.7* (La représentation tautologique).** — Soit V un k-espace vectoriel. Le groupe $\mathrm{GL}(V)$ opère sur V de manière tautologique par automorphismes k-linéaires, *via* la formule $(g, v) \mapsto g(v)$. La représentation correspondante est appelée la *représentation tautologique* de $\mathrm{GL}(V)$; elle correspond à $\mathrm{Id}\colon \mathrm{GL}(V) \to \mathrm{GL}(V)$.

11.1.8. *Sous-représentations.* — Soit V une représentation de G et soit W un sous-espace vectoriel de V, stable sous l'action de G (c'est le cas dès que $gW \subset W$ pour tout $g \in G$, *cf.* 4.3.4.1). Le groupe G opère par restriction sur W, et cette opération se fait évidemment par automorphismes k-linéaires, ce qui permet de voir W comme une représentation de G.

Les représentations ainsi construites seront appelées les *sous-représentations* de V. Une somme de sous-représentations de V est une sous-représentations de V.

Un exemple important de sous-représentation de V est fourni par l'ensemble V^G des vecteurs fixes sous G ; il est en effet immédiat que c'est un sous-espace vectoriel de V, évidemment stable sous G. L'action de G sur V^G est par construction triviale.

Si G est fini alors pour tout vecteur $v \in V$ la somme $\sum_{g \in G} gv$ appartient à V^G ; en effet on a pour tout $h \in G$ l'égalité

$$h(\sum_g gv) = \sum_g (hg)v = \sum_g gv,$$

la dernière égalité provenant du fait que $g \mapsto hg$ est une bijection de G sur lui-même.

11.1.9. *Représentations quotient.* — Soit V une représentation de G et soit W une sous-représentation de V. Soit $v \in V$ et soit $g \in G$. L'application linéaire composée

$$\begin{tikzcd} V \arrow[r, "v \mapsto gv"] \arrow[rr, bend left] & V \arrow[r] & V/W \end{tikzcd}$$

s'annule sur W donc induit par passage au quotient un endomorphisme de V/W. On construit ainsi une application $G \times V/W \to V/W$. Il est immédiat qu'elle fait de V/W une représentation de G ; on a par construction $g\overline{v} = \overline{gv}$ pour tout $g \in G$ et tout $v \in V$. Nous dirons que V/W muni de cette opération de G est la *représentation quotient* de V par W.

11.1.10. *Représentations et morphismes de groupes.* — Soit (V, ρ) une représentation de G et soit $\varphi\colon H \to G$ un morphisme de groupes. La composée $\rho \circ \varphi$ est un morphisme de groupes de H dans $\mathrm{GL}(V)$, qui fait de V une représentation de H ; par construction, on a $hv = \varphi(h)v$ pour tout $h \in H$ (notez le cas important où H est un sous-groupe de G et où φ est l'inclusion : on a alors simplement restreint la représentation initiale à H).

11.1.11. *Sommes et produits de représentations.* — Soit $(V_i)_i$ une famille de représentations du groupe G. La formule $(g, (v_i)_i) \mapsto (gv_i)_i$ fait de $\prod_i V_i$ une représentation de G dont $\bigoplus_i V_i$ est une sous-représentation. Lorsque nous écrirons $\prod V_i$ ou $\bigoplus V_i$ sans autre précision, il sera toujours entendu que nous les considérons comme des représentations de G par les procédés que nous venons de décrire.

Si l'ensemble d'indices est fini, les représentations $\prod_i V_i$ et $\bigoplus_i V_i$ coïncident, et l'on préférera le plus souvent la notation $\bigoplus_i V_i$.

***Exemple 11.1.12* (La représentation signature).** — Soit k un corps et soit n un entier. Soit $\varepsilon_{n,k}$ le morphisme de $\mathfrak{S}_n$ dans $k^\times = \mathrm{GL}(k)$ qui envoie σ sur $\varepsilon(\sigma) = \pm 1$, vue comme élément de $k^\times$. Le morphisme $\varepsilon_{n,k}$ définit une représentation k-linéaire de $\mathfrak{S}_n$ dont l'espace vectoriel sous-jacent est la droite k ; elle vérifie par construction l'égalité $\sigma\lambda = \varepsilon(\sigma)\lambda$ pour tout $(\sigma, \lambda) \in \mathfrak{S}_n \times k$. Cette représentation sera appelée la *signature* sur le corps k) et nous la noterons $\Sigma_{n,k}$.

Si $n \leqslant 1$ ou si la caractéristique de k est égale à 2 le morphisme $\varepsilon_{n,k}$ est trivial, et $\Sigma_{n,k} = \mathbf{1}_{\mathfrak{S}_n,k}$. Dans le cas contraire, on peut choisir une transposition τ dans $\mathfrak{S}_n$, et l'on a alors $\tau\lambda = -\lambda \neq \lambda$ pour tout λ non nul dans k ; en conséquence, $\Sigma_{n,k} \neq \mathbf{1}_{\mathfrak{S}_n,k}$.

11.2. Morphismes de représentations. — Soit G un groupe et soient V et W deux représentations k-linéaires de G.

11.2.1. — Le k-espace vectoriel $\mathrm{Hom}_k(V, W)$ hérite d'une structure naturelle de représentation de G, par la formule

$$(g, \varphi) \mapsto [v \mapsto g(\varphi(g^{-1}v))].$$

11.2.2. — Le sous-espace vectoriel $\mathrm{Hom}_k(V, W)^G$ de $\mathrm{Hom}_k(V, W)$ est l'ensemble des applications k-linéaires φ de V vers W telles que $g\varphi(g^{-1}v) = \varphi(v)$ pour tout $g \in G$ et tout $v \in V$. Comme $g \mapsto g^{-1}$ est une bijection de G sur lui-même, cela revient à demander que $g^{-1}\varphi(gv) = \varphi(v)$ pour tout (g, v), soit encore que $\varphi(gv) = g\varphi(v)$ pour tout (g, v). Autrement dit, $\mathrm{Hom}_k(V, W)^G$ est l'ensemble des applications k-linéaires équivariantes (définition 4.1.12) de V vers W Ces applications seront appelées les *morphismes de représentations* de V vers W. On écrira souvent $\mathrm{Hom}_G(V, W)$ au lieu de $\mathrm{Hom}_k(V, W)^G$.

11.2.2.1. — L'identité de V est un morphisme de représentations. La composée de deux morphismes de représentations est un morphisme de représentations ; si un un morphisme de représentations est bijectif (on parle alors d'*isomorphisme de représentations*), sa réciproque est un morphisme de représentations. Le noyau d'un morphisme de représentations est une sous-représentation de la source ; son image est une sous-représentation du but.

11.2.2.2. — Si V' est une sous-représentation de V, l'inclusion $V' \hookrightarrow V$ et le morphisme quotient $\pi\colon V \to V/V'$ sont des morphismes de représentations. Si $V \to W$ est un morphisme de représentations s'annulant sur V', l'application k-linéaire induite $V/V' \to W$ est un morphisme de représentations ; ainsi, $f \mapsto f \circ \pi$ établit une bijection entre $\mathrm{Hom}_G(V/V', W)$ et l'ensemble des morphismes de représentations de V vers W s'annulant sur V'. Les formules $U \mapsto \pi(U)$ et $\Omega \mapsto \pi^{-1}(\Omega)$ mettent en bijection l'ensemble des sous-représentations de V contenant V' et celui des sous-représentations de V/V'.

Si $(V_i)_i$ est une famille de représentations de G, alors pour tout indice j les applications k-linéaires naturelles $V_j \to \bigoplus_i V_i$ et $\prod_i V_i \to V_j$ sont des morphismes de représentations.

11.2.2.3. — La plupart des propriétés des représentations que nous considérerons (par exemple : être fidèle, être triviale, avoir un sous-espace de vecteurs fixes de dimension donnée) sont trivialement invariantes par isomorphisme de représentations ; nous ferons un usage abondant, et le plus souvent implicite, de cette remarque.

11.2.3. *Représentation contragrédiente.* — Appliquons 11.2.2 lorsque W est la représentation unité de G, c'est-à-dire la droite vectorielle k avec action triviale de G. On en déduit une structure naturelle de représentation de G sur l'espace dual $V^* := \mathrm{Hom}_k(V, k)$. Par construction, on a pour toute forme linéaire φ appartenant à V^* et tout élément g de G l'égalité

$$(*) \quad g\varphi = v \mapsto (\varphi(g^{-1}v)).$$

On dit que V^* est la représentation *contragrédiente* de V. Soit $\rho\colon G \to \mathrm{GL}(V)$ le morphisme décrivant l'opération de G sur V, et soit $\rho^*\colon G \to \mathrm{GL}(V^*)$ le morphisme décrivant l'opération de G sur V^*. Soit $g \in G$. La formule $(*)$ ci-dessus signifie que $\rho^*(g) =^t \rho(g^{-1}) =^t \rho(g)^{-1}$, où ^{t} désigne la transposition *au sens de la dualité* (définition 10.6.14).

Supposons que V est de dimension finie, choisissons une base $(e_1, \ldots, e_n)$ de V, et notons $(e_1^*, \ldots, e_n^*)$ sa base duale. Soit π l'expression de ρ dans la base $(e_1, \ldots, e_n)$, et soit π^* l'expression de ρ^* dans la base $(e_1^*, \ldots, e_n^*)$. On a alors par ce qui précède $\pi^*(g) =^t \pi(g)^{-1}$ pour tout $g \in G$, où ^{t} désigne cette fois-ci la transposition *matricielle.*

11.2.4. *Représentations isomorphes.* — Soit G un groupe et soient (V_1, ρ_1) et (V_2, ρ_2) deux représentations de G.

11.2.4.1. — Les représentations (V_1, ρ_1) et (V_2, ρ_2) sont isomorphes en tant que représentations de G si et seulement si il existe un isomorphisme φ appartenant à $\mathrm{Hom}_k(V_1, V_2)$ qui est équivariant vis-à-vis de l'opération ρ_1 à la source, et de l'opération ρ_2 au but. En revenant aux définitions, on voit que cela signifie que $\rho_2(g) \circ \varphi = \varphi \circ \rho_1(g)$, pour tout $g \in G$, soit encore $\rho_2(g) = \varphi \circ \rho_1(g) \circ \varphi^{-1}$.

11.2.4.2. — Compte-tenu des formules de changement de base, il résulte de ce qui précède que (V_1, ρ_1) et (V_2, ρ_2) sont isomorphes si et seulement si il existe une base $\mathscr{B}_1$ de V_1 et une base $\mathscr{B}_2$ de V_2 telles que l'expression de ρ_1 dans la base $\mathscr{B}_1$ soit égale à celle de ρ_2 dans la base $\mathscr{B}_2$.

11.2.4.3. — Supposons que V_1 et V_2 soient égaux à un même k-espace vectoriel V. On peut reformuler 11.2.4.1 en disant que (V, ρ_1) et (V, ρ_2) sont isomorphes si et seulement si $\rho_1\colon G \to \mathrm{GL}(V)$ et $\rho_2\colon G \to \mathrm{GL}(V)$ sont conjuguées par un élément de $\mathrm{GL}(V)$.

11.2.5. *Classification des représentations de dimension 1.* — Soit G un groupe. Soit V un k-espace vectoriel de dimension 1. L'application $\lambda \mapsto \lambda_{\mathrm{Id}_V}$ induit un isomorphisme de k-algèbres entre k et $\mathrm{End}_k(V)$, et partant un isomorphisme de groupes $k^\times \simeq \mathrm{GL}(V)$. Se donner une représentation de G d'espace sous-jacent V revient donc à se donner un morphisme $\rho\colon G \to k^\times$.

11.2.5.1. — Soient ρ et ρ' deux morphismes de G dans $k^\times$, et soient V et V' deux k-espaces vectoriels dimension 1 ; on voit V (resp. V') comme une représentation de G *via* ρ (resp. ρ'). Soit f un isomorphisme k-linéaire entre V et V'. L'application f est équivariante si et seulement si $f \circ \rho(g)\mathrm{Id}_V \circ f^{-1} = \rho'(g)\mathrm{Id}_{V'}$ pour tout $g \in G$, soit encore si et seulement si $\rho(g)\mathrm{Id}_{V'} = \rho'(g)\mathrm{Id}_{V'}$ pour tout g, c'est-à-dire enfin si et seulement si $\rho = \rho'$. Notez que cette dernière condition ne fait plus intervenir f : si elle est satisfaite, *tout* isomorphisme k-linéaire entre V et V' est donc un isomorphisme de représentations.

11.2.5.2. — Récapitulons : l'ensemble des classes d'isomorphie de représentations k-linéaires de dimension 1 de G est en bijection naturelle avec l'ensemble des morphismes $\rho \colon G \to k^\times$. La classe associée à un tel morphisme est celle de la représentation $(D, g \mapsto \rho(g)\mathrm{Id}_k)$ pour n'importe quelle k-droite vectorielle D (on peut par exemple prendre $D = k$).

11.2.6. *Bidualité et représentations.* — Soit G un groupe et soit V une représentation k-linéaire de G. Soit $\theta_V : V \to V^{**}$ le morphisme canonique.

11.2.6.1. — On peut voir V^{**} de manière naturelle comme une représentation k-linéaire de G : c'est la contragrédiente de la contragrédiente de V (11.2.3).

11.2.6.2. — *Le morphisme $\theta_V \colon V \to V^{**}$ est équivariant, et est donc un isomorphisme de représentations lorsque V est de dimension finie.* En effet, soit $v \in V$ et soit $g \in G$. On a les égalités

$$\begin{aligned} \theta_V(gv) &= \varphi \mapsto \varphi(gv) \\ &= \varphi \mapsto (g^{-1}\varphi)(v) \\ &= \varphi \mapsto \theta_V(v)(g^{-1}\varphi) \\ &= g\theta_V(v), \end{aligned}$$

d'où notre assertion.

11.3. Représentations de permutation. — Nous allons maintenant expliquer comment construire une représentation k-linéaire d'un groupe d'un groupe G à partir d'une opération *ensembliste* de ce dernier ; cela repose sur le le procédé général de «k-linéarisation» décrit en 10.4.8.

11.3.1. — Soit G un groupe opérant sur un ensemble X et soit $\pi \colon G \to \mathfrak{S}_X$ le morphisme correspondant. La formule $\varphi \mapsto \varphi_{k\text{-lin}}$ définit un morphisme de groupes μ de $\mathfrak{S}_X$ vers $\mathrm{GL}(kX)$, qui est injectif (10.4.8.3). La composée $\mu \circ \pi$ est un morphisme de G vers $\mathrm{GL}(kX)$; elle définit donc une représentation k-linéaire de G, d'espace vectoriel sous-jacent kX, caractérisée par la formule $g[x] = [gx]$ pour tout $x \in X$; cette représentation est appelée la *k-linéarisée* de l'opération de G sur X, ou encore la *représentation de permutation* associée à l'opération de G sur X.

Si $g \in G$, la matrice de $g_{k\text{-lin}}$ dans la base $([x])$ est la *matrice de permutation* $(\delta_{x,gy})_{x,y}$.

Comme μ est injective, G opère fidèlement sur kX si et seulement si il opère fidèlement sur X.

Définition 11.3.2 **(La représentation régulière).** — Soit G un groupe. On appelle *représentation régulière de G* (sur le corps k) la k-linéarisation de l'opération de G sur lui-même par translations à gauche. Comme l'opération de G sur lui-même par translations à gauche est fidèle, la représentation régulière est fidèle.

11.3.3. — Soit G un groupe opérant sur un ensemble X. On fait l'hypothèse que X est de cardinal fini $n > 0$; on considère kX comme une représentation de G *via* la linéarisation de l'opération de G sur X.

11.3.3.1. — Soit v le vecteur $\sum_{x\in X}[x]$ de kX, non nul car $n > 0$. Comme $x \mapsto gx$ est pour tout $g \in G$ une bijection de X sur lui-même, on a $gv = v$ pour tout $g \in G$. La droite vectorielle kv de X est donc stable sous G, et l'action de G sur kv est triviale. Autrement dit, kv est une sous-représentation de kX isomorphe à la représentation unité $\mathbf{1}_{G,k}$.

11.3.3.2. — Soit φ la forme linéaire sur kX qui envoie un vecteur $\sum a_x[x]$ sur $\sum a_x$. Elle est non nulle car X est non vide et car $\varphi([x]) = 1$ pour tout $x \in X$. On a de plus $\varphi(gw) = \varphi(w)$ pour tout $g \in G$ et tout $w \in kX$ (il suffit par linéarité de le vérifier lorsque $w = [x]$ pour $x \in X$, auquel cas les deux membres de l'égalité sont égaux à 1). Son noyau H est donc un hyperplan de X stable sous G ; c'est en particulier une sous-représentation de kX.

On a $\varphi(v) = n$. *Supposons que n soit non nul dans k, c'est-à-dire que k soit de caractéristique nulle ou première à n.* On a alors $H \cap kv = \{0\}$; l'hyperplan H et la droite kv sont ainsi en somme directe, ce qui entraîne que $kX = kv \oplus H$ pour des raisons de dimension. La représentation kX de G est en conséquence isomorphe à $\mathbf{1}_{G,k} \oplus H$, et les représentations quotient kX/kv et kX/H sont respectivement isomorphes à H et à $\mathbf{1}_{G,k}$.

11.3.3.3. — Nous utiliserons à plusieurs reprises la construction du 11.3.3.2 ci-dessus lorsque $X = \{1, \ldots, n\}$ et lorsque G est le groupe $\mathfrak{S}_n$ opérant tautologiquement sur X ; la représentation H obtenue dans ce cas particulier sera notée $V_{n,k}$.

11.3.4. *Une vision «affine» de $V_{n,k}$.* — Cette discussion ne sera pas utilisée par la suite. Nous l'incluons simplement pour la curiosité du lecteur ; elle suppose connue les bases de la géométrie affine.

11.3.4.1. — Soit E un k-espace affine et soit R un *repère affine* de E, c'est-à-dire un sous-ensemble de points affinement indépendants qui ne sont pas tous contenus dans un même sous-espace affine *strict* de E (si E est de dimension finie, un repère affine est une famille de points affinement indépendants de cardinal $\dim E + 1$: pensez à deux points distincts d'une droite, ou à trois points non alignés d'un plan, ou à 4 points non coplanaires dans un espace de dimension 3). Soit G le sous-groupe du groupe $\mathrm{GA}(E)$ des bijections affines de E dans lui-même constitué des bijections φ telles que $\varphi(R) = R$. La théorie générale des espaces affines assure que pour toute bijection σ de R dans lui-même, il existe une unique bijection affine $\varphi \in G$ dont la restriction à R est égale à σ ; en d'autres termes, la restriction à R induit un isomorphisme $G \simeq \mathfrak{S}_R$.

11.3.4.2. — Choisissons maintenant un entier n non nul dans k. Pour tout i, posons

$$u_i = \frac{n-1}{n}[i] - \frac{1}{n}\sum_{j\neq i}[j] \in k\{1, \ldots, n\}$$

Il est immédiat que u_i appartient à la sous-représentation $V_{n,k}$ de $k\{1, \ldots, n\}$. Soit R l'ensemble $\{u_1, \ldots, u_n\}$. C'est un repère affine de $V_{n,k}$: pour le voir en minimisant les calculs à effectuer, on peut remarquer que la translation de vecteur $(1/n)([1]+\ldots+[n])$ (qui est une bijection affine) envoie R sur $\{[1], \ldots, [n]\}$, et il est immédiat que les éléments $[i]$ de $V_{n,k}$ sont affinement indépendants.

Soit G le sous-groupe de $\mathrm{GA}(V_{n,k})$ constitué des bijections φ telles que $\varphi(R) = R$. Un calcul élémentaire montre que l'isobarycentre des éléments de R (qui a un sens car $|R| = n$ est non nul dans k) est égal à l'origine. Comme il est évidemment invariant sous G, le groupe G est en réalité contenu dans $G \subset \mathrm{GL}(V_{n,k})$.

On vérifie sans peine que $\sigma u_i = u_{\sigma(i)}$ pour tout i. En vertu de 11.3.4.1 il en résulte que le morphisme $\mathfrak{S}_n \to \mathrm{GL}(V_{n,k})$ n'est autre que le morphisme composé

$$\mathfrak{S}_n \simeq \mathfrak{S}_R \simeq G \subset \mathrm{GL}(V_{n,k})$$

où le premier isomorphisme est déduit de la numérotation $u_1, \ldots, u_n$ des éléments de R et où le second est celui décrit au 11.3.4.1.

11.3.4.3. — Supposons que $k = \mathbf{R}$, et munissons $\mathbf{R}\{1, \ldots, n\}$ du produit scalaire pour lequel la base $([1], \ldots, [n])$ est orthonormée. Comme $\mathfrak{S}_n$ permute les vecteurs de cette base, il opère par isométries sur $\mathbf{R}\{1, \ldots, n\}$. Le morphisme naturel de $\mathfrak{S}_n$ dans $\mathrm{GL}(V_{n,\mathbf{R}})$ induit ainsi un isomorphisme entre $\mathfrak{S}_n$ et le groupe des *isométries* affines de $V_{n,\mathbf{R}}$ préservant R.

Les u_i sont par ailleurs équidistants les uns les autres (pour le voir, le plus simple est d'utiliser la translation par $(1/n)([1]+\ldots+[n])$ comme au 11.3.4.2 ; c'est une isométrie affine qui envoie u_i sur $[i]$ pour tout i, et l'on remarque alors que $||\ [i] - [j]\ || = \sqrt{2}$ pour tout (i, j) avec $i \neq j$). Ainsi lorsque $n = 2$ (resp. 3) on peut décrire R comme un triangle équilatéral (resp. un tétraèdre régulier) centré à l'origine, et vous avez sans doute déjà rencontré dans ce cas particulier l'isomorphisme que nous avons exhibé entre $\mathfrak{S}_3$ (resp. $\mathfrak{S}_4$) et le groupe des isométries préservant R ; en général (*i.e* pour n quelconque) R est ce qu'on appelle un *simplexe régulier*.

11.4. Produit tensoriel de représentations. — Soit G un groupe. Nous allons expliquer ci-dessous comment munir le produit tensoriel de deux représentations k-linéaires de G d'une opération k-linéaire de G.

11.4.1. — Soient (V, ρ) et (W, ϖ) deux représentations k-linéaires de G. Pour tout élément g de G, on pose $(\rho \otimes \varpi)(g) = \rho(g) \otimes \varpi(g)$. L'application $\rho \otimes \varpi$ est un morphisme de groupes de G dans $\mathrm{GL}(V \otimes_k W)$; la représentation $(V \otimes_k W, \rho \otimes \varpi)$ est appelée le *produit tensoriel* des représentations (V, ρ) et (W, ϖ) et est noté $(V, \rho) \otimes_k (W, \varpi)$.

Si l'on omet de mentionner ρ et ϖ et qu'on se contente d'évoquer les représentations V et W de G, le produit tensoriel $V \otimes_k W$ sera toujours considéré, sauf mention expresse du contraire, comme une représentation de G *via* le procédé décrit ci-dessus.

11.4.2. — Soient U, V et W trois représentations k-linéaires de G et soit (V_i) une famille de représentations k-linéaires de G. On dispose (10.7.5, 10.7.6, proposition 10.7.9 et remarque 10.7.11) d'isomorphismes k-linéaires naturels

$$V \otimes_k W \simeq W \otimes_k V, \ \ U \otimes_k (V \otimes_k W) \simeq (U \otimes_k V) \otimes_k W$$

$$\text{et}\ \left(\bigoplus_i V_i\right)\otimes_k W \simeq \bigoplus_i V_i \otimes_k W.$$

On vérifie immédiatement que ce sont des morphismes de représentations (l'équivariance se teste sur les tenseurs purs). Le produit tensoriel de représentations est donc commutatif, associatif, et commute aux sommes directes de représentations.

11.4.3. — Soient V et W deux représentations k-linéaires de G de dimension finie. On a construit au 10.7.15.3 un isomorphisme $u\colon V^*\otimes_k W \simeq \mathrm{Hom}_k(V, W)$. Nous allons vérifier que c'est un isomorphisme de représentations. Il suffit de tester l'équivariance sur les tenseurs purs. Soient donc $\varphi \in V^*$, $w \in W$ et $g \in G$. On a les égalités

$$\begin{aligned} u(g(\varphi\otimes w)) &= u\left([v\mapsto \varphi(g^{-1}v)]\otimes gw\right) \\ &= v\mapsto \varphi(g^{-1}v)gw \\ &= v\mapsto g(\varphi(g^{-1}v)w) \\ &= v\mapsto g(u(\varphi\otimes w)(g^{-1}v)) \\ &= gu(\varphi\otimes w), \end{aligned}$$

d'où l'assertion.

11.4.4. *Produit tensoriel par une représentation de dimension 1.* — Rappelons tout d'abord quelques faits établis en 11.2.5 *et sq.* Modulo l'identification canonique entre $k^\times$ et $\mathrm{GL}(k)$, tout morphisme $\varpi\colon G \to k^\times$ induit une représentation k-linéaire (k, ϖ) de G ; et toute représentation k-linéaire de G de dimension 1 est isomorphe à une telle (k, ϖ) pour un unique ϖ.

Soit (V, ρ) une représentation k-linéaire de G et soit ϖ un morphisme de G vers $k^\times$. Le produit tensoriel $k\otimes_k V$ est canoniquement isomorphe à V comme k-espace vectoriel d'après 10.7.4.1. Le produit tensoriel des représentations (k, ϖ) et (V, ρ) est donc isomorphe à une représentation d'espace sous-jacent V ; le but de ce qui suit est de la décrire.

Soit $g \in G$ et soit $v \in V$. L'élément de $k\otimes_k V$ qui correspond à v est le tenseur pur $1\otimes v$. Par définition du produit tensoriel de deux représentations, on a

$$g\cdot(1\otimes v) = g(1)\otimes g(v) = \varpi(g)\otimes\rho(g)(v)\ ;$$

l'élément de V qui correspond au tenseur pur $\varpi(g)\otimes\rho(g)v$ de $k\otimes_k V$ est $\varpi(g)\rho(g)(v)$.

Il en résulte que $(k, \varpi)\otimes_k(V, \rho)$ est isomorphe à $(V, g\mapsto \varpi(g)\rho(g))$. Informellement, tensoriser par (k, ϖ) revient donc à «tordre» la représentation ρ par le morphisme ϖ.

Soit W un sous-espace vectoriel de V. C'est une sous-représentation de (V, ρ) si et seulement si $\rho(g)(w) \in W$ pour tout $(g, w) \in G\times W$. Comme ϖ est à valeurs dans $k^\times$, cela revient à demander que $\varpi(g)\rho(g)(w) \in W$ pour tout $(g, w) \in G\times W$, c'est-à-dire encore que W soit une sous-représentation de $(V, g\mapsto \varpi(g)\rho(g))$. Ainsi, (V, ρ) et $(V, g\mapsto \varpi(g)\rho(g)) \simeq (k, \varpi)\otimes_k (V, \rho)$ ont les mêmes sous-représentations.

Exemple 11.4.5. — Soit V une représentation k-linéaire de G. On déduit de 11.4.4 que $\mathbf{1}_{G,k}\otimes_k V$ est isomorphe à V (comme représentation de G) ; cela justifie l'expression «représentation unité» : $\mathbf{1}_{G,k}\otimes_k V$ est en quelque sorte le neutre relativement au produit tensoriel des représentations k-linéaires de G.

Remarque 11.4.6. — Comme $k^\times$ est abélien, le groupe $\mathrm{Hom}(G, k^\times)$ a une structure naturelle de groupe abélien ; elle fournit par transport de structure une loi de groupe abélien sur l'ensemble des classes d'isomorphie de représentations de dimension 1 de G. Il résulte facilement de 11.4.4 que celle loi est induite par le *produit tensoriel* de représentations.

Exemple 11.4.7. — Soit n un entier et soit (V, ρ) une représentation k-linéaire de $\mathfrak{S}_n$. En vertu de 11.4.4, on dispose d'un isomorphisme naturel de représentations

$$\Sigma_{n,k} \otimes_k (V, \rho) \simeq (V, \sigma \mapsto \varepsilon(\sigma)\rho(\sigma)).$$

Remarque 11.4.8. — On a signalé au 11.3.4.3 qu'on peut identifier $V_{4,\mathbf{R}}$ à la représentation de $\mathfrak{S}_4$ consistant à voir celui-ci comme le groupe des isométries d'un tétraèdre régulier (centré en l'origine). Nous vous laissons utiliser cette description et l'exemple 11.4.7 ci-dessus pour montrer que $\Sigma_{4,\mathbf{R}} \otimes_{\mathbf{R}} V_4$ s'identifie à la représentation de $\mathfrak{S}_4$ consistant à voir celui-ci comme le groupe des isométries *directes* d'un cube centré en l'origine.

11.5. Caractère d'une représentation. — Soit G un groupe. Nous allons maintenant introduire une fonction attachée à une représentation de dimension finie de G, qu'on appelle son *caractère*. Son intérêt ne vous sautera sans doute pas aux yeux d'emblée, mais c'est un outil essentiel dans l'étude des représentations ; nous en verrons une illustration à la section suivante consacrée aux représentations complexes de dimension finie des groupes finis.

Définition 11.5.1. — Soit (V, ρ) une représentation k-linéaire de G *de dimension finie*. On appelle *caractère* de la représentation (V, ρ) l'application

$$G \to k, g \mapsto \mathrm{Tr}(\rho(g)).$$

On le note souvent $\chi_{(V,\rho)}$, ou parfois simplement χ_V ou χ_ρ si le contexte le permet.

***Exemple 11.5.2* (Caractère de la représentation unité)**

Il résulte immédiatement de la définition que $\chi_{\mathbf{1}_G}$ est la fonction constante égale à 1.

***Exemple 11.5.3* (Caractère d'une représentation de dimension 1)**

Se donner une classe d'isomorphie de représentation k-linéaire de dimension 1 de G revient à se donner un morphisme de G vers $k^\times$ (11.2.5 *et sq.*). Si ρ est un tel morphisme, la classe d'isomorphie correspondante est celle de $(k, g \mapsto \rho(g)\mathrm{Id}_k)$; le caractère associé est $g \mapsto \mathrm{Tr}(\rho(g)\mathrm{Id}_k) = \rho(g)$: *en dimension 1, le caractère d'une représentation coïncide avec le morphisme $G \to k^\times$ qui la définit à isomorphisme près.* Nous verrons à la section suivante que lorsque $k = \mathbf{C}$ et lorsque G est fini, une représentation de dimension finie est encore entièrement déterminée à isomorphisme près par son caractère, mais c'est un phénomène évidemment plus subtil que celui que nous venons de mettre en évidence.

11.5.4. — Soit (V, ρ) une représentation k-linéaire de G de dimension finie. Pour tout couple (h, g) d'éléments de G on a

$$\chi_{(V,\rho)}(hgh^{-1}) = \mathrm{Tr}(\rho(hgh^{-1})) = \mathrm{Tr}(\rho(h)\rho(g)\rho(h)^{-1}) = \mathrm{Tr}(\rho(g)) = \chi_{(V,\rho)}(g).$$

Ainsi, le caractère $\chi_{(V,\rho)}$ est une *fonction centrale*, c'est-à-dire que sa valeur sur un élément g de G ne dépend que de la classe de conjugaison de g. On pourra pour cette raison considérer à l'occasion $\chi_{(V,\rho)}$ comme une fonction de l'ensemble des classes de conjugaison de G vers k ; cela signifie plus précisément que si c est une telle classe, on se permettra de noter $\chi_{(V,\rho)}(c)$ la valeur constante que prend $\chi_{(V,\rho)}$ sur c.

11.5.5. *Caractère et dimension.* — Soit V une représentation k-linéaire de G de dimension finie. Le scalaire $\chi_V(e)$ est égal à $\mathrm{Tr}(\mathrm{Id}_V)$, c'est-à-dire à la dimension de V *vue comme élément de* k.

Si k est de caractéristique nulle l'appplication canonique $\mathbf{Z} \to k$ est injective, et $\chi_V(e)$ peut donc être considéré comme un élément de $\mathbf{Z}$, égal à $\dim V$: on retrouve $\dim V$ à partir de χ_V.

Si k est de caractéristique $p > 0$ l'application canonique $\mathbf{Z} \to k$ a pour noyau $p\mathbf{Z}$, et son image s'identifie à $\mathbf{Z}/p\mathbf{Z}$. On peut donc voir $\chi_V(e)$ comme un élément de $\mathbf{Z}/p\mathbf{Z}$, égal à la classe de $\dim V$ modulo p. On ne retrouve plus *a priori* $\dim V$ à partir de χ_V, mais seulement sa classe modulo p. Il est en fait vain d'espérer davantage. En effet, supposons que la dimension de V est multiple de p, et que G agit trivialement sur V. Dans ce cas, $\chi_V(g) = \mathrm{Tr}(\mathrm{Id}_V) = 0$ pour tout $g \in G$, et le caractère χ_V est donc la fonction nulle, indépendamment de la dimension de V.

11.5.6. *Caractères et isomorphismes.* — Soient (V_1, ρ_1) et (V_2, ρ_2) deux représentations k-linéaires de dimension finie de G. Si elles sont isomorphes, elles ont même caractère : c'est une conséquence directe de 11.2.4.2.

11.5.7. *Caractères et sommes directes.* — Soit $(V_i)_{i\in I}$ une famille finie de représentations k-linéaires de G de dimension finie. On choisit pour tout i une base $(e_{ij})_j$ de V_i, et on note $\mathscr{B}$ la concaténation des (e_{ij}) ; c'est une base de $\bigoplus_i V_i$. Soit $g \in G$; pour tout i, désignons par M_i la matrice dans la base (e_{ij}) de l'opération de g sur V_i ; désignons par M la matrice dans la base $\mathscr{B}$ de l'opération de g sur $\bigoplus V_i$. On a alors $M = \mathrm{Diag}(M_i)_i$, et partant $\mathrm{Tr}(M) = \sum_i \mathrm{Tr}(M_i)$; il vient

$$\chi_{\bigoplus_i V_i} = \sum_i \chi_{V_i}.$$

11.5.8. *Caractères et produit tensoriel.* — Soit W une représentation k-linéaires de G de dimension finie. Fixons une base (e_i) de V et une base (f_r) de W. Soit $g \in G$, soit $M = (a_{ij})$ la matrice de l'opération de g sur V dans la base (e_i), et soit $N = (b_{rs})$ la matrice de l'opération de g sur W dans la base (f_r). Soit $L = (c_{irjs})$ la matrice de l'opération de g dans la base $(e_i \otimes f_r)_{i,r}$ de $V \otimes_k W$. Il résulte de 10.7.14 que l'on a

$c_{irjs} = a_{ij}b_{rs}$ pour tout (i,r,j,s). On a alors

$$\begin{aligned}\mathrm{Tr}(L) &= \sum_{i,r} c_{irir}\\ &= \sum_{i,r} a_{ii}b_{rr}\\ &= (\sum_i a_{ii})(\sum_r b_{rr})\\ &= \mathrm{Tr}(M)\mathrm{Tr}(N).\end{aligned}$$

Il vient

$$\chi_{V\otimes_k W} = \chi_V \cdot \chi_W.$$

11.5.9. *Caractères et contragrédientes.* — Soit V une k-représentation linéaire de dimension finie de G. Soit (e_i) une base de V, et soit (e_i^*) la base duale de (e_i). Soit $g \in G$ et soit M la matrice dans la base (e_i) de l'opération de g sur V. La matrice dans la base (e_i^*) de l'opération de g sur V^* est alors ${}^tM^{-1}$ (11.2.3). On a donc

$$\chi_{V^*}(g) = \mathrm{Tr}({}^tM^{-1}) = \mathrm{Tr}(M^{-1}) = \chi_V(g^{-1}).$$

En général, le scalaire $\chi_{V^*}(g)$ n'a aucune raison de ne dépendre que de $\chi_V(g)$. Mais c'est le cas lorsque $k = \mathbf{C}$ et lorsque G est fini. Pour le voir, plaçons-nous sous cette hypothèse et notons n le cardinal de G. On a alors $g^n = e$, et partant $M^n = \mathrm{Id}$. Comme M annule le polynôme $X^n - 1$, ses valeurs propres sont toutes des racines de l'unité, et en particulier sont toutes de module 1 ; pour une telle valeur propre λ, on a donc $\lambda^{-1} = \overline{\lambda}$. En se plaçant dans une base de trigonalisation de M on en déduit que $\mathrm{Tr}(M^{-1}) = \overline{\mathrm{Tr}(M)}$.

Si G est fini et $k = \mathbf{C}$, on a donc $\chi_{V^*} = \overline{\chi_V}$.

11.5.10. *Caractères et morphismes de groupes.* — Soit (V,ρ) une représentation k-linéaire de G de dimension finie. Soit H un groupe et soit $\varphi\colon H \to G$ un morphisme. Il résulte immédiatement que le caractère de la représentation $(V, \varphi\circ\rho)$ de H est égal à $\chi_{(V,\rho)} \circ \varphi$.

11.5.11. *Caractère d'une représentation de permutation.* — Soit X un ensemble fini sur lequel G opère. On considère l'espace vectoriel kX comme une représentation de G *via* la linéarisation de son opération sur X.

11.5.11.1. — Soit $g \in G$. La matrice M de $g_{k\text{-lin}}$ dans la base (e_x) est la matrice de permutation $(\delta_{x,gy})_{x,y}$. Pour tout $x \in X$, le terme diagonal $\delta_{x,gx}$ de M est égal à 1 si $x \in \mathrm{Fix}(g)$ et à 0 sinon. Il vient

$$\chi(g) = \mathrm{Tr}(M) = |\mathrm{Fix}(g)|,$$

où l'entier $|\mathrm{Fix}(g)|$ est vu comme appartenant à k.

11.5.11.2. — Supposons que $|X|$ est non nul dans k. Posons $v = \sum_x e_x$, et notons H l'ensemble des combinaisons linéaires de la forme $\sum_x a_x e_x$ où $\sum a_x = 0$. Les sous-espaces vectoriels kv et H de kX en sont deux sous-représentations de dimension respective 1 et $n-1$, la représentation kv est isomorphe à $\mathbf{1}_{G,k}$, et kX est égal à $kv \oplus H$ (11.3.3.2).

Soit $g \in G$. Il résulte de ce qui précède que

$$\chi_H(g) = \chi_{kX}(g) - \chi_{\mathbf{1}_{G,k}}(g) = |\mathrm{Fix}(g)| - 1.$$

La dernière égalité résulte de 11.5.11.1, et l'entier $|\mathrm{Fix}(g)| - 1$ est là encore vu comme appartenant à k.

11.5.12. — Déclinons 11.5.11 lorsque $G = \mathfrak{S}_4$, et lorsque X est l'ensemble $\{1, 2, 3, 4\}$ sur lequel $\mathfrak{S}_4$ opère tautologiquement ; le corps k est alors supposé de caractéristique différente de 2, et H est la représentation $V_{4,k}$.

Considérons $\chi_{V_{4,k}}$ comme une fonction sur l'ensemble des classes de conjugaison de $\mathfrak{S}_4$. Nous allons donner ci-dessous la liste de ces classes, et la valeur de $\chi_{V_{4,k}}$ sur chacune d'elle (qui sera un entier, à interpréter comme un élément de k). Nous avons choisi de désigner une classe par le symbole c auquel on adjoint en indice la description du type de décomposition en cycles qui définit la classe en question (prop. 5.3.8).

Liste des classes de conjugaison de $\mathfrak{S}_4$ et des valeurs de $\chi_{V_{4,k}}$.

- ◇ La classe $c_\varnothing$ dont le seul élément est l'identité. Comme $\mathrm{Id}_{\{1,2,3,4\}}$ a quatre points fixes, on a $\chi_{V_{4,k}}(c_\varnothing) = 4 - 1 = 3$.
- ◇ La classe $c_{(**)}$ des transpositions. Une transposition de $\{1, 2, 3, 4\}$ a deux points fixes. Il vient $\chi_{V_{4,k}}(c_{(**)}) = 2 - 1 = 1$.
- ◇ La classe $c_{(***)}$ des 3-cycles. Un 3-cycle de $\{1, 2, 3, 4\}$ a un point fixe. Il vient $\chi_{V_{4,k}}(c_{(***)}) = 1 - 1 = 0$.
- ◇ La classe $c_{(****)}$ des 4-cycles. Un 4-cycle de $\{1, 2, 3, 4\}$ est sans point fixe. Il vient $\chi_{V_{4,k}}(c_{(****)}) = 0 - 1 = -1$.
- ◇ La classe $c_{(**)(**)}$ des produits de 2 transpositions à supports disjoints. Un tel produit est sans point fixe. Il vient $\chi_{V_{4,k}}(c_{(**)(**)}) = 0 - 1 = -1$.

Lemme 11.5.13. — *Supposons que G est fini, et que son cardinal n est non nul dans k. Soit V une représentation k-linéaire de G.*

(1) *L'endomorphisme*

$$q \colon v \mapsto \frac{1}{n} \sum_g gv$$

de V est un projecteur d'image V^G.

(2) *Si V est de dimension finie la somme $\frac{1}{n} \sum_{g \in G} \chi_V(g)$ est égale à l'entier* $\dim V^G$ vu comme élément de k. *Cette somme est donc «vraiment» égale à* $\dim V^G$ *si la caractéristique de k est nulle, et à* $\dim V^G$ *modulo p si la caractéristique de k est un nombre premier p,* cf. *la discussion du 11.5.5.*

Démonstration. — Commençons par établir (1). En vertu du lemme 10.5.2, il suffit de s'assurer que $q(v) \in V^G$ pour tout $v \in V$, et que $q(v) = v$ pour tout $v \in V^G$.

Il résulte de 11.1.8 que $q(v) \in V^G$ pour tout $v \in V$. Soit $v \in V^G$. On a alors $q(v) = \frac{1}{n} \sum_g gv = \frac{1}{n} \sum_{g \in G} v = \frac{1}{n} nv = v$.

Supposons que V est de dimension finie. Par définition de χ_V, l'élément $\frac{1}{n} \sum_{g \in G} \chi_V(g)$ de k est alors égal à la trace du projecteur q, et donc à la dimension de son image V^G (vue comme élément de k). □

11.6. Représentations irréductibles et indécomposables. — Nous allons maintenant aborder, à propos des représentations linéaires, un type de question que nous avons déjà considéré en théorie des groupes (lors de l'étude des suites de Jordan-Hölder) : comment «dévisser» une représentation donnée en représentations aussi simples que possible ? Bien entendu, il faut savoir pour commencer ce que l'on entend rigoureusement par «aussi simple que possible» ; la notion précise répondant à ce cahier des charges sera celle de représentation *irréductible*, que nous allons maintenant définir et étudier ; nous aurons également besoin d'une autre notion, celle de représentation *indécomposable*.

Définitions 11.6.1. — Soit G un groupe et soit V une représentation k-linéaire de G. On dit que V est *irréductible* si $V \neq \{0\}$ et si les seules sous-représentations de V sont $\{0\}$ et V. On dit que V est *indécomposable* si $V \neq \{0\}$ et si V *ne peut pas* s'écrire comme une somme directe de deux sous-représentations strictes. On dit que V est *décomposable* si elle n'est pas indécomposable (notez que la représentation nulle est décomposable).

Le lecteur aimant les facéties ensemblistes pourra remarquer qu'une représentation est décomposable si et seulement si elles peut s'écrire comme une somme directe finie de sous-représentations strictes, la représentation nulle étant la *somme directe vide* de telles sous-représentations.

Remarque 11.6.2. — Soit G un groupe. Si V est une représentation k-linéaire irréductible de V, elle est indécomposable : c'est une conséquence immédiate des définitions.

Remarque 11.6.3. — Soit G un groupe et soit $f\colon V \to W$ un morphisme surjectif de représentations k-linéaires de G. Si V est irréductible et si f est surjective alors ou bien $W = \{0\}$ ou bien f est un isomorphisme ; en effet, $\mathrm{Ker} f$ est une sous-représentation de V, donc est ou bien nulle, ou bien égale à V tout entier, d'où notre assertion.

Notez l'analogie entre cette remarque et la remarque 8.1.7 sur les groupes simples.

Exemple 11.6.4. — Soit G un groupe. Toute représentation k-linéaire de dimension 1 de G est irréductible, et *a fortiori* indécomposable : c'est évident. Ainsi, la représentation unité $\mathbf{1}_{G,k}$ est irréductible ; et pour tout entier n, la représentation signature $\Sigma_{n,k}$ de S_n (exemple 11.1.12) est irréductible.

11.6.5. — Soit G un groupe, soit (V, ρ) une représentation k-linéaire de G et soit D une représentation k-linéaire de G de dimension 1. Il résulte de 11.4.4 que $D \otimes_k (V, \rho)$ est isomorphe à une représentation k-linéaire de la forme (V, ρ') et que (V, ρ) et (V, ρ') ont les mêmes sous-représentations. Par conséquent $D \otimes_k (V, \rho)$ est irréductible (resp. indécomposable) si et seulement si (V, ρ) est irréductible (resp. indécomposable).

11.6.6. — Soit G un groupe et soit V une représentation k-linéaire de dimension finie de G. La représentation V est irréductible (resp. indécomposable) si et seulement si sa contragrédiente V^* est irréductible (resp. indécomposable). Pour le voir, il suffit en vertu de 11.2.6.2 de montrer que si V^* est irréductible (resp. indécomposable) alors V est irréductible (resp. indécomposable).

Supposons donc V^* irréductible. Elle est alors non nulle, et V est donc non nulle. Soit W une sous-représentation de V et soit $W^\perp \subset V^*$ son orthogonal au sens de la dualité. Le sous-espace vectoriel $W^\perp$ de V est stable sous l'action de G : si $\varphi \in W^\perp$ et si $g \in G$ on a pour tout w appartenant à W l'égalité $(g\varphi)(w) = \varphi(g^{-1}w) = 0$ car $g^{-1}w \in W$ et $\varphi \in W^\perp$; ainsi, $g\varphi \in W^\perp$. Par conséquent, $W^\perp$ est une sous-représentation de V^*. Par irréductibilité de V^* on a $W^\perp = \{0\}$ ou $W^\perp = V^*$; il vient $W = V$ ou $W = \{0\}$ et V est irréductible.

Supposons maintenant V^* indécomposable. Elle est alors non nulle, et V est donc non nulle. Supposons que V soit la somme directe $V_1 \bigoplus V_2$ de deux de ses sous-représentations. Comme une forme linéaire s'annulant sur V_1 et sur V_2 est clairement nulle, on a $V_1^\perp \cap V_2^\perp = \{0\}$, et un calcul de dimension montre alors que $V^* = V_1^\perp \oplus V_2^\perp$. Comme $V_1^\perp$ et $V_2^\perp$ sont deux sous-représentations de V^* (*cf. supra*), l'indécomposabilité de cette dernière implique que $V_1^\perp = \{0\}$ ou $V_2^\perp = \{0\}$, c'est-à-dire que $V_1 = V$ ou $V_2 = V$. Par conséquent, V est indécomposable.

11.6.7. — Soit $f\colon G \to H$ un morphisme de groupes, et soit V une représentation k-linéaire de H ; on peut également voir V comme une représentation de G *via* le morphisme f (11.1.10). Soit W un sous-espace vectoriel de V. Si W est stable sous H on a *a fortiori* $f(g)w \in W$ pour tout $g \in G$ et tout $w \in W$, et W est stable sous G. Réciproquement, supposons que W est stable sous G *et que f est surjective.* Soit $h \in H$ et soit $w \in W$. Par hypothèse, il existe $g \in G$ tel que $h = f(g)$; on a alors $hw = f(g)w$, et $f(g)w \in W$ puisque W est stable sous G. Ainsi, $hw \in W$ et W est stable sous H.

Il en résulte que *si f est surjective* les sous-représentations de V vue comme représentation de G sont exactement les sous-représentations de V vu comme représentation de H ; par conséquent, V est irréductible (resp. indécomposable) comme représentation de G si et seulement si V est irréductible (resp. indécomposable) comme représentation de H.

11.6.8. — Soit G un groupe opérant sur un ensemble X de cardinal fini $n \geqslant 2$; la k-linéarisation de cette opération fait de kX une représentation k-linéaire de G. Celle-ci n'est pas irréductible : on a en effet exhibé au 11.3.3.1 un vecteur v non nul de kX tel que la droite vectorielle kv soit une sous-représentation de kX (sur laquelle l'action de G est par ailleurs triviale).

11.6.8.1. *Le cas où n est non nul dans k.* — La représentation kX est alors décomposable : on a en effet exhibé sous cette hypothèse au 11.3.3.2 une sous-représentation H de kX de dimension $n-1$ telle que $kX = kv \oplus H$.

11.6.8.2. — Si on ne fait plus d'hypothèse sur n, il peut arriver que kX soit indécomposable, comme l'illustre l'exemple qui suit.

Supposons que k est de caractéristique 2, et faisons opérer $\mathfrak{S}_2$ sur $\{1,2\}$ tautologiquement. Nous allons vérifier que la représentation $k\{1,2\}$ obtenue en linéarisant cette action est indécomposable.

Soit τ l'automorphisme k-linéaire de $k\{1,2\} = k[1] \oplus k[2]$ induit par la permutation (12) ; par définition, on a $\tau([1]) = [2]$ et $\tau([2]) = [1]$. Une sous-représentation stricte et non nulle de $k\{1,2\}$ est simplement une droite vectorielle stable sous τ, c'est-à-dire une droite de la forme ku où u est un vecteur propre de τ. Pour montrer que $k\{1,2\}$ est indécomposable, il suffit donc de s'assurer que τ n'est pas diagonalisable. Or comme $\tau^2 = \mathrm{Id}$, toute valeur propre λ de τ vérifie l'équation $\lambda^2 = 1$. Puisque k est de caractéristique 2, cela équivaut à $(\lambda - 1)^2 = 0$ et donc à $\lambda = 1$. Si τ était diagonalisable, elle serait en conséquence égale à l'identité, ce qui est absurde puisque τ échange [1] et [2].

11.6.9. *Les représentations standard réelle et complexe du groupe diédral.* — Soit n un entier au moins égal à 3. On a construit aux 6.2.9 *et sq.* un isomorphisme entre le groupe diédral D_n (6.2.7.3) et le groupe des isométries $\mathbf{R}$-linéaires de $\mathbf{R}^2 \simeq \mathbf{C}$ préservant le groupe Γ des racines n-ièmes de l'unité (qui est un polygone régulier à n côtés). On obtient ainsi une représentation $\mathbf{R}$-linéaire $(\mathbf{R}^2, \rho)$ de D_n et, par restriction, une représentation $\mathbf{R}$-linéaire $(\mathbf{R}^2, \rho|_{\mathbf{Z}/n\mathbf{Z}})$ du sous-groupe distingué $\mathbf{Z}/n\mathbf{Z}$ de D_n.

Le choix de la base canonique de $\mathbf{R}^2$ permet d'identifier $\mathrm{GL}(\mathbf{R}^2)$ à $\mathrm{GL}_2(\mathbf{R})$, et de voir ρ comme un morphisme de D_n vers $\mathrm{GL}_2(\mathbf{R})$ (en d'autres termes, on identifie ρ à son expression matricielle dans la base canonique de $\mathbf{R}^2$). Comme $\mathrm{GL}_2(\mathbf{R}) \subset \mathrm{GL}_2(\mathbf{C}) \simeq \mathrm{GL}(\mathbf{C}^2)$ (où $\mathbf{C}^2$ est lui aussi muni de sa base canonique), on peut voir le couple $(\mathbf{C}^2, \rho)$ comme une représentation $\mathbf{C}$-linéaire de D_n, et le couple $(\mathbf{C}^2, \rho|_{Z/n\mathbf{Z}})$ comme une représentation $\mathbf{C}$-linéaire de $\mathbf{Z}/n\mathbf{Z}$.

11.6.9.1. — Il découle des constructions de 6.2.9 *et sq.* que $\rho\colon D_n \to \mathrm{GL}_2(\mathbf{R})$ est donnée par les formules suivantes :

$$\begin{array}{rcl} (\overline{a}, \overline{0}) & \mapsto & \begin{pmatrix} \cos(2a\pi/n) & -\sin(2a\pi/n) \\ \sin(2a\pi/n) & \cos(2a\pi/n) \end{pmatrix} \\ (\overline{a}, \overline{1}) & \mapsto & \begin{pmatrix} \cos(2a\pi/n) & \sin(2a\pi/n) \\ \sin(2a\pi/n) & -\cos(2a\pi/n) \end{pmatrix} \end{array}$$

11.6.9.2. — *La représentation $(\mathbf{R}^2, \rho_{|\mathbf{Z}/n\mathbf{Z}})$ est irréductible. En effet, comme $\mathbf{R}^2$ est de dimension 2, il suffit de s'assurer que $\mathbf{R}^2$ ne contient aucune droite stable sous l'action de $\mathbf{Z}/n\mathbf{Z}$. Or si D était une telle droite, elle ser*ait en particulier stable sous $\rho(\overline{1}, \overline{0})$, qui n'est autre que la rotation de matrice

$$\begin{pmatrix} \cos(2\pi/n) & -\sin(2\pi/n) \\ \sin(2\pi/n) & \cos(2\pi/n) \end{pmatrix}.$$

Son polynôme caractéristique est $X^2 - 2\cos(2\pi/n)X + 1$, qui a deux racines complexes $e^{2i\pi/n}$ et $e^{-2i\pi/n}$ qui ne sont pas réelles puisque $n \geqslant 3$. Cette rotation n'ayant pas de valeur propre réelle, elle n'a pas de droite stable dans $\mathbf{R}^2$, ce qui achève la preuve.

11.6.9.3. — *La représentation $(\mathbf{R}^2, \rho)$ est irréductible.* C'est une conséquence triviale de 11.6.9.2, car une droite de $\mathbf{R}^2$ stable sous l'action de D_n le serait *a fortiori* sous celle de $\mathbf{Z}/n\mathbf{Z}$.

11.6.9.4. — *La représentation $(\mathbf{C}^2, \rho|_{\mathbf{Z}/n\mathbf{Z}})$ est décomposable, et n'est* a fortiori *pas irréductible.* En effet, $\rho(\overline{1}, \overline{0})$ a d'après 11.6.9.2 deux valeurs propres complexes distinctes, et est donc diagonalisable en tant qu'endomorphisme de $\mathbf{C}^2$. Il existe dès lors deux droites vectorielles complexes Δ et Δ' de $\mathbf{C}^2$ stables sous $\rho(\overline{1}, \overline{0})$ telles que

$\mathbf{C}^2 = \Delta \oplus \Delta'$. Comme $\rho(\overline{k}, \overline{0}) = \rho(\overline{1}, \overline{0})^k$ pour tout k, chacune de ces droites est stable sous l'action de $\mathbf{Z}/n\mathbf{Z}$, et est en conséquence une sous-représentation de $(\mathbf{C}^2, \rho|_{\mathbf{Z}/n\mathbf{Z}})$; ainsi, cette dernière est décomposable.

11.6.9.5. — *La représentation* $(\mathbf{C}^2, \rho)$ *est irréductible.* En effet, supposons qu'il existe une droite Λ de $\mathbf{C}^2$ stable sous l'action de D_n. Elle est en particulier stable sous l'action de $\rho(\overline{0}, \overline{1})$, qui est la symétrie de matrice

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Celle-ci étant diagonale avec deux valeurs propres distinctes, ses seules droites stables sont $\mathbf{C} \cdot (0,1)$ et $\mathbf{C} \cdot (1,0)$, et Λ est donc l'une de ces deux droites. Mais comme

$$\underbrace{(\overline{1}, \overline{0})}_{\in D_n} \cdot \underbrace{(0,1)}_{\in \mathbf{C}^2} = (-\sin(2\pi/n), \cos(2\pi/n))$$

et

$$\underbrace{(\overline{1}, \overline{0})}_{\in D_n} \cdot \underbrace{(1,0)}_{\in \mathbf{C}^2} = (\cos(2\pi/n), \sin(2\pi/n))$$

et comme $\cos(2\pi/n)$ et $\sin(2\pi/n)$ sont tous deux non nuls, on voit qu'aucune de ces deux droites n'est stable sous l'action de D_n, et l'on aboutit ainsi à une contradiction.

11.6.10. *Étude de* $V_{3,k}$. — Rappelons la contruction de la représentation k-linéaire $V_{3,k}$ de $\mathfrak{S}_3$ (11.3.3.2, 11.3.3.3). On fait opérer le groupe $\mathfrak{S}_3$ tautologiquement sur $\{1, 2, 3\}$, et $V_{3,k}$ est alors la sous-représentation de $k\{1, 2, 3\}$ constituée des vecteurs de la forme $a[1] + b[2] + c[3]$ avec $a + b + c = 0$; c'est un plan vectoriel. Pour toute permutation $\sigma \in \mathfrak{S}_3$, on note u_σ l'automorphisme k-linéaire de $V_{3,k}$ induit par σ.

11.6.10.1. *Le cas où la caractéristique de k est différente de* 2 *et de* 3. — Nous allons montrer que $V_{3,k}$ est alors irréductible. Le vecteur $[1]-[2]$ de $V_{3,k}$ est un vecteur propre de $u_{(12)}$, associé à la valeur propre -1 ; et le vecteur $[1] + [2] - 2[3]$ de $V_{3,k}$ est un vecteur propre de $u_{(12)}$, associé à la valeur propre 1. Comme $1 \neq -1$ dans le corps k (il est de caractéristique différente de 2), l'endomorphisme $u_{(12)}$ a deux valeurs propres distinctes, 1 et -1 ; il est donc diagonalisable. Par ce qui précède, son sous-espace propre associé à 1 est $k([1] + [2] - 2[3])$; celui associé à -1 est $k([1] - [2])$.

Si la représentation $V_{3,k}$ n'était pas irréductible, elle possèderait une sous-représentation de dimension 1, c'est-à-dire une droite Δ stable sous l'action de $\mathfrak{S}_3$. Étant en particulier stable sous $u_{(12)}$, la droite Δ serait engendrée par un vecteur propre de $u_{(12)}$, et serait donc égale ou bien à $k([1]+[2]-2[3])$, ou bien à $k([1]-[2])$. Mais aucune de ces deux droites n'est stable sous $u_{(23)}$, car $[1]-[3]$ n'est pas colinéaire à $[1] - [2]$ et $[1] - 2[2] + [3]$ n'est pas colinéaire à $[1] + [2] - 2[3]$ (en effet, $-2 \neq 1$ dans k puisque la caractéristique de k est différente de 3) ; on aboutit ainsi à une contradiction.

11.6.10.2. *Le cas où la caractéristique de k est* 3. — On vérifie alors immédiatement que la droite $k([1] + [2] + [3])$ est une sous-représentation de $V_{3,k}$, et cette dernière n'est donc pas irréductible.

11.6.10.3. *Le cas où la caractéristique de k est* 2. — Nous allons montrer que $V_{3,k}$ est irréductible. Les vecteurs $[1]+[2]$ et $[1]+[3]$ appartiennent à $V_{3,k}$ et sont échangés par $u_{(2,3)}$, qui n'est donc pas égal à l'identité. Comme $u_{(2,3)}^2 = \mathrm{Id}_{V_{3,k}}$, l'endomorphisme $u_{(2,3)}$ de $V_{3,k}$ annule $X^2 - 1 = (X-1)^2$ et admet dès lors 1 comme seule valeur propre. Puisqu'il n'est pas égal à l'identité, le sous-espace propre correspondant est une droite, et c'est donc nécessairement $k([2]+[3])$; celle-ci est par conséquent la seule droite de $V_{3,k}$ stable sous $u_{(2,3)}$. Par symétrie, $k([1]+[2])$ est la seule droite de $V_{3,k}$ stable sous $u_{(1,2)}$. Il n'existe donc pas de droite de $V_{3,k}$ stable à la fois sous $u_{(12)}$ et $u_{(23)}$; il n'existe *a fortiori* aucune droite de $V_{3,k}$ stable sous $\mathfrak{S}_3$, et $V_{3,k}$ est dès lors irréductible.

Lemme 11.6.11. — *Soit G un groupe et soient V et W deux représentations k-linéaires irréductibles de G.*

(1) *Tout morphisme de représentations de V vers W est ou bien nul, ou bien un isomorphisme.*

(2) *L'anneau* $\mathrm{End}_G(V)$ *est un* anneau à division, *c'est-à-dire qu'il est non nul et que chacun de ses éléments non nuls est inversible (la différence avec un corps étant qu'on n'exige pas la commutativité) ; cette assertion est appelée le* Lemme de Schur.

(3) *Si V est de dimension finie et si k est algébriquement clos alors* $\mathrm{End}_G(V) = k$ *; autrement dit, tout endomorphisme de représentation de V est une homothétie.*

Démonstration. — Soit φ un morphisme de représentations de V vers W. Supposons que φ soit non nul. Le noyau de φ est une sous-représentation stricte de V ; par irréductibilité de V, il vient $\mathrm{Ker}\varphi = \{0\}$ et φ est injectif. L'image de φ est une sous-représentation non nulle de W ; par irréductiblité de W, il vient $\mathrm{Im}\varphi = W$, et φ est surjectif. Ainsi, φ est bijectif, et (1) est vraie.

Pour montrer (2), on commence par remarquer que comme V est non nul (par définition d'une représentation irréductible), l'anneau $\mathrm{End}_G(V)$ est non nul, puisque Id_V est non nul. Il résulte par ailleurs de (1) que tout élément non nul de $\mathrm{End}_G(V)$ est bijectif, et partant inversible. Ainsi, (2) est vraie.

Montrons maintenant (3). Supposons donc que k est un corps algébriquement clos et que V est de dimension finie. Soit A la k-algèbre à division $\mathrm{End}_G(V)$. En tant que sous-algèbre de $\mathrm{End}_k(V)$, elle est de dimension finie. Soit a un élément de A. La sous-algèbre commutative $k[a]$ de A est de dimension finie sur k ; la famille $(a^n)_{n\in\mathbf{N}}$ est donc liée, ce qui entraîne que a possède un polynôme annulateur non nul. L'idéal I de $k[X]$ formé des polynômes P tels que $P(a) = 0$ est donc non nul ; par principalité de $k(X]$, il est de la forme (Q) où Q est un polynôme unitaire. Si Q était égal à 1 on aurait $1 = 0$ dans $k[a] \subset A$, contredisant la non-nullité de A. Par conséquent $\deg Q > 0$ et Q possède dès lors une racine λ dans k puisque ce dernier est algébriquement clos. On peut écrire $Q = (X-\lambda)R$ avec $\deg R = \deg Q - 1$; le polynôme R n'est pas multiple de Q pour des raisons de degré, si bien que $R(a) \neq 0$. L'égalité $Q(a) = 0$ implique que $(a-\lambda)R(a) = 0$. En multipliant à droite par l'élément $R(a)^{-1}$ de l'anneau à division A on voit que $a - \lambda = 0$, c'est-à-dire que $a = \lambda$. □

Exemple 11.6.12. — Nous allons donner un exemple de représentation irréductible dont l'algèbre des endomorphismes est un anneau à division non commutatif. Pour cela, introduisons brièvement l'anneau $\mathbf{H}$ des *quaternions*, qui se décrit comme suit : c'est un $\mathbf{R}$-espace vectoriel de la forme

$$\mathbf{R} \oplus \mathbf{R}i \oplus \mathbf{R}j \oplus \mathbf{R}k,$$

muni d'une loi interne définie par bilinéarité à partir des formules suivantes :

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

On vérifie que cette loi interne est associative et fait de $\mathbf{H}$ une $\mathbf{R}$-algèbre, et que si a, b, c, d sont quatre réels non tous nuls alors l'élément $a + bi + cj + dk$ de $\mathbf{H}$ est inversible d'inverse

$$\frac{a - bi - cj - dk}{a^2 + b^2 + c^2 + d^2}.$$

En conséquence, $\mathbf{H}$ est une algèbre à division, manifestement non commutative (i et j ne commutent pas, par exemple). On note H_8 l'ensemble $\{1, -1, i, -i, j, -j, k, -k\}$. C'est un sous-groupe de $\mathbf{H}^\times$, appelé le *groupe quaternionique.* Il n'est pas abélien (là encore, par exemple, parce que i et j ne commutent pas) ; le lecteur pourra vérifier qu'en dépit de ce défaut, tous les sous-groupes de H_8 sont distingués (nous ne nous en servirons pas et l'indiquons à titre de curiosité).

On fait opérer H_8 sur $\mathbf{H}$ par multiplication à gauche. Cette opération permet de voir $\mathbf{H}$ comme une représentation $\mathbf{R}$-linéaire de H_8, et nous allons tout d'abord montrer qu'elle est irréductible. Soit V un sous-$\mathbf{R}$-espace vectoriel de $\mathbf{H}$ stable sous l'action de H_8. Il est alors stable par multiplication à gauche par i, j et k, et donc, par linéarité, par multiplication à gauche par n'importe quel élément h de $\mathbf{H}$. Supposons $V \neq 0$ et soit v un élément non nul de V. Pour tout $h \in \mathbf{H}$, on a $h = (hv^{-1})v$, lequel appartient à V par ce qui précède. Ainsi $V = \mathbf{H}$ et $\mathbf{H}$ est irréductible.

L'anneau des endomorphismes de la représentation $\mathbf{H}$ *de* H_8 *s'identifie à l'anneau à division* $\mathbf{H}^{\mathrm{op}}$, où $\mathbf{H}^{\mathrm{op}}$ désigne la $\mathbf{R}$-algèbre qui admet $\mathbf{H}$ comme espace vectoriel sous-jacent et dont la multiplication $*$ est définie par la formule $a * b = ba$. En effet, soit $a \in \mathbf{H}$ et soit d_a la multiplication à droite par a. Cet endomorphisme $\mathbf{R}$-linéaire de $\mathbf{H}$ commute à la multiplication à gauche par les éléments de $\mathbf{H}$, et en particulier à l'action de H_8. Comme $d_{ab} = d_b \circ d_a$, l'application $a \mapsto d_a$ est un morphisme d'algèbres de $\mathbf{H}^{\mathrm{op}}$ dans $\mathrm{End}_{H_8}\mathbf{H}$. Ce morphisme est injectif : en effet si $d_a = 0$ alors $d_a(1) = a = 0$. Et ce morphisme est surjectif. En effet, soit $f \in \mathrm{End}_{H_8}\mathbf{H}$. Posons $a = f(1)$. Comme f commute à l'action de H_8, on a

$$f(i) = \underbrace{f(i \cdot 1) = i \cdot f(1)}_{\text{Le point désigne l'action de } H_8} = ia$$

et de même $f(j) = ja$ et $f(k) = ka$. Par linéarité, $f(h) = ha$ quel que soit $h \in \mathbf{H}$, et l'on a donc $f = d_a$.

Remarque 11.6.13. — Soit $\mathbf{H_C}$ le $\mathbf{C}$-espace vectoriel $\mathbf{C} \oplus \mathbf{C}i \oplus \mathbf{C}j \oplus \mathbf{C}k$, muni de la loi interne définie par bilinéarité à partir des formules $(*)$. Cette loi est associative et fait de $\mathbf{H_C}$ une $\mathbf{C}$-algèbre, mais ce n'est pas une algèbre à division ; on peut montrer plus précisément qu'elle est isomorphe à $\mathrm{M}_2(\mathbf{C})$.

Remarque 11.6.14. — L'anneau $\mathbf{H}^{\mathrm{op}}$ est en fait isomorphe à $\mathbf{H}$, *via* une sorte de conjugaison : l'application qui envoie $z = a + bi + cj + dk$ sur $\overline{z} := a - bi - cj - dk$ est en effet une involution $\mathbf{R}$-linéaire telle que $\overline{z_1 z_2} = \overline{z_2 z_1}$ pour tout $(z_1, z_2) \in \mathbf{H}^2$. Elle induit donc un isomorphisme de $\mathbf{H}$ sur $\mathbf{H}^{\mathrm{op}}$.

11.7. Suites de Jordan-Hölder, représentations semi-simples. — Nous allons maintenant voir qu'on a une théorie des suites de Jordan-Hölder de représentations, tout à fait parallèle à celle que nous avons rencontrée en théorie des groupes – le rôle des groupes simples y est tenu par les représentations irréductibles.

Nous nous intéresserons ensuite à une classe spécifique de représentations admettant une suite de Jordan-Hölder, celle des représentations *semi-simples*.

On fixe un groupe G ; dans toute la sous-section 11.7, le terme *représentation* signifiera «représentation k-linéaire de G».

Définition 11.7.1. — Soit V une représentation. Soit n un entier. On appellera *suite de composition de longueur n de V* toute suite

$$\{0\} = V_n \subset V_{n-1} \subset \ldots \subset V_1 \subset V_0 = V$$

de sous-représentations de V. On ne demande pas que les V_i soient deux à deux distinctes, et que le cas limite $n = 0$ est autorisé (mais n'est possible que si $V = \{0\}$).

Les V_i/V_{i+1} sont appelés les *quotients successifs* de la suite.

Définition 11.7.2. — Soit V une représentation. Une *suite de Jordan-Hölder de V* est une suite de composition de V dont tous les quotients successifs sont des représentations irréductibles.

11.7.3. — Soit V une représentation et soit

$$\mathscr{S} = (\{0\} = V_n \subset V_{n-1} \subset \ldots \subset V_1 \subset V_0 = V)$$

une suite de Jordan-Hölder de V. Pour toute représentation irréductible W, on appelle *multiplicité de W dans $\mathscr{S}$*, et l'on note $\mu(\mathscr{S}, W)$, le nombre d'indices i tels que la représentation quotient V_i/V_{i+1} soit isomorphe à W.

11.7.4. — Nous allons maintenant énoncer les analogues en théorie des représentations des théorèmes sur les suites de Jordan-Hölder vus en théorie des groupes. Nous ne donnerons pas de démonstrations : les preuves sont *mutatis mutandis* les mêmes ici que dans le monde des groupes ; il convient de remplacer partout «groupe» par «représentation», «groupe simple» par «représentation irréductible», «cardinal» par «dimension», et de supprimer l'adjectif «distingué».

Lemme 11.7.5. — *Soit V une représentation de dimension finie. Elle possède une suite de Jordan-Hölder.*

Théorème 11.7.6. — *Soit V une représentation. Pour tout couple $(\mathscr{S}, \mathscr{S}')$ de suites de Jordan-Hölder de V et toute représentation irréductible W, les multiplicités $\mu(\mathscr{S}, W)$ et $\mu(\mathscr{S}', W)$ coïncident ; en particulier, $\mathscr{S}$ et $\mathscr{S}'$ ont même longueur.*

Notation 11.7.7. — Soit V une représentation. Si V admet une suite de Jordan-Hölder $\mathscr{S}$ alors pour toute représentation irréductible W, l'entier $\mu(\mathscr{S}, W)$ est indépendant du choix de $\mathscr{S}$; nous le noterons simplement $\mu(G, W)$ et l'appellerons la *multiplicité de W dans V*.

Proposition 11.7.8. — *Soit V une représentation et soit V' une sous-représentation de V. Les assertions suivantes sont équivalentes :*

(i) *chacune des représentations V' et V/V' possède une suite de Jordan-Hölder ;*

(ii) *V possède une suite de Jordan-Hölder.*

De plus si elles sont satisfaites alors $\mu(V, W) = \mu(V', W) + \mu(V/V', W)$ pour toute représentation irréductible W.

Nous allons maintenant introduire une classe particulièrement simple de représentations admettant une suite de Jordan-Hölder.

Définition 11.7.9. — Soit V une représentation. On dit que V est *semi-simple* si elle est somme directe de sous-représentations irréductibles.

Exemple 11.7.10. — Soit V un k-espace vectoriel sur lequel G agit trivialement. En écrivant V comme une somme directe de droites (il suffit pour ce faire de choisir une base de V), on voit que V est somme directe de sous-représentations toutes isomorphes à $\mathbf{1}_{G,k}$; elle est donc semi-simple.

11.7.11. — Soit V une représentation. Supposons que V s'écrive comme une somme directe *finie* $\bigoplus_{i \in I} V_i$ de sous-représentations irréductibles (c'est par exemple le cas si V est semi-simple et de dimension finie).

Soit n le cardinal de I. Identifions arbiRairement I à $\{1, \ldots, n\}$. La suite

$$V_n = \{0\} \subset V_1 \subset V_1 \oplus V_2 \subset \ldots \subset V_1 \oplus \ldots \oplus V_{n-1} \subset V$$

est alors une suite de composition de V, dont les quotients successifs sont les V_i ; c'est par conséquent une suite de Jordan-Hölder de V.

Si W est une représentations irréductible, la multiplicité $\mu(V, W)$ est égale au nombre d'indices i tel que $V_i \simeq_G W$. *Ce nombre ne dépend donc pas de la décomposition finie $V = \bigoplus V_i$ choisie.*

Remarque 11.7.12. — Soit V une représentation. Supposons que V soit semi-simple, et choisissons une décomposition $V = \bigoplus_{i \in I} V_i$ où les V_i sont des sous-représentations irréductibles. Soit W une représentation irréductible. Il résulte de 11.7.11 que si I est fini, le cardinal de l'ensemble $\{i \in I, V_i \simeq_G W\}$ ne dépend pas de la décomposition. Ce résultat reste en fait vrai sans hypothèse de finitude sur I, mais nous ne le démontrerons pas ici (et nous ne nous en servirons pas).

Lemme-définition 11.7.13. — *Soit V une représentation semi-simple. Choisissons une écriture $V = \bigoplus_{i \in I} V_i$ où les V_i sont des sous-représentations irréductibles. Soit W une représentation irréductible et soit J le sous-ensemble de I constitué des indices i tels que $V_i \simeq_G W$. La sous-représentation $V' := \bigoplus_{i \in J} V_i$ est engendrée (comme sous-espace vectoriel) par les sous-représentations de V de la forme $f(W)$ où $f \colon W \to V$ est un morphisme équivariant. En particulier, V' ne dépend que de W*

et pas de la décomposition $V = \bigoplus V_i$. *On dit que* V' *est la* composante isotypique *relative à* W.

Démonstration. — Soit V'' le sous-espace vectoriel de V engendré par les $f(W)$ où f est comme dans l'énoncé ; c'est une sous-représentation de V et nous allons montrer que $V' = V''$.

Pour tout $i \in J$ il existe un isomorphisme équivariant $u_i : W \to V_i$, que l'on peut voir comme un morphisme équivariant de W dans V d'image V_i. Comme V' est la somme (directe) des $u_i(W)$ il vient $V' \subset V''$.

Réciproquement, soit f un morphisme équivariant de W dans V. Montrons que $f(W) \subset V'$, ce qui prouvera que $V'' \subset V'$ et achèvera la démonstration. Soit i un élément de I qui n'appartient pas à J, et soit p_i la projection de V sur V_i parallèlement à $\bigoplus_{j \in I, j \neq i} V_j$. La composée $p_i \circ f$ est un morphisme équivariant de W vers V_i. Comme W et V_i sont irréductibles, ce morphisme est ou bien nul, ou bien un isomorphisme (lemme 11.6.11) ; comme $i \notin J$ ce ne peut être un isomorphisme, si bien que $p_i \circ f = 0$. Ceci valant pour tout i n'appartenant pas à J, il vient $f(W) \subset \bigoplus_{i \in J} V_i = V'$. □

Remarque 11.7.14. — Soit V une représentation semi-simple. Supposons que V admette une écriture $V = \bigoplus V_i$ comme somme directe de sous-représentations irréductibles, et que les V_i soient deux à deux non isomorphes. Cette écriture de V comme somme directe de sous-représentations irréductibles est alors la seule à permutation des sommandes près.

En effet, donnons-nous une autre décomposition $V = \bigoplus W_\ell$. Fixons i. Comme V_i n'est isomorphe à aucun des V_j pour $j \neq i$, la composante isotypique de V relative à V_i est V_i. Cette composante est somme directe de certains des W_ℓ (ceux qui sont isomorphes à V_i). Mais comme V_i est irréductible, il existe nécessairement un indice $s(i)$ tel que $V_i = W_{s(i)}$. On a alors $V = \bigoplus W_{s(i)}$, ce qui veut dire que tout indice ℓ est de la forme $s(i)$ pour un indice i (qui est unique car on a alors $W_\ell = V_i$), d'où notre assertion.

11.8. Irréductibilité et semi-simplicité dans le cas d'un groupe fini. — Nous allons maintenant établir quelques résultats relatifs à l'irréductibilité et à la semi-simplicité sur les représentations k-linéaires (le plus souvent de dimension finie) d'un groupe fini. Comme nous allons le voir, la situation dépend très fortement du fait que le cardinal du groupe soit on non nul dans le corps k.

Lemme 11.8.1. — *Supposons que* k *est algébriquement clos (exemple :* $k = \mathbf{C}$*), et soit* G *un groupe* abélien *fini dont le cardinal* n *est non nul dans* k. *Soit* V *une représentation* k*-linéaire de* G *de dimension finie.*

(1) *La représentation* V *est somme directe de sous-représentations de dimension 1 ; elle est en particulier semi-simple.*

(2) *Les assertions suivantes sont équivalentes :*

  (i) V *est irréductible ;*

  (ii) V *est indécomposable ;*

(iii) *V est de dimension 1.*

Démonstration. — Soit Γ l'image de G dans $\mathrm{GL}(V)$ par le morphisme qui définit la représentation étudiée. Comme G est abélien, Γ l'est aussi ; comme tout élément de G est annulé par n, on a $\gamma^n = \mathrm{Id}_V$ pour tout $\gamma \in \Gamma$. Puisque k est algébriquement clos, le polynôme $X^n - 1$ est scindé dans k. Ses racines y sont simples : en effet, sa dérivée est égale à nX^{n-1}, et elle ne s'annule qu'en 0 puisque n est non nul dans k ; en particulier, elle n'a pas de racine commune avec $X^n - 1$, d'où notre assertion. Il s'ensuit que tout élément de Γ est diagonalisable. Le groupe Γ étant abélien, il existe alors une base $(e_1, \ldots, e_m)$ de V qui est une base de diagonalisation simultanée de tous les éléments de Γ. Chacune des droites ke_i est stable sous les éléments de Γ, et partant sous l'action de G ; les droites ke_i sont donc des sous-représentations de V, et l'on a $V = \bigoplus_i ke_i$, d'où (1). L'assertion (2) en découle immédiatement. ☐

Lemme 11.8.2. — *Soit p un nombre premier et soit G un p-groupe. Supposons que k est de caractéristique p.*

(1) *Une représentation k-linéaire de G est irréductible si et seulement si elle est isomorphe à $\mathbf{1}_{G,k}$.*
(2) *La représentation unité $\mathbf{1}_{G,k}$ est à isomorphisme près la seule représentation k-linéaire de dimension 1 de G.*
(3) *Une représentation W de G est semi-simple si et seulement si G agit trivialement sur W.*

Démonstration. — Soit V une représentation k-linéaire irréductible de G. L'espace V est par définition non nul, et il résulte alors du lemme 4.4.9 que $V^G \neq \{0\}$. Choisissons un vecteur non nul v de V^G. Puisque G agit trivialement sur V, la droite kv est une sous-représentation de V, isomorphe à $\mathbf{1}_{G,k}$. Puisque V est irréductible, on a $V = kv$. Ceci achève de montrer (1) ; les assertions (2) et (3) en sont des conséquences immédiates. ☐

Théorème 11.8.3. — *Soit G un groupe de cardinal fini n tel que n soit non nul dans k. Soit V une représentation k-linéaire de G et soit W une sous-représentation de V. Il existe alors une sous-représentation W' de V telle que l'on ait $V = W \oplus W'$.*

Démonstration. — On commence par choisir un supplémentaire arbitraire W_0 de W dans W. Soit p la projection de V sur W parallèlement à W_0. On pose

$$q = \frac{1}{n} \sum_{g \in G} gp \in \mathrm{End}_k(V)$$

(la notation gp fait référence à l'opération naturelle de G sur $\mathrm{End}_k(V)$, *cf.* 11.2.1 ; notez que $\frac{1}{n}$ existe bien dans k d'après nos hypothèses). L'endomorphisme q appartient à $\mathrm{End}_k(V)^G$ (*cf.* 11.1.8), ce qui signifie que q est équivariant (11.2.2).

L'endomorphisme q est un projecteur d'image W. En vertu du lemme 10.5.2, il suffit de s'assurer que $q(v) \in W$ pour tout $v \in V$, et que $q(v) = v$ pour tout $v \in W$. Soit $v \in V$. Pour tout $g \in G$ l'élément $w_g := p(g^{-1}v)$ appartient à W (puisque W

est l'image de p) ; comme W est une sous-représentation de G, l'élément gw_g de V appartient encore à W. On en déduit que le vecteur

$$q(v) = \frac{1}{n}\sum_g (gp)(v) = \frac{1}{n}\sum_g gw_g$$

appartient à W. Supposons que $v \in W$. Soit $g \in G$. Le vecteur $g^{-1}v$ appartient encore à W (car ce dernier est une sous-représentation de V) et est donc invariant sous p ; il vient $w_g = g^{-1}v$, puis $gw_g = v$. Ainsi

$$q(v) = \frac{1}{n}\sum_{g\in G} v = \frac{1}{n}\cdot nv = v,$$

ce qui achève la preuve.

Conclusion. Le noyau W' de q est alors un supplémentaire de W, et est une sous-représentation de V puisque q est équivariant. □

Théorème 11.8.4. — *Soit G un groupe fini et soit k un corps dans lequel $|G|$ est non nul. Soit V une représentation k-linéaire de dimension finie de G. La représentation V est semi-simple.*

Démonstration. — Il s'agit de montrer que V est somme directe (nécessairement finie) de sous-représentations irréductibles. On procède par récurrence sur la dimension de V. Si $V = \{0\}$ la famille *vide* de sous-représentations irréductibles de V convient. Supposons maintenant $\dim V > 0$, et l'assertion (1) vraie en dimension strictement inférieure à celle de V.

Si la représentation V est irréductible, l'assertion est vraie en prenant pour (V_i) la famille singleton $\{V\}$. Si V n'est pas irréductible, elle possède une sous-représentation stricte et non nulle V'. Le théorème 11.8.3 affirme l'existence d'une sous-représentation V'' de V telle que $V = V' \oplus V''$ (c'est ici qu'on utilise le fait que $|G|$ est non nul dans k). Comme V' est stricte et non nulle, V'' est stricte (et non nulle). On peut dès lors appliquer l'hypothèse de récurrence à V' et à V'' : chacune d'elle se décompose en une somme directe finie de représentations irréductibles. En concaténant ces deux décompositions, on obtient une décomposition de V de la forme souhaitée. □

Remarque 11.8.5. — Soient G, k et V comme dans l'énoncé du théorème 11.8.4 ci-dessus. On déduit immédiatement de l'assertion (1) que V est indécomposable si et seulement si elle est irréductible. Notons l'importance de l'hypothèse que $|G|$ est non nul dans k : on a vu en 11.6.8 et 11.6.8.2 que si k est de caractéristique 2, la k-linéarisation de l'opération tautologique de $\mathfrak{S}_2$ sur $\{1, 2\}$ est indécomposable, mais pas irréductible.

Remarque 11.8.6. — Soit G un groupe fini et soit k un corps tel que $|G|$ soit non nul dans k. Il résulte du théorème 11.8.4 et de 11.7.11 que pour classer à isomorphisme près les représentations k-linéaires de G de dimension finie, il suffit de savoir classer les représentations k-linéaires de G de dimension finie qui sont *irréductibles*. Lorsque $k = \mathbf{C}$ (auquel cas $|G|$ est non nul dans k quel que soit le groupe fini G), ce sera l'objet de la section suivante.

Exemple 11.8.7. — Soit k un corps tel que 6 soit non nul dans k (en d'autres termes, la caractéristique de k est différente de 2 et 3). D'après 11.3.3 *et sq.*, k-linéarisation $k\{1,2,3\}$ de l'opération tautologique de $\mathfrak{S}_3$ sur $\{1,2,3\}$ est somme directe de deux sous-représentations : l'une est la droite $k(e_1+e_2+e_3)$ et est isomorphe à $\mathbf{1}_{\mathfrak{S}_3,k}$ (en d'autres termes, $\mathfrak{S}_3$ agit trivialement dessus) ; l'autre est le plan $V_{3,k}$ défini comme l'ensemble des vecteurs de la forme $ae_1+be_2+ce_3$ avec $a+b+c=0$; on a vu au 11.6.10.1 que $V_{3,k}$ est irréductible.

L'écriture $k\{1,2,3\}=k(e_1+e_2+e_3)\bigoplus V_{3,k}$ est donc un exemple de décomposition comme somme directe de représentations irréductibles, dont l'existence est prédite par le théorème 11.8.4. Les deux représentations irréductibles $k(e_1+e_2+e_3)\simeq\mathbf{1}_{\mathfrak{S}_3,k}$ et $V_{3,k}$ sont évidemment non isomorphes (elles ne sont pas de même dimension !). Chacune est donc de multiplicité 1 dans $k\{1,2,3\}$, et toute représentation k-linéaire irréductible de $\mathfrak{S}_3$ qui n'est isomorphe ni à $\mathbf{1}_{\mathfrak{S}_3,k}$ ni à $V_{3,k}$ est de multiplicité nulle dans $k\{1,2,3\}$. C'est par exemple le cas de la représentation signature $\Sigma_{3,k}$ (exemple 11.1.12) ; elle n'est pas isomorphe à $V_{3,k}$ car elle est de dimension 1, et n'est pas isomorphe à $\mathbf{1}_{\mathfrak{S}_3,k}$ car $\mathfrak{S}_3$ n'agit pas trivialement sur $\Sigma_{3,k}$, *cf. loc. cit.*

Signalons par ailleurs qu'en vertu de la remarque 11.7.14, cette écriture de $k\{1,2,3\}$ comme somme directe de représentations irréductibles est la seule à échange près des deux sommandes.

12. Représentations complexes des groupes finis

Nous allons désormais nous intéresser aux représentations $\mathbf{C}$-linéaires de dimension finie d'un groupe fini. *Dans toute cette section, l'expression «représentation» signifiera donc, sauf mention expresse du contraire, «représentation* $\mathbf{C}$*-linéaire de dimension finie».*

12.1. Orthogonalité des caractères et conséquences. — On fixe un groupe fini G ; notons que $|G|$ est non nul dans $\mathbf{C}$.

12.1.1. *Rappels sur les représentations de G.* — Soit V une représentation de G ; les faits suivants résultent du théorème 11.8.4. et de 11.7.11 (qui est lui-même une conséquence du théorème 11.7.6).

La représentation V admet une écriture $V = \bigoplus V_i$ comme somme directe finie de sous-représentations irréductibles (et est donc irréductible si et seulement si elle est indécomposable) ; de plus, pour toute représentation irréductible W de G, le cardinal de $\{i, V_i \simeq G\}$ ne dépend que de V et W, et pas du choix de la décomposition $V = \bigoplus V_i$; on l'appelle la multiplicité de W dans V.

12.1.2. *Reformulation.* — Reformulons un peu différemment le résultat rappelé au 12.1.1 ci-dessus. Un *système complet de représentations irréductibles de G* est un ensemble $\mathscr{W}$ de représentations irréductibles de G tel que chaque représentation irréductible de G soit isomorphe à une et une seule représentation appartenant à $\mathscr{W}$. Donnons-nous un tel système $\mathscr{W}$, et soit V une représentation de G. Il existe alors une unique famille $(\mu_W)_{W\in\mathscr{W}}$ d'entiers (nécessairement presque tous nuls) tels que $V \simeq \bigoplus_{W\in\mathscr{W}} W^{\mu_W}$. Pour tout $W \in \mathscr{W}$, l'entier μ_W est la multiplicité de W dans V.

12.1.3. — Pour classer les représentations de G (à isomorphisme près), il suffit par ce qui précède de classer ses représentations irréductibles.

12.1.4. *Un produit hermitien sur $\mathbf{C}^G$.* — L'ensemble $\mathbf{C}^G$ des fonctions de G dans $\mathbf{C}$ est un $\mathbf{C}$-espace vectoriel de dimension $|G|$ (une base est donnée par la famille $(\delta_g)_{g\in G}$, où δ_g envoie g sur 1 et h sur 0 pour tout $h \neq g$). On le munit de la forme sesquilinéaire

$$(\chi, \varpi) \mapsto \langle \chi, \varpi \rangle := \frac{1}{|G|} \sum_{g\in G} \chi(g)\overline{\varpi(g)}.$$

On vérifie aussitôt que c'est un produit hermitien, pour lequel $(\sqrt{|G|}\delta_g)_g$ est une base orthonormée.

12.1.5. — Soit $\mathscr{C}$ l'ensemble des classes de conjugaison de G. On identifie le $\mathbf{C}$-espace vectoriel $\mathbf{C}^{\mathscr{C}}$ au sous-espace vectoriel de $\mathbf{C}^G$ constitué des fonctions *centrales* (11.5.4).

Soient χ et ϖ deux fonctions centrales de G vers $\mathbf{C}$, vues comme appartenant à $\mathbf{C}^{\mathscr{C}}$. On a alors

$$\langle \chi, \varpi \rangle = \frac{1}{|G|} \sum_{c\in\mathscr{C}} |c|\chi(c)\overline{\varpi(c)}.$$

***Théorème 12.1.6* (Orthogonalité des caractères).** — *Soient V et W deux représentations irréductibles de G. Le produit hermitien $\langle \chi_V, \chi_W \rangle$ est égal à 1 si V et W sont isomorphes, et à 0 sinon.*

Démonstration. — On a

$$\begin{aligned}
\langle \chi_V, \chi_W \rangle &= \frac{1}{|G|}\sum_{g\in G}\chi_V(g)\overline{\chi_W(g)} \\
&= \frac{1}{|G|}\sum_{g\in G}\chi_V(g)\chi_{W^*}(g) \qquad \text{(d'après 11.5.9)} \\
&= \frac{1}{|G|}\sum_{g\in G}\chi_{W^*\otimes_k V}(g) \qquad \text{(d'après 11.5.8)} \\
&= \frac{1}{|G|}\sum_{g\in G}\chi_{\mathrm{Hom}_k(W,V)}(g) \qquad \text{(d'après 11.4.3)} \\
&= \dim \mathrm{Hom}_k(W,V)^G \qquad \text{(d'après le lemme 11.5.13)} \\
&= \dim \mathrm{Hom}_G(W,V) \qquad \text{(d'après 11.2.2).}
\end{aligned}$$

Nous allons conclure en distinguant deux cas.

Supposons que W n'est pas isomorphe à V. Comme V et W sont irréductibles, l'assertion (1) du lemme 11.6.11 assure que $\mathrm{Hom}_G(W,V) = 0$; il résulte alors du calcul ci-dessus que $\langle \chi_V, \chi_W \rangle = 0$.

Supposons que $W \simeq V$. Dans ce cas le **C**-espace vectoriel $\mathrm{Hom}_G(W,V)$ est isomorphe à $\mathrm{End}_G(V)$, et comme V est irréductible cette dernière algèbre est égale à **C** d'après l'assertion (3) du lemme 11.6.11 ; il résulte alors du calcul ci-dessus que $\langle \chi_V, \chi_W \rangle = 1$. □

Corollaire 12.1.7. — *Fixons un système complet $\mathscr{W}$ de représentations irréductibles de G. Soit V une représentation de G. Pour toute $W \in \mathscr{W}$ de G, on note μ_W la multiplicité de W dans V.*

1. *On a $\langle \chi_V, \chi_U \rangle = \mu_U$ pour toute $U \in \mathscr{W}$.*
2. *On a $\langle \chi_V, \chi_V \rangle = \sum_{W\in\mathscr{W}} \mu_W^2$. En particulier, $\langle \chi_V, \chi_V \rangle = 1$ si et seulement si V est irréductible.*
3. *La classe d'isomorphie de V est entièrement déterminée par χ_V.*
4. *Si V est la représentation régulière de G (définition 11.3.2) alors $\mu_W = \dim W$ pour tout $W \in \mathscr{W}$.*
5. *L'ensemble $\mathscr{W}$ est fini et $\sum_{W\in\mathscr{W}}(\dim W)^2 = |G|$.*

Démonstration. Elle va reposer sur le fait suivant, conséquence directe du théorème 12.1.6 : si U et W sont deux éléments de $\mathscr{W}$ alors $\langle \chi_W, \chi_U \rangle$ est égal à

δ_{UW}. On a par ailleurs un isomorphisme $V \simeq \bigoplus_{W\in\mathscr{W}} W^{\mu_W}$, et partant l'égalité $\chi_V = \sum_{W\in\mathscr{W}} \mu_W\chi_W$ (11.5.7).

Preuve de (1). On a

$$\begin{aligned}\langle \chi_V, \chi_U\rangle &= \langle \sum_{W\in\mathscr{W}} \mu_W\chi_W, \chi_U\rangle \\ &= \sum_{W\in\mathscr{W}} \langle \mu_W\chi_W, \chi_U\rangle \\ &= \sum_{W\in\mathscr{W}} \mu_W\delta_{WU} \\ &= \mu_U.\end{aligned}$$

Preuve de (2). On a

$$\begin{aligned}\langle \chi_V, \chi_V\rangle &= \langle \sum_{W\in\mathscr{W}} \mu_W\chi_W, \sum_{U\in\mathscr{W}} \mu_U\chi_U\rangle \\ &= \sum_{W\in\mathscr{W}} \sum_{U\in\mathscr{W}} \mu_W\mu_U\langle \chi_W, \chi_U\rangle \\ &= \sum_{W\in\mathscr{W}} \sum_{U\in\mathscr{W}} \mu_W\mu_U\delta_{UW} \\ &= \sum_{W\in\mathscr{W}} \mu_W^2.\end{aligned}$$

Preuve de (3).On déduit de (1) que $V \simeq \bigoplus_{W\in\mathscr{W}} W^{\langle\chi_V,\chi_W\rangle}$, d'où (3).

Preuve de (4). Supposons que V est la représentation régulière de G. Pour tout $h \in G$, la translation $g \mapsto hg$ n'a aucun point fixe si $h \neq e$, et a $|G|$ points fixes si $h = e$; il en résulte que $\chi_V(e) = |G|$ et que $\chi_V(h) = 0$ pour tout $h \neq e$ (11.5.11.1). Soit $W \in \mathscr{W}$. On a alors

$$\mu_W = \langle \chi_V, \chi_W\rangle = \frac{1}{|G|}\sum_{g\in G} \chi_V(g)\overline{\chi_W(g)} = \overline{\chi_W(e)} = \dim W,$$

où la première égalité vient de l'assertion (1).

Preuve de (5). Supposons toujours que V est la représentation régulière de G. Comme V est isomorphe à $\bigoplus_{W\in\mathscr{W}} W^{\mu_W}$ on a

$$|G| = \dim V = \sum_{W\in\mathscr{W}} \mu_W \dim W = \sum_{W\in\mathscr{W}} (\dim W)^2,$$

où la dernière égalité provient de l'assertion (4). □

Proposition 12.1.8. — *Soit $\mathscr{C}$ l'ensemble des classes de conjugaison de G. Soit $\mathscr{W}$ un système complet de représentations irréductibles de G. La famille des χ_W pour W parcourant $\mathscr{W}$ est alors une base orthonormée du sous-espace vectoriel $\mathbf{C}^{\mathscr{C}}$ de l'espace hermitien $\mathbf{C}^G$. En particulier, $|\mathscr{W}| = |\mathscr{C}|$: le nombre de représentations irréductibles de G à isomorphisme près est égal au nombre de classes de conjugaison de G.*

Démonstration. — Soit E le sous-espace vectoriel de $\mathbf{C}^{\mathscr{C}}$ engendré par les χ_W pour W parcourant $\mathscr{W}$. Il résulte du théorème 12.1.6 que la famille $(\chi_W)_{W\in\mathscr{W}}$ est orthonormée ; c'est donc une base orthonormée de E. Il suffit pour conclure de démontrer que l'inclusion $E \subset \mathbf{C}^{\mathscr{C}}$ est une égalité. Nous allons démontrer que l'orthogonal de E dans $\mathbf{C}^{\mathscr{C}}$ est nul, ce qui permettra de conclure. Soit donc $\varphi \in \mathbf{C}^{\mathscr{C}}$ un élément orthogonal à E.

Soit $W \in \mathscr{W}$ et soit f l'endomorphisme $w \mapsto \sum_{g\in G} \overline{\varphi(g)} gw$ de W. Pour tout $h \in G$ l'élément hf de $\mathrm{End}_k(W)$ envoie un vecteur w sur

$$hf(h^{-1}w) = \sum_{g\in G} \overline{\varphi(g)} hgh^{-1}w = \sum_{g\in G} \overline{\varphi(hgh^{-1})} hgh^{-1}w = \sum_{g\in G} \overline{\varphi(g)} gw = f(w),$$

où la deuxième égalité provient du fait que φ est centrale. On a donc $hf = f$, ce qui signifie que f est G-équivariante ; c'est par conséquent un endomorphisme de la représentation W, et partant une homothétie puisque W est irréductible (11.6.11, assertion (3)). Soit λ le rapport de f. On a les égalités

$$\lambda \dim W = \mathrm{Tr}(f) = \sum_{g\in G} \overline{\varphi(g)} \chi_W(g) = |G|\langle \chi_W, \varphi\rangle.$$

Comme φ est orthogonale à E par hypothèse, ce dernier terme est nul ; puisque $\dim W > 0$ (une représentation irréductible est non nulle) il vient $\lambda = 0$, puis $f = 0$.

Soit U une représentation quelconque de G. En écrivant U comme somme directe de représentations irréductibles et en appliquant ce qui précède à chacun des sommandes, on voit que l'endomorphisme $u \mapsto \sum \overline{\varphi(g)} gu$ de U est nul. C'est en particulier le cas lorsque U est la représentation régulière de G. On a alors

$$0 = \sum_{g\in G} \overline{\varphi(g)} g[e] = \sum_{g\in G} \overline{\varphi(g)}[g].$$

Puisque $([g])_{g\in G}$ est une base de U, il vient $\varphi(g) = 0$ pour tout g ; ainsi φ est nulle, ce qu'il fallait démontrer. □

Commentaires 12.1.9. — La proposition 12.1.8 assure que le nombre de classes d'isomorphie de représentations irréductibles de G est égal à $|\mathscr{C}|$. Mais attention : *elle ne fournit aucune bijection naturelle permettant d'associer une représentation irréductible à un élément de $\mathscr{C}$.*

12.2. Interlude : éléments entiers sur un sous-anneau. — Pour démontrer le dernier théorème que nous avons en vue sur les représentations, nous allons procéder d'une manière relativement détournée, fondée sur la notion d'entier algébrique. Nous allons donc commencer par introduire celle-ci, et plus généralement par introduire la notion d'élément d'un anneau entier sur un sous-anneau.

Définition 12.2.1. — Soit B un anneau et soit A un sous-anneau de B. Soit C un sous-anneau de B contenant A. Nous dirons que C est *fini* sur A s'il existe une famille finie $(c_1, \ldots, c_r)$ d'éléments de C telle que tout élément de C s'écrive comme une combinaison A-linéaire des c_i (c'est-à-dire comme une somme $\sum a_i c_i$ où les a_i appartiennent à A).

12.2.2. *Transitivité de la finitude.* — Soit B un anneau et soient A, C et D trois sous-anneaux de B avec $A \subset C \subset D$. Si D est fini sur C et si C est fini sur A alors D est fini sur A. En effet, si l'on choisit une famille $(d_1, \ldots, d_r)$ d'éléments de D telle que tout élément de D soit une combinaison C-linéaire des d_i, puis une famille $(c_1, \ldots, c_s)$ d'éléments de C telle que tout élément de C soit une combinaison C-linéaire des c_j, il est immédiat que tout élément de D est une combinaison A-linéaire des $c_j d_i$ pour (i, j) parcourant $\{1, \ldots, r\} \times \{1, \ldots, s\}$.

Théorème-Définition 12.2.3. — *Soit B un anneau et soit A un sous-anneau de B. Soit $b \in B$. Les assertions suivantes sont équivalentes :*

(i) *Il existe un polynôme* unitaire *P à coefficients dans A tel que $P(b) = 0$.*

(ii) *Le sous-anneau $A[b]$ de B est fini sur A.*

(iii) *Il existe un sous-anneau C de B contenant A, contenant b, et fini sur A.*

Lorsqu'elles sont satisfaites on dit que b est entier *sur A.*

Démonstration. — Supposons que (i) soit vraie, et soit n le degré de P. L'égalité $P(b) = 0$ permet d'écrire b^n comme une combinaison A-linéaire des b^i pour $i \leqslant n-1$ (cela utilise de manière cruciale le fait que P est *unitaire*). Une récurrence immédiate sur m montre alors que b^m est pour tout $m \geqslant n$ une combinaison A-linéaire des b^i pour $i \leqslant n-1$. Il s'ensuit que tout élément de $A[b]$ est combinaison A-linéaire des b^i pour $i \leqslant n-1$, ce qui montre (ii).

L'implication iii)$\Rightarrow$(iii) est évidente. Supposons maintenant que (iii) est vraie. Soit $(c_1, \ldots, c_r)$ une famille d'éléments de C telle que tout élément de C soit combinaison A-linéaire des c_i.

Si u est une application A-linéaire de C dans lui-même, il existe une famille (α_{ij}) d'éléments de A telle que $u(c_j) = \sum_i \alpha_{ij} c_i$ pour tout j ; cette famille n'a aucune raison d'être unique, et nous dirons que la matrice $(\alpha_{ij}) \in \mathrm{M}_r(A)$ est *une* matrice de u dans la famille (c_i). On démontre exactement comme en algèbre linéaire que si u et v sont deux applications A-linéaires de C dans lui-même, si M est une matrice de u dans la famille (c_i) et si N est une matrice de v dans la famille (c_i), alors :

- ◇ pour tout $a \in A$, la matrice $M + aN$ est une matrice de $u + av$ dans la famille (c_i) ;
- ◇ MN est une matrice de $u \circ v$ dans la famille (c_i).

Soit maintenant M une matrice de l'application $u := c \mapsto bc$ dans la famille (c_i). Soit P le polynôme caractéristique de M, qui est unitaire et à coefficients dans A. Le théorème de Cayley-Hamilton (valable sur tout anneau commutatif, voir la remarque 12.2.4 ci-dessous) assure que $P(M) = 0$. Comme $P(M)$ est une matrice de l'application $P(u)$ dans la famille (c_i), il vient $P(u) = 0$. Mais $P(u)$ n'est autre que la multiplication par $P(b)$ et on a donc $P(b) = P(u)(1) = 0$, ce qui montre (i). □

Remarque 12.2.4 **(À propos du théorème de Cayley-Hamilton)**
Vous avez certainement vu des preuves du théorème de Cayley-Hamilton *sur un corps*. Mais sa validité sur un corps entraîne en fait sa validité sur tout anneau commutatif, comme nous allons maintenant l'expliquer.

Soit n un entier et soit $M = (a_{ij})$ une matrice carrée de taille n à coefficients dans un anneau commutatif A ; soit χ_M son polynôme caractéristique. Pour tout $m \leqslant n$, le coefficient de X^m dans χ_M est de la forme $Q_m(a_{ij})_{i,j}$ pour un certain polynôme Q_m à coefficients dans $\mathbf{Z}$ en les indéterminées X_{ij}, homogène de degré m, et qui est indépendant de A et de M. Il s'ensuit que pour tout couple (r, s) le terme d'indice (r, s) de la matrice $\chi_M(M)$ est de la forme $R_{rs}(a_{ij})_{ij}$, où R_{rs} est un polynôme à coefficient dans $\mathbf{Z}$ en les indéterminées X_{ij}, ne dépendant ni de A ni de M.

On veut montrer que $\chi_M(M) = 0$. Ceci revient à montrer que $R_{rs}(a_{ij})_{ij}$ est égal à 0 pour tout (r, s). Nous allons en fait prouver que R_{rs} est le polynôme nul, ce qui permettra de conclure. Pour ce faire, remarquons que si l'on note N la matrice (X_{ij}) appartenant à $\mathrm{M}_n(\mathbf{Z}[X_{ij}])_{i,j}$ alors R_{rs} est le coefficient d'indice (r, s) de $\chi_N(N)$. Or $\chi_N(N) = 0$ par le théorème de Cayley-Hamilton appliqué sur le *corps* $\mathbf{Q}(X_{ij})_{ij}$; on a donc bien $R_{rs} = 0$ pour tout (r, s).

Remarque 12.2.5. — Soient A, B et b comme dans le théorème 12.2.3. *Supposons de plus que A est un corps et que B est intègre.* Si b satisfait les conditions équivalentes du théorème, on dit le plus souvent que b est *algébrique* sur A plutôt qu'entier. Notons que dans ce cas la propriété (i) du théorème équivaut simplement à demander que b annule un polynôme non nul de $A[X]$, car si c'est le cas b annule aussi le quotient de ce polynôme par son coefficient dominant, qui est unitaire.

L'élément b est algébrique sur A si et seulement si $A[b]$ est un corps. En effet, supposons b algébrique sur A. Dans ce cas $A[b]$ est de dimension finie sur A par la propriété (ii) du théorème. Si x est un élément non nul de $A[b]$, la multiplication par x définit un endomorphisme de $A[b]$, qui est injectif car $A[b]$ est intègre (comme sous-anneau de B). Puisque $A[b]$ est de dimension finie, cet endomorphisme est surjectif ; en particulier, son image contient 1, ce qui veut dire que x est inversible ; ainsi, $A[b]$ est un corps. Réciproquement, supposons que $A[b]$ soit un corps. Si $b = 0$ alors b est algébrique sur A (il annule le polynôme X). Si b est non nul, il possède un inverse dans $A[b]$, que l'on peut écrire $\sum a_i b^i$. L'élément b annule alors le polynôme $X(\sum a_i X^i) - 1$, qui est non nul ; ainsi, b est algébrique sur A.

Théorème 12.2.6. — *Soit B un anneau et soit A un sous-anneau de B.*

(1) *L'ensemble C des éléments de B entiers sur A est un sous-anneau de B qui contient A, et c'est un corps si A est un corps et si B est intègre.*

(2) *Tout élément de B entier sur C est entier sur A.*

Démonstration. — Si a est un élément de A il annule le polynôme unitaie $X - a$ et il appartient donc à C. Par conséquent, $A \subset C$. Comme C contient A, il contient 0 et 1. Donnons-nous deux éléments b et c de C. Nous allons montrer que que $b - c$ et bc appartiennent à C, ce qui prouvera que C est un sous-anneau de B. Comme b appartient à C, il est entier sur A et $A[b]$ est donc fini sur A. Comme c appartient à C il est entier sur A, et *a fortiori* sur $A[b]$; par conséquent $A[b, c] = A[b][c]$ est fini sur $A[b]$. Il s'ensuit par transitivité que $A[b, c]$ est fini sur A (12.2.2). Comme $b - c$ et bc vivent dans $A[b, c]$, ils sont entiers sur A et appartiennent donc à C.

Supposons que A soit un corps et que B soit intègre. L'anneau C est alors non nul en tant que sous-anneau de B. Et si b est un élément de C, il résulte de la remarque

12.2.5 que $A[b]$ est un corps ; si b est non nul, il est donc inversible dans $A[b]$ et *a fortiori* dans C, ce qui montre que C est un corps et achève de prouver (1).

Montrons maintenant (2). Soit b un élément de B entier sur C. Il existe $n \geqslant 1$ et des éléments $\beta_0, \ldots, \beta_{n-1}$ dans C tels que $b^n + \beta_{n-1}b^{n-1} + \ldots + \beta_0 = 0$. Pour tout i entre 0 et $n-1$, l'élément β_i est entier sur A et *a fortiori* sur $A[\beta_0, \ldots, \beta_{i-1}]$. L'anneau $A[\beta_0, \ldots, \beta_i] = A[\beta_0, \ldots, \beta_{i-1}][\beta_i]$ est donc fini sur $A[\beta_0, \ldots, \beta_{i-1}]$. Par transitivité et par une récurrence immédiate, il s'ensuit que $A[\beta_0, \ldots, \beta_{n-1}]$ est fini sur A. Comme b est entier sur $A[\beta_0, \ldots, \beta_{n-1}]$, l'anneau $A[\beta_0, \ldots, \beta_{n-1}, b]$ est fini sur $A[\beta_0, \ldots, \beta_{n-1}]$, et est en conséquence fini sur A par transitivité. Ainsi b vit dans un sous-anneau de B fini sur A, et il appartient dès lors à C. $\square$

Définition 12.2.7. — Un *nombre algébrique* est un nombre complexe algébrique sur $\mathbf{Q}$; un *entier algébrique* est un nombre complexe entier sur $\mathbf{Z}$. On note $\overline{\mathbf{Q}}$ l'ensemble des nombres algébriques, et $\overline{\mathbf{Z}}$ celui des entiers algébriques. En vertu du théorème 12.2.6 et de la remarque 12.2.5, $\overline{\mathbf{Z}}$ est un sous-anneau de $\mathbf{C}$ (qui contient $\mathbf{Z}$, mais c'est le cas de tout sous-anneau de $\mathbf{C}$!) et $\overline{\mathbf{Q}}$ est un sous-corps de $\mathbf{C}$ (qui contient $\mathbf{Q}$, mais c'est le cas de tout sous-anneau de $\mathbf{C}$!).

Remarque 12.2.8. — Soit P un polynôme non constant à coefficients dans $\overline{\mathbf{Q}}$. Comme $\mathbf{C}$ est algébriquement clos, P a une racine z dans $\mathbf{C}$; par définition, z est algébrique sur $\overline{\mathbf{Q}}$ et il résulte alors du théorème 12.2.6 que z appartient à $\overline{\mathbf{Q}}$. Ainsi, $\overline{\mathbf{Q}}$ est algébriquement clos.

Exemples 12.2.9. — Soit n un entier $\geqslant 1$. Les nombres complexes $i, j = e^{2i\pi/3}$ ou encore $\sqrt[n]{2}$ sont des entiers algébriques, qui annulent respectivement les polynômes unitaires $X^2 + 1$, $X^2 + X + 1$ et $X^n - 2$. Le nombre complexe $1/\sqrt[n]{2}$ est un nombre algébrique, qui annule le polynôme $X^n - (1/2)$; mais ce n'est pas un entier algébrique, car sinon $1/2$ serait lui aussi un entier algébrique, ce qui est exclu par le lemme ci-dessous.

Lemme 12.2.10. — *Les sous-anneaux* $\overline{\mathbf{Z}}$ *et* $\overline{\mathbf{Q}}$ *de* $\mathbf{C}$ *sont stables sous la conjugaison complexe, et l'on a* $\overline{\mathbf{Z}} \cap \mathbf{Q} = \mathbf{Z}$.

Démonstration. — Soit a un entier algébrique (resp. un nombre algébrique). Par définition, il existe un polynôme unitaire à coefficients entiers (resp. rationnels) P tel que $P(a) = 0$. On a alors $P(\overline{a}) = \overline{P(a)} = 0$, et $\overline{a}$ est un entier algébrique (resp. un nombre algébrique) ; par conséquent, $\overline{\mathbf{Z}}$ et $\overline{\mathbf{Q}}$ sont stables sous la conjugaison complexe.

Montrons maintenant que $\overline{\mathbf{Z}} \cap \mathbf{Q} = \mathbf{Z}$. On sait que $\overline{\mathbf{Z}}$ contient $\mathbf{Z}$; il reste à s'assurer que $\overline{\mathbf{Z}} \cap \mathbf{Q} \subset \mathbf{Z}$. Soit a appartenant à $\overline{\mathbf{Z}} \cap \mathbf{Q}$. Écrivons $a = p/q$ où p et q sont deux éléments premiers entre eux de $\mathbf{Z}$, avec $q > 0$. Soit $P = X^n + a_{n-1}X^{n-1} + \ldots + a_0$ un polynôme unitaire à coefficients dans $\mathbf{Z}$ tel que $P(a) = 0$. On a

$$\frac{p^n}{q^n} + a_{n-1}\frac{p^{n-1}}{q^{n-1}} + \ldots + a_0 = 0.$$

En multipliant par q^n, il vient

$$p^n + a_{n-1}p^{n-1}q + \ldots + a_0q^n = 0.$$

Il en résulte que q divise p^n. Mais comme q est premier avec p, le lemme de Gauß assure que q est premier à p^n ; par conséquent, $q = 1$ et $a \in \mathbf{Z}$. □

12.3. Dimension d'une représentation irréductible. — Nous revenons maintenant aux représentations ($\mathbf{C}$-linéaires et de dimension finie) du groupe fini G, dont nous notons $\mathscr{C}$ l'ensemble des classes de conjugaison. Le but de ce qui suit est de démontrer que la dimension d'une telle représentation divise $|G|$.

12.3.1. *Intégralité des caractères.* — Soit (W, ρ) une représentation de G. Pour tout $g \in G$ on a $g^n = e$ et donc $\rho(g)^n = \mathrm{Id}_W$. Ainsi, $\rho(g)$ annule le polynôme $X^n - 1$. Ses valeurs propres sont donc des racines n-ièmes de l'unité, et sont en particulier entières ; il s'ensuit que $\chi_W(g) = \mathrm{Tr}\,(\rho(g))$ est un entier algébrique.

Théorème 12.3.2. — *La dimension de toute représentation irréductible de G divise $|G|$.*

Démonstration. — Soit n le cardinal de G et soit V une représentation irréductible de G ; on note χ le caractère de V, et m sa dimension. Soit $\mathscr{C}$ l'ensemble des classes de conjugaison de G. Nous allons montrer que m divise n, c'est-à-dire encore que le nombre rationnel n/m appartient à $\mathbf{Z}$; il suffit pour cela de montrer qu'il est entier sur $\mathbf{Z}$ (lemme 12.2.10).

Soit $c \in \mathscr{C}$. Notons f_c l'endomorphisme $v \mapsto \sum_{g \in c} gv$ de V. Pour tout $h \in G$, l'endomorphisme hf_c de V envoie un vecteur v sur

$$hf_c(h^{-1}v) = \sum_{g \in c_c} hgh^{-1}v = \sum_{g \in c} gv = f_c(v),$$

l'avant-dernière égalité provenant du fait que la conjugaison par h induit une bijection de c sur elle-même (puisque c est une classe de conjugaison). On a donc $hf_c = f_c$. Il s'ensuit que f_c est équivariant, et c'est donc une homothétie puisque V est irréductible (lemme 11.6.11, assertion (3)). La trace de f_c est par construction égale à $\sum_{g \in c} \chi(g)$, c'est-à-dire à $|c|\chi(c)$; ainsi, f_c est l'homothétie de rapport $|c|\chi(c)/m$. Soit $c' \in \mathscr{C}$. Le produit $f_c f_{c'}$ envoie un vecteur v sur $\sum_{g \in c} \sum_{h \in c'} ghv$.

Soit $d \in \mathscr{C}$ et soit γ un élément de d. Le cardinal de l'ensemble E_γ des couples $(g, h) \in c \times c'$ tels que $gh = \gamma$ est alors indépendant de γ. En effet, soit $\delta \in d$. Comme γ et δ appartiennent à la même classe de conjugaison, il existe $\omega \in G$ tel que $\omega\gamma\omega^{-1} = \delta$; on vérifie alors aussitôt que $(g, h) \mapsto (\omega g\omega^{-1}, \omega h\omega^{-1})$ induit une bijection de E_γ sur E_δ, de réciproque $(g, h) \mapsto (\omega^{-1}g\omega, \omega^{-1}h\omega)$. Notons $\mu_{cc'd}$ le cardinal commun de tous les E_γ pour $\gamma \in d$.

On a pour tout $v \in V$ les égalités

$$f_c f_{c'}(v) = \sum_{g \in c} \sum_{h \in c'} ghv = \sum_{d \in \mathscr{C}} \sum_{g \in d} \mu_{cc'd}\, gv = \sum_{d \in \mathscr{C}} \mu_{cc'd} f_d(v).$$

Autrement dit, $f_c f_{c'} = \sum_{d \in \mathscr{C}} \mu_{cc'd} f_d$. En passant aux rapports des homothéties en jeu, on obtient l'égalité

$$(|c|\chi(c)/m)(|c'|\chi(c')/m) = \sum_{d \in \mathscr{C}} \mu_{cc'd}(|d|\chi_V(d)/m).$$

Soit N le sous-groupe abélien de $\mathbf{C}$ engendré par $\mathbf{Z}$ et les éléments de la forme $|c|\chi(c)/m$ pour $c \in \mathscr{C}$. Il découle de la formule ci-dessus et du fait que les $\mu_{cc'd}$ sont des entiers (positifs, mais ça n'a pas d'importance ici) que N est stable par multiplication, et il contient 1 par construction. C'est en conséquence un sous-anneau de $\mathbf{C}$ fini sur $\mathbf{Z}$ (c'est-à-dire de type fini comme groupe abélien). Ses éléments sont donc des entiers algébriques ; en particulier, $|c|\chi(c)/m$ est un entier algébrique pour toute $c \in \mathscr{C}$.

L'égalité $\langle \chi, \chi \rangle = 1$ se récrit

$$\begin{aligned} & \frac{1}{n}\sum_{c\in\mathscr{C}} |c|\chi(c)\overline{\chi(c)} &= 1 \\ \iff \quad & \sum_{c\in\mathscr{C}} |c|\chi(c)\overline{\chi(c)} &= n \\ \iff \quad & \sum_{c\in\mathscr{C}} |c|\frac{\chi(c)}{m}\overline{\chi(c)} &= \frac{n}{m}\,. \end{aligned}$$

Or pour toute $c \in \mathscr{C}$, il résulte de 12.3.1 que $\chi(c)$ est un entier algébrique (il en va donc de même de $\overline{\chi(c)}$), et on a vu plus haut que $|c|\chi(c)/m$ est un entier algébrique. On en déduit au vu de la dernière égalité ci-dessus que n/m est un entier algébrique, ce qui termine la démonstration. □

12.4. Table de caractères. — On désigne toujours par G un groupe fini et par $\mathscr{C}$ l'ensemble des classes de conjugaison de G ; posons $n = |G|$ et $r = |\mathscr{C}|$. Grâce au corollaire 12.1.7, on peut considérer que comprendre la théorie des représentations de G revient plus ou moins à comprendre quels sont les caractères des représentations irréductibles de G, que nous appellerons plus brièvement *caractères irréductibles* de G. Il y a exactement r caractères irréductibles (prop. 12.1.8).

Dans ce qui suit, nous considérerons toujours les caractères comme des fonctions de $\mathscr{C}$ dans $\mathbf{C}$. Rappelons que le produit hermitien de deux telles fonctions est donné par la formule

$$\langle \chi, \psi \rangle = \frac{1}{n}\sum_{c\in\mathscr{C}} |c|\chi(c)\overline{\psi(c)}.$$

On a en particulier pour toute fonction centrale χ l'égalité

$$\langle \chi, \chi \rangle = \frac{1}{n}\sum_{c\in\mathscr{C}} |c| \cdot |\chi(c)|^2 \ ;$$

le carré hermitien de χ ne dépend donc que de la fonction $|\chi|$.

Nous utiliserons les notations $\mathbf{1}_G, \Sigma_n$ et V_n au lieu de $\mathbf{1}_{G,\mathbf{C}}$, $\Sigma_{n,\mathbf{C}}$ et $V_{n,\mathbf{C}}$ (la représentation signature $\Sigma_{n,\mathbf{C}}$ a été définie au 11.1.12 ; la représentation $V_{3,\mathbf{C}}$ a été définie au 11.3.3.3).

12.4.1. — On cherchera à construire ce qu'on appelle la *table des caractères* de G, qui se présente sous la forme d'un tableau carré ayant r lignes et r colonnes. Chaque colonne correspond à un élément de $\mathscr{C}$ (on commence en général par la classe singleton $\{e\}$), et chaque ligne à un caractère irréductible, c'est-à-dire à une classe d'isomorphie de représentation irréductible de G. Dans la case correspondant au caractère χ et à la classe c, on place le nombre complexe $\chi(c)$. Nous allons maintenant expliquer les différents outils dont on dispose pour construire cette table.

12.4.1.1. — On peut bien entendu chercher à exhiber directement des représentations irréductibles de G. On commence par faire la liste de celles qu'on connaît déjà : il y a toujours par exemple la représentation unité, et plus généralement toutes les représentations de dimension 1. Notons également que si H est un groupe, si V est une représentation irréductible de H et si $f\colon G \to H$ est un morphisme *surjectif* alors V est encore irréductible lorsqu'on voit V comme une représentation de G *via* f (11.6.7). On peut ensuite, à partir de ce premier stock de représentations, en fabriquer d'autres par les techniques standard vues en cours (*cf. infra*) ; mais il faut faire attention à deux points importants.

(a) On doit s'assurer que les nouvelles représentations ainsi obtenues sont bien irréductibles ; on le fait par exemple en s'assurant que la norme hermitienne de leur caractère est égale à 1 (corollaire 12.1.7).

(b) On doit éviter les redondances : deux représentations données par des descriptions *a priori* différentes peuvent en effet s'avérer isomorphes ; on s'en rend compte en constatant qu'elles ont même caractère.

Donnons deux exemples des «techniques standard» que nous avons évoquées.

- ◇ *Le passage à la contragrédiente.* Soit V une représentation irréductible de G. La contragrédiente V^* de V a pour caractère $\overline{\chi_V}$. Elle est irréductible : cela résulte de 11.6.6, mais on peut également le voir en remarquant que

$$1 = \langle \chi_V, \chi_V \rangle = \langle \overline{\chi_V}, \overline{\chi_V} \rangle$$

puisque χ_V et $\overline{\chi_V}$ ont même module.

Mais notez que si χ_V est à valeurs réelles, $\chi_{V^*} = \chi_V$; la représentation V^* est alors isomorphe à V, et l'on est face à un cas de redondance tel que décrit au (b) ci-dessus.

- ◇ *Le produit tensoriel.* Soient V_1 et V_2 deux représentations irréductibles de G. Le produit tensoriel $V_1 \otimes_k V_2$ a pour caractère $\chi_{V_1}\chi_{V_2}$. Le représentation $V_1 \otimes_k V_2$ est donc irréductible si et seulement si $\chi_{V_1}\chi_{V_2}$ est de norme hermitienne égale à 1. Ce n'est pas toujours le cas (nous verrons des contre-exemples plus bas) ; toutefois cela se produit par exemple dès que V_1 ou V_2 est de dimension 1. C'est en effet une conséquence directe de 11.6.5, mais nous allons donner une autre preuve à l'aide de la théorie des caractères. Supposons par exemple que V_1 est de dimension 1. Dans ce cas χ_{V_1} coïncide avec le morphisme de G vers $\mathbf{C}^\times$ qui définit V_1 à isomorphisme près (exemple 11.5.3). Comme G est fini, $\chi_1(G)$ est un sous-groupe fini de $\mathbf{C}^\times$; il est donc constitué d'éléments de torsion, c'est-à-dire de racines de l'unité qui sont en particulier de module 1. Il en résulte que $|\chi_{V_1}\chi_{V_2}| = |\chi_{V_2}|$; puisque la norme hermitienne de χ_{V_2} est égale à 1, il en va alors de même de celle de $\chi_{V_1}\chi_{V_2}$, et $V_1 \otimes_k V_2$ est bien irréductible.

12.4.1.2. — On peut aussi chercher à obtenir de nouveaux caractères irréductibles sans construire directement les représentations correspondantes. Pour cela, on peut exploiter différentes propriétés de la table de caractères que nous allons maintenant décrire. Pour en faciliter les énoncés et les preuves, fixons une numérotation $c_1, \dots, c_r$ des éléments de $\mathscr{C}$ (avec $c_1 = \{e\}$) et une numérotation $\chi_1, \dots, \chi_r$ des caractères irréductibles de G. Pour tout i on désigne par n_i le cardinal de c_i, par V_i «la» représentation irréductible de caractère χ_i, et par d_i la dimension de V_i. La table des caractères s'interprète alors comme la matrice carrée M de taille r dont le terme général χ_{ij} est égal à $\chi_i(c_j)$.

Indiquons d'emblée une première propriété de la matrice M : on a pour tout i l'égalité $\chi_{i1} = \chi_i(c_1) = \chi_i(e) = d_i$. La première colonne de M est donc constituée d'entiers ; ils divisent tous n (théorème 12.3.2), et la somme de leurs carrés est égale à n en vertu de l'assertion (5) du corollaire 12.1.7.

L'orthogonalité des caractères (théorème 12.1.6) signifie par ailleurs que les lignes de M forment une base orthonormée pour le produit hermitien

$$((\alpha_j)_j, (\beta_j)_j) \mapsto \frac{1}{n} \sum n_j \alpha_j \overline{\beta}_j$$

de $\mathbf{C}^r$. Autrement dit, on a pour tout (i, ℓ) l'égalité

$$\frac{1}{n} \sum_j n_j \chi_{ij} \overline{\chi_{\ell j}} = \delta_{i\ell}.$$

Si N désigne la matrice $\mathrm{Diag}(n_1, \dots, n_r) \in \mathrm{M}_n(\mathbf{C})$, cela peut se récrire

$$\frac{1}{n} MN {}^t\overline{M} = \mathrm{I}_r.$$

La matrice ${}^t\overline{M}$ est ainsi inversible d'inverse $\frac{1}{n}MN$, ce qui entraîne que

$$\frac{1}{n} {}^t\overline{M} MN = \mathrm{I}_r.$$

Pour tout (i, j), le terme d'indice (i, j) de $\frac{1}{n} {}^t\overline{M}MN$ est égal à

$$\frac{n_j}{n} \sum_\ell \overline{\chi_{\ell i}} \chi_{\ell j}.$$

Cette dernière expression vaut donc δ_{ij}, ce qui signifie que pour le produit hermitien *standard* de $\mathbf{C}^r$ (dont la définition ne fait plus intervenir les n_i) :

(a) les colonnes de M sont deux à deux orthogonales ;

(b) pour tout j, le carré de la norme de la j-ème colonne de M est égal à n/n_j.

Lorsqu'on applique (b) pour $j = 1$ (rappelons que $c_1 = \{e\}$ et donc que $n_1 = 1$) on retrouve l'égalité $n = \sum_i |\chi_{i1}|^2$, mentionnée plus haut.

12.4.2. *Caractères d'un groupe abélien.* — Supposons que le groupe fini G est abélien. Le lemme 11.8.1 assure que les représentations irréductibles de G sont exactement les représentations de G *de dimension 1*. Se donner une classe d'isomorphie d'une telle représentation revient à se donner un morphisme de G vers $\mathbf{C}^\times$, qui coïncide alors avec le caractère irréductible correspondant (11.5.3). L'ensemble des caractères irréductibles de G est donc égal à $\mathrm{Hom}(G, \mathbf{C}^\times)$.

Comme G est abélien, les classes de conjugaison de G sont les singletons $\{g\}$ avec $g \in G$. On déduit donc de la théorie générale que l'ensemble des caractères irréductibles de G a pour cardinal $|G|$. Il s'ensuit que $|\mathrm{Hom}(G, \mathbf{C}^\times)| = |G|$. Mais on peut démontrer cette égalité directement, sans faire appel à la théorie générale des représentations ; on peut même être beaucoup plus précis, comme en atteste le lemme suivant.

Lemme 12.4.3. — *Le groupe* $\mathrm{Hom}(G, \mathbf{C}^\times)$ *est isomorphe à G (non canoniquement en général) ; en particulier, son cardinal est égal à $|G|$*

Démonstration. — Notons G additivement. Le théorème 3.9.4 assure l'existence d'une (unique) famille finie $(d_1, \ldots, d_n)$ d'entiers > 1 telle que $d_1|d_2|\ldots|d_n$ et d'un isomorphisme

$$G \simeq \mathbf{Z}/d_1\mathbf{Z} \oplus \mathbf{Z}/d_2\mathbf{Z} \oplus \ldots \oplus \mathbf{Z}/d_n\mathbf{Z}.$$

Il est immédiat que pour toute famille de morphismes $(h_i \colon \mathbf{Z}/d_i\mathbf{Z} \to \mathbf{C}^\times)_{1 \leqslant i \leqslant r}$ il existe un et un seul morphisme $h \colon G \to \mathbf{C}^\times$ tel que $h|_{\mathbf{Z}/d_i\mathbf{Z}} = h_i$ pour tout i, à savoir $(x_1, \ldots, x_r) \mapsto \prod_i h_i(x_i)$; on vérifie aussitôt que $(h_1, \ldots, h_r) \mapsto h$ établit un isomorphisme entre $\bigoplus_i \mathrm{Hom}(\mathbf{Z}/d_i\mathbf{Z}, \mathbf{C}^\times)$ et $\mathrm{Hom}(G, \mathbf{C}^\times)$. Il suffit donc pour conclure de montrer que $\mathrm{Hom}(\mathbf{Z}/d\mathbf{Z}, \mathbf{C}^\times)$ est isomorphe à $\mathbf{Z}/d\mathbf{Z}$ pour tout $d \geqslant 1$.

Soit d un entier $\geqslant 1$. Le groupe $\mathrm{Hom}(\mathbf{Z}/d\mathbf{Z}, \mathbf{C}^\times)$ s'identifie au groupe des éléments de d-torsion de $\mathbf{C}^\times$, c'est-à-dire au groupe des racines d-ièmes de l'unité de $\mathbf{C}^\times$. Or ce dernier est cyclique et de cardinal d (il est engendré par $\exp(2i\pi/d)$), ce qui achève la démonstration. □

12.4.4. *Table des caractères de $\mathfrak{S}_3$; décomposition de quelques représentations.* — Le groupe $\mathfrak{S}_3$ a 3 classes de conjugaison : la classe $c_\varnothing$ de l'identité, qui est un singleton ; la classe $c_{(**)}$ des transpositions, qui est de cardinal 3 ; et la classe $c_{(***)}$ des 3-cycles, qui est de cardinal 2. Il s'ensuit que $\mathfrak{S}_3$ admet trois (classes d'isomorphie de) représentations irréductibles.

12.4.4.1. — Nous avons déjà rencontré deux représentations de $\mathfrak{S}_3$ de dimension 1 (et partant irréductibles), et non isomorphes : la représentation unité $\mathbf{1}_{\mathfrak{S}_3}$, et la représentation signature Σ_3. Le caractère $\chi_{\mathbf{1}_{\mathfrak{S}_3}}$ est trivial, et χ_{Σ_3} envoie $c_\varnothing$ et $c_{(***)}$ sur 1, et $c_{(**)}$ sur (-1). Nous avons démontré par ailleurs au 11.6.10.1 que la représentation V_3, qui est de dimension 2, est irréductible. Calculons son caractère à l'aide de 11.5.11.2. Comme $\mathrm{Id}_{\{1,2,3\}}$ a 3 points fixes sur $\{1, 2, 3\}$, on a $\chi_{V_3}(c_\varnothing) = 3 - 1 = 2$; une transposition de $\{1, 2, 3\}$ ayant 1 point fixe, on a $\chi_{V_3}(c_{(**)}) = 1 - 1 = 0$; et comme un 3-cycle de $\{1, 2, 3\}$ n'a aucun point fixe, on a $\chi_{V_3}(c_{(***)}) = 0 - 1 = -1$. Remarquons que

$$\langle \chi_{V_3}, \chi_{V_3} \rangle = \frac{1}{6}(1 \cdot 2^2 + 3 \cdot 0^2 + 2 \cdot (-1)^2) = 1 :$$

on voit que la théorie des caractères fournit une nouvelle preuve de l'irréductibilité de V_3.

12.4.4.2. *Conclusion.* — On a ainsi exhibé trois représentations irréductibles de $\mathfrak{S}_3$ deux à deux non isomorphes. Il n'y en a donc à isomorphisme près pas d'autres. La table des caractères de $\mathfrak{S}_3$ est par conséquent égale à

$$\begin{array}{cccc} & c_\varnothing & c_{(**)} & c_{(***)} \\ \mathbf{1}_{\mathfrak{S}_3} & 1 & 1 & 1 \\ \Sigma_3 & 1 & -1 & 1 \\ V_3 & 2 & 0 & -1 \end{array}.$$

12.4.4.3. — Montrons à titre d'exemple comment calculer la décomposition de $V_3^{\otimes 2} := V_3 \otimes_{\mathbf{C}} V_3$ en somme directe de représentations irréductibles de $\mathfrak{S}_3$. Le caractère $\chi_{V_3^{\otimes 2}}$ est égal à $(\chi_{V_3})^2$; il envoie donc la classe $c_\varnothing$ sur 4, la classe $c_{(**)}$ sur 0, et la classe $c_{(***)}$ sur 1. Pour déterminer la décomposition de $V_3 \otimes_{\mathbf{C}} V_3$, il suffit de calculer le produit scalaire de $\chi_{V_3^{\otimes 2}}$ avec chacun des trois caractères irréductibles.

$$\langle \chi_{V_3^{\otimes 2}}, \chi_{\mathbf{1}_{\mathfrak{S}_3}} \rangle = \frac{1}{6}(1 \cdot 4 \cdot 1 + 3 \cdot 0 \cdot 1 + 2 \cdot 1 \cdot 1) = 1.$$

$$\langle \chi_{V_3^{\otimes 2}}, \chi_{\Sigma_3} \rangle = \frac{1}{6}(1 \cdot 4 \cdot 1 + 3 \cdot 0 \cdot (-1) + 2 \cdot 1 \cdot 1) = 1.$$

$$\langle \chi_{V_3^{\otimes 2}}, \chi_{V_3} \rangle = \frac{1}{6}(1 \cdot 4 \cdot 2 + 3 \cdot 0 \cdot 0 + 2 \cdot 1 \cdot (-1)) = 1.$$

Ainsi, $V_3^{\otimes 2} \simeq \mathbf{1}_{\mathfrak{S}_3} \oplus \Sigma_3 \oplus V_3$.

12.4.4.4. — Étudions maintenant le produit tensoriel $\Sigma_3 \otimes_{\mathbf{C}} V_3$. Son caractère est égal à $\chi_{\Sigma_3}\chi_{V_3}$; il envoie donc la classe $c_\varnothing$ sur 2, la classe $c_{(**)}$ sur 0, et la classe $c_{(***)}$ sur -1. Autrement dit, $\chi_{V_3}\chi_{\Sigma_3} = \chi_{V_3}$. Par conséquent, $\Sigma_3 \otimes_{\mathbf{C}} V_3 \simeq V_3$.

12.4.5. *Table des caractères de $\mathfrak{S}_4$; décomposition de quelques représentations.* — Le groupe $\mathfrak{S}_4$ a 5 classes de conjugaison : la classe $c_\varnothing$ de l'identité, qui est un singleton ; la classe $c_{(**)}$ des transpositions, qui est de cardinal 6 ; la classe $c_{(**)(**)}$ des produits de deux transpositions, qui est de cardinal 3 ; la classe $c_{(***)}$ des 3-cycles, qui est de cardinal 8 ; et la classe $c_{(****)}$ des 4-cycles, qui est de cardinal 6. Par conséquent, $\mathfrak{S}_4$ possède cinq (classes d'isomorphie de) représentations irréductibles.

12.4.5.1. — On connaît déjà deux représentations irréductibles de $\mathfrak{S}_4$ qui sont de dimension 1 et non isomorphes : la représentation triviale $\mathbf{1}_{\mathfrak{S}_4}$ et la signature Σ_4. Le caractère $\chi_{\mathbf{1}_{\mathfrak{S}_4}}$ est trivial, et χ_{Σ_4} envoie $c_\varnothing, c_{(**)(**)}$ et $c_{(***)}$ sur 1, et $c_{(**)}$ et $c_{(****)}$ sur (-1).

Nous allons maintenant vérifier que la représentation V_4 (qui est de dimension 3) est irréductible. Son caractère a été calculé en 11.5.12 : il envoie la classe $c_\varnothing$ sur 3, la classe $c_{(**)}$ sur 1, la classe $c_{(**)(**)}$ sur (-1), la classe $c_{(***)}$ sur 0, et la classe $c_{(****)}$ sur (-1). Le carré de sa norme hermitienne est égal à

$$\frac{1}{24}(1 \cdot 9 + 6 \cdot 1 + 3 \cdot 1 + 8 \cdot 0 + 6 \cdot 1) = 1.$$

Par conséquent, V_4 est irréductible.

Considérons maintenant la représentation $\Sigma_4 \otimes_{\mathbf{C}} V_4$. Elle est de dimension 3, et son caractère est égal à $\chi_{\Sigma_4}\chi_{V_4}$. Il envoie dès lors la classe $c_\varnothing$ sur 3, la classe $c_{(**)}$ sur -1, la classe $c_{(**)(**)}$ sur (-1), la classe $c_{(***)}$ sur 0, et la classe $c_{(****)}$ sur 1. Il diffère de χ_{V_4} ; la représentation $\Sigma_4 \otimes_{\mathbf{C}} V_4$ n'est pas isomorphe à V_4, et est irréductible puisque Σ_4 est de dimension 1.

On a ainsi exhibé quatre représentations irréductibles (deux à deux non isomorphes) de $\mathfrak{S}_4$; il en manque une, disons W. Si δ désigne sa dimension on a $1^2+1^2+3^2+3^2+\delta^2 = 24$; il vient $\delta = 2$. On a $\chi_W(c_\varnothing) = \dim W = 2$; appelons a, b, c et d les valeurs de χ_W sur $c_{(**)}$, $c_{(**)(**)}$, $c_{(***)}$, et $c_{(****)}$. La table de caractères de $\mathfrak{S}_4$ se présente alors ainsi :

$$\begin{array}{cccccc} & c_\varnothing & c_{(**)} & c_{(**)(**)} & c_{(***)} & c_{(****)} \\ \mathbf{1}_{\mathfrak{S}_4} & 1 & 1 & 1 & 1 & 1 \\ \Sigma_4 & 1 & -1 & 1 & 1 & -1 \\ V_4 & 3 & 1 & -1 & 0 & -1 \\ \Sigma_4 \otimes_{\mathbf{C}} V_4 & 3 & -1 & -1 & 0 & 1 \\ W & 2 & a & b & c & d \end{array}.$$

Comme les colonnes du tableau sont orthogonales pour le produit hermitien usuel, il vient : $a = 0$, $b = 2$, $c = -1$, $d = 0$. Ainsi, la table des caractères de $\mathfrak{S}_4$ est finalement

$$\begin{array}{cccccc} & c_\varnothing & c_{(**)} & c_{(**)(**)} & c_{(***)} & c_{(****)} \\ \mathbf{1}_{\mathfrak{S}_4} & 1 & 1 & 1 & 1 & 1 \\ \Sigma_4 & 1 & -1 & 1 & 1 & -1 \\ V_4 & 3 & 1 & -1 & 0 & -1 \\ \Sigma_4 \otimes_{\mathbf{C}} V_4 & 3 & -1 & -1 & 0 & 1 \\ W & 2 & 0 & 2 & -1 & 0 \end{array}.$$

12.4.5.2. — La théorie assure donc l'existence d'une représentation W de $\mathfrak{S}_4$ irréductible, de degré 2, ayant le caractère indiqué ci-dessus. On peut en donner une description un peu plus tangible, comme suit. On a construit à l'exemple 7.3.2 un morphisme surjectif $\varphi : \mathfrak{S}_4 \to \mathfrak{S}_3$ dont le noyau est le «groupe de Klein» $\{\mathrm{Id}, (12)(34), (13)(24), (14)(23)\}$. Comme $\mathfrak{S}_4$ est de cardinal 24, comme $\mathfrak{S}_3$ est de cardinal 6, et comme le groupe de Klein est de cardinal 4, l'égalité $24/4 = 6$ assure que φ est surjectif.

Dès lors, la représentation V_3 de $\mathfrak{S}_3$, qui est irréductible et de dimension 2, induit *via* φ une représentation de $\mathfrak{S}_4$, qui est également irréductible et de dimension 2. Elle coïncide donc *nécessairement* avec W – mais le lecteur soupçonneux est invité à s'assurer que son caractère $\chi_{V_3} \circ \varphi$ est bien égal à χ_W.

12.4.5.3. — Nous allons à titre d'exemple calculer la décomposition de $V_4^{\otimes 2}$ comme somme directe de représentations irréductibles de $\mathfrak{S}_4$. Son caractère est égal à $\chi_{V_4}^2$. Il envoie donc la classe $c_\varnothing$ sur 9, la classe $c_{(**)}$ sur 1, la classe $c_{(**)(**)}$ sur 1, la classe $c_{(***)}$ sur 0, et la classe $c_{(****)}$ sur 1. Nous allons maintenant calculer son produit hermitien avec les différents caractères irréductibles de $\mathfrak{S}_4$.

$$\langle \chi_{V_4^{\otimes 2}}, \chi_{\mathbf{1}_{\mathfrak{S}_4}} \rangle = \frac{1}{24}(1 \cdot 9 \cdot 1 + 6 \cdot 1 \cdot 1 + 3 \cdot 1 \cdot 1 + 8 \cdot 0 \cdot 1 + 6 \cdot 1 \cdot 1) = 1.$$

$$\langle \chi_{V_4^{\otimes 2}}, \chi_{\Sigma_4} \rangle = \frac{1}{24}(1 \cdot 9 \cdot 1 + 6 \cdot 1 \cdot (-1) + 3 \cdot 1 \cdot 1 + 8 \cdot 0 \cdot 1 + 6 \cdot 1 \cdot (-1)) = 0.$$

$$\langle \chi_{V_4^{\otimes 2}}, \chi_{V_4} \rangle = \frac{1}{24}(1 \cdot 9 \cdot 3 + 6 \cdot 1 \cdot 1 + 3 \cdot 1 \cdot (-1) + 8 \cdot 0 \cdot 0 + 6 \cdot 1 \cdot (-1)) = 1.$$

$$\langle \chi_{V_4^{\otimes 2}}, \chi_{\Sigma_4 \otimes_{\mathbf{C}} V_4} \rangle = \frac{1}{24}(1 \cdot 9 \cdot 3 + 6 \cdot 1 \cdot (-1) + 3 \cdot 1 \cdot (-1) + 8 \cdot 0 \cdot 0 + 6 \cdot 1 \cdot 1) = 1.$$

$$\langle \chi_{V_4^{\otimes 2}}, \chi_W \rangle = \frac{1}{24}(1 \cdot 9 \cdot 2 + 6 \cdot 1 \cdot 0 + 3 \cdot 1 \cdot 2 + 8 \cdot 0 \cdot (-1) + 6 \cdot 1 \cdot 0) = 1.$$

Il vient $V_4^{\otimes 2} \simeq \mathbf{1}_{\mathfrak{S}_4} \oplus V_4 \oplus (\Sigma_4 \otimes_{\mathbf{C}} V_4) \oplus W$.

12.4.5.4. — Nous allons calculer la décomposition de $V_4 \otimes_{\mathbf{C}} W$ comme somme directe de représentations irréductibles de $\mathfrak{S}_4$. Son caractère est égal à $\chi_{V_4}\chi_W$. Il envoie donc la classe $c_\varnothing$ sur 6, la classe $c_{(**)}$ sur 0, la classe $c_{(**)(**)}$ sur (-2), la classe $c_{(***)}$ sur 0, et la classe $c_{(****)}$ sur 0. Nous allons maintenant calculer son produit hermitien avec les différents caractères irréductibles de $\mathfrak{S}_4$.

$$\langle \chi_{V_4 \otimes_{\mathbf{C}} W}, \chi_{\mathbf{1}_{\mathfrak{S}_4}} \rangle = \frac{1}{24}(1 \cdot 6 \cdot 1 + 6 \cdot 0 \cdot 1 + 3 \cdot (-2) \cdot 1 + 8 \cdot 0 \cdot 1 + 6 \cdot 0 \cdot 1) = 0.$$

$$\langle \chi_{V_4 \otimes_{\mathbf{C}} W}, \chi_{\Sigma_4} \rangle = \frac{1}{24}(1 \cdot 6 \cdot 1 + 6 \cdot 0 \cdot (-1) + 3 \cdot (-2) \cdot 1 + 8 \cdot 0 \cdot 1 + 6 \cdot 0 \cdot (-1)) = 0.$$

$$\langle \chi_{V_4 \otimes_{\mathbf{C}} W}, \chi_{V_4} \rangle = \frac{1}{24}(1 \cdot 6 \cdot 3 + 6 \cdot 0 \cdot 1 + 3 \cdot (-2) \cdot (-1) + 8 \cdot 0 \cdot 0 + 6 \cdot 0 \cdot (-1)) = 1.$$

$$\langle \chi_{V_4 \otimes_{\mathbf{C}} W}, \chi_{\Sigma_4 \otimes_{\mathbf{C}} V_4} \rangle = \frac{1}{24}(1 \cdot 6 \cdot 3 + 6 \cdot 0 \cdot (-1) + 3 \cdot (-2) \cdot (-1) + 8 \cdot 0 \cdot 0 + 6 \cdot 0 \cdot 1) = 1.$$

$$\langle \chi_{V_4 \otimes_{\mathbf{C}} W}, \chi_W \rangle = \frac{1}{24}(1 \cdot 6 \cdot 2 + 6 \cdot 0 \cdot 0 + 3 \cdot (-2) \cdot 2 + 8 \cdot 0 \cdot (-1) + 6 \cdot 0 \cdot 0) = 0.$$

Il vient $V_4 \otimes_{\mathbf{C}} W \simeq V_4 \oplus (\Sigma_4 \otimes_{\mathbf{C}} V_4)$.

13. Généralités sur les formes quadratiques

Nous allons maintenant aborder la dernière partie de cours. Son but est l'introduction d'une nouvelle notion, celle de *forme quadratique*, et l'étude de certains groupes matriciels auxquels elle donne lieu. Ce chapitre sera consacré à un certain nombre de résultats généraux ; un certain nombre de questions plus spécifiques seront abordées dans le chapitre suivant (qui est aussi le dernier).

13.1. Définitions générales. — L'essentiel de ce chapitre concernera des espaces vectoriels de dimension finie sur un corps de caractéristique différente de 2, mais les définitions de bases peuvent s'énoncer sans ces hypothèses restrictives. On fixe un corps k, pour le moment quelconque.

Définition 13.1.1. — Soit V un k-espace vectoriel. Une *forme bilinéaire* sur V est une application bilinéaire b de $V \times V$ vers k. On dit que b est *symétrique* (resp. *antisymétrique*) si $b(v,w) = b(w,v)$ (resp. $b(v,w) = -b(w,v)$) pour tout (v,w).

13.1.2. — Soit V un k-espace vectoriel. L'ensemble $\mathrm{Bil}(V)$ des formes bilinéaires sur V a une structure naturelle de k-espace vectoriel : si b et b' sont deux telles formes et si λ est un scalaire, on a

$$(b+b')(v,w) = b(v,w) + b'(v,w) \text{ et } (\lambda b)(v,w) = \lambda b(v,w)$$

pour tout $(v,w) \in V^2$. L'ensemble des formes bilinéaires symétriques (resp. antisymétriques) sur V en est un sous-espace vectoriel. L'espace des formes bilinéaires symétriques sur V sera noté $\mathrm{Sym}(V)$ (nous ne nous servirons guère des formes antisymétriques dans ce cours et n'avons donc pas besoin d'introduire une notation spécifique pour l'espace vectoriel de ces dernières).

13.1.3. *Aspect calculatoire.* — Soit V un k-espace vectoriel et soit $(e_i)_{i\in I}$ une base de V.

L'application linéaire de $\mathrm{Bil}(V)$ dans $k^{I\times I}$ qui envoie une forme b sur la famille $b(e_i,e_j)_{i,j}$ est un isomorphisme. Sa réciproque envoie une famille (a_{ij}) sur la forme bilinéaire

$$\left(\sum x_i e_i, \sum y_i e_i\right) \mapsto \sum_{i,j} a_{ij} x_i y_j.$$

Soit b une forme bilinéaire sur V. Pour tout (i,j), posons $a_{ij} = b(e_i,e_j)$. La forme b est symétrique (resp. antisymétrique) si et seulement si $a_{ij} = a_{ji}$ (resp. $a_{ij} = -a_{ji}$) pour tout (i,j). C'est en effet clairement nécessaire, et suffisant d'après la formule ci-dessus qui reconstitue b à partir des a_{ij}.

13.1.4. *Liens avec la dualité.* — Soit V un k-espace vectoriel.

13.1.4.1. — Si b est une forme bilinéaire sur V, la formule $v \mapsto (w \mapsto b(v,w))$ définit une application k-linéaire b_{g} de V vers V^*. L'application $b \mapsto b_{\mathrm{g}}$ établit une bijection k-linéaire entre l'espace des formes bilinéaires sur V et celui des applications k-linéaires de V vers V^*, de réciproque $\ell \mapsto ((v,w) \mapsto \ell(v)(w))$.

On dispose d'une construction analogue à celle décrite ci-dessus en «échangeant la gauche et la droite». Plus précisément, si b est une forme bilinéaire sur V, la formule

$v \mapsto (w \mapsto b(w, v))$ définit une application k-linéaire b_{d} de V vers V^*. L'application $b \mapsto b_{\mathrm{d}}$ établit une bijection k-linéaire entre l'espace des formes bilinéaires sur V et celui des applications k-linéaires de V vers V^*, de réciproque $\ell \mapsto ((v, w) \mapsto \ell(w)(v))$.

Il est clair que b est symétrique (resp. antisymétrique) si et seulement si $b_{\mathrm{g}} = b_{\mathrm{d}}$ (resp. $b_{\mathrm{g}} = -b_{\mathrm{d}}$).

13.1.4.2. — Soit b une forme bilinéaire sur V. La transposée ${}^{t}b_{\mathrm{g}}$ est une application k-linéaire de V^{**} vers V^*. Si w est un vecteur de V on a les égalités

$$\begin{aligned} {}^{t}b_{\mathrm{g}}(\theta_V(w)) &= \theta_V(w) \circ b_{\mathrm{g}} \\ &= v \mapsto \theta_V(w)(b_{\mathrm{g}}(v)) \\ &= v \mapsto b_{\mathrm{g}}(v)(w) \\ &= v \mapsto b(v, w) \\ &= b_{\mathrm{d}}(w). \end{aligned}$$

On montrerait de même que ${}^{t}b_{\mathrm{d}}(\theta_V(w)) = b_{\mathrm{g}}(w)$.

Supposons que V soit de dimension finie. Dans ce cas θ_V est un isomorphisme, et *modulo l'identification induite entre V et V^{**}* ce qui précède peut simplement se récrire ${}^{t}b_{\mathrm{g}} = b_{\mathrm{d}}$ et ${}^{t}b_{\mathrm{d}} = b_{\mathrm{g}}$. Ainsi, b est symétrique (resp. antisymétrique) si et seulement si ${}^{t}b_{\mathrm{g}} = b_{\mathrm{g}}$ (resp. ${}^{t}b_{\mathrm{g}} = -b_{\mathrm{g}}$), ou encore si et seulement si ${}^{t}b_{\mathrm{d}} = b_{\mathrm{d}}$ (resp. ${}^{t}b_{\mathrm{d}} = -b_{\mathrm{d}}$).

Définition 13.1.5. — Soit V un k-espace vectoriel. Une application $q\colon V \to k$ est appelée une *forme quadratique* sur V s'il existe une forme linéaire b sur V telle que $q(v) = b(v, v)$ pour tout $v \in V$. Nous dirons que q est la forme quadratique *définie par b*.

Remarque 13.1.6. — Soit q une forme quadratique sur V, définie par une forme bilinénaire b. Il est immédiat que $q(\lambda v) = \lambda^2 q(v)$ pour tout $v \in V$ et tout $\lambda \in k$, et que $q(v + w) = q(v) + q(w) + b(v, w) + b(w, v)$ pour tout $(v, w) \in V^2$.

13.1.7. — Soit V un k-espace vectoriel. L'ensemble $\mathrm{Quad}(V)$ des formes quadratiques sur V a une structure naturelle de k-espace vectoriel : si q et q' sont deux telles formes et si λ est un scalaire, on a

$$(q + q')(v) = q(v) + q'(v) \ \text{ et } \ (\lambda q)(v) = \lambda q(v)$$

pour tout $v \in V$.

13.1.8. *Aspect calculatoire.* — Soit V un k-espace vectoriel et soit $(e_i)_{i \in I}$ une base de V.

13.1.8.1. — Soit q une forme quadratique sur V. Il existe une unique famille $((A_i)_i, (A_{\{i,j\}})_{\{i,j\}})$ de scalaires paramétrée d'une part par I et d'autre part par l'ensemble des paires d'éléments de I (par paire, nous entendons partie de cardinal exactement 2) telle que

$$q\left(\sum x_i e_i\right) = \sum_i A_i x_i^2 + \sum_{\{i,j\}} A_{\{i,j\}} x_i x_j$$

pour toute famille $(x_i)_{i \in I}$ de scalaires presque tous nuls.

Justifions cette assertion. Remarquons tout d'abord que si $((A_i)_{i\in I}, (A_{\{i,j\}})_{\{i,j\}})$ est une famille comme ci-dessus on a alors nécessairement $A_i = q(e_i)$ pour tout i et

$$A_{\{i,j\}} = q(e_i + e_j) - A_i - A_j$$

pour toute paire $\{i, j\}$, ce qui assure l'unicité de la famille $((A_i), (A_{\{i,j\}}))$.

Montrons maintenant son existence. Choisissons une forme bilinéaire b qui définit q, et posons $a_{ij} = b(e_i, e_j)$ pour tout i. On a alors pour toute famille (x_i) de scalaires presque tous nuls les égalités

$$\begin{aligned} q\left(\sum x_i e_i\right) &= b(\sum x_i e_i, \sum x_i e_i) \\ &= \sum_{i,j} a_{ij} x_i x_j \\ &= \sum_i a_{ii} x_i^2 + \sum_{\{i,j\}} (a_{ij} + a_{ji}) x_i x_j. \end{aligned}$$

On voit donc que si l'on pose $A_i = a_{ii}$ pour tout i et $A_{\{i,j\}} = a_{ij} + a_{ji}$ pour toute paire $\{i, j\}$, la famille $((A_i), (A_{\{i,j\}}))$ satisfait les conditions requises.

13.1.8.2. — Réciproquement, donnons-nous une famille de scalaires $((A_i)_i, (A_{\{i,j\}})_{\{i,j\}})$. L'application

$$q \colon V \to k, \ \sum x_i e_i \mapsto \sum_i A_i x_i^2 + \sum_{\{i,j\}} A_{\{i,j\}} x_i x_j$$

est alors une forme quadratique sur V.

Pour le voir, il s'agit de montrer l'existence d'une forme bilinénaire b sur V qui définit q. Or la base (e_i) étant fixée, se donner une forme bilinéaire b sur V revient à se donner une famille (a_{ij}) de scalaires (à savoir la famille $(b(e_i, e_j))$, *cf.* 13.1.3). Et si l'on se donne une famille (a_{ij}) de scalaires, il résulte de 13.1.8.1 que la forme bilinéaire associée définit q si et seulement si $a_{ii} = A_i$ pour tout i et $a_{ij} + a_{ji} = A_{\{i,j\}}$ pour toute paire $\{i, j\}$.

Mais il est immédiat qu'il existe une famille (a_{ij}) satisfaisant ces conditions : il suffit de poser $a_{ii} = A_i$ pour tout i et, pour toute paire $\{i, j\}$, de prendre a_{ij} quelconque et a_{ji} égal à $A_{\{i,j\}} - a_{ij}$.

Notons par contre qu'on n'a pas unicité de la famille (a_{ij}) (c'est-à-dire de la forme b définissant q), du moins dès que l'ensemble I contient au moins une paire $\{i, j\}$, c'est-à-dire dès que $|I| \geqslant 2$: en effet dans ce cas on voit que a_{ij} peut être choisi arbitrairement.

Toutefois l'ambiguïté disparaît si la caractéristique de k est différente de 2 et si l'on impose en plus à b d'être symétrique, comme en atteste le lemme qui suit.

Lemme 13.1.9. — *Supposons que la caractéristique de k est différente de* 2*, soit V un k-espace vectoriel et soit q une forme quadratique sur V. Il existe une unique forme bilinéaire* symétrique *b sur V qui définit q ; elle est donnée par la formule*

$$b(v, w) = \frac{q(v + w) - q(v) - q(w)}{2}.$$

Démonstration. — Commençons par l'unicité. Soit b une forme bilinéaire symétrique définissant q. On a alors pour tout $(v,w) \in V^2$ les égalités

$$\begin{aligned} q(v+w) &= b(v+w, v+w) \\ &= b(v,v) + 2b(v,w) + b(w,w) \\ &= q(v) + 2(b,v,w) + q(w). \end{aligned}$$

Ainsi, on a nécessairement $b(v,w) = \frac{1}{2}(q(v+w) - q(v) - q(w))$ ce qui montre à la fois que b est unique et qu'elle est donnée par la formule requise.

Montrons maintenant l'existence de b. Choisissons une forme bilinéaire arbitraire β qui définit q. L'application

$$b\colon (v,w) \mapsto \frac{\beta(v,w) + \beta(w,v)}{2}$$

est une forme bilinéaire sur V, symétrique par construction. Et l'on a pour tout $v \in V$ l'égalité

$$b(v,v) = \frac{\beta(v,v) + \beta(v,v)}{2} = \beta(v,v) = q(v).$$

Ainsi, b définit q. □

Remarque 13.1.10. — Conservons les notations du lemme 13.1.9. Soit (e_i) une base de V et soit $((A_i), (A_{\{i,j\}})$ la famille de scalaires telle que q soit donnée par la formule

$$\sum x_i e_i \mapsto \sum_i A_i x_i^2 + \sum_{\{i,j\}} A_{\{i,j\}} x_i x_j \ ;$$

pour tout (i,j), posons $(a_{ij}) = b(e_i, e_j)$. On a alors $a_{ii} = a_i$ pour tout indice i et $a_{ij} = a_{ji} = A_{\{i,j\}}/2$ pour toute paire $\{i,j\}$ (cela découle par exemple de l'expression de b à partir de q fournie par le lemme ; on peut aussi remarquer directement que la forme bilinéaire associée à la famille (a_{ij}) ci-dessus est symétrique, et définit q en vertu de 13.1.8.1).

Remarque 13.1.11. — Le lemme 13.1.9 peut se reformuler comme suit : si k est de caractéristique différente de 2 et si V est un k-espace vectoriel, $b \mapsto (v \mapsto b(v,v))$ établit une bijection linéaire entre $\mathrm{Sym}(V)$ et $\mathrm{Quad}(V)$, dont la réciproque est donnée par la formule $q \mapsto ((v,w) \mapsto \frac{q(v+w)-q(v)-q(w)}{2})$.

13.2. Formes bilinéaires symétriques en dimension finie. — Après ces généralités, nous allons nous concentrer sur le cas de la dimension finie. Le corps k est encore (pour un moment) supposé quelconque.

13.2.1. *Matrice d'une forme bilinéaire dans une base.* — Soit V un k-espace vectoriel de dimension finie et soit $\mathscr{B} = (e_i)_{i \in I}$ une base de V. Soit b une forme bilinéaire sur V ; pour tout (i,j), posons $a_{ij} = b(e_i, e_j)$. La matrice $M := (a_{ij})$ est appelée *matrice de b dans la base $\mathscr{B}$*. La forme bilinéaire b est symétrique (resp. antisymétrique) si et seulement si c'est le cas de M.

13.2.1.1. *Expression matricielle de b.* — Soient v et w deux vecteurs de V. Écrivons $v = \sum x_i e_i$ et $w = \sum y_i e_i$, et notons X (resp. Y) la famille (x_i) (resp. (y_i)) vue comme vecteur colonne. On a alors

$$b(v, w) = \sum_{ij} a_{ij} x_i y_j = {}^tXMY.$$

13.2.1.2. *Interprétation conceptuelle de M.* — Soit b_{d} l'application linéaire de V vers V^* associée à b et définie en 13.1.4.1. En identifiant V^{**} à V on a pour tout indice j l'égalité

$$e_i^{**}(b_{\mathrm{d}}(e_j)) = b_{\mathrm{d}}(e_j)(e_i) = b(e_i, e_j) = a_{ij}.$$

Par conséquent, M est la matrice de l'application b_{d} dans les bases (e_i) et (e_i^*).

13.2.1.3. *Formules de changement de base.* — Soit $\mathscr{C}$ une autre base de V, et soit P la matrice de passage de $\mathscr{B}$ à $\mathscr{C}$. La matrice P peut s'interpréter comme la matrice de l'identité dans les bases $\mathscr{C}$ (au départ) et $\mathscr{B}$ (à l'arrivée). Par conséquent, ${}^tP$ s'interprète comme la matrice de l'identité dans les bases duales $\mathscr{B}^*$ et $\mathscr{C}^*$.

Soit N la matrice de b dans la base $\mathscr{C}$. Comme M (resp. N) s'interprète comme la matrice de b_{d} dans les bases $\mathscr{B}$ et $\mathscr{B}^*$ (resp. $\mathscr{C}$ et $\mathscr{C}^*$), il vient

$$N = {}^tPMP.$$

13.2.2. *Rang d'une fome bilinéaire.* — Soit b une forme bilinéaire sur un k-espace vectoriel V de dimension finie. Les applications b_{g} et b_{d} de V vers V^* sont transposées l'une de l'autre modulo l'identification canonique entre V et V^{**} (13.1.4 *et sq.*) ; elles ont donc même rang, et ce rang commun est appelé le *rang de la forme bilinéaire b*.

Si $\mathscr{B}$ est une base de V et si M désigne la matrice de b dans $\mathscr{B}$, le rang de b est égal au rang de M : c'est une conséquence de 13.2.1.2.

On dit que b est *non dégénérée* si son rang est égal à $\dim V$. C'est le cas si et seulement si b_{g} est bijective, ou encore si et seulement si b_{d} est bijective. Cela revient également à demander que M soit inversible.

13.2.3. *Noyau d'une forme bilinéaire symétrique.* — Soit b une forme bilinéaire *symétrique* sur un k-espace vectoriel V de dimension finie.

13.2.3.1. — Par symétrie de b, les applications b_{g} et b_{d} coïncident, et leur noyau commun est appelé le *noyau de la forme bilinéaire b* ; il est noté $\mathrm{Ker}(b)$. La forme b est non dégénérée si et seulement si $\mathrm{Ker}(b) = \{0\}$.

Concrètement, un vecteur v de V appartient au noyau de b si et seulement si $b(v, w) = 0$ pour tout $w \in V$, ou encore si et seulement si $b(w, v) = 0$ pour tout $v \in V$.

Si $\mathscr{B}$ est une base de V et si M désigne la matrice de b dans $\mathscr{B}$, un vecteur v de V appartient à $\mathrm{Ker}(b)$ si et seulement si le vecteur colonne C de ses coordonnées dans la base $\mathscr{B}$ satisfait l'égalité $MC = 0$: c'est une simpe conséquence du fait que M est la matrice de l'application b_{d} dans les bases $\mathscr{B}$ et $\mathscr{B}^*$.

13.2.3.2. — Posons $W = \mathrm{Ker}(b)$. Si v_1 et v_2 sont deux éléments de V, et si w_1 et w_2 sont deux éléments de W, il est immédiat que $b(v_1 + w_1, v_2 + w_2)) = b(v_1, v_2)$. Il s'ensuit que b induit par passage au quotient une forme bilinéaire symétrique $\overline{b}$ sur V/W, telle que $\overline{b}(\overline{v_1}, \overline{v_2}) = b(v_1, v_2)$ pour tout couple (v_1, v_2) d'éléments de V. Si v

appartient à V, sa classe $\overline{v}$ appartient au noyau de $\overline{b}$ si et seulement si $b(v,v') = 0$ pour tout $v' \in V$, c'est-à-dire si et seulement si $v \in W$, ou encore si et seulement si $\overline{v} = 0$. Par conséquent, $\overline{b}$ *est non dégénérée.*

13.2.4. *Orthogonalité.* — Soit V un k-espace vectoriel de dimension finie et soit b une forme bilinéaire symétrique sur V. On dit que deux vecteurs v et w de V sont *orthogonaux* (relativement à b) si $b(v,w) = 0$; on dit que deux sous-espaces vectoriels W et W' de V sont orthogonaux si tout vecteur de W est orthogonal à tout vecteur de W'. On appelle orthogonal d'un sous-espace vectoriel W de V, et l'on note $W^\perp$, le sous-espace vectoriel de V formé des vecteurs w qui sont orthogonaux à tout vecteur de W.

Remarquons que par définition, v appartient à l'orthogonal de W relativement à b si et seulement si $b_{\mathrm{g}}(v)$ appartient l'orthogonal de W au sens de la dualité.

Remarquons aussi que $V^\perp$ est par définition égal au noyau de b et plus généralement que $\mathrm{Ker}(b|_W) = W \cap W^\perp$.

Il est clair que $W \subset (W^\perp)^\perp$, et que si U est un sous-espace vectoriel de V tel que $U \subset W$ alors $U^\perp \supset W^\perp$.

Proposition 13.2.5. — *Soit V un k-espace vectoriel de dimension finie et soit b une forme bilinéaire symétrique sur V. Soit W un sous-espace vectoriel de V.*

(1) *On a l'égalité* $\dim W + \dim W^\perp = \dim V + \dim \mathrm{Ker}(b) \cap W$.

(2) *Si b est non dégénérée,* $\dim W + \dim W^\perp = \dim V$ *et* $(W^\perp)^\perp = W$.

(3) *On a $V = W \oplus W^\perp$ si et seulement si $b|_W$ est non dégénérée, c'est-à-dire si et seulement si $W \cap W^\perp = \{0\}$.*

Démonstration. — Commençons par (1). Supposons tout d'abord que b est non dégénérée, et notons W' l'orthogonal de W au sens de la dualité. On a $W^\perp = b_{\mathrm{g}}^{-1}(W')$, et comme b est non dégénérée b_{g} est bijective et il vient

$$\dim W^\perp = \dim W' = \dim V - \dim W,$$

ce qui montre (1) dans ce cas particulier.

Traitons maintenant le cas général. La forme b induit une forme bilinéaire symétrique non dégénérée $\overline{b}$ sur $V/\mathrm{Ker}(b)$ (13.2.3.2). L'image de W dans le quotient $V/\mathrm{Ker}(b)$ s'identifie à $W/(\mathrm{Ker}(b) \cap W)$; comme $\mathrm{Ker}(b) = V^\perp \subset W^\perp$, l'image de $W^\perp$ dans $V/\mathrm{Ker}(b)$ s'identifie à $W^\perp/\mathrm{Ker}(b)$. Soit $v \in V$. Son image $\overline{v}$ dans $V/\mathrm{Ker}(b)$ appartient à l'orthogonal de $W/(\mathrm{Ker}(b)/ \cap W) = \{\overline{w}\}_{w\in W}$ si et seulement si on a $\overline{b}(\overline{v},\overline{w}) = 0$ pour tout w appartenant à W, c'est-à-dire encore si et seulement si $b(v,w) = 0$ pour tout $w \in W$, ce qui revient à demander que $v \in W^\perp$. Par conséquent,

$$(W/(\mathrm{Ker}(b) \cap W))^\perp = \{\overline{v}\}_{v\in W^\perp} = W^\perp/\mathrm{Ker}(b).$$

On a alors les égalités

$$\begin{aligned}
\dim W^\perp - \dim \mathrm{Ker}(b) &= \dim W^\perp/\mathrm{Ker}(b) \\
&= \dim(W/(\mathrm{Ker}(b) \cap W)^\perp \\
&= \dim V/\mathrm{Ker}(b) - \dim W/(\mathrm{Ker}(b) \cap W) \\
&= \dim V - \dim \mathrm{Ker}(b) - \dim W + \dim(\mathrm{Ker}(b) \cap W),
\end{aligned}$$

où la troisième égalité découle du cas non dégénéré déjà traité. Il vient

$$\dim W + \dim W^{\perp} = \dim V + \dim \mathrm{Ker}(b) \cap W,$$

ce qui termine de montrer (1).

Montrons (2). Si b est non dégénérée on a $\mathrm{Ker}(b) = \{0\}$ et (1) entraîne donc que $\dim W + \dim W^{\perp} = \dim V$. En appliquant cette égalité à l'espace $W^{\perp}$ on voit que $\dim (W^{\perp})^{\perp} + \dim W^{\perp} = \dim V$, et partant que $\dim (W^{\perp})^{\perp} = \dim W$. Puisque $W \subset (W^{\perp})^{\perp}$ on a finalement $(W^{\perp})^{\perp} = W$.

Montrons enfin (3). Si $V = W \oplus W^{\perp}$ on a évidemment $W \cap W^{\perp} = \{0\}$. Réciproquement, supposons que $W \cap W^{\perp} = \mathrm{Ker}(b|_W) = \{0\}$. Puisque $\mathrm{Ker}(b) \subset W^{\perp}$ on a *a fortiori* $\mathrm{Ker}(b) \cap W = \{0\}$. En vertu de (1) ceci entraîne que $\dim W + \dim W^{\perp}$ est égal à $\dim V$. Puisque $W \cap W^{\perp} = \{0\}$ il en résulte que $V = W \oplus W^{\perp}$. □

Commentaires 13.2.6. — Soit V un k-espace vectoriel de dimension finie n, et soit b une forme bilinéaire sur V. Si W est un sous-espace vectoriel de V, l'assertion (1) de la proposition 13.2.5 montre que $\dim W + \dim W^{\perp} = n + \dim \mathrm{Ker}(b) \cap W$. En particulier $\dim W + \dim W^{\perp} \geqslant n$, et l'inégalité peut être stricte : par exemple dans le cas extrême où b est la forme bilinéaire nulle, on a $W^{\perp} = V$ et donc $\dim W + \dim W^{\perp} = n + \dim W$, qui est strictement supérieur à n dès que W est non nul.

Si l'on suppose que la forme b est non dégénérée on a $\mathrm{Ker}(b) = \{0\}$, si bien que $\dim W + \dim W^{\perp} = n$. Mais cela ne signifie pas pour autant que $V = W \oplus W^{\perp}$.

En fait, il résulte de l'assertion (3) de la proposition 13.2.5 que $V = W \oplus W^{\perp}$ si et seulement si $W^{\perp} \cap W = \{0\}$ (indépendamment du fait que b soit dégénérée ou non), et les deux exemples qui suivent attestent d'une part qu'il est possible que b soit non dégénérée et qu'il existe un sous-espace W de V tel que $W \cap W^{\perp} \neq \{0\}$, et d'autre part qu'il est possible que b soit dégénérée et qu'il existe un sous-espace vectoriel non trivial W de V tel que $V = W \oplus W^{\perp}$.

Exemple 13.2.7. — Soit V un k-espace vectoriel de dimension 2 et soit (e_1, e_2) une base de V. Soit b la forme bilinéaire symétrique sur V de matrice

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

dans la base (e_1, e_2) ; en d'autres termes, b est définie par la formule

$$b(x_1 e_1 + x_2 e_2, y_1 e_1 + y_2 e_2) = x_1 y_2 + x_2 y_1.$$

La matrice de b étant visiblement inversible, b est non dégénérée.

Puisque b est non dégénérée, $(ke_1)^{\perp}$ est de dimension $2 - 1 = 1$. Mais comme $b(e_1, e_1) = 0$, on a $ke_1 \subset (ke_1)^{\perp}$ et finalement $(ke_1)^{\perp} = ke_1$ (on peut aussi le déduire directement de la formule décrivant b). On montre de même que $(ke_2)^{\perp} = ke_2$.

Exemple 13.2.8. — Soit V un k-espace vectoriel de dimension 2 et soit (e_1, e_2) une base de V. Soit b la forme bilinéaire symétrique sur V de matrice

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

dans la base (e_1, e_2) ; en d'autres termes, b est définie par la formule

$$b(x_1e_1 + x_2e_2, y_1e_1 + y_2e_2) = x_1y_1.$$

La matrice de b n'est pas inversible, et b est donc dégénérée. Son noyau est égal à ke_2 : on le déduit de sa détermination matricielle (13.2.3.1), ou directement de la formule décrivant b que l'on vient de donner. Cette même formule montre aussi que $(ke_1)^\perp = ke_2$; on a donc $V = (ke_1) \oplus (ke_1)^\perp$.

13.3. Diagonalisation des formes quadratiques. — *On suppose à partir de maintenant que k est de caractéristique différente de* 2. Si V est un k-espace on dispose donc d'un isomorphisme canonique entre les espaces $\mathrm{Sym}(V)$ et $\mathrm{Quad}(V)$ (lemme 13.1.9, remarque 13.1.11). Nous parlerons donc de la forme bilinéaire symétrique associée à une forme quadratique, et vice-versa.

13.3.1. — La notion de forme bilinéaire symétrique associée à une forme quadratique permet d'étendre aux formes quadratiques le vocabulaire introduit pour les formes bilinéaires symétriques. Nous appellerons ainsi noyau (resp. rang, resp. matrice dans une base donnée) d'une forme quadratique définie sur un k-espace vectoriel de dimension finie V le noyau (resp. le rang, resp. la matrice dans la base en question) de la forme bilinéaire associée.

Définition 13.3.2. — Soit V un k-espace vectoriel de dimension finie et soit q une forme quadratique sur V. Un vecteur v de V est dit *isotrope* pour la forme q si $q(v) = 0$ et *anisotrope* dans le cas contraire. Le *cône isotrope* de la forme q est l'ensemble des vecteurs de V qui sont isotropes pour q. Un sous-espace vectoriel W de V est dit *anisotrope* (resp. *isotrope*, *totalement isotrope*) pour q si l'intersection de W avec le cône isotrope de q est nulle (resp. non nulle, égale à W tout entier). On dit parfois que q est anisotrope (resp. ...) pour signifier que V est anisotrope (resp. ...) pour q.

Commentaires 13.3.3. — Soit V un k-espace vectoriel de dimension finie et soit q une forme quadratique sur V. Si la forme q est clairement indiquée par le contexte, on parlera simplement de vecteur isotrope, de sous-espace isotrope, de cône isotrope, sans mentionner explicitement q à chaque fois ; et lorsqu'on parlera d'orthogonalité, ce sera sauf mention expresse du contraire en référence à la forme bilinéaire symétrique associée à q.

Mentionnons incidemment qu'un sous-espace W de V est totalement isotrope si et seulement si $q|_W = 0$, et que W est à la fois anisotrope et totalement isotrope si et seulement si $W = \{0\}$.

Remarque 13.3.4. — Soit q une forme quadratique sur un k-espace vectoriel V de dimension finie et soit b la forme bilinéaire associée. On a alors

$$\mathrm{Ker}(q) = \mathrm{Ker}(b) = \{v \in V, \forall w \in V, b(v, w) = 0\} \subset \{v \in V, q(v) = 0\}.$$

Le noyau de q est donc contenu dans son cône isotrope, mais l'inclusion est stricte en générale. En fait, le cône isotrope n'a même aucune raison d'être un sous-espace vectoriel de V ; sa définition montre qu'il est stable sous la multiplication par les scalaires (d'où le terme «cône»), mais il peut très bien ne pas l'être sous l'addition, comme en atteste l'exemple ci-dessous.

Exemple 13.3.5. — Reprenons les notations V, e_1, e_2 et b de l'exemple 13.2.7 ; soit q la forme quadratique associée à b. Elle est donnée par la formule

$$q(xe_1 + ye_2) = 2xy.$$

Le noyau de q est par définition égal au noyau de b qui est trivial (on a vu en *loc. cit.* que b est non dégénérée). Quant à son cône isotrope, c'est l'ensemble des vecteurs de V de la forme $xe_1 + ye_2$ avec $2xy = 0$; c'est donc la *réunion* des deux droites vectorielles ke_1 et ke_2 (rappelons que k est désormais de caractéristique différente de 2), qui n'est évidemment pas stable par addition (par exemple, elle ne contient pas $e_1 + e_2$).

13.3.6. — Soit V un k-espace vectoriel de dimension finie et soit q une forme quadratique sur V. Soit (e_i) une base de V. On sait que la forme q est donnée par une formule du type $\sum x_i e_i \mapsto \sum_i A_i x_i^2 + \sum_{\{i,j\}} A_{\{i,j\}} x_i x_j$ où les A_i et les $A_{\{i,j\}}$ sont des scalaires uniquement déterminés (13.1.8.1). Si l'on désigne par (X_i) la base duale de (e_i) on peut donc écrire $q = \sum_i A_i X_i^2 + \sum_{\{i,j\}} A_{\{i,j\}} X_i X_j$. Il résulte par ailleurs de la remarque 13.1.10 que la matrice de q dans base (e_i) a pour terme général a_{ij} où $a_{i,i} = A_i$ pour tout i, et $a_{ij} = a_{ji} = A_{\{i,j\}}/2$ pour tout (i,j).

Exemple 13.3.7. — Soit V un k-espace vectoriel de dimension 3 et soit $\mathscr{B}$ une base de V de base duale notée (X, Y, Z). Posons

$$q = X^2 + 2Y^2 - 3Z^2 + XY + 2YZ - 5XZ.$$

La matrice de q dans $\mathscr{B}$ est alors égale à

$$\begin{pmatrix} 1 & \frac{1}{2} & -\frac{5}{2} \\ \frac{1}{2} & 2 & 1 \\ -\frac{5}{2} & 1 & -3 \end{pmatrix}.$$

Remarque 13.3.8. — Soit V un k-espace vectoriel de dimension finie et soit q une forme quadratique sur V. Soit v un vecteur non nul de V. La restriction $q|_{kv}$ a pour matrice $[q(v)]$ dans la base v. Par conséquent, $q|_{kv}$ est non dégénérée si et seulement si v est anisotrope. En vertu de l'assertion (3) de la proposition 13.2.5, il s'ensuit que si v est anisotrope, $V = kv \oplus (kv)^{\perp}$.

13.3.9. — Soit q une forme quadratique sur un k-espace vectoriel V de dimension finie. Soit (e_i) une base de V, de base duale (X_i). Les assertions suivantes sont visiblement équivalentes :

(i) q est de la forme $\sum a_i X_i^2$;

(ii) la matrice de q dans la base (e_i) est diagonale ;

(iii) les e_i sont deux à deux orthogonaux.

Si elles sont satisfaites, la matrice de q dans la base (e_i) est $\mathrm{Diag}(a_i)_i$, et le rang de q est donc le cardinal de $\{i, a_i \neq 0\}$.

Théorème 13.3.10 **(diagonalisation des formes quadratiques)**

Soit V un k-espace vectoriel de dimension finie et soit q une forme quadratique sur V. Il existe une base de V dans laquelle la matrice de q est diagonale.

Démonstration. — On raisonne par récurrence sur $n := \dim V$. Si $n = 0$ alors $V = \{0\}$ et la base vide de V (la seule qui existe !) convient. Supposons maintenant $n > 0$ et le résultat vrai en dimension $n - 1$. On distingue deux cas.

Supposons que q est nulle. Dans ce cas, n'importe quelle base de V convient.

Supposons que q est non nulle. Elle possède alors un vecteur anisotrope v. D'après la remarque 13.3.8, on a $V = kv \oplus (kv)^{\perp}$. La dimension de $(kv)^{\perp}$ est alors égale à $n - 1$. En vertu de l'hypothèse de récurrence, il existe une base $(e_2, \ldots, e_n)$ de $(kv)^{\perp}$ dans laquelle la matrice de $q|_{(kv)^{\perp}}$ est diagonale, ce qui revient à dire que les e_i sont deux à deux orthogonaux. Les vecteurs de la base $(v, e_2, \ldots, e_n)$ de V sont alors deux à deux orthogonaux, et la matrice de q dans $(v, e_2, \ldots, e_n)$ est donc diagonale. □

Remarque 13.3.11. — Compte-tenu de l'effet d'un changement de base sur la matrice d'une forme quadratique (13.2.1.3), le théorème ci-dessus peut se reformuler ainsi : *pour tout entier n et toute matrice symétrique M appartenant à $\mathrm{M}_n(k)$, il existe une matrice $P \in \mathrm{GL}_n(k)$ telle que ${}^tPMP$ soit diagonale.*

13.3.12. *Algorithme de Gauß.* — Nous allons maintenant décrire un algorithme rapide de diagonalisation d'une forme quadratique. Plus précisément, cet algorithme permet, partant d'une forme quadratique q, d'obtenir une écriture sous la forme $q = \sum a_\ell \varphi_\ell^2$ où les a_ℓ sont des scalaires et où les φ_ℓ sont des formes linéaires linéairement indépendantes sur k. Pour obtenir une base de diagonalisation de q il suffit alors de prolonger (φ_ℓ) en une base de V^*, puis de prendre sa base antéduale ; dans celle-ci la matrice de q sera diagonale, avec pour coefficients les a_ℓ ainsi qu'un certain nombre de zéros (autant que de formes rajoutées pour compléter (φ_ℓ) en une base). Le rang de q sera donc égal au nombre de a_ℓ non nuls.

Le principe de l'algorithme est le suivant. On suppose donnée une famille libre (X_i) de V^*, et une écriture de q de la forme $\sum_i A_i X_i^2 + \sum_{\{i,j\}} A_{\{i,j\}} X_i X_j$. Nous allons expliquer comment écrire $q = \sum a_\ell \varphi_\ell^2$ où les a_ℓ sont des scalaires et où les φ_ℓ sont linéairement indépendantes et appartiennent à $\mathrm{Vect}(X_i)$. On procède récursivement sur le nombre n de formes X_i. On considère donc qu'on sait faire tourner l'algorithme pour les entiers $< n$, et on distingue plusieurs cas.

Supposons qu'il existe i tel que $A_i \neq 0$. On écrit alors

$$q = A_i X_i \left(X_i + \sum_{j \neq i} \frac{A_{\{i,j\}}}{A_i} X_j \right) + q',$$

où q' s'écrit comme un polynôme homogène de degré 2 en les X_j pour $j \neq i$. On a

$$X_i \left(X_i + \sum_{j \neq i} \frac{A_{\{i,j\}}}{A_i} X_j \right) = \left(X_i + \sum_{j \neq i} \frac{A_{\{i,j\}}}{2A_i} X_j \right)^2 - \frac{1}{4} \left(\sum_{j \neq i} \frac{A_{\{i,j\}}}{A_i} X_j \right)^2,$$

si bien qu'on peut écrire

$$q = A_i \left(X_i + \sum_{j \neq i} \frac{A_{\{i,j\}}}{2A_i} X_j \right)^2 + q'',$$

où q'' s'écrit là encore un polynôme homogène de degré 2 en les X_j pour $j \neq i$. En appliquant l'algorithme à q'' (qui ne met en jeu que $n-1$ formes) on obtient une écriture $q'' = \sum a_\ell \psi_\ell^2$, où les ψ_ℓ sont des formes linéaires linéairement indépendantes appartenant à $\mathrm{Vect}(X_j)_{j\neq i}$. Par ailleurs $\psi := X_i + \sum_{j\neq i} \frac{A_{\{i,j\}}}{2A_i} X_j$ n'appartient pas à $\mathrm{Vect}(X_j)_{j\neq i}$ (en raison de la présence du terme X_i dans son écriture) ; la forme ψ et les ψ_ℓ constituent donc une famille d'éléments linéairement indépendants de $\mathrm{Vect}(X_j)$ (où $j = i$ est désormais autorisé), et $q = A_i\psi^2 + \sum a_\ell \psi_\ell^2$, ce qui termine l'algorithme.

Supposons maintenant que tous les A_i sont nuls. Si tous les $A_{\{i,j\}}$ sont nuls aussi il n'y a rien à faire (on a $q = 0$ et on peut prendre pour φ_ℓ la famille vide de formes). Supposons maintenant qu'il existe une paire $\{i,j\}$ telle que $A_{\{i,j\}} \neq 0$. On a

$$q = A_{\{i,j\}} \left(X_iX_j + X_i \sum_{r\neq i, r\neq j} \frac{A_{\{i,r\}}}{A_{\{i,j\}}} X_r + X_j \sum_{r\neq i, r\neq j} \frac{A_{\{j,r\}}}{A_{\{i,j\}}} X_r \right) + q',$$

où q' s'écrit comme un polynôme homogène de degré 2 en les X_r pour $r \notin \{i,j\}$.

La forme quadratique

$$X_iX_j + X_i \sum_{r\neq i, r\neq j} \frac{A_{\{i,r\}}}{A_{\{i,j\}}} X_r + X_j \sum_{r\neq i, r\neq j} \frac{A_{\{j,r\}}}{A_{\{i,j\}}} X_r$$

peut par ailleurs se récrire

$$\left(X_i + \sum_{r\neq i, r\neq j} \frac{A_{\{j,r\}}}{A_{\{i,j\}}} X_r \right) \left(X_j + \sum_{r\neq i, r\neq j} \frac{A_{\{i,r\}}}{A_{\{i,j\}}} X_r \right) - \left(\sum_{r\neq i, r\neq j} \frac{A_{\{j,r\}}}{A_{\{i,j\}}} X_r \right) \left(\sum_{r\neq i, r\neq j} \frac{A_{\{i,r\}}}{A_{\{i,j\}}} X_r \right),$$

si bien qu'on a

$$q = A_{\{i,j\}} \left(X_i + \sum_{r\neq i, r\neq j} \frac{A_{\{j,r\}}}{A_{\{i,j\}}} X_r \right) \left(X_j + \sum_{r\neq i, r\neq j} \frac{A_{\{i,r\}}}{A_{\{i,j\}}} X_r \right) + q'',$$

où q'' est un certain un polynôme homogène de degré 2 en les X_r pour $r \notin \{i,j\}$.

En posant

$$\varphi = X_i + \sum_{r\neq i, r\neq j} \frac{A_{\{j,r\}}}{A_{\{i,j\}}} X_r \text{ et } \psi = X_j + \sum_{r\neq i, r\neq j} \frac{A_{\{i,r\}}}{A_{\{i,j\}}} X_r,$$

on a

$$q = A_{\{i,j\}}\varphi\psi + q'' = A_{\{i,j\}} \frac{(\varphi+\psi)^2}{4} - A_{\{i,j\}} \frac{(\varphi-\psi)^2}{4} + q''.$$

En appliquant l'algorithme à q'' (qui ne met en jeu que $n-2$ formes) on obtient une écriture $q'' = \sum a_\ell \psi_\ell^2$, où les ψ_ℓ sont des formes linéaires linéairement indépendantes appartenant à $\mathrm{Vect}(X_r)_{r\notin\{i,j\}}$.

Si λ et μ sont deux éléments de k tels que $\lambda(\varphi+\psi) + \mu(\varphi-\psi)$ appartienne à $\mathrm{Vect}(X_r)_{r\notin\{i,j\}}$ on voit en considérant les coefficients de X_i et X_j que $\lambda = \mu = 0$. Par conséquent $\varphi+\psi, \varphi-\psi$ et les ψ_ℓ constituent une famille d'éléments linéairement indépendants de $\mathrm{Vect}(X_r)$ (où $r = i$ et $r = j$ sont désormais autorisés), et

$$q = A_{\{i,j\}} \frac{(\varphi+\psi)^2}{4} - A_{\{i,j\}} \frac{(\varphi-\psi)^2}{4} + \sum a_\ell \psi_\ell^2,$$

ce qui termine l'algorithme.

Exemple 13.3.13. — Soit V un k-plan vectoriel, soit (e_1, e_2) une base de V et soit (X, Y) sa base duale. Notons q la forme quadratique $X^2 + 3XY - 5Y^2$ sur V. La mise en œuvre de l'algorithme de Gauß est très rapide ; en l'appliquant, on obtient les égalités

$$q = X(X + 3Y) - 5Y^2 = (X + 3Y/2)^2 - 9/4Y^2 - 5Y^2.$$

Si l'on pose $S = X + 3Y/2$ alors (S, Y) est une base de V^* et $q = S^2 - (29/4)Y^2$. Si (f_1, f_2) désigne la base antéduale de (S, Y), la matrice de q dans (f_1, f_2) est donc diagonale, et plus précisément égale à

$$\begin{pmatrix} 1 & 0 \\ 0 & -29/4 \end{pmatrix}.$$

Le rang de q vaut 2 si la caractéristique de k est différente de 29, et 1 sinon.

Comment trouver la base (f_1, f_2) ? *Attention : ce n'est pas* $(e_1 + 3/2e_2, e_2)$ *comme on pourrait être tenté de l'écrire*. On raisonne comme suit. Soit P la matrice de passage de (e_1, e_2) à (f_1, f_2). C'est la matrice de l'identité de (f_1, f_2) dans (e_1, e_2). Par conséquent, ${}^tP$ est la matrice de l'identité de (e_1^*, e_2^*) dans (f_1^*, f_2^*), et ${}^tP^{-1}$ est dès lors la matrice de l'identité de (f_1^*, f_2^*) dans (e_1^*, e_2^*), c'est-à-dire la matrice de passage de (e_1^*, e_2^*) à (f_1^*, f_2^*). Or $(e_1^*, e_2^*) = (X, Y)$ et $(f_1^*, f_2^*) = (S, Y) = (X + 3Y/2, Y)$. Il vient

$${}^tP^{-1} = \begin{pmatrix} 1 & 0 \\ 3/2 & 1 \end{pmatrix},$$

c'est-à-dire

$$P^{-1} = \begin{pmatrix} 1 & 3/2 \\ 0 & 1 \end{pmatrix},$$

et finalement

$$P = \begin{pmatrix} 1 & -3/2 \\ 0 & 1 \end{pmatrix},$$

si bien que $f_1 = e_1$ et $f_2 = -(3/2)e_1 + e_2$.

Exemple 13.3.14. — Soit V un k-plan vectoriel, soit (e_1, e_2) une base de V et soit (X, Y) sa base duale. Notons q la forme quadratique XY sur V. La mise en œuvre de l'algorithme de Gauß est très rapide ; en l'appliquant, on obtient les égalités

$$q = XY = \frac{(X + Y)^2}{4} - \frac{(X - Y)^2}{4}.$$

Si l'on pose $S = X+Y$ et $T = X-Y$ alors (S, T) est une base de V^* et $q = S^2/4 - T^2/4$. Si (f_1, f_2) désigne la base antéduale de (S, Y), la matrice de q dans (f_1, f_2) est donc diagonale, et plus précisément égale à

$$\begin{pmatrix} 1/4 & 0 \\ 0 & -1/4 \end{pmatrix}.$$

Le rang de q est égal à 2.

La matrice de passage de (X, Y) à (S, T) est égale à

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Par conséquent, la matrice de passage P de (e_1, e_2) à (f_1, f_2) vérifie l'égalité

$$^tP^{-1} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$

ce qui veut dire que

$$P^{-1} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

puis que

$$P = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{pmatrix}.$$

Par conséquent $f_1 = (1/2)e_1 + (1/2)e_2$ et $f_2 = (1/2)e_1 - (1/2)e_2$.

Exemple 13.3.15. — Soit V un k-espace vectoriel de dimension 3, soit (e_1, e_2, e_3) une base de V et soit (X, Y, Z) sa base duale. Désignons par q la forme quadratique $XY + YZ + XZ$ sur V.

Mettons en œuvre l'algorithme de Gauß. Il vient

$$\begin{aligned} q &= XY + YZ + XZ \\ &= (X + Z)(Y + Z) - Z^2 \\ &= \frac{(X + Y + 2Z)^2}{4} - \frac{(X - Y)^2}{4} - Z^2. \end{aligned}$$

Posons $S = X + Y + 2Z$ et $T = X - Y/$ La famille (S, T, Z) est une base de V^* et $q = S^2/4 - T^2/4 - Z^2$. Si (f_1, f_2, f_3) désigne la base antéduale de (S, T, Z), la matrice de q dans (f_1, f_2, f_3) est diagonale, et plus précisément égale à

$$\begin{pmatrix} 1/4 & 0 & 0 \\ 0 & -1/4 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

La forme q est de rang 3.

La matrice de passage de (X, Y, Z) à (S, T, Z) est égale à

$$\begin{pmatrix} 1/4 & 1/4 & 0 \\ 1/4 & -1/4 & 0 \\ 1/2 & 0 & 1 \end{pmatrix}.$$

Si P désigne la matrice de passage de (e_1, e_2, e_3) à (f_1, f_2, f_3) on a donc

$$^tP^{-1} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 2 & 0 & 1 \end{pmatrix},$$

c'est-à-dire

$$P^{-1} = \begin{pmatrix} 1 & 1 & 2 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

et finalement

$$P = \begin{pmatrix} 1/2 & 1/2 & -1 \\ 1/2 & -1/2 & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$

En conséquence on a :

$$\begin{aligned} f_1 &= (1/2)e_1 + (1/2)e_2 \\ f_2 &= (1/2)e_1 - (1/2)e_2 \\ f_3 &= -e_1 - e_2 + e_3. \end{aligned}$$

13.4. Généralités sur les isométries. — Lorsqu'on s'intéresse aux k-espaces vectoriels munis de formes quadratiques (ou de formes bilinéaires symétriques, ce qui revient au même), il est naturel de considérer les applications linéaires compatibles aux formes en question ; cela va conduire à la notion d'*isométrie*. Avant de la présenter, nous allons introduire un peu de vocabulaire.

13.4.1. — Nous appellerons *k-espace quadratique* un k-espace vectoriel V de dimension finie muni d'une forme quadratique q. S'il n'est pas nécessaire de mentionner explicitement la forme, nous nous contenterons de nommer l'espace ; nous pourrons donc écrire des phrases comme «soit V un k-espace quadratique». S'il nous faut être plus précis, nous écrirons «soit (V, q) un k-espace quadratique», voire «soit (V, q, b)» un espace quadratique pour indiquer que la forme bilinéaire associée à q est notée b (mais la donnée de b est redondante, puisqu'elle est entièrement déterminée par q).

On dira qu'un k-espace quadratique (V, q, b) est non dégénéré si q est non dégénérée, qu'il est totalement isotrope si $q = 0$, et qu'il est anisotrope si le cône isotrope de q est réduit à $\{0\}$ (remarquons que comme le noyau d'une forme quadratique est contenu dans son cône isotrope, tout k-espace quadratique anisotrope est non dégénéré). Lorsqu'on parlera d'orthogonalité ou d'isotropie dans l'espace V, ce sera toujours en référence à q et b.

Tout sous-espace vectoriel d'un sous-espace quadratique hérite (par restriction) d'une structure naturelle d'espace quadratique.

Si (V, q, b) est un k-espace quadratique, $V/\mathrm{Ker}(b)$ hérite *via* la forme $\overline{b}$ induite par b (13.2.3.2) d'une structure de k-espace quadratique non dégénéré.

13.4.2. *Sommes directes orthogonales : le cas interne.* — Soit (V, q, b) un k-espace quadratique. Si (V_i) est une famille finie de sous-espaces vectoriels de V, nous écrirons $V = \overset{\perp}{\bigoplus} V_i$ pour signifier que V est somme directe des V_i et que ces derniers sont deux à deux orthogonaux. On suppose à partir de maintenant que c'est le cas.

13.4.2.1. — Si (v_i) et (w_i) sont deux éléments de $\prod V_i$, on a les égalités

$$b\left(\sum v_i, \sum w_i\right) = \sum_i b(v_i, w_i) \text{ et } q\left(\sum v_i\right) = \sum q(v_i).$$

13.4.2.2. *Aspect matriciel.* — Donnons-nous une base $\mathscr{B}_i$ de V_i pour tout i, et soit $\mathscr{B}$ la base de V obtenue par concaténation des $\mathscr{B}_i$. Pour tout i, désignons par M_i la matrice de $q|_{V_i}$ dans $\mathscr{B}_i$. La matrice de q dans $\mathscr{B}$ est alors la matrice $\mathrm{Diag}(M_i)$ (c'est une conséquence immédiate de 13.4.2.1).

13.4.2.3. — Soit $(v_i) \in \prod V_i$. La formule donnée en 13.4.2.1 montre aussitôt que $\sum v_i \in \mathrm{Ker}(b)$ si et seulement si $v_i \in \mathrm{Ker}(b|_{V_i})$ pour tout i. Autrement dit, on a l'égalité $\mathrm{Ker}(b) = \bigoplus \mathrm{Ker}(b|_{V_i})$; en particulier, V est non dégénéré si et seulement si chacun des V_i est non dégénéré.

Supposons que ce soit le cas, et fixons i. Comme V_i est non dégénéré, $V = V_i \oplus V_i^{\perp}$. Combiné au fait que $V = \overset{\perp}{\bigoplus_j} V_j$, ceci entraîne que $V_i^{\perp} = \bigoplus_{j \neq i} V_j$.

13.4.3. *Sommes directes orthogonales : le cas externe.* — Soit (V_i, q_i, b_i) une famille finie d'espaces quadratiques. On appelle somme directe orthogonale *externe* des espaces quadratiques V_i, et l'on note $\overset{\perp}{\bigoplus} V_i$, l'espace quadratique $(\bigoplus V_i, q, b)$ où b et q sont définies par les formules

$$b((v_i),(w_i)) = \sum_i b_i(v_i, w_i) \text{ et } q((v_i)) = \sum q_i(v_i).$$

Par construction, chacun des V_i s'identifie comme espace quadratique à un sous-espace de $\overset{\perp}{\bigoplus} V_i$, et modulo ces identifications $\overset{\perp}{\bigoplus} V_i$ est la somme directe orthogonale interne des V_i : cette construction force donc en quelque sorte les V_i à être en somme directe orthogonale interne (mais comme d'habitude, il faut faire attention : plusieurs V_i peuvent être égaux à un même espace quadratique W, et ils donnent néanmoins lieu à des sommandes distincts dans la somme directe orthogonale externe, auxquels on pense comme à des copies de W).

13.4.4. — Soit (V, q, b) un espace quadratique et soit W un k-espace vectoriel. Soit $f \colon W \to V$ une application linéaire. L'application de $W \times W$ dans k qui envoie un couple (w, w') sur $b(f(w), f(w'))$ est bilinéaire symétrique ; la forme quadratique correspondante est $w \mapsto b(f(w), f(w)) = q(u(w))$; c'est donc $q \circ f$.

Définition 13.4.5. — Soient (V_1, q_1, b_1) et (V_2, q_2, b_2) deux espaces quadratiques. Si f est une application linéaire de V_1 vers V_2, il résulte de 13.4.4 que les assertions suivantes sont équivalentes :

(i) $q_2(f(v)) = q_1(v)$ pour tout $v \in V_1$;

(ii) $b_2(f(v), f(w)) = b_1(v, w)$ pour tout $(v, w) \in V_1^2$.

On dit que f est une *isométrie* si f est bijective et si elle satisfait (i) et (ii). Notons que pour vérifier (ii) on peut par bilinéarité se contenter de faire varier v et w au sein d'une base donnée de V_1.

Remarque 13.4.6. — On impose par définition à une isométrie d'être bijective. Mais avec les notations de la définition ci-dessus, si b_1 est non dégénérée et si f satisfait (i) et (ii), alors f est automatiquement injective (et donc bijective si $\dim V_2 = \dim V_1$). En effet, supposons donc b_1 non dégénérée et soit v un élément de $\mathrm{Ker}(u)$. Pour tout $w \in V$ on a

$$b_1(v, w) = b_2(f(v), f(w)) = 0,$$

si bien que v appartient à $\mathrm{Ker}(b)$, qui est nul par hypothèse ; par conséquent $v = 0$ et f est injective.

Définition 13.4.7. — On dit que deux k-espaces quadratiques (V, q) et (W, ρ) sont *isométriques* s'il existe une isométrie $f \simeq V \to W$. Cela revient visiblement à demander qu'ils aient même dimension et qu'il existe une base $\mathscr{B}$ de V et une base $\mathscr{C}$ de W telles que $\mathrm{Mat}_{\mathscr{B}}\, q = \mathrm{Mat}_{\mathscr{C}}\, \rho$.

Définition 13.4.8. — Si (V, q) est un k-espace quadratique, l'ensemble des isométries de V sur lui-même est un sous-groupe de $\mathrm{GL}(V)$ qui est appelé le *groupe orthogonal* de V (ou de (V, q), ou de q) et est noté $\mathrm{O}(V)$ (ou $\mathrm{O}(V, q)$, ou $\mathrm{O}(q)$).

Nous considérerons dans la suite deux types de questions à propos de la notion d'isométrie : d'une part la classification des espaces quadratiques à isométrie près (c'est un problème dont la difficulté dépend de façon cruciale du corps de base), et d'autre part l'étude du groupe orthogonal d'un espace quadratique. Avant cela, nous allons donner quelques exemples d'isométries.

Exemple 13.4.9 **(Somme directe orthogonale d'isométries)**
Soient $(V_i)_{i\in I}$ et $(W_i)_{i\in I}$ deux familles finies de k-espaces quadratiques et soit $(f_i \colon V_i \to W_i)$ une famille d'isométries. L'application

$$\bigoplus f_i \colon \overset{\perp}{\bigoplus} V_i \to \overset{\perp}{\bigoplus} W_i,\ (v_i)_i \mapsto (u_i(v_i))_i$$

est une isométrie.

Exemple 13.4.10 **(Symétries orthogonales).** — Soit (V, q) un k-espace quadratique et soit W un sous-espace vectoriel de V tel que $V = W \oplus W^{\perp}$. Soit σ la symétrie par rapport à W et parallèlement à $W^{\perp}$. L'application σ est une isométrie. En effet, soit $v \in V$; écrivons $v = w + w'$ avec $w \in W$ et $w' \in W^{\perp}$. On a $\sigma(v) = w - w'$, si bien que

$$\begin{aligned} q(\sigma(v)) &= q(w - w') \\ &= q(w) + q(w') \\ &= q(w + w') \\ &= q(v) \end{aligned}$$

(les deuxième et troisième égalités utilisent le fait que $w \perp w'$), ce qu'il fallait démontrer.

Une telle symétrie est appelée une *symétrie orthogonale*, et une *réflexion* si W est un hyperplan de V.

Notation 13.4.11. — Soit $(a_1, \dots, a_n)$ une famille finie d'éléments de k. Nous noterons $\langle a_1, \dots, a_n \rangle$ le k-espace quadratique k^n muni de la forme quadratique de matrice $\mathrm{Diag}(a_1, \dots, a_n)$ dans la base canonique, qu'on peut également définir comme étant égale à $\sum a_i X_i^2$, où (X_i) est la base duale de la base canonique, ou encore comme l'application

$$\begin{array}{ccc} k^n & \longrightarrow & k, \\ (x_1, \ldots, x_n) & \mapsto & \sum a_i x_i^2. \end{array}$$

Le rang de cette forme est le cardinal de $\{i, a_i \neq 0\}$.

13.4.11.1. — Il résulte immédiatement de la définition que pour toute famille finie $(a_1, \ldots, a_n, b_1, \ldots, b_m)$ d'éléments de k, le k-espace quadratique

$$\langle a_1, \ldots, a_n \rangle \overset{\perp}{\oplus} \langle b_1, \ldots, b_m \rangle$$

est isométrique à $\langle a_1, \ldots, a_n, b_1, \ldots, b_m \rangle$.

13.4.11.2. — Soit $(a_1, \ldots, a_n)$ une famille d'éléments de k. Un k-espace quadratique (V, q) est isométrique à $\langle a_1, \ldots, a_n \rangle$ si et seulement si V possède une base dans laquelle q a pour matrice $\mathrm{Diag}(a_1, \ldots, a_n)$, c'est-à-dire encore si et seulement si V^* possède une base (X_i) telle que $q = \sum a_i X_i^2$.

En particulier, $\langle a_1, \ldots, a_n \rangle$ est isométrique à $\langle a_{\sigma(1)}, \ldots, a_{\sigma(n)} \rangle$ pour tout $\sigma \in \mathfrak{S}_n$ (permuter les éléments de la base canonique) ; et $\langle a_1, \ldots, a_n \rangle$ est isométrique à $\langle \alpha_1^2 a_1, \ldots, \alpha_n^2 a_n \rangle$ pour toute famille (α_i) d'éléments non nuls de k (dilater les éléments de la base canonique).

13.4.11.3. — On peut reformuler le théorème de diagonalisation des formes quadratiques (th. 13.3.10) en disant que tout k-espace quadratique est isométrique à $\langle a_1, \ldots, a_n \rangle$ pour une certaine famille finie $(a_1, \ldots, a_n)$.

Exemple 13.4.12 **(Classes d'isométrie de droites quadratiques)**
Commençons par une première remarque. Soit (D, q) une k-droite quadratique (c'est-à-dire un k-espace vectoriel quadratique de dimension 1). Soit e un vecteur non nul de D. La matrice de q dans la base e est alors la matrice scalaire $[q(e)]$; par conséquent, (D, q) est isométrique à $\langle q(e) \rangle$.

Soient maintenant a et b deux éléments de k ; nous allons chercher une condition nécessaire et suffisante pour que $\langle a \rangle$ et $\langle b \rangle$ soient isométriques.

Par définition, ces espaces sont isométriques si et seulement s'il existe une bijection k-linéaire $f \colon k \to k$ telle que $bf(x)^2 = ax^2$ pour tout $x \in k$. Comme les automorphismes k-linéaires de k sont exactement les homothéties de rapport inversible, cela revient à demander qu'il existe $\lambda \in k^\times$ tel que $b\lambda^2 x^2 = ax^2$ pour tout $x \in k$; c'est le cas si et seulement s'il existe $\lambda \in k^\times$ tel que $a = b\lambda^2$ (c'est clairement suffisant, et nécessaire comme on le voit en testant l'égalité requise avec $x = 1$).

Ainsi, $a \mapsto \langle a \rangle$ établit une bijection entre l'ensemble des orbites de k sous l'action $(\lambda, \alpha) \mapsto \lambda^2 \alpha$ de $k^\times$ et l'ensemble des classes d'isométries de k-droites quadratiques ; la réciproque envoie une droite quadratique (D, q) sur l'orbite de $q(e)$ pour n'importe quel vecteur non nul e de D.

Comme $\langle a\rangle$ est non dégénéré si et seulement si a est non nul, on voit que la bijection ci-dessus induit une bijection entre $k^\times/(k^\times)^2$ et l'ensemble des classes d'isométries de k-droites quadratiques non dégénérées.

13.4.13. *Groupe orthogonal : aspect matriciel.* — Soit (V, q) un k-espace quadratique ; soit $(e_i)_{i\in I}$ une base de V et soit M la matrice de q dans (e_i).

13.4.13.1. — Soit f un automorphisme k-linéaire de V. Soit N la matrice de f dans la base (e_i). Si $v = \sum x_i e_i$ est un élément de V et si X désigne la famille (x_i) vue comme vecteur colonne, on a $q(v) = {}^tXMX$ et

$$q(f(v))) = {}^t(NX)M(NX) = {}^tX({}^tNMN)X.$$

Par conséquent, $q\circ f$ est la forme quadratique de matrice ${}^tNMN$ dans la base (e_i). Il s'ensuit que pour que f soit une isométrie, c'est-à-dire pour que $q \circ f = q$, il faut et il suffit que ${}^tNMN = M$.

13.4.13.2. — Supposons de plus que q est non dégénérée. Dans ce cas la matrice M est inversible, et si N est une matrice de $\mathrm{GL}_I(k)$ telle que ${}^tNMN = M$ on a $(\det N)^2 \det M = \det M$, ce qui entraîne que $(\det N)^2 = 1$ puisque $\det M \neq 0$. Par conséquent, toute isométrie de V a un déterminant égal à 1 ou à (-1). Nous dirons qu'une isométrie de V est *directe* si son déterminant est égal à 1, et *indirecte* sinon. L'ensemble des isométries directes de V est un sous-groupe distingué de $\mathrm{O}(V)$ (c'est le noyau du déterminant), qui est appelé le *groupe spécial orthogonal* de V (ou de (V, q), ou de q) et est noté $\mathrm{SO}(V)$ (ou $\mathrm{SO}(V, q)$, ou $\mathrm{SO}(q)$).

Si $V = \{0\}$ on a évidemment $\mathrm{SO}(V) = \mathrm{O}(V) = \mathrm{GL}(V) = \{\mathrm{Id}_V\}$. Si $\dim V > 0$ alors $\mathrm{SO}(V)$ est d'indice 2 dans $\mathrm{O}(V)$. En effet, il suffit pour le voir de s'assurer que $\det\colon \mathrm{O}(V) \to \{-1, 1\}$ est surjectif. Or comme $\dim V > 0$ et comme q est non dégénérée, q n'est pas identiquement nul, et il existe donc un vecteur non isotrope v dans V. On a $V = kv \oplus (kv)^\perp$ (remarque 13.3.8), et la réflexion par rapport à $(kv)^\perp$ est alors une isométrie indirecte de V (exemple 13.4.10).

13.4.14. *Adjoint d'un endomorphisme, lien avec les isométries.* — Soit (V, q, b) un espace quadratique non dégénéré.

13.4.14.1. — Soit f un endomorphisme de V. L'application linéaire $b_{\mathrm{g}}\colon V \to V^*$ est bijective. Il existe donc une unique application linéaire $f^*\colon V \to V$ faisant commuter le diagramme

$$\begin{tikzcd} V \arrow[r, "b_{\mathrm{g}}"] \arrow[d, "f^*"'] & V^* \arrow[d, "{}^tf"] \\ V \arrow[r, "b_{\mathrm{g}}"] & V^* \end{tikzcd},$$

à savoir $b_{\mathrm{g}}^{-1} \circ {}^tf \circ b_{\mathrm{g}}$; on dit que f^* est *l'adjoint* de l'endomorphisme f. La commutativité du diagramme ci-dessus peut se reformuler de manière plus tangible. Elle signifie en effet que pour tout $v \in V$ on a les égalités

$$b_{\mathrm{g}}(f^*(v)) = {}^tf(b_{\mathrm{g}}(v)) = b_{\mathrm{g}}(v) \circ f$$

entre applications linéaires de V dans k ; or l'application de gauche envoie un vecteur w de V sur $b(f^*(v), w)$, et celle de droite l'envoie sur $b(v, f(w))$. L'endomorphisme f^* est donc *caractérisé* par le fait que

$$b(f^*(v), w) = b(v, f(w))$$

pour tout couple (v, w) d'éléments de V.

13.4.14.2. — On déduit de la formule

$$f^* = b_{\mathrm{g}}^{-1} \circ {}^t f \circ b_{\mathrm{g}}.$$

que $f \mapsto f^*$ est linéaire, que $\mathrm{Id}_V^* = \mathrm{Id}_V$, que $(f \circ g)^* = g^* \circ f^*$ pour tout couple (f, g) d'endomorphismes de V, et que si f est un automorphisme de V alors f^* est bijectif d'inverse $(f^{-1})^*$.

Par ailleurs, si f est un endomorphisme de V son adjoint f^* est caractérisé par le fait que

$$b(f^*(v), w) = b(v, f(w))$$

pour tout couple (v, w), et cette condition est symétrique en f et f^* ; par conséquent $(f^*)^* = f$.

13.4.14.3. *Aspect matriciel de l'adjonction.* — Soit f un endomorphisme de V et soit (e_i) une base de V. Soient M et N les matrices respectives de q et f dans la base (e_i). Comme N s'interprète également comme la matrice de b_{g} dans les bases (e_i) et (e_i^*), il résulte de l'égalité $f^* = b_{\mathrm{g}}^{-1} \circ {}^t f \circ b_{\mathrm{g}}$ que la matrice de f^* dans la base (e_i) est égale à $(M^{-1})^t NM$.

13.4.14.4. *Adjonction et isométries.* — Soit f un automorphisme linéaire de V. L'application f est une isométrie si et seulement si $b(f(v), f(w)) = b(v, w)$ pour tout $(v, w) \in V \times V$. Puisque f^{-1} est bijective, ceci revient à demander que $b(f(v), f(f^{-1}(u))) = b(v, f^{-1}(u))$ pour tout $(v, u) \in V^2$, c'est-à-dire encore que $b(f(v), u) = b(v, f^{-1}(u))$ pour tout $(v, u) \in V^2$. Par conséquent, f *est une isométrie si et seulement si* $f^* = f^{-1}$.

Si (e_i) est une base de V et si M et N désignent les matrices respectives de q et f dans la base (e_i), on déduit de ce qui précède et de 13.4.14.3 que f est une isométrie si et seulement si

$$N^{-1} = (M^{-1})^t NM,$$

ce qui équivaut bien à la condition ${}^t NMN = M$ qui a été vue plus haut par d'autres méthodes (13.4.13.1).

13.5. Espaces quadratiques à isométrie près : premières réductions. — Nous allons maintenant entamer l'étude générale de la classification des k-espaces quadratiques à isométrie près. Nous allons commencer par effectuer un certain nombre de dévissages pour se ramener tout d'abord au cas des espaces non dégénérés, puis à celui des espaces anisotropes.

Lemme 13.5.1. — *Soit V un k-espace quadratique. Il existe un sous-espace totalement isotrope W et un sous-espace non dégénéré S de V tels que $V = W \overset{\perp}{\oplus} S$. De plus W coïncide nécessairement avec $V^\perp$, et il est en particulier unique ; quant à*

S, il est isométrique au quotient $V/V^{\perp}$, et sa classe d'isométrie est donc uniquement déterminée.

Démonstration. — Soit S un supplémentaire quelconque de $V^{\perp}$ dans V. L'espace S étant orthogonal à $V^{\perp}$ par définition de ce dernier, on a $V = V^{\perp} \oplus S$. L'espace $V^{\perp}$ est totalement isotrope par construction ; et comme tout élément de V est orthogonal à $V^{\perp}$, un élément de V est orthogonal à S si et seulement s'il est orthogonal à V, c'est-à-dire si et seulement s'il appartient à $V^{\perp}$. Autrement dit $S^{\perp} = V^{\perp}$, ce qui entraîne que $S \cap S^{\perp} = \{0\}$, c'est-à-dire que S est non dégénéré. Ceci achève de montrer l'existence de la décomposition annoncée.

Venons-en aux assertions d'unicité. Considérons donc une décomposition

$$V = W \overset{\perp}{\oplus} S$$

comme dans l'énoncé, et notons b la forme bilinéaire symétrique structurale sur l'espace quadratique V. Soit v un vecteur de V ; écrivons $v = w + s$ avec $w \in W$ et $s \in S$. Le vecteur v appartient à $V^{\perp}$ si et seulement si pour tout couple (w', s') de $W \times S$ on a $b(w+s, w'+s') = 0$. Mais $b(w+s, w+s') = b(w, w') + b(s, w') + b(s', w) + b(s, s')$ et compte-tenu du caractère totalement isotrope de W et de son orthogonalité avec S, les trois premiers termes de cette somme sont nuls ; celle-ci est donc égale à $b(s, s')$, d'où il découle que $v \in V^{\perp}$ si et seulement si $b(s, s') = 0$ pour tout $s' \in S$. Puisque S est non dégénéré, c'est le cas si et seulement si $s = 0$, c'est-à-dire si et seulement si $v \in W$. Par conséquent, $W = V^{\perp}$.

Puisque $W = V^{\perp}$, l'application quotient de V vers $V/V^{\perp}$ induit un isomorphisme $S \simeq V/V^{\perp}$, qui est une isométrie par définition de la forme bilinéaire $\overline{b}$ que b induit sur $V/V^{\perp}$. □

En particulier, on voit que tout k-espace quadratique s'écrit comme somme directe orthogonale d'un k-espace quadratique totalement isotrope et d'un k-espace quadratique non dégénéré, et que les deux termes de cette écriture sont uniquement déterminés à isométrie près. La classification des k-espaces quadratiques à isométrie près se ramène ainsi à celle des k-espaces quadratiques non dégénérés, que nous allons maintenant aborder. Nous allons chercher à isoler une «partie anistrope» d'un tel sous-espace ; cela requiert tout d'abord de comprendre «d'où viennent» ses éventuels vecteurs isotropes.

Lemme-Définition 13.5.2. — *Soit (V, q) un k-plan quadratique non dégénéré. Les assertions suivantes sont équivalentes.*

(i) *Le plan quadratique V est isotrope.*

(ii) *Il existe une base (S, T) de V^* telle que $q = ST$.*

(iii) *Il existe une base de V dans laquelle la matrice de q est*

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

(iv) *Le k-espace quadratique V est isométrique à $\langle 1, -1 \rangle$.*

Lorsqu'elles sont satisfaites, on dit que V est un plan hyperbolique.

Démonstration. — On note b la forme bilinéaire symétrique sur V associée à q. Il est immédiat que (ii) $\Longleftrightarrow$ (iii).

Supposons que (i) soit vraie. Choisissons un vecteur v non nul de V tel que $q(v) = 0$, et complétons v en une base (v, w) de V. Comme q est non dégénérée, $(kv)^{\perp}$ est de dimension $2 - 1 = 1$. Or puisque v est isotrope, on a $kv \subset (kv)^{\perp}$, d'où finalement l'égalité $(kv)^{\perp} = kv$; en particulier, w n'appartient pas à $(kv)^{\perp}$, c'est-à-dire que $b(v, w) \neq 0$.

Pour tout $\lambda \in k$ on a $q(w + \lambda v) = q(w) + 2\lambda b(v, w)$ (puisque $q(v) = 0$) ; il s'ensuit que

$$q\left(w - \frac{q(w)}{2b(v,w)}v\right) = 0.$$

Quitte à remplacer w par $w - \frac{q(w)}{2b(v,w)}v$, on peut donc supposer que $q(w) = 0$; puis quitte à diviser w par $b(v, w)$ on peut supposer que $b(v, w) = 1$. La matrice de q dans la base (v, w) est alors égale à

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

d'où (iii).

Supposons que (ii) est vraie. On a alors

$$q = ST = \frac{(S+T)^2 - (S-T)^2}{4} = \left(\frac{S+T}{2}\right)^2 - \left(\frac{S-T}{2}\right)^2.$$

Posons $X = (S + T)/2$ et $Y = (S - T)/2$; la famille (X, Y) est alors une base de V^*, et $q = X^2 - Y^2$. Ainsi, (iv) est vraie.

Supposons enfin que (iv) soit vraie, et soit (e_1, e_2) une base dans laquelle la matrice de q est égale à

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

On a alors $q(e_1 + e_2) = 1^2 - 1^2 = 0$, et le vecteur non nul $e_1 + e_2$ est isotrope, ce qui montre (i) et achève la démonstration. □

Il y a donc à isométrie près un seul plan quadratique non dégénéré isotrope : «le» plan hyperbolique $\langle 1, -1 \rangle$. La situation n'est pas si simple en dimension plus grande, mais les plans hyperboliques y restent tout de même la seule source d'isotropie, comme en atteste la proposition suivante.

Proposition 13.5.3. — *Soit V un k-espace quadratique non dégénéré et soit $(v_1, \ldots, v_r)$ une famille libre d'éléments de V telle que le sous-espace vectoriel engendré par les v_i soit totalement isotrope.*

Il existe une famille $(w_1, \ldots, w_r)$ de vecteurs de V telle que les propriétés suivantes soient satisfaites :

- ◇ *pour tout i, l'espace $H_i := \mathrm{Vect}(v_i, w_i)$ est un plan hyperbolique ;*
- ◇ $\mathrm{Vect}(v_1, w_1, \ldots, v_r, w_r) = \overset{\perp}{\bigoplus} H_i$.

Démonstration. — On note b (resp. q) la formes bilinéaire symétrique (resp. quadratique) structurale de V. On raisonne par récurrence sur r. Si $r = 0$ la proposition est trivialement vraie (on prend évidemment pour $(w_i)_i$ la famille vide). Supposons $r > 0$ et le résultat vrai pour les entiers $< r$. Comme $(v_1, \ldots, v_r)$ est libre, il existe une forme linéaire φ sur V telle que $\varphi(v_1) = 1$ et $\varphi(v_i) = 0$ pour tout $i > 1$ (prolonger (v_i) en une base de V et prendre pour φ la première forme linéaire coordonnée). Comme b est non dégénérée, $b_{\mathrm{g}} \colon V \to V^*$ est bijective ; il existe par conséquent un vecteur $w_1 \in V$ tel que $b_{\mathrm{g}}(w_1) = \varphi$.

Posons $H_1 = \mathrm{Vect}(v_1, w_1)$. Par définition de w_1 on a $b(w_1, v_1) = \varphi(v_1) = 1$; par conséquent, w_1 n'est pas colinéaire à v_1, puisque ce dernier est isotrope par hypothèse, ce qui veut dire que $q(v_1) = b(v_1, v_1) = 0$. L'espace H_1 est donc de dimension 2, et la matrice de $q|_{H_1}$ dans la base (v_1, w_1) est égale à

$$\begin{pmatrix} 0 & 1 \\ 1 & q(w_1) \end{pmatrix}.$$

Comme elle est inversible, H_1 est non dégénéré. Puisque v_1 est isotrope, H_1 est un plan hyperbolique. Par ailleurs, le caractère non dégénéré de H_1 assure que $V = H_1 \oplus H_1^{\perp}$. On a pour tout $i > 1$ l'égalité $b(w_1, v_i) = \varphi(v_i) = 0$ par choix de w_1, et l'égalité $b(v_1, v_i) = 0$ car $\mathrm{Vect}(v_1, \ldots, v_r)$ est totalement isotrope. En conséquence les vecteurs v_i pour $i \geqslant 2$ appartiennent à $H_1^{\perp}$.

L'hypothèse de récurrence appliquée à la famille $(v_2, \ldots, v_r)$ de l'espace $H_1^{\perp}$ (qui est non dégénéré car $V = H_1 \oplus H_1^{\perp}$ est non dégénéré) assure l'existence d'une famille $(w_2, \ldots, w_r)$ de vecteurs de $H_1^{\perp}$ satisfaisant les propriétés suivantes :

- ◇ pour tout $i \geqslant 2$, l'espace $H_i := \mathrm{Vect}(v_i, w_i)$ est un plan hyperbolique ;
- ◇ $\mathrm{Vect}(v_2, w_2, \ldots, v_r, w_r) = \overset{\perp}{\bigoplus\limits_{i \geqslant 2}} H_i$.

La famille $(w_1, \ldots, w_r)$ satisfait alors clairement les propriétés requises. □

Corollaire 13.5.4. — *Soit (V, q) un k-espace quadratique non dégénéré et isotrope. Pour tout $\lambda \in k$ il existe $v \in V$ tel que $q(v) = \lambda$.*

Démonstration. — Puisque V est non dégénéré et isotrope, la proposition 13.5.3 assure que V contient un plan hyperbolique H ; quitte à remplacer V par H, on peut supposer que V lui-même est un plan hyperbolique. Il possède alors une base (e_1, e_2) de V telle que $q(x_1 e_1 + x_2 e_2) = x_1 x_2$ pour tout $(x_1, x_2) \in k^2$. En particulier, $q(\lambda e_1 + e_2) = \lambda$. □

Pour pouvoir utiliser la proposition 13.2.5 afin d'isoler une «partie anisotrope» dans un k-espace quadratique, nous allons avoir besoin d'un «théorème de simplification» qui est par ailleurs très intéressant en lui-même, et a bien d'autres applications dans la théorie.

Lemme 13.5.5. — *Soit (V, q) un k-espace quadratique non dégénéré et soient v et w deux vecteurs anisotropes de V tels que $q(v) = q(w)$. Il existe une isométrie f de V telle que $f(v) = w$. On peut même choisir pour f une réflexion si $v - w$ est anisotrope, et la composée de deux réflexions sinon.*

Démonstration. — Soit b la forme bilinéaire symétrique associée à la forme quadratique q. On a $q(v) - q(w) = 0$, c'est-à-dire $b(v+w, v-w) = 0$. Par ailleurs l'un au moins des deux vecteurs $v+w$ et $v-w$ est anisotrope. En effet, s'ils étaient tous les deux isotropes, on aurait

$$q(v) = \frac{1}{4}q((v+w)+(v-w)) = \frac{1}{4}(q(v+w)+q(v-w)+2b(v+w,v-w)) = 0,$$

ce qui est contradictoire avec nos hypothèses.

Supposons que $v-w$ est anisotrope. On a alors $V = k(v-w) \oplus (k(v-w))^{\perp}$. Soit f la réflexion par rapport à l'hyperplan $F := (k(v-w))^{\perp}$. Comme $v-w$ est orthogonal à F par définition on a $f(v-w) = w-v$. Comme $b(v+w, v-w) = 0$ on a $v+w \in F$ et donc $f(v+w) = v+w$. Il vient

$$f(v) = \frac{1}{2}f((v+w)+(v-w)) = \frac{1}{2}(v+w+w-v) = w.$$

Supposons que $v+w$ est anisotrope. On a alors $V = k(v+w) \oplus (k(v+w))^{\perp}$. Soit g la réflexion par rapport à l'hyperplan $G := (k(v+w))^{\perp}$. Comme $v+w$ est orthogonal à G par définition on a $g(v+w) = -w-v$. Comme $b(v+w, v-w) = 0$ on a $v-w \in G$ et donc $g(v-w) = v-w$. Il vient

$$f(v) = \frac{1}{2}g((v+w)+(v-w)) = \frac{1}{2}(-v-w+v-w) = -w.$$

Puisque w est anisotrope, on a $V = kw \oplus (kw)^{\perp}$. Soit h la réflexion par rapport à l'hyperplan $H := (kw)^{\perp}$. Comme w est orthogonal à h par définition on a $h(w) = -w$. Par conséquent $h \circ g(v) = h(-w) = w$. □

***Théorème 13.5.6* (Théorème de simplification de Witt)**
Soient U, V et W trois k-espaces quadratiques non dégénérés. Supposons que $U \overset{\perp}{\oplus} V$ est isométrique à $U \overset{\perp}{\oplus} W$. Alors V est isométrique à W.

Démonstration. — On raisonne par récurrence sur la dimension de U. Si $U = \{0\}$ le théorème est vide (et vrai). Supposons maintenant $\dim U > 0$ et le résultat vrai en dimension $< \dim U$.

Posons $E = U \overset{\perp}{\oplus} V$; c'est un espace quadratique non dégénéré (puisque U et V sont non dégénérés). Par hypothèse, il existe une isométrie f de $U \overset{\perp}{\oplus} W$ sur E. Posons $U' = f(U)$. À l'intérieur de l'espace quadratique E on a $U^{\perp} = V$ et $(U')^{\perp} = f(W)$. Comme f est une isométrie, U et U' sont isomorphes, et il suffit de montrer que $V = U^{\perp}$ est isométrique à $f(W) = (U')^{\perp}$. Notons que $U^{\perp}$ et $(U')^{\perp}$ sont tous deux non dégénérés (car $U^{\perp} = V$ et car $(U')^{\perp}$ est isométrique à W).

Comme U est non nul et non dégénéré, $q|_U$ est non nulle. Il existe donc un vecteur anisotrope $u \in U$. Par hypothèse, il existe une isométrie $\sigma\colon U \simeq U'$; posons $v = \sigma(u)$ et remarquons que comme σ est une isométrie on a $q(v) = q(u) \neq 0$. Soit F l'orthogonal de u dans U et soit G celui de v dans U'. Comme $v = \sigma(u)$ on a $G = \sigma(F)$, et les espaces quadratiques F et G sont en particulier isométriques. Comme u est anisotrope on a $U = (ku) \overset{\perp}{\oplus} F$ (ce qui entraîne que F est non dégénéré, puisque c'est le cas de

U). De même, $U' = (kv) \overset{\perp}{\oplus} G$, et G est non dégénéré. On a alors $E = (ku) \overset{\perp}{\oplus} F \overset{\perp}{\oplus} U^{\perp}$ et $E = (kv) \overset{\perp}{\oplus} G \overset{\perp}{\oplus} (U')^{\perp}$.

Puisque E est non dégénéré et puisque $q(u) = q(v) \neq 0$, il résulte du lemme 13.5.5 qu'il existe une isométrie τ de E telle que $\tau(u) = v$. L'isométrie τ envoie l'orthogonal de u dans E sur celui de v ; autrement dit, elle envoie $F \overset{\perp}{\oplus} U^{\perp}$ sur $G \overset{\perp}{\oplus} (U')^{\perp}$. Ces deux espaces sont donc isométriques. Comme F et G sont isométriques et de dimension $\dim U - 1$ et comme tous les espaces en jeu sont non dégénérés, notre hypothèse de récurrence assure que $U^{\perp}$ et $(U')^{\perp}$ sont isométriques. □

Théorème-Définition 13.5.7. — *Soit V un k-espace quadratique non dégénéré. Il existe un entier d et des sous-espaces $H_1, \ldots, H_d, W$ de V tels que :*

- *les H_i sont des plans hyperboliques ;*
- *W est anisotrope ;*
- *$V = H_1 \overset{\perp}{\oplus} H_2 \overset{\perp}{\oplus} \ldots \overset{\perp}{\oplus} H_d \overset{\perp}{\oplus} W$.*

L'entier d est uniquement déterminé ; c'est plus précisément la dimension maximale d'un sous-espace totalement isotrope de V, et tout sous-espace totalement isotrope de V est contenu dans un sous-espace totalement isotrope et de dimension d de V.

L'espace W est uniquement déterminé à isométrie près. On l'appelle la partie anisotrope *de V ; si $W = \{0\}$ on dit que V est* hyperbolique.

Démonstration. — Commençons par une remarque générale. Soit U un sous-espace totalement isotrope de V et soit r sa dimension. En appliquant la proposition 13.5.3 à une base arbitraire $(v_1, \ldots, v_r)$ de U, on obtient l'existence de r plans hyperboliques $H_1, \ldots, H_r$ dans V, deux à deux orthogonaux, tels que $U \subset H_1 \overset{\perp}{\oplus} \ldots \overset{\perp}{\oplus} H_r$. Par ailleurs comme les H_i sont non dégénérés il en va de même de leur somme directe orthogonale, et si l'on note F l'orthogonal de $H_1 \overset{\perp}{\oplus} \ldots \overset{\perp}{\oplus} H_r$ on a donc

$$V = H_1 \overset{\perp}{\oplus} \ldots \overset{\perp}{\oplus} H_r \overset{\perp}{\oplus} F.$$

Montrons maintenant le théorème. Commençons par l'existence de la décomposition annoncée. Soit d la dimension maximale d'un sous-espace totalement isotrope de V (cet entier est bien défini puisqu'il y a au moins un sous-espace totalement isotrope de V, à savoir $\{0\}$), et soit U un sous-espace totalement isotrope de V de dimension d. En vertu de la remarque faite en début de preuve, il existe d plans hyperboliques $H_1, \ldots, H_d$ dans V, et un sous-espace W de V tels que $V = H_1 \overset{\perp}{\oplus} \ldots \overset{\perp}{\oplus} H_d \overset{\perp}{\oplus} W$ et $U \subset \overset{\perp}{\bigoplus} H_i$. Il suffit pour conclure la partie du théorème relative à l'existence de montrer que W est anisotrope. Or si w est un vecteur isotrope de W, il est orthogonal à l'espace totalement isotrope U si bien que $U \oplus kw$ est totalement isotrope. Par maximalité de U il vient $U \oplus kw = U$ et $w = 0$; ainsi, W est anisotrope.

Venons-en aux assertions relatives à l'unicité. Supposons donc donnée une décomposition $V = H_1 \overset{\perp}{\oplus} H_2 \overset{\perp}{\oplus} \ldots \overset{\perp}{\oplus} H_d \overset{\perp}{\oplus} W$ comme dans l'énoncé. Nous allons montrer que tout sous-espace totalement isotrope de V est contenu dans un sous-espace totalement isotrope de dimension d (ce qui entraînera formellement que d est la dimension maximale d'un sous-espace totalement anisotrope de V puisqu'il y a comme on l'a déjà remarqué au moins un tel sous-espace, à savoir $\{0\}$).

Soit U un sous-espace totalement isotrope de V et soit r sa dimension. D'après la remarque faite en début de preuve, on peut écrire $V = G_1 \overset{\perp}{\oplus} G_2 \overset{\perp}{\oplus} \dots \overset{\perp}{\oplus} G_r \overset{\perp}{\oplus} F$ où les G_i sont des plans hyperboliques, où $U \subset \overset{\perp}{\bigoplus} G_i$, et où F est un certain sous-espace vectoriel de V. Si l'on avait $r > d$ on pourrait écrire

$$V = H_1 \overset{\perp}{\oplus} H_2 \overset{\perp}{\oplus} \dots \overset{\perp}{\oplus} H_d \overset{\perp}{\oplus} W = G_1 \overset{\perp}{\oplus} G_2 \overset{\perp}{\oplus} \dots \overset{\perp}{\oplus} G_d \overset{\perp}{\oplus} (G_{d+1} \overset{\perp}{\oplus} \dots \overset{\perp}{\oplus} G_r \overset{\perp}{\oplus} F)$$

et comme les G_i et les H_i sont tous des plans hyperboliques, le théorème de simplification de Witt (13.5.6) assure qu'il existe une isométrie

$$f\colon W \simeq G_{d+1} \overset{\perp}{\oplus} \dots \overset{\perp}{\oplus} G_r \overset{\perp}{\oplus} F$$

en particulier, W contient le plan hyperbolique $f^{-1}(G_{d+1})$, contredisant son caractère anisotrope.

On voit donc que $r \leqslant d$. On peut dès lors écrire

$$V = H_1 \overset{\perp}{\oplus} H_2 \overset{\perp}{\oplus} \dots \overset{\perp}{\oplus} H_r \overset{\perp}{\oplus} (H_{r+1} \overset{\perp}{\oplus} \dots \overset{\perp}{\oplus} H_d \overset{\perp}{\oplus} W) = G_1 \overset{\perp}{\oplus} G_2 \overset{\perp}{\oplus} \dots \overset{\perp}{\oplus} G_r \overset{\perp}{\oplus} F.$$

Comme les G_i et les H_i sont tous des plans hyperboliques, le théorème de simplification de Witt assure qu'il existe une isométrie

$$g\colon H_{r+1} \overset{\perp}{\oplus} \dots \overset{\perp}{\oplus} H_d \overset{\perp}{\oplus} W \simeq F.$$

Pour tout i compris entre $r+1$ et d, choisissons un vecteur isotrope non nul v_i dans le plan hyperbolique H_i. Le sous-espace de V engendré par U et les $g(v_i)$ est alors de dimension d, et il est totalement isotrope car U est totalement isotrope et car les $g(v_i)$ sont isotropes, deux à deux orthogonaux et orthogonaux à U ; de plus, il contient U par construction.

On a donc établi l'unicité de d (à travers sa caractérisation annoncée). Il reste à s'assurer de celle de W à isométrie près. Mais une fois que l'unicité de d est connue, c'est une conséquence immédiate du théorème de simplification de Witt. $\square$

Nous allons terminer ce chapitre par l'introduction d'un invariant très utile des formes quadratiques non dégénérées.

Lemme-Définition 13.5.8. — *Soit (V,q) un k-espace quadratique non dégénérée, soit $\mathscr{B}$ une base de V et soit M la matrice de q dans $\mathscr{B}$. L'image de* $\det M$ *dans $k^\times/(k^\times)^2$ ne dépend pas de $\mathscr{B}$, et est appelée le* discriminant *de q (ou de V, ou de (V,q)).*

Démonstration. — Soit $\mathscr{C}$ une (autre) base de V et soit P la matrice de passage de $\mathscr{B}$ à $\mathscr{C}$. La matrice de q dans la base $\mathscr{C}$ est alors égale à ${}^tPMP$, dont le déterminant vaut $(\det P)^2 \det M$. $\square$

Exemple 13.5.9. — Si $(a_1, \ldots, a_n)$ est une famille finie d'éléments de $k^\times$, le discriminant de l'espace non dégénéré $\langle a_1, \ldots, a_n \rangle$ est égale à $a_1 \cdot \ldots \cdot a_n$.

Pour calculer en pratique le discriminant d'un espace quadratique (V, q) non dégénéré quelconque, le plus simple est en général de diagonaliser q puis d'appliquer la formule précédente.

Remarque 13.5.10. — On a vu à l'exemple 13.4.12 que $a \mapsto \langle a \rangle$ établit une bijection entre $k^\times/(k^\times)^2$ et l'ensemble des classes d'isométries de droites quadratiques non dégénérées. La bijection réciproque envoie une telle droite sur son discriminant : cela découle du fait que pour tout $a \in k^\times$, le discriminant de $\langle a \rangle$ est égal à a.

Ainsi, la classe d'isométrie d'une k-droite quadratique non dégénérée est entièrement déterminée par son discriminant.

Remarque 13.5.11. — Pour tout entier n strictement positif et tout élément a de $k^\times$ il existe un k-espace quadratique non dégénéré de discriminant a : il suffit en effet de montrer l'existence d'une matrice symétrique appartenant à $\mathrm{M}_n(k)$ de déterminant a, et on peut par exemple prendre $\mathrm{Diag}(1, \ldots, 1, a)$. (Il n'y a par contre qu'une seule forme quadratique sur $\{0\}$, la forme nulle, qui est de discriminant est 1 puisque la matrice vide est de déterminant 1 – essayez de comprendre pourquoi !)

Exemple 13.5.12. — Soit $((V_i, q_i))_i$ une famille finie de k-espaces quadratiques non dégénérés, et soit (V, q) leur somme directe orthogonale. On déduit aussitôt de la description matricielle de q à partir de celle des q_i (13.4.2.2) que le discriminant de q est égal au produit des discriminants des q_i.

13.5.13. *Discriminant et plans hyperboliques.* — Soit V un k-plan quadratique non dégénéré. Choisissons a et b dans $k^\times$ tel que V soit isométrique à $\langle a, b \rangle$. Posons $\alpha = b/a$. L'espace V est isométrique à $\langle a, a\alpha \rangle$ et son discriminant est dès lors égal à $a^2\alpha$, et donc à α (puisqu'il prend ses valeurs dans $k^\times/(k^\times)^2$).

Si V est hyperbolique il est isométrique à $\langle 1, -1 \rangle$, et son discriminant vaut donc (-1). Réciproquement, supposons que le discriminant de V est égal à (-1) ; cela signifie qu'il existe $\lambda \in k^\times$ tel que $\alpha = -\lambda^2$. Soit (e_1, e_2) une base de V dans laquelle q a pour matrice $\mathrm{Diag}(a, \alpha a)$. On a alors $q(\lambda e_1 + e_2) = \lambda + \alpha^2 = 0$ et l'espace V est donc isotrope ; c'est en conséquence un plan hyperbolique.

14. Théorèmes de classification, étude des groupes orthogonaux

On désigne toujours par k un corps de caractéristique différente de 2. Dans ce dernier chapitre, nous allons tout d'abord décrire la classification des k-espaces quadratiques à isométrie près lorsque k est algébriquement clos, lorsque $k = \mathbf{R}$, et lorsque k est fini. Puis nous passerons à l'étude détaillée du groupe orthogonal d'un espace quadratique (non dégénéré) de dimension 2, avant de démontrer le théorème de Cartan-Dieudonné (qui assure que toute isométrie d'un espace quadratique non dégénéré de dimension n est produit d'au plus n réflexions), et de terminer par quelques considérations sur le groupe orthogonal d'un espace euclidien.

Dans ce qui suit, nous emploierons le symbole d'isomorphisme $\simeq$ pour désigner la relation d'isométrie.

14.1. Classification des formes quadratiques sur un corps quadratiquement clos. — Nous allons maintenant classer à isométrie près les espaces quadratiques sur un corps *quadratiquement clos*, c'est-à-dire un corps dans lequel tout élément est un carré. L'exemple à garder en tête est évidemment celui d'un corps algébriquement clos algébriquement clos, mais il en existe d'autres. On peut par exemple définir récursivement une suite de sous-corps de $\overline{\mathbf{Q}}$ comme suit : on pose $k_0 = \mathbf{Q}$, et l'on note k_{n+1} le sous-corps de $\overline{\mathbf{Q}}$ engendré par k_n et par les racines carrées de tous les éléments de k_n ; la réunion k des k_n est alors quadratiquement close par construction, et elle n'est pas égale à $\overline{\mathbf{Q}}$ tout entier (essayez par exemple de montrer qu'elle ne contient pas $\sqrt[3]{2}$).

Proposition 14.1.1. — *Supposons k quadratiquement clos. Soit (V,q) un k-espace quadratique, soit n la dimension de V et soit r le rang de la forme q ; posons $t = n - r$. L'espace quadratique (V,q) est alors isométrique à*

$$\langle \underbrace{1,\dots,1}_{r\ termes}, \underbrace{0,\dots,0}_{t\ termes} \rangle.$$

Démonstration. — Par diagonalisation des formes quadratiques (13.3.10), on sait qu'il existe une base $(Y_1,\dots,Y_n)$ de V^* et une famille $(a_i)_{1\leqslant i\leqslant r}$ de scalaires non nuls tels que $q = \sum_{1\leqslant i\leqslant r} a_i Y_i^2$. Puisque k est quadratiquement clos, chacun des a_i s'écrit α_i^2 pour un certain $\alpha_i \in k^\times$. On a alors

$$q = \sum_{1\leqslant i\leqslant r} a_i Y_i^2 = \sum_{1\leqslant i\leqslant r} (\alpha_i Y_i)^2.$$

Posons $X_i = \alpha_i Y_i$ si $i \leqslant r$ et $X_i = Y_i$ sinon. La famille (X_i) est encore une base de V^*, et $q = \sum_{1\leqslant i\leqslant r} X_i^2$. □

Commentaires 14.1.2. — Il résulte de la proposition ci-dessus que lorsque k est quadratiquement clos, la classe d'isométrie d'un k-espace quadratique (V,q) est entièrement déterminée par la dimension de V et le rang de q (et donc simplement par la dimension de V si on le suppose de plus non dégénéré).

14.2. Classification des formes quadratiques réelles. — Nous allons maintenant aborder notre second résultat de classification, qui concerne les formes quadratiques sur $\mathbf{R}$.

Définition 14.2.1. — Soit V un $\mathbf{R}$-espace vectoriel et soit q une forme quadratique sur V. On dit que q est *positive* (resp. *négative*) si $q(v) \geqslant 0$ (resp. $q(v) \leqslant 0$) pour tout $v \in V$. On dit qu'elle est *définie positive* (resp. *définie négative*) si $q(v) > 0$ (resp. $q(v) < 0$) pour tout $v \in V \backslash \{0\}$.

Remarque 14.2.2. — Traditionnellement, une forme bilinéaire symétrique sur un espace vectoriel réel est appelée un *produit scalaire* si la forme quadratique associée est définie positive, et on appelle *espace euclidien* un $\mathbf{R}$-espace vectoriel de dimension finie muni d'un produit scalaire. Autrement dit, un espace euclidien est un $\mathbf{R}$-espace quadratique (V, q) avec q définie positive.

Remarque 14.2.3. — Soit V un espace vectoriel réel et soit q une forme quadratique sur V. Il résulte immédiatement de la définition que si q est définie positive ou définie négative, elle est anisotrope (et *a fortiori* non dégénérée si dim $< +\infty$).

Théorème 14.2.4. — *Soit (V, q) un $\mathbf{R}$-espace quadratique et soit n sa dimension. Il existe deux entiers r et s tels que*

$$V \simeq \langle \underbrace{1, \ldots, 1}_{r \textit{ termes}}, \underbrace{-1, \ldots, -1}_{s \textit{ termes}}, \underbrace{0, \ldots, 0}_{t \textit{ termes}} \rangle,$$

en posant $t = n - r - s$.

De plus, les entiers r et s sont uniquement déterminés : l'entier r est la dimension maximale d'un sous-espace W de V tel que $q|_W$ soit définie positive, et l'entier s est la dimension maximale d'un sous-espace W de V tel que $q|_W$ soit définie négative.

Démonstration. — Commençons par montrer l'existence des entiers r et s Soit d le rang de q. Par diagonalisation des formes quadratiques (13.3.10), on sait qu'il existe une base $(Y_1, \ldots, Y_n)$ de V^* et une famille $(a_i)_{1 \leqslant i \leqslant r}$ de réels non nuls tels que $q = \sum_{1 \leqslant i \leqslant d} a_i Y_i^2$. Quitte à réordonner les Y_i, on peut supposer qu'il existe r tel que a_i soit strictement positif si $i \leqslant r$ et strictement négatif sinon ; on pose $s = d - r$. Pour tout i compris entre 1 et $r + s$, il existe un réel non nul α_i tel que $a_i = \alpha_i^2$ si $i \leqslant r$, et $a_i = -\alpha_i^2$ sinon. On a alors

$$q = \sum_{1 \leqslant i \leqslant r+s} a_i Y_i^2 = \sum_{1 \leqslant i \leqslant r} (\alpha_i Y_i)^2 - \sum_{r+1 \leqslant i \leqslant r+s} (\alpha_i Y_i)^2.$$

Posons $X_i = \alpha_i Y_i$ si $i \leqslant d$ et $X_i = Y_i$ sinon. La famille (X_i) est encore une base de V^*, et $q = \sum_{1 \leqslant i \leqslant r} - \sum_{r+1 \leqslant i \leqslant r+s} X_i^2$, si bien que

$$V \simeq \langle \underbrace{1, \ldots, 1}_{r \text{ termes}}, \underbrace{-1, \ldots, -1}_{s \text{ termes}}, \underbrace{0, \ldots, 0}_{t \text{ termes}} \rangle,$$

avec $t = n - r - s$.

Montrons maintenant la caractérisation des entiers r et s fournie par l'énoncé ; soit $(e_1, \ldots, e_n)$ une base de V dans laquelle la matrice de q est égale à

$$\mathrm{Diag}\left(\underbrace{1, \ldots, 1}_{r \text{ termes}}, \underbrace{-1, \ldots, -1}_{s \text{ termes}}, \underbrace{0, \ldots, 0}_{t \text{ termes}}\right)$$

avec $t = n - r - s$. Soit R (resp. S) la plus grande dimension d'un sous-espace W de V tel que $q|_W$ soit définie positive (resp. négative) ; il s'agit de s'assurer que $R = r$ et $S = s$.

Prouvons par exemple que $r = R$; l'égalité $s = S$ se montre de manière analogue (ou se déduit de l'égalité $r = R$ appliquée à la forme quadratique $(-q)$). Remarquons tout d'abord que si $(x_1, \ldots, x_r)$ est une famille de réels alors $q(\sum_{1 \leqslant i \leqslant r} x_i e_i) = \sum x_i^2$; par conséquent $q|_{\mathrm{Vect}(e_1,\ldots,e_r)}$ est définie positive, si bien que $R \leqslant r$. Réciproquement, soit W un sous-espace vectoriel de V tel que que $q|_W$ soit définie positive. Pour toute famille $(x_{r+1}, \ldots, x_{r+s}, \ldots, x_n)$ de réels on a $q(\sum_{r+1 \leqslant i \leqslant n} x_i e_i) = -\sum_{r+1 \leqslant i \leqslant r+s} x_i^2$, si bien que $q|_{\mathrm{Vect}(e_{r+1},\ldots,e_n)}$ est négative. Par conséquent, la restriction de la forme q à l'intersection $W \cap \mathrm{Vect}(e_{r+1}, \ldots, e_n)$ est à la fois définie positive et négative, ce qui entraîne que $W \cap \mathrm{Vect}(e_{r+1}, \ldots, e_n) = \{0\}$. Il en résulte que $\dim W \leqslant n - (n - r) = r$; ainsi $R \leqslant r$, d'où finalement l'égalité $r = R$. □

Commentaires 14.2.5. — Soit (V, q) un $\mathbf{R}$-espace quadratique. Il résulte du théorème 14.2.4 que la classe d'isométrie de (V, q) est entièrement déterminée par la dimension n de V et le couple (r, s) où r (resp. s) est la dimension maximale d'un sous-espace de V en restriction auquel q est définie positive (resp. définie négative).

Le couple (r, s) est appelé la *signature* de q (ou de V, ou de (V, q)), et il résulte immédiatement de sa définition que $r + s$ est égal au rang de q, et donc à n si V est non dégénéré.

La classe d'isométrie d'un $\mathbf{R}$-espace quadratique non dégénéré est donc entièrement déterminé par sa signature.

Remarque 14.2.6. — Soit (V, q) un $\mathbf{R}$-espace quadratique et soit $(X_1, \ldots, X_n)$ une base de V^* telle que q soit de la forme $\sum_i a_i X_i^2$ pour une certaine famille (a_i) de nombres réels. Soit (r, s) la signature de q. La preuve du théorème 14.2.4 montre alors que r (resp. s) est égal au nombre d'indices i tels que a_i soit strictement positif (resp. strictement négatif).

14.2.7. — Mentionnons maintenant quelques types de $\mathbf{R}$-espaces quadratiques non dégénérés qui jouent un rôle particulièrement importants.

14.2.7.1. — Il y a tout d'abord ceux dont la signature est de la forme $(n, 0)$, c'est-à-dire ceux dont la forme quadratique associée est définie positive : ce sont les espaces euclidiens.

14.2.7.2. — Il y a ensuite les espaces *lorentziens*, dont la signature est de la forme $(n, 1)$. La motivation initiale pour accorder un intérêt spécifique à ce type d'espaces est la théorie de la relativité : l'espace-temps est en effet modélisé (au moins localement) par un $\mathbf{R}$-espace vectoriel V de dimension 4 muni de la forme quadratique de signature

$(3,1)$ s'écrivant $X^2 + Y^2 + Z^2 - c^2T^2$ dans une base convenable (X, Y, Z, T) de V^* (où c est la vitesse de la lumière).

14.2.7.3. — En géométrie algébrique, l'étude des surfaces algébriques projectives fait intervenir de manière naturelle, *via* ce qu'on appelle les *formes d'intersection*, des $\mathbf{R}$-espaces quadratiques non dégénérés de signature est de la forme $(1, n)$.

14.3. Classifications des formes quadratiques sur un corps fini. — On suppose que k est fini, et l'on note q son cardinal. Si p désigne la caractéristique de k, le corps $\mathbf{F}_p$ se plonge canoniquement dans k, ce qui fait de celui-ci un $\mathbf{F}_p$-espace vectoriel ; comme k est fini, il est de dimension finie sur $\mathbf{F}_p$; si l'on pose $n = \dim_{\mathbf{F}_p} k$ on a alors $k \simeq \mathbf{F}_p^n$, et partant $q = p^n$. Notons que $n > 0$ (comme $1 \neq 0$ dans un corps, son cardinal vaut toujours au moins 2).

14.3.1. *Les carrés de $k^\times$.* — Le groupe $k^\times$ est cyclique (3.8.6) de cardinal $q - 1$. Comme k est par hypothèse de caractéristique différente de 2, le nombre premier p est impair, et $q-1$ est donc pair. L'élévation au carré est un morphisme de groupes de $k^\times$ dans lui-même, dont le noyau $\{-1, 1\}$ est de cardinal 2 (notez que $1 \neq -1$ puisque la caractéristique de k est différente de 2). Le groupe $(k^\times)^2$, qui est l'image de ce morphisme, est en conséquence de cardinal $(q - 1)/2$. Il s'ensuit que c'est l'unique sous-groupe d'indice 2 de $k^\times$, et qu'il coïncide avec l'ensemble des éléments de $k^\times$ de $(q - 1)/2$-torsion. Or si λ est un élément de $k^\times$, on a $\lambda^{q-1} = 1$, c'est-à-dire encore $(\lambda^{(q-1)/2})^2 = 1$; on en déduit que $\lambda^{(q-1)/2} \in \{-1, 1\}$.

L'élévation à la puissance $(q - 1)/2$ est donc un morphisme de groupes de $k^\times$ dans $\{-1, 1\}$, dont le noyau est le sous-groupe $(k^\times)^2$ de $k^\times$, qui est d'indice 2. Par conséquent, $\lambda \mapsto \lambda^{(q-1)/2}$ induit un isomorphisme $k^\times/(k^\times)^2 \simeq \{-1, 1\}$.

Proposition 14.3.2. — *Soit V un k-espace vectoriel de dimension au moins* 3 *et soit q une forme quadratique sur V. La forme q est isotrope.*

Démonstration. — Quitte à remplacer V par n'importe lequel de ses sous-espace vectoriels de dimension 3, on peut supposer qu'il est de dimension 3, puis l'identifier à $\langle a, b, c\rangle$ pour un certain triplet (a, b, c) d'éléments de k ; autrement dit on s'est ramené au cas où $V = k^3$ et où $q = (x, y, z) \mapsto ax^2 + by^2 + cz^2$.

Si l'un des scalaires a, b ou c est nul, l'isotropie de q est claire : par exemple si $a = 0$ on a $q(1, 0, 0) = 0$, et on raisonne de façon analogue si $b = 0$ ou si $c = 0$. Supposons maintenant que a, b et c sont tous trois non nuls. L'application $\lambda \mapsto a + b\lambda$ est alors une bijection de k sur lui-même. Le cardinal de $E := \{a + by^2\}_{y\in k}$ est donc égal au cardinal de $\{y^2\}_{y\in k} = \{0\} \cup (k^\times)^2$, c'est-à-dire à $1 + (q - 1)/2 = (q + 1)/2$. L'application $\lambda \mapsto -c\lambda$ est encore une bijection de k sur lui-même ; par conséquent, le cardinal de $F := \{-cz^2\}_{z\in k}$ est égal à celui de $\{z^2\}_{z\in k}$, c'est-à-dire à $(q + 1)/2$.

La somme $|E| + |F|$ est égale à $q + 1 > q$, ce qui entraîne que $E \cap F \neq \varnothing$. Il existe donc y et z dans k tels que $a + by^2 = -cz^2$, soit encore $a + by^2 + cz^2 = 0$. Par conséquent $q(1, y, z) = 0$ et q est isotrope. □

Corollaire 14.3.3. — *Soit (V, q) un k-espace quadratique non dégénéré de dimension au moins* 2. *Pour tout $\lambda \in k$ il existe $v \in V$ tel que $q(v) = \lambda$.*

Démonstration. — On sait que c'est vrai si V est isotrope (corollaire 13.5.4), ce qui est notamment le cas si $\dim V \geqslant 3$ (proposition 14.3.2) ; on peut donc supposer que V est de dimension 2 et anisotrope. Le choix d'une base de diagonalisation de q permet de se ramener au cas où (V, q) est égal $\langle a, b\rangle$ pour un certain couple (a, b) d'éléments de k.

En vertu de la proposition 14.3.2, le k-espace quadratique $\langle a, b, -\lambda\rangle$ est isotrope. Il existe donc $(x, y, z) \neq (0, 0, 0)$ tel que $ax^2 + by^2 - \lambda z^2 = 0$. On a nécessairement $z \neq 0$ car sinon on aurait $(x, y) \neq (0, 0)$ et $ax^2 + by^2 = 0$, contredisant l'anisotropie de q. Comme $z \neq 0$ on peut écrire $a(x/z)^2 + b(y/z)^2 = \lambda$, c'est-à-dire $q(x/z, y/z) = \lambda$. □

Proposition 14.3.4. — *Soit (V, q) un k-espace quadratique non dégénéré de dimension $n \geqslant 1$, et soit a un élément de $k^\times$ tel que le discriminant de q soit égal à a modulo $(k^\times)^2$. L'espace quadratique (V, q) est alors isométrique à $\langle 1, \ldots, 1, a\rangle$.*

Démonstration. — On raisonne par récurrence sur n. Si $n = 1$, il découle de la remarque 13.5.10 que V est isométrique à $\langle a\rangle$ (cela n'utilise pas le fait que k est un corps fini), et la proposition est donc vraie dans ce cas.

Supposons maintenant $n \geqslant 2$ et le résultat vrai en dimension $n - 1$. Comme V est non dégénéré et de dimension $\geqslant 2$, l'application $q\colon V \to k$ est surjective (corollaire 14.3.3) ; en particulier, il existe $e_1 \in V$ tel que $q(e_1) = 1$, ce qui entraîne que ke_1 est isométrique à $\langle 1\rangle$ (et que son discriminant est égal à 1). Puisque e_1 est anisotrope, on a $V = ke_1 \oplus (ke_1)^\perp$; le sous-espace $(ke_1)^\perp$ est alors non dégénéré, et le discriminant de $q|_{(ke_1)^\perp}$ est en vertu de la remarque 13.5.12 encore égal à a, puisque le discriminant de $q|_{ke_1}$ vaut 1.

Par hypothèse de récurrence appliquée à la restriction de q à $(ke_1)^\perp$, l'espace quadratique $(ke_1)^\perp$ est isométrique à $\langle \underbrace{1, \ldots, 1, a}_{n-1 \text{ termes}}\rangle$. Par conséquent

$$V = ke_1 \oplus (ke_1)^\perp \simeq \langle 1\rangle \overset{\perp}{\oplus} \langle \underbrace{1, \ldots, 1, a}_{n-1 \text{ termes}}\rangle = \langle \underbrace{1, \ldots, 1, a}_{n \text{ termes}}\rangle.$$

□

14.3.5. — Rappelons que $k^\times/(k^\times)^2$ est de cardinal 2 (*cf.* 14.3.1). Il résulte dès lors de la proposition précédente qu'il y a pour tout entier $n > 0$ exactement deux classes de k-espaces quadratiques non dégénérés de dimension n, chacune d'elles étant entièrement déterminée par son discriminant. (En dimension nulle il y a par contre une seule classe d'isométrie de k-espace quadratique non dégénérée, et le discriminant correspondant vaut 1, *cf.* remarque 13.5.11).

14.4. Parties anisotropes. — Nous venons de décrire les classes d'isométrie de k-espaces quadratiques dans différentes situations. Dans chacune d'elle nous allons maintenant décrire, un k-espace quadratique non dégénéré étant donné, sa «partie anisotrope» fournie par le théorème 13.5.7, qui est bien déterminée à isométrie près.

14.4.1. *Le cas où k est quadratiquement clos.* — Soit n un entier positif ou nul. Il y a d'après la proposition 14.1.1 une seule classe d'isométrie de k-espaces quadratiques non dégénérés de dimension n. Soit (V, q) un tel espace.

14.4.1.1. *Supposons que n est de la forme $2d$.* — Comme $\bigoplus_{1\leqslant i\leqslant d} \langle 1, -1\rangle$ est non dégénéré et de dimension $2d$, il est isométrique à V. Par conséquent, V est hyperbolique.

14.4.1.2. *Supposons que n est de la forme $2d$.* — Comme $\left(\overset{\perp}{\bigoplus_{1\leqslant i\leqslant d}} \langle 1, -1\rangle\right) \overset{\perp}{\oplus} \langle 1\rangle$ est non dégénéré et de dimension $2d+1$, il est isométrique à V. Par conséquent, la partie anisotrope de V est $\langle 1\rangle$.

14.4.2. *Le cas où $k = \mathbf{R}$.* — Soit n un entier positif ou nul, soit (V, q) un $\mathbf{R}$-espace quadratique non dégénéré de dimension n, et soit (r, s) sa signature. Par définition, $r + s = n$ et

$$V \simeq \langle \underbrace{1, \ldots, 1}_{r \text{ termes}}, \underbrace{-1, \ldots, -1}_{s \text{ termes}} \rangle.$$

14.4.2.1. *Supposons que $r \geqslant s$.* — En posant $e = r - s$, on peut écrire

$$V \simeq \langle \underbrace{1, -1, 1, -1, \ldots, 1, -1}_{s \text{ blocs } 1,-1}, \underbrace{1, \ldots, 1,}_{e \text{ termes}} \rangle.$$

Autrement dit,

$$V \simeq \left(\overset{\perp}{\bigoplus_{1\leqslant i\leqslant r}} \langle 1, -1\rangle\right) \overset{\perp}{\oplus} \langle \underbrace{1, \ldots, 1}_{e \text{ termes}} \rangle$$

et sa partie anisotrope est donc le $\mathbf{R}$-espace quadratique $\langle 1, \ldots, 1\rangle$ de dimension e (qui est un espace euclidien).

14.4.2.2. *Supposons que $r \leqslant s$.* — En posant $e = r - s$, on peut écrire

$$V \simeq \langle \underbrace{1, -1, 1, -1, \ldots, 1, -1}_{r \text{ blocs } 1,-1}, \underbrace{-1, \ldots, -1,}_{e \text{ termes}} \rangle.$$

Autrement dit,

$$V \simeq \left(\overset{\perp}{\bigoplus_{1\leqslant i\leqslant s}} \langle 1, -1\rangle\right) \overset{\perp}{\oplus} \langle \underbrace{-1, \ldots, -1}_{e \text{ termes}} \rangle$$

et sa partie anisotrope est donc le $\mathbf{R}$-espace quadratique $\langle -1, \ldots, -1\rangle$ de dimension e (qui est si l'on peut dire «anti-euclidien», c'est-à-dire défini négatif).

14.4.2.3. *Le cas $r = s$.* — C'est l'intersection des deux cas précédents, correspondant à chaque fois à la situation limite dans laquelle $e = 0$; l'espace V est alors hyperbolique.

14.4.3. *Le cas où k est fini.* — Supposons maintenant k fini. Soit n un entier strictement positif, et soit a un élément de $k^\times$. Il y a d'après la proposition 14.3.4 une unique classe d'isométrie de k-espaces quadratiques non dégénérés de dimension n et de discriminant a. Soit (V, q) un tel espace.

14.4.3.1. *Supposons que n est de la forme $2d+1$.* — L'espace quadratique

$$\bigoplus_{1\leqslant i\leqslant d}\langle 1,-1\rangle \overset{\perp}{\oplus} \langle (-1)^d a\rangle$$

est de dimension n, non dégénéré et de discriminant $(-1)^{2d}a = a$; par conséquent, il est isométrique à V, si bien que la partie anisotrope de V est à $\langle (-1)^d a\rangle$.

14.4.3.2. *Supposons que n est de la forme $2d$ et que $a = (-1)^d$ modulo $(k^\times)^2$.* — L'espace quadratique

$$\bigoplus_{1\leqslant i\leqslant d}\langle 1,-1\rangle$$

est de dimension n, non dégénéré et de discriminant $(-1)^d = a \bmod (k^\times)^2$; par conséquent, il est isométrique à V, si bien que V est hyperbolique.

14.4.3.3. *Supposons que n est de la forme $2d$ et que $a \neq (-1)^d$ modulo $(k^\times)^2$.* — L'espace quadratique

$$\bigoplus_{1\leqslant i\leqslant d-1}\langle 1,-1\rangle \overset{\perp}{\oplus} \langle 1,(-1)^{d-1}a\rangle$$

est de dimension n, non dégénéré et de discriminant $(-1)^{2(d-1)} = a$; par conséquent, il est isométrique à V. Par ailleurs, le fait que $a \neq (-1)^d$ modulo $(k^\times)^2$ entraîne que $(-1)^{d-1}a \neq (-1)$ modulo $(k^\times)^2$. Le k-plan quadratique $\langle 1,(-1)^{d-1}a\rangle$ n'est donc pas un plan hyperbolique (puisque le discriminant d'un plan hyperbolique vaut -1) ; il est en conséquence anisotrope, et la partie anisotrope de V est dès lors $\langle 1,(-1)^{d-1}a\rangle$.

14.5. Le cas d'un corps fini : exemples concrets. — Nous allons décliner les résultats précédents (proposition 14.3.4 et 14.4.3 *et sq.*) concernant les corps finis dans différents cas explicites.

14.5.1. *Le cas du corps* $\mathbf{F}_3$. — On a alors $\mathbf{F}_3^\times = \{-1,1\}$, et $(\mathbf{F}_3^\times)^2 = \{1\}$.

14.5.1.1. — Il y a à isométrie près deux $\mathbf{F}_3$-espaces quadratiques de dimension 3 : celui de discriminant 1, à savoir $\langle 1,1,1\rangle$, de partie anisotrope anisotrope $\langle -1\rangle$; et celui de discriminant (-1), à savoir $\langle 1,1,-1\rangle$, de partie anisotrope $\langle 1\rangle$.

14.5.1.2. — Il y a à isométrie près deux $\mathbf{F}_3$-espaces quadratiques de dimension 4 : celui de discriminant 1, à savoir $\langle 1,1,1,1\rangle$, qui est hyperbolique ; et celui de discriminant (-1), à savoir $\langle 1,1,1,-1\rangle$, de partie anisotrope $\langle 1,1\rangle$.

14.5.2. *Le cas du corps* $\mathbf{F}_5$. — On a alors $\mathbf{F}_5^\times = \{-2,-1,1,2\}$, et $(\mathbf{F}_5^\times)^2 = \{1,-1\}$.

14.5.2.1. — Il y a à isométrie près deux $\mathbf{F}_5$-espaces quadratiques de dimension 3 : celui de discriminant 1, à savoir $\langle 1,1,1\rangle$, de partie anisotrope $\langle -1\rangle \simeq \langle 1\rangle$ et celui de discriminant 2, à savoir $\langle 1,1,2\rangle$, de partie anisotrope $\langle -2\rangle \simeq \langle 2\rangle$ (rappelons que (-1) est un carré).

14.5.2.2. — Il y a à isométrie près deux $\mathbf{F}_5$-espaces quadratiques de dimension 4 : celui de discriminant 1, à savoir $\langle 1,1,1,1\rangle$, qui est hyperbolique ; et celui de discriminant 2, à savoir $\langle 1,1,1,2\rangle$, de partie anisotrope $\langle 1,-2\rangle \simeq \langle 1,2\rangle$ (rappelons que (-1) est un carré).

14.6. Le groupe orthogonal en toute petite dimension. — Nous allons brièvement décrire le groupe orthogonal d'un k-espace quadratique de dimension 0 ou 1.

14.6.1. *Le cas de la dimension nulle.* — Il y a évidemment une seule forme quadratique q sur l'espace nul, à savoir la forme nulle. Elle est non dégénérée (son noyau est trivial !), et l'on a

$$\mathrm{O}(q) = \mathrm{SO}(q) = \mathrm{GL}(\{0\}) = \{\mathrm{Id}\}.$$

14.6.2. *Le cas de la dimension 1.* — Soit (D, q) une droite quadratique non dégénérée. Le groupe $\mathrm{GL}(D)$ s'identifie canoniquement à $k^\times$ (tout endomorphisme de la droite vectorielle D est une homothétie).

Soit $\lambda \in k^\times$. L'homothétie de rapport λ appartient à $\mathrm{O}(q)$ si et seulement si $q(\lambda x) = q(x)$ pour tout $x \in D$, c'est-à-dire encore $\lambda^2 q(x) = q(x)$ pour tout $x \in D$. La forme q est non dégénérée, c'est-dire non identiquement nulle. Par conséquent, l'homothétie de rapport λ appartient à $\mathrm{O}(q)$ si et seulement si $\lambda^2 = 1$, soit encore si et seulement si $\lambda \in \{-1, 1\}$. Il vient

$$\mathrm{O}(q) = \{\mathrm{Id}, -\mathrm{Id}\} \text{ et } \mathrm{SO}(q) = \{\mathrm{Id}\}.$$

Remarque 14.6.3. — Si V est un k-espace quadratique non dégénéré de dimension $\leqslant 1$, alors toute isométrie de V est produit de réflexions : c'est en effet évident pour l'identité, qui est le *produit vide* de réflexions. Il reste à considérer le cas de $-\mathrm{Id}$ lorsque V est une droite ; or dans ce cas, $-\mathrm{Id}$ est la réflexion par rapport au sous-espace nul, qui est bien un hyperplan de V.

14.7. Le groupe orthogonal en dimension 2. — On se propose maintenant de s'intéresser en détail au groupe orthogonal d'un k-plan quadratique non dégénéré V ; comme vous le verrez, notre approche évoquera la méthode classique d'utilisation des nombres complexes pour l'étude des isométries d'un plan euclidien.

14.7.1. — On fixe un k-plan quadratique non dégénéré (V, q) ; notre but est de décrire les groupes $\mathrm{SO}(q)$ et $\mathrm{O}(q)$. Choisissons une isométrie entre V et $\langle a_1, a_2\rangle$ où a_1 et a_2 sont deux éléments convenables de $k^\times$. Il est immédiat que pour tout $\lambda \in k^\times$, un automorphisme f de V est une isométrie pour q si et seulement si c'est une isométrie pour λq. On peut par conséquent supposer, quitte à remplacer q par $a_1^{-1}q$, que V est isométrique à $\langle 1, a\rangle$ pour un certain $a \in k^\times$; en fait, il sera plus commode de remplacer a par $-a$, c'est-à-dire de supposer que V est isométrique à $\langle 1, -a\rangle$.

Rappelons que V est hyperbolique (ou isotrope, ce qui revient au même) si et seulement si son discriminant est égal à -1 (13.5.13) ; comme son discriminant est égal à $-a$ on voit que V est hyperbolique (ou isotrope) si et seulement si a est un carré dans k.

14.7.2. — Soit L la k-algèbre quotient $k[T]/T^2 - a$. Notons ω l'élément $\overline{T}$ de L. La théorie de la division euclidienne montre immédiatement que $(1, \omega)$ est une base de L comme k-espace vectoriel ; lorsque nous évoquerons l'écriture d'un élément de L sous la forme $x + y\omega$, il sera toujours sous-entendu que x et y appartiennent à k.

14.7.2.1. — Par construction, on a $\omega^2 = a$, et donc $(-\omega)^2 = a$. La propriété universelle du quotient assure alors l'existence d'un unique morphisme de k-algèbres de $L = k[T]/T^2-a$ dans L qui envoie $\overline{T} = \omega$ sur $-\omega$. L'image d'un élément $z = x+y\omega$ de L sous ce k-morphisme est égale à $x - y\omega$; nous la noterons $\overline{z}$. Il est immédiat que la «conjugaison» $z \mapsto \overline{z}$ est une involution (c'est-à-dire un automorphisme égal à son propre inverse, ou encore de carré égal à l'identité).

Soit $z = x + y\omega$ un élément de L. On a

$$z\overline{z} = (x + y\omega)(x - y\omega) = x^2 - \omega^2 y^2 = x^2 - ay^2 \in k \subset L$$

Ainsi, $z \mapsto z\overline{z}$ peut être vue comme une forme quadratique sur le k-espace vectoriel L, et $(L, z \mapsto \overline{z})$ est isométrique à $\langle 1, -a\rangle$, comme notre espace V de départ. On peut donc supposer que $V = L$ et que $q = (z \mapsto z\overline{z})$; nous n'emploierons plus désormais la notation V.

L'application $z \mapsto q(z) = z\overline{z}$ commute au produit et envoie 1 sur 1 ; elle induit donc un morphisme de groupes de $L^\times$ vers $k^\times$, dont on note U le noyau. Un élément z de L appartient à U si et seulement si $q(z) = 1$, sans avoir besoin de vérifier *a priori* que z est inversible, car c'est une conséquence automatique de l'égalité $z\overline{z} = 1$. Remarquons que $U \cap k^\times$ est formé des éléments u de k tels que $u^2 = 1$; par conséquent, l'intersection $U \cap k^\times$ est égale à $\{-1, 1\}$.

14.7.3. — On a vu plus haut en 14.7.1 que le fait que a soit ou non un carré dans k a une influence importante sur le comportement de q : celle-ci est isotrope si $a \in (k^\times)^2$, et anisotrope dans le cas contraire. Nous allons voir maintenant que cela a également une grande influence sur la structure de L.

14.7.3.1. *Supposons que a n'est pas un carré dans k, c'est-à-dire que la forme q est anisotrope.* — Le polynôme T^2-a est alors irréductible dans l'anneau principal $k[T]$, ce qui entraîne que L est un corps (en effet L est non nul puisque de dimension 2, et si P est un polynôme de $k[T]$ qui n'est pas nul modulo $T^2 - a$ alors P est premier à $T^2 - a$ et il existe donc une relation de Bézout $RP + Q(T^2 - a) = 1$, si bien que $\overline{P}$ est inversible dans L, d'inverse $\overline{R}$).

14.7.3.2. *Supposons que a est un carré dans k, c'est-à-dire que q est isotrope.* — Écrivons $a = \alpha^2$ avec $\alpha \in k^\times$. On a alors $T^2 - a = (T - \alpha)(T + \alpha)$. Comme la caractéristique de k est différente de 2, les éléments α et $-\alpha$ de $k^\times$ sont distincts et le lemme chinois fournit un isomorphisme de k-algèbres

$$L = k[T]/T^2 - a \simeq k[T]/(T - \alpha) \times k[T]/(T + \alpha) \simeq k \times k$$

(la structure de k-algèbre sur $k\times k$ est donnée par le plongement diagonal $\lambda \mapsto (\lambda, \lambda)$). Le premier de ces isomorphismes envoie une classe $\overline{P}$ modulo $T^2 - a$ sur le couple formé des classes de P modulo $T - \alpha$ et modulo $T + \alpha$; le second envoie un couple de classes $(\overline{Q}, \overline{R})$ sur $(Q(\alpha), R(-\alpha))$.

Cet isomorphisme $L \simeq k \times k$ envoie en particulier $\omega = \overline{T}$ sur $(\alpha, -\alpha)$; il envoie donc un élément $z = x + y\omega$ sur $(x + y\alpha, x - y\alpha)$, et envoie dès lors $\overline{z} = x - y\omega$ sur $(x-y\alpha, x+y\alpha)$. Autrement dit, modulo l'isomorphisme $L \simeq k\times k$, la conjugaison $z \mapsto \overline{z}$ correspond simplement à l'automorphisme d'échange des facteurs $(\lambda, \mu) \mapsto (\mu, \lambda)$; et la forme quadratique q correspond donc à

$$(\lambda, \mu) \mapsto (\lambda, \mu)(\mu, \lambda) = (\lambda\mu, \lambda\mu) = \lambda\mu$$

lorsqu'on voit k comme sous-anneau de L. Son cône isotrope est la réunion des deux droites $k(1,0)$ et $k(0,1)$.

Le groupe U s'identifie quant à lui à l'ensemble des couples (λ, μ) appartenant à $k^\times \times k^\times$ tels que $(\lambda, \mu)(\mu, \lambda) = (1,1)$, c'est-à-dire encore tels que $\lambda\mu = 1$. Autrement dit, $\lambda \mapsto (\lambda, \lambda^{-1})$ définit un isomorphisme de $k^\times$ sur U.

14.7.4. — Nous allons maintenant donner deux exemples importants d'isométries de (L, q).

14.7.4.1. — Soit $u \in U$. Notons h_u l'homothétie de rapport u de L dans lui-même. C'est un automorphisme k-linéaire (et même L-linéaire) de L. On a pour tout $z \in L$ les égalités

$$q(h_u(z)) = q(uz) = \overline{uz}uz = \underbrace{\overline{u}u\overline{z}z = \overline{z}z}_{u \in \mathrm{U}} = q(z),$$

et h_u appartient donc à $\mathrm{O}(q)$.

Écrivons $u = x + y\omega$. La matrice de h_u dans la base $(1, \omega)$ de L est alors

$$\begin{pmatrix} x & ay \\ y & x \end{pmatrix},$$

et son déterminant est $x^2 - ay^2 = q(u) = 1$; par conséquent, $h_u \in \mathrm{SO}(q)$.

Comme $h_{uv} = h_u \circ h_v$ pour tout $(u, v) \in \mathrm{U}^2$, l'application $u \mapsto h_u$ définit un morphisme de groupes de U dans $\mathrm{SO}(q)$.

14.7.4.2. — Notons σ la conjugaison $z \mapsto \overline{z}$. On a pour tout $z \in L$ les égalités

$$q(\sigma(z)) = q(\overline{z}) = \overline{z}z = z\overline{z} = q(z),$$

et σ est donc une isométrie. La matrice de σ dans la base $(1, \omega)$ de L est alors

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

et son déterminant est donc -1 ; par conséquent, σ est une isométrie indirecte.

On en déduit au vu de 14.7.4.1 que $h_u \circ \sigma = (z \mapsto u\overline{z})$ est pour tout $u \in \mathrm{U}$ une isométrie indirecte.

Proposition 14.7.5. — *Soit f une isométrie de L.*

(1) *Si f est directe, il existe un unique $u \in \mathrm{U}$ tel que $f = h_u$; autrement dit, $u \mapsto h_u$ définit un isomorphisme de U sur $\mathrm{SO}(q)$.*

(2) *Si f est indirecte, il existe un unique $u \in \mathrm{U}$ tel que $f = h_u \circ \sigma$.*

Démonstration. — L'unicité de u est claire dans les deux cas : il est nécessairement égal à $f(1)$. Réciproquement, posons $u = f(1)$. Comme f est une isométrie on a $q(u) = q(1) = 1$, et u appartient donc à U. Nous allons montrer que $f = h_u$ si f est directe, et que $f = h_u \circ \sigma$ si f est indirecte.

La matrice de q dans la base $(1, \omega)$ est diagonale ; par conséquent, ω est orthogonal à 1. Puisque f est une isométrie, $f(\omega)$ est orthogonal à u. Par ailleurs, comme h_u est une isométrie, $u\omega = h_u(\omega)$ est orthogonal à $u = h_u(1)$. Puisque q est non dégénérée, l'orthogonal de ku est de dimension 1 ; comme on vient de voir qu'il contient $u\omega$, il coïncide avec $ku\omega$.

Puisque $f(\omega)$ est orthogonal à u, il existe en vertu de ce qui précède un scalaire $\lambda \in k$ tel que $f(\omega) = \lambda u\omega$. Puisque f est une isométrie, on a

$$-a = q(\omega) = q(f(\omega)) = q(\lambda u\omega) = \lambda^2 q(u)q(\omega) = -\lambda^2 a.$$

Comme a est non nul, il vient $\lambda^2 = 1$, c'est-à-dire $\lambda = 1$ ou $\lambda = -1$.

Supposons que $\lambda = 1$. On a alors $f(1) = u$ et $f(\omega) = u\omega$. Par conséquent, f et h_u coïncident sur les deux vecteurs d'une base, ce qui entraîne que $f = h_u$, et que f est directe.

Supposons que $\lambda = -1$. On a alors $f(1) = u$ et $f(\omega) = -u\omega$. Par conséquent, f et $h_u \circ \sigma$ coïncident sur les deux vecteurs d'une base, ce qui entraîne que $f = h_u \circ \sigma$, et que f est indirecte. □

14.7.6. — La proposition 14.7.4.2 ci-dessus assure notamment que $u \mapsto h_u$ induit un isomorphisme de U sur $\mathrm{SO}(q)$. On a donc une suite exacte

$$1 \longrightarrow \mathrm{U} \xrightarrow{u \mapsto h_u} \mathrm{O}(q) \xrightarrow{\quad\det\quad} \{-1, 1\} \longrightarrow 1\ .$$

Notons par ailleurs un fait important : étant isomorphe à U, le groupe $\mathrm{SO}(q)$ est en particulier abélien.

Soit $\lambda \colon \{-1, 1\} \to \{0, 1\}$ l'application qui envoie 1 sur 0 et -1 sur 1. Comme σ est une isométrie indirecte d'ordre 2, l'application $\varepsilon \mapsto \sigma^{\lambda(\varepsilon)}$ de $\{-1, 1\}$ dans $\mathrm{O}(q)$ est une section de $\det \colon \mathrm{O}(q) \to \{-1, 1\}$. Cette section induit un isomorphisme

$$(u, \varepsilon) \mapsto h_u \circ \sigma^{\lambda(\varepsilon)}$$

de $\mathrm{U} \rtimes_\varphi \{-1, 1\}$ sur $\mathrm{O}(q)$ compatible avec la suite exacte ci-dessus, où φ est un morphisme de $\{-1, 1\}$ dans $\mathrm{Aut(U)}$ que nous allons déterminer (en particulier, on retrouve l'assertion (2) qu'on pouvait donc déduire de (1) ; mais la preuve de la proposition fournit (1) et (2) simultanément).

Soit $u \in \mathrm{U}$. On a les égalités

$$\sigma \circ h_u \circ \sigma^{-1} = \sigma \circ h_u \circ \sigma = (z \mapsto \overline{u\overline{z}}) = (z \mapsto \overline{u}z) = h_{\overline{u}} = h_{u^{-1}}$$

car $\overline{u} = u^{-1}$ puisque $u \in \mathrm{U}$. Par conséquent le morphisme φ envoie -1 sur $u \mapsto u^{-1}$ (ce dernier est bien un automorphisme de U puisque U est abélien).

14.7.7. *Droites invariantes sous les isométries directes.* — Soit $u \in \mathrm{U}$. Nous nous proposons d'étudier les valeurs propres et vecteurs propres de l'isométrie directe h_u (vue comme application k-linéaire de L dans L) en distinguant trois cas.

14.7.7.1. *Le cas où u appartient à k.* — Cela revient à demander que $u = 1$ ou $u = -1$ (14.7.2.1). Dans le premier cas $h_u = \mathrm{Id}$, son unique valeur propre est 1 et le sous-espace propre correspondant est L tout entier ; dans le second cas $h_u = -\mathrm{Id}$, son unique valeur propre est -1 et là encore, le sous-espace propre correspondant est L tout entier.

14.7.7.2. *Le cas où $u \notin k$ et où a n'est pas un carré dans k.* — On sait alors que q est anisotrope, et que L est un corps. Soit λ un élément de k. Supposons qu'il existe z non nul dans L tel que $h_u(z) = \lambda z$; on aurait alors que $(u-\lambda)z = 0$, et partant $u = \lambda$ puisque L est un corps, ce qui est absurde puisqu'on a supposé que u n'appartient pas à k.

Par conséquent, l'isométrie directe h_u n'a pas de valeur propre, c'est-à-dire encore pas de droite stable.

14.7.7.3. *Le cas où $u \notin k$ et où a est un carré dans k.* — On sait alors que q est isotrope, et qu'on dispose d'une identification des k-algèbres L et $k \times k$ modulo laquelle la conjugaison est l'échange des facteurs, q est l'application $(\lambda, \mu) \mapsto \lambda\mu$ de cône isotrope $k(1,0) \cup k(0,1)$, et U est le groupe $\{(\lambda, \lambda^{-1})\}_{\lambda \in k^\times}$.

Travaillons modulo cette identification. L'élément u est alors de la forme (λ, λ^{-1}) avec $\lambda \in k^\times$, et h_u est l'application $(x, y) \mapsto (\lambda x, \lambda^{-1} y)$. De plus $\lambda \neq \lambda^{-1}$ (ce qui veut dire que $\lambda \notin \{-1, 1\}$) puisqu'on a supposé que u n'appartient pas à k (qui est considéré comme plongé diagonalement dans $k \times k$). Il s'ensuit que h_u est diagonalisable avec deux valeurs propres distinctes λ et λ^{-1}, et que les deux droites propres correspondantes sont les droites isotropes $k(1,0)$ et $k(0,1)$.

14.7.8. *Les isométries indirectes de L sont les réflexions.* — Soit u un élément de U et soit f l'isométrie indirecte $h_u \circ \sigma = z \mapsto u\overline{z}$. On a

$$f \circ f = (z \mapsto u\overline{u\overline{z}}) = (z \mapsto u\overline{u}z) = \mathrm{Id}$$

car $u\overline{u} = 1$; par conséquent, f est une involution. Puisque f annule le polynôme $X^2 - 1 = (X-1)(X+1)$, l'endomorphisme f est diagonalisable et ses valeurs propres appartiennent à $\{-1, 1\}$. Puisque f est une isométrie indirecte, son déterminant est -1, et elle a donc pour matrice $\mathrm{Diag}(1, -1)$ dans une base convenable ; autrement dit, c'est une symétrie par rapport à une droite de L. Nous allons montrer que c'est plus précisément une symétrie orthogonale, c'est-à-dire encore ici (les droites étant de codimension 1) une réflexion.

14.7.8.1. — Montrons tout d'abord qu'aucune des droites stables sous f n'est isotrope. C'est tautologique si q est anisotrope. Supposons q isotrope et utilisons l'identification $L \simeq k \times k$ dont les propriétés ont été rappelées au 14.7.7.3 ci-dessus. On peut alors écrire $u = (\lambda, \lambda^{-1})$ pour un certain $\lambda \in k^\times$, et f est égale à $(x, y) \mapsto (\lambda y, \lambda^{-1} x)$. En particulier, f *échange* les deux seules droites isotropes $k(1,0)$ et $k(0,1)$, et aucune d'elle ne peut donc être une droite propre.

14.7.8.2. — Soit D l'une des droites propres de f. Elle est anisotrope par ce qui précède, donc $D^\perp$ est une droite distincte de D. Comme f est une isométrie, on a $f(D^\perp) = f(D)^\perp = D^\perp$, si bien que $D^\perp$ est également stable. Les deux droites propres de f sont donc orthogonales l'une de l'autre (et anisotropes). Par conséquent, f est bien une réflexion.

Remarque 14.7.9. — Il résulte des discussions qui précèdent, et plus précisément de 14.7.7.3 et 14.7.8.1, que lorsque q est isotrope, on a une caractérisation géométrique très simple des isométries directes et indirectes : une isométrie de L est directe (resp.

indirecte) si et seulement si elle préserve (resp. échange) les deux droites isotropes de L.

Remarque 14.7.10. — Par ce qui précède, toute isométrie indirecte de L est une réflexion. Par ailleurs, soit f une isométrie directe de L. On a $f = (f \circ \sigma) \circ \sigma$, et $f \circ \sigma$ et σ sont indirectes ; par conséquent, f est composée de deux réflexions.

14.7.11. *Aspect matriciel.* — Nous allons conclure cette étude par la variante matricielle de la proposition 14.7.4.2. Soit donc f un endomorphisme du k-ev L.

14.7.11.1. — On déduit de *loc. cit.* que f est une isométrie directe, resp. indirecte, si et seulement si il existe x et y dans k tels que $x^2 - ay^2 = 1$ et tels que la matrice de f dans la base $(1, \omega)$ soit égale à

$$\begin{pmatrix} x & ay \\ y & x \end{pmatrix}, \quad \text{resp.} \begin{pmatrix} x & -ay \\ y & -x \end{pmatrix}.$$

14.7.11.2. — Supposons de plus que q est isotrope, c'est-à-dire que a est un carré dans k, et utilisons l'identification $L \simeq k \times k$ dont les propriétés sont rappelées au 14.7.7.3. On déduit alors de la proposition 14.7.4.2 que f est une isométrie directe, resp. indirecte, si et seulement s'il existe $\lambda \in k^\times$ tel que la matrice de f dans la base canonique de $k \times k$ soit égale à

$$\begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix}, \quad \text{resp.} \begin{pmatrix} 0 & \lambda^{-1} \\ \lambda & 0 \end{pmatrix}.$$

Remarque 14.7.12. — Revenons à la situation initiale d'un k-plan quadratique V isométrique à $\langle 1, -a \rangle$. Donnons-nous une base $\mathscr{B}$ de V dans laquelle la matrice de q est égale à $\mathrm{Diag}(1, -a)$. Il existe alors une isométrie de V sur L envoyant $\mathscr{B}$ sur $(1, \omega)$; par conséquent, un endomorphisme de V est une isométrie directe, resp. indirecte, si et seulement si sa matrice dans la base $\mathscr{B}$ est de la forme correspondante décrite au 13.4.13.1.

Supposons V isotrope et choisissons une base $\mathscr{B}'$ de V dans laquelle la matrice de q est égale à $\mathrm{Diag}(1, -1)$. Il existe alors une isométrie de V sur $k \times k$ envoyant $\mathscr{B}'$ sur la base canonique de $k \times k$; par conséquent, un endomorphisme de V est une isométrie directe, resp. indirecte, si et seulement si sa matrice dans la base $\mathscr{B}'$ est de la forme correspondante décrite au 14.7.11.2.

14.8. Le théorème de Cartan-Dieudonné. — Nous nous proposons maintenant de démontrer le théorème de Cartan-Dieudonné, qui assure que toute isométrie d'un espace quadratique V non dégénéré est produit de réflexions, et l'on dispose en plus d'une borne sur le nombre de réflexions nécessaires : c'est la dimension n de V. En fait, comme nous allons le voir, il y a une première version plus facile de ce théorème qui fournit simplement la borne $2n$ (et tout de même la borne n lorsque V est anisotrope, ce qui couvre par exemple le cas de la géométrie euclidienne). L'obtention de la borne n en toute généralité est plus difficile, et fera l'objet d'un second théorème.

Lemme 14.8.1. — *Soit V un k-espace quadratique non dégénéré, soit v un vecteur anisotrope de V et soit W le sous-espace $(kv)^\perp$ de V. Soit f une isométrie de V telle* *que $f(v) = v$ et telle que $f|_W$ soit la composée d'un nombre fini r de réflexions de W. Alors f est la composée de r réflexions de V.*

Commentaires 14.8.2. — Sous les hypothèses du lemme ci-dessus, $V = kv \oplus W$ et f stabilise kv, donc stabilise son orthogonal W. On peut en conséquence voir $f|_W$ comme une isométrie de W, ce qui est implicitement fait dans l'énoncé.

Démonstration du lemme 14.8.1. — Écrivons $f|_W = s_1 \circ s_2 \circ \ldots \circ s_r$ où s_i est une réflexion de W dont on note H_i l'hyperplan. Pour tout i, notons σ_i l'endomorphisme de V qui envoie v sur v et dont la restriction à W est égale à σ_i ; c'est une réflexion de V d'hyperplan $kv \overset{\perp}{\oplus} H_i$. On a alors par construction $f = \sigma_1 \circ \sigma_2 \circ \ldots \circ \sigma_r$, ce qui termine la preuve. □

***Théorème 14.8.3* (Théorème de Cartan-Dieudonné, version facile)**
Soit V un k-espace quadratique non dégénéré dont on note n la dimension et soit f une isométrie de V. L'application f est la composée d'au plus $2n$ réflexions, et d'au plus n réflexions si V est anisotrope.

Démonstration. — Notons q la forme quadratique structurale sur V. On procède par récurrence sur n. Le résultat est vrai si $n = 0$ (14.6.1 et remarque 14.6.3). On suppose désormais $n > 0$ et le résultat vrai en dimensions $< n$. Comme $V \neq \{0\}$ et comme q est non dégénérée, il existe $v \in V$ tel que $q(v) \neq 0$. Posons $W = (kv)^{\perp}$; on a $V = kv \oplus W$. Par définition d'une isométrie, $q(f(v)) = q(v)$, et la proposition 13.5.5 assure dès lors l'existence d'une isométrie σ de V telle que $\sigma(f(v)) = v$ qui est un produit d'au plus 2 réflexions, et qui est une réflexion si q est anisotrope. Par construction, v est invariant sous l'isométrie $\sigma \circ f$, qui stabilise donc W.

Appliquons l'hypothèse de récurrence à l'isométrie $(\sigma \circ f)|_W$ de l'espace quadratique non dégénéré W qui est de dimension $n - 1$. Elle assure que l'isométrie $(\sigma \circ f)|_W$ est composée d'au plus $2n - 2$ réflexions, et même d'au plus $n - 1$ réflexions si $q|_W$ est anisotrope. En vertu du lemme 14.8.1, il en va de même de $\sigma \circ f$; il s'ensuit que $f = \sigma \circ (\sigma \circ f)$ est produit d'au plus $2n - 2 + 2 = 2n$ réflexions, et même d'au plus $n - 1 + 1 = n$ réflexions si q est anisotrope. □

Avant d'énoncer la version délicate du théorème de Cartan-Dieudonné, nous aurons besoin du lemme suivant.

Lemme 14.8.4. — *Soit V un k-espace quadratique non dégénéré et soit f une isométrie de V. On a l'égalité*

$$\mathrm{Ker}(f - \mathrm{Id})^{\perp} = \mathrm{Im}(f - \mathrm{Id}).$$

Démonstration. — On note b la forme bilinéaire symétrique structurale de V. Soit v un vecteur de V tel que $f(v) = v$, et soit w un vecteur de V. Nous allons utiliser la caractérisation des isométries en terme d'adjonction (13.4.14 et 13.4.14.4). On a

$$\begin{aligned} b(v, f(w)) &= b(f(v), f(w)) \\ &= b(v, f^*(f(w))) \\ &= b(v, w), \end{aligned}$$

où la dernière égalité provient du fait que $f^* = f^{-1}$ puisque f est une isométrie. Par conséquent on a

$$b(v, f(w) - w) = b(v, w) - b(v, w) = 0$$

et v est orthogonal à $f(w) - w$. On en déduit que $\mathrm{Im}(f - \mathrm{Id}) \subset \mathrm{Ker}(f - \mathrm{Id})^{\perp}$. Par ailleurs comme V est non dégénérée on a

$$\dim \mathrm{Ker}(f - \mathrm{Id})^{\perp} = \dim V - \dim \mathrm{Ker}(f - \mathrm{Id}) = \dim \mathrm{Im}(f - \mathrm{Id}),$$

d'où il découle que

$$\mathrm{Im}(f - \mathrm{Id}) = \mathrm{Ker}(f - \mathrm{Id})^{\perp}.$$

□

***Théorème 14.8.5* (Théorème de Cartan-Dieudonné, version délicate)**
Soit V un k-espace quadratique non dégénéré dont on note n la dimension et soit f une isométrie de V. L'application f est la composée d'au plus n réflexions.

Démonstration. — Notons q la forme quadratique structurale sur V. On raisonne par récurrence sur n. Il résulte de 14.6.1, 14.6.2 et remarque 14.6.3 que le théorème est vrai en dimension $\leqslant 1$, et de la remarque 14.7.10 qu'il est vrai en dimension 2. On suppose donc $n \geqslant 3$ et le théorème vrai en dimensions $< \dim V$.

Supposons tout d'abord qu'il existe un vecteur anisotrope v de V tel que $f(v) = v$, et posons $W = (kv)^{\perp}$. Comme v est anisotrope, on a $V = kv \oplus W$; comme kv est stable sous f, l'espace W est également stable sous f, et il résulte de l'hypothèse de récurrence que $f|_W$ est produit d'au plus $n - 1$ réflexions ; en vertu du lemme 14.8.1, f est elle-même produit d'au plus $n - 1$ réflexions, et le théorème est démontré.

Supposons maintenant qu'il existe un vecteur v de V anisotrope tel que $v - f(v)$ soit lui-même anisotrope. Posons $W = (kv)^{\perp}$; on a $V = kv \oplus W$. Par définition d'une isométrie, $q(f(v)) = q(v)$, et puisque $v - f(v)$ est anisotrope, la proposition 13.5.5 assure dès lors l'existence d'une rélfexion σ de V telle que $\sigma(f(v)) = v$. Par construction, v est invariant sous $\sigma \circ f$, qui stabilise donc W.

Appliquons l'hypothèse de récurrence à l'isométrie $(\sigma \circ f)|_W$ de l'espace quadratique non dégénéré W qui est de dimension $n - 1$. Elle assure que l'isométrie $(\sigma \circ f)|_W$ est composée d'au plus $n - 1$ réflexions ; en vertu du lemme 14.8.1, il en va de même de $\sigma \circ f$; il s'ensuit que $f = \sigma \circ (\sigma \circ f)$ est produit d'au plus $n - 1 + 1 = n$ réflexions, ce qui termine la preuve.

Il reste à traiter le cas où f satisfait la propriété suivante :

(P) $\mathrm{Ker}(f - \mathrm{Id})$ *est totalement isotrope, et* $v - f(v)$ *est isotrope pour tout vecteur anisotrope* v *de* V*.*

Montrons tout d'abord que sous l'hypothèse (P), $v - f(v)$ est isotrope pour tout vecteur v de V. Il suffit de le vérifier lorsque v est isotrope et non nul, et l'on se donne donc un tel v. Comme V est non dégénéré, l'orthogonal $(kv)^{\perp}$ de v est de dimension $n - 1$. Puisque $n \geqslant 3$, l'entier $n - 1$ est strictement supérieur à $\lfloor n/2 \rfloor$; or la dimension maximale d'un sous-espace totalement isotrope de V est en vertu du théorème-définition 13.5.7 majorée par $\lfloor n/2 \rfloor$; il s'ensuit que $(kv)^{\perp}$ n'est pas totalement isotrope, et il contient donc un vecteur anisotrope w. Puisque w est orthogonal à v et puisque v est isotrope, on a $q(w - v) = q(w + v) = q(w) \neq 0$. L'hypothèse (P) assure alors que

$f(w) - w$ est isotrope, que $f(w+v) - (w+v) = f(w) - w + f(v) - v$ est isotrope, et que $f(w-v) - f(w-v) = f(w) - w - f(v) + v$ est isotrope. Posons $a = f(w) - w$ et $b = f(v) - v$. Les vecteurs $a, a+b$ et $a-b$ sont isotropes. L'égalité $q(a+b) = q(a-b)$ entraîne alors que a et b sont orthogonaux, puis l'isotropie de a et celle de $a+b$ entraînent celle de b. Ainsi, $f(v) - v$ est isotrope, ce qu'on souhaitait établir.

D'après le lemme 14.8.4, on a $\mathrm{Im}(f - \mathrm{Id}) = \mathrm{Ker}(f - \mathrm{Id})^{\perp}$, ce qui entraîne que $\mathrm{Im}(f - \mathrm{Id})^{\perp} = \mathrm{Ker}(f - \mathrm{Id})$. Par ailleurs l'espace $\mathrm{Ker}(f - \mathrm{Id})$ est totalement isotrope par hypothèse, et $\mathrm{Im}(f - \mathrm{Id})$ est totalement isotrope par ce qui précède. On en déduit que $\mathrm{Im}(f - \mathrm{Id}) \subset \mathrm{Ker}(f - \mathrm{Id})$ et $\mathrm{Ker}(f - \mathrm{Id}) \subset \mathrm{Im}(f - \mathrm{Id})$; autrement dit, on a l'égalité $\mathrm{Im}(f - \mathrm{Id}) = \mathrm{Ker}(f - \mathrm{Id})$. Ceci entraîne que $(f - \mathrm{Id})^2 = 0$ (l'inclusion $\subset$ aurait suffi) et également que $\dim V = 2 \dim \mathrm{Ker}(f - \mathrm{Id})$; en particulier, $n = \dim V$ est pair.

Puisque $f - \mathrm{Id}$ est nilpotente, il existe une base de V dans laquelle sa matrice est triangulaire supérieure avec des 0 sur la diagonale ; la matrice de f dans cette même base est alors triangulaire supérieure avec des 1 sur la diagonale, si bien que $\det f = 1$: l'isométrie f est directe.

Nous allons maintenant conclure. Choisissons une réflexion arbitraire σ dans V (il suffit de choisir un vecteur anisotrope et de considérer la réflexion par rapport à l'orthogonal de la droite qu'il engendre). L'isométrie $\sigma \circ f$ est indirecte. Par ce qui précède, elle ne peut donc satisfaire la propriété (P) ; par conséquent, $\sigma \circ f$ est justiciable de l'un des deux cas particuliers traités en début de preuve, et le théorème vaut pour $\sigma \circ f$; celle-ci est dès lors produit d'au plus n réflexions. Puisque n est pair et puisque $\sigma \circ f$ est indirecte, $\sigma \circ f$ ne peut être produit de n réflexions ; elle est donc produit d'au plus $n-1$ réflexions, et $f = \sigma \circ (\sigma \circ f)$ est produit d'au plus n réflexions. $\square$

14.9. Groupes orthogonaux réels. — Nous nous proposons maintenant de dire quelques mots sur le groupe orthogonal d'un $\mathbf{R}$-espace quadratique non dégénéré, avec une insistance particulière sur le cas des espaces euclidiens. Notons que le cas des plans relève de l'étude générale menée à la sous-section 14.7 avec $k = \mathbf{R}, a = -1$ et $L = \mathbf{C}$.

14.9.1. *À propos des espaces euclidiens.* — Faisons maintenant quelques remarques sur la façon dont certains résultats vus plus haut se déclinent dans le cas des espaces euclidiens.

14.9.1.1. *Bases orthonormées.* — Soit (V, q, b) un espace euclidien et soit n la dimension de V. Une base (e_i) de V est dite *orthonormée* si $b(e_i, e_j) = \delta_{ij}$ pour tout (i, j) ; c'est le cas si et seulement si la matrice de q dans la base (e_i) est l'identité ; si l'on note (X_i) la base duale de (e_i), cela revient à demander que $q = \sum X_i^2$. L'espace V est de signature $(n, 0)$; il possède donc une base orthonormée. Notons que lorsque V est une droite, cela signifie simplement que V possède un vecteur v tel que $q(v) = 1$ (ce qu'on voit très simplement directement, en prenant v non nul arbitraire puis en le divisant si besoin par $\sqrt{q(v)}$).

14.9.1.2. *Orthogonalité.* — Soit V un espace euclidien et soit W un sous-espace vectoriel de V. L'espace W est encore euclidien, et en particulier non dégénéré ; par conséquent, $V = W \oplus W^{\perp}$.

14.9.1.3. *Matrices de rotations.* — Soit V un plan euclidien. Il résulte de 14.7.11.1 (voir aussi la remarque **??**) qu'un endomorphisme f de V est une isométrie directe (qu'on appelle aussi *rotation* dans ce contexte) si et seulement si il existe deux nombres réels x et y tels que $x^2 + y^2 = 1$ et tels que la matrice de f dans la base $\mathscr{B}$ soit égale à

$$\begin{pmatrix} x & -y \\ y & x \end{pmatrix}.$$

Autrement dit, f est une isométrie si et seulement s'il existe un réel θ (uniquement déterminé modulo 2π) tel que la matrice de f dans la base $\mathscr{B}$ soit égale à

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

Cette matrice sera notée R_θ.

14.9.1.4. *Matrices d'isométries indirectes.* — Soit maintenant g une isométrie indirecte de V. On sait que g est une réflexion (14.7.8) ; par conséquent, il existe une base orthogonale (e_1, e_2) de V dans laquelle la matrice de g est égale à $\mathrm{Diag}(1, -1)$; quitte à remplacer e_i par $e_i/\sqrt{q(e_i)}$ pour $i = 1, 2$, on peut supposer que (e_1, e_2) est orthonormée.

14.9.2. *Topologie du groupe orthogonal.* — Soit (V, q) un $\mathbf{R}$-espace quadratique non dégénéré, soit $\mathscr{B} = (e_1, \ldots, e_n)$ une base de V et soit M la matrice de q dans $\mathscr{B}$. En tant que $\mathbf{R}$-espace vectoriel de dimension finie, $\mathrm{End}_{\mathbf{R}}(V)$ a une topologie naturelle, donnée par n'importe laquelle de ses normes (elles sont toutes équivalentes), et tout sous-ensemble de $\mathrm{End}_{\mathbf{R}}(V)$ sera considéré comme muni de la topologie induite.

14.9.2.1. — Les sous-groupes $\mathrm{O}(q)$ et $\mathrm{SO}(q)$ de $\mathrm{GL}(V)$ sont fermés dans ce dernier. En effet, si f est un endomorphisme de V de matrice N dans $\mathscr{B}$ alors f est une isométrie pour q si et seulement si ${}^tNMN = M$ (13.4.13.1), ce qui est visiblement une condition fermée ; par conséquent $\mathrm{O}(q)$ est fermé dans GL(V), et $\mathrm{SO}(q)$ l'est aussi puisque c'est l'intersection de $\mathrm{O}(q)$ et du noyau de l'application continue $\det \colon \mathrm{GL}(V) \to \mathbf{R}^\times$.

14.9.2.2. *Le cas euclidien.* — Supposons de plus que q est définie positive et que $\mathscr{B}$ est orthonormée. On a alors $M = \mathrm{I}_n$ et un endomorphisme f de V appartient à $\mathrm{O}(q)$ si et seulement si sa matrice N dans la base $\mathscr{B}$ vérifie l'égalité ${}^tMM = \mathrm{I}_n$. Remarquons par ailleurs qu'on sait aussi que f est une isométrie si et seulement si $b(f(e_i), f(e_j)) = b(e_i, j) = \delta_{ij}$ pour tout (i, j), ce qui revient à demander que les colonnes de M forment une base orthonormée pour le produit scalaire usuel de $\mathbf{R}^n$; on retrouve cette condition en explicitant coefficient par coefficient ce que signifie l'égalité ${}^tMM = \mathrm{I}_n$.

On en déduit que le groupe $\mathrm{O}(q)$ est une partie non seulement fermée, mais bornée, de l'espace vectoriel $\mathrm{End}_{\mathbf{R}}(V)$; autrement dit, $\mathrm{O}(q)$ est compact, et il en va de même de son sous-groupe fermé $\mathrm{SO}(q)$.

***Proposition 14.9.3* (Compacité et anisotropie).** — *Soit (V, q) un $\mathbf{R}$-espace quadratique non dégénéré. Les assertions suivantes sont équivalentes :*

(i) *la forme q est anisotrope ;*

(ii) *le groupe* $\mathrm{O}(q)$ *est compact.*

Démonstration. — On a vu ci-dessus que si q est définie positive alors $\mathrm{O}(q)$ est compact ; il l'est également si q est définie négative puisque $\mathrm{O}(q) = \mathrm{O}(-q)$. Par conséquent, (i)⇒(ii).

Nous allons montrer réciproquement que (ii)⇒(i). On suppose donc $\mathrm{O}(q)$ compact, et l'on cherche à montrer que q est anisotrope. On raisonne par l'absurde. On suppose donc q isotrope. On peut alors écrire V sous la forme $H \overset{\perp}{\oplus} W$ où H est un plan hyperbolique (théorème-définition 13.5.7). Pour toute isométrie f de H, notons $\rho(f)$ l'isométrie de V qui envoie un élément $v = h + w$ de V (où $h \in H$ et $w \in W$) sur $f(h) + w$. L'application ρ est un morphisme injectif et continu $\mathrm{O}(q|_H) \hookrightarrow \mathrm{O}(q)$. Comme $q|_H$ est isotrope et non dégénérée, il résulte de 14.7.11 qu'il existe une base (e_1, e_2) de H telle qu'un endomorphisme f de H est une isométrie si et seulement si sa matrice dans (e_1, e_2) est de la forme

$$\begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix} \text{ ou } \begin{pmatrix} 0 & \lambda^{-1} \\ \lambda & 0 \end{pmatrix}$$

où $\lambda \in \mathbf{R}^\times$. Soit $(e_3, \ldots, e_n)$ une base de W. Une isométrie g de V appartient à l'image de ρ si et seulement si sa matrice $N := (a_{ij})$ dans la base $(e_1, \ldots, e_n)$ satisfait les conditions suivantes :

- ◇ $a_{ij} = 0$ dès que ($1 \leqslant j \leqslant 2$ et $i > 2$) ou ($j \geqslant 3$ et $i \neq j$)
- ◇ $a_{jj} = 1$ pour tout $j \geqslant 3$.

Par conséquent, le groupe $G := \rho(\mathrm{O}(q|_H))$ est un sous-groupe fermé de $\mathrm{O}(q)$; en particulier, G est compact. L'application de G dans V qui envoie g sur $g(e_1)$ est continue, et $\{g(e_1)\}_{g \in G}$ est donc compact. Mais par ce qui précède on a

$$\{g(e_1)\}_{g \in G} = \{\lambda e_1\}_{\lambda \in \mathbf{R}^\times} \cup \{\lambda e_2\}_{\lambda \in \mathbf{R}^\times}$$

qui n'est pas compact (cet ensemble n'est même en fait ni fermé ni borné dans V). On aboutit ainsi à une contradiction. □

Nous allons avoir besoin du lemme suivant, qui a son intérêt propre.

Lemme 14.9.4. — *Soit V un $\mathbf{R}$-espace vectoriel de dimension finie et soit f un endomorphisme de V. L'endomorphisme f possède une droite stable ou un plan stable.*

Démonstration. — Si f possède une valeur propre réelle alors f possède un vecteur propre, et partant une droite stable. Supposons maintenant que f n'ait pas de valeur propre réelle, et soit χ son polynôme caractéristique. Puisque f n'a pas de valeur propre réelle, χ n'a pas de facteur irréductible de degré 1 dans $\mathbf{R}[X]$; il s'écrit donc $\prod_i \chi_i$ où chacun des χ_i est irréductible de degré 2 (les χ_i ne sont pas forcément deux à deux distincts). Puisque $\chi(f) = 0$ par le théorème de Cayley-Hamilton, et puisque la dimension de V est strictement positive, il existe i tel que $\chi_i(f)$ ne soit pas injective. Écrivons $\chi_i = X^2 + aX + b$. Il existe par choix de χ_i un vecteur v non nul de V tel que $f^2(v) = -af(v) - bv$; ceci entraîne aussitôt que l'espace vectoriel engendré par v et $f(v)$ est stable sous f. Cet espace est non nul (puisque v est non nul), et ce ne peut pas être une droite car sinon $f(v)$ serait colinéaire à v et f aurait une valeur propre réelle ; il est donc dimension 2, et l'on a bien exhibé un plan de V stable sous f. □

Théorème 14.9.5 **(Description générale des isométries d'un espace euclidien)**

Soit V un espace vectoriel euclidien dont on note n la dimension et soit f une isométrie de V. Il existe une base orthonormée de V dans laquelle la matrice de f est de la forme suivante :

- $\diamond$ $\mathrm{Diag}(R_{\theta_1}, \ldots, R_{\theta_d})$ *si $n = 2d$ et si f est directe ;*
- $\diamond$ $\mathrm{Diag}(1, -1, R_{\theta_1}, \ldots, R_{\theta_{d-1}})$ *si $n = 2d$ et si f est indirecte ;*
- $\diamond$ $\mathrm{Diag}(1, R_{\theta_1}, \ldots, R_{\theta_d})$ *si $n = 2d + 1$ et si f est directe ;*
- $\diamond$ $\mathrm{Diag}(-1, R_{\theta_1}, \ldots, R_{\theta_d})$ *si $n = 2d + 1$ et si f est indirecte.*

Démonstration. — Notons q la forme quadratique structurale de V. Il suffit de montrer l'existence d'une base orthonormée de V dans laquelle la matrice de f est de la forme $\mathrm{Diag}(\beta_1, \ldots, \beta_\ell)$ où chaque bloc β_i est ou bien un bloc scalaire égal à 1 ou à -1, ou bien un bloc 2×2 de la forme R_θ ; on passe en effet de cette forme à l'une des quatre décrites dans l'énoncé en remarquant que $\mathrm{Diag}(1, 1) = R_0$ et $\mathrm{Diag}(-1, -1) = R_\pi$ et en permutant convenablement les éléments de la base.

Nous allons procéder par récurrence sur n. Le cas $n = 0$ est évident : la base vide (la seule possible !) a alors la propriété requise. Supposons $n > 0$ et le résultat vrai en dimensions $< n$. En vertu du lemme 14.9.4, l'isométrie f possède une droite ou un plan stable.

Supposons que f possède une droite stable D ; soit e_1 un vecteur de D tel que $q(e_1) = 1$. Puisque $q(f(e_1)) = 1$, il existe $\varepsilon \in \{-1, 1\}$ tel que $f(e_1) = \varepsilon e_1$. On a $V = D \oplus D^\perp$; comme D est stable sous f, son orthogonal l'est aussi. L'hypothèse de récurrence assure alors l'existence d'une base orthonormée $(e_2, \ldots, e_n)$ de $D^\perp$ dans laquelle la matrice de $f|_{D^\perp}$ est de la forme $\mathrm{Diag}(\beta_1, \ldots, \beta_\ell)$ avec les β_ℓ comme plus haut ; la matrice de f dans la base orthonormée $(e_1, \ldots, e_n)$ est donc égale à $\mathrm{Diag}(\varepsilon, \beta_1, \ldots, \beta_\ell)$ et elle est bien de la forme requise.

Supposons que f possède un plan stable P. On a $V = P \oplus P^\perp$; comme P est stable par f, son orthogonal $P^\perp$ l'est aussi. L'hypothèse de récurrence assure alors l'existence d'une base orthonormée $(e_3, \ldots, e_n)$ de $P^\perp$ dans laquelle la matrice de $f|_{P^\perp}$ est de la forme $\mathrm{Diag}(\beta_1, \ldots, \beta_\ell)$ avec les β_ℓ comme plus haut. On distingue maintenant deux cas. Commençons par celui où f est directe. Il existe alors une base orthonormée (e_1, e_2) de P dans laquelle la matrice de f est égale à R_θ pour un certain réel θ (14.9.1.3) ; la matrice de f dans la base orthonormée $(e_1, e_2, e_3, \ldots, e_n)$ de V est alors égale à $\mathrm{Diag}(R_\theta, \beta_1, \ldots, \beta_\ell)$ et elle est bien de la forme requise. Traitons ensuite le cas où f est indirecte. Il existe alors une base orthonormée (e_1, e_2) de P dans laquelle la matrice de f est $\mathrm{Diag}(1, -1)$ (14.9.1.4). La matrice de f dans la base orthonormée $(e_1, e_2, e_3, \ldots, e_n)$ de V est alors égale à $\mathrm{Diag}(1, -1, \beta_1, \ldots, \beta_\ell)$ et elle est bien de la forme requise. □

Année universitaire 2019-2020

ANTOINE DUCROS